ACTA SCIENTIAE

ISSN: 2178-7727

Meta-Heuristic Optimized Dual-Attention Deep Network with Depth-wise Separability for High-Precision Renal Tumor Diagnosis

Sachin D.Shingade[1], Midhun Chakkaravarthy[2], Dimitrios Karras[3], Komal M. Masal[4]

[1]Lincoln University College LUCM, Petaling Jaya, Malaysia,
pdf.sachin@lincoln.edu.my
[2]Lincoln University College LUCM, Petaling Jaya, Malaysia
midhun@lincoln.edu.my
[3]LUCM Malaysia, NKUA Athens Greece , EUT Albania,
dimitrios.karras@gmail.com
[4]Pune Institute of Computer Technology PICT, Pune, Maharashtra, India
kmmasal@pict.edu

ABSTRACT

The human kidney's primary function is to maintain homeostasis by filtering met-abolic byproducts and toxins from the circulatory system; however, the uncon-trolled proliferation of atypical cells can lead to tumor formation, presenting diag-nostic challenges due to the limitations specifically overfitting, high complexity, and low accuracy of current kidney tumor classification methods. This paper in-troduces a novel deep learning architecture "Meta-Heuristic Optimized Dual-Attention Deep Network with Depth-wise Separability for High-Precision Renal Tumor Diagnosis" designed to analyze CT scan data to accurately localize, quanti-fy (size), and map the spatial distribution of renal tumors, thereby significantly improving therapeutic planning. The process begins with crucial pre-processing using the Guided Triple Gaussian Functioning Filter to ensure effective noise reduction and optimal image clarity. Subsequently, a Pyramid Convolutional Co-ordinate Attention-based Residual Autoencoder is utilized for semantic feature extraction, employing its pyramid layers and coordinate attention to focus selec-tively on tumor regions across multiple scales. These refined features are then fed into the Squeeze-and-Excitation based Dual-Attention Module Multi-scale Con-volutional Depthwise Separable Network , a specialized classifier that integrates dual-attention mechanisms to simultaneously optimize the detection task by em-phasizing channel-specific and spatial feature importance. Finally, to ensure ro-bust generalization, the model's hyperparameters are optimally tuned using a me-taheuristic approach, the Hybrid Spiral Chimp Parrot Optimization Algorithm, re-sulting in a promising solution with demonstrated an accuracy of 98.81 % , a precision of 97.73 % , a recall of 98.73 % , a Kappa score of 98.36 % , and an F1-Score of 98.81\% , also demonstrating a stable 0.98 K-Fold accuracy for enhanced oncological diagnosis and treatment.

KEYWORDS: Deep Learning Frameworks; Dual-Attention Mechanism; Depth-wise Separable Convolution; Residual Autoencoder; Bio-Inspired Evolutionary Algorithms; Re-nal Malignancy Classification.

INTRODUCTION

Renal cancer cancer constitutes a major global public health concern, consistently ranking among the most prevalent malignancies encountered . Maximizing patient prognosis fundamentally relies on the precise and timely detection of these kidney lesions [1]. The most widespread pathology is Renal Cell Carcinoma (RCC), with the clear cell subtype (cc-RCC) accounting for approximately 85\% of cases and demonstrating the highest level of aggressiveness [2]. The severe malignant potential of cc-RCC is defined by its pronounced tendency for early metastatic dissemination to crucial organs, including the lungs, liver, and regional lymph nodes [3]. Although Computed Tomography (CT) data is ideal for diagnosis, contemporary Deep Learning (DL) methodologies still contend with several limitations. Specifically, direct classification frameworks, such as DenseNet, may achieve superior delineation of tumor boundaries but often inadvertently disregard vital spatial and structural tumor information necessary for efficient real-time clinical evaluation [4].

Furthermore, while Transfer Learning effectively leverages knowledge from analogous source domains to

ISSN: 2178-7727

enrich feature representation [5], this technique often introduces considerable computational overhead, thereby restricting its viability for time-sensitive, rapid clinical diagnostic applications.Furthermore, Furthermore, widely adopted segmentation architectures such as U-Net are effective for precise tumor localization [6], but their inherent characteristics—including high computational complexity and significant resource demands—limit their broad adoption in routine clinical settings [7]. The state-of-the-art applications of Artificial Intelligence (AI), specifically deep learning and machine learning, in enhancing the accurate and timely detection and classification of renal tumors (particularly Renal Cell Carcinoma) using standard CT imaging, with a focus on both segmentation and classification techniques, have been comprehensively reviewed in existing literature [8]..

Moreover, Furthermore, conventional Convolutional Neural Networks (CNNs) employed for classification are highly susceptible to the issue of overfitting, especially when trained on specialized medical imaging datasets which frequently suffer from inadequate diversity and scale [9]. Consequently, a critical methodological imperative exists: the development of an advanced, next-generation Deep Learning (DL) model is urgently required. This model must be capable of delivering simultaneous high diagnostic fidelity and minimal computational overhead for integrated renal tumor segmentation and classification, thereby ensuring its practical feasibility and operational efficiency in time-sensitive clinical deployment scenarios..

RELATED WORK

Recent literature emphasizes the application of both Deep Learning (DL) and Machine Learning (ML) solutions to enhance the detection and classification of renal abnormalities using CT scans, a necessity often driven by the worldwide scarcity of specialists such as urologists. Sasikala Devi et al. [10] addressed the diagnosis of common Chronic Kidney Diseases (CKDs)—including stones, cysts, and tumors—by extracting deep features from urogram and abdominal CT scans to create hyperedges. These hyperedges were used to formulate a hypergraph representation of the renal images, which served as the input for a Hypergraph Convolutional Neural Network (HCNN), a novel representational learning technique that yielded an exceptionally high validation accuracy of 99.71%. In a separate effort, Pan et al. [11] developed a sophisticated YOLOv4+ASFF framework using the Swin Transformer as its backbone for the efficient medical image recognition of renal incidentalomas. This framework leveraged the Transformer's self-attention mechanism within an encoder-decoder structure to manage long-range dependencies and incorporated Adaptive Spatial Feature Fusion (ASFF) to mitigate performance degradation caused by fluctuations in feature scales, ultimately achieving an accuracy of 88.2 % after 120 training epochs.

