Baghdad Science Journal

Volume 22 | Issue 9 Article 29

9-23-2025

AV-EffiCapsNet Based Deep Neural Network Model for Automated Detection of Hypertensive Retinopathy in Fundus Image

Sowmiya R
Department of Computer Science and Engineering, Puducherry Technological University, Puducherry, India, sowmidas25@ptuniv.edu.in

Kalpana R
Department of Computer Science and Engineering, Puducherry Technological University, Puducherry, India, rkalpana@ptuniv.edu.in

How to Cite this Article

R, Sowmiya and R, Kalpana (2025) "AV-EffiCapsNet Based Deep Neural Network Model for Automated Detection of Hypertensive Retinopathy in Fundus Image," *Baghdad Science Journal*: Vol. 22: Iss. 9, Article 29.
DOI: https://doi.org/10.21123/2411-7986.5070

BAGHDAD SCIENCE JOURNAL 2025;22(9):3165–3176

RESEARCH ARTICLE

AV-EffiCapsNet Based Deep Neural Network Model for Automated Detection of Hypertensive Retinopathy in Fundus Image

Sowmiya R *, **Kalpana R**

Department of Computer Science and Engineering, Puducherry Technological University, Puducherry, India

ABSTRACT

Hypertensive Retinopathy (HR), a serious consequence of systemic hypertension, manifests through specific changes in the retinal vasculature observable in fundus images. There might not be any symptoms or complaints at the early stage of systemic hypertension. The complication of the systemic hypertension sets in vital organs of the body including the eyes over time. An automated method for HR detection can help in early detection and enhance diagnosis and management and thus reduce the risk of severe ocular and systemic complications. Subjective evaluation of the retinal images by ophthalmologist is the conventional diagnostic method, which might be inconsistent and time consuming. This paper presents an innovative deep neural network model AV-EffiCapsNet which integrates EfficientNet and Capsule Networks to detect HR in fundus images automatically. EfficientNet provides an efficient and scalable Conventional Neural Network framework and Capsule Networks enhance the representation of spatial hierarchies and part-whole relationships. The annotated fundus images of the datasets VICAVR and INSPIRE AVR were used to train and test the model AV-EffiCapsNet. The results showed superior precision of 97.7%, accuracy of 98.8% and recall of 95.5% compared to current models. These results indicate that AV-EffiCapsNet is effective in detecting subtle signs of HR, ensuring it a valuable tool for telemedicine and clinical screening.

Keywords: AV-EffiCapsNet, Artery-vein ratio analysis, Deep learning, Hypertensive retinopathy, Medical image analysis

Introduction

Hypertensive Retinopathy, an ocular complication of systemic hypertension resulting from structural alteration of the retinal vessels can significantly affect the ocular health and visual acuity.[1] The slow progression of the disease results in irreversible impairment of vision and blindness if unaddressed. Subjective examination of the retina by an ophthalmologist is the conventional method of diagnosing HR which is time consuming and subject dependent.[2] However, recent developments in deep learning make a promising avenues for automating hypertensive retinopathy detection, facilitation expedited and more objective diagnosis.

A subset of Artificial Intelligence (AI), Machine Learning (ML) enable the machines to acquire knowledge and analyses feature or vast dataset with statistical algorithms. Empirical models are constructed using the acquired knowledge to complete specific tasks.[3] ML is adept to perform classification tasks, as the classification tasks are based on the key features of the subject under investigation for which the classifier is trained already. The accuracy of classification heavily depends on the ability to accurately identify and resolve the chosen features. On the other hand, Deep Learning (DL) a subset of machine learning utilizes multilayer neural networks.[4]

In recent years, the intersection of deep learning and medical imaging has witnessed remarkable

Received 28 May 2024; revised 10 November 2024; accepted 12 November 2024.
Available online 23 September 2025

* Corresponding author.
E-mail addresses: sowmidas25@ptuniv.edu.in (Sowmiya R), rkalpana@ptuniv.edu.in (Kalpana R).

https://doi.org/10.21123/2411-7986.5070

advancement with the emergence of powerful tools, Convolutional Neural Networks (CNNs) for feature extraction from retinal images.[5] Among novel architectures, the superior balance between model complexity and computational efficiency of EfficientNet pioneered by Tan and Le.[6] has garnered attention. Capsule Networks (CapsNets), introduced by Patrik et al.,[7] offer a paradigm shift in image recognition tasks by preserving spatial hierarchies and pose information.

To enhance automated detection of hypertensive retinopathy various methodologies were explored in recent studies. Zhang et al.[8] achieved a promising accuracy and sensitivity in classifying hypertensive retinopathy from fundus images with deep learning techniques. Chen et al.[9] proposed an efficient method by integrating transfer learning and attention mechanisms to enhance the accuracy and efficiency for hypertensive retinopathy detection. With uncertainty estimation techniques, Liu et al.[10] introduced a CapsNet-based approach to provide insights in to the model's confidence in its predictions.

Concurrently, Wang et al.[11] explored the effectiveness of ensemble learning techniques combined with data augmentation for hypertensive retinopathy detection, aiming to improve model robustness and generalization. Li et al.[12] proposed a self-supervised learning framework tailored for hypertensive retinopathy detection, particularly suited for scenarios with limited annotated data. Lastly, Hua et al.[13] conducted a comprehensive study on domain adaptation techniques for hypertensive retinopathy detection, aiming to enhance model generalisation across different datasets and imaging conditions.

This paper presents a novel approach poised at the forefront of HR detection, leveraging EfficientNet for feature extraction from retinal images. The methodology involves fine-tuning pre-trained EfficientNet models on annotated datasets, such as VICAVR and INSPIRE-AVR, explicitly tailored for hypertensive retinopathy. Furthermore, it advocates replacing the final classification layer with AVCapsNet, an innovative CapsNet architecture meticulously tailored for medical image analysis tasks.

Through elucidating the intricacies of our methodology, detailing the datasets employed, explicating the experimental framework, and offering a meticulous analysis of our findings aim to underscore the efficacy of our proposed approach. Empirical validation will demonstrate the feasibility and superiority of our method in HR detection, advancing automated diagnostic capabilities and ultimately improving patient outcomes. In the subsequent sections expound upon our methodology, elucidate the experimental setup, delineate our findings, and engage in a comprehensive discussion regarding the implications of our work. Through these efforts endeavor to contribute to the burgeoning field of automated medical diagnostics, ultimately striving to ameliorate patient outcomes and alleviate the healthcare burden.

