Mallikarjun Gachchannavar. International Journal of Engineering Research and Applications
www.ijera.com
ISSN: 2248-9622, Vol. 13, Issue 10, October 2023, pp 46-53

RESEARCH ARTICLE OPEN ACCESS

Nucleus Detection for Cervical Cancer Using Deep Learning Techniques: A Review

Mallikarjun Gachchannavar
Lecturer
Department of Electronics and communication /Government Polytechnic Hubli
mallikarjung346@gmail.com

ABSTRATCT
The health of the female reproductive system is significantly impacted by cervical cancer. The development of optical coherence tomography (OCT) as a non-invasive, high-resolution imaging method for the diagnosis of cervical illness. However, OCT picture annotation requires a lot of expertise and takes a long time, which makes it difficult to train deep learning-based classification models. The purpose of this work is to build a self-supervised learning-based computer-aided diagnosis (CADx) method for categorizing in-vivo cervical OCT pictures. Methods: The proposed CADx technique develops a contrastive texture learning (CTL) strategy to take use of the texture characteristics in unlabeled cervical OCT images in addition to high-level semantic data recovered using a convolutional neural network (CNN). On the OCT image dataset from a multi-center clinical trial including 733 Chinese patients, we performed ten-fold cross-validation. Results: Our method outperformed two out of four medical experts on the test set and achieved an area-under-the-curve value of 0.9798, 0.0157 with a sensitivity of 91.17%, 4.99% and a specificity of 93.96 4.72% for OCT image patches in a binary classification task for detecting high-risk diseases, including high-grade squamous intraepithelial lesion and cervical cancer. Furthermore, employing a cross-shaped threshold voting mechanism, our method demonstrated 91.53% sensitivity and 97.37% specificity on an external validation dataset made up of 287 3D OCT volumes from 118 Chinese patients at a new hospital.

Keywords: cervical cancer, self-supervised learning, local binary pattern, optical coherence tomography, visualization

I. INTRODUCTION

In 2018, cervical cancer had the fourth-highest incidence and fatality rate among women.1 In 2018, there were over 570,000 new cases and 311,000 deaths from cervical cancer2, and more than 80% of these cases and fatalities took place in underdeveloped and developing nations. It has been shown that the human papillomavirus (HPV) is the primary cause of nearly all occurrences of cervical cancer. Fortunately, cervical cancer has been recognized by the World Health Organization (WHO) as a public health issue, and particular measures have been taken to hasten its eradication. These measures include efficient HPV vaccination, cervical screening, and prompt treatment of precancerous lesions.

The HPV test, thin-prep cytologic test (TCT), and colposcopy are a few examples of common clinical screening techniques for detecting cervical illness. However, each of these approaches has drawbacks of its own. For instance, the HPV test has a comparatively high percentage of false-positive results, while colposcopy and TCT both have a comparatively high risk of missing diagnosis of high-risk conditions. cervical cancer. The gold standard for the diagnosis of cervical lesions is a cervical biopsy with histological confirmation. But it is invasive, time-consuming, and there is a fair amount of missed diagnosis because of poor biopsy sampling. To meet the WHO 2030 objectives, quick, non-invasive, efficient, and intelligent screening and diagnostic methods must be developed [6].

A new non-invasive three-dimensional (3D) imaging method called optical coherence tomography (OCT) was created in the early 1990s. To provide detailed cross-sectional pictures of biological systems, OCT devices employ light waves. OCT is currently used extensively in ophthalmology, cardiology, and other clinical medical specialties. A precise diagnosis may also be made using OCT imaging of the cervix because recent studies5, 8, and 10 have demonstrated that this technology can capture histomorphology characteristics of cervical tissue both exes vivo and in vivo. Cervical OCT pictures, however, are not known to many pathologists and gynecologists, suggesting a steep learning curve. Therefore, to aid them in more effectively analyzing cervical OCT pictures, we need to create intelligent computer-aided diagnosis (CADx) methods.

