Intelligent Web Application based on RF and CNN for the Multiclass Classification of Cognitive Impairment in Older Adults

Joseph Ramirez
Peruvian University of Applied Sciences
Lima, Perú
u20211c828@upc.edu.pe

Sara Estrada
Peruvian University of Applied Sciences
Lima, Perú
u20211a109@upc.edu.pe

Willy Ugarte*
Peruvian University of Applied Sciences
Lima, Perú
willy.ugarte@upc.pe

Abstract - Cognitive impairment remains a growing public health challenge in aging populations, particularly in regions with limited access to specialized diagnosis tools. This study proposed an intelligent web-based system for the multiclass classification of cognitive impairment in older adults, based on structured clinical data and magnetic resonance imaging (MRI). The system was designed to offer a practical, low-cost alternative to traditional methods, enabling healthcare professionals to make accurate early diag-noses using accessible data. The platform includes two predictive modules: one based on Random Forest using clinical and func-tional data, and another based on Convolutional Neural Networks (CNN) for MRI classification. To assess performance, we trained and tested models on validated datasets, optimizing hyperparameters with Optuna. The Random Forest model achieved an accuracy of 81.8% and an F1 score of 0.868 for classifying patients into Control or MCI+Dementia, while the CNN model reached 99% across all metrics using only MRI scans. The system also includes a visualization dashboard to facilitate diagnostic decisions. These results suggest that machine learning models can complement clinical judgment and support timely diagnosis in resource-limited settings. The proposed tool is promising for scalable deployment and can be extended with multimodal data in future research.

Keywords: Cognitive Impairment, Convolutional Neural Networks, Random Forest, MRI Classification, Geriatric Assessment

I. Introduction

Cognitive impairment in older adults is a growing global public health concern. Worldwide, over 60% of dementia cases remain undiagnosed due to limited access to diagnostic tools and specialized personnel, particularly in low- and middle-income countries [1]. In Latin America, this challenge is especially acute, where resource constraints and lack of specialized care hinder timely diagnosis [2]. According to the Ministry of Health (MINSA), in countries such as Peru, cognitive disorders remain significantly underdiagnosed outside major urban centers [3]. Traditional diagnostic approaches often require costly neuroimaging like Magnetic Resonance Imaging (MRI) or expert-administered cognitive tests, creating barriers for early intervention in resource-limited settings. These concerns underscore the urgent need for accessible, intelligent diagnostic solutions to support early and effective detection of cognitive impairment.

*Both authors contributed equally to this research.

Early detection of cognitive impairment enables timely intervention and cognitive rehabilitation strategies that can delay or prevent progression to dementia [4]. However, existing platforms are often inaccessible to general practitioners due to their cost, technical complexity, or lack of integration with clinical workflows. Our goal is to bridge this gap by developing an intelligent web application, **IntelliCog**, capable of automatically classifying cognitive status—Control, Mild Cognitive Impairment (MCI), or Dementia—based on structured clinical data and MRI images. This platform is designed to be scalable across diverse healthcare contexts, while its validation phase focuses on clinical environments in Peru.

Detecting cognitive impairment from structured data and MRI is challenging due to the heterogeneity and complexity of medical records. Conventional rule-based systems are inflexible and fail to capture non-linear relationships between variables. Additionally, while neuroimaging offers valuable insights, converting high-dimensional MRI data into interpretable outputs requires deep learning models with significant computational requirements. Many existing solutions either overfit small datasets or rely on features that are not routinely available in standard clinical practice.

Previous works often rely on expensive or invasive features such as Cerebrospinal Fluid (CSF) biomarkers, Electroencephalograms (EEGs), or high-resolution MRIs from controlled research cohorts [5][6]. Moreover, studies using multimodal machine learning approaches have rarely integrated functional and biochemical data from real-world geriatric populations in Latin America. Unlike these methods, IntelliCog focuses on data routinely collected in outpatient care—Activities of Daily Living (ADL)/Instrumental Activities of Daily Living (IADL) scores, creatinine levels, hemoglobin, age, and comorbidities—paired with open-access neuroimaging data (OASIS-3) [7], to create an interpretable and deployable solution.

