International Journal of Advanced Trends in Engineering and Management
(ISSN 2583-7052)

Vol. IV, No. 11, pp. 8-20, November

Automated Detection of Pharyngitis Using Swin Transformer with Hilber-Huang Feature Extraction

Aparna.V[1*], D. Lakshmi[2], D.Jayalakshmi[3]

[1]Assistant Professor, Department of Electronics and Communication Engineering, Dhirajlal Gandhi College of Technology, Opp Salem Airport, Salem-636309. tejusakthi88@gmail.com

[2]Professor, Department of Electrical and Electronics Engineering, AMET University, Chennai. lakshmiee@gmail.com

[3]Assistant Professor, Department of Electronics and Communication Engineering, Arasu Engineering College, Kumbakonam 612501. jailuxmi09@gmail.com

[1*]**Corresponding Author E-mail:** tejusakthi88@gmail.com

ABSTRACT: Pharyngitis is a common inflammation of pharynx that often goes untreated in its early stages, potentially leading to complications if treatment is delayed. This study combines cutting-edge transformer approach with improved image processing techniques to offer precise detection of pharyngitis. Wavelet Transform (WT) based image preprocessing, is applied to the Pharyngitis dataset for decomposing an image into various frequency components. Bayesian based segmentation is utilized for segmenting the pre-processed images into regions by model uncertainty and unpredictability in pixel data. The Hilbert-Huang Transform (HHT), extract discriminative characteristics from pharyngitis throat images. A proposed Swin Transformer model, a cutting-edge vision transformer architecture that uses shifted windows for effective and hierarchical feature encoding. The Swin Transformer is to improve diagnostic accuracy by allowing the model to collect global structural information as well as fine-grained texture patterns. A Swin transformer framework is proposed to enhance the diagnosis of pharyngitis image. Using Python software, the proposed framework achieved higher accuracy and F1-score of 95%, precision and recall of 96%, is accomplished when compared to other techniques.

Keywords: Pharyngitis Disease Detection, Wavelet Transform, Bayesian based segmentation, Hilbert-Huang Transform, Swin Transformer.

1. Introduction

Pharyngitis, commonly known as sore throats is inflammation affecting throat and neck region. In the UK, sore throats are among the most frequent acute conditions seen in primary care, accounting for approximately 10% of all general visits annually [1]. The term “sore throat” generally encompasses both tonsillitis and pharyngitis. Clinical evaluation frequently makes it challenging to distinguish between bacterial and viral causes of tonsillitis, particularly during times when predominance overlaps [2]. In the USA, acute pharyngitis leads to approximately 15 million physician visits each year [3]. Group a streptococcus is found in 20 to 40% cases of childhood pharyngitis (37% in a recent meta-analysis); the remaining cases are viral [4]. Pharyngitis is usually treated with antibiotics to accelerate symptom relief, prevent complications, and reduce the spread of group a streptococcus, while antibiotics are not indicated in viral cases [5]. Whereas, the standard tests for diagnosing pharyngitis is a throat culture in the microbiology laboratory using a throat swab. However, throat culture is impractical for routine care because the result is only available after 24–48 hrs [6]. To address these challenges, various pre-processing, segmentation, feature extraction, and classification techniques are employed for improved diagnosis and differentiation.

The raw images from the original dataset are pre-processed by some of the following techniques. Pharyngitis image is processed using the pre-processing technique such as, image normalization it adjust the pixel values from sore throat image to improve the image quality. Image normalization reduces the variation caused by camera quality models converge faster and perform better. Nonetheless, normalization distorts the original pixel intensity values; reduce performance [7-8]. Additionally, Random Oversampling (RO) is utilized to increase the number of images in the training set in pharyngitis image. In order to enhance the method's performance an equal number of samples in the positive and negative classes are achieved with good performance. Nevertheless, smaller sample duplication causes the model to attain, particularly in small datasets [9]. To overcome the above limitation proposed WT based pre-processing technique is utilized images into various frequency components from pharyngitis image for further Processing.

Some of the following segmentation techniques are used to segment the pharyngitis on throat image. Segmentation based on Watershed algorithm for segmenting the pharyngeal region**,** identifying swollen areas, lesions in the pharyngitis image. Watershed algorithm require accurate boundary detection, perfect for dividing densely populated areas, directly affects the gradients of pixel intensity, useful for input based on markers. However, watershed tends to split images into too many small regions due to noise or minor gradients [10-11]. Consequently, Canny Edge detection algorithm identifies pore boundaries on pharyngitis in throat image. To provide accurate edge positioning, the gradient's magnitude and direction of pharyngitis area are identified. Nevertheless, it is slower and more complicated than simpler edge detectors [12]. To overcome the above techniques the boundaries, region, areas and lesion are segmented using the Bayesian segmentation. Bayesian based segmentation approach is proposed in this research that effectively segments the pharyngitis image.

To extract the features from the pharyngitis on throat image some of the following techniques are utilized. From the segmented pharyngitis images, the prominent features of pharyngitis are extracted by the distribution of edge in certain areas of an image to identify patterns with good performance. Histogram of Oriented Gradient (HOG) is useful for identifying redness, lesions, or structural abnormalities in the pharyngeal region because it shows the directional intensity patterns. Nonetheless, color and intensity fluctuations which are significant in images that are not captured by HOG, it primarily focuses on edge and gradient patterns [13]. Whereas, Improved Bag of Visual Words used as an extraction technique for pharyngitis image. It is used to improve the extraction of the segmented throat image. Yet, computational cost of training is greatly increased by using high-dimensional descriptors and sophisticated clustering [14]. Moreover, features from the pharyngitis images are extracted using Gray Level Co-event Matrix (GLCM), Gray Level Run Length Matrix (GRLM). GLCM and GRLM is capable of identifying indirect variations in throat tissue texture, characteristics that are helpful for classification problems, including as contrast, correlation, homogeneity, energy, and entropy. However, Co-occurrence data gets distorted by noisy throat images, high computational cost [15]. To overcome the above limitations pharyngitis on throat images are not extracting the discriminative characteristics. The proposed feature extraction technique is HHT; it efficiently captures intrinsic mode functions and adaptive frequency components, to extract discriminative characteristics from pharyngeal images.

