Cervix Type and Cervical Cancer Classification System Using Deep Learning Techniques

Lidiya Wubshet Habtemariam¹, Elbetel Taye Zewde¹,², Gizeaddis Lamesgin Simegn³,²

¹Biomedical Imaging Unit, School of Biomedical Engineering, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia; ²AI and Biomedical Imaging Research unit, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

Correspondence: Gizeaddis Lamesgin Simegn, Tel +251913925481, Email gizeaddis.lamesgin@ju.edu.et

Purpose: Cervical cancer is the 4th most common cancer among women, worldwide. Incidence and mortality rates are consistently increasing, especially in developing countries, due to the shortage of screening facilities, limited skilled professionals, and lack of awareness. Cervical cancer is screened using visual inspection after application of acetic acid (VIA), papanicolaou (Pap) test, human papillomavirus (HPV) test and histopathology test. Inter- and intra-observer variability may occur during the manual diagnosis procedure, resulting in misdiagnosis. The purpose of this study was to develop an integrated and robust system for automatic cervix type and cervical cancer classification using deep learning techniques.

Methods: 4005 colposcopy images and 915 histopathology images were collected from different local health facilities and online public datasets. Different pre-trained models were trained and compared for cervix type classification. Prior to classification, the region of interest (ROI) was extracted from cervix images by training and validating a lightweight MobileNetv2-YOLOv3 model to detect the transformation region. The extracted cervix images were then fed to the EfficientNetb0 model for cervix type classification. For cervical cancer classification, an EfficientNetB0 pre-trained model was trained and validated using histogram matched histopathological images.

Results: Mean average precision (mAP) of 99.88% for the region of interest (ROI) extraction, and test accuracies of 96.84% and 94.5% were achieved for the cervix type and cervical cancer classification, respectively.

Conclusion: The experimental results demonstrate that the proposed system can be used as a decision support tool in the diagnosis of cervical cancer, especially in low resource settings, where the expertise and the means are limited.

Keywords: cervical cancer, cervix type, classification, deep learning, detection, histopathology image, transformation zone

Introduction

The cervix is part of the female reproductive organ structurally found at the lower fibromuscular portion of the uterus. It contains two kinds of cells: rectangular columnar cells and flat scale-like squamous cells. Abnormal cells are developed in the transformation zone, which is an area where columnar cells are constantly changing into squamous cells.¹ The location of the transformation zone varies among women. Based on transformation zone location, cervix is classified as cervix type 1, type 2, and type 3.²³ The cells in the transformation zone develop gradually to abnormal cells and could change into cervical cancer through time.

Cervical cancer is one of the leading causes of cancer-related deaths in women, worldwide, with 80% of the cases occurring in developing countries.⁴⁵ According to the human papillomavirus (HPV) information center report, about 6294 new cervical cancer cases are diagnosed annually in Ethiopia.⁶ In remote areas, which have poor medical conditions with insufficient healthcare accessibility and unqualified medical staff, cervical cancer incidence and mortality are estimated to be higher.

Cervical cancer screening techniques include visual inspection after application of acetic acid (VIA),⁷ papanicolaou test (Pap),³⁸ and human papillomavirus (HPV) test.⁹ The type of cervical cancer is confirmed by visual examination of histopathology images under a microscope. These manual techniques are usually tedious, time-taking, and prone to error.
due to inter- and intra-observer variability. Computer-aided diagnosis (CAD) techniques have the potential to automate the manual screening and diagnosis procedures and help medical professionals in a wide range of decision-making, including disease detection, localization of cervix, cancer grading, and treatment planning.

Cervix type classification has been proposed in the literature using the transfer learning approach.\textsuperscript{10–12} Inception v3, ResNet 50, and ResNet 34 pre-trained models were employed in these literatures to classify cervix into type 1, type 2 and type 3. However, the accuracies reported were generally low. Similarly, a two-stage solution which performs classification of cervix type following region of interest (ROI) extraction were proposed in different literatures.\textsuperscript{13–19} Cruz et al\textsuperscript{13} proposed a color assembly method for ROI extraction and used it to train an artificial neural network (ANN), and an accuracy of 64\% were reported. A similar approach was proposed by Lei et al\textsuperscript{14} using Inception v3 and an accuracy of 74\% were reported. The highest accuracy (96.77\%) was reported by Gorantla et al\textsuperscript{18} using Mask RCNN for ROI extraction and Hierarchical Convolutional Mixture of Experts (HCME) neural network for classification.

