Optimizing Cervical Cancer Classification with SVM and Improved Genetic Algorithm on Pap Smear Images

S. Umamaheswari, Y. Birnica, J. Boobalan, V. S. Akshaya

Abstract

This study presents an approach to optimize cervical cancer classification using Support Vector Machines (SVM) and an improved Genetic Algorithm (GA) on Pap smear images. The proposed methodology involves preprocessing the images, extracting relevant features, and employing a genetic algorithm for feature selection. An SVM classifier is trained using the selected features and optimized using the genetic algorithm. The performance of the optimized model is evaluated, demonstrating improved accuracy and efficiency in cervical cancer classification. The findings hold the potential for assisting healthcare professionals in early cervical cancer diagnosis based on Pap smear images.

Keywords: SVM, Pap smear images, Cervical cancer, GA, Healthcare.

MSC 2020: 62H35.

1 Introduction

Cervical cancer is a serious health problem that affects women all over the world and has a significant impact on morbidity and mortality rates [1]. Early detection through screening plays a serious role in reducing the burden of this disease. However, the interpretation of Pap smear images, a commonly used screening tool, can be subjective and lead to varying levels of diagnostic accuracy. In recent years, there has been an increasing interest in using sophisticated computational methods like

©2024 by Computer Science Journal of Moldova
doi:10.56415/csjm.v32.05

support vector machines and genetic algorithms to optimize cervical cancer classification based on Pap smear images. This study aims to address the challenges associated with cervical cancer classification by proposing an approach that combines SVM and an improved genetic algorithm. By integrating these techniques, the goal is to develop a robust and efficient model capable of accurately analyzing Pap smear images and providing reliable classification outcomes. The optimization process involves preprocessing the images, extracting relevant features, selecting informative features using a genetic algorithm, and training an SVM classifier with optimized parameters. The resulting model holds the potential to enhance the accuracy and efficiency of cervical cancer diagnosis, assisting healthcare professionals in making timely and accurate decisions for improved patient care.

2 Methodology

The proposed methodology for optimizing cervical cancer classification with SVM and Improved Genetic Algorithm on Pap smear images involves several steps:

1. *Pre-processing:* The Pap smear images are first pre-processed to enhance the contrast and eliminate noise. This involves techniques such as histogram equalization and Gaussian filtering.
2. *Segmentation:* The pre-processed images are then segmented to extract the region of interest (ROI), which contains the cervical cells. This involves using a super pixel-based Markov random field segmentation algorithm.
3. *Feature extraction:* Various textures and patterns are extracted from the segmented ROI. These include Gabor filters, Haralick features, and Zernike moments.
4. *Feature selection:* The most appropriate features that improve classification accuracy are determined using an improved genetic algorithm.
5. *Classification:* The selected features are used to train an SVM classifier. The SVM classifier is optimized using the Improved Genetic Algorithm to find the best hyperparameters that maximize the classification accuracy.
6. *Performance evaluation:* Parameters including accuracy, precision,

recall, and F1-Measure are used to assess the proposed system's performance [2]. The results are compared with existing benchmark methods to determine the effectiveness of the proposed approach.

3 Proposed System

3.1 Image pre-processing

To optimize Pap smear images for cervical cancer analysis, image preprocessing [3] is essential. To improve image quality, decrease noise, and extract important information, many techniques are used. Among these methods are scaling the images to a uniform dimension, boosting contrast to improve visibility, lowering noise through smoothing or filtering, normalizing intensity or color values, extracting the region of interest (such as the cervix), segmenting various structures within the image, and registering multiple images for precise comparisons. The quality and appropriateness of Pap smear pictures are enhanced through the use of various pre-processing techniques, enabling more precise and trustworthy analysis and classification of cervical cancer.

Image pre-processing methods encompass a range of techniques used to enhance and optimize images before further analysis or classification. In the context of medical imaging, including Pap smear images for cervical cancer analysis, the following methods are commonly employed: 1. Image resizing, 2. Contrast enhancement, 3. Noise reduction, 4. Image normalization, 5. Edge detection, 6. Image segmentation, 7. Morphological operations. This article mainly focuses on Image normalization and Image segmentation techniques.

3.1.1 Image Normalization

Image normalization is a pre-processing technique used to standardize the intensity or color values of an image. It ensures that images from different sources or under different conditions can be compared and analyzed accurately. By rescaling pixel values to a consistent range or distribution, normalization improves the reliability and comparability of image features, reducing the impact of variations caused by factors

such as lighting or imaging settings. It enhances the quality and consistency of images, making them more suitable for subsequent analysis or classification tasks.

3.2 Image Segmentation

The Region Of Interest (ROI) containing the cervical cells is extracted from the segmented Pap smear images that have undergone pre-processing. Super pixel-based Markov Random Field (MRF) [4] segmentation technique is used to complete this segmentation process.

3.2.1 Super pixel generation

Super pixels are compact and perceptually meaningful image regions that group pixels with similar characteristics together. They provide a higher level of abstraction compared to individual pixels and can facilitate more efficient and accurate segmentation. Super pixels are generated by dividing the image into compact, regular regions while preserving image boundaries.