The pharyngitis and no-pharyngitis on throat image are classified by some of the following techniques. However, to classify the pharyngitis image Deep Learning based on EfficientBO that diagnosis exudative pharyngitis. EfficientBO identifies the pharyngitis and non-pharyngitis image with higher accuracy. EfficientBO its training is adequately calibrated; optimizing accuracy without over fitting, accurate and lightweight. However, being sample-efficient, EfficientBO is costly to evaluate each model, particularly for deep models like EfficientNet [16]. Moreover, an Xception module based CNN for the detection of sore throat or pharyngitis image. Xception based CNN is a depthwise separable convolutions, which improve performance and efficiency by separating spatial and channel-wise feature learning. Reduces the number of parameters whereas preserving high performance by using depthwise separable convolutions [17]. To overcome the limitations of existing classification, a proposed Swin transformer is implemented for transforming the global structural information to get high accuracy and efficiency of pharyngitis prediction.

2. Related Work

Yoo et al (2020) [18] have proposed a Cycle consistency Generative Adversarial Network (CycleGAN) to diagnosis the pharyngitis image. The pragmatic characteristics of pharyngitis were reflected in the synthetic images generated using CycleGAN, increase in the diagnosis accuracy of pharyngitis. Nevertheless, not appropriate for real-time applications without model compression or low-resource situations.

Swathi et al (2024) [19] have presented a combination of Gabor, HOG, Bag of Visual Words and pretrained models SE-ResNet and Inception-V4 (GHBResIncep) for analyzing pharyngitis throat image. Effective treatment planning is possible by early diagnosis, which save lives and reduce complications. GHBResIncep capture multi-scale characteristics, which are helpful for patterns at both the micro and macro levels. Yet, if the dataset is tiny and poorly augmented, there is a greater chance of over fitting.

Demircan *et al* (2025) [20] have proposed a Vision Transformer (ViT) with the K-Nearest Neighbours (K-NN) classifier to enhance the accuracy of pharyngitis in throat disease classification. ViT enhanced global situation, long-range dependencies and improve performance on image classification tasks, especially when trained on large-scale datasets. However, requires large amounts of data to train effectively, it is computationally expensive, demanding significant memory and processing power, particularly during training.

Lai et al (2023) [21] have proposed a TimeSFormer (Time-Space transformer) to classify the pharyngitis of throat image. TimeSFormer treats the spatial and temporal dimensions separately, enabling it to efficiently learn global context across both space and time**.** It is more scalable and interpretable for the pharyngitis region. Nevertheless, computationally intensive, requiring significant memory and processing time.

Warule et al (2024) [22] have proposed a deep neural network-based transformer model for pharyngitis classification. The transformer model efficiently captures the long-term dependencies in the sequence of spectral features using a multi-head attention mechanism. Transformer model have ability to capture long-range dependencies through self-attention mechanisms, allowing the model to understand global context more effectively. However, Transformers are also prone to over fitting when training on small datasets, and their interpretability remains limited.

The contribution of this work is follows:

- WT based pre-processing method shattered into multi-scale frequency components it capture both high-frequency texture variations and low-frequency structure patterns for detecting pharyngitis infection.

- A Bayesian-based segmentation approach is efficiently divide the image into pharyngitis region, by simulating the uncertainty and noise present in pharyngitis images.
- HHT based feature extraction, which efficiently captures intrinsic mode functions and adaptive frequency components, to extract discriminative characteristics from pharyngeal images.
- A Swin Transformer is proposed to improve the classification and detection of pharyngitis images, achieving high accuracy through efficient shifted window based attention.

3. Proposed Work

The proposed block diagram is given in figure 1 for pharyngitis prediction using DL technique with pharyngitis datasets. The proposed work introduces framework for the classification of pharyngitis using throat images by combining advanced signal processing techniques with a state-of-the-art DL model. Initially, the pharyngitis image dataset undergoes pre-processing where the WT is applied to decompose the images into multiple frequency components, effectively capturing both rough structural and fine textural variations. Additionally, a Bayesian-based segmentation approach is utilized to partition the image into clinically meaningful regions. To further enhance feature discrimination, the HHT is employed to extract intrinsic mode functions and adaptive frequency characteristics that reflect complex patterns typically present in inflamed pharyngeal regions.

