Deep Learning based Cervical Cancer Classification and Segmentation from Pap Smears Images using an EfficientNet

Krishna Prasad Battula[1]
Research scholar, School of CSE
VIT-AP University, Amaravati
Andhra Pradesh, India

Dr B. Sai Chandana[2]
School of CSE, VIT-AP University
Amaravati
Andhra Pradesh, India

***Abstract*—One of the most prevalent cancers in the world, cervical cancer claims the lives of many people every year. Since early cancer diagnosis makes it easier for patients to use clinical applications, cancer research is crucial. The Pap smear is a useful tool for early cervical cancer detection, although the human error is always a risk. Additionally, the procedure is laborious and time-consuming. By automatically classifying cervical cancer from Pap smear images, the study's goal was to reduce the risk of misdiagnosis. For picture enhancement in this study, contrast local adaptive histogram equalization (CLAHE) was employed. Then, from this cervical image, features including wavelet, morphological features, and Grey Level Co-occurrence Matrix (GLCM) are extracted. An effective network trains and tests these derived features to distinguish between normal and abnormal cervical images by using EfficientNet. On the aberrant cervical picture, the SegNet method is used to identify and segment the cancer zone. Specificity, accuracy, positive predictive value, Sensitivity, and negative predictive value are all utilized to analyze the suggested cervical cancer detection system performances. When used on the Herlev benchmark Pap smear dataset, results demonstrate that the approach performs better than many of the existing algorithms.**

***Keywords*—Cervical cancer; pap smear; time-consuming; contrast local adaptive histogram equalization (CLAHE); Grey Level Co-occurrence Matrix (GLCM); morphological features; wavelet; SegNet**

I. INTRODUCTION

A thin layer of tissues made up of cells covers the human cervix. Cervical cancer is the term used to describe a condition when a cell is transformed into a malignant cell that can divide and expand quickly to form a tumor [1]. Cervical cancer, which affects women worldwide and ranks as the second leading cause of cancer-related mortality, is fatal. If this cancer is found early enough, it may be treated [2]. Typically, a biopsy and screening procedure is used to make the diagnosis. Techniques for image processing can be used to determine where cancer has spread. The fourth-most frequent cancer-related cause of death in women is cervical cancer [3, 4].

Intelligent systems and medical image processing both contribute to the analysis of cancerous cells. They grow more time and money efficient as new approaches are developed [5-7]. They are currently gaining popularity in place of traditional techniques including Pap smears, colposcopies, and Cervicography [8]. These methods are objective to the human experience, but it's important to note that they don't completely replace the professional doctor's subjective assessment, even though they can greatly aid it [9].

Analyses of the nucleus and cytoplasm are typically necessary for cell classification investigations to take cell type into account. Consequently, it is essential to develop algorithms that would aid in nuclei and cytoplasm segmentation [10-12]. The majority of feature extraction uses the same standards that specialists use to evaluate a cell. However, there is limited knowledge of cervical cytology [13]. Although this has not yet been researched, the cell might possess traits found in higher features. As a result, deep learning techniques have recently made representational learning more well-known. The automatic extraction of characteristics from input photos is a remarkable benefit of deep learning [14].

Analyses of the nucleus and cytoplasm are typically necessary for cell classification investigations to determine the type of cell. As a result, algorithms that can divide the cytoplasm and nuclei into separate parts must be developed [15-17]. The majority of feature extraction considers cells using the same standards as experts. Cervical cytology is, however, not well understood. The cell might possess traits found in higher forms, however, this has not yet been researched [18].

As a result, deep learning techniques have recently made representational learning more well-known. The automatic extraction of characteristics from input photos is a remarkable benefit of deep learning. As a result, automatic screening has made extensive use of deep learning [19]. In particular, DeepPap achieved a predictive performance of 98.6% on the Herlev benchmark Pap smear dataset and comparable effectiveness on the HEMLBC private Pap smear dataset using patch extraction from the nucleus ground truth mask and a transfer learning strategy to initialize weights with a pre-trained model.

The main cause of cervical cancer-related deaths in female patients is that the disease cannot be identified at an earlier stage, and patients do not experience any symptoms until cancer has progressed to its terminal stage [20]. Only if it is

discovered at an earlier stage can the death ratio for female patients be decreased. To prevent patient deaths, this research suggests a mechanism for detecting cervical cancer at an earlier stage.

The major key contributions of the research are as follows,

- The preprocessing stages are used to increase the classification efficiency, even more, here images can be resized, data augmentation strategies are used and CLAHE is employed to improve the image quality.
- Following that, features such as moment invariant features, GLCM features, and wavelet features are extracted from the preprocessed image.
- The EfficientNet classifier is trained using these features to categorize the cervical images as Normal or Abnormal.
- Finally, we propose a SegNet for segmenting the defected areas, it uses multi-stage architecture and attention blocks in each stage.
- The Herlev dataset has undergone various ablation experiments. Our proposed network outperforms the state efficiency concerning all other approaches, according to the experimental data.

The paper is organized as follows. In Section II, a few similar prior efforts are summarized. Section III describes the proposed system. Section IV presents experimental findings and a discussion. Section V contains the work's conclusion.

II. Literature Review

In the literature, there is a lot of research comparing the effectiveness of various methods utilized to treat cervical cancer. ML and DL techniques were used in group studies. It is clear from studies on cervical cancer that deep learning techniques including CNN, stacked autoencoder, VGG19, and LASSO were applied.

Convolutional neural networks (CNNs) were introduced by Ghoneim et al. [21] for the identification and categorization of cervical cancer cells. To extract deep-learned features, a CNNs model is fed the cell pictures. After that, a classifier powered by an extreme learning machine (ELM) classifies the input photographs. Through transfer learning and fine-tuning, CNN's model is applied. Their accuracy is superior to others as compared to the current system. Comparatively, the level of complexity is higher.

