Grenze International Journal of Engineering and Technology, June Issue

XAI-Enhanced MRI Analysis for Alzheimer's Disease Prediction

Sakhineti Aakash Kumar Raj[1], Veluru Manoj[2] and Helen Victoria[3]
[1-3]Department of Networking and Communications, School of Computing, College of Engineering and Technology SRM Institue of Science and Technology, SRM Nagar, Kattankulathur, Tamilnadu,India.
Email: ar0829@srmist.edu.in, vm2792@srmist.edu.in, helenvia@srmist.edu.in

***Abstract*— This research delves into the application of Explainable Artificial Intelligence (XAI) in the healthcare sector, specifically focusing on Alzheimer's disease prediction. Alzheimer's, a debilitating brain dis-ease, impacts daily activities through cognitive and be-havioral changes. Leveraging XAI in medical diagnostics enhances interpretability, providing transparent insights crucial for informed decision-making by healthcare pro-fessionals. The study explores innovative approaches within the complex landscape of machine learning and data utilization, aiming to unravel the intricacies of Alz-heimer's prediction models. Through the interdiscipli-nary lens of interpretable machine learning, this research seeks to enhance our understanding of the disease, paving the way for personalized and reliable diagnostic and treatment strategies.**

***Index Terms*— XAI, CNN, VGG16, Random Forest, MRI, KNN, XG Boost, LIME, Image Processing, Django, Django Server.**

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) has revolutionized the diagnosis of neurological conditions by offering detailed visualization of internal structures, particularly within the brain. Despite its immense utility, accurately identifying anatomical landmarks within MRI images remains chal-lenging, which is crucial for diagnosing conditions such as tumors and neurodegenerative disorders.

To address the challenge of landmark detection in MRI images, there is a growing interest in integrating Explainable Artificial Intelligence (XAI) techniques into medical image analysis, especially within machine learning approaches. These techniques aim to enhance both precision and inter-pretability in the diagnostic process.

Our proposed method combines machine feature learning with XAI principles to improve landmark detection in MRI brain disease diagnosis. By automatically extracting distinc-tive features from brain images, our approach facilitates accurate landmark localization and provides insight into the decision-making process of the model.

We leverage various XAI techniques, including saliency mapping, attention mechanisms, and feature attribution methods, to understand which image features are most rel-evant to the model's predictions. These techniques enhance trust and transparency in the diagnostic process, enabling healthcare professionals to better interpret and validate the model's findings.

By reducing the reliance on manual effort for landmark identification, our innovative technique streamlines the di-agnostic workflow, leading to increased efficiency and im-proved patient care outcomes. Additionally, the integration of XAI into machine learning-based medical image analysis offers a promising avenue for enhancing

Grenze ID: 01.GIJET.10.2.92_1

neurological diag-noses.
Our approach provides both accuracy and interpretability, empowering healthcare professionals with the necessary tools to make informed decisions in neurological diagnosis. This not only improves patient care and outcomes but also enhances the overall efficiency and effectiveness of neu-rology practices.
In summary, the integration of XAI into machine learning-based medical image analysis offers a promising solution for improving neurological diagnoses. By providing accu-rate and interpretable results, our approach enhances trust, transparency, and efficiency in the diagnostic process, ulti-mately benefiting both healthcare professionals and pa-tients in the field of neurology.

