

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Autoencoder-Driven Fiducial Landmark Identification in 3D brain MRI for Neuroimaging Alignment

DEEPALI G¹, ANITHA H², SWATHI B.P² and SUHAS M.V¹

¹Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, 576104, India

²School of Computer Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, 576104, India

Corresponding author: Suhas M.V(e-mail: suhas.mv@manipal.edu) and Swathi B.P(e-mail: swathi.bp@manipal.edu)

This work was supported by the Indian Council for Medical Research (ICMR) under Grant-in-aid Scheme of Department of Health Research (File No: R.11013/23/2021-GIA/HR, dated 15.11.2021)

ABSTRACT Accurate identification of fiducial markers, such as the left pre-auricle (LPA), right pre-auricle (RPA), and nasion, is critical for head localization, source reconstruction, anatomical alignment, and ensuring reproducibility in functional and structural neuroimaging studies. However, manual annotation of these landmarks is time-consuming, prone to human error, and further complicated by variations in MRI acquisition protocols and incomplete head coverage. In this work, we propose a deep learning-based framework using autoencoders to automatically detect fiducial markers in 3D brain MRI scans. The method integrates a structured preprocessing pipeline, including intensity normalization, spatial resampling, and noise reduction, to ensure data consistency. A tailored autoencoder architecture is then trained to reconstruct MRI slices while learning the spatial characteristics of MRI slices at fiducial points across axial, coronal, and sagittal planes. Fiducial localization is determined based on the slice with the lowest reconstruction error, ensuring precise identification. Our approach significantly enhances registration accuracy, achieving an average ground truth Euclidean distance of 1.42 mm for LPA and 1.89 mm for RPA, with low variance across a dataset of 500 MRI volumes. The results demonstrate that automated fiducial detection can reduce manual intervention, improve anatomical alignment, and enhance the robustness of neuroimaging workflows. This method offers a scalable and efficient solution for clinical and research applications, facilitating more reliable head modeling, source localization, and multimodal image integration.

INDEX TERMS Fiducial Marker Detection, 3D Brain MRI, MEG-MRI Registration, Autoencoder, Deep Learning, Anatomical Landmark Localization, Machine Learning, Neuroimaging, Automated Landmark Detection.

I. INTRODUCTION

ACCURATE anatomical landmark detection is critical in medical imaging for tasks such as image registration, segmentation, and surgical planning. For instance, aligning magnetoencephalography (MEG) with magnetic resonance imaging (MRI) depends on precise identification of fiducial markers like the left preauricular (LPA), right preauricular (RPA), and nasion [1], [2]. Similarly, cardiac imaging requires accurate landmark localization for heart chamber segmentation and motion tracking [3]–[6], while orthopedic imaging relies on it for joint alignment assessment and implant positioning [7]–[10]. In neurosurgery, precise fiducial detection ensures accurate electrode placement and tumor localization.

Despite their importance, fiducials are typically annotated manually by clinicians, a process prone to inter-operator variability and errors, especially in cases with incomplete head coverage or structural distortion due to tumors. Furthermore, traditional brain tumor identification methods—such as multi-atlas registration or 2D segmentation networks—often require extensive preprocessing, suffer from low generalizability across patient populations, and lack the ability to model complex 3D spatial relationships [11], [12]. Many also rely on handcrafted features, which are sensitive to noise, anatomical deformation, and varying acquisition protocols.

Recent advances in deep learning have significantly improved landmark detection performance. Convolutional Neural Networks (CNNs) have shown promise in 3D imaging

tasks by capturing hierarchical spatial features and outperforming atlas-based techniques in many scenarios [13]–[15]. However, challenges remain: existing approaches are often landmark-specific, may not generalize across datasets, and many 2D methods lack spatial continuity in 3D volumes. Descriptor-dependent models also struggle with real-world distortions, artifacts, and anatomical variability [16], [17].

To address these limitations, we propose an unsupervised deep learning framework for automated fiducial detection in 3D brain MRI, capable of simultaneously localizing the LPA, RPA, and nasion. Our method integrates a multi-plane, autoencoder-based detection strategy for LPA and RPA, combined with an anatomical-heuristic-driven localization for the nasion, enabling improved robustness under variable imaging conditions. This has potential clinical applications in MEG-MRI co-registration, brain tumor localization, and image-guided neurosurgery.

A. RELATED WORKS

Anatomical landmark detection and MEG-MRI co-registration have been widely studied using various deep learning and machine learning approaches. These methods can be broadly grouped into supervised deep learning models, alternative non-supervised or dimensionality-aware methods, MEG-specific pipelines, and recent self-configuring frameworks. Each group addresses specific challenges but also presents trade-offs in generalizability, precision, or clinical applicability.

1) Supervised CNN-Based Landmark Detection:

Recent landmark detection efforts have primarily leveraged supervised convolutional neural networks (CNNs). Prabhu et al. [18] proposed an R-CNN framework for fiducial detection in brain MRI to support MEG-MRI co-registration. Their multi-task learning approach improved performance but remained dependent on manually labeled fiducials and reported mean localization errors between 3.6–4.0 mm, indicating limited precision and adaptability to anatomical variability. Xue et al. [19] developed a 2D CNN model for cardiac landmark detection, achieving high detection rates (96.6–100%) in long- and short-axis slices. However, the reliance on 2D views restricted spatial continuity, which is essential for localizing 3D fiducials like LPA and RPA. Similarly, Yang et al. [20] applied a multi-task DNN for detecting brain midline features and commissures. While achieving point shift accuracy of 1.05 ± 0.87 pixels, their method was limited to axial and sagittal views, lacking full 3D contextual understanding. Grewal et al. [21] introduced a multi-stage CNN using region-based proposals to refine coarse predictions, but its dependency on handcrafted descriptors made it sensitive to noise and anatomical variability. In our previous work [29], we presented a 3D CNN-based approach for fiducial point detection in MRI. The method significantly improved precision and reliability, demonstrating its potential for clinical applications through extensive validation on MRI datasets.

2) Alternative Approaches for Reducing Supervision and Enhancing 3D Representation:

To overcome reliance on large annotated datasets and 2D constraints, alternative methods have explored unsupervised or geometry-driven strategies. Lin et al. [22] proposed a 3D Generalized Hough Transform (GHT) for MEG-MRI co-registration, achieving a mean RMS error of 2.88 mm without requiring training data. However, its geometric voting mechanism lacked robustness to anatomical irregularities and low-quality scans. Noothout et al. [23] introduced a global-to-local fully convolutional network (FCNN) with a multi-scale approach. Though flexible across imaging modalities, its two-stage regression pipeline struggled with sparsely distributed fiducials exhibiting low contrast. Kumar et al. [24] developed a reinforcement learning method using a Deep Q-Network (DQN) to navigate to landmarks in 3D MRI. Despite being innovative, the model's agent-environment design was sensitive to voxel discretization and computationally expensive.

3) MEG-MRI Co-registration Pipelines:

Some studies directly addressed MEG co-registration workflows. Zheng et al. [25] proposed an AI-based framework combining CNNs with clustering to support epileptic spike detection and MEG source localization. Although this approach improved registration accuracy (median error: 2.34 mm), it did not target sparse fiducial localization or integration with anatomical MRI. Houck et al. [26] evaluated manual versus automated fiducial detection using MNE-Python, achieving a small mean difference (0.313 mm). However, their method was prone to distortions from dental work or hair products, requiring manual correction and reducing clinical reliability.

4) Recent Self-Configuring and Cascaded Deep Learning Frameworks:

More recent work has shifted toward self-configuring and cascaded models. Salari et al. [27] proposed CAMLD, a coarse-to-fine CNN framework for facial landmark localization in MRI, achieving a mean radial error (MRE) of 3.27 mm on HCP T1w data. However, it still required large-scale supervised training data. Ertl et al. [28] introduced nnLandmark, a self-configuring model based on nnU-Net with heatmap regression, reporting an MRE of 1.2 mm on brain MRI. Yet, volumetric heatmap-based detection is less ideal for sparse fiducials like LPA or RPA, where direct coordinate regression is more effective.

B. GAPS AND OPPORTUNITIES

Despite notable advancements in deep learning-based anatomical landmark detection for 3D MRI, several critical limitations persist in existing methods. Most supervised approaches rely heavily on large, manually annotated datasets, which are labor-intensive to create and limit scalability across different clinical centers or imaging protocols [27], [28]. Additionally, while heatmap-based regression models such as nnLandmark [28] have shown promising accuracy for dense

landmark detection, they are less suited for sparse, isolated fiducial points like LPA, RPA, and nasion, where precise coordinate regression is often preferable to coarse volumetric heatmap estimation.

