



Article Toward a Highly Accurate Classification of Underwater Cable Images via Deep Convolutional Neural Network

Guan Wei Thum ^{1,*}, Sai Hong Tang ^{1,*}, Siti Azfanizam Ahmad ¹ and Moath Alrifaey ^{1,2}

- ¹ Department of Mechanical and Manufacturing Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia; s_azfanizam@upm.edu.my (S.A.A.); eng.alrifaiy2005@gmail.com (M.A.)
- ² Department of Mechanical Engineering, Faculty of Engineering Technology and Built Environment, UCSI University, Taman Connaught, Kuala Lumpur 56000, Malaysia
- * Correspondence: alvinthum615@gmail.com (G.W.T.); saihong@upm.edu.my (S.H.T.); Tel.: +03-9769-6332 (S.H.T.)

Received: 1 October 2020; Accepted: 3 November 2020; Published: 16 November 2020



Abstract: Underwater cables or pipelines are commonly utilized elements in ocean research, marine engineering, power transmission, and communication-based activities. Their performance necessitates regularly conducted inspection for maintenance purposes. A vision system is commonly used by autonomous underwater vehicles (AUVs) to track and search for underwater cable. Its traditional methods are characteristically applicable in AUVs, wherein they are equipped with handcrafted features and shallow trainable architectures. However, such methods are subpar or even incapable of tracking underwater cable in fast-changing and complex underwater conditions. In contrast to this, the deep learning method is linked with the capacity to learn semantic, high-level, and deeper features, thus rendering it recommended for performing underwater cable tracking. In this study, several deep Convolutional Neural Network (CNN) models were proposed to classify underwater cable images obtained from a set of underwater images, whereby transfer learning and data augmentation were applied to enhance the classification accuracy. Following a comparison and discussion regarding the performance of these models, MobileNetV2 outperformed among other models and yielded lower computational time and the highest accuracy for classifying underwater cable images at 93.5%. Hence, the main contribution of this study is geared toward developing a deep learning method for underwater cable image classification.

Keywords: underwater cable; pipelines; image classification; deep learning; transfer learning; data augmentation

1. Introduction

Underwater infrastructures such as underwater communication cable, underwater power cable, and subsea pipeline are highly crucial to humankind. In particular, underwater communication cables played an essential role over the last 170 years to connect the whole world. Previously, the Asia-Pacific Economic Cooperation forum revealed that around 97% of all intercontinental data available are transferred via underwater cables [1]. In the current modern era, the global demand for data consistently increases every year, especially in the explosion of mobile device usage and the development of cloud computing, big data, and artificial intelligence. Therefore, new underwater communication cables are highly in demand in order to support higher speed and larger capacity of data, voice, and video transmissions. Accordingly, underwater power cables typically deployed at oil and gas platforms or renewable power projects are generally utilized to connect topside and subsea facilities for power

2 of 23

provision purposes [2]. Meanwhile, an underwater pipeline is implemented to transport important resources such as oil and natural gas, whereby the longest underwater pipeline is used to transport natural gas, namely, the Langeled Pipeline. The cable is about 1200 km long under the North Sea and spans from the Ormen Lange field in Norway to the Easington Gas Terminal in the United Kingdom [3]. From 1986 to 2003, around 70% of faulty communication cables were attributed to benthic fishing and ship anchors, whereby they would occur between the water depths of 0 and 200 m [4]. According to the Submarine Telecoms Industry Report 2018/2019, the average time required for a repair crew to restore the cables has shown a reducing trend from 30 days to 26 days throughout the duration of 2013 until 2019. In general, most of the time required is spent by the crew to find, track, and diagnose the faulty cables [5].

The importance of underwater cables and underwater pipelines as underwater infrastructures renders their protection paramount, whereby the cables are typically protected by covering them with steel wires and burying them under the seabed [6]. Due to the working conditions of rough and aggressive underwater pipeline conditions, they are prone to leakage or failure. As a result, the oil and gas sector has developed reliable leak detection systems to monitor the state of the pipelines [7]. For example, Ortiz et al. [8] have stated that a better alternative is to consistently inspect the current cable state to prevent damages, which include corrosion, crack, or those due to human activities such as marine traffic or fishing. Accordingly, constant monitoring and inspection of the underwater cables are highly recommended for early detection of defects, which are commonly carried out with the used of surface ships, remotely operated underwater vehicles or both together. However, their response and mobilization time are not satisfactorily adequate [9], whereas inspection undertaken by human divers poses health and safety issues, especially due to the difficulties for them to find, track, inspect, and diagnose underwater cables over an extended period [10]. Therefore, an autonomous underwater vehicle (AUV) is strongly recommended for the cable tracking operation and diagnosing its fault, whereby it is capable of collecting information of importance using its sensor and making an appropriate decision in different conditions via its embedded intelligent algorithms [11].

Furthermore, a variety of sensors can be fitted on AUVs to perform underwater cable tracking operations, wherein extensive research efforts have been spent to study and use them, such as vision-based sensor, sonar sensor, and magnetometer sensor. Normally, AUVs are fitted with vision sensors for research purposes, typically in biological, geological, and archaeological surveys, thus rendering them one of the standard equipment [2]. This type of sensor is recommended to be programmed with an inspection framework for underwater cable search and detection [12], whereby it is less expensive, capable of identifying faults in a short distance, and offers a lower power consumption for its operations. Meanwhile, undersea scenes are typically captured by video cameras, following which important information is extracted and analyzed by the vision system to guide the operation of underwater robots. Accordingly, the rough and dynamic underwater conditions may cause underwater images to show a blurring effect, low-contrast environment, and nonuniform illumination, thus increasing their complexity, computational time and cost [13]. Low-contrast and non-illumination environments confuse the object detection algorithm to detect the target object from background environment [14]. Besides, natural properties of water cause the scattering and absorption of light, which influences the quality of images [15]. This further increases the difficulty in object detection. Consequently, researchers have developed a different kind of technique to restore or enhance the degraded image qualities [16], which will specifically reduce the noise in images and thus increase the classification accuracy. Lee et al. [16] applied Jaffe-McGlamery model to restore and recover the degraded color information. With the restoration of the color information, this simplifies the object detection algorithm to identify the interested object.

Some of the commonly implemented cable tracking methods include Hough Transform, Kalman filters, and particle filters, which incorporate an algorithm to search for the main straight line in the underwater images and reveal its position [2]. In line with this, Ortiz et al. [17] have developed a system based on image segmentation, thereby transforming an image to grayscale outcome that can be classified in different gray-level intensities in order to extract the cable from its environment.

