

RESEARCH ARTICLE

OPEN ACCESS

REAL-TIME NAIL-FOLD CAPILLARY DETECTION USING YOLO-RELAN WITH SALIENCY-BASED EXPLAINABILITY MODEL

Santhoshkumar S P¹, Parvathi R², Sivaprakash P³, Vimala Ithayan J⁴, Dilip Charaan R M⁵, Vijayanandh T⁶

¹Assistant Professor (SG), School of Computing, Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Chennai, India.

²Associate Professor Grade 1, School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Katpadi, Vellore, India.

³Professor, School of Electrical and Communication, Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Chennai, India.

^{4,5}Associate Professor, School of Computing, Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Chennai, India.

⁶Associate Professor, School of Electrical and Communication, Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Chennai, India.

¹<https://orcid.org/0000-0001-8531-759X>, ²<https://orcid.org/0000-0001-7109-134X>, ³<https://orcid.org/0000-0002-2777-8889>

⁴<https://orcid.org/0000-0002-4647-0863>, ⁵<https://orcid.org/0009-0002-1332-4592>, ⁶<https://orcid.org/0000-0002-6307-4272>

E-mail: *spsanthoshkumar16@gmail.com, 2005.parvathi@gmail.com, sivabtl85@gmail.com, vimala.ithayan@gmail.com, dilip_charraan@yahoo.co.in, drvijayanandht@veltech.edu.in

ARTICLE INFO

ABSTRACT

Article History

Received on: January 27, 2026

Reviewed on: February 27, 2026

Accepted on: July 01, 2026

Published on: July 31, 2026

Keywords:

Nail-fold capillaries,
YOLOv12,
ReLAN,
Saliency Heatmaps,
DL.

The accurate identification of nail-fold capillaries is essential for the early detection of systemic diseases such as microvascular abnormalities and systemic sclerosis. The complexity of capillary design and variations in image quality present significant challenges for precise identification in automated detection systems. An enhanced Deep Learning (DL) architecture that incorporates morphological preprocessing, shape-based feature enhancement, and a robust object recognition foundation. To emphasize essential vascular features, shape focus and extraction modules are performed after input images have previously undergone preprocessing with morphological filters to decrease noise and maintain capillary structures. YOLOv12 is utilized to accomplish the detection task using an enhanced backbone architecture based on Residual-Expanded Layer Aggregation Network (ReLAN), which improves feature reuse, gradient flow, and spatial representation. While significantly improving detection accuracy, this backbone maintains real-time functionality. To provide transparency and clinical reliability, the system makes use of Explainable AI (XAI) tools, especially saliency heatmaps, which highlight the critical areas influencing model predictions and provide interpretability for medical experts. The proposed approach achieves the highest accuracy of 88%, 90% of precision, 89% of recall, 86% of map-50 and 88% of map50-95. These results of the experimental evaluation demonstrate that a reliable and understandable solution for nail-fold capillary detection in clinical applications were achieved using the proposed approach.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

The microvascular structure in nail-fold capillaries is capable of being directly and noninvasively assessed using Nail-Fold Capillaroscopy (NFC), a conventional technique for health evaluation. In order to sustain microcirculation and enable material exchange throughout the body, microvascular integrity is essential. According to earlier research, microcirculatory alterations result from conditions like diabetes, rheumatoid arthritis, Raynaud's phenomenon, and systemic sclerosis (SSc) [1]. SSc is a disease caused by vascular lesions and autoimmune disorders that limit blood flow and induce organ fibrosis throughout the body. NFC is a non-invasive, easy-to-use, and economical method that offers vital information on microvascular anomalies in living tissues. In the past, the NFC has been used to investigate the microvascular changes in nail-fold capillaries linked to different rheumatologic disorders. The growing body of research on NFC in rheumatic disorders demonstrates the technique's effectiveness in clinical settings as well as the variety of research applications.

Additionally, the NFC has been useful in studying other systemic conditions including schizophrenia and diabetes mellitus that are marked by changes in capillary architecture [2-4]. The nail-fold bed's microvascular involvement is a reflection of the body's general microvascular damage and is directly correlated with the target organs' involvement. In addition to enabling the early diagnosis of connective tissue diseases, primarily SSc by differentiating between primary and secondary Raynaud's phenomenon, NFC also offers information on the disease's prognosis based on the degree of target organ damage in these conditions [18]. Nail-fold capillary abnormalities occur in 70–80% of SSc patients and 9–27% of non-SSc individuals, respectively [5-7]. The NFC method captures precise images of nail-fold capillaries using a capillaroscope. Following the acquisition of these images, clinicians use their critical experience to visually analyze and manually measure a variety of factors, such as the existence of dilated capillaries, capillary morphology, and capillary density, which is measured as the number of loops per millimeter [8]. Although this clinical analysis, which is motivated by their experience, is essential, it is still very difficult to provide uniform and standardized assessments among healthcare providers. Although attempts have been made to automate capillary image analysis, results have not yet improved due to capillary size variability [9-11]. This variation in size restricts the training of machine learning models and ultimately leads to less-than-ideal prediction performance, which cause patients to receive a delayed or incorrect diagnosis [12]. The present research proposed an improved deep learning-based technique for automated nail-fold capillaroscopy image processing in order to overcome this restriction.

This method used object detection algorithms to automatically identify and categorize the morphology of individual nail-fold capillaries, which are important characteristics for medical diagnostics in general and early SSc diagnosis in particular since they indicate the involvement of internal organs [13-15]. Additionally, the proposed method's improved deep learning framework demonstrated that it could identify capillaries in multi-scale, untrained microscopic images. This research work includes preprocessing, model selection and explainable AI stages to evaluate the process for the detection of real time nail-fold capillaries. In order to discover nail-fold capillaries, the image is prepared and refined using preprocessing techniques such shape focus, shape extraction, and morphological filtering, which facilitate the identification and analysis of capillary structures. These procedures are intended to attract attention to the pertinent shapes, improve the image quality, and increase the precision and effectiveness of capillary identification [16-18]. The preprocessed image is then fed to model selection stage where it is capable of accurately identifying and localizing capillary structures in microscopic images [19]. While conventional DL models are capable of detecting nail-fold capillaries that are often suffer from limitations such as restricted accuracy, poor generalization across diverse image qualities and difficulty in capturing the intricate vascular structures and variations present in capillary patterns and table 1 illustrates the DL methods to detect nail-fold capillaries [20].

Table 1: Related works for DL methods to detect Nail-fold capillaries.

S. No	Author/ Year	Methodology	Merits	Demerits	Performance metrics
1.	Hao Yin et al [2023] [21]	YOLO v5	- It leverages the strength of DL to automatically identify small and complex vascular structures. - It is suitable for both clinical and research applications.	The model's performance decreases, when dealing with overlapping, blurred or low-contrast capillaries and it requires high quality annotated datasets for optimal training	YOLOv5 attains an accuracy of 85.2%.
2.	Vincenzo Taormina et al [2023] [22]	Faster R-CNN	- It increases image clarity by boosting contrast and suppressing noise, which is critical for highlighting the delicate and fine structures of nail-fold capillaries. - It reduces the need for manual intervention, enhances reproducibility and supports more objective clinical evaluations.	The computational complexity of Faster R-CNN makes the process slower and more resource-intensive compared to single stage detectors.	The method achieves an accuracy of 96.2%
3.	Shupin Chen et al [2023] [23]	U-Net	- It provides an effective approach for accurately delineating capillary structures in microscopic images. - It allows for precise pixel-level segmentation capturing both global context and fine details of the capillary loops.	The model's performance heavily depends on the availability of high-quality, annotated datasets and it is not generalize well to images captured under different imaging conditions.	The performance metrics of U-Net accomplishes an accuracy of 94.1%
4.	Chiara Filippini et al [2021] [24]	CNN	- CNNs are capable of learning hierarchical features directly from images, enabling them to detect subtle morphological changes in capillaries that is used to indicate microvascular diseases. - It reduces the reliance on manual assessment, minimizes observer bias and allows for rapid and objective analysis.	It requires high quality datasets for effective training and limited data led to overfitting for new patients or imaging conditions.	CNN attains an accuracy of 86%
5.	Indriani et al [2025] [25]	ResNet	- It allows the network to learn intricate vascular patterns and subtle capillary abnormalities with high precision. - It improves diagnostic consistency and reduces manual intervention by clinicians.	The architecture is computationally intensive, requiring high memory and processing power, which limit its use in real-time diagnostic systems.	The performance of ResNet is 87% of accuracy, 88% of precision and recall, and 86% of FI-score.

