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

pp. 211-222, April

Dual-Path Mobile Net-Based Framework for Esophageal Cancer Detection using Enhanced Endoscopic Imaging

P. Mukilan[4], P. Karputha Pandi[5]

[4]Professor, Department of Electronics and Communication Engineering, Dhanalakshmi Srinivasan College of Engineering, Coimbatore-641105. venmukilan@ieee.org

[5]Assistant professor, Department of Electrical and Electronics Engineering, Erode Senguthar Engineering College, Perundurai.

Corresponding Author E-mail: karputhapandi15102@gmail.com

ABSTRACT: Esophageal Cancer (EC) is essential for increasing patient survival in early identification; yet, endoscopists find it difficult to identify the cancer cells. This paper proposed a Dual-Path MobileNet for esophageal cancer screening is proposed to reduce the workload of physician and increase detection accuracy. Firstly, un-sharp mask filter is applied to Esophageal Endoscopy Images to brighten the edge and to increase its sharpness for better quality of image. Next, the processed image is given to segmentation process using Self Organizing Map (SOM) clustering algorithm. Here the SOM convert the high-resolution image into a lower-resolution image. After that image is extracted using Histogram of Oriented Gradient (HOG) it capture the edge and shape of the gradient direction for further analysis. Finally, a Dual-path MobileNet framework is processed to enhance the classification of esophageal cancer diagnosis. Using python software the proposed framework have an improved accuracy of 95% is accomplished when compared to other techniques.

Keywords: Dual-path MobileNet, Histogram of Oriented Gradient (HOG), Self Organizing Map (SOM) clustering, unsharp mask filter, Esophageal Cancer (EC)

1. Introduction

Esophageal Cancer (EC) is the sixth most common cause of cancer-related deaths globally and the ninth most common type of cancer overall [1]. Histologically, EC comprises adenocarcinoma (EAC) and Esophageal Squamous Cell Carcinoma (ESCC). EAC typically affects the lower third or junction of the esophagus, whereas ESCC typically affects the middle or upper third [2]. Consequently, up to 90% of ECs in lower-income nations as well as the "esophageal cancer belt," which includes northern China and Asian republics are caused by ESCC [3]. In addition, EAC makes up about 20% of all ECs in Western nations [4].

Whereas, Barrett's Esophagus (BE), a familiar preneoplastic injury for EAC, it is defined by intestinal metaplasia containing goblet cells replacing the esophageal squamous epithelium. A recent meta-analysis found that BE and EAC were diagnosed concurrently in as much as 57% of patients and around 12% of patients with an EAC diagnosis had previously been diagnosed with BE [5]. The 5-year survival rate for ESCC is 18% overall, but drops to fewer than 5% if distant metastases are detected upon analysis [6]. Similarly, when EAC is detected in its advanced stages, the 5-year survival rate is less than 20%. [7].

Using Image augmentation technique the esophageal cancer is diagnosed. This technique

greatly improves the model's capacity for generalization, particularly when it comes to identifying tissues inside malignant slides. Possibility of over fitting, higher computing costs when using augmentation [8-9]. Although, to detect EC an efficient High-Resolution U-Net (HRU-Net) architecture is used. It is exploited for esophageal cancer and esophageal carcinoma segmentation. Lack of period information, inadequate directionality, and shift sensitivity [10]. However, using Naive Bayes classifier the cancerous areas are detected. To adjust and improve its feature extraction capabilities naïve Bayes classifier is used. Low accuracy on complex issues, especially when working with high-dimensional search spaces [11]. Furthermore, to classify the esophageal cancer and Esophageal Squamous Cell Carcinoma (ESCC) a deep learning technique is used. It improved the sensitivity and accuracy for EC and ESCC. It is difficult to handle unbalanced data and inadequate performance when dealing with noisy data [12].

However, Convolutional Neural Networks (CNN) is used to improve the classification accuracy. Using CNN, the invasion depth of EC under is determined. It is more effective than endoscopists at identifying Esophageal Squamous Cell Carcinoma. To utility well, CNNs need a lot of labeled training data [13]. Moreover, to effectively identify esophageal cancer a hybrid Quantum Convolution Neural Network (QCNN) model is used. Here QCNN the normal objects are detected successfully. QCNN produces less accurate localizations because it requires a fixed-size input [14].

To improve the accuracy a hybrid ResNet-UNet model is used. The advantage of hybrid ResNet-UNet is easy to understand because they represent a solution to a problem step-by-step. With a larger number of parameters, hybrid ResNet-UNet typically needs more training data to function [15]. Nevertheless, mortality falls and the prediction much improves when EC be identified and managed early using Dual-path MobileNet.

2. Related work

Chen *et al* [16] (2021) have proposed an improved Faster RCNN (Faster R-CNN) to detect Esophageal cancer. Here using improved Faster Region based CNN the normal objects are detected successfully. Faster R-CNN produces less accurate localizations because it requires a fixed-size input

Takeuchi *et al* [17] (2021) have proposed a Fine-tuned VGG16 for the detection of esophageal cancer. Using fine-tuned VGG-16 model EC image improve the diagnostic accuracy and lower the overall misdiagnosis rate. Over fitting issue arises due to its complexity is a drawback for fine-tuned VGG-16.

Mubarak, D. M. N [18] (2022) have proposed a CNN model ResNet50 with Support Vector Machine (SVM) to classify the EC. The ResNet50 with SVM model is reliable enough too accurately and efficiently identify and classify benign and maligent. Slow convergence and local minima susceptibility are two drawbacks of ResNet with SVM.

Yousefi *et al* [19] (2021) have proposed a CNN to improve the classification accuracy. EC utilizing spatial and channel attention gates in every dense block, the Dilated Dense Attention Unet (DDAUnet) selectively focuses on determinant feature maps and areas. When identifying cancer areas, several issues are increased; poor identification speed, high computation costs, and inadequate precision.

