INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2026, VOL. 72, NO. 3, PP. 1-9
Manuscript received March 09, 2026; revised June, 2026. doi: 10.24425/ijet.2026.1724

Attention-enhanced DenseNet architecture for robust cervical cytology classification

Betelhem Zewdu Wubineh, Andrzej Rusiecki, and Krzysztof Halawa

***Abstract*—This study presents an attention-enhanced DenseNet201 framework for cervical cytology classification. Four attention mechanisms, SAM, CA, CBAM, and BAM, were integrated into the DenseNet201 backbone and evaluated with four optimizers (Adam, AdamW, RMSProp, and SGD). Experiments were conducted on the Mendeley and Pomeranian multicell dataset using percentage split and 5-fold cross-validation. In multi-class classification, CA with Adam achieved the highest accuracy of 98.75%, while several attention optimizer combinations reached 99.90% in binary classification. On the Pomeranian dataset, cross-validation accuracy reached 90.04%, confirming the model's generalization capability across datasets of different sizes and class distributions. These results demonstrate that attention-enhanced DenseNet architectures effectively emphasize diagnostically relevant regions, providing a reliable and robust approach for automated cervical cytology analysis.**

***Keywords*—Cervical cell classification; Attention mechanisms; Optimizers; Deep learning; DenseNet**

I. Introduction

CERVICAL cancer is one of the most serious health challenges affecting women, particularly in low and middle-income countries, where regular screening is limited [1]. Although the disease is largely preventable, delays in identifying cellular abnormalities remain a major obstacle to reducing incidence and mortality [2]. Therefore, early detection and treatment are critical for improving survival rates and mitigating the disease's impact [3] [4]. The standard approach for early detection is the Papanicolaou (Pap) smear examination [5]. A cytologist performs a procedure in which cervical cells are collected using a spatula and placed on a glass slide to be visually inspected under a microscope [6][7]. While this long-standing screening approach has significantly reduced cervical cancer mortality, it remains highly dependent on expert interpretation. Specialists must review large numbers of slides and distinguish subtle morphological variations between normal and abnormal epithelial cells, differences in nuclear size, contour, texture, or the nucleus-to-cytoplasm ratio [8]. This visual assessment is vulnerable to fatigue and subjectivity [9], especially when thousands of cells must be evaluated within a single smear. As a result, manual cytology can be both labor-intensive and prone to diagnostic inconsistencies [10] [11], underscoring the need for reliable computational tools capable of supporting or partially automating early detection.

In recent years, advances in deep learning have new possibilities for automating cytological analysis. Convolutional neural networks (CNNs) have revealed strong performance in image-based medical classification tasks by learning hierarchical visual patterns directly from data rather than relying on handcrafted features [12][13]. Among these architectures, DenseNet has proven especially effective due to its ability of connectivity pattern, which encourages feature reuse and efficient gradient propagation [1][14]. DenseNet was chosen for this study because its dense connectivity helps preserve subtle morphological cues and supports stable gradient flow, making it particularly suitable for distinguishing fine nuclear abnormalities in cervical cytology images. Such properties are valuable when extracting fine morphological structures from cervical cell images, where diagnostic indicates often appear subtle and dispersed.

Classical CNNs may struggle to focus on task-relevant regions because their feature maps contain irrelevant or noisy information, and attention modules help guide the network to focus on clinically meaningful regions [15]. By guiding the network to focus on informative spatial or channel-wise patterns, attention modules help strengthen discriminative cues while reducing the impact of irrelevant content [16]. Adding attention to DenseNet thus provides a simple yet effective way to improve classification reliability without significantly increasing computational cost. Several mechanisms have been widely explored in the literature, including the spatial attention module (SAM), which enables the model to focus on the most informative regions of the image [17]. The other one is channel attention (CA) adaptively weights feature maps using global pooling and a bottleneck layer to emphasize informative channels [18]. More composite modules, such as the convolutional block attention module (CBAM), which sequentially combine channel and spatial attention [19]. The final attention module is the bottleneck attention module (BAM), which applies channel and spatial attention in parallel, where channel attention adaptively weights feature maps via global pooling and a small neural network, and spatial attention captures broader contextual information using dilated convolutions, further enhancing feature representations [12]. Attention mechanisms have also shown effectiveness in other deep learning tasks, including gesture recognition and few-shot text classification, highlighting their broad applicability [20].

Despite progress in deep learning–based cervical cytology analysis, many existing approaches still struggle to consistently emphasize diagnostically relevant regions, particularly when abnormalities are subtle or when datasets differ in size and distribution. This creates a need for a more robust and

All authors are with Wroclaw University of Science and Technology, Poland (e-mails and orcid: betelhem.wubineh@pwr.edu.pl, https://orcid.org/0000-0002-4790-7449; andrzej.rusiecki@pwr.edu.pl, https://orcid.org/0000-0003-2239-1076; krzysztof.halawa@pwr.edu.pl, https://orcid.org/0000-0001-6508-0468).

generalizable framework capable of extracting and highlighting meaningful nuclear and cytoplasmic features across diverse conditions. Integrating attention mechanisms into a strong backbone, such as DenseNet, offers a promising solution to these limitations.

This study aims to develop and evaluate an attention-enhanced DenseNet architecture capable of delivering robust and accurate cervical cell classification by systematically incorporating and comparing multiple attention mechanisms to determine their effectiveness in improving diagnostic performance.

Despite progress in deep learning-based cervical cytology analysis, existing approaches rarely evaluate how different attention mechanisms influence diagnostic reliability and do not systematically assess robustness across datasets of varying sizes and class distributions. Most prior studies focus on a single attention module or a limited model configuration, leaving uncertainty about which strategy is most effective for highlighting diagnostically relevant regions. To address this gap, this study systematically integrates and compares four attention mechanisms, SAM, CA, CBAM, and BAM, within a DenseNet backbone, using multiple optimizers and two cervical cytology datasets with both percentage split and cross-validation settings. This comprehensive evaluation not only identifies the most stable and diagnostically reliable attention strategy but also provides methodological insight into designing robust, attention-enhanced architectures for cervical cell classification, representing a clear novelty over previous work.

II. MATERIALS AND METHODS

A. Dataset and its Preprocessing Technique

This study evaluates the proposed framework using the publicly available Mendeley dataset and a private dataset from Pomeranian Medical University. The Mendeley dataset consists of 963 multicell images divided into four groups: negative for intraepithelial lesion (Negative), low-grade intraepithelial lesions (LSIL), high-grade intraepithelial lesions (HSIL), and squamous cell carcinoma (SCC) [21]. For binary classification, the Negative class represents normal cells, while the other three categories are considered abnormal. The Pomeranian dataset was collected at Pomeranian Medical University, Szczecin, Poland, and comprises 262 images from three classes: HSIL, LSIL, and negative squamous intraepithelial lesion (NSIL) [22]. A summary of the datasets is presented in Table 1.