Source: Authors, (2026).

Existing DL methods to detect nail-fold capillaries have various limitations, hence the proposed YOLOv12-ReLAN model is implemented, which allows for precise detection of individual capillaries, abnormal loops and morphological changes that are indicative of microvascular diseases. Next it is passed to Explainable AI tools, particularly saliency heatmaps, which emphasise the crucial areas impacting model predictions and offer interpretability for medical experts to ensure transparency and clinical reliability. The main objectives of this research includes:

- To develop an automated detection framework which identifies target features from input image dataset, an advanced DL technique is implemented.
- To enhance image quality and feature visibility, preprocessing operations such as shape focusing, shape extraction and morphological filtering is utilized.
- To achieve an optimal balance between detection accuracy and computational efficiency and for precise object detection, YOLOv12-ReLAN is employed.
- To analyze visual interpretation of model's decision-making process, ensuring transparency and reliability in predictions, an explainable AI, saliency heatmap is used.

II. PROPOSED SYSTEM DESCRIPTION

The proposed approach employs an enhanced deep learning framework to improve object detection accuracy by combining YOLOv12 with the Backbone R-ELAN architecture and explainable AI methods. Figure 1 displays the block diagram of the proposed method.

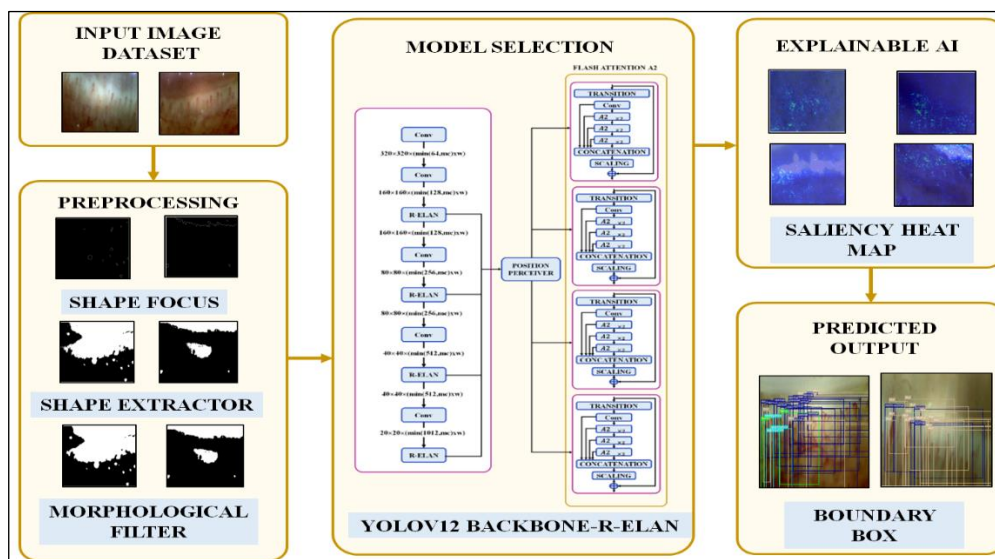


Figure 1: Block diagram of the proposed model.

Source: Authors, (2026).

The input image dataset, which forms the basis for the model's training and evaluation, is where the process commences. Several image improvement and feature refinement approaches, such as Shape Focus, Shape Extraction, and Morphological Filtering, are used during the preprocessing step. By highlighting important structural elements and eliminating unwanted noise, these methods ensure the model receives clear and instructive input data. The YOLOv12 Backbone-R-ELAN architecture is used in the Model Selection stage after the refined images have been preprocessed. This model was selected because of its excellent feature extraction capabilities, effective multi-scale object detection, and resilience when dealing with intricate image fluctuations. Then, to increase interpretability and reliability, the Explainable AI module is incorporated. It provides Saliency Heat Maps, which graphically depict the areas that have the most impact on the model's decision-making. Lastly, the system provides accurate localization and an understandable visual explanation of the detection process by outputting the Predicted Results in the form of Boundary Boxes surrounding the identified objects. Thus, the proposed method provides an effective and transparent DL framework for image-based object detection.

III. PROPOSED SYSTEM MODELLING

III.1 PREPROCESSING

Preprocessing is a crucial first step in the nail-fold capillary detection process that sets up the image for accurate analysis. The nail-fold region's fine capillary structures are enhanced and isolated during this phase using specific techniques.

III.1.1 Shape Focus

The nail-fold image's capillary-like structures are more visible due to the Shape Focus method. This stage seeks to emphasize the thin, elongated, and curved characteristics of capillaries while minimizing background noise and irrelevant data. This phase ensures that the crucial structural elements of capillaries are maintained and made more recognizable for further study by sharpening the emphasis on these particular forms.

III.1.2 Shape Extractor

The improved capillary forms are separated and identified from the image using the Shape Extractor process. This procedure focuses on possible capillary areas and then extracts those utilizing textural and geometric characteristics like continuity, length, and curvature. The capillary structures also be separated from the surrounding tissue. This stage's primary objective is to acquire precise depictions of the capillaries for additional processing and analysis.

III.1.3 Morphological Filter

An essential image processing step in nail-fold capillary detection is the morphological filter, which modifies the forms of capillary structures using computational morphological principles to improve their clarity and continuity. Morphology is concerned with the shape or structure of objects, and morphological filtering is especially helpful for improving segmented images by eliminating undesired noise or microscopic projections, filling in tiny holes or gaps, and smoothing out item outlines. Dilation and erosion are the basic processes in this process, and they both make use of a structural element a predetermined shape that dictates where the image is processed. In addition to these, morphological edge detection uses these basic operations such as erosion, dilation, opening, and closing to extract clear and defined edges of the capillaries. By applying the difference between the processed and the original image, the algorithm effectively highlights capillary boundaries. While erosion and dilation alone can enhance edges, they are less effective in filtering out noise; thus, opening and closing are preferred for achieving smoother, cleaner results. Overall, the morphological filter improves the visibility and accuracy of nail-fold capillary detection by refining the shape and connectivity of capillary structures while minimizing background artifacts.

III.2. MODEL SELECTION-YOLOV12-BACKBONE-R-ELAN

YOLOv12 is particularly proficient at detecting nail fold capillaries because it combines efficient feature extraction with precise localization in a single, end-to-end network. The backbone creates hierarchical feature maps by processing the input image through several convolutional layers. These feature maps capture both fine details, such as individual capillary loops, and more general spatial context, such as the overall arrangement of the nail fold. Because nail fold capillaries are very minute and may appear partially occluded, the network must identify subtle differences while keeping background, which makes these multi-scale feature maps crucial. The diagram of YOLOv12-BACKBONE-R-ELAN is depicted in figure 2.

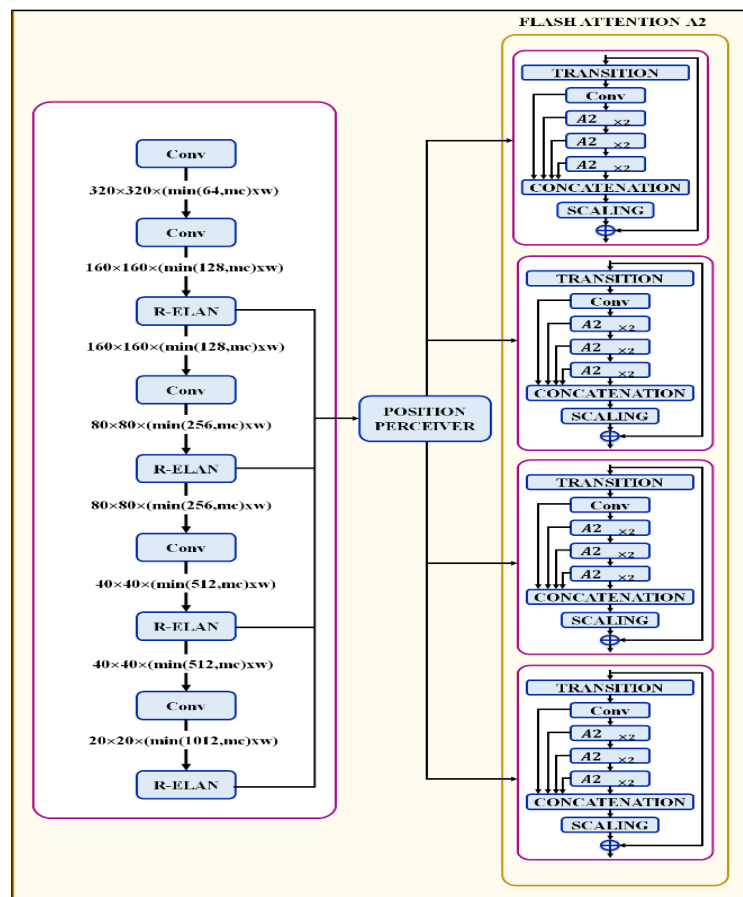


Figure 2: Diagram of YOLOv12-BACKBONE-R-ELAN.

