UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING

Lucija Žužić

Machine Learning Models for Watercraft Trajectory Analysis and Maritime Infrastructure Inspection

DOCTORAL THESIS

Rijeka, 2026.

UNIVERSITY OF RIJEKA
FACULTY OF ENGINEERING

Lucija Žužić

Machine Learning Models for Watercraft Trajectory Analysis and Maritime Infrastructure Inspection

DOCTORAL THESIS

Thesis Supervisor: Prof. Jonatan Lerga, PhD

Rijeka, 2026.

SVEUČILIŠTE U RIJECI
TEHNIČKI FAKULTET

Lucija Žužić

Modeli strojnoga učenja za analizu putanja plovila i inspekciju pomorske infrastrukture

DOKTORSKI RAD

Mentor: prof. dr. sc. Jonatan Lerga

Rijeka, 2026.

Acknowledgements

I thank Prof. Jonatan Lerga, PhD, for acting as my mentor on this Doctoral thesis. I thank Prof. Goran Mauša, PhD, for previously acting as my mentor on my Bachelor's and Master's theses.

I would like to thank Senior Assistant Marko Njirjak, PhD, Assistant Marko Babić, Patrizia Janković Bevandić, PhD, Senior Assistant Erik Otović, PhD, Associate Prof. Daniela Kalafatović, PhD, Associate Prof. Ivan Dražić, PhD, Associate Prof. Loredana Simčić, PhD, Docent Franko Hržić, PhD, Prof. Xiumei Li, PhD, Irena Petrijevčanin, PhD, Luka Škrlj, Prof. Aleksandar Nešković, PhD, Prof. Renato Filjar, PhD, Stella Dumenčić, and Barbara Breš, who acted as my co-authors on published papers. I would also like to thank Deni Klen for creating segmentation masks for maritime infrastructure inspection by hand. Finally, I would like to thank all my colleagues, fellow students, and staff at the Department of Computer Engineering, the Faculty of Engineering, and the broader University of Rijeka.

I would like to thank my family and friends for their advice in my academic work and private life.

A special note goes out to my sister for igniting my interest in Computer Engineering and motivating me to choose it as my research area in my further studies.

I also thank everyone else who supported me but was not mentioned by name here.

Abstract

The rapid expansion of maritime traffic, particularly in the recreational sector involving personal watercrafts, presents significant safety and operational challenges. Conventional monitoring systems often fail to capture the erratic dynamics of small vessels or adequately inspect the aging underwater infrastructure that supports maritime activities. This doctoral dissertation addresses these gaps by developing advanced machine learning frameworks for watercraft trajectory analysis and maritime infrastructure inspection.

The first part of the research focuses on trajectory forecasting and anomaly detection. A comparative analysis is conducted between classical probabilistic methods, such as Markov chains and Bayesian inference, and modern deep learning architectures, including recurrent neural network models, attention mechanisms, and foundation models.

The study demonstrates that while Bayesian methods offer simple interpretability for short-term predictions, location-agnostic deep learning models significantly outperform them in long-term forecasting accuracy. These models are successfully applied to a trajectory intersection detection system, achieving a true negative rate of 95.45% and drastically reducing false alarms compared to traditional rule-based systems.

Additionally, a novel, computationally efficient trajectory fingerprinting method based on inflection points is introduced, offering an easy-to-understand and flexible solution for real-time trajectory anomaly detection on embedded systems.

The second part addresses the visual inspection of underwater infrastructure. A specialized semantic segmentation framework is proposed that uses a custom U-Net pipeline trained with an $F1$ score loss function.

Keywords: trajectory forecasting, deep learning, anomaly detection, semantic segmentation, personal watercraft, underwater inspection

Sažetak

Naglo povećanje pomorskog prometa, posebno u rekreacijskom sektoru koji uključuje osobna plovila, donosi značajne sigurnosne i operativne izazove. Konvencionalni sustavi nadzora često ne uspijevaju pratiti nepredvidivu dinamiku malih plovila niti adekvatno pregledati ostarjelu podvodnu infrastrukturu koja podržava pomorske aktivnosti. Ova doktorska disertacija rješava navedene nedostatke razvojem naprednih okvira strojnog i dubokog učenja za analizu putanja plovila i inspekciju pomorske infrastrukture.

Prvi dio istraživanja usmjeren je na predviđanje putanja i detekciju njihovih anomalija. Provedena je usporedna analiza između klasičnih probabilističkih metoda, poput Markovljevih lanaca i Bayesovog zaključivanja, te modernih arhitektura dubokog učenja, uključujući povratne neuronske mreže, mehanizme pažnje i temeljne modele.

Istraživanje pokazuje da, dok Bayesove metode nude jednostavno tumačenje za kratkoročne procjene, lokacijski agnostički modeli dubokog učenja značajno ih nadmašuju u točnosti dugoročnog predviđanja. Ovi su modeli uspješno primijenjeni na sustav za detekciju presjeka putanja, postižući stopu istinski negativnih rezultata od 95.45% i drastično smanjujući broj lažnih uzbuna u usporedbi s tradicionalnim sustavima temeljenim na pravilima.

Dodatno, predstavljena je nova, računalno učinkovita metoda otiska putanje temeljena na točkama infleksije, koja nudi jednostavno razumljivo i fleksibilno rješenje za detekciju anomalija u stvarnom vremenu na ugradbenim sustavima.

Drugi dio bavi se vizualnom inspekcijom podvodne infrastrukture. Predložen je specijalizirani okvir za semantičku segmentaciju koji koristi prilagođeni U-Net radni okvir treniran s funkcijom gubitka temeljenom na $F1$ rezultatu.

Ključne riječi: predviđanje putanja, duboko učenje, detekcija anomalija, semantička segmentacija, osobna plovila, podvodna inspekcija

Prošireni sažetak

Naglo povećanje pomorskog prometa, posebno u rekreacijskom sektoru koji uključuje osobna plovila, donosi značajne sigurnosne i operativne izazove. Konvencionalni sustavi nadzora često ne uspijevaju pratiti nepredvidivu dinamiku kretanja malih plovila niti adekvatno pregledati ostarjelu podvodnu infrastrukturu. Ova doktorska disertacija doprinosi rješenju navedenih nedostataka razvojem naprednih tehnika strojnog učenja za analizu putanja plovila i inspekciju pomorske infrastrukture.

Prvi dio istraživanja usmjeren je na predviđanje putanja i detekciju njihovih anomalija u pomorskom prometu. Provedena je usporedna analiza između klasičnih probabilističkih metoda, poput Markovljevih lanaca i Bayesovog zaključivanja, te modernih arhitektura dubokog učenja, uključujući povratne neuronske mreže, mehanizme pažnje i temeljne modele.

Pokazano je da Bayesove metode nude jednostavno tumačenje za kratkoročna predviđanja putanje. U disertaciji se položaj plovila procjenjuje pomoću Bayesovog pristupa na temelju smjera, brzine, vremenskih intervala između mjerenja te pomaka u geografskoj širini i dužini. Predstavljen je dodatni pristup korištenjem Markovljevog lanca za obradu podataka o putanji preuzetih iz sustava za praćenje plovila u oblaku koji omogućuje daljinsko upravljanje plovilima. Izvješća o visini valova i meteorološki podaci korišteni su za procjenu utjecaja vremenskih uvjeta na putanje osobnih plovila. Konstrukcija putanje uključuje udružene pomake geografske dužine i širine ili brzinu, smjer i stvarne vremenske intervale između mjerenja. Podjela putanja na segmente koji se ne preklapaju postiže dugoročni prognostički prozor do 10 sekundi. Rezultati su se pokazali usporedivima onima dobivenim kada su okolišne varijable dodane modelima uvjetne vjerojatnosti. U usporedbi s metodom koja je koristila Markovljeve lance, uspješnija se pokazala metoda procjene koja je koristila jednu ili dvije prethodne stvarne vrijednosti i Bayesov pristup, dokazujući da korištenje prethodno predviđenih vrijednosti u lancu akumulira pogrešku. Ograničenja s kojima se ova metoda suočava u dugoročnom predviđanju potakla su prelazak na složenije pristupe koji uključuju strojno učenje.

Za razliku od postojećih radova usmjerenih na definirane pomorske rute, ova disertacija uvodi pristup dubokog učenja koji ne ovisi o lokaciji i moguće ga je generalizirati na različita okruženja. To se postiže inovativnim pristupom pretprocesiranja koji uključuje pomak i skaliranje putanje. Novi dvosmjerni i konvolucijski model dugog kratkoročnog pamćenja (engl. Long

Short-Term Memory (LSTM)) koji je autorica razvila za klasifikaciju binarnih peptida prilagođen je za predviđanje kontinuiranih vrijednosti. Korištenjem arhitektura povratnih neuronskih mreža (engl. Recurrent Neural Network (RNN)) s LSTM i slojevima rekurentne jedinice s vratima (engl. Gated Recurrent Unit (GRU)) te recentnih modela temeljenih na pažnji i temeljnih modela, kao što je ujedinjeni model vremenskih nizova (engl. Unified Time Series (UniTS)), uspoređena je učinkovitost predviđanja putanje koristeći stvarne podatke prikupljene na 1282 najamna mjesta za osobna plovila diljem svijeta. Dodatno, značajan doprinos istraživanja predstavlja to što se proširuje primjenjivost navedenih modela, uključujući originalne dvosmjerne i konvolucijske LSTM pristupe, na procjenu pomorske putanje, čime se uklanja potreba za treniranjem zasebnih modela za različite lokacije uz postizanje iznimne prediktivne točnosti. Rezultati pokazuju da temeljni modeli i mehanizmi pažnje značajno nadmašuju tradicionalne metode.

Modeli koji su predloženi za korištenje u predviđanju uspješno su primijenjeni na sustav za otkrivanje presjeka putanja (mogućih sudara plovila), postižući stopu istinski negativnih rezultata od 95.45% i drastično smanjujući lažne uzbune u usporedbi s tradicionalnim sustavima temeljenim na pravilima. Detekcija presjeka ovdje se definira kao predviđanje putanja osobnih plovila do 30 sekundi unaprijed i identificiranje potencijalnih presjeka putanja koji bi mogli dovesti do nesreća. Predviđanje putanje integrirano je u otkrivanje sudara, zamjenjujući prethodni sustav inteligentne kontrole udaljenosti (engl. Intelligent Distance Control (IDC)) analizom podataka s 527 lokacija rasprostranjenih po različitih kontinentima i prikupljenih tijekom nekoliko godina te analizom ključnih varijabli poput geografske dužine i širine, brzine i smjera. Model pažnje koji koristi pomake geografske dužine i širine, s horizontom predviđanja od 30 sekundi, povećao je stopu istinski negativnih rezultata na 95.45%, što je značajno poboljšanje u odnosu na pristup koji ne koristi strojno učenje. GRU model pažnje vraća rezultate za manje od 12 milisekundi po sekvenci, omogućujući implementaciju u stvarnom vremenu.

U ovoj disertaciji predlaže se i inovativna, računalno učinkovita i prilagodljiva metoda za grupiranje putanja i otkrivanje kinematičkih anomalija temeljena na točkama infleksije. Takvom metodologijom generira se otisak putanje koji omogućuje lako razumljiv i fleksibilan pristup za otkrivanje anomalija. Ovakav pristup može se primijeniti na putanjama u stvarnom vremenu u ugradbenim sustavima i može se prilagoditi promjenama u duljinama putanje i okruženjima. Kodiranjem promjena oblika putanje u kompaktne otiske, tehnika izbjegava računalno zahtjevno duboko učenje. Rezultati pokazuju da predložena metoda otisaka može razlikovati anomalne obrasce vožnje, nudeći održivo rješenje za otkrivanje anomalija u stvarnom vremenu. Putanje su segmentirane, točke infleksije potom izdvojene, a grupiranje se provodi pomoću otisaka putanje. Dodatni eksperimenti ljudske evaluacije potvrđuju da je pristup usklađen s aktivacijom IDC sustava.

Drugi dio istraživanja bavi se vizualnom inspekcijom podvodne infrastrukture. Predložen je specijalizirani okvir za semantičku segmentaciju koji koristi prilagođeni U-Net radni okvir. Isti

je dizajniran za preciznu segmentaciju slika koja postiže visoku preciznost u različitim domenama, uključujući inspekciju i nadzor podvodnih struktura. Za razliku od tradicionalnih modela s više klasa, predloženi pristup odvojeno trenira U-Net model za određenu klasu, omogućujući podešavanje parametara prilagođeno domeni i optimizaciju praga. Predloženi radni okvir uključuje funkciju gubitka temeljenu na Dice ($F1$) rezultatu i prilagođenu za što bolje performanse na velikom broju slika prilikom treniranja modela, predobradu specifičnu za domenu i podjelu prema klasama prije treniranja. Predloženi redoslijed primijenjenih postupaka pokazao je dosljedna poboljšanja performansi u odnosu na najsuvremenije metode, kao što su You Only Look Once (YOLO) model i osnovni U-Net. Rezultati pokazuju da ova modularna strategija segmentacije postiže bolje vrijednosti za Dice koeficijente, presjek nad unijom (engl. Intersection over Union (IoU)) i preciznost pojedine klase, posebno za male ili nejasne strukture.

Contents

Chapter 1

INTRODUCTION

1.1 Overview of challenges in the maritime domain

The volume of global maritime traffic has increased significantly. More than $50,000$ commercial vessels travel the sea daily. Maritime trade represents close to 90% of the world's cargo trade [62]. In the recreational sector, small personal watercraft have increased traffic volume in tourist destinations. The coexistence of heavy commercial traffic and erratic movements of recreational watercraft presents complex safety challenges.

In coastal regions popular with tourists, marinas and rental operators aim to bridge this gap by using proprietary tracking systems such as OtoTrak [98]. Cloud-based remote-control solutions enforce location and speed limits. However, their safety mechanisms are predominantly reactive. For example, the OtoTrak Intelligent Distance Control (IDC) system relies on hard-coded heuristics that assume constant velocity and heading over short time horizons. Static proximity rules may be sufficient for less agile vessels [74, 75]. However, linear assumptions with frequent, rapid maneuvers cause false alarms. Late intervention and missed warnings can also occur in dynamic environments. Models for larger vessel types [9, 66] are thus inadequate if an intersection of trajectories is imminent.

Shifting from reactive, rule-based systems to proactive, predictive models is needed for the next iteration of maritime safety systems. Using advanced Machine Learning (ML) techniques, this research aims to develop robust trajectory forecasting and anomaly detection frameworks that anticipate dangerous situations before they materialize [89, 155].

The primary aim of this research is to develop and validate computational methods for watercraft trajectory prediction, anomaly detection, and underwater infrastructure inspection that outperform existing rule-based and generic ML approaches [178, 181].

Implicit in this aim is the hypothesis that data preprocessing, specifically the normalization of trajectory vectors, is sufficient to create efficient, location-agnostic models that generalize well

to unseen environments. It is further posited that specialized Deep Learning (DL) architectures, such as attention-based Recurrent Neural Networks (RNNs) and modified U-Net pipelines, are required to handle the specific characteristics of maritime data, whether it be the erratic movement of a watercraft or the turbid visual environment of an underwater inspection.

To this end, the dissertation investigates:

- The efficacy of probabilistic versus DL models for variable-horizon forecasting.
- The application of foundation models, specifically Unified Time Series (UniTS) [41], to the maritime domain.
- The development of interpretable, lightweight anomaly detection methods suitable for embedded systems.
- The optimization of semantic segmentation networks for detecting structural defects in low-contrast underwater imagery.

1.1.1 Research aim and research hypotheses

This research aims to develop methods applied in the maritime domain for watercraft trajectory prediction, trajectory anomaly and intersection detection, and underwater infrastructure inspection.

The research hypotheses are:

- Trajectory data preprocessing enables the development of efficient location-agnostic models for watercraft trajectory prediction, trajectory anomaly and intersection detection.
- The proposed U-Net-based semantic segmentation framework outperforms multi-class models for underwater infrastructure inspection.
- Machine learning-based trajectory intersection detection outperforms rule-based detection.
- Inflection-point-based trajectory anomaly detection output matches survey results and rule-based detection.

1.1.2 Expected scientific contribution of the proposed research

The expected scientific contributions include:

- An interpretable inflection point-based trajectory anomaly detection model that matches survey results and a rule-based detection model.

- Bayesian and Markov chain models utilizing environmental data fused with trajectory data for watercraft trajectory prediction.

- Modified deep learning models for watercraft trajectory prediction.

- A U-Net-based semantic segmentation framework for underwater infrastructure inspection.

1.2 The need for advanced trajectory analysis

Drivers can forecast the movement of surrounding entities. Natural processes of observation are a key factor for safety in crowded waterways. Autonomous [87] and Intelligent Transportation Systems (ITSs) [8] struggle to reproduce this capacity [45]. Advanced Driver Assistance System (ADAS) applications in the automotive sector have strongly advanced in this aspect [72, 136], but maritime environments have their own unique challenges. These obstacles include fluid dynamics, weather conditions (wind, waves, and currents) [6, 73, 118], and the lack of structured lanes, which contribute to highly non-linear movement patterns that require specialized clustering methods [92, 106], deviating significantly from established clustering research in other domains [46, 51, 82].

Current forecasting models cannot strike a balance between short-term precision and long-term accuracy. Markov chains [39] and probabilistic approaches are easy to understand but degrade in performance for long-term predictions [38]. Conversely, DL models, such as RNNs models [16, 50] and transformers [56, 142], excel at capturing temporal dependencies but can be computationally intensive and opaque [17, 70, 79].

This research addresses these challenges by developing an array of location-agnostic models trained on a massive dataset of real-world personal watercraft trajectories. By normalizing trajectory data [164] to be independent of specific geographic coordinates, generalized models were deployed across diverse locations without requiring site-specific retraining. This work builds upon foundational studies in maritime pattern extraction [81, 100, 103], replacing rigid Graph Neural Network (GNN) models that require nodes and edges in the graph to be predefined using a structured traffic flow [54, 157]. However, none of the previously proposed methods are location agnostic [25, 141]. The previous findings do not address the dynamic nature of a personal watercraft [68, 107, 163]. Instead, other authors focus their attention on Long Short-Term Memory (LSTM) models in other domains [2, 26, 167]. Trajectory-based studies have been largely focused on human movement [146], cycling [156], and land traffic [78, 138], but there are few papers focusing on agile small watercraft models operating near the coast.

1.3 The need for maritime infrastructure inspection

Image analysis has advanced from probabilistic modeling of complex multi-modal data to lightweight DL architectures for real-time applications. Semantic segmentation is used in real-world systems. Applications vary, from medical diagnostics and environmental monitoring to infrastructure assessment and robotic navigation. In structural health monitoring, the early detection of anomalies such as structural cracks is needed for rapid reaction [67]. Structural defects, such as cracks or indentations, in submerged or deteriorated environments are important in early detection to enhance the safety of maritime and civil infrastructure.

Visual inspection of underwater infrastructure is a vital assignment in ensuring maritime safety. However, divers still perform this task manually. The process that involves humans is hazardous, time-consuming, and subjective. Recent trends in structural health monitoring emphasize computational efficiency and edge-device deployment. Qingyi et al. [108] developed a lightweight Crack Segment You Only Look Once (CS-YOLO) model specifically for concrete crack segmentation. Novel contributions bring model compression and computation demands to the forefront [80]. These innovations show that heavy convolutional architectures can be optimized into lightweight frameworks, such as the Cross Stage Partial Network (CSPNet) [148], to perform high-precision segmentation tasks in resource-constrained environments. Environmental noise, occlusion, and heterogeneous visual characteristics cripple the capabilities of traditional image segmentation models [124] deployed in the medical and civil infrastructure domains. Underwater crack detection, due to sparse data and visibility issues [97, 158], remains underexplored. Automated solutions that combine Remotely Operated Vehicle (ROV) imaging with computer vision are needed to mitigate these risks. Most popular object detection models, including You Only Look Once (YOLO), cannot accurately accommodate the fine-grained segmentation required for irregular cracks in turbid environments [151]. A specialized DL approach, such as the U-Net framework adapted in this research, [64, 143], can be a viable solution for semantic segmentation. The cited studies in unison create a summary of recent segmentation approaches. The research has focused on innovations, including statistical fusion models designed for data-rich, multi-source environments and efficient neural networks optimized for speed and specific feature-extraction tasks.

1.4 Dissertation structure

This dissertation focuses on various models for trajectory analysis for personal watercraft operation and maritime infrastructure inspection. For this purpose, a dataset on personal watercraft rentals was obtained with the permission of OtoTrak [98] to study trajectories, and images capturing submerged concrete walls that must be inspected were sourced from Vectrino [64, 143].

This dissertation comprises four main parts: probabilistic approaches to trajectory forecasting, DL approaches to trajectory forecasting and intersection detection, inflection point-based trajectory fingerprinting, and U-Net semantic segmentation for underwater infrastructure inspection. The flowchart in Figure 1.1 illustrates the connections among the research areas and the individual purposes of each part of the dissertation.

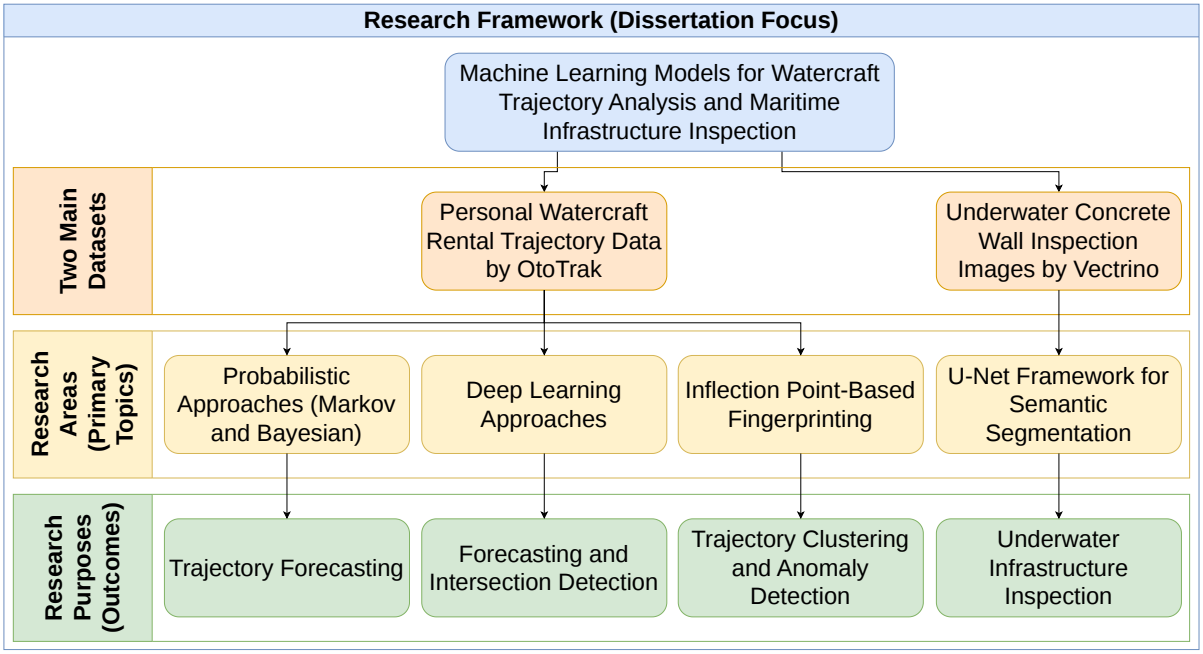

Figure 1.1: A flowchart illustrating the various research domains utilized in the dissertation and the corresponding outcomes resulting from various parts of the research. The blue block indicates the dissertation title, the orange block denotes the two main datasets, the yellow blocks encompass the primary research areas, and the green blocks present the research outcomes.

Chapter 2

DATA

This thesis makes a dual contribution to the maritime domain by using two main datasets that are distinct in nature. The watercraft data, provided by OtoTrak [98], contains messages transmitted during recreational use to study the longitude, latitude, speed, and heading of vessels, as well as other metadata. The OtoTrak [98] data has been utilized in personal watercraft trajectory forecasting [170, 173], trajectory intersection detection [174], and trajectory anomaly detection [176]. The second segment of this thesis addresses image processing and designs a U-Net pipeline [175] tailored to an infrastructure inspection framework yet applicable across diverse domains, as the design of self-driving vessels also requires obstacle recognition and hazard detection for optional avoidance maneuvers. The underwater inspection produced images of concrete walls [64, 143] that must be examined for potential damage, and the U-Net model's output can indicate the need for repair operations.