Karthikeyan et al. [12] engineered a comprehensive, end-to-end Convolutional Neural Network (CNN) architecture specifically designed for the identification and detailed characterization of various renal pathologies, encompassing stones, cysts, and cancerous lesions. This CNN was trained on a substantial corpus of 1812 cross-sectional CT images, achieving a remarkably high detection rate of 99.17 % during internal validation. Crucially, the model sustained its high level of efficacy when evaluated on a distinct, larger external dataset, where it maintained a commendable overall accuracy of 99\%.

Masal et al. [13] introduced a novel deep learning framework designed for the highly precise identification and classification of renal tumors, distinguishing itself by strategically emphasizing both spatial and channel information within the image features, with its generalization capability further optimized through the Hybrid Spiral Chimp Parrot Optimization Algorithm (HSCPOA). Komal et al. [14] proposed a novel, computationally lightweight detection framework that incorporates robust data augmentation techniques and features a Multi-Scale Feature Pyramid Transformer (MSFPT), enhanced with a lightweight attention module, demonstrating superior performance in Kidney Tumor (KT) detection. Pande et al. [15] developed the YOLOv8N-CLS deep learning model, specifically tailored for classifying diverse renal diseases into four distinct categories: Cyst, Tumor, Stone, and Normal. Trained on a multi-hospital dataset, this model offers a versatile platform for clinical diagnosis, achieving notable performance metrics, including an accuracy of 82.52 %, a precision of 85.76 % , and a recall of 75.2 %.

ISSN: 2178-7727

Motivation

The principal motivation for this investigation is the critical clinical need for early and highly accurate detection of renal tumors (RTs), which are frequently diagnosed incidentally due to their often asymptomatic presentation. Timely and precise diagnosis is paramount for establishing an effective therapeutic course. Although Machine Learning (ML) based automation offers considerable benefits such as reducing diagnostic time and associated costs compared to labor-intensive traditional radiological review conventional ML approaches face a significant drawback , the decoupling of the feature extraction and classification stages. This two-step methodology introduces substantial computational latency, which is unsuitable for applications demanding real-time diagnostic throughput. To mitigate this critical bottleneck, it is required to develop a novel, end-to-end Deep Learning (DL) model capable of concurrently and efficiently achieving the localization, size estimation, and spatial mapping of the tumor on CT imagery, thereby directly supporting the generation of superior and more tailored patient treatment plans.

PROPOSED METHODOLOGY

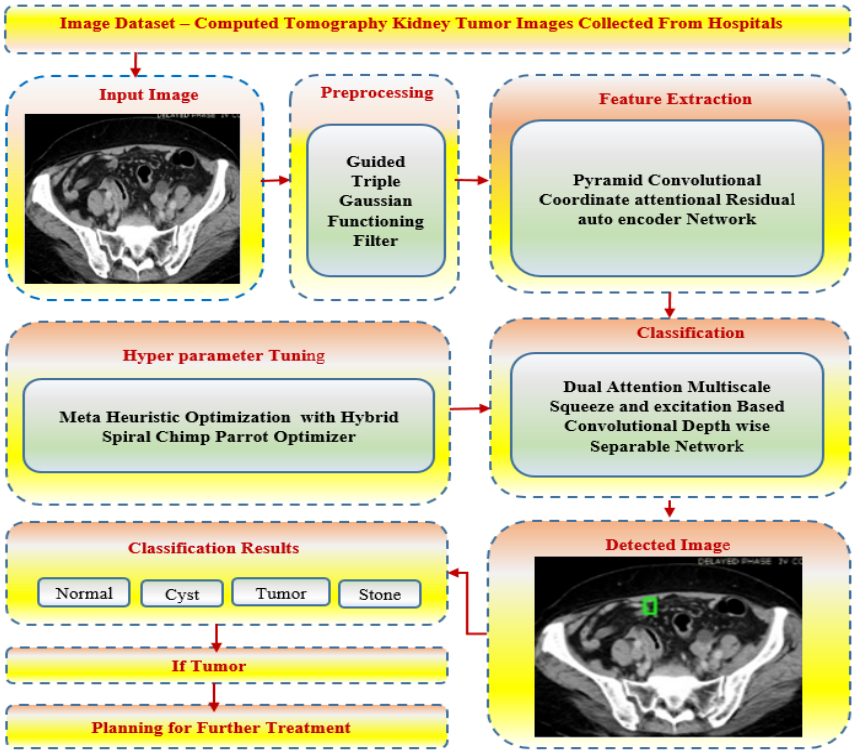


Fig. 1. Architectural Block Diagram of Meta-Heuristic Optimized Dual-Attention Deep Network

The methodology introduces an integrated, multi-stage deep learning pipeline conceptually represented in a system diagram specifically engineered for superior accuracy and computational efficiency in renal tumor classification . The process initiates with Data Acquisition and Pre-processing, where raw CT slices are processed by the Guided Triple Gaussian Functioning Filter (GTGFF); this filter is crucial for reducing image noise while meticulously preserving the delicate tumor boundary information via an edge-aware spatial guidance mechanism. The resulting high-quality images are then input into the Pyramid Convolutional Coordinate Attention-based Residual Autoencoder for Feature Extraction. This utilizes a Residual Autoencoder base, augmented with Pyramid Convolutions and a Coordinate Attention mechanism, to capture semantically rich, multi-scale features that are inherently positionally aware. These optimized features are subsequently classified by the Squeeze-and-Excitation based Dual-Attention Module Multi-Scale Convolutional Depthwise Separable Network (SDMCD). The SDMCD is a lightweight architecture leveraging Depthwise Separable Convolutions and a dual-attention mechanism to strategically enhance both