To classify cervical cancer from Pap smears, William et al. [22] developed an improved fuzzy c-means method. Through the employment of a Trainable Weka Segmentation classifier, cells were segmented, and trash was eliminated sequentially. Wrapper filters were used to select features. The method surpasses several of the current algorithms in terms of false negative rate, false positive rate, and classification error, according to the results. The primary drawback of the recommended automated Pap smear analysis systems is their inability to handle the Pap smear architectures complexities.

Deep learning methods, such as softmax classification with stacked autoencoder, have reportedly been utilized to categorize data sets, according to Adem et al. [23]. The raw data collection is transformed into a lower dimension dataset by applying a stacked autoencoder. In order to minimize the data dimension and create a classifier with high accuracy, the stacked autoencoder model was used. Comparatively speaking to the other machine learning method, it has higher classification success rates for data related to cervical cancer. It is necessary to increase accuracy by removing relevant information.

The VGG19 (TL) model and the CYENET were presented by Chandran et al. [24] to automatically classify cervical tumors from colposcopy pictures. By improving the VGG19 model, which is extensively used for medical image processing, the transfer learning method is applied to forecast accuracy. By utilizing an optimal architecture and an ensemble technique CYENET, the CNN created from scratch is intended to automate the screening of cervical pictures. The outcomes of the experiments demonstrate that the suggested CYENET had high performances. As a result of the dimension reduction, training process is very long.

The classification of cervical biopsy tissue images based on LASSO and ensemble learning-support vector machine (EL-SVM) was first presented by Huang et al. [25]. The average optimization time was decreased by 35.87 seconds while maintaining the classification accuracy when the LASSO technique was used for feature selection. Serial fusion was then carried out. 468 biopsy tissue pictures were identified and classified using the EL-SVM classifier. The ROC curve and error curve were utilized to assess the classifier's generalizability. The results of the experiment indicate that a superior categorization result was obtained. A two-step feature selection process that takes time and makes it challenging to distinguish between individual cells.

A smaller visual Geometry Group-like Network is used to classify the segmented entire cervical cell data by Allehaibi et al. [26] using a Mask R-CNN and VGG-like Net. The ResNet10 network serves as the foundation of the Mask R-CNN, fully utilizing geographical data and past knowledge. Mask R-CNN performs better in precision, recall, and ZSI than the prior segmentation approach during the segmentation phase when applied to the entire cell. The performance of the seven-class problem categorization produces excellent results. The suggested method uses a little volume of data and requires more research to fully understand cervical cells.

Allehaibi et al. [27] presented an Inception v3-based cervical cell categorization system with features that were intentionally extracted. The accuracy of cervical cell recognition has been significantly improved by the use of Inception v3 and artificial characteristics. Additionally, this research inherits the strong learning capability from transfer learning to produce an accurate and efficient classification of cervical cell images while addressing the under-fitting issue with a limited amount of medical data. The suggested algorithm offers great accuracy, good universality, and minimal complexity. The Suggested method had issues with certain cells having overlapping cytoplasm sections. Table I represents the merits and demerits of the existing papers.

TABLE I. MERITS AND DEMERITS OF RELATED WORKS

Reference	Year	Method	Dataset	Merits	Demerits
[21]	2020	CNN& Extreme Machine learning	Herlev database	More scalable and practical	Investigating cervical cells requires more research.
[22]	2019	fuzzy c-means algorithm	DTU/Herlev benchmark Pap smear dataset	Higher precision and smaller data dimensions	Due to the dimension reduction, training time is quite long.
[23]	2019	stacked autoencoder	Cervical cancer dataset	Acquire more complementing features	an increase in image fusion complexity
[24]	2021	VGG19 and CYENET	Intel ODT dataset	Better accuracy Efficient classification	More complexity Need more investigation
[25]	2020	LASSO- (EL-SVM)	Cervical cancer	Better sensitivity and specificity	Extraction of pertinent data is required to increase accuracy.
[26]	2019	Mask R-CNN	Herlev Pap Smear dataset	Improvements in accuracy and feature selection	The selection of features in two phases takes time.
[27]	2019	Inception v3	Herlev dataset	Excellent robustness and best performance	need development to change the parameter

It is possible to draw a conclusion using a few of the restrictions and potential improvements that may be learned from the existing research. The effectiveness of existing approaches was insufficient. The algorithm's complexity and potential are primarily responsible for this. Our proposed approach is used to address the gaps in the current body of research. The segmentation and classification method for cervical cancer that has been suggested is intended to enhance classifier performance.

III. METHODOLOGY

The four essential components of the proposed methodology were preprocessing, feature extraction, classification, and segmentation. The preprocessing processes are used to increase the classification efficiency even more. The raw photos were prepared using the pre-processing method. Data compression and image enhancement were part of the pre-processing. By using the CLAHE approach the image can be enhanced. Then, characteristics including GLCM features, wavelet features, and moment invariant features are extracted from the preprocessed image. By comparing the cervical picture with the taught features, the EfficientNet classifier is trained to determine if the image is normal or abnormal. Finally, SegNet is employed to segment the defected area. After that, the dataset is separated into three parts. The training model is then fed this dataset. A selection of test datasets is utilized to label the training model, which is then used to categorize cervical cancer. Fig. 1 demonstrates the architecture diagram of the proposed framework.

A. Preprocessing

Images are strengthened in resolution during the initial stages of pre-processing utilizing various filters. To prevent further distortions, several data augmentation methods, including shearing, scaling, and rotation, are also used. Images are also resized at various scales. The CLAHE algorithm can be used to improve the image.

