

Genetic Algorithm Optimized Vision Transformer Based Transfer Learning Approach for Early Prediction of Chronic Kidney Disease

Arindam Roy¹; Avishek Gupta²; Ram Prasad Chakraborty³

^{1,2}Assistant Professor, Department of Computer Science & Engineering
Abacus Institute of Engineering & Management, Magra, Hooghly, West Bengal, India

³Assistant Professor, Department of Information Technology
Dr. B. C. Roy Engineering College, Durgapur, West Bengal, India

Publication Date: 2025/11/08

Abstract: Chronic Kidney Disease (CKD) is an ongoing health issue defined by the slow decline in kidney performance over a prolonged period. This condition can escalate to kidney failure, which is life-threatening without the absence of dialysis or a kidney transplant. The common key factors leading to CKD include diabetes, hypertension, Glomerulonephritis. Standard diagnostic practices tend to be laborious, require considerable resources, and are vulnerable to errors made by humans. With the ongoing advancements in artificial intelligence within the healthcare domain, machine learning and deep learning algorithms are playing a pivotal role in the accurate and effective detection of CKD. The purpose of the proposed research is to construct and authenticate a predictive model aimed at diagnosing chronic kidney disease. Image processing-based diagnostic approaches have shown a greater success rate compared to other detection methods. To tackle these challenges, our investigation proposes a new methodology that merges Genetic Algorithms (GA) with a Vision Transformer (ViT) model that employs Hierarchical Attention in a transfer learning framework, thereby improving both feature selection and classification accuracy. In this scholarly work, we have scrutinized the performance of two leading architectures, VGG16 and ResNet50, and have proposed an attention-centric approach that employs the Vision Transformer model optimized by evolutionary algorithm. The attention mechanism has the potential to grasp long-term dependencies in images. Extraction of complex features is to be done using ViT Model, preceded by several hyper parameters like number of epochs, learning rate, batch size, number of layers, layer size, number of attention heads, and attention window size should be optimized through a genetic algorithm to enhance performance and also feature selection is done by GA to get the optimized result. A dataset available to the public consists of four types of image data of kidneys: Cyst, Normal, Tumor, and Stone, which are used in the three architectures discussed above. In total, there are 12,446 images that are segmented for training, testing, and validation purposes. Experimental results demonstrate that our GA optimized ViT Model surpasses state-of-the-art traditional models, achieving a 98.05% F1-score, and the model also shows superiority in terms of trainable parameters

Keywords: Chronic Kidney Disease, Fast ViT (Vision Transformer), Genetic Algorithm (GA), Transfer Learning, ResNet50, VGG16.

How to Cite: Arindam Roy; Avishek Gupta; Ram Prasad Chakraborty (2025) Genetic Algorithm Optimized Vision Transformer Based Transfer Learning Approach for Early Prediction of Chronic Kidney Disease. *International Journal of Innovative Science and Research Technology*, 10(11), 67-76. <https://doi.org/10.38124/ijisrt/25nov076>

I. INTRODUCTION

Chronic kidney disease (CKD) is a progressive and long-term condition characterized by damage to the kidneys, which impairs their ability to filter blood effectively. This impairment results in the collection of waste and fluid within the body. The damage to the kidneys hinders their function of filtering blood and eliminating waste and excess fluid as they are meant to do. Common symptoms of CKD include fatigue, swelling, elevated blood pressure, and itchy or dry in the early

stages, while more advanced stages may present with shortness of breath, vomiting, nausea, and changes in skin color. Cardiovascular disease (CVD) can be defined as a disorder related to heart and blood cells. The main parts of cardiovascular system are the heart, veins, arteries, and blood capillaries. Coronary Artery Disease, hypertension, cardiac arrest, stroke are some of the common types of cardiovascular disease. Among a range of significant reasons such as uncontrolled diabetes, smoking, obesity, depression and high

level of cholesterol CKD also raises the risk of cardiovascular disease.

As per the records on the basis of several surveys which was carried out during 2024 [1], over 10% of our whole population is affected by chronic kidney disease (CKD) and is marked as the 16th leading cause of death. It is anticipated that CKD may rise to the 5th position by 2040 [2]. Most of the times it may be observed that there is no noticeable symptoms, therefore identifying and addressing of CKD at the earlier stage is very challenging that causes permanent harm. Traditional diagnostic methods such as blood test, imaging, urine analysis as a treatment for CKD can be laborious, intrusive and prone to human error. In rural area it is also observed that the delay of proper diagnosing due to scarcity of potential doctors. Based on the types and intensity of renal failure, several alternative treatments exist, such as addressing the root cause, maintaining of proper a diet chart, proper medications, and in critical situation kidney replacement may be required [3].

Although numerous pathological tests, majorly x-ray computed tomography scans, Magnetic Resonance Imaging (MRI), and CT scans, are commonly utilized for the early detection and treatment of kidney ailments like cysts, stones, and tumors, the global shortage of nephrologists and radiologists is a significant concern. The advent of deep learning, a field within artificial intelligence, has increasingly underscored the role of computer-aided diagnosis. With the ongoing development of machine learning and data mining algorithms, a variety of techniques have been applied to medical imaging data for disease analysis and early prediction [4-7]. In the field of biomedical imaging, numerous deep learning models, especially those utilizing Convolutional Neural Networks (CNNs), have surpassed conventional machine learning approaches like support vector machines (SVM), Artificial Neural Networks (ANN), decision trees (DT), random forests (RF), and logistic regression (LR). CNNs excel in image processing tasks due to their capability to analyze hierarchical representations systematically. To optimize their complementary strengths, CNNs and RNNs are combined within a hybrid model.

