Improved Model to Detect Cancer from Cervical Histopathology Images by Optimizing Feature Selection and Ensemble Classifier

R. Baghia Laxmi [1,*], B. Kirubagari [1], and Lakshmana Pandian [2]

[1] Department of Computer Science and Engineering, Faculty of Engineering and Technology, Annamalai University, Chidambaram, India; Email: kirubacdm@gmail.com (B.K.)

[2] Department of Computer Science and Engineering, Puducherry Technological University, Puducherry, India; Email: lpandian72@pec.edu (L.P.)

*Correspondence: rbaghialaxmi2022@gmail.com (R.B.L.)

***Abstract*—Cervical Cancer (CC) remains the fourth most typical cancer internationally. Whole-Slide Images (WSIs) remain a significant benchmark for CC prognosis. Missed prognoses and mis- prognoses frequently happen because of the higher similarity in pathological cervical images, the higher quantity of readings, the lengthy reading duration and the pathologists inadequate experience degrees. The prevailing paradigms possess inadequate Feature Extraction (FE) and portrayal abilities and they are burdened with inadequate pathological classification. Hence, this study initially proffers a new FE network called NASNet alongside a genetic algorithm- based feature selection procedure. Next, the chosen features will be sent as input into the ensemble classifier to classify 4 classes—Negative for Intra-Epithelial Malignancy (NILM), Squamous Cell Carcinoma (SCC), Low Squamous Intra-Epithelial Lesion (LSIL) and High Squamous Intra-Epithelial Lesion (HSIL). The database will be split into a training set (90%) and a test set (10%). The proffered network is called Genetic NASNet Ensemble Classifier (GenNASNet_EC) and is correlated with the prevailing methodologies concerning the Accuracy, Precision, Recall, Specificity, FPR, FNR and F1-Score. Consequently, it is observed that the proffered GenNASNet_EC attains the Accuracy of 98.02%, Precision of 97.56%, Recall of 98.02%, Specificity of 99.34%, FPR of 0.0066%, FNR of 0.0198% and AUC of 98.35%.**

***Keywords*—cervical cancer, squamous cell carcinoma, high squamous intra-epithelial lesion, genetic NASNet ensemble classifier, pathological cervical images**

I. Introduction

Over the past few years, the clinical image prognosis discipline is booming towards artificial intelligence that efficiently enhanced the pathologist's prognostic efficacy and precision and inadequate medical exposure [1]. Cancer consistently remains a chief issue amongst humans, particularly Cervical Cancer (CC) that possesses higher occurrence and death tolls internationally. In the developed nations, the CC occurrence remains less because of the top-order clinical medicament. For instance, in the last thirty years, because of the progressions in diagnostic checking and preventive technology in the USA, the CC occurrence is decreased by about fifty percent [2]. Initial diagnostic checking for CC remains a vital phase in its prevention. Clinical diagnostic checking approaches for CC encompass cervical scraping, Human Papillomavirus (HPV) testing and liquid-based cell identification [3]. Subsequent to identifying Cervical Lesions (CL) via the aforementioned initial diagnostic checking methodologies, a follow-up biopsy remains requisite for pathological prognosis. Pathologists employ circular electric cutters for executing the CL's conization and acquiring tissue parts. Otherwise, the pathologists employ biopsy forceps for acquiring tissue ut of the lesions, create complete slides and employ them for creating Whole-Slide Images (WSIs) by microscopical imaging and later, create glass slides for microscopical investigation [4]. It remains the entrenched norm for CC prognosis. Lately, it turned typical to generate the slides WSIs for the employment in cancer prognosis [5]. Hence, the chief problem in CC prognosis remains in what way to detect WSIs swiftly and precisely in disparate phases.

Presently, the pathologists classify cervical Biopsy Images (BIs) as per the WHO classification of Tumors of Female Reproductive Organs, Fourth Edition chiefly by assessing the abnormal cells quantity within the BI, cell mitosis speed, cell discrepancy, abnormal cells quantity within the nucleus, polarity disorders existence, and surface cell keratinization's level [6]. Nevertheless, fit persons BIs and Low Squamous Intraepithelial Lesions (LSILs) images remain greatly alike, and the classification job remains arduous and relies greatly upon the pathologist's exposure [7]. Hence, the CL's initial detection of images remains a vital competence for clinical organizations with pathologists who possess just some years of experience or nil professional pathologists by any means [8]. Traditional Deep Learning (DL) algorithms execute nicely upon several jobs yet represent

Manuscript received November 17, 2022; revised February 15, 2023; accepted April 13, 2023; published August 11, 2023.

a decay within the generalization capability as the issues size rises. Regarding very intricate issues like computer vision, the traditional techniques incline to flop in catching the intricate functions ingrained within the higher-spatial space for excerption [9]. It inspires us to proffer a new Feature Extraction (FE) network. This study's inputs include the following:

- Classification of cervical classes four classes like Negative for Intra-epithelial Malignancy (NILM), Squamous Cell Carcinoma (SCC), Low Squamous Intra- Epithelial Lesion (LSIL) and High Squamous Intra-Epithelial Lesion (HSIL) that are performed by Ensemble-based Machine Learning (ML) algorithm.
- To embrace a new FE Network (FEN) called NASNet alongside genetic algorithm- based Feature Selection (FS) procedure for resolving the issue of less comprehensive diagnostic checking efficacy.

The remaining of this study is organized as: Section II highlights some prevailing studies, Section III exhibits the proffered technique and methods, and Section IV presents the experimental results and discussion. Lastly, Section V sums up with a conclusion and upcoming study.

II. Related Work

In Image Processing (IP), because of the images broad assortment and the huge quantity of its kinds, the color, texture, figure and spatial association features of one image remain greatly intricate and defining such Image Features (IF) turned into a chief issue. The conventional IF's excerption frequently needs analysts to possess abundant and complete professional knowledge. In the progression and calibration phases of an algorithm, modeling features physically involves great duration and energy for the analysts and the last outcomes frequently be conditional upon experience and luck to some degree. Owing to the Neural Network's (NN) progression, deep convolutional networks have given analysts novel concepts of image FE.

A. Survey upon NN-Based FE

Xue and Zhou *et al.* [10] proposes a localized, fusion-related, hybrid imaging and DL technique for classifying squamous epithelium into Cervical Intraepithelial Neoplasia (CIN) grades for the 83 digitalized histology images database. Diving the epithelium area into ten vertical sections, twenty-seven handmade IF and rectangular patches and sliding window-related Convolutional Neural Network (CNN) features will be calculated for every section. The hybrid methodology attains a 15.51% and 11.61% enhancement upon the DL and imaging techniques only accordingly having an 80.72% complete Epithelium CIN (E-CIN) classification precision exhibiting the optimized E-CIN classification capacity of fusing image and DL features. Guo and Banerjee *et al.* [11] proposes a methodology for automatedly identifying CC. The texture features (GLCM) assessment and excerption have been executed upon the case images. The K-means and marker-controlled watershed algorithms have been employed for segmenting the images for fusing GLCM and pathological features. Lastly, a Support Vector Machine (SVM) has been employed for classification and detection. The accuracy of the detected cancerous and typical classification remains 90%, and the accuracy of the CIN image classification has been just 70%. Even though this study analyzed CBTI with disparate lesions levels, an exhaustive study of LSIL and HSIL classification is lacking. Keenan and Diamond *et al.* [12] introduces an automatic, localized methodology centered upon fusion for analyzing CC tissue atypicality. Subsequent to voting for the vertical stage by employing the SVC and linear discriminant assessment methodologies for sixty-one instances of images, the greatest classification accuracy for CIN attained 88.5%. The comprehensive classification impact of this study is not fine and the CBTIs classification is not completely investigated.