2.1 OtoTrak dataset

OtoTrak [98] system begins with the OtoTrak module, an in-house-designed electronic device that interfaces with a watercraft. The OtoTrak [98] module is easy to install on a personal watercraft, according to instructions for many common watercraft types, and connects to the OtoTrak cloud. Inside a sturdy Ingress Protection (IP) 67 (IP67)-compliant case, it houses features developed for real-world use by a team with years of experience in marine electronics. The module connects to the OtoTrak cloud services [98] via a cellular network and leverages Global System for Mobile Communications (GSM) to ensure continuous availability.

The OtoTrak Control app [98] allows users on any smartphone or tablet to obtain a clear overview of what is happening with the watercrafts. It is easy to manually trigger any action or switch between profiles to automatically execute action plans prepared by the rental supervision systems. Using the OtoTrak mobile app [98] and the Global Positioning System (GPS), the

rental site operator can always see the watercraft on a map and use the plug-and-play controls to limit its speed, turn off the engine, or sound an alarm directly from a smartphone or tablet. The OtoTrak web app [98] allows users to configure their ride zones, manage operator users, and view all collected data in an organized manner through interactive reports. Existing customers can create additional operators who can log in to the app and use it, while OtoTrak handles the rest [98]. The OtoTrak approach [98] is a game-changer in the rental business, allowing operators to worry less about automated rental supervision, draw ride zones directly from the app, and assign a watercraft fleet to them. The support team is available and actively participates in ongoing product development.

The OtoTrak system [98] tracks every ride and prepares action plans to reduce the speed of rented watercrafts leaving the zone or of watercrafts in close proximity, thanks to the IDC. It can sound an Automatic Tour Control (ATC) buzzer to tell the driver their ride is up. The OtoTrak module [98] is designed to enhance safety and convenience by providing scheduling insights.

Ride timers are a standout feature of many OtoTrak-powered rentals [98] because ATC allows users to enter the ride duration at any time, before or after renting the watercraft, and the watercraft's buzzer beeps automatically 3 minutes before the ride ends. The 3 minutes can be adjusted to any time that suits the users and allows riders to return. ATC also accounts for jet-ski distance, ensuring that rental rides conclude on schedule.

2.1.1 Tours and events table structure

Standard tours (rides) use either the "MASTER" or "RENT" mode, as indicated in Table 2.1 by the numbers 1 and 2, which are used in the database entries for trajectories created by the owner of the rental site or by tourists, respectively.

Table 2.1: Operation mode ID, name, and description [98].

ID	Name	Description
1	MASTER	Owner testing
2	RENT	Tourist rental
3	ERROR	Data corruption
4	TOUR	Safari mode participants
6	TOUR_LEAD	Safari mode leader
7	TRACKING	Variable-frequency data
40	TOW	Non-functional vehicle
73	RACE	Full watercraft speed

If data cannot be successfully stored and transmitted at the minimum frequency, for example, when the time delay exceeds 5 seconds, the entries will be marked as unreliable with the word "ERROR" and the number 3. Safari mode enables operators to monitor guided tours and set

group limits with a single button press. The tour guide will generate a ride using the number 6 and the descriptor "TOUR_LEAD," while other participants in the safari mode will generate rides using the number 4 and the descriptor "TOUR." If a reporting frequency other than the standard 1 Hz is desired, the ride will be marked as "TRACKING" and highlighted by the number 7. A vehicle requiring repair will be returned to shore, as indicated by the word "TOW" and the identifier 40. Adrenaline-seeking riders may wish to utilize the full capacity of the watercraft speed via the operation mode titled "RACE" and numbered 73.

The tours table, described in Table 2.2, allows the user to browse the "id", "asset_id," "device_id," "location_id," and "client_id," which uniquely identify a trajectory, watercraft, OtoTrak module, rental site, and OtoTrak customer, respectively [98]. The "operation_mode_id" corresponds to the numbers listed with their corresponding descriptors in Table 2.1. Rental site operators can enter the ride's "expected_duration" in seconds, and the planned "start" and "end" timestamps. The ride "version" starts at 1 and increases by 1 for each update. The ride will be marked by timestamps for the first ("created_at") and the latest version ("updated_at"). The number of "persons" is optional and applies only to safari-mode tours.

Table 2.2: Tours column names, types, descriptions, and example values [98].

Name	Type	Description	Example value
id	Integer	Tour (ride) ID	7643819
asset_id	Integer	Vehicle ID	5237
device_id	Integer	OtoTrak module ID	5737
location_id	Integer	Location ID	20
client_id	Integer	Owner ID	13
operation_mode_id	Integer	operation mode ID	2
expected_duration	Integer	Expected duration (s)	600
version	Integer	Tour (ride) version	2
persons	Integer	Number of persons	3
start	DateTime	Start timestamp	2023-05-27 12:23:39
end	DateTime	End timestamp	2023-05-27 12:35:34
created_at	DateTime	Created timestamp	2023-05-27 12:23:45.651825
updated_at	DateTime	Updated timestamp	2023-05-27 12:35:41.816821

The integer values for "asset_id," "device_id," "location_id," and "client_id" in Table 2.2 correspond to the textual data in "asset_name," "hardware_id," "location_name," and "client_name," in Table 2.3. While the numbers in Table 2.2 must remain stable and cannot be changed to ensure unique and consistent identification, the names in Table 2.3 can be updated at any time upon client request. There is one exception, as the "operation_mode_name" in Table 2.3 is matched with the "operation_mode_id" in Table 2.2 as a fixed system-wide feature that is outside user

control and defined in Table 2.1.

Table 2.3: Tours identifier column names, types, descriptions, and example values [98].

Name	Type	Description	Example value
asset_name	String	Vehicle name	955 DB
hardware_id	String	Module name	359159972073523
location_name	String	Location name	Jet ski rent Dubrovnik
client_name	String	Owner name	Turistički obrt "Hermes 2"
operation_mode_name	String	Operation mode name	RENT

The events table, showcased in Table 2.4, can be linked with the tour data in Table 2.2 and Table 2.3 by the "asset_id," "hardware_id," "location_id," and "client_id," which destinguish a single watercraft, OtoTrak module, rental site, and OtoTrak customer, respectively [98]. The "time" field records timestamps to preserve the order of the trajectory data.

Table 2.4: Events column names, types, descriptions, and example values [98].

Name	Type	Description	Example value
time	DateTime	Point in time when event occurred	2023-05-27 12:23:39.447085
asset_id	Integer	Vehicle ID	5237
hardware_id	String	Module name	359159972073523
location_id	Integer	Location ID	20
client_id	Integer	Owner ID	13

The events table contains a JavaScript Object Notation (JSON) object in the column "fields," and the properties of this object are described in Table 2.5. The "speed" is recorded in km/h, and the watercraft's position is recorded using "latitude" and "longitude" in °. The "direction" starts at 90° (North), and this must be accounted for in trajectory processing and creation. The "operation_mode" is defined in Table 2.1, and the "billing_plan_code" is either BASIC, STANDARD, or PREMIUM due to company standards. The GSM connection level is reported as a percentage via the "gsm_signal" value. The data also reports the "engine_hours," "throttle_limit," "battery_voltage," and "fuel_level_avg" status, as well as other features, such as "outstanding_limit_count" and "lock_geo_alert_distance," which are used to execute commands. OtoTrak [98] incorporates innovative features that make personal watercraft more fun for everyone, enhance safety using the "idc_triggers" and "speed_limit," and reduce costs. Real-time statistics and history in the cloud enable operators to view the individual paths of every ride and obtain a daily summary of how watercrafts were used, including speeds, fuel levels, engine fault codes, and more, in real time or later. System diagnostics enable operators to read active and historical "fault_codes" from the engine and other components to prevent damage to the watercraft.

The boolean values in the object defined by the "fields" column in the events database table are outlined in Table 2.6. Key programming enables users to remove old keys and program new

Table 2.5: Events field names, types, descriptions, and example values [98].

Name	Type	Description	Example value
speed	Float	Speed in km/h	1.9
buzzer	Integer	Sound alarm type (pattern)	4
latitude	Float	GPS latitude	42.653845
longitude	Float	GPS longitude	18.06391
direction	Integer	Heading $[0, \ldots, 359]$	44
gsm_signal	Integer	GSM signal strength (%)	31
speed_limit	Integer	Throttle attenuation (%)	70
idc_triggers	String	Vehicle ID list (speed limit)	['359159975042681']
fault_codes	String array	Fault codes list	['062C', 'D6A2']
operation_mode	String	Operation mode name	RENT
billing_plan_code	String	Price bracket	PREMIUM
battery_voltage	Float	Battery voltage	13.399999618530273
lock_geo_alert_distance	Integer	Lock geo alert distance	23
outstanding_limit_count	Integer	Outstanding	2
		limit count	
engine_hours	Integer	Engine hours	1490
fuel_level_avg	Integer	Fuel level average (%)	48
throttle_limit	Integer	Throttle limit (%)	99

ones in seconds if they have lost an ignition key, and the "key" in the events database indicates if any key exists. The "flip," "buzzer_active," "atc_buzzer_active," "engine," and "ignition" true or false values show when a watercraft is flipped, the buzzer or the ATC buzzer is turned on, and when the engine shuts down, or the ignition is toggled. A positive "sleep_mode" indicates an inactive watercraft, and a true "staff_mode" indicates that the current operation mode is not for the general public. The "zone" calculation uses the vessel's internal map directly, while the "in_zone," "in_primary_zone," and "in_restricted_zone" values are computed on the server. The "onboard_geofencing" option can be enabled or disabled, and the "speed_limit_active" value changes dynamically in response to driver and operator actions. The flags "lock_geo," "lock_engine," "lock_geo_alert," and "lock_engine_alert" grant the operator granular supervision over all vessels.

The data in Table 2.1 capture valuable ride characteristics for filtering trajectories from the final

Table 2.6: Events boolean field names and descriptions.

Name	Description
key	Indicator if key is detected
flip	Indicator if vehicle is upside-down
zone	Indicator if vehicle is within allowed zone (based on internal map)
engine	Engine running status
in_zone	Indicator if vehicle is within allowed zone (based on server calculation)
ignition	Ignition status
sleep_mode	Indicator if OtoTrak module is inactive
staff_mode	Indicator if the current operation mode is only for authorized personnel
buzzer_active	Indicator if sound alarm is active
in_primary_zone	Indicator if vehicle is within primary zone (based on server calculation)
in_restricted_zone	Indicator if vehicle is within restricted zone (based on server calculation)
onboard_geofencing	Indicator if geo fencing option is enabled on the OtoTrak module
speed_limit_active	Indicator if speed limit is active
lock_geo	Indicator if geo fencing is locked
lock_engine	Indicator if engine is locked
lock_geo_alert	Indicator if geo fencing alert is locked
lock_engine_alert	Indicator if engine alert is locked
atc_buzzer_active	Indicator if the ATC buzzer is active

dataset. After this, the metadata in Table 2.2 and Table 2.3 can be linked to events, as defined in Table 2.4. The events are then collected to form a complete trajectory, using columns for longitude, latitude, speed, and direction, listed in Table 2.5. Most of the metadata and boolean flags in Table 2.6 were not part of this study, as they are separate from the problem statement defining trajectory forecasting and trajectory anomaly detection. The raw data from the OtoTrak [98] database were not used directly as model inputs, and the preprocessing is described below.

2.1.2 Processed table structure

Four variables were studied separately using the approaches described in the author's papers published in the Journal of the Franklin Institute [170], Expert Systems with Applications [173], Knowledge-Based Systems [174], and Pattern Recognition [176]. These variables are speed, direction, longitude, and latitude offset. In this work, trajectories are defined by different subsets of the variables under investigation, which are sufficient to describe motion in two-dimensional space. These subsets were either longitude and latitude offset or speed and direction. The elapsed time between two records in the database, or the distance between two points in time,

was not forecast but was used to create a full trajectory from speed and direction.

To boost performance and mitigate problems that may arise due to the varying magnitudes among different parts of the input data [91, 130], several conventions were proposed to unify and normalize the data:

1. The values for latitude and longitude are rounded to ten decimal places to obtain a finite state space.
2. The speed at the point is recorded in kilometers per hour and rounded to a whole number.
3. The direction, representing the offset from the North (90°, positive direction of the y-axis), which has a value between 0° and 360°, recorded by the module and increases from east to west, was also used.
4. The time between two recorded points was also modeled and rounded to milliseconds, using seconds with three decimal places.
5. For each trajectory, the latitudes and longitudes were converted into relative values, which represent distances from the starting point, to ensure the model is location-independent. The latitudes and longitudes were further normalized to a 0.1 ° range, as no trajectory in the data exceeded this range.
6. Trajectories on the negative side of the x-axis are mirrored on the y-axis, and trajectories on the negative side of the y-axis are mirrored on the x-axis.

The data was processed for use in scientific papers published in the Journal of the Franklin Institute [170], Expert Systems with Applications [173], Knowledge-Based Systems [174], and Pattern Recognition [176]. Only the "longitude," "latitude," "speed," and "direction" values, together with their corresponding "time" values, were used. The "time" was processed in Python to use the number of seconds since the epoch (1970-01-01) instead of the date and time text. The value in milliseconds can be formatted as a timestamp if desired. On the advice of Oto-Trak partners [98], only the "MASTER" and "RENT" operation modes from Table 2.1 were used in the Journal of the Franklin Institute [170] and Expert Systems with Applications [173] publications, while the Knowledge-Based Systems [174] paper used only trajectories marked as "RENT." The Pattern Recognition [176] publication loosened the restrictions and also used "TRACKING" trajectory data. Trajectories marked as "ERROR" or where the lag between consecutive transmissions exceeded 5 seconds were omitted in the published research. Processed file fields are described in Table 2.7.

To ensure a globally applicable, location-agnostic approach, the "processed_longitude" and "processed_latitude" for the GPS trajectories were translated so that they start at $(0,0)$, the final point of the trajectories lies on the x-axis, and trajectories are mirrored so that all values

Table 2.7: Processed file fields.

Name	Description
time	Point in time when event occurred
processed_time	Seconds since epoch (1970-01-01)
longitude	GPS longitude
latitude	GPS latitude
processed_longitude	Normalized GPS longitude (translated and mirrored)
processed_latitude	Normalized GPS latitude (translated and mirrored)
scaled_longitude	Scaled GPS longitude (0.1 °)
scaled_latitude	Scaled GPS latitude (0.1 °)
speed	Speed in km/h
direction	Heading starting from the North
real	Is the data real-world or interpolated

are positive (in the first quadrant of the Cartesian coordinate system). As the statistical analysis of the raw data in the rest of the text demonstrates, 0.1 ° is sufficient to cover the longitude and latitude span of one trajectory, so the "scaled_longitude" and "scaled_latitude" data use this range. As requested by the reviewers for Expert Systems with Applications [173], linear interpolation can optionally be applied to fill all time gaps exceeding 1 second, since the expected transmission frequency is 1 Hz.

2.2 Image data

The dataset [64, 143] includes 1209 classified underwater images, separated into three classes representing cracks, indentations, and ropes. In addition, the images were split into training, testing, and validation sets to preserve the class ratio, thereby enabling standard benchmarking of model performance. Infrastructure inspection increasingly relies on automated crack detection, but this approach also has merit in the transportation sector, as computer vision can prevent accidents in autonomous vehicles and rental watercraft if a camera is installed onboard. All data captured underwater [64, 143] were grayscale-normalized, and the original high-resolution files were scaled to 416×416 pixels. The choice of using grayscale conversion for the underwater images [64, 143] containing instances of cracks, indentations, and ropes before inputting them into the U-Net model [175], was motivated by the observed shift in lightness and pixel intensity between these predefined labels.

2.2.1 Image preprocessing

Prior to use in the various U-Net models [175], the minimum grayscale pixel value is subtracted from all images, and the result is then divided by the absolute difference between the minimum and maximum grayscale pixel values. This process ensures that all values are positive and fall

within the range from 0 to 1, as shown in Table 2.8.

Table 2.8: The mean of the grayscale pixel value for each class label after the underwater images [64, 143]. The background represents any pixel not belonging to a particular label, and the last row in the table represents the absolute value of the difference between the mean for the background and foreground [175].

Value	Crack	Indentation	Rope
Background mean	0.617	0.6271	0.3603
Foreground mean	0.7209	0.6728	0.3825
Difference of means	0.1039	0.0457	0.0222

As the grayscale pixel mean in the rope class [64, 143] in the validation data was larger than that of the background, which contains all pixels outside the rope label, all values were replaced by subtracting them from 1. This process makes the model more robust, as the value 1 for foreground pixels now has a larger mean than the class 0 representing the background. A more pronounced change in grayscale pixel mean values between background and foreground improves the ability to properly differentiate between labels, underscoring the value of studying label means for the background and foreground after preprocessing.

The underwater images [64, 143] were resized and scaled to 416×416, which conforms to the YOLO [59, 137] segmentation model input rules. YOLO [59, 137] segmentation only accepts square images with dimensions exactly divisible by 32. If the image dimensions are not divisible by 32, the images are automatically resized according to the settings of the trained model.

2.2.2 Image augmentation

To obtain additional, more diverse imagery for training, the preprocessing augmentation creates three new images from a single original image [175]. In the augmentation process, the setup includes a 50% chance of a horizontal flip and an equal chance of a 90° clockwise rotation, a 90° counter-clockwise rotation, an upside-down rotation, or no rotation. A part of the image representing 0% and 30% of the original was cropped with a uniform probability of 7.5% at each edge. Additionally, a random rotation between -15° and 15° was applied. Hue alterations were randomized over the range -10° to 10°. Saturation and brightness changes use the predefined ranges of -25% to 25% and -20% to 20%, respectively. To implement a Gaussian blur, a kernel size of up to 1 pixel was used.

2.2.3 Image data split

As for data split, 80% of the data was set aside for training, and the remaining 20% was split into 7% for testing and 13% for validation in Table 2.9 for the underwater images [64, 143].

Table 2.9: The data split for training, testing, and validation, as well as the number of images in the entire underwater dataset [64, 143]. All classes marked in the image are listed in a separate column that divides the images into categories [175].

Class combination	Testing	Training	Validation	Entire dataset
Only crack	2	18	3	23
Only crack and indentation	13	146	24	183
Crack, indentation, and rope	34	379	61	474
Only crack and rope	6	67	11	84
Only indentation	7	73	12	92
Only indentation and rope	23	259	42	324
Only rope	2	23	4	29
Crack combined with others	55	610	99	764
Indentation combined with other	77	857	139	1073
Rope combined with others	65	728	118	911
Any class present	87	965	157	1209

The original images were triplicated during augmentation before being used for training. The process was also stratified by the combination of classes present in the image. The division was applied individually for categories of images that contained the same combinations of the crack, indentation, and rope classes.

Table 2.9 shows the training, testing, and validation data size for the underwater images [64, 143], and the images are divided into categories based on the combination of classes that are labeled in that instance. Cracks occur the least often in underwater images [64, 143], as most concrete surfaces studied are undamaged and are structurally sound. Ropes are the second most common class [64, 143], and their presence is not required by the problem under study, but they often occur given the nature of the shared maritime domain. Indentations [64, 143] are the most common class since a large portion of the inspected concrete walls have accumulated some smaller changes in structure, which is a natural part of the material aging process, further accelerated by the properties of sea water. Despite variation in class size, the difference is not significant enough to have a major influence on model performance for each label [64, 143], as evidenced by the results presented in the rest of the text.

The following chapters present the research results based on the described datasets, focusing on personal watercraft trajectory forecasting, anomaly detection, and underwater image analysis.

Chapter 3

PROBABILISTIC APPROACHES FOR TRAJECTORY FORECASTING

3.1 Probabilistic model introduction

A baseline of classical probabilistic methods [120] is used as a prerequisite for exploring complex DL architectures, and it is utilized to establish a baseline. This chapter deals with Bayesian probability and Markov chain applications in watercraft trajectory forecasting. These methods offer high interpretability [32], allowing us to explicitly model transition probabilities among kinematic states such as accelerating, turning, and stopping.

The motivation for this approach lies in the stochastic nature of human operation [11, 84, 132]. Since it lacks a fixed destination, a personal watercraft's future state is a probability distribution conditioned on its current state and intent influenced by different external factors [52, 60, 154]. By modeling these transitions, this research aims to capture the driver behavior inherent in the data [94, 104].

3.2 Probabilistic model methodology

Bayesian and Markov models for short-term and long-term predictions that align spatio-temporal with environmental data were proposed in this thesis as faster and lightweight alternatives to DL based on prior speed, heading, and offsets in longitude and latitude, using conditional probabilities powered by wind speed, temperature, and wave height data. Temperature and wind speed were derived from Meteorological Aerodrome Report (METAR) data [115], and environmental data, including European Union (EU) Copernicus Marine Service Information [30] describing past sea states (wave height) via satellite, were synchronized with GPS data.

3.2.1 Markov chains

The movement of a personal watercraft is defined using a stochastic process $\{X_t | t \in T\}$. Discrete-time Markov chains are used, with the caveat that the Markov property holds, meaning that the probability of transitioning to the next state depends solely on the current state, not the sequence of events that preceded it [10, 39].

$$P(X_{n+1} = j | X_n = i, X_{n-1} = k, \dots) = P(X_{n+1} = j | X_n = i) \tag{3.1}$$

The state of the watercraft is defined using discretized variables. The variable list consists of speed, heading, and relative position offsets (Δx and Δy). The transition matrix Π is created from data points from the OtoTrak dataset [98], calculating the empirical probability of moving from one kinematic state to another. This method aligns with established techniques for vessel pattern knowledge discovery [103].

3.2.2 Bayesian inference

To improve short-term accuracy for state estimation, Bayesian updating is investigated. It is assumed that the Markov property holds over a window of two consecutive states (a second-order Markov chain). The forecast for the next step X_n is derived from the posterior probability distribution given the actual observed values A_{n-1} and A_{n-2}:

$$P((X_n, A_{n-1}) | A_{n-1}, \dots, A_0) \approx P((X_n, A_{n-1}) | A_{n-1}, A_{n-2}) \tag{3.2}$$

This approach uses the ground truth of the latest data to account for the errors that usually accumulate in pure dead-reckoning systems [96, 122].

3.3 Probabilistic model results and discussion

The models were evaluated on a held-out test set using Euclidean distance and Root Mean Squared Error (RMSE) as the primary metric. The results, published in the Journal of the Franklin Institute [170], indicated strong performance on short horizons (1-2 seconds), with the Bayesian model successfully capturing immediate kinematic trends. Interestingly, the model predicted the y-offset (latitude) with slightly higher accuracy than the x-offset (longitude) in Table 3.1, likely attributable to the predominant orientation of coastlines in the training data.

However, the limitations of this approach became evident over longer forecasting horizons (5-10 seconds) in Table 3.2. The Markov chains are recursive, and their very nature causes the

Table 3.1: The average RMSE in $^\circ(\times 10^{-4})$, for the x-offset and y-offset estimated on the testing dataset by different models and short-term forecasting times. The bolded numbers indicate the lowest values and the best models for each forecasting time and variable [170].

Variable	Model	1 s	2 s	3 s	4 s	5 s
x-offset	1-step Bayes	1.776	3.755	**7.474**	8.597	8.999
x-offset	2-step Bayes	**1.504**	**3.735**	7.485	**8.593**	9.04
x-offset	Wave 2-step Markov	7.645	8.414	8.944	9.244	**8.937**
y-offset	1-step Bayes	1.335	2.594	**5.387**	6.431	6.813
y-offset	2-step Bayes	**1.105**	**2.588**	5.411	**6.382**	**6.753**

rapid accumulation of errors. While the model could predict that a vessel currently turning, for example, left would likely continue turning left in the next second, it failed to capture complex, long-term maneuvers or correct for sudden, erratic changes in driver behavior. This finding confirms that while probabilistic methods provide a solid mathematical baseline and are computationally efficient [38], they lack the capacity to model the high-dimensional, non-linear dependencies required for robust long-term forecasting in this domain.

Table 3.2: The average RMSE in $^\circ(\times 10^{-4})$, for the x-offset and y-offset estimated on the testing dataset by different models and long-term forecasting times. The bolded numbers indicate the lowest values and the best models for each forecasting time and variable [170].

