(empty page)

Ship Motion Prediction with Deep Learning using IMU Data and Images

THIS IS A PLACE HOLDER TITLE PAGE

Lance De Waele

2022

(second title page)

(confidentiality notice)

# Preface

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Fusce ultricies, orci et scelerisque volutpat, nibh metus vestibulum ipsum, quis convallis ex orci ut massa. Curabitur felis dolor, tempor eu interdum nec, mattis quis felis. Vestibulum in nibh sit amet quam porta tristique. Fusce eu tortor tempus, tincidunt tortor hendrerit, sollicitudin elit. Cras a tempor urna. Vivamus vel malesuada purus. Sed feugiat egestas dolor, at feugiat lorem. Aliquam erat volutpat. Ut vel suscipit mi, quis vehicula lacus. Duis vitae libero semper, dignissim risus quis, vulputate augue. Praesent libero mauris, pretium id pharetra eget, malesuada et augue. Donec sed tincidunt augue. Nunc condimentum lectus non augue consequat, eget malesuada felis volutpat. Vestibulum ornare ultricies orci. Nulla sit amet dictum justo, non commodo arcu.

# Abstract

(English)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Fusce ultricies, orci et scelerisque volutpat, nibh metus vestibulum ipsum, quis convallis ex orci ut massa. Curabitur felis dolor, tempor eu interdum nec, mattis quis felis. Vestibulum in nibh sit amet quam porta tristique. Fusce eu tortor tempus, tincidunt tortor hendrerit, sollicitudin elit. Cras a tempor urna. Vivamus vel malesuada purus. Sed feugiat egestas dolor, at feugiat lorem. Aliquam erat volutpat. Ut vel suscipit mi, quis vehicula lacus. Duis vitae libero semper, dignissim risus quis, vulputate augue. Praesent libero mauris, pretium id pharetra eget, malesuada et augue. Donec sed tincidunt augue. Nunc condimentum lectus non augue consequat, eget malesuada felis volutpat. Vestibulum ornare ultricies orci. Nulla sit amet dictum justo, non commodo arcu.

# Extended abstract

(Nederlands)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Fusce ultricies, orci et scelerisque volutpat, nibh metus vestibulum ipsum, quis convallis ex orci ut massa. Curabitur felis dolor, tempor eu interdum nec, mattis quis felis. Vestibulum in nibh sit amet quam porta tristique. Fusce eu tortor tempus, tincidunt tortor hendrerit, sollicitudin elit. Cras a tempor urna. Vivamus vel malesuada purus. Sed feugiat egestas dolor, at feugiat lorem. Aliquam erat volutpat. Ut vel suscipit mi, quis vehicula lacus. Duis vitae libero semper, dignissim risus quis, vulputate augue. Praesent libero mauris, pretium id pharetra eget, malesuada et augue. Donec sed tincidunt augue. Nunc condimentum lectus non augue consequat, eget malesuada felis volutpat. Vestibulum ornare ultricies orci. Nulla sit amet dictum justo, non commodo arcu.

# Table of contents

[Preface 4](#_Toc99541024)

[Abstract 5](#_Toc99541025)

[Extended abstract 5](#_Toc99541026)

[Table of contents 6](#_Toc99541027)

[List of Figures 7](#_Toc99541028)

[List of Abbreviations 8](#_Toc99541029)

[Introduction 9](#_Toc99541030)

[ASV sensors 10](#_Toc99541031)

[Ship motion in six degrees of freedom 10](#_Toc99541032)

[Wave dynamics 12](#_Toc99541033)

[Technologies 12](#_Toc99541034)

[Simulation data 13](#_Toc99541035)

[Neural network architectures 13](#_Toc99541036)

[1 Related work 14](#_Toc99541037)

[2 Objectives 14](#_Toc99541038)

[3 Data 15](#_Toc99541039)

[3.1 Simulated data 15](#_Toc99541040)

[3.2 Real data 15](#_Toc99541041)

[3.3 Data analysis 15](#_Toc99541042)

[3.4 Data preprocessing 16](#_Toc99541043)

[3.5 Data loading 16](#_Toc99541044)

[4 Proposed models 16](#_Toc99541045)

[5 References 17](#_Toc99541046)

