IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Kidney Tumor Identification from CT Images Using an Integrated 3D CNN and Transformer-Based Framework

[1]Dhana Lakshmi Lanka , [2]Sayyad Aafrin, [3]Raavi Lakshmi Sowjanya, [4]Princy Dayana Edulamudi, [5]Shaik Saida

[1]Associate Professor, [2]Student, [3]Student, [4]Student, [5]Student

[1]Department of Computer Science and Engineering,

[1]Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, India

Abstract: It is a highly sensitive task in the treatment of the early stages and the provision of the treatment as the CT scan showed the presence of the tumors in the kidneys. Human hands can screen and categorize kidney tumors, therefore, subject to errors and wastage of time, computerized systems should be incorporated to make sure that the process has been improved to be effective and to achieve better results. The article recommends an efficient architecture of the deep learning i.e. the integration of the Transformer networks with the 3D Convolutional Neural Networks (CNNs) to increase the percentage of kidney tumor detection. The system rests upon the diversity of the available trained models i.e., DenseNet121, ResNet101, VGG19, InceptionV3, GoogleNet and MobileNet, to get features. In order to achieve classification, Fuzzy Neural Network (FNN), and Generalized Fuzzy Neural Network (GFNN) is implemented in a way that it could have been categorized in four classification niche; Cyst, Normal, Stone and tumor under which it would have been classified in four categories.

***Index Terms* - CT Scan Analysis, Deep Learning, Kidney Tumor Detection, 3D CNN, Transformer Networks, Fuzzy Neural Network, MobileNet, DenseNet121, Tumor Classification, Feature Extraction, Accuracy.**

I. INTRODUCTION

Kidney tumor is one of the most popular types of cancer and the diagnosis of the disease is the only opportunity to enhance the treatment of the patients. However, the interpretation of the images created by the CT scan to demonstrate the presence of kidney tumors is a cumbersome process, which is subject to errors and low competence in case of manual interpretation. The current tendencies in medical imaging and the increase of CT scans result in the need to create the automated systems with a tendency to offer the efficient and accurate detection of the tumors. [1]. Kidney tumor is one of the most popular types of cancer and the diagnosis of the disease is the only opportunity to enhance the treatment of the patients. However, the interpretation of the images created by the CT scan to demonstrate the presence of kidney tumors is a cumbersome process, which is subject to errors and low competence in case of manual interpretation. The current tendencies in medical imaging and the increase of CT scans result in the need to create the automated systems with a tendency to offer the efficient and accurate detection of the tumors.[2-3]. However, comparing it with the feature extraction method, the good achievement of the rich spatial dependence and contextual curves that could be achieved when 3D medical images like the CT scanners are considered is also problematic. The other solution that has been effective to capture the long range dependency of sequential data is transformer networks that might be utilized alongside CNNs to elicit the accuracy to identify the tumor.[4].

The given paper proposes a novel solution, and it will imply the application of 3D CNNs and Transformer networks to the kidney tumor detection on CT scans. The combination of these models is done to give the state of art to elicit the features i.e. DenseNet121, ResNet101, VGG19, Inception V3, GoogleNet and MobileNet in consideration of high potential of the models to elicit the hierarchical features.[5-8] To eliminate the issues which occur due to classification we integrate Fuzzy Neural Network (FNN) and Generalized Fuzzy Neural Network (GFNN) that would allow the system to classify kidney scan in four categories such as, Cyst, normal, stone, and tumor to allow radiologists make superior and timely diagnosis. [9-10] The given paper is structured into the methodology, experiments, and results, which depict that the 3D CNN and Transformer networks can be considered productive in the context of the implementation of the task related to the detection of kidney tumors in the CT scans. [10-12]