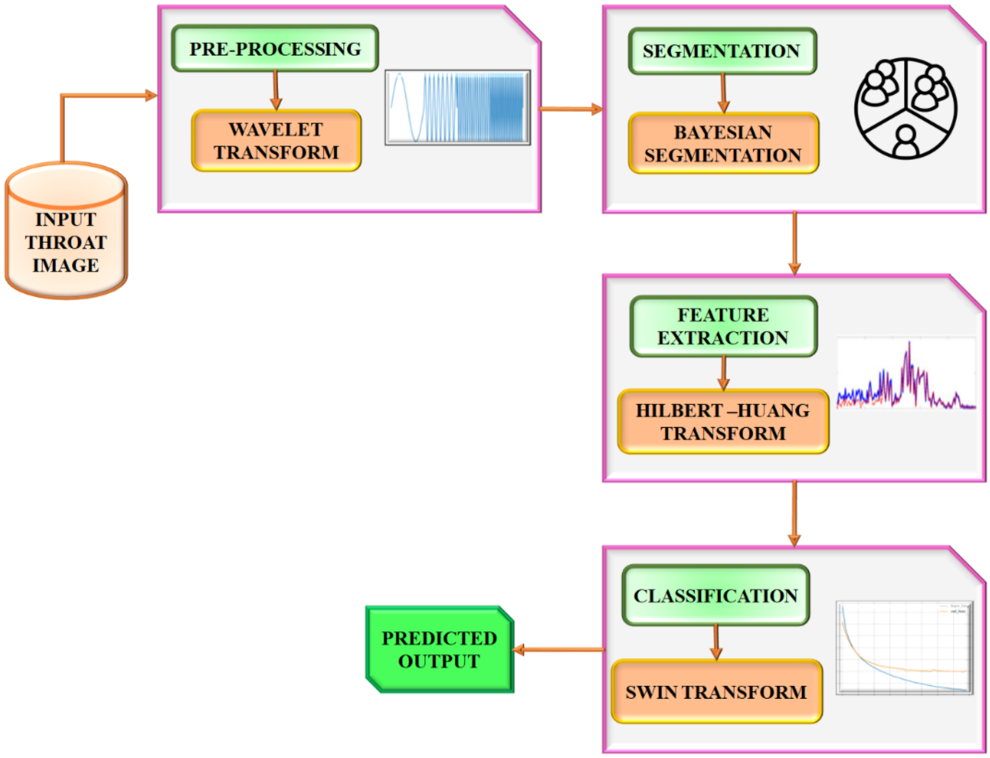

Figure 1: *Block diagram of proposed work*

The resulting features are then fed into a Swin Transformer model for classification. A cutting-edge vision transformer architecture that leverages shifted window-based attention mechanisms. This allows the model to effectively encode both global structural information and local texture patterns in a hierarchical manner. The integration of frequency-aware, uncertaintydriven features with a powerful transformer-based learning framework enables more accurate and robust classification of pharyngitis, making this approach well-suited for automated throat disease screening.

3.1 Pre-Processing for Pharyngitis by Wavelet Transform

Wavelet Transform serves as an effective pre-processing technique for enhancing pharyngitis image analysis. WT decomposes an input throat image into a series of sub-bands at different resolution levels, capturing both spatial and frequency domain information. This multi-resolution decomposition is essential for identifying key visual nods such as inflammation, redness, and tissue texture variations that appear at different scales in pharyngeal images. Unlike traditional filtering methods, the wavelet-based approach preserves both fine-grained details and global structures by isolating low-frequency and high-frequency components. The WT equation is given as

$$W_{\psi}(a,b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}(z)\,dz \qquad (1)$$

Where, $\psi_{a,b}(z) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-a}{b}\right)$. $\psi_{a,b}(z)$ indicates the mother wavelet, a indicates the dilation parameter controlling frequency resolution, and b is the translation parameter controlling time localization. Z Indicates the variable.

The wavelet $\psi_{a,b}(z)$ have been calculated using both translation and dilation from the mother wavelet. This $x(t)$ translation parameter a is called the dilation factor. In the pharyngitis detection, this allows improved localization of pathological patterns that obscured by noise or variability in lighting conditions. The WT output thus enhances the discriminative quality of the image. The processed image is fed to segmentation based on Bayesian segmentation algorithm.

3.2 Segmentation by Bayesian Algorithm for Pharyngitis

Bayesian-based segmentation plays a crucial role in accurately partitioning throat images into pharyngitis regions. The Bayesian segmentation approach models the uncertainty and variability inherent in throat images, particularly those affected by inconsistent lighting, varying color intensity, and anatomical differences. By applying Bayesian inference**,** the algorithm estimates the most probable labelling of each pixel or region based on both prior knowledge and observed image data. This allows the model to differentiate between healthy and swollen areas of the pharynx with greater values in figure 2, even when boundaries are indistinct or textures are irregular.

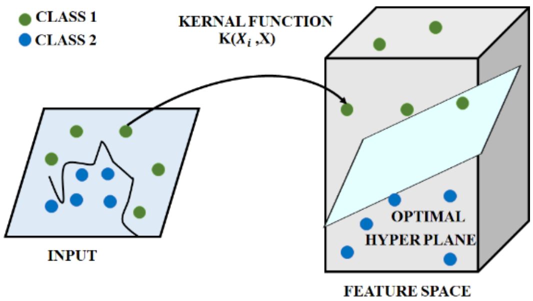

Figure 2: *Basic Bayesian algorithm*

The function f as $y = f(x)$ that, given inputs $X = \{x_1, \cdots, x_n\}$ and their corresponding outputs $Y = \{y_1, \cdots, y_n\}$, produce predictive outputs. A prior distribution be used over the space of functions as $p(f)$, to Bayesian inference. This distribution represents the prior belief of which functions are likely to have generated the data. The likelihood, defined as $(Y|f, X)$, describes the process by which, given a function, an observation is generated. The posterior distribution given the pharyngitis dataset, $p(f|X, Y)$, to the Bayes rule:

The given inputs $X = \{x_1, \cdots, x_n\}$ and their associated output $Y = \{y_1, \cdots, y_n\}$, the function f as $y = f(x)$yields predictive outputs. For Bayesian inference, a prior distribution is applied throughout the space of functions as $p(f)$ by this distribution the prior belief about which functions are most likely to have produced the data is represented. The process by which an observation is produced given a function is described by the likelihood, which is defined as $(Y|f, X)$. The posterior distribution according to the Bayes rule $p(f|X, Y)$, given the pharyngitis dataset:

$$p(f|X, Y) = \frac{p(Y|f,X)p(f)}{\int p(Y|f,X)p(f)df} \quad (2)$$

By integrating across all potential functions f, a new output is anticipated for a new input data set $x*$:

$$p(y*|x*, X, Y) = \int p(y*|f*)p(f*|x*, X, Y)df* \quad (3)$$

By conditioning the model on a finite number of random variables w (the layer weights), the integration operation renders equation (3) unsolvable. Unlike deterministic methods, the Bayesian algorithm adapts to image-specific noise and unpredictability, making it more robust for segmenting complex structures. In the throat of pharyngitis, this leads to more precise isolation of affected regions such as swollen tonsils, red patches, or lesions, which enhances the effectiveness of feature extraction and classification stages. Subsequent, the segmented output is fed to feature extraction technique for pharyngitis in throat image.