Nevertheless, traditional CNN-based models encounter considerable obstacles when it comes to the effective processing of kidney images. This problem primarily occurs due to the variations in kidney structure among various individuals, as well as similarity in grayscale values between the kidney and surrounding organs such as the liver and spleen, which may result in misclassification. Numerous studies also face issues related to class imbalance, where some classes are represented by fewer examples, potentially resulting in biased models that may overfit the more prevalent classes while struggling to accurately predict fewer common conditions. Moreover, CT images often include noise, which disrupts the feature learning capabilities of conventional CNNs, making it difficult to extract the crucial features essential for precise predictions. While the kidney's structure reveals a specific level of spatial dependency, the convolution processes in standard CNNs are limited to capturing solely

local features, which in turn diminishes their predictive capability.

Therefore, in this research, we have proposed a Genetic Algorithms optimized Vision Transformer (ViT) model for the early identification of chronic kidney disease. When compared to traditional CNN, the ViT demonstrates superior capability in capturing global dependencies in kidney CT images, revealing significant benefits in feature representation. The Vision Transformer amalgamates attention mechanisms, feedforward networks, and normalization layers to discern intricate patterns, facilitating faster convergence and more accurate representations of both local and global features. Several types of deep learning models such as Convolution Neural Network (CNN), Recurrent Neural Network (RNN), autoencoder are combined by leveraging the benefits of evolutionary algorithms like Genetic Algorithms, Particle Swarm Optimization. These advantages encompass educational system, complex network design, and hyperparameters optimization [8]. Moreover, after the feature extraction using Vision Transformer, a deep feature selection is conducted by our evolutionary Genetic Algorithm to get the optimal best result. The primary goal of this optimization strategy is to boost the overall performance metrics of the model, which includes precision, accuracy, recall, and F1-score. The proposed approach is validated and tested on a publicly available dataset of kidney tumors [9].

II. LITERATURE REVIEW

There is an urgent need for more precise and effective diagnostic techniques since kidney stones have become much more common in recent decades. If not identified early, kidney stones—which are usually made of calcium oxalate, uric acid, or other crystalline materials—can cause excruciating pain, blockage of urinary channels, and long-term renal problems. Although they are still often employed, traditional diagnostic methods like X-rays and ultrasonography frequently lack the sensitivity to detect minute stones or minute structural irregularities. Since non-contrast computed tomography (NCCT) has almost 100% sensitivity and specificity, it is currently considered the gold standard. Concerns regarding radiation exposure are raised by its widespread use, particularly for patients who need recurrent imaging [10] Khan et al. (2022). This has spurred academics to look at several approaches that strike a compromise between patient safety and diagnostic accuracy.

In order to improve the diagnostic accuracy of grading of Clear cell Renal Cell Carcinoma (ccRCC) using CT images, [11] Yang et al. (2022) suggested a Transformer-based deep learning algorithm. The model's performance was assessed using sensitivity, average classification accuracy, specificity, and Area Under Curve (AUC) on a dataset of 759 patients. By comparing the performance with other deep learning models this transformer-based architecture performs far better, with an average accuracy, sensitivity and specificity of 87.1%, 91.35, 85.3%, and an AUC of 90.3% respectively. Transformer-based network with self-attention mechanism was mainly used for complex feature extraction and a non-linear classifier was employed for final

classification. Further improvement was demonstrated by the integrated model, which included various training models, with an accuracy of 86.5% and an AUC of 91.2%. It indicates that the Transformer-based network is more efficient than normal deep learning algorithms for ccRCC grading, and it demonstrates robustness in handling noise in CT images, suggesting potential applications in other jobs.

In [12] Lakshmi et al. (2023) used Mask R-CNN (Region Based Convolution Neural Network) a two-stage deep learning model for object detection and classification, for kidney tumor classification based on the X-ray image of kidneys. They performed hyperparameter tuning of MASK R-CNN model architecture to get the best result at optimal time. They achieved training accuracy of 98.61% and for testing 93.89%.

In [13] Mehmet et al. (2023) employed a GAN (Generative Adversarial Networks) based approach combined with conventional CNN and BiLSTM for enriching the accuracy of model's transfer learning ability for tumor detection using medical imaging data. They used i2b2/UTHealth 2010 dataset for their research. The GAN network was used for data augmentation and pseudo-data generation algorithms. Main method in this research was divided into mainly two steps –tumor location identification, size calculation, and metastatic tumor sites identification. This initiative seeks to improve accuracy and generalizability in scientific event extraction, particularly with respect to tumor-related events.

In [14] Nagawa et al. (2024) in their work titled "Three-Dimensional Convolutional Neural Network-Based Classification of Chronic Kidney Disease Severity Using MRI" employ MRI images to develop a 3D convolutional neural network (CNN) model that categorizes the severity of chronic kidney disease (CKD). The model achieved an accuracy of 86.2%, demonstrating optimal performance when analyzing images of both kidneys. This approach leverages deep learning techniques to classify the stages of CKD, potentially enhancing disease management.

After evaluating 196 adult patients, Bittencourt, Jalila Andréa Sampaio, et al. (2024) in [15] found that 12.24% of them had chronic kidney disease (CKD), and 45.8% of them also had metabolic syndrome (MS). In order to screen for MS, their study used the KNN algorithm, which showed remarkable sensitivity and specificity.

In order to help in early detection, Satukumati, S. B., & Bhat, M. N. (2024) in [16,17] created a fuzzy neural system for predicting chronic kidney disease (CKD) by evaluating patient data. Although the method makes use of both fuzzy logic and neural networks, one of its main disadvantages is how difficult it is to train and interpret. Practical implementation is hindered by the computational complexity and tuning difficulty, particularly when dealing with huge datasets.

