water

MDPI

Article

# Deep Learning Based Approach to Classify Saline Particles in Sea Water

**Mohammed Alshehri [1,*], Manoj Kumar [2], Akashdeep Bhardwaj [2], Shailendra Mishra [3] and Jayadev Gyani [3]**

1 Department of Information Technology, College of Computer and Computer and Information Sciences, Majmaah University, Majmaah 11952, Saudi Arabia
2 School of Computer Science, University of Petroleum and Energy Studies, Bidholi, Dehradun 248007, India; wss.manojkumar@gmail.com (M.K.); bhrdwh@yahoo.com (A.B.)
3 Department of Computer Engineering, College of Computer and Computer and Information Sciences, Majmaah University, Majmaah 11952, Saudi Arabia; s.mishra@mu.edu.sa (S.M.); je.gyani@mu.edu.sa (J.G.)
* Correspondence: ma.alshehri@mu.edu.sa

**Abstract:** Water is an essential resource that facilitates the existence of human life forms. In recent years, the demand for the consumption of freshwater has substantially increased. Seawater contains a high concentration of salt particles and salinity, making it unfit for consumption and domestic use. Water treatment plants used to treat seawater are less efficient and reliable. Deep learning systems can prove to be efficient and highly accurate in analyzing salt particles in seawater with higher efficiency that can improve the performance of water treatment plants. Therefore, this work classified different concentrations of salt particles in water using convolutional neural networks with the implementation of transfer learning. Salt salinity concentration images were captured using a designed Raspberry Pi based model and these images were further used for training purposes. Moreover, a data augmentation technique was also employed for the state-of-the-art results. Finally, a deep learning neural network was used to classify saline particles of varied concentration range images. The experimental results show that the proposed approach exhibited superior outcomes by achieving an overall accuracy of 90% and f-score of 87% in classifying salt particles. The proposed model was also evaluated using other evaluation metrics such as precision, recall, and specificity, and showed robust results.

**Keywords:** classification; deep learning; convolutional neural networks; transfer learning; saline particles; salinity



**Citation:** Alshehri, M.; Kumar, M.; Bhardwaj, A.; Mishra, S.; Gyani, J. Deep Learning Based Approach to Classify Saline Particles in Sea Water. *Water* **2021**, *13*, 1251. https://doi.org/10.3390/w13091251

Academic Editor: Donghwi Jung

Received: 31 March 2021
Accepted: 27 April 2021
Published: 29 April 2021



## 1. Introduction

Water is an essential element for the survival of different life forms existing on Earth. Freshwater is consumed across a diverse array of applications such as agriculture [1], health care [2], domestic use, industrial applications [3,4], etc. The present freshwater sources such as lakes, ponds, and reservoirs are becoming dehydrated and disappearing at an extreme rate. Furthermore, with the increase in the number of human populations, freshwater consumption has increased every year, leading to scarcity and massive demand for freshwater sources. With advanced industry evolution and globalization, different issues regarding water, pollution, environment, and energy, etc. being overexploited has attracted extensive attention. As important parameters, water resources are assessed at national, international, urban, and industrial scales [5]. With water shortages, desalination treatment as a water production system that can be used for freshwater supply from seawater is therefore a solution for water resource conservation. For seawater desalination, researchers have conducted various studies and focused toward replacing old energy with renewable energy to improve desalination techniques. Membrane process or reverse osmosis (RO) is one of such techniques used for desalination and is widely used. Practices based on RO shows better thermodynamic performances compared to thermal scales.

These solutions use more energy consumption and have a lower recovery rate [6]. More of these works are discussed further in Section 2 of the presented study.

The amount of fresh water has been nearly the same over the years, but due to rapid increase of the population on Earth, it has led to a greater consumption of clean water every year. Human action, competition for resources, and rapid growth in industrialization affect the natural health of the atmosphere. Pollutants are mainly contributed by industrial lime, oil combustion, agriculture waste, and anthropogenic activities. Many pollutants are transferred by the wind directions from the industrial localities and national highways. In dense cities, gas and toxic substances systematically spread in the atmosphere. Although their natural concentrations seldom touch dangerous levels, they can be harmful to human health and to the environment at high concentrations. Pollutants can be transferred in water in dissolved form either in the surface or ground. With fresh water resources dwindling, seawater desalination is expected to become increasingly important in meeting Saudi Arabia's future water needs. Desalination, on the other hand, is a more energy-intensive process than conventional water treatment processes. Saudi Arabia is said to use 25% of domestic oil and gas production in desalination plants, with that percentage expected to rise to 50% by 2020.

The possible sources of deriving freshwater are surface water and seawater bodies. About 97% of the water on Earth exists in the form of seawater and oceans. However, these sources contain a significant amount of total dissolved solids (TDS) that ranges from 3000 mg/L to 4000 mg/L [7], which makes it unfit for human consumption and the agricultural sector. Further to this, seawater contains high mineral content and elements such as $Na^{+}$, $K^{+}$, $Ca^{2+}$, $Mg^{2+}$, and ($SO_4$), major contributors to making seawater saline. This identifies saline particles as a vital task, so that it is appropriately treated with great accuracy using suitable desalinated seawater methods. This research study presents an attempt to use an artificial intelligence deep learning framework to identify saline particles in different salt concentrations using a scattering pattern. Desalination is treated as one of the extreme solutions for water supply. The discussion should be focused on its affordability and cleaning techniques, while some of the techniques in the past have addressed technological problems. The provided solution by this study will focus on reducing the cost to classify the salinity of water. Therefore, the main contribution is to use current edge-based techniques to classify salinity from any form of water with minimal efforts. Even though desalination and water reuse can be used as autonomous solutions to water shortage, the role of desalination classification is an effort to avoid future water scarcity problems [8].

To address these research issues and gaps, a unique deep learning system was implemented to capture the concentration of the salt particles. There are no such works that have discussed the detection of seawater saline particles and classification using deep learning for implementing efficient and accurate desalination techniques. A detailed analysis for addressing salinity in seawater is presented in further sections.

**Objectives:**

Moreover, this paper examines and addresses the following key points:

- Contaminants such as salt particles present in water cause hardness and make it unfit for consumption. Identifying saline particles in the water will help water treatment plants treat water effectively with high performance.
- The proposed approach aims to target seawater, a rich source of salt particles, using a deep learning system.
- The presented study involves experimentation and the analysis of different saltwater concentrations to provide the system's robustness against high saline concentrations.
- The proposed work is evaluated using various classification measuring metrics such as precision, recall, accuracy scores, etc.
- Conventional water treatment plants are less accurate and efficient in treating seawater. Integrating the proposed framework with water treatment plants can boost their operational and functional capabilities, resulting in a better seawater treatment process.

