Est. 1841 | YORK ST JOHN UNIVERSITY

Gornale, Shivanand, Patil, Supriya, Siddalingappa, Rashmi, Deo, Anup W. and Athani, Abdulgaphur (2026) SE-BioNet: A Squeeze-and-Excitation CNN for Non-Invasive Overthinking Detection Using Biofield Imaging. F1000 Research.

Downloaded from: https://ray.yorksj.ac.uk/id/eprint/16241/

The version presented here may differ from the published version or version of record. If you intend to cite from the work you are advised to consult the publisher's version: https://doi.org/10.12688/f1000research.192292.1

Research at York St John (RaY) is an institutional repository. It supports the principles of open access by making the research outputs of the University available in digital form. Copyright of the items stored in RaY reside with the authors and/or other copyright owners. Users may access full text items free of charge, and may download a copy for private study or non-commercial research. For further reuse terms, see licence terms governing individual outputs. Institutional Repositories Policy Statement

RaY

Research at the University of York St John

For more information please contact RaY at ray@yorksj.ac.uk

Check for updates

RESEARCH ARTICLE

SE-BioNet: A Squeeze-and-Excitation CNN for Non-Invasive Overthinking Detection Using Biofield Imaging

[version 1; peer review: 1 approved with reservations]

Shivanand Gornale[1], Supriya Patil[1], Rashmi Siddalingappa[2], Anup W Deo[3], Abdulgaphur Athani[4]

[1]a Department of Computer Science, School of Mathematics and Computing Sciences, Rani Channamma University, Belagavi, Karnataka, India
[2]Department of Computer and Data Science, York St John University, York, England, UK
[3]Institute of Biofield Evaluation and Research Centre, Thane, Maharashtra, India
[4]Department of Engineering Technology, Universiti Teknikal Malaysia Melaka Fakulti Kejuruteraan Pembuatan, Durian Tunggal, Malacca, Malaysia

v1 **First published:** 02 Nov 2026, **15**:1819
https://doi.org/10.12688/f1000research.192292.1
Latest published: 02 Nov 2026, **15**:1819
https://doi.org/10.12688/f1000research.192292.1

Abstract

Background

The human biofield has been explored as a possible indicator of physical and emotional states. Complementary and alternative medicine (CAM) practices, such as Reiki, Therapeutic Touch, and Pranic Healing, incorporate biofield-related information into their assessment and therapeutic approaches. Gas Discharge Visualization (GDV) and Poly Contrast Interference Photography (PIP) are among the imaging techniques used to visualize biofield patterns. However, most existing studies have focused on GDV images, with comparatively limited research using PIP images and deep learning techniques. Overthinking is a growing mental health concern that is often difficult to identify using conventional diagnostic methods. This study explores whether deep learning can identify patterns in PIP biofield images associated with overthinking.

Methods

A convolutional neural network (CNN) integrated with a Squeeze-and-Excitation (SE) attention mechanism was developed for binary classification of biofield images into two classes: overthinking and no overthinking pattern. The SE module enhances discriminative feature learning by adaptively recalibrating channel-wise feature responses. The model was trained and evaluated separately on datasets

Open Peer Review

Approval Status ?

	1
version 1 02 Nov 2026	? view

1. **Gunjan Chhabra**, Swami Rama Himalayan University, Dehradun, India

Any reports and responses or comments on the article can be found at the end of the article.

annotated independently by two experts (Expert-I and Expert-II), and its performance was assessed using 5-fold cross-validation.

Results

The proposed SE-BioNet model achieved classification accuracies of 93.95% and 93.80% on Expert-I and Expert-II annotated datasets, respectively, with corresponding F1-scores of 94.02% and 93.59%. Five-fold cross-validation yielded mean accuracies of 92.22% ± 1.17% and 92.08% ± 1.34%, demonstrating stable generalisation.

Conclusions

The proposed framework demonstrates technical feasibility and consistent performance across cross-validation folds. However, validation using larger, multi-centre datasets is necessary to assess its generalizability. These findings highlight the potential of attention-based model in learning more useful features though further testing with larger datasets and expert validation is necessary for clinical application.

Keywords
Biofield imaging, Overthinking, Polycontrast Interference Photography (PIP), Repetitive negative thinking, Squeeze-and-Excitation (SE) attention

Corresponding author: Rashmi Siddalingappa (r.siddalingappa@yorksj.ac.uk)

Author roles: Gornale S: Conceptualization, Formal Analysis, Methodology, Supervision, Validation, Writing – Review & Editing; **Patil S**: Data Curation, Formal Analysis, Methodology, Resources, Visualization, Writing – Original Draft Preparation; **Siddalingappa R**: Conceptualization, Methodology, Resources, Supervision, Writing – Review & Editing; **W Deo A**: Data Curation, Formal Analysis; **Athani A**: Methodology, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2026 Gornale S *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Gornale S, Patil S, Siddalingappa R *et al.* **SE-BioNet: A Squeeze-and-Excitation CNN for Non-Invasive Overthinking Detection Using Biofield Imaging [version 1; peer review: 1 approved with reservations]** F1000Research 2026, **15**:1819 https://doi.org/10.12688/f1000research.192292.1

First published: 02 Nov 2026, **15**:1819 https://doi.org/10.12688/f1000research.192292.1

1. Introduction

Overthinking, commonly used to describe a pattern of sustained, repetitive, and self-focused thinking that is difficult to control or disengage from, is clinically studied under the constructs of rumination, defined as repetitive, passive focus on distressing thoughts and their causes[1] and repetitive negative thinking (RNT), a broader transdiagnostic term capturing persistent intrusive thought patterns across anxiety, depression, and related conditions.[2] Throughout this paper, the term overthinking refers to this shared cognitive process, using accessible language that reflects its occurrence across both clinical and non-clinical populations. Repetitive and self-focused thinking is recognised as a key transdiagnostic cognitive process associated with anxiety, depression, and stress-related conditions,[3] and according to the World Health Organisation,[4] approximately one in eight people worldwide are affected by such conditions. Findings from the National Mental Health Survey of India (2015–16) further highlight this burden, reporting prevalence rates of approximately 3.5% for anxiety and 2.7% for depression, with a treatment gap exceeding 85% for common mental disorders.[5] Despite its clinical significance and prevalence, overthinking is difficult to identify early because it manifests as an internal cognitive process. At the more resource-intensive end, Electroencephalography (EEG) measures brain wave activity, capturing alpha and beta wave fluctuations associated with excessive cognitive arousal,[6] and functional magnetic resonance imaging (fMRI) visualises brain network activity, particularly hyperactivation of the Default Mode Network comprising the medial prefrontal cortex (mPFC), posterior cingulate cortex (PCC), and related regions during maladaptive rumination and impaired emotional regulation.[7–9] The mPFC, in particular, plays a central role in self-evaluation, emotional regulation, and decision-making,[10] and its hyperactivation at rest is closely associated with excessive self-referential thought. Physiological measures such as heart rate variability (HRV) further reflect autonomic nervous system imbalance, with reduced parasympathetic activity observed during prolonged, repetitive cognitive engagement.[11] Moving toward non-invasive alternatives, Gas Discharge Visualisation (GDV) records energy patterns produced in electric fields, thought to reflect autonomic nervous system activity,[12] with studies showing sensitivity to stress, emotional imbalance, and mental fatigue.[13] At the least invasive end of the spectrum, Polycontrast Interference Photography (PIP) captures biofield energy patterns in real time without physical contact or electrical stimulation. While EEG and fMRI provide neurophysiology-grounded measurements, they are expensive, require specialised infrastructure and trained personnel, and are unsuitable for routine or large-scale screening. HRV monitoring, though more accessible, captures systemic physiological responses rather than cognitive states directly. GDV, while less resource-intensive than neuroimaging, requires direct physical contact, involves electrical stimulation, and is sensitive to variability introduced by skin moisture, pressure, and operator handling, limiting its reproducibility. PIP, by contrast, is a non-contact, non-stimulating modality that is safe for repeated use and potentially suitable for continuous or large-scale psychophysiological assessment, motivating its selection as the imaging technique in this study. The biofield patterns visualised through PIP are hypothesised to reflect biophoton emissions, ultra-weak light signals naturally generated by living systems whose intensity and distribution have been reported to vary with physical, emotional, and mental states.[14,15] In particular, the cranial biofield region captured via PIP is thought to reflect energy changes associated with sustained internal thought and brain network activity.[16] making it a candidate modality for non-invasive assessment of cognitive state. However, the precise biophysical mechanisms linking PIP colour patterns to specific cognitive states, such as overthinking, remain hypothetical and have not been independently established. This study presents an exploratory, hypothesis-generating investigation into whether deep learning-based analysis of cranial PIP biofield images can distinguish patterns putatively associated with overthinking from those that are not. We propose a Convolutional Neural Network integrated with a Squeeze-and-Excitation (SE) attention mechanism to enhance feature learning by emphasising energy-relevant regions within the cranial biofield. The framework is evaluated against SVM, baseline CNN, self-attention, and multi-head attention variants as a proof-of-concept contribution to the emerging intersection of biofield imaging and computational mental health research, with the explicit aim of motivating further investigation through multi-expert labelling protocols and independent neurophysiological validation.

The key contributions of the current paper are as follows: 1. Creation of a labelled dataset comprising 857 PIP-based biofield images of the cranial region, annotated by two independent domain experts with near-perfect inter-rater agreement (Cohen's $\kappa = 0.9532$). 2. A CNN model integrated with the Squeeze-and-Excitation (SE) attention mechanism is proposed to enhance discriminative feature learning, achieving 93.95% accuracy in detecting overthinking-related patterns. 3. Providing an automated, non-invasive approach for identifying potential overthinking patterns, reducing reliance on subjective assessments while demonstrating robust generalisation through five-fold cross-validation (mean accuracy: $92.22\% \pm 1.17\%$ and $92.08\% \pm 1.34\%$). To the best of our knowledge, this is the first study to: i) Apply deep learning with attention mechanisms to cranial PIP biofield images for cognitive state assessment, ii) Create a systematically annotated dataset of biofield images with expert validation, and iii) Demonstrate the feasibility of using non-contact biofield imaging for detecting patterns potentially associated with overthinking.

2. Related work

Integrating biofield imaging techniques with artificial intelligence has opened new avenues for objective mental health assessment. As shown in Table 1, prior investigations have examined Gas Discharge Visualisation (GDV), EEG-based biomarkers, and deep learning architectures incorporating attention mechanisms for disease detection and emotional state classification. This section reviews these approaches to highlight methodological advances and limitations. Misaki et al.[17] examined age-related differences in neural representations of rumination and worry. They analysed neuroimaging datasets using pattern classification methods to assess discriminability across age groups. Li et al.[18] investigated repetitive negative thinking in children within depression and anxiety contexts. Survey-based psychological datasets

Table 1. Summary of recent machine learning and deep learning approaches for rumination and overthinking detection using EEG, fMRI, and behavioural datasets.