Another key challenge is model sensitivity to anatomical outliers or image artifacts. Even in recent deep learning pipelines, fiducial detection performance can degrade in the presence of skull defects, incomplete head coverage, or low-contrast regions near bony landmarks, which frequently occur in clinical MRI datasets [25], [26]. These issues are compounded by the fact that many existing models optimize for densely distributed anatomical points, making sparse detection tasks prone to higher localization error and false positives.

To address these limitations, we propose an unsupervised autoencoder-based framework for fiducial detection in 3D brain MRI. Unlike existing supervised CNN and heatmap-based regression models that rely on large annotated datasets and are optimized for dense landmark detection, our method eliminates annotation dependency by leveraging unsupervised feature learning and reconstruction-error-driven localization. The proposed framework integrates multi-plane information from axial, coronal, and sagittal slices to enable precise 3D coordinate estimation for isolated fiducial points such as LPA and RPA. Additionally, by operating directly on reconstruction error patterns within autoencoder latent spaces, the method demonstrates enhanced robustness to anatomical outliers, incomplete skull coverage, and image artifacts commonly encountered in clinical MRI datasets. This work thus offers a scalable, annotation-independent, and clinically adaptable alternative to current deep learning-based fiducial detection solutions for neuroimaging applications such as MEG-MRI co-registration and surgical navigation.

II. METHODOLOGY

This section outlines the proposed unsupervised framework for automated fiducial marker detection in 3D brain MRI. The methodology includes preprocessing of MRI volumes, slice extraction across anatomical planes, autoencoder-based reconstruction-driven detection of LPA and RPA, and heuristic-based nasion localization. A multi-plane detection strategy improves spatial precision, and the autoencoder architecture integrates regularization and sparsity constraints for enhanced robustness. Key details of the datasets, network design, training process, and detection algorithms are described below. An overview of the proposed system is shown in Figure. 1

While LPA and RPA were identified through an autoencoder-based approach, nasion localization was performed using a coordinate midpoint-based anatomical heuristic combined with cranial contour detection, as detailed in Algorithm 4. A comprehensive description of its prior validation appears in our earlier work [29].

A. DATASET

Three distinct MRI datasets from the OpenNeuro repository [32]–[34] were used to train the proposed autoencoder model. These datasets span a range of clinical and research contexts—including aging, motion artifact correction, and cognitive neurofeedback—providing the heterogeneity essential for robust neural network generalization. Each dataset includes high-resolution T1-weighted MRI scans, and the final curated dataset comprises 500 head MRI volumes, each annotated with fiducial markers. The availability of pre-labeled fiducial points eliminates the need for labor-intensive and error-prone manual annotation, making the data well-suited for training and evaluating automated landmark detection systems. Table 1 summarizes the key characteristics of the three datasets, including participant demographics, imaging specifications, modalities, and public availability.

B. PREPROCESSING PIPELINE

To ensure consistency and effective feature extraction, a structured preprocessing pipeline was applied uniformly to all 500 MRI scans included in the dataset. Each T1-weighted MRI volume underwent the following operations:

1. *Intensity Normalization*: Each scan was scaled using min-max normalization to standardize intensity values across the dataset and reduce inter-scan contrast variation [35].
2. *Morphological Refinement*: After binarization, hole filling was performed using morphological operations to eliminate small gaps within brain structures and to smooth segmentation boundaries [36].
3. *Noise Reduction*: Adaptive thresholding, based on Otsu's method, was employed to remove background noise by automatically selecting an optimal threshold minimizing intra-class variance [37].
4. *Spatial Normalization*: All scans were resampled to an isotropic voxel resolution of 1 mm³ using trilinear interpolation to ensure uniform spatial dimensions across participants [38].
5. *Fiducial Alignment*: Fiducial markers (LPA, RPA, and nasion) in each MRI were transformed into a standardized coordinate space using a rigid-body transformation [39], facilitating consistent anatomical localization across subjects.

These preprocessing steps, summarized in Algorithm 1, were applied to every MRI volume prior to autoencoder training, ensuring that intensity, spatial resolution, and anatomical orientation were harmonized across the dataset. This standardized preparation was critical for ensuring the reliability and generalizability of the proposed fiducial marker detection framework.

C. NETWORK ARCHITECTURE AND ALGORITHM

The autoencoder architecture in this work is designed to learn efficient, compressed representations of MRI slice data while ensuring accurate reconstruction of the input. The neural network follows an encoder-bottleneck-decoder design and comprises three primary components: an input layer, hidden

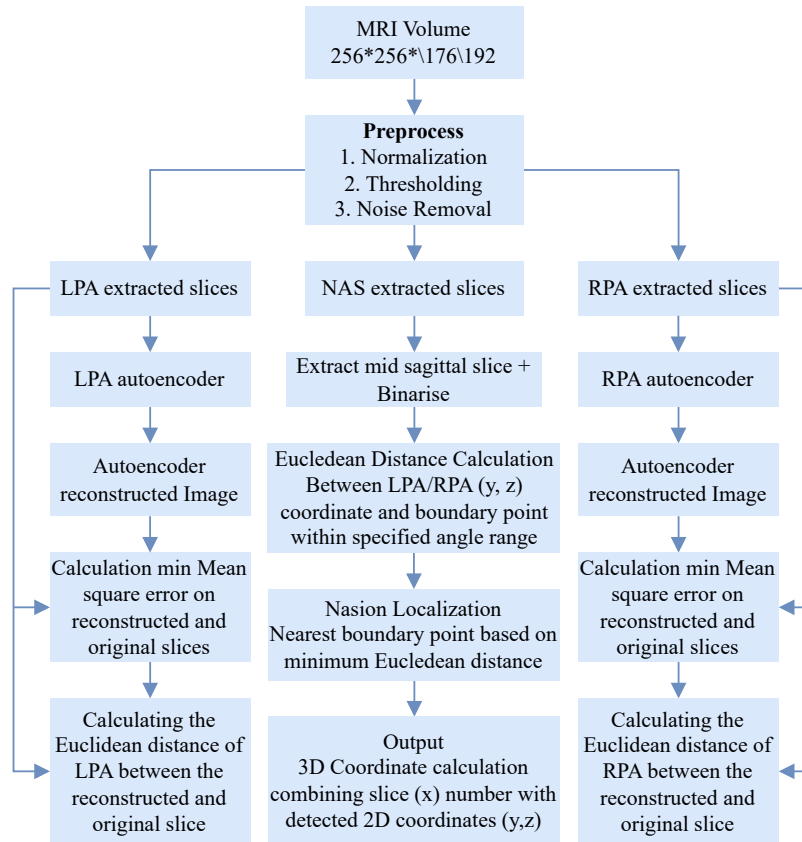


FIGURE 1. System-level overview of the proposed fiducial marker detection framework for 3D brain MRI. The diagram illustrates the complete pipeline from input to output, including (1) MRI volume preprocessing, (2) slice extraction across axial, coronal, and sagittal planes, (3) autoencoder-based reconstruction for detecting LPA and RPA via minimal reconstruction error, and (4) nasion localization using Euclidean distance and angular constraints on mid-sagittal slices. This integrated workflow enables automated, accurate, and efficient fiducial point detection for neuroimaging applications.

TABLE 1. Summary of MRI datasets used in this study

Attribute	MR-ART [32]	Neurocognitive Aging [33]	Neurofeedback Superstition [34]
Modality	T1-weighted MPRAGE	T1-weighted MPRAGE	T1-weighted MPRAGE
Image Dimensions / Protocol	1 mm ³ isotropic resolution 256 × 256 × n TR=2300ms, TE=3ms, TI=900ms	1 mm isotropic TR=1900–2530ms, TE=2.52–3.4ms	1 mm isotropic 256 × 256 × 192 TR=1650ms, TE=1.82ms
Participants / Demographics	148 healthy adults Age: 18–75 (median = 25.16) 95 females	301 adults, 181 younger (mean = 22.6), 120 older (mean = 68.6) 55% female	97 healthy adults Mean age = 27.6 (SD = 7) 50 females
Scanner / Acquisition Site	Siemens Magnetom Prisma 3T Brain Imaging Centre, Hungary	GE Discovery MR750 and Siemens TimTrio 3T, Cornell MRI Facility & York University	Siemens Skyra 3T MRI Lab Graz, Austria
Notes / Availability	440 scans (STAND, HM1, HM2) Public via OpenNeuro (ds004051) Includes motion artifacts	Multi-site aging study Behavioral + fMRI data Public via OpenNeuro (ds004636)	Anatomical + functional scans Neurofeedback study, Public via BioRxiv/OpenNeuro (ds003823)

layers (including the encoder and decoder modules), and an output layer.