Deep-learning-based CADx techniques for medical pictures with various modalities have outperformed conventional machine learning algorithms lately. The same holds true when examining cytology and colposcopy pictures used in cervical screening. For instance, ***Almubarak et al.*** classified histology pictures from cervical tissue samples on a small-scale dataset with 77.25% accuracy by utilizing multi-scale features. Support vector machine (SVM), k-nearest neighbor (k-NN), and artificial neural network (ANN) models were used by ***Bhargava et al.*** to train Pap smear image classifiers using manually created features. On a limited dataset, they discovered that the ANN-based classifier completed the test the best. A hybrid transfer learning strategy for cervical cancer detection was proposed by Kudva et al.13 utilizing two convolutional neural networks (CNNs): AlexNet and VGG-16. A heuristic and ANN-based classification model for early cervical cancer screening using linear SVM was proposed by Priya et al. Two pre-trained deep learning models make up the pipeline ***Alyafeai et al.*** presented for completely automated categorization of cervical cancer from cervigram pictures. Using time-lapsed colposcopic images, Li et al.16 suggested a deep learning framework for recognizing cervical intraepithelial neoplasia (CIN) and cervical cancer. They put the proposed framework to the test and discovered that it performed just like a colposcopist who was already working. The categorization outcomes of a few recent studies for the analysis of cytology and colposcopy pictures are listed in Table I.

Image type	Literature	Model	Performance	Data size
Pap smear test	Ref.[12]	deep learning	accuracy: 95.5%	66 images
Pap smear test	Ref.[21]	deep learning + SVM	accuracy: 98.32% and 97.87% for two datasets	dataset1: 917 images dataset2: 4,049 images
colposcopy	Ref.[13]	deep learning	accuracy: 91.46%	1,644 images
colposcopy	Ref.[16]	deep learning	accuracy: 78.33%	7,668 patients
colposcopy	Ref.[17]	SVM	sensitivity: 81.3% and specificity: 78.6% for detecting CIN+	134 patients
colposcopy	Ref.[18]	deep learning	accuracy: 81.35%	800 images
colposcopy	Ref.[19]	deep learning	accuracy: 73.08%; AUC: 0.75 for detecting CIN2/CIN3+	1,709 patients
colposcopy	Ref.[20]	deep learning	accuracy: 85.5%; AUC: 0.909 for detecting HSIL+	1,400 patients
cytology	Ref.[21]	deep learning + SVM	accuracy: 99.47%	963 images
cytology	Ref.[22]	deep learning	accuracy: 88.84%; F-score: 59.96%	14,432 image patches

Table No. 1: cervical screening images: summary of chosen classification models

Additionally, SSL was employed to reduce the critical need for data annotation for medical photos. Using a little amount of annotated data, Zhu et al.24 suggested a pretext task that classifies slices taken from 3D computed tomography (CT) and magnetic resonance imaging (MRI) volumes. Similar research was done by ***Taleb et al.***, who also made their implementations available as open-source libraries. By learning translation and rotation invariant characteristics from the initial 3D volumetric scans, these 3D SSL algorithms may make use of the whole 3D spatial context, enhancing data efficiency and performance on many downstream tasks. A context restoration technique was also put out by Chen et al.26 to acquire useful semantic information from three different kinds of medical pictures.

In this article, we provide a self-supervised contrastive texture learning technique to address the issue of costly OCT picture annotation. It may use the high-level semantic characteristics and local texture information that a CNN model collected to automatically learn visual representations of OCT images for various cervical illnesses. After pre-training, the CNN model may be fine-tuned for classification using only a small dataset of labeled cervical OCT images, which eases the load on picture annotation. We also created a CADx method to help gynecologists analyze cervical OCT pictures quickly in real-world settings. The CADx technique comprises of a feature visualization module that highlights potential lesions for gynecologists and an image classification model that has already been trained using the suggested SSL procedure.