Deep learning models have also been proposed for diagnosis of cervical cancer by different literatures.\textsuperscript{20–24} Guo et al\textsuperscript{20} proposed a localized fusion-based approach to classify squamous epithelium into four classes, training a support vector machine (SVM) and linear discriminant analysis (LDA) using 61 digitized histopathologic hematoxylin and eosin-stained images. Similarly, Almubarak et al\textsuperscript{21} used 65 digitized histology images to train a CNN for classification of cervical cancer and an accuracy of 77.27\% was claimed. Likewise, Sornapudi et al\textsuperscript{23} proposed an epithelium detection module using 150 whole-slide images (WSI) and reported 85\% accuracy, but this was limited to classifying only WSIs. Wei et al\textsuperscript{24} proposed a method to classify the histopathology image as normal, cervical intraepithelial neoplasia (CIN) lesion, and cancer using 90 histology images based on texture and lesion feature, yielding an accuracy of 90\%. Better cervical cancer classification accuracies were reported by training a deep learning model with a relatively large dataset.\textsuperscript{25,26} However, these approaches were magnification power-dependent and also limited to classifying a few classes of cancer.

Most of the studies proposed in the literature were designed to classify either the cervix type or few types of cervical cancers. Moreover, the generalization abilities of the models are low, and the classifications were limited only to a few classes and applicable for images acquired using specific magnification powers. In this paper, a full-fledged, integrated, magnification power independent and accurate system is proposed for cervical cancer screening and diagnosis by automating both the pre-screening (cervix type classification) and cervical cancer type classification procedures.

**Methods**

**Cervix Type Classification**

*Figure 1* demonstrates the general steps of the proposed cervix type classification system. After data collection, annotation, and labeling, all the cervix images were first fed to the object detection network for cervix localization (ROI extraction). Then, the ROI extracted images were used to train and validate the cervix type classification model.

**Dataset**

A total of 4005 images were collected for cervix type classification. 133 colposcope images were collected from Tercha General Hospital (south west Ethiopia) using a speculum (to access the cervix), and 13 MP and 18 MP Tecno smartphone cameras after application of 5\% acetic acid on the cervix. The collected images were then labeled into three classes (type 1, type 2 and type 3) by three expert gynecologists: an integrated emergency surgical officer, and senior medical doctors. The other images were collected from the publicly available online Kaggle dataset.\textsuperscript{27} Ethical approvals from the institutional research review board of Jimma Institute of Health, Jimma University, and a written informed consent forms from all subjects, were obtained prior to data collection.

**Detection of Cervical Region (ROI Extraction)**

Before feeding images to the detection network, annotation was performed using the Dark Label annotation tool. The data format of annotation was Pascal Visual Object Class (VOC) and was set to detect only one object or cervix.

The MobileNetv2-YOLOv3 lightweight network was used for localizing the cervix (ROI extraction). YOLOv3 is the third version of you only look once (YOLO) which predicts at three different scales. It divides an image into grid cells,
produces the cell’s probabilities, and predicts boxes. The detection layer makes detection at feature maps of three different sizes, having strides 32, 16, and 8 which are factors to down sampling image.\textsuperscript{28} MobileNetv2 is used as a backbone to facilitate migration to the mobile terminal and improve the generalization ability of the model and extract features. Furthermore, it combines the anti-residual module with the depth wise separable convolution. The depth wise separable convolution can effectively improve the computational efficiency and reduce the number of parameters. This enables the network to perform well with the lightweight character.\textsuperscript{24,29}

Prior to training, the network was configured according to the cervix type dataset. The dataset was split into 80\% for training, 10\% for validation, and 10\% for testing with respective annotation files. The annotation files and images were converted to TFRecord file format. Then, the input image, which was originally 416x416, was down-sampled using strides of 32, 16, and 8 resulting in 13×13, 26×26, and 52×52 output images. The detection network was then trained with these data using an Adam optimizer, 0.001 initial learning rate and batch size of 8.