The rest of this paper is organized as follows. Section 2 discusses the related works in this domain. Section 3 presents the proposed framework used for implementing the desired deep learning model. Results and discussion are presented in Section 4 while the conclusions from the work and future directions are discussed in Section 5.

## 2. Literature Work

Numerous methods have been proposed and presented by researchers worldwide to predict salinity and harmful elements in saline water. Some of the best-in-class techniques proposed are discussed below.

Melesse et al. [9] proposed a hybrid approach to predicting river water salinity in the Babol-Rood River. The input variables essential for the machine learning (ML) model were predicted using Pearson's correlation coefficient method. Furthermore, the main predictor used for predicting the salinity in river water was TDS. The designed ML models were evaluated using various regression estimation metrics. The demonstrated results showed that the presented approach outperformed the other popularly known techniques by achieving high-performance rates. Banerjee et al. [10] predicted the salinity forecast of groundwater based on pumping rate using ANN with quick propagation. The advantage of using quick propagation for calculating the derivatives is that it uses an adaptive learning rate for every connection weight used during the training of the network. The implemented model was compared with the conventional statistical saturated-unsaturated (SUTRA) computational model. The model also estimated the standard baseline for consuming groundwater below the pumping rate of 13,000 L/day, so the groundwater salinity should be maintained within 2.5%. Barzegar and Moghaddam [11] estimated groundwater salinity using artificial neural networks (ANN). The proposed model predicted groundwater salinity measured in terms of electrical conductivity over various input concentrations such as $Cl^-$, $Na^+$, $Mg^{2+}$, etc. The benefits of using ANN for predicting salinity is that it handles a non-linear relationship between different variables effectively and detects robust patterns in the data to provide correct model predictions. Furthermore, the weights of the implemented ANN were estimated using the genetic algorithm (GA), an evolutionary optimization approach that optimizes the correct weights. The presented approach was compared with other popularly known techniques. The results demonstrated that the model outperformed the remaining approaches by achieving great results. A hybrid evolutionary approach was presented by Azad et al. [12] for the prediction of water quality at three river stations. Their research was based on predicting the water quality based on five parameters, namely, electrical conductivity (EC), sodium adsorption ratio (SAR), total dissolved solids (TDS), carbonate hardness (CH), and total hardness (TH). The proposed approach was evaluated on regression metrics that demonstrated exemplary performance with high accuracy. Duong et al. [13] optimized the air-gap membrane distillation system for seawater desalination. They reduced the usage of heat and electricity during the desalination process, but they did not provide a solution of classification of the saline particles. This technique directly impacts the energy consumption and therefore poor accuracy of the assessment results. Doornbusch et al. [14] investigated the solution for multivalent ions in seawater on the desalination performance of multistage electrodialysis. They used natural seawater as feed solution, and conventional cation exchange membranes (CEMs) as well as CEMs with preferential removal of multivalent ions were compared. Their technique results were steady for 18 days with an average energy usage of 3 kWh/m$^3$, which showed the effectiveness of multistage electrodialysis seawater desalination.

Furthermore, the proposed approach was compared with other evolutionary approaches and proven to be a reliable technique for estimating water quality. An efficient and reliable solution to measure the water's salinity was proposed by Hussain et al. [15]. They estimated the salinity of water using a smartphone. Ambient light sensor (ALS) and embedded flash lamps were used to detect and detect seawater salinity. Furthermore, their research study involved the analysis of multiple approaches for identifying saline particles. Moreover, its compact size and inexpensive developmental cost can serve as an alternative to less efficient salinometers. The simulated results were verified using various measuring metrics such as sensitivity, reproducibility, and dynamic range, which exhibited superior results. In addition to the prediction of water quality, it is vital to treat the saline water to be utilized in day-to-day tasks. Some of the modern and recent techniques used to model this purpose can be discussed as follows: Tufa et al. [16] presented a membrane-based hybrid technique for water desalination. They used the membrane distillation (MD) process to increase the overall recovery rate of water by lowering the discharged brine volume. Furthermore, MD brine was passed through reverse electro-dialysis (RED) to generate saline gradient power, which produced energy used to derive the high-pressure pumps at the reverse osmosis (RO) stage. The implemented technique serves as a robust and energy-efficient desalination approach that causes less greenhouse gas emissions and the generation of low-grade waste heat that functions as an ideal source of energy for industrial applications. Maia et al. [17] conducted a review of studies related to solar chimneys for water desalination power generation. Solar chimneys use solar energy, which serves as a reliable source of energy that generates hot air that is used to derive turbines and generate power. This makes solar chimneys a cost-effective and reliable solution. However, it suffers from the disadvantages of low efficiency. To resolve this issue, hybrid solar chimney desalination systems can be implemented such as integrating photovoltaic and desalination based solar chimneys that effectively utilize solar power and give robust output. Suwaileh et al. [18] reviewed possible solutions based on membrane desalination and water reclamation for agriculture. Approaches such as reverse osmosis (RO), membrane bioreactor, electro dialysis, etc. are popularly known techniques for treating water. However, the quality of the water produced is not fit for agricultural use. The fusion of membrane-based techniques such as RO and forward osmosis (FO)-based integration systems can provide useful solutions by containing the necessary amount of nutrients required for fertigation. In arid and semi-arid regions such as Saudi Arabia, Iran, and Kuwait where freshwater sources are absent, it is essential to provide methods for water treatment and forecast water demands. Alkhudhiri et al. [19] estimated future demands for wastewater and seawater treatment using exponential smoothening and linear regression techniques. They predicted that the variation of treated water will be about 4% in the years 2025–2050. Additionally, the demand for water supply is expected to rise to 5% due to the growth in industrial sectors. Mansouri and Ghoniem [20] explored nuclear energy as an alternative to fossil fuel for powering water desalination. Nuclear energy serves as an efficient and cost-effective substitute to reduce fossil fuel consumption and emissions of carbon dioxide in the environment. Furthermore, analysis was conducted using the IAEA modeling tool to demonstrate that nuclear energy is cost-efficient and exhibits a more robust performance than fossil fuels. Additionally, analysis of the results showed that nuclear desalination using RO was proven to be an economical and efficient design for seawater desalination. Pan et al. [21] showed a study to estimate the RO and capacitive deionization (CDI) for brackish water desalination. They systematically summarized the technological information of RO and CDI, focusing on the effect of key parameters on desalination performance as well as energy-water efficiency, economic costs, and environmental impacts. This study showed that both RO and CDI play important roles in water and resource recovery from brackish water. Ayers and Westcot [22] proposed a technique for classification of water regarding its risk of toxicity according to the contents of sodium and chloride. This work objective was to study the quality of the irrigation water. The results of 45 complete analyses (electrical conductivity, pH, and concentrations of $Ca^{2+}$, $K^{+}$, $Cl^{-}$, $HCO^{3-}$, and $CO^{3z-}$) of water samples belonging to different places, located in the study region were used. All analyzed water samples presented low sodality and high alkalinity. Rose and Marry [23] showed an alternative by identifying the solvents present in the water body ensuring desalination of the particular ionic compound or metal. The primary objective was to classify the sensor data, viz, the salts in TDS. To assess the conductivity and TDS, 500 samples of various compositions were used. Machine learning was applied to classify the salts with the help of the K-nearest neighbor classifier. Taheran et al. [24] reviewed different techniques for PhAC removal by using membrane separation processes as these are highly used for quality drinking and industrial water. They showed that the osmosis membrane can proficiently eliminate practically all PhACs, however, its operational expense is generally high and nano-filtration (NF) layers are profoundly impacted by electrostatic and hydrophobic association. To improve the performance and robustness, it is proposed to consolidate membrane layers with different frameworks, like activated carbon and enzymatic degradation.