In the preprocessing stage, all images were resized to a size of 224 × 224 pixels to ensure uniform input dimensions. Additionally, pixel intensities were rescaled to a range of 0-1 to normalize the data and stabilize model training.

B. Proposed Architecture

The proposed classification framework is built upon a DenseNet201 backbone, selected for its strong feature-reuse capability and efficient gradient propagation [23]. In the implementation, the pretrained DenseNet201 is loaded with ImageNet weights, and its convolutional layers are frozen to preserve the generic visual features learned during large-scale training. To enhance the model's ability to emphasize diagnostically important regions, four different attention mechanisms were integrated and evaluated. The first is the SAM, which generates two spatial descriptors using average and max pooling across channels. These descriptors are concatenated and convolved with a 7×7 filter to create an attention map that highlights the most informative spatial regions in the feature map, such as the nucleus or its boundary structures. The second variant introduces CA, where global average and max pooling are applied across spatial dimensions to capture channel-wise statistics. These descriptors pass through a shared bottleneck fully connected layer before being re-expanded to the original number of channels, allowing the network to reweight the feature channels according to their diagnostic relevance. The third configuration integrates CBAM, in which the channel attention unit is followed by a spatial attention unit, enabling sequential refinement of feature representations along both channel and spatial dimensions. The final variant applies the BAM, which processes spatial and channel attention in parallel and fuses them multiplicatively with the backbone features. This module aims to suppress irrelevant background textures and emphasize diagnostically meaningful nuclear patterns while maintaining lightweight computational cost.

TABLE I
SUMMARY OF THE DATASET USED IN THE STUDY

Dataset	Class	No of image	Category	Total
Mendeley	HSIL	113	Abnormal	963
	LSIL	163		
	SCC	74		
	Negative	613	Normal	
Pomeranian	HSIL	70		262
	LSIL	35		
	NSIL	157		

After the attention module, the resulting feature tensor undergoes global average pooling and two dense layers with 512 and 256 neurons activated with Mish, followed by a 0.5 dropout regularization and a final softmax classifier. To evaluate how optimization algorithms influence training dynamics, each attention configuration was trained separately using four optimizers. Adam [24], an adaptive optimizer that combines momentum with per-parameter learning rates for faster convergence. AdamW (decoupled weight decay), which separates weight decay from gradient updates to improve generalization [25]. Root Mean Square Propagation (RMSProp), which normalizes gradients using a moving average of squared gradients to stabilize training and prevent oscillations during optimization [26]. Stochastic Gradient Descent (SGD) with momentum updates network weights using small, randomly selected batches of data rather than the entire dataset, which accelerates training and introduces beneficial noise that can help escape local minima [27]. All the optimizers are initialized with the same learning rate of 0.002. This systematic design enables a direct comparison of how spatial- vs. channel-focused attention interacts with different gradient update strategies. By embedding these four attention modules at the top of DenseNet201 and pairing them with multiple optimizers, the study examines the effect of targeted feature enhancement and training stability on the accurate recognition of cervical cell types. An attention-based classification of cervical cells is presented in Fig. 1.

Training was performed using an NVIDIA GeForce RTX 3090 GPU to support efficient model execution. The experiment was carried out using a Jupyter notebook with TensorFlow and

a Keras environment. The data was divided into 64% for training, 20% for testing, and 16% for validation, and 5-fold cross-validation was applied. The model was evaluated using precision, recall, accuracy, and F1 score, and a confusion matrix to provide a comprehensive understanding of classification effectiveness across all cell categories. All experiments were evaluated using 5-fold cross-validation. For each model, the mean and standard deviation of Accuracy, Precision, Recall, and F1-score across folds are reported to reflect both performance and stability. The training configuration is shown in Table II.

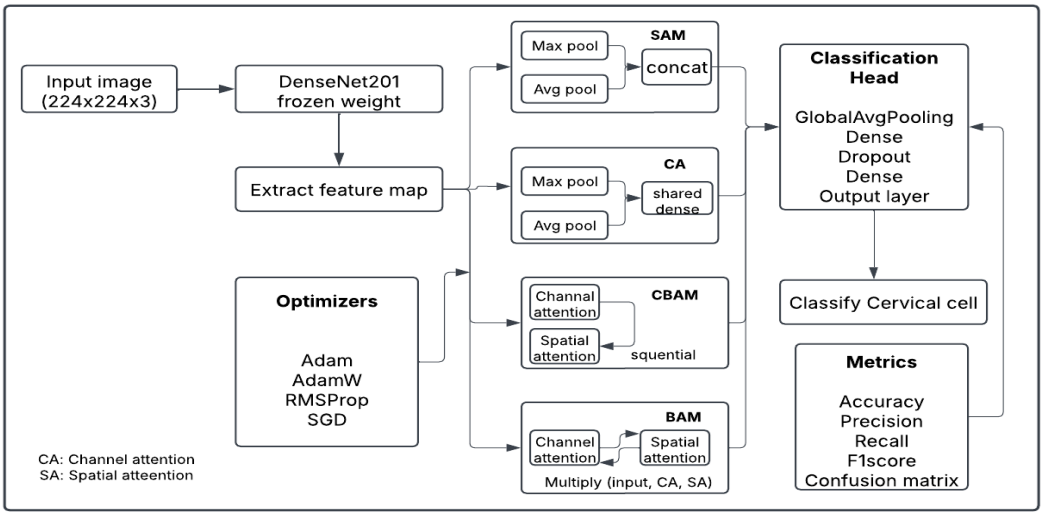

Fig. 1 Attention-based classification of cervical cells

TABLE II
TRAINING CONFIGURATION

Parameter	Value
Base model	DenseNet201 (pretrained on ImageNet, top layers frozen)
Attention modules	SAM, CA, CBAM, BAM
Optimizers	Adam, AdamW, RMSProp, SGD with momentum
Epochs	100
Batch size	16
Loss function	Categorical Cross-Entropy
Early stopping	Used, monitored by validation accuracy
Learning rate	0.002
Dropout	0.5

III. RESULTS AND DISCUSSION

The proposed attention-enhanced DenseNet201 models were evaluated on both the Mendeley and Pomeranian multicell cervical cell datasets. Table III summarizes the overall classification performance of the four attention mechanisms (SAM, CA, CBAM, BAM) paired with different optimizers on the Mendeley dataset. We conducted an ablation study to identify the best-performing combination for classifying cervical cells. All the results in Tables III, IV, V, and VIII were expressed in percentages, where Acc is accuracy, Pre is precision, Rec is recall, and F1 is the F1 score.

