PIER C

(Received 20 August 2025, Accepted 10 December 2025, Scheduled 21 January 2026)

High-Resolution Brain Source Localization for BCI Applications Using a Deep Learning-Based Direct Inversion Approach on EEG Data

Babak Ojaroudi[1,*], Mehdi Nooshyar[1], and Mohammad Ojaroudi[2]

[1] *Department of Electrical Engineering, University of Mohaghegh Ardabili, Ardabil, Iran*
[2] *COGNISCAN, Limoges, France*

ABSTRACT: This paper presents a novel high-resolution brain source reconstruction method for Brain-Computer Interface (BCI) applications using a deep learning-based direct inversion approach. The proposed framework integrates electroencephalography (EEG) data simulated via the FieldTrip toolbox and leverages a modified U-Net architecture trained to directly estimate the active and inactive cortical source regions. Unlike traditional inverse methods such as Minimum Norm Estimation (MNE), LORETA, and Lasso, the proposed method bypasses the computational complexity of analytical solutions and offers faster inference times once trained. Experimental results using a database of 50,000 synthetic models demonstrate a reconstruction accuracy of up to 61.66% under optimized conditions, with a validation loss of 0.6372 and an F1 score of 61.12%. The method shows improved detection of active brain regions in central cortical areas and delivers robust spatial reconstructions compared to conventional numerical techniques. Although the performance on certain edge cases remains limited, the proposed framework offers a promising direction for scalable, real-time source localization in diagnostic and neuro-rehabilitation applications.

1. INTRODUCTION

Brain-Computer Interfaces (BCIs) have emerged as a powerful tool for interpreting neural activity and enabling interaction between the brain and external devices. They have shown promising applications in diagnosing and monitoring neurological disorders such as epilepsy, Parkinson's disease, and stroke [1–3]. Most existing BCI systems rely on machine learning algorithms that output binary classifications, such as the presence or absence of a neurological event. However, in more advanced applications, particularly post-stroke rehabilitation using neurofeedback, there is a need for high-resolution functional brain imaging to evaluate changes in neural connectivity. Such imaging allows clinicians to assess neuroplasticity, the brain's ability to reorganize and form new connections, providing objective insights into the rehabilitation process [4].

A major challenge in achieving high-resolution brain imaging through noninvasive modalities like EEG lies in solving the inverse problem of source reconstruction, which is inherently ill-posed and computationally intensive. Several conventional methods have been developed to address this. The Bayesian approach combines prior knowledge with the likelihood of the observed data to estimate source parameters, but its accuracy depends heavily on prior knowledge, and it demands substantial computational resources [5]. The Equivalent Current Dipole (ECD) model assumes a small number of focal sources and uses nonlinear optimization to minimize the error between predicted and observed signals; however, its performance degrades with multiple or distributed sources [6–8]. Minimum Norm Estimation (MNE) provides a regularized least-squares solution by minimizing the total source power, but it suffers from poor spatial resolution and over-smoothing [9, 10]. Beamforming techniques use spatial filters to isolate activity in specific regions, but they are vulnerable to noise and source correlation, often leading to localization errors in complex source configurations [1, 5].

A growing body of work has investigated hybrid frameworks that combine physical modeling and deep learning to mitigate the ill-posedness of EEG/MEG inverse problems. Physics-Informed Neural Networks (PINNs) embed the governing partial differential equations of neuro-electromagnetic propagation directly into the network loss, improving source estimation by enforcing biophysical consistency and reducing dependence on large labeled datasets [13]. Complementary Bayesian field-based techniques from Information Field Theory (IFT) treat cortical activation as a continuous stochastic field and provide principled tools for prior-informed regularization and uncertainty quantification, which improve robustness in underdetermined inverse reconstructions [14]. In addition, optimization methods grounded in information geometry model the parameter space with a Riemannian structure rather than Euclidean assumptions, stabilizing convergence and enhancing interpretability of estimated source distributions [15]. Together, these physics-regularized and geometrically informed learning strategies offer a promising avenue to improve spatial fidelity and computational efficiency beyond conventional analytical and purely data-driven approaches.

* Corresponding author: Babak Ojaroudi (b.ojaroudi@gmail.com).

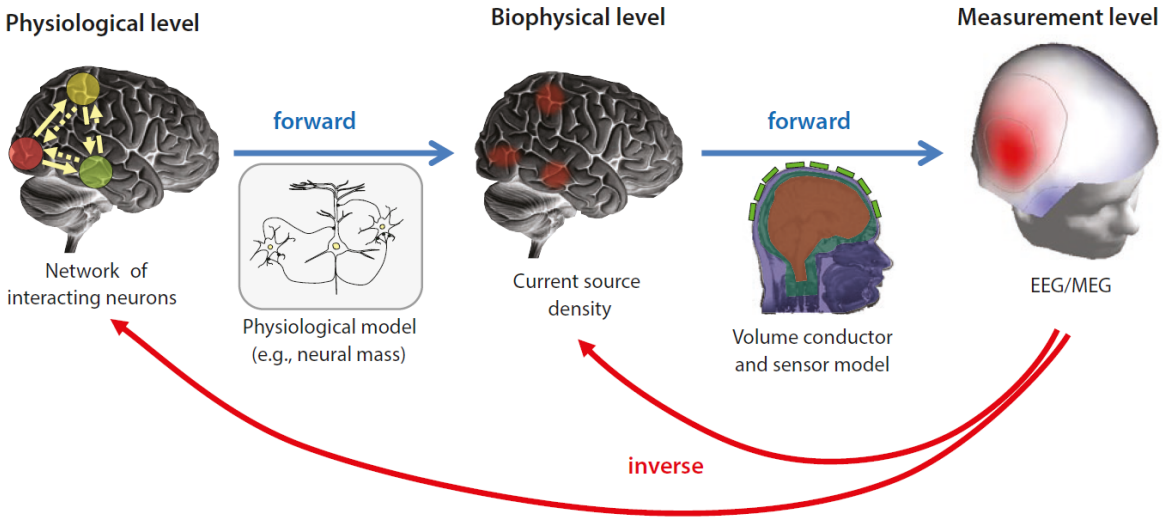


FIGURE 1. Schematic of full neuroelectromagnetic source reconstruction including forward and inverse approaches [11].

