https://doi.org/10.65091/icicset.v2i1.25

Kidney Tumor Detection in Medical Image Processing

Sugitha N
Dept. of Electronics and Communication Engineering
Saveetha Engineering College
Chennai, India
sugithan@saveetha.ac.in

Ragul S
Dept. of Electronics and Communication Engineering
Saveetha Engineering College
Chennai, India
ragul7690@gmail.com

Nithaesh S
Dept. of Electronics and Communication Engineering
Saveetha Engineering College
Chennai, India
nithaeshnishanth6886@gmail.com

***Abstract*—Accurate classification of medical images, particularly CT scans, plays a vital role in diagnosing and detecting renal anomalies, including tumors, stones, and cysts. This work investigates several cutting-edge deep learning techniques for the automated identification and categorization of certain renal conditions. Large datasets comprising thousands of CT images were used to evaluate different methods according to important performance metrics such as recall, accuracy, precision, and F1-score. The proposed approaches demonstrated high classification accuracy, with several models achieving over 99% in sensitivity and specificity. Additionally, The study emphasizes the significance of incorporating explainability techniques to improve model interpretability, Facilitating clinicians' comprehension of the decision-making process. The findings emphasize the potential of these advanced techniques in enhancing medical diagnostics, particularly in environments where computational efficiency and accuracy are critical.**

***Index Terms*—Medical image processing, Kidney tumor detection, Kidney abnormalities, Deep learning, classification algorithms.**

I. Introduction

Kidney tumors are a serious health issue that can be detected early being essential for improving patient outcomes. The objective of this research is to develop a sophisticated image processing system that uses Deep Learning techniques, namely Convolutional Neural Networks (CNN), to detect kidney abnormalities.The aim is to automate the diagnostic process, enhancing accuracy and efficiency while reducing reliance on manual image interpretation by radiologists. The system processes medical images from multiple imaging modalities, such as CT scans, MRI, and ultrasound, applying various preprocessing techniques, including normalization, contrast improvement, and noise reduction to maximize picture quality for analysis. A key component of the system is tumor segmentation, which precisely delineates kidney tumors from surrounding tissues, overcoming challenges like irregular shapes and varying sizes. By leveraging CNNs, the system identifies subtle patterns within the images, distinguishing between benign and malignant tumors. A variety of datasets are used to train the model to ensure high generalization across different patient cases, improving sensitivity and specificity in detecting tumors. The system also integrates clinical data such as patient history and genetic information, offering a more comprehensive diagnostic tool and opening the door for customized treatment. The classification of medical images, especially in radiology, has saw notable progress in the last few years as a result of deep learning model development. CNN, one of these models, have shown to be a very successful way to analyze medical pictures, such CT scans, with exceptional accuracy. These networks are able to automatically extracting hierarchical characteristics from pictures, which makes them invaluable in detecting various diseases and abnormalities. This article's primary focus is on the classification of renal CT images based on the VGG16 architecture in a deep learning model. The classification task involves distinguishing between four key categories of kidney conditions: normal, cyst, tumor, and stone.Its architecture consists of 16 layers and is renowned for being deep and simple, which makes it a great feature extractor for a variety of image categorization applications. Although it was originally trained using a massive dataset like as ImageNet, VGG16 has proven to be an effective base model for transfer learning in specialized activities include classifying medical images. In this study, the already trained VGG16 model is adapted to categorize kidney CT photos into four classifications. The model may use transfer learning to refine its knowledge from a large general-purpose dataset for the particular job of renal disease classification.Where the last few layers of VGG16 model are unfrozen to allow these layers to adapt to the kidney image dataset. Fine-tuning is crucial in transfer learning, because it makes it possible for the model to extract more precise characteristics from the medical pictures that were not present in the original dataset. To avoid large updates that could destabilize the pre-trained weights, a smaller learning rate is employed during this phase. Both the

Adam optimizer and sparse categorical cross-entropy, which are appropriate for multi-class classification issues, are used in the training phase as the loss function

