Breast Cancer Histology Image Classification Using Deep Learning Approaches

Fatemeh Ranjbaran
fatemeh.ranjbaran@studio.unibo.it

Introduction

Breast cancer remains one of the most common types of cancer globally, necessitating accurate and efficient diagnosis methods. I endeavored to streamline this process through deep learning techniques, tackling the BACH Grand Challenge, which focuses on the automated classification of breast cancer histology images.

Dataset and Preprocessing

The BACH challenge dataset comprises of histology images classified into four categories: Normal tissue, Benign lesion, In situ carcinoma, and Invasive carcinoma. A vital step in achieving optimal performance of deep learning models is data preprocessing. Here, each image was appropriately resized, normalized, and augmented to create a robust dataset, helping to improve the model's ability to generalize.

Model and Training

Two distinct deep learning architectures were employed: Vision Transformer (ViT) with Multihead Attention, and ResNet with Attention. The ViT architecture, derived from transformers widely used in Natural Language Processing (NLP), demonstrated efficacy in vision tasks. The addition of the Multihead Attention mechanism allowed the model to focus on different image regions simultaneously, resulting in more nuanced image comprehension.

Conversely, ResNet is a convolutional neural network (CNN) model, known for its efficacy in image classification tasks. Here, an Attention mechanism was incorporated, allowing the model to weigh the importance of features within the images.

For both models, the Adam optimizer was used, characterized by adaptive learning rates for different parameters. A learning rate of 0.0001 was employed, along with Cross-Entropy Loss as the objective function, a standard for multi-class classification problems.

Results

The ViT model and the ResNet Attention model were tested on the unseen data, where they demonstrated the following performance metrics:

For the ViT model:

Class 0 (Normal tissue): Accuracy: 0.8500, Precision: 0.7778, Recall: 0.7000, F1-Score: 0.7368
Class 1 (Benign lesion): Accuracy: 0.8500, Precision: 0.7895, Recall: 0.9375, F1-Score: 0.8571
Class 2 (In situ carcinoma): Accuracy: 0.8500, Precision: 1.0000, Recall: 0.7778, F1-Score: 0.8750
Class 3 (Invasive carcinoma): Accuracy: 0.8500, Precision: 1.0000, Recall: 1.0000, F1-Score: 1.0000

For the ResNet Attention model:

Class 0 (Normal tissue): Accuracy: 0.4500, Precision: 0.3333, Recall: 0.1818, F1-Score: 0.2353
Class 1 (Benign lesion): Accuracy: 0.4500, Precision: 1.0000, Recall: 0.1538, F1-Score: 0.2667
Class 2 (In situ carcinoma): Accuracy: 0.4500, Precision: 0.6000, Recall: 0.8571, F1-Score: 0.7059
Class 3 (Invasive carcinoma): Accuracy: 0.4500, Precision: 0.3636, Recall: 0.8889, F1-Score: 0.5161

In comparison to previous studies such as "Ensemble Network for Region Identification in Breast Histopathology Slides" (Marami et al., 2018) and "Assessment of Breast Cancer Histology Using Densely Connected Convolutional Networks" (Kohl et al., 2018) that achieved an accuracy of 0.84 and 0.83 respectively, the results of the ViT model demonstrated superior performance. Particularly noteworthy is the model's perfect precision and recall for the invasive carcinoma class (Class 3).

Meanwhile, the ResNet Attention model showcased high recall values for the cancerous classes (Classes 2 and 3), indicating its potential for identifying all positive instances, albeit with some false positives.

The following plots illustrate the performance of the models over the training epochs.

