Mustansiriyah Journal of Pure and Applied Sciences, Print ISSN: 2957-9910. Online ISSN: 2957-9929, Vol. 4, No. 3, 2026, Pages 171-181
DOI: 10.47831/mjpas.v4i3.418

MJPAS

MUSTANSIRIYAH JOURNAL OF PURE AND APPLIED SCIENCES

Journal homepage: https://mjpas.uomustansiriyah.edu.iq/index.php/mjpas

RESEARCH ARTICLE - PHYSICS

Using Deep Learning Technology to Detecting Kidney Disease

Ahmed Sami Salman[1]*, Ziad Mohamed Abood[2]

[1] The Directorate of Education for Second Cark in Baghdad, Ministry of Education, Baghdad, Iraq

[2] Department of Physics, College of Education, Mustansiriyah University, Baghdad, Iraq

* Corresponding author's Email: ahmedsami198088@uomustansiriyah.edu.iq

Article Info.

Article history:

Received
23 April 2025

Accepted
4 June 2025

Publishing
30 September 2026

Abstract

Deep learning is gaining significant importance due to data interpretation and its application to a wide range of diseases in general, especially kidney disease detection. Machine learning is widely applied in healthcare. Providing a sufficient number of samples (images) and providing the appropriate statistical software were major research challenges. Image preprocessing was then performed to convert the original medical image data into medical image data free of some unwanted distortions (noise). This was done to enhance the images, which must be of the same size and dimensions, to reveal certain image features, and transform them into high-quality medical images for use in kidney disease detection and classification. Ten statistical image metrics were used: SC, AD, MD, SSIM, RFSIM, FSIM, RMSE, LMSE, MAE, and PCC. These metrics were used to evaluate the performance of the two classification algorithms (CNN and RNN). The classification process used precision, accuracy, recall, and F1 metrics. High accuracy results were achieved, indicating that the resulting results are very good. The evaluation metrics provide the information needed to determine the performance of a classification model based on a given score (correct or incorrect). The study focused on the use of deep learning in the recognition and classification of kidney diseases. The results of the current study indicated that the CNN and RNN algorithms used in the proposed kidney disease recognition and classification system achieved acceptable results. The results achieved high classification accuracy for the proposed method for kidney diseases (90.52%, 64.64%) and for normal cases (94.05%, 77.12%) for the CNN and RNN algorithms, respectively.

The official journal published by the College of Education at Mustansiriyah University

Keywords: Deep Learning, Kidney diseases, Medical imaging, Feature extraction, Classification.

1. Introduction

Deep learning is a branch of artificial neural networks and is widely used in scientific computing. Deep learning techniques are used in a wide range of sectors to solve complex problems. All deep learning algorithms use various forms of neural networks. It is a branch of machine learning. It has been used in diverse applications, including areas such as computer vision, medical applications, image classification, multimedia concept retrieval, and medical information processing. Among the various machine learning algorithms, deep learning (DL) is widely used in these applications [1] [2].

Deep learning (DL) has been considered the gold standard for computing. Furthermore, it has gradually become the most widely used computational approach, achieving impressive results on many complex cognitive tasks, matching or even surpassing human performance. One advantage of deep learning is its ability to learn massive amounts of data [3].

Feature extraction is automated in all deep learning algorithms, discovering their best performance, and extracting discriminative features using minimal human effort and domain knowledge. These algorithms feature a multi-layer data representation structure, where the first layers extract low level features, while the last layers extract high level features [4].

Deep learning, due to its great success, is one of the most prominent applications and various computational tools. Key concepts, structures, and evolutionary matrixes are discussed. Convolutional neural networks (CNNs) are among the most popular and widely used deep learning networks, along with recurrent neural networks (RNNs). Thanks to CNNs and RNNs, deep learning has become very popular. The main advantage of CNNs and RNNs over others is that they automatically detect important features without any human supervision, which has made them the most widely used. Therefore, we delve into CNNs and RNNs by introducing their main components [5] [6].

In this paper, we provide an overview of deep learning algorithms and their various applications in solving big data problems. This paper also contributes to the use of specific algorithms and programs and their application to other diseases.

