TransAugNet: Transformer-Aware Cyclic Augmentation for Biomedical Image Analysis

Dr. S. N. Tirumala Rao[1], Nimmala Ashok[2], Chennareddy Sudheer Reddy[3], Yamarthy Venkata Krishna[4],
Yerraganti KrishnaBhargavi[5], Maddala Seetha[6], Dr.Sireesha Moturi[7]
[1,2,3,4,7]Department of Computer Science and Engineering,
Narasaraopeta Engineering College (Autonomous), Narasaraopet,
Palnadu District,Andhra Pradesh
[5]Department of CSE-DS,GRIET,Hyderabad
[6]Department of CSE, G. Narayanamma Institute of Technology and Science
[5] kittu.bhargavi@gmail.com, [6] maddala.seetha@gnits.ac.in,
[1]nagatirumalarao@gmail.com, [2]ashoknimmala478@gmail.com, [4]krishnayamarthi34@gmail.com,
[3]sudheerreddychennareddy1201@gmail.com, [7]sireeshamoturi@gmail.com

***Abstract*—Accurate identification of medical conditions is an important aspect of disease diagnosis. Over time, the doctors themselves need to analyze the disease of the patient. Doctors might not be able to detect the exact disease quite often. This leads to major damage to the patient, like delaying of diagnosis of the disease. In the worst case, this may lead to life-threatening conditions. To overcome these scenarios a deep learning technique named TransAugNet was introduced for accurate monitoring of disease through medical images. In recent times, due to advancements in AI, we can detect the disease accurately by using these deep learning techniques. It uses an approach of Cyclic Augmentation through the ResNet3D-50 model for precise monitoring and helps in diagnosing the disease. This ResNet3D-50 model mainly converts the three-dimensional image to a stack of two-dimensional image slices. These slices are being worked together and help in the prediction of the disease. By using this deep learning technique, the response time will be low, and it supports the doctor in identifying the disease faster. The response time was much lower when compared to older methods, which were introduced earlier. Our model uses Transformer-Aware Cyclic Augmentation, combining attention with cyclic intensity to better understand biomedical images and improve classification accuracy.**

***Index Terms*—Cyclic Augmentation, deep learning, TransAugNet, ResNet3D-50, disease diagnosing**

I. Introduction

These days, Medical imaging is very useful to diagnose the disease of a patient. Generally, medical imaging mainly includes images like X-ray images and any other resonance medical images. In earlier days, the analysis of medical images was generally conducted by specialists in that specific disease. The accuracy varies whenever the disease is in the early stage, and it will be difficult for human expertise. This may lead to a delay in the diagnosis of the disease and may mislead the patient's treatment plan, which is life-threatening. To overcome this delay, we need to automate the diagnosis of the disease.

A. Early detection vs Late detection

From [1] the table I shows the advantages of early detection of diseases are achieved. In many cases, the detection will be useful for dangerous diseases like brain tumors [2] too. The benefit of early detection of the disease will increase the survival rate in comparison to late detection of the disease, where late identification of the disease will lead to increased mortality and cause severe loss. Early Detection of diseases faces problems due to complex imagery, lack of data, multiple variations in a single disease, complex identification, and working with 3D images and models that require detailed examination and training on huge data. To address these problems, we came up with a deep learning model, ResNet3D-50. In this research, cyclic augmentation has been applied, where we can improve accuracy efficiently and accurately predict the disease.

TABLE I: Comparison of Early and Later Identification Rates for Diseases

Disease Name	Early Identification (%)	Later Identification (%)
Adrenal Masses	70–80	5–15
Vessel Diseases	50–70	40–50

The three-dimensional images are stacked up into 2D slices. These slices are trained together into cross-sectional slices that represent the full volumetric structure. By using this model, clinicians can accurately work through the identification of the diseases. . The fig 1 mainly describes the average disease identification response time by the human expertise and by AI-related expertise [2].