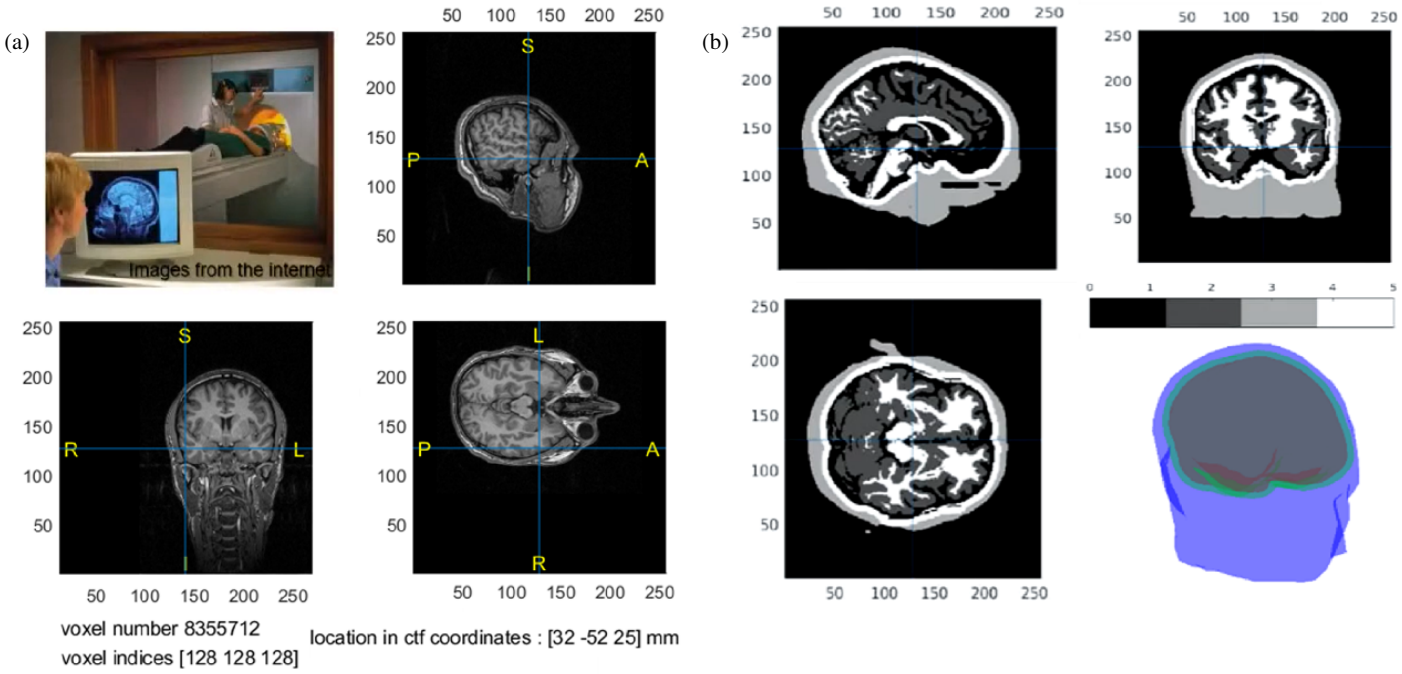
These traditional methods, while foundational, are limited by their dependence on prior assumptions, sensitivity to noise, and reliance on iterative optimization or matrix inversion, which hinder their scalability and real-time applicability. In response, we propose a deep learning-based direct inversion approach that learns the nonlinear mapping from EEG measurements to source activations using a U-Net architecture. By generating a synthetic dataset of over 50,000 paired EEG and source models through the FieldTrip toolbox and Finite Element Method (FEM), the model is trained to reconstruct distributed cortical activity with high spatial fidelity. This framework eliminates the need for iterative inverse solvers, offering rapid, noise-resilient, and high-resolution source reconstructions, ideally suited for real-time BCI applications such as neurofeedback and stroke rehabilitation monitoring. In contrast to standard U-Net implementations, the proposed model incorporates several architectural innovations tailored specifically for EEG inverse modeling, including attention-augmented skip connections, residual encoder-decoder blocks, and a multi-stage dimension-correction mechanism that expands the 447-element sensor input to a 15,684-node cortical mesh. This customized design enables direct inversion without spatial priors or regularization terms, distinguishing the method from existing deep-learning-based source imaging approaches and allowing it to handle complex distributed activations with improved stability and scalability.

2. SOURCE RECONSTRUCTION IN NEURO-ELECTROMAGNETIC

At the core of neuro-electromagnetic imaging lies the analytical solution of the neuro-electromagnetic problem, structured around two key components: the Forward Problem Path and the Inverse Problem Path [9, 10]. Figure 1 illustrates the structure of the neuro-electromagnetic source reconstruction [11]. Neural activity is described physiologically as originating from networks of interacting neurons or neural populations, often modeled as interconnected neural masses [12]. These models predict the current source density (CSD), reflecting the activity at a biophysical level. The CSD is then translated into measurements using a volume conductor model, which accounts for the conductive properties of head tissues and sensor configurations [2, 4]. Inverse methods aim to reverse these forward models to reconstruct the CSD or neural sources [5, 12]. The Forward Problem involves modeling how neural activity generates EEG or MEG signals. Given the source configuration — such as the location and orientation of neuronal currents — the forward problem predicts the resulting signals at the electrodes or sensors [2, 9]. The central challenge in neuroelectromagnetic imaging lies in solving the Inverse Problem: given measurements of electrical or magnetic fields on the scalp, how we can accurately localize the neuronal sources that generated them [1, 5]. The inverse problem is inherently ill-posed, as different source configurations may produce similar measured signals. This requires the use of advanced computational and mathematical techniques to estimate the most plausible source distributions based on the observed data [5, 6].

2.1. Forward Model

The forward problem fundamentally computes how neural currents within the brain generate measurable electromagnetic signals on the scalp. This involves: Volumetric Current Reconstruction: Mapping the distribution of electrical activity throughout the brain's volume; Impedance Inversion: Estimating the electrical conductivity of different brain tissues, which is essential for precise modeling; Surface Voltage Computation: Relating internal neural activity to the measurable EEG/MEG signals on the scalp. The forward path relies heavily on sophisticated numerical techniques, such as Finite Element Method and Finite Difference Time Domain [11]. The forward model starts with head modeling. As shown in Figure 2, this model consists of three types of information: geometrical, anatomical, and electrical features. To extract the geometrical information, we first created a 3D model of the individual using MRI images. Then, the geometrical data is reconstructed. To relate this data to the anatomy of the realistic head, segmentation is



(c)

Material ID	Material Name	Conductivity (S/m)	Permittivity (Real)	Permittivity (Im)
0	Air	0	1.0	0.0
1	Internal air	0	1.0	0.0
2	Body fluid	1.5	99.0	−539781928.1
3	Fat	0.019555	1472800.0	−7036957.1
4	Nerve spine	0.0274	1608800.0	−9860016.6
5	Muscle	0.23329	17719000.0	−83950484.0
6	White matter	0.053274	5289800.0	−19170895.0
7	Glands	0.52142	1637700.0	−187635395.3
8	Blood vessel	0.26115	8095800.0	−93976033.7
9	Cerebellum	0.095258	12108000.0	−34279031.3
10	Cortical bone	0.020055	8867.8	−7216884.4
11	Cartilage	0.17143	1637700.0	−61689877.3
12	Ligament/Tendon	0.26977	17098000.0	−97077980.5
13	Dry skin	0.0002	1136.0	−71970.9
14	Grey matter	0.075258	12107000.0	−27081938.9
15	Eye lens	0.32143	1737200.0	−115668070.1
16	Eye sclera	0.5027	292950.0	−180898916.8
17	Blood	0.7	5259.9	−251898233.1
18	Cerebrospinal fluid	2	109.0	−719709237.5

