

Hyperventilation-induced Cerebral Patterns Analysis for Neurocognitive Disorder Detection via AI Models

Kusum Tara*

Department of Biological and Material Engineering, Graduate School of Science and Engineering, Saga University, Japan

Ruimin Wang

Department of Electrical and Electronic Engineering, Faculty of Science and Engineering, Saga University, Japan

Yoshitaka Matsuda

Institute of Ocean Energy, Saga University, Japan

Satoru Goto

Department of Electrical and Electronic Engineering, Faculty of Science and Engineering, Saga University, Japan

Takako Mitsudo

Minkodo-Minohara Hospital, Fukuoka, Japan. NHO Hizen Psychiatric Medical Center, Saga, Japan

Takao Yamasaki

Minkodo-Minohara Hospital, Fukuoka, Japan. International University of Health and Welfare, Fukuoka, Japan

Takenao Sugi

Department of Electrical and Electronic Engineering, Faculty of Science and Engineering, Saga University, Japan

Abstract

Monitoring hyperventilation (HV)-induced cerebral patterns offers a promising approach for early detection of mild cognitive impairment (MCI) and Alzheimer's disease (AD), which are associated with disrupted cerebral blood flow, neuronal excitability, and brain connectivity. This study analyzed EEG signals collected under HV-induced physiological stress using deep learning convolutional neural network (CNN)-based AI models to observe brain activity changes. Phase-amplitude coupling (PAC) and spectral topographic mapping (STM) images were generated from EEG and applied to two CNN architectures, MobileNetV2 and Xception. Functional connectivity was further examined using Pearson correlation and Granger causality to identify neural alterations. Among these AI models, Xception combined with PAC images achieved the highest classification accuracy of 98.95 %, outperforming MobileNetV2 (96.45 %) by effectively capturing non-linear brain dynamics, impaired phase synchronization, and disrupted neural communication in MCI and AD patients. Gradient-weighted class activation mapping (Grad-CAM) was used to rank EEG channels based on their importance, revealing that parietal and occipital channels contributed most to model decisions. These results demonstrate that the proposed PAC-based Xception model provides a reliable method for identifying neurocognitive dysfunction and neural desynchronization of MCI and AD patients with reduced delta-alpha and theta-alpha couplings.

Contribution of the Paper: This study incorporates Granger causality for analyzing HV-induced functional connectivity changes between EEG channels and phase-amplitude coupling with AI models to detect impaired phase synchronization and disrupted neural communication, enhancing reliability in classifying MCI and AD.

Keywords: neurocognitive disorders, hyperventilation, cerebral patterns, granger causality, phase-amplitude coupling.

IJCVPSP

ISSN: 2186-1390 (Online)
<http://cennser.org/IJCVPSP>

Article History:
 Received: 16/3/2025
 Revised: 14/7/2025
 Accepted: 1/11/2025
 Published Online: 23/11/2025

1. INTRODUCTION

Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) are neurocognitive disorders characterized by progressive cognitive decline, affecting memory, and daily functioning [1]. MCI is an intermediate stage between normal aging and dementia, where individuals experience noticeable cognitive difficulties but can still perform daily activities, whereas AD is the most common cause of dementia, leading to severe memory loss and impaired cognitive function. With an aging global population, concerns about MCI and AD are rising, as these conditions significantly impact healthcare systems and quality of life. According to the World Health Organization (WHO), over 55 million people worldwide suffer from dementia, with AD accounting for 60–70 % of cases, and nearly 15–20 % of people aged 65 or older have MCI [2]. In hospitals, MCI and AD are commonly diagnosed through neuropsychological assessments, magnetic resonance imaging (MRI), positron emission tomography (PET) scans, and cerebrospinal fluid (CSF) analysis, which help to detect structural and functional brain abnormalities. However, EEG is preferable due to its non-invasive nature, and high sensitivity to functional brain changes [3]. EEG can capture early neural dysfunctions, including altered power spectral patterns and disrupted functional connectivity, making it a valuable tool for early-stage detection and AI-driven classification of neurocognitive disorders.

Recent studies have explored neurocognitive disorder detection using eyes-closed (EC), eyes-open (EO), and hyperventilation (HV) induced EEG signals, particularly in MCI and AD. AI-driven deep-learning models have gained prominence for their ability to extract meaningful features and classify cognitive impairments effectively. In 2024, Şeker employed an EEGNet-based CNN model, achieving 96.00 % accuracy in MCI vs. healthy controls (HC) classification under eyes-closed conditions but lacked HV analysis [4]. Similarly, Watanabe in 2024 applied a deep CNN (MNet) for dementia classification, reaching 92.70 % accuracy, yet struggled with interpreting neural activity patterns [5]. Earlier studies by Miraglia in 2023 examined HV-induced EEG alterations using graph theory but did not incorporate deep learning, while Coppola in 2010 analyzed visual-evoked potentials under HV in migraine patients without applying AI models [6-7]. A research gap exists in leveraging HV-induced EEG signals with deep-learning models for enhanced neurocognitive disorder detection. HV-induced cerebral pattern analysis is crucial as HV alters cerebral blood flow, modulates neuronal excitability, and affects brain connectivity, providing unique

biomarkers for MCI and AD. This study addresses this gap by integrating HV-induced EEG with AI models, particularly phase-amplitude coupling (PAC) images in architectures like Xception, to enhance interpretability, and early detection of cognitive decline in MCI and AD patients.

In this study, CNN-based deep learning models with MobileNetV2 and Xception architectures were employed to classify MCI and AD using EEG signals. Spectral topographic mapping (STM) and PAC images, derived from delta, theta, alpha, and beta bands, were used to capture spatial power distributions and cross-frequency interactions between hemispheres. EEG data from Minkodo-Minohara Hospital EEG Database (EEG-MHDB) were collected into pre-HV (resting), during-HV, and post-HV (resting) conditions, enabling visualization of region-specific brain dynamics across frontal, temporal, parietal, and occipital areas. While STM revealed distinct neural activity patterns, PAC showed interaction between low-frequency phase (delta, theta) and high-frequency amplitude (alpha, beta) oscillations crucial for cognitive processes. PAC-based Xception model captured disrupted phase synchronization and non-linear neural connectivity patterns in MCI and AD. Grad-CAM further enhanced model interpretability by ranking EEG channels based on their contribution to predictions, highlighting parietal and occipital channels as most influential. An ablation study using these top-ranked channels confirmed classification accuracy. This integrated approach offers a reliable framework for early detection of neurocognitive disorders

2. MATERIALS AND METHODS

The block diagram of Figure 1 outlines the neurocognitive disorder detection process through HV-induced cerebral patterns in MCI and AD patients through a structured pipeline. The framework begins with selecting MCI and AD subject groups, followed by EEG recording and preprocessing to ensure signal quality. Connectivity between channels is assessed using Pearson correlation and Granger causality to identify key EEG channels. Signals are segmented into delta, theta, alpha, and beta bands, then transformed into STM and PAC images to capture power distribution and neural interactions. These images are input into CNN models using MobileNetV2 and Xception architectures for classification. Model performance is evaluated using precision, recall, F-score, and accuracy, ensuring reliable MCI and AD detection. Grad-CAM is used to enhance interpretability of the selected model by identifying the most influential channels, supporting both accurate and explainable AI-driven diagnosis.

