Motor Imagery EEG Classification using Convolutional Neural Networks with Data Augmentation via Multiple Channel Sets

Ali Yasser Mohammad*
Dr. Ghadi Mahmoudi**
Dr. Hassan Bustani***

(Received 27 / 7 / 2025. Accepted 23 / 6 / 2026)

□ ABSTRACT □

Brain-Computer Interfaces (BCIs) provide a direct communication link between the human brain and external devices, with Electroencephalography (EEG) being a widely adopted non-invasive recording modality. Motor Imagery (MI), the mental simulation of movement, is a common BCI paradigm, though its accurate classification from EEG signals is challenging. This research proposes a Convolutional Neural Network (CNN) model for MI-EEG classification, incorporating a channel-set-based data augmentation strategy using three 6-channel EEG sets. For each original MI trial from the EEG Motor Movement/Imagery Dataset (103 subjects, 11,503 unique trials), three augmented samples were generated from three distinct 6-channel sets (fronto-central, central, and centro-parietal), each sample having dimensions of 6 channels × 640 time points.

After band-pass filtering (8–40 Hz) and subject-wise channel z-score normalization, the augmented dataset (34,509 samples) was used to train and evaluate the CNN. Under a mixed-subject/sample-wise evaluation protocol, the model achieved a validation accuracy of 91.55% for five classes. In addition, a stricter subject-wise 5-fold cross-validation protocol was added to evaluate generalization to unseen subjects; this protocol yielded a trial-level accuracy of 50.52% ± 5.82% and a trial macro-F1 of 50.65% ± 5.83%. These results clarify the difference between within-dataset mixed evaluation and cross-subject generalization, and demonstrate the usefulness and limitations of channel-set-based augmentation for MI-EEG classification.

Keywords: Brain-Computer Interface (BCI), Electroencephalography (EEG), Motor Imagery (MI), Convolutional Neural Network (CNN).

Copyright **:**Latakia University journal (formerly Tishreen) -Syria, The authors retain the copyright under a CC BY-NC-SA 04

*Master's Student, Department of Telecommunication and Electronic Engineering, Faculty of Mechanical and Electrical Engineering, Lattakia University (formerly Tishreen), Latakia, Syria. Email: mailto:ali.y.mohammad@latakia-univ.edu.sy

** Assistant Professor, Department of Telecommunication and Electronic Engineering, Faculty of Mechanical and Electrical Engineering, Lattakia University (formerly Tishreen), Latakia, Syria. Email: g.mahmoudi@tishreen.edu.sy, ghadi.mahmoudi@fra-uas.de.

*** Assistant Professor, Department of Telecommunication and Electronic Engineering, Faculty of Information and Communications Technology Engineering – Tartous University, Tartous, Syria. Email: Hasanalbustani@yahoo.com.

تصنيف تخيل الحركة من التخطيط الكهربائي للدماغ باستخدام الشبكات العصبونية التلافيفية مع توسعة البيانات عبر مجموعات قنوات متعددة

علي ياسر محمد *

د. غادي محمودي **

د. حسن بستاني ***

(تاريخ الإيداع 27 / 7 / 2025. قُبِل للنشر في 23/ 6 / 2026)

□ ملخّص □

توفر واجهات الدماغ والحاسوب (BCI) وصلة مباشرة بين دماغ الإنسان والأجهزة الخارجية، ويُعد تخطيط كهربائية الدماغ (EEG) تقنية تسجيل غير جراحية شائعة. يمثل تخيل الحركة (MI)، وهو المحاكاة الذهنية للحركة، نموذجاً شائعاً في BCI، لكن تصنيفه الدقيق من إشارات EEG يمثل تحدياً. يقترح هذا البحث نموذج شبكة عصبونية تلافيفية (CNN) لتصنيف إشارات MI-EEG، مستخدماً استراتيجية لتوسعة البيانات تعتمد على ثلاث مجموعات قنوات سداسية. لكل تجربة MI أصلية من مجموعة البيانات EEG Motor Movement/Imagery Dataset، وبعد اعتماد 103 مشاركين و 11,503 حِقبة أصلية، تم توليد ثلاث عينات مُعززة من ثلاث مجموعات قنوات مميزة (أمامية–مركزية، مركزية، ومركزية–جدارية)، بأبعاد 6 قنوات × 640 نقطة زمنية للعينة.

بعد تطبيق ترشيح نطاقي 8–40 هرتز وتطبيع z-score خاص بكل مشارك ولكل قناة، استُخدمت البيانات المعززة، وعددها 34,509 عينة، لتدريب نموذج CNN وتقييمه. حقق النموذج دقة تحقق بلغت 91.55% لخمس فئات ضمن بروتوكول التقييم المختلط/المعتمد على العينات. كما أُضيف تقييم أكثر صرامة باستخدام subject-wise 5-fold cross-validation، بحيث لا يظهر أي مشارك في التدريب والتحقق ضمن الطية نفسها، وبلغت الدقة على مستوى الحِقبة الأصلية 50.52% ± 5.82%، بينما بلغت قيمة Trial macro-F1 مقدار 50.65% ± 5.83%.

الكلمات المفتاحية: واجهة الدماغ والحاسوب (BCI)، تخطيط كهربائية الدماغ (EEG)، تخيل الحركة (MI)، الشبكة العصبونية التلافيفية (CNN).

حقوق النشر : مجلة جامعة اللاذقية (تشرين سابقاً)– سورية، يحتفظ المؤلفون بحقوق النشر بموجب الترخيص CC BY–NC–SA 04

* طالب ماجستير، قسم هندسة الاتصالات والالكترونيات، كلية الهندسة الميكانيكية والكهربائية، جامعة اللاذقية (تشرين سابقاً)، اللاذقية، سورية. البريد الالكتروني: ali.y.mohammad@latakia-univ.edu.sy

** مدرس، قسم هندسة الاتصالات والالكترونيات، كلية الهندسة الميكانيكية والكهربائية، جامعة اللاذقية (تشرين سابقاً)، اللاذقية، سورية. البريد الالكتروني: ghadi.mahmoudi@fra-uas.de , g.mahmoudi@tishreen.edu.sy

** مدرس، قسم هندسة الاتصالات والالكترونيات، كلية هندسة تكنولوجيا المعلومات والاتصالات – جامعة طرطوس ، طرطوس، سورية .البريد الالكتروني: .Hasanalbustani@yahoo.com

مقدمة:

تُعد واجهة الدماغ والحاسوب (Brain-Computer Interface - BCI) تقنية ناشئة تهدف إلى إنشاء مسار تواصل مباشر بين الدماغ البشري والأجهزة الخارجية، متجاوزةً المسارات العصبية العضلية التقليدية[1]. تكتسب هذه التقنية أهمية خاصة للأفراد الذين يعانون من إعاقات حركية شديدة، حيث تمنحهم القدرة على التحكم بالأجهزة المساعدة كالكراسي المتحركة والأطراف الاصطناعية أو حتى التواصل عبر الحاسوب. من بين طرق التقاط النشاط الدماغي المختلفة، يبرز تخطيط كهربائية الدماغ (Electroencephalography - EEG) كخيار غير جراحي مفضل في العديد من تطبيقات BCI البحثية والعملية، وذلك بفضل سهولة تطبيقه، وتكلفته المنخفضة نسبياً، وقدرته الجيدة على تتبع التغيرات السريعة في النشاط الدماغي (الدقة الزمنية)[3] ,[2].

تتميز إشارات EEG بأنماط ترددية مختلفة: دلتا δ (0.5–4 هرتز)، وثيتا θ (4–7 هرتز)، وألفا α (8–13 هرتز)، وبيتا β (14–30 هرتز)، وموجات جاما γ (>30 هرتز)، وترتبط هذه الأنماط بحالات دماغية متنوعة. في سياق BCI، يُستخدم نموذج تخيل الحركة (Motor Imagery - MI) بشكل واسع، وهو يعتمد على قدرة الشخص على تعديل نشاط دماغه إرادياً عبر تخيل أداء حركة معينة دون تنفيذها فعلياً، وتظهر الاستجابات المرتبطة به غالباً ضمن نطاقي mu و beta فوق المناطق الحسية–الحركية [2].

على الرغم من التقدم المحرز، لا يزال تصنيف مهام MI بدقة يمثل تحدياً كبيراً بسبب الطبيعة المعقدة لإشارات EEG، وانخفاض نسبة الإشارة إلى الضجيج، والتأثر بالتشوهات، والتباين الكبير بين الأفراد. بينما استكشفت الأبحاث السابقة مناهج متنوعة، بدءاً من الطرق التقليدية المعتمدة على استخلاص السمات يدوياً وصولاً إلى تقنيات التعلم العميق، لا تزال هناك حاجة لتطوير نماذج تحقق توازناً بين الدقة العالية، وسهولة التطبيق، والقدرة على العمل مع معالجة أولية مبسطة. يهدف هذا البحث إلى المساهمة في هذا الاتجاه من خلال تطوير وتقييم نموذج شبكة عصبونية تلافيفية (CNN) لتصنيف إشارات MI-EEG، مع التركيز على استخدام معالجة أولية بسيطة واستراتيجية مبتكرة لتوسعة البيانات تعتمد على تمثيل كل تجربة MI أصلية عبر مجموعات قنوات متعددة، وذلك بهدف تحسين دقة التصنيف ومتانته لأنظمة BCI العملية.

الدراسات السابقة

شهد مجال تصنيف مهام (MI) من إشارات (EEG) تطوراً ملحوظاً في السنوات الأخيرة، مدفوعاً بالحاجة المتزايدة لواجهات دماغ–حاسوب (BCI) فعالة. ركزت العديد من الأبحاث على استخدام تقنيات التعلم العميق، وبخاصة شبكات (CNNs)، نظراً لقدرتها على استخلاص السمات المعقدة تلقائياً من بيانات EEG.