TABLE III
MENDELEY DATASET RESULTS BASED ON PERCENTAGE SPLIT

Model (Attention-Optimizer)	Mendeley percentage split multi-class				Percentage split binary class			
	Acc	Pre	Rec	F1	Acc	Pre	Rec	F1
SAM-Adam	97.93	97.90	97.93	97.89	97.93	97.93	97.93	97.93
SAM-AdamW	97.93	97.93	97.93	97.89	**99.48**	99.49	99.48	99.48
SAM-RMSProp	96.89	97.09	96.89	96.67	99.28	99.19	100	**99.60**
SAM-SGD	63.73	40.62	63.73	49.61	63.73	40.62	63.73	49.61
CA-Adam	**98.45**	98.47	98.45	98.41	98.45	98.45	98.45	98.44
CA-AdamW	**98.45**	98.47	98.45	98.41	98.96	98.98	98.96	98.96
CA-RMSProp	97.93	97.98	97.93	97.85	98.96	98.98	98.96	98.96
CA-SGD	63.73	40.62	63.73	49.61	63.73	40.62	63.73	49.61
CBAM-Adam	97.93	97.90	97.93	97.89	98.96	98.98	98.96	98.96
CBAM-AdamW	96.89	97.09	96.89	96.67	**99.48**	99.49	99.48	99.48
CBAM-RMSProp	97.93	97.98	97.93	97.85	**99.48**	99.49	99.48	99.48
CBAM-SGD	63.73	40.62	63.73	49.61	63.73	40.62	63.73	49.61
BAM-Adam	95.34	95.76	95.34	94.86	98.96	98.98	98.96	98.96
BAM-AdamW	94.82	94.78	94.82	94.38	98.96	98.98	98.96	98.96
BAM-RMSProp	90.16	85.07	90.16	87.24	98.45	98.45	98.45	98.44
BAM-SGD	63.73	40.62	63.73	49.61	63.73	40.62	63.73	49.61

Table III presents the results of the proposed attention-enhanced DenseNet201 models on the Mendeley dataset using the percentage split approach for both multi-class and binary classification. For multi-class classification, the CA module combined with Adam or AdamW optimizers achieved the highest accuracy of 98.45% and an F1-score of 98.41%. SAM

and CBAM modules also performed well with Adam or AdamW, reaching accuracies of 97.93%, whereas BAM achieved slightly lower performance with a maximum accuracy of 95.34%. The SGD optimizer consistently underperformed across all attention modules, with accuracies around 63-64%, highlighting its limited effectiveness. For binary classification, the SAM, CA, and CBAM modules paired with AdamW or RMSProp achieved near-perfect performance, with accuracy and F1-score ranging from 98.96% to 99.60%, while BAM performed slightly lower, and SGD again showed poor results. Results of the study using the 5-fold cross-validation are presented in Table IV.

Table IV summarizes the results obtained using 5-fold cross-validation on the Mendeley dataset. For multi-class classification, CA with Adam achieved the highest accuracy (98.75 ± 1.02%), while other combinations, such as CBAM-Adam (98.54 ± 0.51%), also performed well. For binary classification, several attention-optimizer combinations, including SAM-AdamW, CA-RMSProp, and BAM-AdamW, achieved near-perfect performance (up to 99.90 ± 0.21%) across all metrics, demonstrating the model's robust capability to distinguish normal from abnormal cervical cells. Overall, the results indicate that attention modules, particularly CA, combined with appropriate optimizers, significantly enhance classification performance. The results for the Pomeranian multiclass dataset are presented in Table V.

In Table V, the performance of the attention-enhanced DenseNet201 models is shown on the Pomeranian dataset using both percentage split and 5-fold cross-validation. In the percentage split setting, CA-Adam and CBAM-AdamW achieved the highest accuracy of 88.68%, followed by SAM-RMSProp, which was 84.91%accuracy. SGD again showed the weakest performance across all attention modules. Under 5-fold cross-validation, the highest accuracy was obtained by CA-Adam (90.04 ± 5.39%), followed closely by CBAM-RMSProp (89.67 ± 4.35%) and CA-AdamW (89.29 ± 4.18%), indicating that these configurations not only performed well on average but also maintained reasonable stability across folds. While BAM-based models showed moderate performance, SGD-based configurations consistently underperformed.

TABLE IV
MENDELEY DATASET RESULTS BASED ON 5-FOLD CROSS-VALIDATION

Model (Attention-Optimizer)	Mendeley 5-fold cross-validation multi-class				Mendeley 5-fold cross-validation binary-class			
	Acc (%)	Pre (%)	Rec (%)	F1 (%)	Acc (%)	Pre (%)	Rec (%)	F1 (%)
SAM-Adam	97.30 ± 0.83	97.48 ± 0.70	97.30 ± 0.83	97.19 ± 0.91	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**
SAM-AdamW	97.92 ± 0.66	98.10 ± 0.54	97.92 ± 0.66	97.87 ± 0.73	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**
SAM-RMSProp	97.61 ± 1.07	97.78 ± 0.97	97.61 ± 1.07	97.53 ± 1.13	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**
SAM-SGD	63.62 ± 3.82	40.62 ± 4.87	63.62 ± 3.82	49.54 ± 4.79	63.62 ± 2.52	40.54 ± 3.27	63.62 ± 2.52	49.50 ± 3.19
CA-Adam	**98.75 ± 1.02**	**98.80 ± 0.97**	**98.75 ± 1.02**	**98.75 ± 1.01**	99.79 ± 0.25	99.79 ± 0.25	99.79 ± 0.25	99.79 ± 0.25
CA-AdamW	98.65 ± 0.42	98.69 ± 0.44	98.65 ± 0.42	98.63 ± 0.44	99.79 ± 0.25	99.80 ± 0.25	99.79 ± 0.25	99.79 ± 0.25
CA-RMSProp	98.13 ± 0.71	98.24 ± 0.63	98.13 ± 0.71	98.07 ± 0.75	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**
CA-SGD	63.62 ± 3.82	40.62 ± 4.87	63.62 ± 3.82	49.54 ± 4.79	97.19 ± 2.20	97.36 ± 1.97	97.19 ± 2.20	97.21 ± 2.17
CBAM-Adam	98.54 ± 0.51	98.64 ± 0.44	98.54 ± 0.51	98.55 ± 0.48	99.79 ± 0.41	99.80 ± 0.40	99.79 ± 0.41	99.79 ± 0.41
CBAM-AdamW	98.55 ± 1.29	98.65 ± 1.16	98.55 ± 1.29	98.49 ± 1.38	99.79 ± 0.41	99.80 ± 0.40	99.79 ± 0.41	99.79 ± 0.41
CBAM-RMSProp	98.44 ± 0.47	98.50 ± 0.44	98.44 ± 0.47	98.41 ± 0.48	99.79 ± 0.25	99.79 ± 0.25	99.79 ± 0.25	99.79 ± 0.25
CBAM-SGD	63.62 ± 3.82	40.62 ± 4.87	63.62 ± 3.82	49.54 ± 4.79	63.62 ± 2.52	40.54 ± 3.27	63.62 ± 2.52	49.50 ± 3.19
BAM-Adam	95.74 ± 1.89	95.02 ± 4.18	95.74 ± 1.89	94.91 ± 3.02	99.79 ± 0.42	99.80 ± 0.41	99.79 ± 0.42	99.79 ± 0.42
BAM-AdamW	95.84 ± 1.14	96.07 ± 1.02	95.84 ± 1.14	95.74 ± 1.24	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**	**99.90 ± 0.21**
BAM-RMSProp	96.77 ± 2.27	96.78 ± 2.38	96.77 ± 2.27	96.58 ± 2.63	99.79 ± 0.42	99.79 ± 0.42	99.79 ± 0.42	99.79 ± 0.42
BAM-SGD	63.62 ± 3.82	40.62 ± 4.87	63.62 ± 3.82	49.54 ± 4.79	63.62 ± 2.52	40.54 ± 3.27	63.62 ± 2.52	49.50 ± 3.19