The input layer accepts training data either as a cell array of images of uniform size or as a matrix where each column represents a separate data sample. Each input slice of dimension $256 \times 256 = 65,536$ features is progressively compressed into a lower-dimensional latent space of size 128, which captures the most significant structural information while discarding redundant data.

The proposed autoencoder architecture consists of an encoder and decoder module, each designed to efficiently pro-

cess MRI slice data while preserving essential anatomical features. In the encoder, each fully connected (Dense) layer is sequentially followed by Batch Normalization (BN) and a ReLU activation function, improving training stability, accelerating convergence, and mitigating internal covariate shift during optimization. The latent space representation is obtained at the bottleneck layer without applying any activation, capturing a compressed yet informative feature embedding. The decoder mirrors this structure, progressively expanding the latent representation back to the original input size. Similar to the encoder, each Dense layer in the decoder is followed

Algorithm 1 Preprocessing Pipeline for MRI Volume

```

1: function PreprocessMRI( $V$ )
2:   // Step 1: Intensity normalization
3:   Apply min-max normalization to standardize intensity values
4:   // Step 2: Morphological refinement
5:   Apply morphological hole filling to remove small gaps
6:   // Step 3: Noise reduction
7:   Perform adaptive thresholding using Otsu's method
8:   // Step 4: Spatial normalization
9:   Resample to isotropic voxel size of  $1\text{mm}^3$  via trilinear interpolation
10:  // Step 5: Fiducial alignment
11:  Apply rigid-body transformation to standardize fiducial positions
12:  return Preprocessed MRI volume  $V'$ 
13: end function

```

by Batch Normalization and ReLU, with a final Sigmoid activation applied at the output layer to constrain reconstructed pixel intensities to the range $[0, 1]$. To further enhance generalization and model interpretability, L2 regularization was imposed on the network weights, and a sparsity constraint using Kullback–Leibler (KL) divergence was applied to the latent activations, encouraging compact, meaningful feature encodings.

The model is trained by minimizing the Mean Squared Error (MSE) loss between the input and reconstructed output. The use of Batch Normalization throughout the encoder and decoder contributes to faster convergence and improved robustness against intensity variations and anatomical differences in MRI data.

The choice of an autoencoder architecture over standard CNN, FCN, or transformer-based approaches is motivated by the nature of the task. Our goal is unsupervised fiducial detection—learning to reconstruct image patches and infer landmark locations without requiring manually labeled ground truth for every slice. Unlike CNNs or transformers, which require dense labels or annotated keypoints, autoencoders can capture intrinsic image features by minimizing reconstruction loss. This enables the model to focus on the most relevant anatomical patterns, even when data is partially incomplete or noisy. Moreover, autoencoders offer lower computational overhead compared to transformer-based models and require fewer parameters, making them efficient for volumetric brain data. Similar approaches have been successfully employed for learning from incomplete and noisy medical data in recent literature [30], further validating the suitability of this design choice for landmark localization.

1) Encoder

The encoder compresses the input data into a lower-dimensional latent representation, facilitating dimensionality reduction while preserving essential anatomical features. Let

the input MRI slice be represented as a vector x of dimension 65,536, and the latent representation be denoted as z of dimension 128. The encoding operation through each encoder layer is performed as:

$$h_i = \text{ReLU}(\text{BN}(W_{e_i}h_{i-1} + b_{e_i})) \quad (1)$$

where W_{e_i} and b_{e_i} are the weight matrix and bias vector for the i -th encoder layer, BN denotes Batch Normalization, and ReLU is the activation function applied element-wise:

$$\text{ReLU}(x) = \max(0, x) \quad (2)$$

for $i = 1, 2, \dots, n-1$, with $h_0 = x$.

At the bottleneck layer (the final encoder layer), the latent representation z is computed as:

$$z = W_{e_n}h_{n-1} + b_{e_n} \quad (3)$$

No activation function is applied at this stage, allowing the latent space to preserve a linear encoding of essential features.

2) Latent Representation

The bottleneck layer represents the most compressed form of the data, reducing the 65,536-dimensional input to a compact 128-dimensional latent vector z . This representation acts as a constraint, ensuring that only the most salient structural features are preserved for downstream reconstruction. The latent vector captures global and local structural patterns while filtering out irrelevant noise and redundant details.

3) Decoder

The decoder reconstructs the original input from the latent representation z , progressively expanding it back to its original dimension. The decoding operation through each decoder layer is expressed as:

$$h_j = \text{ReLU}(\text{BN}(W_{d_j}h_{j-1} + b_{d_j})) \quad (4)$$

for $j = 1, 2, \dots, m-1$, with $h_0 = z$.

At the final output layer, a sigmoid activation is applied to constrain the reconstructed pixel intensities to the range $[0, 1]$:

$$\hat{x} = \text{Sigmoid}(W_o h_{out} + b_o) \quad (5)$$

This ensures that the reconstructed MRI slice \hat{x} closely approximates the original input x , effectively minimizing the reconstruction loss during training.

D. AUTOENCODER TRAINING

A total of six autoencoders were trained using axial, coronal, and sagittal slices for LPA and RPA detection. The detailed training procedure is outlined in Algorithm 2, which describes the iterative optimization process, regularization, and sparsity constraints applied during training. The autoencoder was optimized to minimize the reconstruction error using the Mean Squared Error (MSE) loss function:

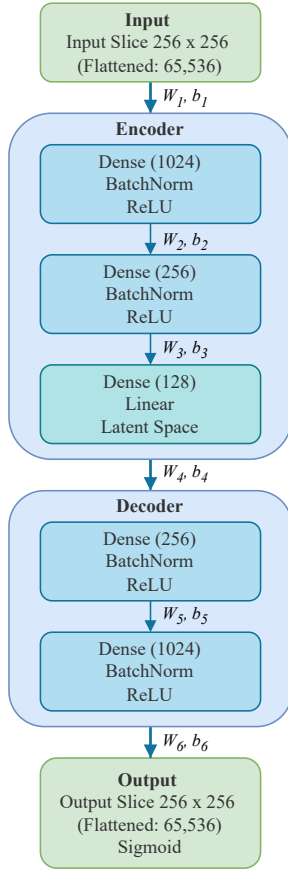


FIGURE 2. Schematic of the autoencoder architecture used for fiducial point detection. The encoder compresses the input data from 65,536 dimensions to a latent space of 128 dimensions, while the decoder reconstructs the data back to its original size. Each stage involves weight matrices (W), biases (b), and activation functions to capture and represent the input features.

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i|^2 \quad (6)$$

where N is the number of training samples, x_i is the original input, and \hat{x}_i is the reconstructed output.

1) Regularization

To enhance generalization and prevent overfitting, regularization techniques were incorporated:

$L2$ Regularization: Penalizes large weight values by adding a term to the loss function:

$$\mathcal{L}_{L2} = \lambda \sum_{i,j} W_{i,j}^2 \quad (7)$$

where λ is the regularization coefficient.

Sparsity Regularization: Encourages selective neuron activation, improving feature selection. This is enforced using a Kullback-Leibler (KL) divergence penalty:

$$\mathcal{L}_{sparsity} = \sum_j \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j} \quad (8)$$

Algorithm 2 Training Regularized Autoencoder for Fiducial Detection

```

1: function TrainAutoencoder( $\mathcal{D}$ )
2:   Initialize encoder and decoder weights with small
   random values
3:   Apply L2 regularization to weights and biases
4:   Apply sparsity constraint on latent layer activations
   (KL divergence)
5:   for each epoch do
6:     for each batch of MRI slices in  $\mathcal{D}$  do
7:       Normalize each slice to  $[0, 1]$ 
8:       Forward pass:
       • Dense + ReLU + BatchNorm in encoder
       • Encode to latent space (Linear activation)
       • Dense + ReLU + BatchNorm in decoder
       • Output reconstruction via Sigmoid activation
9:     Compute total loss:
        $\mathcal{L}_{total} = \text{MSE Loss} + \lambda \cdot \text{L2 Reg.} + \beta \cdot \text{KL Sparsity}$ 
10:    Backpropagate gradients and update parameters
11:   end for
12: end for
13:   return Trained Autoencoder model  $\mathcal{A}^*$ 
14: end function

```

where ρ is the desired average activation, and $\hat{\rho}_j$ is the actual activation of hidden unit j .