Source: Authors, (2026).

The Residual Efficient Layer Aggregation Network (R-ELAN), which forms the backbone of YOLOv12, is intended to convert unprocessed image data into rich, multi-scale feature maps that constitute the basis for accurate object detection. Deeper residual connections are introduced by the R-ELAN structure to alleviate gradient bottlenecks in deep networks and encourage feature reuse. Pseudo-code of YOLOv12 backbone using R-ELAN is shown below,

Pseudo-code 1: YOLOv12 backbone- R-ELAN.

```

Parameters:
Input Image I(H, W, 3)
Output Feature Maps F3, F4, F5
1: For each epoch do:
    # Stage 1 – Initial Convolution Extraction
2:   For each convolutional layer Lc do:
3:     For each sample in I do:
4:       Compute feature map aijm = ConvBNAct(I)
5:       # Dimension of aijm = (H/2, W/2, 64)
6:     End for
7:   End for
# Stage 2 – Recursive ELAN Block 1 (Low-level Features)
8:   Input aijm → R-ELAN(128, 3 layers)
9:   Split feature into X1, X2
10:  For each layer l = 1 to 3 do:
11:    Y = ConvBNAct(X2, 3×3)
12:    X2 = Concat(X2, Y)
13:  End for
14:  F2 = ConvBNAct(Concat(X1, X2), 128, 1×1)
15:  # Dimension of F2 = (H/4, W/4, 128)
# Stage 3 – Recursive ELAN Block 2 (Mid-level Features)
16:  F3 = R_ELAN(F2, 6 layers, 256)
17:  # Dimension of F3 = (H/8, W/8, 256)
# Stage 4 – Recursive ELAN Block 3 (High-level Features)
18:  F4 = R_ELAN(F3, 9 layers, 512)
19:  # Dimension of F4 = (H/16, W/16, 512)
# Stage 5 – Recursive ELAN Block 4 (Deep Feature Extraction)
20:  F5 = R_ELAN(F4, 3 layers, 1024)
21:  # Dimension of F5 = (H/32, W/32, 1024)
# Stage 6 – Multi-Scale Feature Output
22:  Output F3, F4, F5 to the Neck and Detection Head
23: End for
# R-ELAN Function
Function R_ELAN(input, layers, channels):
    Split input into X1, X2
    For i = 1 to layers do:
        Y = ConvBNAct(X2, 3×3)
        X2 = Concat(X2, Y)
    End for
    Output = ConvBNAct(Concat(X1, X2), channels, 1×1)
    return Output
End Function

```

This characteristic is especially useful for identifying nail-fold capillaries, which are tiny, delicate, and frequently partially obscured. The backbone guarantees that even the smallest capillary patterns are represented in the feature hierarchy by maintaining high-resolution spatial information across several scales. The network's convolutional blocks are defined as a series of smaller, parallel convolutions, formalized by the equation,

$$F_{out} = \sum_{i=1}^n W_i * F_{in} + b_i \quad (1)$$

Where, F_{out} is the feature map that is produced, W_i are the Convolutional filters, F_{in} is the input feature map, and b_i is the phrase for bias. Instead of using big, monolithic convolutions, our method divides processing across several lightweight kernels, enabling the network to effectively extract complex characteristics of nail-fold capillaries while preserving computational efficiency. YOLOv12 uses 7×7 separable convolutions in deeper layers to accurately recognize elongated or convoluted capillary structures without incurring excessive processing costs. These convolutions preserve spatial context with fewer parameters. Table 2 shows the enhancement in YOLOv12 backbone-R-ELAN for detecting nail-fold capillaries.

Table 2: YOLOv12 backbone-R-ELAN for detecting nail-fold capillaries.

Component	Functionality	Innovations in YOLOv12
Backbone-R-ELAN	Extracts detailed, fine- scale features from images of nail-fold capillaries which is crucial for detecting tiny, intricate capillary structures.	The innovations enhance the ability to detect small, detailed features like capillaries while preserving context, even in noisy or low resolution images. These innovations reduce computational cost, making the detection process more efficient.

Source: Authors, (2026).

Multi-scale feature pyramids, which produce feature maps at various resolutions, are used within the backbone to improve the detection of small structures. While lower-resolution maps capture contextual information across greater skin areas, high-resolution maps in the early layers capture fine-grained details of capillaries. Because capillaries can differ in size, density, and orientation, this multi-scale representation is crucial. Thus, YOLOv12's enhanced backbone R-ELAN architecture, which is strengthened by residual aggregation and controlled by the convolutional formulation mentioned above, offers a reliable framework for identifying nail-fold capillaries.

III.3. EXPLAINABLE AI –SALIENCY HEATMAP

Explainable AI (XAI) attempts to overcome the opaque aspect of deep learning models by making their decision-making process interpretable. To ensure reliability and clinical relevance in medical image analysis, such as the identification of nail-fold capillaries, it is essential to comprehend which picture regions contribute to the model's predictions. Saliency heatmaps are a popular XAI technique that effectively highlights the structures the model deems most significant by quantifying the impact of each input pixel on the predicted class. A saliency heatmap mathematically defined as,

$$w = \frac{\partial S_n}{\partial I} \quad (2)$$

In this case, w stands for each pixel's importance weight, and S_n is the score for the particular class n derived from the trained model such as the existence of a nail-fold capillary, and I represents the input image's pixel values. Saliency heatmaps make the small vascular structures that the entire system depends on prominent. These heatmaps make it possible to confirm that the model concentrates on real capillaries rather than unimportant background features by identifying which pixels or small patches contribute most to the classification. This pixel-level interpretability offers a tool for clinical validation and increases trust in the automated detection system. Additionally, saliency heatmaps offer a number of extra advantages when it comes to nail-fold capillary detection:

- **Model refinement:** Researchers identify possible biases or misclassifications by analyzing the parts of the image the model incorrectly highlights, background noise. This realization enables better dataset selection and model retraining to ensure that the network precisely focuses on actual capillary structures.
- **Quality control:** Saliency maps serve as a verification step during clinical analysis to ensure that the model's focus is on the nail-fold area and pertinent vascular characteristics rather than irrelevant skin regions. This improves automated capillary detection's dependability prior to producing diagnostic results.
- **Human-AI cooperation:** Clinicians are able to contrast their own observations of capillary morphology with the saliency maps produced by AI. Clinicians are more confident in the system's predictions because of this alignment between human expertise and AI reasoning.

This explainable AI techniques ensure that the model detects nail-fold capillaries accurately and generates data that doctors are able to comprehend and validate.

IV. RESULT AND DISCUSSION

The proposed approach detects and classifies nail fold capillaries in real-time using YOLO-RELAN into five categories: aberrant, aggregation, blur, hemo, and normal. The model is trained on the ACFD_HT dataset, which provides a diverse collection of high-quality images for robust capillary analysis. This technology helps diagnose diseases like systemic sclerosis by precisely identifying capillary characteristics from high-resolution images using object detection. In order to provide transparency and confidence in clinical applications, the model additionally includes an explainability layer with a saliency heat map that emphasizes the locations in the image that affected the algorithm's decision.