2. Deep learning (DL)

Deep learning has evolved as a powerful machine learning method; DL has demonstrated impressive performance in various application areas, particularly image classification. Deep learning techniques provide accurate image classification performance, aiming to distinguish between classes [7].

Healthcare is one of the most important and widespread applications of deep learning. Furthermore, deep learning has demonstrated tremendous performance in healthcare applications. Therefore, we take deep learning applications in medical image analysis as an example to describe them [8].

Medical imaging has evolved as an important part of medical science. Therefore, image resolution and accuracy are highly desirable for rapid and timely diagnosis, and design deep learning algorithms [9].

Deep learning offers significant potential for solving problems that have resisted the efforts of the artificial intelligence community for many years. It has proven highly effective in detecting complex structures in large datasets and is therefore applicable to many areas of medical science. Deep learning has achieved very promising results in various tasks, especially medical image classification, allowing it to easily take advantage of the increasing amount of available data [10] [11].

Various deep learning applications are widely spread. These applications include healthcare, neural network analysis, image processing (e.g., recognition and enhancement), and visual data processing methods (e.g., multimedia data analysis and computer vision), these applications have been classified into five categories: classification, localization, detection, segmentation, and registration. Although each of these tasks has its own goal, there is a fundamental overlap in the pipeline implementation of these applications. Classification is a concept that classifies a set of data into classes. Detection is used to locate objects of interest in an image while taking the background into account [12] [13].

Several performance features are used for deep learning: [14] [15]

1. End-to-end learning: Deep learning has the potential to perform in virtually all application areas.
2. Robustness: In general, fine-tuned features are not required in deep learning techniques. Instead, optimized features are learned automatically. This ensures robustness to typical changes in the input data.
3. Generalization: Different data types or different applications can use the same deep learning technique, an approach often referred to as transfer learning (TL).

Convolutional neural networks (CNNs) and recurrent neural networks (RNNs), along with many available large-scale public visual datasets, have reached a new and distinct stage. Using well-trained CNNs and RNNs, training datasets can learn comprehensive features by leveraging the initial features generated by transfer learning. However, performance is not guaranteed when using these features to build a new model, as there are always differences between the source and target domains. Experimental results efficiently achieve optimal solutions, demonstrating a time-consuming feasibility beyond human capabilities [16] [17].

These classifications operate simultaneously, each constructing distinct models based on a set of training and test data [18].

The classification algorithms are tested on a kidney disease classification dataset.

3. Methodology

In this paper, techniques based on deep learning algorithms that are widely used in distinguishing kidney diseases, such as Convolutional neural networks (CNN) and recurrent neural networks (RNN), were used in the training and testing stages.

The images containing the region of interest in kidney disease were selected. Total image 750 images were classified into 250 images for healthy cases and 500 images for kidney disease cases. The size of the images should be 250×250. Then the resulting data was divided into two stages, training and testing.

