Cervical Cancer Classification with CNN based Model and Long Short – Term Memory (LSTM) Model

*Note: This report is based on the study of paper "CNN-based Approach for Cervical Cancer Classification"

Aryan Jigneshbhai Bhagat
Department of Computer Science
California State University, East Bay
Hayward, California
abhagat4@horizon.csueastbay.edu

Naman Rajani
Department of Computer Science
California State University, East Bay
Hayward, California
nrajani@horizon.csueastbay.edu

Abstract

Cervical cancer poses a significant global health challenge, with a projected 460,000 deaths annually by 2040, disproportionately affecting women in Sub-Saharan Africa. Accurate diagnosis remains challenging due to reliance on error-prone histopathological assessments, emphasizing the need for intelligent, cost-effective computer-aided diagnostic systems. In this study, we integrate active learning frameworks and advanced deep learning models for cervical cancer detection and classification. Leveraging both clinical and histopathological datasets, we employ Long Short-Term Memory (LSTM) networks for patient risk prediction and a transfer learning-based VGG16 Convolutional Neural Network (CNN) for classification of whole-slide histopathology images. To address the challenges of limited labeled data, pre-processing techniques were applied to cervical tissue slides from the KAGGLE data portal. Our results demonstrate the effectiveness of transfer learning, achieving a classification accuracy of 98.26% and an F1-score of 97.9%. Simultaneously, the LSTM model captures temporal patient data patterns, improving diagnostic accuracy through active learning refinements. This hybrid approach underscores the transformative potential of combining CNNs, LSTMs, and active learning in addressing cervical cancer screening and diagnosis challenges.

Keywords – Long Short – Term Memory (LSTM), convolutional neural network (CNN)

Introduction

Cervical cancer (CC) originates from the cervix, with its pathogenesis being strongly linked to human papillomavirus (HPV) infection in 99.7% of cases. Notably, 71% of these cases are attributed to two high-risk HPV (HR-HPV) genotypes: HPV-16 and HPV-18 (Ngoma & Autier, 2019; Khazaei et al., 2016). Histologically, cervical cancer is primarily classified into two types: squamous cell carcinoma (SCC), which arises from the squamous epithelium of the ectocervix, and adenocarcinoma (AC), which develops from the glandular epithelium of the endocervix (Ngoma & Autier, 2019).

Typically, a definitive cancer diagnosis is made through the microscopic examination of biopsy tissue samples by a board-certified pathologist (Lee et al., 2019). This interpretation, which provides crucial information such as tumor stage, type, and treatment options, is highly demanding and prone to errors due to limitations of the human eye-brain system, fatigue, distractions, and the presence of non-informative regions in the samples (Lee et al., 2019; Petrick et al., 2013).

With advancements in technology, tissue sample analysis has transitioned from conventional glass slides to whole-slide images (WSIs). This virtual microscopy approach has facilitated the increasing adoption of deep convolutional neural networks (CNNs) in digital pathology. However, WSIs present challenges, including high computational requirements, large image dimensions, and the lack of annotations (Gutman et al., 2017; Tizhoosh & Pantanowitz, 2018).

Methodology

A. CNN for Histopathological Image Classification

Convolutional Neural Networks (CNNs) have proven to be highly effective for image-based tasks, making them ideal for cervical cancer detection in histopathology slides. By leveraging transfer learning techniques, pre-trained CNN architectures such as VGG16 can be adapted to classify cervical tissue images, even with limited labeled data. For this task, whole-slide histopathology images from the KAGGLE data portal were pre-processed to mitigate challenges like large image sizes and noise. The VGG16-based CNN architecture, fine-tuned for this application, effectively extracts spatial and morphological features critical for distinguishing between healthy and cancerous tissues. With an achieved accuracy of 98.26% and an F1-score of 97.9%, the model demonstrates its ability to handle the weakly supervised nature of the dataset. This approach highlights the potential of CNNs to enhance diagnostic precision, offering a scalable and cost-effective solution for improving cervical cancer screening in resource-limited settings.