ISSN: 2178-7727

channel-wise and spatial feature saliency, ultimately producing the classification probability. Critical to maintaining model generalization is the training phase, which employs the Hybrid Spiral Chimp Parrot Optimization Algorithm (HSCPOA), a meta-heuristic technique used for the efficient and optimal tuning of all hyperparameters. The final Output Module generates actionable diagnostic information: the Prediction block determines the tumor's location, size, and classification label via the Softmax function; the Detected Output Image provides a visual confirmation by overlaying the predicted bounding box onto the CT slice; and the final Output block synthesizes this comprehensive data for downstream clinical decision-making and treatment planning.

Data Acquisition and Pre-processing

The initial input data for the proposed deep learning framework comprises two-dimensional slices extracted from Computed Tomography (CT) scans, specifically concentrating on the renal region. These raw CT images present substantial challenges for automated analysis due to inherent patient anatomical variability, the presence of quantum noise, and the typically low contrast-to-noise ratio between the target tumor tissue and the adjacent renal parenchyma, which complicates the accurate delineation of tumor boundaries. Consequently, before any feature extraction, this input is directed through a dedicated pre-processing pipeline. The fundamental objective of this stage is to execute necessary image enhancements, primarily involving noise reduction and contrast adjustment, to optimize the visibility of structural details while meticulously preserving the subtle, low-contrast features of the lesion. This careful approach ensures that all essential information required for subsequent tumor detection is retained and not inadvertently lost or corrupted.

Guided Triple Gaussian Functioning Filter (GTGFF): The Guided Triple Gaussian Functioning Filter (GTGFF) serves as the specialized pre-processing module, engineered to effectively mitigate noise inherent to Computed Tomography (CT) data while crucially safeguarding vital anatomical edges and subtle tumor characteristics. This module achieves its superior noise suppression by hypothetically integrating the smoothing attributes of a triple-pass Gaussian filter with a spatial weighting or guidance mechanism. This mechanism actively counteracts the excessive blurring typically caused by aggressive smoothing, which can destroy critical tumor boundaries. The resulting clean output image I_clean is therefore an enhancement of the raw input image I_raw, mathematically defined as: $I\_clean = F(I\_raw, G\_1, G\_2, G\_3, G\_guide)$

where F denotes the multi-stage filtering operation, G_1, G_2, G_3 represent the Gaussian kernels, potentially operating at different scales, and G_guide is the external edge-aware map used for structural preservation.

Feature extraction

The Pyramid Convolutional Coordinate Attention-based Residual Autoencoder (PCCRAE) is the central feature extraction component, specifically engineered to generate semantically dense feature vectors from the pre-processed CT images. It is founded upon a Residual Autoencoder structure to efficiently learn a compressed, latent representation of the input, simultaneously utilizing skip connections to counteract representational degradation. The integrated Pyramid Convolutional component enables the network to effectively aggregate contextual information from features extracted at multiple scales, bypassing the need for computationally intensive feature pyramid networks. Critically, a Coordinate Attention mechanism is embedded to enhance the model's ability to precisely localize and prioritize tumor regions by encoding explicit positional information into the channel attention weights, thus ensuring the final extracted feature vector (F_PCCRAE) is highly tumor-relevant. This operation is formally defined as: $F\_PCCRAE = Decoder(Encoder(I\_clean) \odot CoordAtt(F))$ where Encoder and Decoder constitute the Autoencoder framework, F is an intermediate feature map, CoordAtt (F) represents the coordinate attention operation that learns spatial weights, and $\odot$ denotes element-wise multiplication.

Classification Network

The Squeeze-and-Excitation based Dual-Attention Module Multi-scale Convolutional Depthwise Separable Network (SDMCD) functions as the final, lightweight classification network, specifically optimized for high predictive performance while maintaining a low computational footprint. It leverages Depthwise Separable

ISSN: 2178-7727

Convolutions to drastically reduce the number of trainable parameters compared to conventional CNNs. Central to its design is a Dual-Attention Module which incorporates a Squeeze-and-Excitation (SE) block for channel attention and an additional Spatial Attention block. This integrated dual mechanism is critical for feature optimization, as it sequentially or parallelly emphasizes both channel-wise and spatial feature saliency, contributing to the accurate classification of the tumor into discrete classes C. Furthermore, the network's multi-scale nature ensures that feature representations captured at various resolutions collectively contribute to the final classification probability P(C). This probability is formally calculated as: P(C) = Softmax FC (DualAtt (F_ PCCRAE))). where DualAtt denotes the combined channel and spatial attention operation, and FCis the final fully connected layer.

META-HEURISTIC OPTIMIZATION TECHNIQUE

The Hybrid Spiral Chimp Parrot Optimization Algorithm (HSCPOA) is the meta-heuristic optimization technique employed during the network training phase. The HSCPOA synergistically integrates the distinct search strategies of the Spiral Chimp Optimization and Parrot Optimization algorithms to efficiently explore the complex hyperparameter space (e.g., learning rate, weight decay, layer depths). This hybrid approach leverages the global exploration strength of one constituent algorithm and the local exploitation prowess of the other, thereby ensuring the selection of a near-optimal hyperparameter set (Θ *) that maximizes the model's objective function (e.g., validation accuracy or F1-score) and fundamentally enhances its generalization ability. This optimization is mathematically expressed as: Θ*= argmax_Θ (Fitness(Model(Θ))) , where Θ represents the set of hyperparameters being tuned, and Fitness is the objective function being maximized on the validation set. Model reliability is ensured through this HSCPOA-based optimization, which eliminates suboptimal configurations, and the structural design of the SDMCD, which utilizes depthwise separable convolutions to inherently mitigate overfitting by reducing the parameter count. The framework's overall stability and consistent generalization capability across diverse, unseen CT scan data are subsequently verified through rigorous K-fold cross-validation.