Variable	Model	6 s	7 s	8 s	9 s	10 s
x-offset	Wave 1-step Markov	9.114	9.137	**8.942**	9.197	**8.877**
x-offset	Wave 2-step Markov	9.132	**9.091**	9.064	9.153	9.135
x-offset	Wind 2-step Markov	**8.951**	9.312	9.11	**8.947**	9.188
y-offset	2-step Markov	6.903	7.08	7.248	**6.946**	7.08
y-offset	Temperature 1-step Markov	7.06	**6.842**	7.154	7.022	7.016
y-offset	Wind 1-step Bayes	6.978	6.993	7.035	7.015	**6.949**
y-offset	Wind 2-step Bayes	6.948	7.068	**6.958**	7.042	7.042
y-offset	Wind 2-step Markov	**6.867**	7.234	7.134	7.104	7.044

Markov chain models depend on compounded errors in prior forecasts, forming a theoretical basis for the findings.

All variants of probabilistic frameworks were less successful than other authors' work on long-term trajectory prediction, for example, for land-based vehicles [3]. This motivates the next chapter of the thesis, which uses ML approaches to predict future states and evaluates these improved results using the described probabilistic models and the DL methodology patented in prior work.

Traditional methods, such as Kalman filters [11] or an adaptive hybrid Bernoulli filter [34], assume smooth, linear motion and fail to model the abrupt, nonlinear motion of a personal watercraft. ML and probabilistic models use different normality models, such as Kernel Density Estimation (KDE) [81], Gaussian Mixture Model (GMM) [69], and multiple Ornstein-Ühlenbeck (OU) processes [34], and perform better than traditional methods, but do not capture long-term

dependencies or context changes.

This failure of prior methods motivates the shift to DL and RNN models, such as LSTM [3, 93, 99], Gated Recurrent Unit (GRU), transformer-based models [76, 147, 150], GNN models [155], and the UniTS [41] foundation models for time series forecasting. These innovative approaches excel at learning sequential dependencies, and foundation models offer powerful generalization. This thesis aims to adapt advanced DL models to the specific, high-variance setting of GPS-equipped personal watercraft.

Chapter 4

DEEP LEARNING APPROACHES FOR TRAJECTORY FORECASTING AND INTERSECTION DETECTION

4.1 Introduction to deep learning models

To move beyond the restrictions of probabilistic methods listed in Chapter 3, this chapter explores the application of DL to maritime trajectory forecasting. This study focuses on RNN models, which are well equipped to process sequential data. This research investigates their ability to learn complex temporal dependencies. Furthermore, these forecasting capabilities are applied to the critical safety application of trajectory intersection detection, aiming to reduce the False Positive Rate (FPR) of existing rule-based systems.

DL models were used for both short-term and long-term predictions. PyTorch was used to develop RNN, LSTM [3], GRU, and GRU attention [49] models. The input data were normalized and mirrored. The UniTS foundation model [41] was fine-tuned for zero-shot prediction due to its prompt-token architecture. Bidirectional and convolutional LSTM were adapted from the original architecture designed for sequence classification using the Keras library and developed by the author, as published in Nature Machine Intelligence [93]. The original peptide classification model included dropout [40], Scaled Exponential Linear Unit (SELU), and sigmoid activation [93].

4.2 Deep learning model methodology

Figure 4.1 is a diagram illustrating the architectures of several ML models used for forecasting personal watercraft trajectories [168, 172] with variable horizons using four variables (longitude

offset, latitude offset, speed, and heading). It is divided into four main sections, each depicting a specific model type or component.

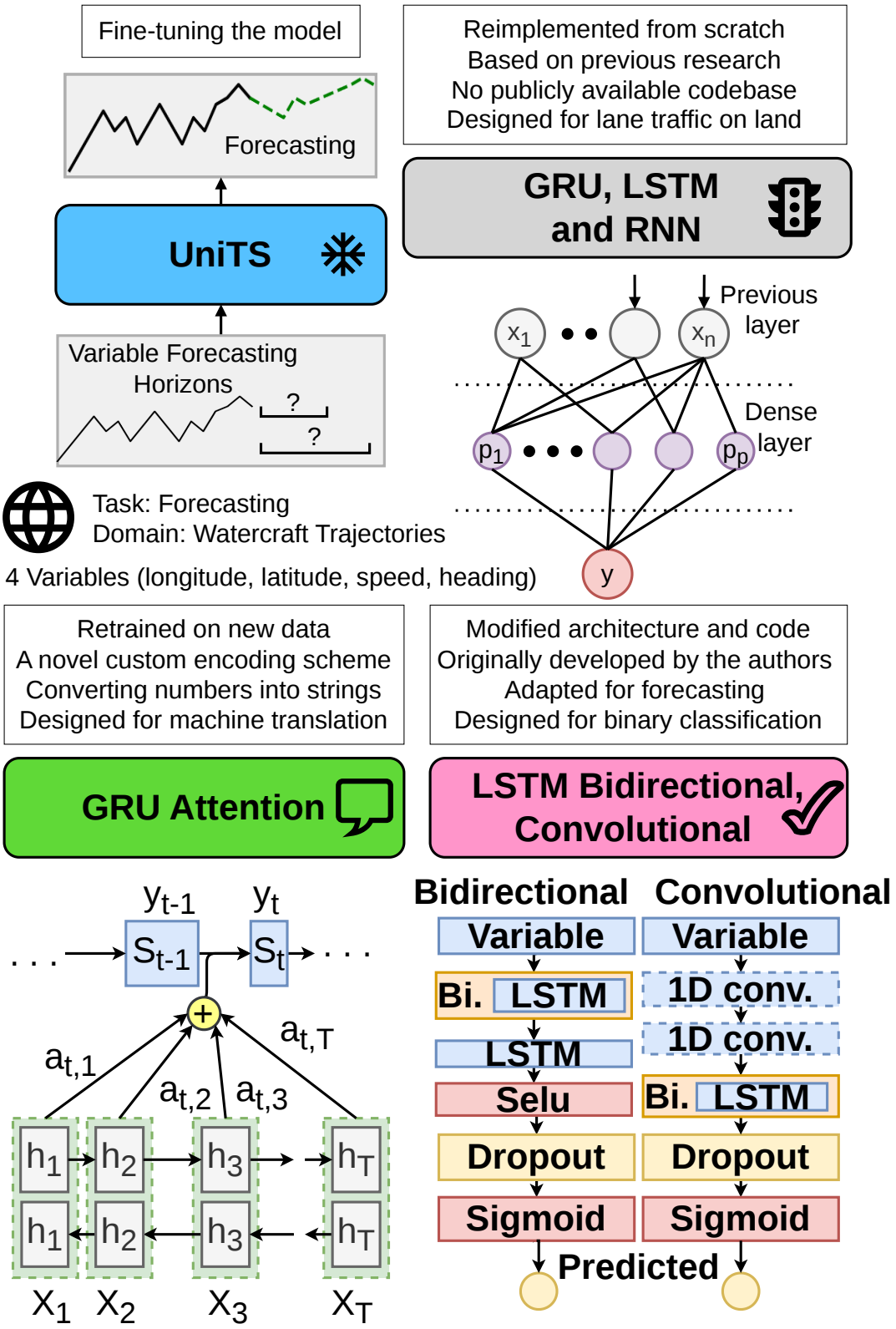

Figure 4.1: Schematic representation of the DL architectures employed, including the GRU, LSTM, and the UniTS foundation model [173, 174].

The top-left section shows the architecture of the UniTS foundation model [41], fine-tuned for forecasting. The structure of the GRU, LSTM [3, 145], and RNN [1] models in the top-right was reimplemented based on the textual description. This architectural design explanation used to write new code is sourced from prior research with proprietary code for land-based lane traffic [3]. The models use a feed-forward architecture in which the input nodes $(x_1 \dots x_n)$ from a previous layer feed into a dense layer $(p_1 \dots p_p)$. The bottom-left part details a GRU attention mechanism with an encoder-decoder structure, where the input states (y_{t-1}, y_t) are transformed into hidden states (S_{t-1}, S_t). This model was retrained on new data using a custom encoding scheme that converts numeric values to strings, since the model was originally designed for machine translation [44, 49, 134]. The bottom-right section presents two parallel architectures, named LSTM Bidirectional and LSTM Convolutional. The LSTM bidirectional architecture includes a bidirectional LSTM layer, SELU activation, dropout [40, 131], and a final layer with a sigmoid activation function to produce a predicted value. The LSTM convolutional architecture features two one-dimensional convolutional layers, a bidirectional LSTM layer, and dropout, and it includes a sigmoid activation function. The architecture and code are novel, as they were intended for binary peptide classification [93, 99] and were adapted for forecasting. The diagram in Figure 4.1 is a schematic overview of the DL models used in the research, highlighting their architectures, origins, and specific adaptations for trajectory forecasting tasks.

4.2.1 Recurrent neural networks and attention

Several RNN models were experimented on and compared. The list includes simple RNN models, LSTMs [3, 50], and GRUs [5, 6, 16]. LSTMs and GRUs approaches are especially good at this assignment, as gating allows information to be retained across long sequences. LSTMs and GRUs architecture design is targeted to address the vanishing gradient problem [42, 105].

To further enhance the model's ability to focus on relevant historical data, an attention mechanism is integrated [4, 49, 142]. In trajectory prediction, attention weights enable the model to dynamically prioritize different time steps in the input sequence. For example, when predicting a turn, the model might attend heavily to the most recent few seconds of heading change, whereas for speed prediction, it might average over a longer history.

4.2.2 Foundation models

The UniTS foundation model [41] was also evaluated. UniTS represents a paradigm shift in time-series forecasting [65], utilizing a unified, pre-trained transformer architecture [140] that can be adapted to new tasks via zero-shot learning or fine-tuning [139]. UniTS was applied to the maritime dataset to test the hypothesis that large-scale foundation models can generalize to

specific physical domains without extensive architectural modification. This aligns with recent trends in optimal starting points for forecasting [165], Propagated HIerarchical Learning Network (PHILNet) [55], financial forecasting [15, 125], weather prediction [73, 90], and broader traffic prediction reviews [14, 28, 128].

4.3 Forecasting performance results and discussion

The models were trained on a dataset covering 1282 rental sites globally. The results, published in Expert Systems with Applications [173], demonstrated a clear performance hierarchy. The UniTS model achieved the lowest RMSE for speed prediction on horizons exceeding 4 seconds and for position offsets (x, y) on long horizons (20-30 seconds), as seen in Table 4.1.

Table 4.1: The RMSE in $^\circ \times 10^{-4}$ for the x-offset and the y-offset, in km/h for the speed, and in $^\circ$ for the heading estimated on the testing dataset by different RNN models and long-term forecasting times. The bold values mark the lowest RMSE and the most successful models for each forecasting time and variable [173].

Variable	Model	5 s	10 s	20 s	30 s
x-offset	GRU Att 4	**4.829** $\times\mathbf{10^{-4}}$ °	**4.922** $\times\mathbf{10^{-4}}$ °	5.97 $\times 10^{-4}$ °	6.653 $\times 10^{-4}$ °
x-offset	UniTS	5.275 $\times 10^{-4}$ °	5.401 $\times 10^{-4}$ °	**5.822** $\times\mathbf{10^{-4}}$ °	**6.103** $\times\mathbf{10^{-4}}$ °
y-offset	GRU Att 4	**1.751** $\times\mathbf{10^{-4}}$ °	2.519 $\times 10^{-4}$ °	3.512 $\times 10^{-4}$ °	3.923 $\times 10^{-4}$ °
y-offset	GRU Reference	1.76 $\times 10^{-4}$ °	**2.438** $\times\mathbf{10^{-4}}$ °	3.158 $\times 10^{-4}$ °	3.441 $\times 10^{-4}$ °
y-offset	UniTS	1.998 $\times 10^{-4}$ °	2.44 $\times 10^{-4}$ °	**3.019** $\times\mathbf{10^{-4}}$ °	**3.339** $\times\mathbf{10^{-4}}$ °
speed	UniTS	**5.24 km/h**	**7.24 km/h**	**9.35 km/h**	**10.51 km/h**
heading	Bi	56.25°	**66.75**°	77.44°	**82.83**°
heading	Conv	**56.08**°	66.77°	77.36°	83.23°
heading	UniTS	57.75°	66.99°	**77.11**°	83.13°

This suggests that self-attention mechanisms in transformers better capture global temporal context than recurrent architectures. Robustness analyses of RNN models [160], distance correlation approaches [119], and Position Weight Deformable former (PWDformer) [150] support these findings. Similar advancements have been seen with Trend and Change-point Detection Transformer (TCD-former) [147], Feature-based Forecast Model Performance Prediction (FFORMPP) [135], Time Series Feature Extraction Library (TSFEL) [7], Time–Frequency enhanced Decomposed Network (TFDNet) [83], and Recurring Concept Drift (RCD) frameworks [43].

However, the superiority of the UniTS model is not guaranteed. The GRU Att 4 model achieved

the lowest *x*-offset RMSE for horizons of 5 and 10 seconds (4.829×10^{-4} ° and 4.922×10^{-4} °, respectively), the GRU Att 4 and the GRU Reference model achieved the lowest *y*-offset RMSE for horizons of 5 and 10 seconds (1.751×10^{-4} ° and 2.438×10^{-4} °, respectively), whereas the Conv and Bi models achieved the lowest RMSE for the heading at forecasting times of 5, 10, and 30 seconds (56.08°, 66.75°, and 82.83°, respectively).

In contrast to the findings on longer horizons, for shorter horizons and specific variable combinations, the GRU model with attention proved highly competitive across the models tested [95], even on highly dynamic input data [121], as demonstrated in Table 4.2. The GRU architecture balances accuracy and computational latency better than the larger UniTS model, making it a strong candidate for deployment on edge devices.

Table 4.2: The RMSE in ° $\times 10^{-4}$ for the *x*-offset and the *y*-offset, in km/h for the speed, and in ° for the heading estimated on the testing dataset by different RNN models and short-term forecasting times. The bold values mark the lowest RMSE and the most successful models for each forecasting time and variable [173].

Variable	Model	2 s	3 s	4 s
x-offset	GRU Att 2	4.482 $\times 10^{-4}$ °	4.582 $\times 10^{-4}$ °	**4.698** $\times$**10^{-4}** °
x-offset	GRU Att 4	**4.449** $\times$**10^{-4}** °	**4.564** $\times$**10^{-4}** °	4.718 $\times 10^{-4}$ °
y-offset	GRU Att 4	**1.146** $\times$**10^{-4}** °	**1.355** $\times$**10^{-4}** °	**1.551** $\times$**10^{-4}** °
speed	Conv	3.13 km/h	**3.85 km/h**	**4.61 km/h**
speed	LSTM Linear	**3.05 km/h**	3.87 km/h	4.67 km/h
heading	Bi	44.46°	**48.83**°	52.59°
heading	Conv	**43.82**°	48.83°	**52.22**°

The possible use of these findings can reach beyond a rental vessel system and is not confined to it. Foundation models, such as UniTS, can be mined for additional use cases across different modes of transport, both on land, at sea, and in the air [184], even for modeling humans and animals [27], increasing safety and enabling autonomous driving and navigation [171]. The research shows promising results and opens possibilities for real-world implementation after further adjustments and the addition of environmental considerations [183], such as wave height, ocean current, wind, and temperature, as well as the current state and flow of traffic density, to enable a further increase in resilient trajectory prediction for any eventuality.

4.4 Trajectory intersection detection

The best-performing forecasting models were applied to the intersection detection task, and False Negative (FN), True Negative (TN), False Positive (FP), and True Positive (TP) classifications were noted. Classification metrics derived from confusion matrices were used in the paper published in Knowledge-Based Systems [174], including accuracy, $F1$ score, sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), and Balanced Accuracy (BA). Statistical tests (McNemar's test [31]) were used to assess whether the models differed significantly from the ground truth, with the Bonferroni correction for multiple comparisons. In addition, the effect size was determined using Cohen's d. The goal was to predict whether two vessels would collide within a 30-second window. The GRU attention model (in the configuration "GRU Att 1") achieved a True Negative Rate (TNR) of 95.45% and a PPV of 99.19%, as can be calculated from the confusion matrix demonstrated in Figure 4.2.

This is a dramatic improvement over the shorter 5-second window used in the baseline IDC system, which, being reactive, generates a high volume of FP results. The IDC had a higher count and percentage of TP detections and a lower count and percentage of FN detections than all tested RNN models with a forecasting time of 30 seconds and 5 seconds for the x-offsets and y-offsets, equaling 192 and 89.72%, and 0 and 0%, for TP and FN detections, respectively. On the other hand, IDC also had a higher count and percentage of FP detections and a lower count and percentage of TN detections than all tested RNN models with a forecasting time of 30 seconds and 5 seconds for the x-offsets and y-offsets, equaling 22 and 10.28%, and 0 and 0%, for FP and TN detections, respectively. This reveals that RNN models can reduce false alerts and increase true alerts compared to basic linear assumptions that do not employ ML.

By accurately predicting that vessels on converging paths would likely miss each other (due to inferred turn rates), the ML model successfully suppressed unnecessary interventions, as shown in Figure 4.3. Statistical significance was assessed using the McNemar test [31] in the paper published in Knowledge-Based Systems [174], which showed distinct performance gains over simpler linear models.

In Figure 4.2, as published in Knowledge-Based Systems [174], the RNN Linear, RNN Reference, GRU Third, and GRU Twice models had the lowest count and percentage of TN detections and the highest count and percentage of FP detections among different RNN models with a prediction horizon of 30 seconds for the x-offsets and y-offsets, equaling 7 and 3.27%, and 15 and 7.01%, respectively. These architectures are much less successful than the GRU attention model variants.

The RNN Linear model had the highest count and percentage of TP detections and the lowest count and percentage of FN detections when compared to the UniTS foundation model and various RNN, GRU, LSTM, and attention-based models, using a forecasting time of 30 seconds

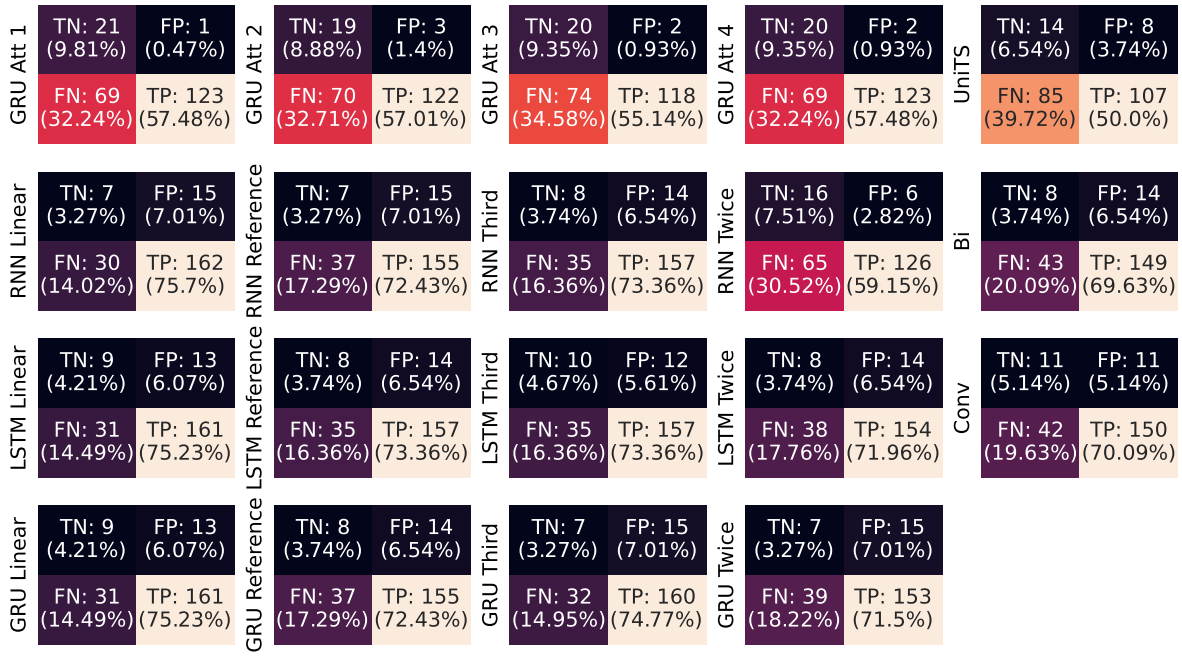

Figure 4.2: The confusion matrix heatmap with the number and percentage (%) of FN, TN, FP, and TP classifications for intersection detection using different RNN models and a forecasting time of 30 seconds for the longitude and latitude offsets [174].

for the x-offsets and y-offsets, equaling 162 and 75.7%, and 30 and 14.02%, respectively. However, the GRU attention model variants are superior for FP and TN instances, yielding fewer false alerts.

In Figure 4.3, as published in Knowledge-Based Systems [174], the LSTM Reference model had the lowest count and percentage of TN detections and the highest count and percentage of FP detections for intersection detection using UniTS, and various RNN, GRU, GRU attention, and LSTM models with a prediction horizon of 5 seconds for the x-offsets and y-offsets, equaling 9 and 4.21%, and 13 and 6.07%, respectively. This means that the LSTM Reference model will produce many false alerts, rendering it unsuitable for the proposed use case.

The UniTS models had the lowest count and percentage of TP detections and the highest count and percentage of FN detections among various RNN, GRU, GRU attention, and LSTM models with a prediction horizon of 30 seconds and 5 seconds for the x-offsets and y-offsets, equaling 107 and 50.0%, and 85 and 39.72%, for 30 s and TP and FN detections, respectively, and 103 and 48.13%, and 89 and 41.59%, for 5 s and TP and FN detections, respectively. This means that the UniTS model will miss many detections, and that the other attention-based approaches are preferable. The UniTS model proved inadequate for addressing the intersection-detection dilemma posed in this thesis, as it does not account for potentially dangerous behavior and would impose unnecessary restrictions on drivers in other scenarios.

The LSTM Reference and GRU Linear models had the highest count and percentage of TP detections and the lowest count and percentage of FN detections for intersection detection when

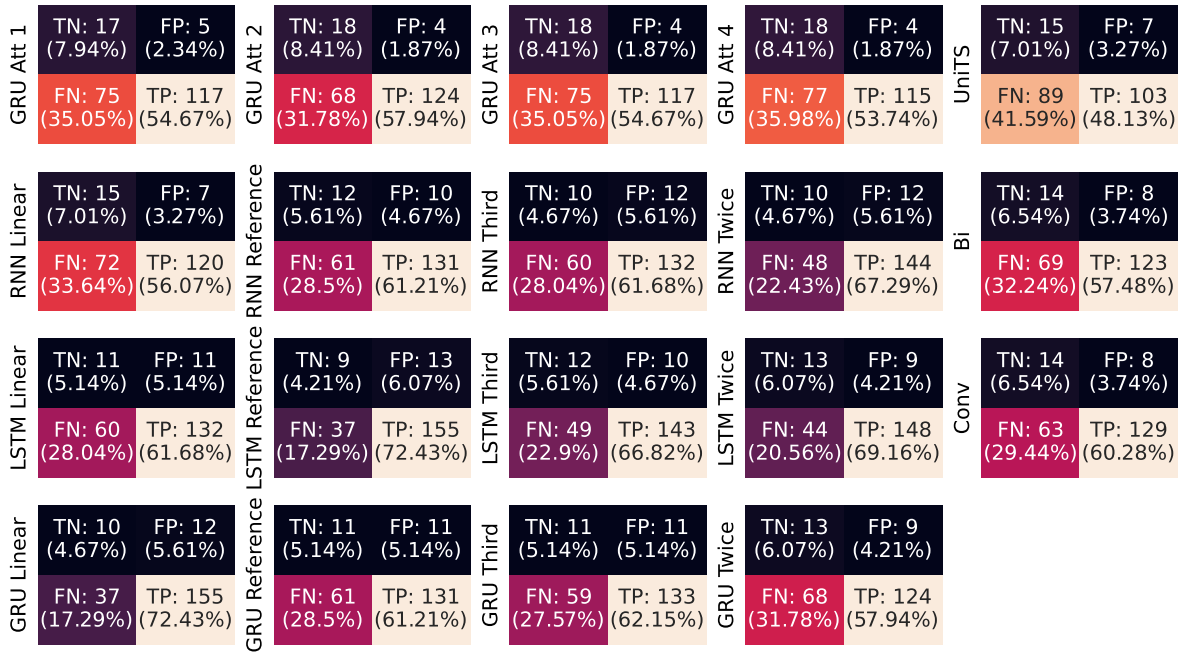

Figure 4.3: The confusion matrix heatmap with the number and percentage (%) of FN, TN, FP, and TP classifications for intersection detection using different RNN models and a forecasting time of 5 seconds for the longitude and latitude offsets [174].

compared to the UniTS foundation model and various RNN, GRU, LSTM, and attention-based models, using a forecasting time of 5 seconds for the x-offsets and y-offsets, equaling 155 and 72.43%, and 37 and 17.29%, respectively. On the other hand, the GRU attention model yields fewer FP and more TN detections, which are highly relevant to the desired application.