II. Literature Review

Dementia is a term used to describe a group of symptoms that includes changes in cognitive abilities. AD is not a natu-ral aspect of the aging process. It occurs due to intricate changes that take place in the brain years before symptoms manifest, which ultimately results in the degeneration of brain cells and their interconnections [1]. In the early stage of this disease patient finds difficulty in walking, speaking, and swallowing, however, in later stage it starts affecting daily routine activities of person as it induces partial memory loss and thinking capabilities. This disease is hard to detect as it initiates in brain decades before symptoms are observed [2]. Early AD therapy is more efficient and results in minor harm in comparison with medication ad-ministered later in the disease's progression. Alzheimer's di-agnosis can be done using different scans such as PET, MRI or CT. Traditional medical testing takes a long time and cannot predict early warning signs. Machine learning mod-els can be used for early detection and to achieve efficient classification accuracy [3]. Orimaye et al. evaluated four ML algorithms, including DT algorithm, Naïve Bayes (NB) algorithm, SVM algorithm with radial basis kernel, and Neu-ral Network algorithm [13]. Using six significant syntactic and lexical features extracted from 242 AD participants and 242 healthy individuals, they found that compared with other ML algorithms, SVM gave a better accuracy score of 74%, a recall rate of 73%, and a precision rate of 75% [4]. Alzheimer's disease is not inevitable in people with Down syndrome. While all people with Down syndrome are at risk, many adults with Down syndrome will not manifest the changes of Alzheimer's disease in their lifetime. Alt-hough risk increases with each decade of life, at no point does it come close to reaching 100% [5]. Transparency and interpretability are emphasized through the application of the Local Interpretable Model-Agnostic Explanations (LIME) method to explain model predictions. Results demonstrate DeepEBTDNet's efficacy in brain tumour de-tection, even across datasets, achieving a validation accu-racy of 98.96% and testing accuracy of 94.0% [6]. The XAI assessment was conducted across 132 brain parcels, ex-tracted from a combination of the Harvard-Oxford and AAL brain atlases, and compared to well-known pathologi-cal regions to measure adherence to domain knowledge. Results highlighted acceptable classification performance as compared to the existing literature (ResNet18: TPRmedi-an = 0.817, TNRmedian = 0.816; BC-GCN-SE: TPRmedian = 0.703, TNRmedian = 0.738) [7]. The utility of structural imaging and other markers will be increased by standardiza-tion of acquisition and analysis methods, and by develop-ment of robust algorithms for automated assessment [8].

III. Motivation

Alzheimer's disease (AD) is a progressive neurodegenerative disorder with devastating consequences. Early detection is crucial for effective intervention and improving patient quality of life. However, current diagnostic methods often involve expensive and invasive procedures.
MRI scans offer a non-invasive way to analyze brain structures potentially affected by AD. However, accurate assessment relies on identifying specific anatomical landmarks within the MRI, which is currently a manual and error-prone process.
Machine learning algorithms hold immense promise for analyzing medical images and potentially predicting AD. However, their "black box" nature can be a significant hurdle. Without understanding the reasoning behind the prediction, healthcare professionals hesitate to trust the results.
By incorporating XAI techniques like saliency maps and feature attribution methods, the project aims to create an interpretable model. This allows healthcare professionals to see which image features like specific brain regions the model finds most relevant for predicting AD.
The project proposes using algorithms like Naive Bayes, XGBoost, Decision Trees, and Random Forests for their effectiveness in classification tasks. By employing these algorithms alongside XAI, the project strives to achieve both accurate AD prediction and transparent reasoning behind those predictions.

LIME is an acronym which stands for Local Interpretable Model-agnostic Explanations. The “Local” refers to the fact that LIME focuses on interpreting the behaviour of one data point at a time, not every data point in the dataset. In the case of image classification, this means that it interprets one image at a time. LIME gives a weight to each generated point, using a Gaussian (RBF) Kernel.

$$RBF(x^{(i)}) = \exp\left(- \|x^{(i)} - x^{(ref)}\|^2 / kw\right)$$

Gaussian Kernel Formula

The Gaussian Kernel attributes a value in the range [0,1], the higher the closer to the reference point. The kernel width kw parameter decides how large is the circle of the meaningful weights around the red dot.
The ultimate goal is to develop a reliable and user-friendly tool that empowers healthcare professionals to make informed decisions regarding AD diagnosis, leading to earlier intervention and improved patient outcomes.

IV. OBJECTIVES

Focus on early identification of tumors and neurodegenera-tive disorders for timely intervention and treatment. Devel-op a solution that works across various disease stages, de-mographics, and imaging conditions for broader applicabil-ity. Using machine learning algorithms, as is the case for Alzheimer's disease prediction, to predict early stage tumors. In the case of cancer patients, this may allow timely intervention and treatment. Fig1 depicts us the Extended project's focus beyond Alzheimer's disease to include the early identification of other neurodegenerative disorders like Parkinson's disease or Huntington's disease. Utilizing deep learning models, as seen in Alzheimer's disease stage identification research can aid in this expansion.

