**A Multi-Task Deep Learning Model for Dementia Presence and Severity Prediction from Brain MRI**


Ayesha Anzer, Anthony Aighobahi. *Gerontechnology* 25(s)


**Purpose** Dementia affects over 55 million people worldwide. Cases are expected to double by 2030, nearly [1]. Early and accurate detection of dementia and its severity from brain MRI is critical for timely intervention and care planning. Deep learning models trained on MRI data have shown promise in automating Alzheimer's disease diagnosis [2]. However, most prior studies address either disease presence or stage classification in isolation. This study aims to develop a multitask deep learning model that can predict both the presence of dementia (binary classification) and clinical severity (ordinal three-class: non-demented, very mild, mild/moderate) from structural MRI. This approach provides a comprehensive, clinically oriented decision-support tool. **Method** We developed a multitask convolutional neural network based on the EfficientNet-B0 architecture. The model includes one output head for binary dementia classification and another for ordinal severity classification. The public OASIS MRI dataset [3] of older adults was used. We performed a stratified subject-wise split to ensure balanced representation of non-demented and demented cases in both training and testing sets. Each subject's T1-weighted MRI scans were preprocessed, and inputs were augmented to improve generalizability. The model was trained end-to-end to optimize both tasks simultaneously. We used a weighted multi-task loss: cross-entropy for presence and an ordinal regression loss for severity. To mimic clinical practice, predictions from multiple MRI slices or sessions for a subject were aggregated by majority vote or confidence averaging to yield subject-level diagnoses. Decision probability thresholds for binary classification were tuned on a validation set to balance sensitivity and specificity for detecting even very mild dementia. Performance was evaluated on a held-out test set at both the image and subject levels. Evaluation metrics included accuracy, area under the ROC curve (AUC) for binary outcomes, macro-average F1 score for ordinal outcomes, and confusion matrices. Key metrics are summarized in Table 1. **Results and Discussion** The EfficientNet-B0 multi-task model achieved high accuracy on both tasks. This highlights the benefit of joint learning. For detecting the presence of dementia, the model achieved a subject-level accuracy of 91% and an image-level accuracy of 88%. AUC was 0.95, and the macro-F1 was approximately 0.90 after threshold tuning. The three-class dementia severity classification achieved 82% subject-level accuracy (image-level 75%) with a macro-F1 of 0.80. The model accurately distinguished non-demented, very mild, and mild/moderate cases. Confusion matrix analysis showed that most misclassifications occurred between adjacent severity categories (very mild vs. mild/moderate). This finding aligns with known clinical challenges in early dementia staging. The model was highly sensitive in detecting very mild cases, suggesting potential utility for early intervention. The slight performance boost at the subject level over single-scan predictions indicates that aggregating multiple images per patient can reduce variability and improve diagnostic confidence. Clinically, these findings show that a lightweight CNN like EfficientNet-B0 can deliver practical results highly relevant to patient care by jointly identifying dementia and grading its severity. A high AUC for dementia detection indicates that the model can reliably flag pathological changes, supporting early recognition, while its severity grades could guide individualized care pathways. Embedding such a model into imaging workflows enables rapid, automated alerts for probable dementia, helping clinicians triage and plan follow-up efficiently. Evaluating subject-level diagnoses using stratified splits mirrors routine clinical practice, where patient-level judgement is critical. By tuning the model's output to prioritize sensitivity, especially for early disease, the tool can better support timely interventions. Overall, this work takes a step toward clinically integrated AI systems, advancing the use of deep learning to streamline dementia screening and enable more proactive patient management.

**References**


1. Alzheimer's Disease International (ADI). (2021). *Dementia statistics*. Retrieved from the ADI website.
2. Carcagnì, P., Leo, M., Del Coco, M., Distante, C., & De Salve, A. (2023). Convolution neural networks and self-attention learners for Alzheimer's dementia diagnosis from brain MRI. *Sensors, 23*(3), 1694.
3. Marcus, D. S., Wang, T. H., Parker, J., Csernansky, J. G., Morris, J. C., & Buckner, R. L. (2007). Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle-aged, non-demented, and demented older adults. *Journal of Cognitive Neuroscience, 19*(9), 1498–1507



**Keywords:** Dementia, Brain MRI, Multi-task Learning, Alzheimer's Disease, Clinical Decision Support



**Affiliation:** Faculty of Engineering and Applied Science, University of Regina, Canada
**Email:** aag833@uregina.ca



**Acknowledgement:** The authors declare that this research received no external funding and was conducted using publicly available data.


***Table 1. Key Performance Metrics by Task and Data Level***

| Task | Data Level | Accuracy | AUC | Macro-F1 |
|---|---|---|---|---|
| Dementia Presence (binary) | Image-level | 88% | 0.93 | 0.88 |
| Dementia Presence (binary) | Subject-level | 91% | 0.95 | 0.90 |
| Dementia Severity (3-class) | Image-level | 75% | 0.85 | 0.73 |
| Dementia Severity (3-class) | Subject-level | 82% | 0.90 | 0.80 |

**Note:** AUC for multi-class severity is reported as a macro-average of one-vs-all ROC curves for the three classes.