

Neuro-AI: An Integrated Deep Learning Framework for Enhanced MRI-Based Alzheimer's Disease Classification

Firdose Fathima Begum Syeda¹, Faridha Dudekua², Joice Vidhya Grace Lakka³, Lakshmi B V S N.⁴

^{1,2,3}Student, ⁴Assistant Professor, Department of Computer Science and Engineering (AI & ML), Rajeev Gandhi Memorial College of Engineering & Technology, Nandyal, India.

E-mail: ¹firdosesyeda32@gmail.com

Abstract

Alzheimer's Disease (AD) is a progressive neurodegenerative condition wherein the early identification of Cognitively Normal (CN), Mild Cognitive Impairment (MCI) and AD is of importance to researchers. The different medical equipment used and data collection settings bring about differences that do not pertain biologically when processing the data using automated analysis tools. This research work proposes, Neuro-AI, a deep learning-based pipeline for MRI image preprocessing, harmonization, super-resolution, explainable classification and automatic report generation. T1-weighted MRI images from 202 selected images of OASIS-3 dataset were used in this study and the obtained data are preprocessed by N4 bias correction, skull stripping and intensity normalization techniques. Then a multi-level enhancement process was employed through FusionNet, CycleGAN and SRGAN methods to maximize the imaging variation and spatial resolution from 128×128 pixels to 256×256 pixels. Then, a pretrained MobileNetV3-Small model was fine-tuned on classifying CN, MCI and AD images. The model achieved an accuracy of 99.95% and a weighted F1-score of 0.9995 on the balanced dataset. The image quality analysis indicated an increase in the average PSNR by 9.60 dB and the average SSIM by 0.0547 with SRGAN processing. GradCAM was used for generating visualization for model predictions, and TinyLLaMA for creating structured reports based on clinical data using Streamlit. While the results indicate the potential of the proposed system, its dependence on the subject level data and lack of multi-site validation prevent any claims about generalization.

Keywords: Alzheimer's Disease, Brain MRI, Deep Learning, MRI Harmonization, Super-Resolution, Explainable AI, GradCAM.

1. Introduction

AD (Alzheimer's Disease) is one of the common neurodegenerative diseases associated with dementia [1]. Early diagnosis of AD and Mild Cognitive Impairment (MCI) is crucial since the effects of AD may show up in brain structure before any severe symptoms appear. MRI (magnetic resonance imaging) enables non-invasive brain structure assessment and is frequently utilized in Alzheimer's disease research studies [2]. Variations in MRI scanners, acquisition protocols, and sites where images are collected can influence MRI-based deep learning models. Variations in images can lead to non-biological differences in intensity and appearance that can make the model learn the scanner artifacts instead of the disease-related information [3]–[8]. Also, an imbalance between the number of CN, MCI, and AD samples can lead to the class bias in models [11], [12].

Deep learning techniques have also been proven to be effective in harmonizing MRIs, translating images, super-resolving images, and classifying Alzheimer's disease [4], [6], [7], [11], [15]. Nonetheless, classification accuracy alone is not very helpful in understanding why a model has made a certain classification. Explainable AI techniques such as GradCAM help provide visual proof by showing what parts of the image contributed to the classification process [13]. In order to solve the aforementioned problems, this research work introduces Neuro-AI, a comprehensive system for MRI pre-processing, harmonizing, super-resolving, three-class classification, visual explanation, and clinical reporting automation. The proposed solution utilizes T1-weighted MRI scans from the OASIS-3 [14] dataset and classifies them into Cognitively Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). The main contributions are:

- A multi-stage pipeline combining FusionNet, CycleGAN, and SRGAN for MRI harmonization and enhancement.
- Session-level clinical label assignment based on the closest clinical assessment.
- Class balancing using undersampling, weighted sampling, and inverse-frequency loss.
- Integration of GradCAM visualization and automated reporting within a Streamlit application.

2. Related Works

MRI data obtained using different scanning devices can contain some systematic biases that would impact automated processing. Some statistical harmonization techniques have been proposed in order to mitigate such issues [3], [5]. With deep learning methods, the image-level harmonization including contrast harmonization and unsupervised learning of scanner-invariant representation is now possible [4], [7], [8]. It has been shown that it is possible to diminish scanner-related effects while maintaining necessary anatomical information.