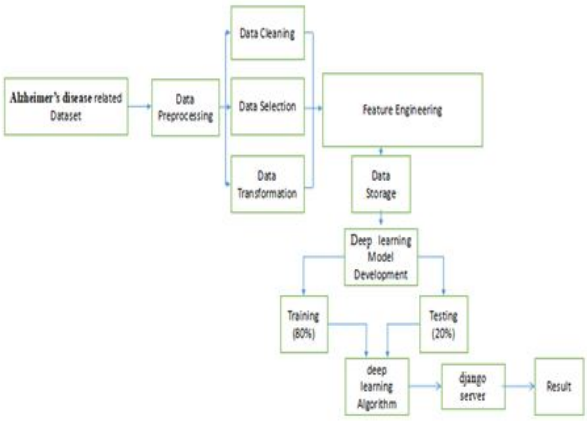

Fig1-System Architecture Diagram

The In Fig2 illustrates us the architecture of Explainable AI focuses on the development of models that are transparent and interpretable in the decision making process. This requires the creation of algorithms and systems capable of explaining their outputs in a human comprehensible way. The aim is to improve the reliability of artificial intelligence systems, allowing users to understand why decisions are made.

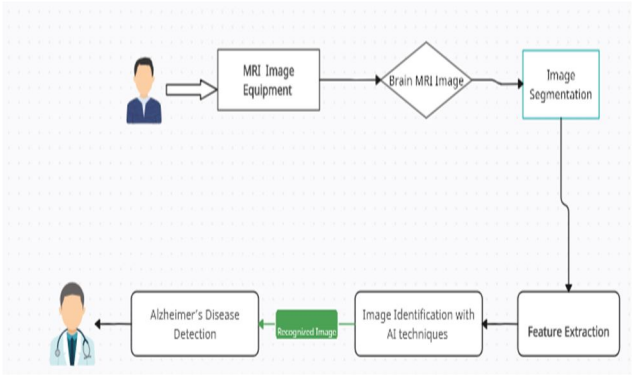

Fig2- XAI Architecture

LIME (Local Interpretable Model-agnostic Explanations) is a technique used for explaining the predictions of machine learning models. Fig3 was generated local approximations of a model's behaviour in relation to the particular data point, which makes it possible to interpret each prediction on its own. LIME creates a simpler, more interpretable model that can explain the complex model's decision locally.

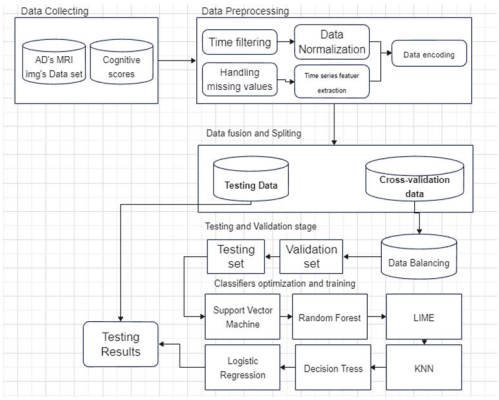

Fig 3-LIME architectuer

V. PROPOSED METHODOLOGY

The proposed methodology for XAI-enhanced MRI analysis for Alzheimer's Disease prediction integrates both traditional machine learning algorithms such as Random Forest, Decision Tree, Naive Bayes, and XG Boost, alongside explainable AI techniques to enhance transparency and interpretability of the predictive models. The methodology shall be described as follows:

1. Data Collection and Preprocessing

- For the purpose of collecting MRI data from individuals with Alzheimer's disease.
- Automate the data in order to guarantee uniformity and removal of noise, including demystification, standardisation or features extraction techniques specifically designed for MRI data.

	Group	Visit	MR Delay	M/F	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
0	Nondemented	1	0	0	87	14	2.0	27.0	0.0	1987	0.696	0.883
1	Nondemented	2	457	0	88	14	2.0	30.0	0.0	2004	0.681	0.876
5	Nondemented	1	0	1	88	18	3.0	28.0	0.0	1215	0.710	1.444
6	Nondemented	2	538	1	90	18	3.0	27.0	0.0	1200	0.718	1.462
7	Nondemented	1	0	0	80	12	4.0	28.0	0.0	1689	0.712	1.039
...	...	...	...	...	...	...	...	...	...	...	...	...
2293	Nondemented	2	1510	0	78	18	2.0	30.0	0.0	1484	0.703	1.183
2294	Nondemented	1	0	1	82	16	3.0	29.0	0.0	1484	0.760	1.183
2295	Nondemented	2	842	1	84	16	3.0	28.0	0.0	1500	0.744	1.170
2296	Converted	1	0	1	85	18	1.0	29.0	0.0	1264	0.701	1.388
2297	Converted	2	846	1	87	18	1.0	24.0	0.5	1275	0.683	1.376

Fig2- Patient Numerical Data

2. Selection of features

- Employ techniques such as Principal Component Analysis (PCA) or feature importance from Random Forest and XG Boost to identify the most relevant features for Alzheimer's Disease prediction.