II. DATASETS

Many of the state-of-the-art techniques, as can be observed in Section 6, do not adequately

perform on bigger datasets of real photos with rather high levels of variability. This is mostly because there isn't a comprehensive dataset of cervical cytology photos and annotations that is openly accessible. To do this, a more difficult collection of actual cervical cytology pictures is offered. The aim of the dataset construction was to create a dataset that is highly representative of typical cervical cytology pictures to reproduce the characteristics of such datasets in practice arXiv:1811.09651v1 [cs.CV] 23 Nov 2022.

The Moffitt Cancer Center (Tampa, FL) donated archival ThinPrep Pap-stained slides for this research. Using an integrated hardware-software microscope system (Stereologer, SRC Biosciences, Tampa, FL), images from these slides were collected in a methodical, random way. The same cytotechnologist inspected each slide under a microscope and assessed them using The Bethesda System (TBS) [2]. Slides were categorized as Negative, Low-grade Squamous Intraepithelial Lesions (LSIL), or High-grade Squamous Intraepithelial Lesions (HSIL) for this study. The top and bottom of each slide at a manually chosen focal plane were manually determined using the automated XYZ stepping motor and Stereology software. Each picture has a resolution of 1280 × 960 pixels. The dataset consists of four primary sections, which are described below.

1. IMAGE STACK: The dataset contains 93 stacks of photos at a 40x magnification. From top to bottom of the slide, each stack has 10–20 photographs that were taken at evenly spaced fields of views (designated frame000–frame092).

2. Extended Dept of the Filed (EDF): By using the process outlined in [12], an EDF image was produced from each stack and added to the EDF folder.

3. Frame/Label: The labels N, L, or H (for Negative, LSIL, or HSIL, respectively) are placed next to each frame. Additionally, the entire set of frames was divided into training and test sets (approximately 25% of the frames from each label were placed in the test set) depending on the slides from which these frames were obtained and their visual difficulty level. Label.csv contains frame labels and training/test partitioning, with 0 and 1 denoting the training and test sets, respectively.

4. Coordinates: Each frame was inspected, and a manual marking was made at a specific location inside each cervical nucleus. Only when more than half of a cervical nucleus is found within a frame and when the center of the nucleus is around 10 pixels away from the boundary is a cervical nucleus identified as touching the border or as partially inside a frame.

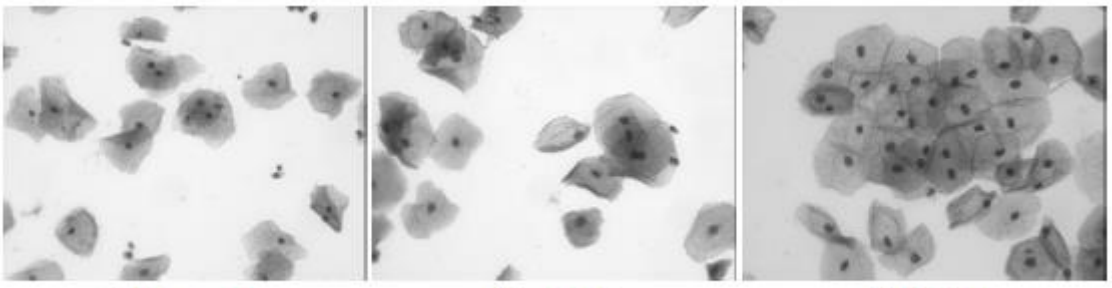

Figure No. 1: Frames from each of the grade categories

The multi-center clinical trial that used OCT to assess cervical lesions in vivo generated the OCT picture collection.5 From August 2017 to December 2019, seven hundred thirty-three gynecological outpatients were enrolled at five hospitals in China. One or both of the cervical screening findings (i.e., the HPV and TCT results) being positive is the standard for patient recruitment. OCT examinations and colposcopy-directed cervical biopsies were performed on each patient enrolled in this multi-center clinical research.