The numerical values presented here shed light on a new way to increase performance, especially in reducing false alarms. In events where swift intervention is key, such as changes in course or heading, the forecasting models accurately detected pairs of trajectories with no intersection, and the static IDC approach started an unnecessary warning. This adds validation to the crucial statement that the model of trajectories generated in this study can incorporate a more detailed representation of vessel dynamics than unchanging linear simplistic models.

The next chapter expands the field of study of trajectory patterns by introducing an anomaly-detection approach in addition to the previously described predictive models. This moves beyond forecasting to define categories that could help explain differences in human operators' driving styles and prevent accidents through early warnings integrated with adaptive systems.

Chapter 5

INFLECTION POINT-BASED TRAJECTORY FINGERPRINTING

5.1 Fingerprinting introduction

The increasing volume of Automatic Identification System (AIS) data has driven significant research into maritime situational awareness, particularly in trajectory mining and anomaly detection. The prior study by Pallotta et al. [102] established the Traffic Route Extraction and Anomaly Detection (TREAD) methodology. TREAD converts raw AIS data into an interpretable graph of maritime routes to identify off-route anomalies. Enhancing the graph-based domain, Coscia et al. used unsupervised learning that models maritime traffic as multiple mean-reverting stochastic processes. These enhancements are aimed at enabling a more granular understanding of vessel behavior across connected traffic networks [19, 20].

Most literature uses kinematic modeling to differentiate between normal navigation and anomalous deviations. Millefiori et al. [88] show that vessel kinematics models using mean-reverting processes are effective. Examples of mean-reverting processes include the OU process for long-term prediction, as further refined by Coraluppi et al. [18], who introduced the Mixed OU process to better exploit context in multi-target tracking. D'Afflisio et al. [22] also used OUs to distinguish anomalous changes and standard lanes. The models all share a statistical approach for marking abnormal maneuvers relative to common behavior.

Recent advancements have added Bayesian filtering approaches that increase detection resilience in the presence of noise and uncertainty. Forti et al. [34] proposed a Hybrid Bernoulli Filtering (HBF) approach for the joint detection and tracking of anomalous path deviations. This work was later elaborated upon to use mean-reverting processes with unknown parameters [35] and Random Finite Set (RFS) tracking to handle clutter in crowded maritime domain data [36]. These efforts finally resulted in a generalized Bayesian filtering framework for dynamic

anomaly detection [37].

Beyond theoretical modeling, these methodologies have been applied to specific security and safety scenarios. D'Afflisio et al. [21] extended detection capabilities to identify malicious data spoofing and stealth deviations by exploiting heterogeneous information sources. Practical applicability is not separated from anomaly detection approaches in theory, as described using the Suez Canal Ever Given grounding incident [33].

Given the abundant data available via the AIS ecosystem, as established in prior text, the information presented exceeds even the most optimistic projections. The largest database by a large margin comes from AIS, but there are also commercial GPS tracking solutions [57, 111]. Even though AIS is a public, detailed dataset for large vessels, usable in Maritime Situational Awareness (MSA) and traffic management [101], it does not adequately capture the dynamics of smaller, agile watercraft. Such vessels often are not required to follow the routes prescribed for large ships and instead navigate more independently of established shipping lanes [47, 133]. Consequently, anomaly detection for identifying criminal activities, navigation errors, or mechanical failures remains a significant gap for this vessel class [12, 85, 86], as confirmed in an extensive analysis of prior studies [69, 114].

DL models are accurate but often represent black boxes with high computational requirements. This chapter addresses the need for interpretable, lightweight anomaly detection methods suitable for embedded deployment on watercraft. A geometric approach based on trajectory fingerprinting is proposed to enable transparent, understandable anomaly detection [61].

5.2 Fingerprinting methodology

The core idea is to reduce a complex trajectory to a sequence of inflection points, representing critical moments when the vessel changes direction along the x or y axis. Trajectories are segmented into overlapping windows, and the sequence of inflection points is encoded into a compact hexadecimal string (the fingerprint) [71, 149]. This process, as outlined in Figure 5.1, greatly lowers the data dimensionality and maintains the crucial original kinematic topology.

The trajectory preprocessing begins by creating a uniform orientation for raw trajectory data. All trajectories are translated so that the starting point aligns with the origin of the Cartesian coordinate system $(0,0)$. The trajectory is rotated to result in the last point being exactly on the positive x-axis. To maintain positive values and consistency, any negative coordinate values are mirrored into the first quadrant.

Once preprocessed, the trajectories are segmented and features extracted. Trajectories are split into overlapping segments of specific window sizes, such as 10 or 20 consecutive points. The system identifies inflection points, defined as points where the direction of motion along either

the x or y-axis changes (from increasing to decreasing, or vice versa), as demonstrated in Figure 5.2, presenting the preprocessing and inflection point extraction, as well as Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and K-means clustering.

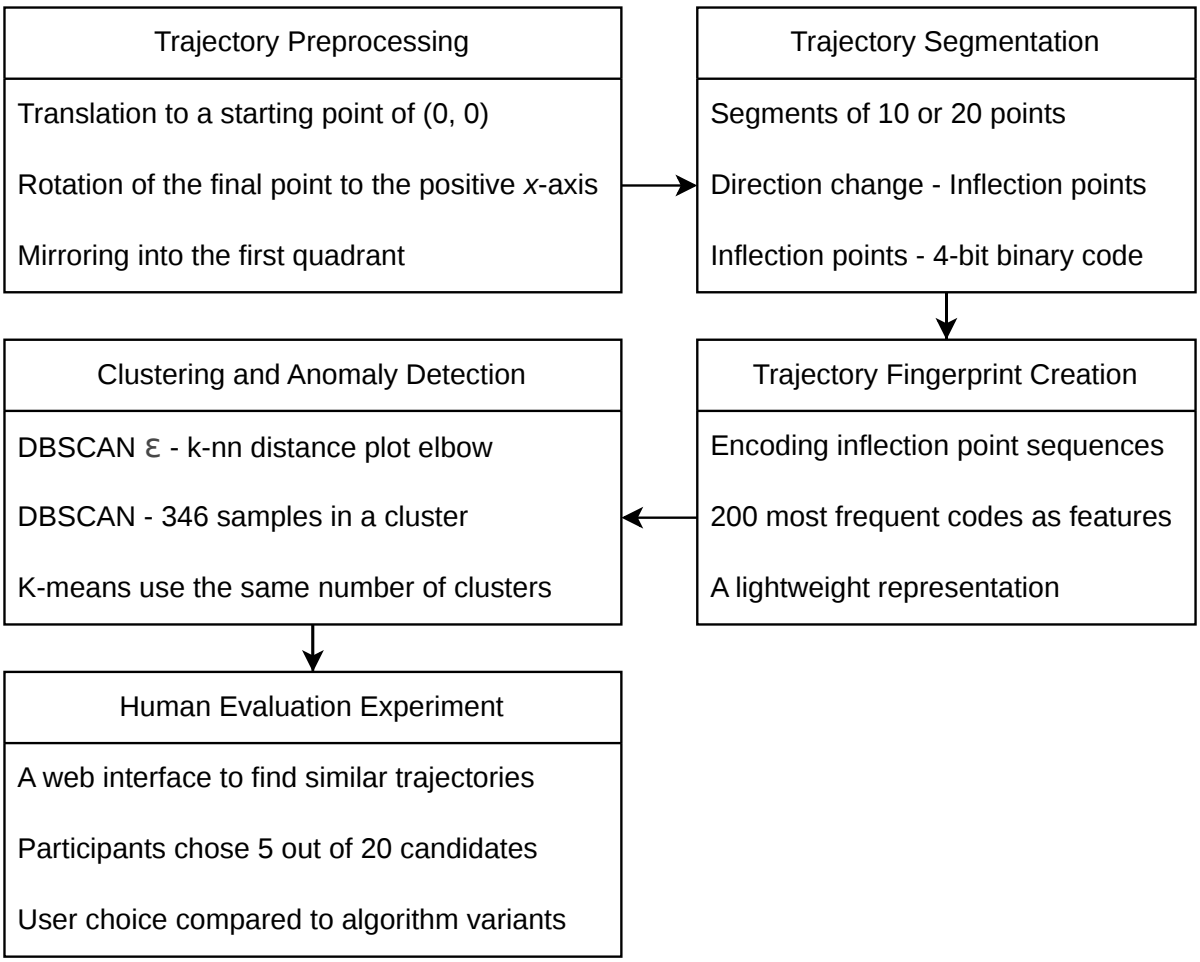

Figure 5.1: The outline for the sequential stages of the inflection point fingerprinting methodology, ranging from raw trajectory preprocessing and segmentation to fingerprint creation, clustering algorithm application (DBSCAN and K-means), and final validation through human evaluation experiments [176].

Every inflection point is transformed into a 4-bit binary code determined by the directional changes before and after the point. The direction at the previous and the next point can be marked as $[x_{before}, x_{after}, y_{before}, y_{after}]$. The extracted features are compactly represented. The 4-bit binary codes are converted into hexadecimal digits. The sequences of these hexadecimal digits form the trajectory fingerprint. This yields a lightweight representation that avoids the computational overhead associated with DL models.

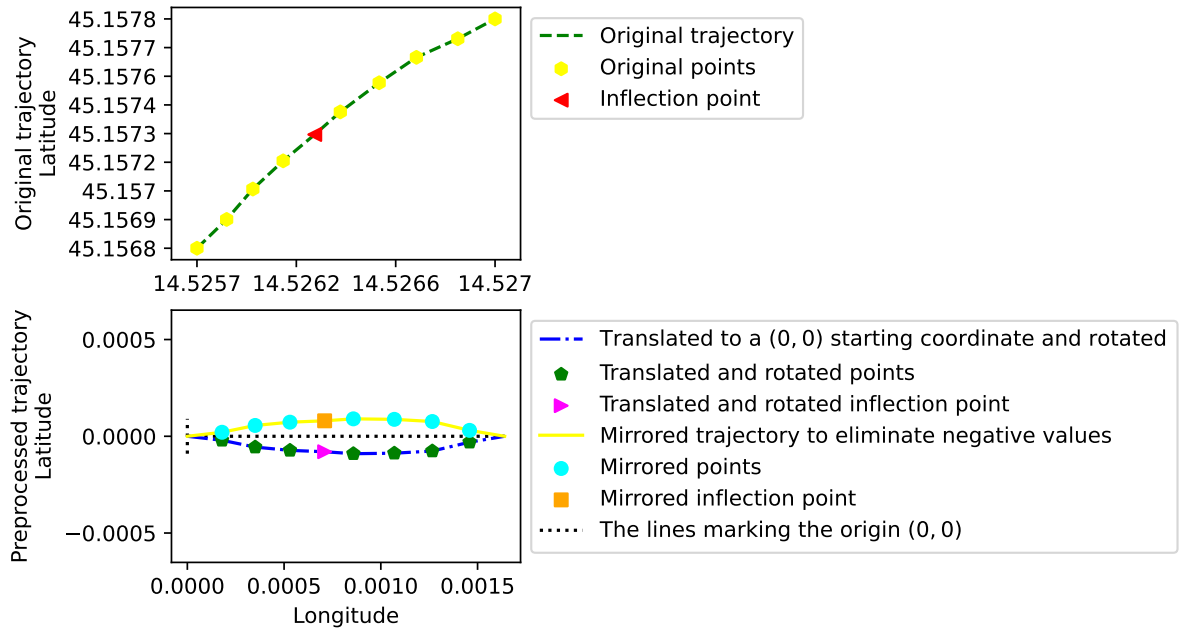

Figure 5.2: An illustration of the trajectory preprocessing and inflection point extraction pipeline. This results in a standardized input for fingerprinting. The raw trajectory is translated to the origin, rotated to align with the axes, and mirrored to eliminate negative values [176].

5.3 Fingerprinting results

The resulting fingerprints are required as input for clustering algorithms that detect kinematic anomalies. Unsupervised clustering algorithms, including K-means [47] and DBSCAN [29, 109, 123], are applied to these fingerprints. Anomalies are identified within the clusters, defined as trajectories where IDC was triggered. Analysis of elbow points in k-Nearest Neighbors (k-NN) distance plots [13] was used to determine optimal ε parameters for DBSCAN clustering of trajectory fingerprints, as shown in Figure 5.3.

To validate the solutions that humans submitted in the web experiment and the clustering and fingerprinting against the ground truth, a comparative experiment was conducted, inspired by the author's prior experience in web design [180, 182], and the results were published in Pattern Recognition [176]. A web interface presents participants with a baseline trajectory and 20 candidate trajectories. Participants select the 5 candidate trajectories most similar to the baseline to validate the method's alignment with human perception of kinematic similarity. User choices are compared against algorithm variants and ground-truth labels using confusion-matrix performance metrics [144].

To validate this approach, the algorithmic anomaly detection was compared with a survey of human experts and with the rule-based IDC logs. The results in Figure 5.4 showed remarkable convergence: the number of TP detections was largest, and the number of FN detections was

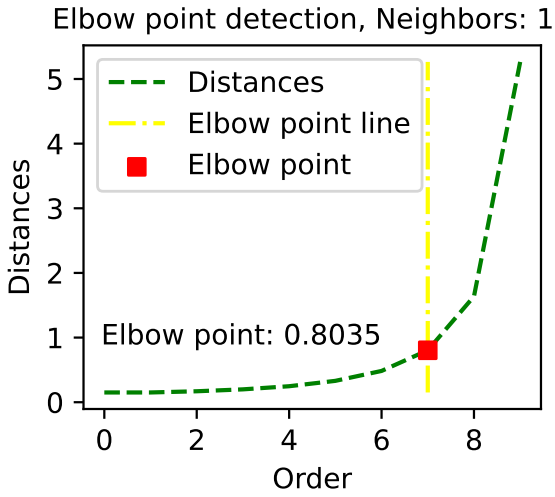

Figure 5.3: Analysis of elbow points in k-NN distance plots, used to determine optimal ε parameters for DBSCAN clustering of trajectory fingerprints [176].

smallest for DBSCAN across both window sizes, using IDC triggers on the personal watercraft dataset [98] as labels. Furthermore, the method successfully identified the same high-risk maneuvers flagged by the IDC system, using a fraction of the computational power required by DL models. This validates the method as a viable, interpretable alternative for real-time edge computing applications [58].

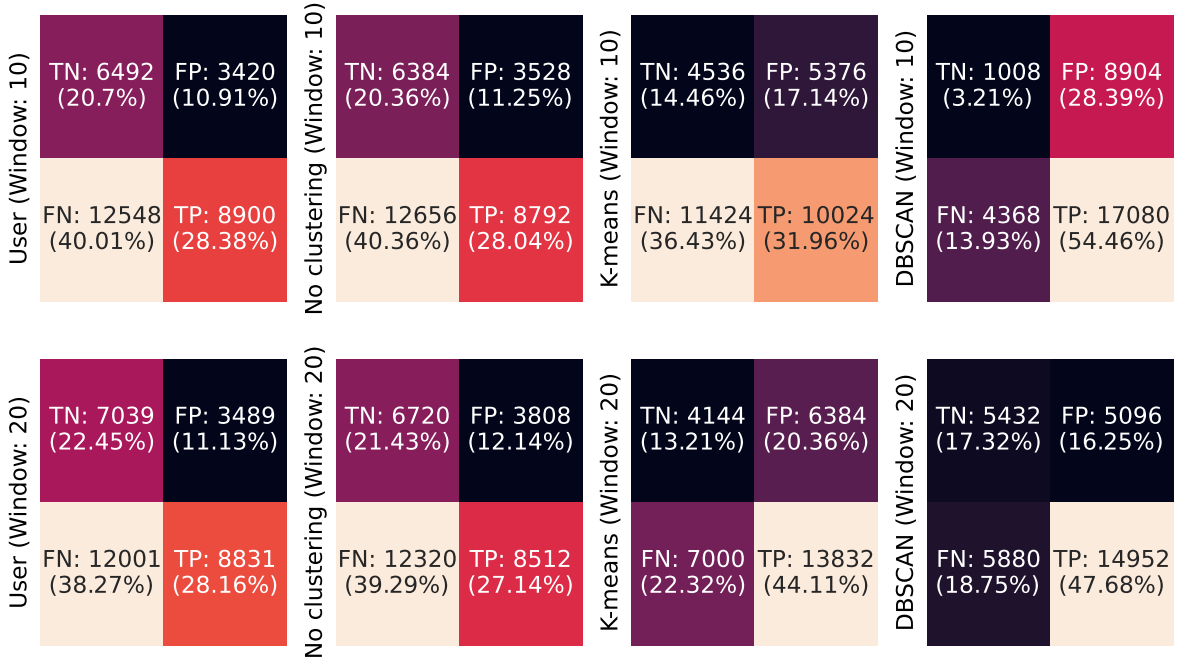

Figure 5.4: The confusion matrix heatmap for window sizes of 10 and 20 and the user's selection, algorithm without clustering, K-means, and DBSCAN compared to IDC triggers on the personal watercraft dataset [98].

This section of the thesis presents a computationally efficient, adaptable trajectory representa-

tion scheme that leverages inflection points for the detection and clustering of vessel kinematic anomalies. The methodology demonstrates enviable accuracy and high-speed operation across locations spanning multiple countries and continents, with the potential to incorporate AIS data and to support larger vessels that use indicators such as rate of turn (ROT) when changing heading. The approach is easy to understand and explain, executes quickly, and is suitable for real-world use on heavily restricted embedded systems. Its resilience was further confirmed via human user labels.

Moving on from the personal watercraft domain, the next chapter of the dissertation presents a methodology for underwater inspections, another avenue to contribute to maritime system operational safety and efficiency.

Chapter 6

U-NET SEMANTIC SEGMENTATION FOR UNDERWATER INFRASTRUCTURE INSPECTION

6.1 U-Net framework introduction

In addition to the vessel movements described in the preceding text, maritime safety also encompasses aspects related to the structural integrity of port infrastructure. The following chapter describes a computer vision framework for the automated inspection of underwater structures [169, 175, 179]. Underwater imagery poses unique challenges: turbidity, low light, and marine growth create a noisy visual environment in which standard terrestrial models, such as YOLO [59, 137], often fail to segment fine defects, such as cracks [127, 158].

A domain-adaptive segmentation framework based on a modified U-Net architecture [113] is proposed to address this issue. The most significant contribution is the separation of the multi-class problem. Separate, class-specific U-Net models are trained to replace a unified network to detect cracks, indentations, and obstacles simultaneously. This allows independent hyperparameter tuning, including decision thresholds and loss weights, for each defect type [116, 124, 152].

The proposed pipeline uses multi-domain data for testing, class-specific U-Net training with a Dice-based loss function compared to GMM and Spatially-Variant Gaussian Mixture Model (SVGMM) alternatives, and a custom post-processing pipeline, as shown in Figure 6.1. This diagram outlines the end-to-end pipeline for the underwater and Roboflow datasets, detailing the stages of preprocessing (grayscale conversion, normalization, and augmentation), class-specific U-Net model training with variable loss functions and regularization, and the final postprocessing and evaluation against benchmark models, such as YOLO.

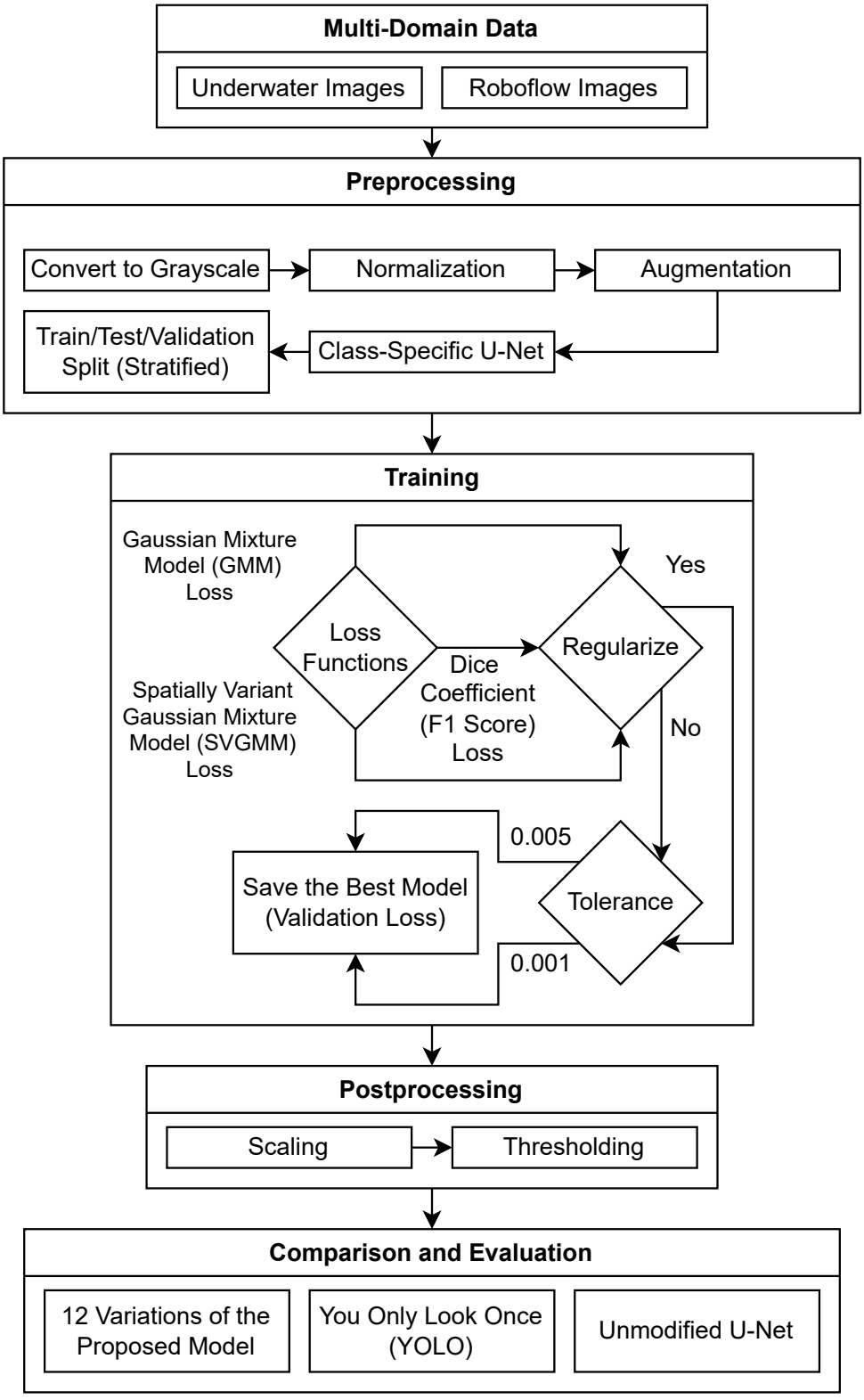

Figure 6.1: The proposed segmentation framework pipeline in the flowchart, which includes multi-domain data input, class-specific U-Net training with a Dice-based loss function, and a custom post-processing step [175].

6.2 U-Net framework methodology

The methodology includes data input, preprocessing [91], training, postprocessing, and evaluation.

The pipeline begins with the ingestion of multi-domain data, specifically underwater images and Roboflow images. Because there are significant differences in lightness and pixel intensity across object classes in these datasets, the images are first converted to grayscale.

Following conversion, the data is next normalized to the $[0,1]$ range [130]. The pipeline then applies augmentation modeled on the Roboflow crack dataset, including random rotations, flips, and adjustments to saturation and brightness, to triple the number of outputs per training example.