Our proposal combines a Random Forest (RF) model trained on clinical data from 283 patients at the Hospital Geriátrico Juan Pablo II in Katowice, Poland, and a Convolutional Neural Network (CNN) trained on sagittal MRI slices from the OASIS-3 dataset. Preprocessing methods include min-max normalization, Pearson correlation for multicollinearity reduction, and image augmentation. RF was trained with 100 trees, max depth 12, and Gini impurity. The CNN uses three convolutional layers with ReLU and max pooling. Although our system provides a scalable solution, limitations include dependency on labeled MRI data and the need for further validation in diverse hospital settings, including those in Peru.

Our main contributions are as follows:

- We construct a multimodal dataset by integrating structured clinical data from a geriatric hospital in Poland and MRI images from the OASIS-3 repository.
- We develop an intelligent platform that uses Random Forest and CNN models to classify cognitive impairment into three categories: Control, MCI, and Dementia.
- We validate the proposed system technically (via cross-validation and performance metrics) and discuss its scalability for deployment in resource-constrained healthcare services.

This paper is organized as follows: Section 2 presents an overview of related work in machine learning applied to cognitive impairment. Section 3 details the datasets used, preprocessing steps, model architecture, and platform design. In Section 4, we present experimental results, model configurations, and performance metrics. Finally, Section 5 concludes the paper and discusses future research directions.

II. Related work

In recent years, various machine learning approaches have been applied to the classification of cognitive impairment using clinical, biochemical, and neuropsychological data. Although significant progress has been made, existing systems often rely on specialized equipment or data sources that are not commonly available in standard clinical practice. The following studies highlight recent contributions and help contextualize the novelty of our proposal.

To begin with, in [8], the authors evaluated several machine learning models - including support vector machines, linear discriminant analysis, and random forests - to classify cognitive states using neuropsychological tests (MMSE, MoCA, KDSQ) from 705 older adults. Their results demonstrated strong performance, with Random Forest achieving the highest accuracy in multiclass classification. However, it is important to note that this study relied exclusively on cognitive test scores. In contrast, our proposal integrates additional biochemical and functional variables, making it more robust and representative of real-world clinical settings where multimodal data is more feasible.

Moreover, [9] proposed a model to classify cognitive impairment in patients with cerebrovascular disease based on dual-task gait analysis. Using machine learning algorithms such as SVM and Random Forest on a dataset of 47 participants, the study reported a balanced accuracy of 82.7%. Although the results are promising, the need for gait measurement systems limits its scalability. Unlike [9], our approach avoids reliance on motion-capture technology and instead leverages routinely collected clinical data, enhancing its applicability in non-specialized geriatric care.

In addition, [10] applied Random Forest and XGBoost algorithms to data from the NHANES cohort, using 64 features including blood biomarkers and audiometry thresholds to classify cognitive impairment. Their model achieved an AUC of 0.834, validating the predictive power of accessible variables. Nevertheless, the inclusion of audiometric data and other complex variables may pose integration challenges in typical clinical workflows. Unlike [10], our system strictly prioritizes clinical, biochemical, and functional features that are standard in outpatient geriatric evaluations, increasing its potential for adoption.

Furthermore, [11] developed a predictive model using XGBoost and SHAP analysis to identify patients at risk of progressing from MCI to Alzheimer's disease. The study introduced a NEAR-MISS undersampling technique to address data imbalance and incorporated high-cost biomarkers such as MRI and CSF data. While the predictive accuracy was improved, the model's dependency on advanced imaging restricts its use to research environments. By contrast, our system is designed for early detection in primary care using low-cost variables, without sacrificing interpretability or clinical relevance.