2.1. Subjects

This study analyzed EEG data from participants MCI and AD from the Minkodo-Minohara Hospital EEG database (EEG-MHDB) in Japan. A total of 34 subjects (22 AD, 12

*Corresponding author

Email addresses: 24808202@edu.cc.saga-u.ac.jp (Kusum Tara), wangrm@cc.saga-u.ac.jp (Ruimin Wang), ymatsuda@cc.saga-u.ac.jp (Yoshitaka Matsuda), goto@cc.saga-u.ac.jp (Satoru Goto), mitsudo.takako.692@m.kyushu-u.ac.jp (Takako Mitsudo), takaoya0317@gmail.com (Takao Yamasaki), sugi@cc.saga-u.ac.jp (Takenao Sugi)

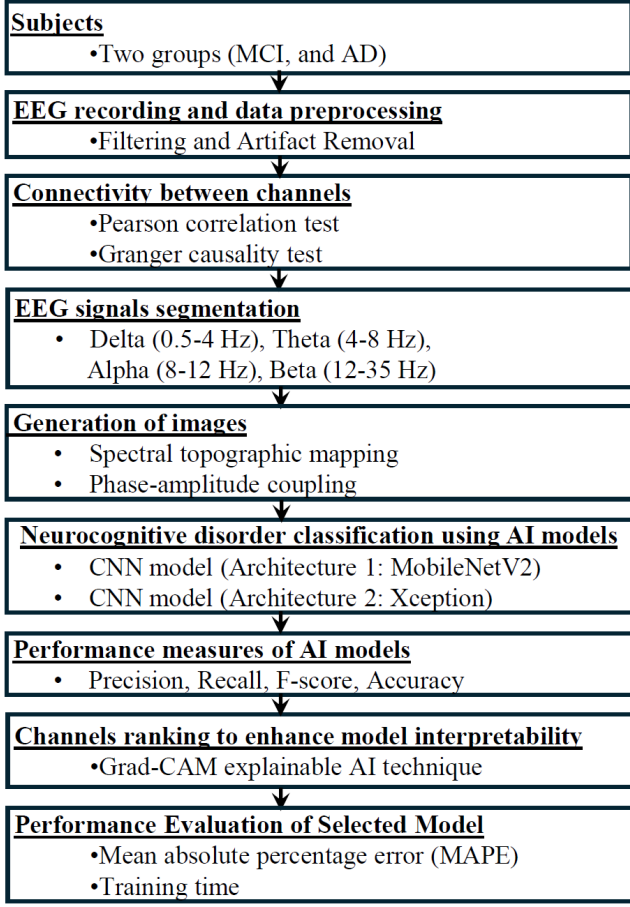


Figure 1: Neurocognitive disorder detection process through HV-induced cerebral patterns in MCI and AD patients.

MCI), aged 60 to 90 years, with normal vision and hearing, were included after providing written informed consent [8]. All underwent neuropsychological assessments, including the mini-mental state examination (MMSE) and clinical dementia rating (CDR) scale, alongside EEG monitoring. Clinical diagnosis by neurologists identified dementia based on cognitive symptoms interfering with daily functions, while MCI was determined using a cut-off score of 1.5 standard deviations below the age-adjusted norm in the Wechsler memory scale (WMS) logical memory test. Participants were classified based on MMSE scores, with MCI (MMSE 24–27, $N = 12$) and AD (MMSE 0–23, $N = 22$). Statistical analysis, including the Bonferroni correction test, revealed significant differences between the groups, showing a notable effect of age ($p < 0.001$), where MCI patients were significantly younger than AD patients (Bonferroni-corrected $p < 0.001$) [9]. Additionally, MMSE scores showed a significant decline from MCI to AD ($p < 0.001$), confirming a progressive cognitive deterioration trend (MCI > AD). These results emphasize the role of age and MMSE scores in distinguishing cognitive dysfunction

and HV-induced response analysis (AD > MCI).

2.2. EEG recording and data preprocessing

All subjects underwent routine clinical EEG recording at Minkodo-Minohara Hospital, Japan, following standard scalp preparation with gel and cotton tips [8]. EEG signals were recorded using 21 electrodes placed according to the international 10–20 system, covering frontal (Fp1, Fp2, F3, F4, F7, F8, Fz), temporal (T3, T4, T5, T6), central (C3, C4, Cz), parietal (P3, P4, Pz), occipital (O1, O2), and ipsilateral earlobes reference (A1, A2) regions, with a sampling rate of 500 Hz. Participants sat with eyes closed in a dimly lit room while EEG data were recorded under three conditions: resting-state (pre-HV), hyperventilation (during-HV), and resting-state (post-HV). The study focused on HV-induced cerebral activity, where subjects performed overbreathing for 2s, followed by a 2s rest before and after overbreathing, with each phase repeated in 2s intervals, separated by a 2s break, as depicted in Figure 2. For each participant, EEG signals were recorded over 30 trials, with each trial lasting 2s, allowing for a comprehensive analysis of HV-induced cerebral patterns in MCI and AD. Data preprocessing involved re-referencing all channels using reference ear electrodes (A1, A2) and filtering the signals to 0.5–35 Hz to remove noise and enhance signal clarity. This study analyzed HV-induced cerebral activity by selecting the most relevant channels from frontal, temporal, central, parietal, and occipital regions using Pearson correlation and Granger causality tests to uncover HV-related neural alterations that differentiate MCI and AD groups.

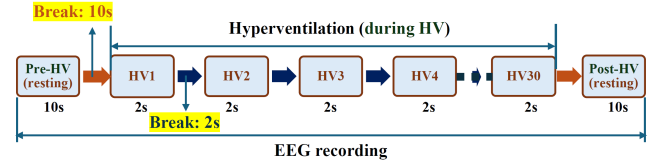


Figure 2: HV performed during clinical EEG examination. HV involved overbreathing for 2s with 2s breaks between HVs. Subjects kept their eyes closed throughout EEG recording for consistent neural response measurement.

2.3. Connectivity between channels

Pearson correlation and Granger causality tests were applied to 19 EEG channels across frontal (Fp1-A1, Fp2-A2, F3-A1, F4-A2, F7-A1, F8-A2, Fz-A1), temporal (T3-A1, T4-A2, T5-A1, T6-A2), central (C3-A1, C4-A2, Cz-A1), parietal (P3-A1, P4-A2, Pz-A1), and occipital (O1-A1, O2-A2) regions to analyze HV-induced functional connectivity and neural communication impairments in MCI and AD patients. These techniques offer complementary perspectives on brain interactions, effectively identifying neural patterns while minimizing redundancy. Each region plays a crucial role: parietal channels (P3-A1, P4-A2) facilitate sensory integration and cognitive processing,

temporal channels (T3-A1, T4-A2, T5-A1, T6-A2) are essential for memory encoding and auditory processing, and occipital channels (O1-A1, O2-A2) support visual processing and stimulus-driven responses. Analyzing connectivity changes across these regions provides critical insights into neural dysfunction, enhancing the detection of MCI and AD-related neurocognitive impairments.