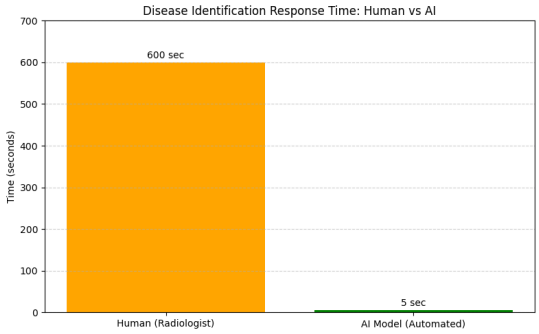

Fig. 1: Average disease identification response time

We are using deep learning techniques to overcome these challenges. We propose a deep learning solution that makes use of ResNet3D-50 with some efficient preprocessing steps to deliver this project efficient disease prediction and classification, leading to a minimal mortality rate. In the current research, we have decided to opt for 3D datasets of the disease, like Adrenal and Vessel high-quality imagery, and trained over 30 epochs.

Accurate disease prediction tasks used to produce minimal results with traditional methods. However, in the recent past, due to new approaches, these tasks have been experiencing drastic improvements over the last couple of years, primarily due to advancements in Deep Learning, especially techniques associated with imagery data. While past methods only use static cyclic augmentation, our approach brings transformer attention into the ResNet3D-50 backbone. This helps the model understand spatial patterns in the images much better, improving diagnostic accuracy.

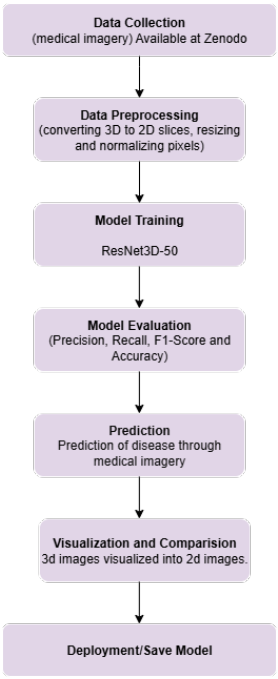

Fig. 2: Work flow

The fig 2 mainly describes the workflow of the model. It follows the steps like collection, preprocessing, training, evaluation, prediction, comparison, and deployment of the model

II. Literature Review

Ahmeed Suliman Farham. [3] came up with PRCnet, which is specially trained for brain MRI images. In this research, a method-oriented combination of MRI was used to augment brain MRI images. But this model sometimes leads to low data diversity, where the model was less trained, leading to incorrect results D. Sun. [4] used the HSMix model for mixing of hard and soft data augmentation for medical image segmentation. This is not entirely preprocessed and also lacks transformer-based segmentations. It leads to an incomplete evaluation of the model. Yulin Wang. [5] Introduced a TUM-Syn model that was capable of generating brain MRI specified by textual imaging metadata from routinely acquired scans. It also focuses on 3d imagery, but the main drawback was that the synthetic artifacts were of low quality. Hafsa Laçi1. [6] presented a deep learning method for classifying medical images such as X-ray, MRI, and ultrasonic sounds. It uses models like VGG, cGAN to classify the medical imagery. But it lacks multi-modal analysis and also consists of insufficient data to train the model. Zahid ur Rahman. [7] came up with DuCo-Net, a Dual-Contrastive Learning Network for Medical Report Retrieval Leveraging Enhanced Encoders and Augmentations. It uses DenseNet121 and BioBERTU-Net models. But it lacks contrastive learning in medical report generation. MIN-JUN KIM. [8] proposed a deep learning technique using models like EfficientNetV2, VIT-B/16, and Mixer-B/16. It mainly focuses on 2D datasets but is limited to three-dimensional data. VALENTINA SÁNCHEZ [9] came up with the technique GraphRNN, EncoderForest, and these are mainly used to produce graphs, time-series for the fMRI images. But it was limited for the exploration of multi–model data augmentation. Simon D. Westfechtel1. [10] proposed a deep learning architecture named nnU-Net. It was mainly used for empirical analysis of augmentation on model robustness against MRI motion artifacts. But the generalization to other tasks was unclear. Wanying Li. [11] came up with a technique to improve Chinese medical accuracy for microscopic images, which faced challenges in detecting rare cases and lacked validation using large-scale image datasets. Panagiotis Alimisis [12] introduced a deep learning model, DDPM, which provides a comprehensive review of diffusion models for image augmentation. But it lacks specialized surveys on DMs in augmentation

Hence, in this model, we have used ResNet3D-50 as the backbone for the disease classification, which also helps in disease diagnosis.