# List of Figures

[Figure 1: Prototype of Autonomous Surface Vessel (ASV) with on-board computer and sensors (MarSur – Robotics & Autonomous Systems, n.d.) 9](file:///D:\1%20Documenten\2021-2022%20Industrieel%20Ir\Thesis\Thesis\Documentation\paper.docx#_Toc99489943)

[Figure 2: ZED-mini IMU stereo camera (StereoLabs, Paris, France) 10](file:///D:\1%20Documenten\2021-2022%20Industrieel%20Ir\Thesis\Thesis\Documentation\paper.docx#_Toc99489944)

[Figure 3: Six degrees of freedom in ship motion(de Masi et al., 2011) 11](file:///D:\1%20Documenten\2021-2022%20Industrieel%20Ir\Thesis\Thesis\Documentation\paper.docx#_Toc99489945)

[Figure 4: Heave (m) in function of time (s) (Ham et al., 2017) 12](file:///D:\1%20Documenten\2021-2022%20Industrieel%20Ir\Thesis\Thesis\Documentation\paper.docx#_Toc99489946)

[Figure 5: structure of a basic neural network 13](file:///D:\1%20Documenten\2021-2022%20Industrieel%20Ir\Thesis\Thesis\Documentation\paper.docx#_Toc99489947)

# List of Abbreviations

AI: Artificial Intelligence

ASV: Autonomous Surface Vessel

IMU: Inertial Measurement Unit

PoV: Point of View

GPU: Graphics Processing Unit

LSTM: Long Short-Term Memory

CNN: Convolutional Neural Network

NN: Neural Network

ReLU: Rectified Linear Unit

# Introduction

In the last few years, the world has seen an exponential increase in technological advancements. This evolution brought with it a new influence of autonomous systems controlled by artificial intelligence (AI). Each of these systems designed with its own goals and characteristics, optimized for its task. More and more of these systems are being deployed as a direct or indirect replacement for tasks humans could do, but also, for tasks humans can’t physically do. And because these autonomous systems are optimized for specific jobs, they can be more accurate and faster at it than humans.

Because autonomous systems can replace the position of a human, they are especially useful in military operations. They can take over the role of a human in dangerous environments such as an active warzone and therefore eliminate the risk of someone’s life. On the other hand, they can also be used as a complimentary asset, providing support and aid in logistics. More and more of these autonomous assets such as drones, surface vessels, tanks and reconnaissance vehicles are being deployed around the world for various objectives. But with this increasing amount of autonomous assets, there is need for communication between them, to allow them to work together and be aware of the state of each other when they need to interact (de Cubber, 2019).

“*Interoperability is the key that acts as the glue among the different units within the team, enabling efficient multi-robot cooperation.”* (*MarSur – Robotics & Autonomous Systems*, n.d.)



Figure 1: Prototype of Autonomous Surface Vessel (ASV) with on-board computer and sensors (MarSur – Robotics & Autonomous Systems, n.d.)

The Robotics & Autonomous Systems lab of the Belgian Royal Military Academy is currently working on two autonomous vehicles in two projects named MarSur (*MarSur – Robotics & Autonomous Systems*, n.d.)and MarLand (*MarLand – Robotics & Autonomous Systems*, n.d.). Project MarSur is developing an autonomous surface vessel (ASV) (figure 1) that will interact with a drone that is being developed by project MarLand. The drone needs to be able to take-off and land on the ASV. This proposes a challenge since the ASV is continuously moving due to sea waves and can therefore not provide a stable landing surface. For a smooth landing to be possible, the ASV must be capable to determine its state in a three-dimensional space and predict its movement in the ocean. These predictions must provide an accurate estimation over a future time series to determine the optimal time for the drone to land. This optimal time is a period where the ASV is as stable as possible so that the impact on the drone will be minimized. To facilitate these predictions, the ASV is equipped with an on-board computer and multiple sensors.

## ASV sensors

The ASV is equipped with a ZED-mini stereo IMU camera (figure 2) (StereoLabs, Paris, France). This is a multipurpose sensor that can provide both video from its cameras and numeric data from its Inertial Measurement Unit (IMU), to accurately describe the state of the sensor and its surroundings. The ZED-mini has two built-in motion sensors: an accelerometer and a gyroscope. These provide a real-time data stream at 800Hz of the movement of the sensor along the rotational and translational axes. These types of movements will be described more in depth in the next part of the introduction.

