Spatial Information Research
https://doi.org/10.1007/s41324-023-00537-x

Deep convolutional neural network to predict ground water level

Abu Sarwar Zamani[1,2] · Aisha Hassan Abdalla Hashim[2] · Arepalli Gopi[3] · Kavita Moholkar[4] · Mohammed Rizwanullah[1] · Rasool Altaee[5]

Received: 5 April 2023 / Revised: 22 June 2023 / Accepted: 23 June 2023

Abstract
In contrast to the atmosphere and fresh surface water, which can only briefly store water, the natural water cycle may use groundwater as a "reservoir" that stores water for extended periods. Even though there is a considerable degree of variation and complexity in the subsurface environment, there is a minimal availability of data from the field. Both of these challenges were faced by those who used models that were based on actual reality. Statistical modelling gradually improved the accuracy of the model's calibration. Groundwater has become an increasingly important resource for supplying the water requirements of a rising global population. The fact that there is such a large stockpile allows it to be used once again, even during dry seasons or droughts. This article presents a deep convolutional neural network-based model for predicting groundwater levels. As part of the experimental setup, 174 satellite pictures of groundwater are included in the input data set. Images are preprocessed using the CLAHE method. The CNN, SVM, and AdaBoost methods make up the classification model. The results have shown that CNN can classify things correctly 98.5 per cent of the time. Precision and Recall rate of Deep CNN is also better for ground water image classification.

Keywords Environment Monitoring System · Ground water level prediction · Deep convolutional neural network · Satellite images

1 Introduction

Water is an essential component for the continued existence of all forms of life on our planet. It is essential for the upkeep of ecosystems on a local and global scale, as well as for the delivery of medical care, the manufacture of food, and the generation of power. Water is a resource that simply must be present for any form of expansion or advancement that takes place on a national level. It's possible that water is the only natural resource that has an effect on every facet of human society, but it's only a theory. Water could be the only natural resource that does that. The availability of drinkable water and sufficient sanitary facilities would be the single most important element in lowering the sickness rate in less developed countries. As a result, it is of the utmost importance to track down reliable water sources,

✉ Abu Sarwar Zamani
a.zamani@psau.edu.sa

Aisha Hassan Abdalla Hashim
aisha@iium.edu.my

Arepalli Gopi
gopi.arepalli400@gmail.com

Kavita Moholkar
kavita.moholkar@gmail.com

Mohammed Rizwanullah
r.mohammed@psau.edu.sa

Rasool Altaee
rasool.azeem@uomus.edu.iq

[1] Department of Computer and Self Development, College of Preparatory Year, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia

[2] Department of Electrical and Computer Engineering, International Islamic University Malaysia, Kuala Lumpur 53100, Malaysia

[3] Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur (D.T), Andhra Pradesh, India

[4] Department of computer science and business systems, JSPM's Rajarshi Shahu College of Engineering, Pune, India

[5] Medical Laboratories Techniques Department, Al-Mustaqbal University College, Hillah, Iraq

Published online: 25 August 2023

make certain that they are always properly maintained, and make effective use of the water that is available. Even if the amount of potable water that is currently available is woefully insufficient, there are still some areas of the world that do not have access to a reliable source of the resource [1].

Surface water [2] and groundwater are the two basic types of water supplies that are available. The term "surface water" refers to any body of water that exists on the surface of the land or that is held above ground. This is the kind of water that may be found in bodies of water such as lakes, streams, reservoirs, and the ocean itself. The term "groundwater" refers to any water that is located under the surface of the earth. Wherever there is a space that is not occupied, such as the crevices between the rocks or the pores in the mud. In contrast to the atmosphere and fresh surface water, which can only temporarily store water, the natural water cycle might use groundwater as a "reservoir" that stores water for longer periods of time. This would be in contrast to the atmosphere and fresh surface water, which can only store water temporarily. When precipitation, streams, and rivers are able to reach that level, it is possible to perform the task of recharging the aquifer that is below the water table. For the purpose of satisfying the thirst of an ever-expanding worldwide population, groundwater has emerged as a resource of increasing significance, and this is true in both economically developed and developing countries. The fact that there is such a large stockpile allows it to be used once again, even during dry seasons or droughts [3]. Due to the lower levels of pollution, the expenses of development are cheaper, and there are less requirements for maintenance. The vast majority of the time, you can just access it whenever you have a need for it. It is less susceptible to the effects of natural calamities.

According to the opinions of several experts, groundwater is the source of more than 95% of the world's potable fresh water. Groundwater is either the major or secondary source of drinking water for almost half of the population in the United States. The majority of rural homes, around 95%, rely on water that comes from the ground as their main supply of water. The subsurface aquifer supplies water to about 50% of all irrigated agricultural land. The use of groundwater helps satisfy about one-third of the industrial industry's water demand. Groundwater is used in many aspects of agricultural production in arid locations, where precipitation is in short supply and difficult to anticipate [4]. So, all these reasons proves that it is necessity of the world to locate dependable water supplies.

