Experimental Methods Draft

Model Types

In this project, we employed several types of machine learning models to detect osteoporosis from medical imaging data. The models used include a custom Convolutional Neural Network (CNN), two pre-trained deep learning models (InceptionResNetV2 and ResNet50), and a Stacked ensemble model that combines the outputs of multiple models. Each model was selected based on its potential to effectively capture the complex features inherent in medical images (LeCun *et al*., 2015).

Architectural Design Choices

1. Custom CNN

The custom CNN model was designed with multiple convolutional layers followed by max-pooling layers to progressively reduce the spatial dimensions while extracting meaningful features (Goodfellow *et al*., 2016). The network consisted of three convolutional layers with 32, 64, and 128 filters, respectively. Each convolutional layer was followed by a Rectified Linear Unit (ReLU) activation function and a max-pooling layer with a pool size of (2, 2). A fully connected dense layer with 256 neurons and ReLU activation was added before the output layer, which used a softmax activation function to classify the images. This architecture was chosen for its balance between complexity and performance, allowing it to learn significant patterns without overfitting.

2. InceptionResNetV2 and ResNet50

Both InceptionResNetV2 and ResNet50 models were utilized due to their proven effectiveness in image classification tasks. These models were pre-trained on the ImageNet dataset and fine-tuned on our custom dataset. InceptionResNetV2 combines the Inception architecture with residual connections, providing a powerful mechanism to handle vanishing gradients and capture a wide range of features at multiple scales. ResNet50, known for its residual learning framework, enables the training of very deep networks by mitigating the vanishing gradient problem through shortcut connections. Both models were used with their convolutional base, followed by a global average pooling layer and a dense output layer with softmax activation for classification.

3. **Stacked Ensemble Model**

The Stacked ensemble model integrated the predictions of the custom CNN, InceptionResNetV2, and ResNet50 models to leverage their individual strengths (Prakash *et al*., 2023). A meta-learner model, trained on the outputs of these base models, was used to make the final predictions. This approach aimed to improve the overall performance by combining the diverse perspectives of different architectures.

Training Procedure

The training dataset was split into training, validation, and test sets with a ratio of 70:20:10, respectively. The data was augmented using techniques such as rotation, zoom, and horizontal flipping to increase the diversity and robustness of the model. The models were trained for a maximum of 20 epochs, with early stopping implemented based on validation accuracy to prevent overfitting. The batch size varied between models: 16 for InceptionResNetV2, 4 for ResNet50, and 32 for the custom CNN. The categorical cross-entropy loss function was used for training, given the classification nature of the task, and accuracy was the primary metric for evaluation.

Optimization Procedures

Hyperparameter tuning and architectural adjustments were critical for optimizing the models. Learning rates were tuned using grid search and varied between 1e-4 and 1e-3. Dropout layers were added in the custom CNN to prevent overfitting, with dropout rates ranging from 0.3 to 0.5. The optimizer used was Adam, known for its adaptive learning rate capabilities (Ogundokun *et al*., 2022). Additionally, the number of filters, kernel sizes, and activation functions were varied to find the optimal configuration. The InceptionResNetV2 and ResNet50 models were fine-tuned by unfreezing the top layers of the convolutional base and allowing them to learn along with the newly added dense layers. Regularization techniques such as L2 regularization were applied to the dense layers to further enhance generalization.

By carefully selecting model architectures, tuning hyperparameters, and employing robust training procedures, the models were optimized to achieve high accuracy in detecting osteoporosis from medical images.

References

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. Available at http://www.deeplearningbook.org

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444. https://doi.org/10.1038/nature14539

Ogundokun, R. O., Maskeliunas, R., Misra, S., & Damaševičius, R. (2022, July). Improved CNN based on batch normalization and adam optimizer. In *International Conference on Computational Science and Its Applications* (pp. 593-604). Cham: Springer International Publishing.

Prakash, J. A., Ravi, V., Sowmya, V., & Soman, K. P. (2023). Stacked ensemble learning based on deep convolutional neural networks for pediatric pneumonia diagnosis using chest X-ray images. *Neural Computing and Applications*, *35*(11), 8259-8279.