

박 사 학 위 논 문
Ph.D. Dissertation

이미지를 이용한 기계학습 기반
해상 상태 예측

Sea state prediction based on machine learning
using images

2022

김 윤 호 (金 允 浩 Kim, Yun-Ho)

한 국 과 학 기 술 원

Korea Advanced Institute of Science and Technology

박 사 학 위 논 문

이미지를 이용한 기계학습 기반
해상 상태 예측

2022

김 윤 호

한 국 과 학 기 술 원

기계항공공학부/기계공학과

이미지를 이용한 기계학습 기반
해상 상태 예측

김 윤 호

위 논문은 한국과학기술원 박사학위논문으로
학위논문 심사위원회의 심사를 통과하였음

2022년 6월 14일

심사위원장 이 필 승

심사위원 김진환

심사위원 윤국진

심사위원 장대준

심사위원 강희진



Sea state prediction based on machine learning using images

Yun-Ho Kim

Advisor: Phill-Seung Lee

A dissertation/thesis submitted to the faculty of
Korea Advanced Institute of Science and Technology in
partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Daejeon, Korea

June 14, 2022

Approved by


Phill-Seung Lee
Professor of Mechanical Engineering

The study was conducted in accordance with Code of Research Ethics¹⁾.

1) Declaration of Ethical Conduct in Research: I, as a graduate student of Korea Advanced Institute of Science and Technology, hereby declare that I have not committed any act that may damage the credibility of my research. This includes, but is not limited to, falsification, thesis written by someone else, distortion of research findings, and plagiarism. I confirm that my dissertation contains honest conclusions based on my own careful research under the guidance of my advisor.

DME

김윤호. 이미지를 이용한 기계학습 기반 해상 상태 예측. 기계 공학과. 2022년. 92+viii 쪽. 지도교수: 이필승. (영문 논문)

Kim, Yun-Ho. Sea state prediction based on machine learning using images. School of Mechanical Engineering. 2022. 92+iii pages. Advisor: Lee, Phill-Seung. (Text in English)

초 록

선박의 안전한 항해는 인류 역사와 더불어 항상 강조되었으며, 이를 위한 엄밀한 계측장비, 고도화된 항해 기법은 나날이 발전하고 있다. 그럼에도 선박의 사고는 지속적으로 발생하고 있으며, 이 중 숙련되지 않은 선원의 판단에 기인한 사고도 유의미하게 발생한다. 또한 대형 선박의 경우 정밀 계측기기의 활용, 전문인력의 승선으로 해상 조건에 대한 판단을 최대한 정확하게 진행하려고 하나 연안을 운항하는 중소형 선박의 경우 중앙기관의 기상 예보 기반으로 전적으로 선장의 판단에 기인하여 선박 운항을 하게 된다.

본 연구에서는 급속도로 발전하고 있는 기계학습법을 이용하여 해상환경을 예측하는 시스템을 제안하였다. 해상 조건을 판단함에 있어 파도의 크기를 판단하는 것은 선장의 주요임무로서 이런 오랜 경험을 기반으로 Beaufort wind scale로서 바람과 파도에 대한 분류가 되어 있다. 이를 충분한 데이터 기반으로 해상 조건을 판단할 수 있는 적합한 기계학습법의 선별, 적용 및 이의 유효성을 판단하고자 하였다.

초기에는 수치 그래픽 데이터 파랑장 기반 학습을 통해 기계학습의 적용가능성 검토 및 학습모델의 특성을 파악하였다. 합성곱 신경망 기반으로 장파정파, 단파정파에 대하여 특정지점의 파고를 식별하는 시도들을 진행하였다. 이미지 전처리 기법 적용, 신경망 구조 변경, 하이퍼 파라미터의 보정 등을 다양하게 적용하였으나 전반적인 정확도를 향상시키기는 어려웠다. 이를 통해 엄밀 계측이 아닌, 해상 상태를 식별함에 있어 기계학습의 적용 가능성을 살펴보고자 하였다.

실제 해역에서의 데이터를 일정 기간 이상 획득하여 학습데이터를 생성하였다. 이 때 이미지 데이터는 한국 서남부 해역에서 카메라를 이용해 획득하였으며, 해상 상태는 기상청 제공 공개 데이터 및 파고 정보는 해저면 설치 파고계로부터 확보하였다. 단일 스냅샷 기반의 학습에서는 단순 합성곱 신경망 기반 학습은 한계를 보였기에, 합성곱 신경망 및 장단기 메모리 기반 딥러닝 네트워크를 적용하였다. 이 때 적절한 데이터 증폭 기법을 함께 사용했을 시, 짧은 영상기반 예측이 실제 해양환경에서도 적용되는 것을 확인하였다. 선박의 운동을 인위적으로 모사하여 적용한 이미지들에 본 네트워크의 유효성을 확인했으며, 이를 개선하고자 이미지 처리 기법들을 통한 화각 보정 및 재예측을 시도하였다.

마지막으로 본 연구에서 활용된 기법의 한계점 및 이를 개선하기 위한 계획들을 정리하였다. 또한 운항 중인 선박에서 획득한 이미지에 대한 예측을 통해 실제 문제에서의 초기 적용을 수행하였다.

핵심 낱말 해상 상태 예측, 기계학습, 합성곱 신경망, 장단기 메모리, 연속 이미지, 평균 파고

Abstract

Safe navigation of ships has always been emphasized along with human history, and for this purpose, rigorous measuring equipment and advanced navigation technologies are developing continuously. However, accidents of ships continue to occur, and among them, accidents due to the judgement from inexperienced ship crews occur meaningfully. In addition, in the case of larger ships, the use of precision measuring instruments and the use of professional personnel to make judgements on sea conditions as accurately as possible are attempted to sail the ship safely.

In this study, a system for predicting the ocean environment using the rapidly developing machine learning method was proposed. In judging sea states, it is the captain's main duty to determine the size and seriousness of waves, and based on this long experience, the winds and waves are classified as the Beaufort wind scale. The purpose of this study is to choose and apply a suitable machine learning method that can predict sea states with sufficient big data, and to evaluate its effectiveness.

Initially, the applicability of machine learning was reviewed and the characteristics of the learning model were identified through numerical graphical wave fields. Attempts were performed to identify the wave height of a certain point for long-crested and short-crested waves based on convolutional neural network. Various simulations with image pre-processing techniques, neural network changes, hyper-parameter tunings were conducted, while it was difficult to improve the prediction accuracy. Though this, we tried to examine the applicability of machine learning in identifying the sea state rather than the strict measurement.

Learning data was collected from the actual site over a certain period of time. The images were obtained using a camera in the southwestern area of Korea, and the sea states were obtained from the public data and a seabed installed wave height meter. In single snapshot-based learning, simple convolutional neural network-based learning showed limitations. Therefore, the combined model with convolutional neural network and long short-term memory was applied. In this case, when an appropriate data augmentation technique was utilized together, it was confirmed that the short video-clip based prediction can be applied to the real ocean environment. The applicability of this network was confirmed on images applied by artificially simulating the motion of a vessel, and to improve this, correction and re-prediction of the angle of view through image processing techniques were attempted.

Lastly, limitations of the present technique are arranged and possible solutions are introduced. In addition, the initial application to real problems was performed through prediction of images acquired from a ship in operation.

Keywords Ocean environment prediction; Machine learning; Convolutional neural network; Long short-term memory; sequential images; Average wave height

Contents

Contents	i
List of Tables	iii
List of Figures	V
Chapter 1. Introduction	1
Chapter 2. Conventional methodologies for sea-state measurement	7
Chapter 3. Machine learning methodologies	12
3.1 Convolutional neural network.....	12
3.2 Long short-term memory	17
3.3 Neural network architecture for movie-clip based machine learning	20
Chapter 4. Wave height estimation with numerical irregular waves	24
4.1 Estimation on long-crested waves.....	24
4.2 Estimation on short-crested waves.....	30
Chapter 5. Data acquisition and classification	47
Chapter 6. Training results with non-prescreened data	56
6.1 Initial training	56
6.2 Applicability evaluation with deformed images	58
Chapter 7. Training results with prescreened data	64
7.1 Single snapshot-based machine learning results	64
7.2 Video-based machine learning results	70
Chapter 8. Applications in marine engineering	79
Chapter 9. Concluding Remarks	86
Bibliography	87

List of Tables

1.1 Exclusive summary on artificial intelligence development	4
2.1 Development in ocean environment measuring instruments using X-band radar	8
3.1 Explanation on machine learning layers	15
3.2 Training options of CNN	16
3.3 Computing hardware specifications	16
3.4 GoogLeNet incarnation of the Inception architecture	22
3.5 Training options of LSTM	23
4.1 Correlation matrix for long-crested waves	27
4.2 Correlation matrix for black & white filtered long-crested waves	30
4.3 Correlation matrix for (f) filtered long-crested waves	30
4.4 Wave height estimation results with different filter applications	31
4.5 Correlation matrix for short-crested waves	36
4.6 Validation accuracy and computational time with various max pooling conditions	40
4.7 Convolution layer specifications in variation of number of convolution layers	40
4.8 Wave conditions with wave train 3 variations	42
5.1 Sea state classification in ITTC Recommended Procedure	50
5.2 Well-conditioned data gathered at ADCP installation location	51
5.3 Numbers of snapshot images utilized for the sea state estimation	52
5.4 Numbers of snapshot images utilized for the average wave height estimation	52
5.5 Normal conditioned data gathered at ADCP installation location	54
6.1 Correlation matrix for sea state estimation with normal conditioned data	57
6.2 Prediction results with image focusing	59
6.3 Prediction results with image rotating	59
6.4 Prediction results with image tilting	59
7.1 Prediction probability of testing data for the sea state	66
7.2 Probability of testing data for the sea state prediction using half of the data	67
7.3 Prediction probability of sample snapshots for the average wave height	69
7.4 Confusion matrix for average wave height classification with snapshots	69
7.5 Results of converting current photo data into video clips	71
7.6 Correlation matrix for the H_{ave} prediction with videos(Length=6s)	73

7.7 Correlation matrix for the H_{ave} prediction with videos(Length=30s).....	73
7.8 Correlation matrix for the H_{ave} prediction with videos(Length=60s).....	73
7.9 Correlation matrix for the H_{ave} prediction with videos(Length=180s).....	74
7.10 Correlation matrix for the H_{ave} prediction with videos(Length=300s).....	74
7.11 Data classification with augmentation in $H_{ave}=2.5m$	76
7.12 Correlation matrix for the H_{ave} prediction with augmented videos(Length=180s).....	78
7.13 Comparison in classification evaluation indicators.....	78
8.1 Statistical values of prediction accuracies on images from ship-mounted camera.....	84

List of Figures

1.1. Major maritime accidents related to the enactment and revision of SOLAS.....	1
1.2. Marine accident status by the type of accident in Korea (2016~2020).....	2
1.3. Marine accident status by ship usage in Korea (2016~2020).....	2
1.4. Historical achievements in artificial intelligence.....	3
1.5. Deep generative model for probabilistic now-casting of precipitation from radar in UK.....	5
2.1. WaveRadar REX installation setup(left) & wave field using X-band radar(right).....	8
2.2. The automated trinocular stereo imaging system (ATSIS) (left) and results (right).	9
2.3. Types of marine meteorological observations(left) and their installation locations in Korea(right).....	10
2.4. Small ship accident induced by insufficient and inadequate marine measuring system.....	11
3.1. Illustration of the AlexNet.....	12
3.2. Schematics of the convolution neural network.....	14
3.3. Proposed architecture of the convolution neural network (CNN).....	12
3.4. Example of a LSTM net.....	17
3.5. Repetitive modules of RNN with one layer.....	18
3.6. LSTM with four interactive layers.....	18
3.7. Repetitive modules of RNN with one layer.....	20
3.8. GoogLeNet network.....	21
3.9. LSTM network.....	22
4.1. Snapshots of long-crested waves.....	25
4.2. Wave elevation at the center of the domain.....	25
4.3. Classification by wave height.....	26
4.4. Estimation of each wave height category for long-crested waves.....	26
4.5. Black & White filtered long-crested wave snapshots.....	28
4.6. Original and cropped images.....	28
4.7. Various filtered images with cropped images.....	29
4.8. Estimation accuracy with different filter applications in long-crested waves.....	31
4.9. Short-crested wave spectrum.....	33

4.10. Incoming directions of wave trains.....	33
4.11. Snapshots of modelled short-crested wave.....	34
4.12. Progress report of global accuracy: short crested wave.....	34
4.13. Progress report of cross entropy loss: short crested wave.....	35
4.14. A sample ill-estimated case: Answer: '1.25' , Estimation: '-4.50'	36
4.15. Variation on max pooling technique application.....	37
4.16. Global accuracy with various max pooling conditions.....	38
4.17. Cross entropy loss with various max pooling conditions.....	38
4.18. Estimation accuracy by categories with various max pooling conditions.....	39
4.19. Global accuracy with various number of convolution layers.....	41
4.20. Cross entropy loss with various number of convolution layers.....	41
4.21. Accuracy by number of layers in categories with various number of convolution layers.....	42
4.22. Sea-state variations for (a) Case C1., (b) Case C3., and (c) Case C5.....	43
4.23. Global accuracy in wave spectrum variation.....	44
4.24. Cross entropy loss in wave spectrum variation.....	44
4.25. Estimation accuracy per category in wave spectrum variation.....	46
5.1. Locations of image data acquisition, wave height measuring instrument.....	47
5.2. Fixed-type wave energy converter installed at Jeju Island and CCTV camera.....	48
5.3. Acoustic Doppler Current Profiler(ADCP).....	48
5.4. Conceptual view of sea state classification by significant wave height.....	49
5.5. Sample snapshot images of well-conditioned for (a) sea state 4, (b) sea state 5, and (c) sea state 6.....	50
5.6. Labelling ratio of normal conditioned data.....	55
5.7. Sample snapshot images of normal conditioned data.....	55
6.1. Training results with normal data: Global accuracy.....	56
6.2. Training results with normal data: Cross entropy loss.....	57
6.3. Image distortion: focusing, rotating, and tilting.....	58
6.4. Image re-rotation process.....	60
6.5. Various vector extraction methods from an image.....	61
6.6. An example of angle detection from an image: Harris model & Minimum Eigenvalue method.....	61
6.7. Estimated angle with Harris model with 10deg rotated images.....	62

6.8. Estimated angle with Harris model with 20deg rotated images.....	62
6.9. Estimated angle with Harris model with 30deg rotated images.....	62
6.10. Comparison between predictions with rotated and re-rotated images.....	63
7.1. 5-fold cross validation for snapshot-based machine learning.....	64
7.2. Training results for the sea state Estimation: Global training accuracy of validating data.....	65
7.3. Training results for the sea state Estimation: Cross entropy loss of validating data.....	65
7.4. Training results for the sea state estimation: Global training accuracy of validating data.....	66
7.5. Training results for the wave height prediction: Global training accuracy of validating data.....	68
7.6. Training results for the wave height prediction: Cross entropy loss of validating data.....	68
7.7. Training results for the Have prediction with different lengths of videos: Global accuracy.....	71
7.8. Training results for the Have prediction with different lengths of videos: Cross entropy loss.....	72
7.9. Precision, recall, and accuracy for Have prediction with different lengths of videos	75
7.10. Data augmentation for ‘Have=2.50m’ category.....	76
7.11. Training results for the Have prediction with augmented videos: Global accuracy...	77
7.12. Training results for the Have prediction with augmented videos: Cross entropy loss.....	77
8.1. Images open to the public in real time.....	79
8.2. Schematic concept of developing plan of national coastal now-casting map.....	80
8.3. Korea government-affiliated marine research institute infrastructure.....	81
8.4. Various measuring instruments utilized in marine research institutes.....	81
8.5. Operation examples of the camera installed at the ship’ s bow.....	82
8.6. A Camera installed at a real ship.....	83
8.7. Ship route for achieving images.....	83
8.8. Sample images achieved from ship-mounted camera.....	83
8.9. Prediction result on images from ship-mounted camera.....	84

8.10. Sample images per each classified category..... 85

Chapter 1. Introduction

Ocean waves periodically generate excessive kinetic energy and are directly exposed to ships or other man-made structures. Forecasting and measuring ocean conditions remain an intensely critical issue throughout human history because the ocean environment has a significant influence on aquaculture, marine transportation, construction of coastal structures, and offshore structures. More precise estimations of the ocean environment allow a safer and more reliable marine engineering for humans. Major maritime accidents that led to enactment and revision to the Safety Of Life At Sea(SOLAS) is shown in **Fig. 1.1**.

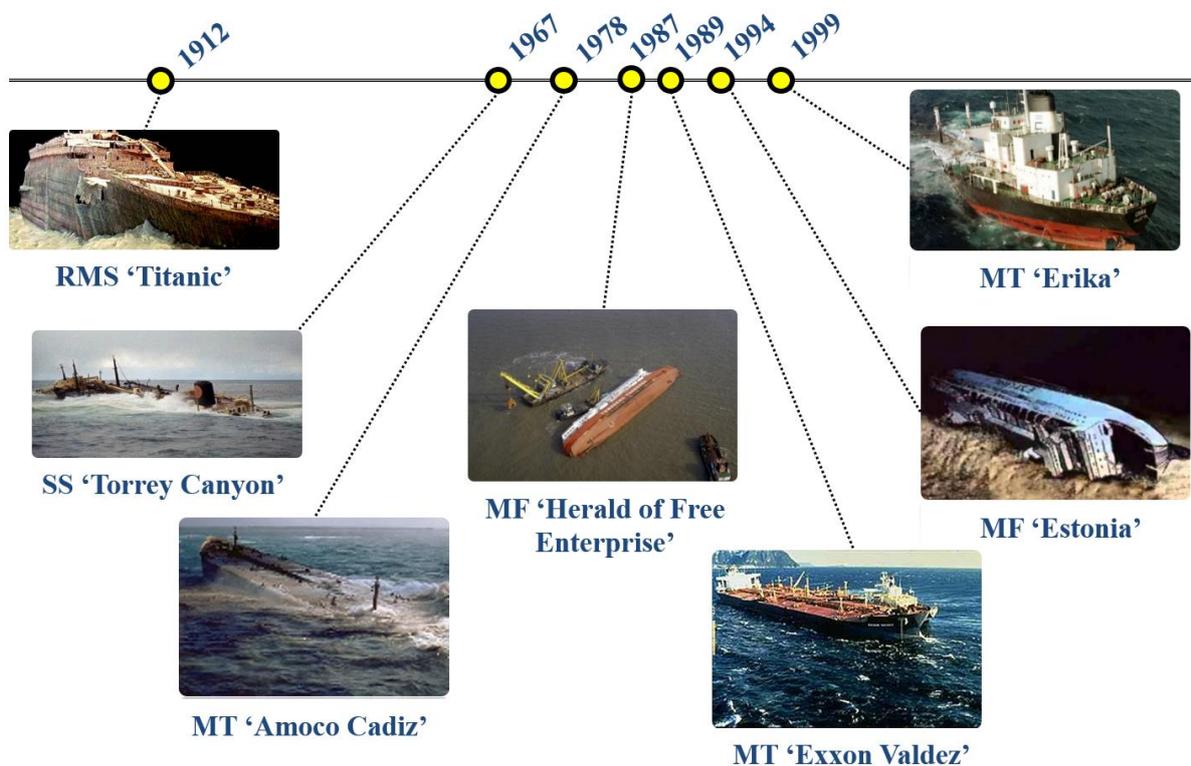


Fig. 1.1. Major maritime accidents related to the enactment and revision of SOLAS.

Marine accident status by type of accident and ship usage for five years(2016-2020) in Korea is listed in **Fig. 1.2** and **Fig. 1.3**. Although they are arranged in terms of simple quantity, a large number of accidents are concentrated on fishing boats. Also major accidents(crash, capsized, sinkage, and fire explosion) that lead to ship sinking account for a large number of accidents. Even, technologies for accurately measuring and monitoring maritime information have been developed continuously, there are lots of accidents nowadays and most accidents are occurred in small ships. Small ships do not have expensive instruments or monitoring systems and rely mostly on the captain's experience for sailing.

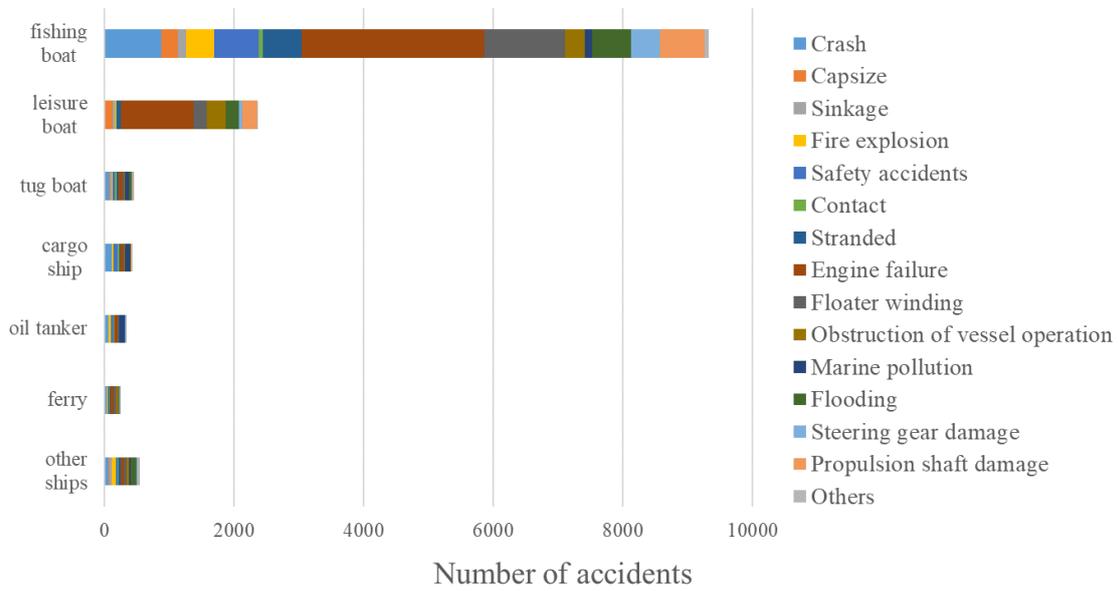


Fig. 1.2. Marine accident status by the type of accident in Korea (2016-2020).

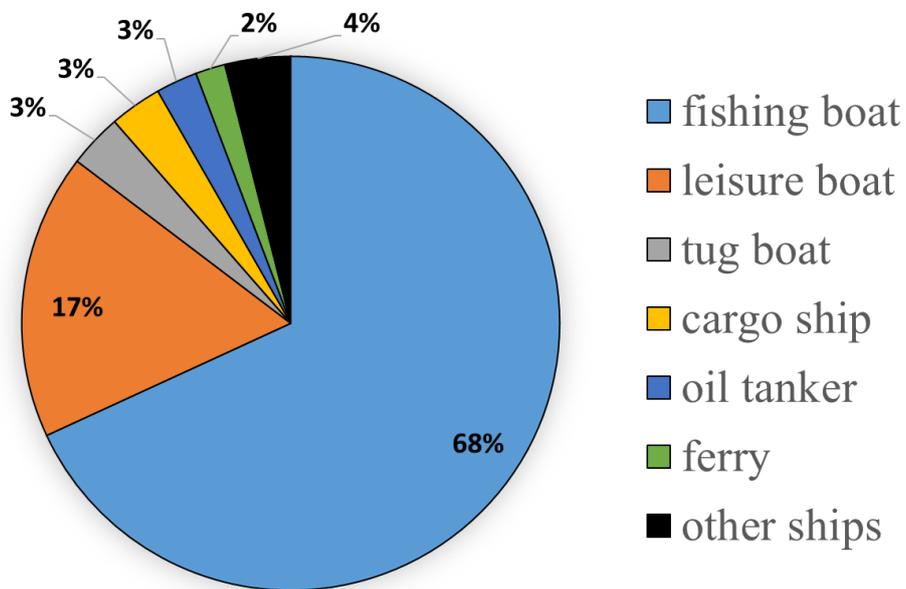


Fig. 1.3. Marine accident status by ship usage in Korea (2016-2020).

Various wave height gauges, which can be categorized into pressure, ultrasonic, microwave, and probe types, have been steadily developed over the years. Furthermore, advanced instruments such as the acoustic Doppler current profiler (ADCP), cross-correlation velocity profiler, and coherent Doppler velocity profiler have been installed near coastal lines worldwide [1]. The observations were carried out using moored buoys, automatic weather stations on lighthouse, research vessels, and base stations of oceanic-meteorological observations [2, 3]. An image post-processing technique with multiple vision cameras for wave height measuring has been applied in a few studies [4–7]. Although measuring instruments and corresponding technologies have been steadily developed over the years, the accuracy and consistency of the measured data still need to be improved. The installation cost of oceanic instruments is relatively high, and the maintenance of these devices is considerably difficult in harsh ocean environments.

Meanwhile, since the pioneering studies by McCulloch and Pitts [8] and Rosenblatt [9], artificial neural networks (ANNs) have shown remarkable improvements, some historical events are shown in Fig. 1.4. Considerable efforts on neural networks have led ANNs into the deep learning stage [10, 11]. The max pooling algorithm was designed in 2011, and a convolution neural network (CNN), comprising the activation functions of ReLU, overlapping pooling, local response normalization, dropout, and data augmentation, was introduced in 2012 [12]. Artificial intelligence was achieved worldwide recognition in 2016 through the Google Deepmind Challenge match [13, 14]. Artificial intelligence first emerged during the 1940s, followed by the machine learning era in the 1980s, with neural networks moving into the deep learning phase after 2011. LeCun et al. [15] comprehensively explained important developments in this field. The exclusive summary on artificial intelligence development is listed in Table 1.1.

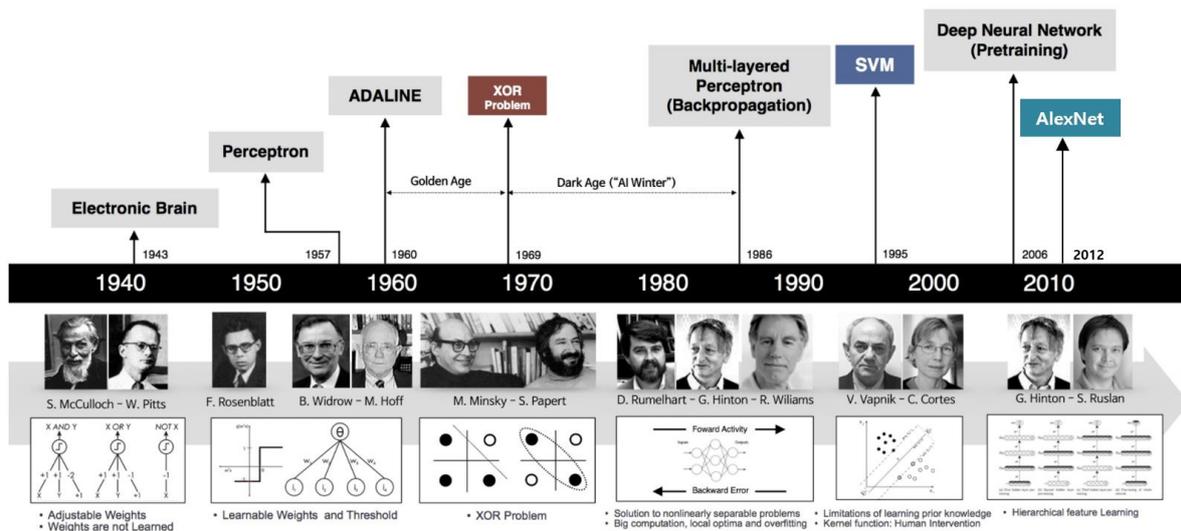


Fig. 1.4. Historical achievements in artificial intelligence [43].

Applications of artificial intelligence in fluid dynamics, coastal engineering, ocean engineering, and oceanography have been well summarized in several articles [16–19]. Performance predictions and fault detections on ships or offshore structures have also been conducted using several machine learning schemes [20–22]. In particular, there have been some attempts to predict the wave characteristics using artificial intelligence, including neural networks and support vector machines (SVMs) [23–30]. In measurement of sea surface height by a satellite radar altimeter, machine learning techniques were utilized in reducing the sea state bias [31–32]. This approach allows for more accurate estimation of sea state bias with significant wave height and wind speed. The CNN algorithm has recently been applied to wave estimation. Liu et al. [33] predicted the wave heights and periods in two-dimensional wave flumes. Three different sea conditions (spilling breaking, plunging breaking, and unbroken) were classified and predicted by Buscombe et al. [34].

Table 1.1. Exclusive summary on artificial intelligence development.

Period	Explanation
1940s	The beginning of neural networks (electronic brain)
1950s and 1960s	The first golden age of neural network (perceptron)
1970s	The winter of neural networks (XOR problem)
1980s	Renewed enthusiasm (multi-layered perceptron, backpropagation)
1990s	Subfield of radial basis function networks was developed
2000s	The power of neural networks ensembles & support vector machine is apparent
2006	Hinton presents the Deep Belief Network(DBN)
2009	Deep Recurrent Neural Network
2010	Convolutional Deep Belief Network (CDBN)
2011	Max-pooling CDBN
2012	ILSVRC(ImageNet Large Scale Visual Recognition Challenge) winner using CNN
2016	AlphaGO issue

In 2021, the most recent, many predictive studies on the marine environment using machine learning were performed. Ocean data classification was conducted in US using unsupervised machine learning for planning hybrid wave-wind offshore energy devices [35]. A study for transformation to nearshore wave from global wave data using the combined model, ANN and Group method of data handling(GMDH), was conducted at a certain port of Japan [29]. Coastal zone significant wave height prediction by supervised machine learning of ANN and decision tree model was also conducted [30]. A remarkable research on the nowcasting using generative models of radar was conducted in UK [36]. The deep generative model for the probabilistic nowcasting of precipitation from radar was developed, which is shown in **Fig. 1.5**. The wave field is around the UK, 1536km X 1280km, and

the forecasting was conducted for further 5-90 minutes. The prediction result showed the improvement forecast quality compared to the traditional numerical simulation.

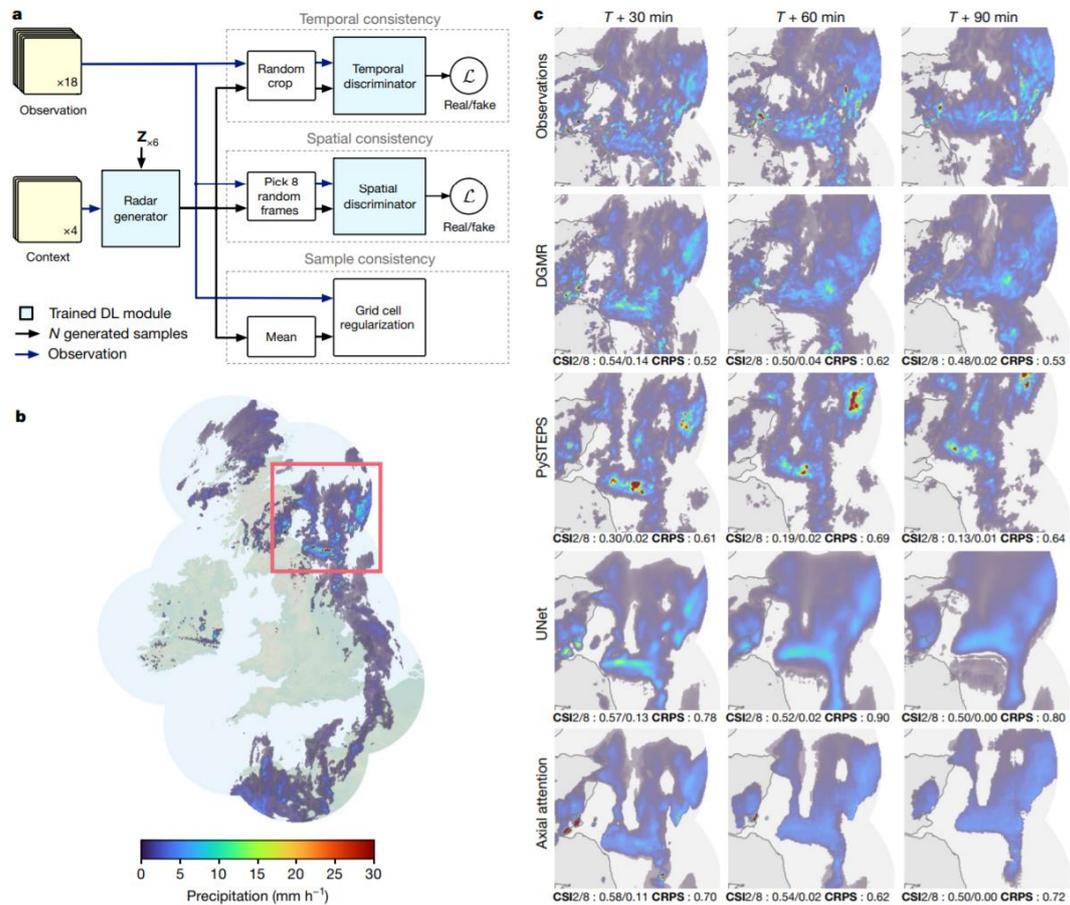


Fig. 1.5. Deep generative model for probabilistic now-casting of precipitation from radar in UK [36].

An auxiliary marine environment information provision service is essential for avoiding dangerous areas and calculating the optimal route for a vessel to navigate. Large merchant ships mostly receive ocean environmental information from satellites or install their own measuring instruments on ships. However, small ships sailing on the coast do not use these expensive equipment or systems and rely on the experience of the captain to the marine environment. It is difficult to systematically manage the operation of small ships, and there is a big difference in experience among captains. For this reason, it is necessary to develop a simple and effective system that can be used universally for smaller coastal ships.

In this study, we have developed an artificial intelligence-based wave field estimating technique to simply monitor the marine environment without applying complicated methodologies or introducing expensive equipment. We utilized deep learning simulations for marine environment prediction. In this study, Convolutional Neural Network(CNN) and Long Short-Term Memory(LSTM) were adopted as machine learning algorithm. Once the specified big image data have been trained, a single snapshot can be used to identify the sea conditions. We initially tested and modified our model with numerically generated snapshots, by computational fluid dynamics and Airy wave super-positions. We then achieved trainees as movie clip snapshots from the southwestern coastal region of Korea, which were then categorized based on the sea state and average wave height. The training performance was measured using the training accuracy.

The remainder of this paper is organized as follows. **Chap. 2** briefly describes the conventional sea-state measurement systems. **Chap. 3** presents the deep learning model based on convolution neural network and long short-term memory. Wave height estimations with numerical irregular waves are written in **Chap. 4**. The deep learning results with numerical waves are described in **Chap. 5**. Real ocean environment snapshots without any pre-processing on data are presented in **Chap. 6**. Trainings with prescreened data in terms of images and videos are summarized in **Chap. 7**. Possible applications of this methodology in marine engineering are written in **Chap. 8**. The conclusions are presented in **Chap. 9**.

Chapter 2. Conventional methodologies for sea-state measurement

The ocean environment is characterized by several properties such as salinity, temperature, wind speed and direction, wave conditions, and so on. While other quantities mostly affect ships and offshore structures in terms of structural reliability, the instantaneous dynamic behavior of ocean floaters is mainly determined by ocean environmental loads. Above all, water waves give a significant amount of inertia periodically to ships and offshore structures, and their station-keeping and maneuvering abilities are strongly determined by the wave conditions. Because the wave energy is roughly proportional to the square of the wave height, a precise measurement of the wave height has been continuously emphasized.

Conventional methodologies for sea-state measurement are using wave measurement devices such as pressure-, ultrasonic-, microwave-, and probe-type gauges mounted on buoys or the offshore platform on the sea. The wave measurement device measures wave elevations especially, and absolute surface elevation. Recent trend is that more fundamental aspects of the wave field are need to be measured and collected to know sea-state conditions. It is essential to measure multiple wave and surface current measurements based on the advanced system for wave field measurement. As following this requirement, there are two main methods for wave field measurement: 1) radar- and 2) image-based methods.

Radar-based wave field measurement can measure wide range of wave fields and not contact with the water, and is easy to maintain and deploy. Specifically, the WaveRadar REX [37] is widely used by the offshore oil and gas industry. The WaveRadar REX based on the frequency-modulated continuous wave (FMCW) method has continuous sampling rate at frequencies up to 10 Hz, providing excellent resolution for the wave processes. The WaveRadar consists of a transmitter head housing, the electronics and a parabolic dish antenna as shown in **Fig. 2.1**. The received signal reflected from the water surface is mixed with the signal that is being transmitted at that moment, and the result to represent wave fields with wave height, direction, and period on the screen. The state of the art of ocean environment measuring instruments using X-band radar is written in **Table 2.1**.