For the purpose of this study, the discipline of hydrology makes use of soft computing. Deep learning is used in many areas related to environment monitoring. Humidity and moisture sensor, soil sensor, gas sensor, rainfall sensor, speed sensor, light sensor, wind speed sensor are some common sensors used in the environment monitoring systems [5].

To put it another way, hydrology is the study of rain as a phenomenon that occurs all over the world [6]. The constant movement of water on, above, and below the surface of the earth is referred to as the hydrological cycle. We can consider precipitation, evaporation, infiltration, transpiration, ground water flow, and surface runoff to be different aspects of the same system when speaking in terms of the hydrological cycle. It is difficult to characterize the processes that have been discussed above since they are very nonlinear across several dimensions (spatial and temporal). Both data-driven models and physical models may be used to investigate and solve hydrological problems. The underlying physical and mathematical assumptions of models that are based on physical numerical flow have a tendency to be simplified. This is done so that the models can produce outcomes that are both practical and attainable. In spite of the fact that the subsurface environment exhibits a significant degree of diversity and complexity, there is an incredibly low availability of data collected from the field. Those individuals who utilized models that were founded on actual reality were the ones who were confronted with both of these issues. Through the application of statistical modeling, the precision of the model's calibration was able to see gradual improvements. The inadequacy of physically based numerical flow models and statistically based modeling methodologies to deal with inconsistency and uncertainty was one of the driving forces for the development of soft computing [7].

Within the field of soft computing [8], one may be able to discover unconventional optimization strategies. There is no cutthroat competition amongst the basic components that make up FL, ANN, GA, and PSO (Particle Swarm Optimization). A convoluted mathematical representation of the problem is not required when one uses the computing approaches known as soft computing. Experimentation is performed to acquire data, which is then utilized for educational purposes. Approximation and interpolation are two techniques that enable these systems to generalize their findings and provide outcomes based on inputs that have never been tested before (by using outputs from previous learned inputs). This indicates that it provides approximative solutions at an affordable price. When it comes to FL, which deals with imprecision and uncertainty in datasets, soft computing is a collection of ways, and each methodology has its own strengths. This is because soft computing is a collection of methodologies. The use of ANNs as a learning and adaptation tool shows promise; nonetheless, GA and PSO stand out as essential search and optimization methodologies.

Section 2 contains literature review of different techniques used for image segmentation, feature selection and classification. Section 3 presents proposed methodology for ground water level classification and detection. Section 4 presents result analysis and discussion. Section 5 contains conclusion and future work.

2 Literature survey

The fuzzy rule-based modeling/qualitative modeling technique, which was invented by Sugeno et al. [9], makes it possible for the behavior of the system to be modeled in a way that is more accurate to how it actually behaves. This is because the technique combines fuzzy logic with traditional modeling techniques. This example demonstrates how fuzzy modeling can be used as a technique for the construction of a system model by making use of a description language that contains fuzzy logic and fuzzy predicates. Fuzzy modelling can be utilized as a technique for the creation of a system model by using a description language. This illustration provides a presentation of the methodology. When combined with a fuzzy model, the language approximation approach has the potential to produce a qualitative model. In order to figure out the internal structure of a fuzzy model, the fuzzy clustering approach was put to use. One step in the right direction would be to advance qualitative modeling by enhancing the practice of language approximation. This would be one step forward in the process. This indicates that scientists have made significant progress toward replacing numerical computing with that of words, which can be understood in a more general way as follows: For the purpose of gaining an understanding of how a system operates on a physical level, FL makes use of if-then principles that are relatively well articulated.

One of the two approximate models that were provided by Bardossy [10] is one that is based on accurate flow principles and makes use of information about the moisture content of surrounding layers as well as the depth of the saturating front. The other model, which adopted a more traditional approach, detailed the procedure for carrying out the penetration. The mathematical solution to the Richards equation provided the source of the data that was used to train the second fuzzy rule-based model. The amount of moisture presents in the layers that are adjacent to one another served as the foundation for the establishment of the standards. Traditional models completed their work more slowly and required more input variables, while both of these techniques concluded their work more rapidly.

Panigrahi [11] constructed a model of the operation of reservoirs by making use of FL. The administration of a reservoir with a singular use was modeled with the help of fuzzy logic principles. A conditional assertion serves as the foundation for the paradigm. The development of the model consists of three primary stages: the construction of fuzzy rules, the making of inferences, and the defuzzification of the results. After these tasks have been completed, it will be possible to construct membership functions for input, output, demand, and supply. We will examine the Malaprabha irrigation region, which is located in the state of Karnataka in India, as an illustration of how the methodology can be applied in practice. It was hypothesized that the reservoir's retaining capacity, the volume of water flowing into the reservoir, and the amount of water that needed to be released all contributed to the discharge. The data on which the adaptable rules are based were obtained through the simulation of a reservoir that was functioning in accordance with a steady state policy that was taken from a stochastic dynamic programming model. This is done in order to ensure that the adaptable guidelines construction requirements are met. This implementation of the random model was made as opposed to the conventional dependence on the extensive knowledge of experienced reservoir administrators.