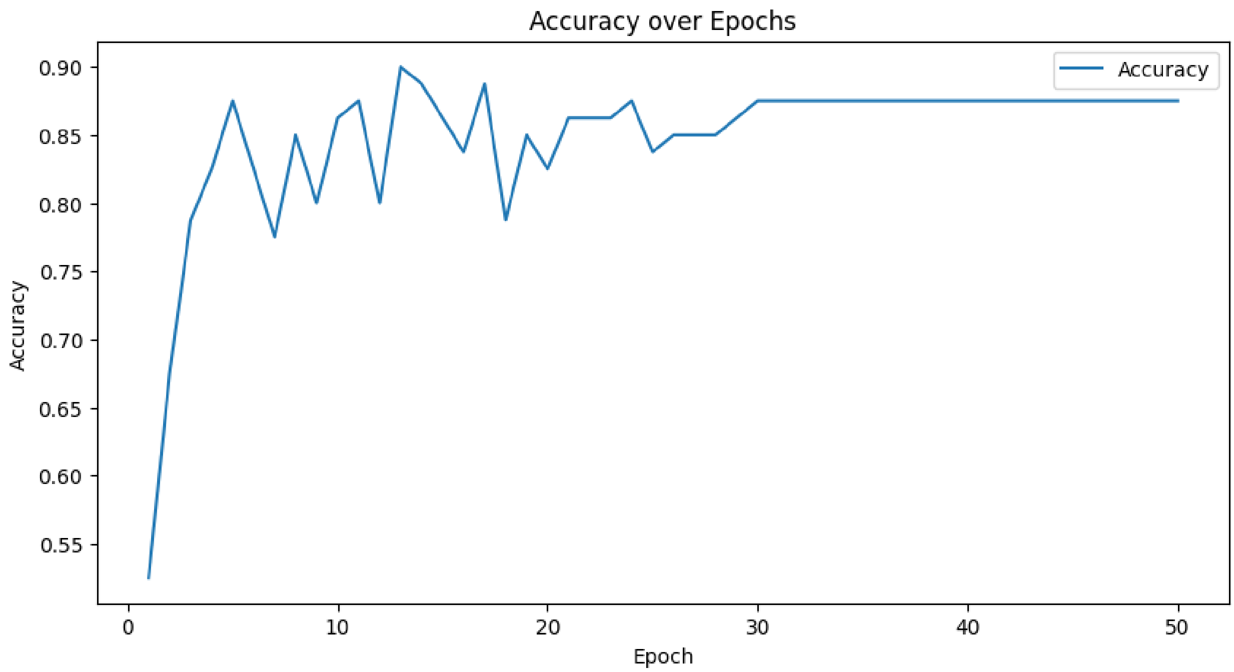

Figure 1. Accuracy of the Vision Transformer (ViT) model over training epochs.

