A Stain-Invariant and Calibrated Deep Learning Framework for Automated Cervical Cytology Classification

V. Gokula Krishnan[1], Arvind Kumar Tiwari[2]

[1] Post-Doctoral Research Fellow, Department of Computer Science and Engineering, Lincoln University College, Malaysia;

[2] Adjunct Professor, Lincoln University College, Malaysia;

Email ID : pdf.gokul@lincoln.edu.my

Abstract: Pap-smear cytology, a method for the screening of cervical cancer, is exposed to a number of issues such as stain variability, morphological alterations, class imbalance, and unavailability of dependable confidence estimation in automated systems. To overcome these problems, the article comes up with a stain-invariant and well-calibrated multi-scale deep learning system named SiCal-CytoNet for cervical cell classification. It not only supports stain separation, domain generalization, adaptive multi-scale feature fusion, prototype-guided classification and probability calibration to improve diagnosis accuracy but also ensures reliability. The SiCal-CytoNet achieved a classification accuracy of 98.9%, macro-F1 score of 98.7%, AUROC of 99.7%, and AUPRC of 99.4% on the SIPaKMeD dataset. It was, at the same time, well-calibrated as shown by the low Brier score of 0.012 and Expected Calibration Error (ECE) of 1.7%. On top of that, the model was able to handle stain illumination blurring, and variations in styles effectively. The introduction of a confidence-based selective prediction process enabled the model to individually manage 88.1% of the cases while ensuring high diagnostic reliability. The findings very much indicate that the triple combination of stain-invariant learning, multi-scale feature representation, and calibrated prediction results in an accurate, robust, and clinically reliable solution for an AI-empowered cervical cancer screening and cytology decision support.

Keywords: Pap-smear Cytology, Optical-density, Brier Loss, Calibration, Prototype-guided Metric Head.

Introduction

Worldwide cervical cancer remains a major health issue, Mostly in low- and middle-income countries where the people are not able to receive a timely screening and diagnosis. Pap-smear cytology is capable of being an excellent technique of screening, but its large-scale global implementation is currently hindered by the heavy workload in analysis, variability of results between different observers, and the increasing diagnostic workloads. Though, deep learning techniques were developed and they showed a good potential for the automated classification of cytology. Yet, model's performance and clinical applicability degradation is still caused by changes in staining, imaging conditions and appearance class imbalance, as well as the unreliable confidence estimates.

The SIPaKMeD dataset contains five categories of cervical cells which are very different with appearance and texture making automated classification a very challenging task. This paper introduces SiCal-CytoNet to deal with these issues, which is a stain invariant and calibrated multi-scale deep learning structure for cervical cytology. Our technique integrates stain-aware preprocessing, domain generalization techniques, adaptive multi-scale feature learning, prototype-guided classification, and probability calibration not only for enhancing classification accuracy but also for boosting prediction reliability. The proposed structure can be employed for cervical cancer screening and decision-making that are clinically dependable by combining powerful feature learning with reliable confidence estimation.

Related work

Recently, developments in artificial intelligence have markedly enhanced cervical cancer diagnosis via imaging and cytology analysis as well as multimodal learning. For instance, attention-based segmentation networks have yielded remarkably good results in identifying cervical tumors as well as anatomical structures from MRI images, which implies the critical role feature attention and image quality enhancement play. Combining imaging features and clinical information in a multimodal way resulted in further improvements in the prediction of early-stage cancer by taking advantage of the complementary nature of different diagnostic evidence.

A number of review publications have documented the increasing use of AI in gynecologic oncology for activities such as screening diagnosis treatment planning, and disease monitoring. Deep learning architectures, Mostly those with multi-scale feature extraction, ensemble learning, and transfer learning, have demonstrated high classification accuracies on cervical cytology datasets like SIPaKMeD. Besides, explainable AI methods have been investigated to increase model transparency and clinician acceptance.

Though, even after these advances, dealing with stain differences, domain shifts between laboratories, class imbalance, and trustworthy confidence estimation still remain as problems. Also, calibration metrics and external validation are seldom mentioned, which further curbs real-world implementation. These constraints call for the creation of excellent and reliable cytology systems combining stain-invariant learning, multi-scale feature extraction, and calibrated predictions to clinically support decisions.

