UNIVERSITÀ
DEGLI STUDI
DI TRIESTE

UNIVERSITÀ
DEGLI STUDI
DI UDINE

Dottorato di Ricerca Interateneo in Ingegneria Civile–Ambientale e Architettura

Curriculum: INGEGNERIA CIVILE-AMBIENTALE
Settore Scientifico Disciplinare: ICAR/04
XXXIV Ciclo

ANALISI E MODELLAZIONE DELLE INTERAZIONI VEICOLO-PEDONE PER LO SVILUPPO DI SISTEMI ATTIVI DI ASSISTENZA ALLA GUIDA E DI PROTEZIONE DEI PEDONI

DOTTORANDO
MATTEO MIANI

COORDINATORE
PROF. ALBERTO SDEGNO

SUPERVISORE DI TESI
PROF. NICOLA BALDO

CO-SUPERVISORE DI TESI
PROF. ANDREA MARINI

ANNO ACCADEMICO 2020/2021

Abstract

La sicurezza e la mobilità dei pedoni sono requisiti basilari che dovrebbero caratterizzare ogni sistema di trasporto urbano. Tuttavia, le morti degli utenti della strada più vulnerabili costituiscono ancora oggi una componente significativa di tutte le vittime della strada nel Mondo. Nonostante gli innumerevoli sforzi compiuti per l'innovazione tecnologica dei veicoli e il riesame degli spazi urbani, le statistiche sull'incidentalità dimostrano la necessità e l'importanza di sviluppare sempre più affidabili sistemi di protezione in grado di diminuire gli impatti sociali ed economici del sistema di trasporto.

Sebbene sul mercato di massa siano stati immessi molti sistemi di frenata automatica di emergenza (o AEB, dall'inglese Automatic Emergency Braking), una misura di sicurezza chiave nei veicoli moderni in grado di evitare o mitigare gli effetti di una collisione, diversi ricercatori hanno individuato una nuova strategia per lo sviluppo efficiente di questi sistemi: migliorare la sicurezza dei pedoni nel traffico urbano richiede sistemi "intelligenti" in grado, non solo di comprendere lo stato attuale dell'interazione veicolo-pedone, ma di anticipare proattivamente il futuro modello di rischio dell'evento. In altre parole, prevedere in anticipo le decisioni degli utenti nella scena di traffico, interpretare i comportamenti dei conducenti e definire accurate metriche di valutazione del rischio sono gli obbiettivi da perseguire per raggiungere nuovi traguardi nell'ambito della mobilità sostenibile.

Questo elaborato discute la natura globale del problema della sicurezza dei pedoni e i diversi approcci che sono stati sviluppati dai gruppi di ricerca nel Mondo per affrontarlo. Inoltre, la tesi presenta nel dettaglio lo studio, l'implementazione e l'analisi di un innovativo modello di valutazione del rischio, recentemente oggetto di pubblicazione su rivista internazionale, per l'efficientamento dei sistemi di assistenza alla guida esistenti. Il modello proposto, basato su moderne tecniche di *Machine Learning* e processi di analisi in linea con la letteratura scientifica più recente, è in grado di predire, fino a tre secondi nel futuro, il livello di rischio atteso negli incontri tra veicolo e pedone sulle strisce pedonali in funzione della rappresentazione attuale della scena di traffico tratta da radar e telecamere esterne al veicolo. Infatti, l'algoritmo prototipato fornisce una previsione sequenziale, su più orizzonti temporali, di indicatori di sicurezza operativi che descrivono in continuo il processo di incontro e permettono di annotare le interazioni conflittuali gravi. L'applicazione è stata ottimizzata attraverso dati di mobilità, acquisiti con un simulatore di guida avanzato ad elevato grado di realismo, su un campione di giovani conducenti. Questi ultimi hanno affrontato diversi conflitti veicolo-pedone su un percorso urbano virtuale pianificato. La conoscenza acquisita dal modello in questo contesto potrà essere sfruttata per facilitare l'adattamento online del sistema a nuove situazioni operative e alle diverse caratteristiche comportamentali degli utenti.

Indice

Abstract **I**
Elenco delle Figure **V**
Elenco delle Tabelle **VII**
Elenco delle Abbreviazioni **IX**
1 Introduzione **1**
2 Comportamenti e Rischio **15**
2.1 Modellazione del Comportamento ... 17
2.1.1 Rappresentazione della scena ... 19
2.1.2 Output predetto ... 21
2.1.3 Approccio di modellazione ... 22
2.2 Valutazione del Rischio ... 30
2.2.1 Definizioni operative di severità ... 32
2.3 Un nuovo sistema di valutazione della sicurezza dei pedoni ... 39
3 Raccolta ed Elaborazione dei Dati **45**
3.1 Il simulatore di guida AS 1000-M ... 47
3.2 La sperimentazione al simulatore AS 1000-M ... 49
3.2.1 Validazione della simulazione ... 52
3.3 Formulazione del problema di modellazione ... 66
3.4 Rappresentazione della situazione ... 67
3.5 Generazione delle classi di severità ... 69
4 Metodologia **73**
4.1 Reti Neurali Artificiali ... 73
4.2 Funzioni di errore e regolarizzazione delle ANN ... 76
4.3 k-fold Cross Validation ... 78
4.4 Gated Recurrent Unit (GRU) ... 79
4.5 Dettagli di implementazione ... 81
5 Presentazione e Discussione dei Risultati **85**
6 Conclusioni e Ricerche future **97**
Appendice A **99**
Calcolo del Time-to-Collision per un punto ed un segmento ... 99
Calcolo del Time Advantage e T_2 per un punto ed un segmento ... 101
Appendice B **103**
Definizione delle metriche di valutazione dei modelli ... 103
Risultati della ricerca a griglia ... 106
Bibliografia **109**

Elenco delle Figure

Figura 1: Distribuzione dei decessi di pedoni per Regione della World Health Organization. 2

Figura 2: Illustrazione dei diversi fattori di rischio per la sicurezza dei pedoni. .. 5

Figura 3: La piramide - le interazioni tra gli utenti della strada come un continuum di eventi di traffico. 9

Figura 4: Struttura di un sistema ADAS. 15

Figura 5: Esempi di previsione secondo i tre diversi modelli di movimento. 23

Figura 6: Il nuovo sistema di valutazione del livello di severità basato sull'interazione veicolo-pedone. 40

Figura 7: Il simulatore di guida AS 1000-M. 47

Figura 8: Vista frontale da veicolo degli scenari simulati: (a) ragazzo che attraversa; (b) ragazza che attraversa. 50

Figura 9: Profili di velocità medi dei conducenti. Legenda: MC - conducenti maschi del gruppo di controllo; ME - conducenti maschi del gruppo sperimentale; FC - conducenti femmine del gruppo di controllo; FE - conducenti femmine del gruppo sperimentale. 56

Figura 10: Schema di riferimento per il calcolo delle distanze longitudinali tra pedone e veicolo. 57

Figura 11: Effetto di genere e distrazione sulla variabile TA. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi della variabile TA, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo. 60

Figura 12: Effetto di genere e distrazione sulle variabili (a) V_i e (b) L_{Vi}. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi delle variabili V_i e L_{Vi}, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo. 61

Figura 13: Effetto di genere e distrazione sulle variabili (a) V_b e (b) L_{Vb}. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi delle variabili V_b e L_{Vb}, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo. 62

Figura 14: Effetto di genere e distrazione sulla variabile L_{Vmin}. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi della variabile L_{Vmin}, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo. 63

Figura 15: Effetto di genere e distrazione sulla variabile $\sigma(V_n)$. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi della variabile $\sigma(V_n)$, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze

statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo. 64

Figura 16: Illustrazione del paradigma di apprendimento e previsione. L'algoritmo di apprendimento mappa la rappresentazione della situazione passata (al tempo $t - T_k$) ed il livello di severità corrente (al tempo t); quindi, il modello addestrato predice il livello di severità futuro (al tempo $t - T_k$) usando la rappresentazione della situazione corrente (al tempo t). 67

Figura 17: Esempio di segmentazione del profilo di velocità in primitive di comportamento per un partecipante della sperimentazione. 68

Figura 18: Esempio di andamento del $T_2(t)$ durante l'interazione tra veicolo e pedone, con segmentazione in base alle soglie statiche di pericolosità, per un partecipante della sperimentazione (il medesimo di Figura 17). 71

Figura 19: Esempi di comuni funzioni di attivazione, l'unità lineare esponenziale (ELU), la tangente iperbolica ($tanh$), l'unità lineare rettificata (ReLU) e la logistica sigmoidale ($logsig$). 74

Figura 20: Rappresentazione schematica delle operazioni eseguite da una unità GRU. 81

Figura 21: Distribuzioni cumulate di frequenza degli errori assoluti di previsione del modello GRU_{T2} sui cinque folds di validazione, per gli orizzonti temporali di (a) $1\ s$, (b) $2\ s$ e (c) $3\ s$. 88

Figura 22: Rappresentazione schematica della matrice di confusione. 89

Figura 23: Curve ROC del modello GRU_{TADV} sui cinque folds di validazione, per gli orizzonti temporali di (a) $1\ s$, (b) $2\ s$ e (c) $3\ s$. 91

Figura 24: Curve ROC del modello $S\text{-}GRU_{SL}$ sui cinque folds di validazione, per gli orizzonti temporali di (a) $1\ s$, (b) $2\ s$ e (c) $3\ s$. 93

Figura A1: a) schema per il calcolo di TTC, T_{ADV} e T_2 per un punto ed un segmento, b) esempio del caso particolare $t_{ln\,1} = 0$. 99

Elenco delle Tabelle

Tabella 1: Definizione di alcuni indicatori SSM. 34

Tabella 2: Frequenze nell'uso del telefono cellulare durante la guida. 54

Tabella 3: Verifica delle ipotesi del test ANOVA. Legenda: le prime quattro righe riportano il risultato del test di normalità di Shapiro-Wilk (valori p) per ognuno dei quattro gruppi considerati nello studio (cioè FE, FC, ME, MC) e separatamente per ogni variabile del profilo di velocità; diversamente, l'ultima riga mostra il risultato (valori p) del test di omogeneità della varianza (o test di Levene) che verifica se la varianza dei punteggi è uguale per ognuno dei quattro sottogruppi, separatamente per ogni variabile del profilo di velocità. 58

Tabella 4: Valore medio (deviazioni standard) delle variabili sul profilo di velocità per i gruppi di controllo e sperimentale. 59

Tabella 5: Statistiche descrittive delle variabili di input e output ai modelli GRU. 72

Tabella 6: Risultati sperimentali (deviazioni standard) per i modelli GRU_{T2} e GRU_{TADV}. 87

Tabella 7: Confronto tra i modelli GRU_{TADV} con e senza applicazione dello SMOTE. 90

Tabella 8: Risultati sperimentali (deviazioni standard) per i modelli $M\text{-}GRU_{SL}$ e $S\text{-}GRU_{SL}$. 92

Tabella 9: Confronto tra i modelli $M\text{-}GRU_{SL}$ e $S\text{-}GRU_{SL}$. 92

Tabella 10: Confronto tra i modelli $S\text{-}GRU_{SL}$ con e senza applicazione dello SMOTE. 94

Tabella B1: Risultati della ricerca a griglia per il modello GRU_{T2}. 106

Tabella B2: Risultati della ricerca a griglia per il modello GRU_{TADV}. 107

Tabella B3: Risultati della ricerca a griglia per il modello $S\text{-}GRU_{SL}$. 108

Elenco delle Abbreviazioni

SSM: *Surrogate Safety Measures*
RNN: *Recurrent Neural Network*
LSTM: *Long Short-Term Memory*
GRU: *Gated Recurrent Unit*
ADAS: *Advanced Driver Assistance Systems*
FCW: *Forward Collision Warning*
AEB: *Automated Emergency Braking*
RAM: *Risk Assessment Model*
EV: *Ego-vehicle*
CNN: *Convolutional Neural Network*
ANN: *Artificial Neural Network*
QI: *Quoziente Intellettivo non verbale*
SRT: *Speed Reduction Time*
SGD: *Stochastic Gradient Descent*
SMOTE: *Synthetic Minority Over-Sampling Technique*
FAR: *False Alarm Rate*
AUC: *Area Under the Curve*
ROC: *Receiver Operating Characteristic*

Capitolo 1

Introduzione

"Walk to be healthy. Walk to be happy." – Charles Dickens

Charles Dickens aveva ragione: camminare può aiutare a sconfiggere la tristezza e avere incredibili benefici sulla salute fisica e mentale [1]. Con la disponibilità di percorsi ciclabili e pedonali, e l'incoraggiamento pubblico alla mobilità attiva, la popolarità di pedoni e ciclisti è in rapida crescita nel Mondo [2]. Per rendere più sicuri gli spostamenti a piedi, gli ingegneri e gli urbanisti cercano continuamente di creare ambienti adatti a chi cammina, anche migliorando i punti di attraversamento delle strade [3], [4]. Tuttavia, se una camminata giornaliera di 20 minuti riduce il rischio di morte precoce di un individuo del 25% [5], i pedoni sono ancora oggi gli utenti della strada più vulnerabili e rappresentano una quota significativa delle vittime della strada nel Mondo. Secondo gli ultimi rapporti delle Organizzazioni Internazionali che si occupano di sicurezza stradale [6], [7], in media il 23% delle 1.35 milioni di vittime annuali da incidente stradale nel Mondo sono pedoni (Figura 1). Inoltre, le vittime tra i pedoni risultano essere soprattutto le fasce più deboli della popolazione, ovvero i bambini (tra 0 e 14 anni) e gli anziani (dai 64 anni in su), a causa del minore livello di motorizzazione e della loro intrinseca vulnerabilità fisica [8]. Infatti, rispetto agli altri utenti della strada, il pedone non è protetto da dispositivi in grado di dissipare energia durante una collisione; in parte per questo motivo, gli incidenti tra veicolo e pedone causano ogni anno ingenti danni umani e finanziari alle comunità. D'altra parte, il numero di vittime della strada non è lo stesso nei diversi Paesi. Infatti, i Paesi a medio reddito annuo pro capite (ovvero compreso in media tra $1'006$ e $12'235$ dollari), che detengono il 59% del totale di veicoli a motore nel Mondo, registrano circa l'80% delle vittime della strada annuali. In Europa, il tasso di mortalità per incidente stradale nei Paesi a medio reddito (14.4 morti per $100'000$ abitanti) è quasi tre volte superiore a quello dei Paesi ad alto reddito (tra cui l'Italia), dove si registrano in media 5.1 morti per $100'000$ abitanti. Lo stesso risultato, ovvero la medesima disuguaglianza tra Paesi, può essere riscontrato considerando i soli pedoni tra le vittime della strada [7]. Ad ogni modo, la quota parte dei pedoni sul numero totale delle vittime della strada varia tra il 14% ed il 40% nei diversi Continenti (Figura 1). Per esempio, i pedoni costituiscono circa il 27% delle vittime di incidenti stradali in Europa, il 22% in Oceania e nelle Americhe, mentre nei Paesi dell'Asia Sud-Orientale il tasso è solo del 14%. Diversamente, i pedoni rappresentano circa il 40% delle vittime della strada in Africa, che registra il più alto tasso tra i Continenti. Traducendo in numeri queste statistiche, si riesce a comprendere l'entità del fenomeno: circa $312'000$ pedoni perdono la vita in incidenti stradali ogni anno nel Mondo [9], di cui l'India è responsabile per quasi $15'800$ vittime, l'Ue28 (ovvero i 28 Stati dell'Unione Europea, incluso il Regno Unito) contribuisce con circa $5'400$ vittime [8], gli Stati Uniti per quasi $6'000$ vittime [10], e così via. Oltre alle conseguenze sociali e culturali, le vittime e i feriti gravi negli incidenti portano a perdite economiche significative in ogni Paese; basti pensare alla perdita di produttività delle vittime e al costo delle cure di cui necessitano i feriti. Infatti, gli incidenti stradali costano alla maggior parte delle

Nazioni il 3% del loro prodotto interno lordo ogni anno [7]. Quelle presentate sono statistiche davvero allarmanti e devono incoraggiare azioni più attente per la protezione degli utenti vulnerabili della strada. Siccome il prezzo pagato per la mobilità è ancora troppo alto, la sicurezza dei pedoni ha la massima priorità e rappresenta una grande sfida per i decisori politici e i progettisti stradali, i quali dovranno proporre e introdurre innovative soluzioni per tendere verso un sistema di trasporto senza incidenti mortali o feriti gravi entro il 2050 [11]. In realtà, la United Nations General Assembly ha fissato l'ambizioso obiettivo, adottato anche dalla UE nell'ambito della Vision Zero [12], di dimezzare il numero globale di morti e feriti gravi per incidenti stradali entro il 2030 $(A/RES/74/299)$.

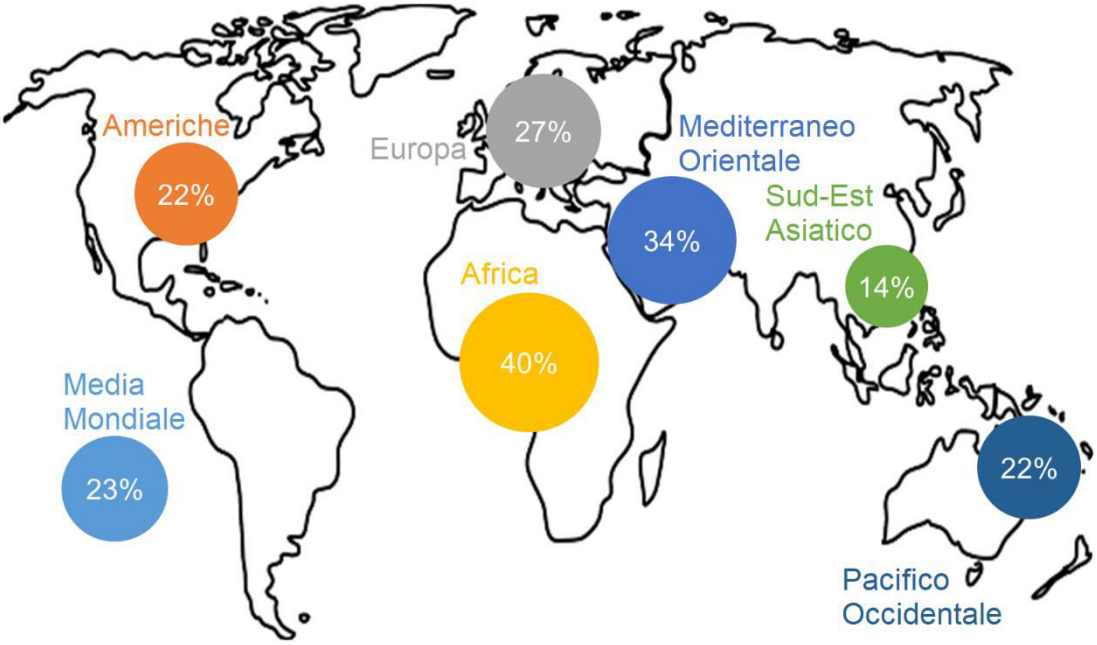


Figura 1: Distribuzione dei decessi di pedoni per Regione della World Health Organization.

In Italia, nel 2019, i pedoni coinvolti in incidenti mortali sono stati 534, ovvero il $16{,}8\%$ delle fatalità registrate sulle vie di comunicazione nazionali [13], di cui 336 uomini e 198 donne. La maggioranza di questi incidenti avviene in ambito urbano, dove i percorsi delle persone e dei veicoli si incontrano più frequentemente. Gli attraversamenti pedonali, come richiesto dal Nuovo Codice della Strada $(D.lgs.285/92\ e\ smi)$, devono essere evidenziati da segnaletica sia orizzontale, le strisce pedonali, che verticale, i cartelli, a cui possono aggiungersi anche segnalazioni semaforiche o luminose, al fine di informare gli automobilisti con largo anticipo della possibile presenza di passanti sulla carreggiata. D'altra parte, i pedoni, per quanto siano tenuti ad utilizzare questi attraversamenti, possono oltrepassare la strada anche al di fuori di essi, compiendo quindi quello che viene definito un attraversamento "non segnalato". Anche questi avvengono in numero maggiore nei centri cittadini, dove l'elevata mobilità pedonale e il grande numero di punti di interesse presenti nel tessuto urbano spingono i passanti ad attraversare la strada dove risulta loro più comodo e non dove sia più idoneo [14], [15]. Tuttavia, a differenza di quanto si possa pensare, la maggior parte degli incidenti in ambito urbano che coinvolgono i pedoni avvengono proprio

sulle strisce pedonali; in Italia se ne possono contare $7^{\cdot}747$ nel solo 2019, di cui 143 mortali, per un indice di letalità dell'1.8%; viceversa, esternamente agli attraversamenti zebrati sono stati registrati $4^{\cdot}535$ incidenti, di cui 124 mortali, per una letalità del 2.7% [16]. Queste statistiche sono state riscontrate anche in altri Paesi [17], [18] e dimostrano che le intersezioni urbane, in cui il tracciato stradale si interseca con i percorsi pedonali, sono le parti della rete stradale più problematiche per la sicurezza dei pedoni. Tuttavia, valutare la relazione tra una specifica caratteristica dell'intersezione stradale (in questo caso l'attraversamento zebrato) ed il suo livello di incidentalità, senza tenere conto del microambiente urbano circostante, dei volumi di traffico, delle caratteristiche della carreggiata e del comportamento degli utenti, può portare a valutazioni distorte [19]. L'efficacia di un attraversamento pedonale varia notevolmente in base alla sua posizione, alla velocità dei veicoli, agli strumenti di controllo del traffico, al numero di corsie, al comportamento e caratteristiche dei diversi utenti della strada [20]–[24]. Infatti, Eisses [25] ha giustificato in parte l'elevata incidentalità sulle strisce pedonali in Europa asserendo che la segnaletica fornisce al pedone un falso senso di sicurezza, nel senso che spinge gli utenti a valicare la strada senza ricercare un *feedback* dall'automobilista in arrivo e con velocità di camminata mediamente maggiori [14], consci del loro diritto di precedenza. Altri fattori potenzialmente importanti associati alle collisioni tra veicolo e pedone nelle intersezioni urbane sono la geometria dell'intersezione (numero di segmenti che si intersecano), la direzione del traffico (a senso unico o bidirezionale) e la presenza di parcheggi per veicoli a lato strada [26]–[29]. Diversamente, al di fuori delle strisce pedonali, i passanti mostrano comportamenti generalmente più cauti (consci, questa volta, di non avere diritto di precedenza), aspettando spazi maggiori tra le autovetture in entrambe le corsie prima di iniziare l'attraversamento e ricercando negli autisti in avvicinamento dei segnali che ne rendano manifesta l'intenzione di fermarsi, evitando anche di passare qualora questi non sembrino rallentare a sufficienza [14], [15], [30]. Dall'altro lato, gli autisti hanno modo di moderare la velocità e frenare in sicurezza solo quando le condizioni di visibilità e il loro livello di attenzione permettono l'individuazione repentina del pedone [31]. Quindi, il fattore cardine alla base della maggiore mortalità degli attraversamenti non segnalati risulta essere quindi l'imprevedibilità degli stessi, causata da una condotta inaspettata da parte dei pedoni spesso in concomitanza con fattori di circostanza quali meteo, ora del giorno, visibilità del tracciato e livelli di distrazione del conducente.

In generale, i fattori che causano gli incidenti stradali sono identificati in base a tre gruppi principali: fattore umano, fattore veicolare e fattore stradale/ambientale. Come fattore più importante negli incidenti, quello umano è il risultato di due diverse classi di aberrazione [32], gli errori e le violazioni, associate a prestazioni e/o comportamenti di guida impropri da parte degli utenti della strada (conducenti, motociclisti, pedoni, ecc.) in situazioni complesse. Nella fattispecie, sono possibili due tipologie psicologicamente distinte di errori: le azioni di un utente possono deviare involontariamente dall'intenzione originale a causa di *deficit* di attenzione (si parla anche di azioni esteriorizzate non pianificate, come rendersi conto di non avere un ricordo nitido della strada appena percorsa); oppure le azioni pianificate di un utente possono deviare da un corretto svolgimento per un fine desiderato a causa delle carenze nei processi

di giudizio e/o inferenziali coinvolti nella selezione di un obiettivo, o dei mezzi per raggiungerlo, o entrambi (come quando un utente in coda per immettersi a sinistra su una via principale si concentra così tanto sul traffico di questa via da tamponare quasi la macchina che lo precede). Diversamente, le violazioni sono state definite come infrazioni volontarie di quei codici sociali o regole legali ritenuti necessari per mantenere il funzionamento sicuro di un sistema potenzialmente pericoloso [32], [33]. Le violazioni hanno origini psicosociali e motivazionali, di conseguenza, possono essere comprese solo analizzando il contesto sociale [32]. Queste aberrazioni possono essere sia di tipo aggressivo, perché associate ad azioni che implicano una componente aggressiva interpersonale di carattere emotivo (come arrabbiarsi con un altro guidatore e inseguirlo o tallonarlo per il solo scopo di infastidirlo), che di tipo ordinario, ovvero azioni intenzionali che si allontanano dalla guida sicura senza avere un obiettivo aggressivo specifico (come uscire da un incrocio così velocemente da obbligare un altro conducente che avrebbe la precedenza a fermarsi). In sintesi, il fattore umano comprende aspetti cognitivi, come la disattenzione o i vuoti di memoria, comportamentali, come la tendenza a guidare in modo rischioso o a commettere violazioni del Codice della Strada (ad esempio, l'uso di alcol e droghe, il mancato uso della cintura di sicurezza, il mancato rispetto dei limiti di velocità, il sorpasso dove è vietato dall'apposito segnale, l'uso del telefono cellulare, ecc.), insieme a quelli demografici, come l'età e il sesso. Dall'altra parte, fattori come i difetti o le capacità intrinseche del veicolo (come la stabilità, il sistema di frenata, etc.), così come le condizioni climatiche avverse, lo stato degradato della pavimentazione stradale e le inadeguate caratteristiche geometriche possono aumentare il rischio di un incidente [34]–[36], anche interagendo tra loro. In particolare, un pedone ha una probabilità del 90% di sopravvivere ad un impatto che avviene a velocità pari o inferiori a $30\ km/h$, ma meno del 50% per collisioni a velocità superiori a $45\ km/h$ [37]. Quindi, il limite di velocità (e la velocità operativa degli utenti che ne consegue) è un fattore cruciale perché ha influenza sulla gravità degli eventi incidentali tra veicolo e pedone. Allo stesso modo, l'illuminazione influisce sulla visibilità e costituisce quindi un fattore critico per la sicurezza: circa il 45% degli incidenti che coinvolgono pedoni in Europa si verificano su strade buie perché prive di illuminazione [8]. Inoltre, i recenti progressi nelle tecnologie dei veicoli hanno introdotto nuove questioni relative alla sicurezza dei pedoni. Per esempio, i veicoli elettrici, che hanno cominciato a popolare le nostre strade, producono pochissimo rumore rispetto a quelli convenzionali basati sulla combustione di carburante, soprattutto a basse velocità. Questa caratteristica non aiuta un pedone a rilevare la loro presenza, specialmente quando il veicolo sopraggiunge alle spalle. Tale problema è particolarmente sentito dagli anziani e dalle persone con problemi di udito [38]. Un'altra questione legata alla sicurezza è sollevata dall'inserimento dei veicoli autonomi sulla strada. Anche se "l'autonomia" del veicolo dovrebbe attenuare gli effetti di errori e violazioni dei conducenti umani, i veicoli autonomi introducono nuovi dilemmi di sicurezza legati agli errori della macchina [39], [40]. Nonostante gli sforzi investiti per migliorare le capacità di questi veicoli, come il rilevamento, il riconoscimento e il *tracking* dei pedoni utilizzando sensori e altri dispositivi [41], [42], i rapporti sugli incidenti hanno dimostrato che i pedoni sono spesso coinvolti in collisioni con tali veicoli [43]. Diversi studi di ricerca sui veicoli autonomi hanno rivelato che la mancanza di

interazione con conducenti umani diminuisce i vantaggi di questa tecnologia in termini di sicurezza dei pedoni [44], [45].

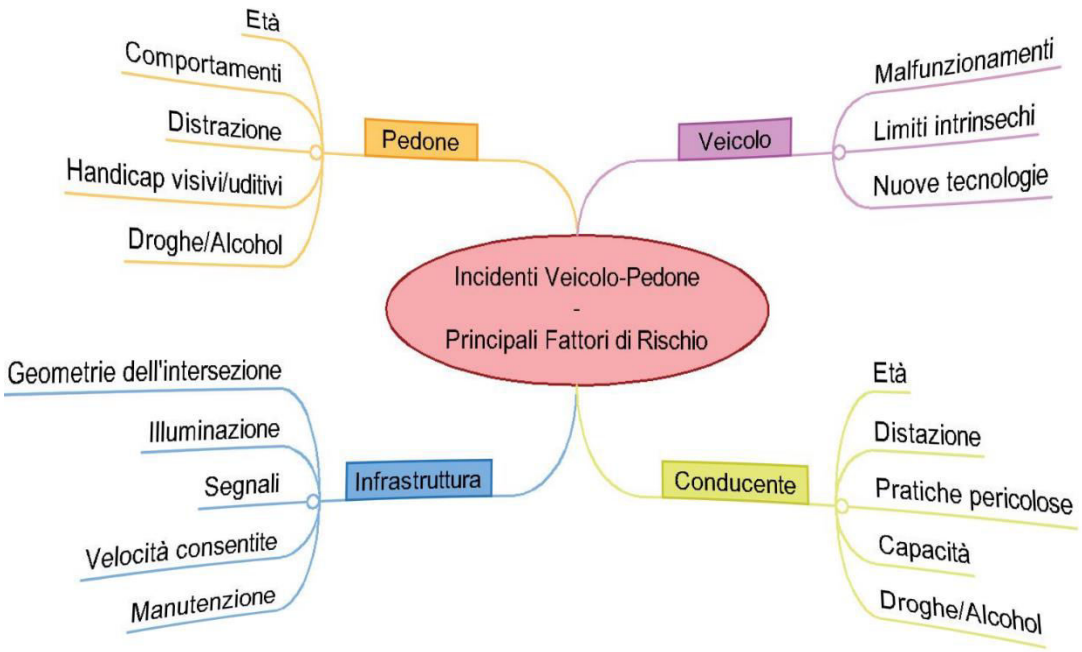


Figura 2: Illustrazione dei diversi fattori di rischio per la sicurezza dei pedoni.

Ad ogni modo, il basso rendimento di uno di questi fattori (Figura 2) causa instabilità nel ciclo interattivo tra gli altri e, in ultima analisi, aumenta la probabilità di un incidente. Siccome il fattore umano contribuisce in maniera significativa all'instabilità di questo ciclo, valutare il comportamento dei diversi utenti della strada al fine di individuare le giuste misure correttive può giocare un ruolo importante nel migliorare la sicurezza del sistema di trasporto e, di conseguenza, nel ridurre il rischio di incidenti.

Diversi studi in passato hanno dimostrato che le violazioni degli utenti sono indicatori significativi di coinvolgimento in incidenti stradali letali o gravi [33], [46], [47], oltre a rappresentare la minaccia più rilevante alla sicurezza dei pedoni [48]. Per questa ragione e siccome gli errori in genere non comportano conseguenze fatali o gravi per gli utenti [33], il termine "comportamenti rischiosi" o "a rischio" viene comunemente utilizzato come implicito riferimento alle violazioni. A tal proposito, alcuni ricercatori hanno osservato che caratteristiche della personalità, come la ricerca di sensazioni, e in particolare la ricerca del brivido e dell'avventura, hanno una forte associazione con comportamenti di guida rischiosi [33], [49], [50]. Alcune variabili demografiche, come l'età, il sesso, l'istruzione, la condizione fisica hanno una forte influenza sui comportamenti di conducenti e pedoni [46], [48], [51]. Per esempio, caratteristiche come la vista, i tempi di reazione, così come la migliore capacità di controllare il veicolo da parte dei giovani conducenti portano questo gruppo a mostrare una guida più pericolosa rispetto agli anziani. Diversamente, questi ultimi compensano i problemi derivanti dell'invecchiamento attraverso comportamenti positivi come la guida a una velocità inferiore rispetto al limite o ad una distanza maggiore dagli altri veicoli, ma allo stesso tempo commettono più errori alla guida rispetto ai

giovani [52]. In modo simile, gli uomini commettono più violazioni delle donne (indipendentemente dalla fascia di età considerata), ma queste sono più soggette ai lapsus alla guida rispetto agli uomini. I pedoni più giovani sono più propensi a commettere una violazione a parità di situazione da affrontare rispetto a quelli più anziani. Infatti, il pedone potrebbe essere visto in alcuni contesti come un "ostacolo", ovvero costituire un rischio per la sicurezza per sé stesso e per gli altri utenti della strada. Esempi di comportamento irrazionale dei pedoni includono l'attraversamento della strada da parte di un individuo in stato di ebbrezza o bambini che corrono dietro ai palloni [10]. Inoltre, come accennato, la sicurezza stradale dipende anche dai livelli di concentrazione del pedone e del guidatore. Un guidatore ha sempre dei carichi mentali dovuti alle conversazioni che intrattiene con i passeggeri, all'uso del telefono cellulare e/o all'elaborazione dei propri pensieri. Lo stesso vale per i pedoni a cui si aggiungono altre fonti di distrazione come occuparsi di bambini o di animali da compagnia al seguito, trasportare borse o pacchi magari pesanti o di grosse dimensioni e tante altre. Tuttavia, è principalmente l'uso del telefono cellulare, sia durante la guida sia mentre si cammina a lato strada, a porre seri problemi di sicurezza, oltre a rappresentare una grave violazione del Codice della Strada. La National Highway Transportation Safety Administration ha stimato che il 9.7% dei conducenti statunitensi, in un qualunque momento diurno della giornata, utilizza un telefono cellulare, tenendolo in mano o parlando in vivavoce [53]; per l'utilizzo del telefono alla mano, questa stima è del 3.2%, corrispondente a circa $472^{\cdot}486$ conducenti per lo più di età compresa tra i 16 e 24 anni. Pochi Paesi dell'UE conducono indagini sistematiche sull'uso del telefono cellulare da parte degli utenti, ma alcune indagini su strada hanno dimostrato che tra l'1% e l'11% (in Italia il 9%) dei conducenti usano il telefono durante la guida [54]. Questa abitudine (e in generale la distrazione alla guida) può compromettere le decisioni e le azioni degli utenti in diversi modi. Hatfield e Murphy [55] hanno analizzato l'effetto dell'uso dello smartphone sul comportamento dei pedoni durante l'attraversamento della strada; questi ricercatori hanno scoperto che i pedoni che parlano al telefono tendono a camminare lentamente, sono meno propensi a guardare il traffico prima di iniziare (o durante) l'attraversamento e ad aspettare che il traffico si fermi. Diversamente, le comuni prestazioni di guida alterate includono un peggioramento del controllo della velocità, difficoltà nel mantenere le appropriate traiettorie e distanze tra i veicoli, un aumento del tempo di reazione nel rispondere ai pericoli, un ritardo nella risposta ai segnali stradali, una minore capacità di percepire stimoli che provengono dalla zona periferica del campo visivo [56]–[59]. Per esempio, i tempi di reazione di un conducente sono del 30% più lenti durante una conversazione al telefono rispetto alla guida con un livello di alcohol nel sangue di $80\ mg/100\ ml$ (cioè al limite legale nel Regno Unito) e del 50% più lenti rispetto a condizioni di guida normali [60]. Infatti, la distrazione del guidatore, definita come una deviazione dell'attenzione ad un'attività concorrente [61], riduce le risorse cognitive assegnate al compito di guida. Un'analisi condotta utilizzando la risonanza magnetica funzionale ha dimostrato che la distrazione indotta dall'uso del telefono cellulare, il quale richiede l'elaborazione di stimoli uditivi complessi, altera in modo significativo l'attività cerebrale associata alla guida [62]. L'aumento del carico cognitivo dovuto ad una conversazione telefonica determina una minore attenzione all'ambiente stradale,

dove non tutte le informazioni che un guidatore osserva vengono elaborate, producendo una cecità da disattenzione [63]. In altre parole, il cervello umano compensa il carico cognitivo indotto dalla conversazione telefonica attraverso una mancata trasmissione alla memoria di lavoro di alcune informazioni visive, determinando una tendenza da parte del guidatore distratto a “vedere” ma non a “guardare” [64]. Quindi, il fattore negativo più importante associato all’uso del telefono cellulare durante la guida, sia in vivavoce che alla mano, è la deviazione dell’attenzione dalla guida alla conversazione stessa e l’entità degli effetti dipendono dalla complessità sia della conversazione che della situazione di guida [54]. A sostegno di questa affermazione, circa un terzo degli incidenti mortali in Giappone che hanno coinvolto un pedone sono causati da conducenti distratti [65], e analisi approfondite di dati di guida naturalistici hanno dimostrato che la causa principale delle collisioni tra veicoli e pedoni è la deviazione dell’attenzione del conducente su qualcosa di diverso dal pedone in pericolo [66].

Quindi, le diverse caratteristiche comportamentali e le capacità fisiologiche degli utenti della strada sono una delle ragioni principali delle differenze nelle loro prestazioni [67]–[70]. Tuttavia, le teorie psicologiche sul comportamento di guida [71], [72] sono concordi nel ritenere che l’utente, durante la marcia, effettua un continuo confronto tra il rischio accettato e quello percepito in funzione delle informazioni fornite dall’ambiente stradale e che, a seguito di tale confronto, adatta il proprio comportamento (ad esempio, regolando la velocità del veicolo o il margine di sicurezza) affinché la marcia avvenga al livello di rischi accettato [73]. Quindi, in questo processo di controllo del rischio e di adattamento delle condizioni di guida, le capacità dell’utente di anticipare e rilevare tempestivamente cambiamenti rilevanti nella situazione di traffico sono essenziali al fine di consentirgli la modulazione dei parametri di controllo della guida ed evitare l’adozione di comportamenti a rischio. Van Der Hulst e coll. [74] hanno dimostrato, in uno scenario di *car-following* (ovvero di veicoli che si susseguono su una strada), che i conducenti aumentano l’intervallo di tempo tra loro e l’auto che li precede se sono in grado di anticipare la decelerazione di tale veicolo. In altre parole, questi automobilisti sono riusciti a mantenere una distanza più sicura (corrispondente ad un minor rischio di tamponamento) rispetto a quei conducenti che non si aspettavano una frenata improvvisa dell’auto davanti a sé. Quindi, la probabilità di collisione è minore se i conducenti si comportano sulla base di una sufficiente anticipazione dei rischi attorno a sé [73]. Pertanto, se le caratteristiche del comportamento di guida associate a un basso (o addirittura alto) rischio di collisione fossero note per le diverse situazioni di traffico, la sicurezza di un evento potrebbe essere valutata confrontando il comportamento attuale misurato con queste caratteristiche comportamentali; di conseguenza, sarebbe possibile progettare soluzioni in grado di richiamare anticipatamente l’attenzione del conducente sulla situazione di potenziale conflitto, così come dotare i veicoli di sistemi in grado di intervenire autonomamente con le giuste misure preventive per evitare una collisione. Questo significa affrontare la sicurezza degli utenti non dalla prospettiva di una collisione, ma dalla prospettiva della probabilità di conflitto tra loro, attraverso studi comportamentali di automobilisti e utenti vulnerabili. Un tale approccio fornisce un contesto per migliorare la sicurezza dei pedoni durante l’attraversamento della strada e, in definitiva, adottare sistemi intelligenti per prevenire gli incidenti tra veicolo e pedone. Infatti, il principale

obiettivo di qualsiasi sistema di sicurezza del trasporto è quello di eliminare o ridurre i fattori che aumentano il rischio di incidenti. Quindi, studiare le diverse situazioni di traffico e le caratteristiche comportamentali degli utenti della strada è diventata la chiave per comprendere meglio le complesse condizioni che si creano tra i conducenti e i pedoni quando si incontrano e definire il modello causale del conflitto tra questi.

Tradizionalmente, le analisi della sicurezza stradale si sono basate principalmente sui dati storici relativi agli incidenti. I database sugli incidenti più dettagliati includono le velocità di viaggio e d'impatto dei veicoli, le manovre di frenata e di sterzata del conducente, nonché schizzi dettagliati delle scene dell'incidente. Combinando queste informazioni è possibile ricostruire l'evento di traffico e, per esempio, ricavare la posizione del pedone rispetto al veicolo in funzione del tempo durante la fase che precede l'incidente tra questi [75]. Tuttavia, nel corso degli anni sono stati discussi numerosi problemi (da un punto di vista analitico) associati ai dati sugli incidenti [76]. Per riassumere, le limitazioni più importanti includono 1) la loro frequenza relativamente bassa, che porta a problemi statistici relativi ai piccoli campioni di dati (gli incidenti sono eventi molto eccezionali, nel senso che sono il risultato di una serie di realizzazioni infelici di tante piccole probabilità legate al trinomio uomo-veicolo-ambiente), 2) la mancata denuncia o sottostima degli eventi incidentali, funzione della gravità dell'evento e del tipo di utenti della strada coinvolti, 3) le limitate informazioni (o l'assenza di informazioni nella maggior parte dei casi) in merito agli aspetti comportamentali precedenti l'incidente, 4) l'impossibilità di compiere analisi in tempo reale, siccome i dati incidentali sono tipicamente rilasciati ogni tre o sei mesi [76]–[78]. Nella pratica, questi problemi limitano le possibilità di trarre conclusioni sulla causalità degli incidenti e su come possano essere prevenuti in futuro [79]–[82]. Per superare queste limitazioni, alcuni ricercatori sostengono che le analisi sulla sicurezza stradale possono trarre beneficio da metodi, noti come *Traffic Conflict Techniques*, che utilizzano eventi di "quasi incidente" (o "quasi infortunio") come dato surrogato complementare agli incidenti veri e propri [76], [82]. Un quasi incidente si riferisce a un evento di traffico che ha il potenziale di evolvere in un incidente ma non porta ad alcuna conseguenza per le persone o danni alla proprietà, grazie alle azioni evasive (come cambiare traiettoria durante l'attraversamento, diminuire la velocità o sterzare per evitare una collisione) degli utenti della strada coinvolti [78], [83]. In altre parole, il traffico può essere visto come una serie di eventi elementari (ovvero interazioni tra due utenti della strada, in arrivo simultaneo su un punto specifico dell'infrastruttura) che differiscono nel loro grado di pericolosità e di gravità (ovvero, in una sola parola, di severità) per gli utenti coinvolti; inoltre, esiste una relazione tra la frequenza di accadimento e la severità degli eventi [77]. Diversi concetti che descrivono questa idea sono emersi nel corso degli anni. Per esempio, Hydén [84] descrive la relazione tra frequenza e rischio con una "piramide della sicurezza", dove la base della piramide è formata da normali incontri di traffico, abbastanza sicuri e frequenti; diversamente, la punta della piramide contiene gli eventi più gravi, come gli incidenti con feriti e/o morti, che sono altamente infrequenti (Figura 3). Tra le normali interazioni e gli incidenti ci sono i conflitti di traffico, che vengono utilizzati per descrivere gli eventi di quasi incidente [84]. A seconda del rischio per gli utenti, i conflitti possono essere classificati in gravi, lievi e potenziali [85].

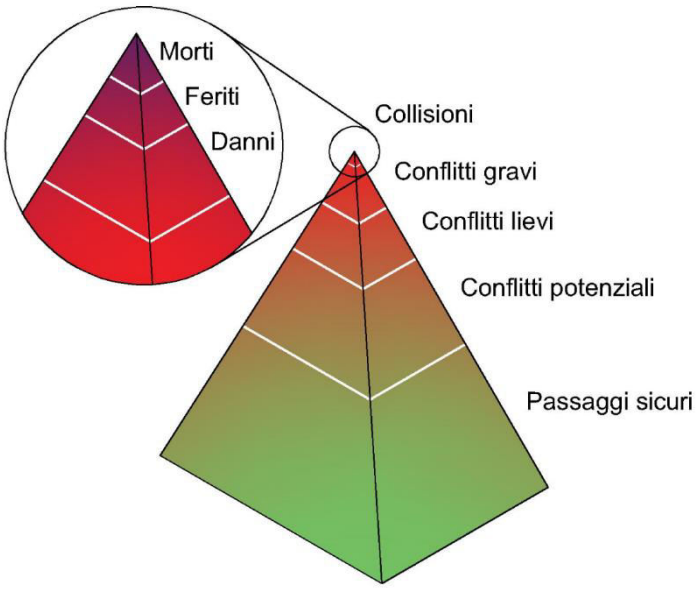


Figura 3: La piramide - le interazioni tra gli utenti della strada come un continuum di eventi di traffico.

Ad ogni modo, un terreno comune a tutti i concetti sviluppati è che i conflitti di traffico sono più frequenti degli incidenti e hanno un'interdipendenza con gli incidenti stessi che, una volta compresa, può essere usata per stimare il rischio e dedurre le cause degli incidenti stradali, senza dover osservare gli incidenti stessi. Esiste, infatti, un'ampia letteratura che dimostra l'esistenza di una forte correlazione tra la frequenza dei conflitti gravi (anche se definiti in modi diversi dagli autori) e la frequenza degli incidenti [86], [87]. In particolare, il modo più comune per l'analisi dei conflitti di traffico è quello di utilizzare Indicatori Surrogati di Sicurezza (SSM, dall'inglese *Surrogate Safety Measures*), che esprimono la gravità di un incontro di traffico come "vicinanza" a un incidente in termini di tempo o di spazio [76]. Nella pratica, queste metriche (dette anche *micro-level behavior indicators*) descrivono principalmente il movimento degli utenti della strada, ad esempio le loro traiettorie (come sequenze di posizioni), le velocità, le derivate della velocità (accelerazione, tasso di decelerazione ed altre) e tutte quelle variabili che possono essere calcolate da questo tipo di dati per comprendere più a fondo le relazioni tra comportamenti e rischio. A differenza dei dati sugli incidenti, che sono tipicamente gestiti e rendicontati dal Dipartimento dei Trasporti di ciascun Paese, i dati sui quasi incidenti richiedono uno sforzo aggiuntivo per la raccolta, la quale deve avvenire attraverso sensori (telecamere, RADAR e/o LiDAR) installati all'esterno dei veicoli o a bordo strada. Infatti, le analisi basate sui dati acquisiti con sensori possono migliorare notevolmente l'accuratezza e l'efficienza del processo rispetto alle osservazioni umane [88], garantendo una visione di dettaglio sulle interazioni. Questo aspetto è molto vantaggioso soprattutto se si considera che, rispetto agli incontri tra veicoli, le interazioni tra pedone e veicolo sono di solito più complesse: i pedoni si muovono a velocità relativamente basse e possono fermarsi, accelerare o cambiare rotta molto rapidamente rispetto ai veicoli [83], [89], [90]. Inoltre, negli ultimi anni, lo sviluppo della *Computer Vision* ha reso possibile l'applicazione

delle tecniche automatizzate di elaborazione video all'analisi dei conflitti, anche per la valutazione della sicurezza dei pedoni [91]–[93]. Ad ogni modo, attraverso l'uso della sensoristica è possibile acquisire ed elaborare i valori continui delle caratteristiche di mobilità degli utenti, generare i profili delle loro traiettorie in serie temporali ed, infine, prevedere le condizioni di rischio attraverso SSM [85].

Per modellare i dati delle traiettorie generate dai video o acquisite con sistemi RADAR o LiDAR, sussiste la necessità di un metodo in grado di elaborare dati con struttura temporale o sequenziale e lunghezza variabile sia per gli input che per gli output. Negli ultimi anni, le tecniche di *Deep learning* si stanno affermando nel campo dei trasporti per la risoluzione di vari problemi (dall'ottimizzazione di sistemi e processi fino al riconoscimento degli oggetti), tra cui quelli legati a dati sequenziali [94]–[96]. In particolare, le reti neurali ricorrenti (RNN, dall'inglese *Recurrent Neural Network*) sono modelli connessionisti che catturano la dinamica delle sequenze tramite cicli nella rete di connessioni [97]. A differenza delle reti neurali *feedforward* standard, le reti ricorrenti mantengono uno stato che può rappresentare informazioni da una finestra di contesto arbitrariamente lunga; in altre parole, queste reti sono capaci di passare selettivamente le informazioni attraverso gli step della sequenza, mentre elaborano i dati sequenziali un elemento alla volta. Inoltre, le RNN possono modellare simultaneamente dipendenze sequenziali e temporali su più scale. Anche se le RNN sono state tradizionalmente difficili da addestrare, e spesso contengono milioni di parametri, i recenti progressi nelle architetture di rete, nelle tecniche di ottimizzazione e nel calcolo in parallelo hanno permesso di migliorare su larga scala il processo di apprendimento. Negli ultimi anni, i sistemi basati su architetture LSTM (dall'inglese *Long Short-Term Memory*) [98] e GRU (*Gated Recurrent Unit*) [99] hanno dimostrato ottime prestazioni in compiti diversi come l'annotazione delle immagini, la traduzione linguistica e il riconoscimento della scrittura. In particolare, la GRU è una versione semplificata di LSTM, più facile da implementare e più rapida da addestrare: in pratica, vengono eliminate alcune componenti che non influiscono sulle prestazioni del modello e, diversamente, rendono la LSTM più complessa per il numero di parametri del modello da apprendere. Come conseguenza, la LSTM può ricordare sequenze più lunghe (cioè catturare dipendenze a più lungo termine nei dati delle serie temporali) rispetto alla GRU. Ad ogni modo, la GRU è un'opzione efficiente per modellare dati sequenziali ed è stata applicata recentemente nel campo dei trasporti per prevedere la situazione del traffico [100], il tempo di viaggio [101] e modellare gli eventi di quasi incidente nelle intersezioni regolate da semaforo [85].

Sebbene le tecniche di *Deep Learning* permettano di modellare la relazione causale tra comportamenti e rischio (o meglio, il rischio di collisione dalle caratteristiche del comportamento di guida), resta il quesito su come sfruttare questi modelli per prevenire gli incidenti, soprattutto tra pedoni e veicoli. È stato precedentemente accennato che se un conducente si comporta sulla base di una sufficiente anticipazione dei rischi, la probabilità di collisione è minore; inoltre, è noto che la gravità degli incidenti che coinvolgono pedoni può essere mitigata riducendo la velocità di collisione e, quindi, la forza di impatto [102], [103]. Infatti, esistono due vantaggi legati alla riduzione della velocità di collisione: da un lato, l'energia di collisione viene ridotta portando ad un minor rischio di lesioni gravi, dall'altro, l'impatto secondario quando il pedone colpisce il terreno viene

attenuato [75]. Pertanto, il miglioramento della sicurezza dei pedoni richiede 1) sistemi in grado di fornire in anticipo al conducente avvisi sulla situazione di traffico e, in casi eccezionali, di controllare i sottosistemi (freni, acceleratore e/o sterzo) del veicolo al fine di prevenire gli incidenti o almeno ridurre la velocità (al livello per cui un pedone non rischia di essere gravemente ferito), e 2) dispositivi in grado di ridurre la gravità delle lesioni in conseguenza ad un impatto inevitabile. Nella pratica, queste tecnologie a livello di veicolo sono chiamate, rispettivamente, sistemi di protezione attiva e passiva per il loro tempo di attivazione rispetto al potenziale incidente. In particolare, le soluzioni passive riguardano l'installazione sul veicolo di componenti in grado di assorbire gli urti (quindi, sono in grado di proteggere il pedone solo al momento della collisione). Per esempio, l'airbag per pedoni [104] e il sistema di sollevamento del cofano [105] proteggono la testa del pedone durante il contatto con il veicolo. Uno studio ha dimostrato che l'intervallo di confidenza al 95% del numero di vite salvate dal sollevamento del cofano è compreso tra 32.8 e 83.6 pedoni [106]. Diversamente, la tecnologia di protezione attiva consiste in sistemi avanzati di assistenza alla guida (ADAS, dall'inglese *Advanced Driver Assistance Systems*) in grado di aiutare il conducente nelle diverse situazioni di conflitto con altri utenti o di guida pericolosa, come sotto fatica o distrazione. Per esempio, in situazioni di potenziale conflitto, i sistemi *Forward Collision Warning* (FCW) possono essere efficaci nel dirigere l'attenzione del guidatore verso la situazione di traffico e quindi aiutare a prevenire o mitigare una collisione. I sistemi FCW fanno uso di interfacce uomo-macchina uditive, visive e tattili per avvisare il conducente della situazione di potenziale pericolo e permettergli di intraprendere una qualsiasi azione correttiva o evasiva. Alcuni studi hanno dimostrato che i sistemi FCW riducono i tempi di reazione dei conducenti in risposta agli eventi di traffico, le velocità di avvicinamento all'utente a rischio e, di conseguenza, il tasso di lesività e fatalità degli incidenti [107], [108]. Tuttavia, poiché il conducente ha bisogno di tempo per reagire dopo l'avvertimento, le ipotesi sui tempi di reazione e gli input di risposta (che riflettono il vero comportamento dei conducenti) sono essenziali per valutare l'efficacia dei sistemi FCW [109], [110]. Inoltre, è fondamentale bilanciare l'efficacia in termini di tempo dell'avviso con l'accettazione del sistema da parte del conducente. Più invasivo e precoce è l'avviso, più efficacemente un sistema può prevenire le collisioni (ovvero, da un altro punto di vista, prima e più forte sono applicati i freni, maggiore è la riduzione della velocità); ma allo stesso tempo, l'accettazione di tali sistemi da parte del conducente diminuisce, ed è stato osservato che in tal caso i conducenti sono più propensi a ignorare il sistema o a spegnerlo [111]. Un sistema FCW può quindi agire come un *Imminent Crash Warning* con un'attivazione piuttosto tardiva, mirando a una reazione immediata del conducente, o come un *Cautionary Crash Warning* con un'attivazione precoce, con l'obbiettivo di dirigere l'attenzione sull'evento di traffico e ad aumentare la consapevolezza del rischio; altrimenti, potrebbe incorporare entrambe le funzioni in un approccio a *triggers* [112], [113]. Ad ogni modo, questi sistemi richiedono non solo un riconoscimento preciso del pedone, ma anche una stima affidabile del rischio di collisione [75]: è fondamentale raggiungere un equilibrio accettabile tra un'alta efficacia del sistema e un basso tasso di false attivazioni. Lo stesso discorso rimane valido anche per l'*Automated Emergency Braking* (AEB), in grado di prevenire o mitigare le collisioni frenando

l'auto automaticamente senza l'intervento del guidatore. Proprio per la natura automatizzata di questi sistemi (che appunto non richiedono alcun input da parte del conducente), molta attenzione e seri sforzi sono stati dedicati al loro sviluppo [38], [114]. Questi sistemi si basano su unità di rilevamento, come telecamere, LiDAR, RADAR, laser scanner o un insieme di più dispositivi, per percepire l'ambiente che circonda il veicolo [115]. Un sistema di bordo analizza i segnali raccolti dai sensori, rileva gli ostacoli esistenti nella traiettoria del veicolo e li classifica sulla base di modelli di aspetto. Se viene riconosciuto un pedone, il sistema può utilizzare modelli di comportamento del pedone per prevederne la reazione e la traiettoria, oppure stimare il livello di rischio corrente attraverso SSM. Se è probabile una collisione con il pedone sulla base delle traiettorie predette oppure si verifica il superamento di determinate soglie critiche per la sicurezza ed, in entrambi i casi, il conducente non intraprende alcuna azione evasiva efficace, il sistema AEB prende il controllo attivo dei freni e, in alcuni casi, di altri sottosistemi del veicolo (volante, acceleratore, sospensione, ecc.) per mettere in atto le azioni necessarie, ad esempio decelerare, cambiare corsia o frenare bruscamente [116]. Lo studio di questi sistemi ha auspicato una significativa riduzione degli incidenti mortali, della gravità delle lesioni per gli impatti inevitabili e dei costi sociali [75], [117].