The GAN models have been extensively used for medical image transformations. The CycleGAN model allows performing the unpaired image-to-image translation due to cycle consistency requirements [6] and thus can be used when paired images are not available. SRGAN utilizes adversarial training to produce high-quality high-resolution images [15]. U-Net represents an effective encoder-decoder architecture for preservation of spatial information [16]. All those ideas are incorporated in the suggested framework sequentially.

Deep convolutional neural networks have been successful in classifying Alzheimer's disease using brain MRI images [11]. Medical and imaging sources, such as ADNI, have also been useful in building automated neuroimaging tools [12]. Class imbalance can influence multi-class classification accuracy especially in case of CN, MCI, and AD classes that are not well-balanced.

Another factor to consider in medical imaging applications is the explainability of models. GradCAM produces heatmaps that demonstrate which parts are used for making a prediction [13]. Inclusion of visual explanations along with classification can offer extra information about the disease category that the model predicts.

Prior researches have mainly concentrated on either MRI harmonization [3]-[8], GANs for translation [6], super-resolution [15], classification of Alzheimer's disease [11] or explainability [13]. Unified approach that combines all these stages together with session-level labeling, balancing of classes and automated reports is less common. Proposed Neuro-AI framework seeks to fill this research gap by integrating all the stages in one MRI analysis pipeline based on the OASIS-3 dataset [14].

3. Dataset

The proposed Neuro-AI architecture is based on the use of the Open Access Series of Imaging Studies 3 (OASIS-3) dataset, which is a longitudinal neuroimaging and clinical data set meant to facilitate research on normal aging and Alzheimer's disease [14]. The OASIS-3 data set offers longitudinal MRIs along with clinical and cognitive data of the study subjects during several visits. The presence of repeat MRI scans and related clinical data makes this dataset appropriate for investigating brain structure changes and its correlation with Alzheimer's disease progression.

4. Methodology

The Neuro-AI architecture has been proposed is a full-fledged deep learning architecture intended for MRI pre-processing, harmonization, super-resolution, AD classification, visual explanation, and report generation. Figure 1 depicts the complete Neuro-AI pipeline. The whole process includes five main steps, namely: (1) pre-processing of the MRI data; (2) harmonization through FusionNet; (3) domain transfer using CycleGAN; (4) super-resolution through SRGAN; and (5) disease classification using GradCAM for explainability and report generation using TinyLLaMA.

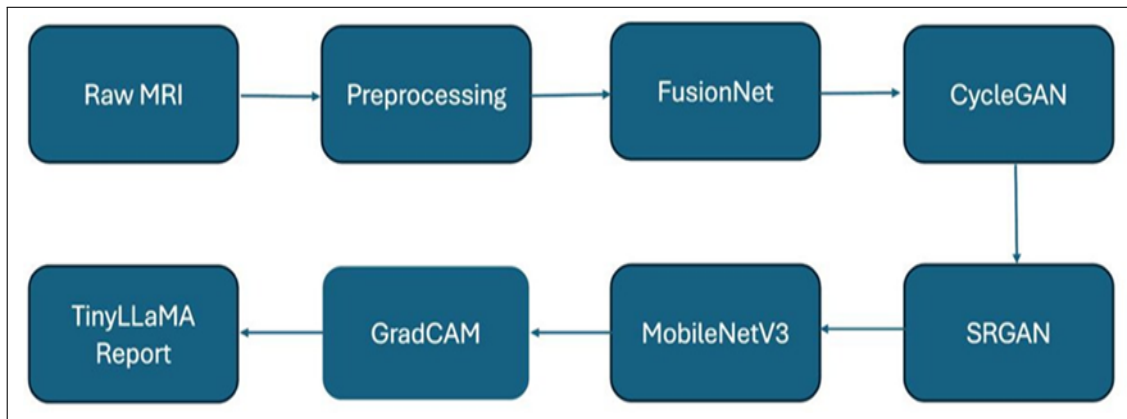


Figure 1. Neuro-AI end-to-end pipeline

4.1 Preprocessing

The MRI image at T1 in NIfTI format is preprocessed following a series of steps to normalize the intensity and achieve uniformity between different images. An example of the preprocessed image is given in Figure 2. Firstly, N4 Bias Field Correction is done to overcome

the problem of low frequency intensity inhomogeneity due to magnetic field variations. Secondly, skull stripping is done to eliminate all the non-brain parts of the image and only keep those parts which have relevance to Alzheimer's disease.