في هذا السياق، تم اقتراح نماذج CNN متنوعة لمعالجة إشارات MI-EEG. تُعد بنية DeepConvNet التي قدمتها دراسة [4] من النماذج المرجعية المهمة في تصنيف إشارات EEG باستخدام الشبكات التلافيفية، إذ أظهرت قدرة الشبكات العميقة على تعلم السمات الزمنية–المكانية مباشرة من الإشارة. كما قدمت دراسة [5] نموذج CMO-CNN، وهو شبكة CNN أحادية البعد (1D-CNN) مدمجة ومتعددة الفروع، مصممة للعمل مباشرة على إشارات EEG الخام مع دمج كتل الانتباه أحادية البعد (1D Squeeze-and-Excitation) ووصلات التخطي (Skip Connections)، وأظهرت نتائج جيدة على مجموعتي البيانات BCI Competition IV 2a و 2b.

استخدمت بعض الدراسات الحديثة قاعدة PhysioNet EEG Motor Movement/Imagery Dataset وبعدد 103 مشاركين بعد استبعاد السجلات ذات الوسوم غير الدقيقة. فقد اعتمد الباحثون في [6]

نموذج EEGNet Fusion على بيانات مقطعة من تجارب مدتها 4 ثوان، ثم قُسمت كل تجربة إلى ثمان نوافذ غير متداخلة بطول 80 عينة مع تطبيق مرشح Notch عند 60 هرتز ومرشح Butterworth Band-pass بين 2 و60 هرتز. وطورت دراسة [7] نموذج EEGNet Fusion V2 بآلية تجهيز مشابهة، واستخدموا تقسيماً عشوائياً للعينات بنسبة 70% للتدريب و10% للتحقق و20% للاختبار. كما قدمت دراسة [8] آلية Feature Reweighting لأربع فئات من التخيل الحركي باستخدام القنوات الـ 64 والمقطع 0–4 ثوان بعد ظهور إشارة التخيل، مع تقييم بخمس طيات دون تصريح واضح بأن الطيات مفصولة على مستوى المشاركين. لذلك تُعد نتائج هذه الأعمال مهمة للسياق، لكنها لا تكافئ مباشرة التقييم العابر للمشاركين المستخدم في هذه الدراسة.

اعتمد نهج آخر على دمج السمات الهندسية مع تلك التي تتعلمها الشبكات. في دراسة [9]، تم تطوير نموذج MS-CNN متعدد المقاييس الذي يعالج نطاقات ترددية مختلفة بشكل متوازٍ، وتم تعزيز أدائه عبر دمج سمات مثل الإنتروبية التفاضلية (DE) وطيف القدرة العصبي (NPS) مع مخرجات الشبكة، بالإضافة إلى تطبيق مجموعة متنوعة من تقنيات توسعة البيانات لتحسين الدقة والمتانة على مجموعة بيانات BCI Competition IV 2b.

يمثل التعامل مع التشوهات، خاصة الناتجة عن الحركة، تحدياً هاماً في تطبيقات BCI العملية. قدمت دراسة [10] نموذجاً تجريبياً يعتمد على بيانات مقياس التسارع وخصائص البيئة لإزالة التشوهات الحركية من إشارات MI-EEG المسجلة من مستخدمي الكراسي المتحركة، ثم استخدمت شبكة CNN معدلة (LM-CNN) لتصنيف الإشارات التي تم حذف التشوهات منها.

في إطار النماذج الهجينة، استكشفت الدراسة [11] دمج شبكة CNN متعددة الطبقات مع مكدس من المرمزات الذاتية المتناثرة (SSAE). هدفت هذه البنية إلى استخلاص تمثيلات مكانية من مجموعات قنوات EEG مختلفة ودمجها عبر SSAE لالتقاط سمات ثابتة عبر المشاركين وتعزيز قدرة التعميم.

أظهرت دراسات أخرى مثل [12] فعالية بنى CNN مخصصة تتضمن كتل تلافيفية زمنية ومكانية متتالية في تصنيف إشارات MI-EEG، مع التركيز على تصميم شبكة قادرة على استخلاص السمات الزمنية–المكانية وتحقيق أداء تنافسي. بينما تناولت دراسة [13] تطوير نموذج CNN أحادي البعد يعمل مباشرة على إشارات EEG الخام مع دمج آلية انتباه ذاتي (self-attention mechanism) لتركيز النموذج على الأجزاء الزمنية والقنوات الأكثر أهمية، مع تقييم شامل لعدة تقنيات توسعة بيانات.

حتى عند تطبيق هياكل CNN قياسية مثل ResNet18 على مهام معرفية مختلفة عن تخيل الحركة، كما في دراسة [14] التي حللت الإشارات المرتبطة بالحدث (ERPs)، تم التأكيد على أهمية اختيار النافذة الزمنية المناسبة للتحليل والتحديات المرتبطة بتصنيف إشارات خام مجمعة من عدة مشاركين.

بشكل عام، تُظهر هذه الأعمال التوجه المتزايد نحو استخدام الشبكات العصبونية التلافيفية بأشكالها المختلفة لتصنيف إشارات MI-EEG، مع استكشاف دمج السمات، وآليات الانتباه، وتقنيات توسعة البيانات، والتعامل مع التشوهات. ومع ذلك، لا تزال هناك حاجة مستمرة لتطوير نماذج تحقق توازناً بين الدقة العالية، والقدرة على التعميم عبر المشاركين، والعمل بفعالية مع معالجة أولية بسيطة، واستخدام تشكيلات قنوات عملية. يهدف بحثنا الحالي إلى المساهمة في هذا المجال من خلال تقييم أداء بنية CNN قياسية مع تطبيق استراتيجية توسعة بيانات محددة تعتمد على مجموعات قنوات متعددة، وذلك على مجموعة بيانات كبيرة من حيث عدد المشاركين، مع الحفاظ على بساطة المعالجة الأولية.

أهمية البحث وأهدافه:

تكتسب (BCI) أهمية متزايدة في مجالات متعددة، لعل أبرزها هو تمكين الأفراد ذوي الإعاقات الحركية الشديدة من استعادة جزء من استقلاليتهم عبر التحكم بالأجهزة الخارجية والتواصل مع البيئة المحيطة باستخدام نشاطهم الدماغي فقط. يعتمد (MI) كنموذج أساسي في العديد من أنظمة BCI، لكن تصميم أنظمة تصنيف دقيقة وموثوقة لإشارات MI-EEG يمثل تحدياً كبيراً.

تواجه أنظمة التصنيف التقليدية صعوبات تتعلق بانخفاض نسبة الإشارة إلى الضجيج في إشارات EEG، والتباين الكبير بين الأفراد، والحاجة غالباً لخطوات معالجة أولية معقدة، ومرحلة استخراج سمات تتطلب خبرة هندسية وطبية حيوية دقيقة وقد تؤدي لفقدان معلومات هامة. كما أن الاعتماد على عدد كبير من الأقطاب الكهربائية يزيد من كلفة وتعقيد هذه الأنظمة.

أظهرت تقنيات التعلم العميق، وبخاصة (CNN)، قدرة واعدة على معالجة البيانات المعقدة مثل إشارات EEG وتعلم السمات التمييزية مباشرةً، مما يفتح المجال لتبسيط عملية التصميم وتحسين الأداء. بالإضافة إلى ذلك، يمكن لتقنيات توسعة البيانات (Data Augmentation) أن تزيد من فعالية نماذج التعلم العميق عبر تعزيز تنوع بيانات التدريب المتاحة.

يهدف هذا البحث بشكل أساسي إلى:

1. تطوير وتقييم نموذج تصنيف لإشارات MI-EEG يعتمد على بنية CNN قادرة على العمل بفعالية مع معالجة أولية بسيطة تقتصر على الترشيح النطاقي.
2. تقييم قدرة النموذج على التمييز بين خمس فئات (أربع مهام تخيل حركة شائعة وحالة الراحة).
3. استخدام استراتيجية لتوسعة البيانات تعتمد على تمثيل كل تجربة MI أصلية عبر ثلاث مجموعات مختلفة من القنوات السداسية (6-channel sets) المتموضعة في مناطق دماغية مختلفة ذات صلة بالنشاط الحركي (الأمامية–المركزية، المركزية، والمركزية–الجدارية)، بهدف تحسين قدرة النموذج على استخلاص سمات مكانية متنوعة، والتغلب على تحديات فرط التلبيق overfitting، وتعزيز قدرته على التعميم عبر المشاركين.
4. تقييم أداء النموذج وفق بروتوكولين مختلفين: تقييم مختلط/معتمد على العينات لتقدير الأداء داخل مجموعة البيانات، وتقييم عابر للمشاركين subject-wise لقياس قدرة النموذج على التعميم على مشاركين غير مرئيين أثناء التدريب.
5. توفير مقارنة مرجعية مباشرة مع نموذج EEGNet [15] على مجموعة البيانات نفسها وبروتوكول التقييم نفسه، إضافةً إلى تقديم تحليل تفسيري أولي باستخدام Grad-CAM.

طرائق البحث ومواده:

اعتمدت الدراسة على منهجية تجريبية قائمة على بناء وتدريب وتقييم نموذج تعلم عميق لتصنيف MI-EEG. تم تنفيذ جميع مراحل العمل باستخدام لغة البرمجة Python (الإصدار 3.12) ضمن بيئة Google Colaboratory، بالاعتماد على المكتبات العلمية المتخصصة التالية:

- MNE-Python (الإصدار 1.9.0): لتحميل بيانات EEG الأولية ومعالجتها وتحليلها وتجزئتها [16].
- TensorFlow مع Keras API: لبناء بنية نموذج CNN، وتدريبه، وتقييمه.
- NumPy: للعمليات الحسابية ومعالجة المصفوفات العددية.
- Scikit-learn: لتقسيم البيانات وحساب مقاييس الأداء التفصيلية (مصفوفة الالتباس، Precision, Recall, F1-Score).
- Matplotlib: لإنشاء الرسوم البيانية وتصوير النتائج ومنحنيات التعلم.
- Pickle: لحفظ وتحميل كائنات البيانات المعالجة مسبقاً.

