UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXVI CICLO DEL DOTTORATO DI RICERCA IN

SCIENZE DELLA TERRA, MATEMATICA E FLUIDO DINAMICA

3D Models of Rocky Coasts using Structure-from-Motion (SfM) Technology Enhanced by Deep Learning

Settore scientifico-disciplinare: **GEO/03/A**

DOTTORANDA
VALERIA VACCHER

COORDINATORE
PROF. STEFANO MASET

SUPERVISORE DI TESI
PROF. STEFANO FURLANI

CO-SUPERVISORE DI TESI
PROF. SYLVIO BARBON JUNIOR

CO-SUPERVISORE DI TESI
PROF. AMERIGO CORRADETTI

ANNO ACCADEMICO 2023/2024

Dedication

Ai miei figli, Jonathan e Rachel,
la mia gioia più grande e la mia ispirazione. Siete il motivo per cui desidero crescere e migliorarmi ogni giorno. Vi auguro di inseguire i vostri sogni, di essere ambiziosi e di credere sempre in voi stessi.

Al mio compagno, Thomas,
per essere sempre al mio fianco, per il tuo amore e per la fiducia che hai in me.

Abstract

Over 80% of the Mediterranean coastline is rocky, consisting of complex erosional landforms where marine and subaerial processes interact.

The primary objective of this thesis is to use an extensive archive of high-resolution images to study the rocky coasts in the Mediterranean Sea by reconstructing detailed 3D models, combining the geomorphological survey with Deep Learning techniques. The research focuses specifically on the intertidal zone, an area between high and low tide levels, where distinctive tidal notches develop within limestone formations. Antonioli et al. (2015) demonstrated that the Mediterranean region contains numerous tidal notches that exhibit considerable variation in morphological characteristics, dimensions, and elevational positions relative to mean sea level.

Through the swim survey, an extensive dataset was collected using a custom-built raft equipped with all the instrumentation. The methodology involves maintaining a constant swimming speed of 2 km/h while positioning the data collection platform as close as possible to the shoreline, keeping it parallel to the coast. This enables the systematic acquisition of time-lapse images and 360-degree videos above and below the waterline.

Subsequently, an organizational framework was established for the extensive dataset comprising 9 terabytes of coastal images, creating "cancelled" folders dedicated to each group of images, where unsuitable images were placed. This structured classification system provides a robust foundation for subsequent analytical procedures. To enhance data accessibility and collaborative research capabilities, the Geoswim Image Viewer platform was developed, able to display syncronyzed above and below waterline image with integrated geolocation metadata. Leveraging this organized dataset structure, two Convolutional Neural Networks (CNNs) were developed. The first network achieved exceptional performance with 98.85% accuracy in discriminating between useful and non-useful coastal images, while the second network demonstrated 80.6% accuracy in identifying five distinct coastal geomorphotypes including plunging cliffs, sloping coasts, shore platforms, screes, and pocket beaches, following the classification system, following the classification proposed by Biolchi et al. (2016a).

The most significant contribution of this research involves the generation of integrated three-dimensional models of the intertidal zone through the fusion of above and below waterline photogrammetric techniques. The method has been tested in several study areas, but this thesis is presented through three study areas whose results are most significant, including Cala Uccello in Lampedusa, Mġarr ix-Xini in Gozo, and Halq it-Tafal in Malta. The merged 3D models enabled quantitative and semi-qualitative analyzes of tidal notch morphology across the study sites. At Mġarr ix-Xini, statistical analysis of twenty measurement points revealed mean notch dimensions of 2.34 ± 0.074 m in width and 1.26 ± 0.060 m in depth. The estimated eroded volume was 35.87 m^3 along 15.5 m of coastline. In contrast, Halq it-Tafal exhibited morphological variability, with depth measurements ranging from 0.14 to 0.95 m and a mean depth of 0.47 ± 0.052 m. The approach revealed some methodological limitations when applied to certain coastal morphologies, especially in the area of Cala Uccello where the coast is sloping coast type.

This research topic addresses a relevant and current issue in coastal geomorphology. It proposes a new method of studying rocky coasts that integrates SfM technologies with artificial intelligence to automatically analyze coastal morphologies. Although it is not possible to systematically validate the data obtained from the 3D models, and their application is still experimental, this approach lays the groundwork for future automatic coastal mapping and monitoring tools.

Summary

1. Introduction

Rocky coastlines consist of intricate erosional formations resulting from the interaction of marine and terrestrial processes at the air-water interface. These coastlines often have steep vertical profiles that descend directly into deep marine waters, creating cliffs (Kennedy et al., 2014). Despite their geological significance, these coastal environments remained poorly studied until the late 1980s (Trenhaile, 2002), when pioneering works by Sunamura (1992) and Trenhaile (1987) established the framework for rocky coast classification. The recent digital transition and democratization of increasingly advanced digital tools (Tavani et al., 2022) had let to a more semi-automated and quantitative characterization of forms.

This work focuses on the study of rocky coastlines, particularly the application of digital photogrammetry in the intertidal and nearshore zones. It addresses the technical aspects of swim surveys, leveraging recent artificial intelligence advancements to minimize initial data screening. The resulting data is used to generate three-dimensional reconstructions of the intertidal and nearshore zones at three sites in the Mediterranean Sea.

The Mediterranean Sea provides an ideal setting for advancing field rocky coast research, being a semi-enclosed and sheltered basin characterized by temperate waters and relatively stable climatic conditions and particularly suitable for swimming surveys along rocky coastlines (Furlani, 2020; Furlani & Antonioli, 2023). The Geoswim program aims to study the geology and geomorphology of rocky Mediterranean coasts by conducting swimming surveys with the support of a raft that houses all the necessary instruments (Furlani et al., 2014a, 2017a, 2017b, 2018). Visual analysis of the images collected during the field surveys (Furlani et al., 2018; Vaccher, 2019) confirmed the effectiveness of the method in assessing horizontal changes in the coastal shape, but this approach requires a very time-consuming post-processing, given the large number of images collected, both above and below the waterline (9 Terabyte). To address these computational challenges and unlock the full potential of this extensive dataset, the methodological framework builds on the significant advances in computer power and photogrammetric techniques over the past decade (Tavani et al., 2022), particularly the development of Structure-from-Motion (SfM) algorithms. These advancements have fundamentally changed the capabilities of three-dimensional topographic modeling (Fraser & Cronk, 2009; Matthews, 2008; Micheletti et al., 2014; Remondino & El-Hakim, 2006), and have proven valuable in coastal geomorphic research, with applications including documentation and analysis of coastal landslides (Bennett et al., 2012; Devoto et al., 2020, 2021; Esposito et al., 2017), and the study of coastal processes (Swirad et al., 2019).

Recent coastal geomorphology studies have highlighted the advantages of using uncrewed aerial vehicles (UAVs) for photogrammetry. This approach provides several benefits, including efficient monitoring of areas that are vulnerable to degradation and instability (Goncalves & Henriques, 2015), as well as rapid and cost-effective mapping of extensive coastal sections (Piras et al., 2017), revolutionizing the capabilities of three-dimensional topographic modelling (Fraser & Cronk, 2009; Matthews, 2008; Remondino & El-Hakim, 2006).

Additionally, it allows for the production of high-resolution orthophotos, which are essential for accurate three-dimensional reconstructions (Levin et al., 2019; Papakonstantinou et al., 2016). Therefore, it was decided to use UAV images stored in the Geoswim archive to complement the data collected during the field survey, in order to observe the differences between the 3D models created with and without the use of aerial photos. The test is limited to just one case study. In addition to these spatial datasets, temporal analysis is facilitated by video documentation techniques. 'Frame-grabbed' images and time-lapse sequences enable the study of dynamic geomorphological processes, such as beach sediment transport and morphological evolution (Pitman, 2014). However, none of these data alone can provide reliable documentation of the intertidal zone along the coast.

Similarly, the below waterline interface requires the use of underwater photogrammetry, which provides a cost-effective alternative to traditional survey methods such as LiDAR and sonar, especially in shallow water environments (Teague & Scott, 2017).

While photogrammetric methods have been widely applied to geomorphological studies using UAV technologies (Remondino et al., 2011) and dive surveys (Westoby et al., 2012), their application to tidal zone mapping using swim surveys has only recently been explored as part of the Geoswim expeditions (Furlani, 2020) and technical complications arising from different refractive indices between air and seawater.

The research employs an interdisciplinary approach, which, as defined in recent literature (Lyall, 2019: NAS, 2005), involves integrating information, data, techniques, tools, and perspectives from multiple disciplines to address problems that cannot be solved within the confines of a single discipline. By combining field geological and geomorphological survey methods, i.e. field mapping, photographic collection and geomorphological observations with a measuring tape, with Deep Learning techniques, this research aims to establish a protocol of modeling rocky coasts, particularly the intertidal zone.

The emergence of Artificial Intelligence (AI) has greatly expanded the scope and potential of interdisciplinary research, particularly in the geosciences. As demonstrated by Jianping et al. (2016), the development of quantitative geoscience and geological big data analysis represents an emerging trend in research. AI's ability to process and analyze large volumes of data enables researchers to uncover patterns and insights that would otherwise remain hidden using traditional methods of manual data analysis.

This integration of AI with physical approaches proves particularly valuable in tackling complex environmental challenges (Jiang et al., 2020). Several studies confirm the use of DL techniques in coastal geomorphology. These techniques are used for monitoring coastal erosion (Brose et al., 2025; Blais & Akhloufi, 2025), coastal mapping (Calkoen et al., 2025), and extracting coastlines from remote sensing images (Khurram et al., 2025), detecting and classifying marine morphological features in nearshore areas (Sozio et al., 2024; Sozio et al., 2025), and measuring tides and surges automatically in coastal areas (Sabato et al., 2024).

This Ph.D. research is based on the extensive data archive of three study areas within the Geoswim project (Pelagie islands, Malta and Gozo), including both instrumental and observational data collected along Mediterranean coastal sectors, but the main focus has been on the photographic data. Primary data collection involved capturing dual-media images using GoPro cameras positioned both above and below the water surface, in some cases, along with 360-degree videos and aerial images taken with the support of a UAV.

The main objective is to create 3D models of the rocky coastline, first by obtaining separate models that reflect the landforms above and below the water level, and then by merging the two models. The first step involves establishing a systematic approach to image analysis by organizing the extensive collection of coastal photographs in terms of time and space and developing a tailored swim survey protocol to enhance the pre- and post-survey phases. This foundational step creates a structured framework that is essential for subsequent analytical processes. Building on this organization, to enhance data accessibility and promote collaborative research opportunities, a web-based interface was developed, allowing for synchronized visualization of aerial and submerged images acquired along the same intertidal zone.

A convolutional neural network (CNN) was developed to improve the quality of the dataset by efficiently distinguishing between images useful for constructing 3D models and those containing extraneous elements irrelevant to modeling rocky coasts. The CNN uses images captured during the surveys above and below the waterline.

The capabilities of the network are further extended to perform automated geomorphological classification of coastal landforms, following the classification proposed by Biolchi et al. (2016a). In this step, only the above waterline images were used because they are more suitable for recognizing coastal types, as these images provide better contrast between air and coast, with less blur and no influence from water turbidity (Furlani et al., 2017b, Furlani et al., 2021). The entire dataset was used for reorganizing the dataset and training the CNN.

The final step is the creation of integrated three-dimensional models of the rocky coast. This process synthesizes multiple image sets to produce detailed representations of coastal landform both above and below the waterline. Despite a very large archive of photographic data, it was necessary to identify areas of study within it that would allow the proposed method to be tested for creating 3D models of the coastline along the intertidal zone. The areas were selected based on geomorphological interest (the presence of a tidal notch), the type of coastline (sloping or plunging), and the quality of the images collected above and below the water (GPS signal, not blurred, free of turbidity, and without interruption from other objects).

Thanks to this thesis, this phase demonstrates the practical application of the developed framework in 3D model reconstruction, favouring a new way to study the rocky coast in the intertidal zone, with the possibility to analyze, measure, and store the point zero for future geomorphological surveys in the same areas.

2. Thesis structure

This Ph.D. research focuses on several activities, as shown in Figure 2.1, integrating field geomorphological analysis with technological applications like deep learning techniques.

First, images are collected during a field acquisition phase using a swim-survey method, including both time-lapse images captured by GoPro and continuous video captured by 360 cameras. This phase was realized by standardizing the geomorphological swim-survey protocol.

Second, the curation of the extensive Geoswim dataset is a fundamental step that involves the development of a coherent framework for cataloguing and organizing the vast collection of time-lapse images and associated data.

Third, the development of a web-based visualization system that allows the simultaneous display of above and below water image pairs, integrated with a real-time map.

Fourth, the implementation of artificial intelligence through a CNN for automated image analysis, capable of distinguishing between useful and not useful images for analysis. Another CNN, using the same model, was implemented to perform multi-class classification of coastal morphological features.

Finally, the creation and integration of three-dimensional models covering the intertidal zone, merging above and below the waterline photogrammetric models to address the complex physical constraints of the air-water interface. The combination of the data acquisition method and the technological application to the data creates a functioning system in which the collected data can be checked and validated.

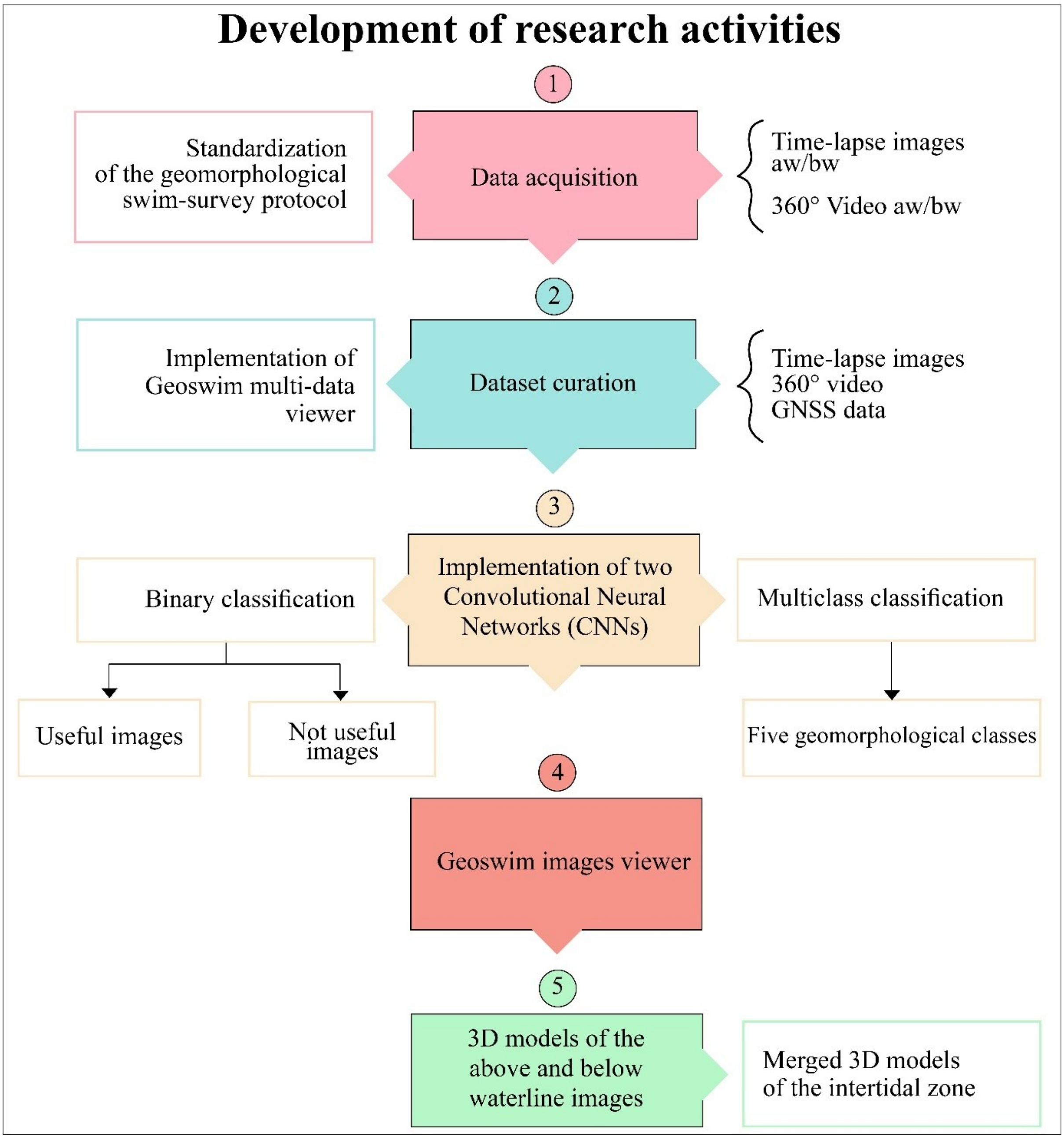

Figure 2.1: Sketch of the PhD work that illustrates the procedure applied for images manipulation. The central boxes from 1 to 5 refer to the main points developed in the thesis, the side boxes summarize what was done for each step, and in curly brackets the type of data.

3. The Mediterranean Sea

The Mediterranean Sea (Figure 3.1) is a complex intercontinental basin located between Europe, Africa and Asia. This semi-enclosed sea is connected to the Atlantic Ocean by the Strait of Gibraltar in the west and to the Black Sea by the Dardanelles and Bosporus Straits in the east, while an artificial connection to the Red Sea has existed since 1869 through the Suez Canal (Millot & Taupier-Letage, 2005). The Mediterranean Sea covers an area of approximately 2.5 million km² and can be divided into two main regions: the western Mediterranean, characterized by extensive abyssal plains and bounded by the Sicilian Channel, and the eastern Mediterranean, dominated by the Mediterranean Ridge System (Mather, 2009).

Some maps identify the Tyrrhenian and Ionian Seas as forming the central Mediterranean region. The Ionian Sea itself covers 616,000 km² and contains the most extensive abyssal plain in the Mediterranean, reaching a maximum depth of 4,270 m in the Calypso Deep off the Peloponnese coast (Millot & Taupier-Letage, 2005).

The Mediterranean coastlines reflect a complex geological history shaped by the long-term interaction between the Eurasian and African plates (Mather, 2009). This tectonic relationship dates back some 250 million years to the fragmentation of Pangea and the subsequent formation of the Tethys Ocean, the precursor to the Mediterranean Sea. The region occupies the western part of the Alpine-Himalayan orogenic belt, and its collisional history has progressively isolated the Mediterranean from other ocean basins. The current geodynamic configuration represents the culmination of several major processes: seafloor spreading along the Mid-Atlantic Ridge, closure of the Tethys, and the final stages of the Alpine orogeny (Mather, 2009). This dynamic history has produced a basin characterized by deformed microplates, mobile zones between these plates that manifest as mountain ranges, and a mosaic of ancient and more recent oceanic crusts.

One of the most dramatic episodes in the geological history of the Mediterranean occurred during the late Miocene (about 6-7 million years ago), when tectonic movements closed the Strait of Gibraltar. This interrupted the connection with the Atlantic Ocean and turned the Mediterranean into a large salt lake. The resulting Messinian salinity crisis led to substantial evaporation and deposition of thick evaporite sequences composed of carbonates, calcium sulphate (gypsum and anhydrite), and halite (Pinet, 2006). The crisis ended when the strait reopened about 5 million years ago, allowing Atlantic waters and marine organisms to repopulate the basin.

The Mediterranean region remains tectonically active, with Africa moving northwards at a rate of about 5 mm per year (Calais et al., 2003; Nocquet & Calais, 2004). This ongoing convergence generates significant seismic activity throughout the basin. The region contains both active subduction zones associated with volcanic activity and older, quiescent subduction zones (Mather, 2009). The geodynamic characteristics of the Mediterranean are dictated by lithospheric blocks that exhibit various structural and kinematic interactions, including collision, subduction, back-arc spreading, and fold-and-thrust belt development. This complex configuration is largely due to the original geometry of the opposing plate margins and the presence of continental blocks within the western Tethys (Jolivet & Faccenna, 2000; Royden & Papanikolaou, 2011; Serpelloni et al., 2007).

The Mediterranean basin experiences a transitional climate characterized by wet winters and dry summers (Bonada & Resh, 2013; Rohling et al., 2009). This climate zone lies between the storm belt of northern Europe and the tropical zone of northern Africa. Although generally mild, significant storms can develop, especially during the winter months (Cavaleri, 2000).

These weather systems originate either from eastward moving Atlantic fronts or within regional cyclogenetic areas such as the Gulf of Lion and the Gulf of Genoa. Summer conditions are dominated by high pressure systems and descending air masses, leading to dry conditions, especially in the southern Mediterranean.

A distinctive hydrological feature of the Mediterranean coasts is the presence of karst groundwater systems (Ford & Williams, 2007). These springs, which can discharge in excess of 1000 liters per second in some locations, represent the outlet points of extensive groundwater pipe networks (Civita, 2008). Their flow regimes typically follow seasonal rainfall patterns but can be dramatically activated during intense precipitation events (Bonacci et al., 2006).

The Mediterranean Sea experiences significantly smaller tidal ranges compared to oceanic environments. The average tidal amplitude is about 40 cm, with exceptions in the Gulf of Gabes and parts of the northern Adriatic, where amplitudes can reach up to 1.0-1.2 m. In contrast, areas near amphidromic points, such as parts of Greece and Sicily, experience minimal tidal variation. Atlantic tidal influences penetrate through the Strait of Gibraltar, but rapidly decrease eastward (Antonioli et al., 2015).

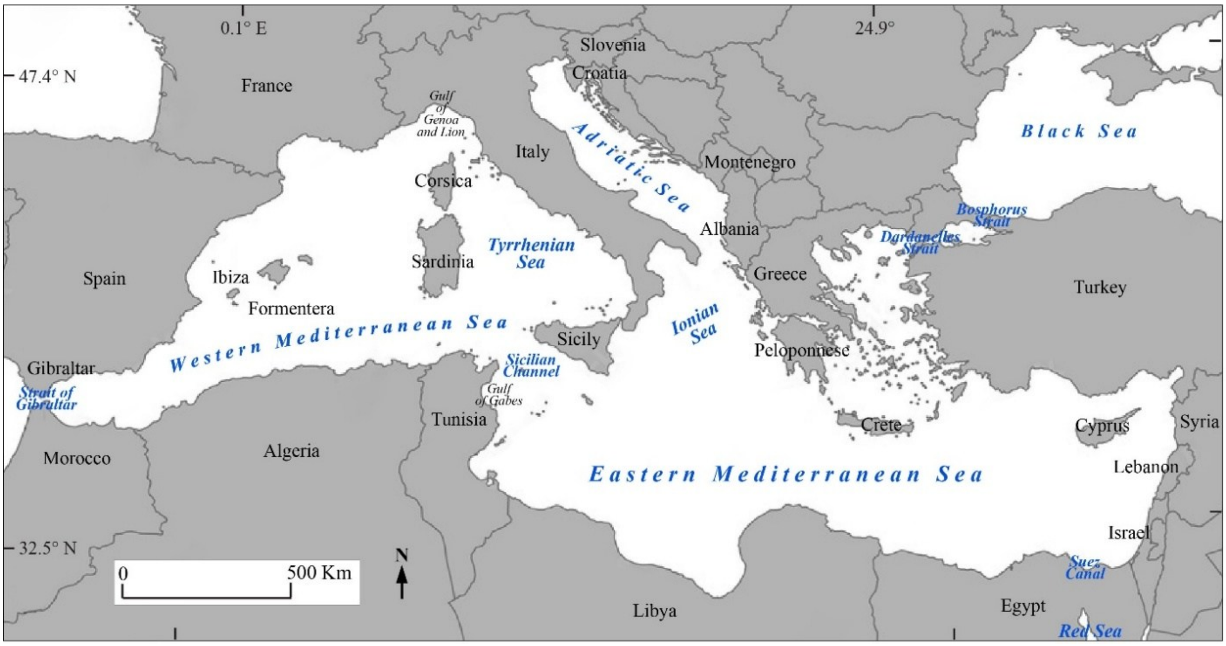

Figure 3.1: Map of the Mediterranean Sea.

4. The rocky coasts of the Mediterranean

More than half of the Mediterranean coastline is dominated by rocky shorelines, ranging from the granite cliffs of Corsica and Sardinia to the limestone karst coasts of the Adriatic and the volcanic shores of the Aegean Islands (Furlani et al., 2014a; Figure 4.1). These coasts develop through the complex interaction of wave action, tectonic processes, and the region's distinctive microtidal conditions. This creates landforms that reflect the unique geological and oceanographic characteristics of the basin. The semi-enclosed nature of the Mediterranean and its limited connectivity with global oceans produce coastal environments with morphodynamic processes that distinguish them from other rocky coastlines worldwide.

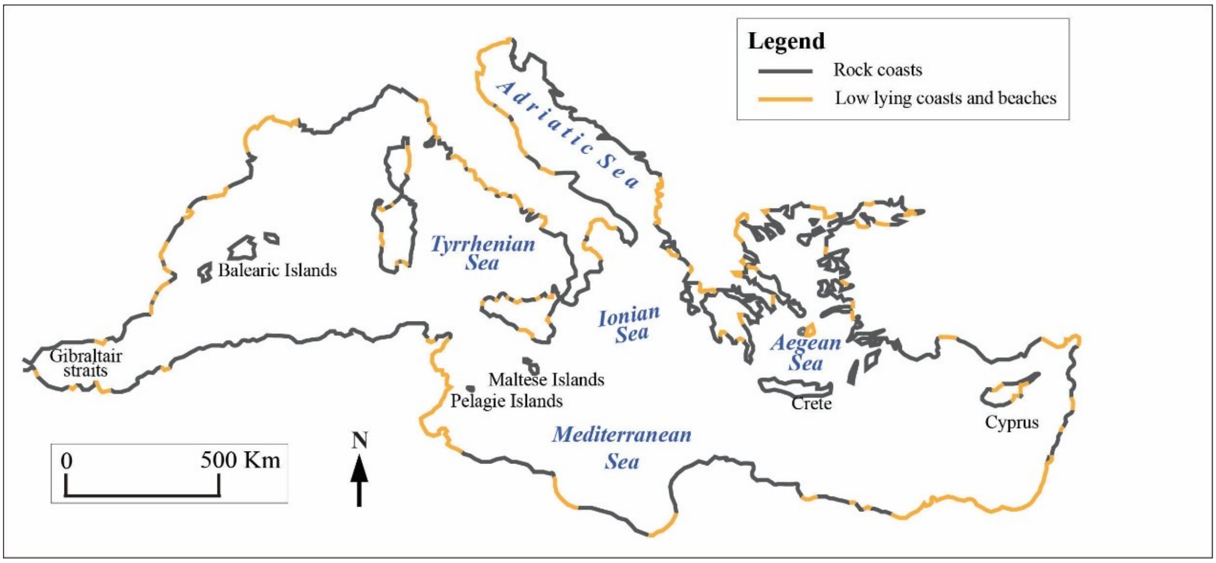

Figure 4.1: Map of the Mediterranean Sea area showing rock coasts and low-lying beaches.

Research on the rocky coasts of the Mediterranean has evolved from early qualitative descriptions of classic sites, such as the Amalfi Coast and the Greek archipelagos, to modern quantitative studies that utilize advanced surveying technologies. Light detection and ranging (LiDAR) techniques and terrestrial laser scanners have proven particularly valuable for documenting the notch systems and shore platforms that are characteristic of microtidal environments in the Mediterranean (Kennedy et al., 2014; Lim et al., 2005). These technologies enable measurements at the centimeter-to-millimeter scale, which are essential for monitoring tectonic coastal movements in seismically active regions of the Mediterranean, such as southern Italy, the Aegean, and the Turkish coast. This quantitative shift in Mediterranean coastal research has revealed the intricate relationship between regional tectonic activity, sea-level fluctuations, and the evolution of coastal morphology (Stephenson, 2015).

The evolution of Mediterranean rocky coasts is the result of several region-specific factors that distinguish these environments from other rocky coastlines around the world. The microtidal regime, with average amplitudes of 40 cm and exceptional ranges reaching 1.0–1.2 m only in the Gulf of Gabès and the northern Adriatic, concentrates wave energy within narrow vertical zones (Antonioli et al., 2015). This creates the pronounced horizontal notches and shore platforms visible throughout the basin. The Mediterranean climate, characterized by wet winters and dry summers, produces distinctive weathering patterns. Intense winter storms from

Atlantic fronts and regional cyclogenesis areas, such as the Gulf of Lions, alternate with prolonged dry periods that enhance chemical weathering in carbonate lithologies (Cavaleri, 2000).

The geological diversity of Mediterranean coastlines, shaped by the region's complex tectonic history within the Alpine-Himalayan orogenic belt, provides varied substrates for coastal landform development. Examples from the Mediterranean illustrate this lithological control: the resistant volcanic rocks of Stromboli and other volcanic islands form plunging cliffs, while the widespread Mesozoic carbonates of the Adriatic and eastern Mediterranean develop extensive, karst-influenced shore platforms. The Tyrrhenian coast exhibits pronounced notch development in limestone sequences. In contrast, the tectonically active Aegean islands display uplifted marine terraces at various elevations, reflecting ongoing vertical crustal movements (Mather, 2009).

The legacy of the Messinian salinity crisis significantly influences present-day Mediterranean coastal morphology. Evaporite sequences, composed of carbonates, gypsum, anhydrite, and halite, were deposited during this late Miocene event. These sequences create distinctive coastal lithologies that are particularly susceptible to chemical weathering and dissolution processes (Pinet, 2006). These evaporite formations contribute to the development of coastal caves, submarine springs, and rapid retreat rates in affected coastal areas.

Tectonic processes profoundly influence Mediterranean coastal morphology, producing diagnostic features such as marine notches at multiple elevations and extensive uplifted coastal platforms. The ongoing northward movement of Africa at a rate of approximately 5 mm per year generates significant seismic activity throughout the basin and creates rapid spatial transitions between subsiding, uplifting, and stable coastal segments (Calais et al., 2003; Nocquet & Calais, 2004). Evidence from ancient shorelines shows that since the last interglacial period (MIS 5.5), Mediterranean coastal evolution has exhibited marked regional variations. Western Mediterranean coasts have generally remained stable over the past 125,000 years, while large areas of Italy, Greece, and Turkey have experienced complex patterns of vertical movement (Antonioli et al., 2009; Anzidei et al., 2014; Ferranti et al., 2006, 2010).

The distinctive hydrological characteristics of Mediterranean coasts also influence the development of rocky shorelines. Karst groundwater systems, capable of discharging over 1,000 liters per second, create submarine springs that affect local coastal erosion patterns and produce distinctive coastal cave systems (Civita, 2008; Ford & Williams, 2007). These springs follow seasonal rainfall patterns, but intense precipitation events can dramatically activate them, contributing to episodic coastal erosion and the development of coastal karst landforms unique to Mediterranean carbonate coastlines (Bonacci et al., 2006).

Regional variations in Mediterranean rocky coast morphology reflect the complex interaction of tectonic, climatic, and oceanographic factors. For example, the gently sloping platforms along the eastern Adriatic coast contrast sharply with the uplifted, steeply sloping cliffs and narrow shore platforms that are characteristic of the Greek and Turkish coasts. The French and Spanish Mediterranean shores exhibit extensive horizontal shore platform development in resistant crystalline rock, whereas the Italian peninsula exhibits rapid morphological transitions related to active tectonics and diverse lithological controls (Antonioli et al., 2015).

The microtidal environment of the Mediterranean Sea fundamentally shapes coastal landform preservation and development patterns throughout the basin. Unlike in macrotidal environments, where extensive intertidal zones develop, Mediterranean rocky coasts concentrate geomorphic processes within narrow elevation ranges. This often results in sharp morphological boundaries and well-preserved relict features from previous sea-level positions.

The combination of this preservation potential with the region's tectonic activity and diverse geological heritage makes Mediterranean rocky coasts invaluable archives of Quaternary environmental change and examples of coastal evolution under boundary conditions unique to semi-enclosed basins.

4.1. Mediterranean coastal landforms

The Mediterranean basin is known for its distinctive coastal landforms, which reflect the unique microtidal regime, complex geological heritage, and ongoing tectonic activity of the region. These landforms distinguish the rocky coasts of the Mediterranean from other coastal environments around the world and provide insights into the evolutionary history of the basin (Furlani et al., 2014a).

Marine notches are one of the most characteristic features of Mediterranean rocky coasts. These horizontal erosional notches, formed by wave action at mean sea level, are well-developed due to the narrow tidal range that concentrates erosive forces within a restricted vertical zone (Trenhaile, 2002). In tectonically active regions like southern Italy, western Greece, and the Turkish coast, multiple notch levels document episodes of relative sea-level change. This creates distinctive, stepped cliff profiles that chronicle eustatic and tectonic movements over millennial timescales (Antonioli et al., 2015; Ferranti et al., 2006).

Coastal karst landforms dominate large areas of Mediterranean rocky shores, especially along the Adriatic Sea and the eastern Mediterranean, where Mesozoic carbonates comprise the coastal substrate (Ford & Williams, 2007). These landforms include coastal dolines, solution pans carved into shore platforms, and extensive networks of sea caves. Submarine karst springs capable of discharging over 1,000 liters per second create distinctive coastal morphologies through focused erosion. These springs represent outlets of extensive underground drainage networks that reflect the complex hydrogeological systems of the region (Bonacci et al., 2006; Civita, 2008).

Vermetid reef platforms are unique biogenic coastal formations found exclusively in warm Mediterranean waters. Built by sessile gastropods (*Dendropoma petraeum* and *Vermetus triquetrus*), these organogenic structures create distinctive, rim-like platforms that extend seaward from rocky shores, particularly along the coasts of Sicily, southern Italy, and the Levantine basin (Chemello & Silenzi, 2011). These reefs modify wave energy patterns, create ecological niches, and contribute to coastal protection by dissipating wave energy.

Volcanic coastal landforms characterize significant portions of Mediterranean rocky shores, reflecting the basin's position within the Alpine-Himalayan orogenic belt (Mather, 2009). The Tyrrhenian coast displays distinctive lava platforms, particularly around active volcanic centers such as Stromboli and Etna. Volcanic tuff cliffs, formed from explosive eruptions, create distinctive coastal morphologies with high erosion rates and characteristic weathering patterns.

Pillow lava exposures along submarine volcanic coasts provide insights into past eruptive environments and sea-level positions (Cas & Wright, 1987).

Uplifted marine terraces are prominent coastal features in tectonically active Mediterranean regions. These step-like platforms were carved during previous interglacial periods and were subsequently elevated by tectonic processes. They are particularly extensive along the Calabrian coast, western Crete, and the Turkish Mediterranean shore (Antonioli et al., 2009; Ferranti et al., 2006). Each terrace level corresponds to a specific interglacial period, creating a natural archive of Quaternary Sea level and tectonic history extending back several hundred thousand years (Anzidei et al., 2014).

Evaporite coastal formations reflect the legacy of the Messinian salinity crisis on Mediterranean coastal morphology (Pinet, 2006). Gypsum and anhydrite cliffs, which are particularly prominent along the Sicilian and Spanish Mediterranean coasts, exhibit rapid erosion rates and distinctive solution weathering patterns. These formations create coastal caves, natural arches, and rapid cliff retreat rates, distinguishing them from more resistant carbonate lithologies.

Coralligenous formations are distinctive Mediterranean bioconstructions that develop on rocky substrates in the circalittoral zone at depths ranging from 20 to 120 meters. Dominated by coralline algae, these communities create complex, three-dimensional structures that modify hydrodynamics, provide coastal protection, and support exceptional levels of biodiversity unique to Mediterranean marine ecosystems (Ballesteros, 2006).

Lithophaga galleries, which are carved by boring mollusks (primarily *Lithophaga lithophaga*), are another distinctive Mediterranean coastal feature, particularly on limestone shores. These bioerosional structures create distinctive honeycomb patterns that significantly accelerate coastal retreat rates and create textural patterns on wave-cut platforms and cliff faces (Devescovi & Iveša, 2007).

These Mediterranean-specific landforms interact to create coastal morphologies that serve as natural laboratories for understanding coastal evolution under distinctive environmental conditions. Their distribution patterns reflect the complex interplay of regional climate, geology, tectonics, and oceanographic processes that characterize this unique semi-enclosed basin (Furlani et al., 2018). The preservation potential of these features in the microtidal environment of the Mediterranean provides exceptional opportunities for reconstructing coastal evolution and environmental change over multiple temporal scales.

4.1.1. Tidal notches

Tidal notches are horizontal depressions in coastal cliffs that typically occur at mid tide and serve as geologic indicators of historical sea level (Trenhaile, 2015). These coastal landforms are characterized as laterally extending depressions at the base of a cliff or at sea level, where the width exceeds the depth (Drago, 2009; Sunamura, 1992; Trenhaile, 2015).

The genesis of these distinctive coastal features is primarily due to increased erosion rates in the intertidal zone compared to adjacent supratidal or subtidal areas (Furlani et al., 2009; Furlani & Cucchi, 2013; Moses, 2013; Moses et al., 2014). Their formation results from a combination of erosive wave action and chemical and biological weathering processes acting on the cliff material (Trenhaile, 2015). In the central Mediterranean, these notches predominantly have a U-shaped profile, as documented by Antonioli et al. (2015), Carobene (2014), and Pirazzoli (1986). Marine tidal notches specifically refer to undercuts ranging from a few centimeters to several meters in depth, carved into steep calcareous cliff faces at or near sea level (Carobene, 1972; Kelletat, 2005; Pirazzoli, 1986; Figure 4.2b, c). When the base or floor of the notch is absent, as observed in locations such as Istria, Croatia (Benac et al., 2004, 2008) and Gozo, Malta (Furlani et al., 2017b), the formation is referred to as a roof notch (Figure 4.2d, e). Both structural types are distributed throughout the Mediterranean coastline, where approximately half of the rocky shores consist of Mesozoic to Quaternary carbonate formations (Furlani et al., 2014b).

Modern tidal notches are particularly well developed throughout the Mediterranean Sea, due to its characteristic microtidal regime, and can be observed along most of its perimeter. The development of these notches in Mediterranean carbonate coastlines involves a complex interplay of processes. Four main mechanisms drive notch formation: biological agents, cyclic wetting and drying associated with salt weathering, hyperkarst processes, and mechanical erosion (Antonioli et al., 2015). In contrast to other global regions, wave abrasion contributes minimally to notch development in the Mediterranean environment (Antonioli et al., 2015; Trenhaile, 2015). Instead, chemical dissolution and biological weathering are the dominant factors shaping sedimentary carbonate shores in this region (De Waele & Furlani, 2013).

A landmark study by Antonioli et al. (2015) examined 73 coastal sites in the central Mediterranean, providing a synthesis of previous research on tidal notches. This work included detailed measurements of notch dimensions, documentation of biological edge characteristics, and established correlations between notch morphology and environmental factors such as wave energy, tidal range, and rock lithology. The researchers postulated that over the past 6.8 thousand years, marine tidal notches have undergone morphological evolution in response to relative sea level rise, contributing to the progressive retreat of sea cliffs (Antonioli et al., 2015).

For methodological approaches to study these coastal features, this research follows the protocols described by Antonioli et al. (2015). Notch dimensions are measured using invar rods, while geographic positions are recorded using GPS devices, with all spatial data referenced to the WGS84 coordinate system. Standard measurements capture key morphological parameters, including mean notch width (vertical extent), inward notch depth (horizontal extent), inward bottom depth (horizontal extent of the notch base, corresponding to the extent of the biological edge, if present), characteristics of the biological edge (thickness and dominant species) at the notch base, if present, and the depth of the cliff toe (Figure 4.2a).

Tidal notches have considerable value as sea level markers and have been used for this purpose in numerous studies (Antonioli et al., 2017; Benac et al., 2004, 2008; Faivre et al., 2011; Pirazzoli et al., 1996).

Two major challenges remain in the study of tidal notches. The first concerns understanding their precise formation mechanisms, which likely combine intertidal chemical dissolution processes, wetting and drying cycles, biological erosion, wave action, or most likely an interaction of several factors (Antonioli et al., 2015). The second challenge relates to dating, as tidal notches cannot be dated directly, requiring researchers to estimate age by dating organisms that make up the biological rim that partially covers the notch (Faivre et al., 2013; Pirazzoli et al., 1994) or by correlation with other datable indicators. This dating limitation has led to an ongoing scientific debate about the origin of notches. The conventional view is that tidal notches form around mean sea level, with any deviation from their characteristic semi-ellipsoidal morphology (Carobene, 1972; Pirazzoli, 1986) or position beyond the tidal range indicating rapid (coseismic or volcano-tectonic) or gradual (due to regional tectonic processes) land movement. However, alternative models suggest that notch formation occurs only during periods of relative climate and sea level stability, when bioerosion processes can match the rate of sea level rise (Cooper et al., 2007; Evelpidou et al., 2012).

In the Mediterranean basin in particular, the interpretation of tidal notches as paleo-sea level indicators presents both opportunities and challenges due to the complex interactions between tectonic, glacio-hydro-isostatic, and eustatic processes that collectively influence relative sea-level fluctuations in the region (Lambeck et al., 2004). This complexity enhances the importance of tidal notches for understanding the dynamic coastal evolution of the Mediterranean during the Quaternary.

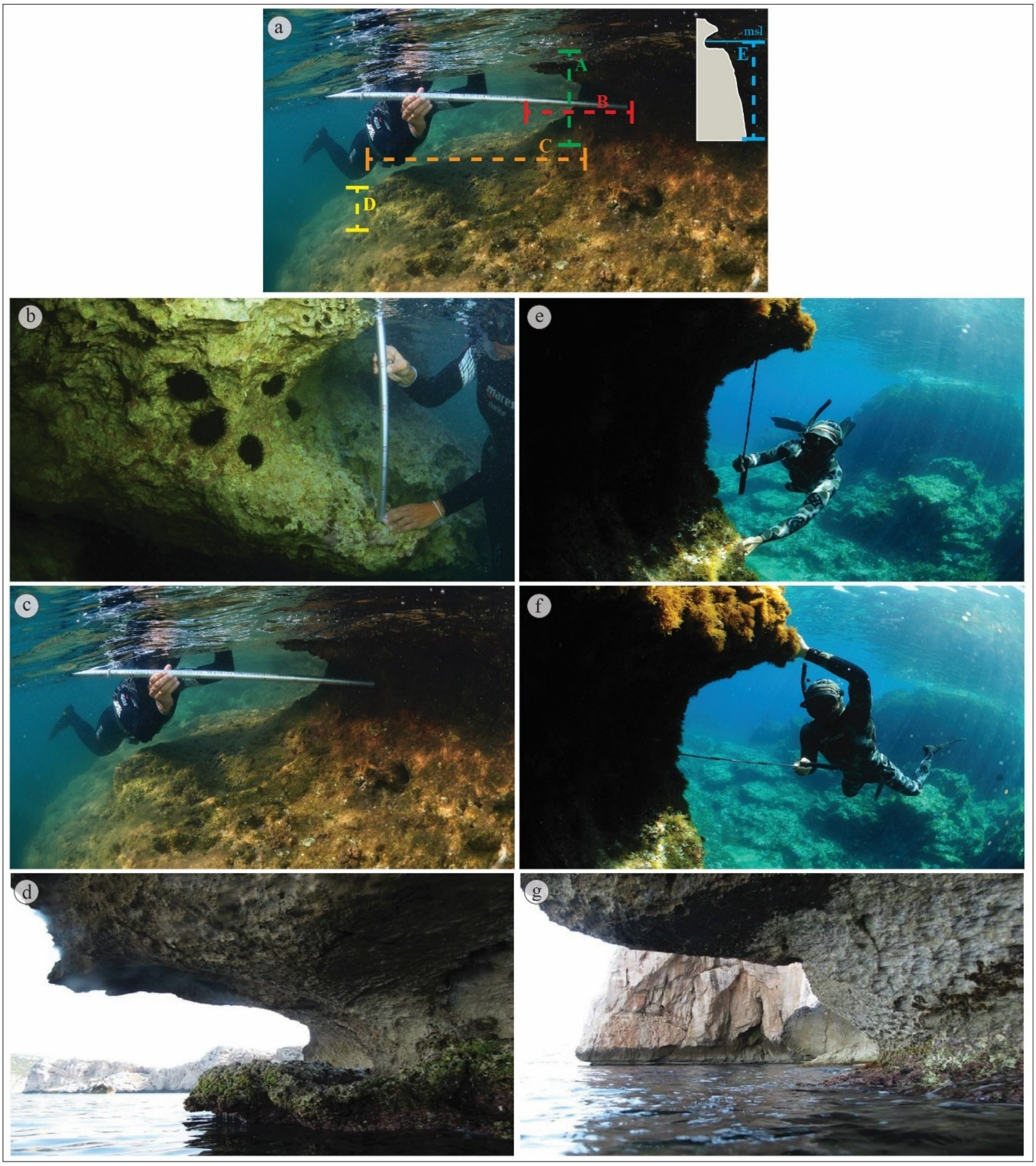

Figure 4.2: Example of U-shaped and roof notch tidal notches. a) Standardized terminology defining the morphological features of notches proposed by Antonioli et al. (2015), where (A) is the average width of the notch, (B) is the inward depth of the notch, (C) is the horizontal extent of the base of the notch, (D) is the thickness of the biological rim, and (E) is the depth of the cliff toe; b) measure of the vertical extent of the submerged marine tidal notch in Lampedusa, Pelagie Islands (A); c) measure of the horizontal extent of the submerged marine tidal notch in Lampedusa, Pelagie Islands (B); d) example of actual C-shaped tidal notch in Capo Caccia, Sardinia; e) measure of the vertical extent of the roof notch in Paros, Greece; e) measure of the horizontal extent of the roof notch in Paros, Greece; f) example of actual C-shaped tidal notch in Capo Caccia, Sardinia.

4.1.2. Bioconstruction

Marine bioconstructions, alternatively referred to as "biogenic reefs" in scientific literature (Ingrosso et al., 2018), represent complex three-dimensional structures of biological origin that enhance biodiversity and regulate the ecological functions in benthic ecosystems worldwide.

Throughout the Mediterranean, these formations manifest as diverse structures, including coralligenous assemblages (Figure 4.3a), vermetid reefs, cold-water coral communities in deep water, *Lithophyllum byssoides* trottoirs (Figure 4.3), shallow-water coral aggregations composed of *Cladocora caespitosa* or *Astroides calycularis* (Figure 4.3c), and reef formations created by serpulid tubeworms. These biological organisms alter the morphological, chemical, and physical properties of their underlying substrates, creating habitats that support numerous organisms and play critical roles in ecosystem processes (Ingrosso et al., 2018). The Mediterranean Sea is characterized by bioconstructions of different configurations and dimensions, distributed throughout its bathymetric range, from surface waters to abyssal depths. Among these formations, vermetid reefs stand out as distinctive structures constructed by colonial snails inhabiting the warmer parts of the Mediterranean Sea. These reefs develop predominantly in coastal regions characterized by seasonal sea surface temperatures that remain above 14°C during winter months and reach approximately 24°C during summer periods (Furlani et al., 2021). Their geographic distribution appears to be limited by a northern boundary near 38° N latitude. Researchers value these formations as excellent biological markers of sea level change over the last two millennia due to their rather restricted bathymetric distribution and considerable extent (Chemello & Silenzi, 2011). These structures built by gastropod are found exclusively in the lower medio-littoral and upper infra-littoral zones of rocky shorelines, with architectural complexity varying according to substrate composition (most elaborate on calcarenites, followed by limestones, dolomites, basalts, and least developed on flysch) (Chemello, 1989; Chemello et al., 1990). The development of fully formed reefs requires an abrasion platform and active surface hydrodynamics, while coastal slope determines their morphology and dimensions, with optimal development occurring at angles between 15° and 40° relative to the horizontal (Chemello, 2009).

Figure 4.3 shows some examples of the most frequently observed organisms during the survey campaigns along the Mediterranean coasts.

Figure 4.3: Organism of the intertidal zone along the Mediterranean coasts: a) *Corallina* red algae (Paros, Greece); b) *Lithophyllum lichenoides* (Lampedusa, Pelagie Islands); c) *Astroides calycularis* (Lampedusa, Pelagie Islands); d) *Lithophaga lithophaga* (Ustica, Sicily); e) *Patella* (Aegadian Islands, Sicily); f) *Echinus* (Paros, Greece); g) *Cystoseira* (La Maddalena Archipelago, Sardinia); h) *Dendropoma* (Naxos, Greece); i) *Mytilus* (Paros, Greece); l) *Balanus* (Lampedusa, Pelagie Islands).

4.2. Geomorphological classification of Mediterranean coasts

Several classification systems have been developed to categorize Mediterranean coastal landforms. These systems have evolved from global approaches to regional methodologies that address the characteristics of this semi-enclosed basin.

Sunamura (1992) established early quantitative approaches to coastal classification by developing a framework centered on the dynamic relationship between wave energy and rock resistance. Using mathematical models, Sunamura predicted erosion rates and coastal evolution through categories based on energy-resistance ratios. This work provided a foundation for understanding coastal morphodynamics applicable to Mediterranean environments with appropriate regional modifications.

Subsequently, Bird (2008) provided a comprehensive morphological classification system with global applicability. This system focuses on descriptive form analysis that considers structural geological influences, variations in rock types, and inherited landforms from past environmental conditions. This system creates broad categories, such as plunging cliffs, cliff-and-shore platforms, and sloping rocky shores, based primarily on observable morphological characteristics identifiable throughout Mediterranean coastal environments.

Recognizing the need for regional specialization, Furlani et al. (2014b) developed a specific classification for the Mediterranean. Their approach incorporates distinctive Mediterranean factors: microtidal conditions with limited tidal ranges (average 40 cm), wave environments limited by fetch and dominated by seasonal storm events, carbonate-dominated geology with extensive karst processes, a complex Quaternary sea-level history creating multiple inherited erosional levels, and significant human impact spanning millennia of coastal modification.

Biolchi et al. (2016a) refined the practical implementation of Mediterranean coastal classification by developing a framework designed specifically for systematic coastal surveys and quantitative spatial analysis. Their methodology establishes seven distinct geomorphotypes optimized for Mediterranean conditions: plunging cliffs, sloping coasts, shore platforms, pocket beaches, cliffs with shore platforms, screes, and built-up coasts. This classification system provides clear, quantitative thresholds, including specific slope-angle ranges (0–5° for platforms, 5–45° for sloping coasts, and >45° for cliffs), platform width criteria, water depth parameters, and anthropogenic modification percentages.

Plunging cliffs (Figure 4.4a) are characterized by vertical or sub-vertical faces (>80° slope) that drop directly to depths exceeding 5 meters below low tide without developing shore platforms or transitional zones (Biolchi et al., 2016a). They typically occur in resistant rock formations, such as the limestone cliffs of the Amalfi Coast or the granite shores of Corsica, where high wave energy meets deep nearshore waters.

Sloping coasts (Figure 4.4b) are rocky coastlines with intermediate gradients between 5° and 45°. They often develop in areas with moderate wave energy or variable rock resistance (Biolchi et al., 2016a). These transitional morphologies are common along the volcanic shores of the Aeolian Islands and parts of the Turkish Mediterranean coast.

Shore platforms (Figure 4.4c) are horizontal to sub-horizontal rock surfaces with a slope of 0–5° that extend seaward from cliff bases. They are formed through marine erosional processes, including wave action, biochemical weathering, and biogeochemical dissolution (Biolchi et al., 2016a; Trenhaile, 2002). In Mediterranean contexts, these platforms often exhibit well-developed tidal notches at mean sea level due to concentrated erosional energy within the narrow tidal range.

Pocket beaches (Figure 4.4d) consist of sediment accumulations (sand, gravel, or pebbles) confined within rocky embayments. They are important sediment environments within predominantly rocky Mediterranean coastlines (Biolchi et al., 2016a). These features are particularly common along the French Riviera and the Balearic Islands.

Cliffs (Figure 4.4e) with shore platforms combine vertical cliff faces with basal platform development. This represents a morphological continuum influenced by lithological resistance and wave energy patterns (Biolchi et al., 2016a). This morphotype is widespread along the Adriatic limestone coasts and parts of the Greek archipelago.

Screes (Figure 4.4f) are characterized by extensive deposits of rock debris resulting from gravity, often associated with rapid mass wasting in areas where steep coastal slopes meet the sea (Biolchi et al., 2016a). These features are particularly prominent along tectonically active Mediterranean coasts that are subject to seismic triggering of slope failures.

Built-up coasts (Figure 4.4g) are anthropogenically modified shorelines where the natural morphology has been significantly altered by coastal engineering, urban development, or industrial activities (Biolchi et al., 2016a). These environments are increasingly prevalent throughout Mediterranean coastal areas, reflecting centuries to millennia of human coastal occupation and modification.

The systematic application of these classification schemes enables the quantitative analysis of Mediterranean coastal morphology and facilitates comparative studies across the diverse geological basins and environmental settings. The regional specificity of Mediterranean classification frameworks accounts for the unique processes and boundary conditions of this coastal environment, providing essential tools for coastal research, management, and hazard assessment (Furlani et al., 2018).

For this thesis, Biolchi et al.'s (2016a) classification was used. However, only five classes were considered. The "cliff" and "built-up coast" classes were excluded because they were not present in the studied areas.

Figure 4.4: Coastal geomorphotypes: a) plunging cliff (Siracusa, Sicily); b) sloping coast (Lampedusa, Pelagie Islands); c) shore platform (Banjole, Croatia); d) pocket beach (Lampedusa, Pelagie Islands); e) cliff (El Hierro, Canarias); f) scree (Lampedusa, Pelagie Islands); g) built-up coast (Brač, Croatia).

4.3. The Geoswim programme

The Geoswim program started in 2012, with the first one-man expedition around the Istrian coast to conduct a comprehensive geomorphological survey of the rocky coasts, with particular emphasis on the intertidal and nearshore zone (Furlani, 2012; Furlani et al, 2014b). The initiative was developed with the primary objective of collecting continuous data, specifically images and observational information, using a snorkelling approach along the rocky coasts of the Mediterranean basin (Furlani et al., 2014a). This methodology has proven to be effective for the detailed study of kilometres of coastline and the identification of specific coastal features, such as tidal notches (Furlani et al., 2014b, 2017a, 2017b, 2018), which are considered among the most reliable markers of current and past sea level due to their close relationship with intertidal processes, including wave action, biological activity, and chemical weathering (Antonioli et al., 2015).

The rocky coasts of the Mediterranean provide an ideal environment for the Geoswim approach due to their favourable wave and tidal conditions (Furlani et al., 2014a; Morucci et al., 2016). Surveys typically cover large stretches of coastline, up to 10 km per day, following a standardized protocol described by Furlani (2020). Since its inception, the program has expanded significantly, with Furlani and Antonioli (2023) documenting more than 20 expeditions conducted between 2012 and 2022, covering more than 530 km of Mediterranean coastline (Figure 4.5).

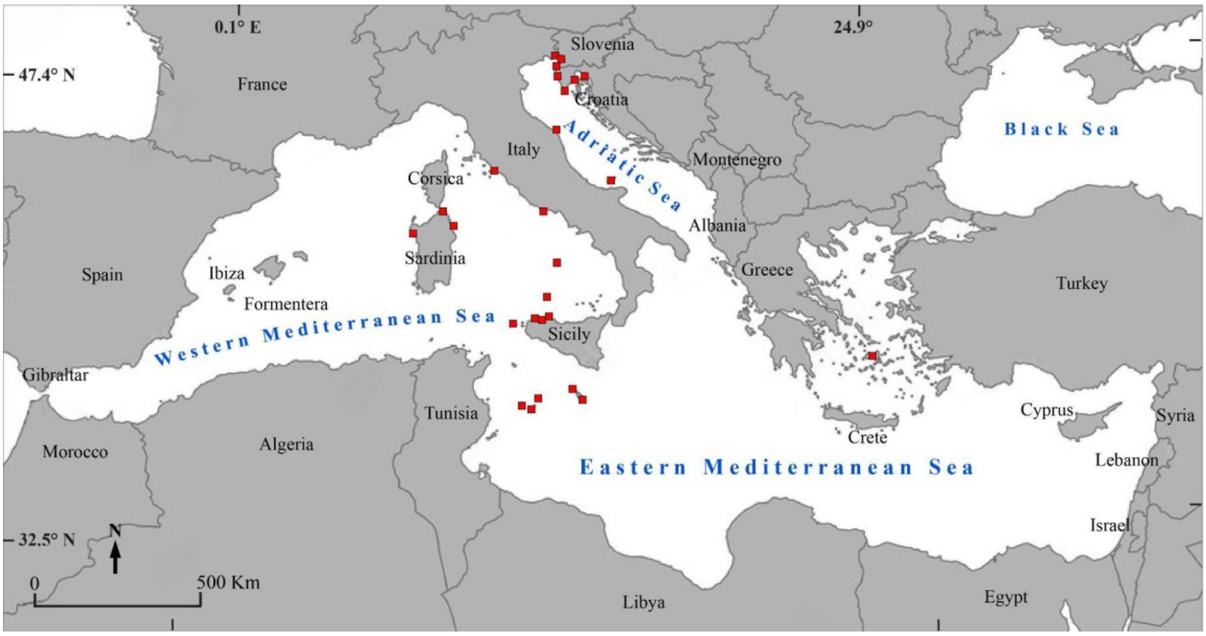

Figure 4.5: The areas surveyed in the Mediterranean Sea within the Geoswim programme (Modified from Furlani and Antonioli, 2023).

The project focuses on specific morphotypes of rocky coasts, mainly plunging cliffs or sloping coastal formations (Biolchi et al., 2016a, 2016b). Researchers conducting these surveys observe and document lateral changes in coastal landforms, paying particular attention to tidal notches, submarine freshwater springs, and ecological indicators such as the presence of barnacles, limpets, worm shells, and various red algal species. In addition to observational data, the Geoswim methodology includes the collection of instrumental measurements during coastal surveys. These include time-lapse images of the surveyed coastline, photos documenting significant landforms or other notable features, continuous video documentation, electrical conductivity (EC) and temperature (T) measurements to identify submarine freshwater springs, and bathymetric readings to determine the depth of cliff bottoms or to create transects perpendicular to the coast (Furlani, 2020; Furlani & Antonioli, 2023; Figure 4.6).

