

Doctoral dissertation review
titled

Segmentation and Classification of Pap Smear Images to Detect Cervical Cancer using Deep Learning

This review of the doctoral dissertation, authored by Betelhem Zewdu Wubineh, was commissioned by the Chairman of the Disciplinary Council for Technical Information Technology and Telecommunications, prof. Wojciech Bożejko (Letter No. W4N-KP-SW.4110.7.2026 dated 25.06.2026). The supervisor of this dissertation is prof. Andrzej Rusiecki, and the auxiliary supervisor is dr Krzysztof Halawa. The doctoral dissertation is written in English.

1. The research problem and its significance

What is the most important issue considered in the dissertation?

The dissertation's central issue is how to develop a reliable, accurate, and clinically interpretable deep-learning pipeline for early cervical cancer detection from Pap smear images, despite limited, imbalanced, heterogeneous, and often difficult-to-annotate cytology datasets. More specifically, the work addresses the interrelated problem of whether cervical-cell segmentation, synthetic-data augmentation, attention/transformer mechanisms, and ensemble learning can improve the classification of normal and abnormal cervical cells. The dissertation aims to construct and assess a computer-aided diagnosis framework that is more robust than conventional single-model, classification-only approaches.

There are mainly four connected technical challenges:

- Accurate delineation of cervical-cell components, principally cytoplasm and nuclei, against background and in more complex cellular images.
- Reliable discrimination of normal and abnormal cells, including multiclass and binary classification scenarios.
- Mitigation of limited sample size and class imbalance through realistic synthetic image generation.
- Determination of whether segmentation-guided classification genuinely adds predictive value relative to direct classification from original images, and which representation of the segmented cell is most useful.

Is it scientific in nature?

The dissertation is clearly scientific in nature, particularly within information and communication technology, medical image analysis, and applied machine learning. It identifies a defined research problem, states objectives and hypotheses, proposes methodological contributions, performs systematic experiments, and interprets results against established baselines and prior research. It addresses a clearly defined and socially important research problem in the field of medical image

WPLYNĘŁO

02-09-2026

analysis: the development of robust deep-learning methods for the segmentation and classification of Pap-smear images to support early cervical-cancer detection. The candidate formulates research objectives, questions, and hypotheses; proposes and evaluates original architectural and pipeline modifications; conducts extensive quantitative experiments using recognised metrics and several datasets; compares the results with baseline and literature methods; and discusses limitations and future research directions. The work therefore fulfils the substantive criteria of empirical and methodological research in information and communication technology. However, the practical clinical implications should be treated cautiously pending validation using larger, independently annotated, multi-institutional clinical datasets and expert assessment of the proposed explanatory outputs.

Does it have any practical significance?

The dissertation has meaningful practical significance because it develops and evaluates tools that could support computer-aided cervical-cancer screening from Pap-smear images; particularly by helping clinicians localize relevant cell regions, prioritize suspicious cases, and obtain more consistent image-based assessments. Its practical relevance is strongest as a decision-support and workflow-support technology, not as a replacement for cytologists or pathologists. The practical motivation is substantial. Manual Pap-smear analysis requires specialist expertise, is time-consuming, and can vary between readers. The dissertation targets automated analysis of cervical cytology images to make screening more consistent and potentially more scalable, especially where access to trained cytologists and pathology infrastructure is limited.

In practical terms, a mature version of the proposed system could support a screening workflow by:

- Detecting and segmenting relevant cellular structures, including cell boundaries, cytoplasm, and nuclei.
- Classifying images or detected cell regions into normal and abnormal categories, or into more detailed diagnostic classes depending on the dataset.
- Flagging difficult or low-confidence cases for human review rather than presenting every prediction with equal certainty.
- Providing visual localization through segmentation masks, Grad-CAM, and saliency maps, allowing the user to inspect regions that contributed to a prediction.
- Reducing the manual burden of reviewing large numbers of images by enabling prioritization rather than fully automated diagnosis.

At the same time, the practical utility should be assessed as prospective rather than fully demonstrated. Before clinical deployment, the methods require validation on larger, independently annotated, multi-institutional datasets; expert assessment of the explanation outputs; evaluation of robustness to routine slide-preparation variability; and analysis of workflow, safety, computational, and regulatory requirements.

2. Author's contribution

What is the author's most important contribution to obtaining the results described in the dissertation?

If necessary, the reviewer may distinguish between the contribution claimed by the author of the thesis and the contribution the reviewer considers most important. In this case, the reviewer should provide reasons why the reviewer disagrees with the author's claims (e.g., someone has already proposed the idea, or it is original but incorrect for the reasons described in section 3). We also ask for comments on the practicality of the proposed solutions (it may be that the problem is very practical, but the proposed solution is not). If appropriate, the reviewer may refer to other quality indicators (e.g., quality of the publication, the author's patents, citations, implemented applications, etc.).

