***DEEPLEARNING PROJECT***

*Comparative Study of Image Classification Using VGG19 and EfficientViT-B0 for Benign or Malignant Mouth ulcer Detection*

*Krutika Deshmukh*

***ABSTRACT:***

*This project presents a comparative study of two pre-trained deep learning models—VGG19 and EfficientViT-B0—applied to binary image classification using an oral disease image dataset containing two classes: benign and malignant. The goal is to classify the images into these categories and evaluate the models using key metrics such as accuracy, precision, recall, F1-score, MCC, and confusion matrix. Data augmentation techniques were employed to enhance model generalization. In addition to training on the provided dataset, random images were sourced from Google and tested for predictions to assess the models' ability to generalize to unseen data. This report discusses the implementation, performance comparison, and detailed analysis of the two models.*

***DATA SET DESCRIPTION:***

*My dataset contains a total of* ***323 images****, with* ***165 benign lesions*** *and* ***158 malignant lesions****. These images were divided into* ***80% for training*** *and* ***20% for testing****. Additionally,* ***10 random images from Google*** *were taken and tested for predictions to evaluate the models' performance on unseen data.*

***MODEL TRAINING PROCESS FOR EFFICIENTVIT-B0 AND VGG19:***

1. *DATA PREPROCESSING AND AUGMENTATION****:*** *To improve model generalization, various data augmentation techniques were applied to the training dataset. The same augmentation techniques were used for both models*

***A screenshot of a computer program

Description automatically generated***

1. *MODEL ARCHITECTURES:*

***VGG19*** *is a deep convolutional neural network consisting of 19 layers. For this study, we used a pre-trained VGG19 model and modified the classifier to handle binary classification (benign vs. malignant). For VGG 19 Last 3 layers retrained to adapt to the dataset.*

A screenshot of a computer

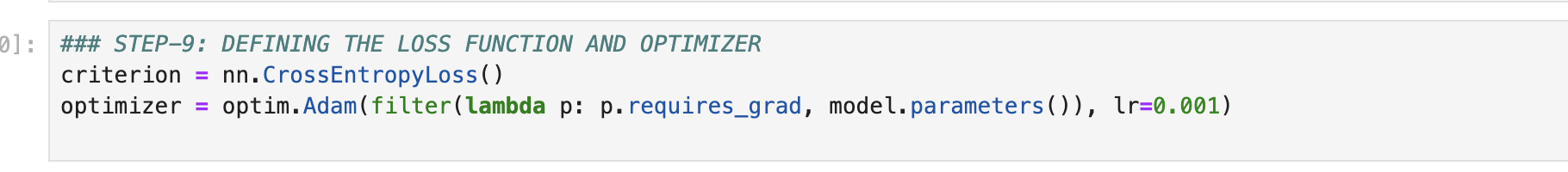
Description automatically generated

**EfficientViT-B0** is a modern, lightweight transformer-based architecture optimized for efficiency. It was also modified for binary classification. *All layers frozen except the final classifier. Pre-trained weights loaded, and only the classifier layer trained for binary classification (benign vs. malignant).*

A screenshot of a computer program

Description automatically generated

3)*TRAINING AND EVALUATION: Both models were trained using the Adam optimizer with a learning rate of 0.001. The cross-entropy loss function was used for both models. Training and evaluation were conducted using the same training and testing dataset split (80% training, 20% testing)*

**

*4. Results****:***

***Loss and Accuracy Over Epochs:***

*For* ***VGG19****, training was conducted over* ***10 epochs****, while for* ***EfficientViT-B0****, it was conducted over* ***5 epochs****. The training and validation loss and accuracy were recorded at each epoch for both models*

***Loss Graphs for VGG19 and EfficientViT-B0****:*

1. *VGG-19*

*A graph with blue and orange lines

Description automatically generated*

1. *FOR EFFICIENTVIT-B0:*

*A graph with a line and a line

Description automatically generated*

* *The VGG19 model exhibited a more stable training process with decreasing training and test losses, which suggests that it generalized well to the test set.*
* *The EfficientViT-B0 model, while quickly reducing the training loss, struggled with generalization, as seen in the increase in test loss after initial epochs, likely due to overfitting.*

***Accuracy Graphs for VGG19 and EfficientViT-B0****:*

*EfficientViT-B0 Accuracy :*

***A graph with a line

Description automatically generated***

* *The* ***training accuracy*** *consistently increased over the 5 epochs, approaching near perfect accuracy by the final epoch.*
* *The* ***test accuracy****, however, followed an irregular pattern. After an initial increase up to epoch 3, the test accuracy declined in later epochs, indicating that the model may have started overfitting to the training data.*