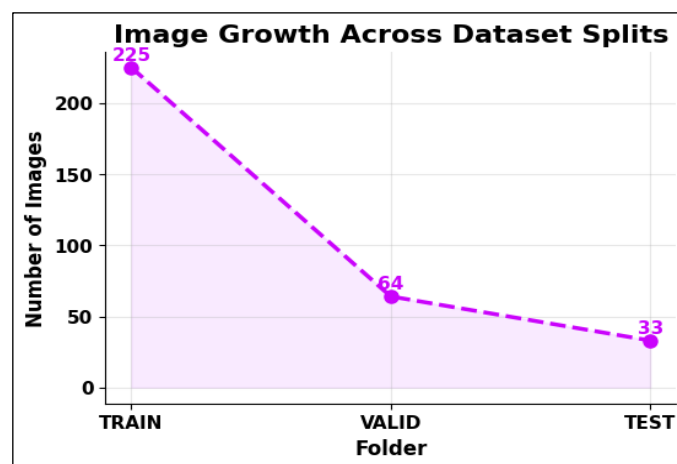


Figure 3: Image growth across dataset splits.

Source: Authors, (2026).

The distribution of images over the three dataset splits such as Train, Validation, and Test is depicted in Figure 3. There are 225 images in the training set, 64 in the validation set, and just 33 in the test set. This illustrates a typical dataset split pattern, in which smaller amounts are set aside for testing and validation to assess performance and generalization, while the majority of the data is used for training to ensure efficient model learning.

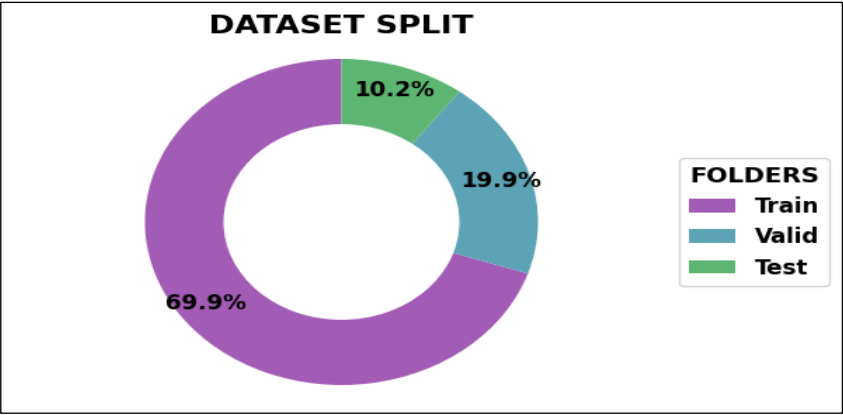


Figure 4: Dataset Split.
Source: Authors, (2026).

The dataset split is depicted in Figure 4, with 69.9% of the photos utilized for training, 19.9% for validation, and 10.2% for testing. While lesser amounts are set aside for performance assessment and final testing, this distribution ensures that the majority of the data enables model learning.

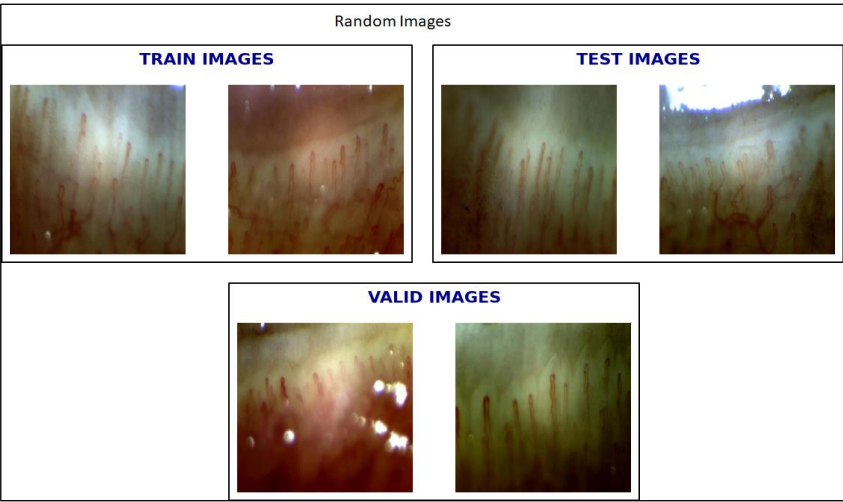


Figure 5: Random images.
Source: Authors, (2026).

Sample images from the training, validation, and testing datasets are displayed in the Figure 5. Clear microscopic images with comparable structural patterns are present in each collection, suggesting consistent data quality across splits. These random samples show that all subsets reflect comparable image properties, ensuring the model's learning and evaluation are balanced.

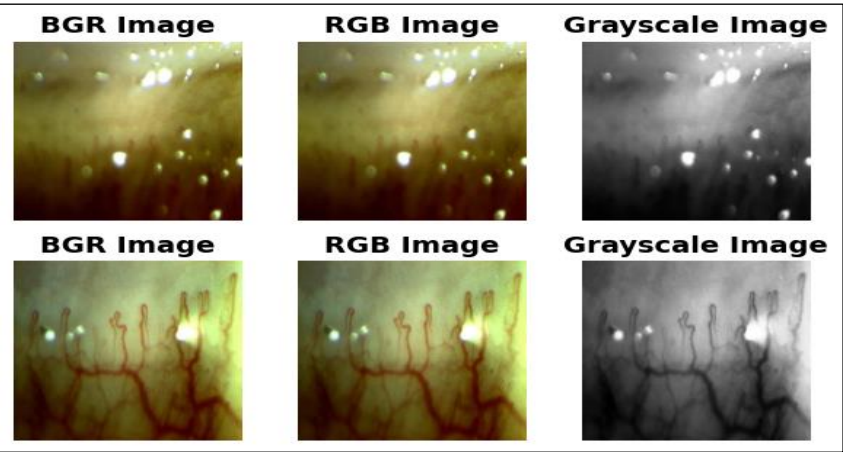


Figure 6: Comparison of BGR, RGB and Grayscale images.
Source: Authors, (2026).

Figure 6 illustrates the visual distinctions between tiny images in BGR, RGB, and grayscale. Grayscale conversion is beneficial for feature extraction and structural analysis since it simplifies the image by simply representing intensity variations, whereas BGR and RGB preserve complete color information.

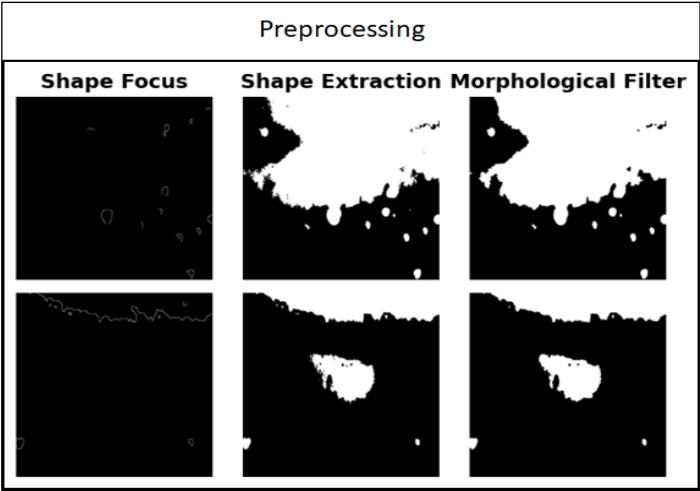


Figure 7: Preprocessing images.
Source: Authors, (2026).

The preprocessing steps, such as shape focus, shape extraction, and morphological filtering, are depicted in Figure 7. By separating pertinent shapes and eliminating noise, these procedures improve the image's structural properties and produce clearer, more defined areas for further analysis.

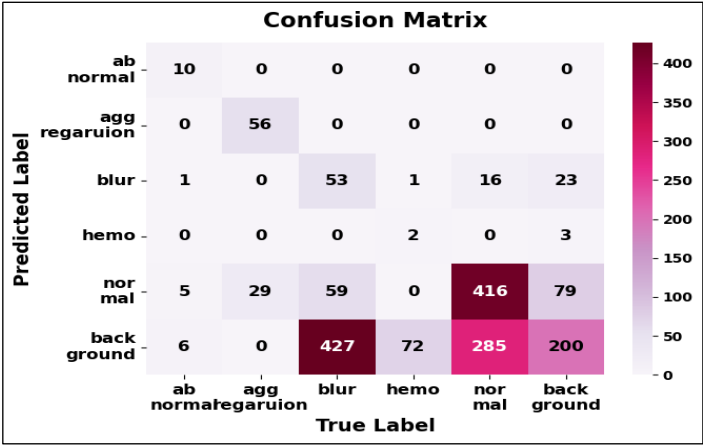


Figure 8: Confusion matrix for YOLOv8.
Source: Authors, (2026).