Image-based measurement non-intrusively measures wave fields for the temporal evolution of wave characteristics with wave height, period, and direction along the three dimensions. The automated trinocular stereo imaging system (ATSIS) [7] which is one of the spatial wave measurement system is available for measuring dynamical, spatial wave measurement. The ATSIS system uses digital video cameras to measure temporal evolution of three-dimensional wave characteristics as shown in **Fig. 2.2**.

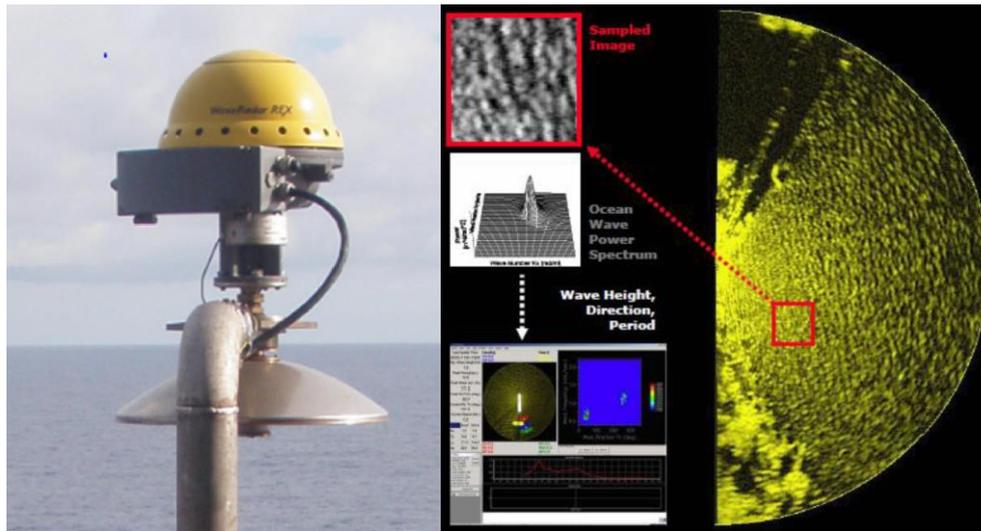


Fig. 2.1. WaveRadar REX installation setup(left) & wave field using X-band radar(right) [37].

Table 2.1. Development in ocean environment measuring instruments using X-band radar.

Item (X-band radar)	Unit	Country	Institute	Technical characteristics			
				Range	Resolution	STD	
Wave	height	Hs [m]	Norway, Denmark	MIROS OceanWaves	0.5 ~ 20	0.1	10%
	period	Tp [s]	Norway, Denmark	MIROS OceanWaves	3.0 ~ 20	0.1	5%
	direction	[deg]	Norway, Denmark	MIROS OceanWaves	0 ~360	1	10%
Wind	direction	[deg]	Denmark	GKSS	Corr. = 0.99		14.24°
	speed	[m/s]	Denmark	GKSS	Corr. = 0.97		0.85m/s
Current	direction	[deg]	Norway	MIROS	1 ~ 360	1	7%
	speed	[m/s]	Norway	MIROS	0.0 ~ 2.5	0.01	0.05m/s
Bilateral flow	number	US, Japan	Oregon univ. Tsukuba univ.				in developing

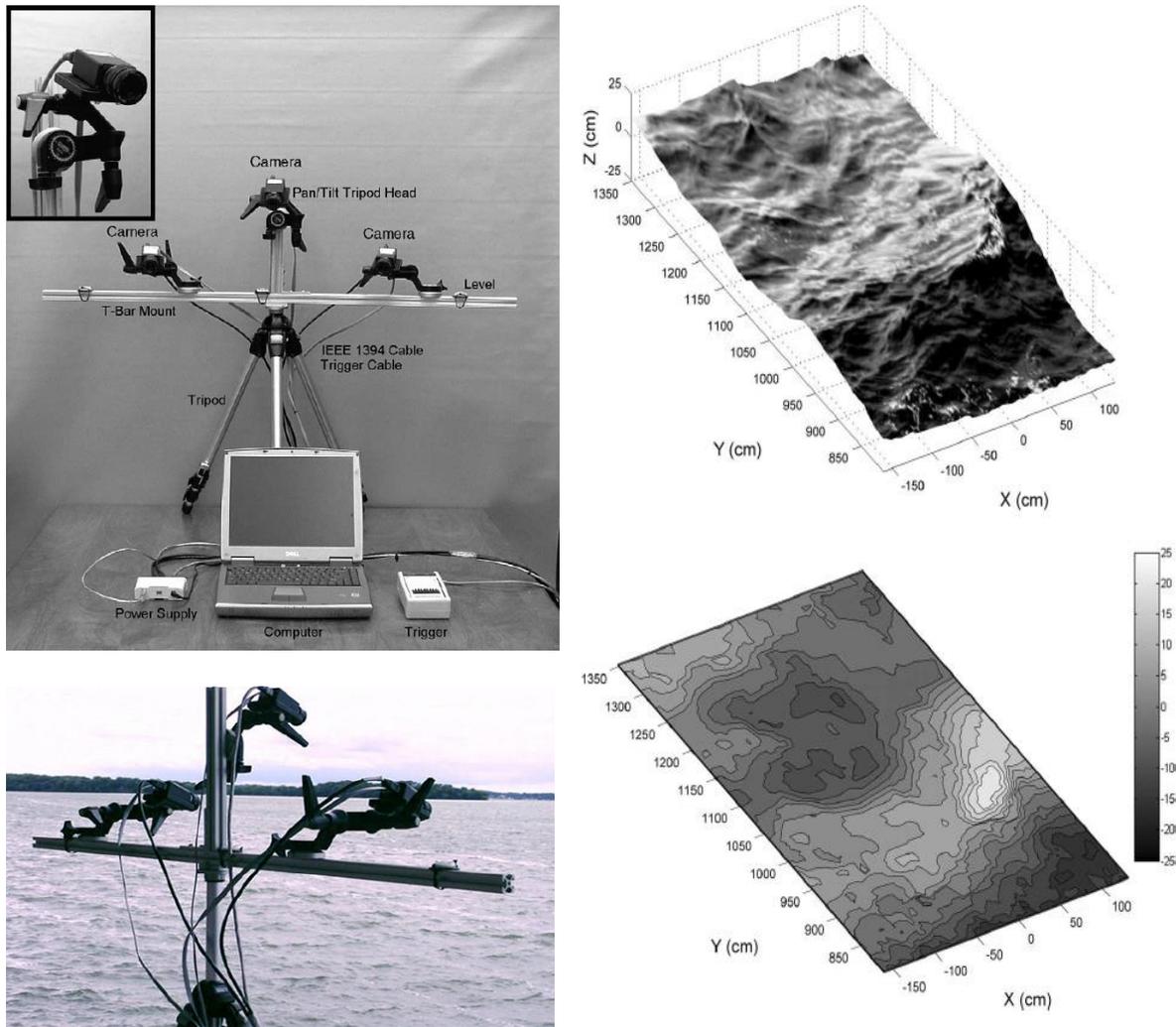


Fig. 2.2. The automated trinocular stereo imaging system (ATSIS) (left) and results (right) [7].

In South Korea, there are numerous marine meteorological observations as 46 tidal observatories, 3 ocean observatories, 39 ocean observation buoys, 10 current observations, and 3 ocean science research stations, shown in **Fig. 2.3** [3]. They measure principal ocean environmental quantities such as tidal level, water temperature, salinity, wave height, ambient temperature, atmosphere pressure, wind speed and direction. While the instruments and corresponding technologies have been developed steadily, the accuracy and consistency of measured data need to be improved. Also the installation cost of oceanic instruments is relatively expensive and the maintenance of these devices is quite difficult with harsh ocean environment.

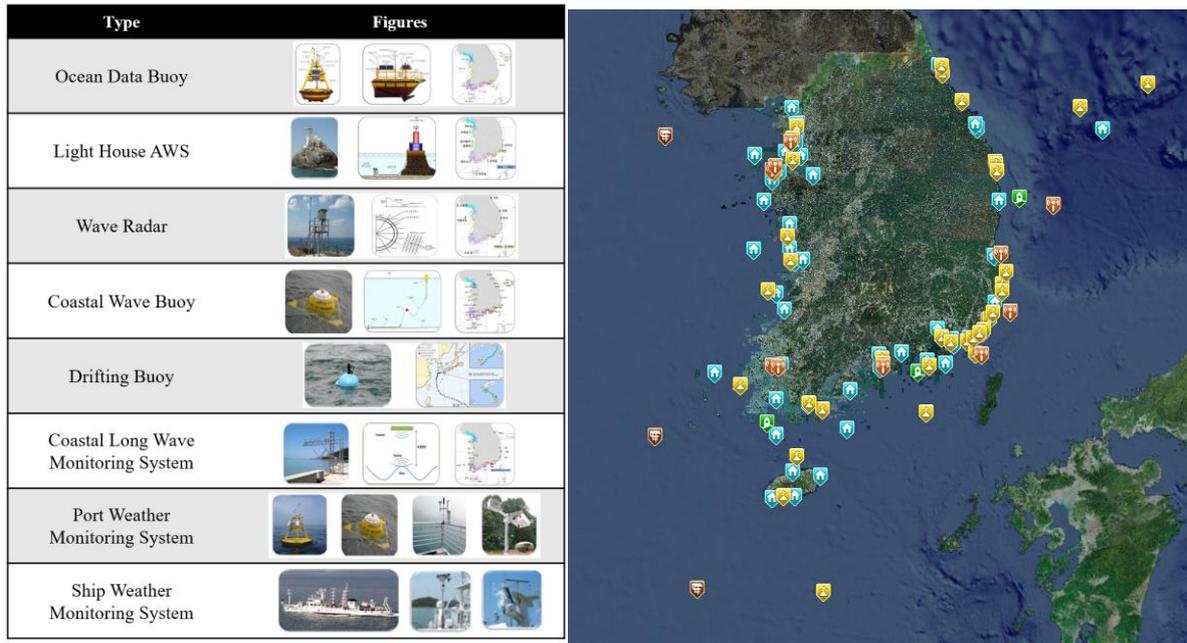


Fig. 2.3. Types of marine meteorological observations(left) and their installation locations in Korea(right) [3].

Currently, marine measuring equipment requires professional experience and high cost in installation and operation. In the case of small ships, even that is properly received and mostly depends on the experience of the captain for judging the ocean environments. As a result, capsizing accidents of small ships on the coast continue to occur, and among them, some accidents due to sudden changes in waves continuously occur. The inadequate marine environment evaluation system and corresponding accident cases in small ships are illustrated in **Fig. 2.4.**

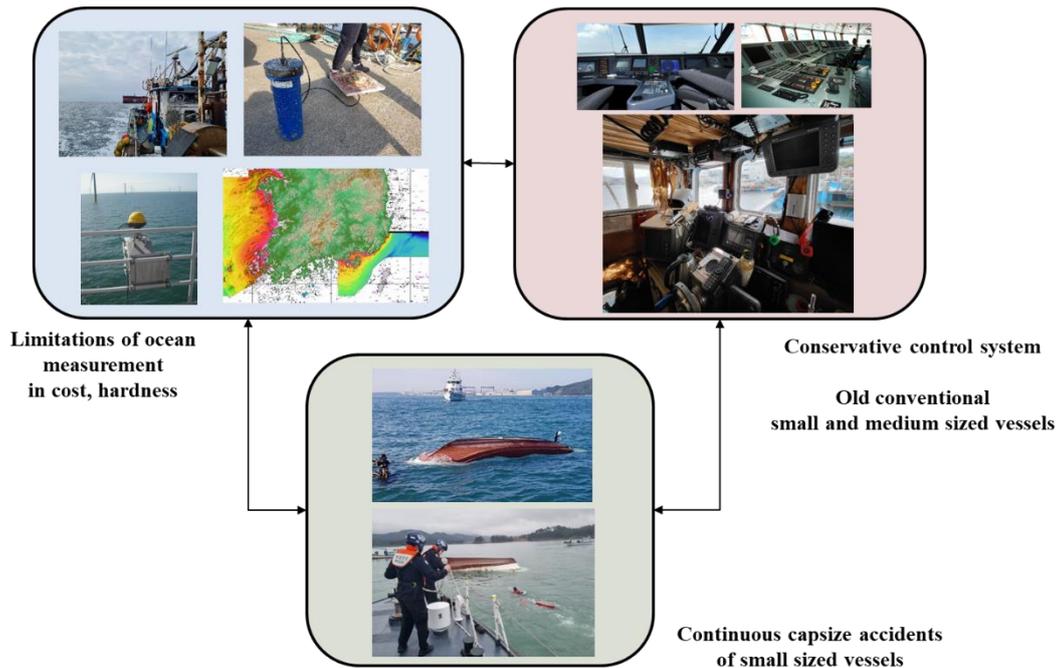


Fig. 2.4. Small ship accident induced by insufficient and inadequate marine measuring system.

Therefore, in this study, a practical now-casting system for ocean environment was developed that is cheaper and easier to operate than the current measurement forecasting system. Different from the classical approaches of meteorology, metrology, and oceanography, the data science was applied. This was set in accordance with the development of new technologies to overcome the shortcomings of existing data science and increase the speed of high-speed calculation for big data with the rapid development of computing hardware. In summary, the considered system was aimed to be developed of a low-cost but generally reliable marine environment estimation system by applying artificial intelligence, which has recently developed rapidly among data science.

Chapter 3. Machine learning methodologies

The comprehensive explanation on utilized machine learning methodologies are written in this chapter. First, the applied convolutional neural network in snapshot-based learning is explained. The utilized long short-term memory is then explained in sequential inputs. The video-based learning is conducted with the combined model with convolutional neural network and long short-term memory, which is lastly written in this chapter.

3.1 Convolutional neural network

A convolutional neural network (CNN) is a deep neural network, most commonly used in the application of computer vision and image processing. It is a powerful artificial intelligence tool used for image classification. In the case of an image, three-dimensional data consisting of vertical, horizontal, and channel (color) should be normalized to one-dimensional data used in the network. Convolutional neural networks use convolution layers and pooling layers, unlike the existing fully connected neural networks. This prevents the loss of spatial information in the process of flattening the photographic data and enables effective extraction and learning of features. In the convolution layer, weighting per pixel is applied using the kernel to extract or highlight the domain features. In the pooling layer, after receiving data from the convolution layer, it is used to reduce the size and emphasize specific data at the same time. In particular, in the case of Convolution Neural Networks, since filters are used as shared parameters, very few learning parameters are utilized compared to general artificial intelligence neural networks.

In the development history of a CNN, the invention of AlexNet, which is shown in **Fig. 3.1**, was a revolution in terms of advancements [12]. Several CNN architectures have been launched since then, such as VGG [38], ResNet [39], GoogLeNet [40], InceptionV3 [41], Xception [42], Inception-ResnetV2 [43], and MobilenetV2 [44]. The differences among these networks are the adoption of depth-wise convolutional layers, the applicable range of filters, and residual connections. In this study, the initial CNN algorithm, identical to that used in AlexNet, was applied in all training cases, as shown in the schematics of the concept in **Fig. 3.2**, for sea state estimations using snapshot images.

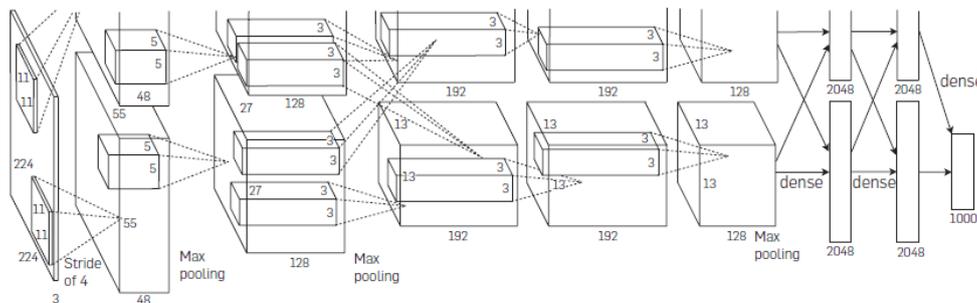


Fig. 3.1. Illustration of the AlexNet [12].

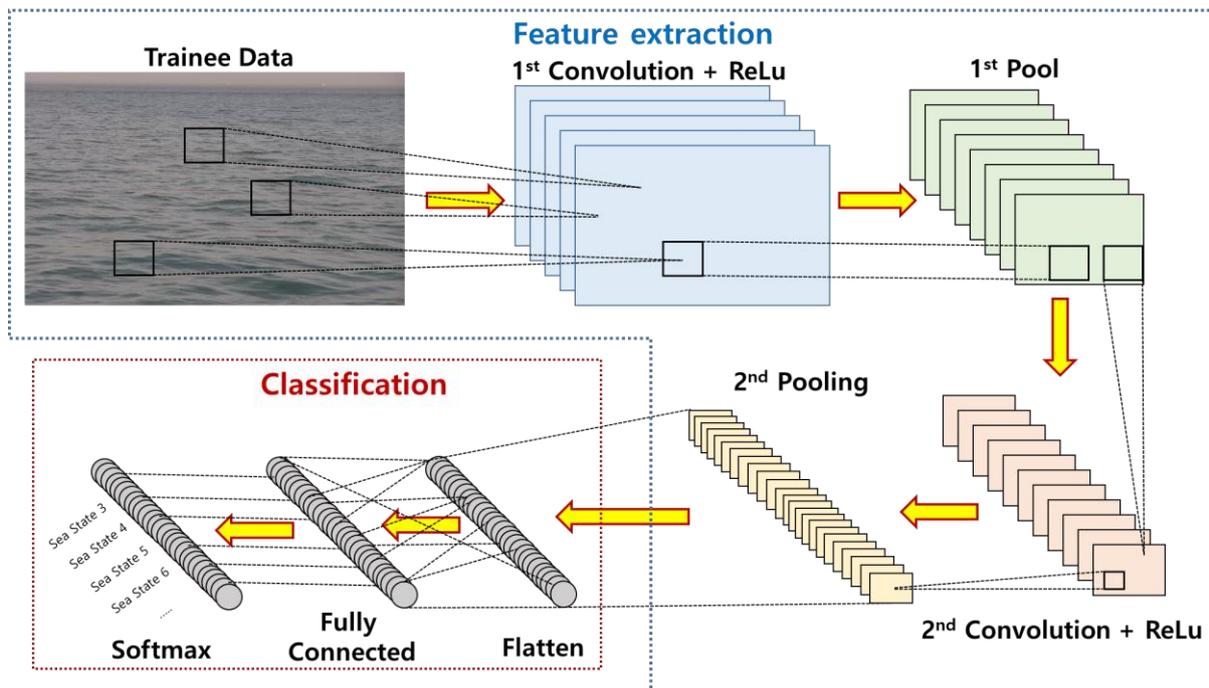


Fig. 3.2. Schematics of the convolution neural network.

The CNN creates a hierarchy of layers with various weights and biases applications, applying the category activation function, and the image feature extraction in each layer. This procedure allows a set of filters to extract local characteristics, and then sequentially transfer features to sub-divided smaller data set. The weights of filters are adjusted using back-propagation. After passing the designed convolution layers, a fully connected layer is constructed as a multiplication of the inputs by a weight matrix and a bias vector. The normalization on outputs from the fully connected layer is performed, the classifications with calculating the loss function is then done, and these results become the inputs of the next machine learning sequence.

We constructed the CNN structure for this research as **Fig. 3.3**. Four training layers are exist, and each layer has a 2D convolutional layer, a batch normalization layer, an activation function of reLU(rectified linear unit), and a 2D max pooling layer. The max pooling technique is eliminated in the last layer. The number of sliding convolutional filters in each 2D convolutional layer is sequentially 16, 32, 64, and 64, while the padding size of 3 by 3 and the padding option of same values along all edges of the layer input are identical in all layers. After the batch normalization layer for parallel learning, the rectified linear unit(reLU) is applied as an activation function to perform a nonlinear threshold operation. The max pooling size is 4 by 4 with 4 pixels moving interval for all layers, which does not allow the overlapped pooling. Adaptive moment estimation(ADAM) is utilized in updating parameters. The number of mini-batch size is same as the number of categories, which allows the number of adjusting training parameters are same as the number of trainees per category. The entire training data is utilized 8 times, which is defined as 8 epochs. All deep CNN learning simulations in this study have been conducted in

Deep Learning Toolbox of MATLAB [45], the specific explanations for training options can be found in [12]. Explanation on each layer and main applied training options are listed in **Table 3.1** and **3.2**, respectively.

All deep learning simulations were conducted with a personnel computer, which has twelve central processing units(CPUs) and one graphics processing unit(GPU) under ‘Window’ operation system, listed in **Table 3.3**. Even multi-GPUs computing is quite powerful in big data trainings, the single GPU, GeForce RTX 2080 SUPER, was enough to conduct several training cases in this research. All training simulation cases did not take more than a day, which are illustrated for each case in the next chapter.

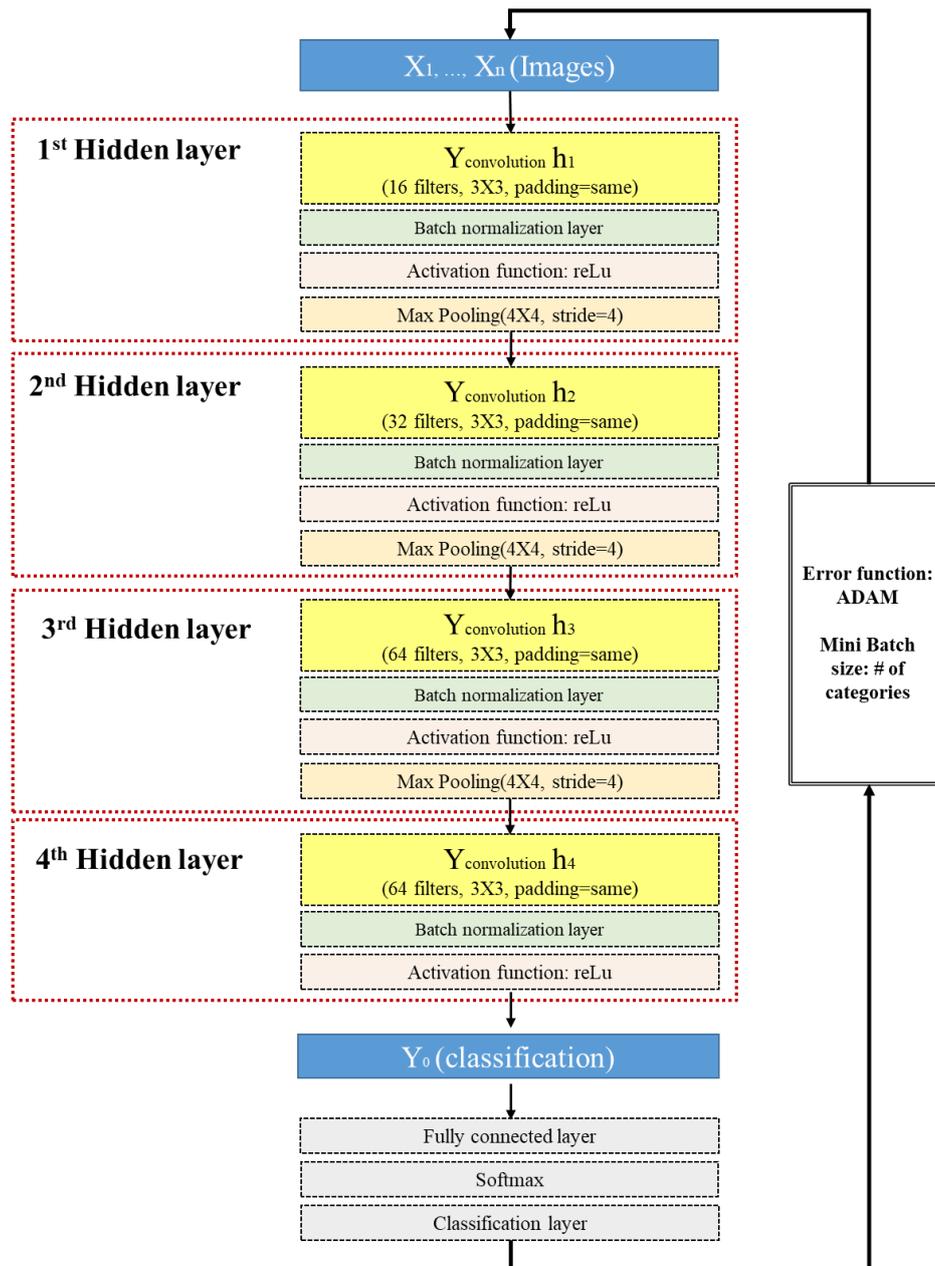


Fig. 3.3. Proposed architecture of the convolution neural network (CNN).

Table 3.1. Explanation on machine learning layers.

Layer	Explanation
Convolution	The layer convolves the input by moving filters along the input vertically and horizontally and computing the dot product of the weights and the input, and then adding a bias term.
Batch normalization	It normalizes each input channel across a mini-batch for speeding up training of convolutional neural networks and reducing the sensitivity to network initialization.
Activation function	It performs a threshold operation to each element of the input, where any value less than zero is set to zero.
Max pooling	It performs down-sampling by dividing the input into rectangular pooling regions, and computing the maximum of each region.
Fully connected	It multiplies the input by a weight matrix and then adds a bias vector.
Softmax	It applies a softmax function to the input. Softmax function converts a vector of numbers into a vector of possibilities.
Classification	It computes the cross entropy loss for multi-class classification problems with mutually exclusive classes. The layer infers the number of classes from the output size of the previous layer.
Error function	Adaptive moment estimation(ADAM) is used as the error function. An element-wise moving average of both the parameter gradients and their squared values.

The corresponding functions in activation layer, softmax layer, and adaptive moment estimation(ADAM) error calculating layer are written in Eq. 3.1, 3.2, and 3.3, respectively.

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (3.1)$$

$$S(y_i) = \frac{e^{y_i}}{\sum_{j=1}^i e^{y_j}} \quad (3.2)$$

$$\begin{aligned} m_i &= \beta_1 m_{i-1} + (1 - \beta_1) \nabla E(\theta_i) \\ v_i &= \beta_2 v_{i-1} + (1 - \beta_2) [\nabla E(\theta_i)]^2 \end{aligned} \quad (3.3)$$

$$\theta_{i+1} = \theta_i - \frac{\alpha m_i}{\sqrt{v_i} + \varepsilon}$$

Table 3.2. Training options of CNN.

Item	Option
Learning method	CNN(Convolution Neural Network)
Number of convolution layer	4
Solver	ADAM(Adaptive Moment Estimation)
Activation function	ReLU
Number of filters per convolution layer	16 / 32 / 64 / 64
Max pooling	1 st , 2 nd , 3 rd : 4X4 with 4 stride
Gradient decay factor	0.9
Squared gradient decay factor	0.999
Gradient threshold method	L ₂ norm
Epoch	8 for initial sea state training 2 for following trainings
Mini-batch size	Same as number of categories

Table 3.3. Computing hardware specifications.

Hardware	Specification
Operating System(OS)	Windows 10 Enterprise
Processor	AMD Ryzen 9 3900X 12-Core Processor/3.80GHz
RAM [GB]	63.9
System Type	64 bit OS / x64 based processor
Graphics Processing Unit(GPU)	GeForce RTX 2080 SUPER

3.2 Long short-term memory

Long Short-Term Memory(LSTM) is one of specified recurrent neural network(RNN) architecture which has the ability to perform learning that requires a long dependency period. LSTM was introduced by Hochreiter & Schmidhuber [46], and has since continued to develop and become famous through following studies. LSTM has solved lots of problems well, and is still widely used nowadays.

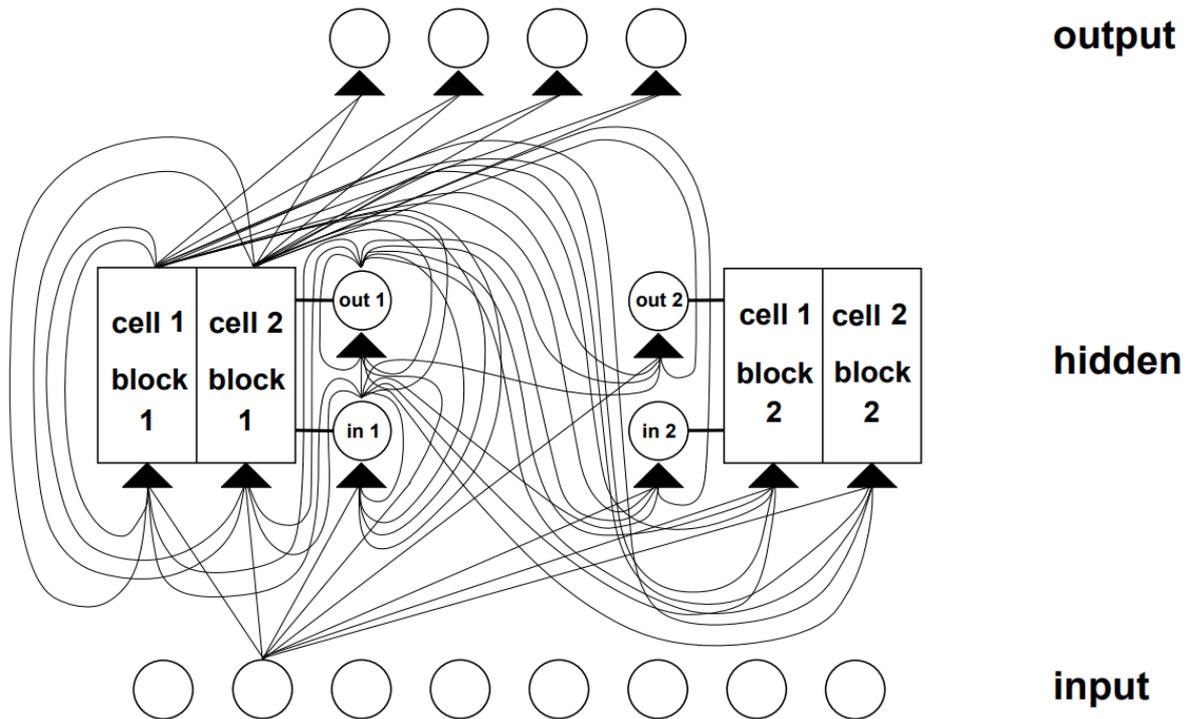


Fig. 3.4. Example of a LSTM net [46].

The basic concept of LSTM is to explicitly design to avoid the problem of long dependency periods. Remembering information for long periods of time should be a basic behavior of the model, so that the model does not struggle to learn it. All Recurrent Neural Network(RNN)s are chains of iterative neural network modules, which is shown in **Fig. 3.5**. Basically, this repeating module has a quite simple structure in RNN, such as the tanh layer.

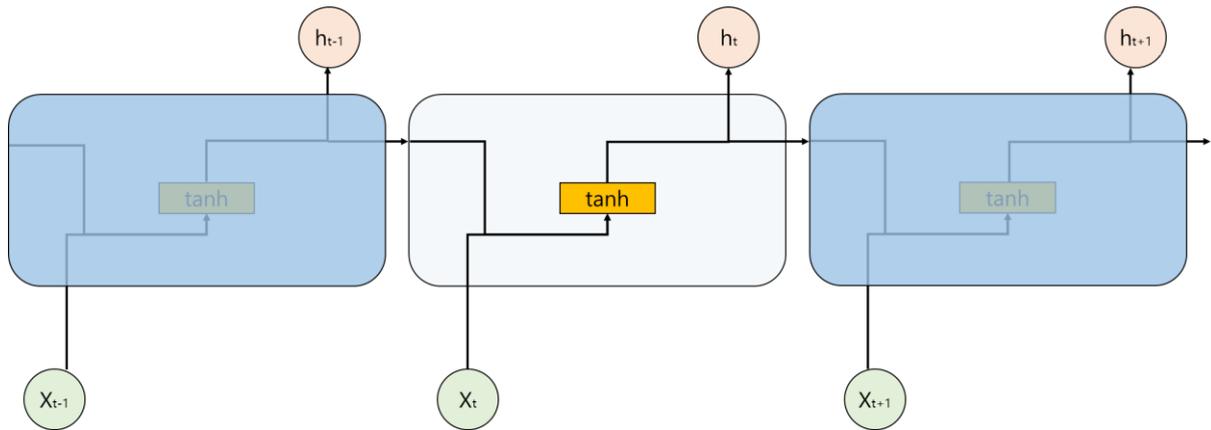


Fig. 3.5. Repetitive modules of RNN with one layer.

LSTM has the same chain-like structure, while each repeating module has a different structure. Instead of a simple neural network layer, four layers are designed to communicate with each other in a special way. The fundamental architecture of one repeating module of LSTM is shown in **Fig. 3.6**. In this figure, circles mean ‘pointwise operation’, rectangles indicate ‘neural network layer’, and arrows show the information translations.

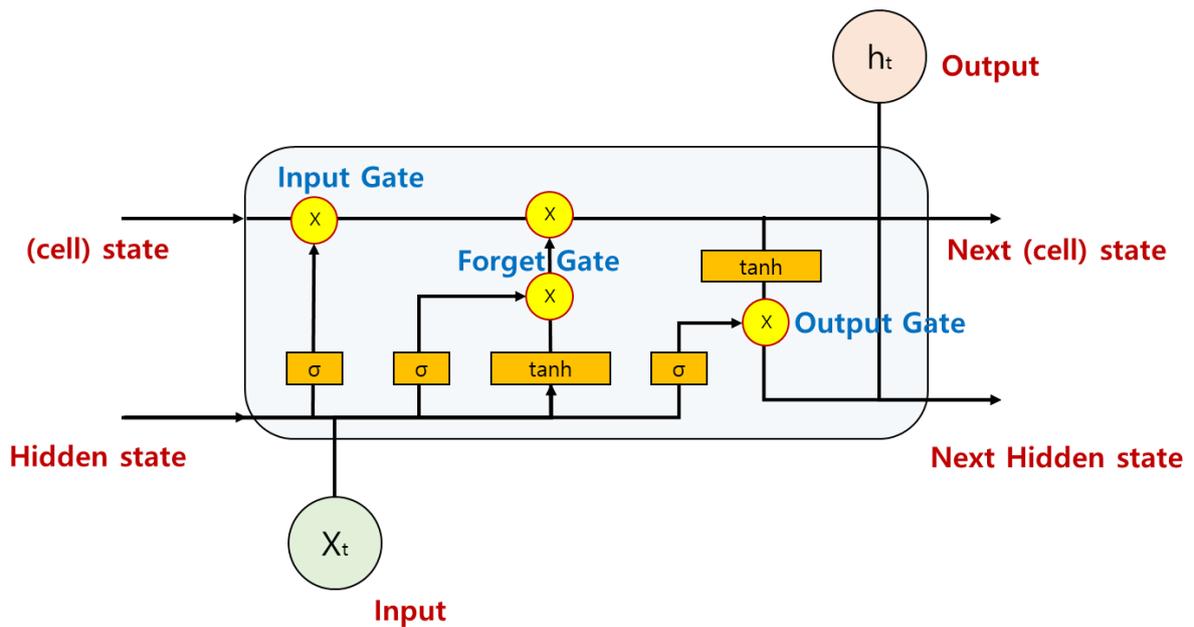


Fig. 3.6. LSTM with four interactive layers.

The core of LSTM is the ‘cell state’, which corresponds to the upper line drawn horizontally in the figure. The cell state acts as a conveyor belt, which continues to drive the entire chain while applying only small linear interactions. It is quite easy to just pass the information flow without interruptions. LSTM has the ability to add or remove some features from the cell state, which is carefully controlled by a structure called a ‘gate’. Gates are additional ways in which information can be conveyed, consisting of a sigmoid layer or a pointwise multiplication.

LSTM has three gates, and these gates protect and control the cell state.

The first step in LSTM is to decide which information to discard from the cell state, which is determined by the sigmoid layer. So, the gate at this stage is called 'forget gate layer'. In this stage, the values of h_{t-1} and x_t are taken and sent to C_{t-1} as values between 0 and 1. If the value is 1, it becomes "Keep all information", and if the value is 0, it becomes "Take it all away". The next step is to decide which of the new information to be stored in the cell state. In this stage, the sigmoid layer called 'input gate layer' initially determines which values to be updated. Then, the tanh layer generates a vector, C_t , which is a new candidate, and prepares to add it to the cell state. By combining the information from these two steps, we create a material that will update the state.

Thirdly, we will update the old state, C_{t-1} , to create a new cell state, C_t . Since we have already decided which values to be updated and corresponding weights, all we need to do here is to practice them. In this stage, we multiply the previous state by F_t so that we really forget what we decided to forget in the initial first step. Then add $i_t \cdot C_t$. This added value becomes a scaled value by how much to update the value set to be updated in the next step. Finally, it remains to decide what to export as output. This output will be a filtered value based on the cell state. In this stage, it firstly decides which part of the cell state to output by burning the input data in the sigmoid layer. Then the cell state is burned into the tanh layer, and after receiving a value between -1.0 and 1.0, it is multiplied by the output of the sigmoid gate calculated earlier. That way, we can export only the part we want to send as output.