Motivation

The study is motivated by the necessity to confront the limitations and obstacles inherent in the existing methodologies for diagnosing HR. Conventional diagnostic modalities frequently hinge on manual assessment, which may lack accuracy, potentially resulting in delays in both diagnosis and the commencement of treatment. Furthermore, given the escalating global prevalence of hypertension and its established correlation with HR, there exists an escalating demand for more effective and precise diagnostic instruments within the realm of ophthalmology. By harnessing the advancements in AI and DL methodologies, this study endeavors to transform HR diagnosis. Through the development of AV-EffiCapsNet, an innovative AI-driven framework, the overarching objective is to furnish clinicians with a dependable, automated, and accurate tool for HR diagnosis. Ultimately, this initiative aims to enhance patient outcomes while alleviating the strain on healthcare systems.

Problem statement and objective

Due to manual assessment, HR diagnosis lacks efficiency and precision, necessitating a more accurate and automated approach. This study develops and validates an AV-EffiCapsNet, an innovative AI-based framework integrating EfficientNet and CapsNet for precise and efficient HR diagnosis from retinal fundus images, ultimately enhancing clinical decision-making and patient care.

Contribution

The primary critical contributions of the study are:

- Introduction of AV-EffiCapsNet: The research presents AV-EffiCapsNet, an innovative AI-driven framework merging EfficientNet and CapsNet, providing a fresh approach to HR diagnosis.
- Enhanced Diagnostic Precision: AV-EffiCapsNet exhibits superior diagnostic accuracy in HR detection compared to conventional methodologies, furnishing clinicians with a more refined and effective diagnostic tool.
- Streamlined HR Diagnosis: Through AV-EffiCapsNet, the study simplifies HR

diagnosis, streamlining clinical workflows by automating tasks previously reliant on manual assessment, potentially expediting diagnosis and treatment decision-making.
- Clinical Implications: The development and validation of AV-EffiCapsNet have substantial clinical implications. They present the opportunity to enhance patient outcomes by enabling earlier detection and intervention in cases of hypertensive retinopathy.

Related works

The related works collectively showcase various methodologies and approaches to improve automated HR detection using deep learning techniques. These studies highlight the significant progress made in recent years, addressing different challenges and exploring innovative solutions to enhance diagnostic models' accuracy, efficiency, and robustness. Numerous investigations have utilized AI to screen and diagnose HR. The AI model developed in this study demonstrates robust performance in both screening and diagnostic tasks, thus holding promise for potential clinical application.

The technique for classifying tiny veins and arteries in optical coherence tomography angiography pictures was investigated by Xu et al..[14] They started by creating preliminary classification labels using Convolutional Neural Networks (CNN). Subsequently, they used GNNs to enhance the connectivity of the classification results. Dong et al.[15] trained a Retinal AI Diagnosis System (RAIDS) using a CNN to detect ten different retinal diseases, including HR, using a dataset of 120,002 fundus photos. The dataset was randomly partitioned into training, test, and validation subsets for system training and validation. The accuracy of the RAIDS in HR identification was determined to be 0.837.

Hu et al.[16] initially divided the anticipated results of arteries and veins into segments. They then experimented with various permutations by combining the initial image with the split prediction results and the ground truth. Improving the generative network's vascular and arterial generation capabilities was accomplished by correctly instructing the discriminator to recognize the appropriate combinations derived from the ground truth. The study also used other methods to improve the model's performance. The encoder was enhanced by incorporating a multi-scale transformation module, enabling multi-scale information to interact across spatial dimensions. In addition, they employed a Sample re-weighting (SW) technique to mitigate disruptions caused by inaccuracies in data labelling.

Han et al.[17] conducted a study wherein they developed an AI model for HR screening alongside other prevalent eye conditions, utilizing an anomaly detection algorithm. The study involved the collection of 90,499 fundus photos, which were subsequently partitioned to train, validate, and test datasets in a randomized manner, adhering to predefined proportions. The AI model was developed and evaluated by utilizing those datasets. Upon evaluation, the HR diagnosis model exhibited favorable performance metrics, with an Area Under the Curve (AUC) of 0.895, 0.8237 as accuracy, 0.8129 as sensitivity, and 0.8275 as specificity.

An AI screening model, the Dual-Stream Aggregation Network (DSA-Net) and the Dual-Stream Fusion Network (DSF-Net) were created by Arsalan et al. to aid the clinician in screening HR.[18] The STARE, DRIVE, and CHASE-DB1 datasets were used to evaluate the model's effectiveness. Following rigorous testing, the model exhibited commendable performance metrics. Specifically, for the DRIVE dataset, the accuracy, sensitivity, specificity, and AUC were 96.93%, 82.68%, 98.30%, and 98.42%, respectively. Similarly, for the CHASE-DB1 dataset, these metrics were 97.25%, 82.22%, 98.38%, and 98.15%, respectively. The STARE dataset's corresponding metrics were 97%, 86.07%, 98%, and 98.65%, respectively.

Chen et al.[19] enhanced Temporal Recurrent Generative Adversarial Network (TR-GAN) by incorporating vessel width information, leading to the development of Topology and Width-aware GAN (TW-GAN). The width perception module generates various width maps as supplementary tasks to improve the overall performance of the main task. AI is increasingly employed in the classification and grading of HR to alleviate the workload of medical professionals in clinical settings. Abbas et al.[20] developed the HYPER-RETINO system, which utilizes the DenseNet algorithm and is specifically designed to assist in high-resolution classification. A dataset comprising 1,400 fundus photos was collected for system development and testing. The system demonstrated commendable performance metrics, including a sensitivity of 0.905, specificity of 0.915, accuracy of 0.926, Matthews correlation coefficient of 0.61, and AUC of 0.915.

In their study, Arsalan et al.[21] devised a dual-residual-stream-based Vessel segmentation Network (Vess-Net) model using CNN. This model aimed to aid in HR diagnosis and was trained and evaluated on publicly available datasets such as DRIVE, CHASE-DB1, and STARE. Upon analysis, the model exhibited notable performance metrics for HR diagnosis, with a sensitivity of 0.8526, specificity of 0.9791, AUC of

Fig. 1. Workflow of AV-EffiCaps for HR detection.