Numerosi sistemi FCW e AEB, progettati con lo scopo di proteggere i pedoni, sono stati recentemente introdotti sul mercato automobilistico [111]. Tuttavia, questi sistemi sono generalmente costituiti da tre sottosistemi o livelli, chiamati "percezione", "decisione" ed "esecuzione", che possono avere prestazioni molto diverse a seconda del fornitore che li ha sviluppati e della tecnologia di sensori utilizzata per acquisire i dati della scena [2]. Per questi motivi e per garantire un funzionamento adeguato in una vasta gamma di scenari di traffico, i sistemi FCW e AEB sono in continuo sviluppo. Ad ogni modo, i sistemi esistenti si basano su livelli di decisone, ovvero modelli di valutazione del rischio (RAM, dall'inglese *Risk Assessment Model*), che valutano istante per istante la situazione corrente di traffico tratta dai sensori e, in base a prestabiliti tempi di reazione dei conducenti, decidono l'azione da mettere in campo (come, per esempio, l'avviso uditivo o la frenata di emergenza, concretizzate dal livello di esecuzione) [65]. Diversamente, come accennato in precedenza, alcuni ricercatori sostengono che se i veicoli fossero in grado di prevedere i comportamenti dei conducenti e degli utenti della strada vicini con un solo secondo di anticipo, la maggior parte degli incidenti stradali potrebbe essere evitata [73], [118]. In altre parole, se i veicoli fossero in grado di preludere la probabile evoluzione della situazione attuale di traffico e valutare quanto pericolosa quella situazione futura potrebbe essere, sarebbe possibile prendere delle decisioni di guida appropriate in anticipo, come regolarizzare le traiettorie agendo sui sottosistemi del veicolo (anche in forma automatica) prima che l'interazione diventi grave oppure adattare i sistemi di assistenza al comportamento del conducente (cioè, per determinare, se e quando avviare o interrompere l'esecuzione di un intervento, evitando interferenze di guida non necessarie). Questo approccio richiede in genere modelli di previsione del movimento degli utenti (*Motion Prediction Model*) che devono essere abbinati ad un RAM: nella pratica, la valutazione del rischio si divide in una prima fase di previsione delle potenziali traiettorie future per tutti gli utenti in movimento sulla scena e una seconda fase di stima del rischio basata

sulla probabilità di collisione tra ogni possibile coppia di traiettorie. Tuttavia, i modelli di movimento possono essere anche molto difficili da definire per la complessità delle interazioni che caratterizzano il traffico [119]. Per questa ragione, la maggior parte dei sistemi ADAS di massa non ha una capacità predittiva a medio termine delle intenzioni degli utenti della strada [120] e, diversamente, sono progettati per agire in modo reattivo in situazioni ad alto rischio (in particolare i sistemi AEB).

Scopo della tesi

In questa tesi viene presentato un innovativo sistema di valutazione dei rischi che predice, fino a tre secondi nel futuro, il livello di pericolosità degli incontri veicolo-pedone nel traffico urbano (cioè, gli incontri tra un'auto e un pedone su un passaggio pedonale). Tale sistema, implementato da un modello GRU, definisce la situazione futura dell'interazione sulla base della rappresentazione della scena attuale tratta dai segnali di tre RADAR esterni al veicolo. Un esperimento di simulazione di guida è stato progettato per raccogliere caratteristiche di mobilità sequenziali su un campione di 65 studenti universitari e generare una coppia di indicatori SSM che permettono l'annotazione accurata delle condizioni di sicurezza dei pedoni durante l'intero processo di incontro. Il campione è composto da 19 femmine e 41 maschi (cinque partecipanti sono stati esclusi dalle analisi, in quanto non hanno potuto completare l'esperimento per l'eccessivo disagio), di età compresa tra 20 e 30 anni, in possesso di una patente di guida in corso di validità, che percorrono almeno $5000\ km$ all'anno alla guida di un veicolo. Sulla base delle soglie critiche per la sicurezza tratte dalla più recente letteratura scientifica, il modello presentato può essere utilizzato per potenziare i sistemi FCW e AEB esistenti e contribuire significativamente al miglioramento della sostenibilità dei mezzi di trasporto.

Differenziazione rispetto alla letteratura scientifica

L'algoritmo prototipato differisce dalla recente letteratura in materia per la capacità di predire direttamente il livello di rischio delle interazioni veicolo-pedone senza ricorrere alla previsione delle traiettorie degli utenti, il cui processo di generazione è computazionalmente costoso e poco compatibile con la valutazione del rischio in tempo reale [119]. Inoltre, per la prima volta la predizione del livello di rischio tra veicolo e pedone viene effettuata su più orizzonti temporali nel futuro (precisamente a $1, 2$ e $3\ s$, rispetto alla situazione corrente rilevata dai sensori) e attraverso una rappresentazione minima ed efficiente della situazione attuale di traffico. Infatti, il sistema dipende soltanto da sei variabili relative alle caratteristiche di mobilità del conducente e alle proprietà della scena di traffico al tempo presente, estratte dai dati grezzi dei sensori esterni al veicolo: la primitiva di comportamento del conducente (una variabile categorica che descrive il comportamento elementare a cui può essere ricondotta la manovra del conducente in avvicinamento alle strisce pedonali), il tempo atteso di arrivo del conducente al potenziale punto di conflitto, la velocità del veicolo, la velocità del pedone e le coordinate polari (angolo azimutale orizzontale e distanza relativa) del pedone rispetto al fronte del veicolo.

Principali contributi alla conoscenza scientifica

Questa tesi raccoglie il lavoro di quattro articoli scientifici, due dei quali (Articoli B e D) sono stati pubblicati su riviste internazionali e i restanti come atti di conferenze (Articoli A e C). Di seguito, vengono riportati gli estremi delle quattro pubblicazioni:

Articolo A:
N. Baldo, A. Marini e M. Miani, «Effects of Cognitive Distraction on Driver's Stopping Behaviour: A Virtual Car Driving Simulator Study», *IOP Conference Series: Materials Science and Engineering*, vol. 960, n.2, ID 022082, dic. 2020, doi: 10.1088/1757 − 899X/960/2/022082

Articolo B:
N. Baldo, A. Marini e M. Miani, «Drivers' Braking Behavior Affected by Cognitive Distractions: An Experimental Investigation with a Virtual Car Simulator», *Behavioral Sciences*, vol. 10, n. 10, ID 150, ott. 2020, doi: 10.3390/bs10100150

Questi primi due lavori hanno contribuito alla validazione del simulatore di guida (presentata al Capitolo 3) per lo studio dei comportamenti dei conducenti e fornito una maggiore comprensione degli effetti indotti da una conversazione telefonica attraverso altoparlanti sulla guida di giovani utenti della strada.

Articolo C:
M. Miani, M. Dunnhofer, C. Micheloni, A. Marini e N. Baldo, «Young drivers' pedestrian anti-collision braking operation data modelling for ADAS development», *Transportation Research Procedia*, vol. 60, pagg. 432 − 439, gen. 2022, doi: 10.1016/j. trpro. 2021.12.056

Articolo D:
M. Miani, M. Dunnhofer, C. Micheloni, A. Marini e N. Baldo, «Surrogate Safety Measures Prediction at Multiple Timescales in V2P Conflicts Based on Gated Recurrent Unit», *Sustainability*, vol. 13, n. 17, ID 9681, ago. 2021, doi: 10.3390/su13179681

A differenza dei precedenti lavori, questi ultimi hanno contribuito alla concettualizzazione, sviluppo ed implementazione del modello di valutazione del rischio, oggetto del presente lavoro di tesi.

Capitolo 2

Comportamenti e Rischio

Nel Capitolo 1, la sicurezza dei diversi utenti della strada è stata attestata come fondamento di ogni sistema ADAS. Un concetto strettamente correlato è quello di rischio, che può essere intuitivamente inteso come la probabilità e la gravità del danno (ovvero delle lesioni) che un veicolo (un pedone) di interesse può subire nel prossimo futuro. Da questa definizione è chiaro che, per valutare il rischio associato ad una particolare situazione, è necessario disporre di modelli matematici che permettano di prevedere come questa situazione si evolverà in futuro.

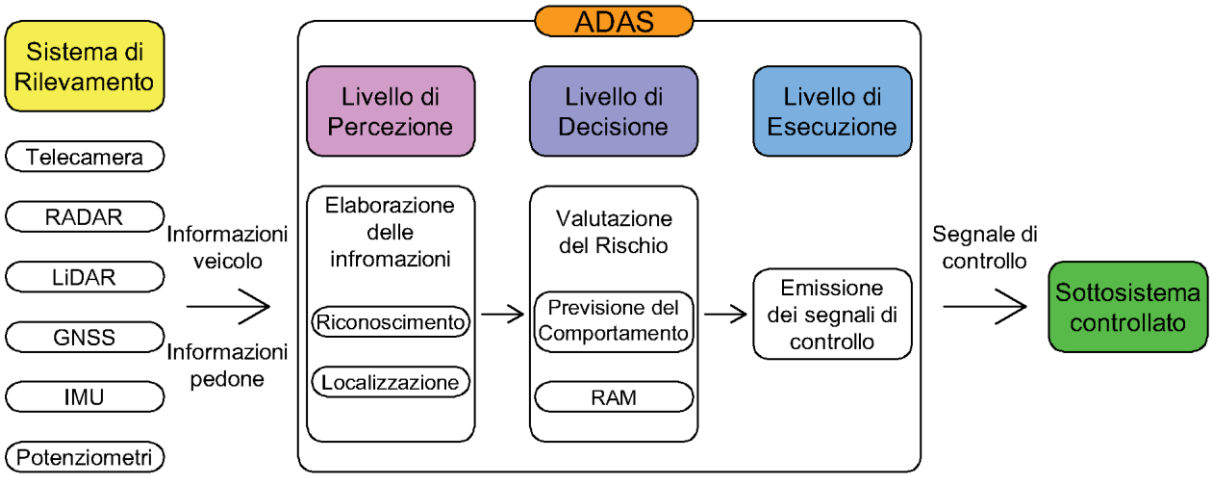


Figura 4: Struttura di un sistema ADAS.

Inoltre, è stato accennato che i sistemi FCW e AEB sono generalmente costituiti da tre sottosistemi o livelli, chiamati “percezione”, “decisione” ed “esecuzione”, che lavorano in sequenza. Come mostrato in Figura 4, il livello di percezione riceve i segnali grezzi dal sistema di rilevamento esterno al veicolo e li elabora per riconoscere e localizzare i diversi utenti che partecipano alla situazione di traffico. Diversamente, il livello di decisione fornisce una comprensione delle dinamiche tra gli utenti, ovvero determina il livello di rischio sulla base delle informazioni ottenute e formula un giudizio sullo stato di sicurezza presente e/o futuro, il quale detta l’azione da attuare. L’ultimo livello, quello di esecuzione, è responsabile dell’emissione dei segnali di controllo (per esempio, i segnali di avvertimento o di frenata) corrispondenti al livello di rischio stimato e della loro trasmissione ai diversi sottosistemi di controllo (ovvero il sistema di preallarme o il sistema di frenata del veicolo).

Esistono diverse possibili configurazioni del livello di decisione che dipendono dall’esigenza o meno di anticipare l’evoluzione futura degli eventi e da come questa previsione viene effettuata. Infatti, come già accennato, la maggior parte dei sistemi FCW e AEB attualmente in commercio non anticipa gli eventi ma si basa sulla valutazione della situazione corrente di traffico, tratta dai sensori, per poter agire in modo reattivo in situazioni ad alto rischio; in altre parole, il

sistema viene attivato solo quando il livello stimato di rischio della scena attuale di traffico supera una determinata soglia critica per la sicurezza. Ad ogni modo, il livello di decisione di qualsiasi sistema ADAS si basa sempre su un modello di valutazione del rischio (RAM), che può dipendere o meno dalla previsione delle "traiettorie", "manovre" o "comportamenti" (questi termini sono usati nella letteratura scientifica in modo intercambiabile per indicare le posizioni, presenti e future, degli utenti) [121]–[123]. Nel caso di anticipazione degli eventi, si parla di "*Risk based on colliding future trajectories*" quando il rischio di una situazione (inteso come rischio di collisione) è stimato prevedendo le traiettorie future degli utenti attraverso un modello di movimento, e di "*Risk based on unexpected behavior*" quando il rischio è legato al rilevamento delle intenzioni di manovra o degli atteggiamenti dei conducenti che deviano dal comportamento nominale atteso (in base al contesto) sulla strada [119].

Nel seguito di questo Capitolo viene descritto, al più alto livello di astrazione, lo stato dell'arte sulle tecniche di modellazione del movimento degli utenti e di valutazione del rischio, per consentire al lettore una migliore comprensione del contesto scientifico in cui il presente lavoro di tesi si inserisce. In particolare, la trattazione si concentra sui metodi di *Deep Learning* e di rappresentazione dei dati di input comunemente adottati per la predizione delle intenzioni e/o traiettorie degli utenti. Sul tema della valutazione della sicurezza, l'attenzione si concentra principalmente sugli indicatori SSM che appartengono alla categoria "*Risk based on colliding future trajectories*".

Le recenti tendenze delle pubblicazioni nella letteratura scientifica sui sistemi ADAS e di guida autonoma mostrano che dal 2018 il numero di studi che utilizzano tecniche di apprendimento è in forte aumento [124]. Una tra le molteplici ragioni di questo fatto è il minor sforzo richiesto per la progettazione del sistema, soprattutto quando si lavora con scenari intrinsecamente complessi, come il traffico urbano [125]. Il prezzo da pagare per questo vantaggio è un aumento del lavoro inziale per la configurazione della struttura del modello e la raccolta dei dati per il *training* del sistema, ovvero la fase di formazione per l'apprendimento dei parametri che definiscono la funzione del modello. Sebbene esistano diversi sistemi che utilizzando modelli matematici ed euristiche per stimare le traiettorie degli utenti della strada, il numero di situazioni e/o comportamenti in ambienti complessi è diventato così elevato che gli approcci di apprendimento, in grado di incorporare sia l'incertezza sulle traiettorie future che le interazioni tra gli agenti, continueranno ad affermarsi come tecnica predominante nella modellazione degli eventi di traffico. Ad ogni modo, la letteratura annovera un gran numero di strategie alternative per apprendere le relazioni tra un insieme di caratteristiche estratte dall'elaborazione di dati sensoriali (cioè acquisiti da telecamere, RADAR, etc.) e il comportamento degli utenti.

Il Capitolo 2 è dedicato a queste tematiche perché molti aspetti della modellazione effettuata in questa tesi richiamano approcci di apprendimento frequentemente utilizzati per la previsione dei comportamenti di pedone o veicolo target (ovvero, il cui comportamento sulla scena è di interesse). Tuttavia, sebbene entrambi i metodi siano finalizzati all'anticipazione degli eventi futuri, vale la pena ricordare che lo scopo della tesi non è lo sviluppo di un modello di movimento degli utenti ma la modellazione del livello di rischio nelle interazioni

veicolo-pedone sulla base della rappresentazione attuale della situazione di traffico. In altre parole, l'obbiettivo è creare un sistema che, invece di predire le singole traiettorie degli utenti interagenti per stimare la probabilità di collisione tra queste, apprenda implicitamente i modelli di movimento, osservando come evolve l'interazione tra gli agenti dinamici sulla scena (ovvero, il conducente e il pedone), e deduca il rischio di tale situazione. In questo modo, si evita l'uso di modelli classici di movimento che, diversamente, introducono ipotesi semplificative sul problema di interazione, data la complessità del comportamento umano e la varietà di stimoli interni ed esterni alla base del movimento [124]. Tuttavia, siccome la valutazione accurata del rischio richiede il calcolo di indicatori SSM, che dipendono dalle posizioni relative e dalle direzioni di movimento di pedone e veicolo, è necessario osservare come generalmente gli utenti interagiscono e a posteriori, sulla base dei comportamenti osservati, ricostruire l'evoluzione dei livelli di sicurezza durante l'intero processo di incontro; quindi, attraverso tecniche di apprendimento supervisionato, il legame tra i livelli di rischio e le caratteristiche della situazione di traffico, in diversi istanti di tempo o orizzonti temporali, può essere acquisito direttamente dai dati e questa conoscenza appresa può essere sfruttata per facilitare l'adattamento del sistema a nuove situazioni operative non considerate nello sviluppo del modello. In questi termini, il sistema proposto basa il suo funzionamento sul paradigma *Sense–Reason–Predict* (Sezione 2.1.3), e può essere facilmente adattato per un apprendimento online (ovvero su un veicolo in movimento), attraverso l'acquisizione di informazioni in tempo reale e la loro post-elaborazione. Quindi, sebbene le traiettorie $2D$ per pedone e veicolo non vengano considerate nella modellazione, il modello prototipato (il quale rappresenta a tutti gli effetti un innovativo RAM) costituisce un approccio alternativo a quelli presentati in questo Capitolo ed ibrido, siccome è in grado di valutare il comportamento per inferire il livello di rischio futuro (come avviene nei metodi di "*Risk based on unexpected behavior*").

2.1 Modellazione del Comportamento

I conducenti di un veicolo compiono traiettorie che sono principalmente governate dalle regole di guida e dalle situazioni ambientali. Inoltre, a differenza dei pedoni che hanno la capacità di fermarsi o accelerare rapidamente, questi utenti non possono cambiare le loro traiettorie istantaneamente in quelle desiderate (a causa dalla maggiore inerzia dei veicoli). Tuttavia, la previsione del comportamento dei veicoli non è un compito banale a causa di diverse sfide. In primo luogo, c'è un'interdipendenza tra i comportamenti dei conducenti, nel senso che il comportamento di un veicolo influenza quello degli altri veicoli sulla strada e viceversa. Pertanto, prevedere le traiettorie di un veicolo richiede l'osservazione di tutti i veicoli circostanti. In secondo luogo, la geometria della strada e le regole del traffico possono modificare le azioni dei conducenti. Per esempio, il posizionamento del segnale di precedenza ad un incrocio determina un comportamento del conducente in avvicinamento diverso rispetto a quello imposto da un cartello di stop. Quindi, senza considerare questi elementi, un modello sviluppato in un ambiente di guida specifico avrebbe prestazioni limitate in altri contesti. In terzo luogo, il comportamento futuro dei veicoli è multimodale,

ovvero possono esistere all'istante corrente diverse possibili manovre future per un veicolo, data la sua storia di movimento pregressa (cioè le posizioni agli istanti precedenti a quello attuale). Per esempio, quando un veicolo rallenta ad un incrocio senza cambiare la sua direzione di marcia, sono possibili nel caso più semplice sia manovre di svolta a destra che a sinistra.

Un'altra parte altrettanto importante del problema di anticipazione proattiva delle traiettorie degli utenti della strada è la previsione del comportamento dei pedoni (o in generale, degli utenti della strada vulnerabili). La sfida posta dall'esigenza di produrre previsioni accurate del movimento dei pedoni deriva dalla complessità del comportamento umano e dalla varietà dei suoi stimoli interni ed esterni. Infatti, in analogia a quanto osservato per il comportamento dei veicoli sulla strada, il movimento delle persone può essere guidato dagli obbiettivi individuali, dalla presenza e dalle azioni degli agenti circostanti (il termine agente indica gli elementi dinamici di interesse sulla scena come pedoni, ciclisti, automobili etc.), dalle relazioni sociali tra gli agenti, dalle regole e norme sulla circolazione, o dall'ambiente con la sua topologia, geometria e semantica. Inoltre, la maggior parte di questi fattori non sono direttamente osservabili e devono essere dedotti da informazioni di contesto.

Oltre alle sfide intrinseche del problema di predizione del movimento, l'implementazione di un tale modulo nei veicoli presenta diverse limitazioni pratiche. Per esempio, i computer di bordo hanno risorse computazionali limitate e i veicoli moderni possono osservare solo parzialmente l'ambiente circostante a causa dell'occlusione da ostacoli o della portata, campo visivo e precisione della sensoristica di bordo. Per superare questi limiti, molti degli studi esistenti presuppongono l'accesso ad un'ampia vista dall'alto senza ostacoli dell'ambiente di guida (si parla di *Bird's Eye View* nel contesto scientifico), che può essere ottenuta dai sensori dell'infrastruttura (per esempio da telecamere di sorveglianza dell'infrastruttura). Nella realtà, questo dato può essere disponibile solo se esiste un canale di comunicazione con capacità sufficiente tra l'infrastruttura e il veicolo ma, d'altra parte, non è conveniente coprire ogni tratto dell'infrastruttura con tali sensori. Quindi, un modulo di previsione del comportamento non può fare sempre affidamento sulla *Bird's Eye View* di un sensore dell'infrastruttura.

Di conseguenza, per quanto esposto, un modulo completo di previsione del comportamento in un veicolo dovrebbe identificare tutti i possibili movimenti futuri degli agenti, essere robusto e operare in tempo reale per permettere al sistema ADAS o di guida autonoma, che lo implementa, di agire nella pratica in modo affidabile ed efficace.

Prima di procedere alla discussione sui metodi di predizione del comportamento, è importante chiarire alcuni termini e concetti alla base di tale problema. Il termine "agente target" viene usato per indicare l'utente (per esempio il pedone) o l'elemento dinamico (il veicolo) del quale si vuole prevedere il comportamento. L'ego-veicolo (EV) è il veicolo che osserva l'ambiente circostante per prevedere il comportamento dell'agente target. I "veicoli circostanti" sono i veicoli il cui comportamento viene considerato dal modello di previsione in quanto possono potenzialmente influenzare il comportamento futuro del veicolo target. Studi diversi possono adottare criteri differenti per selezionare questi veicoli in base alle loro ipotesi di modellazione. Si assume che

il comportamento dell'agente target sia regolare e diretto all'obbiettivo, altrimenti il problema della predizione del movimento diventa molto arduo o addirittura mal posto. Si definisce "percorso" una sequenza di posizioni (x, y) e "traiettoria" un percorso associato a una legge temporale o a un profilo di velocità. Si parla di previsione "a breve termine" e "a lungo termine" per caratterizzare, rispettivamente, orizzonti di previsione di 1 - $2\ s$ e fino a $20\ s$ nel futuro.

In generale, il problema della predizione del movimento si articola nei seguenti tre elementi: rappresentazione della scena, approccio di modellazione e tipo di output predetto. Nelle Sezioni 2.1.1, 2.1.2 e 2.1.3 vengono descritti alcuni aspetti che possono caratterizzare questi tre elementi e che hanno avuto una forte influenza sullo sviluppo del modello presentato in questa tesi.

2.1.1 Rappresentazione della scena

I dati di input ad un modello di movimento definiscono il modo in cui viene rappresentata la situazione di traffico e, quindi, gli stimoli interni ed esterni che determinano il comportamento degli agenti sulla strada. La maggior parte degli algoritmi utilizza la posizione e la velocità attuali del pedone target [126]–[128], spesso considerando anche la storia degli stati recenti. Queste informazioni sono fornite da un sistema di tracciamento (*tracking*) degli agenti che ne assicura una corretta identificazione durante l'intero periodo di osservazione. Ovviamente, lo stato attuale dell'agente può essere acquisito solo se esso è osservabile dai sensori dell'EV. Posizione e velocità (talvolta assieme ad accelerazione e orientamento) sono anche i principali attributi dell'agente target nei compiti di previsione del movimento del veicolo [129], [130].

I metodi per prevedere l'attraversamento dei pedoni [131], [132] e il movimento dei ciclisti [133] trattano tipicamente ogni agente individualmente. Tuttavia, il più delle volte, diversi agenti condividono lo stesso ambiente stradale, e adattano le proprie azioni in base alla presenza e al comportamento degli altri. Pertanto, per prevedere in maniera più efficace il movimento è utile considerare l'interazione tra gli agenti in movimento sulla scena. Un approccio convenzionale è quello di fornire esplicitamente al modello di previsione la storia degli stati recenti dell'agente target e di quelli circostanti ad esso. In questo caso il problema si sposta su come identificare i veicoli o pedoni sulla scena che hanno un impatto effettivo sull'agente target. Alcuni ricercatori [134]–[136] utilizzano una distanza limite, ovvero solo le interazioni dei veicoli entro una certa distanza sono considerate nel modello di previsione. Altché e de La Fortelle [137] considerano solo i veicoli che possono essere rilevati senza occlusioni dal LiDAR dell'EV e utilizzano un diverso set di caratteristiche per il veicolo target (corrispondente all'EV, trattandosi di un problema di guida autonoma) e i veicoli circostanti al fine di replicare le informazioni su cui un guidatore umano basa le sue decisioni. In particolare, gli attributi degli agenti rilevati sono relative all'EV (ovvero, la velocità longitudinale relativa al target, le distanze laterale e longitudinali dal target, il *Time-to-Collision* con il target), poiché i conducenti solitamente prendano decisioni basate sulle distanze percepite e sulle velocità relative piuttosto che sui loro valori in un riferimento assoluto. Inoltre, la decisione di includere esplicitamente il *Time-to-Collision* (uno dei più importanti SSM, il quale quantifica il tempo rimanente prima che avvenga una collisione tra due utenti se le loro traiettorie e velocità rimanessero immutate) come caratteristica dell'interazione

deriva dalla rilevanza di questa metrica sulle decisioni del conducente, in questo caso per il cambio di corsia [138]; infatti, le neuroscienze mostrano che il cervello animale e umano si basa in maniera rilevante sulle stime del *Time-to-Collision* per eseguire compiti motori [139]. Similmente, Kooij e coll. [140] hanno dimostrato che è possibile raggiungere predizioni più accurate sul percorso degli utenti vulnerabili della strada (pedoni e ciclisti, i quali possono cambiare rapidamente il loro stato dinamico) quando vengono fornite informazioni di contesto al modello. In particolare, gli autori sostengono che la decisione di un pedone di fermarsi sul marciapiede prima di attraversare può essere prevista da tre indizi, estratti da video attraverso *Computer Vision*: l'esistenza di un veicolo su una rotta di collisione (in questo caso viene utilizzato, come indicatore di interazione, la distanza minima tra il pedone e il veicolo se le loro traiettorie fossero estrapolate nel tempo con velocità fissa), la consapevolezza del pedone della situazione di potenziale pericolo (riscontrata attraverso rilevamento dell'orientazione del volto) e la sua posizione relativa rispetto al bordo del marciapiede. Inoltre, diverse applicazioni di *Machine Learning* (che utilizzano LSTM) hanno dimostrato che è anche possibile imparare implicitamente come più agenti si muovono in presenza di altri [141], [142], anche dalla prospettiva del veicolo [143], [144].

Diversamente, gli approcci chiamati "*End-to-End"* si basano su sequenze di dati grezzi acquisiti dai sensori. In questo modo, i dati di input al modello di previsione contengono tutta la conoscenza disponibile sull'ambiente circostante. Questo permette ad un modello basato sull'apprendimento di imparare estraendo le caratteristiche utili da tutti i dati sensoriali disponibili. I segnali grezzi, rispetto alle rappresentazioni di input precedentemente esposte, hanno spesso una dimensione maggiore (nel senso che il modello *End-to-End* ha molte più variabili in ingresso, rispetto ad un approccio standard). Questo comporta maggiori risorse computazionali per la loro elaborazione e, conseguentemente, una minor praticità per l'implementazione a bordo dei veicoli. Una soluzione a questo problema è quella di ripartire le risorse computazionali tra più funzioni di uno stesso modello, ovvero i predittori *multi-task*. Infatti, nel contesto del *Deep Learning*, è comune addestrare un modello per più compiti [145]. Inoltre, in un veicolo dotato di sistemi di automazione, il modulo di rilevamento delle entità dinamiche (*object detection*) sfrutta i dati grezzi dei sensori e, di solito, si basa su un modello con milioni di parametri [146]. Pertanto, esso rappresenta un buon candidato per la condivisione di parametri con il modulo di previsione del comportamento. Luo e coll. [147] usano una rete neurale profonda (*Deep Neural Network*) per risolvere congiuntamente i problemi di rilevamento $3D$, tracciamento e previsione del movimento su un veicolo autonomo. Gli autori sfruttano le nuvole di punti $3D$ acquisite da un sensore su diversi fotogrammi temporali per generare un'immagine *Bird's Eye View* dell'ambiente, che è molto efficiente sia in termini di memoria che di calcolo. Lo stesso approccio è utilizzato da Casas e coll. [148]; tuttavia, questi autori hanno sviluppato un predittore multi-task che sfrutta sia le nuvole di punti $3D$ prodotte da un sensore LiDAR, sia le mappe dinamiche dell'ambiente (ovvero una *Bird's Eye View*), per ridurre il tempo di reazione nelle applicazioni di guida autonoma.

2.1.2 Output predetto

Metodi diversi producono rappresentazioni differenti del comportamento futuro dell'agente target, come output del modello di previsione. Tuttavia, le tecniche esistenti per l'apprendimento del comportamento possono essere divise in due macro-classi: metodi di classificazione e di regressione.

I problemi di classificazione riguardano la previsione delle intenzioni di movimento (o di manovra), talvolta chiamate anche "primitive di comportamento" [125]. La previsione dell'intenzione ha il compito di stimare quale manovra (ovvero la classe di manovra) il conducente di un veicolo intende fare nei prossimi intervalli di tempo [119], oppure quale tra i modi dinamici (per esempio, attraversare, fermarsi o andare nella stessa direzione) un pedone intende assumere nel prossimo futuro [149]. Per quanto riguarda i veicoli, le primitive segmentano un comportamento di guida complesso in una sequenza di intenzioni di base: per esempio, nella guida in autostrada, l'insieme delle manovre potrebbe essere il mantenimento della corsia e il cambio di corsia a sinistra/destra; mentre, in una intersezione, potrebbe essere svoltare a sinistra/destra e andare dritto. In quest'ultimo contesto, Khairdoost e coll. [150] hanno implementato un sistema di *Deep Learning* che può anticipare (di $3.6\ s$ in media) le manovre del conducente (cambio di corsia a sinistra/destra, svolta a sinistra/destra e andare dritto), sfruttando la direzione dello sguardo e la posizione della testa del conducente, assieme alle informazioni sullo stato dinamico del veicolo. Ortiz e coll. [125] hanno diviso il comportamento dei conducenti in avvicinamento ad un semaforo nelle tre primitive di mantenimento della velocità, frenata e arresto, ed addestrato un semplice modello di rete neurale (chiamato *multi-layer perceptron*) per effettuare previsioni temporali su orizzonti temporali fino a 6 secondi.

Tuttavia, gli studi esistenti prevedono generalmente l'intenzione dell'agente target utilizzando un insieme di poche classi. Talvolta, questa pratica è considerata uno svantaggio poiché un numero ridotto di classi può fornire solo una comprensione di alto livello del comportamento. Ad ogni modo, questo problema può essere risolto suddividendo le manovre complesse in tante sottoclassi che descrivano il comportamento in modo più preciso. Per esempio, in uno scenario di guida in autostrada, le classi di cambio di corsia potrebbero essere suddivise in cambio di corsia "brusco" e "normale". Un altro svantaggio è la specificità delle classi per un singolo ambiente di guida. Questo problema può essere risolto definendo un insieme che contiene le primitive di comportamento in un contesto di guida più ampio (ovvero, che considera molteplici situazioni di traffico). Tuttavia, per prevedere le intenzioni utilizzando una dettagliata classificazione, è necessario un dataset di *training* piuttosto ampio e diversificato che includa campioni sufficienti in ogni classe, ed un modello con notevoli risorse computazionali per imparare l'annotazione dei dati di input all'insieme delle intenzioni.

Diversamente, i problemi di regressione riguardano la previsione delle posizioni e velocità degli agenti target, quali veicoli [136], [137], [148], ciclisti [133], [140], [151] e pedoni [140], [142], in una definita finestra temporale nel futuro, attraverso una comprensione generale delle loro dinamiche di movimento ed interazione. In generale, conoscere le posizioni future permette di avere informazioni più precise sul comportamento dell'agente. Tuttavia, trattare un

output continuo, come i percorsi o le traiettorie, aggiunge complessità al problema di predizione rispetto all'output discreto dei modelli di classificazione dell'intenzione. Per esempio, riconosciuta una specifica situazione di traffico e osservata la storia di movimento per il veicolo target, esistono molteplici traiettorie che quel veicolo potrebbe compiere in base alle intenzioni del conducente. Per questa ragione, i predittori di traiettoria si dividono in approcci unimodali e multimodali [152]: i primi predicono solo la traiettoria più probabile, in maniera indipendente o condizionata dalla stima dell'intenzione di manovra, mentre nel secondo caso un modello predice una traiettoria per ogni classe di comportamento, appresa dinamicamente o predefinita, assieme alla sua probabilità. Tuttavia, anche questi approcci non sono esenti da errori e svantaggi: per esempio, i modelli di previsione delle traiettorie unimodali indipendenti dalle intenzioni possono convergere a soluzioni che non rappresentano validi comportamenti futuri [130]; i modelli, unimodali e multimodali, basati su un insieme di manovre predefinite non possono prevedere accuratamente le traiettorie del veicolo se l'intenzione è diversa da quelle memorizzate e, in genere, hanno bisogno di annotare manualmente le intenzioni di manovra nel dataset di *training* (un operazione che richiede tempo ed è soggetta ad errori). Al contrario, i modelli di predizione delle traiettorie multimodali, in grado di appendere dinamicamente le intenzioni in base allo scenario di guida, sono esenti dai problemi degli altri metodi ma, allo stesso tempo, sono inclini a convergere verso un solo tipo di manovra [130] oppure non riescono ad esplorare tutti le manovre esistenti.

2.1.3 Approccio di modellazione

Esistono diversi metodi per produrre previsioni sul movimento ma questi vengono in genere raggruppati nelle tre seguenti categorie [119], [124], in base alle ipotesi fatte sulle entità modellate (si può notare che la categorizzazione differisce anche nel livello di "cognizione" tipicamente coinvolto nel processo di previsione):

1. *Metodi basati sulla fisica* (nel seguito utilizzeremo anche il termine inglese *physics-based* in riferimento a tali metodi): producono i modelli più semplici e assumono che il movimento dell'agente dipende solo dalle leggi della fisica. Pertanto, questi approcci definiscono un modello dinamico esplicito, basato sulla legge del moto di Newton, e seguono uno schema detto *Sense-Predict*, siccome rilevano la situazione e predicono il movimento atteso sotto i vincoli cineto-dinamici.

2. *Metodi basati su schemi* (o *pattern-based*): sono più avanzati dei metodi basati sulla fisica in quanto presuppongono che il movimento di un agente sia il risultato di una serie di schemi tipici di comportamento (nel caso di veicoli, le manovre) che possono essere acquisiti a priori e riconosciuti dalle traiettorie parziali osservate dell'agente stesso. Pertanto, questi approcci sono basati su uno schema detto *Sense-Learn-Predict*, perché devono apprendere gli schemi di movimento dai dati delle traiettorie degli agenti osservati.

3. *Metodi consapevoli delle interazioni* (o *interaction-aware*): tengono conto delle interdipendenze tra le azioni degli agenti e dell'effetto delle regole del traffico su tali interazioni per ragionare sulle intenzioni di movimento degli

agenti stessi. Questi modelli seguono uno schema detto *Sense-Reason-Predict*.

In sintesi, le differenze tra le tre classi possono essere illustrate considerando la situazione in Figura 5. In questo evento di traffico, un modello basato sulla fisica assume una velocità e un orientamento costante per le auto; un modello basato sulle manovre assume che l'auto blu vada dritta e l'auto gialla, per esempio, giri a sinistra; un modello consapevole dell'interazione assume che l'auto blu vada dritta, che l'auto gialla giri a sinistra e che il movimento congiunto delle auto sia vincolato dalle regole del traffico.

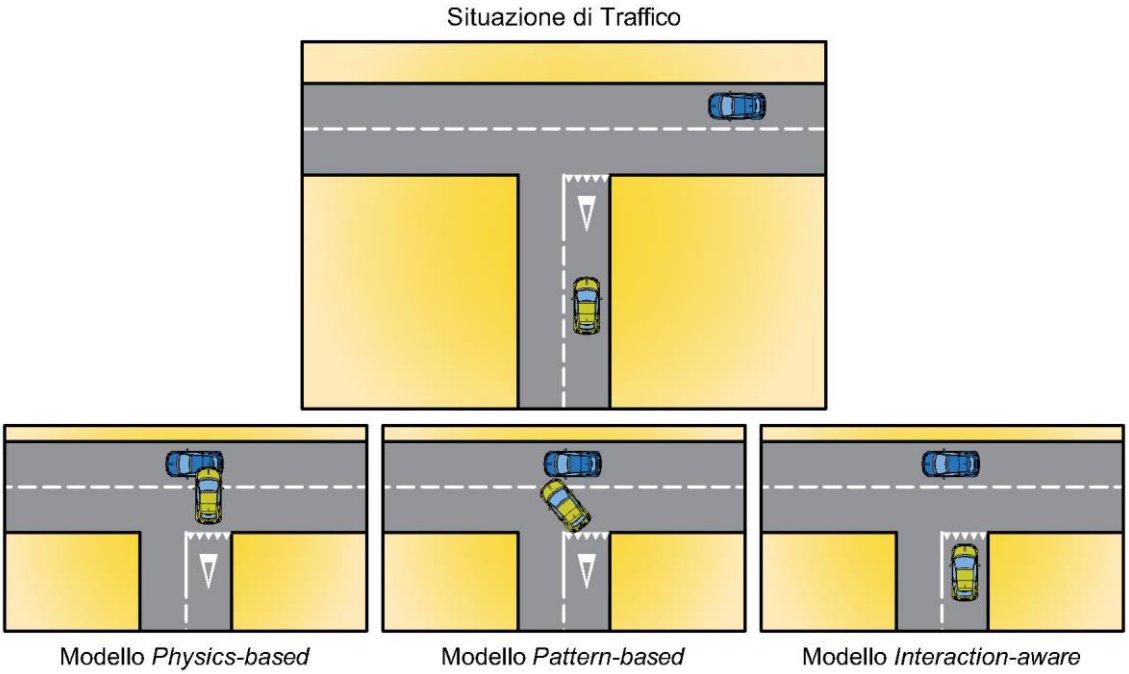


Figura 5: Esempi di previsione secondo i tre diversi modelli di movimento.

I modelli *physics-based* rappresentano gli agenti come entità dinamiche governate dalle leggi della fisica. Quindi, il movimento futuro è previsto dalla cinematica dell'agente nella situazione corrente, definita per esempio attraverso velocità, accelerazione, direzione e posizione attuale. Nel caso dei veicoli, si considerano in genere gli input di controllo (per esempio, la rotazione dello sterzo, la pressione sui pedali), le proprietà del mezzo (come il peso e l'aerodinamicità), le condizioni climatiche esterne (con le quali si stima il coefficiente di attrito della superficie stradale) e l'evoluzione dello stato del veicolo (per esempio posizione, direzione, velocità). Questi modelli sono più o meno complessi a seconda di quanto sia dettagliata la rappresentazione della dinamica e della cinematica dell'agente, di come vengano gestite le incertezze e se sia considerata o meno la geometria dell'infrastruttura. Con la scelta di una funzione di transizione appropriata, gli approcci basati sulla fisica possono essere facilmente applicati a più ambienti, senza la necessità di un data set di allenamento (alcuni dati per la stima dei parametri del modello sono comunque necessari). Lo svantaggio nell'uso di modelli di movimento espliciti e deterministici è che potrebbero non cogliere la complessità del mondo reale. Le funzioni di transizione mancano di

informazioni riguardanti il “contesto generale”, sia sulla scala spaziale che su quella temporale, portando a soluzioni che posso rappresentare minimi locali. In genere, siccome i modelli *physics-based* si basano solo sulle proprietà minime ed essenziali del movimento senza considerare regole di traffico, non sono in grado di anticipare i cambiamenti nello stato dell’agente causati, per esempio, dall’esecuzione di una particolare manovra (come rallentare ad una intersezione oppure girare a sinistra/destra a velocità costate) o da fattori esterni (come, rallentare a causa di un veicolo di fronte o dare la precedenza ad un pedone); di conseguenza, si può incorrere in previsioni errate nei contesi urbani, come si vede nel caso di Figura 5. In pratica, questo limita l’uso dei metodi basati sulla fisica a brevissimi orizzonti di previsione (meno di $1\,s$) e ad ambienti relativamente privi di ostacoli. Tuttavia, la possibilità di produrre un’inferenza approssimativa ma veloce, l’applicabilità a più scenari, e la facile interpretabilità hanno reso questi approcci un’opzione popolare per evitare le collisioni di piattaforme mobili (principalmente dei robot di servizio/supporto, ad esempio domestico, ma esistono anche applicazioni per veicoli a guida autonoma in ambienti circoscritti e predefiniti) e il tracciamento delle persone [124].

I modelli di movimento basati su schemi assumono che il movimento di un agente sulla strada corrisponda ad una serie di comportamenti o manovre eseguite indipendentemente dagli altri utenti. In questo contesto, i termini “manovra” e “comportamento” sono usati indifferentemente per indicare “un movimento fisico o una serie di movimenti che richiedono abilità e attenzione” [119]. Nella pratica, i modelli *pattern-based* approssimano una funzione dinamica/cinematica arbitraria a partire da dati di addestramento, ovvero apprendono schemi comportamentali da traiettorie di movimento osservate. Il comportamento può essere appreso adottando diversi approssimatori di funzioni, come le ANN, i modelli markoviani nascosti o i processi gaussiani. Ad ogni modo, la previsione della traiettoria si fonda sul riconoscimento anticipato del comportamento che l’agente intende eseguire: identificata l’intenzione di manovra di un conducente (o il comportamento di un pedone), si può assumere che il movimento futuro del veicolo (ovvero del pedone) corrisponderà con buona probabilità a quell’azione. Grazie a questa capacità di anticipare gli eventi, le traiettorie predette da uno schema *pattern-based* sono più rilevanti e affidabili a lungo termine di quelle prodotte da modelli *physics-based*. In particolare, esistono modelli *pattern-based* che sfruttano “traiettorie prototipo” e quelli che stimano intenzioni di manovra. Nel primo caso, l’idea alla base è che le traiettorie degli agenti possano essere raggruppate in un certo numero finito di *cluster*: ogni *cluster* corrisponde a uno schema caratteristico di movimento, ovvero una traiettoria “prototipo”, in quanto rappresentativo del comportamento medio osservato degli agenti in situazioni tipo. Infatti, i modelli di movimento, che sono concretamente definiti dalle traiettorie prototipo, sono appresi dai dati durante una fase di addestramento. Successivamente, la predizione può essere eseguita online, data una traiettoria parziale (ovvero una sequenza di stati passati dell’agente), trovando lo schema di comportamento più probabile e usando le traiettorie prototipo come modello per il movimento futuro. L’alternativa a questo approccio è quella di stimare a priori la possibile intenzione dell’agente (per esempio, nel caso di un veicolo target, arrestarsi ad uno stop, dare la precedenza ad un pedone, seguire un altro veicolo oppure eseguire una svolta a

sinistra/destra) e prevedere di conseguenza gli stati fisici futuri che corrispondono alla probabile esecuzione di quel comportamento. Quindi, non sussiste la necessità di trovare la corrispondenza tra una traiettoria parziale osservata dell'agente e una traiettoria prototipo precedentemente definita; diversamente, le caratteristiche rilevanti (dette anche di "alto livello", perché intrinseche del processo) che legano intenzioni e comportamenti sono acquisite dai dati di *training* e utilizzate per riconoscere i movimenti. In questo modo, è più facile generalizzare il modello appreso a contesti arbitrari. Tuttavia, nella pratica, l'ipotesi che gli agenti si muovano indipendentemente l'uno dall'altro non regge, in quanto la strada è un ambiente condiviso e il comportamento eseguito da un utente influenza necessariamente quello degli altri individui presenti. Le interdipendenze diventano particolarmente rilevanti nelle intersezioni stradali, dove la regola delle precedenze e priorità costringe i veicoli a tenere conto delle manovre effettuate dagli altri veicoli e delle intenzioni dei pedoni. Trascurare queste dipendenze può portare a interpretazioni errate delle situazioni e, di conseguenza, influenzare la valutazione del rischio.

Azioni come rallentare o cambiare corsia per evitare una collisione, dare la precedenza all'agente che ne ha diritto, così come rispettare le regole del traffico imposte dalla cartellonistica verticale ed orizzontale, sono esempi di interazione tra utenti, e tra conducenti ed ambiente stradale nell'ultimo caso proposto. I modelli di movimento che prendono in considerazione le interazioni tra gli agenti e l'ambiente sono chiamati *interaction-aware*, ovvero letteralmente "consapevoli dell'interazione". Infatti, questi modelli incorporano il concetto di "agente razionale", secondo cui l'agente pianifica e ragiona sugli obiettivi di movimento a lungo termine e stabilisce il comportamento che gli permette di raggiungere tali obiettivi. Di conseguenza, ogni sua azione e/o interazione con altri utenti ha un impatto sul futuro ed il modello in questione deve tenerne conto. Infatti, a differenza degli altri approcci presentati in questa Sezione, i metodi di *interaction-aware* sfruttano informazioni sul contesto, come posizione e velocità relativa tra gli agenti interagenti, per determinare cambiamenti nelle traiettorie che potrebbero manifestarsi a breve o a lungo termine. Per esempio, in un contesto autostradale, un modello *physics-based* potrebbe prevedere erroneamente che l'ego-veicolo si manterrà nella propria corsia quando non viene osservata alcuna variazione rilevante nella posizione laterale o nella direzione; diversamente, un modello *interaction-aware* potrebbe essere consapevole dell'intenzione dell'ego-veicolo di chiudere la distanza da un veicolo più lento che lo precede per un probabile cambio di corsia, non osservando alcun veicolo sopraggiungere dalla corsia adiacente. Quindi, prendere in considerazione le dipendenze tra gli agenti porta ad una migliore interpretazione del loro movimento e ad una valutazione più affidabile del rischio. Inoltre, i modelli *interaction-aware* permettono previsioni a più lungo termine rispetto ai modelli *physics-based* e si generalizzano a nuovi contesti meglio degli approcci *pattern-based* (rispetto ai quali risultano anche più affidabili) [124]. Tuttavia, gli approcci *interaction-aware* funzionano correttamente se gli obbiettivi che gli agenti intendono realizzare possono essere esplicitamente definiti dai dati di contesto, ovvero dalla rappresentazione dell'ambiente circostante tratta dall'ego-veicolo sulla base dei dati grezzi dei sensori di bordo. Inoltre, l'esaustività del metodo ha un grosso svantaggio: il tempo di esecuzione di questi modelli, i quali devono calcolare le potenziali traiettorie dell'agente target

sulla base di tutti gli agenti influenti, è computazionalmente costoso e varia esponenzialmente con il numero di agenti considerati, la complessità dell'ambiente e l'orizzonte di previsione [153], [154]. Per questa ragione, a volte (soprattutto quando non si utilizzano approcci basati sull'apprendimento) questi modelli non sono compatibili con la valutazione del rischio in tempo reale. Ad ogni modo, i modelli *interaction-aware* sono i più esaustivi proposti finora in letteratura ed il modello presentato in questa tesi può essere classificato all'interno di questa categoria.

Nelle Sezioni 2.1.1 e 2.1.2 sono state descritte le diverse modalità per definire ed incorporare le interazioni, assieme alle tipologie di output predetti dai modelli di movimento (in particolare di quelli i modelli *interaction-aware*). Tuttavia, vale la pena ricordare che un sistema di previsione non può essere sempre accurato e preciso, nel senso che talvolta la struttura dell'ambiente stradale può consentire molteplici manovre future e, allo stesso tempo, il comportamento dei conducenti umani aggiunge un ulteriore grado di incertezza al problema. Per questa ragione, in molte situazioni ha poco valore produrre una risposta deterministica, mentre sarebbe più utile incorporare le incertezze producendo una distribuzione su possibili traiettorie o intenzioni future. Sebbene si possano utilizzare diversi approssimatori di funzioni, i metodi di *Deep Learning* sono in grado di incorporare sia l'incertezza che le interazioni tra gli agenti e sono meno soggetti (ma non esenti) a problemi di scala dei tempi di esecuzione rispetto ad altri approcci. Nel seguito di questa Sezione vengono presentati gli approcci di modellazione più utilizzati e più rilevanti ai fini di questa tesi per lo sviluppo di modelli *interaction-aware*, prevalentemente finalizzati ad applicazioni di guida autonoma o assistita. Non saranno trattate le applicazioni riguardanti la previsione di comportamenti e/o intenzioni dei pedoni siccome aggiungono ulteriore complessità al problema; tuttavia, queste previsioni possono essere effettuate sfruttando gli stessi approcci delle applicazioni di guida autonoma e, spesso, la previsione per i diversi agenti viene gestita da più modelli separatamente per poi, in ultima analisi, raggruppare le predizioni ed effettuare le valutazioni di interesse [124].

In particolare, gli studi esistenti posso essere classificati in base ai metodi di modellazione in tre categorie, vale a dire reti neurali ricorrenti (RNN), reti neurali convoluzionali (CNN, dall'inglese *Convolutional Neural Network*) e altri metodi. In quest'ultimo caso sono particolarmente rilevanti le combinazioni di RNN e CNN.

La più semplice RNN, anche nota come *Vanilla* RNN, può essere considerata l'estensione di una rete neurale artificiale *feedforward* a due strati in cui il secondo strato nascosto è collegato ricorsivamente al primo (Sezione 4.4). Questa piccola modifica permette all rete di modellare dati sequenziali in modo efficiente. Ad ogni passo della sequenza, la *Vanilla* RNN elabora i dati di input correnti assieme all'informazione rilevante memorizzata nei passi temporali precedenti e trasmessa dai neuroni nascosti del precedente strato. In linea di principio, una *Vanilla* RNN con un numero sufficiente di unità nascoste può approssimare qualsiasi funzione che lega una sequenza di input ad una di output [124]. Tuttavia, è difficile addestrare nella pratica questo tipo di rete ad apprendere lunghe sequenze, a causa del fenomeno di *vanishing* del gradiente [155], ed è per questo che sono state recentemente implementate nuove architetture RNN in grado di ovviare a tale fenomeno, come la LSTM [98] e la GRU [156]. Nel

campo della previsione dei comportamenti degli agenti, le LSTM sono i modelli di apprendimento profondo più utilizzati [124], [152]. Di seguito, gli studi recenti sono categorizzati in base alla complessità dell'architettura di rete, in *Single* RNN e *Multiple* RNN.

Nella classe delle *Single* RNN sono inclusi i modelli in cui una RNN viene utilizzata secondo l'approccio *pattern-based* per la predizione dell'intenzione o della traiettoria unimodale del singolo agente, e i modelli RNN secondari che si accostano a un approccio principale per supportare funzioni più sofisticate, come l'interazione dell'agente con altri utenti oppure la previsione multimodale. Zyner e coll. [157] hanno sfruttato l'architettura LSTM per implementare una RNN a tre strati come classificatore di sequenze e prevedere le intenzioni dei conducenti nelle intersezioni urbane. A tal fine, una sequenza di caratteristiche sulla posizione, direzione e velocità dell'EV relative all'istante corrente e ai quattro passi temporali precedenti (corrispondenti ciascuno a $0.6\,s$) è fornita in successione alle unità ricorrenti LSTM. Applicando il classificatore *softmax* all'ultimo strato nascosto (Sezione 4.2), la dimensione dell'output della RNN è ricondotta al numero di classi, ovvero le possibili destinazioni offerte dall'intersezione (in questo caso Est, Nord e Sud). Il sistema ha prodotto risultati affidabili su una finestra di previsione di $1.3\,s$. Phillips e coll. [158] in modo molto simile sono riusciti a predire se un conducente girerà a sinistra, a destra o continuerà dritto in intersezioni a tre e a quattro braccia fino a $150\,m$ prima di raggiungere l'intersezione stessa, con una *Accuracy* (definizione al Capitolo 5) media sulle 9 partizioni di validazione (Sezione 4.3) superiore a 0.850. Altché e de La Fortelle [137] hanno utilizzato una rete LSTM con un solo strato nascosto come regressore per prevedere, su un orizzonte temporale fino a $10\,s$ nel futuro, la posizione laterale e la velocità longitudinale del veicolo target che circonda l'EV in un contesto autostradale. Tuttavia, i risultati in termini di errore quadratico medio tra target osservati e output predetti si possono considerare soddisfacenti solo per gli orizzonti più brevi (fino a 3-$4\,s$), mentre per gli orizzonti temporali più lunghi la predizione ha reagito con un ritardo (questo ritardo può arrivare anche a $8\,s$ per un orizzonte di previsione di $10\,s$), dimostrando che il modello non è stato in grado di interpretare i comportamenti dell'agente target nel lungo periodo. Talvolta, le reti LSTM con un solo strato nascosto hanno invece ottenuto risultati competitivi rispetto alla controparte multistrato, come per esempio nella predizione del comportamento umano [137]. Per prevedere una traiettoria basata sulle intenzioni del conducente, Ding e Shen [136] hanno usato una rete LSTM per anticipare la politica di guida dell'EV (come procedere dritto, girare a sinistra o a destra, dare la precedenza) attraverso una sequenza dei propri stati passati. Quindi, l'intenzione predetta e le informazioni acquisite sul contesto (inerenti alla geometria dell'ambiente stradale, alle regolamentazioni del traffico e agli agenti in movimento) vengono sfruttate per generare una previsione iniziale della traiettoria futura dell'EV. Infine, un metodo di ottimizzazione non lineare è applicato per affinare la previsione iniziale in modo che la traiettoria futura predetta sia basata sull'intenzione del conducente così come sulle regole di traffico, sulla geometria del contesto stradale e sulle interazioni con gli altri agenti. Gli autori hanno dimostrato che è essenziale incorporare l'anticipazione della politica di movimento a causa della natura multimodale della traiettoria futura. Nel loro sistema di previsione online a due livelli, l'output finale è una traiettoria

continua che si adatta automaticamente alle diverse situazioni di traffico e ha dimostrato di essere accurata anche in contesti urbani. In particolare, è rilevante osservare che le prestazioni del sistema proposto sono state analizzate e validate attraverso simulazioni di guida autonoma.