3.2.2 Markov Random Field (MRF) Modelling

MRF is a probabilistic graphical model broadly used in image segmentation. It models the relationship between neighboring pixels by incorporating contextual information to improve the accuracy.
Let's denote the pre-processed image as I and the corresponding super pixels as S. Each super pixel S consists of a set of pixels $S = p_1, p_2, ..., p_n$. The goal is to assign a label (foreground or background) to each superpixel. The MRF formulation is as follows,

$$E(S) = \sum \left(D(S) + V(S) \right), \tag{1}$$

where $E(S)$ is the energy function, D(S) represents the data term, and $V(S)$ denotes the regularization term.

3.2.3 Data Term

The data term measures the compatibility between each super pixel and the estimated class labels. It is based on the colour and texture features of the super pixels. The data term is defined as follows,

$$D(S) = \sum (D_{data(S)}), \tag{2}$$

where $D_{data(S)}$ represents the data cost for each super pixel S.

3.2.4 Regularization Term

The regularization term encourages spatial coherence and smoothness in the segmentation results. It penalizes sharp transitions between neighboring super pixels. The regularization term is defined as follows,

$$V(S) = \sum (V_{reg(S)}), \tag{3}$$

where $V_{reg(S)}$ represents the regularization cost for each super pixel S.

3.2.5 Optimization

The goal is to find the optimal labeling configuration S* that minimizes the energy function $E(S)$. This optimization problem is typically solved using optimization techniques such as graph cuts or belief propagation.

The pre-processed pictures are segmented into areas that correspond to the cervical cells, which form the ROI using the super pixel-based MRF. This segmentation step is crucial for isolating the cells and providing accurate input for the subsequent stages of the classification process.

3.3 Feature Extraction

Feature extraction is the process that takes the segmented ROI from the previous module and extracts texture and shape features from it. These qualities offer crucial details regarding the traits and patterns visible in cervical cell images which can be utilized for classification. Gabor filters, Haralick features, and Zernike moments are used to extract texture and shape features. Gabor filters are spatial frequency

filters that capture texture information at different scales and orientations, while Haralick features describe the statistical properties of pixel intensities in an image. Zernike moments capture the shape characteristics of the segmented ROI, including its boundary and internal structure.

The extraction of these features is performed using AlexNET's pretrained CNN model. The extracted features form a feature vector representing the unique characteristics of the segmented ROI, which serves as input data for the subsequent classification module. Machine learning algorithms or classifiers can be trained to distinguish between different types of cervical cells and classify them accordingly.

The feature extraction module plays a vital role in capturing relevant information from segmented ROIs and transforming it into a suitable format for classification. By utilizing Gabor filters, Haralick features, and Zernike moments, the system effectively captures both texture and shape characteristics, enabling accurate classification and diagnosis of cervical cancer cells.

3.4 Feature Selection

The fourth module of the proposed cervical cancer classification project is featuring selection. In this module, an Improved Genetic Algorithm (IGA) is employed to identify and select the most related features from the feature vector obtained in the previous module. The goal of feature selection is to keep just the most useful features that considerably improve classification accuracy while reducing the dimensionality of the feature space.

Feature selection using an Improved Genetic Algorithm involves the following steps:

1. Initial Population: A population of potential feature subsets is randomly generated. Everyone in the population represents a candidate feature subset, where each feature is represented by a binary value (0 or 1) indicating its inclusion or exclusion in the subset.

2. Fitness Evaluation: The fitness of everyone in the population is obtained using a Fitness Function (FF). The FF assesses the classification accuracy achieved by using the corresponding feature subset. This

evaluation is typically done by training and testing a classifier (SVM) on the selected features and measuring its performance using metrics like accuracy, precision, recall, or F1-Measure.

$$fitness(individual) = performance\ (classifier,\ selected\ features) \tag{4}$$

3. Selection: A selection process is applied to select individuals with higher fitness values. The individuals with better performance (higher classification accuracy) have a higher probability of being selected for reproduction. Common selection techniques include tournament selection, roulette wheel selection, or rank-based selection [5]. In this research, Roulette wheel selection is used because it gives better results compared with other methods.

4. Crossover: Crossover is performed by selecting pairs of individuals from the selected population and combining their feature subsets to create new offspring. This process simulates the genetic recombination that occurs in natural evolution. Different crossover techniques can be used, such as 1-point crossover, 2-point crossover, or uniform crossover.

5. Mutation: Mutation introduces random changes in the feature subsets of the offspring. This facilitates the exploration of new search space areas while preventing an early convergence. Mutation randomly flips the binary values (0 to 1 or 1 to 0) of certain features in the offspring.

6. Fitness Evaluation (Offspring): The fitness of the newly created offspring is evaluated using the fitness function, similar to the initial population. This step determines the classification accuracy achieved by the offspring using their modified feature subsets.

7. Replacement: The offspring are selected to replace individuals in the population based on their fitness values. This process ensures the survival of individuals with higher fitness and discards those with lower fitness, maintaining the population size.

8. Termination: The Improved Genetic Algorithm iteratively performs the steps of selection, crossover, mutation, and fitness evaluation. This can be a fixed number of generations, reaching a specific fitness threshold, or convergence of the algorithm.

3.5 Classification Using SVM

The fifth module of the cervical cancer classification project uses features from previous modules to train the SVM classifier. The classifier is optimized using the Improved Genetic Algorithm to find optimal hyperparameters for maximum accuracy. SVM is a widely used supervised learning algorithm in medical image analysis. There are six steps involved in classification.

1. Data Preparation: The feature vector, consisting of the selected features obtained from the feature selection module, serves as the input data for the SVM classifier. The feature vector is typically represented as a matrix, where each row represents an instance or sample, and each column represents a feature.