3.3 Feature Extraction by Hilbert-Huang Transform for Pharyngitis

The HHT is a powerful diagnostic method for extracting features from non-linear and non-stationary data, it especially effective for pharyngitis detection. HHT involves two major steps: Empirical Mode Decomposition (EMD) and Hilbert Spectral Analysis (HSA). The process begins with EMD, which breaks down the image data into a set of Intrinsic Mode Functions (IMFs) that represent different components. These IMFs are then analysed using the HHT to extract instantaneous pharyngitis features. In EMD, a throat image is decomposed into a finite number of IMFs, each representing oscillatory modes at different image components. The equation is expressed as

$$x(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \quad (4)$$

Where $IMF_i(t)$ is the intrinsic mode function and $r_n(t)$is the final residual. Each IMF is then subjected to the Hilbert Transform to extract instantaneous pharyngitis image components. The Hilbert Transform $H(t)$ of an $IMF_c(t)$ is given by:

$$H(t) = \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{C(T)}{t-T} dT \quad (5)$$

These IMFs are then analysed using the HHT to extract instantaneous pharyngitis throat image features. This provides a detailed pharyngitis representation of image textures, allowing the system to detect important visual signs such as redness, swelling, or irregular surface patterns associated with pharyngitis. By offering an adaptive and data-driven approach to feature extraction, HHT improves the quality and relevance of information passed to the classifier, ultimately enhancing diagnostic accuracy. The extracted throat image is given to the proposed Swin transformer for further classification.

3.4 Classification by Swin Transformer

For pharyngitis diagnosis, the proposed Swin Transformer serves as a powerful transformer model designed to effectively capture both local texture features and global structural information in throat images. Unlike traditional Convolutional Neural Networks (CNNs), the Swin Transformer utilizes a shifted window attention mechanism that processes in figure3 the image in non-overlapping windows, and then shifts them to allow cross-window connections. This approach allows the model to build hierarchical feature representations, making it capable of learning fine-grained details such as inflammation, redness, or lesions, as well as larger anatomical patterns like the overall structure of the pharynx.

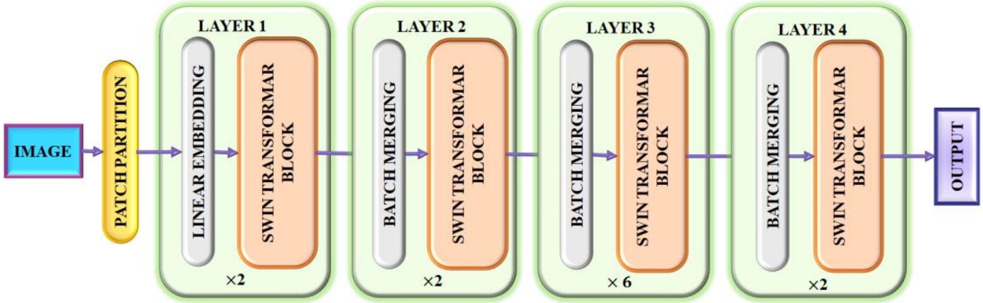

Figure 3: *Basic Swin transformer block for pharyngitis*

In the pharyngitis detection, the Swin Transformer receives features extracted through earlier pre-processing steps (e.g., WT, HHT, Bayesian segmentation) and encodes them in a way that captures both contextual relationships and texture variations. Its multi-stage architecture allows it to operate at different scales, enhancing its ability to identify subtle pathological signs that might be missed by conventional models. As a result, the Swin Transformer significantly improves the accuracy and robustness of pharyngitis classification, making it a highly effective tool in automated medical image analysis systems. The extracted throat image's standard backbone of separate vision is the Swin transformer. Figure 4 illustrates the complexity increases linearly with image sizes as the Swin transformer creates hierarchical feature maps from tiny patches, with some patches remaining fixed in each window.

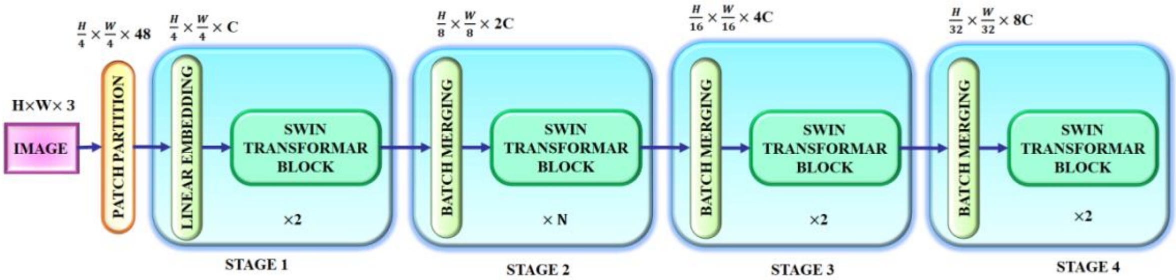

Figure 4: *Swin transformer builds hierarchal feature maps*

Figure 5 shows the shifted windows serve as the foundation for the Swin transformer's construction. The Layer Normalization (LN) layer, multi-head self-attention module, residual connection, and 2-layer Multilayer Perceptron (MLP) with GELU non-linearity are the primary parts of each Swin transformer. Both the shifting window-based multi-head self-attention (SW-MSA) and window-based multi-head self-attention (W-MSA) modules.