III. Methodology

This model was trained on high-resolution medical imagery, which the dataset was obtained from MedMNIST. has multiple. It consists of 12 two-dimensional subsets and

6 three-dimensional subsets. The data set of this research paper was available at https://zenodo.org/records/10519652 [13] and consists of over 9000 high-resolution images. High-resolution imagery is perfect for identifying disease and is used to diagnose the disease. The 3D datasets are of adrenalMNIST, vesselMNIST, fractureMNIST, noduleMNIST, synapseMNIST, and organMNIST. High-resolution imagery enables the model to distinguish between diseases, which helps the model to improve it's accuracy eventually. The data set not only includes a large number of images it also has have lot of diversity and variability of data. The data set consists of high-resolution imagery that identifies the disease from different three-dimensional subsets, which enables the model to generalize more and perform well on real-life imagery as well, which is both essential and commendable for practical deployment.

A. Sample images from dataset

Fig. 3: Sample images obtained from the MedMNIST dataset

The fig 3 refers to the images that are obtained from the MedMNIST dataset of disease Adrenal and vessel.

B. Data preprocessing

Data set consists of high-resolution medical imagery of different sizes, as deep learning models expect it's inputs to have a fixed shape to run filters on the images. The following preprocessing steps are performed before model training.

- The 3D images are converted into single-channel grayscale images, working with the model EfficientNetV2-M.
- All the images are resized to 224 x 224 to fulfill the input requirements.
- The pixel values are normalized using ImageNet standards.

Mean = [0.485, 0.456, 0.406]
Std = [0.229, 0.224, 0.224]

C. Augmentation techniques

1. Original (Without Augmentation): In this phase, no augmentation was applied, where general preprocessing like resizing, converting into grey scale, and normalization

2. RandAugment: In this phase, random transformations like random rotations, random horizontal, brightness/contrast changes. RandAugmentation is mainly used to improve generalization and reduce overfitting.

$$I' = a_{i_N}^{(M)} \circ a_{i_{N-1}}^{(M)} \circ \cdots \circ a_{i_1}^{(M)}(I)$$

- $a_{i_N}^{(M)}$: This represents the M-th transformation applied at the index i_N.
- $\circ$: denotes the function composition.

3. Cyclic Augmentation: In this phase, the augmentations are fixed rounds. By using the Cyclic Augmentation, we get better control over the data sets. The augmentation will get applied In a cyclic manner. Firstly, low-intensity augmentations will be performed, and then high-intensity augmentations will be performed, and this process will be continued in a cyclic manner to control the performance of the model. It generalizes the data and avoids model overfitting.

$$m(t) = m_{\min} + (m_{\max} - m_{\min}) \cdot \frac{1 + \sin\left(\frac{2\pi t}{T}\right)}{2}$$

- m(t) represents how strong the augmentation is at time t.
- $m_{\min}$ = minimum value of magnitude
- $m_{\max}$ = maximum value of magnitude

Generally, cyclic augmentation mainly varies the data by making changes through rotating the image or making variations in the perspective of brightness or contrast.

4. Cyclic Augmentation + ILR: In this phase, the cyclic augmentation will be done along with Intensity Level Augmentation. ILR mainly focuses on the change in the intensity level to train the model to train through different conditions by adjusting the brightness/contrast.

$$\eta(t) = \frac{\eta_0}{1 + \lambda t}$$

- η_0 is the initial learning rate.
- λ is the decay constant.

Fig. 4: Before and After preprocessing

The fig 4 refers to the image before and after preprocessing. In this three-dimensional image was sliced into multiple parts of two-dimensional images, where the two-dimensional images

are individual cross-sectional slices that represent the full volumetric structure.

D. Model Architecture Overview

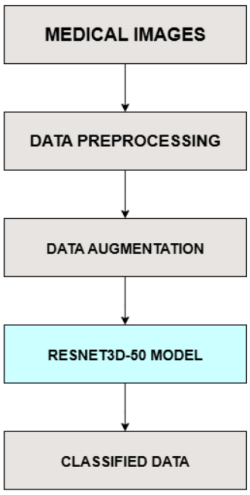

Fig. 5: Model architecture

The workflow of the proposed model begins with medical images as input. These images firstly undergo preprocessing steps like Grayscale conversion, Normalization, and Resizing, then those preprocessed images will enter the phase of data Augmentation, where the images are transformed based on the different augmentation techniques. Next, it enters the phase of the ResNet3D-50 model, which extracts the volumetric features from the images, and finally, the output will be the classified image.