The confusion matrix displays the performance of a multi-class classification model in six different categories: aberrant, aggregation, blur, hemo, normal, and background in figure 8. Most predictions are focused along the diagonal, particularly for normal (416) and aggregation (56), showing high accuracy in these categories. However, there is significant misclassification between "background" and other classifications such as abnormal (427) and normal (285), indicating that the model struggles to discriminate background data from these categories.

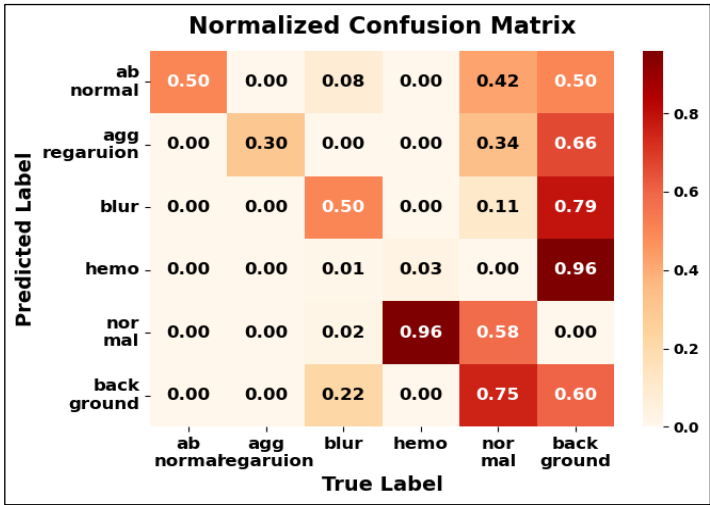


Figure 9: Normalized confusion matrix for YOLOv8.
Source: Authors, (2026).

Figure 9 indicates the normalized confusion matrix which shows the proper categorization for hemo (0.96) and blur (0.79), while background and normal classes display moderate confusion with other categories, indicating that the model performs well in general and needs to improve its ability to differentiate comparable classes.

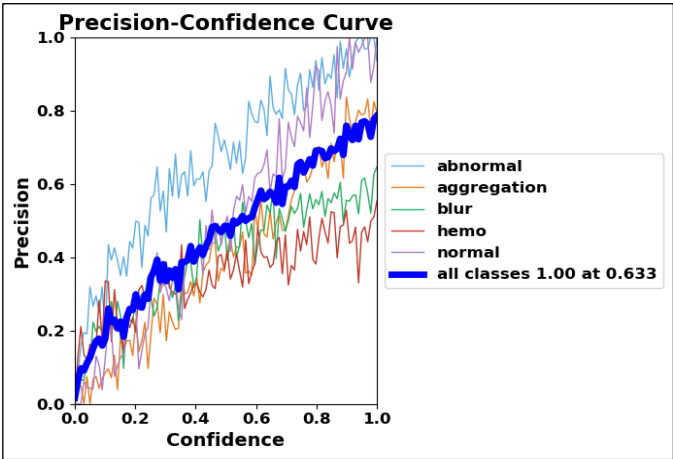


Figure 10: Precision-Confidence curve for YOLOv8.
Source: Authors, (2026).

Figure 10 shows the precision-Confidence Curve across different classes. The x-axis represents the confidence level, which ranges from 0 to 1, while the y-axis denotes precision, which similarly ranges from 0 to one. Each line in the graph represents a separate class: aberrant, aggregation, blur, hemo, and normal. The thick blue line represents the total precision for all classes, with a value of 1.00 at a confidence level of 0.633.

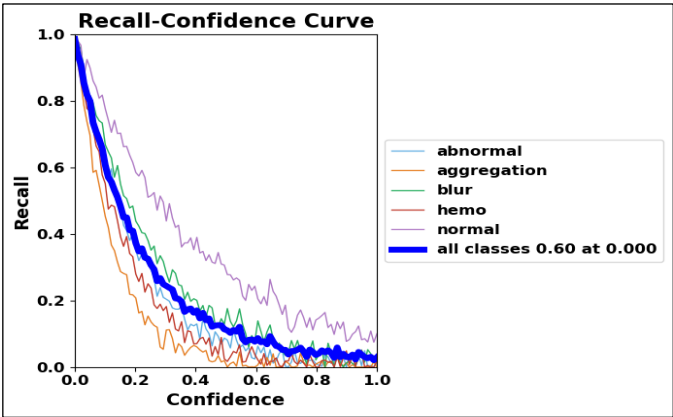


Figure 11: Recall-Confidence curve for YOLOv8.
Source: Authors, (2026).

A Recall-Confidence Curve is displayed in figure 11, showing that recall varies with confidence for various classes. Confidence is represented by the x-axis, and recall is represented by the y-axis. A class is shown by each line: aberrant, aggregation, blur, hemo, and normal. The total recall for all classes is represented by the thick blue line, which has a value of 0.60 at a confidence level of 0.000. Recall generally declines as confidence rises, suggesting that fewer but more accurate predictions are made at increasing confidence levels.

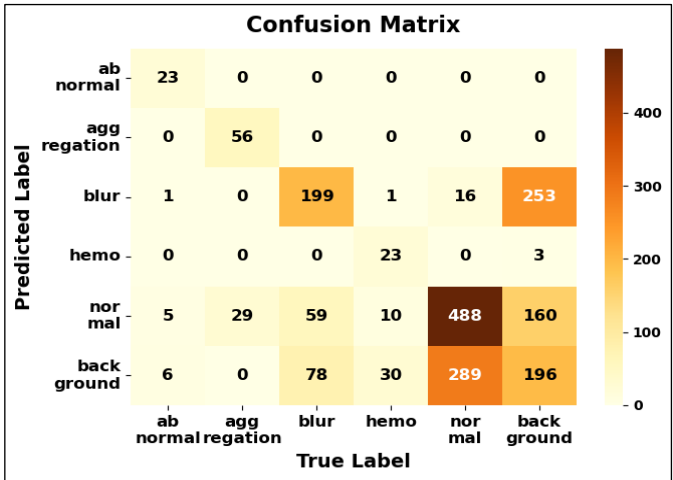


Figure 12: Confusion matrix for YOLOv10.
Source: Authors, (2026).

A confusion matrix depicting a classification model's performance across six classes abnormal, aggregation, blur, hemo, normal, and background is displayed in figure 12. The diagonal cells show samples that have been successfully identified; the normal class has the highest accuracy (488 accurate). Misclassifications, which show some overlap or similarity between these categories, are primarily observed between blur and background classes as well as between normal and background.

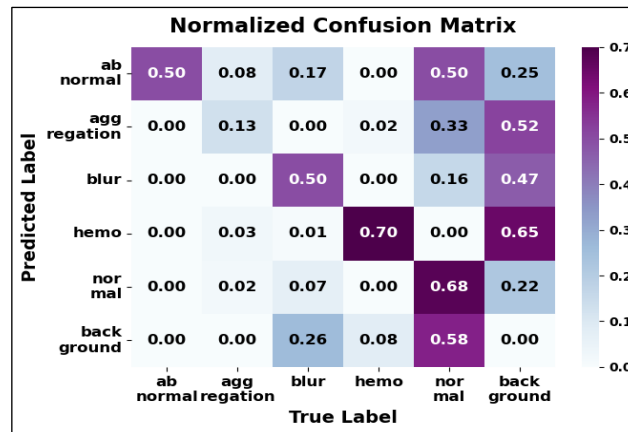


Figure 13: Normalized Confusion matrix for YOLOv10.
Source: Authors, (2026).

The accuracy in which a classification model predicts five categories like abnormal, aggregation, blur, bleeding, and background is displayed in figure 13. It draws attention to the fact that while the model does well for some categories, such as bleeding, it misclassifies others, such as background and abnormal. Prediction accuracy is reflected in color intensity, with darker hues denoting superior performance.