The examination of saline particles in seawater is a complicated task. Deep learning and various machine learning techniques are employed in many image processing and classification tasks and have provided superior results compared to the conventional approaches. Some of the studies involving the use of deep learning are discussed as follows. Huang et al. [25] proposed a deep learning system based on CNN with transfer learning to identify and detect underwater marine organisms. The proposed approach was tested against various factors such as marine turbulence, shooting angle variations, and illumination variations. Furthermore, three types of data augmentation techniques were used to evaluate the proposed technique. Some of the other best machine learning methods includes [26–31].

## 3. Research Methodology

The proposed research study involves the classification of saline particles in two stages. The first stage involves the preparation of scattered images for various ranges of salt salinity concentration using a Raspberry Pi device. These scattered images will serve as our training dataset. A detailed discussion regarding the creation and simulation of the dataset is discussed in Section 3.2. The second stage explained in Section 3.3 involves preprocessing and using the generated dataset to train deep learning neural network models to classify saline particles of varied ranges in appropriate classes. The steps that we followed as an algorithm is shown below (Algorithm 1):

**Algorithm 1**

Step 1: Salinity captured images are taken as input.
Step 2: Images are divided into training and testing of 7:3 proportion.
Step 3: Transfer learning based VGG16 deep learning model for salinity particles classification is designed with 19 weighted layers.
Step 4: Classification for 10 salt concentration particles is evaluated.
Step 5: Results are plotted using various performance ML based parameters.
Predictive analysis is done with other ML models.

### 3.1. Sample Preparation

The sample solution used for generating the scattering pattern was prepared by adding sodium chloride to a water-filled cuvette made up of glass in intervals of 10 ppt (parts per thousands). The minimum concentration of the salt solution was 0 ppt while the maximum saline solution was 100 ppt. Ten classes as tabulated in Table 1 were generated using the prepared samples. Since the concentration of salt particles in oceans or seawater ranges to about 35 ppt [32], therefore, the prepared solution acts as seawater, which can be used to classify the concentration of salt particles in seawater, which in turn could prove to be advantageous for desalinated water treatment plants.

### 3.2. Stage 1: Experimental Analysis

The simulation of stage 1 for generating scattered pattern images is demonstrated in Figure 1. The experiment used a ~1 mW red pointer laser diode that operated on a wavelength of 660 nm. Furthermore, the diode's light was focused on a water-filled cuvette that contained aa saline salt solution using a lens of 2.5 cm focal length. The scattered light from the saline particles was projected onto a white screen placed at a distance of 1.5 cm away from the cuvette. Finally, the scattered pattern was captured using a complementary metal-oxide-semiconductor (CMOS) camera device placed at a distance of 5 cm away from the screen. Furthermore, the camera was connected to a Raspberry Pi 3 Model B+ device with a 64-bit quad-core processor running at 1.4 GHz with 1 GB RAM. The advantage of using it in this simulation was that it has a powerful processor that is faster than other devices, which can handle multiple tasks simultaneously and control the temperature of the connected devices so that they can run longer without heating. It is also compatible and ideal for connecting interfaces and components such as HDMI, USB, and Ethernets. It is widely used in commercial applications such as biometrics systems [33], cloud based applications [34], etc. Finally, the images were transferred to a laptop with a hardware configuration consisting of an i5-7200U processor running at 3.2 GHz speed and having 8 GB RAM. In addition to this, the device also contained an external NVIDIA GPU 960MX for better performance. The training dataset consisted of 1000 images in total with a distribution of 100 images in ten individual classes.

**Table 1.** Distribution of salt concentrations in the respective classes.

| Classes | Range (in ppt) |
|---|---|
| 1 | 0–10 |
| 2 | 10–20 |
| 3 | 20–30 |
| 4 | 30–40 |
| 5 | 40–50 |
| 6 | 50–60 |
| 7 | 60–70 |
| 8 | 70–80 |
| 9 | 80–90 |
| 10 | 90–100 |

