Journal of Theoretical and Applied Information Technology
31st December 2025. Vol.103. No.24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

AN HHO OPTIMIZED CNN VIT FRAMEWORK FOR CERVICAL CELL CLASSIFICATION

REMYA R[1,2*], KUMUDHA RAIMOND[1]

[1]Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India

[2]Federal Institute of Science and Technology, Ernakulam, Kerala, India

E-mail: *remyaraghavan@gmail.com

ABSTRACT

Accurate identification and classification of cervical cells is crucial for the early identification and treatment of cervical cancer. In image classification, the ability to capture both local texture and global contextual features is crucial for improving model performance. Convolutional Neural Networks (CNN) are effective in extracting local features using hierarchical filters whereas Vision Transformers (ViT) are good at capturing long range dependencies using self-attention mechanisms. This study presents a hybrid method that combines the features extracted by ResNet50 and ViT architectures to integrate their complementary strengths. The process incorporates Harris Hawks Optimization (HHO) to extract discriminative features and reduce their dimensionality. Finally, an SVM was used to classify the optimized features. The method was evaluated on two benchmarked datasets, SIPaKMeD and Mendeley, achieving impressive accuracies on both datasets. The results demonstrate that the hybrid method significantly enhances cervical cell classification and can support more reliable clinical decisions in cervical cancer detection.

Keywords: *Cervical cancer, ResNet 50, ViT, Harris Hawks Optimization (HHO), SVM*

1. INTRODUCTION

Cancer arises when cells in a particular organ grow improperly. Aberrant proliferation of cells may have an effect on the breast and cervix in women [1]. Cervical cancer is the 3rd most deadly form of cancer in developing nations [2]. Approximately 500,000 individuals with cervical carcinoma have been recorded globally in 2012, half of which were fatal [3]. It is projected that by 2030, there will be 400000 deaths a year, so these statistics appear to be only going up [4]. he procedure, known as "cervical cancer screening" involves examining the cervix for any aberrant tissues or cancerous cells [5].

Pap test is the most commonly used test to identify cervical cancer. Here the cells are observed under a microscope to identify abnormalities. Presence of debris, blood, and occluded and cut away cells etc. makes the diagnostic process difficult. Also, the presence of huge number of cells in a single smear makes the manual screening time consuming. Pap smear tests, vital for early cervical cancer detection, reduce deaths by 60%, yet manual cytology is costly, time-consuming, and error-prone. For the past two decades several research studies have been done in the field of computer aided diagnosis of medical images.

An early-stage cervical cancer diagnosis can be made rapidly with ML techniques, making them a promising option for system implementation [6]. Although there has been progress in using ML techniques, like ensemble models and sophisticated techniques for sampling to forecast cervical cancer, more research is still needed to fully evaluate how these techniques work together to improve accuracy, handle missing data, correct class imbalances, and extract complex features [7]. Furthermore, Deep Learning (DL) enriches the procedure of diagnosing cervical cancer by examining colposcopy and cervical cytology images by reducing human error and increasing accuracy. This development has enormous promise in the fight against cervical cancer in terms of early detection and, eventually, better patient outcomes [8].

DL can systematically extract additional discriminating features from pathological images to traditional manual feature extraction methods. Its advantages include a simplified extraction process and the ability to easily and systematically adjust the performance [9]. Multiple layers are used by DL models to obtain attributes at the general, intermediate, and high levels from images, thereby learning hierarchical features. This approach can lead to accurate image detection, segmentation, and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

classification. DL methods have been successfully used in several medical specialties, including radiology, pathology, and medical imaging, particularly in the identification and detection of cancer [10-12].

The work's primary contribution is

- The proposed approach integrates ResNet50 and ViT to extract robust and discriminative features from cervical cell images, effectively capturing both local and global patterns for improved diagnostic accuracy.
- Principal Component Analysis (PCA) is used to reduce the dimensionality of the combined features.
- The use of HHO ensures the selection of optimal features, reducing dimensionality and selecting non redundant features.
- The optimized features are classified using SVM, achieving reliable differentiation of various cervical cell types.

The article's remaining content is arranged as follows: The literature is reviewed in section 2, the proposed methodology which includes a block diagram and algorithm is elaborated in section 3, the findings of the proposed method are discussed in section 4, and the article is finally concluded in section 5.

2. LITERATURE SURVEY

In order to solve the issues associated with manual classification cervical cells many automated methods have been developed over the past two decades. Machine learning techniques mainly depends on manually extracted hand crafted features which were fed in to different classifiers. Development of deep learning models has substantially boosted the classification performance.

A Hybrid model which is making use of the CNN features that were fed in to Extreme learning machine classifier was proposed [13]. Performance of the model was good comparing with traditional handcrafted features. A Fisher Score CNN (FS-CNN)[14] was used to detect and categorize cervical cancer. After gathering the data from an open-source database, a central filter was used to eliminate the extraneous noise from the photos. The purpose of this research is to ensure early cervical cancer diagnosis and classification, thereby mitigating the problems associated with delayed diagnosis. It increases the weight of the FS-CNN algorithm to classify the types of abnormal cells as well as normal and abnormal cells. The suggested classification strategy is selected to maximize the use of CNN design parameters with the help of the Fisher score. The study demonstrated the value of DL methods for the early recognition and classification of images related to cervical cancer. One potential challenge in developing a diagnostic system for newly emerging cervical precancerous data is the lack of immediate evaluation on a wider range of databases and the reliance on reliable image processing technology and CNN algorithms

Alquran et al. developed an automated approach for classifying cervical cancer that combines deep learning (DL) and machine learning (ML) techniques [15]. Features extracted from a pre-trained Shuffle Net model combine with a new deep learning architecture called Cervical Net. Following preprocessing and augmentation, PCA is used to decrease the features from both networks, and canonical correlation analysis (CCA) is used to fuse them. Five machine learning classifiers are employed to evaluate the resulting fused feature set, and the support vector machine (SVM) achieves the highest accuracy.

