

AGMC-TU Pap-Smear Cytological Image Dataset: Creation, Annotation, and Analysis towards Early Detection of Cervical Cancer

¹Sourav Dey Roy, ^{1,*}Priya Saha, ^{2,*}Niharika Nath, ³Abhijit Datta, and ^{1,#}Mrinal Kanti Bhowmik

¹Department of Computer Science and Engineering, Tripura University (A Central University), Suryamaninagar-799022, Tripura (W), India

²Department of Biological and Chemical Sciences, New York Institute of Technology, New York, NY 10023, USA

³Department of Pathology, Agartala Government Medical College, Agartala-799006, Tripura (W), India
souravdeyroy49@gmail.com, priyasaha.cse@gmail.com; nnath@nyit.edu; dattaabhijit@hotmail.com; and
mrinalkantibhowmik@tripurauniv.ac.in

Abstract—Automatic cervical cancer screening based on pap-smear images is a highly effective tool where the cells are categorized into normal and abnormal. However, success of most automation tool depends on the accurate extraction of features from the pap-smear images that represent some discriminative characteristics between these two categories of cells. In this paper, we described the designing protocols for creation of a new pap-smear image dataset entitled as AGMC-TU Pap-Smear Cytological Image Dataset. The dataset comprises of 50 normal and 50 abnormal pap-smear images belonging to ethnic and non-ethnic populations of low resource cervical cancer prone regions. Moreover, ground truths of suspicious nucleus regions are annotated in terms of pixel oriented binary masks are also provided with the dataset. Analysis of our dataset includes a conventional (i.e., shape features) and deep feature based study of pap-smear images by dividing them into two major groups: normal and abnormal. Outcome of the analysis clearly differentiates normal and abnormal pap-smear images.

Index Terms—Cervical Cancer; AGMC-TU Pap-Smear Cytological Image Dataset; Feature Analysis; Classification; Performance Evaluation

I. INTRODUCTION

Cancer incidence and mortality rates continue to rise globally. Among all cancers, cervical cancer or cervix uteri carcinoma is the second most frequent cancer in women after breast cancer representing 6.6 % of all female cancer across the globe [1]. It is due to the abnormal growth of cell tissues in the cervix region of female that are spread to the other part of the body. The risk of developing these abnormal changes is connected with an infectious virus called Human Papilloma Virus (HPV) [2]. According to a report on Cancer Burden in North-East (NE) India, the death incidence rate of cervical cancer is still high as compared to the rest of India [3]. The reason for this is low awareness of lifestyle choices and failure of early detection. Different types of cancer control programs are organized in different places of NE India. But these programs are not sufficient to cure cancer from the root and few rural places are still not covered by these government organized programs [4]. Apart from this, there is a severe lack of oncologist experts and hospital infrastructures for people who are affected by cervical cancer.

In general, any cervix abnormality takes 10-12 years for pre-cancerous stage to progress to cervical cancer [5]. Thus cervical cancer is curable if the patient is treated properly at an early stage. Due to the improving medical imaging technology, cervical cytology (i.e., conventional pap test) is considered to be one of the supreme popular screening test for early detection

of cervical cancer and has significant impact in developing countries for effective reduction in death rate [6]. However, clinical examination of cervical cytological images (i.e., pap-smear images) are performed manually which are usually labour-intensive, time consuming and strongly subjective biased. Therefore, automatic detection methods can effectively relieve the laborious tasks required by clinical experts and meanwhile reducing the subjective human errors (i.e., observer independent). Some other benefits for automation of cervical cancer screening are:

- **Cost Effective:** The automatic detection system is cost effective compared to other medical processes because no pathologists or clinical experts are required for detection process.
- **Less Time Consuming:** Generally, manual screening of vast amount of medical data requires more time. This problem can be solve by automating screening technique, which is more accurate, efficient and reliable result.

To automate the screening methods, detection and localization of cell nucleus from the holistic pap-smear images is an evitable task because there is a significant changes in the structural information of the abnormal cells in contrast to the normal cells. Numerous approaches [6]-[13] are proposed in the literature for automatic detection and classification of pap smear images. In contrast, now a days, automatic screening of cervical cancer are stucked by the limitations of feature design and selection of discriminative features. This is because pap-smear images often suffers from low contrast, noise and ambiguity. Moreover, depending upon the fact that detection of abnormality in cervical cells is a challenging task for human experts, potentiality of describing cell abnormality may lie in high level features of pap-smear images but this has been less understood. Recognizing the impact of vision based screening of cervical cancer to the research community and medical experts, our contributions are summarized below:

1. The paper describes designing protocol and characteristics of AGMC-TU Pap-Smear Cytological Image Dataset collected from asymptomatic female population of low resource region of India (i.e., Tripura). The dataset can be used by research community for non-commercial use on request available at [14].
2. Along with the proposed dataset, we annotated the ground truths of the suspicious nucleus regions of normal and abnormal cell images in the form of binary masks.