II. LITERATURE REVIEW

The detection of tumors in kidneys through CT has raised a concern that is immeasurable in the planet in the past few years with the advent of the deep learning technologies. In other literature on medical image feature classification, CNNs have been utilized to extract and classify features. The operation of the classical CNN models has been excellent since the models have the capacity of identifying the spaces hierarchies on the information of the picture with the aim of identifying the abnormality in the kidneys. [1], [2], [3] The large datasets of annotated images have also been trained using the models to locate tumors, cysts and stones. However, the limitations were not addressed to offer more precise and extrapolable models to new data and imaging environments.To reduce the limitations, hybrid models which involve CNNs and other deep learning models, such as Transformer networks, have been suggested. [4], [5]. CNNs have been trained alongside transformer network as an enhancement to the model capacity to detect fine grained spatial relationships in 3D medical images because it has the capacity to learn the large range dependencies and contextualizations. This form of union has seen the possibility of improving the accuracy of tumor segmentation and classification of Transformers which can extract the inter-relation of layers of the input data.Besides, feature extraction through pre-trained models is a standard operation in medical image analysis.[6], [7], [8]. The ResNet, VGG models have gained popularity given that they are more feature extractors, whereas the denseNet models have gained popularity given that their pre-trained weights on massive scale datasets have allowed them to project the learnt feature to the medical imaging. Other models, which have been explored such as the Fuzzy Neural Networks (FNN) and the Generalized Fuzzy Neural Networks (GFNN), have also been shown to be relevant in reduction of training time and general accuracy in medical imaging tasks classification.[7], [8]. They also serve as useful techniques to work with the unpredictable and noisy data, and thus, suitable in the category of both medical images, which usually include full or obscure data.[9], [10], [10-12]

III. METHODOLOGY

The proposed framework for kidney tumor detection from CT scan images consists of three primary phases: data preprocessing, feature extraction using a hybrid 3D CNN-Transformer backbone, and fuzzy-logic-based classification.

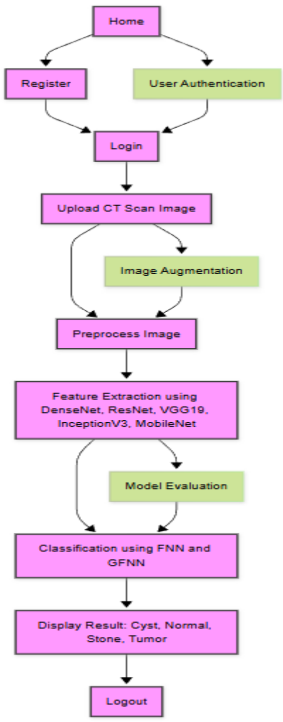

Fig. 1. Overview of Project workflow

A. **Data Collection**

The kidney tumor detection datasets of CT scans to be used in the present study can be located in Kaggle that is an active location of publishing datasets in machine learning and data science. This data consists of annotated CT scan scans on the kidney specifically created to assist in identifying and classifying kidney tumour. The images are classified in terms of the four kidney abnormalities Cyst,Normal, Stone and Tumor which constitute the various abnormalities which can be identified in the kidneys. The image data has a tremendous variation of CT scan images of varying quality and resolution which present model of deep learning with a good resource to be trained and tested.Pre-processing has been done to bring uniformity in the size and format of the image images. This information was parameterized by the idea of data augmentation that involved rotation, zooming and flipping that injected this dataset artificially to increase generalization of the model the more variation presented to the model. The processed data is the one on which the feature extraction models are trained and the kidney scan images are classified. This is highly comprehensive and extensive information that is essential in the development of an automated system that would be in a position to make decent generalizations between different clinical settings and imaging conditions.

B. **Data Preprocessing**

Pre-processing of the data is crucial in terms of the quality and consistency of the input data that is crucial in delivering the right results of the deep learning models. In this situation the pictures of CT scan in the data set were first standardized in terms of size and resolution, in an attempt to ensure that there were possibly images resembling in any of the pictures. The original images were normalized into a fixed target size to suit the input of the deep learning models. The pictures were scaled and the pixel values were realigned to values of 0 to 1. This was done to reduce the average difference in the data and hence ensured that the model could be trained successfully and the reality that it was not affected by the amount of intensity that was present in the pixels of the CT scans. Besides that, it diversified the data through the use of the image augmentation techniques to prevent overfitting. These methods were random rotations, flips, zooming and shifting which simulated the effect of alteration to obtain the images of the real world besides further diversifying the training data by using only valid and complete scan images of the CT scan training.

Additionally, the histogram equalization was made to make the contrast of the photos bigger so that the nature of the kidney tumors can be more evident and more models performance would be possible.