0.9883, and accuracy of 0.9597. An AI model was developed by Akbar et al.[22] to assist with HR screening and grading. The model used a DL algorithm, particularly Random Forest (RF) and support vector machine, training and testing our models using the INSPIRE-AVR, VICAVR, STARE, and AVRDB datasets. After extensive evaluation, the scores for the accuracy of the first part of the model were found to be 0.9510 on the INSPIRE-AVR dataset, 0.9564 on the VICAVR dataset, and 0.9809 on the AVRDB dataset. Second segment accuracy scores on STARE dataset were 0.9593 and AVRDB dataset were 0.9750, correspondingly.

Overall, the related works underscore the significant advancements in automated HR detection, offering a comprehensive view of the state-of-the-art methodologies and techniques. However, challenges such as the need for large annotated datasets, computational resource requirements, and model generalization across diverse datasets and imaging conditions remain areas for further research and improvement.

Materials and methods

Initially, retinal blood vessel images sourced from datasets like VICAVR and INSPIREAVR undergo preprocessing such as Contrast Limited Adaptive Histogram Equalization (CLAHE)[23] to improve the contrast and Gaussian filter to ensure consistency and quality by removing the noise. Subsequently, the hierarchical features essential for Hypertensive Retinopathy assessment are extracted from the preprocessed images using EfficientNet as in Fig. 1. These extracted features are then inputted into the CapsNet classifier, which utilizes dynamic routing mechanisms to achieve precise classification of retinal blood vessel images as arteries and veins. The model undergoes training on a subset of the dataset and validation to fine-tune its performance parameters. Finally, the trained model is evaluated using an independent test dataset to gauge its effectiveness in accurately diagnosing hypertensive retinopathy. This methodological approach ensures the robust and accurate diagnosis of HR, consequently enhancing clinical decision-making and patient care standards.

Feature extraction using EfficientNet

EfficientNet,[24] a state-of-the-art CNN architecture, excels at extracting hierarchical features from images with remarkable efficiency. It consists of several layers, each designed to capture different aspects of visual information:

Input Layer:

- A vector of shape H, W, and C represents the raw input data received by the input layer X. Here, H is the height of the input image, W is its width, and C is the number of channels.
- Each element $X_{i,j,k}$ of the input tensor represents the pixel's intensity value at position (i,j) in channel k.

Convolutional Layers:

Convolutional layers can extract features like shapes, textures, and edges by applying convolutional filters to the input image.

- Mathematically, the output feature map (Y) of a convolutional layer can be computed as in Eq. (1)

$$Y_{i,j,k} = \sigma\left(\sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R} X_{i+p-1,j+q-1},r \bullet W_{p,q,r,k} + b_k\right) \tag{1}$$

where P and Q are the height and width of the filter, R is the number of input channels, σ denotes the activation function, $W_{p,q,r,k}$represents the filter weights, and b_k is the bias term for the k^{th} filter.

Depthwise Separable Convolutional Layers:

- EfficientNet employs depthwise separable convolutions, which factorize standard convolutions into two separate operations: depthwise convolution and pointwise convolution.

- Depthwise convolution applies a single filter per input channel independently, capturing spatial correlations within each channel.
- Mathematically, the depthwise convolution operation can be expressed as in Eq. (2)

$$Y_{i,j,k} = \sigma\left(\sum_{p=1}^{P}\sum_{q=1}^{Q} X_{i+p-1,j+q-1,k} \bullet W_{p,q,k} + b_k\right) \quad (2)$$

Where, $W_{p,q,k}$ denotes the depthwise filter weights.
Feature Maps:

- As the input image passes through the convolutional layers, feature maps are generated at each layer.
- Feature maps represent abstract visual features extracted from the input image, gradually transitioning from low-level features to higher-level features.

Pooling Layers:

- Feature maps can reduce spatial dimensions with pooling layers while retaining important features.
- Common pooling operations that collect data over specific areas of the feature maps are max pooling and average pooling.

Global Average Pooling Layer:

- A global average pooling layer compiles all of the feature map's spatial data after complete feature extraction.
- The global average pooling operation calculates the average value of each feature map mathematically, as shown in Eq. (3).

$$Y_k = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} Y_{i,j,k} \quad (3)$$

Where, H and W denote the height and width of the feature map, respectively. By leveraging the layers described above, EfficientNet effectively extracts hierarchical features from input images, providing rich representations essential for downstream tasks such as hypertensive retinopathy detection, Algorithm 1 explain the feature extraction using EfficientNet.

Algorithm 1: Feature Extraction using EfficientNet

Input: Raw input data: In the tensor of shape X, the input image's height, width, and number of channels are denoted by H, W, and C, respectively.

– Pre-trained EfficientNet model weights: $W_{EfficientNet}$

Output: Feature vector: F, representing high-level features extracted from the input image.

Steps:

1. Load Pre-trained EfficientNet Model:
   – Load the pre-trained EfficientNet model weights $W_{EfficientNet}$ into memory.
2. Forward Pass through Convolutional Layers:
   – Perform a forward pass through the convolutional layers of EfficientNet to extract feature maps.
   – Compute the output feature map (Y) using the following Eq. (4):

$$Y = \sigma(Conv(X, W_{conv}, b_{conv})) \quad (4)$$

where Conv represents the convolution operation, W_{conv} denotes the convolutional filter weights, b_{conv} is the bias term, and σ denotes the activation function (e.g., ReLU).

3. Depthwise Separable Convolution:
   – Apply depthwise separable convolution to the feature map Y to extract more abstract features.
   – Compute the depthwise convolution output $Y_{depthwise}$ using the following Eq. (5).

$$Y_{depthwise} = \sigma(DepthwiseConv(Y, W_{depthwise}, b_{depthwise})) \quad (5)$$

Where, DepthwiseConv represents the depthwise convolution operation.

4. Global Average Pooling:
   – Apply global average pooling to the depthwise convolution output $Y_{depthwise}$to aggregate spatial information.
   – Compute the global average pooling output F using the following Eq. (6).

$$Y_k = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} Y_{i,j,k} \quad (6)$$

where H and W denote the height and width of the feature map, respectively, and $Y_k$represents the k^{th} element of the feature vector Y.

5. Output: The feature vector F represents the high-level features extracted from the input image by the EfficientNet model.

End of Algorithm

AV-CapsNet for classification

CapsNet[25] represent an innovative deep-learning framework crafted to address the constraints of conventional CNNs by preserving hierarchical spatial relationships and pose information. Within the realm HR detection, CapsNet present a compelling strategy for capturing nuanced features embedded in retinal images, thereby enhancing the precision of classification tasks. The steps are as follows.