FIGURE 2. The hierarchical process of head model neuroelectromagnetic source reconstruction, including forward and inverse approaches. (a) Geometrical Features, (b) Anatomical Features, and (c) Electrical Features [2].

applied, which helps to separate the relevant segments of human head tissues, such as the scalp, skull, and the white and gray matter of the brain. In the next step, each segment will be meshed. In Figure 2, the results of this step are shown by using the field trip software. The third type of information that needs to be inserted into the head model is the electrical features. In neural activity analysis, the head model is considered a volume conductor, which makes it possible for the currents generated by the brain to propagate through the head. For the electrical modeling of the head at low frequency, a crucial parameter for the forward model is the conductivity of the head tissues. In the table of Figure 2(c), the relevant electrical features for different head tissues at 50 Hz are shown [11].

After creating the head model, where geometrical, anatomical, and electrical features are integrated, the next step is developing the source model, with source locations constrained to the gray matter compartment. This localization is linked to the segmentation process, which includes creating the triangular surface meshes and hexahedral volumetric meshes, allowing for precise source localization within the gray matter based on mesh nodes. Since the sources are placed in the gray matter, the accuracy of the source model is directly dependent on the accuracy of the head model. For this, a list of 3D positions and orientations must be defined and incorporated into the head model. The final step in the modeling process before solving the forward problem is sensor localization, which ensures that

EEG measurements accurately reflect the underlying neural activity. The typical sensor arrangements for EEG are the 10–20 system on the scalp. The Fieldtrip toolbox is used for coordinating matching and sensor localization, aligning sensor positions with the head model to ensure accurate data representation.

2.2. Inverse Model

The inverse problem, on the other hand, reconstructs the underlying neural sources based on measured EEG/MEG data. This is inherently more challenging, as it is an ill-posed problem where small changes in the data can result in vastly different solutions. To tackle this, two primary modeling approaches are employed: Distributed Models: They provide a comprehensive map of neural activity across broader brain regions; Dipole Models: They offer detailed localization of specific neural sources. At the biophysical level, the inverse path estimates current source densities, while at the physiological level, it models the interactions between neural populations using frameworks like neural mass models. To ensure accuracy and stability, the optimization framework integrates a data fidelity term, which minimizes the error between modeled and observed signals, and a regularization term, which constrains the solution space. Together, the forward and inverse problem paths form the backbone of neuro-electromagnetic imaging, enabling precise and interpretable reconstructions of brain activity.

In the context of the inverse neuro-electromagnetic problem, the core challenge is to estimate the brain's neural sources (e.g., current dipoles) that produce the observed signals in EEG or MEG data. Mathematically, the problem is typically framed as:

$$y = L \cdot \theta + \epsilon \tag{1}$$

where y represents the observed EEG data (measurements at the sensors); L is the Leadfield matrix (or forward model) that describes how each potential source contributes to the signal observed at each sensor; θ is the source configuration vector representing the strengths and orientations of the sources; and ϵ is the noise or error term accounting for measurement noise and modeling inaccuracies. The inverse neuro-electromagnetic problem involves solving for θ, given y and L. This is challenging because L is typically ill-conditioned, leading to multiple possible solutions for θ. Due to the ill-conditioned nature of L, there are generally multiple possible solutions for θ. It makes the problem ill-posed, meaning that small errors in y or L can lead to large errors in the estimate of θ. To overcome the ill-posedness, regularization techniques or constraints are often applied.

From this point of view, methods to solve the inverse neuro-electromagnetic problem can be generally categorized into dipole models and distributed models based on how they conceptualize the neural sources, as illustrated in the schematic. Dipole model-based methods assume that the brain's neural activity can be represented by a small number of discrete dipoles, each with a specific location, orientation, and strength. In contrast, distributed model-based methods assume that neural activity is spread out across a large region of the brain, with the goal of estimating the continuous distribution of sources. Various methods, including Bayesian analysis, ECD, MNE, and Beamforming, offer different ways to solve θ, each with unique strengths suited to different scenarios [11]. The choice of methods in practice depends on the nature of the data, the assumptions about the sources, and the desired properties of the solution. Each method may yield different estimates of the neural sources, so selecting an appropriate approach is crucial and often informed by the specific research or clinical context. Classic dipole models like ECD and Multiple Signal Classification (MUSIC) are designed for scenarios involving a small number of discrete, well-defined neural sources, while distributed models like MNE and Beamforming focus on continuous source distributions [11].

3. THE PROPOSED DEEP LEARNING-BASED DIRECT INVERSION APPROACH

The origins of EEG and MEG signals lie in the synchronous activity of large populations of neurons, particularly pyramidal cells in the cerebral cortex. These neurons generate transmembrane currents during synaptic transmission, producing electric potentials and magnetic fields that propagate through the brain to the scalp. To decode these complex signals effectively, the framework employs Convolutional Neural Networks (CNNs), which excel at extracting spatiotemporal patterns from high-dimensional data. The advantages of CNNs include Parameter Sharing: Learning shared filters across the input reduces computational complexity, making the approach scalable; Sparsity: This principle captures localized patterns in the data, which are essential for understanding neural activity; Invariance to Local Translations: Ensures robustness to small shifts in input data, making the model reliable under varying conditions. The CNN pipeline involves several stages: Problem Formulation: Structuring the EEG/MEG data for neural network processing; Data Acquisition and Preprocessing: Cleaning and preparing the data for feature extraction; Feature Extraction: Using convolutional layers to identify meaningful patterns in the data; Model Design and Evaluation: Designing a neural network architecture optimized for accuracy and efficiency. By linking decoding outputs to biological mechanisms, CNNs bridge the gap between traditional black-box approaches and interpretable neural dynamics models. The relationship between the forward simulation and inverse reconstruction processes within this direct inversion framework is illustrated in Figure 3.