The total loss function used for training was:

$$\mathcal{L}_{total} = \mathcal{L}_{MSE} + \lambda \mathcal{L}_{L2} + \beta \mathcal{L}_{sparsity} \quad (9)$$

where β controls the sparsity penalty.

2) Training Procedure

The extracted MRI dataset was split into training, validation, and test sets in a 70%-15%-15% ratio. Each MRI slice was preprocessed to a uniform resolution of 256×256 pixels, preserving fiducial markers across axial, coronal, and sagittal planes.

Using MATLAB's `trainAutoencoder` function, training was conducted (Algorithm 2) with the following hyperparameters:

- **HiddenSize:** Latent space dimension, set to 128.
- **MaxEpochs:** Number of training iterations, set to 100.
- **$L2$ Regularization:** Weight decay to mitigate overfitting.
- **Optimization Algorithm:** Scaled Conjugate Gradient (SCG).

SCG is a second-order optimization method that approximates Newton's method without explicitly computing the Hessian, offering faster and more stable convergence than standard gradient descent, particularly for networks incorporating regularization and sparsity constraints. Unlike conventional gradient descent, SCG automatically adjusts the effective step size during optimization without requiring a fixed learning rate.

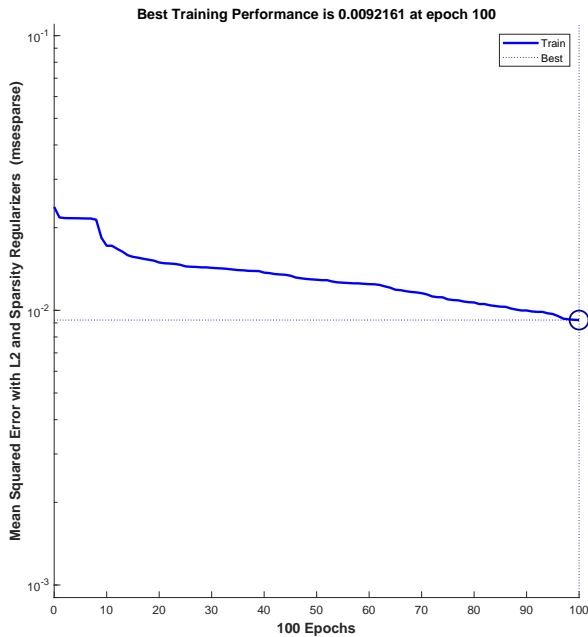


FIGURE 3. Training loss progression over 100 epochs. The Mean Squared Error (MSE) steadily decreases, indicating effective learning and convergence.

The network weights were updated using:

$$W_{t+1} = W_t - \eta \nabla \mathcal{L}_{\text{total}}(W_t) \quad (10)$$

where η is the adaptive step size and $\nabla \mathcal{L}_{\text{total}}(W_t)$ is the gradient of the total loss function, with respect to the weights at iteration t .

3) Training Convergence

The autoencoder was trained for 100 epochs, with loss values monitored at each iteration. As shown in Figure 3, the MSE consistently decreased, indicating successful training. The best training performance was achieved at epoch 100, with a final MSE of 0.0092161.

The use of L2 regularization and sparsity constraints contributed to stable weight optimization and prevented overfitting. Convergence was confirmed using the validation set, with early stopping applied if performance degradation was detected. The complete fiducial detection process, including slice-wise reconstruction error computation and coordinate selection, is presented in Algorithm 3. The detected fiducial points from the slice-level reconstructions were combined to obtain full 3D coordinates, ensuring precise anatomical alignment and robust landmark localization.

4) Multi-Plane Detection Strategy

While the autoencoder architecture processes individual 2D MRI slices, the proposed framework achieves 3D spatial localization through a structured multi-plane detection strategy. Specifically, six independent autoencoders were trained — one each for the axial, coronal, and sagittal slices corresponding to LPA and RPA localization. For each fiducial, the slice

Algorithm 3 Fiducial Detection using Trained Autoencoder

```

1: function DetectFiducial( $V', \mathcal{A}^*$ )
2:   for all planes  $p \in \{\text{axial, coronal, sagittal}\}$  do
3:     Extract slice stack  $S_p = \{s_1^p, \dots, s_n^p\}$ 
4:     for each slice  $s_i^p$  do
5:       Normalize  $s_i^p$ 
6:       Compute reconstructed slice  $\hat{s}_i^p = \mathcal{A}^*(s_i^p)$ 
7:       Compute reconstruction error
8:        $e_i^p \leftarrow \text{MSE}(s_i^p, \hat{s}_i^p)$ 
9:     end for
10:    Identify slice with minimum error:  $i_p^* \leftarrow \arg \min_i e_i^p$ 
11:  end for
12:  return 3D coordinate:  $\hat{p} = (i_{\text{coronal}}^*, i_{\text{sagittal}}^*, i_{\text{axial}}^*)$ 
13: end function

```

Algorithm 4 Nasion Detection from LPA and RPA

```

1: function DetectNasion( $V', \hat{p}_{\text{LPA}}, \hat{p}_{\text{RPA}}$ )
2:   Compute midpoint:  $m = \frac{\hat{p}_{\text{LPA}} + \hat{p}_{\text{RPA}}}{2}$ 
3:   Extract mid-sagittal slice at coordinate  $m$ 
4:   Perform edge detection to get cranial contour
5:   Identify the most anterior point (closest to forehead)
6:   return Nasion coordinate  $\hat{p}_{\text{NAS}}$ 
7: end function

```

with the minimum reconstruction error in each plane was selected as the likely anatomical location. These detected indices from orthogonal planes were then aggregated to compute the final 3D coordinate of the fiducial marker. This multi-plane fusion approach effectively mitigates the limitations of slice-wise processing by integrating complementary spatial cues from multiple anatomical orientations, enabling volumetric localization without requiring a 3D network architecture. This strategy reduces computational overhead while maintaining spatial consistency in 3D fiducial estimation.

The nasion detection was performed following a midline-based strategy, wherein the midpoint between the detected LPA and RPA coordinates was used to identify the mid-sagittal plane. The most anterior point on the cranial contour within this plane was selected as the nasion. The detailed computational procedure for this detection is outlined in Algorithm 4.

III. EXPERIMENTS AND RESULTS

The detection of the LPA and RPA points is carried out independently by the 6 distinct autoencoders. The detected slices in 3 planes corresponding to LPA and RPA and the calculated difference between the ground truth and estimated coordinates is shown in figure 4 and 5 respectively. Once estimated, the fiducial points are displayed over the 3D MRI for visual observation as shown in figure 6.

Table 2 presents a sample of detected fiducial points, their corresponding ground truth coordinates, and the Euclidean distance error in millimeters (mm). These results confirm the precision of the method, as demonstrated in Figures 4 and 5,

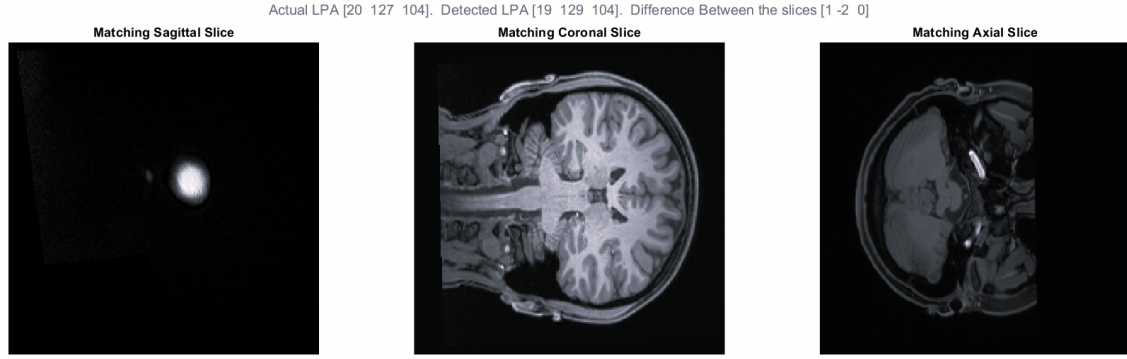


FIGURE 4. Detected slices of LPA in 3 planes with the Euclidean distance calculated.