Alyafeai and Ghouti [13] tries to establish an objective Scoring System (SS) employing automated Machine Vision (MV). An MV system has been established employing the KS400 macro programming language for 230 CIN images in all degrees. The classification impact of typical and CIN3 within the SS has been as higher as 98.7%. The accuracy of cytosis and CIN1 has been 76.5%. This study has not in any way analyzed the association betwixt CC phases and CIN, typical and tiny cell BIs. Tubishat and Alswaitti *et al.* [14] presents a pipeline comprising two pre-trained DL paradigms for automated cervix identification and cervical tumor classification. The initial paradigm identifies the cervix area a thousand times quicker than the advanced data-driven paradigms when attaining an identification accuracy of 0.68 concerning the intersection of union (IoU) standard. Self-excerpted features will be employed by the next paradigm for classifying the cervix tumors. Such features will be learnt employing 2 lightweight paradigms centered upon CNNs.

B. Survey upon Optimization-Based FS

Jia and Xing *et al.* [15] proffers a Dynamic Butterfly Optimization Algorithm (DBOA) as an enhanced version of the Butterfly Optimization Algorithm (BOA) for FS issues. The BOA portrays one of the latest proffered Optimization Algorithms (OAs). It established its capability for resolving disparate kinds of issues with competing outcomes when correlated with the rest of the OAs. Nevertheless, the initial BOA possesses issues while optimizing large-sized issues. These problems encompass inactivity into Local Optima (LO) and mission solutions diverseness while doing the optimization procedure. Jia and Li *et al.* [16] introduces a new metaheuristic optimizer called the Chaotic Harris Hawks Optimization (CHHO). In this, two chief enhancements have been proposed to the conventional HHO algorithm. The initial enhancement remains to implement the chaotic maps at the HHO's beginning stage for optimizing the populace diverseness in the Search Space (SSp). The next enhancement remains to

employ the Simulated Annealing (SA) algorithm to the present best resolution for optimizing HHO exploitation. For validating the execution of this algorithm, CHHO has been implemented upon fourteen clinical standard databases out of the UCI ML depository. Zheng and Li *et al*. [17] models two disparate hybrid paradigms centered upon Spotted Hyena Optimization (SHO) for the FS issue. This algorithm could search the optimal or almost optimal Feature Subset (FSt) in the FSp for lessening the provided Fitness Function (FF). In the initial paradigm, the Simulated Annealing (SA) algorithm will be embedded within the SHO algorithm (SHOA), named SHOSA-1, for optimizing the optimal solution observed by the SHOA following every repetition. In the next paradigm, the SA optimizes the last resolution acquired by the SHOA, named SHOSA-2.

Wan and Ma *et al.* [18] puts forth a hybrid FSt choosing algorithm known as the maximum Pearson maximum distance improved whale optimization algorithm (MPMDIWOA). Initially, centered upon Pearson's correlation coefficient and correlation distance, a filter algorithm has been proffered called Maximum Pearson Maximum Distance (MPMD). In MPMD, two criteria have been proffered for calibrating the weights of the pertinence and iteration. Next, the altered whale optimization algorithm could function as a wrapper algorithm. Mulmule and Kanphade *et al.* [19] presents a new discrete sine cosine algorithm (SCA)-based multi objective feature selection (MOSCA_FS) technique for hyperspectral imagery. In this methodology, a new and efficient architecture of multi-objective hyperspectral FS has been modeled. In this architecture, the proportion betwixt the Jeffries-Matusita (JM) distance and mutual information (MI) has been designed for lessening the iteration and increasing the pertinence of the chosen FSt. Furthermore, one more metric—the variance (Var)—has been implemented to optimize the data quantity. A Cervical Histopathology Image Classification (CHIC) framework based on Multilayer Hidden Conditional Random Fields (MHCRFs) is suggested to categories cervical cancer in its well-differentiated, moderately differentiated, and malignant lesions phases.

The prevailing paradigms generate an important quantity of Feature Data (FD). Since the paradigm framework rises with its input data, it really enhances the calculation intricacy and turns the paradigm very susceptible. In the prevailing paradigms, the FS job has been examined in a constricted manner upon physically cropped cancerous areas while completely automatic implementation causes calculably costly in whole slide clinical image processing. Thus, the proffered NASNet alongside genetic algorithm-based FS procedure has been employed for choosing valid features, removing unnecessary data, and providing chosen features to the classifier.

III. System Approach

In this section, it is explained the results of research and at the same time is given the comprehensive discussion. Results can be presented in figures, graphs, tables and others that make the reader understand easily [14, 15]. The discussion can be made in several sub-sections.

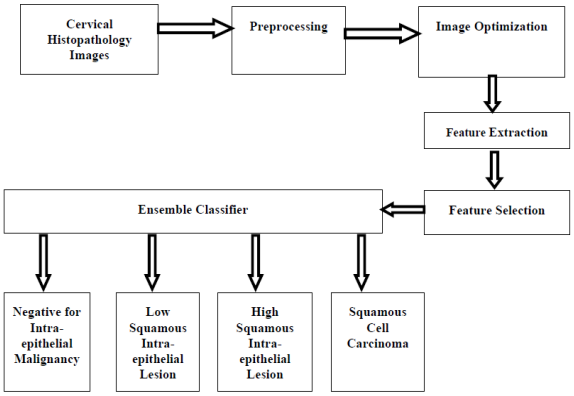

Figure 1. System framework for classifying the 4 Classes of CC.

Fig. 1 illustrates the proffered FS and selection-based classification approach for Histopathology Images (HIs) within the detection classes in CC. The CHI will be obtained out of the dataset and the images will be preprocessed by employing an effectual filter by Adaptive Median Filter (AMF) and Macenko-Stain Normalization (MSN) for eliminating redundant noise. The pre-processed image will be optimized for augmenting the initial image's quantity for the FE. The optimized image experiences additional procedures like de-texturized, de-colorized, edge-optimized and find rotate. Then, FE is performed in the optimized image by implementing the DL paradigm NASNet. The FD will be feature chosen by a metaheuristic enhanced genetic technique. The Ensemble Classifier (EC) like Random Forest (RF), SVM and Naïve Bayes will be employed to classify 4 classes—NILM, SCC, LSIL and HSIL—for optimizing the outcomes.