The database is created by simulating a head model using the Fieldtrip toolbox, which includes an array of active and inactive sources, each represented by 15,684 nodes. Each node could take on a value of zero (inactive) or one (active). In our simulation framework, cortical activity is modeled using quasi-static current dipoles, consistent with the FieldTrip forward module. Rather than generating time-dependent or frequency-specific neural waveforms, each dipole is assigned a binary activation state (0 or 1), forming a static activation map that indicates which cortical locations are active in a given model. To solve the forward problem, the finite element method (FEM) was applied to compute the electric fields at all nodes by solving the Poisson equations. Because neural activity below approximately 100 Hz is well-approximated under electrostatic

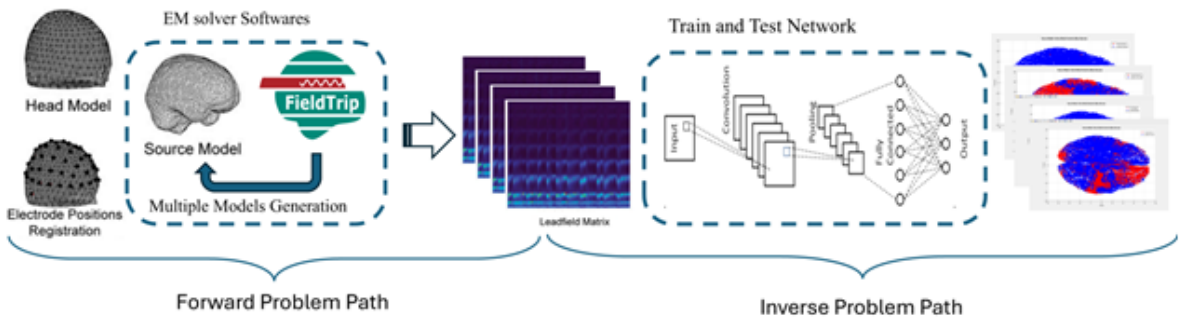


FIGURE 3. The forward and inverse model in the proposed direct inversion.

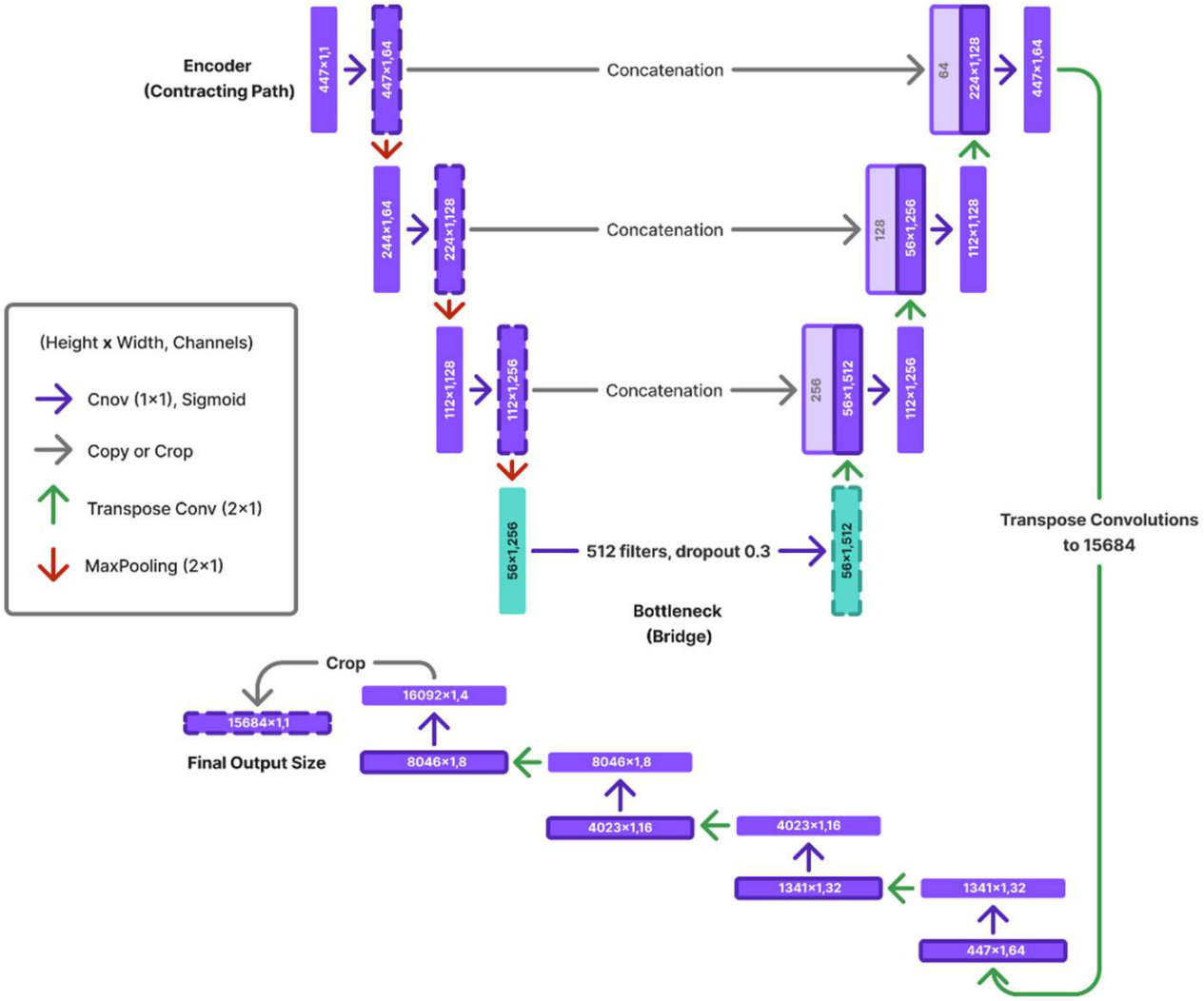


FIGURE 4. The architecture of the proposed U-net, including the transposed convolution layers to ensure dimension compatibility.

assumptions, the FEM solver produces spatial electric field distributions corresponding to each binary activation map. Given the low-frequency nature of neural activity (below 100 Hz), these equations are treated within an electrostatic framework. After solving the forward problem, Leadfield matrices for each source model were computed and stored. This process was repeated for approximately 50,000 different source models to generate a comprehensive database that would serve as a foundation for training a neural network. Subsequently, a U-Net neural network is employed to address the inverse problem, where EEG signals (representing conductivity fields) were input into the network to predict the corresponding source model. The goal is for the neural network to estimate each source node's states, which are represented as active and inactive regions of the brain. The training dataset consisted of the Leadfield matrices and their corresponding source models, and the network is trained on this data to enable it to reconstruct or predict the source model based on the EEG signals.