II. Related Work

While recent developments in deep learning have significantly enhanced the categorization of kidney abnormalities in CT images, several notable gaps persist in the existing literature, particularly when analyzed against the backdrop performance of our model, the findings of relevant studies. Firstly, although our model achieved a remarkable validation accuracy of 94%, comparable studies, such as those by
Maqsood et al. (2024) and Bingol et al. (2023), indicate that hybrid models employing complex architectures often require substantial computational resources. These requirements can impede their deployment in clinical settings, especially in low resource environments where accessibility to high-performance computing is limited. Our model, built on the VGG16 architecture, provides a simpler yet effective alternative, suggesting that future research should explore optimizing existing models to achieve similar accuracies without excessive computational demands. Additionally, while our approach highlights the importance of precision and recall, many studies, including Asif et al. (2023), often focus solely on accuracy metrics without fully addressing the consequences of false negatives and false positives in medical diagnostics. In the context of kidney disease detection, a false negative may lead to missed diagnoses, while a false positive could subject patients to unnecessary anxiety and further invasive testing. Therefore, it is imperative for future research to incorporate a broader array of metrics, such as the region beneath the ROC curve (AUC-ROC), to offer a more sophisticated comprehension of model performance. Moreover, while our model's design emphasizes interpretability through precision and recall metrics, the integration of explainable AI techniques remains limited in the literature. Studies utilizing models like the Swin Transformer, as noted by Islam et al. (2022), have shown that explainability is crucial for gaining trust from clinicians. Future research should focus on developing frameworks that not only offer high classification accuracy but also enhance the interpretability of model predictions to facilitate clinical decision-making. Furthermore, while datasets like those employed in the research conducted by Alzu'bi et al. (2022) and Bhandari et al. (2023) demonstrate substantial representation of various kidney conditions, there is a lack of diverse demographic and clinical data. This limitation raises concerns regarding the generalizability of models to broader patient populations. Our study underscores the necessity for diverse datasets that account for variations in age, gender, and comorbidities to ensure that models can perform effectively across different demographic groups. Lastly, while advanced models such as the Vision Transformer (ViT) have shown promise, the complexity and training time associated with these architectures can present significant barriers to implementation in clinical environments. Future research should evaluate the trade-offs between model complexity and performance, aiming to identify efficient architectures that can deliver robust performance with lower training and inference times. By addressing these gaps, future studies can enhance the applicability and effectiveness in kidney disease detection, eventually enhancing medical imaging research and patient outcomes.

TABLE I: COMPARISIOIN OF MODEL PERFORMANCE IN RECENT RESEARCH

Reference	Model	Metrics
Maqsood, F., et al. (2024)	Hybrid SpinalZFNet	Accuracy: 99.8%
Bingol, H., et al. (2023)	Hybrid CNN	Accuracy: 99.37%
Asif, S., et al. (2023)	IR-CNN	Accuracy: 99.38%
Bhattacharjee, A., et al. (2023)	ViT	Accuracy: 93%
Bhandari, M., et al. (2023)	CNN	Accuracy: 99.52%
Alzu'bi, D., et al. (2022)	CNN-6	Accuracy: 97%
Islam, M. N., et al. (2022)	Swin Transformer	Accuracy: 99.30%

Article(2022-2024)

III. Problem statement

Kidney tumors are a significant health concern worldwide, with early and accurate detection being critical for effective treatment and improved patient outcomes. However, analyzing medical images for tumor detection requires specialized expertise and is time-consuming, which can delay diagnosis and treatment in resource-limited settings.This project aims to develop an efficient and accurate computer-aided detection system for kidney tumor identification using medical image processing techniques and deep learning models. By leveraging Convolutional Neural Networks (CNNs), the system will classify medical images into categories such as normal, cyst, tumor, and stone, aiding healthcare professionals in faster and more precise diagnosis. This solution seeks to enhance diagnostic accuracy, reduce human error, and support early intervention efforts in kidney tumor management.