An automated system using SMOTE features for effective management of missing values was proposed [16]. The system makes use of features with SMOTE up-sampled for managing missing values, KNN Imputer, and a stacked ensemble voting classifier model which incorporates three ML models. However, may pose challenges in terms of interpretability and computational efficiency.

In order to examine the possible relationship between the data in the whole slide image (WSI) and the prognosis for cervical cancer, a DL-based pathological risk score (RS) is being developed and validated to predict patients' prognosis [17]. A CNN with an autoencoder was used to extract high-dimensional pathological features to build a prognosis-associate RS. Overall survival (OS) and disease-free survival (DFS) forecasting effectiveness of RS in training and testing datasets, as well as different medical subdivisions, was confirmed using Kaplan–Meier survival analysis, using the score threshold chosen by X-tile. It created a useful signature for cervical cancer patients' OS and DFS stratification and prognosis prediction. The possible complexity and scalability problems that could occur from combining manual features in an attempt to improve interpretability, could result in higher computational overhead and make model maintenance more challenging.

"CerCan.Net," an effective CAD for cervical cancer recognition through automation was suggested [18].To lower the classification complexity, three lightweight CNNs with smaller factors and deep layers are employed by CerCan.Net. It then explores the effects of producing a reduced number of profound features that might distinguish discrete cervical cancer subtypes while adhering to an FS methodology. Recent DL models, noise, picture ambiguity, and optimization strategies for DL hyperparameters are not included in this study. Another method for cervical cell segmentation involves marker controlled watershed segmentation for cytoplasm and nuclei [19]. Feature extraction is done using Random Forest. Classification is done using bagging ensemble classification. A feature fusion based method was proposed [20] that utilizes CNN based and wavelet based features are fused and a correlation based feature selection is done. The classification is done using a Random Forest Classifier.

Fine grained cervical cell classification using morphology based CNN was proposed [21]. A five dimensional input consisting of the RGB patches with nuclei and cytoplasm segmentation mask is provided as the input. For high precision it is challenging to use diagnosis with deep learning based cervical cell classification.

Several ensemble based methods are also used for the categorisation of cervical cells. Afuzzy rank based ensemble classification using three deep learning claffifiers was performed for the classification of cervical cells[22]. Another ensemble model combining pretrained models DenseNet-169, VGG-19, and Xception and a Swin transformer model for the classification of cervical cell image [23]. The method obtained an accuracy of 95.50% on the SIPaKMeD dataset and 98.65% on the Mendeley LBC dataset. ViT based features are extracted from cervical images [24]. Particle Swarm Optimization (PSO) was used to select the features and then SVM used for the classification. Reinforcement learning approach is combined with ReNet50 architecture for the diagnosis of cervical images [25]. Extracted features from the CNN model are refined using reward functions .Evaluation of the model is done using publicly available datasets.

The development of CNN has improved the performance of the classification model. Ensemble learning can improve output, but it often demands increased resources and computing power. The combination of different algorithms introduces the risk of overfitting if a model operates effectively with the data used for training but finds it difficult to generalize to new, untested data. In the context of cervical precancerous data, challenges include limited immediate evaluation of diverse databases, reliance on reliable image processing technology and CNN algorithms, and potential issues related to interpretability and computational efficiency of the models. Incorporating manual features for interpretability may lead to complexity and scalability problems, resulting in higher computational overhead and challenging model maintenance.

The proposed hybrid framework effectively captures deep features and ViT features and employs optimization algorithm for the feature selection that results in a reduction in the number of features and computational cost. The optimal features are passed in to machine learning classifier for the accurate classification of cervical cell types.

3. PROPOSED METHODOLOGY

Cervical cancer, identified by abnormal cervical cells, remains a principal cause of fatal injuries in women, especially in developing countries, despite reduced rates in developed nations due to routine screening. Several deep learning based methods have been used for the classification of cervical cells. This study proposes a hybrid method that integrates deep learning for high level feature extraction, HHO for selecting discriminative features and machine learning for the final classification.

The proposed method begins with a preprocessing stage, where the cell images are resized and augmented to ensure consistency and generalization. These preprocessed images are then passed through a hybrid deep neural network combining ResNet-50 and ViT, powerful feature extraction models known for their ability to capture semantic and fine grained features. Feature reduction is performed using PCA. The reduced features are subsequently optimized using the HHO algorithm, which selects the most relevant and discriminative features to reduce dimensionality and improve model performance. Finally, the optimized feature set is fed into a SVM classifier, which categorizes various cervical cell types with high accuracy. This integrated pipeline effectively balances computational efficiency, model interpretability, and classification accuracy, addressing the limitations of traditional manual feature combination methods and resource-intensive ensemble learning. The schematic representation for the proposed hybrid approach is offered in Figure 1.

3.1 Preprocessing

For the improved generalization of the model the input images undergo a series of preprocessing steps before training. Initially image augmentation is

Journal of Theoretical and Applied Information Technology
31st December 2025. Vol.103. No.24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

performed by randomly flipping the images both horizontally and vertically as well as rotating by a specified angle θ which makes the model robust about orientation changes. Following augmentation all the images are resized to a spatial dimension of 224X224 to ensure consistent input size required for the subsequent deep learning architectures.

