Mammogram Report

Sahand Hamzehei

In this project, the primary objective is to evaluate the image history of patients over time by comparing prior and current images of the same patient. To achieve this, the project consists of two main components. Initially, a baseline model was defined on a public dataset and categorized into "normal" and "abnormal" classes. This baseline model serves as the foundation for subsequent analyses. Following this, a twin network strategy should be used to assess the presence of tumors in patients. This approach involves comparing the Euclidean distance between prior and current images, utilizing convolutional neural network (CNN) layers to facilitate this comparison and prediction. The integration of these components allows for a comprehensive assessment of changes in patient images over time, highlighting the project's innovative approach to medical imaging analysis.

Evidence shows that the baseline model encounters difficulties in distinguishing between calcification (Calc) and architectural distortion (ARCH) classes. This challenge arises from the inherent variability and distinct definitions of tumors across these classes. To enhance the model's discriminative power, it is suggested that after initial training on "normal" and "abnormal" classifications, a fine-tuning phase be introduced. This phase would further categorize the "Abnormal" class into three more specific classes: "MASS," "Calc," and other cancer classifications, as permitted by the public dataset. It is also known that the level of complexity of different datasets can change the sensitivity of the baseline model. This shows how important it is to do pre-processing that is specific to each dataset to get the best results. Additionally, incorporating artifacts in approximately 10% of the dataset has been observed to significantly enhance the robustness of the model. This strategic approach aims to reduce the model's initial limitations and improve its overall ability to accurately classify a broader spectrum of tumor types.

In addition to the previously discussed strategies, it has been observed that employing the ResNet family of pre-trained models tends to guarantee better results compared to other pre-trained models. However, the choice between ResNet50 and ResNet101 remains a subject of uncertainty, with no clear consensus on which model is definitively more effective for this project's twin network. Factors such as available time, dataset size, complexity, and the extent of fine-tuning required on the twin network play a crucial role in this decision. ResNet50 is thought to be better because it is simpler and doesn't need as many features to be extracted. This makes it better for situations where time is limited, datasets are small, or the complexity is low and fine-tuning is not needed. This makes its performance reliable, though not necessarily superior in every context.

Conversely, ResNet101, with its greater capacity for extracting a higher level of features, is considered potentially more useful for situations demanding more extensive data, greater complexity, and deeper fine-tuning efforts. This model's ability to capture more detailed aspects of tumors in some images suggests that, particularly when the baseline is fine-tuned on the different classes of "abnormal," a closer examination of ResNet101's performance is highly recommended. Such an examination could reveal whether the additional complexity and feature extraction capabilities of ResNet101 translate into significant advantages for the project.

In the end, the decision on which model is more appropriate—ResNet50 or ResNet101—has not been conclusively determined and requires further study. The balance between these two models involves weighing their reliability and potential for capturing detailed tumor information against the specific needs and constraints of the project. As such, if the baseline is to be fine-tuned on varied "abnormal" classifications, it is advisable to evaluate the performance of ResNet101 closely. This approach underscores the necessity for ongoing research and experimentation to determine the most effective pre-trained model for enhancing the project's medical imaging analysis capabilities.

The utility of using pre-trained models has been proven, especially when layers are unfrozen carefully to tailor the model to specific needs. The selective unfreezing of appropriate layers from the pre-trained model allows for more effective customization and optimization of the model for the task at hand. This approach enhances the model's ability to learn from the dataset, ensuring that the nuances of the data are captured more accurately, which is critical for the success of the twin network and the overall project.

Within the twin network framework, a comparison study of Manhattan distance and Euclidean distance was carried out to see how well they work at finding similarities between images. The Manhattan distance, calculated as the sum of the absolute differences between points in a vector space, offers a linear metric of similarity. In contrast, the Euclidean distance, which measures the shortest path between two points in a multidimensional space, provides a geometric perspective of similarity.

The experimental findings from the project indicate that the Euclidean distance yields better results in the context of this twin network. The benefit of Euclidean distance can be attributed to its ability to capture the nuanced geometric disparities between image pairs, which is critical for accurately detecting tumor presence and evolution over time. Because of this, the code that is given supports both distance metrics and gives you the choice of training the model with either contrastive loss or the ADAM optimization algorithm.

The contrastive loss function, which learns embeddings by making the distance between similar pairs smaller and the distance between dissimilar pairs larger, is a better architecture for this application, according to more research.

This loss function aligns well with the project's objectives, enhancing the model's ability to discern subtle differences between image pairs. So, using the Euclidean distance metric along with the contrastive loss function is thought to be the best way to use this twin network model. This will allow for a more accurate and detailed analysis of the patient's image history.

In conclusion, the project's findings underscore the importance of aligning current and prior images of a patient to effectively isolate tumor presence prior to the model training process. This approach not only enhances the accuracy of tumor detection but also lays a foundational framework for a more informed analysis by the model. Such a strategy is critical in ensuring that temporal changes in the images are accurately captured and analyzed, thereby facilitating a more robust identification of tumor development or regression over time.

To further refine this process, the development of a new loss function emerges as a requisite advancement. This new loss function must be adept at differentiating not just between the broad categories of "normal" and "abnormal," but also within the nuanced subclasses of abnormalities. The key requirements for this loss function include the ability to handle the intricacies of medical imaging data, such as the varying appearances of tumors and the subtle distinctions between different types of abnormalities. Moreover, this loss function should optimize the model's sensitivity and specificity, thereby ensuring that it can accurately classify a wide range of tumor characteristics. The necessity for such a specialized loss function stems from the current limitations observed in generic loss functions, which may not fully capture the complexity and variability inherent in medical imaging, especially in a context that requires detailed temporal analysis.

Lastly, the exploration of image pairings, specifically Normal-Normal and Normal-Abnormal combinations, has revealed a potential challenge in the model's learning process. These pairings, while useful in certain contexts, may introduce confusion and prevent the model from learning the full spectrum of possibilities, including Abnormal-Abnormal and Abnormal-Normal scenarios. Such limitations highlight the need for a strategic approach in selecting and pairing images for training, ensuring that the model is exposed to a comprehensive range of image variations. This exposure is crucial for the model to learn and understand all possible combinations and their implications, ultimately enhancing its predictive accuracy and generalizability.