Author	ML/DL method	Dataset	Key findings
Misaki et al. (2026)[17]	Whole-brain activation pattern decoding using individualized AutoML models (Auto-sklearn).	Primary dataset-[fMRI]	Neural differentiation between rumination and worry increases with age, reflecting improved cognitive control
Li et al. (2025)[18]	Random forest	Primary dataset	Core features of Repetitive Negative Thinking (RNT) are pivotal for symptom propagation between anxiety and depression.
Ahire, N. (2025)[19]	1D-CNN, SVM, Logistic Regression;	Primary dataset- [EEG]	Task-condition EEG outperforms resting EEG for MDD (incl. ruminative) classification; 1D-CNN captures temporal EEG patterns most effectively; frontal theta key discriminator
(Kapitany M et al. 2024)[20]	EEGNET	Primary dataset- [EEG]	EEG-based depression detection under inner and overt speech conditions using emotion-related stimuli
Welkerling J et al, 2025[21]	Linear SVM decoder; features: alpha/beta/theta power + inter-electrode connectivity;	Primary dataset- [EEG]	Proof-of-concept: EEG decodes induced rumination above chance;
Wei L et al., 2025[22]	NBS-Predict (ML-based network edge selection + cross-validated ridge regression); tested FPN-DMN and FPN-SN edge sets separately	Primary dataset-[fMRI]	Reflective rumination (adaptive) has unique neural fingerprint separable from brooding via FPN coupling; FPN-DMN and FPN-SN edges provide best prediction; healthy rumination is detectable from rs-fMRI
(Kaushik P et al. 2023)[23]	1D-CNN, LSTM, SVM, RF	Primary dataset- [t rs-EEG + SART task-based EEG]	rs-EEG outperforms task-EEG for vulnerability prediction; task-based EEG is better for understanding rumination mechanisms ('stickiness'); left frontal increased amplitude marks high vulnerability
(Aydin et al., 2022)[24]	Cost Regulation Extreme Learning Machine	LEMON	Frequency-specific complexity features can effectively distinguish between maladaptive rumination and cognitive distraction
Bailey & Kang (2022)[25]	Portable EEG Monitoring	Primary dataset	Frontal beta activity serves as an objective biomarker for rumination; a reduction in this activity indicates lower cognitive hyperactivity and overthinking
Silveira E 2020[26]	Linear support vector machine	Clinical and blood biomarkers variables	Identified rumination patterns across a transdiagnostic sample

were analysed using statistical and computational models to understand transdiagnostic effects. Ahire et al.[19] proposed a hybrid framework combining machine learning and deep learning for depression diagnosis using EEG-derived brain-wave patterns. Clinical EEG datasets were utilised to improve classification accuracy by integrating models. Kapitany-Foveny et al.[20] investigated EEG-based depression detection under inner and overt speech conditions using emotion-related stimuli. EEG data were analysed with machine learning techniques to identify which condition yields more discriminative biomarkers. The study by Welkerling et al.[21] focused on decoding ruminative states from neurophysiological signals. They analysed experimental neurophysiological datasets using machine learning to identify patterns associated with rumination. Wei et al.[22] explored rumination using triple-network brain connectivity analysis of fMRI data. The study highlights the role of large-scale brain network interactions in sustaining ruminative thought patterns. Kaushik et al.[23] explored using EEG to predict vulnerability to depression in a non-clinical population by comparing resting-state and task-based recordings. Results demonstrated that task-based EEG provided more discriminative features for identifying depression-prone individuals. Aydin et al.[24] have analysed frequency-specific complexity features of EEG signals to classify maladaptive rumination and cognitive distraction, demonstrating the effectiveness of traditional machine learning classifiers. Bailey & Kang[25] analysed the effects of outdoor walking and sitting on cognitive performance and mental well-being using behavioural measures. Their findings suggest that improved mood and cognitive engagement during walking may reduce repetitive negative thinking or overthinking. Silveira et al.[26] applied machine learning to decode rumination across patients with anxiety, mood, and psychotic disorders. They used clinical data from a transdiagnostic outpatient sample to build predictive models of rumination. Biofield-based approaches have been increasingly explored as non-invasive tools for understanding physical and mental health conditions. Bryndin et al.[27] and Gao et al.[28] have studied the use of GDV technology for assessing depression. They found clear differences in biofield energy patterns between depressed and healthy individuals. They have also diagnosed health conditions using biofield visualisation and highlighted various biofield therapies to balance energy for physical and mental well-being. Mangione et al.[29] reviewed biofield therapies such as Reiki and Healing Touch and their effects on mental health. The study reported reduced stress and anxiety, along with improved mood, suggesting that these therapies may help reduce overthinking and support mental well-being. Neurophysiological evidence further supports the link between mental health and biofield-based interventions. A study by Olbrich et al.[30] examining EEG-based brain signal biomarkers for Major Depressive Disorder (MDD) identified alterations in frontal alpha activity, event-related potentials, and functional connectivity as key indicators of depression. Uchida et al.[31] examined the effects of Okada Purifying Therapy, a form of biofield therapy, on brain activity in healthy adults using EEG. Their findings showed a significant increase in frontal and central alpha-band power, which is commonly associated with relaxation and reduced cognitive load. Increased alpha activity suggests reduced excessive self-focused thinking, a core feature of overthinking and default mode network (DMN) overactivity. Clinical evidence also supports the effectiveness of biofield energy interventions. Trivedi et al.[32] Conducted a clinical trial involving 77 adults with mental health conditions to evaluate the impact of biofield energy treatment. The results demonstrated significant reductions in stress, anxiety, and psychological symptoms, with the therapy being safe, well-tolerated, and suitable for long-term use.

With advances in image processing and machine learning, several studies have explored automated analysis of biofield images. Chhabra et al.[33] developed a CNN-based emotion recognition framework combined with aura colour-space visualisation to detect emotional states. Based on the identified emotions, they applied sound therapy to address negative emotional patterns. Analysis of biofield images before and after therapy showed a 57% increase in positive emotional states. In another study by Gornale et al.,[34] biofield images of the knee region were analyzed to detect energy imbalances and classify them into three categories. The model achieved an accuracy of 87.14% using Support Vector Machines (SVM) and 94.28% using the K-Nearest Neighbors (KNN) classifier. Veerabrahmachar et al.[35] Conducted a pilot study to examine the effects of Nada Yoga meditation on energy levels and the alignment of seven chakras. Biofield images captured using electro-photonic imaging before and after meditation showed notable improvements in energy levels and chakra alignment. While prior studies have primarily focused on EEG and neuroimaging signals for detecting depression and rumination, the application of machine learning to biofield imagery remains limited. Existing biofield-based computational approaches have explored emotion recognition from facial signals, physical health assessment using Gas Discharge Visualisation (GDV) imagery, and musculoskeletal condition detection from PIP images. However, these studies do not specifically address cognitive mental states or focus on the cranial biofield. To the best of our knowledge, the use of deep learning to analyse cranial PIP biofield images in the context of sustained, repetitive cognitive engagement has not been systematically investigated. Recent studies have shown that attention mechanisms improve deep learning performance by enabling models to focus on the most informative features. SE-Net, introduced by Hu et al. (2020),[36] is a channel-wise attention mechanism that adaptively recalibrates feature responses with minimal computational overhead. In medical image analysis, Roy et al.[37] demonstrated consistent improvements in whole-brain MRI and organ CT segmentation using SE variants. For brain tumour analysis, Zhou et al.[9] proposed scSE-NL V-Net for MRI segmentation, while Huang et al.[38] reported improved MRI-based tumour classification using SE-enhanced ResNet50V2. Similarly, Gencer et al.[39] developed a hybrid SE-based model for retinal disease classification from OCT images. These studies

highlight the effectiveness of SE attention in enhancing feature representation across diverse medical imaging tasks. Motivated by these findings, this work proposes a deep learning framework for classifying cranial PIP biofield images as overthinking or no-overthinking using SE attention. The proposed approach aims to improve feature learning from biofield images and support the identification of overthinking-related patterns, complementing existing neurophysiological assessment techniques.

3. Materials and methods

In this study, biofield images are collected and processed to develop an automated classification. This section explains the overall experimental setup of the proposed system. As shown in Fig. 1, the complete pipeline consists of image acquisition, image preprocessing, feature extraction using a CNN backbone, energy-guided channel attention using SE blocks, and binary classification.

3.1 Image acquisition

The images used in this study were acquired using the Poly contrast Interference Photography (PIP) technique at the Institute of Biofield Evaluation and Research Centre (IBERC), Thane, Maharashtra, India. The institute holds an Indian patent (No. 528035) covering the methods of image acquisition, image analysis, and clinical interpretation based on colour scheme variations. In this technique, a customised digital camera is employed within a specially designed imaging setup (Fig. 2(a)) that minimises electromagnetic interference and ensures controlled lighting conditions. Written informed consent for participation in the study was obtained from all participants prior to image acquisition. Before biofield images were captured, each participant was informed about the purpose, procedure, and intended use of the data, and signed a written consent form. Fig. 2(b) shows a sample image as a colour spectrum. Health-related interpretations of the biofield are commonly derived from the colour patterns observed on and around the body[33–34]). The PIP procedure involves capturing approximately 35 images of each subject from multiple angles to record the biofield patterns surrounding the human body. From this dataset, we selected and stored the frontal profile images of 857 subjects separately for subsequent processing and analysis.

3.2 Image preprocessing

All biofield images are resized to a fixed spatial resolution of 224 × 224 pixels to maintain uniformity across the dataset. Since region-of-interest (ROI) images are already extracted, no additional segmentation is applied. This method aims to identify images such as either Overthinking or No Overthinking based on energy-related patterns observed in the region above the face. Face detection is performed using the Haar Cascade Classifier,[40] a machine learning–based method for efficient identification of facial features. Based on the detected face region and relevant anatomical landmarks, the region of interest (ROI) is extracted and stored separately. Fig. 3 shows the stepwise extraction of the ROI. The ROI images were annotated under the guidance of a domain expert affiliated with the IBERC, who has over 15 years of experience in biofield image interpretation and holds a patent related to this methodology. Another expert is a radiologist affiliated with the Panvel Polyclinic in Panvel, Mumbai. The experts followed a consistent visual assessment protocol based on the distribution, intensity, and spatial extent of red and dark pink colour patterns in the cranial region. In this framework,

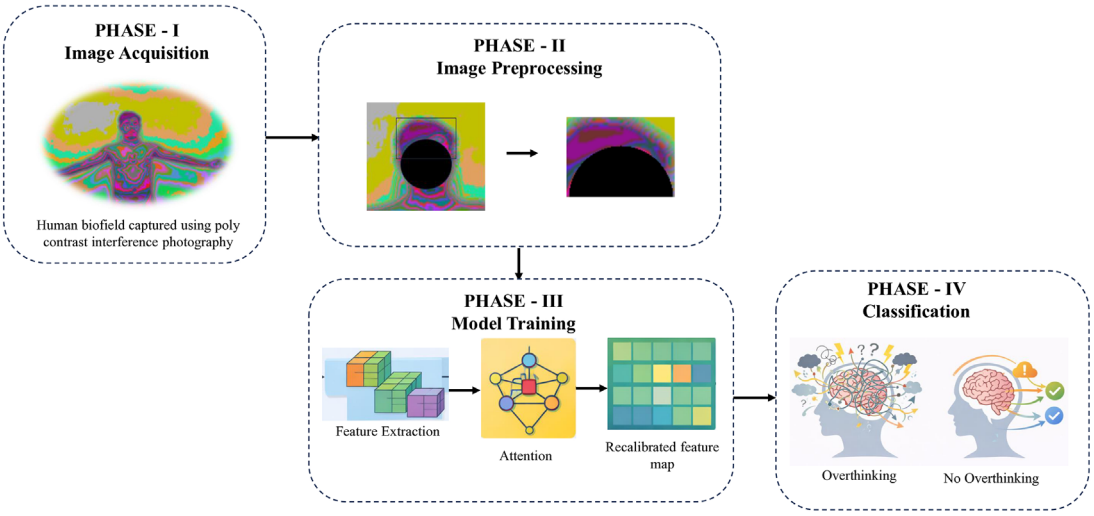


Fig. 1. Block diagram of the proposed methodology illustrating the sequential workflow, including image acquisition and preprocessing, image classification into overthinking and no overthinking.