Unlike standard approaches that train a single network for all classes, this framework uses a class-specific U-Net, training a separate model for each class to accommodate overlapping labels and improve flexibility. The data is then split using a stratified train/test/validation split to preserve the class ratio and ensure consistent model evaluation.

The core training phase involves a decision process regarding loss functions and hyperparameters. The framework selects GMM and SVGMM, and the Dice Coefficient ($F1$ score) loss, which is designed to align directly with segmentation evaluation metrics. Unlike pixel-wise cross-entropy, which can be overwhelmed by the vast background class (healthy concrete), the Dice loss directly optimizes for the overlap between the predicted defect mask and the ground truth:

$$\mathscr{L}_{Dice} = 1 - \frac{2\sum_i p_i g_i}{\sum_i p_i + \sum_i g_i} \tag{6.1}$$

where p_i and g_i are the predicted and ground truth values for pixel i. This aligns with recent advancements in segmentation architectures, including Segment Anything Model (SAM) [63, 110], Deep Residual Learning [48], MobileNets [53], and specialized crack networks like Depth CrackNet [117].

The pipeline then decides whether to regularize the model using the mean pixel values from validation images. The tolerance threshold for early stopping can be set at either 0.001 or 0.005. The system is designed to save the best model based on the validation loss monitored at the end of each epoch.

Next, the model outputs undergo postprocessing following the success of model training and testing. This involves scaling the predictions to the $[0,1]$ range. Subsequently, the optimal threshold is used, as tuned on the Precision-Recall (PR) curve for each class to maximize the $F1$ score.

6.3 U-Net framework results and discussion

The final stage involves a comparative analysis, as published in Expert Systems with Applications [175]. Performance is evaluated using 12 variants of the proposed model (obtained by combining three loss functions, regularization options, and two tolerance thresholds), the YOLO segmentation model, and the unmodified U-Net (the baseline architecture).

The framework was evaluated on a proprietary underwater inspection dataset [64, 143] and the public Roboflow crack dataset [112] to test domain generalization. The results show that the proposed deepF1 model significantly outperforms YOLO and standard U-Net implementations.

For underwater images [143] in Figure 6.2, and Roboflow cracks [112] in Figure 6.3, the deepF1 model (with an early stopping tolerance of 0.001) achieved more TN and fewer FP classifications, a substantial improvement over the YOLO result. The class-specific training strategy proved particularly effective for indentations, a class with high visual variance that often confused the multi-class models. These findings confirm that in specialized, high-noise domains such as underwater inspection, tailored architectures with metric-aligned loss functions outperform generic object detectors. The work complements existing studies on Bayesian U-Net [23], dual encoder networks [159], Universal 3D Anomaly Detection (Uni-3DAD) [77], hybrid solar cell defect detection [24], ultra-lightweight Mamba networks [161], lightweight YOLO for cracks [153], and self-training approaches [129].

Figure 6.2, as presented in Expert Systems with Applications [175], shows a confusion matrix heatmap and annotates the average count and percentage of TP, TN, FP, and FN detections on cracks, indentations, and ropes in the proprietary underwater data [64, 143] for each tested U-Net variant, and the YOLO model [59].

The $deepSVG$ model in Figure 6.2, with regularization and a tolerance of 0.005, as presented in Expert Systems with Applications [175], has the highest average number and percentage of TP instances and the lowest average number and percentage of FN instances for all classes in the proprietary underwater dataset [64, 143], at 11133 and 6.43% and 690 and 0.4%, respectively, but it also has the lowest average number and percentage of TN instances and the highest average number and percentage of FP instances for all classes in the proprietary underwater dataset [64, 143], at 19667 and 11.36% and 141566 and 81.8%, respectively. This means that the $deepF1$ model is preferable because it yields fewer false alerts.

The $deepG$ model in Figure 6.2, with no regularization and a tolerance of 0.005, as presented in Expert Systems with Applications [175], has the lowest average number and percentage of TP instances and the highest average number and percentage of FN instances for all classes in the proprietary underwater dataset [64, 143], at 4138 and 2.39% and 7685 and 4.44%, respectively. This additionally validates the superiority of the $deepF1$ model, which has fewer missed detections.

Figure 6.2: A confusion matrix heatmap with the average count and percentage of TP, TN, FP, and FN pixel classifications for the YOLO model [59] and the proposed models [175] on the underwater image dataset [143].

Figure 6.3, as presented in Expert Systems with Applications [175], shows a confusion matrix heatmap and annotates the average count and percentage of TP, TN, FP, and FN detections on the public Roboflow above-water crack images [112] for each tested U-Net variant, and the YOLO model [59].

The *deepSVG* model in Figure 6.3, with regularization and a tolerance of 0.001, as presented in Expert Systems with Applications [175], has no FN instances and has the highest number and percentage of TP instances (3057 and 1.77%) on the Roboflow crack dataset [112], but it also has no TN instances and has the highest number and percentage of FP instances (169999 and 98.23%), so it should not be taken into account, given that the negative class was never assigned.

Among models that did not completely ignore negative labels in Figure 6.3, as presented in Expert Systems with Applications [175], YOLO has the highest average number and percentage of TP instances and the lowest average number and percentage of FN instances on the Roboflow surface crack images [112], at 2402 and 1.39%, and 655 and 0.38%, respectively. However, the

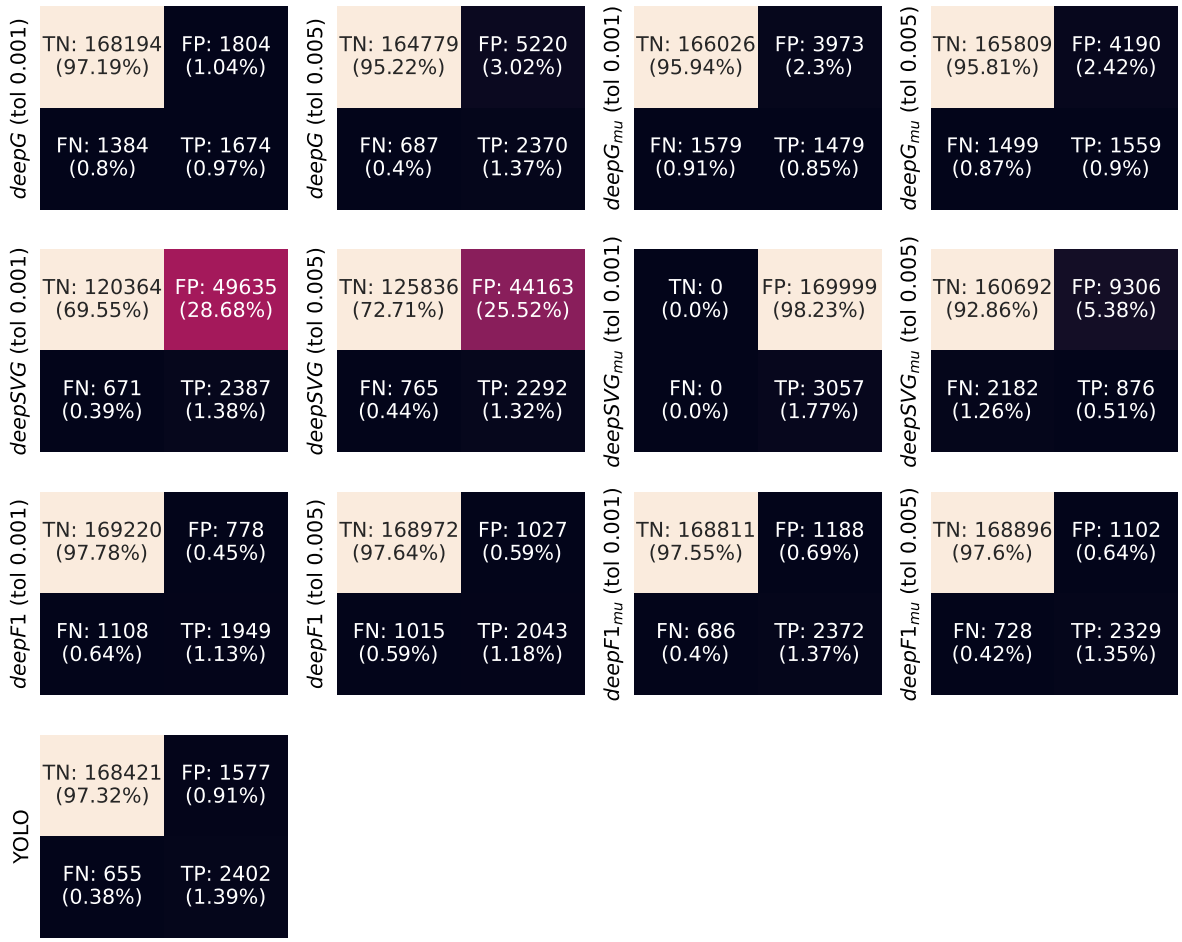

Figure 6.3: A confusion matrix heatmap with the average count and percentage of TP, TN, FP, and FN pixel classifications for the YOLO model [59] and the proposed models [175] on the Roboflow crack dataset [112].

proposed *deepF*1 model performs better in eliminating false alerts.

When ignoring the models that failed to detect the negative class in Figure 6.3, as presented in Expert Systems with Applications [175], the *deepSVG* model, with no regularization and a tolerance of 0.001, has the lowest average number and percentage of TN instances and the highest average number and percentage of FP instances on the Roboflow surface crack images [112], at 120364 and 69.55% and 49635 and 28.68%, respectively. This further supports the claim that the *deepF*1 model is the best choice, with fewer false detections.

The *deepG* model in Figure 6.3, with regularization and a tolerance of 0.001, as presented in Expert Systems with Applications [175], has the lowest average number and percentage of TP instances and the highest average number and percentage of FN instances on the Roboflow surface crack images [112], at 1479 and 0.85% and 1579 and 0.91%, respectively. This is additional evidence for the hypothesis that the *deepF*1 model is beneficial, as it has fewer missed detections.

The showcased numbers for diverse domains help pinpoint the model's contribution and the

implications of the novel pipeline, especially for the Dice ($F1$) score loss function used in model training. These results are the consequence of multiple crucial additions to the workflow, such as separating models for each class, facilitating independent training and preprocessing/postprocessing, and using a loss function that improves results for the utilized segmentation metrics.

The novel strategy is better than YOLO [59] in segmenting defect classes in underwater concrete inspection [64, 143] and is usable for real-world cases, and YOLO [59] degrades for small areas or structures that are not well defined. The increase in the $F1$ score and Intersection over Union (IoU) results demonstrates the novel pipeline's resilience, even without high-contrast inputs and despite significant noise, which are all inherent challenges in underwater images [64, 143].

The above-water scenario from Roboflow [112] is a vastly different domain with changes in contrast and material design. The proposed solution is better than YOLO [59] even in these new conditions [112], showing the possibility to extend across a single problem definition. These findings lead to the conclusion that the workflow and training pipeline outlined in this study are significantly modifiable and do not depend on a single narrow input definition.

Chapter 7

CONCLUSION

Most existing maritime research focuses on large vessels used for commercial shipping, which involve fixed routes and rich AIS data, whereas the focus of this thesis is on personal watercraft [98], which are fundamentally different. Operations are in free-navigation areas, not fixed lanes [101, 111]. Movement is irregular, reactive, and short-range, often influenced by human interaction. Local weather [115], or sea state features [30], play a significant role. Personal watercraft lack AIS and rely only on GPS data. The thesis aims to develop novel approaches that work in low-context environments without predefined routes.

In addition to predicting the movement of a single vessel, this dissertation tackles the issue of trajectory intersection detection to enhance maritime safety. Multi-agent prediction models [126, 162] allow for complex trajectory-interaction models, but existing approaches often rely on video or images of vessels in movement, which were not available. However, attention mechanisms [4] enhance focus on relevant time steps and sequence features, and encoder-decoder architectures [49] enhance the interpretability and precision, so these approaches were implemented in the conducted research. Existing models for structured land traffic [3] and older models for maritime shipping lanes [74, 106] were combined to form the inspiration for the proposed classification of trajectories as intersecting or non-intersecting based on paired trajectory prediction.

Additionally, to identify unsafe deviations from expected patterns, this thesis uses DBSCAN [81] to group similar trajectories, successfully overcoming the difficulties of parameter tuning. Other authors have used Bayesian networks [86] to detect deviations, but these methods may require manual route encoding. Similarly, context-based approaches [101] assess route deviation, but they do not account for free navigation in personal watercraft models. This thesis aims to develop an interpretable anomaly-detection system using inflection-point encoding that aligns with human judgment and operates without predefined rules.

The study uses GPS trajectories from the OtoTrak [98] system across global locations, with a dataset comprising timestamps, geolocation, speed, heading, and IDC markers indicating

safety-system activations and speed throttling to prevent accidents. Trajectories were normalized, and further steps involved rotating the endpoint to align with the x-axis and mirroring to enable location-agnostic global model development. Removing trajectories with large time gaps and segmenting into forecasting windows was also necessary in the preprocessing approach.

Predictions relied on forecasts with different time steps, and the next step in the proposed approach is benchmarking Bayesian and Markov models with RNN, LSTM [3], and GRU models, GRU attention architectures [49], LSTM convolutional and bidirectional model [93], and the UniTS foundation model [41]. Forecasting models were evaluated based on RMSE and the R^2 coefficient.

Predictions of the positions of two vessels across various forecasting windows, along with the spatial intersection of their predicted paths, were used to identify potential collisions, and validation was conducted using IDC activations and actual outcomes [177].

The classification of a single trajectory as normal or anomalous [101] proposed in this thesis uses inflection point encoding (changes in the x-direction or y-direction), and 4-bit binary strings that later form a hexadecimal trajectory fingerprint, and trajectories are finally clustered using DBSCAN [81] and K-means, while also determining optimal ε values for DBSCAN using the elbow point in k-NN graphs. The ground-truth values use IDC trigger events as labels for anomalies, and a human web experiment, in which users mark similar trajectories, was incorporated to assess alignment with human judgment.

The final component of this thesis uses computer vision for infrastructure inspection and benchmarks the standard U-Net [113] and YOLO [59] architectures, both of which are powerful tools for image segmentation. As monolithic models struggle with generalizability [124], and different water conditions and lighting pose significant challenges for underwater imaging, modular class-specific U-Net models were proposed in this study to achieve higher precision and robustness in both underwater [64, 143] and surface [112] data, two vastly different data domains. Training and testing images in total covered proprietary underwater images [64, 143] and Roboflow images with surface cracks [112]. Image preprocessing involved grayscale conversion, normalization, and resizing [112, 166], as well as augmentation, including flips, rotations, crops, hue, saturation, brightness, and blurring. Image postprocessing involved scaling the output and applying the proposed classification threshold for each class, as derived from the validation Precision-Recall curve, to maximize the $F1$ score. The U-Net [113] model was trained using the Dice ($F1$ score) loss function, as well as GMM and SVGMM [124] loss.

This thesis has successfully demonstrated that the proposed data preprocessing approach enables location-agnostic modeling, allowing trajectory forecasting models [3] to generalize across diverse locations without requiring site-specific retraining. The developed Bayesian and Markov chain models rely on trajectory data while also incorporating environmental variables. These include meteorological and ocean-state information, such as wind speed from METAR [115]

and sea-state data from the Copernicus Marine Service [30]. By integrating these otherwise distinct datasets, the models are able to generate conditional probabilities that reflect changing natural conditions, including wave height and temperature. When evaluating the x-offset and y-offset estimates with a forecasting horizon of 1 second, 2 seconds, or 4 seconds, the Bayesian model with two previous states with real instead of estimated values as input achieved the lowest RMSE (1.504×10^{-4} °, 3.735×10^{-4} °, and 8.593×10^{-4} ° for the x-offset, and 1.105×10^{-4} °, 2.588×10^{-4} °, 6.382×10^{-4} ° for the y-offset), while one prior state was the best for the Bayesian model and a forecasting horizon of 3 seconds (7.474×10^{-4} ° for the x-offset and 5.387×10^{-4} ° for the y-offset), indicating that Markov chain approaches underperfom due to error accumulation and that environmental variables did not significantly contribute to model performance [170] in short-term forecasting in Table 3.1. Long-term forecasting in Table 3.2 was inadequate in all cases tested in the paper published in the Journal of the Franklin Institute [170], motivating the shift towards DL.

The achieved results showed that applying the proposed data preprocessing steps, including normalization, rotation, and mirroring, helped develop efficient location-agnostic models for watercraft trajectory prediction. This approach enabled the deployment of DL models, including RNN variants [3, 49] and the UniTS [41] foundation model across global locations without site-specific retraining. The UniTS model achieves the lowest RMSE for the x-offset and y-offset when the forecasting time is at least 20 seconds (5.822×10^{-4} ° and 6.103×10^{-4} ° for the x-offset and 20 and 30 seconds, respectively, and 3.019×10^{-4} ° and 3.339×10^{-4} ° for the y-offset and 20 and 30 seconds, respectively), for the speed when the forecasting time is at least 5 seconds (5.24 km/h, 7.24 km/h, 9.35 km/h and 10.51 km/h for 5, 10, 20, and 30 seconds, respectively), and for the heading when the forecasting time is 20 seconds (77.11°). The results published in Expert Systems with Applications demonstrate [173] that the UniTS model should be used instead of the proposed alternative approaches in further development, particularly for long-term forecasting (Table 4.1), which is the ultimate goal, rather than short-term forecasting (Table 4.2).

The presented research validates that DL-based intersection detection, by using attention mechanisms and encoder-decoder architectures, greatly improves upon prior rule-based systems [98]. The DL-based intersection detection results, which strongly favor the variant using attention mechanisms [49], demonstrate that it significantly outperforms traditional rule-based systems by reducing FP occurrences, thereby enhancing user experience without compromising safety. This approach enhances precision in identifying intersecting courses by learning sequential dependencies in complex, high-variance personal watercraft trajectories [98]. The GRU Att 1 model in Figure 4.2, as published in Knowledge-Based Systems [174], had the highest count and percentage of TN intersection detections and the lowest count and percentage of FP intersection detections when compared to the UniTS foundation model and various RNN, GRU, LSTM, and attention-based models, using a forecasting time of 30 seconds for the x-offsets and

y-offsets, equaling 21 and 9.81%, and 1 and 0.47%, respectively. The GRU Att 2, GRU Att 3, and GRU Att 4 models in Figure 4.3, as published in Knowledge-Based Systems [174], had the highest count and percentage of TN intersection detections and the lowest count and percentage of FP intersection detections for intersection detection among different RNN models with a forecasting time of 5 seconds for the *x*-offsets and *y*-offsets, equaling 18 and 8.41%, and 4 and 1.87%, respectively. As reducing the number of redundant speed-throttling events was the main objective of this research, the GRU attention models are all strong candidates, outperforming the basic IDC approach that yielded 22 FP detections.

This thesis also introduces an interpretable trajectory fingerprinting method that lowers the demands of complex ML and brings the processing speed down to the requirements of lightweight embedded logic. The inflection-point methodology pioneered by the authors for real-time anomaly detection is applicable to new locations and unexplored environments not present in the training data and can be employed to achieve expanded coverage with no additional adjustments, ensuring that the framework is ready for any eventuality an untrained and risk-prone tourist with an adventurous spirit might encounter in a new geographic area. The research produced a novel, interpretable trajectory fingerprinting method based on inflection point encoding by representing directional changes as hexadecimal notation. This method reduces the cost of complex ML and adapts to lightweight embedded logic, offering a robust, real-time anomaly-detection solution that aligns with human judgment and operates without predefined routes. The confusion matrix heatmap for window sizes of 10 and 20 and the user's selection, algorithm without clustering, K-means, and DBSCAN in Figure 5.4, as presented in Patter Recognition [176], proves the superiority of the proposed DBSCAN clustering based on trajectory fingerprints leveraging inflection points [176], as there are more TP labels and fewer FN labels for window sizes of 10 (17080 and 54.46% for TP, 4368 and 13.93% for FN) and 20 (14952 and 47.68% for TP, 5880 and 18.75% for FN) points, with IDC activation on the OtoTrak dataset [98] used as the ground truth.

A novel specialized U-Net segmentation framework establishes a new baseline for underwater infrastructure inspection [158], outperforming universal models thanks to class-specific training and Dice-based loss functions. A custom U-Net segmentation framework was developed, and this creates a new benchmark for underwater infrastructure inspection [158]. Deviating from monolithic models, this pipeline uses modular, class-specific U-Net models trained with a Dice ($F1$) score loss function to handle varying water conditions and lighting, outperforming generic multi-class models [59, 137]. The $deepF1$ model in Figure 6.2, with regularization and a tolerance of 0.001, as presented in Expert Systems with Applications [175], has the highest average number and percentage of TN instances and the lowest average number and percentage of FP instances for all classes in the proprietary underwater dataset [64, 143], at 158165 and 91.4% and 3068 and 1.77%, respectively, outperforming YOLO and GMM and SVGMM loss U-Net variants. The $deepF1$ model in Figure 6.3, with no regularization and a tolerance of 0.001, as

presented in Expert Systems with Applications [175], has the highest average number and percentage of TN instances and the lowest average number and percentage of FP instances for the above-water crack images [112], at 169220 and 97.78% and 778 and 0.45%, respectively, once again outshining YOLO and GMM and SVGMM loss U-Net variants in a different domain.

Future work

Future studies may address the applicability and the use of these models in real-world maritime safety systems. The problems that remain unsolved include variable transmission delays in real-time streams. Another area for improvement is expanding datasets to include extreme weather conditions. Additionally, the concept of continual learning, where models update themselves based on new data without forgetting old patterns, remains a frontier for deployment in the ever-changing maritime environment. The options for further development include the listed crucial considerations:

- **Real-time system integration:** Moving from offline evaluation to inference benchmarking and full integration into real-time monitoring systems. The greatest effort is required to solve the problems relating to variable transmission delays in real-time streams and optimizing latency for rapid warnings in case of danger.
- **Location agnostic model generalization:** Further exploring the transferability of the prediction capabilities of all tested models, including prompt-token foundation models like UniTS. Future work should test the architecture's ability to generalize to unseen vessel types and behaviors without extensive fine-tuning.
- **Dataset expansion and extreme conditions:** Expanding the datasets to encompass extreme weather conditions and a wider range of anomaly types beyond the present scope of IDC trigger events. This includes validating models against a wider range of environmental data sources.
- **Continual learning deployment:** Continual learning frameworks implement models that update periodically based on new trajectory data without disregarding previously established patterns. This is needed to preserve accuracy in the constantly evolving maritime environment and adapt to new traffic regulations or recreational zones.

Abbreviations

The following abbreviations are used in this dissertation:

ADAS Advanced Driver Assistance System

AIS Automatic Identification System

ATC Automatic Tour Control

BA Balanced Accuracy

CS-YOLO Crack Segment You Only Look Once

CSPNet Cross Stage Partial Network

DBSCAN Density-Based Spatial Clustering of Applications with Noise

DL Deep Learning

EU European Union

FFORMPP Feature-based Forecast Model Performance Prediction

FN False Negative

FP False Positive

FPR False Positive Rate

GMM Gaussian Mixture Model

GNN Graph Neural Network

GPS Global Positioning System

GRU Gated Recurrent Unit

GSM Global System for Mobile Communications

HBF Hybrid Bernoulli Filtering

IDC Intelligent Distance Control

IoU Intersection over Union

IP Ingress Protection

ITS Intelligent Transportation System

k-NN k-Nearest Neighbors

KDE Kernel Density Estimation

LSTM Long Short-Term Memory

METAR Meteorological Aerodrome Report

ML Machine Learning

MSA Maritime Situational Awareness

NPV Negative Predictive Value

OU Ornstein-Ühlenbeck

PHILNet Propagated HIerarchical Learning Network

PPV Positive Predictive Value

PR Precision-Recall

PWDformer Position Weight Deformable former

RCD Recurring Concept Drift

RFS Random Finite Set

RMSE Root Mean Squared Error

RNN Recurrent Neural Network

ROT rate of turn

ROV Remotely Operated Vehicle

SAM Segment Anything Model

SELU Scaled Exponential Linear Unit

SVGMM Spatially-Variant Gaussian Mixture Model

TCD-former Trend and Change-point Detection Transformer

TFDNet Time–Frequency enhanced Decomposed Network

TN True Negative

TNR True Negative Rate

TP True Positive

TREAD Traffic Route Extraction and Anomaly Detection

TSFEL Time Series Feature Extraction Library

Uni-3DAD Universal 3D Anomaly Detection

UniTS Unified Time Series

YOLO You Only Look Once

Bibliography

[1] Agarap, A. F. Statistical Analysis on E-Commerce Reviews, with Sentiment Classification using Bidirectional Recurrent Neural Network (RNN). *arXiv preprint arXiv:1805.03687* (June 2020). eprint: https://doi.org/10.48550/arXiv.1805.03687.