Lastly, [12] introduced a DenseNet-based deep learning framework trained on EEG-derived images for dementia classification, achieving a notable accuracy of 94.17%. Although this approach demonstrates the power of physiological signals and deep models, it depends on the availability of EEG equipment and complex preprocessing pipelines. In contrast, our system will adopt a multimodal architecture that combines structured clinical, biochemical, and functional data with EEG-derived features. This integration aims to improve diagnostic accuracy while maintaining interpretability through the use of ensemble models such as Random Forest. Moreover, by incorporating low-cost EEG headsets compatible with outpatient settings, our proposal bridges the gap between signal-based precision and real-world usability, thus expanding accessibility to early cognitive screening technologies.

III. Contribution

3.1 Context

3.1.1 Preliminary Concepts

In this section, we present the main concepts underpinning our work. Our intelligent platform aims to classify cognitive impairment in older adults using a Random Forest-based model. The system integrates biometric data, clinical history, cognitive test results, comorbidities, and functional assessments to support accurate and explainable decision-making.

3.1.2 Key Definitions

Definition 3.1 (Cognitive Impairment [13].] Cognitive impairment refers to when a person has trouble remembering, learning new things, concentrating, or making decisions that affect their everyday life. It can range from mild to severe and is often assessed through standardized cognitive tests.

Definition 3.2 (Random Forest [14].] Random Forest is a supervised machine learning algorithm that builds multiple decision trees during training and outputs the mode of the classes (for classification) as the final prediction. It is known for handling large datasets with high dimensionality and for its robustness to overfitting.

Definition 3.3 (Convolutional Neural Network[15].] A Convolutional Neural Network (CNN) is a class of deep neural networks commonly used for processing data that has a grid-like topology, such as images. CNNs are composed of multiple layers, including convolutional layers, pooling layers, and fully connected layers, which automatically and adaptively learn spatial hierarchies of features from input data.

3.2 Methods

This study will propose the design and implementation of **IntelliCog**, a smart web platform for the automated classification of cognitive impairment in older adults using machine learning algorithms, specifically Random Forest. The proposed method will be structured in six stages: selection of the classification algorithm, data preprocessing and partitioning, architectural design, model training, interface development, and clinical-technical validation of the platform.

3.2.1 Classification Algorithm Selection

To classify cognitive states, we will adopt the Random Forest (RF) algorithm, an ensemble model based on decision trees that has shown strong robustness and accuracy in multiclass clinical tasks, particularly when working with heterogeneous and correlated features[16]. Compared to SVM or deep neural networks, RF offers key advantages in medical environments:

- Handling of missing and categorical data without complex transformations;
- Interpretability through feature importance rankings;
- Capability to perform well on small or medium-sized datasets with good generalization.

RF has been successfully applied in previous work, such as [16], who classified cognitive states using NHANES data. In contrast, IntelliCog will integrate functional (ADL, IADL), biochemical (urea, creatinine), and localized sociodemographic variables, delivering a contextualized approach suited to geriatric clinical environments.

3.2.2 Data Preparation and Preprocessing

The development of the IntelliCog system will rely on two distinct datasets corresponding to the two classification models proposed: a clinical-based model and an image-based model using MRI.

1. Clinical Dataset. The clinical model will be trained on a publicly available dataset containing structured medical data from 283 older adults collected at the John Paul II Geriatric Hospital in Katowice, Poland, between 2015 and 2019 [17]. This dataset includes patients labeled as Control, Mild Cognitive Impairment (MCI), or Dementia based on cognitive screening tools such as MMSE, MoCA, and CDT. It comprises sociodemographic data (age, sex, education), clinical scores (ADL, IADL, BERG), lifestyle indicators (stress, BMI), biochemical markers (creatinine, hemoglobin, vitamin D and B12, TSH, sodium, etc.), and comorbidities (e.g., hypertension, diabetes, atherosclerosis). The dataset is openly available at:

- **DOI:** 10.1038/s41598-025-90460-y

The preprocessing steps for this dataset include:

- Missing value imputation using k-NN (k=3);
- Normalization of continuous features using Min-Max scaling;
- One-Hot Encoding for categorical variables;
- Multicollinearity reduction using Pearson correlation ($\rho > 0.85$).