A. Histopathological Images Preprocessing

In the preprocessing, entire data regarding the image background texture within the image will be removed. Just the cells and cell edges will be present within the image. Lessening numerous data regarding the tissue might remain beneficial for FE methodologies. The AMF functions in a rectangular xyspace. This differs the R_{xy}dimension within the filtering procedure centered upon the requirements undermentioned. The median within the 3-by-3 neighborhood close to the correlating pel in the gathered image has been employed for generating every output pel. The image edges, contrastingly, remain substituted with zeros. The filter output contains just a single value, which substitutes the current pel value at (x, y) in which the point where R remains centered at time. Let $(A \times B)$ image, $I(p, q)$ $with$ $I(p, q) \in \{1,2, \dots, A\} \times \{1,2, \dots, B\}$, a bidimensional $(x \times x)$ median filter will be provided as,

$$I(p, q) = median\{I(p + u, q + v)\} \quad (1)$$

in which

$$u, v\epsilon \left(-\frac{(s_{\min}+s_{\max})}{s_{\mathrm{mad}}}, \ldots, \frac{(s_{\min}+s_{\max x})}{s_{\mathrm{mad}}}\right) \cdot S_{xy} + R_{\max} \quad (2)$$

and (I, j) indicate the pel point at (i, j).
S_{min}= minimum pel value of R_{xy}
S_{max}= maximum pel value of R_{xy}
S_{med} = median pel value of R_{xy}
S_{xy} = pel value at coordinates (x, y)
R_{max} = maximum permitted R_{xy} dimension.

Therefore, AMF will be employed for smoothening the non-repulsive noise evolving out of two-dimensional edges devoid of blurring boundaries and preserved images. The preprocessing approach alongside AMFs produces smoothened images when sustaining the edges that remain a vital feature. Subsequent to the filtration procedure, normalization will be performed employing MSN. For normalizing the source image's $s$color appearance (CA) to that of a target image t, initially, their CA and stain density maps (SDMs) will be analyzed by factorizing V_s into $W_s V_s$ and V_t into $W_t V_t$ employing the proffered MSN expression undermentioned. Next, the density map's scaled variant of source Hs will be joined with the CA of the target $W_t$rather than that of the source W_s for producing the normalized source image. It sustains the framework concerning the stain density (SD) H and just modifies the semblance concerning $W$and could be explained by,

$$H^{nor}(j,:) = \underline{Hs\,(j,:)}\; H^{RM}(j,:)\, j=1,2,\ldots,r \quad (3)$$

$$Vnorm = W\, Hnorm$$

$$s \qquad t \quad s$$

$$I^{norm} = I \exp(-V^{norm})$$

$$s \qquad 0 \qquad s$$

In which $H^{RM} = R(H_i) \in R^{r\times 1}$, $i = (s, t)$ and $RM(\cdot)$ calculate every row vector's strong pseudo maximal at 99%. Dissimilar to Non-Linear (NL) mapping betwixt stats of H_s and H_t, herein $H_s$will be just multiplied by a scalar and, thus, maintain the source image's relative SDMs complete. In such a manner, when the precise stain partition is performed, the color normalization approach solely modifies the stain CA (basis) when sustaining the source's framework.

B. Image Optimization

This segment provides an effectual technique employing Gaussian Laplacian Pyramid Blending (GLPB) for HI data optimization, and this is illustrated in Fig. 2. Picture optimization involves minimizing the file size of your photographs without compromising quality in order to maintain load time. The hypothesis remains that the pyramid blending appears as the very appropriate resolution for dealing with this problem and generates a natural-looking HIs. The procedure comprises decaying every image into an array of spatial-frequency bands. Subsequent to this, a band-pass composite could be built in every band through a transformation region.

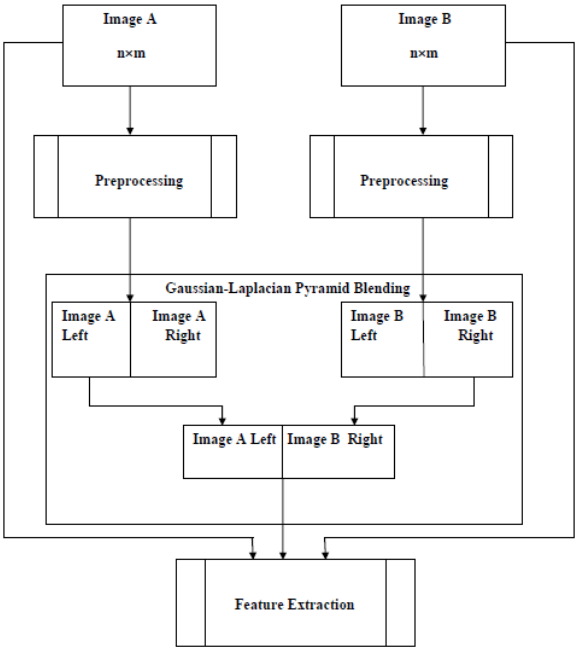

Figure 2. Image optimization procedure.

For joining $I_{A,eft}$ and $I_{B,right}$ into a composite image having minimal or nil obvious junction borders, the Laplacian Pyramid (LP) will be employed for smoothening the border in a scale- dependent manner for preventing boundary artifacts. Presuming that I_A and I_B contain a similar solution, the Gaussian Laplacian Pyramid Blending (GLPB) will be modeled by the ensuing phases:

- Input: Pre-processed images I_A and I_B, and a(binary) mask R, which indicates the blend (0 = I_A, 1 = I_B), alongside $A \neq B$
- Construct I_{Ai}'s (AL) LP, $i \in 0...N$; I_B 's (BL) LP, $j \in 0,...,N$ and R_p's (RG) Gaussian pyramid $p \in 0,...,N$.
- Construct an LP for the outcome F, employing linear interpolation through each pel having a blend mask provided at disparate degrees of particular for each pyramid degree k: $FL_K(r, c) = (1 - RG_K(r, c)) \times AL_K(r, c) + RG_K(r, c) \times BL_K(r, c)$in which r and c represent row and column accordingly.
- Rebuild the complete-resolution image for F by constructing F Gaussian (FG) and F Laplacian (FL) by,

$$F = FG_0 \text{ from } FL_i,\ i \in 0,1,2,\ldots, N$$

$$FG_N = FL_N$$

$$FG_K = FL_K + EXPAND\,(FG_{K+1}) \quad (4)$$

The above scheme focuses on preventing a paradigm that learns sick persons color portrayals that, indeed, remain associated yet to the staining procedure.

C. FE Employing NASNet

In this segment, a pre-trained NASNet paradigm will be imported. Wide-range database deficit requires the employ of a pre-trained paradigm. Next, the dense layers

of 128 × 1 substitute the paradigm's FE portion, that is, the paradigm's head, 3 × 1 and 128 × 1, 2 × 1 for binary and tertiary classification accordingly. Rather than searching the whole cells, just the finest cells will be chosen in the NASNet framework. It would turn the search quicker and, hence, additional normalized features can be acquired. The NAS SSp possesses a controller-child framework in which the controller will be a usual cell and the child will be a decrement cell, as illustrated in Fig. 3.