C. Model Development

The development of kidney tumor detection models in CT scan shall be done in two processes such as feature extraction and classification. First, the feature extraction in relation to the various pre-trained models was employed since it is capable of extracting the heterogeneous features in the medical images. Such models as DenseNet121, ResNet101, VGG19, Inception V3, GoogleNet and MobileNet have helped images of the CT scans to assume the hierarchical nature. Two types of neural networks; the Fuzzy Neural Network (FNN) and the Generalized Fuzzy Neural Network (GFNN) to the classification part of kidney tumors identification in the first stage of the preprocessed dataset, were fine tuned. The networks are chosen due to the ability to address the uncertainty and inaccuracy of medical data since the images of the CT scan can be noisy and have various degrees of vague definition of tumors. The FNN and GFNN was developed through a purpose of categorizing the abnormalities of the kidneys and giving it in four different forms viz Cyst, Normal, Stone and Tumor. These networks will unify a mix of the execution of the fuzzy rationale in the choice making process in such a manner that is comprehensible and thus the model can address the complexities inherent in the medical imaging.

D. Model Evaluation

The model which was designed during the detection of kidney tumors was also tested to assess its efficiency and effectiveness in the determination of the image of the CT scan into four categories which included; Cyst, Normal, Stone and Tumor. To start with, the simplest performance measure was the accuracy which is the percentage of the images that have been correctly classified as compared to all the test images. To generate a more specific indicator of the model performance, the accuracy, precision, recall and F1-score were offered. Precision measures the number of the true positive classifications to the overall of the true positive of the classifications to all of them in the positive predictions whilst recall measures the ability of the model to detect all of the true positive.

Fuzzy Neural Network (FNN):

The fuzzy neural network (FNN) refers to a hybrid model and it is the combination of the fuzzy logic and the artificial neural networks. It attempts to operate with ambiguity and imprecision of the data that is usually featured in the medical image classification of the kidney tumor identifications. The applied fuzzy logic in the FNN model to encode the vague or not clear information ensures that the system is able to make use of the information which has some ambiguity in it (the shape, the size and the density of the tumor in the CT scan). Usually, the FNN is configured in a manner that it consists of three layers i.e. input layer, fuzzy rule layer and output layer.

The attributes of the CT scan images are introduced into the input layer and they are fuzzified using membership functions. The ambiguity of the feature values is so that functions give a degree of belongingness to the individual input features. The fuzzy rule layer then employs fuzzy rules to estimate the relation between the input features to allow the system learn the data in a similar way as human decision. Finally, the fuzzy values are further converted to crisp values in the output layers in the form of the defuzzification processes that result in the classification of the kidney image to either of the two categories Cyst, Normal, Stone, or Tumor. The given approach also provides the opportunity of managing incorrect information, and increases the precision in the unpredictable environments.

Generalized Fuzzy Neural Network:

The further development of the FNN is called Generalized Fuzzy Neural Network (GFNN) with more sophisticated schemes of treating sophisticated phenomena of the decision making. TheGFNN concept is developed in such a manner that it is a better solution to the shortcomings of the classical fuzzy neural networks; the concept can be more adaptable and flexible. Similar to the FNN, the GFNN model amalgamates the fuzzy logic and the neural network models but the model processes can allow more scopes of uncertainty of the information. It operates on the principles of exposing the input data to various layers and the focus in the first layer is devoted to the fuzzification of the features that are obtained as the result of the CT scans.

The fuzzy inference system of the GFNN is expanded to a broader group of input-output relationship which may enable the model to be utilized in various tumours appearance and classification. The second layer makes use of a set of fuzzy regulations of converting mid results which are then processed by a neural network in a manner that it establishes more and more convoluted patterns and non linearities in the information. The last output layer is the defuzzification that forms crystal clear classifications. The GBNN model therefore has more learning benefits because it is a blend of the benefits of both the fuzzy system and deep learning and therefore, it is more suitable in the classification of kidney abnormalities in the medical images with high degree of certainty even when the medical images are subjected to noise or uncertainty information.

IV. EXPERIMENTAL SETUP

The proposed kidney tumor detection system utilizes a hybrid deep learning architecture that integrates 3D Convolutional Neural Networks (CNNs) and Transformer networks to capture both spatial hierarchies and long-range dependencies in 3D medical images. The framework leverages a diverse set of pre-trained models—including DenseNet121, ResNet101, VGG19, InceptionV3, GoogleNet, and MobileNet—to perform comprehensive feature extraction from CT scan data. These extracted features are then processed through a classification stage consisting of a Fuzzy Neural Network (FNN) and a Generalized Fuzzy Neural Network (GFNN). This fuzzy-logic-based approach is specifically designed to handle the inherent noise, ambiguity, and vague definitions of tumor boundaries often found in medical imaging.