**Cervix Type Classification Model**

For the cervix type classification, four pre-trained models were selected, trained, validated and compared. The selected models were VGGNet, ResNet, EfficientNet, and Resnet50 ensembled with MobileNetV2. The models were trained and evaluated twice using two types of datasets: (1) ROI extracted images and (2) raw (ROI unextracted) images. Adam and stochastic gradient descent optimizers with learning rates of 0.001–0.0001 and batch size of 32 were used for all models. A 10-fold cross-validation technique was used to partition the data. The dataset was split 10-folds and set to 90\% training and 10\% validation in each fold. Moreover, the L2 regularization technique\textsuperscript{30} was applied in each layer to reduce overfitting.

**Cervical Cancer Classification**

**Figure 2** demonstrates the procedures used to train and validate the cervical cancer classification system. These include data collection, augmentation, preprocessing, model training, model validation and testing.

**Dataset**

A total of 915 hematoxylin and eosin (H&E)-stained histopathology images labeled with four cancer classes (normal, precancer, adenocarcinoma, and squamous cell carcinoma), were gathered from Jimma University Medical Center (JUMC) and St. Paul Hospital for the cervical cancer classification. A smartphone camera and digital camera placed on an OPTIKA light microscope, as well as a digital slide scanner, were used to acquire the images. Microscopic images were acquired by 4x, 10x, 40x and 100x magnification powers with a resolution as small as 419×407 and as large as 2048×1536.
H&E stain is one of the primary tissue stains for histology. This is because H stain makes nuclei easily visible in blue against a pink background of cytoplasm and other tissue regions. This enables a pathologist to easily identify and evaluate the tissue; a highly manual process. For automated image analysis, these H&E stained images need to be normalized. This is because of the significant variation in image color arising from both sample preparation and imaging conditions. Therefore, normalizing H&E images is important for classification. Histogram matching technique was used as a color normalizer to reduce the color variation of the images.

Cervical Cancer Classification Model
For cervical cancer classification, the state-of-the-art pre-trained model called EfficientNetB0 was trained and validated. The SoftMax layer of the model was trained with raw histopathology image (without applying histogram matching) and preprocessed image (histogram matched image) for comparison. Adam optimizer with a learning rate of 0.0001, 50 number of epochs and batch size 32 were used for training. The comparison was performed to evaluate the generalization ability of the model with color normalized and raw images. Finally, unseen (new) data were used to test the performance of the model.

Performance Evaluation Metrics
Accuracy, precision, recall, F1 score, kappa score, and ROC-AUC plot were used to evaluate the performance of the models. The metrics were calculated from the model’s confusion matrix based on TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative) values as demonstrated in Equations (1–6). AUC-ROC plot was used to visualize how well the model can distinguish between classes.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{2}
\]

\[
\text{Precision} = \frac{TN}{TN + FP} \tag{3}
\]

\[
\text{Specificity} = \frac{TP}{TN + FP} \tag{4}
\]
\[ F - \text{measure} = \frac{2 \times \text{Recall} \times \text{precision}}{\text{Recall} + \text{precision}} \] \hfill (5)

\[ \text{Kappa score} = \frac{\text{Observed Agreement} - \text{Expected Agreement}}{1 - \text{Expected Agreement}} \] \hfill (6)

**Results**

**ROI Detection and Cervix Type Classification**

**ROI Detection**

Figure 3 demonstrates the loss versus epoch graph of the cervix ROI detection model. Both the training and validation losses decreased inversely with epoch, yielding a training loss of 0.48 and validation loss of 0.39 at the 60th epoch. The network mean average precision (mAP) was 99.89%. Figure 4 show the sample images generated from the detection model. The results demonstrate that, the detection model was able to remove the irrelevant region from the images preserving the cervix region. Figure 5 shows the intersection over union (IoU) score for some selected test images generated by the model, demonstrating how accurately the model predicted the ROI with respect to the ground truth. The model performs well and the IoU score was greater than 0.5.

![Figure 3 Performance of ROI detection model (A) Training loss (B) Validation loss.](https://doi.org/10.2147/MDER.S366303)
Result of Cervix Type Classification

Table 1 demonstrates the training results of VGG16, ResNet50, EfficientNetB0 and the hybrid model (ResNet50 with MobilenetV2) for cervix type classification trained using raw data and ROI extracted data. As indicated, EfficientNetB0 model trained with ROI extracted image outperformed the other models with an accuracy of 92.19%, and it was selected for cervix type classification. Better results were obtained for all models trained with an ROI extracted dataset compared to models trained with a raw dataset. 