Currently, various modified models of LSTM have been developed. The suggested model by Gers & Schmidhuber [47] has added a 'peephole connection', which is named as it makes the gate layers look at the cell state. As a slightly more modified LSTM, the Gated Recurrent Unit (GRU) was introduced in Cho et al. [48]. This model combines the forget gate and the input gate into one 'update gate', the cell state and the hidden state, and there are several other changes. Consequently, the GRU has a simpler structure than the existing LSTM and is becoming more popular. There is also a depth-gated RNN introduced by Yao et al. [49]. There are also models that resolve long dependency periods in a way that differs from LSTMs, such as Koutnik et al. [50].

3.3 Neural network architecture for movie-clip based machine learning

In this study, machine learning was also performed using movie clips. CNN is used to extract image features, and LSTM is used for classification learning on continuous sequence data. The procedure is shown in **Fig. 3.7**.

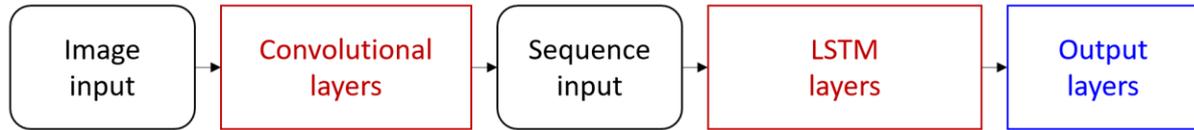


Fig. 3.7. Conceptual process for video-based deep learning.

First, transforming the video into a sequence of feature vectors using a pretrained convolutional neural network to extract features from each frame. We used the GoogLeNet [39] as a CNN algorithm. This algorithm won the 2014 ImageNet Large Scale Visual Recognition Challenge(ILSVRC). It has the 22-layers deep model and four features are important; 1X1 convolution for reducing feature map, Inception module for feature extraction, Global average pooling for flattening with lower computing compared to fully connected layer, and Auxiliary classifier for avoiding vanishing gradient. The network is illustrated in **Fig. 3.8** and is described in **Table 3.4**.

We use a convolutional neural network as a feature extractor by taking the activation values when inputting video frames into the neural network. In this stage, video clips are converted to sequences of feature vectors. Here, the feature vector is the output of the activations function applied to the last pooling layer(“pool5-7X7_s1”) of the GoogLeNet neural network. After terminating through the CNN, the data is arranged in a size of ‘1024Xn’.

Then, we create an LSTM neural network that can classify a sequence of feature vectors representing the video. The LSTM neural network architecture is defined with the following specified layers: A sequence input layer whose input size is equal to the feature dimension of the feature vector, Bi-LSTM layer with 2000 hidden units followed by a dropout(0.5 for this study), fully connected layer, softmax layer, and classification layer. The schematic architecture of LSTM in this study is shown in **Fig. 3.9**, and the training options in LSTM are described in **Table 3.5**.

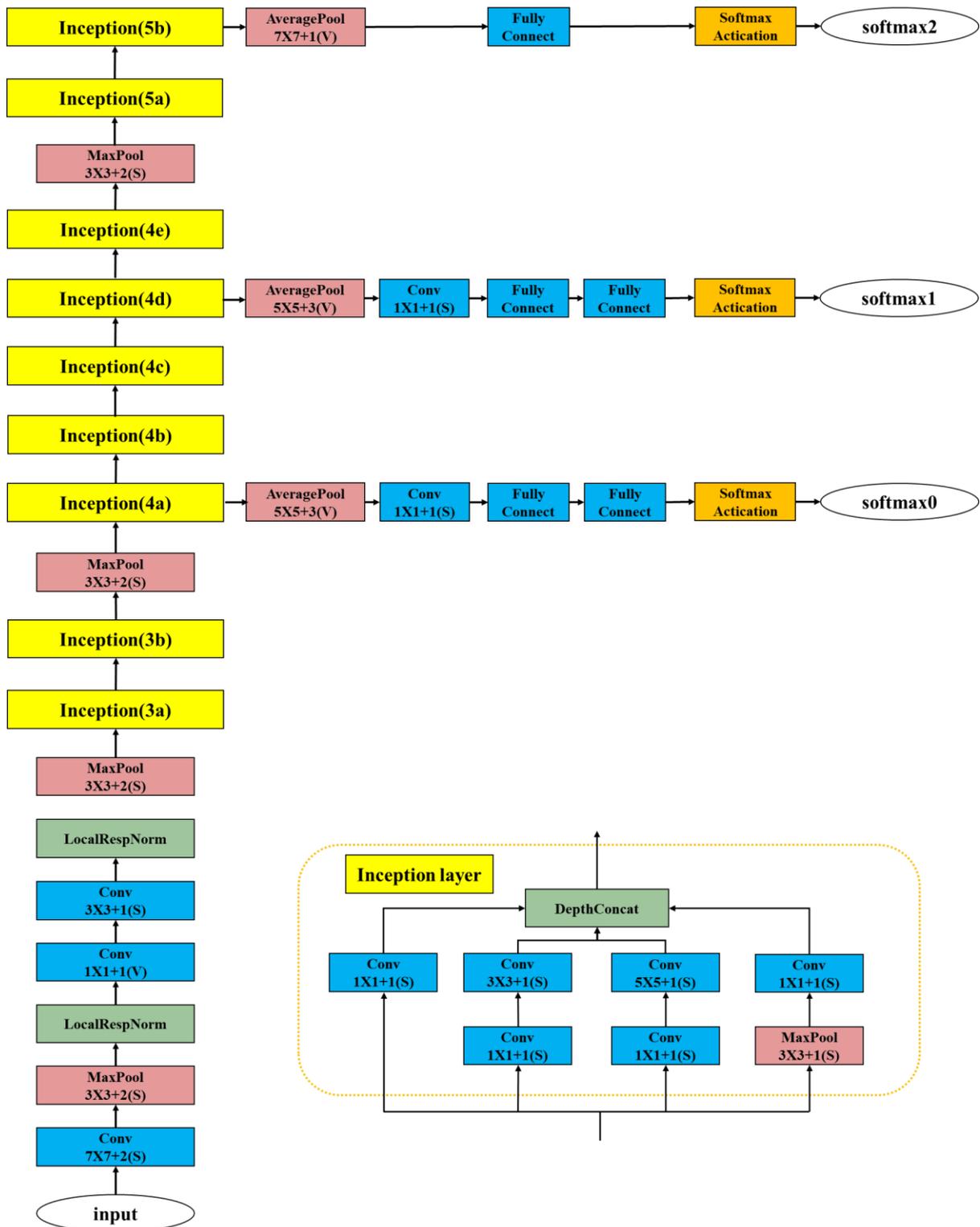


Fig. 3.8. GoogLeNet network [40].

Table 3.4. GoogLeNet incarnation of the Inception architecture [40].

type	patch size/stride	output size	depth	#1X1	#3X3 reduce	#3X3	#5X5 reduce	#5X5	pool proj	params	ops
convolution	7x7/2	112x112x64	1							2.7k	34m
max pool	3x3/2	56x56x64	0								
convolution	3x3/1	56x56x192	2		64	192				112k	360m
max pool	3x3/2	28x28x192	0								
inception(3a)		28x28x256	2	64	96	128	16	32	32	159k	128m
inception(3b)		28x28x480	2	128	128	192	32	96	64	380k	304m
max pool	3x3/2	14x14x480	0								
inception(4a)		14x14x512	2	192	96	208	16	48	64	364k	73m
inception(4b)		14x14x512	2	160	112	224	24	64	64	437k	88m
inception(4c)		14x14x512	2	128	128	256	24	64	64	463k	100m
inception(4d)		14x14x528	2	112	144	288	32	64	64	580k	119m
inception(4e)		14x14x832	2	256	160	320	32	128	128	840k	170m
max pool	3x3/2	7x7x832	0								
inception(5a)		7x7x832	2	256	160	320	32	128	128	1072k	54m
inception(5b)		7x7x1024	2	384	192	384	48	128	128	1388k	71m
avg pool	7x7/1	1x1x1024	0								
dropout(40%)		1x1x1024	0								
linear		1x1x1000	1							1000k	1m
softmax		1x1x1000	0								

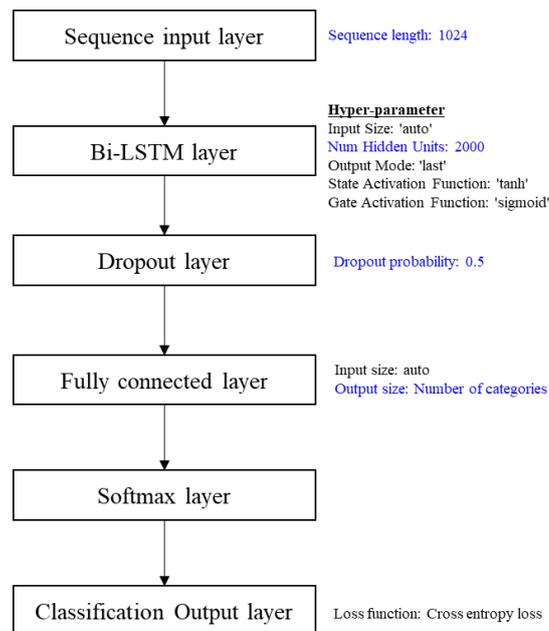


Fig. 3.9. LSTM network [46].

Table 3.5. Training options of LSTM.

Item	Option
Learning method	LSTM(Long Short-Term Memory)
Optimizer	ADAM(Adaptive Moment Estimation)
Gradient Decay Factor	0.9000
Squared Gradient Decay Factor	0.9990
Epsilon	1.0000e-08
Initial Learn Rate	1.0000e-04
Learn Rate Schedule	None
Learn Rate Drop Factor	0.1000
Learn Rate Drop Period	10
L ₂ Regularization	1.0000e-04
Gradient Threshold Method	L ₂ norm
Gradient Threshold	2
Max Epochs	60
Mini Batch Size	4
Verbose	1
Verbose Frequency	20
Validation Data	{{37X1 cell} [37X1 categorical]}
Validation Frequency	81
Validation Patience	Inf
Shuffle	'every-epoch'
Execution Environment	'auto'
Plots	'training-progress'
Sequence Length	'longest'
Sequence Padding Value	0
Sequence Padding Direction	'right'
Dispatch in Background	0
Reset Input Normalization	1

Chapter 4. Wave height estimation with numerical irregular waves

In this chapter, machine learnings on simultaneous wave height estimation were conducted. Learning data was generated numerically as Airy wave theory, in terms of long-crested waves and short-crested waves. The captured snapshots were then categorized with a certain wave height division. Several image filters were applied in long-crested wave data and the results were compared. Some attempts to increase the prediction performance were conducted for short-crested wave learnings.

4.1 Estimation on long-crested waves

To check the applicability of the machine learning into sea state classification, the initial learning was performed with numerical wave field data. We generated a long-crested wave field with Airy wave super-positions. The JONSWAP spectrum was chosen and the detailed explanation can be found in DNVGL [51], and the formulas are illustrated as follow.

$$S(\omega) = \frac{5}{16} (1 - 0.287 \ln(\gamma)) \cdot H_s^2 \cdot \omega_p^{-4} \cdot \omega^{-5} \cdot e^{\left[-1.25 \left(\frac{\omega}{\omega_p} \right)^4 \right]} \cdot \gamma^{\left[\frac{-(\omega - \omega_p)^2}{2\sigma^2 \omega_p^2} \right]} \quad (4.1)$$

$S(\omega)$ is the wave spectral density, γ is the non-dimensional peak parameter, H_s is the significant wave height, ω_p is the peak wave frequency, ω is the wave frequency, σ is the spectral width parameter. Based on the previous measuring data from the northern sea of Europe, the widely used value of σ is 0.07 for $\omega \leq \omega_p$ and 0.09 for $\omega > \omega_p$. The continuous wave spectrum can be discretized with the uniform interval, $\Delta\omega$. The corresponding wave amplitude($A(\omega_i)$) and phase($\varepsilon(\omega_i)$) for each frequency component can be written as **Eqs. 4.2** and **4.3**.

$$A(\omega_i) = \sqrt{2S(\omega_i)\Delta\omega} \quad (4.2)$$

$$\varepsilon(\omega_i) = \text{randomly distributed} \quad (4.3)$$

The time series of wave elevations can be achieved by the wave amplitudes and phases in frequency domain as **Eq. 4.4**. The discretized number of the wave spectrum is N .

$$\eta(t) = \sum_{i=1}^N A(\omega_i) \cos(\omega_i t + \varepsilon(\omega_i)) \quad (4.4)$$

For a long-crested wave field, a single wave train is used.

The sea state 4 condition was modelled, which has 2.4m of a significant wave height, 6.6s of a peak period, and 2.5 of a peak parameter. Four snapshots of the long-crested wave field are shown in **Fig. 4.1**. The y-axis labelling was eliminated to avoid unexpected interruptions for deep learnings. The size of images was a height of 656 pixels and a width of 875 pixels, which represents the 400m by 400m of the physical domain. The time series of the center and the classification in terms of wave heights are shown in **Fig. 4.2** and **Fig. 4.3**, respectively. The maximum wave elevation was 2.133m and the minimum one was -2.262m. The snapshots were then supervised in ten categories between -1.0m and +1.0m with 0.2m intervals, larger waves than +1.0m were labelled as '+1.5' and lower wave than -1.0m were labelled as '-1.5'.

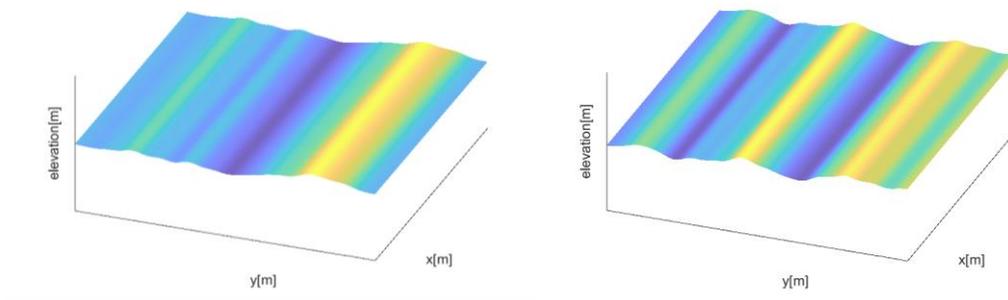


Fig. 4.1. Snapshots of long-crested waves.

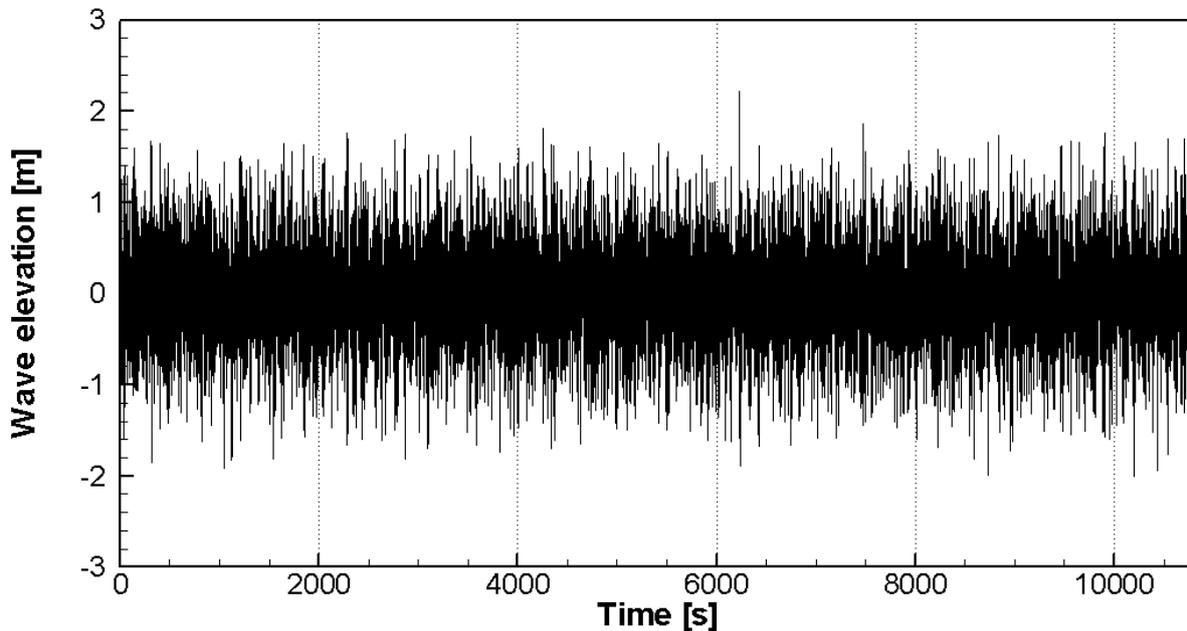


Fig. 4.2. Wave elevation at the center of the domain.

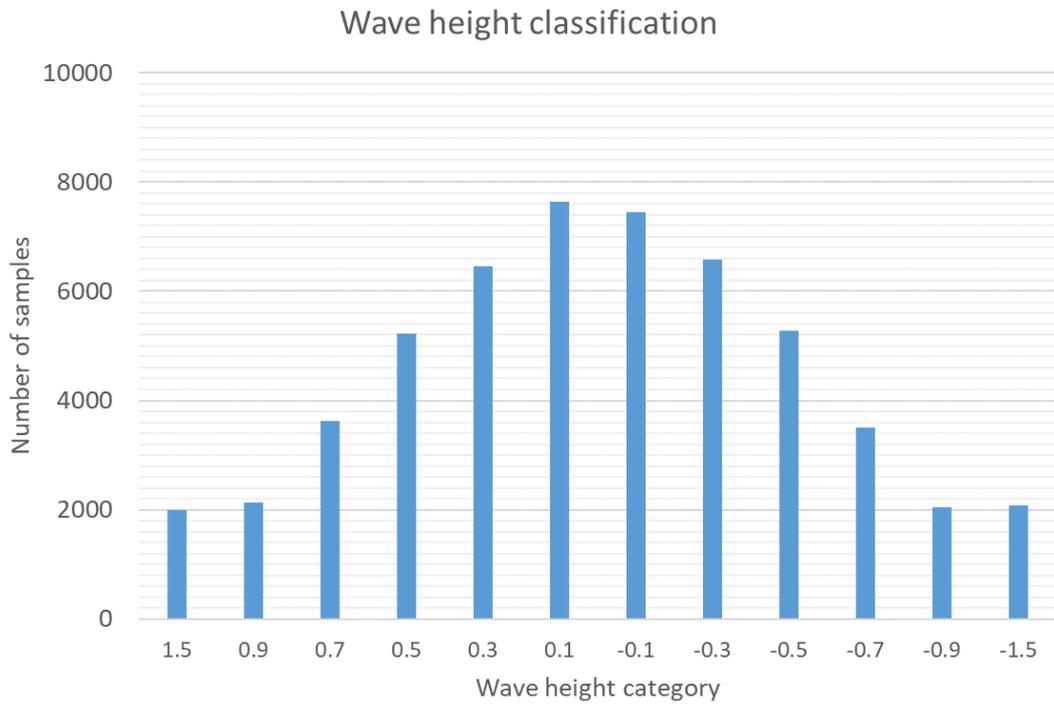


Fig. 4.3. Classification by wave height.

We conducted the wave height estimation with the prescribed deep learning scheme in **Chap. 3.1**. The number of trainees per category was 1500 and the number of validation data per category was 300. The computational time consumption was 4 hours 50 minutes 54 seconds, and the global accuracy reached at 91.94%. The estimation accuracies of each wave height class for designed long-crested waves are shown in **Fig. 4.4** and the correlation matrix between original data and estimated data was illustrated in **Table 4.1**.

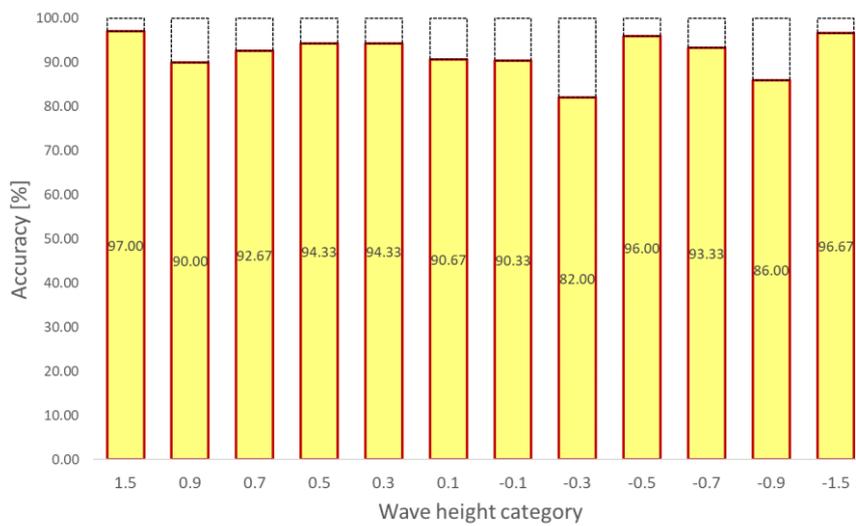


Fig. 4.4. Estimation of each wave height category for long-crested waves.

Table 4.1. Correlation matrix for long-crested waves.

	PREDICTED DATA												
	1.5	0.9	0.7	0.5	0.3	0.1	-0.1	-0.3	-0.5	-0.7	-0.9	-1.5	
ORIGINAL DATA	1.5	97.000	2.667						0.333				
	0.9	3.667	90.000	6.000	0.333								
	0.7		1.000	92.667	6.000	0.333							
	0.5	0.333		1.000	94.333	4.333							
	0.3				4.667	94.333	1.000						
	0.1					7.000	90.667	2.333					
	-0.1					0.333	3.667	90.333	5.333	0.333			
	-0.3					0.667	0.333	2.333	82.000	14.667			
	-0.5								1.000	96.000	3.000		
	-0.7									4.000	93.333	2.667	
	-0.9										6.333	86.000	7.667
	-1.5											3.333	96.667

In validating data, all categories showed the relatively precise accuracies over 90.0% except for the ‘-0.3’ category with 82.0% estimation accuracy and the ‘-0.9’ category with 86.0% estimation accuracy. In machine learning through CNN, fairly complicated feature extractions and classifications are performed with numerous artificial neurons, so it is almost impossible to identify the estimation cause at the training stage. Recently, the field of Explainable AI(EAI) has been researched to find out the estimation reasons, but this study was limited to the application of CNN’s wave field estimation.

The largest and the lowest wave height groups, which delegates the wave field condition, were correctly estimated over 95.0%. In the correlation matrix, the dominance of diagonal terms was strongly shown, even a singular estimation in the ‘1.5’ category with the estimation of ‘-0.5’ was exist. In other words, the band-width was quite narrow and estimations were mostly adjacent to the diagonal terms.

We then adopted numerous image filtering techniques to evaluate their effectiveness in wave estimation. The black & white filtering applications were shown in **Fig. 4.5**.

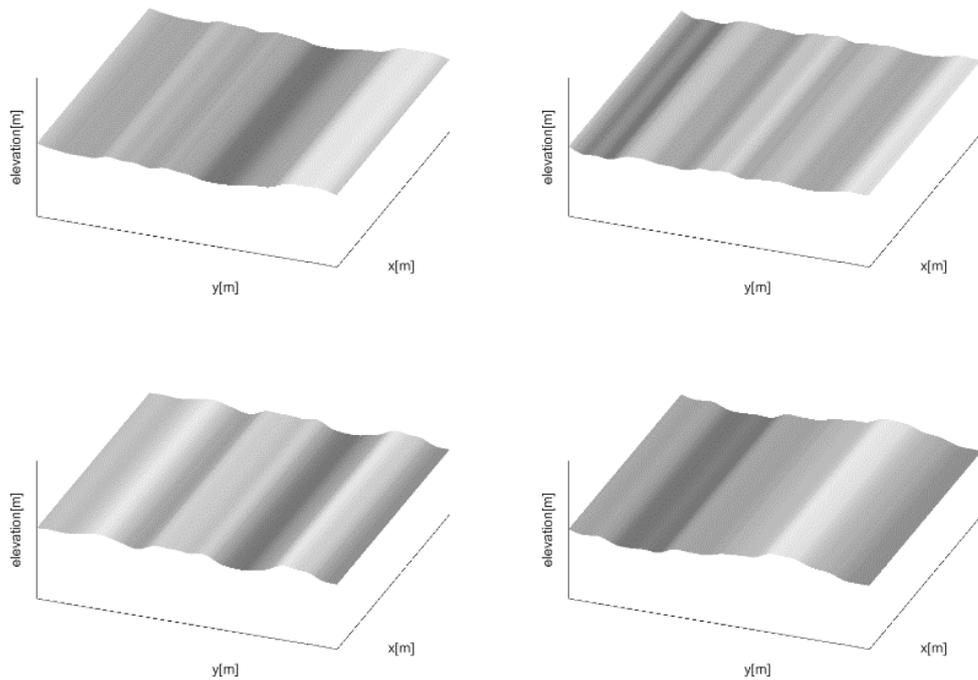


Fig. 4.5. Black & White filtered long-crested wave snapshots.

The cropping technique was also adopted for utilizing the central parts of images, which is shown in **Fig. 4.6**. After cropping the images, other widely used techniques were adopted, some filtered images are shown in **Fig. 4.7**. As explaining these techniques briefly, the blurring technique subtracts the high frequency noise in images, the edge enhancement technique emphasizes edges with the kernel matrices such as $[-1, -1, -1; -1, 9, -1; -1, -1, -1]$ 3 by 3 matrix, the embossing technique turns images into gray patterns and remains highlights only, and the edge detection figures out the high gradient parts. These filters can be applied separately or simultaneously.

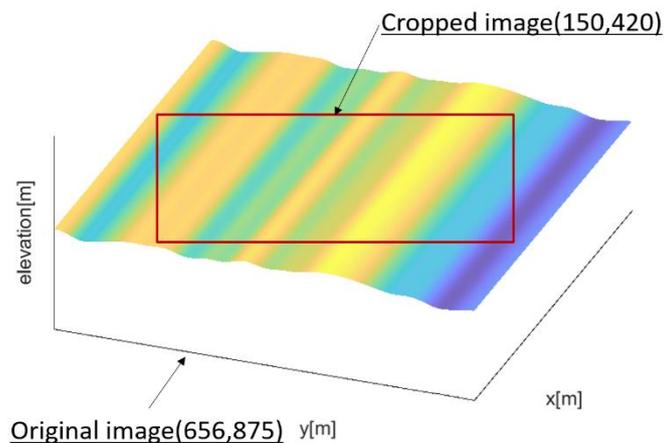


Fig. 4.6. Original and cropped images.

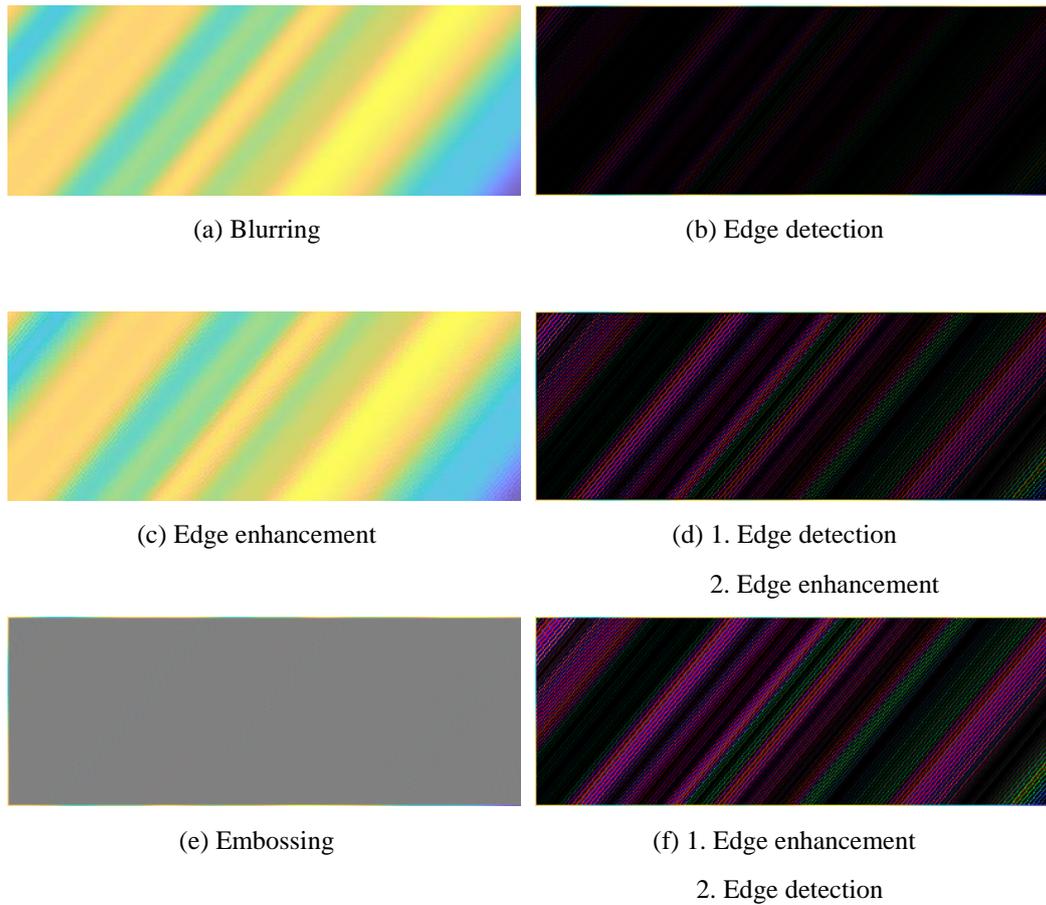


Fig. 4.7. Various filtered images with cropped images.

The deep learning simulations were conducted with the same CNN structure and numerical schemes under the identical computational environment with different training data. Two estimation results using image filtering techniques are shown in **Table 4.2** and **Table 4.3**. The former result was achieved from the black & white filter application only and the latter one was carried out from the sequentially applications of image filters as cropping, edge enhancement, and edge detection. In **Table 4.2**, the estimation results at diagonal terms were slightly lower, the global tendency of the estimation result of black & white filtered data training was similar to that of original RGB based data training. Two wrong estimations which are separated far away from the answer were shown in the '-0.1' training category, while all estimations were exist within two adjacent categories of answers.

While the estimation of training data with multiple filters(cropping, edge enhancement, and edge detection sequentially, hereinafter called 'method f') showed a different tendency, shown in **Table 4.3**. The band-width of the matrix was expanded and the estimation results of diagonal terms were dramatically reduced in moderate wave height groups. The lowest estimation result was 45.0% in the '-0.7' validation groups, and results of three categories were lower than 50.0%. It is believed that applying too many image filters crushes or disappears the features of the image, resulting in low classification accuracy by machine learning. A more meticulous approach

to applying a filter that can more clearly emphasize the characteristics of an image is needed.

Table 4.2. Correlation matrix for black & white filtered long-crested waves.

	PREDICTED DATA												
	1.5	0.9	0.7	0.5	0.3	0.1	-0.1	-0.3	-0.5	-0.7	-0.9	-1.5	
ORIGINAL DATA	1.5	97.67	2.33										
	0.9	7.00	90.00	2.67	0.33								
	0.7	0.33	4.00	89.67	6.00								
	0.5			5.00	90.00	5.00							
	0.3				7.33	90.00	2.67						
	0.1				0.33	10.33	88.00	1.33					
	-0.1	0.33		0.33		1.67	4.33	80.00	13.33				
	-0.3							0.67	96.00	3.33			
	-0.5								7.00	86.00	7.00		
	-0.7									3.67	94.00	2.00	0.33
	-0.9										6.00	90.33	3.67
	-1.5											7.00	93.00

Table 4.3. Correlation matrix for (f) filtered long-crested waves.

	PREDICTED DATA												
	1.5	0.9	0.7	0.5	0.3	0.1	-0.1	-0.3	-0.5	-0.7	-0.9	-1.5	
ORIGINAL DATA	1.5	98.33	1.67										
	0.9	4.67	95.00	0.33									
	0.7		4.33	75.33	20.00	0.33							
	0.5			7.00	76.67	15.00	1.33						
	0.3			0.67	25.00	47.33	25.00	1.67	0.33				
	0.1				3.33	18.67	60.33	15.33	2.00			0.33	
	-0.1				0.33	0.67	18.33	47.00	30.67	2.67	0.33		
	-0.3			0.33			1.33	12.00	53.67	30.67	0.33	1.00	0.67
	-0.5							0.33	22.67	53.00	19.00	4.33	0.67
	-0.7								2.67	26.67	45.00	24.00	1.67
	-0.9									2.00	14.67	67.00	16.33
	-1.5										0.33	16.00	83.67

The estimations per category of three trainings with different data sets, RGB images, black & white images, data using method f, are illustrated in **Fig. 4.8**. The main estimation results comparison is written in **Table 4.4**.

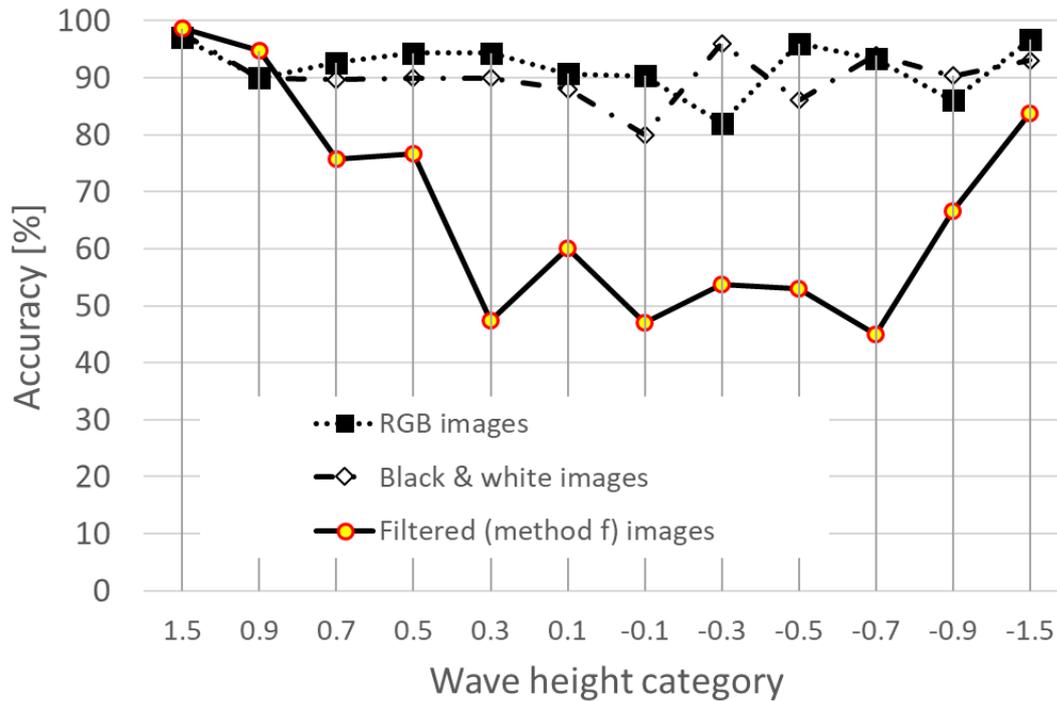


Fig. 4.8. Estimation accuracy with different filter applications in long-crested waves.

Table 4.4. Wave height estimation results with different filter applications.

Data	Original	Black & white	(f) filtered
Computation time	5hr 50m 54s	5h 8m 10s	21m 37s
Validation accuracy	91.94%	90.39%	66.86%
Highest accurately estimated category(accuracy)	1.5(97.0%)	1.5(97.7%)	1.5(98.3%)
Lowest accurately estimated category(accuracy)	-0.3(82.0%)	-0.1(80.0%)	-0.7(45.0%)
Standard deviation of accuracy results per category	4.29	4.45	18.02

The RGB images drew the most accurate estimation result, which took about 5 hours and 50 minutes. The training result with black & white images gave the similar global accuracy with that with RGB images, while the computing time was reduced about 12.2%. The training with method f applied data took less than 22 minutes, which is 6.2% of RGB images training, the global accuracy reached at only 66.86%. The highest accurate estimation category was ‘1.5’ for all trainings, the lowest one was changed among three trainings. We inferred that numerous filters considerably squashed the characteristics of long-crested waves.