The effectiveness of ML techniques for CKD identification was emphasized by Ghosh et al. (2024) [18] a hybrid model and a comprehensive analysis of different machine learning techniques utilizing the UCI CKD failure dataset. The accuracy achieved by the hybrid technique was 94.99%. The study demonstrated the possibilities of data mining in the healthcare industry. Limitations were also mentioned, such as the use of a single dataset and the absence of clinical validation. A DL model for CKD initial prediction and detection was proposed by Singh et al. [19], who also evaluated its performance against other modern machine learning methods. Recursive Feature Elimination (RFE) was used to select the key features, serum creatinine and hemoglobin, and missing data were substituted using averages. The DL model outperformed the other classifiers in terms of accuracy. However, the short dataset employed hindered the model's evaluation.

III. PROPOSED METHODOLOGY

A diagrammatic representation of the various steps is given in Figure 1 that describes the total workflow of early prediction of chronic kidney disease using dual level GA optimized Vision Transformer model. Each of the steps has been illustrated in details in the below sub-sections.

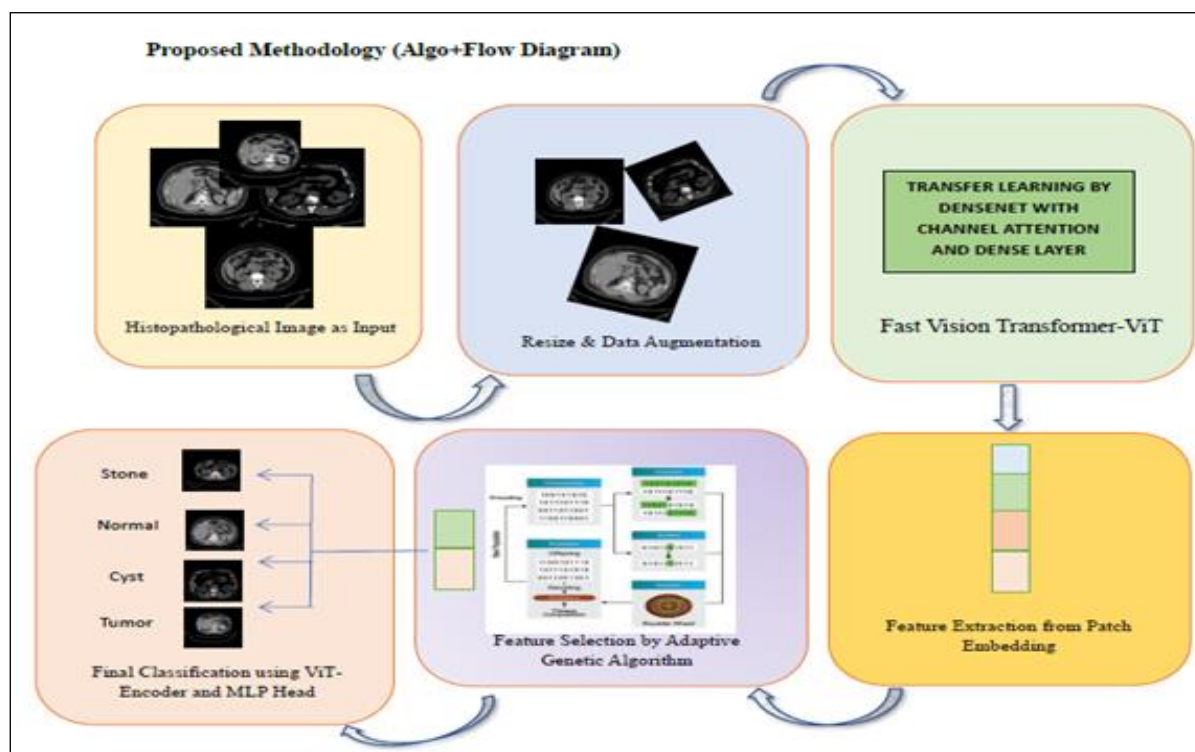


Fig 1 Workflow of Proposed Methodology of Early Prediction of CKD using GA-ViT

➤ Data Preprocessing:

In this research, we conduct data preprocessing on all CT images of CT KIDNEY DATASET [20] publicly available a benchmark dataset. The dataset is divided into training, validation and test sets in the ratio of 80:10:10. In order to boost the model performance and enhance its generalization capabilities we adopt several preprocessing techniques. At the outset of the training data preparation, we adjust the pixel values by dividing them by 255, which ensures all images poses consistent pixel intensity values. We utilize a range of data augmentation techniques, including random rotation at the range of 20 to 30 degrees, horizontal flipping, noise removal, shear transformation, contrast alteration, reflection and scaling operation. By disabling smoothing, one eliminates smooth transitions in intensity. The median filter is instrumental in noise reduction and the enhancement of image quality. The purpose of this augmentation is to boost the diversity of the training data and facilitate the model's ability to learn strong features.

➤ Hyperparameter Optimization of ViT Model:

Hyperparameter optimization is essential for the proper functioning deep learning models. The automation of this process has gained traction recently, employing advance methodologies like genetic algorithm. The GA based ViT framework employs Genetic Algorithm (GA) to systemically optimize model configurations by exploring a predefined hyperparameter search space (Table I).