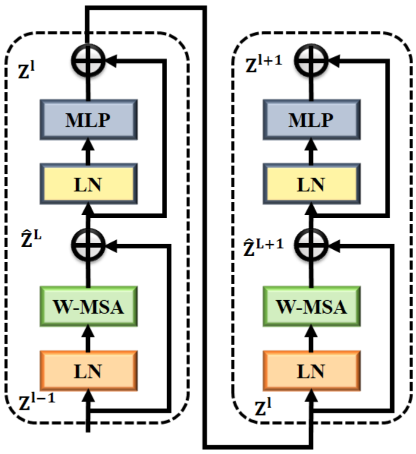

Figure 5: *Construction of Swin transformer is based on the shifted windows*

The outputs of the Multilayer Perceptron (MLP) and shifting window-based multi-head self-attention (S/W-MSA) modules are given in equation

$$Z^l = W - MSA\left(LN(Z^{l-1})\right) + Z^{l-1} \quad (6)$$

$$Z^l = MLP\left(LN(\hat{Z}^l)\right) + \hat{Z}^l \quad (7)$$

Where, W indicates the window, Z^l indicates the output of the current transformer layer l, Z^l indicates intermediate representation, typically the output from the self-attention block and residual connection. LN normalizes the input across the feature dimensions. It utilizes a hierarchical classification approach by splitting features into patches and applying transformer blocks at different levels to handle large image sizes effectively. This makes it perfect for tasks requiring fast and precise processing of high quality images, like object detection and image categorization. Thus the proposed Swin transform performs high classification accuracy and good performance for the throat image.

4. Result and Discussion

In result and discussion section the pharyngitis disease in throat images are predicted using the proposed Swin transform. For identifying pharyngitis disease the pharyngitis dataset is taken from the kaggle.com implemented using Python software. The total train images have the count of 380, and the total test images have the count of 50.

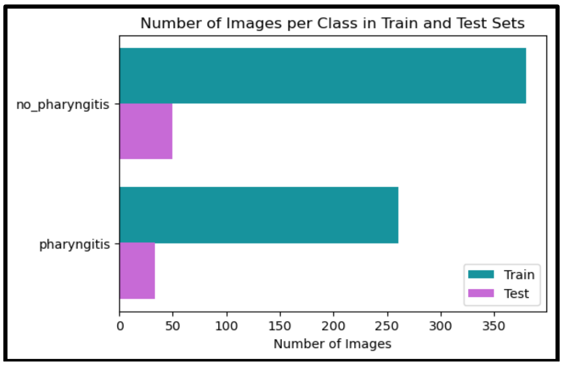

Figure 6: *Number of images per class in train and test sets*

Figure 6 shows the number of images per class in train and test sets for throat images. The total train images have the count of 350, where the no-pharyngitis have the train image of 350 images and test have 50 images. And the pharyngitis have total train images of 261 and test images of 33 images.

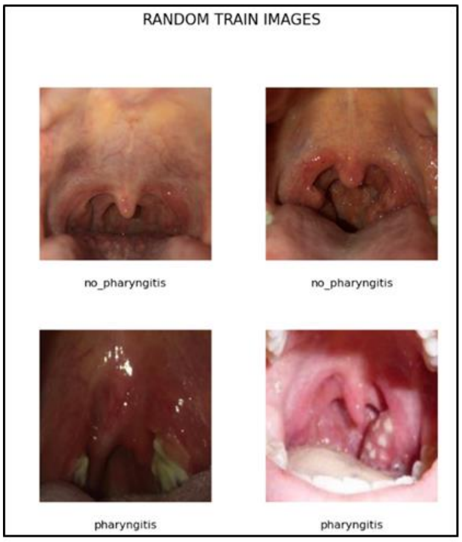

Figure 7: *Random train images*

Figure 7 shows random images for test of pharyngitis and throat image. The random images are taken from the pharyngitis dataset for further process. Here, the throat image have two classes pharyngitis and no-pharyngitis images are randomly chosen from pharyngitis dataset to train for further analysis.

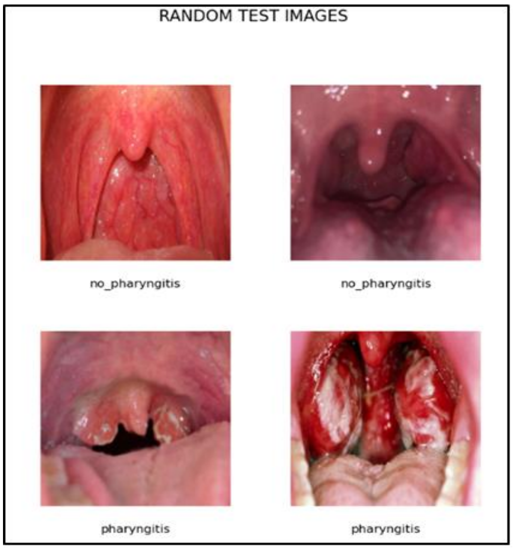

Figure 8*: Random image test*

Figure 8 shows random images for test dataset of pharyngitis and throat image. The random images are taken from the pharyngitis dataset for further process. Here, the throat image have two classes pharyngitis and no-pharyngitis images are randomly chosen from pharyngitis dataset to test for further analysis.

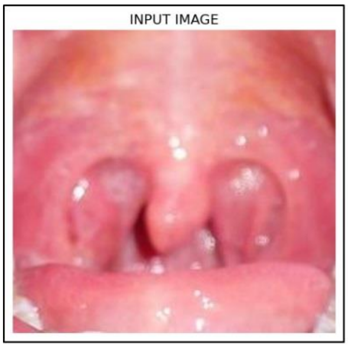

Figure 9: *Input throat image*

Figure 9 shows the input throat image. The input image is from the pharyngitis dataset randomly chosen test and train images. The input throat image contains large pixel size, raw image and blurry. The throat image is a raw image from pharyngitis dataset from kaggle.com uses the following methods such as pre-processing, segmentation, feature extraction and classification for further processes.