* Both the second and third authors have equally contributed in the manuscript.
Corresponding Author

- The paper compares and contrast the efficiency of the segmentation methods for nucleus regions (i.e., region of interests in pap-smear images) segmentation so as to identify and select the most outer performed method for abstraction of conventional shape features from the holistic pap-smear images.
- Moreover, the potentiality of conventional shape features and Deep Convolutional Neural Networks (CNNs) as a fixed feature extraction method for pap-smear cells classification (i.e., in terms of normal and abnormal cells) from the holistic pap-smear images and segmented nucleus region of pap-smear images have been evaluated.

Paper Outline. Section II elaborately describes our proposed AGMC-TU Pap-Smear Cytological Image Dataset. Section III describes the analysis of pap-smear images for normal and abnormal pap-smear cell image classification. And finally, the paper is concluded in section IV.

II. MATERIALS: AGMC-TU PAP-SMEAR CYTOLOGICAL IMAGE DATASET

In this section, we will provide detailed description of designing issues and other acquisition factors of the AGMC-TU Pap-Smear Cytological Image Dataset.

A. Designing Issues and Acquisition Factors

To gather significant information from health related data, it is very much essential to maintain a standard protocol for dataset acquisition. For effective acquisition, we have considered several necessary components such as pre-analytical patient preparation, patient information form, collection of pap-smear cells, cell staining and the microscopic imagery system. Each of this components is described elaborately.

Pre-analytical Patient Preparation. To ensure the qualitative processing of cell specimens, female patients from the North-East Indian population those are suspect to have cervical abnormality are instructed to follow guidelines and necessary steps before the pap test, such as:

- Avoid pap-test during the menstrual cycle. To get the best result, pap-test was performed in the middle of the patient's cycle, 10-20 days after the first day of patient's menstrual period.
- Avoid sexual intercourse and wash of vagina/ cervix region 48 hours prior the pap-test.
- For women undergoing treatments, the procedure may cause cervical or vaginal infection and a wait period is recommended to perform the pap-test.

Patient Information Form. On the day of pap-test, a form is filled up by each patient which includes personal information (i.e., name, age, height, weight, contact address, etc.) and background information in connection to the disease (i.e., duration of disease, indications (if any), history of cervical

TABLE I
OVERALL STATISTICS AND CLINICO-DEMOGRAPHIC INFORMATION OF AGMC-TU PAP-SMEAR CYTOLOGICAL IMAGE DATASET

Patient Parameters	Normal (10)	Abnormal (10)
Age: <35	7	1
35-50	2	6
>50	1	3
Tobacco Consumption	2	6
Menarche Age: <12	1	2
At 12	6	2
>12	3	6
Age of Marriage: <18 Years	1	7
>=18 Years	9	3
Number of Children: 1-2	8	3
>2	2	7
Oral Contraceptives Intake	3	5
Overweight	4	6
Multiple Sexual Partners	--	4
Family History of Cervical Cancer	1	5

cancer in her family, medical and pathological tests undergone previously, surgeries (if any), etc.).

Collection of Pap smear Cells. During the Pap test, medical expert uses a vaginal speculum (i.e., metallic Cusco's speculum) which separates the walls of the vagina for careful inspection of cervix region. Then, using an Ayre's spatula, medical expert gently fixes the cervix surface to collect the cells and a cytobrush is used to collect the cells from the inner part of the cervix region that flows towards uterus.

Cells Fixation and Staining. After collecting the cells, medical expert spreads the cells onto a glass slide and staining is performed for identification of features related to nucleus region. Additionally, rapid fixation of the pap-smear slides are needed so as to retain the cell details and further avoid the drying out of the smears. In our study, Ethanol or spray fixative in polyethylene glycol formulation is used for cell fixation. After that Haematoxylin and Polychromic stains are used to stain the nucleus and cytoplasm regions respectively. After staining, clearing is performed in xylol so as to produce cellular transparency before mounting. And finally, DPX is used for cell mounting that is a combination of plasticizer, xylene, and distyrene.

Capturing Pap Smear Images using Light Microscope. After the cell staining process, samples are observed under the light microscope, equipped with a compact digital camera. In this study, sample images are taken randomly through a OLYMPUS SP 350 digital camera which has 8 megapixels resolution with an aperture range of F2.8 to F4.9 and 8mm (38 mm equivalent) zoom wide, adapted to a OLYMPUS CX41 light microscope which has eyepiece of F.N. 20 along with 6V/30W halogen light source and UIS2 (Universal Infinity-corrected) optical system. Finally, images are stored in .jpg format of 132×158 pixels.