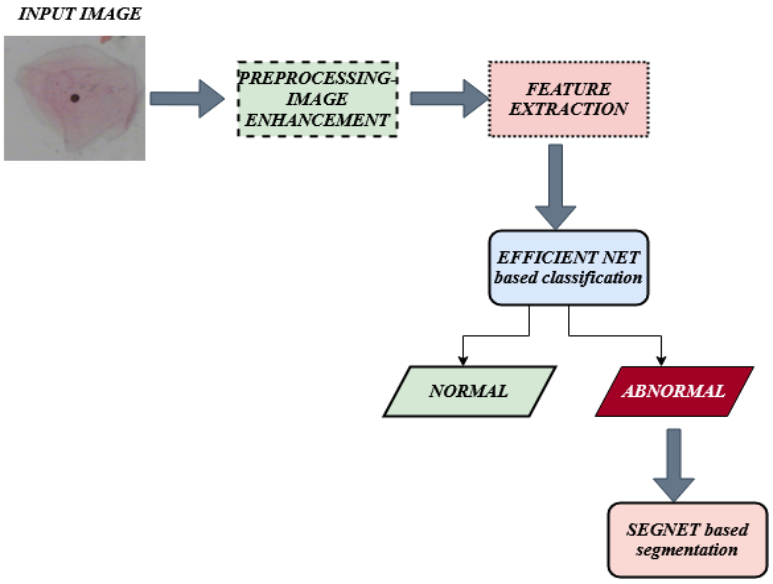

Fig. 1. Architecture Diagram of the Proposed Methodology.

1) Contrast Local Adaptive Histogram Equalization (CLAHE): With the help of histogram clipping, the histogram-based image enhancement technique known as CLAHE can only amplify images to a certain degree. It is a technique that is efficient for assigning projected intensity levels in medical data. In order to adjust the brightness of a pixel's display to reflect where that pixel ranks in terms of intensity in its histogram, the method analyses an intensities histogram in a contextual region that is focused on each individual pixel. The histogram that is generated shows the image contrast created by the technique at every luminance. It is a modified version of the standard histogram.

It lessens contrast enhancement, which is typically achieved through the histogram equalization method, which also makes more noise. As a consequence, by reducing contrast augmentation in Histogram Equalization, the expected result was attained in situations where noise became particularly obvious by enhancing contrast, such as medical photos. To lessen contrast, one can restrict the slope of the linked function. The use of CLAHE in our work has helped to increase the accuracy rate overall.

B. Feature Extraction

The nucleus' and cytoplasm's essential characteristics, including texture, shape, and color, were retrieved at this stage. In an image, features stand for the traits of the pixel pattern. In this study, morphological, wavelet, and GLCM characteristics are retrieved from cervical images to distinguish between normal and pathological images. GLCM was employed to extract eight texture features. Normal and abnormal cells appear very differently in the cervical Pap smear image in terms of color and form distribution.

2) Wavelet features: Due to their speed and superior transformation capabilities compared to other transforms like Contourlet and Curvelet, Wavelets are beneficial in multi-resolution analysis of cervical images. This study decomposes the magnitude response Gabor image using the Discrete Wavelet Transform (DWT), which is applied to every row and column.

The LL, HL, LH, and HH sub-bands are produced by the first level decomposition. L and H stand for low and high frequencies, respectively. Additionally, the second-level decomposition of DWT is applied to the LL sub-band to create four additional subbands. The feature pattern for the cervical image classifications uses each of these subbands.

3) GLCM features: The feature extraction method known as GLCM is employed to extract the energy features of the cervical picture. Any single channel image can have one built. The GLCM is a square matrix with the same number of rows and columns as there were in the original image's grayscale. The GLCM matrix is created by counting the number of times a grayscale intensity pixel will be found next to a pixel with the value of the fused cervical picture at various orientations, such as 0°, 45°, 90°, and 135°. The GLCM matrix is built in this work at a 45° angle. The contrast, energy, entropy, and correlation characteristics of the GLCM features are employed to distinguish the cancerous image from the healthy cervical image.

Contrast: It's described as,

$$Contrast = \sum(|\, i - j \,|^2 \times p(i, j)) \tag{1}$$

This texture feature calculates the difference in grayscale between adjacent pixels.

Correlation: It's described as,

$$Correlation(R) = \sum(i - \mu i)(j - \mu j)\frac{p(i, j)}{[\sigma_i, \sigma_j]} \tag{2}$$

The correlation between brightness in adjacent pixels is measured by R. The GLCM's row average i and column average (j) are respectively. The GLCM's row μ_i and column μ_j corresponding standard deviations are denoted by σ_i and σ_j.

Energy: It is a second angular moment, calculated by adding the squares of all the GLCM's components.

$$Energy(E) = \sum p(i, j)^2 \tag{3}$$

Energy, a unit of measurement for homogeneity, runs from zero to unity for a picture.

Entropy: It is described as,

$$Entropy = -\sum p(i, j)[\log_2 p(i, j)] \tag{4}$$

The degree to which the GLCM's elements are near the diagonal is gauged by this metric. The value is between 0 and 1.

4) Morphological features: Cell size and form make up the morphological characteristics. For this study, eight connected chain codes were used to determine the morphological characteristics of cells. The eight pixels around the eight-linked chain code are the connected pixels.

a) Area: the number of pixels that each cell has taken up;

b) Circumference: The cell's circumference is equal to one week.

c) Nuclear to cytoplasm ratio: ratio of nuclear to cellular size:

$$\frac{N}{C} = \frac{Nucl_{area}}{Nucl_{area} + Cyto_{area}} \tag{5}$$

C. Classification based on EfficientNet

The EfficientNet group has eight variants, spanning from B0 to B7, and the quantity of estimated parameters does not increase much as the amount of models increases, even though accuracy increases. In contrast to previous CNN models, EfficientNet employs the Swish activation function rather than the Rectifier Linear Unit (ReLU) activation function. Deep learning frameworks seek to uncover more efficient methods using fewer methods. EfficientNet, unlike other state-of-the-

art approaches, achieves more efficient results by scaling depth, width, and resolution evenly while reducing the strategy's size. When resources are limited, the initial step in the compound scaling strategy is to discover the relationship among the various scaling dimensions of the network. This method determines an appropriate scaling factor for the depth, breadth, and resolution dimensions. After that, these coefficients are used to scale the baseline network to the target network. Table II represents the EfficientNet architecture.