Per affrontare una previsione multimodale e/o il concetto di "agente razionale" all'interno di una RNN, gli studi esistenti fanno solitamente uso di diverse RNN che operano in parallelo. Ding e coll. [143] impiegano un gruppo di codificatori GRU per modellare a coppie l'interazione tra il veicolo target e i veicoli circostanti in un contesto autostradale e sfruttano questa consapevolezza per predire accuratamente (*Recall* $= 0.967$, per la definizione di questa metrica si rimanda al Capitolo 5) le intenzioni del veicolo target su orizzonti temporali fino a $5\,s$. Diversamente, siccome i modelli LSTM non possono descrivere simultaneamente le interazioni spaziali tra i diversi veicoli e le relazioni temporali tra le sequenze di una traiettoria, Dai e coll. [144] hanno implementato un modello LSTM spazio-temporale per la previsione della traiettoria attesa del veicolo target, composto dall'unione di due gruppi di reti LSTM che si occupano rispettivamente di modellare l'interazione tra gli agenti e di anticipare le sequenze temporali delle traiettorie di ciascun veicolo. Similmente, Xin e coll. [159] hanno presentato un algoritmo strutturato in due blocchi LSTM in serie, in cui alla prima unità viene fornita la sequenza passata di posizioni laterali del veicolo target per riconoscere la sua intenzione di cambio corsia mentre la seconda unità predice la posizione futura dalla politica di movimento predetta e dalla sequenza passata di posizioni longitudinali dell'agente. Su orizzonti temporali di $5\,s$, l'errore quadratico medio del modello per la previsione longitudinale e laterale è inferiore a $5.80\,m$ e $0.50\,m$ ($RMSE$ complessivo, senza separare le due componenti di movimento, è pari a $5.79\,m$), rispettivamente. Deo e Trivedi [160] hanno proposto un modello *interaction-aware* composto da diverse LSTM di codifica e decodifica per la previsione del movimento dei veicoli che circondano l'EV in un contesto autostradale. Questo modello assegna valori di confidenza alle intenzioni di manovra degli agenti e produce una distribuzione multimodale sul movimento futuro di ciascun agente sulla base di questi valori. Confrontando questo approccio con lo stato dell'arte, gli autori hanno ottenuto risultati migliori in termini di errore quadratico medio di previsione ($RMSE = 4.66\,m$) su orizzonti temporali fino a $5\,s$.

Sebbene le RNN siano una delle principali reti neurali impiegate per l'analisi delle dipendenze temporali e la previsione di sequenze, sono carenti nella modellazione delle relazioni spaziali (come le interazioni tra gli agenti) e dei dati immagine (talvolta rilevanti per l'acquisizione di informazioni sul contesto di guida). Questo spiega perché le soluzioni sofisticate che utilizzano RNN di solito sfruttano metodi aggiuntivi per compensare la debolezza della *Single* RNN.

Rispetto alle reti *feedforward* convenzionali, le CNN includono 1) strati di convoluzione, nei quali un filtro con pesi apprendibili è convoluto sullo strato di input, 2) strati di *pooling*, i quali riducono la dimensione spaziale del proprio input attraverso sottocampionamento, e 3) uno strato finale *fully-connected*, che mappa il proprio input alla dimensione dell'output desiderato. Questa struttura è particolarmente utile per individuare e riconoscere oggetti, volti e scene nelle immagini [161] e, per questa ragione, le CNN sono comunemente impiegate per estrarre informazioni contestuali da dati immagine. Quindi, nelle applicazioni di

guida autonoma o assistita, le CNN sono intese come modelli di *object detection*, ovvero per la classificazione dei diversi utenti della strada nei fotogrammi delle immagini acquisite da telecamere fisse o installate a bordo dei veicoli. Lo stato dell'arte in questo settore è costituito dal modello *Mask* R-CNN [162], che sta *Region-based Convolutional Neural Network*. Si tratta di un modello di rilevamento avanzato, adattato dal *Faster*-RCNN [163], in grado di annotare automaticamente diversi tipi di oggetti in un fotogramma e generare una sagoma (detta "maschera di segmentazione") per ogni istanza rilevata. L'algoritmo, come prima operazione, scansiona l'intera immagine e stima le aree che hanno più probabilità di contenere un'istanza; quindi, classifica i diversi oggetti in ciascuna area e parallelamente produce una maschera di segmentazione per ciascuno. Questo meccanismo ha dimostrato prestazioni migliori rispetto agli esistenti modelli di *object detection*, specialmente sui piccoli oggetti nella scena che sono generalmente difficili da rilevare [164]. Tuttavia, un modello di *object detection* ha bisogno di un modello di *object tracking* per tracciare e seguire il movimento di ogni utente rilevato nell'ambiente stradale. Un metodo di *tracking* comune e all'avanguardia è il DCF-CSR *Tracker* (*Discriminative Correlation Filter with Channel and Spatial Reliability*) [165], il quale in accoppiamento ad un modello di rilevamento permette di generare le traiettorie degli utenti dai video.

Ad ogni modo, esistono architetture CNN in grado di affrontare simultaneamente i compiti di rilevamento degli oggetti e di previsione del comportamento [152]. Questi modelli sono apprezzati per le loro capacità di lavorare sui fotogrammi dei video, generare output in formato immagine (chiamati "mappe di occupazione" perché rappresentano graficamente la probabilità di occupazione di ogni cella di una *Bird's Eye View* dell'ambiente di guida nei passi temporali futuri) e conservare le relazioni spaziali intrinseche nei dati di input mentre questi vengono processati. Grazie a queste peculiarità, le CNN hanno permesso di produrre mappe di occupazione che considerano sia l'interazione tra gli agenti sia il contesto di guida. Tuttavia, queste architetture mancano di un meccanismo per modellare le sequenze di dati, che è fondamentale nella previsione dei comportamenti per acquisire le dipendenze tra gli stati degli agenti nel tempo. Per questa ragione, si trovano sempre più spesso nella letteratura modelli che combinano RNN e CNN per trarre vantaggio dalle capacità di entrambe, ovvero per elaborare sia la dimensione spaziale che quella temporale caratterizzanti i dati sul comportamento di guida. Per esempio, Deo e Trivedi [166] hanno proposto un modello a doppio strato LSTM che utilizza tra questi una CNN per apprendere in modo robusto le interdipendenze nel movimento di tutti i veicoli che circondano l'EV. In pratica, un codificatore LSTM cattura la dinamica di movimento per ciascun veicolo circostante sulla base degli stati passati osservati. Quindi, uno strato convoluzionale, detto di *social pooling*, impara l'interdipendenza spaziale attraverso un *social tensor*, ovvero una griglia *Bird's Eye View* centrata sull'EV e popolata dagli stati LSTM di tutti gli agenti circostanti secondo la configurazione spaziale della scena. Infine, un decodificatore LSTM basata su classi di manovra produce una distribuzione multimodale sulle traiettorie future dell'EV. Questa configurazione ha mostrato risultati migliorativi rispetto allo stato dell'arte in termini di errore quadratico medio di previsione ($RMSE = 4.37\ m$ sull'orizzonte temporale massimo di $5s$), perché il modello ha ora accesso agli stati di movimento degli agenti circostanti e alla loro

configurazione spaziale. Nel contesto di modellazione attraverso combinazione di CNN e RNN, l'algoritmo implementato da Zhang e coll. [85], finalizzato allo sviluppo di sistemi FWC nell'ambito dei veicoli connessi, ha rappresentato uno spunto interessante per il sistema presentato in questa tesi. Infatti, gli autori hanno utilizzato il *Mask* R-CNN e il DCF-CSR *Tracker* per ricostruire caratteristiche di mobilità sequenziale di pedoni e veicoli in intersezioni regolate da semaforo e, successivamente, hanno addestrato una GRU a prevedere gli eventi di quasi incidente dallo stato attuale della scena di traffico. A differenza dei modelli presentati in questa Sezione, gli autori non hanno effettuato predizioni in avanti (perché si sono limitati ad analizzare la situazione corrente) ma hanno saputo sfruttare il potere di entrambe le architetture neurali CNN e GRU per annotare l'interazione tra gli utenti e produrre una valutazione del rischio in tempo reale con risultati soddisfacenti (*Recall* $= 0.856$ sulle 10 partizioni di validazione).

2.2 Valutazione del Rischio

Nella Sezione precedente sono stati trattati gli approcci di modellazione più comunemente utilizzati nella letteratura scientifica per apprendere e prevedere il movimento futuro dei veicoli e dei pedoni. Tuttavia, la questione su come queste previsioni possano essere sfruttate per valutare l'effettivo rischio (o la severità) di una situazione è rimasto irrisolto.

All'inizio del Capitolo 2 è stato dimostrato come il RAM sia una garanzia fondamentale per la sicurezza funzionale dei sistemi FCW e AEB. Infatti, questo modello non solo è responsabile della stima dello stato di sicurezza della guida e della previsione del rischio di collisione, ma determina la tempista di attivazione dei sistemi di assistenza. Di conseguenza, il modello di valutazione del rischio, come elemento fondante il livello di decisione dei sistemi FCW e AEB, influisce direttamente sul risultato finale di prevenzione delle potenziali collisioni e, quindi, dovrebbe soddisfare i seguenti requisiti [167], almeno per quanto riguarda gli eventi di incontro tra veicolo e pedone:

1. Il RAM dovrebbe potersi adattare alle diverse e complesse situazioni di incontro. Le condizioni per lo sviluppo e la verifica del modello si dovrebbero basare sulla simulazione di diversi intervalli di velocità del conducente nel contesto del traffico urbano; inoltre, la velocità trasversale del pedone, così come la distanza laterale relativa e il luogo della potenziale collisione, dovrebbero essere definiti di conseguenza (ovvero sulla base delle medesime condizioni di traffico).

2. Le informazioni relative a veicolo e pedone necessarie alla valutazione del rischio dovrebbero essere facilmente accessibili. Il RAM dovrebbe poter acquisire (sulla base del sistema di rilevamento stabilito) e analizzare prontamente le informazioni operative (principalmente la posizione relativa e l'azimut del pedone rispetto al veicolo) per giudicare lo stato di sicurezza dell'interazione.

3. Il RAM dovrebbe non solo determinare accuratamente (in termini quantitativi) lo stato di sicurezza dell'evento ma classificarlo per livelli di rischio (ovvero in termini qualitativi). Date le diverse situazioni di incontro, il RAM dovrebbe essere in grado di classificare il rischio in modo ragionevole, affinché il

sistema di assistenza che lo implementa possa emettere, solo nelle condizioni effettivamente necessarie, il segnale di controllo più appropriato per prevenire l'incidente.

4. Il RAM non dovrebbe interferire con il normale comportamento di guida del conducente, ovvero il conducente non dovrebbe essere sottoposto a interferenze eccessive durante il normale processo di guida. Se il segnale di controllo emesso o la tempistica di attivazione del sistema risultassero irragionevoli, questi potrebbero causare disagio ai passeggeri del veicolo o disorientare il conducente sul proprio modo di guidare o di giudicare le situazioni.

Pertanto, un RAM che possa catturare rapidamente, oggettivamente ed accuratamente il livello di rischio nel processo di incontro tra veicolo e pedone (e, quindi, in grado di comprendere nella sua completezza le relazioni tra comportamento e rischio) è essenziale per giudicare adeguatamente la situazione ed emettere, nelle tempistiche opportune, il giusto segnale di assistenza o controllo della guida per prevenire potenziali collisioni. Per arrivare a questo risultato è necessario per prima cosa chiarire alcuni concetti fondamentali.

Un incontro tra due utenti della strada (ovvero il loro arrivo simultaneo in una certa porzione limitata della superficie stradale) può essere visto come un evento elementare nel processo di traffico che ha un potenziale di evolvere in un incidente (attraverso una collisione tra i due utenti). Hydén [84] ha suggerito l'esistenza di una dimensione di "severità", comune a tutti gli eventi di traffico, e ha proposto un modello piramidale per descrivere la relazione tra la gravità e la frequenza degli eventi (Figura 3). Secondo questo modello, maggiore è la severità (presentata come la posizione verticale nella piramide), minore è la frequenza (il volume della porzione di piramide fino a quella altezza) degli eventi. Tuttavia, il concetto di severità richiede qualche chiarimento. La gravità di un incidente è determinata dalle conseguenze della collisione tra gli utenti (per esempio, il numero di morti e feriti oppure la perdita totale in unità monetarie). Questa definizione è alquanto problematica per gli incontri che non si concludono con una collisione, poiché, a rigore, un "quasi incidente" con pochi centimetri di margine tra i veicoli coinvolti e un passaggio con sufficiente margine di sicurezza per la stessa area dei medesimi veicoli avrebbero le stesse conseguenze (se non per le differenze nei livelli di adrenalina nel sangue dei conducenti). Quindi, due aspetti devono essere considerati nella definizione di severità: il potenziale di un incontro di evolvere in un incidente (che si può confondere con il concetto di rischio) e la gravità delle conseguenze derivanti dal suo accadimento. Il potenziale di collisione può essere descritto nel seguente modo: gli incidenti sono eventi stocastici e, sebbene un particolare incidente possa essere spiegato sulla base di un certo numero di fattori che lo hanno causato, la collisone può anche essere considerata come la coincidenza sfortunata dei medesimi fattori nel luogo e tempo specifici dell'incontro. Se alcuni di questi fattori contribuenti non fossero stati presenti, l'incidente si sarebbe potuto evitare. In altre parole, ogni incontro può trasformarsi in un incidente qualora si presentino nuovi fattori sfavorevoli. In ogni caso, un quasi incidente ha un minor margine per tollerare un ulteriore fattore "sfortunato" rispetto a un passaggio in sicurezza e, di conseguenza, è

caratterizzato da un potenziale di incidente più elevato. Diversamente, la gravità di un incidente è generalmente influenzata da fattori come il tipo di utenti coinvolti, l'angolo di collisione tra questi, la velocità di impatto, etc. (Capitolo 1).

Trascurando, per il momento, la questione se si debba o meno tenere separate la probabilità di collisione e la gravità delle conseguenze, si assuma che la severità integri entrambi i concetti. Quando un livello di severità è assegnato agli incontri di traffico, questi possono essere collocati all'interno di una certa distribuzione, al pari della piramide di Hydén (queste distribuzioni sono chiamate gerarchie di severità da Svensson [168]). Ovviamente, è ragionevole supporre che esista una gerarchia "vera" che riflette la severità oggettiva ma il modo in cui viene definita la severità per i diversi incontri determina la forma effettiva della gerarchia. In particolare, l'introduzione di diverse misure operative per descrivere la severità ha creato gerarchie abbastanza diverse in cui uno stesso evento non è collocato esattamente allo stesso livello [83], [85], [89], [169]. Inoltre, la porzione nella quale si collocano la maggior parte degli eventi, in una gerarchia stabilita per un sito specifico (per esempio, una intersezione), riflette il comportamento predominante degli utenti della strada [76]; quest'ultimo può essere interpretato come il risultato dell'ottimizzazione, compiuta dagli utenti, tra il desiderio di raggiungere elevati standard di mobilità, e quello di mantenere certi margini di sicurezza e comfort (i quali dipendono anche da considerazioni diverse dalla sicurezza [72], come la volontà di mantenere una certa velocità). Ad ogni modo, queste gerarchie garantiscono una migliore comprensione delle situazioni del traffico da un punto di vista della sicurezza rispetto alle statistiche sugli incidenti, che rappresentano solo la parte superiore della distribuzione.

Dunque, le questioni da affrontare (e che ancora oggi rappresentano un punto in sospeso) sono relative a 1) come definire nel miglior modo operativo la severità di un incontro e 2) come classificare gli incontri in base al livello di severità.

2.2.1 Definizioni operative di severità

In merito alla prima questione, in genere si assume per la severità la sola dimensione relativa al rischio di collisione perché molti dei fattori che hanno implicazioni dirette sulla gravità degli incidenti (come l'uso del casco, l'età degli utenti coinvolti, il contatto visivo tra il conducente dell'EV e l'agente target) sono difficili da rilevare o da estrarre, per esempio, dall'analisi automatica dei video [76]. Talvolta, come nella soluzione proposta da Hydén [84] (nota come *Swedish Conflict Technique*), questo problema è aggirato definendo la severità come il potenziale di un incidente con feriti o morti. Di conseguenza, è naturale considerare la collisione come principale fonte di rischio e basare le valutazioni della severità (intesa quindi come rischio di collisione) unicamente sulla previsione delle collisioni. Questo approccio, che è stato adottato anche in molti lavori sui sistemi di trasporto intelligenti [119], è noto come "*Risk based on colliding future trajectories*" e si compone tipicamente di due fasi: 1) la previsione delle potenziali traiettorie future per tutte le entità in movimento nella scena di traffico e 2) la stima del rischio basata sulla probabilità complessiva di collisione tra ogni coppia di traiettorie predette (poiché in genere le situazioni di traffico sono trattate come interazioni multiple tra coppie di agenti, al fine di semplificare la complessità del problema reale). Le tecniche per prevedere le traiettorie sono

state presentate nella Sezione 2.1.3. Date le traiettorie future dell'EV e dell'agente target, esistono diversi metodi per estrarre informazioni riguardanti il verificarsi di una collisione (alcuni dei quali permettono di stabilire una probabilità per tipo di collisione) ma, al fine di stabilire una gerarchia della sicurezza, le *Traffic Conflict Techniques* basate sulle *Surrogate Safety Measures* rappresentano la soluzione comunemente adottata [170]. In particolare, esistono diversi indicatori SSM che possono integrare il rischio di collisione: questi indicatori esprimono la severità di un incontro come una misura di "prossimità" nello spazio o nel tempo ad un potenziale incidente [171]. Tuttavia, per applicare gli indicatori SSM alla valutazione della sicurezza è necessario rispondere ad alcuni quesiti. In primo luogo, è lecito chiedersi "quali indicatori si devono usare?" oppure "è necessario utilizzare un solo indicatore o una combinazione di più SSM?". Trovare un indicatore universale che sia applicabile a qualsiasi tipo di situazione e che rifletta tutti gli aspetti rilevanti nel processo di incontro rappresenta ancora oggi una sfida per i ricercatori. In genere, è necessario un insieme di SSM, il quale può variare a seconda dell'ambiente di traffico e del tipo di utenti coinvolti nell'evento di interesse. Infatti, l'utilizzo di un solo parametro (come nella maggior parte degli studi esistenti) non permette una valutazione accurata e completa della situazione reale, soprattutto quando si impiegano indicatori di prossimità nello spazio: per esempio, la breve distanza tra due utenti della strada su rotte collidenti non è una misura adeguata del livello di sicurezza nell'interazione, a meno che questa grandezza non sia accompagnata da informazioni sulle loro velocità. Pertanto, da un punto di vista metodologico, le misure di prossimità nel tempo sono più informative della controparte nello spazio, poiché sono il risultato di una combinazione di velocità e distanze degli utenti in conflitto [76], e per questo motivo risultano più idonee per l'analisi dei conflitti di traffico. Tra queste, le più utilizzate nel contesto scientifico [83]–[85], [169] sono il *Time-to-Collision* (TTC), il *Post Encroachment time* (PET), il *Time-to-Accident* (TA) e il *Time Gap* (TG). I dettagli di questi SSM sono riportati in Tabella 1, mentre la procedura di calcolo del TTC è esposta in dettaglio nell'Appendice A.

Tuttavia, la necessità di definire un insieme di indicatori temporali e/o misure complementari continua a sussiste per una ulteriore questione: "quali interazioni devono essere studiate?". Per esempio, un ricercatore potrebbe essere interessato alla valutazione delle sole interazioni gravi, all'apice della gerarchia, per lo sviluppo di un sistema di assistenza reattivo in sole situazioni ad alto rischio; oppure potrebbe ritenere più rilevanti le interazioni che hanno comportato un'azione evasiva da parte degli utenti, per duplicare tale comportamento in un veicolo autonomo. Nel caso generale, in cui sono di interesse tutti gli eventi potenzialmente pericolosi, esistono due scuole di pensiero. La prima è basata sulla tecnica olandese, nota come DOCTOR (*Dutch Objective Conflict Technique for Operation and Research*) e sviluppata dall'*Institute of Road Safety Research* nei Paesi Bassi [172], secondo cui la severità del conflitto di traffico dipende dal valore più basso del TTC raggiunto durante l'incontro e dal PET; diversamente, il valore del TTC al momento dell'inizio di un'azione evasiva (corrispondente al *Time-to-Accident*), assieme alla velocità di guida nel medesimo istante, sono considerati nella tecnica svedese [77]. La scelta di utilizzare solo due parametri è legata principalmente a motivazioni pratiche perché, al fine di rendere il metodo

operativo, è importante mantenere il numero di indicatori il più basso possibile (a rischio di perdere alcune informazioni). In merito alla scelta degli indicatori negli approcci olandese e svedese, è necessario esporre alcune osservazioni.

Tabella 1: Definizione di alcuni indicatori SSM.

Indicatore	Definizione	Note
Time-to-Collision (TTC)	Il tempo impiegato da due utenti della strada per collidere qualora mantenessero l'attuale traiettoria e velocità relativa.	Insieme di valori calcolati in modo continuo nel tempo. I due utenti in conflitto devono essere su rotte collidenti (ovvero rotte di collisione). Molto usato nella pratica.
Post Encroachment time (PET)	L'intervallo di tempo tra il momento in cui il primo utente lascia la zona di conflitto e il momento di arrivo del secondo nel medesimo luogo.	Valore discreto. Il primo utente ha già attraversato la zona di conflitto. Particolarmente utile nel caso di traiettorie trasversali (collisioni prossime all'angolo retto).
Time-to-Accident (TA)	Il tempo rimanente al verificarsi di una collisione dal momento in cui uno dei due utenti coinvolti inizia un'azione evasiva.	Valore discreto. È legato all'azione evasiva compiuta dagli utenti della strada.
Time Gap (TG)	Il tempo che verrebbe misurato tra il momento in cui la parte posteriore del primo veicolo passa il punto di conflitto e il momento in cui la parte anteriore del secondo veicolo arriva al medesimo punto se entrambi mantenessero la stessa velocità e traiettoria.	Insieme di valori calcolati in modo continuo nel tempo. TG riflette la probabilità che il TTC raggiunga rapidamente valori bassi se gli utenti della strada entrassero in rotta di collisione. Il concetto si può estendere anche agli incontri tra veicolo e pedone/ciclista. La differenza con il PET è rappresentata dal punto di conflitto in questione.
T_2 *Indicator* (T_2)	Il tempo atteso affinché il secondo utente arrivi al punto di conflitto, mantenendo l'attuale velocità e la traiettoria "pianificata".	Insieme di valori calcolati in modo continuo nel tempo. Gli utenti non devono essere necessariamente su una rotta di collisione per il calcolo del parametro. In questo caso, il T_2 eguaglia il TTC.
Time Advantage (T_{ADV})	In qualsiasi istante durante un incontro, rappresenta il valore PET atteso qualora gli utenti della strada mantenessero l'attuale traiettoria e velocità relativa.	Insieme di valori calcolati in modo continuo nel tempo. Rappresenta l'estensione temporale del PET.

In primo luogo, la rotta di collisione tra i due utenti interagenti alla fine di un evento è una condizione necessaria affinché si verifichi una collisione. Tuttavia, anche gli incontri tra utenti che non hanno rotte collidenti possono avere un potenziale di collisione, in quanto possono manifestarsi cambiamenti nelle relazioni spaziali o temporali tra questi da condurre a una rotta di collisione. Queste relazioni in un determinato momento dell'interazione possono essere

classificate in tre tipi di percorso, chiamati da Laureshyn e coll. [76], "*collision course*", "*crossing course*" e "*non-crossing course*". Nel primo tipo di relazione, gli utenti della strada sono in rotta di collisione e si scontreranno se non verrà presa alcuna azione evasiva. La situazione di "*crossing course*" rappresenta la condizione in cui i percorsi pianificati dagli utenti si sovrappongono, ma la collisone sarà evitata poiché questi passano per la zona spaziale comune alle traiettorie in tempi diversi. Affinché la collisione diventi possibile, è necessaria una correzione dei parametri di guida nel tempo, ovvero uno o entrambi gli utenti devono cambiare la propria velocità. In altre parole, la situazione deve prima trasformarsi in una relazione del tipo "*collision course*". Diversamente, quando i percorsi degli utenti non si sovrappongono in alcun modo si parla di "*non-crossing course*". Questa situazione si verifica quando gli utenti si muovono in parallelo in due corsie adiacenti o hanno percorsi divergenti. Tuttavia, questo non significa che il rischio di collisione sia nullo, poiché correzioni del percorso o aggiustamenti dei parametri compiuti da uno o entrambi gli utenti posso portare alla sovrapposizione delle traiettorie e, in certe condizioni, a rotte collidenti. Per esempio, un pedone che cammina sul marciapiede a lato strada è a pochi passi dalle auto che viaggiano sulla corsia adiacente; quindi, esiste il rischio che la situazione si sviluppi in un conflitto molto severo se il pedone cambiasse improvvisamente la direzione di camminata. Tuttavia, anche in una tale situazione, prima che avvenga una collisone tra veicolo e pedone, la relazione deve trasformarsi da "*non-crossing course*" a "*collision course*", possibilmente attraverso una condizione di "*crossing course*". Ovviamente, l'evoluzione di un evento può avvenire anche nel verso opposto, ovvero da "*collision course*" a "*non-crossing course*", soprattutto a seguito di un'azione evasiva. In partica, anche se queste tre relazioni sono teoricamente diverse, creano un continuum in cui il passaggio da un tipo all'altro avviene uniformemente, perché il comportamento degli utenti non cambia bruscamente in queste transizioni; anzi, il cambio di relazione è il risultato di piccole (e reversibili) variazioni di velocità e/o percorso [76]. In questi termini, l'uso del solo TTC non è sufficiente per rilevare tutte le situazioni potenzialmente pericolose, come supportato dalle osservazioni dei conflitti reali nel traffico [173]. Per esempio, Svensson [168] ha notato che nelle situazioni in cui due conducenti si evitano l'un l'altro con un margine di tempo molto breve (e quindi, in senso stretto, la loro relazione ha carattere di tipo "*crossing course*"), il loro comportamento evasivo è analogo a quello che avrebbero assunto se fossero stati in una condizione di "*collision course*". La spiegazione più ovvia è che, con piccoli margini di sicurezza, anche un piccolo cambiamento di velocità (o un piccolo disturbo nel processo) potrebbe facilmente portare a una collisione effettiva e, di conseguenza, la situazione viene percepita e vissuta come se fosse ad alto rischio [174]. La tecnica olandese DOCTOR include nella sua definizione di conflitto sia le situazioni del tipo "*collision course*" che quelle "*crossing course*" con margine di sicurezza molto piccolo: infatti, il TTC coglie le situazioni in cui è presente una rotta di collisione (perché, solo sotto questa condizione, l'indicatore può essere calcolato), mentre il valore del PET permette di includere anche le "quasi collisioni". In questa tecnica, si assume che valori di *TTC* minori od uguali di $1.5\,s$ e valori di PET minori od uguali ad $1.0\,s$ costituiscono una situazione potenzialmente pericolosa in un contesto urbano. Diversamente, nella tecnica svedese, i conflitti di interesse sono solo quelli più

severi, ovvero in grado di generare feriti e/o morti qualora non venga presa alcuna azione evasiva da uno o entrambi gli utenti interagenti. Tuttavia, entrambi i metodi si basano solo su un determinato momento durante il processo di incontro, che può essere idoneo per le analisi in post elaborazione della sicurezza di un certo ambiente di traffico ma non per le valutazioni in tempo reale su un veicolo in movimento.

Per quanto visto, l'incontro tra due utenti della strada è un processo continuo e le misure di severità dovrebbero essere abbastanza flessibili da includere tutti gli stadi dell'interazione, permettendo una transizione continua anche nella descrizione del livello di rischio dell'evento. Gli indicatori *Time Advantage* (T_{ADV}) e T_2, suggeriti per la prima volta in combinazione da Laureshyn e coll. [76], rappresentano un tentativo di colmare questa lacuna. La descrizione dettagliata degli indicatori T_{ADV} e T_2 è fornita in Tabella 1, mentre l'Appendice A illustra la procedura per il loro calcolo considerando la dimensione degli utenti. T_{ADV} è un indicatore usato per descrivere le situazioni in cui due utenti della strada passano una zona spaziale comune, ma in tempi diversi [175]. T_{ADV} amplia il concetto di PET, calcolando per ogni istante durante l'interazione il valore atteso del PET se la coppia di utenti mantenesse l'attuale velocità e traiettoria; quindi, T_{ADV} può essere valutato in continuo durante l'intero processo di interazione, a differenza del PET che assume un solo valore durante l'incontro. La specificità del T_{ADV} è che, mentre i suoi bassi valori possono riflettere gli aspetti di sicurezza ($T_{ADV} = 0$ indica nessun margine, ovvero "*collision course*"), i valori più alti (sopra 2-3 s) descrivono le normali situazioni di traffico e possono essere interpretati come una misura del vantaggio di un utente sull'altro in una competizione per il passaggio sulla medesima porzione limitata dell'infrastruttura [175], [176]. Infatti, l'utente che ha il maggiore vantaggio temporale è molto probabile che sia quello che passa per primo sulla zona comune. Tuttavia, se il vantaggio temporale è piccolo, il secondo utente (ovvero l'ultimo a passare per la zona comune) può accelerare con l'obiettivo di passare per primo; questa situazione si verifica principalmente quando uno degli utenti della strada ha più "potere" sull'altro, per esempio, nel caso di conflitto tra un veicolo ed un pedone [177]. Ad ogni modo, l'aspetto rilevante è la capacità dell'indicatore T_{ADV} di descrivere sia la sicurezza che l'efficienza dei processi di traffico; questa peculiarità potrebbe permettere una migliore comprensione di come questi due aspetti siano bilanciati dagli utenti della strada [77]. Tuttavia, il vantaggio temporale non è di per sé sufficiente per descrivere il rischio di collisione, poiché è anche importante sapere quando avverrà l'effettiva invasione della zona comune di passaggio: anche se il T_{ADV} è piccolo in un certo momento, gli utenti della strada potrebbero ancora avere abbastanza tempo per regolare le proprie velocità e traiettorie. Per questa ragione, Laureshyn e coll. [76] hanno proposto l'indicatore T_2 per includere, nella descrizione complessiva dell'evento, il tempo di arrivo del secondo utente al potenziale punto di collisione (Tabella 1). Infatti, dal punto di vista della sicurezza, l'ultima condizione necessaria al verificarsi di una collisione è proprio l'arrivo del secondo utente al potenziale punto di conflitto, qualunque siano le azioni del primo utente. Nel caso in cui i due utenti della strada passano il punto di conflitto con un margine di tempo, T_2 riflette il tempo massimo disponibile per intraprendere azioni evasive (poiché è colui che arriva per ultimo che ha il margine temporale più ampio) ed alleviare la severità della situazione. Inoltre,

l'indicatore T_2 estende il concetto di TTC, poiché il suo calcolo non richiede una rotta di collisione: diversamente, questo indicatore permette una transizione continua tra le situazioni di "*non-crossing course*" e "*collision course*" (e viceversa, passando sempre attraverso la condizione di "*crossing course*") all'interno della stessa interazione (senza bisogno di cambiare indicatore, a differenza della dicotomia tradizionale TTC contro PET). Se gli utenti della strada entrano in rotta di collisione, T_2 è ovviamente uguale al TTC. In particolare, Borsos e coll. [178] hanno recentemente dimostrato che i concetti di TTC e T_2 sono trasferibili per la stima della probabilità di collisone, con valori di soglia più severi per T_2 nell'annotazione dei livelli di severità. Inoltre, come il TTC, il T_2 è un indicatore continuo e può essere calcolato per qualsiasi momento, purché entrambi gli utenti della strada siano diretti verso la comune zona di conflitto. Infatti, il T_2 non può più essere calcolato dopo che il primo utente della strada ha lasciato la zona di conflitto (poiché la collisione non è più possibile). In particolare, sotto questa condizione, l'ultimo valore possibile di T_2 durante un'interazione ha praticamente lo stesso significato del PET e riflette il momento in cui i due utenti della strada sono più vicini nello spazio l'uno all'altro. Diversamente, il valore più piccolo del T_2 registrato durante l'evento di incontro riflette il momento in cui gli utenti sono più vicini nel tempo. Nella maggior parte dei casi, questi due valori coincidono (poiché T_2 normalmente diminuisce man mano che gli utenti della strada si avvicinano l'uno all'altro), ma nel caso di cambiamenti significativi di velocità durante l'interazione, ad esempio a causa di una brusca frenata, potrebbero rappresentare istanze temporali diverse [169].

Per quanto osservato, la soluzione proposta da Laureshyn e coll. [76], ovvero l'impiego di T_{ADV} e T_2 per la definizione del rischio di collisione, rappresenta attualmente la modalità operativa più esaustiva per cogliere la severità degli incontri di traffico. La capacità di tale approccio di descrivere l'intero processo di incontro come un'interazione continua tra gli utenti garantisce accuratezza al processo di valutazione, ma per essere concretamente implementato in una RAM è necessario stabilire i valori di soglia degli indicatori che definiscono le situazioni potenzialmente pericolose, ovvero permettono di classificare le situazioni in differenti livelli di rischio.

2.2.2 Annotazioni delle classi di severità

Per annotare gli eventi di quasi incidente (ovvero classificare gli incontri in base al livello di rischio), i valori critici o di soglia degli indicatori SSM devono essere stabiliti. Non esiste una definizione precisa e univoca di tali valori, poiché essi dipendono dal comportamento degli utenti nei contesti di guida specifici e dalle finalità del processo di classificazione. Mahmud e coll. [179] hanno fornito una panoramica dei valori di soglia per il minimo valore del TTC durante un incontro, da una selezione di studi per condizioni diversificate: per esempio, nel caso di intersezioni urbane regolate da semaforo, gli autori hanno citato due riferimenti, uno per le interazioni tra veicoli [180] e l'altro per quelle tra conducenti e ciclisti [181], nelle quali sono stati definiti auspicabili per la valutazione della sicurezza i valori di soglia pari a 1.6 e $3\ s$, rispettivamente. Nel primo caso, ovvero nell'interazione tra veicoli, Huang e coll. [180] hanno impiegato anche il valore minimo del TG durante le manovre di incontro, definendone un valore critico di $3\ s$. È stato precedentemente accennato che nella tecnica DOCTOR le situazioni

potenzialmente pericolose nelle aree urbane sono identificate da $TTC \leq 1.5\ s$ e $PET \leq 1.0\ s$. Laureshyn e coll. [169] hanno osservato che le situazioni più severe nelle interazioni di traffico tra veicoli in svolta in una intersezione regolata da semaforo sono caratterizzate da un valore minimo di T_2 di $1.5\ s$ o inferiore. Ad ogni modo, esistono diversi approcci che possono essere utilizzati per classificare gli incontri in base alla severità.

La maggior parte degli studi esistenti utilizza valori di soglia definiti arbitrariamente sulla base di considerazioni relative ai tempi di percezione e reazione degli utenti [179]. Nel caso di applicazioni a sistemi FCW e AEB, la definizione dei livelli potrebbe tenere conto anche del processo di frenata [182], il quale viene accuratamente analizzato per determinare o un tempo di avviso ragionevole (nel caso di sistemi FCW) o il tempo effettivo di azione del freno (oltre al tempo di reazione del conducente, esiste sempre un tempo di ritardo idraulico nella risposta dei freni). In questo contesto, Yang e call. [167] hanno proposto un algoritmo gerarchico per classificare i livelli di sicurezza nelle interazioni veicolo-pedone in differenti stati cineto-dinamici di guida. In pratica, tenendo conto della velocità del conducente, delle distanze percorse durante le fasi di percezione del pericolo, reazione ed applicazione dei freni fino all'arresto del veicolo, assieme a una distanza di sicurezza ridondante tra veicolo e pedone al momento del completo arresto (fissata a $2\ m$), gli autori hanno stabilito come i valori di soglia del TTC debbano variare nell'intervallo tra 20 e $60\ km/h$, tipici del contesto urbano. In particolare, il loro sistema RAM definisce tre livelli di sicurezza corrispondenti a tre diversi valori di soglia del TTC: il primo livello rappresenta le condizioni di guida in sicurezza (la RAM non emette alcun segnale di controllo), il secondo livello indica la presenza di un potenziale rischio di collisione (la RAM emette un segnale di avviso al conducente), e il terzo livello denota un imminente collisione (la RAM emette un segnale di controllo dei freni). Ovviamente, maggiore è la velocità iniziale, maggiori sono i valori di soglia corrispondenti ai tre livelli. Nel caso di velocità iniziale di $50\ km/h$, il primo livello si ha per $TTC > 3.0\ s$, il secondo livello per $1.5\ s < TTC \leq 3.0\ s$, mentre il terzo livello per $TTC \leq 1.5\ s$.

L'alternativa a questo approccio è quello di analizzare la forma dei profili di indicatori SSM continui durante le interazioni, come per esempio le curve del TTC e/o T_{ADV}, ecc. Un'analisi dettagliata di queste curve può rivelare le forme "tipiche" che caratterizzano le situazioni critiche. Allo stesso modo, si possono trovare forme che riflettono processi "normali" (non critici). Tipicamente, questo tipo di analisi sfrutta metodi di *pattern recognition* come discusso da Laureshyn e coll. [183]. Per esempio, Kathuria e Vedagiri [83] hanno analizzato la dinamica degli incontri tra veicolo e pedone sulle strisce pedonali in intersezioni non regolate da semaforo e suddiviso le interazioni in due modelli di comportamento, definiti "*responsive*" e "*un-responsive*" in base alla presenza di azioni evasive di uno o entrambi gli utenti, in base ai profili della velocità, del TTC e del TG di entrambi gli utenti interagenti. A seguito di un test di importanza delle variabili, gli autori hanno applicato una *Import Vector Machine* per classificare i livelli di severità di 1486 eventi sulla base del solo TTC per il modello "*responsive*", ed entrambi TTC e PET per il caso "*un-responsive*". In particolare, sono stati identificati tre livelli di severità, chiamati "*safe passage*", "*mild interaction*" e "*critical interaction*": per il modello "*responsive*", l'evento è un annotato come "*safe passage*" se $TTC >$

$2.5\ s$, "*mild interaction*" se $1.2\ s < TTC \leq 2.5\ s$ e "*critical interaction*" se $TTC \leq 1.2\ s$; diversamente, per il modello "*un-responsive*", l'evento è un "*safe passage*" se $TTC > 2.3\ s$ e $PET > 2.6\ s$, "*critical interaction*" se $TTC < 1.3\ s$ e $PET < 1.0\ s$, in tutti gli altri casi un "*mild interaction*". Similmente, Ni e coll. [89] hanno proposto il concetto di tre modelli di interazione, chiamati "*hard interaction*", "*no interaction*" e "*soft-interaction*", osservabili analizzando i profili delle velocità e degli indicatori TTC e PET durante gli incontri veicolo-pedone nelle intersezioni regolate da semaforo. Attraverso l'approccio *Support Vector Machine* (in grado di trattare problemi multi-classe), gli autori sono riusciti a classificare tre livelli di severità, caratterizzati da valori di soglia pari a $1.5\ s$ e $3.0\ s$ sul TTC, e $1.0\ s$ e $3.0\ s$ sul PET.

Un ulteriore approccio per selezionare i valori di soglia, che si sta diffusamente affermando nel contesto scientifico, è rappresentato dalla teoria degli eventi estremi [184]. Si tratta di una branca della statistica finalizzata alla modellazione dei comportamenti stocastici degli eventi estremi. Questi ultimi hanno valori insolitamente grandi o piccoli, come i valori nelle code di una distribuzione normale. In particolare, attraverso il confronto tra le stime fornite degli indicatori SSM e le frequenze effettivamente osservate degli incidenti, la teoria degli eventi estremi ha dimostrato di essere uno strumento efficace per valutare la sicurezza del traffico [178], [185]–[187]. Tra le esperienze più recenti, Zhang e coll. [85] hanno applicato questa statistica per modellare i valori di PET e TTC, e selezionare i valori di soglia più appropriati per annotare gli eventi di quasi incidente tra veicolo e pedone in intersezioni regolate da semaforo. I valori che gli autori hanno individuato sono pari a $6\ s$ per il PET e $3\ s$ per il TTC. Le classi di severità sono state definite di conseguenza: quando $PET \leq 6\ s$ e $TTC \leq 3\ s$, l'interazione è definita "*serious conflict*"; quando uno degli indicatori è inferiore al valore di soglia, l'interazione è del tipo "*slight conflict*"; quando nessuno dei due indicatori è inferiore ai valori di soglia, l'evento è considerato "*safe*", ovvero sicuro. Anche Borsos e coll. [178] hanno utilizzato la teoria degli eventi estremi per confrontare gli indicatori TTC e T_2, e studiare sotto quali condizioni questi SSM risultassero più adatti per valutare la sicurezza del traffico. Utilizzando i dati acquisti da telecamere in una intersezione regolata da semaforo, relativi alle interazioni tra veicoli in svolta a sinistra e quelli che procedono sulla strada principale, le analisi condotte dagli autori hanno concluso che i due indicatori sono trasferibili in quanto conducono alle medesime stime della probabilità di collisione, ma con valori di soglia più bassi per T_2 rispetto al TTC: per esempio, la stessa probabilità di collisione corrispondente ad un valore di soglia pari a $5\ s$ per il TTC, si ottiene con un valore di circa $2.5\ s$ sul T_2.

2.3 Un nuovo sistema di valutazione della sicurezza dei pedoni

Sulla base della letteratura presentata nelle Sezioni precedenti, il lavoro di ricerca esposto in questa tesi ha affrontato l'elaborazione di un nuovo sistema in grado di prevedere l'evoluzione dei livelli di rischio nel corso delle interazioni veicolo-pedone nel traffico urbano. Lo schema funzionale di riferimento è rappresentato in Figura 6. Il sistema, nella sua formulazione concettuale, prevede all'origine l'estrazione e l'elaborazione delle traiettorie degli utenti della strada interagenti dalle nuvole di punti acquisite da uno o più RADAR a onde millimetriche ad alta risoluzione [188].

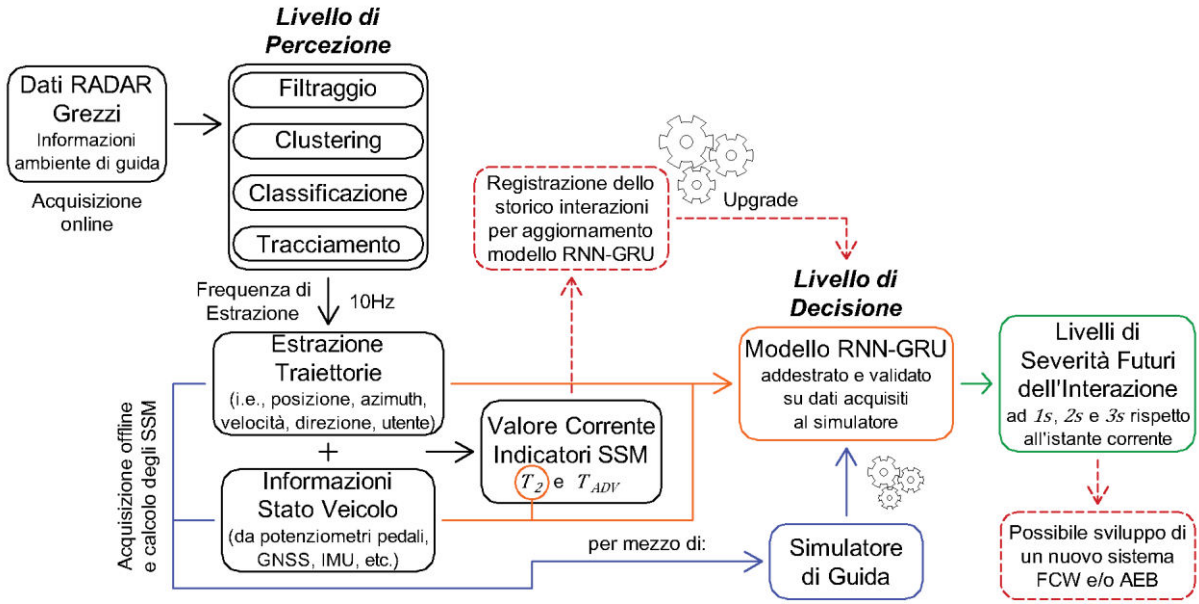


Figura 6: Il nuovo sistema di valutazione del livello di severità basato sull'interazione veicolo-pedone.

I sensori RADAR, installati nella parte anteriore del veicolo, eseguono la scansione dell'ambiente di guida, rilevando contemporaneamente sia gli utenti in movimento che gli oggetti sullo sfondo (ad esempio, alberi, edifici, barriere di protezione, etc.). L'operazione di elaborazione dei dati grezzi, che definisce il livello di percezione del sistema e la cui finalità è l'estrazione delle traiettorie, si compone di ulteriori quattro fasi: il filtraggio dell'ambiente, il *clustering* degli oggetti, la classificazione ed il tracciamento dei pedoni. L'obiettivo del filtraggio è quello di escludere, dalla nuvola di punti acquisita, i dati che caratterizzano lo "sfondo", come alberi, edifici e punti del terreno, ma allo stesso tempo preservando le informazioni relative agli utenti della strada. Il *clustering* mira a identificare i punti che appartengono allo stesso oggetto per crearne una sagoma. Siccome, nel caso generale, diversi utenti (come veicoli, pedoni e ciclisti) possono interessare l'ambiente di guida simultaneamente, è fondamentale distinguere i pedoni dal risultato del *clustering* attraverso classificazione; infatti, il processo di *clustering* non permette l'annotazione delle sagome rilevate. Infine, il *tracking* dei pedoni viene effettuato per determinare le traiettorie, le quali, al termine del processo di elaborazione, contengono informazioni come la posizione relativa, l'azimut orizzontale, la velocità relativa e la direzione di movimento del pedone rispetto al veicolo. Queste grandezze possono essere espresse in termini assoluti attraverso una trasformazione prospettica e grazie alle informazioni sullo stato cineto-dinamico del EV. Ovviamente, la frequenza di acquisizione delle traiettorie dipende sia dalle capacità intrinseche di rilevamento dei RADAR che degli algoritmi di supporto alle operazioni di elaborazione dei dati grezzi. Successivamente all'estrazione della traiettoria del pedone e all'acquisizione delle informazioni sull'EV, gli indicatori T_{ADV} e T_2 sono calcolati per l'istante di guida corrente secondo la procedura descritta in Appendice A. Quindi, le informazioni rilevanti del processo di incontro (estratte dalle traiettorie) vengono passate ad una RNN di tipo GRU per stimare e anticipare il livello di

rischio degli incontri veicolo-pedone, grazie all'ottimizzazione preliminarmente effettuata su dati acquisiti attraverso un simulatore di guida avanzato ad elevato grado di realismo. Ovviamente, la raccolta di dati durante la guida su strada ed il loro successivo invio ad un server dedicato per l'annotazione dei livelli di gravità effettivi e l'aggiornamento del sistema, potrebbe rendere gli algoritmi proposti più efficienti e facilitare l'adattamento a nuove situazioni operative. Il modello GRU costituisce il fulcro del livello di decisione del sistema e, una volta stabilite le regole per l'algoritmo di decisione (ovvero in quali condizioni debba essere emesso il segnale di "avviso al conducente" e in quali quello di "frenata di emergenza") potrebbe essere sfruttato per l'annotazione anticipata degli eventi di quasi-incidente e, di conseguenza, per il potenziamento dei sistemi FCW e/o AEB esistenti.

I dettagli del processo sopra esposto saranno trattati nei successivi Capitoli. Tuttavia, vale la pena evidenziare in questa Sezione alcuni aspetti rilevanti che differenziano il sistema proposto dalla letteratura esistente in materia, e giustificare le scelte metodologiche effettuate.

In primo luogo, la tesi presenta principalmente un prototipo del modello GRU per la previsione su più orizzonti temporali (a 1, 2 e 3 s rispetto all'istante di guida corrente) dei livelli di rischio nelle interazioni tra un veicolo e un pedone in un contesto urbano. Questa previsione viene effettuata rispetto alla rappresentazione corrente della situazione di traffico che sarebbe tratta dai dati grezzi di tre RADAR ad onde millimetriche, installati sul fronte anteriore del veicolo. A tal fine, un esperimento di simulazione di guida è stato progettato e validato per raccogliere caratteristiche di mobilità sequenziali su un campione di 65 studenti universitari nei confronti di incontri ad elevato rischio. Osservando il comportamento dei partecipanti a questi eventi è stato possibile acquisire profili comportamentali in termini di T_{ADV} e T_2, i quali sono stati sfruttati per l'annotazione accurata delle condizioni di sicurezza dei pedoni durante l'intero processo di incontro sulla base di valori soglia tratti dalla più recente letteratura scientifica. In particolare, gli eventi di quasi-incidente tra pedone e veicolo sono stati classificati in due sole categorie di severità, chiamate "*safe*" e "*unsafe*", poiché l'intenzione è quella di cogliere in anticipo le situazioni di conflitto effettivamente gravi; questo permetterebbe di migliorare il *timing* di attivazione dei sistemi esistenti, evitando inutili interferenze con il normale comportamento di guida. Quindi, definito un insieme minimo ed efficiente di caratteristiche del processo di incontro e traslando opportunamente nel tempo i livelli di severità, sono stati generati i vettori di input e di target per il processo di addestramento e validazione del modello neurale. La conoscenza acquisita dal modello in un tale contesto potrà essere sfruttata per facilitare l'adattamento online a nuove situazioni operative specifiche.

Le diverse fasi che caratterizzano il livello di percezione del sistema proposto (Figura 6) non sono oggetto del presente lavoro di tesi in quanto esistono diverse soluzioni alternative, validate e calibrate anche in contesti reali di guida, che permettono l'estrazione delle traiettorie dei pedoni (si ricorda, intese come informazioni dettagliate su posizione, azimut orizzontale, velocità e direzione di movimento) garantendo frequenze di campionamento di $10Hz$ o superiori [78], utilizzando sensori RADAR (o LIDAR, i quali generano nuvole di punti con i medesimi problemi di estrazione) e gli approcci di modellazione neurale

presentati alla Sezione 2.1.3. Alcune recentissime implementazioni di algoritmi per l'elaborazione delle nuvole di punti acquisite su un veicolo in movimento sono state sviluppate da Cennamo e coll. [189], Isele e coll. [190], Meyer e coll. [191], Lin e coll. [192]. Per le ragioni esposte, l'elaborazione presentata al Capitolo 3 si riferisce alle operazioni da compiere per passare dalle informazioni estratte dai dati grezzi (le traiettorie) alle variabili di input e target considerate per la modellazione. Ovviamente, nella simulazione di guida sono state riprodotte le condizioni reali di estrazione delle traiettorie più gravose; per cui i dati acquisiti durate le simulazioni sono registrati ogni $0.1\ s$.

La motivazione che giustifica la scelta della sensoristica RADAR [188] è legata alla migliore qualità del rilievo di tali sensori rispetto alle telecamere (sebbene queste garantiscano estrazioni anche a $30\ Hz$ [85]) e alla maggiore efficienza in condizioni meteorologiche avverse (nebbia, pioggia, neve) rispetto ai LiDAR (che sfruttano un segnale luminoso con una lunghezza d'onda minore rispetto alle onde radio). Tuttavia, non è esclusa la possibilità che nelle implementazioni future del sistema si possa considerare l'aggregazione delle informazioni acquisite da più sensori, tra cui telecamere, LiDAR e RADAR.

Rispetto alle procedure convenzionali di valutazione della sicurezza dei pedoni basate principalmente sulla previsione delle traiettorie o intenzioni future dell'EV e dell'agente target, l'algoritmo prototipato si differenzia per la capacità di comprendere ed anticipare l'evoluzione del rischio di collisione durante i processi di incontro. In altre parole, si evita di considerare il complesso problema multimodale della previsione delle traiettorie (si ricorda che per uno stesso comportamento esistono molteplici traiettorie che un utente può compiere) e, piuttosto, si cerca di acquisire direttamente dai dati il modo in cui gli utenti si comportano (o, meglio, interagiscono) in situazioni ad alto rischio per evitare una potenziale collisione. Questo è possibile grazie all'impiego di indicatori SSM in grado di descrivere in modo continuo il processo di interazione tra veicolo e pedone, e architetture neurali che possono lavorare con sequenze temporali. Infatti, il sistema prototipato rappresenta una delle prime esperienze di modellazione di parametri direttamente legati alla valutazione della sicurezza delle interazioni e la prima in assoluto ad effettuare una previsione su più orizzonti temporali di livelli di severità (precisamente a 1, 2 e $3\ s$, rispetto alla situazione corrente rilevata dai sensori). In particolare, come ulteriore aspetto rilevante, questa predizione multilivello dipende da una rappresentazione minima ed efficiente della situazione attuale di traffico: solo sei sono le variabili che compongono il vettore di input al modello neurale, le quali sono legate allo stato corrente degli agenti e alla loro prossimità temporale (come viene dettagliatamente descritto alla Sezione 3.4). Infatti, Ortiz e coll. [125] hanno dimostrato che non è necessario impiegare lo stato degli attuatori come caratteristiche rilevanti per prevedere il comportamento futuro di un EV (situazione molto comune negli studi sulla previsione delle traiettorie, presentati in Sezione 2.1.3): gli autori hanno predetto con semplici algoritmi di apprendimento (cioè, una *multi-layer perceptron neural network*) il comportamento di frenata (le primitive di comportamento) di conducenti in avvicinamento ad un semaforo, ottenendo ottime precisione su scale temporali fino a $3\ s$, usando come caratteristiche di input la velocità del veicolo, assieme alla colore della luce e alla distanza del semaforo.

Infine, vale la pena porre all'attenzione del lettore il fatto che l'approccio presentato in questa tesi ha alcuni aspetti in comune con la recente ricerca sviluppata da Zhang e coll. [85], soprattutto in merito alla modellazione attraverso GRU, ma gli autori prevedono il livello di severità attuale (basato su TTC e PET) delle interazioni veicolo-pedone in intersezioni regolate da semaforo, catturando la dinamica dell'evento attraverso telecamere in posizione fissa, mentre lo scopo del lavoro di tesi è una previsione su più orizzonti temporali di indicatori continui attraverso dati che potrebbero essere effettivamente acquisiti da un veicolo in movimento.

Capitolo 3

Raccolta ed Elaborazione dei Dati

Per raggiungere gli obbiettivi della ricerca (ed in generale di qualsiasi sistema FCW e/o AEB), è necessaria l'acquisizione di dati di mobilità relativi a molteplici incontri veicolo-pedone in scenari urbani reali. Inoltre, queste informazioni devono permettere la corretta valutazione degli indicatori di sicurezza per l'interazione durante l'applicazione online del sistema, ed essere facilmente acquisite da sensori di bordo (per esempio, potenziometri sui pedali, IMU, GNSS, così come RADAR, LiDAR e telecamere). In particolare, l'analisi della letteratura in materia [152], ha permesso di identificare tre gruppi principali di caratteristiche per lo studio del problema in questione: gli stati degli attuatori e del volante, le informazioni sulla dinamica del veicolo, la velocità ed il vettore direzione del pedone. Tuttavia, come discusso nella Sezione 2.3, è possibile (e preferibile) utilizzare una rappresentazione minima ed efficiente della situazione di traffico per effettuare predizioni su orizzonti temporali non troppo lunghi [125]; infatti, un modello che impegna caratteristiche irrilevanti, scarsamente correlate con la variabile target, può avere difficoltà di apprendimento, maggiori requisiti di memoria e complessità, oltre ad una minore capacità di generalizzazione [193]. Pertanto, in questo studio, i dati relativi allo stato degli attuatori e alla rotazione del volante non sono utilizzati in forma "grezza" ma accorpate per estrarre l'unica informazione davvero rilevante da essi celata, ovvero le intenzioni del conducente. Ad ogni modo, per raccogliere tutte queste informazioni garantendo la sicurezza dei conducenti, è stato reclutato un campione di 65 studenti universitari, dotati di patente in corso di validità, per partecipare ad un esperimento di simulazione di guida presso il Laboratorio Strade del Dipartimento Politecnico di Ingegneria e Architettura (DPIA) dell'Università di Udine. Infatti, i simulatori di guida possiedono alcuni vantaggi sugli studi di ricerca che implicano la guida su strada [194], tra cui:

1. Versatilità: il software di simulazione permette di ricreare ambienti di guida e scenari diversificati per soddisfare le esigenze della ricerca, riproducendo all'occorrenza situazioni atmosferiche avverse, come neve, vento e nebbia secondo diverse intensità;
2. Ambiente controllato e ripetibile: si tratta di un requisito fondamentale per confrontare i risultati di una indagine e/o migliorare l'accuratezza delle misure perché, mentre una sperimentazione in un contesto reale è del tutto casuale, i partecipanti ad una simulazione sono sottoposti alle stesse condizioni di traffico ed atmosferiche, ed affrontano i medesimi eventi progettati dall'ingegnere ricercatore (quindi, è possibile isolare le variabili sperimentali da altri fattori che potrebbero influenzare il comportamento di guida);
3. Sicurezza per il conducente: rispetto agli studi di guida naturalistici che osservano il comportamento su strada dei partecipanti senza preoccuparsi della loro sicurezza (perché l'obbiettivo è proprio la raccolta di informazioni su eventi incidentali o di quasi incidente che occorrono spontaneamente a causa dei comportamenti degli agenti e/o di fattori ambientali), i conducenti

sottoposti alla simulazione non corrono alcun rischio anche quando l'obbiettivo è l'analisi dei comportamenti in situazioni potenzialmente pericolose;

4. Costi ridotti: gli studi di simulazione non richiedono gli stessi costi di una guida su strada; basta pensare alle spese per il carburante, la manutenzione del veicolo, l'assicurazione e l'eventuale pedaggio se lo studio deve avvenire in un contesto autostradale.