IV. Data Preparation and Augmentation

In the kidney image classification project, effective data preparation is critical for ensuring that the model trains efficiently and accurately. Four types of images Normal, Stone, Tumor and Cyst are included in collection. These images are organized within a directory structure that facilitates easy loading using TensorFlow's function. This function automates the process of reading the images and creating TensorFlow datasets. To prepare the data for training, the images are

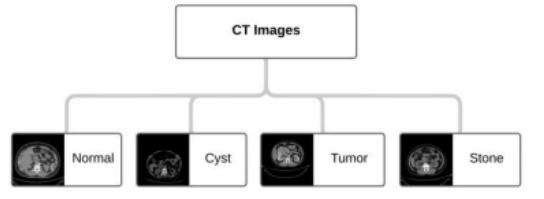

Fig. 1. Kidney CT Image Classification.

scaled to 150x150 pixels, which is their standard size. This standardization is essential, as deep learning models require input images of consistent size for processing. Furthermore,

10% for validation and 90% of the images are put aside for training, dividing a datasets in to validation and training subsets. During training, this stratified divide aids in tracking the model's performance, revealing information about its capacity to extrapolate to unknown facts. A random seed is specified (seed=123) to ensure reproducibility of the dataset splits. Furthermore, the model employs the pre-trained VGG16 architecture, which significantly enhances feature extraction from the images, resulting in improved classification accuracy. The inclusion of dropout layers helps prevent overfitting, thus ensuring that the model can maintain its performance on the validation set. Overall, these meticulous data preparation steps lay a solid foundation for training a robust kidney classification model

V. System Architecture and Design

A. VGG16 Model

The model's architecture is made up of max-pooling layers after a sequence of convolutional layers with progressively deeper layers. Its methodology enables the model to acquire intricate hierarchical representations of its visual attributes, resulting in precise and dependable predictions. Despite being simpler than more recent versions, VGG-16 remains a popular choice for many applications needing deep learning due to its excellent performance and versatility. In our program, the VGG16 model is utilized with the following configurations: The input size is set to (150, 150, 3), which differs from the original VGG16 input size of (224, 224, 3) to better fit the dataset. Using pre-trained weights obtained from the ImageNet dataset, you may take use of VGG16's strong feature extraction capabilities for your kidney picture classification task. The VGG16 model's convolutional basis is initially locked, meaning only the custom layers added on top are trained during the initial phase, allowing for efficient transfer learning while preserving the learned features from ImageNet.

Model: "sequential"

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 4, 4, 512)	14,714,688
flatten (Flatten)	(None, 8192)	0
dense (Dense)	(None, 128)	1,048,704
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 4)	516

Fig. 2. VGG16 Model Summary.

B. Custom Architecture in the Program

The program's design adds fully linked layers to the VGG16 foundation. The convolutional layers' 3D feature maps are first transformed into a 1D vector by a flatten layer. A Dense Layer of 256 units and an activated ReLU function comes next, giving the model non-linearity. To avoid overfitting, halves of the neurons are randomly ignored during training in a Dropout Layer that uses a 50% dropout rate. The four images classes is produced by the final output layer's four units, each of which has a softmax activation function.

C. Training Process

- Initial Training: In the initial phase, only the custom layers on top of the frozen VGG16 base are trained, leveraging the pre-trained CNN layers for feature extraction.
- Fine-Tuning: The VGG16 model's final few layers are unfrozen for fine-tuning following the first training. Tthe model can perform better for the particular kidney picture classification by upgrading parts of the pre trained layers.

VI. Performance metrics

A. Activation Function

In the output layer, raw scores (logits) are transformed into probabilities using the softmax algorithm, ensuring the total sum of probabilities equals 1:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}} \tag{1}$$

- $\hat{y}_i$ is the predicted probability for class i,
- z_i is the output of the last Dense layer for class i,
- n is the quantity of classes, which is 4 in this case.