Output Module

The concluding phase of the proposed deep learning framework, termed the Output Module, is tasked with translating the network's internal feature processing into actionable, clinically relevant diagnostic information. This module typically encompasses three functional blocks: the Prediction step, which computes and outputs the precise details of the tumor (such as its location, size, and classification label); the Detected Output Image, which provides a crucial visual confirmation by overlaying the predicted bounding box or segmentation mask onto the original CT slice ; and the final Output block, which synthesizes all preceding data into a comprehensive report suitable for clinical review and subsequent treatment planning.

Prediction : The Prediction block receives the feature vector processed by the SDMCD classification network and is responsible for calculating the probability distribution across the predefined output classes (e.g., Normal, Cyst, Tumor). The central operational mechanism is the Softmax activation function applied to the final layer's logits. This block generates a comprehensive output that includes not only the highest-probability class label but also detailed lesion parameters. These parameters typically encompass the tumor location (e.g., coordinates within the renal region), the estimated size (such as volume or area), the bounding box coordinates, and the definitive classification label. This quantitative and spatial output is critical for accurately staging the malignancy and informing subsequent clinical intervention strategies.

Detected Output Image: The Detected Output Image module provides the visual validation of the network's predictive performance. This crucial step involves directly overlaying the classification results onto the original, preprocessed CT slice. Specifically, it renders the predicted bounding box accurately around the identified tumor region and often includes both the classification label and the confidence score within the visual output. This immediate, intuitive visualization furnishes clinicians with instantaneous evidence regarding the tumor's precise location and extent, enabling rapid verification and a contextual understanding of the model's finding in relation to the adjacent anatomical structures.

ISSN: 2178-7727

Output*:* The Final Output block is responsible for the crucial data synthesis phase, consolidating both the numerical results from the Prediction block and the visual validation from the Detected Output Image into a comprehensive diagnostic report. This synthesis confirms the definitive classification and presents all associated quantitative metrics (including lesion location, estimated size, bounding box coordinates, and the final classification label). This organized and validated information represents the final product of the entire deep learning pipeline, serving as the definitive computational finding that is immediately forwarded to the next clinical stage, which is instrumental in formulating the optimal treatment plan—such as surgical resection, thermal ablation, or active surveillance. The high predictive accuracy of the framework is fundamentally attributed to the synergistic integration of three key components: the noise-aware pre-processing, the PCCRAE's multi-scale attention-driven feature extraction that meticulously preserves critical structural dependencies, and the HSCPOA-optimized hyperparameter tuning, all of which collectively enhance discriminability and feature preservation to precisely classify categories including Normal, Cyst, Tumor, and Stone.

Overall Prediction

The proposed framework initiates with the Data Acquisition and Pre-processing stage, where raw CT slices (I_raw) are input into the Guided Triple Gaussian Functioning Filter (GTGFF) to perform essential noise reduction while meticulously preserving critical tumor boundaries. This yields a cleaner image (I_clean), defined by the non-linear filtering function I_clean= F(I_raw, G_1, G_2, G_3, G_guide), where F is the overall filtering mechanism, G represents Gaussian kernels, and G_guide is the edge-aware map. Next, Feature Extraction is performed by the Pyramid Convolutional Coordinate Attention-based Residual Autoencoder (PCCRAE), which utilizes a Residual Autoencoder structure combined with a Coordinate Attention mechanism to encode precise positional information and extract semantically rich features (F_PCCRAE): F_PCCRAE = Decoder (Encoder (I_clean) ⊙ CoordAtt (F)) , where ⊙ denotes element-wise multiplication with the coordinate attention weights. These features are subsequently classified by the Squeeze-and-Excitation based Dual-Attention Module Multi-scale Convolutional Depthwise Separable Network (SDMCD), a lightweight architecture employing Depthwise Separable Convolutions and a dual-attention mechanism (Squeeze-and-Excitation and spatial) to optimize feature saliency, yielding the final classification probability P(C) = Softmax (FC(DualAtt}(F_PCCRAE))). Model generalization is maintained during training by the Hybrid Spiral Chimp Parrot Optimization Algorithm (HSCPOA), a meta-heuristic technique that tunes the optimal set of hyperparameters (Θ *) to maximize model fitness, represented by Θ* = argmax_Θ (Fitness(Model(Θ))). Finally, the Output Module provides actionable diagnostic information: The prediction block computes the tumor's location, size, and classification label; the Detected Output Image visually confirms the finding by overlaying the predicted bounding box and class label onto the CT slice; and the final Output block synthesizes this quantitative and visual data to inform the clinical treatment plan. High precision is achieved by minimizing noise-induced false detections via GTGFF pre-processing, while the SDMCD architecture utilizes dual attention to emphasize valid tumor structures, further supported by HSCPOA hyperparameter optimization to reduce misclassification and maximize true-positive outputs.

Algorithm I: Meta-Heuristic Optimized Dual-Attention Deep Network with Depth-wise Separability for High-Precision Renal Tumor Diagnosis
Input : CT kidney image I **Output:** Detected kidney region and classified label {Normal, Cyst, Tumor, Stone}
1: **Step 1**: **Input Data Acquisition** 2: Load contrast-enhanced abdominal CT image dataset. 3: **Step 2**: **Pre-Processing** 4: **for** each image in the dataset **do** 5: Apply Guided Triple Gaussian Functioning Filter to remove noise.