4.2 Estimation on short-crested waves

In order to evaluate the applicability of the convolution neural network for complicated ocean conditions, short crested waves with three different wave trains were produced. For a short-crested wave field, the wave spectral density is expressed as **Eq. 4.5**. It can be represented by the summation of the products of wave spectrum and directionality function.

$$S(\omega, \theta) = \sum_{i=1}^n S_i(\omega) D_i(\omega, \theta) \quad (4.5)$$

The short-crested wave spectrum, $S(\omega, \theta)$, is the sum of i-th uni-directional wave spectra, $S_i(\omega)$ with i-th directionality functions, $D_i(\omega, \theta)$. The directionality function fulfils the requirement as **Eq. 4.6**. The summation in all directions should be equal as 1.

$$\int_{\theta} D(\theta, \omega) d\theta = 1 \quad (4.6)$$

In this study, the directionality function was equally applied to each spectrum for the sake of simplicity. The time domain conversion for each spectral densities is given by **Eqs. 4.1-4.4**, and time series are superimposed.

The third wave train was the strongest, which had 7.7m of a significant wave height, 17.0s of a peak period, 3.3 of a peak parameter, and 45 degrees of a wave incoming direction. The first and second wave train parameters were chosen as widely used values for sea state 4 and sea state 3, their wave incoming directions were 90 degrees and 130 degrees, respectively. The wave spectrum of the short-crested waves and the wave incoming directions are shown in **Fig. 4.9** and **Fig. 4.10**. A few samples are illustrated in **Fig. 4.11**.

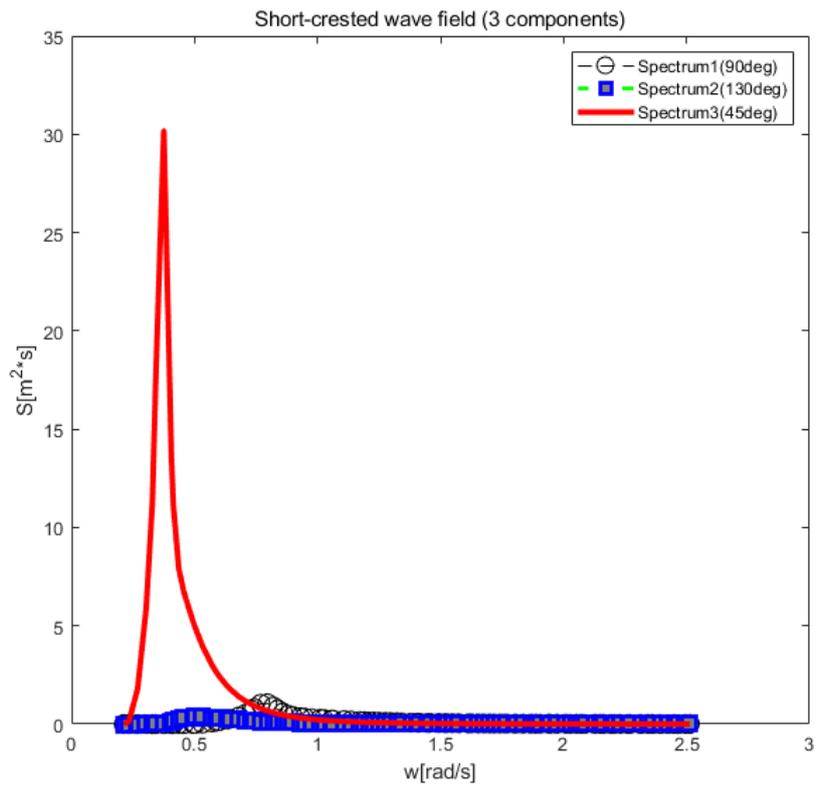


Fig. 4.9. Short-crested wave spectrum.

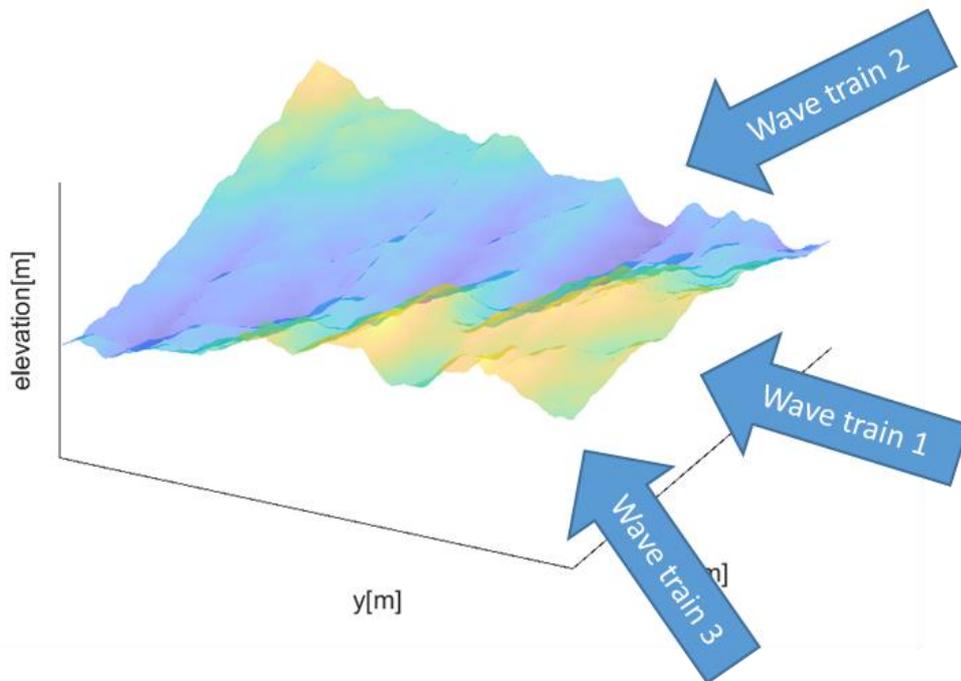


Fig. 4.10. Incoming directions of wave trains.

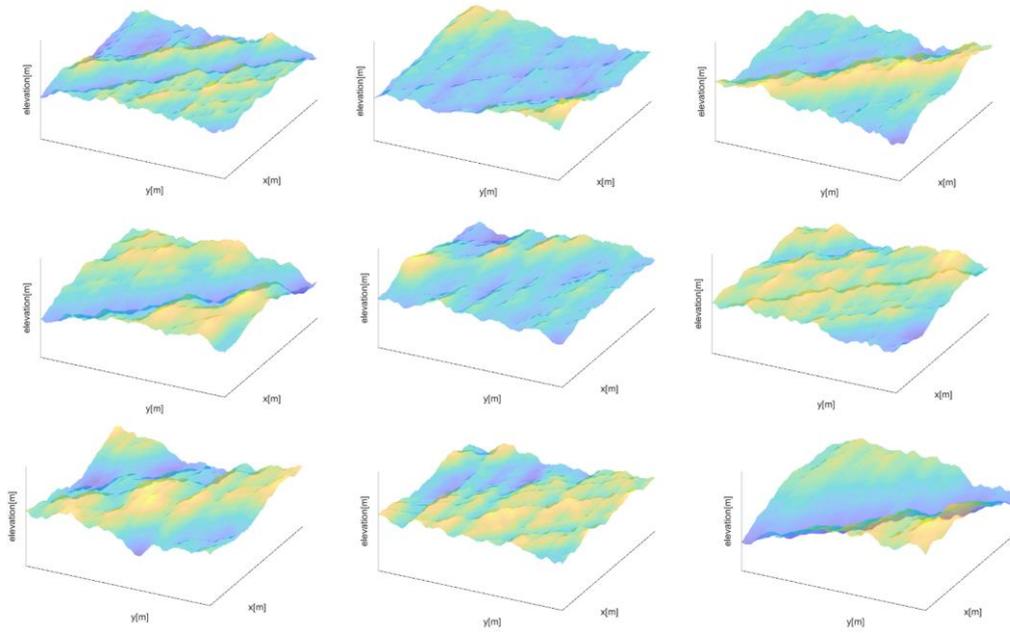


Fig. 4.11. Snapshots of modelled short-crested wave.

The wave data were supervised with fourteen categories with the same strategy of the long-crested waves training, where the maximum amplitude was 6.88m and the minimum amplitude was -6.86m. A number of training data per category was 1500, and that of validation data per category was 300. All training setups were identical to those of the long-crested waves training. The total computational time was 6 hours 30 minutes 42 seconds, which is slightly longer than that of the long crested waves training as two more categories were added. The global accuracy for the short crested waves training is shown in **Fig. 4.12**, and the cross entropy loss is shown in **Fig. 4.13**.

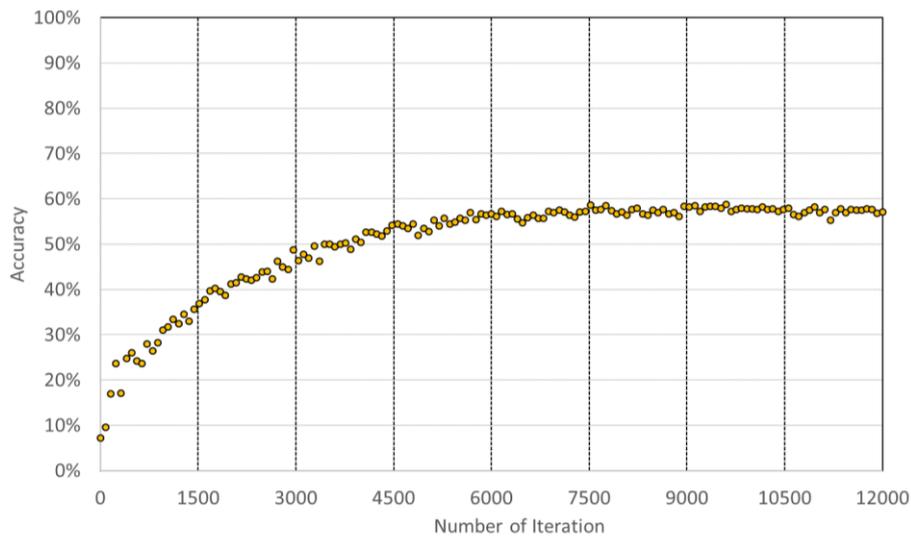


Fig. 4.12. Progress report of global accuracy: short crested wave.

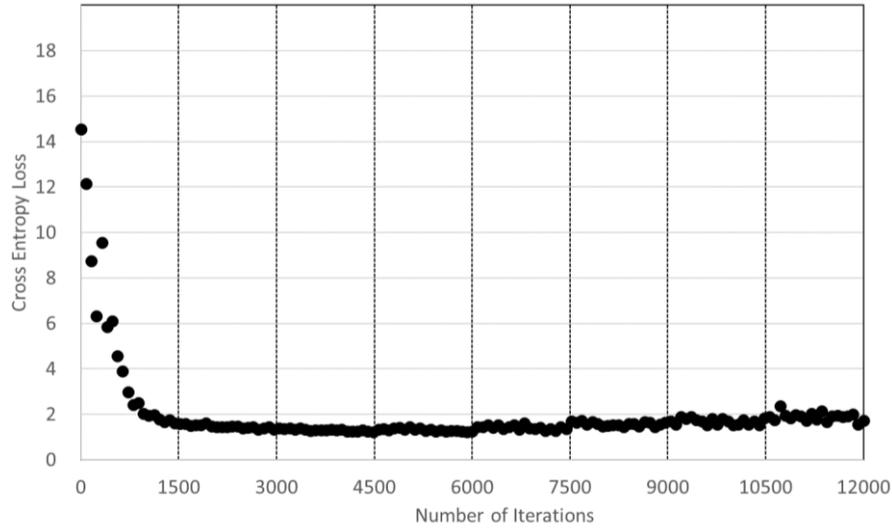


Fig. 4.13. Progress report of cross entropy loss: short crested wave.

The global accuracy reached at 57.10% and the cross entropy loss arrived at 1.7219, which the estimation result was quite poor compared to long crested waves trainings. Even the whole data set was penetrated 8 times, the cross entropy loss never decreased under 1.5. It can be rephrased that a certain amounts of discrepancies were continuously through the whole training stage. We checked the correlation matrix shown in **Table 4.5** to examine the estimation result more closely.

The estimation accuracies of the largest and the lowest categories were 87.67 and 90.33, relatively. While those of other categories were drawn between 39.67 and 67.00. The band width of the correlation matrix was quite enlarged than those in **Table 4.1** to **4.3**. The sample snapshot of wrong estimations was shown in **Fig. 4.14**, where the answer was '1.25' and the estimation was '-4.50'. The estimation results per category for this snapshot were 61.44% for '-4.5', and 18.44% for '1.25'. While two steep wave trains were found in this sample, which could be a reason of the wrong estimation. However the reasoning of the wrong estimation with a simple convolution neural network was not possible as we cannot examine the whole variations on weightings and biases during the training. Even the short crested waves were modelling with just three wave trains, we thought the training data set could not present the reliable estimation in wave height classification.

Table 4.5. Correlation matrix for short-crested wave.

	PREDICTED DATA																		
	4.5	3	2.25	1.75	1.25	0.75	0.25	-0.25	-0.75	-1.25	-1.75	-2.25	-3	-4.5					
ORIGINAL DATA	87.67	10.67	1.00	0.33	0.33														
4.5		4.00	67.00	21.33	5.33	2.33													
3			0.67	8.67	64.00	21.33	3.67	0.67	0.67	0.33									
2.25				0.33	26.33	48.33	16.67	4.67	1.33	1.33	0.67	0.33							
1.75					1.00	6.33	22.00	44.33	16.00	5.00	3.33	0.67	1.00	0.33					
1.25						0.33	2.00	8.33	20.67	39.67	20.00	5.00	1.67	2.00	0.33				
0.75							0.33	2.33	5.00	12.00	47.00	25.00	4.00	3.33	0.67	0.33			
0.25								0.33	3.67	3.33	22.00	45.67	13.67	9.33	1.00	0.67	0.33		
-0.25									1.67	7.00	16.00	43.67	26.00	2.67	1.33	1.33	0.33		
-0.75										0.33	0.67	3.00	4.67	17.00	56.33	13.33	3.33	1.00	0.33
-1.25											1.00	2.33	3.67	24.67	41.67	21.33	3.67	1.67	1.67
-1.75												1.67	9.00	13.00	60.67	13.00	2.67	2.67	2.67
-2.25													2.33	1.67	17.33	63.00	15.67	15.67	15.67
-3														0.33	2.33	7.00	90.33	90.33	90.33
-4.5																			

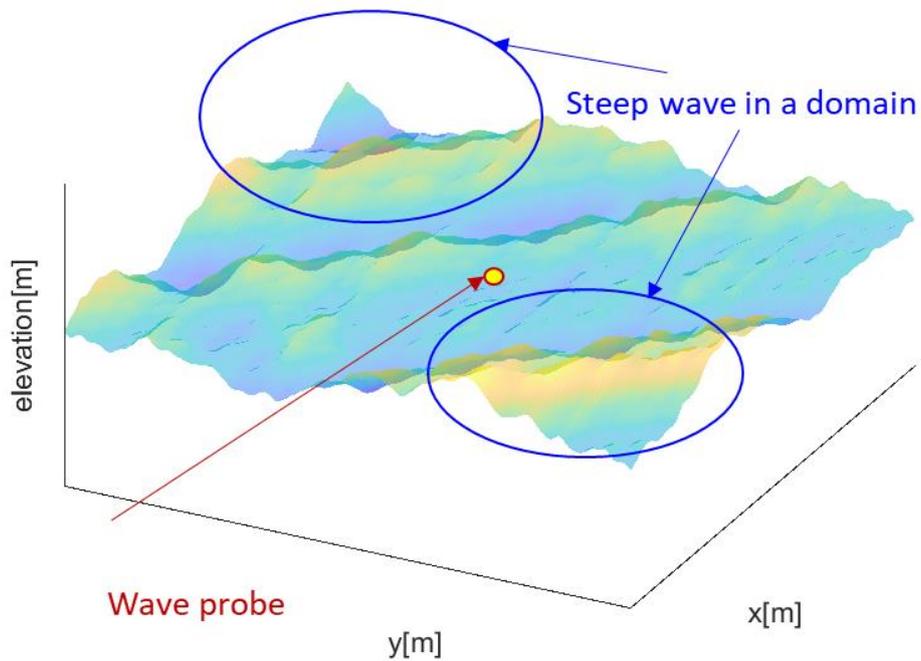


Fig. 4.14. A sample ill-estimated case: Answer: '1.25', Estimation: '-4.50'.

We then varied some training options or input files to increase the estimation accuracy with short crested waves. A series of parametric studies were conducted, while we added three cases in this paper. The variation on the max pooling applications was defined as ‘Case. A’, the confirmation on the number of convolutional layers was defined as ‘Case. B’, and the modification on the dominant wave sea state was defined as ‘Case. C’.

The concept of the parametric study case A is illustrated in **Fig. 4.15**. We changed the max pooling size as 2 by 2(Case. A1), 4 by 4(Case. A2), and 8 by 8(Case. A3) without the overlapped pooling. Lastly, one additional case(Case. A4) was added in the 8 by 8 pooling case allowing the overlapped pooling with the stride distance of 4 by 4. All max pooling techniques in Fig. 2 were modified as above settings. **Fig. 4.16** shows the global accuracy and **Fig. 4.17** shows the corresponding cross entropy loss in Case. A. Estimations per each category are represented in Fig. 4.18.

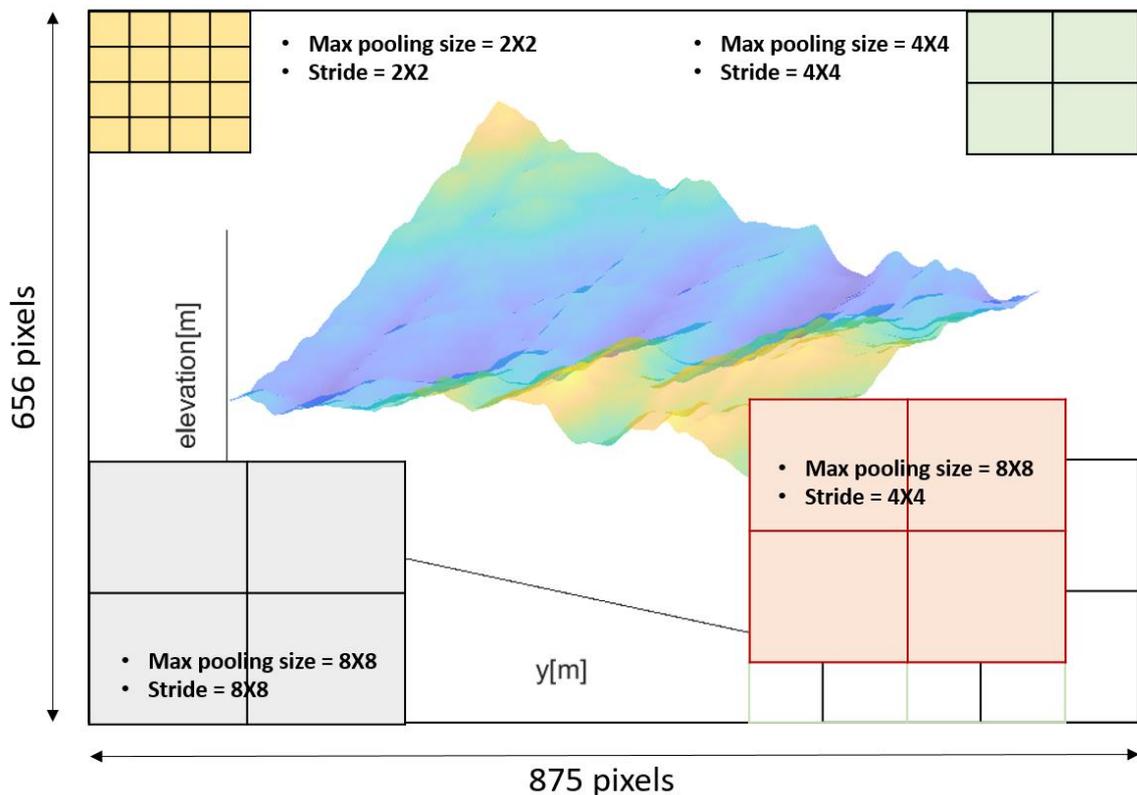


Fig. 4.15. Variation on max pooling technique application.

The global accuracy of the Case. A3 reached at 45.29%, while those of other cases were over 50.0%. The Case. A1 and the Case. A3 were seemed to be stable after the 4 epochs, the Case. A2 and the Case. A4 were still increased at the end of iterations. Except for the Case. A1, the cross entropy loss of three cases were extremely reduced at the beginning of trainings. While the cross entropy loss of the Case. A1 also followed other quantities’ tendency after 1 epoch iteration, even there was a fluctuation at the end of iterations. All cross entropy losses had

a certain amount of value more than 1.0 at the end of iterations.

In terms of estimation results per category, shown in **Fig. 4.18**, all cases showed the highest estimation results at the largest and lowest wave height categories except for the Case. A4. The ill-estimation in the middle size wave height categories in all cases. Additionally, the lowest estimations of the Case. A3 and the Case. A4 did not exceed 20.0%. The Case. A1 and the Case. A2 seemed to have a better estimation performance, while all results were not appropriate for wave height estimation, differently from the long crested wave deep learning was. The consequent validation accuracies and computational times are listed in **Table 4.6**. In aspect of the estimation accuracy and computational time effectiveness, the Case. A2 was the most suitable with short crested wave data in this research.

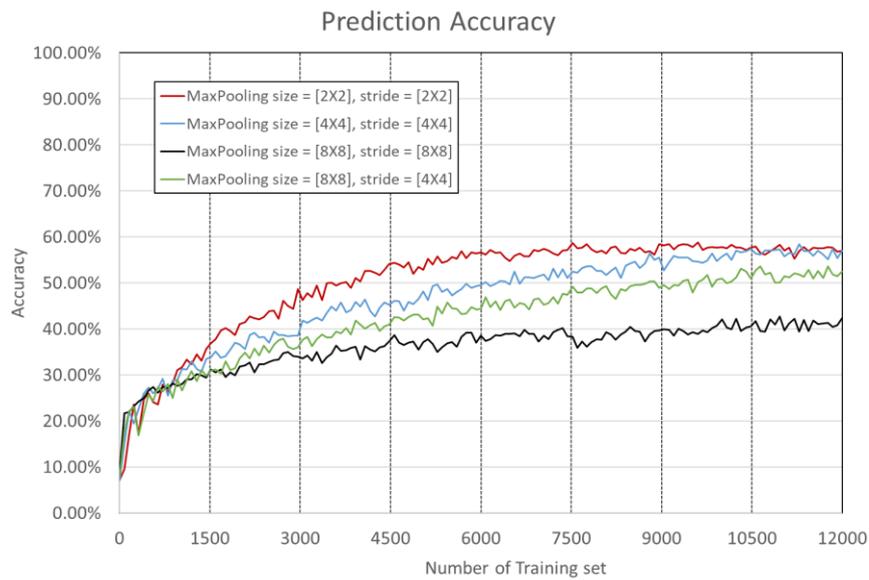


Fig. 4.16. Global accuracy with various max pooling conditions.

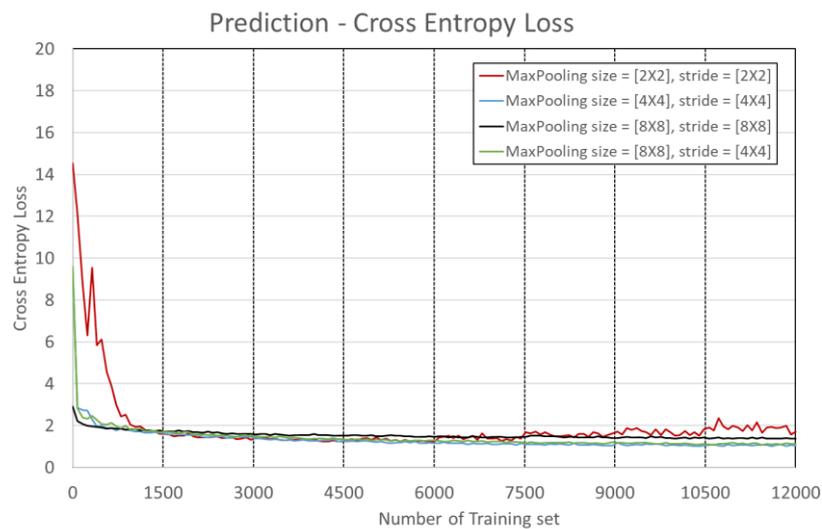
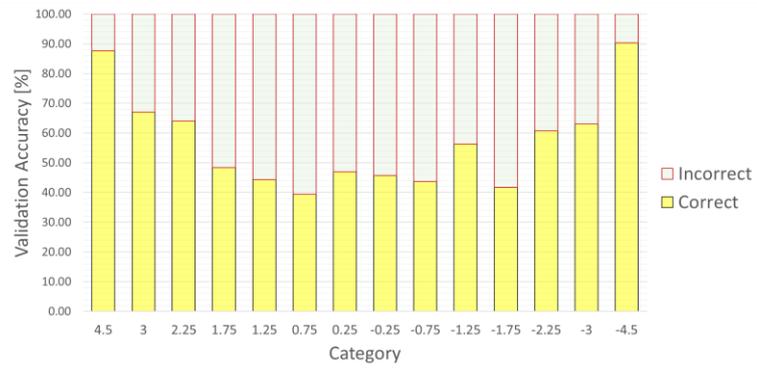
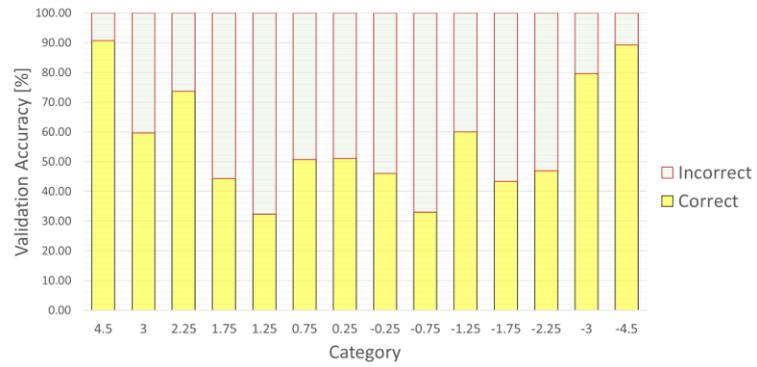


Fig. 4.17. Cross entropy loss with various max pooling conditions.

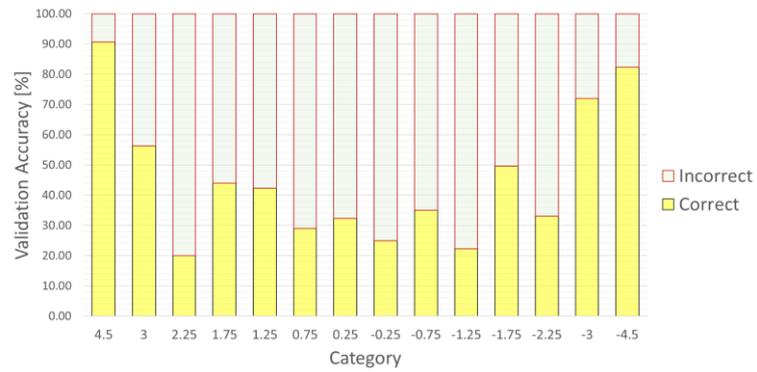
Case. A1.



Case. A2.



Case. A3.



Case. A4.

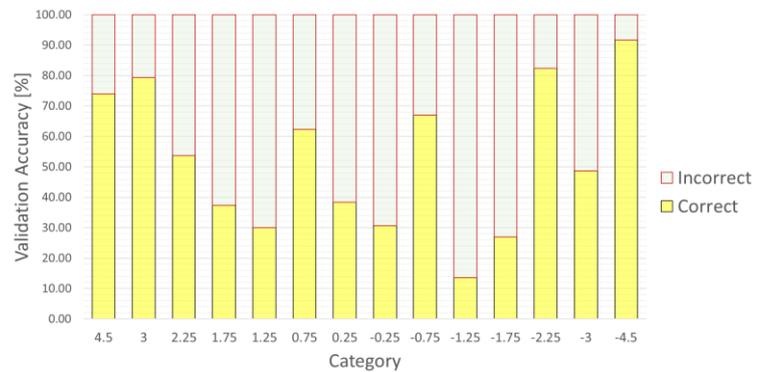


Fig. 4.18. Estimation accuracy by categories with various max pooling conditions.

Table 4.6. Validation accuracy and computational time with various max pooling conditions.

	<u>Case.A1</u>	<u>Case.A2</u>	<u>Case.A3</u>	<u>Case.A4</u>
	Max pooling size =2×2, Stride=2×2	Max pooling size =4×4, Stride=4×4	Max pooling size =8×8, Stride=8×8	Max pooling size =8×8, Stride=4×4
Validation accuracy [%]	57.10	57.21	45.29	52.57
Data training time	390m 42s	209m 45s	192m 9s	211m 42s

The second attempt was the variation on the number of the convolution layers. Our initial model used four layers shown in **Fig. 3.3**, and changed the number of convolution layers from one to five. The detailed information of the Case. B is listed in Table 6. In order to suppress the complexity, the filter size was not changed and the number of filters followed the initial training structure. Also the global accuracy and the computational time are included in **Table 4.7**.

Table 4.7. Convolution layer specifications in variation of number of convolution layers.

	Number of convolution layers	Filter size	Number of filters	Validation accuracy[%]	Data training time
Case. B1	1	3×3	16	45.62	344m 31s
Case. B2	2	3×3	16	53.07	369m 19s
		3×3	32		
Case. B3	3	3×3	16	56.40	380m 7s
		3×3	32		
		3×3	64		
Case. B4	4	3×3	16	57.10	390m 42s
		3×3	32		
		3×3	64		
		3×3	64		
Case. B5	5	3×3	16	60.52	516m 23s
		3×3	32		
		3×3	64		
		3×3	64		
		3×3	64		

The global estimation accuracy, the cross entropy loss, and the estimation accuracy per category in the Case. B are illustrated in **Fig. 4.19**, **Fig. 4.20**, and **Fig. 41**, respectively. The Case. B1 showed the minimum consequent accuracy of 45.62%, while other accuracies were obtained between 53.07% and 60.52%. All global accuracies were stabilized after 4 epochs. In terms of cross entropy loss, the cross entropy loss of the Case. B1 was steadily decreased in the whole iterations, while those of other cases showed a dramatic falling within 1 epoch and remained at the certain level. Even we eliminated two layers in the Case.B2 form the initial condition, the Case. B4, the estimation tendency was not changed significantly. The computational time was not quite increased up to the training with four convolutional layers, while there was a huge increase between four layers and five layers. In this parametric study, still the high estimation results were concentrated to the highest and lowest wave height categories. Similarly to the variation on the max pooling conditions, the variation on the number of convolution layers showed the insignificant effect in wave height estimation of the short crested waves.

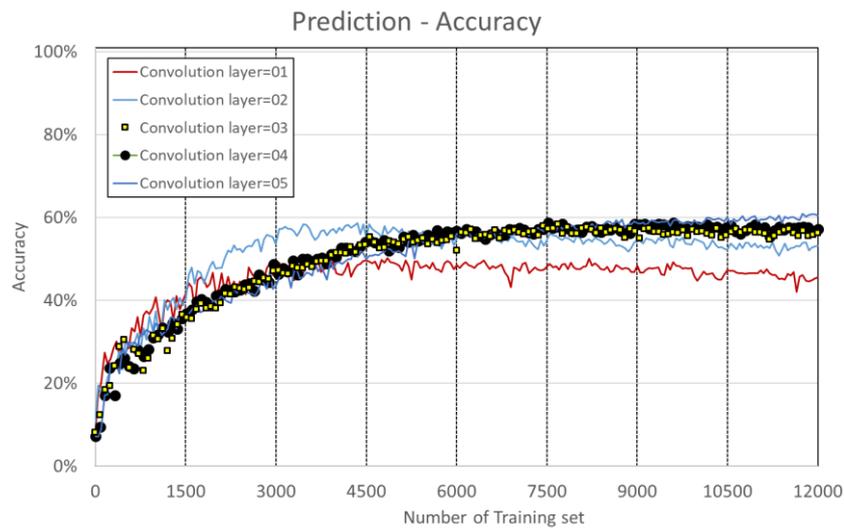


Fig. 4.19. Global accuracy with various number of convolution layers.

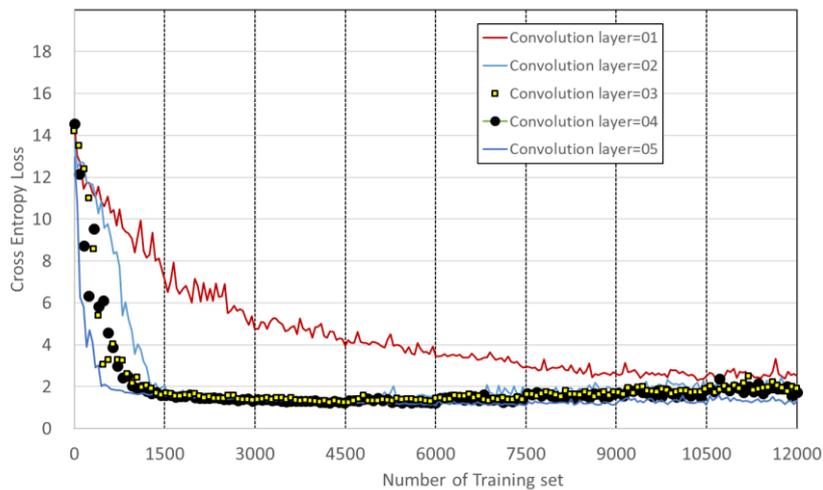


Fig. 4.20. Cross entropy loss with various number of convolution layers.

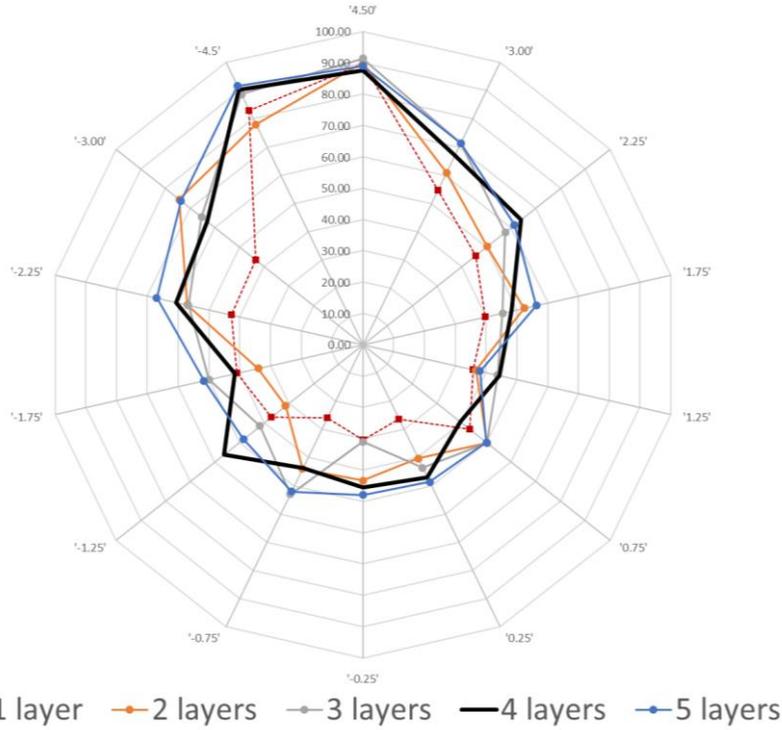


Fig. 4.21. Accuracy by number of layers in categories with various number of convolution layers.

We lastly changed the short crested wave conditions with variation on the wave train 3 to figure out the data dependency in short crested wave deep learnings. The wave spectrum of the train 3 was varied from sea state 3 to 7. The detailed wave conditions are listed in **Table 4.8** and some snapshots of sea state 7, 5, and 3 are shown in **Fig. 4.22**.

Table 4.8. Wave conditions with wave train 3 variations.