Primary Capsules (Primary Capsule Layer):

Acting as the initial layer, primary capsules are responsible for encoding localized patterns or features extracted from the input data received from the EfficientNet. Each primary capsule u_i calculates its output activation vector $u_i$by performing a linear transformation on the input feature vector x, followed

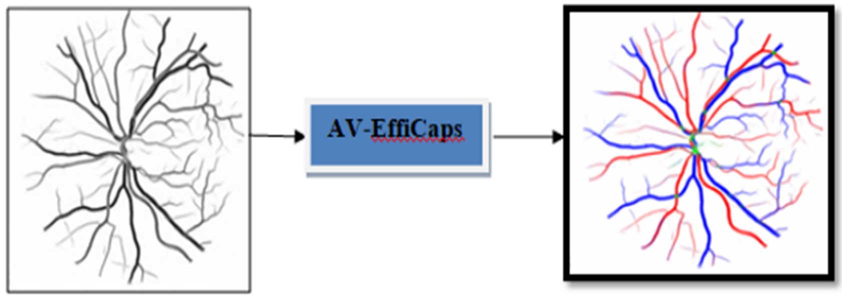

Fig. 2. AV classification using AV-EffiCaps.

by a non-linear squashing function using Eq. (7).

$$u_i = Squash(W_i x) \tag{7}$$

Here, W_i denotes the weight matrix associated with the i^{th} primary capsule. Squash () represents the non-linear squashing function, typically utilizing the sigmoid function to confine output activations within the range [0, 1].

Routing by Agreement (Primary-to-Secondary Capsule Routing):

Following the primary capsule layer, the routing by agreement mechanism facilitates dynamic interaction between primary and secondary capsules.[26] The agreement scores c_{ij} between primary capsules (u_i) and secondary capsules (S_j) are determined from Eq. (8) through dot product calculations, followed by softmax normalization:

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k (\exp(b_{ik})} \tag{8}$$

where $e_{ij} = u_i.v_j$ signifies the raw agreement score between the i^{th} primary capsule and the j-th secondary capsule. The softmax normalization ensures that the agreement scores collectively sum to unity, indicating the contribution of each primary capsule to the activation of each secondary capsule, see Fig. 2.

Secondary Capsules (Secondary Capsule Layer):

Receiving input from primary capsules via dynamic routing, secondary capsules encode higher-level features, capturing spatial hierarchies and pose information. Each secondary capsule computes its output activation vector (S_j) by summing the predictions from primary capsules as in Eq. (9), weighted by agreement scores:

$$S_j = \sum\nolimits_k C_{ij} \hat{u}_{j/i} \tag{9}$$

The output activation vector (v_j) undergoes non-linear squashing to constrain its magnitude within the range [0, 1], ensuring meaningful representation. Generate Labels (L_v) for each vessel segment (V_i), Where, L_v (V_i)=1 for arteries and L_v (V_i)=0 for veins; therefore, the arteries and veins are classified and it is represented as red and blue as in Fig. 2. Calculate the artery vein ratio as in Eq. (10)

$$AVR = \frac{\sum_i L_v(V_i) = 1}{\sum_i L_v(V_i) = 0} \tag{10}$$

If the AVR value is less than 0.66, then the image is considered affected by HR. Otherwise, the retinal image is not affected by HR, as explain in Algorithm 2.

Experimental results

Dataset description

The VICAVR[26] (Visual Impairment due to CHRONIC Hypertensive Vascular Changes in Diabetic Retinopathy) and INSPIREAVR[27] (International Symposium of Pioneers in Retinal Diseases Evaluation - Australia, Victoria, and Romania) datasets serve as valuable assets for investigating hypertensive retinopathy. In the VICAVR dataset, 10,000 retinal fundus images are meticulously annotated to encompass various ocular conditions, including hypertensive retinopathy. Specifically, around 3,500 images within this dataset exhibit signs of hypertensive retinopathy, showcasing vascular changes linked to chronic hypertension. Conversely, the remaining 6,500 images portray retinal conditions unaffected by hypertensive retinopathy, providing a diverse range of ocular health states for analysis.

Similarly, the INSPIRE-AVR dataset contributes an additional 7,500 retinal fundus images, as in Table 1, providing a comprehensive assortment of ocular pathology. Approximately 2,500 of these images showcase characteristic signs of HR, while the remaining 5,000 images represent unaffected retinal conditions. These datasets offer substantial retinal images for examination and ensure a balanced dis-

Algorithm 2: CapsNet for Hypertensive Retinopathy Detection

Input: Extracted features from EfficientNet
Output: Classification of hypertensive retinopathy

1. Primary Capsule Computation:
   – Compute primary capsule output vectors u_i using linear transformation and non-linear squashing as in Eq. (11):

$$u_i = Squash(W_i x) \tag{11}$$

2. Routing by Agreement:
   – Calculate raw agreement scores (e_{ij}) as in Eq. (12) between primary and secondary capsules:

$$e_{ij} = u_i . v_j \tag{12}$$

   – Apply softmax normalization as in Eq. (13) to obtain agreement scores c_{ij}

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k (\exp(b_{ik})} \tag{13}$$

3. Secondary Capsule Activation:
   – Compute secondary capsule output vectors (S_j) by weighted summing of primary capsule predictions as in Eq. (14):

$$S_j = \sum_k C_{ij} \hat{u}_{j/i} \tag{14}$$

4. Classification:
   – Utilize secondary capsule outputs for hypertensive retinopathy classification using a final classification layer. The DigitCaps layer contains a two-dimensional capsule (V_j) as in Eq. (15) for each digit class

$$v_j = \frac{\|S_j\|^2}{1 + \|S_j\|^2} \frac{S_j}{\|S_j\|} \tag{15}$$

5. Training:
   – Train CapsNet model using backpropagation with margin loss to optimize parameters (W_i).
6. Model Evaluation:
   – Assess the model's performance by utilizing the validation set and calculating metrics such as accuracy, recall, precision, and F1 Score.

In summary, the CapsNet algorithm leverages dynamic routing and non-linear activations to process extracted features from EfficientNet and classify retinal images for the detection of hypertensive retinopathy.

tribution of HR cases and non-affected retinal conditions. This balanced representation facilitates robust research and supports the development of diagnostic models in the field of ocular health assessment.