Meanwhile, the Kalman filter is used to predict and identify the pose of cable in the subsequent images, which is a method employed by Antich and Ortiz [18] to predict the cable. However, the process of segmentation is different from previous research efforts, wherein contour extraction and line extraction are included to search for the main straight line in the images and thus greatly minimizing the computation time. Regardless, both approaches require one to decide the number of segments by partitioning the images manually. Such static parameters may lead to false results due to the constantly changing environment and the nature of the obstacles.

Accordingly, Balasuriya and Ura [19] have utilized multi-sensors that are fused together to search for underwater cables, whereby the dead reckoning position is employed to predict the cable in the images, which is then located via Hough Transform. Meanwhile, Chen et al. [20] have proposed an algorithm that applies a probabilistic Hough Transform for line detection and increases the detection speed, wherein the line detection requires good visibility of the underwater images and 40% of the edge point. Similarly, Fatan et al. [11] have proposed the use of Hough Transform for cable tracking purposes by using the Multilayer Perceptron (MLP) neural network and Support Vector Machine (SVM) to extract texture information from the images. This combines machine learning (MLP and SVM) and Hough Transform, the latter of which is capable of tracking a straight line in the images based on the extracted information. It should be noted that the Hough Transform approach needs a controlled environment for cable tracking and its limiting features include sediment-covered pipes, non-uniform illumination, or spurious edge detection from other pipes or elements, which will reduce its performance [12].

Nevertheless, Wirth et al. [21] have proposed a probabilistic approach for vision-based cable tracking with particle filters, which is tested by using a group of videos that record the state of power cable that has been installed about 30 years ago. Meanwhile, Ortiz et al. [8] have developed a cable tracking system which used the particle filters to counter the complexity of undersea environments. For every video frame in the video, the previously obtained probability density function data of the undersea cable parameters are then applied to predict the position of cable in a subsequent frame. However, a fast change of environment would affect the algorithm to predict the position of underwater cable.

Over the last few decades, a deep learning method has been successfully used in different kinds of fields, technologies, and mechanisms requiring a huge amount of data for training purposes, providing useful information. The improvement of its methods has been found to be a remarkable success in the context of image identification, object detection, image classification, and face identification tasks [22]. Besides, deep learning is employed in face detection [23], pedestrian detection [24], and underwater object detection as well [25,26]. According to O' Byrne et al. [27], it lies in its highly repurposable nature. When a model is trained with a huge amount of marine growth images, it can be reprogrammed to become a crack detection model specifically by training the model with crack images.

In the era of the Internet of Things (IoT), the technology of deep learning has grown rapidly in the context of computer vision, which is a task highly influenced by video analysis and image understanding methods. The development of deep learning has attracted much research attention in current times. Image classification or recognition is the primary domain in the deep learning field, which learns the important feature from the images and based on that information, images will be classified [22,28]. Object detection or localization locates and provides spatial information for the target object in images. Object tracking needs to obtain sufficient feature data during the image classification stage and combine with a deep-learning-based classifier for tracking the target in a real-time scenario [28].

Commonly, traditional object detection models can roughly be categorized into three stages, namely the informative region selection, feature extraction, and classification [29]. Accordingly, they are less effective in constructing a complex situation when classifying multiple low-level image features with high-level context [12]. However, deep learning may address these issues typically present in traditional architectures [29]. This is due to the natural properties of the deep learning method,

large availability dataset and high powerful hardware in the current market [30]. Valdenegro-Toro [31] and Kvasic et al. [32] have applied deep Convolutional Neural Network (CNN) for underwater object recognition by using the underwater sonar images, their study showed outstanding performance when compared to the current state-of-the-art method. Jalal et al. [33] have developed a deep learning method to locate the position of the fish and perform species classification from images. The Jalal et al. [33] method has achieved remarkable results compared to other methods.

Figure 1 describes the application and process of deep Convolutional Neural Network (CNN) in object detection. Deep CNN consists of various neural networks in object detection, such as Convolutional Neural Network, Region-based Convolutional Neural Network, and Fast Region-based Convolutional Neural Network.



Figure 1. Use of convolutional neural network for object detection.

In particular, Buetti-Dinh et al. [34] have developed deep CNNs to classify the bacterial biofilm composition, whereby its usage outperforms the human experts given the results obtained: 90% (via CNN) compared to 50% (by human experts). Similarly, Villon et al. [26] have employed deep CNN to identify fish species from underwater images, wherein its accuracy is as high as 94.9% and greater than humans (89.3%). Therefore, the proposed deep CNN is able to identify fishes in complex conditions and more effective than human methods even for the smallest or blurry images. In 2019, Gómez-Ríos et al. [35] created and proved that deep CNN was an excellent technique for the classification of underwater coral images accurate coral classification model by using several deep CNN models which are Inception V3, ResNet-50, ResNet-152, DenseNet-121, DenseNet-161.

Previous used vision systems are using an algorithm to identify two straight lines from images and calculate the probability of underwater cable in images based on the previous frame. The existing knowledge gap in those methods is the complex and fast change of underwater environments that might affect the detection of a line from images and influence the prediction of the system. Based on the aforementioned, the author believed deep CNN to be able to minimize the gap of knowledge in the classification of underwater cable images. Hence, the current study employs different types of deep CNN models to perform the classification of underwater cable images and varying optimization techniques are applied to increase their performance.

2. Materials and Methods

Nowadays, deep learning techniques have been used and studied across different research areas due to their outstanding performance. In particular, deep CNNs are one of the artificial neural networks (ANNs) formed by a stack of convolutional layers, activation function, pooling layers, and fully connected layers. Their natural procedure entails the learning of low-level and high-level features such as edges and curves, and shapes and different patterns from input image data, respectively [14]. Driven by such achievements across various research areas, this study employed the deep CNN method to perform the classification of underwater cable images.

Figure 2 details the overall framework in the development of deep CNN model for the classification of underwater cable images from underwater images. Input data, deep Convolutional Neural Network

model, and experiment setting are important for this particular task. A group of underwater cable videos were collected from different underwater cable service companies. Those videos were converted to video frames. After that, the video frames are standardized to 75×75 pixels and arranged accordingly before being fed into the deep CNN models for training purposes.



Figure 2. Flow chart of deep Convolutional Neural Network (CNN) model development for the classification of underwater cable images.