2. MRI Image Dataset. For the CNN-based classification model, we will utilize the MRI subset from the OASIS-3 dataset[18], which includes thousands of T1-weighted sagittal MRI scans of older adults diagnosed as cognitively normal, MCI, or Alzheimer's disease. These images are labeled with diagnostic outcomes and matched with relevant cognitive and demographic metadata, offering an ideal source for training deep learning models.

- **Dataset:** OASIS-3
- **Access:** https://www.oasis-brains.org

The preprocessing pipeline for the MRI data will consist of:

- Selection of T1-weighted sagittal slices as input to the CNN;
- Intensity normalization and resizing to 224×224 pixels;
- Data augmentation (rotation, flipping, contrast adjustment);
- Conversion to grayscale tensor arrays;

By integrating these datasets, IntelliCog aims to support both structured-data and image-based classification, enhancing diagnostic precision and model robustness.

3.2.3 Model Training with Random Forest

The Random Forest model will be trained with 100 trees, maximum depth of 12, Gini as the split criterion, and random feature selection per split. Hyperparameter tuning will be performed using 5-fold cross-validation and grid search on the following:

3.2.4 IntelliCog System Architecture

The web-based platform will adopt a modular, scalable architecture composed of five layers, as shown in Figure 5.

- **User Layer**: Authentication, user management, and patient tracking for clinicians;
- **Web Interface Layer**: Built with React and TailwindCSS to enable clinical data entry, visualization of results, and report downloads;
- **Business Logic Layer**: Handles validation, invokes the trained RF model, and generates interpretation outputs;
- **Backend/API Layer**: Developed in Python with FastAPI to expose REST endpoints for model prediction;
- **Storage Layer**: PostgreSQL relational database stores users, patient records, model predictions, and reports.

This architecture will support deployment via Heroku, AWS, or GCP and allow horizontal scaling.

3.2.5 Physical Architecture of the System

The physical architecture of the system is organized into three main layers: presentation, processing, and data. This structure enables a clear separation of concerns, which facilitates system maintenance and scalability.

- **Presentation Layer:** Represented by the web interface used by the geriatric physician. Actions such as registering patient information, viewing profiles, performing cognitive assessments, and uploading MRI images are carried out here.
- **Processing Layer:** Contains APIs responsible for managing patient data and performing cognitive detection using a classification model. It also manages structured data (in JSON format), allowing seamless interaction between layers.
- **Data Layer:** Includes the server database (BBDD) for storing structured patient information and the S3 server for storing MRI images in PNG format.

3.2.6 Web Interface and Interactive Modules

The UI will be designed in Figma and built with React to offer a geriatrician-centered UX. The key modules to be implemented include:

- User login and registration;
- Patient record registration and editing;
- Entry of clinical, functional, and biochemical data;
- Cognitive status classification and output visualization;
- Report generation in PDF format and email delivery;
- External validation panel for comparison with clinical diagnosis.

The platform will be entirely in Spanish and include accessibility features and online help (see Figure 1).

Fig. 1. Landing page of the IntelliCog platform

3.2.7 Clinical and Technical Validation

The system's effectiveness will be validated in two stages:

- **Technical Validation:** Model evaluation using local clinical datasets with standard metrics;
- **Clinical Validation:** Testing with anonymized international data and feedback from six geriatricians to assess usability and interpretability.

These efforts will ensure the model's robustness and the platform's practical utility in supporting cognitive diagnosis.