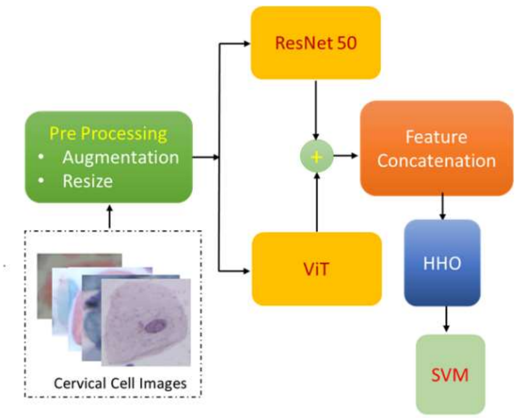

Figure 1: Block Diagram of the Hybrid Method

3.2 ResNet 50 Architecture

ResNet50 [26] is a deep convolutional neural network architecture based on the concept of residual learning. It consists of 50 layers with a series of convolutional identity blocks that incorporate skip connections. Figure 2 shows the architectural diagram of the ResNet50 model. The architecture begins with initial convolution layer followed by max pooling operation. This is followed by four sequential stages containing multiple residual blocks. Resnet 50 achieves strong performance in image recognition tasks while maintaining computational efficiency. In this work Resnet 50 is used to extract the deep hierarchical features from the preprocessed cervical cell images.

3.3 ViT Architecture

ViT is a deep learning architecture that applies the principle of transformer model to image classification tasks. Unlike CNN which rely on hierarchical feature extraction through convolutional layers, ViT processes image data using self attention mechanisms to model global relationship between image regions. Figure 3 illustrates the architecture of ViT.

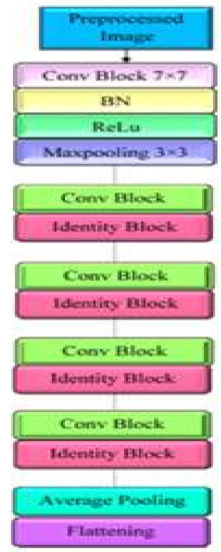

Figure 2: Overview of the ResNet50 Architecture

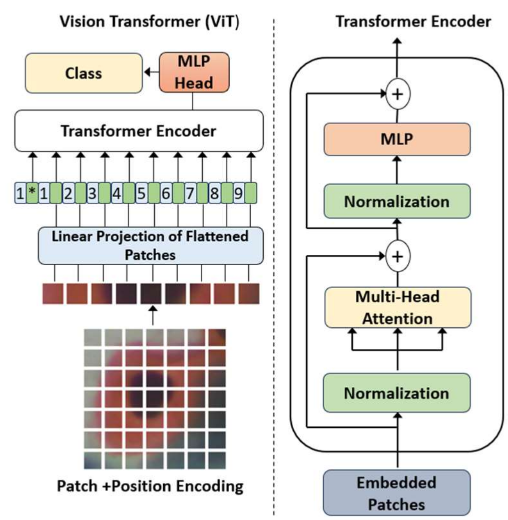

Figure 3: Architecture of the ViT model

The ViT model operates by dividing each input image $I \in \mathbb{R}^{H \times W \times C}$ in to fixed size non overlapping patches of PXP size [27]. Each patch is then flattened and linearly projected in to an embedding space. Positional encodings are added to these embeddings to retain spatial information. A learnable classification (CLS) token is prepended to the sequence which is then processed through multiple transformer encoder layers that include self attention and feed forward sublayers. The output from the CLS token in the last encoder layer serves as the global feature descriptor for the input image. In this

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

work we have used ViT Base 16 configuration which divides the image in to 16 X16 patches and maps them in to a 768 dimensional embedding space.

3.4 Harris Hawks Optimization (HHO)

HHO [28] is a recent metaheuristic algorithm that was created to resemble Harris Hawk behavior in natural environments, which employ distinctive pursuit strategies to apprehend their prey. It takes inspiration from swarm intelligence. HHO can be described as a population-based algorithm where a hawk pack cooperates while pursuing prey in different ways. The Harris Hawk has two main phases: the exploitative phase and the explorative phase. During the exploitative phase, hawks locate their target, make sudden jumps, and employ a variety of attack techniques. Any optimization problem can be implemented with the HHO algorithm.

3.4.1 Exploration Phase

Although the Harris Hawks often have trouble finding their prey, they have remarkable vision for tracking and locating it. As a result, the hawks find the location and search for prey. As a result, the hawks use two unofficial observational techniques to watch their prey after settling on a spot. Hawks perch where prey and other hawks are when $q < 0.5$; when $q \geq 0.5$, they perch at random on any largest tree. The probability or threshold parameter that determines hawks' perching behavior is represented by the symbol q. There are equivalent possibilities for every strategy. The Equation is given as,

$$A(i_t+1) = \begin{cases} A_{rand}(i_t) - r_1|A_{rand}(i_t) - 2r_2A(i_t)| \; q \geq 0.5 \\ \left(A_{prey}(i_t) - A_m(i_t)\right) - r_3\left(LB + r_4(UB - LB)\right), \; otherwise \end{cases} \quad (1)$$

where $A(i_t+1)$ indicates Hawks' position vector in the subsequent iteration i_t, $A_{prey}(i_t)$ indicates the position of prey, $A(i_t)$ signifies the hawks' current vector position, Each iteration updates the random numbers r_1, r_2, r_3, r_4 and q inside (0,1). LB and UB denote the lower and upper limits of the parameters. The current hawk population's average location is denoted by A_m, and a randomly chosen population's hawk is represented by $A_{rand}(i_t)$. Average location of the hawks is found using equation (2).

$$A_m(i_t) = \frac{1}{N}\sum_{i=1}^{N} A_i(i_t) \quad (2)$$

Where there are N hawks in total, and each hawk's location in each iteration i_t is indicated by $A_i(i_t)$. Although there are several methods to determine the average location, the simplest rule was used. HHO approach enables the prey to move from an exploring phase to an exploitation phase based on energy levels.

$$E = 2E_0\frac{1-i_t}{I_t} \quad (3)$$

According to equation (3) above, where I_t is the total number of iterations, i_t is the iteration of the show, and the prey's energy (E) decreases with each iteration, where E_0 indicates first energy of the prey.