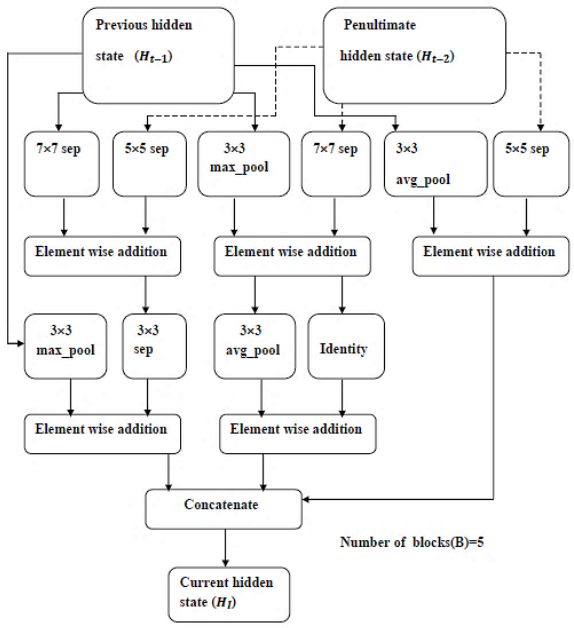

Figure 3. NASNet's framework for FE.

Hidden state: It remains the pooling layer (PL) and is called the pool. In NASNet, max pooling will be solely employed and normally the pooling kernel's dimension will be 2 × 2 through the stride of 2.

Component-wise inclusion: This layer will be performed within NASNet alongside the convolution's assistance. Its dimension's configuration remains n1×n2 in which n1 and n2 represent the dimensions of input and output tensors respectively. n1 remains a triplet (7 × 7 × 512) while n2 remains normally an integer.

Dropout: This layer called "drop" will be employed for DL procedure optimization. This positions some counts, which are connected with the specific node percentage network towards zero and MVGG_16 fixes this as 0.5 within the two dropout layers. It is A Straightforward Method That Avoid Overfitting in Neural Networks. During activation, it is possible to disregard some neurons at randomly.

ReLU layer (ReLUL) that consistently ensures the convolution layer (CnL) within NASNet enriches the NASNet's nonlinearity. The CnLs exist inside both the PLs have the equivalent channel number, kernel dimension, and stride. Indeed, gathering twain 3 × 3 CnLs and triad 3 × 3 convolution kernels (CKs) remains equivalent to a solo 5 × 5 and 7 × 7 CnLs accordingly. Stacking two or three tiny CKs functions more swifter when compared with a solo big CK. Furthermore, criteria numbers will be lessened. ReLUL that will be included amid under- dimensioned CnLs will be actually beneficial. The input frame images (FIs) and their correlating map video FIs remain S = (S(1),…,S(N)) and M = (M(1),…, M(N)) accordingly. The chief intention remains in crafting a paradigm that maps S toM using a few training data (TrD). It is designed as a probabilistic technique by learning the dispensation paradigm through the labels that are portrayed by,

$$P(n(M, i, w_m)|n(S, i, w_s)) \quad (5)$$

in which n (I,i,w) indicates a patch having w×w dimension for the image I, concentrated upon pel i. In this, w_s remains favored to be greater in order thereby additional contextual data could be excerpted. Its functional format f can be provided by,

$$f_i(s) = \sigma(a_i(s)) = P(m_i = 1|s) \quad (6)$$

in which a_i and f_i portray the input's total for the i^{th} output and i^{th} output element's importance accordingly. $\sigma(x)$, a logistic utility, can be described as,

$$\sigma(x) = \frac{1}{1+\exp(-x)} \quad (7)$$

NASNet with softmax output (SO) module will be employed for multi-class marking (MCM). The SO remains vector of dimension L that defines the conveyance greater than potential marks of pel i. Alongside those lines for MCM, when the path out of pel I towards output module l remains considered, the recomposed condition will be:

$$f_{il}(S) = \frac{\exp(a_{il}(S))}{z} = P(m_i = l|s) \quad (8)$$

in which $f_i(s)$ portrays the prognosis probability wherein pel i will be mapped to label j. The proffered methodology's benefits will be expressed as follows:

- Initially, NASNet feasibly deals with a lot of labeled data out of diverse domains.
- Next, this remains quicker while paralleled with the Graphics Processing Unit (GPU). Thus, it remains as well expanded for additional pel quantity. The TrD will be simulated by lessening the kernel dimension via the proffered methodology's calculative learning procedure.
- Each patch within the TrD will be provided with initiative sigma. Because of the training patches' huge quantity, optimization turns intricate. It could be performed by employing a binary classifier, which employs minimal patches. Some hyper-parameters are modified to a few levels. This assessment upon sensitivity is described by hyper-parameters for them to be adjusted with greater precision.

Concatenate Layer: Fig. 3 exhibits the residual extended skip (RES) block's framework. The input to

this framework has been provided to five parallel connections. In the initial four of them, two CnLs have been implemented. In every connection with CnLs, $N \times 1$ filter dimension (FDn) has been employed for the initial CnL and $1 \times N$ FDn for the next CnL. We employed two cascaded convolution layers (CCL) instead of employing a solo CnL having the FDn of $N \times N$. Employing two CnLs produces a lower criteria quantity that advantages the comprehensive framework. Furthermore, while doing the experimentations, the observance has been done in which the influence of the CCLs having the lesser criteria's quantity remains the same as the convolution's solo layer possessing a greater criteria's quantity. The final connection remains the skip connection in which the input remains as the forwarded one. Entire outputs out of the four connections have been summarized for obtaining a solo output. The three CnLs consecutively have been implemented upon the summarized output. The three CnLs possess FDns of 3×3, 3×3 and 1×1. The RES block produces the mid-range features out of the lower-range features that assist in managing data deterioration. The cancer areas possess enormous dimension differences wherefore RES executes contextual aggregation upon multi-scales that turn this scale-invariant. The RES rises the valid receptive field and permits the NASNet to possess finer FE.

Penultimate Hidden State Layer (PHSL): The input towards PHSL has been provided to two parallel connections. These connections possess two CnLs. In the initial one, the two CnLs employ $N \times 1$ and $1 \times N$ accordingly. The next one initially employs the $1 \times N$FDn, and, later, another CnL possesses FDn of $N \times 1$. The modification in amalgamation in these two connections creates a fine feature set that could provide the last outcome. The outputs out of the two connections have been summarized and processed as PHSL's output.

D. FS Employing a Genetic Algorithm (GA)

The FS objective remains to enhance the accuracy level, lessen size and training duration, and optimize normalization by lessening overfitting. FS approaches remain a subset of FE's normal discipline. FE is employed for generating novel features out of initial features' functions, while FS gives back an array of options.

GA has been employed in this research for selecting the Optimal Features (OF) through the tournament choosing approach, and the tournament's dimension, herein, remains two partakers as illustrated in Fig. 4. Populace dimension, populace kind, and the productions' quantity have been entirely provided the value Bit String (BS) and respectively in the input value task. While it is performed, the procedures of even mutation and computation crossover have been executed having the mutation probability and crossover probability of 0.10 and 0.8 accordingly.