*VGG19 Accuracy:*

***A graph with blue and orange lines

Description automatically generated***

* *The* ***training accuracy*** *for the VGG19 model showed a steady and smooth increase, reaching around 85% by epoch 10.*
* *The* ***test accuracy*** *followed a more stable trend compared to EfficientViT-B0, peaking around epoch 6 and showing a minor decrease in later epochs. However, the gap between the training and test accuracy was narrower than EfficientViT-B0, indicating a more balanced performance and less overfitting.*

***CONFUSION MATRIX***

*To further analyze the model’s performance, confusion matrices were generated for both models, highlighting the number of correctly and incorrectly classified images.*

*For VGG19:*

*A diagram of a confusion matrix

Description automatically generated*

*For EfficientViT-B0:*

*A blue squares with white text

Description automatically generated*

*VGG19 is better at detecting malignant lesions, with fewer false negatives but more false positives. EfficientViT-B0 excels at identifying benign cases, with fewer false positives but struggles more with malignant lesions, showing more false negatives. Each model has strengths in different aspects of classification.*

***PERFORMANCE METRICS*** *:*

*Metric VGG19 EfficientViT-B0*

*Accuracy 0.7385 0.7077*

*Precision 0.7179 0.8125*

*Recall 0.8235 0.4483*

*F1 Score 0.7671 0.5778*

*MCC 0.4778 0.4211*

*VGG19 shows a better balance between precision and recall, with a higher F1 score and MCC (0.4778), indicating more reliable performance in identifying malignant cases and overall classification consistency. EfficientViT-B0, while having higher precision, has a lower recall and F1 score, and a lower MCC (0.4211), suggesting it is more conservative in classifying malignant cases but misses many, making it less balanced than VGG19.*

***Results of Prediction on 10 Images:***

*Ten random benign and malignant images were collected, and the VGG19 model correctly predicted 5 out of 10 images, whereas the EfficientViT-B0 model only predicted 1 out of 10 images correctly.*

*For EfficientvitB0:*

*Close-up of a person's tongue with a sore

Description automatically generated*

*The performance difference between VGG19 and EfficientViT-B0 can be attributed to their architectures and the small dataset size. VGG19, with its deeper and more complex structure, excels at capturing detailed features, enabling better generalization despite the limited data. In contrast, EfficientViT-B0, optimized for efficiency, struggles to capture the same level of detail, resulting in more missed malignant cases. The small dataset further exacerbates this, as it restricts the models' ability to learn diverse patterns. VGG19's higher MCC score and its better balance between precision and recall demonstrate its superior performance, especially when handling new, unseen images.*

*Close-up of a person's tongue with a sore

Description automatically generated*

***LIMITATIONS OF MY PROJECT :***

***Small Dataset Size****: With only 323 images, the dataset limits the models' ability to generalize, leading to potential overfitting, especially in deeper models like VGG19.*

***Limited Variability****: Lack of diverse conditions (e.g., lighting, angle) in the dataset restricts the models’ ability to learn robust features for accurate classification.*

***EfficientViT-B0 Underperformance****: EfficientViT-B0's simpler architecture struggles with the complexity of the dataset, resulting in lower recall and accuracy, particularly for malignant cases.*

***Random Image Testing:*** *Google-sourced random images may come from different environments, causing prediction inaccuracies as the models were not trained on such uncontrolled data.*

***IMPROVEMENTS FOR MY PROJECT :***

***Preprocessing Enhancement****: Advanced normalization and noise reduction techniques could help the models focus better on lesions, improving accuracy.*

***Transfer Learning:*** *Fine-tuning the models on a larger, similar medical dataset could enhance their ability to learn more detailed features before applying them to the smaller dataset.*

***Exploring Other Models****: Trying other architectures like ResNet or DenseNet could improve performance, especially where EfficientViT-B0 struggles****.***

***REFERENCES:***

[*https://pytorch.org/vision/stable/models.html*](https://pytorch.org/vision/stable/models.html)

[*https://github.com/Priyanshu9898/Oral-Disease-Classification/blob/main/best\_model/efficientvit\_b0\_oral\_disease\_classifier.pth*](https://github.com/Priyanshu9898/Oral-Disease-Classification/blob/main/best_model/efficientvit_b0_oral_disease_classifier.pth)

[*https://data.mendeley.com/datasets/mhjyrn35p4/2*](https://data.mendeley.com/datasets/mhjyrn35p4/2) *--------dataset link*

[*https://www.mdpi.com/2075-4418/13/21/3360*](https://www.mdpi.com/2075-4418/13/21/3360) *-----------article link*

[*https://www.geeksforgeeks.org/sklearn-classification-metrics/*](https://www.geeksforgeeks.org/sklearn-classification-metrics/)

[*https://www.geeksforgeeks.org/vgg-net-architecture-explained/*](https://www.geeksforgeeks.org/vgg-net-architecture-explained/)

[*https://github.com/huggingface/pytorch-image-models*](https://github.com/huggingface/pytorch-image-models)

*----also took the help of Cht gpt*