CNN belongs to the EfficientNet group. In terms of layer width, layer depth, input resolution, and a combination of these criteria, EfficientNet methods scale well. EfficientNet is a recent deep learning approach that aims to improve model efficiency while also improving accuracy. From B0 to B7, there are various variants. The inversion bottleneck MB Conv is the basic building piece for EfficientNet. It was first presented in MobileNetV2, however, it is used significantly further than MobileNetV2 because of the larger FLOPS budget. Blocks in MBConv are made up of layers that expanded and then compress the channels, hence direct connections are employed among bottlenecks that connected far fewer channels than expansion layers. When compared to typical layers, this design has in-depth separate convolutions that minimize computation by almost a k2 factor. The 2D convolution window's height and breadth are determined by the kernel size, or k, which is the opposite.

EfficientNet presents a novel compound scaling model that scales network width, depth, and image size uniformly using a compound coefficient ϕ.

$$depth: d = \alpha^{\phi} \tag{6}$$

$$width: w = \beta^{\phi} \tag{7}$$

$$resolution: r = \gamma^{\phi} \tag{8}$$

$$s.t.\ \alpha.\beta^2.\gamma^2 \approx 2 \qquad \alpha \geq 1, \beta \geq 1, \gamma \geq 1 \tag{9}$$

FLOPS are proportional to d, w2, and r2 in a standard convolution process. Because convolution operations account for the majority of the cost of computation in convolution networks, growing the network as indicated in Eq. (3.5) boosts the network's FLOPS by about (α, β2, γ2) φ (α, β2, γ2) φ in total.

The batch normalization (BN) restricts the final layer's outcome to a range, requiring a mean of zero and one SD. This adjustment shortens the training period and improves the model's stability. The compound scaling approach scales this system in two phases, starting with the baseline EfficientNet-B0.

Step 1: Considering double as many materials are allocated, a grid search with φ = 1 is used to find the best values for α, β, γ.

Step 2: The generated α, β, γ values are set as constants, and the baseline network is scaled up using Equation. (3.5) with varied values φ to generate EfficientNet-B1 through B7.

D. Segmentation

Using morphological techniques, the cancerous regions of an aberrant cervical picture are segmented. An encoder network, a decoder network, and a final layer for pixel-wise categorization make up SegNet [28]. This configuration consists of four blocks, the final block of which does not execute pooling. To create feature vectors that correspond to each input, features from the input image are extracted using the encoder layers. A decoder network that consists of four blocks of upsampling, convolution, and batch normalizing layers is then applied after that.

TABLE II. EFFICIENTNET FRAMEWORK

Level	Operator	Resolution	Channels	Layers
EfficientNetB0 architecture, the network baseline				
1	Conv1×1/Pool/FC	7×7	1,280	1
2	MBConv6, k3×3	7×7	320	1
3	MBConv6, k6×6	14×14	192	4
4	MBConv6, k3×2	14×14	112	3
5	MBConv6, k5×4	28×28	80	3
6	MBConv6, k5×5	56×56	40	2
7	MBConv6, k3×3	112×112	24	2
8	MBConv1, k3×3	112×112	16	1
9	Conv 3×3	224×224	32	1
Additional layers				
10	FC/Softmax	1	NC	1
11	FC/BN/Swish	1	128	1
12	FC/BN/Swish/Dropout	1	512	1
13	B.N./Dropout	7×7	1280	1

Segmentation mask is created by the decoder network using feature vectors, and output at high resolution is produced by upsampling layers using low features. The final layer is a classification layer that uses 2D convolution with a 1x1 filter size to do pixel-by-pixel classification. Throughout the training, a stride of one and a filter size of 3x3 are employed. Except for the final layer, where the sigmoid activation function is employed, ReLU is employed as the activation function. To reduce the size of the feature map, max-pooling layers with a pool size of 2x2 are employed [29]. After each batch, the weights are optimized by the Adam optimizer with a learning rate of 1e-4.

The model is trained throughout 10 epochs with a batch size of 32. For pixel-wise segmentation, a Softmax classifier is given the final decoder output feature maps. For quick and precise image segmentation, the decoder recovered spatial dimensions. Due to its memory and processing speed, the SegNet architecture is generally superior to other systems like U-Net and FCN.

IV. Result and Discussions

To illustrate the conclusion, using a benchmark dataset to compare the proposed method to existing techniques in terms of NPV, sensitivity, specificity, PPV, and accuracy. The materials and metrics that were employed to achieve the intended results will be described in this paper. The proposed experiment's performance was evaluated in PYTHON.

A. Dataset Description

Herlev dataset, which is made available to the public for disease detection, was obtained from Denmark Hospital to detect cervical disease. There are 917 total Pap smear images in the complete data set. The dataset is split into a training set and a testing set, with the training set including 643 photos and the testing set including 274 images. The classifications of Normal and Abnormal have been taken into consideration.

With a total cell count of 917, we have taken into account the various cervical cell types, including epithelial and dysplastic, which are divided into normal and abnormal classes. The sample smear cervical cells from the dataset shown in Fig. 2 were chosen at random.

B. Evaluation Metrics in Cervical Cancer Diagnosis

The effectiveness of the categorization model is explained by several evaluation criteria used by various authors. Confusion matrices are used to assess the model's efficacy for the majority of medical image classification. Sensitivity, Specificity, Accuracy, and F1 score are some of the distinctive metrics employed for the analysis. The prediction output includes four results: true positive, true negative, false positive, and false negative. The various performance criteria used to evaluate the mode are displayed in Table III.

Accuracy: The number of correctly identified images determines a technique's classification accuracy, which is evaluated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

Sensitivity (Recall): It measures the percentage of positive samples that are accurately categorized. Sensitivity has a value between 0 and 1.

$$Sensitivity = \frac{TP}{TP + FN} \quad (11)$$

Specificity: It is a measurement of the percentage of incorrectly identified negative samples.

$$Specify = \frac{TN}{TN + FP} \quad (12)$$

Positive Predictive Value (PPV): It counts the number of pixels that are positive and accurately identify cancerous regions.

$$PPV = \frac{TP}{TP + FP} \quad (13)$$

Negative Predictive Value (NPV): It counts the number of negative pixels that are associated with incorrectly identified cancer region pixels.

$$NPV = \frac{TN}{TN + FN} \quad (14)$$

Confusion Matrix: The Confusion Matrix summarizes the categorization problem's prediction results. The confusion matrix reveals not just the classifier's errors, but also the sorts of errors. Fig. 3 represents the output of cervical cancer.