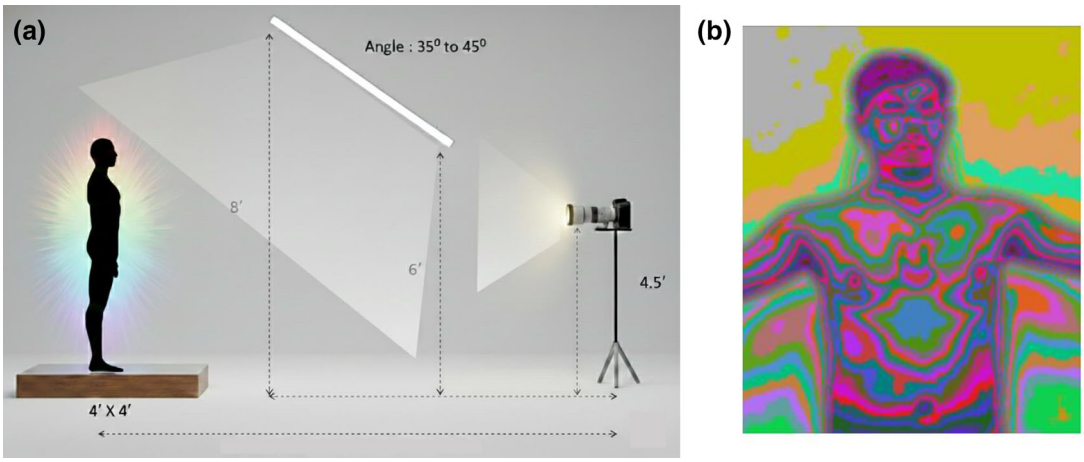


Fig. 2. (a): Environment Setup for biofield scan, which includes alignment of camera, light source and person to be scanned. (b) shows a sample Biofield Image captured using PIP.

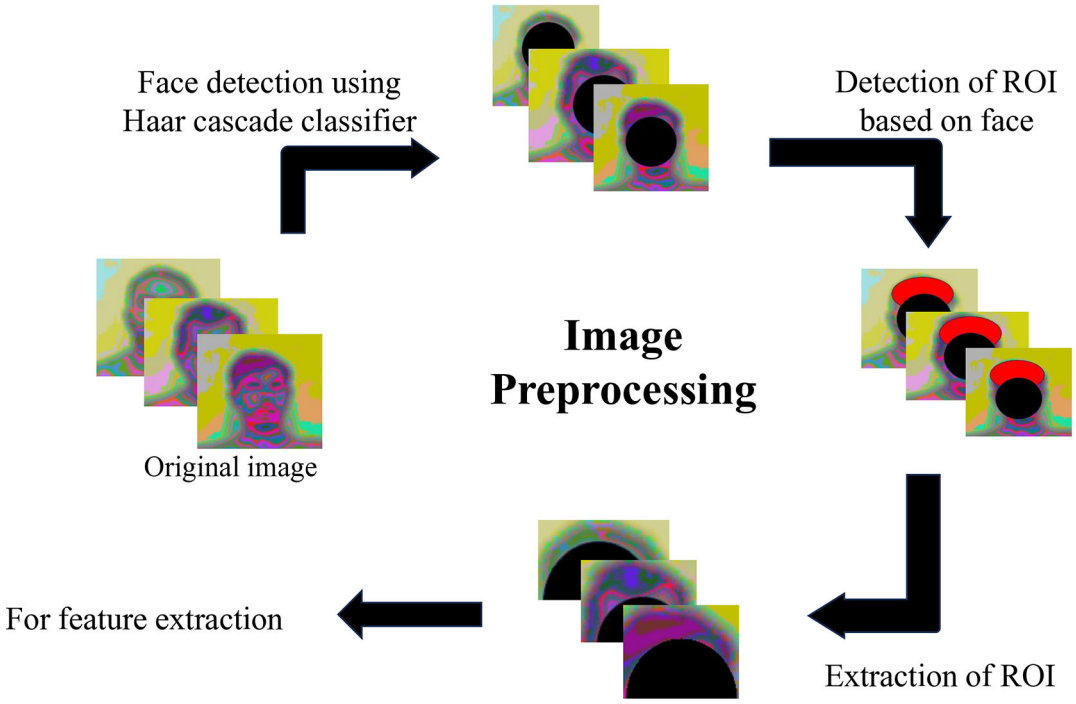


Fig. 3. Stepwise ROI extraction process. The face region is first detected using the Haar cascade classifier. Based on the detected face location, the region above the face is selected as the ROI for further analysis.

relatively wider and more concentrated patches of these colour patterns are interpreted as indicative of localised variations in biofield activity, which are hypothesised to be associated with sustained cognitive engagement referred to in this study as overthinking. Fig. 4 (a) and (b) show some sample images of both classes.

3.3 Image augmentation

Image augmentation is a technique that artificially expands the training dataset by applying controlled transformations to existing images. It greatly enhances the performance of deep learning models, especially in medical imaging, where data is frequently scarce.[41–42] To address the risk of data leakage, augmentation was applied exclusively to the training set after the dataset was partitioned into train, validation, and test splits. Augmentation techniques such as rotation, shearing, vertical flipping, and horizontal flipping were used in this work. Each image was independently flipped along the horizontal and vertical axes with probability $p = 0.5$. To account for slight variations in orientation, a random rotation was

applied within ±15°. A random affine transformation with shear in the range of ±10° with rotation fixed at 0° was used to simulate minor geometric distortions. Fig. 5 shows the result of these transformations on a sample image. After augmentation, the train dataset grew from 599 images to 2995 images. By increasing variability in the input data without additional data collection, these augmentation strategies simulate real-world image variations and help reduce the risk of overfitting.

3.4 Model architecture

In this work, an attention-enhanced deep learning framework, as shown in Fig. 6, is proposed to classify biofield images into Overthinking and No-overthinking categories. The model SE-BioNet is based on a convolutional neural network (CNN) integrated with a Squeeze-and-Excitation (SE) attention mechanism. This enables recalibration of channel-wise feature responses,[43,44] emphasising energetically significant regions in biofield images discriminative for overthinking vs no-overthinking classification.

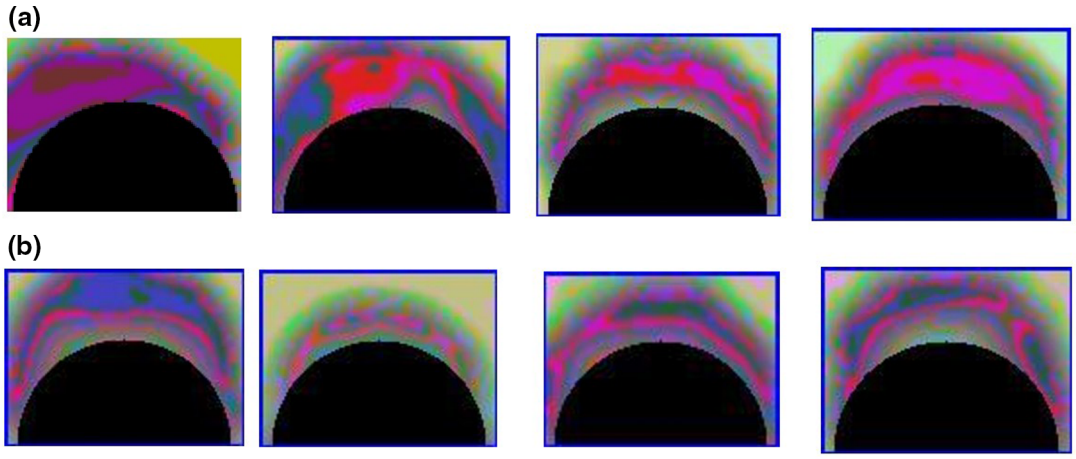


Fig. 4. (a). Represents sample images containing overthinking patterns. Fig. 4(b). Represents images with the absence of overthinking pattern.

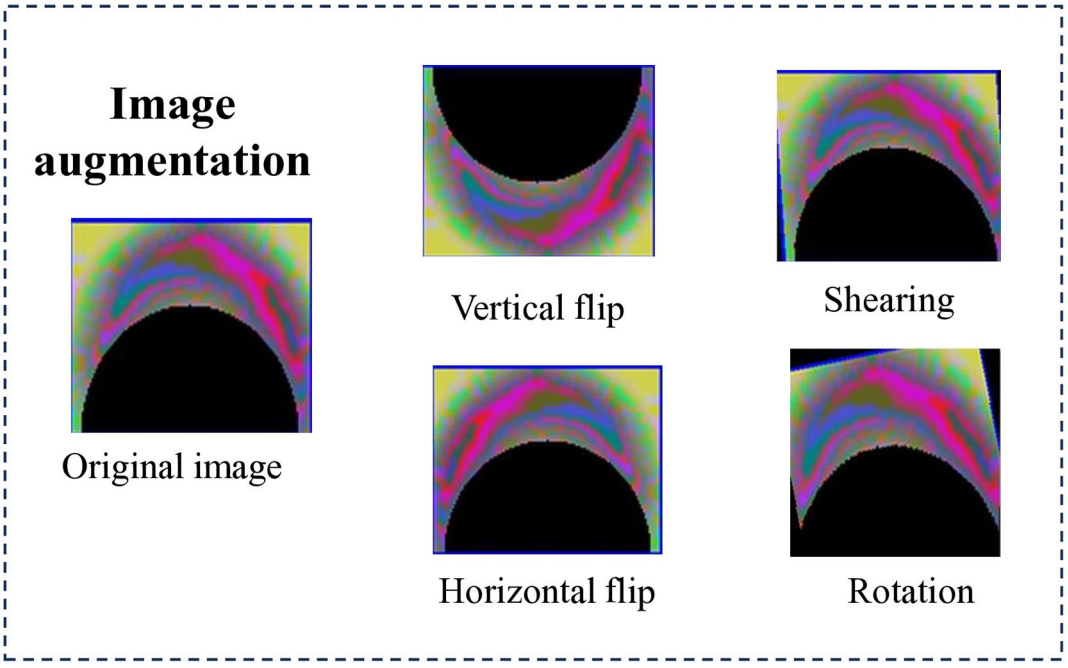


Fig. 5. Illustration of image augmentation techniques applied to biofield images, including vertical flip, horizontal flip, shearing, and rotation.

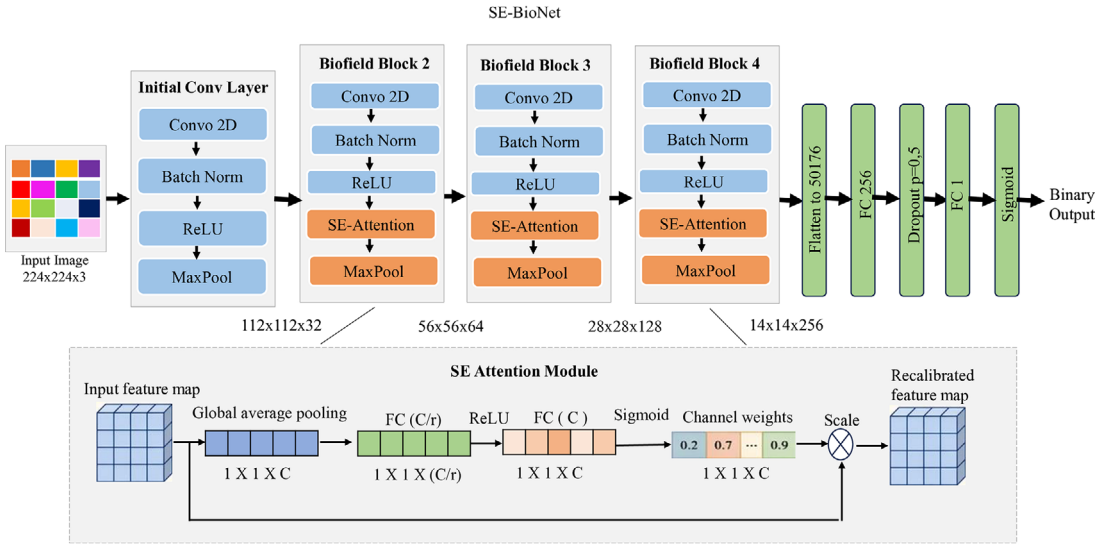


Fig. 6. Architecture diagram of the SE-CNN Block that includes convolution, batch normalization, ReLU activation, global average pooling, and channel-wise feature recalibration.