Wu *et al* [20] (2021) have proposed a Deep Convolutional Neural Network (DCNN) for automatic esophageal lesion classification and segmentation. An Esophageal Lesion Network (ELNet) is used to segment each lesion by its class. Low accuracy in complex issues, especially when working with high-dimensional search spaces.

The contribution of this study is summarized as follows:

- Introduces un-sharp mask filter, to brighten the edges of an image for better image quality.
- For better segmentation the high resolution image is converted into low resolution using Self Organizing Map clustering.
- To capture the edge and shape of the gradient direction in EC image Histogram of Oriented Gradient is used.
- Develops a Dual-path MobileNET classifier to enhance EC accuracy and efficiency.
- Validates the proposed model using metrics like accuracy, precision, recall, F1-score, and AUC, demonstrating significant improvement over existing methods.

3. Proposed work

The proposed block diagram in figure 1 for esophageal cancer disease classification by deep learning technique is applied to Esophageal Endoscopy Images datasets. Initially the input esophageal cancer image overlaying an original image with a high-quality form using unsharp mask filtering technique. Then the processed esophageal cancer image is given as input to SOM clustering here it convert high-resolution image data into a lower-resolution.

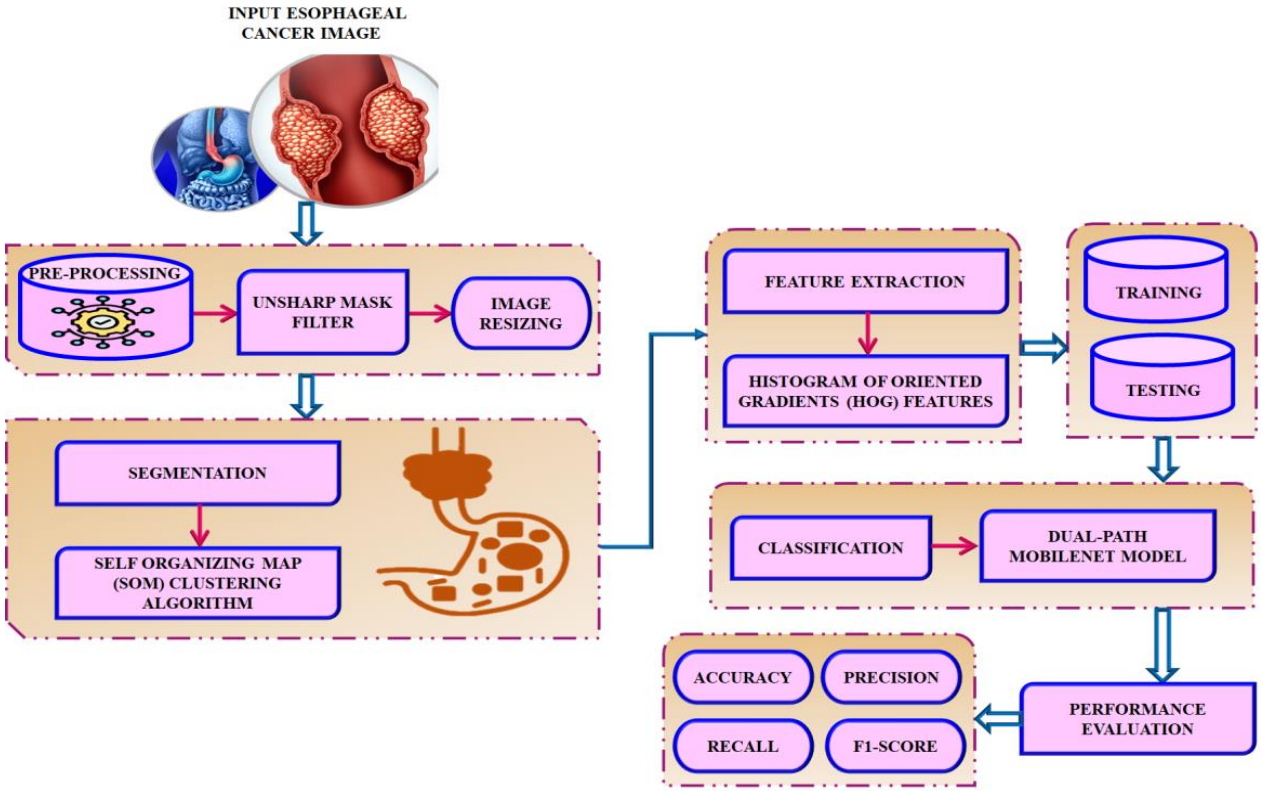

Figure 1: *Block diagram of proposed work*

Next, by using feature extraction technique called HOG it captures edge and shape information of gradient directions in certain areas of the EC image. In classification stage, a Dual-path MobileNet is utilized to increase the accuracy and efficiency. To improve model accuracy, performance of EC the classification technique is used.