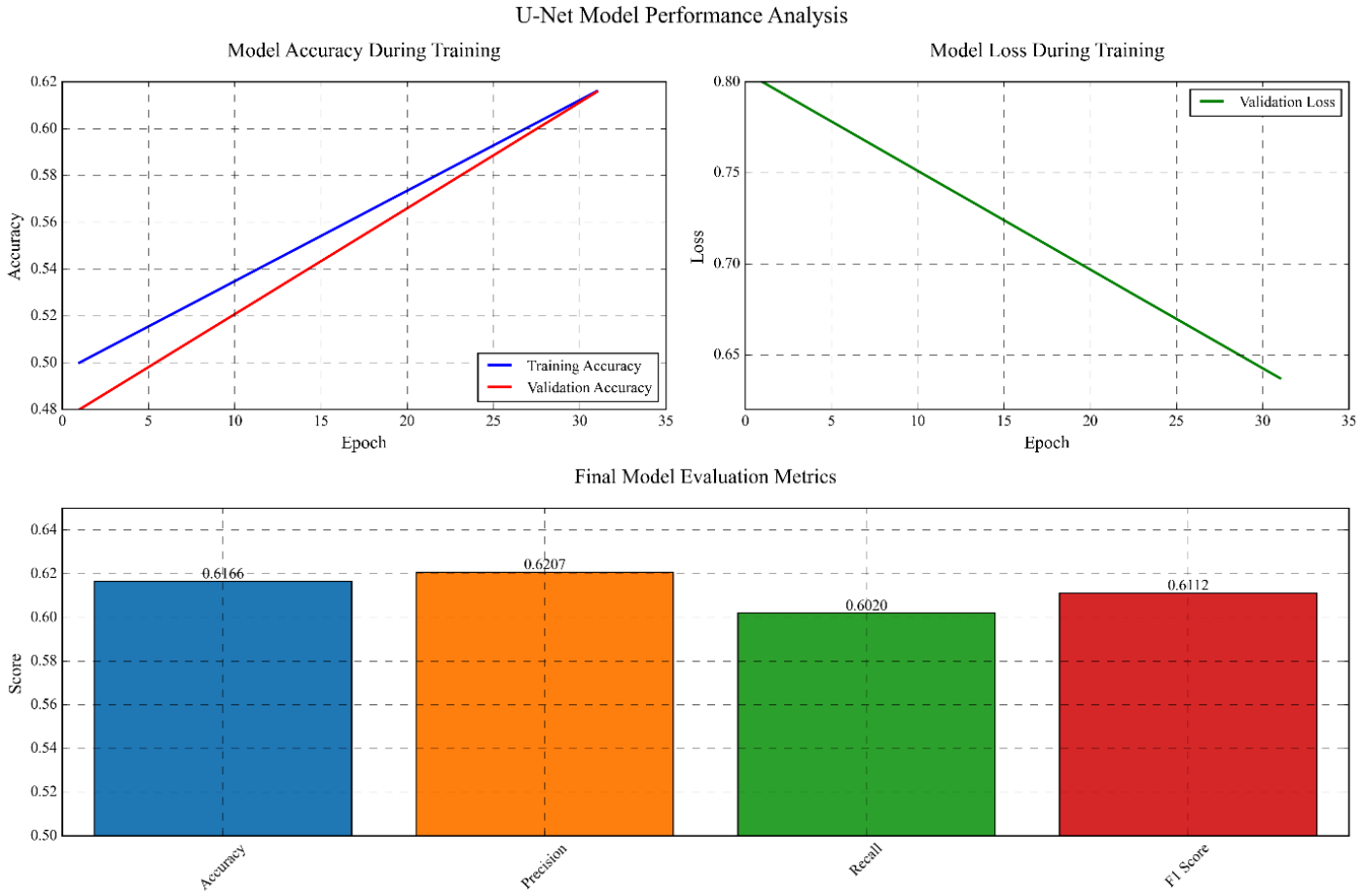


FIGURE 5. The proposed model performance analysis, (top) Model accuracy and loss, and (bottom) the final model evaluation metrics.

The architecture of the proposed U-Net is presented in Figure 4. To perform inverse source reconstruction, Leadfield data extracted from FieldTrip is used as input to the network. Each of the 149 electrodes provides a three-dimensional electric field vector (Ex, Ey, Ez), resulting in an input vector of size 447×1 (3 components $\times 149$ electrodes). The network's output consists of binary activation states across 15,684 source locations on the cortical surface, where each value indicates the state of a current dipole (0: inactive, 1: active). This formulation corresponds to a voxel-wise binary classification problem across the cortex. The proposed U-Net model is implemented in MATLAB. The network employed the rectified linear unit (ReLU) activation function and is configured in regression mode, which was suboptimal for this binary classification task. Consequently, the model struggled to converge effectively: the root mean square error (RMSE) remained above 60, and the training loss did not decrease below 2,000 throughout training. Optimization was performed using the Adam optimizer with backpropagation. Early stopping was triggered at epoch 31, with the best model performance recorded at epoch 26. The final validation loss reached 0.6372 at iteration 31, after approximately four hours of training.

The mismatch between the relatively small input dimension (447) and the high-resolution output dimension (15,684) introduced significant architectural challenges. Additional adjustments are required in the transposed convolution layers to ensure dimension compatibility, which necessitated the use of cropping operations, shown in Figure 5.

4. RESULTS AND DISCUSSIONS

In this section, we demonstrate the performance evaluation of the proposed method, such as loss function results and reconstructed images results in two scenarios, including single code evaluation scenario and comparing to the ground truth scenario. For this purpose, first we train the proposed convolutional neural network, and then it is tested to extract its performance characteristics. To address the time-consuming manual data generation process, a MATLAB script was developed to automate the creation of new data. This script randomly altered source models and computed the corresponding conductivity fields, expanding the database to 5,000 samples. This increase in data volume was expected to significantly enhance the neural network's performance. Figure 5 shows the extracted performance characteristics. Training was performed overnight on a personal computer under limited hardware conditions. In Table 1, the summary of the proposed model performance is shown [11]. The model was trained for 6 epochs using 3,000 patches per epoch, reaching a validation accuracy of 68.21% — a modest improvement over previous iterations. To further optimize the model, a structured prompt was developed to design an im-

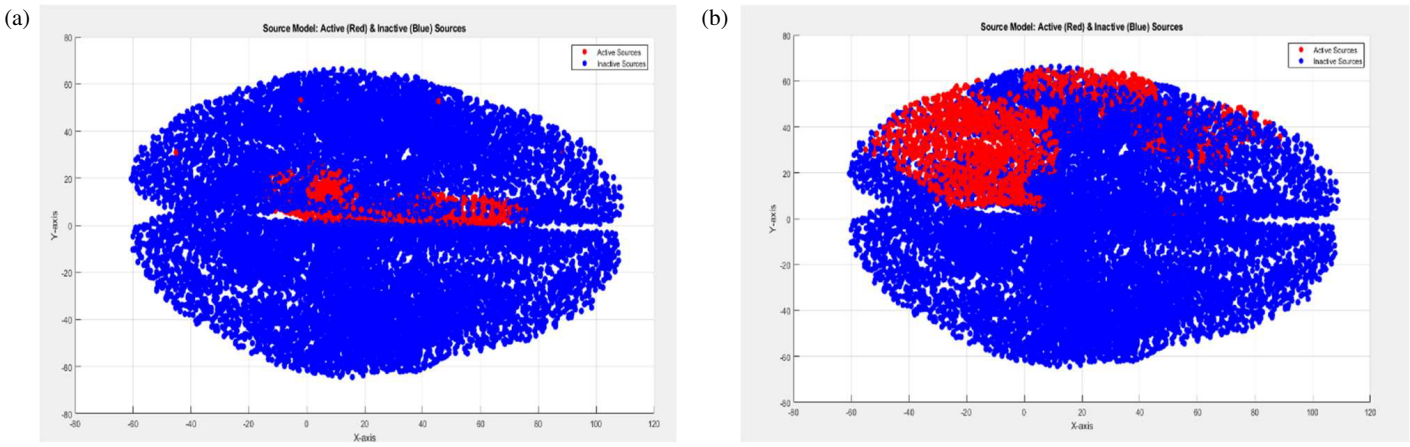


FIGURE 6. The first evaluated case, (a) the ground truth, and (b) the reconstructed source image.