CNN with energy attention using Squeeze-and-Excitation (SE) Attention Mechanism

To improve classification accuracy, SE attention is integrated after selected convolutional blocks, thereby enhancing the discriminative power of the extracted features. The SE mechanism explicitly models interdependencies between feature channels by performing feature recalibration through three main steps: Squeeze, Excitation and Recalibration. The squeeze operation in equation 1 aims to condense spatial information from each feature map into a compact channel-wise descriptor that captures the input's global context. Given a feature map F, produced by a convolutional layer, global average pooling (GAP) is applied independently to each channel:

$$z_c = \frac{1}{H X W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_c(i.j), c = 1,2,\ldots,C \tag{1}$$

Where $F_c(i,j)$ denotes the feature value at the spatial location (i,j) in the c^{th} channel. H and W denote the height and width of the feature map, respectively, while i and j are the spatial indices. The output Z_C is a scalar representing the global descriptor of the c^{th} channel, and c = 1,2, …,C indicates the total number of channels. This operation converts each two-dimensional feature map into a single scalar value, resulting in a vector $z \in R^{H \times W \times C}$ that represents the global energy distribution of the feature channels.[45] The squeeze step captures the average energy intensity present in each channel across the entire biofield image, enabling the network to understand which channels encode stronger or weaker energetic responses. The excitation operation learns nonlinear relationships among feature channels and determines their relative importance. The squeezed vector z is passed through a lightweight gating mechanism composed of two fully connected (FC) layers, producing a vector of activations as given in equation 2, which is then used to rescale the feature maps.[45] This step allows the network to learn how different energy-related channels interact and which channels are more relevant for distinguishing overthinking from no-overthinking states.

$$s = \sigma(W_2.\delta(W_1.z)) \tag{2}$$

Here, z is the input vector from the squeeze operation. W_1 and W_2 are weight matrices of fully connected layers, where W_1 reduces the dimensionality and W_2 restores it. The symbol δ denotes a non-linear activation function, and σ is the sigmoid activation function that produces channel-wise attention weights. In the recalibration step, the original feature maps are adaptively reweighted using the learned attention coefficients. Each channel is scaled by its corresponding excitation weight, as shown in equation 3. This operation preserves the spatial structure of the feature maps while emphasising channels that are more discriminative for classification.[45] Channels corresponding to significant energy patterns are amplified, while less informative or noisy channels are suppressed. This selective emphasis improves the network's ability to capture subtle biofield variations associated with overthinking.

$$\widehat{F_c} = s_c . F_c \quad (3)$$

Where, F_c denotes the original feature map of the c^{th} channel, and s_c is the corresponding attention weight. The output $\widehat{F_c}$ is the recalibrated feature map, scaled by its importance.

4. Experiments and results

This section presents the experimental validation of the proposed SE-BioNet for detecting overthinking patterns from biofield images. A comprehensive evaluation was conducted to assess model performance, including implementation details, ablation analysis of attention mechanisms, class-wise classification results, and overall performance metrics.

4.1 Dataset description

Polycontrast Interference Photography (PIP) technology is relatively niche within mainstream biomedical research and is primarily used by institutions specialising in integrative and energy-based diagnostic approaches. Consequently, standardised, annotated datasets for machine learning–based analysis are limited, and no publicly available benchmark datasets currently exist for PIP-based imaging. The dataset used in this study was obtained from the Institute of Biofield Evaluation and Research Center (IBERC), Thane, India. Biofield images were captured using the PIP technique and stored as 1152 × 768 pixel PNG images. This work used 857 volunteer scans, including 398 female and 459 male participants aged 20 to 75 years. We obtained written informed consent from all participants before data collection. Participants ranged in age from 20 to 75 years (mean age: 45.3 ± 12.7 years), with 398 females (46.4%) and 459 males (53.6%). Exclusion criteria were minimal, limited to acute medical emergencies and the inability to provide informed consent. Since biofield imaging is non-invasive and suitable for all age groups, minimal exclusion criteria were applied. Health conditions and medication history were not used for pre-screening; instead, assessments were performed after imaging through expert interpretation and subject interviews. The generated reports also included recommendations for further clinical investigations when required. Currently, experienced domain experts interpret images manually. These expert assessments serve as the ground truth for this study, while ongoing research aims to progressively automate the interpretation process using machine learning techniques. Based on observed biofield characteristics, the images are annotated as overthinking or no-overthinking, as shown in Table 2 by the two experts in biofield evaluation: Expert I from IBERC, Thane, and Expert II, a radiologist from Panvel. To facilitate early identification of overthinking patterns, a region of interest (ROI) is extracted from the original images and stored separately for further analysis.

4.2 Ethics approval statement

The dataset used in this study was created using human biofield images collected from IBERC, Thane, Maharashtra, India with permission (Permission Letter Ref: RCU/BELAGAVI/PHY/2023–24, dated 27 October 2023). This study involved retrospective analysis of these previously acquired biofield images of human participants. All personally identifiable information was removed before preprocessing and model development, ensuring participant confidentiality and privacy.

Ethics approval:

The research proposal, titled “Analysis of Human Biofield Scan Images for Physical, Mental, Emotional Health Issues Using Image Processing Techniques,” was approved by the Institutional Ethics Committee (IEC) of Rani Channamma University, Belagavi, Karnataka, India (Reference No. RCU/BELAGAVI/PHY/2023–24, dated 27 October 2023). Approval covered the use of biofield images for academic and research purposes, on the condition that participant confidentiality was maintained and the approved protocol was followed.

4.3 Implementation details

All experiments were performed in Google Colab using PyTorch with CUDA acceleration on an NVIDIA Tesla T4 GPU. The dataset was divided into training (70%), validation (15%), and testing (15%) subsets. Input images were

Table 2. The annotations details given by the two experts.

	Expert - I	Expert - II
Overthinking	413	417
No overthinking	444	440

resized to 224 × 224. The model was trained for 50 epochs using a batch size of 16. The proposed architecture depicted in Table 3 consists of convolutional layers with Batch Normalisation and MaxPooling, integrated with Squeeze-Excitation Attention modules using a channel reduction ratio of 8. A dropout layer with a rate of 0.5 was applied before the final fully connected layer to reduce overfitting. The final classification layer uses a sigmoid activation function. Binary Cross-Entropy (BCE) loss was used for optimisation, using the Adam optimiser with a learning rate of 1×10^{-4}. A decision threshold of 0.5 was applied to convert predicted probabilities into binary class labels.

The proposed SE-BioNet contains approximately 2.8 million trainable parameters, which is comparable to the base CNN (2.1 million) but significantly fewer than the Multi-Head Attention variant (4.3 million). The SE attention blocks add only 0.7 million parameters, representing a minimal computational overhead (3.2% increase) while providing substantial performance improvements.

4.4 Ablation study

An ablation study was conducted to analyse the contribution of different attention mechanisms to classification performance. The results show that the proposed SE-BioNet model outperformed SVM, CNN Attention, Self-Attention, and Multi-Head Attention on both the Expert-I and Expert-II annotated datasets. Before augmentation, SE-BioNet achieved the highest accuracies of 81.58% and 85.92% for Expert-I and Expert-II, respectively. After applying data augmentation, the performance of all models improved; however, SE-BioNet continued to achieve the best results with accuracies of 93.95% (Expert-I) and 93.80% (Expert-II). It also obtained the highest F1-scores of 94.02% and 93.59% on the respective datasets. These results indicate that the SE attention mechanism helps the model learn more useful features, leading to better classification performance. The improvement after augmentation further shows that data augmentation enhances the model's robustness and generalisation ability.

4.5 Classification results

The effectiveness of different machine learning and deep learning models for detecting overthinking from biofield images was evaluated on datasets annotated by Expert-I and Expert-II. The models included SVM, CNN, Self-Attention, Multi-

Table 3. Detailed layer-wise architecture of SE-BioNet model.

Layer	Type	Input shape	Output shape	Kernel Size
conv1	Conv2d	(B, 3, 224, 224)	(B, 32, 224, 224)	3x3
bn1	BatchNorm2d	(B, 32, 224, 224)	(B, 32, 224, 224)	-
pool1	MaxPool2d	(B, 32, 224, 224)	(B, 32, 112, 112)	2x2
block2.conv	Conv2d	(B, 32, 112, 112)	(B, 64, 112, 112)	3x3
block2.bn	BatchNorm2d	(B, 64, 112, 112)	(B, 64, 112, 112)	-
block2.att	SEAttention	(B, 64, 112, 112)	(B, 64, 112, 112)	-
block2.pool	MaxPool2d	(B, 64, 112, 112)	(B, 64, 56, 56)	2x2
block3.conv	Conv2d	(B, 64, 56, 56)	(B, 128, 56, 56)	3x3
block3.bn	BatchNorm2d	(B, 128, 56, 56)	(B, 128, 56, 56)	-
block3.att	SEAttention	(B, 128, 56, 56)	(B, 128, 56, 56)	-
block3.pool	MaxPool2d	(B, 128, 56, 56)	(B, 128, 28, 28)	2x2
block4.conv	Conv2d	(B, 128, 28, 28)	(B, 256, 28, 28)	3x3
block4.bn	BatchNorm2d	(B, 256, 28, 28)	(B, 256, 28, 28)	-
block4.att	SEAttention	(B, 256, 28, 28)	(B, 256, 28, 28)	-
block4.pool	MaxPool2d	(B, 256, 28, 28)	(B, 256, 14, 14)	2x2
flatten	Flatten	(B, 256, 14, 14)	(B, 50176)	-
fc1	Linear	(B, 50176)	(B, 256)	-
drop	Dropout	(B, 256)	(B, 256)	-
fc2	Linear	(B, 256)	(B, 1)	-
sigmoid	Sigmoid	(B, 1)	(B, 1)	-

Head Attention, and the proposed SE-BioNet. Tables 4 and 5 present the results before and after data augmentation, respectively. As given in Table 5, before augmentation, the performance of all models was comparatively lower and less balanced across the two classes. For the Expert-I annotated dataset, the proposed SE-BioNet achieved the highest overall accuracy of 81.56%, with class-wise accuracies of 87.50% for overthinking and 75.63% for no-overthinking samples. The Multi-Head Attention model achieved 82.70% overall accuracy but showed a large difference between the two classes. For the Expert-II annotated dataset, SE-BioNet again produced the best performance with an overall accuracy of 85.91%, correctly identifying 96.83% of overthinking samples and 75.00% of no-overthinking samples. These results indicate that although meaningful patterns were learned, the limited training data restricted the models' generalisation capability. After applying augmentation to the training dataset, as shown in Table 6, a substantial improvement was observed across all deep learning models. For the Expert-I annotated dataset, the proposed SE-BioNet achieved the highest overall accuracy of 93.95%, with balanced class-wise accuracies of 95.24% for overthinking and 92.65% for

Table 4. Performance analysis of attention mechanisms for biofield-based overthinking classification on Expert-I and Expert-II annotated data.