3.1 Pre-processing using unsharp mask filter

Using unsharp mask filter the EC is pre-processed. This filter is used to improve the acutance in a number of imaging applications. It requires removing a blurry image from the original and then adding the back to the original. Following is the definition of an unsharp mask filter:

$$U(x,y) = P(x,y) + \lambda[P(x,y) - G(x,y)] \quad (1)$$

The coordinates are (x,y), the original image is P, the image later being filtered by a low pass Gaussian filter is G, the sharp image is U, and the sharpening parameter is λ. The degree of amplification in the high-frequency band is controlled and picture details are adaptively enhanced by using a dynamic sharpening parameter λ rather than a fixed one. This technique dynamically adjusts the sharpness value by using image variance to compute the dissimilarity among every pixel as well as its surrounding pixels. This is how the unsharp mask filter is calculated:

$$N\_d = P(i-1,\ j) - P(i,j) \quad (2)$$

$$S\_d = P(i+1,\ j) - P(i,j) \quad (3)$$

$$E\_d = P(i\ ,j+1) - P(i,j) \quad (4)$$

$$W\_d = P(i,\ j-1) - P(i,j) \quad (5)$$

$$T\_diff(x,y) = |N\_d + S\_d + E\_d + W\_d| \quad (6)$$

$$\lambda(x,y) = exp(-\sigma_2)0.18 + T\_diff \quad (7)$$

$$U(x,y) = P(x,y) + \lambda(x,y)[P(x,y) - Q(x,y)] \quad (8)$$

$N\_d, S\_d, E\_d, and\ W\_d$ stand for the variations seen in the east, west, south, and north, respectively, whereas (x,y) are the coordinates, σ^2denoted variance of P, Q denotes output of the better showed filter, and T_diff denotes total total dissimilarity among every pixel as well as its neighboring pixels. Following to pre-processing technique the segmentation is done by using SOM clustering.

3.2 Segmentation using Self Organizing Map (SOM) clustering:

Using SOM clustering the EC is segmented. SOM is a significant unsupervised machine-learning technique based on neural networks that reflects the distribution of input processed image using a competitive learning algorithm. There are two layers to it: input and output. A class that requires clustering is represented by a neuron in the output layer. Each input sample locates its winning neuron, or ideal partner neuron, in the output layer through "competitive learning." Through training, the active neuron's parameters are continuously modified.

Both the winning and neighbouring neurons' weights are nearer the original sample value. After training, the winning neurons imprinted on the SOM output layer will also be near each other if the subsamples in the original high-dimensional sample space. High-dimensional space's topological relationship is mirrored in SOM's low-dimensional space. SOM aims to identify the topological connections between data samples.

Self-organizing map algorithm: In SOM, all of the input layer's neurons are fully connected by the output layer's neurons, which are arranged in an array. Clustering is implemented using a competitive learning technique.

A procedure for competitive learning: The "winning" neuron, sometimes referred to as the Best Matching Unit (BMU), have the smallest Euclidean distance between each neuron weight W and the input pattern $X \subseteq R\ d$. The following is the formula:

$$D(X, C_k) = min\{||X - W||\} \quad (9)$$

WhereC_k denotes BMU, W denotes he neuron's weight vector k denotes output layer.

Set $k\ is$ the winning neuron , the winning neuron and its neighbours' weight update formula is as follows:

$$\omega_k(new) = \omega_k(old) + \rho\phi(k,n)(x - \omega_k) \quad (10)$$

Where $\phi(k,n)$ denotes neighbourhood purpose then drops by the distance of $||W\,k - W\,n||$ between $k\ and\ n$. Neuron n is the neighbour of winning neuron $k.\rho$ is the learning rate and decreases over time. The neighbourhood function is:

$$\phi(k,n) = \{\, e\,(-||W_k - W_k||^2)/2\sigma^2\,, k \neq n \quad (11)$$

$$\phi(k,n) = 1, k = n \quad (12)$$

Where σ^2denotes the parameter for width that progressively shrinks over time. Next, feature extraction by using Histogram of Oriented Gradient (HOG) feature.

3.3 Feature extraction using HOG Algorithm

The Histogram of Oriented Gradient (HOG) feature is a type of descriptor that detects object attributes using image processing and computer vision technologies. By computing a statistical histogram of the directional gradient in a particular region of the image, features are extracted. In the realm of image identification, Hog feature extractions have been used periodically.

3.4 Method for HOG

(1) Detection Window: HOG cut the image through the block and window for detection. Calculate the pixel values of a region in an image using mathematical units called cells. The terms window, block, and cell, as well as their relationships, are first introduced for detecting the EC.

Window: split the picture into several identical windows based on a specific slide and size.

Block: split every window into multiple identical blocks based on a specific slide and size;

Cell: Each window is separated into several identical cells of a specific size that are part of the feature extraction unit and stay in place.

(2) Normalized images: Gamma and color value normalization are examples of normalizing an image in figure 2. The impact of lighting conditions be successfully mitigated by normalizing the entire image. Additionally, normalization prevent a significant amount of the external exposure contribution to the grain intensity of the image. Standard formula for gamma compression:

$$I(x,y) = I(x,y)^{gamma} \quad (13)$$

I(x,y) takes values based on the effect.

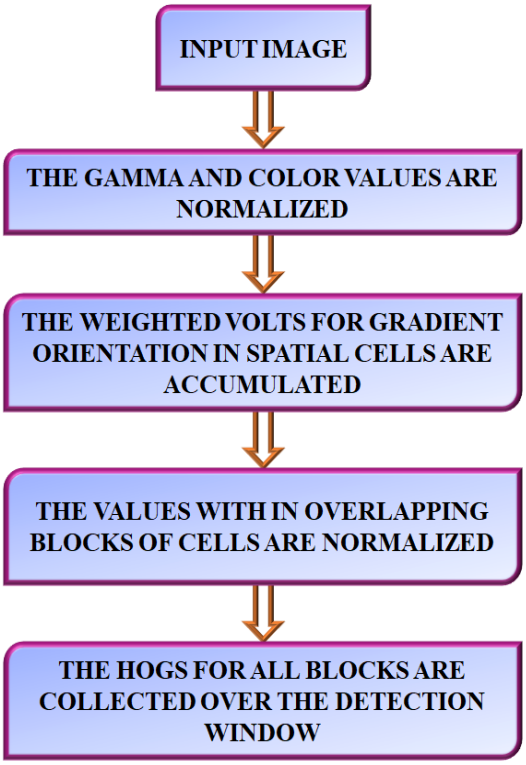

Figure 2*: Flow chart of HOG*

(3) Calculated gradient: In order to determine the gradient orientation, the gradient value in both the horizontal and vertical coordinate orientations must first be determined. The following is the formula:

$$f_x(x,y) = I(x+1,y) - I(x-1,y) \quad (14)$$

$$f_x(x,y) = I(x,y+1) - I(x,y-1) \quad (15)$$

At pixel (x, y), the gradient direction and amplitude values are:

$$G(x,y) = \sqrt{f_x^2(x,y) + f_y^2(x,y)} \quad (16)$$

(4) Building a gradient column diagram: Bins (number of divisions) define the orientation division. Bins are typically used, and the gradient direction is divided into discrete segments.

