Uncertainty-Aware Multi-Task Deep Learning Framework for Cervical Lesion Segmentation and Classification

Betelhem Zewdu Wubineh [1,2][0000-0002-4790-7449] , Andrzej Rusiecki [1][0000-0003-2239-1076] and Krzysztof Halawa [1][0000-0003-2239-1076]

[1] Wroclaw University of Science and Technology, Faculty of Information and Communication Technology, Wroclaw, Poland
[2] College of Engineering and Technology, Wachemo University, Hossana, Ethiopia

betelhem.wubineh@pwr.edu.pl, andrzej.rusiecki@pwr.edu.pl, and krzysztof.halawa@pwr.edu.pl

Abstract. Accurate detection and classification of cervical lesions in Pap smear images is essential for effective cervical cancer screening. Conventional deep learning methods often treat segmentation and classification as separate tasks and tend to produce overconfident predictions. We propose a unified uncertainty-aware multi-task deep learning framework that performs simultaneous cervical lesion segmentation and classification. The architecture combines a pretrained ResNet50 encoder with a U-Net-style decoder for pixel-level segmentation and an attention-guided evidential classification head for uncertainty-aware diagnosis. Dirichlet-based evidential learning provides predictive uncertainty, while a homoscedastic uncertainty-based loss dynamically balances the two tasks. Evaluation on Pomeranian, Herlev, and SIPaKMeD datasets demonstrates strong performance, with segmentation accuracy of 99.19, 87.84%, and 93.85%, respectively, and classification accuracy of 99.14% in binary and 97.04% in multiclass for SIPaKMeD. Uncertainty correlates negatively with segmentation quality, and low Expected Calibration Error (ECE) indicates well-calibrated predictions. These results suggest that the proposed framework provides both accurate and reliable predictions, making it suitable for clinical decision support.

Keywords: Multi-task, Cervical cancer, segmentation, Evidential deep learning, Uncertainty, classification.

1 Introduction

Cervical cancer remains one of the leading causes of cancer-related mortality among women worldwide [1]. According to global health reports, early detection through routine screening significantly reduces both incidence and mortality rates. The Papanicolaou (Pap) smear test is widely used to detect abnormal cellular changes before they

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

progress to invasive cancer [2][3]. However, manual microscopic examination of Pap smear images is time-consuming, labor-intensive, and subject to inter-observer variability [4][5][6]. Subtle morphological differences between lesion grades (cervical cell types) increase diagnostic complexity.

To assist pathologists, computer-aided diagnosis (CAD) systems based on deep learning (DL) have been increasingly explored [5][7][8]. Convolutional Neural Networks (CNNs) have demonstrated strong performance in medical image analysis, particularly in segmentation and classification tasks [9]. Architectures such as U-Net proposed by Ronneberger et al. have been widely adopted for biomedical segmentation due to their ability to preserve spatial information through skip connections [10][11]. Similarly, pretrained backbone networks such as ResNet have achieved high accuracy in image-level classification problems [12]. However, most existing models focus on accuracy and ignore the reliability of their predictions. For clinical adoption, estimating predictive uncertainty is crucial, as it can flag cases where human review is needed, improving safety and trust in AI-assisted diagnosis [13], [14].

Despite these advances, several limitations remain in current cervical cytology analysis systems. Albzour and Lam proposed a segmentation and classification of pap smear images for cervical cancer detection using deep learning [15], combining a U-Net segmentation stage with a separate classification model and comparing performance with and without segmentation, highlighting a modular pipeline rather than joint training. In addition, Aaseegha & Venkataramana [16] noted that conventional methods use for separate segmentation (e.g., SegNet) and classification (ResNet50) models rather than a unified architecture, limiting feature sharing. Joint learning can enhance representation, especially when annotated medical data are scarce. Furthermore, most DL classifiers rely on softmax probabilities, which often produce overconfident predictions even when uncertain [17][18]. In high-risk applications like cervical cancer screening, predictive uncertainty helps flag cases for human review, and fixed loss weights in multi-task frameworks can be suboptimal due to varying task difficulty and convergence [19].

To address these challenges, this paper proposes a unified uncertainty-aware multi-task deep learning framework for cervical lesion segmentation and classification. The proposed architecture uses a pretrained ResNet50 encoder to extract hierarchical features shared between tasks. A U-Net style decoder reconstructs pixel-wise lesion masks for segmentation, while an attention-guided classification head utilizes deep semantic features for lesion type (cervical cell type). Instead of conventional softmax, the classification branch adopts Evidential Deep Learning (EDL) to model class predictions as Dirichlet distributions. This formulation enables the estimation of predictive uncertainty in addition to class probabilities. Furthermore, we incorporate homoscedastic uncertainty-based loss weighting to dynamically balance segmentation and classification losses during training. This allows the network to automatically learn the relative

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

importance of each task, improving overall optimization stability. The main contributions of this work are summarized as follows:

- A unified multi-task architecture that jointly performs cervical lesion segmentation and image-level classification within a single end-to-end framework.
- An attention-guided evidential classification head that provides both class probabilities and uncertainty estimates using Dirichlet modeling.
- A learnable homoscedastic uncertainty-based loss formulation that dynamically balances multi-task objectives.