2.3.1. Pearson correlation test

Pearson correlation test captures linear connectivity and synchrony across brain regions, including frontal-central, frontal-parietal, frontal-temporal, occipital-frontal, occipital-parietal, occipital-temporal, temporal-parietal, and intra-regional correlations [10]. It measures the strength of association between EEG channel pairs by analyzing signal similarity, providing valuable insights into functional connectivity and neural interactions across different brain regions. Pearson correlation coefficient, ($|r|$) is calculated and shown in Eq. (1):

$$|r| = \frac{\sum_{i=1}^N (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X_i - \bar{X})^2 \sum_{i=1}^N (Y_i - \bar{Y})^2}} \quad (1)$$

where X_i and Y_i are individual sample points, $\bar{X}$ and $\bar{Y}$ are the means of the datasets X and Y . The pairs of channels are considered highly correlated if the correlation coefficient, $|r|$ is close to 1 and uncorrelated if the correlation coefficient, $|r|$ is 0 or below 0.3. Pearson correlation coefficients ($|r|$) identify highly correlated EEG channels that exhibit consistent connectivity changes in MCI and AD groups, while eliminating weakly correlated or redundant channels. By analyzing variations in $|r|$, this approach highlights relevant neural connections between EEG channels, enhancing the detection of HV-induced cerebral patterns for MCI and AD classification.

2.3.2. Granger causality test

Granger Causality test reveals directional connectivity by assessing how one channel in brain region connects another channel, such as occipital influencing parietal regions [11]. This helps analyze neural communication dynamics, offering insights into information flow impairments in MCI and AD. If X_t and Y_t are two channels, to check if X_t Granger-causes Y_t , compute Granger causality F-statistic as shown in Eq. (2):

$$F - \text{statistic} = \frac{(\sigma_1^2 - \sigma_2^2)/p}{\sigma_2^2/(N - 2p)} \quad (2)$$

Where, σ_1^2 and σ_2^2 are variance of residuals from univariate autoregressive model without causal influence and bivariate autoregressive model with causal influence respectively, p is number of lagged observations, N is number of observations. The univariate autoregressive model without causal influence is represented in Eq. (3):

$$Y_t = \sum_{i=1}^p \alpha_i Y_{t-i} + \epsilon_t \quad (3)$$

Where, Y_t is a channel, α_i are the autoregressive coefficients, ϵ_t is the error term. The bivariate autoregressive model with causal influence is depicted in Eq. (4):

$$Y_t = \sum_{i=1}^p \alpha_i Y_{t-i} + \sum_{j=1}^p b_j X_{t-j} + \epsilon'_t \quad (4)$$

Where, X_t is another channel, b_j are the autoregressive coefficients for X_t , ϵ'_t is the new error term. If the F-statistic is significant, X_t Granger-causes Y_t , indicating directional connectivity between EEG channels; otherwise, no causal influence exists. In channel connectivity analysis, an F-statistic < 0.10 indicates weak or no causality, $0.10-0.20$ suggests significant functional connectivity, and > 0.20 implies strong Granger causality, meaning one channel strongly connects another. Higher values reflect neural interactions in cognition or disruptions in MCI and AD, while lower values indicate weak relationships between channels.

2.4. EEG signals segmentation

EEG signals were recorded across three conditions such as pre-HV (resting), during-HV, and post-HV (resting) from 19 channels of every patient to analyze HV-induced cerebral patterns in MCI and AD patients. For each participant, HV trials consisted of 30 trials (2s each), while pre-HV and post-HV trials lasted 10s each. To enhance analysis resolution, each HV trial was divided into two 1s segments, resulting in 60 segments per subject of 19 channels (during-HV: 30 trials $\times$ 2 segments $\times$ 19 channels). Similarly, pre-HV and post-HV trials were each divided into ten 1s segments, yielding 10 segments per subject per condition (pre-HV: 10 segments $\times$ 19 channels, post-HV: 10 segments $\times$ 19 channels). Each segment was divided into four frequency bands (delta: 0.5–4 Hz, theta: 4–8 Hz, alpha: 8–12 Hz, beta: 12–35 Hz) for further analysis. STM and PAC images were generated using a 1s time window from the selected EEG channels. Considering all 34 subjects and 19 EEG channels, a total of 51680 segments were obtained for MCI and AD groups. Each test set for MCI and AD groups contained 10336 segments. Then images were generated to ensure reliable CNN model performance in detecting neurocognitive disorders.

2.5. Generation of images

STM and PAC images were used to represent power distribution across selected brain regions during pre-HV, during-HV, and post-HV conditions. These images provide reliable measures of neural oscillatory responses, offering valuable insights into HV-induced brain dynamics and neurocognitive alterations.

2.5.1. STM

STM provides a visual representation of power distribution across four frequency bands over the scalp, highlighting spatial variations in brain activity during pre-HV, during-HV, and post-HV conditions [12]. Generated from segmented EEG data, STMs are derived by computing power spectral density (PSD) using the Welch method for each 1s segment, as formulated in Eq. (5).

$$P(f) = \frac{1}{N} \sum_{i=1}^N |X_i(f)|^2 \quad (5)$$

where $P(f)$ is the power at frequency f , and $X_i(f)$ is the Fourier transform of the signal at channel, i . These mappings capture asymmetrical power distributions between the left and right hemispheres across different frequency bands, offering insights into cognitive decline. They facilitate the characterization of brain oscillatory responses, providing a deeper understanding of neural dynamics and stimulus-driven cortical activity, ultimately aiding in the differentiation of neurocognitive disorders.

2.5.2. PAC

PAC quantifies the interaction between low-frequency phase (e.g., delta, theta) and high-frequency amplitude (e.g., alpha, beta) oscillations in the brain [13]. This coupling evaluates how low-frequency oscillations (delta and theta) influence high-frequency oscillations (alpha and beta), which are critical for assessing cognitive decline. The mathematical representation of PAC is given in Eq. (6):

$$C(f_p, f_a) = \frac{1}{N} \sum_{n=1}^N A(f_a, t_n) e^{j\phi(f_p, t_n)} \quad (6)$$

Where $C(f_p, f_a)$ represents the coupling strength between phase frequency f_p (delta, theta) and amplitude frequency f_a (alpha, beta), $A(f_a, t_n)$ is the amplitude envelope of the high-frequency signal, $\phi(f_p, t_n)$ is the instantaneous phase of the low-frequency signal at time t_n . The significance of PAC strength lies in its ability to detect disrupted neural communication across brain regions in MCI and AD. Strong PAC strength in the frontal and parietal regions supports executive functions and working memory, while PAC strength in the occipital region plays a crucial role in visual processing. A reduction in PAC strength, particularly in delta-alpha and theta-alpha interactions, leads to impaired phase synchronization, weakening interregional communication and neural coordination, which are essential for cognitive function.

2.6. Neurocognitive disorders classification using AI models

The classification of neurocognitive disorders was performed using CNN deep-learning models, specifically MobileNetV2 and Xception architectures, to analyze HV-induced cerebral patterns of MCI and AD patients utilizing EEG-based STM and PAC images [10, 14]. According to Table

1, MobileNetV2 consists of 53 total layers, while Xception has 71 layers, both using a batch size of 64 and the Adam optimizer with a learning rate of 0.0001. MobileNetV2 was trained for 90 epochs, whereas Xception was trained for 80 epochs. As shown in Figure 3(a), MobileNetV2 initially processes 224×224 input images, reducing them to 7×7 feature maps through a series of convolutional and pooling layers, culminating in 1280 neurons in the Fully Connected (FC) layer before classification into two groups (MCI and AD). In contrast, Xception depicted in Figure 3 (b) follows a deeper structure, refining 224×224 inputs into 7×7 feature maps, but with a higher depth and 2048 neurons in the FC layer before classification. These deep learning architectures provide classification for neurocognitive disorder detection.