	Negative	LSIL	HSIL	Total
Training	12	34	23	69
Test	4	12	8	24
Total	16	46	31	93

Table No. 2: Number of Frames within Each Grade Category

	Negative	LSIL	HSIL	Total
Training	179	1125	679	1983
Test	59	411	252	722
Total	238	1536	931	2705

Table No. 3: Number of Nuclei within Each Grade Category

Ten or twenty frames, each of which corresponds to an in-vivo cervical OCT picture, make up a 3D OCT volume taken from the position of a 12-hour clock on the cervix. Each 3D OCT volume's class label was established in accordance with the biopsy specimen's associated histopathology-confirmed diagnosis. We chose 1,256 3D OCT volumes from 699 patients that matched the appropriate hematoxylin and eosin (H&E) stained pictures because of mistakes in biopsy collection and image quality problems. Using the accompanying H&E-stained picture as a guide, pathologists then isolated patches of 600600 pixels that comprise the lesion region from each chosen OCT image. This is done because normal cervical tissues and local lesions frequently coexist in an OCT image (i.e., one frame of an OCT volume).

III. EVALUATION MEASURES

A cell of binary segmentation masks (for each segmented nucleus) or a list of the coordinates of identified points for nuclei are three options for the assessment code in MATLAB that is provided with the dataset. A detection is considered True Positive (TP) in the case of a list of coordinates if it is close to a point (less than 10 pixels away) in the ground truth. There can only be one TP produced by each split region. A point in the ground truth that does not lie within any segmented areas is referred to as a False Negative (FN). If more than one point is included inside a segmented region and the points are not covered by any other segmented region, all but one point will be considered as FN. A detected point that is more than 10 pixels away from a manual point or any segmented zone that excludes a manual point is referred to as a False Positive (FP). The assessment measures employed to present the findings are.

$$Precision = \frac{TP}{TP + FT}$$

$$Recall = \frac{TP}{TP + FN}$$

The algorithm also calculates the Standard Deviation (STD) of the Precision and Recall metrics inside each individual picture.

IV. BASELINE METHOD AND TRAINING

We use an updated version of the nucleus segmentation algorithm put out in [13] as a foundational technique for dividing the cervical region of the proposed dataset. The main differences are that the parameters are set using the training set rather than choosing them empirically because there is a large training dataset, that the check for regions with a relatively low inside-to-outside average intensity difference is removed because it is more of an image-specific characteristic, and that the removal of regions that overlap the boundary is added because the majority of the nuclei touching the boundary are not annotated except that they are mostly inside the boundary. The technique initially applies a 2-D adaptive noise-removal filter to blur the picture [14] before repeatedly binarizing the resultant image using a series of rising thresholds. Finding seed spots inside the nuclei and subsequently growing the areas by raising the threshold in consecutive thresholdings were the objectives of the iterative thresholding. In the iterative process, regions that were too tiny or too concave (by gauging a region's solidity) were eliminated. Additionally, a permissible range of the average intensity inside each nucleus was established, and areas might merge if the newly combined, bigger region was more solid than all the individual smaller regions.