1. مجموعة البيانات(Dataset) :

تم الاعتماد على مجموعة البيانات المتاحة للعموم "EEG Motor Movement/Imagery Dataset" (الإصدار 1.0.0)، والمقدمة عبر منصة PhysioNet[17]. تم جمع هذه البيانات من 109 مشتركين أصحاء باستخدام نظام BCI2000، وتضمنت تسجيلات EEG باستخدام 64 قطباً كهربائياً موزعة وفقاً للنظام الدولي 10–10، وبمعدل أخذ عينات 160 هرتز.

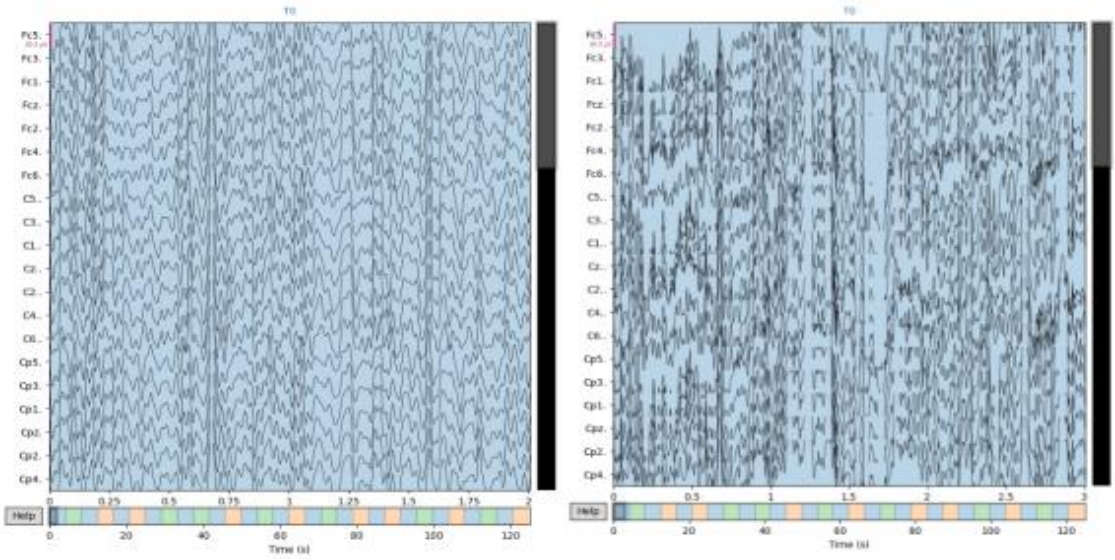
2. المشاركون والمهام المختارة:

تم استبعاد بيانات 6 مشاركين (الأرقام 38، 88، 89، 92، 100، 104) بناءً على توصيات مرتبطة باستخدام هذه المجموعة في دراسات سابقة، ليصبح العدد النهائي للمشتركين المعتمدين 103. تم التركيز في هذه الدراسة على الأشواط التجريبية (Runs) التي تضمنت مهام (MI) وفترات الراحة (Baseline)، وتحديداً الأشواط 4، 8، 12 (التي تحتوي على مهام: Baseline, Left Fist MI, Right Fist MI) والأشواط 6، 10، 14 (التي تحتوي على مهام Baseline, Both Fists MI, Both Feet MI).

3. المعالجة الأولية للإشارات وتجميع البيانات:

خضعت بيانات (EEG) الخام، المسجلة بصيغة EDF، لسلسلة من خطوات المعالجة الأولية بهدف تنقيتها وتجهيزها للتحليل اللاحق. تم تنفيذ هذه الخطوات باستخدام أدوات متخصصة لمعالجة إشارات EEG ضمن بيئة البرمجة بايثون.

أولاً، تم تطبيق مرشح تمرير حزمة رقمي (Digital Band-pass Filter) من نوع الاستجابة النبضية المحدودة (FIR) على بيانات كل تسجيل (شوط تجريبي). تم تصميم هذا المرشح للسماح بمرور الترددات الواقعة ضمن النطاق 8 هرتز إلى 40 هرتز فقط، وهو النطاق الذي يحتوي على إيقاعات الدماغ (مثل mu و beta) الأكثر ارتباطاً بنشاط تخيل الحركة، مع التخفيف من أثر الترددات المنخفضة جداً والترددات العالية. يمكن تصور تأثير عملية الترشيح هذه من خلال مقارنة شكل الإشارة قبل وبعد تطبيق المرشح (الشكل 1).



(أ) قبل الترشيح (ب) بعد الترشيح

الشكل 1. عينة من إشارات EEG مدتها 3 ثوان قبل وبعد تطبيق مرشح تمرير الحزمة (8–40 هرتز)

ثانياً، تم تنظيف التعليقات التوضيحية (Annotations) المرفقة مع البيانات، وذلك بحذف أي علامات تشير إلى مقاطع غير صالحة أو حدود زائفة قد تكون ناتجة عن عمليات تسجيل أو معالجة سابقة.

ثالثاً، ولتوحيد مرجعية مواقع الأقطاب الكهربائية عبر جميع المشاركين والتسجيلات، تم تطبيق مونتاج قياسي (Standard Montage) هو standard_1005. تضمن ذلك أيضاً توحيد أسماء الأقطاب الكهربائية وفقاً للمعيار المستخدم.

رابعاً، تم قص (Cropping) أي شوط تجريبي يتجاوز طوله الإجمالي 124 ثانية إلى هذه المدة القصوى لضمان اتساق مدة التسجيلات المستخدمة.

خامساً، تم إعادة تسمية الأحداث (Events) المسجلة في التعليقات التوضيحية إلى رموز دلالية تمثل المهام الفعلية: حالة الراحة ('B' - Baseline')، تخيل حركة القبضة اليسرى ('L')، تخيل حركة القبضة اليمنى ('R')، تخيل حركة كلتا القبضتين ('LR')، وتخيل حركة كلتا القدمين ('F').

سادساً، تم تقطيع (Epoching) التسجيلات المستمرة والمرشحة إلى مقاطع زمنية منفصلة تسمى حِقَب أو تجارب (Epochs/Trials). تم استخلاص حِقبة بطول 4 ثوانٍ لكل حدث مُعاد تسميته. تم تحديد بداية كل حِقبة بعد فترة قصيرة جداً (160/1 من الثانية) من ظهور علامة الحدث. لم يتم تطبيق أي تصحيح لخط الأساس (Baseline Correction) في هذه المرحلة.

أخيراً، تم تجميع جميع الحِقَب الناتجة من الأشواط التجريبية المختارة ولكافة المشاركين المعتمدين وعددهم 103 مشاركين. في هذه المرحلة، احتفظت كل حِقبة ببيانات القنوات الـ 64 الأصلية، ثم نُظمت الحِقَب وفق خمس فئات رئيسية تمثل المهام قيد الدراسة: Right Hand، Left Hand، Both Hands، Both Feet، وBaseline. شكّلت هذه الحِقَب الأصلية قاعدة البيانات الأولية المشتركة التي انطلقت منها خطوات اختيار القنوات، وتوسعة البيانات، وتحضير البيانات النهائية وفق بروتوكولي التقييم المعتمدين في الدراسة.

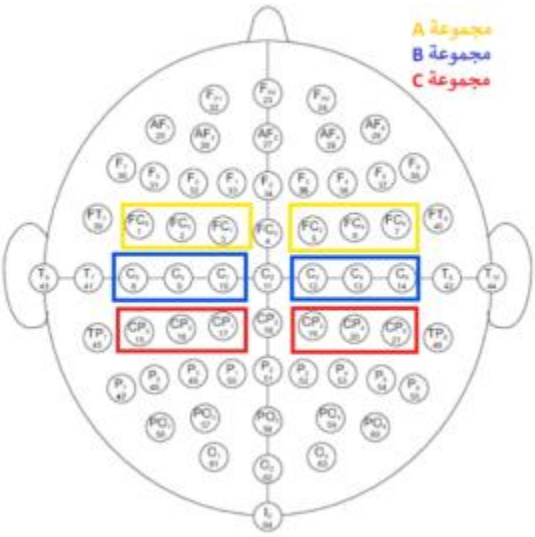
4. اختيار القنوات، توسعة البيانات، وتحضير البيانات النهائية للتدريب:

في هذه المرحلة، تم تطبيق استراتيجية مبتكرة لاختيار مجموعات قنوات محددة وتنفيذ عملية توسعة للبيانات (Data Augmentation) بهدف تحسين أداء النموذج وقدرته على التعميم.

تم تحديد ثلاث مجموعات قنوات سداسية (6-channel sets)، تم اختيارها لتمثيل مناطق دماغية مختلفة يُعتقد أنها نشطة أثناء مهام تخيل الحركة، وهي:

1. المجموعة "a" (المنطقة الأمامية–المركزية – Fronto-Central): تشمل القنوات FC1, FC2, FC3, FC4, FC5, FC6.
2. المجموعة "b" (المنطقة المركزية – Central): تشمل القنوات C5, C6, C3, C4, C1, C2.
3. المجموعة "c" (المنطقة المركزية–الجدارية – Centro-Parietal): تشمل القنوات CP1, CP2, CP3, CP4, CP5, CP6.

يوضح الشكل 2 مواقع القنوات المستخدمة ضمن المجموعات الثلاث "a", "b", "c" على فروة الرأس.