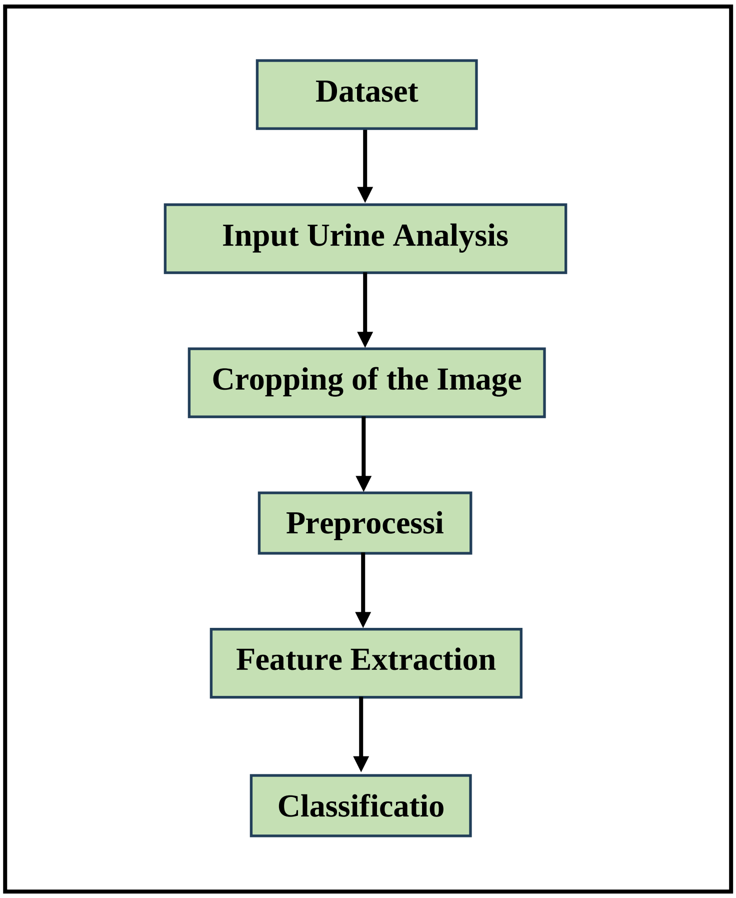

Figure (1): A diagram of the classification method

3.1 Data Collection

To evaluate the performance of the three classification algorithms, we collected (1000) medical images related to kidney diseases obtained from Al-Karamah Hospital, it consists of (670) healthy images and (330) images with kidney disease (kidney stone disease, hronic kidney disease) used in the proposed classification system.

Figure (3) shows examples of healthy medical cases and cases with two different kidney diseases, namely kidney stone disease, chronic kidney disease.

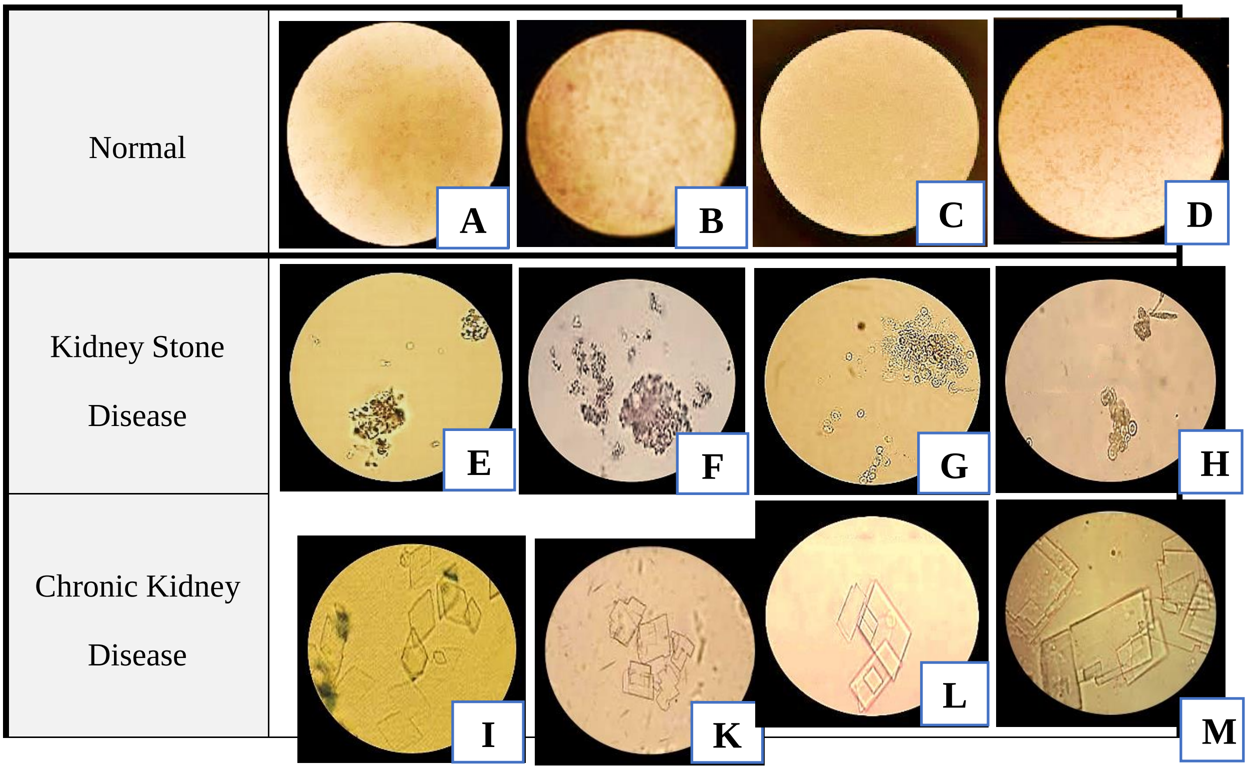

Figure (2): Represents medical images of normal medical cases (A-D) and cases with kidney diseases (E-M).