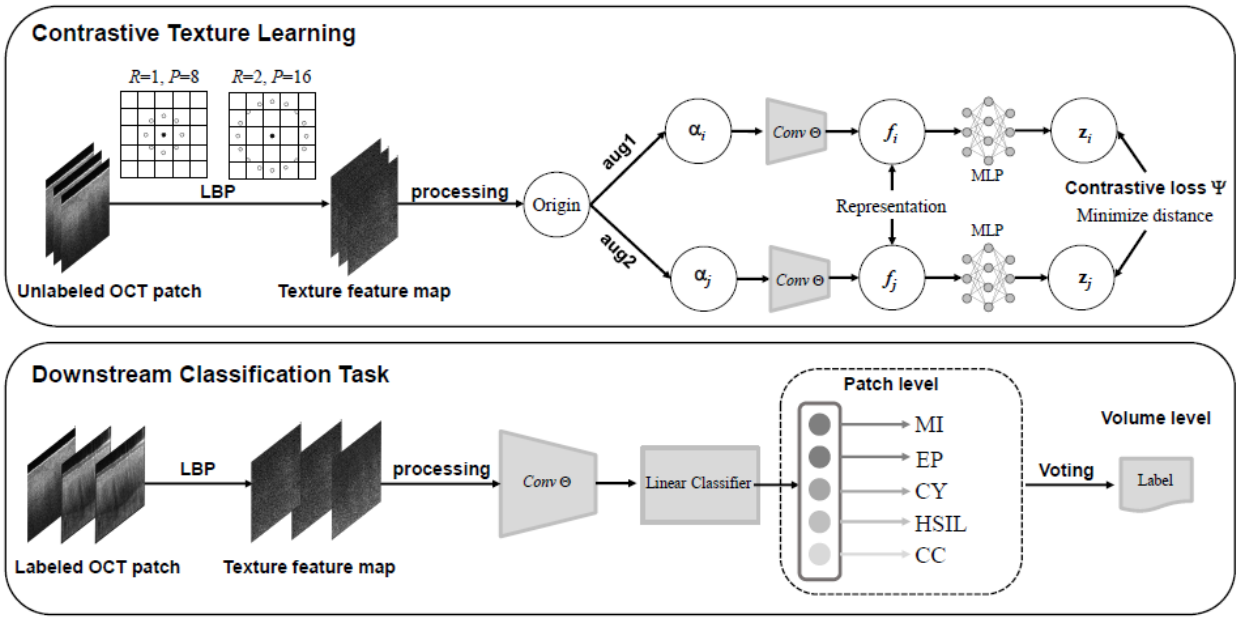

Figure No. 2: Training Methods of Classification Model

A multi-granularity OCT image classification model for patches, pictures, and 3D volumes forms the basis of our CADx strategy. This categorization model has been set up and trained on its own. Typically, SSL comprises two phases. To start, it establishes a particular pretext job for pre-training a model on a sizable unlabeled dataset. Learning representations is the goal of the pretext task. Contrastive Texture Learning Contrary loss LBP processing Task for Downstream Classification Linear Classifier Processing LBPC on vBlock level MIEPCYHSILCCR = 1, P = 8, R = 2, P = 16a blank OCT patch with texture feature MLPMLP Representation reduce distance Labeled OCT patch, Conv, Conv Feature texture map Label Volume level 6 for tasks that come after. Second, the pre-trained model is adjusted for certain downstream tasks using a small-scale annotated dataset. An overview of the development and training of the categorization model is shown in Figure 2.

The training set in the dataset was used to train and determine the least and maximum acceptable values for the average intensity, as well as the minimum size and the minimum acceptable solidity. Finally, the areas that overlapped the boundary were eliminated. As opposed to the technique in [13], the parameters were chosen based on the training set evaluation of the algorithm. A grid search was used to investigate various parameter combinations. The combination of parameters with the greatest F measure,

$$\frac{2.Precision.Recall}{Precision + Recall}$$

To evaluate the effectiveness of the approach on the test set, one training set was chosen. The approach was tested on the test set using the settings that produced the highest F measure during training.

V. DEEP LEARNING METHOD

To identify cervical nuclei in the dataset, we trained and evaluated a convolutional neural network (CNN). Two fully connected layers (FCL) and two convolutional layers (CL) make up the intended CNN architecture. Max pooling layers (MPL) came after each CL, whereas ReLu activation function (RL) came after each FCL. To create the probability of each patch being a nucleus or not, the output layer (OL) containing two neurons was followed by a soft max loss layer (SFL).

The grayscale frames in the training dataset were extracted to create image patches of 75x75 pixels. Patches were evenly sampled every 15 pixels, and if they fell within 15 pixels of a manually defined point, they were classified as positive; if not, they were labeled as negative. In 100 epochs, the network was trained using a constant learning rate of 0.001. Only 1.71% of the total 654,948 patches that were retrieved and utilized to train the network were classified as positive.

The test sets frames' extracted patches were classified using the trained CNN. In comparison to the training dataset, the patches were retrieved at a closer interval of 3 pixels. A threshold of 0.5 was used to construct, dilate, and threshold the hit map. Smaller than 100-pixel portions that didn't correspond to the nuclei were eliminated.