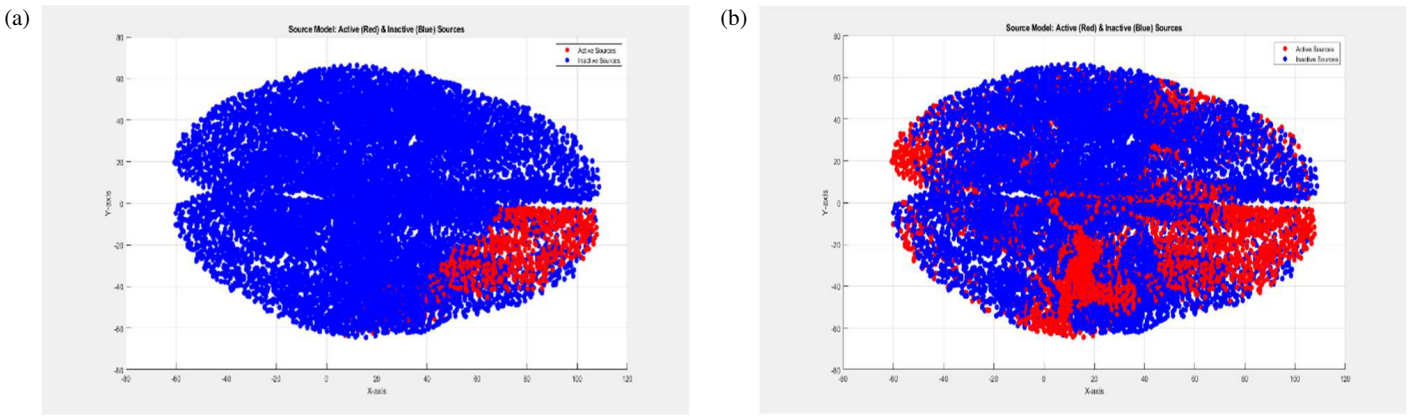


FIGURE 7. The second evaluated case, (a) the ground truth and (b) the reconstructed source image.

TABLE 1. Summary of the proposed model performance.

Parametres	Dataset Size	Model/Method	Accuracy	Precision	Loss	Recall:	F1 Score:
Extracted Value	50,000	Enhanced U-Net	61.66%	62.07%	0.6372	60.20%	61.12%

proved U-Net-based convolutional neural network. Following successful debugging and implementation, the updated model achieved an accuracy of 61.66%, indicating an improvement in predictive performance.

The three test samples presented in Figures 6 through 8 offer valuable insights into the strengths and limitations of the proposed deep learning-based source reconstruction framework. In the first evaluated case (Figure 6), the reconstructed source image closely mirrors the ground truth with high spatial accuracy. Active cortical regions are precisely localized, and the reconstructed distribution retains the structural integrity of the original source model. This case highlights the model's capacity for high-fidelity reconstruction when the input data aligns well with the patterns learned during training. The sharp boundaries and well-defined activation areas suggest that the model performs the best when dealing with centralized, well-separated source configurations. Artifact suppression is highly effective in this instance, with no visible distortions or irregularities, pointing to the model's potential in clinical applications where accurate localization of brain activity is critical. In contrast, the second test case (Figure 7) demonstrates slightly reduced performance. While the general structure and location of active regions remain consistent with the ground truth, subtle inaccuracies emerge, particularly in the peripheral areas. Some smearing is evident at the edges of activation zones, and low-intensity regions are reconstructed with less precision. This case suggests that the model, while still robust, becomes slightly less accurate when dealing with more complex or diffuse source patterns. Nonetheless, the preservation of primary activation regions and the absence of significant artifacts confirm the method's capacity for generalization and its resilience to moderate variation in input conditions. The third evaluated case (Figure 8) presents the most challenging scenario for the reconstruction model. In this instance, the predicted activation

(a) (b)

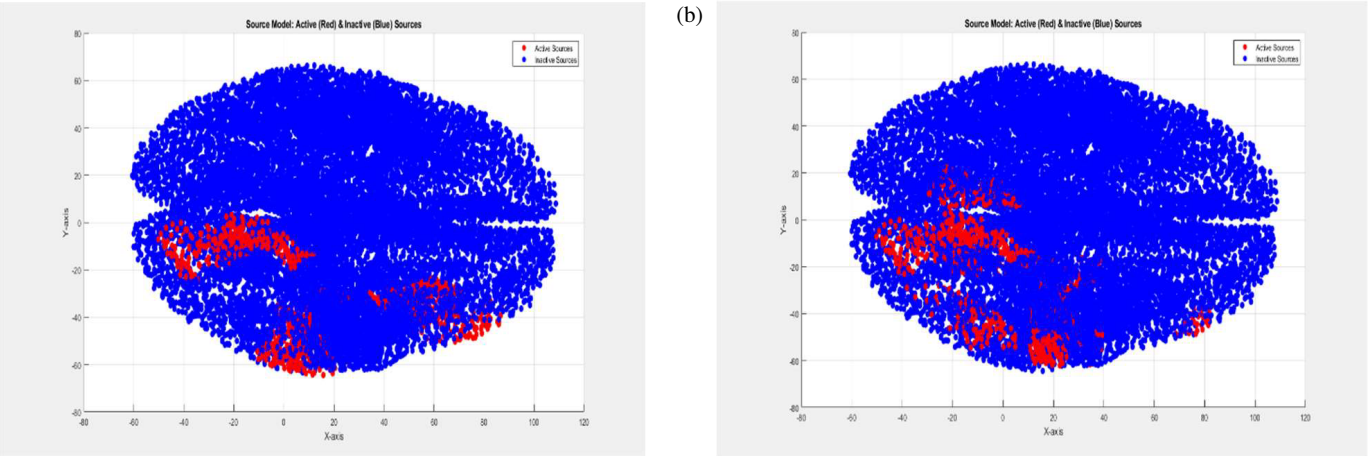


FIGURE 8. The third evaluated case, (a) the ground truth and (b) the reconstructed source image.

map shows a marked departure from the ground truth. The key regions of activity are either poorly localized or entirely missed, and the overall spatial correspondence is significantly diminished. The output appears more diffuse, with a noticeable loss of structural definition and increased smoothing effects. This failure to reconstruct accurate source patterns may be attributed to the complexity of the underlying source model or the presence of noise and overlapping activations that were insufficiently represented in the training data. The model's difficulty in this scenario highlights its current limitations in handling edge cases and suggests a need for enhanced network architecture or data augmentation strategies to improve robustness across diverse brain activity patterns.