After that, intensity normalization is done on the brain image. Z-Score Normalization is done during the preprocessing step in order to normalize the intensity of the voxels. After that, appropriate axial slices are selected from the preprocessed image and they are resampled to achieve a resolution of 128×128 pixels.

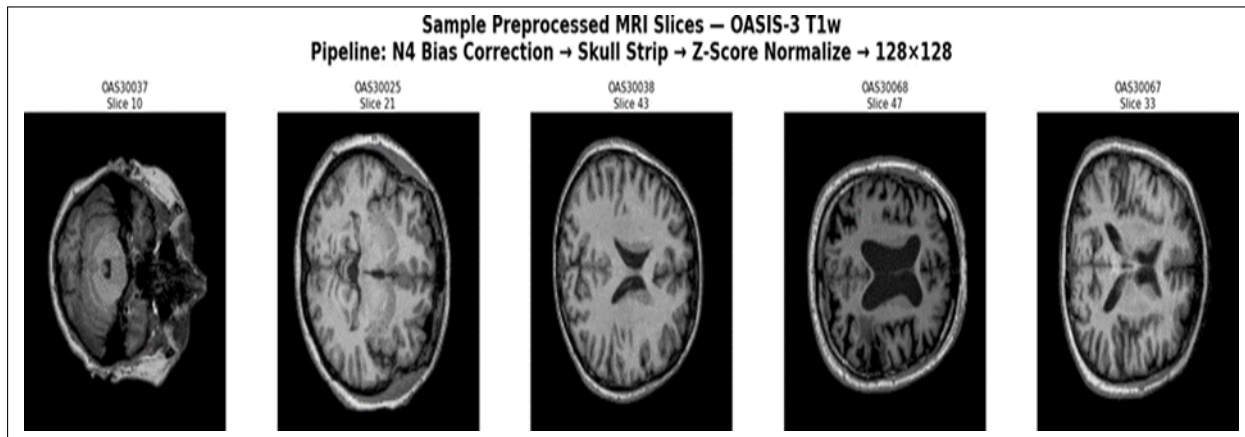


Figure 2. Sample preprocessed MRI slices

4.2 FusionNet-based MRI Harmonization

FusionNet serves as the initial deep learning harmonization step. The neural network utilizes the U-Net encoder-decoder design framework with channels of [16, 32, 64, 128] dimensions. The goal is to minimize the scanner-specific and acquisition-specific intensity discrepancies while maintaining anatomically significant features.

The FusionNet optimization objective is defined as:

$$L_{\text{fusion}} = 0.8L_{L1} + 0.2L_{\text{SSIM}} \quad (1)$$

where L_{L1} represents the L1 reconstruction loss and L_{SSIM} denotes the structural similarity loss. The combined objective encourages both pixel-level fidelity and preservation of structural information.

FusionNet is trained using the Adam optimizer with a learning rate of 104, a batch size of 4, and 30 training epochs.

4.3 CycleGAN-based Domain Translation

After FusionNet harmonization, CycleGAN is applied to generate images between unpaired CN and AD domains. This method contains two generators, GAB and GBA, and PatchGAN discriminators. In both generators, there are four residual blocks.

The CycleGAN objective is expressed as:

$$L = L_{\text{GAN}} + 10L_{\text{cycle}} + 5L_{\text{id}} \quad (2)$$

where L_{GAN} represents the adversarial loss, L_{cycle} denotes the cycle-consistency loss, and L_{id} represents the identity loss. The cycle-consistency component encourages the translated image to preserve the underlying anatomical content, while the identity term helps prevent unnecessary alterations.

A replay buffer with 50 images from the generated set is employed for improving the training stability. In case of inference, CN images are transformed using GAB and AD images using GBA, whereas MCI images remain unaltered. This is the labeling-invariant inference approach employed for retaining labels during harmonization.

4.4 SRGAN-based Super-Resolution

The 128×128 harmonized images are then further processed using a Super-Resolution Generative Adversarial Network (SRGAN). The purpose of this step is to enhance the spatial resolution and retain perceptually significant anatomical structures.

The SRGAN produces images of size 256×256 through a generator comprising six residual blocks along with a $2\times$ upsampling process using PixelShuffle. The optimization objective is given by:

$$L_{\text{SR}} = L_{\text{pixel}} + 10^{-3}L_{\text{adv}} \quad (3)$$

where L_{pixel} represents the pixel reconstruction component and L_{adv} denotes the adversarial loss. The resulting high-resolution images are used as inputs to the disease-classification stage.