	Expert - I				Expert - II			
Model	**Accuracy**	**Precision**	**Recall**	**F1Score**	**Accuracy**	**Precision**	**Recall**	**F1Score**
Before augmentation								
SVM	77.44%	78.46%	76.12%	77.27%	76.15%	75.38%	76.56%	75.97%
CNN attention	71.43%	71.43%	86.21%	78.13%	73.46%	75.0%	77.77%	76.36%
Self-attention	75.51%	75.0%	85.71%	80.0%	75.51%	74.28%	89.65%	81.25%
Multi-head attention	79.59%	95.65%	70.96%	81.48%	73.46%	80.0%	77.41%	78.68%
SE-BioNet	81.58%	78.31%	87.50%	82.65%	85.92%	79.48%	96.83%	87.30%
After augmentation								
SVM	88.89%	89.36%	84.0%	86.60%	85.38%	83.58%	87.50%	85.50%
CNN attention	87.33%	82.22%	95.24%	88.25%	90.27%	86.62%	95.24%	90.73%
Self-attention	90.04%	88.38%	92.19%	90.25%	89.30%	86.09%	93.75%	89.76%
Multi-head attention	91.51%	93.80%	88.89%	91.28%	92.24%	95.27%	88.89%	91.97%
SE-BioNet	93.95%	92.84%	95.24%	94.02%	93.80%	96.76%	90.62%	93.59%

Table 5. Class-wise accuracy and overall accuracy of various attention mechanisms on the Expert-I and Expert II annotated dataset before augmentation.

	Expert – I			Expert – II		
Model	**Overthinking**	**No overthinking**	**Overall accuracy**	**Overthinking**	**No overthinking**	**Overall accuracy**
SVM	76.11%	78.78%	77.45%	76.56%	75.75%	76.16%
CNN Attention	86.20%	50%	68.1%	77.77%	68.18%	72.97%
Self-Attention	85.71%	61.90%	73.8%	89.65%	55%	72.32%
Multi-Head Attention	70.96%	94.44%	82.70%	77.41%	66.66%	72.04%
SE-BioNet	87.5%	75.63%	81.56%	96.83%	75%	85.91%

no-overthinking samples. Similarly, for the Expert-II annotated dataset, SE-BioNet achieved the best overall accuracy of 93.80%, with class-wise accuracies of 90.62% and 96.97% for overthinking and no-overthinking classes, respectively. Multi-Head Attention also showed strong performance, achieving 91.50% and 92.24% accuracy on the Expert-I and Expert-II datasets, respectively, while Self-Attention and CNN obtained comparatively lower results. To assess whether the observed performance improvements were statistically significant, McNemar's test was conducted to compare the proposed SE-BioNet's predictions with those of the other models on the test set. The proposed SE-BioNet demonstrated significantly better performance than CNN Attention ($p < 0.01$), Self-Attention ($p < 0.05$), and Multi-Head Attention ($p < 0.05$) across both expert-annotated datasets. These results confirm that the performance gains achieved by the SE attention mechanism are not attributable to chance and provide statistical evidence of its effectiveness in learning discriminative biofield features (see Table 6A). Fig. 7 shows the confusion matrix and ROC curves for the results obtained. Overall, the results demonstrate that data augmentation significantly improves classification performance and model generalisation. Among all evaluated approaches, the proposed SE-BioNet consistently achieved the highest overall accuracy on both expert-annotated datasets. The balanced class-wise performance further suggests that the SE attention mechanism effectively enhances discriminative feature learning, enabling better separation of overthinking and no-overthinking biofield patterns than the other attention mechanisms considered in this study.

4.6 Impact of attention mechanism on classification

t-SNE visualisations show how clearly the model separates the two classes in a simple 2D plot. Fig. 8 shows the t-SNE visualisations of features learned by the baseline CNN and the proposed SE-BioNet on the Expert-I and Expert-II annotated datasets. In both datasets, the CNN features exhibit some overlap between the overthinking and no-overthinking classes, indicating limited class discrimination. In contrast, the SE-BioNet produces more compact and clearly separated clusters with reduced overlap. This improved feature separability is consistent with the higher classification accuracy achieved by SE-BioNet and demonstrates the effectiveness of the SE attention mechanism in learning discriminative biofield features.

4.7 Model evaluation

To evaluate the generalisation capability of the proposed SE-BioNet model, 5-fold stratified cross-validation was performed separately on the datasets annotated by Expert-I and Expert-II. As shown in Table 7, the model achieved

Table 6. Class-wise accuracy and overall accuracy of various attention mechanisms on Expert-I and Expert II annotated dataset (in %) after augmentation.

	Expert – I			Expert – II		
Model	**Overthinking**	**No overthinking**	**Overall accuracy**	**Overthinking**	**No overthinking**	**Overall accuracy**
SVM	84%	92.5%	88.27%	87.5%	83.33%	85.41%
CNN Attention	95.24%	79.41%	87.32%	95.24%	85.29%	90.26%
Self-Attention	92.19%	87.88%	90.03%	93.75%	84.85%	89.3%
Multi-Head Attention	88.89%	94.12%	91.50%	88.89%	95.59%	92.24%
SE-BioNet	95.24%	92.65%	93.95%	90.62%	96.97%	93.80%

Table 6. A. Statistical significance (p-values) of performance differences between SE-BioNet and other models using McNemar's test.

Comparison	Expert-I (p-value)	Expert-II (p-value)
SE-BioNet vs. CNN Attention	0.007**	0.009**
SE-BioNet vs. Self-Attention	0.023*	0.031*
SE-BioNet vs. Multi-Head Attention	0.018*	0.026*

Note: **$p < 0.01$, *$p < 0.05$ indicate statistically significant differences*

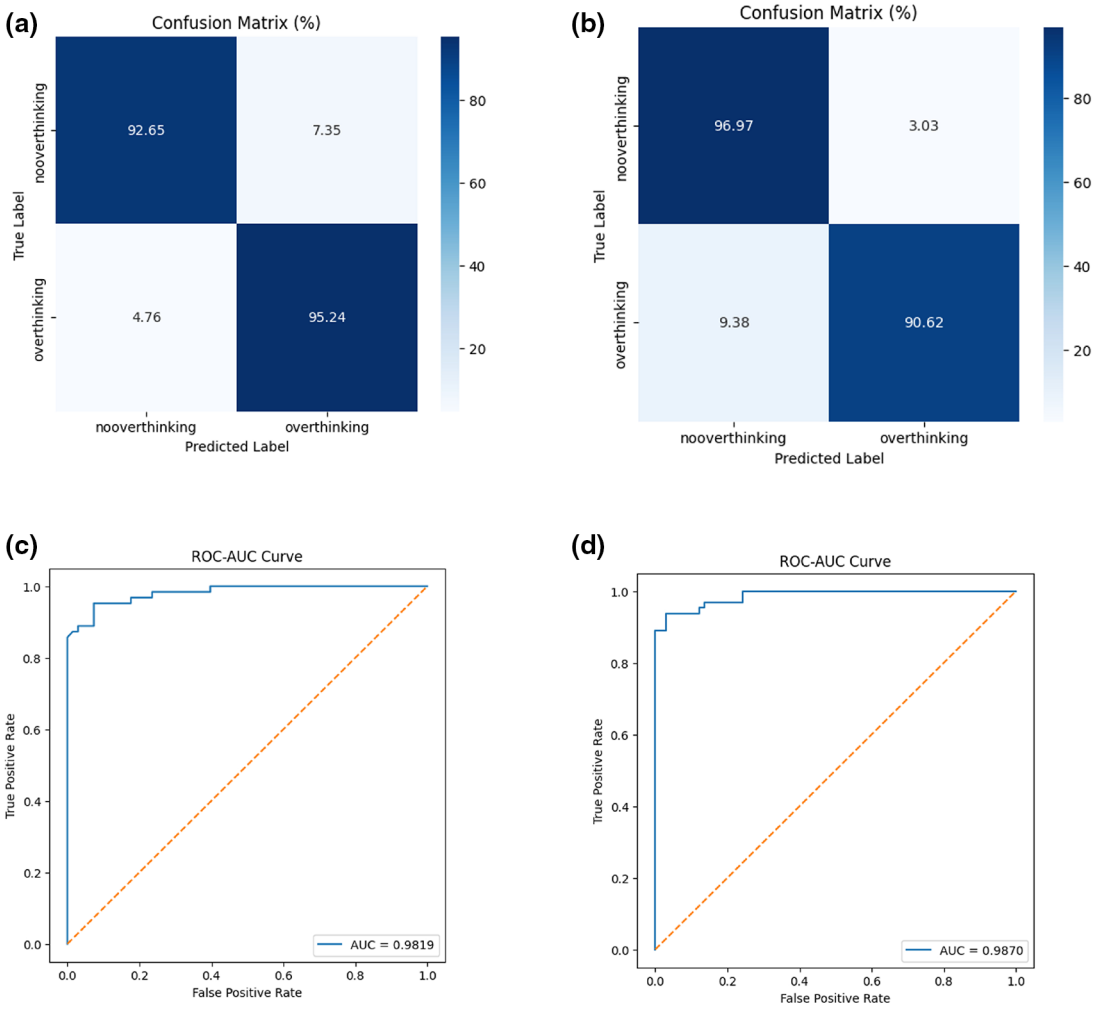


Fig. 7. Confusion matrices and ROC–AUC curves of the proposed SE-BioNet model: (a) Expert-I confusion matrix, Fig. 7 (b) Expert-II confusion matrix, Fig. 7 (c) Expert-I ROC–AUC curve, Fig. 7 (d) Expert-II ROC–AUC curve.

accuracy ranging from 90.65% to 93.66% for Expert-I and 90.32% to 93.66% for Expert-II across five folds. The mean cross-validation accuracy was 92.22% ± 1.17% for Expert-I and 92.08% ± 1.34% for Expert-II. The small standard deviation and limited variation in accuracy across folds indicate that the proposed model performs stably and consistently across different data partitions. These results demonstrate the robustness of the learned feature representations and suggest good generalisation ability across both expert-annotated datasets. The inference time per image was approximately 0.023 seconds on a Tesla T4 GPU, making the model suitable for real-time screening applications. Training for 50 epochs took approximately 45 minutes.

4.8 Inter- rater reliability

Inter-rater reliability was evaluated to assess the consistency of the annotation process. Two independent experts classified 857 samples into the Overthinking and No Overthinking categories. The level of agreement between the experts was measured using Cohen's Kappa coefficient (K), which quantifies agreement beyond chance and is widely used to evaluate the reliability of categorical annotations.[46,47] Cohen's Kappa is defined using equation 4:

$$k = \frac{P_0 - P_e}{1 - P_e} \tag{4}$$

where P_0 is the observed agreement and P_e is the expected agreement by chance. The observed agreement was computed as:

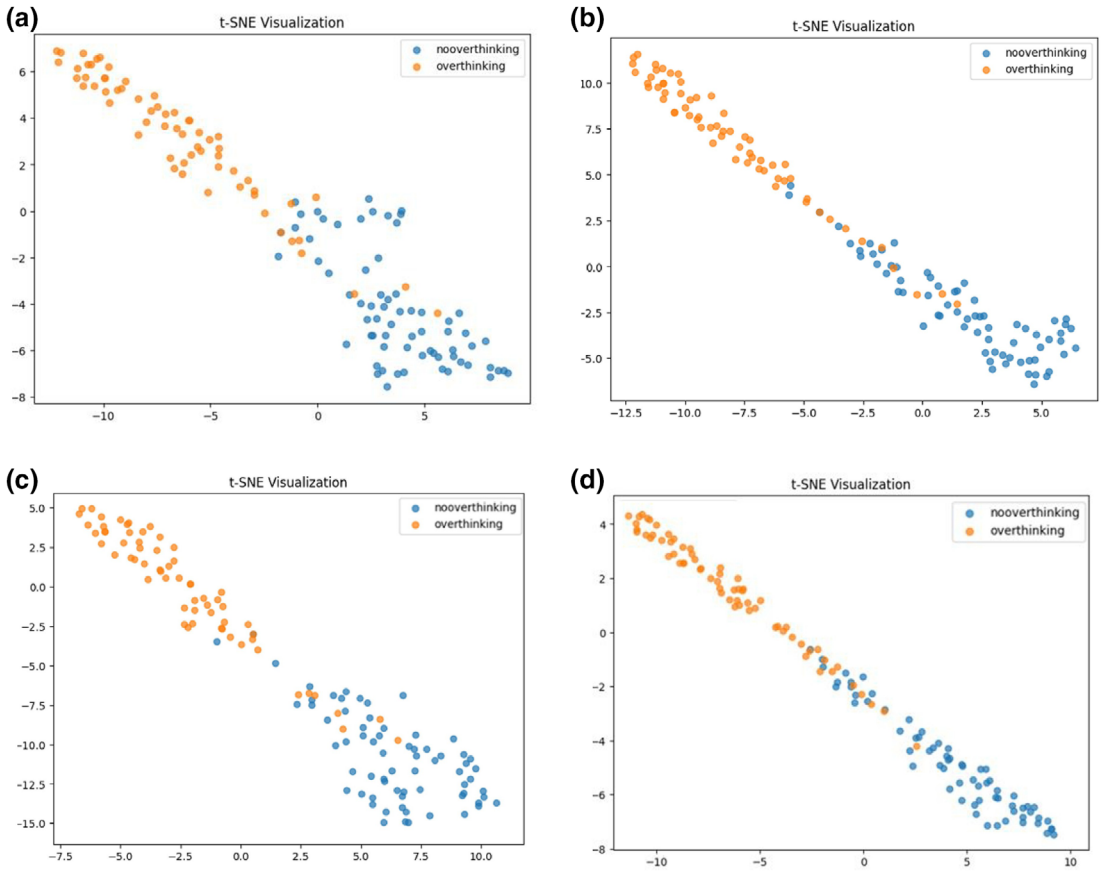


Fig. 8. t-SNE visualization (perplexity = 30, iterations = 1000) comparing feature separability achieved by the baseline CNN and the proposed SE-BioNet on Expert-I and Expert-II annotated datasets. Fig. 8(a). CNN (Expert-I) shows overlapping clusters, Fig. 8(b) SE-BioNet (Expert-I) demonstrates clear separation, Fig. 8 (c) CNN (Expert-II) exhibits limited separability, and Fig. 8 (d) SE-BioNet (Expert-II) achieves well-separated clusters, indicating the effectiveness of the SE attention mechanism in learning discriminative features.

Table 7. Five-fold cross-validation results of the proposed SE-BioNet model (in %).

	Accuracy	
Fold	**Expert – I**	**Expert - II**
1	92.65%	92.99%
2	93.66%	90.32%
3	90.65%	93.78%
4	92.65%	91.49%
5	91.49%	91.82%
Mean ± SD	92.22% ± 1.17%	92.08% ± 1.34%

$$P_0 = \frac{\text{Total agreed samples}}{\text{Total samples}} = \frac{432 + 405}{857} = \frac{837}{857} = 0.9766$$

The expected agreement by chance P_e was calculated based on the marginal label distributions of both annotators. Annotator 1 assigned 444 samples to NoOverthinking and 413 to Overthinking, while Annotator 2 assigned 440 and 417, respectively, out of 857 total samples. The chance agreement for each class was:

$$P_e = \left(\frac{444}{857} \times \frac{440}{857}\right) + \left(\frac{413}{857} \times \frac{417}{857}\right)$$

$$P_e = (0.5181 \times 0.5135) + (0.4819 \times 0.4865) = 0.2660 + 0.2344 = 0.5004$$

Substituting into the kappa formula:

$$k = \frac{0.9766 - 0.5004}{1 - 0.5004} = \frac{0.4762}{0.4996} = 0.9532$$

The resulting k = 0.9532 indicates almost perfect agreement (k > 0.80) between the two annotators,[48] confirming the high reliability of the annotation process. The 95% confidence interval for Cohen's κ was calculated as [0.932, 0.974], confirming the robustness of the inter-rater agreement.

4.9 Statistical test of significance

The Chi-square test was performed to assess the statistical significance of the agreement between the proposed SE-BioNet model's predictions and the annotations from the two medical experts. The test was conducted separately for Expert-I and Expert-II. The null hypothesis (H_1) states that there is a significant association (agreement) between the SE-BioNet predictions and the expert annotations, whereas the alternative hypothesis (H_1) states that no significant association exists. The degree of freedom is 1, and the critical χ^2 value at a 5% significance level is 3.841. The Chi-square statistics were computed using equation 5:

$$\sum \frac{(f_o - f_e)^2}{f_e} \tag{5}$$

Where f_o and f_e represent the observed and expected frequencies, respectively.

Table 8 presents the details of the Chi-square test for Medical Experts I and II. The computed values of the Chi-square statistic are 0.015393 for Expert-I and 0.246679 for Expert-II, and are observed to be lower than the critical value of 3.841 at 5% significance level (df = 1). Hence, the null hypothesis is accepted. It indicates significant agreement between the findings of the proposed SE-BioNet algorithm and the annotations from both medical Expert I and Expert II.

4.10 Comparative study with State-of-the-Art Methods

A comparative analysis of the proposed model against state-of-the-art methods is shown in Table 9. Most prior work has relied on EEG signals as the primary input modality. Aydin et al.[24] achieved 99.47% accuracy using a Cost Regulation Extreme Learning Machine on the LEMON dataset, and Kaushik et al.[23] reported up to 98.06% accuracy with a 1D-CNN on resting-state EEG. These accuracies largely reflect the well-structured, noise-controlled nature of EEG recordings under laboratory conditions. Ahire[19] achieved 90.21% accuracy using a 1D-CNN on task-based EEG data, while Kapitany et al.[20] reported 69.5% accuracy with EEGNet, highlighting variability in EEG-based classification depending

Table 8. Chi-square test between the results of proposed SE-BioNet algorithm and annotations by medical Expert I and Expert II.

Class	SE-BioNet algorithm (f_0)	Expert annotation (f_0)	Total	Expected f_e (SE-BioNet)	Expected f_e (medical expert)
Expert - I					
Positive	64	63	127	63.5000	63.5000
Negative	66	67	133	66.5000	66.5000
Total	130	130	260		
Expert - II					
Positive	60	64	124	62.0000	62.0000
Negative	70	66	136	68.0000	68.0000
Total	130	130	260		

Table 9. Comparison of state-of-the-art methods for detecting overthinking or rumination and related cognitive states using EEG, fMRI, and machine learning techniques.

Author	ML/DL method	Dataset	Performance metrics
Misaki et al. (2026)[17]	Whole-brain activation pattern decoding using individualized AutoML models (Auto-sklearn).	Primary dataset-[fMRI]	ROC-AUC: 0.91
Li et al. (2025)[18]	Random forest	Primary dataset	accuracy = 0.82, recall = 0.80, precision = 0.69
Ahire, N. (2025)[19]	1D-CNN, SVM, Logistic Regression;	Primary dataset- [EEG]	Best: 1D-CNN on TASK-EEG accuracy = 90.21%
(Kapitany M et al. 2024)[20]	EEGNET	Primary dataset- [EEG]	69.5%
Welkerling J et al, 2025[21]	Linear SVM decoder; features: alpha/beta/theta power + inter-electrode connectivity;	Primary dataset- [EEG]	Mean accuracy = 54.39% (SD = 8.76%) $t(23) = 1.74$, $p = 0.048$, Cohen's $d = 0.35$;
Wei L et al., 2025[22]	NBS-Predict (ML-based network edge selection + cross-validated ridge regression); tested FPN-DMN and FPN-SN edge sets separately	Primary dataset-[fMRI]	Cross-validated prediction: $r = 0.49$, $R^2 = 0.24$, $p = 0.001$ (FPN-DMN + SN edges); FPN-DMN: $r = 0.43$; FPN-SN: $r = 0.41$;
(Kaushik P et al. 2023)[23]	1D-CNN, LSTM, SVM, RF	Primary dataset- [rs-EEG + SART task-based EEG]	Best rs-EEG: 1D-CNN accuracy = 98.06%; Best task-EEG: LSTM accuracy = 91.42%
(Aydin et al., 2022)[24]	Cost Regulation Extreme Learning Machine	LEMON	accuracy = 99.47%
Bailey & Kang (2022)[25]	Portable EEG Monitoring	Primary dataset	Decrease in frontal beta activity
Silveira E 2020[26]	Linear support vector machine	Clinical and blood biomarkers variables	accuracy = 0.83%, sensitivity = 75%, specificity = 71%
Proposed model SE-BioNet	CNN with SE Attention	Primary dataset [biofield]	Expert- I accuracy = 93.95%, precision = 92.84%, recall of = 95.24%, F1-score = 94.02% Expert- II accuracy = 93.80%, precision = 96.76%, recall of = 90.62%, F1-score = 93.59%

on task design and feature representation. In the neuroimaging domain, Misaki et al.[17]) reported an ROC-AUC of 0.91 using individualised AutoML models on fMRI data, and Silveira et al.[26] achieved an AUC of 0.83 using a linear SVM on clinical biomarker variables. The proposed SE-BioNet integrates a Squeeze–Excitation attention mechanism to adaptively recalibrate channel-wise feature responses in biofield images, enabling the model to emphasise the most discriminative patterns associated with overthinking. A key distinction of the proposed framework is its input modality: biofield scan images captured using PIP. While PIP technology requires a controlled environment for accurate energy capture, it remains significantly less costly than EEG setups and neuroimaging systems such as fMRI or MRI scanners. Crucially, PIP-based biofield imaging is entirely non-invasive and radiation-free, making it a safer and more accessible alternative for real-world cognitive state monitoring.

5 Discussion

Despite the high classification performance achieved by the proposed model, a small number of samples were still misclassified. Fig. 9 presents representative examples of such misclassified cases from the test dataset. These errors may be attributed to the limited size of the dataset and the presence of overlapping or visually similar biofield patterns between

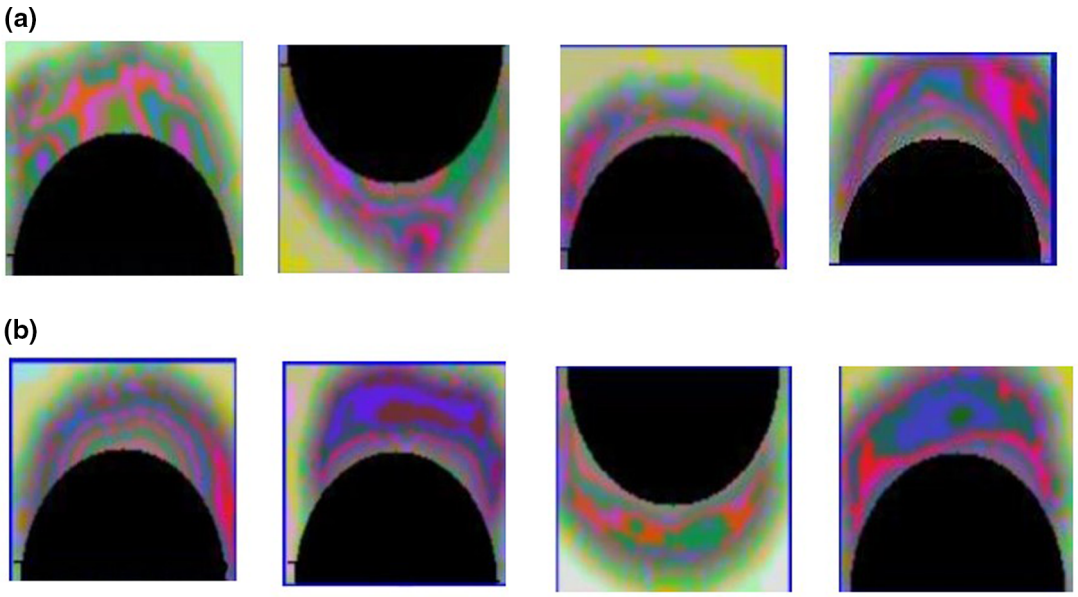


Fig. 9. (a) shows overthinking samples misclassified as no overthinking. Fig. 9 (b) shows no overthinking misclassified as overthinking samples.