3.4.2 Exploitation Phase

Prey usually finds it easy to get out of dangerous situations. The hawks employ diverse chasing strategies as a result. The Hawks plan states that four primary strategies are used during the exploitation stage. Let p be the likelihood that neither the prey may escape away from it ($p < 0.5$) nor not be able to ($p \geq 0.5$). Both soft and hard besieges are used to encircle the target. The location of the hawks' surroundings depends on the prey's energy (E). Hawks attack prey in groups to maximize the likelihood of capturing the target. Hawks intensify their attack to seize their prey as soon as they release energy. Soft besiege is used when $|E| \geq 0.5$, and hard besiege is used when $|E|<0.5$.

During Soft besiege the prey has enough energy to flee via a chance bounce when $|E| \geq 0.5$ and $p \geq 0.5$. The rules that follow serve as models for this behavior as given in equations (4) and (5).

$$A(i_t+1) = \Delta A(i_t) - E\left|JA_{prey}(i_t) - A(i_t)\right| \quad (4)$$

$$\Delta A(i_t) = A_{rand}(i_t) - A(i_t) \quad (5)$$

where $J = 2(1 - r_5)$ indicates the prey's random jump strength in the process of escaping, where the position vector of the prey and its present location in each iteration i_t are denoted by $\Delta A(i_t)$ and $r_5$is a random number within (0,1). To resemble the characteristics of prey motions, with every iteration, the J value fluctuates at random.

Simultaneously, the prey is surrounded by Harris Hawks in silence, wearing it down until they are ready to pounce. A limited selection of features is created from the original dataset, and the various prey features are replicated in the selected hawks.

During Hard besiege, the prey cannot run comfortably because it is at the lowest energy level when $|E| < 0.5$ and $p \geq 0.5$. Using equation (8), the current position of the Harris Hawk is improved.

$$A(i_t + 1) = A_{prey}(i_t) - E|\Delta A(i_t)| \quad (6)$$

The suggested model for the hard besiege replicates the prey's single characteristic for the current hawk.

During Progressively fast dives combined with soft besiege (SB) before the unexpected attack, the prey can run quickly and the SB can be carried out when $p < 0.5$ and $|E| \geq 0.5$. This prey sample is designed by the HHO algorithm using the levy distribution with outstanding perturbation. Choose the characteristics of the current hawk that are different by using the provided solution, keeping in mind the prey's energy level is given as equation (7)

$$B = A_{prey}(i_t) - E\left|JA_{prey}(i_t) - A(i_t)\right| \quad (7)$$

They then assess whether it will be a successful dive by comparing the potential outcome like a change to the previous dive. Additionally, it starts to dive quickly, erratically, and suddenly when the approach the prey based on the Levy Flight (LF)-based patterns by applying equation (10) if it was not reasonable:

$$C = B + S \times LF(D) \quad (8)$$

Where D is the dimension of the issue and LF is the levy flight function, which is found using Equation (9). A vector with size $1 \times D$ is called S.

$$LF(a) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \sigma = \left(\frac{\Gamma(1+\beta) \times \sin\frac{\pi\beta}{2}}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)}} \right)^{\frac{1}{\beta}} \quad (9)$$

where u and v are arbitrary values within (0,1) and $\beta$is a default constant set to 1.5. Thus, Eq. (12) can be used as the ultimate method for upgrading hawk positions during the soft besiege phase:

$$A(i_t + 1) = \begin{cases} B & if \ F(B) < F(A(i_t)) \\ C & if \ F(C) < F(A(i_t)) \end{cases} \quad (10)$$

where B and C are obtained using equations. (7) and (8) respectively.

Progressive rapid dives with hard besiege (HB is used when the prey cannot escape and $p < 0.5$ and $|E|$<0.5. The SB is connected to this scenario as a whole. Hawks get closer to their prey over time. Choose the various prey features which are duplicated to a hawk chosen randomly from the population based on the prey's energy status. The equation (11) is given as,

$$A(i_t + 1) = \begin{cases} B & if \ F(B) < F(A(i_t)) \\ C & if \ F(C) < F(A(i_t)) \end{cases} \quad (11)$$

where B and C are obtained using new equations. (12) and (13)

$$B = A_{prey}(i_t) - E\left|JA_{prey}(i_t) - A_m(i_t)\right| \quad (12)$$

$$C = B + S \times LF(D) \quad (13)$$

where $A_m(i_t)$ is obtained using Equation (2).

Algorithm 1: HHO algorithm

Input: Features Extracted from Pap smear images
Output: The best feature for evaluating the model's efficiency

Set the hawk population Ai (i=1,2,3, ⋯, N) to a starting point.
While(end)
After determining the new fitness value of the hawk, the ideal position was identified and marked as the prey location.
For (every hawk (Ai))
    E_0= 2 rand () − 1
    Using Equation (1), update E_0 and the prey's energy level.
    if ($|E| \geq 1$)
        Then (Execute the Exploration phase)
    if (($|E| < 1$)
        if ($|E| \geq 0.5 and\ p \geq 0.5$)
            Update the position using (6)
        Else if ($|E| < 0.5 and\ p \geq 0.5$)
            use Equation (6) to update the hawk's position
        Else if ($|E| \geq 0.5 and\ p < 0.5$)
            use Equation (7) to update the hawk's position.
        Else if ($|E| < 0.5 and\ p < 0.5$)
            use Equation (11) to update the hawk's position
        The updated fitness value of the hawks is determined.
.
Result ← Best feature subset
Return Result

Journal of Theoretical and Applied Information Technology
31st December 2025. Vol.103. No.24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4 RESULT AND DISCUSSION

This section demonstrates the results of the proposed hybrid method along with assessment metrics and performance comparison. The proposed hybrid method was executed using NVIDIA A100 GPU. The dataset is divided in to 70%,20%,10% for training, validation and testing respectively. For the HHO algorithm the population size was set to 30, and the algorithm was executed for 100 iterations. Two datasets the SIPaKMeD, and Mendeley datasets are used in the present research for assessment.