4.5 Disease Classification

A pretrained MobileNetV3-Small model is fine-tuned on a classification task with three classes: CN, MCI, and AD. The initial slices distribution is significantly imbalanced: there are

838 slices with CN class label, 278 with MCI, and 12,540 with AD class label. In order to decrease the effect of this imbalanced distribution, three approaches are used.

First, the samples from the most populated classes are undersampled to have 278 samples in each class. The second approach is the use of `WeightedRandomSampler` during the training. Third, the inverse-frequency weighting is applied in the `CrossEntropyLoss` function.

The classifier is trained for 30 epochs by using the Adam optimizer with learning rate 3×10^{-4} and cosine annealing learning rate scheduling. The data set is split into train, validation, and test sets in proportion 70/15/15.

4.6 GradCAM Explainability and Clinical Report Generation

GradCAM is used to help explain visually the predictions made by MobileNetV3-Small. The gradient-weighted activation maps are derived from the last layer of features, which is `model.features[-1]`, and are then superimposed on the respective slices of the MRI images. This will serve to identify parts of the image that are contributing significantly to the model's classification process.

The platform also integrates the TinyLLaMA 1.1B via Ollama, which generates automatically an output report. The output reports for each prediction are comprised of five items namely the disease explanation, GradCAM explanation, clinical test recommendations, prognosis, and clinical conclusion together with an AI disclaimer. This entire process of MRI prediction, visualization, and report generation is incorporated in the interactive Streamlit platform.

5. Results and Discussion

5.1 Preprocessing and MRI Harmonization

Preprocessing successfully performed on all 202 chosen MRI scans. N4 Bias Field Correction and skull stripping were performed to normalize intensity and eliminate structures outside the brain, respectively. After session-level labeling, the obtained slice numbers were 838 CN slices, 278 MCI slices, and 12,540 AD slices.

FusionNet-based harmonization step provided a stable transformation of images without losing their anatomy. Replay buffer used in CycleGAN training helped to ensure stable training for domain translation and minimize optimization oscillations. Label-preserving method was used for inference, which preserved three categories.

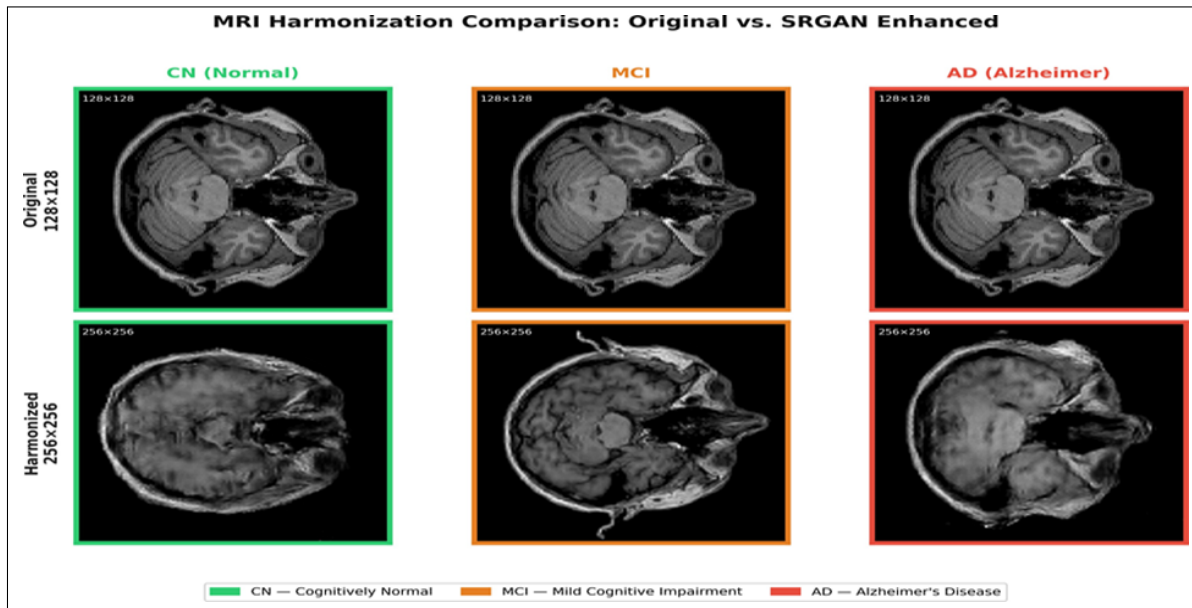


Figure 3. MRI harmonization comparison

The SRGAN model resulted in a higher level of sharpness in comparison to the usual bilinear upsampling. The anatomical landmarks such as gyri and sulci were more prominent in the 256×256 generated images. Figure 3 provides visual representation for the comparison of the 128×128 slice images against their counterparts generated using SRGAN for CN, MCI, and AD groups.

