Alzheimer Prediction

1st Caim Alexandru Gabriel
Computer Science
Politehnica University of Bucharest
Bucharest, Romania
alexandru.caim@stud.acs.upb.ro

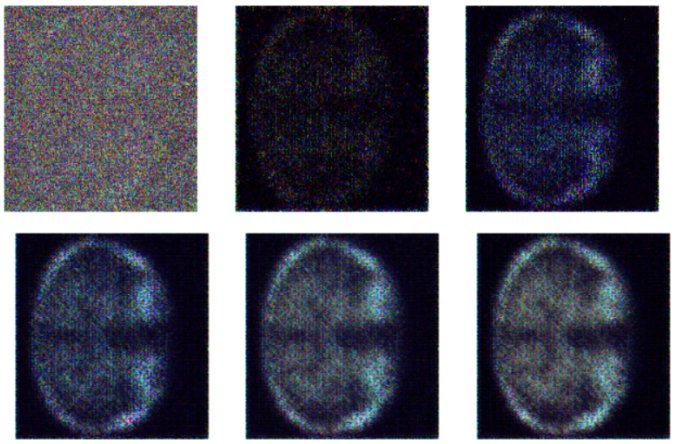
2nd Darius Mihail Calota
Computer Science
Politehnica University of Bucharest
Bucharest, Romania
darius.calota@stud.acs.upb.ro

I. Dataset Improvements

For this dataset, several architectures and models with satisfactory results have been proposed, with metrics nearing perfection. However, the dataset is imbalanced, and there is a significant lack of examples for the 'Moderate Dementia' class. A first step in mitigating this shortcoming is the use of augmentation techniques. When dealing with MRI images, the angle by which we can rotate the frame must fall within the range [-15, 15]; otherwise, the resulting image will not correspond to reality. We will use a range of [-10, 10]. Additionally, we have observed that after scaling, the displacement of the portion representing the brain becomes more noticeable. Decentered images help prevent overfitting, so we aim to shift by 5% horizontally to obtain new frames.

We also attempted an approach based on GANs to generate synthetic images. A Pix2Pix GAN using Mean Squared Error as the loss function was trained on the minority class images. Due to the lack of sufficient examples, the noise in the output is considerable, and fine details are missing.

Fig. 1. Evolution of GAN images threw epochs



II. Ablation Study

Building on the architecture proposed by Wang et al.(2018) [3] and the comparison made by Saratxaga et al. [2] between 3D and 2D CNNs, we designed a network composed of 3D convolutions. The model took as input a volume with a depth of 61 and a height and width of 224. As a result of transforming a 3D volume, only 8 examples for the 'Moderate Dementia' category were obtained, which is insufficient. Thus, the representatives of this class in the test set have reduced the accuracy of the classification. The results showed an accuracy of 0.81 on the test set.

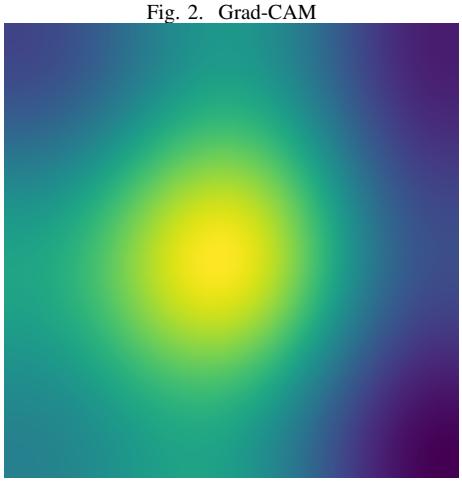
Another modification to Wang's original model involved training in two stages:

I) Classic training and the creation of a GRAD-CAM to highlight the regions that contributed most to the classification process.

II) Splitting the initial images into patches, training them independently, and during testing, determining the probability of class membership by taking the weighted average of the patch probabilities, with weights derived from the GRAD-CAM values.

The idea started from the visualization of the Grad-CAM, where we noticed that the areas near the corners of the image are not as relevant, so we focused our attention on the rest. The issues encountered resulted from the variation in relevant regions for prediction. The representation of the brain across the 61 slices in the transverse plane introduces multiple regions that the model considers for classification, depending on their coordinates along the Z-axis. In the end, the weights for the parts of the image did not vary as much as expected. The obtained accuracy was 0.72.

Fig. 2. Grad-CAM



The use of a Focal Loss function to prioritize the classification of the minority class did not bring improvements.

This was because, in attempting to increase the precision of detecting more severe cases, we significantly reduced the accuracy in detecting the early stages.