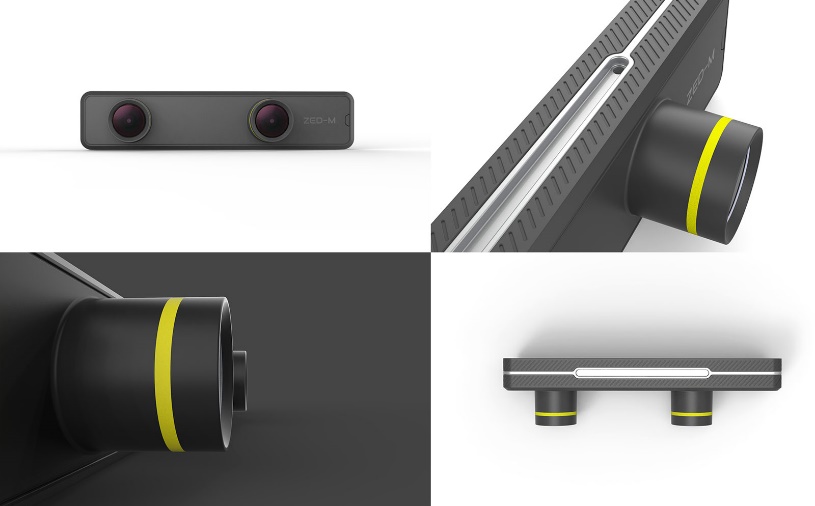


Figure 2: ZED-mini IMU stereo camera (StereoLabs, Paris, France)

The two forward facing lenses on the ZED-mini provide stereo video. This can be used to map the objects in front of the ZED-mini in a three-dimensional space. A stereo image is a combination of two separate images that are captured from two slightly offset point-of-views (PoV) such as the lenses on the ZED-mini. These two PoV’s imitate the left and right human eye and create a perception of depth when the two images are fused together to create one stereo image. This process is called stereoscopy. In computer vision, these two images can be compared to each other to extract three-dimensional information from two dimensional images.

The data from the motion sensors and the camera’s will be used as input to predict the future movement of the vessel. The deep learning models will digest a sequence of data and images of the past to predict a sequence of data in the future. To do this, the model will try to find trends in the data and continue these trends with regards to the information gathered from the images with incoming waves. Two important parameters will be how much historic data the model takes as input and how much data it can accurately predict in the future.

## Ship motion in six degrees of freedom

The motion of a ship or any rigid object in a three-dimensional space can be described in six degrees of freedom. These six degrees can further be divided into two categories: translational and rotational movement. Where translational movement is movement along one of the three axes in a three-dimensional space, rotational movement is the rotation of an object around these same three axes. These three reference axes run through the center of mass of the ship and are oriented as follows:

* Vertical Z-axis runs vertically through the vessel
* Transverse Y-axis runs horizontally across the vessel
* Longitudinal X-axis runs horizontally through the length of the ship

Each type of motion, translation or rotation, among each of the three axes has a different impact on the movement of the vessel (figure 3). The translational movements are expressed in linear units such as meters and are named as followed:

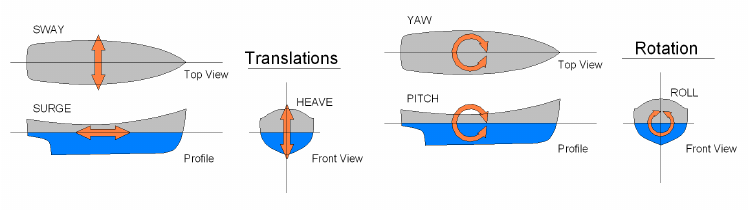


Figure 3: Six degrees of freedom in ship motion(de Masi et al., 2011)

* **Sway**: side to side movement along the transverse Y-axis
* **Surge**: forward and backwards movement along the longitudinal X-axis
* **Heave**: upward and downward movement along the vertical Z-axis

The rotational movements are expressed in angular units and are named as followed:

* **Yaw**: rotational movement around the vertical Z-axis
* **Pitch**: rotational movement around the transverse Y-axis
* **Roll**: rotational movement around the longitudinal X-axis

To predict the motion of a ship, one must differentiate between these different motions. Together they form the complete three-dimensional orientation of a ship. But not all of them need to be predicted. Surge and yaw are controlled by the ASV’s autonomous systems and are respectively controlled by the amount of thrust and the rudder position – steering the ship. Surge and yaw will also not change very drastically during the landing or take-off of a drone since this behavior would directly impede our main goal of providing a smooth landing. On the other hand, the sway of a ship, also referred to as drift, is primarily caused by sideways winds or currents in the water and will have minor impact on the stability of the ship whenever the drone needs to take-off or land. If the drone aims for a GPS-tracker present in the ASV, it will follow the vessels movement no matter the sway.