3. Model Training

- Use pre-processed data to train prediction models using machine learning methods such as Random Forest, Decision Tree, Naive Bayes, and XG Boost.
-Improve model performance by fine-tuning hyperparameters using approaches such as grid search and Bayesian optimization.

4. Explanatory Model Analysis

- Use AI approaches to provide insights into the model's decision-making process and increase transparency.
- For Random Forest and Decision Tree models, feature importance can be calculated to determine which MRI features are most important in Alzheimer's disease prediction.

- Naive Bayes models can provide probabilistic reasons for classifications, such as the chance of an individual having Alzheimer's disease based on MRI findings.
XG Boost models provide interpretability through the presentation of feature significance scores and decision routes.

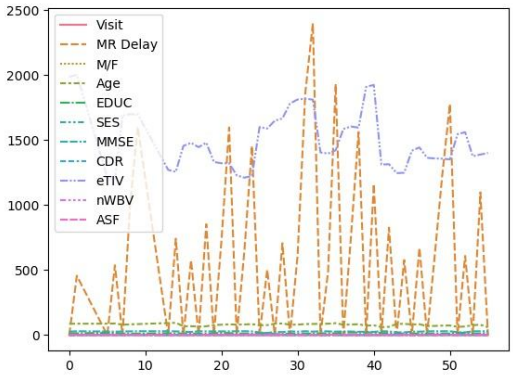

Fig-4 Data Training & Testing graphs

5. Model Evaluation

- Evaluate each model's performance using metrics such as accuracy, precision, recall, and F1-score.
- Conduct cross-validation to ensure model resilience and reduce overfitting.

6. Comparison and interpretation

- Compare different algorithms' predicted accuracy and interpretability.
- Analyse the XAI approaches' explanations to learn how the models create predictions and uncover potential biomarkers for Alzheimer's disease.

7. Deployment and validation

- Use the best-performing model in clinical situations to forecast Alzheimer's disease.
- Validate the model's performance with new, previously unknown data to guarantee generalizability and reliability.

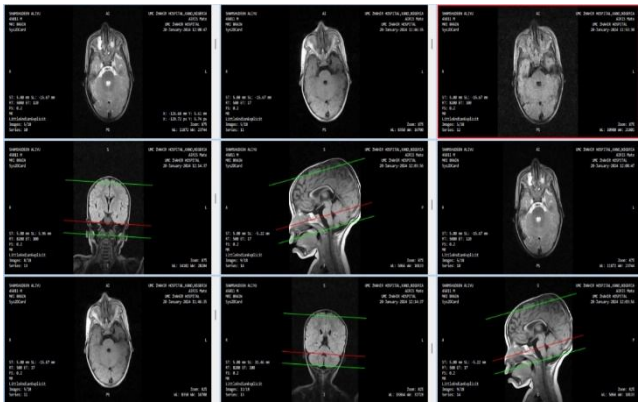

Fig3- MRI Data set images

Incorporating explainable AI into this methodology improves the interpretability of the predictive models, giving physicians and researchers new insights into the underlying causes driving Alzheimer's disease prediction using MRI data. This transparency is critical for increasing trust in predictive models and allowing their use in clinical practice.

VI. Results And Metrics

These are the results of the following researched project on XAI based AD's prediction, the following image represents the final outcomes of the epoch's loss, testing, and training accuracy.