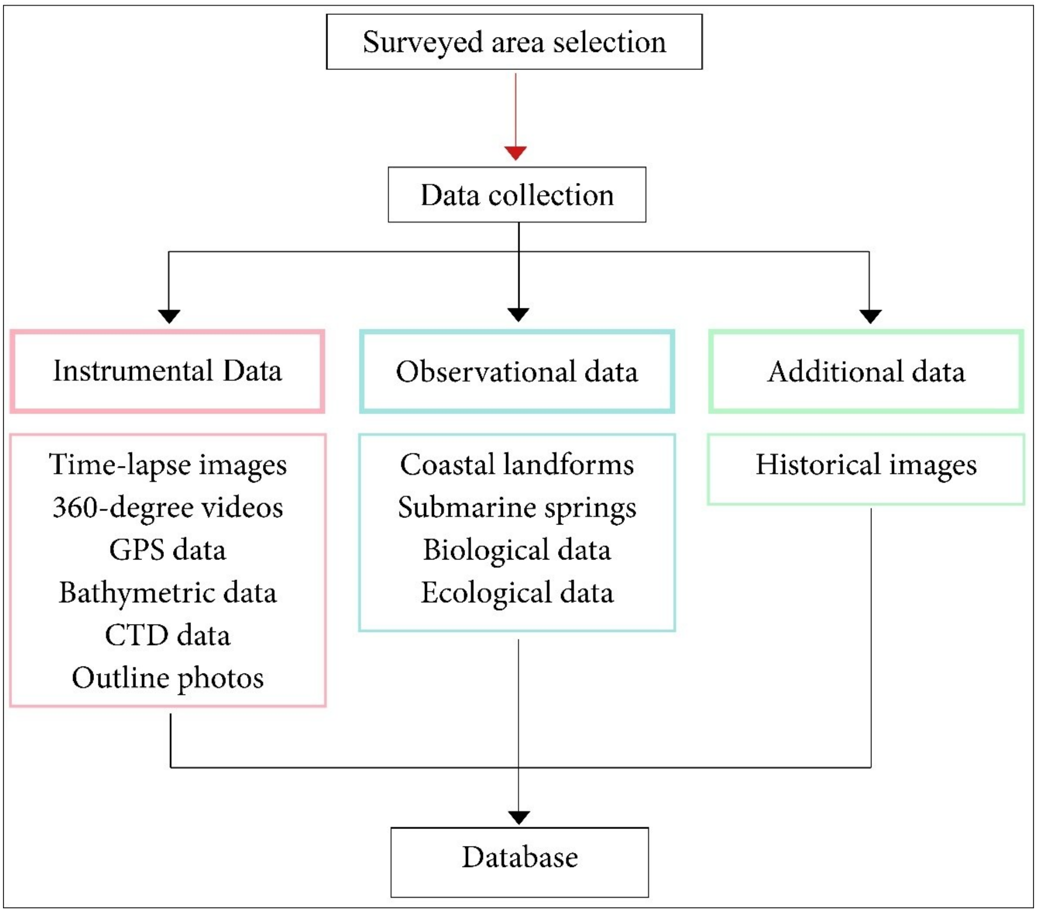

Figure 4.6: Summary of the data collected during the Geoswim survey (modified from Furlani and Antonioli, 2023).

Geoswim surveys are typically conducted by a team of five or six underwater researchers and assistants, though they can also be conducted by a single researcher. A survey team is necessary, especially for long stretches, where some researchers work at sea while others provide support from land or a boat.

The technical equipment used during these surveys has been extensively described by Furlani (2020). The instrumentation, including GPS units, cameras, CTD sensors, and echosounders, is mounted on a specially designed raft known as an Instrumental-Supported Raft (ISR). Depending on the expedition objectives, all or a subset of these instruments may be mounted on the raft. One or more cameras housed in waterproof enclosures continuously capture videos and time-lapse images of the coastline. These cameras are positioned on the side of the raft to take time-lapse images at intervals of one second or less in the direction of travel.

These images collected during these surveys are used to construct 3D models of the surveyed coastline (Furlani et al., 2020). In geomorphological studies, these images often contribute to the development of 3D models both above and below the waterline (e.g., Abadie et al., 2018; Barker et al., 1997; Chandler, 1999; Fawcett et al., 2019; Fraser and Cronk, 2009).

With respect to T and EC measurements, data collected during surveys at depths ranging from -0.15 m to -0.50 m allow researchers to identify freshwater discharges from submarine springs (e.g., Furlani et al., 2014a, 2014b, 2017a). These springs originate from artesian groundwater flow in response to pressure gradients in aquifers and fractures in confining layers (Moore et al., 2009). Several submarine springs have been documented by Furlani et al. (2014b) in Istria (Croatia) and by Furlani et al. (2017a) on the islands of Gozo and Comino in the Maltese archipelago through the analysis of lateral variations in T and EC along the coastline.

The Geoswim program has systematically documented the coastal landforms at several Mediterranean sites (Furlani, 2020), creating a huge dataset that allows detailed regional comparisons. By studying the distribution and characteristics of coastal morphologies in the Mediterranean context, evolutionary models have been developed that take into account the combination of geological, climatic and oceanographic conditions in the region studied.

5. Materials and methods

This chapter presents the methodological framework developed for the acquisition, processing, and analysis of photogrammetric data in rocky coastal environments, with a particular focus on the intertidal and nearshore zone where air and water interfaces converge. This methodology combines geomorphological swim survey techniques with computational approaches to overcome the difficulties associated with coastal monitoring and analysis.

The chapter begins with pre-survey planning protocols and field data acquisition procedures (Sections 5.1.1-5.1.2), detailing the specialized equipment configuration and survey parameters optimized for the Geoswim programme. Section 5.1.3 outlines the systematic post-survey data management protocols essential for maintaining research continuity and data integrity across multiple field campaigns.

Section 5.2 introduces the Geoswim Image Viewer development for double images visualization. Sections 5.3-5.5 address the implementation of artificial intelligence for coastal analysis, covering the theoretical foundations of Convolutional Neural Networks (Section 5.3.2), the specialized ResNet architecture used in this research (Section 5.3.3), training methods (Section 5.4), and Explainable AI approaches (Section 5.5).

The final sections focus on three-dimensional reconstruction techniques, introducing the Structure from Motion and Multi-View Stereo (SfM-MVS) workflow (Section 5.6.1) and the critical underwater camera calibration process (Section 5.6.2). Special attention is given to the technical challenges arising from the different physical properties of air and water, in particular the effects of refraction and lighting conditions on image acquisition and processing.

Additionally, Sections 5.6.3 and 5.6.4 provide an overview of the software used in the post-processing phase.

5.1. Swim survey protocol

This thesis describes the protocol used for data collection during the swim surveys, which was designed to meet its main goal. The detection method is based on the Geoswim program but has been adapted to detect photographic data with optimal settings,

Each swim survey expedition consists of three main phases. The initial setup phase includes choosing the survey area, setting up logistics, selecting the survey team, choosing the equipment, and briefing the team on the precautions to take during the survey. The second phase defines the photogrammetric acquisition process. The final phase includes activities at the end of each survey, such as data archiving and subsequent research analysis.

5.1.1. Pre-survey planning

The pre-survey planning includes both the selection of the study areas and the environmental assessment (e.g. meteo-marine conditions, waves, wind) to ensure optimal data collection.

Equipment preparation and positioning are the main components of the pre-survey protocol. The dual recording system consisted of GoPro 6 and Insta360 cameras mounted above and below the waterline at equal distances from the centre of the raft (Figure 5.1). These units operate in time-lapse mode with GPS functionality enabled. Regular monitoring of battery life at approximately 2-hour intervals and memory capacity ensures uninterrupted data collection.

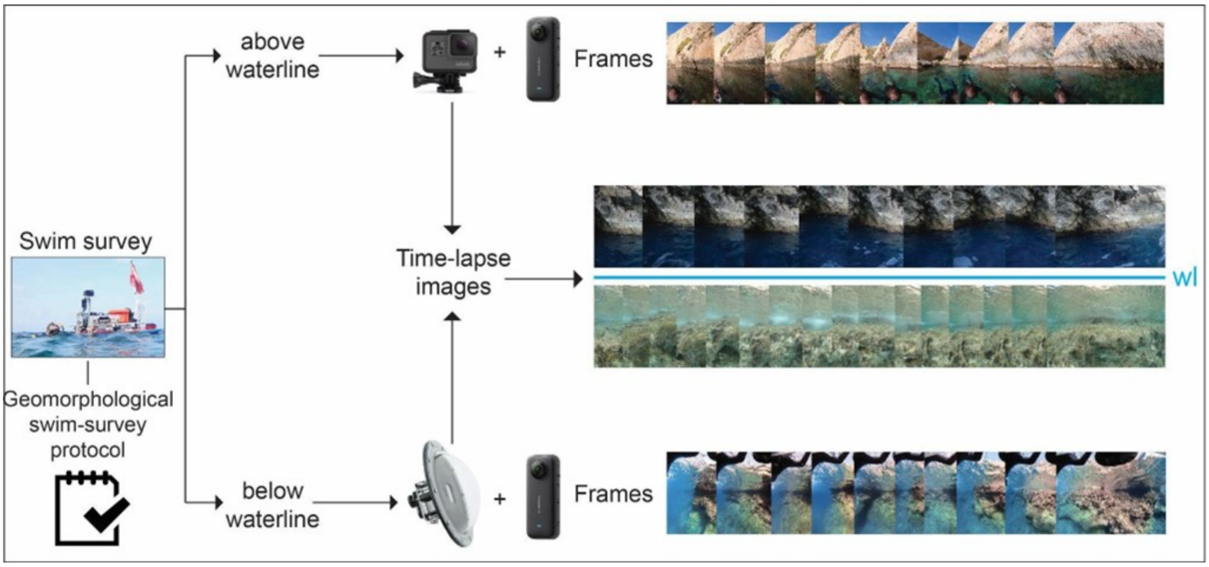

Figure 5.1: Instrumentation setup for photogrammetric data acquisition during swim surveys. Setting of the ISR with the dual camera systems: GoPro Hero 6 cameras mounted 0.5 m above and below the waterline for time-lapse imaging (2 frames/second) and Insta360 cameras for simultaneous 360-degree video capture.

The initialization phase includes a few steps that affect the accuracy of the survey. Full camera connectivity must be established before the survey begins, while GPS waypoint marking at survey endpoints enables accurate tracking. The comprehensive equipment checklist for field surveys is detailed in Table 5.1 and ensures that all necessary tools are properly prepared and functional.

Survey operations require precise control of the raft during the survey phase, maintaining a consistent, moderate speed provides optimal overlap between time-lapse images (minimum 75%). Safe navigation requires maintaining several meters from the coasts to cover and detect an adequate stretch of coastline, while minimizing the distance along plunging cliffs.

In configurations where a second raft is used, it serves as a platform for additional instrumentation, enabling detailed coastline observations and GPS point collection. This dual raft configuration requires strategic task distribution to avoid data redundancy while ensuring continuous, comprehensive image acquisition.

Table 5.1: Check list of the equipment for the swim survey expeditions.

Instruments	**Options**	**Check**
GoPro aw/bw	Time/Date	Y/N
	SD	Y/N
	GPS	Y/N
	Battery	Y/N
	Dome	Y/N

	Charger	Y/N
	Synchronizator	Y/N
360-degree camera	Time/Date	Y/N
	SD	Y/N
	Charger	Y/N
	Battery	Y/N
CTD Divers		Y/N
pH Divers		Y/N
Echosounder	Battery	Y/N
	Cables	Y/N
Others	Garmin	Y/N
	Torch	Y/N
	Radios	Y/N
	Metric rod	Y/N
	Iron bar	Y/N
	Field booklet	Y/N
	Pencil	Y/N
	Flag	Y/N
	Calibration grid	Y/N
	Extendable metre	Y/N

5.1.2. Photo acquisition

Data were acquired through three systematic field campaigns from September 2020 to October 2023, targeting three sectors of the Mediterranean rocky coastline. A total of nine terabytes of images were collected.

Weather and sea conditions are basic to the success of the survey, requiring a thorough weather assessment with attention to wind direction, currents, waves and atmospheric conditions. These factors have a significant impact on data quality, especially for image and video capture. While turbidity from rough seas affects image clarity and wave action can distort photographic subjects, cloudy conditions often provide superior image quality through diffused lighting compared to direct sunlight, which can create difficult exposure variations.

The ISR was central to the data collection strategy, enabling precise positioning and stable image acquisition while maintaining the necessary proximity to coastal features for high-resolution documentation. Made of ethylene vinyl acetate (EVA), the ISR's dimensions (1.15 m length, 0.57 m width, 0.13 m height) and minimal weight (3 kg) were carefully optimized for manoeuvrability while maintaining stability during data collection. The construction material takes advantage of its low density (approximately 0.98 g/cm^3) and UV resistance. The raft's unsinkability certification by the manufacturers further enhanced its reliability for marine surveying operations (Furlani et al., 2021).

The instrumentation configuration consisted of two GoPro Hero 6 cameras mounted in waterproof housings and strategically positioned in vertical orientation on the side of the ISR (Figure 5.2). The above-water camera was positioned 0.5 m above the waterline, while its underwater counterpart was mounted 0.5 m below, both oriented perpendicular to the coastline. These cameras were set to capture time-lapse images at 0.5 second intervals. The underwater camera was equipped with a special 6-inch diameter TELESIN plastic dome optimized for superior optical performance in seawater conditions (Furlani et al., 2021). The technical specifications of the GoPro Hero 6 cameras include a maximum resolution of 3000x4000 pixels, CMOS sensor technology, and a field of view ranging from 15 mm to 30 mm. The integrated GNSS receiver in the GoPro Hero 6 uses a u-blox M8 chipset, providing multi-constellation positioning capabilities including GPS, GLONASS, Galileo, BeiDou and QZSS, with a typical positioning accuracy of approximately 2.5 m.

In all three case studies, the photogrammetric capabilities were further enhanced by the integration of two Insta360 degree cameras providing 5.7 K spherical capture with advanced motion stabilization technology. In the third case study, the standard raft-based images were also integrated with aerial images captured by a DJI Air 2S drone with a resolution capacity of 5472x3648 pixels.

Field surveys were conducted by swimming at a consistent speed of 1-2 km/h along predetermined routes. No motors were used on the raft during the snorkel surveys. The distance from the coast was dynamically adjusted between 1 and 7 meters, responding to local topographic variations and ensuring optimal documentation of lateral variations in coastal erosional landforms. This methodological approach facilitated the comprehensive documentation of various coastal features, including tidal notches (Furlani & Antonioli, 2023). The resulting dataset provides a robust baseline for future comparative analyzes (Furlani, 2020) and establishes a systematic framework for monitoring temporal changes in coastal morphology.

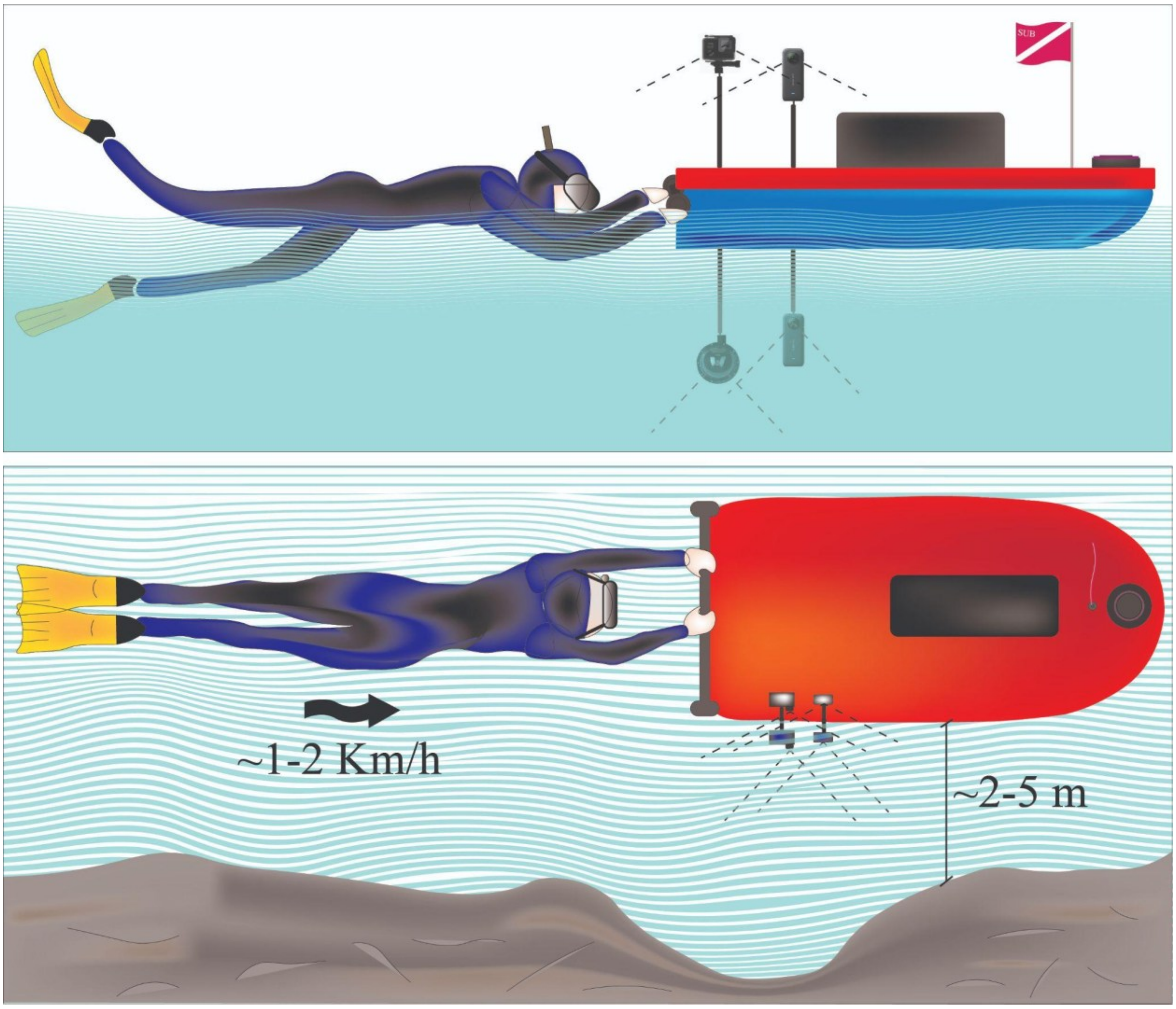

Figure 5.2: The small raft typically used configuration; a) side view of the raft with 2 GoPro cameras and 2 360-degree cameras, both above and below the waterline; b) top view of the raft, showing how the surveys are typically conducted, swimming at 1 to 2 Km/h and 2 to 5 m maximum far from the coastline.

5.1.3. Post-survey data management

Systematic management of field equipment and data is fundamental for maintaining research continuity and data integrity. Immediate post-survey equipment maintenance requires thorough cleaning and appropriate storage of all instruments. Storage conditions for sensitive components, especially SD cards and batteries, must maintain controlled temperature and humidity levels to prevent data corruption and equipment degradation.

Data collection and organization follows a standardized workflow to ensure consistency across field survey. The primary data set includes multiple data streams: time-lapse images from above and below the waterline, 360-degree videos, CTD diver measurements (high and low positions), bathymetric sonar data, GPS waypoints and tracks, and field observations. Daily data transfer to local storage systems prevents data loss and allows immediate quality assessment.

Hierarchical data structure implements systematic organization through standardized naming conventions that include day codes and dates. Time-lapse images are categorized by survey area, with separate categorization for above waterline (aw) and below waterline (bw) images, including a separate archive for temporarily excluded images. Files downloaded from the 360-degree cameras are saved as videos that can be read using the Insta360 Studio program. Raft observations and UAV-acquired images are stored separately, maintaining their original resolution and metadata.

Spatial reference data, consisting of waypoints and survey paths, follow the data code sector format convention, allowing for efficient cross-referencing with other data sets.

This approach to managing post-survey data ensures data integrity and facilitates analysis. It also ensures that equipment is ready for future surveys. The archive structure supports long-term storage of the data and enables easier collaboration between research team members.

5.2. Geoswim Image Viewer

The Geoswim Image Viewer provides a unique solution for simultaneous visualization and analysis of aw and bw images.

Geomorphological survey methods often involve separate analysis of terrestrial and underwater images, leading to potential gaps in understanding the holistic nature of these environments (Micheletti et al., 2015; Westoby et al., 2012). The Geoswim project aims to bridge this gap by providing a unified platform for image comparison.

The main objectives of the Geoswim Image Viewer are to facilitate the seamless comparison of aw and bw images, to integrate geolocation data for spatial context, to preserve and present relevant metadata for each image pair, to provide an intuitive interface for researchers across different devices, and finally to optimize the management of large datasets for efficient analysis.

The Geoswim Image Viewer system is based on ad hoc architecture that combines efficient data processing with a user-friendly front-end. The following section describes the main components and their interactions.

5.2.1. Key functions

The structure used to create the viewer described in this paragraph highlights the functionality of the system created to simultaneously visualize the images above and below the waterline, including a map with the geolocation and a panel with the EXIF data associated with the displayed images.

The Figure 5.3 describes all the steps of the Geoswim image system. The first step to implement the multi-image visualization is to execute a Python script that performs several essential preprocessing operations on each image. This script handles image resizing and compression to optimize file size while maintaining visual quality, embeds a watermark to establish provenance, and crucially, preserves the EXIF metadata from the original files.

EXIF (Exchangeable Image File Format) is a standardized specification for image metadata embedded within digital photograph files. Developed in the early 1990s by the Japan Electronic Industries Development Association (JEIDA), EXIF data contains critical technical information about the image capture conditions, including camera settings (aperture, shutter speed, ISO), timestamp information, geolocation coordinates when available, camera and lens model identifiers, and even thumbnail previews. For scientific visualization applications, preserving this metadata is essential as it provides contextual information about image acquisition parameters that may impact subsequent analysis and interpretation. Maintaining the integrity of EXIF data ensures that no valuable information is lost during the pre-processing workflow.

To efficiently manage the large Geoswim dataset, the system implements parallel processing, which significantly reduces the processing time using all available CPU cores.

Image compression and resizing uses the Pillow image manipulation library, with a 0.25 resize ratio to balance detail preservation and file size, using the Lanczos filter for high-quality image resizing, and JPEG compression at 80% of quality to further optimize file size. The addition of the semi-transparent watermark "GEOSWIM" to each image (opacity 40%) for minimal image obstruction, both above and below the waterline, was implemented with a function that creates a repeating pattern across the image, ensures visibility without compromising image details. The final step before processing the images was to extract the EXIF data from the original images using the 'piexif' library. This ensured that all relevant metadata, including GPS coordinates, were transferred to the processed images.

To geolocate the images on the map, Geoswim Image viewer uses the Leaflet.js library (*Leaflet*, 2024) which is one of the most used JavaScript open-source library for mobile interactive maps, allowing for real-time display of image locations and interactive exploration of the study area. Usually, maps are loaded lazily to optimize the initial load time and reduce unnecessary data transfer. Leaflet.js supports loading only the visible area and nearby areas of the map.

To further improve network performance, even on slower connections, we implemented progressive image loading in this work. This approach enables the simultaneous loading of image pairs and provides better control over the process. Since image loading consumes a lot of bandwidth, a client-side caching system was implemented to store metadata and EXIF data. This reduces the need for repeated server requests and improves response times for frequently requested data. While users browse the image collection, the viewer uses a smart preloading system to minimize waiting times. When a user moves forward through the sequence (for example, from image 1 to image 2), the system automatically loads images 3 and 4 into the browser cache in the background. Similarly, when navigating backward, it preloads the preceding images. This approach ensures smooth navigation through the large dataset of high-resolution image pairs, even with limited bandwidth. The preloading happens invisibly to users, allowing them to focus on comparing the above and below waterline images without interruption.

As shown in Figure 5.3, the images have been processed and copied to the web server folder, where the three core components for the web display of the images are located. The first is the HTML page (index.html) that defines the layout of the image display, the map integration, and the metadata presentation. The second is the CSS (styles.css), which implements a responsive design using flexbox, a very popular CSS property used to place elements of our web page conveniently and quickly inside a parent box, and media queries. The third is the JavaScript commands (script.js) that dynamically manage content updates, image navigation, and user interaction, implementing asynchronous loading techniques for fluid performance.

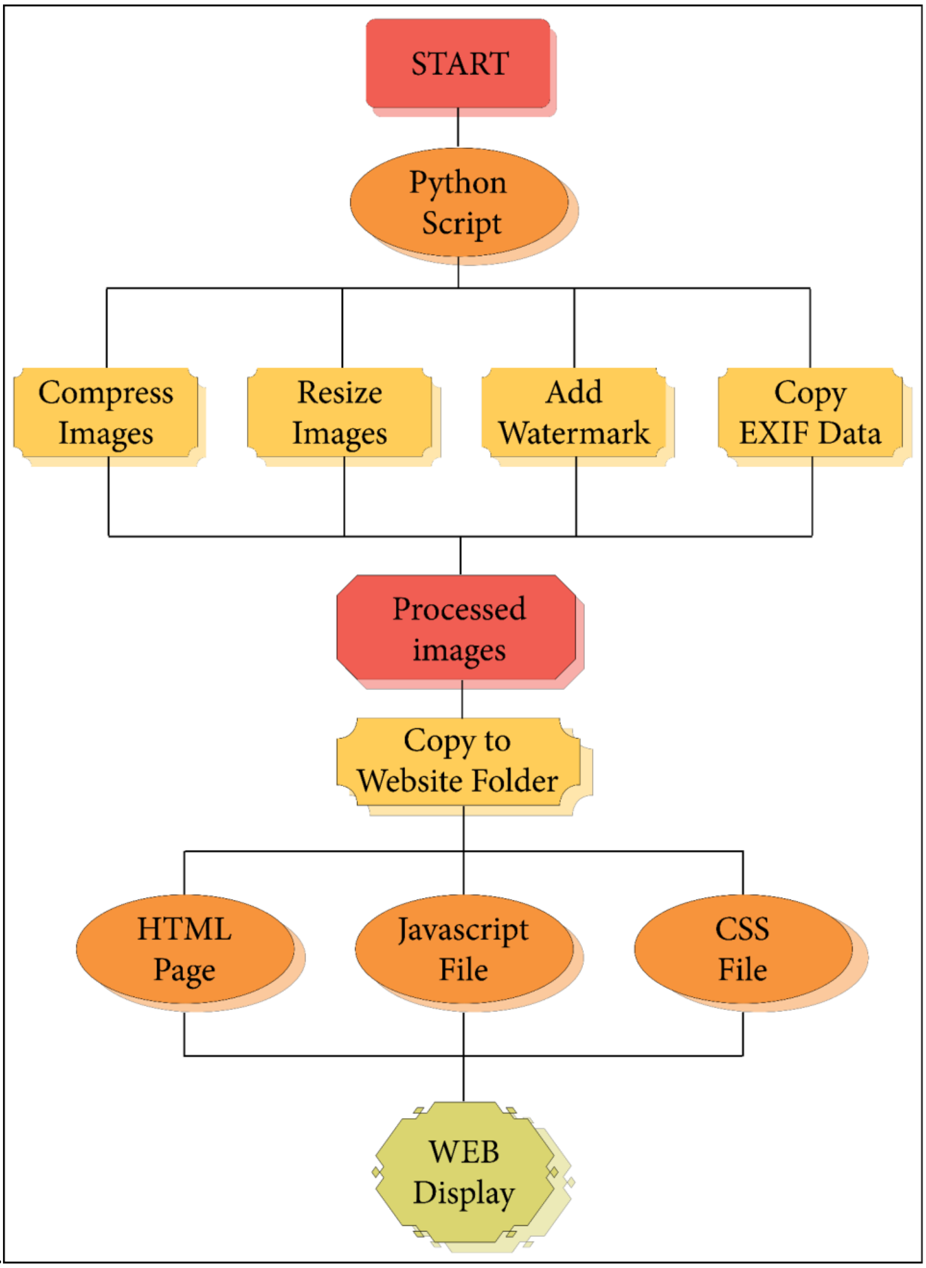

Figure 5.3: Flowchart showing step-by-step how the image processing system works.

5.3. Artificial Neural Network (ANN)

5.3.1. The structure of an ANN

Artificial Neural Networks (ANNs) are computational systems inspired by the structural and functional organization of the cerebral cortex, particularly its neurons and synapses (Goodfellow et al., 2016). The objective of these systems is to emulate the processing of information by biological neurons. This is achieved by accepting multiple input signals from neighbouring neurons, processing them, and producing an output signal in the form of a real or binary number (Zupan, 1994). Through this processing, ANNs are capable of inferring rules and patterns to facilitate decision-making processes.

In biological systems, a neuron represents the fundamental cellular unit of the brain's nervous system. It receives and combines signals from other neurons via dendrites, which contain synaptic junctions. The junctions contain neurotransmitter fluid that controls the flow of electrical signals (Uhrig, 1995). These connections are dynamic, undergoing changes in strength over time. New connections may form, while others may weaken or disappear (Nagyfi, 2018).

The node is the central processing element of the artificial neuron, and it contains an activation number. Upon receipt of signals, the artificial neuron processes them and produces an output that is transmitted to connected neurons in the subsequent layer. Each connection, or edge, is assigned a weight, which is a real value representing the strength of the connection during processing. Additionally, neurons may possess an associated bias parameter, denoted by the letter b.

The neuron's output can be expressed mathematically as shown in (1), where y represents the neuron's output, b is the neuron's bias, x_i is the input signal from the i-th edge, w_i is the weight of the i-th edge, and f is the activation or transfer function:

$$y = f(\sum_i w_i x_i + b) \tag{1}$$

Figure 5.4 illustrates the parallel between biological and artificial neurons (Nagyfi, 2018). The artificial neuron processes information through two principal functions. Initially, a linear combination of the input variables (x_0, x_1, x_i) is multiplied by their respective weights (w_i). Subsequently, an activation function transfers the signals to other neurons (Zupan, 1994).

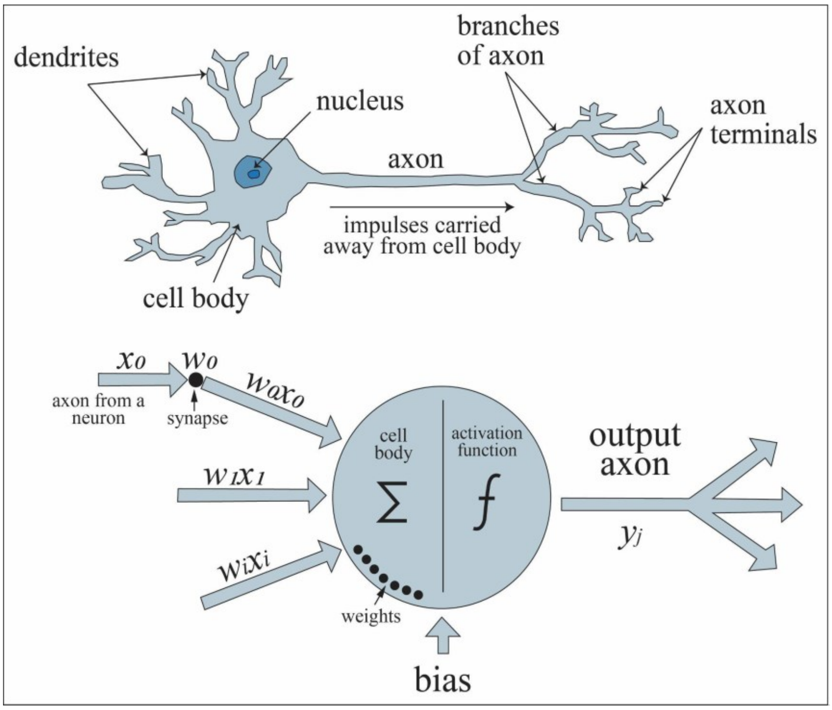

Figure 5.4: Comparison between biological and artificial neuron (modified from Nagyfi, 2018). The circle simulating the cell body of the neuron represents a simple mathematical procedure that generates n output signals from the set of input signals represented by the multivariate vector.

A typical neural network comprises an input layer that receives external data, hidden layers for processing, and an output layer that produces results. In fully connected networks, each neuron connects to all neurons in adjacent layers, with information flowing forward through successive layers, classified as feedforward ANNs (Nagyfi, 2018). Figure 5.5 shows neurons as circles and connections as lines between layers. The input layer serves as a non-active distribution point, passing signals to the first active layer (Zupan, 1994). This example displays a four-layer network with three inputs, two hidden layers containing eight total nodes, and two outputs (Graff et al., 2014). This multi-layered structure allows ANNs to process complex information patterns and learn from examples, thereby making them powerful tools for various applications in pattern recognition, data analysis, and decision-making tasks.

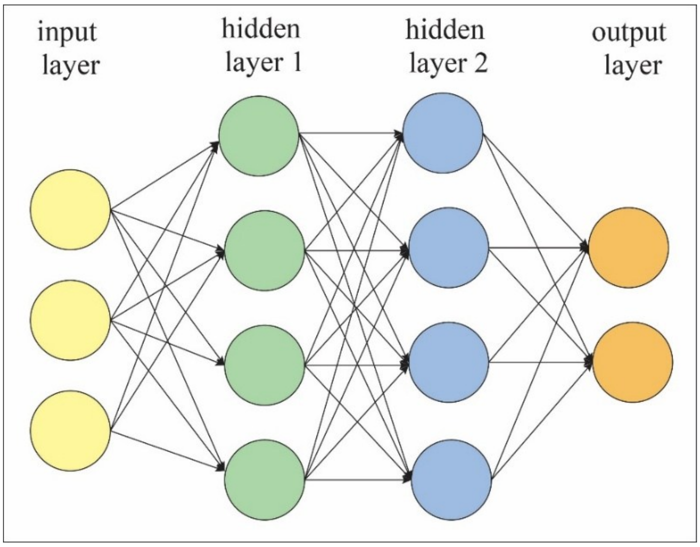

Figure 5.5: Example of a 4-layer neural network with 3 inputs, 8 hidden nodes, 2 hidden layers, and 2 outputs (modified from Graff et al., 2014).

5.3.2. Architecture of a CNN

The architecture of a CNN, a specialized type of ANN, consists of several specialized layers, each of which performs a different function in the processing pipeline.

The convolutional layer serves as the basic building block of CNNs. This layer contains a set of trainable filters (or kernels) that are optimized during the training process. Each filter slides over the dimensions of the input image, computing a cross-correlation between the input and the filter to generate an activation map.

Two essential hyperparameters control the behaviour of convolutional layers: padding and stride. Padding addresses boundary conditions by adding values around the input edges when the filter extends beyond the input boundaries, indicating, how many parameters are added to each edge of the image (Goodfellow, 2016). Stride determines the step size of the filter as it moves across the input image - smaller stride values capture finer details in the activation map.

Figure 5.6 shows the calculation of activation map elements. In this example, a 5x5 input grid contains a highlighted 3x3 region in the upper left corner, representing the current receptive field. A 3x3 filter containing the pattern [1, 0, -1] repeated over three rows processes this region. The convolution multiplies the corresponding values between the receptive field and the filter and sums the results to produce the activation map value. This process continues as the filter traverses the input, generating a complete 3x3 activation map, with each element representing a neuron's response to its local input region.

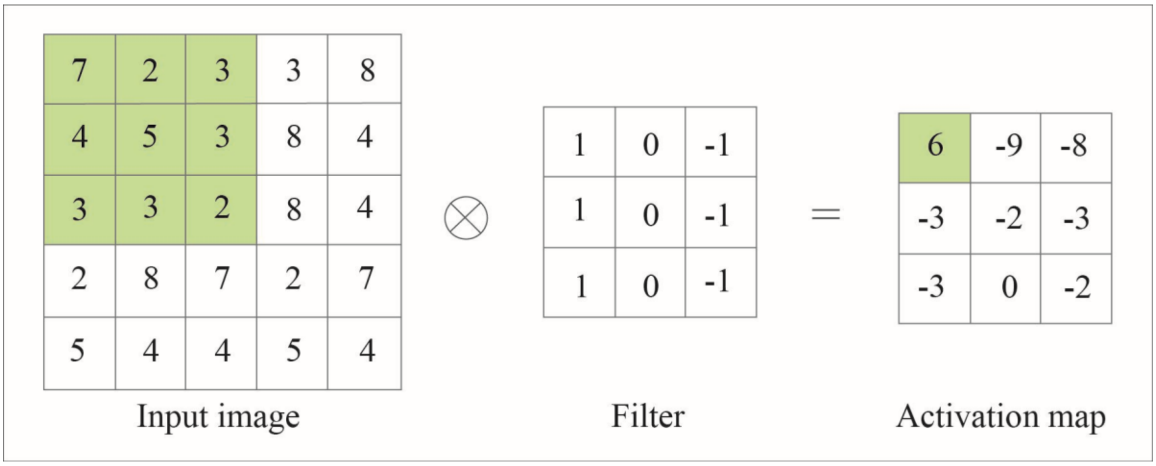

Figure 5.6: Graphical example of a convolutional process.

The Rectified Linear Unit (ReLU) layer introduces nonlinearity into the network by applying an activation function that replaces negative values with zero, as shown in Figure 5.7.

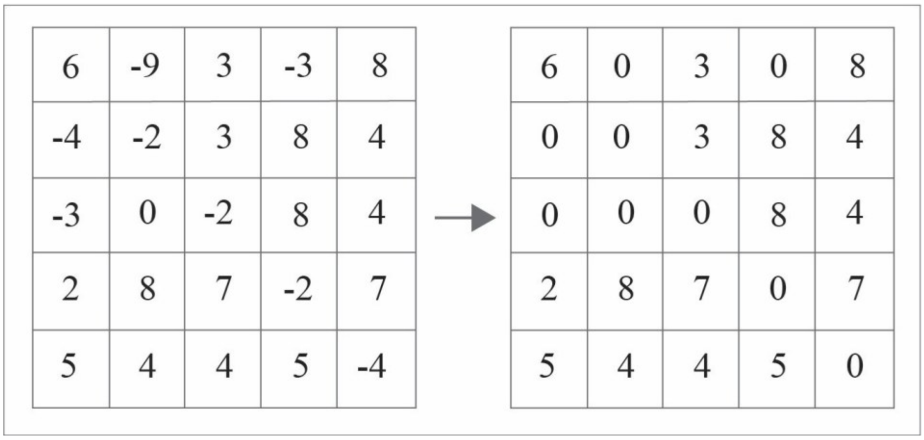

Figure 5.7: Graphical example of a ReLU process.

Goodfellow et al. (2016) describe the pooling layer as a critical component for achieving spatial invariance and dimension reduction. The pooling operation involves sliding a filter over the input and applying specific operations to each region. Like convolutional layers, pooling operations use stride and padding parameters to control filter movement and boundary handling. There are two primary types of pooling: Max Pooling and Average Pooling. Max Pooling (Figure 5.8) selects the maximum value from each filter region, creating a feature map of the most prominent features. Average Pooling (Figure 5.9) calculates the average value within each filter region, creating a feature map that represents the average feature activation.

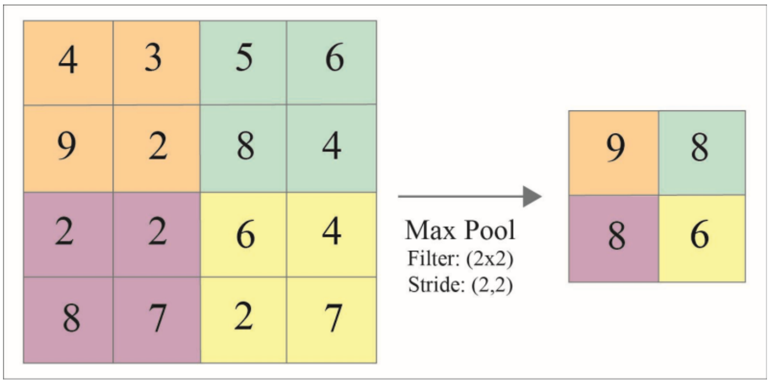

Figure 5.8: Graphical example of the max pool process.

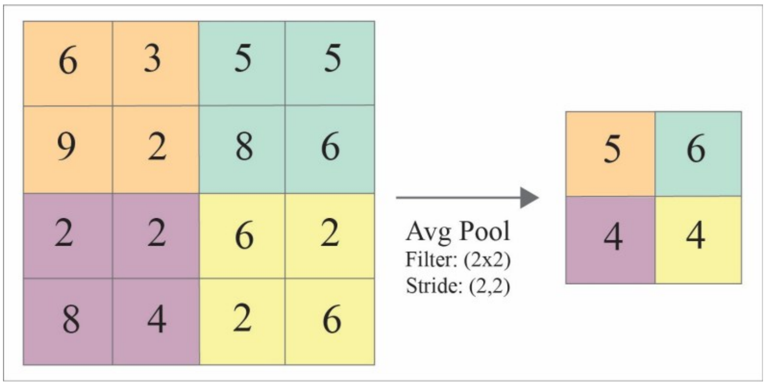

Figure 5.9: Graphical example of the Average Pool process.

The batch normalization layer, introduced by Garbin et al. (2020), addresses internal covariate shift by normalizing input distributions within training mini batches. In fact, during the training of the network, the continuous change of the distribution of the input to the layers after each mini batch, when the weights are uploaded, creates a recurring problem. This layer transforms inputs to maintain zero mean and unit variance by calculating and applying batch-specific statistical parameters. This normalization process significantly reduces the training time and increases the learning stability. By providing consistently scaled features to subsequent layers, batch normalization enables more efficient data processing regardless of the original feature distributions (Ioffe & Szegedy, 2015).

5.3.3. ResNet

The ResNet (Residual Network) architecture uses residual blocks to facilitate network training (Ložnjak et al., 2020). Each residual block implements a dual-pathway processing mechanism. Within a residual block, the input data takes two different paths. The main pathway processes the input through a series of convolutional layers with associated weights and activation functions. Meanwhile, a parallel skip connection preserves the original input through an identity mapping. This skip connection, also known as a shortcut connection, allows the input to bypass one or more layers and be added directly to the output of the main pathway (Arrighi et al., 2023).

In this study, the ResNet-152 variant consisting of 152 layers organized into residual blocks of varying configurations was employed. The network architecture concludes with a fully connected layer that performs the final classification task by mapping the learned feature representations to the desired output classes. Figure 5.10 illustrates the fundamental residual learning building block and shows how input x is processed through the weighted transformation pathway F(x) and the identity skip connection. The outputs are then combined through element-wise addition.

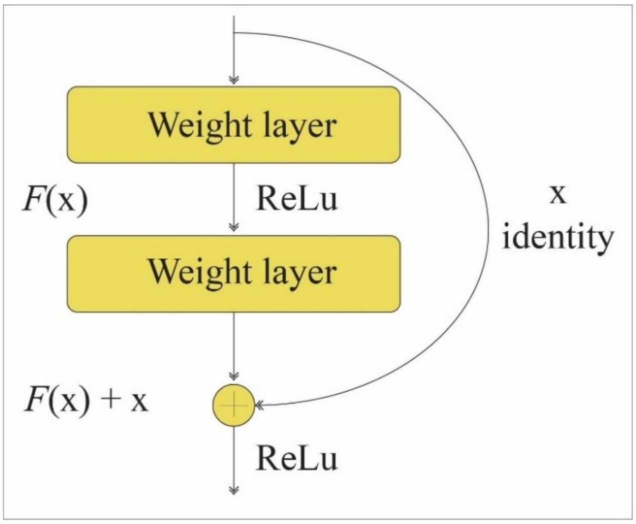

Figure 5.10: Structure of a residual block, modified from Zhang et al., 2022.

The core principle of residual learning is expressed by the mathematical formula:

$$F(x) = H(x) - x$$

where H(x) represents the desired underlying mapping and x is the input to the first subset of layers. This formulation assumes that the input and output dimensions are identical, which allows for the addition operation represented by the circled plus symbol in the figure.

While this reformulation may seem counterintuitive at first, He et al. (2016) demonstrated its remarkable efficiency. The approach effectively presupposes the problem by using the identity function as a reference point, rather than requiring the network to learn a zero mapping as in traditional CNNs. This makes it easier for the network to learn and recognize perturbations if they are viewed as deviations from the identity rather than as entirely new functions. The residual learning technique is applied periodically every few stacked layers, creating a deep architecture that maintains efficient training characteristics despite its depth.

5.3.4. ANN in geosciences

The methodology for implementing artificial intelligence in coastal geomorphological analysis focuses on CNNs for automated feature detection and classification. This approach builds on established machine learning techniques previously applied to geological mapping and coastal zone analysis (Hoonhout et al., 2015; Valentini et al., 2017). The methodology integrates structured support vector machines for coastal image classification, achieving significant accuracy in distinguishing between water, sand, and vegetation features (Hoonhout et al., 2015).

The framework incorporates interactive AI techniques adapted from geological mapping applications, where automated image segmentation is combined with expert guidance to efficiently delineate features (Vasuki et al., 2014, 2017). This approach is particularly valuable for processing large remote sensing datasets acquired by UAV and marine survey platforms. The integration of CNN architectures enables automated processing of time-lapse images acquired both above and below the waterline, following methods established by Van Veen (2017) and further refined for marine applications. Data preprocessing follows protocols developed for underwater image analysis, incorporating recent advances in the removal of water column effects using the Sea-thru method (Akkaynak and Treibitz, 2019).

The implementation utilizes DL architectures optimized for handling the variable environmental conditions found along the coastline. Model training incorporates techniques to improve generalization across diverse coastal environments, addressing problems identified in previous coastal monitoring applications (Buscombe & Ritchie, 2018; Valentini & Balouin, 2020). The methodology also integrates quantitative metrics for assessing model performance and reliability, which are essential for validating automated classifications against expert interpretations.

Additional methodological considerations address the integration of multiple data sets, including bathymetric measurements and environmental parameters (Bergsma et al., 2016). This comprehensive approach enables systematic analysis of coastal morphodynamics while maintaining the high standards of accuracy required for scientific research and coastal management applications (Chen et al., 2019).

5.4. Training process of the ANNs

In recent years, there has been a notable increase in interest surrounding the field of ANN, largely due to the remarkable capabilities that these networks possess. To fully exploit the potential of these systems, it is vital to gain an understanding of the mechanisms by which they learn. Although the networks were initially inspired by biological brains, their operational principles differ substantially from biological neural processes (Yusiong, 2012). An ANN functions as an adaptive system with distinctive properties including the ability to adapt, learn, and generalize. ANNs demonstrate high accuracy in classification and prediction tasks through their massively parallel processing architecture, fault tolerance, self-organization, and adaptive capabilities. These characteristics enable them to address complex problems by modifying their network structure during the learning process (Huanping et al., 2011).

Training an ANN involves trying to achieve the best set of parameters. The supervised learning is the most widely adopted learning paradigm and forms the foundation of the methodology employed in this thesis. In these frameworks, each input data sample is paired with a target label or ground truth value.

The algorithm is associated with an error function. The learning algorithm transforms the error signal into modifications to the connection weights that regulate the inputs to the various processing elements (Uhrig, 1995). The main goal of this learning process is to determine the optimal set of connection weights that align the ANN's output with the desired target values (Yusiong, 2012; Figure 5.11). The training is considered complete when the error function corresponds to the local minimum.

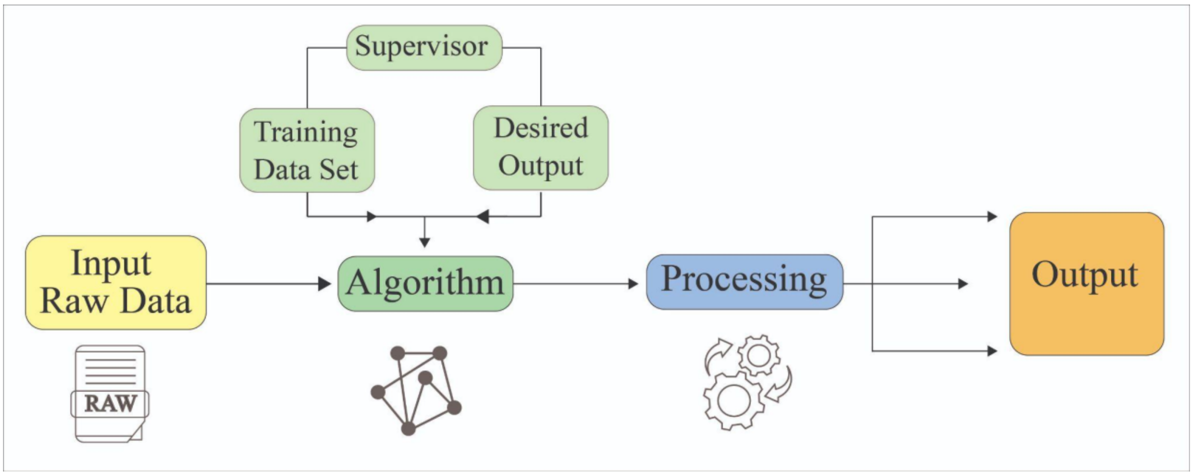

Figure 5.11: How the supervised learning system works (modified from Mehta, 2019).

This thesis deals with the topic of binary and multi-class classification, which involves the categorization of new elements (images) into two or five different classes, respectively, based on a training set of input images and associated output classes. As Van Veen (2017) describes, numerous network models exist, differentiated by their topology, data flow, neuron types, and intrinsic characteristics such as density and layer configuration. This study focuses on CNNs, characterized by full connectivity and at least one convolutional layer, with information flowing unidirectionally from the input layer through hidden layers to the output layer, without feedback loops or cyclic connections. CNNs have demonstrated efficacy in processing pixel data for image recognition and visual analysis applications (Van Veen, 2017). Their diverse architectures and layer configurations (Section 5.3.3) render them particularly well-suited for image-based tasks. These networks represent a notable advancement in machine learning approaches to visual data processing and analysis.

This Chapter focuses on the more technical aspects of the CNNs, explaining the approaches chosen and used in this work.

5.4.1. Activation functions

The activation function plays a crucial role in defining how a CNN layer operates, specifically determining the output of the node given a set of inputs (Hinkelmann, 2018). Figure 5.12 illustrates the most implemented activation functions in CNN. The linear activation function (Figure 5.12a) represents the simplest type, defining a linear neuron whose output is calculated as (1).

The step activation function (Figure 5.12b) operates as a binary function defining the binary threshold neuron, where the output is calculated based on a threshold corresponding to the bias of the neuron (Avinash, 2017):

$$g(x) = \begin{cases} 1 \; if \; \sum_i w_i x_i \geq -b \\ \quad 0 \text{ otherwise} \end{cases}$$

The sigmoid or logistic function (Figure 5.12c) defines the sigmoid neuron, which produces a smooth, bounded real value output (Avinash, 2017):

$$g(x) = \frac{1}{1 + e^{-x}}$$

Where $x = b + \sum_i w_i x_i$.

The rectified linear function (ReLU) (Figure 5.12d) operates linearly for positive values and outputs zero for negative values (Avinash, 2017):

$$g(x) = \begin{cases} x \; if \; x > 0 \\ 0 \text{ otherwise} \end{cases}$$

ReLU functions are computationally efficient and can generate true zero outputs for negative inputs, resulting in sparse activation of hidden units. In addition, they behave like linear activation functions, optimizing and reducing vanishing gradient problems since gradients remain proportional to node activations. This property is particularly important for hardware implementation and training of deep multilayer networks with nonlinear activation functions using backpropagation (Glorot et al., 2011). These activation functions form the basis of the network operations, each serving specific purposes in different network architectures and applications. Their choice significantly influences the network's learning capacity and overall performance in different computational tasks.

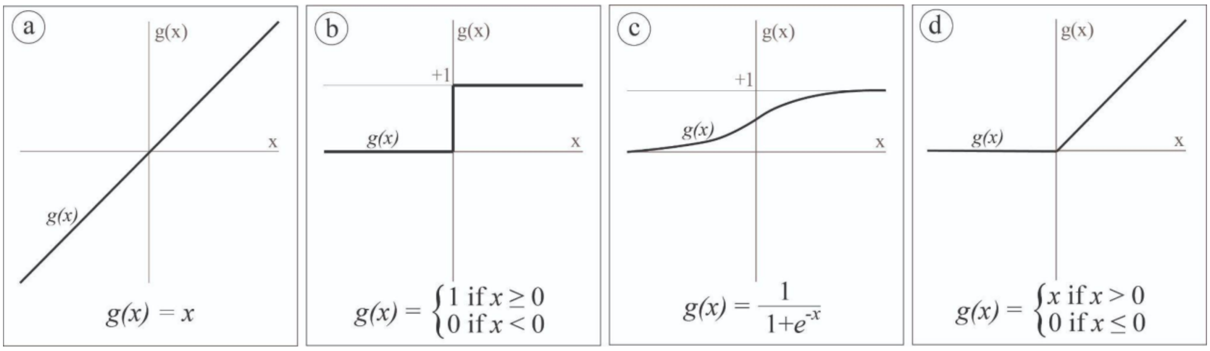

Figure 5.12: Activation functions. a) Linear activation function, which would either output as the function; b) Step activation function, which would output either 0 or 1 if the sum of the weights was larger than the threshold; c) Sigmoid or logistic activation function, which always outputs values between 0 and 1; d) Rectified linear function, which always outputs values of 0.

5.4.2. Loss Function

The CNN learning process requires optimizing its weights and biases using a loss function, which quantifies the error between predicted and desired outputs. For binary classification tasks, the binary cross-entropy loss function is particularly effective, providing faster convergence and better generalization compared to alternatives such as mean squared error (Goodfellow et al., 2016). Binary cross-entropy is based on the information theory concept of entropy, which measures uncertainty in probability distributions (Cover, 1999):

$$H(q) = -\sum_{c=1}^{C} q(y_c) \cdot \log\big(q(y_c)\big)$$

(2)

Where q is the probably distribution, y_c is the class label, and C is the number of classes. The cross-entropy between the true distribution q and the predicted distribution p is defined as:

$$H_p(q) = -\sum_{c=1}^{C} q(y_c) \cdot \log\big(p(y_c)\big)$$

This measures the dissimilarity between q and p, with lower values indicating better matching (Bishop & Nasrabadi, 2006). During training, the network minimizes the cross-entropy over N training examples. Thus, given $q(y_i) = \frac{1}{N}$:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} \log\big(p(y_i)\big).$$

For binary classification with classes {0,1}, when $\hat{y}_i$ represents the output probability of the model:

For class 1: $y_i = 1 \rightarrow -\log(\hat{y}i)$
For class 0: $y_i = 0 \rightarrow -\log(1 - \hat{y}i)$

Let y^0 and y^1 be the mean predictions for each class.

For multiclass classification tasks where C>2 (C as a real number), the cross-entropy loss function provides a theoretical framework that naturally extends from binary classification. The logarithmic loss function, which is mathematically equivalent to cross-entropy, serves as the primary optimization objective for multiclass classification problems (Goodfellow et al., 2016). A critical component of multiclass classification is the Softmax function, which transforms the raw network outputs into an appropriate probability distribution. The Softmax function takes a vector z of C real numbers and normalizes them into a probability distribution of C probabilities. While the input logits can be negative or positive and do not sum to 1, the Softmax transformation ensures that each output component is in the interval [0,1] and all components sum to 1, making them valid probabilities. Importantly, larger input values correspond to larger output probabilities. This transformation is defined as:

$$\sigma(z_i) = \frac{e^{z_i}}{\sum_{j=i}^{C} e^{z_j}}$$

where z_i is the raw network power (logit) for class C and the resulting probability for class.

The multiclass cross-entropy loss for a dataset of N samples across C classes is then defined as:

$$L = \frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} t_i log y_{i,c}$$

where $y_{i,c}$ is 1 for the correct class and 0 otherwise, and L is the model's predicted probability for class c of sample i after Softmax activation.

This formulation provides numerical stability and computational efficiency due to the simplicity of its derivatives, making it particularly suitable for first-order gradient descent optimization methods (Goodfellow et al., 2016). In addition, the function does not encode any prior assumptions about class relationships, so it serves as a versatile, general purpose objective function that can be applied to various classification scenarios. Modern DL frameworks implement this loss by combining Softmax activation and cross-entropy computation into a single, numerically stable operation (*CrossEntropyLoss - PyTorch Documentation,* 2024). The resulting gradients provide efficient learning signals that help the model develop well-calibrated probability estimates across classes.

5.4.3. Gradient descent algorithm

Gradient descent is a fundamental optimization method in DL that aims to minimize the loss function through iterative parameter updates. The algorithm computes the gradient of the loss function, which represents the direction of steepest ascent at any given point. For a differentiable function f: Rn → R at point x = ($x_1,..,x_n$), the gradient represents the direction of steepest slope and measures how the output of the function varies with input parameters. If the loss function is differentiable and has a minimum point, the gradient decreases as we approach this point, since the slope of the function becomes smaller. At each iteration, the parameters are updated in the opposite direction of the gradient, with the step size determined by the learning rate η (Durán, 2019):

$$\theta = \theta - \eta \cdot \nabla_\theta J(\theta)$$

(3)

Where J is the loss function, θ represents the network parameters (weights and biases). The choice of learning rate is critical. Too small a value will result in slow convergence and computation, while too large a value may cause the model to overshoot the minimum (Figure 5.13). This figure illustrates the relationship between learning rate and neural network training performance through four scenarios (Franco, 2023). Setting the learning rate too high, as shown by the red curve, makes the training process highly unstable. The loss function exhibits erratic oscillations that may lead to divergence rather than convergence. This instability prevents the model from learning effectively and can render the training process ineffective. A moderately high learning rate, represented by the pink curve, produces somewhat more controlled training behavior. Although this approach may eventually converge, the irregular progress and potential for instability make it suboptimal for reliable model training. Conversely, low learning rate, depicted by the blue curve, results in extremely slow convergence, despite providing stable, smooth progress. While this approach guarantees stability and eventual convergence, the computational cost and time requirements render it impractical for real-world applications. The optimal learning rate configuration, illustrated by the green curve, achieves an ideal balance between convergence speed and training stability. The corresponding optimization path shows appropriately sized steps that consistently move toward the global minimum, enabling effective and efficient model training within a reasonable timeframe.