Table 1: Hyperparameters for different architectures.

Hyperparameters	Deep-learning architectures	
	MobileNetV2	Xception
Total layers	53	71
Batch size	64	64
Epochs	90	80
Learning rate	0.0001	0.0001
Optimizer	Adam	Adam

A total of 51680 segments (MCI: 18240 segments, AD: 33440 segments) from 34 subjects were used, with each subject contributing 1520 spectral topographic mapping images and 1520 phase-amplitude coupling images. The dataset was divided into 70 % for training (36176 spectral topographic mapping images, 36176 phase-amplitude coupling images), 10 % for validation (5168 spectral topographic mapping images, 5168 phase-amplitude coupling images), and 20 % for testing (10336 spectral topographic mapping images, 10336 phase-amplitude coupling images), ensuring reliable model evaluation. The CNN models processed $224 \times 224 \times 3$ images using this dataset split. CNN model's performance was assessed using precision, recall, F-score, and accuracy, as defined in Eq. (7) to Eq. (10): precision measured positive prediction accuracy, recall (sensitivity) evaluated true positive detection, F-score balanced precision and recall, and accuracy reflected overall prediction correctness [10].

These measures confirmed the effectiveness and reliability of MobileNetV2 and Xception deep learning architectures in classifying neurocognitive disorders such as MCI and AD.

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (8)$$

$$\text{F-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% \quad (9)$$

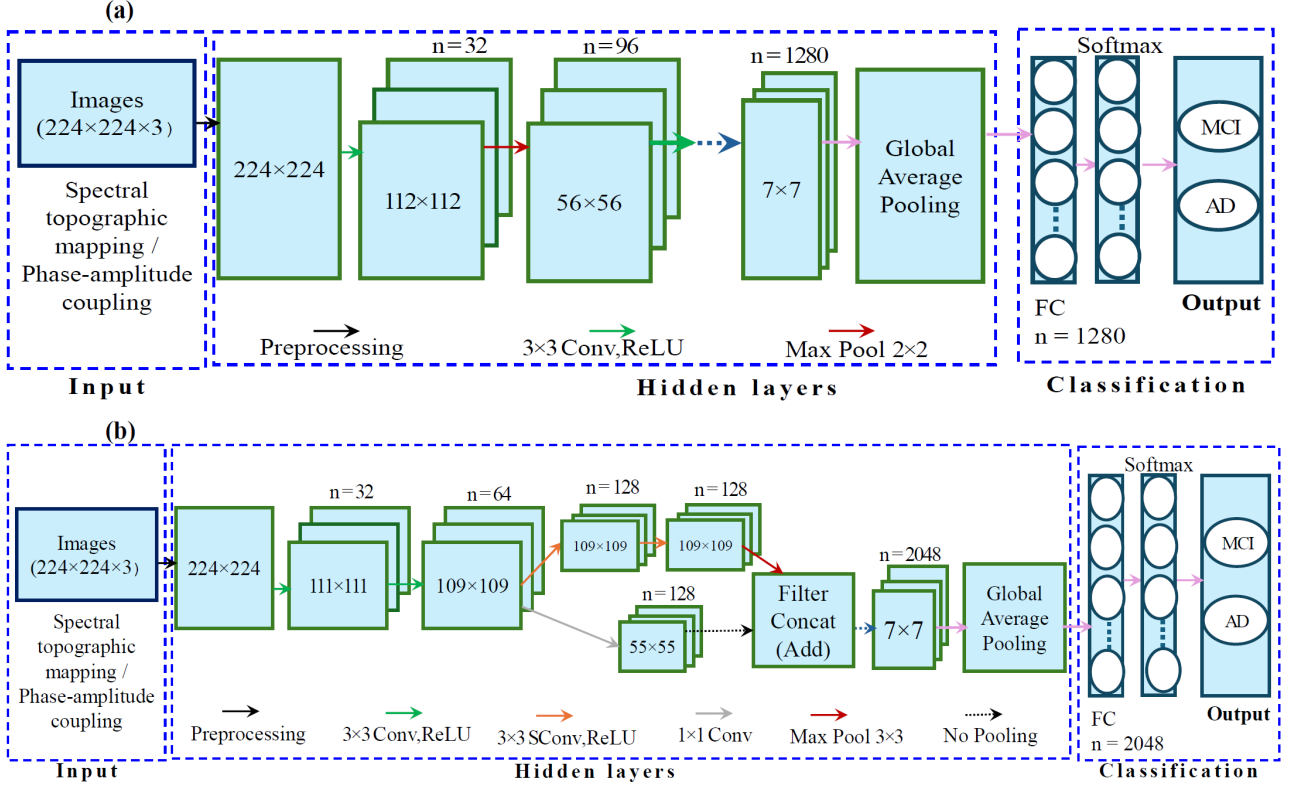


Figure 3: Deep learning architectures of CNN models; (a) MobileNetV2, (b) Xception.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (10)$$

Where, TP = True Positive, TN = True Negative, FP = False Positive, and FN = False Negative.

2.7. Channels ranking to enhance model interpretability

Grad-CAM is a neural network interpretability technique that identifies the most influential input regions contributing to a model's prediction by utilizing the gradients of target outputs with respect to convolutional feature maps [15]. In the context of EEG-based neurocognitive disorder detection, PAC images are generated for each EEG channel and passed through Xception based CNN model. For each EEG channel j , Grad-CAM computes feature attributions Ω_j by a weighted sum of feature maps from the last convolutional layer in Eq. (11) [15]:

$$\Omega_j = \text{ReLU}\left(\sum_{k=1}^K w_j^k A_j^k\right) \quad (11)$$

Here, A_j^k represents the k -th feature map for the j -th EEG channel and w_j^k is the weight capturing the importance of feature map A_j^k , calculated as Eq. (12) [15]:

$$w_j^k = \frac{1}{N} \sum_{i=1}^N \max(0, \frac{\partial y}{\partial A_{i,j}^k}) \quad (12)$$

Where N is the number of elements in feature map and $\partial y / \partial A_{i,j}^k$ is the gradient of the output with respect to each activation. This formulation ensures that only positively contributing features are considered, enhancing focus on discriminative patterns. By averaging the resulting Ω_j scores across samples, EEG channels are ranked based on their relative contribution to the final classification, allowing for explainable and data-driven selection of the most relevant channels for detecting neurocognitive disorders.