IV. Experiments

4.1 Protocol

4.1.1 Environment Setup

The experiments were conducted on a virtual machine provided by **Google Colab**, with the following technical configuration:

- **RAM:** 12.6 GB
- **GPU:** NVIDIA Tesla T4
- **CPU:** 2 virtual Intel Xeon cores at 2.20GHz
- **Operating System:** Ubuntu 18.04 LTS
- **Programming Language:** Python 3.10

4.1.2 Libraries Used

The following Python libraries were employed:

- `scikit-learn` – for classification models, preprocessing, metrics, and feature selection.
- `imbalanced-learn` – for handling imbalanced classes using SMOTE (Synthetic Minority Over-sampling Technique).
- `optuna` – for hyperparameter optimization.
- `pandas`, `numpy`, `matplotlib`, `seaborn` – for data manipulation and visualization.
- `torch` and `torchvision` – for the design, training, and evaluation of neural networks.
- `PIL` – for image loading and manipulation.
- `os` – for file and directory path management.

4.1.3 Random Forest Training Process

A preprocessing *pipeline* was built with the following steps:

(1) Imputation of missing values using the mean with `SimpleImputer`.
(2) Feature scaling using `StandardScaler`.
(3) Selection of the 10 most relevant features using `SelectKBest` with `f_classif`.

The selected variables were:

```
['Age', 'Years of study', 'D3', 'Uric
Acid', 'HA', 'ADL', 'IADL', 'BERG', 'Stress',
'BMI']
```

Ten classical classification models were trained and evaluated with their default configurations. Subsequently, the **Random Forest** model was selected due to its strong performance, and its hyperparameters were fine-tuned using the `Optuna` library with 50 iterations.

4.1.4 CNN Training Process

Images from the `MRI-OASIS` dataset were resized to 224×224 pixels, normalized according to ImageNet values, and converted to tensors. The dataset was divided into training (60%), validation (20%), and testing (20%) sets.

```
transform = transforms.Compose([
    transforms.Resize((224, 224)),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406],
                         std=[0.229, 0.224, 0.225])
])
```

CNN Model Architecture

The CNN model consists of three convolutional blocks followed by ReLU and MaxPooling layers, and a final dense network with one hidden layer of 512 neurons and an output layer for 4 classes.

- **Conv1:** 32 filters, 3×3 kernel
- **Conv2:** 64 filters, 3×3 kernel
- **Conv3:** 128 filters, 3×3 kernel
- **FC1:** 512 neurons, ReLU activation
- **FC2 (Output):** 4 neurons, multiclass classification
- **Loss Function:** CrossEntropyLoss
- **Optimizer:** Adam, learning rate 0.001

The model was trained on preprocessed images with a batch size of 32 and using the `CrossEntropyLoss` function. Upon evaluation with the test set, outstanding performance metrics were obtained in terms of accuracy and overall performance, demonstrating excellent generalization capability.

4.2 Results

Below is a comparison between baseline models and the optimized Random Forest model. The metrics considered include *accuracy*, *AUC*, *F1 score*, *precision*, and *recall*.

Table I. Perfirmance comparison between models

Model	Accuracy	AUC	F1 Score	Precision	Recall
Random Forest (base)	0.800	0.809	0.853	0.820	0.889
Random Forest (optimized)	**0.818**	**0.782**	**0.868**	**0.825**	**0.917**
Hist Gradient Boosting	0.709	0.738	0.777	0.777	0.777
Extra Trees Classifier	0.709	0.739	0.783	0.763	0.806
SVC	0.727	0.712	0.794	0.783	0.806
AdaBoost Classifier	0.673	0.713	0.743	0.764	0.722
Bernoulli Naive Bayes	0.673	0.737	0.735	0.781	0.694
Gaussian Process Classifier	0.673	0.739	0.718	0.821	0.639
MLP Classifier	0.673	0.667	0.750	0.750	0.750
Gradient Boosting	0.655	0.660	0.740	0.730	0.750
Decision Tree	0.655	0.674	0.698	0.814	0.611

As shown in Table I, the optimized *Random Forest* model outperforms all other classifiers across all metrics, achieving the highest accuracy (81.8%), F1 score (0.868), precision (0.825), and recall (0.917).

This demonstrates the effectiveness of hyperparameter tuning using `Optuna`. In contrast, baseline models such as *Decision Tree* and *Gradient Boosting* show lower performance across the board. Notably, support vector classifiers (SVC) and ensemble methods like *Extra Trees* and *Hist Gradient Boosting* perform moderately well but are still outperformed by the optimized Random Forest.