The amount of groundwater that is refilled was calculated by Awasthi et al. [12] utilizing variables such as temperature, rainfall, and stream velocity as their data sources. With the help of water budget models, we were able to determine the volume of groundwater that was successfully recharged. It is possible that it will be challenging to precisely compute the factors that are crucial to these models. Some of these variables include insufficient soil moisture, actual evapotranspiration, direct draining, and others. The fact that the values of these variables can change depending on where you are in the world makes it that much more difficult to appropriately analyze them in any given scenario. When extraordinary events do happen, the standard deviation of the prediction increases to considerably higher levels than it would have been otherwise. One more thing that we found out was that the model for the water budget does not take into account the replenishment processes that take place during severe but brief downpours. This was one of the things that we observed. The estimate that was obtained from the water budget was around a quarter of a million liters greater than the approximation, on average; however, there was a disparity between the two of almost one million liters. It was expected that the amount of data that was fed into the model when it was being trained would have a direct bearing on the precision of the results that the model would provide.

A data-driven water level prediction model was demonstrated by Alvisi et al. [13] through the utilization of the ANN method, the Mamdani method, and the Takagi-Sugeno FL method. The committee came up with three completely different variations of the proposal. Each of the three models was evaluated using the same inputs and consequences for the evaluation. It was decided to use the Reno River in Casalecchio di Reno as a reference point in order to guarantee that the investigation was carried out with the highest level of caution and accuracy possible. It has been observed that models constructed with fuzzy logic methods might have unanticipated constraints in the amount of forecasting potential they possess. Due to the fact that Alvisi and his team were utilizing the ANN procedure, they were not in the position to recognize errors of this nature.

As part of the GLF monitoring technique, Affandi et al. [14] have published daily groundwater level fluctuations (GLF) projections. Although the LM (Levenberg Marquardt) and RBF (Radial Basis Function) were our go-to ANN models for this investigation, we also made use of the ANFIS. (Adaptive Neuro Fuzzy Inference System). (RBF). The primary objective of this research was to come up with a method that could be used regularly, was reliable, and could be used to forecast the GLF over an extended period of time. The findings indicated that the efficiency of all three approaches decreased significantly as the time increment grew longer, but apart from that, there were no other differences that were particularly noticeable.

An artificial neural network was used in the research conducted by Solaimani [15] in a semiarid region of Iran to investigate the association between rainfall and flow in a catchment area. This action was taken with forethought as its objective. By applying the feed forward back propagation technique, we demonstrated in this study how a number of different algorithms, each of which possesses multi-layer observational performance, can be used in conjunction with one another to address the problem of forecasting when it will rain. This was done to demonstrate the vast number of different applications that could potentially be used with these techniques. This investigation's objectives were to determine whether or not it is possible to use ANN to anticipate stream movements and to evaluate its performance in comparison to that of more conventional approaches. We investigated gradient descent (GDX), conjugate gradient, and LM in order to determine which training method resulted in the most substantial increases in cognitive abilities. Gradient descent (GDX) was the clear winner. Our objective was to identify the teaching method that resulted in the greatest increase in arithmetic competence. When compared side by side, it was obvious that the Artificial Neural Network method was more accurate than the conventional regression model in estimating the amount of water that waterways would transport. This was the case because the ANN method used a neural network rather than a regression model.

According to the findings of Fernandez et al. [16], those who use the Magdalena River in Colombia (South America) experience significant financial losses because of the frequent changes in navigation options brought about by variations in river levels and waterway conditions. These changes are being caused by natural calamities like hurricanes and wildfires, among other things. These shifts are an unavoidable consequence of the constantly shifting water movements and environmental conditions in the canal. Predicting how water levels will change in the future is the primary objective of this research, which employs neural fuzzy logic to evaluate the problem. The information collected from the limn graphical locations that were utilized in this investigation was supplied by the Colombian Institute of Hydrological, Meteorological, and Environmental Studies. (IDEAM).

Artificial neural networks (ANN) and fuzzy logic were utilized in the study that Mayilvaganan et al. [17] conducted in order to evaluate the groundwater level prediction within a region. (FL). A couple of examples of abbreviations are artificial neural networks (ANN) and fuzzy logic (FL). We were successful in developing a three-layer artificial neural network (ANN) that is capable of feed-forward processing with the assistance of the sigmoid function and back propagation. Input and output variables were able to be successfully incorporated into the FL model thanks to the implementation of Gaussian Fuzzy Membership Functions, which were utilized at every stage of the construction process. It's conceivable that the data accumulated can lend a hand in formulating some hazy guidelines for action. The findings demonstrated that ANNs performed significantly better than FLs in every parameter that was evaluated.