TABLE V
POMERANIAN DATASET RESULTS BASED ON PERCENTAGE SPLIT AND 5-FOLD CROSS-VALIDATION

Model (Attention-Optimizer)	Pomeranian percentage split multi-class				Pomeranian 5-fold cross-validation multi-class			
	Acc (%)	Pre (%)	Rec (%)	F1 (%)	Acc (%)	Pre (%)	Rec (%)	F1 (%)
SAM-Adam	83.02	78.97	83.02	80.17	86.61 ± 4.27	83.91 ± 6.70	86.61 ± 4.27	83.99 ± 5.35
SAM-AdamW	83.02	83.33	83.02	80.99	84.35 ± 4.06	79.11 ± 3.79	84.35 ± 4.06	80.69 ± 3.82
SAM-RMSProp	84.91	81.40	84.91	81.62	84.71 ± 5.32	80.92 ± 8.48	84.71 ± 5.32	81.71 ± 6.61
SAM-SGD	60.38	36.45	60.38	45.46	59.93 ± 5.79	36.26 ± 6.81	59.93 ± 5.79	45.09 ± 6.98
CA-Adam	**88.68**	89.73	88.68	88.95	**90.04 ± 5.39**	**89.70 ± 5.90**	**90.04 ± 5.39**	**89.73 ± 5.76**
CA-AdamW	86.79	86.39	86.79	86.41	89.29 ± 4.18	88.76 ± 4.25	89.29 ± 4.18	88.17 ± 4.50
CA-RMSProp	84.91	81.40	84.91	81.62	88.91 ± 3.57	89.06 ± 3.69	88.91 ± 3.57	88.65 ± 3.52
CA-SGD	60.38	36.45	60.38	45.46	59.93 ± 5.79	36.26 ± 6.81	59.93 ± 5.79	45.09 ± 6.98
CBAM-Adam	81.13	78.62	81.13	79.06	88.15 ± 5.25	85.41 ± 8.91	88.15 ± 5.25	86.06 ± 7.18
CBAM-AdamW	**88.68**	87.09	88.68	86.72	87.39 ± 2.67	86.17 ± 2.63	87.39 ± 2.67	86.23 ± 2.45
CBAM-RMSProp	81.13	78.21	81.13	78.98	89.67 ± 4.35	90.02 ± 4.71	89.67 ± 4.35	89.19 ± 4.48
CBAM-SGD	60.38	36.45	60.38	45.46	59.93 ± 5.79	36.26 ± 6.81	59.93 ± 5.79	45.09 ± 6.98
BAM-Adam	83.02	76.46	83.02	78.42	84.33 ± 4.34	75.14 ± 6.39	84.33 ± 4.34	79.02 ± 5.48
BAM-AdamW	84.91	74.28	84.91	79.14	84.71 ± 4.43	76.08 ± 6.37	84.71 ± 4.43	79.60 ± 5.46
BAM-RMSProp	83.02	75.24	83.02	78.05	85.09 ± 4.49	79.34 ± 6.55	85.09 ± 4.49	80.98 ± 5.47
BAM-SGD	60.38	36.45	60.38	45.46	59.93 ± 5.79	36.26 ± 6.81	59.93 ± 5.79	45.09 ± 6.98

To statistically verify whether the performance differences between the attention optimizer combinations were meaningful, we employed the non-parametric Friedman test, which is suitable for comparing multiple classifiers evaluated on the same dataset across repeated folds [28]. The test was conducted on the 5-fold cross-validation accuracy results. When the Friedman test yielded statistical significance ($p < 0.05$), we applied the Nemenyi post hoc test to analyze pairwise differences. This analysis ensured that the reported performance improvements reflect true model superiority rather than random variation. For both datasets, the Friedman test was performed on the four attention modules (SAM, CA, CBAM, BAM) using the Adam optimizer, because Adam consistently produced superior performance and thus provides a fair comparison framework across attention mechanisms.

For each fold, the models were ranked according to their performance, and the Friedman test statistic (χ^2) was computed using the standard formula.

$$\chi^2 = \frac{12}{Nk(k+1)}\sum R_i^2 - 3N(k+1) \quad (1)$$

where k is the number of models (4 in our model), N is the number of folds (in our case, 5), and R_i is the sum of the ranks of the i model. The χ^2 statistic follows a chi-square distribution with $k-1$ degrees of freedom, and the corresponding p-value was obtained from this distribution to determine whether significant differences existed among the compared models.

The four attention modules (SAM, CA, CBAM, BAM) were ranked in each fold based on their accuracy performance. The best-performing model in each fold received the lowest rank value, while the lowest-performing model received the highest rank. This procedure generated a ranking matrix across five folds, and for each attention module, the ranks were summed to obtain its total rank value R_i. The average rank of each model was then obtained by dividing R_i by the number of folds, which produced the following mean rank values for both datasets, and is presented in Table VI.

TABLE VI
FRIEDMAN RANKING RESULTS FOR BOTH DATASETS

Attention method	Mendeley dataset			Pomeranian dataset		
	Rank value	Rank*fold	$\sum R_i^2$	Rank value	Rank*fold	$\sum R_i^2$
SAM	3.0	3*5= 15	15^2 =225	2.4	2.4*5= 12	12^2 =144
CA	1.6	1.6*5= 8	8^2 =64	1.4	1.4*5= 7	7^2 =49
CBAM	1.8	1.8*5=9	9^2 =81	2.4	2.4*5=12	12^2 =144
BAM	3.6	3.6*5=18	18^2 =324	3.8	3.8*5=19	19^2 =361
Sum			694			698

Using these rank totals, the Friedman chi-square statistic was computed according to equation 1:

$$\chi^2 = \frac{12}{5*4(4+1)}\sum R_i^2 - 3*5(4+1) = \frac{12}{5*4(5)}\sum R_i^2 - 3(5)(5)$$

$$\chi^2 = \frac{12}{100}\sum R_i^2 - 75$$

For the Mendeley dataset, the sum of squared rank totals was $\sum R_i^2 = 694$, resulting in

$\chi 2$=0.12(694) -75 = 83.2 – 75 = **8.28** in software 8.38.