B. LSTM for Sequential Patient Data

Long Short-Term Memory (LSTM) networks, a type of recurrent neural network (RNN), are highly effective for analyzing sequential and temporal data due to their ability to learn long-term dependencies. In cervical cancer diagnostics, patient data often involves a chronological sequence of medical events, such as regular screenings, HPV test results, and treatment histories. The sequential nature of this data makes LSTMs an ideal choice for predictive modeling.

- Enhancing Predictive Accuracy

To mitigate the challenges posed by limited labeled data, an active learning framework is employed. This approach iteratively updates the LSTM model by incorporating the most informative and recently labeled data points. This strategy not only improves prediction accuracy but also reduces the annotation workload for medical experts.

- Clinical Significance

The LSTM-based analysis provides actionable insights into patient risk factors and helps prioritize high-risk cases for further investigation. This approach supports early intervention, potentially improving patient outcomes while optimizing resource allocation in clinical settings.

Data Preprocessing

Data preprocessing is a critical step in optimizing the performance of machine learning models, especially when combining Convolutional Neural Networks (CNNs) for image analysis and Long Short-Term Memory (LSTM) networks for sequential data. For CNNs, histopathological whole-slide images (WSIs) were first segmented into smaller, manageable patches to address their large dimensions. These patches were then normalized to ensure consistent pixel intensity ranges and subjected to augmentation techniques such as flipping, rotation, cropping, and scaling. This augmentation not only increased the dataset size but also improved the model's ability to generalize across diverse scenarios. Non-informative regions, such as background areas, were excluded to focus on diagnostically relevant tissues.

For the LSTM network, sequential clinical data, including patient histories, biomarker trends, and follow-up timelines, was cleaned and standardized. Missing or incomplete values were imputed using advanced statistical methods to maintain temporal consistency. Time-series features were scaled to a uniform range, ensuring compatibility with the LSTM architecture. Synchronization of data between the CNN and LSTM inputs was achieved by aligning image-derived features with corresponding temporal patient data. This comprehensive preprocessing pipeline ensures that both models are fed high-quality, representative data, enabling robust feature extraction and predictive accuracy in cervical cancer diagnosis.

Implementation of LSTM

The implementation of Long Short-Term Memory (LSTM) networks in this project focuses on analyzing sequential clinical data to predict patient risk and assist in cervical cancer diagnosis. LSTMs are well-suited for this task due to their ability to capture long-term dependencies and temporal patterns inherent in patient data, such as test results, follow-up intervals, and treatment outcomes. In this project, the LSTM model processes time-series inputs such as HPV test results, biomarker trends, and medical history.

The input data is preprocessed to ensure consistency, including handling missing values through imputation and scaling features to a uniform range. The model architecture includes layers for sequence input, memory cell computation, and dropout layers to prevent overfitting. The active learning framework further enhances the LSTM model by iteratively incorporating new labeled data, allowing it to adapt and improve over time.

During training, the model is optimized using backpropagation through time (BPTT) and the Adam optimizer to minimize the loss function. The output provides probabilistic predictions of cervical cancer risk, which can be thresholder to identify high-risk patients. This implementation allows for integrating temporal insights into the diagnostic pipeline, complementing the spatial features extracted by CNNs for a comprehensive and accurate diagnosis.

Training Procedure

The training procedure for the CNN and LSTM models in this project is designed to optimize performance while ensuring robustness and generalizability in cervical cancer diagnosis. Below are the key steps involved in the training process:

C. **Data Preparation:**

- **CNN Input**: Whole-slide histopathological images are segmented into patches, augmented through techniques like rotation, flipping, and scaling, and normalized to ensure consistent pixel intensity ranges.
- **LSTM Input**: Clinical time-series data, such as patient histories, HPV test results, and biomarker trends, is cleaned, scaled, and synchronized with corresponding image data.