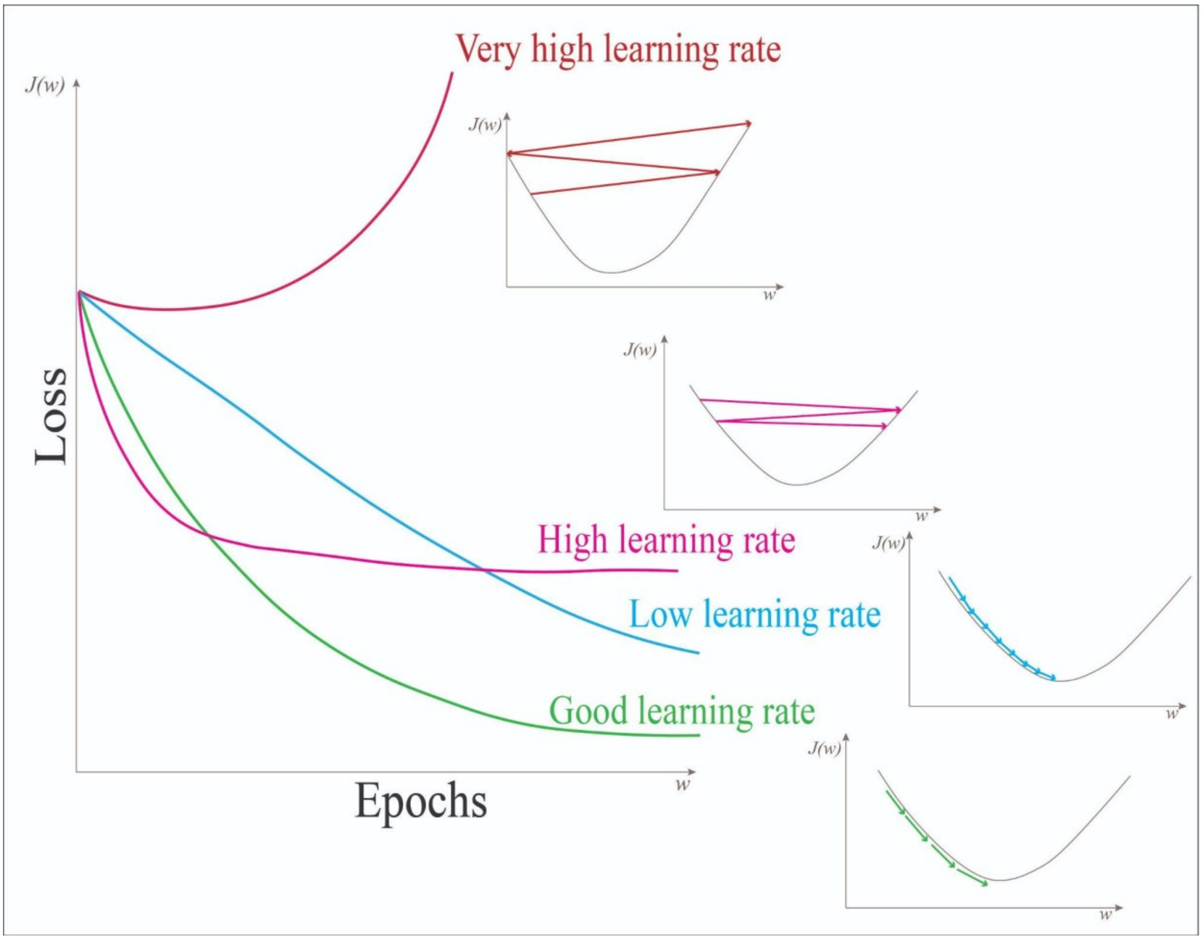

Figure 5.13: Graphical example of how the algorithm depends on the chosen learning rate (modified from Franco, 2023).

Ruder (2016) outlines three main approaches to gradient descent optimization. Batch gradient descent (BGD) computes gradients over the entire dataset. Stochastic gradient descent (SGD) updates parameters for each training example. Mini-batch gradient descent computes updates for small batches of *n* examples.

As specified by Bengio et al. (2017), modern implementations typically use mini-batch SGD to balance computational efficiency and convergence stability (Figure 5.14). Batch gradient descent (shown in purple) uses the entire training dataset to compute gradients at each iteration. This results in a smooth and direct path toward the minimum. Mini-batch gradient descent, depicted in orange, strikes a balance between computational efficiency and convergence stability by using subsets of the training data to estimate gradients. This approach produces a moderately smooth path with minor fluctuations, balancing the stability of batch gradient descent with improved computational efficiency. Stochastic gradient descent, in green, computes gradients using individual training samples, resulting in a highly erratic and noisy optimization path. Although this method provides the fastest computational speed per iteration and can potentially escape local minima, its irregular trajectory demonstrates significant variance in gradient estimates.

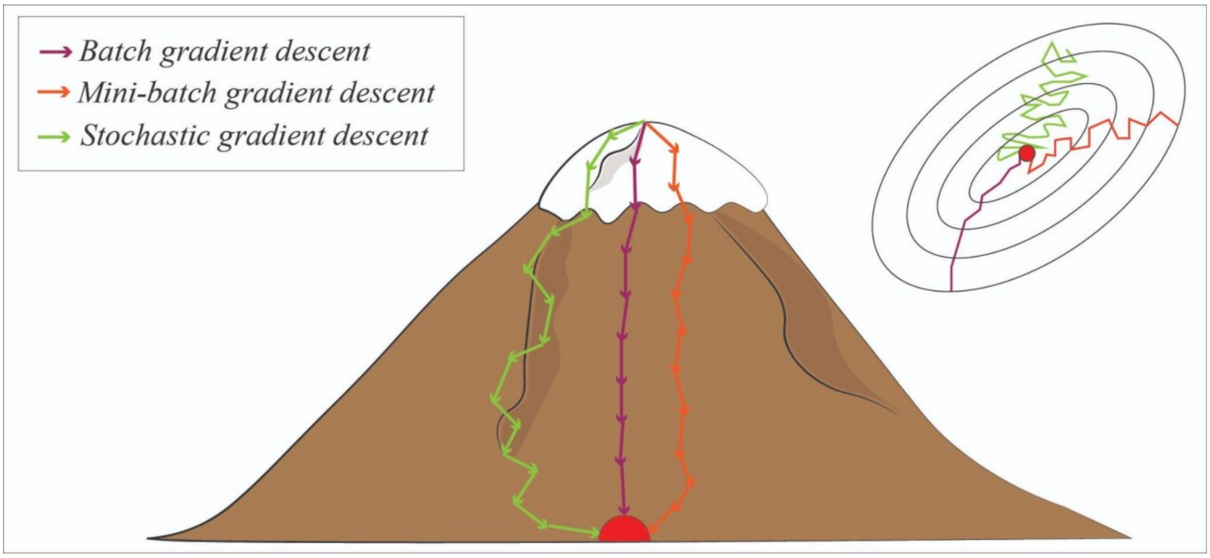

Figure 5.14: Gradient descent variants' trajectory towards local minimum (modified from Dabbura, 2017).

Qian (1999) improved the algorithm by introducing momentum helps dampen oscillations and navigate areas of high curvature more effectively (Figure 5.15).

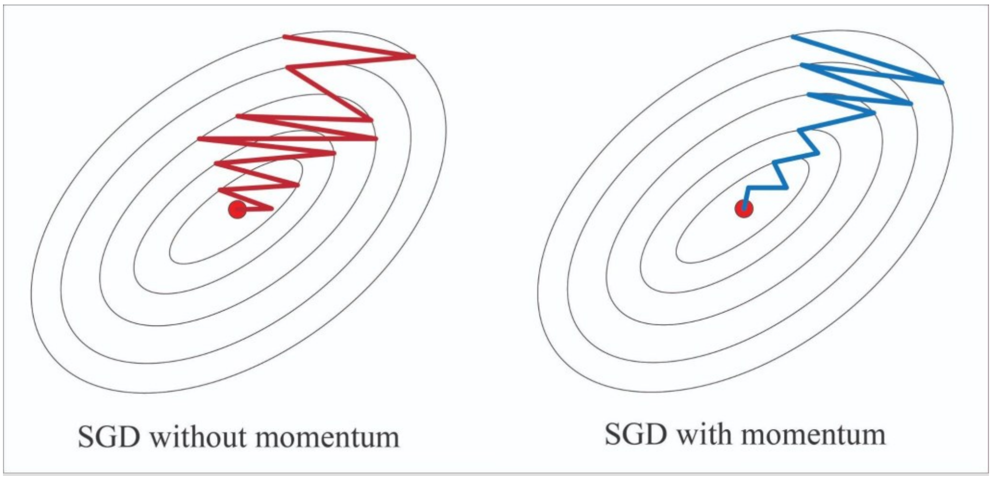

Figure 5.15: Stochastic Gradient Descent without momentum (red) and Stochastic Gradient Descent with momentum (blue) (modified from Hirol, 2023).

Further developments include the Adaptive Gradient Algorithm (Adagrad) (Duchi et al., 2011), which automatically adjusts learning rates for each parameter, and RMSProp (Tieleman, 2012), which addresses Adagrad's aggressively decreasing learning rate by exponentially decaying averages of squared gradients. Kingma & Ba (2017) introduced the Adam optimizer, which combines RMSProp with momentum using first and second moment estimates of the gradient (Figure 5.16).

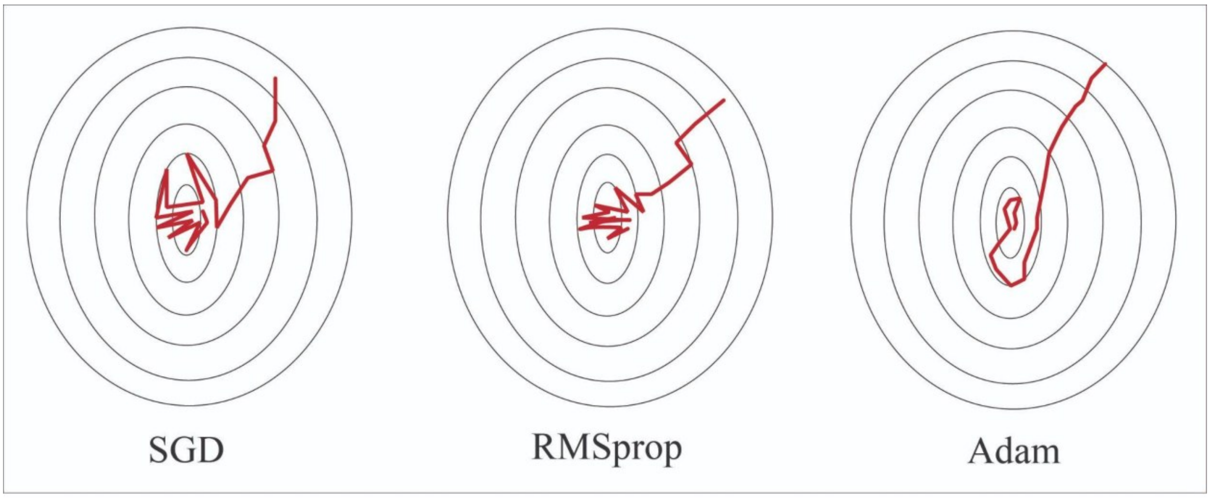

Figure 5.16: Comparison between SGD, RMSProp and Adam optimization algorithms (https://github.com/Atcold/torch-Video-Tutorials/?tab=readme-ov-file/).

5.4.4. Backpropagation

The backpropagation algorithm, introduced by Rumelhart et al. (1986), is fundamental to the learning phase of the networks and provides an efficient method to computer gradients from the loss function. The learning process in feedforward networks occurs through a two-phase mechanism. During the forward pass, data propagates through the network from input to output, with each layer of neurons processing information using their current weights and activation functions to generate predictions. The network then quantifies the prediction error using a loss function. In the backward pass, through backpropagation, the network analyzes how each connection contributed to this error, enabling appropriate adjustment of internal parameters for improved future predictions.

The architecture of a CNN consists of a linear chain of stacked layers (Figure 5.17), where o represents the output and x represents the input. With each layer containing optional parameters ($\theta_1, \ldots, \theta_4$), the network can be defined as a composition of functions:

$$f = f_4 \circ f_3 \circ f_2 \circ f_1$$

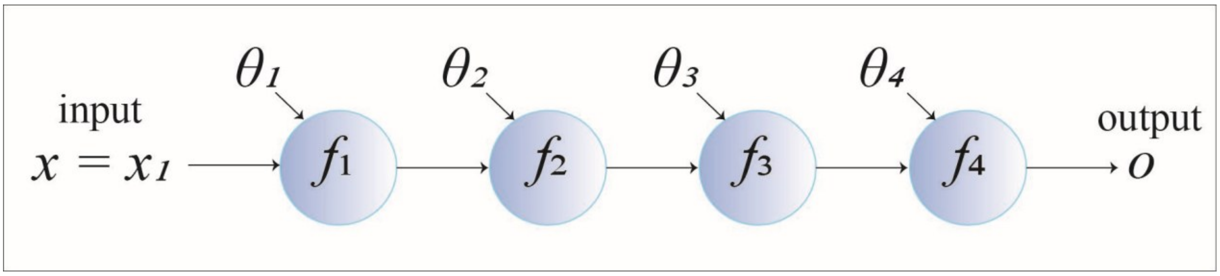

Figure 5.17: A simple four-layer linear chain feedforward model, where o is the output and x is the input (modified from Murphy, 2022).

At the heart of backpropagation is the chain rule, which allows evaluation of the gradient of the loss function with respect to the parameters θ. Using Leibniz notation, the chain rule states that for a function g differentiable at point x, and a function f differentiable at point $u = g(x)$:

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

As detailed by Murphy (2022), backpropagation functions as a dynamic programming algorithm, with specific implementations for different layer types and generalization across DL models. Essentially, it is a process of re-evaluating the weights of the neurons in the network that rebuilds the model as a function of the error evaluation.

5.4.5. Train, validation and test

The CNN learning process involves three distinct phases, each of which uses different subsets of the available data to ensure robust model development and evaluation. During the training phase, the network learns from the input data by adjusting its internal parameters through the optimization process, described by Rumelhart (1986).

The data set is strategically divided into three subsets to optimize the model and evaluate its performance. The training set is used to adjust the model parameters through iterative learning.

The validation set serves as an independent check during training, allowing evaluation of the model's performance on unseen data and helping to prevent overfitting. The test set, which is completely hidden during training and validation, provides the final assessment of the model's generalization capabilities.

During model evaluation, the process is reversed - weights remain fixed as inputs vary, and the network produces outputs that can be evaluated for accuracy. This approach, with clear separation between training, validation, and testing phases, ensures reliable evaluation of the true performance and generalization ability of the network.

5.4.6. Dropout

Dropout is an innovative regularization technique that addresses one of the main problems in DL: overfitting. As described by Srivastava et al. (2015), overfitting occurs when a network becomes overly specialized to its training data, compromising its ability to generalize and accurately classify new and unseen data in the test set. The dropout mechanism operates during the training phase through a systematic but randomized approach. For each layer in the network, a predetermined proportion of neurons are temporarily deactivated or "dropped out" of participation in the forward and backward passes. This selective deactivation process, illustrated in Figure 5.18, is based on a probabilistic framework in which each neuron has a specified probability of being temporarily removed from the network, controlled by a hyperparameter.

The resulting training process is inherently noisy, as the network must learn to process information through different pathways and neuronal combinations. This architectural flexibility proves advantageous, promoting the development of more robust and adaptive models that can better generalize to new data patterns.

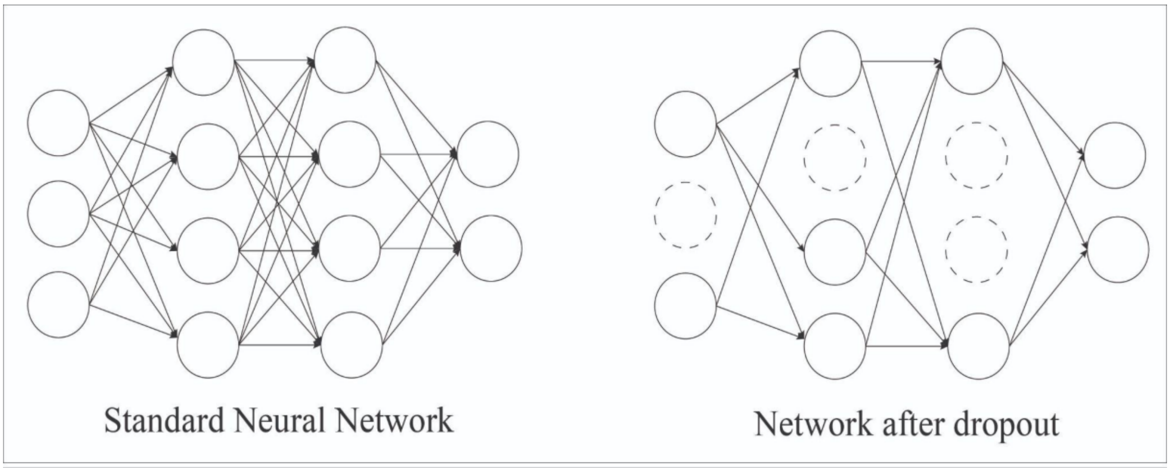

Figure 5.18: Graphical example of the Dropout network mode, where the image on the left shows a standard CNN with two hidden layers, and the image on the right shows an example of a thinned network created by applying dropout to the network on the left.

5.4.7. Transfer Learning

Transfer learning is one of the most important downstream applications of trained image classification models (He et al., 2019). The fundamental goal of transfer learning is to improve understanding of current tasks by connecting them to related tasks performed in different contexts (Hosna et al., 2022). By building bridges between source and target domains, transfer learning finds potentially superior solutions that would otherwise be difficult to discover (Zhuang et al., 2021).

The underlying assumption is that many factors that explain variation in one problem are relevant to variation in another problem, especially in supervised learning contexts where inputs remain consistent while targets may differ (Goodfellow et al., 2016).

Transfer learning creates an efficient communication channel between source and target tasks, making the learning process more dynamic and adaptive (Pan & Yang, 2013). Its strategic value extends beyond individual task improvement to knowledge sharing across domains (Taylor & Stone, 2009). This cross-domain capability is particularly valuable when target training data is scarce, as it reduces the reliance on large amounts of domain-specific data to construct effective learning models (Zhuang et al., 2021).

However, successful implementation requires careful consideration of domain compatibility. When transfer occurs between incompatible domains, performance on the target task may deteriorate, a phenomenon known as negative transfer. The relationship between source and target tasks must be sufficiently aligned for knowledge transfer to yield positive results.

In practical applications, pre-trained models such as those based on ImageNet have demonstrated the effectiveness of transfer learning. ImageNet provides an extensive image database organized according to the WordNet hierarchy, with each node represented by hundreds or thousands of images. Models trained on this large dataset can transfer learned visual representations to specialized domains with minimal fine-tuning, illustrating why transfer learning has become such a promising area of machine learning research and application (Zhuang et al., 2021).

5.5. EXplainable Artificial Intelligence (XAI)

As DL models grow in complexity, the need for interpretability becomes increasingly important. EXplainable Artificial Intelligence (XAI), as discussed by Lipton (2018), is emerging as an important framework for understanding, interpreting, and refining DL models. This study implements a proof-of-concept using XAI tools, specifically techniques based on Class Activation Mapping (CAM) (Zhou et al., 2016) and Gradient-weighted CAM (GradCAM) (Selvaraju et al., 2017a, b).

In the field of XAI, numerous techniques have been proposed to improve the transparency, explicability, and interpretability of ANN (Angelov et al., 2021; Arrieta et al., 2020; Schwalbe & Finzel, 2024). In the context of CNN, particular attention has been paid to methods that focus on the visualization of internal representations and network dynamics. Bach et al. (2015) emphasized the importance of understanding classification decisions in automated image classification systems and proposed a layer-wise relevance propagation technique that generates heat maps to visualize input features relevant to the decision process (Arrighi et al., 2023).

Since the introduction of CAM, researchers have proposed various techniques to generalize the approach across different Deep Neural Networks (DNNs) or to improve its accuracy (Chattopadhay et al., 2018; Draelos & Carin, 2020; Fu et al., 2020; Jiang et al., 2021; Muhammad & Yeasin, 2020; Ramaswamy et al., 2020; Selvaraju et al., 2017; Srinivas & Fleuret, 2019). CAM-based techniques have shown utility in object localization within CNNs trained for image classification, even in the absence of explicit position information (Zhou et al., 2016). This concept has been further developed in studies of weakly supervised object detection (Cheng et al., 2020; Inbaraj & Jeng, 2021).

However, the quantitative evaluation of XAI tools presents significant problems. Many studies applying these techniques to CNNs rely primarily on qualitative considerations and visual evaluation of subset datasets (Joo & Kim, 2019; Lopes et al., 2022). Nauta et al. (2023) addresses this limitation by proposing a unified framework for evaluating these tools. Researchers have identified several concerns with CAM-based techniques, including low fidelity (Draelos & Carin, 2020; Li et al., 2021), difficulties in handling multiple object instances (Singh & Sharma, 2022), and challenges in precise instance localization (Xiao et al., 2021).

According to Dwivedi et al. (2023), CAM and related techniques are classified as local XAI tools, meaning that they are applied after model training without parameter modification and produce explanations for individual data points rather than global model interpretations.

5.5.1. Gradient-weighted Class Activation Map (GradCAM)

GradCAM, introduced by Selvaraju et al. (2016), generates an activation map G_c through a linear combination of feature map activations from the last convolutional layer of the network (A_k), weighted by the loss function gradients relative to the neurons of this layer (w_k^C):

$$G_C = \text{ReLU}\,(\sum_k w_k^C \cdot A_k)$$

The technique builds on the work of Zhou et al. (2016) by generating heatmaps that highlight image regions that are most important for network task performance. GradCAM's implementation follows a systematic process, starting with modifying the model architecture by cutting it at the target convolutional layer and adding a fully connected layer for final classification. This modification allows the system to retain spatial information while allowing class-specific gradient computation.

The visualization process consists of many steps. First, the input image is processed by the modified architecture. Then, the system computes gradients of the chosen class score with respect to the feature maps of the target convolutional layer. These gradients are normalized, retaining positive values and suppressing negative values, to construct the final heat map. The resulting visualization is scaled to the original image dimensions and provides a normalized score between 0 and 1, indicating the relative importance of each region in the network's decision process (Arrighi et al., 2023).

The advantage of GradCAM is its ability to preserve spatial information through the convolutional layers, allowing precise localization of features that influence classification decisions.

5.5.2. GradCAM++

GradCAM++ (Chattopadhay et al., 2018) represents a significant evolution in visualization techniques, solving some limitations of the original GradCAM approach. This advanced iteration uses a different framework to compute weights w_k, incorporating weighted combinations of positive partial derivatives, including higher-order derivatives (https://github.com/jacobgil/pytorch-grad-cam). This enhanced approach allows for more nuanced gradient weighting, resulting in superior localization capabilities and improved handling of complex visual scenarios.

The main innovation of GradCAM++ is its pixel-wise weighting mechanism, which offers several distinct advantages over its predecessor. First, it offers improved ability to handle multiple instances of the same object within a single image, a scenario where the original GradCAM often had problems. Second, it provides more precise object localization, particularly for smaller or less prominent features that might be missed by traditional GradCAM. This approach allows for a more refined attribution of importance to different regions of the input image, resulting in a more accurate and detailed visualization of the attention patterns of the network (Arrighi et al., 2023).

However, this improved performance comes at the cost of increased computational complexity due to the computation of higher order derivatives. This trade-off is particularly relevant in real-world applications where processing time and computational resources are limited. However, the improved accuracy and detail provided by GradCAM++ often justifies the additional computational overhead, especially in applications such as medical image analysis, architectural studies, or any scenario where precise feature localization is critical for decision making.

5.6. Three-dimensional model construction in geomorphology

The development of 3D modeling applications in coastal geomorphology has been remarkable over the last decades. Traditional aerial and terrestrial photogrammetric techniques have been successfully adapted to coastal environments using stereoscopic principles (Fraser & Cronk, 2009; Matthews, 2008; Remondino & El-Hakim, 2006). Underwater photogrammetry has made significant progress, particularly in archaeological surveys, marine biological studies, and industrial inspections (Capra, 1993; Teague & Scott, 2017). Recent improvements in underwater photographic equipment and automation of photogrammetric software have increased interest in developing simple, cost-effective solutions for reconstructing accurate 3D object geometries in aquatic environments (Furlani, 2020; Guo et al., 2016).

However, creating complete 3D models of the intertidal zone presents technical problems, particularly in integrating data from above and bw, due to the different physical properties of air and water. A key technical obstacle is dealing with the different refractive indices of air and water, which affect the optical properties of camera lenses, especially when using popular tools such as GoPro cameras for bathymetric measurements (Schmidt & Rzhanov, 2012). Recent studies have shown that some of these technical difficulties can be overcome by specific data acquisition and processing techniques (Furlani et al., 2020), including the use of action cameras and 360-degree cameras both above and below the waterline to analyze the intertidal zone by creating 3D models of artificial and natural vertical walls. But, as Furlani et al. (2020) explained, when combining two models, it is common to use benchmarks, which are fixed points with known latitude and longitude values. These markers are present in both models to be combined. In the absence of reference points during the image acquisition phase, other tests were carried out in which an attempt was made to merge the models using markers fixed on the models themselves. These tests looked for points in common between the upper and lower models, but this method was not successful. For these reasons, the purpose of this thesis is to propose a new method of combining 3D models.

Appendix 1 summarizes all the steps taken to create the combined 3D model of the intertidal zone.

5.6.1. Structure from Motion and Multi-View Stereo (SfM-MVS) photogrammetry workflow

Structure-from-Motion (SfM) photogrammetry has revolutionized three-dimensional modeling in the geosciences, providing an efficient approach to characterize topographic surfaces and analyze spatiotemporal changes in natural environments (Mosbrucker et al., 2017). This technique has become particularly valuable for marine geomorphological studies, providing detailed digital representations of reality through a systematic process of data acquisition and processing (Remondino & El-Hakim, 2006).

The main advantage of SfM over traditional photogrammetric methods is its ability to automatically determine camera geometry and orientation parameters while processing overlapping images from multiple viewpoints (Westoby et al., 2012). Unlike conventional approaches that require explicit definition of interior and exterior orientation parameters, SfM employs advanced algorithms for automatic feature detection and correspondence matching across photographs (Lowe, 1999). This automation significantly reduces the technical expertise required while maintaining high data quality.

Traditional topographic research has relied largely on Digital Elevation Models (DEMs) produced by conventional photogrammetry (Barker et al., 1997; Chandler, 1999; Lane et al., 1994) and differential global positioning systems (dGPS) (Brasington et al., 2000; Fix & Burt, 1995). While airborne and terrestrial laser scanners later emerged as powerful tools for high-resolution data collection (Alho et al., 2009; Heritage & Hetherington, 2007), these methods often require expensive equipment and specialized expertise. SfM technology has democratized high-quality 3D data collection by reducing both cost and technical barriers (James & Robson, 2012).

Three-dimensional model construction for each data chunk follows standardized procedures within Metashape (version 2.0.2 build 16404 64 bit) (e.g., Carrivick et al., 2016; Tavani et al., 2014; Vaccher et al., 2024), with specific user-adjustable parameters optimized for coastal environments, chosen for its robust processing capabilities and user-friendly interface. The software integrates various algorithms to reconstruct three-dimensional structures by simultaneously determining extrinsic camera calibration parameters (position and orientation), intrinsic camera calibration parameters (focal length, distortion coefficients, and sensor size), and tie point locations to generate a sparse 3D point cloud (James & Robson, 2012; Ullman, 1979). The underwater calibration requires a specific procedure that is highlighted in section 5.6.2.

In marine environments, successful SfM applications require careful consideration of image acquisition parameters. Image quality has a significant impact on reconstruction accuracy (Grüen, 2012), requiring consistent lighting conditions and minimal water turbidity. The underwater applications benefit from a minimum of sixty images for optimal accuracy (Abadie et al., 2018), especially when considering refraction effects that alter focal length by 30-33%. Each point of interest must appear in at least three images acquired from different spatial locations, with convergence angles influencing mesh geometry (Eltner et al., 2016).

The workflow begins with elaborate algorithms (Ullman, 1979) that analyze 2D pixel coordinates across multiple images (Grüen et al., 2004; Szeliski, 2011) (Figure 5.19). During this critical first phase, algorithms like the Scale Invariant Feature Transform (SIFT) identify and match common points across images (Lowe, 1999), creating a network of tie points fundamental to 3D reconstruction. Using Sparse Bundle Adjustment (Snavely et al., 2008), the software determines camera positions and generates an initial sparse point cloud, establishing the basic geometric framework. The photo alignment process, which generates the initial sparse point cloud representation, used carefully selected parameters to optimize reconstruction quality: high accuracy setting, enabled generic pre-selection, key point limit of 40,000, and tie point limit of 4,000.

This process represents a significant advancement over traditional photogrammetric methods by automatically determining camera geometry and orientation parameters from overlapping images taken from multiple viewpoints (Westoby et al., 2012).

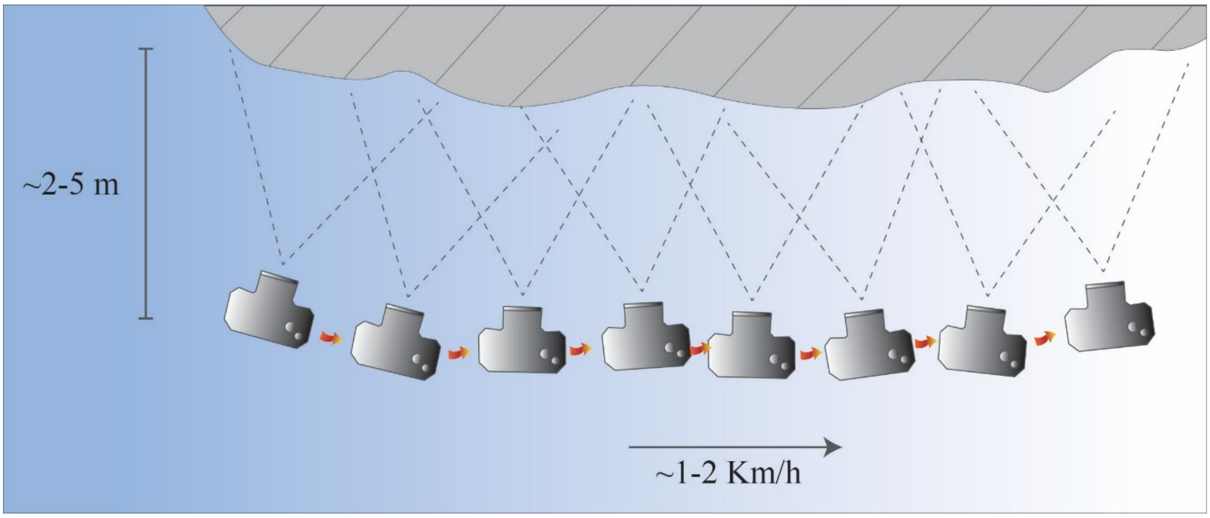

Figure 5.19: Basics of photogrammetry and the SfM technique, where the central object is photographed from different angles and the corresponding points are identified and correlated to create a three-dimensional image (modified from Fagerström, 2018).

In the second phase, Multi-View Stereo (MVS) algorithms (Seitz et al., 2018; Vaccher et al., 2024) are used to generate a dense point cloud, thus significantly increasing spatial resolution (Figure 5.20). This computationally intensive process projects individual points from multiple images to determine precise spatial coordinates, incorporating depth filtering to manage outliers and ensure model clarity. The resulting dense cloud is transformed into a mesh by triangular or quadrangular meshes, enabling volume measurements and detailed analysis.

The final texturing process applies photographic detail to create a visually accurate representation of the subject.

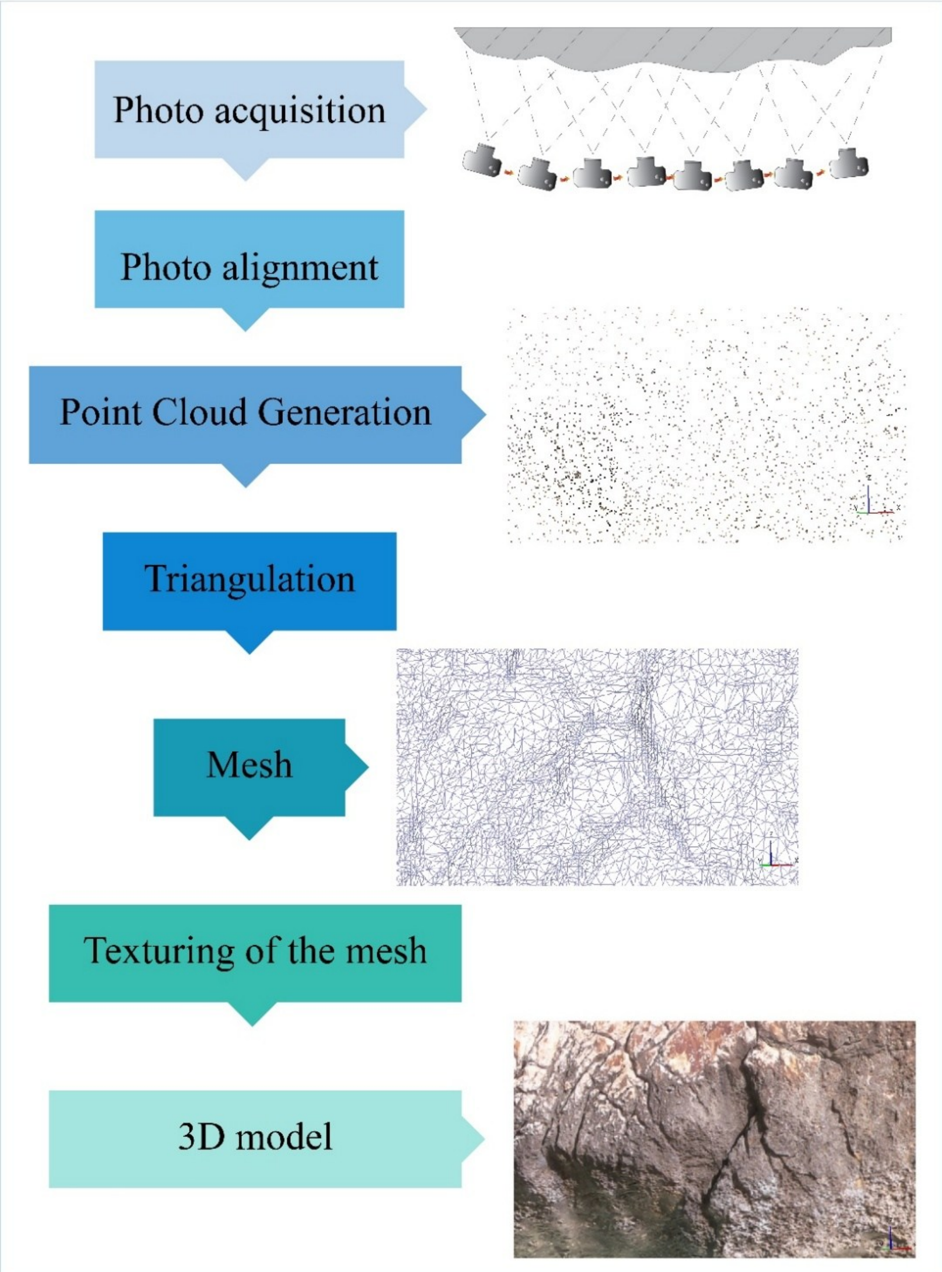

Figure 5.20: SfM-MVS workflow.

5.6.2. Underwater camera calibration

Accuracy assessment of underwater 3D modeling takes some specific problems due to the complexity of measuring Ground Control Points (GCPs) and checkpoints in underwater environments. Both physical properties of water and camera characteristics affect image acquisition by altering optical ray paths (Nocerino et al., 2016). As a result, the intrinsic parameters of underwater cameras differ significantly from their aw counterparts (Guo et al., 2016), primarily due to refraction at the air/water interface when using air-filled camera housings (Agrafiotis & Georgopoulos, 2015; Maas, 2015). Therefore, accurate lens distortion calibration through underwater camera system calibration is essential to optimize image alignment processes.

The successful implementation of underwater photogrammetric projects requires careful consideration of various environmental factors. The physical properties of water, which is approximately 800 times denser than air, vary with temperature, salinity, and pressure (Nocerino et al., 2016). The complex interplay of different refractive indices and lighting conditions, compounded by the reflection and absorption of sunlight, creates a difficult detection environment. Additional complications arise from water turbidity, wave action, currents, and light scattering, requiring specialized equipment such as high-tech waterproof domes for GoPro cameras.

The lens calibration process, a critical component of the SfM photogrammetry workflow, requires precise definition of distortion parameters through either manual or self-calibration approaches (Carrivick et al., 2016; Shortis, 2015). While aerial photography typically relies on Metashape's self-calibration capabilities for intrinsic and extrinsic corrections (Remondino & Fraser, 2006), underwater applications benefit significantly from manual calibration procedures performed under actual underwater conditions.

The manual calibration protocol uses a rectangular chessboard-like calibration frame included with Metashape that is permanently fixed to a 1 cm thick plastic table using waterproof adhesive (Figure 5.21). The calibration frame measures 40.9 cm by 27.25 cm, with individual squares measuring 1.65 cm per side (Figure 5.21a). The calibration process requires that this checkerboard be photographed underwater at distances approximating those used in coastal surveys. While previous research suggests a minimum of 20 images for adequate lens correction (Guo et al., 2016), this study used 34 photographs in Agisoft Metashape, activating all parameters in the "Calibrate Lens" section to ensure optimal accuracy, especially given the challenges posed by fisheye lenses.

The fisheye lens of the GoPro camera introduces significant radial distortion, which is particularly pronounced in underwater conditions, although accuracies of about 0.5 mm are still achievable (Guo et al., 2016). The calibration process requires setting the acquisition mode to "fisheye" and carefully monitoring the calibration results through Metashape's camera calibration menu. Successful calibration is indicated by correctly detected corners, which are marked with red and blue dots to distinguish them from unused grey dots (Figure 5.21b). The system also provides a visualization of reprojection errors for each detected corner, illustrating the discrepancy between detected and computed estimated positions (Figure 5.21c).

Table 5.2 presents a comparative analysis of camera calibration parameters obtained through this post-processing approach, contrasting underwater parameters with those derived from aerial applications. These parameters include focal length (f), main point offsets (c_x, c_y), symmetric radial distortion coefficients (k_1, k_2, k_3), decentration distortion coefficients (p_1, p_2), and affinity and non-orthogonality distortion coefficients (b_1, b_2), demonstrating the significant differences in optical characteristics between underwater and aerial environments. Note the presence of additional coefficients (k_4, b_1, b_2) required for underwater calibration due to the increased optical complexity in the aquatic medium. Values were obtained by manual calibration for underwater parameters and self-calibration for aerial parameters using Agisoft Metashape Professional software.

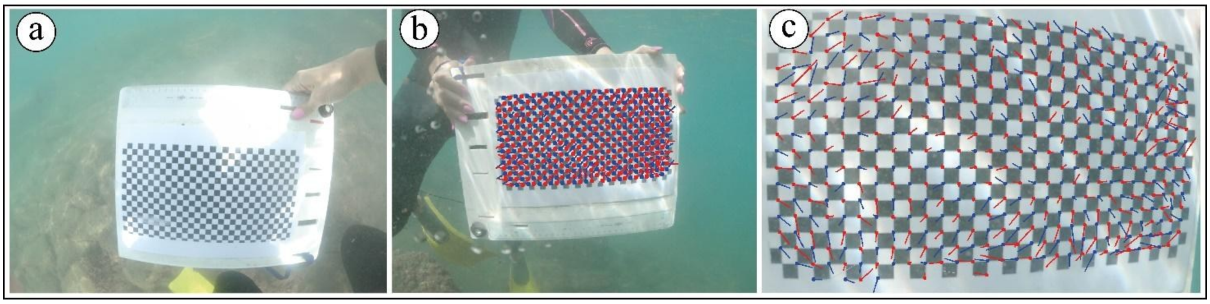

Figure 5.21: Underwater camera calibration and quality assessment. a) Physical setup of the Metashape calibration chessboard mounted on a 1 cm thick plastic table for underwater camera calibration. The chessboard dimensions are 40.9 cm × 27.25 cm, with square sides of 1.65 cm; b) software visualization of the calibration process showing the detected chessboard corners: red and blue dots indicate successfully detected corners used for calibration calculations, while grey dots represent unused detections. The alternating colours ensure that each point can be visually distinguished from its neighbours for quality control;(c) detailed view of a detected chessboard corner illustrating the reprojection error calculation, which represents the spatial difference between the detected corner position and its calculated estimated position.

Table 5.2: Comparative analysis of camera calibration parameters obtained from the GoPro Hero 6 camera system under different operating conditions (underwater versus air). Parameters include focal length (f) in millimetres, centre point offsets (c_x, c_y) in pixels, symmetric radial distortion coefficients (k_1, k_2, k_3, k_4), decentration distortion coefficients (p_1, p_2), and affinity and non-orthogonality distortion coefficients (b_1, b_2).

Calibration parameters	**GoPro Hero 6 Underwater**	**GoPro Hero 6 Air**
f (mm)	1767.30016	1755.1672
c_x (pixel)	-10.9266	-3.04698
c_y (pixel)	-3.14622	11.3176
k_1	-0.0486126	-0.229273
k_2	0.573766	0.0580051
k_3	-1.26848	-0.00672646
k_4	0.979334	-
p_1	0.000139304	0.000139304
p_2	0.000548803	-0.000815481
b_1	0.419889	-
b_2	-2.90859	-

5.6.3. Insta360 Studio software

Insta360 Studio is a software used for processing and editing 360-degree video content captured by Insta360 cameras (https://www.insta360.com/support/supportcourse?post_id=20713). In the case of scientific use, the software provides researchers with the essential tools to manipulate spherical shots for a range of scientific applications, including environmental monitoring, spatial analysis, and documentation. Insta360 Studio provides a simplified workflow for converting raw 360-degree footage into usable formats. Once the footage is imported, the software automatically merges video feeds from multiple lenses to create a cohesive spherical environment. For analysis purposes, the software allows conventional flat-format video to be extracted from the spherical content through a process called reframing. This feature allows researchers to generate multiple traditional perspective views from a single 360-degree capture.

The software's colour-correction tools provide options for adjusting brightness, contrast, saturation, and white balance throughout the spherical environment. These adjustments can be applied uniformly or selectively to resolve lighting inconsistencies commonly found in outdoor field recordings or challenging underwater conditions.

Export functionality supports multiple output formats and resolution options, including standard video files (MP4), equirectangular projections for specialized applications, and optimized versions for virtual reality platforms. This versatility ensures compatibility with different analysis tools and presentation methods commonly used in academic research.

In order to exploit the videos recorded during the survey campaigns, the start and stop points in the video were selected and compared with the time-lapse images taken in the same area and at the same time. Considering that these videos are not associated with GPS coordinates, they were used as a reference to associate pieces of video with time-lapse images of clearly recognizable objects in the photos or easily distinguishable parts of the rocky coastline. The videos were then cropped to 4:3 format. The video is then loaded directly into the Metashape software, where the frames are extracted.

5.6.4. CloudCompare software

CloudCompare is a 3D processing platform designed for the manipulation and analysis of point clouds and triangulated meshes. Initially developed for comparative analysis between dense 3D point clouds from laser scanners or between point clouds and triangular meshes, the software uses a specialized octree structure optimized for these operations (https://www.cloudcompare.org/presentation.html). Over time, it has evolved into a comprehensive processing tool, incorporating numerous advanced algorithms for registration, resampling, attribute management, statistical computation, sensor integration, segmentation, and visualization enhancement (Girardeau-Montaut, 2024).

In this research, CloudCompare served as the primary analytical tool for extracting topographic profiles from integrated 3D coastal models. The workflow began with exporting finished models from Agisoft Metashape in .obj format, preserving texture information and spatial coordinates. Once these files were imported into CloudCompare, the software's visualization capabilities facilitated inspection of the models.

The profile extraction process followed a systematic procedure in which the integrated 3D model, combining both above and below water components, was first prepared for analysis by vertex selection. The "Extract cloud section along polylines" function was used to create custom transect paths to capture key morphological features of the coastal area. The software's "Generate orthogonal section along polylines" function was then used with specific step distance and width parameters, followed by point extraction along active sections with appropriate thickness settings and edge length constraints. Finally, the "Extract points along active sections" function allows you to define the width and edge length of the final topographic profiles of the model. The profiles can be selected or not in the left dashboard of the software interface.

Also, the methodology for defining eroded areas of the tidal notch used a systematic approach to isolate and analyze it. A polyline was created by selecting four points around the area of interest, from which a best-fit plane was generated. This plane served as the reference surface for calculating the perpendicular distances of each point in the model. The parameters of the plane, including slope and size, were adjusted to obtain a plane that was as cross-sectional as possible to the model. The segmentation process isolated the portion of the model containing the tidal notch while preserving the original model data. The Cloud to Mesh Distance Calculator was then applied to both the segmented model and the reference plane, producing a color-coded visualization with distance values displayed on an RGB scale.

6. Study areas

This thesis investigates three case studies of three different Mediterranean regions, shown in Figure 6.1: the carbonate coast of Cala Uccello (Lampedusa, Pelagie Islands) (Section 6.1) and the limestone reefs of the islands of Mġarr ix-Xini (Gozo) and Halq it-Tafal, Malta (Malta) (Section 6.2). These areas were selected based on their rocky coastal setting and interesting geomorphological features, including tidal notches. Furthermore, the coastal features are well preserved and accessible for data collection using GoPro cameras, 360-degree recording systems and, for the Malta study area, UAV photogrammetry.

These sites are characterized by plunging cliffs and sloping coasts, in line with the classification of coastal geomorphotypes proposed by Biolchi et al. (2016a), which characterizes coastal features based on topographic, lithostructural, and geomorphological characteristics, extending Bird's (2008) classification to include sloping coasts, boulders, reefs, pocket beaches, and built-up coasts.

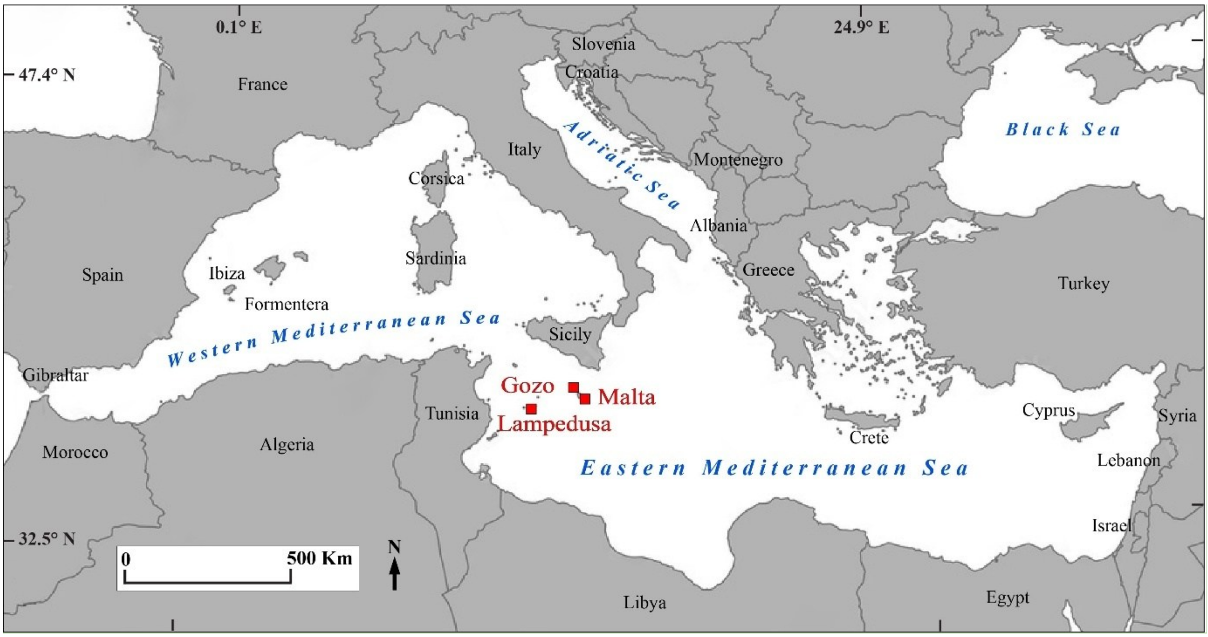

Figure 6.1: Mediterranean map showing the study areas in Lampedusa (Pelagie Islands), Gozo and Malta (Maltese Archipelago).

6.1. Cala Uccello, Lampedusa (Pelagie Islands)

Lampedusa, the southernmost Italian island, is positioned in the Sicilian Channel of the central Mediterranean Sea at 35°30'N, 12°36'E (Figure 6.2). The 20.2 km² island extends 10 km E-W and has approximately 33 km of predominantly rocky coastline. Its diverse geomorphological features reflect the complex geological history and distinctive marine climate setting of the island (Antonioli et al., 2011).

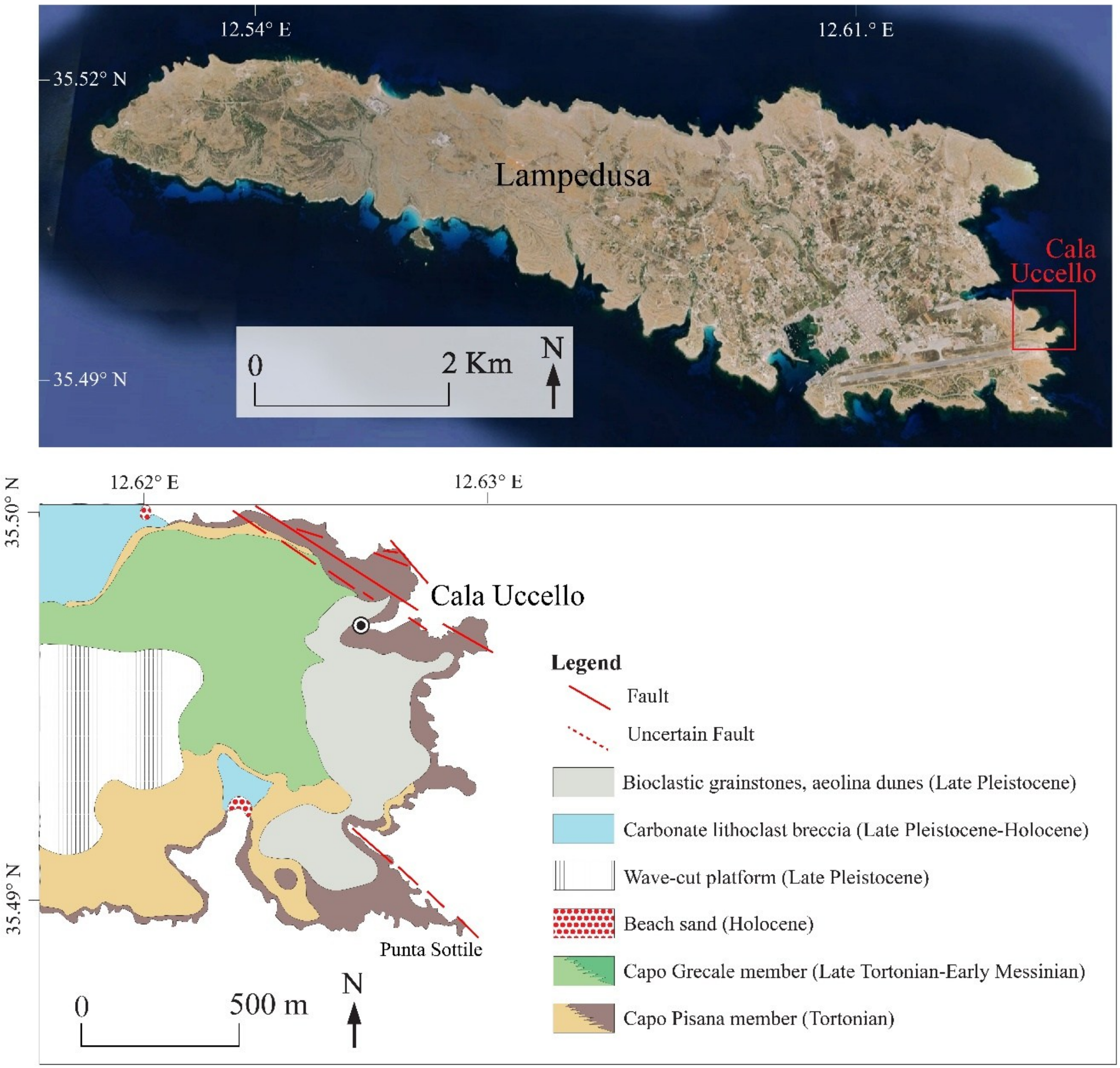

Figure 6.2: The study site of Cala Uccello (Lampedusa) and the geological map (redrawn from Grasso and Pedley, 1988).

Lampedusa lies on the African lithosphere within the Pelagian Block. This structural unit experienced significant crustal stretching during the Neogene-Quaternary period. This resulted in an intraplate rift system that extends across the central Mediterranean (Civile et al., 2010; Lentini et al., 1995). Its position on the undeformed African foreland has resulted in remarkable tectonic stability throughout the Quaternary period. Significant vertical displacements occurred primarily during the Early Pleistocene Calabrian stage (Catalano et al., 1993; Civile et al., 2021).

The geological composition of Lampedusa is characterized by a well-preserved Tortonian-Messinian succession of porite coral boundstones and bioclastic calcarenites (Figure 6.2). These deposits are rich in marine fossils, including mollusks, bryozoans, echinoids, and algal fragments (Grasso et al., 1985; Grasso & Pedley, 1988). The island is the emergent part of a larger carbonate plateau that trends from northwest to southeast. This plateau is structurally organized as a horst composed of Neogene-Quaternary carbonate sequences (Pedley & Grasso, 1993).

This geological framework has created a coastal morphology characterized by alternating resistant limestone headlands and more erodible calcarenite sections exhibiting differential erosion patterns.

The Cala Uccello study area, located on the eastern part of the island, exemplifies the diverse coastal geomorphotypes characteristic of Mediterranean carbonate environments. Geologically, the site is composed of Pleistocene-Holocene carbonate lithoclast breccias and Late Pleistocene wave-cut platforms belonging to the Capo Grecale and Cala Pisana members.

Major tectonic lineaments create structural controls on coastal morphology here (Grasso & Pedley, 1988).

The dominant geomorphotype of Lampedusa consists of plunging cliffs with vertical to sub-vertical limestone faces (slope angles of 75–90°) that descend directly into deep water (>8 m depth) without developing extensive shore platforms. These cliffs reach maximum heights of 45 meters above mean sea level and exhibit well-developed joint systems oriented northwest-southeast and northeast-southwest (Antonioli et al., 2011). The cliff faces show evidence of multiple erosional processes, including mechanical wave action, salt weathering, and biochemical dissolution processes, which are characteristic of Mediterranean carbonate coasts.

As Figure 6.3 shows, Cala Uccello bay has a sloping geomorphology with well-preserved tidal notches carved at mean sea level (0.0 ± 0.05 m). These notches exhibit typical Mediterranean microtidal characteristics.

Figure 6.3: The Cala Uccello Bay, Lampedusa.

In the more sheltered portions of Lampedusa, limited shore platform development occurs where horizontal to sub-horizontal surfaces (slopes of 0–3°) extend seaward from cliff bases for maximum distances of 15–20 m. These platforms are carved from the more resistant Capo Grecale limestone member and show typical Mediterranean shore platform characteristics, including well-developed solution pans, biogeochemical weathering features, and extensive bioerosion galleries created by the boring activities of *Lithophaga lithophaga* (Furlani et al., 2011).

The structural geology of Lampedusa has facilitated the development of coastal cave systems, with the largest extending approximately 25 meters inland from the present shoreline. These caves contain marine and subaerial speleothems, indicating the complex interactions between sea-level fluctuations and groundwater processes during Quaternary climatic cycles (Antonioli et al., 2011). Storm-deposited carbonate boulders ranging from 0.5 to 2.5 meters in maximum diameter occur sporadically along the cliff base and within coves. These boulders represent extreme wave events that transported them from the adjacent shore platform and cliff face. The largest boulders show evidence of multiple transport episodes, and biogeochemical weathering patterns indicate residence times of several decades to centuries (Borzì et al., 2024). Due to its position in the central Mediterranean, Lampedusa is exposed to a complex wave climate influenced by regional wind patterns and the Sicilian Channel's distinctive hydrodynamic characteristics. The surrounding marine environment is a high-energy setting in which modified Atlantic Mediterranean Water flows eastward in a 200-meter-thick surface layer that subsequently divides into the Atlantic Ionian Stream and the Atlantic Tunisia Current (Astraldi et al., 2001; Poulain et al., 2012).

Long-term wave statistics for the Lampedusa area, derived from ERA5 reanalysis data and validated against nearby wave buoy measurements, indicate a complex, directional wave climate with significant seasonal variations (Barbariol et al., 2019). During the winter months (October–March), the dominant wave directions are from the northwest (300–330°). These directions are associated with the passage of Atlantic frontal systems and regional cyclogenesis in the western Mediterranean. The summer months (May–September) are characterized by lower wave energy conditions, with predominant directions from the northeast (30–60°) and southeast (120–150°).

Significant wave height statistics demonstrate pronounced seasonal variability. Winter mean values range from 1.2 to 1.8 meters, with maximum recorded heights exceeding 6.5 meters during extreme storm events. Summer conditions are characterized by mean values ranging from 0.6 to 1.0 meters (De Leo et al., 2021). Peak wave periods range from four to six seconds during local wind-generated conditions to eight to twelve seconds for Atlantic swell penetrating through the Strait of Gibraltar. Cala Uccello's eastern orientation provides partial protection from dominant northwestern storm waves but leaves the site exposed to intense wave action during southeastern (scirocco) wind events and storms approaching from the Ionian Sea.

Historical storm analyses indicate that significant wave heights exceeding 4 meters occur with return periods of approximately two to three years, while extreme events ($H_s > 6$ meters) have return periods of 15 to 20 years (Lionello et al., 2006). The semi-enclosed nature of the Mediterranean Sea creates distinctive fetch-limited conditions that influence wave development around Lampedusa. Maximum effective fetch distances range from 800 km for northwestern approaches to 350 km for southeastern directions. This results in wave energy distributions that differ significantly from those in open-ocean environments (Cavaleri & Bertotti, 2004).

Lampedusa shows typical Mediterranean climate, which is characterized by pronounced seasonality in temperature and precipitation patterns. Its southern latitude (35°30'N) places the site at the interface of European and African climatic influences, creating the environmental conditions for coastal landform development. Meteorological data from the ENEA climate monitoring station (2010–2023) reveal significant seasonal variations in precipitation. December records the most rainfall (61 ± 12 mm), while July typically receives no measurable precipitation (ENEA, 2023). Annual precipitation averages 315 ± 45 mm, concentrated primarily from October through March.