By integrating segmentation, classification, and principled uncertainty estimation into a single framework, the proposed method aims to improve both diagnostic performance and reliability, supporting safer deployment of AI systems in cervical cancer screening.

2 Materials and Methods

2.1 Dataset Collection and Preprocessing

The Pomeranian dataset was collected at the Pomeranian Medical University in Szczecin, Poland. It includes 419 Pap smear images in BMP format, each with a resolution of 1130 x 1130 pixels. The dataset contains 124 HSIL images, 61 LSIL images, and 234 normal cases [20]. The Herlev dataset consists of 917 Pap smear images divided into seven different cervical cell classes. Those are 150 squamous cell carcinoma in situ, 146 moderate squamous non-keratinizing dysplasia, 197 severe squamous non-keratinizing dysplasia, 182 mild squamous non-keratinizing dysplasia, 70 intermediate squamous epithelial, 98 normal columnar epithelial, and 74 superficial squamous epithelial cells [21]. The first four classes are abnormal, while the last three are treated as normal for binary classification. The SIPaKMeD dataset contains 4,049 single-cell images that were manually cropped from 966 cluster images of Pap smear slides [22]. These images are grouped into five categories: 831 superficial intermediate, 787 parabasal, 793 metaplastic, 825 koilocytotic, and 813 dyskeratotic cells. The first two are normal, and the last three are abnormal.

In the preprocessing stage, we resized all images to 224 x 224 pixels and normalized the pixel values to the range 0 to 1. For the segmentation task, masks for the Pomeranian dataset were generated automatically using OpenCV, without manual review by medical experts. We preprocessed all images by converting them to grayscale, applying Otsu's adaptive thresholding, and performing morphological operations (opening and dilation with an elliptical kernel) to create binary masks. These masks were used to approximate whole-cell boundaries for experimental purposes. However, since they were not expert-annotated, this may affect the segmentation accuracy and introduce potential bias. For the SIPaKMeD dataset, the corresponding segmentation masks are publicly available online. In the case of the Herlev dataset, the original multiclass color-

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

coded masks were converted into binary masks to perform whole-cell segmentation. The dataset is divided into 64% training, 16% validations, and 20% testing.

2.2 Proposed Method

The proposed framework is a unified multi-task DL architecture designed to simultaneously perform cervical lesion segmentation, image-level classification, and predictive uncertainty estimation. The architecture comprises a shared ResNet50 encoder for hierarchical feature extraction, followed by two task-specific branches: a U-Net-style decoder for pixel-wise segmentation and an attention-guided evidential classification head for uncertainty-aware lesion labeling. This design enables shared knowledge between tasks while allowing task-specific specialization. A trainable multi-task loss balances segmentation and classification objectives dynamically, inspired by recent works in multi-task learning with uncertainty weighting [23]–[25].

Images are processed by a pretrained ResNet50 encoder. Hierarchical feature maps are extracted from multiple depths, preserving both low-level spatial details and high-level semantic information. These multi-scale features are shared between the segmentation and classification branches, promoting cross-task feature learning.

Segmentation Branch. The segmentation branch reconstructs the original image resolution via a U-Net-style decoder, which consists of five up-sampling blocks using transposed convolutions. Each block takes as input the feature map from the previous decoder stage and the corresponding encoder feature map via skip connections. These skip connections concatenate the encoder features (from ResNet50) with the upsampled decoder features to preserve fine-grained spatial information lost during downsampling. The combined features are then refined through convolutional and batch normalization layers before being passed to the next decoder stage to improve training stability. The final layer applies a 1×1 convolution with sigmoid activation, producing a binary segmentation mask. To mitigate foreground-background imbalance, Dice loss is used

$$L_{seg} = 1 - \frac{2|P \cap G|}{|P| + |G|}, \tag{1}$$

where P and G are predicted and ground-truth masks.

The combination of a ResNet50 encoder and U-Net-style decoder is particularly suitable for cervical cytology images, as it captures both fine-grained cellular boundaries through skip connections and high-level morphological features necessary for distinguishing subtle differences between lesion types.