ANNs have been used successfully to reproduce a wide variety of non-linear hydrological processes [18, 19], including rainfall-runoff, stream flow, ground-water management, water quality modeling, and precipitation. As described by Maier and Dandy [20], one of the climatic occurrences that has been approximated using ANNs is the capacity to recognize complicated non-linear relationships between input and output data. This is one of the climatic occurrences that has been approximated using ANNs. The use of artificial neural networks (ANNs) to forecast rainfall is just one illustration of how this technology has been applied to the study of hydrology. The use of artificial neural networks (ANNs) for the purposes of forecasting and prediction in relation to variables that affect water availability is becoming increasingly common. Efficiency considerations had an impact on every stage of the process, including data classification and preparation, model input and network design selection, optimization of training-time link weight, and model validation.

Lohani [21] made use of FL and ANN in order to gain a better understanding of the relationship that exists between the stages. The direct measurement of the flow rate of the stream was a difficult task that required a lot of time and money. It was demonstrated through a variety of statistical and graphical approaches that there is a connection between the stages and the corresponding documented emissions. The establishment of a link between the two was the intent of this particular course of action. The presence of hysteresis in the data indicated that conventional techniques based on least squares regression analysis were unable to adequately characterize the non-linear link that existed between stage and the number of discharges that were recorded. Along the length of the Narmada River in India, multiple monitoring sites were put up, and the data from these sites was used by the authors to correlate the model to the gauge and flow that was actually observed. An investigation similar to this one was carried out, as demonstrated by the following illustration. In addition, a step-by-step process for developing the TS fuzzy model has been laid out for readers. It is abundantly obvious, in light of the findings, that the TS fuzzy modeling strategy performed significantly better than both the conventional and the ANN approaches. When it came to dealing with the hysteresis effect, the modeling findings demonstrated that the FL-based model performed significantly better than the ANN techniques by a significant margin. This was true for every single category of information that could possibly exist. (loop rating curve).

Because of its rising notoriety, the PSO is now being utilized as a research tool in an increasing number of fields, such as medicine and engineering, amongst others. Yan et al. tackle the notorious "traveling salesman problem" (TSP) in their article [22], and they reach the conclusion that the application of changing techniques was successful in solving the problem. Local minima prohibited conventional developing algorithms from tackling well-known optimization problems like the traveling salesman problem in many different cases. (TSP). Comparatively, the rate at which the PSO algorithm converges on a solution is much quicker than that of the developing algorithm.

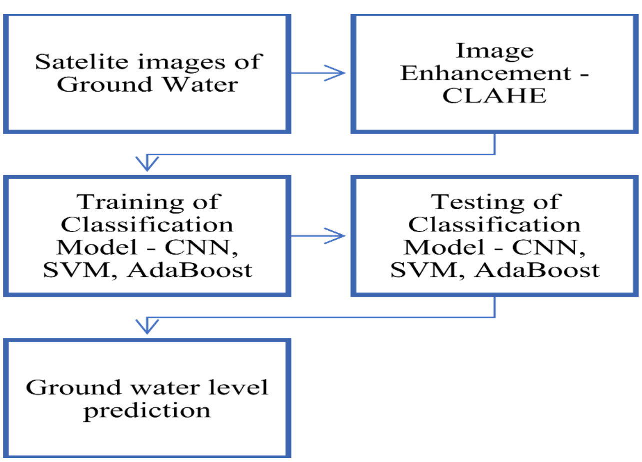

Fig. 1 Deep convolutional neural network enabled environment monitoring system for ground water level prediction

According to Chintalapati [23], SVM-PSO was able to generate more accurate projections of the amount of groundwater when compared to ARMA, ANN, and ANFIS. This conclusion can be attributed to the outcomes of the SVM-PSO research. Support vector networks were utilized in the process of developing a technique that can accurately forecast the level of groundwater. (SVMs). The suggested SVM-PSO technique was utilized in order to determine the depth of the groundwater. After sifting through a large number of potential input structures to find the one that offered the most value, the SVMPSO model used k-fold cross validation to evaluate how well it performed. The PSO algorithm underwent some changes in order to facilitate the process of locating the optimal point at which to configure the SVM's settings. The results obtained by the SVM-PSO model, the results obtained by the ARMA model, the results obtained by the ANN model, and the results obtained by an adaptive neural fuzzy inference system were compared and contrasted. The SVM-PSO model rapidly established itself as the undisputed frontrunner in the race to forecast the path that future groundwater levels will take.

PSO was initially presented by Mahnam and Ghomi [24] as a technique for the generation of time-variant fuzzy series predictions. This form of prediction makes use of a wide variety of methodologies, some of which include high-order time-invariant fuzzy time series as one example. The challenge of providing accurate forecasts has been overcome by the creation of a blended methodology. This technique uses a time-variant fuzzy time series in addition to PSO, which is an innovative strategy to evolutionary computing. The methodology that is explained in this article can be utilized to determine not only the duration of each interval but also its position in relation to the subject matter that is being discussed. Two different datasets were selected in order to demonstrate the technique and evaluate its forecasting effectiveness in comparison to that of four other fuzzy time series algorithms. The research indicates that the proposed algorithm is competitive with regard to the other techniques that fall into its category.