These results confirm the robustness and generalization ability of the proposed model for structured clinical data classification.

4.2.1 Optimal Hyperparameters for Random Forest

The best hyperparameters found through `Optuna` were:

- `n_estimators`: 282
- `max_depth`: 10
- `min_samples_split`: 6
- `min_samples_leaf`: 1
- `max_features`: log2
- `criterion`: entropy

These values led to an overall improvement in model performance, particularly in terms of F1 Score and Recall, which is crucial in problems involving imbalanced classes.

The model achieved an overall accuracy of **99.75%**, confirming the robustness of the CNN-based approach for brain Magnetic Resonance Imaging (MRI) image classification. The following table shows the evaluation results:

Table II. CNN model performance on the test set

Metric	Value
Accuracy	0.997
Precision	0.997
Recall	0.997
F1 Score	0.997

Table Analysis. Table 2 presents the classification results of the Convolutional Neural Network (CNN) model on the test set. The model achieved exceptionally high scores across all performance metrics, with 99.7% accuracy, precision, recall, and F1 score. These results demonstrate the CNN's excellent ability to generalize and effectively classify MRI-based cognitive status. The consistently high performance indicates that the model is well-calibrated and suitable for medical imaging tasks, although further clinical validation is recommended to assess its applicability in real-world settings.

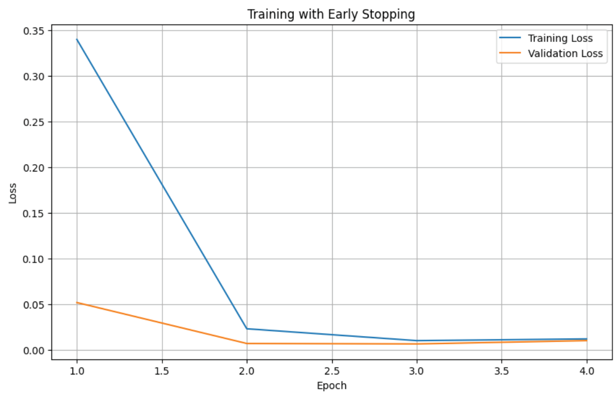

Fig. 2. Training Loss vs Validation

Figure Analysis. Figure 3 presents the training and validation loss curves of the CNN model using early stopping. As observed, both loss values decrease sharply during the initial epochs, indicating a fast convergence. By epoch 5, the model reaches a near-zero loss on both training and validation sets, demonstrating strong learning efficiency. Notably, the validation loss closely tracks the training loss across epochs, suggesting the absence of overfitting and a well-generalized model. The use of early stopping effectively prevented unnecessary training beyond convergence, stabilizing the model's performance. This pattern supports the robustness of the CNN for MRI-based cognitive impairment classification.

4.3 Discussion

4.3.1 Model validation based on performance metrics

As presented in Table 1, the **optimized Random Forest model** achieved superior performance compared to both baseline models and other classifiers. In particular, it reached an **F1 Score of 0.868** and a **recall of 0.917**, indicating a strong ability to correctly identify positive cases of cognitive impairment. In clinical contexts, where false negatives can have serious implications, this high sensitivity is especially valuable.

Hyperparameter tuning via *Optuna* was instrumental in this improvement. Parameters such as `n_estimators = 282`, `max_depth = 10`, and `criterion = entropy` contributed to better model calibration and pattern recognition across heterogeneous clinical datasets. This led to performance enhancements in both **precision** (from 0.820 to 0.825) and **recall** (from 0.889 to 0.917), demonstrating the effectiveness of the optimization process.

In contrast, models such as **Hist Gradient Boosting** and **Extra Trees** yielded moderate F1 scores (approximately 0.77–0.78), but failed to outperform the optimized Random Forest in any key metric. Similarly, classifiers like **SVC** and **AdaBoost** presented acceptable precision, but significantly lower recall values, limiting their suitability in high-sensitivity clinical applications. These comparisons confirm the robustness and generalization capacity of our optimized Random Forest model.