Encoding Solutions: The initial phase for executing GA remains the solutions' encoding. In GA, All potential solutions (PSs) are completely portrayed by a numerical vector. The historical encoding remains BSs; nevertheless, actual encoding possesses numerous usages and advantages. Herein, each solution should possess the peaks' labels and the correlating threshold values.

Original populace: As in whatsoever phase-by-phase optimization disadvantage, the data of fine starting criteria benefits the algorithm's convergence speed. Yet, this type of data scarcely exists. It results in covering a huge portion of Solution Space (SSp) because of the haphazard original populace's production. Hence, for turning the SSp's exploration simpler, a more heterogeneous original populace remains appropriate. Herein, criteria to be fixed remain the ensuing: for every s_{pop} independent of the original populace, there remains the peaks' quantity α, the peaks' set employed, and peaks' correlating thresholds. Furthermore, a peak possessing huge amplitude amid the diverse spectra remains probably to remain very intriguing for discrimination.

Fitness values (FVs): Since discrimination of diverse spectra remains the aim, the FVs are considered for the fine classification rate r attained by each PS. Furthermore, normally, a council generated of diverse peaks remains apparently for executing fine discrimination than a little one. Nevertheless, employing numerous Decision Stumps (DS) can result in overfitting the training set and missing its normality. Hence, a Parsimony Term (PT) regarding the peaks' quantity αwithin the council has been included. Since there remains nil former data of the DS' optimal range needed, the PT $\rho(a)$ has been described as the council dimension's Linear Function (LF) where tiny dimensions remain preferable:

$$\rho(a) = \alpha a + \beta \tag{9}$$

The FF (F_fun) will be provided as,

$$F_{fun} = r + c \times \rho(a) \tag{10}$$

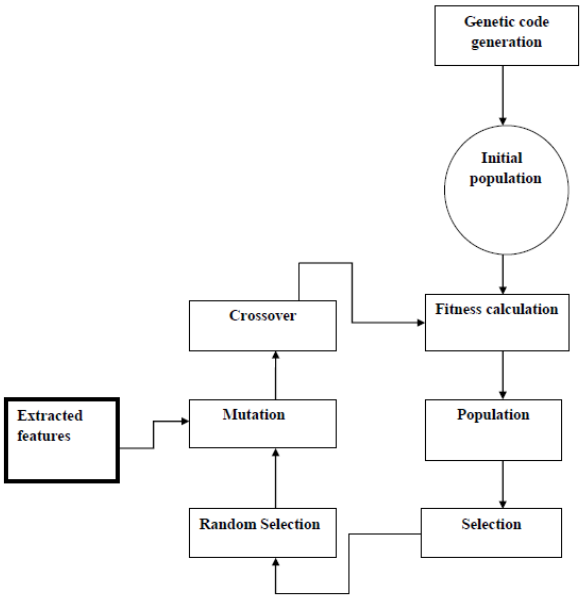

Figure 4. GA's flowchart for FS.

Selection Phase: This phase promoted the FV. As per the FV, independents have been ranked and the top rank has been provided to the finest one. Next, for having a solution in the subsequent production, its probability remains:

P (choosing k^{th} stratified solution) = $\partial + \mu \times k$, in which ∂ and μ are selected, hence

$$\sum_{k=1}^{Spop} \partial + \mu \times k = 1 \quad (11)$$

Crossover Phase: This phase's aim remains to collect intriguing features (peaks and thresholds) of numerous solutions within novel independents by creating an amalgamation of the formerly sustained solutions. It becomes requisite to note that this phase remains separated from the optimization, i.e., a crossover could generate fine and bad solutions equitably. Just the choosing phase has been employed for removing the worst solutions.

Mutation Phase: This phase creates the requisite risk for effectively exploring the SSp. Whatsoever this space's point could be achieved, and this can be ensured. Furthermore, when an LO is attained, a very swift convergence towards this LO would be prevented by mutation. This mutation rate (the solutions' ratio that would experience mutation) πm can be described for each production. Here in, mutations would be split into three kinds:

- Peak removal: Haphazardly a peak will be selected and eliminated out of the solution (that is, the council),
- Peak inclusion: A novel peak will be selected, included, and, later, the optimal threshold will be related,
- Threshold reposition: Haphazardly one of the thresholds will be eliminated and substituted by any other one.

E. Ensemble Classifier

Ensemble Learning (EL) remains fundamentally employed for enhancing the paradigm's execution (for instance, classification, prediction, and function approximation), or lessening the similarity of an inappropriate choice to a bad one. The rest of EL's implementations encompass designating confidence to the decision done by the paradigm, choosing OF (or near optima), data fusion, incremental learning, non-stationary learning and error-correcting.

RF: The RF determines the last comprehensive classification centered upon the preponderance acquired by tree voting. The RF formation could be defined in the ensuing phases.

- Produces N number of bootstrap samples out of the database.
- Every node obtains the features haphazard sample of dimension m in which $m < M$ (M indicates the features complete quantity).
- Builds a split by employing the m features chosen in phase 2 and computes the k node employing the finest split point (k indicates subsequent node).
- Redo splitting the tree till just one leaf node is achieved and the tree remains finished.
- The algorithm will be trained individually upon every bootstrapped.
- Employs the trees classification voting for gathering the prognosis data out of the (n) trained trees.
- Employs the greatest voted features for constructing the last RF paradigm.

SVM: SVM learns the classification via a training database of the format

$$\{x', y_i\},\ x' \in R^n,\ y_i \in \{-1,1\},\ i = 1,2,\ldots,\ l$$

Every one of the TrD'sl instances has an n-dimensional vector x, which defines the particular instance's features, and a label y which classifies the instances that are affiliated with 1 among the 2 classes — 1 or −1 (positive or negative) accordingly. Provided adequate training instances, the SVM will be capable of classifying formerly unknown instances (data examples), having nil predetermined label, into 1 of the 2 classes.

Naïve Bayes: Posterior Probability (PP) exists for every class. Every class' PP can be computed. The following expression exhibits the naïve Bayes formulation.

$$\mathrm{P}(\mathrm{ci}/\mathrm{v1}, \mathrm{v2}, \ldots, \mathrm{Vn}) = \frac{p(ci)\prod_{j=1}^{n} p(\mathrm{Q}_{ei})}{p(v1,v2,\ldots,vn)} \quad (12)$$

P($ci/v1$,$v2\ldots.vn$) signifies the feature excerpted image alongside its specific class ci; ci must be inferior to v, in which v=$v1$,$v2\ldots.vn$ that remains noticed as an instance. Then, the samples have been split into 10 haphazardly produced subsets. Additionally, a tenfold cross-validation methodology reliant upon the naïve Bayes paradigm has been employed for training the algorithm ninefold and testing this upon the rest onefold instead.