Figure 6 illustrates the selected EfficientNetB0 model accuracy and loss curves.

The selected model (EfficientNetB0) was further evaluated using a test dataset. Test results are demonstrated using the confusion matrix and ROC-AUC plot in Figure 7. The precision, recall, F1-score and Kappa score are calculated from

Figure 4 Sample results of the Cervix ROI extraction model (A) input images (B) ROI extracted images.

Figure 5 Sample test images and their respective IoU score of ROI detected images.
the values of the confusion matrix. An overall test accuracy of 96.84% and kappa score of 0.95 were achieved using the EfficientNetB0 model for cervix type classification. Table 2 illustrates the summary of classification results.

**Cervical Cancer Classification Results**

**Preprocessing Results**
The first step of cervical cancer classification was preprocessing histopathologic images. Image resizing and color normalization using a histogram matching technique to all the dataset, and data augmentation to the training dataset, were

<table>
<thead>
<tr>
<th>Dataset Used</th>
<th>Model</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw cervix image dataset</td>
<td>ResNet50</td>
<td>84.1</td>
</tr>
<tr>
<td></td>
<td>VGG16</td>
<td>79.67</td>
</tr>
<tr>
<td></td>
<td>Hybrid ResNet50 and MobileNetV2</td>
<td>88.1</td>
</tr>
<tr>
<td></td>
<td>EfficientNetB0</td>
<td>89</td>
</tr>
<tr>
<td>ROI extracted cervix image dataset</td>
<td>ResNet50</td>
<td>90.73</td>
</tr>
<tr>
<td></td>
<td>VGG16</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Hybrid ResNet50 with MobileNetV2</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>EfficientNetB0</td>
<td>92.19</td>
</tr>
</tbody>
</table>

**Table 2** Summary of classification results.
applied prior to model training. Initially, the images were resized to 224×224. Then augmentation technique using image transformation (45°, 90°, 180°, and 270° rotation) was applied to all images. Figure 8 demonstrates sample image and augmented images using affine image transformation. The effect of histogram matching on sample images is demonstrated in Figure 9.

Cervical Cancer Classification

The training result of EfficientNetB0 model for classifying histopathological images into normal, precancer, adenocarcinoma, and squamous cell carcinoma is shown using the accuracy and loss curves in Figure 10. The highest validation accuracy, 97.8%, was achieved at the 50th epoch.

The model was also further tested using a test dataset. The confusion matrix and ROC-AUC plot (Figure 11) demonstrates the test results of the model. The AUC scores for all classes were close to 1 confirming the model’s good measure of separability. Table 3 presents the summary of the model’s test performance. The model was found to be 94.5% accurate in classifying the test dataset.

![Confusion matrix](image1.png)  ![ROC-AUC plot](image2.png)

**Figure 7** Test results of EfficientNetB0 model for cervix type classification (A) Confusion matrix (B) ROC-AUC plot.

<table>
<thead>
<tr>
<th>Classes</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-Score (%)</th>
<th>Kappa Score</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td>94.06</td>
<td>97.82</td>
<td>96</td>
<td>0.95</td>
<td>96.84</td>
</tr>
<tr>
<td>Type 2</td>
<td>96.37</td>
<td>97.79</td>
<td>97</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 3</td>
<td>100</td>
<td>94.92</td>
<td>97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2** Result of Cervix Type Classification with Different Metrics for the Final Model
A user interface has been developed for ease of use of the proposed system (sample snapshot demonstrated on Figure 12). The UI includes a “Browse” button to browse images from the local machine and the “Preprocess” button that allows the preprocessed images to be displayed. The “Result” button will return the models predicted cervix/cancer type.

User Interface (UI)

A user interface has been developed for ease of use of the proposed system (sample snapshot demonstrated on Figure 12). The UI includes a “Browse” button to browse images from the local machine and the “Preprocess” button that allows the preprocessed images to be displayed. The “Result” button will return the models predicted cervix/cancer type.

Discussion

In this paper, a deep learning approach has been employed for cervix transformation zone detection, and for cervix type and cervical cancer classification. A total of 4005 cervix images were collected from local health facilities and online public datasets for training and validating deep learning models for the cervix type classification. For the cervical cancer classification, a total of 915 cervical tissue biopsy images were collected from Jimma University Medical Center and St. Paul Hospital.