For the Pomeranian dataset, the sum of squared rank totals was $\sum R_i^2 = 698$, resulting in

$\chi 2$=0.12(698) -75 = 83.76 – 75 = **8.76** in software 6.51.

The difference is that we used the average ranks, which are rounded to 1 decimal, and the software used exact rank values from each fold, so its χ^2 is more accurate. This manual calculation was presented only to illustrate the computation procedure. The p-values were obtained by comparing the Friedman chi-square statistic to the chi-square distribution with k-1=3 degrees of freedom. For the Mendeley dataset, x^2 = 8.39 which exceeded the 0.05 critical threshold ($\chi^2_{0.05,3}$= 7.815, yielding p = 0.039 < 0.05; therefore, the performance differences were statistically significant. In contrast, for the Pomeranian dataset, x^2 = 6.51, which was below the same threshold, resulting in p = 0.089 > 0.05, indicating that the observed performance differences were not statistically significant.

When the Friedman test indicated statistical significance ($p < 0.05$), a Nemenyi post-hoc test was applied to perform pairwise comparisons between the models. The post-hoc analysis was conducted only for the Mendeley dataset, as it was the only case where the Friedman test yielded p< 0.05. The Nemenyi test evaluates whether the difference between the average ranks of two models exceeds the Critical Difference (CD), which is computed as

$$CD = q_\alpha\sqrt{\frac{k(k+1)}{6N}} \quad (2)$$

where $q_\alpha$is the critical value of the Studentized range distribution for a given significance level. If the absolute difference between two models' mean ranks is greater than the CD, the performance difference between them is considered statistically significant. Pairwise relationships were additionally summarized using Nemenyi p-value matrices and Critical Difference diagrams to visually highlight whether any model achieved statistically superior performance over others.

$CD = q_\alpha\sqrt{\frac{4(4+1)}{6(5)}} = \sqrt{\frac{4(5)}{6(5)}} = \sqrt{\frac{20}{30}} = \sqrt{0.6667} =$ 0.8165, *0.8165*2.569=* ***2.09***

The critical Studentized value was q0.05k=4=2.569. Therefore, multiplying this value by 0.8165 resulted in a Critical Difference (CD) of 2.09, which corresponded to approximately 2.10 in the software output. If the difference between two models' mean ranks exceeds the CD, the difference is statistically significant; otherwise, it is not. The Nemenyi post-hoc test p-values are presented in Table VII.

TABLE VII
NEMENYI POST-HOC P-VALUE MATRIX FOR PAIRWISE COMPARISONS OF ATTENTION MECHANISMS

	SAM	CA	CBAM	BAM
SAM	1.0000	0.3831	0.3831	0.8831
CA	0.3831	1.0000	1.0000	0.0919
CBAM	0.3831	1.0000	1.0000	0.0919
BAM	0.8831	0.0919	0.0919	1.0000

For example, the p-value for the comparison between SAM and BAM was 0.8831, which is greater than 0.05, indicating no statistically significant difference. Similarly, the p-value between CA and SAM was 0.3831, which is also greater than 0.05, confirming non-significance. In fact, none of the pairwise comparisons yielded p-values below 0.05, meaning that no attention mechanism demonstrated statistically superior performance over the others. The corresponding Critical Difference (CD) diagram for the Mendeley dataset is presented in Fig. 2.

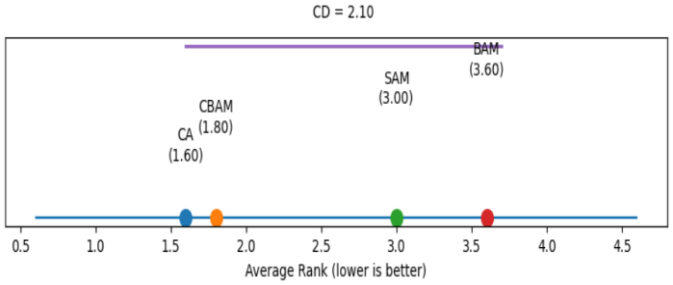

Fig. 2 Critical difference diagram using the Mendeley dataset

Fig. 2 illustrates the CD diagram obtained from the Nemenyi post-hoc analysis for the Mendeley dataset. The x-axis represents the average ranks of the four attention mechanisms, where a lower value indicates better performance. Among the methods, CA achieved the best mean rank (1.60), followed closely by CBAM (1.80), whereas SAM (3.00) and BAM (3.60) obtained comparatively higher ranks, indicating lower performance. The CD value was 2.10, meaning that two methods would be considered statistically different only if the distance between their average ranks exceeded this threshold. As shown in the diagram, none of the models are separated by a difference greater than the Critical Difference (CD) value. For example, the rank differences between BAM and CA (3.6 − 1.6 = 2.0), SAM and CA (3.0 − 1.6 = 1.4), and CBAM and CA (1.8 − 1.6 = 0.2) are all below the CD threshold. This indicates that although CA exhibits the best ranking trend, no attention mechanism demonstrates statistically significant superiority over the others. Overall, while the Friedman test indicates some variation among the modules, the pairwise comparisons suggest that these differences are not statistically significant, and no single attention module can be conclusively identified as superior based on the 5-fold evaluation.

The confusion matrix of CA in Adam for multiclass and binary classification is shown in Fig. 3. As shown in Fig. 3, the Mendeley dataset had three samples, and no misclassifications occurred in the multiclass and binary classification, respectively. In addition, on the Pomeranian dataset, twenty-six samples were misclassified on the mean accuracy of the 5-fold cross validations. The training and validation accuracy and loss for the SAM module with the Adam optimizer in the Mendeley multi-class percentage split are presented in Fig. 4.