Key Contribution

This research paper introduces SiCal-CytoNet which is a multi-scale deep learning model capable of stain-invariant and calibrated cervical cytology classification. Our method combines different techniques related to stain-aware preprocessing, learning domain generalization, adaptive multi-scale feature fusion, prototype-guided simplified representation learning, and probability calibration to enhance classification accuracy and prediction reliability. Besides, a confidence-based selective prediction system is added to provide reliable support for clinical decision-making by pointing out uncertain cases which require specialist's review. Extensive experiments performed on the SIPaKMeD dataset demonstrate the robustness of the method despite the differences in staining, lighting, and image quality, So implicating that the technique is capable of providing efficient, dependable, and clinically significant cervical cancer screening.

Method, Experiments and Results

This article that presents a novel multi-scale deep learning system called SiCal-CytoNet which is stain-invariant and calibrated for automated classification of cervical cytology. The five major stages of the proposed workflow are described here: At first, Pap smear images are transformed into the optical density space and then their deconvolution and augmentation are done to reduce stain variations. The authors go a step further and investigate domain generalization techniques like Fourier Domain Adaptation (FDA), MixStyle, and feature alignment. These methods help them to effectively solve the problem of domain shift resulting from laboratory-specific imaging differences. The remaining part of the system focuses First and foremost on feature extraction and classification. In fact, it not only extracts morphological features at various levels of resolution changes, but also employs a learnable

resolution gate for adaptive fusion of these features. Increasing class separability is achieved with a prototype-guided metric head, while two different methods, evidential learning and temperature scaling, are used for model confidence to accurately reflect actual likelihoods. Figure 1 illustrates the entire system design and operation.