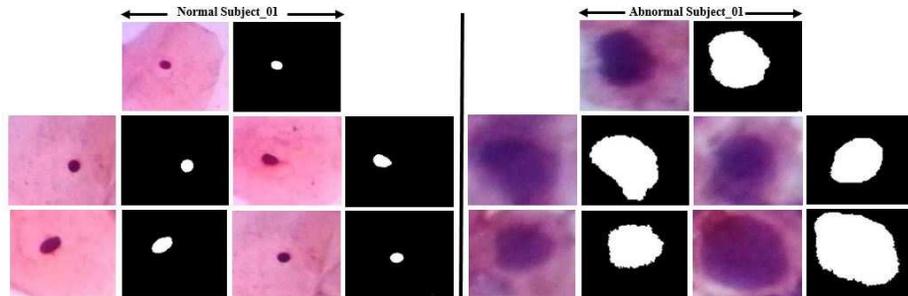


Fig. 1. AGMC-TU Pap-Smear Cytological Image Dataset (a) Original Pap-Smear Cells; (b) Corresponding Ground Truth Images

TABLE II
PERFORMANCE EVALUATION OF THE SEGMENTATION ALGORITHMS FOR ACCURATE SEGMENTATION OF REGION OF INTEREST (NUCLEUS REGION)

Segmentation Methods	Parameters Setting	Similarity Measurement						
		JI	DSC	MSE	Re	Pr	FM	ACC
K-Means [17]	c=2	0.6969	0.7316	0.1024	0.5838	0.7747	0.8335	0.8117
Fuzzy C-Means [18]	c=2; m=1.5	0.7039	0.7612	0.0962	0.7419	0.9201	0.8393	0.9050
Region Growing [19]	T=1.3	0.7315	0.7631	0.0819	0.7779	0.9237	0.8521	0.9241
Mean Shift [20]	S _R =23; R _R =13; w= 15	0.7563	0.8521	0.0758	0.7953	0.9548	0.9008	0.9406
Otsu's Thresholding [21]	T=50	0.3941	0.4359	0.1563	0.5558	0.6921	0.6752	0.8093
Active Contour Model [22]	$\alpha=1$; $\beta=1$; $\gamma=1.2$	0.8955	0.9259	0.0030	0.9228	0.9511	0.9454	0.9768
FODPSO [23]	L=3	0.7142	0.7398	0.1024	0.6053	0.7852	0.8563	0.8703
EM [24]	c=2	0.6902	0.7103	0.1078	0.5794	0.7639	0.8157	0.8117
KT [25]	T= 10	0.4832	0.5289	0.1503	0.5623	0.7428	0.7026	0.8095
SegNet [26]	Models are used for Testing	0.7001	0.7493	0.0800	0.7380	0.8983	0.8103	0.8967
U-Net [27]	Models are used for Testing	0.7385	0.8503	0.0679	0.7890	0.9401	0.8579	0.9376

Bold Face- Most Outer Performed Segmentation Method; JI- Jaccard Index; DSC- Dice Similarity Coefficient; MSE- Mean Square Error; Re- Recall; Pr- Precision; FM- F-Measure; ACC- Accuracy; c- Cluster Number; m- Degree of Fuzziness; T- Threshold; S_R- Spatial Resolution; R_R- Range Resolution; w- Size of the Smallest Segment; α - Elastic constant; β - Curvature Constant, γ - Image Energy Constant; L- Level

B. Overall Statistics

By maintaining above mentioned standard protocol suite, the pap-smear images are taken randomly from 20 female subjects including 10 normal and 10 abnormal subjects. The dataset comprises of total 100 single cell pap-smear images i.e., 5 images from each subject. The analysis of patient characteristics based on clinico-demographic information of the collected pap-smear images is demonstrated in TABLE I.

C. Ground Truth Annotation

Ground truth annotation of medical images is performed by the medical experts using manual segmentation for abnormality identification. However, manual segmentation may result in subjective biasness and uncertainty in the annotation. To alleviate with the difficult of subjective biasness, in our study, five sets of suspicious nucleus region are manual segmented by five research members using the help of well-known GNU Image Manipulation Program (GIMP) [15] annotation tool. The foreground pixels (i.e., nucleus region) are represented with intensity value as 1 and the background pixels (i.e., cytoplasm and remaining part) are represented with intensity value 0. The final ground truth image is generalized depending upon on the maximum voting policy scheme. This is generally performed using a threshold value, T. Thus pixel in the (i, j) location of resultant ground truth image is considered as foreground pixel if at least T (in our analysis we have considered the value of T as 3) number of research members included it as a foreground pixel. As a result, the final ground truth image is the result of the most comprehensive agreement among all technicians. The final ground truth images of suspicious nucleus region of each pap-smear image obtained using maximum voting policy scheme are provided in .jpg format with the dataset. The sample images of proposed dataset are displayed in Fig. 1.

III. METHOD: ANALYSIS OF AGMC-TU PAP SMEAR CYTOLOGICAL IMAGE DATASET

In this section, analysis of our AGMC-TU Pap Smear Cytological Image Dataset is performed in two major groups: normal and abnormal. The detailed description of the analysis is provided below.