[2] Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., and Savarese, S. Social LSTM: Human Trajectory Prediction in Crowded Spaces. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Las Vegas, NV, USA, June 2016), IEEE, pp. 961–971. eprint: https://doi.org/10.1109/CVPR.2016.110.

[3] Altché, F., and de La Fortelle, A. An LSTM network for highway trajectory prediction. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)* (Yokohama, Japan, Oct. 2017), IEEE, pp. 353–359. eprint: https://doi.org/10.1109/ITSC.2017.8317913.

[4] Bahdanau, D., Cho, K., and Bengio, Y. Neural Machine Translation by Jointly Learning to Align and Translate. *arXiv preprint arXiv:1409.0473* (May 2016). eprint: https://doi.org/10.48550/arXiv.1409.0473.

[5] Bao, K., Bi, J., Gao, M., Sun, Y., Zhang, X., and Zhang, W. An Improved Ship Trajectory Prediction Based on AIS Data Using MHA-BiGRU. *Journal of Marine Science and Engineering 10*, 6 (June 2022), 804. eprint: https://doi.org/10.3390/jmse10060804.

[6] Bao, K., Bi, J., Ma, R., Sun, Y., Zhang, W., and Wang, Y. A Spatial-Reduction Attention-Based BiGRU Network for Water Level Prediction. *Water 15*, 7 (Mar. 2023), 1306. eprint: https://doi.org//10.3390/w15071306.

[7] Barandas, M., Folgado, D., Fernandes, L., Santos, S., Abreu, M., Bota, P., Liu, H., Schultz, T., and Gamboa, H. TSFEL: Time Series Feature Extraction Library. *SoftwareX 11* (Jan. 2020), 100456. eprint: https://doi.org/10.1016/j.softx.2020.100456.

[8] Bertazzi, L., Chagas, G. O., Coelho, L. C., Laganà, D., and Vocaturo, F. Online algorithms for the multi-vehicle inventory-routing problem with real-time demands. *Trans-*

portation Research Part C: Emerging Technologies 170 (Jan. 2025), 104892. eprint: https://doi.org/10.1016/j.trc.2024.104892.

[9] Botts, C. H. A Novel Metric for Detecting Anomalous Ship Behavior Using a Variation of the DBSCAN Clustering Algorithm. *SN Computer Science 2*, 5 (Aug. 2021), 412. eprint: https://doi.org/10.1007/s42979-021-00804-4.

[10] Bradley, J. V. Complete Counterbalancing of Immediate Sequential Effects in a Latin Square Design. *Journal of the American Statistical Association 53*, 282 (June 1958), 525–528. eprint: https://doi.org/10.1080/01621459.1958.10501456.

[11] Carvalho, A., Gao, Y., Lefevre, S., and Borrelli, F. Stochastic Predictive Control of Autonomous Vehicles in Uncertain Environments. In *12th International Symposium on Advanced Vehicle Control (AVEC)* (Tokyo, Japan, Sept. 2014), AVEC, pp. 1–8. eprint: https://www.researchgate.net/publication/263308694_Stochastic_predictive_control_of_autonomous_vehicles_in_uncertain_environments.

[12] Chandola, V., Banerjee, A., and Kumar, V. Anomaly detection: A survey. *ACM Computing Surveys 41*, 3 (July 2009), 1–58. eprint: https://doi.org/10.1145/1541880.1541882.

[13] Chen, J., Zhang, X., and Yuan, Z. Feature selections based on two-type overlap degrees and three-view granulation measures for k-nearest-neighbor rough sets. *Pattern Recognition 156* (Dec. 2024), 110837. eprint: https://doi.org/10.1016/j.patcog.2024.110837.

[14] Chen, X., and Chen, R. A Review on Traffic Prediction Methods for Intelligent Transportation System in Smart Cities. In *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (Suzhou, China, Oct. 2019), IEEE, pp. 1–5. eprint: https://doi.org/10.1109/CISP-BMEI48845.2019.8965742.

[15] Cheng, D., Yang, F., Xiang, S., and Liu, J. Financial time series forecasting with multi-modality graph neural network. *Pattern Recognition 121* (Jan. 2022), 108218. eprint: https://doi.org/10.1016/j.patcog.2021.108218.

[16] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Doha, Qatar, Oct. 2014), Association for Computational Linguistics, pp. 1724–1734. eprint: https://doi.org/10.3115/v1/D14-1179.

[17] Chondrodima, E., Pelekis, N., Pikrakis, A., and Theodoridis, Y. An Efficient LSTM Neural Network-Based Framework for Vessel Location Forecasting. *IEEE Transactions on Intelligent Transportation Systems 24*, 5 (May 2023), 4872–4888. eprint: https://doi.org/10.1109/TITS.2023.3247993.

[18] Coraluppi, S., Carthel, C., Braca, P., and Millefiori, L. The Mixed Ornstein-Uhlenbeck Process and context exploitation in multi-target tracking. In *2016 IEEE 19th International Conference on Information Fusion* (Heidelberg, Germany, July 2016), IEEE, pp. 217–224. eprint: https://ieeexplore.ieee.org/document/7527891.

[19] Coscia, P., Braca, P., Millefiori, L. M., Palmieri, F. A. N., and Willett, P. Multiple Ornstein–Uhlenbeck Processes for Maritime Traffic Graph Representation. *IEEE Transactions on Aerospace and Electronic Systems 54*, 5 (Feb. 2018), 2158–2170. eprint: https://doi.org/10.1109/TAES.2018.2808098.

[20] Coscia, P., Palmieri, F. A. N., Braca, P., Millefiori, L. M., and Willett, P. Unsupervised Maritime Traffic Graph Learning with Mean-Reverting Stochastic Processes. In *2018 IEEE 21st International Conference on Information Fusion* (Cambridge, UK, July 2018), IEEE, pp. 1822–1828. eprint: https://doi.org/10.23919/ICIF.2018.8455392.

[21] D'Afflisio, E., Braca, P., Chisci, L., Battistelli, G., and Willett, P. Maritime Anomaly Detection of Malicious Data Spoofing and Stealth Deviations from Nominal Route Exploiting Heterogeneous Sources of Information. In *2021 IEEE 24th International Conference on Information Fusion* (Sun City, South Africa, Nov. 2021), IEEE, pp. 1–7. eprint: https://doi.org/10.23919/FUSION49465.2021.9627049.

[22] D'Afflisio, E., Braca, P., Millefiori, L. M., and Willett, P. Detecting Anomalous Deviations From Standard Maritime Routes Using the Ornstein–Uhlenbeck Process. *IEEE Transactions on Signal Processing 66*, 24 (Oct. 2018), 6474–6487. eprint: https://doi.org/10.1109/TSP.2018.2875887.

[23] Dechesne, C., Lassalle, P., and Lefèvre, S. Bayesian U-Net: Estimating Uncertainty in Semantic Segmentation of Earth Observation Images. *Remote Sensing 13*, 19 (Sept. 2021), 3836. eprint: https://doi.org/10.3390/rs13193836.

[24] Demirci, M. Y., Beşli, N., and Gümüşçü, A. An improved hybrid solar cell defect detection approach using Generative Adversarial Networks and weighted classification. *Expert Systems with Applications 252*, Part A (Oct. 2024), 124230. eprint: https://doi.org/10.1016/j.eswa.2024.124230.

[25] Ding, C., Wang, W., Wang, X., and Baumann, M. A Neural Network Model for Driver's Lane-Changing Trajectory Prediction in Urban Traffic Flow. *Mathematical Problems*

in Engineering 2013 (Feb. 2013), 967358. eprint: https://doi.org/10.1155/2013/967358.

[26] Duan, Y., Yisheng, L. V., and Wang, F. Y. Travel time prediction with LSTM neural network. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)* (Rio de Janeiro, Brazil, Nov. 2016), IEEE, pp. 1053–1058. eprint: https://doi.org/10.1109/ITSC.2016.7795686.

[27] Dumenčić, S., Žužić, L., and Breš, B. Trajectory Analysis of Griffon Vulture in Kvarner Bay. *The Journal of CIEES 2*, 2 (May 2023), 36–43. eprint: https://doi.org/10.48149/jciees.2022.2.2.6.

[28] Elmazi, K., Elmazi, D., Musta, E., Mehmeti, F., and Hidri, F. An Intelligent Transportation Systems-Based Machine Learning-Enhanced Traffic Prediction Model using Time Series Analysis and Regression Techniques*. In *2024 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)* (Craiova, Romania, Sept. 2024), IEEE, pp. 1–6. eprint: https://doi.org/10.1109/INISTA62901.2024.10683864.

[29] Ester, M., Kriegel, H. P., Sander, J., and Xu, X. A density-based algorithm for discovering clusters in large spatial databases with noise. In *KDD'96: Proceedings of the Second International Conference on Knowledge Discovery and Data Minin* (Portland, OR, USA, Aug. 1996), AAAI Press, pp. 226–231. eprint: https://dl.acm.org/doi/10.5555/3001460.3001507.

[30] European Union-Copernicus Marine Service. Global Ocean Waves Reanalysis WAVERYS. https://doi.org/10.48670/MOI-00022, Nov. 2024. Accessed: 22.05.2025.

[31] Fagerland, M. W., Lydersen, S., and Laake, P. The McNemar test for binary matched-pairs data: mid-p and asymptotic are better than exact conditional. *BMC Medical Research Methodology 13* (July 2013), 91. eprint: https://doi.org/10.1186/1471-2288-13-91.

[32] Fisher, T., Gibson, H., Liu, Y., Abdar, M., Posa, M., Salimi-Khorshidi, G., Hassaine, A., Cai, Y., Rahimi, K., and Mamouei, M. Uncertainty-aware interpretable deep learning for slum mapping and monitoring. *Remote Sensing 14*, 13 (June 2022), 3072. eprint: https://doi.org/10.3390/rs14133072.

[33] Forti, N., D'Afflisio, E., Braca, P., Millefiori, L. M., Willett, P., and Carniel, S. Maritime Anomaly Detection in a Real-World Scenario: Ever Given Grounding in the Suez Canal. *IEEE Transactions on Intelligent Transportation Systems 23*, 8 (Nov. 2022), 13904–13910. eprint: https://doi.org/10.1109/TITS.2021.3123890.

[34] Forti, N., Millefiori, L. M., and Braca, P. Hybrid Bernoulli Filtering for Detection and Tracking of Anomalous Path Deviations. In *2018 IEEE 21st International Conference on Information Fusion* (Cambridge, UK, July 2018), IEEE, pp. 1178–1184. eprint: https://doi.org/10.23919/ICIF.2018.8455567.

[35] Forti, N., Millefiori, L. M., Braca, P., and Willett, P. Anomaly Detection and Tracking Based on Mean–Reverting Processes with Unknown Parameters. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing* (Brighton, UK, May 2019), IEEE, pp. 8434–8438. eprint: https://doi.org/10.1109/ICASSP.2019.8683428.

[36] Forti, N., Millefiori, L. M., Braca, P., and Willett, P. Random Finite Set Tracking for Anomaly Detection in the Presence of Clutter. In *2020 IEEE International Radar Conference (RADAR)* (Florence, Italy, Sept. 2020), IEEE, pp. 1–6. eprint: https://doi.org/10.1109/RadarConf2043947.2020.9266705.

[37] Forti, N., Millefiori, L. M., Braca, P., and Willett, P. Bayesian Filtering for Dynamic Anomaly Detection and Tracking. *IEEE Transactions on Aerospace and Electronic Systems 58*, 3 (Nov. 2021), 1528–1544. eprint: https://doi.org/10.1109/TAES.2021.3122888.

[38] Fushiki, T. Estimation of prediction error by using K-fold cross-validation. *Statistics and Computing 21*, 2 (Oct. 2011), 137–146. eprint: https://doi.org/10.1007/s11222-009-9153-8.

[39] Gagniuc, P. A. *Markov Chains: From Theory to Implementation and Experimentation*. Wiley, Hoboken, NJ, USA, June 2017. eprint: https://doi.org/10.1002/9781119387596.

[40] Gal, Y., and Ghahramani, Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)* (New York, NY, USA, June 2016), vol. 48, PMLR, pp. 1050–1059. eprint: https://proceedings.mlr.press/v48/gal16.html.

[41] Gao, S., Koker, T., Queen, O., Hartvigsen, T., Tsiligkaridis, T., and Zitnik, M. UniTS: A Unified Multi-Task Time Series Model. *arXiv preprint arXiv:2403.00131* (Nov. 2024). eprint: https://doi.org/10.48550/arXiv.2403.00131.

[42] Gers, F. A., Schmidhuber, J., and Cummins, F. Learning to forget: Continual prediction with LSTM. *Neural Computation 12*, 10 (Oct. 2000), 2451–2471. eprint: https://doi.org/10.1162/089976600300015015.

[43] Gonçalves Jr, P. M., and de Barros, R. S. M. RCD: A recurring concept drift framework. *Pattern Recognition Letters 34*, 9 (July 2013), 1018–1025. eprint: https://doi.org/10.1016/j.patrec.2013.02.005.

[44] Gowda, T., and May, J. Finding the Optimal Vocabulary Size for Neural Machine Translation. In *Findings of the Association for Computational Linguistics: EMNLP 2020* (May 2020), pp. 3955–3964. eprint: https://doi.org/10.18653/v1/2020.findings-emnlp.352.

[45] Gregov, G. Modeling and predictive analysis of the hydraulic GEROLER motor based on artificial neural network. *Engineering Review 42*, 2 (Feb. 2022), 91–100. eprint: https://doi.org/10.30765/er.1813.

[46] Guo, Y., and Wu, G. A restarted large-scale spectral clustering with self-guiding and block diagonal representation. *Pattern Recognition 156* (Dec. 2024), 110746. eprint: https://doi.org/10.1016/j.patcog.2024.110746.

[47] Hanyang, Z., Xin, S., and Zhenguo, Y. Vessel Sailing Patterns Analysis from S-AIS Data Based on K-means Clustering Algorithm. In *2019 IEEE 4th International Conference on Big Data Analytics (ICBDA)* (Suzhou, China, Mar. 2019), IEEE, pp. 10–13. eprint: https://doi.org/10.1109/ICBDA.2019.8713231.

[48] He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Las Vegas, NV, USA, June 2016), IEEE, pp. 770–778. eprint: https://doi.org/10.1109/CVPR.2016.90.

[49] Hiemsch, P. Incorporating Attention Mechanisms in RNN-based Encoder-Decoder Models. Technical report, Lebanese International University, June 2023. eprint: https://github.com/PatrickSVM/Seq2Seq-with-Attention/blob/main/report/732A81-2023-PRA1-pathi619.pdf.

[50] Hochreiter, S., and Schmidhuber, J. Long Short-Term Memory. *Neural Computation 9*, 8 (Nov. 1997), 1735–1780. eprint: https://doi.org/10.1162/neco.1997.9.8.1735.

[51] Hou, J., Lin, H., Yuan, H., and Pelillo, M. Flexible density peak clustering for real-world data. *Pattern Recognition 156* (Dec. 2024), 110772. eprint: https://doi.org/10.1016/j.patcog.2024.110772.

[52] Houenou, A., Bonnifait, P., Cherfaoui, V., and Yao, W. Vehicle trajectory prediction based on motion model and maneuver recognition. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Tokyo, Japan, Nov. 2013), IEEE, pp. 4363–4369. eprint: https://doi.org/10.1109/IROS.2013.6696982.

[53] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv preprint arXiv:1704.04861* (Apr. 2017). eprint: https://doi.org/10.48550/arXiv.1704.04861.

[54] Jiang, W., Luo, J., He, M., and Gu, W. Graph Neural Network for Traffic Forecasting: The Research Progress. *ISPRS International Journal of Geo-Information 12*, 3 (Feb. 2023), 100. eprint: https://doi.org/10.3390/ijgi12030100.

[55] Jiménez-Navarro, M. J., Martínez-Ballesteros, M., Martínez-Álvarez, F., and Asencio-Cortés, G. PHILNet: A novel efficient approach for time series forecasting using deep learning. *Information Sciences 632* (June 2023), 815–832. eprint: https://doi.org/10.1016/j.ins.2023.03.021.

[56] Kang, M., Jeong, J., Lee, S., Cho, J., and Hwang, S. J. Distilling LLM Agent into Small Models with Retrieval and Code Tools. *arXiv preprint arXiv:2505.17612* (Nov. 2025). eprint: https://doi.org/10.48550/arXiv.2505.17612.

[57] Kazemi, S., Abghari, S., Lavesson, N., Johnson, H., and Ryman, P. Open data for anomaly detection in maritime surveillance. *Expert Systems with Applications 40*, 14 (Oct. 2013), 5719–5729. eprint: https://doi.org/10.1016/j.eswa.2013.04.029.

[58] Khan, A., Bil, C., Marion, K., and Crozier, M. Real-time prediction of ship motion and attitude. In *Proceedings of the 24th International Congress of the Aeronautical Sciences (ICAS 2004)* (Yokohama, Japan, Sept. 2004), ICAS, p. 489. eprint: https://www.icas.org/icas_archive/ICAS2004/PAPERS/489.PDF.

[59] Khanam, R., and Hussain, M. YOLOv11: An Overview of the Key Architectural Enhancements. *arXiv preprint arXiv:2410.17725* (Oct. 2024). eprint: https://doi.org/10.48550/arXiv.2410.17725.

[60] Khosroshahi, A., Ohn-Bar, E., and Trivedi, M. M. Surround vehicles trajectory analysis with recurrent neural networks. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)* (Rio de Janeiro, Brazil, Nov. 2016), IEEE, pp. 2267–2272. eprint: https://doi.org/10.1109/ITSC.2016.7795922.

[61] Kim, D., Park, J., Chung, H. C., and Jeong, S. Unsupervised outlier detection using random subspace and subsampling ensembles of Dirichlet process mixtures. *Pattern Recognition 156* (Dec. 2024), 110846. eprint: https://doi.org/10.1016/j.patcog.2024.110846.

[62] Kim, Y. J., Lee, J. S., Pititto, A., Falco, L., Lee, M. S., Yoon, K. K., and Cho, I. S. Maritime Traffic Evaluation Using Spatial-Temporal Density Analysis Based on Big AIS

Data. *Applied Sciences 12*, 21 (Nov. 2022), 11246. eprint: https://doi.org/10.3390/app122111246.

[63] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W. Y., Dollár, P., and Girshick, R. Segment Anything. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)* (Paris, France, Oct. 2023), IEEE, pp. 3992–4003. eprint: https://doi.org/10.1109/ICCV51070.2023.00371.

[64] Klen, D., and Lerga, J. Comprehensive annotation of underwater data for image segmentation. In *MY FIRST CONFERENCE 2024 8th Annual PhD Conference on Engineering and Technology* (Rijeka, Croatia, Sept. 2024), University of Rijeka. eprint: https://mfc.uniri.hr/wp-content/uploads/2024/09/MFC_2024_Book_of_Abstracts.pdf.

[65] Kolambe, M., and Arora, S. Forecasting the Future: A Comprehensive Review of Time Series Prediction Techniques. *Journal of Electrical Systems 20*, 2s (Apr. 2024), 575–586. eprint: https://doi.org/10.52783/jes.1478.

[66] Kowalska, K., and Peel, L. Maritime anomaly detection using Gaussian Process active learning. In *2012 IEEE 15th International Conference on Information Fusion* (Singapore, Singapore, July 2012), IEEE, pp. 1164–1171. eprint: https://ieeexplore.ieee.org/document/6289940.

[67] Kulkarni, S., Singh, S., Balakrishnan, D., Sharma, S., Devunuri, S., and Korlapati, S. C. R. CrackSeg9k: A Collection and Benchmark for Crack Segmentation Datasets and Frameworks. In *Computer Vision – ECCV 2022 Workshops* (Tel Aviv, Israel, Oct. 2022), Springer Nature Switzerland, pp. 179–195. eprint: https://doi.org/10.1007/978-3-031-25082-8_12.

[68] Kumar, P., Perrollaz, M., Lefèvre, S., and Laugier, C. Learning-based approach for online lane change intention prediction. In *2013 IEEE Intelligent Vehicles Symposium (IV)* (Gold Coast, QLD, Australia, June 2013), IEEE, pp. 797–802. eprint: https://doi.org/10.1109/IVS.2013.6629564.

[69] Laxhammar, R. Anomaly detection for sea surveillance. In *2008 IEEE 11th International Conference on Information Fusion* (Cologne, Germany, July 2008), IEEE, pp. 1–8. eprint: https://ieeexplore.ieee.org/document/4632192.

[70] Lee, D., Malacarne, S., and Aune, E. Explainable time series anomaly detection using masked latent generative modeling. *Pattern Recognition 156* (Dec. 2024), 110826. eprint: https://doi.org/10.1016/j.patcog.2024.110826.

[71] Lee, J. G., Han, J., and Whang, K. Y. Trajectory clustering: a partition-and-group framework. In *SIGMOD '07: Proceedings of the 2007 ACM SIGMOD international conference on Management of data* (Beijing, China, June 2007), ACM, pp. 593–604. eprint: https://doi.org/10.1145/1247480.1247546.

[72] Lefevre, S., Vasquez, D., and Laugier, C. A survey on motion prediction and risk assessment for intelligent vehicles. *ROBOMECH Journal 1*, 1 (July 2014), 1. eprint: https://doi.org/10.1186/s40648-014-0001-z.

[73] Li, B., and Qian, Y. Weather Prediction Using CNN-LSTM for Time Series Analysis: A Case Study on Delhi Temperature Data. *arXiv preprint arXiv:2409.09414* (Sept. 2024). eprint: https://doi.org/10.48550/arXiv.2409.09414.

[74] Li, J., Liu, J., Zhang, X., Li, X., Wang, J., and Wu, Z. A Novel Hybrid Approach for Detecting Abnormal Vessel Behavior in Maritime Traffic. In *2023 7th International Conference on Transportation Information and Safety (ICTIS)* (Xi'an, China, Aug. 2023), IEEE, pp. 1–7. eprint: https://doi.org/10.1109/ICTIS60134.2023.10243728.

[75] Liu, B., de Souza, E. N., Hilliard, C., and Matwin, S. Ship movement anomaly detection using specialized distance measures. In *2015 IEEE 18th International Conference on Information Fusion* (Washington, DC, USA, July 2015), IEEE, pp. 1113–1120. eprint: https://ieeexplore.ieee.org/document/7266683.

[76] Liu, H., Wang, Z., Dong, X., and Du, J. OnsitNet: A memory-capable online time series forecasting model incorporating a self-attention mechanism. *Expert Systems with Applications 259* (Jan. 2025), 125231. eprint: https://doi.org/10.1016/j.eswa.2024.125231.

[77] Liu, J., Mou, S., Gaw, N., and Wang, Y. Uni-3DAD: Gan-inversion aided universal 3D anomaly detection on model-free products. *Expert Systems with Applications 272* (May 2025), 126665. eprint: https://doi.org/10.1016/j.eswa.2025.126665.

[78] Liu, Q., Lathrop, B., and Butakov, V. Vehicle lateral position prediction: A small step towards risk assessment. In *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (Qingdao, China, Oct. 2014), IEEE, pp. 667–672. eprint: https://doi.org/10.1109/ITSC.2014.6957766.

[79] Liu, Y., Lin, X., Chen, Y., and Cheng, R. Multi-order graph clustering with adaptive node-level weight learning. *Pattern Recognition 156* (Dec. 2024), 110843. eprint: https://doi.org/10.1016/j.patcog.2024.110843.

[80] Livojević, A., Žužić, L., and Lerga, J. Image Compression Using Base64, BWT, RLE and LZW Coding. In *Proceedings of the Baška SIF (Spatial Intelligence Forum) Meeting 2024* (Baška, Croatia, June 2024).