2. SVM Training: Using the labelled training data, the SVM classifier is trained. The feature vectors with corresponding class labels (such as "positive" or "negative" for cervical cancer) make up the training data. Finding the best hyperplane to maximally separate the positive and negative instances in the feature space is the goal of the SVM method.

3. Hyper parameter Optimization: IGA is involved in optimizing the hyper parameters of the SVM classifier. The hyper parameters are settings that control the behavior and performance of the SVM algorithm. Examples of SVM hyper parameters include the kernel type and regulation parameter (C). The Improved Genetic Algorithm investigates several hyperparameter combinations to determine which ones produce the best classification accuracy.

4. Fitness Evaluation: The fitness function in the Improved Genetic Algorithm evaluates the classification accuracy achieved by the SVM classifier using the selected hyper parameters. The classification accuracy is typically measured using metrics such as accuracy, precision, recall, or F1-Measure [6]. The Improved Genetic Algorithm chooses individuals for reproduction in the next generation who have higher fitness values.

5. Crossover and Mutation: Crossover and mutation procedures are carried out using the Improved Genetic Algorithm to produce new candidate solutions in the population. Crossover involves integrating

the hyper parameters of chosen individuals to produce a new generation with various combinations of hyper parameters. For exploring new areas of the search space and preventing premature convergence, mutation introduces random modifications to the hyperparameters.

6. Replacement and Termination: The offspring with their modified hyper parameters replace individuals in the population based on their fitness values. This process ensures the survival of individuals with better classification accuracy. The Improved Genetic Algorithm iteratively performs the steps of fitness evaluation, crossover, mutation, replacement, and termination.

3.6 Performance Evaluation

Performance evaluation is a crucial step in assessing the effectiveness of the proposed cervical cancer classification system. It involves the use of various metrics to quantitatively measure the system's performance and compare it with existing benchmark methods.

1. Accuracy: Accuracy measures the overall correctness of the classification system and is defined as the ratio of correctly classified instances to the total number of instances [6]:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{5}$$

2. Precision: Precision is the ability of the system that measures how well the system can pick out positive examples from the sum of instances that are projected to be positive. It is calculated as [6]:

$$Precision = \frac{TP}{TP + FP}. \tag{6}$$

3. Recall: The capacity of the system to accurately identify positive cases out of all real positive instances is measured by recall, also known as sensitivity or true positive rate. It is defined as [6]:

$$Recall = \frac{TP}{TP + FN}. \tag{7}$$

4. F1-Measure: The harmonic mean of precision and recall is the F1-Measure, which offers a single measure to balance them both. It is

calculated as [6]:

$$F1 - Measure = 2 * \frac{Precision * Recall}{Precision + Recall}. \quad (8)$$

By considering both precision and recall, the F1-Measure provides a comprehensive evaluation of the system's performance, especially when dealing with imbalanced datasets.

The performance assessment is normally carried out using an appropriate dataset that contains labeled samples of various cervical cell types. The dataset is divided into training and testing sets, with the former being used to train the classification model and the latter to assess the model's efficacy.

The proposed system's performance is compared with recent techniques by applying the same evaluation metrics to their results. The comparison helps determine whether the proposed approach achieves superior or comparable performance compared to other methods.

4 Dataset

Dataset Name: Cervical Cancer Pap Smear Images Dataset. There are four classes available in this data set, and each class has its own image samples as listed below [6].

1. Carcinoma In Situ (200 samples)
2. Light Dysplastic (300 samples)
3. Moderate Dysplastic (250 samples)
4. Severe Dysplastic (350 samples)

Carcinoma In Situ

Number of Samples: 200

Description: Carcinoma in situ shown in Figure 1 states the abnormal cells that are found only in the innermost lining of the cervix. An aberrant cell cluster known as a carcinoma in situ is one that has not yet spread from the site where it initially developed, yet it has the potential to do so in the future to transform into cancer. Carcinoma in situ is considered a pre-cancerous condition and is an early stage of cervical cancer.

Figure 1. Carcinoma In Situ

Light Dysplastic

Number of Samples: 300

Light dysplastic in Figure 2 states the presence of mildly anomalous cells in the cervical tissue. These cells show some changes but are not considered cancerous. Light dysplastic cells may indicate the primary phases of cervical cancer development and require further monitoring and treatment.

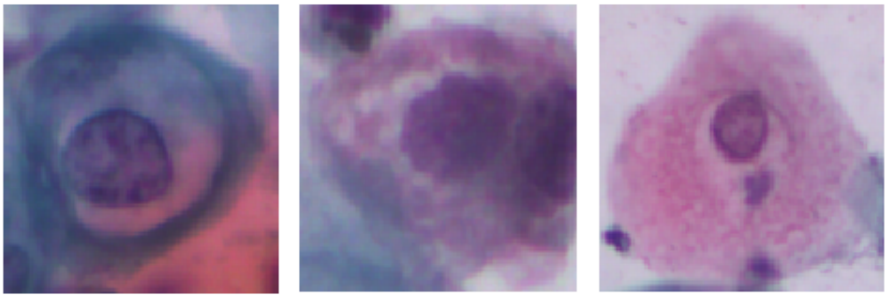

Figure 2. Light Dysplastic

Moderate Dysplastic

Number of Samples: 250

Moderate dysplastic shown in Figure 3 describes the presence of moderately anomalous cells in the cervical tissue. These cells exhibit more pronounced changes than light dysplastic cells but are still not classified as cancerous. Moderate dysplastic cells indicate a higher risk of

cervical cancer and may require closer monitoring and treatment.