(5) Cell-normalized gradient histogram in the block: normalization is required because local lighting and foreground–background contrast significantly impact the growing range of gradient intensity.

(6) Create the HOG: feature vector by combining all of the components at last to find EC.

Next, the featured image is given as the input to the proposed Dual-path MobileNet for classification.

3.5 Classification using Dual-path MobileNet

In this proposed work Deep Learning method called Dual-path MobileNet is utilized for effective classification of esophageal cancer. The Dual-path MobileNet model is a type of deep learning model proposed to classify EC and no EC. The Convolutional Neural Network (CNN) is the basis of all state-of-the-art image classification systems. They are a type of Neural Network (NN) in which matrix multiplication is replaced with convolution for at least one layer. Convolution considers neighboring pixels, which significantly improves the network's efficacy in contrast to the NN model, which treats each element of the original image as a separate input. MobileNets are based on a reduced architecture that uses depth-wise separable convolution layers to produce lightweight network.

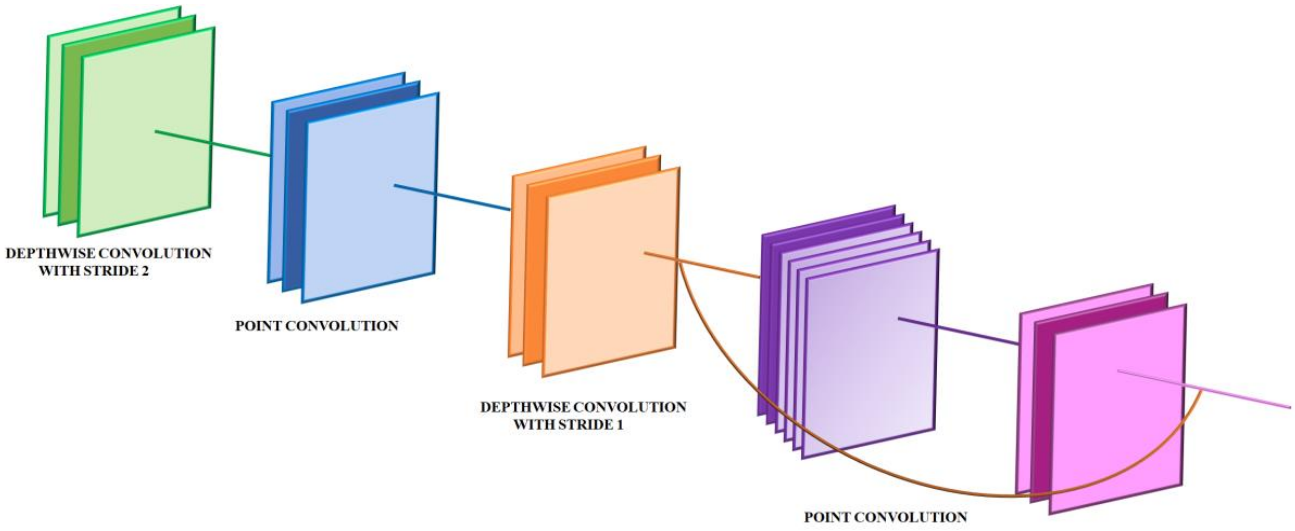

Figure 3: *Dual-path mobileNet architecture*

Eventually, the two-layer combination is dubbed depthwise separable convolution. Feature map convolutions, which form the basis of the MobileNet framework, have a major influence on converting a normal convolution into a Depth-Wise Convolution (DWC) and an 11-convolution called a pointwise convolution. For each network interface, MobileNets' depth-wise convolution uses a single filter. The pointwise convolution then uses an 11 convolution to combine the DWC results.