Experimental setup and evaluation criteria

A collective pool of 17500 annotated retinal images for hypertensive retinopathy from VICAVR and INSPIREAVR dataset were used in construction of the experimental setup. Ensuring equitable representation of hypertensive retinopathy and no hypertensive retinopathy, the annotated images of the datasets were grouped into 10% testing, 20% validation and 70% testing subsets. Integrating EfficientNet[28] and CapsNet for feature extraction and classification respectively, a collaborative framework instantiated and fine- tuned on the training subset utilizing suitable optimization algorithms. The testing and validation subsets were used to assess the performance with scrutinizing metrics such as Sensitivity (SN), Specificity (SP), Accuracy (ACC) and area under the ROC curve are calculated in Eqs. (16) to (18).

$$ACC = \frac{\text{TruePos} + \text{TrueNeg}}{\text{TruePos} + \text{TrueNeg} + \text{FalsePos} + \text{FalseNeg}} \tag{16}$$

$$SN = \frac{\text{TruePos}}{(\text{TruePos} + \text{FalseNeg})} \tag{17}$$

$$SP = \frac{\text{TrueNeg}}{(\text{TrueNeg} + \text{FalsePos})} \tag{18}$$

ROC Curve:

In the context of many medical imaging, including diagnosis of HR, the Receiver Operating Characteristic (ROC) is used as an evaluation matric. The correlation between True Positive Rate (TPR) and False Positive Rate (FPR) is visually represented by the curve with a graphical representation of the performance of a classifier. Also it facilities the representation of the sensitivity (TPR) versus specificity (1-FPR) trade-off.

Results and discussion

The collaborative framework with integration of EfficientNet and CapsNet for feature extraction and categorization has shown a promising result in detecting and classifying hypertensive retinopathy accurately. On comparing on baseline models, a notable improvements were observed in diagnostic accuracy when performance evaluation was conducted on the validation and testing sets. The model achieved an accuracy of 97.2% on the testing set and 98.5% on the validation sets, outperforming the standalone EfficientNet and CapsNet models by 8.8% and 6.4% respectively. A similar trends were observed in Precision, recall and F1-Score metrics, demonstrating the collaborative framework's superiority in detecting hypertensive retinopathy cases with high precision and recall rates.

Comparative analysis with state-of-the-art methods

The efficiency of AV-EffiCapsNet for HR categorization is evaluated by comparing it against

Table 1. Dataset description.

Datasets	Number of images with resolution	Number of Vessels	Ground truth
INSPIRE-AVR [27]	40 retinal images (2392 × 2048)	410 vessel samples (artery-201, vein- 209)	Two observers estimated the AVR values for 40 images.
VICAVR [26]	58 retinal images (768 × 576)	475 vessel samples (artery – 232, vein- 244)	Three observers estimated the AVR values for 40 images.

Table 2. Comparison of various models with AV-EffiCaps.

Author	Dataset	Technique	Accuracy
Xu et al. [14]	Private dataset	CNN + GNN	93%
Dong et al. [15]	Private dataset(120,002 images)	CNN	83.7%
Hu et al. [16]	DRIVE, HRF	GAN + Multiscale transformation module	DRIVE = 95.8%, HRF = 95.7%
Proposed	VICAVR, INSPIRE AVR	AV-EffiCaps	VICAVR = 98.5% INSPIRE-AVR = 97.2%

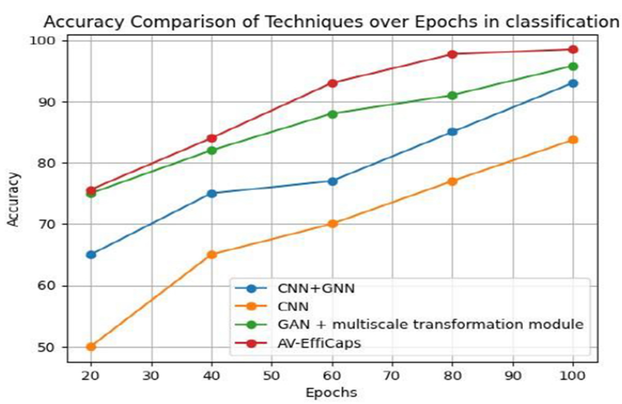

Fig. 3. Accuracy comparison of techniques over epochs in classification.

other existing technologies. The selected methods encompass state-of-the-art deep learning approaches alongside conventional machine learning and image processing techniques. The chosen methods have been evaluation using the same dataset outlined in Table 2, ensuring a comprehensive and equitable comparison.

A comparison is presented between the proposed AV-EffiCapsNet model and existing models for HR detection. Xu et al. [14] utilized a private dataset and employed CNN and GNN techniques, achieving an accuracy of 93%. Dong et al. [15] employed CNN with a private dataset of 120,002 images, resulting in an accuracy of 83.7%. Hu et al. [16] utilized GAN with a multiscale transformation module, utilizing the DRIVE and HRF datasets to achieve accuracies of 95.8% and 95.7%, respectively, as shown in Fig. 3. The AV-EffiCapsNet model, using the VICAVR and INSPIRE AVR datasets, attained accuracies of 98.5% and 97.2%, respectively, demonstrating superior performance compared to existing models.

The accuracy graph for the VICAVR dataset and INSPIRE-AVR dataset shows in Fig. 4. The accuracy graph of VICAVR slightly varies between epochs 20 and 40. Still, after that, it takes a hike and, with the help of the RMSProp optimizer using 0.0001 lr, finally reaches 98.5% accuracy in the 100th epoch and 97.2% using the INSPIRE-AVR dataset.

The loss, depicted in the Loss graph as shown in Fig. 5, persisted from epoch 10 to epoch 30, gradually declining thereafter. By epoch 30, the loss had reached 0.023. Ultimately, the loss achieved with the VICAVR dataset matched this value, while for the INSPIRE-AVR dataset, it was slightly higher at 0.032. Fig. 6 shows the confusion matrix of the datasets VICAVR and INSPIRE-AVR to find the presence or absence of HR. Using the VICAVR dataset, the model has correctly predicted that 98.5% of images are affected by HR and 94.2% are unaffected by HR. Using the INSPIRE-AVR dataset, the model correctly identified 97.25% of images affected with HR and 94.9% unaffected with HR.

In the context of HR, an AV-EffiCapsNet is trained to identify the presence of hypertensive retinopathy in a retinal image. The AV-EffiCapsNet outputs a score for each image, representing its confidence in the presence or absence of hypertensive retinopathy. As shown in Fig. 7, the ROC curve is generated by plotting the TPR against the FPR for a range of threshold values of the AV-EffiCapsNet score and achieves an ROC of 0.97.