This precipitation seasonality creates distinctive weathering cycles, with intense winter rainfall promoting chemical weathering and karst processes. This is followed by prolonged summer drought periods that enhance salt crystallization and thermal expansion processes in coastal rock formations. The wind regime at Lampedusa is dominated by northwestern (mistral) winds during the winter months. Mean velocities are 15–25 km/h, and maximum gusts exceed 80 km/h during storm passages (Pirazzoli et al., 1997). Summer conditions are characterized by more variable wind patterns, including thermal land-sea breeze circulation and periodic scirocco events from the southeast (Figure 6.4).

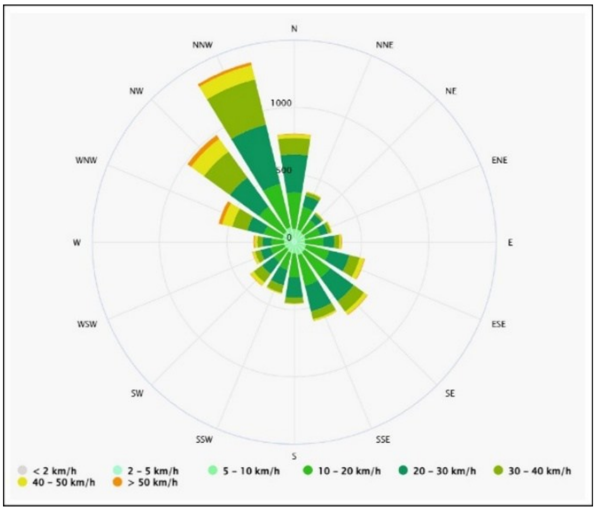

Figure 6.4: The Lampedusa wind rose (35.50°N, 12.61°E, 16 m asl. ERA5T) (https://www.meteoblue.com/en/weather/historyclimate/climatemodelled/lampedusa_italy_2524459).

The mean annual air temperature is 19.5°C, ranging from a mean of 13.2°C in January to a mean of 26.8°C in August. Sea surface temperatures range from winter minimums of 15–16 °C to summer maximums of 26–27 °C, creating thermal stress cycles that contribute to rock weathering processes through expansion-contraction mechanisms (Antonioli et al., 2011). The maximum tidal amplitude is approximately 40 cm, which is typical of microtidal conditions in the Mediterranean. These conditions concentrate erosional processes within narrow vertical zones and facilitate the development of distinctive tidal notch morphologies observed at the site.

6.2. The Maltese Archipelago Sites: Mġarr ix-Xini and Halq it-Tafal

The Maltese archipelago, which is located approximately 96 km south of Sicily and 290 km from North Africa, hosts two study sites: Mġarr ix-Xini in Gozo (36°01'48"N, 14°20'12"E) and Halq it-Tafal in Malta (35°52'36"N, 14°22'45"E) (Figure 6.5). These two sites were chosen because they both have a plunging coastline with a characteristic tidal notch.

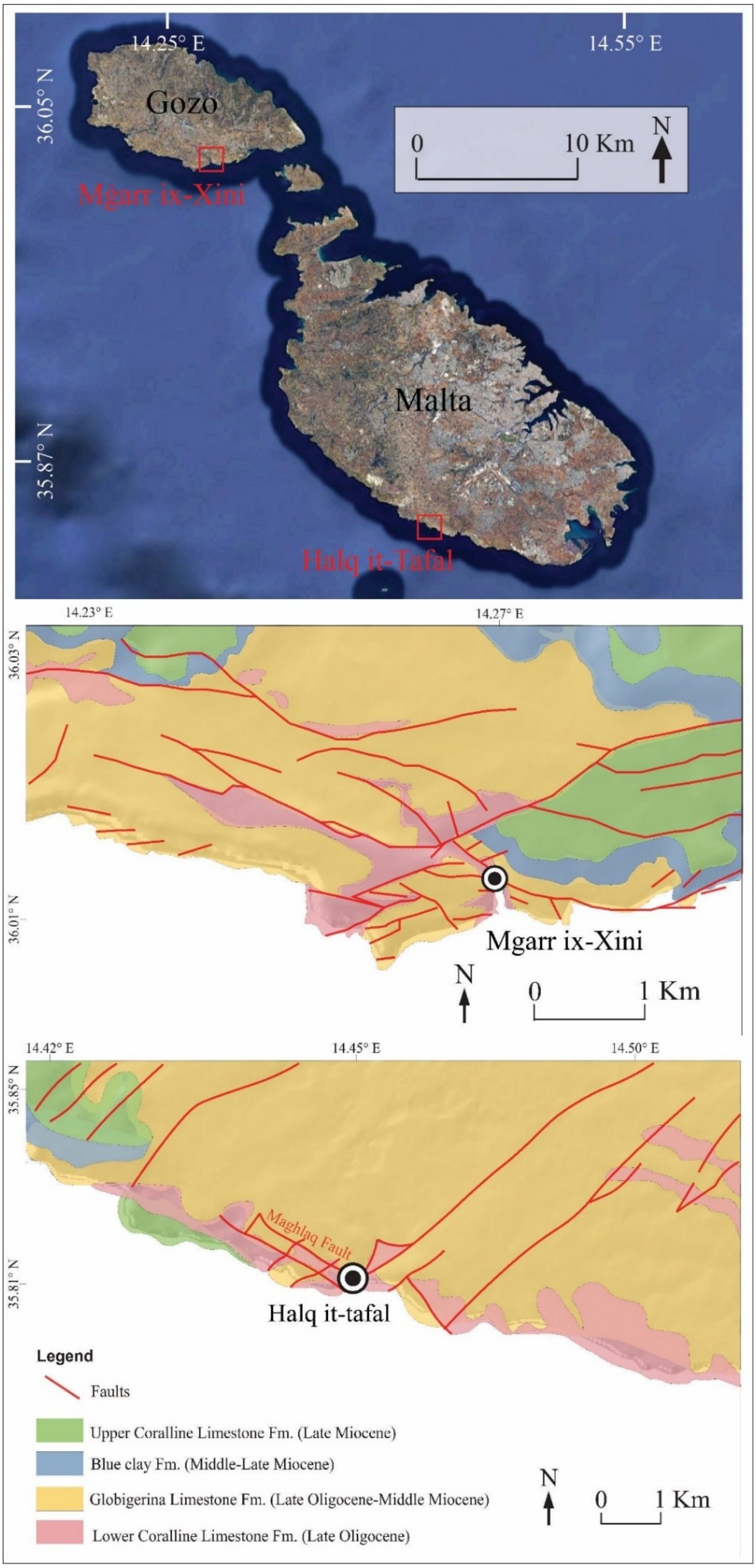

Figure 6.5: The study sites of Halq it Tafal (Gozo) and Mġarr ix-Xini (Malta) (redrawn after Directorate. O.E., 1993) and the respectively geological maps.

The archipelago is located on the Malta-Ragusa Rise, a significant structural feature that is part of the submarine ridge system that extends from the Ragusa peninsula in Sicily to the continental shelf of North Africa (Schembri et al., 2009). This geological setting places the islands on the northern flank of the northwest-southeast-trending Pantelleria rift system, a major extensional structure that experienced significant continental rifting during the Neogene-Quaternary period (Dart et al., 1993; Finetti, 1984). The complex tectonic history has created distinctive structural controls on coastal morphology that vary significantly between the two study sites.

The Maltese Islands' geological foundation comprises a well-defined sedimentary sequence spanning the Upper Oligocene to Upper Miocene periods. This sequence is important for investigating coastal evolution in carbonate-dominated lithologies (Baldassini & Di Stefano, 2017). The stratigraphic succession consists of five principal formations that exert fundamental control over coastal geomorphotype development. The Lower Coralline Limestone (Late Oligocene) consists of bioclastic, bedded, gray limestones that are highly resistant to wave erosion. The Globigerina Limestone (Late Oligocene to Middle Miocene) consists of soft, yellowish, massive, fine-grained biomicrites that erode rapidly and exhibit distinctive weathering patterns. The Blue Clay Formation (Middle to Late Miocene) consists of alternating layers of dark gray and light gray marls that create slope instability conditions when exposed along the coast. The Greensand Formation (Late Miocene) consists of marly, clayey sands and arenites with intermediate erosional resistance. The Upper Coralline Limestone (Late Miocene) consists of highly resistant bioclastic limestones that form the most prominent coastal cliffs and headlands throughout the archipelago (Figure 6.5).

The structural framework that controls coastal morphology reflects two distinct tectonic episodes that have profoundly influenced landscape evolution (Martinelli et al., 2019). The older structural system is oriented ENE-WSW and has been active since the Lower Miocene. It has created a distinctive pattern of ridges and grabens that controls the orientation of major coastal embayments and headland systems. The younger structural system is dominated by the northwest-southeast trending Maghlaq fault system, which is directly associated with Pantelleria Rift activity (Alexander, 1988). This complex tectonic history produced the archipelago's characteristic southwest-northeast tilt, creating asymmetric morphology. This morphology is characterized by spectacular high cliffs (up to 250 m) along the southern and western coasts, which contrast with the low-lying coastal platforms and embayments along the eastern shores.

The Mġarr ix-Xini study site, located on the southern coast of Gozo, occupies a major structural valley carved by ENE-WSW trending fault systems, which create a distinctive, funnel-shaped coastal geometry. The site extends approximately 800 m along the coastline and has a maximum inland penetration of 400 m, creating a semi-enclosed marine environment. The Mġarr ix-Xini area is dominated by Upper Coralline Limestone formations, which form vertical to sub-vertical cliffs reaching heights of 45–60 m (148–197 ft) above mean sea level along the headland margins (Figure 6.6).

Figure 6.6: The study sites of Mġarr ix-Xini, Malta.

The coastal geomorphotypes at Mġarr ix-Xini demonstrate the complex interaction between structural geological controls and Mediterranean coastal processes. Plunging cliffs characterize the headland areas, where the formations create vertical faces with slope angles of 85-90° that descend directly into deep water (8-15 m deep) without significant shore platform development. These cliffs exhibit well-developed joint systems oriented parallel to regional structural trends. These systems create distinctive rectangular joint block patterns that influence mechanical wave erosion and gravitational slope processes (Devoto et al., 2013).

The site hosts tidal notch development that provides precise indicators of recent sea-level positions and coastal evolution rates (Furlani et al., 2024).

The Halq it-Tafal study site is located on Malta's northeastern coast, where NW-SE trending fault systems intersect with the regional ENE-WSW structural grain. This creates a complex coastal geometry characterized by alternating headlands and embayments. Intensive fracturing related to the intersection of multiple fault systems is evident in the structural geology, creating rectangular joint patterns with spacings of 0.5–2.0 m that exert fundamental control over coastal erosion processes (Martinelli et al., 2019).

The coastal geomorphotype at Halq it-Tafal reflects this complex geological framework. Areas where Globigerina limestone formations are exposed exhibit extensive shore platform development, creating horizontal surfaces with slope angles of 0-3° that extend up to 80 m seaward from cliff bases. These platforms exhibit the most extensive development observed in the Mediterranean region, with total platform areas exceeding 15,000 m² at the main study site. These platforms display characteristic Mediterranean features, including well-developed solution pan networks, intensive bioerosion galleries, and distinctive, rim-like vermetid reef formations that create natural structures that dissipate wave energy (Chemello & Silenzi, 2011).

Several submarine freshwater springs were identified along the Gozo coastline during the 2013 Geoswim campaign (Furlani et al., 2017b), indicating active groundwater discharge that influences coastal processes. The combination of these environmental factors with the carbonate lithology and structural framework has resulted in diverse coastal geomorphotypes characterized by plunging cliffs, wave-cut platforms, and structurally controlled bays. The southwestern coast is characterized by continuous limestone plunging cliffs with well-developed tidal notches (Biolchi et al., 2016b; Furlani et al., 2013), while the northwestern coast is characterized by extensive landslides, primarily rock spreads and block slides, forming a distinctive coastal landform known locally as "rdum" (Devoto et al., 2012, 2013).

The contrasting wave energy environments experienced by the two study sites reflect their different coastal orientations and exposure to regional wave climate patterns. Mġarr ix-Xini, with its southern exposure, receives maximum wave energy during scirocco (southeastern) wind events and Atlantic swell propagation from the west. It experiences partial shelter from the dominant northwestern storm waves. Wave height statistics for the southern coast of Gozo, derived from ERA5 reanalysis data and validated against regional wave buoy measurements, indicate mean significant wave heights of 1.5–2.2 m in winter, with maximum values exceeding 7.5 m during extreme storm events (Barbariol et al., 2019). Summer conditions are characterized by lower energy, with mean significant wave heights of 0.8–1.2 m and predominant wave directions from the southeast and southwest.

In contrast, Halq it-Tafal exhibits more complex wave exposure patterns due to its northeastern orientation and location within a large embayment system. The site receives maximum wave energy during northeastern storm events associated with Bora winds and regional cyclogenesis in the central Mediterranean. Meanwhile, it remains relatively sheltered from Atlantic swell and scirocco events. Wave statistics indicate winter mean significant wave heights of 1.0–1.6 m, with maximum events reaching 5.5–6.0 m. Summer conditions are characterized by very low energy, with mean values of 0.4–0.8 m (De Leo et al., 2021).

The distinctive marine environmental conditions affecting both sites reflect the semi-enclosed nature of the central Mediterranean Sea and the complex bathymetric controls surrounding the Maltese archipelago. Water depths increase rapidly from both coastlines, reaching 50–80 m within 1–2 km of the shore. This creates distinctive, steep nearshore bathymetric profiles that influence wave shoaling and energy dissipation patterns. The regional oceanographic setting is characterized by complex current systems, including the eastward-flowing Atlantic Ionian Stream along the northern margins of the archipelago, and seasonal thermal stratification, which influences physical and biochemical coastal processes (Poulain et al., 2012).

Tidal conditions at both sites reflect the characteristic microtidal Mediterranean regime, with maximum spring tidal ranges of 0.206 meters and neap tidal ranges of 0.046 meters (Drago, 2009). This limited tidal range concentrates erosional processes within narrow vertical zones, facilitating the development of distinctive horizontal tidal notches that precisely indicate sea-level position and coastal evolution rates. However, the astronomical tidal signal is often modified by meteorological effects, including storm surge events, which can raise water levels by 0.3–0.5 meters during intense cyclonic storms.

Climate conditions at both study sites reflect classic Mediterranean characteristics, including pronounced seasonal variations that influence coastal geomorphological processes. Mean annual air temperatures range from 18.5 to 19.2°C, with winter minimums of 12 to 14°C and summer maximums of 26 to 28°C. These temperatures create thermal stress cycles that contribute to mechanical weathering through expansion and contraction processes (Furlani et al., 2017b). Sea surface temperatures exhibit similar seasonal patterns, ranging from 15-17°C in winter to 25-27°C in summer.

Precipitation patterns show marked seasonality, with annual totals of 530 ± 80 mm concentrated primarily from October through March. December is the wettest month, with an average of 95 ± 25 mm of rainfall, while July and August typically receive less than 5 mm of precipitation per month. These seasonal precipitation patterns create distinctive weathering cycles. Intense winter rainfall promotes karst processes and chemical weathering. This is followed by prolonged summer droughts that enhance salt crystallization and biochemical weathering processes in the coastal zone.

Wind climate patterns exhibit significant seasonal variations that directly influence wave generation and coastal energy regimes. The winter months are dominated by northwestern (Mistral) and northeastern (Bora) winds, which have mean velocities of 20–30 km/h and maximum gusts that exceed 100 km/h during storm passages. Summer conditions are characterized by more variable wind patterns, including thermal land-sea breeze circulation, with typical velocities of 10–20 km/h, and periodic scirocco events from the southeast (Lionello et al., 2006; see Figure 6.7).

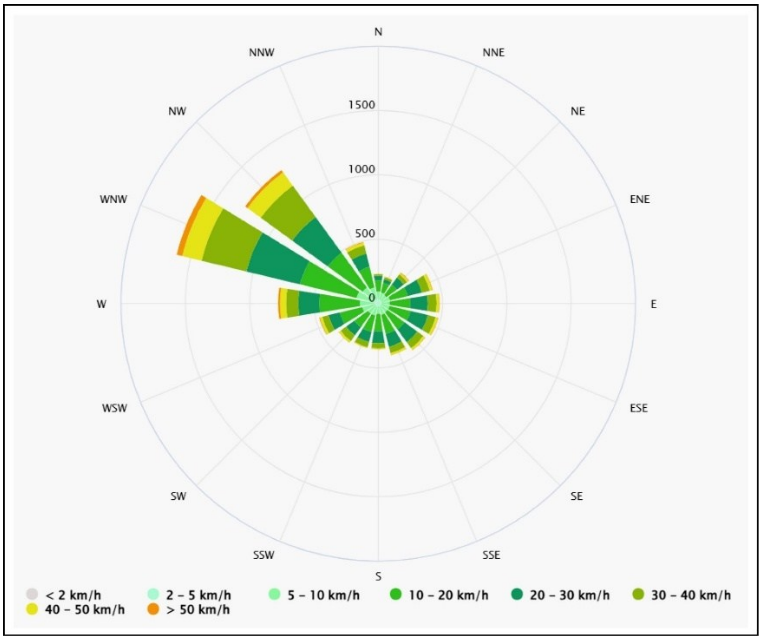

Figure 6.7: The Malta islands wind rose (35.88°N, 14.45°E, 60 m asl. ERA5T) (https://www.meteoblue.com/en/weather/historyclimate/climatemodelled/island-of-malta_malta_2562772).

While the Maltese archipelago is generally characterized by low to moderate seismicity, it remains subject to the effects of regional earthquake activity originating in the Sicilian Channel and eastern Sicily (Armigliato et al., 2007; Galea, 2007). The tectonic setting and seismic history have contributed to the development of distinctive coastal features, including extensive cave systems, natural arches, and gully networks, which characterize both study sites.

7. Results

This chapter outlines the results and findings of this research project, detailing a systematic approach to coastal image analysis and three-dimensional reconstruction.

The result section begins with an overview of the basic data preparation process in Section 7.1, which explains in detail the methodical organization and structuring of the coastal image dataset.

This careful organization provides the essential basis for all subsequent analyzes. Section 7.2 presents the newly organized Geoswim dataset and the development and implementation of the Geoswim Image Viewer. This is a tool designed to enhance the accessibility and visualization of coastal image collection.

The computational aspects of the research are detailed in Section 7.3, which outlines the methodology for building the CNNs model. This section includes information on data partitioning, normalization, and augmentation techniques. The results are presented in Section 7.4 for the binary classification and Section 7.5 for the multi-class classification. The results of the binary classification demonstrate the model's ability to automatically discriminate between coastal images that are useful and not useful for coastal analysis and modeling. In contrast, the multi-class classification extends this capability by identifying specific geomorphological features within the coastal environment, categorizing the images according to their coastal geomorphotype.

The final sections (7.6 and 7.7) of this chapter focus on three-dimensional reconstruction techniques. Section 7.6 examines the process of merging 3D models above and below the water surface, addressing the technical issues presented by the water-air interface. Section 7.7 then presents detailed case studies demonstrating the practical application and validation of our methods in different coastal environments.

7.1. Geomorphological survey and dataset curation

All results obtained during the survey campaigns are photographic (GoPro) or 360° video (Insta360). During the Lampedusa campaign, observational data was collected and summarized in an unavailable table. However, this table does not contain important data about the Cala Uccello site.

The process of organizing and categorizing the extensive Geoswim image dataset used for this thesis comprises approximately nine terabytes of visual data. The same dataset was then used for CNNs training, but only three folders were selected for 3D modeling and to demonstrate how the Geoswim image viewer works. These folders were chosen based on three different study areas.

This chapter delves into the manual curation approach employed to structure the collection of images, which forms the foundation for future automated classification using CNNs.

The dataset consists of images captured both above and below the waterline. The primary objective of the curation process was to create a logical and intuitive structure that would facilitate easy access, efficient compression, and seamless linking between related images (Figure 7.1).

The initial step in the curation strategy involved a comprehensive visual inspection of each image. This process required careful observation and selection based on specific criteria, such as recognizing pairs of corresponding images above and bw. In the end, each image captured aw corresponds to the images captured bw on a one-to-one basis.

Images were first organized by surveyed area, then by the day of the survey, and finally divided into aw (or asl) and bw (or bsl) categories. This hierarchical structure mirrors the spatial and temporal progression of the data collection. To further refine the dataset and prepare it for machine learning applications, an additional layer of categorization was introduced. Within each package of images (defined by location and survey date), we created a folder named "cancelled". Specifically, during field surveys, it can happen that "useless" photos are collected for the study of rocky coasts, such as blurry photos, photos with boats or ships, photos too far from the shore to allow good inspection, underwater photos with very cloudy water, etc. (Figure 7.2). All these images have been removed from the main folders above and below the water and moved to the deleted images folder. The images in these folders play a crucial role in teaching binary classification CNN to automatically distinguish between useful and useless images.

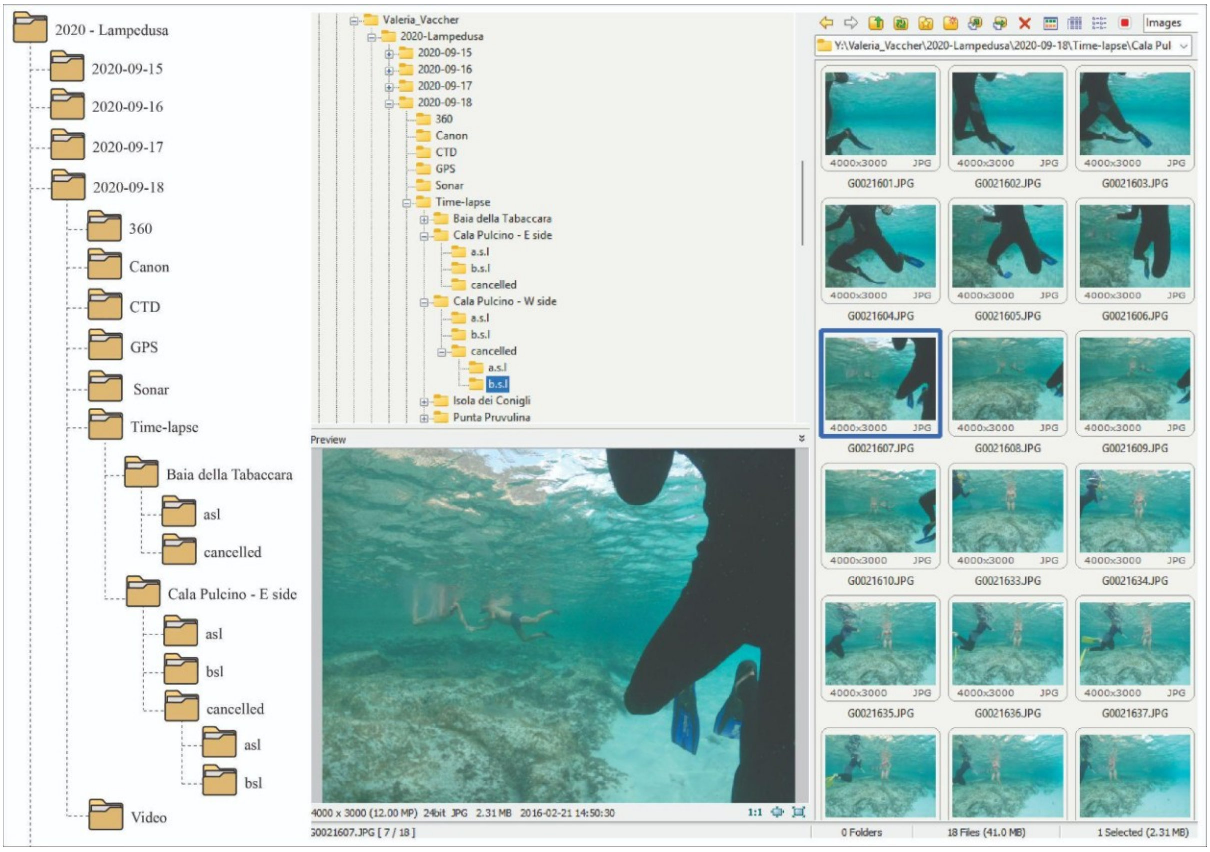

Figure 7.1: Example of Lampedusa dataset curation.

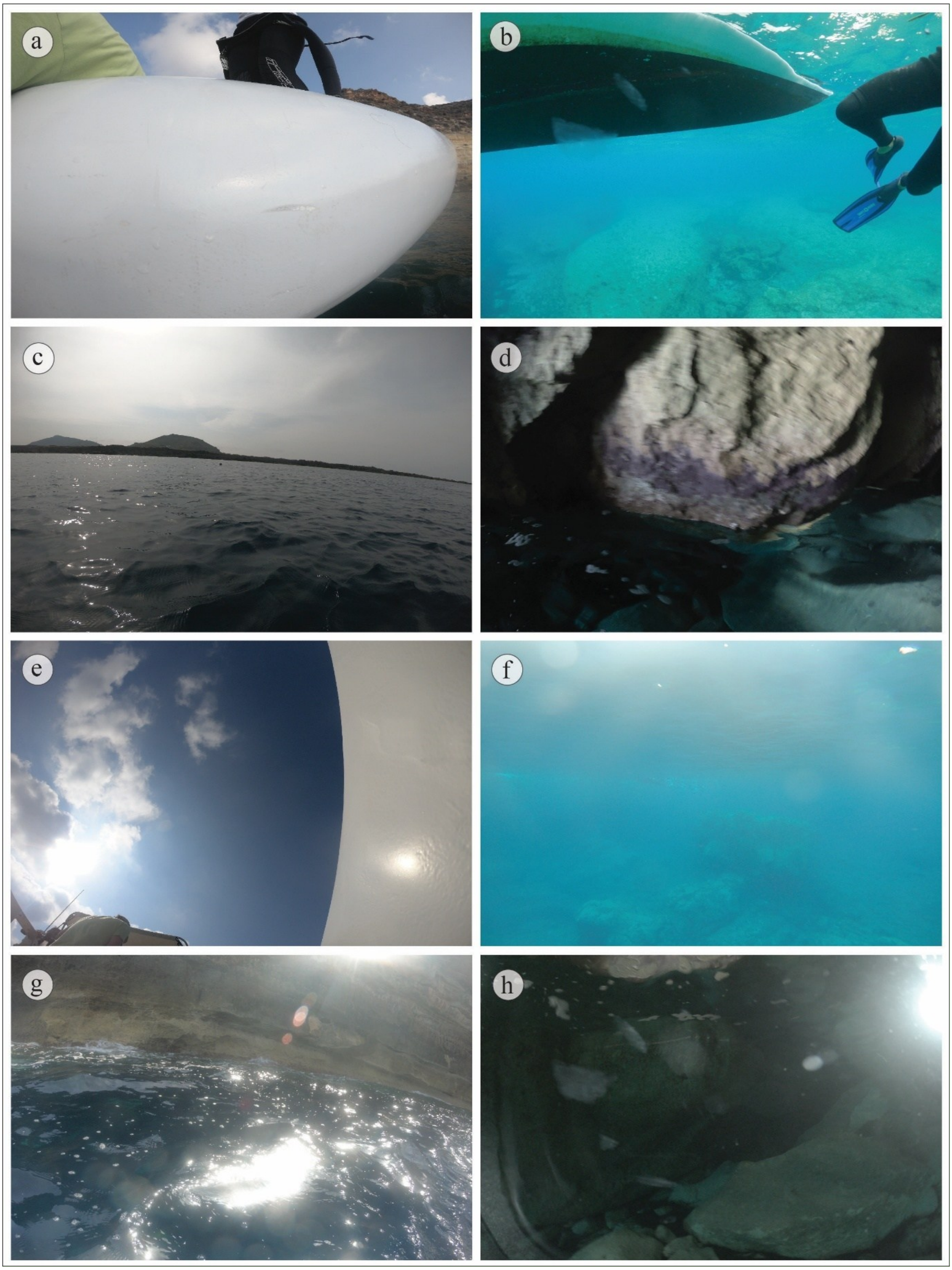

Figure 7.2: Example of images that are not useful in the Geoswim image archive and have been moved to the "cancelled" folder. a) Boat is the main feature; b) bw image with boat and person; c) too far away; d) blurred and dark image; e) sky and boat; f) cloudy bw image; g) overexposed aw image; h) blurred image.

7.2. Geoswim Images Viewer

This paragraph presents the Geoswim Images Viewer, a web interface designed primarily to simultaneously observe and analyze Geoswim images collected above and below water.

The images are linked sequentially and associated with a map that shows precisely where each was collected during the field survey, along with their corresponding EXIF data (date, time, latitude, longitude, altitude).

The Geoswim Images Viewer can be accessed via the URL https://valeriavaccher.com/geoswim. From this landing page, researchers can select specific geographical survey sites from the available dataset repositories. For expedited access to this resource, Figure 7.3 provides a corresponding QR code that, when scanned with an appropriate device, automatically navigates to the aforementioned web address.

As shown in Figure 7.3, the interface is divided into three main panels. The left panel (Map) shows the geographic context of the images, the middle panel (Images) shows the image pair, and the right panel (Information) shows EXIF data and navigation controls. The images retain their original aspect ratios to avoid distortion. This layout ensures that all information is displayed simultaneously for quick analysis and comparison.

The responsive design adapts to different screen sizes for desktop, tablet, and mobile by changing the displayed configuration of the three main columns.

The viewer provides intuitive navigation controls for efficiently browsing large image datasets. Users can move through images individually or in larger increments, select specific step sizes, jump directly to any image, and use arrow keys for quick navigation.

Geolocation data is provided with spatial context for each image pair, including an interactive map showing the location of the images currently being viewed and a marker that automatically updates, reflecting the current image location.

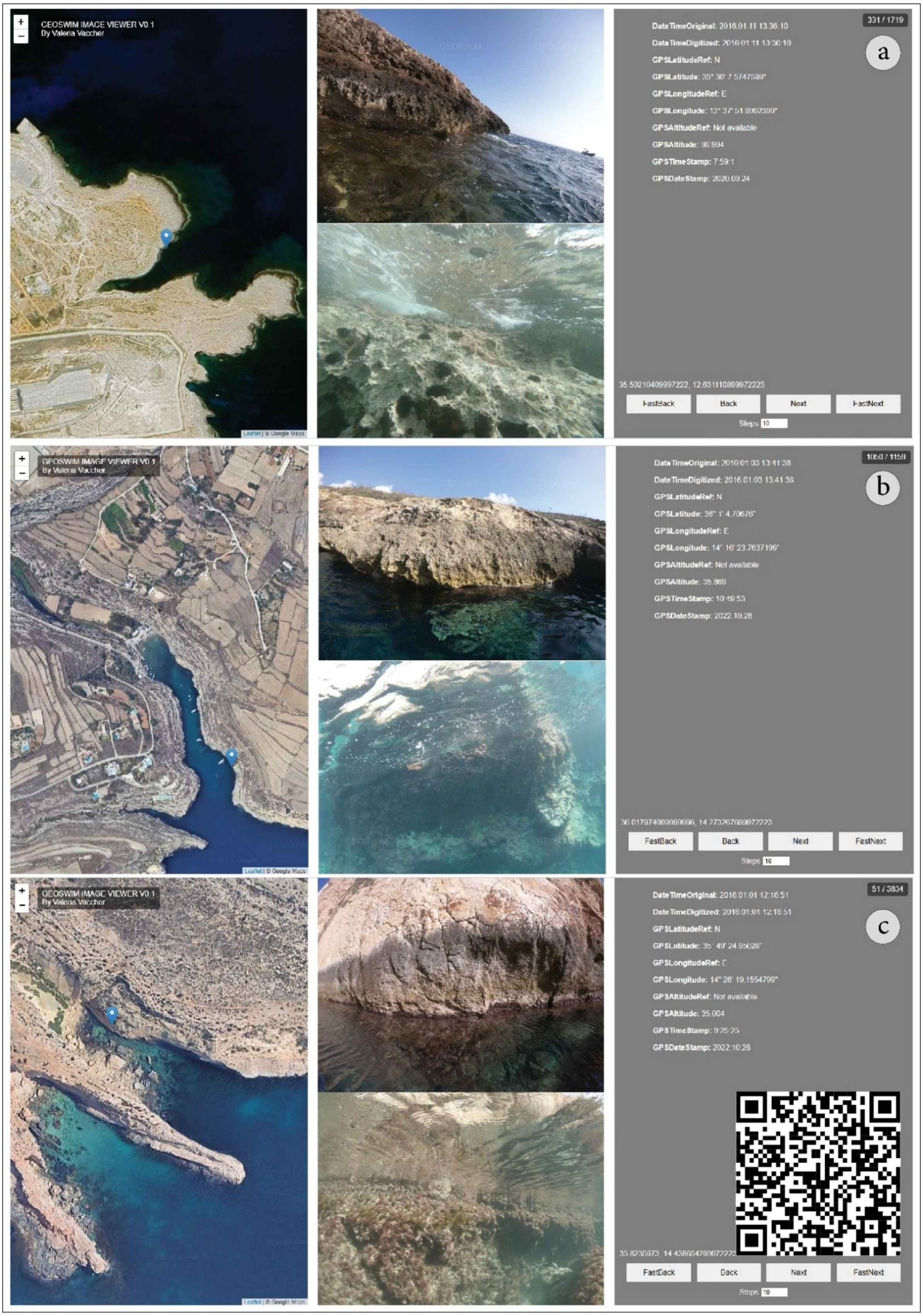

Figure 7.3: The interface features the map, images collected above and bw, and the EXIF data of Geoswim Image Viewer of the study area of a) Cala Uccello (Lampedusa), b) Mġarr ix-Xini (Gozo) and c) Halq it-Tafal (Malta). The QR code is the direct link to the web page that accesses the three sites studied.

7.3. Model training setup

This research develops and implements two different ResNet-152d, each addressing specific challenges in coastal image analysis and classification.

The first one implements a binary classification system that distinguishes between "useful" and "not useful" images within large time-lapse datasets. This automated filtering system significantly streamlines the data preparation process by efficiently identifying images suitable for further analysis. By automatically eliminating images compromised by poor visibility, obstructions, or technical issues, the system greatly reduces the manual effort required by researchers. This pre-filtering step not only accelerates the research workflow but also ensures the quality of subsequent analyzes by producing a refined dataset of high-quality images.

The second one goes beyond basic filtering to perform morphological classification of coastal landscapes. The classification proposed by Biolchi et al. (2016a) consists of seven geomorphological classes, as explained in Chapter 4.1. However, only the five classes present in the analyzed study areas were used in this thesis. This excludes classes such as cliffs and built-up coasts.

This CNN system can identify and categorize five distinct coastal morphological classes: plunging cliffs, sloping coasts, scree, pocket beaches, and shore platforms. This automated classification capability enables rapid qualitative and quantitative assessment of coastal morphology over large study areas.

7.3.1. Dataset organization for binary classification

The binary classification CNN was implemented and trained on the Marconi100 high-performance computing infrastructure. The training methodology involved 100 epochs, with each epoch representing a complete iteration through the training dataset. To enhance the generalization capabilities of the model, we implemented a data augmentation framework that incorporates stochastic transformations that generate distinct variations of the training samples during each epoch.

The implementation used mini-batch optimization with a carefully calibrated batch size of 16 samples. This specific batch size configuration was determined through empirical analysis to achieve an optimal balance between computational efficiency and model convergence characteristics. The mini-batch approach controls two critical aspects of the training dynamics: the frequency of network parameter updates and the resolution of loss function calculations during the backpropagation phase.

The first step for binary classification is the selection and setup of the dataset before feeding it into the network.

Dataset preparation is based on the organizational structure described in Section 7.1, which is the basis for training the network. The image collection consists of photographs taken with GoPro cameras mounted on a vertical pole on the ISR. These photographs were taken both underwater and above water, as described in detail in Section 5.1.2.

For binary classification, the images were systematically archived in the Geoswim database with extensive metadata including date, time, and location coordinates. The archive implements a hierarchical structure with three primary categories: above waterline (aw), below waterline (bw), and cancelled images. The cancelled category is further subdivided into two subcategories: cancelled aw and cancelled bw images. This classification was performed through careful manual inspection of individual images, creating a robust ground truth dataset for training the network to distinguish between “useful” and “not-useful” images (Figure 7.4).

To ensure a balanced representation across all classes, we randomly sampled 2500 images from each category (2500 aw images, 2500 bw images, 2500 aw cancelled images, and 2500 bw cancelled images), using a Python script (see Appendix 2), resulting in a total dataset of 10,000 images. Each image retains its original resolution of 4000×3000 pixels. The final dataset consists of 2500 valid “aw” images, 2500 valid “bw” images, 2500 “cancelled aw” images, and 2500 “cancelled bw” images. This balanced dataset structure was designed to avoid class bias during the training process and to ensure robust classification performance across all categories.

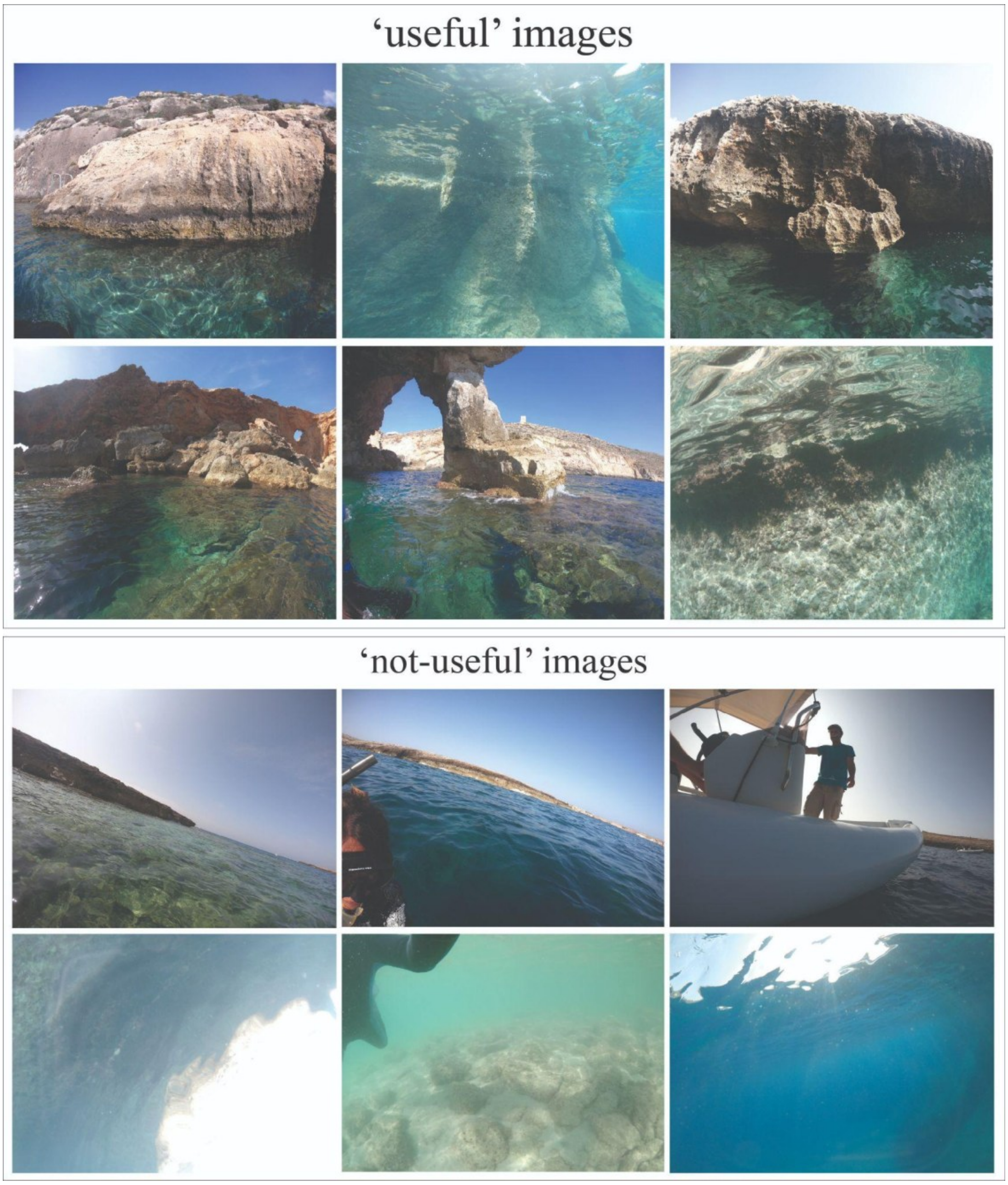

Figure 7.4: Example of images randomly selected as 'useful' and 'not useful' from the dataset.

7.3.2. Dataset organization for multi-class geomorphological classification

The multi-class classification CNN was implemented and trained on the Marconi100 high-performance computing infrastructure. The training methodology involved 500 epochs, each corresponding to a full pass over the entire training dataset. An iteration is an individual step during training in which the network updates its weights after processing a batch of data. The number of iterations is obtained by dividing the total number of examples in the dataset by the batch size, then multiplying that number by the number of epochs.

The implementation used mini-batch optimization with a carefully calibrated batch size of 4 samples. Various configurations were tested until the parameter for the best configuration was identified as 4.

Building on the foundation of binary classification, we developed a comprehensive multi-class classification system to classify coastal geomorphology into five categories. The dataset includes images of the five aforementioned classes: plunging cliff (380 images), pocket beach (14 images), scree (284 images), shore platform (65 images), and sloping coast (267 images), all from the same Geoswim dataset used in the binary classification phase (Figure 7.5).

The image selection and labelling process prioritized morphological distinctiveness over numerical balance between classes. While this approach resulted in an unbalanced dataset - the images are not equal in number for each class - it ensured that each class contained high-quality examples that clearly demonstrated the defining characteristics of their respective geomorphological features.

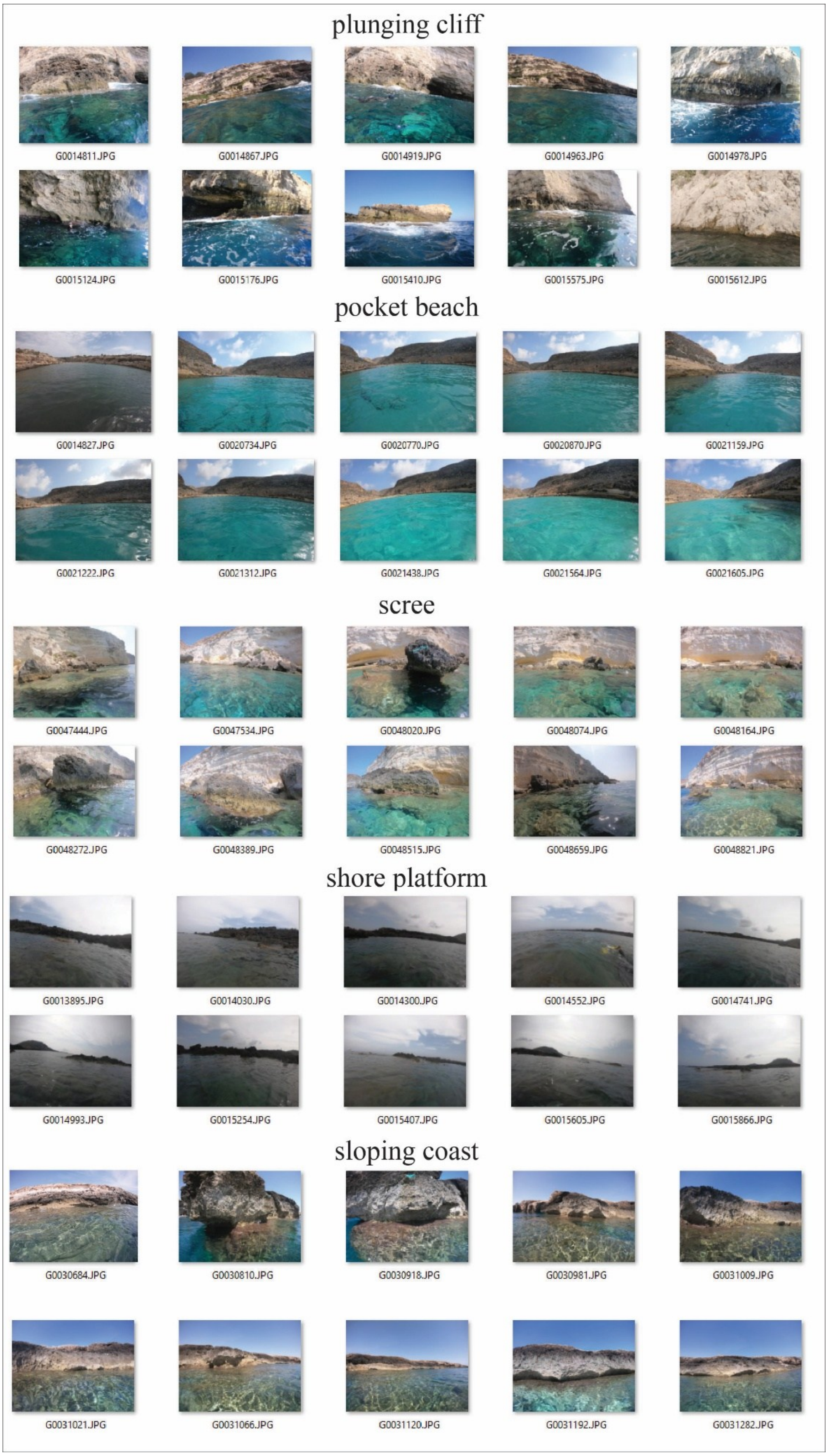

Figure 7.5: Example of manually selected images for each of the five geomorphological classes from the Geoswim dataset.

The machine learning framework shown in Figure 7.6 implements a systematic approach divided into three main components. The data management structure follows standard machine learning practices, partitioning the dataset into training, validation, and test sets. In addition, a separate "NEW" pathway was included, designed to evaluate previously unseen images, which enhances the practical applicability of the model.

For the development of the DL part, two ResNet configurations were implemented: a primary network for training and general performance evaluation, and a secondary network equipped with GradCAM capabilities for new image classification and model interpretation. This dual-network approach enables both robust classification performance and transparent decision making.

The performance evaluation framework uses F1 scores as the primary metric for assessing classification accuracy across all five geomorphological classes (C1 to C5). Our model achieved significant success across all categories, demonstrating its effectiveness in distinguishing between different coastal morphotypes despite the inherent challenges of an unbalanced class distribution.

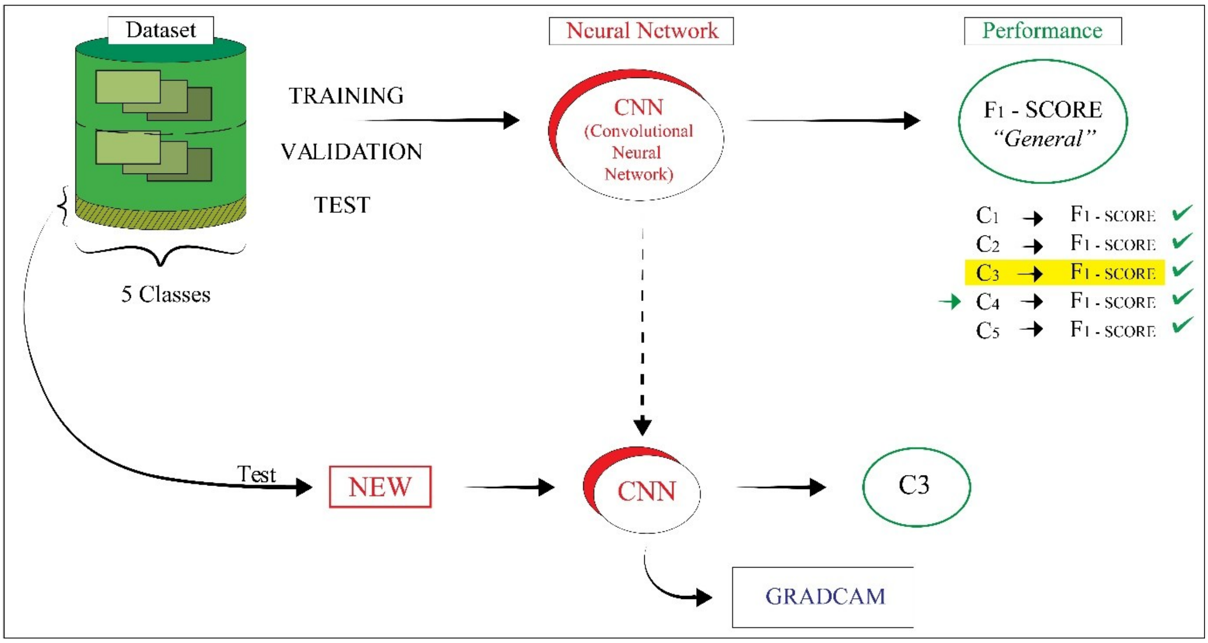

Figure 7.6: Workflow of the CNN for multi-class image classification. The diagram is divided into three main sections that represent a complete machine learning framework, from data input, through training and evaluation, to the final classification and interpretation tools of a dataset of images of the intertidal zone.

7.3.3. Dataset Split and Data Normalization

The dataset pre-processing pipeline was implemented using the transformation functions of the Torch Vision library, incorporating both normalization and data partitioning strategies to optimize the network training. The pre-processing workflow was designed to ensure data consistency, reduce computational complexity, and improve model learning capabilities.

The normalization process standardizes the input data by transforming pixel values across all images to achieve a mean of 0.0 and a standard deviation of 1.0. Each image in our dataset is represented as a tensor of shape (*C, H, W*), where *C* represents the colour channels (RGB), H represents the image height, and W represents the image width.

The normalization procedure is applied to each colour channel independently according to the formula:

$$\omega_{C_i} = \frac{x_{C_i} - \mu_{C_i}}{\sigma_{C_i}}$$

where ω_{C_i} represents the normalized pixel values, x_{C_i} denotes the original pixel values, and μ_{C_i} and σ_{C_i} represent the mean and standard deviation of channel i, respectively. The specific normalization parameters were determined for both developed CNNs. For the first binary classification, mean = [0.3704, 0.5237, 0.5401] and std = [0.2422, 0.2196, 0.2341]. For the second multi-class classification, mean = [0.4278, 0.4638, 0.4692] and std = [0.2542, 0.2307, 0.2524]. These values reflect the inherent characteristics of our image dataset, with the variation in means and standard deviations across channels indicating different intensity distributions in RGB colour space (e.g., 0.4278 is the average of the values contained in the red channels of all pixels of all images).

Both datasets were systematically partitioned following established machine learning practices, with an 80-10-10 partitioning ratio implemented for the training, validation, and test sets, respectively. This partitioning resulted in 8,000 images for training, 1,000 for validation, and 1,000 for testing. The training set configuration was optimized for efficient processing, implementing batch-wise loading with shuffle enabled for the training set to ensure random sampling during training iterations, while keeping the validation and test sets in their original order to ensure reproducible evaluation metrics.

The implemented normalization strategy provides more advantages for network training. Data normalization effectively constrains input values within a consistent range, significantly reducing the impact of data outliers and extreme values that could otherwise dominate the learning process. The reduction in data skewness achieved through normalization leads to improved model convergence rates, which is particularly important when dealing with all the images, where lighting conditions and water turbidity can cause significant variations in pixel intensity distributions.

In addition, normalized inputs play a critical role in overcoming the vanishing and exploding gradient problems commonly encountered in deep networks. By maintaining controlled input ranges, the normalization process helps ensure that gradient values remain within manageable limits during backpropagation, contributing to more stable and efficient training dynamics. The process also helps reduce internal covariate shift, where the distribution of network activations changes during training as parameters of previous layers are updated. This stabilization of internal representations allows for higher learning rates and faster convergence during the training process.

The pre-processing pipeline was implemented using PyTorch's transform capabilities, specifically the 'torchvision.transforms' module. Normalization parameters were computed over the entire dataset prior to partitioning to ensure consistent normalization across all data partitions.

7.3.4. Data augmentation

The implementation of data augmentation techniques proved essential to improve the robustness and generalization capabilities of the model. This methodology expands the effective training dataset through the systematic application of geometric transformations, while preserving the fundamental features that define each image class.

To perform the data augmentation was used the Pytorch library. Data augmentation includes a set of geometric transformations such as rotation and bidirectional flipping, spatial adjustments through translational shifts, perspective modifications, and basic image adjustments including brightness, contrast, and saturation variations. In particular, the perspective transformation, a specific form of planar homography, played an important role in our augmentation strategy. This transformation establishes a mathematical relationship between different projections of the same three-dimensional object into different projective planes. The implementation of perspective transformations effectively simulates the natural variations in camera positions and orientations that occur during underwater surveys. This approach mimics human visual perception, where object size varies with distance, thereby creating a more naturalistic training dataset.

The enhancement protocol was carefully calibrated to maintain image validity while introducing meaningful diversity into the training data. The transformed images retained their essential features and class labels, while presenting the network with variations that reflect real-world underwater imaging conditions. This controlled introduction of variability enhances the model's ability to generalize across viewing conditions, reduces overfitting by exposing the network to a wider range of input variations, improves resilience to the variable conditions encountered during surveys, and simulates natural perspective variations.

The effectiveness of our augmentation strategy is particularly relevant given the unique challenges of underwater images, where viewing angles and positions can vary significantly due to water conditions, currents, and operational constraints. By incorporating these variations into the training process, the model develops robust feature recognition capabilities that are more applicable to real-world applications.

7.4. Binary classification results

7.4.1. Model performance analysis

The optimization framework follows an iterative gradient descent pattern, as evidenced by the learning curves shown in Figure 7.7. The evolution of both the training and validation losses shows distinct patterns in the learning dynamics of the model. Initially, both curves start with high loss values, with the training loss starting at approximately 0.47 and the validation loss starting at 0.37. The first 20 epochs show a steep decline in both curves, indicating a period of rapid learning where the model quickly improves its performance on both training and validation data sets. After this initial phase, the learning trajectory enters a more gradual convergence pattern. The training loss shows a relatively smooth and consistent decrease throughout the remaining epochs, eventually converging to approximately 0.07. In contrast, the validation loss shows more significant variations, while maintaining a general downward trend until stabilizing around 0.15. The divergence between the training and validation losses is particularly noticeable after epoch 40, where the training loss continues its gradual decline while the validation loss exhibits more pronounced oscillations. Despite these fluctuations, the overall trend indicates successful model convergence, with both metrics showing significant improvement from their initial values. This pattern of behaviour is typical of DL models and provides valuable insight into the model's generalization capabilities and learning efficiency. The recursive optimization process continues until the model achieves convergence, effectively identifying the parameter configuration that minimizes classification error while maintaining its ability to generalize.

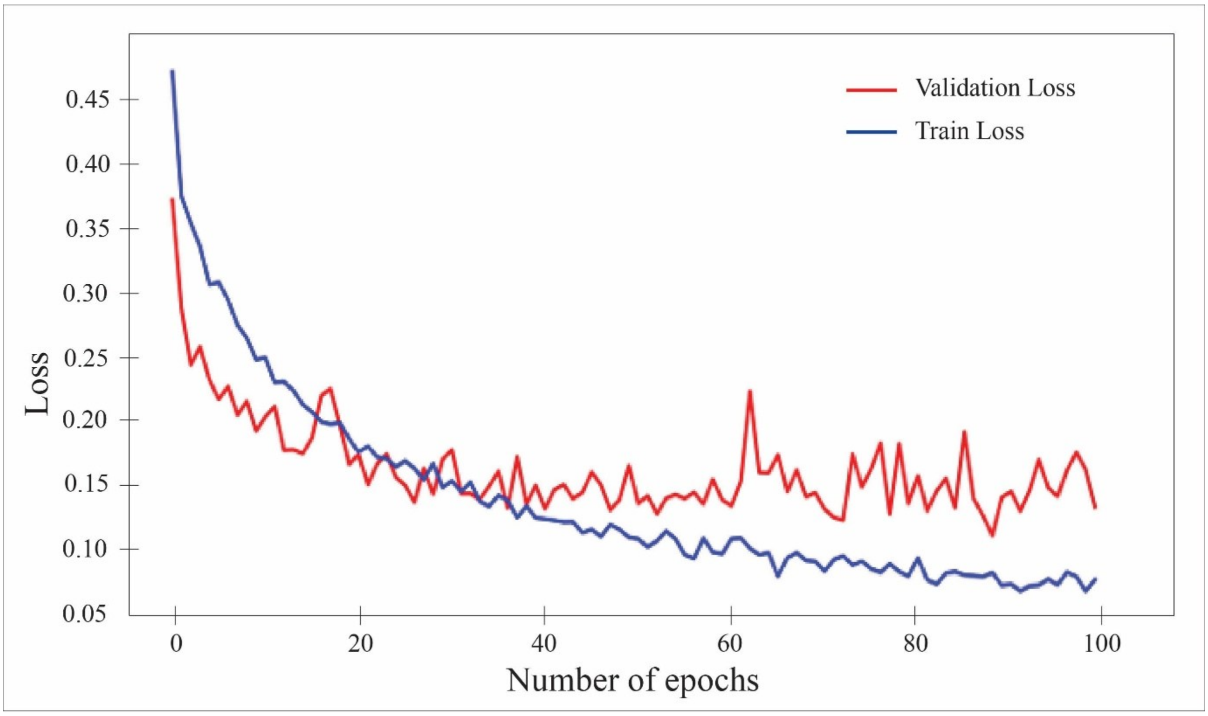

Figure 7.7: Learning curve of the binary classification network model, showing both training loss and validation loss over 100 iterations or epochs of training.

The CNN is evaluated on accuracy, that corresponds to the fraction of images correctly classified by the model, and F1-score, a key metric in machine learning that measures a model's accuracy by combining precision and recall into a single value. In addition, the confusion matrix provides a quick quantification of the number of false positives and negatives. All these quantities were calculated on the validation split of the dataset.

First, the confusion matrix is introduced, a 2x2 table which is useful to visualize and summarize the performances of a classification algorithm (Figure 7.8). In the case of the binary classification, where the classes are 'useful' and 'not-useful', respectively indicated as positive or negative, the elements of the table are: True Negative (TN), the number of cases predicted by the model as "useful", which are correctly classified (3854 images); False Negative (FN), the number of cases predicted by the model as 'useful', which are wrongly classified (46 images); True Positive (TP), the number of cases predicted by the model as 'Not- useful', which are correctly classified (1281 images); False Positives (FP), the number of cases predicted by the model as 'Not-useful', which are wrongly classified (14). Confusion matrix elements are the basis for the metrics evaluated (see Table 7.1).