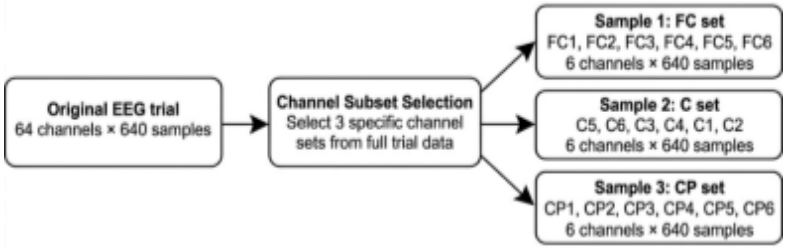


الشكل 2. مخطط القنوات والمجموعات

استند اختيار مجموعات القنوات السداسية (الأمامية–المركزية، المركزية، والمركزية–الجدارية) إلى نتائج دراسات سابقة أوضحت أهمية هذه المناطق القشرية في معالجة تخيل الحركة. على سبيل المثال، تؤكد الدراسة التي أجراها Mustile وزملاؤه في عام 2024 [18] على أن عملية تخيل الحركة تُفعِّل شبكة دماغية واسعة تتخطى حدود المناطق الحركية الأولية لتشمل مناطق حسية حركية وأخرى جدارية، وتستلزم هذه العملية تعبئة موارد قشرية إضافية. فالمنطقة المركزية (المجموعة 'b')، التي تغطي القشرة الحسية الحركية، تُعرف على نطاق واسع بأنها مركز لتعديلات ألفا وبيتا (α و β) المرتبطين بتمثيل الحركة وتخيلها. وبالمثل، فإن إدراج المناطق الأمامية–المركزية (المجموعة 'a') يأتي ليعكس الدور الحيوي للشبكات العصبية أمام الحركية في عمليات التخطيط والإعداد للحركة. أما اختيار المناطق المركزية–الجدارية (المجموعة 'c') فيعزى إلى الأهمية الثابتة للفص الجداري في عمليات التكامل الحسي–الحركي والمعالجة المكانية، وهي وظائف أساسية ضمن الشبكة العصبية الكلية المسؤولة عن تخيل الحركة. انطلاقاً من هذه المعطيات، فإن الهدف من استخدام عينات بيانات مُعززة ومُنتقاة من هذه

المناطق الدماغية المتميزة والمشاركة بفعالية، هو تزويد نموذج الشبكة العصبونية التلافيفية (CNN) بمنظورات مكانية متنوعة للنشاط الدماغي، الأمر الذي يُتوقع أن يُعزز من قدرة النموذج على التعلم الفعال والتعميم الدقيق.

تم تطبيق عملية توسعة البيانات على النحو التالي: لكل حِقبة أصلية تحتوي على 64 قناة، تم إنشاء ثلاث نسخ تمثل ثلاث رؤى مكانية مختلفة لنفس الحدث الدماغي، وذلك عبر اختيار قنوات المجموعة الأمامية–المركزية FC، والمجموعة المركزية C، والمجموعة المركزية–الجدارية CP. أصبحت كل عينة بيانات عبارة عن مصفوفة بأبعاد 6 قنوات × 640 نقطة زمنية. يوضح الشكل 3 هذه الآلية.



الشكل 3. مخطط استراتيجية توسعة البيانات عبر مجموعات القنوات الثلاث FC وC وCP.

يوضح الجدول 1 عدد العينات لكل فئة وبلغ الحجم النهائي للبيانات 11,503 حِقبة أصلية و34,509 عينة بعد التوسعة. لا تُستخدم هذه التوسعة قبل التقسيم في بروتوكول التقييم العابر للمشاركين، بل تُطبّق بعد تحديد مشاركي التدريب والتحقق ضمن كل طية، وذلك لمنع أي تسرب محتمل بين مجموعتي التدريب والتحقق.

جدول 1. التوزيع العددي للحِقَب الأصلية والعينات المعززة لكل فئة.

الفئة	عدد الحِقَب الأصلية	عدد العينات بعد التوسعة FC+C+CP
Right Hand	**2271**	**6813**
Left Hand	**2313**	**6939**
Both Hands	**2304**	**6912**
Both Feet	**2280**	**6840**
Baseline	**2335**	**7005**
Total	**11503**	**34509**

بعد ذلك، تم تحضير البيانات النهائية لتدريب النموذج وتقييمه وفق بروتوكولين مختلفين:

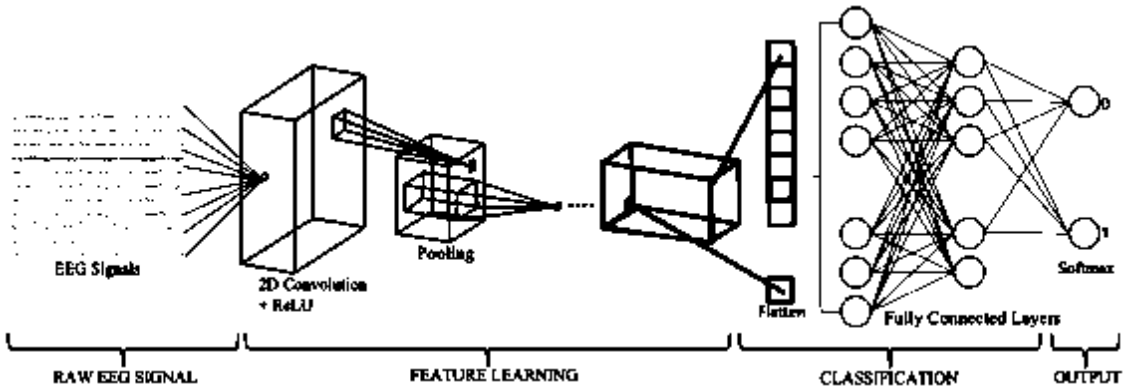
1. في بروتوكول التقييم المختلط/المعتمد على العينات، تم توليد العينات المعززة من الحِقَب الأصلية عبر مجموعات القنوات الثلاث، ثم خُلطت العينات الناتجة وقُسمت عشوائياً بنسبة 80% للتدريب و20% للتحقق. يهدف هذا البروتوكول إلى تقدير أداء النموذج داخل مجموعة البيانات، لكنه لا يُعد دليلاً كافياً على قدرة النموذج على التعميم على مشاركين غير مرئيين، بسبب احتمال وجود عينات مشتقة من مشاركين أو حِقَب أصلية متقاربة في مجموعتي التدريب والتحقق.
2. في بروتوكول التقييم العابر للمشاركين subject-wise 5-fold cross-validation، قُسم المشاركون إلى خمس طيات، بحيث تظهر بيانات كل مشارك إما في مجموعة التدريب أو في مجموعة التحقق ضمن الطية الواحدة، دون أي تداخل بين المشاركين. في هذا البروتوكول، طُبقت عملية اختيار القنوات وتوسعة البيانات بعد تقسيم المشاركين داخل كل طية، أي بعد تحديد مشاركي التدريب والتحقق، وذلك لمنع انتقال أي تمثيلات مشتقة من المشارك نفسه بين المجموعتين.

3. استُخدم subject-wise channel z-score normalization حيث حُسب المتوسط والانحراف المعياري لكل قناة ولكل مشارك على حدة من بيانات ذلك المشارك فقط ثم طُبّق التطبيع على بياناته. لذلك أصبح عامل التحجيم في النسخة المعدّلة SCALE_FACTOR = 1.0.
4. تحويل تسميات الفئات العددية (0، 1، 2، 3، 4) إلى صيغة الترميز الأحادي (One-Hot Encoding) حيث يتم تمثيل كل فئة بمتجه ثنائي يحتوي على '1' في الموضع المقابل للفئة و '0' في المواضع الأخرى (مثال: الفئة 2 تصبح [0، 0، 1، 0، 0]). هذا هو الشكل المطلوب لمخرجات الشبكة عند استخدام دالة خسارة التصنيف متعدد الفئات (categorical cross-entropy).
5. تحويل قوائم بيانات التدريب والتحقق إلى مصفوفات NumPy، وتم تعديل أبعادها لتناسب متطلبات الدخل لطبقات CNN في Keras، ليصبح شكل كل عينة هو (6، 640، 1). البعد الأخير (1) يشير إلى أن البيانات تمثل ميزة واحدة (الجهد) لكل نقطة زمنية وقناة.

لتقييم مدى فعالية استراتيجية توسعة البيانات المقترحة، أُجريت تجارب مقارنة باستخدام مجموعات القنوات المفردة، كما أُضيف نموذج EEGNet كنموذج مرجعي يُدرّب على مجموعة البيانات نفسها وببروتوكول subject-wise 5-fold cross-validation نفسه. عُرضت النتائج باستخدام المتوسط والانحراف المعياري وفترة الثقة 95% عبر الطيات الخمس.

5. بنية نموذج الشبكة العصبونية التلافيفية (CNN Model Architecture):

لتصنيف إشارات MI-EEG (ذات الأبعاد 6 قنوات × 640 نقطة زمنية × 1 ميزة)، تم بناء نموذج شبكة عصبونية (CNN) تسلسلية. حُددت البنية النهائية تجريبياً بعد اختبار عدة تكوينات من حيث عدد الكتل التلافيفية، وعدد المرشحات، ومعدلات Dropout، مع اعتماد البنية التي حققت أفضل توازن بين دقة التحقق والحد من فرط التلبيق. تتكون البنية من أربع كتل تلافيفية متتالية، كل منها تحتوي على طبقة Conv2D (بمرشحات متزايدة 16، 32، 64، 128، وحجم نواة 3×3)، تليها طبقة MaxPooling2D (2x2) لتقليل الأبعاد، ثم طبقة Dropout (بمعدل 0.2) للحد من فرط التلبيق.



الشكل 4. شكل توضيحي لآلية عمل شبكة CNN

تم اختيار دالة التنشيط elu (Exponential Linear Unit) في الطبقات التلافيفية والكثيفة نظراً لقدرتها على التعامل مع القيم السالبة الموجودة في بيانات EEG المدخلة، مما قد يساعد في تمثيل أفضل للبيانات وتجنب مشكلة "العصبونات الميتة".

