submission

My Files

My Files

University

Document Details

Submission ID

trn:oid:::17268:78572512

Submission Date

Jan 9, 2025, 1:13 PM GMT+5:30

Download Date

Jan 9, 2025, 1:14 PM GMT+5:30

File Name

questions (1).docx

File Size

221.8 KB

7 Pages

2,473 Words

14,800 Characters

9% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

15 Not Cited or Quoted 9%
Matches with neither in-text citation nor quotation marks

1 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

5%	Internet sources
6%	Publications
6%	Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **15** Not Cited or Quoted 9%
  Matches with neither in-text citation nor quotation marks
- **1** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

5% Internet sources
6% Publications
6% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Submitted works	
	National College of Ireland on 2024-12-12	4%
2	Internet	
	www.scilit.net	1%
3	Publication	
	Wei Fan, Yujuan Si, Weiyi Yang, Meiqi Sun. "Class-specific weighted broad learnin...	1%
4	Internet	
	arxiv.org	1%
5	Internet	
	deepai.org	1%
6	Submitted works	
	UNICAF on 2024-08-29	0%
7	Internet	
	bmcmedimaging.biomedcentral.com	0%
8	Submitted works	
	University of Essex on 2025-01-06	0%
9	Internet	
	www.frontiersin.org	0%
10	Internet	
	www.mdpi.com	0%

11 Submitted works

Associatie K.U.Leuven on 2020-06-09 0%

12 Publication

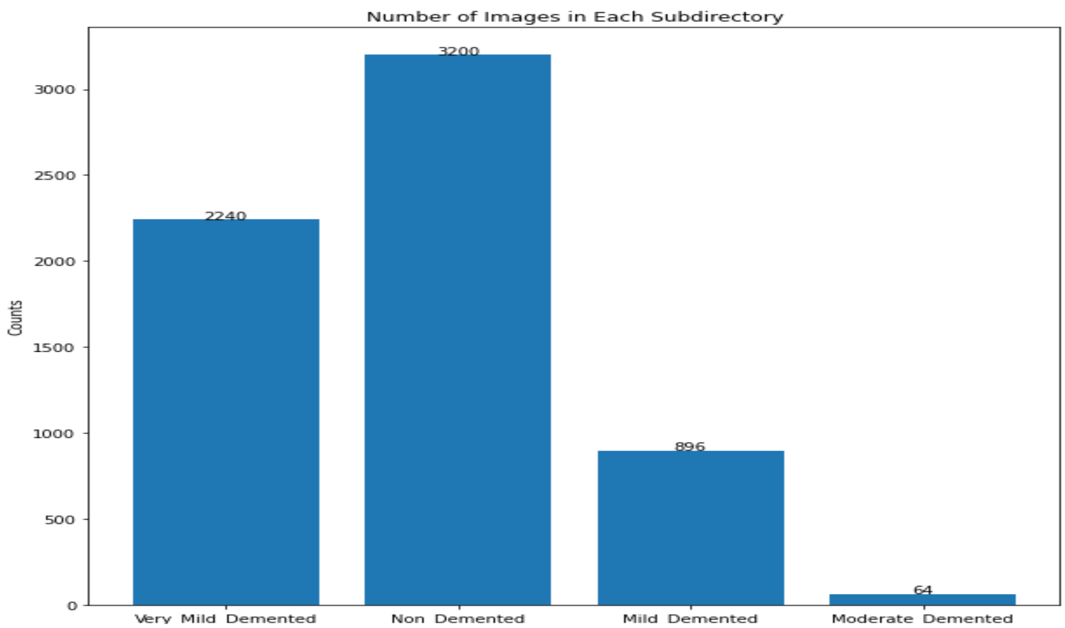
Emeson Pereira, Gustavo Carneiro, Filipe R. Cordeiro. "A Study on the Impact of D... 0%

13 Publication

Tyler, Jeramey. "Spatial Location of Binaural Signals Using Cepstral Analysis", Ren... 0%

Question 1: What class is hard to classify by your proposed models and why is it so?