3 Methodology

Deep convolutional neural network enabled environment monitoring system for ground water level prediction is presented in Fig. 1 below. This environment monitoring system consists of input data set of satellite images related to ground water level. These images are enhanced by CLAHE algorithm. Then classification model is trained and tested. Classification model consists of three classification techniques- Support Vector Machine, AdaBoost and Convolutional Neural Network.

The photographs of the retina that are obtained during the tests include a variety of flaws, including poor contrast, irregular lighting, and other such issues. Images of the fundus retina are more luminous in the centre of the retina than they are on the edges or sides. The regions that make up the picture's centre are brightly lighted, however the regions that make up the sides and borders of the image are not. Because of this, poor contrast amplification of the retinal picture might allow some lesions to go undetected. This is due to the fact that the central part of the image and the OD are excessively enhanced. CLAHE, which stands for contrast limited adaptive histogram equalization, is a technique that is used to increase the contrast of the picture over the whole of the retina. CLAHE, which is an acronym that stands for contrast limited adaptive histogram equalization, is a technique for enhancing contrast that involves splitting the histogram into two halves that are equal in size. To improve the contrast of an image using the technique known as contrast limited adaptive histogram equalization (CLAHE), the area that needs to be improved is first segmented into a large number of smaller sections that are all of the same size. Then, the contrast of each of these sections is improved independently. Histogram equalization is applied to each region, and as a result, the distribution of grey values is altered, which enables previously hidden features in the image to become more visible. Using this approach leads to an improvement in the picture's overall quality. After applying CLAHE to the green channel, finer details in the fundus image, such as blood vessels, microaneurysms, haemorrhages, and exudates, may be seen with greater clarity [25].

Support Vector Machine (SVM) is a cutting-edge technique that offers the possibility of making groundwater investigation more straightforward. Images are often shown in computer vision demonstrations in the form of non-linear matrices that are composed of square pixels. This is done to ensure that the image is shown in the most accurate manner possible. The method of image classification may then make use of the characteristics of the target item that have been retrieved from the picture. When seeking to gather information from photos, one may make use of the tried-and-true techniques of feature extraction. The SVM does this task by first constructing a hyperplane between the non-linear data points included inside the pictures. After doing so, it then assigns each picture pixel to one of two classes. It is feasible that the classification accuracy may improve if a greater distance was placed between the hyperplane and the data point. In addition to this, the support vector machine (SVM) [26] may be used to derive valuable characteristics from photographic images.

Researchers at the University of Michigan came up with the innovative technique known as AdaBoost to increase gradients. This approach works well for judgments with a yes or no answer. After constructing the first tree decision tree and analyzing how well it did on the training data, the effectiveness of the tree is evaluated. In this section, the utility of the tree decision tree will be evaluated. This technique brings together a number of distinct categorization strategies under the umbrella of a single overarching plan. The training data are used to construct the first model, and if there are flaws with that model, subsequent models are constructed to solve those problems. Model development is considered to be finished after all possible models have been constructed or if accurate predictions can be made using data from the training set. The result of merging numerous different models is the best classification model that has been developed up to this point [27].

When applied to enormous datasets, the effectiveness of machine learning methods is dramatically reduced as a result of challenges such as underfitting, model complexity, and an inability to make efficient use of resources. The efficiency of the techniques is diminished as a direct result of these impediments. Deep learning networks, when applied to large datasets, have the potential to disclose insights that were previously unknown, generate accurate predictions, and even automate processes that were previously performed manually. Computer models are now able to learn from both graphical and written data thanks to the application of "deep learning." As the amount of data available has increased, a number of different deep learning models have been developed to accomplish extraordinary performance in comparison to more conventional machine learning techniques.

The convolutional neural network is one of the most widespread forms of the deep neural network architecture that is used in the field of computer vision (CNN) [28]. The components that make up a CNN are referred to as the convolutional layer, the pooling layer, the activation layer, and the connected convolutional layer. A deep CNN makes use of a large number of convolutional layers that are linked in order to carry out operations from beginning to finish. Filtering is at the heart of the convolutional layer. At the convolutional layer, just a small number of the input picture's pixels—say let's 3×3—are allowed to get through the filter.

A "dot" operation is carried out on the pixel values by the filter, and a weight is utilized in order to determine the degree to which this has an impact on the outcome. Because of this, the size of the image's data point matrices has decreased as a result of the application of the convolutional layer. Back propagation is used to train the network with the activation layer's input vectors, and the activation layer's matrix is used to incorporate nonlinearity into the network. When samples are combined into a single group, both the number of levels present in the samples and the amount of space occupied by the filter matrix are decreased. This particular kind of layer is referred to as a minimal layer because it only selects one quality for each category. The result of the max layers is used by the connected layer to generate a probability distribution over the potential names. This distribution is compiled by the connected layer. In order to determine these chances, the connected layer is utilized. When classifying something, the result that is expected to occur most frequently is taken into consideration.