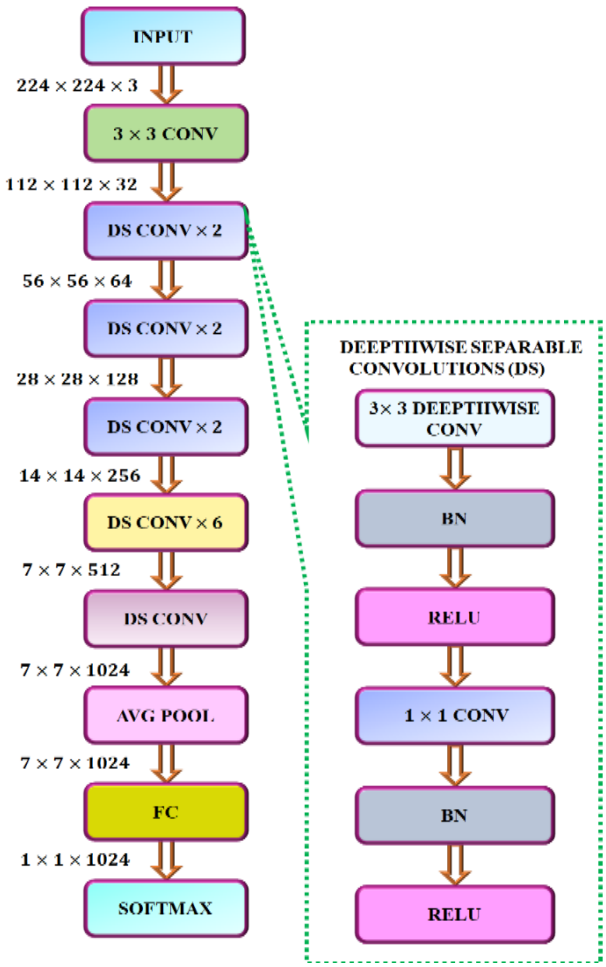

Figure 4: *MobileNet network layer*

Conventional convolutions filter the inputs and combine them into a new set of outputs in a single iteration. The computational and framework size is significantly reduced as a result of this simplification. The MobileNet network layer served as the model for design in figure 4, which flinches with the first convolutional layer then evolutions over thirteen depthwise convolution layers, each which is prospered by pointwise convolution layers.

3.6 Result and discussion

The aim of this work is to predict the esophagus and no-esophagus from EC images. Dataset is taken from kaggle.com for EC identification. Deep learning models are then applied to train Esophageal Endoscopy Images. Using python software the esophagus and no-esophagus is predicted. The train images have the count of 6000 and train images have the count of 8300.

Figure 5: *Train a test dataset counts area plot for EC*

Figure 5 shows the train a test dataset counts area plot for EC. Here total number of images is 8300 and the class labels are esophagus and no-esophagus. The train images have the count of 6000 and train images have the count of 8300.

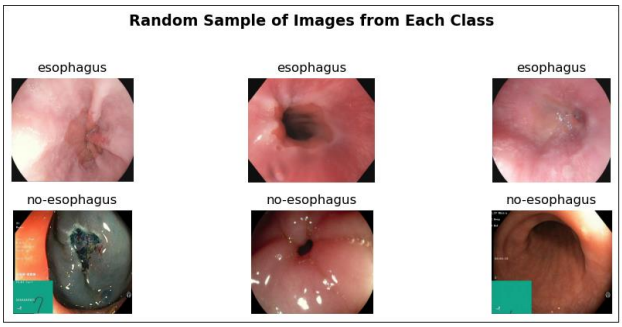

Figure 6: *Random samples of images in test and train*

In figure 6 random samples of images in test and train are shown. Here three esophagus and no-esophagus image for each class of random images. The random images in test and train are taken from the Esophageal Endoscopy images dataset. Here the two different classes of esophagus and no- esophagus image are present.

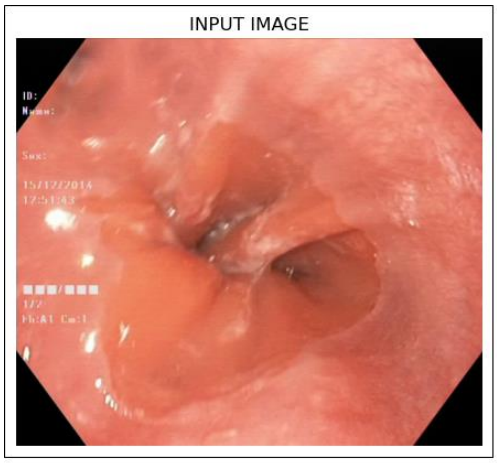

Figure 7: *Input esophageal image*

Figure 7 shows the input esophageal image. Here the input image is taken from the esophageal endoscopic image dataset. Using the pre-processing techniques the input image is processed.

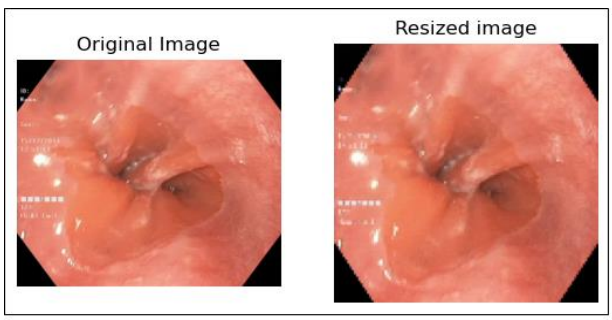

Figure 8: *Resized image*

Figure 8 shows the resized image EC image. Here the input EC image is resized for processing. Preserve visual quality by mapping the original image's pixel values to the resized image.

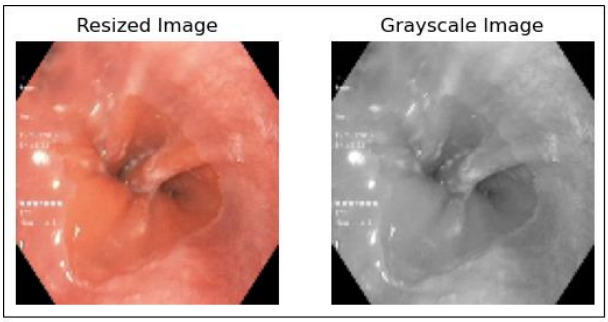

Figure 9: *Gray scale image*

The second steps of pre-processing for EC images are shown in Figure 9. Here the resized image is converted into grayscale image using grayscale conversion technique.

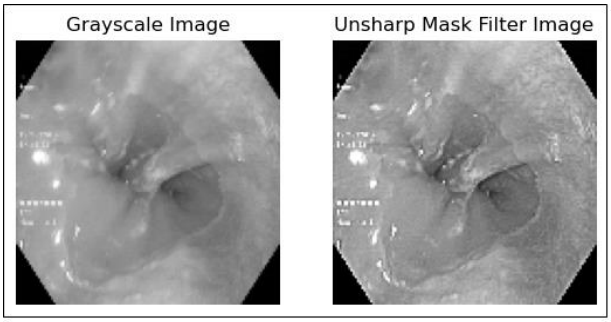

Figure 10: *Un-sharp mask filter*

Figure 10 shows the un-sharp mask filter EC image. Here the grayscale image is given as an input to un-sharp mask filter and then the edges are sharpened using the un-sharp mask filtering technique for good quality image.