بعد الجزء التلافيفي، تقوم طبقة Flatten بتحويل السمات إلى متجه يُمرر إلى طبقتين Dense (256 ثم 128 وحدة) مع طبقات Dropout (بمعدل 0.5) بينهما. أخيراً، تحتوي طبقة المخرجات Dense على 5 وحدات مع دالة تنشيط softmax لإنتاج التوزيع الاحتمالي على الفئات الخمس. يحتوي النموذج على 1,441,669 معامل قابل للتدريب. يوضح الشكل 5 مخطط هذه البنية.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 4, 638, 16)        160

 max_pooling2d (MaxPooling2D  (None, 2, 319, 16)       0
 )

 dropout (Dropout)           (None, 2, 319, 16)        0

 conv2d_1 (Conv2D)           (None, 2, 319, 32)        4640

 max_pooling2d_1 (MaxPooling  (None, 1, 159, 32)       0
 2D)

 dropout_1 (Dropout)         (None, 1, 159, 32)        0

 conv2d_2 (Conv2D)           (None, 1, 159, 64)        18496

 max_pooling2d_2 (MaxPooling  (None, 1, 80, 64)        0
 2D)

 dropout_2 (Dropout)         (None, 1, 80, 64)         0

 conv2d_3 (Conv2D)           (None, 1, 80, 128)        73856

 max_pooling2d_3 (MaxPooling  (None, 1, 40, 128)       0
 2D)