Wave Train	Train 1	Train 2	Train 3				
Case	-	-	Case.C1	Case.C2	Case.C3	Case.C4	Case.C5
Sea State	SS3	SS4	SS7	SS6	SS5	SS4	SS3
H_S [m]	1.22	1.98	7.50	5.00	3.25	1.88	0.88
T_p [s]	8.00	9.50	15.00	12.40	9.70	8.80	7.50
γ [-]	2.5	1.0	3.3	3.3	3.3	3.3	3.3
Heading [deg]	160.0	180.0	225.0	225.0	225.0	225.0	225.0
Number of categories			14	12	8	6	8
Training data per category			1400	1100	2300	3000	3900
Total number of trainees			19600	13200	18400	18000	31200
Validation accuracy[%]			57.86	59.42	66.08	78.17	68.13
Data training time			322m 39s	202m 39s	305m 27s	212m 50s	541m 34s

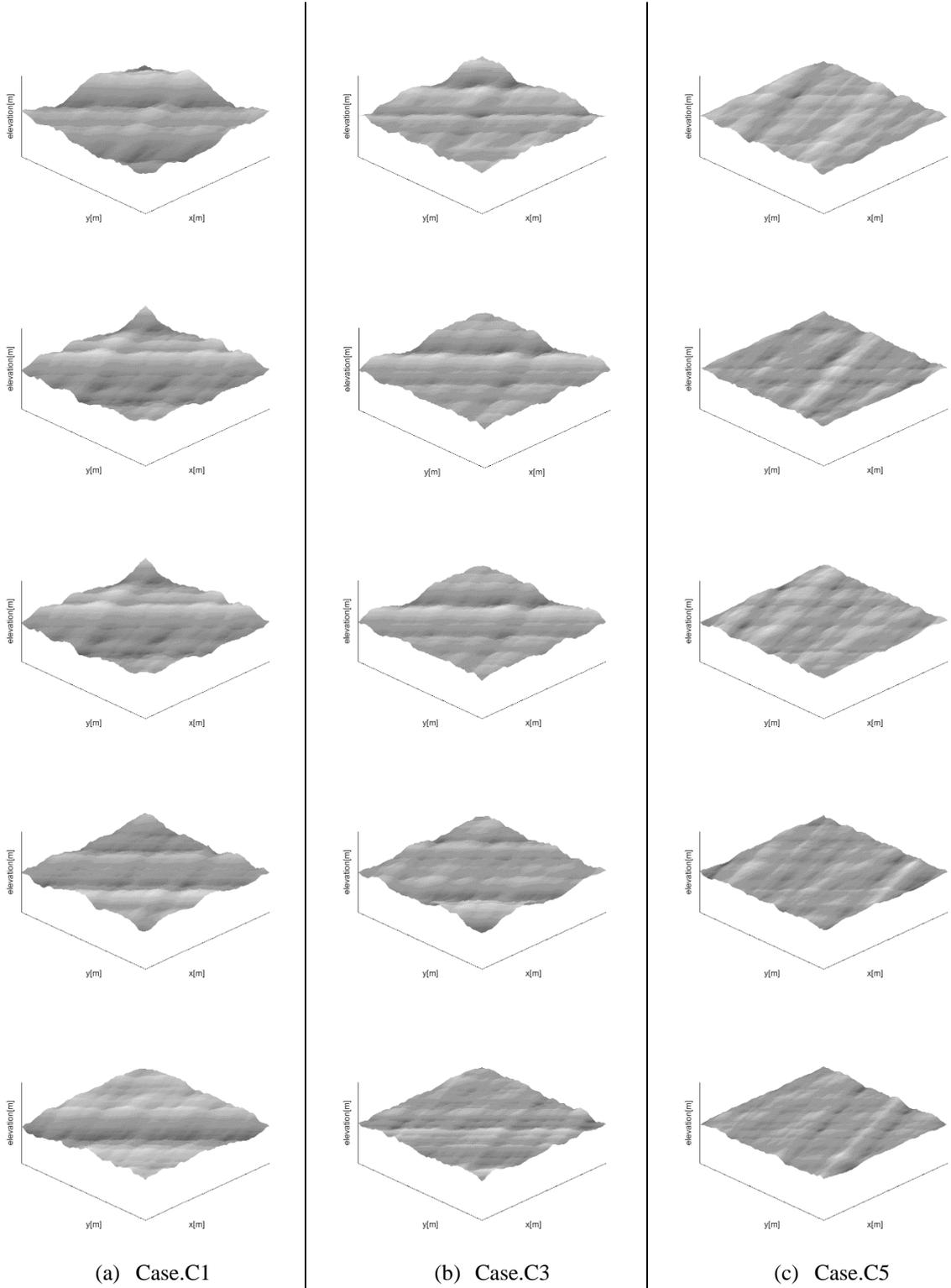


Fig. 4.22. Sea-state variations for (a) Case. C1, (b) Case. C3, and (c) Case. C5.

As the wave train 3 became smaller waves from Case. C1 to Case. C5, the wave height classifications were changed with maxima and minima of each wave field. As Case. C3, Case. C4, and Case. C5 were similar to two unchanged wave trains, categories of those conditions were smaller and the number of trainees per category was increased. **Fig. 4.23** shows the global accuracy and **Fig. 4.24** shows the cross entropy loss of the wave spectrum variations. The increase of computational time was proportional to the increase of the total number of trainees generally.

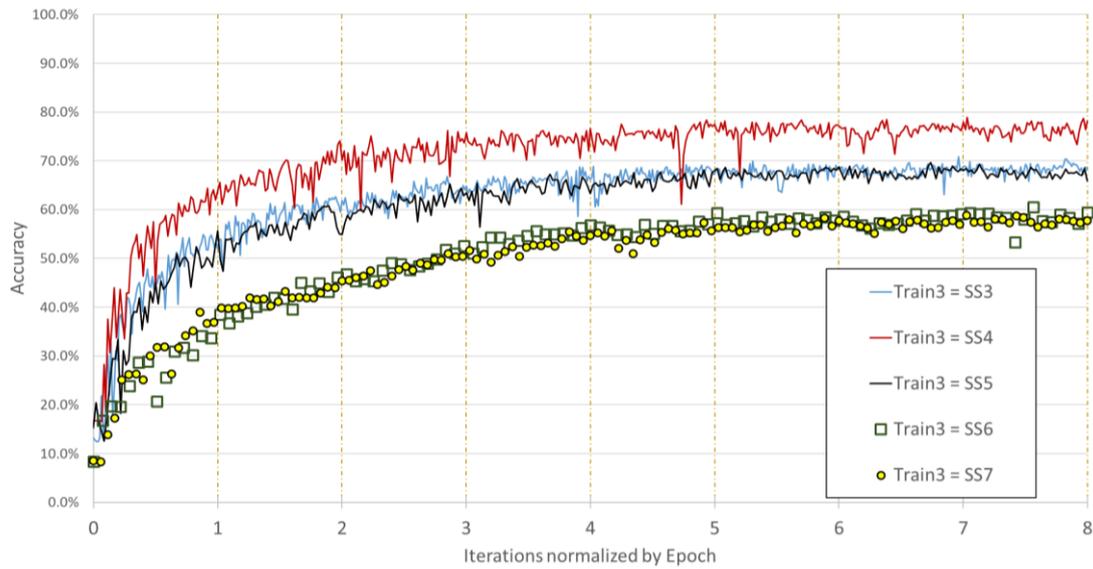


Fig. 4.23. Global accuracy in wave spectrum variation.

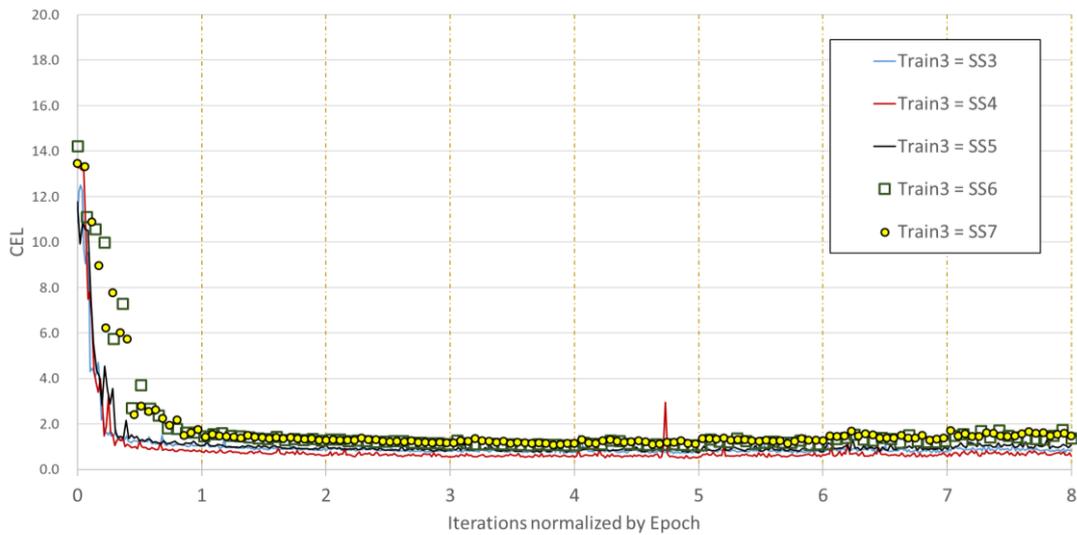


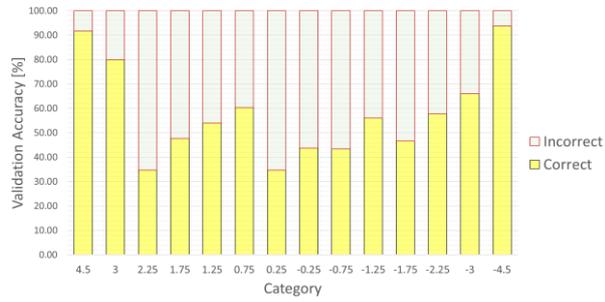
Fig. 4.24. Cross entropy loss in wave spectrum variation.

The Case. C4 showed the highest global accuracy of 78.17%, while other estimation results were obtained from 57.86% to 68.13%. The tendency between the Case. C1 and the Case. C2 was quite similar, also the Case. C3 and the Case. C5 showed analogous results. We inferred the highest estimation in the Case. A4 was achieved as this case had the smallest number of categories and the sufficient number of trainees per category. While none of cases exceeded 80.0% and all cross entropy losses had a certain amount of values at the end of iterations.

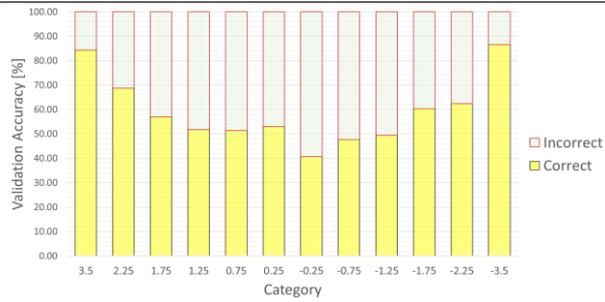
In order to examine the results precisely, the estimation accuracies per each category and the correlation matrices are illustrated in **Fig. 4.25**. The lowest diagonal term values were 34.67%, 40.67%, 51.33%, 64.00%, and 46.67% from the Case. C1 to C5, respectively. Differently from the Case. C1 and the Case. C2, the diagonal dependencies were increased in the Case. C3 to C5. We indirectly found that the estimation accuracy can be increased by the increase of the number of trainees per category. Still the wave height classifications were inappropriate for the median range of wave heights. While we surmised that simplifying the categories, such as sea states or broaden wave height categories with enough data seemed to be appropriate in wave field estimation from the Case. C4 result.

With various attempts of deep learnings with short crested waves, we changed our strategy from wave height classification precisely to estimation of overall sea state conditions. As lots of trainings with short crested waves showed relatively high estimation results in the highest and lowest wave categories, which are representative groups in wave fields.

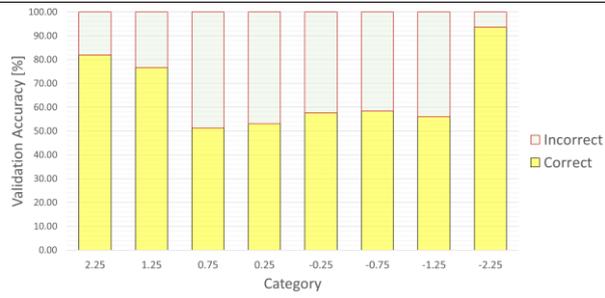
Case.C1.



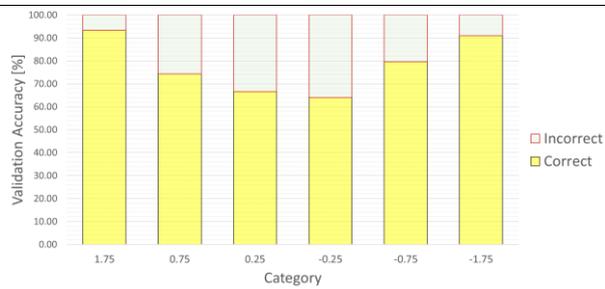
Case.C2.



Case.C3.



Case.C4.



Case.C5.

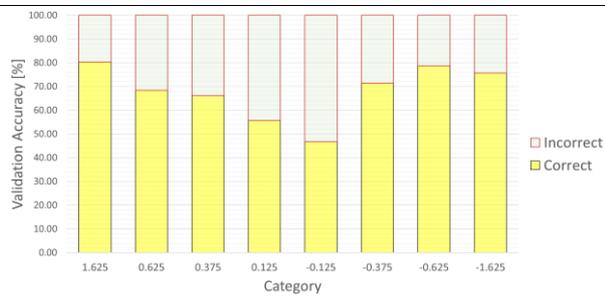


Fig. 4.25. Estimation accuracy per category in wave spectrum variation.

Chapter 5. Data acquisition and classification

We then obtained image-based real wave field data in the southwestern coastal area of Jeju Island in Korea. **Fig. 5.1** shows the data acquisition location including an ocean structure and a wave height gauge on Google Maps.

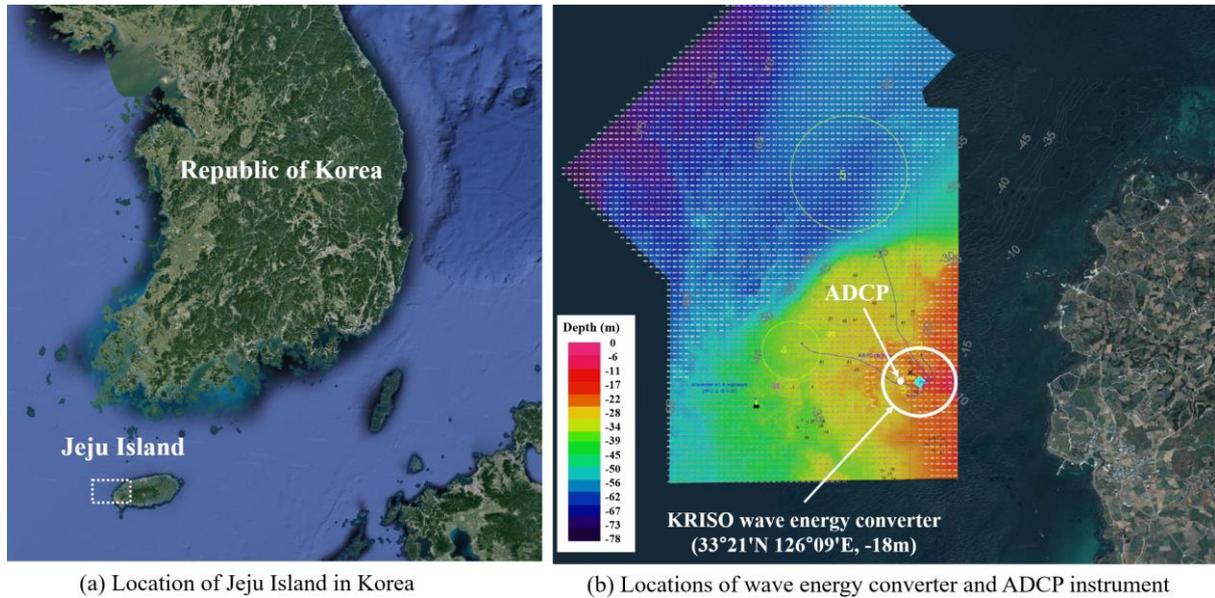


Fig. 5.1. Locations of image data acquisition, wave height measuring instrument.

The data required for the CNN were snapshot images obtained from the CCTV camera and wave heights measured from the ADCP instrument. A CCTV camera was installed on a fixed-type wave energy converter located in the southwestern coastal area of Jeju Island. **Fig. 5.2** illustrates the fixed-type wave energy converter and the CCTV camera on its top. In addition, in order to measure wave heights, we utilized the ADCP instrument made by AWAC [36], shown in **Fig. 5.3**, to label the ocean environment images. It can measure real-time current profiles and waves within a 50m water depth range. The wave height measuring range is from -15.0 m to $+15.0$ m and the wave period measuring range is from 1.0s to 50.0s. This device was installed at the designated sea bottom location of water depth 18.0m.

The well-conditioned data collection time was restricted and set from 7 a.m. to 11 a.m. because diffused reflections of sunlight at the sea surface did not occur in the morning. In addition, all images were converted into black and white to minimize the effect of light intensity. The CCTV camera was used for 1 hour every day to monitor the ocean surface and snapshot images were extracted from CCTV movie clips with intervals of 0.5 s. Extreme foggy and rainy movie clips were excluded, as were some snapshots with other obstacles such as ships.



Fig. 5.2. Fixed-type wave energy converter installed at Jeju Island and CCTV camera.



Fig. 5.3. Acoustic Doppler Current Profiler(ADCP).

The well-conditioned data collection time was restricted and set from 7 a.m. to 11 a.m. because diffused reflections of sunlight at the sea surface did not occur in the morning. In addition, all images were converted into black and white to minimize the effect of light intensity. The CCTV camera was used for 1 hour every day to monitor the ocean surface and snapshot images were extracted from CCTV movie clips with intervals of 0.5 s. Extreme foggy and rainy movie clips were excluded, as were some snapshots with other obstacles such as ships.

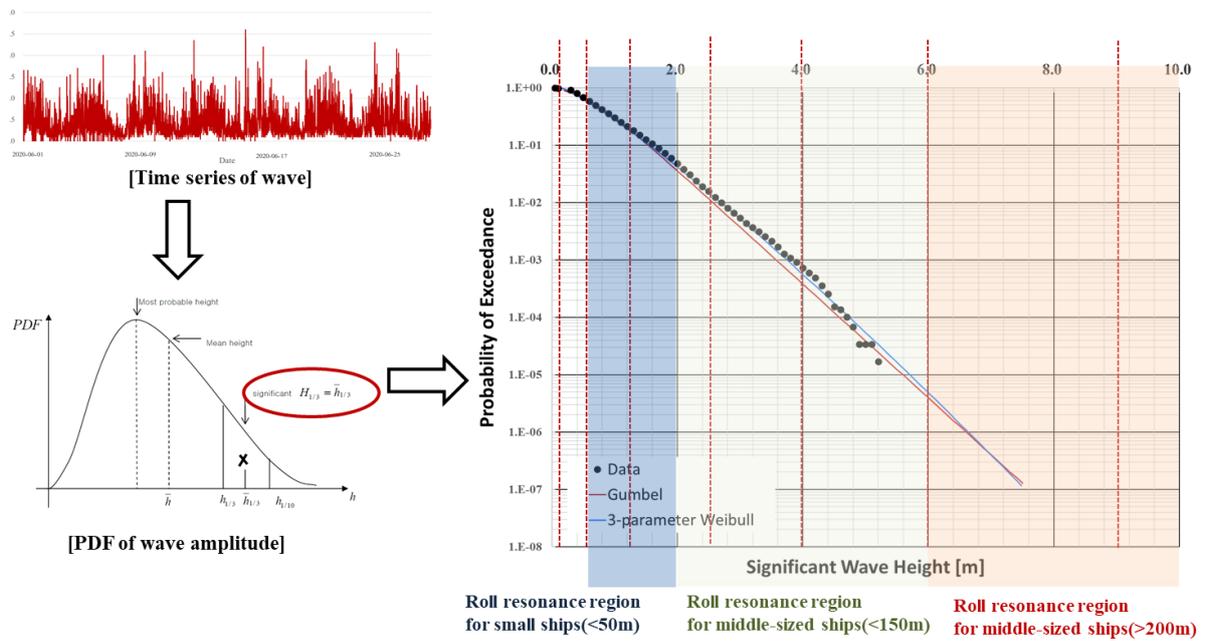


Fig. 5.4. Conceptual view of sea state classification by significant wave height.

Among many factors that determine the ocean environment, three primary environmental loads are winds, currents, and waves. Except for winds as wide-banded spectrum and currents as fairly long average periods (typically, 12 hours, 3 months, 18.6 years), return periods of waves are overlapped with the natural periods of ships. The average wave height in the upper third of the wave spectrum is defined as the significant wave height. **Fig. 5.4** shows the conceptual graphs of the probability of exceedance according to significant wave heights, which are obtained by wave spectrum. Small waves have a high occurrence frequency, and due to the dispersion relation, small waves correspond to short periods. At this time, periods of small waves overlaps with the natural periods of small vessels, with a short length of less than 50m. Therefore, in the case of small ships, it is important to predict large waves, but even small waves can cause large displacements due to the resonance.

The sea state can be classified as **Table 5.1** with the significant wave height, proposed by ITTC. The average wave and significant wave heights were obtained by analyzing the data recorded using the ADCP instrument. In **Table 5.2**, the well-conditioned data collected while monitoring the ocean environment are described, and the corresponding sea states are listed. General weather information and wind directions were obtained from the public information service from Korean Hydrographic and Oceanographic Agency (KHOA) [3]. Note that the average and significant wave heights were calculated from the recorded wave height time series, and the corresponding sea states were categorized by the significant wave height ranges according to the International Towing Tank Conference (ITTC) recommended procedures and guidelines [52]. Snapshots of each sea state are shown in **Fig. 5.5**.

Table 5.1. Sea state classification in ITTC Recommended Procedure [52]

Degree	Height [m]	Description
0	-	Calm (Glassy)
1	0~0.10	Calm (rippled)
2	0.10~0.50	Smooth
3	0.50~1.25	Slight
4	1.25~2.50	Moderate
5	2.50~4.00	Rough
6	4.00~6.00	Very rough
7	6.00~9.00	High
8	9.00~14.00	Very high
9	14.00~	Phenomenal

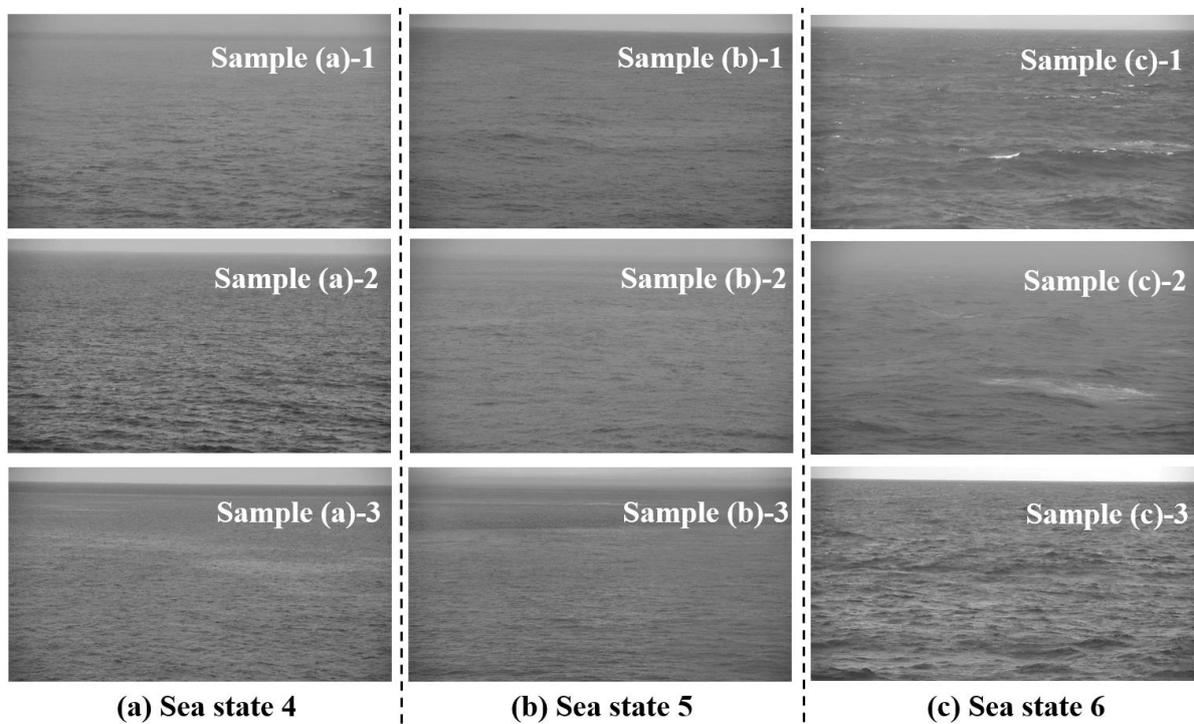


Fig. 5.5. Sample snapshot images of well-conditioned for (a) sea state 4, (b) sea state 5, and (c) sea state 6.

Table 5.2. Well-conditioned data gathered at ADCP installation location.

Date	Start time	End time	Condition	Average wave height[m]	Significant wave height[m]	Sea state	
May, 2020	17	10:30	11:30	Strongly cloudy	1.25	2.00	4
	19	10:00	11:00	Strongly cloudy	3.00	4.79	6
	20	8:00	9:00	Fine	1.50	2.39	4
	24	10:00	11:00	Strongly cloudy	1.00	1.60	4
	25	10:00	11:00	Cloudy	1.00	1.60	4
	26	10:00	11:00	Cloudy	1.00	1.60	4
	28	10:00	11:00	Cloudy	1.00	1.60	4
	29	10:00	11:00	Strongly cloudy	1.00	1.60	4
	31	10:00	11:00	Cloudy & rainy	1.25	2.00	4
June, 2020	2	10:00	11:00	Fine	1.25	2.00	4
	3	10:00	11:00	Cloudy & weakly rainy	1.00	1.60	4
	4	9:00	10:00	Strongly cloudy	1.50	2.39	4
	6	10:00	11:00	Strongly cloudy	1.50	2.39	4
	7	10:00	11:00	Fine	1.25	2.00	4
	8	10:00	11:00	Fine	1.25	2.00	4
	9	10:00	11:00	Cloudy	1.50	2.39	4
	14	10:00	11:00	Cloudy	2.50	3.99	5
	15	10:00	11:00	cloudy & weakly rainy	1.50	2.39	4
	16	10:00	11:00	Cloudy	1.25	2.00	4
	19	10:00	11:00	Cloudy	2.00	3.19	5
	20	10:00	11:00	Strongly cloudy	1.50	2.39	4
	21	10:00	11:00	Fine	1.50	2.39	4
	22	8:00	9:00	Fine	1.50	2.39	4
	26	8:00	9:00	Fine	1.50	2.39	4
	28	10:00	11:00	Strongly cloudy	1.50	2.39	4
July, 2020	1	8:00	9:00	Fine	1.50	2.39	4
	4	8:00	9:00	Cloudy	1.50	2.39	4
	5	9:00	10:00	Cloudy	2.00	3.19	5
	6	8:00	9:00	Cloudy & rainy	2.00	3.19	5
	7	7:00	8:00	Cloudy & rainy	2.00	3.19	5
	8	8:00	9:00	Strongly cloudy	2.00	3.19	5
	12	8:00	9:00	Cloudy & rainy	1.50	2.39	4
	13	8:00	9:00	Cloudy & rainy	3.00	4.79	6
	14	8:00	9:00	Cloudy	3.00	4.79	6
	15	8:00	9:00	Cloudy	1.50	2.39	4
	16	8:00	9:00	Cloudy	1.00	1.60	4
	17	7:30	8:30	Fine	1.25	2.00	4

Tables 5.3 and **5.4** present the number of datasets corrected according to the classification of the sea states and average wave heights, respectively. Only three sea states (4–6) were observed during the data recording period. About three-quarters of the collected data were defined as sea state 4. The number of samples in sea state 6 was the least among the three collected data. The average wave heights were categorized from 1.0 to 3.0 m with 0.25 m intervals. Three average wave heights (1.00, 1.25, and 1.50 m) belong to sea state 4, two average wave heights (2.00 m and 2.50 m) belong to sea state 5, and the average wave height of 3.00 m belongs to sea state 6. The highest number of samples was counted in the 1.50m category and the lowest one was counted in the 2.50m category. As a result, there are three and six categories in the sea state and average wave heights, respectively.

All categories in **Tables 5.3** and **5.4** were divided into three groups: training, validation, and test data. Training and validation data were used in the learning stage to produce a trained network. In the estimating stage, test data were utilized.

In all categories, the same number of data used for training was used, and random selection was made to avoid bias in the results. Note that the number of data obtained by category could vary considerably. Some studies have been conducted on the effect of disparity in these data categories on training results [53].

Table 5.3. Numbers of snapshot images utilized for the sea state estimation.

Sea state	Number of days	Ratio[%]	Number of snapshots
SS4	28	75.68	181,578
SS5	6	16.22	41,562
SS6	3	8.11	20,802

Table 5.4. Numbers of snapshot images utilized for the average wave height estimation.

Average wave height	Number of days	Ratio[%]	Number of snapshots
1.00m (SS4)	7	18.92	38,458
1.25m (SS4)	7	18.92	47,838
1.50m (SS4)	14	37.84	95,282
2.00m (SS5)	5	13.51	34,376
2.50m (SS5)	1	2.70	7,186
3.00m (SS6)	3	8.11	20,802

Meanwhile, there is a limit to the learning result using a clear image of a specific time period. Accordingly, all data for a certain period (September, 2021) were collected and learned without restrictions on the weather and time zone. Details of normal conditioned data achieving is shown in **Table 5.5** and the labelling ratio is shown in **Fig.**

5.6. It includes data from foggy, rainy weathers, unclear images, and data with other artificial obstacles such as ships. Two labels were assigned to each day, before- and after-noon. V_w means the wind speed, H_{ave} is the average wave height, and H_s is the significant wave height. Measuring methods and Information sources are identical to those of **Table. 5.2**.

Since computer hardware specifications were insufficient to collect and analyze data for all time data, five time zones were selected; night(03:00~03:05), sunrise(10 minutes before and after sunrise), morning(10:00~10:10), afternoon(15:00~15:10), and sunset(10 minutes before and after sunset). Although we tried to keep these conditions as much as possible, there are cases in which data in near time period is inevitably collected or not collected at all due to the in-situ problems.

There are two notable events in the data collection period. One was that there was a period during which medium-sized hurricane 'Chan-thu' passed through the relevant sea area. Samples of sea state 8 conditions, which were not found in the well-conditioned data set, were obtained in normal conditioned data by the influence of the hurricane. The other is that when data was collected at night, there was usually indirect lighting by fishing boats, while there was no fishing during Korea's thanks giving day. Accordingly, in the case of a night photography, there was a period in which the light hardly enters.

Some images of normal conditioned are shown in **Fig. 5.7**. As the data acquisition periods are spread among a day, the brightness, saturation, diffuse reflection of light appearance, and so on, appear in various ways. In addition, since the movement of ships is also included, all data of specified time periods were utilized without abbreviated artificially selected sea photos.

Table 5.5. Non-prescreened data gathered at ADCP installation location.

Sep. 2021		Weather	Wind Dir.	Vw [m/s]	Have [m]	Hs [m]	Labeling	Remarks
01	Morning	Cloudy	S-SW	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Cloudy and occasional rain	S-SW	9 — 13	1.5 2.5 2.4 4.0	Sea state 4		
02	Morning	Cloudy	NW-N	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Cloudy and rain	NW-N	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
03	Morning	Cloudy and occasional rain	N-NE	8 — 13	1.0 2.5 1.6 4.0	Sea state 5		
	Afternoon	Cloudy	NE-E	9 — 14	1.5 3.0 2.4 4.8	Sea state 5		
04	Morning	Cloudy	NE-E	10 — 14	2.0 3.0 3.2 4.8	Sea state 5		
	Afternoon	Cloudy	NE-E	9 — 13	1.5 2.5 2.4 4.0	Sea state 5		
05	Morning	Cloudy	NE-E	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Cloudy and occasional rain	NE-E	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
06	Morning	Cloudy and occasional rain	E-SE	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Cloudy and occasional rain	SE-S	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
07	Morning	Cloudy and occasional rain	NW-N	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Cloudy	W-NW	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
08	Morning	Too cloudy	W-NW	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Sunny	NW-N	7 — 11	1.0 1.5 1.6 2.4	Sea state 4		
09	Morning	Too cloudy	NW-N	6 — 11	0.5 1.5 0.8 2.4	Sea state 4		
	Afternoon	Too cloudy	NW-N	6 — 9	0.5 1.0 0.8 1.6	Sea state 3		
10	Morning	Cloudy and rain	NW-N	6 — 11	0.5 1.5 0.8 2.4	Sea state 4		
	Afternoon	Cloudy and rain	NW-N	6 — 11	0.5 1.5 0.8 2.4	Sea state 4		
11	Morning	Cloudy and sometimes rain	E-SE	6 — 11	0.5 1.5 0.8 2.4	Sea state 4		
	Afternoon	Too cloudy	E-SE	6 — 11	0.5 1.5 0.8 2.4	Sea state 4		
12	Morning	Too cloudy	NE-E	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Cloudy and sometimes rain	E-SE	8 — 13	1.0 2.5 1.6 4.0	Sea state 5		
13	Morning	Cloudy and rain	E-SE	10 — 16	2.0 4.0 3.2 6.4	Sea state 6	Hurricane 'Chan-thu'	
	Afternoon	Cloudy and rain	E-SE	10 — 16	2.0 4.0 3.2 6.4	Sea state 6	Hurricane 'Chan-thu'	
14	Morning	Cloudy and rain	E-SE	12 — 18	2.0 5.0 3.2 8.0	Sea state 6	Hurricane 'Chan-thu'	
	Afternoon	Cloudy and rain	E-SE	12 — 18	2.0 5.0 3.2 8.0	Sea state 6	Hurricane 'Chan-thu'	
15	Morning	Cloudy and rain	E-SE	10 — 18	2.0 5.0 3.2 8.0	Sea state 6	Hurricane 'Chan-thu'	
	Afternoon	Cloudy and occasional rain	NE-E	10 — 18	2.0 5.0 3.2 8.0	Sea state 6	Hurricane 'Chan-thu'	
16	Morning	Cloudy and rain	E-SE	14 — 20	3.0 5.0 4.8 8.0	Sea state 7	No data	
	Afternoon	Cloudy and rain	E-SE	16 — 22	4.0 6.0 6.4 9.6	Sea state 7	No data	
17	Morning	Cloudy and rain	W-NW	18 — 26	5.0 8.0 8.0 12.8	Sea state 8	Hurricane 'Chan-thu'	
	Afternoon	Too cloudy	NW-N	10 — 16	2.0 4.0 3.2 6.4	Sea state 6	Hurricane 'Chan-thu'	
18	Morning	Cloudy	NW-N	9 — 13	1.5 2.5 2.4 4.0	Sea state 5	No fishing	
	Afternoon	Too cloudy	N-NE	9 — 13	1.5 2.5 2.4 4.0	Sea state 5		
19	Morning	Too cloudy	NE-E	7 — 11	0.5 1.5 0.8 2.4	Sea state 4	No fishing	
	Afternoon	Too cloudy	E-SE	7 — 11	0.5 1.5 0.8 2.4	Sea state 4		
20	Morning	Too cloudy	E-SE	7 — 11	0.5 1.5 0.8 2.4	Sea state 4	No fishing	
	Afternoon	Too cloudy	SE-S	8 — 12	1.0 2.0 1.6 3.2	Sea state 4		
21	Morning	Cloudy and rain	S-SW	9 — 13	1.5 2.5 2.4 4.0	Sea state 5	No fishing	
	Afternoon	Sunny	SW-W	8 — 13	1.0 2.5 1.6 4.0	Sea state 5		
22	Morning	Sunny	SW-W	8 — 12	1.0 2.0 1.6 3.2	Sea state 4	No fishing	
	Afternoon	Sunny	W-NW	7 — 11	1.0 1.5 1.6 2.4	Sea state 4		
23	Morning	Sunny	NW-N	7 — 11	1.0 1.5 1.6 2.4	Sea state 4	No fishing	
	Afternoon	Sunny	NW-N	6 — 9	0.5 1.0 0.8 1.6	Sea state 3		
24	Morning	Too cloudy	N-NE	6 — 9	0.5 1.0 0.8 1.6	Sea state 3		
	Afternoon	Too cloudy	NE-E	6 — 11	0.5 1.5 0.8 2.4	Sea state 4		
25	Morning	Too cloudy	NE-E	8 — 13	1.0 2.5 1.6 4.0	Sea state 5		
	Afternoon	Too cloudy	NE-E	8 — 13	1.0 2.5 1.6 4.0	Sea state 5		
26	Morning	Too cloudy	NE-E	10 — 14	2.0 3.0 3.2 4.8	Sea state 5		
	Afternoon	Too cloudy	NE-E	10 — 14	2.0 3.0 3.2 4.8	Sea state 5		
27	Morning	Too cloudy	NE-E	8 — 13	1.0 2.0 1.6 3.2	Sea state 4		
	Afternoon	Too cloudy	NE-E	8 — 13	1.0 2.5 1.6 4.0	Sea state 5		
28	Morning	Cloudy	E-SE	9 — 13	1.5 2.5 2.4 4.0	Sea state 5		
	Afternoon	Cloudy	SE-S	10 — 14	2.0 3.0 3.2 4.8	Sea state 5		
29	Morning	Cloudy and occasional rain	S-SW	10 — 14	2.0 3.0 3.2 4.8	Sea state 5		
	Afternoon	Cloudy and occasional rain	SE-S	10 — 14	2.0 3.0 3.2 4.8	Sea state 5		
30	Morning	Cloudy and occasional rain	N-NE	10 — 14	2.0 3.0 3.2 4.8	Sea state 5	No data at night, sunrise	
	Afternoon	Cloudy and occasional rain	N-NE	9 — 13	1.5 2.5 2.4 4.0	Sea state 5		

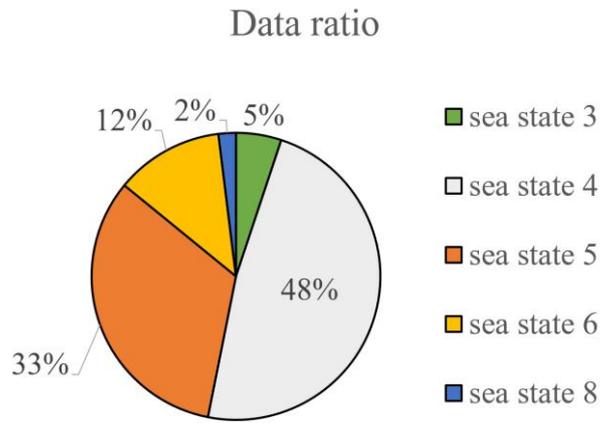


Fig. 5.6. Labelling ratio of normal conditioned data.