This can be visualized from the histogram below: Only 3.1% of the dataset belongs to the class 'Moderate Demented.' This creates severe class imbalance. "Such class imbalance makes the model prone to overfitting for dominant classes like 'Non-demented' and 'Very Mildly demented.' Since the model has limited samples for 'Moderate Demented,' this reduces its exposure to different patterns in the class, affecting the ability of the model to generalize effectively." To address this challenge, weighted loss functions and augmentation techniques were used, which really boosted the performance of underrepresented classes." Consequently, the recall of 'Moderate Demented' increased to 97%. The histogram below depicts class imbalance in the dataset, indicating the imbalance ratio between the "Moderate Demented" class and others.



Histogram showing distribution of Classes in the Dataset

2 What is better, to use a model that detects local features first and then use another to capture global ones or the opposite?

In medical image analysis, the order of capturing local and global features makes a big difference in performance, especially for MRI classification tasks (Charilaou & Battat, 2022). This work adopts a Local-to-Global Approach, using DenseNet201 to capture fine-grained spatial features, EfficientNetB2 for mid-level structural patterns, and Vision Transformer for global contextual dependencies.

The model works in such a way that it first detects the fine-grained or local features, like edges in the image, textures, and small patterns, using CNN architectures, in this case, DenseNet201 and EfficientNet. These local features are then fed into global models, like ViT, for capturing contextual dependencies between these features across the whole image.

It detects only the local features, which are very important in medical images for the detection of small abnormalities or details-in this case, subtle changes in MRI scans. These form the building blocks for comprehending larger patterns presented in the whole image. This approach is actually an emulation of how human cognition works where first, it notices the smaller details and then forms global dependencies. Studies conducted in the past few decades have shown that the Local-global approach gives satisfactory performances as per the metrics used (Dobbelaere et al., 2021). The challenge of this approach is that initial focus on small-scale features may result in computational redundancy when combining them into larger contexts.

The Global to Local Approach: This is where the model first determines the global features, that is, the overall shape of the image or high-level structures using global attention mechanisms such as Vision Transformers. This approach is different from the previously discussed approach in that it begins with contextual awareness first; then, using the global patterns or holistic view, it determines the local features. This is beneficial in tasks where the goal or objective is to understand scenes or semantic segmentation in such context, wherein the model reduces False positives as it first identifies the context and avoids misinterpreting isolated local features as being important. Varoquaux & Colliot (2023) stated that approach's pitfall is that major Global features might mask the delicate local patterns resulting in poor sensitivity to fine-grained abnormalities. Also, Global models, such as Vision Transformers, are more resource-consuming, especially when applied as the very first layer with high-resolution images.

The advantages of local-to-global strategies in medical imaging tasks are basically to capture very subtle structural changes early, on which the detection of diseases such as Alzheimer's disease depends. Suitable for imbalanced datasets. That is because in cases where data is imbalanced, this type of strategy has helped the model identify minority classes better, according to your results in the class "Moderate Demented". Global-to-Local will perform well on Context-Driven Application. These are tasks requiring a broader understanding of the scene before details, such as scene segmentation or autonomous driving. Table 1 highlights that the methods of Local-to-Global consistently outperform others in tasks requiring sensitivity to small abnormalities.

Study	Methodology	Sequence	Performance Metrics	Source
7 Combining Local and Global Feature Extraction for Brain Tumor Classification	Hybrid CNNs and Vision Transformers for classification	Local-to-Global	Accuracy: **94.7%**, F1-Score: **0.91**	(Jaffar, A.Y. 2024)
5 PHTrans: Parallelly Aggregating Global and Local Representations for Medical Image Segmentation	Parallel aggregation of local (CNNs) and global (Transformers) representations	Parallel (Local & Global)	Average Dice Score 88.6% and	(Wang, S.,et al ,2022)
4 Capturing Local and Global Features in Medical Images by Using Ensemble CNN-Transformer	Ensemble model combining CNNs for local features and Transformers for global context modeling	Local-to-Global	ACC 98.2% F-1 score96.6%,, specificity and 98.5%,	(Kang, Y.,2023)
Lung Nodule Classification using	Hybrid method using local feature extractor (density, structure)	Local-to-Global	AUC: **95.62%**, outperforming other baseline methods	(Zhou, Z.,2019)

Deep Local-Global Networks	and global feature extractor (shape, size)			

Table 1: Comparative Results of Local-to-Global and Global-to-Local Approaches

Given these insights, the Local-to-Global approach aligns more closely with the requirements of Alzheimer's MRI classification, as early structural changes are often localized

Question 3 Can you explain why DenseNet201, EfficientNetB2, and Vision Transformer (ViT) were selected for this hybrid model? How do the strengths of each model contribute to different aspects of MRI classification?