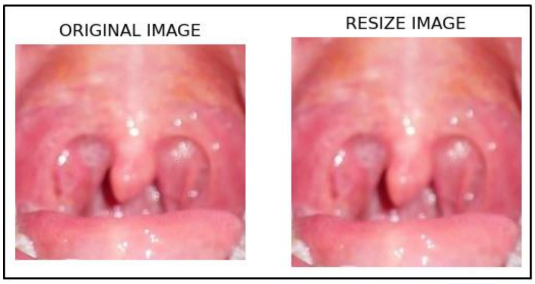

Figure 10: *Input image into resize image*

Figure 10 shows the input image to resize image. Here the input throat image is resized using image resizing technique. The raw input image is converted in to resize image for further process. The noises present in the original image are removed and develop the better quality image. The better quality pharyngitis image is given as an input to grayscale conversion technique; their resize image is converted in to grayscale.

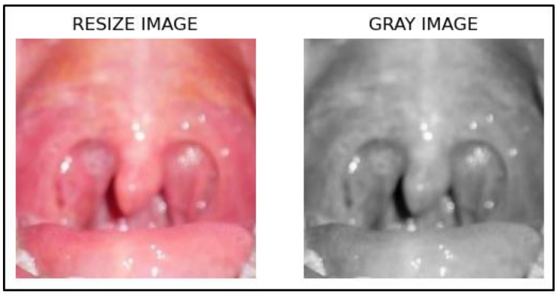

Figure 11: *Resized image into gray scale image*

Figure 11 shows the resized image into gray scale image for throat approach. Here resized image is transformed into grayscale image for enhancing the pharyngitis region. Thus by the gray scale conversion technique the resize image is converted in to gray image. The throat image is converted in to a gray scale or black and white representation for this gray scale image the grayscale conversion technique is used.

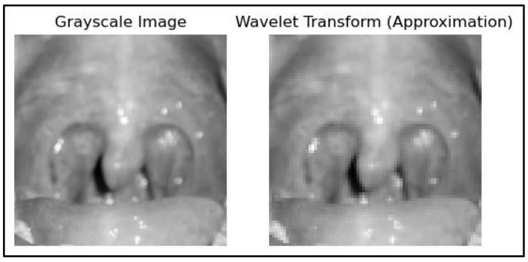

Figure 12: *Grayscale image into wavelet transform*

Figure 12 shows the grayscale image into wavelet transform for pharyngitis in throat image. This process decomposes the image into multiple sub-bands that capture different frequency and orientation information. The captured WT image is given as an input to the segmentation technique. Here the pharyngitis throat image is transformed into frequency component.

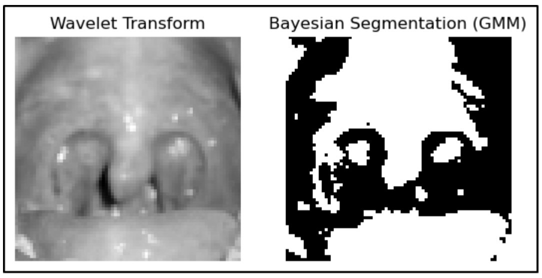

Figure 13: *Wavelet transform to Bayesian segmentation*

Figure 13 shows the wavelet transform to Bayesian segmentation. Local frequency features are extracted via wavelets and are strong to illumination and noise. A probabilistic framework for modelling uncertainty and mixed pixel classes is offered by Bayesian segmentation. The segmented pharyngitis throat image is fed as an input to the feature extraction technique.

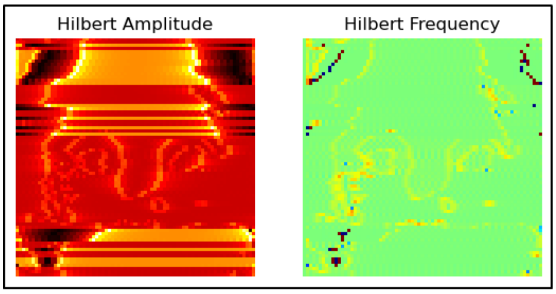

Figure 14: *Hilbert amplitude and Hilbert frequency*

Figure 14 shows the Hilbert amplitude and frequency for pharyngitis in throat image. The amplitude and frequency of HHtis calculated for pharyngitis in throat image. The HHT amplitude is calculated and it is shown in red color, the HHT frequency is calculated and is shown in green color. Both the frequency and amplitude value for HHT is fed to the classification technique.

4.1 Performance Metrics

Performance metrics like accuracy, sensitivity and specificity which indicate the likelihood that pharyngitis be correctly detected and appropriately diagnosed as the probabilities calculated by the swin transform. The performance metrics in this study explained in table 1.

Table 1: *Performance Metrics*

Performance metrics	Values
Accuracy	95
Precision	96
Recall	96
$F1-score$	95

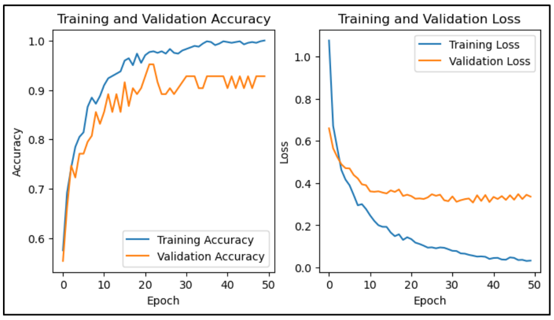

Figure 15: *Training and validation accuracy and loss*

Figure 15 shows the training and validation accuracy and loss. The X-axis denotes the epochs and Y-axis denotes accuracy for throat image. The model accuracy graph illustrates increasing accuracy of 95%, with the model achieving high performance and minimal over fitting. The X-axis indicates the epochs and Y-axis indicates the loss. Training and validation loss is reliably declining in training and validation loss, representing improved learning and reduced error.