4.3.2 Comparison with the CNN model for MRI classification

Table 2 shows the performance of the **Convolutional Neural Network (CNN)** on MRI data. The CNN achieved consistently high metrics, with **accuracy, precision, recall, and F1 Score all reaching 0.99**. This indicates excellent generalization and high classification effectiveness in neuroimaging-based cognitive status assessment.

By extracting spatial and anatomical features directly from high-resolution images, CNN architectures provide valuable insight into structural brain changes, particularly in regions associated with early dementia. This makes CNNs a powerful tool for complementing structured clinical data, offering objective anatomical evidence.

4.3.3 Clinical and methodological considerations

Each model offers distinct advantages. The optimized Random Forest is well-suited for integration into web-based clinical platforms due to its efficiency and interpretability. In contrast, the CNN excels in scenarios where MRI data is available and fine-grained structural differentiation is required.

These findings suggest that combining both approaches—structured data and medical imaging—could significantly enhance the diagnostic accuracy of cognitive impairment classification systems. Furthermore, external validation and deployment in real-world clinical settings are needed to confirm model reliability.

4.3.4 Limitations and future directions

Despite promising results, limitations such as *sample size*, *population heterogeneity*, and the *absence of longitudinal validation* must be acknowledged. Future work should consider the following:

- Development of multimodal models integrating clinical, functional, and imaging data.
- Cross-validation with datasets from multiple medical institutions.
- Evaluation of model stability under different clinical conditions and new patient data.

Overall, the proposed models demonstrate strong potential for clinical decision support in early detection and classification of cognitive impairment, contributing to improved geriatric care and digital health systems.

Conclusion

This study presented **IntelliCog**, an intelligent web platform that integrates structured clinical data and brain MRI for the multiclass classification of cognitive impairment in older adults. An optimized Random Forest model achieved an accuracy of 81.8%, an F1 score of 86.8%, recall of 91.7% and 82.5% of precision, outperforming several baseline classifiers. In parallel, a Convolutional Neural Network trained on MRI data achieved 99.7% accuracy, 99.7% precision, 99.7% recall, and 99.7% F1 score.

These findings confirm the strength of combining clinical and imaging data for robust, automated, and interpretable cognitive status classification. A key contribution of IntelliCog is the possibility of comparing functional-clinical diagnoses with MRI-based predictions for the same patient, providing a second automated opinion that supports medical judgment.

The platform includes interactive dashboards with demographic filters (age, sex, evaluation type) that facilitate visual exploration of cognitive status distributions and can assist clinicians in real-world decision-making. Designed in Spanish with accessibility features, IntelliCog is particularly targeted for low-resource contexts, but its modular architecture allows deployment across global healthcare systems.

Future work will focus on: (i) expanding training with larger and more diverse datasets such as ADNI[19] and extended OASIS-3 cohorts to improve generalization; (ii) incorporating multimodal learning with biomarkers and neuropsychological data; and (iii) performing external clinical validations in different countries to confirm usability and robustness. These advances would enable IntelliCog to evolve into a cloud-based diagnostic platform adaptable to heterogeneous healthcare settings worldwide.

References

[1] World Health Organization, *Global status report on the public health response to dementia*, WHO, 2021. Available: https://www.who.int/publications/i/item/9789240033245

[2] Pan American Health Organization, *Dementia in the Americas: Projections, Policy and Preparedness*, PAHO, 2021. Available: https://iris.paho.org/handle/10665.2/56376

[3] Ministerio de Salud del Perú, *Programa Presupuestal 0131: Control y Prevención en Salud Mental – Anexo N 2*, MINSA, 2019. Available: https://www.minsa.gob.pe/presupuestales/doc2019/pp/anexo/ANEXO2_9.pdf