Classification Branch. The classification branch operates from the encoder feature map (ResNet50). A 1×1 convolution with sigmoid activation produces an attention map, which weights the encoder features to highlight diagnostically relevant regions. The attention-weighted features are processed through Global Average Pooling and

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

Dropout (0.4). followed by a dense layer with ReLU activation to produce non-negative evidence e_i for each class. Instead of softmax, the network predicts evidence values e_i for each class, converted to Dirichlet parameters:

$$\alpha_i = e_i + 1 \tag{2}$$

and class probabilities are obtained as:

$$p_i = \frac{\alpha_i}{\sum_j \alpha_j} \tag{3}$$

Total evidence $S = \sum_i \alpha_i$ is used to compute predictive uncertainty:

$$U = \frac{K}{S}, \tag{4}$$

where K is the number of classes, and S is the total evidence. The classification evidential loss is calculated as:

$$L_{cls} = \text{MSE}(y, p) + 0.1U, \tag{5}$$

where y is the true label, p is the predicted label, U is uncertainty, and MSE means squared error between true label and predicted probability. This penalizes misclassifications while discouraging overconfident predictions under low-evidence conditions.

2.3 Multi-task Loss Optimization and Metrics

Segmentation and classification losses are combined using homoscedastic uncertainty-based weighting:

$$L_{total} = e^{-s_1} L_{seg} + e^{-s_2} L_{cls} + s_1 + s_2, \tag{6}$$

where L_{seg} is segmentation loss, L_{cls} is classification loss, s_1 and s_2 are the learnable log-uncertainties for the segmentation and classification tasks, respectively. These are defined as $s_1 = \log \sigma_{seg}^2$ and $s_2 = \log \sigma_{cls}^2$. Where σ_{seg}^2 and σ_{cls} represent the standard deviations of the errors in each task. Intuitively, the sigmas encode task-specific uncertainty, allowing the model to weight each loss according to its confidence: tasks with higher uncertainty contribute less to the total loss. These parameters are learned during training to balance the relative contribution of each task.

The network is trained end-to-end using the Adam optimizer (learning rate of 0.0001), batch size 8, for 100 epochs. To evaluate the segmentation model, accuracy, precision, recall, dice score, and intersection over union (IoU). To evaluate the classification model, we used accuracy, precision, recall, F1-score, confusion matrix, and Expected Calibration Error (ECE). The ECE quantifies the model's confidence calibration, comparing predicted probabilities to actual accuracy, or to evaluate whether the model's confidence scores are reliable. The proposed method is shown in Fig. 1.

The experiments were carried out in Python on a Windows workstation featuring an NVIDIA GeForce RTX 3090 GPU. The system was powered by an AMD Ryzen 9 5950X processor with 16 cores operating at 3.40 GHz and supported by 128 GB of RAM. The proposed framework was implemented using the TensorFlow and Keras deep learning libraries.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

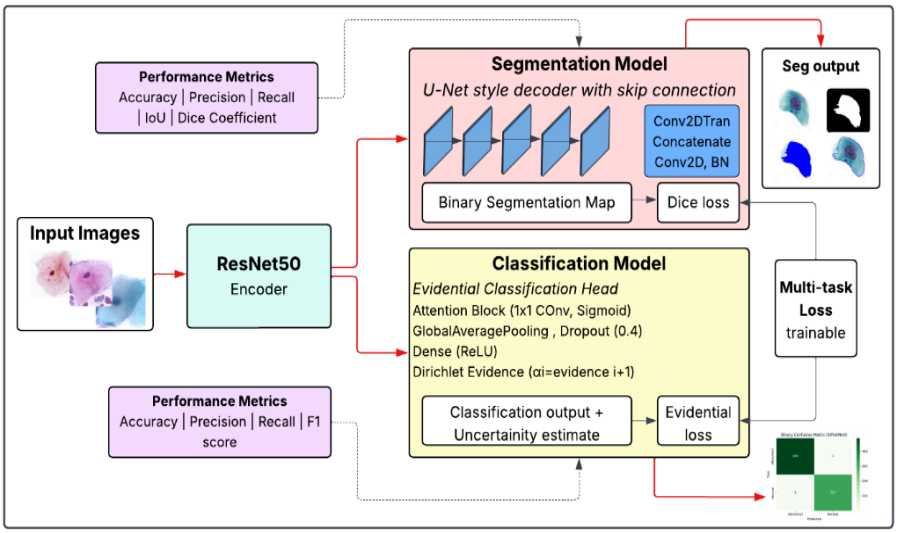

Fig 1. Multi-task DL model for segmentation and evidential classification

3 Results and Discussion

This section presents the performance evaluation of the proposed uncertainty-aware multi-task framework for cervical lesion segmentation and classification. We assess the model on three benchmark Pap smear datasets, Pomeranian, Herlev, and SIPaKMeD, using a comprehensive set of metrics for both segmentation and classification tasks. Additionally, we examine the model's ability to estimate uncertainty, highlighting its potential to identify ambiguous cases requiring expert review. The segmentation using the ResNet50 encoder with U-style decoder is shown in Table 1.