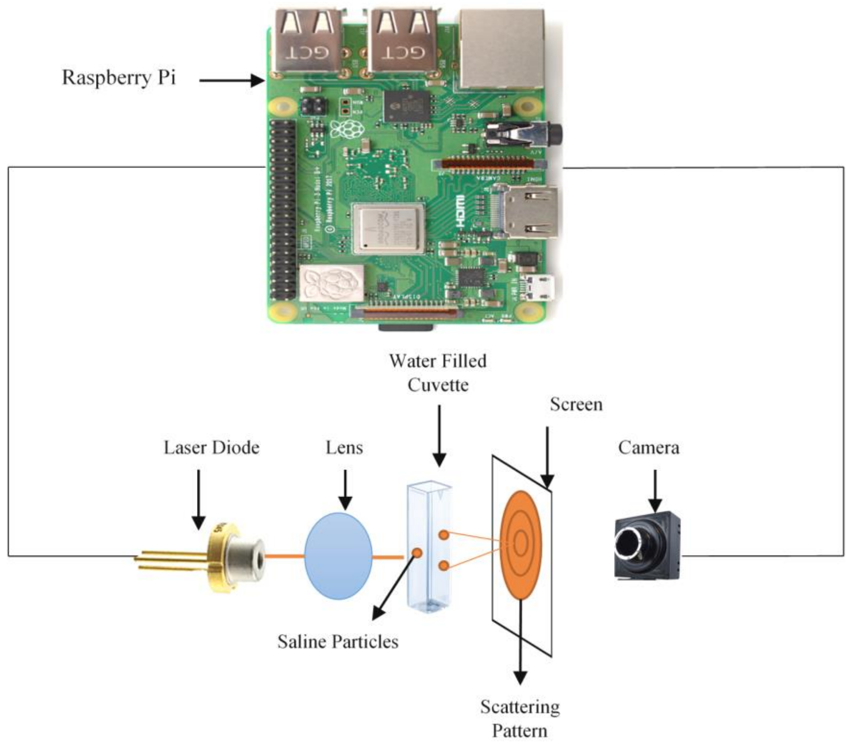


**Figure 1.** Schematic representation for determining the scattering image pattern.

### 3.3. *Stage 2: Classification of Concentration of Saline Particles*

The classification of a scattered pattern is done using a deep learning technique. In this research work, VGG16 [35], a popular CNN model, was used to classify the images. The sole reason for using it to implement the proposed system is that it has a very deep network architecture consisting of up to 19 weighted layers that facilitates the capture of robust and unique features, making it ideal for classifying images with great accuracy. Furthermore, transfer learning was used in conjunction with the CNN model since it did not contain enough training images. In addition to this, fine-tuning was integrated on the top of the model to predict ten classes. The implementation of the model was done using Keras and Tensor flow frameworks. Figure 2 shows the architecture of the fine-tuned VGG16 model with weights and knowledge transferred from the ImageNet database [36].

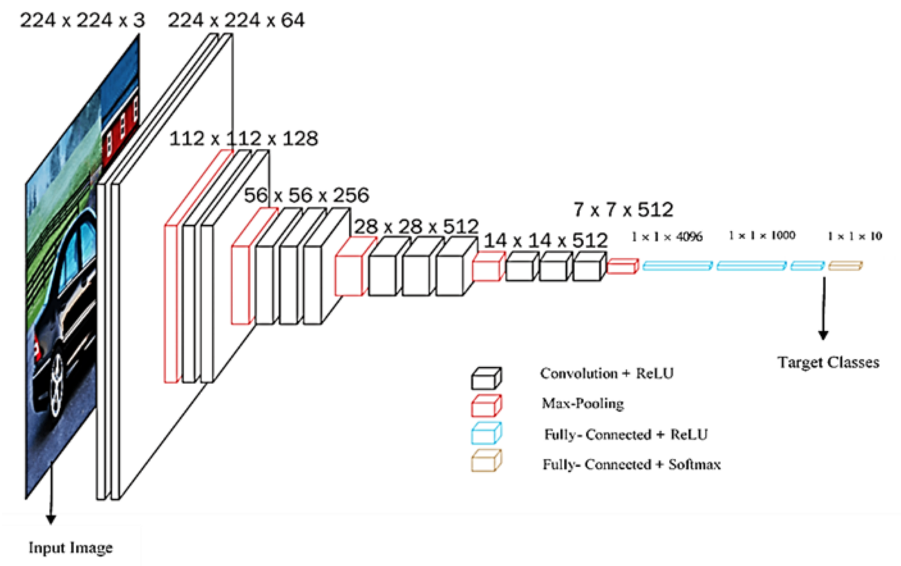


**Figure 2.** The architecture of the fine-tuned VGG-16 model.

Initially, the training images were converted to grayscale images to simplify the computation task. Then, the input images were reshaped to the required dimensions of 224 × 224 × 3 for the input of the model. The layers of the model were frozen so that the model only trained the added layers. The data augmentation technique was applied to the training set due to its limited size. Table 2 formulates the various transformations applied to the training images. Furthermore, the images were split into training and testing data in the ratio 70:30, respectively. Customized parameters tabulated in Table 3 were used to perform network training and intricate pattern and design learning. According to Table 3, the Adam optimizer, represented using Equations (1) and (2), is used for performing optimization of the network since it outputs robust results.

$$W_i = W_{i-1} - \alpha \frac{V_{dw}}{\sqrt{S_{dw} + \in}} \quad (1)$$

$$b_i = b_{i-1} - \alpha \frac{V_{db}}{\sqrt{S_{dw} + \in}} \quad (2)$$

**Table 2.** Data augmentation techniques applied to the model.

| Type of Transformation | Value |
|---|---|
| Rotation | 20 |
| Zooming | 0.2 |
| Width Shifting | 0.2 |
| Height Shifting | 0.2 |
| Rescaling | 1.0/255 |
| Sample wise Centering | True |
| Horizontal Flipping | True |

**Table 3.** List of parameters used for ADAM optimization.

| Parameter Name | Default Value |
|---|---|
| $\alpha$ | 0.001 |
| $\beta_1$ | 0.9 |
| $\beta_2$ | 0.999 |
| $\in$ | $10^{-8}$ |

In Equations (1) and (2), $W_i$ denotes the weights at $i^{th}$ iteration; $\alpha$ is the tuned step size sequence known as learning rate; and $b_i$ denotes the bias at $i^{th}$ iteration. In the equation, $V_{dW}$ and $S_{dW}$ are computed using Equations (3) and (4).

$$V_{dW} = \frac{\beta_1 V_{dW-1} + (1 - \beta_1)dW}{(1 - \beta_1{}^t)} \tag{3}$$

$$S_{dW} = \frac{\beta_2 S_{dW-1} + (1 - \beta_2)dW^2}{(1 - \beta_2{}^t)} \tag{4}$$

where $\beta_1$ and $\beta_2$ are the hyperparameters used to control the first-order momentum and second-order momentum, respectively. Similarly, analogous to Equations (3) and (4), $V_{db}$ and $S_{db}$ can be computed. The default setting to test the machine learning problems for various parameters to perform the optimization are tabulated in Table 3 under the list of parameters used in designing the model.

Since the scattered images were complicated, to learn the robust patterns, we therefore used a low learning rate of $1 \times 10^{-5}$ for model training as per Table 4. Moreover, the particular value for this hyper-parameter was determined by experimenting and evaluating various ANN models. As the implemented deep learning model was trained with lower learning rate, the models were trained on high epochs. Categorical cross entropy loss function was used for the modeling and multi-classification task. Moreover, since image data were used in the proposed study for implementing CNN that contained lots of features and patterns, we trained the model on a batch of 32 images for easy computation.