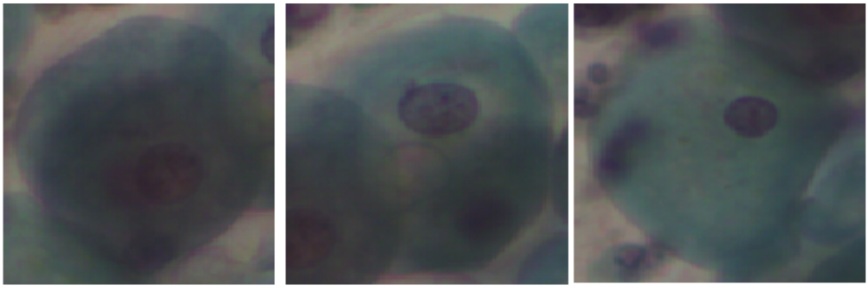

Figure 3. Moderate Dysplastic

Severe Dysplastic

Number of Samples: 350

Severe dysplastic in Figure 4 shows significantly abnormal cells in the cervical tissue. These cells show significant changes and have a higher possibility of developing cervical cancer if left untreated. Severe dysplastic cells require prompt medical intervention and treatment.

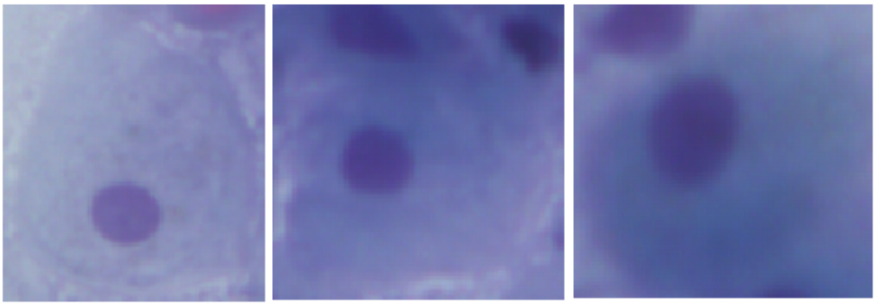

Figure 4. Severe Dysplastic

Data split: 70% (770 samples) of the samples from the data set are being used for training, 15% (165 samples) of the data is used for validation, and the remaining 15% (165 samples) is used for testing [7].

5 Results and Discussions

The graphical user interface of the segmentation process in the proposed cervical cancer classification system is presented in Figure 5. It provides users with a user-friendly platform to interact with the system and perform image segmentation on the selected Pap smear image.

1. Segmentation Options: The GUI includes controls related to segmentation techniques available in the system.

2. Segmentation Visualization: The GUI includes a visualization area where users can see the segmented regions overlaid on the original image. This allows users to visually inspect the segmentation results and assess the quality of the segmentation before proceeding to further analysis.

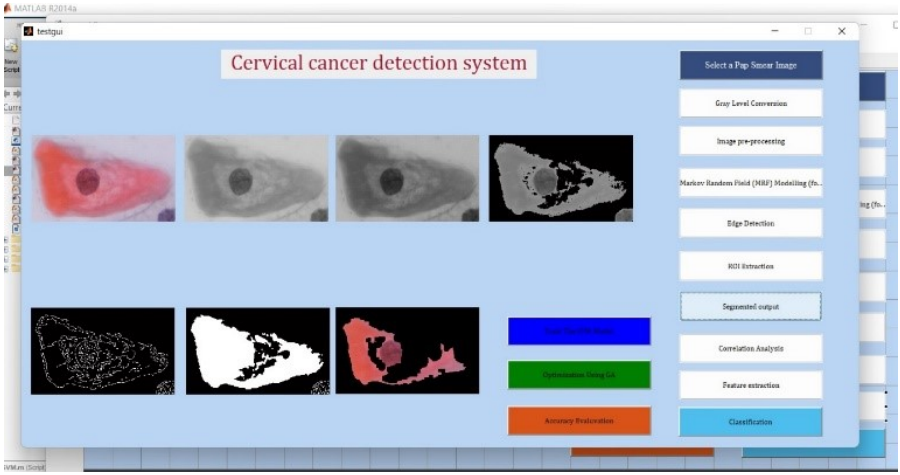

Figure 5. Segmentation process of the proposed cervical cancer classification system

The classification efficiency comparison for different methods in Cervical Cancer Classification is presented. The proposed method, IGA-SVM, is compared with existing machine learning algorithms such as Linear Regression, Logistic Regression, Decision Trees, Random Forests, and Naive Bayes. The evaluation metrics used for comparison include Accuracy, Precision, Recall, and F1-Measure [25].

The proposed method, IGA-SVM obtains the optimum Accuracy of 97.14% compared to other machine learning algorithms. It also demonstrates superior Precision (99.50 %) and F1-Measure (97.90%), indicating the ability to correctly classify the Carcinoma In Situ class with high precision and balance between precision and recall. The Recall (96.30%) is slightly lower compared to Naive Bayes, Random Forests, and Decision Trees, but still at a high level. Classification efficiency comparison for Carcinoma In Situ with other learning models is shown in Figure 6.