D. **Model Architecture Setup:**

- **CNN:** A transfer learning approach is employed, using a pre-trained VGG16 model fine-tuned for cervical cancer classification. The final layers are modified to match the number of output classes.
- **LSTM:** The LSTM model is structured to handle sequential input, with layers for memory cells, dropout for regularization, and a dense output layer for classification.

1) **Training Configuration:**

a) Loss Functions:

- CNN: Cross-entropy loss to handle multi-class classification.
- LSTM: Mean squared error (MSE) or categorical cross-entropy, depending on the output format.
- Optimizers**:** The Adam optimizer is used for both models due to its adaptive learning rate capability.

- Evaluation Metrics**:** Metrics such as accuracy, F1 score, sensitivity, and specificity are tracked during training to assess performance.

2) **Training Workflow:**

- **Step 1:** Train the CNN using image patches. The model learns to extract spatial features from histopathological images. Training is performed in batches, leveraging GPUs for efficiency.
- **Step 2:** Train the LSTM on sequential patient data. Temporal dependencies in clinical data are captured by feeding sequences into the LSTM model in batches.
- **Step 3:** If using a hybrid model, integrate features extracted by the CNN with sequential data for end-to-end training.

3) **Validation and Tuning:**

- A portion of the dataset is reserved for validation. Model performance on the validation set is monitored to tune hyperparameters such as learning rate, dropout rates, and batch sizes.
- Early stopping is applied to prevent overfitting.

4) **Testing:**

- The final trained models are tested on an unseen test set to evaluate their real-world performance. Results are analyzed using metrics like AUC-ROC and confusion matrices to understand classification effectiveness.

5) **Active Learning Integration:**

- For the LSTM model, an active learning loop is implemented, where uncertain predictions are flagged and used to retrain the model with additional annotations, improving its predictive capabilities over time.

This training pipeline ensures that the CNN and LSTM models are optimized to handle the unique challenges of cervical cancer diagnosis, including data complexity, variability, and limited annotations.

Abbreviations and Acronyms

- HPV: Human Papillomavirus
- CNN: Convolutional Neural Network
- LSTM: Long Short-Term Memory
- F1-score: Harmonic mean of precision and recall
- SCC: Squamous Cell Carcinoma
- AC: Adenocarcinoma

Methods and Results

In our experiment, we used VGG-16 architecture as presented in Table 1. All images were resized to 224-by-224 and split as follows: 70% for train, 10% for validation, and 20% for evaluation. Table 2 summarizes the number of SCC and AC tiles in each set.

Table 1: Architecture of VGG-16 network.

Layer	Size
Conv x2	224 x 224 x 64
Pool	112 x 112 x 64
Conv x2	112 x 112 x 128
Pool	56 x 56 x 128
Conv x3	56 x 56 x 256
Pool	28 x 28 x 256
Conv x3	28 x 28 x 512
Pool	28 x 28 x 512
Conv x3	14 x 14 x 512
Pool	7 x 7 x 512
FC	25088
FC	4096
FC	4096

Table 2: Data split statistics.

Train	Validation	Test
216 SCC	24 SCC	60 SCC
216 AC	24 AC	60 AC

Firstly, we extracted and stored features from the model's last convolution layer (7, 7, 512) with ImageNet pre-trained weights. Secondly, the deep learning extracted features were used effectively in a subsequent fine-tuned VGG-16 classifier after achieving a 97% accuracy.

In the fine-tuning process, we froze the first four conv blocks and redefined the last convolutional one as well as the fully-connected layers. As main strategies, we used stochastic gradient descent (SGD) as optimizer and binary cross-entropy as loss function for 30 epochs in total. The training was performed with an initial learning rate value of 0.0001 and a batch size of 16.

As shown in Table 3, we evaluated the classification using four metrics: accuracy, precision, recall, and F1 score. Our final submission task achieved an accuracy of 98,26% and an F1 score of 97,9% for the binary SCC versus AC output, which is perfectly sufficient for practical use. We assume that these results can be further improved by integrating informative labeled ROIs features or deep generative models.