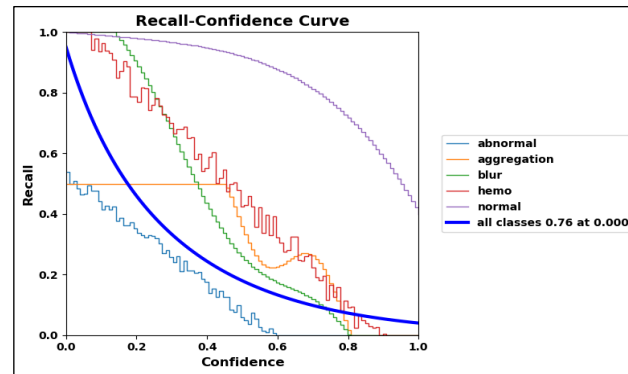


Figure 14: Recall-Confidence curve for YOLOv10.
Source: Authors, (2026).

The connection between model confidence and retrieval for the several classes such as abnormal, aggregation, blur, hemo, and normal is depicted in figure 14. Recall typically declines with increasing confidence, making the model more certain but missing more real cases. The overall performance is summarized by the all classes line, which has an average recall of 0.76 at 0% confidence.

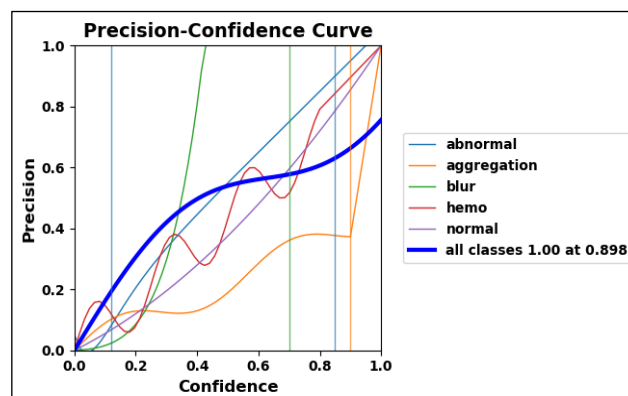


Figure 15: Precision-Confidence curve for YOLOv10.
Source: Authors, (2026).

The precision-confidence curve for the several classes abnormal, aggregation, blur, hemo, and normal is shown in figure 15. As confidence grows, precision typically climbs as well, suggesting more trustworthy forecasts at greater confidence levels. Overall performance is represented by the thick blue line, which indicates that all classes achieve an accuracy of 1.00 at a confidence level of 0.898.

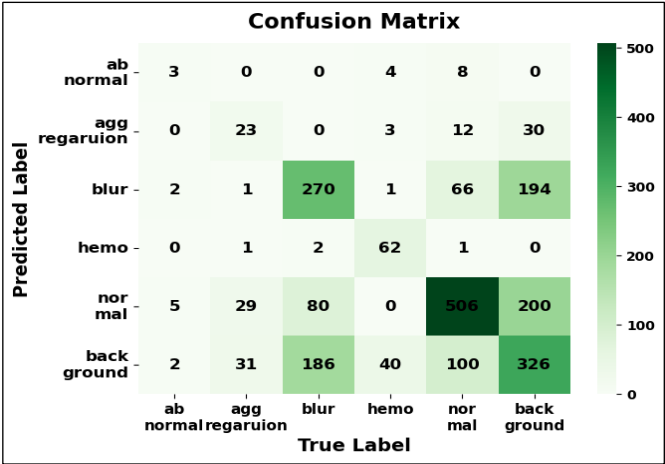


Figure 16: Confusion matrix for YOLOv12.
Source: Authors, (2026).

A classification model's confusion matrix, which shows the accuracy with which it predicted five categories like abnormal, aggregation, blur, hemo, and normal are displayed in Figure 16. Higher values are displayed in darker green in the matrix, which contrasts anticipated and true labels. 270 is the highest score for accurately predicting blur. This matrix aids in assessing the model's accuracy and performance across several classes.

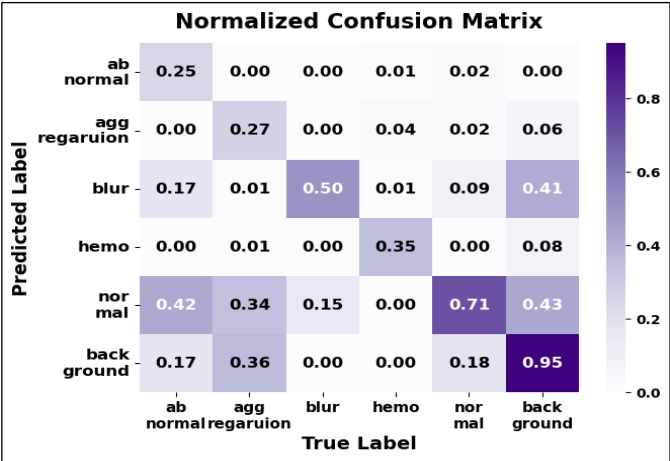


Figure 17: Normalized confusion matrix for YOLOv12.
Source: Authors, (2026).

Figure 17 show the percentage of accurate predictions for each class. Higher proportions are indicated by deeper hues on a scale of 0 to 1. The "background" class, for instance, has a high value of 0.95, indicating precise predictions. By highlighting areas of strong and poor predictions, such as lower values for the abnormal and hemo classes, the matrix aids in evaluating the model's relative performance across classes.

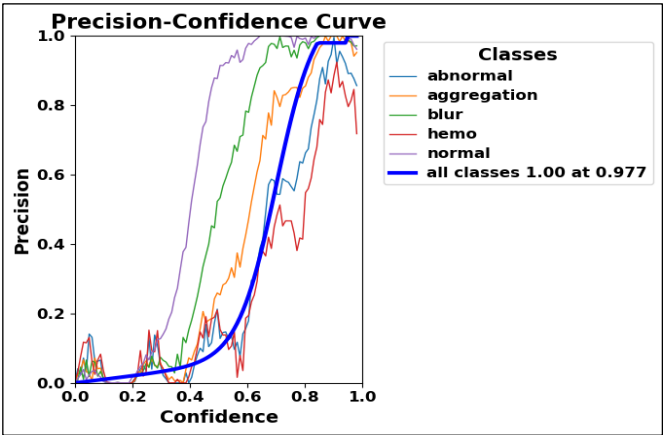


Figure 18: Precision-confident curve for various classes of YOLOv12.
Source: Authors, (2026).

The precision-confidence relationship for five distinct classes such as abnormal, aggregation, blur, hemo, and normal are displayed in Figure 18. The precision for each class is displayed on the y-axis, while the x-axis indicates the confidence level. The blue curve, which reaches a perfect score of 1 at a confidence level of 0.977, shows the overall precision across all classes.

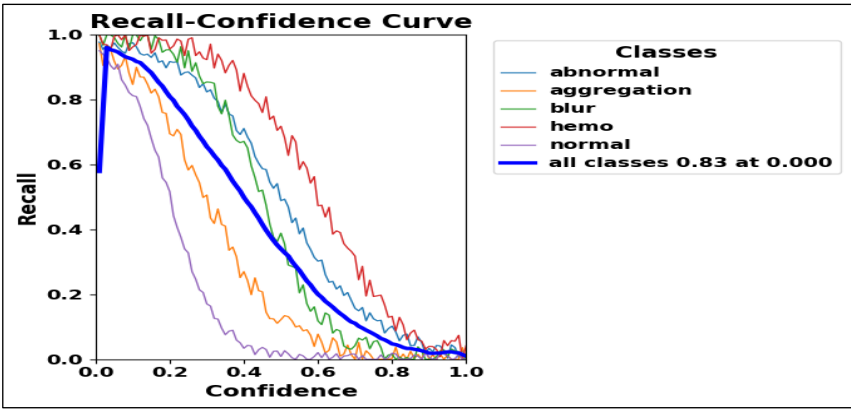


Figure 19: Recall- Confidence curve for different classes of YOLOv12.
Source: Authors, (2026).

The recall-performance for five classes such as abnormal, aggregation, blur, hemo, and normal across various confidence levels is shown in Figure 19. The x-axis displays the degree of confidence, and the y-axis depicts recall. The overall recall for all classes is represented by the blue curve, which begins high and progressively falls as confidence rises before settling at about 0.83 when confidence is zero.