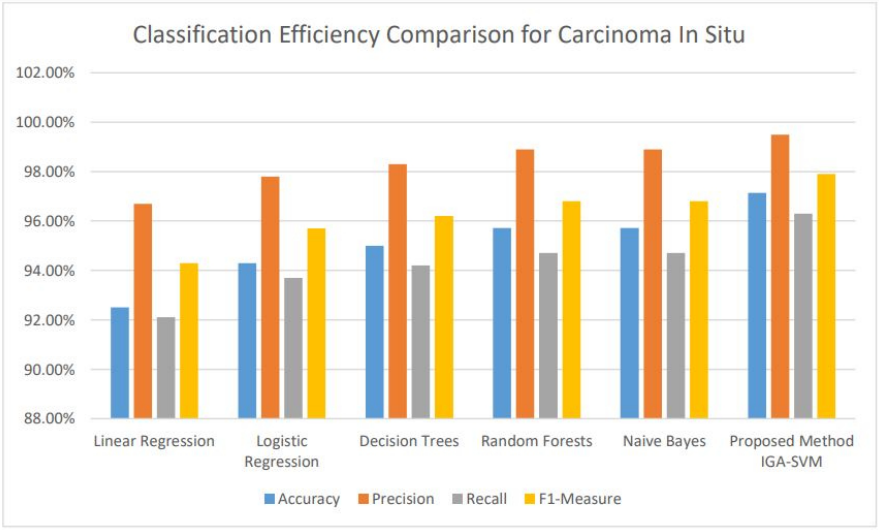

Figure 6. Classification Efficiency Comparison for Carcinoma In Situ Classification with machine learning models

Classification Efficiency Comparison for Light Dysplastic Classification with machine learning models is shown in Figure 7. According to the results, the IGA-SVM outperforms all other methods with an Accuracy of 98.1%. It also achieves high Precision (98.0%) and F1-Measure (98.7%), indicating the model's ability to accurately classify the Light Dysplastic class.

Additionally, the Recall (99.4%) is the highest among all methods, suggesting that the proposed method effectively captures the positive

instances of the Light Dysplastic class.

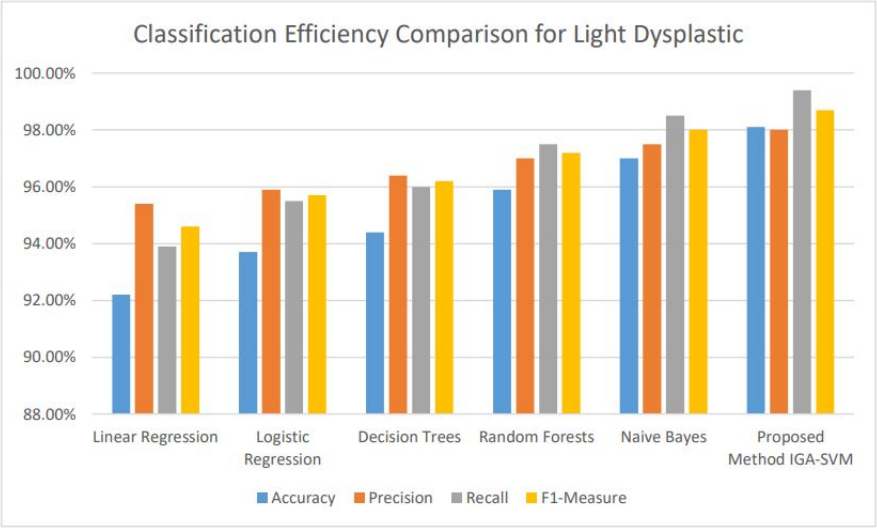

Figure 7. Classification Efficiency Comparison for Light Dysplastic Classification with machine learning models

Classification Efficiency Comparison for Moderate Dysplastic Classification with machine learning models is presented in Figure 8. Similar to previous cases, IGA-SVM achieves the highest Accuracy (98.9%), Precision (100%), and F1-Measure (99.2%) among all methods. It shows perfect Precision, indicating no false positive predictions in the Moderate Dysplastic class. The Recall (98.4%) is slightly lower compared to Naive Bayes but still at an excellent level.

IGA-SVM attains the highest Accuracy (96.1%) and F1-Measure (96.9%) among all methods in Severe Dysplastic classification model as depicted in Figure 9. It demonstrates high-precision (96.7%) and Recall (97.2%), indicating its ability to effectively classify the Severe Dysplastic class. It outperforms the other methods consistently in terms of Accuracy and F1-Measure.

According to the results, the proposed method, IGA-SVM, consistently achieves the highest or competitive classification performance across all dysplastic classes. It outperforms other machine learning al-