5.2 Classification Performance

The performance evaluation of the MobileNetV3-Small model was conducted using a balanced dataset with 250 data instances. The details of precision, recall, F1-score, and support values for each category are given in Table 1.

Table 1. Classification performance on the balanced test set

Class	Precision	Recall	F1	Support
CN (Normal)	1.00	1.00	1.00	83
MCI	0.99	1.00	0.99	84
AD (Alzheimer)	1.00	0.99	1.00	83
Wtd. Avg	1.00	1.00	1.00	250

The reported classification accuracy of Neuro-AI system is 99.95%, and the weighted F1-score is 0.9995. Confusion matrix provided in Figure 4 shows the prediction performance of

the three diagnostic categories.

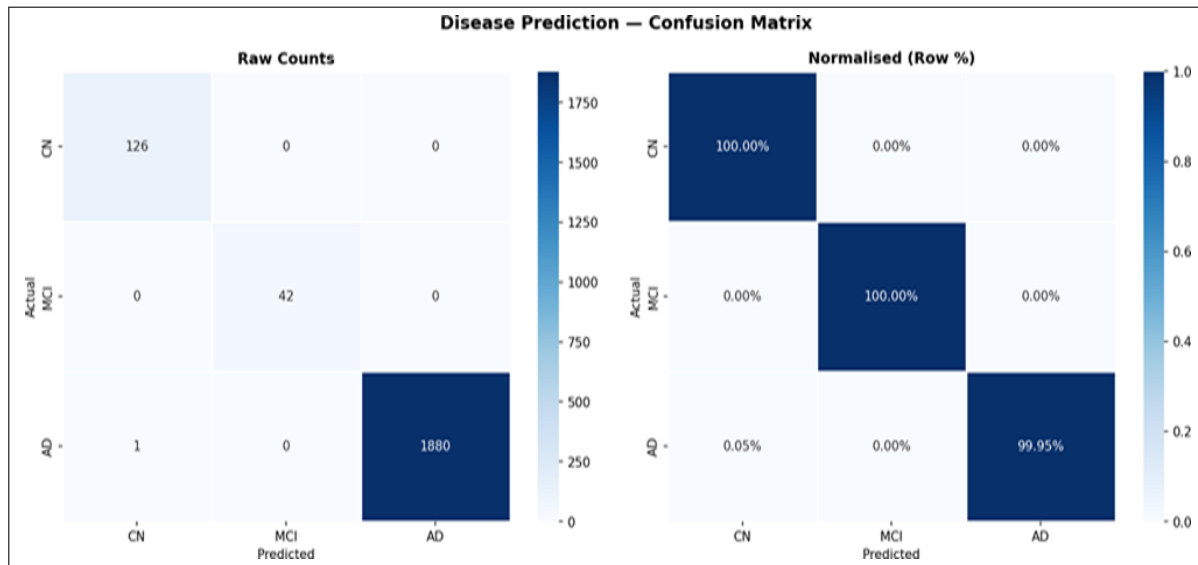


Figure 4. Confusion matrix for the Neuro-AI model.

5.3 Image Quality Evaluation

Table 2 presents the numerical comparison of the quality of images generated by the CycleGAN algorithm and SRGAN. The latter consistently improved both PSNR and SSIM measures compared to CycleGAN generated images.

In the case of CN samples, PSNR rose from 32.00 dB to 41.43 dB, which corresponds to 9.42 dB gain, while SSIM from 0.9363 to 0.9931. In case of MCI samples, PSNR rose from 31.47 dB to 40.86 dB and SSIM from 0.9309 to 0.9930. In case of AD samples, PSNR rose from 32.38 dB to 42.37 dB and SSIM from 0.9500 to 0.9953.

Overall, in the case of all 150 images analyzed, PSNR improved on average from 31.95 dB to 41.55 dB, which corresponds to overall improvement of 9.60 dB. On the other hand, SSIM measure on average improved from 0.9391 to 0.9938, which corresponds to an increase of 0.0547.