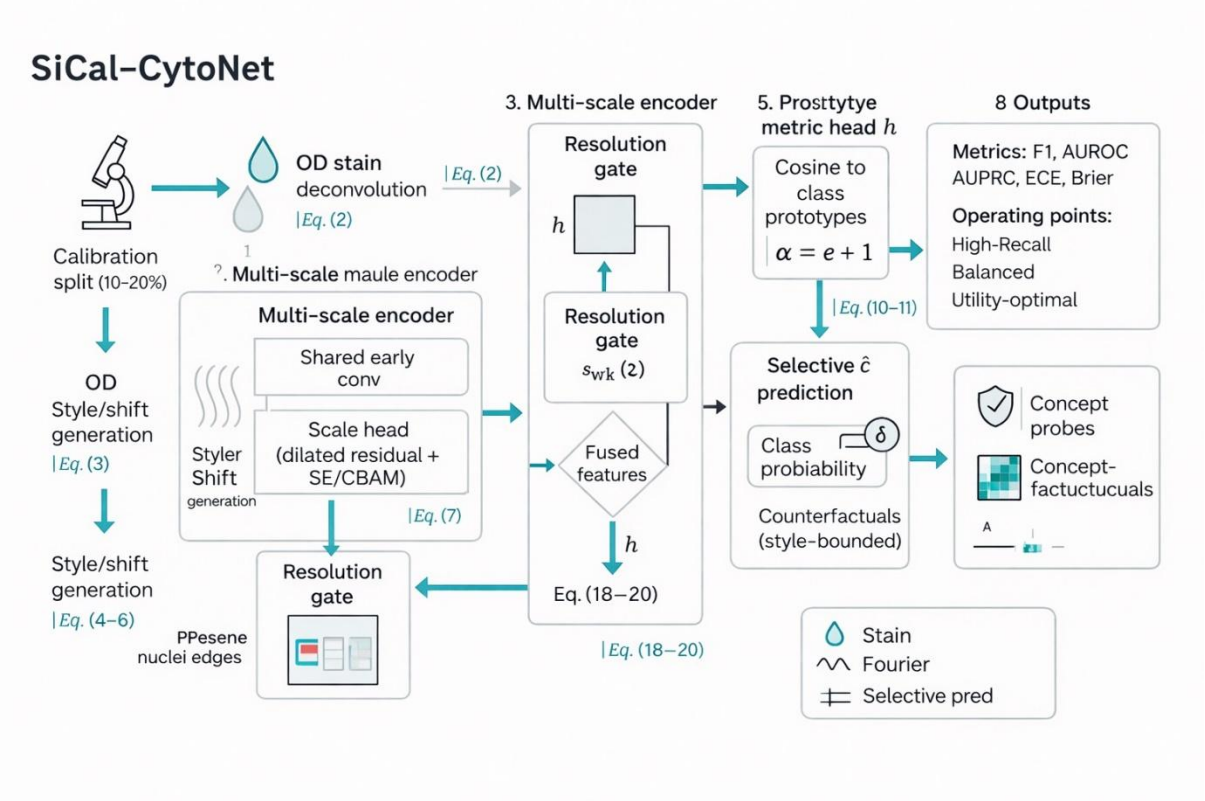

Figure 1: Workflow of the proposed model

We performed our research on the SIPaKMeD collection, which is a publicly available collection of data containing 4,049 annotated images of cervical cells divided into five diagnostic groups. To evaluate the model's performance, we used Accuracy, Macro-F1 score AUROC AUPRC, Brier Score, and Expected Calibration Error (ECE) as evaluation metrics. The proposed system has produced result of 98.9% accuracy, a Macro-F1 score of 98.7%, an AUROC of 99.7%, and an AUPRC of 99.4%. Also, the model was very well calibrated as evidenced by a Brier score of 0.012 and an ECE of 1.7%. Evaluation of robustness to variations in stain, style shifts, blur, and changes in illumination, confirmed very high performance in all cases, while ablation studies showed the strength of multi-scale fusion, domain generalization, and calibration components. The implementation of confidence-based selective prediction strategy made possible the automated classification of 88.1% of patients with a sustained high level of diagnostic reliability.

Discussions

We experimentally found that cervical cell recognition performance can be dramatically improved by combining stain-invariant preprocessing, adaptive multi-scale feature learning, and prototype-guided classification. The resolution-gated fusion not only captures delicate nuclear details but also considers the overall cellular context. At the same time, domain generalization methods help maintain robustness

to variations in staining and imaging. By using calibration techniques such as evidential learning and temperature scaling, we obtain reliable confidence scores that are essential for clinical decision support. Also, the selective prediction method increases the level of trustworthiness by referring the doubtful cases to the experts for a second opinion. Taken together, the proposed SiCal-CytoNet system is a high-performing, reliable, and clinically trustworthy tool for AI-powered cervical cancer screening and cytology automatization.

Conclusions

In this paper, we proposed SiCal-CytoNet, a stain-invariant and calibrated multi-scale deep learning model for automated cervical cytology classification on the SIPaKMeD dataset. Our model combined stain-aware preprocessing, adaptive multi-scale feature fusion, prototype-guided learning, and probability calibration, resulting in outstanding classification performance: 98.9% accuracy, 98.7% Macro-F1 score, 99.7% AUROC, and 99.4%AUPRC. Besides, it remained well-calibrated (Brier score = 0.012) and was consistent with Expected Calibration Error of 1.7%. We also showed that the proposed structure is resilient to staining, illumination and image quality changes which makes it a good clinical triage tool by reliability through confidence-based selective prediction. We plan to improve classification of more difficult cell types and to test the system on whole-slide images and other cytology datasets. We shall also consider federated learning, semi-supervised learning, and inference efficiency for real-time deployment. Finally, we will work on interpretable visualizations for clinicians that would help explainability. All these will help develop SiCal-CytoNet further as reliable, scalable AI-assisted cervical cancer screening solution.

References

1. S. K. Mathivanan, D. Francis, S. Srinivasan, V. P. K. Khatavkar, and M. A. Shah, “Enhancing cervical cancer detection and robust classification through a fusion of deep learning models,” *Scientific Reports*, vol. 14, art. no. 10812, 2024. https://doi.org/10.1038/s41598-024-61063-w
2. M. K. Nour, I. Issaoui, A. Edris, A. Mahmud, M. Assiri, and S. S. Ibrahim, “Computer aided cervical cancer diagnosis using gazelle optimization algorithm with deep learning model,” *IEEE Access*, vol. 12, pp. 13046–13054, 2024. https://doi.org/10.1109/ACCESS.2024.3351883
3. I. Pacal, “MaxCerVixT: A novel lightweight vision transformer-based approach for precise cervical cancer detection,” Knowledge-Based Systems, vol. 289, art. no. 111482, 2024. https://doi.org/10.1016/j.knosys.2024.111482
4. J. Singh, R. K. Kaliyar, R. Kumari, N. Sharma, and P. D. Prasad, “Advancements in early detection of cervical cancer using machine learning and deep learning models for cervicography analysis,” *International Conference on Emerging Innovations and Advanced Computing (INNOCOMP)*, Bengaluru, India, pp. 252–256, 2024. https://doi.org/10.1109/INNOCOMP63224.2024.00049
5. H. A. Sarhangi, D. Beigifard, E. Farmani, and H. Bolhasani, “Deep learning techniques for cervical cancer diagnosis based on pathology and colposcopy images,” *Informatics in Medicine Unlocked*, vol. 47, art. no. 101503, 2024. https://doi.org/10.1016/j.imu.2024.101503