The arithmetical formula for the proffered ensemble methodology has been defined. Consider the confidence scores for C class quantity provided by the Base Learner (BL) i are $p_1^i, p_2^i, p_3^i, \ldots, p_n^i$ in this i=1,2,3. Initially the entire C has been collected out of every BL. Since ($p_1^i, p_2^i, p_3^i, .. p_n^i$) portraying the probabilities,

$$\sum_{k=1}^{c} p = 1, \forall i = 1,2,3 \quad (13)$$

Consider $(R_1^{i1}, R_2^{i2}, R_3^{i3}, R_c^{i1})$ and $(R_1^{i2}, R_2^{i2}, R_3^{i2}, R_c^{i2})$ remain the fuzzy ranks (FR) produced by employing the two NL functions. The FR can be computed as,

$$R_k^{i1} = 1 - \tanh\left(\frac{p_k^i}{2}\right) \quad (14)$$

The domain's description for the functions computing NL rankings would be [0, 1] as (p_k^i) $[0, 1]$. Owing to EC, cancer kinds like NILM, SCC, LSIL, and HSIL will be classified.

IV. Performance Analysis

The experimental outcome has been performed employing the criteria employed for assessment remains Accuracy, Precision, Recall, Specificity, FNR, FPR and AUC. Those criteria will be correlated with the advanced methodologies like AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble and RN_RB_AN with the proffered Genetic NASNet Ensemble Classifier (GenNASNet_EC).

A. Dataset Explanation and System Configuration Settings

The whole application of this proffered FE and classification has been carried out in the Python tool and the configurations regarded for the experiment include PC with Windows 10 Pro, 8GB RAM and Intel i3 processor CPU @1.70GHz, 64-bit operating system. The Kaggle database has been employed for inferring the experimental data. The colposcopy CC's database has been split into 10%, 8% and 10% of authentication, assessment and testing accordingly. The layer depth, initial learning rate, momentum value, optimizer and L2 value has been calculated out of Bayesian optimization. Fifty remains the fixed epochs quantity in the training paradigm. The multiple-GPU setting, 0.0001 as the initial learning and 64 as the batch dimension are employed in training.

- Staining method: IHC Staining, AQP, HIF, VEGF.
- Magnification: 400×.
- Microscope: Nikon (Japan).
- Acquisition software: NIS-Elements F 3.2.
- Image size: 1280 × 960 pixels.
- Image format: "*.tiff" or "*.png".
- Image types: Well differentiation: The tumour cells are closer to normal cells, cell heteromorphism is relatively small, cell sizes and morphology are similar; Moderate differentiation: Most cancer cells are concentrated in moderately differentiated, the characteristic is between well differentiated and poorly differentiated cervical cancer cells; Poorly differentiation: The cell structure is not visible, and the topological structure is disordered.

B. Metrics

- **Accuracy** signifies the projected network paradigm's normal prognosis ability. True Positive (TP) and True Negative (TN) calculate the classifier format's ability for computing the presence and absence of the sickness. False Positive (FP) and False Negative (FN) compute the false prognoses quantity generated by the paradigm.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (15)$$

- **Precision** signifies the comprehensive attainment of the leaf sickness classification paradigm accordingly. This remains the classification function's similarity that remains the prediction of the result as TP rate at the existence of sickness. This will be as well identified as TP quantity and could be calculated by,

$$Precision\ (P) = \frac{\mathrm{TP}}{\mathrm{FP+TP}} \quad (16)$$

- **Recall** signifies the classifier's similarity, which attains result as negative at the non-existence of sickness. This, alternatively, is called TN rate and could be computed by,

$$Recall\ (R) = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \quad (17)$$

- **Specificity** calculate the efficiency of individual classifiers for identifying negative labels as

$$Specificity(S) = \frac{\mathrm{TN}}{\mathrm{FP+TN}} \quad (18)$$

- **False positive rate** calculate the efficiency of individual classifiers for identifying false positive labels as

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP+TN}} \quad (19)$$

- **False negative rate** calculate the efficiency of individual classifiers for identifying false-negative labels as

$$\mathrm{FNR} = \frac{\mathrm{FN}}{\mathrm{FN+TP}} \quad (20)$$

- **Area under curve** calculate the effectiveness for individual classifiers to avoid the false classification as

$$\text{Area under Curve} = \frac{1}{2}\left(\frac{\mathrm{TP}}{\mathrm{TP+TN}} + \frac{\mathrm{TN}}{\mathrm{TN+FP}}\right) \quad (21)$$

Fig. 5 illustrates the confusion matrix where the rows indicate the anticipated class (output class), and the columns indicate the real class (target class) for classification. The crosswise cells indicate the tested networks, which will be rightly and wrongly classified. The right column portrays each anticipated class whereas the bottom row portrays each real class' execution. The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble, RN_RB_AN with the proffered GenNASNet_EC are exhibited in Table I defines the performance analysis of proposed method with different parameters.