The segmentation performance of the proposed multi-task uncertainty-aware framework demonstrates strong and consistent results across all three datasets. On the Pomeranian dataset, the model achieved a good performance, with an accuracy of 99.19% and a Dice score of 96.68%, indicating highly precise delineation of cervical lesions and effective handling of fine-grained structures. The Herlev dataset showed slightly lower metrics, with an accuracy of 87.84% and a Dice score of 91.53%, reflecting the increased variability and complexity of cell types in this dataset. For the SIPaKMeD dataset, the model maintained robust performance, achieving a Dice score of 94.30% and an IoU of 89.22%, highlighting its ability to generalize well to single-cell images. Overall, these results illustrate that the framework reliably captures lesion regions with high precision and overlap with ground-truth masks, confirming the effectiveness of the U-Net-style decoder combined with multi-scale features from the shared ResNet50 encoder. The segmentation mask is shown in Fig. 2.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

Table 1. Segmentation results using the proposed multi-task uncertainty-aware method

Datasets	Accuracy	Precision	Recall	Dice score	IoU
Pomeranian	99.19%	95.58%	97.81%	96.68%	93.57%
Herlev	87.84%	90.95%	92.13%	91.53%	84.39%
SIPaKMeD	93.85%	93.62%	95.00%	94.30%	89.22%

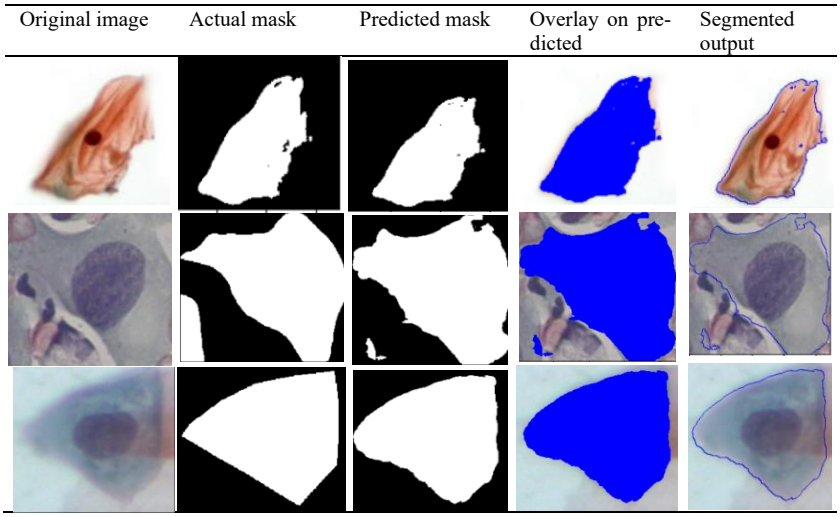

Fig. 2. Segmentation output using the proposed multi-task uncertainty-aware method

The segmentation performance of the proposed uncertainty-aware multi-task model was qualitatively evaluated on three benchmark datasets: Pomeranian (top row), Herlev (middle row), and SIPaKMeD (bottom row). As shown in Fig. 2, the model accurately localizes the primary cellular structures, with predicted masks (third column) closely matching the ground truth masks (second column). In the Pomeranian dataset, the model demonstrates high morphological fidelity, effectively capturing cell boundaries. Minor deviations are observed in the Herlev and SIPaKMeD results; in Herlev, the predicted boundaries show slight misalignments with the actual masks, leading to a modest reduction in quantitative metrics such as the Dice score. For SIPaKMeD, while the ground-truth masks exhibit sharp, well-defined edges, the model's predictions are slightly smoother. This smoothing arises from the U-Net-style decoder and ResNet50 backbone, which emphasize global structural consistency and classification-relevant features, optimized through the integrated evidential loss, over pixel-perfect edge precision. Despite these minor variations, the final segmented outputs (fifth column) show that the predicted contours remain accurately centered on the target lesions, offering

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

reliable localization for potential clinical applications. The classification results of the ResNet50 encoder with the evidential classification head are shown in Table 2.