Table 2. Comparison of CycleGAN and SRGAN performance

Class	Samples	CycleGAN PSNR (dB)	CycleGAN SSIM	SRGAN PSNR (dB)	SRGAN SSIM	PSNR Gain (dB)	SSIM Gain
CN	50	32.00	0.9363	41.43	0.9931	9.42	0.0569
MCI	50	31.47	0.9309	40.86	0.9930	9.39	0.0620
AD	50	32.38	0.9500	42.37	0.9953	9.99	0.0453
Overall	150	31.95	0.9391	41.55	0.9938	9.60	0.0547

5.4 GradCAM Explainability and Streamlit Interface

GradCAM visualizations were applied to reveal the areas in the images that contribute towards the predictions made by the model. Activation maps produced different spatial distributions for the various diagnostic categories. Activation maps of the AD prediction had high levels of activation within the areas related to changes brought about by Alzheimer disease, such as the hippocampus and entorhinal region.

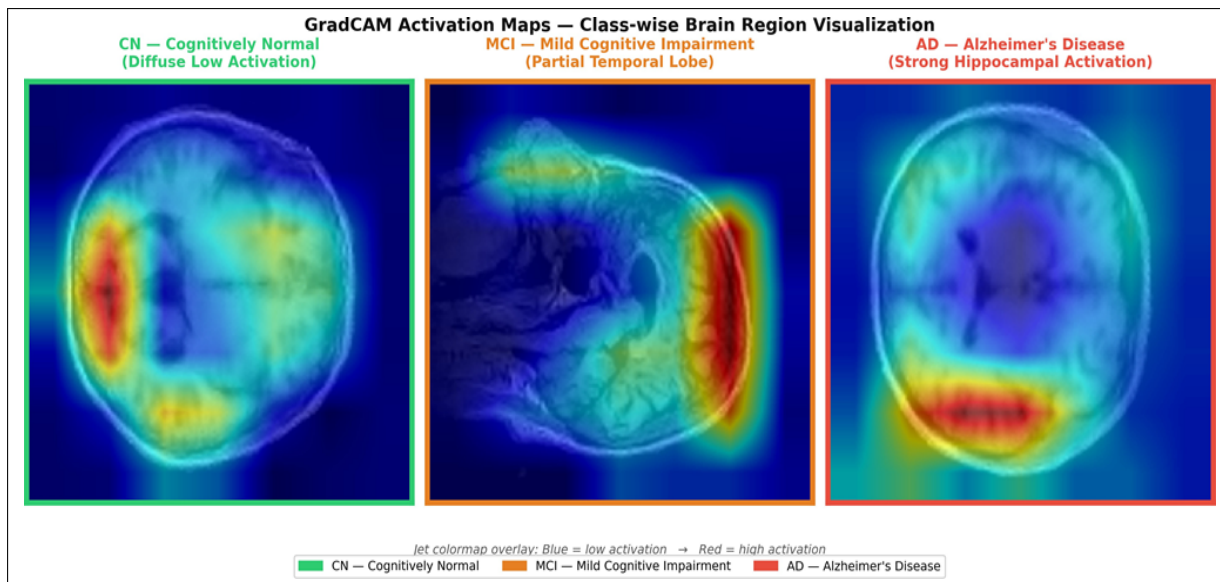
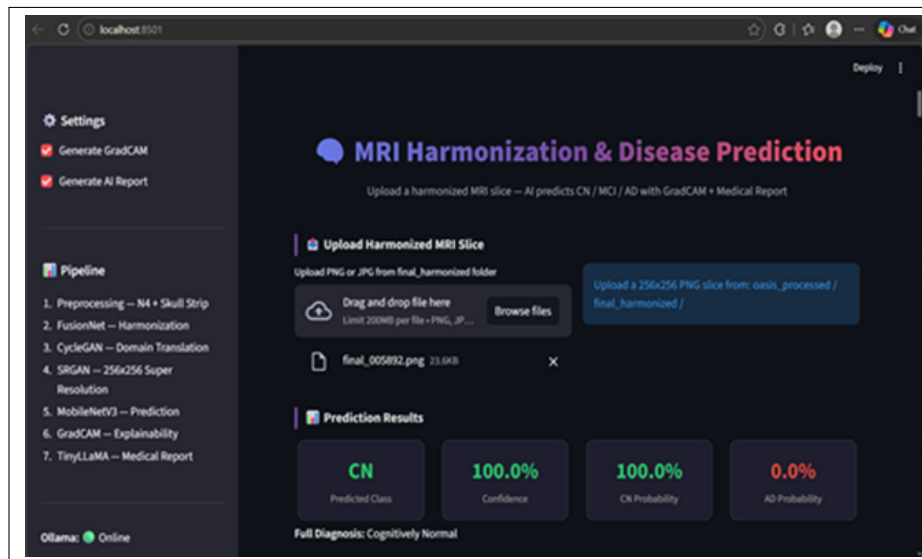
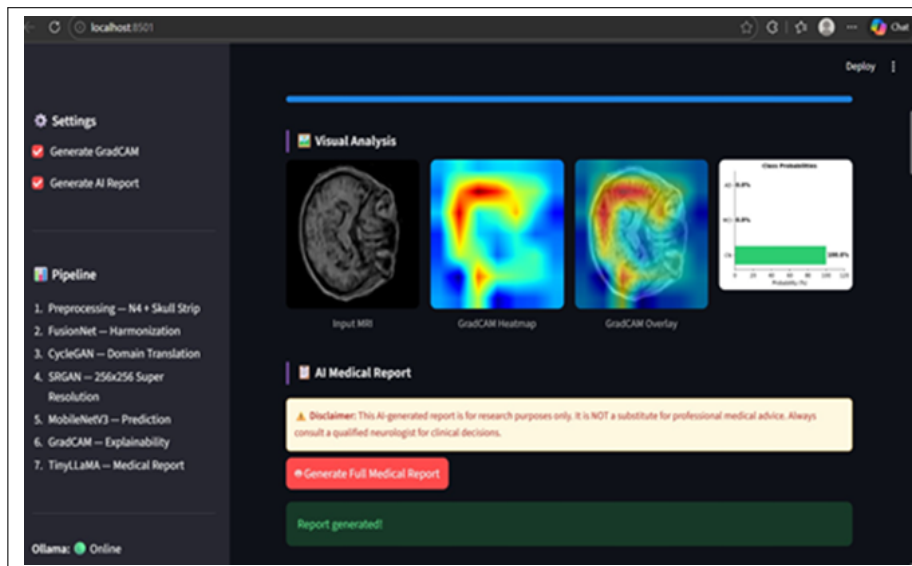
**Figure 5.** GradCAM overlays.