TABLE I. PERFORMANCE ANALYSIS

Accuracy Correlation									
Number of Epochs	**AlexNet**	**VGG-16**	**VGG-19**	**ResNet-50**	**ResNet-101**	**GoogleNet**	**Ensemble**	**RN_RB_AN**	**GenNASNet_EC (proposed)**
10	81.0	81.3	**81.5**	86.2	88.1	84.9	92.6	94.3	97.3
20	82.6	83.1	**83.2**	86.7	88.6	85.4	92.9	94.8	97.1
30	84.0	84.3	**84.6**	88.1	89.3	86.3	93.2	95.2	98.2
40	84.5	84.9	**85.1**	88.5	89.6	86.6	93.9	95.6	98.3
50	85.1	85.2	**85.3**	88.8	89.9	87.1	94.1	95.9	98.1
Precision Correlation									
Number of Epochs	**AlexNet**	**VGG-16**	**VGG-19**	**ResNet-50**	**ResNet-101**	**GoogleNet**	**Ensemble**	**RN_RB_AN**	**GenNASNet_EC (proposed)**
10	79.4	81.0	84.6	85.7	83.4	84.6	85.4	96.4	97.6
20	79.6	81.7	85.0	86.1	83.8	84.9	85.9	96.9	97.1
30	80.0	82.0	85.3	86.4	84.1	85.2	86.3	97.14	97.2
40	80.5	82.5	85.9	86.9	84.6	85.6	86.9	97.5	97.3
50	80.9	83.1	86.1	87.2	84.9	85.9	87.2	97.9	97.2
Recall Correlation									
Number of Epochs	**AlexNet**	**VGG-16**	**VGG-19**	**ResNet-50**	**ResNet-101**	**GoogleNet**	**Ensemble**	**RN_RB_AN**	**GenNASNet_EC (proposed)**
100	80.4	79.2	81.6	85.9	86.5	87.5	86.1	94.2	98.1
200	80.9	79.6	81.9	86.1	86.9	87.9	86.5	94.9	98.0
300	81.2	80.4	83.2	86.3	87.1	88.1	87.5	95.2	98.2
400	81.6	80.9	83.6	86.9	87.6	88.5	87.9	95.6	98.5
500	81.9	81.2	83.9	87.1	88.1	88.9	88.1	95.9	98.4
Specificity Correlation									
Number of Epochs	**AlexNet**	**VGG-16**	**VGG-19**	**Res net-50**	**ResNet-101**	**GoogleNet**	**Ensemble**	**RN_RB_AN**	**GenNASNet_EC (proposed)**
10	81.6	80.1	78.1	85.1	87.2	85.9	86.5	97.9	99.3
20	81.9	80.5	78.6	85.5	87.9	86.2	86.9	98.2	99.2
30	82.3	81.4	79.4	86.4	88.4	86.5	87.2	98.4	99.1
40	82.6	81.9	79.9	86.9	88.6	86.9	87.6	98.6	99.5
50	82.9	82.1	80.2	87.2	88.9	87.2	87.9	98.9	99.4
FPR Correlation									
Number of Epochs	**AlexNet**	**VGG -16**	**VGG-19**	**Res net-50**	**ResNet-101**	**GoogleNet**	**Ensemble**	**RN_RB_AN**	**GenNASNet_EC (proposed)**
10	0.25	0.35	0.12	0.018	0.019	0.07	0.05	0.013	0.005
20	0.29	0.37	0.14	0.019	0.019	0.09	0.07	0.014	0.006
30	0.31	0.38	0.15	0.10	0.10	0.10	0.10	0.015	0.005
40	0.33	0.41	0.18	0.12	0.11	0.13	0.15	0.016	0.006
50	0.35	0.45	0.21	0.14	0.15	0.15	0.18	0.017	0.006
FNR Correlation									
Number of Epochs	**AlexNet**	**VGG -16**	**VGG-19**	**Res net-50**	**ResNet-101**	**GoogleNet**	**Ensemble**	**RN_RB_AN**	**GenNASNet_EC (proposed)**
10	0.13	0.21	0.09	0.10	0.11	0.12	0.11	0.043	0.01
20	0.15	0.22	0.10	0.11	0.12	0.13	0.12	0.045	0.012
30	0.18	0.23	0.11	0.13	0.13	0.13	0.13	0.047	0.011
40	0.20	0.24	0.13	0.14	0.15	0.14	0.15	0.049	0.012
50	0.21	0.25	0.14	0.16	0.16	0.15	0.17	0.051	0.013
AUC Correlation									
Number of Epochs	**AlexNet**	**VGG -16**	**VGG-19**	**Res net-50**	**ResNet-101**	**GoogleNet**	**Ensemble**	**RN_R B_AN**	**GenNASNet_EC (proposed)**
10	69.5	72.5	79.5	80.2	85.4	86.1	88.4	94.3	98.2
20	69.9	72.6	79.9	80.6	85.6	86.5	88.6	94.8	98.1
30	70.0	73.0	80.0	81.0	86.0	87.0	89.0	95.2	98.3
40	70.2	73.1	80.2	81.4	86.3	87.3	89.2	95.6	98.1
50	70.6	73.9	80.5	81.6	86.5	87.7	89.5	95.9	98.0

TABLE II. COMPREHENSIVE CORRELATIVE OUTCOMES

Methodology	**Accuracy**	**Precision**	**Recall**	**Specificity**	**FPR**	**FNR**	**AUC**
AlexNet	0.840	0.800	0.812	0.823	0.31	0.18	0.70
VGG-16	0.843	0.82	0.804	0.814	0.38	0.23	0.73
VGG-19	0.846	0.853	0.832	0.794	0.15	0.11	0.80
ResNet-50	0.881	0.864	0.863	0.864	0.10	0.13	0.81
ResNet-101	0.893	0.841	0.871	0.884	0.10	0.13	0.86
GoogLeNet	0.863	0.852	0.881	0.865	0.10	0.13	0.87
Ensemble	0.932	0.863	0.875	0.872	0.10	0.13	0.89
RN_RB_AN	0.950	0.970	0.950	0.980	0.02	0.05	0.93
GenNASNet_EC	0.9802	0.9756	0.9802	0.9934	0.0066	0.0198	0.9835

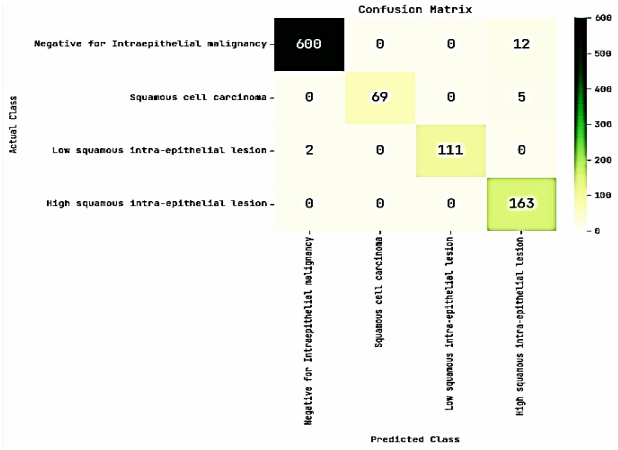

Figure 5. Confusion matrix portrayal.

The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble, RN_RB_AN with the proffered GenNASNet_EC are exhibited in Table I for accuracy While correlated, the prevailing methodologies attain 84%, 84.3%, 84.6%, 88.1%, 89.3%, 86.3%, 93.2% and 95%, whereas the proffered methodology attains 98.02% accordingly. The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble, RN_RB_AN with the proffered GenNASNet_EC are exhibited in Table I for precision. While correlated, the prevailing methodologies attain 80.08%, 82.06 %, 85.38%, 86.46%, 84.16%, 85.24%, 86.34%, and 97%, whereas the proffered methodology attains 97.56% accordingly. The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble, RN_RB_AN with the proffered GenNASNet_EC are exhibited in Table I for recall. While correlated, the prevailing methodologies attain 81.2%, 80.26%, 82.84%, 86.46, 87.24%, 88.18%, 87.22% and 98%, whereas the proffered methodology attains 98.02% accordingly. The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble and RN_RB_AN with the proffered GenNASNet_EC are exhibited in Table I for specificity. While correlated, the prevailing methodologies attain 81.2%, 80.26%, 82.84%, 86.46, 87.24%, 88.18%, 87.22%, and 98%, whereas the proffered methodology attains 99.34% accordingly. The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble and RN_RB_AN with the proffered GenNASNet_EC are exhibited in Table I for the FPR. While correlated, the prevailing methodologies attain 0.31, 0.38, 0.15, 0.1, 0.1, 0.1, 0.1and 0.02, whereas the proffered methodology attains 0.006 accordingly. The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble and RN_RB_AN with the proffered GenNASNet_ECare exhibited in Table I for the FNR. While correlated, the prevailing methodologies 0.18, 0.23, 0.11, 0.13, 0.13, 0.13, 0.13 and 0.05, whereas the proffered methodology attains 0.0198 accordingly. The execution measurements of diverse approaches of AlexNet, VGG-16, VGG-19, ResNet-50, ResNet-10, GoogLeNet, Ensemble and RN_RB_AN with the proffered GenNASNet_EC are exhibited in Table I for the AUC. While correlated, the prevailing methodologies 70%, 73%, 80%, 81%, 86%, 87%, 89% and 93%, whereas the proffered methodology attains 98.35% accordingly.