A. Pre-Processing of Cytological Pap-Smear Images

For abnormality detection of cervico-vaginal pap-smear images, sharp boundaries between nucleus and cytoplasm are required to be defined. The microscopic images are usually poor quality color images which are stained to color the cells and are thus affected by noise. To adjust the contrast between cell nucleus and cytoplasm for subsequent processing, the original pap-smear images are converted into gray scale images and then Contrast-Limited Adaptive Histogram Equalization (CLAHE) method [16] is used for further enhancement.

TABLE III
OVER AND UNDER SEGMENTATION MEASUREMENTS FOR ACCURATE EXTRACTION OF NUCLEUS REGION

Segmentation Methods	Oseg	Useg
Active Contour Model [22]	0.0510	0.5632
Mean Shift [20]	0.1922	0.5890
U-Net [27]	0.2001	0.5883

Oseg- Over Segmentation; Useg- Under Segmentation

B. Suspicious Nucleus Region (Region of Interest) Segmentation

For shape features extraction from the proposed dataset, effective nucleus segmentation from the holistic pap-smear images is an important task. In our study, we have done comparison of some segmentation methods so as to identify the most efficient method for segmentation of nucleus regions from the holistic pap-smear images. The nine well-known segmentation methods used in our study are: K-Means clustering method (KM) [17]; Fuzzy C-Means clustering method (FCM) [18]; Region Growing method (RG) [19]; Mean Shift method (MS) [20]; Otsu's Thresholding method (OT) [21]; Active Contour method (ACM) [22], Fractional-Order Darwinian Particle Swarm Optimization method (FODPSO) [23]; Expectation Maximization method (EM) [24] and Kapur Thresholding method (KT) [25], SegNet [26], and U-Net [27].

TABLE II reports the experiment results of the above mentioned segmentation methods on nucleus regions segmentation from the pap-smear images. The performance of the aforementioned segmentation methods is measured in terms of seven pixel level measurement metrics with respect to the ground truth images. These metrics are: Recall (Re), Precision (Pr), Jaccard Index (JI), F-Measure (FM), Dice Similarity Coefficient (DSC), and Accuracy (ACC). The associated parameters against each of the segmentation methods are also listed in Table 2. The parameters fixation of the clustering based segmentation methods (i.e., K-Means [17], Fuzzy C-Means [18], Mean Shift methods [20] and Expectation Maximization method (EM) [24]) are selected against the value(s) for which the I-index is maximum [42]. Conversely for the remaining six segmentation methods, the parameters are empirically set based on trial and error method. From the experimental result as shown in TABLE II, it can be observed that although Mean shift segmentation method [20] obtained noticeable result as per precision value but in comparison to all the similarity metrics it can be concluded that active contour model [22] with parameters Elastic constant (α) = 1; Curvature Constant (β) = 1 and Energy Constant (γ) = 1.2 have more similarity and less error difference between the ground truths and machine segmented outputs. Also, we have noticed that active contour model [22], Mean Shift [20], and U-Net [27] are the three outer performed methods. Furthermore for effective segmentation of the nucleus regions, we have further compared these three outer performed methods so as indicated above in terms of over segmentation and under segmentation measurement [43]. TABLE III reports the over segmentation

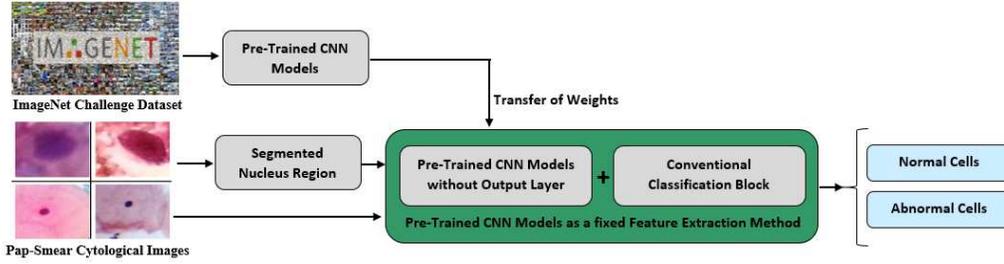


Fig. 2 Framework for the State-of-the-art CNN Architecture as a Fixed feature Extraction Method and Classification

(Oseg) and under segmentation (Useg) measurement of the three outer performed segmentation techniques. From TABLE III, it can be observed that active contour model [22] outperforms the remaining two segmentation methods in terms of lower values of Oseg and Useg metrics. It, can be considered as the outperformed method for accurate extraction of nucleus region.

C. Feature Extraction and Asymmetric Analysis

In our proposed analysis, we have extracted two types of features i.e., conventional shape features and deep based features. The details of these two types of features extracted from the pap-smear images are provided below.