5.1: IMAGE CLASSIFICATION

A linear classifier is utilized in place of the MLP used for the pretext job in order to acquire more generic feature information for the downstream task. This linear classifier specifically consists of a fully connected layer and a softmax function. By utilizing the characteristics discovered during the pre-training procedure, the network is refined on a small number of annotated cervical OCT image patches. Equation (5) formulates the cross-entropy loss function for classification. After the fine-tuning procedure is complete, it creates a specific label for each input patch and outputs the appropriate probability distributions for each of the five categories (i.e., MI, EP, CY, HSIL, and CC). By adding the probabilities for each category inside the general class, it is possible to determine the chance that the input patch belongs to a binary class (high-risk or low-risk).

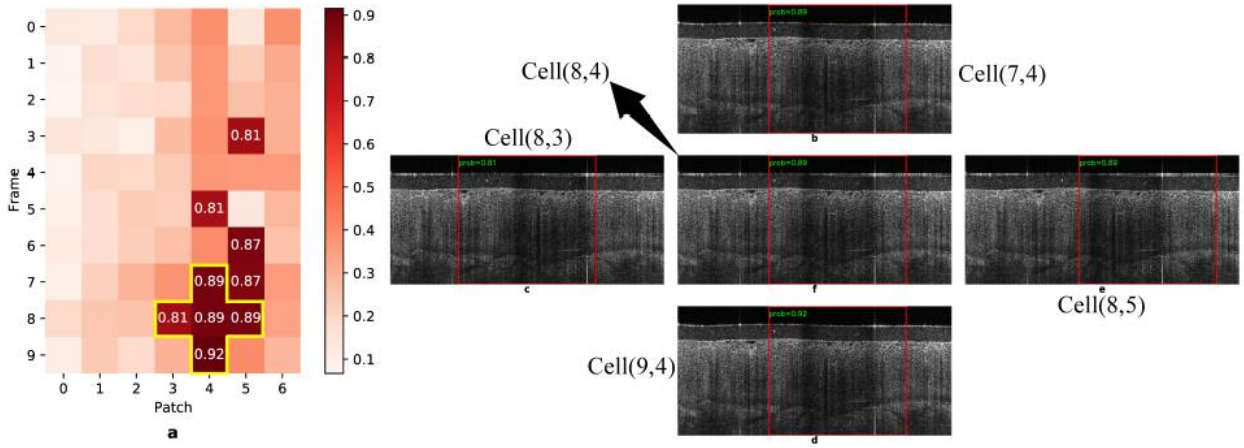

Figure No. 3: An example of the threshold voting method using a cross

We employ slicing windows to retrieve n patches from a single frame in an OCT volume of mm frames. Therefore, we can get a mmn matrix of patch predictions. Additionally, in this work, we develop a voting approach called cross-shaped threshold voting to combine patch-level predictions into a single outcome for the entire OCT volume in accordance with the clinical experience of medical specialists. This method's main goal is to find a series of high-risk patches that may resemble a cross in both the axial and cross-sectional axes. In addition, we establish a high-risk (or positive) threshold value of 0.8 to distribute the false-positive and false-negative rates more evenly. The cross-shaped threshold voting approach is shown in Figure 2. The fifth patch was in the eighth, ninth, and tenth frames, as well as the fourth, fifth, and sixth patches in the ninth frame. This volume had five high-risk patches that were found by our approach.

5.2: FEATURE VISUALIZATION

To help gynecologists identify cervical lesions, we additionally provide volume- and patch-level feature visualization since the classification model is constructed using the CNN architecture. For an OCT volume, for instance, we provide a heat matrix of size mm to aid in the 3D localization of a potential lesion (see the left portion of Figure 2). Additionally, to improve the interpretation of patch-level diagnoses, we employ the class activation map (CAM)30, which creates a heat map emphasizing areas most pertinent to the associated category based on OCT images' texture properties.