L'uso di simulatori di guida avanzati per l'analisi dei comportamenti del conducente e lo sviluppo di sistemi FCW o AEB è una pratica accettata e ampiamente riconosciuta [195], poiché il comportamento del conducente osservato nella simulazione di guida mostra gli stessi modelli della guida nel mondo reale [196]: questo concetto è noto come validità relativa e sarà approfondito più dettagliatamente nella Sezione 3.1. In questo contesto, Saito e Raksincharoensak [197] hanno progettato e validato, con un simulatore di guida, un sistema AEB per la protezione dei pedoni che controlla i sottosistemi del veicolo (freno e acceleratore) in situazioni potenzialmente pericolose o incerte: in tali scenari, il sistema permette di decelerare il veicolo e mantenere una velocità di guida "sicura". Hou e coll. [198], studiando i comportamenti dei conducenti in tipici conflitti con ciclisti (riprodotti attraverso un simulatore di guida), hanno proposto un metodo per migliorare i tempi di attivazione e l'emissione dei segnali di controllo/avviso dei sistemi AEB per la protezione dei ciclisti. Bella e coll. [195] hanno recentemente utilizzato un simulatore di guida a base fissa (significa che il sistema non possiede la strumentazione hardware per riprodurre rollio e beccheggio durante la guida simulata) per valutare la funzionalità e l'efficacia di due differenti sistemi di avviso (ovvero attraverso allarme acustico o visivo) al conducente della presenza di un pedone che attraversa all'interno o all'esterno delle strisce pedonali. Gli autori hanno scoperto che entrambi i sistemi di assistenza hanno influenze positive sulla risposta dei conducenti (l'interazione veicolo-pedone ha un carattere meno severo e le manovre del conducente sono più rilassate), ma non sono state riscontrate differenze statisticamente significative sugli effetti delle due tipologie di avviso. Diversi anni prima, Choi e coll. [199] avevano condotto esperimenti di simulazione di guida per valutare la risposta dei conducenti, in termini di variazione di velocità del veicolo, ad un sistema integrato (attivo e passivo) di protezione dei pedoni (e, quindi, per definire il giusto *timing* di attivazione del sistema). Gli autori, attraverso una simulazione agli elementi finiti, hanno dimostrato che il loro sistema potrebbe ridurre di circa il 90% la fatalità dei pedoni in situazioni di attraversamento all'esterno delle strisce pedonali. In merito alla quantificazione dei benefici (in termini di vite salvate e riduzione della lesività) offerti dai sistemi AEB per la protezione degli utenti vulnerabili della strada (ciclisti e pedoni), Kovaceva e coll. [200] hanno presentato un approccio per combinare sistematicamente le informazioni acquisite con metodi diversi, sotto l'ipotesi che i simulatori di guida siano lo strumento preliminare per lo sviluppo dei sistemi ADAS: gli autori propongono di applicare l'inferenza Bayesiana per aggiornare le valutazioni basate sulla simulazione di guida con nuovi riscontri derivanti dall'applicazione nel mondo reale (attraverso studi di guida naturalistici) di un prototipo del sistema AEB sviluppato.

Il simulatore di guida impiegato per lo sviluppo del sistema oggetto di tesi è il modello *AS 1000-M*, a due gradi di libertà, prodotto dalla compagnia norvegese AutoSim AS©.

3.1 Il simulatore di guida AS 1000-M

La guida di un veicolo è un'attività complessa che il simulatore di guida *AS 1000-M* (Figura 7) del Laboratorio Strade è in grado di riprodurre con un elevato grado di realismo, fornendo un'esperienza sensoriale di livello superiore.

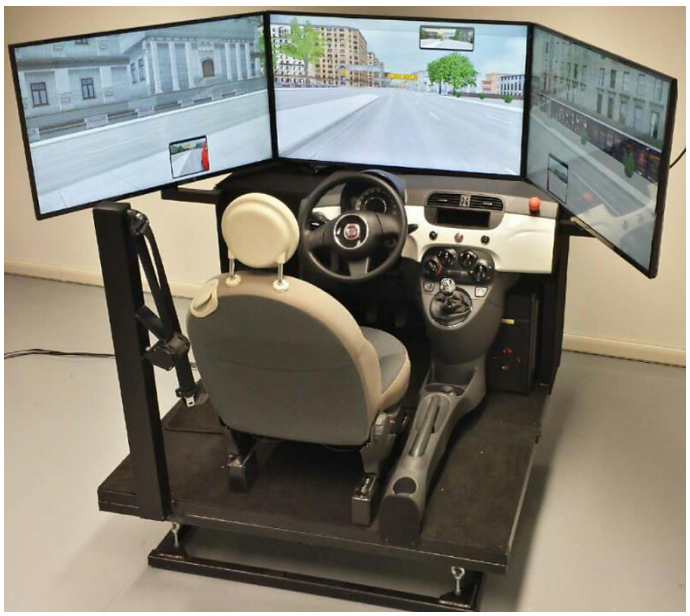


Figura 7: Il simulatore di guida AS 1000-M.

Questo dispositivo sofisticato è stato progettato utilizzando componenti reali della *city car* FIAT 500 (in particolare, il cruscotto, il volante, i pedali, la leva del cambio, il freno a mano, il sedile del conducente, i tappetini e la cintura di sicurezza), trasduttori sensoriali di vibrazione del motore ed attuatori estremamente rapidi di rollio (oscillazione attorno all'asse longitudinale) e beccheggio (oscillazione attorno all'asse trasversale) del veicolo nell'ambiente stradale virtuale. Infatti, l'intera cabina di pilotaggio del simulatore è montata su una base mobile a due gradi di libertà (una innovazione importante rispetto ai simulatori statici), la quale è in grado di riprodurre le forze apparenti a cui un guidatore sarebbe sottoposto durante la guida su strada (nei casi più semplici di puro rollio, la piattaforma si inclina verso sinistra nelle curve a destra, ovvero a destra nelle curve a sinistra, oppure nel caso di puro beccheggio, la piattaforma si inclina in avanti durante la frenata e all'indietro nelle accelerazioni). Quindi, tutto questo equipaggiamento, insieme al *force feedback* dello sterzo e agli effetti sonori (del motore, dei veicoli in movimento circostanti, del clacson, etc.) riprodotti da un sistema audio Hi-Fi con effetto 3D e Doppler, consentono al guidatore di avere risposte veloci, movimenti precisi e sensazioni realistiche. Inoltre, la possibilità del simulatore di muoversi (per lo meno in due dimensioni) permette di ridurre durante la sperimentazione il numero di casi di *cybersickness* [194], ovvero un malessere (nelle forme di nausea, vertigini ed affaticamento)

che può manifestare il partecipante a seguito dell'esposizione più o meno prolungata ad un ambiente virtuale. Un partecipante che presenta questo tipo di malessere può alterare i dati raccolti e deve in ogni caso essere escluso dal campione.

Tuttavia, la tecnologia di simulazione non si limita alle console e ai controller: le innovazioni ingegneristiche nel campo della modellazione 3D e *rendering* del mondo reale consentono di proiettarsi ovunque. Infatti, attraverso il software Unity® e gli *asset* dedicati alla costruzione di strade virtuali (per esempio, *EasyRoads3D*©), è possibile riprodurre qualsiasi contesto o ambiente di guida, anche nei dettagli del mondo fisico. I 3 monitor LCD da 43 pollici che coprono un campo visivo di quasi 180° rendono concreta l'immersione nel mondo virtuale riprodotto. Lo sperimentatore, attraverso una stazione di controllo dedicata (esterna al simulatore), può modificare le condizioni al contorno che caratterizzano il trinomio veicolo-strada-ambiente, azionare trigger di scenari, monitorare in tempo reale il comportamento di guida dei partecipanti alla simulazione e registrare i loro interventi sui pedali e sul volante (assieme ai parametri cineto-dinamici del veicolo) con passi temporali anche dell'ordine del centesimo di secondo (la frequenza più alta di campionamento impostabile è di $30\ Hz$, corrispondente a 3 centesimi di secondo); tutto questo viene controllato attraverso codici nel linguaggio di programmazione *Common Lisp* all'interno di una interfaccia grafica progettata ed implementata dalla stessa compagnia norvegese AutoSim, chiamata *Scenario Builder*©, la quale appunto permette la creazione di scenari di guida complessi in modo semplificato. Al termine della simulazione, questo software è in grado di produrre un file di riepilogo della simulazione che può essere memorizzato per ulteriori analisi di ricerca. Ogni *record* nel file rappresenta un dato sullo stato del veicolo (e sulle proprietà della scena, se opportunamente specificato attraverso il codice) all'istante di campionamento, e può includere il tempo trascorso dall'inizio della simulazione, la posizione del veicolo (nel sistema di riferimento della mappa virtuale), le componenti di accelerazione e di velocità del veicolo, la posizione rispetto al centro della corsia, la rotazione del volante, la forza sul pedale del freno e/o acceleratore, la marcia inserita, l'uso degli indicatori di direzione, le coordinate polari (lunghezza del segmento congiungente e angolo azimutale orizzontale rispetto alla direzione del veicolo) di un altro utente virtuale, etc.

Nonostante l'elevato grado di realismo attualmente raggiunto nella simulazione, affinché un simulatore di guida possa essere considerato un utile strumento di ricerca, è molto importante verificare l'affidabilità delle misure registrate durante la guida in realtà virtuale; questo processo è noto come validazione del simulatore. Infatti, bisogna riconoscere che una delle maggiori questioni inerenti all'uso di simulatori di guida è legata alla necessità di stabilire se il comportamento del partecipante alla sperimentazione possa differire o meno dalla guida nel mondo reale. Generalmente, si effettua una distinzione tra due livelli di validazione [201], detti "validità assoluta" e "validità relativa". La validità assoluta si riferisce alla corrispondenza numerica (quantitativa) tra il comportamento (in termini di parametri cineto-dinamici) del conducente nell'ambiente simulato e nel mondo reale. Diversamente, la validità relativa si riferisce alla corrispondenza qualitativa tra gli effetti indotti da variabili indipendenti sul comportamento di guida (come gli effetti delle violazioni alla

guida, della presenza di un pedone a bordo strada che intende attraversare, di un semaforo che improvvisamente diventa rosso). Una recente sperimentazione al simulatore del Laboratorio Strade ha permesso di validarne, in termini relativi, l'utilizzo per lo studio dei comportamenti dei conducenti durante interazioni potenzialmente pericolose con un pedone su strade urbane a quattro corsie (due corsie per senso di marcia), sia sotto gli usuali livelli attentivi sia sotto influenza di distrazioni cognitive [59]. Ovviamente, questo risultato non è generalizzabile a contesti di guida molto diversi o per situazioni di attraversamento inusuali (le quali richiederebbero una validazione ad hoc). Inoltre, uno studio di validazione rigoroso del simulatore di guida *AS 1000-M*, che confronti in situazioni analoghe i comportamenti dei conducenti durante una simulazione con dati acquisiti nel mondo reale, non è ancora stato effettuato nel contesto italiano (validazioni, sia in termini relativi che assoluti, sono state effettuate a livello aziendale dalla casa costruttrice del simulatore ma i risultati non sono fruibili pubblicamente). Tuttavia, considerando lo scopo della ricerca presentata in questa tesi, è richiesta solo la validità relativa (alla base di tutti gli studi presentati all'inizio del Capitolo 3): infatti, l'obbiettivo del sistema prototipato era quello di apprendere le dipendenze temporali intrinseche, in eventi potenzialmente pericolosi, tra le caratteristiche di mobilità osservate nell'istante corrente ed i livelli di rischio su diversi orizzonti nel futuro. Se consideriamo la letteratura presentata alla Sezione 2.2.2, che ha analizzato gli schemi di interazione tra pedone e veicolo sulla base dell'osservazione diretta del comportamento su strada, è possibile affermare che le dipendenze di interesse sono qualitativamente già note, anche considerando semplicemente lo schema teorico presentato da Laureshyn e coll. [76]. Inoltre, avendo validato una procedura per lo studio dei comportamenti in un contesto urbano [59], è possibile estendere questo risultato (il quale ha dimostrato che il comportamento dei conducenti in situazioni potenzialmente pericolose, osservato al simulatore, è pienamente coerente con il modello comportamentale atteso) a tutti gli studi che si svolgono nelle medesime condizioni di sperimentazione [202]. Ad ogni modo, si ricorda al lettore che l'intenzione della ricerca qui presentata è lo sviluppo di un prototipo come verifica di fattibilità del sistema di Figura 6 (Sezione 2.3), il quale sarà confermato ed aggiornato attraverso nuovi dati derivanti dalla sua applicazione nel contesto di guida reale (che necessariamente dovrà sfruttare la conoscenza appresa attraverso i dati del simulatore, come avviene in altri studi della letteratura presentati all'inizio della Capitolo 3). Infine, considerando l'affidabilità dei risultati sul comportamento dei conducenti nei confronti degli utenti vulnerabili della strada negli studi [15], [195], [197]–[199], [202] che hanno impiegato un simulatore di guida con caratteristiche simili o inferiori al modello *AS 1000-M*, si può affermare che ci sono sufficienti garanzie sulla validità del metodo utilizzato. Pertanto, nella Sezione 3.2 viene descritta in dettaglio la procedura di sperimentazione e ripostati i risultati dello studio di validazione [59].

3.2 La sperimentazione al simulatore AS 1000-M

La ricerca presentata in questa tesi ha impiegato due scenari stradali, con finalità diverse, che rappresentano rispettivamente un contesto suburbano e urbano, puramente immaginari ma ispirati ad alcune località norvegesi e

soprattutto a quartieri della città di Tromsø. Infatti, tali scenari (o *terrain* nel gergo tecnico) sono stati realizzati dalla compagnia norvegese AutoSim (rispettando gli standard europei) e resi disponibili per studi di ricerca a tutti coloro che possiedono un simulatore dell'azienda. In particolare, la sperimentazione vera e propria che ha permesso di acquisire le informazioni necessarie allo sviluppo del modello GRU si è svolta nel *terrain* urbano, osservando il comportamento dei partecipanti nei confronti di due incontri di traffico pianificati (ed opportunamente progettati): un ragazzo, nel primo (Figura 8a), ed una ragazza, nel secondo (Figura 8b), entrano in un attraversamento pedonale zebrato, partendo dal bordo del marciapiede.

(a) ***(b)***

Figura 8: Vista frontale da veicolo degli scenari simulati: (a) ragazzo che attraversa; (b) ragazza che attraversa.

Queste situazioni di traffico sono state pianificate su un percorso all'interno dell'ambiente urbano, indicato al partecipante attraverso frecce di colore verde che apparivano sullo schermo centrale, il quale poteva richiedere circa 15 minuti per essere completato, considerando il limite di velocità di $50\ km/h$. Su tale percorso, i partecipanti hanno affrontato diverse intersezioni regolate da semaforo, curve strette (a $90°$), brevi rettilinei e attraversamenti occasionali di pedoni, alcuni dei quali sono avvenuti all'esterno delle strisce pedonali. Gli incontri pianificati, che sono stati utilizzati per registrare i dati (e calcolare gli indicatori SSM per ogni partecipante in modo offline, seguendo la procedura descritta nell'Appendice A) sono stati impostati su una strada a quattro corsie (due corsie per senso di marcia), con segnaletica verticale ed orizzontale conformi agli standard europei (Figura 8). Un flusso di traffico, definito come densità di veicoli attorno all'EV, è stato fissato a 6 veicoli in un raggio di $300\ m$: si tratta di una condizione necessaria per popolare l'ambiente di guida ma molto modesta per evitare qualsiasi influenza del traffico sul comportamento del conducente (come riscontrato nella letteratura pertinente [4], [202]).

Gli scenari sono stati progettati come segue: 1) il partecipante arriva ad un semaforo rosso a circa $200\ m$ dalle strisce pedonali, sulle quali ha una chiara visuale, 2) quando il partecipante riprende la guida, il pedone (inizialmente nascosto) cammina con un angolo di $90°$ verso la strada per poi fermarsi sul bordo del marciapiede, 3) nel momento in cui il veicolo, in base alla sua velocità corrente, si trova a circa $3\ s$ dalle strisce pedonali (il concetto è analogo a quello di T_2), il pedone inizia l'attraversamento e mantiene una velocità pari ad $1/10$ di

quella del veicolo (corrispondente a circa $1.4\ m/s$ per una velocità del veicolo di $50\ km/h$). Queste condizioni, coerenti con la letteratura in materia [31], [85], impongono al partecipante di reagire repentinamente per rallentare e dare la precedenza al pedone, altrimenti sarebbe inevitabile una collisione. In questo modo, i dati raccolti si riferiscono a manovre di arresto che hanno evitato una collisione con il pedone, per pochi centimetri o diversi metri, a seconda della capacità ed attenzione di ciascun partecipante (perché il conducente di un veicolo generalmente cambia il proprio comportamento di guida solo quando percepisce il potenziale conflitto con il pedone, che non necessariamente coincide con il momento in cui il pedone inizia l'attraversamento); in altre parole, è stata ricreata una situazione potenzialmente pericolosa ma, sebbene sia stato imposto un *trigger* basato sulla vicinanza temporale, in realtà la risposta del guidatore dipende dalle sue capacità percettive e abilità di guida.

Sessantacinque studenti universitari sono stati reclutati per l'esperimento. Cinque partecipanti sono stati esclusi dalle analisi, in quanto non hanno potuto completare l'esperimento per l'eccessivo disagio (*cybersickness*). Il reclutamento è avvenuto inviando una richiesta di partecipazione agli indirizzi di posta elettronica di tutti gli studenti iscritti per l'A.A. 2019/2020 ad uno dei seguenti corsi di laurea triennale, presso l'Università degli Studi di Udine: Ingegneria Civile, Scienze Agrarie, Giurisprudenza e Relazioni Pubbliche. Per partecipare all'esperimento, lo studente doveva avere un'età compresa tra i 20 e i 30 anni, essere in possesso di una patente di guida in corso di validità, aver percorso almeno $5000\ km$ alla guida di un veicolo nell'ultimo anno e non soffrire di patologie che avrebbero potuto compromettere l'attività di guida. Vale la pena sottolineare che la partecipazione all'esperimento è stata volontaria, non c'era alcuna ricompensa monetaria, e tutti i partecipanti hanno dato il proprio consenso informato dopo essere stati istruiti sulla procedura della prova di guida simulata. Al contrario, non sono stati informati sugli obiettivi della ricerca, per evitare effetti di influenza sul comportamento. Infine, lo studio è stato condotto in conformità con la Dichiarazione di Helsinki, e il protocollo è stato approvato dal Comitato Etico Locale dell'Università di Udine.

In merito alla procedura per l'esecuzione della sperimentazione, tutti i partecipanti sono stati adeguatamente addestrati all'uso degli attuatori del simulatore prima della guida sperimentale e hanno potuto testare le loro capacità su un percorso della durava di circa 5 minuti, all'interno del *terrain* suburbano (in cui il limite di velocità varia tra i 50 e i $60\ km/h$). Per ripristinare condizioni psicologiche simili a quelle iniziali e per limitare la possibile assuefazione o stanchezza dei partecipanti, al termine della guida suburbana e prima della simulazione urbana, è stata inserita una pausa di 5 minuti durante la quale ogni partecipante ha compilato un questionario volto a raccogliere i suoi dati personali (ad esempio, età, esperienza di guida, etc.) e ha completato le Matrici Progressive Colorate di Raven [203] per acquisire indirettamente informazioni sulle sue capacità intellettive non verbali. Al contrario, alla fine della sperimentazione, un secondo questionario sul disagio percepito durante la guida è stato compilato da ogni partecipante per identificare ed eliminare dal campione quei guidatori che avevano sperimentato un eccessivo disagio. La procedura appena presentata è stata in parte ispirata ad altri studi di letteratura finalizzati all'analisi delle interazioni veicolo-pedone [31], [195]. Quindi, il campione

effettivamente utilizzato per lo sviluppo del modello GRU comprende 19 femmine e 41 maschi, con un'età media di 24.0 (deviazione standard ± 3.43) e 24.1 (± 1.91) anni ed un quoziente intellettivo non verbale (QI) medio di 33.4 (± 2.50) e 33.9 (± 1.66), rispettivamente. Pertanto, il campione di maschi e femmine può essere considerato approssimativamente bilanciato per età e QI.

Nella Sezione 3.2.1 seguente, vengono presentati i risultati dello studio [59] di validazione del simulatore per l'analisi dei comportamenti di guida in ambito urbano.

3.2.1 Validazione della simulazione

In primo luogo, è fondamentale chiarire al lettore che il fine ultimo dello studio [59] era esaminare gli effetti di una conversazione telefonica, per mezzo di altoparlanti, sul comportamento di frenata/arresto di giovani utenti della strada (in questo caso studenti universitari di età fra i 20 e i 30 anni) nei confronti di un pedone che attraversa sulle strisce pedonali (nelle medesime condizioni descritte alla Sezione 3.2). Questo argomento è ancora poco rappresentato nella letteratura scientifica ma le poche evidenze sperimentali hanno verificato che l'uso del telefono cellulare in vivavoce (nel contesto scientifico si usa spesso il termine "*hands-free*") durante la guida influisce significativamente sulle prestazioni dei conducenti. Siccome questo studio ha rappresentato la prima esperienza di utilizzo del simulatore *AS 1000-M* del Laboratorio Strade per l'analisi dei comportamenti, il tema della sperimentazione ha permesso di stabilire una condizione di studio per la validazione di questo innovativo strumento tecnologico; questa condizione è rappresentata dal confronto qualitativo e dall'analisi delle manovre dei conducenti. In particolare, le manovre acquisite con il simulatore sono in linea con i risultati delle ricerche sui profili di velocità dei conducenti in avvicinamento ad un attraversamento zebrato, sia in condizioni di normale attenzione alla guida [4], [31], [177], [202], sia quando i livelli cognitivi sono ridotti dal coinvolgimento in una conversazione telefonica [58]. Questo confronto ha confermato la validità relativa, già dichiarata dalla casa costruttrice del simulatore. Pertanto, nel seguito saranno descritti i risultati della sperimentazione sugli effetti indotti da un uso legale del telefono cellulare alla guida (ovvero per mezzo di altoparlanti); tuttavia, prima di tutto, vengono brevemente riportati gli aspetti che differenziano la sperimentazione corrente con quella relativa allo sviluppo del modello GRU (Sezione 3.2).

Caratteristiche della sperimentazione

Per questo esperimento, procedendo al pari di quanto descritto alla Sezione 3.2, sono stati reclutati ottanta studenti universitari e solo due sono stati esclusi dalle analisi per aver sperimentato *cybersickness*. Tuttavia, per lo scopo dello studio, i partecipanti sono stati divisi in due gruppi, uno di controllo ed uno sperimentale, e per entrambi le analisi sono state concentrate sul momento in cui una bambina entra in un attraversamento zebrato dal marciapiede (Figura 8b).

Il gruppo di controllo ha guidato attraverso lo scenario sperimentale senza alcun compito cognitivo imposto. In questo modo, sono stati ottenuti dati di riferimento sul comportamento in condizioni di "normale" attenzione alla guida (con le fluttuazioni attese dei livelli attentivi in percorsi monotoni). Questi dati, acquisiti ad intervalli di $0.25\ s$, includevano: il tempo dall'inizio della simulazione,

posizione di veicolo e pedone nel sistema cartesiano della mappa virtuale, il modulo del vettore di velocità di entrambi gli utenti, la pressione sui pedali del freno ed acceleratore, l'orientamento spaziale, assieme ad informazioni accessorie per la verifica delle condizioni di marcia, quali la componente di accelerazione (decelerazione) longitudinale, la marcia corrente inserita e la pressione sul pedale della frizione.

Diversamente, al gruppo sperimentale è stato chiesto di guidare effettuando una conversazione al telefono, la quale è stata studiata per diminuire la quantità di risorse cognitive assegnate al compito di guida. In particolare, durante la pausa di 5 minuti tra la prova in ambito suburbano e la sperimentazione urbana, solo ai partecipanti del gruppo sperimentale è stata mostrata una storia a fumetti, composta da sei scene. Le storie possibili erano tre, ovvero la Storia del Vaso di Fiori [204], la Storia del Nido [205] e la Storia del Litigio [206]. Queste storie sono bilanciate per numero di concetti, parole e frasi che possono suscitare. L'ordine di somministrazione di queste storie è stato ruotato da partecipante a partecipante al fine di minimizzare il possibile bias costituito da storie con strutture concettuali differenti. In particolare, le storie sono state mostrate ai partecipanti del gruppo sperimentale su un PC portatile, girato verso di loro in modo che l'esaminatore potesse dichiarare di non conoscerne il contenuto. In questo modo, è stato minimizzato il possibile effetto esercitato dalla condivisione del referente. Ai partecipanti è stato chiesto di immaginarsi mentalmente la storia riportata nelle vignette stimolo e di non riferirla in quel momento. All'inizio della guida in ambito urbano, lo sperimentatore ha chiamato ogni partecipante del gruppo sperimentale, ha chiesto di attivare la funzione vivavoce e di posizionare il telefono in una posizione comoda e sicura; quindi, è rimasto in linea per tutta la durata del percorso. Per circa 10 minuti, gli studenti hanno guidato indisturbati sul percorso urbano indicatogli da frecce verdi sullo schermo centrale. Giunti al semaforo rosso che precede l'attraversamento, lo sperimentatore ha chiesto al partecipante collegato di raccontare la storia visionata in precedenza. Durante la narrazione della storia (circa un minuto), i conducenti del gruppo sperimentale hanno affrontato l'evento di traffico pianificato (come descritto alla Sezione 3.2). Infine, gli ultimi minuti di guida sono stati privi di distrazioni.

Caratteristiche dei partecipanti

Nella pausa di 5 minuti, tra la prova in ambito suburbano e la sperimentazione urbana, ai partecipanti di entrambi i gruppi è stato fornito un questionario, che oltre a raccogliere i dati personali, ha indagato l'uso del telefono cellulare durante la guida.

Il gruppo di controllo era composto da 15 maschi e 11 femmine, con un'età media di 23.5 ($\pm$1.55) e 22.6 ($\pm$1.96) anni ed un QI non verbale medio di 33.7 ($\pm$1.63) e 32.3 ($\pm$3.06), rispettivamente. Il gruppo sperimentale comprendeva 30 maschi e 22 femmine, con un'età media di 24.4 (2.14) e 23.8 (3.36) anni ed un QI non verbale medio di 33.9 (1.72) e 33.2 (2.24), rispettivamente. Di conseguenza, i due gruppi sono stati equilibrati rispetto all'età e al QI.

Dei 78 partecipanti, l'80.8% era in possesso di una patente di guida da almeno tre anni; quindi, il restante 19.2% era costituito da neopatentati che, per legge, devono rispettare regole di guida più severe rispetto ai conducenti esperti.

Il 52.6% degli studenti reclutati percorreva mediamente meno di $10^{\cdot}000\ km$ all'anno, il 41.0% tra $10^{\cdot}000$ e $20^{\cdot}000\ km$ all'anno, la restante percentuale più di $20^{\cdot}000\ km$ in un anno tipo. Negli ultimi tre anni, il 20.5% era stato coinvolto in un incidente stradale mentre il 11.5% aveva ricevuto una contravvenzione per superamento dei limiti di velocità ($Art.142$ del Nuovo Codice della Strada). Diversamente, nessuno dei partecipanti aveva ricevuto un avviso di infrazione per passaggio con il semaforo rosso ($Art.146$), uso del telefono cellulare ($Art.173$) o guida sotto l'effetto di alcol e/o stupefacenti ($Art.186-187$).

Alla domanda "Usa il cellulare durante la guida?", il 71.8% dei partecipanti ha risposto di sì. Questi 56 giovani conducenti hanno risposto a domande più specifiche sul tipo e frequenza di utilizzo del telefono cellulare alla guida, di cui sono riportati solo i dati più allarmanti (una panoramica completa è fornita in Tabella 2): Il 7.1% dei partecipanti ha riferito di utilizzare il proprio telefono durante la guida, tenendolo alla mano e vicino all'orecchio; almeno una volta al giorno, il 28.6% utilizza il cellulare durante la guida per rispondere ed effettuare chiamate, il 25.0% per leggere ed inviare messaggi di testo e il 5.4% per leggere e-mail o navigare sul Web; infine, il 3.6% ha conversazioni telefoniche durante la guida che durano più di 10 minuti.

Tabella 2: Frequenze nell'uso del telefono cellulare durante la guida.

Violazioni alla guida	**Conteggio**	**Percentuale***
Utilizzo del telefono cellulare alla guida:		
Si	56	71.8
No	22	28.2
Frequenza nell'uso del cellulare per effettuate o rispondere a una chiamata durante la guida:		
Almeno una volta al giorno (frequente)	16	28.6
Una o due volte alla settimana (moderatamente frequente)	22	39.3
Una o due volte al mese o all'anno (poco frequente)	16	28.6
Mai	2	3.6
Tempo medio trascorso al telefono durante la guida per le chiamate:		
<1 minuto	27	48.2
2-5 minuti	23	41.1
6-10 minuti	4	7.1
>10 minuti	2	3.6
Utilizzo di altoparlanti o auricolari durante la guida per rispondere/effettuare chiamate:		
Sempre	34	60.7
A volte	18	32.1
Mai (cellulare usato alla mano)	4	7.1
Frequenza nell'uso del cellulare per leggere o scrivere messaggi durante la guida:		
Almeno una volta al giorno (frequente)	14	25.0
Una o due volte alla settimana (moderatamente frequente)	12	21.4
Una o due volte al mese o all'anno (poco frequente)	14	25.0
Mai	16	28.6
Frequenza nell'uso del cellulare per leggere e-mail o navigare in internet durante la guida:		
Almeno una volta al giorno (frequente)	3	5.4
Una o due volte alla settimana (moderatamente frequente)	6	10.7
Una o due volte al mese o all'anno (poco frequente)	3	5.4
Mai	44	78.6

* le percentuali sono calcolate sul totale dei partecipanti (78) per la domanda relativa all'utilizzo del cellulare alla guida, mentre sono calcolate rispetto al numero di soggetti che ne hanno ammesso l'uso (56) per tutte le domande successive.

Elaborazione preliminare dei dati acquisiti

L'impatto della distrazione telefonica e il potenziale effetto di genere sul comportamento dei conducenti sono stati valutati considerando la velocità del veicolo e le sue variazioni nelle vicinanze dell'attraversamento pedonale. Inoltre, sulla base del metodo svedese (Sezione 2.2.1), l'interazione tra veicolo-pedone è stata analizzata calcolando il parametro TA (Tabella 1), il quale è legato al momento di inizio dell'azione evasiva compiuta dal conducente (nel caso specifico della simulazione di traffico progettata). Una pratica comune negli studi di letteratura [31], [58], [202], [207] è quella di considerare l'inizio della manovra di frenata/arresto del veicolo (l'azione evasiva in questione) con il momento in cui il partecipante inizia a diminuire la sua velocità, rilasciando completamente il pedale dell'acceleratore; questa condizione viene fatta coincidere anche con il momento in cui il guidatore rileva la presenza del pedone sul bordo del marciapiede e percepisce il potenziale rischio (una condizione necessaria per l'inizio dell'azione evasiva).

In particolare, il comportamento di frenata/arresto è stato caratterizzato osservando i profili di velocità dei partecipanti nel tratto di circa $100\,m$ che precede l'attraversamento zebrato [4] ed acquisendo (o calcolando) gli indicatori operativi suggeriti dalla letteratura [31], [202], [208]:

- V_i e L_{Vi}: la velocità (o "velocità iniziale") e la relativa distanza longitudinale dalle strisce pedonali quando il conducente decide di rilasciare il pedale dell'acceleratore e diminuire la velocità del veicolo, avendo percepito il pedone sul bordo del marciapiede;
- $TA = L_{Vi}/V_i$: Il tempo rimanente al verificarsi di una collisione dal momento in cui il conducente inizia l'azione evasiva;
- V_b e L_{Vb}: la velocità e la relativa distanza longitudinale dalle strisce pedonali quando il conducente ha azionato i freni;
- L_{Vmin}: distanza longitudinale dal punto di conflitto in cui è stata registrata la minima velocità V_{min} o l'arresto del veicolo;
- $\sigma(V_n)$: la deviazione standard della velocità del veicolo durante la manovra di frenata/arresto, chiamata anche "fluttuazione della velocità" [208].

Vale la pena ricordare al lettore che la condizione di simulazione, per cui il pedone parte quando il veicolo si trova a circa $3\,s$ dall'attraversamento pedonale in base alla sua velocità corrente, costringe necessariamente il partecipante a fermarsi o a ridurre drasticamente la velocità iniziale del veicolo. Per questo motivo, la minima velocità registrata durante la manovra non è stata considerata tra le variabili di studio, mentre risulta particolarmente rilevante la vicinanza spaziale al termine della manovra definita da L_{Vmin}. La Figura 9 mostra i profili di velocità media dei conducenti (abbozzati per mezzo di V_i e L_{Vi}, V_b e L_{Vb}, V_{min} e L_{Vmin}) per ciascuno dei quattro gruppi considerati nello studio: MC - conducenti maschi nel gruppo di controllo, ME - conducenti maschi nel gruppo sperimentale, FC - conducenti femmine nel gruppo di controllo, FE - conducenti femmine nel gruppo sperimentale.

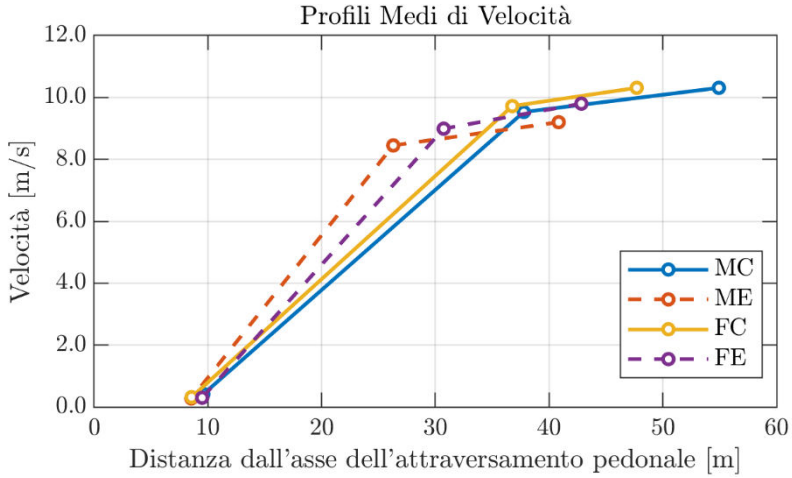


Figura 9: Profili di velocità medi dei conducenti. Legenda: MC - conducenti maschi del gruppo di controllo; ME - conducenti maschi del gruppo sperimentale; FC - conducenti femmine del gruppo di controllo; FE - conducenti femmine del gruppo sperimentale.

In merito all'elaborazione dei dati sulla guida simulata per l'acquisizione dei parametri sopra esposti, il profilo di velocità di ciascun partecipante è stato analizzato per individuare le condizioni di marcia, ovvero i tre momenti caratteristici della manovra: l'inizio della decelerazione, l'azionamento dei freni, quindi l'arresto del veicolo o il raggiungimento del minimo locale di velocità. L'individuazione dell'istante in cui il conducente ha alzato completamente (e in modo definitivo) il piede dal pedale dell'acceleratore non è stata banale in quanto i conducenti possono rimuovere il piede dall'acceleratore anche per cambiare marcia; di conseguenza, oltre ad osservare il comportamento sul pedale dell'acceleratore, è stato necessario porre particolare attenzione sia alla pressione sulla frizione che alla marcia correntemente inserita. Diversamente, per il calcolo delle distanze dal punto di conflitto nelle diverse condizioni di marcia e del TA, le posizioni di veicolo e pedone nel sistema di riferimento cartesiano (x, y) della mappa virtuale sono state elaborate per ricavare le coordinate polari del vettore di posizione del pedone rispetto al fronte del veicolo (r_0, θ): r_0 è la posizione relativa tra veicolo e pedone, mentre θ rappresenta l'angolo azimutale orizzontale del pedone rispetto alla direzione di marcia del veicolo. Uno schema esemplificativo è riportato in Figura 10. Inoltre, bisogna considerare che le posizioni del veicolo acquisite durante la simulazione si riferiscono al baricentro D del veicolo (Figura 10); questo punto è di scarso interesse poiché la parte del veicolo che può collidere con il pedone è il suo fronte anteriore. Pertanto, la prima operazione è stata quella di passare dalle coordinate del punto D a quelle del punto A (Figura 10): si tratta di un calcolo trigonometrico piuttosto semplice che dipende dalla distanza fissa (pari ad $1.00\ m$, per il modello di veicolo simulato) tra baricentro e paraurti frontale, e dall'orientamento spaziale del veicolo. Di conseguenza, nota la posizione del punto A e del punto B (Figura 10) è possibile calcolare r_0 come distanza euclidea:

$$r_0 = \sqrt{(x_B - x_A)^2 + (y_B - y_A)^2}, \tag{3.1}$$

e l'angolo azimutale θ in base alla formula trigonometrica:

$$\theta = -\arctan\left(\frac{y_B - y_A}{x_B - x_A}\right). \tag{3.2}$$

In genere, la deviazione della direzione di marcia rispetto all'asse della corsia è rappresentata da un angolo piuttosto piccolo che produce errori dell'ordine del millimetro sulle distanze calcolate; per questa ragione, può essere trascurato nei calcoli seguenti [167].

La distanza longitudinale L_1 (Figura 10) tra il potenziale punto di conflitto ed il veicolo in una qualsiasi condizione di marcia è espressa da:

$$L_1 = r_0 \cos\theta. \tag{3.3}$$

In questo modo è stato possibile associare alle velocità V_i, V_b e V_{min} le rispettive distanze longitudinali L_{Vi}, L_{Vb} e L_{Vmin} dall'attraversamento zebrato, e quindi, definire i profili di velocità in Figura 9, mediando il risultato ottenuto per ciascun gruppo.

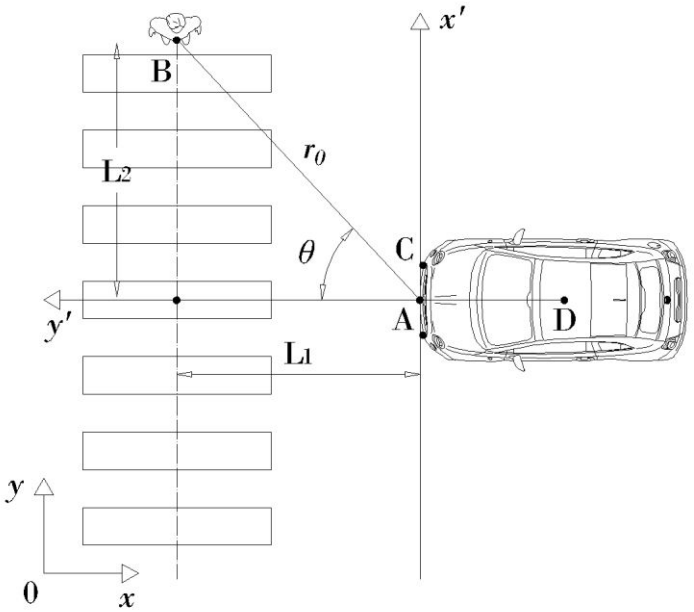


Figura 10: Schema di riferimento per il calcolo delle distanze longitudinali tra pedone e veicolo.

Verifica delle ipotesi

L'analisi della varianza a due vie (o ANOVA a due vie) permette di testare l'impatto di due variabili indipendenti su una o più variabili dipendenti. Il vantaggio di usare un'ANOVA a due vie, rispetto ad altri approcci statistici, è legato alla possibilità di testare, oltre agli "effetti principali" (cioè l'effetto complessivo di ogni variabile indipendente, come il genere, il livello attentivo, l'età, etc.), anche un effetto di interazione, cioè quando l'effetto di una variabile indipendente è influenzato da un'altra grandezza; per esempio, si potrebbe sospettare che le interazioni tra un veicolo ed un pedone siano più pericolose in presenza di distrazione ma solo per i conducenti maschi. Esistono due diverse versioni

dell'ANOVA a due vie: ANOVA tra gruppi (quando esistono due gruppi distinti di persone) e ANOVA a misure ripetute (quando le stesse persone vengono testate in più di una occasione).

Le variabili selezionate sono state analizzate utilizzando il metodo ANOVA a due vie tra gruppi, assumendo l'appartenenza al gruppo (1 – gruppo di controllo; 2 – gruppo sperimentale) e il genere (1 – conducenti maschi; 2 – conducenti femmine) come fattori tra partecipanti, e i parametri di guida (TA, V_i, L_{Vi}, V_b, L_{Vb}, L_{Vmin}, $\sigma(V_n)$) come variabili dipendenti.

L'applicazione del test ANOVA comporta la verifica di alcune ipotesi di base: l'assenza di outlier nel campione, la normalità delle variabili dipendenti e l'omogeneità della varianza tra i gruppi. In particolare, i box-plot non hanno individuato alcun outlier nei dati acquisiti (o calcolati) per ciascuna variabile nei quattro sottogruppi. La normalità è stata valutata mediante il test di Shapiro-Wilk (valore di significatività a livello $p > 0{,}05$): ad eccezione del parametro L_{Vb} associato al gruppo di conducenti maschi del gruppo sperimentale (gruppo ME), tutte le altre variabili dipendenti hanno soddisfatto l'ipotesi di distribuzione normale dei valori registrati per ciascuno dei quattro sottogruppi (Tabella 3). Tuttavia, la dimensione del campione (il gruppo ME è composto da 30 studenti) e la robustezza del test ANOVA nei confronti della violazione di questo presupposto [209] hanno reso possibile l'applicazione del metodo anche per la variabile L_{Vb}. Infine, il test di Levene (valore di significatività a livello $p > 0{,}05$) ha dimostrato che la varianza di ciascuna variabile dipendente era approssimativamente uguale tra i gruppi (ultima riga di Tabella 3). La significatività degli effetti di gruppo e di genere è stata valutata utilizzando medie non ponderate e somme dei quadrati di tipo III [210].

Tabella 3: Verifica delle ipotesi del test ANOVA. Legenda: le prime quattro righe riportano il risultato del test di normalità di Shapiro-Wilk (valori p) per ognuno dei quattro gruppi considerati nello studio (cioè FE, FC, ME, MC) e separatamente per ogni variabile del profilo di velocità; diversamente, l'ultima riga mostra il risultato (valori p) del test di omogeneità della varianza (o test di Levene) che verifica se la varianza dei punteggi è uguale per ognuno dei quattro sottogruppi, separatamente per ogni variabile del profilo di velocità.

Gruppo	Test	Risultati in termini di valore p (valore di significatività per $p > 0.05$)						
		TA	V_i	L_{Vi}	V_b	L_{Vb}	L_{Vmin}	$\sigma(V_n)$
FE	Shapiro-Wilk	0.505	0.545	0.888	0.867	0.950	0.944	0.178
FC	Shapiro-Wilk	0.936	0.956	0.502	0.981	0.565	0.143	0.379
ME	Shapiro-Wilk	0.835	0.183	0.130	0.159	0.000	0.208	0.727
MC	Shapiro-Wilk	0.144	0.921	0.225	0.759	0.851	0.359	0.198
Tutti	Levene	0.192	0.511	0.812	0.463	0.463	0.991	0.294

Oltre ai parametri considerati del profilo di velocità, la letteratura in materia suggerisce ulteriori parametri di interesse per caratterizzare le manovre di fermata/arresto dei partecipanti alla simulazione come il tasso medio di decelerazione (ovvero la decelerazione media per passare da V_i a V_{min}) [31], [202], [211] e le fluttuazioni di posizione [208] rispetto al centro della corsia (cioè le deviazioni standard della distanza tra il centro della corsia ed il baricentro del veicolo). Tuttavia, nessuna di queste variabili supplementari ha soddisfatto i requisiti per l'applicabilità dell'analisi ANOVA, né in termini di normalità né di

omogeneità della varianza tra i gruppi; per questo motivo, non sono state considerate nella seguente discussione.

Risultati e loro discussione

La Tabella 4 riporta i valori medi delle variabili considerate per caratterizzare il comportamento dei conducenti.

Tabella 4: Valore medio (deviazioni standard) delle variabili sul profilo di velocità per i gruppi di controllo e sperimentale.

Gruppo	$TA\ (s)$	$V_i\ (m/s)$	$L_{Vi}\ (m)$	$V_b\ (m/s)$	$L_{Vb}\ (m)$	$L_{Vmin}\ (m)$	$\sigma(V_n)\ (-)$
Controllo	5.76 [a] (1.91)	10.61 [a] (1.62)	59.28 [a] (15.68)	9.61 [a] (2.05)	37.35 [a] (14.46)	9.17 [a] (2.89)	3.85 [b] (0.82)
Sperimentale	5.23 [a] (1.52)	9.73 [b] (1.51)	50.90 [b] (16.01)	8.67 [b] (1.63)	28.17 [b] (11.35)	8.95 [a] (2.72)	4.21 [a] (0.68)

Nota: lettere diverse in apice indicano differenze statisticamente significative tra i gruppi ($p < 0{,}05$).

Oltre alle informazioni riportate in Tabella 4, è interessante notare che i partecipanti del gruppo di controllo e di quello sperimentale hanno impiegato in media un tempo simile per passare dalla V_i a V_{min}: un t-test a campioni indipendenti è stato condotto per confrontare il tempo di riduzione della velocità (SRT, dall'inglese *Speed Reduction Time*, ovvero il tempo impiegato per passare da V_i a V_{min}) tra conducenti distratti e non distratti. Il risultato del t-test ha confermato che non c'è stata una differenza significativa nei punteggi ($t(39.02) = -1.57, p = 0.13$) tra gruppo di controllo ($9.86\ s$) e quello sperimentale ($8.42\ s$). Pertanto, in media, i partecipanti hanno percepito la situazione di pericolo molto prima dell'attraversamento effettivo del pedone ($\mathrm{SRT} >> 6s$, [207], [211]).

Nel seguito vengono descritti ed analizzati con maggiore dettaglio i risultati per ciascuna delle grandezze considerate.

TA – *Tempo effettivo disponibile per evitare una collisione*

In merito al parametro di interazione TA tra veicolo e pedone al momento dell'azione evasiva compiuta dal conducente, l'analisi ANOVA (seconda colonna di Tabella 4 ha evidenziato l'assenza di una differenza statisticamente significativa nei comportamenti dei partecipanti, qualsiasi sia il gruppo di appartenenza ($F(1,74) = 1.67, p = 0.20$) o il loro genere ($F(1,74) = 0.02, p = 0.89$). Inoltre, non è stato individuato alcun effetto di interazione gruppo-genere nel campione osservato ($F(1,74) = 0.00, p = 0.98$). Di conseguenza, il valore medio di TA per i due gruppi è molto simile (gruppo di controllo: $5.76\ s$; gruppo sperimentale: $5.23\ s$, vedi Tabella 4). Il grafico a barre che rappresenta i valori medi e le corrispondenti deviazioni standard della variabile TA, per i maschi e le femmine in ciascun gruppo, è riportato in Figura 11.

Questo risultato ha confermato quanto osservato in precedenza per l'SRT: la reazione del conducente è tale per cui al pedone viene concesso in media un margine di sicurezza sufficiente per passare prima dell'auto; pertanto, il comportamento di guida osservato ($TA > 5\ s$) ha garantito in media la sicurezza del pedone [177]. Si tratta di una conclusione che aggiunge prove ad un crescente filone di ricerca sulla distrazione cognitiva: il risultato in termini di TA

dimostra che conversare su un telefono cellulare in vivavoce durante la guida può non avere un aumento associato del rischio di collisione [212] e, in particolare per questo studio, nessun aumento in media del rischio per la sicurezza del pedone. Tuttavia, questo riscontro non esclude la potenziale presenza di effetti sulle prestazioni di guida del conducente, in particolare di natura compensativa, come si può riscontrare nella letteratura esistente [213], [214]. Per una descrizione esaustiva dei potenziali motivi per cui prestazioni di guida alterate non si traducano in un corrispondente aumento del rischio di incidente, si rimanda a Dingus e coll. [212].

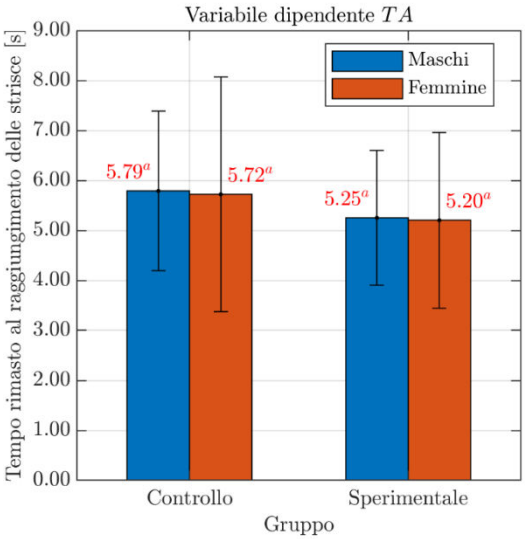


Figura 11: Effetto di genere e distrazione sulla variabile TA. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi della variabile TA, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo.

Pertanto, è possibile procedere ad un confronto (a parità di rischio per la sicurezza dei pedoni) tra i profili di velocità dei partecipanti, allo scopo di rilevare la presenza di effetti significativi della distrazione cognitiva sulle capacità di guida e, di conseguenza, validare ulteriormente il simulatore di guida, assieme allo "*story retelling*" (ovvero la narrazione di una storia) come metodo efficace per aumentare il carico di lavoro mentale in un test di simulazione di guida. Per effettuare tale confronto, la velocità e la distanza dei veicoli dal potenziale punto di conflitto sulle strisce pedonali sono stati analizzati separatamente, come suggerito (ed effettuato) da Bella e coll. [31], [202], [214] e Saifuzzaman e coll. [208].

V_i, L_{Vi} – Velocità iniziale del conducente e distanza associata dal punto di conflitto

Per quanto riguarda la velocità iniziale V_i, corrispondente al momento in cui il conducente decide di rilasciare il pedale dell'acceleratore per rallentare, l'analisi ha evidenziato la presenza di una differenza statisticamente significativa tra i gruppi di controllo e sperimentale ($F(1,74) = 4.94, p < 0.03$), caratterizzata da una dimensione "media" dell'effetto ($\eta_p^2 = 0.06$, secondo il criterio di Cohen [215]). È

interessante notare che il test ANOVA non abbia riscontrato alcun effetto principale legato al genere ($F(1,74) = 1.21, p = 0.28$), o effetto di interazione gruppo-genere ($F(1,74) = 0.46, p = 0.50$). Similmente, per la distanza L_{Vi} dal punto di conflitto associata a V_i, la differenza tra i gruppi è risultata statisticamente significativa con una dimensione media dell'effetto ($F(1,74) = 4.22, p < 0.04; \eta_p^2 = 0.05$), mentre nessun effetto di genere ($F(1,74) = 0.02, p = 0.90$) o di interazione ($F(1,74) = 0.36, p = 0.55$) è stato riscontrato. I grafici a barre che rappresentano i valori medi e le corrispondenti deviazioni standard delle variabili V_i e L_{Vi}, per i maschi e le femmine in ciascun gruppo, sono presentati in Figura 12.

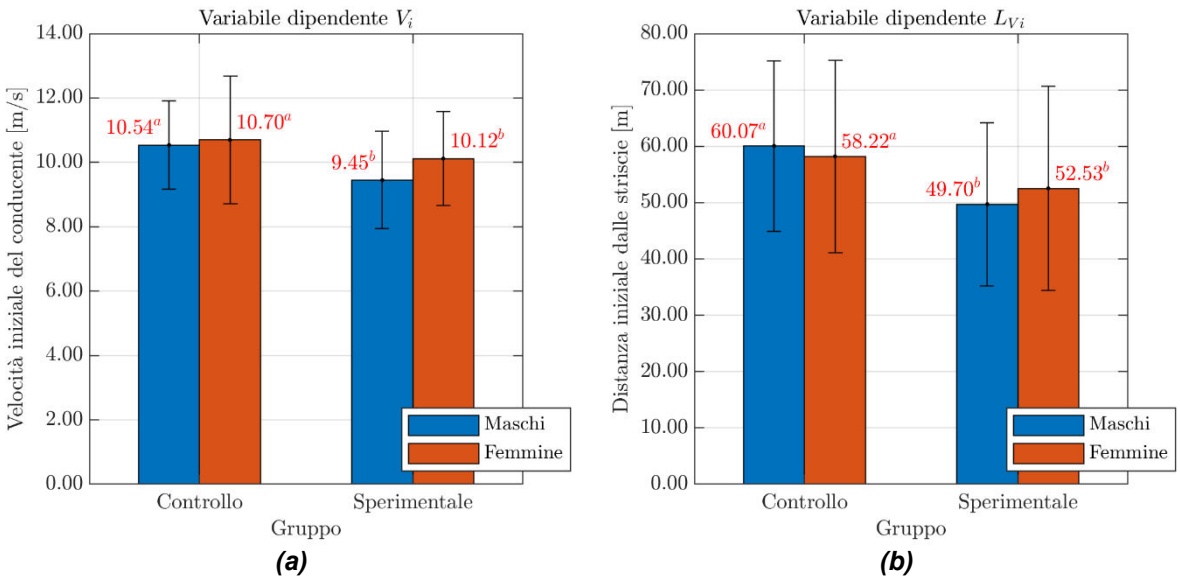
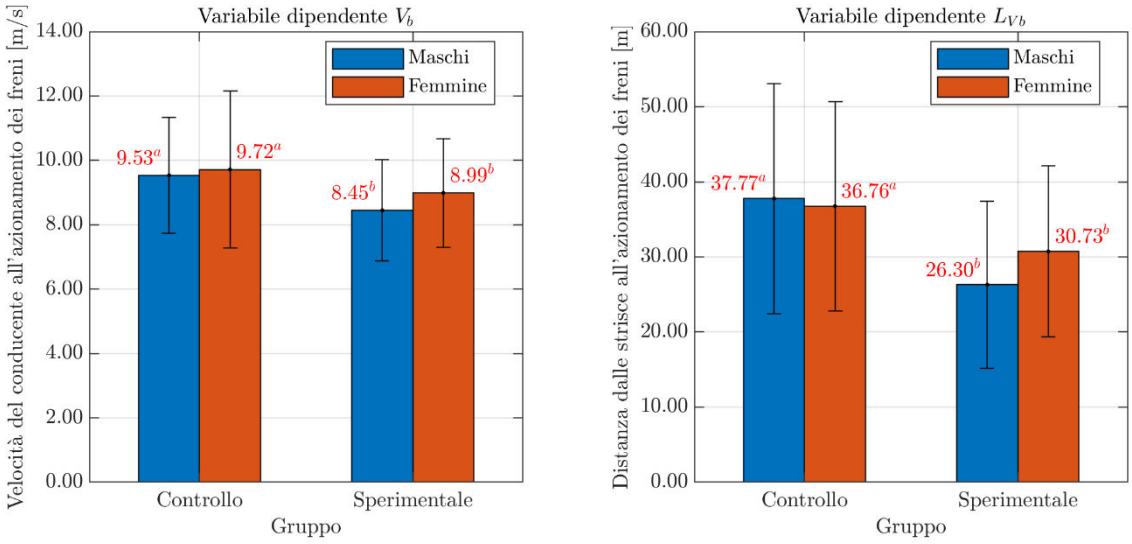


Figura 12: Effetto di genere e distrazione sulle variabili (a) V_i e (b) L_{Vi}. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi delle variabili V_i e L_{Vi}, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo.