Several deep CNN models were chosen in this study and subjected to training and performance evaluation, whereby transfer learning and data augmentation were applied to optimize their performance and attain accuracy above 90%. Then, the suitable deep CNN model and optimization techniques were proposed for the classification of underwater cable images in which its training was carried out by using TensorFlow-GPU in Jupyter Notebook on an Intel Core i7 7500U/2.7 GHz using an NVIDIA GeForce 940MX laptop.

2.1. Data Acquisition

In this research, all underwater images employed were converted from a video of underwater cable tracking task, whereby about 13 underwater cable tracking and inspection video clips had been obtained from the website of underwater service companies. Ten videos clips were selected for training and validation purpose while another three videos were used to test the performance of the models. Accordingly, the video clips were selected and extracted to yield video frame images, following which a total of 2000 underwater images were collected. Then, the images were manually selected and categorized into two groups, which were labeled as 'images with underwater cable' and 'images without underwater cable' accordingly. Much of the conventional marine research employs color images to train the deep CNN models as they can extract more information from such images [36]. Recently, the deep learning techniques are strongly recommended in solving problems with low-cost

devices and less public resources [37,38]. The higher-resolution images require higher computational time and expensive hardware for processing the information, hence low-resolution images are decided to be used in this study Therefore, all underwater images in this study were set in the RGB color mode and resized to the dimension of 75 × 75 pixels, which were the minimum input size of the deep CNN models. All of the images were divided randomly into either the training set (70%), validation set (20%), and test set (10%) before they were fed to the deep CNN model. Figure 3 shows the overall framework for data acquisition and data pre-processing, whereas Table 1 depicts the categorization of image data for this particular study.



Figure 3. Overall framework for data acquisition and data pre-processing.

	Traini	ng Set	Validat	tion Set	Test Set			
	Images with Underwater Cable	Images without Underwater Cable	Images with Underwater Cable	Images without Underwater Cable	Images with Underwater Cable	Images without Underwater Cable		
Number of Images Data	670	730	200	200	100	100		

Table 1. Categorization of image data for training CNN models.

Figures 4 and 5 show some examples of images with underwater cable and images without underwater cable. With the use of lower-resolution images, the performance of deep CNN models in the classification of underwater cable images might be affected. The study of Kannojia and Jaiswal [39], in particular, has shown that such performance is decreased when the degradation of image resolution from a higher to lower grade occurs. Regardless, the use of several optimization techniques such as transfer learning and data augmentation can minimize the problem of low-resolution images [14,35,40].



Figure 4. Sample images: Images with underwater cable $(3 \times 75 \times 75 \text{ pixels})$.



Figure 5. Sample images: Images without underwater cable $(3 \times 75 \times 75 \text{ pixels})$.

2.2. Deep Convolutional Neural Network Model

In general, deep learning architectures can be categorized into Deep Belief Network (DBN), Boltzmann Machine (BM), Restricted Boltzmann Machine (RBM), Deep Auto-Encoder (DAE), and Convolutional Neural Network (CNN) accordingly. Much of the recent scholarly efforts have underlined the superior performance of deep CNN in the learning feature from images, whereby they are used for image classification problems [41]. In this study, several deep CNN models were selected based on the consideration of the learning parameter, layer of the models, computational cost, and performance of deep CNN models. As a result, the following were chosen: MobileNet, MobileNet V2, Inception V3, Xception, and Inception-ResNet-V2.

- MobileNet—MobileNet is built for a very small, low-latency model and is purposely designed to
 use in low-cost applications. MobileNets is designed based on a streamlined architecture that
 applies depthwise separable convolutions. This is to develop a less computational cost and lower
 parameter of deep neural network. MobileNet has been applied in large-scale geolocalization,
 face attributes, object detection, face embedding [38]. The minimum input layer of MobileNet
 accepts images of 32 × 32 pixels.
- MobileNetV2—MobileNetV2 is a new mobile architecture introduced by a new technique that is inverted residual with a linear bottleneck. This further increases its performance and reduces the need for main memory from the hardware [42]. The minimum input layer of MobileNetV2 accepts images of 32 × 32 pixels.
- Inception V3—The architecture of Inception was introduced as GoogLeNet, named Inception V1. InceptionV3 is a variant of GoogleNet that was refined by adding factorization ideas [43]. The computational cost of Inception is lower than VGGNet as the number of parameters is less. The minimum input layer of InceptionV3 accepts images of 75 × 75 pixels.
- Xception—Xception stands for "Extreme Inception" which outperforms than Inception V3. The architectures of Xception are stimulated by the idea of Inception, where replaced Inception modules with depthwise separable convolutions [44] and Xception outperforms Inception V3 due to higher model efficiency. The minimum input layer of Xception accepts images of 71 × 71 pixels.
- Inception-ResNet-V2—Inception-ResNet-V2 is the combination of the idea of Inception model and Resnet model for obtaining high performance at low computational cost compare to other models. The Inception model tends to develop deeper layers to achieve good performance. Resnet model performs better for training very deep architecture by using residual block to inherent importance data. The combination of both models means that the Inception model is able to reap all the advantages of the Resnet model while maintaining its computational efficiency [45]. The minimum input layer of the Inception-ResNet-V2 model is 75 × 75 pixels.

2.3. Transfer Learning

Transfer learning is a machine learning technique where a pre-trained deep CNN model is subjected to retraining with new input data for a different task. Instead of rebuilding a new model, which requires a lot of time and cost, transfer learning technique is thus found to be useful for reusing the model without influencing its performance [46]. In this article, fine-tuning and deep feature learning were applied to train the models by using the underwater images collected. Generally, a huge amount of data is required for the deep CNN models to learn from scratch. However, in the case of insufficient data for training the model in a specific problem, the fine-tuning approach is decidedly helpful. It is carried out by retraining the last layer of the model with a new dataset to classify the images with or without a cable. Here, the weights of the early layers are frozen as they are used to learn low-level features such as edges and lines, whereas the last layer is specifically employed for the classification task. Meanwhile, another approach to transfer learning is deep feature learning in which new input data are provided to the pre-trained models, which will learn the feature from the input data. Such models will then be retrained with new input data and the weight values in the layers are

thus updated. Here, the learning rate of the pre-trained deep CNN models is faster than a new deep CNN model due to all of the weights stored in the layers [46]. Figure 6 shows the general framework for both fine-tuning and deep feature learning.



Figure 6. General framework of fine-tuning and deep feature learning.