III. Proposed Method

The initial model has a very good classification accuracy, and most of the erroneous predictions involve confusion between the 'Moderate Dementia' and 'Very Mild Dementia' classes. In reality, the difference between these two is minor, and this is also reflected in the resulting logits.

The proposed method relies on optimizing the classification based on the model's outputs. Thus, we use an SVM with an RBF kernel, which is trained and makes predictions at a frequency equal to the validation epochs.

Test results (weighted)		
Metrics	Baseline	Proposed Method
Accuracy	0.97	0.98
Precision	0.97	0.98
Recall	0.97	0.98
F1 Score	0.97	0.98

Validation results		
Metrics	Classic	SVM
Accuracy	0.9773	0.9807
Precision	0.9734	0.9735
Recall	0.9547	0.9339
F1 Score	0.9637	0.9521

Fig. 3. Architecture

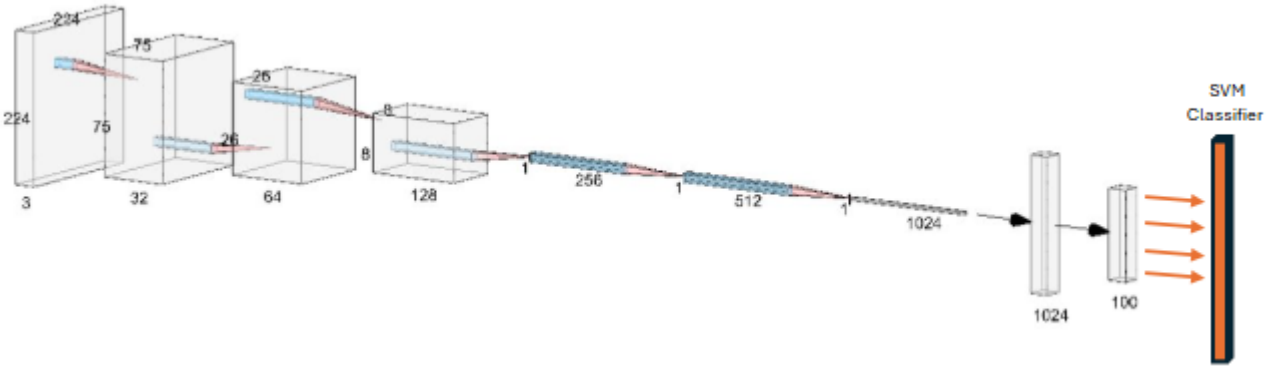
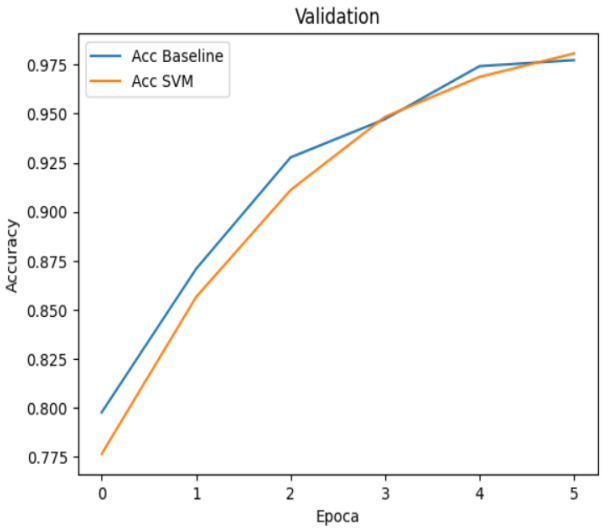


Fig. 4. Accuracy Evolution



We conducted a comparison between the initial model and the proposed variant by training both models for 12 epochs, performing validation 6 times.

The test set consisted of 738 images (Mild Dementia), 10,069 images (Non Demented), 67 images (Moderate Dementia), and 2,092 images (Very Mild Dementia).

The tests were conducted on a P100 graphics card provided by Kaggle.

The results showed that the inclusion of an SVM could improve the model's performance. The use of 3D convolutions was not beneficial, nor was the attempt to split the data into multiple segments. 2D convolutions proved to be the optimal alternative, and Wang's model was very well-suited for this dataset. Expanding the data volume could help the model, but this can largely be achieved with real images of patients, a common issue in artificial intelligence applications in medicine.

REFERENCES

[1] *OASIS Alzheimer's Detection*. https://www.kaggle.com.

[2] Cristina L Saratxaga et al. "MRI deep learning-based solution for Alzheimer's disease prediction". In: *Journal of personalized medicine* 11.9 (2021), p. 902.

[3] Shui-Hua Wang et al. "Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling". In: *Journal of medical systems* 42 (2018), pp. 1–11.