3.2 Preprocessing stage

Image preprocessing is a method of converting the original medical image data into medical image data free of some unwanted artifacts (noise) that are used in the process of disease detection and classification.

The purpose of preprocessing is to enhance images that should be of the same size and dimensions to show some image features and convert them into a high-quality medical image to be fed into the classification model.

Image preprocessing is a smart idea, because it changes or modifies the true nature of the raw data. However, the smart application of image preprocessing can provide benefits and address issues that ultimately lead to improved medical images.

Image preprocessing has beneficial effects on the quality of feature extraction from images and the results of medical image analysis because it reduces the training time of the model and improves it without affecting its performance.

3.3 Proposed Method

The proposed method consists of two main stages, namely the training stage and the testing stage. The training stage and the testing stage used medical images of normal healthy cases and sick healthy cases with medical symptoms of kidney diseases, which were carefully diagnosed by a nephrologist.

The next Third step was to crop the region of interest from a set of images captured by a microscope linked to a camera connected to a computer to capture images of kidney disease imaging samples. The fourth step involved preprocessing, which involved cropping the region of interest with different dimensions and sizes. There is a slight difference between one image and another in size and thus there is a difference in the number of pixels, so the size of all images was standardized from 515×515 to 250×250 pixels. This gives accurate data results.

Image features were extracted in the fifth stage by defining image quality metrics for the data, part of which was fed as training data into the classification algorithms, followed by the final step, which is the testing stage, where accuracy requirements were defined.

