Survey on Automatic Kidney Lesion Detection using Deep Learning

Mrs. Jyothi M Patil[1], Shreyas M[2], Shwetha KL[3], Tejana Patel H B[4], Thejaswini L[5]

[1]Assistant Professor, Department of CSE, ATME College of Engineering Mysore, Karnataka, India
[2-5]Student, Department of CSE, ATME College of Engineering Mysore, Karnataka, India

---

Abstract - *Kidney lesion detection is a crucial step in the diagnosis and management of kidney disorders. Deep learning methods have showed potential in enhancing kidney lesion detection's precision and effectiveness. Using medical imaging modalities such ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), and X-rays, we discuss recent research that have investigated the application of deep learning algorithms for kidney lesion diagnosis. We examine the effectiveness of several deep learning architectures and algorithms as well as the difficulties and restrictions associated with using deep learning to the identification of kidney lesions. Our study demonstrates that kidney lesion diagnosis using a variety of imaging modalities can be done with great accuracy and efficiency using deep learning algorithms. However, there are significant restrictions that must be overcome, including the necessity for sizable labelled datasets and the possibility of bias. We also highlight future objectives for deep learning research in kidney lesion identification, such as the creation of comprehensible deep learning models and the incorporation of deep learning with additional clinical data. Overall, this study emphasizes the potential of deep learning algorithms in enhancing kidney lesion identification and its contribution to improving the diagnosis and treatment of renal illness.*

Key Words: Kidney Lesion Detection, Deep Learning, Medical Imaging, Magnetic Resonance Imaging, Computed Tomography, Convolutional Neural Network, Segmentation

1. INTRODUCTION

The kidneys are important organs in the human body that filter waste and extra fluid from the blood in order to maintain the body's equilibrium. Abnormal growths or lumps on the kidneys are referred to as kidney lesions, also known as renal lesions. Early discovery is essential for successful treatment of these lesions, which can either be benign or malignant.

A medical procedure known as renal lesion identification includes locating and assessing the existence and severity of kidney lesions. The detection process commonly makes use of medical imaging techniques including ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), and X-rays. With these methods, medical professionals may see the kidneys and spot any anomalies like tumours or cysts.

To prevent benign kidney lesions from developing into malignant ones, early detection and treatment of kidney lesions are essential. Also, early discovery enables medical professionals to choose the best course of action, which, depending on the nature and severity of the lesion, may entail surgery, chemotherapy, or radiation therapy.

Kidney lesions can take many different forms, such as cysts, tumours, and abscesses. Cysts are fluid-filled sacs that can form on the kidney's outside or inside of it. They are often benign and don't need to be treated, but if they are too big or start to hurt, they could need to be surgically removed or drained.

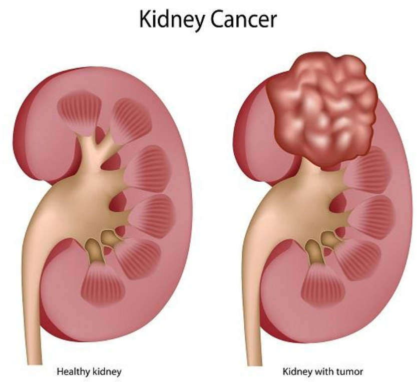

Figure - 1: Healthy Kidney VS Kidney with tumor

The identification of kidney lesions is a critical step in the diagnosis and management of renal illnesses. Using medical imaging methods including ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), and X-rays, the detection phase entails locating and assessing the existence and severity of kidney lesions. The manual analysis of medical photographs, however, can be time-consuming and prone to mistake. In order to increase the precision and effectiveness of kidney lesion diagnosis, researchers have investigated the use of deep learning algorithms.

There are various benefits of using deep learning algorithms to find kidney lesions. Secondly, deep learning algorithms can rapidly and reliably process massive datasets of medical pictures, which can save time and lower the possibility of human mistake. Second, deep learning algorithms can gain accuracy and generalisation skills by learning from big datasets of medical pictures. Lastly, the sensitivity and specificity of kidney lesion identification may be increased by training deep learning algorithms to recognise subtle signs and patterns that may be challenging for humans to notice.