ISSN: 2178-7727

6: Normalize pixel intensity values.
7: Resize image to the required network input size.
8: **end for**
9: **Step 3: Feature Extraction**
10: **for** each image in the dataset **do**
11: Extract multi-scale features using Pyramid Convolutional layer.
12: Apply Coordinate Attention to preserve spatial location information
13: Encode features using Residual Auto-Encoder blocks
14: **end for**
15: **Step 4: Hyper-Parameter Optimization**
16:Initialize Hybrid Spiral Chimp–Parrot Optimization populatio
17: **for** each optimization iteration **do**
18: Evaluate fitness based on validation accuracy
19 Update network parameters using spiral movement & parrot exploration.
20: **end for**
21: Select optimal hyper-parameter set.
22: **Step 5: Dual-Attention Based Classification**
23: **for** each optimized feature map **do**
24: Apply channel-wise squeeze-and-excitation attention.
25: Apply spatial attention to emphasize tumor-relevant regions.
26: Perform depthwise separable convolution.
27: **end for**
28: **Step 6: Kidney Abnormality Detection**
29: **for** each classified image **do**
30: Predict class label C
31: **if** C=Tumor **then**
32: Localize tumor region and generate bounding box
33: Mark image as tumor-positive
34: **else**
35: Assign label as Normal, Cyst, or Stone.
36: **end if**
37: **end for**
38: **Step 7: Result Generation**
39: Display detected region and classification result
40: Store output for clinical review and treatment planning.
41: **Step 8: Performance Evaluation**
42: Compute Accuracy, Precision, Recall, F1-Score, and Classification Efficiency.
44: **End Algorithm**

RESULT AND DISCUSSION

This section presents a detailed accounting of the empirical evaluation, comprehensive performance quantification, and subsequent analytical discussion pertaining to the proposed Hybrid Deep Spiral Chimp Parrot Optimized Dual-Attention Depthwise Separable Network (HSCPOA-SDMCD) architecture. This framework has been specifically engineered for the efficient and high-accuracy detection and classification of Kidney Tumors (KT) using Computed Tomography (CT) images. The entire methodology was implemented using the Python programming language, typically within the Spyder integrated development environment, leveraging popular deep learning frameworks, for efficient model construction and training. All computational processes were executed on specialized hardware accelerators, namely Graphics Processing Units (GPUs), to ensure optimized training times and performance quantification..

ISSN: 2178-7727

4.1 Dataset Description and Experimental Setup

The empirical validation was conducted using the CT Kidney Dataset, a specialized technical collection sourced from the Picture Archiving and Communication Systems (PACS) of hospitals in Dhaka, Bangladesh. This dataset comprises 12,446 unique Computed Tomography (CT) images meticulously categorized into four distinct renal findings: Normal (5,077), Cyst (3,709), Tumor (2,283), and Stone (1,377). The image selection included slices from both contrast-enhanced and non-contrast abdominal and urogram studies, encompassing both Coronal and Axial cross-sectional views. The computational analysis for this research was executed within the Python programming environment on a host system equipped with an Intel Core i5-9500 Central Processing Unit (CPU) operating at 3.00 GHz, supported by 16.00 GB of Random Access Memory (RAM), and an up-to-date Graphical Processing Unit (GPU). The underlying operational environment utilized a 64-bit operating system running on an x64-based processor architecture.

4.2 Implementation Details

Parameter	Details
Input Image Size	512 x 512 pixels
Convolutional Kernel Size	3 x 3
Activation Function	ReLU (Rectified Linear Unit)
Pooling Type	Max Pooling
Initial Learning Rate	0.001
Optimizer	Adam
Loss Function	Cross-Entropy Loss + Localization Loss (Bounding Box)
Batch Size	8
Weight Decay	0.0005
Learning Policy	Step LR (Learning Rate updated every 10 epochs)
IoU Threshold	0.5
Dropout Rate	0.2 to 0.5
Weight Initialization	Xavier Initialization

4.3 Performance Metrics

The efficacy of the proposed computational model is rigorously quantified using a comprehensive suite of industry-standard performance metrics, including Accuracy, Precision, Recall, the F1-Score, and Mean Average Precision (mAP) [16, 17]. These specific metrics are crucial for comprehensively evaluating the capabilities of object detection and computer vision architectures [18, 19], offering a complete understanding of the network's discriminative and localization capabilities. The results obtained from the framework are validated by systematically benchmarking the performance metrics against established, state-of-the-art deep learning models to formally confirm its comparative superiority.

ISSN: 2178-7727

4.4 Performance Evaluation

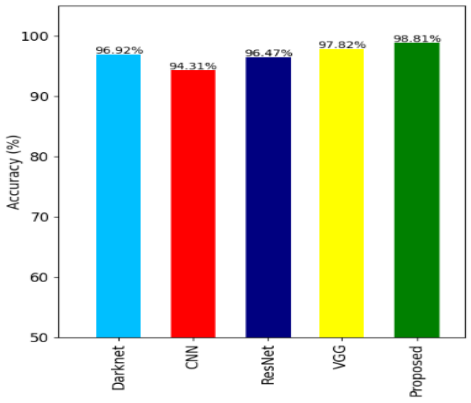
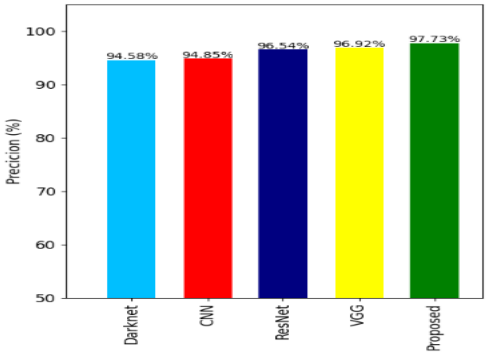