 dropout_3 (Dropout)         (None, 1, 40, 128)        0

 flatten (Flatten)           (None, 5120)              0

 dense (Dense)               (None, 256)               1310976

 dropout_4 (Dropout)         (None, 256)               0

 dense_1 (Dense)             (None, 128)               32896

 dropout_5 (Dropout)         (None, 128)               0

 dense_2 (Dense)             (None, 5)                 645

=================================================================
Total params: 1,441,669
Trainable params: 1,441,669
Non-trainable params: 0
```

الشكل 5. مخطط تفصيلي لبنية نموذج CNN المستخدم

النتائج والمناقشة:

لتقييم أداء نموذج CNN المقترح على مجموعة التحقق، تم اعتماد عدة مقاييس كمية. تم حساب الدقة الإجمالية (Overall Accuracy) كنسبة التصنيفات الصحيحة إلى إجمالي العينات. كما تم تتبع خسارة التحقق (Validation Loss) (categorical_crossentropy) لتقدير مدى خطأ النموذج. تم تحليل منحنيات التعلم (الدقة والخسارة مقابل الحقب) لمراقبة سلوك التدريب. تم إنشاء مصفوفة الالتباس (Confusion Matrix) لتوضيح أداء التصنيف لكل فئة، ومنها تم استخلاص قيم التصنيفات الإيجابية الصحيحة (TP)، والسلبية الصحيحة (TN)، والإيجابية الخاطئة (FP)، والسلبية الخاطئة (FN) لكل فئة. بناءً على هذه القيم، تم حساب المقاييس التالية لكل فئة:

- الدقة (Precision):

$$Precision = TP/(TP + FP) \quad (1)$$

- الاستدعاء (Recall):

$$Recall = TP/(TP + FN) \quad (2)$$

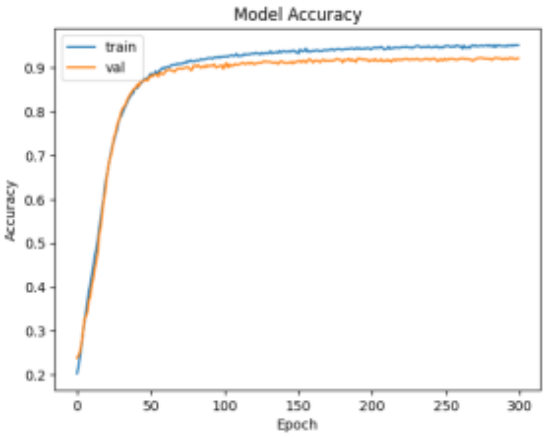
- مقياس F1 (F1-Score):

$$F1_Score = 2 * (Precision * Recall)/(Precision + Recall) \quad (3)$$

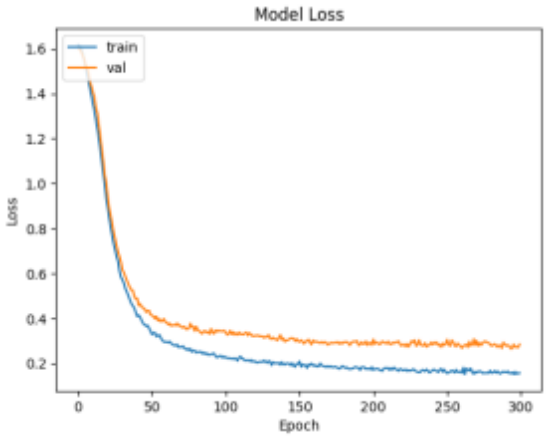
1. أداء التصنيف الإجمالي ومنحنيات التعلم:

بعد تدريب نموذج CNN المقترح لمدة 300 حقبة زمنية (epoch) على مجموعة بيانات التدريب المُعززة وضمن بروتوكول التقييم المختلط/المعتمد على العينات، وباستخدام مجموعة التحقق لتقييم الأداء عند نهاية كل حقبة، حقق النموذج دقة تحقق (validation accuracy) إجمالية بلغت 91.55%. بلغت قيمة خسارة التحقق (validation loss) النهائية 0.2879.

يوضح الشكل 6 والشكل 7 منحنيات التعلم للنموذج. يُظهر منحنى دقة التدريب ارتفاعاً مستمراً ليصل إلى ما يقارب 95%، بينما يرتفع منحنى دقة التحقق بشكل ملحوظ في الحقب الأولى ثم يستقر تدريجياً حول 91.55%. بالمثل، تنخفض خسارة التدريب بشكل مطرد، بينما تنخفض خسارة التحقق بشكل كبير في البداية ثم تتقارب إلى قيمة منخفضة نسبياً (حوالي 0.28). يشير التقارب الجيد بين أداء التدريب والتحقق، مع وجود فجوة طفيفة، إلى أن النموذج قد تعلم بشكل فعال وأن استخدام طبقات التسرب (Dropout) ساهم في التخفيف من فرط التلبيق بشكل كبير مع ضرورة عدم تفسير هذه النتيجة بوصفها تعميماً مباشراً على مشاركين غير مرئيين.



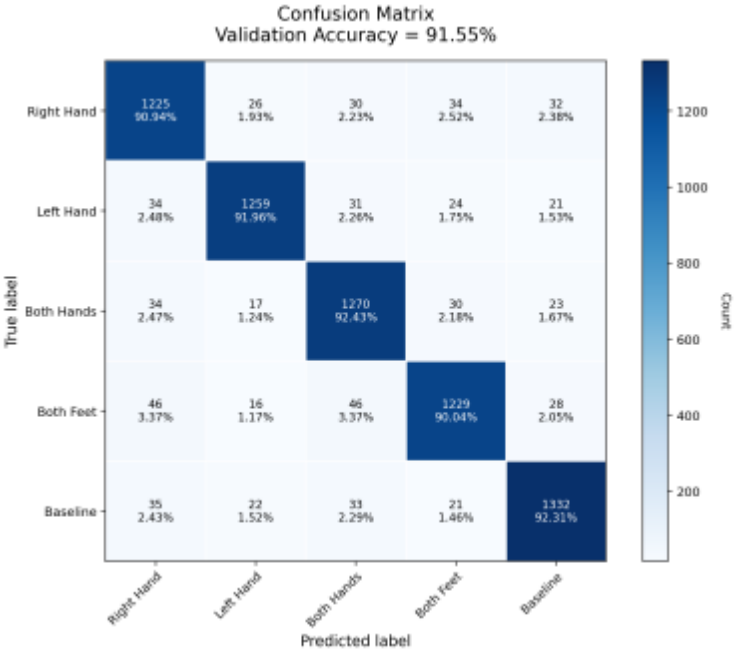
الشكل 6. منحنيات التعلم – الدقة مقابل الحقب



الشكل 7. منحنيات التعلم – الخسارة مقابل الحقب

2. تحليل مصفوفة الالتباس ومقاييس الأداء لكل فئة:

تم حساب مصفوفة الالتباس على مجموعة التحقق لتقديم رؤية مفصلة لأداء التصنيف عبر الفئات الخمس (تخيل اليد اليمنى، تخيل اليد اليسرى، تخيل كلتا اليدين، تخيل كلا القدمين، وحالة الراحة/خط الأساس). يوضح الشكل 8 مصفوفة الالتباس الناتجة.



الشكل 8 مصفوفة الالتباس لنتائج التصنيف على مجموعة التحقق (5 فئات)

من مصفوفة الالتباس، تم استخلاص قيم TP, TN, FP, FN لكل فئة، ومن ثم تم حساب مقاييس الدقة (Precision)، الاستدعاء (Recall)، ومقياس F1 (F1-Score). يلخص الجدول 2 هذه المقاييس لكل فئة.

جدول 2. مقاييس الأداء التفصيلية لكل فئة على مجموعة التحقق

Class	Precision	Recall	F1-Score
Right Hand (R)	**0.8916**	**0.9094**	**0.9004**
Left Hand (L)	**0.9396**	**0.9196**	**0.9295**
Both Hands (LR)	**0.9007**	**0.9243**	**0.9124**
Both Feet (F)	**0.9185**	**0.9004**	**0.9094**
Baseline (B)	**0.9276**	**0.9231**	**0.9253**

تُظهر القيم في الجدول 2 أداءً مستقراً عبر الفئات الخمس ضمن بروتوكول التقييم المختلط، إذ تراوحت قيم F1-score بين 0.9004 و0.9295، وبلغت الدقة الكلية 91.55%. ومع ذلك لا تُستخدم هذه النتيجة وحدها لإثبات التعميم على مشاركين جدد، لذلك أُضيف تقييم عابر للمشاركين.

3. التقييم العابر للمشاركين والمقارنة مع EEGNet:

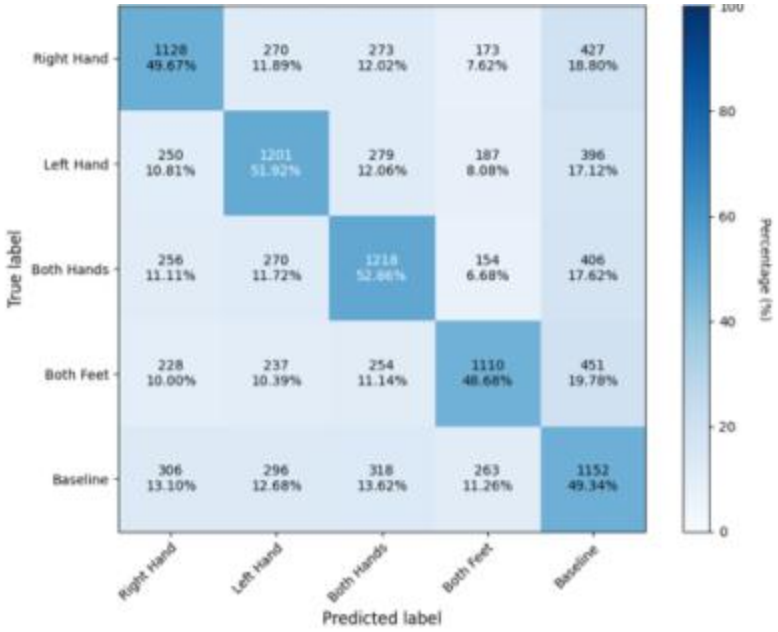
أُعيد تقييم النموذج باستخدام بروتوكول subject-wise 5-fold cross-validation، بحيث لا تظهر بيانات أي مشارك في مجموعتي التدريب والتحقق ضمن الطية نفسها. كما تم تطبيق اختيار القنوات وتوسعة البيانات بعد تقسيم المشاركين داخل كل طية، وذلك لمنع تسرب التمثيلات المشتقة من الحِقَب الأصلية بين المجموعتين. يوضح الجدول 3 نتائج الطيات الخمس للنموذج المقترح.

جدول 3. نتائج طيات التقييم العابر للمشاركين للنموذج المقترح.

البند	القيمة (%)
Fold 1	56.26
Fold 2	46.13
Fold 3	46.65
Fold 4	46.06
Fold 5	57.50
المتوسط	50.52
الانحراف المعياري SD	5.82
CI %95	5.10

تُظهر نتائج الجدول 3 أن أداء النموذج المقترح تراوح بين 46.06% و57.50% عبر الطيات الخمس، بمتوسط دقة قدره 50.52% وانحراف معياري 5.82%. يشير هذا التباين بين الطيات إلى وجود اختلاف واضح في صعوبة تصنيف بيانات المشاركين غير المرئيين، وهو أمر متوقع في إشارات EEG بسبب التباين الكبير بين الأفراد. وعلى الرغم من أن الدقة أقل بكثير من نتيجة التقييم المختلط، فإنها تبقى أعلى من مستوى التخمين العشوائي لخمس فئات، والذي يبلغ تقريباً 20%.

يوضح الشكل 9 مصفوفة الالتباس المجمعة للنموذج المقترح ضمن بروتوكول التقييم العابر للمشاركين، وذلك بهدف تحليل أنماط التصنيف الخاطئ بين الفئات الخمس.



الشكل 9. مصفوفة الالتباس للنموذج المقترح وفق بروتوكول subject-wise evaluation.

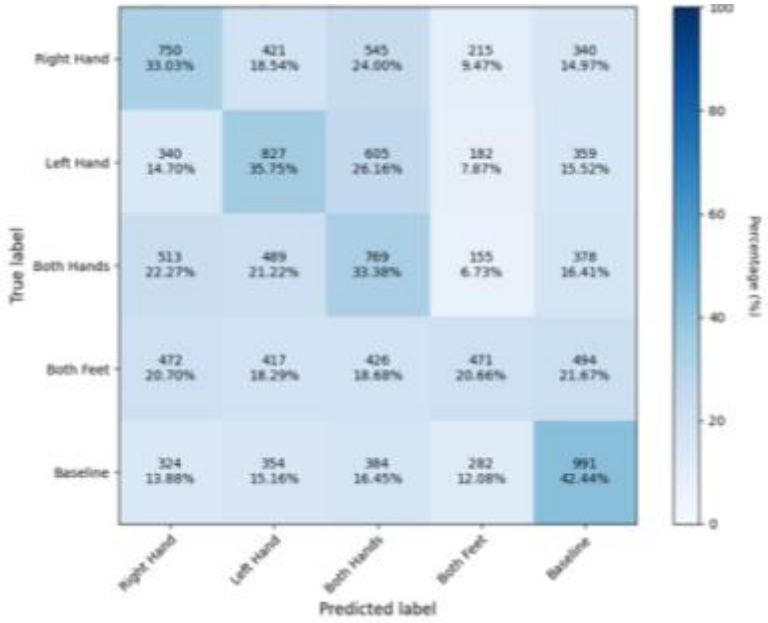
تبيّن مصفوفة الالتباس أن النموذج احتفظ بقدرة تمييزية مقبولة بين الفئات، إلا أن بعض فئات تخيل الحركة أظهرت تداخلاً أكبر فيما بينها مقارنة بنتائج التقييم المختلط. ويعكس ذلك صعوبة تعلم سمات ثابتة عبر المشاركين، خصوصاً عندما تختلف شدة ونمط استجابة EEG المرتبطة بتخيل الحركة من مشارك إلى آخر. ولإعطاء مقارنة مرجعية عادلة، تم تدريب نموذج EEGNet وتقييمه على مجموعة البيانات نفسها وباستخدام بروتوكول subject-wise 5-fold cross-validation نفسه. يعرض الجدول 4 مقارنة بين النموذج المقترح ونموذج EEGNet من حيث دقة التصنيف و macro-F1 على مستوى الحِقبة الأصلية.

جدول 4. مقارنة النموذج المقترح مع EEGNet ضمن بروتوكول التقييم نفسه.

النموذج	بروتوكول التقييم	Trial accuracy mean (%)	Trial accuracy SD (%)	Trial macro-F1 mean (%)	Trial macro-F1 SD (%)	CI %95 (%)
CNN المقترح	Subject-wise 5-fold CV	50.52	5.82	50.65	5.83	5.10
EEGNet	Subject-wise 5-fold CV	33.13	1.22	32.38	1.56	1.07

توضح نتائج الجدول 4 أن النموذج المقترح حقق دقة 50.52% ± 5.82% وقيمة Trial macro-F1 مقدارها 50.65% ± 5.83%، في حين حقق EEGNet دقة 33.13% ± 1.22% وقيمة Trial macro-F1 مقدارها 32.38% ± 1.56%. وبذلك تفوق النموذج المقترح على EEGNet ضمن الشروط التجريبية نفسها، مما يشير إلى أن البنية المقترحة واستراتيجية التمثيل عبر مجموعات القنوات الثلاث كانتا أكثر ملاءمة لهذه البيانات مقارنة بالبنية المرجعية المستخدمة.

يوضح الشكل 10 مصفوفة الالتباس الخاصة بنموذج EEGNet ضمن بروتوكول التقييم العابر للمشاركين نفسه، مما يسمح بمقارنة نمط الأخطاء بين النموذجين.



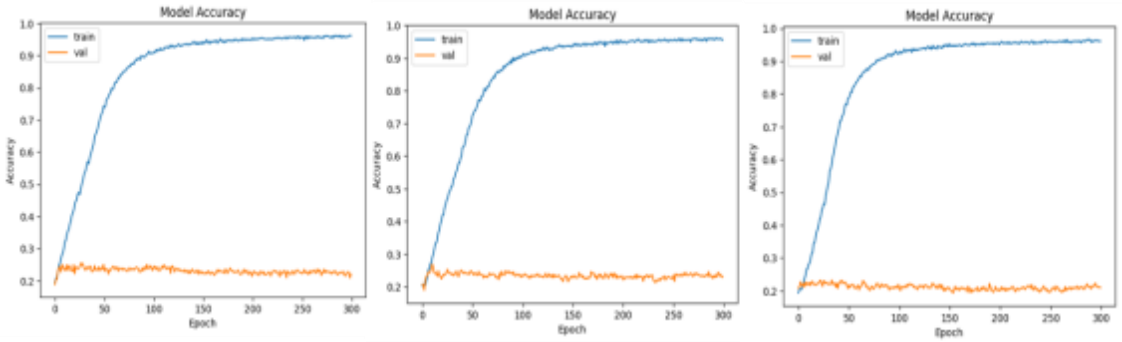
الشكل 10. مصفوفة الالتباس لنموذج EEGNet المرجعي ضمن بروتوكول subject-wise 5-fold CV.

بالمقارنة مع مصفوفة الالتباس للنموذج المقترح، تُظهر نتائج EEGNet تداخلاً أكبر بين الفئات وانخفاضاً أوضح في القدرة على الفصل بينها. تؤكد هذه النتيجة أن التعميم على مشاركين غير مرئيين يمثل تحدياً أساسياً في تصنيف إشارات MI-EEG، لكنها في الوقت نفسه تبيّن أن النموذج المقترح حقق أداءً أفضل من النموذج المرجعي عند تثبيت مجموعة البيانات وبروتوكول التقييم.

4. تقييم تأثير استراتيجية توسعة البيانات وأداء مجموعات القنوات الفردية:

لتقييم أثر استراتيجية توسعة البيانات المقترحة، أُجريت تجارب إضافية باستخدام كل مجموعة قنوات بشكل منفرد: FC، وC، وCP. في كل تجربة، استُخدمت البنية نفسها والمعاملات الفائقة نفسها للنموذج الرئيسي، لكن دون دمج مجموعات القنوات الثلاث، بحيث مثّلت كل حِقبة أصلية بمجموعة قنوات واحدة فقط.

توضح منحنيات التعلم في الشكل 11 أن النماذج المدربة على مجموعات القنوات الفردية عانت من فرط تخصيص واضح؛ إذ تجاوزت دقة التدريب 95% في معظم الحالات، بينما بقيت دقة التحقق منخفضة بين 21.7% و24.3%. فقد حققت مجموعة FC دقة تحقق 22.5%، ومجموعة C دقة 24.3%، ومجموعة CP دقة 21.7%. تشير هذه النتائج إلى أن استخدام مجموعة قنوات واحدة لم يكن كافياً لاستخلاص تمثيلات قابلة للتعميم، مقارنةً باستخدام التمثيلات المكانية المتعددة الناتجة عن دمج المجموعات الثلاث.



(أ) مجموعة القنوات الأمامية–المركزية (FC) (ب) مجموعة القنوات المركزية (C) (ج) مجموعة القنوات المركزية–الجدارية (CP)

الشكل 11. منحنيات تعلم الدقة للنماذج المدربة على مجموعات قنوات فردية دون توسعة بيانات

بناءً على ذلك، دعمت استراتيجية التوسعة عبر المجموعات الثلاث تحسين الأداء ضمن التقييم المختلط، بينما أظهر التقييم العابر للمشاركين أن التعميم على مشاركين غير مرئيين لا يزال تحدياً أساسياً.

5. مقارنة مع الأعمال السابقة ومناقشة النتائج:

تضع النتائج المعروضة في الجدول 5 أداء النموذج المقترح ضمن السياق العام للدراسات السابقة في تصنيف إشارات MI-EEG. ويجدر بالذكر أن الدراسات تختلف في مجموعة البيانات وعدد المشاركين وعدد الفئات وعدد القنوات ومعدل أخذ العينات وبروتوكول التقييم. لذلك فإن الجدول لا يهدف إلى إثبات تفوق مباشر لنموذج على آخر بل إلى توضيح موقع أداء النموذج المقترح مقارنة بالطرائق المنشورة ضمن شروط تجريبية مختلفة.

ضمن بروتوكول التقييم المختلط، حقق النموذج المقترح دقة 91.55% لخمس فئات باستخدام 6 قنوات فقط لكل عينة من مجموعة PhysioNet EEG Motor Movement/Imagery، التي تضم 103 مشاركين بعد الاستبعاد وبمعدل أخذ عينات 160 هرتز. وبالمقارنة مع الدراسات التي استخدمت مجموعة PhysioNet نفسها، حققت الدراستان [6] و[7] دقة 83.8% و87.8% على الترتيب لمهام التخيل الحركي. وبذلك جاءت دقة النموذج المقترح أعلى بمقدار 7.75% مقارنة بالدراسة [6]، وأعلى بمقدار 3.75% مقارنة بالدراسة [7]. ومع ذلك، تبقى هذه المقارنة سياقية وليست مباشرة تماماً، لأن الدراستين [6] و[7] اعتمدتا تصنيفاً ثنائياً واستخدمتا 64 قناة، في حين اعتمدت الدراسة الحالية خمس فئات واستخدمت 6 قنوات فقط لكل عينة، أي أقل بنسبة تقارب 90.6% من عدد القنوات المستخدمة في تلك الدراسات. كما أن بروتوكول التقييم في هذه الدراسات أقرب إلى التقييم المختلط أو غير مفصول صراحة على مستوى المشاركين، لذلك لا يمكن اعتباره مكافئاً للتقييم العابر للمشاركين المستخدم في هذه الدراسة.

جدول 5. مقارنة أداء النموذج المقترح مع دراسات سابقة على مجموعات بيانات MI-EEG

Study	Model Architecture	Dataset	Classes	Ch. (Input)	Accuracy (%)
الدراسة المقترحة (subject-mixed)	CNN	PhysioNet	5	6	91.55
[6]	EEGNet Fusion	PhysioNet	2	64	83.8 (MI)
[7]	EEGNet Fusion V2	PhysioNet	2	64	87.8 (MI)
[8]	CNN + FRM/MSF	PhysioNet	4	64	95.05
[9]	MS-CNN + Feat. Integ. + Aug.	BCI Comp. IV 2b	2	3	93.74
[5]	CMO-CNN (1D, Multi-Br) + Aug.	BCI Comp. IV 2a	4	22	83.92
	CMO-CNN (1D, Multi-Br) + Aug.	BCI Comp. IV 2b	2	3	87.19
[11]	CNN (2D) + SSAE	Private (D1)	2	Grouped	84.7
	CNN (2D) + SSAE	Public (D2)	2	Grouped	86.4
[10]	LM-CNN + Empirical Model + Feat.	Private (Motion)	4	8	94.04
[13]	CNN (1D) + Attention + Aug. (GAN)	BCI Comp. IV 2a	4	22	93.6
	CNN (1D) + Attention + Aug. (SW)	BCI Comp. IV 2b	2	3	87.83
[14]	ResNet18	Private (Cognitive)	3	19	72
[12]	Custom CNN (Temp+Spat) + Aug.	BCI Comp. IV 2a	2	22	87.34
	Custom CNN (Temp+Spat) + Aug.	BCI Comp. IV 2b	2	22	86.29

في دراسة [8] فقد استخدمت أيضاً مجموعة PhysioNet وحققت دقة 95.05% لأربع فئات باستخدام 64 قناة ونموذج يعتمد على دمج سمات متعددة. وتزيد هذه النتيجة على دقة النموذج المقترح بمقدار 3.50%، إلا أن المقارنة المباشرة معها تبقى محدودة بسبب اختلاف عدد الفئات وعدد القنوات وطريقة بناء السمات وعدم التصريح الواضح بأن الطيات كانت مفصولة على مستوى المشاركين. وبذلك تُظهر هذه الدراسة أن استخدام عدد أكبر من القنوات وطرائق دمج سمات متقدمة قد يرفع الدقة لكنه يزيد أيضاً من تعقيد النموذج مقارنة بالنهج الحالي القائم على عدد محدود من القنوات ومعالجة أولية أبسط.

بالنسبة الدراسات التي استخدمت مجموعات بيانات BCI Competition IV 2a و2b، مثل [5] و[9] و[12] و[13]، فقد حققت نتائج جيدة باستخدام بنى متعددة الفروع، أو آليات انتباه، أو دمج سمات، أو تقنيات توسعة بيانات مختلفة. غير أن هذه المجموعات تختلف عن PhysioNet من حيث عدد المشاركين وعدد القنوات ومعدل أخذ العينات وعدد الفئات، كما أنها تضم عادة عدداً أقل من المشاركين مقارنة بمجموعة PhysioNet المستخدمة في هذه الدراسة. لذلك لا يمكن اعتبار الفروق العددية في الدقة مقارنة مباشرة، بل مؤشراً عاماً على موقع النموذج ضمن الدراسات السابقة.

ولتعزيز المقارنة تم تدريب نموذج EEGNet المرجعي على مجموعة البيانات نفسها وباستخدام بروتوكول subject-wise 5-fold cross-validation نفسه. ضمن هذا التقييم الصارم، حقق EEGNet دقة 33.13% ± 1.22% وقيمة Trial macro-F1 مقدارها 32.38% ± 1.56%، بينما حقق النموذج المقترح دقة 50.52% ± 5.82% وقيمة Trial macro-F1 مقدارها 50.65% ± 5.83%. وهذا يمثل زيادة مقدارها 17.39% في الدقة مقارنة بـ EEGNet، أي تحسناً نسبياً يقارب 52.5% وفق البروتوكول نفسه. كما حقق النموذج المقترح زيادة مقدارها 18.27% في Trial macro-F1، أي تحسناً نسبياً يقارب 56.4%. تشير هذه النتائج إلى أن البنية

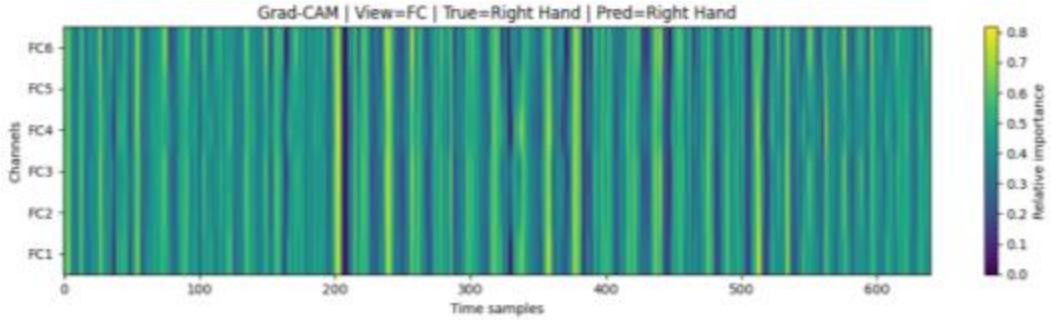
المقترحة كانت أكثر ملاءمة من EEGNet ضمن شروط هذه الدراسة واستراتيجية التوسعة المعتمدة على مجموعات القنوات الثلاث.

بناءً على ذلك، يحقق النموذج المقترح أداءً مناسباً ضمن التقييم المختلط مقارنة بالدراسات ذات البروتوكولات المشابهة أو غير المفصولة بوضوح على مستوى المشاركين، مع استخدام عدد أقل من القنوات. وفي المقابل، يوضح التقييم العابر للمشاركين أن التعميم على مشاركين غير مرئيين يبقى تحدياً أساسياً في تصنيف إشارات MI-EEG، ولا سيما عند استخدام مجموعة بيانات كبيرة نسبياً ومتنوعة بين الأفراد. لذلك تبرز الحاجة إلى تطوير نماذج أكثر قدرة على تعلم سمات مستقرة بين المشاركين، مع الحفاظ على بساطة المعالجة الأولية وتقليل عدد القنوات المطلوبة للتطبيق العملي.

6. تحليل قابلية التفسير باستخدام Grad-CAM:

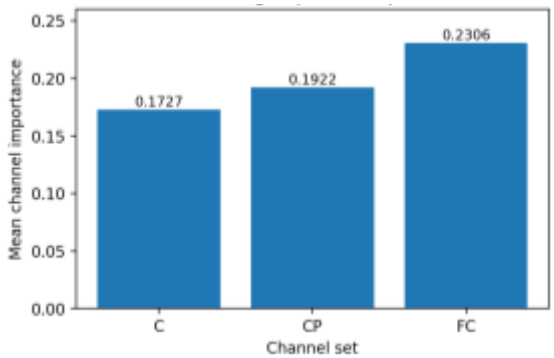
أُضيف تحليل تفسيري باستخدام Grad-CAM بهدف تحديد المناطق الزمنية–القنوية التي أسهمت في قرارات التصنيف داخل النموذج المقترح. استُخدم هذا التحليل على عينات مصنفة بشكل صحيح، ثم حُسبت أهمية القنوات ضمن مجموعات القنوات الثلاث المستخدمة في استراتيجية التوسعة، وهي: المجموعة الأمامية–المركزية FC، والمجموعة المركزية C، والمجموعة المركزية–الجدارية CP.

يوضح الشكل 12 مثالاً لخريطة Grad-CAM لعينة من فئة Right Hand صُنفت بشكل صحيح ضمن مجموعة القنوات FC. تمثل الألوان في الخريطة درجة الأهمية النسبية عبر القنوات والزمن، حيث تشير القيم الأعلى إلى مناطق كان لها تأثير أكبر في قرار النموذج بالنسبة لهذه العينة.



الشكل 12. خريطة Grad-CAM لعينة مصنفة بشكل صحيح من فئة Right Hand ضمن مجموعة القنوات FC.

وللانتقال من تفسير عينة مفردة إلى مؤشر تجميعي على مستوى مجموعات القنوات، تم حساب متوسط أهمية كل مجموعة من مجموعات القنوات الثلاث.



الشكل 13. متوسط أهمية مجموعات القنوات حسب Grad-CAM.

يوضح الشكل 13 أن المجموعة الأمامية–المركزية FC امتلكت أعلى متوسط أهمية (0.2306)، تلتها المجموعة المركزية–الجدارية CP بمتوسط (0.1922)، ثم المجموعة المركزية C بمتوسط (0.1727).

تشير هذه النتائج إلى أن ترتيب أهمية مجموعات القنوات في التحليل المدروس كان FC > CP > C. ويمكن تفسير ذلك بأن القنوات الأمامية–المركزية كانت الأكثر مساهمة في قرار التصنيف ضمن هذا التحليل، مع وجود مساهمة واضحة أيضاً للمناطق المركزية–الجدارية والمركزية.

الاستنتاجات والتوصيات:

تُظهر النتائج أن بنية CNN المقترحة حققت دقة 91.55% ضمن التقييم المختلط/المعتمد على العينات بعد اعتماد subject-wise channel z-score normalization، بينما بلغت الدقة في التقييم العابر للمشاركين 50.52% ± 5.82%. وبالتالي فإن التعميم على مشاركين جدد يبقى تحدياً أساسياً في تصنيف إشارات MI-EEG.

الاستنتاجات:

1. أثبت النموذج المقترح قدرته على تصنيف إشارات MI-EEG ضمن التقييم المختلط بدقة تحقق 91.55% لخمس فئات، إلا أن هذه النتيجة لا تُفسر بوصفها دليلاً مباشراً على التعميم العابر للمشاركين.
2. أظهرت استراتيجية توسعة البيانات عبر ثلاث مجموعات قنوات سداسية أنها توفر تمثيلات مكانية متعددة للإشارة، وقد دعمت نتائج Grad-CAM أن هذه المجموعات تسهم بدرجات مختلفة في قرار التصنيف.
3. أظهر التقييم العابر للمشاركين subject-wise 5-fold cross-validation دقة 50.52% ± 5.82%، ما يبرز صعوبة التعميم على مشاركين جدد ويؤكد ضرورة استخدام هذا النوع من التقييم في الدراسات المستقبلية.
4. تفوق النموذج المقترح على EEGNet المرجعي عند تدريبهما وتقييمهما على مجموعة البيانات نفسها وببروتوكول التقييم العابر للمشاركين نفسه، إذ بلغت دقة EEGNet مقدار 33.13% ± 1.22%.
5. تؤكد الدراسة أن بنى CNN قادرة على استخلاص سمات تمييزية من إشارات EEG ذات المعالجة الأولية البسيطة، لكنها تحتاج إلى تطويرات إضافية لتحسين ثبات السمات والتعميم بين المشاركين.

التوصيات:

بناءً على استنتاجات هذا العمل والآفاق المستقبلية، نقترح التوصيات التالية:

1. تطوير بنى هجينة تجمع بين CNN والمرمزات الذاتية Autoencoders، بهدف الاستفادة من التعلم غير الخاضع للإشراف لاستخراج تمثيلات مدمجة من إشارات EEG، ولا سيما السمات الناتجة عن طبقة العنق Bottleneck layer، والتي يمكن استخدامها كمدخلات غنية ومضغوطة قبل مرحلة التصنيف.
2. دراسة إضافة طبقات انتباه مكانية–زمانية spatio-temporal attention layers إلى البنية المقترحة، بحيث يتمكن النموذج من إعطاء وزن أكبر للقنوات والفترات الزمنية الأكثر ارتباطاً بمهام تخيل الحركة.
3. إجراء تجارب أوسع للتقييم العابر للمشاركين باستخدام بروتوكولات تحقق خارجية أو عدداً أكبر من الطيات، مع مقارنة إضافية مع بنى زمنية متقدمة مثل TCN.

4. تبسيط النموذج وتقليل تعقيده الحسابي، مع الحفاظ على بروتوكولات تقييم صارمة على مستوى المشاركين، بهدف تسهيل دمجه في أنظمة BCI تعمل في الوقت الفعلي.

References:

[1]X. Wang, V. Liesaputra, Z. Liu, Y. Wang, and Z. Huang, "An in-depth survey on Deep Learning-based Motor Imagery Electroencephalogram (EEG) classification," *Artif. Intell. Med.*, vol. 147, p. 102738, Jan. 2024, doi: 10.1016/J.ARTMED.2023.102738.

[2]J. S. Kumar and P. Bhuvaneswari, "Analysis of Electroencephalography (EEG) Signals and Its Categorization–A Study," *Procedia Eng.*, vol. 38, pp. 2525–2536, 2012, doi: 10.1016/j.proeng.2012.06.298.

[3]S. Sanei and J. A. Chambers, "Fundamentals of EEG Signal Processing," *EEG Signal Processing and Machine Learning*, pp. 77–113, Sep. 2021, doi: 10.1002/9781119386957.CH4.

[4]R. T. Schirrmeister *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapp.*, vol. 38, no. 11, p. 5391, Nov. 2017, doi: 10.1002/HBM.23730.

[5]X. Liu, S. Xiong, X. Wang, T. Liang, H. Wang, and X. Liu, "A compact multi-branch 1D convolutional neural network for EEG-based motor imagery classification," *Biomed. Signal Process. Control*, vol. 81, p. 104456, Mar. 2023, doi: 10.1016/J.BSPC.2022.104456.

[6]K. Roots, Y. Muhammad, and N. Muhammad, "Fusion Convolutional Neural Network for Cross-Subject EEG Motor Imagery Classification," *Computers 2020, Vol. 9, Page 72*, vol. 9, no. 3, p. 72, Sep. 2020, doi: 10.3390/COMPUTERS9030072.

[7]R. R. Chowdhury, Y. Muhammad, and U. Adeel, "Enhancing Cross-Subject Motor Imagery Classification in EEG-Based Brain–Computer Interfaces by Using Multi-Branch CNN," *Sensors 2023, Vol. 23, Page 7908*, vol. 23, no. 18, p. 7908, Sep. 2023, doi: 10.3390/S23187908.

[8]T. Lotey, P. Keserwani, D. P. Dogra, and P. P. Roy, "Feature reweighting for EEG-based motor imagery classification," *Biomed. Signal Process. Control*, vol. 110, Dec. 2025, doi: 10.1016/j.bspc.2025.108215.

[9]A. M. Roy, "A CNN model with feature integration for MI EEG subject classification in BMI," *bioRxiv*, p. 2022.01.05.475058, Feb. 2022, doi: 10.1101/2022.01.05.475058.

[10] R. K. Megalingam, K. S. Sankardas, and S. K. Manoharan, "An Empirical Model-Based Algorithm for Removing Motion-Caused Artifacts in Motor Imagery EEG Data for Classification Using an Optimized CNN Model," *Sensors*, vol. 24, no. 23, Dec. 2024, doi: 10.3390/S24237690,.

[11] J. Yang, Z. Ma, J. Wang, and Y. Fu, "A Novel Deep Learning Scheme for Motor Imagery EEG Decoding Based on Spatial Representation Fusion," *IEEE Access*, vol. 8, pp. 202100–202110, 2020, doi: 10.1109/ACCESS.2020.3035347.

[12] O. Tarahi, S. Hamou, M. Moufassih, S. Agounad, and H. I. Azami, "Decoding brain signals: A convolutional neural network approach for motor imagery classification," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 7, p. 100451, Mar. 2024, doi: 10.1016/J.PRIME.2024.100451.

[13] E. Lashgari, J. Ott, A. Connelly, P. Baldi, and U. Maoz, "An end-to-end CNN with attentional mechanism applied to raw EEG in a BCI classification task," *J. Neural Eng.*, vol. 18, no. 4, p. 0460e3, Aug. 2021, doi: 10.1088/1741-2552/AC1ADE.

[14] M. O. Petrov, Z. V. Nagornova, N. V. Andreeva, and N. V. Shemyakina, "Classification of Raw Single-Trial EEG/ERP Signals of a Combined Subjects' Data Sample in Response to the Presentation of Mathematical Sums, Correct and Incorrect Answers to them, Using CNNs," *Proceedings of 2024 5th International Conference on Neural Networks and Neurotechnologies, NeuroNT 2024*, pp. 128–131, 2024, doi: 10.1109/NEURONT62606.2024.10585340.

[15] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: a compact convolutional neural network for EEG-based brain-computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Jul. 2018, doi: 10.1088/1741-2552/AACE8C.

[16] "MNE — MNE 1.9.0 documentation." Accessed: May 09, 2025. [Online]. Available: https://mne.tools/stable/index.html

[17] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.," *Circulation*, vol. 101, no. 23, 2000, doi: 10.1161/01.CIR.101.23.E215,.

[18] M. Mustile, D. Kourtis, M. G. Edwards, D. I. Donaldson, and M. Ietswaart, "Neural correlates of motor imagery and execution in real-world dynamic behavior: evidence for similarities and differences," *Front. Hum. Neurosci.*, vol. 18, p. 1412307, 2024, doi: 10.3389/FNHUM.2024.1412307.