Illuminating CNN Decisions: Class Activation Maps for Cervical Cancer Analysis

ABSTRACT

Convolutional Neural Networks (CNNs) are popular in medical imaging for disease classification, but their non-interpretability affects trust and dependability. Class Activation Maps (CAMs) offer a means to visualize the parts of an image that contribute the most to a model's decision, making it more explainable. By using a Global Average Pooling (GAP) layer following the last convolutional layer, CAM produces class-discriminative saliency maps, providing model prediction insight. Grad-CAM and Grad-CAM++ build upon CAM by employing gradient-based methods to enhance explainability without altering the architecture of CNN. In this project, we implement CAM, Grad-CAM, and Grad-CAM++ on a cervical cancer dataset with the goal of emphasizing cellular abnormalities in Pap smear images. This ensures that the model is attending to informative medical features and not uninformative artifacts, thus enhancing confidence in automatic cervical cancer classification. These approaches, however, have their drawbacks. CAM needs architectural changes, whereas Grad-CAM and Grad-CAM++ depend on backpropagation, thus being computationally costly. Grad-CAM++ improves localization but can still fail to handle overlapping cell structures and fine morphological variations. These methods also fail to explain fully high-level decisions, at times resulting in deceptive activations.

PROJECT OBJECTIVES:

This project aims to improve the interpretability of CNN-based cervical cancer classification by utilizing Class Activation Maps (CAMs), Grad-CAM, and Grad-CAM++. These methods offer visual explanations by highlighting the most important regions in medical images. By incorporating explainability techniques, it intends to enhance trust and transparency in AI-assisted cervical cancer diagnosis, allowing clinicians to validate model predictions and confirm that AI systems align with medical expertise. The implementation will be conducted using PyTorch, showcasing how CNN explainability can be applied to real-world cancer detection tasks.

METHODOLOGY:

Tech Stack for the Project

1. Programming Language & Platform: Python, Google Colab
2. Libraries for Visualization: Matplotlib, Seaborn, OpenCV
3. Model Interpretability: Captum (for PyTorch), tf-explain (for TensorFlow)
4. Dataset: Cervical cancer dataset (e.g., Pap smear images)

Step 1: Curate high-quality cervical cancer image datasets, enhance them through augmentation, and prepare them for deep learning models.

Step 2: Choose a robust CNN architecture, fine-tune hyperparameters, and train it to classify cervical cancer images with high precision.

Step 3: Extract deep feature representations to generate heat maps highlighting the most influential regions in CNN decision-making.

Step 4: Overlay CAM heatmaps on original images to gain insights into model focus areas, improving trust and transparency in AI-driven diagnosis.

Step 5: Refine the CNN by adjusting hyperparameters, integrating advanced interpretability techniques, and enhancing predictive accuracy.

Step 6: Test the model in real-world scenarios, develop a clinician-friendly tool, and deploy it for aiding cervical cancer screening and diagnosis.

KEY FINDINGS:

- Class Activation Map (CAM) uses Global Average Pooling (GAP) to generate class-specific saliency maps by applying learned weights to feature maps.
- Grad-CAM improves interpretability by utilizing gradients from the last convolutional layer to highlight important regions without modifying the model.
- Grad-CAM++ enhances localization by assigning higher scaling factors to more relevant pixels using higher-order gradients, making it more precise for fine-grained feature detection.

REFERENCES:

Bueno-Crespo, A. *et al.* (2024). Diagnosis of Cervical Cancer Using a Deep Learning Explainable Fusion Model. In: Ferrández Vicente, J.M., Val Calvo, M., Adeli, H. (eds) Bioinspired Systems for Translational Applications: From Robotics to Social Engineering. IWINAC 2024. Lecture Notes in Computer Science, vol 14675. Springer, Cham. https://doi.org/10.1007/978-3-031-61137-7_42

TEAM:

1. Ishika Singhal(E22CSEU1542)
2. Anushka Shishodia(E22CSEU1014)