Table 2. Classification results using the proposed multi-task uncertainty-aware method

Datasets	Accuracy	Precision	Recall	F1 score
Pomeranian	94.05%	94.27%	94.05%	93.87%
Herlev	95.11%	95.39%	95.11%	94.82%
SIPaKMeD multiclass	97.04%	97.05%	97.04%	97.03%
SIPaKMeD binary	99.14%	99.14%	99.14%	99.13%

As can be seen in Table 2, the model demonstrates strong predictive capability across all datasets. For the Pomeranian dataset, the model achieves an accuracy and recall of 94.05%, and an F1 score of 93.87%. On the Herlev dataset, the performance improves slightly, with an accuracy of 95.11% and a corresponding F1 score of 94.82%, reflecting robust classification even in a dataset with more diverse cellular morphologies. The SIPaKMeD dataset, evaluated both in multiclass and binary classification, shows the highest performance. In the multiclass scenario, the model achieves an accuracy of 97.04% with an F1 score of 97.03%, demonstrating reliable differentiation among multiple cell types. When simplified to binary classification, the model attains 99.14% accuracy and an F1 score of 99.13%, highlighting its exceptional capability to correctly identify abnormal cells with minimal misclassification. The confusion matrix using the proposed method is illustrated in Fig. 3.

In Fig. 3, for the Pomeranian dataset, only 5 out of 84 samples were misclassified, resulting in an error rate of 5.95%. Most of the errors occurred when LSIL images were misclassified as HSIL. In the Herlev dataset, 9 normal cells were incorrectly labeled as abnormal, leading to a 4.89% error rate. The SIPaKMeD binary task showed very high reliability, with only 7 misclassifications out of 810 samples, corresponding to an error rate of 0.86%. For SIPaKMeD multiclass, 24 samples were misclassified, resulting in a 2.10% error rate, mostly among morphologically similar Dyskeratotic, Koilocytotic, and Metaplastic cells. Overall, the model demonstrates strong and consistent performance across both binary and multiclass tasks, with minimal errors even in challenging fine-grained classifications. The predicted classification across all the datasets is shown in Fig. 4.

Fig. 4 presents representative test samples along with their corresponding true and predicted labels. Most of the examples demonstrate correct classification, such as the NSIL and HSIL cells in the Pomeranian dataset and all examples in the SIPaKMeD dataset, where the predicted labels match the actual classes. However, some misclassifications are also observed and are highlighted in red. For example, an LSIL sample is predicted as HSIL in the Pomeranian dataset, and a normal cell is incorrectly predicted as abnormal in the Herlev dataset. These visual examples

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

provide a clearer understanding of the model's prediction behavior across different datasets.

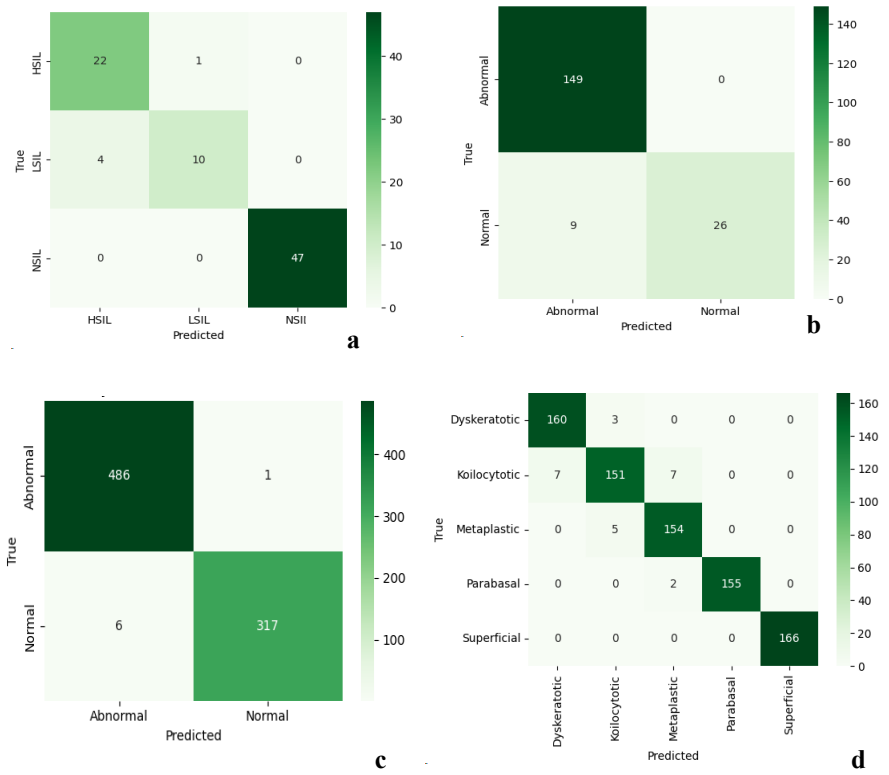

Fig. 3 Confusion matrix: a) Pomeranian, b) Herlev, c) SIPaKMeD binary, d) SIPaKMeD multiclass