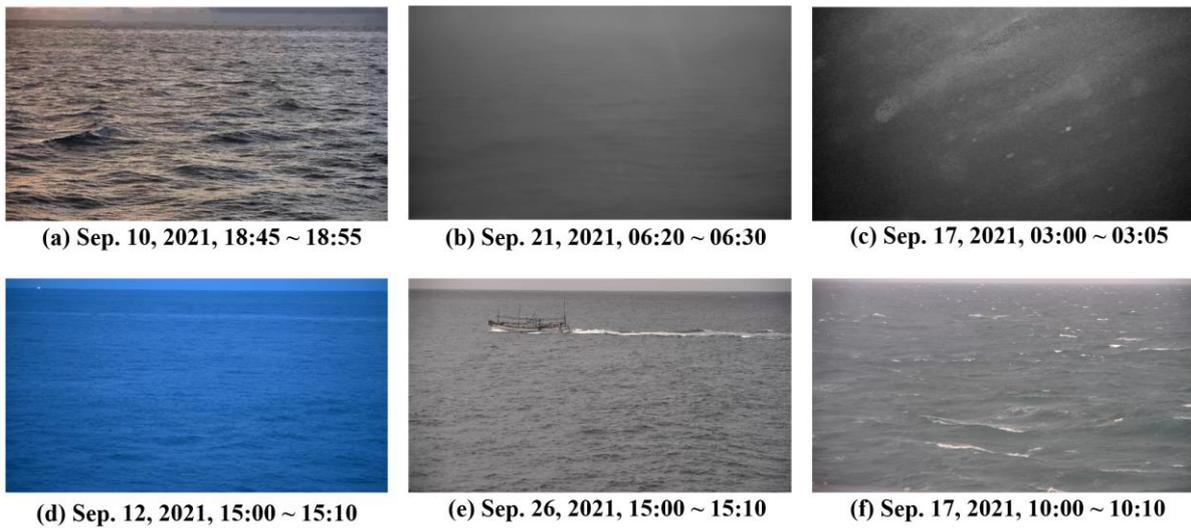


Fig. 5.7. Sample snapshot images of normal conditioned data.

Chapter 6. Training results with non-prescreened data

Learning was also performed on data including different time ranges that were not prescreened, listed in **Table 5.5**. In this chapter, there is no pre-processing of data such as artificial eliminating of images with obstacles, foggy weathers.

6.1 Initial training

Snapshot-based machine learning was performed with the described algorithm in **Chap. 3.1**. **Fig. 6.1** shows the global training accuracy and **Fig. 6.2** shows the cross entropy loss with data divided into five sea state categories. Data were augmented for the insufficient categories among the acquired data. The number of training data per category is 5000, the number of validation data per category is 300, and the number of testing data is 300. The computing time was 253m 25s.

In the case of the global accuracy, it is increased rapidly and reached the stable point after 20 epochs, even some fluctuations are occasionally occurred. In the end, an accuracy of 75.20% is obtained. Unlike previous learnings, it shows the phenomenon of suppressing accuracy at a certain value. In the case of the loss function, the tendency to decrease at the very beginning is the same as for other learnings. However, strangely, it was confirmed that the value of the loss function started to increase and stabilized around 3.0. To find out the cause of this phenomenon, predictions were made on the testing data.

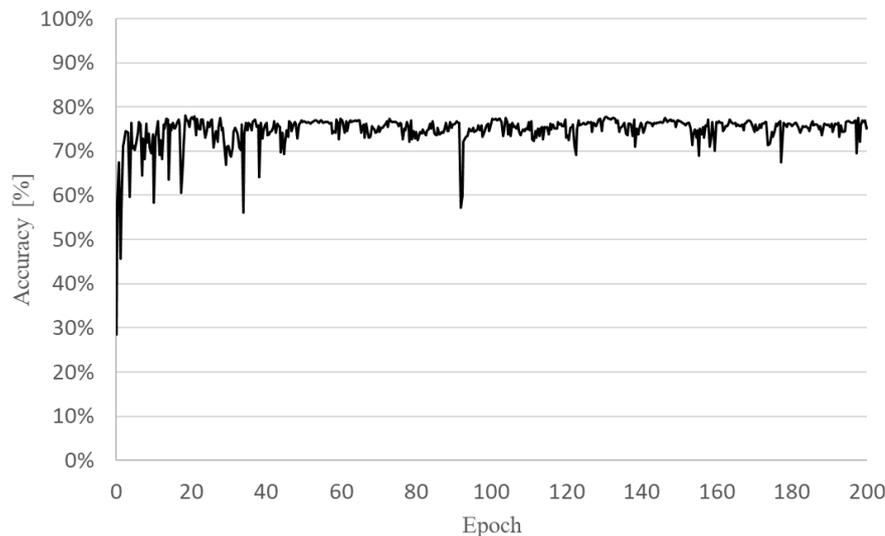


Fig. 6.1. Training results with normal data: Global accuracy.

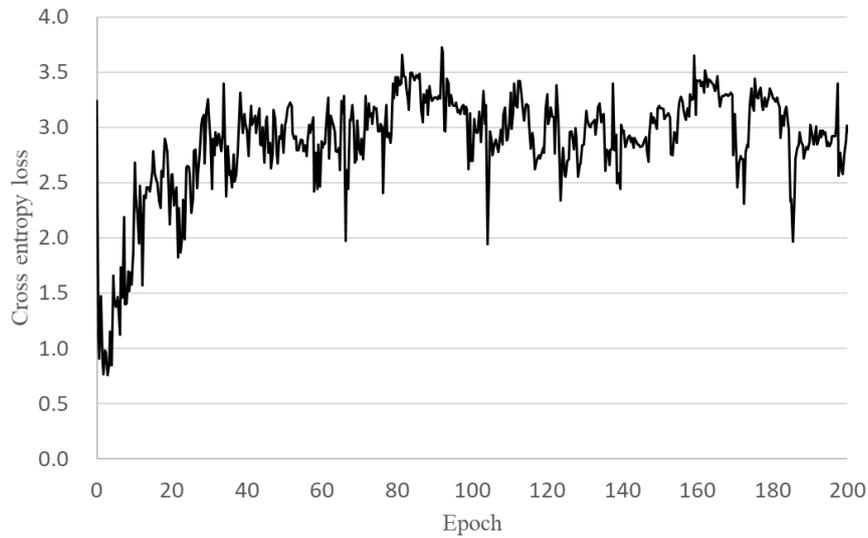


Fig. 6.2. Training results with normal data: Cross entropy loss.

The correlation matrix for the estimation accuracy in testing data is shown in **Table 6.1**. In the correlation table, predictions on sea state 3, 5, and 8 showed high accuracies. However, in the case of predictions for sea state 4, there were much more predictions with neighboring sea state 3 and 5 than for the correct answer. This confirms once again that it is difficult to accurately predict the dense categories among the low wave height categories as shown in the previous predictions. Predictions for sea state 6 have almost no predictions for correct answers, and predictions are made with sea state 4 and 8 mostly. All data of sea state 6 were obtained when the hurricane passed the target area, so it includes both the calm sea condition before the hurricane and the rough sea condition in the hurricane’s influence. For calculating the characteristic wave quantities, since it is derived after spectral analysis of the sea for one hour, it cannot properly represent the abrupt change of the sea surface within that time.

Table 6.1. Correlation matrix for sea state estimation with normal conditioned data.

	SS3 (predicted)	SS4 (predicted)	SS5 (predicted)	SS6 (predicted)	SS8 (predicted)
SS3 (observed)	99.996%	0.004%	0.000%	0.000%	0.000%
SS4 (observed)	36.953%	11.416%	51.607%	0.024%	0.000%
SS5 (observed)	0.661%	21.785%	76.239%	0.000%	1.315%
SS6 (observed)	0.000%	34.719%	4.066%	0.016%	61.198%
SS8 (observed)	0.000%	0.000%	0.000%	0.001%	99.999%

6.2 Applicability evaluation with deformed images

Until now, learnings have been carried out on images acquired from a fixed angle of view. However, it is practically difficult to obtain the testing data from the identical condition (angle of view, distance, lightness, and so on) as the network acquisition, and in particular, it should be confirmed whether the technology can be applied to a continuously moving vessel. In this study, three modes of deformation (focusing, rotating, and tilting) are considered, described in Eqs. 6.1 to 6.3.

$$\text{Size} = [656; 875; 3] \quad (6.1)$$

$$\text{n\% focusing} = [\text{n\% of range; n\% of range; 3}] = 0$$

$$R = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.2)$$

$$Tt = \begin{bmatrix} 1 & \alpha & 0 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.3)$$

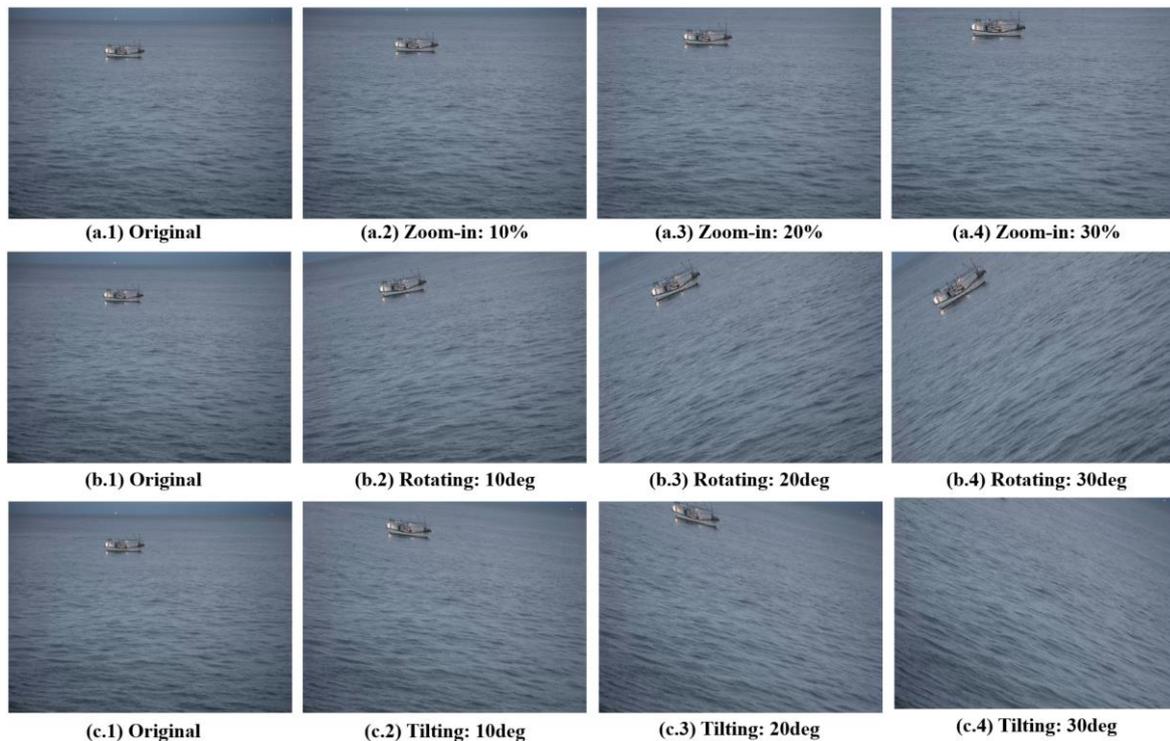


Fig. 6.3. Image distortion: focusing, rotating, and tilting.

The images obtained through the focusing, rotating, and tilting are shown in **Fig. 6.3**. In the case of rotating and tilting, pixel empty spaces are created at the edges after image change, and since only the data necessary for learning is cut out as a rectangular frame, some focusing effects are also included. These image deformations are considered as the camera's forward and backward motion, roll motion, and coupled rotational motion, so it can be seen that the basic motion mode of a ship is indirectly represented.

The prediction results are summarized in **Table 6.2-4**. The results excludes results for sea state 4 and 6, which were not well predicted. As checking the accuracy change with image focusing, the accuracy does not decrease significantly at 10% zoom-in. However, as the focusing increases, the accuracy decreases and the decrease in accuracy is much greater at small wave heights. This phenomenon also appears in rotating and tilting, but in these deformation modes, it can be seen that the accuracy reduction is quite large even at 10deg or 10% deformation.

Table 6.2. Prediction results with image focusing.

	Original	10% zoom-in	20% zoom-in	30% zoom-in
SS3	1.000	0.861	0.284	0.166
SS5	0.762	0.674	0.573	0.481
SS8	1.000	0.998	0.880	0.651

Table 6.3. Prediction results with image rotating.

	Original	10deg rotation	20deg rotation	30deg rotation
SS3	1.000	0.327	0.248	0.126
SS5	0.762	0.341	0.229	0.184
SS8	1.000	0.550	0.493	0.479

Table 6.4. Prediction results with image tilting.

	Original	10% tilting	20% tilting	30% tilting
SS3	1.000	0.099	0.009	0.000
SS5	0.762	0.497	0.382	0.286
SS8	1.000	0.740	0.674	0.483

In order to broaden the effective range of such image estimation performance, the following studies are considered. One way is to construct a network by increasing the total amount of data through data augmentation of the first image sets. The latter is about rotational motion, and it is a method to learn by synchronizing angles of training data and testing data by calculating main vector directions by converting the shape of the wave crest into a vector on the image.

By linking the measuring instruments, the inclined angle can be easily known and images can be compensated to the reference angle. However, it is not easy to specify an angle from only actual images without a baseline or

given information. Here, an attempt for vector extraction from the image was conducted. After knowing the inclined angle, images were then rotated again to match the training network frame. In this case, the generated black space was eliminated.

The concept of this re-rotation is illustrated in **Fig. 6.4**. The rotation images in which the results are summarized in **Table 6.3** are for the red square region in the upper center of **Fig. 6.4**.

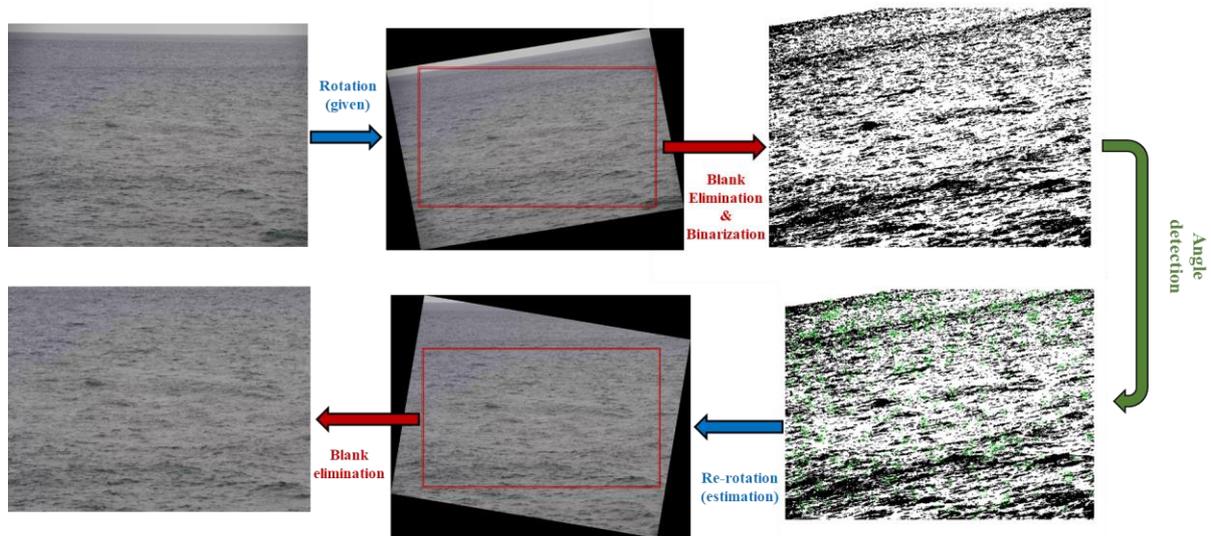


Fig. 6.4. Image re-rotation process.

Many studies of extracting the vector for the gradient change from an image have been conducted in vision engineering. Here, a total of seven techniques were applied to estimate the inclined angle in the photograph; BRISK [54], FAST [55], Harris [56], KAZE [57], Minimum Eigenvalue [58], ORB [59], and SURF [60]. To make the input information value more concise, the angle was estimated from the black & white converted image.

An example of applying seven techniques to one image is shown in **Fig. 6.5**. Here, the green squares highlight what were found as representative vectors in each technique. When the FAST method is applied, it can be seen that the vectors are mostly found in the upper part of the image. While the vectors, which are found by other methods, are scattered throughout the whole image.

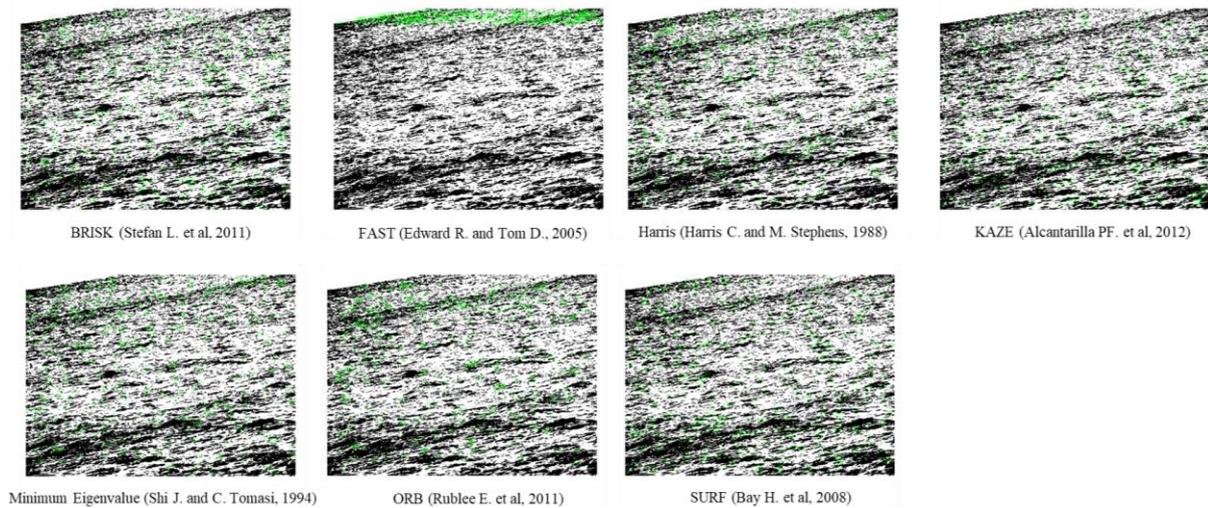


Fig. 6.5. Various vector extraction methods from an image.

Estimation of the angle for the sea state 3 data was conducted with these techniques. Except for the Harris method and Minimum Eigenvalue method, others did not generally converge to a constant value. Accordingly, the angle estimation for the 10degrees rotated sea state 3 images of two models(Harris and Minimum Eigenvalue) is illustrated in **Fig. 6.6**. Relatively classical methods estimated a relatively constant values. In particular, it was confirmed that the Harris model estimated close to 10 degrees except for night time images.

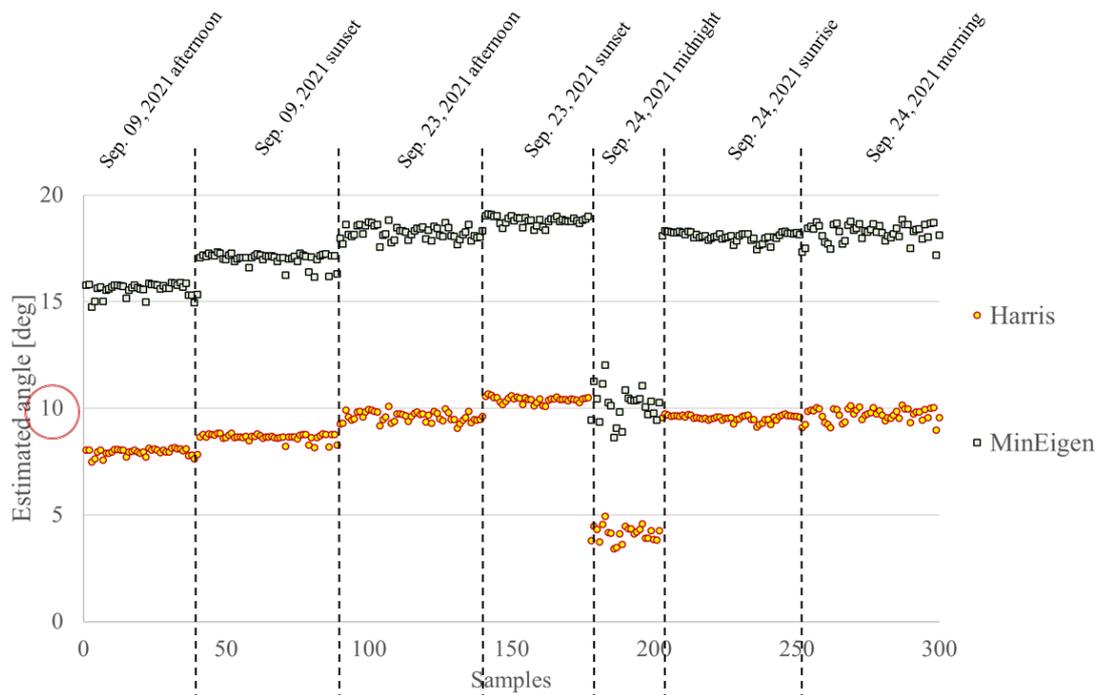


Fig. 6.6. An example of angle detection from an image: Harris model [56] & Minimum Eigenvalue model [58].

The angle estimation for the entire rotated data using the Harris model is shown in **Fig. 6.7-9**. Generally, although it shows a deviation of about 5 degrees, it is estimated well as the given angles.

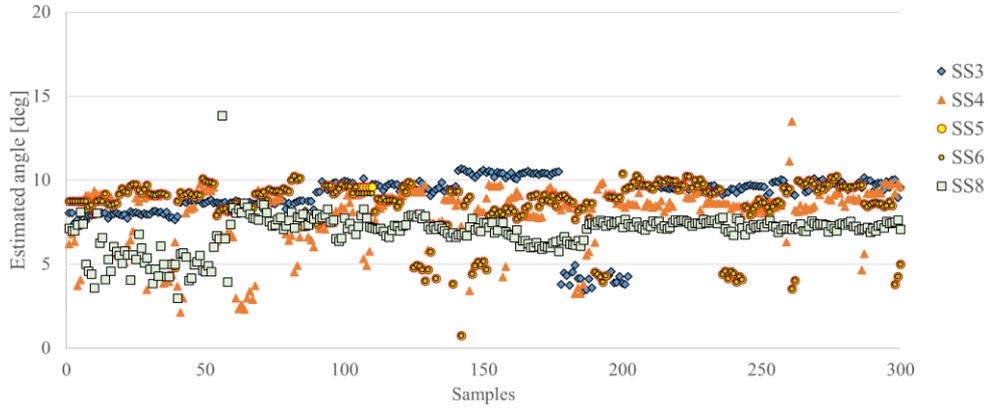


Fig. 6.7. Estimated angle with Harris method with 10deg rotated images.

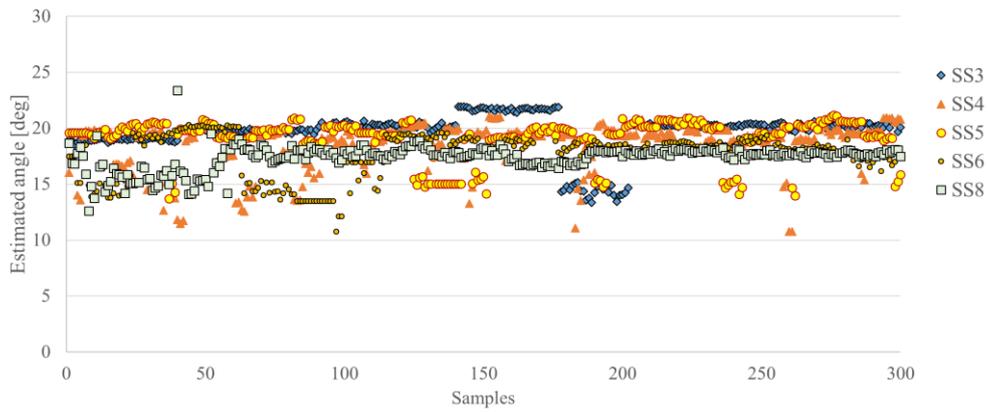


Fig. 6.8. Estimated angle with Harris method with 20deg rotated images.

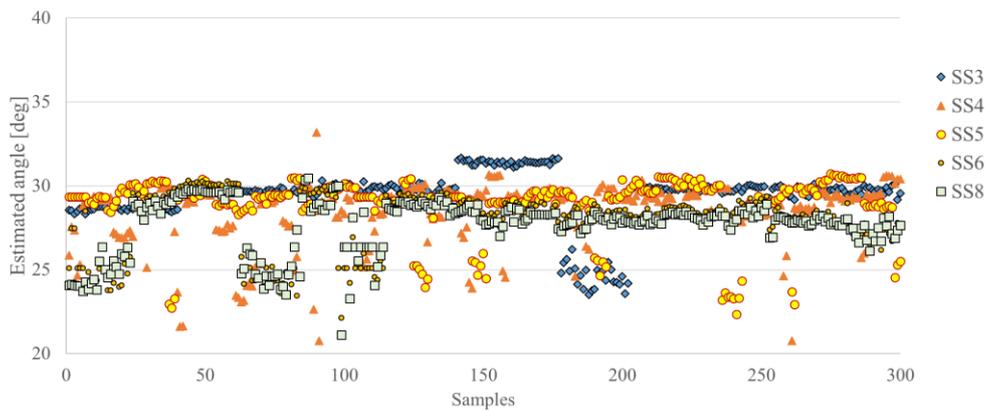


Fig. 6.9. Estimated angle with Harris method with 30deg rotated images.

Prediction was performed lastly on the images rotated back to the estimated angle. The prediction accuracy for re-rotated images is shown in **Fig. 6.10** along with the results for the initially rotated images in **Table 6.3**. In the image rotated by 10 degrees in sea state 3, the accuracy is the mostly increased. Accuracy is also improved for images rotated 20 degrees in sea state 3, but the deviation is not large. For the image rotated by 30 degrees in sea state 3, there was almost no effect from re-rotation. In sea state 5, there was an insignificant increase by each angle. In the sea state 8 condition, there was a slight increase in small angle changes, while the accuracy decreased in the image rotated by 30 degrees. As the applied vector detecting algorithm is simple and the slopes of representative vectors are averaged, the precision of angle estimation is partially lacking. In addition, focusing was inevitably accompanied to obtain a rectangular frame excluding blank space while applying two sequential rotations.

A simple attempt was made to solve this problem, but it was difficult to generalize. Therefore, we recognized the need to construct a more comprehensive network by using precision measuring instruments for angle detection and augmenting data in various modes such as rotation, focusing, and so on.

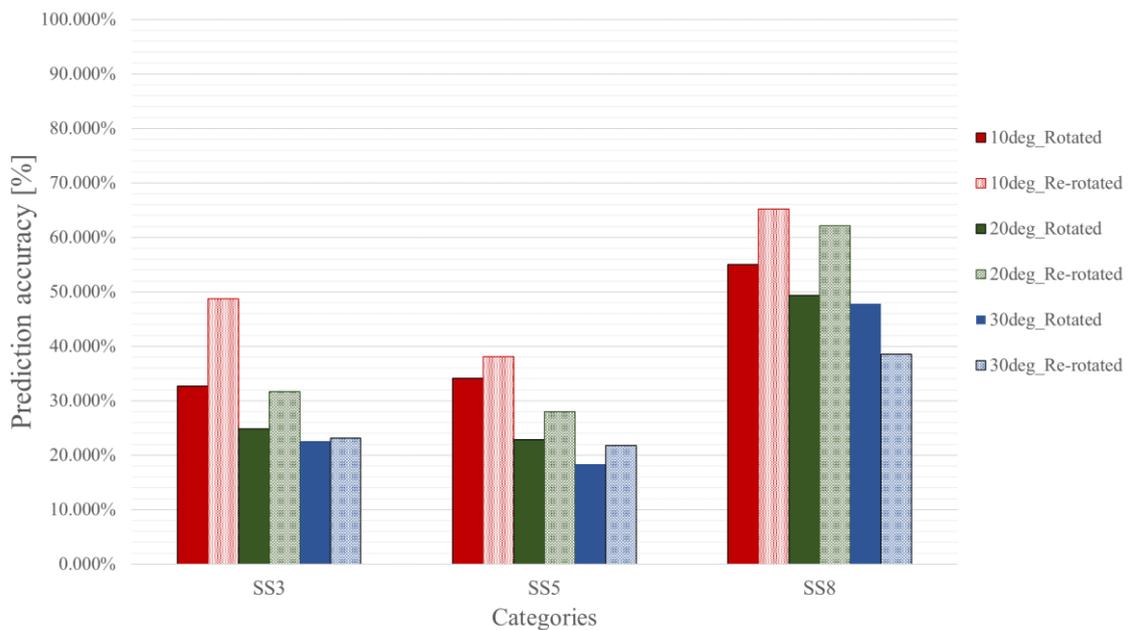


Fig. 6.10. Comparison between predictions with rotated and re-rotated images.

Chapter 7. Training results with prescreened data

The well-conditioned data(May~July, 2020), which is listed in **Table 5.2**, were trained, validated, and tested through a deep learning procedure to estimate sea states and average wave heights.

7.1 Single snapshot-based machine learning results

In order to evaluate this learning model in statistics, ‘K-fold cross validation’ was adopted. The problem with evaluating model performance using a part of the entire data as a validation set(called held-out validation), is that if the size of the dataset is small, the reliability of performance evaluation on the test set decreases. If the performance is different depending on how you hold the test set, the model evaluation index will be biased due to the effect of chance.

Fig. 7.1 shows the data set separation as training(including validating) and testing data. The number of training, validating, and testing data is 11000, 1000, and 3000, respectively. After dividing the whole data set into five folds, and data was utilized at least once for testing. This allows us to predict performance on unseen data, choose a better model, and tune the hyper-parameters.



Fig. 7.1. 5-fold cross validation for snapshot-based machine learning.

We initially predict the sea state with machine learning methodology, which is written in **Chap. 3.1**. **Fig. 7.2** shows the training results of deep learning in the sea state classification with snapshot images as the global training accuracy and **Fig.7.3** shows the corresponding cross-entropy loss. The training time was about 1 hour and 35 minutes. The final accuracies in five folds were from 73.0% to 90.17%. In the case of accuracy, the learning result tends to be maintained with a fluctuation range of about 20% after a rapid increase in the initial stage. The corresponding loss function also varies, but ultimately has a certain amount of loss until the end of learning.

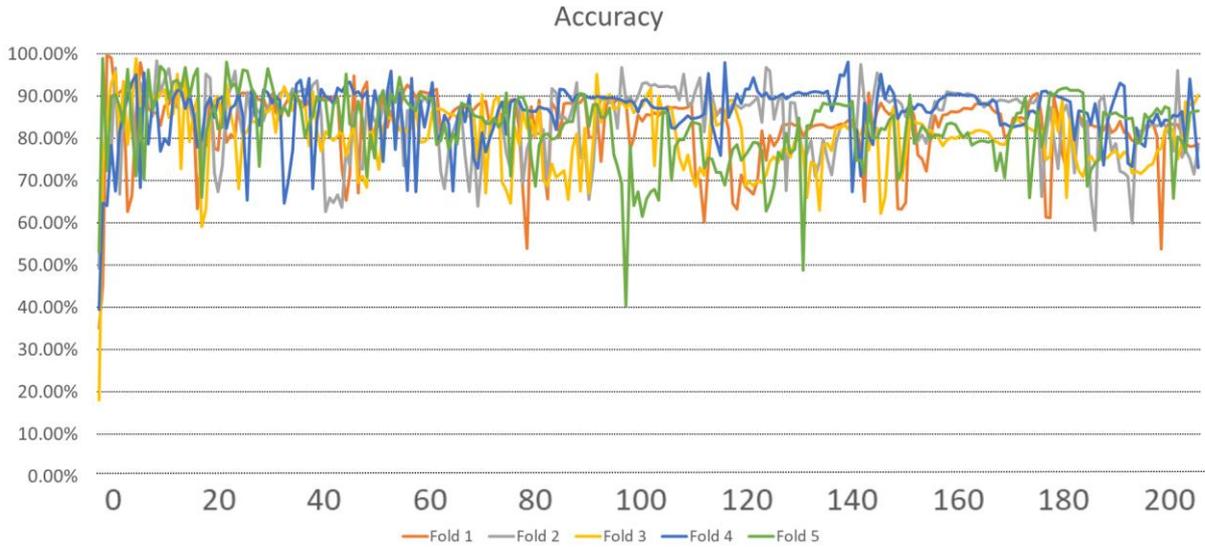


Fig. 7.2. Training results for the sea state estimation: Global training accuracy of validating data.

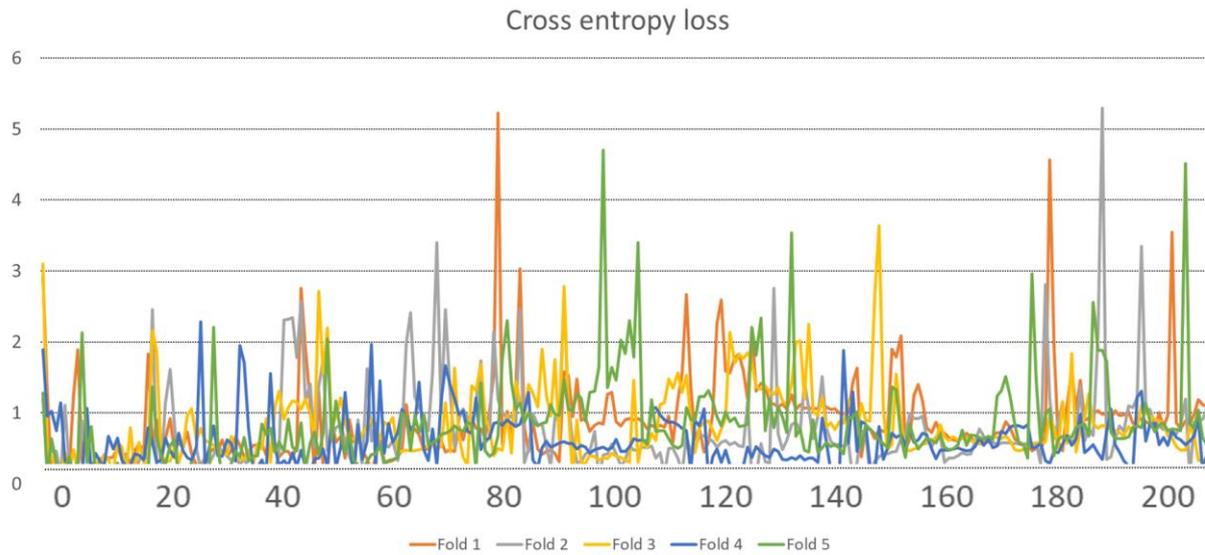


Fig. 7.3. Training results for the sea state estimation: Cross entropy loss of validating data.

Table 7.1 lists the prediction probabilities of testing data using the trained network. It can be seen that the group corresponding to sea state 6 predicts the correct answer fairly accurately. It was confirmed that while sea state 4 was predicted relatively accurately, the prediction of sea state 5 was failed and the prediction results was concluded with sea state 4. As can be seen in **Fig. 5.4**, it is difficult to classify sea state 4 and 5 even with the naked eye. Therefore, it is thought that simple snapshot-based machine learning does not perform well in cases where wave height changes are insignificant, such as sea state 4 and 5.