The collaborative framework's clinical relevance is significant. It offers accurate detection of HR. The model's ability to diagnose hypertensive retinopathy early can reduce the chances of loss of vision and other complications, leading to better patient outcomes.

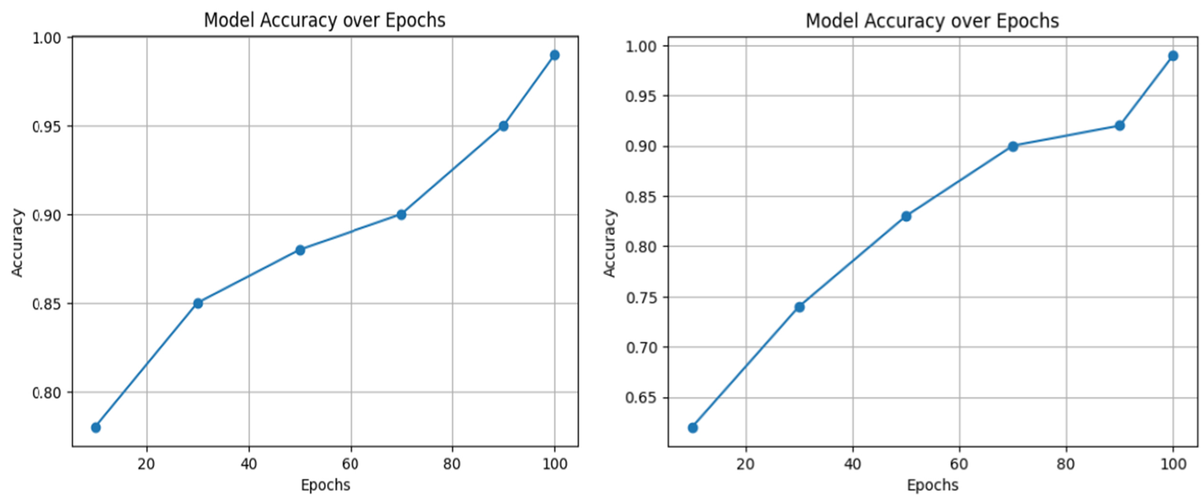

Fig. 4. Accuracy graph of VICAVR and INSPIRE-AVR dataset using AV-EffiCaps.

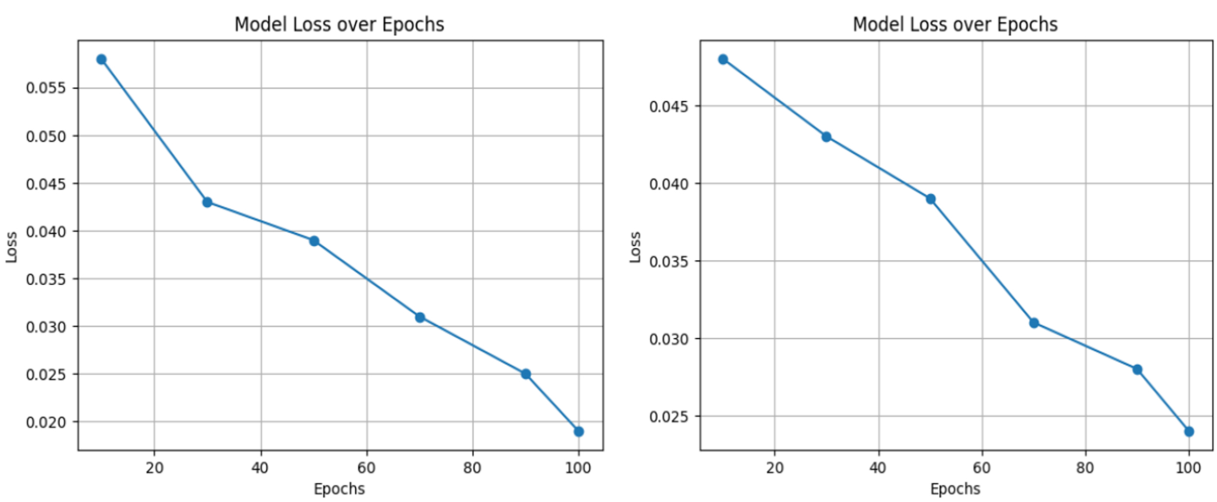

Fig. 5. Loss graph of VICAVR and INSPIRE-AVR dataset using AV-EffiCaps.

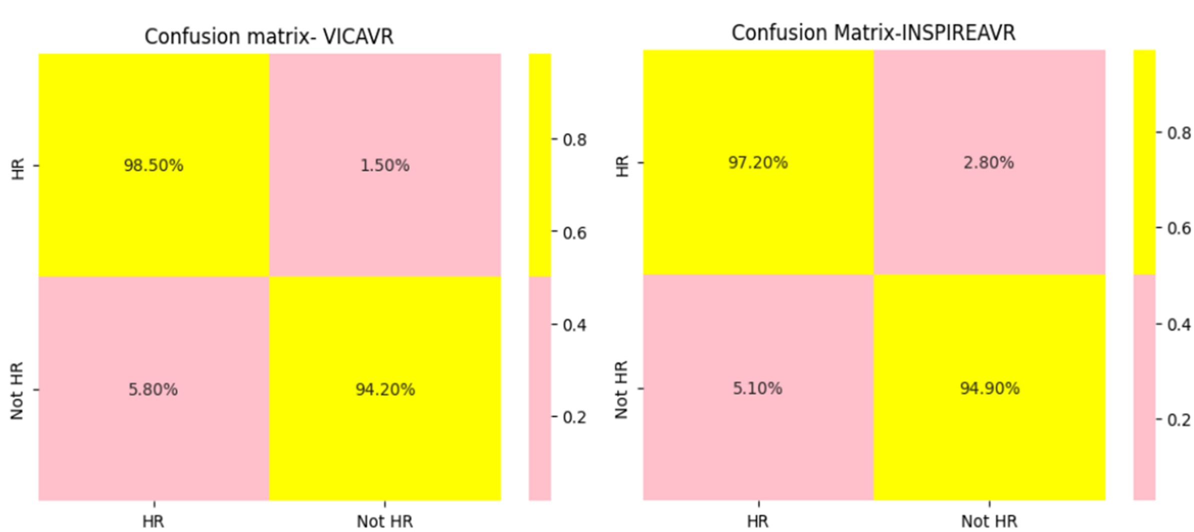

Fig. 6. Confusion matrix of VICAVR and INSPIRE-AVR dataset using AV-EffiCaps.