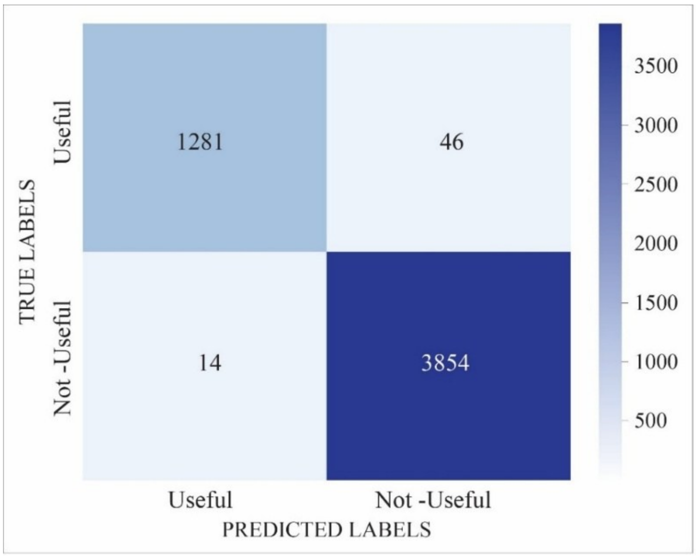

Figure 7.8: Confusion matrix in the binary classification.

Table 7.1: Confusion matrix in binary classification.

	Prediction	
Ground truth	Useful	Not useful
Useful	1281	46
Not useful	46	3854

By classification accuracy we mean the ratio of the number of corrected predictions to the total number of observations:

$$Accuracy = \frac{Number\ of\ corrected\ predictions}{Total\ number\ of\ predictions\ made} = \frac{TP + TN}{TP + TN + FP + FN}$$

A higher value of accuracy corresponds to an efficient model. The accuracy of our model is about 98.85% (Table 7.2).

F1-score combines two other metrics and can be defines as the harmonic mean of precision and recall, where recall is defined like:

$$Recall = \frac{TP}{TP + FN}$$

And precision is defined like:

$$Precision = \frac{TP}{TP + FP}$$

Our calculated values of recall and precision are, respectively, 0.9892 and 0.9653. Then, the value of F1-score could be calculated like:

$$F1 = \frac{2 \cdot (Recall\ \cdot\ recision)}{Recall + Precision}$$

The calculated vale of F1-score is 97.71% (Tabel 7.2).

Table 7.2: Metric evaluation in binary classification.

Metric	Value (%)
Accuracy	98.85
F1-score	97.71

The model achieved a remarkable classification accuracy of 98.85% on the test set after completing the training regime. Our evaluation protocol used a standard 80-10 split, where 80% of the dataset was used for training and 10% for testing, randomly selected to allow for unbiased evaluation. These test examples are therefore not seen by the network and can be used to evaluate it. The training phase demonstrated efficient convergence, although our observations suggest that comparable performance could be achieved with fewer epochs.

The test results indicate the model's robust ability to discriminate between acceptable and unusable photos (Figure 7.9). However, certain edge cases have been identified, particularly images containing human subjects or dominant non-geological objects, where the model's classification confidence was lower. These cases highlight potential areas for improvement in the model's feature extraction capabilities.

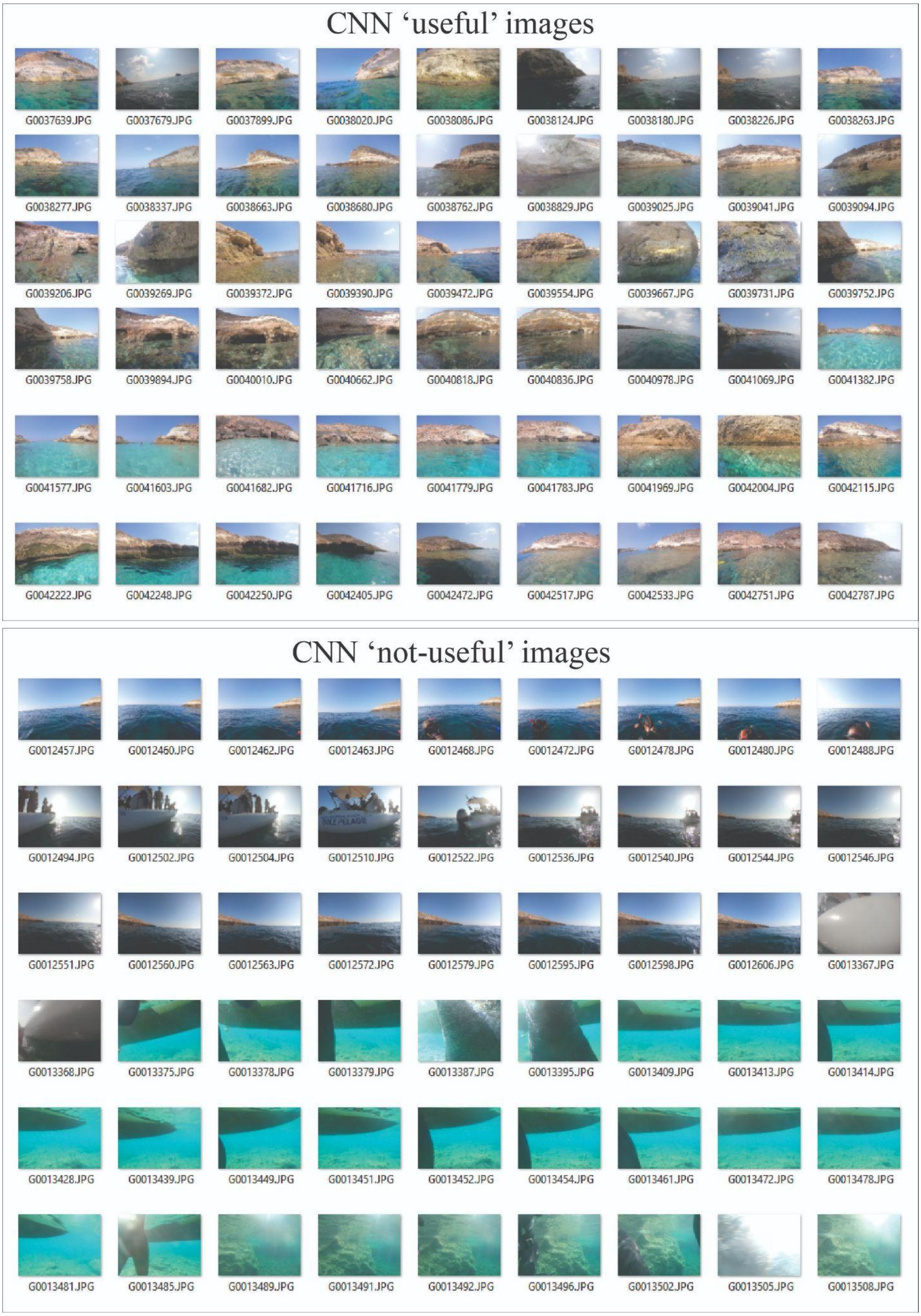

Figure 7.9: Photos classified by CNN using the binary classification with 98.85% of accuracy.

7.4.2. GradCAM and GradCAM ++ results

The analysis of the predicted area and the Ground Truth area provided significant insights through both quantitative metrics and qualitative examination of the generated heatmaps (Figure 7.10).

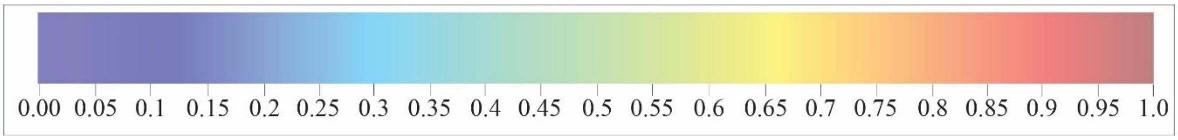

Figure 7.10: Visualization of the correlation between value and colour used by GradCAM to generate the heatmap.

To validate the model's learning process and interpretation capabilities, a comprehensive evaluation was performed using eight carefully selected test images. Each image was analyzed using both the GradCAM and GradCAM++ visualization techniques. The model's classification system assigns a suffix to each image filename: "_1" indicates "not useful" and "_0" indicates "useful" images (e.g., G0012345_1).

The visualization analysis showed that both GradCAM and GradCAM++ are successful in identifying critical regions that influence the model's classification decisions.

Figure 7.11 shows all the photos used from our test set, consisting of three views: the original image, the GradCAM visualization, and the GradCAM++ result. In this visualization set, the main patterns in the eight test images and their respective heatmap generations can be observed. Let's systematically analyze these results.

For the images classified as not useful ('_1'), the first image (G0016034) shows a diver in underwater conditions. Both GradCAM and GradCAM++ emphasize the human figure and the diving equipment, correctly identifying them as elements that make the image unusable for the dataset. GradCAM++ provides a more focused activation on the diver's body and equipment. Image G0016012 shows part of a boat and possibly a person. Both methods highlight the human presence and the structure of the boat, but GradCAM++ shows additional activation points, indicating that it detected more potentially problematic features that led to the "not useful" classification. G0016187 shows surface water scenes with various interfering elements. The heatmaps show that both methods identified problematic areas, with GradCAM++ showing a more precise localization of these features.

In image G0013657, which shows an underwater scene with some geological formations, there is indeed a noticeable difference between the GradCAM and GradCAM++ visualizations. GradCAM shows more scattered activation points (red-yellow hotspots) distributed across the image, with small areas of high focus scattered throughout the scene. This scattered pattern may indicate that the model is responding to multiple features or potential perturbations in the water or rock formations. In contrast, GradCAM++ produces a more focused and consolidated activation map, with a primary area of strong activation in the central region of the image. This suggests that GradCAM++ is more selective in identifying the key features that led to the image being classified as not useful ('_1'). For images classified as useful ('_0'), images G0010688, G0012181, and G0012653 show underwater rock formations and seafloor features. In these cases, both GradCAM and GradCAM++ highlight geological features and water clarity characteristics that make these images suitable for the dataset. GradCAM++ tends to provide more concentrated hotspots around specific geological features, while GradCAM shows broader activation patterns.

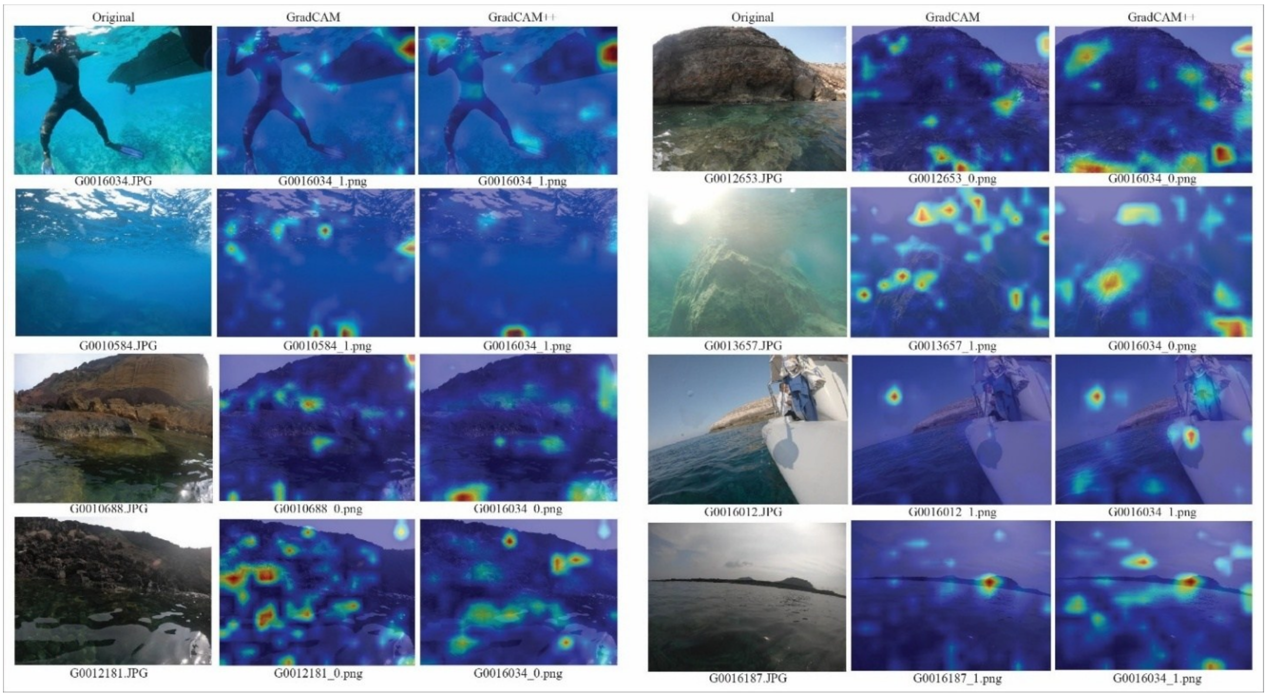

Figure 7.11: Visualization of the original picture and the final heatmap obtain using GradCAM and GradCAM++ techniques in binary classification.

7.5. Multi-class classification results

7.5.1. Model Performance analysis

The training loss (blue line in Figure 7.12) shows a consistent downward trend throughout the training process. It begins at approximately 1.2 and decreases rapidly during the first 100 iterations, followed by a slower, more gradual decrease. After 200 iterations, the training loss stabilizes between 0.1 and 0.2, suggesting that the model is effectively learning the patterns present in the training data.

The validation loss (red line in Figure 7.12) exhibits a different pattern than the training loss. Initially, it decreases alongside the training loss (during approximately the first 20 iterations). However, it then stabilizes at a higher value (between 0.6 and 0.8) while showing significant fluctuations throughout the training process. Due to these fluctuations in the validation loss and the growing gap between the training and validation loss curves, it appears that overfitting is occurring. This phenomenon, along with its underlying causes, will be discussed in detail in the discussion section (see Section 8.3.2).

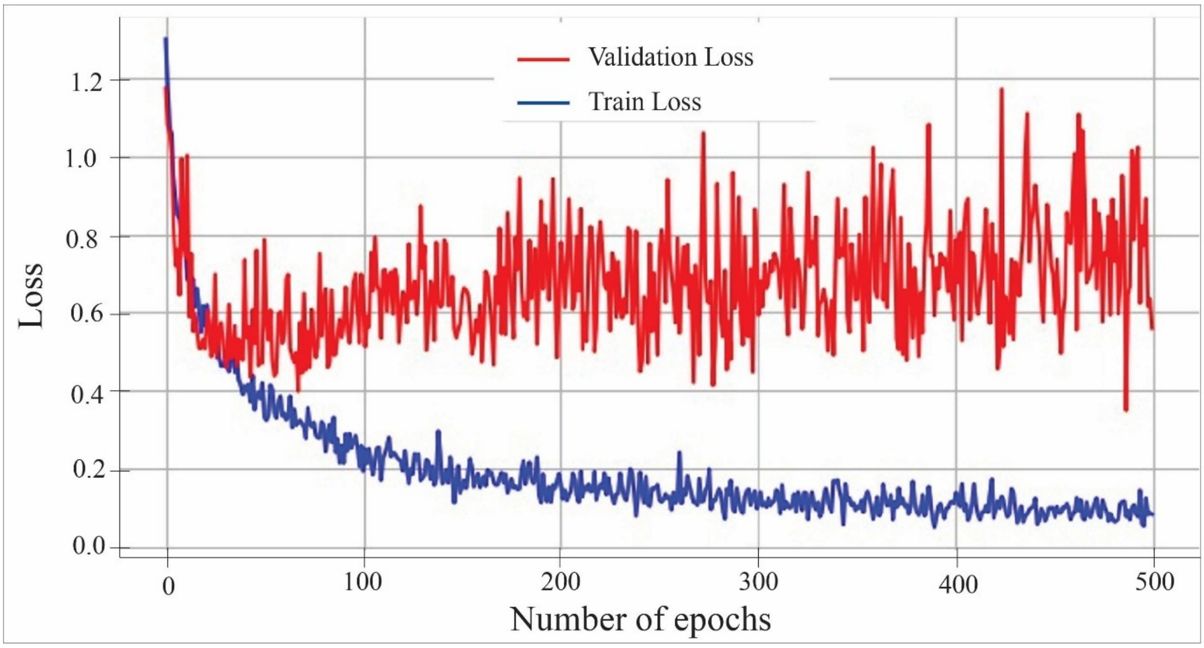

Figure 7.12: Learning curve of the multi-class classification network model, showing both training loss and validation loss over 500 epochs of training.

The model was then evaluated in terms of accuracy and F1 score, as well as through a case study related to the performance of the binary classification model. In addition, the confusion matrix was included to quantify the number of false positives and false negatives effectively. All these metrics were calculated based on the validation split of the dataset.

First, we introduce the confusion matrix, which is presented as a 5x5 table. This matrix is a valuable tool for visualizing and summarizing the performance of our classification algorithm (Figure 7.13). The confusion matrix illustrates the classification performance across five distinct coastal geomorphic categories: 'plunging cliff' (_0), 'pocket beach' (_1), 'scree' (_2), 'shore platform' (_3), and 'sloping coast' (_4).

For each class, we can analyze the true positives (TP), false positives (FP), and false negatives (FN) to derive related metrics (Table 7.3).

From the analyzed classes, 55 images are correctly identified (55 TP) in the plunging cliffs category. However, in the same category, 9 images are misclassified (9 FP of which 2 are classified as scree and 7 are sloping coast). Additionally, a total of 15 images of plunging cliffs were misclassified into other categories (15 FN of which 12 are classified as scree and 3 as sloping coast).

For the scree category, 34 images are correctly identified (34 TP). However, 21 images are misclassified (21 FP, of which 12 are classified as plunging cliffs, 1 as pocket beach, 4 as shore platform, and 4 as sloping coast). Additionally, 7 images of scree are misclassified into other categories (7 FN, of which 2 are classified as plunging cliffs, 1 as shore platform, and 4 as sloping coast). In the shore platform category, 17 images are correctly identified (17 TP). However, 1 image is misclassified (1 FP, classified as scree). Additionally, 5 images of shore platform are misclassified into other categories (5 FN, of which 4 are classified as scree and 1 as sloping coast). The sloping coast category shows strong performance with 54 images correctly identified (54 TP). However, 8 images are misclassified (8 FP, of which 3 are classified as plunging cliffs, 4 as scree, and 1 as shore platform). Additionally, 11 images of sloping coast are misclassified into other categories (11 FN, of which 7 are classified as plunging cliffs and 4 as scree).

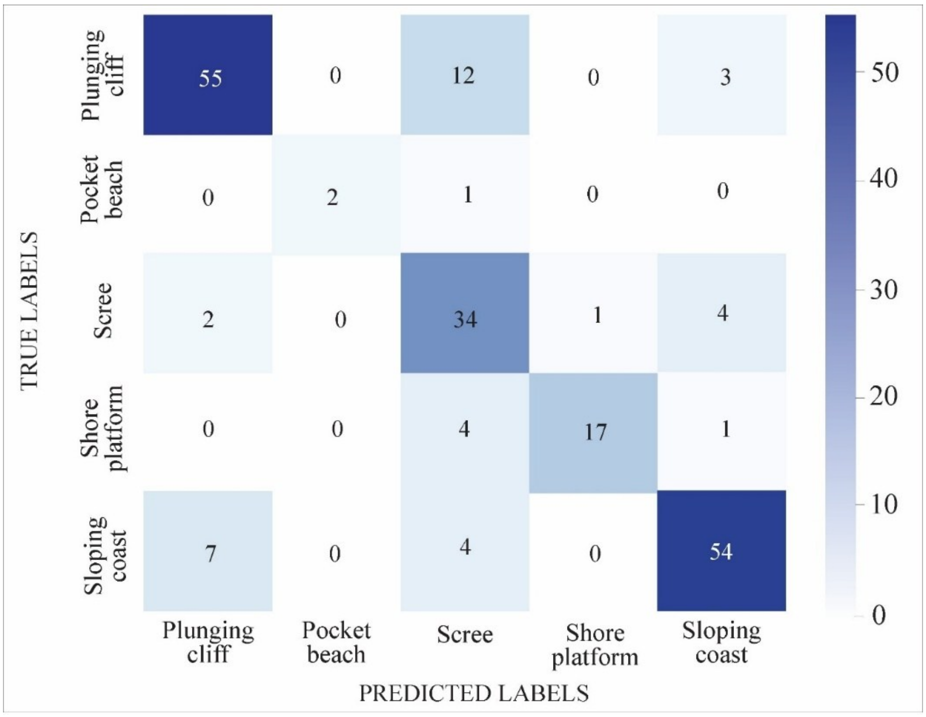

Figure 7.13: Confusion matrix in multi-class classification.

Table 7.3: Confusion matrix in multi-class geomorphological classification.

	Prediction				
Ground truth	Plunging cliff_0	Pocket beach_1	Scree_2	Shore platform_3	Sloping coast_4
Plunging cliff_0	55	0	12	0	3
Pocket beach_1	0	2	1	0	0
Scree_2	2	0	34	1	4
Shore platform_3	0	0	4	17	1
Sloping coast_4	7	0	4	0	54

By classification accuracy, this refers to the ratio of the number of correct predictions to the total number of observations. The formula for accuracy can be expressed as:

$$Accuracy = \frac{\sum TP}{Total\ samples}$$

Confusion matrix analysis revealed a strong overall performance, with a balanced accuracy of 80.6%. This demonstrates the ability of the model to effectively discriminate between five different coastal morphological classes (Table 7.4).

The F1-score in multi-class classification is calculated step by step for each class based on the confusion matrix and is defined as the harmonic mean of precision and recall. It provides insight into different levels of classification performance (see Section 7.4.1).

The model showed particularly strong performance in identifying plunging cliffs, achieving an F1 score of 0.821 (precision = 0.859, recall = 0.786). This robust performance can be attributed to the distinctive morphological characteristics of plunging cliffs and the adequate representation of this class in the training dataset. Similarly, sloping coast classifications showed excellent reliability with an F1 score of 0.851 (precision = 0.871, recall = 0.831), indicating the model's strong ability to identify this type of coastal formation.

The classification of land platforms also showed an F1 score of 0.850 (precision = 0.944, recall = 0.773). The remarkably high precision indicates that when the model identifies a shore platform, it does so with high confidence and accuracy. The slightly lower recall suggests that some instances of shore platforms were wrongly classified into other categories, particularly as scree formations.

The pocket beach category, despite limited representation in the dataset, achieved a respectable F1 score of 0.800 (precision = 1.000, recall = 0.667). The precision score indicates that the model made no false positive predictions for this class, although the lower recall suggests that some pocket beaches were missed and classified as other formations.

The most challenging category proved to be scree formations, with an F1 score of 0.708 (precision = 0.618, recall = 0.829). The lower precision indicates that the model incorrectly classified other coastal formations as scree, particularly in cases where morphological features showed similar characteristics. However, the higher recall suggests that when scree formations were present, the model was generally successful in identifying them.

The macro-average F1 score of 80.6% across all classes indicates robust overall performance, particularly when considering the inherent difficulties of multi-class coastal geomorphological classification. These problems include varying lighting conditions, perspective differences in images, and the natural gradation between different coastal formation types.

Table 7.4: Metric evaluation in multi-class geomorphological classification

Metric	Value (%)
Accuracy	80.6
F1-score	80.6

7.5.2. GradCAM and GradCAM ++ results

The visualization and interpretation of model behaviour using GradCAM and GradCAM++ techniques provide valuable insight into how the DL model identifies and distinguishes between different coastal landforms. Following the same visualization approach used in the binary classification (Figure 7.10), the consistency in the interpretation of the heatmap colour scale is maintained, where warmer colours (red and yellow) indicate regions of higher importance to the model's decision process, while cooler colours (blue) represent areas of lower importance. This standardized visualization approach allows direct comparisons between different landform classes and helps validate the model's learning process.

For each coastal landform class, plunging cliff, pocket beach, scree, shore platform, and sloping coast, we analyze how the model interprets and distinguishes the characteristic geomorphological features. The following analysis examines the activation patterns for each class, revealing the specific features and regions that the model considers most important for classification.

Starting with the class of plunging cliffs (_0), the GradCAM and GradCAM++ visualizations reveal significant insights into the model's learning process (Figure 7.14). The original images consistently show steep, near-vertical cliff faces dropping directly into relatively deep water, which is characteristic of plunging cliffs. The water typically appears green or turquoise to indicate depth, and the cliffs have minimal to no visible platform at water level.

The GradCAM visualizations (middle column) show that the model focuses primarily on the intersection between the cliff face and the water, as indicated by the concentrated heat maps in these areas. This suggests that the model has learned that this rapid transition from cliff to deep water is a significant diagnostic feature of plunging cliffs.

The GradCAM++ results (right column) provide additional insight, showing more refined and specific activation patterns. These visualizations typically show stronger and more focus to the cliff-water interface, with additional focus on the upper portions of the cliff face. This suggests that the model also considers the overall verticality and structure of the cliff face in its classification process.

Notably, in images such as G0014059.JPG and G0014964.JPG, both GradCAM and GradCAM++ show strong activations at the base of the cliff where it meets the water, confirming the model's focus to this feature.

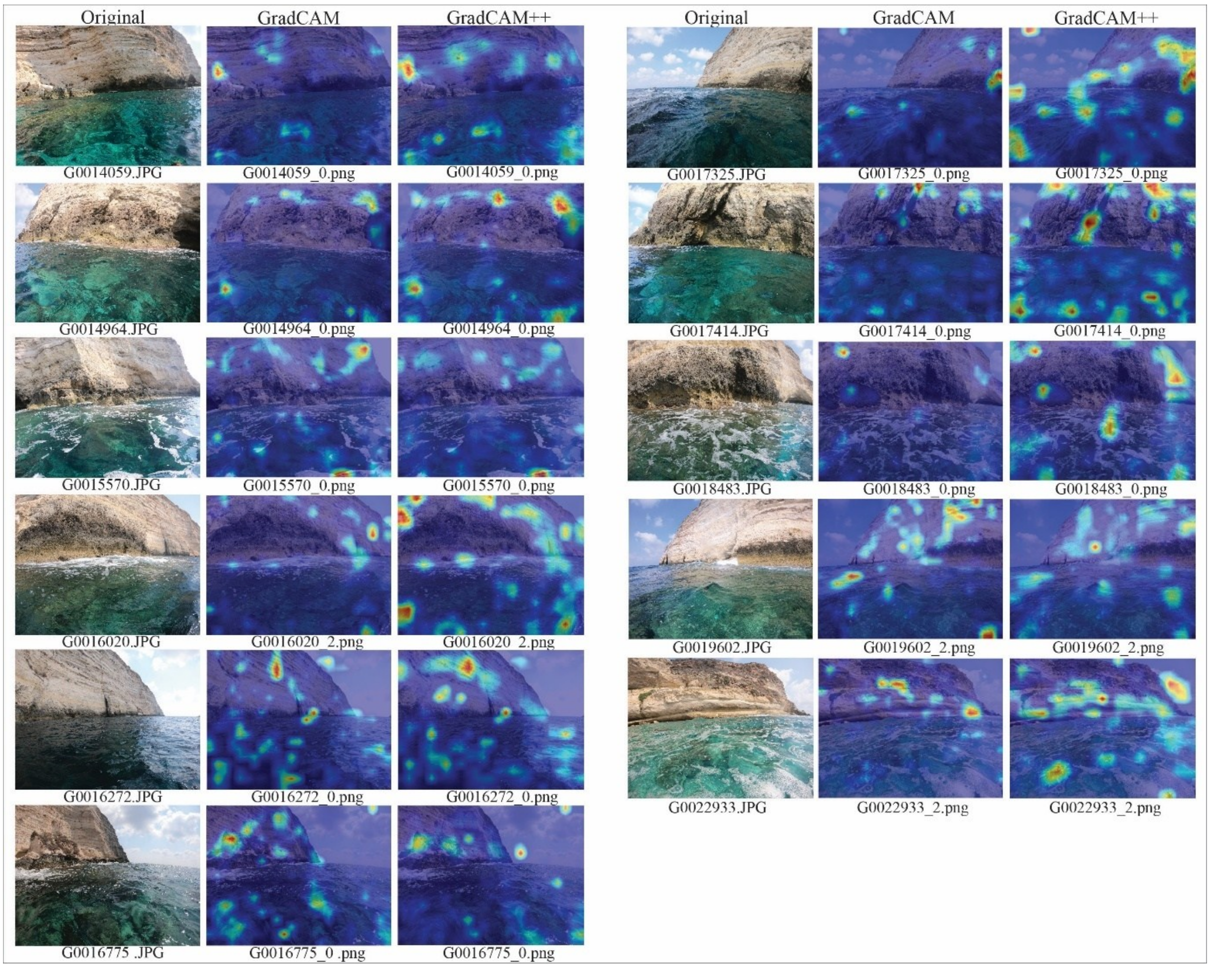

Figure 7.14: Visualization of the original picture and the final heatmap obtain using GradCAM and GradCAM++ techniques in plunging cliff classification.

Looking at the GradCAM and GradCAM++ visualizations for pocket beaches (marked with _1), it can be observed how the model identifies the distinctive features of this coastal landform (Figure 7.15). The original images show the typical morphology of pocket beaches, characterized by small, enclosed beach systems bounded by headlands on either side.

The GradCAM visualizations show that the model pays particular focus to the lateral boundaries of the pocket beaches where the confining headlands meet the water. This is evidenced by the concentrated activation patterns at these transition points, as clearly seen in the images G0020682.JPG and G0021270.JPG. The model appears to recognize that these lateral confinements are specific features that distinguish pocket beaches from other coastal landforms.

GradCAM++ results provide more detailed views, showing increased focus to both headland boundaries and the beach-water interface. In images such as G0021423.JPG, activation patterns are particularly strong at the curved intersection of the beach with the water, suggesting that the model has learned to identify the characteristic shape of pocket beaches. The improved resolution of GradCAM++ also shows that the model considers the overall geometric configuration of the landform, with focus distributed over the confining promontory and the central beach area.

Interestingly, in images where the beach material is visible (such as G0022611.JPG), both visualization techniques show activation patterns that include the beach sediment area, indicating that the model also considers the presence of accumulated sediment as a defining feature of pocket beaches.

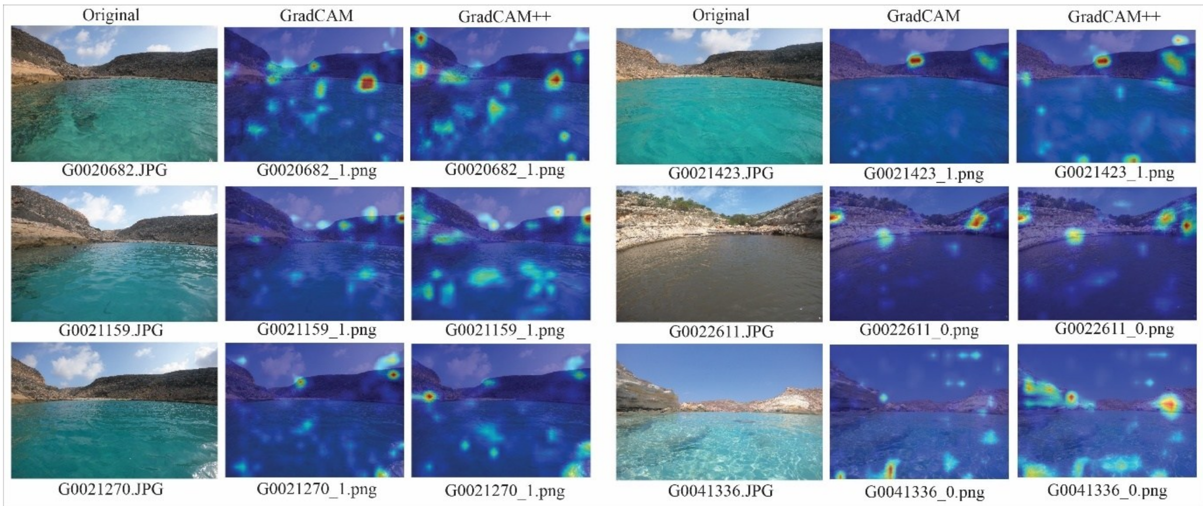

Figure 7.15: Visualization of the original picture and the final heatmap obtain using GradCAM and GradCAM++ techniques in pocket beach classification.

Analyzing the GradCAM and GradCAM++ visualizations for the scree class (marked with _2), it can be observed how the model identifies these distinctive accumulations of loose rock fragments at the base of cliffs (Figure 7.16). The original images show the characteristic features of coastal scree: accumulated rock debris at the base of the cliff, creating a distinctive transition zone between the cliff face and the water.

The GradCAM visualizations show that the model focuses heavily on the interface where the debris meets the water, as evidenced by the concentrated heat maps in these areas. This is particularly evident in the images G0020664.JPG and G0021253JPG, where the activation patterns highlight the irregular, rough texture characteristic of boulder deposits.

The GradCAM++ results provide more detailed activation patterns, showing an increased focus on the boulder deposits and their relationship to the cliff face. In images such as G0020276.JPG and G0021997.JPG, the activation patterns extend up the cliff face above the boulders. This more refined visualization technique shows that the model has learned to recognize not only the presence of loose material, but also its contextual relationship within the broader coastal system. Both visualization techniques show strong activations at points where the boulder material creates irregular protrusions into the water, as seen in G0023804.JPG and G0024174.JPG.

The visualizations show that the model considers the landform's overall morphological configuration, focusing on both the accumulated material and its relationship to adjacent features.

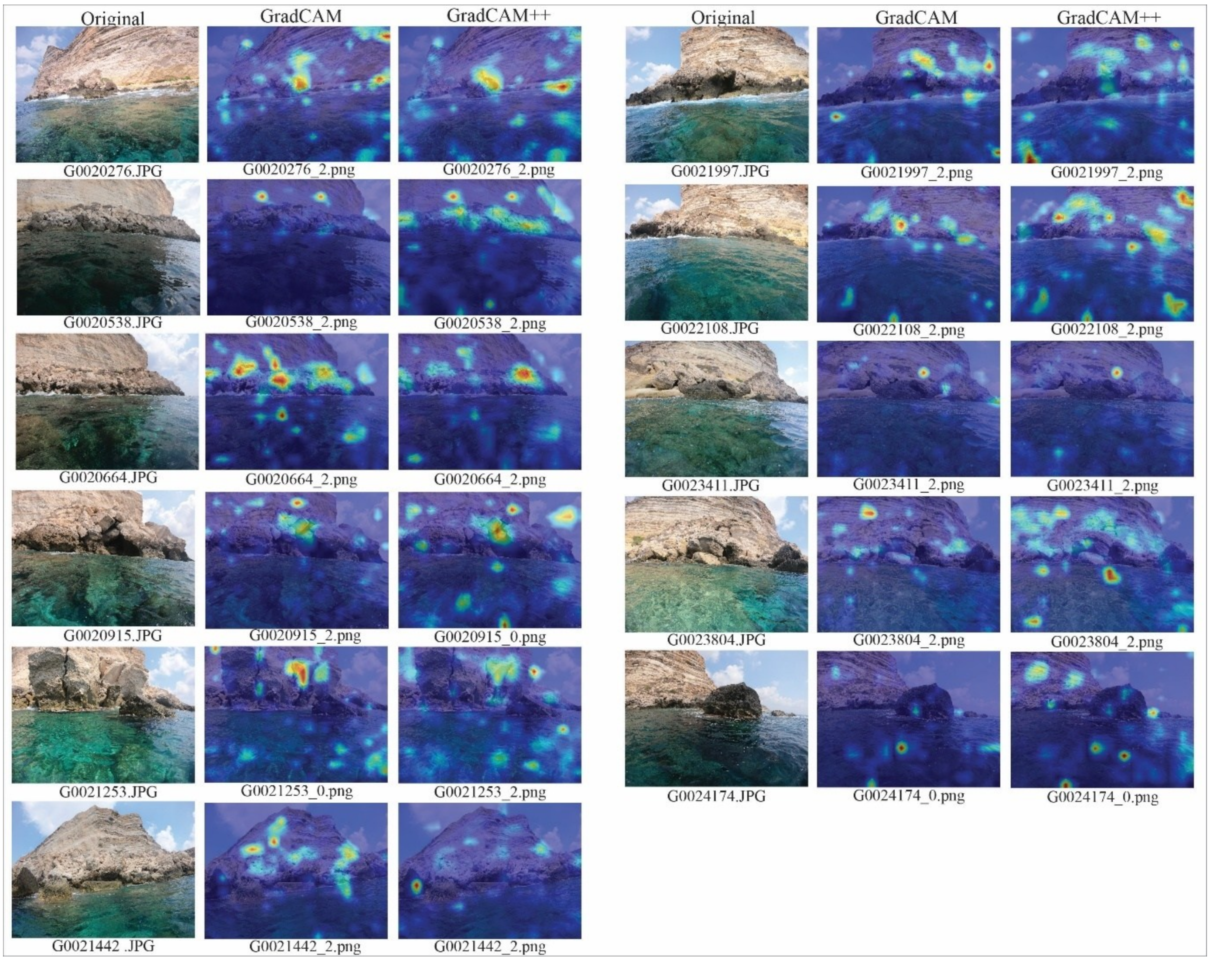

Figure 7.16: Visualization of the original picture and the final heatmap obtain using GradCAM and GradCAM++ techniques in scree classification.

Analyzing the GradCAM and GradCAM++ visualizations for the shore platform class (marked _3), it can be observed how the model identifies these distinctive horizontal or sub-horizontal rock surfaces at or near sea level (Figure 7.17).

The GradCAM visualizations show that the model focuses intensely on the horizontal extent of the platform surface, especially where it intersects with deeper water at its seaward edge. This is clearly seen in the images G0041731.JPG and G0042884.JPG, where the activation patterns highlight the transition zone between the platform and deeper water.

The GradCAM++ results provide a more detailed view, showing an increased focus on the platform surface and its relationship to the underlying cliff. In images such as G0024732.JPG, the activation patterns across the platform surface are particularly strong, suggesting that the model has learned to recognize the flat or gently sloping morphology characteristic of shore platforms and that distinguishes them from other coastal landforms.

Both visualization techniques show strong activations at the seaward edge of the platforms, as seen in G0040417.JPG and G0041731.JPG.

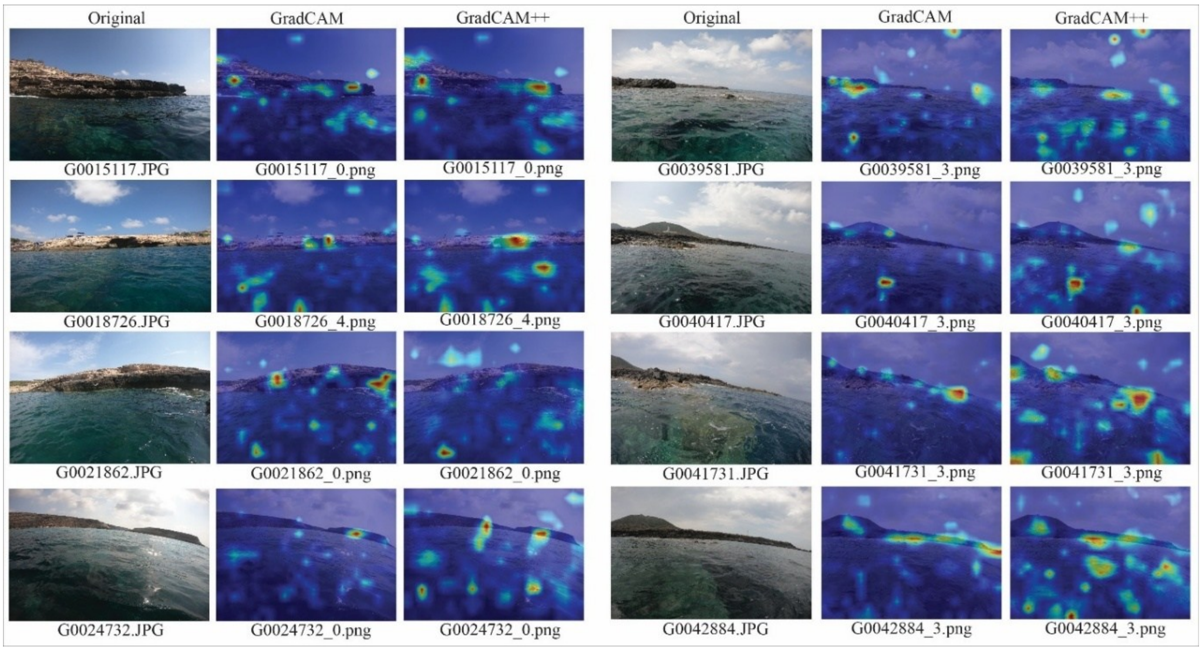

Figure 7.17: Visualization of the original picture and the final heatmap obtain using GradCAM and GradCAM++ techniques in shore platform classification.

Looking at the GradCAM and GradCAM++ visualizations for the sloping coast class (marked with _4), the original images show the characteristic features of sloping coasts: moderately sloped rock surfaces that gradually descend into the water, creating a softer transition from land to sea compared to plunging cliffs (Figure 7.18). This is particularly evident in the images G0016615.JPG and G0023049.JPG, where the coastal profile shows a consistent, moderate angle of slope.

The GradCAM visualizations show that the model focuses strongly on the overall slope gradient, especially at the water interface. This is clearly seen in the images G0022550.JPG and G0051169.JPG, where the activation patterns highlight the continuous sloping surface from the upper parts of the coast down to the waterline.

The GradCAM++ results provide more refined insights, showing an increased focus on both the slope angle and its continuity. In images such as G0024148.JPG and G0051436.JPG, the activation patterns are distributed more evenly across the sloped surface. This indicates that the model focuses on the texture and structure of the rock surface along the slope.

Of note is that both visualization techniques show activation patterns that follow the angle of the slope, as shown in G0022937.JPG, where the heat maps align with the natural angle of the slope.

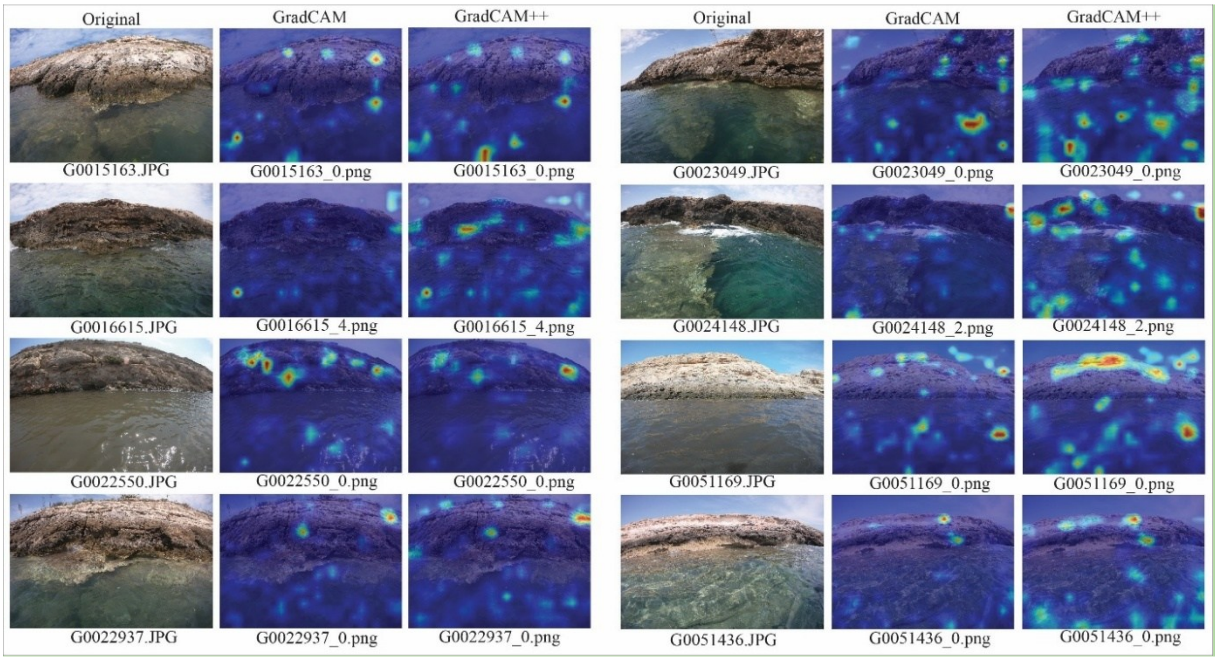

Figure 7.18: Visualization of the original picture and the final heatmap obtain using GradCAM and GradCAM++ techniques in sloping coast classification.

7.6. Building of integrated 3D models

The photogrammetric reconstruction for building integrated coastal 3D models typically follows a structured workflow that progresses through several distinct stages. The process begins with photo acquisition, followed by image alignment, where common points across multiple photos are identified and matched. This initial alignment produces a sparse point cloud that provides the basic spatial framework.

The workflow then progresses to dense cloud generation, where more detailed point data is extracted from the aligned images. This dense point cloud is used as the basis for mesh construction, where triangulation algorithms convert discrete points into a continuous surface representation. The final processing step is texture mapping, where photographic details are projected onto the geometric mesh to create a visually accurate representation. The entire 3D modeling process is described in detail in Section 5.6.

When modeling coastal environments, special consideration must be given to the interface between the above-water and below-water domains. The integration phase typically addresses the challenge of establishing common reference points and coordinate systems.

The photogrammetric workflow was performed using previously established underwater camera calibration parameters (Section 5.6.2) and adjusted camera position data. Mesh construction proceeded with the point cloud as the basic source data, maintaining high quality settings and face count parameters to ensure detailed surface reconstruction. Texture generation used diffuse mapping with the generic mapping mode enabled to optimize the visual representation of coastal features.

The final integration phase addressed the specific challenge of combining the emergent and submerged coastline models. This was accomplished using Metashape's "merge chunks" functionality, with special attention to depth maps and asset parameters. This approach ensured very good integration of the above and below waterline model components, while maintaining data integrity and spatial accuracy throughout the reconstruction process (Figure 7.19).

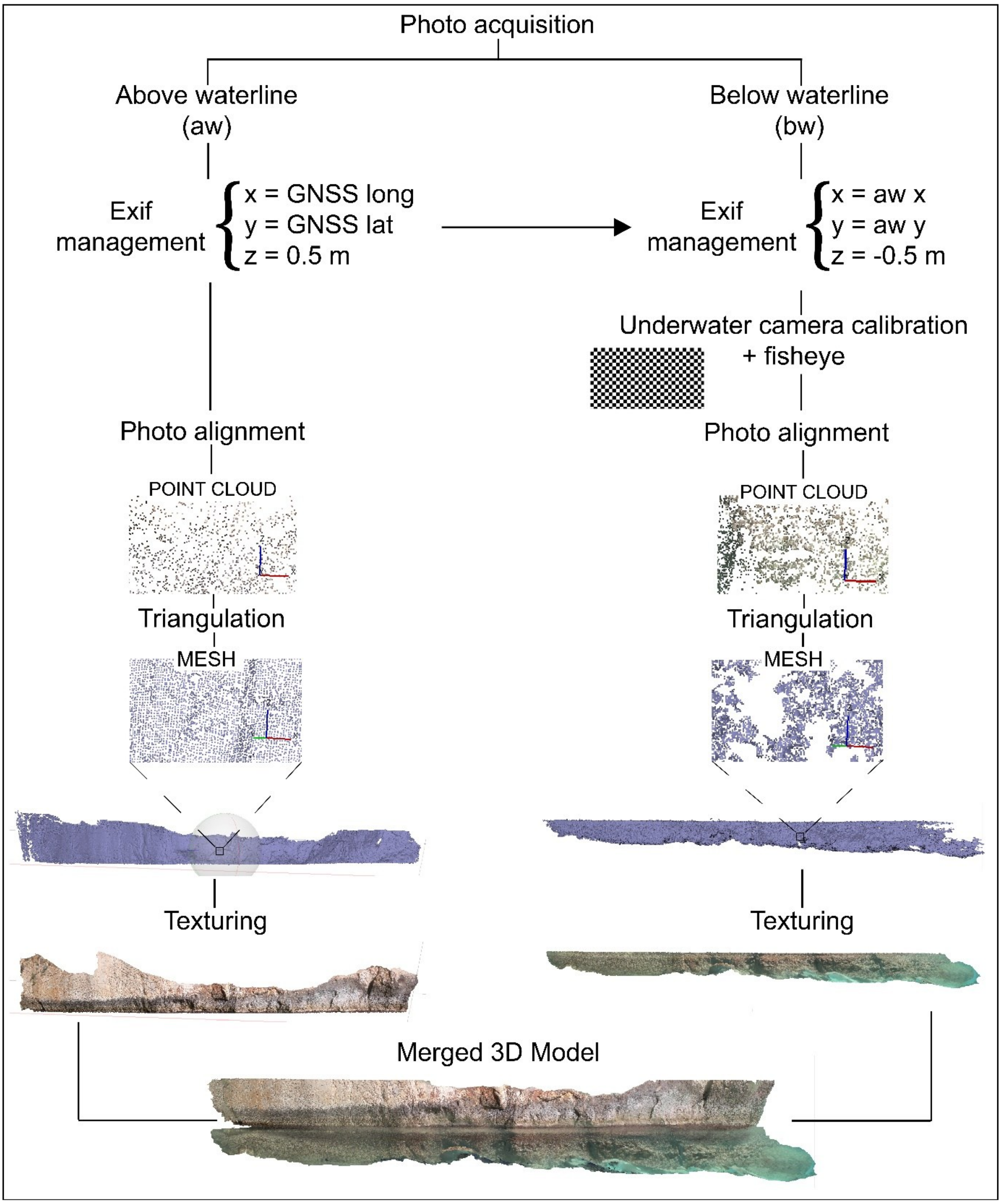

Figure 7.19: Integrated workflow for three-dimensional modeling of intertidal and nearshore environments. The workflow begins with parallel processing streams for above and below water images in Agisoft Metashape, using fixed camera for precise spatial referencing. Critical to the process is the initial alignment of the aw images, which provides the spatial framework for subsequent integration.

7.6.1. GNSS integration

The integration of above and below waterline 3D models requires precise management of spatial data beyond the underwater camera calibration process. This integration follows a precise order within the project workflow.

The process begins with the creation of two distinct blocks within a single project file, separating the aw and bw images. For the aw dataset, all photo heights are standardized to 0.5 meters prior to alignment, as this is the effective height at which the action camera is placed on the iron bar on the raft, perpendicular to the waterline, during the acquisition phase. The alignment process produces revised latitude and longitude coordinates for the aw dataset, which are then transferred to the bw dataset to compensate for the lack of underwater GNSS data. The underwater images are assigned an elevation of -0.5 meters, reflecting their approximate depth relative to mean sea level at the time the data were acquired. This methodological approach assumes relative consistency in elevation and depth during image acquisition, an assumption that is reasonable for short acquisition periods where tidal variations and wave effects are minimal. Systematic management of spatial data ensures accurate integration of the aw and bw model components, resulting in a representation of the shape of the coast within the intertidal zone.

7.7. 3D model

This paragraph presents comprehensive results from the implementation of the acquisition protocol and three-dimensional modeling process at three different study sites in the Mediterranean Sea. Each case study shows the application of the integration of 3D models under different conditions (i.e. weather, waves, wind, and type of coast - sloping coast or plunging cliff), providing insight into the effectiveness and limitations of the approach. The analysis includes a detailed examination of image acquisition, alignment parameters, and three-dimensional modeling processes for both above and below water components. Additionally, this study presents comparative analysis between satellite-derived images and three-dimensional model representations, with particular focus to landforms such as tidal notches (Furlani et al., 2024). Selected three-dimensional models were exported from the photogrammetric processing software in .obj format and subsequently imported into CloudCompare software (see Section 5.6.4) to enable comprehensive quantitative morphological analysis of the reconstructed coastal features.

7.7.1. Cala Uccello, Lampedusa (Italy)

The survey at Cala Uccello included two sectors, designated Zone 1 and Zone 2 (Figure 7.20), spanning 65 meters and 23 meters, respectively. The environmental conditions during the survey showed wave heights ranging from 0.5 to 1.5 meters and prevailing wind speeds of 4-5 m/s from the southwest (*Lampedusa Boa Oceanografica* - ENEA, 2024).

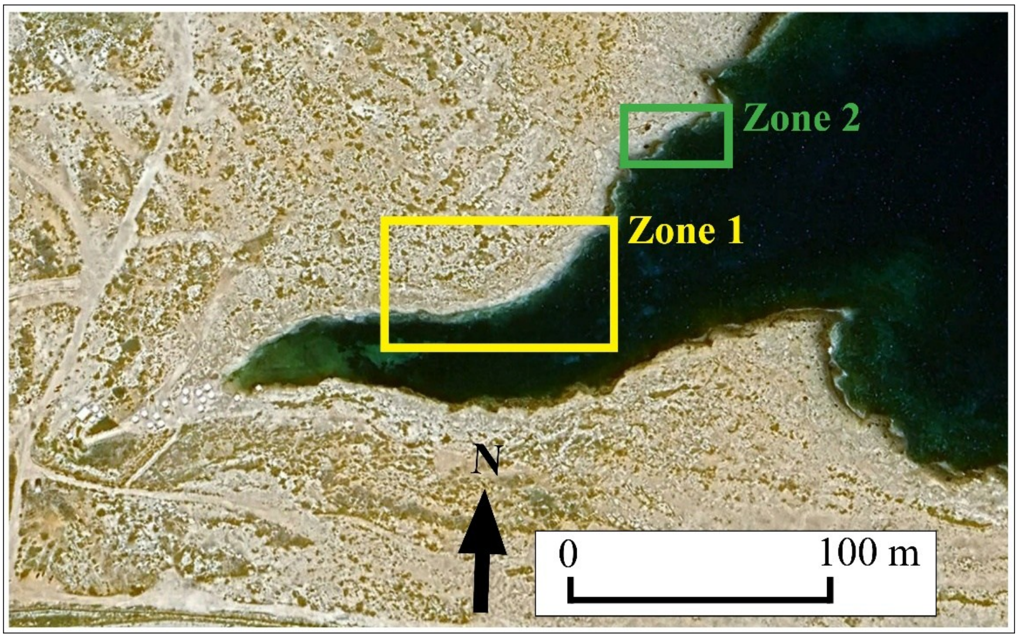

Figure 7.20: Spatial distribution of the two survey zones along the limestone coastline. Zone 1 extends for 65 meters, while Zone 2 encompasses 23 meters of coastline.

The post-processing phase of the data collected in the two areas produced two three-dimensional models with different characteristics (Table 7.5, Figure 7.21). The Zone 1 model (Figure 7.21a) used only GoPro images, with a total of 480 images: 240 aw and 240 bw. The alignment process successfully merged all 240 aw images, while 227 of the 240 underwater images were successfully aligned. The resulting merged model contained 287,000 tie points, generating a mesh of 4,422,584 faces at a resolution of 0.832 mm/pixel.

The Zone 2 model (Figures 7.21b, 7.21c) incorporated GoPro and 360-degree camera images. The aw component used 187 images (80 GoPro, 107 360-degree camera), while the bw component used only 80 GoPro images. The alignment process successfully integrated all of the aw images and 66 of the 80 bw images. The final merged model contained 280,877 tie points, resulting in a mesh of 3,718,300 faces at a resolution of 2.18 mm/pixel.

Table 7.5: Parameters of 3D model construction over two different survey areas, both aw and bw components and their merged outputs. 1) Survey area designation; 2) Cross-reference to associated figures; 3) 3D model dataset; 4) Total number of images acquired; 5) Images successfully aligned; 6) Tie points generated during alignment; 7) Total faces in final mesh construction.

(1) Study area	(2) Figure	(3) 3D model	(4) n° images	(5) n° aligned images	(6) Tie points (points)	(7) Mesh (faces)
Cala Uccello (Lampedusa)	4.5a	aw model (GoPro)	240	240	115 229	1 831 055
		bw model (GoPro)	240	227	171 765	7 863 055
		merged model (GoPro)	480	467	287 000	4 422 584
	4.5b, c	aw model 2 (GoPro+360)	187	187	226 318	1 569 546
		bw model 2 (GoPro)	80	66	54 559	667 353
		merged model 2 (GoPro+360)	267	253	280 877	3 718 300

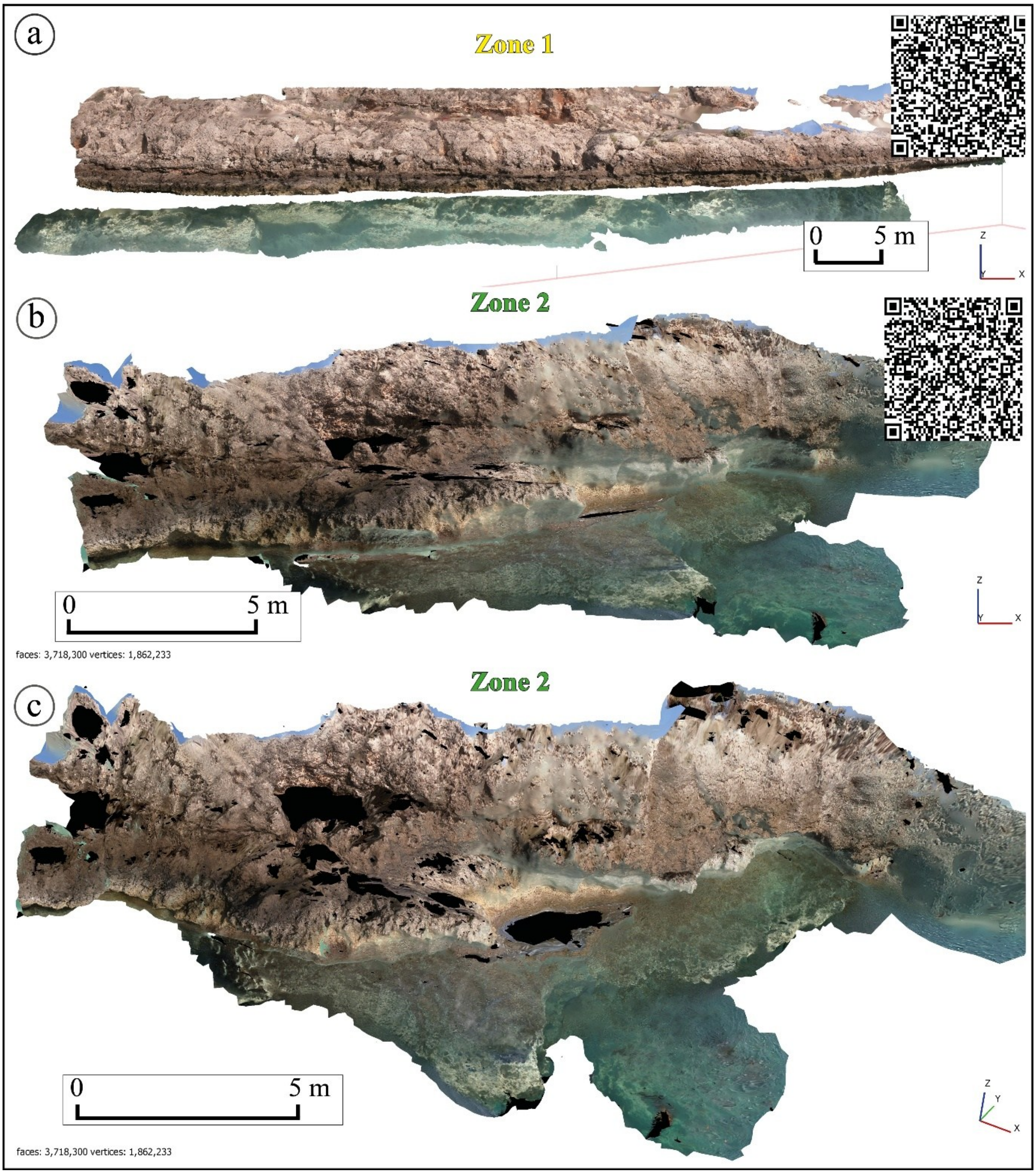

Figure 7.21: Multi-perspective visualization of the 3D model at the Cala Uccello rias site. a) Frontal view (xz plane) of the Zone 1 model of the intertidal and nearshore zone, above to below the waterline; b) Frontal view (xz plane) of the Zone 2 model; c) Elevated perspective view of the Zone 2 model in arbitrary xyz orientation. The QR-codes allow to view the corresponding 3D model on the sketchfab.com web platform.

