Alzheimer's Detection from Handwritten Drawings Using Multimodal Deep Learning

Ramazan Turan

AI 4 Alzheimer's Hackathon
2025

1 Introduction

The goal of this project is to provide an efficient and accessible method for estimating alzheimer disease risk without requiring laboratory-based tests, enabling low cost and scalable screening for widespread use. While exploring real world AI applications that are both practical and easily deployable, the idea of detecting alzheimer disease from handwritten drawings emerged. Hand drawn tasks are simple and inexpensive, making them suited for real life cognitive screening scenarios. During the search for a suitable dataset, the DARWIN dataset was identified as a strong fit for handwriting based alzheimer detection. While exploring access options, I found a Kaggle dataset containing a subset of the DARWIN dataset, uploaded by Francesco Fontanella, one of the authors of the original DARWIN dataset paper. This increased my confidence in the data source. Based on this, the project was developed to estimate alzheimer risk from handwritten drawings using multimodal deep learning.

2 Model Architecture

2.1 Visual Feature Extraction

Handwritten circle drawing images are processed using EfficientNetB0 as the visual feature extractor. EfficientNetB0 was selected after evaluating multiple backbone architectures, including Resnet, Mobilenet, and COnvNext, due to its superior performance on the validation set and balance between model capacity and efficiency. The convolutional feature maps produced by EfficientNetB0 are further refined using a Convolutional Block Attention Module (CBAM), which applies both channel and spatial attention.

2.2 Demographic Feature Encoding

In addition to visual information, demographic features specifically age and years of education are fed to the model. These features are processed through a lightweight multilayer perceptron (MLP) consisting of a linear layer, activation function, batch normalization, and dropout.

2.3 Multimodal Feature Fusion and Classification

The visual embeddings obtained from the attention-enhanced image encoder are concatenated with the encoded demographic features to form a multimodal representation. This fused feature vector is passed through a fully connected classification head to predict alzheimer disease risk. The final output is a binary classification indicating Healthy Control or alzheimer disease, providing efficient risk estimation.

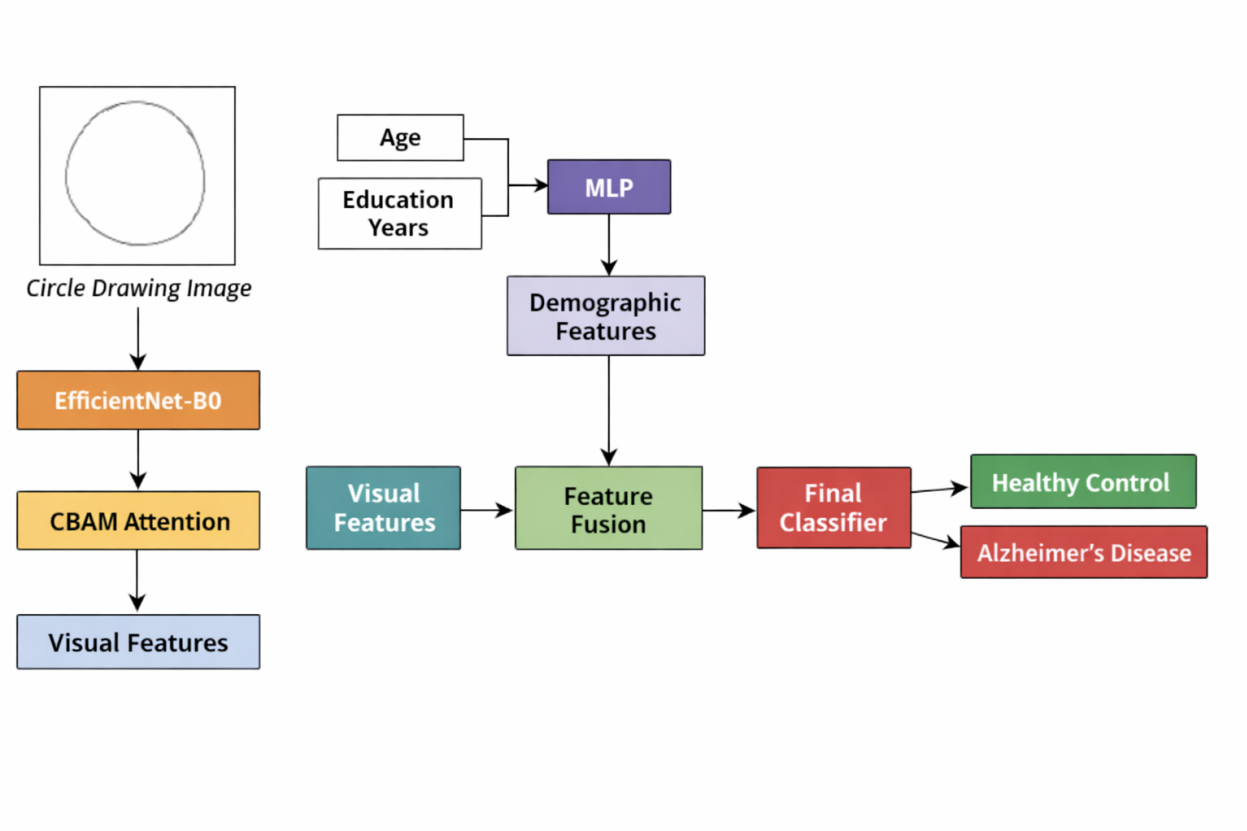

Figure 1: The overall architecture of the proposed multimodal model.

3 Results and Conclusion

3.1 Results

The multimodal model was evaluated on a held-out validation set using accuracy, ROC-AUC, sensitivity metrics. The model achieved a final validation accuracy of 92.5 and a ROC-AUC score of 0.91, demonstrating a good performance. The confusion matrix analysis revealed zero false negatives, resulting in a sensitivity (recall) of 1.00. The model achieved a specificity of 0.84, indicating a reasonable balance between false positives and true negatives.

3.2 Feature Importance Analysis

Feature importance analysis demonstrates the importance of the demographic data. I used two of the features provided in the dataset metadata, ”Age” and ”Education years”. ”Age” feature demonstrated a very strong positive correlation (0.63) with alzheimer risk. It shows, as age increases, the risk of alzheimer also increases. ”Education years” feature demonstrated a negative correlation 0.43 with alzheimer risk. It shows, as the ”Education years” increases, the risk of alzheimer decreases. These results match what is commonly observed in clinical research, showing that the model learns natural and realistic risk factors.

3.3 Attention and Model Interpretability

Grad-CAM visualizations were used to analyze model attention patterns. The results indicate that the network focuses on structural irregularities, distortions, and incomplete shapes in handwritten circle drawings when predicting Alzheimer disease. Well formed and symmetric drawings are more frequently associated with Healthy Control predictions. These findings enhance the interpretability of the model.

3.4 Conclusion

This work demonstrates that alzheimer disease risk can be effectively estimated using handwritten drawing tasks combined with basic demographic information, without requiring laboratory-based or invasive diagnostic procedures. By integrating visual features extracted using EfficientNetB0, attention mechanisms (CBAM), and clinical metadata through multimodal fusion, the proposed approach achieves strong performance on a limited dataset. Overall, the results suggest that handwriting-based analysis offers a low-cost, accessible, and scalable solution for early alzheimer risk screening.