Table 1: Hyperparameter Search Space for Ga Based VIT Architecture

Hyperparameter	Search Range	Description
patch size	[16, 28, 32]	Input image size divisors e.g. 255
num_layers	[8, 16]	Number of layers (min:8, max:16)
embedding_dim	[192, 528]	Embedding dimension (min:192, max:528)
num_heads	[3,8]	Number of attention heads (min:3, max:8)
mlp_dim	[768, 1536]	MLP dimension (min:768, max:1536)
dropout_rate	[0.0, 0.5]	Dropout rate (min:0.0, max:0.5)

➤ Feature Extraction Using Transfer Learning of Vision Transformer:

The methodology utilizes a transfer learning strategy that incorporates ResNet50 within a Vision Transformer (ViT) to classify chronic kidney disease in CT scan images. This framework capitalizes on the robust feature extraction abilities of a pre-trained ResNet50 model to improve the efficiency of a ViT, especially in contexts with limited data or specific tasks. ResNet50 is a prominent deep learning architecture recognized for its depth and residual learning features, which enhance feature extraction and address the vanishing gradient issue. In our methodology, we employ ResNet50 in its pre-trained state, taking advantage of its capacity to identify critical patterns in CT scan images.

In this methodology, the pre-trained ResNet50 model, trained on ImageNet, is utilized as a fixed feature extractor. The convolutional layers of ResNet50 extract comprehensive, hierarchical features from the input images. These extracted features are then fed as input to the Vision Transformer, replacing the conventional patch embedding layer. The self-attention mechanism of the ViT, along with its subsequent

layers, processes these pre-extracted features for classification or other tasks. This approach is advantageous when computational resources are limited or when a strong, generic feature representation is needed. Combining the strengths of CNNs (like ResNet50) to extract local features and ViTs for global context learning can lead to more robust and accurate models.

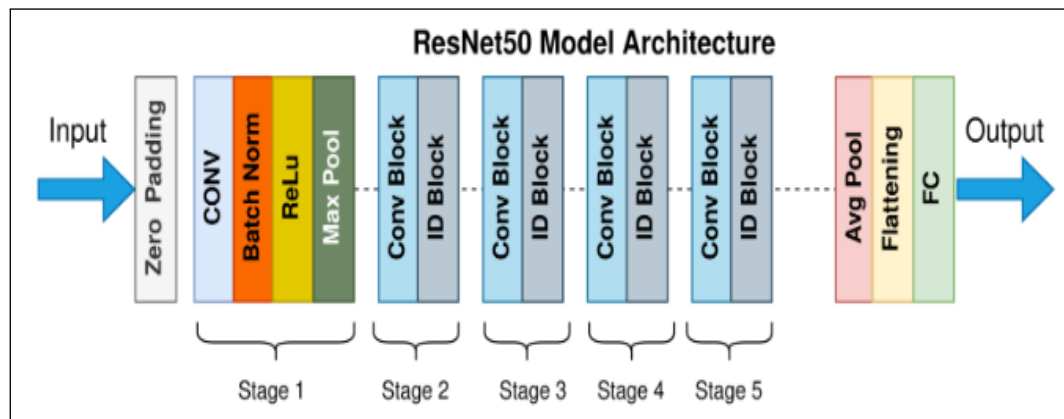


Fig 2: ResNet 50 Architecture
Source: Adapted from [18]

➤ GA Based Adaptive Feature Selection

A wide array of features can be produced through the feature extraction process of any deep learning model, which the classifier is required to analyze. Nevertheless, not all of these features are beneficial for classification tasks. The leftover components are either redundant or irrelevant, merely contributing to increased computation times and occupying additional physical space. Moreover, the presence of these non-informative features diminishes classification accuracy. In the context of the features generated from the

previously mentioned Transfer Learning method utilizing ResNet50, feature selection has been performed to eliminate these issues. In the proposed approach, the fitness of each chromosome within the GA population is assessed using two separate filter methods, rendering the GA an adaptive feature selection algorithm.

GA is a meta heuristic search approach that mimics the natural evaluation. In fact, a GA implementation involved with the realization of the following operations.