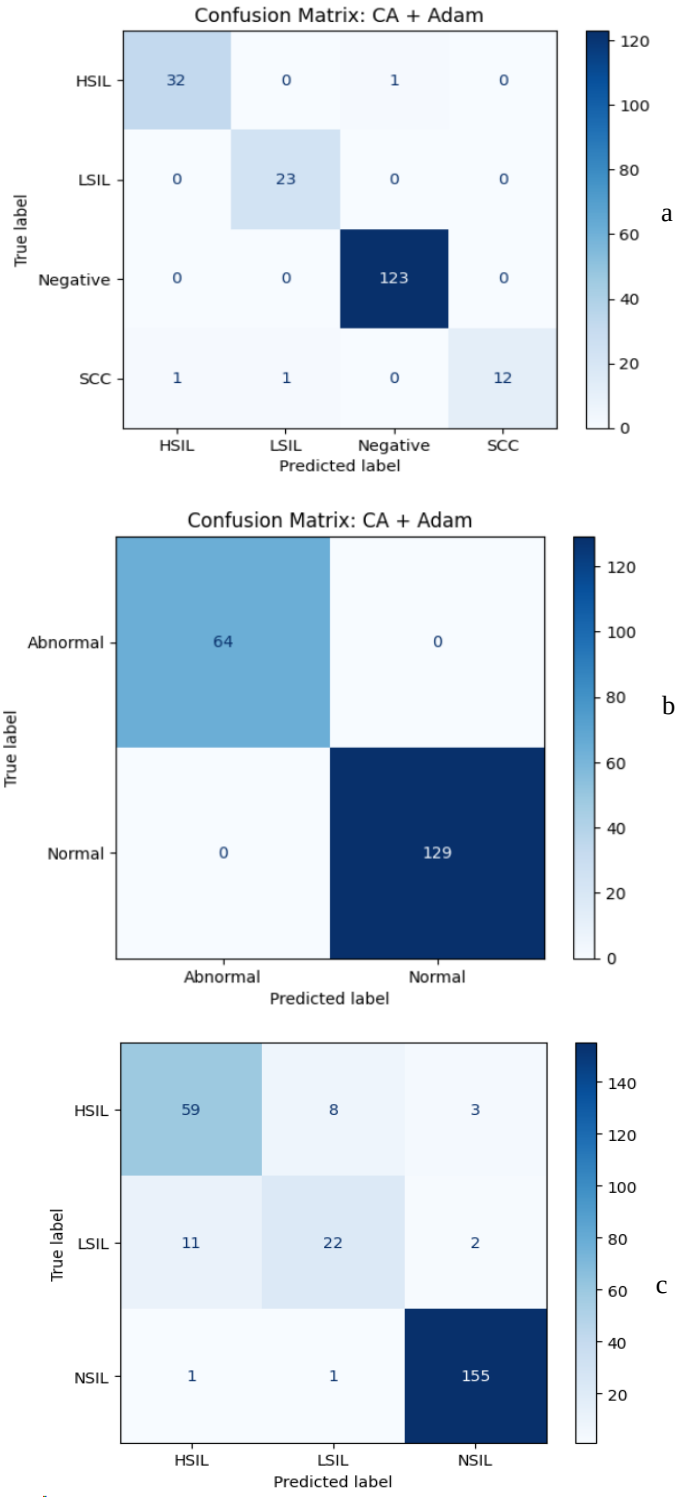

Fig. 3 Confusion matrix of CA with Adam optimizer: a) In percentage split of multiclass and b) 5-fold cross-validation of binary classification of Mendeley dataset, and c) 5-fold cross-validation of the Pomeranian dataset

Fig. 4 illustrates the training and validation accuracy and loss curves for the SAM module using the Adam optimizer over 18 epochs in the multi-class percentage split case. Both accuracy curves rise steadily and closely follow each other, reaching values near 1.0, indicating strong convergence and good generalization. Similarly, the training and validation loss curves decrease rapidly and stabilize at very low values, with only minor fluctuations. The close alignment between training and

validation curves across both metrics demonstrates stable learning with minimal overfitting.

Overall, the results indicate that the attention-enhanced DenseNet201 models perform robustly in cervical cell classification across both datasets.

Channel Attention (CA) consistently achieved the highest accuracy for multi-class classification, while SAM and CBAM provided strong alternative configurations. Binary classification was generally easier, with several attention-optimizer combinations achieving perfect performance on the Mendeley dataset. While perfect accuracy can sometimes indicate overfitting, the consistent cross-validation performance and closely aligned training-validation curves (Fig. 4) suggest that the model generalizes well. Moreover, the binary task distinguishes morphologically distinct normal and abnormal cells, which explains the observed high accuracy

The Pomeranian dataset, being smaller, produced slightly lower accuracies compared to the Mendeley dataset. The consistent underperformance of SGD highlights the importance of adaptive optimizers when training attention-augmented networks. These findings demonstrate that incorporating attention modules into DenseNet201 supports focused learning on diagnostically relevant regions, enhancing feature discriminability and enabling reliable automated cervical cytology screening.

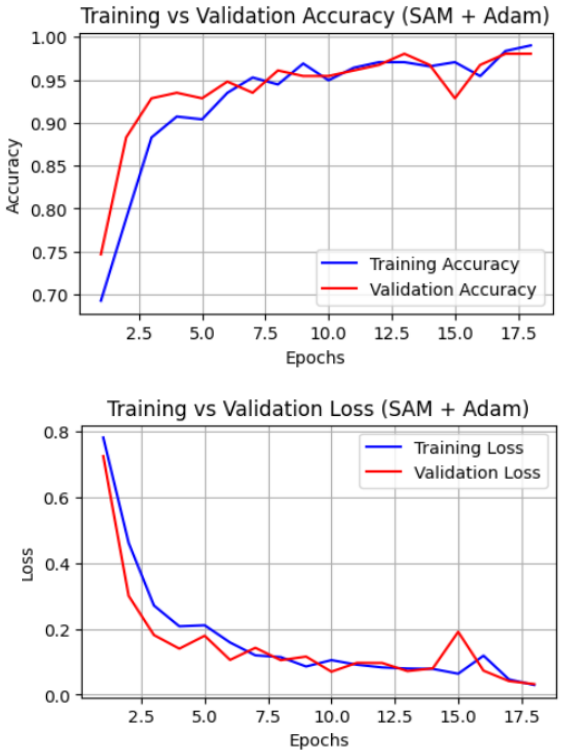

Fig. 4 Training and validation accuracy and loss using the Mendeley dataset

The comparison of our results in the Mendeley dataset from the previous work is presented in Table 8. As can be seen from the table, in [29] A CNN served as a feature extractor, followed by Linear Discriminant Analysis (LDA) for dimensionality reduction. A Support Vector Machine (SVM) was then employed for binary classification, achieving an accuracy of 97.94%. In [7] DenseNet201 was applied for multiclass classification and reached an accuracy of 98.3%. Other studies using DenseNet variants or ensembles, such as DenseNet201 alone [30] or DenseNet169 combined with MobileNetV2 [31], achieved accuracies ranging from 91.75% to 97.31%, while ensemble methods [32] achieved perfect performance (100%) in multiclass classification.