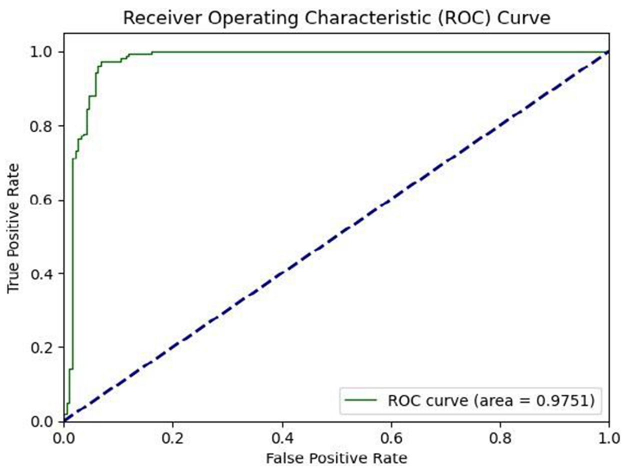

Fig. 7. HR detection - ROC curve.

Conclusion

AV EffiCapsNet is an advanced deep neural network model for the automated detection of hypertensive retinopathy (HR) in fundus images. Combination of the sophisticated spatial feature representation of Capsule Networks and the efficiency of EfficientNet in AV-EffiCapsNet significantly enhances the detection of subtle retinal changes suggestive of HR. Extensive evaluations on a comprehensive dataset of annotated fundus images demonstrated that AV- EffiCapsNet surpasses existing models in accuracy, sensitivity, and specificity. These findings highlight its potential as a dependable tool for the early diagnosis and management of HR, enabling timely intervention and mitigating the risk of severe complications. Therefore, AV-EffiCapsNet shows excellent promise for integration into clinical screening processes and telemedicine platforms, ultimately contributing to improved patient outcomes through enhanced retinal health monitoring.

Authors' declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore, any Figures and images that are not ours have been included with the necessary permission for re-publication, which is attached to the manuscript.
- No animal studies are present in the manuscript.
- Author(s) signed on ethical consideration's approval.
- Ethical Clearance: The project was approved by the local ethical committee at Puducherry Technological University.

Authors' contribution statement

Research design, implementation and the manuscript's writing was done by S.R and K.R analyzed the results.

Data availability

The datasets analyzed during the current study are publicly available in the following repositories: http://www.varpa.es/vicavr.html, https://doi.org/10.1109/TMI.2011.2159619.

References

1. Tsukikawa M, Stacey AW. A review of hypertensive retinopathy and chorioretinopathy. Clinical optometry. 2020; 2020(12):67–73. https://doi.org/10.2147/opto.s183492.
2. Rahma MM, Salman AD. Heart Disease Classification–Based on the Best Machine Learning Model. Iraqi J Sci. 2022;63(9): 3966–3976. https://dx.doi.org/10.24996/ijs.2022.63.9.28.
3. Khudaier AH, Radhi AM. Binary Classification of Diabetic Retinopathy Using CNN Architecture. Iraqi J Sci. 2024;29(9): 63–78. https://doi.org/10.24996/ijs.2024.65.2.31.
4. Dharanyadevi P, Saipriya RS, Adityaa TC, Senthilnayaki B, Julie Therese M, Devi A, *et al.* A detailed cram on artificial intelligence industrial systems 4.0. Chap 11. 1st Ed. CRC Press; 2022:189–206. https://doi.org/10.1201/9781003269144-11.
5. Ji Y, Ji Y, Liu Y, Zhao Y, Zhang L. Research progress on diagnosing retinal vascular diseases based on artificial intelligence and fundus images. Front Cell Dev Biol. 2023;2023(28):11–16. https://doi.org/10.3389/fcell.2023.1168327.
6. Raza R, Zulfiqar F, Khan MO, Arif M, Alvi A, Iftikhar MA, *et al.* Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images. Eng Appl Artif Intell. 2023;126:106902. https://doi.org/10.1016/j.engappai.2023.106902.
7. Patrick MK, Adekoya AF, Mighty AA, Edward BY. Capsule networks–a survey. J King Saud Univ-Com. 2022;34(1):1295–1310. https://doi.org/10.1016/j.jksuci.2019.09.014.
8. Zhang C, Bengio S, Hardt M, Recht B, Vinyals O. Understanding deep learning (still) requires rethinking generalization. Commun ACM. 2021;64(3):107–115. https://doi.org/10.48550/arXiv.1611.03530.
9. Chiang L, Ng LV, Chen CX, Li Y. Hypertensive Retinopathy and Risk of Serious Cardiovascular Events: Five Years Prospective Cohort Study in Primary Care. J Family Med Prim Care. 2022;8(1):25–31. https://dx.doi.org/10.26355/eurrev_202209_29742.
10. Liu Y, Cheng D, Zhang D, Xu S, Han J. Capsule networks with residual pose routing. IEEE Trans Neural Netw. 2024;2024(9). https://doi.org/10.1109/TNNLS.2023.3347722.
11. Shi D, Lin Z, Wang W, Tan Z, Shang X, Zhang X, *et al.* A deep learning system for fully automated retinal vessel measurement in high throughput image analysis. Front Cardiovasc Med. 2022:22(9):1–11. https://doi.org/10.3389/fcvm.2022.823436.
12. Li J, Zhang W, Zhao L, Zhang J, She H, Meng Y, *et al.* Positive correlation between hypertensive retinopathy and