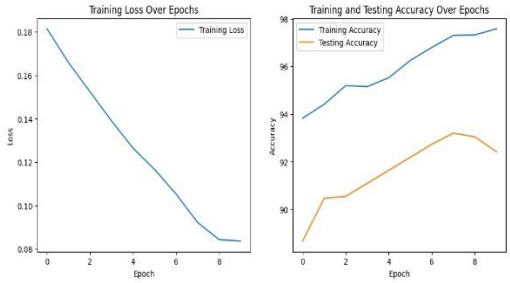

Fig4- Epochs loss and Testing & Training accuracy

The image mentioned below was the confusion matrix that represents the Confusion matrix, a set of tables used for machine learning to visualize the algorithm's performance during classification tasks. In this case, it appears that the confusion matrix is evaluating a classification model for dementia. The rows are the actual diagnosis of dementia-Very_Mild_Demented,Non_Demented,Mild_Demented,Moderate_Demented, while the columns represent the diagnos-tic predictions made by the model. The number of people in this category is represented by the number of cells in each cell. For example, the cell "Very_Mild_Demented" and the column "Non_Demented" are in the row "Very_Mild_Demented".

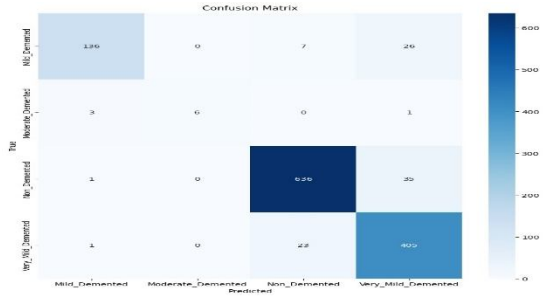

Fig5- Confusion matrix for classification.

These images represent the classification of the ADs in three different ways namely Normal Image Classification, Nu-merical Data based classification, and XAI Based classifi-cation.

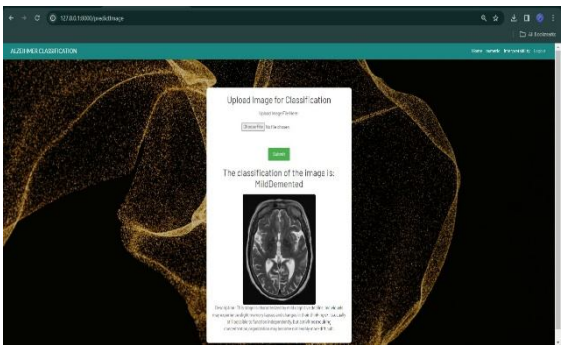

Fig 6- Normal Image Classification

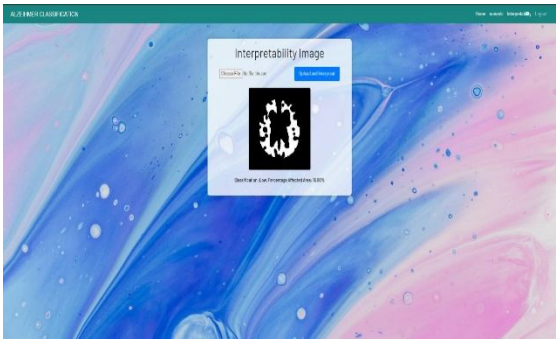

Fig7- XAI Based Classification of AD's

The above-mentioned image illustrates the XAI based Out-put which clearly shows us the affected area of the brain by Alzheimer's Disease. This output is generated with the help of LIME Algorithm usage in the backed which has been train with 3200 images, 32 images per batch with a less epochs loss.

VII. Conculsion And Future Scope

In conclusion, our project merges convolutional neural net-works (CNNs), improved landmark-based deep feature learning, and Explainable Artificial Intelligence (XAI) to innovate neuroimaging for better diagnosis of brain diseas-es, notably Alzheimer's. With CNNs and deep feature learn-ing, we enhance model interpretability using XAI, providing insights into decision-making processes. This integration not only boosts diagnostic accuracy but

also instills trust in pre-dictions. Additionally, the incorporation of k-Nearest Neighbors (KNN) and Naive Bayes algorithms further en-riches our approach. KNN employs similarity measures to classify new data points based on proximity to existing ones, while Naive Bayes offers a probabilistic method, as-suming feature independence for class likelihood calcula-tion. These additions expand our analysis, fortifying the robustness and adaptability of our diagnostic framework for brain diseases.

Integrating IoT into this project with XAI for Alzheimer's prediction using MRI holds promise. This future scope can be explored in the following ways:

Wearable sensors and cognitive monitors can develop wearable sensors that track daily activities, sleep patterns, and cognitive function. These sensors could collect data on gait, tremors, memory lapses, and reaction times. To antici-pate cognitive decline, or to identify a first sign of Alzhei-mer's disease, use AI models that have been tested on these sensor data. Explainable artificial intelligence can then be used to understand how the model arrived at its conclusion based on readings from sensors.