E. Detailed architecture of ResNet3d-50

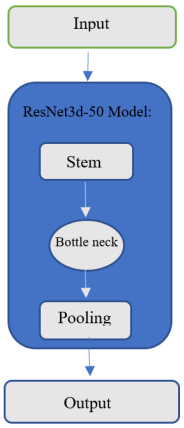

Fig. 6: ResNet3D-50 architecture overview

The fig 6 mainly describes how the model works. Firstly, the input data from the dataset is a three-dimensional image volume. It is a grayscale image, meaning it has only one channel. The volume consists of 64 slices stacked together, where each slice is like a single 2D image. The volumetric data is then used as the input to the ResNet3d-50 model for the classification tasks. Whenever you call the resnet50() function in your code, you are actually using a predefined 3D version provided by the MONAI library. MONAI internally builds a 3D variant, often called ResNet3D-50, by converting all 2D operations like Conv2D and MaxPool2D into their 3D counterparts like Conv3D and MaxPool3D. Now let's dive into a more detailed low-level design of ResNet3d-50 Architecture

Input layer: The input of our model consists of coloured images with dimensions 256X256X3, where 3 represents the RGB colour channels.

Stem layer: The stem layer is the first stage of the model. The fig 7 mainly shows the flow of stem layer. It mainly consists of:

- Conv3D (7X7X7): This layer uses 3D filters of size 7 X 7 X 7 to look at the image volume through chunks.
- BatchNorm3D: After that, we normalize the values so that learning becomes smoother and faster. It mainly steps in to even out the data, which makes the model learned more clearly and precisely.
- ReLU: Adds non-linearity, meaning it helps the model understand complex things by turning off negative values. It mainly kicks in to add more flexibility. It mainly focuses on important features and blocks unnecessary things.
- MaxPool3D: It reduces the volume size while keeping the most important features, like zooming out to see the bigger picture.

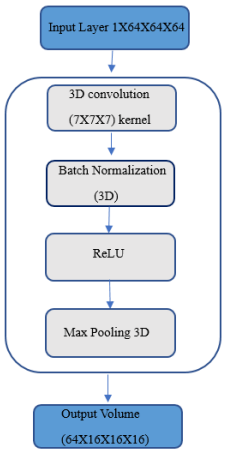

Fig. 7: Stem layer architecture

F. ResNet3D-50 encoder

Stage 1: First Convolution and Activation (conv1_relu) The initial stage of ResNet3D-50 begins with the raw 3D input volume and uses a 3D convolution with a fairly large kernel. A ReLU activation, which adds non-linearity so that the network can learn intricate patterns. This process decreases the spatial size by a small margin but enhances the number of channels to set up deeper feature extraction.

Stage 2: Residual Block Group 1 (conv2_block3_out): During this phase, the network employs a set of residual blocks—residual blocks enable the model to learn identity mappings, and it is simpler to train extremely deep networks without information loss. There are several 3D convolution layers with shortcut connections skipping over them in each residual block, which prevents vanishing gradient issues.

Stage 3: Residual Block Group 2 (conv3_block4_out): This is the third stage, features by viewing larger spatial contexts in the 3D data. This allows the model to recognize more abstract and complicated patterns, like vessel shapes or pathological patterns, in the volumetric images.

Stage 4: Residual Block Group 3 (conv4_block6_out): This is the fourth stage. Efficient gradient flow is sustained through residual connections, and the model strikes a balance between preserving crucial spatial information and abridging the input information.

Stage 5: Residual Block Group 4 (conv5_block3_out): The last encoder stage processes the highest abstracted features with a limited spatial extent but extremely rich feature representation. In this stage, the network maximally captures the sophisticated 3D patterns The output of this stage is used as the foundation for the classification head or any decoder stages, should it be applied to segmentation.

G. Bottleneck

Bottleneck is the final and deepest powerful layer of the entire network, with the processing of the encoder's final output, whose spatial size is (8×8×2048), which refines it into a more compact but highly informative form. Unlike segmentation models, ResNet3D-50 does not use a decoder. Instead, after this bottleneck, it uses global average pooling. This implies that the model doesn't attempt to recreate the original input but instead tries to comprehend its innermost semantic patterns.

H. Global Average Pooling

In our vessel disease classification model using ResNet3D-50, Global Average Pooling (GAP) plays a key role in converting the rich but compressed 3D feature maps into a simple, meaningful summary. After the encoder and bottleneck compress the volumetric vessel images into abstract feature representations, GAP takes each feature channel and calculates its average value across the entire spatial volume. This effectively summarizes complex 3D vessel characteristics into a concise vector that still captures important patterns relevant to disease presence.