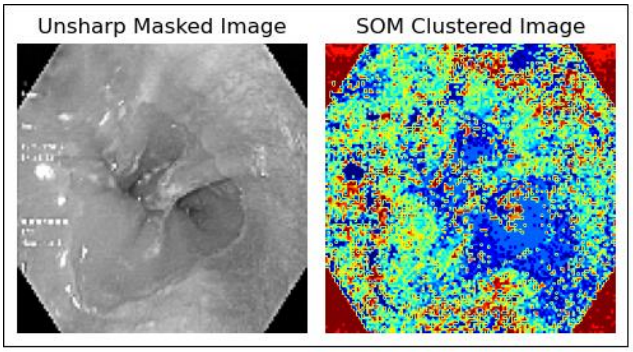

Figure 11: *SOM clustering image*

Figure 11 shows the self-organizing Map clustering image for EC. The unsharp mask filtered image is given as input to segmentation stage where SOM is used. Here the image is converted as high resolution image into low resolution clustered image which shows as color image.

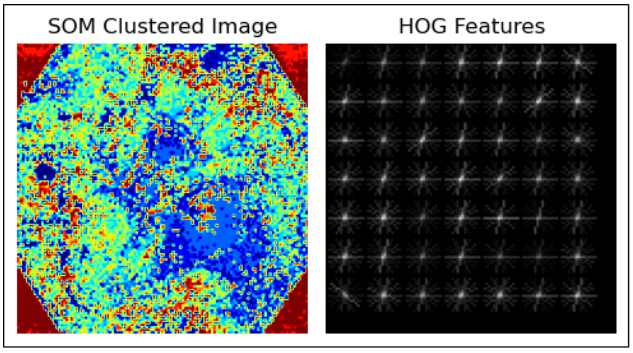

Figure 12: *HOG image*

Figure 12 shows the Histogram of Oriented Gradient for EC image. Using HOG each gradient directions in certain areas of the image, captured about its edges and shapes for EC image.

3.7 Performance metrics

For performance metrics EC image is identified by Dual-path MobileNet and classified into four types: True Positive (TP), which occurs when the model correctly detects cancer; True Negative (TN), which occurs when the model properly detects non-cancer; False Positive (FP), which occurs when the model detects cancer wrongly; as well as False Negative (FN), which occurs when the model wrongly detects non-cancer. The performance evaluation metric, precision calculated and defined as $TP/(TP + FP)$. To assess the diagnostic performance, four additional measurements: sensitivity$[TP/(TP + FN)]$, specificity $[TN/(TN + FP)]$, F value $[2 \times (recall \times precision)/ (recall + precision)]$, and accuracy $[(TP + TN)/(TP + FP + FN + TN)]$.

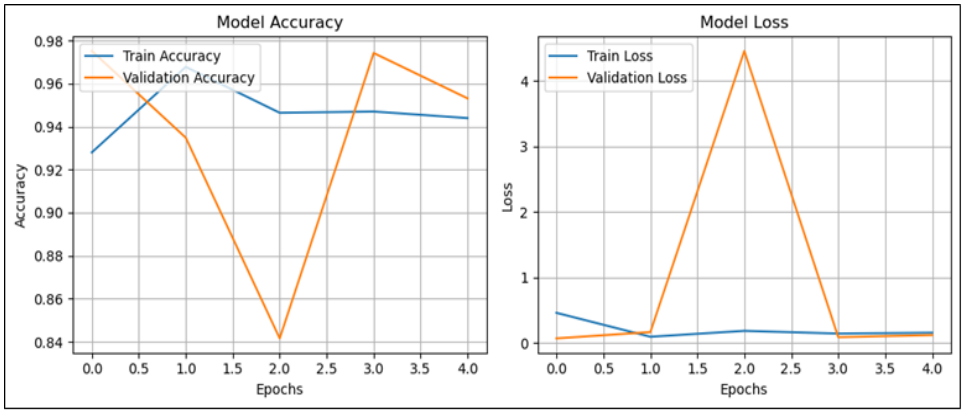

Figure 13: *Model loss and model accuracy*

The Model Loss graph figure 13 shows a reliable decline in training and validation loss, representing improved learning and reduced error. The X-axis denotes the epochs and y-axis is accuracy for model accuracy simultaneously for model loss X-axis for epochs and Y-axis for loss. The Model Accuracy graph illustrates increasing accuracy of 95%, with the model achieving high performance and minimal over fitting.

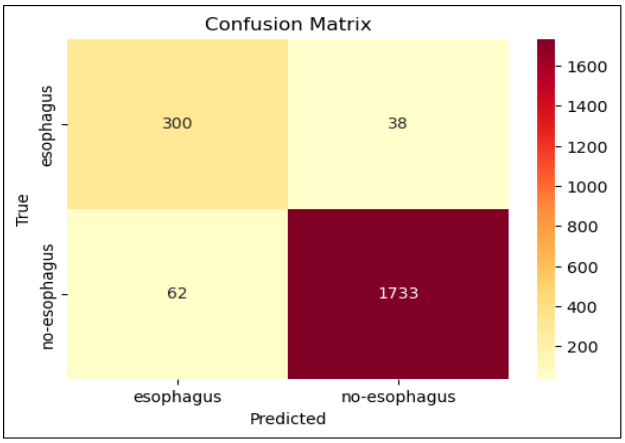

Figure 14: *Confusion matrix for proposed work*

Figure 14 shows the confusion matrix for proposed Dual-path MobileNet. For this investigation, the following definitions are given: False Negative (FN), False Positive (FP), True Positive (TP), and True Negative (TN) are calculated for esophagus and no-esophagus.