4 Results and discussion

For the purpose of this study, we made use of a dataset consisting of 174 satellite images of monthly groundwater levels taken between March 2002 and May 2019. Each image has a resolution of 360 by 180 pixels and uses a color-coded system to depict the TWS of the landmass of the Earth. These images were first obtained from the GRACE study carried out by the National Aeronautics and Space Administration of the United States (NASA). The original data set was obtained from the online database maintained by the Physical Oceanography Distributed Active Archive Center (PO. DAAC) [29]. Figure 2 below shows the ground water satellite image.

Classification Accuracy, Precision, Recall, and F1 Score are the most important variables to consider while evaluating performance. When attempting to explain the performance of the models based on the test data, a representation approach known as a confusion matrix is often used. The performance of the machine learning techniques can be visualised with the help of the confusion matrix. The confusion matrix is capable of calculating a total of four different values, including True Positive (TP), False Negative (FN), True Negative (TN), and False Positive (FP). The term “True Positive” refers to the fact that the model has determined a certain amount of information to be positive and that information itself is positive. The term “True Negative” refers to the situation in which the whole amount of categorised data is negative and has been marked as such.

In addition, a False Positive is defined as the number of data that has been identified as positive while having a negative value. The term “type 1 mistake” may also be used to refer to a “false positive.“ In conclusion, the False Negative is a representation of the fact that the quantity for the data is negative and is categorised as negative. The False Negative is a sort of mistake that is also known as type 2. The classification accuracy, precision, recall, and F1 score of deep convolutional neural networks, as well as those of current machine learning and transfer learning approaches, can all be calculated using these variables.Results are shown below in Figs. 3, 4 and 5, figure and Fig. 6.

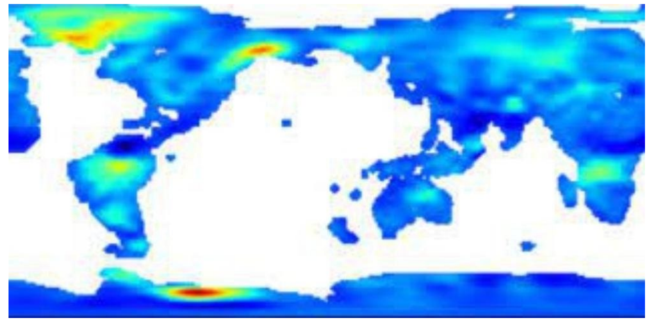

Fig. 2 Ground water satellite image

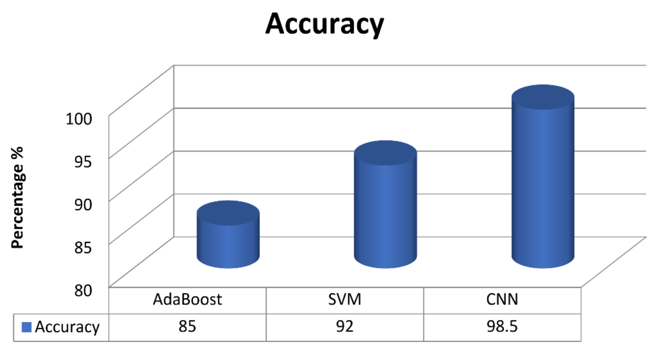

Fig. 3 Accuracy of Classifiers for Ground Water Level Prediction

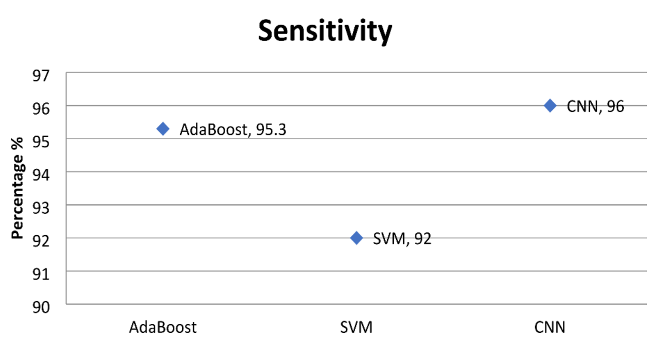

Fig. 4 Sensitivity of Classifiers forGround Water Level Prediction

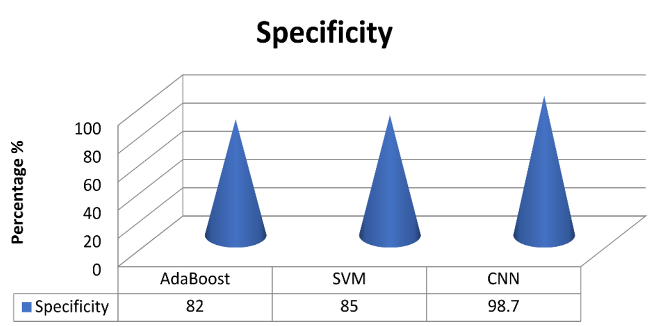

Fig. 5 Specificity of Classifiers for Ground Water Level Prediction

Accuracy of CNN algorithm is 98.5%. It is 6.5% higher than the accuracy of SVM technique. Accuracy of AdaBoost technique is 85%. Sensitivity of CNN algorithm is 96%. Sensitivity of SVM is 92%, it is 4% less than the sensitivity