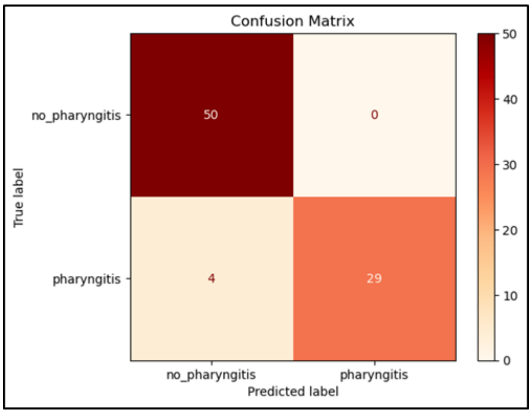

Figure 16: *Confusion matrix for throat image*

Figure 16 shows the confusion matrix for throat image. Proposed method evaluates the model's performance for two classes 'pharyngitis and no-pharyngitis. No-pharyngitis have the value of 50 and pharyngitis have the value of 29 respectively. Here confusion matrix for throat image shows the maroon color indicates the no-pharyngitis and red color indicates the pharyngitis value.

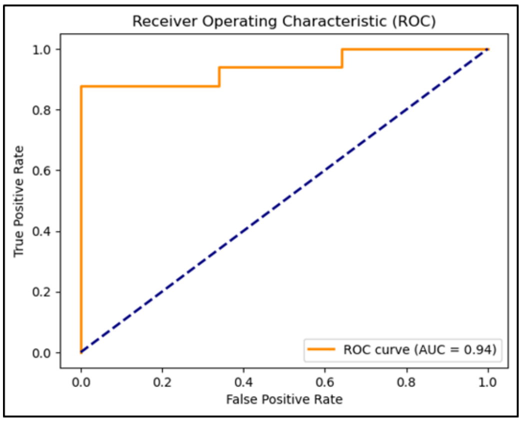

Figure 17: *ROC curve*

Figure 17 shows the multi-class ROC curve for proposed swin transform continues to have the highest AUC value of 0.94 demonstrating the unique overview and flexibility of throat image approach. The X-axis indicates the false positive rate and Y-axis indicates the true positive rates. The ROC curve shows the orange color curve it have the value 0.94 for pharyngitis in throat image.

4.2 Comparison for the suggested method

The comparison of the suggested method with existing method is presented in table 2. The TimeSFormer have the classification accuracy of 90% [21], ViT-KNN have the classification accuracy of 94% [20] and the proposed Swin transform have the higher accuracy of 95% compared to the existing method in table 2. The proposed swin transform have better performance for pharyngitis in throat image.

Table 2: *Comparison for the proposed method*

Method	Accuracy
Times Former	90%
ViT-KNN	94%
Swin transform	95%

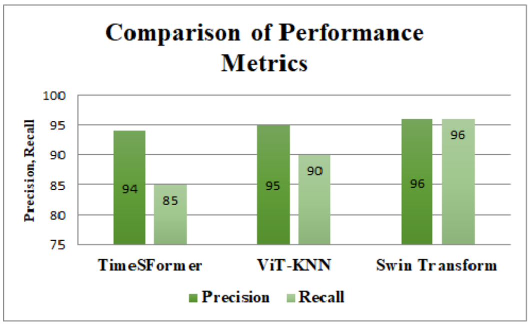

Figure 18: *Comparsion of performance metrics*

Figure 18 shows the comparison of performance metrics for precision and recall value of proposed work and the existing work. The TimeSFormer have precision value of 94% and recall value of 85% [21], ViT-KNN have the precision value of 95% and recall of 90% [20] and proposed swin transform have the precision value of 96% and recall of 96% for identifying pharyngitis in throat image. The proposed precision and recall value is high compared to the existing method.

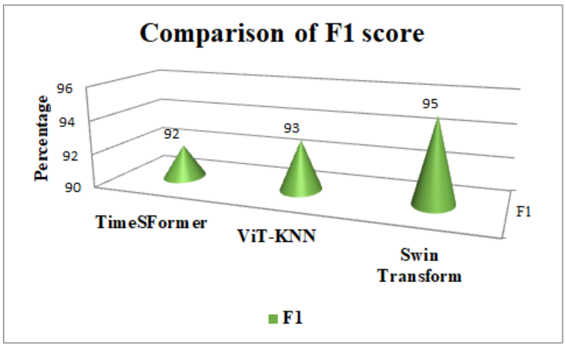

Figure 19: *Comparison of F1 score*

Figure 19 shows the comparison of f1 score value of proposed work and the existing work. The TimeSFormer have F1-score value of 92% [21], ViT-KNN have F1-score value of 93%[20] and proposed swin transform have the F1- score value of 95% for identifying pharyngitis in throat image. The proposed f1 score value is high compared to the existing method.

5. Conclusion

In this paper, swin transformer based model is proposed for the detection and classification of pharyngitis in throat image. The two classes of pharyngitis and no-pharyngitis images are classified successfully using the proposed swin transformer. The preprocessing stage decomposes images into various frequency components using WT it transform image to frequency. Next, Bayesian based segmentation separate a throat image into pharyngitis region for pharyngitis throat image. Furthermore, HHT transforms the efficiently captures intrinsic mode functions and adaptive frequency components from the throat images. Finally, the proposed swin transform classifies the pharyngitis throat image with better accuracy. The performance evaluation, conducted on the pharyngitis dataset, demonstrates the superior capabilities of the swin transform method. The improved accuracy and F1-score of 95%, precision and recall value of 96%, is achieved as a greater accuracy and performance in contrast to the existing methods.

References

1. Nicole Rebusi; Molly L. Paras; Miriam B. Barshak, Year: 2024, ‘Acute Pharyngitis, Tonsillitis, and Peritonsillar Abscess” Infections of the Ears, Nose, Throat, and Sinuses, pp. 253– 273.

2. Lei Zhou; Lineng Zhang; Feng Xu, Year: 2025, “The combined manifestations of dramatically sore throat, congested and edematous mucosa, no-swelling tonsil are specific in acute Omicron pharyngitis” BMC Infectious Diseases, Vol: 25, No/: 1, pp.29.