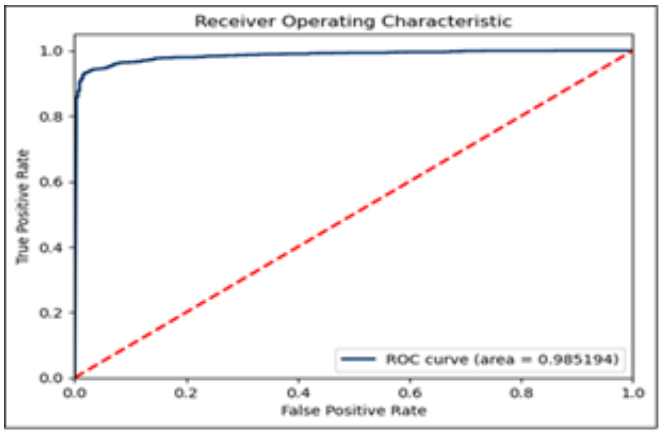

Figure 15: *ROC curve*

Figure 15 shows the ROC curve for esophagus and no-esophagus. Here the ROC value for proposed Dual-path MobileNet is 0.98.

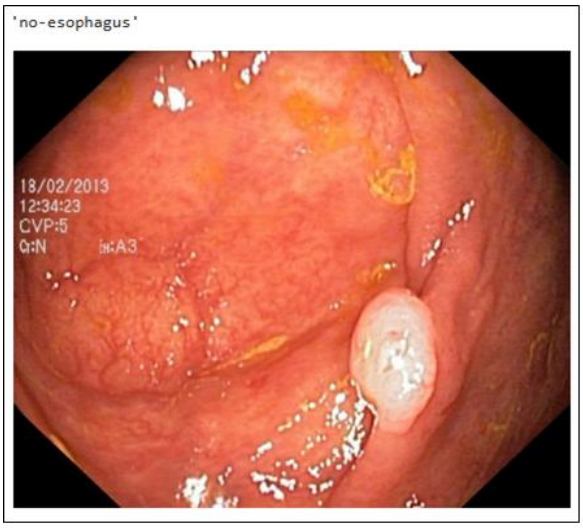

Figure 16: *Predicted output of no-esophagus*

Figure 16 shows the predicted output of no-esophagus. Here the no-esophagus image is predicted using the endoscope with better quality.

3.8 Comparison for the proposed method

The comparison of proposed method with existing method is presented in table 1. The CNN-VGG-16 have the classification accuracy of 84.2% [17] and CNN-ResNet 50 have the accuracy of 93.5% [18]. The proposed Dual-path MobileNet have the higher accuracy of 95% compared to the existing method in table 1. The proposed Dual-path MobileNet have better performance.

Table 1: *Comparison of proposed method*

Study	Method	Accuracy
Takeuchi *et al*. [17]	CNN-VGG-16 with SVM	84.2%
Mubarak *et al.*[18]	CNN-ResNet50	93.5%
Proposed approach	Dual-path MobileNet	95%

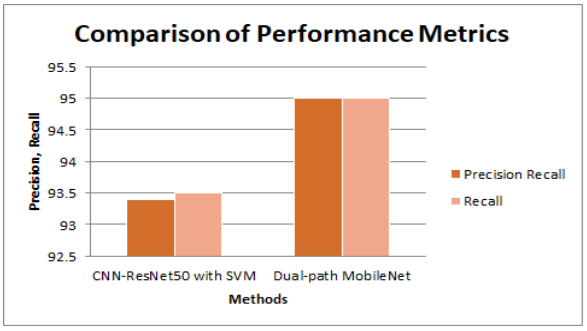

Figure 17: *Comparison for precision and recall*

The Precision and recall value for classifiers CNN-VGG-16 with SVM the proposed is presented in Figure 17. The CNN-VGG-16 with SVM have the value of 93.5% and 93.4% [17] and proposed Dual-path MobileNet method have the precision value of 95% and recall of 95% for EC.

4. Conclusion

In this study, Dual-path MobileNet is proposed for the classification of EC image. The preprocessing stage effectively enhances image quality by sharpen the edges of an image through un-sharp mask filtering. The segmentation process is refined using SOM clustering which enhanced the low resolution image. Furthermore, HOG-based feature extraction captures the structure and texture of EC images. The performance evaluation, conducted on the esophageal endoscopic image dataset, demonstrates the superior capabilities of the Dual-path MobileNet related to existing methods. The improved accuracy of 95% is achieved as a

greater accuracy in contrast to the existing techniques.

References

1. Vikrant Rai; Joe Abdo; Devendra K. Agrawal, Year: 2023, "Biomarkers for early detection, prognosis, and therapeutics of esophageal cancers", International Journal of Molecular Sciences, Vol; 24, No: 4, pp. 3316.

2. Cho-Lun Tsai; Arvind Mukundan; Chen-Shuan Chung; Yi-Hsun Chen; Yao-Kuang Wang; Tsung-Hsien Chen; Yu-Sheng Tseng; Chien-Wei Huang; I-Chen Wu; Hsiang-Chen Wang, Year: 2021, "Hyperspectral imaging combined with artificial intelligence in the early detection of esophageal cancer", Cancers , Vol:13, No:18, pp. 4593.

3. Mahdi Sheikh; Gholamreza Roshandel; Valerie McCormack; Reza Malekzadeh, Year: 2023, "Current status and future prospects for esophageal cancer", Cancers, Vol: 15, No: 3, pp. 765.