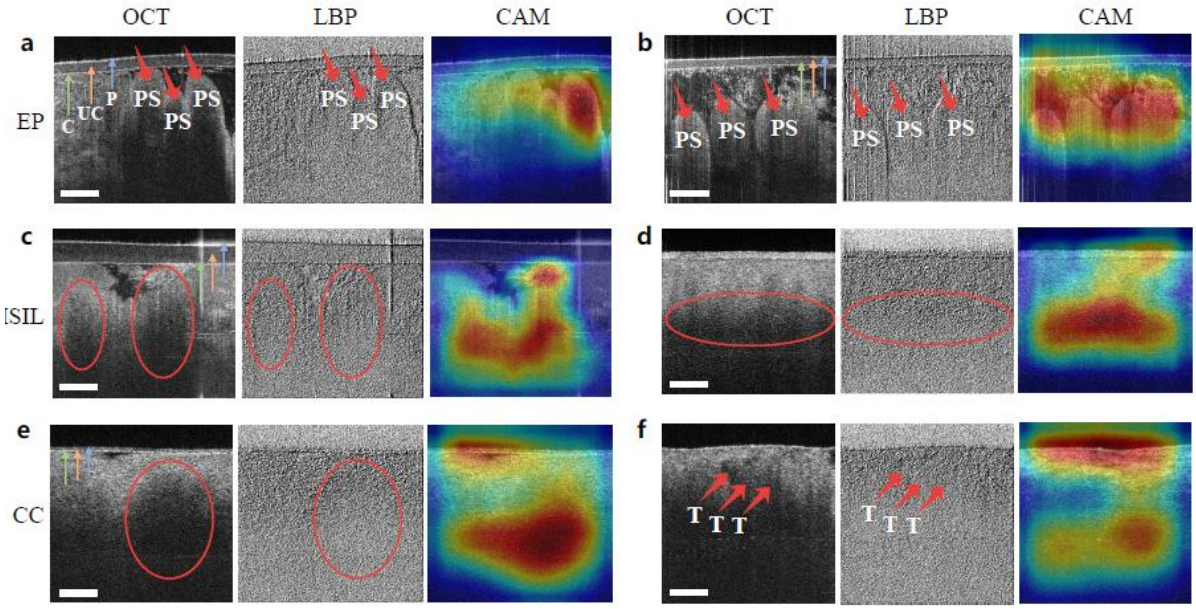

Figure No. 4: Pixel-level histomorphology feature visualization.

Two examples of HSIL were seen in the OCT images in the second row of Figure 5, where the characteristic layered structure and basement membrane are absent. Additionally, as the tissue depth grew, the image intensity rapidly decreased, creating hyper-scattering icicle-like forms (note the areas encircled by the two red eclipses in Figure 5c). The stroma below was no longer visible, and the OCT picture in Figure 5d now showed distinct patterns of light and dark made up of these icicle-like forms (note the area encircled by the red eclipse). These histomorphology features5 were emphasized in the relevant OCT images by the CAM heat maps. Unfortunately, the human-readable texture information needed to comprehend HSIL is not readily available in the LBP maps, which calls for more research.

VI. CONCLUSION

The newly released dataset includes 93 actual cervical cytology pictures that were taken from ThinPrep cytology slides that were classed as Normal, LSIL, and HSIL. The dataset includes picture labels and 2705 hand designated locations inside nuclei that are used as ground truth. The dataset may be used to evaluate how well picture categorization, nucleus segmentation, and detection techniques work in this field of study. This collection is more complete in that it includes more genuine photos, real nuclei that have been manually labelled, and slides of various degrees that were used to capture the photographs. On this dataset, two additional cutting-edge techniques that perform well on the ISBI 2014 and 2015 overlapping cervical cell segmentation challenge [10, 11] datasets struggle. It demonstrates how much more difficult the new dataset is compared to the previous two datasets for the job of nucleus recognition and segmentation. Although the approach in [13] produces better results than the other two ways, it still falls short of the outcomes that have been previously reported for the other datasets.