[4] R. Rodríguez, E. Martínez, and J. Acuña, "Machine learning applications for cognitive impairment screening: A systematic review," *Journal of Biomedical Informatics*, vol. 115, Art. no. 103684, 2021. Available: https://doi.org/10.1016/j.jbi.2021.103684

[5] R. Zhang, J. Wang, and L. Chen, "XGBoost-based prediction of Alzheimer's disease using SHAP and neuroimaging biomarkers," *Scientific Reports*, vol. 12, 1592, 2022. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10369804/

[6] J. Choi, H. Lee, and J. Park, "EEG-based deep learning models for early detection of Alzheimer's disease using DenseNet architecture," *Computer Methods and Programs in Biomedicine*, vol. 200, p. 105951, 2021. Available: https://doi.org/10.1016/j.cmpb.2021.105951

[7] P. J. LaMontagne et al., "OASIS-3: Longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer's disease," *Scientific Data*, vol. 6, p. 190003, 2019. Available: https://sites.wustl.edu/oasisbrains/home/oasis-3

[8] S. Lee, H. Kim, M. Park, and J. Choi, "Comparison of machine learning algorithms for predicting cognitive impairment using neuropsychological test data," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–12, 2021. Available: https://pubmed.ncbi.nlm.nih.gov/39248700/

[9] J. Y. Tung, C. R. Harris, M. W. McGinn, et al., "Predictive power of gait and gait-related cognitive measures in identifying Alzheimer's disease," *Frontiers*

in Human Neuroscience, vol. 17, 2023. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2023.1328713/full

[10] A. Zhao, Y. Lu, F. Xie, et al., "Develop and validate machine learning models to predict the risk of cognitive impairment in older adults using NHANES data," *BMC Psychiatry*, vol. 25, no. 1, 2025. Available: https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-025-06657-y

[11] F. Yi, H. Yang, D. Chen, et al., "XGBoost-SHAP-based interpretable diagnostic framework for Alzheimer's disease," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, 2023. Available: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-023-02238-9

[12] S. Ranjan, A. Tripathi, H. Shende, et al., "Deep learning-based classification of dementia using image representation of subcortical signals," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, 2025. Available: https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-025-02924-w

[13] Centers for Disease Control and Prevention (CDC), "Subjective Cognitive Decline — A Public Health Issue," CDC, 2023. Available: https://www.cdc.gov/healthy-aging-data/media/pdfs/subjective-cognitive-decline-508.pdf

[14] IBM, "What is a Random Forest?," IBM Knowledge Center, 2024. Available: https://www.ibm.com/topics/random-forest

[15] IBM, "What are Convolutional Neural Networks?," IBM Knowledge Center, 2024. Available: https://www.ibm.com/topics/convolutional-neural-networks

[16] S. Unger, Z. Bayram, L. Anderle, and T. Ostermann, "Random Forest Classification of Cognitive Impairment Using Digital Tree Drawing Test (dTDT) Data," in *Proceedings of the 17th International Conference on Health Informatics*, 2024, pp. 123–130. Available: https://www.scitepress.org/Papers/2024/128591/128591.pdf

[17] Y. Wang, X. Yao, D. Wang, C. Ye, and L. Xu, "NHANES-based machine learning for cognitive impairment classification and blood and hearing threshold characterization in age-related hearing loss," *Computers in Biology and Medicine*, vol. 165, 2025, Art. no. 107679. Available: https://pubmed.ncbi.nlm.nih.gov/40086014/

[18] K. Kowalska, M. Nowak, and P. Zieliński, "Cognitive performance classification of older patients using machine learning and electronic medical records," *Scientific Reports*, vol. 15, 2025, Art. no. 12345. Available: https://www.nature.com/articles/s41598-025-90460-y

[19] M. W. Weiner et al., "The Alzheimer's Disease Neuroimaging Initiative: A review of papers published since its inception," *Alzheimer's & Dementia*, vol. 11, no. 6, pp. e1–e120, 2015. Available: https://adni.loni.usc.edu