Table 7.1. Prediction probability of testing data for the sea state.

		Sea state 4 (observed)	Sea state 5 (observed)	Sea state 6 (observed)
Sea state 4 (predicted)	fold 1	0.989	0.011	0.000
	fold 2	0.715	0.285	0.000
	fold 3	1.000	0.000	0.000
	fold 4	0.997	0.003	0.000
	fold 5	1.000	0.000	0.000
Sea state 5 (predicted)	fold 1	0.998	0.002	0.000
	fold 2	0.734	0.266	0.000
	fold 3	0.533	0.467	0.000
	fold 4	0.648	0.352	0.000
	fold 5	0.949	0.051	0.000
Sea state 6 (predicted)	fold 1	0.000	0.000	1.000
	fold 2	0.000	0.000	1.000
	fold 3	0.000	0.000	1.000
	fold 4	0.000	0.000	1.000
	fold 5	0.000	0.001	0.999

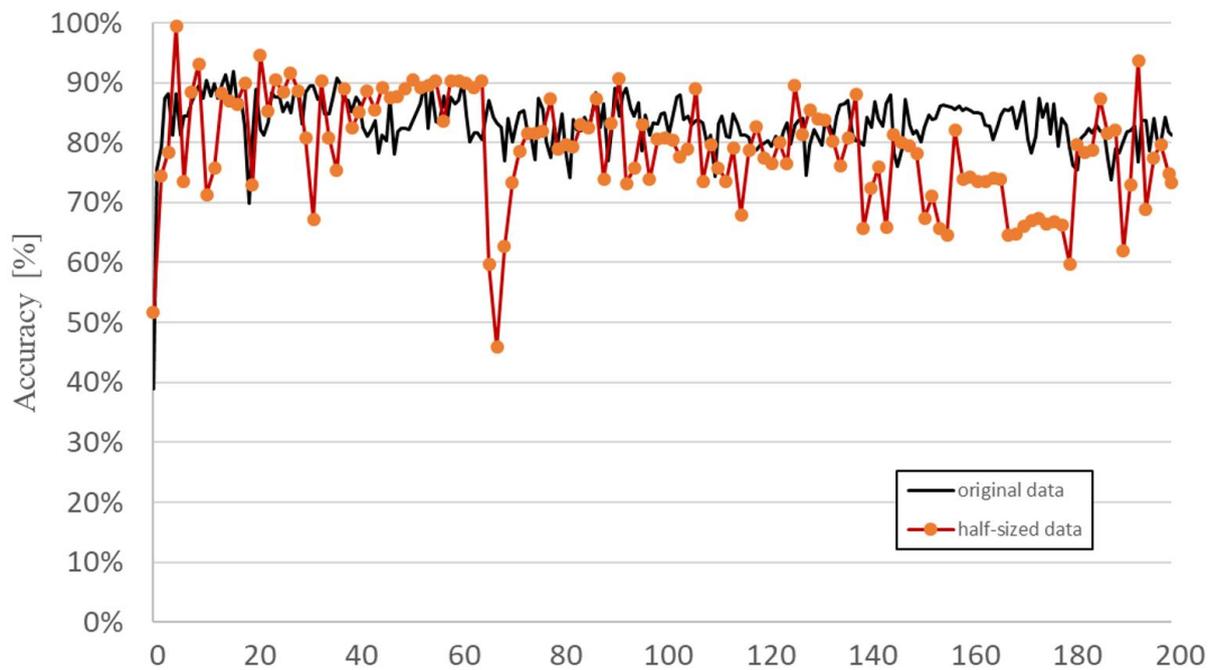


Fig. 7.4. Training results for the sea state estimation: Global training accuracy of validating data.

In order to check the dependence on the number of data, the same training was performed with only the data quantity reduced by half, which is shown in **Fig. 7.4**. The black line represent the prediction accuracy with original data, the red line with circles represent the prediction accuracy with half-sized data. All learning from now on are the results of 5-fold cross validation. Although the range of variation is large in the result of half-sized data, both final accuracies are similar.

However, the predictions for the testing data came out differently. **Table 7.2** describes the probability of testing data for the sea state prediction using half of the data. When the correct answer group was sea state 4, the prediction accuracy was somewhat lower, and at this time, many predictions were made with sea state 5. Many of the predictions in sea state 5 are similar to those in sea state 4, but the amount is reduced. It seems that the amount of data that can determine the distinct characteristics between sea state 4 and 5 is not secured. Finally the accuracy in sea state 6 also decreased significantly.

Table 7.2. Probability of testing data for the sea state prediction using half of the data.

	Sea state 4 (observed)	Sea state 5 (observed)	Sea state 6 (observed)
Sea state 4 (predicted)	69.171%	30.829%	0.000%
Sea state 5 (predicted)	57.489%	42.511%	0.000%
Sea state 6 (predicted)	0.012%	28.883%	71.105%

Using the photo snapshots, we again trained a network in order to classify the average wave heights in **Table 5.4**. The global training accuracy for the wave height classification with snapshots is shown in **Fig. 7.5** and the corresponding loss function is shown in **Fig. 7.6**. It can be seen that the accuracy increases steadily until the second epoch, and then converges near a certain value. The tendency of the loss function is similar to that in sea state classification.

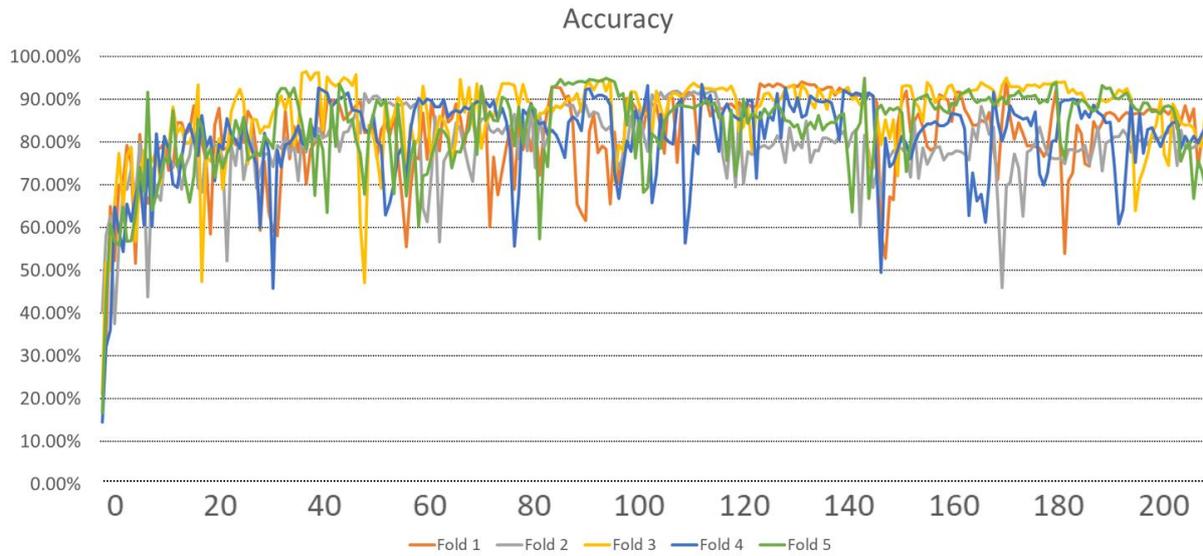


Fig. 7.5. Training results for the wave height prediction: Global training accuracy of validating data.

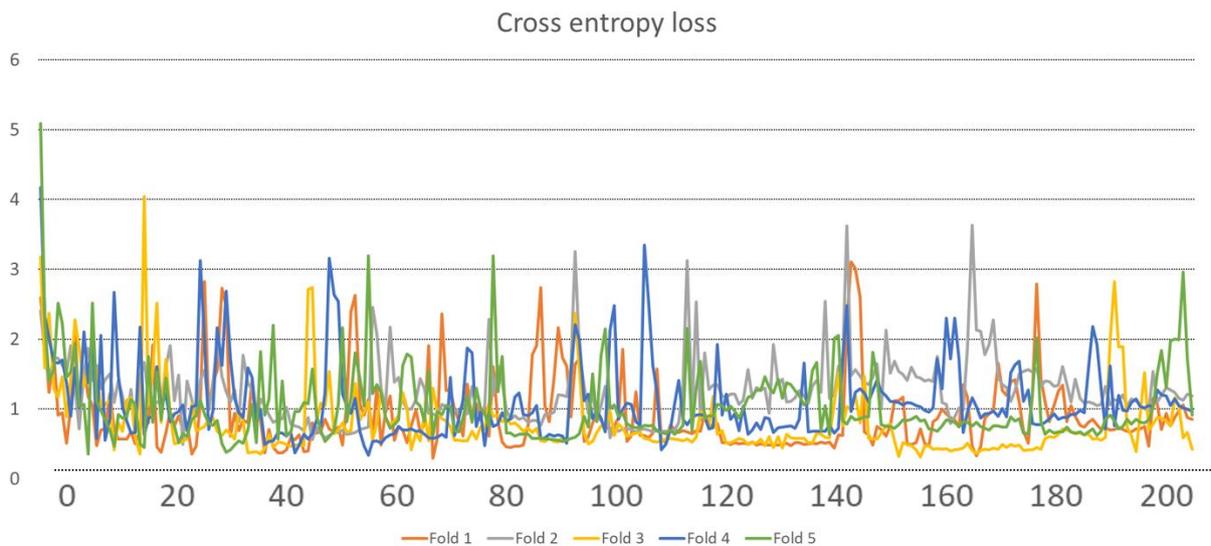


Fig. 7.6. Training results for the wave height prediction: Cross entropy loss of validating data.

The correlation matrix in average wave height prediction for testing data is described in **Table 7.3**. The overall prediction performance is generally ill in the range of $H_{ave}=1.0\sim 2.0\text{m}$. In this range, the prediction shows that it is focused at 1.25m and 1.50m. However, at relatively high wave heights of 2.5m and 3.0m, the prediction accuracy is quite high. As seen in the sea state classification, this section is characterized with low wave height and no significant features, such as wave breaking at crests, which is difficult for experts to judge.

Table 7.3. Prediction probability of testing data for the average wave height.

	H _{ave} =1.00m (predicted)	H _{ave} =1.25m (predicted)	H _{ave} =1.50m (predicted)	H _{ave} =2.00m (predicted)	H _{ave} =2.50m (predicted)	H _{ave} =3.00m (predicted)
H _{ave} =1.00m (observed)	0.04%	3.73%	96.23%	0.00%	0.00%	0.00%
H _{ave} =1.25m (observed)	0.88%	77.64%	4.15%	17.33%	0.00%	0.00%
H _{ave} =1.50m (observed)	14.11%	0.00%	65.89%	15.34%	4.50%	0.15%
H _{ave} =2.00m (observed)	0.00%	0.29%	55.40%	44.30%	0.02%	0.00%
H _{ave} =2.50m (observed)	0.00%	0.00%	1.02%	0.00%	98.98%	0.00%
H _{ave} =3.00m (observed)	0.00%	0.00%	0.00%	0.00%	0.00%	100.00%

In order to more closely evaluate the classification performance, the confusion matrix for average wave height classification with snapshots is sum up as **Table 7.4**. There are four values; (1) True Positive(TP): Predict the correct answer as true(correct answer) = 64.474%, (2) False Positive(FP): Predicting a true false answer as true(wrong answer) = 30.741%, (3) False Negative(FN): Predicting a true answer as false(wrong answer) = 35.525%, and (4) True Negative(TN): Predicting a correct answer that is actually false as false(correct answer) = 69.259%.

Table 7.4. Confusion matrix for average wave height classification with snapshots.

		Actual answer	
		TRUE	FALSE
Classification result	TRUE	64.474%	30.741%
	FALSE	35.525%	69.259%

We evaluated the wave height classification performance in terms of precision, recall, and accuracy. Precision is the proportion of what the model classifies as true, which is actually true, can be expressed as **Eq. 7.1**. Recall is the ratio of what the model predicts to be true out of what is actually true, can be expressed as **Eq. 7.2**. Accuracy is the most intuitive way to describe the performance of a model, and it is a metric that takes into account both the prediction of True values as True and the prediction of False value as False, is written in **Eq. 7.3**.

$$(\textit{precision}) = \frac{TP}{TP + FP} \quad (7.1)$$

$$(\textit{Recall}) = \frac{TP}{TP + FN} \quad (7.2)$$

$$(\textit{Accuracy}) = \frac{TP + TN}{TP + FN + FP + TN} \quad (7.3)$$

In this problem, precision, recall, and accuracy are 67.714%, 64.475%, and 66.867%, respectively. Since the training was carried out after setting the same number of data per each category, there is no significant difference between indicators. The poor learning results in the lower average wave height range are the primary reason that the overall accuracy is less than 70%.

7.2 Video-based machine learning results

In the case of ocean waves, although classified as narrow-banded due to their high inertia, it is still characterized as spectrum. Therefore, in the case of a single snapshot, high waves can be detected even in calm ocean conditions, and vice versa. In the single snapshot-based machine learning, the learning result for low wave height was not good, so another approach was performed to overcome this problem. An attempt was made to enable machine learning of the spectrum by inserting a continuous sequence rather than a single photo as input.

Table 7.5 describes the results of converting the current photo data into video clips. In the case of video length, a total of 5 cases were encoded in 6, 30, 60, 180, 300 s. Among the total number of data, the data of the low average wave height categories, from 1.0m to 2.0m, accounted for a large proportion, and the data corresponding to 2.5 and 3.0m were relatively few. In the beginning, this data bias was not taken into account, so the video was created by dividing the data in batches without augmentation. The percentage of testing data per each category was 10%, and 10-fold cross validation was applied to increase the reliability of the learning results.

Machine learning was performed again based on the prepared video clips. CNN was applied to generate continuous sequences while extracting features, and LSTM was applied to the subsequent classification problem. It is described in **Chap. 3.3**.

Table 7.5. Results of converting current photo data into video clips.

		# of training data		# of validating data		# of testing data	
Movie clip I (Length=6s)	H _{ave} =1.00m	2596	81.00%	288	9.00%	320	10.00%
	H _{ave} =1.25m	3229	81.00%	359	9.00%	398	10.00%
	H _{ave} =1.50m	6431	81.00%	715	9.00%	794	10.00%
	H _{ave} =2.00m	2320	81.00%	258	9.00%	286	10.00%
	H _{ave} =2.50m	485	81.10%	54	9.00%	59	9.90%
	H _{ave} =3.00m	1404	81.00%	156	9.00%	173	10.00%
Movie clip II (Length=30s)	H _{ave} =1.00m	518	81.00%	58	9.00%	64	10.00%
	H _{ave} =1.25m	646	81.10%	72	9.00%	79	9.90%
	H _{ave} =1.50m	1287	81.00%	143	9.00%	158	9.90%
	H _{ave} =2.00m	464	81.00%	52	9.00%	57	10.00%
	H _{ave} =2.50m	97	81.70%	11	9.10%	11	9.20%
	H _{ave} =3.00m	281	81.20%	31	9.00%	34	9.80%
Movie clip III (Length=60s)	H _{ave} =1.00m	259	81.00%	29	9.00%	32	10.00%
	H _{ave} =1.25m	323	81.20%	36	9.00%	39	9.80%
	H _{ave} =1.50m	644	81.00%	72	9.00%	79	9.90%
	H _{ave} =2.00m	232	81.20%	26	9.00%	28	9.80%
	H _{ave} =2.50m	49	82.40%	5	9.20%	5	8.50%
	H _{ave} =3.00m	140	81.20%	16	9.00%	17	9.80%
Movie clip IV (Length=180s)	H _{ave} =1.00m	86	81.50%	10	9.10%	10	9.40%
	H _{ave} =1.25m	107	81.10%	12	9.00%	13	9.80%
	H _{ave} =1.50m	214	81.10%	24	9.00%	26	9.80%
	H _{ave} =2.00m	77	81.50%	9	9.10%	9	9.50%
	H _{ave} =2.50m	16	85.30%	2	9.50%	1	5.30%
	H _{ave} =3.00m	47	82.10%	5	9.10%	5	8.80%
Movie clip V (Length=300s)	H _{ave} =1.00m	52	81.60%	6	9.10%	6	9.40%
	H _{ave} =1.25m	64	80.90%	7	9.00%	8	10.10%
	H _{ave} =1.50m	129	81.50%	14	9.10%	15	9.50%
	H _{ave} =2.00m	47	82.10%	5	9.10%	5	8.80%
	H _{ave} =2.50m	8	73.60%	1	8.20%	2	18.20%
	H _{ave} =3.00m	27	79.40%	3	8.80%	4	11.80%

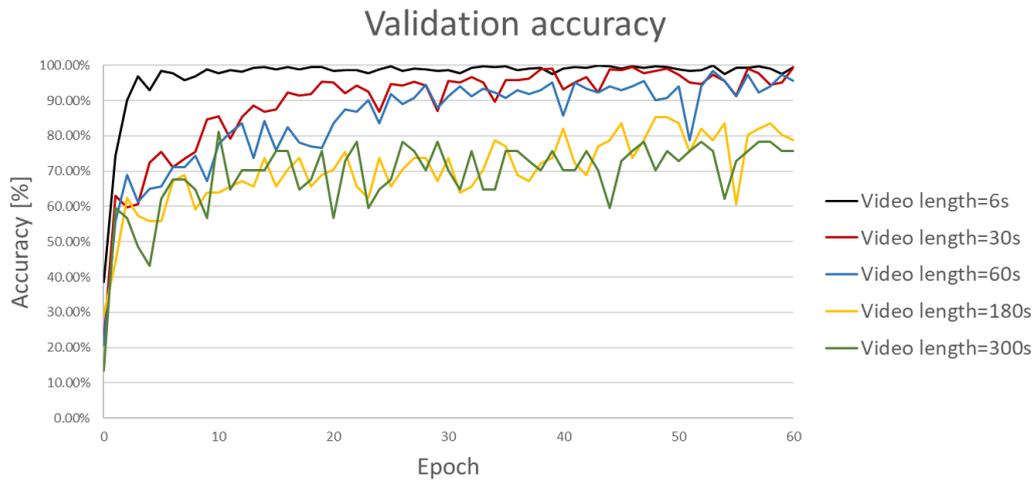


Fig. 7.7. Training results for the H_{ave} prediction with different lengths of videos: Global accuracy.

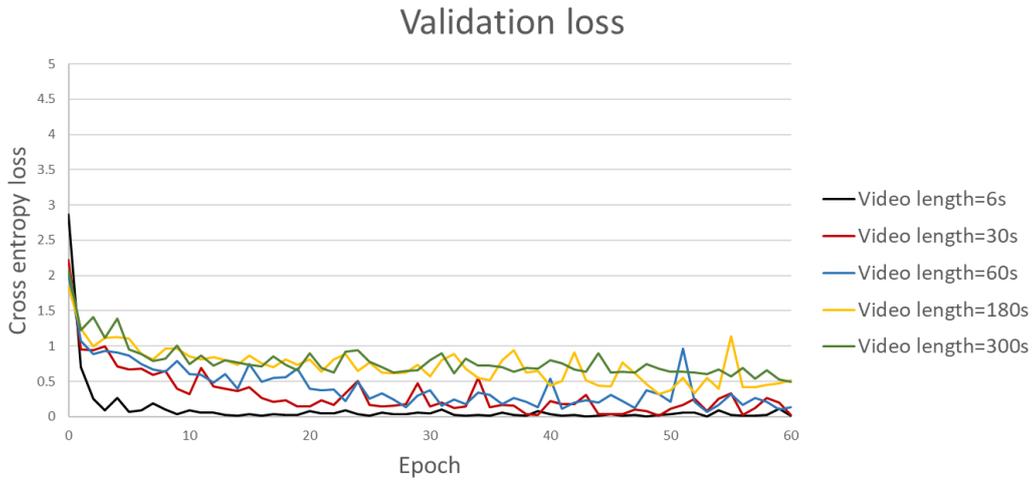


Fig. 7.8. Training results for the H_{ave} prediction with different lengths of videos: Cross entropy loss.

Fig. 7.7 shows the training results in terms of global accuracy for the average wave height prediction with different lengths of video clips, and **Fig. 7.8** shows the corresponding cross entropy loss. In the case of video-based learning with the length of 6s, which has a largest number of data, the global learning rate increases significantly in the initial stage. It can be seen that the results of learning with 30s and 60s videos show a similar trends, which are slowly increased up to the end of the training. The final results in validating data show a high accuracy of over 90% in learnings with 6s, 30s, and 90s videos. However, in the case of trainings with longer videos(180s, 300s video based learnings), the results are similarly increased up to the first 10 epochs, but after that, the increase almost disappears. The accuracies are finally calculated in the 70-80% section. This trend is also directly inverse proportion to the graph of the loss function.

The prediction accuracy for testing data in video based learnings is then summarized as correlation matrices. In **Table 7.6**, it was confirmed that most of predictions in 1.0m and 1.25m categories were made at 1.50m. This phenomenon was also observed in 30s video based learning, described in **Table 7.7**, lots of data predicted to be 1.50m from the 1.25m and 2.0m categories. Also, the bias of this prediction was found to be the same in 60s video based learning, as seen in **Table 7.8**.

Table 7.6. Correlation matrix for the H_{ave} prediction with videos(Length=6s).

	H _{ave} 1.00m (predicted)	H _{ave} 1.25m (predicted)	H _{ave} 1.50m (predicted)	H _{ave} 2.00m (predicted)	H _{ave} 2.50m (predicted)	H _{ave} 3.00m (predicted)
H _{ave} 1.00m (observed)	21.01%	0.80%	78.10%	0.08%	0.00%	0.00%
H _{ave} 1.25m (observed)	0.00%	3.62%	95.84%	0.53%	0.00%	0.00%
H _{ave} 1.50m (observed)	0.83%	3.85%	75.62%	17.35%	2.24%	0.11%
H _{ave} 2.00m (observed)	0.00%	0.01%	4.75%	95.23%	0.01%	0.00%
H _{ave} 2.50m (observed)	0.00%	0.00%	0.03%	0.00%	99.96%	0.00%
H _{ave} 3.00m (observed)	0.00%	0.00%	2.02%	0.00%	0.00%	97.97%

Table 7.7. Correlation matrix for the H_{ave} prediction with videos(Length=30s).

	H _{ave} 1.00m (predicted)	H _{ave} 1.25m (predicted)	H _{ave} 1.50m (predicted)	H _{ave} 2.00m (predicted)	H _{ave} 2.50m (predicted)	H _{ave} 3.00m (predicted)
H _{ave} 1.00m (observed)	1.62%	79.39%	18.98%	0.01%	0.00%	0.00%
H _{ave} 1.25m (observed)	0.00%	1.32%	98.66%	0.01%	0.00%	0.00%
H _{ave} 1.50m (observed)	0.01%	3.68%	88.26%	1.02%	7.01%	0.01%
H _{ave} 2.00m (observed)	0.00%	0.19%	70.38%	26.52%	2.90%	0.00%
H _{ave} 2.50m (observed)	0.00%	0.16%	0.47%	0.01%	99.35%	0.01%
H _{ave} 3.00m (observed)	0.00%	0.00%	0.01%	0.00%	0.00%	99.99%

Table 7.8. Correlation matrix for the H_{ave} prediction with videos(Length=60s).

	H _{ave} 1.00m (predicted)	H _{ave} 1.25m (predicted)	H _{ave} 1.50m (predicted)	H _{ave} 2.00m (predicted)	H _{ave} 2.50m (predicted)	H _{ave} 3.00m (predicted)
H _{ave} 1.00m (observed)	0.95%	93.46%	5.49%	0.05%	0.05%	0.01%
H _{ave} 1.25m (observed)	0.00%	0.31%	99.55%	0.12%	0.01%	0.00%
H _{ave} 1.50m (observed)	0.16%	20.01%	79.44%	0.31%	0.05%	0.03%
H _{ave} 2.00m (observed)	0.01%	1.03%	31.82%	66.91%	0.22%	0.00%
H _{ave} 2.50m (observed)	0.00%	0.01%	57.27%	0.04%	42.65%	0.03%
H _{ave} 3.00m (observed)	0.00%	0.00%	0.02%	0.00%	0.00%	99.98%

Table 7.9. Correlation matrix for the H_{ave} prediction with videos(Length=180s).

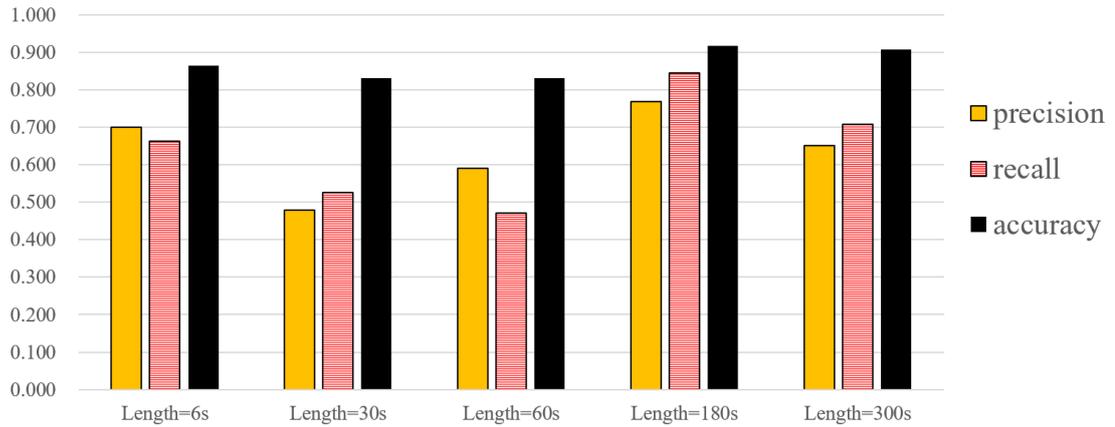
	H _{ave} 1.00m (predicted)	H _{ave} 1.25m (predicted)	H _{ave} 1.50m (predicted)	H _{ave} 2.00m (predicted)	H _{ave} 2.50m (predicted)	H _{ave} 3.00m (predicted)
H _{ave} 1.00m (observed)	77.69%	14.34%	0.88%	7.03%	0.06%	0.00%
H _{ave} 1.25m (observed)	1.50%	75.33%	3.43%	18.71%	1.03%	0.00%
H _{ave} 1.50m (observed)	0.45%	15.62%	58.54%	23.00%	2.12%	0.27%
H _{ave} 2.00m (observed)	0.19%	1.59%	1.86%	84.67%	11.69%	0.00%
H _{ave} 2.50m (observed)	0.00%	0.00%	0.07%	0.08%	99.84%	0.00%
H _{ave} 3.00m (observed)	0.00%	0.00%	0.11%	0.00%	0.00%	99.89%

Table 7.10. Correlation matrix for the H_{ave} prediction with videos(Length=300s).

	H _{ave} 1.00m (predicted)	H _{ave} 1.25m (predicted)	H _{ave} 1.50m (predicted)	H _{ave} 2.00m (predicted)	H _{ave} 2.50m (predicted)	H _{ave} 3.00m (predicted)
H _{ave} 1.00m (observed)	73.80%	18.18%	6.76%	0.43%	0.81%	0.02%
H _{ave} 1.25m (observed)	0.91%	60.63%	29.74%	8.34%	0.36%	0.01%
H _{ave} 1.50m (observed)	1.00%	20.02%	66.35%	7.90%	1.13%	3.60%
H _{ave} 2.00m (observed)	1.08%	36.58%	15.68%	45.87%	0.77%	0.02%
H _{ave} 2.50m (observed)	0.87%	9.87%	39.95%	23.60%	25.55%	0.16%
H _{ave} 3.00m (observed)	0.00%	0.00%	7.85%	0.03%	0.01%	92.11%

However, this bias is drastically reduced in 180s video based learning, which is described in **Table 7.9**. Although the accuracy of one of the diagonal components was 58.54%, all diagonal terms show the highest percentage in all categories. This is similar to the 300s video based learning, represented in **Table 7.10**, but it was confirmed that the prediction for the 2.50m data was not performed well. It is considered that the learning was not performed properly because the number of data corresponding to 2.5m is extremely small in 300s video based learning.

To check the suitability of the video length, classification performance indicators(precision, recall, and accuracy) were calculated for each training, as shown in **Fig. 7.9**. The cases where the accuracy exceeds 90% are two cases; with 180s videos and 300 videos. When learnings with shorter videos, values in precision and recall are significantly low even if their accuracies are still over 80%. It was concluded that the average wave height was predicted accurately with 180s videos, in the corresponding dataset.



	Length=6s	Length=30s	Length=60s	Length=180s	Length=300s
precision	0.700	0.479	0.590	0.769	0.650
recall	0.662	0.526	0.471	0.846	0.700
accuracy	0.864	0.832	0.832	0.917	0.900

Fig. 7.9. Precision, recall, and accuracy for H_{ave} prediction with different lengths of videos.

Two attempts were adjusted to improve the learning results based on video data. First, the data of the 2.5m category, which has significantly fewer samples than other categories, was artificially augmented. After cropping multiple images from an original image, the number of data is increased by converting them into videos. The zone was set to have an angle of view composition similar to that of other categories as much as possible. An example of cropping multiple images from one image is shown in **Fig. 7.10**. By increasing the number of data in 2.5m category, it was possible to similarly match the order with the number of data in other categories.

In the second attempt, the data classification interval was adjusted constant. The current data are classified at 0.25m intervals, but the actual data were mostly obtained at 0.5 intervals. In addition, subdivided classification at small wave heights acts as a major cause of poor learning results. Accordingly, the 1.25m category was excluded and each category was adjusted to be classified only at 0.5m interval. The modified categories and the corresponding number of data are listed in **Table 7.11**.

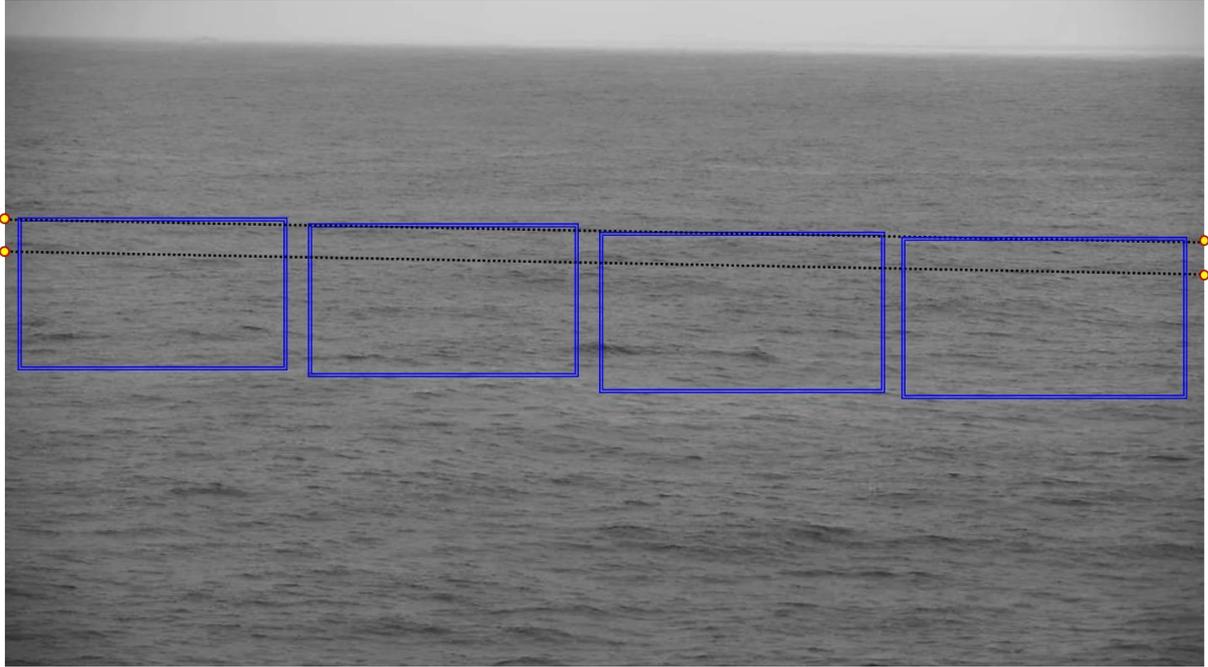


Fig. 7.10. Data augmentation for ‘Have =2.50m’ category.

Table 7.11. Data classification with augmentation in $H_{ave}=2.5m$.

		# of training data		# of validating data		# of testing data	
Movie clip (180s, 360frames)	$H_{ave}=1.00m$	86	81.1%	10	9.4%	10	9.4%
	$H_{ave}=1.50m$	214	81.1%	24	9.1%	26	9.8%
	$H_{ave}=2.00m$	77	81.1%	9	9.5%	9	9.5%
	$H_{ave}=2.50m$	62	81.6%	7	9.2%	7	9.2%
	$H_{ave}=3.00m$	47	82.1%	5	8.8%	5	8.8%

The learning results using the augmented and re-classified data are shown in **Fig. 7.11** as a global accuracy and **Fig. 7.12** as a loss function. In this case, the training was increased up to 200 epochs. The computation time was taken as 63m for the CNN module and 123m for the LSTM module. In the case of accuracy, it slowly increased up to 100 epochs and eventually ended with a high value of 94.7%. In addition, it was confirmed that the loss function gradually decreased, leaving only a fairly small value in the end. In the stage of the training, it was confirmed that the problem was suitable for the network configuration by checking the validating data.

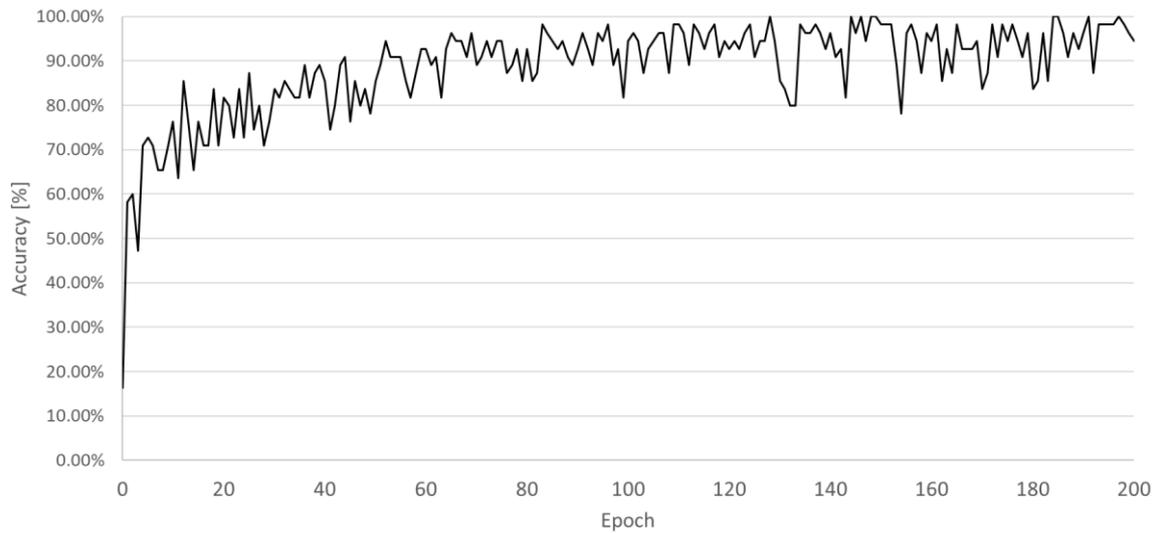


Fig. 7.11. Training results for the H_{ave} prediction with augmented videos: Global accuracy.

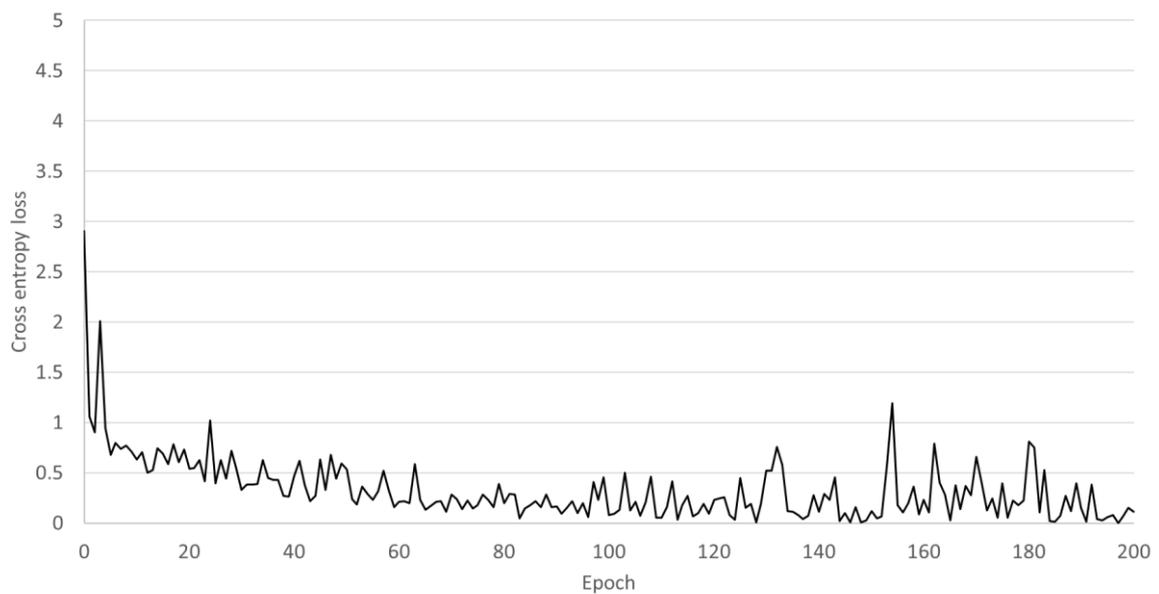


Fig. 7.12. Training results for the H_{ave} prediction with augmented videos: Cross entropy loss.