This leaves three main factors remaining which have the most impact on the stability of the vessel: roll, pitch and heave. These three movements have one thing in common, they are all directly caused by the waves in the ocean and are very hard to control. Different methods exist to dampen these movements and keep the vessel as stable as possible such as bilge keels and antiroll tanks. However, most of them are either infeasible or ineffective or don’t provide the required stabilization on smaller vessels (Perez & Blanke, 2017). In this case, predicting these movements instead of trying to dampen them, can be an alternative solution. Although it should be noted that using them together, will most likely yield the best performance. Pitch, roll and heave can be divided in two categories based on the effect they have on the landing and take-off of the drone. Pitch and roll are responsible for the stability of the landing surface and heave is responsible for the impact on the drone when landing.

To provide a stable landing zone for the drone, the pitch and roll of the vessel should remain constant and as close to zero as possible. Depending on the characteristics of the drone, the model should be able to analyze its prediction sequence and find a window where the desired circumstances to land/take-off are met. To determine this window of landing/take-off opportunity, different parameters need to be defined such as the maximum difference in consequent prediction values, the length of the window and the interval in which all predicted values should lie. For example, the roll and pitch values should all remain in a [-3°, 3°] interval, the stable window duration must be at least five seconds and there should be no difference larger than two degrees between consecutive predicted values.

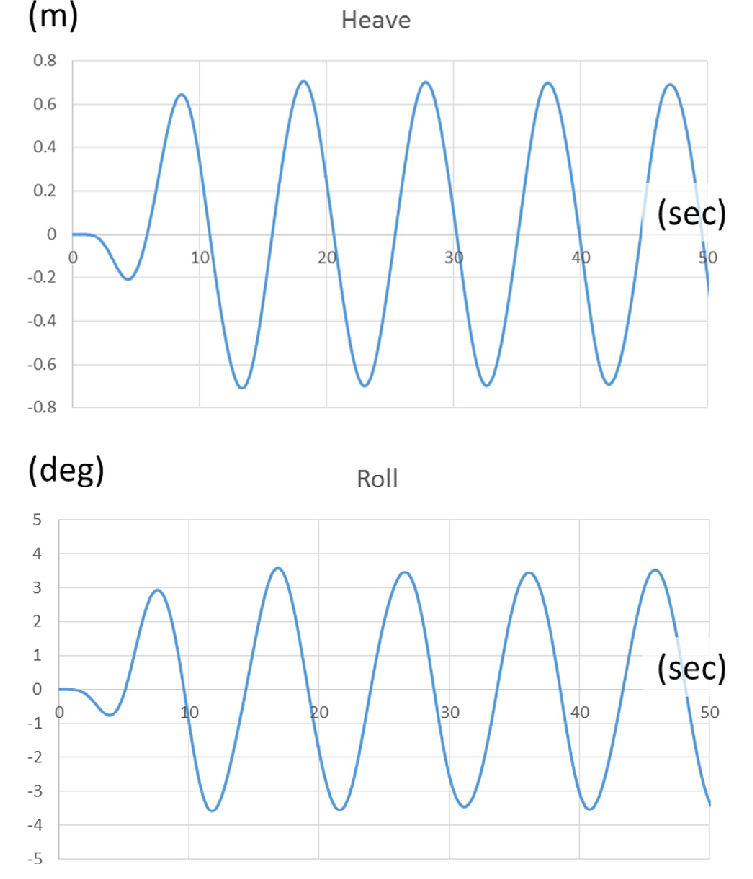


Figure 4: Heave (m) in function of time (s) (Ham et al., 2017)

To minimize the impact on the drone when landing, the heave needs to be constant or decreasing. This means that the vessel is either not moving up or down, or it is slowly moving downwards following the motion of the descending drone. In regular waves, the heave of a vessel follows a wave-like function, alternating between upwards and downwards motion (figure 4). In this case, the window of take-off/landing opportunity can be defined as the points where the vessel transitions from upwards to downwards motion, or vice versa, and thus has an acceleration of zero.