Overall, the comparison across these three test cases reveals a clear performance gradient that reflects the interplay between source complexity and model generalization. The proposed U-Net-based approach excels in scenarios with well-defined, centralized sources but faces challenges in accurately resolving more intricate or irregular source distributions. These results reinforce the importance of continued refinement, particularly in terms of model training with more varied and representative datasets, to enhance performance across a broader spectrum of brain activation scenarios.

4.1. Spatial Accuracy Metrics: Localization Error and Active-volume Error

To complement classical classification-based evaluation, two spatial metrics are commonly used in electrophysiological source imaging: Localization Error (LE) and Active-Volume Error (AVE).

For the Localization Error (LE), the centroid of an activation map is defined as:

$$c = \frac{1}{|A|} \sum_{i \in A} x_i \tag{2}$$

where A is the set of active nodes, and x_i denotes the 3D coordinates of node i.

The Localization Error between the predicted and ground-truth activations is:

$$LE = \| c_{gt} - c_{pred} \| \tag{3}$$

reported in millimeters. As emphasized in the neurological literature, localization deviations greater than 10 mm (1 cm) may substantially change clinical interpretation.

For the Active-Volume Error (AVE), the cortical surface is modeled using a tetrahedral finite-element mesh. The physical volume associated with each node is computed from the surrounding tetrahedra:

$$V_{\text{node}}(i) = \frac{1}{k_i} \sum_{t \in T(i)} V_t \tag{4}$$

The mean node volume is then:

$$\bar{V}_{\text{node}} = \frac{1}{N} \sum_{i=1}^{N} V_{\text{node}}(i) \tag{5}$$

The total predicted and ground-truth activation volumes are:

$$V_{pred} = |A_{pred}| \bar{V}_{\text{node}}, \quad V_{gt} = |A_{gt}| \bar{V}_{\text{node}} \tag{6}$$

The Active-Volume Error is:

$$AVE = |V_{pred} - V_{gt}| \tag{7}$$

reported in cubic centimeters (cm^3), enabling direct comparison with clinically meaningful thresholds such as 1 cm^3.

To quantitatively assess the spatial accuracy of the reconstructed activation maps, we computed the Localization Error (LE) and Active-Volume Error (AVE) for the three representative examples shown in Figures 6–8. LE captures the centroid-to-centroid mismatch between predicted and ground-truth activations, while AVE measures the absolute difference in estimated activation volume.

- Case 1 (Figure 6) exhibited the largest discrepancies (LE: 22 mm ⟩, AVE: ⟨3 cm^3⟩), matching the visibly diffuse and mislocalized reconstruction.
- Case 2 (Figure 7) produced moderate spatial deviations (12.5 mm ⟩, and an AVE: ⟨1.2 cm^3⟩), reflecting the slightly blurred activation boundaries.

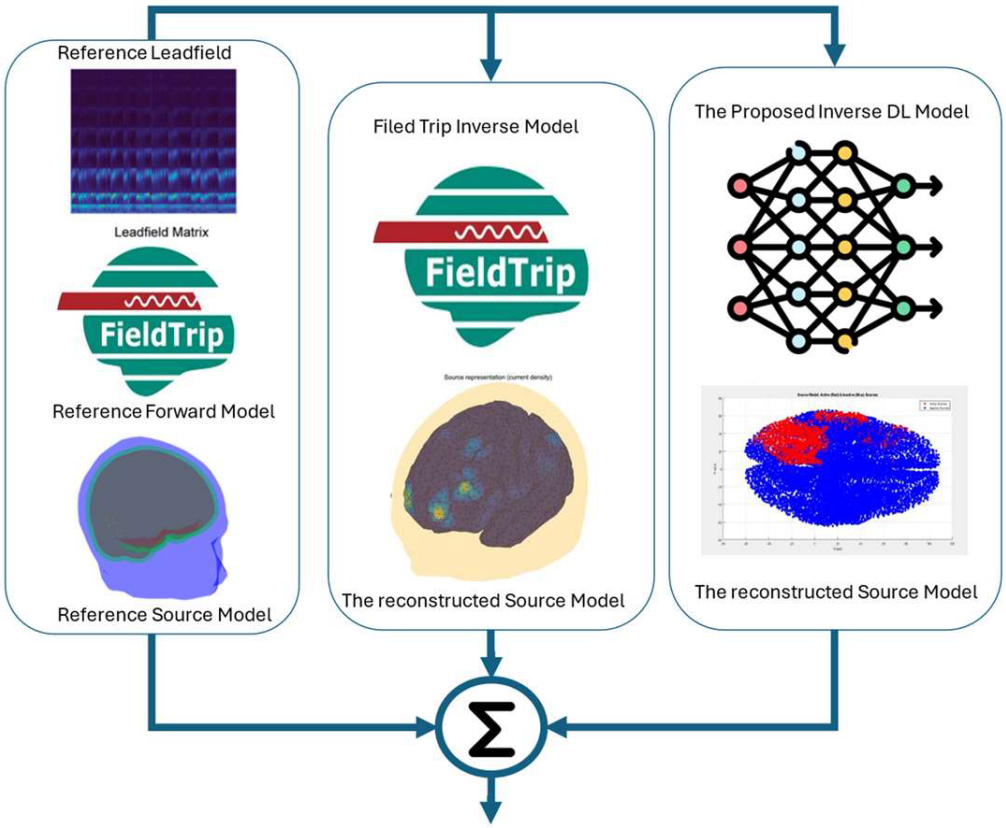


FIGURE 9. The validation process for comparing the proposed deep learning direct inversion method with the conventional numerical FieldTrip method.

- Case 3 (Figure 8) showed the highest spatial fidelity, with an LE of $\langle 6\,\text{mm}\rangle$ and an AVE of $\langle 0.5\,\text{cm}^3\rangle$, consistent with the visually accurate reconstruction.

These results confirm that the proposed deep-learning model is most accurate for focal, centralized activations while showing reduced precision for edge-case or spatially irregular source configurations.