Individual cell images that have been cropped independently from 966 cluster cell images of Pap smear samples are included in the SIPaKMeD dataset [29]. The cells are divided into five groups by skilled cytopathologists. The two categories of normal cells are parabasal and superficial-intermediate cells; koilocytes and dyskeratotic cells are the two groups of aberrant but non-cancer cells; benign (metaplastic) cells make up the last group. Single Cell Images (SCI) were used separately for this investigation. Table 1 displays the distribution of images in the SIPaKMeD dataset collection. Figure 4 shows sample images from SIPaKMeD dataset.

Table 1: Distribution of images in SIPaKMeD dataset

Category	Class	No. of Images
Normal	Superficial - Intermediate	831
	Parabasal	787
Abnormal	Koilocytotic	825
	Dyskeratotic	813
Benign	Metaplastic	793
	Total	**4049**

Table 2: Class distribution in Mendeley dataset

Category	Class	No. of images
Abnormal	High Squamous Intraepithelial Lesion (HSIL)	113
	Low Squamous Intraepithelial Lesion (LSIL)	163
	Squamous Cell Carcinoma (SCC)	74
Normal	Negative for Intra epithelial Malignancy (NILM)	613
	Total	**963**

The 963-image Mendeley LBC dataset [30] was developed by the Obstetrics and Gynecology Department at Guwahati Medical College and Hospital. Four distinct groups comprise this data set. Details of the dataset distribution is given in Table 2. The liquid-based cytology technique was used to create the images from the cells of 460 patients. Figure 5 shows the sample images from the Mendeley LBC dataset.

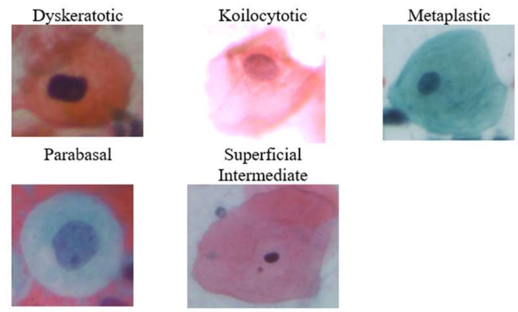

Figure 4: Sample from SIPaKMeD dataset

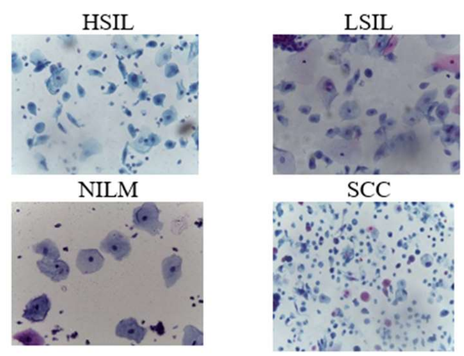

Figure 5: Sample from Mendeley dataset

4.1 Experimental Results

The input images were processed through the Resnet 50 model and resulting feature space of dimension 100353 was obtained. Parallelly from the ViT model a feature space of dimension 768 was obtained. Feature concatenation resulted in a total of 101121 features. PCA is applied to the concatenated features and the reduced features were passed to HHO algorithm for optimal feature selection. The optimal features selected from SIPaKMed dataset is 132 while for the Mendeley dataset 116 features were obtained. These features were subsequently used to train SVM classifier. Table 3 and 4 shows the result of the proposed methodology for both the datasets. Confusion matrix for the datasets is shown in Figure 6 and 7. To assess the classification performance of the model ROC curves were generated for each class as illustrated in Figure 8 and 9. These curves depict the relationship between the True Positive Rate against the False Positive Rate.

Journal of Theoretical and Applied Information Technology
31st December 2025. Vol.103. No.24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: Classification Performance of SIPaKMeD dataset

Model	Accuracy	Precision	Recall	F1 score
ResNet50+ SVM	94.04	94.04	94.04	94.04
ViT+SVM	96.19	96.20	96.16	96.18
ResNet50+ ViT+ SVM	97.11	97.14	97.11	97.11
ResNet50+ ViT+ PCA HHO+SVM	98.52	98.52	98.52	98.52

Table 4: Classification Performance of Mendeley dataset

Model	Accuracy	Precision	Recall	F1 score
ResNet50+ SVM	97.14	97.20	97.14	97.13
ViT + SVM	96.23	96.3	96.23	96.23
ResNet50+ ViT + SVM	97.79	97.80	97.79	97.79
ResNet50+ ViT+ PCA+ HHO+SVM	99.48	99.48	99.48	99.48

Table 5 t-test result of the proposed model

Dataset	Model	P-value	T-statistic
SIPaKMeD	ResNet50+ViT+ PCA+ HHO+SVM	0.00056520	-5.512736
Mendeley	ResNet50+ViT+ PCA+ HHO+SVM	0.00016649	-6.616566

The research problem is the effect of selecting optimized features from the ResNet50 and ViT features. The selection of optimized features extracted from ResNet50 and ViT significantly improves the classification performance of the proposed model compared to the base model without feature optimization. The statistical significance of this improvement can be validated through a paired t-test using 5-fold cross-validation results. It is hypothesized that the proposed model will yield higher performance metrics than the baseline for both datasets, as presented in Table 5.