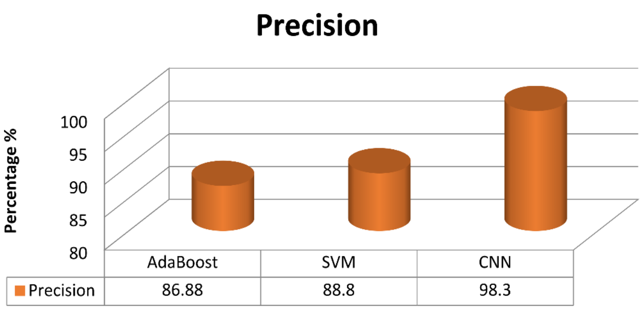

Fig. 6 Precision of Classifiers for Ground Water Level Prediction

of CNN algorithm. Sensitivity of AdaBoost techniques is slightly less than the sensitivity of CNN algorithm.

Specificity of CNN method is 98.7%, It is 13.7% higher than the specificity of SVM technique and 16.7% higher than the specificity of AdaBoost technique.

Precision of CNN is 98.3%. It is highest among CNN, SVM and AdaBoost techniques. Recall rate of SVM algorithm is 98.2%, which is 1.6% less than the recall rate of CNN method.

5 Conclusion

In contrast to the atmosphere and fresh surface water, which can only temporarily store water, the natural water cycle might use groundwater as a "reservoir" that stores water for extended periods. This would contrast the atmosphere and fresh surface water, which can only store water temporarily. Even though the subsurface environment exhibits a significant degree of diversity and complexity, there is incredibly low availability of data collected from the field. Those individuals who used models that were founded on actual reality were the ones who were confronted with both of these issues. To meet the water demands of a growing worldwide population, groundwater has emerged as a resource of increasing significance, and this is true in both economically developed and developing countries. Even when there is a dry spell or a drought, it is possible to utilise it again since there is such a vast store. This article presents a model for forecasting the level of groundwater that is built on a deep convolutional neural network. As part of the experimental setup, 174 satellite pictures of groundwater are included in the input data set. Images are preprocessed using the CLAHE method. The CNN, SVM, and AdaBoost methods make up the classification model. The results have shown that CNN is able to classify things correctly 98.5 per cent of the time. The recall rate of CNN is also 98.2 per cent.

Acknowledgements This study is supported via funding from Prince Sattam bin Abdulaziz University project number(PSAU/2023/R/1444).

Data availability Data shall be made available on request.

Declarations

Compliance with ethical standards The authors have no conflicts of interest. No human or animal participation is involved in this research.

References

1. Yadav, B., Ch, S., Mathur, S., & Adamowski, J. (2017). Assessing the suitability of extreme learning machines (ELM) for groundwater level prediction. *J Water Land Dev*, *32*, 103–112.
2. Zhou, T., Wang, F., & Yang, Z. (2017). Comparative analysis of ANN and SVM models combined with wavelet preprocess for groundwater depth prediction. *Water*, *9*, 781.
3. Chen, Z., Cong, B., Hua, Z., Cengiz, K., & Shabaz, M. (2021). Application of clustering algorithm in complex landscape farmland synthetic aperture radar image segmentation. In Journal of Intelligent Systems (Vol. 30, Issue 1, pp. 1014–1025). Walter de Gruyter GmbH. https://doi.org/10.1515/jisys-2021-0096.
4. Mirzavand, M., & Ghazavi, R. (2015). A stochastic modelling technique for groundwater level forecasting in an arid environment using time series methods. *Water Resource Management*, *29*, 1315–1328.
5. Rahaman, M. M., Thakur, B., Kalra, A., Li, R., & Maheshwari, P. (2019). Estimating High-Resolution Groundwater Storage from GRACE: A Random Forest Approach. *Environments*, *6*, 63.
6. Jing, W., Yao, L., Zhao, X., Zhang, P., Liu, Y., Xia, X., Song, J., Yang, J., Li, Y., & Zhou, C. (2019). Understanding terrestrial water storage declining trends in the Yellow River Basin. *J Geophys Res Atmos*, *124*, 12963–12984.
7. Sahour, H., Sultan, M., Vazifedan, M., Abdelmohsen, K., Karki, S., Yellich, J. A., Gebremichael, E., Alshehri, F., & Elbayoumi, T. M. (2020). Statistical applications to downscale GRACE-derived terrestrial water storage data and to fill temporal gaps. *Remote Sens*, *12*, 533.
8. Mukherjee, A., & Ramachandran, P. (2018). Prediction of GWL with the help of GRACE TWS for unevenly spaced time series data in India: Analysis of comparative performances of SVR, ANN and LRM. *Journal Of Hydrology*, *558*, 647–658.
9. Sugeno, M., & Yasukawa, T. (1993). A fuzzy-logic-based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems*, *1*(1), 7.
10. B´ardossy, A., & Disse, M. (1993). Fuzzy rule-based models for infiltration. *Water Resources Research*, *29*(2), 373–382.
11. Panigrahi, D. P. (2000). .Mujumdar,"Reservoir operation modelling with fuzzy logic. *Water Resources Management*, *14*(2), 89–109.
12. Awasthi, A. K., Dubey, O. P., Awasthi, A., & Sharma, S. (2005). "A fuzzy logic model for estimation of groundwater recharge", in Proceedings of NAFIPS 2005–2005 Annual Meeting of the North American Fuzzy Information Processing Society, pp. 809–813.
13. Alvisi, S., Mascellani, G., & Franchini, M. (2006). and A. Bardossy,"Water level forecasting through fuzzy logic and artificial neural network approaches", Hydrology and Earth System Sciences Discussions, vol. 10, no. 1, pp.1–7.
14. Affandi, A. K. (2007). Watanabe,"Daily groundwater level fluctuation forecasting using soft computing technique. *Nature and Science*, *5*(2), 1–10.
15. Solaimani, K. (2009). A study of rainfall forecasting models based on artificial neural network. *Asian Journal of Applied Sciences*, *2*(6), 486–498.
16. Fernandez, N., & Jaimes, W. (2010). Altamiranda," neuro-fuzzy modeling for level prediction for the navigation sector on the