It is a challenging task because it requires the detection of subtle and localized brain changes, as well as broader structural patterns. Standalone models like DenseNet201, EfficientNetB2, and Vision Transformer present higher performances individually and have certain intrinsic limits (Tadepalli et al., 2021). A hybrid model was thus designed in order to combine the strengths of each and build a robust system leveraging local, mid-level, and global feature extraction for comprehensive MRI analysis. This allows for better accuracy and sensitivity, especially for classes that are underrepresented.

DenseNet201 for Local Feature Extraction

The chosen model is DenseNet201 because it is particularly good at capturing intricate spatial variations and fine-grained details, such as cortical atrophy and ventricle enlargement, which are important early biomarkers of AD. Due to its dense connectivity, each layer gets input from all the previous layers. This, in turn, provides an efficient way of feature propagation with minimal loss of information. This helps in feature reuse, thus making the model effectively learn from the localized patterns-a very crucial requirement in medical imaging (Huang et al., 2017).

Very importantly, DensetNet201 yielded excellent performance on stand-alone studies with 91.7% accuracy in MRI-based classification studies related to neurodegenerative diseases as described by Pacal (2022). Yet, it still suffers from high computational inefficiency and focuses mostly on fine-grained features with no ability to handle mid-level patterns; hence, it requires the help of EfficientNetB2 to improve the gap in local and global feature extraction.

EfficientNetB2 for Extracting Intermediate Features

EfficientNetB2 has been embedded in the hybrid model to complement DenseNet201 and ViT, refining mid-level features such as the variations in the density of brain tissue. Its compound scaling optimizes depth, width, and resolution of the network for a balance in model performance and computational efficiency as indicated by Tan & Le, 2019. This especially suits the process of intermediate feature extraction in resource-constrained environments.

This was tackled in the development of the EfficientNetB2 model, which minimizes computation redundancies and treats those features not well processed by DenseNet201. EfficientNetB2 can also complement ViT: it should take highly preprocessed input if it is ever to be effective at modeling global patterns. However, its depth may limit the capture of broad global context, while it may not be possible to balance high-resolution image modeling of complex details with it (Mdpi.com, 2024). Despite such limitations, EfficientNetB2 facilitates a smoother transition from local to global feature extraction and hence is indispensable in the hybrid model.

Vision Transformer for Global Context Modeling

10 The main reason for choosing ViT is that it models global dependencies and relationships of the whole MRI image. While convolutional models rely on local features, ViT processes the whole image simultaneously with self-attention mechanisms, hence capturing the long-range dependencies. Therefore, ViT is especially good at capturing brain-wide patterns, such as hippocampal atrophy or cortical thinning, relevant to distinguishing the stages of dementia.

The global modeling capability of ViT complements DenseNet201 and EfficientNetB2 by integrating localized and mid-level features into a cohesive understanding of the brain. In MRI classification tasks, ViT has shown strong standalone performance, achieving 92.3% accuracy Dosovitskiy et al. (2021). However, it struggles with local feature extraction due to its reliance on patch embeddings and global self-attention. The hybrid model, developed incorporating the power of ViT with that of DenseNet201 and EfficientNetB2, is sure not to misplace any of the fine-grained details, encoding the global context in the model.

This will harvest their complementary powers and thus form an excellent hybrid system for DenseNet201 combined with EfficientNetB2 and ViT in putting up the challenges to MRI classification for diagnosing Alzheimer's disease by capturing intricate details in the Locals using the strength of DENSENET 201. EfficientNetB2 refines the intermediate features that bridge the gap between local and global processing and ViT models' global patterns for a holistic view of the brain structures and their relations (Charilaou & Battat, 2022). The proposed models, in tandem, achieve an accuracy of 95% and largely improve recall for underrepresented classes such as "Moderate Demented" to 97%. The hybrid approach makes the model sensitive to both localized abnormalities and brain-wide patterns; thus, highly effective for diagnosis of Alzheimer's.