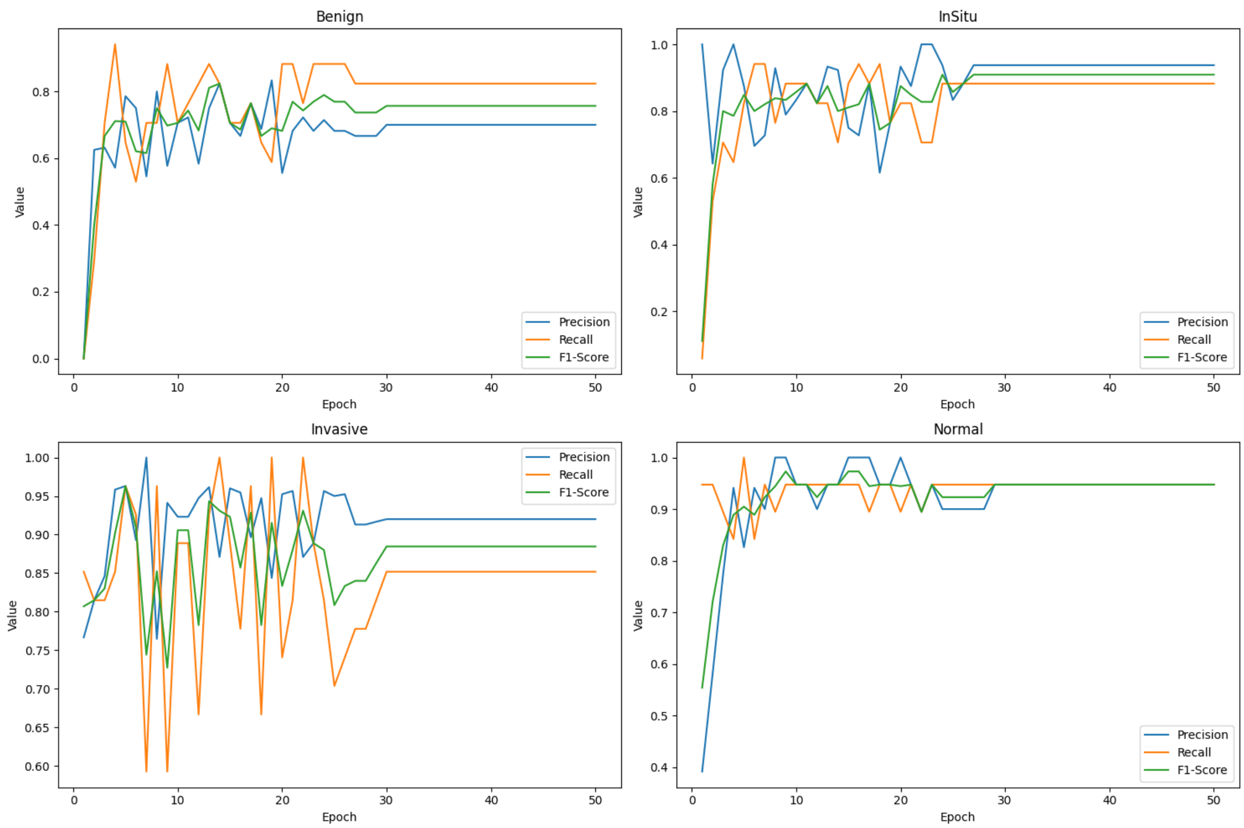

Figure 2. Precision, Recall, F1-Score: Figure Y. Precision, Recall, and F1-Score of the Vision Transformer (ViT) model over training epochs for each class.

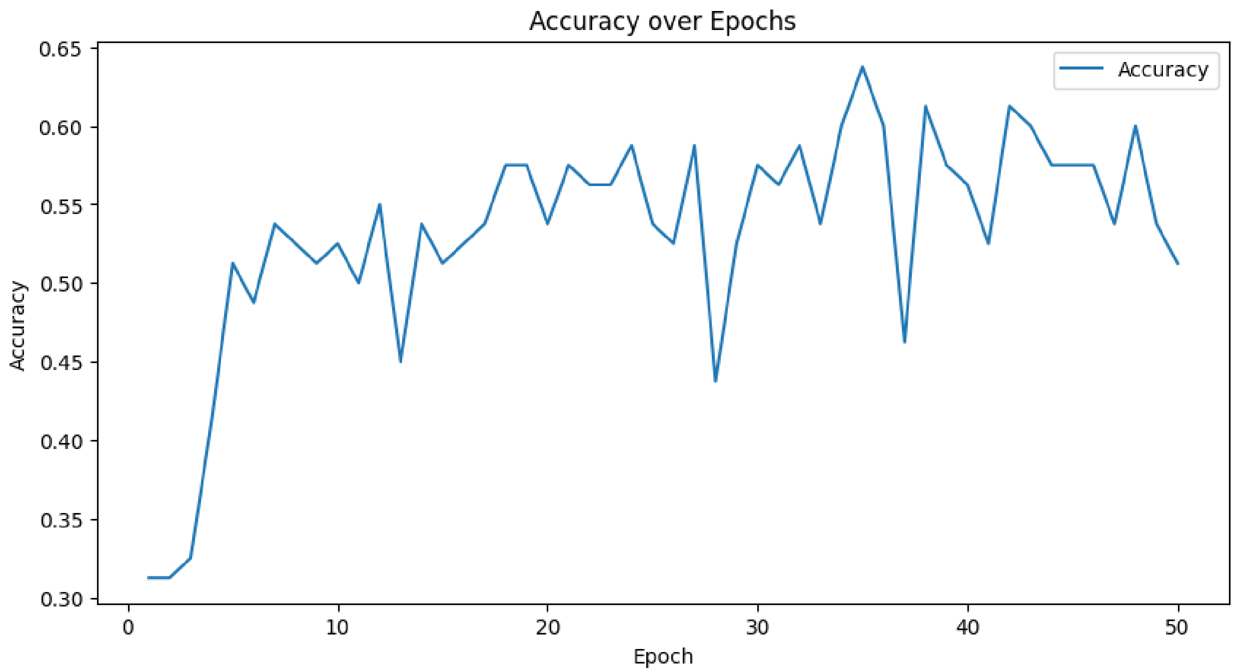

Figure 1. Accuracy of the Residual Attention Network model over training epochs.