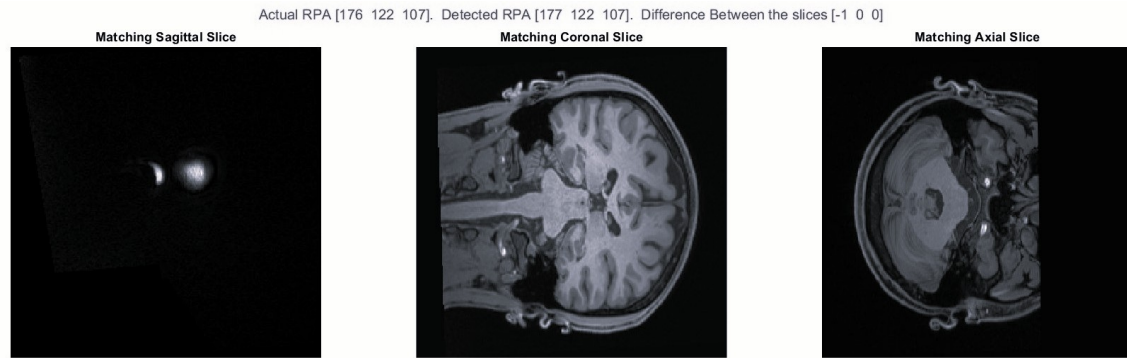


FIGURE 5. Detected slices of RPA in 3 planes with the Euclidean distance calculated.

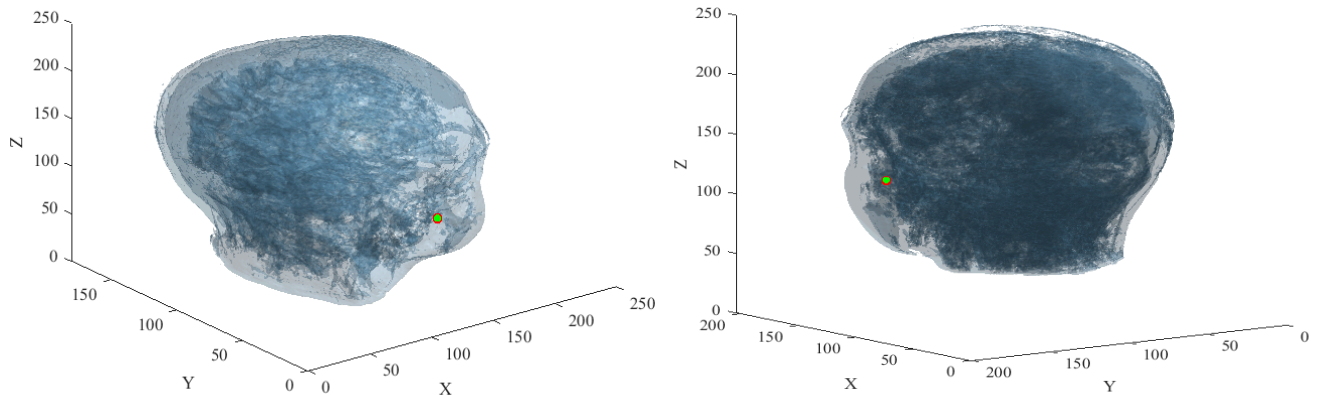


FIGURE 6. Visualization of LPA (left) and RPA (right) point detection on a 3D MRI. The red dot represents the ground truth, and the green dot represents the detected points,

where the detected fiducial points closely match their ground truth locations.

A. ANALYSIS OF PERFORMANCE METRICS

Once the fiducial marker is detected, its estimated coordinates (x_d, y_d, z_d) are compared against the ground truth coordinates (x_t, y_t, z_t) from the dataset. The accuracy of the detection is evaluated using the Euclidean distance which provides an objective measure of the deviation between the detected and actual fiducial points:

$$d = \sqrt{(x_t - x_d)^2 + (y_t - y_d)^2 + (z_t - z_d)^2}. \quad (11)$$

To further assess the accuracy of the proposed fiducial detection method, we computed key statistical metrics, including the mean Euclidean distance (ED), standard deviation of ED, mean squared error (MSE), standard deviation of squared error (SSE), and 95% confidence intervals (CI). Additionally, effect size (Cohen's d) was used to compare LPA and RPA detection accuracy, and outlier analysis was conducted to evaluate the influence of extreme values.

TABLE 2. Sample list of ground truth and detected fiducial points along with their difference

Fiducial	Ground truth	Detected	Difference
LPA	[14 118 97]	[14 118 97]	[0 0 0]
	[18 130 114]	[18 129 114]	[0 1 0]
	[15 120 96]	[16 120 95]	[-1 0 1]
	[19 126 109]	[21 126 109]	[-2 0 0]
	[20 127 104]	[21 129 104]	[-1 -2 0]
	[22 132 104]	[21 129 103]	[1 3 1]
RPA	[176 118 97]	[176 118 97]	[0 0 0]
	[175 117 112]	[175 117 111]	[0 0 1]
	[175 117 91]	[175 117 93]	[0 0 -2]
	[172 121 109]	[169 121 109]	[3 0 0]
	[172 132 114]	[175 132 115]	[-3 0 -1]
	[174 129 107]	[178 129 107]	[-4 0 0]

TABLE 3. Performance evaluation metrics for LPA and RPA detection.

Metrics	LPA (mm)	RPA (mm)
Mean ED	1.66	2.19
95% CI (ED)	[1.42, 1.90]	[1.92, 2.46]
Standard Deviation	1.04	1.36
Mean of MSE	1.28	2.22
Standard Deviation of Squared Error	0.67	0.807
Cohen's d (Effect Size)	-0.44 (Moderate)	—

Table 3 provides a comprehensive summary of detection accuracy across the entire dataset. The lower Euclidean distances for LPA detection indicate higher precision, while slightly higher values for RPA suggest minor variability. The performance metrics for LPA and RPA, based on the Euclidean distances provided, show that the LPA has a mean Euclidean distance of 1.66 mm with a standard deviation of 1.04 mm. This indicates that the detected points for LPA are generally close to the ground truth with relatively low variability. On the other hand, the RPA has a mean Euclidean distance of 2.19 mm and a standard deviation of 1.36 mm, suggesting that the detected points for RPA are slightly further from the ground truth and exhibit more variability compared to LPA. The Euclidean distance error distributions for LPA and RPA detection was visualized across 500 MRI datasets. The histogram in Figure 7 provides insight into the overall accuracy and error variability of the method. From the histogram, we observe that the majority of LPA errors fall within the 0–2 mm range, with a gradual decline in frequency as the error increases. This suggests that the LPA detection method is highly precise for most subjects. In contrast, RPA errors show a wider spread, with a noticeable proportion of subjects exhibiting errors in the 2–4 mm range, indicating slightly higher variability in detection. The presence of a rightward skew in both distributions highlights occasional cases where errors exceed 4 mm. These larger errors may be due to individual anatomical differences, limitations in image preprocessing, or challenging fiducial placements.

1) Confidence Intervals (CI) for Euclidean Distance

The 95% confidence interval (CI) for the LPA detection ED ranges from 1.42 mm to 1.90 mm, while the CI for RPA detection extends from 1.92 mm to 2.46 mm. The non-overlapping

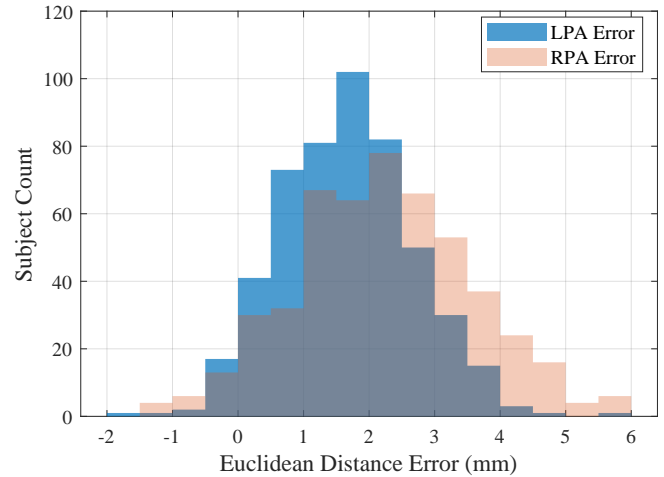


FIGURE 7. Distribution of Euclidean Distance Errors for LPA and RPA Detection. This histogram illustrates the distribution of Euclidean distance errors (in mm) for the detected LPA and RPA across 500 MRI volumes. The x-axis represents the Euclidean distance error (mm), while the y-axis shows the number of subjects exhibiting that error. The LPA errors (blue bars) tend to cluster around lower values, indicating higher precision, whereas the RPA errors (orange bars) show slightly more variability.