The author's most important contribution is the development and systematic experimental validation of an integrated deep-learning framework for cervical cytology analysis, combining cell segmentation, classification, synthetic-data augmentation, ensemble learning, transformer-based feature modelling, and uncertainty-aware prediction. The principal scientific value of the work lies not in the isolated use of individual established mechanisms, such as U-Net variants, attention modules, residual connections, GAN augmentation, or pretrained CNN ensembles, but in their coherent integration and comparative evaluation in segmentation-guided cervical-cell classification.

In my opinion, the most important original result is the demonstration that segmentation-derived input must retain a controlled amount of pericellular context. The results show that tightly masked ROI images may reduce classification performance by removing useful information, whereas BBox-based representations, with a limited margin around the segmented cell, often provide better and more stable classification results. This observation is supported across several datasets and model configurations and constitutes a practically useful design principle for future computer-aided cervical-cancer screening systems.

The proposed deep-learning models are technically sound and well matched to the analysed problem. They cover complementary aspects of cervical-cytology image analysis, including multiscale segmentation, attention-based feature selection, global contextual modelling through Vision Transformers, generative augmentation, ensemble classification, and uncertainty-aware multitask learning. Their main scientific value lies in the coherent integration and comparative evaluation of established deep-learning mechanisms, rather than in the introduction of entirely new architectural principles. I like the proposed DL models, however, they are largely incremental architectural extensions and integrations of known components. U-Net, dilated convolutions, Group Normalization, residual blocks, SPP, squeeze-and-excitation, DeepLabV3+, Vision Transformers, GANs, pretrained CNN backbones, attention modules, and ensembles were already established before this work. The originality lies primarily in the integration, adaptation, and systematic comparison of established deep-learning mechanisms for cervical-cytology analysis, rather than in the invention of wholly new neural-network principles.

Particularly valuable is the author's experimental demonstration that segmentation-derived BBox inputs preserving limited pericellular context can outperform tightly masked ROI images in classification. This result has direct practical implications for the design of future segmentation-guided computer-aided screening systems.

The models should nevertheless be regarded as validated research prototypes. Before clinical deployment, the preferred architecture should be simplified and optimised, and its performance should be confirmed using large, independently annotated, multi-centre clinical datasets and expert-validated explanation and uncertainty analyses. Some proposed models are computationally heavy.

The proposed solutions have clear practical relevance as components of clinical decision-support systems: they can support the localisation of cellular structures, prioritisation of suspicious cases, robust classification, and uncertainty-aware referral of difficult cases for expert assessment. Nevertheless, the work should be regarded as a validated research prototype rather than a clinically deployable autonomous diagnostic tool. Further multi-institutional validation, expert annotation of segmentation and explanatory regions, prospective workflow studies, robustness testing under routine laboratory variability, and computational optimisation are necessary before implementation in clinical practice.

The scientific quality of the work is additionally supported by a substantial publication record in peer-reviewed journals and conference proceedings directly related to the doctoral topic. The thesis would

be further strengthened by independently documented information on authorship contributions, reproducible software or data availability, citation impact, intellectual-property outcomes, and clinical or industrial implementation.

3. Correctness

Are the claims made in the thesis trustworthy? Are the justifications sound? Point out any weaknesses and errors you've noticed. Also, highlight the aspects of validity that are most valuable (elegance of evidence, experimental design, analysis of empirical data, quality of prototype software/hardware, etc.).

The claims made in the dissertation are generally credible within the experimental scope of the study. The author provides a broad and coherent empirical evaluation of segmentation, classification, augmentation, ensemble learning, transformer-based modelling, and uncertainty-aware prediction for cervical cytology images. Particularly valuable are the comparisons of original-image, ROI-based, and BBox-based inputs, which demonstrate that exact segmentation masks do not necessarily maximise classification performance and that preserving limited local context around the cervical cell can be advantageous. The use of several datasets, multiple performance metrics, selected five-fold cross-validation, cross-dataset evaluation, confusion matrices, ranking analyses, and explicit discussion of limitations strengthens the validity of the conclusions.

The proposed methods should be regarded as strong research prototypes with substantial potential for computer-aided decision support, rather than as clinically ready autonomous diagnostic systems. Further work should include independently annotated multi-institutional data, prospective or reader-study validation, expert assessment of synthetic images and explanatory maps, formal analysis of distribution shift and calibration, and optimisation of the computationally expensive ensemble architectures for deployment.