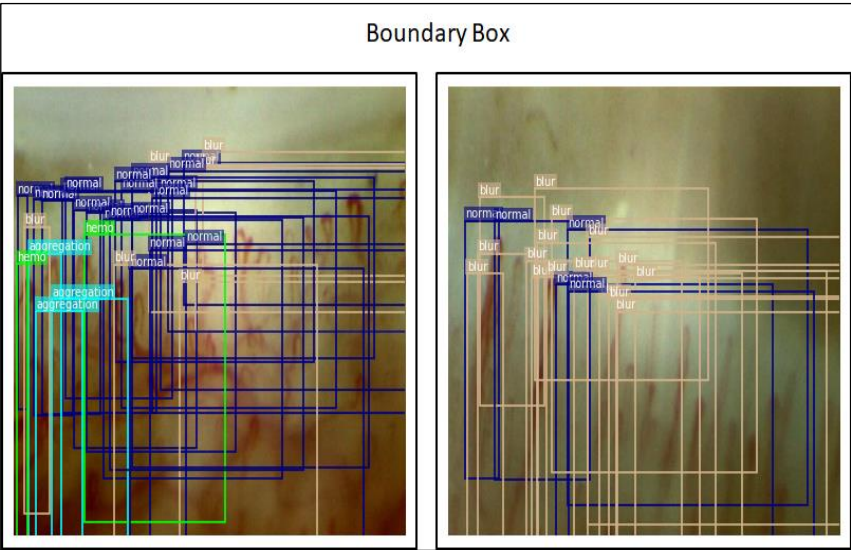


Figure 20: Boundary box.
Source: Authors, (2026).

Two microscopic outlooks Boundary Box are displayed in Figure 20, with labeled bounding boxes delineating various detected locations. The left image includes several overlapping boxes in various types that indicate properties like normal, blur, aggregation, and hemo. The labels in the right image are mostly blurry and have fewer variants. The object detection outputs used to categorize and localize features in digital blood smear images are represented by these bounding boxes.



Figure 21: Image annotations with class labels.
Source: Authors, (2026).

Bounding boxes and class labels are superimposed on a number of sample images in Figure 21. Each bounding box, which is labeled with one of the following classes normal, blur, aggregation, or hemo identifies regions inside the image. The model's predictions for different image segments are shown by the labels and boxes, with some overlapping or inaccurate classifications.

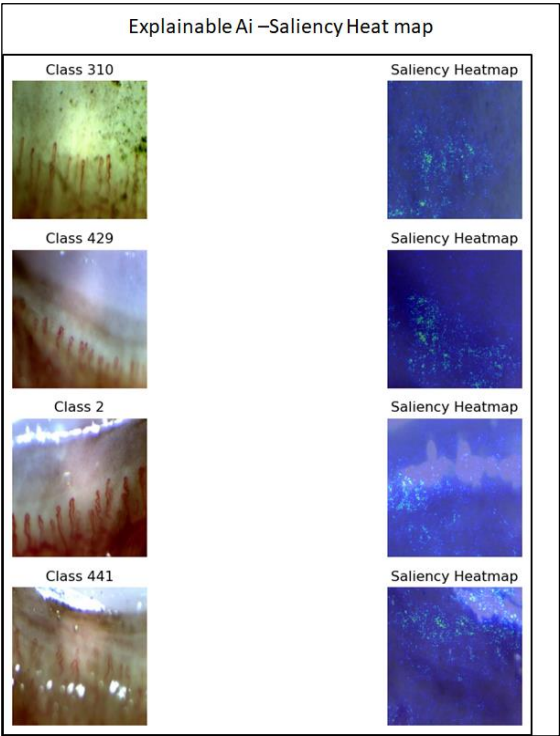


Figure 22: Explainable AI-Saliency heatmap.
Source: Authors, (2026).

Explainable AI-Saliency Heatmaps for four distinct classes are displayed in Figure 22. Every class image on the left has a matching saliency heatmap on the right. Important image regions are highlighted in blue and green hues in the heatmaps, which show the places the AI model concentrated on for categorization. Higher model attention or feature relevance in the decision-making process is represented by brighter patches.

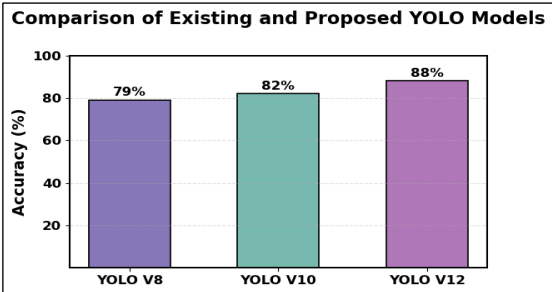


Figure 23: Comparison of existing and proposed YOLO models.
Source: Authors, (2026).

The accuracy of several YOLO models is contrasted in Figure 23. YOLO V8 scores 79%, YOLO V10 improves to 82%, and the suggested YOLO V12 achieves the greatest accuracy of 88%, demonstrating a definite improvement in the model's performance.

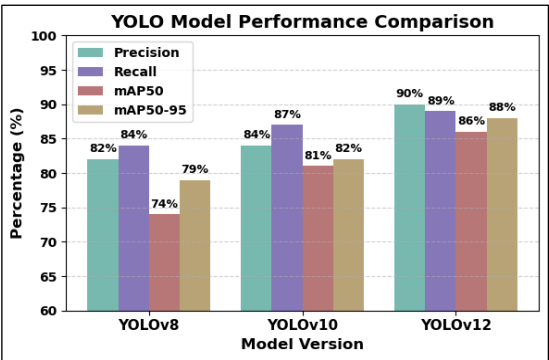


Figure 24: Performance comparison of YOLO models.
Source: Authors, (2026).

The YOLOv8, YOLOv10, and YOLOv12 models are contrasted in Figure 24 according to Precision, Recall, mAP50, and mAP50–95. With 90% precision, 89% recall, 86% mAP50, and 88% mAP50–95, YOLOv12 outperforms YOLOv8 and YOLOv10 in terms of detection accuracy and consistency.

Table 3: Performance metrics of the proposed system.

Model	Accuracy	Precision	Recall	map50	map50-95
PROPOSED	88%	90%	89%	86%	88%

Source: Authors, (2026).

Table 3 shows the proposed approach performs effectively on a number of important metrics. With an accuracy of 88%, it accurately classifies a significant percentage of cases. Additionally, the model has good recall (89%) and precision (90%), indicating that it is both successful in detecting the majority of true positives and accurate in its positive predictions. At an interaction threshold of 0.50, the mean Average Precision (mAP50) is 86%; at different interaction thresholds, the mAP50-95 ranges from 0.50 to 0.95, reaching 88%.

Table 4: Comparative analysis of accuracy for various models.

Model	Accuracy
YOLOv5 [21]	85.2 %
CNN [24]	86 %
ResNet [25]	87 %
PROPOSED	88 %

Source: Authors, (2026).

A comparison of model accuracy for recognizing items and categorization is shown in table 4. With an accuracy of 85.2%, the YOLOv5 model outperforms CNN with 86% and ResNet with 87%. When compared to other deep learning architectures, the suggested model achieves the greatest accuracy of 88%, indicating higher feature extraction and detection performance.

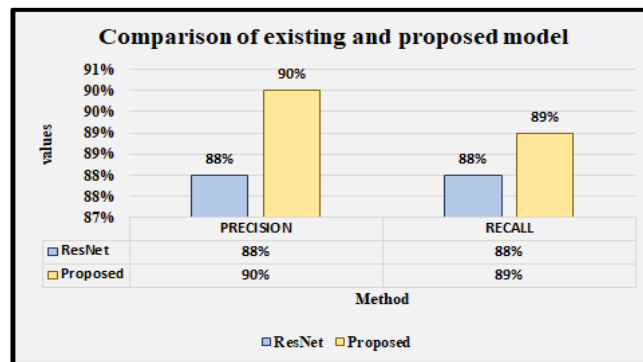


Figure 25: Comparison of existing and proposed model.

Source: Authors, (2026).

The precision and recall levels of the proposed model and the ResNet model are compared in Figure 25. The proposed approach outperforms ResNet, which records 88% for both measures, with 90% precision and 89% recall. This suggests that the proposed framework outperforms the current ResNet architecture in terms of feature identification and classification accuracy and dependability.