In comparison, our proposed model integrates channel attention (CA) into DenseNet201, allowing the network to emphasize diagnostically relevant regions of cervical cell images. For multiclass classification, our model achieved 98.75% accuracy, 98.80% precision, 98.75% recall, and 98.75% F1-score. For binary classification, it achieved 99.90% across all metrics. This demonstrates that incorporating attention mechanisms improves feature representation and classification performance, achieving results that are comparable or superior to previous approaches while providing enhanced interpretability of discriminative regions.

TABLE VIII
COMPARISON OF OUR RESULTS FROM PREVIOUS WORK

Ref	Dataset	Methods	Classification type	Acc	Pre	Rec	F1
[30]	Mendeley	DenseNet201	Multiclass	97.31	-	-	-
[31]	Mendeley	DenseNet169 + MobileNetV2	Multiclass	97.12	-	-	-
		DenseNet201	Multiclass	91.75	-	-	-
[29]	Mendeley	CNN-LDA+SVM	Binary	97.94	-	-	-
[7]	Mendeley	DenseNet201	Multiclass	98.3	-	-	-
[32]	Mendeley	Ensemble	Multiclass	100	100	100	100
Our	Mendeley	CA with DenseNet201	Multiclass	98.75	98.80	98.75	98.75
Our	Mendeley	CA with DenseNet201	Binary	99.90	99.90	99.90	99.90

The comparison of the Mendeley dataset across different attention mechanisms and optimizers is presented in Fig. 5.

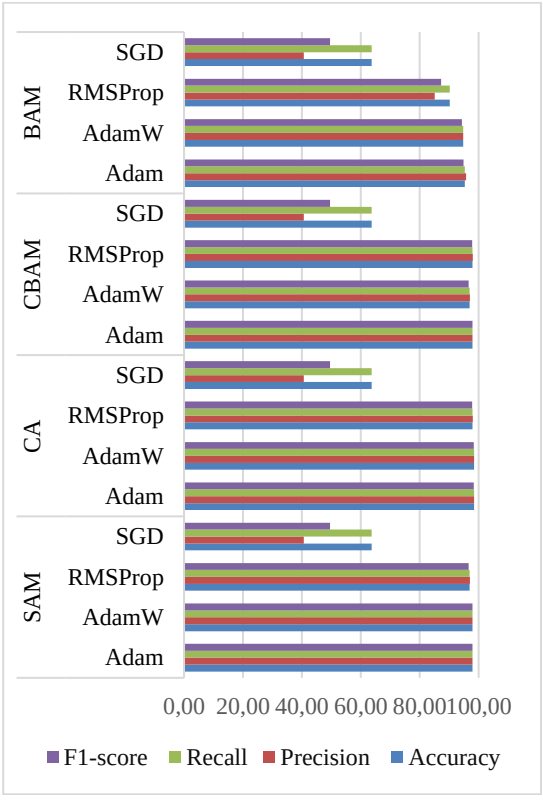

Fig. 5 Comparison of the attention-optimizer module on the Mendeley dataset

Although the proposed model achieved high performance, particularly in binary classification, the dataset size is relatively limited. Therefore, further validation on larger and independent external datasets is necessary before clinical deployment to ensure generalizability. Additionally, the study focused on two datasets and specific attention mechanisms, which may not capture all variability in cervical cytology images or generalize to different populations or imaging conditions. Future work should explore more diverse datasets and investigate additional architectures to strengthen robustness and clinical applicability

IV. CONCLUSIONS

This study evaluated four attention mechanisms integrated into DenseNet201 to enhance cervical cell classification. Channel Attention (CA) consistently delivered the best multi-class performance, while SAM and CBAM also showed strong results. Binary classification on the Mendeley dataset achieved perfect performance with multiple attention–optimizer combinations. Testing on the Pomeranian dataset confirmed the model's ability to generalize to a different dataset, although multi-class accuracy was slightly lower (maximum 90.57%), likely due to the smaller number of images and limited class samples. Training curves indicated stable convergence and minimal overfitting. These findings demonstrate that attention modules enable DenseNet201 to focus on diagnostically relevant regions, enhancing feature discrimination and overall classification accuracy. The proposed framework offers a reliable and effective tool for automated cervical cytology analysis, especially on the Mendeley dataset. Future work will include validation on larger, more diverse datasets and exploration of integration into clinical screening workflows.

REFERENCES

[1] Y. Zhang, C. Ning, and W. Yang, "An automatic cervical cell classification model based on improved DenseNet121," *Sci. Rep.*, vol. 15, no. 1, pp. 1–18, 2025, doi: 10.1038/s41598-025-87953-1.

[2] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Segmentation of Cytology Images to Detect Cervical Cancer Using Deep Learning Techniques," in *International Conference on Computational Science*, Cham: Springer Nature Switzerland, 2024, pp. 270–278.

[3] M. Fang, M. Fu, B. Liao, X. Lei, and F. X. Wu, "Deep integrated fusion of local and global features for cervical cell classification," *Comput. Biol. Med.*, vol. 171, no. February, p. 108153, 2024, doi: 10.1016/j.compbiomed.2024.108153.

[4] B. Z. Wubineh and A. Rusiecki, "SE-DeepLabV3 +: Cervical Cell Segmentation and Classification Using a Novel SE-Based DeepLabV3 + and Ensemble Method," *IEEE Access*, vol. 13, no. July, pp. 116430–116441, 2025, doi: 10.1109/ACCESS.2025.3586764.

[5] A. S. Shaik, S. M. Aswatha, and R. J. Pandya, "Deep Learning Enabled Segmentation, Classification and Risk Assessment of Cervical Cancer," 2025, [Online]. Available: http://arxiv.org/abs/2505.15505

[6] N. A. Alias *et al.*, "Image Dataset for Cervical Cell Diagnosis - a Review," *6th Iraqi Int. Conf. Eng. Technol. its Appl. IICETA 2023*, pp. 480–484, 2023, doi: 10.1109/IICETA57613.2023.10351193.

[7] O. Yaman and T. Tuncer, "Exemplar pyramid deep feature extraction based cervical cancer image classification model using pap-smear images," *Biomed. Signal Process. Control*, vol. 73, no. October 2021, p. 103428, 2022, doi: 10.1016/j.bspc.2021.103428.

[8] N. Gnanavel, P. Inparaj, N. Sritharan, D. Meedeniya, and P. Yogarajah, "Interpretable Cervical Cell Classification: A Comparative Analysis," *ICARC 2024 - 4th Int. Conf. Adv. Res. Comput. Smart Innov. Trends Next Gener. Comput. Technol.*, pp. 7–12, 2024, doi: 10.1109/ICARC61713.2024.10499737.