Uncertainty and Calibration Analysis. To further evaluate the reliability of the proposed evidential framework, we analyzed the relationship between predictive uncertainty and segmentation quality and model calibration using ECE. For the Pomeranian dataset, a negative correlation of -0.33 was observed between uncertainty and Dice score. Similarly, the Herlev dataset showed a correlation of -0.30. This negative relationship indicates that higher uncertainty is generally associated with lower segmentation accuracy, which confirms that the model expresses greater uncertainty when predictions are less reliable. For the SIPaKMeD dataset, the correlation was weaker (-0.08 for multiclass and 0.018 for binary classification). This suggests that segmentation performance in this dataset is more stable, and uncertainty values vary less with the Dice score. For the Pomeranian and Herlev datasets, the negative correlation shows that the model becomes more uncertain when its predictions are less accurate, which is a desirable behavior. For the SIPaKMeD dataset, the correlation is close to zero (slightly positive in the binary case), indicating that

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

the model performs consistently and its uncertainty does not change much with accuracy. Overall, negative or low correlations are desirable because they show that uncertainty can help identify difficult or ambiguous cases that may need review by a medical expert.

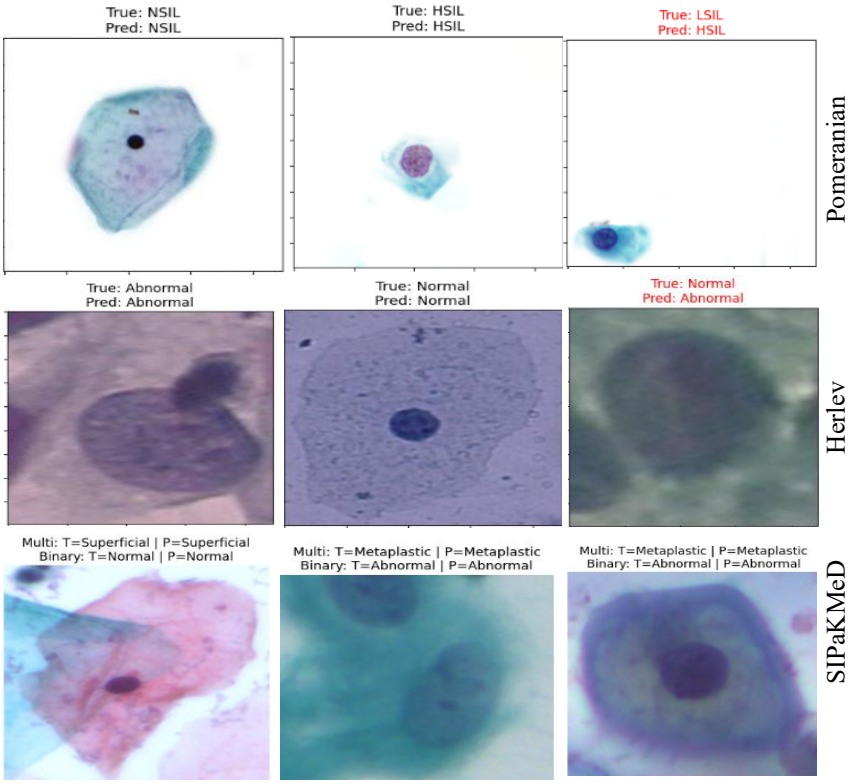

Fig. 4 True and predicted labels for representative test samples (errors in red)

In terms of calibration, the ECE values were low across all datasets. The Pomeranian dataset achieved an ECE of 0.026, Herlev 0.019, and SIPaKMeD showed very low calibration errors of 0.018 (multiclass) and 0.007 (binary). These small ECE values indicate that the predicted probabilities are well aligned with actual accuracy, meaning the model's confidence estimates are reliable. Overall, these results demonstrate that the evidential classification head not only provides accurate predictions but also produces meaningful and well-calibrated uncertainty estimates, which are especially important for clinical decision support systems.

While the proposed method achieves high segmentation accuracy on the Pomeranian dataset (99.19%), it is important to note that the segmentation masks for this dataset were generated automatically using OpenCV without expert review. This may have introduced some bias, potentially inflating performance compared to

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

expert-annotated masks. Future work will validate these results on expert-annotated or multi-center datasets to assess generalizability. Table 3 shows the comparison with the previous work.