Nel complesso, questi risultati suggeriscono che i conducenti (sia maschi che femmine) distratti dalla conversazione telefonica a vivavoce procedono ad una velocità mediamente inferiore ($9.73 < 10.71\, m/s$, terza colonna di Tabella 4) rispetto a quelli con usuali livelli attentivi alla guida, ma iniziano la frenata significativamente più tardi rispetto al gruppo di controllo ($50.90 < 59.28\, m$, quarta colonna di Tabella 4). Gli effetti sulla velocità iniziale potrebbero essere interpretati come un effetto compensativo [216] per la riduzione delle risorse cognitive destinate all'esperienza di guida [217]. Inoltre, le posizioni medie L_{Vi} suggeriscono che l'impegno nel compito cognitivo ha indotto una percezione ritardata del pedone sul marciapiede [58], poiché il punto in cui il conducente ha iniziato a rallentare è stato più lontano dalle strisce pedonali per il gruppo di controllo ($59.28\, m$, quarta colonna prima riga di Tabella 4) rispetto a quello sperimentale ($50.90\, m$, quarta colonna seconda riga di Tabella 4).

V_b, L_{Vb} – Velocità e distanza associata quando il conducente ha azionato i freni

Per quanto riguarda la velocità del veicolo V_b ($8.67 < 9.61\ m/s$, quinta colonna di Tabella 4) e la sua distanza dal punto di conflitto L_{Vb} ($28.17 < 37.35\ m$, sesta colonna di Tabella 4) al momento di azionamento dei freni da parte del conducente, le analisi hanno evidenziato la presenza di una differenza significativa legata al gruppo con un effetto medio per V_b ($F(1,74) = 4.38, p < 0.04$; $\eta_p^2 = 0.06$) e grande per L_{Vb} ($F(1,74) = 8.33, p < 0.01$; $\eta_p^2 = 0.10$). In entrambi i casi, non è stato riscontrato alcun effetto di genere (V_b: $F(1,74) = 0.69, p = 0.41$; L_{Vb}: $F(1,74) = 0.32, p = 0.58$) o di interazione gruppo-genere (V_b: $F(1,74) = 0.17, p = 0.68$; L_{Vb}: $F(1,74) = 0.80, p = 0.37$). I grafici a barre che rappresentano i valori medi e le corrispondenti deviazioni standard delle variabili V_b e L_{Vb}, per i maschi e le femmine in ciascun gruppo, sono presentati in Figura 13.



(a) ***(b)***

Figura 13: Effetto di genere e distrazione sulle variabili (a) V_b e (b) L_{Vb}. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi delle variabili V_b e L_{Vb}, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo.

Questi risultati hanno confermato quanto riscontrato in precedenza per le variabili V_i e L_{Vi}, mostrando che la distrazione indotta influenza negativamente la percezione e risposta media dei conducenti all'evento di traffico pianificato.

L_{Vmin} – Distanza dal punto di conflitto al termine della manovra

In merito alla distanza dal punto di conflitto L_{Vmin} alla fine della manovra di frenata/arresto, l'analisi non ha riscontrato alcuna differenza significativa legata al gruppo ($F(1,74) = 0.01, p = 0.92$), al genere ($F(1,74) = 0.00, p = 0.95$) o all'interazione gruppo-genere ($F(1,74) = 2.20, p = 0.14$). Il grafico a barre che rappresenta i valori medi e le corrispondenti deviazioni standard della variabile L_{Vmin}, per i maschi e le femmine in ciascun gruppo, è riportato in Figura 14.

Questo risultato suggerisce che, sebbene la distrazione cognitiva abbia un effetto significativo sulle prime fasi della manovra, il comportamento di compensazione del rischio attuato dai conducenti nel gruppo sperimentale è stato efficace per l'arresto del veicolo (oppure, per una drastica riduzione della velocità) ad una distanza media dal passaggio pedonale ($9.17\ m$, settima colonna prima riga di Tabella 4) paragonabile a quella dei partecipanti nel gruppo di controllo ($8.95\ m$, settima colonna seconda riga di Tabella 4). Quindi, si conferma ulteriormente che la sicurezza dei pedoni è stata in media garantita.

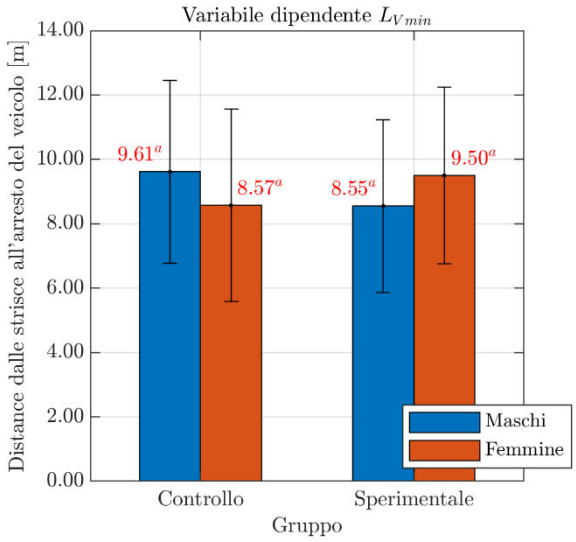


Figura 14: Effetto di genere e distrazione sulla variabile L_{Vmin}. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi della variabile L_{Vmin}, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo.

I risultati in termini di distanze iniziali L_{Vi} e finali L_{Vmin} dei conducenti dal potenziale punto di conflitto portano all'ovvia conclusione che gli spazi di arresto sono stati mediamente diversi tra i gruppi di controllo e sperimentale: un'analisi della varianza a due vie tra i gruppi sullo spazio di arresto (derivante dalla somma delle distanze percorse dal conducente dalla percezione del pedone al termine della manovra di frenata) ha evidenziato la presenza di una differenza statisticamente significativa legata al gruppo di appartenenza, con un effetto di media grandezza ($F(1,74) = 4.01, p = 0.05, \eta_p^2 = 0.05$). Non è stato riscontrato alcun effetto di genere ($F(1,74) = 0.02, p = 0.89$) o di interazione gruppo-genere ($F(1,74) = 0.11, p = 0.74$). Il valore medio dello spazio di arresto per i conducenti distratti ($41.90\ m$) è stato nettamente inferiore a quello dei conducenti non distratti ($50.10\ m$).

$\sigma(V_n)$ – *Fluttuazione della velocità*

Il grafico a barre che rappresenta i valori medi e le corrispondenti deviazioni standard della variabile $\sigma(V_n)$, per i maschi e le femmine in ciascun gruppo, è riportato in Figura 15.

Un effetto principale significativo, con una dimensione media dell'effetto, indotto dalla distrazione telefonica sulla fluttuazione della velocità è stato osservato tra i gruppi $(F(1,74) = 4.29, p < 0.04;\ \eta_p^2 = 0.06)$ ma non è stato evidenziato alcun effetto di genere $(F(1,74) = 3.25, p = 0.08)$ o di interazione gruppo-genere $(F(1,74) = 0.09, p = 0.76)$. Questo suggerisce che $\sigma(V_n)$ aumenta durante il coinvolgimento in un compito cognitivo: la fluttuazione della velocità (ultima colonna di Tabella 4) per i conducenti del gruppo sperimentale è stata del 7.1% superiore a quella dei partecipanti del gruppo di controllo. Infatti, Saifuzzaman e coll. [208] avevano evidenziato che i conducenti distratti possono avere difficoltà a tenere sotto controllo le variazioni di velocità e dimostrato che le fluttuazioni di velocità in un evento di *car-following* erano superiori in tale condizione di guida.

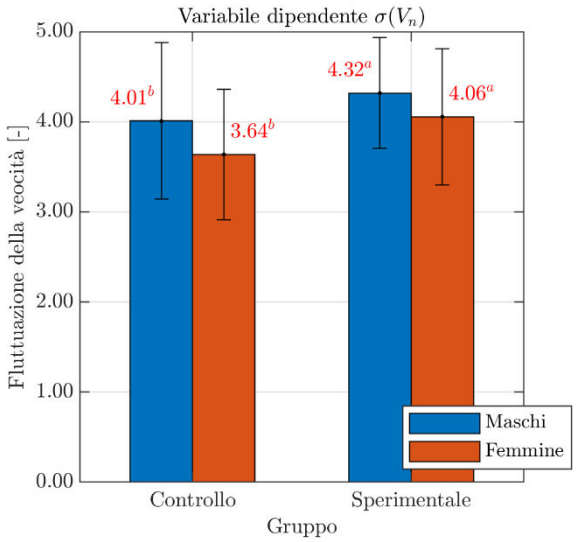


Figura 15: Effetto di genere e distrazione sulla variabile $\sigma(V_n)$. Legenda: i numeri in rosso sopra le barre rappresentano i valori medi della variabile $\sigma(V_n)$, per i maschi (barra blu) e le femmine (barra rossa) in ciascun gruppo. Le differenze statisticamente significative sono evidenziate da lettere diverse in apice. Le barre di errore della deviazione standard sono riportate sui grafici per ogni sottogruppo.

Conclusioni

Lo studio di validazione [59], che ha confrontato le manovre di frenata/arresto tra conducenti impegnati in una conversazione in vivavoce al telefono e quelli con normali livelli attentivi alla guida, ha confermato i risultati di diverse ricerche sui profili di velocità degli utenti in risposta ad un evento di traffico potenzialmente pericoloso. Infatti, i risultati dell'analisi statistica hanno confermato che le conversazioni al cellulare del tipo "*hands-free*" possono avere influenze significative sui comportamenti e le capacità percettive dei conducenti. Nello specifico, l'effetto indotto dalla conversazione è stato statisticamente significativo sui processi di selezione (in riferimento a V_i) e di mantenimento $(\sigma(V_n))$ della velocità, così come su quelli di percezione del rischio (L_{Vi}) e di decisione (V_b, L_{Vb}) dell'azione da attuare. Per ricapitolare brevemente i risultati dell'analisi:

1. I conducenti appartenenti al gruppo sperimentale hanno raggiunto in media una velocità iniziale inferiore a quella dei conducenti del gruppo di controllo, che sono stati lasciati liberi di guidare sul percorso urbano senza alcun compito cognitivo imposto. Quello osservato potrebbe essere un comportamento compensatorio per l'aumento del rischio associato al coinvolgimento in un compito secondario, ovvero la conversazione in vivavoce. Questo comportamento è stato infatti osservato e spiegato in tal senso in diversi studi che hanno analizzato i profili di velocità di conducenti distratti dal cellulare [216], [218].
2. L'aumento della deviazione standard associata al profilo di velocità medio dei conducenti del gruppo sperimentale, rispetto a quella di controllo, dimostra che la distrazione indotta dall'uso del cellulare alla guida compromette la capacità di mantenere un adeguato o desiderato andamento delle velocità durante l'interazione con un pedone che attraversa la strada, come affermato da Saifuzzaman e coll. [208].
3. Le distanze dal potenziale punto di conflitto nei due momenti fondamentali di una manovra di arresto, ovvero la percezione del pedone (L_{Vi}) e l'azionamento dei freni (L_{Vb}), tenute in media dai gruppi di controllo e sperimentale avvalorano l'ipotesi che i conducenti distratti abbiano percepito la presenza del pedone sul marciapiede (e quindi, la sua intenzione di attraversare la strada) in ritardo rispetto ai conducenti di controllo. Questo ritardo si è protratto fino all'azionamento dei freni che, di conseguenza è avvenuto significativamente più vicino al potenziale punto di conflitto [31].
4. A differenza della letteratura esistente, che raramente si è espressa sulle differenze di comportamento tra uomini e donne, lo studio ha evidenziato la compatibilità dei comportamenti tra conducenti di diverso genere, i quali sembrano avere le stesse prestazioni alterate dalle conversazioni telefoniche in risposta all'evento di traffico pianificato.

Sulla base dei risultati di cui sopra, è possibile affermare che esistono prove sufficienti che avvalorano la validità relativa del simulatore *AS 1000-M* e, contemporaneamente, lo *story retelling* come metodo efficace per riprodurre gli effetti della distrazione cognitiva sulle prestazioni di guida in un esperimento di simulazione. Quest'ultimo risultato non è banale perché, al momento, non è ancora stato identificato in maniera univoca e definitiva un protocollo per l'esecuzione di studi sulla guida simulata, volti a valutare le prestazioni di guida influenzate dal coinvolgimento in un compito cognitivo secondario. Pertanto, la procedura presentata nello studio [59] (e progettata principalmente da un punto di vista psicologico) può essere un utile riferimento per tutti i ricercatori che intendono studiare il comportamento di conducenti distratti in risposta a nuove soluzioni di progettazione stradale o ad eventi di traffico specifici.

Infine, lo studio [59] ha contribuito alla crescente evidenza scientifica che anche un uso legale del telefono cellulare alla guida (ovvero in vivavoce) può influenzare in maniera significativa le prestazioni di guida. Tuttavia, sono necessarie ulteriori valutazioni su un campione più ampio di conducenti, non solo per confermare ulteriormente le conclusioni presentate, ma per indagare e confrontare l'effetto della distrazione cognitiva dello *story retelling* su gruppi

diversi, come i conducenti più giovani e più anziani, considerando anche componenti comportamentali come l'aggressività.

3.3 Formulazione del problema di modellazione

In questa Sezione si riprende la trattazione relativa al problema originario, ovvero lo sviluppo di un modello GRU in grado di prevedere l'evoluzione della severità di incontri tra veicolo e pedone nel traffico urbano, ed in particolare viene affrontata la strategia di apprendimento e previsione alla base del processo di implementazione del modello.

Formalmente, il sistema di previsione della severità deve eseguire una "mappatura" (ovvero una associazione) tra la rappresentazione della situazione di traffico al tempo corrente di guida t, cioè le proprietà della scena e lo stato del veicolo tratti dai dati grezzi dei sensori di bordo, e il livello di rischio atteso dell'incontro veicolo-pedone, definito dagli indicatori T_2 e T_{ADV}, per i tempi $t + 1\ s$, $t + 2\ s$ e $t + 3\ s$. Tuttavia, si vuole ricordare al lettore che l'obbiettivo ultimo della ricerca è quello di eseguire l'apprendimento e la previsione in un veicolo reale in movimento, una condizione che si definisce "online" (perché deve avvenire in tempo reale in un veicolo che viaggia su strada). Questo implica che l'algoritmo di apprendimento scelto debba essere addestrato per un dato tempo t ed una scala temporale T_k al fine di rappresentare la relazione tra la situazione a $t - T_k$ ed il parametro target al tempo t, poiché non è possibile guardare nel futuro [125]. In pratica, con riferimento al problema in questione, una RNN del tipo GRU viene addestrata a prevedere accuratamente gli indicatori di sicurezza surrogati (o direttamente i livelli di severità, come descritto alla Sezione 3.5) agli istanti $t - 2\ s$, $t - 1\ s$ e t utilizzando le osservazioni del sistema di rilevamento del veicolo disponibili all'istante $t - 3\ s$. Una volta raggiunta una convergenza accettabile tra gli output predetti e i target sperimentali, il modello addestrato rappresenta il sistema di previsione a medio termine desiderato e può essere utilizzato per fare previsioni nel futuro basate sulla rappresentazione della scena attuale (ovvero, per prevedere i livelli di severità al tempo $t + T_k$ utilizzando la rappresentazione della situazione al tempo t). Questo processo è illustrato in Figura 16.

Ovviamente, alla base di questo discorso vi è l'assunzione che le diverse caratteristiche di input al modello neurale possano essere acquisite tutte simultaneamente e ad un intervallo di tempo regolato dalla più bassa frequenza di campionamento tra quelle dei sensori coinvolti (nel problema in questione, la frequenza è fissata a $10\ Hz$ per le ragioni esposte alla Sezione 2.3). Infine, considerando che la ricerca presentata in questa tesi si basa su un esperimento di simulatore di guida e l'intenzione dichiarata è quella di sviluppare un prototipo del modello GRU come verifica di fattibilità del sistema di Figura 6 (Sezione 2.3), il calcolo degli indicatori di sicurezza surrogati per ogni partecipante e l'addestramento dell'algoritmo di apprendimento sono stati entrambi eseguiti in modo offline, alla fine degli esperimenti di guida, registrando le opportune informazioni durante la sperimentazione (come se fossero acquisite da un insieme di più RADAR in un veicolo in movimento). Di conseguenza, dato che la procedura appena presentata permette di "guardare nel futuro", le fasi di apprendimento e previsione sono rese più semplici ma le prestazioni del processo di apprendimento non vengono influenzate.

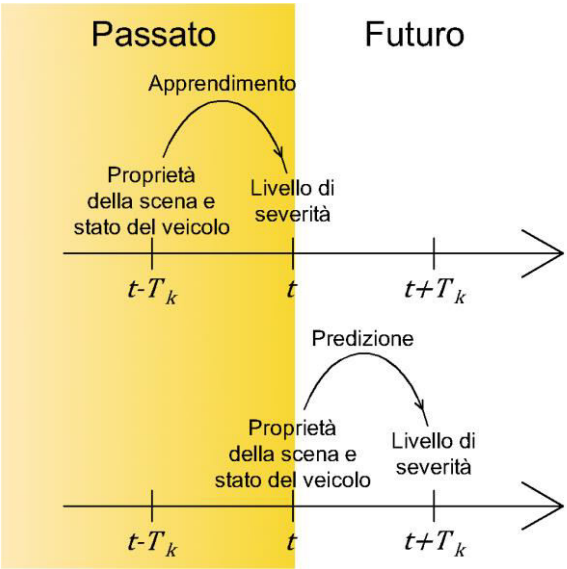


Figura 16: Illustrazione del paradigma di apprendimento e previsione. L'algoritmo di apprendimento mappa la rappresentazione della situazione passata (al tempo $t - T_k$) ed il livello di severità corrente (al tempo t); quindi, il modello addestrato predice il livello di severità futuro (al tempo $t + T_k$) usando la rappresentazione della situazione corrente (al tempo t).

3.4 Rappresentazione della situazione

Il sistema di coordinate (x', y') stabilito per il sistema di rilevamento del veicolo è mostrato in Figura 10. Le seguenti caratteristiche, estratte dai dati grezzi dei sensori, formano le sei componenti del vettore di input del modello di apprendimento e descrivono il processo di interazione tra veicolo e pedone al tempo presente t per ogni partecipante nel data set sperimentale: la primitiva di comportamento del conducente $A(t)$, il valore corrente di $T_2(t)$, la velocità dell'EV $v_v(t)$, la velocità del pedone $v_p(t)$ ed il vettore di posizione del pedone rispetto al fronte del veicolo $(r_0(t), \theta(t))$. La motivazione che spiega l'assunzione di queste grandezze è riportata di seguito, ma vale la pena evidenziare che tale scelta è stata anche convalidata da una analisi preliminare delle prestazioni del modello di apprendimento conseguente.

Come anticipato nella Sezione 2.1.2, le primitive di comportamento sono l'insieme di comportamenti elementari in cui la manovra di avvicinamento alle strisce pedonali del conducente (nella dimensione longitudinale dell'evento) può essere segmentata in base alle informazioni sulla velocità del veicolo e sulla componente longitudinale di accelerazione (per escludere eventuali condizioni di cambio marcia), assieme alle pressioni sui pedali dell'acceleratore e del freno (ricondotte ad un valore tra 0 e 1, per esprimere la massima corsa dei pedali) [125]. Questi dati, facilmente acquisibili sui veicoli moderni (ad esempio, attraverso potenziometri sui pedali, IMU e GNSS), vengono elaborati per definire una variabile categorica, i cui valori nel caso di conflitti veicolo-pedone possono essere 0 per la condizione definita "*stopped behavior*", 1 per "*braking behavior*" e 2 per "*keep speed behavior*". La Figura 17 riporta un esempio di segmentazione del comportamento per un conducente che ha partecipato alla sperimentazione

di guida simulata. Sulla base delle considerazioni offerte dallo studio di Ortiz e coll. [125], le porzioni del profilo in cui la velocità del veicolo v_v è inferiore a $2\,km/h$ (circa $0.56\,m/s$) sono annotate come "*stopped behavior*", in quanto il veicolo è praticamente fermo. La fase di "*braking behavior*" inizia nel momento in cui il conducente rilascia completamente il pedale dell'acceleratore (e, ovviamente, termina con la condizione $v_v(t) \leq 0.56\,m/s$), poiché nel traffico urbano questa condizione rappresenta l'inizio del rallentamento in risposta a un evento. Tutti gli altri istanti della manovra che non appartengono alle due categorie illustrate (ovvero corrispondenti alla condizione per cui $v_v(t) > 0.56\,m/s$, la pressione sul pedale dell'acceleratore è 0 mentre quella sul pedale dell'accelerazione nell'intervallo $(0,1]$, e la componente longitudinale di accelerazione è positiva) sono annotati come "*keep speed behavior*". La variabile categorica $A(t)$ relativa alle primitive di comportamento rappresenta un'informazione importante da fornire al modello GRU in quanto permette di mappare la sequenza dei livelli di severità futuri in base al comportamento attuale del conducente [125].

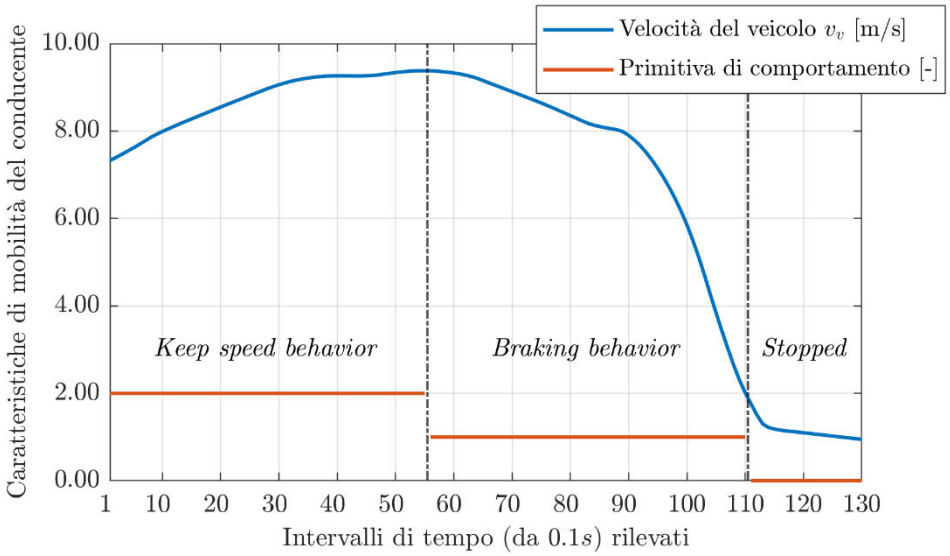


Figura 17: Esempio di segmentazione del profilo di velocità in primitive di comportamento per un partecipante della sperimentazione.

Allo stesso modo, il valore corrente del parametro T_2, il quale può essere calcolato in tempo reale usando le caratteristiche acquisite dai sensori (sulla base delle considerazioni presentate in Sezione 3.2.1 e alla procedura di calcolo descritta in Appendice A), riveste anch'esso un ruolo di primaria importanza poiché le neuroscienze hanno dimostrato che il cervello umano si basa attentamente sui giudizi del TTC (e, di conseguenza, del T_2 nel continuum, poiché i due indicatori sono trasferibili [178]) per eseguire azioni coordinate [139]. Si ricorda che Altché e de La Fortelle [137] hanno sfruttato il medesimo parametro per valutare la decisione del conducente di cambiare corsia in un contesto autostradale e, quindi, migliorare la previsione futura della traiettoria dell'EV.

Le componenti del vettore di posizione sono l'angolo azimutale orizzontale θ del pedone rispetto alla direzione di marcia del veicolo e la distanza relativa r_0 tra l'EV (punto A o C) ed il pedone (punto B, Figura 10). In particolare, v_p, r_0 e θ

sono i dati sullo stato del pedone acquisiti dai RADAR a onde millimetriche installati sulla parte anteriore del veicolo (punti A e C, Figura 10). In particolare, è stato ipotizzato un sistema di rilevamento dell'ambiente di guida costituito da un RADAR a lungo raggio (che opera nella banda di frequenza $77\,GHz$) con distanza massima di rilevamento pari a $160\,m$ e angoli azimutali nel range $\pm 10°$, installato in posizione centrale (nel punto A, Figura 10) sul fronte del veicolo, e da un ulteriore RADAR a medio raggio (nella banda di frequenza $79\,GHz$) su entrambi i lati del fronte anteriore (punto C, Figura 10), caratterizzato da una distanza massima di rilevamento pari a $50\,m$ ed angoli azimutali nel range $\pm 45°$; entrambi i RADAR hanno una risoluzione di $0.2\,m$ e una minima distanza di rilevamento pari a $2\,m$. Questa soluzione garantisce una copertura ottimale della scena per l'applicazione di sistemi AEB o FCW [167], [188]. Tuttavia, non si esclude la possibilità di utilizzare un unico RADAR *Multi-Mode* o *Full-Range* (alcuni dei quali, presenti attualmente sul mercato [219], possono raggiungere distanze di rilevamento massime di $350\,m$ e campi visivi di $120°$) in grado di adattare il campo visivo con la velocità del veicolo.

Come anticipato nella Sezione 2.3, le diverse fasi che caratterizzano l'elaborazione dei dati grezzi (Figura 6) acquisiti dai RADAR non sono oggetto del presente lavoro di tesi in quanto esistono diverse soluzioni alternative, validate e calibrate in contesti reali di guida, che permettono l'estrazione delle traiettorie del pedone (intese come $v_p(t)$, $r_0(t)$ e $\theta(t)$), garantendo frequenze di campionamento di $10\,Hz$ o superiori [78], [189]–[192]. Pertanto, al fine di riprodurre le condizioni reali di estrazione delle traiettorie più gravose, la simulazione è stata programmata per acquisire e registrare le sei componenti del vettore di input con un passo temporale di $0.1\,s$. Le sequenze temporali delle manovre hanno origine con il riconoscimento e tracciamento iniziali dello stato del pedone, eseguite sui dati grezzi del RADAR a lungo raggio, a circa $100\,m$ dalle strisce pedonali: è stata considerata una distanza di rilevamento massima inferiore a quella reale (ipotizzata di $160\,m$) per tenere conto delle complessità pratiche implicite nel processo di rilevamento di un pedone in un ambiente di guida complesso, come il contesto urbano [85]. Diversamente, le medesime sequenze temporali terminano quando il primo utente della strada (il pedone) lascia la zona di potenziale conflitto o il secondo utente della strada (il conducente) si ferma completamente: in entrambe le condizioni, nessuno degli indicatori SSM considerati può essere più calcolato (poiché la collisione non è più possibile).

3.5 Generazione delle classi di severità

Per lo sviluppo del livello di decisione del sistema teorizzato (Figura 6), sono stati considerati e implementati due approcci equivalenti (in termini di risultato finale) [152], ovvero 1) utilizzare più reti GRU in parallelo per modellare separatamente gli SSM (T_2 e T_{ADV}) che permettono di stimare la severità dell'interazione, o 2) addestrare una singola RNN come classificatore di sequenze per annotare direttamente gli eventi di quasi incidente, sulla base delle caratteristiche di mobilità rilevanti negli incontri di traffico veicolo-pedone (Sezione 3.4). Infatti, non possiamo affermare con certezza a priori quale dei due approcci conduca alle migliori prestazioni di classificazione.

Nel primo caso (Approccio 1), i valori continui del T_2 e le categorie definite per il T_{ADV} su tre orizzonti temporali futuri (ovvero, 1, 2 e $3\,s$ nel futuro) costituiscono il vettore di output e l'obiettivo di apprendimento di due modelli GRU distinti. Infatti, sebbene entrambi gli indicatori possano essere calcolati continuamente nel tempo (Sezione 2.2.1), T_{ADV} costituisce una funzione continua ma non derivabile nell'istante in cui avviene il passaggio tra le situazioni di "*crossing course*" e "*collision course*" (e viceversa): in questo passaggio, il valore assunto dal T_{ADV} diventa rapidamente zero e si mantiene tale fintanto che i due utenti della strada rimangono su rotte di collisione (poiché, per definizione, non esiste un vantaggio temporale in una situazione di "*collision course*") [76]. Solo quando si ha il passaggio dalla condizione di "*collision course*" a quella di "*crossing course*" a seguito dell'azione evasiva del conducente (corrispondente alla frenata o all'arresto del veicolo), il valore del T_{ADV} torna a crescere gradualmente per riflettere il vantaggio del pedone sul veicolo. Un tale comportamento induce punti angolosi nell'andamento del T_{ADV} durante lo svolgimento della manovra (cioè, se l'indicatore T_{ADV} venisse tracciato su un grafico in funzione del tempo, si otterrebbe una curva con almeno un punto di non derivabilità), i quali sono piuttosto difficili da rappresentare, anche con le moderne tecniche di regressione per apprendimento, a meno di sovra-adattare il modello neurale (causando un problema di *overfitting*, Sezione 4.2). Per evitare di incorrere in un tale problema pratico [220], è stato deciso di trattare la previsione degli stati futuri del T_{ADV} come un problema di classificazione binaria, mentre quella del T_2 come un semplice problema di regressione multi-obiettivo, siccome quest'ultimo indicatore rappresenta una funzione continua e derivabile in ogni punto del dominio temporale di estensione della manovra (per definizione, si ricorda che il T_2 fornisce un informazione sulla prossimità temporale di passaggio del secondo utente, il conducente del veicolo, per la zona di conflitto). Per quanto riguarda la classificazione del T_{ADV}, è necessario ribadire che la peculiarità di questo indicatore è la sua capacità di riflettere sia la sicurezza che l'efficienza dei processi di incontro, siccome i suoi bassi valori descrivono aspetti di sicurezza ($T_{ADV} = 0$ indica nessun margine, ovvero "*collision course*") mentre valori superiori a $2 - 3\,s$ rappresentano normali situazioni di traffico [76], [175], [176]. Allo stesso tempo, come anticipato nella Sezione 2.2.2, situazione critiche o potenzialmente pericolose di interazioni veicolo-pedone in un contesto urbano sono identificate da valori di $PET \leq 1.5\,s$ o, in maniera ancora più stringente secondo il metodo DOCTOR [172], $PET \leq 1.0\,s$ (si ricorda che il T_{ADV} amplia il concetto di PET, calcolando per ogni istante durante l'interazione il valore atteso del PET se la coppia di utenti mantenesse l'attuale velocità e traiettoria). Per queste ragioni e considerando la procedura di calcolo in Appendice A che considera tutti i potenziali punti del fronte del veicolo che possono entrare in collisione durante l'interazione, in questo lavoro di tesi i valori di T_{ADV} inferiori a $1\,s$ sono stati annotati come "*collision course*" (proprio perché rappresentano una situazione critica in cui il vantaggio di un utente sull'altro è troppo breve o del tutto assente) e quelli superiori a $1\,s$ come "*crossing course*". Quindi, le classi di T_{ADV} individuate permettono di identificare le condizioni future di reale pericolo durante l'interazione, soprattutto quando sono confrontate con la predizione del valore continuo di T_2. La Figura 18 mostra l'andamento del T_2, per un partecipante alla

sperimentazione (il medesimo di Figura 17), durante l'interazione con segmentazione del profilo in base alle soglie di pericolosità selezionate.

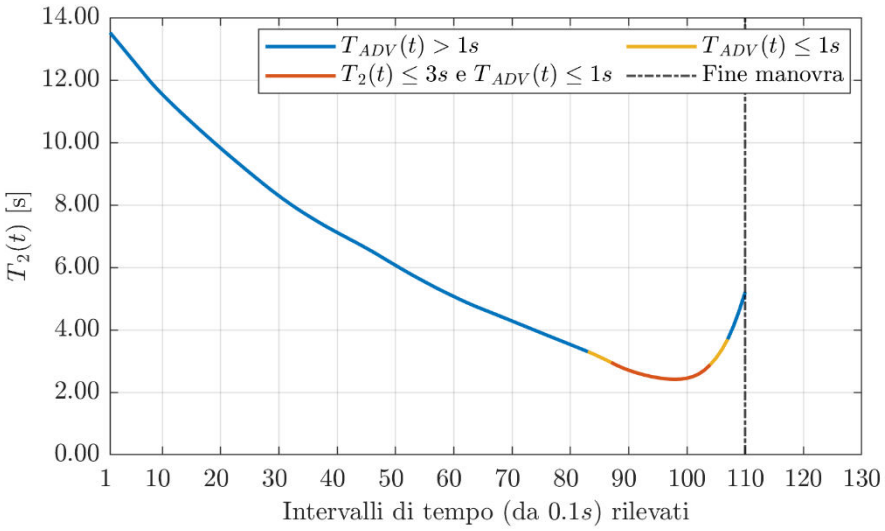


Figura 18: Esempio di andamento del $T_2(t)$ durante l'interazione tra veicolo e pedone, con segmentazione in base alle soglie statiche di pericolosità, per un partecipante della sperimentazione (il medesimo di Figura 17).

Una volta addestrati i due modelli (di classificazione e regressione) individualmente, le loro previsioni sono state utilizzate per classificare le interazioni conflittuali sugli orizzonti a 1, 2 e $3\,s$ nel futuro, distinguendo due categorie di severità, ovvero le classi "*safe*" e "*unsafe*". Queste categorie sono state definite in base ai valori di soglia statici riportati in letteratura (Sezione 2.2.2) e alle condizioni in cui è stato eseguito l'esperimento di simulazione (Sezione 3.2): quando la classe di T_{ADV} è "*collision course*" e il valore T_2 è inferiore a $3\,s$ sullo stesso orizzonte temporale (Figura 18), l'interazione è definita come "*unsafe*", in tutti gli altri casi è annotata come "*safe*". Ulteriori classi intermedie sono state evitate (ad esempio, il caso in cui T_{ADV} segnala "*collision course*" ma il valore di T_2 è maggiore di $3\,s$) poiché, siccome la predizione viene fatta su orizzonti temporali futuri, l'intenzione è quella di cogliere in anticipo le situazioni di conflitto effettivamente gravi; questo permetterebbe di migliorare il *timing* di attivazione dei sistemi esistenti, evitando inutili interferenze con il normale comportamento di guida.

In merito all'Approccio 2), una RNN del tipo GRU è stata implementata per eseguire una mappatura diretta tra i livelli di severità (ovvero le classi "*safe*" e "*unsafe*", sopra descritte) futuri degli incontri veicolo-pedone e le caratteristiche di input che determinano la rappresentazione della situazione di traffico.

Ovviamente, confrontando i risultati derivanti dall'accoppiamento delle previsioni dei due modelli con i target sperimentali, è stato possibile valutare l'efficacia di ciascun approccio nel classificare gli eventi di quasi incidente tra veicolo e pedone. Il confronto, a parità di metriche, tra i due approcci (presentato al Capitolo 5) ha permesso di stabilire quale tra questi garantisce le migliori capacità di generalizzazione. In particolare, vale la pena sottolineare che il modello di classificazione della severità ha come fine ultimo quello di imitare il

comportamento umano in situazioni potenzialmente pericolose e/o gravi che hanno evitato la collisione con il pedone, perché solo in questo modo il sistema teorizzato (Figura 6) potrà riconoscere i comportamenti anomali dei conducenti ed attuare una strategia evasiva. Per questa ragione e considerando i risultati dello studio di validazione [59] relativi al gruppo di controllo, la procedura e condizione di sperimentazione (Sezione 3.2) hanno permesso di cogliere proprio i comportamenti desiderati.

Per riassumere i risultati della sperimentazione e le elaborazioni compiute, in Tabella 5 sono elencate le statistiche descrittive delle variabili considerate in questo studio di ricerca.

Tabella 5: Statistiche descrittive delle variabili di input e output ai modelli GRU.

Variabile	**Descrizione**	**Intervallo valori - Conteggio elementi**	**Unità di misura**
$v_v(t)$	Velocità del veicolo	$(0.00, 15.62)$	$\mathrm{m/s}$
$v_p(t)$	Velocità del pedone	$(0.00, 1.06)$	$\mathrm{m/s}$
$r_0(t)$	Distanza relativa veicolo-pedone	$(5.06, 100.00)$	m
$\theta(t)$	Angolo azimutale orizzontale	$(-0.235, 0.507)$	rad
$A(t)$	Primitiva di comportamento	*Keep speed b.* (2): 3907 *Braking b.* (1): 4544 *Stopped b.* (0): 1756	-
$T_2(t)$	Tempo di arrivo del secondo utente alla zona di conflitto (valore numerico)	$(1.64, 17.69)$	s
$T_{ADV}(t)$	Tempo di "vantaggio" di un utente sull'altro per il passaggio sulla zona di conflitto (categorizzato attraverso il valore di soglia $T_{ADV} = 1\ s$)	*Crossing c.* (0): 8633 *Collision c.* (1): 1574	-
$S(t)$	Categorie di severità degli eventi (soglia critica per $T_{ADV} \leq 1\ s$ e $T_2 \leq 3\ s$)	*Safe* (0): 9121 *Unsafe* (1): 1574	-

In conclusione, è opportuno ricordare che la procedura riportata nell'Appendice A è stata applicata per il calcolo offline degli indicatori SSM e le dimensioni del veicolo simulato sono state considerate nei calcoli in modo che, tra tutti i punti di contatto tra fronte anteriore del veicolo e pedone in una potenziale collisione, venga selezionato quello che porta ai valori più bassi di T_2 e T_{ADV}. Quindi, i valori ottenuti sono stati ricampionati a 1, 2 e $3\ s$ per comporre la matrice di sequenze target (Sezione 3.3).

Capitolo 4

Metodologia

In questa sezione viene affrontato e discusso il metodo di modellazione utilizzato per l'implementazione del modello oggetto della presente tesi. In particolare, viene descritto il funzionamento algoritmico delle più generali reti neurali artificiali, le quali permettono di introdurre le più complesse reti neurali ricorrenti e loro varianti.

4.1 Reti Neurali Artificiali

Le Reti Neurali Artificiali (ANN, dall'inglese *Artificial Neural Networks*) sono modelli matematici che cercano di replicare lo schema di elaborazione delle informazioni che avviene nei sistemi neurali biologici [221]. Imitando la struttura del cervello, una ANN è composta da insiemi di neuroni artificiali (celle elementari di calcolo della rete) che sono collegati tra loro da connessioni pesate. Attraverso queste, i neuroni possono inviare segnali l'uno all'altro al fine di elaborare le informazioni e, infine, produrre un risultato. Dal punto di vista dei problemi di *pattern recognition* (ovvero riconoscimento di schemi funzionali nei dati), le ANN sono una classe generale di funzioni parametriche non lineari in cui i neuroni artificiali sono disposti in più strati sequenziali. Questi, chiamati "strati nascosti", sono organizzati in sequenza tra uno strato di input e uno di output. I neuroni sono implementati come modelli di regressione logistica con funzioni di attivazione non lineari. Ognuno di essi è collegato, attraverso connessioni pesate, ad ogni neurone presente nello strato precedente della rete. La struttura appena descritta identifica un tipo di rete neurale molto semplice e largamente diffusa chiamata *feedforward*, perché le connessioni tra le unità non formano cicli (a differenza delle RNN). Durante un processo di addestramento supervisionato, i pesi delle connessioni sono modificati per mezzo di una regola di apprendimento, al fine di replicare il target sperimentale associato alla rappresentazione di input. Aumentando il numero di strati nascosti e/o il numero di neuroni, è possibile controllare la complessità e la qualità delle relazioni tra le variabili di input e di output.

In termini formali, dato un set di dati D formato da coppie di vettori input-target $D = \{(\boldsymbol{x}^{(d)}, \boldsymbol{y}^{(d)})\}_{d=1}^{|D|}$, in cui $\boldsymbol{x}^{(d)} = [x_0^{(d)}, \dots, x_{F-1}^{(d)}] \in \mathbb{R}^F$ e $\boldsymbol{y}^{(d)} = [y_0^{(d)}, \dots, y_{G-1}^{(d)}] \in \mathbb{R}^G$, lo strato di input di una ANN si riferisce al vettore delle caratteristiche di input $\boldsymbol{x}^{(d)}$ associato alla d-esima osservazione. L'output del primo strato nascosto è il vettore $\boldsymbol{h}^{(1)} = [h_0^{(1)}, \dots, h_{N-1}^{(1)}] \in \mathbb{R}^N$ in cui N è il numero di neuroni che compongono lo strato nascosto e ciascun elemento $h_j^{(1)}$ è ottenuto come:

$$h_j^{(1)} = \phi\left(\sum_{i=0}^{F-1} w_{ij}^{(1)} \cdot x_i^{(d)} + b_j^{(1)}\right). \tag{4.1}$$

In questa espressione, $w_{ij}^{(1)}$ sono i pesi delle connessioni tra i neuroni $x_i^{(d)}$ e $h_j^{(1)}$, mentre $b_j^{(1)}$ sono gli errori aggiunti e ϕ è la funzione di attivazione. Lo strato

successivo, $\boldsymbol{h}^{(2)}$, è calcolato in modo simile, considerando gli elementi del vettore $\boldsymbol{h}^{(1)}$ come strato di ingresso, con i pesi $w_{ij}^{(2)}$ e gli errori $b_j^{(2)}$ corrispondenti, ossia:

$$h_j^{(2)} = \phi\left(\sum_{i=0}^{N-1} w_{ij}^{(2)} \cdot h_i^{(1)} + b_j^{(2)}\right). \tag{4.2}$$

Questo processo viene ripetuto per calcolare le attivazioni di ogni strato $\boldsymbol{h}^{(l)}$, $l \in \{1,2,\dots,L\}$ e l'output della rete $\hat{\boldsymbol{y}}^{(d)} = \phi^{(out)}(\boldsymbol{h}^{(L)})$, con $\phi^{(out)}$ la funzione di attivazione dello strato di output. I parametri della rete da regolare formano l'insieme $W = \left\{w_{ij}^{(l)}, b_j^{(l)} \middle| l \in \{1,2,\dots,L\}\right\}$. Scelte comuni per la funzione di attivazione includono l'unità lineare esponenziale (ELU), la tangente iperbolica ($tanh$), l'unità lineare rettificata (ReLU) e la logistica sigmoidale ($logsig$); una rappresentazione grafica di queste funzioni è fornita in Figura 19.

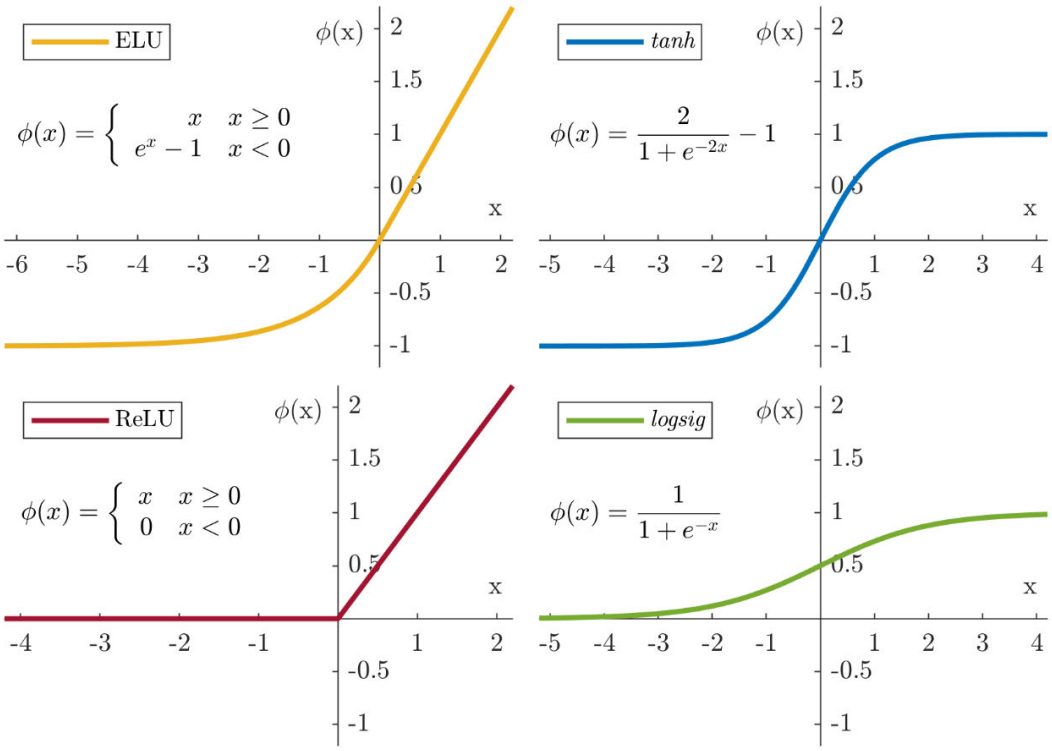


Figura 19: Esempi di comuni funzioni di attivazione, l'unità lineare esponenziale (ELU), la tangente iperbolica ($tanh$), l'unità lineare rettificata (ReLU) e la logistica sigmoidale ($logsig$).

I parametri W sono appresi attraverso un processo di addestramento (o *training*) supervisionato, che si divide nei seguenti passi: per primo, viene eseguito il *forward pass*, ovvero il vettore $\boldsymbol{x}^{(d)}$ è presentato come strato di ingresso della rete e le attivazioni $\boldsymbol{h}^{(l)}$ sono ottenute strato per strato finché l'output di rete viene calcolato $\hat{\boldsymbol{y}}^{(d)}$; a questo punto, si procede al *backward pass* in cui l'output di rete $\hat{\boldsymbol{y}}^{(d)}$ viene confrontato con il target atteso $\boldsymbol{y}^{(d)}$ per mezzo di una funzione obiettivo (o funzione di errore) $L(\hat{\boldsymbol{y}}^{(d)}, \boldsymbol{y}^{(d)})$. Quindi, l'algoritmo di *backpropagation* [222], tuttora uno dei metodi di addestramento più diffusi, calcola il gradiente della funzione obiettivo $L(\cdot)$ rispetto ai parametri W; il termine "*backpropagation*" è

legato essenzialmente alla tecnica utilizzata per il calcolo delle derivate della funzione di errore, basata sulle regole di derivazione delle funzioni composte. Queste derivate sono sfruttate da una regola di apprendimento per rinnovare i pesi e gli errori della rete, ovvero:

$$W^{(e)} = W^{(e-1)} - \alpha \nabla L\left(W^{(e-1)}\right), e \in \{0, \dots, E-1\} \tag{4.3}$$

in cui e è il numero dell'iterazione corrente, W^{e-1} sono i valori dei parametri all'iterazione $e-1$, mentre $\nabla L\left(W^{(e-1)}\right)$ è l'approssimazione al primo ordine delle derivate parziali di $L(\cdot)$ rispetto ai parametri $W^{(e-1)}$. Il processo viene ripetuto per un numero fisso di iterazioni E o fino a quando il valore di perdita è ridotto al minimo. $\alpha > 0$ è il tasso di apprendimento (o *learning rate*) ed è usato per controllare la dimensione del passo verso tale punto di minimo (ovvero lungo l'antigradiente).

Quelli appena elencati sono alcuni degli iperparametri (o *hyperparameters*) del modello neurale che definiscono il funzionamento dell'algoritmo di apprendimento e devono essere attentamente definiti dal ricercatore; quest'ultimo processo è spesso considerato un'arte in quanto richiede tempo, conoscenze specifiche e a volte persino una ricerca esaustiva. Le metodologie standard includono la ricerca casuale o a griglia, le quali richiedono comunque una definizione precisa degli intervalli degli iperparametri e delle strategie di campionamento.

Ad ogni modo, dopo che l'ANN è stata addestrata, i pesi $w_{ij}^{(l)}$ e gli errori $b_j^{(l)}$ ottenuti sono mantenuti fissi mentre nuovi dati per la validazione del modello sono processati solamente nel modo *forward*.

Il descritto processo di addestramento, con la regola di aggiornamento introdotta nell'Eq. 4.3, è noto come discesa stocastica del gradiente (SGD, dall'inglese *Stochastic Gradient Descent*). Ad ogni iterazione, l'algoritmo SGD valuta il gradiente e aggiorna i parametri usando solo un sottoinsieme dei dati di *training*, chiamato *mini-batch*. Un ciclo completo dell'algoritmo di addestramento sull'intero dataset di addestramento (ovvero su tutti i *mini-batch*) è chiamato "epoca". Quindi, l'algoritmo SGD è stocastico perché gli aggiornamenti dei parametri W calcolati usando un *mini-batch* rappresentano una stima di quelli che risulterebbe dall'uso dell'intero set di dati. Tuttavia, l'algoritmo SGD può oscillare nel proprio percorso verso il minimo lungo la superficie di errore, aumentando di conseguenza i tempi di convergenza. Questa situazione viene aggravata dall'impiego comune di un *learning rate* fisso, il quale se non impostato correttamente può ridurre ulteriormente le velocità di convergenza e addirittura condurre a minimi locali. Quindi, la scelta dell'algoritmo di ottimizzazione (e dei parametri ad esso associati) per un modello di *Deep Learning* può fare la differenza tra buoni risultati in minuti, ore e giorni.

Al fine di migliorare la velocità di convergenza e la qualità delle ANN, negli ultimi anni, sono stati proposti diversi miglioramenti a questo algoritmo. In primo luogo, aggiungere un termine (chiamato *momentum*) all'aggiornamento dei parametri, il quale determini il contributo dei passi precedenti sull'iterazione corrente, è un modo per ridurre le oscillazioni dell'algoritmo di *training*. Altri algoritmi di ottimizzazione cercano di migliorare l'addestramento della rete

utilizzando tassi di apprendimento che differiscono per parametro e possono adattarsi automaticamente alla funzione di perdita da ottimizzare.

L'algoritmo Adam (il nome è stato ricavato da *Adaptive Moment Estimation*) [223] implementa entrambi questi miglioramenti. Infatti, questo algoritmo calcola le medie mobili sia del gradiente (m, o momento primo del gradiente) sia del quadrato del gradiente (v, o momento secondo del gradiente) della funzione $L(\cdot)$ sul *mini-batch*:

$$m^{(e-1)} = \beta_1 m^{(e-2)} + (1-\beta_1)\nabla L\left(W^{(e-1)}\right), e \in \{0, \dots, E-1\}, \tag{4.4}$$

$$v^{(e-1)} = \beta_2 v^{(e-2)} + (1-\beta_2)\left[\nabla L\left(W^{(e-1)}\right)\right]^2, e \in \{0, \dots, E-1\}, \tag{4.5}$$

e li usa per aggiornare i parametri della rete nel seguente modo:

$$W^{(e)} = W^{(e-1)} - \frac{\alpha m^{(e-1)}}{\sqrt{v^{(e-1)}} + \epsilon}, e \in \{0, \dots, E-1\}. \tag{4.6}$$

in cui la divisione viene eseguita elemento per elemento.

Nelle espressioni precedenti $\beta_1, \beta_2 \in [0,1)$ sono i nuovi iperparametri introdotti dall'algoritmo e controllano i tassi di decadimento esponenziale delle medie mobili; invece, ϵ è una piccola costante aggiunta per evitare la divisione per zero.

In pratica, senza aggiungere ulteriore complessità alla trattazione, l'algoritmo Adam è un'estensione o miglioramento della SGD perché calcola un *learning rate* adattivo individuale per ogni parametro di una ANN a partire dalle stime dei momenti primo e secondo del gradiente di $L(\cdot)$; come tale, l'algoritmo in questione ha recentemente visto un'ampia adozione e una crescente popolarità nelle applicazioni di *Deep Learning* perché ha permesso di raggiungere buoni risultati velocemente. Infatti, nell'introdurre l'algoritmo, gli autori [223] elencano tra i vantaggi riscontrati la semplicità di implementazione, l'efficienza computazionale, i ridotti requisiti di memoria, l'interpretazione intuitiva degli iperparametri i quali non richiedono un impostazione troppo accurata. Inoltre, risultati empirici hanno dimostrato che Adam non solo funziona bene nella pratica (anche su grandi data set) ma si confronta positivamente con altri metodi di ottimizzazione stocastica: Sebastian Ruder [224] ha sviluppato una revisione completa dei moderni algoritmi di ottimizzazione a discesa del gradiente e nella sezione intitolata "*Which optimizer to use*?" ha raccomandato di utilizzare Adam perché più veloce ed efficiente nella maggioranza delle applicazioni. Inoltre, nell'articolo originale, l'algoritmo Adam era stato applicato su diversi data set e per molteplici compiti dimostrando che la convergenza soddisfa le aspettative dell'analisi teorica e che esso possa risolvere in modo efficiente i problemi pratici di *Deep Learning*.

In questo lavoro di tesi, tale algoritmo di ottimizzazione è stato adottato per calcolare i pesi ottimali del modello neurale implementato.

4.2 Funzioni di errore e regolarizzazione delle ANN

Siccome le reti *feedforward* possono essere utilizzate per compiti di apprendimento supervisionato sia di regressione che di classificazione, le funzioni di errore possono avere formulazioni diverse a seconda del problema da

affrontare. Inoltre, quando si lavora con RNN, bisogna considerare che l'input e/o il target sono sequenze.

Riprendendo la notazione della Sezione 4.1 e usando una terminologia temporale, una sequenza di input $(\boldsymbol{x}^{(d,1)}, \boldsymbol{x}^{(d,2)}, \dots, \boldsymbol{x}^{(d,T)})$ può essere vista come un insieme di dati $\boldsymbol{x}^{(d,t)} = [x_0^{(d,t)}, \dots, x_{F-1}^{(d,t)}] \in \mathbb{R}^F$, ovvero vettori con F valori reali (uno per ogni variabile di input) associati alla d-esima osservazione del data set ed al passo temporale discreto t di campionamento della sequenza. L'indice temporale massimo della sequenza, ovvero la sua lunghezza, è indicato con la lettera T. Allo stesso modo, una sequenza di target è rappresentata dalla notazione $\boldsymbol{y}^{(d,t)} = [y_0^{(d,t)}, \dots, y_{G-1}^{(d,t)}] \in \mathbb{R}^G$. Invece, gli output sequenziali prodotti da un modello sono indentificati con $\hat{\boldsymbol{y}}^{(d,t)}$.