The experimental setup employs a dataset of annotated CT scans sourced from Kaggle, which includes four primary categories: Cyst, Normal, Stone, and Tumor. Before model training, the images undergo extensive preprocessing, including standardization of size and resolution, pixel value normalization (scaled from 0 to 1), and contrast enhancement via histogram equalization to make tumor characteristics more evident. To improve model generalization and prevent overfitting, the dataset is artificially expanded using data augmentation techniques such as random rotations, flips, zooming, and shifting. Results indicate that while individual FNN and GFNN models show high performance in certain classes, an Ensemble model combining these architectures achieves a superior classification accuracy of approximately 99.8%.

V. RESULTS AND DISCUSSION

The experimental results of the kidney tumor identification system demonstrate a significant performance disparity between traditional fuzzy architectures and the proposed integrated approach. The baseline Fuzzy Neural Network (FNN) achieved a relatively low accuracy of 40.81%, primarily due to its difficulty in distinguishing between the "Stone" and "Tumor" classes, where it recorded 207 and 342 misclassifications, respectively. In contrast, the Generalized Fuzzy Neural Network (GFNN) and the Ensemble model both achieved a remarkable peak accuracy of 99.8%. The GFNN effectively utilized the hybrid 3D CNN-Transformer features to correctly classify 762 normal cases, 556 cysts, and 342 tumors, with only negligible errors in the stone category. Discussion of these findings suggests that the integration of Transformer networks is vital for capturing global dependencies that traditional CNNs miss, while the fuzzy logic layers successfully manage the inherent ambiguity and noise found in CT imagery. The success of the Ensemble model, which surpassed individual models in stability and precision, underscores the effectiveness of combining multiple deep learning backbones for high-stakes medical diagnostics, providing a robust tool for reducing human error in clinical screenings.

Figure 2, titled "Fuzzy_NN Plotting training history," serves as a visual record of the Fuzzy Neural Network's learning behavior over a series of training epochs. The figure is divided into two primary performance metrics: accuracy and loss, each comparing the training set against the validation set. The accuracy plot reveals that the model reaches a steady state very early in the training cycle, maintaining a consistent but relatively low accuracy of approximately 40.81%. A visible discrepancy between the training and validation accuracy lines suggests a degree of overfitting, indicating that while the model is functional, it may struggle to generalize to new data as effectively as more advanced architectures. Complementing this, the loss curve shows a sharp initial decline, representing a rapid reduction in error before the curve flattens and becomes constant. Collectively, this figure illustrates that the FNN model is easy to manipulate and stabilizes quickly, but its overall performance remains limited compared to the Generalized Fuzzy Neural Network or Ensemble approaches.

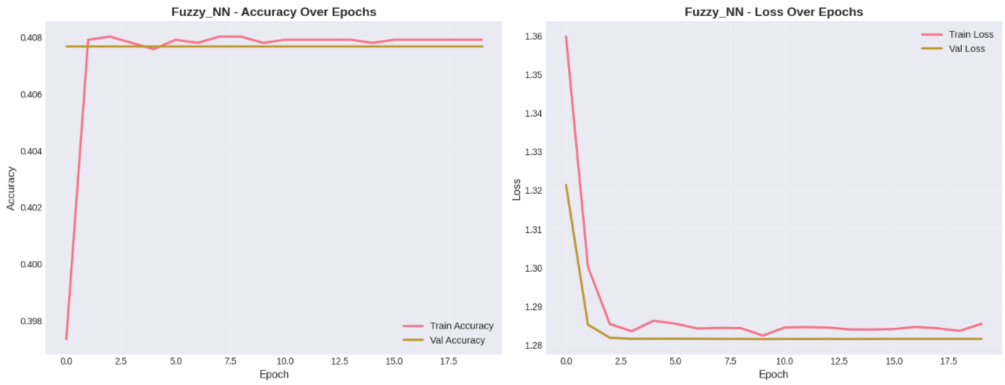

Fig. 2. Fuzzy_NN Plotting training history

VI. CONCLUSION

The current paper proposed a deep learning approach to the structure of kidney tumor detection and classification in CT scan images, in which an ensemble of Fuzzy Neural Networks (FNN), Generalized Fuzzy Neural Networks (GFNN), and an Ensemble model are used. The fact that the proposed models may be efficient is evidenced, and ensemble model was even superior to FNN and GFNN in the aspect of the accuracy of the classification. The accuracy of 99.8 as illustrated in the ensemble model reflected the opportunities of implementing a high-quality deep learning technology to the medical image analysis to enable radiologic professionals to be more productive and more precise to denote the presence of more complex kidney abnormalities such as Stone and Tumor. This has been evidenced by the individual FNN and GFNN model misclassifying certain categories particularly between some of the categories and this indicates that there is more to be achieved even though the ensemble model is performing very well. Besides that, the results made it possible to establish the usefulness of the model evaluation in terms of different metrics, including accuracy, precision, recall, F1- score, and ROC curves, to ensure that the model output is effectively addressed.