From a qualitative and semi-quantitative point of view, it can be said that the Zone 1 model has a lack of data of about six meters along its profile, especially within the intertidal zone. Conversely, the Zone 2 model shows complete coverage of the intertidal zone when viewed orthographically (Figure 7.21b), although elevated perspective analysis reveals some areas without data (show as black areas in Figure 7.21c). Topographic profiles extracted using

CloudCompare software (Figure 7.22) show distinct tidal notch characteristics within the intertidal zone.

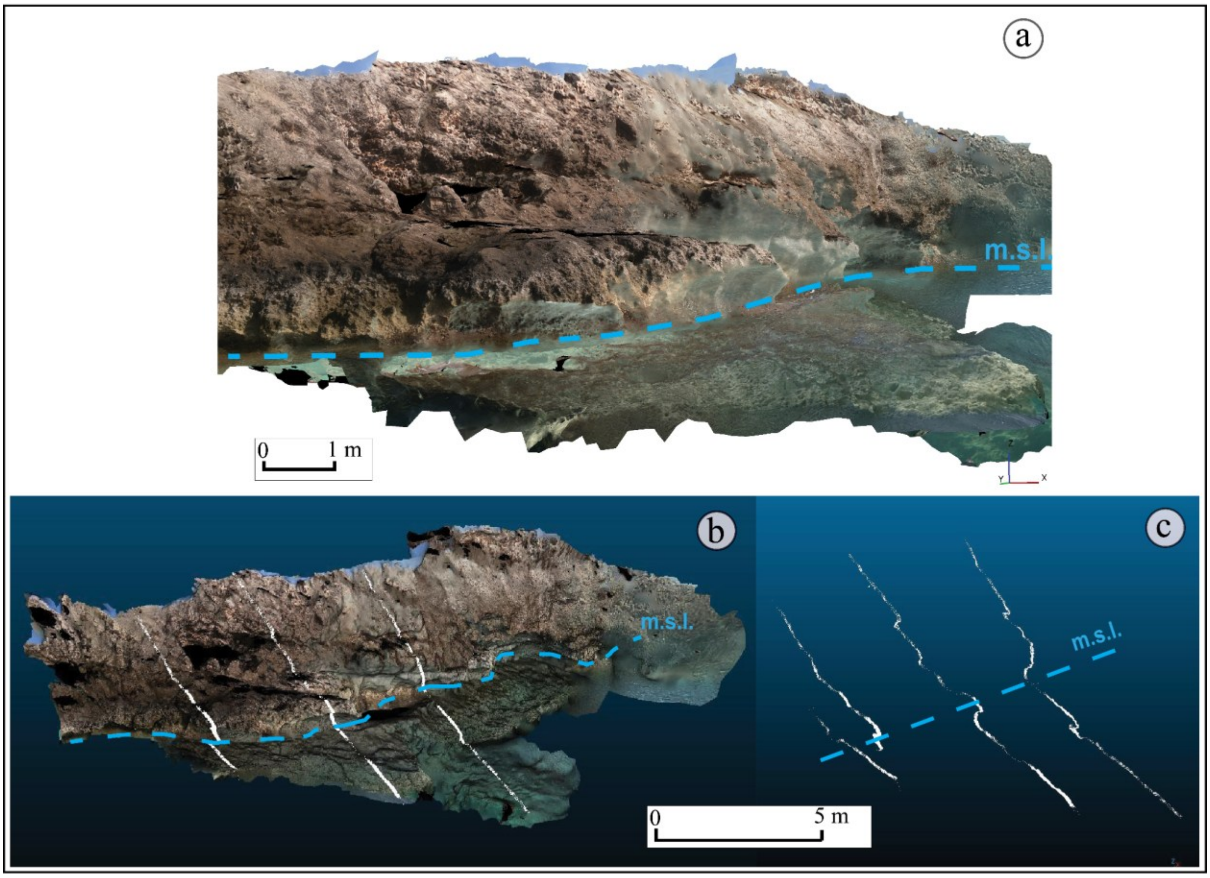

Figure 7.22: Mean sea level display on vertical profiles of three-dimensional model analysis of Zone 2: a) Three-dimensional model of the study area with profile extraction; b) and c) topographic profiles generated using Cloud Compare.

7.7.2. Mġarr ix-Xini, Gozo (Malta)

The survey focused on the eastern sector of the ria of Mġarr ix-Xini (Figure 7.23), which covers approximately 100 meters of coastline. Environmental conditions during the survey were characterized by minimal wave activity and moderate wind speeds averaging 2.1 m/s from the SSE (*Maltese Islands Weather*, 2023).

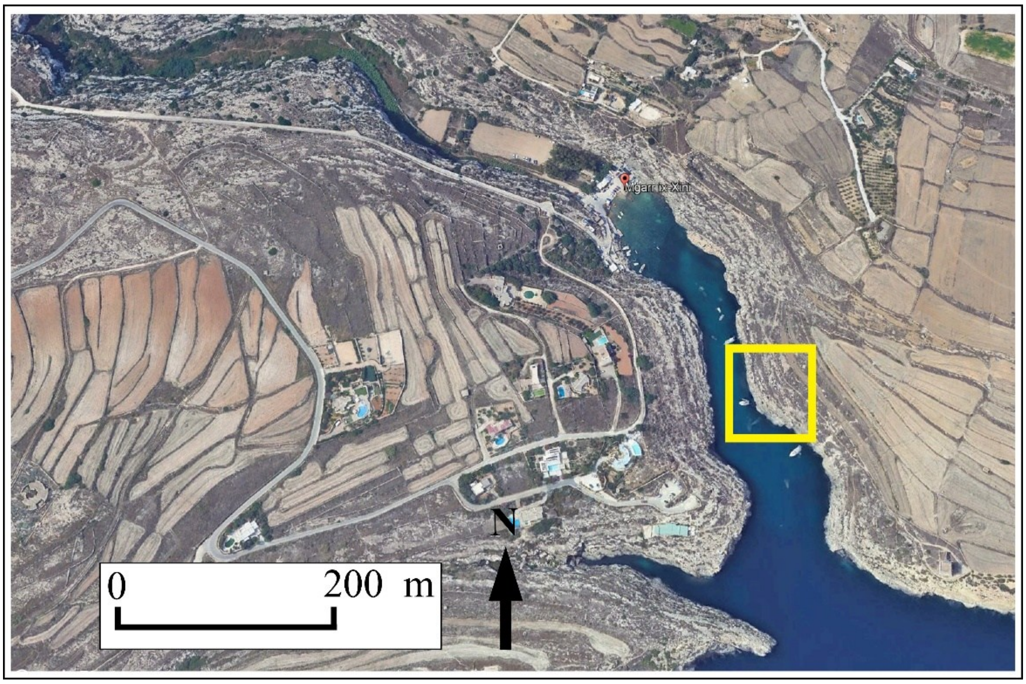

Figure 7.23: Detailed view of the surveyed coastal sector within the Mġarr ix-Xini ria. The study area encompasses a about 100 meter in length.

Two three-dimensional models were generated for the Mġarr ix-Xini ria, each using different image acquisition strategies (Table 7.6, Figure 7.24). The first modeling approach utilized only GoPro images (Figures 7.24a, 7.24b), incorporating a total of 1,068 images equally distributed between aw and bw positions (534 each). The alignment process successfully integrated 521 aw and all 534 bw images. The resulting merged model contained 626,358 tie points, generating a mesh of 12,500,233 faces with a resolution of 1.33 mm/pixel.

The second modeling iteration (Figures 7.24c, 7.24d) employed a more diverse data acquisition scheme. The aw component included 608 images, including 129 GoPro shots. The bw dataset combined the 129 GoPro images with 350 frames (1440x1080 pixels) extracted from the 360-degree camera video, for a total of 479 images. All images were successfully aligned in the final model. This integrated approach generated 620,951,000 tie points, resulting in a mesh of 5,115,318 faces at a resolution of 4.03 mm/pixel.

Table 7.6: Parameters of 3D model construction over distinct modeling approaches, detailing both component and merged model characteristics. 1) Survey area designation; 2) Cross-reference to associated figures; 3) 3D model dataset; 4) Total number of images acquired; 5) Images successfully aligned; 6) Tie points generated during alignment; 7) Total faces in final mesh construction.

(1) Study area	(2) Figure	(3) 3D model	(4) n° images	(5) n° aligned images	(6) Tie points (points)	(7) Mesh (faces)
Mġarr ix-Xini (Gozo)	7.24a, b	aw model (GoPro)	534	521	118 478	13 502 204
		bw model (GoPro)	534	534	507 880	9 714 486
		merged model (GoPro)	1 068	1 055	626 358	12 500 233
	7.24c, d	aw model (GoPro)	129	129	381 219	2 917 118
		bw model (GoPro+360)	479	479	321 758	2 328 373
		merged model (GoPro+360)	608	608	620 951	5 115 318

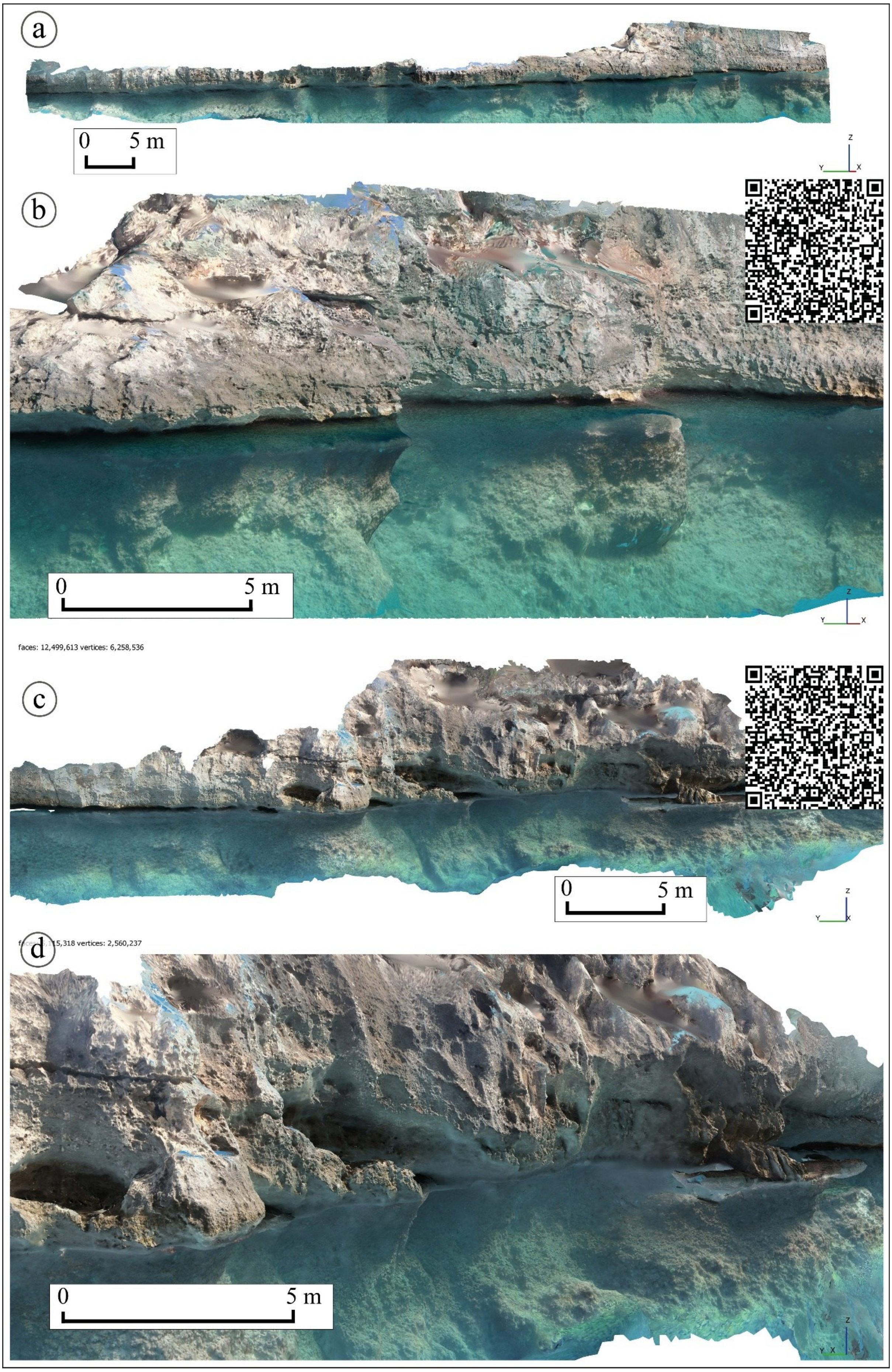
a
0 5 m
b
0 5 m
faces: 12,499,613 vertices: 6,258,536
c
0 5 m
d
0 5 m

Figure 7.24: Multi-prospective representation of the coast at Mġarr ix-Xini through two modeling approaches. a) Three-dimensional view (xyz orientation) of the first model constructed exclusively from GoPro images, demonstrating complete coastal coverage; b) Elevated perspective view of the first model, highlighting complex morphological features and spatial relationships; c) Frontal view (yz plane) of the second model, demonstrating integration of GoPro and 360-degree camera data; d) Alternative perspective (xz view) of the second model, emphasizing enhanced detail capture through multi-sensor integration. The QR-codes allow to view the corresponding 3D model on the sketchfab.com web platform.

From a quantitative point of view, both resulting models show complete coverage of the intertidal zone (Figure 7.24), although there are notable differences in texture rendering, particularly in the emergent areas where the characteristics of the underwater images influenced the colour rendering. The first model covers an extensive area of 326 m², with continuous coverage along 100 m of coastline. The vertical extent ranges from 1.40 meters above to 4.50 meters below (minimum depth -0.40 m).

The second model, integrating GoPro and 360-degree images, covers approximately 43 meters of coastline with a total area of 408 square meters. Vertical coverage extends from 3.50-4 meters above water to 4.40 meters bw (minimum depth -0.50 m). While this model has improved spatial coverage, the influence of underwater images on texture rendering results in a characteristic blue tint throughout the model.

The quantitative assessment of the tidal notch morphology was performed using a combination of photogrammetric modeling and distance analysis. After generating the initial 3D model in Metashape, the model of Figure 7.24a, b was exported in .obj format and imported into CloudCompare. The extracted topographic profiles (Figure 7.25) reveal detailed features of the morphology of the tidal notch within the intertidal zone. The profiles extraction is described in the Section 5.6.4.

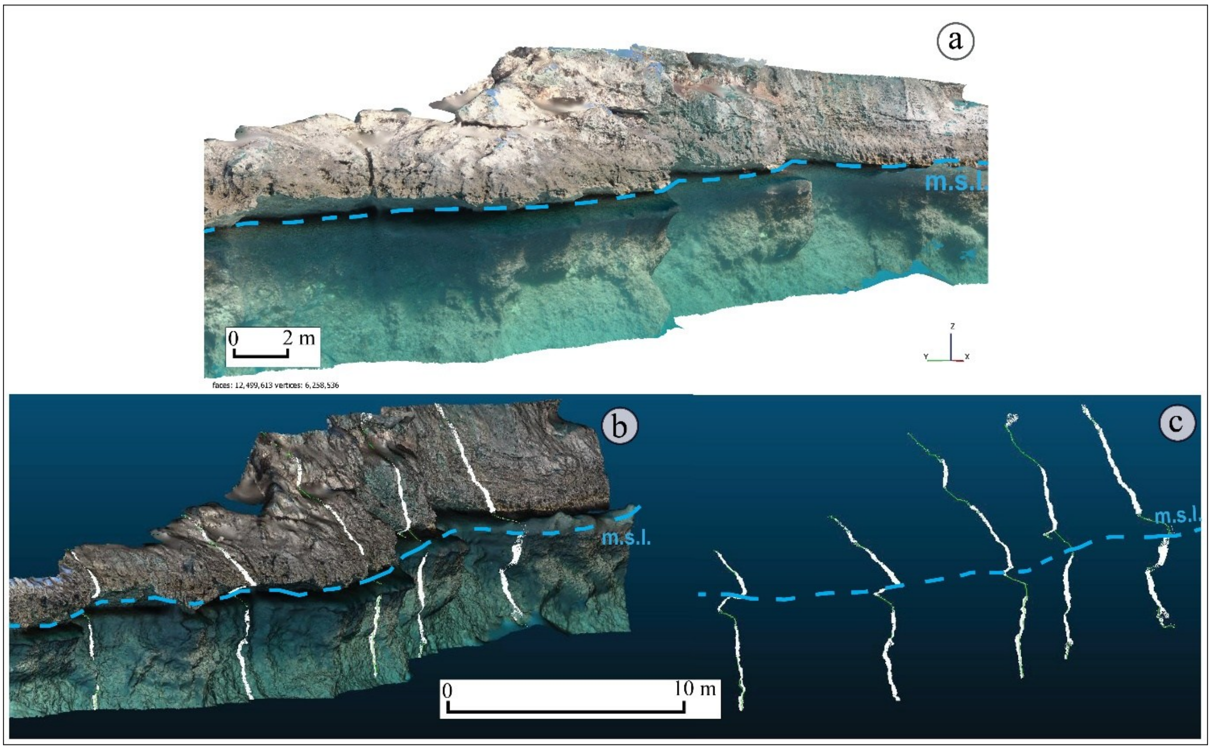

Figure 7.25: Detailed study of coastal erosion features through three-dimensional model analysis. a) Three-dimensional reconstruction with profile extraction locations indicated; b) and c) topographic profiles generated by Cloud Compare software illustrating the shape of the morphology of tidal notch.

The second elaboration of the model in CloudCompare was performed for the analysis of the notch dimensions (Figure 7.26), and the methodology used is describe in Section 5.6.4. In this visualization scheme, red areas indicate maximum distances from the plane (typically representing the deepest portions of the notch), while blue areas represent the closest points to the reference plane. The coastal morphology at Mġarr ix-Xini exhibits a plunging cliff with the visible tidal notch at the intersection of two different sections of the model, above and below the waterline (Figure 7.26). The red/orange areas are the farthest from the built plane (across the model) and the blue areas are the closest. By manipulating the histogram display parameters, the color scale was optimized to emphasize specific distance ranges, allowing for more accurate visualization of the notch morphology (Figure 7.26d). Selective grayscale masking provided additional control over which distance values were visualized in the model analysis.

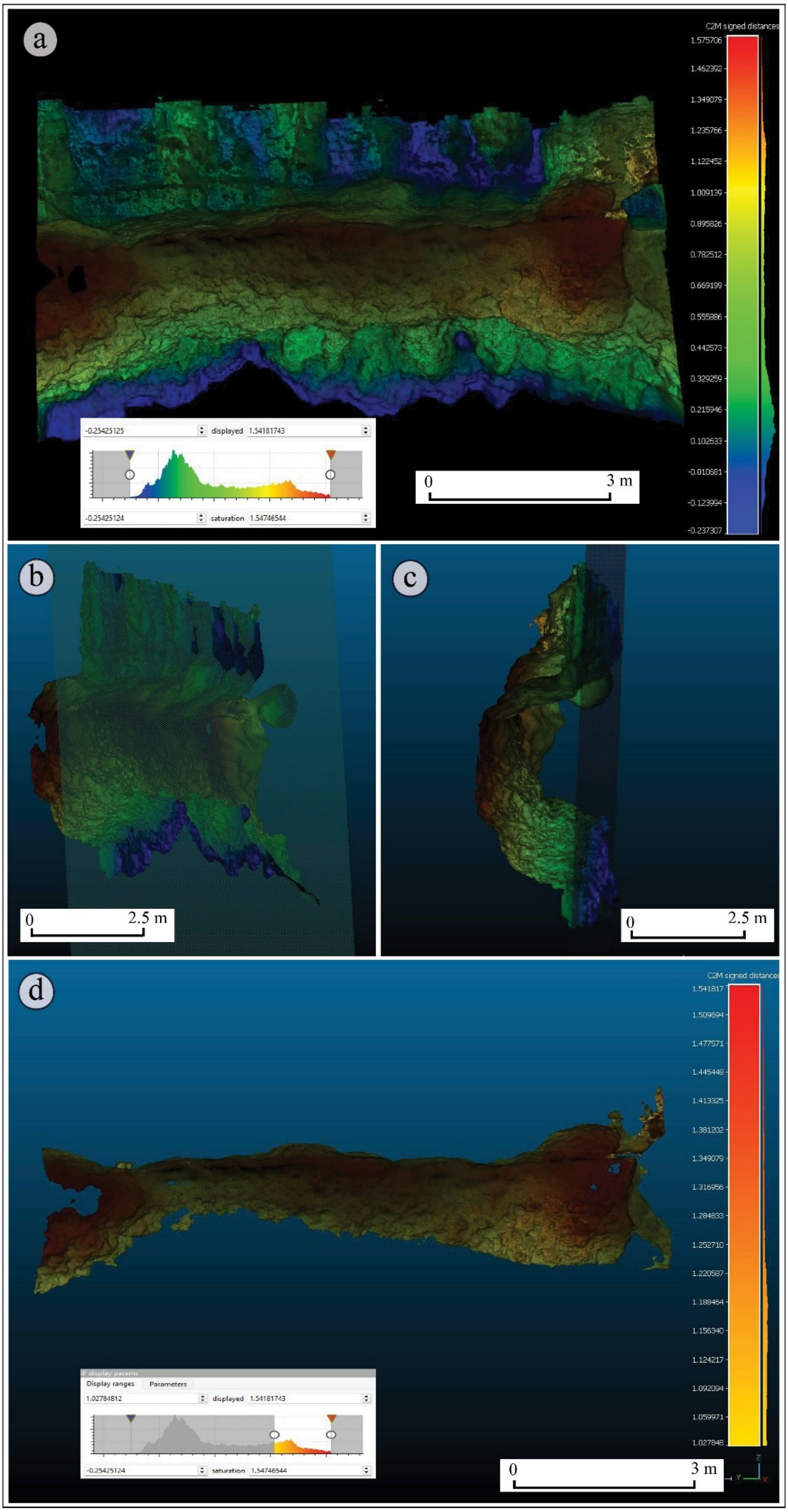
a
C2M signed distances
1.575706
1.462392
1.349079
1.235766
1.122452
1.009139
0.895826
0.782512
0.669199
0.555886
0.442573
0.329259
0.215946
0.102633
-0.010681
-0.123994
-0.237307
0
3 m
b
0
2.5 m
c
0
2.5 m
d
C2M signed distances
1.541817
1.509694
1.477571
1.445448
1.413325
1.381202
1.349079
1.316956
1.284833
1.252710
1.220587
1.188464
1.156340
1.124217
1.092094
1.059971
1.027848
0
3 m

Figure 7.26: Quantitative analysis of the tidal notch by manipulating the 3D model of the Mġarr ix-Xini intertidal zone using CloudCompare software. a), b) and c) front and lateral view of the 3D model, with the corresponding color histogram; d) front view of the 3D model, narrowing the color range to show only the areas furthest from the plane.

Based on the data presented in Table 7.7, a comprehensive analysis of both width and inward depth measurements was performed at twenty measurement points along the tidal notch, randomly selected from the 3D model.

Table 7.7: Descriptive statistics of the width and the inward depth measurements of the tidal notch of Mġarr ix-Xini (n = 20).

	Width (m)	**Inward depth (m)**
	2.45	1.52
	2.22	1.40
	2.57	1.74
	2.41	1.06
	2.69	0.97
	2.97	1.08
	2.53	1.19
	1.83	1.02
	2.00	1.29
	3.09	1.61
	2.04	1.04
	2.12	1.14
	2.39	1.00
	2.35	1.27
	2.16	1.75
	1.93	1.72
	1.96	0.98
	2.31	1.10
	2.47	1.08
	2.34	1.21
$\bar{x}$	2.34	1.26
σ	0.331	0.270
SEM	0.074	0.060

The width measurements of the C-shaped notch show considerable variation, ranging from 1.83 m to 3.09 m. The arithmetic mean ($\bar{x}$), that is the average of the n values, was calculated using the formula:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{5}$$

where n is the number of measurements and x_i represents each individual measurement value. This yielded a mean width of 2.34 m. The standard deviation (σ), that is the spread of data or the dispersion, was determined using:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

(6)

resulting in a value of 0.331 m for width measurements, which quantifies the degree of dispersion in the dataset.

The standard error of the mean (SEM), i.e. the uncertainty in the estimation of the mean, was calculated:

$$SEM = \frac{\sigma}{\sqrt{n}}$$

(7)

This yielded a value of 0.074 m, resulting in a final notch width expression of 2.34 ± 0.074 m.

Similarly, for the inward depth measurements, which ranged from 0.97 m to 1.75 m, the same statistical formulas were applied. The mean inward depth was determined to be 1.26 m, with a standard deviation of 0.270 m. The standard error of the mean for the depth measurements was calculated to be 0.060 m, resulting in a final expression of 1.26 ± 0.060 m.

Using the data obtained, it is possible to make a mathematical estimate of the total eroded volume where the scour marks can now be observed. To do this, is possible to estimate the cross-sectional area as a circular segment, then multiply by the length to get volume. A circular segment area A is calculated as:

$$A = \frac{r^2}{2}(\theta - \sin(\theta))$$

Where r is the radius of the circle and θ is the central angle in radians subtended by the segment. However, since it is not possible to obtain these parameters using the available ones, the following formula can be used as an alternative, approximating a semi-elliptical cross section, similar to a C-shape:

$$A = \frac{\pi}{4} \cdot width \cdot depth$$

$$V = A \cdot length = \frac{\pi}{4} \cdot width \cdot depth \cdot length$$

(8)

By measuring the length of the tidal notch analyzed by the 3D model, a value of 15.5 m is obtained. Therefore, using the formula given in (8), the estimated volume with a width of 2.34 m and a depth of 1.26 m is equal to 35.87 m^3.

7.7.3. Halq it-Tafal, Malta (Malta)

The survey at Halq it-Tafal (Figure 7.27) focused on a 31 meter coastal sector characterized by a prominent high-angle fault structure known as the Maghlaq fault, which defines the shoreline morphology. Environmental conditions were optimal during the survey period, with minimal wave activity and moderate wind speeds averaging 2.1 m/s from the SSE (*Maltese Islands Weather*, 2023).

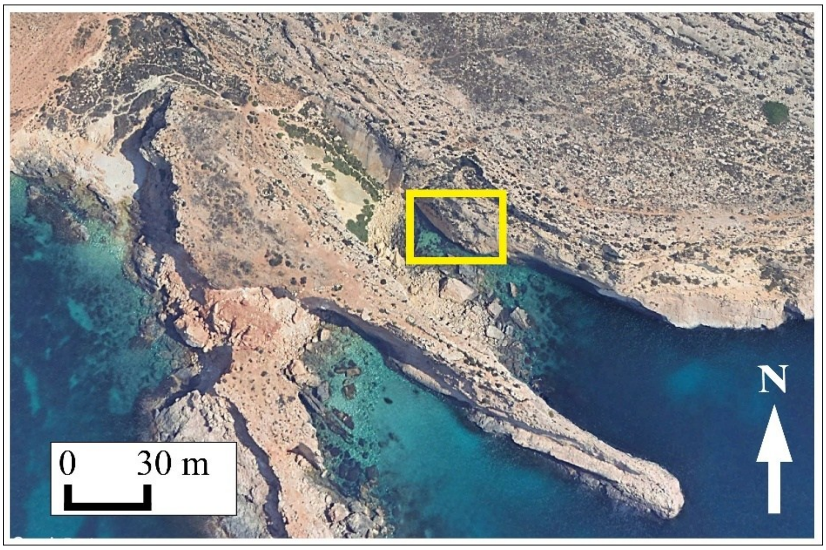

Figure 7.27: Detailed view of the 31 meter surveyed coastal sector, highlighting the prominent Maghlaq fault system that characterizes the coastline.

Three 3D models were created for the Halq it-Tafal site, each using different data integration strategies (Table 7.8, Figure 7.28).

The first modeling iteration used only GoPro images (Figure 7.28a), incorporating 110 aw and 110 bw images. All 220 images were successfully aligned, generating 138,175 tie points and producing a mesh of 5,957,057 faces at a resolution of 0.902 mm/pixel.

The second modeling iteration (Figure 7.28b) expanded the dataset through UAV integration, incorporating 39 aerial images with the existing GoPro data, resulting in 149 aw and 110 bw images. Full alignment was achieved for all 259 images, resulting in 147,509 tie points and a refined mesh of 9,964,161 faces, with a resolution of 0.985 mm/pixel.

The third modeling iteration (Figures 7.28c, 7.28d) integrated Insta360 frame data with the GoPro images, incorporating 127 Insta360 frames (4000x3000 pixels) into the aw dataset. The final model successfully aligned all 347 images, generating 147,809 tie points and a mesh of 691,527 faces, achieving a resolution of 1.16 mm/pixel.

Table 7.8: Parameters of 3D model construction across three different methodological approaches, illustrating the evolution of data integration strategies. 1) Survey area designation; 2) Cross-reference to associated figures; 3) 3D model dataset; 4) Total number of images acquired; 5) Images successfully aligned; 6) Tie points generated during alignment; 7) Total faces in final mesh construction.

(1) Study area	(2) Figure	(3) 3D model	(4) n° images	(5) n° aligned images	(6) Tie points (points)	(7) Mesh (faces)
Halq it-Tafal (Malta)	7.28a,b	aw model (GoPro)	110	110	43 800	3 103 022
		bw model (GoPro)	110	110	94 375	2 852 721
		merged model (GoPro)	220	220	138 175	5 957 057
	7.28c	aw model (GoPro+UAV)	149	149	54 638	7 657 225
		bw model (GoPro)	110	110	92 871	2 972 720
		merged model (GoPro+UAV)	259	259	147 509	9 964 161
	7.28d,e	aw model (GoPro+360)	237	237	105 106	3 780 094
		bw model (GoPro)	110	110	42 703	423 709
		merged model (GoPro+360)	347	347	147 809	691 527

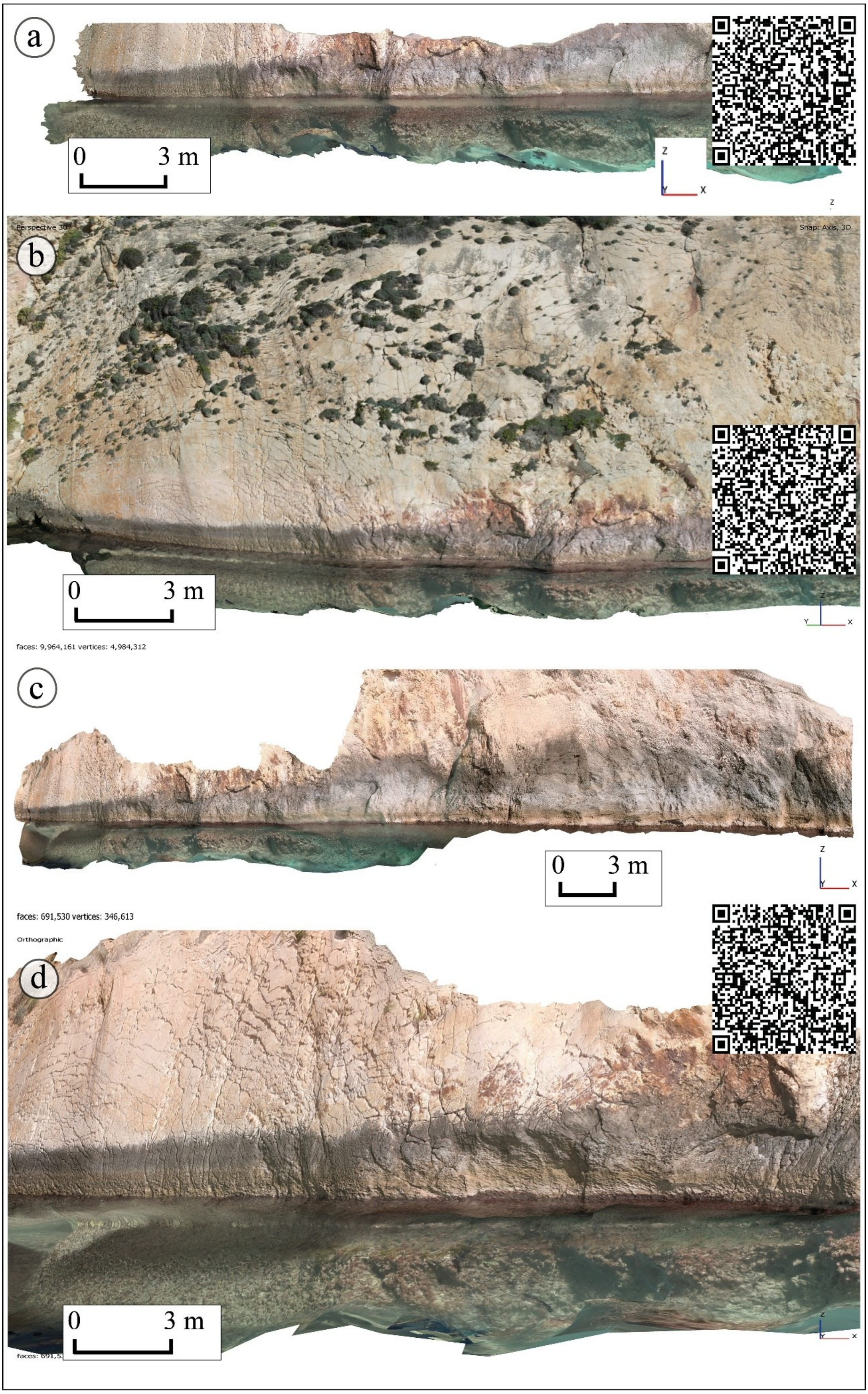
a
0 3 m
b
0 3 m
faces: 9,964,161 vertices: 4,984,312
c
0 3 m
faces: 691,530 vertices: 346,613
Orthographic
d
0 3 m

Figure 7.28: Multi-perspective representation of three-dimensional models using different data integration strategies. a) Frontal view (xz plane) of the baseline model constructed solely from GoPro imagery, demonstrating high-resolution capture of intertidal features; b) Perspective view of the UAV-integrated model, showing expanded vertical coverage of the cliff; c) and d) frontal view and detailed section of the Insta360-integrated model, highlighting improved colour fidelity and natural texture representation. The QR-codes allow to view the corresponding 3D model on the sketchfab.com web platform.

From a qualitative and semi-quantitative point of view, all three models showed complete coverage of the intertidal zone (Figure 7.28), although each had different characteristics based on their data integration approaches. The model built using only the GoPro covered 115 m² and captured the shoreline profile from 2 m above to 1.5 m bw (minimum depth -0.70 m), clearly showing the plunging cliff morphology.

The UAV-integrated model significantly expanded the coverage to 4,970 m², achieving extensive vertical coverage from 38.5 m aw to 1.7 m bw (minimum depth -0.6 m). While this approach improved overall coverage, particularly in the aw areas, the intertidal resolution showed only marginal improvement over the baseline model.

The Insta360-integrated model encompassed 317 m² and provided coverage from 2.3-5 m aw to 1.7 m below (minimum depth -0.60 m).

Similarly, for the roof notch shown in Figure 7.29a, a morphometric analysis was performed using the same methodological approach. The model was processed through CloudCompare to generate the profiles extraction (Figure 7.29), that is described in the paragraph 5.6.4., and the distance measurements from a reference plane, with the color gradient visualizing the spatial variation in notch depth (Figure 7.30).

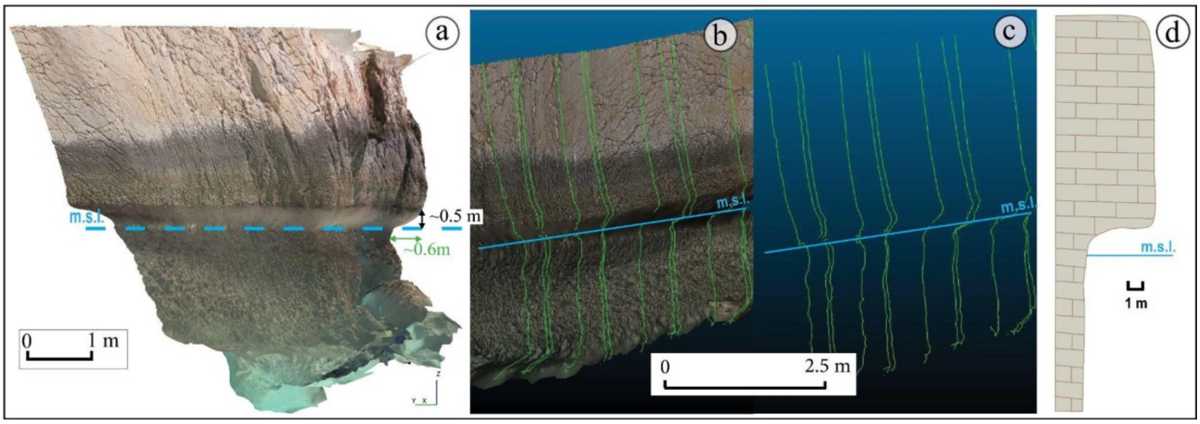

Figure 7.29: Detailed study of erosional features through multi-perspective analysis. a) Complete three-dimensional model with strategic profile extraction locations; b) and c) quantitative topographic profiles generated by Cloud Compare software illustrating the distinctive roof morphology of the tidal notch; d) schematic representation of typical roof notch structure providing context for morphometric interpretation.

In Figure 7.29a, the front view shows the morphometry of the tidal notch. The color gradient ranges from blue (areas closest to the reference plane) to red (areas furthest from the reference plane), with the histogram inset showing the distribution of distance values. The model used for the morphological analysis of the groove shows a slight doming, indicating that the areas of greatest interest in the model are those in green rather than red. In fact, the areas in red, which are those farthest from the transverse plane of the model, correspond to the most distal areas of the curve. The central areas, where the tidal notch is located, are green. By manipulating the color scale (Figure 7.29c), the area represented by the color green was isolated, which, as mentioned, is the area of greatest interest because it contains the tidal notch. Within this area, the depth measurements of the roof notch were randomly selected.

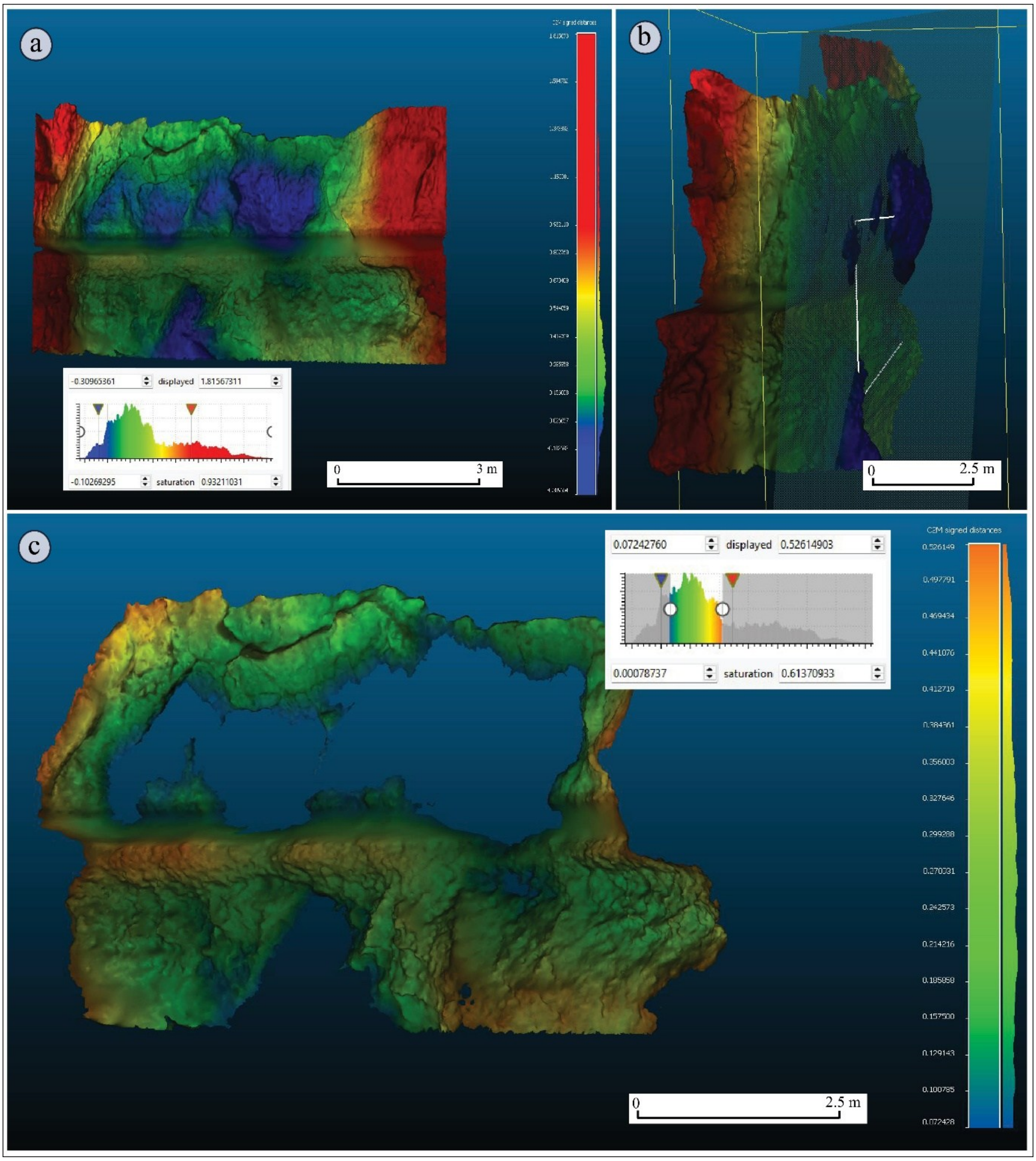

Figure 7.30: Quantitative analysis of the tidal notch by manipulating the 3D model of the Halq it-Tafal intertidal zone using CloudCompare software. a) and b) front and lateral view of the 3D model, with the corresponding color histogram; d) front view of the 3D model, narrowing the color range to show only the areas of interest.

The second quantitative analysis corresponds to the measurements of the roof notch were systematically collected from the model, focusing on the maximum internal depths. Some doming effects were observed in the model, focusing on the green areas rather than relying solely on the red areas for accurate depth measurements. Twenty discrete measurements were taken randomly, ranging from 0.14 m to 0.95 m in depth. Statistical analysis of these measurements yielded a mean maximum internal depth of 0.47 m (5), with a standard deviation of 0.232 m (6). The standard error of the mean (7) was calculated to be 0.052 m, resulting in a final measurement expression of 0.47 ± 0.052 m for the notch depth (Table 7.9).

Table 7.9: Descriptive statistics of the inward depth measurements of the tidal notch of Halq it-Tafal (n = 20).

	Inward depth (m)
	0.60
	0.32
	0.14
	0.75
	0.95
	0.24
	0.32
	0.27
	0.41
	0.37
	0.33
	0.45
	0.67
	0.43
	0.85
	0.88
	0.48
	0.37
	0.27
	0.35
$\bar{x}$	0.47
σ	0.232
SEM	0.052

8. Discussions

8.1. Dataset curation

Geomorphological surveys were conducted in the three study areas in September 2020 and October 2023, respectively. The acquired data were collected according to the protocol explained in Chapter 5.1. Specifically, instrumental data, such as photographic data collected with GoPros and Insta360s, GPS data, bathymetry data collected with an echo sounder, and data on electrical conductivity and water temperature, as well as observational data, were acquired. The observational data from Lampedusa was associated with notes recorded in a field notebook containing measurements of sea caves, tidal notches, and observations of marine life along the coast. Swim surveys are brief with few stops to ensure the continuity of the collected images. Unfortunately, there are no data on the measurements of the tidal notches described below because the campaigns were not specifically carried out for this doctoral thesis. Instead, this thesis stems from the desire to make the most of the photographic data collected during these three swim surveys.

The curation and standardization of the Geoswim image archive has several benefits. First, it ensures that the dataset is easy to use, and the logical structure allows for quick navigation through the image sets or other data.

Secondly, the organization method facilitates data compression. By grouping similar images together, we can implement compression algorithms more effectively, reducing the overall storage images of the dataset. This compression not only reduces storage space but also makes the dataset more manageable for sharing among research teams.

Another key benefit of the curation approach is the linking of images captured above and below the waterline in the same area and at the same time. This connection is fundamental for comprehensive analysis of marine environments, allowing researchers to correlate surface conditions with underwater observations seamlessly.

Perhaps most importantly, the curation strategy lays the groundwork for future implementation of artificial intelligence in image discrimination. As introduced in more detail in Subchapter 7.1, the organized dataset, including the carefully curated "cancelled" folders, serve as a robust training set for CNN models. This enables the development of AI tools capable of automatically categorizing new images, significantly speeding up the process of data organization and analysis in future surveys.

In summary, the manual curation of our extensive image dataset, while time-consuming, has resulted in a well-structured and easily navigable dataset. This organized dataset not only enhances current research capabilities, but also allows for various AI-driven analyzes, such as the identification of coastal features.

The results demonstrate the progression from image data organization, through advanced machine learning applications, to final three-dimensional coastal modelling.

8.2. Geoswim Image Viewer

The development of the Geoswim Image Viewer presented several challenges, each requiring different solutions to ensure greater efficiency.

First, the management and processing of a large number of high-resolution image pairs, implement multitasking computation in Python was used for faster image processing, and image compression was optimized to balance image quality and file size, as well as implement a system to load images very quickly.

Second, the creation of a responsive design that works well on devices ranging from mobile phones to large desktop monitors using a CSS grid has been performed.

Third, the accuracy and consistency of EXIF data throughout the image processing was ensured. This viewer can be easily distributed by simply sharing a link leading to the server (https://valeriavaccher.com/geoswim/).

The Geoswim Image Viewer allows simultaneous visualization of images above and below the waterline. Specifically:

1. It facilitates a direct comparison of the above and below images. The viewer allows to analyze the sequence of images collected during the survey;
2. It facilitates to easily navigate through large datasets of georeferenced image pairs;
3. The synchronization of the aw or bw images provides a unique perspective on the transition between aerial and underwater intertidal landforms;
4. The system's architecture lays the groundwork for future integration of AI and machine learning algorithms. It could be useful for the identification and classification of geological targets and landforms in the intertidal zone, significantly speeding up data analysis and pattern recognition;
5. The integration of geolocation data and interactive mapping capabilities allows the viewer to better understand the spatial relationships between images captured above and bw.

The Geoswim Image Viewer has potential for future development, such as implementing machine learning algorithms to automatically identify and label geological targets or specific landforms in both aw and bw images, developing AI models to analyze temporal changes in matched images, highlighting areas of significant change, integrate species recognition capabilities to support biodiversity studies, implement features to create annotations and plots, develop a system to track changes and annotations made to images over time, develop a login page for many users to work simultaneously, extend the system to allow users to directly upload and download new images and data during and after field surveys.

8.3. The Convolutional Neural Networks

Image recognition using AI and computer vision technology is improving all the time. These technologies are revolutionizing the way to analyze coastal images and extract valuable information to help monitor and manage coastal environments.

At the forefront of this technological evolution are DL models, particularly CNNs. These algorithms have shown remarkable ability to automatically classify and segment various elements within coastal images. From distinguishing between water, sand, vegetation, and man-made structures to enabling automatic coastline detection, CNNs are proving invaluable in quantifying coastal features (Buscombe & Ritchie, 2018; Brose et al., 2025; Sabato et al., 2025; Sozio et al., 2024; Valentini et al., 2017).

In the context of geomorphology, the applications of AI in coastal image analysis are diverse and far-reaching, and not innovative. Other authors, such as Blais and Akhloufi (2025), Khurram et al. (2025), Sozio et al. (2024, 2025), and Sabato et al. (2025), have used convolutional networks and machine learning techniques to recognize geomorphological features in the nearshore area, detect waste in coastal environments, facilitating automated monitoring of coastal erosion and accretion (Valentini et al., 2017). They have also shown promise in estimating water depth from optical images, providing a cost-effective approach to mapping nearshore bathymetry (Bergsma et al., 2016). In addition, AI techniques are being used to analyze water quality parameters such as algal blooms, turbidity, and floating debris (Hoonhout et al., 2015), and to classify different coastal habitats and ecosystems from aerial and satellite images (Chen et al., 2020).

There are numerous benefits to using AI for coastal image analysis. These systems can efficiently process large amounts of image data, extract quantitative metrics, and provide consistent and objective analysis. Developing robust AI models that can handle the variability inherent in coastal environments and image conditions remains an ongoing effort. Current research focuses on improving model generalization, reducing reliance on large, labelled datasets, and developing explainable AI systems.

As AI techniques continue to advance, their potential to enhance coastal monitoring capabilities and support coastal management and conservation efforts is growing (Jianping et al., 2016). The integration of AI with other emerging technologies, such as drones and IoT (Internet of Things) sensors, promises to further expand coastal applications. Innovations such as the Sea-thru method, which aims to remove water column effects from underwater images, could unlock vast underwater datasets for powerful computer vision and machine learning algorithms, opening up possibilities for underwater exploration and conservation (Akkaynak and Treibitz, 2019).

In the context of the Geoswim project, the application of AI techniques is particularly promising. As part of this thesis, AI was useful to understand the limitations and advantages of this technology in the recognition of coastal images, in order to discriminate, from a large set of images, those that are useful or not useful for the study and analysis of rocky coasts. In fact, thanks to AI, it has been tested that the time taken by the machine to clean the dataset from images containing boats or people, blurred images, or images where the coast is very far away and not very visible, is significantly less than the time taken by a human operator. In fact, for a dataset of this size (about 9 terabytes), it would take months of work, compared to just a few minutes for the computer.

Further testing was done using AI to recognize the type of coastline, i.e. whether it is a coastline with cliffs facing the sea, rather than a low coastline with a beach (see Chapter 4.1 for the classification used). Although the dataset used for supervised training was unbalanced, i.e. not providing the same number of images for each class, and the images were very similar to each other, it did not allow the network to develop a fully efficient model. Nevertheless, with the prospect of increasing our dataset with images collected from different field surveys, it still produced a satisfactory result, showing its ability to effectively recognize coastline features as opposed to areas covered by water, and to learn the diagnostic characteristics for each class type.

Both models contribute to the development of a system that aims to replicate the decision-making capabilities of experienced geomorphologists. While human expertise remains valuable, these automated systems offer advantages in processing speed, consistency, and scalability. Understanding the performance limitations and potential sources of error in these systems is critical to their continued refinement and practical application in coastal research and management.

8.3.1. Binary classification considerations and future optimization

The binary classification system demonstrated effectiveness in discriminating between useful and useless coastal images. This performance was achieved through a methodically structured implementation strategy applied to randomly selected images from the Geoswim dataset, including photographs collected both above and below the waterline. The system effectively identified useless images characterized by the presence of swimmers, boats, flippers, or other unwanted objects, as well as images where the shoreline was too far away for meaningful analysis.

The implementation architecture provided a framework for the classification task. This development environment ensured consistency in package versions and dependencies, which is critical for reproducibility in scientific computing. This pattern suggests effective feature extraction during early training phases, with subsequent fine-tuning of feature weights contributing to improved classification accuracy.

Confusion matrix analysis (Figure 7.8., Table 7.1) quantitatively validated discriminative capabilities of the model, with accuracy reaching 98.8%, reflecting balanced performance across both positive and negative classes. The precision and recall metrics further support the robust performance of the model, indicating minimal false positives and false negatives.

The mini-batch optimization approach, using a carefully calibrated batch size of 16 samples, balanced computational efficiency with model convergence, contributing to stable learning behaviour. This batch size proved to be optimal for the available computational resources while maintaining sufficient gradient estimation accuracy.

The comparative analysis between GradCAM and GradCAM++ (Figure 7.11) revealed different characteristics in their visualization approaches. GradCAM++ demonstrated greater focus in identifying critical areas of images corresponding to objects. Its heatmaps showed clearer boundaries and more focused activation regions around significant objects, while GradCAM showed broader sensitivity to image variations with more diffuse activation patterns. This improved localization capability of GradCAM++ is particularly valuable for understanding which specific image features contribute most to classification decisions.

GradCAM++ consistently produced more focused and localized activation maps compared to the broader patterns of GradCAM. This difference was particularly evident in image G0013657.JPG, where GradCAM++ demonstrated a superior ability to reduce noise and focus on relevant features, providing a clearer indication of the areas influencing the model's classification decisions.

The performance of the model on the difficult cases provided interesting insights into its learning behavior. Edge cases involving partial human presence or ambiguous non-geological features highlighted areas where the model's decision boundaries could be further refined. For example, in some cases the network was unable to distinguish the coast because the rock face was too close to the target, or because it detected very rough water and identified the image as one to be cancelled. In these cases, the operator would certainly not have made the mistake, but as mentioned above, the network output can be modified.

Future improvements will focus on expanding the variety of training images, both above and below the waterline, and incorporating additional test images do not present in the original dataset.

8.3.2. Multiclass classification

The multi-class classification analysis revealed complex patterns in model behaviour and performance that provide important insights into the capabilities and limitations of DL approaches to coastal geomorphology classification. The model's overall accuracy of 80.6% shows promising potential, while the underlying patterns in performance across different landform types suggest important considerations for future development (see Section 7.5).

The training dynamics, observed over 500 epochs, showed distinctive patterns that merit careful interpretation (Figure 7.12). The training loss exhibited expected behaviour, starting at approximately 1.2 and improving rapidly over the first 100 iterations before stabilizing between 0.1-0.2. However, the validation loss pattern diverged significantly, stabilizing at a higher range (0.6-0.8) with notable fluctuations. This discrepancy between training and validation performance suggests potential overfitting, due to the fact that some classes have limited data and the dataset is unbalanced.

The model showed particularly robust performance in identifying plunging cliffs (F1 score = 0.821) and sloping coasts (F1 score = 0.851). This can be attributed to their distinctive morphological characteristics and adequate representation in the training dataset. The confusion matrix showed 55 true positives for plunging cliffs and 54 for sloping coasts, indicating a strong discriminative ability for these classes. Shore platforms achieved remarkable precision (0.944), indicating exceptional reliability in positive identifications, although the lower recall (0.773) suggests that some cases were incorrectly classified. This pattern suggests that while the model is very conservative in classifying shore platforms, it occasionally misses valid features, particularly confusing them with scree formations, as evidenced by 4 false negatives in this category.

The mini-batch optimization, using a carefully calibrated batch size of 4 samples as the best configuration as possible after trying several different configurations, effectively balanced computational efficiency with model convergence, contributing to stable learning behaviour.

The GradCAM and GradCAM++ visualizations provided insight into the model's decision-making process. For plunging cliffs, the model consistently focused on the cliff-water interface, particularly evident in images G0014059.JPG and G0014964.JPG (Figure 7.14). Activation patterns showed more focus to the characteristic vertical drop into deep water, consistent with geomorphological knowledge. Both visualization techniques showed concentrated heat maps at these critical transition zones, validating the model's learning of key diagnostic features. For pocket beach classification (Figure 7.15), the model achieved perfect accuracy (1.000) despite limited dataset representation (only 3 instances in the test set). The visualizations showed feature recognition, with the model identifying both the rocky coastlines on either side and the distinctive crescent-shaped interface between the beach and the water. Images G0020682.JPG and G0021270.JPG demonstrated the model's focus to these defining morphological elements.

Scree formations posed the greatest problem (F1 score = 0.708), with higher recall (0.829) but lower precision (0.618) (Figure 7.16). Visualization analysis revealed complex patterns of feature, focusing on both the irregular texture of the accumulated debris and its relationship to the underlying cliff face. This dual focus likely contributed to both the strengths and weaknesses of the classification, as these features often occur in multiple landform types.

The observed performance patterns suggest several critical areas for improvement in dataset composition and balance. The perfect precision but lower recall in some classes suggests overly conservative classification behaviour, particularly for underrepresented landforms.

Future iterations should focus on expanding the dataset, with particular focus to balanced representation across classes. The visualization analysis showed that the model successfully learned to identify key morphological features but sometimes had difficulty with feature combinations that occur across multiple landform types. This suggests the need for more feature level recognition in the model architecture. Model performance varied significantly depending on the clarity of landform boundaries and feature discriminability, suggesting the potential benefit of implementing the focus that can better handle transitional zones between landform types.

The problem is compounded by the inherent complexity of coastal landform classification, where features often overlap and share common elements. For example, the presence of cliffs in multiple categories ("plunging cliffs," " scree," and "sloping coasts") requires the model to learn subtle distinctions rather than relying on the mere presence or absence of features. The current dataset structure may not provide enough examples of these subtle variations, causing the model to overfit to the specific examples present in the training set.