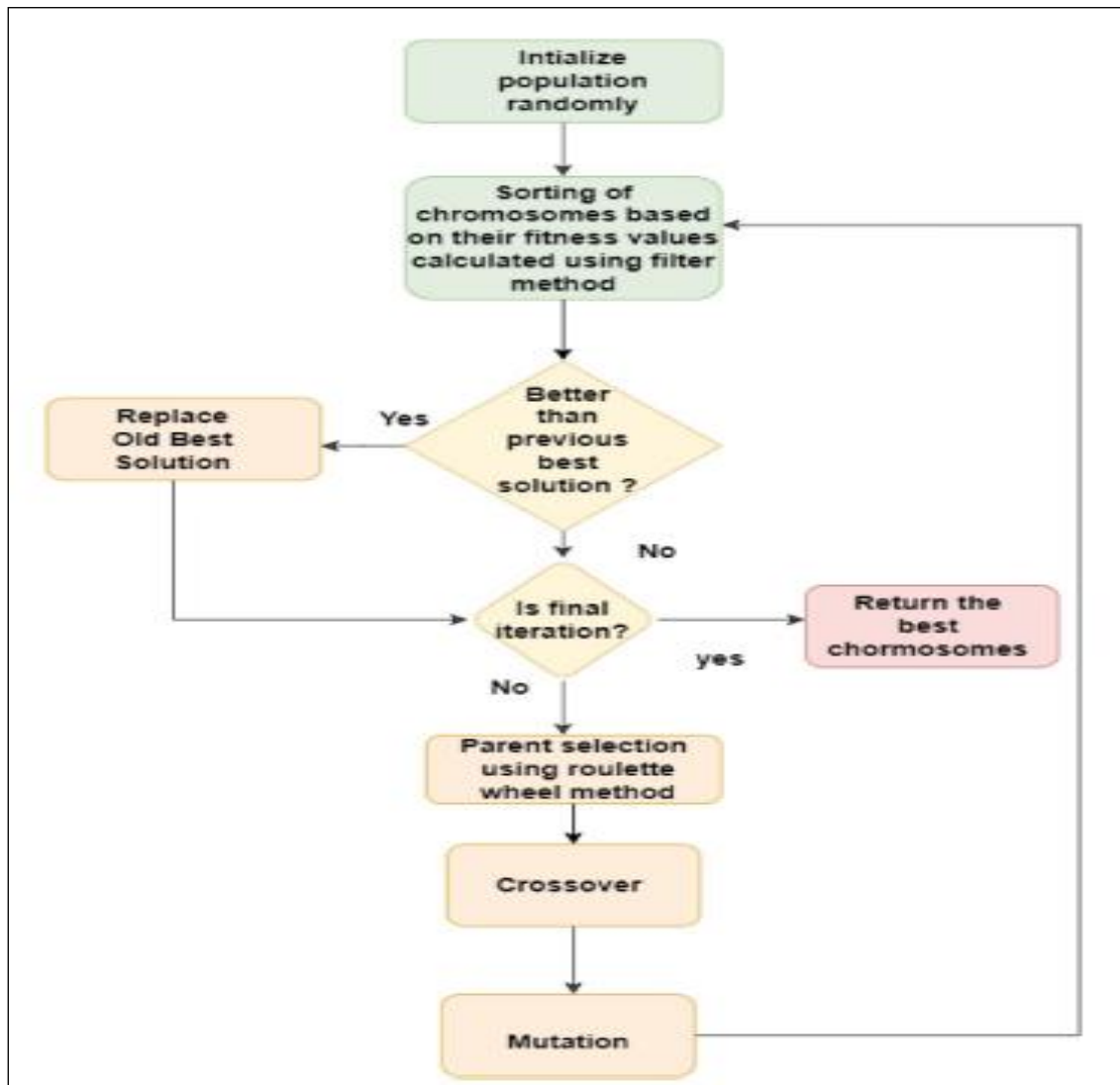


Fig 3: Various Stages of Genetic Algorithm

- **Chromosome Representation:** After feature extraction using pre-trained ResNet50 all the features are encoded as a binary coded array comprises of 0 (if not selected) and 1 (if selected) randomly that represents a chromosome. If n is the total number of features extracted and represented in the form of a binary vector $f = [f_1, f_2, f_3, \dots, f_n]$, where each f_i denotes the i^{th} feature.
- **Initial Population Generation:** Each generation consists of a population of individual chromosome that is randomly generated, each representing subset of potential solution features.
- **Fitness Score Evaluation:** Here in our research, we have used a single objective function to calculate the fitness score of each chromosome.

$$\text{fitness} = \alpha * \text{acc} + (1 - \alpha) * \left(\frac{\text{num}_{\text{features}} - \sum_i \text{feature}_i}{\text{num}_{\text{features}}} \right)$$

- **Selection:** Individuals with higher fitness are selected to become parents for the next generation. Here we have

used Roulette Wheel selection method as a selection process.

- **Crossover:** Selected parent chromosome are combined to create offspring that represents a new solution. In crossover technique basically exchanging of features between a pair of parent chromosome occurs to extend the diverse feature.

0	1	1	0	0	1	0
1	0	1	1	1	0	0

Before crossover

0	1	1	1	0	1	0
1	0	1	0	1	0	0

After crossover

- a) **Mutation:** Small random changes are introduced to offspring chromosome to enhance the diversity in the

population preventing premature convergence to sub optimal solutions. Basically with the probability μ 0 is replaced by 1 and 1 is replaced by 0 in some offsprings for some features.

b) **Termination:** The entire process (steps b-f) is repeated for a specified number of generations or until a predefined termination criterion is met. This criterion could be reaching a satisfactory fitness level, a maximum number of iterations, or a lack of significant improvement over a certain period.

➤ Classification

The feature selection phase is succeeded by the transformer encoder, which is then followed by classification performed by the MLP head of the GA optimized ViT Model. A distinctive, learnable [CLS] token is appended to the sequence of embedded image patches. This token is crafted to gather global information regarding the entire image. The combined sequence (embedded patches + [CLS] token) is processed through a series of Transformer encoder layers. These layers utilize multi-head self-attention and feed-forward neural networks to discern contextual relationships between the patches and the [CLS] token. After traversing the Transformer encoder, the final

representation of the [CLS] token is obtained. This token now encapsulates the aggregated information required for classification. The representation of the extracted [CLS] token is then directed through the MLP head.

- **Architecture:** The MLP head typically consists of one or more fully connected (dense) layers, frequently with non-linear activation functions like GELU (Gaussian Error Linear Unit) in between.
- **Output:** The final layer of the MLP head outputs a vector of logits, where each logit corresponds to a specific class.
- **Softmax Application:** A softmax function is applied to these logits to convert them into a probability distribution over the classes, indicating the likelihood of the image belonging to each class.
- **Prediction:** The class with the highest probability is then chosen as the predicted class for the input image.

➤ Proposed GA Optimized ViT Architecture

- **Image Input:** Input images are fed into the model.
- **Patch Embedding:** Images are divided into patches, and each patch is embedded into a vector representation.