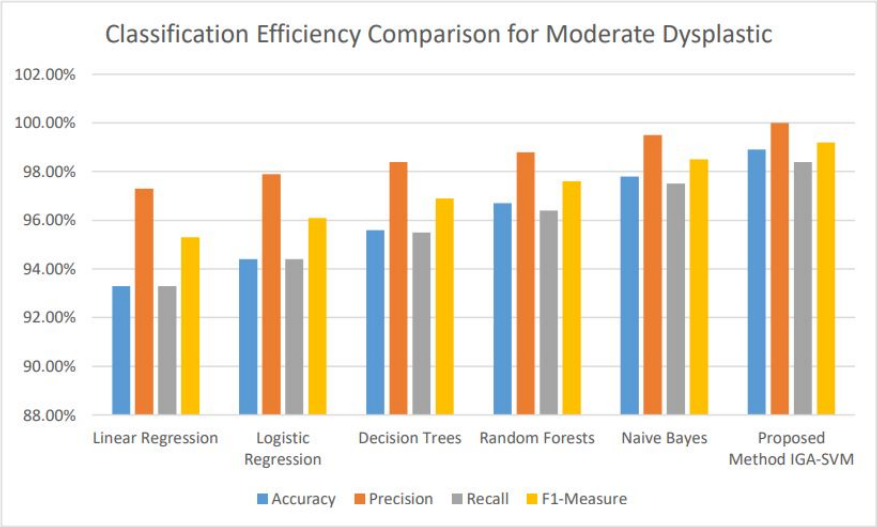

Figure 8. Classification Efficiency Comparison for Moderate Dysplastic Classification with machine learning models

gorithms in terms of Accuracy, Precision, Recall, and F1-Measure in most cases. This suggests that the combination of Improved Genetic Algorithm for feature selection and Support Vector Machine for classification provides an effective framework for cervical cancer classification based on Pap smear images.

6 Related Works

A CAD framework that classifies cytology images uses an ensemble of three standard CNN-based classifiers. The proposed ensemble model generates ranks of the classifiers using two non-linear functions which help to take into account the confidence in predictions of the base learners [8]. A CNN-based ThinPrep cytologic test (TCT) cervical cancer screening model was established through a retrospective study of multicenter TCT images. This model shows improved speed and accuracy for cervical cancer screening, and helps overcome the shortage of med-

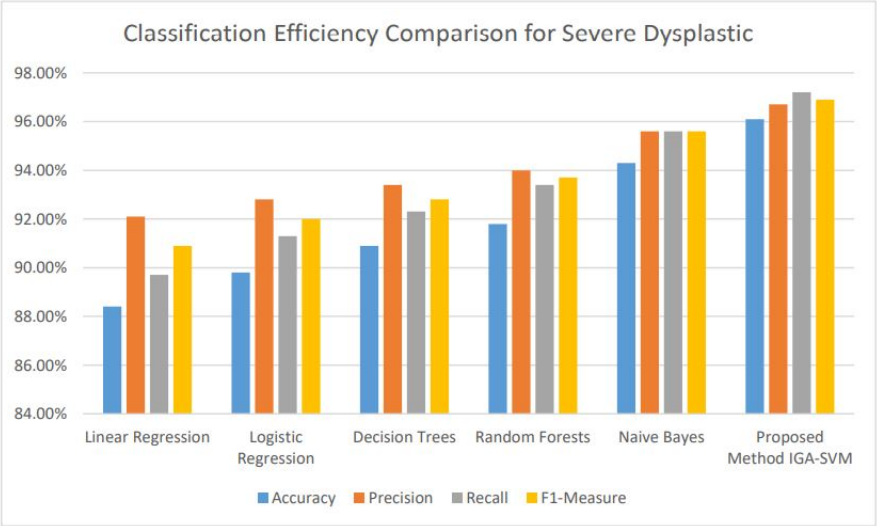

Figure 9. Classification Efficiency Comparison for Severe Dysplastic Classification with machine learning models

ical resources required for cervical cancer screening [9]. An automatic cervical cancer screening system using convolutional neural network was proposed in [10]. A novel deep learning method named AttFPN as an automated detection model for abnormal cervical cells in cervical cancer screening was proposed in [11]. The model was guided by clinical knowledge and attention mechanism, consisting of a multi scale feature fusion structure and an attention module. The proposed method outperformed the related stat-of-the-art deep learning methods and was comparable to a pathologist with 10 years of experience. DGCA-RCNN framework for the detection of abnormal cervical cells in Pap smear images was presented in [12]. It is an extended version of the Faster RCNN-FPN model by introducing deformable convolution layers into FPN to improve scalability and adding a GCA module alongside RPN to enhance the spatial context information. Cervical cell image generation model based on taming transformers (CCGtaming transformers) to provide high quality cervical cancer datasets with sufficient samples and balanced weights was developed [13].

The structure of the encoder was improved by introducing SEblock and MultiRes-block to improve the ability to extract information from cervical cancer cells images. An efficient and totally segmentation-free method for automated cervical cell screening that utilizes a modern object detector to directly detect cervical cells or clumps, without the design of a specific handcrafted feature was proposed in [14] to investigate the presence of unreliable annotations and cope with them by smoothing the distribution of noisy labels. A new framework based on a strong feature Convolutional Neural Networks (CNN) Support Vector Machine (SVM) model was proposed to accurately classify the cervical cells. A method fusing the strong features extracted by Gray-Level Cooccurrence Matrix (GLCM) and Gabor with abstract features from the hidden layers of CNN was conducted, meanwhile the fused ones were input into the SVM for classification. An effective dataset amplification method was designed to improve the robustness of the model [15].

The technique used in hospital laboratories involves the manual numeration of blood cells victimization using a device referred to as Haemocytometer. Using this method can be monotonous, sometimes produces inaccurate results and also is a time consuming process. So as to beat the complications, this analysis presents an absolute automatic systemized system to associate with nursing platelet cells within the blood samples and to classify various forms of Leukaemia [16]. The Pap smear test is a manual screening procedure that is used to detect precancerous changes in cervical cells based on color and shape properties of their nuclei and cytoplasms. Automating this procedure is still an open problem due to the complexities of cell structures. An unsupervised approach for the segmentation and classification of cervical cells was proposed. The segmentation process involves automatic thresholding to separate the cell regions from the background, a multi-scale hierarchical segmentation algorithm to partition these regions based on homogeneity and circularity, and a binary classifier to finalize the separation of nuclei from cytoplasm within the cell regions [17].