Several approaches could be implemented in future work to address these overfitting issues. One basic solution would be to expand the dataset not only in terms of quantity, but more importantly, in terms of diversity. This could be achieved by including images from different geographic locations, different weather conditions, different tidal conditions, and different seasons. Such diversity would help the model learn more robust and generalizable features, rather than memorizing specific instances.

In addition, implementing more data augmentation techniques specifically designed for geological features could help address the limited diversity in the training set. Unlike standard image augmentation techniques, these would need to preserve the geomorphologically significant features while introducing meaningful variation. This could include techniques such as realistic texture synthesis for rock surfaces, simulation of different water conditions, and careful geometric transformations that preserve the physical plausibility of the landforms.

The architecture of the model itself could also be modified to better handle the inherent complexity of coastal landform classification. The implementation of regularization techniques such as dropout with optimized rates could help prevent the model from becoming too dependent on specific features.

Another promising direction would be to implement a learning curriculum approach, where the model is initially trained on unambiguous examples, before gradually introducing more ambiguous cases. For example, the image G0024174.jpg (Figure 7.14) is identified as a plunging cliff (_0) rather than a scree (_2) because it more easily identifies the presence of the cliff behind rather than the boulder in front. The decision to include ambiguous images made it possible to identify the limitations of the network.

This could help establish a stronger base for feature recognition before tackling the more challenging aspects of landform classification. In addition, incorporating physical constraints and domain knowledge into the learning process, perhaps through custom loss functions that penalize physically improbable classifications, could help guide the model toward more geomorphologically predictions.

These results have significant implications for the application of DL to coastal geomorphology. The model's success in identifying distinctive landform features suggests potential applications in automated coastal monitoring and change detection. In fact, if the coastal geomorhotypes are identified with a high degree of accuracy, it is possible to obtain an almost perfect class output, allowing statistical analysis for the specific area surveyed.

However, the variation in performance across different landform types indicates the need for careful consideration of confidence thresholds in practical applications. The visualization results demonstrate that DL models can successfully learn geomorphologically relevant features without explicit programming.

The successful application of GradCAM and GradCAM++ visualization techniques demonstrate the potential for developing more interpretable and geomorphologically informed classification systems. These findings not only validate the model's learning process but also provide valuable direction for improving future iterations of coastal landform classification systems.

8.3.3. Comparison of manual selection vs. AI-based approach

The manual classification process, while providing valuable initial insights and training data, has significant limitations when dealing with our extensive 9 terabyte dataset. This approach requires a significant investment of human resources and time, potentially extending the analysis phase over weeks or months.

The implementation of CNNs for image classification shows remarkable potential for transforming this workflow. Our results show that automated classification can dramatically reduce processing time while maintaining high levels of accuracy. The CNN-based approach achieved an accuracy rate of 96% in binary classification tasks, demonstrating its ability to effectively discriminate between useful and not-useful coastal images. This level of performance suggests that AI-based systems can reliably handle the initial screening of large image datasets, significantly reducing manual workload while maintaining scientific rigor.

Consistency emerges as a significant advantage of the AI-based approach. While human classification can vary due to fatigue, subjective interpretation, or varying levels of expertise, our CNN model demonstrates remarkable consistency in the application of classification criteria. This standardization is particularly valuable for large-scale coastal studies, where maintaining consistent classification standards across thousands of images is essential for reliable scientific analysis.

The scalability of the AI approach is a strong argument for its adoption in coastal research. As our image database continues to grow, the manual approach becomes increasingly unsustainable, requiring a proportional increase in human resources. In contrast, our CNN model demonstrates excellent scalability, processing larger datasets with minimal additional resource requirements once properly trained. This capability is particularly valuable for long-term coastal monitoring programs, where data volumes are constantly increasing.

However, the transition to AI-based classification is not without its challenges. Developing and training effective CNN models requires a significant initial investment in both computational resources and expertise. Our experience shows that careful preparation of training data, extensive model development, and ongoing refinement are essential to achieving reliable results. The "cancelled" folders created during our manual classification process proved invaluable in training the CNN to discriminate between useful and not-useful images, highlighting the complementary nature of manual and automated approaches.

The detection of coastal features presents an interesting comparison between human and machine capabilities. While experienced researchers excel at identifying complex geomorphological features based on contextual understanding, the CNN model shows the ability to detect patterns and relationships that may be missed by manual classification.

This methodology allows to leverage the speed and consistency of automated classification while maintaining the high accuracy standards required for scientific research. The system flags uncertain classifications for expert review, ensuring that complex or unusual coastal features receive appropriate focus. Quality control considerations led to implement a hybrid approach that combines the efficiency of AI-based classification with human oversight for challenging cases. In fact, as mentioned above, the output data from the models can be viewed and modified by the operator at any time. In the case of binary classification, some images that were not correctly classified can be manually moved or deleted and placed in the correct class.

Looking ahead, the development of AI technologies promises to further enhance our ability to monitor and understand coastal environments. However, it remains critical to recognize that these tools augment rather than replace human expertise. The knowledge and interpretive skills of geomorphologists and coastal scientists remain essential for meaningful analysis of the patterns and trends identified by automated systems.

8.4. 3D modeling

The implementation of useful approaches for generating high-resolution three-dimensional models of rocky coasts has demonstrated both significant potential and notable challenges in coastal geomorphological research (Furlani et al. 2020). Through a detailed analysis of three Mediterranean case studies - Cala Uccello (Lampedusa), Mġarr ix-Xini (Gozo), and Halq it-Tafal (Malta) - the integration of above and below the water surface imaging devices was quantitatively evaluated, providing insights into methodological optimizations and limitations. Respectively, the Sections 8.4.1, 8.4.2 and 8.4.3 highlighted the 3D models results for the three case studies.

All the reported results suggest several critical considerations for future research and methodological refinement. The development of improved methods for integrating data from different sensors could simultaneously optimize coverage and resolution. Applying this acquisition methodology along rocky coastlines in different locations would allow the image dataset to be expanded and diversified.

The analysis of 3D models realized not only supported the identification of morphometric elements but also the identification of landforms useful for investigating the morphoevolution of an areas.

8.4.1. Cala Uccello, Lampedusa

The method used to collect the data presented some problems due to the sloping coastal topography at the Cala Uccello site (Figure 7.21a, b). While surveying the area, a technical problem occurred where the return of the wave pushed the equipment against the coast and moved the underwater GoPro. This required the GoPro to be repositioned correctly, but the change in camera orientation caused some problems with image matching.

First, the camera orientation was not constant during the surveying along the sloping coast, mainly due to the waves. This variability affected both the image quality and the assumed constant elevation (+0.5 meters aw and -0.5 meters bw) imposed on the datasets during the processing. Second, a technical problem (the underwater equipment touched the seabed) forced the camera orientation to be changed during the survey, further complicating the image alignment process.

Finally, the presence of waves obscured clear views of the coastline, particularly the clarity of the intertidal zone.

Despite these limitations, the GNSS data from the GoPro cameras allowed coherent scaling and orientation of the models. This is consistent with previous observations in applications of similar scale using comparable GNSS chipsets (e.g., Corradetti et al., 2022).

The integration of 360-degree camera frames proved to be beneficial for coastline reconstruction, especially in Zone 2 (Figure 7.21b, c). However, it is noteworthy that the models generated from GoPro images alone (Zone 1: 0.832 mm/pixel, Zone 2: 0.941 mm/pixel) achieved higher resolution compared to the model incorporating 360-degree images (2.18 mm/pixel). As can be seen in Figure 8.1, a comparison was made between the models generated in Zone 1 (Figure 8.1a), where only the GoPro images were used, and Zone 2 (Figure 8.1b), where both the GoPro and 360-degree camera images were used. Although the scale of the models in the figure is larger for Zone 2 than for Zone 1, it can be seen that the quality in Figure 8.1b is lower and the image is less defined.

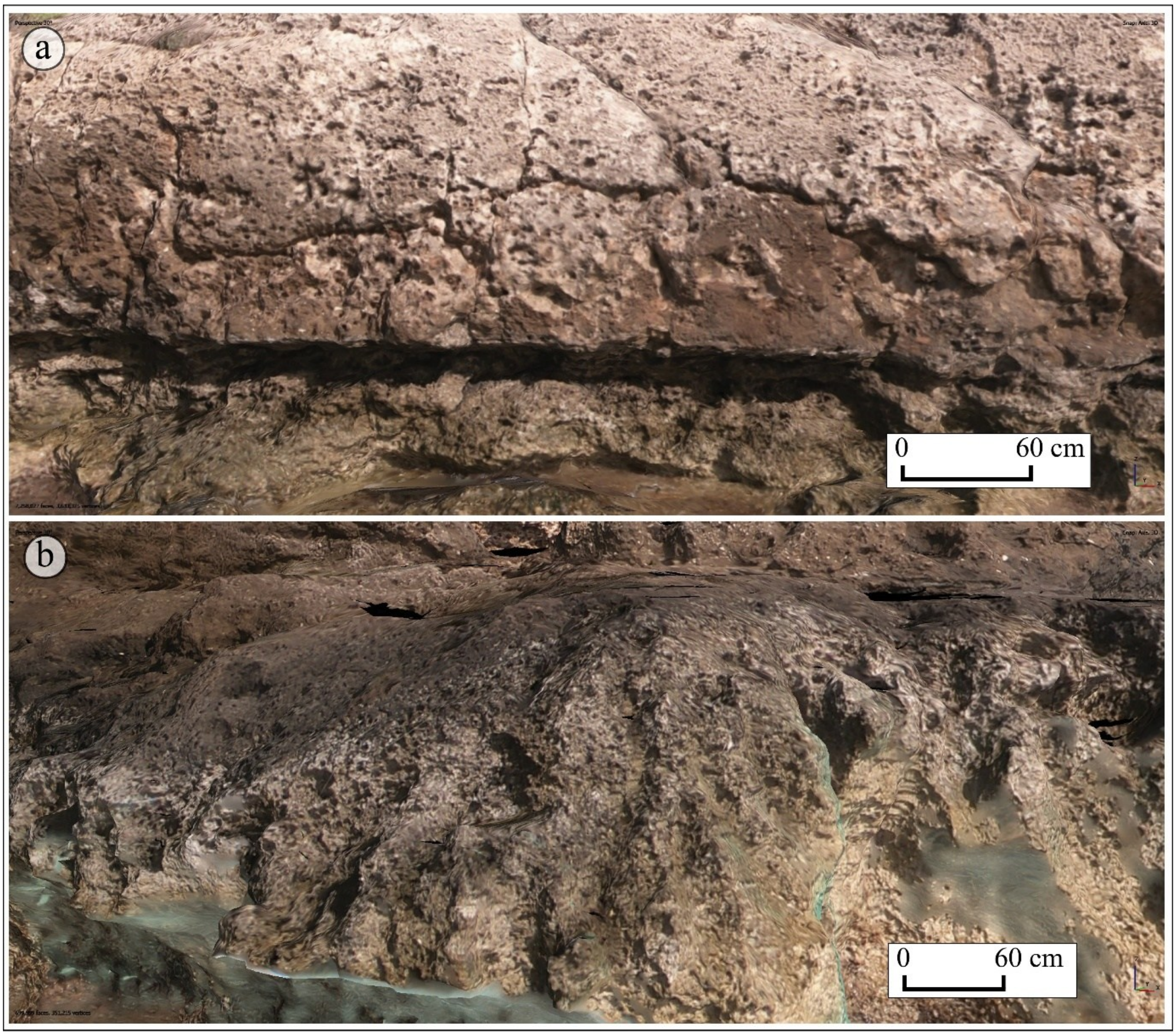

Figure 8.1: Comparison between the resolution a) of the zone 1 model (GoPro only) and b) of the zone 2 model (GoPro + 360 camera) of the area of Cala Uccello.

The continuous and linear gap observed in the 3D model of the Zone 1 of Cala Uccello (Figure 7.21a) was caused by wave motion at the time of the acquisition. Since aw photos cannot faithfully reproduce the underwater environment and bw photos have the same problem for the surface environment, the gap is created. This highlighted a fundamental limitation: above-water photos cannot faithfully reproduce the underwater environment, and below-water photos face similar constraints for surface features. Additional gaps observed in the 3D model of Zone 2 resulted from visual occlusion between photos and target areas in over articulated geometries, suggesting that the sea level concentrated acquisition method (+/- 0.5 meters) is less effective for complex geometries, while remaining highly suitable for cliff studies. These gaps observed in the 3D model of the Zone 2 of Cala Uccello (Figure 7.21b, c) are due to the visual occlusion between the photos and the target, for over-articulated geometries.

Therefore, the acquisition method, which is concentrated at sea level (± 0.5 meters) is less effective for complex geometries, while it is much more effective for cliff studies. For all these reasons, it is necessary to carry out the survey with the sea as calm as possible. In addition, Zone 1, which is further inland and less sloping, requires surveying from a greater distance from the shore. In contrast, Zone 2 is more exposed and less sloping, allowing for a closer instrument approach. However, the increased wave action in Zone 2 resulted in greater water turbidity, resulting in less detailed underwater models.

A comparative analysis was also performed between the coastlines derived from satellite data (image ©2024 Airbus, acquired on June 30, 2023) and those generated by the model for both areas (Figure 8.2). While the overall orientation and scale are consistent on an observational scale, the model for Zone 1 shows a shift of eight meters south of the expected position, while the coastline for Zone 2 shows greater positional accuracy. In fact, given that the GPS values were taken from the GoPros used to collect the data, the yellow coastline in Zone 1 refers to the need to keep the equipment away from the sloping coast. On the other hand, the green line in Zone 2 is much closer because the coast, while still sloping, was a little deeper and allowed for closer acquisition. Despite the closer acquisition, the water turbulence caused by waves and the nature of the coastline are not ideal for creating a 3D model to observe the intertidal zone.

The results from Cala Uccello highlight the need for adapting methods when surveying sloping coastlines. The limitations of horizontal information could potentially be mitigated by incorporating UAV data, which provides a complementary perspective. However, as noted by Furlani et al. (2020), UAV acquisition typically provides less detail in the vertical direction, highlighting the value of this approach for intertidal coverage.

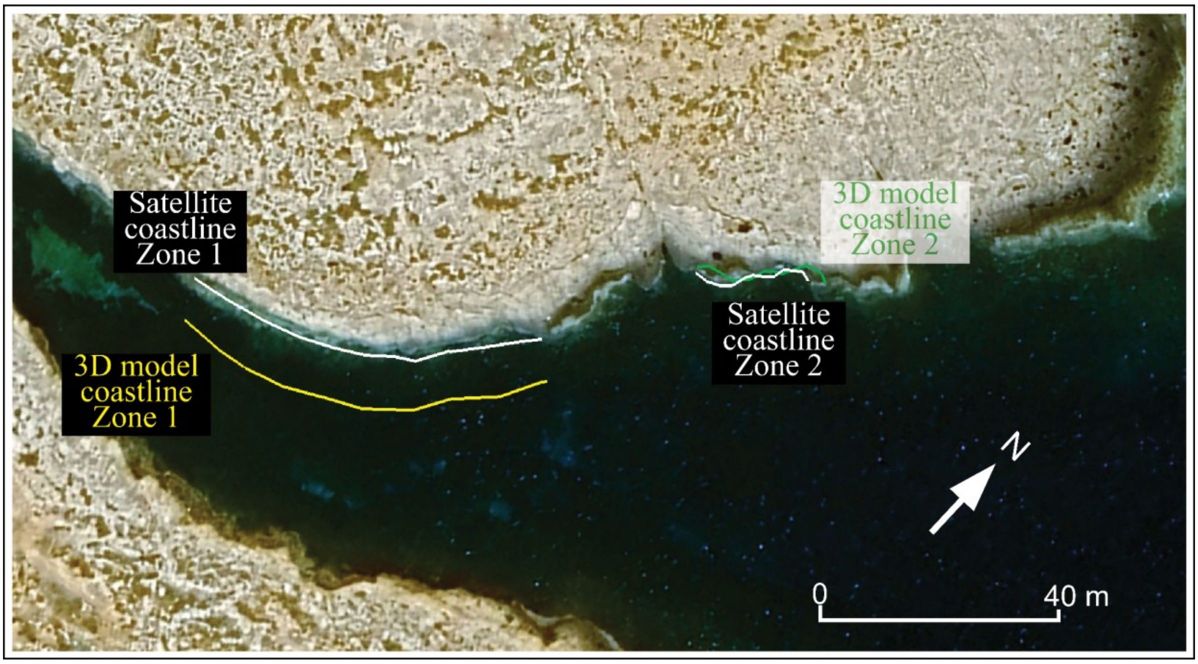

Figure 8.2: Google Earth Image ©2024 Airbus (acquisition 30th, June 2023) view of the sloping coast at Cala Uccello. The GE line-drawing coastlines are represented by a white polyline. The line-drawing of Zone 1 3D model is shown in yellow, while the coastline of Zone 2 model is shown in green.

Several authors (Antonioli et al., 2015; Furlani et al., 2020; Furlani et al., 2024) have stated that the Lampedusa area features tidal notches that cover most of the island's rocky coastline.

The tidal notch at the Cala Madonna site was particularly studied for this thesis due to its characteristics being quite similar to those observed by Furlani et al. (2024). Although there is good observational data and in situ measurements for the Cala Madonna site, the images collected are insufficient for a 3D reconstruction of the area. The opposite is true for the Cala Uccello site; the photos collected were insufficient for creating a good 3D model.

As Furlani et al. (2024) described, calcifying red algae were observed in this area. Numerous scarring from bioerosion caused by sea urchins is present in the subtidal zone.

8.4.2. Mġarr ix-Xini, Gozo

The Mġarr ix-Xini case study produced comprehensive 3D models of the coastal area, successfully capturing the shape of the intertidal coast. However, a notable color normalization issue emerged during the texturing process: the merged model exhibited a bluish tint in the above-water sections, mirroring the tones observed in the underwater photographs. Several factors contributed to this problem: a) different camera calibrations for above and underwater conditions, b) irregular lighting conditions between surface and underwater areas, and c) variations in image quality due to these environmental factors.

Similar to the case of Cala Uccello, the integration of 360-degree camera data resulted in lower resolution models, with the model generated by the GoPro alone having a resolution of 1.33 mm/pixel (Figure 7.24a, b) and the model generated by the GoPro and 360-degree having a resolution of 4.03 mm/pixel (Figure 7.24c, d).

This consistent pattern across sites suggests that while 360-degree cameras provide extended coverage, they may compromise the final model resolution.

Despite color normalization issues, the models shown in Figure 7.25 allow observation of the morphology of the tidal notch with measurements of depth and vertical extent, defining like as a roof notch. In addition, a red algae trottoir can be observed along the intertidal zone.

The implementation of CloudCompare software provided clarity in the visualization, with color-coded distance mapping highlighting the dimensional characteristics of the notch (Figure 7.26). This technique clearly distinguished the deepest parts of the notch, shown in red and orange, from the areas closest to the reference plane, shown in blue (Figure 7.26a, b, c). This visual differentiation facilitated the precise identification of the boundaries of the notch and its internal morphology. The manipulation of the histogram parameters allowed for closer observation of the area bounded by the notch and more accurate depth measurements, which were used to develop a statistical analysis with 20 values (Figure 7.26d). The average width of 2.34 ± 0.074 m and the average depth towards the interior of 1.26 ± 0.060 m show remarkable consistency, despite the natural variability inherent in the characteristics of coastal erosion (Table 7.7). The standard deviation values (0.331 m for width and 0.270 m for depth) suggest that although there are some variations along the 15.5 m of studied coastline, erosive processes have operated with relative uniformity. This consistency may reflect similar exposure to wave energy and rock resistance properties throughout the study area, contributing to the formation of a relatively homogeneous profile.

The quantitative calculation of the volume of 35.87 m^3 represents a value that defines the amount of removed material.

The tidal notch along the Maltese coast has already been studied. Furlani et al. (2024), for example, present a table showing data related to the site. However, these data do not match those analyzed in this thesis. Furlani et al. (2017b) studied the Maltese island coasts and reported in a table the tidal notch values of other Gozo areas. However, these values are not directly comparable to the measurements made in this thesis using 3D models.

Comparative analysis between the satellite-derived coastline (Image ©2024 Airbus, acquired June 30, 2023) and the first model coastline (Figure 8.3) shows consistent overall orientation and scaling. However, a systematic offset of about 5 meters is observed, due to the distance maintained between the survey raft and the coastline during data acquisition. The yellow coastline defined by the model reflects the distance maintained between the raft and the coastline in order to obtain a good photogrammetric acquisition in that area.

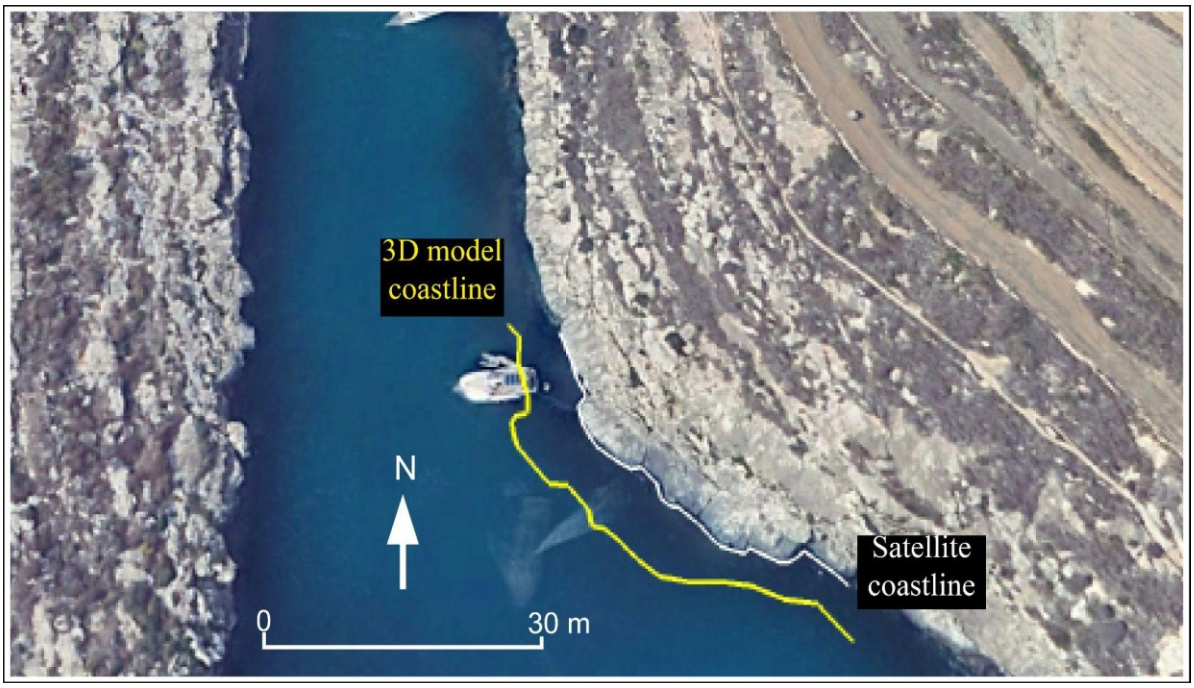

Figure 8.3: Google Earth Image ©2024 Airbus (acquisition 30th, June 2023) view of the plunging cliff at Mġarr ix-Xini. The GE line-drawing coastlines are represented by a white polyline, while the line-drawing of 3D model is shown in yellow.

8.4.3. Halq it-Tafal, Malta

The Halq it-Tafal site allowed for a comprehensive comparison of models produced using different data combinations. The GoPro-only model achieved the highest resolution (0.902 mm/pixel) and provided excellent detail of the intertidal zone (Figure 7.28a); the GoPro-integrated UAV model provided extended coverage aw (up to 38.5 m) while maintaining good resolution (0.985 mm/pixel) (Figure 7.28b); the GoPro-integrated 360-degree model provided intermediate resolution (1.16 mm/pixel) with improved color representation of the cliff wall (Figure 7.28c, d).

The addition of UAV data significantly increased the vertical coverage of the model, but showed some limitations: a) at larger scales, the UAV-enhanced model showed lower resolution and more blur than the GoPro-only model, b) the intertidal zone in the UAV-enhanced model showed a lower concentration of points, possibly due to the difficulty of capturing this area from an aerial perspective. To better observe these differences, a comparative analysis of point cloud density and texture quality was performed between the models. The GoPro-only approach consistently demonstrated superior performance in maintaining high point cloud density and texture sharpness. When examined at higher magnification (Figure 8.4a, b), the base model maintained exceptional clarity and detail, while the UAV-integrated model showed significant degradation in resolution and dense cloud density, particularly pronounced in the intertidal zone where fusion of underwater and aerial datasets proved challenging (Figure 8.4c, d).

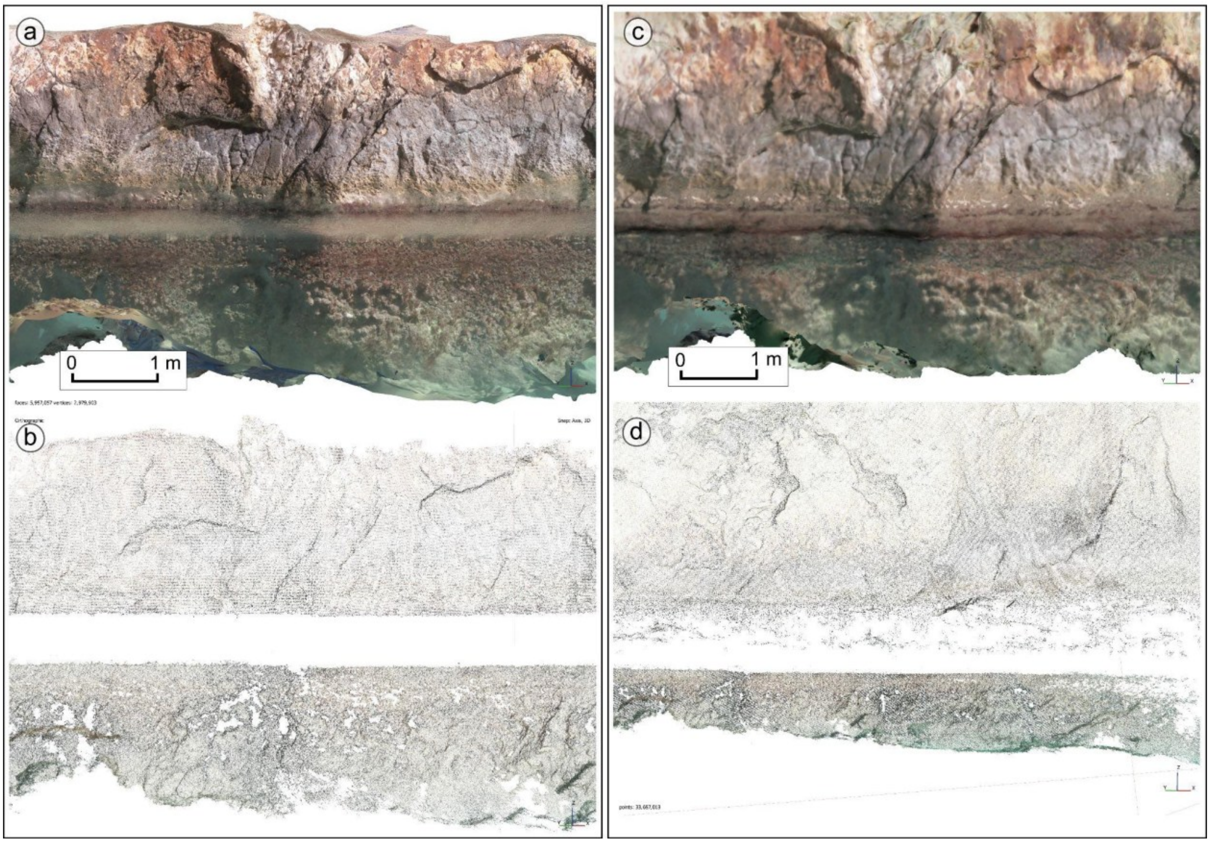

Figure 8.4: Comparison between the 3D models of Halq it-Tafal obtained from GoPro images and aerial photographs. a) and b) zoom in on the 3D model and the dense cloud created by GoPro images; c) and d) zoom in on the 3D model and the dense cloud created by both GoPro images and aerial photographs.

The models highlight several geological features, as shown in Figure 7.29. First, the detailed representation of the morphology of the tidal notch (Antonioli et al., 2015), with measurements of depth (~0.6 m) and vertical extent (~0.5 m above mean sea level). Second, the fracture patterns are shown, with identifiable orientations (NW-SE and NE-SW) and spacings between 0.5 and 1.5 m.

Extraction and analysis of topographic profiles using Cloud Compare allowed detailed characterization of the tidal notch morphology (Figure 7.29) and the interpretation and analysis of the distance mapping (Figure 7.30). The profiles revealed consistent structural patterns across the study area, with the roof notch exhibiting remarkable regularity in its dimensional characteristics. This regularity suggests uniform erosional processes along this stretch of coastline, providing valuable insight into long-term coastal evolutionary patterns (Figure 7.29). This quantitative assessment revealed the distinctive morphology of the roof notch, characterized by its overhanging profile and variable inward penetration, as illustrated by the transition from blue (minimum distance) to red/orange areas (maximum distance from the reference plane) (Figure 7.30).

The slight dome effect affected the interpretation of the distance mapping. In fact, while the standard methodology associates red areas with maximum distances from the reference plane, the dome effect at Halq it-Tafal shifted the areas of greatest interest toward the green spectrum rather than the red. The most distal areas of the curve, shown in red, did not correspond to the position of the tidal notch as would have been expected in a more linear formation, forcing to focus on the central green areas where the actual notch was located (Figure 7.30a, b). By modifying the color scale of the histogram to isolate the green areas, it was possible to isolate the tidal notch (Figure 7.30c) and collect twenty random depth measurements that formed the basis of the quantitative analysis. The considerable variation in depth measurements, from a minimum of 0.14 m to a maximum of 0.95 m, indicates much greater variability than that observed at Mġarr ix-Xini. This variability is further reflected in the standard deviation of 0.232 m, which is approximately 49% of the mean (0.47 m), indicating significant heterogeneity in the erosion processes (Table 7.9). The final measured value of 0.47 ± 0.052 m for the depth of the notch provides a quantitative reference that contrasts with the characteristics at Mġarr ix-Xini (1.26 ± 0.060 m). This substantial difference in cavity depth between the two sites suggests different intensities of erosion processes in the two areas. The relatively high standard error relative to the mean depth at Halq it-Tafal highlights the intrinsic variability of the coastline.

The batter notch values analyzed and measured in this study are comparable to those in Table 2 of Antonioli et al. (2015). Additionally, Figure 6 (Antonioli et al., 2015) confirms that the Malta tidal notch is of the roof type.

The assessment of position accuracy through comparison with satellite images (Image ©2024 Airbus, acquired June 30, 2023) revealed minor systematic distortions in the model-derived coastline (Figure 8.5). The observed "doming" effect, characteristic of vertical wall documentation, resulted in a displacement of approximately 2 meters from the reference coastline. This shift, while significant for absolute positioning purposes, remained consistent throughout the model, suggesting a systematic rather than random error distribution. The distortion is primarily due to the operational constraints of maintaining safe survey distances during data collection, rather than inherent limitations of the modeling methodology.

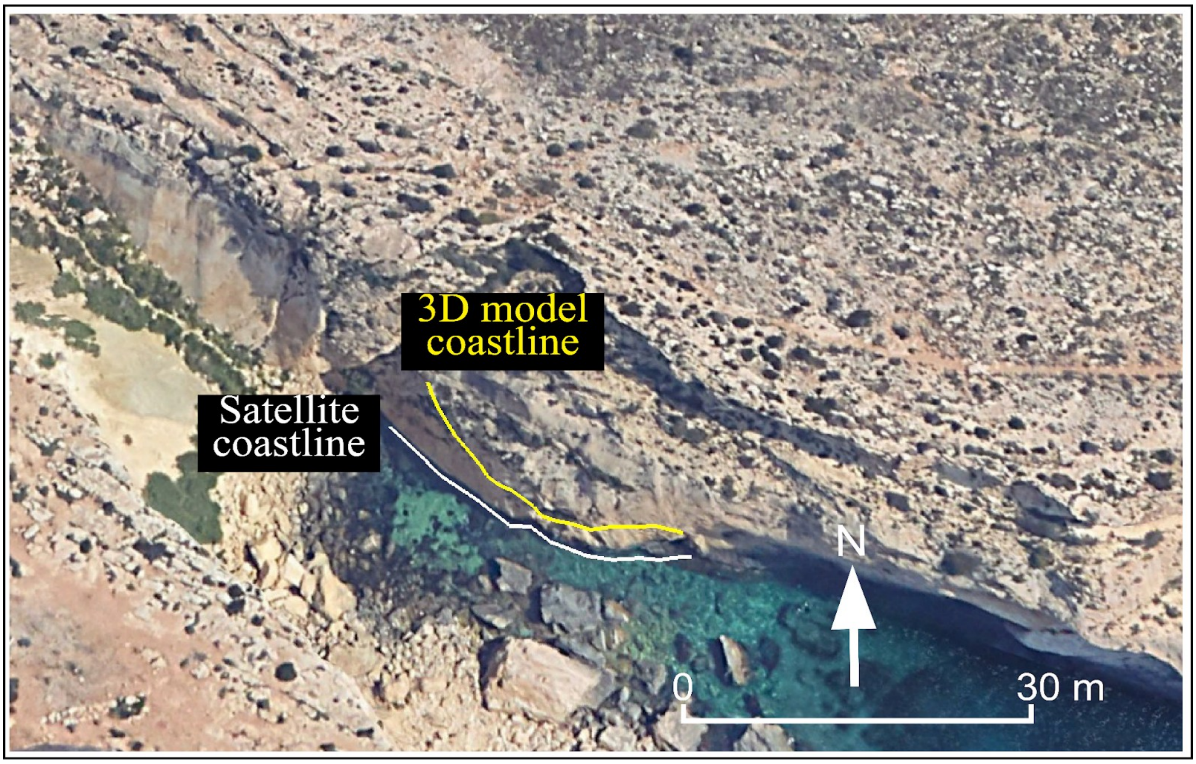

Figure 8.5: Google Earth Image ©2024 Airbus (acquisition 30th, June 2023) view of the plunging cliff at Halq it-Tafal. The GE line-drawing coastlines are represented by a white polyline, while the line-drawing of 3D model is shown in yellow.

8.5. Limitations and future developments

The proposed approach faces several significant limitations that affect its current implementation and validation. The primary constraints are associated with the data acquisition phase, where synchronization between above and below water cameras remains problematic. This technical limitation currently requires operator intervention to control the start and stop of each acquisition phase, as cameras often do not turn on or off simultaneously, batteries may have different lifespans, or technical problems may occur during acquisition. If images are not correctly synchronized, it becomes complicated to study the coastal zone at both interfaces or to adequately reconstruct 3D models. A further limitation lies in the method's dependence on weather and sea conditions.

Before selecting these three case studies, tests were carried out in other surveyed areas, always using images from the same dataset. The other tests were discarded and are not presented in this thesis for various reasons. For instance, in the Cala Calandra (Lampedusa) test model, 185 photos were taken above and 185 below the water. The images depict a sloping coastline with tidal notches and a Lithophyllum platform. However, problems arose during the alignment phase. In the Capo Mulak - Grottacce (Lampedusa) test model, 160 images were taken above the water and 160 below. These images present the same morphology as the previous case. One problem is that not all images are georeferenced, which distorts the model.

Ultimately, only tests that produced at least a sufficient result were considered, and images from the dataset that were georeferenced were selected to give the model a spatial dimension.

The package of photos collected during the Lampedusa campaign is very large but not very usable, as many images are not georeferenced or are distant, blurred, or dark. In some cases, an iron bar was used as a benchmark in the absence of GPS, but since it was only positioned once, this was insufficient. In fact, an infinite number of points pass through a straight line, so the model would still not be oriented.

A systematic validation of the results through model comparison was not possible because this method of 3D model reconstruction in these study areas was performed for the first time. During the doctoral research, it was not feasible to return to the field for a second campaign in the same areas to enable comparison with the models or with in situ tidal notch measurements, due to both logistical constraints and cost limitations. Additionally, the multiclass convolutional neural network requires significant improvement to achieve reliable geomorphological classification, as the current model has limited training data and accuracy.

Despite these current limitations, the method offers substantial potential for future coastal geomorphological studies. These limitations could be addressed by integrating other photogrammetric acquisition techniques such as UAV deployment and, where necessary, LiDAR technology, creating a more robust multi-platform approach for coastal monitoring.

The multiple image visualization system (Geoswim Image Viewer) can be utilized for any type of images that need to be observed simultaneously on screen, with the map showing image locations remaining optional and adaptable to specific research needs.

The binary convolutional neural network development represents an excellent future tool for rapidly eliminating all images that fall outside the scope of surveys, streamlining the data processing workflow.

The development of the multiclass convolutional neural network presents interesting future geomorphological implications. By conducting additional geomorphological surveys, many new images would be acquired that the network has never processed. Using these new images for network training could achieve greater accuracy, and the model would be able to discriminate more easily between classes, rapidly providing numerical responses about coastal types (how many photos correspond to one class rather than another) and giving statistical value to the survey conducted.

Data collection at different time intervals, such as annual surveys, would enable systematic comparison of results from 3D models in future applications. This approach allows direct observation of coastal changes on screen and facilitates comparative measurements. These measurements could provide estimates of coastal erosion rates or assess hazard levels in specific areas. The method could be applied to areas experiencing rapid erosion processes, monitor sudden cliff collapses that pose risks to coastal populations and beach users, or work alongside drone surveys to track storm-deposited boulders along coastlines or on the nearshore seabed. While 3D models remain in development, this approach was designed to provide automated tools for environmental mapping and monitoring with broad applications across various geomorphological studies.

9. Conclusions

This research is focused on the acquisition and manipulation of images collected from swim surveys for the construction of three-dimensional models of rocky coasts above and below sea level, integrating field surveys and Deep learning techniques to develop automatic mapping and coastal monitoring tools. The research produced three main results.

First, the development of an image Viewer for the simultaneous visualization of images above and below the waterline. This platform integrates a geographic map to provide a spatial context for image analysis while facilitating efficient data sharing and collaborative research opportunities. Second, the implementation of CNN achieved an accuracy of 98.85% in the binary classification of coastal images and 80.6% in the identification of five coastal morphotypes. The automated identification of extraneous elements is aimed to image pre-processing, particularly where manual sorting would be prohibitively time-consuming.

Third, the refinement of three-dimensional modeling techniques to integrate above and below water photogrammetry enabled the creation of 3D models of rocky coasts focused in the intertidal zone.

The case studies of Cala Uccello (Lampedusa), Mġarr ix-Xini (Gozo), and Halq it-Tafal (Malta) allowed us to test the acquisition method and the post-processing method using images acquired through the dual camera system. The integration of images above and below the waterline enabled resolutions of up to 0.832 mm/pixel to be achieved, allowing the study of tidal notches. These results demonstrate that low-cost photogrammetric techniques can produce high-quality data for the study of coastal evolution.

For future improvements, the CNN framework should be further refined through extensive testing with new images, and post-processing thresholds should be implemented to optimize classification accuracy. The capabilities of the system should be extended to the detection of specific geomorphological features, such as tidal notches, red algal terraces, or collapsed blocks, through extensive image labelling and retraining of the network.

Technical improvements in image processing and three-dimensional reconstruction are also needed. The blue cast of underwater images requires systematic correction through advanced filtering techniques to improve colour accuracy and feature recognition. Although 360-degree cameras have shown promise in producing high-resolution 3D models, their lack of GPS integration currently limits spatial capture capabilities, a challenge that must be addressed through alternative positioning solutions or hardware modifications. In addition, the quality of the final models, likely due to the camera model used, does not exceed that of models created using GoPro images alone.

The three-dimensional models produced by this research have shown that using CloudCompare software, it was possible to quantitatively analyze the models by observing the topographic profiles of the tidal notch and performing a statistical analysis of their depths and widths. Future work should focus on developing specialized analysis tools and workflows to extract more accurate quantitative measurements and track temporal changes in coastal features. Through the analysis of the three case studies, several critical factors were identified that significantly affect the quality and accuracy of three-dimensional coastal models: a) image quality emerges as a fundamental requirement for the generation of accurate models, with focus to the elimination of blur, haze, and other visual distortions that can compromise feature detection and matching processes; b) the acquisition of accurate GNSS data for above waterline images is essential for establishing the correct model orientation and scale, in order to preserve the spatial accuracy of the final reconstructions; c) environmental conditions play a key role in the success of data acquisition, in which weather and sea conditions have a direct impact on data quality, while lighting conditions and water turbidity have a significant impact on the clarity of underwater images; d) the presence of reflective or homogeneous surfaces presents challenges during image processing, potentially compromising feature matching algorithms and resulting model accuracy.

A comparative analysis of different data acquisition methods reveals distinct advantages and limitations: a) GoPro images consistently provides the highest resolution models, establishing itself as the primary data source for detailed coastal documentation; b) the integration of 360-degree camera data, while increasing spatial coverage, generally results in reduced overall model resolution; c) UAV data integration significantly improves the above water coverage but has limitations in capturing fine intertidal detail.

This methodology's integration of multiple data collection techniques, combined with its systematic approach to survey execution, provides a comprehensive basis for detailed geomorphological analysis and long-term assessment of coastal change. The use of Agisoft Metashape Professional software facilitates automated resolution of merging issues and distortion correction, improving the reproducibility of the method. Analysis using CloudCompare enabled the extraction of precisely positioned topographic profiles that captured the continuous transition between emergent and submerged coastal features.

The data extracted from the profiles allowed for quantitative measurements of coastal morphology along the intertidal zone. The 3D models were also used to extract cross-sections useful for testing the possibility of measuring the depth and width of the tidal notch. In one case study, the total volume eroded near the notch could be calculated. While the values obtained are statistical and not absolute, they can serve as a good starting point for understanding the potential applications of 3D models in studying coastal evolution.

This research demonstrated the potential of low-cost photogrammetric techniques for coastal studies. However, key limitations include camera synchronization issues and the lack of systematic validation, as this is the first time these techniques have been applied to these study areas. These constraints, however, indicate clear pathways for future development, such as integration with UAV and LiDAR technologies, expansion of CNN training datasets, and broader applications beyond tidal notch documentation.

The thesis suggests applying artificial intelligence techniques to coastal geomorphology, a field where these tools remain largely unexplored. However, its contribution to advancing knowledge is limited due to the absence of a systematic comparison with existing approaches. This work represents a first step toward more structured future developments, particularly in monitoring and managing rocky coasts.

Though 3D model application is still experimental, developing this low-cost, high-resolution approach provides automated tools adaptable to diverse Mediterranean coastal environments and geomorphological research applications. Automating morphological analysis using SfM techniques and artificial intelligence could reduce operating times and costs, making the method useful for practical applications in research, planning, and environmental protection while offering automatic coastal mapping and monitoring capabilities.

Abbreviations

Adagrad (Adaptive Gradient Algorithm)

AI (Artificial Intelligence)

ANN (Artificial Neural Networks)

aw (above waterline)

BGD (Batch Gradient Descent)

bw (below waterline)

CNNs (Convolutional Neural Networks)

CTD (Conductivity Temperature Depth)

DL (Deep Learning)

FoV (Field of View)

GCPs (Ground Control Points)

GNSS (Global Navigation Satellite System)

Grad-CAM (Gradient - weighted class activation map)

IBM (Image-based Modeling)

IoT (Internet of Things)

ISR (Instrumental-supported Raft)

MBES (MultiBeam EchoSounder)

RAdam (Rectified Adam)

ReLU (Rectified Linear Function)

RMSProp (Root Mean Squared Propagation)

ROV (Remote Operated Vehicle)

SfM (Structure from Motion)

SGD (Stochastic Gradient Descent)

UAV (Uncrewed Aerial Veicle)

XAI (Explainable Artificial Intelligence)

Bibliography

Abadie, A., Boissery, P., & Viala, C. (2018). Georeferenced underwater photogrammetry to map marine habitats and submerged artificial structures. *The Photogrammetric Record*, *33*(164), 448–469. https://doi.org/10.1111/phor.12263

Agrafiotis, P., & Georgopoulos, A. (2015). Camera constant in the case of two media photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*, 1–6. https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-5-W5/1/2015/isprsarchives-XL-5-W5-1-2015.pdf

Akkaynak, D., & Treibitz, T. (2019). Sea-thru: A method for removing water from underwater images. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1682–1691. http://openaccess.thecvf.com/content_CVPR_2019/html/Akkaynak_Sea-Thru_A_Method_for_Removing_Water_From_Underwater_Images_CVPR_2019_paper.html

Alexander, D. (1988). A review of the physical geography of Malta and its significance for tectonic geomorphology. *Quaternary Science Reviews*, *7*(1), 41–53. https://www.sciencedirect.com/science/article/pii/0277379188900923

Alho, P., Kukko, A., Hyyppä, H., Kaartinen, H., Hyyppä, J., & Jaakkola, A. (2009). Application of boat-based laser scanning for river survey. *Earth Surface Processes and Landforms*, *34*(13), 1831–1838. https://doi.org/10.1002/esp.1879

Angelov, P. P., Soares, E. A., Jiang, R., Arnold, N. I., & Atkinson, P. M. (2021). Explainable artificial intelligence: An analytical review. *WIREs Data Mining and Knowledge Discovery*, *11*(5), e1424. https://doi.org/10.1002/widm.1424

Antonioli, F., Ferranti, L., Lambeck, K., Kershaw, S., Verrubbi, V., & Dai Pra, G. (2006). Late Pleistocene to Holocene record of changing uplift rates in southern Calabria and northeastern Sicily (southern Italy, Central Mediterranean Sea). *Tectonophysics*, *422*(1–4), 23–40.

Antonioli, F., Anzidei, M., Lambeck, K., Auriemma, R., Gaddi, D., Furlani, S., Orrù, P., Solinas, E., Gaspari, A., & Karinja, S. (2007). Sea-level change during the Holocene in Sardinia and in the northeastern Adriatic (central Mediterranean Sea) from archaeological and geomorphological data. *Quaternary Science Reviews*, *26*(19–21), 2463–2486.

Antonioli, F., Ferranti, L., Fontana, A., Amorosi, A., Bondesan, A., Braitenberg, C., Dutton, A., Fontolan, G., Furlani, S., & Lambeck, K. (2009). Holocene relative sea-level changes and vertical movements along the Italian and Istrian coastlines. *Quaternary International*, *206*(1–2), 102–133. https://www.sciencedirect.com/science/article/pii/S1040618208003388?casa_token=CqMOePCtd0EAAAAA:enO6gxTMItESIkn0BmffLsteFpcynZMWUSSWWMNu5Hp78TBWE4oQfcwHfqZ3GqHLtY_CaGG59A

Antonioli, F., Silenzi, S., & Frisia, S. (2001). Tyrrhenian Holocene palaeoclimate trends from spelean serpulids. *Quaternary Science Reviews*, *20*(15), 1661–1670.

Antonioli, F., Lo Presti, V., Rovere, A., Ferranti, L., Anzidei, M., Furlani, S., Mastronuzzi, G., Orru, P. E., Scicchitano, G., Sannino, G., Spampinato, C. R., Pagliarulo, R., Deiana, G., de Sabata, E., Sansò, P., Vacchi, M., & Vecchio, A. (2015). Tidal notches in Mediterranean Sea: A comprehensive analysis. *Quaternary Science Reviews*, *119*, 66–84. https://doi.org/10.1016/j.quascirev.2015.03.016

Antonioli, F., Anzidei, M., Lo Presti, V., Scicchitano, G., Spampinato, C. R., Trainito, E., & Furlani, S. (2017). Anomolous multi-origin marine notch sites: Three case studies in the central Mediterranean Sea. *Quaternary International*, *439*, 4–16. https://doi.org/10.1016/j.quaint.2017.02.037

Anzidei, M., Lambeck, K., Antonioli, F., Furlani, S., Mastronuzzi, G., Serpelloni, E., & Vannucci, G. (2014). Coastal structure, sea-level changes and vertical motion of the land in the Mediterranean. *Geological Society, London, Special Publications*, *388*(1), 453–479. https://doi.org/10.1144/SP388.20

Armigliato, A., Tinti, S., Zaniboni, F., Pagnoni, G., & Argnani, A. (2007). New contributions to the debate on the cause of the January 11th, 1693 tsunami in eastern Sicily (Italy): Earthquake or offshore landslide source (or may be both)? *AGU Fall Meeting Abstracts*, *2007*, S53A-1019. https://ui.adsabs.harvard.edu/abs/2007AGUFM.S53A1019A/abstract

Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., & Benjamins, R. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82–115. https://www.sciencedirect.com/science/article/pii/S1566253519308103

Arrighi, L., Barbon Junior, S., Pellegrino, F. A., Simonato, M., & Zullich, M. (2023). Explainable Automated Anomaly Recognition in Failure Analysis: Is Deep Learning Doing it Correctly? In L. Longo (Ed.), *Explainable Artificial Intelligence* (Vol. 1902, pp. 420–432). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-44067-0_22

Astraldi, M., Gasparini, G. P., Gervasio, L., & Salusti, E. (2001). Dense water dynamics along the Strait of Sicily (Mediterranean Sea). *Journal of Physical Oceanography*, *31*(12), 3457–3475. Scopus. https://doi.org/10.1175/1520-0485(2001)031<3457:DWDATS>2.0.CO;2

Astraldi, M., Conversano, F., Civitarese, G., Gasparini, G. P., d'Alcalà, M. R., & Vetrano, A. (2002). Water mass properties and chemical signatures in the central Mediterranean region. *Journal of Marine Systems*, *33*, 155–177.

Avinash, S. V. (2017, March 30). Understanding activation functions in Neural Networks. *The Theory of Everything*. https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PloS One*, *10*(7), e0130140. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0130140

Baldassini, N., & Di Stefano, A. (2017). Stratigraphic features of the Maltese Archipelago: A synthesis. *Natural Hazards*, *86*(S2), 203–231. https://doi.org/10.1007/s11069-016-2334-9

Ballesteros, E. (2006). Mediterranean coralligenous assemblages: A synthesis of present knowledge. *Oceanography and Marine Biology: An Annual Review*, *44*, 123–195.

Barbariol, F., Bidlot, J.-R., Cavaleri, L., Sclavo, M., Thomson, J., & Benetazzo, A. (2019). Maximum wave heights from global model reanalysis. *Progress in Oceanography*, *175*, 139–160.

Barker, R., Dixon, L., & Hooke, J. (1997). Use of terrestrial photogrammetry for monitoring and measuring bank erosion. *Earth Surface Processes and Landforms*, *22*(13), 1217–1227. https://doi.org/10.1002/(SICI)1096-9837(199724)22:13<1217::AID-ESP819>3.0.CO;2-U

Benac, Č., Juračić, M., & Bakran-Petricioli, T. (2004). Submerged tidal notches in the Rijeka Bay NE Adriatic Sea: Indicators of relative sea-level change and of recent tectonic movements. *Marine Geology*, *212*(1), 21–33. https://doi.org/10.1016/j.margeo.2004.09.002

Benac, Č., Juračić, M., & Blašković, I. (2008). Tidal notches in Vinodol Channel and Bakar Bay, NE Adriatic Sea: Indicators of recent tectonics. *Marine Geology*, *248*(3), 151–160. https://doi.org/10.1016/j.margeo.2007.10.010

Bengio, Y., Goodfellow, I., & Courville, A. (2017). Deep learning (Vol. 1). MIT press Cambridge, MA, USA. https://www.academia.edu/download/62266271/Deep_Learning20200303-80130-1s42zvt.pdf

Bennett, G. L., Molnar, P., Eisenbeiss, H., & McArdell, B. W. (2012). Photogrammetric analysis of slope failures feeding the head of the Illgraben debris flow channel. *EGU General Assembly Conference Abstracts*, 10393. https://ui.adsabs.harvard.edu/abs/2012EGUGA..1410393B/abstract

Bergsma, E. W. J., Conley, D. C., Davidson, M. A., & O'Hare, T. J. (2016). Video-based nearshore bathymetry estimation in macro-tidal environments. *Marine Geology*, *374*, 31–41. https://doi.org/10.1016/j.margeo.2016.02.001

Biolchi, S., Furlani, S., Devoto, S., Gauci, R., Castaldini, D., & Soldati, M. (2016a). Geomorphological identification, classification and spatial distribution of coastal landforms of Malta (Mediterranean Sea). *Journal of Maps*, *12*(1), 87–99. https://doi.org/10.1080/17445647.2014.984001

Biolchi, S., Furlani, S., Covelli, S., Busetti, M., & Cucchi, F. (2016b). Morphoneotectonics and lithology of the eastern sector of the Gulf of Trieste (NE Italy). *Journal of Maps*, *12*(5), 936–946. https://doi.org/10.1080/17445647.2015.1099572

Bird, E. C. F. (2008). *Coastal Geomorphology: An Introduction*. John Wiley & Sons.

Bishop, C., & Nasrabadi, N. (2006). Pattern recognition and machine learning (Vol. 4). Springer. https://link.springer.com/book/9780387310732

Blais, M.-A., & Akhloufi, M. A. (2025). Advances in Remote Sensing and Deep Learning in Coastal Boundary Extraction for Erosion Monitoring. *Geomatics*, *5*(1), 9. https://doi.org/10.3390/geomatics5010009

Bonacci, O., Ljubenkov, I., & Roje-Bonacci, T. (2006). Karst flash floods: An example from the Dinaric karst (Croatia). *Natural Hazards and Earth System Sciences*, *6*(2), 195–203. https://doi.org/10.5194/nhess-6-195-2006

Bonada, N., & Resh, V. H. (2013). Mediterranean-climate streams and rivers: Geographically separated but ecologically comparable freshwater systems. *Hydrobiologia*, *719*(1), 1–29. https://doi.org/10.1007/s10750-013-1634-2

Borzì, L., Scala, P., Distefano, S., Laksono, F. X. A. T., Manno, G., Innangi, S., Gamberi, F., Kovács, J., Ciraolo, G., & Di Stefano, A. (2024). Tsunami propagation and flooding maps: An application for the Island of Lampedusa, Sicily Channel, Italy. *Earth Surface Processes and Landforms*, *49*(14), 4842–4861. https://doi.org/10.1002/esp.5996

Brasington, J., Rumsby, B. T., & McVey, R. A. (2000). Monitoring and modelling morphological change in a braided gravel-bed river using high resolution GPS-based survey. *Earth Surface Processes and Landforms*, *25*(9), 973–990. https://doi.org/10.1002/1096-9837(200008)25:9<973::AID-ESP111>3.0.CO;2-Y

Brose, S. M., Rettelbach, T., Heidler, K., & Grosse, G. (2025). Leveraging multi-decadal, multi-modal high-resolution optical imagery towards automated Arctic coastal erosion monitoring with deep learning. *Geomorphology*, *487*, 109914. https://doi.org/10.1016/j.geomorph.2025.109914

Buscombe, D., & Ritchie, A. C. (2018). Landscape classification with deep neural networks. *Geosciences,* 8(7), 244. https://www.mdpi.com/2076-3263/8/7/244

Calais, E., DeMets, C., & Nocquet, J.-M. (2003). Evidence for a post-3.16-Ma change in Nubia–Eurasia–North America plate motions? *Earth and Planetary Science Letters*, *216*(1), 81–92. https://doi.org/10.1016/S0012-821X(03)00482-5

Calkoen, F. R., Luijendijk, A. P., Hanson, S., Nicholls, R. J., Moreno-Rodenas, A., De Heer, H., & Baart, F. (2025). Mapping the world's coast: A global 100-m coastal typology derived from satellite data using deep learning. *Earth System Science Data Discussions*, 1–37. https://doi.org/10.5194/essd-2025-388

Capra, A. (1993). Non-conventional system in underwater photogrammetry. *International Archives of Photogrammetry and Remote Sensing*, *29*, 234–234. https://www.isprs.org/proceedings/xxix/congress/part5/234_XXIX-part5.pdf

Carobene, L. (1972). *Osservazioni sui solchi di battente attuali ed antichi nel Golfo di Orosei in Sardegna*.

Carobene, L. (2014). Marine Notches and Sea-Cave Bioerosional Grooves in Microtidal Areas: Examples from the Tyrrhenian and Ligurian Coasts—Italy. *Journal of Coastal Research*, *31*(3), 536–556. https://doi.org/10.2112/JCOASTRES-D-14-00068.1

Carrivick, J. L., Smith, M. W., & Quincey, D. J. (2016). Structure from Motion in the Geosciences (1st ed.). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118895818

Cas, R. A. F., & Wright, J. V. (1987). Volcanic successions: Ancient and modern. *Allen and Unin, London*.

Catalano, S., Monaco, C., Tortorici, L., & Tansi, C. (1993). Pleistocene strike-slip tectonics in the Lucanian Apennine (southern Italy). *Tectonics*, *12*(3), 656–665. https://doi.org/10.1029/92TC02251

Cavaleri, L. (2000). The oceanographic tower *Acqua Alta*—Activity and prediction of sea states at Venice. *Coastal Engineering*, *39*(1), 29–70. https://doi.org/10.1016/S0378-3839(99)00053-8

Cavaleri, L., & Bertotti, L. (2004). Accuracy of the modelled wind and wave fields in enclosed seas. *Tellus A: Dynamic Meteorology and Oceanography*, *56*(2), 167. https://doi.org/10.3402/tellusa.v56i2.14398

Chandler, J. (1999). Effective application of automated digital photogrammetry for geomorphological research. *Earth Surface Processes and Landforms*, *24*(1), 51–63. https://doi.org/10.1002/(SICI)1096-9837(199901)24:1<51::AID-ESP948>3.0.CO;2-H

Chattopadhay, A., Sarkar, A., Howlader, P., & Balasubramanian, V. N. (2018). Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. *2018 IEEE Winter Conference on*

Applications of Computer Vision (WACV), 839–847. https://ieeexplore.ieee.org/abstract/document/8354201/

Chemello, R. (1989). La bionomia bentonica ed i molluschi: 5. *Il Piano Infralitorale: Il Marciapiede a Vermeti. Not. SIM*, *7*, 167–170.

Chemello, R., Pandolfo, A., & Riggio, S. (1990). Le biocostruzioni a Molluschi Vermetidi nella Sicilia nord-occidentale. *Atti 53 Congresso UZI*.