2.4. Data Augmentation

Data augmentation is a very useful approach in generating an abundance of data from the original data while preserving the important information in the newly generated data [35,40,47–50]. Therefore, high-quality and a huge amount of data are important for improving the performance of various deep learning models. In this article, data augmentation was applied, wherein an Augmentor software package was employed to generate an abundance of image data for deep CNN model training [51]. To ensure a high-performing deep CNN model, the existing data were extended by applying three different augmentation techniques, namely, rotation, flipping, and random distortion. Meanwhile, other techniques such as random cropping, random zoom, and color augmentation were not suggested since they might impact the deep CNN model to learn the features of an underwater cable. The aforementioned three operations alone would allow for the generation of augmented images, specifically bypassing the input image through the pipeline multiple times. The operation is either applied or skipped based on a user-defined probability parameter; if an operation is applied, its parameters are chosen randomly within the user-specified range. The operation pipeline employed to generate data is shown in Figure 7.



Figure 7. Operation pipeline of Augmentor to generate image.

Here, a total of 23,200 images were generated using the data augmentation techniques, which were then combined with the original data and fed into all the deep CNN models. Figure 8 shows the sample images generated by Augmentor.



Figure 8. Original image and augmented images generated by Augmentor listed as (**a**) original image, an image with underwater cable $(3 \times 75 \times 75 \text{ pixels})$, and (**b**) augmented image, an image generated by Augmentor $(3 \times 75 \times 75 \text{ pixels})$.

2.5. Training Settings

In this study, transfer learning was implemented to train the deep CNN models, which were initialized by using the pre-trained weights obtained from ImageNet. In the context of the fine-tuning approach, all of the layers were frozen, except for the last layer, which was removed as it was employed to classify the images in ImageNet. Hence, the last layers of all deep CNN models in this study were replaced with a dense layer of two neurons via the SoftMax activation function. The dense layer with two neurons was then utilized to classify the images, and then validated using 400 images and tested using 200 images. Meanwhile, the Deep Feature Learning process was similar: the last layer was removed and replaced with a dense layer of two neurons by using the SoftMax activation function. However, all of the previous layers were not frozen and the weight was consistently updated during the training. Afterward, the models were subjected to training, validation, and testing by using 1400, 400, and 200 images, respectively.

Then, the data augmentation technique was applied to generate a total of 20,000 images used for training and 5000 images subjected to validation. Both fine-tuning and deep feature learning techniques were applied for all models, which were then trained with 20,000 images, validated with 5000 images, and tested with 200 images.

An ADAM optimizer with a suitable learning rate value would aid them in learning for the training data set without losing any of the useful features and ensure the learning process in obtaining a good local minima [52]. As transfer learning was utilized in this study, the Adaptive Moment Estimator Optimizer (ADAM) with a learning rate of 0.00001 was thus suggested for the deep CNN models [52]. Meanwhile, categorical cross-entropy loss was applied for any loss function. Throughout the study, the batch size was limited to 10 and iterated for 100 epochs, following which the deep CNN models were implemented by using the Keras API in Jupyter Notebook.

2.6. Testing the Model Performance

All five deep CNN models utilized in this study were trained with 1400 images and validated using 400 images by implementing different techniques, wherein their respective performance was next tested with 200 images. This would ensure the attainment of classification accuracy, precision value, recall value, and f1 score. Next, the classification performance of the models was visualized by using a confusion matrix, which is a table detailing the number of correct and incorrect predictions. Figure 9 shows a sample of the confusion matrix. Following this, the classification accuracy, precision value,

recall value, and f1 score of all five deep CNN models were compared. In particular, classification accuracy is the ratio between the number of true positive samples and the total number of samples, while precision is the closeness of true positive prediction. Meanwhile, recall is also known as sensitivity, which is the true positive rate, whereas the f1 score is used to measure the weighted average of precision and recall values. In a classification problem with more than two classes, average precision, recall, and f1 score are calculated to show the performance of model.

Confusion Matrix



Prediction

Figure 9. Confusion Matrix, a matrix that explain the prediction outcomes.

The accuracy, precision, recall, and F1 score are calculated as follows:

$$Accuracy = (True Positives + True Negatives)/N,$$
(1)

where N is the total number of instances,

3. Results

3.1. Performance of Deep CNN Models with Applied of Transfer Learning

3.1.1. Fine-Tuning

Due to the difference present between the images in ImageNet and ocean images, fine-tuning was applied for all layers of the deep CNN models in this study to train them [53]. The optimization approach was applied to transfer data from the deep CNN model trained by the ImageNet database, which was then retrained using the underwater cable images. Table 2 shows the results of all five deep CNN models, whereby the classification accuracy is low and yields the following values: MobileNet (65.50%), MobileNet V2 (67.50%), Inception V3 (53.00%), Xception (49.00%), and Inception-ResNet-V2 (50.00%). It can be observed from the table that the computational time of all models was below 1 h, which shows in the following values: MobileNet (0.11 h), MobileNet V2 (0.16 h), Inception V3 (0.33 h), Xception (0.39 h), and Inception-ResNet-V2 (0.72 h).

Figure 10 shows the confusion matrix employed for classifying an image with or without an underwater cable. Therefore, MobileNet V2 correctly classified 74 images with underwater cable and 61 images without underwater cable out of 200 testing images.

	Accuracy (%)	Precision Value			Recall Value				F1 Score		
CNN Models		Cable	Without Cable	Average	Cable	Without Cable	Average	Cable	Without Cable	Average	Time (Hours)
MobileNet	65.50	0.653	0.657	0.655	0.66	0.660	0.650	0.657	0.658	0.657	0.11
MobileNet V2	67.50	0.655	0.701	0.678	0.74	0.740	0.610	0.695	0.720	0.707	0.16
Inception V3	53.00	0.528	0.533	0.530	0.57	0.570	0.490	0.548	0.551	0.549	0.33
Xception	49.00	0.493	0.485	0.489	0.66	0.660	0.320	0.564	0.559	0.562	0.39
Inception-ResNet-V2	50.00	0.000	0.500	0.250	0.00	1.000	0.500	0.000	0.667	0.333	0.72

Table 2. Table of overall accuracy, precision value, recall value, f1 score, and computational time of five deep CNN models in underwater cable tracking by using the fine-tuning technique.



Figure 10. Confusion matrix of five deep CNN models in underwater cable tracking by using the fine-tuning technique listed as: (**a**) MobileNet; (**b**) MobileNet V2; (**c**) Inception V3; (**d**) Xception; (**e**) Inception-ResNet-V2.