The test set's baseline technique, as shown in Table 4, may produce results with accuracy and recall of 0.803 and 0.838, respectively. This is far less than what the original technique in [13] was able to do using the dataset made available for the ISBI 2014 and ISBI 2015 competitions. Due to the more challenging job of nucleus recognition and segmentation in this dataset, the technique was able to obtain 0.874 and 0.930 accuracy and recall, respectively, on the ISBI 2014 test set. The test set for the ISBI 2014 challenge consisted of artificial, low-variability pictures. In contrast to the ISBI 2014 and ISBI 2015 datasets, all the photos in this dataset are actual photographs, and the nuclei in various images exhibit far greater variance. The suggested CNN produced the best results. It more than 5% bettered the F measure obtained using the baseline technique. Only around 10% of the nuclei were missed. Most of the missing nuclei have been found to be near the border or to have overlapped nuclei. A more sophisticated post-processing phase may be able to lessen these problems.

REFERENCE

[1]. American Cancer Society, "Cancer prevention & early detection facts & figures 2015-2016," Atlanta: American Cancer Society, 2015.

[2]. Pranav Kumar, SL Happy, Swarnadip Chatterjee, Debdoot Sheet, and Aurobinda Routray, "An unsupervised approach for overlapping cervical cell cytoplasm segmentation," in Biomedical Engineering and Sciences (IECBES), 2016 IEEE EMBS Conference on. IEEE, 2020, pp. 106–109.

[3]. Hady Ahmady Phoulady, Dmitry B. Goldgof, Lawrence O. Hall, and Peter R. Mouton, "A new approach to detect and segment overlapping cells in multi-layer cervical cell volume images," in 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), April 2021, pp. 201–204.

[4]. Jae S. Lim, Two-dimensional Signal and Image Processing, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2021.

[5]. Hady Ahmady Phoulady, Dmitry Goldgof, Lawrence O Hall, and Peter R Mouton, "A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images," Computerized Medical Imaging and Graphics, vol. 59, pp. 38–49, 2022.

[6]. Almubarak HA, Stanley RJ, Long R, et al. Convolutional neural network based localized classification of uterine cervical cancer digital histology images. Procedia Comput Sci. 2017; 114:281–287. https://doi.org/10.1016/j.procs.2017.09.044.

[7]. Bhargava A, Gairola P, Vyas G, Bhan A. Computer aided diagnosis of cervical cancer using HOG features and multi-Classifiers. Proceedings of the 2017

[8]. Alyafeai Z, Ghouti L. A fully automated deep learning pipeline for cervical cancer classification. Expert Syst Appl. 2020; 141:112951. https://doi.org/10.1016/j.eswa.2019.112951.

[9]. Zeng X, Zhang X, Li C, et al. Ultrahigh-resolution optical coherence microscopy accurately classifies precancerous and cancerous human cervix free of labeling. Theranostics. 2018;8(11):3099–3110. https://doi.org/10.7150/thno.24599.

[10]. Asiedu MN, Simhal AK, Chaudhary U, et al. Development of algorithms for automated detection of cervical pre-cancers with a low-cost, point-of-care, pocket colposcope. IEEE Trans Biomed Eng. 2019;66(8):2306–2318. https://doi.org/10.1109/tbme.2018.2887208.

[11]. Alyafeai Z, Ghouti L. A fully automated deep learning pipeline for cervical cancer classification. Expert Syst Appl. 2020; 141:112951. https://doi.org/10.1016/j.eswa.2019.112951.

[12]. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Proceedings of the 3rd International Conference on Learning Representations. 2015. https://arxiv.org/abs/1409.1556.

[13]. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society; 2016; 770–778. https://doi.org/10.1109/CVPR.2016.90.

[14]. Chitra B, Kumar SS. Recent advancement in cervical cancer diagnosis for automated screening: a detailed review. J Ambient Intell Human Comput. 2022;13(1):251–269. https://doi.org/10.1007/s12652-021-02899-2.

[15]. Wu J, Sheng VS, Zhang J, et al. multi-label active learning algorithms for image classification: Overview and future promise. ACM Comput Surv. 2020;53(2):28. https://doi.org/10.1145/3379504.