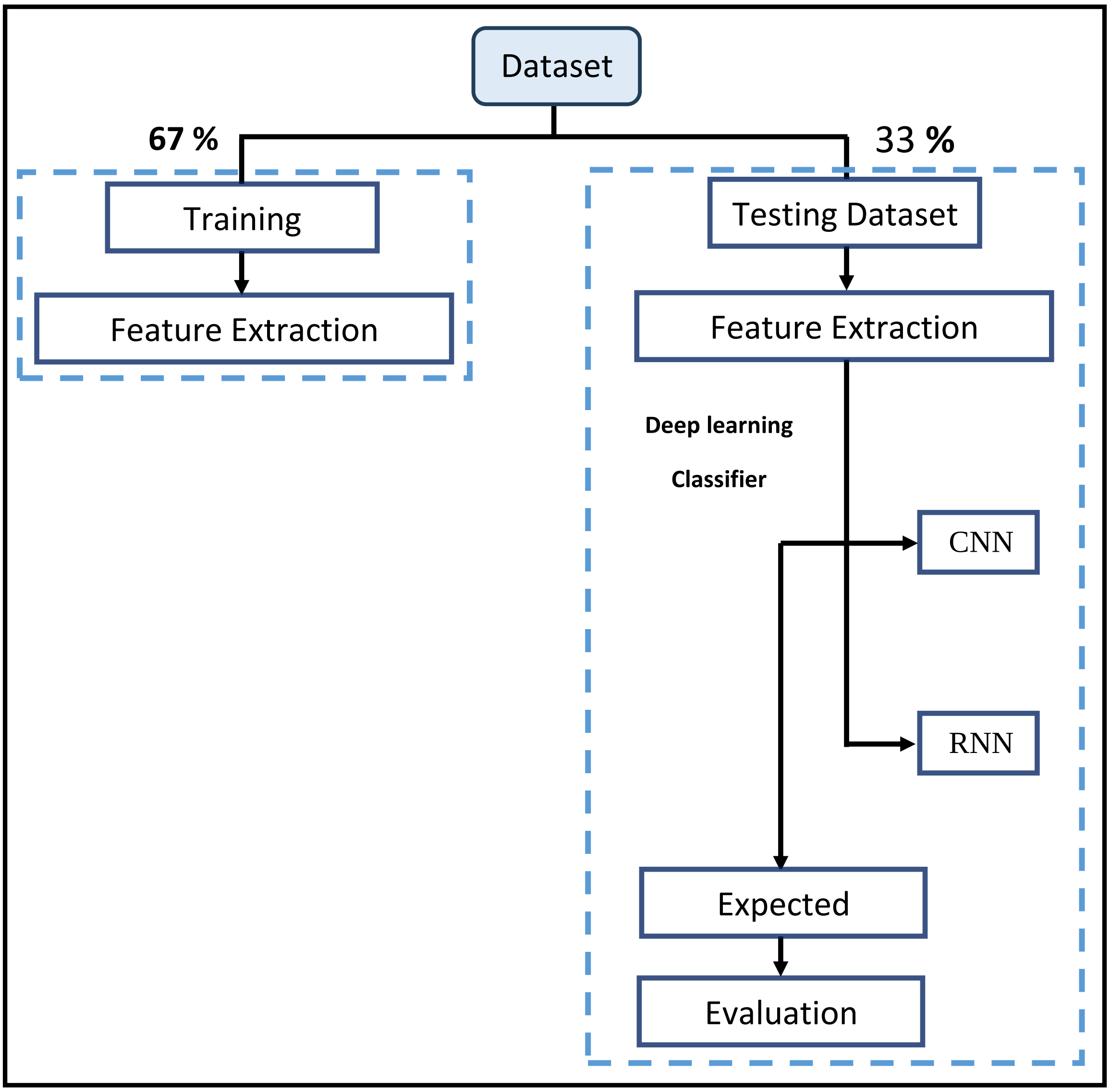

Figure (3): A diagram of the proposed system

3.4 Training stage

This research used (1000) images, 67% of which were data collected from the study samples for the training phase. This phase consists of (670) images used to train the proposed classification system, including (221) images of healthy cases and (449) images of cases with kidney disease, representing (67%) of the dataset. This phase is important in building the research platform. The basic rules are then constructed, as shown in Figure (2), followed by the testing phase.

3.5 Testing stage

This phase includes data testing, which represents 33% of the total data collected from the study samples. This phase consists of (330) images used to test the proposed classification system, including (109) images of healthy cases and (221) images of cases with kidney disease, representing (33%) of the dataset. This is followed by the classification phase, where the system analyzes actual data to achieve the highest accuracy for the proposed system and demonstrate its ability to learn accurately.

Testing the dataset is important because it ensures that the model can be successfully generalized to any data that was unknown, and as a result the model is verified and it is ensured that the dataset does not contain any training or validation data, as shown in Figure (2).

3.6 Feature Extraction

In this study, 10 image statistical metrics were used to measure image quality: (SC, AD, MD, SSIM, RFSIM, FSIM, RMSE, LMSE, MAE, and PCC).

3.7 Classification

This study tested several machine learning algorithms. Because the most accurate methods were selected convolutional neural networks (CNN) and recurrent neural networks (RNN), they were compared to determine the best. Classification accuracy, recall, precision, and the f1-measure were used to select the best model.

3.7.1 Convolutional Neural Networks (CNN)

In the field of deep learning, CNN is the most popular and widely used algorithm. CNNs also known as ConvNets, are multilayer neural networks that are primarily used for mostly employed in image processing and the interpret medical imaging. The main advantage of CNN over others is that it automatically identifies relevant features without any human supervision. CNN has been widely applied in a variety of different fields, including computer vision, face recognition, and most notably in medicine. The architecture of CNN is inspired by neurons in the brains of humans, similar to a traditional neural network. This chain is simulated by CNN. [19]

3.7.2 Recurrent Neural Networks (RNN)

RNNs are a commonly employed and familiar algorithm in the discipline of DL. RNN is mainly applied in the area of Medical. Since the embedded structure in the sequence of the data delivers valuable information, this feature is fundamental to a range of different applications. Tus, it is possible to consider the RNN as a unit of short-term memory. RNNs hidden layers have the ability to remember information. RNN may be used to predict time series since it has Long Short-Term Memory, which allows it to remember inputs. [20]