C. Evaluation of Classification Performances

A comparison of classification techniques and current classifiers is provided in this section. Four classification methods AlexNet, LeNet, VGG and Inception V3 are used in this study. Table III displays the effectiveness of many strategies, including the one that is suggested.

The existing approaches like Lenet, AlexNet, VGG and Inception V3 are compared with the proposed approach. When comparing with the sensitivity metrics it achieves 89.73% in LeNet, 90.16% in AlexNet, 91.37% in VGG, 99.44% in Inception V3 and our proposed approach yield a greater solution which is 99.67%. The next comparison can be made in Specificity LeNet achieves 84.33%, AlexNet yields 87.34%, VGG gains 86.88%, 96.73% in Inception V3 and the proposed approach yields 98.39%.

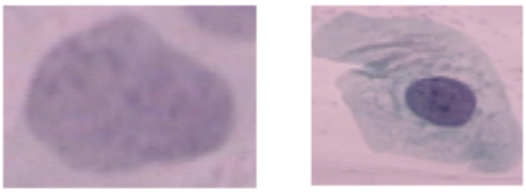

Fig. 2. Sample Images from the Dataset.

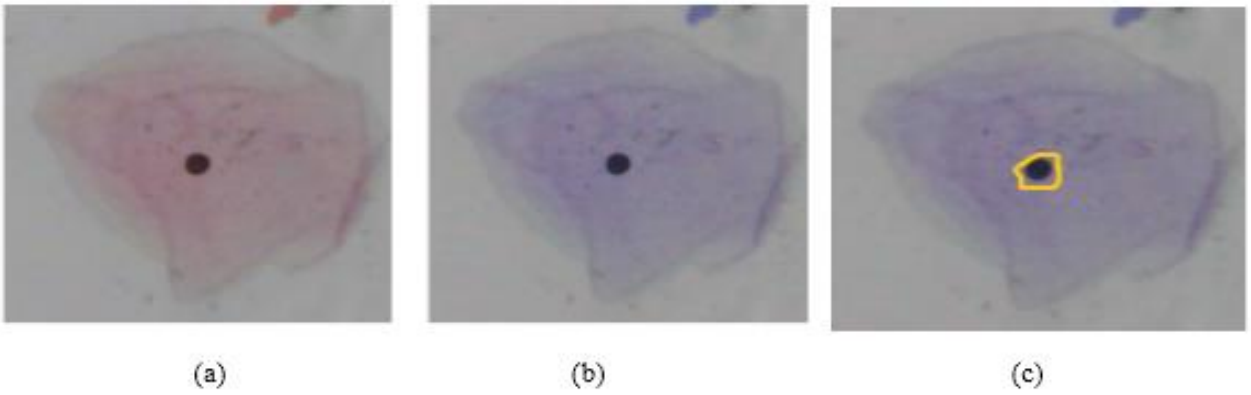

Fig. 3. Output of Cervical Cancer Affected Image (a) Input Image (b) Contrast-enhanced (c) Segmented Defected Area.

TABLE III. PERFORMANCES COMPARISON OF PROPOSED WITH EXISTING APPROACHES

Approaches	Sensitivity	Specificity	Accuracy
LeNet	89.73	84.33	86.76
AlexNet	90.16	87.34	89.57
VGG	91.37	86.88	91.47
Inception V3	99.44	96.73	98.23
EfficientNet	99.67	98.39	99.05

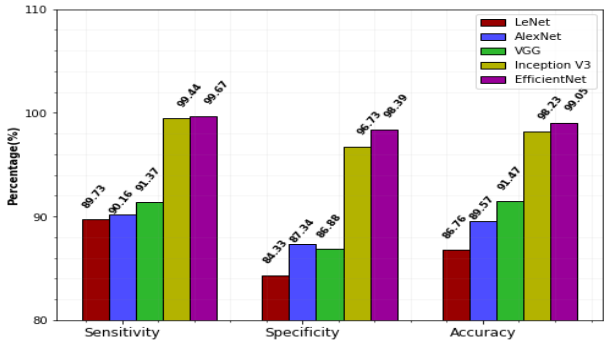

Fig. 4. Performances Comparison of Proposed with Existing Approaches.

Finally, a comparison can be made with accuracy metrics our proposed approach yields 99.05% which is the best solution then the worst solution appear in LeNet which is 86.76%. Performances comparison of proposed with existing approaches is represented in Fig. 4.

The metrics like Accuracy, Specificity, Sensitivity, PPV and NPV are used to compare the performances. To compare our proposed approach performances, the existing approach includes CYENET, DenseNet-121, DenseNet-169 and SVM utilized (Table IV).

Comparison can be made with the Accuracy metrics CYENET achieves 92.30% of accuracy, DenseNet-121 yields 72.42% of accuracy, DenseNet-169 gains 69.79% of accuracy, SVM achieves 63.27% of accuracy finally our proposed approach gains 99.67% of accuracy which is the greater one. Fig. 5 represents the Accuracy comparison of the proposed with existing approaches.

TABLE IV. COMPARISON OF PERFORMANCES OF THE CLASSIFIERS

Approaches	Accuracy (%)	Specificity (%)	Sensitivity (%)	PPV (%)	NPV (%)
CYENET	92.30	96.20	92.40	92	95
DenseNet-121	72.42	76.83	59.86	48.39	84.52
DenseNet-169	69.79	71.48	65	44.84	85.31
SVM	63.27	71.85	78.46	70	76.87
Proposed	99.67	98.39	99.67	94.67	96.13

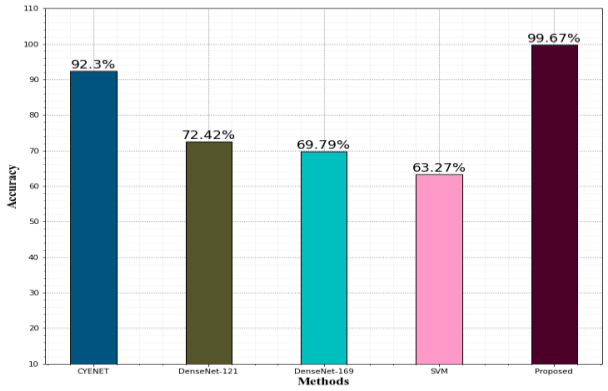

Fig. 5. The Accuracy Comparison of Proposed with Existing Approaches.