Cervical cancer is one of the leading causes of cancer death in females worldwide. The disease can be cured if the patient is diagnosed in the pre-cancerous lesion stage or earlier. A common physical ex-

amination technique widely used in the screening is Papanicolaou test or Pap test. In this automated cervical cancer cell segmentation and classification method [18], a single-cell image is segmented into the nucleus, cytoplasm, and background, using the fuzzy C-means (FCM) clustering technique.

Accurate classification of Pap smear images becomes the challenging task in medical image processing. This can be improved in two ways. One way is by selecting suitable well defined specific features and the other is by selecting the best classifier. To enhance accuracy, the earlier detection and, for diverse perspectives, the proposed research analyzed the techniques used in [19], [21], and [23] to apply in medical imaging analysis, including the classification of Pap smear images for cervical cancer detection. A nominated texture-based cervical cancer (NTCC) classification system which classifies the Pap smear images into any one of the seven classes was presented in [20]. This can be achieved by extracting well defined texture features and selecting best classifier. Pap smear test has been broadly used for detection of cervical cancer. However, the conventional Pap smear test has several shortcomings including: subjective nature (dependent on individual interpretation), low sensitivity (i.e., ability to detect abnormal changes), and the need for frequent retesting. There has been a great effort to automate Pap smear tests, and it is one of the critical fields of medical image processing [22]. A method for detecting overlapping cell nuclei in Pap smear samples was presented in [24]. The extraction of overlapping cell nuclei is a critical issue in automated diagnosis systems. Due to the similarities between overlapping and malignant nuclei, misclassification of the overlapped regions can affect the automated systems final decision.

7 Conclusions

A novel approach for optimizing cervical cancer classification using SVM and Improved Genetic Algorithm on Pap smear images has been proposed. The methodology involves image pre-processing, segmentation, feature extraction, feature selection, classification, and performance evaluation. The results show high accuracy, precision, recall, and F1-Measure in classifying cervical cancer cells. The Improved Ge-

netic Algorithm is effective for the SVM classifier's feature selection and hyper-parameter tuning. Future work will extend the approach to other imaging modalities and explore deep-learning techniques for improved cervical cancer classification. Additional research is required to confirm the methodology's efficacy on larger data sets and more clinical investigations.

References

[1] M.A. Yalda, I. Y. Abdulmalek, and H. R. M. Ali, "Women's Knowledge Regarding Pap Smear and Cervical Cancer in Duhok City in Respect to Related Educational Perspective Session," *History of Medicine*, vol. 9, no. 1, pp. 960–967, 2023.

[2] Cheon et al., "Feature Importance Analysis of a Deep Learning Model for Predicting Late Bladder Toxicity Occurrence in Uterine Cervical Cancer Patients," *Cancers*, vol. 15, no. 13, Article no. 3463, Jul. 2023, DOI: 10.3390/cancers15133463.

[3] D. R. Sarvamangala and R. V. Kulkarni, "Convolutional neural networks in medical image understanding: a survey," *Evol. Intel.*, vol. 15, no. 1, pp. 1–22, Mar. 2022, DOI: 10.1007/s12065-020-00540-3.

[4] H. Yao et al., "Semantic Segmentation for Remote Sensing Image Using the Multigranularity Object-Based Markov Random Field With Blinking Coefficient," *IEEE Trans. Geosci. Remote Sensing*, vol. 61, pp. 1–22, 2023, DOI: 10.1109/TGRS.2023.3301494.

[5] Z. Alyafeai and L. Ghouti, "A fully-automated deep learning pipeline for cervical cancer classification," *Expert Systems with Applications*, vol. 141, Article no. 112951, 2020. DOI: 10.1016/j.eswa.2019.112951.

[6] M. Alsalatie et al., "A New Weighted Deep Learning Feature Using Particle Swarm and Ant Lion Optimization for Cervical Cancer Diagnosis on Pap Smear Images," *Diagnostics*, vol. 13, no. 17, Article no. 2762, Aug. 2023, DOI: 10.3390/diagnostics13172762.

[7] M. Sumathi and S. P. Raja, "Machine learning algorithm-based spam detection in social networks," *Soc. Netw. Anal.*, vol. 13, Article no. 104, 2023, https://doi.org/10.1007/s13278-023-01108-6.

[8] A. Manna, R. Kundu, D. Kaplun, A. Sinitca, and R. Sarkar, "A fuzzy rank-based ensemble of CNN models for classification of cervical cytology," *Sci. Rep.*, vol. 11, Article no. 14538, 2021. https://doi.org/10.1038/s41598-021-93783-8.

[9] X. Tan et al. "Automatic model for cervical cancer screening based on convolutional neural network: a retrospective, multicohort, multicenter study," *Cancer Cell Int*, vol. 21, Article no. 35, 2021, DOI: 10.1186/s12935-020-01742-6.