The cervix type classification system was developed, initially by extracting the ROI from raw cervix images. MobileNetv2-YOLOv3 end-to-end object detection model was used for extracting the ROI. The model input size was 416×416 and the detection was conducted at three different scales by downsampling the image by 32, 16, and 8. Then the model crops the images preserving the required ROI using bounding box vertices as demonstrated in Figures 4 and 5. A mAP of 99.89% was obtained using the proposed detection model. Then, deep learning models including VGG16, ResNet 50, an ensemble of ResNet50 and Mobilenetv2, and EfficientNet families were trained, validated and compared.
for cervix type classification using raw and ROI extracted cervix images. The models were trained with an initial learning rate of 0.00001, and by applying L2 weight decay regularizes in each layer. A better result was obtained using the EfficientNetB0 model trained with ROI extracted cervix images, yielding a validation accuracy of 92.19% and test accuracy of 96.84%.

For the cervical cancer classification, the same EfficientNetB0 model was trained with an 0.0001 learning rate, an Adam optimizer and batch size of 32. A validation accuracy of 97.8%, and test accuracy of 94.5% were achieved for classifying histopathology images into four classes (normal, precancer, adenocarcinoma, and squamous cell carcinoma).

Figure 10  Training and validation performance of EfficientNetB0 model for cervical cancer type classification (A) Accuracy and (B) Loss.
The current work presents an integrated cervical cancer screening system with better performance compared to previous literatures proposed for the same purpose, which were designed to classify either only the cervix type or a few types of cervical cancers. Moreover, the generalization abilities of the models are low, and the classifications were limited only for microscopic images acquired using specific magnification powers. In summary, in this paper, a full-fledged, integrated, magnification power independent and robust system is proposed for cervical cancer screening and diagnosis system by automating both the pre-screening (cervix type classification) and cervical cancer type classification tasks. We acknowledge that the staging of cervical cancer has not been included in this work. Moreover, adding more data may improve the accuracies of the proposed systems.

**Conclusion**

This paper presents two clinically related image classification systems: cervix type and cervical cancer classification. The cervix type classification model classifies cervix images into type 1, type 2, and type 3 after detecting a ROI, and the cervical cancer classification model classifies histopathology images into normal, precancer, adenocarcinoma, and squamous cell carcinoma types. A mean average precision (mAP) of 99.88% for the ROI detection and, average classification test accuracies of 96.84% and 94.5% were achieved for the cervix type and cervical cancer classification, respectively. The results obtained demonstrate that, the proposed system has the potential to be used as a decision support system for physicians, especially those in low resource setting, where both the experts and the means are limited.

**Table 3** Result of Cervical Cancer Classification with Different Metrics

<table>
<thead>
<tr>
<th>Classes</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
<th>Kappa Score</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenocarcinoma</td>
<td>95.65</td>
<td>100</td>
<td>98</td>
<td>0.92</td>
<td>94.5</td>
</tr>
<tr>
<td>Normal</td>
<td>78.85</td>
<td>91.67</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precancer</td>
<td>92.86</td>
<td>92.85</td>
<td>93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squamous cell carcinoma</td>
<td>100</td>
<td>93.02</td>
<td>96</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Figure 11](https://example.com/figure11.png)

**Figure 11** Test results of EfficientNetB0 model for cervical cancer type classification (A) Confusion matrix (B) ROC-AUC plot (0: adenocarcinoma, 1: normal, 2: precancer and 3: squamous cell carcinoma, respectively).
Data Sharing Statement
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate
This research has been approved by Jimma University’s research ethics institutional review board (IRB). This research has been performed in accordance with the ethical standards as laid down in the 1964/Declaration of Helsinki and its later amendments or comparable ethical standards.

Acknowledgment
The resources required for this research were provided by the school of Biomedical Engineering, Jimma Institute of Technology, Jimma University and Jimma University Medical Center (JUMC), St. Paul Millennium Hospital and Tercha General Hospital.

Author Contributions
All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically
reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Resources required for this study were provided by the School of Biomedical Engineering, Jimma Institute of Technology, Jimma University, Ethiopia.

Disclosure

The authors declare that they have no competing interests.

References