When comparing with specificity our proposed approach gains 98.39%, CYENET achieves 96.20%, DenseNet-121 yields 76.83%, DenseNet-169 gains 71.48%, and SVM achieves 71.85%. The figure represents the specificity comparison of the proposed with the existing. The Specificity performance comparison of the proposed with existing approaches is shown in Fig. 6.

Comparison can be made with the Sensitivity metrics CYENET achieves 92.4% of Sensitivity, DenseNet-121 yields 59.86% of Sensitivity, DenseNet-169 gains 65% of Sensitivity, SVM achieves 78.46% of Sensitivity finally our proposed approach gains 99.67% of Sensitivity which is the greater one. Fig. 7 represents the sensitivity comparison of the proposed with existing.

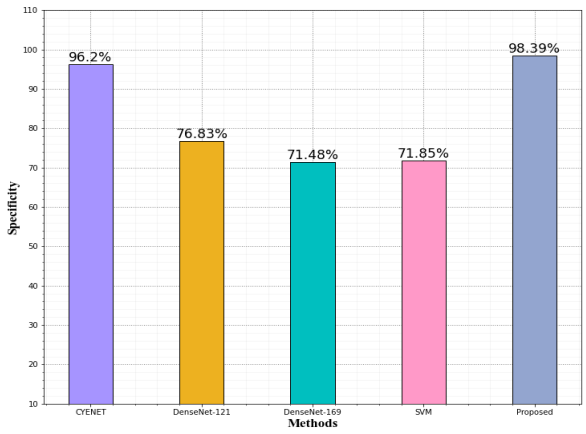

Fig. 6. The Specificity Performances Comparison of Proposed with Existing Approaches.

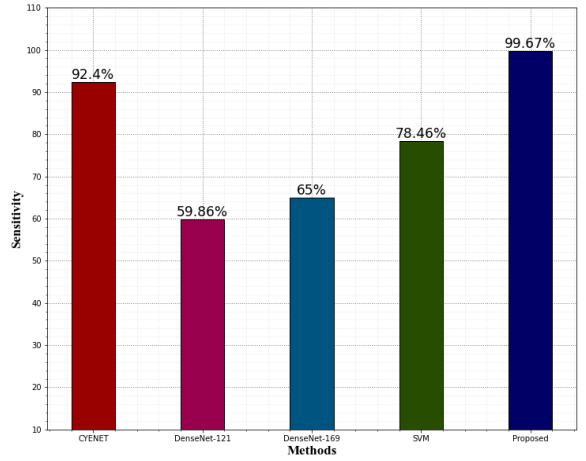

Fig. 7. The Sensitivity Performances Comparison of the Proposed with Existing Approaches.

Comparison can be made with the Sensitivity metrics CYENET achieves 92% of PPV, DenseNet-121 yields 48.39% of PPV, DenseNet-169 gains 44.84% of PPV, SVM achieves 70% of PPV finally our proposed approach gains 94.67% of PPV which is the greater one. Fig. 8 represents the PPV comparison of proposed with existing.

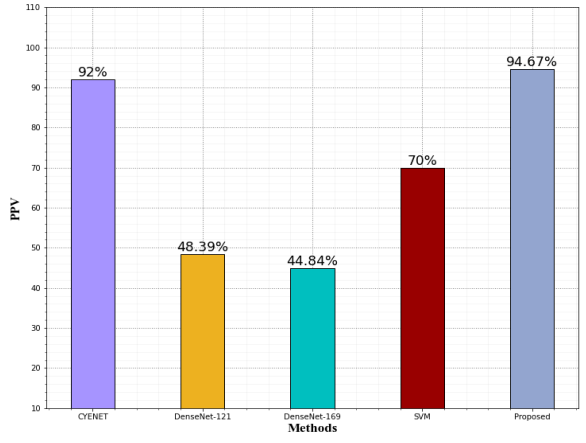

Fig. 8. The PPV Metrics Comparison of Proposed with Existing Approaches.

Comparison can be made with the Sensitivity metrics CYENET achieves 95% of NPV, DenseNet-121 yields 84.52% of NPV, DenseNet-169 gains 85.31% of NPV, SVM achieves 76.87% of NPV finally our proposed approach gains 94.67% of NPV which is the greater one. The Fig. 9 represents the NPV comparison of the proposed with the existing.

The confusion matrix for the end-to-end trained proposed method is shown in Fig. 10. 3% of Normal, 2% of abnormal samples were misclassified, while 97 % of benign samples, 98% of cancerous samples were classified correctly. As a result, the proposed approach acquires the best result.

D. Evaluation of Training and Testing

Train Accuracy and Validation Accuracy curves converge in the end, and after 50 epochs we received an accuracy of 99.56%, which is quite good. The validation Loss curve jumps up and down a bit. It means it would be nice to have more validation data. Fig. 11 represents the proposed training accuracy versus testing accuracy.

After about 25 epochs Validation Loss exceeds Train Loss, which means we have a bit of overfitting here. But the curve doesn't go up over epochs, and the difference between Validation and Train Loss is not that big, so this could be accepted. Fig. 12 represents the proposed training loss versus testing loss

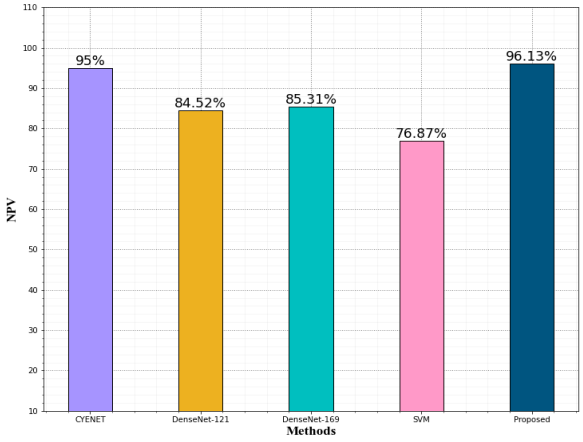

Fig. 9. The NPV Metrics Comparison of Proposed with Existing Approaches.