Figure 5 represents the GradCAM overlay images for CN, MCI, and AD patients. GradCAM visualizations help in understanding the working mechanism of the classifier along with numerical classification results. The entire Neuro-AI model has been embedded into a

Streamlit interface. Figure 6 represents the application which provides MRI image upload, disease prediction, GradCAM visualization, and AI-based clinical report generation. The generated reports include suggestions such as MMSE, amyloid PET, and CSF biomarkers along with AI disclaimer.



(a)



(b)

Figure 6. Interactive streamlit interface—(a) MRI harmonization and disease prediction interface, (b) GradCAM visual analysis and AI medical report interface for the neuro-AI system

5.5 Discussion

The results show that Neuro-AI pipeline offers a reliable method of MRI-based classification of CN, MCI, and AD patients. The session-level label assignment is a crucial part of the pipeline since it allowed the available patient data to be linked with their individual MRI sessions, thus ensuring representation of all three diagnoses when creating the models. Image processing stages provided additional operations in the pipeline. The harmonization step was applied in order to minimize the influence of acquisition variance, while the super-resolution improved the spatial image representations. The image-quality enhancement achieved in the process indicates that the enhancement step can offer better representations for further processing. Improved quality does not necessarily mean improved diagnosis, however, and should be evaluated separately. The classification results prove to be highly effective on the chosen balanced test data. The use of undersampling, weighted sampling, and inverse frequency loss contributed to dealing with the class imbalance problem inherent to the initial data. GradCAM serves as an additional feature of the classifier, adding a visual explanation of the region corresponding to each prediction outcome, thereby improving the interpretability of the outcome.

6. Conclusion

Neuro-AI provided an integrated framework for Alzheimer's disease diagnosis based on MRI data via the integration of image processing, classification, model explainability, and automated reporting. The experimental results showed the possibility of applying the framework for discrimination of cognitively normal, mild cognitive impairment, and Alzheimer's disease patients with provision of visual explanations to the models. The proposed framework also provides an effective way of reporting the obtained predictions and analytical results. The current analysis is limited by the use of specific OASIS-3 sessions and interdependencies that can exist between the repeated measurements of the same subjects. The lack of independent multi-site validation of the obtained results does not allow making any conclusions about the generalizability of the proposed framework. Further research will include the consideration of subject-level split, analysis of the entire OASIS-3 dataset, and testing on independent multi-site datasets. Other imaging modalities such as PET and CSF biomarkers will also be analyzed.