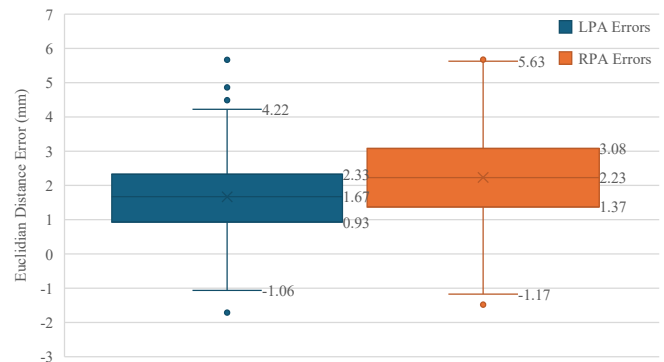


FIGURE 8. Boxplot comparison of Euclidean distance errors for LPA and RPA detection. The box represents the interquartile range (IQR), with the median shown as a horizontal line. Whiskers extend to 1.5 times the IQR, and outliers are displayed as individual points. This visualization highlights the variability in detection accuracy between LPA and RPA fiducial points.

confidence intervals suggest a statistically significant difference in the accuracy of LPA and RPA detection.

2) Effect Size Analysis (Cohen's d)

To quantify the difference in detection accuracy between LPA and RPA, Cohen's $d = -0.44$ was calculated, indicating a moderate effect size. This result suggests that the LPA detection method achieves more precise localization compared to RPA.

3) Outlier Analysis

While the standard deviations indicate a controlled variability in the results, further examination of the Euclidean Distance distribution showed a few outliers (> 3 SD from the mean). These outliers were attributed to anatomical variations in specific MRI volumes. However, their impact on the mean ED values was minimal, as confirmed by sensitivity analysis.

Figure 8 presents a boxplot comparison of Euclidean distance errors for LPA and RPA detection. The LPA errors exhibit a lower median and interquartile range, suggesting more consistent and precise detection compared to RPA. The wider spread and presence of outliers in the RPA error distribution indicate higher variability, which may be attributed to anatomical differences or algorithmic challenges in detecting this fiducial point. The presence of outliers underscores the importance of further refining the detection algorithm or implementing additional preprocessing steps to enhance robustness.

IV. DISCUSSION

This study introduced an unsupervised autoencoder-based framework for automated fiducial marker detection in 3D brain MRI, focusing on LPA and RPA localization. The method achieved promising detection accuracy, with mean Euclidean distances of 1.66 mm for LPA and 2.19 mm for RPA, confirming the reliability of reconstruction-error-driven localization even in the absence of annotated training data.

A notable observation was the performance disparity between LPA and RPA detection. LPA points consistently exhibited lower mean errors and narrower interquartile ranges, while RPA detections demonstrated increased variability and broader error distributions. This discrepancy likely reflects inherent anatomical asymmetries, occasional incomplete head coverage, and MRI acquisition artifacts affecting one auricular region more prominently. Visualizations of error distributions, including histograms and boxplots, reinforced these findings by revealing rightward-skewed distributions and a small number of outliers.

Another key strength of this framework is its adaptability to unlabeled, artifact-prone clinical datasets — a challenge for conventional supervised CNN, R-CNN, and multi-atlas registration-based approaches that require extensive annotations. By leveraging unsupervised reconstruction loss, multi-plane analysis, and anatomical heuristics for nasion detection, the proposed system offers a scalable, efficient, and practical solution for integration into neuroimaging workflows such as MEG-MRI co-registration, neurosurgical planning, and intraoperative guidance. Its scalability stems from the slice-wise autoencoder architecture, which eliminates the need for large annotated datasets, reduces computational overhead compared to volumetric CNNs, and generalizes effectively across diverse MRI protocols and anatomical variations, as demonstrated on multi-site, multi-scanner data.

Comprehensive statistical analyses, including confidence intervals and Cohen's *d* effect sizes, confirmed consistent detection accuracy across cases. Outlier sensitivity analyses further demonstrated minimal influence of extreme values on overall performance, supporting the robustness of this unsupervised approach for anatomical landmark detection in high-dimensional brain MRI.

A. COMPARATIVE ANALYSIS

Table 4 presents a summary of recent automated techniques for fiducial marker detection and MEG-MRI co-registration reported in the literature. It is important to note that the performance metrics listed are drawn from different datasets, imaging modalities, and clinical applications, and are therefore not intended for direct quantitative comparison. Rather, this table serves as a qualitative overview to contextualize the diversity of methodological approaches and reported outcomes in the field.

Conventional supervised approaches, including CNN and R-CNN models, achieve reliable results but rely heavily on annotated datasets. Several recent methods such as CAMLD [27] and nnLandmark [28] have explored self-supervised and heatmap regression strategies for improved landmark detection. In contrast, our proposed unsupervised, reconstruction-error-driven autoencoder framework demonstrates clinically relevant localization accuracy for fiducial marker detection in unlabeled 3D MRI data. Its annotation-free nature, lower computational overhead, and multi-plane detection strategy distinguish it as a practical alternative for neuroimaging workflows such as MEG-MRI co-registration, especially in data-scarce clinical settings.

B. LIMITATIONS AND FUTURE DIRECTIONS

The use of a simple, fully connected autoencoder architecture in this study was a deliberate design decision, prioritizing computational simplicity, unsupervised learning, and ease of integration into annotation-scarce clinical workflows over architectural novelty. While the proposed framework demonstrates strong performance in automated fiducial marker detection, several limitations must be acknowledged. A primary constraint is the loss of intra-slice spatial information due to the flattening of 2D MRI slices into 1D vectors for processing by fully connected layers. This design choice, though computationally efficient, inherently discards local spatial dependencies that could improve landmark localization accuracy, especially in ambiguous or artifact-prone scans.

Another observed limitation lies in the detection performance disparity between LPA and RPA points. The increased variability and broader error distribution associated with RPA detection can be attributed to inherent anatomical asymmetries, inconsistent head coverage, and imaging artifacts affecting one auricular region more prominently than the other.

Additionally, while the multi-plane slice-wise detection strategy partially compensates for volumetric context loss, a fully volumetric model might further enhance robustness and precision. Future work should therefore prioritize the development of volumetric autoencoder architectures or spatially-aware convolutional autoencoders capable of preserving intra-slice and inter-slice structural relationships. Incorporating attention mechanisms, multi-scale feature extraction, and context-aware 3D feature encoding could also improve landmark localization, particularly for anatomically ambiguous regions.

TABLE 4. Summary of automated techniques for fiducial marker detection and MEG-MRI co-registration from recent literature. This table provides a qualitative reference to contextualize methodological trends and reported outcomes in the field.

Work	Objective	Methodology	Performance Metrics
Yang et al. (2020) [20]	Multi-task deep learning for anatomical landmark detection in brain MRI	Multi-task deep neural network trained to detect multiple landmarks by leveraging shared features.	Accuracy for point-detection: 2.09 ± 2.13 pixels; angle-detection: 0.95 ± 0.92 degrees. Multi-task model improved accuracy with a point shift of 1.05 ± 0.87 pixels and angle shift of 0.78 ± 0.69 degrees.
Lin et al. (2020) [22]	MEG-MRI co-registration using 3D Generalized Hough Transform (GHT)	3D GHT-based approach for MEG-MRI alignment, optimizing transformation parameters for co-registration.	MEG-MRI alignment achieved an average RMS error of 2.88 mm, demonstrating reliable co-registration performance.
Xue et al. (2021) [19]	CNN-based landmark detection in cardiac MRI	Supervised learning with MSE loss, optimized using Adam. Performance evaluated using Euclidean distance and localization accuracy.	Detection accuracy: 99.7–100% (cine) and 99.2–99.5% (LGE) for long-axis images; 96.6% (cine), 97.6% (LGE), and 98.7% (T1 mapping) for short-axis images. Euclidean distance: 2–3.5 mm.
Prabhu et al. (2022) [18]	Automated landmark detection for MEG-MRI registration using R-CNN	R-CNN model trained for landmark detection, combined with image processing for nasion localization.	Mean landmark detection errors ranged from 3.6028 ± 1.4037 mm to 4.0512 ± 1.736 mm, indicating challenges in precise localization.
Zheng et al. (2023) [25]	AI-based MEG source localization using CNN and clustering techniques	AI-based autolabeling framework combining CNNs and clustering for detecting epileptic sources.	MEG source localization improved co-registration accuracy with a median error of 2.34 mm, significantly outperforming manual methods.
Suhas et al. (2024) [29]	Fiducial marker detection using a 3D CNN	3D CNN model trained on a labeled MRI dataset with manually annotated fiducial points, applying supervised learning with an MSE loss function.	Accuracy for fiducial point detection: LPA – 96.88% (F1 score: 97%), RPA – 96.55% (F1 score: 96.67%). Mean Euclidean distances: LPA: 2.88 ± 1.11 mm, RPA: 2.82 ± 1.01 mm.
Salari et al. (2024) [27]	Contrast-agnostic anatomical landmark detection	Self-supervised 3D CNN framework (CAMLD) with landmark consistency loss and contrast augmentation for unlabeled MRI datasets	HCP T1w: MRE 3.27 ± 2.24 mm, SDR 54.48% (3 mm), 93.69% (6 mm), 98.94% (9 mm). Statistically outperformed supervised 3D CNN and DL baselines ($p < 0.05$)
Ertl et al. (2025) [28]	3D medical landmark detection	Self-configuring deep learning framework (nnLandmark) adapted from nnU-Net, utilizing heatmap-based regression without manual parameter tuning.	Achieved MRE of 1.5 mm on the MML dental CT dataset and 1.2 mm on the AFIDs brain MRI dataset, aligning with inter-rater variability
Present Work	Automated fiducial marker detection in 3D brain MRI using slice detection and autoencoder framework	Unsupervised autoencoder-based framework trained on MRI slice datasets. Fiducial markers identified by minimizing MSE between predicted and true fiducial locations.	Mean Euclidean distances: LPA: 1.66 ± 1.04 mm, RPA: 2.19 ± 1.36 mm. Demonstrates improved accuracy compared to previous fiducial detection methods.