[81] Loi, N. V., Kien, T. T., Hop, T. V., Son, L. T., and Khuong, N. V. Abnormal Moving Speed Detection Using Combination of Kernel Density Estimator and DBSCAN for Coastal Surveillance Radars. In *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)* (Noida, India, Feb. 2020), IEEE, pp. 143–147. eprint: https://doi.org/10.1109/SPIN48934.2020.9070885.

[82] Long, Z., Gao, Y., Meng, H., Chen, Y., and Kou, H. Semi-supervised clustering guided by pairwise constraints and local density structures. *Pattern Recognition 156* (Dec. 2024), 110751. eprint: https://doi.org/10.1016/j.patcog.2024.110751.

[83] Luo, Y., Zhang, S., Lyu, Z., and Hu, Y. TFDNet: Time–Frequency enhanced Decomposed Network for long-term time series forecasting. *Pattern Recognition 162* (June 2025), 111412. eprint: https://doi.org/10.1016/j.patcog.2025.111412.

[84] Mandalia, H. M., and Salvucci, M. D. D. Using Support Vector Machines for Lane-Change Detection. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Orlando, Florida, Sept. 2005), vol. 49, SAGE, pp. 1965–1969. eprint: https://doi.org/10.1177/154193120504902217.

[85] Martineau, E., and Roy, J. Maritime Anomaly Detection: Domain Introduction and Review. Technical report, Defence Research and Development Canada (DRDC), Oct. 2011. eprint: https://apps.dtic.mil/sti/tr/pdf/ADA554310.pdf.

[86] Mascaro, S., Nicholson, A. N., and Korb, K. B. Anomaly detection in vessel tracks using Bayesian networks. *International Journal of Approximate Reasoning 55*, 1 (Jan. 2014), 84–98. eprint: https://doi.org/10.1016/j.ijar.2013.03.012.

[87] McGregor, S. Incident Number 71: Google admits its self driving car got it wrong: Bus crash was caused by software. https://incidentdatabase.ai/cite/71, Sept. 2016. Accessed: 22.05.2025.

[88] Millefiori, L. M., Braca, P., Bryan, K., and Willett, P. Modeling vessel kinematics using a stochastic mean-reverting process for long-term prediction. *IEEE Transactions on Aerospace and Electronic Systems 52*, 5 (Oct. 2016), 2313–2330. eprint: https://doi.org/10.1109/TAES.2016.150596.

[89] Moriano, P., Berres, A., Xu, H., and Sanyal, J. Spatiotemporal features of traffic help reduce automatic accident detection time. *Expert Systems with Applications 244* (June 2024), 122813. eprint: https://doi.org/10.1016/j.eswa.2023.122813.

[90] Mung, P. S., and Phyu, S. Time Series Weather Data Forecasting Using Deep Learning. In *2023 IEEE Conference on Computer Applications (ICCA)* (Yangon, Myanmar, Feb. 2023), IEEE, pp. 254–259. eprint: https://doi.org/10.1109/ICCA51723.2023.10182058.

[91] Nawi, N. M., Atomi, W. H., and Rehman, M. Z. The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks. *Procedia Technology 11* (June 2013), 32–39. eprint: https://doi.org/10.1016/j.protcy.2013.12.159.

[92] Ning, J. Study on maritime traffic density calculation. *Applied Sciences 6*, 3 (Mar. 2016), 85. eprint: https://doi.org/10.3390/app6030085.

[93] Njirjak, M., Žužić, L., Babić, M., Janković, P., Otović, E., Kalafatović, D., and Mauša, G. Reshaping the discovery of self-assembling peptides with generative AI guided by hybrid deep learning. *Nature Machine Intelligence 6* (Dec. 2024), 1487–1500. eprint: https://doi.org/10.1038/s42256-024-00928-1.

[94] Øksendal, B. *Stochastic Differential Equations: An Introduction with Applications*. Springer Berlin, July 2003. eprint: https://doi.org/10.1007/978-3-642-14394-6.

[95] Olah, C. Neural Networks, Types, and Functional Programming. http://colah.github.io/posts/2015-09-NN-Types-FP, May 2015. Accessed: 01.01.2024.

[96] Olesen, K. V., Boubekki, A., Kampffmeyer, M. C., Jenssen, R., Christensen, A. N., Hørlück, S., and Clemmensen, L. H. A Contextually Supported Abnormality Detector for Maritime Trajectories. *Journal of Marine Science and Engineering 11*, 11 (Oct. 2023), 2085. eprint: https://doi.org/10.3390/jmse11112085.

[97] Orinaitė, U., Karaliūtė, V., Pal, M., and Ragulskis, M. Detecting Underwater Concrete Cracks with Machine Learning: A Clear Vision of a Murky Problem. *Applied Sciences 13*, 12 (June 2023), 7335. eprint: https://doi.org/10.3390/app13127335.

[98] OtoTrak. OtoTrak - Track your watercraft — ototrak.com. https://www.ototrak.com/en-us, July 2024. Accessed: 10.07.2025.

[99] Otović, E., Njirjak, M., Kalafatović, D., and Mauša, G. Sequential Properties Representation Scheme for Recurrent Neural Network-Based Prediction of Therapeutic Peptides. *Journal of Chemical Information and Modeling 62*, 12 (June 2022), 2961–2972. eprint: https://doi.org/10.1021/acs.jcim.2c00526.

[100] Pallotta, G., Horn, S., Braca, P., and Bryan, K. Context-enhanced vessel prediction based on Ornstein-Uhlenbeck processes using historical AIS traffic patterns: Real-world experimental results. In *17th IEEE International Conference on Information Fusion* (Salamanca, Spain, July 2014), IEEE, pp. 1–7. eprint: https://ieeexplore.ieee.org/document/6916016.

[101] Pallotta, G., and Jousselme, A. L. Data-driven detection and context-based classification of maritime anomalies. In *2015 IEEE 18th International Conference on Information*

Fusion (Washington, DC, USA, July 2015), IEEE, pp. 1152–1159. eprint: https://doi.org/10.1109/ICIF.2015.7238618.

[102] Pallotta, G., Vespe, M., and Bryan, K. Traffic Route Extraction and Anomaly Detection from AIS Data. In *Proceedings of the International COST MOVE Workshop on Moving Objects at Sea* (Brest, France, June 2013), COST MOVE. eprint: https://www.researchgate.net/publication/258630626_Traffic_Route_Extraction_and_Anomaly_Detection_from_AIS_Data.

[103] Pallotta, G., Vespe, M., and Bryan, K. Vessel Pattern Knowledge Discovery from AIS Data: A Framework for Anomaly Detection and Route Prediction. *Entropy 15*, 6 (June 2013), 2218–2245. eprint: https://doi.org/10.3390/e15062218.

[104] Parzen, E. *Stochastic Processes*. Dover Publications, Garden City, NY, USA, May 2015.

[105] Pascanu, R., Mikolov, T., and Bengio, Y. On the difficulty of training Recurrent Neural Networks. In *Proceedings of the 30th International Conference on Machine Learning* (Atlanta, GA, USA, June 2013), vol. 28, PMLR, pp. 1310–1318. eprint: https://proceedings.mlr.press/v28/pascanu13.html.

[106] Pedroche, D. S., Herrero, D. A., Herrero, J. G., and Lopez, J. M. M. Clustering of maritime trajectories with AIS features for context learning. In *2021 IEEE 24th International Conference on Information Fusion* (Sun City, South Africa, Nov. 2021), IEEE, pp. 1–8. eprint: https://doi.org/10.23919/FUSION49465.2021.9626956.

[107] Phillips, D. J., Wheeler, T. A., and Kochenderfer, M. J. Generalizable intention prediction of human drivers at intersections. In *2017 IEEE Intelligent Vehicles Symposium (IV)* (Los Angeles, CA, USA, June 2017), IEEE, pp. 1665–1670. eprint: https://doi.org/10.1109/IVS.2017.7995948.

[108] Qingyi, W., Lingyun, G., and Bo, C. The lightweight CS-YOLO model applied for concrete crack segmentation. *Expert Systems with Applications 277* (June 2025), 127298. eprint: https://doi.org/10.1016/j.eswa.2025.127298.

[109] Rahmah, N., and Sitanggang, I. S. Determination of Optimal Epsilon (Eps) Value on DBSCAN Algorithm to Clustering Data on Peatland Hotspots in Sumatra. In *Workshop and International Seminar on Science of Complex Natural Systems* (Bogor, Indonesia, Oct. 2015), vol. 31, IOP Publishing, p. 012012. eprint: https://doi.org/10.1088/1755-1315/31/1/012012.

[110] Ravi, N., Gabeur, V., Hu, Y. T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K. V., Carion, N., Wu, C. Y., Girshick, R., Dollár, P., and Feichtenhofer, C. SAM 2: Segment Anything in Images and Videos.

arXiv preprint arXiv:2408.00714 (Oct. 2024). eprint: https://doi.org/10.48550/arXiv.2408.00714.

[111] Ray, C., Dréo, R., Camossi, E., Jousselme, A. L., and Iphar, C. Heterogeneous integrated dataset for Maritime Intelligence, surveillance, and reconnaissance. *Data in Brief 25* (Aug. 2019), 104141. eprint: https://doi.org/10.1016/j.dib.2019.104141.

[112] Roboflow. Crack Dataset. https://universe.roboflow.com/university-bswxt/crack-bphdr, May 2025. Accessed: 22.05.2025.

[113] Ronneberger, O., Fischer, P., and Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Munich, Germany, Oct. 2015), Springer Nature Switzerland, pp. 234–241. eprint: https://doi.org/10.1007/978-3-319-24574-4_28.

[114] Roy, J. Anomaly detection in the maritime domain. In *SPIE Defense and Security Symposium* (Orlando, FL, USA, Apr. 2008), vol. 6945, SPIE, p. 69450W. eprint: https://doi.org/10.1117/12.776230.

[115] rp5. Definitions — rp5.ru. https://rp5.ru/docs/descript/en, May 2025. Accessed: 22.05.2025.

[116] Russel, N., and Selvaraj, A. MultiScaleCrackNet: A parallel multiscale deep CNN architecture for concrete crack classification. *Expert Systems with Applications 249*, Part B (Sept. 2024), 123658. eprint: https://doi.org/10.1016/j.eswa.2024.123658.

[117] Saberironaghi, A., and Ren, J. DepthCrackNet: A Deep Learning Model for Automatic Pavement Crack Detection. *Journal of Imaging 10*, 5 (Apr. 2024), 100. eprint: https://doi.org/10.3390/jimaging10050100.

[118] Salami, A. A., Agbessi, P. A., Boureima, S., and Ajavon, A. S. A. Wind energy potential estimation using neural network and SVR approaches. *Engineering Review 42*, 3 (Dec. 2022), 32–49. eprint: https://doi.org/10.30765/er.1632.

[119] Salazar, C., and Banerjee, A. G. A distance correlation-based approach to characterize the effectiveness of recurrent neural networks for time series forecasting. *Neurocomputing 629* (May 2025), 129641. eprint: https://doi.org/10.1016/j.neucom.2025.129641.

[120] Salman, A. G., and Kanigoro, B. Visibility Forecasting Using Autoregressive Integrated Moving Average (ARIMA) Models. *Procedia Computer Science 179* (Jan. 2021), 252–259. eprint: https://doi.org/10.1016/j.procs.2021.01.004.

[121] Sangiorgio, M., and Dercole, F. Robustness of LSTM neural networks for multi-step forecasting of chaotic time series. *Chaos, Solitons & Fractals 139* (Oct. 2020), 110045. eprint: https://doi.org/10.1016/j.chaos.2020.110045.

[122] Sanjay-Gopal, S., and Hebert, T. J. Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm. *IEEE Transactions on Image Processing 7*, 7 (July 1998), 1014–1028. eprint: https://doi.org/10.1109/83.701161.

[123] Satopaa, V., Albrecht, J., Irwin, D., and Raghavan, B. Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior. In *2011 31st International Conference on Distributed Computing Systems Workshops* (Minneapolis, MN, USA, June 2011), IEEE, pp. 166–171. eprint: https://doi.org/10.1109/ICDCSW.2011.20.

[124] Schwab, M., Mayr, A., and Haltmeier, M. Deep Gaussian Mixture Model for Unsupervised Image Segmentation. In *Machine Learning, Optimization, and Data Science: 10th International Conference, LOD 2024* (Castiglione della Pescaia, Italy, Sept. 2024), Springer Nature Switzerland, pp. 339–352. eprint: https://doi.org/10.1007/978-3-031-82484-5_25.

[125] Sezer, O. B., Gudelek, M. U., and Ozbayoglu, A. M. Financial time series forecasting with deep learning : A systematic literature review: 2005–2019. *Applied Soft Computing 90* (May 2020), 106181. eprint: https://doi.org/10.1016/j.asoc.2020.106181.

[126] Shaik, T., Tao, X., Li, L., Xie, H., Acharya, U. R., Gururajan, R., and Zhou, X. Predictive deep reinforcement learning with multi-agent systems for adaptive time series forecasting. *Knowledge-Based Systems 326* (Sept. 2025), 113941. eprint: https://doi.org/10.1016/j.knosys.2025.113941.

[127] Shan, J., Huang, Y., and Jiang, W. DCUFormer: Enhancing pavement crack segmentation in complex environments with dual-cross/upsampling attention. *Expert Systems with Applications 264* (Mar. 2025), 125891. eprint: https://doi.org/10.1016/j.eswa.2024.125891.

[128] Sharma, A., and Ranjan, P. Traffic Prediction Model Using Machine Learning in Intelligent Transportation Systems. In *2023 International Conference on Circuit Power and Computing Technologies (ICCPCT)* (Kollam, India, Aug. 2023), IEEE, pp. 1165–1173. eprint: https://doi.org/10.1109/ICCPCT58313.2023.10245139.

[129] Shim, S. Self-training approach for crack detection using synthesized crack images based on conditional generative adversarial network. *Computer-Aided Civil and Infrastructure Engineering 39*, 7 (Apr. 2024), 1019—1041. eprint: https://doi.org/10.1111/mice.13119.

[130] Singh, D., and Singh, B. Investigating the impact of data normalization on classification performance. *Applied Soft Computing 97*, Part B (Dec. 2020), 105524. eprint: https://doi.org/10.1016/j.asoc.2019.105524.

[131] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research 15*, 56 (Jan. 2014), 1929–1958. eprint: https://jmlr.org/papers/v15/srivastava14a.html.

[132] Streubel, T., and Hoffmann, K. H. Prediction of driver intended path at intersections. In *2014 IEEE Intelligent Vehicles Symposium Proceedings* (Dearborn, MI, USA, June 2014), IEEE, pp. 134–139. eprint: https://doi.org/10.1109/IVS.2014.6856508.

[133] Sun, F., Deng, Y., Deng, F., Zhu, Q., and Chu, H. Unsupervised maritime traffic pattern extraction from spatio-temporal data. In *2015 11th International Conference on Natural Computation (ICNC)* (Zhangjiajie, China, Aug. 2015), IEEE, pp. 1218–1223. eprint: https://doi.org/10.1109/ICNC.2015.7378165.

[134] Sutskever, I., Vinyals, O., and Le, Q. V. Sequence to Sequence Learning with Neural Networks. In *Advances in Neural Information Processing Systems 27 (NIPS 2014)* (Montreal, QB, Canada, Dec. 2014), vol. 27, Curran Associates, Inc., pp. 3104–3112. eprint: https://proceedings.neurips.cc/paper_files/paper/2014/hash/5a18e133cbf9f257297f410bb7eca942-Abstract.html.

[135] Talagala, T., Li, F., and Kang, Y. FFORMPP: Feature-based forecast model performance prediction. *International Journal of Forecasting 38*, 3 (July 2022), 920–943. eprint: https://doi.org/10.1016/j.ijforecast.2021.07.002.

[136] Tay, C., Mekhnacha, K., and Laugier, C. *Probabilistic Vehicle Motion Modeling and Risk Estimation*. Springer London, London, Mar. 2012, pp. 1479—1516. eprint: https://doi.org/10.1007/978-0-85729-085-4_57.

[137] Tian, Y., Ye, Q., and Doermann, D. YOLOv12: Attention-Centric Real-Time Object Detectors. *arXiv preprint arXiv:2502.12524* (Feb. 2025). eprint: https://doi.org/10.48550/arXiv.2502.12524.

[138] Tomar, R., and Verma, S. Safety of Lane Change Maneuver Through A Priori Prediction of Trajectory Using Neural Networks. *Network Protocols and Algorithms 4*, 1 (May 2012), 4–21. eprint: https://doi.org/10.5296/npa.v4i1.1240.

[139] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., and Lample, G. LLaMA: Open and Efficient Foundation Language Models. *arXiv preprint*

arXiv:2302.13971 (Feb. 2023). eprint: https://doi.org/10.48550/arXiv.2302.13971.

[140] Valle, J., and Bruno, O. M. Forecasting chaotic time series: Comparative performance of LSTM-based and Transformer-based neural network. *Chaos, Solitons & Fractals 192* (Mar. 2025), 116034. eprint: https://doi.org/10.1016/j.chaos.2025.116034.

[141] Varlamis, I., Tserpes, K., Etemad, M., Soares Júnior, A., and Matwin, S. A Network Abstraction of Multi-vessel Trajectory Data for Detecting Anomalies. In *Proceedings of the Workshops of the EDBT/ICDT 2019 Joint Conference* (Lisbon, Portugal, Mar. 2019), vol. 2322, CEUR Workshop Proceedings. eprint: https://ceur-ws.org/Vol-2322/BMDA_5.pdf.

[142] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. Attention is All you Need. In *Advances in Neural Information Processing Systems 30 (NIPS 2017)* (Long Beach, CA, USA, Dec. 2017), vol. 30, Curran Associates, Inc., pp. 5998–6008. eprint: https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

[143] Vectrino. Homepage Vectrino. https://www.vectrino.hr, May 2025. Accessed: 22.05.2025.

[144] Velez, D. R., White, B. C., Motsinger, A. A., Bush, W. S., Ritchie, M. D., Williams, S. M., and Moore, J. H. A balanced accuracy function for epistasis modeling in imbalanced datasets using multifactor dimensionality reduction. *Genetic Epidemiology 31*, 4 (Feb. 2007), 306–315. eprint: https://doi.org/10.1002/gepi.20211.

[145] Venskus, J., Treigys, P., and Markevičiūtė, J. Unsupervised marine vessel trajectory prediction using LSTM network and wild bootstrapping techniques. *Nonlinear Analysis: Modelling and Control 26*, 4 (July 2021), 718–737. eprint: https://www.redalyc.org/journal/6941/694173200010/html.

[146] Völz, B., Mielenz, H., Siegwart, R., and Nieto, J. Predicting pedestrian crossing using Quantile Regression forests. In *2016 IEEE Intelligent Vehicles Symposium (IV)* (Gothenburg, Sweden, June 2016), IEEE, pp. 426–432. eprint: https://doi.org/10.1109/IVS.2016.7535421.

[147] Wan, J., Xia, N., Yin, Y., Pan, X., Hu, J., and Yi, J. TCDformer: A transformer framework for non-stationary time series forecasting based on trend and change-point detection. *Neural Networks 173* (May 2024), 106196. eprint: https://doi.org/10.1016/j.neunet.2024.106196.

[148] Wang, C. Y., Liao, H. Y. M., Yeh, I. H., Wu, Y. H., Chen, P. Y., and Hsieh, J. W. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. *arXiv preprint arXiv:1911.11929* (Nov. 2019). eprint: https://doi.org/10.48550/arXiv.1911.11929.

[149] Wang, L., Chen, P., Chen, L., and Mou, J. Ship AIS Trajectory Clustering: An HDBSCAN-Based Approach. *Journal of Marine Science and Engineering 9*, 6 (May 2021), 566. eprint: https://doi.org/10.3390/jmse9060566.

[150] Wang, Z., Ran, H., Ren, J., and Sun, M. PWDformer: Deformable transformer for long-term series forecasting. *Pattern Recognition 147* (Mar. 2024), 110118. eprint: https://doi.org/10.1016/j.patcog.2023.110118.

[151] Wu, J., and Zhang, X. Tunnel Crack Detection Method and Crack Image Processing Algorithm Based on Improved Retinex and Deep Learning. *Sensors 23*, 22 (Nov. 2023), 9140. eprint: https://doi.org/10.3390/s23229140.

[152] Xiao, T., Pang, R., Liu, H., Yang, C., Li, A., Niu, C., Ruan, Z., Xu, L., and Ge, Y. Domain adaptation and knowledge distillation for lightweight pavement crack detection. *Expert Systems with Applications 263* (Mar. 2025), 125734. eprint: https://doi.org/10.1016/j.eswa.2024.125734.

[153] Xu, J., Wang, S., Han, R., Wu, X., Zhao, D., Zeng, X., Yin, R., Han, Z., Liu, Y., and Shu, S. Crack segmentation and quantification in concrete structures using a lightweight YOLO model based on pruning and knowledge distillation. *Expert Systems with Applications 283* (July 2025), 127834. eprint: https://doi.org/10.1016/j.eswa.2025.127834.

[154] Yoon, S., and Kum, D. The multilayer perceptron approach to lateral motion prediction of surrounding vehicles for autonomous vehicles. In *2016 IEEE Intelligent Vehicles Symposium (IV)* (Gothenburg, Sweden, June 2016), IEEE, pp. 1307–1312. eprint: https://doi.org/10.1109/IVS.2016.7535559.

[155] Zeb, A., Zhang, S., Wei, X., and Yu, J. J. A generalized feature projection scheme for multi-step traffic forecasting. *Expert Systems with Applications 244* (June 2024), 122962. eprint: https://doi.org/10.1016/j.eswa.2023.122962.

[156] Zernetsch, S., Kohnen, S., Goldhammer, M., Doll, K., and Sick, B. Trajectory prediction of cyclists using a physical model and an artificial neural network. In *2016 IEEE Intelligent Vehicles Symposium (IV)* (Gothenburg, Sweden, June 2016), IEEE, pp. 833–838. eprint: https://doi.org/10.1109/IVS.2016.7535484.

[157] Zhang, C., Chen, H., Li, H., and Chen, C. Learning latent disentangled embeddings and graphs for multi-view clustering. *Pattern Recognition 156* (Dec. 2024), 110839. eprint: https://doi.org/10.1016/j.patcog.2024.110839.

[158] Zhang, F., Gu, Y., Yin, L., Song, J., Qiu, C., Ye, Z., Chen, X., and Wu, J. Research on the generation and evaluation of bridge defect datasets for underwater environments utilizing CycleGAN networks. *Expert Systems with Applications 262* (Mar. 2025), 125576. eprint: https://doi.org/10.1016/j.eswa.2024.125576.

[159] Zhang, J., Zeng, Z., Sharma, P. K., Alfarraj, O., Tolba, A., and Wang, J. A dual encoder crack segmentation network with Haar wavelet-based high–low frequency attention. *Expert Systems with Applications 256* (Dec. 2024), 124950. eprint: https://doi.org/10.1016/j.eswa.2024.124950.

[160] Zhang, X., Zhong, C., Zhang, J., Wang, T., and Ng, W. W. Y. Robust recurrent neural networks for time series forecasting. *Neurocomputing 526* (Mar. 2023), 143–157. eprint: https://doi.org/10.1016/j.neucom.2023.01.037.

[161] Zhang, Z., Peng, B., and Zhao, T. An ultra-lightweight network combining Mamba and frequency-domain feature extraction for pavement tiny-crack segmentation. *Expert Systems with Applications 264* (Mar. 2025), 125941. eprint: https://doi.org/10.1016/j.eswa.2024.125941.

[162] Zhao, C., Song, A., Zeng, Z., Ji, Y., and Du, Y. Multi-modal trajectory forecasting with Multi-scale Interactions and Multi-pseudo-target Supervision . *Knowledge-Based Systems 296* (July 2024), 111903. eprint: https://doi.org/10.1016/j.knosys.2024.111903.

[163] Zheng, J., Suzuki, K., and Fujita, M. Predicting driver's lane-changing decisions using a neural network model. *Simulation Modelling Practice and Theory 42* (Mar. 2014), 73–83. eprint: https://doi.org/10.1016/j.simpat.2013.12.007.

[164] Zheng, Y., and Zhou, X., Eds. *Computing with Spatial Trajectories*. Springer New York, New York, NY, USA, Jan. 2011. eprint: https://doi.org/10.1007/978-1-4614-1629-6.