Conventional Shape Feature Extraction from Segmented Nucleus Region. In our work, after segmenting the regions of interest (i.e., nucleus regions) from the holistic pap-smear images using greedy active contour model (ACM) [22], eleven features based on the nucleus shapes are extracted from the segmented images. These features with Feature IDs are: F1: Nucleus Area; F2: Nucleus Perimeter; F3: Equivalent Diameter; F4: Major Axis Length; F5: Minor Axis Length; F6: Nucleus Roundness; F7: Convex Area; F8: Elongation; F9: Eccentricity; F10: Solidity; F11: Extent. These features are together named as combination of all shape feature set (FSCOM) and is represented as:

$$FSCOM = \{F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9, F_{10}, F_{11}\} \quad (1)$$

The mean and standard deviation of the aforementioned extracted feature set (FSCOM) for normal (50 samples) and abnormal (50 samples) cell images are listed in TABLE IV. After extraction of these 11 shape feature set (FSCOM), statistical significant test based asymmetric analysis has been done between two categories of cells (i.e., normal and abnormal cells) for discriminating features selection that can characterize the difference between these cells. To figure out the most relevant discriminative features, Mann-Whitney-Wilcoxon (MWW) test [28] has been performed between these two categories of cells with significance levels of 0.1%. In TABLE IV, the significance of each considered shape feature values are approximated in terms of the null hypothesis. Depending upon the phenomenon that the nucleus regions are comparatively higher with respect to the normal cells, the null hypothesis is defined as the shape features of the abnormal cell images are approximately larger than the shape features of the normal cell images. The shape features whose p-value < 0.001 are considered to be significant statistically and the features, p-value > 0.001 are considered to be irrelevant. Among these eleven extracted shape features, the null hypothesis is accepted by the six shape based features and are together named as statistically significant feature set (FSSSIG). These features are represented as:

$$FSSSIG = \{F_1, F_2, F_3, F_4, F_5, F_7\} \quad (2)$$

TABLE IV
ASYMMETRIC ANALYSIS OF PAP-SMEAR CELLS BASED ON SHAPE FEATURES EXTRACTED FROM THE REGION OF INTEREST (NUCLEUS REGION) FOR STATISTICAL SIGNIFICANT FEATURE SELECTION

Shape Features	Asymmetric Analysis		Significant Test	
	Normal cell (Mean ± SD)	Abnormal cell (Mean ± SD)	p value	p < 0.001
F ₁ (NA)	456.94±250.10	4029.78±1410.06	2.3480e-06	SSig
F ₂ (NP)	243.32±125.46	404.24± 228.73	1.3491e-07	SSig
F ₃ (ED)	99.80±26.57	304.52 ± 83.42	2.4573-05	SSig
F ₄ (MAJ)	21.89 ± 5.99	69.18 ± 19.37	1.7568-03	SSig
F ₅ (MIN)	26.28 ± 7.35	84.53 ± 24.69	3.2311e-02	SSig
F ₆ (NR)	0.11 ± 0.39	0.02 ± 0.01	0.3410	SISig
F ₇ (CA)	18.87 ± 5.37	58.52 ± 17.71	8.8345e-03	SSig
F ₈ (EL)	0.59 ± 0.21	0.7262 ± 0.13	0.3001	SISig
F ₉ (EC)	0.76 ± 0.17	0.66 ± 0.16	0.3478	SISig
F ₁₀ (SO)	0.03 ± 0.02	0.01 ± 0.02	0.9352	SISig
F ₁₁ (EX)	0.02± 0.01	0.01 ± 0.02	0.5673	SISig

NA- Nucleus Area; NP- Nucleus Perimeter; ED- Equivalent Diameter; MAJ- Major Axis Length; MIN- Minor Axis Length; CA- Convex Area; NR- Nucleus Roundness; EL- Elongation; EC- Eccentricity; So- Solidity; EX- Extent; SSig- Statistically Significant Feature; SISig- Statistically Insignificant Feature

On the otherhand, for selection of differentiating features automatically, relief based feature selection method [29] has been used. This method usually computes the score against each extracted shape feature and consider the top most scored features as the differentiating features. It has been noticed from effective study that considering the value 5 (five) as the rank value beginning from 1 (one) observed to be generating maximum accuracy for cell classification. In our analysis, we have considered these 5 features as automatically significant feature set (FSRBSIG). These features are represented as:

$$FSRBSIG = \{F_1, F_2, F_3, F_4, F_5\} \quad (3)$$

Thus we have extracted three conventional feature sets from the segmented nucleus region i.e., FSCOM, FSSSIG and FSRBSIG.