The low classification accuracy observed for the five deep CNN models is attributable to their inability to learn the underwater cable images from scratch. When fine-tuning was applied, all of their layers were frozen, except for the last layer that was used to classify the underwater cable images. These frozen layers are important to perform low-level, mid-level, and high-level extraction from the underwater cable images as vast differences are present between the ImageNet images and underwater images used in this study. When the layers are frozen, the last layer is unable to get useful information from its previous layers to perform cable tracking, thereby yielding the overall performance that is less than 90% accurate. According to Cetinic et al. [54], the performance of deep CNN model is the lowest among all transfer learning techniques when fine-tuning is applied as it freezes all layers except the last layer, thereby resulting in the low similarity of the source and target domains. Consequently, the fine-tuning technique is noted to be weak in classifying such underwater cable images.

3.1.2. Deep Feature Learning

In this section, the results of deep feature learning technique implementation to train five deep CNN models in classifying underwater images are provided in Table 3. In general, the classification accuracy for the models was improved compared to the previous approach, yielding the following values: MobileNet (89.50%), MobileNetV2 (88.50%), Inception V3 (85.50%), Xception (88.50%), and Inception-ResNet-V2 (87.50%). Furthermore, the overall computational time for the deep learning feature was higher than the fine-tuning technique, whereby MobileNet revealed the highest classification accuracy in the classification of images, among others. It successfully and correctly classified 84 images with underwater cable and 95 images without underwater cable out of 200 testing images. All the confusion matrix of deep CNN models can be observed in Figure A1.

The superior performance of the deep feature learning compared to the fine-tuning technique in the classification of underwater cable images was attributable to the models' ability to learn from scratch. Besides, the technique utilized all layers of the five deep CNN models to learn useful information and features from the training data, allowing them to learn the patterns and classify the underwater cable from the underwater images. Based on these results, the deep feature learning technique is found to be better than fine-tuning in classifying underwater cable images. However, its longer computation time is a specific drawback. In the next subsection, the results of fine-tuning and deep feature learning with data augmentation are presented.

3.2. Performance of Deep CNN Models with Data Augmentation

In general, 20,000 training data were generated via data augmentation and then used to train the deep CNN models, following which fine-tuning and deep feature learning were both applied to train them using the augmented data. Accordingly, the model performance was slightly improved when the fine-tuning technique with data augmentation was applied for all models, except for Inception-ResNet-V2, whose performance was unchanged. The classification accuracy of all deep CNN models is presented in Table 4. Meanwhile, the computational time for training the models yielded a large increment when subjected to more data. In particular, deep feature learning when combined with augmented data was applied to train the deep CNN models, thus yielding superior results compared to other techniques. In fact, the classification accuracy for the deep CNN models was the highest among all experiments done prior, resulting in the following values: MobileNet (91.50%), MobileNetV2 (93.50%), Inception V3 (90.50%), Xception (91.00%), and Inception-ResNet-V2 (91.50%). However, the training time for the models also increases by a huge margin shown in Table 4. All the visual performances of deep CNN models trained with fine-tuning with augmented data and deep feature learning with augmented data were in Figures A2 and A3.

	Accuracy (%)	Precision Value			Recall Value				F1 Score		
CNN Models		Cable	Without Cable	Average	Cable	Without Cable	Average	Cable	Without Cable	Average	Time (Hours)
MobileNet	89.50	0.944	0.856	0.900	0.840	0.950	0.895	0.889	0.900	0.895	0.55
MobileNet V2	88.50	0.905	0.867	0.886	0.860	0.910	0.885	0.882	0.888	0.885	0.58
Inception V3	85.50	0.933	0.847	0.890	0.830	0.940	0.885	0.878	0.891	0.885	2.08
Xception	88.50	0.832	0.882	0.857	0.890	0.820	0.855	0.860	0.850	0.855	2.53
Inception-ResNet-V2	87.50	0.921	0.838	0.880	0.820	0.930	0.875	0.868	0.882	0.875	4.67

Table 3. Table of accuracy, precision value, recall value, f1 score, and computational time of five deep CNN models in underwater cable tracking with the deep feature learning technique.

Table 4. Table of accuracy, precision value, recall value, f1 score, and computational time of five deep CNN models in underwater cable tracking by using fine-tuning and deep feature learning with data augmentation.

	Deep CNN Models	Accuracy (%)	Precision Value			Recall Value			F1 Score			
Technique			Cable	Without Cable	Average	Cable	Without Cable	Average	Cable	Without Cable	Average	Time (Hours)
	MobileNet	69.00	0.679	0.702	0.691	0.720	0.660	0.690	0.699	0.680	0.690	5.00
Fine-Tuning	MobileNet V2	68.50	0.703	0.670	0.687	0.640	0.730	0.685	0.670	0.699	0.684	5.55
	Inception V3	49.50	0.000	0.497	0.249	0.000	0.990	0.495	0.000	0.662	0.331	6.67
	Xception	53.00	0.517	0.607	0.562	0.890	0.170	0.530	0.654	0.266	0.460	6.94
	Inception-ResNet-V2	50.00	0.000	0.500	0.250	0.000	1.000	0.500	0.000	0.667	0.333	10.56
	MobileNet	91.50	0.928	0.903	0.915	0.900	0.930	0.915	0.914	0.916	0.915	7.89
Deep Feature Learning	MobileNet V2	93.50	0.939	0.931	0.935	0.930	0.940	0.935	0.935	0.935	0.935	8.75
	Inception V3	90.50	0.893	0.918	0.905	0.920	0.890	0.905	0.906	0.904	0.905	28.10
	Xception	91.00	0.902	0.918	0.910	0.920	0.900	0.910	0.911	0.909	0.910	36.00
	Inception-ResNet-V2	91.50	0.928	0.903	0.915	0.900	0.930	0.915	0.914	0.916	0.915	66.60

In this study, 20,000 training datasets were generated via data augmentation and used to train the five deep CNN models for learning the difference between images with and without underwater cables, which could further improve the performance of cable tracking despite requiring more computational time. Here, the classification accuracy of deep CNN models trained with augmented data was improved in comparison with the use of fine-tuning or deep feature learning singular. Regardless, their accuracy was still lower than 90%, which was the accuracy for the fine-tuning with the augmented data approach. This is due to the technique's limitation that causes the deep CNN models to be unable to learn important information even though they are trained with that amount of data. In the context of deep feature learning, the models were pre-trained with the data from ImageNet before being trained with the underwater images, thus showing better performance than fine-tuning in classifying the images. Therefore, this proves the ability of data augmentation to optimize the performance of deep CNN models. In addition, some experiments have been done to optimize the model performance in recognizing underwater objects, which is done by increasing the underwater images. The results obtained have proven the concept by yielding an improved performance of the deep CNN models for underwater object recognition [14,55].