Table 3. Comparison of our study with the previous work

Ref	Datasets	Method	Accuracy	Precision	Recall	F1 score
[27]	SIPaKMeD multi	ResNet50	95%	-	-	-
[29]	SIPaKMeD multi	MLP	91.72	91	91.6	91.7
[28]	SIPaKMeD binary	ResNet50	93.7%	-	-	-
[30]	Herlev	DIFF	94.55%	94.10%	91.60%	92.70%
[31]	Herlev	VGG16	94%	94%	94%	94%
[26]	Herlev	Ensemble	99.7	-	-	-
Our	SIPaKMeD multiclass	ResNet50	97.04%	97.05%	97.04%	97.03%
Our	SIPaKMeD binary	ResNet50	99.14%	99.14%	99.14%	99.13%

Table 3 compares the proposed framework with previous studies on SIPaKMeD and Herlev datasets. While some prior works, such as an ensemble approach on Herlev [26], achieve very high accuracy, they rely on multiple models, increasing computational complexity. In contrast, our uncertainty-aware multi-task model uses a single ResNet50 backbone and still achieves 97.04% accuracy for multiclass and 99.14% for binary classification on SIPaKMeD, outperforming other single-model approaches such as ResNet50 [27][28] and MLP [29]. Similarly, on the Herlev dataset, our model demonstrates competitive performance compared to DIFF [30] and VGG16 [31] while providing well-calibrated confidence estimates through the evidential classification head. These results indicate that our framework delivers both high predictive performance and reliable uncertainty quantification without the added complexity of model ensembles.

While individual components such as multi-task learning, attention mechanisms, and evidential deep learning have been previously explored, this work integrates them into a unified framework specifically designed for simultaneous cervical lesion segmentation, classification, and uncertainty estimation. In addition, the proposed method incorporates homoscedastic uncertainty-based loss weighting to dynamically balance tasks and provides a comprehensive evaluation of both predictive performance and calibration across multiple Pap smear datasets. Although the proposed method is evaluated on cervical cytology datasets, its architecture is general and can be applied to other medical imaging tasks involving joint segmentation and classification, such as histopathology or radiology images.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

4 Conclusions

This study presented an uncertainty-aware multi-task deep learning framework for joint cervical lesion segmentation and classification. By combining a shared ResNet50 encoder with a U-Net-style decoder and an evidential classification head, the model performs simultaneous pixel-level localization and image-level diagnosis. Dirichlet-based evidential learning enables reliable uncertainty estimation, mitigating overconfident predictions. Experiments on the Pomeranian, Herlev, and SIPaKMeD datasets show strong and consistent segmentation and classification performance, with high Dice scores, accuracy, and low calibration errors. The negative correlation between uncertainty and segmentation quality confirms that the model expresses uncertainty appropriately for less reliable predictions. While the lack of expert-validated masks in the Pomeranian dataset may introduce bias, future work will explore transformer-based encoders to better capture global context, as well as advanced uncertainty estimation techniques such as Monte Carlo dropout and deep ensembles to improve predictive reliability further. In addition, validation on larger multi-center datasets will be conducted to assess generalizability and robustness.

References

[1] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Deep learning-based automatic segmentation and classification for cervical cancer detection using an improved U-Net and ensemble methods," *Sci. Rep.*, pp. 1–15, 2026, doi: 10.1038/s41598-026-35299-7.

[2] P. Screening and P. Board, "Cervical Cancer Screening (PDQ®)," *PDQ Cancer Inf. Summ. [Internet]*, vol. National C, 2024.

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

[4] H. A. Phoulady, D. Goldgof, L. O. Hall, and P. R. Mouton, "A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images," *Comput. Med. Imaging Graph.*, vol. 59, pp. 38–49, 2017, doi: 10.1016/j.compmedimag.2017.06.007.

[5] S. Yu *et al.*, "Automatic classification of cervical cells using deep learning method," *IEEE Access*, vol. 9, pp. 32559–32568, 2021, doi: 10.1109/ACCESS.2021.3060447.

[6] B. Zewdu Wubineh, Ł. Jeleń, and A. Rusiecki, "DCGAN-based Cytology Image Augmentation for Cervical Cancer Cell Classification," *IEEE Trans. Med. Imaging*, vol. xx, no. 50, pp. 1003–1011, 2020, doi: 10.1016/j.procs.2025.02.206.

[7] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, and X. Hu, "Automatic cell nuclei segmentation and classification of cervical Pap smear images," *Biomed. Signal Process. Control*, vol. 48, pp. 93–103, 2019, doi: 10.1016/j.bspc.2018.09.008.

[8] B. Z. Wubineh and A. Rusiecki, "SE-DeepLabV3 +: Cervical Cell

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

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.

[9] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "ViT-SE_Res: A Hybrid Vision Transformer and ResNet50V2 with Squeeze-And-Excitation Block for Cervical Cell Classification," in *International Conference on Computational Science*, Springer Nature Switzerland, 2025, pp. 11–23.

[10] O. Ronneberger, P. Fischer, and T. Brox., "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference*, Munich, Germany: Springer International Publishing., 2015, pp. 234–241.

[11] N. S. Punn and S. Agarwal, *Modality specific U-Net variants for biomedical image segmentation: a survey*, vol. 55, no. 7. Springer Netherlands, 2022. doi: 10.1007/s10462-022-10152-1.