[9] B. Z. Wubineh and A. Rusiecki, "SPP-SegNet and SE-DenseNet201 : A Dual-Model Approach for Cervical Cell Segmentation and Classification," 2025.

[10] D. Kupas, A. Hajdu, I. Kovacs, Z. Hargitai, Z. Szombathy, and B. Harangi, "Annotated Pap cell images and smear slices for cell classification," *Sci. Data*, vol. 11, no. 1, pp. 1–8, 2024, doi: 10.1038/s41597-024-03596-3.

[11] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Classification of cervical cells from the Pap smear image using the RES_DCGAN data augmentation and ResNet50V2 with self-attention architecture," *Neural Comput. Appl.*, vol. 0123456789, 2024, doi: 10.1007/s00521-024-10404-x.

[12] T. Bashir, H. Wang, M. Tahir, and Y. Zhang, "Wind and solar power forecasting based on hybrid CNN-ABiLSTM, CNN-transformer-MLP models," *Renew. Energy*, vol. 239, no. August 2024, p. 122055, 2025, doi: 10.1016/j.renene.2024.122055.

[13] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Segmentation and Classification Techniques for Pap smear Images in Detecting Cervical Cancer: A Systematic Review," *IEEE Access*, vol. 12, no. August, pp. 118195–118213, 2024, doi: 10.1109/ACCESS.2024.3447887.

[14] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Enhancing Cervical Cell Classification with DenseNet201 and Spatial Attention Mechanism," in *International Conference on Dependability of Computer Systems*, 2025, pp. 229–236.

[15] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas," no. Midl, 2018, [Online]. Available: http://arxiv.org/abs/1804.03999

[16] P. Ghosal, A. Roy, R. Agarwal, K. Purkayastha, A. L. Sharma, and A. Kumar, "Compound attention embedded dual channel encoder-decoder for ms lesion segmentation from brain MRI," *Multimed. Tools Appl.*, vol. 84, no. 26, pp. 31139–31171, 2025, doi: 10.1007/s11042-024-20416-3.

[17] C. M. Mylonakis, P. Velanas, P. I. Lazaridis, P. Sarigiannidis, S. K. Goudos, and Z. D. Zaharis, "Deep Learning Framework Using Spatial Attention Mechanisms for Adaptable Angle Estimation Across Diverse Array Configurations †," *Technologies*, vol. 13, no. 2, pp. 1–21, 2025, doi: 10.3390/technologies13020046.

[18] S. R. Klomp, R. G. J. Wijnhoven, and P. H. N. de With, "Performance-Efficiency Comparisons of Channel Attention Modules for ResNets," *Neural Process. Lett.*, vol. 55, no. 5, pp. 6797–6813, 2023, doi: 10.1007/s11063-023-11161-z.

[19] H. Cui, J. Liang, C. Li, and X. Tian, "Improved Convolutional Neural Network with Attention Mechanisms for River Extraction," *Water (Switzerland)*, vol. 17, no. 12, pp. 1–15, 2025, doi: 10.3390/w17121762.

[20] Y. Wu, Y. Lin, T. Xu, T. Kang, X. Meng, and H. Liu, "Multi-Scale Feature Integration and Spatial Attention for Accurate Lesion Segmentation," *2025 6th Int. Conf. Electron. Commun. Artif. Intell.*, pp. 736–740, 2025, doi: 10.1109/icecai66283.2025.11170575.

[21] E. Hussain, L. B. Mahanta, H. Borah, and C. R. Das, "Liquid based-cytology Pap smear dataset for automated multi-class diagnosis of pre-cancerous and cervical cancer lesions," *Data Br.*, vol. 30, 2020, doi: 10.1016/j.dib.2020.105589.

[22] Ł. Jeleń, I. Stankiewicz-Antosz, M. Chosia, and M. Jeleń, "Optimizing Cervical Cancer Diagnosis with Feature Selection and Deep Learning," *Appl. Sci.*, vol. 15, no. 3, pp. 1–21, 2025, doi: 10.3390/app15031458.

[23] Y. Hou, Z. Wu, X. Cai, and T. Zhu, "The application of improved densenet algorithm in accurate image recognition," *Sci. Rep.*, vol. 14, no. 1, pp. 1–14, 2024, doi: 10.1038/s41598-024-58421-z.

[24] Z. Zhang, "Improved adam optimizer for deep neural networks," in *IEEE/ACM 26th international symposium on quality of service (IWQoS)*, 2018, pp. 1–2.

[25] S. Xie and Z. Li, "Implicit Bias of AdamW: ℓ∞-Norm Constrained Optimization," *Proc. Mach. Learn. Res.*, vol. 235, pp. 54488–54510, 2024.

[26] F. A. Ahda, A. P. Wibawa, D. D. Prasetya, and D. A. Sulistyo, "Comparison of Adam Optimization and RMSprop in Minangkabau-Indonesian Bidirectional Translation with Neural Machine Translation," *Int. J. Informatics Vis.*, vol. 8, no. 1, pp. 231–238, 2024, doi:10.62527/joiv.8.1.1818.

[27] C. Venkatachalam and P. Shah, "A Dual-Branch Deep Learning Framework Combining Xception and ResNet for Accurate Lung and Colon Cancer Detection," *IEEE Access*, vol. 13, no. July, pp. 131483–131497, 2025, doi: 10.1109/ACCESS.2025.3589390.

[28] J. Demˇ, "Statistical Comparisons of Classifiers over Multiple Data Sets," vol. 7, pp. 1–30, 2006.

[29] M. Sholik, C. Fatichah, and B. Amaliah, "Classification of Cervical Cell Images into Healthy or Cancer Using Convolution Neural Network and Linear Discriminant Analysis," *Proc. 2023 IEEE Int. Conf. Ind. 4.0, Artif. Intell. Commun. Technol. IAICT 2023*, pp. 383–389, 2023, doi: 10.1109/IAICT59002.2023.10205826.

[30] T. Sood, Padmavati, and R. Bhatia, "Enhancing pap smear image classification: integrating transfer learning and attention mechanisms for improved detection of cervical abnormalities," *Biomed. Phys. Eng. Express*, vol. 10, no. 6, 2024, doi: 10.1088/2057-1976/ad7bc0.

[31] O. Bilal, S. Asif, M. Zhao, Y. Li, F. Tang, and Y. Zhu, "Differential evolution optimization based ensemble framework for accurate cervical cancer diagnosis," *Appl. Soft Comput.*, vol. 167, no. PB, p. 112366, 2024, doi: 10.1016/j.asoc.2024.112366.

[32] O. Attallah, "CerCan·Net: Cervical cancer classification model via multi-layer feature ensembles of lightweight CNNs and transfer learning," *Expert Syst. Appl.*, vol. 229, no. PB, p. 120624, 2023, doi: 10.1016/j.eswa.2023.120624.