Sotto questa premessa, la funzione di errore per un problema di regressione di sequenze è generalmente l'errore quadratico medio delle risposte previste per ogni passo temporale, ovvero:

$$L(\hat{\boldsymbol{y}}^{(d,t)}, \boldsymbol{y}^{(d,t)}) = \frac{1}{D}\sum_{d=1}^{D}\left(\frac{1}{2T}\left\|\hat{\boldsymbol{y}}^{(d,t)} - \boldsymbol{y}^{(d,t)}\right\|_2^2\right) = \frac{1}{D}\sum_{d=1}^{D}\left(\frac{1}{2T}\sum_{t=1}^{T}\sum_{g=1}^{G}\left(\hat{y}_g^{(d,t)} - y_g^{(d,t)}\right)^2\right), \tag{4.7}$$

in cui D è il numero di osservazioni nel data set di *training* (o nel *mini-batch*), G è il numero di variabili target predette, mentre l'operatore $\|\cdot\|_2^2$ rappresenta il quadrato della norma 2 (o norma euclidea) che nello spazio euclideo G-dimensionale $\mathbb{R}^G$ assume l'espressione riportata nell'ultimo termine dell'Eq.4.7.

Diversamente, in una tipica rete di classificazione, l'ultimo stato nascosto (detto *fully connected layer*, poiché combina tutte le informazioni apprese dagli strati precedenti per la loro annotazione finale e, quindi, la dimensione del suo output è uguale al numero di classi da predire) è seguito da uno strato *softmax* e uno strato di classificazione. Lo strato *softmax* può essere considerata una funzione di attivazione in quanto applica la funzione esponenziale normalizzata (detta anche *softmax*) alle informazioni ricevute dal *fully connected layer*; questa funzione è una generalizzazione multi-classe della funzione logistica sigmoidale [225] che converte il vettore di output del *fully connected layer* di valori reali arbitrari in un vettore di pari dimensione ma con valori compresi nell'intervallo $[0,1]$ e la cui somma è 1. In altre parole, la funzione *softmax* converte valori reali in probabilità di appartenenza ad una classe [225]. Quindi, queste probabilità sono passate allo strato di classificazione che associa ciascun input ad una delle M classi mutuamente esclusive attraverso uno schema di codifica 1-*of*-M (ovvero la risposta della rete è un vettore binario con elementi uguali a zero tranne l'm-esimo, che invece è uguale a 1, considerando in questo caso l'ipotesi che il vettore target appartenga alla classe m-esima) e calcola il valore della funzione di errore per problemi di classificazione multi classe, detta *cross-entropy*:

$$L(\hat{\boldsymbol{Y}}^{(d,t)}, \boldsymbol{Y}^{(d,t)}) = -\frac{1}{D}\sum_{d=1}^{D}\sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{g=1}^{G}\left(Y_{g,m}^{(d,t)}\log\left(\hat{Y}_{g,m}^{(d,t)}\right) + \left(1 - Y_{g,m}^{(d,t)}\right)\log\left(1 - \hat{Y}_{g,m}^{(d,t)}\right)\right), \tag{4.8}$$

in cui D e T sono il numero e la lunghezza delle sequenze nel data set di *training* (o nel *mini-batch*), mentre M è il numero delle classi indipendenti e G il numero di variabili target predette. Nell'Eq. 4.8 gli output del modello e i target

sperimentali sono rappresentati rispettivamente da $Y_{g,m}^{(d,t)}$ e $\hat{Y}_{g,m}^{(d,t)}$, al fine di distinguere queste grandezze da quelle relative ai problemi di regressione.

L'*overfitting* è la situazione in cui un modello di *Machine Learning* si adatta in maniera eccessiva ai dati di addestramento, con conseguenti scarse prestazioni sui dati di validazione o, in generale, su nuovi dati. Per superare questo problema, sono state introdotte diverse soluzioni note come tecniche di regolarizzazione. Tra queste, la tecnica del *weight decay* (o regolarizzazione L2) [226] è la più diffusa ed è stata implementata anche nello studio di questa tesi. Questa tecnica consiste nell'aggiungere un termine addizionale alla funzione obiettivo da ottimizzare che è calcolato come il quadrato della norma euclidea dei pesi della rete; infatti, i pesi delle connessioni determinano la complessità della soluzione offerta dalla rete e, per evitare l'*overfitting*, è necessario contenere (o meglio minimizzare) i valori di questi parametri. In questo modo, l'obbiettivo di ottimizzazione globale diventa:

$$L_{opt}\left(\hat{\boldsymbol{y}}^{(d,t)}, \boldsymbol{y}^{(d,t)}, W^{(e-1)}\right) = L\left(\hat{\boldsymbol{y}}^{(d,t)}, \boldsymbol{y}^{(d,t)}\right) + \frac{\lambda}{2}\left\|W^{(e-1)}\right\|_2^2, \tag{4.9}$$

in cui il coefficiente λ è un iperparametro, chiamato fattore di regolarizzazione, che controlla l'ordine di grandezza del termine aggiunto e, di conseguenza, la soluzione finale.

4.3 k-fold Cross Validation

Comunemente, il data set disponibile per la modellazione neurale viene diviso in due sottoinsiemi casuali che sono impiegati separatamente per l'addestramento e la validazione della ANN; questa pratica è nota come metodo *hold-out*. Con il termine "validazione" si intende il processo di controllo della validità e della correttezza (ovvero dei requisiti pratici per l'utilizzo) del modello neurale addestrato, attraverso la sua applicazione su nuovi dati (non considerati durante il processo di *training*, ma che ovviamente devono rispettare la struttura del vettore di input al modello); in questo contesto, si parla anche di valutazione delle capacità di generalizzazione del modello sulla base delle informazioni acquisite durante il *training*.

Tuttavia, il metodo *hold-out* può portare a valutazioni errate delle capacità di generalizzazione di una ANN a causa della diversa distribuzione dei dati all'interno dei due sottoinsiemi (l'errore di validazione può essere molto diverso a seconda di quali osservazioni sono incluse in ciascun sottoinsieme) e per il rischio di escludere alcuni schemi o andamenti rilevanti dai dati di addestramento (che, di conseguenza, sarebbero evidenti solo nel data set di validazione). Questi effetti sono particolarmente evidenti quando il data set disponibile è relativamente piccolo [227].

Per ottenere una valutazione adeguata delle prestazioni di un modello, la soluzione più efficace è offerta dalla *k-fold Cross Validation*. Questo metodo suggerisce di dividere il data set di interesse in k partizioni di uguali dimensioni, chiamati "*folds*". Quindi, vengono eseguiti k esperimenti in cui ciascun *fold* di dati è assunto a turno come sottoinsieme di validazione, mentre i restanti $k-1$ *folds* sono uniti per formare il sottoinsieme di addestramento. Una volta completata l'intera procedura, si ottengono k punteggi di validazione (attraverso metriche di

performance predefinite) e il loro valore medio viene assunto come rappresentativo delle effettive capacità di generalizzazione del modello neurale.

Per questa ragione, nelle impostazioni dello studio presentato in questa tesi, è stato implementato il metodo di validazione *k-fold* con $k = 5$.

4.4 Gated Recurrent Unit (GRU)

Le reti neurali ricorrenti (RNN) sono un'estensione delle reti neurali *feedforward* convenzionali (Sezione 4.1), capaci di gestire una sequenza di input di lunghezza variabile (e, opzionalmente, può avere anche una sequenza di output di lunghezza variabile). La RNN gestisce la sequenza grazie ad uno stato nascosto ricorrente la cui attivazione ad ogni passo temporale dipende da quella del passo precedente. Formalmente, data una sequenza $\boldsymbol{x} = (\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \dots, \boldsymbol{x}^{(T)})$, la RNN aggiorna il suo stato nascosto ricorrente $\boldsymbol{h}^{(t)}$ in questo modo:

$$\boldsymbol{h}^{(t)} = \begin{cases} 0, & t = 0 \\ \phi\left(W\boldsymbol{x}^{(t)} + U\boldsymbol{h}^{(t-1)}\right), & \text{altrimenti} \end{cases} \tag{4.10}$$

in cui ϕ è una funzione di attivazione non lineare (tipicamente la sigmoidea o la tangente iperbolica), $\boldsymbol{h}^{(t-1)}$ è il vettore di output del precedente stato nascosto mentre W e U sono i pesi associati rispettivamente a $\boldsymbol{x}^{(t)}$ e $\boldsymbol{h}^{(t-1)}$. Quindi, l'output è prodotto unendo l'input all'istante corrente t con il vettore di output dello stato nascosto relativo al passo temporale precedente $t - 1$. Questa è la ragione per cui una RNN può memorizzare le informazioni sequenziali che si trovano nei dati di una serie temporale. Le RNN che utilizzano l'Eq. 4.10 sono anche chiamate *Vanilla* RNN.

Sfortunatamente, è stato osservato, per esempio, da Bengio e coll. [155] che è difficile addestrare una *Vanilla* RNN per catturare le dipendenze a lungo termine a causa del problema di *vanishing* (cioè "scomparsa") del gradiente. Una delle principali cause è l'uso delle classiche funzioni di attivazione non lineari, come la tangente iperbolica o la funzione logistica sigmoidale, che hanno gradiente a valori nell'intervallo $[0,1]$. Questo rende il metodo di ottimizzazione basato sul gradiente problematico, non solo a causa delle variazioni nell'ordine di grandezza del gradiente, ma perché l'effetto delle dipendenze a lungo termine è nascosto (essendo esponenzialmente più piccolo rispetto alla lunghezza della sequenza) dall'effetto delle dipendenze a breve termine. Diversi approcci sono stati sviluppati da ricercatori in tutto il Mondo per ridurre l'impatto negativo del problema di *vanishing*. Tra queste, l'approccio che ha riscontrato maggiore successo è stato quello di progettare sofisticate funzioni di attivazione ed il primo tentativo in questa direzione ha portato ad una unità ricorrente chiamata *Long Short-Term Memory* (LSTM) [98]. Più recentemente, un altro tipo di unità ricorrente, nota come *Gated Recurrent Unit* (GRU), è stato proposto da Cho e coll. [156]. Le RNN che impiegano una di queste unità hanno ottenuto buoni risultati in compiti molto diversi [228]–[230], accomunati dalla necessità di acquisire dipendenze a lungo termine nei dati.

Le RNN sono state impiegate anche per l'implementazione del sistema di apprendimento presentato in questa tesi. In particolare, l'architettura GRU è stata selezionata per la sua maggiore robustezza rispetto alle *Vanilla* RNN e per la sua potenza espressiva che eguaglia la più sofisticata LSTM, ma che si ottiene con

un minor sforzo computazionale; infatti, Chung e coll. [99] hanno dimostrato empiricamente che, utilizzando un numero fisso di parametri per tutti i modelli su diversi data set, la GRU può raggiungere prestazioni migliori dell'unità LSTM sia in termini di tempo convergenza che di aggiornamento dei parametri e di capacità di generalizzazione.

Il meccanismo di funzionamento delle reti GRU viene di seguito descritto in forma generale [97], [99]. Sulla base del vettore delle caratteristiche di input e dell'output dello stato precedente, il neurone GRU esegue diverse operazioni usando operatori chiamati "*gate*". Un *update gate* stabilisce la parte dell'informazione passata da conservare e trasmettere allo stato futuro, mentre un *forget gate* si occupa della parte da dimenticare. In dettaglio, si consideri $\boldsymbol{x}^{(t)}$ il vettore delle caratteristiche disponibile al passo t della serie temporale, $\boldsymbol{h}^{(t)}$ lo stato nascosto e $\boldsymbol{y}^{(t)}$ il vettore di output. Questi vettori vengono utilizzati dalla GRU nelle seguenti operazioni (le formulazioni qui riportate tengono conto che una RNN è costituita da più unità GRU):

$$\boldsymbol{z}^{(t)} = logsig\left(W_z \boldsymbol{x}^{(t)} + U_z \boldsymbol{h}^{(t-1)}\right), \tag{4.11}$$

$$\boldsymbol{f}^{(t)} = logsig\left(W_f \boldsymbol{x}^{(t)} + U_f \boldsymbol{h}^{(t-1)}\right), \tag{4.12}$$

$$\widehat{\boldsymbol{h}}^{(t)} = tanh\left(W_c \boldsymbol{x}^{(t)} + U_c\left(\boldsymbol{f}^{(t)} \odot \boldsymbol{h}^{(t-1)}\right)\right), \tag{4.13}$$

$$\boldsymbol{h}^{(t)} = \left(1 - \boldsymbol{f}^{(t)}\right) \odot \boldsymbol{h}^{(t-1)} + \boldsymbol{z}^{(t)} \odot \widehat{\boldsymbol{h}}^{(t)}, \tag{4.14}$$

le quali sono applicate per tutti gli stati temporali t che si susseguono durante la serie, cioè da $t = 0$ a $t = T$ con T la lunghezza della sequenza (per $t = 0$, si ricorda che $\boldsymbol{h}^{(0)}$ è un vettore di soli zeri). L'*update gate* e il *forget gate* sono implementati dalle Eq. 4.11 e 4.12, rispettivamente. W_z, U_z, W_f, U_f, W_c e U_c sono le matrici dei pesi ottimizzate durante la fase di addestramento della GRU, logsig e tanh sono le funzioni di attivazione presentate alla Sezione 4.1 mentre il simbolo $\odot$ rappresenta il prodotto di Hadamard tra due vettori (si tratta di un'operazione binaria che, applicata a due vettori, produce un altro vettore delle stesse dimensioni degli operandi in cui ogni elemento i è il prodotto degli elementi i dei due vettori originali). Per una più facile comprensione del processo, il flusso delle operazioni compiute dall'unità GRU è rappresentato in Figura 20.

In sintesi, quando $\boldsymbol{x}^{(t)}$ viene immesso nell'unità di rete, questo vettore viene moltiplicato per la matrice dei pesi W_z; allo stesso tempo, $\boldsymbol{h}^{(t-1)}$, prodotto dall'ultimo stato nascosto, viene moltiplicato per il corrispondente peso U_z. Quindi, i due risultati sono sommati e la logistica sigmoidale $logsig$ viene impiegata come funzione di attivazione per generare un valore probabilistico tra 0 e 1. Considerando i contributi di ciascuna unità GRU nella RNN, vengono generati le risposte dell'*update gate* $\boldsymbol{z}^{(t)}$ e del *forget gate* $\boldsymbol{f}^{(t)}$. Il vettore $\widehat{\boldsymbol{h}}^{(t)}$ è chiamato *memory unit* perché memorizza le informazioni rilevanti usando l'output del *forget gate* $\boldsymbol{f}^{(t)}$. L'output dello stato nascosto $\boldsymbol{h}^{(t)}$ al passo temporale t è calcolato dall'Eq. 4.14. Queste equazioni sono processate iterativamente dal primo all'ultimo passo temporale della serie per generare in conclusione l'output della rete GRU, $\boldsymbol{y}^{(t)}$.

Queste particolari operazioni permettono ad una rete di produrre gradienti più significativi rispetto alle *Vanilla* RNN durante la fase di *training* e, in ultima analisi, rendono le unità GRU capaci di acquisire migliori relazioni a lungo termine tra le caratteristiche. Come anticipato, per incrementare ulteriormente la qualità di tali relazioni tra le caratteristiche di input, il modello di apprendimento può essere organizzato come strati sovrapposti di unità GRU.

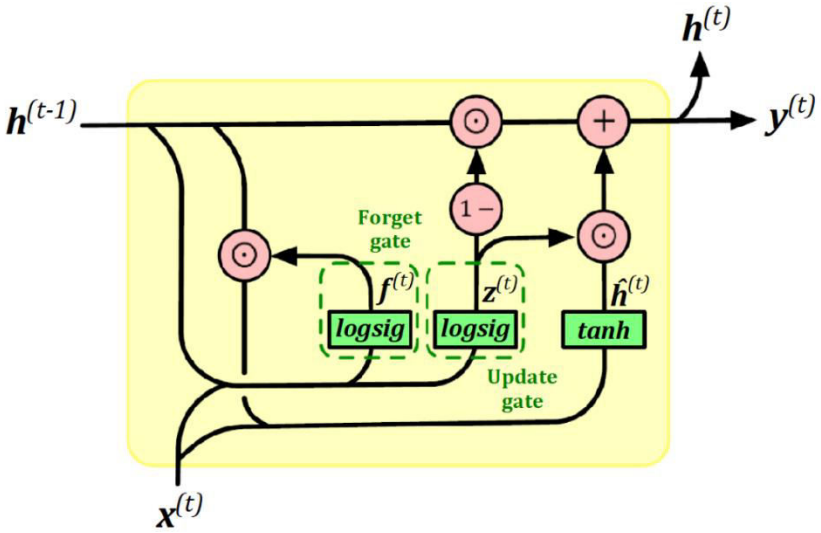


Figura 20: Rappresentazione schematica delle operazioni eseguite da una unità GRU.

Nel contesto dello studio, una rete GRU composta da due strati nascosti ciascuno con N unità nervose è stata utilizzata per l'implementazione del sistema. Il primo strato GRU riceve in ingresso il vettore delle caratteristiche $\boldsymbol{x}^{(t)} = [A(t), T_2(t), v_v(t), v_p(t), r_0(t), \theta(t)]$, il quale rappresenta l'informazione sensoriale (descritta nella Sezione 3.4) disponibile al passo temporale t della sequenza di interazione tra conducente e pedone. Nel medesimo passo temporale, il secondo strato GRU riceve la rappresentazione prodotta dal primo strato. Quindi, un *fully connected layer*, con tanti neuroni quanti sono gli output predetti della sequenza futura, è applicato per produrre la risposta della rete. Nel complesso, il ruolo degli strati GRU è quello di astrarre una rappresentazione significativa che riassuma il processo di interazione tra veicolo e pedone fino al passo temporale t. A sua volta, questa caratteristica di livello superiore è sfruttata dallo strato di output per produrre la previsione degli stati futuri.

4.5 Dettagli di implementazione

In questa Sezione, vengono forniti i dettagli della procedura implementata per addestrare il modello GRU proposto.

Una strategia di ricerca a griglia è stata impiegata per determinare il valore dei più importanti iperparametri strutturali e di addestramento della RNN come il numero di neuroni N della GRU, il valore iniziale del tasso di apprendimento globale α e il fattore di riduzione del tasso di apprendimento δ [231]. Quest'ultimo iperparametro è stato aggiunto per consentire un miglior adattamento dell'algoritmo di *training* alla superficie di errore; in particolare, il valore di α è stato ridotto ogni 50 epoche della quantità δ, un fattore moltiplicativo con valori

tra $(0, 1]$. Nella ricerca a griglia, il dominio degli iperparametri viene diviso in una griglia discreta e ogni combinazione di valori in questa griglia viene testata calcolando le prestazioni del modello che ne consegue attraverso la procedura di *k-fold Cross Validation*. Il punto della griglia che conduce al miglior risultato medio sulle differenti distribuzioni del data set impiegato viene assunto come migliore configurazione del modello neurale (ovvero combinazione ottima di valori per gli iperparametri). La ricerca a griglia è un algoritmo esaustivo che abbraccia tutte le combinazioni di interesse; quindi, può effettivamente trovare il punto migliore nel dominio. Il grande svantaggio è un maggior costo computazionale rispetto ad altri approcci (per esempio, la ricerca casuale, invece di usare tutti i punti della griglia, verifica solo un sottoinsieme di questi punti selezionato casualmente), perché controllare ogni combinazione richiede molto tempo (ogni iterazione della ricerca a griglia richiede k processi di *training* e validazione). Tuttavia, per ridurre i tempi del processo, è stato deciso di fissare alcuni valori degli iperparametri sulla base dei suggerimenti offerti da Kingma e Ba [223] sull'algoritmo di ottimizzazione Adam, e concentrare la ricerca a griglia solo su N, α e δ: i valori di prova considerati per N sono stati $64, 128, 256$ unità; α è stato testato per i valori $0.05, 0.01$ e 0.005, mentre δ per i fattori $0.40, 0.60, 0.80$. Le quantità assunte sono in linea con altri studi di letterature che hanno impiegato architetture GRU, come per esempio in [85]. Per tutti gli esperimenti, il processo di apprendimento è stato condotto per $E = 1000$ epoche attraverso l'ottimizzatore Adam, caratterizzato da $\beta_1 = 0.99$ e $\beta_2 = 0.90$ [223]. La dimensione del *mini-batch* è stata assunta pari al numero di sequenze presenti nel data set di *training* definito dalla *k-fold Cross Validation* con $k = 5$, e la validazione è stata eseguita per ogni *mini-batch* sul corrispondente *fold* di verifica. Il valore fissato di k si associa ad un rapporto di suddivisione $4{:}\,1$ dei partecipanti alla sperimentazione, corrispondente a 52 manovre per l'addestramento e 13 per la validazione. Quindi, in un'epoca dell'algoritmo di addestramento viene compiuto l'intero processo di *Cross Validation*. Il processo di *training* è stato arrestato anticipatamente nel caso in cui i valori della funzione di errore sui sottoinsiemi di validazione risultassero superiori od uguali al minimo registrato nel processo per 100 epoche consecutive; questo metodo è chiamato *early-stopping* e rappresenta una precauzione contro il problema di *overfitting*, in quanto evita l'eccessivo discostamento tra i valori della funzione di errore sui sottoinsiemi di *training* e validazione. Sempre in questo contesto, la tecnica del *weight decay* con fattore di regolarizzazione $\lambda = 0.005$ è stata implementata per evitare l'eccessivo adattamento del modello ai dati di addestramento.

Lo scopo dello studio presentato in questa tesi è lo sviluppo di un modello GRU che possa prevedere, fino a $3\ s$ nel futuro, il livello di rischio degli incontri veicolo-pedone nel traffico urbano. A questo proposito, sono stati identificati due approcci equivalenti [152], cioè 1) utilizzare più reti GRU in parallelo per modellare separatamente gli SSM (T_2 e T_{ADV}) che permettono di stimare la gravità dell'interazione, o 2) addestrare una singola RNN come classificatore di sequenze per annotare direttamente gli eventi di quasi incidente, sulla base delle caratteristiche di mobilità rilevanti degli incontri. Vale la pena sottolineare che, anche se i due approcci considerati sono equivalenti dal punto di vista del risultato finale (cioè la previsione dei livelli di gravità futuri), l'Approccio 1) consentirebbe una maggiore flessibilità nell'annotazione degli eventi quasi-

incidentali, poiché valori di soglia degli SSM diversi da quelli considerati in questo studio potrebbero essere selezionati per classificare il livello di rischio delle interazioni veicolo-pedone (ad esempio, utilizzando valori di soglia specifici per i contesti geografici di funzionamento del sistema, derivati da studi sui comportamenti di guida locali e/o nazionali) [85]. Ad ogni modo, per la modellazione dell'indicatore T_2, siccome questo rappresenta un problema di regressione, la rete GRU corrispondente è stata addestrata per minimizzare la funzione di errore di Eq. 4.7. Diversamente, per i compiti di classificazione del T_{ADV} e di annotazione degli eventi di quasi incidente secondo l'Approccio 2), il modello GRU è stato ottimizzato attraverso la minimizzazione della *cross-entropy* presentata in Eq. 4.8. Ovviamente, si è stato tenuto conto del termine di regolarizzazione secondo l'Eq. 4.9.

Il data set acquisito attraverso la sperimentazione al simulatore conta $10^{\cdot}207$ elementi, con un rapporto di circa $8.4:1$ tra le classi di severità "*safe*" e "*unsafe*" e di circa $5.5:1$ tra le classi del T_{ADV} "*collision course*" e "*crossing course*" (Tabella 5). Tuttavia, questo sbilanciamento tra le classi può portare ad un problema chiamato *accuracy paradox*, nel senso che un modello addestrato su un data set sbilanciato tende a riflettere la distribuzione delle classi predicendo solamente la classe più numerosa, indipendentemente dalle informazioni fornite in input, per garantire una buona accuratezza. Le classificazioni sbilanciate rappresentano una sfida per la modellazione predittiva, poiché la maggior parte degli algoritmi di apprendimento automatico sono stati progettati con il presupposto di un numero uguale di elementi per ciascuna classe. Questo si traduce in modelli che hanno scarse prestazioni predittive, in particolare per la classe di minoranza, che tipicamente è la più importante. Per risolvere questo possibile problema, è stato applicato un metodo di sovra-campionamento, noto come *Synthetic Minority Over-Sampling Technique* (SMOTE) [232], capace di aumentare il numero di elementi nel data set di *training* per la classe in minoranza, cioè processo "*unsafe*". SMOTE è un metodo popolare per l'aumento dei dati, il quale genera nuovi campioni interpolando le caratteristiche tra diversi esempi di classe minoritaria, presenti nel medesimo sottoinsieme di addestramento. In questo modo, sono prodotti nuovi elementi sintetici che conservano la semantica dei dati originali. Per questo lavoro, SMOTE ha permesso di raggiungere un rapporto $1:1$ tra le classi in ciascun sottoinsieme di addestramento definito attraverso 5-*fold Cross Validation*; diversamente, i sottoinsiemi di validazione includono solo elementi originali.

Ogni caratteristica del vettore di input $\boldsymbol{x}^{(t)} = [A(t), T_2(t), v_v(t), v_p(t), r_0(t), \theta(t)]$ al modello GRU è stata mappata nell'intervallo $[-1,1]$, i cui estremi corrispondo ai valori massimi di ciascuna variabile sull'intero data set D. La stessa operazione è stata applicata alla variabile target T_2, mentre T_{ADV} e le classi di severità a $1, 2$ e $3\ s$ sono state rappresentate da matrici a valori binari, in cui 1 corrisponde rispettivamente alle classi "*collision course*" e "*unsafe*" mentre 0 alle classi "*crossing course*" e "*safe*" (Tabella 5). L'operazione di mappatura qui presentata è una pratica comunemente adottata per aumentare i tempi di convergenza ed in generale migliorare il processo di *training*.

Tutto il codice sorgente necessario per questo lavoro è stato scritto con linguaggio MATLAB®. Le RNN sono state implementate utilizzando il framework della The MathWorks Inc., chiamato *Deep Learning Toolbox™* [233], mentre la

procedura SMOTE è stata realizzata attraverso l'*Oversampling-Imbalanced-Data package*, implementato e reso fruibile gratuitamente da Michio Inoue [234]. Il codice è stato eseguito su un personal computer dotato di CPU Intel® Core™ $i7-6700HQ\ 2.60GHz$ e $16GB$ di RAM con sistema operativo Windows 10 Home. Ogni esperimento è durato circa 12 ore.

Nel seguente Capitolo vengono presentati i risultati della ricerca a griglia e il confronto tra i modelli ottimali per gli Approcci 1) e 2) descritti in questa Sezione.

Capitolo 5

Presentazione e Discussione dei Risultati

Nel seguito, gli acronimi $\mathrm{GRU_{T2}}$ e $\mathrm{GRU_{TADV}}$ sono utilizzati con riferimento ai modelli GRU che predicono i valori continui del T_2 e le classi di T_{ADV}, rispettivamente. Diversamente, per eseguire i confronti tra i due approcci implementati, il modello di classificazione dei livelli di severità risultante dall'Approccio 1) è presentato con l'acronimo $\mathrm{M\text{-}GRU_{SL}}$, mentre quello relativo all'Approccio 2) come $\mathrm{S\text{-}GRU_{SL}}$.

In questa sezione, le capacità di generalizzazione dei modelli addestrati sono stimate e analizzate, per ogni orizzonte temporale di previsione (cioè, 1, 2 e $3\ s$ nel futuro), sulla base di metriche di valutazione comunemente utilizzate nel contesto scientifico di riferimento e mediando i loro punteggi sui cinque *folds* di validazione. Queste metriche, diverse per i relativi problemi di classificazione (ovvero, *Accuracy*, *Precision*, *Recall*, *Specificity*, *False Alarm Rate* o FAR, e *Area Under the Curve* o AUC) e regressione (cioè, *Mean Absolute Error* o MAE, *Mean Absolute Percentage Error* o $MAPE$, *Root Mean Squared Error* o $RMSE$), sono presentate brevemente durante la discussione dei risultati, per rendere la valutazione dei modelli più chiara; tuttavia, descrizioni più approfondite e di dettaglio possono essere trovate in [85], [152], [235]. In Appendice B, sono riportate le formulazioni analitiche di queste metriche, riferite alle osservazioni del profilo del singolo partecipante.

In particolate, Altché e de La Fortelle [137] hanno evidenziato che spesso le predizioni a lungo termine delle traiettorie di un EV (problema di regressione) sono qualitativamente corrette ma manifestano una traslazione in avanti rispetto al target, ovvero una specie di "ritardo" che è difficile cogliere con le metriche tradizionali (quelle sopra elencate). Tuttavia, gli autori non forniscono una soluzione per evitare di incorrere in questo problema. Diversamente, in questo lavoro di tesi, il problema è stato affrontato considerando una ulteriore metrica per valutare la regressione delle serie temporali del parametro T_2 e selezionare, in combinazione alle altre, il modello $\mathrm{GRU_{T2}}$ più idoneo; si tratta di un parametro usualmente impiegato nel contesto idrologico per problemi di previsione su più orizzonti temporali nel futuro [235], chiamato *Modified Index of Agreement* (md) [236]. In particolare, md è in grado di considerare contemporaneamente le differenze nelle medie e nelle varianze osservate e previste, fornendo una migliore valutazione delle previsioni del modello rispetto alle metriche tradizionali:

$$md = 1 - \left[\sum_{t=1}^{T} \left| \hat{y}_g^{(t)} - y_g^{(t)} \right| \Big/ \sum_{t=1}^{T} \left| \hat{y}_g^{(t)} - \bar{y}_g \right| + \left| y_g^{(t)} - \bar{y}_g \right| \right] \tag{5.1}$$

con $\bar{y}_g = 1/T \left(\sum_t y_g^{(t)} \right)$ la media delle T osservazioni di un partecipante per l'orizzonte temporale g-esimo. L'Eq. 5.1 deve essere applicata a ciascun partecipante appartenente ad un *fold* di validazione e per ogni orizzonte temporale di previsione; quindi, md può essere mediato sui cinque *folds* per fornire una stima esaustiva delle prestazioni di un modello di regressione, sia sul g-esimo orizzonte di previsione che, complessivamente, sui tre istanti futuri

considerati. Ovviamente, quanto più md è prossimo ad 1, migliore è l'adattamento delle predizioni ai target sperimentali.

I punteggi delle metriche di valutazione, conseguiti da ciascuna combinazione di iperparametri nel processo di ricerca a griglia, sono riportati nelle Tabelle B1, B2 e B3 dell'Appendice B, rispettivamente per i modelli $\mathrm{GRU}_{\mathrm{T2}}$, $\mathrm{GRU}_{\mathrm{TADV}}$ e $\mathrm{S\text{-}GRU}_{\mathrm{SL}}$; i modelli ottimali, ovvero quelli con i migliori risultati medi sui cinque *folds* di validazione (considerando tutti e tre gli orizzonti temporali di previsione), sono evidenziati in grassetto e sono oggetto di maggiori analisi in questa Sezione. La selezione di tali modelli è stata basata principalmente sulla metrica md per la previsione del parametro T_2, e sulla metrica AUC acquista dalla *Receiver Operating Characteristic Curve* (o curva ROC) per ciascun problema di classificazione. Questa curva (maggiori dettagli sono riportati in Appendice B), che traccia il *Recall* in funzione del FAR, è una metrica completa ed esauriente per valutare le prestazioni dei modelli di classificazione [85], poiché più l'area sottesa dalla curva ROC (corrispondente al parametro AUC) è prossima ad 1, migliore è la qualità della previsione (quindi maggiore è la corrispondenza tra le classi previste e quelle attese). In particolare, i risultati riportati nelle Tabelle B2 e B3 sono stati ottenuti senza l'applicazione del metodo SMOTE, per rendere il processo di ricerca a griglia più rapido e meno complesso dal punto di vista computazionale; tuttavia, in questa Sezione sono riportati i risultati derivanti dall'applicazione del metodo SMOTE ai modelli ottimali identificati dalla ricerca a griglia (in grassetto nelle Tabelle B2 e B3) e sono valutati i benefici di tale metodo attraverso l'analisi dei punteggi di validazione.

In merito ai modelli ottimali, il confronto delle prestazioni (md per la regressione, AUC per la classificazione) delle unità GRU derivanti dalle combinazioni di valori per N, α e δ (Appendice B), ha permesso di individuare i valori più appropriati di questi iperparametri per ognuno dei tre modello in questione: il valore iniziale del tasso di apprendimento globale α è fissato a 0.001 per il modello $\mathrm{GRU}_{\mathrm{TADV}}$ ed a 0.005 per entrambi i modelli $\mathrm{GRU}_{\mathrm{T2}}$ e $\mathrm{S\text{-}GRU}_{\mathrm{SL}}$, mentre il fattore di riduzione del tasso di apprendimento δ è pari a 0.80 per tutti i modelli considerati. Il numero di neuroni N dell'unità di memoria GRU è 64 per $\mathrm{GRU}_{\mathrm{T2}}$, 256 per $\mathrm{GRU}_{\mathrm{TADV}}$ e 128 per $\mathrm{S\text{-}GRU}_{\mathrm{SL}}$. Le Tabelle 6 e 8 presentano i punteggi delle metriche di valutazione, validati sui cinque *folds*, distinguendo i risultati per i tre modelli ottimali ($\mathrm{GRU}_{\mathrm{T2}}$ e $\mathrm{GRU}_{\mathrm{TADV}}$ in Tabella 6; $\mathrm{S\text{-}GRU}_{\mathrm{SL}}$ in Tabella 8) e per ciascun orizzonte temporale di previsione (le deviazioni standard sui cinque *folds* sono riportate tra parentesi). Inoltre, la Figura 21 mostra la distribuzione cumulata di frequenza dell'errore (in valore assoluto) di predizione su ciascun *fold* per il modello $\mathrm{GRU}_{\mathrm{T2}}$. Diversamente, le Figure 23 e 24 rappresentano le curve ROC per i modelli $\mathrm{GRU}_{\mathrm{TADV}}$ e $\mathrm{S\text{-}GRU}_{\mathrm{SL}}$, rispettivamente. Infine, mentre le Tabelle 7 e 10 definiscono il contributo del metodo SMOTE al miglioramento delle prestazioni delle unità $\mathrm{GRU}_{\mathrm{TADV}}$ e $\mathrm{S\text{-}GRU}_{\mathrm{SL}}$, la Tabella 9 riassume le capacità di generalizzazione dei modelli di classificazione della severità, $\mathrm{M\text{-}GRU}_{\mathrm{SL}}$ e $\mathrm{S\text{-}GRU}_{\mathrm{SL}}$, derivanti dall'Approccio 1) e 2), attraverso i valori medi delle metriche sui *folds* di validazione (indipendentemente dall'orizzonte temporale di previsione).

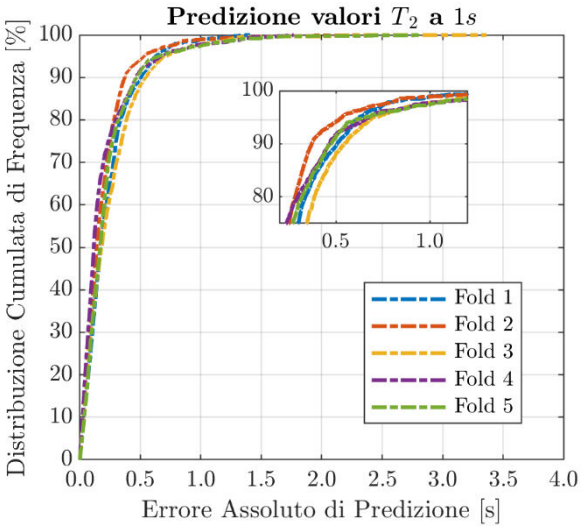
Come atteso, la qualità delle previsioni aumenta se l'orizzonte temporale di previsione diminuisce, indipendentemente dal modello considerato. Questo

risultato suggerisce l'esistenza di una forte correlazione tra le caratteristiche di mobilità selezionate al tempo t e le variabili dipendenti selezionate al tempo $t + 1\,s$. Al contrario, questa correlazione diventa più debole per orizzonti temporali più lunghi e, di conseguenza, distinguere il target atteso diventa più difficile.

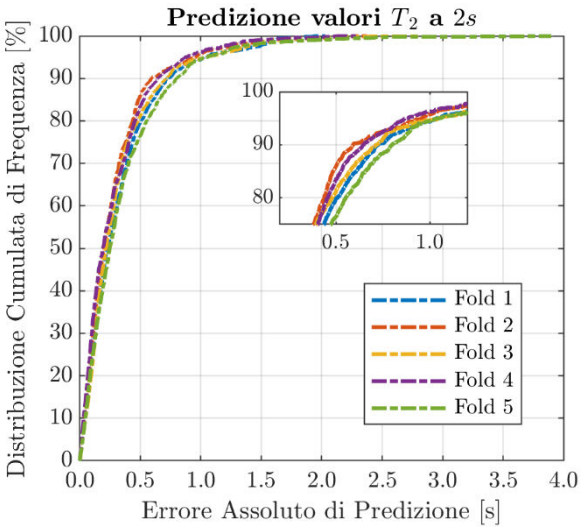
Per esempio, considerando il modello $\mathrm{GRU_{T2}}$ (Tabella 6), tutte le metriche di valutazione peggiorano gradualmente, seppur in maniera molto lieve, man mano che l'orizzonte temporale di previsione si allunga, sia in fase di addestramento che di validazione. Infatti, concentrando l'attenzione sulla metrica $RMSE$ in validazione (quinta colonna, Tabella 6), cioè la radice quadrata della differenza quadratica media tra i valori predetti e quelli osservati al simulatore, il suo peggioramento è caratterizzato da uno scarto prossimo al decimo di secondo: lo scarto massimo, pari a $0.138\,s$, si misura passando dall'orizzonte di previsione di $1\,s$ ($RMSE = 0.327\,s$) a quello di $2\,s$ ($RMSE = 0.465\,s$), mentre la qualità della previsione peggiora di $0.122\,s$ passando da $2\,s$ a $3\,s$ ($RMSE = 0.587\,s$). Pertanto, i risultati ottenuti dal modello $\mathrm{GRU_{T2}}$ sono nel complesso soddisfacenti, come dimostrato anche dalle altre metriche: il MAE in validazione (terza colonna, Tabella 6), cioè l'ordine di grandezza dell'errore medio di previsione in valore assoluto, è risultato su ogni scala temporale inferiore al mezzo secondo; il $90°$ percentile dell'errore assoluto (Figura 21), cioè il valore dell'errore di previsione (in valore assoluto) che viene superato in non più del 10% delle sequenze temporali, è pari a $0.472\,s$ (media sulle cinque *folds*) sull'orizzonte di $1\,s$, $0.696\,s$ su $2\,s$ e $0.967\,s$ su $3\,s$; il parametro md è sempre maggiore di 0.850 qualunque sia l'orizzonte temporale di previsione, dimostrando la stretta corrispondenza tra le curve osservate e quelle predette per l'indicatore T_2. Inoltre, l'ordine di grandezza e la prossimità dei valori dell'$RMSE$ tra l'addestramento (seconda colonna, Tabella 6) e la validazione (quinta colonna) del modello ottimale $\mathrm{GRU_{T2}}$, evidenziano che il problema dell'*overfitting* è stato evitato, come risultato dell'applicazione delle strategie di regolarizzazione L2 ed *early stopping* (illustrate alle Sezioni 4.2 e 4.5).

Tabella 6: Risultati sperimentali (deviazioni standard) per i modelli $\mathrm{GRU_{T2}}$ e $\mathrm{GRU_{TADV}}$.

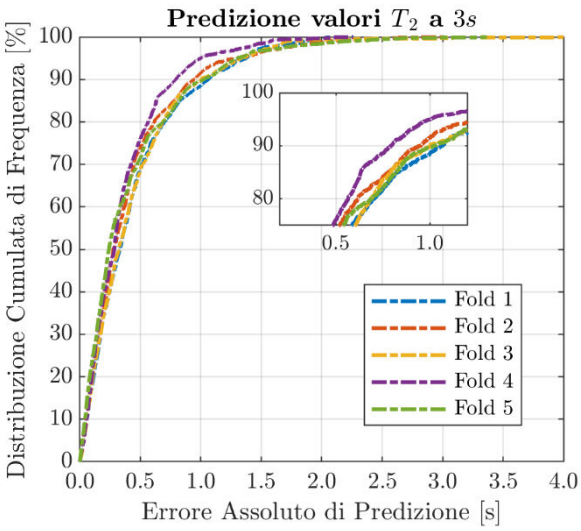
$\mathrm{GRU_{T2}}$	*Training*	**Validazione**					
Orizzonte	$RMSE$	$MAE\ [s]$	$MAPE\ [\%]$	$RMSE\ [s]$	$md\ [-]$		
1s	0.282 (0.017)	0.221 (0.024)	4.539 (0.371)	0.327 (0.033)	0.943 (0.003)		
2s	0.386 (0.014)	0.319 (0.029)	6.617 (0.437)	0.465 (0.042)	0.902 (0.007)		
3s	0.497 (0.036)	0.415 (0.034)	9.048 (0.483)	0.587 (0.054)	0.852 (0.007)		
$\mathrm{GRU_{TADV}}$	*Training*	*Validazione*					
Orizzonte	AUC	*Accuracy*	*Precision*	*Recall*	*Specificity*	FAR	AUC
1s	0.997 (0.001)	0.974 (0.003)	0.919 (0.017)	0.914 (0.023)	0.985 (0.004)	0.015 (0.004)	0.997 (0.001)
2s	0.995 (0.001)	0.968 (0.004)	0.892 (0.027)	0.902 (0.020)	0.980 (0.005)	0.020 (0.005)	0.994 (0.001)
3s	0.992 (0.002)	0.962 (0.007)	0.880 (0.029)	0.874 (0.054)	0.978 (0.006)	0.022 (0.006)	0.991 (0.004)



(a)



(b)



(c)

Figura 21: Distribuzioni cumulate di frequenza degli errori assoluti di previsione del modello GRU_{T2} sui cinque folds di validazione, per gli orizzonti temporali di (a) $1\ s$, (b) $2\ s$ e (c) $3\ s$.

Per quanto riguarda la qualità della previsione offerta dai modelli ottimali di classificazione (dopo l'applicazione del metodo SMOTE), è necessario ricordare al lettore che, in un problema di classificazione binaria, le osservazioni sono annotate come "positive" e "negative" affinché le metriche di valutazione possano essere calcolata direttamente attraverso la matrice di confusione (Appendice B), detta anche tabella di errata classificazione. Quest'ultima è una rappresentazione dell'accuratezza del classificatore in quanto rende visivamente chiaro se il modello in questione annota erroneamente una classe con un'altra: ogni riga della matrice di confusione rappresenta le istanze in una classe "reale" (perché effettivamente riflette l'obbiettivo della classificazione) mentre ogni colonna rappresenta quelle in una classe prevista. Quindi, l'elemento sulla riga i-esima e colonna j-esima rappresenta il numero di volte in cui il classificatore ha annotato la classe "vera" i come quella j. Così, il numero di "veri positivi" (positivi correttamente classificati), "veri negativi" (negativi correttamente classificati), "falsi positivi" (negativi reali classificati in modo errato) e "falsi negativi" (positivi reali classificati in modo errato) può essere definito. Per una migliore comprensione, una rappresentazione schematica della matrice di confusione è mostrata in Figura 22.

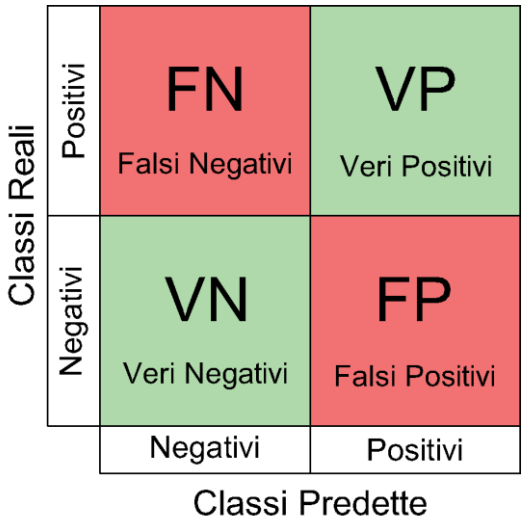


Figura 22: Rappresentazione schematica della matrice di confusione.

Per valutare adeguatamente il modello di classificazione del T_{ADV}, i campioni della classe "*collision course*" sono stati annotati come positivi, siccome questa classe rappresenta quella di maggior rischio per il pedone e, allo stesso tempo, quella minoritaria nel problema generale di classificazione in questione (sulla quale è importante focalizzare maggiormente l'attenzione). Pertanto, il *Recall* (colonna 5, Tabella 6) del modello $\mathrm{GRU_{TADV}}$, che rappresenta la proporzione di campioni positivi effettivi classificati correttamente, dimostra che il sistema è in grado di prevedere molto accuratamente (i valori sono maggiori di 0.870) se l'interazione tra veicolo e pedone si sta muovendo o meno su una rotta di collisione, qualunque sia l'orizzonte di previsione (nonostante il lieve peggioramento discusso in precedenza). Allo stesso modo, i valori AUC (ultima colonna di Tabella 6), superiori a 0.995 per ogni orizzonte temporale futuro,

confermano l'elevata precisione del classificatore binario. Diversamente, la Figura 23 mostra l'effetto della variabilità nella distribuzione dei dati sui *folds* di validazione attraverso le curve ROC: estendendo la finestra temporale di previsione, le differenze tra le partizioni sono più evidenti, cioè la distribuzione dei dati diventa più sparsa, e il modello perde leggermente in capacità di generalizzazione. Tuttavia, la prestazione del modello è ancora più che accettabile. Per quanto riguarda il potenziale problema dell'*overfitting*, i valori assunti dall'AUC in addestramento (seconda colonna, Tabella 6) ed in validazioni (ottava colonna, Tabella 6) sono quasi coincidenti, a dimostrazione che l'unità $\mathrm{GRU_{TADV}}$ è capace di generalizzare in maniera adeguata. Il confronto tra le prestazioni del modello $\mathrm{GRU_{TADV}}$ identificato dalla ricerca a griglia e quello derivante dalla conseguente applicazione del metodo SMOTE (Tabella 6) è presentato in Tabella 7. In particolare, per dimostrare l'efficacia dello SMOTE, l'ultima riga di Tabella 7 riporta la variazione percentuale $VP_{\%}$ di ciascuna metrica di valutazione del classificatore $\mathrm{GRU_{TADV}}$ a seguito dell'applicazione del metodo in questione, definita nel seguente modo:

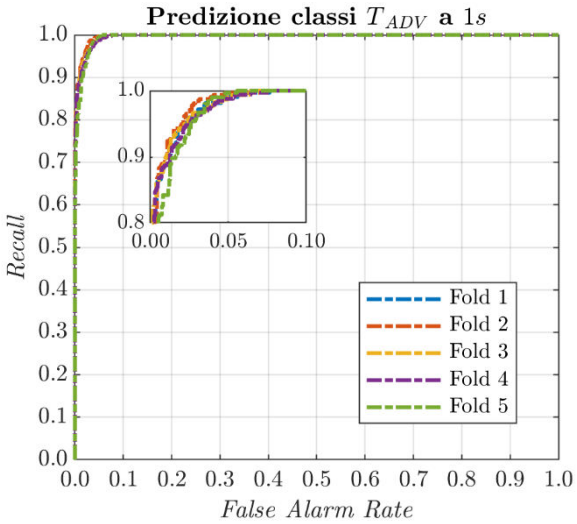
$$VP_{\%} = 100\frac{M_{CON} - M_{SENZA}}{M_{SENZA}}, \tag{5.2}$$

nella quale M_{senza} ed M_{con} rappresentano la medesima metrica di valutazione ma il loro valore è riferito alla seconda e terza riga di Tabella 7, rispettivamente. È facile notare, analizzando la variazione percentuale dell'AUC di validazione (pari al $+0.22\%$), che il metodo SMOTE non ha contribuito in maniera significativa all'incremento delle prestazioni globali del modello $\mathrm{GRU_{TADV}}$, probabilmente sia a causa della forte correlazione delle classi target alle caratteristiche di mobilità selezionate sia per lo sbilanciamento non troppo eccessivo registrato tra le classi del T_{ADV} (il rapporto iniziale tra le classi "*crossing course*" e "*collision course*" era $5.5{:}\,1$, Tabella 5); tuttavia, come atteso, la capacità dell'unità $\mathrm{GRU_{TADV}}$ di discriminare la classe "*collision course*" da quella "*crossing course*" è aumentata del $+1.24\%$, come testimoniato dal parametro *Recall* (quarta colonna, Tabella 7).

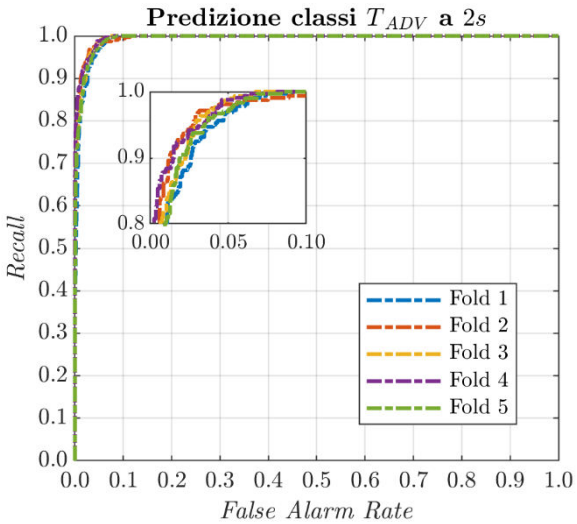
Tabella 7: Confronto tra i modelli $\mathrm{GRU_{TADV}}$ con e senza applicazione dello SMOTE.

Confronto risultati	**Valori medi (deviazioni standard) sui *folds* di validazione**					
	Accuracy	*Precision*	*Recall*	*Specificity*	FAR	AUC
$\mathrm{GRU_{TADV}}$ senza SMOTE	0.964 (0.007)	0.881 (0.029)	0.886 (0.046)	0.978 (0.006)	0.022 (0.006)	0.992 (0.004)
$\mathrm{GRU_{TADV}}$ con SMOTE	0.968 (0.007)	0.897 (0.028)	0.897 (0.037)	0.981 (0.006)	0.019 (0.006)	0.994 (0.003)
Variazione percentuale	$+0.41\%$	$+1.82\%$	$+1.24\%$	$+0.31\%$	-13.64%	$+0.22\%$

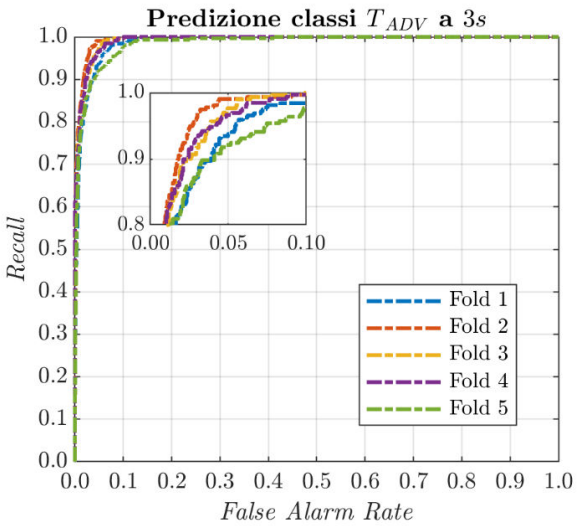
Per i modelli $\mathrm{S\text{-}GRU_{SL}}$ e $\mathrm{M\text{-}GRU_{SL}}$, le istanze all'interno della classe di severità "*unsafe*" sono state annotate come positivi, siccome nelle applicazioni FCW e AEB è necessario ridurre al minimo le false attivazioni del sistema. Nel seguito, poiché il calcolo dell'AUC per il modello derivante dalla combinazione delle predizioni dei sottosistemi $\mathrm{GRU_{T2}}$ e $\mathrm{GRU_{TADV}}$ non può essere eseguito (Appendice B), il confronto tra gli Approcci 1) e 2) è stato basato principalmente sulle metriche *Accuracy* e *Recall* (terza e quinta colonna, Tabella 8).



(a)



(b)



(c)

Figura 23: Curve ROC del modello $\mathrm{GRU}_{\mathrm{TADV}}$ ***sui cinque folds di validazione, per gli orizzonti temporali di (a)*** $1\ s$***, (b)*** $2\ s$ ***e (c)*** $3\ s$***.***

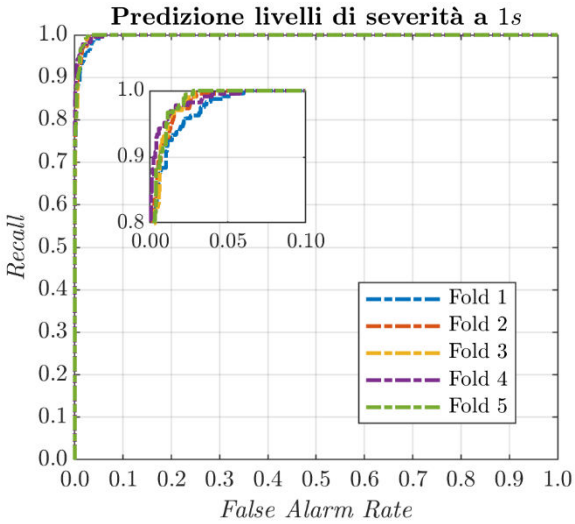
L'analisi dei risultati sulle partizioni di validazione ha evidenziato una migliore capacità del modello S-GRU_{SL} di annotare le interazioni conflittuali gravi negli eventi di quasi incidente tra veicolo e pedone, specialmente sugli orizzonti di previsione più lunghi. Infatti, l'*Accuracy* del modello S-GRU_{SL}, cioè la proporzione di campioni (positivi e negativi) classificati correttamente tra tutte le istanze della serie, è superiore a 0.970 (terza colonna, Tabella 8) su ogni scala temporale di previsione, a differenza del modello M-GRU_{SL} che solo sull'orizzonte temporale più breve ($1\ s$) raggiunge il valore di 0.971. La situazione peggiora ulteriormente per il modello M-GRU_{SL}, a vantaggio dell'unità S-GRU_{SL}, in termini di *Recall* (quinta colonna, Tabella 8): infatti, confrontando i valori medi sui *folds* di validazione (indipendentemente dall'orizzonte temporale di predizione), il guadagno percentuale sulla metrica *Recall* che si acquisisce passando dall'unità M-GRU_{SL} alla S-GRU_{SL} è pari ad un $+15.40\%$ (quarta colonna, Tabella 9). Di conseguenza, si osserva una drastica riduzione (-15.38%, sesta colonna, Tabella 9) del FAR, ovvero la proporzione di campioni effettivamente negativi classificati erroneamente, mentre le restanti metriche (*Precision* e *Specificity*, Appendice B) migliorano grazie alla riduzione del numero di falsi positivi (terza e quinta colonna, Tabella 9). Si tratta di risultati significativi considerando una futura applicazione online del sistema, in quanto prevedere accuratamente l'evoluzione dell'evento (e, come dimostrato, il modello S-GRU_{SL} ha capacità nettamente superiori al M-GRU_{SL}) permette di evitare eccessive interferenze del sistema con il normale comportamento di guida.

Tabella 8: Risultati sperimentali (deviazioni standard) per i modelli M-GRU_{SL} e S-GRU_{SL}.