[10] Aziz-ur-Rehman, Nabeel Ali, Imtiaz. A. Taj, Muhammad Sajid, Khasan S. Karimov, "An Automatic Mass Screening System for Cervical Cancer Detection Based on Convolutional Neural Network", *Mathematical Problems in Engineering*, vol. 2020, Article ID 4864835, 14 pages, 2020. https://doi.org/10.1155/2020/4864835.

[11] Lei Cao et al, "A novel attention-guided convolutional network for the detection of abnormal cervical cells in cervical cancer screening," *Medical Image Analysis*, vol. 73, Article no. 102197, October 2021, DOI: 10.1016/j.media.2021.102197.

[12] Xia Li, Zhenhao Xu, Xi Shen, Yongxia Zhou, Binggang Xiao and Tie-Qiang Li, "Detection of Cervical Cancer Cells in Whole Slide Images Using Deformable and Global Context Aware Faster RCNN-FPN," *Curr. Oncol.*, vol. 28, no. 5, pp. 3585–3601, 2021 Sep 16, DOI: 10.3390/curroncol28050307. PMID: 34590614; PMCID: PMC8482136.

[13]] Chen Zhao, Renjun Shuai, Li Ma, Wenjia Liu, and Menglin Wu, "Improving cervical cancer classification with imbalanced datasets combining taming transformers with T2T-ViT," *Multimed. Tools Appl.*, vol. 81, pp. 24265–24300, 2022. https://doi.org/10.1007/s11042-022-12670-0.

[14] Yao Xiang et al, "A novel automation-assisted cervical cancer reading method based on convolutional neural network", *Biocybernetics and Biomedical Engineering*, vol. 40, no. 2, pp. 611–623, April–June 2020.

[15] A. Dongyao Jia, B. Zhengyi Li, and C. Chuanwang Zhang, "Detection of cervical cancer cells based on strong feature CNN-SVM

network," *Neurocomputing*, vol. 411, pp. 112–127, 21 October 2020, https://doi.org/10.1016/j.neucom.2020.06.006.

[16] K. Balakumar, Anand T. Gokul, G. Naveenkumar, and S. Umamaheswari, "Improving the Performance of Leukemia Detection using Machine Learning Techniques," in *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, (Coimbatore, India), 2022, pp. 867–872, DOI: 10.1109/ICESC54411.2022.9885461.

[17] A. Gençtav, S. Aksoy, and S. Önde, "Unsupervised segmentation and classification of cervical cell images," *Pattern Recognit*, vol. 45, no. 12, pp. 4151–4168, 2012.

[18] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in pap smears," *Comput Meth Prog Bio*, vol. 113, no. 2, pp. 539–556, 2014.

[19] B. V. Dharani Krishna, C. Kavin Prabhu, S. Harish, and S. Umamaheswari, "Towards Building of a Robust Organic Fruit Tester," in *2022 8th International conference on Advanced computing and Communication systems(ICACCS)*, 2022, pp. 631–634.

[20] Edwin Jayasingh Mariarputham and Allwin Stephen, "Nominated Texture based Cervical Cancer Classification," *Computational and Mathematical Methods in Medicine*, Article no. 586928, 2015, DOI: 10.1155/2015/586928.

[21] S. Umamaheswari, S. Aartisha, J. Kanimozhi, and R. Suhashini, "Building accurate legal case outcome prediction models," in *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 16 June 2023, DOI: 10.1109/ICAECA56562.2023.10200651.

[22] T. A. Sajeena and A. S. Jereesh, "Automated cervical cancer detection through RGVF segmentation and SVM classification," in *2015 International Conference on Computing and Network Communications (CoCoNet)*, (Trivandrum, India), pp. 663–669, 2015.

[23] S. Umamaheswari, K. Harikumar, and D. Allinjoe, "Customer Relationship Management using Sentimental Analysis," in *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 2021, DOI: 10.1109/ICAECA52838.2021.9675766.

[24] M. Guven, C. Cengizler, "Data cluster analysis-based classification of overlapping nuclei in Pap smear samples," *Biomed Eng Online*, vol. 13, Article no. 159, 2014, DOI: 10.1186/1475-925X-13-159.

[25] A. Rehman, T. Saba, M. Mujahid, F. S. Alamri, and N. ElHakim, "Parkinson's Disease Detection Using Hybrid LSTM-GRU Deep Learning Model," *Electronics*, vol. 12, no. 13, Article no. 2856, Jun. 2023.

S. Umamaheswari, Y. Birnica,
J. Boobalan, V. S. Akshaya

Received October 11, 2023
Revised February 2, 2024
Accepted February 4, 2024

S. Umamaheswari
ORCID: https://orcid.org/0000-0001-6590-2521
Associate Professor, Dept.of ECE, Kumaraguru College of Technology
Coimbatore, Tamilnadu, India.
E–mail: umamaheswari.s.ece@kct.ac.in

Y. Birnica
Dept. Dept.of ECE, Kumaraguru College of Technology
Coimbatore, Tamilnadu, India.
E–mail: birnica.21mco@kct.ac.in

J. Boobalan
ORCID: https://orcid.org/0000-0002-9655-5435
Assistant Professor, Dept.of ECE, Kumaraguru College of Technology
Coimbatore, Tamilnadu, India.
E–mail: boobalan.j.ece@kct.ac.in

V. S. Akshaya
ORCID: https://orcid.org/0000-0001-7120-3006
Professor, Dept. of. CSE, Sri Eshwar College of Engineering
Coimbatore, Tamilnadu, India.
E–mail: vsakshaya@gmail.com