Chemello, R. (2009). Le biocostruzioni marine in Mediterraneo. Lo stato delle conoscenze sui reef a vermeti. *Biologia Marina Mediterranea*, *16*(1), 2–18. https://iris.unipa.it/retrieve/handle/10447/74688/76055/Vermeti%20Chemello%202009%20BiolMarMedit.pdf

Chemello, R., & Silenzi, S. (2011). Vermetid reefs in the Mediterranean Sea as archives of sea-level and surface temperature changes. *Chemistry and Ecology*. https://www.tandfonline.com/doi/abs/10.1080/02757540.2011.554405

Chen, X., Chen, H. Y. H., Chen, C., Ma, Z., Searle, E. B., Yu, Z., & Huang, Z. (2019). Effects of plant diversity on soil carbon in diverse ecosystems: A global meta-analysis. *Biological Reviews*, *95*(1), 167–183. https://doi.org/10.1111/brv.12554

Cheng, G., Yang, J., Gao, D., Guo, L., & Han, J. (2020). High-quality proposals for weakly supervised object detection. *IEEE Transactions on Image Processing*, *29*, 5794–5804. https://ieeexplore.ieee.org/abstract/document/9069411/

Civile, D., Lodolo, E., Accettella, D., Geletti, R., Ben-Avraham, Z., Deponte, M., Facchin, L., Ramella, R., & Romeo, R. (2010). The Pantelleria graben (Sicily Channel, Central Mediterranean): An example of intraplate 'passive' rift. *Tectonophysics*, *490*(3), 173–183. https://doi.org/10.1016/j.tecto.2010.05.008

Civile, D., Brancolini, G., Lodolo, E., Forlin, E., Accaino, F., Zecchin, M., & Brancatelli, G. (2021). Morphostructural setting and tectonic evolution of the central part of the Sicilian Channel (central Mediterranean). *Lithosphere*, *2021*(1), 7866771. https://pubs.geoscienceworld.org/gsa/lithosphere/article-abstract/2021/1/7866771/594452

Civita, M. (2008). *L'assetto idrogeologico del territorio italiano: Risorse e problematiche*. https://iris.polito.it/handle/11583/1718667

Coombes, M. A. (2014). The rock coast of the British Isles: Weathering and biogenic processes. In D. M. Kennedy, W. J. Stephenson, & L. A. Naylor (Eds.), *Rock Coast Geomorphology: A Global Synthesis* (Vol. 40, p. 0). Geological Society of London. https://doi.org/10.1144/M40.5

Cooper, F. J., Roberts, G. P., & Underwood, C. J. (2007). A comparison of 103-105 year uplift rates on the South Alkyonides Fault, Central Greece: Holocene climate stability and the formation of coastal notches. *Geophysical Research Letters*, *34*(14). Scopus. https://doi.org/10.1029/2007GL030673

Corradetti, A., Seers, T., Mercuri, M., Calligaris, C., Busetti, A., & Zini, L. (2022). Benchmarking different SfM-MVS photogrammetric and iOS LiDAR acquisition methods for the digital preservation of a short-lived excavation: A case study from an area of sinkhole related subsidence. *Remote Sensing*, *14*, 5187. https://doi.org/10.3390/rs14205187

Cover, T. M. (1999). Elements of information theory. John Wiley & Sons.

CrossEntropyLoss—PyTorch 2.5 documentation. (n.d.). Retrieved 14 January 204, from https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html

Dabbura, I. (2017). Gradient descent algorithm and its variants. *Towards Data Science*, 15.

Dart, C. J., Bosence, D. W. J., & McCLAY, K. R. (1993). Stratigraphy and structure of the Maltese graben system. *Journal of the Geological Society*, *150*(6), 1153–1166. https://doi.org/10.1144/gsjgs.150.6.1153

De Leo, F., Besio, G., & Mentaschi, L. (2021). Trends and variability of ocean waves under RCP8.5 emission scenario in the Mediterranean Sea. *Ocean Dynamics*, *71*(1), 97–117. https://doi.org/10.1007/s10236-020-01419-8

Devescovi, M., & Iveša, L. (2007). Short term impact of planktonic mucilage aggregates on macrobenthos along the Istrian rocky coast (Northern Adriatic, Croatia). *Marine Pollution Bulletin*, *54*(7), 887–893.

Devoto, S., Biolchi, S., Bruschi, V. M., Furlani, S., Mantovani, M., Piacentini, D., Pasuto, A., & Soldati, M. (2012). Geomorphological map of the NW Coast of the Island of Malta (Mediterranean Sea). *Journal of Maps*, *8*(1), 33–40. https://doi.org/10.1080/17445647.2012.668425

Devoto, S., Biolchi, S., Bruschi, V. M., Díez, A. G., Mantovani, M., Pasuto, A., Piacentini, D., Schembri, J. A., & Soldati, M. (2013). Landslides Along the North-West Coast of the Island of Malta. In C. Margottini, P. Canuti, & K. Sassa (Eds.), *Landslide Science and Practice* (pp. 57–63). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31325-7_7

Devoto, S., Macovaz, V., Mantovani, M., Soldati, M., & Furlani, S. (2020). Advantages of using UAV digital photogrammetry in the study of slow-moving coastal landslides. *Remote Sensing*, *12*(21), 3566. https://www.mdpi.com/2072-4292/12/21/3566

Devoto, S., Hastewell, L. J., Prampolini, M., & Furlani, S. (2021). Dataset of gravity-induced landforms and sinkholes of the Northeast Coast of Malta (Central Mediterranean Sea). *Data*, *6*(8), 81. https://www.mdpi.com/2306-5729/6/8/81

De Waele, J., & Furlani, S. (2013). *Seawater and Biokarst Effects on Coastal Limestones*. USA. https://doi.org/10.1016/B978-0-12-374739-6.00109-3

Dickson, M. E., & Stephenson, W. J. (2014). The rock coast of New Zealand. In D. M. Kennedy, W. J. Stephenson, & L. A. Naylor (Eds.), *Rock Coast Geomorphology: A Global Synthesis* (Vol. 40, p. 0). Geological Society of London. https://doi.org/10.1144/M40.13

Directorate, O. E. (1993). Geological map of the Maltese Islands. *Office of the Prime Minister: Valletta, Malta*.

Draelos, R. L., & Carin, L. (2020). Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks. *arXiv Preprint arXiv:2011.08891*. https://arxiv.org/abs/2011.08891

Drago, A. (2009). Sea level variability and the 'Milghuba' seiche oscillations in the northern coast of Malta, Central Mediterranean. *Physics and Chemistry of the Earth, Parts A/B/C*, *34*(17), 948–970. https://doi.org/10.1016/j.pce.2009.10.002

Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, *12*(7). https://www.jmlr.org/papers/volume12/duchi11a/duchi11a.pdf

Durán, J. (2019). Everything you need to know about gradient descent applied to neural networks. Medium. https://medium.com/yottabytes/everything-you-need-to-know-about-gradient-descent-applied-to-neural-networks-d70f85e0cc14

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., & Ranjan, R. (2023). Explainable AI (XAI): Core Ideas, Techniques, and Solutions. *ACM Computing Surveys*, *55*(9), 1–33. https://doi.org/10.1145/3561048

Eltner, A., Kaiser, A., Castillo, C., Rock, G., Neugirg, F., & Abellán, A. (2016). Image-based surface reconstruction in geomorphometry – merits, limits and developments. *Earth Surface Dynamics*, *4*(2), 359–389. https://doi.org/10.5194/esurf-4-359-2016

Enea Station for climate observations. (2023). https://www.lampedusa.enea.it/dataaccess/ocean_meteo/index.php?lang=it+%28accessed+on+29+February+2024%29.&step=2&from=2020-09-23&to=2020-09-24&WS=&WD=&vars=Wind+Speed+%28m%2Fs%29%2C+Wind+Direction+%28%C2%B0%29&ok=&name=Valeria+Vaccher&mail=valeria.vaccher%40phd.units.it&office=University+of+Trieste

Esposito, G., Salvini, R., Matano, F., Sacchi, M., Danzi, M., Somma, R., & Troise, C. (2017). Multitemporal monitoring of a coastal landslide through SfM-derived point cloud comparison. *The Photogrammetric Record*, *32*(160), 459–479. https://doi.org/10.1111/phor.12218

Evelpidou, N., Kampolis, I., Pirazzoli, P. A., & Vassilopoulos, A. (2012). Global sea-level rise and the disappearance of tidal notches. *Global and Planetary Change*, *92–93*, 248–256. https://doi.org/10.1016/j.gloplacha.2012.05.013

Fagerström, V. (2018). Structure from Motion, a Cheaper Alternative for Three-Dimensional Modeling in Earth Science. https://www.diva-portal.org/smash/record.jsf?pid=diva2:1213300

Faivre, S., Fouache, E., Ghilardi, M., Antonioli, F., Furlani, S., & Kovačić, V. (2011). Relative sea level change in western Istria (Croatia) during the last millennium. *Quaternary International*, *232*(1–2), 132–143. Scopus. https://doi.org/10.1016/j.quaint.2010.05.027

Faivre, S., Bakran-Petricioli, T., Horvatinčić, N., & Sironić, A. (2013). Distinct phases of relative sea level changes in the central Adriatic during the last 1500years—Influence of climatic variations? *Palaeogeography, Palaeoclimatology, Palaeoecology*, *369*, 163–174. Scopus. https://doi.org/10.1016/j.palaeo.2012.10.016

Fawcett, D., Blanco-Sacristán, J., & Benaud, P. (2019). Two decades of digital photogrammetry: Revisiting Chandler's 1999 paper on "Effective application of automated digital photogrammetry for geomorphological research" – a synthesis. *Progress in Physical Geography: Earth and Environment*, *43*(2), 299–312. https://doi.org/10.1177/0309133319832863

Ferranti, L., Antonioli, F., Mauz, B., Amorosi, A., Dai Pra, G., Mastronuzzi, G., Monaco, C., Orrù, P., Pappalardo, M., & Radtke, U. (2006). Markers of the last interglacial sea-level high stand along the coast of Italy: Tectonic implications. *Quaternary International*, *145*, 30–54.

https://www.sciencedirect.com/science/article/pii/S1040618205001138?casa_token=Vfl5JLiBq_YAAAAA:paPs-bKMd6egXQ2y4dyy08IvKHie1YgIUWjYWvryadED1bryGgVCz6HEG4kuQvGVDZDTeBk9kA

Ferranti, L., Antonioli, F., Anzidei, M., Monaco, C., & Stocchi, P. (2010). The timescale and spatial extent of vertical tectonic motions in Italy: Insights from relative sea-level changes studies. *Journal of the Virtual Explorer*, *36*(Paper 30). https://www.researchgate.net/profile/Luigi-Ferranti/publication/48329978_The_timescale_and_spatial_extent_of_vertical_tectonic_motions_in_Italy_Insights_from_relative_sea-level_changes_studies/links/0fcfd50f4219c3ad99000000/The-timescale-and-spatial-extent-of-vertical-tectonic-motions-in-Italy-Insights-from-relative-sea-level-changes-studies.pdf?_sg%5B0%5D=started_experiment_milestone&origin=journalDetail

Finetti, I. (1984). Geophysical study of the Sicily Channel rift zone. https://ricerca.unityfvg.it/entities/publication/5de37f13-805b-4c67-ac03-815187929b99/details

Fix, R. E., & Burt, T. P. (1995). Global positioning system: An effective way to map a small area or catchment. *Earth Surface Processes and Landforms*, *20*(9), 817–827. https://doi.org/10.1002/esp.3290200907

Ford, D., & Williams, P. D. (2007). *Karst hydrogeology and geomorphology*. John Wiley & Sons. https://books.google.it/books?hl=en&lr=&id=E-J9EAAAQBAJ&oi=fnd&pg=PR11&dq=Ford,+D.,+%26+Williams,+P.+(2007).+Karst+Hydrogeology+and+Geomorphology.+John+Wiley+%26+Sons.&ots=sQMpNtvxFl&sig=RLi2nIkh_SPitKtaN2tJWqNbFkI

Franco, F. (2023, August 15). Linear Regression Algorithm. *Medium*. https://medium.com/@francescofranco_39234/linear-regression-algorithm-696ea1e6544b

Fraser, C. S., & Cronk, S. (2009). A hybrid measurement approach for close-range photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(3), 328–333. https://doi.org/10.1016/j.isprsjprs.2008.09.009

Fu, R., Hu, Q., Dong, X., Guo, Y., Gao, Y., & Li, B. (2020). *Axiom-based Grad-CAM: Towards Accurate Visualization and Explanation of CNNs* (No. arXiv:2008.02312). arXiv. https://doi.org/10.48550/arXiv.2008.02312

Furlani, S., Cucchi, F., Forti, F., & Rossi, A. (2009). Comparison between coastal and inland Karst limestone lowering rates in the northeastern Adriatic Region (Italy and Croatia). *Geomorphology*, *104*(1), 73–81. https://doi.org/10.1016/j.geomorph.2008.05.015

Furlani, S., Biolchi, S., Cucchi, F., Antonioli, F., Busetti, M., & Melis, R. (2011). Tectonic effects on Late Holocene sea level changes in the Gulf of Trieste (NE Adriatic Sea, Italy). *Quaternary International*, *232*(1–2), 144–157.

Furlani, S. (2012). The Geoswim project: Snorkel-surveying along 250 kilometres of the Southern and Western Istrian Coast. *Alpine and Mediterranean Quaternary*, *25*(2), vii–ix.

Furlani, S., Antonioli, F., Biolchi, S., Gambin, T., Gauci, R., Lo Presti, V., Anzidei, M., Devoto, S., Palombo, M., & Sulli, A. (2013). Holocene sea level change in Malta. *Quaternary International, 288,* 146–157.

Furlani, S., & Cucchi, F. (2013). Downwearing rates of vertical limestone surfaces in the intertidal zone (Gulf of Trieste, Italy). *Marine Geology*, *343*, 92–98. https://doi.org/10.1016/j.margeo.2013.06.005

Furlani, S., Pappalardo, M., Gómez-Pujol, L., & Chelli, A. (2014a). The rock coast of the Mediterranean and Black seas. Geological Society, London, Memoirs 40: 89-123.

Furlani, S., Ninfo, A., Zavagno, E., Paganini, P., Zini, L., Biolchi, S., Antonioli, F., Coren, F., & Cucchi, F. (2014b). Submerged notches in Istria and the Gulf of Trieste: Results from the Geoswim project. *Quaternary International, 332*, 37–47.

Furlani, S., Antonioli, F., Cavallaro, D., Chirco, P., Caldareri, F., Martin, F. F., Morticelli, M. G., Monaco, C., Sulli, A., & Quarta, G. (2017a). Tidal notches, coastal landforms and relative sea-level changes during the Late Quaternary at Ustica Island (Tyrrhenian Sea, Italy). *Geomorphology*, *299*, 94–106.

Furlani, S., Antonioli, F., Gambin, T., Gauci, R., Ninfo, A., Zavagno, E., Micallef, A., & Cucchi, F. (2017b). Marine notches in the Maltese islands (central Mediterranean Sea). *Quaternary International*, *439*, 158–168.

Furlani, S., Antonioli, F., Gambin, T., Biolchi, S., Formosa, S., Lo Presti, V., Mantovani, M., Anzidei, M., Calcagnile, L., & Quarta, G. (2018). Submerged speleothem in Malta indicates tectonic stability throughout the Holocene. *The Holocene*, *28*(10), 1588–1597. https://doi.org/10.1177/0959683618782613

Furlani, S. (2020). Integrating observational targets and instrumental data on rock coasts through snorkel surveys: A methodological approach. *Marine Geology*, *425*, 106191. https://doi.org/10.1016/j.margeo.2020.106191

Furlani, S., Vaccher, V., Macovaz, V., & Devoto, S. (2020). A cost-effective method to reproduce the morphology of the nearshore and intertidal zone in microtidal environments. *Remote Sensing*, *12*(11), Article 11. https://doi.org/10.3390/rs12111880

Furlani, S., Vaccher, V., Antonioli, F., Agate, M., Biolchi, S., Boccali, C., Busetti, A., Caldareri, F., Canziani, F., Chemello, R., Causon Deguara, J., Dal Bo, E., Dean, S., Deiana, G., De Sabata, E., Donno, Y., Gauci, R., Giaccone, T., Lo Presti, V., Montagna, P., Navone, A., Orrù, P.E., Porqueddu, A., Schembri, J.A., Tavani, M., Torricella, F., Trainito, E., Vacchi, M., Venturini, E. (2021). Preservation of Modern and MIS 5.5 Erosional Landforms and Biological Structures as Sea Level Markers: A Matter of Luck? *Water*, *13*(15), Article 15. https://doi.org/10.3390/w13152127

Furlani, S., & Antonioli, F. (2023). The swim-survey archive of the Mediterranean rocky coasts: Potentials and future perspectives. *Geomorphology*, *421*, 108529. https://doi.org/10.1016/j.geomorph.2022.108529

Furlani, S., Agate, M., de Sabata, E., Chemello, R., Vaccher, V., Visconti, G., & Antonioli, F. (2024). Dipping Tidal Notch (DTN): Exposed vs. Sheltered Morphometry. *Geosciences*, *14*(6), Article 6. https://doi.org/10.3390/geosciences14060157

Galea, P. (2007). Seismic history of the Maltese islands and considerations on seismic risk. *Annals of Geophysics*. https://www.earth-prints.org/handle/2122/4360

Garbin, C., Zhu, X., & Marques, O. (2020). Dropout vs. Batch normalization: An empirical study of their impact to deep learning. *Multimedia Tools and Applications*, *79*(19), 12777–12815. https://doi.org/10.1007/s11042-019-08453-9

Girardeau-Montaut, D. (2024). *CloudCompare—3D Point Cloud and Mesh Processing Software*. Retrieved 20 March 2025, from https://www.danielgm.net/cc/

Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, 315–323. https://proceedings.mlr.press/v15/glorot11a.html

Goncalves, J. A., & Henriques, R. (2015). UAV photogrammetry for topographic monitoring of coastal areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *104*, 101–111.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. https://scholar.google.com/scholar_lookup?title=Deep%20Learning&publication_year=2016&author=I.%20Goodfellow&author=Y.%20Bengio&author=A.%20Courville

Graff, P., Feroz, F., Hobson, M. P., & Lasenby, A. (2014). SKYNET: An efficient and robust neural network training tool for machine learning in astronomy. *Monthly Notices of the Royal Astronomical Society*, *441*(2), 1741–1759. https://academic.oup.com/mnras/article-abstract/441/2/1741/1071156

Grasso, M., Pedley, H. M., & Reuther, C. D. (1985). The geology of the Pelagian Islands and their structural setting related to the Pantelleria rift (central Mediterranean Sea). https://www.um.edu.mt/library/oar/handle/123456789/37470

Grasso, M., & Pedley, M. H. (1988). The sedimentology and development of Terravecchia Formation carbonates (Upper Miocene) of North Central Sicily: Possible eustatic influence on facies development. *Sedimentary Geology*, *57*(1), 131–149. https://doi.org/10.1016/0037-0738(88)90022-X

Grüen, A., Remondino, F., & Zhang, L. (2004). 3D modeling and visualization of large cultural heritage sites at very high resolution: The Bamiyan valley and its standing Buddhas. *Beiträge Zum XX. Kongress Der ISPRS in Istanbul*, 29. https://ethz.ch/content/dam/ethz/special-interest/baug/igp/igp-dam/documents/Reports/298.pdf#page=34

Grüen, A. (2012). Development and Status of Image Matching in Photogrammetry. *The Photogrammetric Record*, *27*(137), 36–57. https://doi.org/10.1111/j.1477-9730.2011.00671.x

Guilcher, A. (1953). Essai Sur La Zonation Et La Distribution Des Formes Littorales De Dissolution Du Calcaire. *Annales de Géographie*, *62*(331), 161–179. https://www.jstor.org/stable/23442897

Guo, T., Capra, A., Troyer, M., Grün, A., Brooks, A. J., Hench, J. L., Schmitt, R. J., Holbrook, S. J., & Dubbini, M. (2016). Accuracy assessment of underwater photogrammetric three dimensional modelling for coral reefs. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *41*(B5), 821–828. https://www.research-collection.ethz.ch/handle/20.500.11850/118990

Hansom, J., Forbes, D., & S.Etienne. (2014). Chapter 16 The rock coasts of polar and sub-polar regions. In *Geological Society Memoir* (Vol. 40, pp. 263–281). https://doi.org/10.1144/M40.16

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 770–778. https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., & Li, M. (2019). Bag of tricks for image classification with convolutional neural networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 558–567. http://openaccess.thecvf.com/content_CVPR_2019/html/He_Bag_of_Tricks_for_Image_Classification_with_Convolutional_Neural_Networks_CVPR_2019_paper.html

Heritage, G., & Hetherington, D. (2007). Towards a protocol for laser scanning in fluvial geomorphology. *Earth Surface Processes and Landforms*, *32*(1), 66–74. https://doi.org/10.1002/esp.1375

Hinkelmann, K. (2018). Neural Networks. https://web.archive.org/web/20181006235506/http://didattica.cs.unicam.it/lib/exe/fetch.php?media=didattica:magistrale:kebi:ay_1718:ke-11_neural_networks.pdf

Hirol, S. (2023, September 15). StatusNeo. https://statusneo.com/accelerate-convergence-mini-batch-momentum-in-deep-learning/

Hoonhout, B. M., Radermacher, M., Baart, F., & van der Maaten, L. J. P. (2015). An automated method for semantic classification of regions in coastal images. *Coastal Engineering*, *105*, 1–12. https://doi.org/10.1016/j.coastaleng.2015.07.010

Hosna, A., Merry, E., Gyalmo, J., Alom, Z., Aung, Z., & Azim, M. A. (2022). Transfer learning: A friendly introduction. *Journal of Big Data*, *9*(1), 102. https://doi.org/10.1186/s40537-022-00652-w

Huanping, Z., Congying, L., & Xinfeng, Y. (2011). Optimization research on Artificial Neural Network Model. *Proceedings of 2011 International Conference on Computer Science and Network Technology*, *3*, 1724–1727. https://ieeexplore.ieee.org/abstract/document/6182301/

Inbaraj, X. A., & Jeng, J.-H. (2021). Mask-GradCAM: Object identification and localization of visual presentation for deep convolutional network. *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, 1171–1178. https://ieeexplore.ieee.org/abstract/document/9358569/

Ingrosso, G., Abbiati, M., Badalamenti, F., Bavestrello, G., Belmonte, G., Cannas, R., Benedetti-Cecchi, L., Bertolino, M., Bevilacqua, S., Bianchi, C. N., Bo, M., Boscari, E., Cardone, F., Cattaneo-Vietti, R., Cau, A., Cerrano, C., Chemello, R., Chimienti, G., Congiu, L., … Boero, F. (2018). Chapter Three—Mediterranean Bioconstructions Along the Italian Coast. In C. Sheppard (Ed.), *Advances in Marine Biology* (Vol. 79, pp. 61–136). Academic Press. https://doi.org/10.1016/bs.amb.2018.05.001

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv Preprint arXiv:1502.03167*. https://asvk.cs.msu.ru/~sveta/%D1%80%D0%B5%D1%84%D0%B5%D1%80%D0%B0%D1%82/batch_normalization.pdf

James, M. R., & Robson, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface*, *117*(F3). https://doi.org/10.1029/2011JF002289

Jiang, S., Zheng, Y., & Solomatine, D. (2020). Improving AI system awareness of geoscience knowledge: Symbiotic integration of physical approaches and deep learning. *Geophysical Research Letters*, *47*(13), e2020GL088229. https://doi.org/10.1029/2020GL088229

Jiang, P.-T., Zhang, C.-B., Hou, Q., Cheng, M.-M., & Wei, Y. (2021). Layercam: Exploring hierarchical class activation maps for localization. *IEEE Transactions on Image Processing*, *30*, 5875–5888. https://ieeexplore.ieee.org/abstract/document/9462463/

Jianping, C., Jie, X., Qiao, H., Wei, Y., Zili, L., Bin, H., & Wei, W. (2016). Quantitative geoscience and geological big data development: A review. *Acta Geologica Sinica - English Edition*, *90*(4), 1490–1515. https://doi.org/10.1111/1755-6724.12782

Jolivet, L., & Faccenna, C. (2000). Mediterranean extension and the Africa-Eurasia collision. *Tectonics*, *19*(6), 1095–1106. https://doi.org/10.1029/2000TC900018

Joo, H.-T., & Kim, K.-J. (2019). Visualization of deep reinforcement learning using grad-CAM: How AI plays atari games? *2019 IEEE Conference on Games (CoG)*, 1–2. https://ieeexplore.ieee.org/abstract/document/8847950/

Kelletat, D. H. (2005). Notches. *Encyclopedia of Earth Sciences Series*, *14*, 728–729. Scopus.

Kennedy, D. M., & Dickson, M. E. (2006). Lithological control on the elevation of shore platforms in a microtidal setting. *Earth Surface Processes and Landforms*, *31*(12), 1575–1584. https://doi.org/10.1002/esp.1358

Kennedy, D. M. (2013). *Topographic field surveying in geomorphology. Treatise on Geomorphology. 14*, 110–118.

Kennedy, D. M., Stephenson, W. J., & Naylor, L. A. (2014). Chapter 1. Introduction to the rock coasts of the world. *Geological Society, London, Memoirs*, *40*(1), 1–5. https://doi.org/10.1144/M40.1

Khurram, S., Pour, A. B., Bagheri, M., Ariffin, E. H., Akhir, M. F., & Hamzah, S. B. (2025). Developments in deep learning algorithms for coastline extraction from remote sensing imagery: A systematic review. *Earth Science Informatics*, *18*(3), 292. https://doi.org/10.1007/s12145-025-01805-0

Kingma, D. P., & Ba, J. (2017). *Adam: A Method for Stochastic Optimization* (No. arXiv:1412.6980). arXiv. https://doi.org/10.48550/arXiv.1412.6980

Lambeck, K., Antonioli, F., Purcell, A., & Silenzi, S. (2004). Sea-level change along the Italian coast for the past 10,000 yr. *Quaternary Science Reviews*, *23*(14–15), 1567–1598. https://www.sciencedirect.com/science/article/pii/S0277379104000526?casa_token=kJeVKZHNU-MAAAAA:FuON_Zpz69-zekAHvoMAMDge7NxwLDJXplVenKNYXT0O0QDDPv1s0qM1xqXHpt160plTykC3Tg

Lampedusa Boa oceanografica—ENEA. (2024). https://www.lampedusa.enea.it/dati/boa/index.php?lang=it

Lane, S. N., Richards, K. S., & Chandler, J. H. (1994). Developments in monitoring and modelling small-scale river bed topography. *Earth Surface Processes and Landforms*, *19*(4), 349–368. https://doi.org/10.1002/esp.3290190406

Leaflet—An open-source JavaScript library for interactive maps. Retrieved 16 July 2024, from https://leafletjs.com/

Lentini, F., Carbone, S., Catalano, S., & Grasso, M. (1995). Principali lineamenti strutturali della Sicilia nord-orientale. *Studi Geol. Camerti*, *2*(1995), 319–929.

Levin, L. A., Bett, B. J., Gates, A. R., Heimbach, P., Howe, B. M., Janssen, F., McCurdy, A., Ruhl, H. A., Snelgrove, P., & Stocks, K. I. (2019). Global observing needs in the deep ocean. *Frontiers in Marine Science*, *6*, 241. https://www.frontiersin.org/articles/10.3389/fmars.2019.00241/full

Li, J., Lin, D., Wang, Y., Xu, G., & Ding, C. (2021). Towards a reliable evaluation of local interpretation methods. *Applied Sciences*, *11*(6), 2732. https://www.mdpi.com/2076-3417/11/6/2732

Lim, M., Petley, D. N., Rosser, N. J., Allison, R. J., Long, A. J., & Pybus, D. (2005). Combined Digital Photogrammetry and Time-of-Flight Laser Scanning for Monitoring Cliff Evolution. *The Photogrammetric Record*, *20*(110), 109–129. https://doi.org/10.1111/j.1477-9730.2005.00315.x

Lim, M. (2014). Chapter 3 The rock coast of the British Isles: Cliffs. *Geological Society, London, Memoirs*, *40*(1), 19–38. https://doi.org/10.1144/M40.3

Lionello, P., Bhend, J., Buzzi, A., Della-Marta, P. M., Krichak, S. O., Jansa, A., Maheras, P., Sanna, A., Trigo, I. F., & Trigo, R. (2006). Cyclones in the Mediterranean region: Climatology and effects on the environment. In *Developments in earth and environmental sciences* (Vol. 4, pp. 325–372). Elsevier. https://www.sciencedirect.com/science/article/pii/S1571919706800091

Lipton, Z. C. (2018). The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, *16*(3), 31–57. https://doi.org/10.1145/3236386.3241340

Lopes, J. F., Da Costa, V. G. T., Barbin, D. F., Cruz-Tirado, L. J. P., Baeten, V., & Barbon Junior, S. (2022). Deep computer vision system for cocoa classification. *Multimedia Tools and Applications*, *81*(28), 41059–41077. https://doi.org/10.1007/s11042-022-13097-3

Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, *2*, 1150–1157. https://ieeexplore.ieee.org/abstract/document/790410/

Ložnjak, S., Kramberger, T., Cesar, I., & Kovačević, R. (2020). Automobile Classification Using Transfer Learning on ResNet Neural Network Architecture. https://doi.org/10.19279/TVZ.PD.2020-8-1-18

Lyall, C. (2019). Introduction: Mixed messages for the interdisciplinary research community. In C. Lyall, *Being an Interdisciplinary Academic,* 1-77. Springer International Publishing. https://doi.org/10.1007/978-3-030-18659-3_1

Maas, H. G. (2015). A modular geometric model for underwater photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*, 139–141. https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-5-W5/139/2015/isprsarchives-XL-5-W5-139-2015.html

Maltese Islands Weather. (2023, November 9). https://www.malteseislandsweather.com/october-2023-was-extremely-warm-and-the-driest-in-a-century/

Martinelli, M., Bistacchi, A., Balsamo, F., & Meda, M. (2019). Late Oligocene to Pliocene Extension in the Maltese Islands and Implications for Geodynamics of the Pantelleria Rift and Pelagian Platform. *Tectonics*, *38*(9), 3394–3415. https://doi.org/10.1029/2019TC005627

Mather, A. E. (2009). Tectonic setting and landscape development. *The Physical Geography of the Mediterranean*, 5–32.

Matthews, N. A. (2008). Aerial and close-range photogrammetric technology: Providing resource documentation, interpretation, and preservation. https://agris.fao.org/search/en/providers/122376/records/647472cb425ec3c088f30143

Mehta, A. (2019). An ultimate guide to understanding supervised learning. *Digital Vidya.* https://www. digitalvidya. com/blog/supervised-learning.

Micheletti, N., Foresti, L., Robert, S., Leuenberger, M., Pedrazzini, A., Jaboyedoff, M., & Kanevski, M. (2014). Machine Learning: Feature selection methods for landslide susceptibility mapping. *Mathematical Geosciences*, *46*(1), 33–57. https://doi.org/10.1007/s11004-013-9511-0

Micheletti, N., Chandler, J. H., & Lane, S. N. (2015). Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms*, *40*(4), 473–486. https://doi.org/10.1002/esp.3648

Millot, C., & Taupier-Letage, I. (2005). Circulation in the Mediterranean Sea. In A. Saliot (Ed.), *The Mediterranean Sea* (Vol. 5K, pp. 29–66). Springer Berlin Heidelberg. https://doi.org/10.1007/b107143

Moore, P. J., Martin, J. B., & Screaton, E. J. (2009). Geochemical and statistical evidence of recharge, mixing, and controls on spring discharge in an eogenetic karst aquifer. *Journal of Hydrology*, *376*(3–4), 443–455. https://www.sciencedirect.com/science/article/pii/S0022169409004545

Morucci, S., Picone, M., Nardone, G., & Arena, G. (2016). Tides and waves in the Central Mediterranean Sea. *Journal of Operational Oceanography*, *9*(sup1), s10–s17. https://doi.org/10.1080/1755876X.2015.1120960

Mosbrucker, A. R., Major, J. J., Spicer, K. R., & Pitlick, J. (2017). Camera system considerations for geomorphic applications of SfM photogrammetry. *Earth Surface Processes and Landforms*, *42*(6), 969–986. https://doi.org/10.1002/esp.4066

Moses, C. A. (2013). Tropical rock coasts: Cliff, notch and platform erosion dynamics. *Progress in Physical Geography: Earth and Environment*, *37*(2), 206–226. https://doi.org/10.1177/0309133312460073

Moses, C. A., Robinson, D., Kazmer, M., & Williams, R. (2014). Towards an improved understanding of erosion rates and tidal notch development on limestone coasts in the Tropics: 10 years of micro-erosion meter measurements, Phang Nga Bay, Thailand. *Earth Surface Processes and Landforms*, *40*(6), 771–782. https://doi.org/10.1002/esp.3683

Muhammad, M. B., & Yeasin, M. (2020). Eigen-cam: Class activation map using principal components. *2020 International Joint Conference on Neural Networks (IJCNN)*, 1–7. https://ieeexplore.ieee.org/abstract/document/9206626/

Murphy, K. P. (2022). Probabilistic machine learning: An introduction. MIT press. https://books.google.it/books?hl=en&lr=&id=wrZNEAAAQBAJ&oi=fnd&pg=PR27&dq=Probabilistic+machine+learning:+an+introduction&ots=LaodKDfoPg&sig=mapk_KN6KlYQSz_QjGgg-CGoxNY

Nagyfi, R. (2018, September 4). The differences between Artificial and Biological Neural Networks. *Medium.* https://towardsdatascience.com/the-differences-between-artificial-and-biological-neural-networks-a8b46db828b7

(NAS) National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. 2005. Facilitating Interdisciplinary Research. Washington, DC: The National Academies Press. https://doi.org/10.17226/11153.

Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., Schlötterer, J., Van Keulen, M., & Seifert, C. (2023). From Anecdotal Evidence to Quantitative Evaluation Methods: A Systematic Review on Evaluating Explainable AI. *ACM Computing Surveys*, *55*(13s), 1–42. https://doi.org/10.1145/3583558

Naylor, L. A., Stephenson, W. J., & Trenhaile, A. S. (2010). Rock coast geomorphology: Recent advances and future research directions. *Geomorphology*, *114*(1), 3–11. https://doi.org/10.1016/j.geomorph.2009.02.004

Nocerino, E., Menna, F., Fassi, F., & Remondino, F. (2016). Underwater calibration of dome port pressure housings. *XL-3/W4*. https://re.public.polimi.it/handle/11311/1017085

Nocquet, J.-M., & Calais, E. (2004). Geodetic measurements of crustal deformation in the Western Mediterranean and Europe. *Pure and Applied Geophysics*, *161*(3), 661–681. https://doi.org/10.1007/s00024-003-2468-z

Palamara, D. R., Dickson, M. E., & Kennedy, D. M. (2007). Defining shore platform boundaries using airborne laser scan data: A preliminary investigation. *Earth Surface Processes and Landforms*, *32*(6), 945–953. https://doi.org/10.1002/esp.1485

Pan, W., & Yang, Q. (2013). Transfer learning in heterogeneous collaborative filtering domains. *Artificial Intelligence*, *197*, 39–55. https://doi.org/10.1016/j.artint.2013.01.003

Papakonstantinou, A., Topouzelis, K., & Pavlogeorgatos, G. (2016). Coastline zones identification and 3D coastal mapping using UAV spatial data. *ISPRS International Journal of Geo-Information*, *5*(6), Article 6. https://doi.org/10.3390/ijgi5060075

Pedley, H. M., & Grasso, M. (1993). Controls on faunal and sediment cyclicity within the Tripoli and Calcare di Base basins (Late Miocene) of central Sicily. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *105*(3–4), 337–360.

Pinet, P. R. (2006). Invitation to oceanography. Jones & Bartlett Learning. https://books.google.it/books?hl=en&lr=&id=GgtEiPSrvH8C&oi=fnd&pg=PP16&dq=Pinet,+P.R.,+(1996).+Invitation+of+Oceanography,+&ots=BDAo3EWqEh&sig=fA_gzqCCrtgjjaCaYCyV15J4iGs

Piras, M., Taddia, G., Forno, M. G., Gattiglio, M., Aicardi, I., Dabove, P., Russo, S. L., & Lingua, A. (2017). Detailed geological mapping in mountain areas using an unmanned aerial vehicle: Application to the

Rodoretto Valley, NW Italian Alps. *Geomatics, Natural Hazards and Risk*, *8*(1), 137–149. https://doi.org/10.1080/19475705.2016.1225228

Pirazzoli, P. A. (1986). Marine notches. *Sea-Level Research*, 361–400. Scopus. https://doi.org/10.1007/978-94-009-4215-8_12

Pirazzoli, P. A., Laborel, J., Saliège, J. F., Erol, O., Kayan, I., & Person, A. (1991). Holocene raised shorelines on the Hatay coasts (Turkey): Palaeoecological and tectonic implications. *Marine Geology*, *96*(3–4), 295–311.

Pirazzoli, P. A., Stiros, S. C., Laborel, J., Laborel-Deguen, F., Arnold, M., Papageorgiou, S., & Morhangel, C. (1994). Late-Holocene shoreline changes related to palaeoseismic events in the Ionian Islands, Greece. *Holocene*, *4*(4), 397–405. Scopus. https://doi.org/10.1177/095968369400400407

Pirazzoli, P. A., Laborel, J., & Stiros, S. C. (1996). Coastal indicators of rapid uplift and subsidence: Examples from Crete and other eastern Mediterranean sites. *Zeitschrift Für Geomorphologie. Supplementband*, *102*, 21–35. https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=6318960

Pitman, S. J. (2014). Methods for field measurement and remote sensing of the swash zone. https://eprints.soton.ac.uk/364270/

Poulain, P.-M., Barbanti, R., Font, J., Cruzado, A., Millot, C., Gertman, I., Griffa, A., Molcard, A., Rupolo, V., & Le Bras, S. (2007). MedArgo: A drifting profiler program in the Mediterranean Sea. *Ocean Science*, *3*(3), 379–395.

Poulain, P. M., Menna, M., & Mauri, E. (2012). Surface geostrophic circulation of the Mediterranean Sea derived from drifter and satellite altimeter data. *Journal of Physical Oceanography*, *42*(6), 973–990. https://journals.ametsoc.org/view/journals/phoc/42/6/jpo-d-11-0159.1.xml

Qian, N. (1999). On the momentum term in gradient descent learning algorithms. *Neural Networks*, *12*(1), 145–151. https://www.sciencedirect.com/science/article/pii/S0893608098001166

Ramaswamy, H. G. (2020). Ablation-cam: Visual explanations for deep convolutional network via gradient-free localization. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 983–991. http://openaccess.thecvf.com/content_WACV_2020/html/Desai_Ablation-CAM_Visual_Explanations_for_Deep_Convolutional_Network_via_Gradient-free_Localization_WACV_2020_paper.html

Remondino, F., & Fraser, C. (2006). Digital camera calibration methods: Considerations and comparisons. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *36*(5), 266–272. https://www.research-collection.ethz.ch/handle/20.500.11850/158067

Remondino, F., & El-Hakim, S. (2006). Image-based 3D Modelling: A Review. *The Photogrammetric Record*, *21*(115), 269–291. https://doi.org/10.1111/j.1477-9730.2006.00383.x

Remondino, F., Barazzetti, L., Nex, F. C., Scaioni, M., & Sarazzi, D. (2011). UAV photogrammetry for mapping and 3D modeling: Current status and future perspectives. *Proceedings of the International Conference on Unmanned Aerial Vehicle in Geomatics (UAV-g): 14-16 September 2011, Zurich, Switzerland*, 25–31. https://research.utwente.nl/en/publications/uav-photogrammetry-for-mapping-and-3d-modeling-current-status-and

Rohling, E. J., Grant, K., Bolshaw, M., Roberts, A. P., Siddall, M., Hemleben, C., & Kucera, M. (2009). Antarctic temperature and global sea level closely coupled over the past five glacial cycles. *Nature Geoscience*, *2*(7), 500–504. https://www.nature.com/articles/ngeo557

Royden, L. H., & Papanikolaou, D. J. (2011). Slab segmentation and late Cenozoic disruption of the Hellenic arc. *Geochemistry, Geophysics, Geosystems*, *12*(3). https://doi.org/10.1029/2010GC003280

Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv Preprint arXiv:1609.04747*.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Sabato, G., Scardino, G., Kushabaha, A., Ciraolo, G., Scala, P., Manno, G., & Scicchitano, G. (2025). Optical Flow for Wave Characterization in a 2D Water Flume Using Video Analysis: A Cutting-Edge Tool for Coastal Monitoring. *EGU25-424, Copernicus Meetings, 2025*. https://iris.unipa.it/handle/10447/677567

Schembri, P. J., Dimech, M., Camilleri, M., & Page, R. (2007). *Living deep-water Lophelia and Madrepora corals in Maltese waters (Strait of Sicily, Mediterranean Sea)*. https://www.um.edu.mt/library/oar/handle/123456789/5507

Schembri, P. J., Hunt, C., Pedley, H. M., Malone, C., & Stoddart, S. (2009). The environment of the Maltese Islands. https://www.um.edu.mt/library/oar/handle/123456789/47464

Schmidt, V. E., & Rzhanov, Y. (2012). Measurement of micro-bathymetry with a GOPRO underwater stereo camera pair. *2012 Oceans*, 1–6. https://ieeexplore.ieee.org/abstract/document/6404786/

Schwalbe, G., & Finzel, B. (2024). A comprehensive taxonomy for explainable artificial intelligence: A systematic survey of surveys on methods and concepts. *Data Mining and Knowledge Discovery*, *38*(5), 3043–3101. https://doi.org/10.1007/s10618-022-00867-8

Seitz, L., Haas, C., Noack, M., & Wieprecht, S. (2018). From picture to porosity of river bed material using Structure-from-Motion with Multi-View-Stereo. *Geomorphology*, *306*, 80–89. https://doi.org/10.1016/j.geomorph.2018.01.014

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017a). Grad-CAM: Visual Explanations From Deep Networks via Gradient-Based Localization. 618–626. https://openaccess.thecvf.com/content_iccv_2017/html/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.html

Selvaraju, R. R., Das, A., Vedantam, R., Cogswell, M., Parikh, D., & Batra, D. (2017b). Grad-CAM: Why did you say that? *arXiv:1611.07450 Preprint arXiv*. https://doi.org/10.48550/arXiv.1611.07450

Serpelloni, E., Vannucci, G., Pondrelli, S., Argnani, A., Casula, G., Anzidei, M., Baldi, P., & Gasperini, P. (2007). Kinematics of the Western Africa-Eurasia plate boundary from focal mechanisms and GPS data. *Geophysical Journal International*, *169*(3), 1180–1200. https://doi.org/10.1111/j.1365-246X.2007.03367.x

Shortis, M. (2015). Calibration techniques for accurate measurements by underwater camera systems. *Sensors*, *15*(12). https://doi.org/10.3390/s151229831

Singh, P., & Sharma, A. (2022). Interpretation and classification of arrhythmia using deep convolutional network. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–12. https://ieeexplore.ieee.org/abstract/document/9877905/

Snavely, N., Seitz, S. M., & Szeliski, R. (2008). Modeling the World from Internet Photo Collections. *International Journal of Computer Vision*, *80*(2), 189–210. https://doi.org/10.1007/s11263-007-0107-3

Sozio, Angelo & Scardino, Giovanni & Parisi, Francesca & Pirulli, Giuseppe & Fiscarelli, Alessandro & Barracane, Giovanni & Scicchitano, Giovanni. (2025). Machine Learning techniques for the detection of geomorphological features in nearshore environments.

Srinivas, S., & Fleuret, F. (2019). Full-gradient representation for neural network visualization. *Advances in Neural Information Processing Systems*, *32*. https://proceedings.neurips.cc/paper/2019/hash/80537a945c7aaa788ccfcdf1b99b5d8f-Abstract.html

Srivastava, R. K., Greff, K., & Schmidhuber, J. (2015). Highway Networks. *arXiv:1505.00387 Preprint arXiv*. https://doi.org/10.48550/arXiv.1505.00387

Stephenson, W. J. (2000). Shore platforms: Remain a neglected coastal feature. *PROGRESS IN PHYSICAL GEOGRAPHY*, *24*(4), U6–U6.

Stephenson, W. J., & Naylor, L. A. (2011). Geological controls on boulder production in a rock coast setting: Insights from South Wales, UK. *Marine Geology*, *283*(1), 12–24. https://doi.org/10.1016/j.margeo.2010.07.001

Stephenson, W. (2015). Rock Coasts. In G. Masselink & R. Gehrels (Eds.), *Coastal Environments and Global Change* (1st ed., pp. 356–379). Wiley. https://doi.org/10.1002/9781119117261.ch15

Sunamura, T. (1992). Geomorphology of rocky coasts. *3*. Wiley.

Sunamura, T., Tsujimoto, H., & Aoki, H. (2014). The rock coast of Japan. In D. M. Kennedy, W. J. Stephenson, & L. A. Naylor (Eds.), *Rock Coast Geomorphology: A Global Synthesis* (Vol. 40, p. 0). Geological Society of London. https://doi.org/10.1144/M40.12

Swirad, Z. M., Rosser, N. J., & Brain, M. J. (2019). Identifying mechanisms of shore platform erosion using Structure-from-Motion (SfM) photogrammetry. *Earth Surface Processes and Landforms*, *44*(8), 1542–1558. https://doi.org/10.1002/esp.4591

Szeliski, R. (2011). Computer Graphics Achievement Award. *ACM SIGGRAPH 2011 Awards*, 1–1. https://doi.org/10.1145/2018996.2018997

Tavani, S., Granado, P., Corradetti, A., Girundo, M., Iannace, A., Arbués, P., Muñoz, J. A., & Mazzoli, S. (2014). Building a virtual outcrop, extracting geological information from it, and sharing the results in Google Earth via OpenPlot and Photoscan: An example from the Khaviz Anticline (Iran). *Computers & Geosciences*, *63*, 44–53. https://doi.org/10.1016/j.cageo.2013.10.013

Tavani, S., Billi, A., Corradetti, A., Mercuri, M., Bosman, A., Cuffaro, M., Seers, T., & Carminati, E. (2022). Smartphone assisted fieldwork: Towards the digital transition of geoscience fieldwork using LiDAR-equipped iPhones. *Earth-Science Reviews*, *227*, 103969. https://doi.org/10.1016/j.earscirev.2022.103969

Taylor, M. E., & Stone, P. (2009). Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, *10*(7). http://www.jmlr.org/papers/volume10/taylor09a/taylor09a.pdf

Teague, J., & Scott, T. (2017). Underwater photogrammetry and 3D reconstruction of submerged objects in shallow environments by ROV and underwater GPS. *Journal of Marine Science Research and Technology*.

Tieleman, T. (2012). Lecture 6.5-RMSProp: Divide the gradient by a running average of its recent magnitude. In *COURSERA: Neural Networks for Machine Learning, 4* (2), 26. https://cir.nii.ac.jp/crid/1370017282431050757

Trenhaile, A. S. (1980). Shore platforms: A neglected coastal feature. *Progress in Physical Geography: Earth and Environment*, *4*(1), 1–23. https://doi.org/10.1177/030913338000400101

Trenhaile, A. S. (1987). The geomorphology of rock coasts. Clarendon Press and Oxford University Press. https://cir.nii.ac.jp/crid/1130282271255904256

Trenhaile, A. S. (2001). Modelling the Quaternary evolution of shore platforms and erosional continental shelves. *Earth Surface Processes and Landforms*, *26*(10), 1103–1128. https://doi.org/10.1002/esp.255

Trenhaile, A. S. (2002). Rock coasts, with particular emphasis on shore platforms. *Geomorphology*, *48*(1), 7–22. https://doi.org/10.1016/S0169-555X(02)00173-3

Trenhaile, A. S. (2010). The effect of Holocene changes in relative sea level on the morphology of rocky coasts. *Geomorphology*, *114*(1–2), 30–41. https://www.sciencedirect.com/science/article/pii/S0169555X09000695

Trenhaile, A. S. (2014). Climate change and its impact on rock coasts. In D. M. Kennedy, W. J. Stephenson, & L. A. Naylor (Eds.), *Rock Coast Geomorphology: A Global Synthesis* (Vol. 40, p. 0). Geological Society of London. https://doi.org/10.1144/M40.2

Trenhaile, A. S. (2015). Coastal notches: Their morphology, formation, and function. *Earth-Science Reviews*, *150*, 285–304. https://doi.org/10.1016/j.earscirev.2015.08.003

Uhrig, R. E. (1995). Introduction to artificial neural networks. *Proceedings of IECON '95 - 21st Annual Conference on IEEE Industrial Electronics*, *1*, 33–37. https://doi.org/10.1109/IECON.1995.483329

Ullman, S. (1979). The interpretation of structure from motion. *203*(1153), 405–426. https://royalsocietypublishing.org/doi/abs/10.1098/rspb.1979.0006

Vaccher, V. (2019). SfM photogrammetry of rock coasts: Acquisition and elaboration of time-lapse images to create 3d models. Master's Thesis. Department of Mathematic and Geoscience, University of Trieste.

Vaccher, V., Hastewell, L., Devoto, S., Corradetti, A., Mantovani, M., Korbar, T., & Furlani, S. (2024). The application of UAV-derived SfM-MVS photogrammetry for the investigation of storm wave boulder deposits on a small rocky island in the semi-enclosed Northern Adriatic Sea. *Geomatics, Natural Hazards and Risk*, *15*(1), 2295817. https://doi.org/10.1080/19475705.2023.2295817

Vacchi, M., Montefalcone, M., Bianchi, C. N., Morri, C., & Ferrari, M. (2012). Hydrodynamic constraints to the seaward development of Posidonia oceanica meadows. *Estuarine, Coastal and Shelf Science*, *97*, 58–65. https://www.sciencedirect.com/science/article/pii/S0272771411004859?casa_token=NB4_iDF2N_YAAAAA:4mcrspQR7G1vLHQ0OQFwIrEnmGY2nJIllOUFm5a-fDbQRRJgMKA0f-eOTcHiP_HoSMxPwFXQaw

Valentini, N., Saponieri, A., Molfetta, M. G., & Damiani, L. (2017). New algorithms for shoreline monitoring from coastal video systems. *Earth Science Informatics*, *10*(4), 495–506. https://doi.org/10.1007/s12145-017-0302-x

Valentini, N., & Balouin, Y. (2020). Assessment of a smartphone-based camera system for coastal image segmentation and sargassum monitoring. *Journal of Marine Science and Engineering*, *8*. https://doi.org/10.3390/jmse8010023

Van Veen, F. (2017). Neural network zoo prequel: Cells and layers. Retried from https://Www. Asimovinstitute. Org/Author/Fjodorvanveen.

Vasuki, Y., Holden, E. J., Kovesi, P., & Micklethwaite, S. (2014). Semi-automatic mapping of geological Structures using UAV-based photogrammetric data: An image analysis approach. *Computers & Geosciences*, *69*, 22–32. https://doi.org/10.1016/j.cageo.2014.04.012

Vasuki, Y., Holden, E. J., Kovesi, P., & Micklethwaite, S. (2017). An interactive image segmentation method for lithological boundary detection: A rapid mapping tool for geologists. *Computers & Geosciences*, *100*, 27–40. https://doi.org/10.1016/j.cageo.2016.12.001

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*, *179*, 300–314. https://doi.org/10.1016/j.geomorph.2012.08.021

Xiao, M., Zhang, L., Shi, W., Liu, J., He, W., & Jiang, Z. (2021). A visualization method based on the Grad-CAM for medical image segmentation model. *2021 International Conference on Electronic Information Engineering and Computer Science (EIECS)*, 242–247. https://ieeexplore.ieee.org/abstract/document/9587953/

Yusiong, J. P. T. (2012). Optimizing artificial neural networks using cat swarm optimization algorithm. *International Journal of Intelligent Systems and Applications*, *5*(1), 69. https://www.mecs-press.org/ijisa/ijisa-v5-n1/IJISA-V5-N1-7.pdf

Zhang, Y., Han, J., Chen, B., Chang, L., Song, T., & Cai, G. (2022). Classification of Microcalcification Clusters Using Bilateral Features Based on Graph Convolutional Network. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/fonc.2022.871662

Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning Deep Features for Discriminative Localization. 2921–2929. https://openaccess.thecvf.com/content_cvpr_2016/html/Zhou_Learning_Deep_Features_CVPR_2016_paper.html

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, *109*(1), 43–76. Proceedings of the IEEE. https://doi.org/10.1109/JPROC.2020.3004555

Zupan, J. (1994). Introduction to artificial neural network (ANN) methods: What they are and how to use them. *Acta Chimica Slovenica*, *41*(3), 327. https://www.academia.edu/download/52003225/ANN.pdf

Appendix 1

This comprehensive guide provides detailed instructions for creating integrated 3D models using synchronized above-water and underwater images in Agisoft Metashape Professional.

Pre-processing setup

- Print the chessboard calibration pattern from Metashape (Tools → Show Chessboard).
- Mount the printed pattern on 3-4mm plexiglass to take underwater photos. The structure on which the chessboard is mounted must be rigid and resistant to bending in water.
- Take at least 20 underwater images of the grid from approximately the same distance used during the swim survey and with the same camera settings.
- Create separate chunk for the calibration grid images. Align grid images in **High** quality mode; use these calibration parameters for processing the underwater images. Export the calibration parameters (Camera calibration → Adjusted → Save)

Main processing workflow

- Create two separate folders, each containing the same number of images above and below the waterline. Rename the underwater photos with the same names as the corresponding aw images (using a specific program, i.e. Bulk Renames Utility) – this maintains image correspondence between chunks.
- Create a chunk (C1) and import the above waterline images.
- Align the images in **High** quality mode – this establishes the position and orientation reference.
- Create a second chunk (C2) and import the renamed below waterline images.
- Load the saved calibration parameters in C2 and disable **adaptive camera model fitting**. Apply and save.
- From C1, in the left panel with the uploaded images, select **Convert References** → select UTM Zone 33N to use metric values. From the same panel, select **Export References**, considering latitude, longitude, altitude, accuracy and the relative errors.
- Open the downloaded file using Excel, reorder the columns and change the Z value to 0.5 m for the images above and -0.5 for the images below the waterline (the cameras distance is about 1 m). Pitch and roll corrections are unnecessary as Metashape doesn't use these parameters for horizontal-facing cameras.
- From C2, select **Import References** button to upload the modified Exif data. In camera calibration, select **fisheye** mode. Then select **Align** images.
- Create the models, both above and below the water. Apply tie point filtering to clean the point cloud before building mesh, to avoid unnecessary processing. Build **mesh** based on depth

maps (Workflow → Built model → Depth maps - quality High). Create the texture of both models (Workflow → Built texture → Diffuse map – images – generic)

- Then, merge the 3D models (Workflow → **Merge chunks** → select C1 and C2 – select **depth map**s – merge **assets**). Manually clean up the point cloud, build the mesh, and apply the texture to this model as well.

Important factors

- GPS availability for images above the waterline.

- Clear and well-exposed images.

- Proper image correspondence between above and bw images.

- Good overlap between consecutive photos. (Minimum 75%)

- Fisheye lens for underwater processing.

- Depth maps and assets for successfully merging models.

- Proper EXIF data management and altitude corrections.

- Manual cleanup of the merged model.

Appendix 2

The python script used to randomly select a certain number of images from the dataset

```
import os
import shutil
import random

source_folder = '\\\\172.30.121.5\\GeoSwim\\Valeria_Vaccher\\2020-Lampedusa'

destination_folder = 'F:\\cancellate\\bsl'

# Number of images to copy from each ‘asl’ subfolder
number_of_images_to_copy = 250

# Function to copy images from one ‘asl’ folder to another folder
def copy_images_from_asl (source_folder_asl, destination_folder_asl, number_of_images_to_copy):
    if 'cancelled' not in source_folder_asl:
        try:
            images_to_copy = random.sample(os.listdir(source_folder_asl), number_of_images_to_copy)

            for image in images_to_copy:
                origin = os.path.join(source_folder_asl, image)
                destination = os.path.join(destination_folder_asl, image)

                # shutil.copy(origin, destination)
        except:
            print(source_folder_asl)

# browse through folders
for root, dirs, files in os.walk(source_folder):
    for dir_name in dirs:
        if 'b.s.l' == dir_name:

            # Get the complete path of the ‘asl’ folders
            cartella_sorgente_asl = os.path.join(root, dir_name)

            # Create a corresponding folder in the destination folder
            Destination_folder_asl = os.path.join(destination_folder, os.path.relpath(source_folder_asl,
source_folder))
            # os.makedirs(destination_folder_asl, exist_ok=True)

            # Copy random images from the ‘asl’ folder
            Copy_images_from_asl(source_folder_asl, destination_folder_asl,
number_of_images_to_copy)
```

Acknowledgments

Il mio percorso accademico è iniziato nel 2015, durante il tirocinio della laurea triennale, quando ho partecipato per la prima volta a una campagna di nuoto geomorfologico nell'ambito del progetto Geoswim. Quell'esperienza mi ha appassionata fin da subito e mi ha portata fino a qui.

Il programma di dottorato, inizialmente previsto per tre anni e poi prolungato a quattro, è stato impegnativo ma emozionante. L'anno di pausa mi ha offerto la preziosa opportunità di dedicarmi a mio figlio Jonathan. Gli anni più intensi di lavoro sono stati una perfetta combinazione di ufficio, campagne e famiglia. L'ultima parte del dottorato, invece, l'ho affrontata con mia figlia Rachel in grembo e, dopo la sua nascita, grazie al prezioso aiuto dei nonni e di due giovani e brave babysitter, sono riuscita a concludere la tesi nel miglior modo possibile. Mi sento molto fortunata per l'opportunità che ho avuto, oltre che immensamente grata alla mia famiglia, che ogni giorno mi rende più forte e fiera.

Desidero ringraziare tutti i professori, e non solo, che mi hanno offerto il loro tempo, per la loro consulenza tecnica ed il loro sostegno. In particolare, un grazie speciale va a Leonardo.

Grazie di cuore alle mie amiche Glenda, Giulia e Diana, per avermi ascoltata, capita e distratta.

Infine, un ringraziamento unico e speciale a Thomas, il mio compagno: grazie per l'amore e la fiducia che mi dai ogni giorno, e per avermi incoraggiata, giorno dopo giorno, a non mollare mai.