Review of: "An Improved Hybrid Transfer Learning-Based Deep Learning Model for Alzheimer's Disease Detection Using CT and MRI Scans"

Roseline Ogundokun[1]

1 Landmark University

Potential competing interests: No potential competing interests to declare.

The research on detecting Alzheimer's disease using pre-trained deep learning models showcases notable progress, although it also highlights certain aspects that should be improved. These include:

1. The research used the ADNI dataset of 3400 photos, which may not accurately reflect the larger population impacted by Alzheimer's disease. This constraint has the potential to provide partial outcomes that may not apply to diverse demographic cohorts or clinical environments. To improve the strength and applicability of the model, it is advisable to integrate a more comprehensive dataset, including photos from multiple demographic groups and clinical contexts.
2. The paper discusses the problems of underfitting and overfitting, but it does not provide precise information on the measures used to reduce overfitting. Overfitting may have a substantial influence on the performance of a model when it is applied to fresh, unexplored data. The authors must provide a more comprehensive elucidation of the methodologies used to mitigate overfitting. Possible avenues for exploration and discussion include techniques like cross-validation, regularization, and dropout layers.
3. The research employs data augmentation methods, including horizontal flipping and rotating photos by 5 degrees. Nevertheless, the degree to which these strategies enhance model performance remains unquantified. Subsequent versions of the research should include an examination of the influence of various data augmentation approaches on the precision and applicability of the model.
4. The research mainly depends on conventional criteria like accuracy, precision, and recall. Although essential, these indicators do not comprehensively represent model performance, particularly in datasets with unbalanced classes. Supplemental assessment measures, such as the AUROC and AUPRC curves, may provide a more thorough comprehension of the model's effectiveness.
5. The research evaluates the effectiveness of several pre-trained models but does not compare them to the most advanced models currently used for Alzheimer's detection. To determine the effectiveness of the suggested method, it would be advantageous to evaluate its performance compared to established cutting-edge models in identifying Alzheimer's disease.
6. The research lacks empirical verification of the results. Validating the model in a clinical setting is essential for determining its practicality in real-world scenarios. Subsequent investigations should focus on confirming the model's discoveries in a clinical environment, maybe by partnering with healthcare experts and organizations.
7. The research does not examine the model's capacity to explain its conclusions, a crucial aspect in medical applications

for establishing confidence among healthcare practitioners. Incorporating and reviewing methodologies for model interpretability, such as Layer-wise Relevance Propagation (LRP) or SHAP values, would augment the research's impact in the domain.