4. Candidate knowledge

Which chapters (or sections within chapters) of the dissertation discuss the existing state of knowledge and thus confirm the candidate's general knowledge in Technical Information Technology and Telecommunications? What areas of these disciplines are covered in these chapters/sections? What is the reviewer's opinion of the quality of these chapters/sections? What is the reviewer's opinion of the bibliography? How complete is the bibliography? Please provide other arguments for or against stating that the candidate has general knowledge in Technical Information Technology and Telecommunications.

The thesis's principal claims are plausible and partly well supported, especially the claims that architectural refinements can improve segmentation, that ensembles often outperform individual classifiers, and that BBox-based inputs can preserve useful context better than tightly masked ROIs. However, the strength of the claims varies: the evidence supports experimental effectiveness on the selected datasets, but it is not yet sufficient to support strong claims of clinical readiness, broad real-world generalization, or superiority over prior work in an unqualified sense.

The dissertation has a good empirical basis. It contains explicit research questions, detailed model descriptions, quantitative comparisons with baselines, several types of validation, ablation-like comparisons of input representations, multiple datasets, and reported limitations. These features make its main experimental conclusions trustworthy within the scope of the data and protocols used. The most important conclusions are:

- Modified U-Net-based models and other proposed segmentation approaches can achieve high performance on the evaluated cervical-cytology datasets.

- Ensembles, attention mechanisms, and hybrid CNN-ViT models often outperform individual backbone models in the reported experiments.
- RES_DCGAN augmentation can improve some outcomes and produces lower reported FID values than the baseline DCGAN.
- Segmentation can improve localisation and interpretability, but it is not consistently necessary for high classification accuracy.
- BBox crops retaining limited context often outperform exact ROI masks; the amount of preserved context matters and overly large margins can reduce performance.
- Uncertainty-aware multitask learning can jointly produce segmentation, classification, and confidence-related estimates under the reported protocol.

The work is evaluated on several cervical datasets. It also includes supplementary transferability experiments on HAM10000 skin-lesion segmentation and brain-tumour MRI classification. The dissertation uses percentage splits, selected five-fold cross-validation, and some cross-dataset evaluations. The thesis goes beyond classification accuracy by introducing an uncertainty-aware multitask model and by using Grad-CAM and saliency visualisations.

By contrast, statements such as “suitable for real-world clinical applications,” “reliable decision support,” or “significantly outperforms” need more careful qualification. These claims may be valid as descriptions of results under the thesis’s benchmark settings, but are not fully established as claims about routine clinical performance.

The most important methodological weakness is the stated use of OpenCV-based processing to generate segmentation masks for the private Pomeranian dataset, rather than comprehensive manual annotations by expert cytologists.

5. Other notes

The thesis follows a conventional and understandable doctoral structure: introduction, literature review, methodology, experiments, integrated analysis, conclusions, and future work. The research questions, objectives, and stated contributions are presented explicitly. The thesis does not hide limitations. The dissertation is generally correctly typeset and follows a clear formal structure.

Possible errors:

- weekly supervised -> weakly supervised,
- Bounding BBox -> Bounding Box,
- wrong captions for Figure 3. 33 and Table 6. 14.
- Typos, e.g.: DCAGN, Segmetnation
- RES-DCGAN and RES_DCGAN

6. Summary

Taking into account the opinions presented in the previous points and the requirements defined by Article 187 of the Act of 20 July 2018 - Law on Higher Education and Science (as amended)¹ my assessment of the dissertation in terms of three basic criteria is as follows:

A. Does the dissertation contain an original solution to a scientific problem? (select one option by marking X)

¹<http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20190000276>

☒

Definitely YES

☐

Probably YES

☐

It's hard to say

☐

I don't think so

☐

Definitely NOT

B. After reading the dissertation, do you agree that the candidate has general theoretical knowledge in the discipline of Technical Information Technology and Telecommunications?

☒

Definitely YES

☐

Rather YES

☐

It's hard to say

☐

I don't think so

☐

Definitely NOT

C. Does the candidate have the ability to conduct research independently?

☒

Definitely YES

☐

Rather YES

☐

It's hard to say

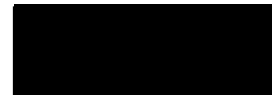
☐

I don't think so

☐

Definitely NOT

Moreover, taking into account that the dissertation presents a substantial, original, and practically relevant contribution to deep-learning-based cervical cytology analysis, supported by extensive experimental studies, multiple datasets, and rigorous evaluation, and its high scientific quality is additionally confirmed by the candidate's strong publication record, including several papers in highly ranked peer-reviewed journals such as Scientific Reports, Cancers, I recommend awarding distinction to the doctoral dissertation.



Signature