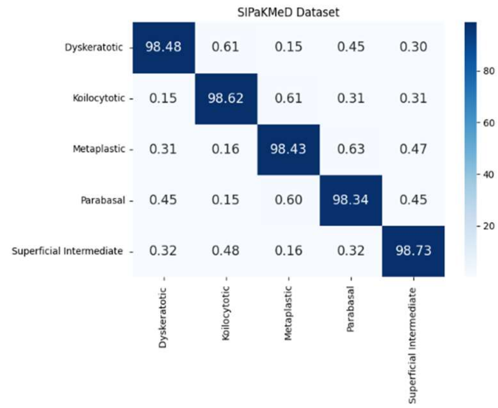

Figure 6: Confusion matrix for SIPaKMeD Dataset

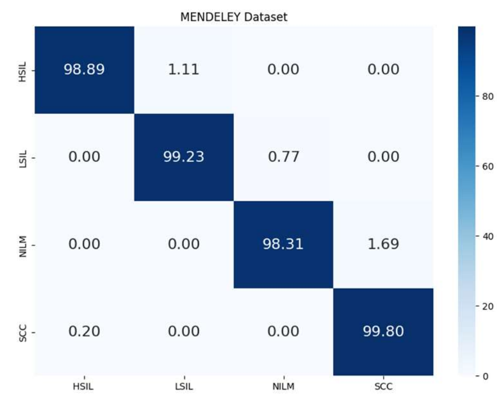

Figure 7: Confusion matrix for the Mendeley LBC Dataset

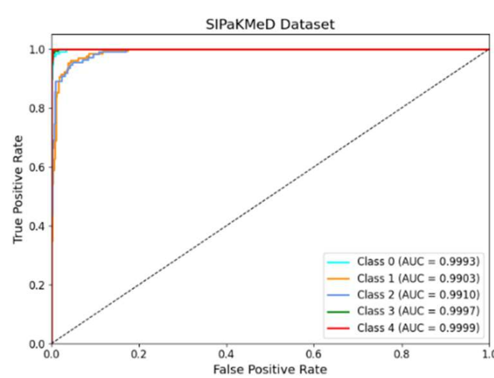

Figure 8: ROC Curve for SIPaKMeD Dataset

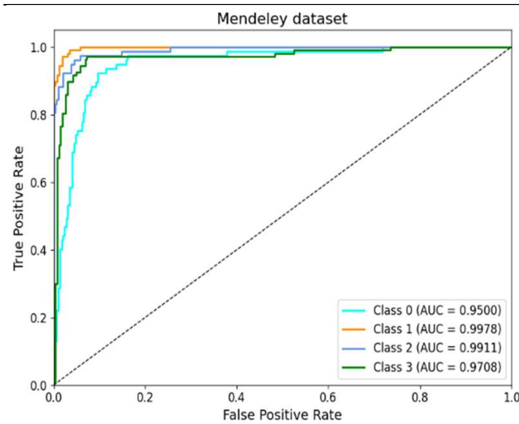

Figure 9: ROC Curve for Mendeley Dataset

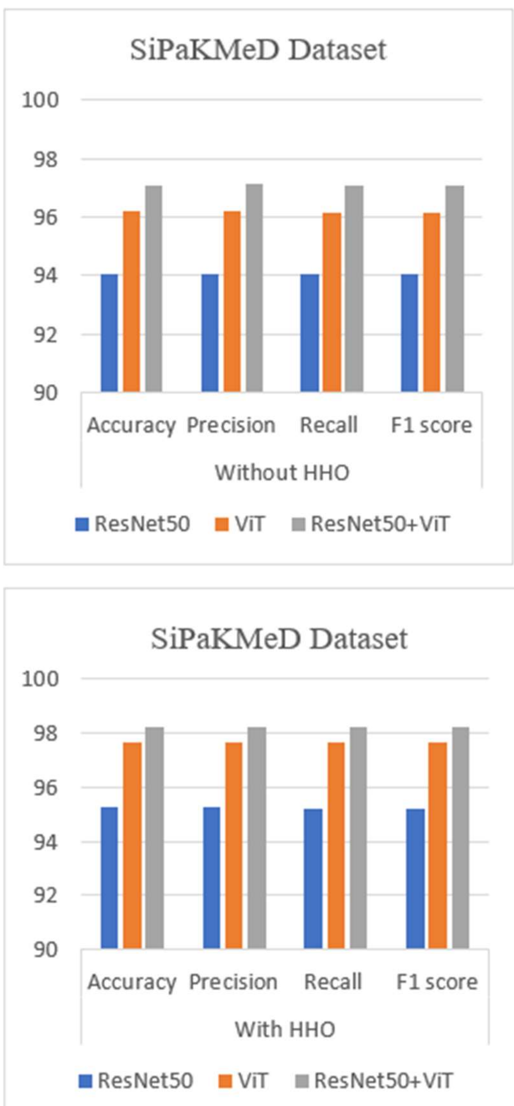

Figure 10: Results of SIPaKMed Dataset

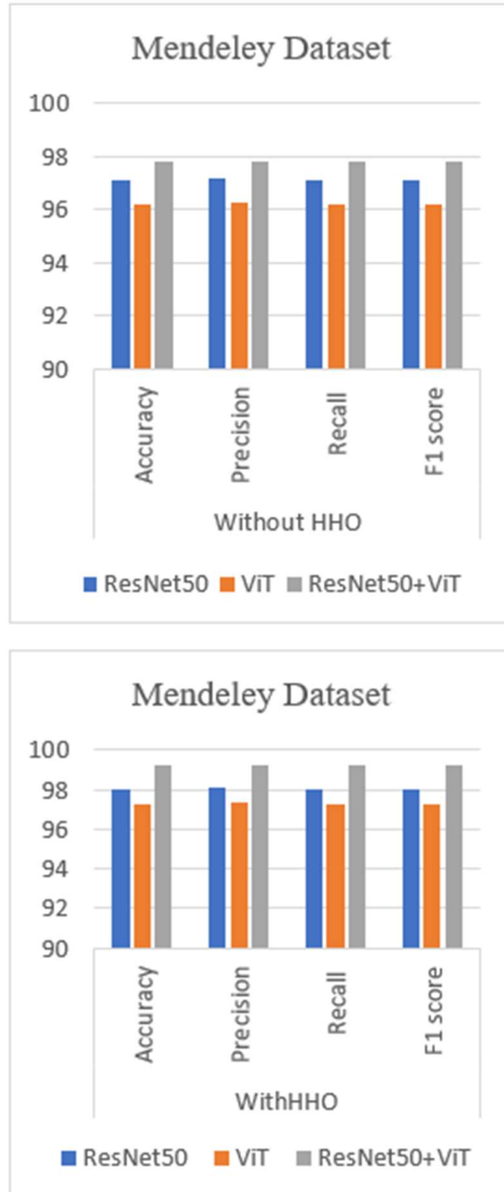

Figure 11: Results of Mendeley Dataset

Figure 10 and 11 demonstrates the result of the proposed methodology using SIPaKMeD and Mendeley datasets highlighting the performance of the model using ResNet50, ViT and their combined results. The results presented clearly demonstrates that the combination of ResNet50 and ViT features enhances the overall performance, while the application of HHO further optimizes and also boosts the model's performance. For the SIPaKMeD dataset obtained an accuracy of 98.52%, precision of 98.52%, recall of 98.52% and F1 score of 98.52 % for 5 class classification. For Mendeley dataset attained an accuracy of 99.48%, precision of 99.48%, recall of 99.48%, and F1-score of 99.48%.

4.2 Comparison with Existing Literature

A comparison with the existing methods for cervical cell classification is demonstrated in Table 6. In [19] segmented nuclei features are used by ensemble classifier and obtained an accuracy of 94.09%.CNN and wavelet based features selected are passed through RF classifier [20] and have obtained accuracy of 97.02% for 5 class classification of SIPaKMeD dataset and 98.12% for Mendeley 4 class classification. Two diffferent ensemble classfification methods [22-23] have obtained accuracy of 95.43% and 95.5% respectively for the 5 class classification of SIPaKMeD dataset. In [24] optimized ViT features are passed through SVM and has obtained an accuracy of 97.23% for 5 class classification of SIPaKMeD dataset. The proposed method has obtained a high accuracy of 98.52% for SIPaKMeD five class classification and an accuracy of 99.48% for 4 class classification of Mendeley LBC dataset.

Although the model has good performance compared to the existing methods, no clinical evaluation of the model has been performed in real world medical settings.