Magdalena River (Colombia). *Journal of Hydroinformatics*, *12*(1), 36–50.
17. Mayilvaganan, M. K. (2011). Naidu,"Comparison of membership functions in adaptive-network-based fuzzy inference system (ANFIS) for the prediction of groundwater level of a watershed. *J Computer Appl Res Dev*, *1*, 35–42.
18. Gharde, K. D., & Kothari, M. (2016). and D. M. Mahale," forecasting runoff and sediment yield by ANN and fuzzy logic algorithms for Kal river, India", Current World Environment, vol. 11, no. 3, pp. 892.
19. Ghazavi, R., & Babaei, S. (2018). Erfanian,"Recharge wells site selection for artificial groundwater recharge in an urban area using fuzzy logic technique. *Water Resources Management*, *32*(12), 3821–3834.
20. Maier, H., & Dandy, G., "Neural networks for the prediction and forecasting of water resources variables: A review of modeling issues and applications", *Environ Modell Software*, vol. 15, no. 1, pp. 101–124.
21. Lohani, A. K., Goel, N. K., & Bhatia, K. K. S. (2006). Takagi–Sugeno fuzzy inference system for modeling stage–discharge relationship. *Journal Of Hydrology*, *331*(1), 146–160.
22. Yan, X., Zhang, C., Luo, W., Li, W., Chen, W., & Liu, H. (2012). Solve traveling salesman problem using particle swarm optimization algorithm. *International Journal of Computer Science Issues*, *9*, 264–271.
23. S. Chintalapati,"Groundwater level forecasting using SVM-PSO", *International Journal of Hydrology Science and Technology*, vol. 2, no. 2, pp. 202–218.
24. Mahnam, M., & Ghomi, S. M. T. F. (2012). A particle swarm optimization algorithm for forecasting based on time variant fuzzy time series. *International Journal of Industrial Engineering and Production Research*, *23*(4), 269–276.
25. Ravikumar, M., Rachana, P. G., Shivaprasad, B. J., & Guru, D. S. (2021). Enhancement of Mammogram images using CLAHE and bilateral Filter Approaches. In V. K. Gunjan, P. N. Suganthan, J. Haase, & A. Kumar (Eds.), *Cybernetics, Cognition and Machine Learning Applications*. Singapore: Algorithms for Intelligent Systems.Springer.
26. Lohani, T. K., Ayana, M. T., Mohammed, A. K., Shabaz, M., Dhiman, G., & Jagota, V. (2023). A comprehensive approach of hydrological issues related to ground water using GIS in the hindu holy city of Gaya, India. *World Journal of Engineering*, *20 No*(2), 283–288. https://doi.org/10.1108/WJE-04-2021-0223.
27. Mitra, A., Jain, A., Kishore, A., et al. (2022). A comparative study of demand forecasting models for a Multi-Channel Retail Company: A Novel Hybrid Machine Learning Approach. *Oper Res Forum*, *3*, 58. https://doi.org/10.1007/s43069-022-00166-4.
28. Durairaj, D. M., & Mohan, B. H. K. (2022). A convolutional neural network based approach to financial time series prediction. *Neural Comput&Applic*, *34*, 13319–13337. https://doi.org/10.1007/s00521-022-07143-2.
29. Felix Landerer (2020). JPL TELLUS GRACE Level-3 Monthly Land Water-Equivalent-Thickness Surface Mass Anomaly Release 6.0 Version 03 in netCDF/ASCII/GeoTIFF Formats; Ver. RL06 v03; PO.DAAC: Pasadena, CA, USA.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.