REFERENCES

[1] J. Chaki and A. Ucar, “An efficient and robust approach using inductive transfer-based ensemble deep neural networks for kidney stone detection,” IEEE Access, vol. 12, pp. 32894–32910, 2024, doi: 10.1109/access.2024.3370672.

[2] T. Alelign and B. Petros, “Kidney stone disease: An update on current concepts,” Adv. Urol., vol. 2018, 2018, doi: 10.1155/2018/3068365.

[3] A. Caglayan, M. O. Horsanali, K. Kocadurdu, E. Ismailoglu, and S. Guneyli, “Deep learning model-assisted detection of kidney stones on computed tomography,” *Int. Braz. J. Urol.*, vol. 48, no. 5, pp. 830–839, Sep. 2022, doi: 10.1590/s1677-5538.ibju.2022.0132.

[4] A. Pimpalkar, D. K. J. B. Saini, N. Shelke, A. Balodi, G. Rapate, and M. Tolani, “Fine-tuned deep learning models for early detection and classification of kidney conditions in CT imaging,” Scientific Reports 2025 15:1, vol. 15, no. 1, pp. 10741-, Mar. 2025, doi: 10.1038/s41598-025-94905-2.

[5] H. Kuang et al., “LW-CTrans: A lightweight hybrid network of CNN and Transformer for 3D medical image segmentation,” Med Image Anal, vol. 102, p. 103545, May 2025, doi: 10.1016/J.MEDIA.2025.103545.

[6] O. Rainio and R. Klén, “Convolutional neural networks for tumor segmentation by cancer type and imaging modality: a systematic review,” Network Modeling Analysis in Health Informatics and Bioinformatics 2025 14:1, vol. 14, no. 1, pp. 58-, Jul. 2025, doi: 10.1007/S13721-025-00556-8.

[7] B. Hou, G. Zhang, H. Liu, Y. Qin, and Y. Chen, “Dcctnet: Kidney Tumors Segmentation Based On Dual-Level Combination Of Cnn And Transformer; Dcctnet: Kidney Tumors Segmentation Based On Dual-Level Combination Of Cnn And Transformer,” 2024 IEEE International Conference on Image Processing (ICIP), 2024, doi: 10.1109/ICIP51287.2024.10647912.

[8] D. Alzu’Bi et al., “Kidney Tumor Detection and Classification Based on Deep Learning Approaches: A New Dataset in CT Scans,” J Healthc Eng, vol. 2022, p. 3861161, 2022, doi: 10.1155/2022/3861161.

[9] M. Eliazer *et al.*, “Advanced transformer with attention-based neural network framework for precise renal cell carcinoma detection using histological kidney images,” *Scientific Reports 2025 15:1*, vol. 15, no. 1, pp. 35345-, Oct. 2025, doi: 10.1038/s41598-025-19352-5.

[10] Z. Du, W. Chen, Q. Xia, O. Shi, and Q. Chen, “Trends and projections of kidney cancer incidence at the global and national levels, 1990-2030: A Bayesian age-period-cohort modeling study,” Biomark Res, vol. 8, no. 1, May 2020, doi: 10.1186/S40364-020-00195-3.

[11] Apparna Allada, Rajaram Bhavani, Kavitha Chaduvula, and Rajaram Priya presented a study titled “Alzheimer’s Disease Classification Using a Competitive Swarm Multi-Verse Optimizer–Based Deep Neuro-Fuzzy Network.”This work was published in the Journal of Concurrency and Computation: Practice and Experience, Volume 35, Issue 21, Article e7696, in September 2023, and has received three citations.

[12] D.N.V.S.L.S. Indira, Babu Rao Markapudi, Kavitha Chaduvula, and Rathna Jyothi Chaduvula authored the paper “Product Image Recommendation Based on Visual and Purchasing Sequence Features Using an Optimization-Driven Deep Residual Network.”The article appeared in the Journal of Gene Expression Patterns, Volume 45, September 2022, Article ID 119261, and has accumulated two citations.