In order to judge the performance of the network obtained through training, the prediction correlation matrix for the testing data is summarized as shown in **Table 7.12**. The values of all diagonal components exceeded 80%, and the accuracy was greatly improved, especially in the small average wave height range. Also, the dominance of the diagonal components was quite high.

Table 7.12. Correlation matrix for the H_{ave} prediction with augmented videos(Length=180s).

	H _{ave} 1.00m (predicted)	H _{ave} 1.50m (predicted)	H _{ave} 2.00m (predicted)	H _{ave} 2.50m (predicted)	H _{ave} 3.00m (predicted)
H _{ave} 1.00m (observed)	99.95%	0.00%	0.04%	0.00%	0.00%
H _{ave} 1.50m (observed)	4.24%	80.11%	15.32%	0.04%	0.29%
H _{ave} 2.00m (observed)	11.02%	0.29%	88.65%	0.05%	0.00%
H _{ave} 2.50m (observed)	0.00%	0.47%	0.09%	99.44%	0.00%
H _{ave} 3.00m (observed)	0.00%	1.93%	0.00%	0.00%	98.07%

In order to judge the performance of the current learning model, it was compared with previous learnings as classification evaluation indicators, which is summarized in **Table 7.13**. Among all learning in this chapter, precision and recall were over 0.9 in this learning. It was confirmed that by applying the appropriate augmentation method to a small data category and modifying the classification criteria, the high learning performance can be achieved without additionally applying other specialized techniques.

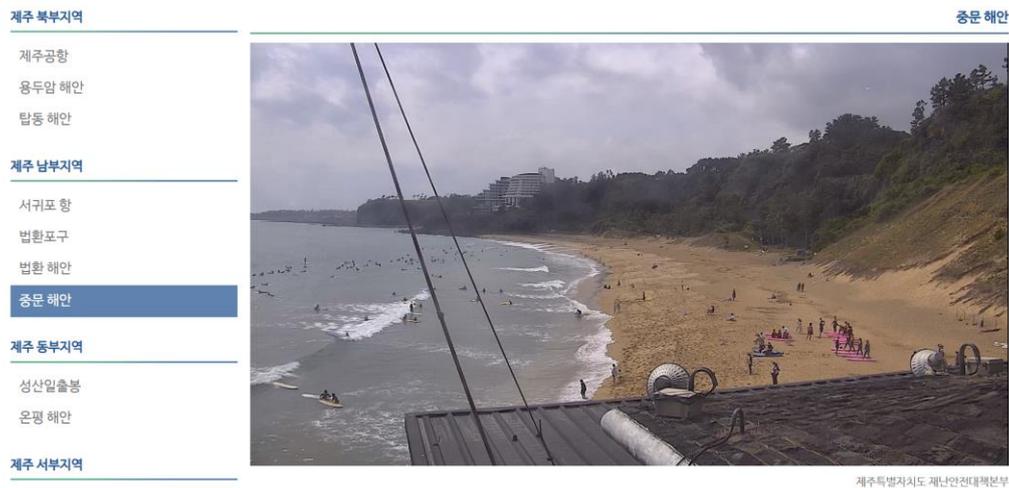
Table 7.13. Comparison in classification evaluation indicators.

	Precision	Recall	Accuracy
Snapshot-based learning	0.677	0.645	0.669
Video-based learning	0.769	0.846	0.917
Video-based learning (0.5 intervals and augmentation)	0.932	0.969	0.972

Chapter 8. Applications in marine engineering

In this study, the possibility of predicting the sea conditions through images was checked. When composing data by choosing only clean images as maintaining a fixed angle of view, it was confirmed that the average wave height can be predicted in 0.5m intervals. Although pre-processing for image quality should be conducted, and it needs to increase the prediction accuracy in deformation modes such rotation should be developed in the future.

One of the most essential points in machine learning is to achieve meaningful big data. Recently, there are numerous CCTV cameras installed along the coast in Korea. They are being used for various purposes, such as ensuring a safe circumstance for people, constructing the national defense system, monitoring coastal erosion, some research purposes, and so on. As utilizing these cameras, the meaningful big data which covers a wide area can be achieved, without additional infrastructure. **Fig. 8.1** shows snapshots that are open to the public in real time.



(a) CCTV images in Southern coastal region in Korea



(b) CCTV images in Eastern coastal beach in Korea.

Fig. 8.1. Images open to the public in real time.

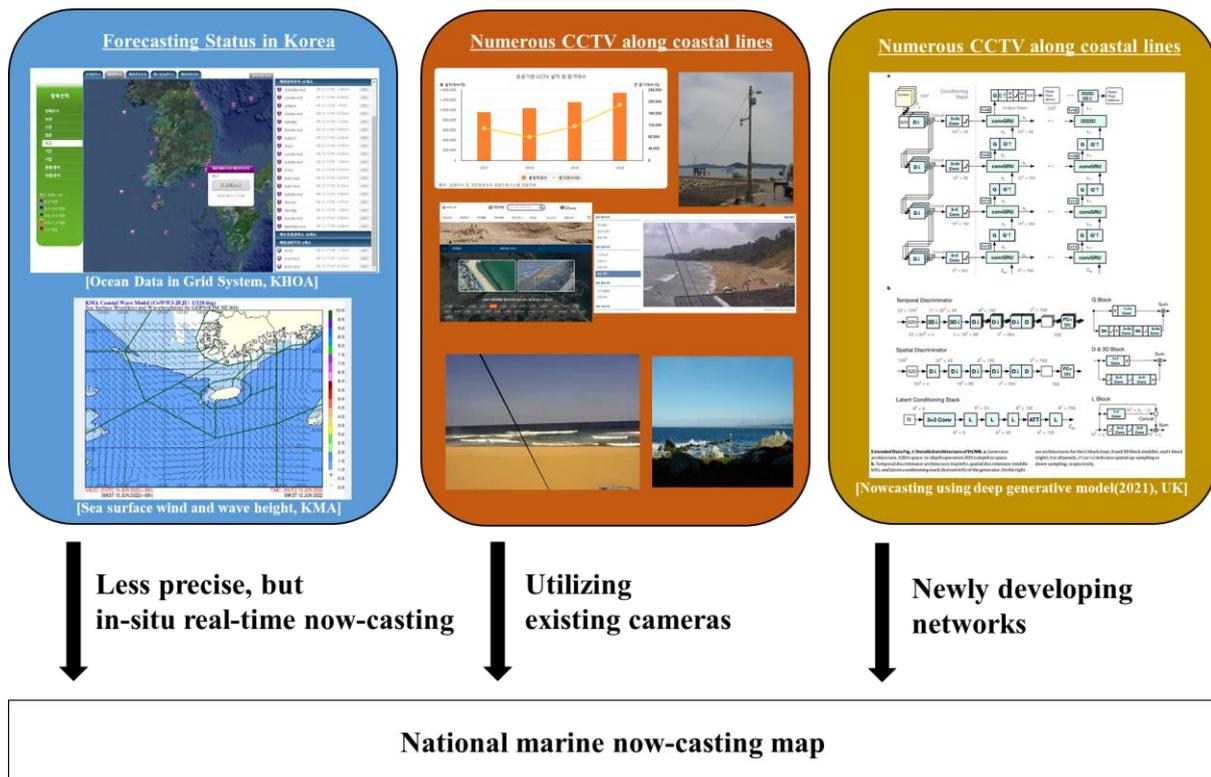


Fig. 8.2. Schematic concept of developing plan of national coastal now-casting map.

It might be possible to configure the wave height now-casting system using CCTV cameras in major coastal areas or ports. Although there are localized characteristics of each region, it is thought that since it is a narrow-banded ocean wave with a large inertia, it is considerable to conduct the following study; after achieving the trained networks at major locations, then predicting the ocean environment in the vicinity of those positions. Also, if we quantitatively analyze the coverage of the acquisition network, it would be possible to guide how many networks we need to acquire along the actual shoreline. The schematic concept of developing plan of national coastal now-casting map based on optical images and deep learning networks is illustrated in **Fig. 8.2**.

However, the management institutes of the CCTV cameras are different and it is rarely possible to utilize these measurements for solely research purpose. Therefore, it is more appropriate to first collect data in certain regions by utilizing the infrastructure of government-affiliated research institutes. Some research infrastructure of Korea government are illustrated in **Fig. 8.3**. Infrastructure exists in various forms, such as ships, centers near the coast, or fixed-type offshore structures. They are also not concentrated in one specified area, but are widely distributed along the coast of Korea. As they are non-profit organizations, it is possible to gather enough big data through these facilities for research purposes. **Fig. 8.4** shows few measuring instruments utilized in these institutes.

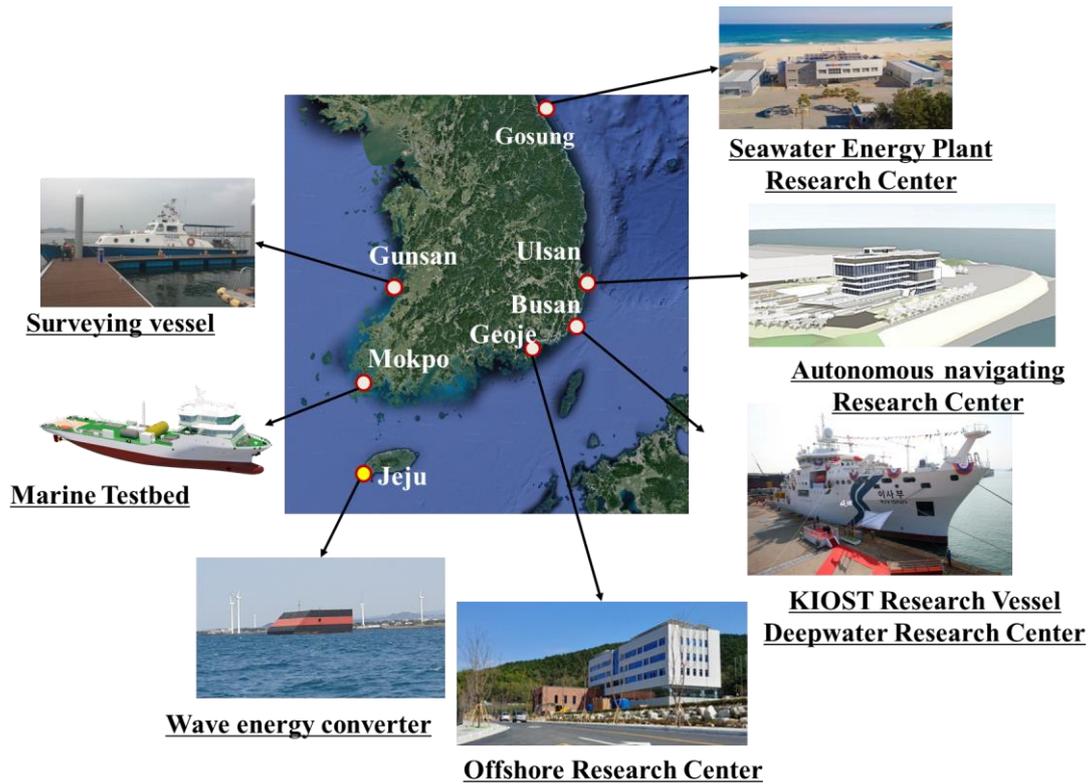


Fig. 8.3. Korea government-affiliated marine research institute infrastructure.

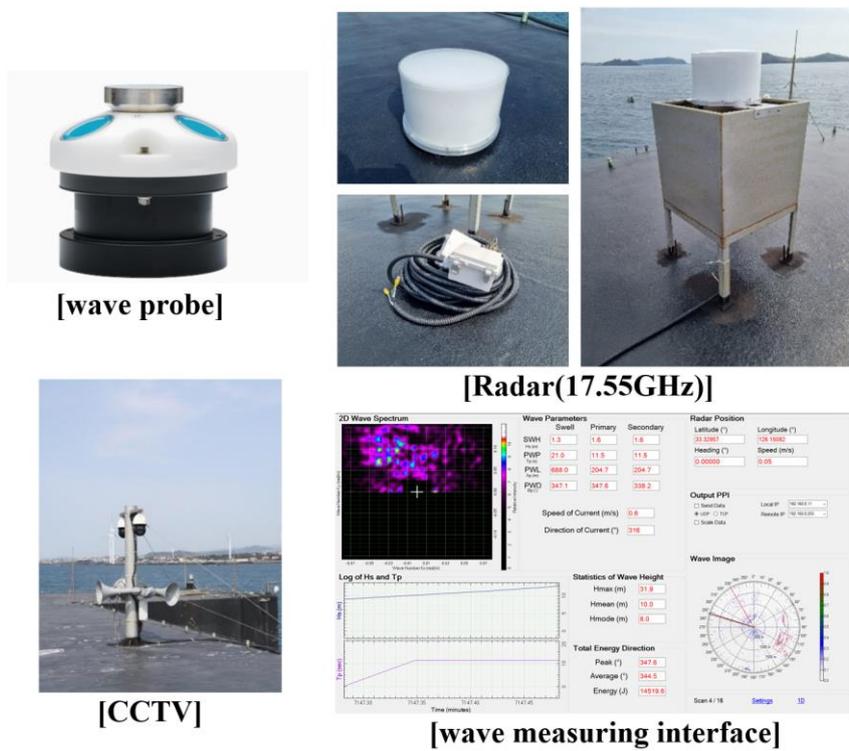


Fig. 8.4. Various measuring instruments utilized in marine research institutes.

When images are acquired by mounting on a real ship, the angle of view of the camera continuously changes due to the wave induced motions of the ship. In **Fig. 8.5**, how much the angle of view installed at the ship's bow changes due to the heave and pitch motions of the vessel is schematically illustrated. At this time, to keep the angle of view relatively constant, a gimbal can be considered to compensate the motions of a ship.

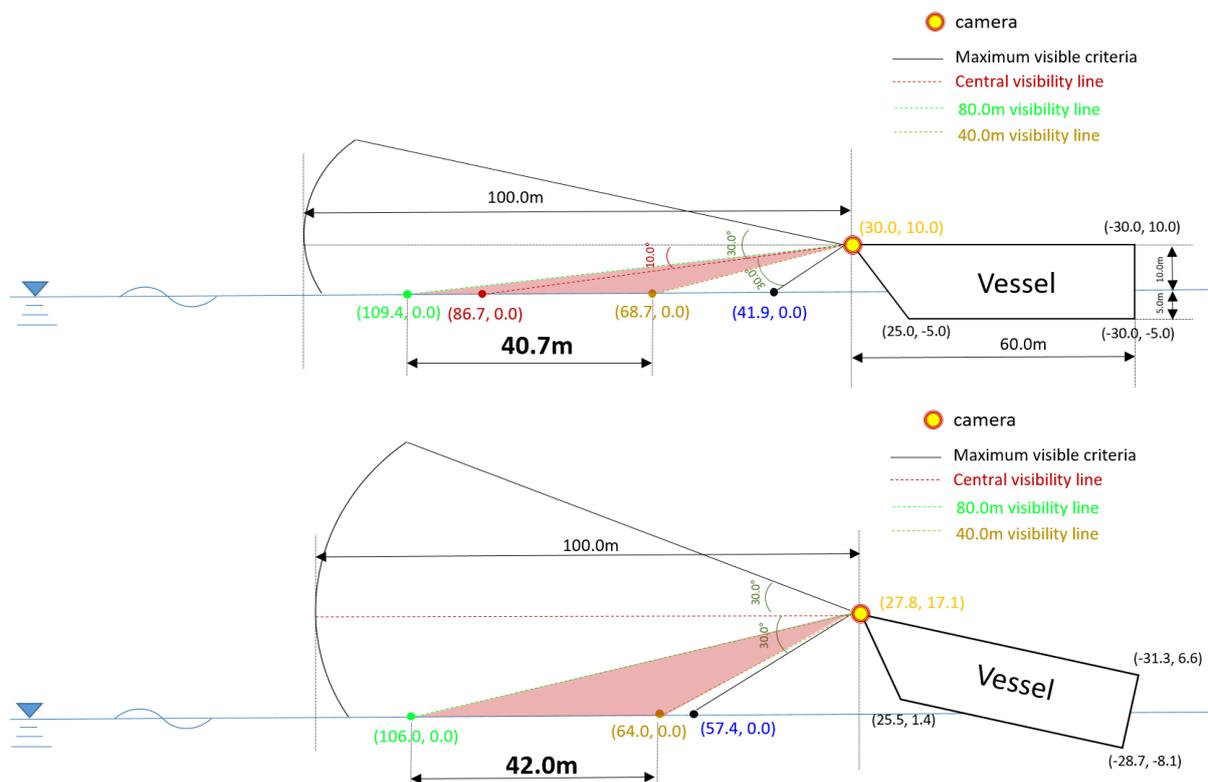


Fig. 8.5. Operation examples of the camera installed at the ship's bow.

In order to check the applicability of the trained network to images from the ship-mounted camera, images were collected from a real ship. A camera was installed near the mid-ship section of a ship operating near Mokpo. **Fig. 8.6** shows the actual installation of the camera at the ship. The operating area at the data collecting time(31th, August, 2020) is quite calm as the vessel, because it is not operating open sea and navigates between islands near the shore. The ship's route of the ship is shown in **Fig. 8.7**.

About 6500 snapshots were acquired in a total sailing itinerary of about 2 hours. According to the information from the Korea Meteorological Administration on the data achieving time, the significant wave height was in the range of 0.1-0.2m. This corresponds to sea state 2, which is a fairly serene level, and is outside the range of the network obtained in this study. Only few photos actually gathered are shown in **Fig. 8.8**.



Fig. 8.6. A camera installed at a real ship.



Fig. 8.7. Ship route for achieving images.



Fig. 8.8. Sample images achieved from ship-mounted camera.

The prediction result for the acquired data from the ship-mounted camera is shown in **Fig. 8.9** and **Table 8.1**. We defined the prediction as ‘ill-estimated’ if all predictions on each category were less than 50%. The prediction of sea state 4, sea state 5, sea state 6, and ill-estimated are 82%, 1%, 17%, and 0.3%, respectively.

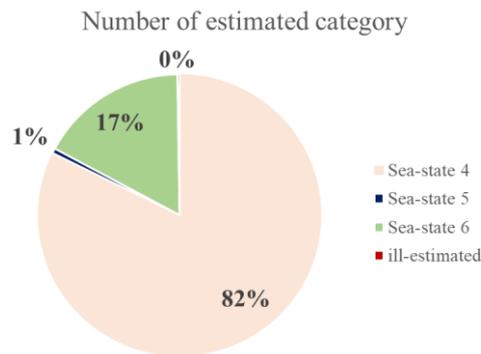


Fig. 8.9. Prediction result on images from ship-mounted camera.

Table 8.1. Statistical values of prediction accuracies on images from ship-mounted camera.

	Sea-state 4 (n=5350)		Sea-state 5 (n=37)		Sea-state 6 (n=1100)		ill-predicted (n=16)	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Sea-state 4	94.99%	0.098	6.22%	0.121	14.52%	0.153	35.38%	0.104
Sea-state 5	0.10%	0.012	83.38%	0.157	1.26%	0.049	25.86%	0.153
Sea-state 6	4.91%	0.097	10.40%	0.120	84.23%	0.155	38.76%	0.102

Samples per each classified category are shown in **Fig. 8.10**. Data acquired near Mokpo corresponded to sea state 2, while it was confirmed that the estimation results converge to sea state 4, which is a relatively calm sea. There is also a significant amount of data estimated with sea state 6, which confirms that the estimation result is different as the waves caused by the movement of the ship enters the data. Even in the case of ill-estimated, it is considered that proper estimation was not achieved because ripples caused by the motion of other buoys or ships were unusually generated.

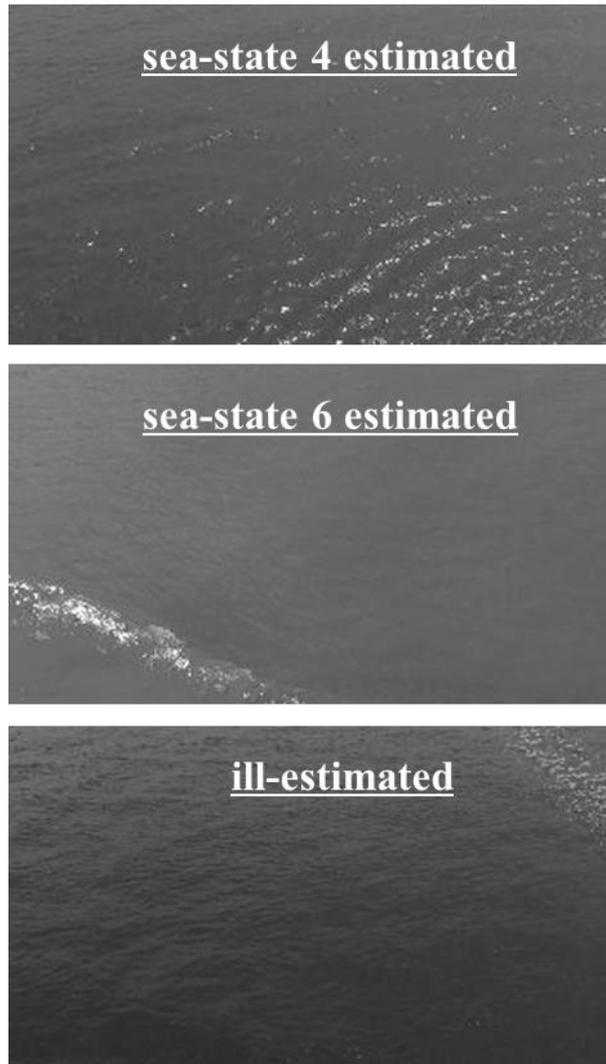


Fig. 8.10. Sample images per each classified category.

Although compensation for the motion of the vessel and preparation for long data acquisition should be conducted, from the basic perspective of judging the sea and corresponding navigating by the captain's experience, it seems to be possible to study the applicability of artificial intelligence in on-board marine now-casting system that can replace or supplement the present decision making system.

Chapter 9. Concluding Remarks

Several deep learning simulations were conducted for sea state and average wave height predictions. Existing wave measuring methodologies were briefly reviewed, and the potential benefits of the proposed wave estimating technique using deep learning were highlighted. We initially trained artificially generated irregular wave fields by Airy wave superposition in order to find the suitable training model.

The snapshot images were then gathered from the southwestern coastal region of Korea and were categorized and trained. Sea-state estimating using the designed convolution neural network showed a relatively good estimation accuracy with prescreened-conditioned data. However, categories of low wave height conditions were not predicted well. To reflect the characteristics of the continuous sea condition as input, the single snapshot-based learning was switched to the video-based one. At this time, instead of applying only the convolutional neural network, a learning model combined with convolutional neural network and long short-term memory was applied. Also, for categories with insufficient data, videos were artificially augmented. The overall prediction accuracy was significantly improved with a new approach, and it was confirmed that the sea condition could be identified above a certain level.

After that, instead of pre-classifying and learning data with good image quality, learning was carried out after collecting data for a certain period of time under general conditions without any restrictions. Compared with the estimation performance of the previous well-conditioned data, the accuracy could not increase above a certain level due to the appearance of many obstacles and severe changes in lightness. Predicting the low wave height categories was still difficult. In particular, when there is a phenomenon with strong nonlinearity such as a typhoon, the prediction was poor. To check whether it is effective when used in real vessels, the prediction performance for artificially deformed images was checked. For small simple deformations, the prediction performance degradation was relatively small. While in general, it was confirmed that the prediction performance deteriorated significantly according to the image deformation. Image angle correction with image feature extraction was adopted to weaken this degradation, but the effect was insignificant.

Few scenarios of problem solving in ocean engineering is presented by applying this technique. It includes the joint researches with autonomous ships, operations of offshore structures. A research plan for predicting sea conditions in small ships is also suggested.

In this study, optical data were acquired only for a limited period of time in a stationary offshore structure. Further studies need to be conducted for training and estimation with longer-term data, in various situations, at other regions. In addition, it is intended to develop a system that makes universal and high-precision estimation by simultaneously applying multiple mechanisms, rather than relying solely on machine learning, through linking with image processing techniques and advanced measurement equipment.

Bibliography

- [1] van Rijn, L.C., 2007. Manual sediment transport measurements in rivers, estuaries and coastal seas - Chapter 12, in: *Measuring Instruments for Fluid Velocity, Pressure and Wave Height*. Aquapublications, Netherlands.
- [2] Korea Meteorological Administrator, 2020, KMA business introduction on marine observation, https://web.kma.go.kr/eng/biz/observation_07.jsp.
- [3] Korea Hydrographic and Oceanographic Agency, 2020. Ocean data in grid framework (Korean website), <http://www.khoa.go.kr/oceangrid/khoa/koofs.do>.
- [4] Jahne, B., Klinker, J., Waas, S., 1994. Imaging of short ocean wind waves: A critical theoretical review. *J. Opt. Soc. Am. A: Opt., Image Sci., and Vis.* 11(8), 2197-2209. <https://doi.org/10.1364/JOSAA.11.002197>.
- [5] Chou, C.R., Yim, J.Z., Huang, W.P., 2004. Determining the hydrographic parameters of the surface of water from the image sequences of a CCD camera. *Exp. Fluids.* 36, 515-527. <https://doi.org/10.1007/s00348-003-0664-3>.
- [6] Viriyakijja, K., Chinnarasri, C., 2015. Wave flume measurement using image analysis. *Acquat. Proced.* 4, 522-531. <https://doi.org/10.1016/j.aqpro.2015.02.068>.
- [7] Piepmeyer, J., Waters, J., Broussard, R., 2006. 2006-2364: A stereo vision-based wave surface measurement project. 2006 Annual Conference & Exposition, Chicago, USA. DOI: 10.18260/1-2-1332.
- [8] McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* 5, 115-133.
- [9] Rosenblatt, F., 1958. The perception: A probabilistic model for information storage and organization in the brain. *Psychol.* 386.
- [10] Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature.* 323, 533-536.
- [11] Hinton, G.E., Osindero, S., Teh, Y.W., 2006. A fast learning algorithm for deep belief nets. *Neural Comput.* 18, 1527-1554.

- [12] Alex, K., Ilya, S., Geoffrey, E.H., 2012. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 1097-1105.
- [13] Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D., 2016. Mastering the game of Go with deep neural networks and tree search. *Nature*. 529, 484.
- [14] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., Van den Driessche, G., Graepel, T., Hassabis, D., 2017. Mastering the game of Go without human knowledge. *Nature*. 550, 354.
- [15] LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature*. 521, 436-444.
- [16] Deo, M.C., 2010. Artificial neural networks in coastal and ocean engineering. *Ind. J. of Geo-Mar. Sci.* 39(4), 589-596. <http://nopr.niscair.res.in/handle/123456789/10807>.
- [17] Kutz, J.N., 2017. Deep learning in fluid dynamics. *J. Fluid Mech.* 814, 1-4. <https://doi.org/10.1017/jfm.2016.803>.
- [18] Sclavounos, P.D., Ma, Y., 2018. Artificial intelligence machine learning in marine hydrodynamics. Proceedings of the ASME 2018 37th International Conference on Ocean, Offshore and Arctic Engineering, Madrid, Spain, OMAE2018-77599. <https://doi.org/10.1115/OMAE2018-77599>.
- [19] Ahmad, H., 2019. Machine learning applications in oceanography. *Aquat. Res.* 2(3), 161-169. <https://doi.org/10.3153/AR19014>.
- [20] Ramirez, W.A., Leong, Z.Q., Nguyen, H.D., Jayasinghe, S.G., 2020. Machine learning post processing of underwater vehicle pressure sensor array for speed measurement. *Ocean Eng.* 213, 107771, 1-6. <https://doi.org/10.1016/j.oceaneng.2020.107771>.
- [21] Gheliotis, M., Lazakis, I., Theotokatos, G., 2020. Machine learning and data-driven fault detection for ship systems operations. *Ocean Eng.* 216, 107968, 1-17. <https://doi.org/10.1016/j.oceaneng.2020.107968>.

- [22] Berghout, T., Mouss, L.H., Bentrucia, T., Elbouchikhi, E., 2021. A deep supervised learning approach for condition-based maintenance of naval propulsion systems. *Ocean Eng.* 221, 108525, 1-11. <https://doi.org/10.1016/j.oceaneng.2020.108525>.
- [23] Zamani, A., Solomatine, D., Azimian, A., Heemink, A., 2008. Learning from data for wind-wave forecasting. *Ocean Eng.* 35, 953-962. <https://doi.org/10.1016/j.oceaneng.2008.03.007>.
- [24] Mahjoobi, J., Mosabbel, E.A., 2009. Prediction of significant wave height using regressive support vector machines. *Ocean Eng.* 36, 339-347. <https://doi.org/10.1016/j.oceaneng.2009.01.001>.
- [25] Wei, C.C., 2017. Nearshore wave prediction using data mining techniques during typhoons: A case study near Taiwan's northeastern coast. *Energ.* 11(1), 11. <https://doi.org/10.3390/en11010011>.
- [26] James, S.C., Zhang, Y., O' Donncha, F., 2018. A machine learning framework to forecast wave conditions. *Coast. Eng.* 137, 1-10. <https://doi.org/10.1016/j.coastaleng.2018.03.004>.
- [27] Sarkar, D., Osborne, M.A., Adcock, T.A.A., 2018. Prediction of tidal currents using Bayesian machine learning. *Ocean Eng.* 158, 221-231. <https://doi.org/10.1016/j.oceaneng.2018.03.007>.
- [28] Stringari, G.E., Harris, D.L., Power, H.E., 2019. A novel machine learning algorithm for tracking remotely sensed waves in the surf zone. *Coast. Eng.* 147, 149-158. <https://doi.org/10.1016/j.coastaleng.2019.02.002>.
- [29] Kim, S., Tom, T.H.A., Takeda, M., Mase, H., 2021. A framework for transformation to nearshore wave from global wave data using machine learning techniques: Validation at the Port of Hitachinaka, Japan. *Ocean Eng.* 221, 108516, 1-13. <https://doi.org/10.1016/j.oceaneng.2020.108516>.
- [30] Demetriou, D., Michailides, C., Papanastasiou, G., Onoufriou, T., 2021. Coastal zone significant wave height prediction by supervised machine learning classification algorithms. *Ocean Eng.* 221, 108592, 1-15. <https://doi.org/10.1016/j.oceaneng.2021.108592>.
- [31] Zhong, G., Liu, B., Guo, Y., and Miao, H., 2018. Sea State Bias Estimation with Least Absolute Shrinkage and Selection Operator (LASSO). *J. Ocean Univ. China.* 17(5), 1019-1025. <https://doi.org/10.1017/s11802-018-3572-0>.

- [32] Zhong, G., Qu, J., Wang, H., Liu, B., Jiao, W., Fan, Z., Miao, H., and Hedjam, R., 2020. Trace-Norm Regularized Multi-Task Learning for Sea State Bias Estimation. *J. Ocean Univ. China*. 19(6), 1292-1298. <https://doi.org/10.1017/s11802-020-4267-x>.
- [33] Liu, T., Zhang, Y., Qi, L., Dong, J., Lv, M., Wen, Q., 2019. WaveNet: learning to predict wave height and period from accelerometer data using convolutional neural network. *Proceedings of International Conference on Environment and Ocean Engineering*, Xiamen, China. <https://doi.org/10.1088/1755-1315/369/1/012001>.
- [34] Buscombe, D., Carini, R.J., 2019. A data-driven approach to classifying wave breaking in infrared imagery. *Remote Sens.* 11(7), 859. <https://doi.org/10.3390/rs11070859>.
- [35] Masoumi, M., 2021. Ocean data classification using unsupervised machine learning: Planning for hybrid wave-wind offshore energy devices. *Ocean Eng.* 219, 108387, 1-15. <https://doi.org/10.1016/j.oceaneng.2020.108387>.
- [36] Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clandy, E., Arribas, A., Mohamed, S., 2021. Skilful precipitation nowcasting using deep generative models of radar. *Nature*. 597, 672-677.
- [37] Ewans, K., Feld, G., Jonathan, P., 2014. On wave radar measurement. *Ocean Dynamics*. 64(9), 1281-303.
- [38] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv*, arXiv:1409.1556.
- [39] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 770-778.
- [40] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, 2014. Going Deeper with Convolutions. *arXiv*, arXiv:1409.4842v1.
- [41] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the inception architecture for computer vision. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2818-2826.

- [42] Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 1251-1258.
- [43] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A., 2017. Inception-v4, Inception-ResNet and the impact of residual connections on learning. Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4278-4284.
- [44] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C., 2018. MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510-4520.
- [45] Beale, M.H., Hagan, M.T., Demuth, H.B., 2020. Deep Learning Toolbox™ User' s Guide. MathWorks, Natick, MA, USA, <https://kr.mathworks.com/help/deeplearning/>.
- [46] Hochreiter, S., Schmidhuber, J., 1997. Long Short-Term Memory. Neural Computation. 9(8), 1735-1780. <http://www.bioinf.jku.at/publications/older/2604.pdf>.
- [47] Gers, F.A., Schmidhuber, J., 2000. Recurrent nets that time and cost. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, Como, Italy, 6697377. DOI: 10.1109/IJCNN.2000.861302.
- [48] Cho, K., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y., 2014. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv, arXiv:1406.1078v3.
- [49] Yao, K., Cohn, T., Vylomova, K., Duh, K., Dyer, C., 2015. Depth-Gated Recurrent Neural Networks. arXiv, arXiv:1508.03790v2.
- [50] Koutnik, J., Greff, K., Gomez, F., Schmidhuber, J., 2014. A Clockwork RNN. arXiv, arXiv:1402.3511v1.
- [51] DNVGL, 2019. DNVGL Recommended Practice: Environmental Conditions and Environmental Loads. DNVGL-RP-C205, Oslo, Norway, <https://oilgasstandards.dnvgl.com/download/dnvgl-rp-c205-environmental-conditions-and-environmental-loads>.
- [52] International Towing Tank Conference, 2012. ITTC - Recommended Procedures and Guidelines: Speed and Power Trials, Part 1 Preparation and Conduct. 7.5-04-01-01.1, <https://itcc.info/media/4208/75-04-01-011.pdf>.

- [53] Johnson, J.M. and Khoshgoftaar, T.M., 2019. Survey on deep learning with class imbalance. *Journal of Big Data*. 6:27, 1-54. <https://doi.org/10/1186/s40537-019-0192-5>.
- [54] Leutenegger, S., Chli, M., Siegwart, R.Y., 2011. BRISK: Binary Robust invariant scalagle keypoints. *Proceedings of the 2011 International Conference on Computer Vison*, Barcelona, Spain, <https://doi.org/10.1109/ICCV.2011.6126542>.
- [55] Rosten, F. and Drummond, T., 2005. Fusing points and lines for high performance tracking. *The 10th IEEE International Conference on Computer Vision*. 2. 1508-1515.
- [56] Harris, C. and Stephens, M., 1988. A Combined Corner and Edge Detector. *Proceedings of the 4th Alvey Vision Conference*, Manchester, UK, 147-151.
- [57] Alcantarilla, P.F., Bartoli, A., Davison, A.J., 2012. KAZE Features. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds) *Computer Vision - ECCV 2012*. ECCV 2012. *Lecture Notes in Computer Science*, vol 7577. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33783-3_16.
- [58] Shi, J. and Tomasi, C., 1994. Good Features to Track. *IEEE Conference on Computer Vision and Pattern Recognition(CVPR94)*, Seattle, Washington, USA, 593-600.
- [59] Rublee, E., Rabaud, V., Konolige, K., Bradski, G. 2011. ORB: An Efficient Alternative to SIFT or SURF. *Proceedings of the 2011 International Conference on Computer Vison*, Barcelona, Spain, 2564-2571.
- [60] Bay, H., ESS, H., Tuytelaars, T., van Gool, L. 2008. SURF: Speeded Up Robust Features. *Computer Vision and Image Understanding(CVIU)*. 110(3). 346-359.