## Wave dynamics

?

## Technologies

To develop the deep learning neural networks, Python[[1]](#footnote-1) will be used. Python is a high-level, general-purpose programming language that is in - most cases - most suited for machine learning applications. To facilitate the development of the deep learning models, different packages were used. The most important of which are briefly discussed below.

Pandas, Seaborn and NumPy were used to process and manipulate data. NumPy provides functionality to perform mathematical functions on large datasets with multiple dimensions. Pandas provides the functionality to visualize and order the data in its Data frames. Seaborn was mainly used to perform data analysis on the numeric data such as pitch and roll, to assess and visualize statistical information between features. To plot training and test results, Matplotlib was used.

For the implementation of the deep learning models, PyTorch was used. PyTorch is an open-source framework for machine learning applications that allows fast development and ease of use. Pytorch was used together with the cudatoolkit extension. This enabled the models to be trained and tested on a dedicated graphics processing unit (GPU) to increase computing performance.

Blender was used to generate a dataset from an ocean wave simulation. Blender is an open-source three-dimensional creation suite that can be used for many purposes. In this case, it was used to simulate incoming waves on a vessel. The data itself will be discussed more in detail bellow.

All code was written Jupyter Notebook for its ease of use and simple debugging. Whenever parts of the code in the notebooks were tested thoroughly, they are extracted to standalone Python files and accessed via import statements.

## Simulation data

During most of the development, a simulation dataset was used from an ocean wave simulation made in Blender (*Nazotron1923/Ship-Ocean\_simulation\_BLENDER: Python Scripts for Generation Sea States and Ship Motion Using 3D Blender Simulator*, n.d.). This dataset was made by Nazar-Mykola Kaminskyi and is publicly available through GitHub. It was not made by or in cooperation with the research group of this thesis. Simulation data was needed for training and testing purposes during the initial phases of development when real data from the ASV was not yet collected and available.

For this simulation, a standard model of a vessel was used which is floating on the simulated sea surface and moves along with the waves. This presents the issue that all models trained with this data will be biased to this vessel’s characteristics. A large, heavy vessel will behave very different opposed to a small, lightweight vessel. Therefore, all models should be retrained and re-evaluated with real data from the vessel it will be used on, as this vessel will most likely have different characteristics. However, this issue can also form a way to compare how well one model can predict the motion for different vessels. If there are only minor increases in performance when the model is retrained with data from the specific vessel, it might be more useful to use one general purpose model for all vessels with similar characteristics.

Using a simulation also presents a second issue: perfect conditions. The data taken from the simulation is from a perfect scenario, meaning that there are no obstructions or other objects on the images. The images are also perfectly stabilized on the moving vessel. This can cause models that are trained on the simulation data to perform very well yet fail to meet desired expectations in a real-world scenario. To minimize this effect, data can be augmented to include more variation, or adjustments can be made to the simulation. Adjustments such as including passing vessels in the images can make it more realistic.

## Neural network architectures

Neural networks (figure 5), also known as artificial neural networks (ANN) are the building blocks of deep learning problems, which are a subset of machine learning problems and artificial intelligence. Neural networks are structures inspired by the human brain, more specifically the neurons within and how they pass signals from one to another. A neural network consists of different layers built with multiple neurons. These neurons are connected to neurons and pass data forward over these connections. The output of one neuron is the input of all the neurons it connects to. This way data is fed forward through the network and updated in every neuron. There are three main parts in a neural network, the input layer, output layer and hidden layers. The input layer has the same number of neurons as the features in the input data. The output layer has the same number of neurons as the predicted output features. In our case, the output features will be pitch, roll and heave.

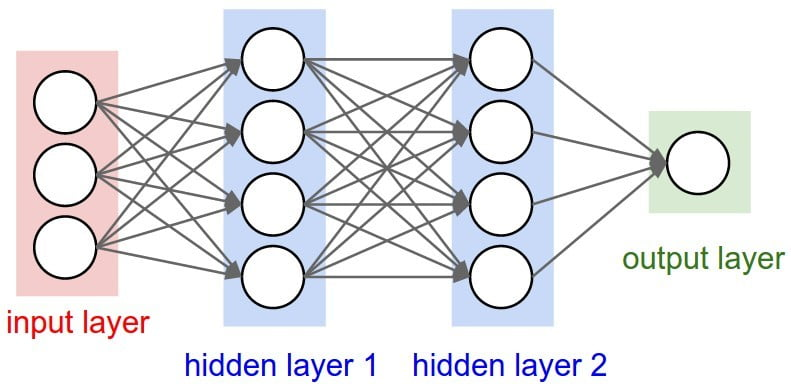


Figure 5: structure of a basic neural network

Intro???????????????????????????