**Table 4.** List of parameters used in designing the model.

| Model Parameters | Values |
|---|---|
| Optimization Algorithm | Adam |
| Learning Rate | $1 \times 10^{-5}$ |
| Epochs | 600 |
| Loss Function | Categorical Cross Entropy |
| Batch Size | 32 |
| Final Activation Function | Softmax |
| Steps Per Epoch | Batch Size |
| Validation Steps | Batch Size/2 |

## 4. Experimental Results and Discussion

This section discusses the classification metrics used to evaluate the robustness of the deep CNN model. This section's primary objective was to visualize the proposed study's effectiveness on different saline concentration images. Furthermore, different measuring metrics were used to evaluate the observations, providing evidence of the system's high accuracy. The experiment was conducted on a machine with technical specifications as follows: Intel-i5 7200U, 8 GB Ram, 1 TB hard disk, and NVIDIA GTX 760MX Graphics. Additionally, the implementation of the proposed CNN model was simulated in Python version: 3.5 environments. To establish the results of the proposed method, we performed a statistical analysis by assessing predictive performance. This was evaluated by estimating the accuracy and Kappa [37] (represented mathematically using Equation (5)) on our training dataset with respect to the proposed model, Bayesian generalized linear model [37] and rotation forest [37] models. The statistical analysis was done and parameters were obtained as shown in Table 5. A 5-fold cross validation was applied to optimize the best model test with the best variable subset = 2.

$$k = \frac{Accuracy - P_e}{1 - P_e} \tag{5}$$

where $k$ is kappa whose value ranges between $(0, 1)$ and $P_e$ represents the expected agreement between the experimental and observed value.

**Table 5.** Predictive analysis on the training dataset.

| Statistics | Bayesian Generalized Linear Model [37] | Rotation Forest [38] | Proposed Model |
|---|---|---|---|
| Accuracy | 87.6% | 87.6% | 90% |
| Kappa | 0.81 | 0.81 | 0.83 |

Compared to these models, the accuracy for saline water particle detection was greater for the proposed ML model while Kappa was same for the other techniques when tested with our dataset. Other performance measuring metrics used for evaluation are tabulated in Figures 4 and 5, which are discussed as follows:

To effectively visualize and understand the implemented classification model, it is necessary to understand the below parameters:

i. True Positives ($T_p$)—This evaluates to true if the actual class labels match with the predicted class labels.
ii. True Negatives ($T_n$)—This evaluates to false if the actual class labels match with the predicted class labels.
iii. False Positives ($F_p$) and False Negatives ($F_n$)—These parameters signify that the actual class labels do not comply with the class predicted by the model or vice-versa.

Confusion Matrix

It is a powerful classification measurement metric that helps to visualize and judge the performance of the classifier. In the confusion matrix represented in Figure 3, the x-axis denotes labels predicted by the classifier while the y-axis denotes actual or true labels. Furthermore, the diagonal values in Figure 4 represent the samples correctly classified by the classifier while the non-diagonal samples denote misclassified samples.

Precision

This estimates the proportion of total positive predictions predicted by our classification model that are actually positive. This is represented using Equation (6).

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{6}$$

Recall

This estimates the proportion of correctly identified positive predictions to all the relevant samples identified as positive by the network model. Equation (7) calculates the recall for the proposed model.

$$\text{Recall} = \frac{T_p}{T_p + F_n} \tag{7}$$

F-Score

This calculates the combined harmonic mean of precision and recall by considering both measuring metrics. Equation (8) estimates the f-score for the proposed classification model.

$$\mathrm{F-score} = 2 \times \left( \frac{Precision \times Recall}{Precision + \_Recall} \right) \tag{8}$$

Accuracy

This estimates the proportion of true positive and negative class labels to the total number of class instances $(T_s)$ in the training data. Equation (9) estimates the f-score for the proposed classification model.

$$\mathrm{Accuracy} = \frac{T_p + T_n}{T_s} \tag{9}$$

Specificity

This estimates the proportion of correctly identified negative predictions to all the relevant samples identified as negative by the network model. Equation (10) calculates the recall for the proposed model.

$$\mathrm{Specificity} = \frac{T_n}{T_n + F_p} \tag{10}$$

All these parameters help to analyze the classification of salinity from the captured images. The proposed work is a kind of novel work where water salinity from the images was classified using the machine learning based model. From the results, it is evident that the method is capable of producing 90% accuracy in terms of water salt classification, while other parameters are also impressive at the initial stage results. Furthermore, the work is unique and a comparison is not feasible with other methods as this type of work has not been used as of today in this domain. To establish the proposed method for water salinity classification, the results are convincible.