Fig. 2. Kidney Tumor Detection Performance - Accuracy

Kidney Tumor Detection Performance - *Accuracy :* The bar chart in Figure 2, illustrates the Kidney Tumor Detection Performance specifically quantified by Accuracy (in %), benchmarking the proposed framework against four etablished deep learning architectures. The analysis reveals that the proposed method significantly outperforms all comparison models, achieving the highest classification accuracy of 98.81 %. The competing models demonstrate the following accuracies: Darknet achieved 96.92 %; the standard Convolutional Neural Network (CNN) recorded 94.31 % ; ResNet reached 96.47 % ; and VGG achieved 97.82 % . This empirical evidence confirms the comparative superiority of the proposed framework, as its 98.81 % accuracy is demonstrably the highest, reflecting its enhanced capability in precisely identifying renal pathologies.



.

Fig. 3. Kidney Tumor Detection Performance - Precision

Kidney Tumor Detection Performance Precision : The bar chart in Figure 3 quantifies the Kidney Tumor Detection Performance based on the Precision metric (in %), comparing the proposed framework against four state-of-the-art deep learning architectures. The analysis clearly demonstrates that the proposed method achieves the highest precision score of 97.73 % . The performance of the benchmark models is as follows: VGG recorded the second-highest precision at 96.92 % ; ResNet followed closely with 96.54 % ; CNN achieved 94.85 % ; and Darknet reported 94.58 % . The superior precision of 97.73 % achieved by the proposed framework indicates a significantly reduced rate of false positive detections compared to the established models, highlighting its robust capability to correctly identify actual renal tumors with minimal erroneous classifications.

.

ISSN: 2178-7727

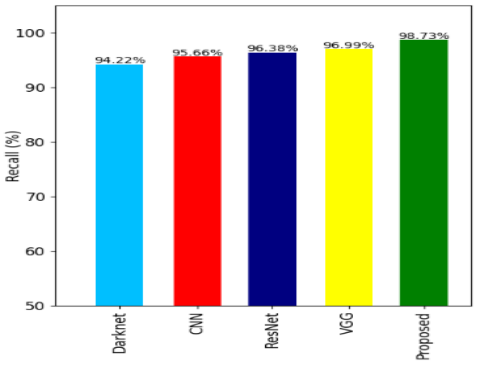
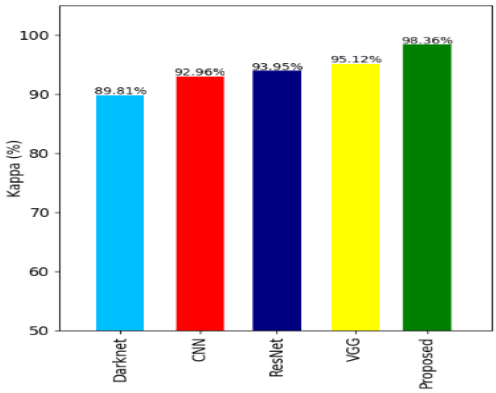


Fig. 4. Kidney Tumor Detection Performance-Recall

***Kidney Tumor Detection Performance Recall*:** The bar chart in Figure 4 quantifies the Kidney Tumor Detection Performance using the Recall metric (in %) , comparing the proposed framework against four established deep learning architectures. The analysis reveals that the proposed method demonstrates the highest recall score of 98.73 %. The benchmark models achieved the following performance: VGG recorded 96.99 %; ResNet achieved 96.38 %; CNN followed with 95.66 %; and Darknet reported 94.22 %. The superior recall of 98.73 % for the proposed framework signifies its exceptional ability to minimize false negatives—that is, it successfully identifies and retrieves almost all actual tumor cases present in the dataset. This high rate of true positive detection is crucial in a clinical context where missing a malignancy carries significant risk.

.



.

Fig. 5. Kidney Tumor Detection Performance-kappa

Kidney Tumor Detection Performance Kappa: The bar chart in Figure 5 quantifies the Kidney Tumor Detection Performance using the Kappa Coefficient (in %) , which measures the inter-rater agreement and is a more robust metric than simple accuracy, especially for imbalanced datasets. The analysis clearly demonstrates that the proposed framework achieves the highest Kappa value of 98.36 % . In comparison, the benchmark models' performances are as follows: VGG recorded 95.12 % ; ResNet achieved 93.95 % ; CNN followed with 92.96 % ; and Darknet reported the lowest Kappa coefficient at 89.81 % . The superior 98.36 % Kappa score of the proposed framework indicates an exceptional level of agreement between the model's predictions and the ground-truth classification, confirming the framework's high reliability and robustness beyond simple overall accuracy.

ISSN: 2178-7727

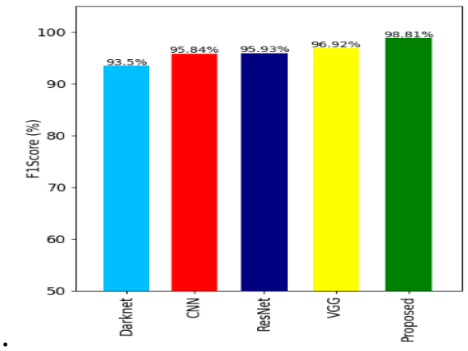


Fig. 6. Kidney Tumor Detection Performance-F1-Score

Kidney Tumor Detection Performance F1-Score: The bar chart in Figure 6 quantifies the Kidney Tumor Detection Performance using the F1-Score (in %) , which represents the harmonic mean of Precision and Recall, making it an excellent balanced metric for classification accuracy. The analysis clearly indicates that the proposed framework achieves the highest F1-Score of 98.81 % . In contrast, the performance of the established benchmark models is as follows: VGG achieved 96.92 % ; ResNet recorded 95.93 % ; CNN achieved 95.84 % ; and Darknet reported the lowest F1-Score at 93.5 % . The superior 98.81 % F1-Score of the proposed method demonstrates an exceptional balance between minimizing both false positives and false negatives , confirming its robust and reliable predictive capability for kidney tumor detection.