Heavily used in related work.

# Related work

A big source of information for this this thesis was the work of Nazar-Mykola Kaminskyi (Kaminskyi, 2019). Kaminskyi also tried to create different neural networks that can predict the motion of a vessel based on pitch, roll and incoming wave images. He created the dataset that was used for a big portion of this thesis and provided research and results from his research project. His results were used to set up criteria to which our models had to comply. This thesis should not be seen as a replacement for Kaminskyi’s work though, but more as continuation and expansion of his work. Kaminskyi never tested his models on real life data and only used pitch and roll as input together with the images. With the ZED-mini, our models will have access to more data and should therefore potentially perform better.

# Objectives

The main goal of this research is to develop a neural network capable of predicting the motion of a surface vessel. To do this, the model ingests a sequence of values describing the state of the vessel and images of incoming waves taken from a stabilized camera pointing to the front of the vessel. As an output, the model should provide a sequence of predicted pitch, roll and heave values for every predicted second. The sequence duration for these predictions should be at minimum thirty seconds to provide the drone with ample time to complete a take-off or landing procedure. Bigger drones will need more time for this procedure so the sequence duration should be maximized within the model’s capabilities.

Since the model will be deployed in real-time scenario’s and will be making predictions in real-time based on the continuous data stream of the ZED-mini, it should be lightweight and not require large amounts of computational power. This means that the inference time of the different proposed models should also be considered when comparing different models’ performances. Both prediction errors and inference time should thus be minimized.

To test the viability of each model, it should be subjected to a set of minimum criteria in the simulation environment. A model is fit to be used with real data only when it meets these criteria. To make this selection, the following attributes of the model must comply with their respective requirements. The root mean square error (RMSE) for the predicted angles for pitch and roll at the 10th second should be no more than three degrees and no more than five degrees at the 30th predicted second. These values are derived from the performance of the best models from Kaminskyi’s work. The RMSE of the heave of the vessel should never exceed twenty centimeters. Finally, the inference time of the model should be no more than fifty milliseconds.

Whenever a prediction sequence is calculated, it should be analyzed by an algorithm. This algorithm will search the sequence for a window of landing/take-off opportunity. This window is a ten second time frame. However, the duration of this window should be easily changeable, to compensate for different sized drones. During this window, heave, roll and pitch should remain close to constant. To define this constant behavior, the next set of criteria will be used. Pitch and roll angles should remain in a [-3°, 3°] interval and consecutive values should have no difference bigger than two degrees. Consecutive heave values should have no difference bigger than one centimeter. For heave, there is no interval in which all predicted values should remain since this is highly dependent on the height of the waves.

# Data

The first step of any machine learning operation is collecting data. For this thesis, data was collected from two sources as discussed in the introduction. A dataset from simulation and real-world data was used. This poses some challenges as the simulation data is captured in an ideal scenario and contains less motion parameters as the real-world counterpart. This can cause the models to perform differently and have slightly different architectures due to the different inputs and outputs based on the available data. In the following two sections, both datasets will be discussed in more detail.

## Simulated data

The simulated dataset contains images of incoming waves (figure 6) and the pitch and roll values of a vessel floating on these incoming waves. The data was generated in 540 different episodes, each episode containing 400 frames of data. The frames were captured at two frames per second to minimize data overlap in consecutive images. Each frame contains one image together with the state of the vessel at the time of the image in the form of a pitch- and roll-value tuple. In its totality, this dataset contains 216.000 frames, which translates to thirty hours of simulated data at two frames per seconds.