Moreover, expanding the framework's adaptability to accommodate diverse MRI acquisition protocols, imaging resolutions, and scanner platforms is essential for broader clinical applicability. Fine-tuning the system on multi-institutional and multi-scanner datasets would enhance its generalizability across heterogeneous clinical environments.

Lastly, future research should focus on integrating this system into real-world clinical neuroimaging pipelines, including MEG-MRI co-registration, neurosurgical planning, and intraoperative navigation. Prospective validation studies in clinical workflows will be crucial to confirm its robustness, scalability, and impact in routine clinical practice.

V. CONCLUSION

This study presented an unsupervised, reconstruction-error-driven framework for automated fiducial marker detection in 3D brain MRI, eliminating the need for manual annotation. Through multi-plane detection and anatomical heuristics, the method achieved clinically relevant localization accuracy on a diverse set of T1-weighted MRI scans. While occasional outliers affected error distributions, the overall detection performance remained consistent and reliable. Future work will explore volumetric and spatially-aware architectures to enhance robustness and extend applicability to multi-institutional clinical workflows.

REFERENCES

- [1] K. D. Singh, I. E. Holliday, P. L. Furlong, G. F. A. & Harding, "Evaluation of MRI-MEG/EEG co-registration strategies using Monte Carlo simulation," *Electroencephalography and Clinical Neurophysiology*, vol. 102, no. 2, pp. 81–85, Feb. 1997. doi: 10.1016/s0921-884x(96)96570-4
- [2] M. Theiß, A. Wollbrink, C. H. Wolters, and H. Brinck, "Impact of different registration methods in MEG source analysis," *Current Directions in Biomedical Engineering*, vol. 2, no. 1, pp. 463–466, Sep. 2016. doi: 10.1515/cdbme-2016-0102
- [3] W. Ding, L. Li, L. Huang, and X. Zhuang, "Unsupervised multi-modality registration network based on spatially encoded gradient information," in *International Workshop on Statistical Atlases and Computational Models of the Heart*, vol. 13131, pp. 151–159, Sep. 2021. doi: 10.48550/arXiv.2105.07392
- [4] Z. L. Sandoval and J.-L. Dillenseger, "Evaluation of computed tomography to ultrasound 2D image registration for atrial fibrillation treatment," in *Computing in Cardiology 2013*, pp. 245–248, Sep. 2013.
- [5] Khalil, A., Faisal, A., Lai, K. W., Ng, S. C., and Liew, Y. M. "2D to 3D fusion of echocardiography and cardiac CT for TAVR and TAVI image guidance," *Medical & Biological Engineering & Computing*, vol. 55, pp. 1317–1326, Aug. 2017. doi: 10.1007/s11517-016-1594-6
- [6] A. Khalil, A. Faisal, S. C. Ng, Y. M. Liew, and K. W. Lai, "Multimodality registration of two-dimensional echocardiography and cardiac CT for mitral valve diagnosis and surgical planning," *Journal of Medical Imaging*, vol. 4, no. 3, pp. 037001–037001. Jul. 2017. doi: 10.1117/1.JMI.4.3.037001
- [7] B. Schlatterer, I. Suedhoff, X. Bonnet, Y. Catonne, M. Maestro, and W. Skalli, "Skeletal landmarks for TKR implantations: evaluation of their accuracy using EOS imaging acquisition system," *Orthopaedics & Traumatology: Surgery & Research*, vol. 95, no. 1, pp. 2–11, Feb. 2009. doi: 10.1016/j.otsr.2008.05.001
- [8] S. E. Kim, J. W. Nam, J. I. Kim, J. K. Kim, and D. H. Ro, "Enhanced deep learning model enables accurate alignment measurement across diverse