IV. Model Parameters

The table II defines the model parameters, which consist of the dimensions, classes, channels, size, loss, techniques, and batches present in the model.

TABLE II: Comparison of Early and Later Identification Rates for Diseases

Parameter	**Value**
Input shape	$(1 \times 64 \times 64 \times 64)$
Learning Rate	0.0001
Optimizer	Adam
Loss	Cross Entropy
Epochs	10
Batch Size	16
Spatial Dimension	3
Output Classes	2
Input Channels	1
Smoothing Techniques	Gaussian Blur

A. Mathematical notations for ResNet3S-50 architecture

1. Input: We have N examples of 3D medical images. The size of each image is 64×64×64 and it has 1 channel (similar to grayscale). The pixel values are normalized to 0 and 1.

$$X \in \mathbb{R}^{N\times 1\times 64\times 64\times 64}, \quad X \in [0,1]$$

2. Encoder (ResNet3D-50): The encoder is comprised of several residual blocks. Every block receives the previous block's feature map, applies convolutional layers to it, and adds the initial input back via a shortcut connection. This improves the model's ability to learn by concentrating on the difference (residual) instead of the whole transformation.

$$F^{(l)} = H^{(l)}\big(F^{(l-1)}\big) + F^{(l-1)}$$

3. Classifier Head: Following the final residual block, the feature map is reduced in dimension using global average pooling. This spatially downsizes by averaging the feature channels, leaving a dense feature vector. This vector is then input into a linear (fully connected) layer that outputs logits per class. A softmax function then transforms these logits to probabilities.

$$Z = \mathrm{GlobalAvgPool}\big(F^{(L)}\big)$$

$$\hat{Y} = \mathrm{Softmax}\big(W_{cls}\cdot Z + b_{cls}\big)$$

4. Loss Function (Cross Entropy): For binary classification, we utilize the Cross Entropy Loss, which is a measure of how well predicted probabilities match the actual labels. The loss heavily penalizes incorrect predictions.

$$L_{CE} = -\sum_{i=1}^{N} y_i \cdot \log(\hat{y}_i)$$

5. Output: The model produces class probabilities for every sample. The class with the greatest predicted probability is the predicted class.

$$\hat{Y} = \mathrm{Softmax}\big(W_{cls}\cdot Z + b_{cls}\big) \in \mathbb{R}^{N\times 2}$$

V. Result and Evaluation metrics

The fig 8 describes about the major outcome of this project was to predict whether the person was diseased or not. By this model, we can predict it accurately.

Fig. 8: disease prediction of 3D images

1. Accuracy: Accuracy informs us about the number of pixels our model identified as positive, which actually were correct. Essentially, it inspects how cautious the model is in its decisions; we refer to it as pixel-wise prediction - i.e., how well the model indicates each pixel as belonging to the disease.

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

2. Recall: Recall refers to the fact that the model is able to capture most of the important features in the image. High recall is particularly crucial in medical segmentation, where failing to capture vital regions results in misdiagnosis.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

3. F1-Score: F1-Score refers to the model's precision in identifying the positive areas without excessive errors. A good F1-score indicates that the model is sensitive and precise, hence suitable for use in real-world medical settings.

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

4. Precision: Precision tells us how many of the pixels our model predicted as positive were actually correct. In simple terms, it checks how careful the model is in making decisions, we call it pixel-wise prediction, meaning how accurately the model marks each pixel as part of the disease.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

A. Performance Visualization

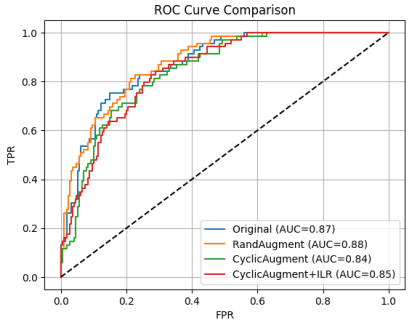

Fig. 9: ROC curve of AdrenalMnist

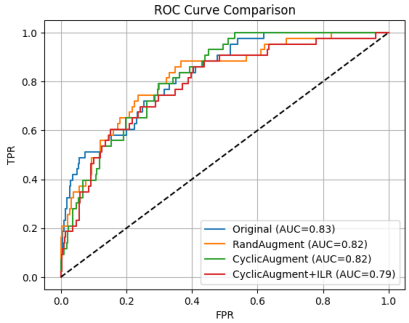

Fig. 10: ROC curve of VesselMnist

The fig 9, 10 shows how the models – Original, RandAugment, CyclicAugment, and CyclicAugment + ILR – performed using Resnet3d-50 for detecting Vessel and Adrenal diseases.