Deep Feature Extraction from Holistic and Segmented Pap-Smear Images. From literature, we have adopted seven state-of-the-art CNN models pre-trained on ImageNet dataset [30] and used them as a fixed feature extraction method where the fully connected layer from the architecture is eliminated while retaining the remaining portion of the network. These CNN models are: VGG16 [31], VGG 19 [31], AlexNet [32], GoogleNet [33], Inception-V3 [34] and ResNet101 [35]. In the present scope, we have used these aforementioned CNN models for extraction of features from both the holistic pap-smear cytological images (DEEP_{HFS}) and segmented pap-smear cytological images (DEEP_{SFS}) in RGB palette for understanding the perception capability of deep CNN models for effective classification of abnormal pap-smear images from the normal cell images. Fig. 2 sketches the overall diagrammatic representation of the adoption of the CNN-architectures as a fixed feature extraction method in our proposed work. From analysis, we have observed that VGG16 [29] provides an improvement (i.e., in terms of classification accuracy) on our dataset over the remaining six pre-trained

TABLE V
CLASSIFICATION ACCURACY OF CELL IMAGES (NORMAL AND ABNORMAL) BASED ON EXTRACTED CONVENTIONAL FEATURE SET

Classifier	Conventional Shape Feature Set								
	FSCOM			FSSSIG			FSRBSIG		
	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity
SVM [36]	89.97%	89.10%	88.14%	96.37%	95.70%	96.32%	95.88%	96.75%	95.93%
RF [37]	84.52%	84.24%	84.31%	93.56%	95.32%	94.83%	91.00%	95.92%	90.20%
DT [38]	82.76%	82.61%	82.32%	89.64%	84.31%	83.67%	88.65%	83.96%	81.45%
KNN [39]	81.36%	81.84%	81.41%	95.16%	92.45%	94.86%	92.21%	91.84%	92.16%
NB [40]	85.54%	84.44%	85.26%	96.24%	95.37%	93.75%	94.16%	91.76%	92.16%
ANN [41]	87.25%	87.47%	87.00%	96.45%	95.70%	96.36%	95.64%	95.12%	96.14%

Boldface and Underlined: Most Outer Performed Conventional Shape Feature Set

TABLE VI
CLASSIFICATION ACCURACY OF CELL IMAGES (PAP-SMEAR) BASED ON EXTRACTED DEEP FEATURE SET FROM HOLISTIC IMAGES (DEEP_{HFS}) AND SEGMENTED NUCLEUS REGION (DEEP_{SFS})

Classifier	Deep Extracted Shape Feature Set					
	DEEP _{HFS}			DEEP _{SFS}		
	Accuracy	Specificity	Sensitivity	Accuracy	Specificity	Sensitivity
SVM [36]	98.22%	98.39%	98.05%	93.76%	98.00%	90.32%
RF [37]	93.33%	90.00%	96.18%	84.56%	94.12%	84.22%
DT [38]	89.12%	88.56%	86.49%	87.56%	83.32%	90.41%
KNN [39]	96.11%	96.87%	96.00%	92.49%	96.21%	86.54%
NB [40]	97.32%	98.43%	96.54%	91.02%	95.52%	86.10%
ANN [41]	96.73%	95.31%	97.27%	93.21%	96.43%	93.00%

Boldface and Underlined- Most Outer Performed Deep Feature Set

CNN models. Thus, in our present study, we have reported the performance of VGG16 [31] pre-trained model as a feature extraction module for classification of two types of pap-smear images (i.e., normal and abnormal).

D. Classification of Normal and Abnormal Cells

Automatic normal and abnormal cells classification are essential for automatic prediction of any disease abnormality. In this section, we have used the mentioned feature sets for cells classification: FSCOM, FSSSIG, FSRBSIG, DEEP_{HFS}, and DEEP_{SFS}.

In our experiment, we have used 50 normal and 50 abnormal pap-smear images from our proposed dataset. Each of the aforementioned feature sets were independently fed to six well-known classifier models i.e., Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), K-Nearest Neighbour (KNN) algorithm, Naïve Bayes (NB) and Artificial Neural Network (ANN) [36]-[41]. In this study, we used a linear kernel to evaluate the performance of an SVM classifier. For the KNN classifier, the best number of neighbours $K=2$ is used to determine classification accuracy. The LDA and NB classifiers, on the other hand, do not require any parameter adjustment. Furthermore, the maximum number of splits for the DT classifier is set to 4, while the maximum number of decision trees for the RF classifier is set to 20. Moreover, for ANN, prediction accuracy is performed for 10 hidden layers. The accuracy, sensitivity and specificity of the aforementioned feature sets for the best suited parameters as mentioned above are shown in TABLE V and VI. In comparison to three conventional feature set (i.e. FSCOM, FSSSIG and FSRBSIG), different classifiers have shown their best performances with Statistically Significant Feature Set (FSSSIG) with an average accuracy, specificity and sensitivity of 94.57%, 93.14%, and 93.30% respectively for all the used classifiers. Moreover Automatically Selected Significant Feature Set (FSRBSIG) have also shown noticeable classification performance as compared to combined Shape feature Set (FSCOM) with an average accuracy, specificity and sensitivity of 92.93%, 92.56% and 91.34% respectively. Also it can be observed that among these five feature sets, deep learning based feature extraction (i.e., DEEP_{HFS} and DEEP_{SFS}) shows noticeable classification performance as compared to conventional feature set (i.e., FSCOM, FSSSIG, and FSRBSIG). Conversely, in comparison to deep based feature sets, features based on holistic cytological images (DEEP_{HFS}) shows best classification performance with

respect to all the classifiers with an average accuracy, specificity and sensitivity of 95.13%, 94.59% and 95.09% respectively.