3.8 Related works

- Fuzhe, and et al., [21] IN (2020), the prevalence of chronic kidney disease (CKD) increases annually. One of the sources for further therapy is the CKD prediction where the Machine learning techniques become more important in medical diagnosis due to their high accuracy classification ability. The accuracy of classification algorithms depends on the proper use of algorithms for feature selection to reduce the data size. Heterogeneous Modified Artifical Neural Network (HMANN) has been proposed for the early detection, segmentation, and diagnosis of chronic renal failure on the Internet of Medical Things (IoMT) platform. Furthermore, the proposed HMANN is classified as a Support Vector Machine and Multilayer

Perceptron (MLP) with a Backpropagation (BP) algorithm. The proposed algorithm works based on an ultrasound image which is denoted as a preprocessing step and the region of kidney interest is segmented in the ultrasound image. In kidney segmentation, the proposed HMANN method achieves high accuracy and significantly reducing the time to delineate the contour.

- Kumar, and et al., [22] IN (2023) chronic kidney disease (CKD) is a gradual decline in renal function that can lead to kidney damage or failure. Using routine doctor consultation data to evaluate various stages of CKD could aid in early detection and prompt intervention. Advanced machine learning algorithms have been developed to detect the presence of kidney disease. Presents a novel deep learning model, which combines a fuzzy deep neural network, for the recognition and prediction of kidney disease. The results show that the proposed model has an accuracy of 99.23%, which is better than existing methods. Furthermore, the accuracy of detecting chronic disease can be confirmed without doctor involvement as future work. Compared to existing information mining classifications, the proposed approach shows improved accuracy in classification, precision, F-measure, and sensitivity metrics.

- Saif, and et al., [23] IN (2024) the medical sector, it has been utilized for predicting the likelihood of getting a health condition such as chronic kidney disease (CKD). The area of early CKD prediction through a high-performing deep learning and ensemble framework. We bridge the gap between existing detection methods and preventive interventions by: developing and comparing deep learning models like CNN, LSTM, and LSTM-BLSTM for (6–12) month CKD prediction; addressing data imbalance, feature selection, and optimizer optimization; and building an ensemble model combining the best individual models (CNN-Adamax, LSTM-Adam, and LSTM-BLSTM-Adamax). Our framework achieves significantly higher accuracy (98% and 97% for 6 and 12 months), paving the way for earlier diagnosis and improved patient outcomes.

3.9 Evaluation Metrics

Various evaluation metrics are presented, such as precision, recall, F-score, development, training, and evaluation sets. These metrics are measured based on a confusion matrix and statistical significance testing and are then used to evaluate machine learning ensembles applied to data [24].

Because evaluation metrics provide the information needed to determine the performance of a classification model, summarizing this information into a single number makes it easier to compare the relative performance of different models. This can be done using a metric such as accuracy [25]. Which is calculated as follows:

Table (1): Evaluation Metrics Prediction [25]

Class	Predicted class	
	Positive	Negative
Positive	True Positive	False Negative
Negative	False Positive	True Negative

The cases shown in Table (1) can be classified using the following terms: [26], [27]

1) True Positive Rate (TP): The percentage of positive cases that were accurately identified. This indicates that the system has detected them correctly.

2) False Negative Rate (FN): The percentage of negative cases that were identified as positive. This indicates the presence of the disease and the system determines that the image is normal.

3) False Positive Rate (FP): The percentage of positive cases that were identified as negative. This indicates that there is no disease and that the system has detected the presence of diseases.

4) True Negative Rate (TN): The percentage of negative cases that were correctly identified. This indicates that there is no disease, and this confirms that the system has detected the presence of diseases.

3.10 Classification Measure

Several performance measures are used to assess the effectiveness of the suggested technique: [28]

1. Accuracy (Acc): The accuracy can be high, which means that the system works well and the results produced are highly accurate, which are defined as:

$$\text{Accuracy (Acc)} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

2. Precision (Pre): the correlation between real positive predicted values and complete positive predicted values, represented as:

$$\text{Precision (Pre)} = \frac{TP}{TP + FP} \qquad (2)$$

3. Recall (Rec): Recall measures the number of correct instances retrieved divided by all correct instances. A confusion matrix, as:

$$\text{Recall (Rec)} = \frac{TP}{TP + FN} \qquad (3)$$

4. F1-Measure (F1): The F1-Measure means the harmonic mean between precision and recall, The F1-Measure can have different indices giving different weights to precision and recall, which is written as:

$$\text{F1-Measure (F1)} = \frac{2 \text{ x } (PR * RC)}{PR + RC} \qquad (4)$$

4. Results

The image quality metrics used in the current study showed significant differences in diagnosing and distinguishing kidney diseases. Based on the image quality metrics, the study results showed that the algorithms used in the proposed kidney disease classification system provide acceptable results, through performance metrics that included (accuracy rate, precision, recall, and F1-measure) for normal cases and kidney diseases, as shown in Table (2).