V. CONCLUSION

The proposed method effectively combines deep learning, explainable AI, and sophisticated preprocessing methods for the precise and comprehensible identification of nail-fold capillaries. The technique improves the visibility of tiny vascular features that are essential for analysis by immediately preprocessing the input image using shape-focused extraction and morphological filtering. With the R-ELAN backbone, the YOLOv12 model effectively locates and detects nail-fold capillaries in the processed images, producing accurate boundary boxes surrounding the identified areas. Saliency heatmaps are an explainable AI component that emphasizes the particular image regions impacting model predictions, ensuring transparency and clinical dependability. In addition to achieving strong detection performance, the complete approach ensures interpretability and reliability, which are essential for real-world use in clinical and diagnostic applications. The proposed approach achieves the highest accuracy of 88%, 90% of precision, 89% of recall, 86% of map50 and 88% of map50-95.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Methodology: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Investigation: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Discussion of results: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Writing – Original Draft: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Writing – Review and Editing: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Resources: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Supervision: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

Approval of the final text: Santhoshkumar S P, Parvathi R, Sivaprakash P, Vimala Ithayan J, Dilip Charaan R M⁵, Vijayanandh T.

VII. REFERENCES

- [1] M. Ebadi Jalal, O. S. Emam, C. Castillo-Olea, B. García-Zapirain, and A. Elmaghraby, "Abnormality detection in nailfold capillary images using deep learning with EfficientNet and cascade transfer learning", *Scientific Reports*, vol. 15, no. 1, pp. 2068, 2025.
- [2] H.T.P. Nguyen and H. Jeong, "Automated Morphology Detection of Nail-Fold Capillaries Through Enhanced Object Detection Network", *IEEE Access*, 2024.
- [3] M. Kim, "Nail fold capillaroscopy as a potential tool to evaluate breast tumor", *Journal of Analytical Science and Technology*, vol. 15, no. 1, pp. 35, 2024.
- [4] P. M. Jesi, S. Kamatchi, and D. Madhivadhani, "HRSHO: A hybrid rain optimized spotted hyena optimizer for efficient feature selection in CNN-based sinusitis classification", *Biomedical Signal Processing and Control*, vol. 87, pp.105441, 2024.
- [5] M. Pakbin, H. S. MARJANEH, and S. R. Najafizadeh, "Quantitative Nail Fold Capillary Blood Flow Using Capillaroscopy System and ImageJ Software in Healthy Individuals", *Frontiers in Biomedical Technologies*, 2022.
- [6] P. C. Hsu, P. Y. Liao, S. W. Huang, H. H. Chang, J. Y. Chiang, and L. C. Lo, "Nailfold capillary abnormalities as indicators of diabetic nephropathy progression: a cross-sectional study in type 2 diabetes", *Annals of Medicine*, vol. 57, no. 1, pp. 2458766, 2025.
- [7] M. Mahajan, T. Kaur, K. Singh, and B. B. Mahajan, "Evaluation of nail fold capillaroscopy changes in patients with diabetic retinopathy and healthy controls, and its correlation with disease duration, HbA1c levels and severity of diabetic retinopathy: An observational study", *Indian Journal of Dermatology, Venereology and Leprology*, vol. 90, no. 6, pp. 782-788, 2024.
- [8] L. Arya, Y. K. Sharma, S. Nayak, J. Sivakumar, "Comparative Evaluation of Deep Learning Models in Alzheimer's Disease Diagnosis", *Procedia Computer Science*, vol. 258, pp. 2352-2361, 2025.
- [9] K. Miyoshi, M. Chikamori, T. Ando, K. Nakata, T. Aoyama, Y.T. Matsunaga, and T. Yamauchi, "Quantitative image analysis of nailfold capillaries during an in-hospital education program for type 2 diabetes or obesity", *Microvascular Research*, pp.104830, 2025.
- [10] J. Fan, J. Cai, and S. Chen, "Fingernail analysis in diagnosis and management of diabetes mellitus," *Clinica Chimica Acta*, pp. 120490, 2025.
- [11] M. L. Forst, G. Rincon, J. H. Levy, D. N. Cornfield, S. R. Quake, "An inflatable 'finger-lock' for stabilizing nailfold capillary videos and modulating blood flow velocities", *medRxiv*, pp. 2025-02, 2025.
- [12] H. Yin, Y. Wu, P. Guo, J. Luo, J. Lin, B. Zhou, Q. Ye, L. Lin, H. Li, D. Zou, and X. Li, "Panoramic Nailfold Flow Velocity Measurement Method Based on Enhanced Plasma Gap Information", *Journal of Imaging Informatics in Medicine*, pp. 1-16, 2025.
- [13] S.R. Gadu, C.S. Potala, "Advancements in Imaging Techniques for Accurate Identification of VCF in Patients with Scoliosis", In *International Conference on Internet of Things and Connected Technologies*, pp. 233-244. Singapore: Springer Nature Singapore, 2023.
- [14] G.A. Kağızmanlı, T. Aydın, K.Y. Acinikli, R. İşgüder, Z.K. Karabacak, K. Demir, E. Böber, S.E. Ünsal, and A. Abacı, "Nailfold Capillaroscopy: A Non-Invasive Tool for Early Detection of Microvascular Alterations in Children with Type 1 Diabetes Mellitus", *J. Clin. Res. Pediatr. Endocrinol*, 2025.
- [15] A. Assialioui, C. Marco-Pascual, V. Torrente-Segarra, R. Domínguez, N. Santos, J. Peñafiel, X. Juanola, M. Povedano, I. Ferrer, "Microvascular abnormalities in skin capillaries of individuals with amyotrophic lateral sclerosis", *Scientific reports*, vol. 14, no. 1, pp. 24648, 2024.
- [16] A. Asaithambi and V. Thamilarasi, "Classification of lung chest X-ray images using deep learning with efficient optimizers", In *2023 IEEE 13th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0465-0469. IEEE, 2023.
- [17] A. Corrado, N. Mansueto N, M. Correale, V. Rella, L. Tricarico, A. Altomare, N. D. Brunetti, F.P. Cantatore, and C. Rotondo, "Flow Mediated Dilation in Systemic Sclerosis: Association with clinical findings, capillaroscopic patterns and endothelial circulating markers", *Vascular Pharmacology*, vol. 154, pp. 107252, 2024.
- [18] L.R.L. Green, "Clinical presentation and impact of calcinosis in Systemic Sclerosis. An introduction of ultrasound as a valuable outcome measure", PhD diss., University of Leeds, 2024.
- [19] Y. Norimatsu, T. Fukasawa, Y. Kabeya, S. Toyama, K.M. Matsuda, A. Kuzumi A, A. Yoshizaki-Ogawa, H. Ichimura, S. Yonezawa, H. Nakano, and S. Sato, "The number of nail fold capillaries and nail fold bleedings reflects the clinical manifestations of systemic sclerosis", *Journal of Investigative Dermatology*, vol. 145, no. 4, pp. 959-961, 2025.
- [20] L. Zhao, J. Tang, D. Chen, X. Liu, Y. Zhou, Y. Shi, G. Wang, and Y. Wang, "A Comprehensive Dataset and Automated Pipeline for Nailfold Capillary Analysis", In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1-5. IEEE, 2024.
- [21] H. Yin, Z. Wu, A. Huang, J. Luo, J. Liang, J. Lin, Q. Ye, M. Xie, C. Ye, X. Li, and Y. Wu, "Automated nail-fold capillary density measurement method based on improved YOLOv5", *Microvascular research*, vol. 150, pp. 104593, 2023.
- [22] V. Taormina, G. Raso, V. Gentile, L. Abbene, A. Buttacavoli, G. Bonsignore, C. Valenti, P. Messina, G. A. Scardina, and D. Cascio, "Automated stabilization, enhancement and capillaries segmentation in videocapillaroscopy", *Sensors*, vol. 23, no. 18, pp. 7674, 2023.
- [23] S. Chen, D. Wei, S. Gu, and Z. Yang, "Nailfold capillary segmentation based on U-Net and attention mechanism", In *Proceedings of the 2023 3rd International Conference on Bioinformatics and Intelligent Computing*, pp. 337-343, 2023.
- [24] C. Filippini, D. Cardone, D. Perpetuini, A.M. Chiarelli, G. Gualdi, P. Amerio, and A. Merla, "Convolutional neural networks for differential diagnosis of raynaud's phenomenon based on hands thermal patterns", *Applied Sciences*, vol. 11, no. 8, pp. 3614, 2021.
- [25] I. Indriani, "Nail Disease Classification Using Graph Attention Network (GAT) and Resnet", *Ranah Research: Journal of Multidisciplinary Research and Development*, vol. 7, no. 4, pp. 2522-2529, 2025.