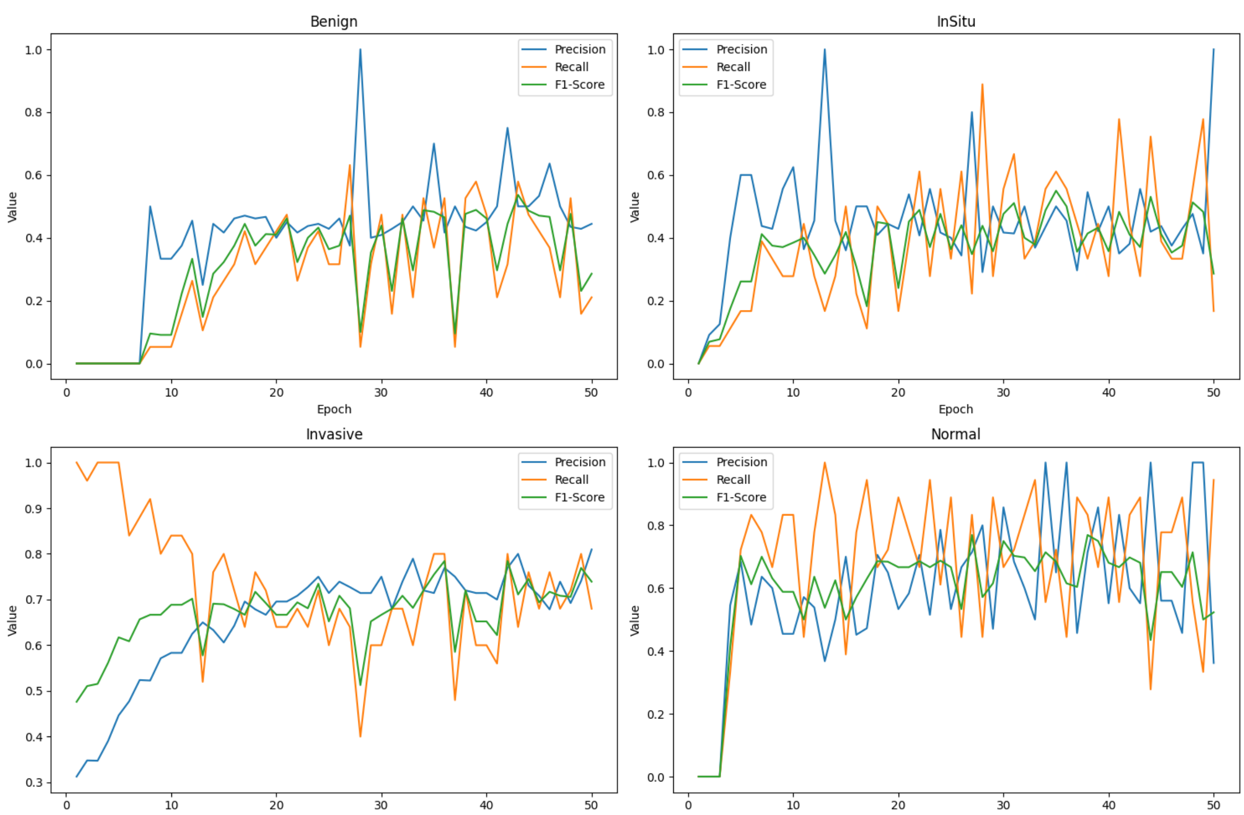

Figure 3. Precision, Recall, F1-Score: Figure W. Precision, Recall, and F1-Score of the Residual Attention Network model over training epochs for each class.

These plots provide a visual representation of the models' learning progression, with metrics such as accuracy, precision, recall, and F1-score evaluated at each epoch for each class. The varying performances across different epochs and classes underscore the complexity of the classification task and the models' adaptability in learning from the training data.

Conclusion

Both architectures demonstrated promise for this classification task, with the ViT model achieving the highest accuracy. The models' performance underscores the potential of deep learning models in automating the classification of breast cancer histology images, contributing to improved diagnosis. Future investigations may look into alternative pre-training strategies or model architectures and extended training to further enhance model performance.

Reference:

1.https://link.springer.com/chapter/10.1007/978-3-319-93000-8_103
2.https://link.springer.com/chapter/10.1007/978-3-319-93000-8_98