2.8. Performance evaluation of selected model

To evaluate the classification performance of the selected deep learning model (Xception) using PAC images derived from EEG channels, mean absolute percentage error (MAPE) is employed to assess how accurately the model predicts neurocognitive disorder outcomes relative to the true values [15]. It is calculated as Eq. (13):

$$\text{MAPE} = \frac{1}{T} \sum_{r=1}^T \left| \frac{y_r - \hat{y}_r}{y_r} \right| \times 100\% \quad (13)$$

where T is the total number of predictions, y_r is the actual value for the r -th instance, and $\hat{y}_r$ is the predicted output from the model. In this study, feature attributions for each EEG channel are obtained using the Grad-CAM technique, which highlights spatial-frequency regions that contribute most to the classification. Based on these

scores, channels are ranked, and a sequential feature ablation study is conducted. The model is retrained with progressively reduced EEG channels, starting from all channels and gradually removing those with the lowest Grad-CAM attribution. The resulting MAPE values from each retraining phase are compared to quantify the significance of each channel. A lower MAPE when using top-ranked channels confirms their relevance in capturing discriminative neural patterns for accurate detection of neurocognitive disorders.

3. RESULTS AND ANALYSIS

3.1. Analysis of channels connectivity

Figure 4 illustrates connectivity between EEG channels using Pearson correlation and Granger causality tests. In the Pearson correlation test of Figure 4(a), channels with significant correlation (0.30–1.00) are represented in green, yellow, and red, indicating moderate to strong connectivity. Notably, occipital-parietal correlations (O1-A1 to P3-A1, O2-A2 to P4-A2), seen in yellow to red regions, reflect sensory-visual integration and HV-induced excitability, crucial for classifying MCI and AD. Similarly, occipital-frontal (O1-A1 to Fp1-A1, O2-A2 to Fp2-A2, O1-A1 to F3-A1, O2-A2 to F4-A2, O1-A1 to F7-A1, O2-A2 to F8-A2) connections, marked in green to blue, highlight top-down modulation of visual attention, while occipital-temporal (O1-A1 to T3-A1, O2-A2 to T4-A2, O1-A1 to T5-A1, O2-A2 to T6-A2) interactions, appearing in yellow to red, play a role in visual memory. The Granger causality test shown in Figure 4(b) indicates significant directional influence (0.20–0.50), with green, yellow, and red regions representing stronger causality. Enhanced occipital-parietal and occipital-frontal causality, particularly in yellow to red, suggests functional integration, while weaker intra-regional causality in occipital lobe (blue regions) indicates neurodegeneration, reinforcing occipital connectivity for MCI and AD classification.

3.2. Analysis of generated images

Figure 5 presents STM images of MCI and AD groups in the alpha (8–12 Hz) frequency band across pre-HV, during-HV, and post-HV conditions. The color distribution indicates power intensity, with blue (low power, 0–50 $\mu V^2/Hz$), green-yellow (moderate power, 50–150 $\mu V^2/Hz$), and red (high power, 150–200 $\mu V^2/Hz$). In pre-HV (Figure 5 (a), Figure 5 (d)), both groups exhibit higher occipital and parietal alpha power, particularly in the MCI group, where posterior activity remains more symmetrical. In contrast, AD patients show greater occipital-temporal asymmetry, with lower power in hemisphere. In during-HV (Figure 5 (b), Figure 5 (e)), a power reduction is observed across both groups, but AD patients experience a more pronounced drop and slow wave activity, particularly in occipital and parietal regions, indicating weaker cortical responsiveness. In the post-HV condition (Figure 5

(c), Figure 5 (f)), MCI patients show partial alpha recovery, particularly in the parietal-occipital regions, while AD patients exhibit persistent alpha suppression, reflecting reduced neural resilience. The asymmetry between hemispheres in AD patients remains evident across all conditions, whereas MCI individuals display a more balanced recovery, supporting the hypothesis of greater neurodegenerative impact in AD.

PAC images for MCI and AD groups in occipital regions across pre-HV, during-HV, and post-HV conditions, showing delta-alpha, theta-alpha, delta-beta, and theta-beta couplings are illustrated in Figure 6. The color scale represents PAC strength, with blue (low coupling, 0.00–0.05), yellow (moderate coupling, 0.05–0.12), and red (high coupling, 0.12–0.20). In pre-HV (Figure 6(a), Figure 6(d)), MCI patients show stronger delta-alpha and theta-alpha coupling (yellow-red regions) in the low-frequency range (0.5–4 Hz phase, 8–12 Hz amplitude), while AD patients exhibit weaker PAC strength, particularly in delta-beta and theta-beta couplings, suggesting reduced neural coordination. In during-HV (Figure 6(b), Figure 6(e)), MCI PAC strength increases, particularly in theta-alpha and delta-beta couplings (red regions), indicating adaptive neural responses, whereas AD patients show further PAC reduction, especially in higher frequencies (theta-beta coupling remains weak in blue regions), suggesting impaired cortical adaptability. In post-HV (Figure 6(c), Figure 6(f)), MCI patients retain moderate PAC recovery (yellow regions in theta-alpha and delta-beta coupling), while AD patients show persistently reduced PAC strength, particularly in delta-alpha and theta-alpha couplings, highlighting cognitive dysfunction and neural desynchronization. The distinct PAC variations indicate weaker neural connectivity in AD, reinforcing PAC as a biomarker for distinguishing MCI from AD.

3.3. Classification results

Table 2 presents performance measures such as precision, recall, F-score, and overall accuracy for classifying neurocognitive disorders (MCI vs AD) using CNN deep learning models, specifically MobileNetV2 and Xception architectures, applied to STM and PAC images.

Table 2: Performance measures to classify NCDs (MCI vs. AD).

D. L. arch.	Input imgs.	Grps.	Prc. (%)	Rec. (%)	F-sc. (%)	Acc. (%)
MobileNetV2	STM	MCI	89.39	92.19	90.77	92.26
		AD	95.67	94.03	94.84	
	PAC	MCI	93.67	94.87	94.41	94.65
		AD	97.26	97.19	97.22	
Xception	STM	MCI	91.25	92.53	91.39	94.53
		AD	95.79	94.77	95.28	
	PAC	MCI	97.51	98.65	98.08	98.95
		AD	99.19	98.97	99.07	

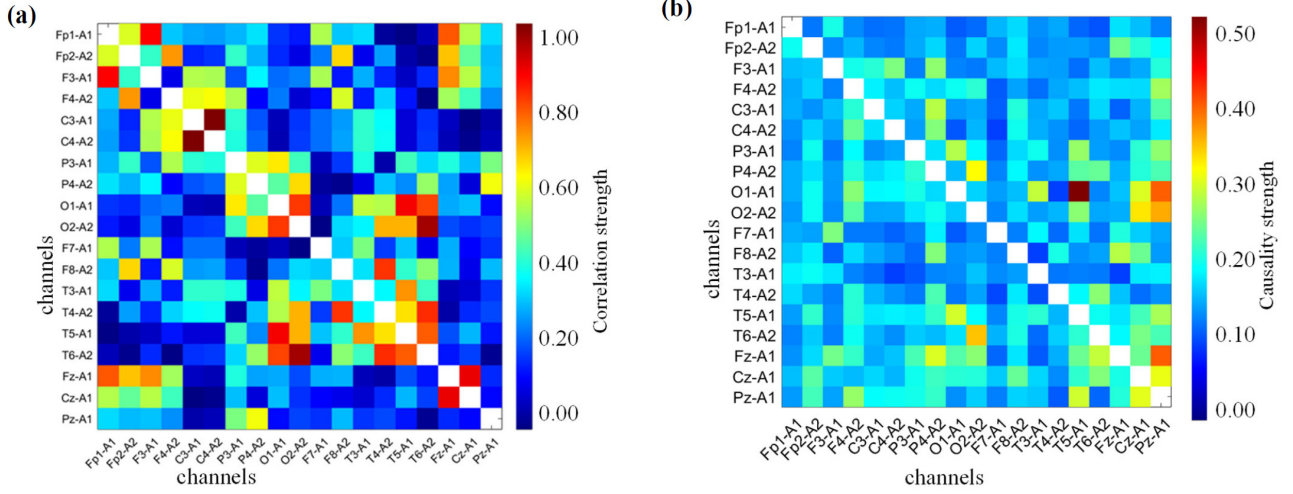


Figure 4: Connectivity between EEG channels; (a) Pearson correlation test, (b) Granger causality test.