[165] Zhong, Y., Ren, Y., Cao, G., Li, F., and Qi, H. Optimal starting point for time series forecasting. *Expert Systems with Applications 273* (May 2025), 126798. eprint: https://doi.org/10.1016/j.eswa.2025.126798.

[166] Zhuang, X. Multivariate Mixture Model for Myocardial Segmentation Combining Multi-Source Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence 41*, 12 (Dec. 2019), 2933–2946. eprint: https://doi.org/10.1109/TPAMI.2018.2869576.

[167] Zyner, A., Worrall, S., Ward, J., and Nebot, E. Long short term memory for driver intent prediction. In *2017 IEEE Intelligent Vehicles Symposium (IV)* (Los Angeles, CA, USA, June 2017), IEEE, pp. 1484–1489. eprint: https://doi.org/10.1109/IVS.2017.7995919.

[168] Žužić, L. Predviđanje putanje osobnih plovila korištenjem modela temeljenih na neuronskim mrežama s povratnom vezom. In *Proceedings of the Scientific conference of doctoral students: INFCON24* (Rijeka, Croatia, Sept. 2024), University of Rijeka, Faculty of Informatics and Digital Technologies, pp. 68–69. eprint: https://inf.uniri.hr/images/studiji/poslijediplomski/INFCON/2024/Infcon_proc_2024.pdf.

[169] Žužić, L. Segmentacija slika korištenjem modela specifičnih za pojedinu klasu temeljenih na dubokom učenju, Gaussovoj mješavini i Bayesovoj vjerojatnosti. In *Proceedings of the Scientific conference of doctoral students: INFCON25* (Rijeka, Croatia, Sept. 2025), University of Rijeka, Faculty of Informatics and Digital Technologies, pp. 47–48. eprint: https://www.inf.uniri.hr/images/studiji/poslijediplomski/INFCON/2025/Infcon_proc_2025_FIN.pdf.

[170] Žužić, L., Dražić, I., Simčić, L., Hržić, F., and Lerga, J. A Bayesian and Markov chain approach to short-term and long-term personal watercraft trajectory forecasting. *Journal of the Franklin Institute 362*, 3 (Feb. 2025), 107509. eprint: https://doi.org/10.1016/j.jfranklin.2025.107509.

[171] Žužić, L., and Filjar, R. Circular Analysis of Long-Term GNSS Horizontal Positioning Error Time Series Observed in the Sub-Equatorial Region. In *Proceedings of the 2025 33rd Telecommunications Forum (TELFOR)* (Belgrade, Serbia, Nov. 2025), IEEE, pp. 1–4. eprint: https://doi.org/10.1109/TELFOR67910.2025.11314378.

[172] Žužić, L., Hržić, F., and Lerga, J. AI-based prediction of trajectories for personal watercrafts. In *Proceedings of the 8th Annual PhD Conference on Engineering and Technology MFC 2024* (Rijeka, Croatia, Sept. 2024), University of Rijeka, Faculty of Maritime Studies, p. 16. eprint: https://mfc.uniri.hr/wp-content/uploads/2024/09/MFC_2024_Book_of_Abstracts.pdf.

[173] Žužić, L., Hržić, F., and Lerga, J. Forecasting the trajectory of personal watercrafts using models based on recurrent neural networks. *Expert Systems with Applications 284* (July 2025), 127964. eprint: https://doi.org/10.1016/j.eswa.2025.127964.

[174] Žužić, L., Hržić, F., and Lerga, J. Collision course detection for personal watercrafts using models based on recurrent neural networks. *Knowledge-Based Systems 333* (Jan. 2026), 114974. eprint: https://doi.org/10.1016/j.knosys.2025.114974.

[175] Žužić, L., Hržić, F., Li, X., and Lerga, J. Class-specific image segmentation across multiple domains using customized U-Net pipelines. *Expert Systems with Applications 296*, Part D, 15 (Jan. 2026), 129203. eprint: https://doi.org/10.1016/j.eswa.2025.129203.

[176] Žužić, L., Hržić, F., Petrijevčanin, I., and Lerga, J. Inflection point-based trajectory fingerprinting for clustering and anomaly detection. *Pattern Recognition 174* (June 2026), 112977. eprint: https://doi.org/10.1016/j.patcog.2025.112977.

[177] Žužić, L., Hržić, F., Šimic, M., Otulić, I., and Lerga, J. Forecasting Horizons for Watercraft Trajectory Intersection Detection. In *Proceedings of the 12th EAI International Conference on Smart Objects and Technologies for Social Good* (Dubrovnik, Croatia, April 2026), Springer Nature Switzerland.

[178] Žužić, L., Klen, D., Špoljar, D., and Filjar, R. Machine Learning Algorithm for the Influence of Space Weather Events of Geomagnetic Activity on GNSS Services. *The Journal of CIEES 6*, 1 (Apr. 2026), 7–15. eprint: https://doi.org/10.48149/jciees.2026.6.1.1.

[179] Žužić, L., and Lerga, J. A Versatile Image Annotation and Masking Tool. In *Proceedings of the 9th Annual PhD Conference on Engineering and Technology MFC 2025* (Rijeka, Croatia, Sept. 2025), University of Rijeka, Faculty of Civil Engineering, p. 25. eprint: https://mfc.uniri.hr/wp-content/uploads/2025/09/MFC2025_BoA_final.pdf.

[180] Žužić, L., and Lerga, J. A Web-Based Card Game Simulation for Optimal Strategy Development. In *Proceedings of the 2025 MIPRO 48th ICT and Electronics Convention* (Opatija, Croatia, June 2025), IEEE, pp. 1–6. eprint: https://doi.org/10.1109/mipro65660.2025.11131986.

[181] Žužić, L., and Lerga, J. Automatic Image Segment Classification Using Contour Features. In *Proceedings of the Baška SIF (Spatial Intelligence Forum) Meeting 2025* (Baška, Croatia, Sept. 2025), University of Rijeka, Faculty of Engineering. eprint: https://riteh.uniri.hr/wp-content/uploads/Baska-SIF-Meeting-2025-Proceedings-1.pdf.

[182] Žužić, L., and Lerga, J. Enhancing Greenhouse Gas Emission Awareness Through Web-based Tools. In *Proceedings of the 2026 MIPRO 49th ICT and Electronics Convention* (Opatija, Croatia, May 2026), IEEE.

[183] Žužić, L., Škrlj, L., Nešković, A., and Filjar, R. Characterisation of GPS Horizontal Positioning Errors and Dst Using Recurrence Plot Analysis in Sub-Equatorial Ionospheric

Conditions. *Urban Science 9*, 11 (Oct. 2025), 451. eprint: https://doi.org/10.3390/urbansci9110451.

[184] Žužić, L., Štajduhar, I., Lerga, J., and Filjar, R. Non-Parametric Model for Curvature Classification of Departure Flight Trajectory Segments. In *Proceedings of the 2025 6th International Conference on Communications, Information, Electronic and Energy Systems (CIEES)* (Ruse, Bulgaria, Nov. 2025), MDPI Engineering Proceedings, pp. 1–14. eprint: https://doi.org/10.3390/engproc2026122001.

List of figures

List of tables

SUMMARY OF THE SCIENTIFIC CONTRIBUTIONS OF PUBLISHED PAPERS

A A Bayesian and Markov chain approach to short-term and long-term personal watercraft trajectory forecasting

In this work, vessel position is estimated using a Bayesian approach based on heading, speed, time intervals, and offsets of latitude and longitude. An additional approach using a Markov chain is presented. The trajectory data comes from a cloud-based marine watercraft tracking system that enables remote control of the vessels. Wave height and meteorological reports were used to evaluate the impact of weather on personal watercraft trajectories. One proposed approach to trajectory estimation uses the longitude and latitude offsets, while another uses the speed, heading, and actual time intervals. A long-term forecasting window of up to ten seconds is achieved by dividing trajectories into segments that do not overlap. The limitation that this method faces in long-term forecasting inspires more sophisticated machine-learning approaches. The most successful estimation method used one or two previous actual values and a Bayesian approach, proving that using previously predicted values in a chain accumulates errors. Considering environmental variables did not improve the model, highlighting that small watercraft operate well even in unstable sea states. This occurs because they generate and ride waves, having a larger impact than some sea conditions.

Žužić L., Dražić I., Simčić L., Hržić F., Lerga J. A Bayesian and Markov chain approach to short-term and long-term personal watercraft trajectory forecasting. Journal of the Franklin Institute. 2025; 362(3):107509. https://doi.org/10.1016/j.jfranklin.2025.107509

B Forecasting the trajectory of personal watercrafts using models based on recurrent neural networks

Monitoring and predicting personal watercraft trajectories is a novel and largely unexplored research area where any development is valuable for various rental services. Unlike existing work focused on specific maritime routes, this study introduces a location-agnostic deep-learning approach capable of generalizing across diverse environments. This is achieved by using an innovative preprocessing approach, including offset and scaling. A novel LSTM bidirectional and convolutional model developed by the authors for binary peptide classification was adapted to accommodate the forecasting of continuous values. By leveraging RNN architectures with LSTM and GRU layers, and cutting-edge attention-based and foundation models, the study benchmarks trajectory forecasting performance using real-world data from 1282 rental sites worldwide. Most notably, the applicability of the listed models and the original LSTM bidirectional and convolutional models is extended to maritime trajectory estimation, eliminating the need for training separate models for different locations while achieving superior predictive accuracy. The results demonstrate that foundation models and attention mechanisms significantly outperform traditional methods, offering a robust and scalable solution for watercraft trajectory forecasting. These findings pave the way for intelligent monitoring systems that enhance maritime safety and operational efficiency.

Žužić L., Hržić F., Lerga J. Forecasting the trajectory of personal watercrafts using models based on recurrent neural networks. Expert Systems with Applications. 2025; 284:127964. https://doi.org/10.1016/j.eswa.2025.127964

C Collision course detection for personal watercrafts using models based on recurrent neural networks

Intersection detection plays a vital role in enhancing safety in personal watercraft rental services, particularly in tourist regions where inexperienced operators frequently navigate busy waterways. Intersection detection is defined here as forecasting personal watercraft trajectories up to 30 seconds ahead and identifying potential intersections that may result in accidents. This work builds on previous research by the authors, introducing DL approaches for personal watercraft trajectory forecasting on real-world data with RNN, LSTM, and GRU architectures, an attention-based model, and the foundation model UniTS for time-series data. Trajectory forecasting is integrated into intersection detection, replacing the previous IDC system by analyzing data from 527 locations across continents collected over several years, and analyzing key variables such as longitude, latitude, speed, and heading. The attention model using longitude and latitude offsets, with the longest forecasting horizon of thirty seconds, increased the true negative rate to 95.45%, which is a significant improvement compared to the non-machine learning approach. This study applies forecasting in real-time safety systems, replacing static proximity-based rules with advanced learning for adaptive and accurate accident prevention. The GRU attention model achieves inference under 12 milliseconds per sequence, enabling real-time deployment.

Žužić L., Hržić F., Lerga J. Collision course detection for personal watercrafts using models based on recurrent neural networks. Knowledge-Based Systems. 2026; 333:114974. https://doi.org/10.1016/j.knosys.2025.114974

D Inflection point-based trajectory fingerprinting for clustering and anomaly detection

Identifying kinematic anomalies is critical for maritime safety, especially in high-traffic areas. This paper presents a novel, lightweight, scalable method for trajectory clustering and kinematic anomaly detection using inflection-point sequences across varying trajectory lengths and environments. By encoding trajectory shape changes into compact fingerprints, the technique circumvents the computational burdens of DL. Results indicate that the proposed fingerprinting method can differentiate anomalous driving patterns, offering a viable solution for real-time anomaly detection. This study presents a simple yet robust trajectory analysis framework applicable to personal watercraft monitoring and to larger maritime vessels, paving the way for improved safety and real-time decision-support systems. Trajectories are segmented, inflection points are extracted, and clustering is performed using trajectory fingerprints. Human evaluation experiments validate the approach aligned with IDC, and a larger window size improves accuracy to 85% or higher when aligned with the ROT.

Žužić, L., Hržić, F., Petrijevčanin, I., Lerga, J. Inflection point-based trajectory fingerprinting for clustering and anomaly detection. Pattern Recognition. 2026; 174:112977.
https://doi.org/10.1016/j.patcog.2025.112977

E Class-specific image segmentation across multiple domains using customized U-Net pipelines

This paper presents a class-specific U-Net-based segmentation framework designed for high-precision segmentation across diverse image domains, including underwater inspection. Unlike traditional multi-class models, the proposed approach trains a separate U-Net for each class, enabling domain-specific tuning and threshold optimization. The proposed framework includes a Dice ($F1$ score) loss function tailored for multi-image performance, domain-customized preprocessing, and a stratified training pipeline. Evaluations across three distinct datasets demonstrate consistent performance improvements over state-of-the-art methods, such as the YOLO model and the baseline U-Net. The results show that this modular segmentation strategy achieves superior Dice coefficients, IoU, and class precision, particularly for small or ambiguous structures. The computational trade-offs and the method's potential for real-world deployment in infrastructure monitoring are discussed to prove the model's applicability.

Žužić L., Hržić F., Li X., Lerga J. Class-specific image segmentation across multiple domains using customized U-Net pipelines. Expert Systems with Applications. 2026; 296(Part D):129203. https://doi.org/10.1016/j.eswa.2025.129203

Curriculum vitae

Lucija Žužić was born in Rijeka, Croatia. She completed her elementary and high school education in Rijeka. She received recognition from the Ministry of Science and Education of the Republic of Croatia for extraordinary success in the high school graduation examinations, and received an award on the 31^{st} of July 2018 for a 100% matriculation final test score in both Mathematics and Informatics.

During her college education, she participated in the Huawei Croatia Seeds for the Future 2020 program and was selected as one of three participants from the Faculty of Engineering at the University of Rijeka. The program lasted from the 16^{th} to the 20^{th} of November 2020, and Lucija Žužić was among the top 3 students with the highest grades among 19 Croatian participants, learning topics such as 5G, artificial intelligence, and cloud computing.

Lucija Žužić also took part in Google's Get Ahead Program, and was selected as one of 500 participants included between the 29^{th} of June and the 22^{nd} of August 2020. She also participated in the Google Hash Code Online Qualification Round at the local hub at the Faculty of Engineering, University of Rijeka, from 2019 to 2021.

Lucija Žužić also competed in a coding competition for teams from Croatian universities, solving algorithmic tasks with two teammates from the Faculty of Engineering, University of Rijeka, from 2018 to 2020, and qualified for the International Collegiate Programming Contest (ICPC) Central Europe Regional Contest 2020, finishing as the 53^{nd} team out of 86.

Her coding skills include intermediate knowledge of C/C++, Python, and Java, as well as basic knowledge of JavaScript, HTML, PHP, SQL, CSS, XML, and web frontend technologies (React, Vue, and Angular). Her interests include competitive programming, algorithms, web development, and Android application development. Her internship with Sedam IT d.o.o. during her master's and bachelor's studies included web frontend development with Vue.js, using Google Firebase and Firestore for the backend database.

During her high school education, she took part in the "FIRST" robotics competition, Long Island Regional #2, at the Hofstra University, Hempstead, New York, the United States of America, from the 12^{th} to the 14^{th} April 2018, as a member of the team Crobotics, which was the first Croatian team to participate in this contest. The team of high school students designed, built,

and programmed a remotely controlled robot using the LabVIEW visual integrated development environment (IDE) and a robotics development kit. The team received the Highest Rookie Seed Award (highest-placed rookie team) and the Rookie Inspiration Award (rookie team with the greatest team spirit).

In 2017, she participated in the competition by IRIM and Croatian Makers Associations, "Creative projects with the BBC micro:bit: Smart piggy bank", as a member of the team that won the 2^{nd} place at the national level.

She took part in six rounds of solving algorithmic tasks every year from 2012 to 2017 in the Croatian Open Competition in Informatics (COCI). She achieved notable results in the Croatian high school computer science competition (Infokup) and participated once each year from 2010 to 2017. In the category of Theoretical Computer Science, she placed 14^{th} at the national level (highest out of three competition levels) in 2016. In the Algorithms category, coding in C++, she qualified for the county level (2^{nd} level out of three competition levels) every year from 2010 to 2017.

She earned her Bachelor's and Master's degrees in Computer Engineering from the Faculty of Engineering at the University of Rijeka. During her graduate studies, she demonstrated exceptional academic success, earning the Rector's Award for Excellence. In 2023, she was honored with the prestigious title of "Student of the Generation" by the University of Rijeka, recognizing her as the most outstanding student of her cohort.

Currently, she is employed as a Research Assistant at the Department of Computer Engineering, Faculty of Engineering, University of Rijeka, where she is also pursuing her doctoral studies under the supervision of Prof. Jonatan Lerga, PhD. Her primary research interests lie in machine learning and deep learning, with a specific focus on time series data, trajectory forecasting, anomaly detection, signal processing, computer vision applications in maritime environments, and artificial intelligence in biology and medicine. In addition to her maritime research, she actively collaborates on interdisciplinary projects, notably applying generative methods in bioinformatics to peptide discovery as part of her master's and bachelor's theses, and contributing to aerospace and ionospheric research.

Her research output is significant, having co-authored papers in high-impact international journals such as *Nature Machine Intelligence*, *Pattern Recognition*, *Knowledge-Based Systems*, *Expert Systems with Applications*, *Urban Science*, and the *Journal of the Franklin Institute*.

List of publications

[1] Dumenčić, S., Žužić, L., and Breš, B. Trajectory Analysis of Griffon Vulture in Kvarner Bay. *The Journal of CIEES 2*, 2 (May 2023), 36–43. eprint: https://doi.org/10.48149/jciees.2022.2.2.6.

[2] Livojević, A., Žužić, L., and Lerga, J. Image Compression Using Base64, BWT, RLE and LZW Coding. In *Proceedings of the Baška SIF (Spatial Intelligence Forum) Meeting 2024* (Baška, Croatia, June 2024).

[3] Njirjak, M., Žužić, L., Babić, M., Janković, P., Otović, E., Kalafatović, D., and Mauša, G. Reshaping the discovery of self-assembling peptides with generative AI guided by hybrid deep learning. *Nature Machine Intelligence 6* (Dec. 2024), 1487–1500. eprint: https://doi.org/10.1038/s42256-024-00928-1.

[4] Žužić, L. Predviđanje putanje osobnih plovila korištenjem modela temeljenih na neuronskim mrežama s povratnom vezom. In *Proceedings of the Scientific conference of doctoral students: INFCON24* (Rijeka, Croatia, Sept. 2024), University of Rijeka, Faculty of Informatics and Digital Technologies, pp. 68–69. eprint: https://inf.uniri.hr/images/studiji/poslijediplomski/INFCON/2024/Infcon_proc_2024.pdf.

[5] Žužić, L. Segmentacija slika korištenjem modela specifičnih za pojedinu klasu temeljenih na dubokom učenju, Gaussovoj mješavini i Bayesovoj vjerojatnosti. In *Proceedings of the Scientific conference of doctoral students: INFCON25* (Rijeka, Croatia, Sept. 2025), University of Rijeka, Faculty of Informatics and Digital Technologies, pp. 47–48. eprint: https://www.inf.uniri.hr/images/studiji/poslijediplomski/INFCON/2025/Infcon_proc_2025_FIN.pdf.

[6] Žužić, L., Dražić, I., Simčić, L., Hržić, F., and Lerga, J. A Bayesian and Markov chain approach to short-term and long-term personal watercraft trajectory forecasting. *Journal of the Franklin Institute 362*, 3 (Feb. 2025), 107509. eprint: https://doi.org/10.1016/j.jfranklin.2025.107509.

[7] Žužić, L., and Filjar, R. Circular Analysis of Long-Term GNSS Horizontal Positioning Error Time Series Observed in the Sub-Equatorial Region. In *Proceedings of the 2025*

33rd Telecommunications Forum (TELFOR) (Belgrade, Serbia, Nov. 2025), IEEE, pp. 1–4. eprint: https://doi.org/10.1109/TELFOR67910.2025.11314378.

[8] Žužić, L., Hržić, F., and Lerga, J. AI-based prediction of trajectories for personal watercrafts. In *Proceedings of the 8th Annual PhD Conference on Engineering and Technology MFC 2024* (Rijeka, Croatia, Sept. 2024), University of Rijeka, Faculty of Maritime Studies, p. 16. eprint: https://mfc.uniri.hr/wp-content/uploads/2024/09/MFC_2024_Book_of_Abstracts.pdf.

[9] Žužić, L., Hržić, F., and Lerga, J. Forecasting the trajectory of personal watercrafts using models based on recurrent neural networks. *Expert Systems with Applications 284* (July 2025), 127964. eprint: https://doi.org/10.1016/j.eswa.2025.127964.

[10] Žužić, L., Hržić, F., and Lerga, J. Collision course detection for personal watercrafts using models based on recurrent neural networks. *Knowledge-Based Systems 333* (Jan. 2026), 114974. eprint: https://doi.org/10.1016/j.knosys.2025.114974.

[11] Žužić, L., Hržić, F., Li, X., and Lerga, J. Class-specific image segmentation across multiple domains using customized U-Net pipelines. *Expert Systems with Applications 296*, Part D, 15 (Jan. 2026), 129203. eprint: https://doi.org/10.1016/j.eswa.2025.129203.

[12] Žužić, L., Hržić, F., Petrijevčanin, I., and Lerga, J. Inflection point-based trajectory fingerprinting for clustering and anomaly detection. *Pattern Recognition 174* (June 2026), 112977. eprint: https://doi.org/10.1016/j.patcog.2025.112977.

[13] Žužić, L., Hržić, F., Šimic, M., Otulić, I., and Lerga, J. Forecasting Horizons for Watercraft Trajectory Intersection Detection. In *Proceedings of the 12th EAI International Conference on Smart Objects and Technologies for Social Good* (Dubrovnik, Croatia, April 2026), Springer Nature Switzerland.

[14] Žužić, L., Klen, D., Špoljar, D., and Filjar, R. Machine Learning Algorithm for the Influence of Space Weather Events of Geomagnetic Activity on GNSS Services. *The Journal of CIEES 6*, 1 (Apr. 2026), 7–15. eprint: https://doi.org/10.48149/jciees.2026.6.1.1.

[15] Žužić, L., and Lerga, J. A Versatile Image Annotation and Masking Tool. In *Proceedings of the 9th Annual PhD Conference on Engineering and Technology MFC 2025* (Rijeka, Croatia, Sept. 2025), University of Rijeka, Faculty of Civil Engineering, p. 25. eprint: https://mfc.uniri.hr/wp-content/uploads/2025/09/MFC2025_BoA_final.pdf.

[16] Žužić, L., and Lerga, J. A Web-Based Card Game Simulation for Optimal Strategy Development. In *Proceedings of the 2025 MIPRO 48th ICT and Electronics Convention* (Opatija, Croatia, June 2025), IEEE, pp. 1–6. eprint: https://doi.org/10.1109/mipro65660.2025.11131986.

[17] Žužić, L., and Lerga, J. Automatic Image Segment Classification Using Contour Features. In *Proceedings of the Baška SIF (Spatial Intelligence Forum) Meeting 2025* (Baška, Croatia, Sept. 2025), University of Rijeka, Faculty of Engineering. eprint: https://riteh.uniri.hr/wp-content/uploads/Baska-SIF-Meeting-2025-Proceedings-1.pdf.

[18] Žužić, L., and Lerga, J. Enhancing Greenhouse Gas Emission Awareness Through Web-based Tools. In *Proceedings of the 2026 MIPRO 49th ICT and Electronics Convention* (Opatija, Croatia, May 2026), IEEE.

[19] Žužić, L., Škrlj, L., Nešković, A., and Filjar, R. Characterisation of GPS Horizontal Positioning Errors and Dst Using Recurrence Plot Analysis in Sub-Equatorial Ionospheric Conditions. *Urban Science 9*, 11 (Oct. 2025), 451. eprint: https://doi.org/10.3390/urbansci9110451.

[20] Žužić, L., Štajduhar, I., Lerga, J., and Filjar, R. Non-Parametric Model for Curvature Classification of Departure Flight Trajectory Segments. In *Proceedings of the 2025 6th International Conference on Communications, Information, Electronic and Energy Systems (CIEES)* (Ruse, Bulgaria, Nov. 2025), MDPI Engineering Proceedings, pp. 1–14. eprint: https://doi.org/10.3390/engproc2026122001.