Figure 11 shows the training accuracy and validation accuracy of Inception-Resnet-V2, whereas those of the remaining models are included in Figure A4. The figure clearly shows the improved training and validation accuracy in the presence of an abundance of data. Another important observation from Figures 11 and A4 was the improved stability of the model when more data were used in training and validation. For example, fine-tuning with an abundance of data, the training and validation accuracy of Inception-ResNet-V2 have not fluctuated as fine-tuning singular. The higher validation accuracy indicates the higher expected performance of deep CNN models when subjected to the testing dataset, whereas the higher training and validation accuracy both also prove the importance of data for deep-learning-based techniques. In the study of Sajjad et al. [47], a comparison of validation accuracy with data augmentation is shown to be higher than that without data augmentation. Therefore, the inclusion of the data augmentation technique improves the performance of deep CNN models in classifying the underwater cable but it requires a longer computational time.



Figure 11. Graph of training and validation accuracy versus the number of epochs (Inception-ResNet-V2). Note: FT, DFL, and DA stand for fine-tuning, deep feature learning, and data augmentation.

3.3. Proposal of a Suitable Deep CNN Model for the Classification of Underwater Cable from Images

This study proposed several deep CNN models to perform the classification of underwater cable, whereby different optimization techniques were applied to increase their respective performance. Figure 12 shows that deep CNN models trained with deep feature learning and data augmentation

yield the highest performance in classifying underwater cable images compared to other techniques. Meanwhile, Figure 12 details the longer computational time obtained by deep feature learning compared to fine-tuning. Besides, data augmentation inclusion increased computational time significantly. Furthermore, the performance of MobileNetV2 was improved from 67.50% (i.e., fine-tuning) to 68.50% (i.e., fine-tuning with data augmentation) and from 88.50% (i.e., deep feature learning) to 93.50% (i.e., deep feature learning with data augmentation). Meanwhile, its computational time was increased from 0.16 h (i.e., fine-tuning) to 5.55 h (i.e., fine-tuning with data augmentation) and from 0.58 h (i.e., deep feature learning) to 8.75 h (i.e., deep feature learning with data augmentation).



■ MobileNet ■ MobileNet V2 ■ Inception V3 ■ Xception ■ Inception Resnet V2

Figure 12. Graph of classification accuracy and computational time of deep CNN models against optimization technique. Note: FT, DFL, and DA stand for fine-tuning, deep feature learning, and data augmentation.

Optimization Techniques

Among all deep CNN models proposed in this study, it showed the highest performance in classifying underwater cable images when trained with deep feature learning and data augmentation. Hence, MobileNetV2 with deep feature learning and data augmentation is thus suggested for the classification of underwater cable images. Previously, the study of Valentini and Balouin [37] has also utilized the same model with transfer learning and data augmentation to perform algae detection by using low-cost smartphone-based images.

In another experiment, the deep CNN models have been trained and optimized via hyperparameter tuning. The training of such models involves the selection of many critical and essential hyperparameters, which can impact the model performance in the context of its accuracy and computational time to a significant extent [56]. Therefore, deep CNN models should have a suitable set up of hyperparameter configuration to obtain a relatively low training time while conserving their high classification performance concurrently. Accordingly, the selected hyperparameters are learning rate, number of patience for early stopping, and batch size, whereby a comparison of the performance metrics to construct a suitable configuration of hyperparameters will lead to better performance and lower computational time. Here, the suggested ranges for the learning rate, number of patience for early stopping, and batch size are within the range of 0.1 to 0.00001, 0 to 50, and 10 to 50, respectively. However, their selection is dependent on the data and hardware used for the experiment.

4. Conclusions

This study proposed using a deep learning method for the classification of underwater cable images from underwater images, which arose due to the challenges of underwater cable tracking by

using the traditional methods. This is due to the large number of different underwater cable types available, variation of lighting underwater causing difficulties for the camera to capture underwater cable images, and the presence of algae and sand covering the underwater cables that increase the chance of misclassification. However, deep learning methods have typically performed greatly in the underwater conditions for identifying underwater cable.

In this study, several deep CNN models were chosen to perform the classification of underwater cable images, whereby transfer learning and data augmentation were implemented to enhance the model performance in classifying underwater cable images from underwater images. Among the deep CNN models evaluated, MobileNetV2 yielded the best performance valued at 93.5% when deep feature learning and data augmentation techniques were applied. Furthermore, the advantages of the deep learning method include its capability to identify underwater cable in any situation when a large volume of images related to the cable is provided to train the deep CNN model. In contrast, its drawbacks consist of more computational power required and expensive GPUs necessary to process the large amount of data and complex data models. Based on the experiment results, it can be concluded that the deep learning method is powerful and highly accurate in the classification of underwater cable images. Accordingly, the contribution of this study lies in the development of a deep learning method to perform underwater cable image classification. For future works, researchers may opt to further improve the deep CNN models by localizing the position of underwater cable in the images. It is suggested for them to apply other types of deep learning approaches, such as Region-based CNN, Fast Region-based CNN, and Mask Region-based CNN. In some of the fields, deep learning had been used to perform defect detection such as road crack, building crack, etc. It is also suggested for researchers to collect a new image dataset to train the deep CNN model to perform an inspection of underwater cable, power cable, or underwater pipeline. The constraint for both suggestions is the collection of suitable image data as deep learning required a huge amount of data for it to capture those features that researchers are interested in.

Author Contributions: Conceptualization, G.W.T., S.H.T., S.A.A. and M.A.; Methodology, G.W.T.; Software, G.W.T.; Supervision, S.H.T. and S.A.A.; Validation, S.H.T. and S.A.A.; Writing—original draft, G.W.T. and M.A.; Writing—review & editing, S.H.T., S.A.A. and M.A. All authors have read and agreed to the published version of the manuscript.

Funding: The open-access preparation of this article was funded by the Research Centre of University Putra Malaysia. **Conflicts of Interest:** The authors declare no conflict of interest.

Appendix A



Figure A1. Cont.



Figure A1. Confusion Matrix of five deep CNN models in underwater cable tracking by using the deep feature learning technique listed as: (a) MobileNet; (b) MobileNet V2; (c) Inception V3; (d) Xception; (e) Inception-ResNet-V2.



Figure A2. Cont.



Figure A2. Confusion matrix of five deep CNN models in underwater cable tracking by using fine-tuning technique with data augmentation listed as: (a) MobileNet; (b) MobileNet V2; (c) Inception V3; (d) Xception; (e) Inception-ResNet-V2.



Figure A3. Cont.