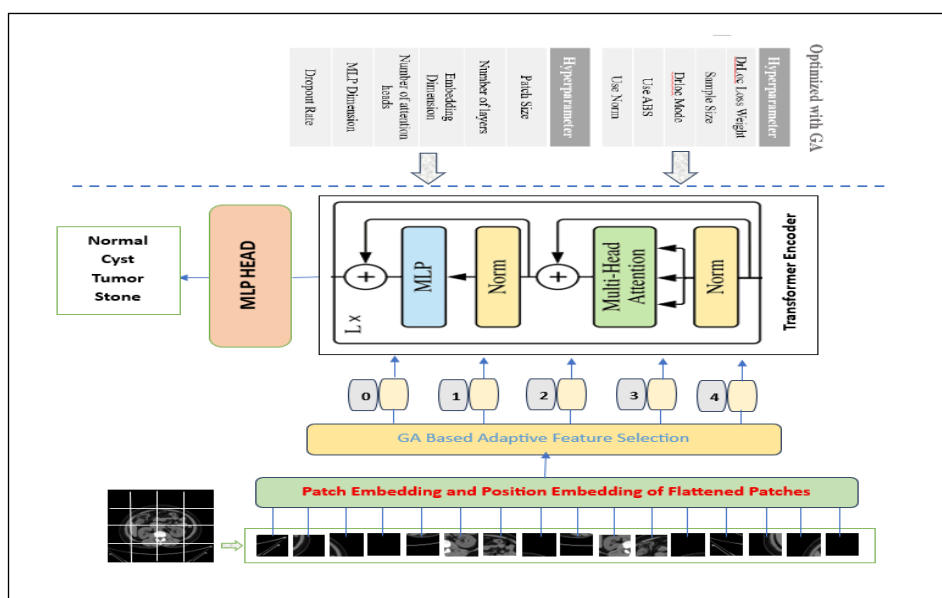


Fig 3: GA Optimized ViT Architecture

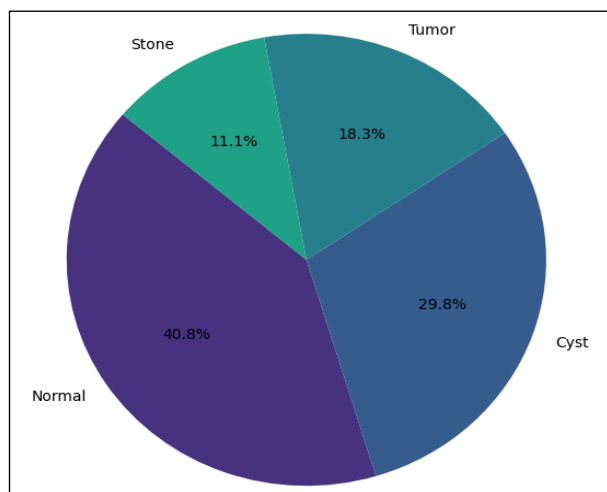
Source: author

- **GA Based Adaptive Feature Selection:** A genetic algorithm is used to select the most relevant features from the patch embeddings.
- **Population Initialization:** An initial population of candidate solutions (feature subsets) is generated randomly.
- **Fitness Evaluation:** The fitness of each candidate solution is evaluated using a predefined objective function (e.g., classification accuracy).
- **Selection:** The fittest candidate solutions are selected to reproduce and form the next generation.
- **Crossover and Mutation:** The selected candidate solutions undergo crossover and mutation operations to generate new offspring.
- **Termination:** The genetic algorithm terminates when a stopping criterion is met (e.g., maximum number of generations).
- **Vision Transformer Encoder:** The identified features are inputted into the encoder of Vision Transformer, which consists of several layers of self-attention and feed-forward neural networks.
- **Classification Head:** The output of the Vision Transformer encoder is fed into a classification head, which produces the final classification output.

IV. DATASET DESCRIPTION

Table 2 Dataset Description

Class	Number of Samples
Kidney Stone	1,377
Kidney Tumor	2,283
Kidney cyst	3,709
Kidney Normal	5,077
Total	12,446



V. EXPERIMENTAL RESULTS

➤ Performance Evaluation Matrices

This research utilizes several matrices to describe the accuracy of our proposed model.

- Accuracy: The ratio of correctly predicted samples to the total.

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

- Precision: This metric indicates the ability of the model to measure the positive predictions.

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

- Recall: The sensitivity or recall of a test is the ability to find the positive cases correctly

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

- F1-score: F-measure (F1-score) is the harmonic mean of recall and precision. F1-score equation is shown below.

$$\text{F1-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

- Confusion Matrix: The contingency matrix is called a confusion matrix. A confusion matrix is a matrix that provides a summary of the predictions made by the model

against actual outcomes when the ground truth is available.

Table 2: Confusion Matrix

	Actually Positive	Actually Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

Result Analysis

In this proposed research on CKD detection utilizing the GA optimized ViT model, we have conducted a comparison of the results with the classical Vision Transformer (ViT) and the Swin Transformer models, respectively.

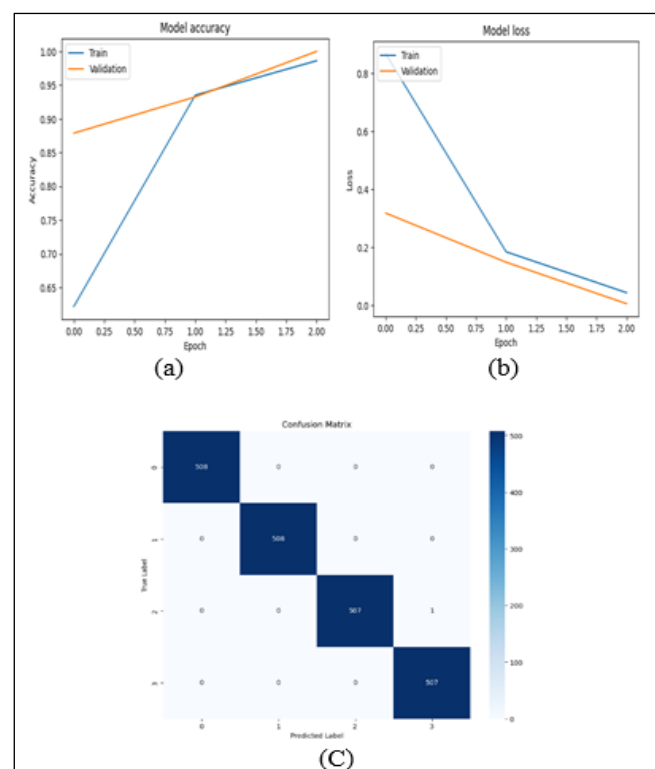


Fig. 5: Accuracy and Loss of Training and Validation Progress for ViT
Source: author