Table II exhibit the correlative assessment for diverse approaches concerning the accuracy, precision, recall, specificity, FPR, FNR, and AUC. Centered upon atop correlation, the proffered approach attained optimal outcomes in CC identification out of HIs. The proffered GenNASNet_EC approach attained an accuracy of 98.02%, precision of 97.56%, recall of 98.02%, specificity of 99.34%, FPR of 0.0066%, FNR of 0.0198% and AUC of 98.35% that are optimized while correlated with the prevailing approaches.

V. Conclusions

This study observes that CC's precise detection and classification depends upon the pathologists professional knowledge and diagnostic experience. Based on the conventional background, we analyzed the fundamental causes for the less detection accuracy of the prevailing computer-aided methodologies of the identification of pathological cervical tissue images and the inadequate identification of the objects; we resolved the issue of prevailing paradigms missing FE, expression abilities and DL having tiny samples. This study initially proffered a new FE network called NASNet with GA-related FS procedure. Next, the chosen features have been inputted into the EC for classifying four classes — NILM, SCC, LSIL and HSIL. The proffered GenNASNet_EC EL-SVM and cGANs are correlated and observed that the proffered GenNASNet_EC attains the accuracy of 98.02%, precision of 97.56%, recall of 98.02%, specificity of 99.34%, FPR of 0.0066%, FNR of 0.0198% and AUC of 98.35%. Nevertheless, because of the medical experimentations complexity and adversity in acquiring pathological data, the algorithm proffered in this study is not made susceptible to correlating medical double-blind authentication experimentations and the comprehensive detection accuracy incorporates additional enhancements.

Conflict of Interest

The authors state that there are no conflicts of interest in the publication of this research.

Author Contributions

R. Baghia Laxmi carried out the research, wrote the paper, design of the algorithm, simulated, and evaluated the results. B. Kirubagari and have been involved in revising the manuscript for important intellectual content. Lakshmana Pandian have contributed to the manuscript's creation and critical revision for key intellectual substance. Lakshmana Pandian have contributed to

proofreading and language editing. All authors approved the final version.

ACKNOWLEDGMENT

The author would like to thank the Department of Computer Science and Engineering, Annamalai University, India for supporting this work.

REFERENCES

[1] The American College of Obstetricians and Gynecologists Cervical Cancer Screening and Prevention, *ASCCP and the Society of Gynecologic Oncology Endorse this Practice Advisory*, Bulletin No. 168, October 2016, pp. 1–20.

[2] R. J. Kurman, M. L. Carcangiu, C. S. Herrington, and R. H. Young, *WHO Classification of Tumours of Female Reproductive Organs*, 4th ed. IARC; Lyon, France, pp. 172–184.

[3] J. Liu, Y. Peng, and Y. Zhang, "A fuzzy reasoning model for cervical intraepithelial neoplasia classification using temporal grayscale change and textures of cervical images during acetic acid tests," *IEEE Access*, vol. 7, pp. 13536–13545, 2019.

[4] P. J. Toliman, S. Phillipsm, S. D. Jong, *et al.*, "Evaluation of p16/Ki-67 dual-stain cytology performed on self-collected vaginal and clinician-collected cervical specimens for the detection of cervical pre-cancer," *Clinical Microbiol Infections*, vol. 26, no. 6, pp. 748–752, Jun 2020.

[5] H. Liang, M. Fu, J. Zhou, *et al.*, "Evaluation of 3D-CPA, HR-HPV, and TCT joint detection on cervical disease screening," *Oncol. Lett*., vol. 12, no. 2, pp. 887–892, 2016.

[6] R. Rivera and J. Madera, "Diagnostic accuracy of conventional cervical cytology (papanicolau smear), liquid based cytology and visual inspection with acetic acid in detecting premalignant and malignant cervical lesions among Filipino women in a tertiary hospital," *International Journal of Gynecologic Cancer*, vol. 29, p. A569, 2019.

[7] Y. Liu, L. Zhang, G. Zhao, *et al.*, "The clinical research of Thinprep Cytology Test (TCT) combined with HPV-DNA detection in screening cervical cancer," *Cell. Mol. Biol*, vol. 63, no. 2, pp. 92–95, 2017.

[8] K. Bryan, "Histopathological evaluation of colposcopic biopsies, LLETZ and cold knife cone biopsies of the uterine cervix in post-menopausal women: Considerations in the setting of the new Australian cervical HPV DNA screening program," *Pathology*, vol. 51, no. 7, pp. 752–755, 2019.

[9] G. Huang, Z. Liu, L. Van Der Maaten, *et al.*, "Densely connected convolutional networks," in *Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, H.I., USA, 2017, pp. 2261–2269.

[10] Y. Xue, Q. Zhou, J. Ye, *et al.*, "Synthetic augmentation and feature-based filtering for improved cervical histopathology image classification," in *Proc. Medical Image Computing and Computer Assisted Intervention — MICCAI 2019, Lecture Notes in Computer Science*, vol. 11764. Springer, 2019.

[11] P. Guo, K. Banerjee, R. J. Stanley, *et al.*, "Nuclei-based features for uterine cervical cancer histology image analysis with fusion-based classification," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 6, pp. 1595–1607, 2016.

[12] S. J. Keenan, J. Diamond, W. G. Mccluggage, *et al.*, "An automated machine vision system for the histological grading of cervical intraepithelial neoplasia (CIN)," *J. Pathology*, vol. 192, no. 3, pp. 351–362, 2020.

[13] Z. Alyafeai and L. Ghouti, "A fully-automated deep learning pipeline for cervical cancer classification," *Expert Systems with Applications*, vol. 141, 112951, 2020.

[14] M. Tubishat, M. Alswaitti, S. Mirjalili, *et al.*, "Dynamic butterfly optimization algorithm for feature selection," *IEEE Access*, vol. 8, pp. 194303–194314, 2020.

[15] H. Jia, Z. Xing, and W. Song, "A new hybrid seagull optimization algorithm for feature selection," *IEEE Access*, vol. 7, pp. 49614–49631, 2019.

[16] H. Jia, J. Li, W. Song, *et al.*, "Spotted hyena optimization algorithm with simulated annealing for feature selection," *IEEE Access*, vol. 7, pp. 71943–71962, 2019.

[17] Y. Zheng, Y. Li, G. Wang, *et al.*, "A novel hybrid algorithm for feature selection based on whale optimization algorithm," *IEEE Access*, vol. 7, pp. 14908–14923, 2019.

[18] Y. Wan, A. Ma, Y. Zhong, *et al.*, "Multiobjective hyperspectral feature selection based on discrete sine cosine algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3601–3618, May 2020.

[19] P. V. Mulmule, R. D. Kanphade, and D. M. Dhane, "Artificial intelligence-assisted cervical dysplasia detection using papanicolaou smear images," *Vis Comput*, pp. 1–12, 2022.