Figure A3. Confusion Matrix of five CNN models in underwater cable tracking by using deep feature learning with data augmentation listed as: (a) MobileNet; (b) MobileNet V2; (c) Inception V3; (d) Xception; (e) Inception-ResNet-V2.



Figure A4. Graph of training and validation accuracy versus number of epochs. Note: FT, DFL and DA stand for fine-tuning, deep feature learning and data augmentation. (a) MobileNet; (b) MobileNet V2; (c) Inception V3; (d) Xception.

References

- 1. Detecon Asia-Pacific Ltd.; Christof Gerlach, R.S.A.C. *Economic Impact of Submarine Cable Disruptions*; APEC Secretariat: Singapore, 2013; p. 96.
- Xu, C.; Chen, J.; Yan, D.; Ji, J. Review of underwater cable shape detection. *J. Atmos. Ocean. Technol.* 2016, 33, 597–606. [CrossRef]
- 3. Mohamed, N.; Jawhar, I.; Al-Jaroodi, J.; Zhang, L. Sensor network architectures for monitoring underwater pipelines. *Sensors* **2011**, *11*, 10738–10764. [CrossRef] [PubMed]
- 4. Kraus, C.; Carter, L. Seabed recovery following protective burial of subsea cables—Observations from the continental margin. *Ocean Eng.* **2018**, *157*, 251–261. [CrossRef]
- 5. Clark, K. Submarine Telecoms Industry Report; SubTel Forum Press: Sterling, VA, USA, 2019; p. 118.
- 6. Carter, L.; Gavey, R.; Talling, P.J.; Liu, J.T. Insights into submarine geohazards from breaks in subsea telecommunication cables. *Oceanography* **2014**, 27, 58–67. [CrossRef]
- 7. Torres, L.; Jiménez-Cabas, J.; González, O.; Molina, L.; López-Estrada, F.R. Kalman filters for leak diagnosis in pipelines: Brief history and future research. *J. Mar. Sci. Eng.* **2020**, *8*, 173. [CrossRef]
- 8. Ortiz, A.; Antich, J.; Oliver, G. A particle filter-based approach for tracking undersea narrow telecommunication cables. *Mach. Vis. Appl.* **2011**, *22*, 283–302. [CrossRef]
- 9. Allibert, G.; Hua, M.D.; Krupínski, S.; Hamel, T. Pipeline following by visual servoing for autonomous underwater vehicles. *Control Eng. Pract.* **2019**, *82*, 151–160. [CrossRef]
- Khan, A.; Ali, S.S.A.; Meriaudeau, F.; Malik, A.S.; Soon, L.S.; Seng, T.N. Visual feedback–based heading control of autonomous underwater vehicle for pipeline corrosion inspection. *Int. J. Adv. Robot. Syst.* 2017, 14. [CrossRef]
- 11. Fatan, M.; Daliri, M.R.; Mohammad Shahri, A. Underwater cable detection in the images using edge classification based on texture information. *Meas. J. Int. Meas. Confed.* **2016**, *91*, 309–317. [CrossRef]
- Horgan, J.; Toal, D. Review of machine vision applications in unmanned underwater vehicles. In Proceedings of the 9th International Conference on Control, Automation, Robotics and Vision, Singapore, 5–8 December 2006. [CrossRef]
- Kuhn, V.N.; Drews, P.L.J.; Gomes, S.C.P.; Cunha, M.A.B.; Botelho, S.S.d.C. Automatic control of a ROV for inspection of underwater structures using a low-cost sensing. *J. Braz. Soc. Mech. Sci. Eng.* 2014, 37, 361–374. [CrossRef]
- 14. Sun, X.; Shi, J.; Liu, L.; Dong, J.; Plant, C.; Wang, X.; Zhou, H. Transferring deep knowledge for object recognition in Low-quality underwater videos. *Neurocomputing* **2018**, 275, 897–908. [CrossRef]
- 15. Wang, K.; Hu, Y.; Chen, J.; Wu, X.; Zhao, X.; Li, Y. Underwater image restoration based on a parallel convolutional neural network. *Remote Sens.* **2019**, *11*, 1591. [CrossRef]
- 16. Lee, D.; Kim, G.; Kim, D.; Myung, H.; Choi, H.T. Vision-based object detection and tracking for autonomous navigation of underwater robots. *Ocean Eng.* **2012**, *48*, 59–68. [CrossRef]
- 17. Ortiz, A.; Simó, M.; Oliver, G. A vision system for an underwater cable tracker. *Mach. Vis. Appl.* **2002**, *13*, 129–140. [CrossRef]
- 18. Antich, J.; Ortiz, A. Underwater cable tracking by visual feedback. *Lect. Notes Comput. Sci.* **2003**, 2652, 53–61. [CrossRef]
- 19. Balasuriya, A.; Ura, T. Vision-based underwater cable detection and following using AUVs. In Proceedings of the OCEANS '02 MTS/IEEE, Biloxi, MI, USA, 29–31 October 2002; Volume 3, pp. 1582–1587. [CrossRef]
- 20. Chen, H.H.; Chuang, W.N.; Wang, C.C. Vision-based line detection for underwater inspection of breakwater construction using an ROV. *Ocean Eng.* **2015**, *109*, 20–33. [CrossRef]
- 21. Wirth, S.; Ortiz, A.; Paulus, D.; Oliver, G. Using particle filters for autonomous underwater cable tracking*. *IFAC Proc. Vol.* **2008**, *41*, 161–166. [CrossRef]
- 22. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [CrossRef]
- Ohn-Bar, E.; Trivedi, M.M. To boost or not to boost? On the limits of boosted trees for object detection. In Proceedings of the 23rd International Conference on Pattern Recognition (ICPR), Cancún, Mexico, 4–8 December 2016; pp. 3350–3355. [CrossRef]
- 24. Yu, P.; Zhao, Y.; Zhang, J.; Xie, X. Pedestrian detection using multi-channel visual feature fusion by learning deep quality model. *J. Vis. Commun. Image Represent.* **2019**, *63*, 102579. [CrossRef]