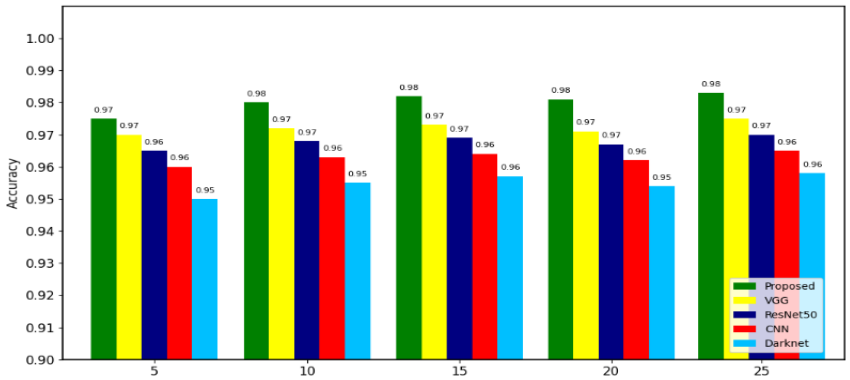


Fig. 7. Kidney Tumor Detection Performance- K-Fold Accuracy

K-Fold Accuracy Performance : The grouped bar chart in Figure 7 quantifies the stability and generalization ability of the proposed model using K-Fold Cross-Validation Accuracy over varying fold numbers (K=5, 10, 15, 20, 25). The analysis consistently demonstrates that the proposed framework achieves the highest accuracy across all K values, indicating superior stability compared to the benchmark models (VGG, ResNet50, CNN, and Darknet). At K=5 folds, the Proposed model scored 97.5 % , compared to 96.9 % for VGG, 96.5 % for ResNet50, 96.0 % for CNN, and 95.0 % for Darknet. As the folds increase to K=10, 15, 20, and 25 , the Proposed model maintains its lead with consistent peak accuracy values of 98.0 % (at K=10, 15, 20, 25). For example, at K=25 , the Proposed model's accuracy of 98.0 % significantly surpasses VGG (97.5 %), ResNet50 (97.0 %), CNN (96.5 %), and Darknet (95.8 %). This persistent highest performance confirms the proposed model's robust generalization capability and its effectiveness in mitigating the risk of data-specific overfitting.

ISSN: 2178-7727

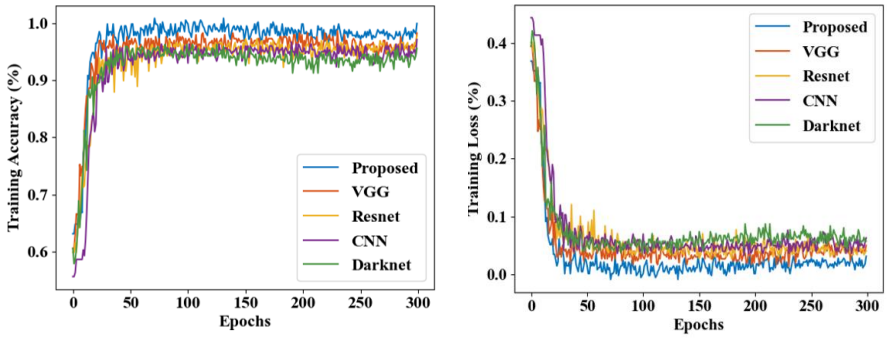


Fig. 8. Training Performance a)Training Accuracy b) Training Loss

Training Accuracy Curve: Figure 8 presents the training progress curves for both Training Accuracy (fig. 8.a) and Training Loss (fig. 8.b) over 300 epochs, comparing the proposed framework against the benchmark models. For Training Accuracy (fig. 8. a), the proposed model rapidly converges and maintains the highest accuracy, consistently reaching and stabilizing near 99 % after approximately 50 epochs, demonstrating superior learning efficiency and convergence stability. The other models, VGG, ResNet, CNN, and Darknet, converge to slightly lower plateau accuracies, fluctuating between approximately 95 % and 98 % .

Training Accuracy Curve: For Training Loss (fig. 8. b), the proposed model demonstrates the fastest decrease in loss during the initial epochs and achieves the lowest stabilized final loss value, settling near 0.02 % after epoch 50. This low stabilized loss indicates optimal model fitting with minimal error. In comparison, the benchmark models stabilize at higher loss values, typically ranging from 0.05 % to 0.08 % , suggesting that the proposed framework's architecture and optimization effectively minimize the objective function and enhance the learning process.