3. Wesam Mulla; Anan Younis; Sana Zahalka, Anat Wieder; Dafna Yahav; Amitai Segev; Israel Mazin et al, Year: 2025, “Clinical characteristics and outcomes of patients with nonrheumatic streptococcal pharyngitis associated myocarditis” The American Journal of Medicine.

4. Anuk D. Indraratna; Sacha Mytton; Alison Ricafrente; Doug Millar; Jody Gorman; Kristy I. Azzopardi; Hannah R. Frost et al, Year: 2024, “A highly sensitive 3base™ assay for detecting Streptococcus pyogenes in saliva during controlled human pharyngitis”, Talanta, Vol: 276, pp. 126221.

5. Rob Daniels; Esther Miles; Karen Button, Year: 2024, “Does the Addition of Point-of-Care Testing Alter Antibiotic Prescribing Decisions When Patients Present with Acute Sore Throat to Primary Care? A Prospective Test of Change”, Diagnostics, Vol: 14, No: 11, pp. 1104.

6. Johanna Jääskeläinen; Marjo Renko; Ilari Kuitunen, Year: 2025, “Centor scores associated poorly with rapid antigen test findings in children with sore throat”, European Journal of Pediatrics, Vol: 184, No: 1, pp. 1– 5.

7. Nimisha Roy; J. David Frost; M. Mahdi Roozbahani, Year: 2023, “Quantifying three-dimensional bodies and throats of particulate system pore space”, Powder Technology, Vol: 415, pp. 118160.

8. Sara Hosseinzadeh Kassania; Peyman Hosseinzadeh Kassanib; Michal J. Wesolowskic; Kevin A. Schneidera; Ralph Detersa, Year: 2021, “Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach” Biocybernetics and biomedical engineering, Vol: 41, No: 3, pp. 867– 879.

9. Edna Marquez; Eira Valeria Barrón-Palma; Katya Rodríguez; Jesus Savage; Ana Laura Sanchez-Sandoval, Year: 2023, “Supervised machine learning methods for seasonal influenza diagnosis” Diagnostics, Vol: 13, No: 21, pp. 3352.

10. Anindityo Patmonoaji; Kento Tsuji; Tetsuya Suekane, Year: 2020, “Pore-throat characterization of unconsolidated porous media using watershed-segmentation algorithm”, Powder Technology, Vol: 362, pp. 635– 644.

11. Yang Liu; Wenbo Gong; Yu Zhao; Xu Jin; and Moran Wang, Year: 2022, “A pore-throat segmentation method based on local hydraulic resistance equivalence for pore-network modeling”,

Water Resources Research, Vol: 58, No: 12, pp. e2022WR033142.

12. K. Dixit; P. Gupta; S. Kamle; N. Sinha, Year: 2020, "Structural analysis of porous bioactive glass scaffolds using micro-computed tomographic images", Journal of Materials Science, Vol: 55, pp. 12705– 12724.

13. Chiranjit Dutta; Prasad Sandhya; Kandasamy Vidhya; Ramanathan Rajalakshmi; Devasahayam Ramya; Kotakonda Madhubabu, Year: 2024, "Effectiveness of deep learning in early-stage oral cancer detections and classification using histogram of oriented gradients" Expert Systems, Vol: 41, No: 6, pp. e13439.

14. M. Swathi; Rajeshkannan Regunathan, Year: 2024, "RETRACTED ARTICLE: A novel feature fusion based approach for detecting pharyngitis, oral cancer, and tonsillitis using improved bag of visual words", Soft Computing, Vol: 28, No: 2, pp. 741–741.

15. N. Suganthi; K. Sarojini, Year: 2023, "Feature extraction based on glcm and glrm methods on covid-19 dataset", In Computer Vision and Machine Intelligence Paradigms for SDGs: Select Proceedings of ICRTAC-CVMIP 2021, pp. 271– 279.

16. Seo Yi Chng; Paul Jie Wen Tern; Matthew Rui Xian Kan; Lionel Tim-Ee Cheng, Year: 2024, "Deep Learning Model and its Application for the Diagnosis of Exudative Pharyngitis", Healthcare Informatics Research, Vol: 30, No: 1, pp. 42– 48.

17. Seong Jae Lee; Joo Young Ko; Hyun Il Kim; Sang-Il Choi, Year: 2020, "Automatic detection of airway invasion from videofluoroscopy via deep learning technology", Applied Sciences, Vol: 10, No: 18, pp. 6179.

18. Tae Keun Yoo; Joon Yul Choi; Younil Jang; Ik Hee Ryu, Year: 2020, "Toward automated severe pharyngitis detection with smartphone camera using deep learning networks", Computers in biology and medicine, Vol: 125, pp. 103980.

19. M. Swathi; Rajeshkannan Regunathan, Year: 2024, "A multi-fusion approach to classify pharyngitis, tonsillitis and oral cancer with iterative relief feature weighting", Alexandria Engineering Journal, Vol: 104, pp. 359– 370.

20. Furkancan Demircan; Murat Ekinci; Zafer Cömert, Year: 2025, "Enhancing intra-aural disease classification with attention-based deep learning models", Neural Computing and Applications, pp.1–16.

21. Derek Ka-Hei Lai; Ethan Shiu-Wang Cheng; Bryan Pak-Hei So; Ye-Jiao Mao; Sophia Ming-Yan Cheung; Daphne Sze Ki Cheung; Duo Wai-Chi Wong; James Chung-Wai Cheung, Year: 2023, "Transformer models and convolutional networks with different activation functions for swallow classification using depth video data", Mathematics, Vol: 11, No: 14, pp. 3081.

22. Pankaj Warule; Snigdha Chandratre; Siba Prasad Mishra; Suman Deb, Year: 2024, "Detection of the common cold from speech signals using transformer model and spectral features", Biomedical Signal Processing and Control, Vol: 93, pp. 106158.