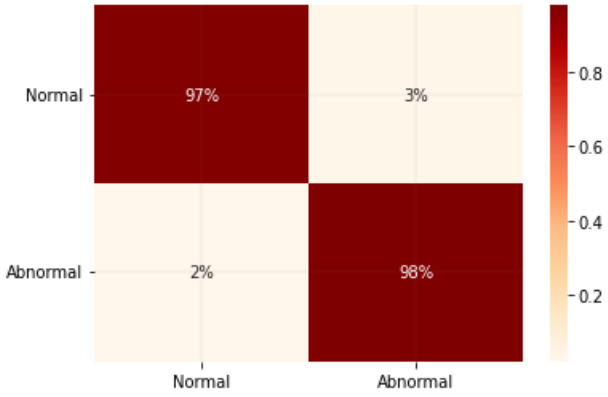

Fig. 10. Confusion Matrix.

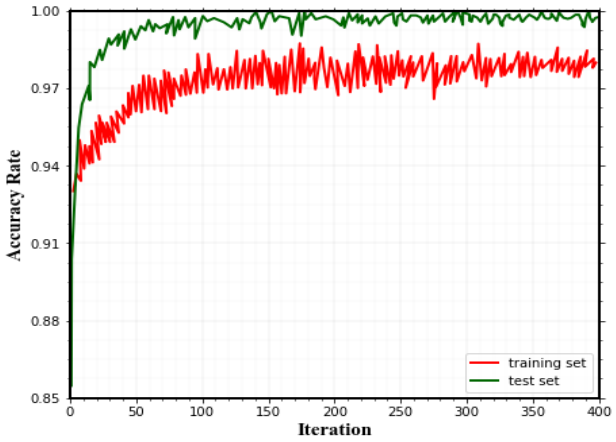

Fig. 11. Proposed Training Accuracy Versus Testing Accuracy.

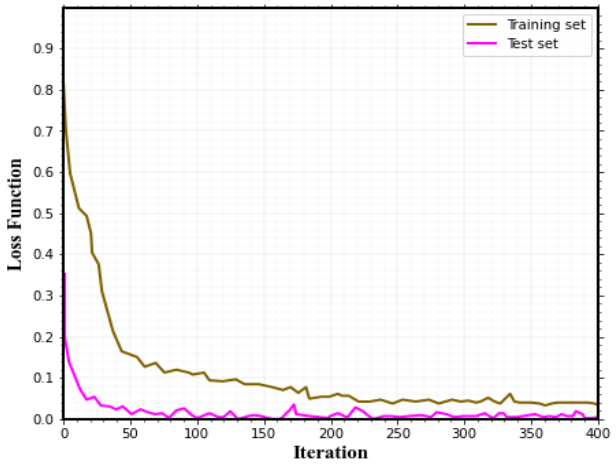

Fig. 12. Proposed Training Loss Versus Testing Loss.

E. Discussions

The success rate of treatment for cervical cancer is greatly impacted by early identification. The pathological study of microscopic Pap smear slide pictures is the primary cervical cancer screening tool. Automatic image analysis techniques are likely to defeat subjective justifications and lighten the workload. Due to the considerable unpredictability of cervical cell pictures, including overlapping cells, dust, contaminants, and uneven irradiation, effective nucleus image segmentation remains a difficult challenge. Furthermore, restrictions in feature design and selection make it difficult to classify cervical smear images.

The quantitative analysis of microscopic Pap smear slide pictures is difficult as a result. The absence of uninvolved photos from the public database of cervical images is crucial for the early detection of cervical cancer. Additionally, there are few cervical smear photos that have been labelled. The paper proposed an automatic cervical smear image categorization system based on EfficientNet to address the existing issues mentioned above. The usefulness of the proposed approach, which takes time to apply, is shown by experimental findings. Future research might focus on improving the method's efficiency and lowering computational complexity.

V. Conclusion

In conclusion, a set of clinically relevant and biologically understandable features are used to offer an automated detection and classification approach for the identification of cancer from cervical pictures. The proposed methodology is based on a network for segmenting and classifying cancer. The CLAHE-based approach is utilized to improve the cervical pictures. EfficientNet classifier is employed to divide cervical pictures into normal and abnormal images. According to the simulation results, the suggested cervical cancer segmentation method can identify both normal and abnormal areas in images of the cervical region. The cervical cancer detection system's performance metrics are 97.42 percent sensitivity, 99.36 percent specificity, 98.29 percent accuracy, 97.28 percent PPV, and 92.17 percent NPV. The theoretical deep learning model will be tested on other datasets in the future. Combining a few sophisticated image processing techniques with the method can also improve it.

Acknowledgment

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

References

[1] Wentzensen, N., Lahrmann, B., Clarke, M. A., Kinney, W., Tokugawa, D., Poitras, N., ... & Grabe, N. (2021). Accuracy and efficiency of deep-learning–based automation of dual stain cytology in cervical Cancer screening. JNCI: Journal of the National Cancer Institute, 113(1), 72-79.

[2] Mohammadi, R., Shokatian, I., Salehi, M., Arabi, H., Shiri, I., & Zaidi, H. (2021). Deep learning-based auto-segmentation of organs at risk in high-dose rate brachytherapy of cervical cancer. Radiotherapy and Oncology, 159, 231-240.

[3] Alyafeai, Z., & Ghouti, L. (2020). A fully-automated deep learning pipeline for cervical cancer classification. Expert Systems with Applications, 141, 112951.