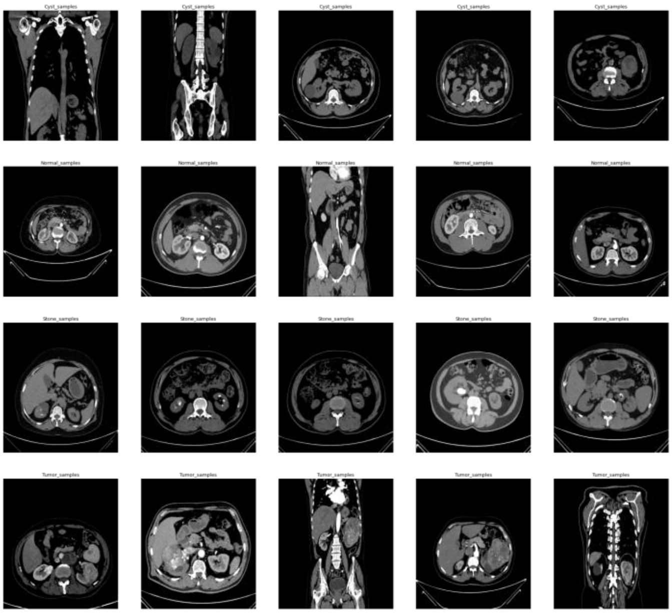

Figure - 2: Kidney tumor, cyst, normal or stone findings.

2. LITERATURE SURVEY

Hadjiyski [1] suggest that the automatic identification of kidney cancer will benefit from the building of a classifier that can differentiate between normal or benign CT pictures of the kidney and the cancer images. Investigation will be done on how varied cropped picture scale affects DLNN accuracy. The goal of this project is to develop a deep learning neural network-based system that will accurately estimate the stage of kidney cancer. The TensorFlow framework was utilised in this research together with the Inception V3 deep learning network structure. The DLNN-based artificial intelligence platform successfully distinguished kidney cancer Stage 1 from Stages 2, 3, and 4. 227 individuals from the Cancer Imaging with various stages of kidney cancer participated in the research. For the training and validation sets, respectively, the attained accuracy was 92% and 86%. The system has the potential to help doctors stage kidney cancer more precisely.

N Saranya's team [2] at Sri Eshwar College of Engineering's Subhanki B Department of Computer Science and Engineering has demonstrated how crucial machine learning algorithms are in the diagnosis of chronic kidney disease (2021). They postulate that using this approach to the practical diagnosis of CKD would have an alluring effect. Initially, there are frequently no symptoms; subsequently, side effects might include leg growth, fatigue, regurgitation, lack of appetite, and disarray. Hypertension, Diabetes, polycystic kidney disease and glomerulonephritis are causes of chronic kidney disease. Family history of chronic renal disease is one of the risk factors. The null values were then filled by running the proposed model a predetermined number of times. The finished product demonstrates how effective the suggested model is. By using random forest imputation, LR, DT might surpass the existing model.

The main manifestations of chronic kidney illness will be examined in this article [3], as well as how weka and machine learning approaches may be used to identify it. The risk of cardiovascular disease and end-stage renal disease is increased by chronic kidney disease. When chronic renal disease reaches an advanced level, the body may begin to accumulate electrolytes and waste. Multilayer perceptron is a general word for any feed-forward ANN that is used ambiguously and often. The number of persons who suffer from chronic renal disease is enormous. The number of those who are afflicted by the sickness is growing daily. It will make chronic kidney disease more predictable. Machines are capable of both illness detection and disease prediction. The accuracy, ROC, precision, recall, and f measure have been determined in this study using a variety of machine learning classifiers. Nevertheless, random forest has the best ROC value and 99% accuracy.

Researchers have published an article [4] on the classification of chronic kidney disease (CKD) using a number of algorithms. A dangerous condition that affects people all over the world, chronic kidney disease (CKD) is a major factor in adverse health consequences. Millions of individuals die each year as a result of poor treatment. The team was able to evaluate how well different ML algorithms worked. Logistic regression offers the highest accuracy and recall, whereas decision trees have the best precision. If detected early and appropriately, CKD can help patients in a variety of ways. It lengthens the patient's life and raises the likelihood of a successful therapy. 400 participants participated in the study. The accuracy of the suggested approach, which consists of decision trees, random forests, and logistic regression, is 98.48, 94.16, and 99.24, respectively. Recall of 97.61, 96.29, and 100, and precision of 100, 95.12, and 98.82. Using the advantages of each feature selection approach, two feature selection strategies are merged. In a comparison, they found that logistic regression had the best accuracy and recall, whereas decision trees have the most precision.