Table 6: Comparison of the model with existing methods

Dataset	Reference	Accuracy
SIPaKMeD	[19]	94.09%
	[22]	95.43%
	[20]	97.02%
	[23]	95.5%
	[24]	97.23%
	Proposed	98.52%
Mendeley	[20]	98.12%
	[23]	98.65%
	Proposed	99.48%

5 CONCLUSION

Cervical cancer is the fourth deadliest cancer among women whose morbidity rate can be reduced if detected in the early stage. This paper suggests an automated method for the classification of cervical cells. The deep features extracted from the images are optimized using HHO algorithm and classification is done using SVM classifier. The application of HHO has enhanced the performance of the classifier by optimizing feature selection and improving overall performance matrix. The suggested methodology has evaluated using two datasets namely SIPaKMeD, and Mendeley LBC and has achieved a high performance for both the datasets.

However, there is still chances for improvement of the performance by considering more enhanced models and more advanced feature selection methods. A comparative study of different optimization methods can be performed. Additionally future works should focus on evaluating model's performance in real time scenarios to assess its performance in practical applications, especially in resource constrained environments.

REFERENCES:

[1] Battula KP, Bolem SC, Harikiran J. A Comparative Review: Detection and Classification of Cervical Cancer from Pap-Smear Images. In2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) 2022 Dec 13 (pp. 932-938). IEEE.

[2] Alyafeai Z, Ghouti L. A fully-automated deep learning pipeline for cervical cancer classification. Expert Systems with Applications. 2020 Mar 1;141:112951.

[3] N. Malhotra, N. Taneja, R. Shankar, A. Pal, B. Chawla, A.A. Awasthi, and R. Janardhanan, "Knowledge, Attitude Practice Towards Cervical Cancer Screening Among Women Globally A Systematic Review", Journal of Women's Health and Development, Vol. 3, No. 3, 2020, pp. 194–210.

[4] Malhotra N, Taneja N, Shankar R, Pal A, Chawla B, Awasthi AA, Janardhanan R. Knowledge, attitude & practice towards cervical cancer screening among women globally–a systematic review. J Womens Health Dev. 2020 Jul;3(3):194-210.

[5] Shinde S, Kalbhor M, Shinde A. Cervical Cancer Screening Methods: Comprehensive Survey. Applied Artificial Intelligence. 2023 Sep 29:301-10.