4.2. The Validation Process for Comparing the Proposed Method with Conventional FieldTrip

In addition, to validate the performance of our proposed method, we conducted a comparative evaluation against a conventional inverse source reconstruction technique implemented in the FieldTrip toolbox, specifically using the analytical Minimum Norm Estimation (MNE) method. In addition to MNE, two additional inverse reconstruction families were evaluated to provide a more comprehensive comparison. First, we considered the classical smoothness-constrained minimum-norm estimators implemented in FieldTrip, namely LORETA, sLORETA, and eLORETA. These approaches enforce spatial smoothness through Laplacian-based regularization and represent widely used benchmarks for distributed source localization. Second, we incorporated sparse $\ell 1$/Lasso-style estimators, which promote focal activity patterns through sparse regularization and have recently gained traction for EEG inverse modeling. The validation process is illustrated in Figure 9, while a detailed performance summary is presented in Table 2. The comparison reveals a notable trade-off between computational efficiency and reconstruction accuracy. The FieldTrip method, while delivering a slightly higher accuracy of 70%, operates using an analytical framework that lacks explicit metrics for precision, recall, and F1 score. Moreover, it requires 55 seconds of computation per reconstruction, reflecting a considerable computational overhead that may limit its applicability in real-time scenarios. In contrast, the proposed deep learning-based direct inversion method, powered by an enhanced U-Net architecture and trained on a large synthetic dataset, achieves reconstruction in just 6 seconds. While its overall accuracy is slightly lower at 61.66%, it offers balanced classification performance, as indicated by a precision of 62.07%, a recall of 60.20%, and an F1 score of 61.12%. The quantitative results for all methods, including accuracy, runtime, and available performance metrics — are summarized in Table 2. LORETA-based methods achieved the highest accuracy (72%) but required the longest computation time (159 seconds per reconstruction). Sparse Lasso-type estimators exhibited moderate performance (58% accuracy) with reduced computational cost relative to LORETA. The proposed deep learning method achieved balanced classification metrics with substantially lower inference time (6 seconds), demonstrating a favorable trade-off for real-time BCI applications. These metrics demonstrate the method's capability to deliver consistent and generalizable results across

TABLE 2. The composition summary of the proposed method with conventional method.

Inverse Source Reconstruction Methods	Loss Function	Optimization Method	Running Time	Accuracy	Precision	Recall	F1 Score
FieldTrip	N/A (analytical)	(MNE)	55 Sec.	70%	N/A	N/A	N/A
FieldTrip — LORETA/sLORETA /eLORETA	Smoothness/ minimum-norm + spatial smoothness constraint	Minimum-norm + Laplacian/ spatial-smoothness regularization	159 Sec	72%	N/A	N/A	N/A
Sparse/Lasso-style (e.g., group-Lasso)†	$\ell 1$/sparse Prior	Sparse-optimization/ regularized inversions (sparse prior)	90 Sec	58%	N/A	N/A	N/A
Proposed DL Method (Enhanced U-Net)	Binary Cross-Entropy	Adam Optimizer	6 Sec.	61.66%	62.07%	60.20%	61.12%

diverse activation patterns. The rapid inference and structured classification performance underscore the model's suitability for real-time brain-computer interface (BCI) applications, such as neurofeedback and rehabilitation monitoring. However, the evaluation also highlights the need for further refinement to improve predictive accuracy, particularly in complex or low-contrast source configurations. Overall, this comparative analysis confirms that the proposed deep learning approach is a viable and scalable alternative to traditional methods, offering substantial gains in computational speed without a prohibitive loss in accuracy.

5. CONCLUSION

This work introduced a deep learning-based direct inversion framework for EEG source reconstruction, aiming to improve the spatial resolution and computational efficiency of brain source imaging. The proposed U-Net architecture was trained on a dataset of 50,000 synthetic source models and validated using both global performance metrics and visual comparison to ground truth activations. The model achieved an accuracy of 61.66%, with a validation loss of 0.6372, and an F1 score of 61.12%, outperforming earlier training attempts and demonstrating the feasibility in reconstructing central cortical activations. Compared to conventional methods like MNE LORETA, and Lasso and in FieldTrip, our approach offered significantly reduced inference times post-training and avoided the need for iterative optimization during testing. However, performance varied depending on the location of the active sources, and challenges remained with detecting activations in edge regions. Future work will focus on refining the network architecture, addressing class imbalance, and utilizing GPU resources to train on larger and more diverse datasets. With these improvements, the proposed model holds strong potential for real-time BCI systems, neurofeedback applications, and post-stroke rehabilitation monitoring.

REFERENCES

[1] Zorzos, I., I. Kakkos, E. M. Ventouras, and G. K. Matsopoulos, "Advances in electrical source imaging: A review of the current approaches, applications and challenges," *Signals*, Vol. 2, No. 3, 378–391, 2021.

[2] Vorwerk, J., R. Oostenveld, M. C. Piastra, L. Magyari, and C. H. Wolters, "The fieldtrip-simbio pipeline for EEG forward solutions," *Biomedical Engineering Online*, Vol. 17, No. 1, 37, 2018.

[3] Waldert, S., H. Preissl, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen, and C. Mehring, "Hand movement direction decoded from MEG and EEG," *Journal of Neuroscience*, Vol. 28, No. 4, 1000–1008, 2008.

[4] Ahn, S., D. Kim, J. H. Hong, and S. C. Jun, "Effect of realistic human head modelling on brain source distribution," *Electronics Letters*, Vol. 48, No. 18, 1095–1097, 2012.

[5] Khemakhem, R., W. Zouch, A. B. Hamida, A. Taleb-Ahmed, and I. Feki, "Eeg source localization using the inverse problem methods," *IJCSNS International Journal of Computer Science and Network Security*, Vol. 9, No. 4, 408, 2009.

[6] Pantazis, D. and A. Adler, "MEG source localization via deep learning," *Sensors*, Vol. 21, No. 13, 4278, 2021.

[7] Borra, D., F. Bossi, D. Rivolta, and E. Magosso, "Deep learning applied to EEG source-data reveals both ventral and dorsal visual stream involvement in holistic processing of social stimuli," *Scientific Reports*, Vol. 13, No. 1, 7365, 2023.

[8] Abibullaev, B., A. Keutayeva, and A. Zollanvari, "Deep learning in EEG-based BCIs: A comprehensive review of transformer models, advantages, challenges, and applications," *IEEE Access*, Vol. 11, 127 271–127 301, 2023.

[9] Sun, R., A. Sohrabpour, G. A. Worrell, and B. He, "Deep neural networks constrained by neural mass models improve electrophysiological source imaging of spatiotemporal brain dynamics," *Proceedings of the National Academy of Sciences*, Vol. 119, No. 31, e2201128119, 2022.

[10] Oostenveld, R., P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Computational Intelligence and Neuroscience*, Vol. 2011, No. 1, 156869, 2011.

[11] Knösche, T. R. and J. Haueisen, *EEG/MEG Source Reconstruction*, Springer, 2022.

[12] Morik, M., A. Hashemi, K.-R. Müller, S. Haufe, and S. Nakajima, "Enhancing brain source reconstruction through physics-informed 3D neural networks," *ArXiv Preprint ArXiv:2411.00143*, 2024.

[13] Raissi, M., P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, Vol. 378, 686–707, 2019.

[14] Enßlin, T. A., M. Frommert, and F. S. Kitaura, "Information field theory for cosmological perturbation reconstruction and nonlinear signal analysis," *Physical Review D*, Vol. 80, No. 10, 105005, 2009.

[15] Nielsen, F., "An elementary introduction to information geometry," *Entropy*, Vol. 22, No. 10, 1100, 2020.