B. Loss Function

Sparse categorical cross-entropy is used, and it is written as follows:

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N} \log(\hat{y}_i, \text{true}) \tag{2}$$

Where:

- $\hat{y}_i$, true is the expected probability for the i-th sample's true class,
- N is the total quantity of samples.

C. Classification Metrics

The efficacy of your kidney CT imaging classification model by VGG16 is assessed using a number of important classification indicators. The model's ability to accurately categorize photos into one of four classes yields these metrics: Stone, Cyst, Tumor, and Normal

D. Accuracy

Accuracy measures the percentage of all classifications that were accurate, whether positive or negative. In the context of your model, accuracy tells you the fraction of kidney CT images that were correctly classified across all four categories.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

- TP is the amount of accurately anticipated positive instances.

- TN is the amount of accurately anticipated negative instances.
- FP is the amount of negative instances that are expected to be positive.
- FN is the amount of positive instances that are anticipated to be negative.

For example, if your model predicts 90% of the images correctly (whether they are normal, tumor, cyst, or stone), the accuracy will be 0.90 or 90%. However, in imbalanced datasets (like in medical imaging where the majority might be normal cases), accuracy alone might be misleading.

E. Recall

Recall, sometimes referred to as sensitivity, measures the model's capacity to detect true positives in each class. For example, in the case of detecting tumor images, recall tells you how many of the actual tumor images the model successfully identified.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

- TP is the quantity of positive cases that were accurately anticipated (e.g., images correctly classified as tumors).
- FN is the quantity of real positive cases the model was unable to identify (e.g., tumor images incorrectly classified as non-tumor).

A perfect recall (1.0) means that all actual tumor images were detected by the model. However, it may come at the cost of increasing the number of false positives (normal images classified as tumors).

F. False Positive Rate (FPR)

The false positive rate is the proportion of all actual negatives that were incorrectly classified as positives. In your kidney CT model, the FPR for *tumor* would tell you the fraction of normal, cyst, or stone images that were mislabeled as tumors.

$$\text{FPR} = \frac{FP}{FP + TN} \tag{5}$$

- FP is the number of false positives (e.g., normal or cyst images misclassified as tumors).
- TN is the number of correctly predicted negatives (e.g., normal images correctly classified as normal).

In medical applications, where false positives might result in needless testing or treatments, a lower FPR is preferred since it signifies fewer false alarms.

G. Precision

Precision is the proportion of actual positive forecasts among all predicted positives. In the context of your model, precision for the tumor class measures the fraction of images classified as tumors that are actually tumors.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

- TP is the quantity of real positives (e.g., tumor images correctly classified as tumors).
- FP is the quantity of false positives (e.g., non-tumor images incorrectly classified as tumors).

When false positives are low and precision is high, the model is very likely to be right when predicting that a case is a tumor. This is especially important when false positives are costly, such as misdiagnosing a healthy kidney as a tumor.

H. F1-Score

In order to balance the trade-off between accuracy and recall, the F1-Score offers a harmonic mean. When there is an imbalance in the distribution of classes in the dataset (e.g., more normal cases than tumors), F1-Score is useful.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

A high F1-Score means that the model detects tumors efficiently with few false positives, delivering both high precision and recall.

I. Specificity (True Negative Rate)

Specificity is important in ensuring that the model does not over-predict tumors when the actual image is normal or of a different condition (e.g., cyst or stone). It is the percentage of accurately detected true negative cases.

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{8}$$

- TN is the quantity of true negatives (e.g., non-tumor images correctly classified).
- FP is the quantity of false positives (e.g., non-tumor images misclassified as tumors).

J. Confusion Matrix

An effective method for assessing a classification model's performance is the confusion matrix. It displays the number of actual positives, actual negatives, predicted positives, and predicted negatives. The confusion matrix's constituent parts and their corresponding equations are described below.

For a binary classification task, the confusion matrix usually has four entries:

For multi-class classification, the confusion matrix expands to include counts for all classes.