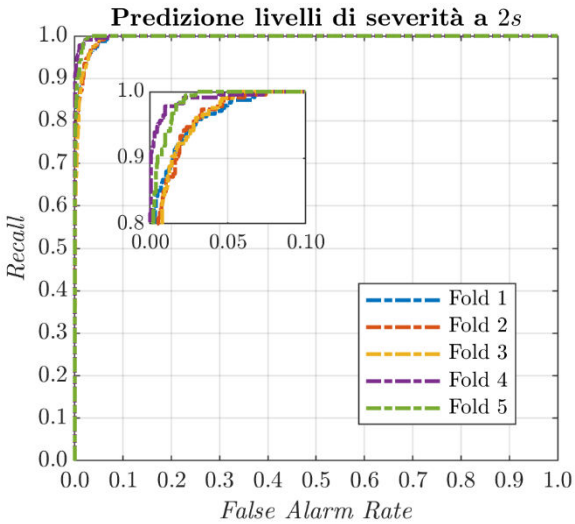
	M-GRU_{SL}	**Validazione**					
	Orizzonte	*Accuracy*	*Precision*	*Recall*	*Specificity*	FAR	
	$1s$	0.971 (0.003)	0.879 (0.049)	0.844 (0.068)	0.986 (0.006)	0.014 (0.006)	
	$2s$	0.967 (0.005)	0.884 (0.052)	0.792 (0.107)	0.987 (0.007)	0.013 (0.007)	
	$3s$	0.958 (0.009)	0.873 (0.036)	0.700 (0.095)	0.988 (0.002)	0.012 (0.002)	
S-GRU_{SL}	***Training***	**Validazione**					
Orizzonte	AUC	*Accuracy*	*Precision*	*Recall*	*Specificity*	FAR	AUC
$1s$	0.998 (0.001)	0.984 (0.004)	0.932 (0.029)	0.917 (0.042)	0.992 (0.005)	0.008 (0.005)	0.998 (0.001)
$2s$	0.997 (0.001)	0.981 (0.006)	0.911 (0.043)	0.907 (0.027)	0.989 (0.006)	0.011 (0.006)	0.997 (0.002)
$3s$	0.996 (0.002)	0.974 (0.006)	0.887 (0.038)	0.873 (0.057)	0.987 (0.006)	0.014 (0.006)	0.995 (0.002)

Tabella 9: Confronto tra i modelli M-GRU_{SL} e S-GRU_{SL}.

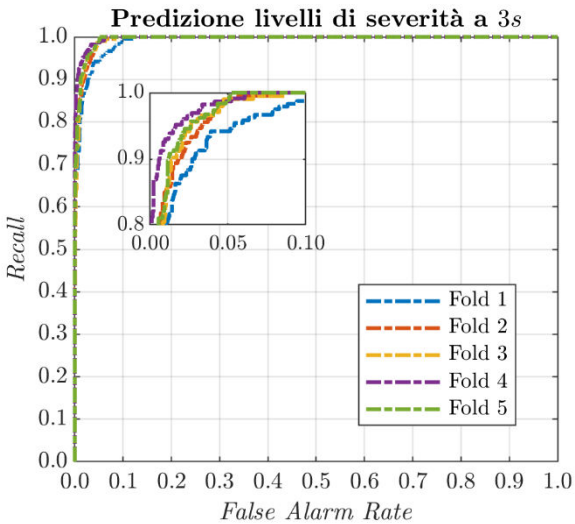
Confronto risultati	**Valori medi (deviazioni standard) sui *folds* di validazione**					
	Accuracy	*Precision*	*Recall*	*Specificity*	FAR	AUC
Prestazioni M-GRU_{SL}	0.965 (0.008)	0.879 (0.043)	0.779 (0.105)	0.987 (0.005)	0.013 (0.005)	-
Prestazioni S-GRU_{SL}	0.980 (0.007)	0.910 (0.039)	0.899 (0.045)	0.989 (0.006)	0.011 (0.006)	0.996 (0.002)
Variazione percentuale	$+1.55\%$	$+3.53\%$	$+15.40\%$	$+0.20\%$	-15.38%	-



(a)



(b)



(c)

Figura 24: Curve ROC del modello S-GRUSL sui cinque folds di validazione, per gli orizzonti temporali di (a) 1 s, (b) 2 s e (c) 3 s.

Escludendo il potenziale problema dell'*overfitting* per il modello singolo $\mathrm{S\text{-}GRU_{SL}}$ (confrontando i risultati AUC in *training* e validazione, seconda e ottava colonna di Tabella 8, si osservano valori praticamente coincidenti), l'insoddisfacente capacità di generalizzazione del modello combinato $\mathrm{M\text{-}GRU_{SL}}$ potrebbe essere legata ad un effetto di propagazione degli errori commessi dai modelli singoli $\mathrm{GRU_{T2}}$ e $\mathrm{GRU_{TADV}}$. Infatti, nonostante gli errori siano di modesta entità (Tabella 6), la loro combinazione potrebbe aver determinato prestazioni per l'unità $\mathrm{M\text{-}GRU_{SL}}$ nettamente inferiori a $\mathrm{S\text{-}GRU_{SL}}$ (Tabelle 8 e 9); tuttavia, siccome il problema di classificazione della severità ha carattere non lineare la determinazione dell'effetto di propagazione dell'errore è difficile da cogliere analiticamente. Ad ogni modo, questa constatazione non è banale in quanto, in molti problemi di previsione delle traiettorie su più orizzonti temporali futuri (Sezione 2.1.3), soluzioni combinate (ad esempio, i *Multiple* RNN) hanno ottenuto risultati migliori rispetti agli equivalenti modelli "compatti", come i *Single* RNN [152]. Di conseguenza, è probabile che il problema di predizione degli indicatori singoli, legati alla valutazione della severità, debba essere affrontato in maniera alternativa oppure più dettagliata, considerando per esempio nuove e/o diverse variabili di input (come la sequenza degli stati passati del veicolo e del pedone) in grado di migliorare le capacità predittive delle unità $\mathrm{GRU_{T2}}$ e $\mathrm{GRU_{TADV}}$. Infatti, si vuole ribadire che l'Approccio 1) consentirebbe una maggiore flessibilità nell'annotazione degli eventi quasi-incidentali durante l'applicazione del sistema online, poiché valori di soglia degli SSM specifici per i contesti geografici di funzionamento del sistema potrebbero essere adottati per classificare il livello di rischio delle interazioni veicolo-pedone [85].

Tabella 10: Confronto tra i modelli $\mathrm{S\text{-}GRU_{SL}}$ con e senza applicazione dello SMOTE.

Confronto risultati	**Valori medi (deviazioni standard) sui *folds* di validazione**					
	Accuracy	*Precision*	*Recall*	*Specificity*	FAR	AUC
$\mathrm{S\text{-}GRU_{SL}}$ senza SMOTE	0.962 (0.010)	0.833 (0.065)	0.802 (0.074)	0.980 (0.009)	0.020 (0.009)	0.987 (0.007)
$\mathrm{S\text{-}GRU_{SL}}$ con SMOTE	0.980 (0.007)	0.910 (0.039)	0.899 (0.045)	0.989 (0.006)	0.011 (0.006)	0.996 (0.002)
Variazione percentuale	+1.87%	+9.24%	+12.09%	+0.92%	-45.00%	+0.91%

Ad ogni modo, l'analisi condotta in questa tesi ha dimostrato che è possibile predire in modo soddisfacente e soprattutto diretto i livelli di severità da una rappresentazione minima ed efficiente della situazione di traffico, attraverso un unico modello neurale ricorrente. Inoltre, rispetto a quanto precedentemente evidenziato per il modello $\mathrm{GRU_{TADV}}$, il metodo SMOTE ha contribuito in maniera significativa all'incremento delle prestazioni dell'unità $\mathrm{S\text{-}GRU_{SL}}$ (Tabella 10), risultato atteso dato lo sbilanciamento iniziale tra le classi "*safe*" e "*unsafe*" (pari a $8.4{:}\,1$, Tabella 5) era molto superiore a quello registrato tra le classi del T_{ADV}. In particolare, sebbene il guadagno percentuale sul parametro AUC di validazione sembri modesto ($+0.91\%$, ultima colonna, Tabella 10), la capacità del modello di classificare correttamente la classe "*unsafe*" è aumentata del $+12.09\%$ (*Recall*, quarta colonna, Tabella 10); conseguentemente, anche le restanti metriche (*Precision* e *Specificity*) migliorano grazie alla riduzione del numero di falsi positivi (terza e quinta colonna, Tabella 10). Diversamente, la Figura 24 mostra l'effetto

della variabilità nella distribuzione dei dati sui *folds* di validazione attraverso le curve ROC, che rispetto alla Figura 23 sembra qualitativamente più contenuto sull'orizzonte temporale più lungo (come dimostrano anche quantitativamente le deviazioni standard dell'AUC, ultima colonna di Tabella 6 e 8).

Per riassumere, nonostante i buoni risultati ottenuti nella modellazione individuale dei parametri di sicurezza surrogati T_2 e T_{ADV}, il modello $\mathrm{S\text{-}GRU_{SL}}$ ha ottenuto le migliori prestazioni, con una *Accuracy* di 0.980, *Recall* di 0.899 e AUC di 0.996 (in media sulle partizioni di validazione, indipendentemente dall'orizzonte temporale di previsione, Tabella 8), nella previsione degli eventi quasi incidentali tra un veicolo ed un pedone nel traffico urbano. In conclusione, i risultati ottenuti costituiscono evidenze sufficienti, ed altrettanto promettenti (se confrontate con i punteggi dei modelli di previsione delle traiettorie in Sezione 2.1.3), per supportare lo sviluppo del sistema proposto e la sua applicazione nella guida su strada.

Capitolo 6

Conclusioni e Ricerche future

Gli studi di ricerca a supporto di questa tesi hanno dimostrato che è possibile elaborare un sistema di *Deep Learning*, basato su unità GRU, capace di prevedere in maniera affidabile l'evoluzione del livello di rischio nel corso delle interazioni tra un veicolo ed un pedone nel traffico urbano. In particolare, il sistema proposto, nella sua formulazione concettuale, prevede l'estrazione e l'elaborazione delle traiettorie degli utenti dalle nuvole di punti acquisite in tempo reale da uno o più RADAR a onde millimetriche ad alta risoluzione; quindi, attraverso un ottimizzatore del tipo *gradient-descent*, tale sistema sfrutta la conoscenza acquisita dalle simulazioni in ambiente virtuale per annotare le caratteristiche individuali di guida e le proprietà della scena in livelli di severità dell'interazione.

Rispetto alle procedure convenzionali di valutazione della sicurezza dei pedoni, l'algoritmo proposto evita il complesso problema multimodale della previsione delle traiettorie degli utenti e si focalizza sull'obbiettivo di acquisire direttamente dai dati il modo in cui gli utenti si comportano in situazioni potenzialmente pericolose per evitare una collisione. Questo è possibile grazie all'impiego di indicatori SSM in grado di descrivere in modo continuo il processo di interazione tra veicolo e pedone, e architetture neurali che possono lavorare con output sequenziali.

Il sistema $\mathrm{S\text{-}GRU_{SL}}$, prototipato con dati acquisiti durante un esperimento di simulazione di guida, ha fornito risultati soddisfacenti per la sua generalizzazione (AUC in validazione pari a 0.996) e promettenti per il miglioramento dei sistemi ADAS esistenti (*Accuracy* pari a 0.980, *Recall* pari a 0.899 in media sulle cinque partizioni di validazione). L'algoritmo prevede su tre orizzonti temporali futuri, precisamente a $t + 1\ s$, $t + 2\ s$ e $t + 3\ s$ rispetto alla situazione corrente t, le due classi di severità "*safe*" e "*unsafe*" dell'interazione, definite sulla base di valori di soglia statici per gli indicatori T_2 e T_{ADV}. Stabilite le regole per l'algoritmo gerarchico di preallarme e/o intervento sui pedali (il livello di esecuzione) e validato l'algoritmo anche per complessi scenari di interazione, questo sistema potrebbe essere sfruttato per avvisare (con il giusto *timing*) i conducenti di un'interazione anomala o pericolosa con un pedone, evitando inutili interferenze con il normale comportamento di guida, oppure per anticipare una manovra di frenata, assicurando una decelerazione del veicolo graduale e meno brusca.

Le prospettive di ricerca futura, aperte dalle evidenze sperimentali ottenute, includono: 1) il confronto dell'algoritmo prototipato con un equivalente sistema basato sulla previsione delle traiettorie, da elaborare sulla base dei medesimi dati acquisiti al simulatore; 2) la generalizzazione del modello $\mathrm{S\text{-}GRU_{SL}}$ a diversi e più complessi scenari di incontro; 3) lo studio dei valori di soglia per gli indicatori più appropriati per il contesto geografico di applicazione del sistema; 4) la modifica del sistema per l'apprendimento ed il funzionamento online, eventualmente riconsiderando l'opzione di scindere il problema nella predizione dei singoli indicatori SSM; 5) la validazione del sistema attraverso dati acquisiti durante incontri veicolo-pedone su strada.

In effetti, si dovrebbe in primo luogo stabilire un riferimento più appropriato per confrontare l'approccio presentato con quelli ampiamente utilizzati in letteratura (i modelli di previsione delle traiettorie). Inoltre, sono necessari ulteriori sforzi di ricerca su una campione più ampio di conducenti, così come su diversi gruppi di utenti (più o meno esperti, anziani, ecc.), in diversi ambienti urbani e situazioni di guida, per considerare il maggior numero possibile di componenti comportamentali (come l'aggressività, l'ansia, la ricerca di sensazioni alla guida) e schemi di interazione. La definizione di valori soglia concretamente legati ai comportamenti assunti sulla strada dai conducenti in un determinato contesto geografico, per esempio applicando la teoria degli eventi estremi a dati acquisiti da telecamere (in posizione fissa ai margini delle intersezioni oppure installate su veicoli per uno studio di guida naturalistico), potrebbe favorire una annotazione più accurata delle classi di severità, favorendo l'adattamento per imitazione del sistema. Ad ogni modo, prima della validazione online del sistema (cioè la validazione del sistema proposto su dati acquisiti in scenari di guida reali), è necessario integrare il modello $\mathrm{S\text{-}GRU_{SL}}$ di valutazione del rischio con un livello di percezione che permetta di effettuare in automatico i processi di filtraggio e *clustering* dei dati grezzi sensoriali, così come la classificazione e tracciamento degli oggetti in movimento: in particolare, deve essere testato il funzionamento coordinato in tempo reale tra il livello di percezione ed il sistema di rilevamento stabiliti per il veicolo prototipo, e la capacità di acquisire dati di mobilità affidabili, sia durante preliminari simulazioni in ambiente virtuale che successivamente in condizioni di guida reali. In questa fase, potrebbe essere necessario introdurre sottosistemi di apprendimento specifici per la valutazione degli scenari più complessi (ad esempio, l'interazione simultanea dell'EV con i veicoli circostanti e i pedoni), congiuntamente ad un metodo per unire le previsioni fornite da ciascun sottosistema. Tuttavia, il successivo passaggio verso l'applicazione su strada dell'algoritmo richiede la definizione di un ulteriore requisito funzionale: sebbene il modello di valutazione del rischio sia essenziale per la sicurezza funzionale di un sistema AEB, la chiave per ottenere l'evitamento attivo della collisione è il controllo della dinamica del veicolo. A questo scopo, è indispensabile progettare i *controller* di livello superiore e inferiore, i quali, ricevuto il segnale di controllo dal livello di esecuzione del RAM (a seguito del probabile verificarsi di una collisione con un pedone), sono incaricati rispettivamente di stimare il valore di decelerazione richiesto per un arresto in sicurezza e di azionare i sottosistemi del veicolo (cioè, la chiusura della valvola a farfalla e la regolazione della pressione sui freni) per realizzare il controllo attivo del veicolo. La strategia di controllo proposta da Yang e coll. [167], basata sulla rete neurale fuzzy e sulla teoria PID (*Proportional Integral Derivative controller*), potrebbe rappresentare un pratico riferimento per l'implementazione dei *controller*. Solo a questo punto, utilizzando un veicolo adeguatamente equipaggiato, il sistema sviluppato di prevenzione delle collisioni tra veicolo e pedone potrà (e dovrà) essere convalidato online in scenari test stabiliti dalle norme Euro-NCAP [237].

Per concludere, la ricerca esposta in questa tesi ha rappresentato un valido contributo allo studio e alla comprensione dei processi di interazione tra gli utenti della strada e, di conseguenza, al miglioramento della sostenibilità dei trasporti all'interno del traffico urbano.

Appendice A

Calcolo del Time-to-Collision per un punto ed un segmento

Analizzando le collisioni tra veicoli, si osserva che è sempre un angolo di uno dei due veicoli coinvolti che incontra un lato dell'altro (supponendo che i veicoli abbiano forme rettangolari, per semplicità). Poiché, nel caso generale, non si sa quale angolo incontra quale lato, si devono analizzare tutte le possibili combinazioni (in totale 32 combinazioni). La procedura per calcolare il TTC per un segmento mobile ed un punto (cioè un lato e un angolo dei veicoli coinvolti) è presentato di seguito. Il valore TTC più basso trovato tra tutte le combinazioni angolo-lato deve essere considerato, poiché rappresenta il lato e l'angolo che entreranno in contatto per primi in una potenziale collisione. Ovviamente, lo stesso discorso continua a valere anche per gli incontri tra veicolo e pedone, dove il pedone è il punto e il frontale del veicolo è il segmento. Il TTC è un parametro continuo e può essere calcolato per qualsiasi istante finché gli utenti della strada sono in rotta di collisione.

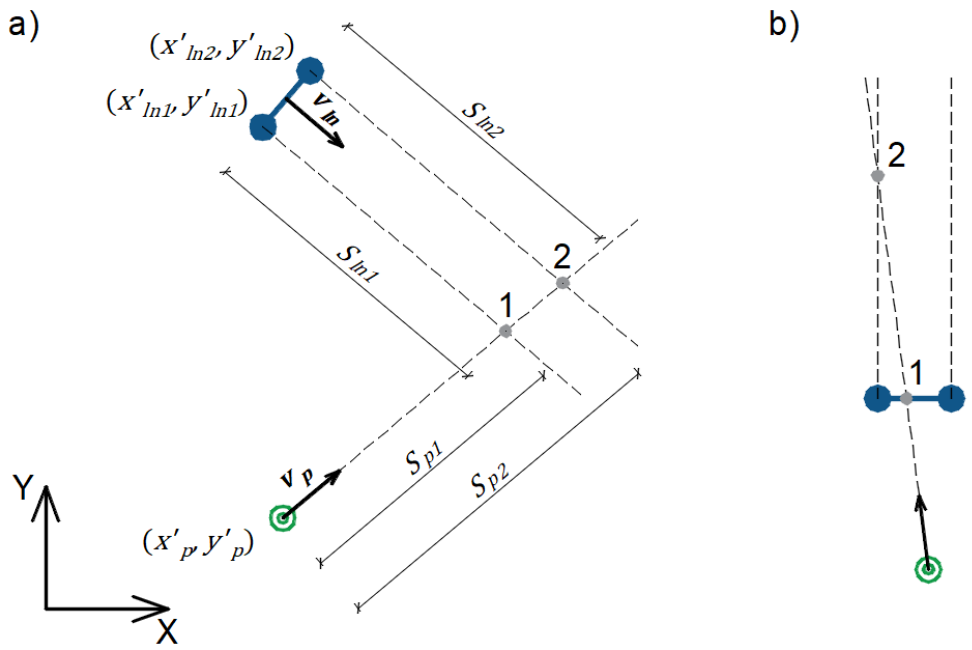


Figura A1: a) schema per il calcolo di TTC, T_{ADV} e T_2 per un punto ed un segmento, b) esempio del caso particolare $t_{ln\,1} = 0$.

Sia (x_p, y_p) la coppia di coordinate attuali del punto e $\boldsymbol{v_p}$ il suo vettore di velocità (Figura A1a). Di conseguenza, la posizione del punto all'istante di tempo t può essere descritta dall'equazione:

$$\begin{cases} x_p = x'_p + v_{p\,x} t \\ y_p = y'_p + v_{p\,y} t \end{cases} \tag{A1}$$

in cui (x'_p, y'_p) è la posizione iniziale del punto, mentre $v_{p\,x}$ e $v_{p\,y}$ sono le componenti del vettore velocità sugli assi X e Y del sistema di riferimento. Allo

stesso modo, la posizione delle estremità del segmento $(x_{ln\,1}, y_{ln\,1})$ e $(x_{ln\,2}, y_{ln\,2})$ è descritta dall'equazione:

$$\begin{cases} x_{ln\,1} = x'_{ln\,1} + v_{ln\,x}t \\ y_{ln\,1} = y'_{ln\,1} + v_{ln\,y}t \end{cases} \text{ e } \begin{cases} x_{ln\,2} = x'_{ln\,2} + v_{ln\,x}t \\ y_{ln\,2} = y'_{ln\,2} + v_{ln\,y}t \end{cases} \tag{A2}$$

in cui $(x'_{ln\,1}, y'_{ln\,1})$ e $(x'_{ln\,2}, y'_{ln\,2})$ sono le posizioni iniziali delle estremità del segmento, mentre $v_{ln\,x}$ e $v_{ln\,y}$ sono le componenti del vettore velocità $\boldsymbol{v_{ln}}$ del segmento sugli assi X e Y.

L'equazione di una retta nella sua forma canonica, passante per il punto $(x_{ln\,1}, y_{ln\,1})$, è:

$$x - x_{ln\,1} = \frac{y - y_{ln\,1}}{k} \tag{A3}$$

con k il parametro che descrive la pendenza del segmento. Nel caso in cui il segmento si muova parallelamente a sé stesso, il coefficiente k rimane costante e, quindi, può essere ricavato dalle posizioni iniziali delle estremità del segmento come segue:

$$k = \frac{y'_{ln\,2} - y'_{ln\,1}}{x'_{ln\,2} - x'_{ln\,1}} \tag{A4}$$

Sostituendo le coordinate x e y generiche dell'Eq. (A3) con quelle del punto in Eq. (A1), il TTC può essere calcolato come:

$$TTC = -\frac{(y'_p - y'_{ln\,1}) - k(x'_p - x'_{ln\,1})}{(v_{p\,y} - v_{ln\,y}) - k(v_{p\,x} - v_{ln\,x})} \tag{A5}$$

oppure, nel caso in cui il denominatore dell'Eq. (A4) sia zero ($k \to \infty$):

$$TTC = \frac{x'_p - x'_{ln\,1}}{v_{p\,x} - v_{ln\,x}} \tag{A6}$$

ovviamente, solo i valori positivi sono rilevanti. La condizione affinché il punto intersechi il segmento è che nell'istante $t = TTC$:

$$\begin{cases} x_{ln\,1} \le x_p \le x_{ln\,2}, \text{ if } x_{ln\,2} \ge x_{ln\,1} \\ x_{ln\,2} \le x_p \le x_{ln\,1}, \text{ if } x_{ln\,1} \ge x_{ln\,2} \end{cases} \text{ e } \begin{cases} y_{ln\,1} \le y_p \le y_{ln\,2}, \text{ if } y_{ln\,2} \ge y_{ln\,1} \\ y_{ln\,2} \le y_p \le y_{ln\,1}, \text{ if } y_{ln\,1} \ge y_{ln\,2} \end{cases} \tag{A7}$$

Il punto di collisione (x_{coll}, y_{coll}) coincide con la posizione del punto (x_p, y_p) nell'istante $t = TTC$.

Calcolo del Time Advantage e T_2 per un punto ed un segmento

Per trovare la distanza temporale che intercorre tra un segmento ed un punto, è sufficiente calcolare le differenze nei tempi con cui il punto e le estremità del segmento passano i punti comuni 1 e 2 (Figura A1a). Per il punto 1, la differenza di tempo Δt_1 è:

$$\Delta t_1 = \left|t_{p\,1} - t_{ln\,1}\right| = \left|\frac{S_{p\,1}}{v_p} - \frac{S_{ln\,1}}{v_{ln}}\right| \tag{A8}$$

in cui $t_{p\,1}$ e $t_{ln\,1}$ sono i tempi impiegati dal punto e dall'estremità $ln\,1$ del segmento per raggiungere il punto 1; $S_{p\,1}$ e $S_{ln\,1}$ sono le distanze dal punto o estremità $ln\,1$ al punto comune 1; infine, v_p e v_{ln} sono i moduli dei vettori velocità associati al punto e al segmento, rispettivamente. Il calcolo della differenza Δt_2 si effettua allo stesso modo ma con riferimento al punto 2. Quindi, il T_{ADV} è il valore minore tra Δt_1 e Δt_2:

$$T_{ADV} = \min(\Delta t_1; \Delta t_2) = \min\left(\left|\frac{S_{p\,1}}{v_p} - \frac{S_{ln\,1}}{v_{ln}}\right|, \left|\frac{S_{p\,2}}{v_p} - \frac{S_{ln\,2}}{v_{ln}}\right|\right) \tag{A9}$$

Un caso particolare, in cui le traiettorie del punto e di una estremità del segmento non si intersecano è mostrato in Figura A1b. In questo esempio $t_{ln\,1} = 0$ e $\Delta t_1 = t_{p\,1}$. Allo stesso modo, bisogna considerare anche i casi in cui $t_{p\,1}$, $t_{p\,2}$ e $t_{ln\,2}$ sono uguali a zero.

Nel caso in cui i due utenti della strada superino il punto di conflitto con un margine di tempo, T_2 riflette il tempo massimo disponibile per intraprendere azioni evasive e alleviare la gravità della situazione. Per trovare il T_2, ancora una volta l'attenzione deve essere focalizzata sui punti 1 e 2. Il "secondo" utente della strada è quello che impiega un tempo più lungo per arrivare ad un punto comune. Per esempio, il tempo necessario al secondo utente per arrivare al punto 1 è:

$$T_{2-1} = \max\left(t_{p\,1}; t_{ln\,1}\right) = \max\left(\frac{S_{p\,1}}{v_p}; \frac{S_{ln\,1}}{v_{ln}}\right) \tag{A10}$$

Il tempo T_{2-2} necessario al secondo utente per arrivare al punto 2 è calcolato nello stesso modo. Dunque, il T_2 è il valore minore tra T_{2-1} e T_{2-2}:

$$T_2 = \min(T_{2-1}; T_{2-2}) = \min\left(\max\left(\frac{S_{p\,1}}{v_p}; \frac{S_{ln\,1}}{v_{ln}}\right), \max\left(\frac{S_{p\,2}}{v_p}; \frac{S_{ln\,2}}{v_{ln}}\right)\right) \tag{A11}$$

Ovviamente, devono essere considerati anche i casi particolari come quello mostrato in Figura A1b. Se gli utenti della strada sono in rotta di collisione, T_2 è uguale a TTC. T_2 è continuo, quindi può essere calcolato per qualsiasi istanza temporale, ma l'ultimo valore possibile per T_2 si ha quando il primo utente della strada lascia la zona di potenziale conflitto (in tal caso l'incidente non è più possibile).

Appendice B

Definizione delle metriche di valutazione dei modelli

In questa Sezione dell'Appendice B, le metriche di valutazione delle prestazioni dei modelli neurali sono definite matematicamente per agevolare la comprensione dei risultati.

Per i problemi di regressione, si consideri la serie temporale di T osservazioni di un partecipante alla sperimentazione, composta da $G = 3$ orizzonti temporali futuri. Inoltre, si assuma $y_g^{(t)}$ e $\hat{y}_g^{(t)}$ come rappresentativi delle osservazioni e previsioni di un partecipante per l'orizzonte temporale g-esimo, rispettivamente; mentre $\bar{y}_g = 1/T\left(\sum_t y_g^{(t)}\right)$ è usato per la media delle T osservazioni del medesimo partecipante per l'orizzonte temporale g-esimo.

Diversamente, per i problemi di classificazione binaria, i campioni sono annotati come positivi (in questo caso sono le classi di maggiore rischio, ovvero "*collision course*" per $\mathrm{GRU_{TADV}}$ e "*unsafe*" per $\mathrm{S\text{-}GRU_{SL}}$) e negativi (le classi "*crossing course*" e "*safe*") in modo che le metriche di valutazione possano essere calcolate attraverso la matrice di confusione, mostrata in Figura 22. Quest'ultima è una rappresentazione dell'accuratezza di classificazione del modello in quanto evidenzia le situazioni in cui il modello annota erroneamente una classe con un'altra: ogni riga della matrice di confusione rappresenta le istanze in una classe reale (ovvero la classe effettiva) mentre ogni colonna rappresenta quelle in una classe predetta. Così, il numero di "veri positivi" (VP - positivi correttamente classificati), "veri negativi" (VN - negativi correttamente classificati), "falsi positivi" (FP - negativi reali classificati in modo errato) e "falsi negativi" (FN - positivi reali classificati in modo errato) può essere definito.

Ovviamente, le equazioni qui riportate devono essere applicate a ciascun partecipante appartenente ad un *fold* di validazione e per ogni orizzonte temporale di previsione, e mediate al fine di ottenere la metrica complessiva di valutazione del modello.

Metriche di valutazione dei modelli di regressione

Il *Mean Absolute Error* (MAE), ovvero l'errore medio di predizione in valore assoluto, è definito nel seguente modo:

$$MAE := \frac{1}{T}\sum_{t=1}^{T}\left|\hat{y}_g^{(t)} - y_g^{(t)}\right| \tag{B1}$$

Il *Mean Absolute Percentage Error* ($MAPE$) è l'errore medio di predizione espresso come variazione percentuale rispetto all'osservazione; pertanto, è definito nel seguente modo:

$$MAPE := \frac{1}{T}\sum_{t=1}^{T}\left|100\frac{\left(\hat{y}_g^{(t)} - y_g^{(t)}\right)}{y_g^{(t)}}\right| \tag{B2}$$

Il *Root Mean Squared Error* ($RMSE$), ovvero la radice quadrata della differenza quadratica media tra i valori predetti e quelli osservati, è definito nel seguente modo:

$$RMSE := \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\hat{y}_g^{(t)} - y_g^{(t)}\right)^2} \tag{B3}$$

Il *Modified Index of Agreement* (md) è una metrica in grado di considerare contemporaneamente le differenze nelle medie e nelle varianze delle osservazioni e delle predizioni [236]; la sua definizione matematica è la seguente:

$$md := 1 - \left[\sum_{t=1}^{T}\left|\hat{y}_g^{(t)} - y_g^{(t)}\right| \Big/ \sum_{t=1}^{T}\left|\hat{y}_g^{(t)} - \bar{y}_g\right| + \left|y_g^{(t)} - \bar{y}_g\right|\right] \tag{B4}$$

Metriche di valutazione dei modelli di classificazione binaria

L'*Accuracy* è la frazione di campioni (positivi e negativi) classificati correttamente tra tutte le istanze (fare affidamento solo sull' *Accuracy* può essere fuorviante quando si lavora su un data set sbilanciato), e viene calcolato nel seguente modo:

$$Accuracy := \frac{VP + VN}{VP + FP + VN + FN} \tag{B5}$$

La *Precision* è la frazione di campioni classificati correttamente tra le istanze annotate come positive (un basso valore della *Precision* indica un numero consistente di istanze annotate erroneamente come positive), e viene calcolata nel seguente modo:

$$Precision := \frac{VP}{VP + FP} \tag{B6}$$

La *Recall*, detta anche *Sensitivity*, è la proporzione di campioni effettivamente positivi classificati correttamente (un basso valore della *Recall* indica un numero consistente di istanze positive annotate erroneamente in altre classi). Questa metrica è particolarmente rilevante nelle applicazioni FCW e AEB, siccome un alto valore di *Recall* è richiesto per evitare false attivazioni del sistema. Dalla definizione, l'espressione matematica è la seguente:

$$Recall := \frac{VP}{VP + FN} \tag{B7}$$

La *Specificity* è la proporzione di campioni effettivamente negativi classificati correttamente, e viene definita matematicamente nel seguente modo:

$$Specificity := \frac{VN}{VN + FP} \tag{B8}$$

Il *False Alarm Rate* (FAR) è la proporzione di campioni effettivamente negativi classificati erroneamente, ed è rappresentata dal complementare a 1 della *Specificity*:

$$FAR := 1 - Specificity = \frac{FP}{VN + FP} \tag{B9}$$

Infine, l'*Area Under the Curve* (AUC) è l'area sottesa dalla curva ROC la quale rappresenta la relazione tra il tasso di veri positivi (*Recall*) e quello di falsi positivi (FAR) per il classificatore binario al variare della soglia di classificazione. Infatti, i modelli di apprendimento per la classificazione binaria generano sempre un valore continuo per la previsione che rientra nell'intervallo $[0,1]$; l'ingegnere deve fissare una soglia di classificazione (un qualsiasi valore compreso nell'intervallo $[0,1]$, per esempio in questo lavoro di tesi uguale a 0.5), al pari di una soglia di probabilità, per annotare tutti i valori uguali o superiori alla soglia come 1, ed i restanti (ovvero i valori al di sotto della soglia selezionata) come 0. Quindi, i punti sulla curva ROC rappresentano la *Recall* ed il FAR per diversi valori di questa soglia. Correlato all'AUC, l'aspetto rilevante ed interessante nella valutazione delle prestazioni di un modello di classificazione binaria è l'andamento della curva ROC: in generale, più questa curva si avvicina all'angolo superiore sinistro del grafico (considerando uno spazio cartesiano delimitato tra 0 ed 1 su entrambi gli assi), migliore è la qualità della previsione offerta dal classificatore (e maggiore ovviamente risulterà l'AUC).

Risultati della ricerca a griglia

***Tabella B1: Risultati della ricerca a griglia per il modello* GRU_{T2}.**

Iperparametri			**Valori medi (deviazione standard) sui *folds* di validazione**			
N	α	δ	$MAE\ [s]$	$MAPE\ [\%]$	$RMSE\ [s]$	$md\ [-]$
64	0.050	0.40	0.371 (0.113)	7.846 (2.436)	0.547 (0.147)	0.882 (0.047)
64	0.050	0.60	0.389 (0.119)	8.048 (2.459)	0.588 (0.187)	0.878 (0.048)
64	0.050	0.80	0.377 (0.090)	7.945 (1.842)	0.556 (0.133)	0.882 (0.039)
64	0.010	0.40	0.431 (0.140)	9.156 (3.189)	0.606 (0.175)	0.860 (0.063)
64	0.010	0.60	0.352 (0.104)	7.386 (2.413)	0.521 (0.134)	0.888 (0.044)
64	0.010	0.80	0.332 (0.092)	6.997 (2.182)	0.492 (0.125)	0.895 (0.042)
64	0.005	0.40	0.529 (0.150)	11.847 (3.589)	0.735 (0.178)	0.828 (0.072)
64	0.005	0.60	0.406 (0.127)	8.679 (3.047)	0.576 (0.154)	0.868 (0.060)
64	**0.005**	**0.80**	**0.318 (0.086)**	**6.735 (1.949)**	**0.460 (0.117)**	**0.899 (0.039)**
128	0.050	0.40	0.748 (0.261)	15.318 (4.579)	1.004 (0.350)	0.801 (0.077)
128	0.050	0.60	0.821 (0.217)	18.295 (5.199)	1.087 (0.278)	0.732 (0.099)
128	0.050	0.80	0.994 (0.444)	22.695 (10.751)	1.283 (0.525)	0.658 (0.178)
128	0.010	0.40	0.419 (0.137)	8.854 (3.123)	0.602 (0.181)	0.865 (0.060)
128	0.010	0.60	0.376 (0.136)	7.780 (2.915)	0.558 (0.182)	0.881 (0.053)
128	0.010	0.80	0.351 (0.123)	7.385 (2.632)	0.521 (0.169)	0.889 (0.048)
128	0.005	0.40	0.468 (0.131)	10.188 (3.107)	0.649 (0.167)	0.848 (0.063)
128	0.005	0.60	0.367 (0.115)	7.629 (2.583)	0.534 (0.151)	0.883 (0.049)
128	0.005	0.80	0.328 (0.104)	6.923 (2.369)	0.489 (0.133)	0.896 (0.044)
256	0.050	0.40	0.823 (0.469)	17.981 (10.301)	1.091 (0.547)	0.730 (0.153)
256	0.050	0.60	0.619 (0.260)	13.464 (5.831)	0.836 (0.311)	0.795 (0.113)
256	0.050	0.80	0.592 (0.216)	12.471 (3.847)	0.808 (0.270)	0.814 (0.068)
256	0.010	0.40	0.388 (0.125)	8.129 (2.739)	0.563 (0.165)	0.875 (0.056)
256	0.010	0.60	0.341 (0.109)	7.093 (2.463)	0.508 (0.144)	0.892 (0.045)
256	0.010	0.80	0.327 (0.093)	6.892 (2.033)	0.485 (0.129)	0.897 (0.040)
256	0.005	0.40	0.438 (0.149)	9.291 (3.487)	0.617 (0.187)	0.858 (0.065)
256	0.005	0.60	0.391 (0.134)	8.229 (3.104)	0.558 (0.171)	0.875 (0.057)
256	0.005	0.80	0.370 (0.139)	7.716 (3.162)	0.535 (0.175)	0.882 (0.054)

***Tabella B2: Risultati della ricerca a griglia per il modello* GRU$_{TADV}$.**

Iperparametri			**Valori medi (deviazione standard) sui *folds* di validazione**					
N	α	δ	*Accuracy*	*Precision*	*Recall*	*Specificity*	FAR	AUC
64	0.010	0.40	0.952 (0.014)	0.863 (0.050)	0.822 (0.064)	0.976 (0.009)	0.024 (0.009)	0.983 (0.012)
64	0.010	0.60	0.962 (0.011)	0.884 (0.035)	0.868 (0.056)	0.979 (0.006)	0.021 (0.006)	0.991 (0.006)
64	0.010	0.80	0.963 (0.012)	0.885 (0.039)	0.871 (0.054)	0.979 (0.007)	0.021 (0.007)	0.991 (0.005)
64	0.005	0.40	0.929 (0.021)	0.786 (0.079)	0.741 (0.075)	0.963 (0.014)	0.037 (0.014)	0.967 (0.018)
64	0.005	0.60	0.947 (0.014)	0.845 (0.056)	0.807 (0.067)	0.973 (0.010)	0.027 (0.010)	0.981 (0.012)
64	0.005	0.80	0.961 (0.013)	0.873 (0.045)	0.870 (0.061)	0.977 (0.008)	0.023 (0.008)	0.990 (0.007)
64	0.001	0.40	0.908 (0.019)	0.681 (0.059)	0.757 (0.118)	0.935 (0.016)	0.065 (0.016)	0.949 (0.014)
64	0.001	0.60	0.915 (0.015)	0.732 (0.054)	0.714 (0.049)	0.952 (0.012)	0.048 (0.012)	0.955 (0.011)
64	0.001	0.80	0.921 (0.021)	0.760 (0.081)	0.722 (0.090)	0.957 (0.021)	0.043 (0.021)	0.960 (0.019)
128	0.010	0.40	0.962 (0.010)	0.878 (0.037)	0.874 (0.049)	0.978 (0.007)	0.022 (0.007)	0.991 (0.005)
128	0.010	0.60	0.962 (0.011)	0.873 (0.041)	0.880 (0.056)	0.977 (0.008)	0.023 (0.008)	0.991 (0.006)
128	0.010	0.80	0.961 (0.010)	0.880 (0.039)	0.864 (0.049)	0.978 (0.007)	0.022 (0.007)	0.990 (0.006)
128	0.005	0.40	0.961 (0.010)	0.876 (0.036)	0.865 (0.047)	0.978 (0.006)	0.022 (0.006)	0.990 (0.007)
128	0.005	0.60	0.963 (0.008)	0.878 (0.037)	0.885 (0.037)	0.977 (0.007)	0.023 (0.007)	0.991 (0.004)
128	0.005	0.80	0.963 (0.009)	0.873 (0.041)	0.888 (0.043)	0.976 (0.008)	0.024 (0.008)	0.992 (0.004)
128	0.001	0.40	0.909 (0.030)	0.746 (0.080)	0.615 (0.227)	0.962 (0.017)	0.038 (0.017)	0.950 (0.025)
128	0.001	0.60	0.920 (0.023)	0.743 (0.096)	0.749 (0.059)	0.951 (0.021)	0.049 (0.021)	0.959 (0.023)
128	0.001	0.80	0.948 (0.018)	0.834 (0.065)	0.822 (0.084)	0.970 (0.011)	0.030 (0.011)	0.980 (0.017)
256	0.010	0.40	0.955 (0.014)	0.864 (0.043)	0.843 (0.066)	0.976 (0.007)	0.024 (0.007)	0.988 (0.008)
256	0.010	0.60	0.953 (0.013)	0.867 (0.039)	0.823 (0.069)	0.977 (0.007)	0.023 (0.007)	0.988 (0.007)
256	0.010	0.80	0.958 (0.012)	0.873 (0.047)	0.852 (0.065)	0.977 (0.009)	0.023 (0.009)	0.990 (0.005)
256	0.005	0.40	0.960 (0.010)	0.878 (0.039)	0.858 (0.060)	0.978 (0.008)	0.022 (0.008)	0.990 (0.006)
256	0.005	0.60	0.962 (0.010)	0.877 (0.033)	0.875 (0.057)	0.978 (0.006)	0.022 (0.006)	0.991 (0.005)
256	0.005	0.80	0.961 (0.009)	0.863 (0.044)	0.888 (0.034)	0.974 (0.009)	0.026 (0.009)	0.990 (0.005)
256	0.001	0.40	0.925 (0.021)	0.766 (0.077)	0.743 (0.068)	0.959 (0.014)	0.041 (0.014)	0.963 (0.020)
256	0.001	0.60	0.956 (0.010)	0.862 (0.033)	0.848 (0.045)	0.975 (0.005)	0.025 (0.005)	0.986 (0.008)
256	**0.001**	**0.80**	**0.964 (0.007)**	**0.881 (0.029)**	**0.886 (0.046)**	**0.978 (0.006)**	**0.022 (0.006)**	**0.992 (0.004)**

***Tabella B3: Risultati della ricerca a griglia per il modello* S-GRU$_{SL}$.**

Iperparametri			**Valori medi (deviazione standard) sui *folds* di validazione**					
N	α	δ	*Accuracy*	*Precision*	*Recall*	*Specificity*	FAR	AUC
64	0.050	0.40	0.939 (0.017)	0.707 (0.104)	0.704 (0.124)	0.966 (0.011)	0.034 (0.011)	0.969 (0.017)
64	0.050	0.60	0.940 (0.017)	0.714 (0.106)	0.721 (0.114)	0.965 (0.014)	0.035 (0.014)	0.969 (0.016)
64	0.050	0.80	0.931 (0.014)	0.699 (0.118)	0.631 (0.181)	0.966 (0.018)	0.034 (0.018)	0.962 (0.021)
64	0.010	0.40	0.930 (0.018)	0.694 (0.113)	0.581 (0.175)	0.971 (0.009)	0.029 (0.009)	0.957 (0.024)
64	0.010	0.60	0.948 (0.015)	0.763 (0.092)	0.734 (0.138)	0.973 (0.010)	0.027 (0.010)	0.977 (0.016)
64	0.010	0.80	0.953 (0.013)	0.788 (0.084)	0.755 (0.101)	0.976 (0.009)	0.024 (0.009)	0.982 (0.010)
64	0.005	0.40	0.920 (0.016)	0.635 (0.109)	0.515 (0.173)	0.967 (0.005)	0.033 (0.005)	0.947 (0.024)
64	0.005	0.60	0.924 (0.016)	0.646 (0.109)	0.577 (0.167)	0.964 (0.008)	0.036 (0.008)	0.952 (0.023)
64	0.005	0.80	0.952 (0.014)	0.781 (0.090)	0.756 (0.108)	0.975 (0.010)	0.025 (0.010)	0.980 (0.012)
128	0.050	0.40	0.929 (0.018)	0.673 (0.088)	0.627 (0.148)	0.964 (0.009)	0.036 (0.009)	0.961 (0.019)
128	0.050	0.60	0.929 (0.017)	0.666 (0.086)	0.662 (0.13)	0.961 (0.012)	0.039 (0.012)	0.960 (0.020)
128	0.050	0.80	0.925 (0.011)	0.639 (0.096)	0.630 (0.151)	0.959 (0.011)	0.042 (0.011)	0.954 (0.015)
128	0.010	0.40	0.951 (0.013)	0.771 (0.084)	0.762 (0.085)	0.973 (0.010)	0.027 (0.010)	0.980 (0.010)
128	0.010	0.60	0.951 (0.012)	0.775 (0.082)	0.759 (0.071)	0.973 (0.011)	0.027 (0.011)	0.982 (0.009)
128	0.010	0.80	0.954 (0.011)	0.799 (0.067)	0.750 (0.090)	0.977 (0.008)	0.023 (0.008)	0.983 (0.008)
128	0.005	0.40	0.939 (0.018)	0.729 (0.115)	0.652 (0.166)	0.972 (0.008)	0.028 (0.008)	0.967 (0.021)
128	0.005	0.60	0.953 (0.014)	0.779 (0.084)	0.773 (0.098)	0.974 (0.010)	0.026 (0.010)	0.982 (0.011)
128	**0.005**	**0.80**	**0.962 (0.010)**	**0.833 (0.065)**	**0.802 (0.074)**	**0.980 (0.009)**	**0.020 (0.009)**	**0.987 (0.007)**
256	0.050	0.40	0.921 (0.015)	0.644 (0.118)	0.520 (0.200)	0.967 (0.014)	0.033 (0.014)	0.949 (0.022)
256	0.050	0.60	0.920 (0.019)	0.646 (0.136)	0.546 (0.218)	0.964 (0.026)	0.036 (0.026)	0.947 (0.025)
256	0.050	0.80	0.922 (0.019)	0.668 (0.131)	0.549 (0.168)	0.965 (0.017)	0.035 (0.017)	0.947 (0.028)
256	0.010	0.40	0.950 (0.013)	0.766 (0.087)	0.760 (0.084)	0.972 (0.011)	0.028 (0.011)	0.980 (0.010)
256	0.010	0.60	0.951 (0.015)	0.771 (0.081)	0.772 (0.076)	0.972 (0.011)	0.028 (0.011)	0.981 (0.010)
256	0.010	0.80	0.952 (0.012)	0.791 (0.085)	0.752 (0.098)	0.976 (0.011)	0.024 (0.011)	0.982 (0.009)
256	0.005	0.40	0.952 (0.014)	0.773 (0.084)	0.771 (0.107)	0.973 (0.009)	0.027 (0.009)	0.982 (0.011)
256	0.005	0.60	0.958 (0.011)	0.807 (0.064)	0.794 (0.076)	0.977 (0.009)	0.023 (0.009)	0.985 (0.007)
256	0.005	0.80	0.957 (0.009)	0.811 (0.055)	0.779 (0.067)	0.978 (0.008)	0.022 (0.008)	0.985 (0.007)

Bibliografia

[1] E. M. Murtagh, L. Nichols, M. A. Mohammed, R. Holder, A. M. Nevill, e M. H. Murphy, «Walking to improve cardiovascular health: a meta-analysis of randomised control trials», *The Lancet*, vol. 384 S2, ID S54, nov. 2014, doi: 10.1016/S0140-6736(14)62180-2.

[2] Z. Song, L. Cao, e C. C. Chou, «Development of Test Equipment for Pedestrian-Automatic Emergency Braking Based on C-NCAP (2018)», *Sensors*, vol. 20, n. 21, ID 6206, ott. 2020, doi: 10.3390/S20216206.

[3] Y. Rydin *et al.*, «Shaping cities for health: complexity and the planning of urban environments in the 21st century», *The Lancet*, vol. 379, n. 9831, pagg. 2079–2108, giu. 2012, doi: 10.1016/S0140-6736(12)60435-8.

[4] L. Domenichini, V. Branzi, e M. Smorti, «Influence of drivers' psychological risk profiles on the effectiveness of traffic calming measures», *Accident Analysis & Prevention*, vol. 123, pagg. 243–255, feb. 2019, doi: 10.1016/J.AAP.2018.11.025.

[5] U. Ekelund *et al.*, «Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC)», *The American Journal of Clinical Nutrition*, vol. 101, n. 3, pagg. 613–621, mar. 2015, doi: 10.3945/AJCN.114.100065.

[6] Organisation for Economic Co-operation and Development, «Road Safety Annual Report 2017», OECD, Parigi, Francia, ott. 2017. doi: 10.1787/irtad-2017-en.

[7] World Health Organization, «Global Status Report on Road Safety 2018», WHO, Ginevra, Svizzera, giu. 2018.

[8] European Road Safety Observatory, «Traffic Safety Basic Facts 2018 - Pedestrians», European Commission, Bruxelles, Belgio, apr. 2018.

[9] World Health Organization, «WHO | Death on the roads», 2018. https://extranet.who.int/roadsafety/death-on-the-roads/#ticker (consultato ott. 19, 2021).

[10] National Highway Traffic Safety Administration, «Traffic safety Facts - 2016 Data - Pedestrians», NHTSA, Washington, DC, USA, mar. 2018.

[11] R. Johansson, «Vision Zero – Implementing a policy for traffic safety», *Safety Science*, vol. 47, n. 6, pagg. 826–831, lug. 2009, doi: 10.1016/J.SSCI.2008.10.023.

[12] European Commission, «EU Road Safety Policy Framework 2021-2030 - Next steps towards Vision Zero», EC, Strasburgo, Francia, ott. 2021.

[13] Istituto Nazionale di Statistica, «Incidenti Stradali - Anno 2019», ISTAT, Roma, Italia, lug. 2020.

[14] Y. Zheng, T. Chase, L. Elefteriadou, B. Schroeder, e V. P. Sisiopiku, «Modeling vehicle–pedestrian interactions outside of crosswalks», *Simulation Modelling Practice and Theory*, vol. 59, pagg. 89–101, dic. 2015, doi: 10.1016/J.SIMPAT.2015.08.005.

[15] F. Bella e F. Nobili, «Driver-pedestrian interaction under legal and illegal pedestrian crossings», *Transportation Research Procedia*, vol. 45, pagg. 451–458, mar. 2020, doi: 10.1016/J.TRPRO.2020.03.038.

[16] Associazione Sostenitori ed Amici della Polizia Stradale, «3° Rapporto Annuale sull'Incidentalità con il Coinvolgimento dei Pedoni - Anno 2019 - Analisi Dati ACI-ISTAT - A cura Ufficio Studi ASAPS», set. 2020. https://www.asaps.it/70913-_3_rapporto_annuale_sull_incidentalita__con_il_coinvolgimento_dei_pedoni_anno_20.html (consultato ott. 21, 2021).

[17] D. A. Quistberg *et al.*, «Multilevel models for evaluating the risk of pedestrian–motor vehicle collisions at intersections and mid-blocks», *Accident Analysis & Prevention*, vol. 84, pagg. 99–111, nov. 2015, doi: 10.1016/J.AAP.2015.08.013.

[18] T. Koepsell *et al.*, «Crosswalk Markings and the Risk of Pedestrian–Motor Vehicle Collisions in Older Pedestrians», *JAMA*, vol. 288, n. 17, pagg. 2136–2143, nov. 2002, doi: 10.1001/JAMA.288.17.2136.

[19] F. L. Mannering e C. R. Bhat, «Analytic methods in accident research: Methodological frontier and future directions», *Analytic Methods in Accident Research*, vol. 1, pagg. 1–22, gen. 2014, doi: 10.1016/J.AMAR.2013.09.001.

[20] C. Lee e M. Abdel-Aty, «Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida», *Accident Analysis & Prevention*, vol. 37, n. 4, pagg. 775–786, lug. 2005, doi: 10.1016/J.AAP.2005.03.019.

[21] M. Wier, J. Weintraub, E. H. Humphreys, E. Seto, e R. Bhatia, «An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning», *Accident Analysis & Prevention*, vol. 41, n. 1, pagg. 137–145, gen. 2009, doi: 10.1016/J.AAP.2008.10.001.

[22] L. Chen, C. Chen, R. Ewing, C. E. McKnight, R. Srinivasan, e M. Roe, «Safety countermeasures and crash reduction in New York City—Experience and lessons learned», *Accident Analysis & Prevention*, vol. 50, pagg. 312–322, gen. 2013, doi: 10.1016/J.AAP.2012.05.009.

[23] S. Yao, B. P. Y. Loo, e W. W. Y. Lam, «Measures of activity-based pedestrian exposure to the risk of vehicle-pedestrian collisions: Space-time path vs. potential path tree methods», *Accident Analysis & Prevention*, vol. 75, pagg. 320–332, feb. 2015, doi: 10.1016/J.AAP.2014.12.005.

[24] X. Wang, J. Yang, C. Lee, Z. Ji, e S. You, «Macro-level safety analysis of pedestrian crashes in Shanghai, China», *Accident Analysis & Prevention*, vol. 96, pagg. 12–21, nov. 2016, doi: 10.1016/J.AAP.2016.07.028.

[25] S. Eisses, «ITS Action Plan - Framework service contract TREN/G4/FV-2008/475/01», EC, Amsterdam, Paesi Bassi, mag. 2011.

[26] L. F. Miranda-Moreno, P. Morency, e A. M. El-Geneidy, «The link between built environment, pedestrian activity and pedestrian–vehicle collision occurrence at signalized intersections», *Accident Analysis & Prevention*, vol. 43, n. 5, pagg. 1624–1634, set. 2011, doi: 10.1016/J.AAP.2011.02.005.

[27] S. Ukkusuri, L. F. Miranda-Moreno, G. Ramadurai, e J. Isa-Tavarez, «The role of built environment on pedestrian crash frequency», *Safety Science*, vol. 50, n. 4, pagg. 1141–1151, apr. 2012, doi: 10.1016/J.SSCI.2011.09.012.

[28] C. Dong, S. H. Richards, D. B. Clarke, X. Zhou, e Z. Ma, «Examining signalized intersection crash frequency using multivariate zero-inflated Poisson regression», *Safety Science*, vol. 70, pagg. 63–69, dic. 2014, doi: 10.1016/J.SSCI.2014.05.006.

[29] J. P. Schorr e S. H. Hamdar, «Safety propensity index for signalized and unsignalized intersections: Exploration and assessment», *Accident Analysis & Prevention*, vol. 71, pagg. 93–105, ott. 2014, doi: 10.1016/J.AAP.2014.05.008.

[30] E. Du Yingzi *et al.*, «Pedestrian Behavior Analysis Using Naturalistic Driving Data in USA», negli Atti della *23rd International Technical Conference on the Enhanced Safety of Vehicles (ESV): Research Collaboration to Benefit Safety of All Road Users*, Seoul, Corea del Sud, mag. 2013, ID 13–0291.

[31] F. Bella, V. Borrelli, M. Silvestri, e F. Nobili, «Effects on Driver's Behavior of Illegal Pedestrian Crossings», negli Atti della *International Conference on Applied Human Factors and Ergonomics (AHFE 2018): Advances in Intelligent Systems and Computing*, Orlando, FL, USA, lug. 2018, vol. 786, pagg. 802–812, doi: 10.1007/978-3-319-93885-1_74.

[32] J. Reason, A. Manstead, S. Stradling, J. Baxter, e K. Campbell, «Errors and violations on the roads: a real distinction?», *Ergonomics*, vol. 33, n. 10–11, pagg. 1315–1332, ott. 1990, doi: 10.1080/00140139008925335.

[33] M. Smorti e S. Guarnieri, «Exploring the factor structure and psychometric properties of the Manchester Driver Behavior Questionnaire (DBQ) in an Italian sample», *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, vol. 23, n. 2, pagg. 185–202, giu. 2016, doi: 10.4473/TPM23.2.4.