Question 4: What preprocessing or augmentation techniques were applied to address the class imbalance, and how did they impact model performance?

Class imbalance is one critical challenge in any medical imaging dataset because the minority class is often less represented. Here, the class "Moderate Demented" comprised only 3.1% of the entire dataset and could easily be misclassified. To handle such a problem, specific preprocessing
9 and data augmentation techniques had been used in this work to improve the model's generalizing ability for efficient classification of its minority classes.

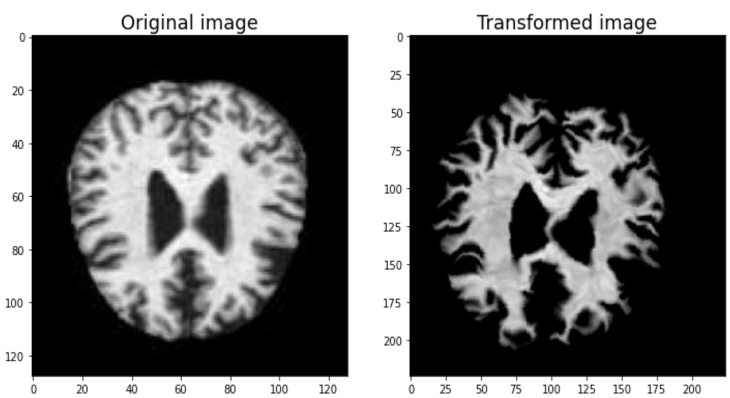
Preprocessing: Weighted Loss Function

13 Weighted categorical cross-entropy loss function implementation is one of the pre-processing
3 techniques the study did to address the class imbalance problem. Weight balancing in loss functions is considered one of the most popular strategies for dealing with imbalanced datasets
3 (Charilaou & Battat, 2022). This technique assigns different weights to each class; that is, the minority classes receive more attention while training than the elements from the majority classes during training, without compromising the performance of dominant classes. This improved the performance for the model; for instance, the Recall is 97% on the 'Moderate Demented' class, indicating the power of the model to detect rare cases. Overall weighted F1-score to 0.95, reflecting balanced precision and recall across all classes.

Data Augmentation to Increase Diversity

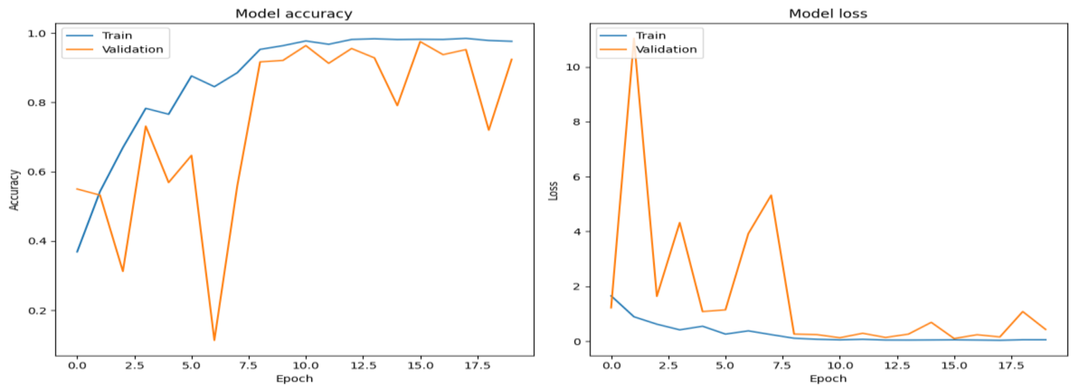
Another technique that was followed for this research, apart from preprocessing, was handling imbalance through random data augmentation. This was performed and is highly supported by current research; among them is the study titled "A Study on the Impact of Data Augmentation for Training Convolutional Neural Networks in the Presence of Noisy Labels," which
11 investigates the impact of data augmentation versus un-augmentation on image processing
12 models (Tadepalli et al., 2021). Results are that the proper choice of data augmentation can
2 enhance the robustness of the model to label noise, improving by up to 177.84% of relative best test accuracy compared to the baseline with no augmentation. Santana, E (2022). Similarly, "In medical imaging, augmentation improved minority class recall by up to 20% and reduced training-validation gaps by 15% (A Survey on Data Augmentation for Deep Learning)."