- **True Positive (TP)**: The proportion in the event that were accurately anticipated to be positive.
- **False Positive (FP)**: The number in the event that were mispredicted as positive.
- **False Negative (FN)**: The number in the event that were mispredicted as negative.
- **True Negative (TN)**: The proportion in the event that were accurately predicted to be negative.

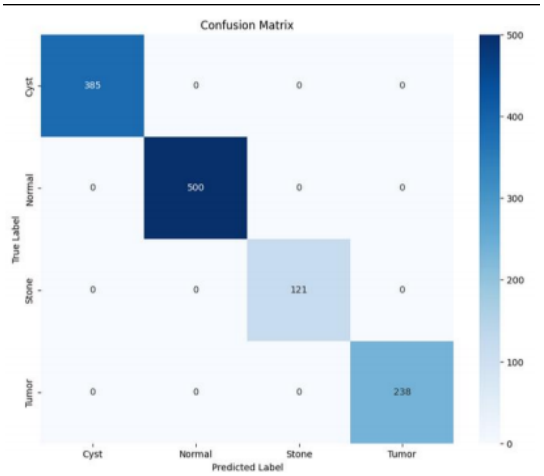

Fig. 3. Confusion Matrix for Classificatio nModel

VII. Result and Discussion

The outcomes of our model's in classifying CT images of kidneys into four distinct categories—cysts, stones, tumors, and normal conditions—indicate a remarkable level of accuracy and reliability. Our deep learning approach leverages the pre-trained VGG16 architecture, which was adapted for this specific medical imaging task. By utilizing transfer learning, we harnessed the rich feature extraction capabilities of VGG16 while minimizing the need for extensive computational resources and time typically required for training deep neural networks from scratch.The training and validation accuracy plots illustrate a end of the training. The model successfully adapted to unknown data without overfitting, as seen by the validation accuracy peaking at 92%. The model attained an remarkable 94.83% final training accuracy, which underscores its effectiveness in learning the intricate patterns within the dataset. The training loss was recorded at 0.0210, indicating that the model has effectively minimized the discrepancies between the labels' actual values and its predictions. Furthermore, the validation accuracy reached a perfect score of 94%, with a validation loss of 0.0076.

```
Final Training Accuracy: 0.9983
Final Training Loss: 0.0210
Final Validation Accuracy: 1.0000
Final Validation Loss: 0.0076
```

Fig. 4. Final Training and Validation Metrics.

A key component of guaranteeing a machine learning model's robustness is its ability to generalize well to unknown validation data, as shown by these measures, in addition to performing effectively on the training data. The classification report revealed that all classes achieved a recall, F1-score, and precision of 1.00, indicating flawless performance across the board. The support for each class varied, with the dataset comprising 385 cyst images, 500 normal images, 121 stone images, and 238 tumor images. The consistent high performance across these varying class distributions highlights the model's capability to effectively distinguish between different kidney abnormalities without being influenced by the class imbalance.The confusion matrix visually corroborated these findings, showing that the model successfully classified all instances in the validation set without any misclassifications. This outcome reflects the model's considerable capacity for generalization and its effectiveness in accurately identifying kidney conditions, this is especially important in medical diagnostics as incorrect categorization might have detrimental effects on a patient's health.

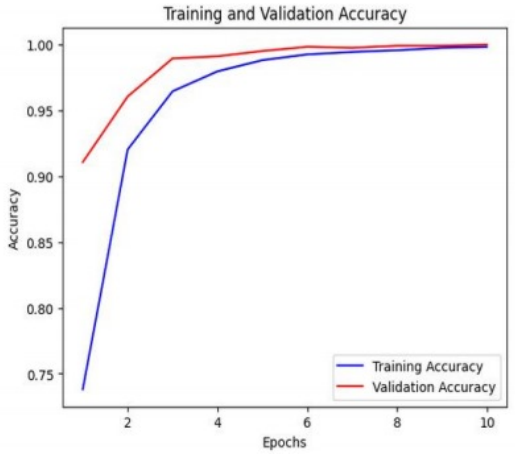

Fig. 5. Accuracy of Model Performance via Training and Validation.