4. Pierfrancesco Visaggi; Brigida Barberio; Matteo Ghisa; Mentore Ribolsi; Vincenzo Savarino; Matteo Fassan; Michele Valmasoni; Santino Marchi; Nicola de Bortoli; Edoardo Savarino, Year: 2021, "Modern diagnosis of early esophageal cancer: from blood biomarkers to advanced endoscopy and artificial intelligence", Cancers, Vol: 13, No: 13, pp. 3162.

5. Tsung-Jung Tsai; Arvind Mukundan; Yu-Sheng Chi; Yu-Ming Tsao; Yao-Kuang Wang; Tsung-Hsien Chen; I-Chen Wu; Chien-Wei Huang; Hsiang-Chen Wang, Year: 2022, "Intelligent identification of early esophageal cancer by band-selective hyperspectral imaging", Cancers, Vol: 14, No: 17, pp. 4292.

6. Xiaokun Li; Lingmin Chen; Siyuan Luan; Jianfeng Zhou; Xin Xiao; Yushang Yang; Chengyi Mao et al, Year: 2022, "The development and progress of nanomedicine for esophageal cancer diagnosis and treatment", In Seminars in cancer biology, Vol: 86, pp. 873–885.

7. Hang Yang; Bing Hu, Year: 2021, "Recent advances in early esophageal cancer: diagnosis and treatment based on endoscopy", Postgraduate Medicine, Vol: 133, No: 6, pp. 665-673.

8. Jianpeng An; Wenqi Li; Yunhao Bai; Huazhen Chen; Gang Zhao; Qing Cai; Zhongke Gao, Year: 2024, "MTECC: A Multi-Task Learning Framework for Esophageal Cancer Analysis", IEEE Transactions on Artificial Intelligence.

9. Yu-Jen Fang; Arvind Mukundan; Yu-Ming Tsao; Chien-Wei Huang; Hsiang-Chen Wang, Year: 2022, "Identification of early esophageal cancer by semantic segmentation", Journal of Personalized Medicine, Vol:12; No: 8, pp. 1204.

10. Muwei Jian; Chen Tao; Ronghua Wu; Haoran Zhang; Xiaoguang Li; Rui Wang; Yanlei Wang; Lizhi Peng; Jian Zhu, Year: 2024, "HRU-Net: A high-resolution convolutional neural network for esophageal cancer radiotherapy target segmentation", Computer Methods and Programs in Biomedicine, Vol:250, pp. 108177.

11. Nazlı Pınar Karahan Şen; Ayşegül Aksu; Gamze Çapa Kaya, Year: 2021, "A different overview of staging PET/CT images in patients

with esophageal cancer: the role of textural analysis with machine learning methods", Annals of Nuclear Medicine, Vol: 35, No: 9, pp. 1030-1037.

12. Hiromu Fukuda; Ryu Ishihara; Yusuke Kato; Takashi Matsunaga; Tsutomu Nishida; Takuya Yamada; Hideharu Ogiyama; Mai Horie; Kazuo Kinoshita; Tomohiro Tada, Year: 2020, "Comparison of performances of artificial intelligence versus expert endoscopists for real-time assisted diagnosis of esophageal squamous cell carcinoma (with video)", Gastrointestinal endoscopy, Vol:92, No: 4, pp. 848-855.

13. Yoshitaka Tokai; Toshiyuki Yoshio; Kazuharu Aoyama; Yoshimasa Horie; Shoichi Yoshimizu; Yusuke Horiuchi; Akiyoshi Ishiyama et al, Year: 2020, "Application of artificial intelligence using convolutional neural networks in determining the invasion depth of esophageal squamous cell carcinoma", Esophagus, Vol: 17, pp. 250– 256.

14. R. I. Minu; Martin Margala; S. Siva Shankar; Prasun Chakrabarti; G. Nagarajan, Year: 2023, "Early-stage esophageal cancer detection using hybrid quantum CNN", Soft Computing, Vol: 1–6.

15. Albert J.de Groof; Maarten R. Struyvenberg; Kiki N. Fockens; Joost van der Putten; Fons van der Sommen; Tim G. Boers; Sveta Zinger et al, Year: 2020, "Deep learning algorithm detection of Barrett's neoplasia with high accuracy during live endoscopic procedures: a pilot study (with video)", Gastrointestinal endoscopy, Vol: 91, No: 6, pp. 1242– 1250.

16. Kuan-bing Chen; Ying Xuan; Ai-jun Lin; Shao-hua Guo, Year: 2021, "Esophageal cancer detection based on classification of gastrointestinal CT images using improved Faster RCNN", Computer Methods and Programs in Biomedicine, Vol: 207, pp. 106172.

17. Masashi Takeuchi; Takumi Seto; Masahiro Hashimoto; Nao Ichihara; Yosuke Morimoto; Hirofumi Kawakubo; Tatsuya Suzuki et al, Year: 2021, "Performance of a deep learning-based identification system for esophageal cancer from CT images", Esophagus, Vol: 18, pp. 612– 620.

18. D. M. N. Mubarak, Year: 2022, "Classification of early stages of esophageal cancer using transfer learning", Irbm, Vol: 43, No: 4, pp. 251– 258.

19. Sahar Yousefi; Hessam Sokooti; Mohamed S. Elmahdy; Irene M. Lips; Mohammad T. Manzuri Shalmani; Roel T. Zinkstok; Frank JWM Dankers; Marius Staring, Year: 2021, "Esophageal tumor segmentation in CT images using a dilated dense attention Unet (DDAUnet)", IEEE Access, Vol: 9, pp. 99235–99248.

20. Zhan Wu; Rongjun Ge; Minli Wen; Gaoshuang Liu; Yang Chen; Pinzheng Zhang; Xiaopu He; Jie Hua; Limin Luo; Shuo Li, Year: 2021, "ELNet: Automatic classification and segmentation for esophageal lesions using convolutional neural network", Medical Image Analysis, Vol: 67, pp. 101838.