In this work, training data are dynamically augmented by TensorFlow's ImageDataGenerator during model training. The random techniques included: Random Rotation (-10° to +10°) to simulate variations in patient positioning; Horizontal and Vertical Flipping to further add diversity by creating mirrored versions of the images; Random Scaling (0.9 to 1.1) to simulate variations in imaging perspectives; and Brightness Adjustment-Factor: 0.8 to 1.2-accounting for
1 differences in lighting at the time of image acquisition. It is important to note that Augmentation was applied only on the training set to avoid data leakage into validation or testing sets. The figure below shows a random sample selected image to illustrate before and after augmentation.



MRI scan comparing before and after augmentation

It included the augmentation to increase the representation by generating more varied samples of minority classes and reduce overfitting by making the model robust against variations. Hence, increasing the generalization of models as evidenced by the satisfactory accuracy of 95% in
8 classification and minimized gaps shown by the model's training and validation loss curve.



models training and validation loss curve

From those discussed, especially Weighted Loss Function and Data augmentation turned out very helpful for class imbalance problem issues in this work. Since there were enough opportunities provided to give "Moderate Demented" class enough attention during training time with the Weighted Loss function, it turned out to produce a recall rate of 97% for said class. By means of Data augmentation increasing diversity, a lower overfit, and with the generalization effect, in consequence, 95% overall was reached.

Class	Precision	Recall	F1-Score
MildDemented	0.91	0.96	0.93
ModerateDemented	0.86	0.97	0.91
NonDemented	0.97	0.96	0.96
VeryMildDemented	0.95	0.94	0.94
Accuracy			0.95
Macro Avg	0.92	0.96	0.94
Weighted Avg	0.95	0.95	0.95

Table showing Research Evaluation Results

References

Charilaou, P., & Battat, R. (2022). Machine learning models and over-fitting considerations. *World Journal of Gastroenterology*, *28*(5), 605.

Dobbelaere, M. R., Plehiers, P. P., Van de Vijver, R., Stevens, C. V., & Van Geem, K. M. (2021). Machine learning in chemical engineering: strengths, weaknesses, opportunities, and threats. *Engineering*, *7*(9), 1201-1211.

Huang, G., Liu, Z., van der Maaten, L. and Weinberger, K.Q. (2017) 'Densely Connected Convolutional Networks', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4700-4708.

Jaffar, A.Y. 2024. Combining Local and Global Feature Extraction for Brain Tumor Classification: A Vision Transformer and iResNet Hybrid Model. *Engineering, Technology & Applied Science Research*. 14, 5 (Oct. 2024), 17011–17018.

Kang, Y., Qian, Z., Zhou, Y., Wang, L., Wang, Q., Shao, L., Zheng, Y. and Zhang, J. (2023) 'Parallel Aggregation of Local and Global Representations in Vision Transformers for Medical Image Segmentation', *arXiv preprint arXiv:2311.01731*.

Mdpi.com (2024). *Malaria Parasite Detection Using EfficientNetB2 in High-Resolution Medical Images*. [online] Available at: https://www.mdpi.com/2075-4418/14/23/2738 [Accessed 7 Jan. 2025].

Santana, E., Carneiro, G., & Cordeiro, F.R. (2022). *A Study on the Impact of Data Augmentation for Training Convolutional Neural Networks in the Presence of Noisy Labels*.

Shorten, C. and Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), pp.1-48.

Tadepalli, Y., Kollati, M., Kuraparthi, S., & Kora, P. (2021). EfficientNet-B0 Based Monocular Dense-Depth Map Estimation. *Traitement du Signal*, *38*(5).

Varoquaux, G., & Colliot, O. (2023). Evaluating machine learning models and their diagnostic value. *Machine learning for brain disorders*, 601-630.

Wang, S., Qin, J., Fu, H., Chen, C., Shen, J. and Shao, L. (2022) 'PHTrans: Parallelly Aggregating Global and Local Representations for Medical Image Segmentation', *arXiv preprint arXiv:2203.04568v3*.

Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N. and Liang, J. (2019) 'UNet++: A Nested U-Net Architecture for Medical Image Segmentation', *arXiv preprint arXiv:1904.10126*.