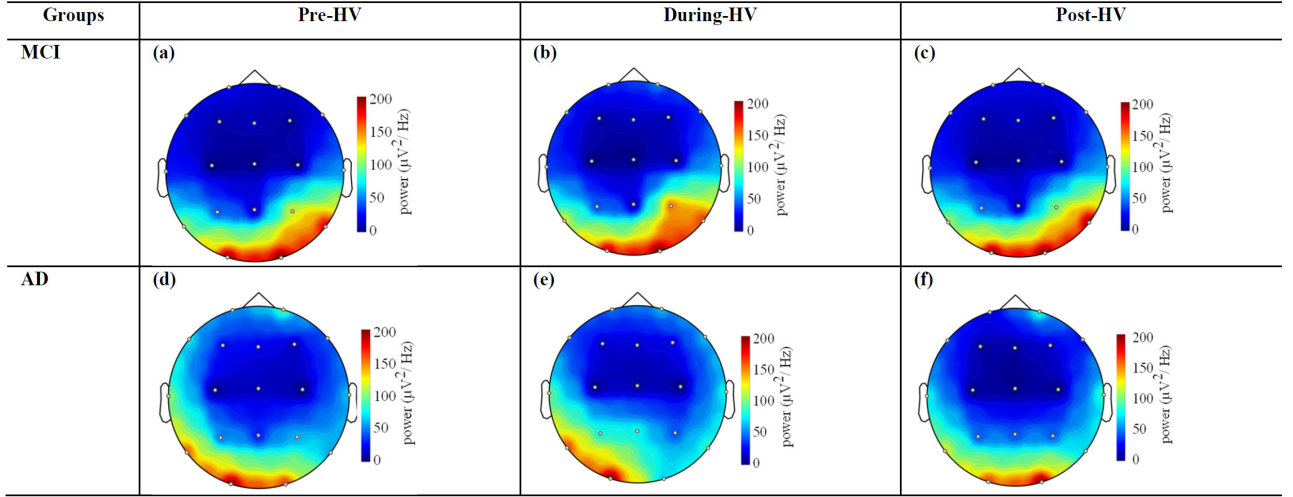


Figure 5: Spectral topographic mapping images for MCI and AD groups in alpha (8-12 Hz) frequency bands; (a) pre-HV, (b) during-HV, (c) post-HV.

For MobileNetV2, using STM, the model achieved precision of 89.39 % (MCI) and 95.67 % (AD), with recall values of 92.19 % (MCI) and 94.03 % (AD), resulting in F-scores of 90.77 % (MCI) and 94.84 % (AD) and an overall accuracy of 92.26 %. When using PAC images, precision improved to 93.97 % (MCI) and 97.26 % (AD), with recall of 94.87 % (MCI) and 97.19 % (AD), leading to F-scores of 94.41 % (MCI) and 97.22 % (AD), and a higher accuracy of 96.45 %. The Xception model outperformed MobileNetV2, achieving precision of 91.25 % (MCI) and 95.79 % (AD) with recall values of 92.53 % (MCI) and 94.77 % (AD) when using STM, resulting in F-scores of 91.39 % (MCI) and 95.28 % (AD) and an overall accuracy of 94.53 %. The highest performance was observed using Xception with PAC, where precision reached 97.51 % (MCI) and 99.19 % (AD), recall values were 98.65 % (MCI) and 98.97

% (AD), F-scores were 98.08 % (MCI) and 99.07 % (AD), with a peak accuracy of 98.95 %. These results indicate that Xception with PAC images provides the most reliable classification of MCI and AD, highlighting its effectiveness in distinguishing neurocognitive disorders.

To address real-time and clinical deployment feasibility for neurocognitive disorder detection, this study evaluated MobileNetV2 and Xception models based on inference time, model size, and suitability for edge devices. MobileNetV2, with a model size of ~ 20 MB and an inference time of ~ 25 ms per image on a standard CPU, achieved 96.45 % accuracy using PAC images, making it well-suited for low-resource, real-time applications such as wearable EEG headsets. Xception, though more computationally intensive with a model size of ~ 35 MB and ~ 30 ms inference time, delivered a higher accuracy of 98.95 %, ef-

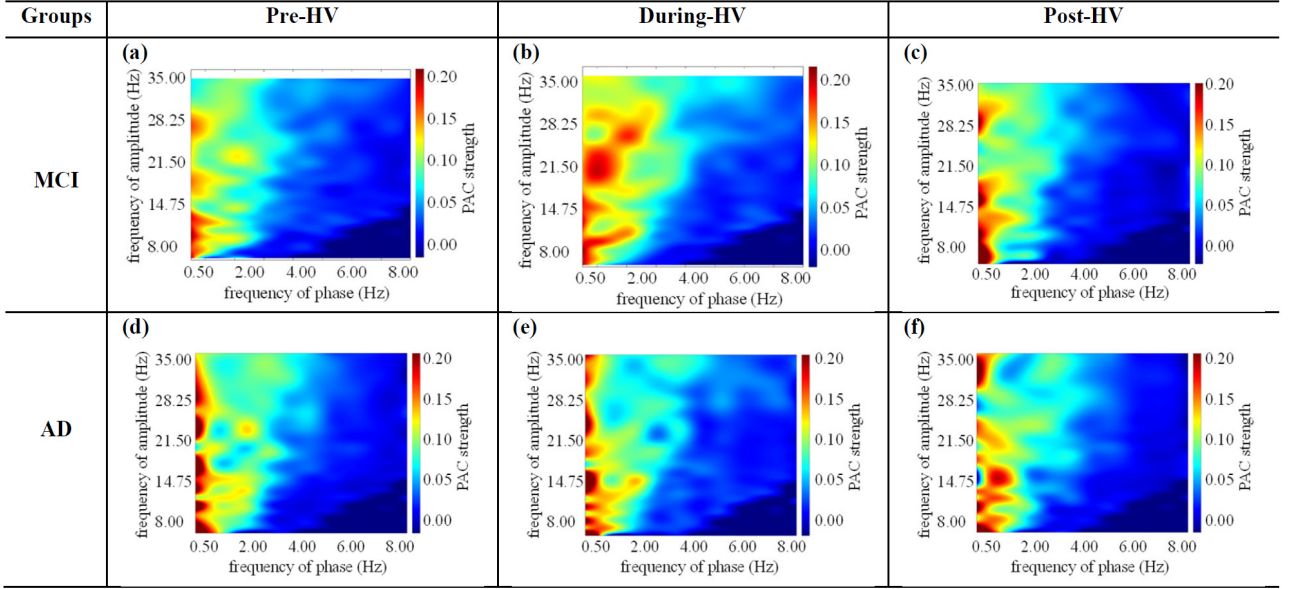


Figure 6: Phase-amplitude coupling images for MCI and AD groups in occipital regions; (a) pre-HV, (b) during-HV, (c) post-HV

fectively capturing complex PAC features crucial for early MCI and AD detection. While MobileNetV2 is ideal for edge deployment, Xception is preferred for hospital-based systems where diagnostic precision is paramount.