IV. CONCLUSION

Cervical cancer is considered to be the pervasive forms of cancer mostly suffered by the female population worldwide. In this paper, role of feature extraction and selection from pap-smear images in improving the accuracy for pap-smear cell image classification has been highlighted. The analysis has been carried out through creation of a new pap-smear image dataset from female populations of low resource cervical cancer prone regions. Experimental results reveal that in comparison to all the feature sets, deep features extracted from both holistic pap-smear images (Deep_{HFS}) have shown noticeable results in comparison to the other considered four sets of features and are considered to be efficiently classify the pap-smear cells in two major categories (i.e., normal and abnormal).

ACKNOWLEDGMENT

The work presented here is being conducted in the Bio-Medical Infrared Image Processing Laboratory of Computer Science and Engineering Department of Tripura University (A Central University), Tripura, Suryamaninagar-799022. The work is supported by Short Term ICMR-DHR International Fellowship Programme (Grant No. INDO/FRC/452/S-69/2019-20-IHD, Dated: 09/12/2019).

REFERENCES

- [1] National Cancer Registry Programme, Three year report of population based cancer registries 2012-2014. [Online]. Available: http://www.icmr.nic.in/ncrp/pber_2012-14/ALL_CONTENT/PDF_Printed_Version/Preliminary_Pages_Printed.pdf
- [2] WHO/ICO. (2013). Information Centre on HPV and Cervical Cancer (HPV Information Centre), Human Papillomavirus and Related Diseases. Report in China. [Online]. Available: www.who.int/hpcenter.
- [3] A report on cancer burden in north eastern states of india [Online]. Available: http://www.ncdirindia.org/All_Reports/Reports_Ne/NE2012_2014/Files/NE_2012_14.pdf
- [4] S.C. Shekhar, S.T. Sharma, N. Raina, "Cervical cancer screening: knowledge, attitude and practices among nursing staff in a tertiary level teaching institution of rural india," Asian Pacific Journal of Cancer Prevention, Vol. 14. No. 6, pp. 3641-3645, 2013.
- [5] Cervical pre-invasive-diagnosis and treatment [Online]. Available: <https://www.cancertherapyadvisor.com/home/decision-support-in->