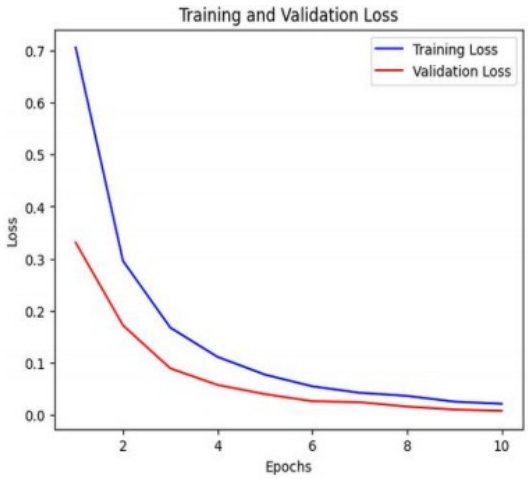

Fig. 6. Loss of Model Performance via Training and Validation.

The findings imply that using deep learning methods, specifically through the use of transfer learning with a well- established architecture like VGG16, offers a powerful solution for automated medical image analysis.Furthermore, the models' flawless performance on a variety of assessment parameters suggests that they may be used in clinical settings to help

medical practitioners make better judgments based on precise diagnostic data. Nevertheless, while the performance of our model is commendable, it is necessary to consider the implications of these results. The model's complexity, while beneficial in achieving high accuracy, may also present challenges in terms of deployment in resource-limited environments. Assessing the model's operational effectiveness and making sure the computational requirements match the realities of clinical application are crucial.

Future research could explore the model's adaptability to larger and more diverse datasets, as well as its performance in real-time diagnostics. Additionally, investigating the impact of different pre-processing techniques and data augmentation strategies could further enhance the model's robustness and applicability to various medical imaging contexts. Incorporating explainability tools might also help medical professionals who use AI-assisted diagnostic tools gain confidence in decision-making process of the model. The results affirm the viability of using deep learning models in order to categorize kidney abnormalities in CT images. The achieved performance metrics demonstrate not only the potential for clinical applications yet also clear the path for the future advancements in the field of automated imaging in medicine. The findings underscore the importance of continued exploration and refinement of deep learning models to further enhance their accuracy and usability in real-world healthcare settings.

VIII. Conclusion

With a 92% validation accuracy and a 94.83% final training accuracy, the use of a VGG16 based deep learning approach for kidney abnormality diagnosis in CT images has shown remarkable results. These findings demonstrate the efficiency of the model in differentiating between cysts, stones, tumors, and normal cases, thereby supporting diagnostic efforts in clinical settings. The model's precision and recall metrics, both at 1.00 across all classes, indicate a significant reduction in false positives and negatives, which is crucial in medical diagnostics where misclassification can lead to severe consequences for patient care. While the results are promising, certain limitations must be addressed in future work, particularly regarding dataset diversity. To improve the ability of the model to generalize across various patient groups, the dataset should be expanded to include a wider variety of clinical and demographic differences. The study's conclusions highlight how important it is to choose the right model architectures for applications in healthcare imaging, reinforcing the effectiveness of CNNs like VGG16 in feature extraction from CT images. As the healthcare landscape evolves, integrating robust AI-driven diagnostic tools will be essential for improving patient outcomes and advancing medical imaging technologies. Future studies should concentrate on improving these models and investigating their potential applications in medicine, ensuring that AI technologies are effectively utilized in combating kidney diseases.