the overthinking and no-overthinking classes, which reduces the model's discriminative capability and generalisation performance. Another key limitation is that the ground-truth labels rely on expert visual interpretation of biofield images rather than independent neurophysiological measurements such as EEG or fMRI. This reliance on subjective annotation may introduce inter-observer variability and limit the biological validation of the labels. In addition, the use of a patented biofield image acquisition technology restricts independent replication and validation by other research groups without access or permission, further affecting reproducibility. To address these limitations, future work should focus on constructing larger, multi-centre datasets to improve model robustness and generalisation. Performance may also improve by incorporating transfer learning with pretrained convolutional backbones and exploring hybrid attention mechanisms that combine channel and spatial recalibration. Furthermore, integrating multimodal physiological signals, such as heart rate variability (HRV) and EEG, could improve biological interpretability and strengthen the validity of the findings. The misclassified samples shown in Figure 9 were further analysed to understand the potential sources of error. Visual inspection revealed that misclassified overthinking samples (Figure 9a) often exhibited subtle colour variations in the upper cranial region, qualitatively similar to no-overthinking patterns, suggesting that these cases lie near the decision boundary. Similarly, misclassified no-overthinking samples (Figure 9b) occasionally displayed scattered dark pink patterns that the model may have interpreted as indicative of cognitive engagement. This suggests that a larger dataset with more diverse examples, particularly near the decision boundary, could further improve classification performance. Beyond the limitations already discussed, this study is constrained by its single-centre design, which may limit generalizability to diverse populations and imaging setups. The binary classification framework (overthinking vs no overthinking) does not capture the potential gradations or subtypes of repetitive negative thinking. Furthermore, the lack of external validation on an independent dataset from a different institution or imaging system limits the robustness of the reported performance metrics. Future multi-centre studies with larger, more diverse cohorts are needed to establish the generalizability of this approach. Overall, the proposed SE-BioNet demonstrates that deep learning with channel-wise attention can effectively learn discriminative features from PIP-based biofield images that align with expert-assigned labels. The balanced class-wise performance and robust cross-validation results suggest that the SE attention mechanism provides a meaningful advantage over baseline and alternative attention mechanisms. However, as discussed in the following section, these findings must be interpreted in light of several important limitations that define the scope of the current contribution.

6 Limitations and future work

Despite the promising classification performance achieved by the proposed SE-BioNet model, this study has several important limitations that must be acknowledged. These limitations define the boundaries of the current contribution and inform a concrete roadmap for future research.

6.1 Absence of neurophysiological validation

The most significant limitation of this study is the lack of independent neurophysiological validation. The ground-truth labels used for training and evaluation were derived from expert visual interpretation of PIP biofield images rather than from established neurophysiological measurements such as Electroencephalography (EEG) or functional Magnetic Resonance Imaging (fMRI). While EEG and fMRI provide direct, biologically grounded measures of brain activity—capturing alpha/beta wave fluctuations, Default Mode Network (DMN) activation, and medial prefrontal cortex (mPFC) engagement associated with rumination and repetitive negative thinking—the relationship between PIP colour patterns and these underlying neural processes remains hypothetical. Consequently, the high classification accuracies reported in this study (93.95% and 93.80%) reflect the model's ability to match expert-assigned labels, not its ability to detect a validated neurophysiological state. Without concurrent EEG or fMRI recordings, it is not possible to confirm whether the biofield patterns the model identifies correspond to genuine cognitive processes associated with overthinking or to other physiological or environmental factors. This limitation is fundamental and must be addressed before making any clinical claims.

6.2 Patented, single-centre dataset and generalisability

A second major limitation concerns the dataset itself. All biofield images were acquired at a single institution—the Institute of Biofield Evaluation and Research Centre (IBERC), Thane, India—using a patented acquisition methodology (Indian Patent No. 528035). This has three important consequences. First, the dataset reflects a single imaging setup, lighting environment, camera configuration, and operator protocol, which may introduce centre-specific biases that the model could inadvertently learn. Second, the patented nature of the technology restricts independent replication and validation by other research groups without access to the proprietary system or permission from the patent holder, limiting reproducibility. Third, the dataset, while substantial for an exploratory study (857 labelled images), remains modest for deep learning applications and may not capture the full diversity of biofield patterns across different populations, age groups, geographical regions, or acquisition conditions. These factors collectively limit the generalisability of the reported performance metrics beyond the specific context in which the data were collected.

6.3 Exploratory proof-of-concept status

This study is a hypothesis-generating proof of concept, not a validated clinical tool. The central hypothesis—that sustained, repetitive cognitive activity (referred to here as overthinking) manifests as detectable patterns in PIP-based cranial biofield images—has not been independently established. The current work demonstrates that a deep learning model with Squeeze-and-Excitation attention can learn discriminative features that align with expert-assigned labels, but this alignment does not constitute biological validation. The framework should therefore be viewed as a methodological contribution that motivates further investigation, rather than as a diagnostic or screening instrument ready for clinical deployment. Any clinical interpretation of the results would be premature and potentially misleading.

6.4 Roadmap for future work

Addressing the above limitations requires a structured, multi-stage research programme. We propose the following concrete roadmap:

(i) **Multi-centre data collection with concurrent neurophysiological recording.** The highest priority is to conduct a multi-centre study in which PIP biofield images are acquired simultaneously with EEG and/or fMRI recordings. This would allow direct correlation between biofield patterns and established neural markers of rumination and repetitive negative thinking, such as frontal alpha asymmetry, theta/beta ratio, DMN hyperactivation, and mPFC connectivity. Concurrent HRV monitoring could further capture autonomic correlates. Such a design would provide the first direct test of whether PIP colour patterns reflect the cognitive states they are hypothesised to represent.

(ii) **Multi-expert labelling with standardised protocols.** Future studies should employ a larger panel of independent experts and a formalised, quantifiable annotation protocol. Inter-rater reliability should be reported using Cohen's κ or intraclass correlation coefficients, and disagreements should be adjudicated through consensus or third-party review. This would reduce the risk of expert-specific bias and improve the reliability of ground-truth labels.

(iii) **Independent external validation.** The model should be evaluated on an independent dataset acquired at a different institution, using a different PIP system or acquisition protocol. External validation on unseen data

from a different centre is the gold standard for assessing generalisability in medical AI and would substantially strengthen the credibility of the framework.

(iv) **Exploration of transfer learning and hybrid attention mechanisms.** Performance may improve by incorporating pretrained convolutional backbones (e.g., ResNet, EfficientNet) through transfer learning and by exploring hybrid attention mechanisms that combine channel-wise recalibration (as in SE) with spatial attention. This could enhance feature learning from limited data and improve robustness.

(v) **Longitudinal and multi-class studies.** Future work should move beyond binary classification to differentiate between subtypes or severity levels of repetitive negative thinking and should include longitudinal monitoring to assess whether biofield patterns change over time or in response to intervention.

(vi) **Progressive automation of the expert-interpretation pipeline.** As the dataset grows and validation improves, the expert-interpretation process itself could be progressively automated, reducing reliance on subjective assessment and enabling large-scale screening applications.

(vii) **Open science and reproducibility.** Where possible, future studies should aim to release de-identified datasets, standardised acquisition protocols, and source code to facilitate independent replication. While the patented acquisition technology presents a constraint, collaboration agreements or shared-access frameworks could enable broader validation.

In summary, the current study provides a methodological foundation and preliminary evidence that warrants further investigation, but it does not yet establish a validated link between biofield imaging and cognitive states. The roadmap outlined above offers a clear path toward stronger scientific validation and, ultimately, toward the development of a reliable, non-invasive tool for cognitive state assessment.

Ethics statement

All participants voluntarily participated in the study and provided informed consent for image acquisition and the use of anonymised data for research purposes. Participant confidentiality and privacy were maintained throughout the study.

Software availability

Source code available from https://github.com/Supriya-Patil84/Paper4_code

Archived software available from: https://doi.org/10.5281/zenodo.23282085

License: [MIT License].

Data availability statement

The biofield images used in this study contain sensitive personal content, as participants were photographed in minimal clothing to allow accurate visualization of energy emissions. Owing to ethical, cultural, and legal obligations to protect participant privacy and identity, the raw image dataset cannot be shared publicly. Access to de-identified data may be considered upon reasonable request to the corresponding author, Dr. Shivanand S. Gornale (shivanand1971@rcub.ac.in) and Supriya Patil (supriyapatil8409@gmail.com) subject to approval by the institutional ethics committee and a data use agreement.

Acknowledgment

The authors sincerely acknowledge the Institute of Biofield Evaluation and Research Centre (IBERC), Thane West, Maharashtra, India, for their invaluable support in facilitating the experimental validation and providing access to the annotated dataset used in this study. The authors are especially grateful to Dr. Arvind K. Takle, Radiologist, Panvel Polyclinic, Panvel, Mumbai, Maharashtra, for his expert contribution in annotating the dataset. The authors also extend their sincere appreciation to the reviewers for their constructive comments and valuable suggestions, which helped improve the quality and presentation of the work.

References

1. Nolen-Hoeksema S: **Responses to depression and their effects on the duration of depressive episodes.** *J. Abnorm. Psychol.* 1991; **100**(4): 569–582.
Publisher Full Text

2. Ehring T, Watkins ER: **Repetitive negative thinking as a transdiagnostic process.** *Int. J. Cogn. Ther.* 2008; **1**(3): 192–205.
Publisher Full Text

3. Spinhoven P, van Hemert AM, Penninx BW: **Repetitive negative thinking as a predictor of depression and anxiety: A longitudinal cohort study.** *J. Affect. Disord.* 2018; **241**: 216–225.
PubMed Abstract | Publisher Full Text

4. World Health Organization: *WHO's response to health emergencies: Annual report 2023.* Geneva, Switzerland: WHO; 2024.
Reference Source

5. Murthy RS: **National mental health survey of india 2015–2016.** *Indian J. Psychiatry.* 2017; **59**(1): 21–26.
PubMed Abstract | Publisher Full Text | Free Full Text

6. Putman P, van Peer J, Maimari I, *et al.*: **EEG theta/beta ratio in relation to fear-modulated response inhibition, attentional control, and affective traits.** *Biol. Psychol.* 2010; **83**(1): 73–78.
PubMed Abstract | Publisher Full Text

7. Hamilton JP, Farmer M, Fogelman P, *et al.*: **Depressive rumination, the default-mode network, and the dark matter of clinical neuroscience.** *Biol. Psychiatry.* 2015; **78**(4): 224–230.
Publisher Full Text

8. Whitfield-Gabrieli S, Ford JM: **Default mode network activity and connectivity in psychopathology.** *Annu. Rev. Clin. Psychol.* 2012; **8**: 49–76.
Publisher Full Text

9. Zhou Y, Friston KJ, Zeidman P, *et al.*: **The hierarchical organization of the default, dorsal attention and salience networks in adolescents and young adults.** *Cereb. Cortex.* 2020; **30**(2): 824–835.
PubMed Abstract | Publisher Full Text | Free Full Text

10. Euston DR, Gruber AJ, McNaughton BL: **The role of medial prefrontal cortex in memory and decision making.** *Neuron.* 2012; **76**(6): 1057–1070.
Publisher Full Text

11. Kim HG, Cheon EJ, Bai DS, *et al.*: **Stress and heart rate variability: A meta-analysis and review of the literature.** *Psychiatry Investig.* 2018; **15**(3): 235–245.
PubMed Abstract | Publisher Full Text | Free Full Text

12. Korotkov K: *The Energy of Consciousness: How to Measure the Immeasurable.* Moscow, Russia: Russian Ministry of Health Publishing; 2014.