A group of researchers [5] has created a machine learning technique to foretell renal illness in patients and those who care for them. A life-threatening condition known as chronic kidney disease (CKD) affects around 14% of people worldwide. Even though the whole CKD domain has been sufficiently covered by the data distribution, general traits like hunger, anemia, and pedal oedema are biased in favor of CKD. A number of incredibly intricate excretion and re-absorption mechanisms result in the production of urine.

This mechanism is necessary for the body's chemical makeup to remain stable. The proposed technique consists of three key steps: model training, model selection, and data preparation. This work proposes a system that incorporates data prepossessing, a mechanism for handling missing values, collaborative filtering, and attribute selection in order to predict CKD status using clinical data. Out of the 11 machine learning algorithms considered, it is shown that the extra tree classifier and random forest classifier generate the greatest accuracy and least degree of bias to the attributes. The study shows the significance of using domain expertise when utilizing machine learning for CKD status prediction as well as the practical elements of data collecting.

Pedro Moreno-Sanchez (2021) [6] remarked on the significance of characteristics in enhancing the readability of early diagnosis of chronic renal disease. The incidence, prevalence, and high financial burden on health systems of chronic kidney disease (CKD) make it a global public health issue. The all-age mortality rate rose to 41.5% in 2017 because to the 1.2 million deaths caused by CKD since 1990. The primary goal of treating CKD is to slow the course of kidney impairment, generally by addressing the underlying causes. The creation of the classifier model has utilised the CRISP-DM technique. The scikit-learn package GridSearchCV was used to train and validate the classifier using 5-fold cross-validation. The Apollo Hospitals in Karaikudi, India provided the dataset used in this study with a total of 400 samples over the course of about two months in 2015. Among the 400 samples, 250 come from the group with CKD, and 150 come from the non-CKD group. As stated in the authors' conclusions, GridSearchCV "obtained results of 100% accuracy, precision, sensivity, specificity, and f1-score." The research conducted by Van Eyck et al. produced the most accurate results to date when compared to the findings from previous relevant investigations, according to Moreno-Sanchez.

A group of researchers [7] has devised an intelligent system that can predict kidney-related disorders with 98.5% accuracy. Deep belief networks were used by a team from the Department of Information Systems, led by Shahinda Elkholy (2021), to study the early detection of chronic kidney disease. Early kidney disease detection protects the patient from life-threatening consequences. 400 patients were included in the analysis. The variables that cause renal illnesses must be properly examined in order to forecast them. While Naive Bayes took less time, it produced findings that were more accurate than those of an Artificial Neural Network. The Categorical Cross-entropy loss function and Deep Belief Network with SoftMax classifier are used to build the model. To address the missing values, they utilised a dataset from the machine learning database at UCI. The effectiveness of the suggested model is assessed and contrasted with the models already in use. In comparison to the current models, the suggested model performs better and has an accuracy of 98.52%.

This study's [8] goal was to use DLM to quantify the connection between CKD and air pollution. Hence, the datasets may be easily separated into two categories: (1) air pollution statistics, and (2) records of patient health education or inspections. This is an introduction to the two datasets utilised in this study. This issue is solved by the deep learning framework created in this study. We included the time periods of CKD patient data and air pollution data into the framework and retrieved the time-series characteristics from the air pollution data. In order to accurately categorise the CKD patient stage, they then extracted the temporal feature information of these characteristics using an LSTM model. Finally, in our experiments, they used real CKD data and air pollution data from Taiwan to evaluate the effectiveness of the recommended technique in predicting the CKD stage of patients.

According to this study [9], the application of artificial intelligence to electronic health data can provide doctors with knowledge that will help them make better informed decisions about prognoses or therapies. In this study, machine learning was used to examine the medical data of patients with CKD and CVD. At the beginning, we predicted whether patients would develop severe CKD, both with and without taking the year it would happen into account or the date of the most recent visit. Our methods generated a mean Matthews correlation coefficient (MCC) of +0.499 in the first case and a mean MCC of +0.469 in the second. Age, eGFR, and creatinine are the most important clinical variables when the temporal component is absent; hypertension, smoking, and diabetes are clinical variables when it is present. We then performed a feature ranking analysis to identify the most important clinical characteristics. They compared our findings to those reported in the most recent scientific literature, discussing the variations in results when the temporal characteristic is added or eliminated.