Expanding Detectable Conditions Train, the model to dis-tinguish between Alzheimer's and other forms of dementia, such as Parkinson's or frontotemporal dementia. To under-stand the decision-making process behind these differences, explore the possibility of using XAI.

Predicting Disease Progression Develop models that can predict the progression rate of Alzheimer's disease based on initial MRI scans. To enable personalized treatment plans, the XAI can be used to describe factors that influence the anticipated progression.

Identifies risk factors with XAI this can help detect particu-lar data elements from sensors or scans that contribute most to the model's predictions. It may also help researchers to understand the causes of memory loss.

REFERENCES

[1] D. K. Moorthy, N. P and S. J. Subhashini, "A Review on Alzheimer's Disease Detection using Machine Learning," 2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS),Trichy,India,2023,pp.573-581,doi: 10.1109/ICAISS58487.2023.10250457.

[2] S. Pallawi and D. K. Singh, "Detection of Alzheimer's Disease Stages Using Pre-Trained Deep Learning Approaches," 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), Hamburg, Germany, 2023, pp.252-256,doi: 10.1109/ICCCMLA58983.2023.10346730.

[3] A. Amrutesh, G. B. C G, A. A, A. R. KP and G. S, "Alzheimer's Disease Prediction using Machine Learning and Transfer Learning Models," 2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/CSITSS57437.2022.10026365.

[4] G. Lyu, "A Review of Alzheimer's Disease Classification Using Neuropsychological Data and Machine Learning," 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 2018, pp. 1-5, doi: 10.1109/CISP-BMEI.2018.8633126.

[5] R. Anitha, Prakash and S. Jyothi, "A segmentation technique to detect the Alzheimer's disease using image processing," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), Chennai, India, 2016, pp. 3800-3801, doi: 10.1109/ICEEOT.2016.7755424.

[6] Ullah, Naeem & Hassan, Muhammad & Khan, Javed & Anwar, M. & Khursheed, Khursheed. (2023). Enhancing explainability in brain tumor detection: A novel DeepEBTDNet model with LIME on MRI images. International Journal of Imaging Systems and Technology. 34. 10.1002/ima.23012.

[7] Coluzzi, Davide & Bordin, Valentina & Rivolta, Massimo & Fortel, Igor & Zhang, Liang & Leow, Alex & Baselli, Giuseppe. (2023). Biomarker Investigation using Multiple Brain Measures from MRI through XAI in Alzheimer's Disease Classification.

[8] Frisoni GB, Fox NC, Jack CR Jr, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol. 2010 Feb;6(2):67-77. doi: 10.1038/nrneurol.2009.215. PMID: 20139996; PMCID: PMC2938772.

[9] Hajamohideen, Faizal & Shaffi, Noushath & Mahmud, Mufti & Subramanian, Karthikeyan & Al Sariri, Arwa & Viswan, Vimbi & Abdesselam, Abdelhamid. (2023). Four-way classification of Alzheimer's disease using deep Siamese convolutional neural network with triplet-loss function. Brain Informatics. 10. 10.1186/s40708-023-00184-w.

[10] Bau D, Zhou B, Khosla A, Oliva A and Torralba A 2017 Network dissection: Quantifying interpretability of deep visual representations Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition pp 6541–9

[11] Böhle M, Eitel F, Weygandt M and Ritter K 2019a Layer-wise relevance propagation for explaining deep neural network decisions in MRI-Based Alzheimer's disease classification Frontiers Aging Neurosci. 11 194

[12] [Achilleos et al., 2020] Kleo G Achilleos, Stephanos Leandrou, Nicoletta Prentzas, Panayiotis A Kyriacou, Antonis C Kakas, and Constantinos S Pattichis. 2020. Extracting explainable assessments of Alzheimer's disease via machine learning on brain MRI imaging data. In 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE, 1036–1041.

[13] [Talo et al., 2019] Muhammed Talo, Ulas Baran Baloglu, Özal Yıldırım, and U Rajendra Acharya. 2019. Application of deep transfer learning for automated brain abnormality classification using MR images. Cognitive Systems Research 54 (2019), 176–188.