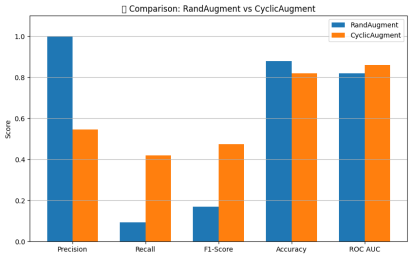

Fig. 11: comparison of RandAugment and CyclicAugment

The fig 11 shows the comparison of accuracy between RandAugment and CyclicAugment. In the concept of recall, precision, accuracy, and f1-score.

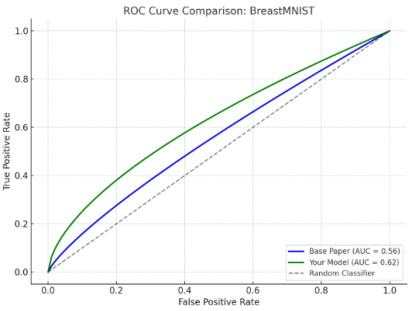

Fig. 12: comparison of breastmnist with roc-curve

The fig 12 mainly describes the comparison of breastMnist data of the present research with previous research. The ROC curve of present research improves slighter than the previous one.

VI. Overall performance table

A. Evaluation metrics for Vessel and Adrenal

TABLE III: EValues of ROC-AUC, F1-Score for Adrenal and Vessel

Setup	ROC AUC		F1-Score	
	Vessel	Adrenal	Vessel	Adrenal
Original	0.87	0.88	0.81	0.74
RandAugment	0.91	0.86	0.76	0.58
CyclicAugment	0.87	0.89	0.74	0.56
CyclicAug + ILR	0.89	0.91	0.77	0.67

TABLE IV: EValues of Recall, Precision for Adrenal and Vessel

Setup	Recall		Precision	
	Vessel	Adrenal	Vessel	Adrenal
Original	0.53	0.71	0.56	0.77
RandAugment	0.75	0.84	0.78	0.88
CyclicAugment	0.84	0.64	0.69	0.69
CyclicAug + ILR	0.66	0.63	0.79	0.71

The table III, IV mainly evaluates the performance of the models that were tested under original, RandAugment, CyclicAugment, and CyclicAugment + ILR. The results are mainly compared on the criteria of ROC-AUC Curve, F1-Score, Recall, and Precision for both Vessel and Adrenal datasets.

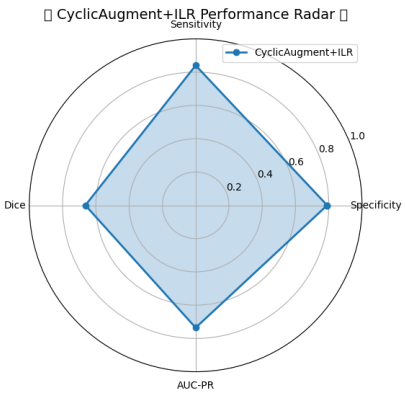

Fig. 13: Performance Radar across evaluation metrics: Sensitivity, Dice, Specificity, and AUC-PR

The fig 13 shows the balanced performance across key metrics. With high Sensitivity(0.85) and good Specificity(0.80), which reduces the false-positives and false-negatives in disease detection. The AUC-PR (0.75) which highlights its robustness in the criteria of imbalance datasets. Dice Score(0.67) indicates the segmentation overlap of the model.

VII. Conclusion

In this research, a deep learning technique was used to accurately identify disease from high-resolution medical imagery [14] using the ResNet3D-50 model. This model demonstrated robust performance across Adrenal and Vessel-related diseases with three-dimensional images. This study introduced several advancements that have improved the accuracy in the identification of diseases. The proposed method significantly enhances capabilities in key domains such as:

- Faster diagnosis: This model decreases the time required for the identification of the disease.
- Clinical decision support: This model helps doctors to identify the disease precisely and helps with faster diagnosis.
- Better 3D understanding: unlike two-dimensional models, this model enhances it to three-dimensional and allows for understanding spatial relationships between slices for any imagery [15].