[34] J. Xu, Y. Ge, W. Qu, X. Sun, e K. Zhang, «The mediating effect of traffic safety climate between pedestrian inconvenience and pedestrian behavior», *Accident Analysis & Prevention*, vol. 119, pagg. 155–161, ott. 2018, doi: 10.1016/J.AAP.2018.07.020.

[35] A. Sheykhfard, F. Haghighi, T. Nordfjærn, e M. Soltaninejad, «Structural equation modelling of potential risk factors for pedestrian accidents in rural and urban roads», *International Journal of Injury Control and Safety Promotion*, vol. 28, n. 1, pagg. 46–57, 2020, doi: 10.1080/17457300.2020.1835991.

[36] A. Sheykhfard, F. Haghighi, E. Papadimitriou, e P. Van Gelder, «Analysis of the occurrence and severity of vehicle-pedestrian conflicts in marked and unmarked crosswalks through naturalistic driving study», *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 76, pagg. 178–192, gen. 2021, doi: 10.1016/J.TRF.2020.11.008.

[37] World Health Organization, «World report on road traffic injury prevention», WHO, Ginevra, Svizzera, feb. 2004.

[38] S. El Hamdani, N. Benamar, e M. Younis, «Pedestrian Support in Intelligent Transportation Systems: Challenges, Solutions and Open issues», *Transportation Research Part C: Emerging Technologies*, vol. 121, ID 102856, dic. 2020, doi: 10.1016/J.TRC.2020.102856.

[39] S. A. Bagloee, M. Tavana, M. Asadi, e T. Oliver, «Autonomous vehicles: challenges, opportunities, and future implications for transportation policies», *Journal of Modern Transportation*, vol. 24, n. 4, pagg. 284–303, ago. 2016, doi: 10.1007/S40534-016-0117-3.

[40] R. McAllister *et al.*, «Concrete problems for autonomous vehicle safety: Advantages of Bayesian deep learning», negli Atti della *26th International Joint Conference on Artificial Intelligence*, Melbourne, Australia, ago. 2017, pagg. 4745–4753, doi: 10.24963/IJCAI.2017/661.

[41] P. Koopman e M. Wagner, «Challenges in Autonomous Vehicle Testing and Validation», *SAE International Journal of Transportation Safety*, vol. 4, n. 1, pagg. 15–24, apr. 2016, doi: 10.4271/2016-01-0128.

[42] A. Sarmento, B. Garcia, L. Coriteac, e L. Navarenho, «The Autonomous Vehicle Challenges for Emergent Market», *SAE Technical Papers*, vol. 2017-November, n. November, nov. 2017, doi: 10.4271/2017-36-0436.

[43] F. M. Favarò, N. Nader, S. O. Eurich, M. Tripp, e N. Varadaraju, «Examining accident reports involving autonomous vehicles in California», *PLOS ONE*, vol. 12, n. 9, ID e0184952, set. 2017, doi: 10.1371/JOURNAL.PONE.0184952.

[44] D. J. Fagnant e K. Kockelman, «Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations», *Transportation Research Part A: Policy and Practice*, vol. 77, pagg. 167–181, lug. 2015, doi: 10.1016/J.TRA.2015.04.003.

[45] J. Li, X. Zhao, M. J. Cho, W. Ju, e B. F. Malle, «From Trolley to Autonomous Vehicle: Perceptions of Responsibility and Moral Norms in Traffic Accidents with Self-Driving Cars», *SAE Technical Papers*, vol. 2016-April, n. April, apr. 2016, doi: 10.4271/2016-01-0164.

[46] J. C. F. de Winter e D. Dodou, «The Driver Behaviour Questionnaire as a predictor of accidents: A meta-analysis», *Journal of Safety Research*, vol. 41, n. 6, pagg. 463–470, dic. 2010, doi: 10.1016/j.jsr.2010.10.007.

[47] A. af Wåhlberg, L. Dorn, e T. Kline, «The Manchester Driver Behaviour Questionnaire as a predictor of road traffic accidents», *Theoretical Issues in Ergonomics Science*, vol. 12, n. 1, pagg. 66–86, gen. 2011, doi: 10.1080/14639220903023376.

[48] A. Sheykhfard, F. Haghighi, E. Papadimitriou, e P. Van Gelder, «Review and assessment of different perspectives of vehicle-pedestrian conflicts and crashes: Passive and active analysis approaches», *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 8, n. 5, pagg. 681–702, ott. 2021, doi: 10.1016/j.jtte.2021.08.001.

[49] H. Iversen e T. Rundmo, «Personality, risky driving and accident involvement among Norwegian drivers», *Personality and Individual Differences*, vol. 33, n. 8, pagg. 1251–1263, dic. 2002, doi: 10.1016/S0191-8869(02)00010-7.

[50] D. C. Schwebel, J. Severson, K. K. Ball, e M. Rizzo, «Individual difference factors in risky driving: The roles of anger/hostility, conscientiousness, and sensation-seeking», *Accident Analysis & Prevention*, vol. 38, n. 4, pagg. 801–810, lug. 2006, doi: 10.1016/j.aap.2006.02.004.

[51] T. Özkan e T. Lajunen, «What causes the differences in driving between young men and women? The effects of gender roles and sex on young drivers' driving behaviour and self-assessment of skills», *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 9, n. 4, pagg. 269–277, lug. 2006, doi: 10.1016/j.trf.2006.01.005.

[52] U.K. Department for Transport, «Older Car Drivers Road Safety Factsheet (2016)», U.K. DfT, London, Regno Unito, mag. 2018.

[53] National Highway Traffic Safety Administration, «Driver Electronic Device Use in 2018», NHTSA, Washington, DC, USA., ott. 2019.

[54] European Road Safety Observatory, «Cell phone use while driving - 2018», EC, Bruxelles, Belgio, feb. 2018.

[55] J. Hatfield e S. Murphy, «The effects of mobile phone use on pedestrian crossing behaviour at signalised and unsignalised intersections», *Accident Analysis & Prevention*, vol. 39, n. 1, pagg. 197–205, gen. 2007, doi: 10.1016/J.AAP.2006.07.001.

[56] W. J. Horrey e C. D. Wickens, «Examining the Impact of Cell Phone Conversations on Driving Using Meta-Analytic Techniques», *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 48, n. 1, pagg. 196–205, mar. 2006, doi: 10.1518/001872006776412135.

[57] J. K. Caird, C. R. Willness, P. Steel, e C. Scialfa, «A meta-analysis of the effects of cell phones on driver performance», *Accident Analysis & Prevention*, vol. 40, n. 4, pagg. 1282–1293, lug. 2008, doi: 10.1016/J.AAP.2008.01.009.

[58] M. M. Haque e S. Washington, «A parametric duration model of the reaction times of drivers distracted by mobile phone conversations», *Accident Analysis & Prevention*, vol. 62, pagg. 42–53, gen. 2014, doi: 10.1016/J.AAP.2013.09.010.

[59] N. Baldo, A. Marini, e M. Miani, «Drivers' braking behavior affected by cognitive distractions: An experimental investigation with a virtual car simulator», *Behavioral Sciences*, vol. 10, n. 10, ID 150, ott. 2020, doi: 10.3390/BS10100150.

[60] P. C. Burns, A. M. Parkes, S. Burton, R. K. Smith, e D. Burch, «How dangerous is driving with a mobile phone? Benchmarking the impairment to alcohol», TRL, Wokingham, Regno Unito, Report TRL547, gen. 2002.

[61] J. D. Lee, K. L. Young, e M. A. Regan, «Defining driver distraction», in *Driver Distraction: Theory, Effects, and Mitigation*. Boca Raton, FL, USA: CRC Press - Taylor & Francis Group, oct. 2008, pagg. 31–40.

[62] M. A. Just, T. A. Keller, e J. Cynkar, «A decrease in brain activation associated with driving when listening to someone speak», *Brain Research*, vol. 1205, pagg. 70–80, apr. 2008, doi: 10.1016/J.BRAINRES.2007.12.075.

[63] D. L. Strayer, F. A. Drews, e W. A. Johnston, «Cell Phone-Induced Failures of Visual Attention During Simulated Driving», *Journal of Experimental Psychology: Applied*, vol. 9, n. 1, pagg. 23–32, mar. 2003, doi: 10.1037/1076-898X.9.1.23.

[64] National Safety Council, «Understanding the distracted brain - Why driving while using

hands-free cell phones is risky Behavior?», NSC, Itasca, IL, USA, apr. 2012.

[65] N. Lubbe, «Brake reactions of distracted drivers to pedestrian Forward Collision Warning systems», *Journal of Safety Research*, vol. 61, pagg. 23–32, giu. 2017, doi: 10.1016/J.JSR.2017.02.002.

[66] A. Habibovic, E. Tivesten, N. Uchida, J. Bärgman, e M. Ljung Aust, «Driver behavior in car-to-pedestrian incidents: An application of the Driving Reliability and Error Analysis Method (DREAM)», *Accident Analysis & Prevention*, vol. 50, pagg. 554–565, gen. 2013, doi: 10.1016/J.AAP.2012.05.034.

[67] E. Papadimitriou, A. Theofilatos, e G. Yannis, «Patterns of pedestrian attitudes, perceptions and behaviour in Europe», *Safety Science*, vol. 53, pagg. 114–122, mar. 2013, doi: 10.1016/J.SSCI.2012.09.008.

[68] E. Papadimitriou, S. Lassarre, e G. Yannis, «Introducing human factors in pedestrian crossing behaviour models», *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 36, pagg. 69–82, gen. 2016, doi: 10.1016/J.TRF.2015.11.003.

[69] J. Wu e H. Xu, «Driver behavior analysis for right-turn drivers at signalized intersections using SHRP 2 naturalistic driving study data», *Journal of Safety Research*, vol. 63, pagg. 177–185, dic. 2017, doi: 10.1016/J.JSR.2017.10.010.

[70] S. Bakhtiari *et al.*, «Effect of Visual and Auditory Alerts on Older Drivers' Glances toward Latent Hazards while Turning Left at Intersections»:, *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, n. 9, pagg. 117–126, mag. 2019, doi: 10.1177/0361198119844244.

[71] G. J. S. Wilde, «Beyond the concept of risk homeostatis: Suggestions for research and application towards the prevention of accidents and lifestyle-related disease», *Accident Analysis & Prevention*, vol. 18, n. 5, pagg. 377–401, ott. 1986, doi: 10.1016/0001-4575(86)90012-6.

[72] G. J. S. Wilde, «Risk homeostasis theory and traffic accidents: propositions, deductions and discussion of dissension in recent reactions», *Ergonomics*, vol. 31, n. 4, pagg. 441–468, 1988, doi: 10.1080/00140138808966691.

[73] H. Yoshitake e M. Shino, «Risk assessment based on driving behavior for preventing collisions with pedestrians when making across-traffic turns at intersections», *IATSS Research*, vol. 42, n. 4, pagg. 240–247, dic. 2018, doi: 10.1016/J.IATSSR.2018.02.001.

[74] M. Van Der Hulst, T. Meijman, e T. Rothengatter, «Anticipation and the adaptive control of safety margins in driving», *Ergonomics*, vol. 42, n. 2, pagg. 336–345, feb. 1999, doi: 10.1080/001401399185694.

[75] E. Rosén, J. E. Källhammer, D. Eriksson, M. Nentwich, R. Fredriksson, e K. Smith, «Pedestrian injury mitigation by autonomous braking», *Accident Analysis & Prevention*, vol. 42, n. 6, pagg. 1949–1957, nov. 2010, doi: 10.1016/J.AAP.2010.05.018.

[76] A. Laureshyn, Å. Svensson, e C. Hydén, «Evaluation of traffic safety, based on micro-level behavioural data: Theoretical framework and first implementation», *Accident Analysis & Prevention*, vol. 42, n. 6, pagg. 1637–1646, nov. 2010, doi: 10.1016/J.AAP.2010.03.021.

[77] Å. Svensson e C. Hydén, «Estimating the severity of safety related behaviour», *Accident Analysis & Prevention*, vol. 38, n. 2, pagg. 379–385, mar. 2006, doi: 10.1016/J.AAP.2005.10.009.

[78] J. Wu, H. Xu, Y. Zhang, e R. Sun, «An improved vehicle-pedestrian near-crash identification method with a roadside LiDAR sensor», *Journal of Safety Research*, vol. 73, pagg. 211–224, giu. 2020, doi: 10.1016/J.JSR.2020.03.006.

[79] G. A. Davis, «Possible aggregation biases in road safety research and a mechanism approach to accident modeling», *Accident Analysis & Prevention*, vol. 36, n. 6, pagg. 1119–1127, nov. 2004, doi: 10.1016/J.AAP.2004.04.002.

[80] R. Elvik, «Operational Criteria of Causality for Observational Road Safety Evaluation

Studies», *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2019, n. 1, pagg. 74–81, gen. 2007, doi: 10.3141/2019-10.

[81] E. Hauer, «Cause, effect and regression in road safety: A case study», *Accident Analysis & Prevention*, vol. 42, n. 4, pagg. 1128–1135, lug. 2010, doi: 10.1016/J.AAP.2009.12.027.

[82] A. P. Tarko, «Use of crash surrogates and exceedance statistics to estimate road safety», *Accident Analysis & Prevention*, vol. 45, pagg. 230–240, mar. 2012, doi: 10.1016/J.AAP.2011.07.008.

[83] A. Kathuria e P. Vedagiri, «Evaluating pedestrian vehicle interaction dynamics at un-signalized intersections: A proactive approach for safety analysis», *Accident Analysis & Prevention*, vol. 134, ID 105316, gen. 2020, doi: 10.1016/J.AAP.2019.105316.

[84] C. Hydén, «The Development of a Method for Traffic Safety Evaluation: the Swedish Traffic Conflicts Technique», *Bulletin Lund Institute of Technology, Department of Traffic Planning and Engineering*, n. 70, Lund, Svezia, lug. 1987.

[85] S. Zhang, M. Abdel-Aty, Y. Wu, e O. Zheng, «Modeling pedestrians' near-accident events at signalized intersections using gated recurrent unit (GRU)», *Accident Analysis & Prevention*, vol. 148, ID 105844, dic. 2020, doi: 10.1016/J.AAP.2020.105844.

[86] K. El-Basyouny e T. Sayed, «Safety performance functions using traffic conflicts», *Safety Science*, vol. 51, n. 1, pagg. 160–164, gen. 2013, doi: 10.1016/J.SSCI.2012.04.015.

[87] E. Sacchi, T. Sayed, e P. Deleur, «A comparison of collision-based and conflict-based safety evaluations: The case of right-turn smart channels», *Accident Analysis & Prevention*, vol. 59, pagg. 260–266, ott. 2013, doi: 10.1016/J.AAP.2013.06.002.

[88] A. Laureshyn, H. LastNameArdö, Å. Svensson, e T. Jonsson, «Application of automated video analysis for behavioural studies: Concept and experience», *IET Intelligent Transport Systems*, vol. 3, n. 3, pagg. 345–357, set. 2009, doi: 10.1049/IET-ITS.2008.0077.

[89] Y. Ni, M. Wang, J. Sun, e K. Li, «Evaluation of pedestrian safety at intersections: A theoretical framework based on pedestrian-vehicle interaction patterns», *Accident Analysis & Prevention*, vol. 96, pagg. 118–129, nov. 2016, doi: 10.1016/J.AAP.2016.07.030.

[90] L. Yue, M. Abdel-Aty, Y. Wu, O. Zheng, e J. Yuan, «In-depth approach for identifying crash causation patterns and its implications for pedestrian crash prevention», *Journal of Safety Research*, vol. 73, pagg. 119–132, giu. 2020, doi: 10.1016/J.JSR.2020.02.020.

[91] K. Ismail, T. Sayed, e N. Saunier, «Automated Analysis of Pedestrian–Vehicle Conflicts: Context for Before-and-After Studies», *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2198, n. 1, pagg. 52–64, gen. 2010, doi: 10.3141/2198-07.

[92] M. H. Zaki e T. Sayed, «Automated Analysis of Pedestrians' Nonconforming Behavior and Data Collection at an Urban Crossing», *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2443, n. 1, pagg. 123–133, gen. 2014, doi: 10.3141/2443-14.

[93] Y. Wu, M. Abdel-Aty, O. Zheng, Q. Cai, e S. Zhang, «Automated Safety Diagnosis Based on Unmanned Aerial Vehicle Video and Deep Learning Algorithm», *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, n. 8, pagg. 350–359, giu. 2020, doi: 10.1177/0361198120925808.

[94] Y. Wang, D. Zhang, Y. Liu, B. Dai, e L. H. Lee, «Enhancing transportation systems via deep learning: A survey», *Transportation Research Part C: Emerging Technologies*, vol. 99, pagg. 144–163, feb. 2019, doi: 10.1016/J.TRC.2018.12.004.

[95] Y. Gong, M. Abdel-Aty, J. Yuan, e Q. Cai, «Multi-Objective reinforcement learning approach for improving safety at intersections with adaptive traffic signal control», *Accident Analysis & Prevention*, vol. 144, ID 105655, set. 2020, doi: 10.1016/J.AAP.2020.105655.

[96] P. Li, M. Abdel-Aty, e J. Yuan, «Real-time crash risk prediction on arterials based on LSTM-CNN», *Accident Analysis & Prevention*, vol. 135, ID 105371, feb. 2020, doi: 10.1016/J.AAP.2019.105371.

[97] Z. C. Lipton, J. Berkowitz, e C. Elkan, «A Critical Review of Recurrent Neural Networks for Sequence Learning», *ArXiv pre-prints,* mag. 2015, ArXiv ID: 1506.00019.

[98] S. Hochreiter e J. Schmidhuber, «Long Short-Term Memory», *Neural Computation*, vol. 9, n. 8, pagg. 1735–1780, nov. 1997, doi: 10.1162/NECO.1997.9.8.1735.

[99] J. Chung, C. Gulcehre, K. Cho, e Y. Bengio, «Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling», *ArXiv pre-prints,* dic. 2014, ArXiv ID*:* 1412.3555.

[100] L. Zhao *et al.*, «T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction», *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, n. 9, pagg. 3848–3858, ago. 2019, doi: 10.1109/TITS.2019.2935152.

[101] J. Zhao, Y. Gao, Y. Qu, H. Yin, Y. Liu, e H. Sun, «Travel Time Prediction: Based on Gated Recurrent Unit Method and Data Fusion», *IEEE Access*, vol. 6, pagg. 70463–70472, ott. 2018, doi: 10.1109/ACCESS.2018.2878799.

[102] E. Rosén e U. Sander, «Pedestrian fatality risk as a function of car impact speed», *Accident Analysis & Prevention*, vol. 41, n. 3, pagg. 536–542, mag. 2009, doi: 10.1016/J.AAP.2009.02.002.

[103] B. C. Tefft, «Impact speed and a pedestrian's risk of severe injury or death», *Accident Analysis & Prevention*, vol. 50, pagg. 871-878, gen. 2013, doi: 10.1016/j.aap.2012.07.022.

[104] L. Jakobsson *et al.*, «Pedestrian airbag technology: a production system», in Atti della *23rd International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, Seoul, Corea del Sud, mag. 2013, ID 13-0447.

[105] E. Ames e P. Martin, «Pop-up hood pedestrian protection», negli Atti della *24th International Technical Conference on the Enhanced Safety of Vehicles (ESV)*, Gothenburg, Svezia, giu. 2015, ID 15-0111.

[106] C. Oh, Y. soo Kang, e W. Kim, «Assessing the safety benefits of an advanced vehicular technology for protecting pedestrians», *Accident Analysis & Prevention*, vol. 40, n. 3, pagg. 935-942, mag. 2008, doi: 10.1016/j.aap.2007.10.010.

[107] O. M. J. Carsten e F. N. Tate, «Intelligent speed adaptation: Accident savings and cost-benefit analysis», *Accident Analysis & Prevention*, vol. 37, n. 3, pagg. 407-416, mag. 2005, doi: 10.1016/j.aap.2004.02.007.

[108] S. H. Chang, C. Y. Lin, C. C. Hsu, C. P. Fung, e J. R. Hwang, «The effect of a collision warning system on the driving performance of young drivers at intersections», *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, n. 5, pagg. 371-380, set. 2009, doi: 10.1016/j.trf.2009.05.001.

[109] T. L. Brown, J. D. Lee, e D. V. McGehee, «Human performance models and rear-end collision avoidance algorithms», *Human Factors*, vol. 43, n. 3, pagg. 462–482, set. 2001, doi: 10.1518/001872001775898250.

[110] S. B. McLaughlin, J. M. Hankey, e T. A. Dingus, «A method for evaluating collision avoidance systems using naturalistic driving data», *Accident Analysis & Prevention*, vol. 40, n. 1, pagg. 8-16, gen. 2008, doi: 10.1016/j.aap.2007.03.016.

[111] N. Lubbe e J. Davidsson, «Drivers' comfort boundaries in pedestrian crossings: A study in driver braking characteristics as a function of pedestrian walking speed», *Safety Science*, vol. 75, pagg. 100-106, giu. 2015, doi: 10.1016/j.ssci.2015.01.019.

[112] J. L. Campbell, C. M. Richard, J. L. Brown, e M. McCallum, «Crash warning system interfaces: Human Factors insight and lessons learned», NHTSA, Washington, DC, USA, Report DOT HS 810 697, gen. 2007.

[113] F. Naujoks, H. Grattenthaler, e A. Neukum, «Zeitliche Gestaltung effektiver Fahrerinformationen zur Kollisionsvermeidung auf der Basis kooperativer Perzeption», negli Atti della *8. Workshop Fahrerassistenzsysteme – FAS 2012*, Walting im Altmühltal, Germania, feb. 2012, pagg. 107–117.

[114] T. Gandhi e M. M. Trivedi, «Pedestrian protection systems: Issues, survey, and challenges», *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, n. 3, pagg. 413–430, set. 2007, doi: 10.1109/TITS.2007.903444.

[115] D. F. Llorca *et al.*, «Stereo regions-of-interest selection for pedestrian protection: A survey», *Transportation Research Part C: Emerging Technologies*, vol. 25, pagg. 226–237, dic. 2012, doi: 10.1016/j.trc.2012.06.006.

[116] B. Musleh, F. García, J. Otamendi, J. M. Armingol, e A. D. la Escalera, «Identifying and Tracking Pedestrians Based on Sensor Fusion and Motion Stability Predictions», *Sensors*, vol. 10, n. 9, pagg. 8028–8053, ago. 2010, doi: 10.3390/S100908028.

[117] A. Badea-Romero, F. Javier Páez, A. Furones, J. M. Barrios, e J. L. de-Miguel, «Assessing the benefit of the brake assist system for pedestrian injury mitigation through real-world accident investigations», *Safety Science*, vol. 53, pagg. 193-201, mar. 2013, doi: 10.1016/j.ssci.2012.10.004.

[118] R. Wu *et al.*, «Modified driving safety field based on trajectory prediction model for pedestrian-vehicle collision», *Sustainability (Switzerland)*, vol. 11, n. 22, ID 6254, nov. 2019, doi: 10.3390/su11226254.

[119] S. Lefèvre, D. Vasquez, e C. Laugier, «A survey on motion prediction and risk assessment for intelligent vehicles», *ROBOMECH Journal*, vol. 1, n. 1, pagg. 1–14, lug. 2014.

[120] A. Rasouli e J. K. Tsotsos, «Autonomous vehicles that interact with pedestrians: A survey of theory and practice», *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, n. 3, pagg. 900–918, mar. 2020, doi: 10.1109/TITS.2019.2901817.

[121] I. Dagli e D. Reichardt, «Motivation-based approach to behavior prediction», negli Atti della *IEEE Intelligent Vehicles Symposium 2002*, Versailles, Francia, giu. 2002, vol. 1, pagg. 227–233, doi: 10.1109/IVS.2002.1187956.

[122] T. Gindele, S. Brechtel, e R. Dillmann, «A probabilistic model for estimating driver behaviors and vehicle trajectories in traffic environments», negli Atti della *13th International IEEE Conference on Intelligent Transportation Systems*, Funchal, Portogallo, set. 2010, pagg. 1625–1631, doi: 10.1109/ITSC.2010.5625262.

[123] G. S. Aoude, V. R. Desaraju, L. H. Stephens, e J. P. How, «Driver behavior classification at intersections and validation on large naturalistic data set», *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, n. 2, pagg. 724–736, feb. 2012, doi: 10.1109/TITS.2011.2179537.

[124] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, e K. O. Arras, «Human Motion Trajectory Prediction: A Survey», *International Journal of Robotics Research*, vol. 39, n. 8, pagg. 895–935, mag. 2019, doi: 10.1177/0278364920917446.

[125] M. Garcia Ortiz, J. Fritsch, F. Kummert, e A. Gepperth, «Behavior prediction at multiple time-scales in inner-city scenarios», negli Atti della *2011 IEEE Intelligent Vehicles Symposium (IV)*, Baden-Baden, Germania, giu. 2011, pagg. 1068–1073, doi: 10.1109/IVS.2011.5940524.

[126] N. Rhinehart, K. M. Kitani, e P. Vernaza, «R2P2: A reparameterized pushforward policy for diverse, precise generative path forecasting», negli Atti della *15th European Conference on Computer Vision (ECCV 2018)*, Monaco, Germania, set. 2018, vol. 112117 LNCS, pagg. 794–811, doi: 10.1007/978-3-030-01261-8_47.

[127] Y. Luo, P. Cai, D. Hsu, e W. S. Lee, «GAMMA: A General Agent Motion Prediction Model for Autonomous Driving», *ArXiv pre-prints*, giu. 2019, ArXiv ID: 1906.01566.

[128] S. Bansal, A. Bajcsy, E. Ratner, A. D. Dragan, e C. J. Tomlin, «A Hamilton-Jacobi

Reachability-Based Framework for Predicting and Analyzing Human Motion for Safe Planning», negli Atti della *2020 International IEEE Conference on Robotics and Automation (ICRA)*, Parigi, Francia, aug. 2020, pagg. 7149–7155, doi: 10.1109/ICRA40945.2020.9197257.

[129] E. Käfer, C. Hermes, C. Wöhler, H. Ritter, e F. Kummert, «Recognition of situation classes at road intersections», negli Atti della *2010 International IEEE Conference on Robotics and Automation (ICRA)*, Anchorage, AK, USA, mag. 2010, pagg. 3960–3965, doi: 10.1109/ROBOT.2010.5509919.

[130] A. Zyner, S. Worrall, e E. Nebot, «Naturalistic Driver Intention and Path Prediction using Recurrent Neural Networks», *ArXiv pre-prints*, lug. 2018, ArXiv ID: 1807.09995.

[131] Y. Gu, Y. Hashimoto, L. T. Hsu, e S. Kamijo, «Motion planning based on learning models of pedestrian and driver behaviors», negli Atti della *19th Iternational IEEE Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, Brasile, nov. 2016, pagg. 808–813, doi: 10.1109/ITSC.2016.7795648.

[132] R. Q. Minguez, I. P. Alonso, D. Fernandez-Llorca, e M. A. Sotelo, «Pedestrian Path, Pose, and Intention Prediction Through Gaussian Process Dynamical Models and Pedestrian Activity Recognition», *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, n. 5, pagg. 1803–1814, mag. 2018, doi: 10.1109/TITS.2018.2836305.

[133] E. A. I. Pool, J. F. P. Kooij, e D. M. Gavrila, «Context-based cyclist path prediction using Recurrent Neural Networks», negli Atti della *2019 IEEE Intelligent Vehicles Symposium (IV)*, Parigi, Francia, giu. 2019, vol. 2019-June, pagg. 824–830, doi: 10.1109/IVS.2019.8813889.

[134] X. Li, X. Ying, e M. C. Chuah, «GRIP: Graph-based Interaction-aware Trajectory Prediction», negli Atti della *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, Auckland, Nuova Zelanda, ott. 2019, pagg. 3960–3966, doi: 10.1109/ITSC.2019.8917228.

[135] F. Diehl, T. Brunner, M. T. Le, e A. Knoll, «Graph neural networks for modelling traffic participant interaction», negli Atti della *2019 IEEE Intelligent Vehicles Symposium (IV)*, Parigi, Francia, giu. 2019, vol. 2019-June, pagg. 695–701, doi: 10.1109/IVS.2019.8814066.

[136] W. Ding e S. Shen, «Online vehicle trajectory prediction using policy anticipation network and optimization-based context reasoning», negli Atti della *2019 International IEEE Conference on Robotics and Automation*, Montréal, Canada, mag. 2019, vol. 2019-May, pagg. 9610–9616, doi: 10.1109/ICRA.2019.8793568.

[137] F. Altché e A. de La Fortelle, «An LSTM network for highway trajectory prediction», negli Atti della *20th International IEEE Conference on Intelligent Transportation Systems*, Yokohama, Giappone, ott. 2017, vol. 2018-March, pagg. 353–359, doi: 10.1109/ITSC.2017.8317913.

[138] J. Schlechtriemen, F. Wirthmueller, A. Wedel, G. Breuel, e K. D. Kuhnert, «When will it change the lane? A probabilistic regression approach for rarely occurring events», negli Atti della *2015 IEEE Intelligent Vehicles Symposium (IV)*, Seoul, Corea del Sud, giu. 2015, vol. 2015-August, pagg. 1373–1379, doi: 10.1109/IVS.2015.7225907.

[139] D. T. Field e J. P. Wann, «Perceiving time to collision activates the sensorimotor cortex», *Current Biology*, vol. 15, n. 5, pagg. 453–458, mar. 2005, doi: 10.1016/J.CUB.2004.12.081.

[140] J. F. P. Kooij, F. Flohr, E. A. I. Pool, e D. M. Gavrila, «Context-Based Path Prediction for Targets with Switching Dynamics», *International Journal of Computer Vision*, vol. 127, n. 3, pagg. 239–262, mar. 2019, doi: 10.1007/S11263-018-1104-4.

[141] S. Yi, H. Li, e X. Wang, «Pedestrian Behavior Understanding and Prediction with Deep Neural Networks», negli Atti della *14th European Conference on Computer Vision (ECCV 2016)*, Amsterdam, Paesi Bassi, ott. 2016, vol. 9905 LNCS, pagg. 263–279, doi:

10.1007/978-3-319-46448-0_16.

[142] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, e S. Savarese, «Social LSTM: Human trajectory prediction in crowded spaces», negli Atti della *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, giu. 2016, vol. 2016-Decem, pagg. 961–971, doi: 10.1109/CVPR.2016.110.

[143] W. Ding, J. Chen, e S. Shen, «Predicting vehicle behaviors over an extended horizon using behavior interaction network», negli Atti della *2019 International IEEE Conference on Robotics and Automation (ICRA)*, Montréal, Canada, mag. 2019, vol. 2019-May, pagg. 8634–8640, doi: 10.1109/ICRA.2019.8794146.

[144] S. Dai, L. Li, e Z. Li, «Modeling Vehicle Interactions via Modified LSTM Models for Trajectory Prediction», *IEEE Access*, vol. 7, pagg. 38287–38296, mar. 2019, doi: 10.1109/ACCESS.2019.2907000.

[145] S. Ruder, «An Overview of Multi-Task Learning in Deep Neural Networks», *ArXiv pre-prints*, giu. 2017, ArXiv ID: 1706.05098.

[146] E. Arnold, O. Y. Al-Jarrah, M. Dianati, S. Fallah, D. Oxtoby, e A. Mouzakitis, «A Survey on 3D Object Detection Methods for Autonomous Driving Applications», *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, n. 10, pagg. 3782–3795, gen. 2019, doi: 10.1109/TITS.2019.2892405.

[147] W. Luo, B. Yang, e R. Urtasun, «Fast and Furious: Real Time End-to-End 3D Detection, Tracking and Motion Forecasting with a Single Convolutional Net», negli Atti della *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, giu. 2018, pagg. 3569–3577, doi: 10.1109/CVPR.2018.00376.

[148] S. Casas, W. Luo, e R. Urtasun, «IntentNet: Learning to Predict Intention from Raw Sensor Data», *ArXiv pre-prints*, gen. 2021, ArXiv ID: 2101.07907.

[149] D. Ridel, E. Rehder, M. Lauer, C. Stiller, e D. Wolf, «A Literature Review on the Prediction of Pedestrian Behavior in Urban Scenarios», negli Atti della *21th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI, USA, nov. 2018, pagg. 3105–3112, doi: 10.1109/ITSC.2018.8569415.

[150] N. Khairdoost, M. Shirpour, M. A. Bauer, e S. S. Beauchemin, «Real-Time Driver Maneuver Prediction Using LSTM», *IEEE Transactions on Intelligent Vehicles*, vol. 5, n. 4, pagg. 714–724, dic. 2020, doi: 10.1109/TIV.2020.3003889.

[151] Z. Huang, J. Wang, L. Pi, X. Song, e L. Yang, «LSTM based trajectory prediction model for cyclist utilizing multiple interactions with environment», *Pattern Recognition*, vol. 112, ID 107800, apr. 2021, doi: 10.1016/J.PATCOG.2020.107800.

[152] S. Mozaffari, O. Y. Al-Jarrah, M. Dianati, P. Jennings, e A. Mouzakitis, «Deep Learning-Based Vehicle Behavior Prediction for Autonomous Driving Applications: A Review», *IEEE Transactions on Intelligent Transportation Systems*, vol 23, n. 1, pagg. 33–47, ago. 2020, doi: 10.1109/TITS.2020.3012034.

[153] F. Kuhnt, J. Schulz, T. Schamm, e J. M. Zollner, «Understanding interactions between traffic participants based on learned behaviors», negli Atti della *2016 IEEE Intelligent Vehicles Symposium (IV)*, Gothenburg, Svezia, giu. 2016, vol. 2016-August, pagg. 1271–1278, doi: 10.1109/IVS.2016.7535554.

[154] C. Hubmann, M. Becker, D. Althoff, D. Lenz, e C. Stiller, «Decision making for autonomous driving considering interaction and uncertain prediction of surrounding vehicles», negli Atti della *2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, USA, giu. 2017, pagg. 1671–1678, doi: 10.1109/IVS.2017.7995949.

[155] Y. Bengio, P. Simard, e P. Frasconi, «Learning Long-Term Dependencies with Gradient Descent is Difficult», *IEEE Transactions on Neural Networks*, vol. 5, n. 2, pagg. 157–166, mar. 1994, doi: 10.1109/72.279181.

[156] K. Cho *et al.*, «Learning phrase representations using RNN encoder-decoder for statistical

machine translation», negli Atti della *2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014)*, Doha, Qatar, ott. 2014, pagg. 1724–1734, doi: 10.3115/v1/d14-1179.

[157] A. Zyner, S. Worrall, e E. Nebot, «A Recurrent Neural Network Solution for Predicting Driver Intention at Unsignalized Intersections», *IEEE Robotics and Automation Letters*, vol. 3, n. 3, pagg. 1759–1764, feb. 2018, doi: 10.1109/LRA.2018.2805314.

[158] D. J. Phillips, T. A. Wheeler, e M. J. Kochenderfer, «Generalizable intention prediction of human drivers at intersections», negli Atti della *2017 IEEE Intelligent Vehicles Symposium (IV)*, Los Angeles, CA, USA, giu. 2017, pagg. 1665–1670, doi: 10.1109/IVS.2017.7995948.

[159] L. Xin, P. Wang, C. Y. Chan, J. Chen, S. E. Li, e B. Cheng, «Intention-aware Long Horizon Trajectory Prediction of Surrounding Vehicles using Dual LSTM Networks», negli Atti della *21th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Maui, HI, USA, nov. 2018, vol. 2018-November, pagg. 1441–1446, doi: 10.1109/ITSC.2018.8569595.

[160] N. Deo e M. M. Trivedi, «Multi-Modal Trajectory Prediction of Surrounding Vehicles with Maneuver based LSTMs», negli Atti della *2018 IEEE Intelligent Vehicles Symposium (IV)*, Changshu, Cina, giu. 2018, vol. 2018-June, pagg. 1179–1184, doi: 10.1109/IVS.2018.8500493.

[161] A. Krizhevsky, I. Sutskever, e G. E. Hinton, «ImageNet classification with deep convolutional neural networks», *Communications of the ACM*, vol. 60, n. 6, pagg. 84–90, mag. 2017, doi: 10.1145/3065386.

[162] K. He, G. Gkioxari, P. Dollar, e R. Girshick, «Mask R-CNN», *ArXiv pre-prints*, gen. 2018, ArXiv ID: 1703.06870v3.

[163] S. Ren, K. He, R. Girshick, e J. Sun, «Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks», *ArXiv pre-prints*, gen. 2016, ArXiv ID: 1506.01497v3.

[164] O. Zheng, M. Abdel-Aty, e Y. Wu, «A.R.C.I.S - Automated Roadway Conflicts Identification System», UCF SST, Orlando, FL, USA, gen. 2019. http://safesmartsst.com/arcis.html (consultato dic. 15, 2021).

[165] A. Lukezic, T. Vojir, L. Cehovin Zajc, J. Matas, e M. Kristan, «Discriminative Correlation Filter With Channel and Spatial Reliability», negli Atti della *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, lug. 2017, pagg. 6309–6318, doi: 10.1109/CVPR.2017.515.

[166] N. Deo e M. M. Trivedi, «Convolutional social pooling for vehicle trajectory prediction», negli Atti della *2018 IEEE/CVF Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Salt Lake City, UT, USA, giu. 2018, vol. 2018-June, pagg. 1549–1557, doi: 10.1109/CVPRW.2018.00196.

[167] W. Yang, X. Zhang, Q. Lei, e X. Cheng, «Research on longitudinal active collision avoidance of autonomous emergency braking pedestrian system (AEB-P)», *Sensors (Switzerland)*, vol. 19, n. 21, ID 4671, ott. 2019, doi: 10.3390/s19214671.

[168] Å. Svensson, «A method for analysing the traffic process in a safety perspective», Tesi di Dottorato, Department of Traffic Planning and Engineering, Lund Institute of Technology, Lund, Sweden, ott. 1998.

[169] A. Laureshyn, T. De Ceunynck, C. Karlsson, Å. Svensson, e S. Daniels, «In search of the severity dimension of traffic events: Extended Delta-V as a traffic conflict indicator», *Accident Analysis & Prevention*, vol. 98, pagg. 46–56, gen. 2017, doi: 10.1016/J.AAP.2016.09.026.

[170] M. Mohamed e N. Saunier, «Motion Prediction Methods for Surrogate Safety Analysis»:, *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2386, n. 1, pagg. 168–178, gen. 2013, doi: 10.3141/2386-19.

[171] L. Zheng, K. Ismail, e X. Meng, «Traffic conflict techniques for road safety analysis: open questions and some insights», *Canadian Journal of Civil Engineering*, vol. 41, n. 7, pagg. 633–641, mag. 2014, doi: 10.1139/CJCE-2013-0558.

[172] J. H. Kraay, A. R. A. van der Horst, e S. Oppe, «Manual conflict observation technique DOCTOR (Dutch Objective Conflict Technique for Operation and Research)», FRSFA, Voorburg, Paesi Bassi, Report 2013-1, set. 2013.

[173] A. R. A. Van der Horst, «A time-based analysis of road user behaviour in normal and critical encounters», Tesi di Dottorato, Department of Electrical Engineering, Mathematics and Computer Science, Technische Universiteit Delft, Delft, Paesi Bassi, apr. 1990.

[174] A. Laureshyn, M. de Goede, N. Saunier, e A. Fyhri, «Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway», *Accident Analysis & Prevention*, vol. 105, pagg. 11–20, ago. 2017, doi: 10.1016/J.AAP.2016.04.035.

[175] A. Hansson, «Studies in Driver Behaviour, with Applications in Traffic Design and Planning. Two Examples», *Bulletin Lund Institute of Technology*, n. 9, gen. 1975.

[176] O. Hagring, «Relation of Traffic Efficiency and Safety: A Pilot Study», *Bulletin Lund Institute of Technology*, n. 192, Lund, Svezia, dec. 2000.

[177] A. Várhelyi, «Drivers' speed behaviour at a zebra crossing: a case study», *Accident Analysis & Prevention*, vol. 30, n. 6, pagg. 731–743, nov. 1998, doi: 10.1016/S0001-4575(98)00026-8.

[178] A. Borsos, H. Farah, A. Laureshyn, e M. Hagenzieker, «Are collision and crossing course surrogate safety indicators transferable? A probability based approach using extreme value theory», *Accident Analysis & Prevention*, vol. 143, ID 105517, ago. 2020, doi: 10.1016/J.AAP.2020.105517.

[179] S. M. S. Mahmud, L. Ferreira, M. S. Hoque, e A. Tavassoli, «Application of proximal surrogate indicators for safety evaluation: A review of recent developments and research needs», *IATSS Research*, vol. 41, n. 4, pagg. 153–163, dic. 2017, doi: 10.1016/J.IATSSR.2017.02.001.

[180] F. Huang, P. Liu, H. Yu, e W. Wang, «Identifying if VISSIM simulation model and SSAM provide reasonable estimates for field measured traffic conflicts at signalized intersections», *Accident Analysis & Prevention*, vol. 50, pagg. 1014–1024, gen. 2013, doi: 10.1016/J.AAP.2012.08.018.

[181] T. Sayed, M. H. Zaki, e J. Autey, «Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis», *Safety Science*, vol. 59, pagg. 163–172, nov. 2013, doi: 10.1016/J.SSCI.2013.05.009.

[182] Y. Zhang, E. K. Antonsson, e K. Grote, «A new threat assessment measure for collision avoidance systems», negli Atti della *2006 IEEE Conference on Intelligent Transportation Systems (ITSC)*, Toronto, Canada, set. 2006, pagg. 968–975, doi: 10.1109/ITSC.2006.1706870.

[183] A. Laureshyn, K. Åström, e K. Brundell-Freij, «FROM SPEED PROFILE DATA TO ANALYSIS OF BEHAVIOUR: Classification by Pattern Recognition Techniques», *IATSS Research*, vol. 33, n. 2, pagg. 88–98, gen. 2009, doi: 10.1016/S0386-1112(14)60247-8.

[184] C. W. Anderson, M. R. Leadbetter, G. Lindgren, e H. Rootzen, «Extremes and Related Properties of Random Sequences and Processes», *Journal of the Royal Statistical Society. Series A (General)*, vol. 147, n. 5, 1984, doi: 10.2307/2981706.

[185] C. Wang, C. Xu, J. Xia, Z. Qian, e L. Lu, «A combined use of microscopic traffic simulation and extreme value methods for traffic safety evaluation», *Transportation Research Part C: Emerging Technologies*, vol. 90, pagg. 281–291, mag. 2018, doi: 10.1016/J.TRC.2018.03.011.

[186] L. Zheng e T. Sayed, «A full Bayes approach for traffic conflict-based before–after safety

evaluation using extreme value theory», *Accident Analysis & Prevention*, vol. 131, pagg. 308–315, ott. 2019, doi: 10.1016/J.AAP.2019.07.014.

[187] F. Orsini, G. Gecchele, M. Gastaldi, e R. Rossi, «Collision prediction in roundabouts: a comparative study of extreme value theory approaches», *Transportmetrica A: Transport Science*, vol. 15, n. 2, pagg. 556–572, set. 2019, doi: 10.1080/23249935.2018.1515271.

[188] K. Masayuki, O. Katsuyuki, e S. Masayoshi, «79GHz-band high resolution millimeter-wave radar», *Fujitsu Scientific & Technical Journal*, vol. 51, n. 4, ott. 2015.

[189] A. Cennamo, F. Kaestner, e A. Kummert, «A Neural Network Based System for Efficient Semantic Segmentation of Radar Point Clouds», *Neural Processing Letters*, vol. 53, n. 5, pagg. 3217–3235, apr. 2021, doi: 10.1007/S11063-021-10544-4.

[190] S. T. Isele, F. Klein, M. Brosowsky, e J. M. Zollner, «Learning semantics on radar point-clouds», negli Atti della *2021 IEEE Intelligent Vehicles Symposium (IV)*, Nagoya, Giappone, lug. 2021, vol. 2021-July, pagg. 810–817, doi: 10.1109/IV48863.2021.9575779.

[191] M. Meyer, G. Kuschk, e S. Tomforde, «Graph Convolutional Networks for 3D Object Detection on Radar Data», negli Atti della *2021 IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, ott. 2021, pagg. 3060–3069.

[192] C.-C. Lin, C.-H. Kuo, e H.-T. Chiang, «CNN-Based Classification for Point Cloud Object With Bearing Angle Image», *IEEE Sensors Journal*, vol. 22, n. 1, pagg. 1003-1011, nov. 2021, doi: 10.1109/JSEN.2021.3130268.

[193] T. Hastie, R. Tibshirani, e J. Friedman, *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, ago. 2009. doi: 10.1007/978-0-387-84858-7

[194] M. F. M. Siam, N. Borhan, e A. Sukardi, «Driving Simulator Development with Two Degrees of Freedom Motion for Driver Behavior Study», *Journal of the Society of Automotive Engineers Malaysia*, vol. 1, n. 1, pagg. 4–11, gen. 2017.

[195] F. Bella e M. Silvestri, «Vehicle–pedestrian interactions into and outside of crosswalks: effects of driver assistance systems», *Transport*, vol. 36, n. 2, pagg. 98–109, apr. 2021, doi: 10.3846/TRANSPORT.2021.14739.

[196] Y. Zhang, Z. Guo, Z. Sun, e Z. Sun, «Driving Simulator Validity of Driving Behavior in Work Zones», *Journal of Advanced Transportation*, vol. 2020, ID 4629132, giu. 2020, doi: 10.1155/2020/4629132.

[197] Y. Saito e P. Raksincharoensak, «Shared Control in Risk Predictive Braking Maneuver for Preventing Collisions with Pedestrians», *IEEE Transactions on Intelligent Vehicles*, vol. 1, n. 4, pagg. 314–324, dic. 2016, doi: 10.1109/TIV.2017.2700210.

[198] L. Hou, J. Duan, W. Wang, R. Li, G. Li, e B. Cheng, «Drivers' Braking Behaviors in Different Motion Patterns of Vehicle-Bicycle Conflicts», *Journal of Advanced Transportation*, vol. 2019, ID 4023970, mar. 2019, doi: 10.1155/2019/4023970.

[199] S. Choi, J. Jang, C. Oh, e G. Park, «Safety benefits of integrated pedestrian protection systems», *International Journal of Automotive Technology*, vol. 17, n. 3, pagg. 473–482, apr. 2016, doi: 10.1007/S12239-016-0049-2.

[200] J. Kovaceva, A. Bálint, R. Schindler, e A. Schneider, «Safety benefit assessment of autonomous emergency braking and steering systems for the protection of cyclists and pedestrians based on a combination of computer simulation and real-world test results», *Accident Analysis & Prevention*, vol. 136, ID 105352, mar. 2020, doi: 10.1016/J.AAP.2019.105352.

[201] F. Bella, «Can Driving Simulators Contribute to Solving Critical Issues in Geometric Design?»:, *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2138, n. 1, pagg. 120–126, gen. 2009, doi: 10.3141/2138-16.

[202] F. Bella e M. Silvestri, «Effects of safety measures on driver's speed behavior at

pedestrian crossings», *Accident Analysis & Prevention*, vol. 83, pagg. 111–124, ott. 2015, doi: 10.1016/J.AAP.2015.07.016.

[203] H. Prakken, «On the problem of making autonomous vehicles conform to traffic law», *Artificial Intelligence and Law*, vol. 25, n. 3, pagg. 341–363, ago. 2017, doi: 10.1007/S10506-017-9210-0.

[204] W. Huber e J. Gleber, «Linguistic and nonlinguistic processing of narratives in aphasia», *Brain and Language*, vol. 16, n. 1, pagg. 1–18, mag. 1982, doi: 10.1016/0093-934X(82)90069-4.

[205] M. Paradis e G. Libben, *The Assessment of Bilingual Aphasia*, 1st ed. New York, NY, USA: Psychology Press, mag. 1987.

[206] L. E. Nicholas e R. H. Brookshire, «A System for Quantifying the Informativeness and Efficiency of the Connected Speech of Adults With Aphasia», *Journal of Speech and Hearing Research*, vol. 36, n. 2, pagg. 338–350, apr. 1993, doi: 10.1044/JSHR.3602.338.

[207] F. Bella, V. Natale, e M. Silvestri, «Driver-pedestrian interaction under different road environments», *Transportation Research Procedia*, vol. 27, pagg. 148–155, gen. 2017, doi: 10.1016/J.TRPRO.2017.12.093.

[208] M. Saifuzzaman, M. M. Haque, Z. Zheng, e S. Washington, «Impact of mobile phone use on car-following behaviour of young drivers», *Accident Analysis & Prevention*, vol. 82, pagg. 10–19, set. 2015, doi: 10.1016/J.AAP.2015.05.001.

[209] F. J. Gravetter e L. B. Wallnau, *Statistics for the behavioral sciences*, 10th ed. Boston, MA, USA: Cengage Learning, gen. 2017.

[210] D. M. Lane, «Online Statistics Education: An Interactive Multimedia Course of Study». https://onlinestatbook.com/ (consultato gen. 10, 2022).

[211] H. Obeid, H. Abkarian, M. Abou-Zeid, e I. Kaysi, «Analyzing driver-pedestrian interaction in a mixed-street environment using a driving simulator», *Accident Analysis & Prevention*, vol. 108, pagg. 56–65, nov. 2017, doi: 10.1016/J.AAP.2017.08.005.

[212] T. A. Dingus *et al.*, «The prevalence of and crash risk associated with primarily cognitive secondary tasks», *Safety Science*, vol. 119, pagg. 98–105, nov. 2019, doi: 10.1016/J.SSCI.2019.01.005.

[213] J. M. Cooper, N. Medeiros-Ward, e D. L. Strayer, «The impact of eye movements and cognitive workload on lateral position variability in driving», *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 55, n. 5, pagg. 1001–1014, ott. 2013, doi: 10.1177/0018720813480177.

[214] F. Bella, V. Natale, M. Silvestri, e F. Nobili, «Drivers' Behavior in Pedestrian Detection: Effects of Road Types», negli Atti della *International Conference on Human Factors and in Transportation (AHFE 2018)*, Orlando, FL, USA, lug. 2019, vol. 786, pagg. 813–822, doi: 10.1007/978-3-319-93885-1_75.

[215] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. New York, NY, USA: Routledge - Taylor & Francis Group, gen. 2013, doi: 10.4324/9780203771587.

[216] J. Törnros e A. Bolling, «Mobile phone use – effects of conversation on mental workload and driving speed in rural and urban environments», *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 9, n. 4, pagg. 298–306, lug. 2006, doi: 10.1016/J.TRF.2006.01.008.

[217] M. A. Recarte e L. M. Nunes, «Mental Workload While Driving: Effects on Visual Search, Discrimination, and Decision Making», *Journal of Experimental Psychology: Applied*, vol. 9, n. 2, pagg. 119–137, giu. 2003, doi: 10.1037/1076-898X.9.2.119.

[218] D. L. Strayer, J. Cooper, e F. A. Drews, «Profiles in cell phone-induced driver distraction», in *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*. Boca Raton, FL, USA: CRC Press - Taylor & Francis Group, apr. 2011, pagg. 27/1-27/10.

[219] ZF - Products for Cars, «Automated Driving Functions: ZF Captures Contract for 4D Full-Range Radar with Chinese OEM - ZF», apr. 2021. https://press.zf.com/press/en/releases/release_25856.html (consultato dic. 29, 2021).

[220] N. Baldo, A. Marini, e M. Miani, «Effects of cognitive distraction on driver's stopping behaviour: A virtual car driving simulator study», *IOP Conference Series: Materials Science and Engineering*, vol. 960, n. 2, ID 022082, dic. 2020, doi: 10.1088/1757-899X/960/2/022082.

[221] W. S. McCulloch e W. Pitts, «A logical calculus of the ideas immanent in nervous activity», *The Bulletin of Mathematical Biophysics*, vol. 5, n. 4, pagg. 115–133, dic. 1943, doi: 10.1007/BF02478259.

[222] D. E. Rumelhart, G. E. Hinton, e R. J. Williams, «Learning representations by back-propagating errors», *Nature*, vol. 323, pagg. 533–536, ott. 1986, doi: 10.1038/323533a0.

[223] D. P. Kingma e J. L. Ba, «Adam: A method for stochastic optimization», *ArXiv pre-prints*, gen. 2017, ArXiv ID: 1412.6980v9.

[224] S. Ruder, «An overview of gradient descent optimization algorithms», *ArXiv pre-prints*, gen. 2017, ArXiv ID: 1609.04747v2.

[225] C. M. Bishop, «Probabilistic Generative Models», in *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, gen. 2006, pagg. 196–198.

[226] A. Y. Ng, «Feature selection, L1 vs. L2 regularization, and rotational invariance», negli Atti della *21th International Conference on Machine Learning (ICML 2004)*, Banff, Canada, lug. 2004, pagg. 615–622, doi: 10.1145/1015330.1015435.

[227] G. James, D. Witten, T. Hastie, e R. Tibshirani, «Cross-Validation», in *An Introduction to Statistical Learning: With Applications in R*. New York, NY, USA: Springer, giu. 2013, pagg. 176–186.

[228] R. Fu, Z. Zhang, e L. Li, «Using LSTM and GRU neural network methods for traffic flow prediction», negli Atti della *31st Youth Academic Annual Conference of Chinese Association of Automation (YAC 2016)*, Wuhan, Cina, nov. 2016, pagg. 324–328, doi: 10.1109/YAC.2016.7804912.

[229] Z. Karevan e J. A. K. Suykens, «Transductive LSTM for time-series prediction: An application to weather forecasting», *Neural Networks*, vol. 125, pagg. 1–9, mag. 2020, doi: 10.1016/J.NEUNET.2019.12.030.

[230] M. Dunnhofer, N. Martinel, e C. Micheloni, «Weakly-supervised domain adaptation of deep regression trackers via reinforced knowledge distillation», *IEEE Robotics and Automation Letters*, vol. 6, n. 3, pagg. 5016–5023, apr. 2021, doi: 10.1109/LRA.2021.3070816.

[231] The MathWorks Inc., «TrainingOptions - Options for Training Deep Learning Neural Network». https://it.mathworks.com/help/deeplearning/ref/trainingoptions.html (consultato dic. 10, 2021).

[232] N. V. Chawla, K. W. Bowyer, L. O. Hall, e W. P. Kegelmeyer, «SMOTE: Synthetic Minority Over-sampling Technique», *Journal of Artificial Intelligence Research*, vol. 16, pagg. 321–357, giu. 2002, doi: 10.1613/JAIR.953.

[233] The MathWorks Inc., «Deep Learning Toolbox». https://it.mathworks.com/products/deep-learning.html (consultato dic. 10, 2021).

[234] M. Inoue, «MATLAB Implementation of SMOTE-related algorithms». https://github.com/minoue-xx/Oversampling-Imbalanced-Data (consultato dic. 10, 2021).

[235] G. Papacharalampous, H. Tyralis, e D. Koutsoyiannis, «Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes», *Stochastic Environmental Research and Risk Assessment*, vol. 33, n. 2, pagg. 481–514, gen. 2019, doi: 10.1007/S00477-018-1638-6.

[236] P. Krause, D. P. Boyle, e F. Bäse, «Comparison of different efficiency criteria for

hydrological model assessment», *Advances in Geosciences*, vol. 5, pagg. 89–97, dic. 2005, doi: 10.5194/ADGEO-5-89-2005.

[237] Euro NCAP, «Vulnerable Road User (VRU) Protection», 2018. https://www.euroncap.com/en/for-engineers/protocols/vulnerable-road-user-vru-protection/ (consultato gen. 20, 2022).