References

- [1] Jack Jr, Clifford R., David A. Bennett, Kaj Blennow, Maria C. Carrillo, Billy Dunn, Samantha Budd Haeberlein, David M. Holtzman "NIA-AA Research Framework: Toward a Biological Definition of Alzheimer's Disease." *Alzheimer's & Dementia* 2018, vol. 14, no. 4, 535-562.
- [2] Brown, Robert W., Y-C. Norman Cheng, E. Mark Haacke, Michael R. Thompson, and Ramesh Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. John Wiley & Sons, 2014.
- [3] Fortin, Jean-Philippe, Drew Parker, Birkan Tunç, Takanori Watanabe, Mark A. Elliott, Kosha Ruparel, David R. Roalf "Harmonization of Multi-site Diffusion Tensor Imaging Data." *Neuroimage* 2017, vol. 161, 149-170.
- [4] Zuo, Lianrui, Blake E. Dewey, Yihao Liu, Yufan He, Scott D. Newsome, Ellen M. Mowry, Susan M. Resnick, Jerry L. Prince, and Aaron Carass. "Unsupervised MR Harmonization by Learning Disentangled Representations using Information Bottleneck Theory." *NeuroImage* 2021, vol. 243, 118569.
- [5] Fortin, Jean-Philippe, Nicholas Cullen, Yvette I. Sheline, Warren D. Taylor, Irem Aselcioglu, Philip A. Cook, Phil Adams "Harmonization of Cortical Thickness Measurements Across Scanners and Sites." *Neuroimage* 2018, vol. 167, 104-120.
- [6] Zhu, Jun-Yan, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks." *IEEE International Conference on Computer Vision (ICCV)*, 2017, 2242-2251.
- [7] Dewey, Blake E., Can Zhao, Jacob C. Reinhold, Aaron Carass, Kathryn C. Fitzgerald, Elias S. Sotirchos, Shiv Saidha "DeepHarmony: A Deep Learning Approach to Contrast Harmonization Across Scanner Changes." *Magnetic Resonance Imaging* 2019, vol. 64, 160-170.
- [8] Dewey, Blake E., Lianrui Zuo, Aaron Carass, Yufan He, Yihao Liu, Ellen M. Mowry, Scott Newsome, Jiwon Oh, Peter A. Calabresi, and Jerry L. Prince. "A Disentangled Latent Space for Cross-Site MRI Harmonization." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Cham: Springer International Publishing, 2020, 720-729.

- [9] Tishby, Naftali, Fernando C. Pereira, and William Bialek. "The Information Bottleneck Method." arXiv Preprint Physics/0004057 (2000).
- [10] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>.
- [11] Islam, Jyoti, and Yanqing Zhang. "Brain MRI analysis for Alzheimer's Disease Diagnosis using an Ensemble System of Deep Convolutional Neural Networks." Brain Informatics 2018, vol. 5, no. 2, 2.
- [12] Petersen, Ronald Carl, Paul S. Aisen, Laurel A. Beckett, Michael C. Donohue, Anthony Collins Gamst, Danielle J. Harvey, Clifford R. Jack Jr "Alzheimer's Disease Neuroimaging Initiative (ADNI) Clinical Characterization." Neurology 2010, vol. 74, no. 3, 201-209.
- [13] Selvaraju, Ramprasaath R., Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. "Grad-cam: Visual Explanations from Deep Networks via Gradient-based Localization." In Proceedings of the IEEE International Conference on Computer Vision, 2017, 618-626.
- [14] LaMontagne, Pamela J., Tammie LS Benzinger, John C. Morris, Sarah Keefe, Russ Hornbeck, Chengjie Xiong, Elizabeth Grant "OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease." Medrxiv (2019), 2019-12.
- [15] Ledig, Christian, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken "Photo-Realistic Single Image Super-Resolution using a Generative Adversarial Network." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, 105-114.
- [16] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In International Conference on Medical Image Computing and Computer-Assisted Intervention, Cham: Springer International Publishing, 2015, 234-241.