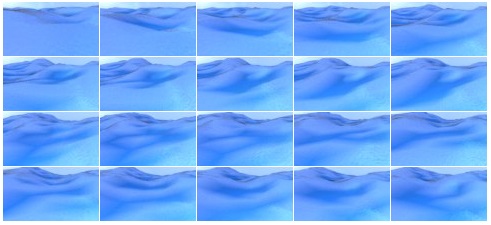


Figure 6: Generated images of incoming waves

With this simulated data we only have access to pitch and roll as input data. This means that during testing in the simulation environment, heave cannot be predicted. However, with the neural networks that will be proposed later, this should not be a big issue for two main reasons. First off, heave follows a somewhat predictable pattern in regular conditions as depicted by figure 4. Because of this, it was assumed not necessary to regenerate the full dataset for just one extra output. It was also assumed that if the model performed well on pitch and roll, it should consequently perform well on heave. Secondly, adding extra input and/or output features to a neural network architecture, is not an expensive operation. All models will have to be retrained anyway when switching from the simulation environment to a real-world environment to adapt to the new vessels’ characteristics and the additional data available from the ZED-mini.

To simulate this data, different parameters were used to finetune the simulation environment. The effects of these parameters on the simulation were not tested as the used dataset was already generated with fixed parameters. However, assumptions can be made on how they would affect the simulation. The most important parameters will be discussed briefly.

As mentioned above, the images were taken at two frames per second by a virtual camera. The position of the camera on the simulated vessel was set with the following parameters:

* Height: *5 meters*
* Rotation around x-axis: *76 degrees (slightly tilted downwards)*

The images were captured in a low resolution to decrease the memory needed to load the dataset. The following resolution parameters were used:

* Height: *54 pixels*
* Width:  *96 pixels*

## Real data

At this moment, data has been captured from the ASV but is not yet available.

The real data will be fed into the model at real-time. This means that the data stream will have to be cleaned and filtered to the desired input format for the model. Numeric motion parameters such as pitch and roll, are output by the ZED-mini at 800Hz. This data stream will have to be reduced to minimize data overlap. Since the simulation data was generated at two frames per second, 2Hz should be a good starting frequency to test the models. Later on, this frequency can be increased or further decreased depending on the model’s performance and inference time.

Similarly, the video framerate from the stereo camera should also be reduced to minimize data overlap between images and reduce computational requirements. The video footage will also have to be compressed to reduce the memory needed to process the images.

## Data analysis

Correlation

## Data preprocessing

Normalize

## Data loading

The simulated data is generated in episodes. Each episode containing 400 frames that chronologically follow each other as can be seen in figure 6. However, the episodes itself do not follow each other chronologically. Different episodes are generated in different simulation sessions and should therefore not be seen as continuous. When using this data with deep learning models that ingest and predict sequences, the data must be sequenced correctly to not include data from different episodes. More specifically, all frames in the input and output sequence, should always be from the same episode.

# Proposed models

In this chapter, all models will be discussed that were created.

# References

de Cubber, G. (2019). *OPPORTUNITIES AND SECURITY THREATS POSED BY NEW TECHNOLOGIES*.

de Masi, G., Gaggiotti, F., Bruschi, R., & Venturi, M. (2011). Ship motion prediction by radial basis neural networks. *IEEE SSCI 2011 - Symposium Series on Computational Intelligence - HIMA 2011: 2011 IEEE Workshop on Hybrid Intelligent Models and Applications*, 28–32. https://doi.org/10.1109/HIMA.2011.5953967

Ham, S.-H., Roh, M.-I., & Zhao, L. (2017). *Integrated method of analysis, visualization, and hardware for ship motion simulation*. https://doi.org/10.1016/j.jcde.2017.12.005

Kaminskyi, N.-M. (2019). *Deep learning models for ship motion prediction from images*.

*MarLand – Robotics & Autonomous Systems*. (n.d.). Retrieved March 29, 2022, from https://mecatron.rma.ac.be/index.php/projects/marland/

*MarSur – Robotics & Autonomous Systems*. (n.d.). Retrieved March 29, 2022, from https://mecatron.rma.ac.be/index.php/projects/marsur/

*Nazotron1923/ship-ocean\_simulation\_BLENDER: Python scripts for generation sea states and ship motion using 3D Blender simulator*. (n.d.). Retrieved March 29, 2022, from https://github.com/Nazotron1923/ship-ocean\_simulation\_BLENDER

Perez, T. ;, & Blanke, M. (2017). Ship Roll Damping Control. *Annual Reviews in Control*, *36*(1), 129–147. https://doi.org/10.1016/j.arcontrol.2012.03.010

1. https://www.python.org/ [↑](#footnote-ref-1)