- institutional imaging protocols," *Knee Surgery & Related Research*, vol. 36, no. 1, p. 4, Jan. 2024. doi: 10.1186/s43019-023-00209-y
- [9] T. A. Nees, C. T. Mueller, M. M. Innmann, D. M. Spranz, F. Westhauser, T. Renkawitz, T. Reiner and T. Walker, "Assessment accuracy of 2D vs. 3D imaging for custom-made acetabular implants in revision hip arthroplasty," *Journal of Personalized Medicine*, vol. 14, no. 8, p. 808. Jul 2024. doi: 10.3390/jpm14080808
- [10] J. Negrillo-Cárdenas, J. R. Jiménez-Pérez, H. Cañada-Oya, F. R. Feito, and A. D. Delgado-Martínez, "Automatic detection of landmarks for the analysis of a reduction of supracondylar fractures of the humerus," *Medical Image Analysis*, vol. 64, p. 101729, Aug. 2020. doi: 10.1016/j.media.2020.101729
- [11] L. Sun, L. Zhang, and D. Zhang, "Multi-atlas based methods in brain MR image segmentation," *Chinese Medical Sciences Journal*, vol. 34, no. 2, pp. 110–119, Jun. 2019. doi: 10.24920/003576
- [12] J. Zhang, M. Liu, and D. Shen, "Detecting anatomical landmarks from limited medical imaging data using two-stage task-oriented deep neural networks," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4753–4764, Jun. 2017. doi: 10.1109/TIP.2017.2721106
- [13] Y. Zheng, D. Liu, B. Georgescu, H. Nguyen, and D. Comaniciu, "3D deep learning for efficient and robust landmark detection in volumetric data," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany*, pp. 565–572, Oct. 2015. doi: 10.1007/978-3-319-24553-9_69
- [14] A. Criminisi, D. Robertson, O. Pauly, B. Glocker, E. Konukoglu, J. Shotton, D. Mateus, A. Martinez Möller, S. G. Nekolla, and N. Navab, "Anatomy detection and localization in 3D medical images," in *Decision Forests for Computer Vision and Medical Image Analysis*, pp. 193–209, 2013. doi: 10.1007/978-1-4471-4929-3_14
- [15] M. Grewal, T. M. Deist, J. Wiersma, P. A. N. Bosman, and T. Alderliesten, "An end-to-end deep learning approach for landmark detection and matching in medical images," *Medical Imaging 2020: Image Processing*, vol. 11313, pp. 548–557, Mar. 2020. doi: 10.1117/12.2549302
- [16] Y. Gao and D. Shen, "Collaborative regression-based anatomical landmark detection," *Physics in Medicine & Biology*, vol. 60, no. 24, p. 9377, Nov. 2015. doi: 10.1088/0031-9155/60/24/9377
- [17] A. Fotaki, E. Puyol-Antón, A. Chiribiri, R. Botnar, K. Pushparajah, and C. Prieto, "Artificial intelligence in cardiac MRI: is clinical adoption forthcoming?," *Frontiers in Cardiovascular Medicine*, vol. 8, p. 818765, Jan. 2022. doi: 10.3389/fcvm.2021.818765
- [18] P. Prabhu, K. A. Kotegar, N. Mariyappa, G. K. Bhargava, J. Saini, and S. Sinha, "A novel approach to detect anatomical landmarks using R-CNN for MEG-MRI registration," Available at <http://dx.doi.org/10.2139/ssrn.4139903>. 2022
- [19] H. Xue, J. Artico, M. Fontana, J. C. Moon, R. H. Davies, and P. Kellman, "Landmark detection in cardiac MRI by using a convolutional neural network," *Radiology: Artificial Intelligence*, vol. 3, no. 5, p. e200197, Jul. 2021. doi: 10.1148/ryai.2021200197
- [20] X. Yang, W. T. Tang, G. Tjio, S. Y. Yeo, and Y. Su, "Automatic detection of anatomical landmarks in brain MR scanning using multi-task deep neural networks," *Neurocomputing*, vol. 396, pp. 514–521, Jul. 2020. doi: 10.1016/j.neucom.2018.10.105
- [21] M. Grewal, J. Wiersma, H. Westerveld, P. A. N. Bosman, and T. Alderliesten, "Automatic landmark correspondence detection in medical images with an application to deformable image registration," *Journal of Medical Imaging*, vol. 10, no. 1, pp. 014007–014007, Jan. 2023. doi: 10.1117/1.JMI.10.1.014007
- [22] S. K. Lin, R. C. Lo, and R. G. Lee, "MEG-MRI co-registration using 3D generalized Hough transform," *Biomedical Engineering: Applications, Basis and Communications*, vol. 32, no. 4, p. 2050028, Aug. 2020. doi: 10.4015/S1016237220500283
- [23] J. M. H. Niothout, B. D. De Vos, J. M. Wolterink, E. M. Postma, P. A. M. Smeets, R. A. P. Takx, T. Leiner, and I. Išgum, "Deep learning-based regression and classification for automatic landmark localization in medical images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 4011–4022, Jul. 2020. doi: 10.1109/TMI.2020.3009002
- [24] Kumar, Y., and Kumar, P. "Anatomical landmark detection in 3D MRI scan using deep neuro-dynamic programming," *Procedia Computer Science*, vol. 235, pp. 1713–1721 Jan. 2024. doi: 10.1016/j.procs.2024.04.162
- [25] L. Zheng, P. Liao, X. Wu, M. Cao, W. Cui, L. Lu, H. Xu, L. Zhu, B. Lyu, and X. Wang, "An artificial intelligence-based pipeline for automated detection and localisation of epileptic sources from magnetoencephalography," *Journal of Neural Engineering*, vol. 20, no. 4, p. 046036, Aug. 2023. doi: 10.1088/1741-2552/acef92
- [26] J. M. Houck and E. D. Claus, "A comparison of automated and manual co-registration for magnetoencephalography," *PLoS One*, vol. 15, no. 4, p. 0232100, Apr. 2020. doi: 10.1371/journal.pone.0232100
- [27] S. Salari, A. Haripour, H. Rivaz, and Y. Xiao, "CAMLD: Contrast-Agnostic Medical Landmark Detection with Consistency-Based Regularization," *arXiv preprint arXiv:2411.17845*, 2024.
- [28] A. Ertl, S. Xiao, S. Denner, R. Peretzke, D. Zimmerer, P. Neher, F. Isensee, and K. Maier-Hein, "nnLandmark: A Self-Configuring Method for 3D Medical Landmark Detection," *arXiv preprint arXiv:2504.06742*, 2025.
- [29] M. V. Suhas, S. Sinha, N. Mariyappa, H. Anitha, and K. A. Kotegar, "Enhancing Precision in Medical Imaging: A 3D CNN Approach for Fiducial Point Detection in MRI Data," *IEEE Access*, vol. 12, pp. 52086–52096, Apr. 2024. doi: 10.1109/ACCESS.2024.3385573.
- [30] X. Lai, X. Wu, and L. Zhang, "Autoencoder-based multi-task learning for imputation and classification of incomplete data," *Applied Soft Computing*, vol. 98, p. 106838, 2021. doi: 10.1016/j.asoc.2020.106838.
- [31] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013. doi: 10.1109/TPAMI.2013.50.
- [32] Á. Nárai, P. Hermann, T. Auer, P. Kemenczky, J. Szalma, I. Homolya, E. Somogyi, P. Vakli, B. Weiss, and Z. Vidnyánszky, "Movement-related artefacts (MR-ART) dataset of matched motion-corrupted and clean structural MRI brain scans," *Scientific Data*, vol. 9, no. 1, p. 630, Oct. 2022. doi: 10.1038/s41597-022-01694-8
- [33] R. N. Spreng, R. Setton, U. Alter, B. N. Cassidy, B. Darboh, E. DuPre, K. Kantarovich, A. W. Lockrow, L. Mwilambwe-Tshilobo, W.-M. Luh, "Neurocognitive aging data release with behavioral, structural and multi-echo functional MRI measures," *Scientific Data*, vol. 9, no. 1, p. 119, Mar. 2022. doi: 10.1038/s41597-022-01231-7
- [34] D. Grössinger, F. P. S. Fischmeister, M. Witte, K. Koschutnig, M. Ninaus, C. Neuper, S. E. Kober, and G. Wood, "The role of superstition of cognitive control during neurofeedback training," *bioRxiv*, pp. 2021–09, Sep. 2021. doi: 10.1101/2021.09.14.460252
- [35] L. G. Nyúl and J. K. Udupa, "On standardizing the MR image intensity scale," *Magn. Reson. Med.*, vol. 42, no. 6, pp. 1072–1081, Dec. 1999. doi: 10.1002/(sici)1522-2594(199912)42:6<1072::aid-mrm11>3.0.co;2-m
- [36] J. Serra, *Image Analysis and Mathematical Morphology*. Academic Press, Inc., 1983.
- [37] N. Otsu et al., "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285–296, pp. 23–27, Jun. 1975.
- [38] P. Th'evenaz, T. Blu, and M. Unser, "Interpolation revisited [medical images application]," *IEEE Trans. Med. Imaging*, vol. 19, no. 7, pp. 739–758, Aug. 2002. doi: 10.1109/42.875199
- [39] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer Jr., R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assist. Tomogr.*, vol. 21, no. 4, pp. 554–568, Jul. 1997.



DEEPAI G. (F) received the B.Tech. degree in Telecommunications from Dr. Ambedkar Institute of Technology, Bengaluru, India, in 2019 and is currently pursuing the M.Tech. degree in Microelectronics at the Manipal Institute of Technology, Manipal, India. From 2019 to 2023, she worked as a Software Tester with Tata Consultancy Services (TCS). Since 2024, she has been interning at Intel, India, working on Layout Verification. Her research interests include Machine Learning, Signal Processing, and Low Power Electronics, and Physical Design.



ANITHA H. (F') is working as a Professor in the department of Electronics and Communication Engineering Manipal Institute of Technology, Manipal, and has over eighteen years of experience. She has received Bachelor of Engineering in Electronics and Communication Engineering from Kuvempu University in 2001 and received Master of Technology in Computer Science and Engineering from the Manipal University in 2005. She has received her Doctoral degree from the Department of Biomedical Engineering under Manipal University in 2012. Her area of interests are Biomedical signal and Medical Image Processing, and Medical Image Registration. She has been awarded six research grants from bodies such as DST and ICMR.



SWATHI B. P received the M.Tech. degree in Software Engineering from the Manipal Institute of Technology, MAHE, Manipal, India, in 2015. She is currently pursuing the Ph.D. degree in the Natural Language Processing and Semantic Web. She is also an Assistant Professor with the Department of Information & Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education,. Her research interests include Natural Language Processing, Semantic Web and Deep Learning. She has presented several papers in national and international conferences and her work has been published in various international journals.



SUHAS M. V. (M') received the B.Tech. degree in Electronics and Communication Engineering from Visvesvaraya Technological University, Belgaum, India, in 2009, and the M.Tech. degree in Digital Electronics and Advanced Communication from Manipal Academy of Higher Education, India, in 2011. He is currently pursuing the Ph.D. degree in Biomedical Signal Processing and Analysis at Manipal Academy of Higher Education. Since 2011, he has been an Assistant Professor with the Department of Electronics and Communication Engineering, Manipal Institute of Technology, Manipal, India. His research interests include medical image processing, biomedical signal analysis, and dynamic brain connectivity analysis using magnetoencephalography (MEG), with a current focus on resting-state network dynamics in temporal lobe epilepsy. Mr. Suhas M. V. is a member of The Institution of Engineers (India) (IEI).

...