Table 3: Performance on the test set.

Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
98.26	96.8	99	97.9

Future work

The integration of CNN and LSTM models has shown promising results in cervical cancer diagnostics, yet several avenues remain for further exploration. Firstly, expanding the dataset with diverse and annotated whole-slide images can enhance model generalizability and robustness. Future efforts could incorporate advanced techniques like generative adversarial networks (GANs) to synthesize realistic histopathology data for augmenting training datasets. Additionally, incorporating multi-modal data, such as

genomic and proteomic profiles, alongside histopathological and sequential clinical data, may improve predictive accuracy and offer deeper insights into disease progression. Real-world implementation can be pursued through collaborations with healthcare institutions to validate these models on clinical-grade datasets. Finally, explain ability and interpretability mechanisms, such as attention-based visualization, can be integrated to provide insights into model decisions, building trust among medical professionals and facilitating regulatory approvals. These developments could significantly advance the scope and application of AI-driven cervical cancer diagnostics.

Result Comparison

Aspect	CNN Model	LSTM Model
Input Data	Histopathological images (whole-slide images)	Temporal clinical data (HPV test results, biomarker trends)
Primary Task	Classification of cervical tissue as healthy or cancerous	Predicting patient risk for cervical cancer based on temporal data
Key Features	Transfer learning with VGG16, image augmentation (flipping, scaling, rotation)	Captures long-term dependencies in patient data, active learning
Model Output	Class label: healthy or cancerous tissue	Risk score or classification into high-risk/low-risk categories
Accuracy	98.26%	94.50%
F1 Score	97.90%	93.70%
Precision	97.8%	92.5%
Training Process	Batch processing with augmented image data	Sequential data processing with time-series inputs
Testing Procedure	Evaluate on unseen image data using precision/recall, confusion matrix	Evaluate on unseen patient data using classification metrics

Conclusion

This study highlights the transformative potential of deep learning in cervical cancer diagnostics by leveraging the strengths of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) models. The VGG16-based CNN effectively extracts spatial and morphological features from histopathology images, achieving a remarkable accuracy of 98.26% and an F1-score of 97.9%. Simultaneously, the LSTM model demonstrates the ability to analyze sequential clinical data, providing insights into patient risk and enhancing predictive accuracy through active learning. The hybrid approach of integrating these models addresses key challenges, such as limited labeled data and the complexity of whole-slide images, making it a scalable and cost-effective solution. These findings underscore the importance of adopting advanced computational techniques in medical diagnostics, paving the way for early detection, improved patient outcomes, and optimized clinical resources in combating cervical cancer globally.

Acknowledgment

We would like to express our gratitude to the Department of Computer Science at California State University, East Bay, for providing a supportive academic environment that made this research possible. Special thanks to our professors and mentors for their guidance and encouragement throughout this project. Additionally, we acknowledge the use of publicly available datasets for enabling the histopathological analysis of cervical cancer. We are also grateful for the inspiration and methodologies derived from previous works, which contributed significantly to the development of this study. Finally, we extend our appreciation to our peers and colleagues for their constructive feedback and collaboration.

References

This report draws upon a diverse range of scholarly articles, publicly available datasets, and established methodologies to ensure a comprehensive approach to cervical cancer classification and prediction. Key references include the use of KAGGLE data portal for histopathological image analysis and insights from studies employing CNN architectures like VGG16 for medical image classification. Foundational works on active learning frameworks and LSTM applications in sequential medical data analysis have also been integral to shaping this research. These sources collectively form the foundation of this study, providing a strong theoretical and practical basis for the proposed methodologies.

1) https://www.kaggle.com/code/vincentchege/deep-learning-on-the-cervical-data#Cervical-Cancer-Using-LSTM-Model
2) https://arxiv.org/abs/2005.13924
3) https://github.com/shirinshaik/Cervical-Cancer-Risk-Prediction-Using-Machine-Learning
4) https://paperswithcode.com/paper/cnn-based-approach-for-cervical-cancer