- medicine/obstetrics-and-gynecology/cervical-pre-invasive-diagnosis-and-treatment/
- [6] L. Zhang, L. Lu, I. Noguez, R.M. Summers, S. Liu and J. Yao, "DeepPap: deep convolutional networks for cervical cell classification," *IEEE Journal of Biomedical and Health Informatics*, Vol. 21, No. 6, pp. 1633-1643, 2017.
 - [7] M. Rohmatillah, S.H. Pramono, H. Suyono and S.A. Sena, "Automatic Cervical Cell Classification Using Features Extracted by Convolutional Neural Network," In 2018 Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS), IEEE, pp. 382-386, 2018.
 - [8] K. Teeyapan, N.T. Umpon and S. Auephanwiriyaikul, "Application of support vector based methods for cervical cancer cell classification," In 2015 IEEE international conference on control system, computing and engineering (ICCSCE), IEEE, pp. 514-519, 2015.
 - [9] Y. Song, T.E. Leng, X. Jiang, J.Z. Cheng, D. Ni, S. Chen, B. Lei and T. Wang, "Accurate cervical cell segmentation from overlapping clumps in pap smear images," *IEEE Transactions on Medical Imaging*, Vol. 36, No. 1, pp. 288-300, 2016.
 - [10] A. Khadidos, V. Sanchez and C.T. Li, "Patch-based segmentation of overlapping cervical cells using active contour with local edge information," In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, pp. 1058-1062, 2017.
 - [11] Y.F. Chen, P.C. Huang, K.C. Lin, H.H. Lin, L.E. Wang, C.C. Cheng, T.P. Chen, Y.K. Chan and J.Y. Chiang, "Semi-automatic segmentation and classification of pap smear cells," *IEEE Journal of Biomedical and Health Informatics*, Vol. 18, No. 1, pp. 94-108, 2013.
 - [12] Z. Lu, G. Carneiro, A.P. Bradley, D. Ushizima, M.S. Nosrati, A.G. Bianchi, C.M. Carneiro and G. Hamarneh, "Evaluation of three algorithms for the segmentation of overlapping cervical cells," *IEEE Journal of Biomedical and Health Informatics*, Vol. 21, No. 2, pp. 441-450, 2016.
 - [13] M.E. Plissiti and C. Nikou, "A review of automated techniques for cervical cell image analysis and classification," In *Biomedical imaging and computational modeling in biomechanics*, Springer, pp. 1-18, 2013.
 - [14] AGMC-TU Pap Smear Cytological Image Dataset. [Online]. Available: <http://mkbhowmik.in/AGMC-tuDb.aspx>
 - [15] E. Oliver, J. Ruiz, S. She, and J. Wang, "The Software Architecture of the GIMP," 2006.
 - [16] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics gems*, pp. 474-485, 1994.
 - [17] U. Snehalatha, M. Anburajan, V. Sowmiya, B. Venkatraman, M. Menaka, "Automated hand thermal image segmentation and feature extraction in the evaluation of rheumatoid arthritis," in *Proc. Institution of Mechanical Engineers Part H: Journal of Engineering in Medicine*, Vol. 229, No. 4, pp. 319-331, 2015.
 - [18] S. Araki, H. Nomura, N. Wakami, "Segmentation of thermal images using the fuzzy c-means algorithm," in *Proc. IEEE Second International Conference on Fuzzy Systems*, pp. 719-724, 1993.
 - [19] W. K. Pratt, *Digital Image Processing 4th Edition*, John Wiley & Sons, Inc., Los Altos, California, 2007.
 - [20] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 603-619, 2002.
 - [21] W. Wang, L. Duan and Y. Wang, "Fast image segmentation using two-dimensional Otsu based on estimation of distribution algorithm," *Journal of Electrical and Computer Engineering*, 2017.
 - [22] D.J. Williams, and M. Shah, "A fast algorithm for active contours and curvature estimation," *CVGIP: Image understanding*, Elsevier, Vol.55, No.1, pp. 14-26, Jan. 1992.
 - [23] M.S. Couceiro, N.M. Ferreira, and J.A. Tenreiro Machado, "Fractional order Darwinian particle swarm optimization," in *Proc. Symposium on fractional signals and systems*, pp. 127-136, 2011.
 - [24] T.J. Ramirez-Rozo, J.C. Garcia-Alvarez, and C.G. Castellanos-Dominguez, "Infrared thermal image segmentation using expectation-maximization-based clustering," in *Proc. 2012 XVII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA)*, pp. 223-226, 2012.
 - [25] J.N. Kapur P.K. Sahoo, and A.K.C. Wong "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer Vision, Graphics, and Image Processing*, Elsevier, Vol. 29, No. 3, pp. 273-85, 1985.
 - [26] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481-2495, 2017.
 - [27] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. International Conference on Medical image computing and computer-assisted intervention*, Springer, Cham, pp. 234-241, 2015.
 - [28] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics*, vol. 1, no. 6, pp. 80-83, 1945.
 - [29] R.J. Urbanowicz, M. Meeker, W.L. Cava, R.S. Olson and J.H. Moore. "Relief-based feature selection: introduction and review," *Journal of Biomedical Informatics*, Elsevier, Vol. 85, pp. 189-203, 2018.
 - [30] J. Deng, W. Dong, R. Socher, L. Li, K. Li and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 248-255, 2009.
 - [31] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein and A.C. Berg, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, Springer, vol. 15, no. 3, pp. 211-252, 2015.
 - [32] A. Krizhevsky, I. Sutskever and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," In *Advances in Neural Information Processing Systems*, pp. 1097-1105, 2012.
 - [33] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 1-9, 2015.
 - [34] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the inception architecture for computer vision," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 2818-2826, 2016.
 - [35] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 770-778, 2016.
 - [36] V. N. Vapnik, "The nature of statistical learning theory," Springer- New York, NY, USA, 1995.
 - [37] L. Breiman, "Random Forests", *Journal of Machine Learning*, Springer, Vol. 45, No. 1, pp. 5-32, 2001.
 - [38] K. Cernak, "A comparison of decision tree classifiers for automatic diagnosis of speech recognition errors," *Computing and Informatics* vol. 29, No. 3, pp. 489-501, 2012.
 - [39] Thirumuruganathan, S. (2010) Resource Document: A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm. [Online]. Available: <https://saravananthirumuruganathan.wordpress.com/2010/05/17/adetailed-introduction-to-k-nearest-neighbor-knn-algorithm/>.
 - [40] Y. Huang and L. Li, "Naive Bayes classification algorithm based on small sample set", *IEEE International Conference on Cloud Computing and Intelligence Systems (CCIS)*, IEEE, pp. 34-39, 2011.
 - [41] F. Amato, A. López, E. M. Peña-Méndez, P. Vanhara, A. Hampl, & J. Havel, "Artificial neural networks in medical diagnosis," *Journal of applied biomedicine*, vol. 11, no. 2, pp. 47-58, 2013.
 - [42] U. Maulik and S. Bandyopadhyay, "Performance evaluation of some clustering algorithms and validity indices," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 12, pp. 1650-1654, 2002.
 - [43] M. Belgiu, and L. Dragut, "Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 96, pp. 67-75, 2014.