Table (2) Results of the proposed performance measures

Ways	Performance Metrics							
	Acc		Pre		Rec		F1	
	D	N	D	N	D	N	D	N
CNN	90.52%	94.05%	92.74%	95.13%	92..20%	96.82%	94.37%	93.19%
RNN	64.64%	77.12%	59.33%	69.64%	59.75%	81.88%	60.29%	71.94%

5. Discussion of Results

This study was conducted to identify kidney disease cases using computer-aided analysis of medical image data obtained from Al-Karama Hospital in Baghdad. More than 750 medical images were collected, one-third of which were normal cases, which constituted 33%, and about two-thirds were kidney disease cases, which constituted 67% of the total. In methods used to detect kidney disease, choosing the right features is more important than choosing a classifier. This was achieved by taking advantage of a number of features.

The efficiency of the proposed method was evaluated by comparing it with related work. Our proposed method outperforms other related methods based on traditional machine learning methods.

This study use matlab (2023b) in the proposed method. Core i5 processor and (8 GB) of RAM constitute the system resources. Processing and extraction (IQM) to be performed. Windows application created in (64-bit) Python programming language with library is the other operating system presented. Random Forests (RF), K-Nearest Neighbor (KNN), and Support Vector Machines (SVM) were the classification approaches used for training.

Convolutional neural networks (CNN) and recurrent neural networks (RNN), were the classification approaches used for training. Convolutional neural networks (CNN) showed the best performance in the final comparison of the results obtained by different algorithms used for prediction, followed by recurrent neural networks (RNN), which produced the lowest result.

6. Conclusion

This research describes a deep learning-based method for classifying kidney diseases. The results demonstrated the effectiveness of the proposed technique in accurately classifying kidney diseases. Furthermore, the results showed that convolutional neural networks (CNNs) performed better than recurrent neural networks (RNNs). The proposed techniques using deep learning methods performed well and achieved high accuracy. This suggests that the proposed method can be effectively used to diagnose and classify kidney diseases. As a result, medical examinations will be less expensive and diagnoses will be of higher quality. Future research will focus on using these techniques to detect various other diseases.

7. Acknowledgments

This study conducted at the College of Education, Mustansiriyah University. The authors would like to express their gratitude to Al-Karamah Hospital, Baghdad, Iraq, for providing logistical support by providing the necessary study samples to conduct the research.

References

[1] T. E. Potok, et al, "A study of complex deep learning networks on high-performance, neuromorphic, and quantum computers." *ACM Journal on Emerging Technologies in Computing Systems (JETC)* 14, no. 2, pp. 1-21, 2018.

[2] X. Youzi, Z. Tian, J. Yu, Y. Zhang, S. Liu, S. Du, and X. Lan, "A review of object detection based on deep learning." Multimedia Tools and Applications 79, pp. 23729-23791, 2020.

[3] A. M. Shhatha, and O. I. Alsaif, "A comprehensive Analysis of Approaches and Difficulties for Cybersecurity Threats-Article Review." *Mustansiriyah Journal of Pure and Applied Sciences 2, no. 4*, pp. 126-139, 2024.

[4] L. Yann, Y. Bengio, and G. Hinton, "Deep learning" *nature journal 521*, no. 7553: pp. 436-444, 2015.

[5] Y. Guangle, T. Lei, and J. Zhong, "A review of convolutional-neural-network-based action recognition" Pattern Recognition Letters 118, pp. 14-22, 2019.

[6] D. Anamika, and G. K. Verma, "Convolutional neural network: a review of models, methodologies and applications to object detection" Progress in Artificial Intelligence 9, no. 2, pp. 85-112, 2020.