3.4. Analysis of performance evaluation of selected model

Figure 7 illustrates the ranking of EEG channels based on their attribution scores derived from the Grad-CAM technique, highlighting their relative importance in detecting neurocognitive disorders. The occipital channels (O1-A1, O2-A2) and parietal channels (Pz-A1, P3-A1, P4-A2) show the highest attribution scores, indicating their dominant influence in the model's decision-making process. These regions reflect early cognitive impairment and disrupted brain oscillations, supporting their clinical relevance. In contrast, frontal and temporal channels such as F8-A2 and T4-A2 contribute less to classification. This ranking enables selection of the most informative EEG channels, improving interpretability and guiding efficient model design.

Figure 8 presents Grad-CAM overlay results that visualize the specific spatial-frequency regions within PAC images that the Xception model emphasized while classifying neurocognitive disorders. The input PAC images (top row) represent the phase-amplitude coupling patterns for MCI and AD, while the bottom row highlights the corresponding Grad-CAM activations. The model's attention in MCI cases is centered around moderate delta-alpha and theta-beta couplings, reflecting adaptive changes, whereas in AD, the focus shifts toward impaired delta-alpha and theta-gamma regions, especially under stress conditions like hyperventilation. These overlays confirm that Xception effectively captures clinically relevant PAC features, demonstrating its strong interpretability and diagnostic utility in distinguishing between MCI and AD.

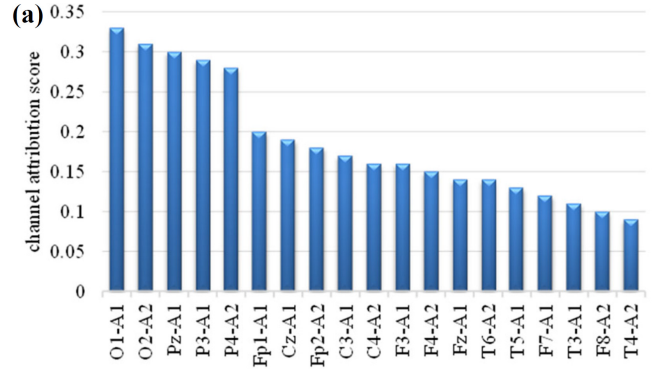


Figure 7: EEG channel attribution scores for Disorder Detection that rank channels in descending order of importance using Grad-CAM.

Figure 9 presents the results of an ablation study evaluating the Xception model's performance when retrained with a progressively increasing number of EEG channels ranked by Grad-CAM. In Figure 9 (a), the MAPE decreases significantly as the top-ranked channels are added, reaching its lowest around the 5 channels including occipital and parietal regions, which highlights their strong contribution to accurate neurocognitive disorder detection. Figure 9 (b) shows a gradual, near-exponential increase in training time as more channels are included, indicating a trade-off between accuracy and computational cost. These findings support the efficient selection of key EEG channels to optimize both model performance and training efficiency.

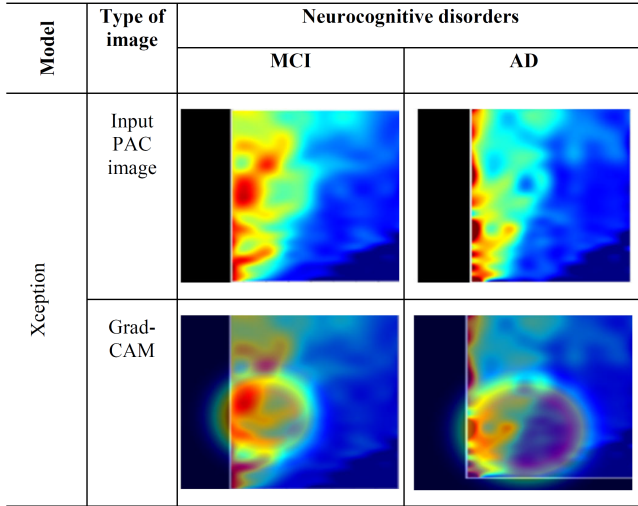


Figure 8: Grad-CAM overlays on PAC images showing the spatial-frequency regions the Xception model focused on for classifying MCI and AD, highlighting key discriminative patterns in each case.

4. DISCUSSION

4.1. Accuracy of detection

The Xception architecture-based CNN model with phase-amplitude coupling images achieved 98.95 % accuracy, outperforming MobileNetV2 by 2.50 % accuracy, as shown in Table 2. Xception's superior feature extraction, deeper spatial analysis, and efficient depthwise separable convolutions capture better EEG-based cerebral patterns recognition, enabling higher accuracy in detecting MCI and AD-related neural dysfunction as represented in Figure 3.

PAC images outperform STM images in neurocognitive disorder detection by capturing non-linear interactions between low-frequency phase (delta, theta) and high-frequency amplitude (alpha, beta) oscillations, revealing functional connectivity disruptions as illustrated in Figure 5 and Figure 6. During HV, cerebral vasoconstriction reduces oxygen delivery, disrupting neural activity in MCI and AD patients, leading to reduced alpha power and increased slow-wave activity. Unlike STM, which maps power distribution in different regions of the brain, PAC highlights phase synchronization deficits, making it a more reliable tool for early detection of cognitive decline.

Pearson correlation and Granger causality tests were applied to select EEG channels for MCI and AD detection by analyzing linear and directional connectivity as depicted in Figure 4. Pearson correlation identified reduced occipital-parietal and temporal connectivity, indicating neural dysfunction, memory decline, and inter-hemispheric asymmetry. Granger causality revealed weakened frontal-to-occipital and occipital-to-parietal causal flow, reflecting predictive connectivity loss and cognitive deficits. These tests highlighted memory disruptions, aiding in EEG-based early neurocognitive disorder detection.

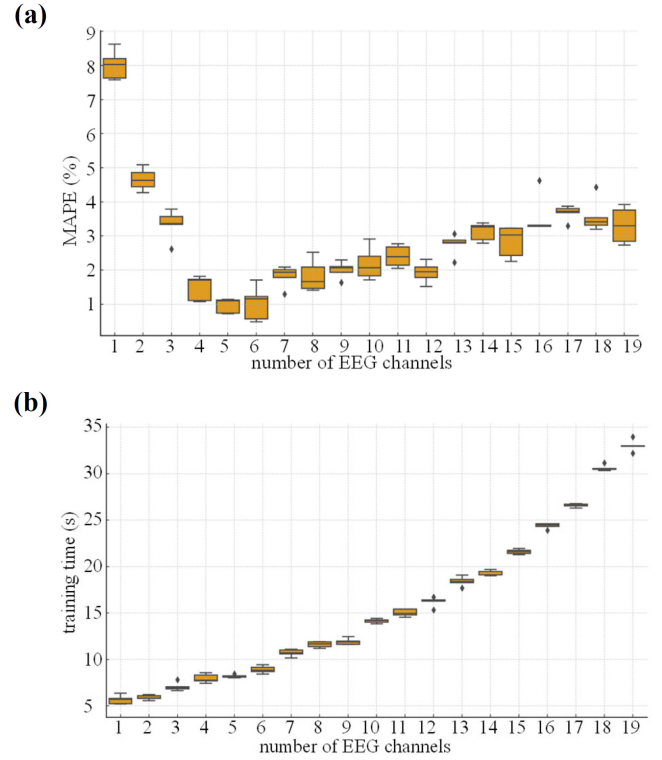


Figure 9: Ablation study — evaluation of Xception model to retrain with the variation of channels; (a) MAPE, (b) training time.