[12] L. Wen, Z. Xiao, X. Xu, and B. Liu, "Disaster Recognition and Classification Based on Improved ResNet-50 Neural Network," *Appl. Sci.*, vol. 15, no. 9, 2025, doi: 10.3390/app15095143.

[13] M. Dong, A. Yang, Z. Wang, D. Li, J. Yang, and R. Zhao, "Uncertainty-aware consistency learning for semi-supervised medical image segmentation," *Knowledge-Based Syst.*, vol. 309, no. December 2024, p. 112890, 2025, doi: 10.1016/j.knosys.2024.112890.

[14] S. Ojha, "Uncertainty Quantification in DL Models for Cervical Cytology," pp. 1–4, 2024.

[15] N. Albzour and S. S. Lam, "Segmentation and Classification of Pap Smear Images for Cervical Cancer Detection Using Deep Learning," *IISE Annu. Conf. Expo 2025, Conf. Proc.*, pp. 597–602, 2025, doi: 10.21872/2025IISE_6101.

[16] M. D. Aaseegha and B. Venkataramana, "SegResDeiT: a hybrid SegNet–ResNet-50–DeiT framework for automated cervical cancer segmentation and classification," *BMC Med. Imaging*, vol. 25, no. 1, 2025, doi: 10.1186/s12880-025-02001-8.

[17] Kurnianingsih *et al.*, "Segmentation and classification of cervical cells using deep learning," *IEEE Access*, vol. 7, pp. 116925–116941, 2019, doi: 10.1109/ACCESS.2019.2936017.

[18] 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.

[19] K. Zou, Z. Chen, X. Yuan, X. Shen, M. Wang, and H. Fu, "A review of uncertainty estimation and its application in medical imaging," *Meta-Radiology*, vol. 1, no. 1, p. 100003, 2023, doi: 10.1016/j.metrad.2023.100003.

[20] 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.

[21] "PAP-SMEAR (DTU/HERLEV) DATABASES & RELATED STUDIES, Herlev University Medical Center." Accessed: Aug. 31, 2024. [Online]. Available: https://mde-lab.aegean.gr/index.php/downloads/

[22] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, and A.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34

Charchanti, “Sipakmed: A New Dataset for Feature and Image Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images,” *Proc. - Int. Conf. Image Process. ICIP*, no. October, pp. 3144–3148, 2018, doi: 10.1109/ICIP.2018.8451588.

[23] S. Sreelatha and V. Shivashetty, “Deep ensemble learning with uncertainty aware prediction ranking for cervical cancer detection using Pap smear images,” *IAES Int. J. Artif. Intell.*, vol. 14, no. 2, p. 1450, 2025, doi: 10.11591/ijai.v14.i2.pp1450-1460.

[24] A. Kasukurthi and R. L. Davuluri, “AMTLUS: Attention-guided multi-task learning with uncertainty estimation in skin lesion segmentation and classification,” *Multimed. Tools Appl.*, vol. 83, no. 37, pp. 84885–84909, 2024, doi: 10.1007/s11042-024-19360-z.

[25] Z. Liu, Z. Yan, and G. Li, “FDMNet: A Multi-Task Network for Joint Detection and Segmentation of Three Fish Diseases,” *J. Imaging*, vol. 11, no. 9, 2025, doi: 10.3390/jimaging11090305.

[26] S. K and G. C, “A hybrid model for efficient cervical cell classification,” *Biomed. Signal Process. Control*, vol. 72, no. PA, p. 103288, 2022, doi: 10.1016/j.bspc.2021.103288.

[27] C. Ozcan, “Deep Learning Based Cervical Cells Classification from Pap Smear Images,” no. December, 2024, doi: 10.54381/itta2024.08.

[28] A. Tripathi, A. Arora, and A. Bhan, “Classification of cervical cancer using Deep Learning Algorithm,” *Proc. - 5th Int. Conf. Intell. Comput. Control Syst. ICICCS 2021*, no. Iciccs, pp. 1210–1218, 2021, doi: 10.1109/ICICCS51141.2021.9432382.

[29] W. Liu *et al.*, “CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN, visual transformer and multilayer perceptron,” *Pattern Recognit.*, vol. 130, p. 108829, 2022, doi: 10.1016/j.patcog.2022.108829.

[30] 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, 2024, doi: 10.1016/j.compbiomed.2024.108153.

[31] N. Sritharan, N. Gnanavel, P. Inparaj, D. Meedeniya, and P. Yogarajah, “Explainable Artificial Intelligence Driven Segmentation for Cervical Cancer Screening,” *IEEE Access*, vol. 13, no. February, pp. 71306–71322, 2025, doi: 10.1109/ACCESS.2025.3561178.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29912-3_34