[4] Matsuo, K., Purushotham, S., Jiang, B., Mandelbaum, R. S., Takiuchi, T., Liu, Y., & Roman, L. D. (2019). Survival outcome prediction in cervical cancer: Cox models vs deep-learning model. American journal of obstetrics and gynecology, 220(4), 381-e1.

[5] Chandran, V., Sumithra, M. G., Karthick, A., George, T., Deivakani, M., Elakkiya, B., ... & Manoharan, S. (2021). Diagnosis of cervical cancer based on ensemble deep learning network using colposcopy images. BioMed Research International, 2021.

[6] Jiang, X., Li, J., Kan, Y., Yu, T., Chang, S., Sha, X., ... & Wang, S. (2020). MRI-based radiomics approach with deep learning for prediction of vessel invasion in early-stage cervical cancer. IEEE/ACM transactions on computational biology and bioinformatics, 18(3), 995-1002.

[7] Kudva, V., Prasad, K., & Guruvare, S. (2017). Detection of specular reflection and segmentation of cervix region in uterine cervix images for cervical cancer screening. Irbm, 38(5), 281-291.

[8] Wu, M., Yan, C., Liu, H., Liu, Q., & Yin, Y. (2018). Automatic classification of cervical cancer from cytological images by using convolutional neural network. Bioscience reports, 38(6).

[9] Ch, P. N., Gurram, L., Chopra, S., & Mahantshetty, U. (2018). The management of locally advanced cervical cancer. Current opinion in oncology, 30(5), 323-329.

[10] Lee, J., Chang, C. L., Lin, J. B., Wu, M. H., Sun, F. J., Jan, Y. T., ... & Chen, Y. J. (2018). Skeletal Muscle Loss Is an Imaging Biomarker of Outcome after Definitive Chemoradiotherapy for Locally Advanced Cervical CancerSkeletal Muscle Loss in Cervical Cancer. Clinical Cancer Research, 24(20), 5028-5036.

[11] Kudva, V., Prasad, K., & Guruvare, S. (2020). Transfer learning for classification of uterine cervix images for cervical cancer screening. In

Advances in Communication, Signal Processing, VLSI, and Embedded Systems (pp. 299-312). Springer, Singapore.

[12] Melamed, A., Margul, D. J., Chen, L., Keating, N. L., Del Carmen, M. G., Yang, J., ... & Rauh-Hain, J. A. (2018). Survival after minimally invasive radical hysterectomy for early-stage cervical cancer. New England Journal of Medicine, 379(20), 1905-1914.

[13] Asadi, F., Salehnasab, C., & Ajori, L. (2020). Supervised algorithms of machine learning for the prediction of cervical cancer. Journal of biomedical physics & engineering, 10(4), 513.

[14] Singh, S. K., & Goyal, A. (2020). Performance analysis of machine learning algorithms for cervical cancer detection. International Journal of Healthcare Information Systems and Informatics (IJHISI), 15(2), 1-21.

[15] Kudva, V., Prasad, K., & Guruvare, S. (2020). Hybrid transfer learning for classification of uterine cervix images for cervical cancer screening. Journal of digital imaging, 33(3), 619-631.

[16] Kan, Y., Dong, D., Zhang, Y., Jiang, W., Zhao, N., Han, L., ... & Luo, Y. (2019). Radiomic signature as a predictive factor for lymph node metastasis in early-stage cervical cancer. Journal of Magnetic Resonance Imaging, 49(1), 304-310.

[17] Matsuo, K., Machida, H., Shoupe, D., Melamed, A., Muderspach, L. I., Roman, L. D., & Wright, J. D. (2017). Ovarian conservation and overall survival in young women with early-stage cervical cancer. Obstetrics and gynecology, 129(1), 139.

[18] Jia, A. D., Li, B. Z., & Zhang, C. C. (2020). Detection of cervical cancer cells based on strong feature CNN-SVM network. Neurocomputing, 411, 112-127.

[19] Nirmal Jith, O. U., Harinarayanan, K. K., Gautam, S., Bhavsar, A., & Sao, A. K. (2018). DeepCerv: Deep neural network for segmentation free robust cervical cell classification. In Computational Pathology and Ophthalmic Medical Image Analysis (pp. 86-94). Springer, Cham.

[20] Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems, 102, 643-649.

[21] Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems, 102, 643-649.

[22] William, W., Ware, A., Basaza-Ejiri, A. H., & Obungoloch, J. (2019). Cervical cancer classification from Pap-smears using an enhanced fuzzy C-means algorithm. Informatics in Medicine Unlocked, 14, 23-33.

[23] Adem, K., Kiliçarslan, S., & Cömert, O. (2019). Classification and diagnosis of cervical cancer with stacked autoencoder and softmax classification. Expert Systems with Applications, 115, 557-564.

[24] Chandran, V., Sumithra, M. G., Karthick, A., George, T., Deivakani, M., Elakkiya, B., ... & Manoharan, S. (2021). Diagnosis of cervical cancer based on ensemble deep learning network using colposcopy images. BioMed Research International, 2021.

[25] Huang, P., Zhang, S., Li, M., Wang, J., Ma, C., Wang, B., & Lv, X. (2020). Classification of cervical biopsy images based on LASSO and EL-SVM. IEEE Access, 8, 24219-24228.

[26] Jia, A. D., Li, B. Z., & Zhang, C. C. (2020). Detection of cervical cancer cells based on strong feature CNN-SVM network. *Neurocomputing*, *411*, 112-127.

[27] Allehaibi, K. H. S., Nugroho, L. E., Lazuardi, L., Prabuwono, A. S., & Mantoro, T. (2019). Segmentation and classification of cervical cells using deep learning. IEEE Access, 7, 116925-116941.

[28] Almotairi, S., Kareem, G., Aouf, M., Almutairi, B., & Salem, M. A. M. (2020). Liver tumor segmentation in CT scans using modified SegNet. Sensors, 20(5), 1516.

[29] Weng, L., Xu, Y., Xia, M., Zhang, Y., Liu, J., & Xu, Y. (2020). Water areas segmentation from remote sensing images using a separable residual segnet network. ISPRS International Journal of Geo-Information, 9(4), 256.