This research study examines the consequences [10] of utilizing clinical factors and the support vector machines algorithm to classify individuals with chronic renal illness. The chronic kidney disease dataset is built on the clinical history, physical examinations, and laboratory tests. Using the three performance parameters of accuracy, sensitivity, and specificity, renal illness patients may be classified with a success rate of over 93%, according to experimental results. The tests they utilised in this study were inexpensive, straightforward, and non-invasive since we employed machine learning approaches. The information is supplied further. The data was taken from a dataset that was obtained from the UCI machine learning repository for CDK patients. By employing this tactic, scientists seek to "down-stage" (increase the percentage of CDK discovered at an early stage) the illness and get it to a stage when curative treatment is more likely to be effective.

Sl No.	Title	Author	Methodology	Limitation
1	Kidney Cancer Staging: Deep Learning Neural Network Based Approach	Nathan Hadjiyski Year: 2020	Inception V3 deep learning network structure was used in this project within the TensorFlow platform.	The DLNN model isn't perfect though and sometimes it can incorrectly classify the kidney cancer stage.
2	Diagnosing Chronic Kidney Disease using KNN Algorithm	Saranya N, Sakthi Samyuktha M, Sharon Isaac, Subhanki B Year: 2021	We analyzed and classified KNN and Logistic Regression algorithms with Chronic Kidney Disease dataset.	In any case, during the time spent setting up the model, because of the restrictions of the conditions.
3	Performance Analysis of Chronic Kidney Disease through Machine Learning Approaches	Minhaz Uddin Emon, Al Mahmud Imran, Rakibul Islam, Maria Sultana Keya, Raihana Zannat, Ohidujjaman Year: 2021	The proposed model has 8 ML classifiers are used namely: LR, NB, MLP, SGD, Adaboost, Bagging, DT, RF classifier are used.	A forest is less interpretable than a single decision tree. Single trees may be visualized as a sequence of decisions.
4	Performance Analysis of Machine Learning Classifier for Predicting Chronic Kidney Disease	Rahul Gupta, Nidhi Koli, Niharika Mahor, N Tejashri Year: 2021	The classification techniques, that is, tree-based decision tree, random forest, and logistic regression have been analyzed.	LR assumes linearity between the predicted (dependent) variable and the predictor (independent) variables.
5	Chronic Kidney Disease Prediction Using Machine Learning Methods	Imesh Udara Ekanayake, Damayanthi Herath Year: 2020	In this work, 11 classification models were used.	Doesn't have a large amount of dataset.
6	Features Importance to Improve Interpretability of Chronic Kidney Disease Early Diagnosis	Pedro A. Moreno-Sanchez Year: 2020	AdaBoost is selected as the best classifier with a 100% in terms of accuracy, precision, sensitivity, specificity, and f1-score;	Noisy data and outliers have to be avoided before adopting an Adaboost algorithm.
7	Early Prediction of Chronic Kidney Disease Using Deep Belief Network	Shahinda Mohamed Mostafa Elkholy, Amira Rezk, Ahmed Abo El Fetoh Saleh Year: 2021	Uses modified Deep Belief Network (DBN) as classification algorithm.	A lot of training data is needed for the model to be effective and that they fail to encode the position and orientation of objects.
8	Deep learning for Etiology of Chronic Kidney Disease in Taiwan	Sheng-Min Chiu, Feng-Jung Yang, Yi-Chung Chen, Chiang Lee Year: 2020	The DLMs used in this study were long and short-term memory (LSTM) models.	DLMs take longer to train and also requires more memory.

9	A Machine Learning Analysis of Health Records of Patients With Chronic Kidney Disease at Risk of Cardiovascular Disease	Davide Chicco, Christopher A. Lovejoy, Luca Oneto Year: 2021	In this work, classification models were RF, Gaussian SVM, Neural Network, Linear SVM, Decision Tree and XGBoost.	They performed the analysis only on a single dataset.
10	Analysis of Chronic Kidney Disease Dataset by Applying Machine Learning Methods	Yedilkhan Amirgaliyev, Shahriar Shamiluulu, Azamat Serek Year: 2019	In this work, the classification model is SVM.	Doesn't have a large amount of dataset and the accuracy is low.