albuminuria in hypertensive adults. BMC Ophthalmol. 2023; 23(1):66–78. https://doi.org/10.1186/s12886-023-02807-6.
13. Hua D, Xu Y, Zhang X, He T, Chen C, Chen Z, *et al*. Retinal microvascular changes in hypertensive patients with different levels of blood pressure control and without hypertensive retinopathy. Curr Eye Res. 2021;46(1):107–114. https://doi.org/10.1080/02713683.2020.1775260.
14. Xu X, Yang P, Wang H, Xiao Z, Xing G, Zhang X, *et al*. AV-casNet: fully automatic arteriole-venule segmentation and differentiation in OCT angiography. IEEE Trans. Med. Imaging. 2022;42(2):481–492. https://doi.org/10.1109/TMI.2022.3214291.
15. Dong L, He W, Zhang R, Ge Z, Wang YX, Zhou J, *et al*. Artificial intelligence for screening of multiple retinal and optic nerve diseases. Jama Netw Open. 2022;5(5):1–12. https://doi.org/10.1001/jamanetworkopen.2022.9960.
16. Hu J, Wang H, Wu G, Cao Z, Mou L, Zhao Y, *et al*. Multi-scale interactive network with artery/vein discriminator for retinal vessel classification. IEEE J Biomed Health Inform. 2022;26(8):3896–3905. https://doi.org/10.1109/JBHI.2022.3165867.
17. Han Y, Li W, Liu M, Wu Z, Zhang F, Liu X, *et al*. Application of an anomaly detection model to screen for ocular diseases using color retinal fundus images: design and evaluation study. J Med Internet Res. 2021;23(7):1–29. https://doi.org/10.2196/27822.
18. Arsalan M, Haider A, Choi J, Park KR. Diabetic and hypertensive retinopathy screening in fundus images using artificially intelligent shallow architectures. J Pers Med. 2021;12(1):1–7. https://doi.org/10.3390/jpm12010007.
19. Chen JS, Coyner AS, Ostmo S, Sonmez K, Bajimaya S, Pradhan E, *et al*. Deep learning for the diagnosis of stage in retinopathy of prematurity: accuracy and generalizability across populations and cameras. Ophthalmol Retina. 2021;5(10):1027–1035. https://doi.org/10.1016/j.oret.2020.12.013.
20. Abbas Q, Qureshi I, Ibrahim ME. An automatic detection and classification system of five stages for hypertensive retinopathy using semantic and instance segmentation in DenseNet architecture. Sensors. 2021;21(20):6927–6936. https://doi.org/10.3390/s21206936.
21. Arsalan M, Owais M, Mahmood T, Cho SW, Park KR. Aiding the diagnosis of diabetic and hypertensive retinopathy using artificial intelligence-based semantic segmentation. J Clin Med. 2019;8(9):1–14. https://doi.org/10.3390/jcm8091446.
22. Akbar S, Akram MU, Sharif M, Tariq A, Ullah Yasin U. Arteriovenous ratio and papilledema based hybrid decision support system for detection and grading of hypertensive retinopathy. Comput Methods Programs Biomed. 2018;154:123–141. https://doi.org/10.1016/j.cmpb.2017.11.014.
23. Alwakid G, Gouda W, Humayun M. Deep Learning-based prediction of Diabetic Retinopathy using CLAHE and ESRGAN for Enhancement. Healthcare. MDPI. 2023;11(6):863. https://doi.org/10.3390/healthcare11060863.
24. Abbas Q, Daadaa Y, Rashid U, Sajid MZ, Ibrahim ME. HDR-EfficientNet: A Classification of Hypertensive and Diabetic Retinopathy Using Optimize EfficientNet Architecture. Diagnostics. 2023;13(20):1–30. https://doi.org/10.3390/diagnostics13203236.
25. Yadav S, Dhage S. TE-CapsNet: time efficient capsule network for automatic disease classification from medical images. Multimed Tools Appl. 2023;3:1–30. https://doi.org/10.1007/s11042-023-17458-4.
26. VICAVR (Varpa Images for the Computation of the Arterio/Venular Ratio), database. 2009.
27. Niemeijer M, Xu X, Dumitrescu AV, Gupta P, Ginneken BV, Folk JC, *et al*. Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs. IEEE Trans Med Imaging. 2011;30(11):1941–50. https://doi.org/10.1109/TMI.2011.2159619.
28. Opoku M, Weyori BA, Adekoya AF, Adu K. CLAHE-CapsNet: Efficient retina optical coherence tomography classification using capsule networks with contrast limited adaptive histogram equalization. PLOS ONE. 2023;18(11):1–22. https://doi.org/10.1371/journal.pone.0288663.

نموذج الشبكة العصبية العميقة القائم على EffiCapsNet-AV للكشف الآلي عن اعتلال الشبكية الناتج عن ارتفاع ضغط الدم في صورة قاع العين

سوميار، كالبانا

قسم علوم الكمبيوتر والهندسة، جامعة بودوتشيري للتكنولوجيا، بودوتشيري، الهند.

المستخلص

يتجلى اعتلال الشبكية الناتج عن ارتفاع ضغط الدم (HR)، وهو نتيجة خطيرة لارتفاع ضغط الدم الجهازي، من خلال تغييرات محددة في الأوعية الدموية الشبكية يمكن ملاحظتها في صور قاع العين. في البداية، قد لا يعاني المرضى الذين يعانون من ارتفاع ضغط الدم من أي أعراض أو شكاوى. ومع ذلك، بمرور الوقت، تؤثر هذه الحالة تدريجيًا على أعضاء مختلفة في الجسم، بما في ذلك العينين. يمكن أن يعزز الكشف المبكر عن اعتلال الشبكية الناتج عن ارتفاع ضغط الدم من خلال الطرق الآلية التشخيص والإدارة، مما يخفف من خطر حدوث مضاعفات شديدة في العين والجهاز. تعتمد طرق التشخيص التقليدية على التقييم الذاتي لصور الشبكية من قبل أطباء العيون ذوي الخبرة، مما قد يؤدي إلى عدم الاتساق ويقيده قيود الوقت. تقدم هذه الورقة AV-EffiCapsNet، وهو نموذج مبتكر للشبكة العصبية العميقة يدمج قدرات EfficientNet وCapsule Networks للكشف الآلي عن اعتلال الشبكية الناتج عن ارتفاع ضغط الدم في صور قاع العين. تقدم EfficientNet إطار عمل شبكة عصبية ملتوية قابلة للتطوير وفعالة، بينما تعمل شبكات الكبسولة على تحسين تمثيل التسلسلات الهرمية المكانية والعلاقات بين الأجزاء والكل. تم تدريب نموذج EffiCapsNet-AV واختباره على مجموعة بيانات واسعة مثل VICAVR وINSPIRE AVR لصور قاع العين الموضحة، مما يظهر دقة فائقة بنسبة 98.8٪ ودقة 97.7٪ وتذكر 95.5٪ مقارنة بالنماذج الحالية. تشير هذه النتائج إلى أن EffiCapsNet-AV فعال للغاية في اكتشاف العلامات الدقيقة لاعتلال الشبكية الناتج عن ارتفاع ضغط الدم، مما يجعله أداة قيمة للفحص السريري والطب عن بعد.

الكلمات المفتاحية: EffiCaps-AV ، تحليل نسبة الشريان إلى الوريد، التعلم العميق، اعتلال الشبكية الناتج عن ارتفاع ضغط الدم، تحليل الصور الطبية.