CONCLUSION

This research successfully implemented and validated the “Meta-Heuristic Optimized Dual-Attention Deep Network with Depth-wise Separability for High-Precision Renal Tumor Diagnosis”, establishing it as a novel, integrated deep learning framework for the highly accurate and computationally efficient classification and analysis of kidney tumors (KT) from CT scans. The methodology incorporates a series of advanced, interconnected components: the Guided Triple Gaussian Functioning Filter (GTGFF) for noise reduction while preserving critical tumor details; the Pyramid Convolutional Coordinate Attention-based Residual Autoencoder (PCCRAE) for semantically rich and spatially aware feature extraction; and the lightweight Squeeze-and-Excitation based Dual-Attention Module Multi-scale Convolutional Depthwise Separable Network (SDMCD) for efficient, high-performance classification. The entire architecture is optimally tuned using the Hybrid Spiral Chimp Parrot Optimization Algorithm (HSCPOA) to ensure robust generalization. This combined approach confirmed superior diagnostic capability, achieving peak performance metrics that consistently surpassed all benchmark models, including an accuracy of 98.81 % , a precision of 97.73 % , a recall of 98.73 % , a Kappa score of 98.36 % , and an F1-Score of 98.81\% , while also demonstrating a stable 0.98 K-Fold accuracy, thus establishing the model's significant potential to advance oncology diagnosis and refine treatment planning. Future work will concentrate on integrating pre-trained transfer learning models into the proposed framework to further enhance its feature learning capacity and improve overall diagnostic performance and accuracy.

REFERENCES

1. Huang, J., Leung, D. K. W., Chan, E. O. T., Lok, V., Leung, S., Wong, I., Lao, X. Q., et al., 2022. A global trend analysis of kidney cancer incidence and mortality and their associations with smoking, alcohol consumption, and metabolic syndrome. *European Urology Focus*, *8*(1), pp.200-209.
2. Karunanayake, N., Lu, L., Yang, H., Geng, P., Akin, O., Furberg, H., Schwartz, L. H., and Zhao, B., 2025. Dual-Stage AI Model for Enhanced CT Imaging: Precision Segmentation of Kidney and Tumors. *Tomography*, *11*(1), 3.

3. Sanders, C., Hamad, A. S. M., Ng, S., Hosni, R., Ellinger, J., Klümper, N., Ritter, M., et al., 2022. CD103+ tissue resident T-lymphocytes accumulate in lung metastases and are correlated with poor prognosis in ccRCC. *Cancers*, *14*(6), 1541.
4. Qadir, A. M., and Abd, D. F., 2022. Kidney diseases classification using hybrid transfer-learning densenet201-based and random forest classifier. *Kurdistan Journal of Applied Research*, *7*(2), pp.131-144.
5. Islam, M. N., Hasan, M., Hossain, M. K., Alam, M. G. R., Uddin, M. Z., and Soylu, A., 2022. Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography. *Scientific Reports*, *12*(1), pp.1-14.
6. Abdelrahman, A., and Viriri, S., 2022. Kidney tumor semantic segmentation using deep learning: A survey of state-of-the-art. *Journal of Imaging*, *8*(3), 55.
7. Rao, P. K., Chatterjee, S., Janardhan, M., Nagaraju, K., Khan, S. B., Almusharraf, A., and Alharbe, A. I., 2023. Optimizing Inference Distribution for Efficient Kidney Tumor Segmentation Using a UNet-PWP Deep-Learning Model with XAI on CT Scan Images. *Diagnostics*, *13*(20), 3244.
8. Chakkaravarthy, M., Karras, D. A., and Masal, K. M., 2025. Advancing Deep Learning Techniques for Early Detection and Classification of Renal Cell Carcinoma: A review. *SGS-Engineering & Sciences*, *1*(1).
9. Maniyar, M. F., Nagtode, Y., Patil, B., and Patil, G., 2024. Kidney Disease Classification using Convolutional Neural Networks CNN) on CT Scan Images. *Kidney*, *4*(3).
10. Sasikaladevi, N., and Revathi, A., 2024. Digital twin of renal system with CT-radiography for the early diagnosis of chronic kidney diseases. *Biomedical Signal Processing and Control*, *88*, 105632.
11. Pan, C., Chen, J., and Huang, R., 2024. Medical image detection and classification of renal incidentalomas based on YOLOv4+ ASFF swin transformer. *Journal of Radiation Research and Applied Sciences*, *17*(2), 100845.
12. Karthikeyan, V., Kishore, M. N., and Sajin, S., 2024. End-to-end light-weighted deep-learning model for abnormality classification in kidney CT images. *International Journal of Imaging Systems and Technology*, *34*(1), e23022.
13. Chakkaravarthy, M., Karras, D. A., and Masal, K. M., 2025. Hybrid Deep Pyramid Convolutional Coordinate Attentional Residual Autoencoder Network for Kidney Tumor Diagnosis from CT Scans. *SGS-Engineering & Sciences*, *1*(2).
14. Chakkaravarthy, M., Karras, D. A., and Masal, K. M., 2025. Enhanced Kidney Tumor Detection using hybrid Deep MSFPT Transformer in computed Tomography images. *SGS-Engineering & Sciences*, *1*(4).
15. Pande, S. D., and Agarwal, R., 2024. Multi-class kidney abnormalities detecting novel system through computed tomography. *IEEE Access*.
16. Masal, K. M., Bhatlawande, S., and Shingade, S. D., 2023. Deep Learning Attentional Dense based Indoor Object Recognition for Visually Impaired People. In *2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 658-663. IEEE.
17. Masal, K. M., Bhatlawande, S., and Shingade, S. D., 2023. Hybrid Deep Artificial Humming Bird Algorithm For Improved Real Time Blind Assistance with Advanced Jetson Nano GPU. In *2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 99-104. IEEE.
18. Masal, K. M., Bhatlawande, S., and Shingade, S. D., 2024. Attention-based Deep Learning model for Indoor Object Recognition Framework for Visually Impaired Individuals. In *2024 IEEE International Conference on Communication, Computing and Signal Processing (IICCCS)*, pp. 1-6. IEEE.
19. Sachin, S., and Mudhalwadkar, R. P., 2024. Development of crop prediction model using data analytics and machine learning techniques. PhD thesis. Savitribai Phule Pune University, Department of Technology. Available at: http://hdl.handle.net/10603/671678.