[6] Al Qathrady M, Shaf A, Ali T, Farooq U, Rehman A, Alqhtani SM, Alshehri MS, Almakdi S, Irfan M, Rahman S, Eljak LA. A novel web framework for cervical cancer detection system: A machine learning breakthrough. IEEE Access. 2024 Mar 18;12:41542-56.

[7] Munshi RM. Novel ensemble learning approach with SVM-imputed ADASYN features for

enhanced cervical cancer prediction. Plos one. 2024 Jan 10;19(1): e0296107.

[8] Sarhangi HA, Beigifard D, Farmani E, Bolhasani H. Deep learning techniques for cervical cancer diagnosis based on pathology and colposcopy images. Informatics in Medicine Unlocked. 2024 Jan 1;47:101503.

[9] Cong L, Feng W , Yao Z, Zhou X, Xiao W. Deeplearning model as a new trend in computer-aided diagnosis of tumor pathology for lung cancer. Journal of Cancer. 2020 Mar 26;11(12):3615.

[10] R.C. Joshi, D. Singh, V. Tiwari, and M.K.Dutta “An Efficient Deep Neural Network Based Abnormality Detection and Multi Class Breast Tumor Classification”, Multimedia Tools and Applications, Vol. 81, No. 10, 2022, pp. 13691–13711.

[11] Pacal I, Karaman A, Karaboga D, Akay B, Basturk A, Nalbantoglu U, Coskun S. An efficient real-time colonic polyp detection with YOLO algorithms trained by using negative samples and large datasets. Computers in biology and medicine. 2022 Feb 1;141:105031.

[12] M. Souaidi and M. El Ansari, “A New Automated Polyp Detection Network MPFSSD in WCE and Colonoscopy Images Based Fusion Single Shot Multibox Detector and Transfer Learning”, IEEE Access, Vol. 10, 2022, pp. 47124–47140.

[13] Ghoneim A, Muhammad G, Hossain MS. Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems. 2020 Jan 1;102:643-9..

[14] Kumari CM, Bhavani R, Padmashree S, Priya R. Identification and classification of cervical cancer using convolutional neural network based on Fisher score. Journal of Data Acquisition and Processing. 2023;38(2):2118.

[15] Alquran H, Alsalatie M, Mustafa WA, Abdi RA, Ismail AR. Cervical net: A novel cervical cancer classification using feature fusion. Bioengineering. 2022 Oct 19;9(10):578.

[16] Karamti H, Alharthi R, Anizi AA, Alhebshi RM, Eshmawi AA, Alsubai S, Umer M. Improving prediction of cervical cancer using knn imputed smote features and multi-model ensemble learning approach. Cancers. 2023 Sep 4;15(17):4412.

[17] Chen C, Cao Y, Li W, Liu Z, Liu P, Tian X, Sun C, Wang W, Gao H, Kang S, Wang S. The pathological risk score: a new deep learning-based signature for predicting survival in cervical cancer. Cancer Medicine. 2023 Jan;12(2):1051-63.

[18] Attallah O. CerCan· Net: Cervical cancer classification model via multi-layer feature ensembles of lightweight CNNs and transfer learning. Expert Systems with Applications. 2023 Nov 1;229:120624.

[19] Win KP, Kitjaidure Y, Hamamoto K, Myo Aung T. Computer-assisted screening for cervical cancer using digital image processing of pap smear images. Applied Sciences. 2020 Mar 5;10(5):1800.

[20] Karim Kutty S, Chandrasekhara Menon G. Enhancing convolutional neural network model with spectral features for the identification of cervical dysplasia. International Journal of Imaging Systems and Technology. 2022 Nov;32(6):1916-27.

[21] Lin H, Hu Y, Chen S, Yao J, Zhang L. Fine-grained classification of cervical cells using morphological and appearance based convolutional neural networks. IEEE Access. 2019 May 27;7:71541-9.

[22] Manna A, Kundu R, Kaplun D, Sinitca A, Sarkar R. A fuzzy rank-based ensemble of CNN models for classification of cervical cytology. Scientific Reports. 2021 Jul 15;11(1):14538.

[23] Liu Y, Wang Z, Dai A, Gu W. A novel ensemble framework based on CNN models and Swin transformer for cervical cytology image classification. InSeventh International Conference on Computer Graphics and Virtuality (ICCGV 2024) 2024 May 13 (Vol. 13158, pp. 150-161). SPIE.

[24] AlMohimeed A, Shehata M, El-Rashidy N, Mostafa S, Samy Talaat A, Saleh H. ViT-PSO-SVM: cervical cancer predication based on integrating vision transformer with particle swarm optimization and support vector machine. Bioengineering. 2024 Jul 18;11(7):729.

[25] Muksimova S, Umirzakova S, Baltayev J, Cho YI. Rl-cervix. net: a hybrid lightweight model integrating reinforcement learning for cervical cell classification. Diagnostics. 2025 Feb 4;15(3):364.

[26] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770-778).

[27] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J. An image is worth 16x16 words:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929. 2020 Oct 22.

[28] Abbas S, Sampedro GA, Abisado M, Almadhor A, Yousaf I, Hong SP. Harris-hawk-optimization-based deep recurrent neural network for securing the internet of medical things. Electronics. 2023 Jun 9;12(12):2612.

[29] Plissiti ME, Dimitrakopoulos P, Sfikas G, Nikou C, Krikoni O, Charchanti A. Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images. In2018 25th IEEE international conference on image processing (ICIP) 2018 Oct 7 (pp. 3144-3148). IEEE.

[30] Hussain E, Mahanta LB, Borah H, Das CR. Liquid based-cytology Pap smear dataset for automated multi-class diagnosis of pre-cancerous and cervical cancer lesions. Data in brief. 2020 Jun 1;30:105589.