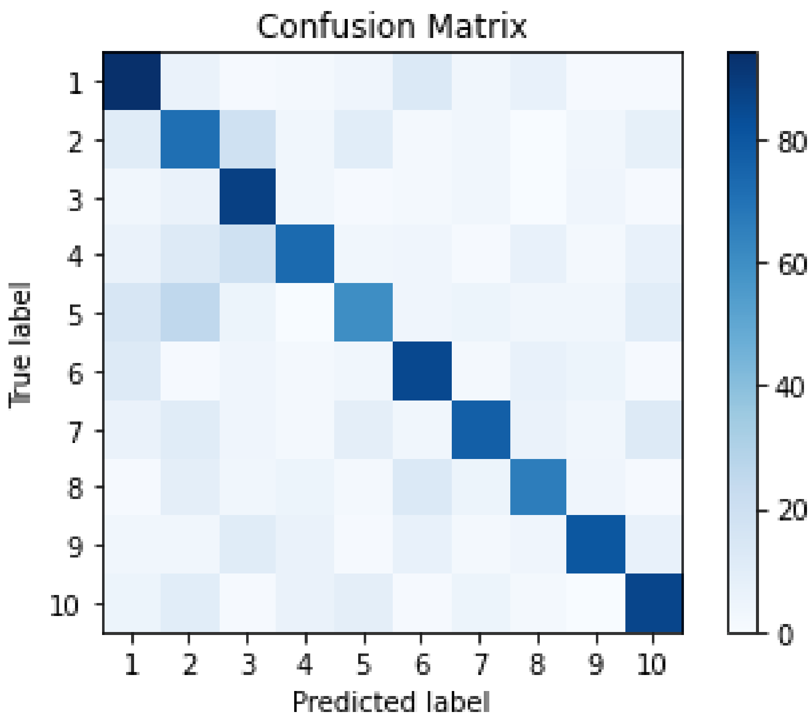


**Figure 3.** Graphical representation of the confusion matrix.

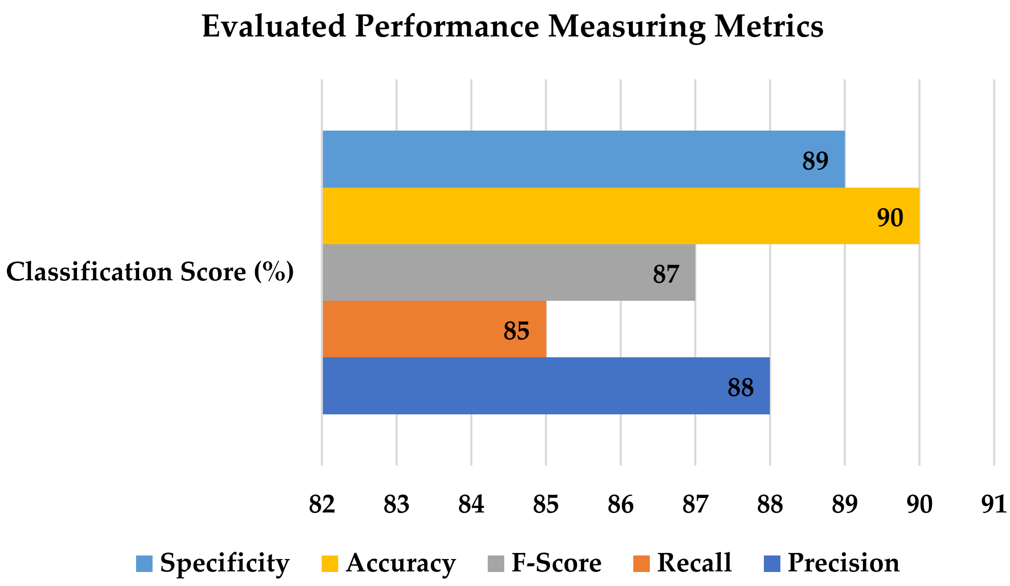


**Figure 4.** Classification results of the proposed approach.

*Discussion*

It can be observed from Figure 3 that our classification model classified most of the samples correctly, while misclassifying only aa small percentage of samples incorrectly. Furthermore, from Figure 4, it was observed that the overall accuracy of the model may increase if there are more training data present for training the model. The specificity metrics of the model showed that it classified most of the negative predictions as correct by exhibiting a true negative ratio of 0.89. On the other hand, the model overall classified 85% correct positive predictions on the test data. Moreover, from Figure 4, it can be seen that the model classified more positive predictions over false misclassification by demonstrating a precision score of about 88%, thereby making the model robust and efficient. Finally, an f-score of 87% showed that the proposed model maintained between precision and recall. To demonstrate state-of-the-art results and provide a detailed analysis report for evaluating the proposed approach, the individual classification scores of target classes are presented in Figure 5.

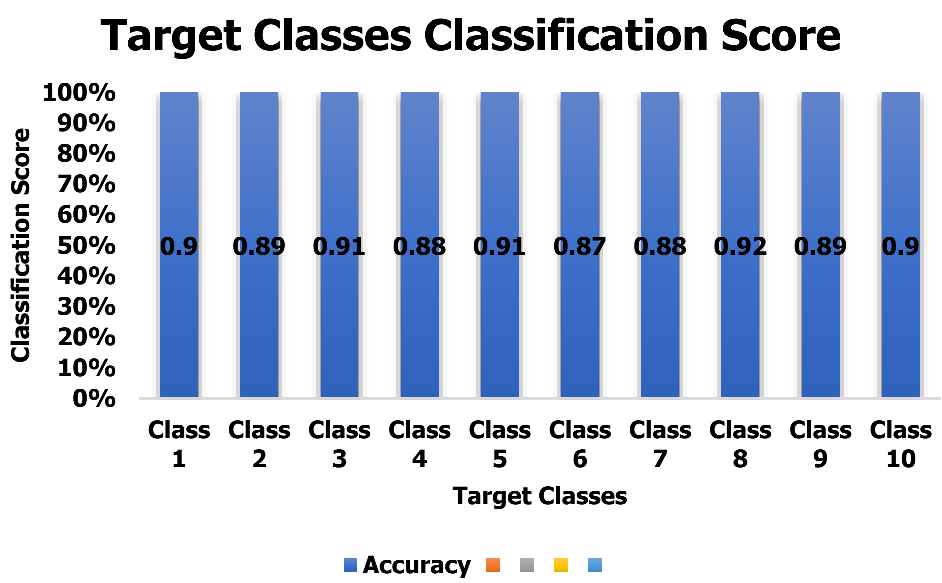


**Figure 5.** Classification score for the individual target classes.

These metric scores demonstrate that our classification deep learning model make accurate classifications and predictions by classifying salt particles. To prove that our classification model is robust and achieves state-of-the art results, the implemented model was compared with similar salinity based prediction systems. Since very little research has been done in the identification of seawater salinity concentration, it is not possible to compare it with other approaches implemented in this area of research, but we have tried to compare it with the best machine learning techniques implemented to estimate salinity. Table 6 tabulates the comparison of the proposed approach with other salinity-based estimation systems. It is evident from Table 6 that our proposed approach outperformed the popularly known salinity based classification approaches by achieving an overall accuracy of 90%. Moreover, the approach proposed by [39] also gave good results by reaching an accuracy of about 86.7%, while the approaches presented by [40,41] produced fair results in comparison to the remaining discussed approaches. The proposed method outperformed compared to these approaches. On comparing our proposed approach with the techniques presented in [23,42], which aims to predict the salinity and estimates TDS in seawater using various machine learning and deep learning techniques, it was observed that our implemented technique was proved to be powerful and accurate in terms of accuracy. Furthermore, the performance and robustness of the proposed technique was also tested with [43], who targeted the salinity of seawater using mixtures of machine learning models on a real world dataset.

**Table 6.** Comparison of various salinity based classification approaches.

| Salinity Prediction Approaches | Accuracy (%) |
|---|---|
| Classifier Ensemble + Bayesian Learning (Mosavi, Hosseini, Choubin, Goodarzi, & Dineva, 2020) [39] | 86.7 |
| Decision Tree (Vermeulen & Niekerk, 2017) [40] | 75 |
| Random Forest (Ivushkin et al., 2019) [41] | 70 |
| **Proposed Approach** | **90** |

This classification implemented model can further be utilized to design the best water treatment plants and devices so that they can treat seawater with high accuracy and exhibit reliable performance. Since nutrients such as $Na^{+}$, $K^{+}$, $Ca^{2+}$, $Mg^{2+}$, and ($SO_4$) present in seawater are very hard to detect and remove, seawater is unfit for consumption. The integration of the proposed model with modern treatment plants could improve their performance considerably, so that the harmful nutrients can be eliminated from seawater, making it suitable for consumption in various sectors such as health care, agriculture, domestic use, etc. In addition to harmful elements, TDS and saline salt hardness can be eradicated, which also acts as one of the major contributors of salinity in seawater.