References

[1] "eMedicine.medscape.com," accessed: 30 July 2025. [Online]. Available: https://emedicine.medscape.com

[2] X. Song, W. Liu, F. Xue, J. Zhong, Y. Yang, Y. Liu, J. Xie, E. Wu, L. Zhang, J. Shi, and R. Yang, "Real-world analysis of haemophilia patients in China: A single centre's experience," *Haemophilia*, vol. 26, no. 4, pp. 584–590, Jul. 2020. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/32432832/

[3] A. S. Farhan, M. Khalid, and U. Manzoor, "Combined oriented data augmentation method for brain mri images," *IEEE Access*, vol. 13, pp. 9981–9994, 2025.

[4] D. Sun, F. Dornaika, and N. Barrena, "Hsmix: Hard and soft mixing data augmentation for medical image segmentation," *Information Fusion*, vol. 115, p. 102741, 2025.

[5] Y. Wang, H. Xiong, K. Sun, S. Bai, L. Dai, Z. Ding, J. Liu, Q. Wang, Q. Liu, and D. Shen, "Toward general text-guided multimodal brain mri synthesis for diagnosis and medical image analysis," *Cell Reports Medicine*, vol. 6, no. 6, p. 102182, 2025.

[6] H. Laçi, K. Sevrani, and S. Iqbal, "Deep learning approaches for classification tasks in medical x-ray, mri, and ultrasound images: a scoping review," *BMC Medical Imaging*, vol. 25, no. 1, p. 156, 2025.

[7] Z. U. Rahman, J.-H. Lee, D. T. Vu, I. Murtza, and J.-Y. Kim, "Duco-net: Dual-contrastive learning network for medical report retrieval leveraging enhanced encoders and augmentations," *IEEE Access*, vol. 13, pp. 27 462–27 476, 2025.

[8] M.-J. Kim, J.-W. Chae, and H.-C. Cho, "Cyclicaugment: Optimized medical image analysis via adaptive augmentation intensity," *IEEE Access*, vol. 13, pp. 86 562–86 575, 2025.

[9] V. Sánchez, Güven, G. Nápoles, and M. Postma, "Data augmentation techniques for fmri data: A technical survey," *IEEE Access*, vol. 13, pp. 66 529–66 556, 2025.

[10] S. D. Westfechtel, K. Kußmann, C. Aßmann, M. S. Huppertz, R. M. Siepmann, T. Lemainque, V. R. Winter, A. Barabasch, C. K. Kuhl, D. Truhn, and S. Nebelung, "Ai in motion: the impact of data augmentation strategies on mitigating mri motion artifacts," *European Radiology*, p. in press, 2025.

[11] W. Li, L. Yang, G. Peng, G. Pang, Z. Yu, and X. Zhu, "An effective microscopic image augmentation approach," *Scientific Reports*, vol. 15, no. 1, p. 10247, 2025.

[12] P. Alimisis, I. Mademlis, P. Radoglou-Grammatikis, P. Sarigiannidis, and G. T. Papadopoulos, "Advances in diffusion models for image data augmentation: a review of methods, models, evaluation metrics and future research directions," *Artificial Intelligence Review*, vol. 58, no. 11, pp. 1–56, 2025.

[13] MIN-JUN KIM 1 , JUNG-WOO CHAE2 , AND HYUN-CHONG CHO, "Cyclicaugment: Optimised medical image analysis via adaptive augmentation intensity," https://zenodo.org/records/10519652, 2025, accessed: [Date Accessed, e.g., Jul. 30, 2025].

[14] S. Moturi, M. Ainavolu, N. R. Dokku, P. Kasula, N. Yaragani, and V. Dodda, "Enhanced lung cancer detection using deep learning ensemble approach," in *2024 1st International Conference for Women in Computing, InCoWoCo 2024 - Proceedings*, IEEE. IEEE, 2024.

[15] S. Rao, T. Dulla, V. Kolla, G. Kurakula, M. Suneetha, S. Moturi, and D. Reddy, "Deeplearning-based tomato leaf disease identification: Enhancing classification with alexnet," in *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation Iatmsi 2025*, IEEE. IEEE, 2025.