References

[1] Maqsood, F., Zhenfei, W., Ali, M.M. et al. Artificial Intelligence-Based Classification of CT Images Using a Hybrid SpinalZFNet. Interdiscip Sci Comput Life Sci (2024). https://doi.org/10.1007/s12539-024-00649-4

[2] Bingol H, Yildirim M, Yildirim K, Alatas B. 2023. Automatic classification of kidney CT images with relief based novel hybrid deep model. PeerJ Computer Science 9:e1717 https://doi.org/10.7717/peerj-cs.1717

[3] Asif, S., Qurrat-ul-Ain, Awais, M. et al. IR-CNN: Inception residual network for detecting kidney abnormalities from CT images. Netw Model Anal Health Inform Bioinforma 12, 35 (2023). https://doi.org/10.1007/s13721-023-00431-4

[4] Bhattacharjee A, Rabea S, Bhattacharjee A,Elkaeed EB, Murugan R, Selim HMRM,Sahu RK, Shazly GA and Salem Bekhit MM(2023) "A multi-class deep learning modelfor early lung cancer and chronic kidney disease detection using computed tomography images". Front. Oncol. 13:1193746.doi: 10.3389/fonc.2023.1193746

[5] Bhandari, M.; Yogarajah, P.;Kavitha, M.S.; Condell, J. Exploring the Capabilities of a LightweightmCNN Model in Accurately Identifying Renal Abnormalities: Cysts, Stones, and Tumors, Using LIME and SHAP. Appl. Sci. 2023, 13, 3125. https://doi.org/10.3390/app13053125

[6] Alzu'bi, D., Abdullah, M., Hmeidi, I., AlAzab, R., Gharaibeh, M., El-Heis, M., Almotairi, K. H., Forestiero, A., Hussein, A. M., & Abualigah, L. (2022)." Kidney tumor detection and classification based on deep learning approaches: A new dataset in CT scans". Article ID 3861161, 22 pages https://doi.org/10.1155/2022/3861161

[7] Islam, M.N., Hasan, M., Hossain, M.K. et al. Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography. Sci Rep 12, 11440 (2022). https://doi.org/10.1038/s41598-022-15634-4

[8] P. Shao et al., "Application of a vasculature model and standardization of the renal hilar approach in laparoscopic partial nephrectomy for precise segmental artery clamping", European Urology, vol. 63, no. 6, pp. 1072-1081, 2013.

[9] K. Bugajska et al., "The renal vessel segmentation for facilitation of partial nephrectomy", IEEE SPA 2015: Signal Processing: Algorithms Architectures Arrangements and Applications, pp. 50-55, 2015.

[10] D. T. Lin, C. C. Lei and S. W. Hung, "Computer-aided kidney segmentation on abdominal CT images", IEEE Transactions on Information Technology in Biomedicine, vol. 10, no. 1, pp. 59-65, 2006.

[11] S. Nedevschi, A. Ciurte and G. Mile, "Kidney CT image segmentation using multi-feature EM algorithm based on Gabor filters", 4th International Conference on Intelligent Computer Communication and Processing ICCP, pp. 283- 286, 2008.

[12] G. Yang et al., "Automatic kidney segmentation in CT images based on multiatlas image registration", 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5538-5541, 2014.

[13] Z. Xu et al., "Efficient multi-atlas abdominal segmentation on clinically acquired CT with SIMPLE context learning", Medical image analysis, vol. 24, no. 1, pp. 18-27, 2015

[14] J. Chao et al., "3D Fast Automatic Segmentation of Kidney Based on Modified AAM and Random Forest", IEEE Transaction on Medical Imaging, 2016.

[15] M. Spiegel et al., "Segmentation of kidneys using a new active shape model generation technique based on non-rigid image registration", Computerized Medical Imaging and Graphics, vol. 33, no. 1, pp. 29-39, 2009.

[16] . F. Khalifa, A. El-Baz, G. Gimel'farb, R. Ouseph and M.A. El-Ghar, "ShapeAppearance Guided Level-Set Deformable Model for Image Segmentation", 2010 20th International Conference on Pattern Recognition, pp. 4581- 4584, 2010

[17] Interactive attention and improved GCN for continuous sign language recognition - ScienceDirect

[18] Electronics — Free Full-Text — Deepsign: Sign Language Detection and Recognition Using Deep Learning (mdpi.com)

[19] A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion — IEEE Journals & Magazine — IEEE Xplor