## 5. Conclusions and Future Works

This paper detected and identified saline particles in different concentrations of the saline solution. The scattering images of particles were collected using the experimental setup shown in Figure 1. Furthermore, we applied the CNN model with transfer learning to classify the saline particles. In addition, the data augmentation technique was also used to generalize well and output robust results. Using the proposed approach, the accuracy of the treatment plants will increase to a great extent, leading to better results for water treatment. The experimental results demonstrated that the deep learning system outputted a higher accuracy of 90%, while the f-score metrics outputted 87%, which shows that the system considers all the classes in a balanced ratio. Furthermore, the remaining classification metrics such as precision, recall, and specificity also exhibited robust and state-of-the-art results. However, from Figure 4, it appears that due to the limited amount of training data, our deep learning system suffered from the limitation of misclassification of saline salt particles. Usage of techniques such as transfer learning and data augmentation increases the training time and thus, in turn, increases the complexity of the system. Moreover, the system could not classify the salinity of salts greater than 100 ppt due to the limited range of class labels in the training data.

In future works, we plan to integrate our proposed approach with modern water treatment plants to study the practicality of deep learning systems in the water treatment process. It would also be interesting to visualize and evaluate the proposed approach with a large training set. We also plan to implement our proposed study with hybrid deep learning models to examine the problem of water treatment more vividly and design highly accurate water treatment systems in the future.

**Author Contributions:** Conceptualization, Methodology, Supervision and Investigation, M.A.; Conceptualization, Methodology, Data curation, Writing—review & editing and Supervision, M.K.; Methodology, Software, Data curation, Validation, and Writing, A.B.; Formal analysis, Investigation, Visualization, and Writing reviews, S.M.; Resources, Writing—review & editing, Software, Daft study, and Data analysis, J.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors extend their appreciation to the deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number (IFP-2020-21).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Patil, V.C.; Al-Gaadi, K.A.; Madugundu, R.; Tola, E.H.M.; Marey, S.; Aldosari, A.; Biradar, C.M.; Gowda, P.H. Assessing Agricultural Water Productivity in Desert Farming System of Saudi Arabia. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *8*, 284–297. [CrossRef]
2. Pivokonsky, M.; Cermakova, L.; Novotna, K.; Peer, P.; Cajthaml, T.; Janda, V. Occurrence of microplastics in raw and treated drinking water. *Sci. Total Environ.* **2018**, *643*, 1644–1651. [CrossRef]
3. Bhojwani, S.; Topolski, K.; Mukherjee, R.; Sengupta, D.; El-Halwagi, M.M. Technology review and data analysis for cost assessment of water treatment systems. *Sci. Total Environ.* **2019**, *651*, 2749–2761. [CrossRef] [PubMed]
4. Ghernaout, D. Increasing Trends Towards Drinking Water Reclamation from Treated Wastewater. *World J. Appl. Chem.* **2017**, *3*, 1. [CrossRef]
5. Liponi, A.; Wieland, C.; Baccioli, A. Multi-effect distillation plants for small-scale seawater desalination: Thermodynamic and economic improvement. *Energy Convers. Manag.* **2020**, *205*, 112337. [CrossRef]
6. Liu, S.; Wang, Z.; Han, M.; Wang, G.; Hayat, T.; Chen, G. Energy-water nexus in seawater desalination project: A typical water production system in China. *J. Clean. Prod.* **2021**, *279*, 123412. [CrossRef]
7. Khan, M.S.; Lal, B.; Sabil, K.M.; Ahmed, I. Desalination of Seawater through Gas Hydrate Process: An Overview. *J. Adv. Res. Fluid Mech. Therm. Sci.* **2019**, *55*, 65–73.
8. Pistocchi, A.; Bleninger, T.; Breyer, C.; Caldera, U.; Dorati, C.; Ganora, D.; Millán, M.; Paton, C.; Poullis, D.; Herrero, F.S.; et al. Can seawater desalination be a win-win fix to our water cycle? *Water Res.* **2020**, *182*, 115906. [CrossRef] [PubMed]
9. Melesse, A.M.; Khosravi, K.; Tiefenbacher, J.P.; Heddam, S.; Kim, S.; Mosavi, A.; Pham, B.T. River Water Salinity Prediction Using Hybrid Machine Learning Models. *Water* **2020**, *12*, 2951. [CrossRef]
10. Banerjee, P.; Singh, V.; Chatttopadhyay, K.; Chandra, P.; Singh, B. Artificial neural network model as a potential alternative for groundwater salinity forecasting. *J. Hydrol.* **2011**, *398*, 212–220. [CrossRef]
11. Barzegar, R.; Moghaddam, A.A. Combining the advantages of neural networks using the concept of committee machine in the groundwater salinity prediction. *Model. Earth Syst. Environ.* **2016**, *2*, 26. [CrossRef]
12. Azad, A.; Karami, H.; Farzin, S.; Mousavi, S.-F.; Kisi, O. Modeling river water quality parameters using modified adaptive neuro fuzzy inference system. *Water Sci. Eng.* **2019**, *12*, 45–54. [CrossRef]
13. Duong, H.C.; Cooper, P.; Nelemans, B.; Cath, T.Y.; Nghiem, L.D. Evaluating energy consumption of air gap membrane distillation for seawater desalination at pilot scale level. *Sep. Purif. Technol.* **2016**, *166*, 55–62. [CrossRef]
14. Doornbusch, G.; van der Wal, M.; Tedesco, M.; Post, J.; Nijmeijer, K.; Borneman, Z. Multistage electrodialysis for desalination of natural seawater. *Desalination* **2021**, *505*, 114973. [CrossRef]
15. Hussain, I.; Das, M.; Ahamad, K.U.; Nath, P. Water salinity detection using a smartphone. *Sens. Actuators B Chem.* **2017**, *239*, 1042–1050. [CrossRef]
16. Tufa, R.A.; Noviello, Y.; Di Profio, G.; Macedonio, F.; Ali, A.; Drioli, E.; Fontananova, E.; Bouzek, K.; Curcio, E. Integrated membrane distillation-reverse electrodialysis system for energy-efficient seawater desalination. *Appl. Energy* **2019**, *253*, 113551. [CrossRef]