[7] T. Haiman, S.C. Chen, and M.L. Shyu, "Evolutionary programming based deep learning feature selection and network construction for visual data classification" Information systems frontiers 22, no. 5, pp. 1053-1066, 2020.

[8] L. Geert, et al., "A survey on deep learning in medical image analysis" *journal medical image analysis 42*, pp. 60-88, 2017.

[9] M. Debadyuti, R. Mondal, P. K. Singh, R. Sarkar, and D. Bhattacharjee, "EnsemConvNet: a deep learning approach for human activity recognition using smartphone sensors for healthcare applications" Multimedia Tools and Applications 79, pp. 31663-31690, 2020.

[10] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R.l Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics" *Journal of big data 2*, pp.1-21, 2015.

[11] V. J. Sravya, and S. Malathi, "Deep learning for multi grade brain tumor classification in smart healthcare systems: a prospective survey" In Smart Intelligent Computing and Communication Technology, pp. 321-325. IOS Press, 2021.

[12] C. Pauley, and J. C. Mandell. "Current applications and future directions of deep learning in musculoskeletal radiology" Skeletal radiology 49, no. 2, pp. 183-197, 2020.

[13] F. A. Al-sudani, and H. A. Al-Afare, "Age estimation from face image using Hybrid Representation for Deep Learning", *Mustansiriyah journal of pure and Applied Sciences (MJPAS)*, vol. 23, no. 2, pp. 35-48, 2022.

[14] H. Kaiming, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition" *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016.

[15] V. E. Brian, H. Kim, R. Pearce, K. Boakye, and B. Chen, "LBANN: Livermore big artificial neural network HPC toolkit" In Proceedings of the workshop on machine learning in high-performance computing environments, pp. 1-6, 2015.

[16] Q. V. Le, "A tutorial on deep learning part 2: Autoencoders, convolutional neural networks and recurrent neural networks", Google Brain, vol. 20, pp. 1-20, 2015.

[17] Y. Hang, L. T. Yang, Q. Zhang, D. Armstrong, and M. J. Deen. "Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives" Neurocomputing 444, pp. 92-110, 2021.

[18] S. Korsuk, et al., "Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images" IEEE transactions on medical imaging 35, no. 5, pp. 1196-1206, 2016.

[19] J. Gu, Z Wang, J. Kuen, and et. al., "Recent advances in convolutional neural networks, Pattern Recognition" vol. 77, pp. 354-377, 2018.

[20] H. Hewamalage, C. Bergmeir, and K. Bandara, "Recurrent neural networks for time series forecasting: current status and future directions", Int J Forecast., 37(1): pp. 388–427, 2020.

[21] M. Fuzhe, T. Sun, L. Liu, and H. Jing. "Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network" *Future Generation Computer Systems 111*, pp. 17-26, 2020.

[22] K. Kumar, M. Pradeepa, M. Mahdal, S. Verma, M. V. L. N. RajaRao, and J. V. N. Ramesh. "A deep learning approach for kidney disease recognition and prediction through image processing" Applied Sciences 13, no. 6, 3621, 2023.

[23] D. Saif, A. M. Sarhan, and N. M. Elshennawy. "Early prediction of chronic kidney disease based on ensemble of deep learning models and optimizers" *Journal of electrical systems and information technology 11*, no. 1, 17, 2024.

[24] M. S. Rana, S. M. Rahman, and A. H. Sung, "Evaluation of tree based machine learning classifiers for android malware detection", *Springer International Publishing*, vol. 11056: 377-385, 2018. Doi: org/10.1007/978-3-319-98446-9_35

[25] F. G. Mohammed, N. N. Thamir, "Esophageal Cancer Detection Using Feed-Forward Neural Network", Webology, 19(1): 6121, 2022.

[26] Y. Liu, Y. Zhou, S. Wen, and C. Tang, "A Strategy on Selecting Performance Metrics for Classifier Evaluation", Int. J. Mob. Comput. Multimed. Commun., vol. 6, no. 4, pp. 20–35, 2014. DOI: 10.4018/IJMCMC.2014100102

[27] E.D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images", Tue. Vol.1, (18):21, 2020. doi: org/10.48550/arXiv.2003.11055

[28] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation", arXiv preprint arXiv: 2010.16061, 2020.