The performance evaluation of the Xception model using Grad-CAM reveals strong interpretability for detecting MCI and AD, as shown in Figure 7, Figure 8, and Figure 9. EEG channel attribution highlights occipital and parietal regions as critical, indicating disrupted posterior-anterior brain connectivity and altered neuronal excitability. Grad-CAM overlays on PAC images show the model focuses on low-to-mid frequency patterns for MCI and broader regions for AD, reflecting distinct coupling abnormalities. Ablation analysis confirms that using top-ranked channels significantly reduces MAPE and training time, demonstrating the model's efficiency and ability to detect disease-specific neurophysiological markers with fewer inputs.

4.2. Comparison of the proposed model with previous studies

The proposed Xception-based AI model with phase-amplitude coupling images achieves 98.95 % accuracy in MCI vs. AD classification, outperforming previous studies as shown in Table 3. Prior models, including EEGNet-based CNN (96.00 % accuracy, HC vs. MCI) and MNet-based Deep CNN (92.70 % accuracy, HC vs. AD), primarily utilized raw EEG with eyes-closed states. Unlike these approaches, the proposed model leverages HV-induced EEG changes, effectively capturing non-linear brain dynamics, enhancing neurocognitive disorders detection and its practical applicability.

Table 3: Comparison with previous studies to classify NCDs.

Study	Model	Input	State	Results
[4]	CNN (EEGNet)	Raw EEG	EC	HC vs. MCI (2 classes) F1-sc.: 96.00 % Acc.: 96.00 %
[5]	CNN (MNet)	Raw EEG	EC	MCI vs. AD (2 classes) Acc.: 92.70 %
Prop.	CNN (Xception)	PAC image	HV	MCI vs. AD (2 classes) Acc.: 98.95 %

4.3. Further development

The AI model with Xception deep-learning architecture by utilizing phase-amplitude coupling images demonstrates reliable performance in classifying MCI and AD patients' cerebral patterns during hyperventilation. To enhance generalizability and clinical applicability, future research will focus on validating the Xception-based AI model using larger, more diverse cohorts through multi-hospital collaborations. Although current dataset expansion is limited by hospital access, subject-level diversity remains a priority. Future work will also include HC data to better distinguish normal brain function from pathological brain activity and improve triage capability. Addressing sample imbalance and incorporating variability in age, sex, and comorbidities will further refine model performance. Additionally, integrating complementary stress-inducing tasks such as EO/EC states and photic stimulation, along with combining CNNs with other deep learning models and applying repeated-measure protocols, will help reduce variability and strengthen its practical application for early neurocognitive disorder detection. Future efforts will focus on optimizing Xception through compression and quantization to support real-time, point-of-care clinical screening.

5. CONCLUSION

This study proposed a phase-amplitude coupling image-based deep learning AI model with Xception architecture to detect impaired phase synchronization and disrupted neural communication across brain regions before, during, and after HV in MCI and AD patients. By comparing MobileNetV2 and Xception architectures with spectral topographic mapping and phase-amplitude coupling images, the Xception architecture-based AI model effectively captures HV-induced neural disruptions and inter-hemispheric asymmetry for MCI and AD classification. The findings highlight phase-amplitude coupling images as key biomarkers for detecting cognitive dysfunction, offering a simple and accurate approach for HV-induced neurocognitive disorder detection.

References

- [1] C. J. Huggins, J. Escudero, et al., "Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing," *J. Neural. Eng.*, vol. 18, no. 4, 2021.
- [2] World Health Organization, Dementia: Key facts, Accessed: Feb 20, 2025. Available at: <https://www.who.int/news-room/fact-sheets/detail/dementia>.
- [3] J. P. Coles, T. D. Fryer, et al., "Incidence and Mechanisms of Cerebral Ischemia in Early Clinical Head Injury," *Journal of Cerebral Blood Flow & Metabolism*, vol. 24, pp. 202-211, 2004.
- [4] Y. Watanabe, Y. Miyazaki Y, et al., "A deep learning model for the detection of various dementia and MCI pathologies based on resting-state electroencephalography data: A retrospective multicentre study," *Neural Networks*, vol. 171, pp. 242-250, 2024.
- [5] M. Şeker, et al., "Deep insights into MCI diagnosis: A comparative deep learning analysis of EEG time series," *J. Neurosci. Methods.*, vol. 403, 2024.
- [6] F. Miraglia, C. Pappaletta, et al., "The combination of hyperventilation test and graph theory parameters to characterize EEG changes in mild cognitive impairment (MCI) condition," *Geroscience*, vol. 45, pp. 1857-1867, 2023.
- [7] G. Coppola, A. Currà, et al., "Changes in visual-evoked potential habituation induced by hyperventilation in migraine," *J. Headache Pain.*, vol. 11, pp. 497-503, 2010.
- [8] T. Fujimatsu, T. Sugi, K. Goto, Y. Matsuda, et al., "Improvement of an EEG analysis system for patients with dementia," *The Twenty-Seventh International Symposium on Artificial Life and Robotics 2022 (AROB 27th 2022)*, Japan, 2022.
- [9] W. Y. Yu, I. Low, et al., "Brain Dynamics Altered by Photic Stimulation in Patients with Alzheimer's Disease and Mild Cognitive Impairment," *Entropy (Basel)*, vol. 23, 2021.
- [10] K. Tara, M. H. Islam, T. Sugi, "ECG based human activity-specific cardiac pattern detection using machine-learning and deep-learning models," *Journal of Electrocardiology*, vol. 90, 2025.
- [11] M. Wang, Z. Liao, et al., "Application of Granger Causality Analysis of the Directed Functional Connection in Alzheimer's Disease and Mild Cognitive Impairment," *J. Vis. Exp.*, vol. 126, 2017.
- [12] X. Bi, et al., "Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning," *Neural Networks*, vol. 114, pp. 119-135, 2019.
- [13] J. Poza, A. Bachiller, et al., "Phase-amplitude coupling analysis of spontaneous EEG activity in Alzheimer's disease," *Annu. Int. Conf IEEE Eng. Med. Biol. Soc.*, 2017, pp. 2259-2262.
- [14] M. Aparna, B. S. Rao, "Xception-Fractalnet: Hybrid Deep Learning Based Multi-Class Classification of Alzheimer's Disease," *Computers, Materials and Continua.*, vol. 74, pp. 6909-6932, 2022.
- [15] C. V. Zyl, X. Ye, R. Naidoo, "Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of Grad-CAM and SHAP," *Applied Energy*, vol. 353, 2024.