17. Maia, C.B.; Silva, F.V.; Oliveira, V.L.; Kazmerski, L.L. An overview of the use of solar chimneys for desalination. *Sol. Energy* **2019**, *183*, 83–95. [CrossRef]
18. Suwaileh, W.; Johnson, D.; Hilal, N. Membrane desalination and water re-use for agriculture: State of the art and future outlook. *Desalination* **2020**, *491*, 114559. [CrossRef]
19. Alkhudhiri, A.; Bin Darwish, N.; Hilal, N. Analytical and forecasting study for wastewater treatment and water resources in Saudi Arabia. *J. Water Process. Eng.* **2019**, *32*, 100915. [CrossRef]
20. Mansouri, N.Y.; Ghoniem, A.F. Does nuclear desalination make sense for Saudi Arabia? *Desalination* **2017**, *406*, 37–43. [CrossRef]
21. Pan, S.-Y.; Haddad, A.Z.; Kumar, A.; Wang, S.-W. Brackish water desalination using reverse osmosis and capacitive deionization at the water-energy nexus. *Water Res.* **2020**, *183*, 116064. [CrossRef]
22. De Medeiros, J.F.; Lisboa, R.D.A.; De Oliveira, M.; da Silva Júnior, M.J.; Alves, L.P. Caracterização das águas subterrâneas usadas para irrigação na área produtora de melão da Chapada do Apodi. *Rev. Bras. Eng. Agríc. Ambient.* **2003**, *7*, 469–472. [CrossRef]
23. Rose, L.; Mary, X.A. Sensor data classification using machine learning algorithm. *J. Stat. Manag. Syst.* **2020**, *23*, 363–371. [CrossRef]
24. Taheran, M.; Brar, S.K.; Verma, M.; Surampalli, R.; Zhang, T.; Valero, J. Membrane processes for removal of pharmaceutically active compounds (PhACs) from water and wastewaters. *Sci. Total Environ.* **2016**, *547*, 60–77. [CrossRef] [PubMed]
25. 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]
26. Aggarwal, A.; Kumar, M. Image surface texture analysis and classification using deep learning. *Multimed. Tools Appl.* **2021**, *80*, 1289–1309. [CrossRef]
27. Aggarwal, A.; AlShehri, M.; Kumar, M.; Alfarraj, O.; Sharma, P.; Pardasani, K.R. Landslide data analysis using various time-series forecasting models. *Comput. Electr. Eng.* **2020**, *88*, 106858. [CrossRef]
28. Aggarwal, A.; Alshehri, M.; Kumar, M.; Sharma, P.; Alfarraj, O.; Deep, V. Principal component analysis, hidden Markov model, and artificial neural network inspired techniques to recognize faces. *Concurr. Comput. Pract. Exp.* **2021**, *33*, e6157. [CrossRef]
29. Aggarwal, A.; Rani, A.; Kumar, M. A robust method to authenticate car license plates using segmentation and ROI based approach. *Smart Sustain. Built Environ.* **2019**, *9*, 737–747. [CrossRef]
30. Chen, Y.; Fan, R.; Yang, X.; Wang, J.; Latif, A. Extraction of Urban Water Bodies from High-Resolution Remote-Sensing Imagery Using Deep Learning. *Water* **2018**, *10*, 585. [CrossRef]
31. Hu, C.; Wu, Q.; Li, H.; Jian, S.; Li, N.; Lou, Z. Deep Learning with a Long Short-Term Memory Networks Approach for Rainfall-Runoff Simulation. *Water* **2018**, *10*, 1543. [CrossRef]
32. Chatterjee, A.; Hu, X.; Lam, F.L.-Y. Modified coal fly ash waste as an efficient heterogeneous catalyst for dehydration of xylose to furfural in biphasic medium. *Fuel* **2019**, *239*, 726–736. [CrossRef]
33. Shah, D.; Haradi, V. IoT Based Biometrics Implementation on Raspberry Pi. *Procedia Comput. Sci.* **2016**, *79*, 328–336. [CrossRef]
34. Hajji, W.; Tso, F.P. Understanding the Performance of Low Power Raspberry Pi Cloud for Big Data. *Electronics* **2016**, *5*, 29. [CrossRef]
35. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2015**, arXiv:1409.1556. Available online: http://arxiv.org/abs/1409.1556 (accessed on 18 March 2021).
36. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [CrossRef]
37. Hosseini, F.S.; Choubin, B.; Mosavi, A.; Nabipour, N.; Shamshirband, S.; Darabi, H.; Haghighi, A.T. Flash-flood hazard assessment using ensembles and Bayesian-based machine learning models: Application of the simulated annealing feature selection method. *Sci. Total Environ.* **2019**, *711*, 135161. [CrossRef] [PubMed]
38. Gelman, A.; Pardoe, I. Bayesian Measures of Explained Variance and Pooling in Multilevel (Hierarchical) Models. University Library of Munich, Germany. April 2004. Available online: https://ideas.repec.org/p/wpa/wuwpem/0404002.html (accessed on 24 March 2021).
39. Mosavi, A.; Hosseini, F.S.; Choubin, B.; Goodarzi, M.; Dineva, A.A. Groundwater Salinity Susceptibility Mapping Using Classifier Ensemble and Bayesian Machine Learning Models. *IEEE Access* **2020**, *8*, 145564–145576. [CrossRef]
40. Vermeulen, D.; Van Niekerk, A. Machine learning performance for predicting soil salinity using different combinations of geomorphometric covariates. *Geoderma* **2017**, *299*, 1–12. [CrossRef]
41. Ivushkin, K.; Bartholomeus, H.; Bregt, A.K.; Pulatov, A.; Kempen, B.; de Sousa, L. Global mapping of soil salinity change. *Remote Sens. Environ.* **2019**, *231*, 111260. [CrossRef]
42. Alshehri, M.; Bhardwaj, A.; Kumar, M.; Mishra, S.; Gyani, J. Cloud and IoT based smart architecture for desalination water treatment. *Environ. Res.* **2021**, *195*, 110812. [CrossRef] [PubMed]
43. Muharemi, F.; Logofătu, D.; Leon, F. Machine learning approaches for anomaly detection of water quality on a real-world data set. *J. Inf. Telecommun.* **2019**, *3*, 294–307. [CrossRef]