To classify into four categories, the ViT model is initially trained, and its performance is assessed using the test data. A confusion matrix is generated, and plots illustrating Training Accuracy versus Validation Accuracy, as well as Training Loss versus Validation Loss, have also been produced. In Fig. 5 (a) and (b) represents the accuracy and loss plots for the classical ViT model for CKD classification. Fig. 5 (c) represents the confusion matrix.

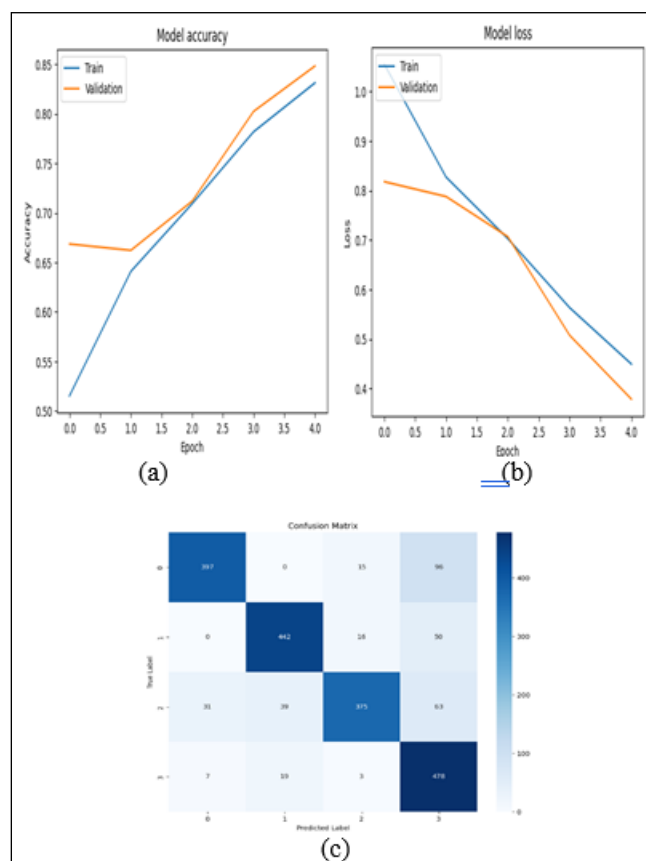


Fig. 6: Accuracy and Loss of Training and Validation Progress for Swin Transformer
Source: author

In Fig. 6 (a) and (b) represents the accuracy and loss plots for the classical Swin Transformer model for CKD classification. Fig. 6 (c) represents the corresponding confusion matrix.

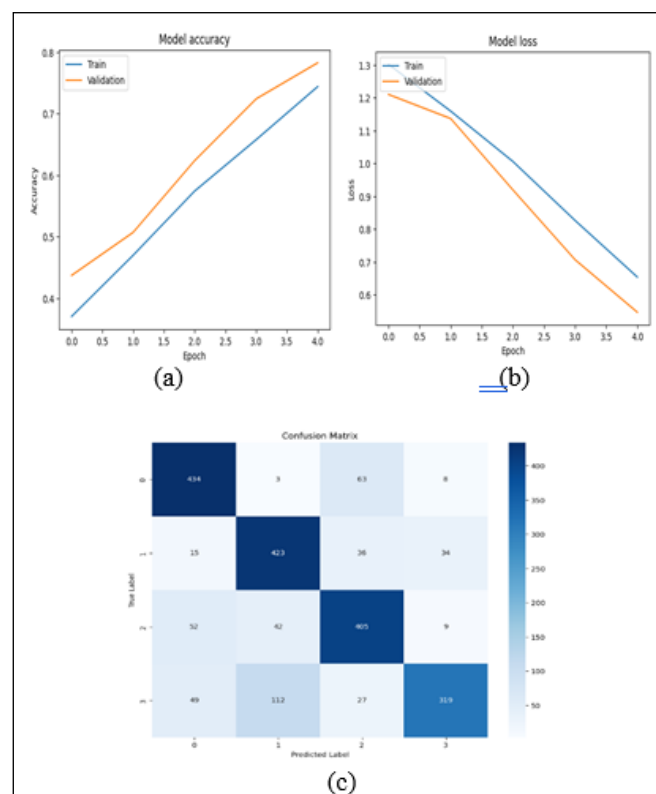


Fig. 7: Accuracy and Loss of Training and Validation Progress for Proposed GA optimized ViT
Source: author

In Fig. 7 (a) and (b) represents the accuracy and loss plots for our proposed Genetic Algorithm induced vision transformer GA optimized ViT model for CKD classification. Fig. 7 (c) represents the corresponding confusion matrix. Our GA optimized ViT has tiny training time per epoch of 11.80 seconds while ViT and Swin Transformer model have a training time of 184.67 and 160.25 seconds.