- Jeon, M.; Lee, Y.; Shin, Y.S.; Jang, H.; Kim, A. Underwater Object Detection and Pose Estimation using Deep Learning. In *IFAC-PapersOnLine*; Elsevier: Amsterdam, The Netherlands, 2019; Volume 52, pp. 78–81. [CrossRef]
- Villon, S.; Mouillot, D.; Chaumont, M.; Darling, E.S.; Subsol, G.; Claverie, T.; Villéger, S. A Deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecol. Inform.* 2018, 48, 238–244. [CrossRef]
- 27. O'Byrne, M.; Pakrashi, V.; Schoefs, F.; Ghosh, B. Semantic segmentation of underwater imagery using deep networks trained on synthetic imagery. *J. Mar. Sci. Eng.* **2018**, *6*, 93. [CrossRef]
- 28. Kaushal, M.; Khehra, B.S.; Sharma, A. Soft Computing based object detection and tracking approaches: State-of-the-Art survey. *Appl. Soft Comput.* **2018**, *70*, 423–464. [CrossRef]
- 29. Zhao, Z.Q.; Zheng, P.; Xu, S.T.; Wu, X. Object Detection with Deep Learning: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [CrossRef] [PubMed]
- Pathak, A.R.; Pandey, M.; Rautaray, S. Application of deep learning for object detection. *Procedia Comput. Sci.* 2018, 132, 1706–1717. [CrossRef]
- Valdenegro-Toro, M. Object recognition in forward-looking sonar images with convolutional neural networks. In Proceedings of the OCEANS 2016 MTS/IEEE Monterey, Monterey, CA, USA, 19–23 September 2016; pp. 1–6. [CrossRef]
- Kvasić, I.; Mišković, N.; Vukić, Z. Convolutional Neural Network Architectures for Sonar-Based Diver Detection and Tracking. In Proceedings of the OCEANS 2019, Marseille, France, 17–20 June 2019. [CrossRef]
- 33. Jalal, A.; Salman, A.; Mian, A.; Shortis, M.; Shafait, F. Fish detection and species classification in underwater environments using deep learning with temporal information. *Ecol. Inform.* **2020**, 57. [CrossRef]
- 34. Buetti-Dinh, A.; Galli, V.; Bellenberg, S.; Ilie, O.; Herold, M.; Christel, S.; Boretska, M.; Pivkin, I.V.; Wilmes, P.; Sand, W.; et al. Deep neural networks outperform human expert's capacity in characterizing bioleaching bacterial biofilm composition. *Biotechnol. Rep.* **2019**, *22*, e00321. [CrossRef]
- 35. Gómez-Ríos, A.; Tabik, S.; Luengo, J.; Shihavuddin, A.S.M.; Krawczyk, B.; Herrera, F. Towards highly accurate coral texture images classification using deep convolutional neural networks and data augmentation. *Expert Syst. Appl.* **2019**, *118*, 315–328. [CrossRef]
- 36. Tamou, A.B.; Benzinou, A.; Nasreddine, K.; Ballihi, L. Transfer learning with deep convolutional neural network for underwater live fish recognition. In Proceedings of the IEEE International Conference on Image Processing, Applications and Systems (IPAS), Sophia Antipolis, France, 12–14 December 2018. [CrossRef]
- 37. Valentini, N.; Balouin, Y. Assessment of a smartphone-based camera system for coastal image segmentation and sargassum monitoring. *J. Mar. Sci. Eng.* **2020**, *8*, 23. [CrossRef]
- Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *Comput. Sci.* 2017, arXiv:1704.04861.
- 39. Kannojia, S.P.; Jaiswal, G. Effects of varying resolution on performance of CNN based image classification an experimental study. *Int. J. Comput. Sci. Eng.* **2018**, *6*, 451–456. [CrossRef]
- 40. Wang, N.; Wang, Y.; Er, M.J. Review on deep learning techniques for marine object recognition: Architectures and algorithms. *Control Eng. Pract.* **2020**. [CrossRef]
- 41. Fu, J.; Rui, Y. Advances in deep learning approaches for image tagging. *APSIPA Trans. Signal Inf. Process.* **2017**, *6*. [CrossRef]
- Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520. [CrossRef]
- Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826. [CrossRef]
- Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807. [CrossRef]
- Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-ResNet and the impact of residual connections on learning. In Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI 2017), San Francisco, CA, USA, 4–9 February 2017; pp. 4278–4284.

- Kaya, A.; Keceli, A.S.; Catal, C.; Yalic, H.Y.; Temucin, H.; Tekinerdogan, B. Analysis of transfer learning for deep neural network based plant classification models. *Comput. Electron. Agric.* 2019, 158, 20–29. [CrossRef]
- 47. Sajjad, M.; Khan, S.; Muhammad, K.; Wu, W.; Ullah, A.; Baik, S.W. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *J. Comput. Sci.* **2019**, *30*, 174–182. [CrossRef]
- 48. Xu, Y.; Zhang, Y.; Wang, H.; Liu, X. Underwater image classification using deep convolutional neural networks and data augmentation. In Proceedings of the 2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Xiamen, China, 22–25 October 2017. [CrossRef]
- Noh, J.M.; Jang, G.R.; Ha, K.N.; Park, J.H. Data Augmentation Method for Object Detection in Underwater Environments. In Proceedings of the 19th International Conference on Control, Automation and Systems, Jeju, Korea, 15–18 October 2019; pp. 324–328. [CrossRef]
- 50. Huang, H.; Zhou, H.; Yang, X.; Zhang, L.; Qi, L.; Zang, A.-Y. Faster R-CNN for marine organisms detection and recognition using data augmentation. *Neurocomputing* **2019**, *337*, *372–384*. [CrossRef]
- 51. Bloice, M.D.; Stocker, C.; Holzinger, A. Augmentor: An image augmentation library for machine learning. *J. Open Source Softw.* **2017**, *2*, 432. [CrossRef]
- 52. Pattanayak, S. Intelligent Projects Using Python: 9 Real-World AI Projects Leveraging Machine Learning and Deep Learning with TensorFlow and Keras; Packt Publishing: Birmingham, UK, 2019; pp. 64–65.
- 53. Choi, H.; Park, M.; Son, G.; Jeong, J.; Park, J.; Mo, K.; Kang, P. Real-time significant wave height estimation from raw ocean images based on 2D and 3D deep neural networks. *Ocean Eng.* **2020**, 201, 107129. [CrossRef]
- 54. Cetinic, E.; Lipic, T.; Grgic, S. Fine-tuning convolutional neural networks for fine art classification. *Expert Syst. Appl.* **2018**, *114*, 107–118. [CrossRef]
- 55. O'Byrne, M.; Ghosh, B.; Schoefs, F.; Pakrashi, V. Applications of virtual data in subsea inspections. *J. Mar. Sci. Eng.* **2020**, *8*, 328. [CrossRef]
- 56. Zhou, S.; Song, W. Deep learning-based roadway crack classification using laser-scanned range images: A comparative study on hyperparameter selection. *Autom. Constr.* **2020**, *114*, 103171. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).