Table - 1: Relevant studies on advantages and disadvantages of approaches to detect kidney disease

3. CONCLUSIONS

In conclusion, using deep learning algorithms to the diagnosis of kidney lesions has demonstrated significant promise for increasing the precision and effectiveness of the detection procedure. Recent studies that were analyzed in this study looked at the use of deep learning algorithms on a variety of imaging modalities, including ultrasound, computed tomography (CT), magnetic resonance imaging (MRI), and X-rays. The findings demonstrated that kidney lesion identification using deep learning algorithms was very accurate and effective, with promising results for tumor and lesion segmentation.

Nevertheless, using deep learning methods has drawbacks as well, such as the need for sizable labelled datasets, the possibility of bias, and the models' lack of explainability. To make sure that deep learning algorithms are dependable and efficient in clinical practice, these restrictions must be overcome.

Future research will focus on the creation of explainable deep learning models that can shed light on the algorithms' decision-making process as well as the integration of deep learning with additional clinical data to increase the precision and efficacy of kidney disease diagnosis and treatment.

Overall, this study emphasizes the potential of deep learning algorithms to enhance kidney lesion identification and their contribution to the advancement of the diagnosis and treatment of renal illness. Deep learning is an area that is still evolving, so more study and development are required before it can completely fulfil its promise for increasing kidney lesion identification and optimizing patient care.

REFERENCES

[1] Hadjiyski, Nathan. "Kidney cancer staging: Deep learning neural network based approach." In *2020 International Conference on e-Health and Bioengineering (EHB)*, pp. 1-4. IEEE, 2020.M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] Saranya, N., M. Sakthi Samyuktha, Sharon Isaac, and B. Subhanki. "Diagnosing chronic kidney disease using KNN algorithm." In *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, pp. 2038-2041. IEEE, 2021.K. Elissa, "Title of paper if known," unpublished.

[3] Emon, Minhaz Uddin, Rakibul Islam, Maria Sultana Keya, and Raihana Zannat. "Performance Analysis of Chronic Kidney Disease through Machine Learning Approaches." In *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, pp. 713-719. IEEE, 2021.

[4] Gupta, Rahul, Nidhi Koli, Niharika Mahor, and N. Tejashri. "Performance analysis of machine learning classifier for predicting chronic kidney disease." In *2020 International Conference for Emerging Technology (INCET)*, pp. 1-4. IEEE, 2020.

[5] Ekanayake, Imesh Udara, and Damayanthi Herath. "Chronic kidney disease prediction using machine learning methods." In *2020 Moratuwa Engineering Research Conference (MERCon)*, pp. 260-265. IEEE, 2020.

[6] Moreno-Sanchez, Pedro A. "Features importance to improve interpretability of chronic kidney disease early diagnosis." In *2020 IEEE International Conference on Big Data (Big Data)*, pp. 3786-3792. IEEE, 2020.

[7] Elkholy, Shahinda Mohamed Mostafa, Amira Rezk, and Ahmed Abo El Fetoh Saleh. "Early prediction of chronic kidney disease using deep belief network." *IEEE Access* 9 (2021): 135542-135549.

[8] Chiu, Sheng-Min, Feng-Jung Yang, Yi-Chung Chen, and Chiang Lee. "Deep learning for Etiology of Chronic Kidney Disease in Taiwan." In *2020 IEEE Eurasia*

Conference on IOT, Communication and Engineering (ECICE), pp. 322-325. IEEE, 2020.

[9] Chicco, Davide, Christopher A. Lovejoy, and Luca Oneto. "A machine learning analysis of health records of patients with chronic kidney disease at risk of cardiovascular disease." *IEEE Access* 9 (2021): 165132-165144.

[10] Amirgaliyev, Yedilkhan, Shahriar Shamiluulu, and Azamat Serek. "Analysis of chronic kidney disease dataset by applying machine learning methods." In *2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT)*, pp. 1-4. IEEE, 2018.