CONCLUSION

In conclusion, we adopted a Vision Transformer-based strategy with Genetic Algorithm-optimized hyperparameters to accurately classify chronic kidney disease from kidney CT images. The training dataset was composed of 80% of the total 12,446 images. The performance of each model was evaluated using various metrics, and the results were compared. Among the models, our GA optimized ViT model exhibited superior performance, achieving an impressive accuracy of 99.2% and a sensitivity value of 0.9928. The findings indicate that our GA optimized ViT model provides a highly effective solution for CKD classification, outperforming other models and emphasizing its potential for supporting clinical diagnosis.

REFERENCES

- [1]. A. K. Bello, I. G. Okpechi, A. Levin, F. Ye, S. Damster, S. Arruebo, J.-A. Donner, F. J. Caskey, Y. Cho, M. R. Davids et al., "An update on the global disparities in kidney disease burden and care across world countries

- and regions,” *The Lancet Global Health*, vol. 12, no. 3, pp.e382–e395, 2024.
- [2]. A. C. Webster, E. V. Nagler, R. L. Morton, and P. Masson, “Chronic kidney disease,” *The Lancet*, vol. 389, no. 10075, pp. 1238–1252, 2017.
 - [3]. N. Bhaskar and M. Suchetha, “Analysis of salivary components as non-invasive biomarkers for monitoring chronic kidney disease,” *International Journal of Medical Engineering and Informatics*, vol. 12, no. 2, pp. 95–107, 2020, doi: 10.1504/IJMEI.2020.106896.
 - [4]. K. Suzuki, Overview of deep learning in medical imaging, *Radiol. Phys. Technol.* 10 (3) (2017) 257–273.
 - [5]. J.-G. Lee, S. Jun, Y.-W. Cho, H. Lee, G.B. Kim, J.B. Seo, N. Kim, Deep learning in medical imaging: general overview, *Korean J. Radiol.* 18 (4) (2017) 570–584.
 - [6]. G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, *Med. Image Anal.* 42 (2017) 60–88.
 - [7]. G. Wang, W. Li, M.A. Zuluaga, R. Pratt, P.A. Patel, M. Aertsen, T. Doel, A.L. David, J. Deprest, S. Ourselin, et al., Interactive medical image segmentation using deep learning with image-specific fine tuning, *IEEE Trans. Med. Imaging* 37 (7) (2018) 1562–1573.
 - [8]. Maha, Gharaibeh., Dalia, Alzu'bi., Malak, Abdullah., Ismail, Hmeidi., Mohammad, Rustom, Al, Nasar., Laith, Abualigah., Amir, H., Gandomi. (2022). Radiology Imaging Scans for Early Diagnosis of Kidney Tumors: A Review of Data Analytics-Based Machine Learning and Deep Learning Approaches. *Big data and cognitive computing*, doi: 10.3390/bdcc6010029
 - [9]. Agarwal, Anush., Gaikar, Rohini., Schieda, Nicola., Elfaal, Mohamed, WalaaEldin., Ukwatta, Eranga. (2023). Deep-learning-based ensemble method for fully automated detection of renal masses on magnetic resonance images. *Journal of medical imaging*, doi:10.1117/1.JMI.10.2.024501
 - [10]. Khan, A., Das, R., & Parameshwara, M. C. (2022). “Detection of kidney stones using digital image processing: A holistic approach. *Engineering Research Express*,” 4(3), 035040. <https://doi.org/10.1088/2634-Fire>
 - [11]. M. Yang, X. He, L. Xu, et al., “Ct-based transformer model for non-invasively predicting the fuhrman nuclear grade of clear cell renal cell carcinoma,” *Frontiers in Oncology*, vol. 12, p. 961 779, 2022.
 - [12]. Sri, V. S., & Lakshmi, G. J. (2023, April). Detection Analysis of Abnormality in Kidney using Deep Learning Techniques and its Optimization. In *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)* (pp.1-6). IEEE.
 - [13]. Mehmet, Çifci., Sadiq, Hussain., Peren, Jerfi, Canatalay.(2023). Hybrid Deep Learning Approach for Accurate Tumor Detection in Medical Imaging Data. *Diagnostics*, doi: 10.3390/diagnostics13061025
 - [14]. K. Nagawa et al., 'Three-dimensional convolutional neural network-based classification of chronic kidney disease severity using kidney MRI,' *Scientific Reports*, 2024.
 - [15]. Bittencourt, J. A. S., Sousa, C. M., Santana, E. E. C., Moraes, Y. A. C. D., Carneiro, E. C. R. D. L., Fontes, A. J. C., ... & Nascimento, M. D. D. S. B. (2024). Prediction of metabolic syndrome and its associated risk factors in patients with chronic kidney disease using machine learning techniques. *Brazilian Journal of Nephrology*, 46(4), e20230135.
 - [16]. Satukumati, S. B., & Bhat, M. N. (2024). Early detection of chronic kidney disease using data mining methods. *International Journal of Advanced Intelligence Paradigms*, 28(1-2), 86-99.
 - [17]. L. Maria Michael Visuwasam, “NMA: Integrating Big Data Into A Novel Mobile Application Using Knowledge Extraction for Big Data Analytics Cluster Computing: The Journal Of Networks, Software Tools And Applications,22 (1),2018.
 - [18]. Ghosh, B. P., Imam, T., Anjum, N., Mia, M. T., Siddiqua, C. U., Sharif, K. S., ... & Hossain, M. Z. (2024). Advancing chronic kidney disease prediction: Comparative analysis of machine learning algorithms and a hybrid model. *Journal of Computer Science and Technology Studies*, 6(3), 15-21.
 - [19]. Singh, V., Asari, V. K., & Rajasekaran, R. (2022). A deep neural network for early detection and prediction of chronic kidney disease. *Diagnostics*, 12(1), 116.
 - [20]. <https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone>