

Легковесный и объяснимый фреймворк ЭЭГ для детекции внимания в задачах быстрого последовательного визуального предъявления

Кевин А. Родригес Санчес
СПбГЭТУ «ЛЭТИ»
rodriguez@internet.ru

Пэй Ютин
СПбГЭТУ «ЛЭТИ»
peiyuting597@gmail.com

Эммануэль Элочукву
СПбГЭТУ «ЛЭТИ»
kayceelochukwu10@gmail.com

Надим Рехмат
СПбГЭТУ «ЛЭТИ»
nadeemassi169@gmail.com

Рагхавендра Метан
СПбГЭТУ «ЛЭТИ»
metanrv@gmail.com

Анагхим Ибрахим
СПбГЭТУ «ЛЭТИ»
ianagheem@gmail.com

Аннотация. Парадигмы быстрого последовательного визуального предъявления (RSVP) широко используются для изучения зрительного внимания и детекции целевых стимулов в электроэнцефалографии (ЭЭГ). Однако многие исследования по декодированию ЭЭГ в основном сосредоточены на предсказательной эффективности, пренебрегая нейрофизиологической интерпретируемостью либо полагаясь на вычислительно затратные модели. В данной работе предлагается лёгковесный и интерпретируемый ЭЭГ-фреймворк для детекции целевых стимулов в RSVP-задачах, основанный на динамике задних компонентов вызванных потенциалов (ERP). Предложенный конвейер обработки данных включает минимальную предобработку, разреженный набор задних электродов, признаки на основе средних амплитуд ERP во временных окнах, а также классические алгоритмы машинного обучения. Эксперименты проведены на открытом датасете PhysioNet RSVP EEG, содержащем 62 записи от 11 испытуемых и 20 534 пригодные для анализа эпохи. Лучшая модель по метрике precision-recall — SVM с радиально-базисной функцией и взвешиванием классов — достигла сбалансированной точности 0.6729 и PR AUC 0.3467, превзойдя LDA при оценке с учётом дисбаланса классов. Анализ показал, что наиболее информативная активность наблюдается в интервале 400–500 мс в задних каналах PO4, PO7 и O1. Полученные результаты демонстрируют, что ERP-ориентированные и вычислительно эффективные конвейеры могут служить интерпретируемой базовой моделью для детекции внимания в RSVP-задачах в нейроинженерии и системах интерфейса мозг–компьютер (BCI).

Ключевые слова: ЭЭГ, RSVP, компоненты вызванных потенциалов, детекция внимания, интерфейс мозг–компьютер, нейроинженерия

I. ВВЕДЕНИЕ

Электроэнцефалография (ЭЭГ) остаётся одной из наиболее практичных модальностей для неинвазивных систем интерфейса мозг–компьютер (BCI) благодаря высокому временному разрешению, портативности и относительно низкой стоимости регистрации данных [2], [7]. Среди парадигм ЭЭГ метод быстрого последовательного визуального предъявления (RSVP) приобрёл особую значимость для изучения зрительного внимания и быстрой детекции целевых стимулов в условиях ограниченного времени [1], [9]. В экспери-

ментах RSVP визуальные стимулы предъявляются быстро последовательно, и испытуемые должны обнаруживать редкие целевые элементы, встроенные среди частых нецелевых. Эти события обычно вызывают компоненты вызванных потенциалов (ERP) в задних областях коры головы, особенно в интервале 300–600 мс [2], [4].

BCI-системы на основе RSVP широко исследовались для задач визуального поиска и быстрой сортировки изображений. Предыдущие исследования показали, что ЭЭГ-ответы в RSVP-задачах могут поддерживать автоматическую локализацию целевых стимулов, а коллективные подходы BCI могут дополнительно улучшать точность детекции [1]. С нейрофизиологической точки зрения декодирование RSVP в большей степени связано с транзиторной динамикой ERP, а не с устойчивой спектральной активностью, что делает признаки, ориентированные на ERP, особенно подходящими для интерпретируемого анализа одиночных проб [4].

Однако многие исследования по декодированию RSVP уделяют приоритетное внимание исключительно точности классификации, часто используя сложные архитектуры глубокого обучения, такие как сверточные или трансформерные модели [8], [13]. Хотя эти подходы способны захватывать богатые пространственно-временные паттерны, они нередко снижают физиологическую интерпретируемость и увеличивают вычислительную сложность. Для задач нейроинженерии лёгковесные конвейеры, сохраняющие прозрачность, воспроизводимость и вычислительную эффективность, остаются крайне ценными [3], [11].

Ещё одной методологической проблемой является выраженный дисбаланс классов, присущий RSVP-парадигмам, поскольку целевые события по своей природе редки. В таких условиях простая точность становится вводящей в заблуждение и должна дополняться метриками, такими как сбалансированная точность и анализ precision-recall [10]. Кроме того, надёжное декодирование RSVP требует моделей, устойчивых к ограниченному числу целевых примеров и межсубъектной вариабельности.

Для решения этих задач в данной работе предлагается легковесный и объяснимый ЭЭГ-фреймворк для детекции внимания в RSVP-задачах. Предложенный конвейер использует разреженный набор задних электродов, минимальную предобработку, признаки на основе средних амплитуд ERP во временных окнах, а также классические алгоритмы машинного обучения. Вместо приоритета сложности модели, данный подход делает акцент на трёх ключевых свойствах: физиологической интерпретируемости, вычислительной эффективности и воспроизводимости.

Основные вклады работы включают: (1) интерпретируемый базовый подход на основе ERP-временных окон для детекции целевых стимулов RSVP с использованием разреженных задних каналов ЭЭГ; (2) оценку с учётом дисбаланса классов в условиях объединённых и субъектно-ориентированных протоколов; (3) нейрофизиологическую валидацию, связывающую выходы классификатора с динамикой задних ERP; и (4) анализ задержек, демонстрирующий вычислительную эффективность предложенного конвейера.

II. МАТЕРИАЛЫ И МЕТОДЫ

A. Набор данных и экспериментальная парадигма

Эксперименты были проведены с использованием открытого датасета PhysioNet EEG Signals from an RSVP Task [5], [6]. Обработанный датасет содержит 62 записи в формате EDF от 11 испытуемых, выполнявших задачу быстрой визуальной детекции целевых стимулов. Участникам предъявлялись аэрофотоснимки с частотой 5, 6 и 10 Гц, и им было необходимо обнаруживать изображения, содержащие самолёты, среди нецелевых стимулов [1].

Аннотации событий были приведены к задаче бинарной классификации (целевой / нецелевой). После предобработки и выделения эпох датасет содержал 20534 пригодные для анализа эпохи, включая 2041 целевую и 18493 нецелевые пробы, что приводит к сильно несбалансированному сценарию детекции.

B. Монтаж ЭЭГ и предобработка

Анализ был сосредоточен на восьми задних электродах: PO8, PO7, PO3, PO4, P7, P8, O1 и O2. Эти каналы отражают заднюю зрительную обработку и активность ERP, связанную с целевыми стимулами, при этом сохраняя лёгкую и интерпретируемую конфигурацию.

Предобработка была намеренно минимальной для сохранения вычислительной эффективности. Непрерывные сигналы ЭЭГ фильтровались полосовым фильтром в диапазоне 0.5–40 Гц для удаления медленного дрейфа и высокочастотного шума [3]. Эпохи, синхронизированные со стимулом, извлекались в интервале от 0 до 0.8 с относительно каждого события. Методы декомпозиции артефактов, такие как ICA, не применялись для сохранения лёгковесного конвейера обработки.

C. Извлечение признаков на основе ERP-временных окон

Поскольку детекция в RSVP в основном определяется транзиторными ERP-ответами, признаки формировались на основе средних амплитуд ERP во временных окнах. Для каждой эпохи и каждого канала

вычислялось среднее значение амплитуды в шести временных окнах: 0–100, 100–200, 200–300, 300–400, 400–500 и 500–600 мс. При использовании восьми каналов и шести временных окон каждая эпоха представляется 48 признаками. Такое представление сохраняет временную интерпретируемость, оставаясь при этом вычислительно более лёгким по сравнению с полноволновыми или глубокими моделями [4], [8].

D. Модели классификации

Были оценены два классических классификатора: линейный дискриминантный анализ (LDA) и метод опорных векторов с радиально-базисной функцией (RBF-SVM). Для уменьшения влияния дисбаланса классов в модели SVM использовалось взвешивание классов. Все признаки были стандартизированы с использованием статистик обучающей выборки.

E. Протокол оценки

Использовались две схемы оценки: (i) объединённое стратифицированное случайное разбиение и (ii) субъектно-ориентированное групповое разбиение. Из-за сильного дисбаланса классов производительность оценивалась с использованием следующих метрик: сбалансированная точность, precision, recall, F1-мера, PR AUC, ROC AUC, специфичность и коэффициент корреляции Мэттьюса (MCC) [10].

F. Нейрофизиологическая интерпретация и анализ задержек

Для соотнесения выходов классификатора с физиологией ЭЭГ проводился анализ, включающий: усреднённые ERP по задней области интереса (ROI), разностные волны (целевой минус нецелевой), разреженные топографические карты задней области (300–500 мс), карты важности признаков по каналам и времени, а также анализ precision–recall.

Анализ задержек включал измерение времени предобработки, сегментации эпох, извлечения признаков и инференса. Общее время обработки определяется как

$$L_{\text{total}} = L_{\text{pre}} + L_{\text{epoch}} + L_{\text{feat}} + L_{\text{inf}}. \quad (1)$$

III. РЕЗУЛЬТАТЫ

A. Производительность классификации

Обработанный датасет дал 20534 пригодные для анализа эпохи из 62 записей и 11 испытуемых, при этом наблюдается выраженный дисбаланс классов между целевыми и нецелевыми пробами. В условиях объединённой оценки модель RBF-SVM с взвешиванием классов достигла наибольшего значения PR AUC (0.3467), что делает её предпочтительной в условиях дисбаланса классов. LDA показал немного более высокую сбалансированную точность (0.6771), но более низкое значение PR AUC (0.2960), что указывает на менее информативный компромисс между precision и recall.

При субъектно-ориентированной групповой оценке производительность снизилась, как и ожидалось, из-за межсубъектной вариабельности. LDA достиг сбалансированной точности 0.6259 и PR AUC 0.1568, тогда как SVM с взвешиванием классов показал сбалансированную точность 0.6201 и PR AUC 0.2620.

Эти результаты указывают на то, что детекция целевых стимулов в RSVP остаётся сложной задачей в условиях независимости от субъекта, хотя признаки на основе ERP-временных окон по-прежнему отражают значимую структуру, связанную с целевыми событиями.

ТАБЛИЦА I. ПРОИЗВОДИТЕЛЬНОСТЬ КЛАССИФИКАЦИИ ПРИ ДИСБАЛАНСЕ КЛАССОВ

Split	Model	Bal. Acc.	Prec.	Recall	F1	PR AUC
Pooled	LDA	0.6771	0.2100	0.6054	0.3119	0.2960
Pooled	SVM	0.6729	0.2480	0.5196	0.3357	0.3467
Subject-aware	LDA	0.6259	0.1634	0.5775	0.2547	0.1568
Subject-aware	SVM	0.6201	0.2276	0.3838	0.2857	0.2620

В. Динамика ERP

Рис. 1 показывает усреднённый ERP по задней области интереса (ROI). Ответы на целевые и нецелевые стимулы начинают расходиться примерно после 250 мс, при этом наибольшее различие наблюдается в интервале 300–500 мс. Этот паттерн согласуется с активностью ERP, связанной с целевыми стимулами в RSVP-парадигмах.

Разностная волна на рис. 2 достигает пика примерно на 400 мс с амплитудой -2.30 мкВ, что подтверждает, что наиболее информативный временной интервал находится в позднем ERP-окне.

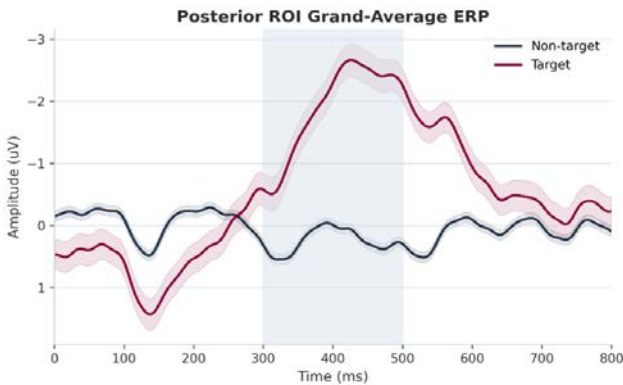


Рис. 1. Усреднённый ERP по задней области интереса (ROI), показывающий наибольшее различие между целевыми и нецелевыми стимулами в интервале 300–500 мс.

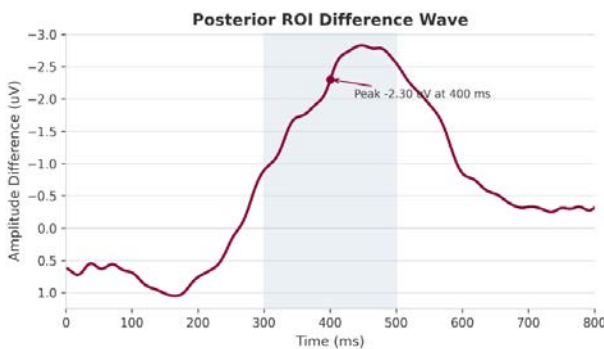


Рис. 2. Разностная волна (целевой минус нецелевой) с максимальным экстремумом около 400 мс.

С. Пространственная и признаковая значимость

Разреженная топографическая карта задней области (рис. 3), усреднённая в интервале 300–500 мс, показывает, что дискриминирующая активность сосредоточена в задних областях коры головы.

Target Minus Non-target (300-500 ms)

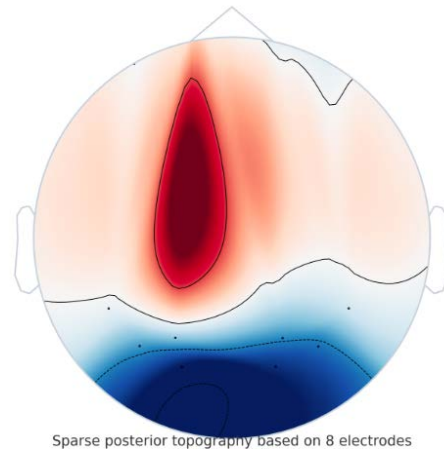


Рис. 3. Топографическая карта задней области разности ERP (целевой минус нецелевой) в интервале 300–500 мс.

Тепловая карта важности признаков по каналам и времени (рис. 4) дополнительно показывает, что наиболее информативные признаки наблюдаются в поздних ERP-окнах, особенно в интервале 400–500 мс, преимущественно в каналах PO4, PO7 и O1.

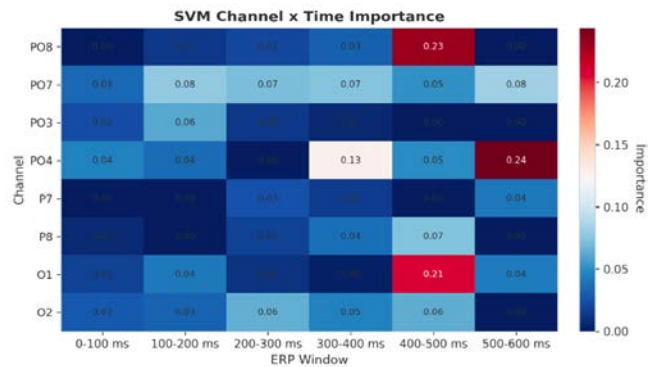


Рис. 4. Тепловая карта важности признаков по каналам и времени, выделяющая поздние задние ERP-окна.

D. Precision-Recall и задержки

Из-за сильного дисбаланса классов анализ precision-recall обеспечивает более информативную оценку, чем точность. Рис. 5 показывает, что средняя точность 0.3467 превышает базовый уровень распространённости, что указывает на значимую чувствительность к целевым стимулам.

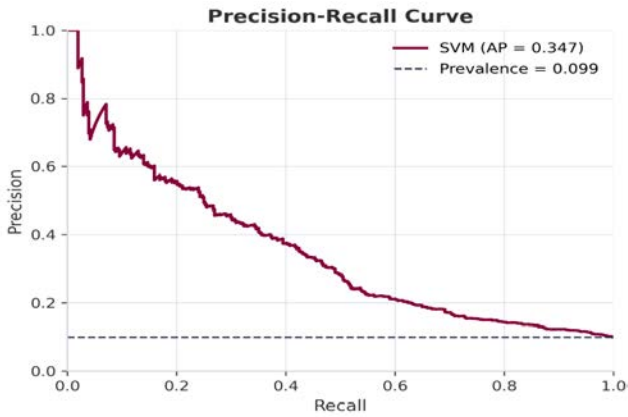


Рис. 5. Кривая precision-recall для лучшего классификатора при объединённой оценке.

Анализ задержек показал, что предобработка доминирует в вычислительных затратах (490.9 мс), за ней следует сегментация эпох (110.4 мс), тогда как извлечение признаков и инференс остаются малозатратными (< 7 мс суммарно). Общее время работы конвейера (L_{total}) составило приблизительно 609 мс, что подтверждает вычислительную эффективность предложенного фреймворка.

ТАБЛИЦА II. РАЗЛОЖЕНИЕ ЗАДЕРЖЕК ПРЕДЛОЖЕННОГО КОНВЕЙЕРА

Компонент	Время (мс)
Preprocessing	490.8787
Epoching	110.3853
Feature Extraction	6.0806
LDA Inference	0.3452
SVM Inference	1.3230

IV. ОБСУЖДЕНИЕ

Полученные результаты подтверждают три основных вывода. Во-первых, детекция целевых стимулов в RSVP лучше описывается с использованием признаков на основе ERP-временных окон, чем с помощью обобщённых нетемпоральных представлений. Наиболее выраженные различия стабильно наблюдаются в интервале 400–500 мс, а наиболее значимыми каналами являются PO4, PO7 и O1. Данный паттерн имеет нейрофизиологический смысл и согласуется с динамикой ERP, связанной с целевыми стимулами в задних областях при RSVP-парадигмах [2], [4].

Во-вторых, абсолютные значения классификации следует интерпретировать в контексте самой задачи. Датасет характеризуется выраженным дисбалансом классов, целевые события по своей природе редки, а используемый монтаж является разреженным и ограничен задними каналами. В таких условиях сбалансированная точность, PR AUC и субъектно-ориентированная оценка дают более корректную характеристику, чем обычная точность. С этой точки зрения результаты SVM при объединённой оценке указывают на значимую, но всё ещё сложную задачу восстановления целевых событий, а не на тривиальное поведение модели, ориентированной на доминирующий класс [10]. Такая интерпретация согласуется с предыдущими исследованиями RSVP, показывающими, что производительность детекции редких событий

должна оцениваться с использованием метрик, учитывающих дисбаланс, а не только точности [1], [9].

В-третьих, предложенный фреймворк остаётся ценным, поскольку сочетает воспроизводимость, интерпретируемость и вычислительную эффективность. Вместо использования сложной «чёрной коробки», данная работа предлагает лёгковесную базовую модель, решения которой напрямую связаны с интерпретируемой структурой ERP по каналам и времени. Это делает предложенный подход особенно актуальным для задач нейроинженерии, где важны как физиологическая прозрачность, так и инженерная простота [3], [11].

Полученные результаты также следует рассматривать в контексте современных тенденций моделирования. Архитектуры глубокого обучения, такие как EEGNet и более новые модели внимания, ориентированные на RSVP, способны захватывать более сложные пространственно-временные структуры, однако они обычно снижают прямую физиологическую интерпретируемость и часто требуют более тщательной настройки [8], [13], [14]. В отличие от этого, предложенный подход на основе ERP-временных окон сохраняет явное соответствие между представлением признаков и временной структурой ERP.

Следует также отметить ряд ограничений. Использовались только восемь задних каналов, что ограничивает пространственное разрешение. Кроме того, обобщающая способность модели в субъектно-ориентированной постановке остаётся умеренной: сбалансированная точность снижается с 0.6729 при объединённой оценке до 0.6201 при субъектно-ориентированной оценке для SVM, что отражает межсубъектную вариабельность EEG-ответов в RSVP-задачах. В будущих исследованиях приоритет может быть отдан стратегиям адаптации к субъекту и методам переноса обучения, поскольку межсубъектная вариабельность, по-видимому, является ключевым ограничением для обобщения [12]. Также могут быть рассмотрены методы пространственной фильтрации и гибридные лёгковесные модели глубокого обучения при сохранении достигнутой интерпретируемости.

В дальнейшем планируется сравнение предложенного фреймворка с компактными моделями глубокого обучения, такими как EEGNet [8], для оценки компромисса между интерпретируемостью и сквозным обучением.

V. ЗАКЛЮЧЕНИЕ

В данной работе представлен лёгковесный и объяснимый ЭЭГ-фреймворк для детекции внимания в задачах быстрого последовательного визуального предъявления. Используя разреженные задние электроды, минимальную предобработку, признаки на основе средних амплитуд ERP во временных окнах и классические классификаторы, предложенный подход продемонстрировал значимую производительность детекции целевых стимулов на датасете PhysioNet RSVP EEG. Лучшая модель при объединённой оценке — RBF-SVM с взвешиванием классов — достигла сбалансированной точности 0.6729 и PR AUC 0.3467. При субъектно-ориентированной оценке производительность умеренно снизилась

(сбалансированная точность 0.6201, PR AUC 0.2620 для SVM), что отражает межсубъектную вариабельность как основную проблему обобщения.

Полученные результаты показывают, что ERP-ориентированный и вычислительно эффективный конвейер RSVP может служить воспроизводимой и интерпретируемой базовой моделью для приложений в области интерфейсов мозг–компьютер (BCI) и нейроинженерии.

Публично доступная реализация предложенного фреймворка доступна по адресу:

<https://github.com/kevinsanchezr/erp-rsvp-attention-eeg> и включает полный экспериментальный конвейер и код анализа.

БЛАГОДАРНОСТИ

Авторы выражают благодарность Санкт-Петербургскому государственному электротехническому университету «ЛЭТИ» им. В.И. Ульянова (Ленина) за академическую поддержку при подготовке и публикации данной работы.

СПИСОК ЛИТЕРАТУРЫ

- [1] A. Matran-Fernandez and R. Poli, "Towards the automated localisation of targets in rapid image-sifting by collaborative brain-computer interfaces," *PLOS ONE*, vol. 12, no. 5, e0178498, 2017.
- [2] S. J. Luck, *An Introduction to the Event-Related Potential Technique*. 2nd ed. Cambridge, MA, USA: MIT Press, 2014.
- [3] A. Chaddad et al., "Electroencephalography signal processing: A comprehensive review and analysis of methods and techniques," *Sensors*, vol. 23, no. 14, Art. no. 6434, 2023.
- [4] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components a tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814-825, 2011.
- [5] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215-e220, 2000.
- [6] A. Matran-Fernandez and R. Poli, "EEG Signals from an RSVP Task," *PhysioNet*, 2017, doi: 10.13026/C2RC7Z.
- [7] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, and F. Yger, "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, 031005, 2018.
- [8] 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," *Journal of Neural Engineering*, vol. 15, no. 5, 056013, 2018.
- [9] L. Acqualagna and B. Blankertz, "Gaze-independent BCI-spelling using rapid serial visual presentation (RSVP): a review," *Frontiers in Neuroscience*, vol. 11, Art. no. 78, Feb. 2017, doi: 10.3389/fnins.2017.00078.
- [10] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLoS ONE*, vol. 10, no. 3, e0118432, 2015.
- [11] T. Tjoa and C. Guan, "A survey on explainable artificial intelligence for medical diagnosis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 2, pp. 520-536, Feb. 2021.
- [12] S. Kilani, S. N. Aghili, Y. Fathi, and A. I. Sburlea, "Optimization of transfer learning based on source sample selection in Euclidean space for P300-based brain-computer interfaces," *Frontiers in Neuroscience*, vol. 18, Art. no. 1360709, Jul. 2024.
- [13] J. Luo et al., "A cross-scale transformer and triple-view attention based domain-rectified transfer learning for EEG classification in RSVP tasks," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 1425-1435, 2024, doi: 10.1109/TNSRE.2024.3359191.
- [14] N. R. Waytowich et al., "Compact convolutional neural networks for classification of asynchronous steady-state visual evoked potentials," *Journal of Neural Engineering*, vol. 15, no. 6, Art. no. 066030, 2018, doi: 10.1088/1741-2552/aae5d8.