13. Rubik B, Muehsam D, Hammerschlag R, *et al.*: **Biofield science and healing: History, terminology, and concepts.** *Global Advances in Health and Medicine.* 2015; **4**(Suppl): 8–14.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Schwabl H, Klima H: **Spontaneous ultraweak photon emission from biological systems and the endogenous light field.** *Research in Complementary and Natural Classical Medicine.* 2005; **12**(2): 84–89.
PubMed Abstract | Publisher Full Text

15. van Wijk R: *Light in Shaping Life: Biophotons in Biology and Medicine.* Geldermalsen, Netherlands: Meluna Research; 2014.

16. Deo A: *Aabhamandal.* Pune, India: Biofield Evaluation;2nd ed 2017.

17. Misaki M, Tsuchiyagaito A, Guinjoan SM: **Age differences in neural discriminability of rumination and worry.** *Commun. Biol.* 2026; **9**, **9**: Article 223.
Publisher Full Text

18. Li K, Ren L, Li X, *et al.*: **Exploring the role of repetitive negative thinking in the transdiagnostic context of depression and anxiety in children.** *BMC Psychol.* 2025; **13**(1): Article 902.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Ahire N: **EEG-derived brainwave patterns for depression diagnosis via hybrid machine learning and deep learning frameworks.** *Appl. Neuropsychol. Adult.* 2025; **32**(1): 1–10.
Publisher Full Text

20. Kapitany-Foveny M, Vetro M, Revy G, *et al.*: **EEG based depression detection by machine learning: Does inner or overt speech condition provide better biomarkers when using emotion words as experimental cues?.** *J. Psychiatr. Res.* 2024; **178**: 66–76.
Publisher Full Text

21. Welkerling J, Schneeweiss P, Wolf S, *et al.*: **Decoding ruminative states from neurophysiological patterns.** *medRxiv.* 2024. (Preprint).
Publisher Full Text

22. Wei L, Dong H, Zhang Z, *et al.*: **Decoding ruminative reflection in healthy individuals: The role of triple network connectivity.** *Int. J. Clin. Health Psychol.* 2024; **24**(4): 100508.
PubMed Abstract | Publisher Full Text | Free Full Text

23. Kaushik P, Yang H, Roy PP, *et al.*: **Comparing resting state and task-based EEG using machine learning to predict vulnerability to depression in a non-clinical population.** *Sci. Rep.* 2023; **13**: Article 7467.
PubMed Abstract | Publisher Full Text | Free Full Text

24. Aydin S, Akin B: **Machine learning classification of maladaptive rumination and cognitive distraction in terms of frequency specific complexity.** *Biomedical Signal Processing and Control.* 2022; **77**: Article 103740, 103740.
Publisher Full Text

25. Bailey AW, Kang HK: **Walking and sitting outdoors: Which is better for cognitive performance and mental states?.** *Int. J. Environ. Res. Public Health.* 2022; **19**(24): Article 16638.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Silveira EdM, Passos C I, Scott J, *et al.*: **Decoding rumination: A machine learning approach to a transdiagnostic sample of outpatients with anxiety, mood and psychotic disorders.** *J. Psychiatr. Res.* 2020; **121**: 207–213.
Publisher Full Text

27. Bryndin E, Bryndina I: **Technological diagnostics of human condition according to spectral analysis of biofield.** *Advances in Bioscience and Bioengineering.* 2019; **7**(3): 50–54.
Publisher Full Text

28. Gao F, Li L, Hu Q: **Study on auxiliary diagnosis of depression based on gas discharge visualization technology.** *Integrative Medicine Reports.* 2024; **3**(1): 132–144.
Publisher Full Text

29. Mangione L, Swengros D, Anderson JG: **Mental health wellness and biofield therapies: An integrative review.** *Issues Ment. Health Nurs.* 2017; **38**(11): 930–944.
PubMed Abstract | Publisher Full Text

30. Olbrich S, Arns M: **EEG biomarkers in major depressive disorder: Discriminative power and prediction of treatment response.** *Int. Rev. Psychiatry.* 2013; **25**(5): 604–618.
Publisher Full Text

31. Uchida S, Iha T, Yamaoka K, *et al.*: **Effect of biofield therapy in human brain.** *J. Altern. Complement. Med.* 2012; **18**(9): 875–879.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Trivedi MK, Branton A, Trivedi D, *et al.*: **The role of biofield energy treatment on psychological symptoms, mental health disorders, and stress-related quality of life in adult subjects: A randomized controlled clinical trial.** *Journal of General and Family Medicine.* 2023; **24**(3): 154–163.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Chhabra G, Onyema EM, Kumar S, *et al.*: **Human emotions recognition, analysis and transformation by the bioenergy field in smart grid using image processing.** *Electronics.* 2022; **11**(23): Article 4059.
Publisher Full Text

34. Gornale SS, Patil SS, Deo A, *et al.*: **Quantitative analysis of biofield energy imbalance through color segmentation and classification using machine learning techniques.** *Proc. the 2024 IEEE 9th International Conference for Convergence in Technology (I2CT).* Pune, India: Apr. 2024; pp.1–5.
Publisher Full Text

35. Veerabrahmachar R, Bista S, Bokde R, *et al.*: **Immediate effect of Nada yoga meditation on energy levels and alignment of seven chakras as assessed by electro-photonic imaging: A randomized controlled crossover pilot study.** *Adv. Mind Body Med.* 2023; **37**(1): 11–16.
PubMed Abstract

36. Hu J, Shen L, Albanie S, *et al.*: **Squeeze-and-excitation networks.** *IEEE Trans. Pattern Anal. Mach. Intell.* 2020; **42**(8): 2011–2023.
PubMed Abstract | Publisher Full Text

37. Roy AG, Navab N, Wachinger C: **Concurrent spatial and channel 'squeeze & excitation' in fully convolutional networks.** *Proc. the 2018 International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI).* Granada, Spain: Sep. 2018; pp.421–429.
Publisher Full Text

38. Huang KA, Alkadri A, Prakash N: **Employing Squeeze-and-Excitation Architecture in a Fine-Tuned CNN for MRI Tumor Classification.** *Cureus.* 2025; **17**: e80084.
PubMed Abstract | Publisher Full Text | Free Full Text

39. Gencer G, Gencer K: **Advanced retinal disease detection from oct images using a hybrid squeeze and excitation enhanced model.** *PLOS ONE.* 2025; **20**(2): Article e0318657.
Publisher Full Text

40. Cuimei L, Zhiliang Q, Nan J, *et al.*: **Human face detection algorithm via Haar cascade classifier combined with three additional classifiers.** *Proc. the 2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI).* Yangzhou, China: Oct. 2017; pp.483–487.
Publisher Full Text

41. Perez L, Wang J: **The effectiveness of data augmentation in image classification using deep learning.** *arXiv (Preprint), arXiv: 1712.04621.* 2017.
Reference Source

42. Shorten C, Khoshgoftaar TM: **A survey on image data augmentation for deep learning.** *J. Big Data.* 2019; **6**: Article 60.
Publisher Full Text

43. Iqbal MI, Avianto D: **Squeeze-and-excitation networks and attention mechanism in automatic detection of coffee leaf diseases based on images.** *Journal of Soft Computing Exploration.* 2024; **5**(4): 320–331.
Publisher Full Text

44. Reddy VKM, Raghavendar Raju L, Sai Prasad K, *et al.*: **Enhanced effective convolutional attention network with squeeze-and-excitation inception module for multi-label clinical document classification.** *Sci. Rep.* 2025; **15**: Article 16988.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

45. Hayat M: **Squeeze & Excitation joint with combined channel and spatial attention for pathology image super-resolution.** *Franklin Open.* 2024; **8**: Article 100170.
Publisher Full Text

46. Cohen J: **A coefficient of agreement for nominal scales.** *Educ. Psychol. Meas.* 1960; **20**(1): 37–46.
Publisher Full Text

47. McHugh ML: **Interrater reliability: The kappa statistic.** *Biochem. Med.* 2012; **22**(3): 276–282.
PubMed Abstract | **Publisher Full Text**

48. Landis JR, Koch GG: **The measurement of observer agreement for categorical data.** *Biometrics.* 1977; **33**(1): 159–174.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ?

Version 1

Reviewer Report 04 November 2026

https://doi.org/10.5256/f1000research.212262.r537488

© **2026 Chhabra G.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Gunjan Chhabra
Department of CSE, School of Science and Technology, Swami Rama Himalayan University, Dehradun, Uttarakhand, India

Summary of the work
The authors train a small convolutional neural network with Squeeze-and-Excitation (SE) blocks, called SE-BioNet, to classify cropped cranial regions of Polycontrast Interference Photography (PIP) biofield images as “overthinking” or “no overthinking”. The dataset contains 857 subjects from a single centre (IBERC, Thane). Two experts labelled the images independently (Cohen’s κ = 0.953). On a 15% test split the model reaches about 93.9% (Expert-I) and 93.8% (Expert-II) accuracy, and 5-fold cross-validation gives about 92%. It is compared with an SVM and with CNN, self-attention and multi-head attention variants. The authors describe the study as exploratory and hypothesis-generating, and state that the link between PIP colour patterns and cognition has not been validated.
I appreciate the candour of the limitations section, the release of the code, and the decision to split the data before augmentation. However, the design cannot support the claim made in the title, and the statistical analysis and several reported numbers need substantial correction. My comments rely partly on my own arithmetic from the printed values, so I ask the authors to verify them against the raw outputs.

1. Presentation and literature
Must address

- SE-BioNet’s pre-augmentation accuracy is 81.58% in Table 4 but 81.56% in Table 5 and the text. Other pre-augmentation values also differ between the two tables (for example CNN Attention 71.43% vs 68.1%, Multi-Head Attention 79.59% vs 82.70%). Please reconcile all tables with the raw outputs and state the test-set size used for each.
- Table 3 contains a 50,176 × 256 dense layer, which alone has about 12.8 million parameters. This contradicts the stated total of 2.8 million and the claimed 3.2% SE overhead. Please recheck.
- “Overall accuracy” appears to equal the mean of the two class-wise accuracies (for example (95.24 + 92.65)/2 = 93.95), which is balanced accuracy. Please relabel it or report plain

accuracy.
- Please check several citations. Ref. 25 is a walking-versus-sitting behavioural study but is listed as “portable EEG monitoring”. Ref. 9 is cited both for default-mode network organisation and for SE-based segmentation. The original SE-Net paper dates from 2018.

Should address
- Add literature that critically examines the validity of biofield, GDV and PIP methods, and literature on shortcut learning in small medical-imaging datasets.
- Table 9 places the present results next to EEG and fMRI studies that have different targets and ground truths. Please remove this comparison or qualify it heavily.

2. Study design and technical soundness

Must address
- There is no independent ground truth. Labels come from expert visual judgement of red and dark-pink colour patches. The authors should either (a) collect a validated measure from the same participants, such as a rumination or repetitive-thinking questionnaire or EEG, or (b) reframe the study explicitly as learning expert annotations of PIP colour patterns, with a matching change to the title and abstract.
- Please add control experiments to rule out shortcut learning: the same model on non-cranial or background regions, a randomised-label control, and checks for association with age, sex, acquisition date and operator.
- Please declare the annotators’ relationship to the institute that holds the patent, and revise the “no competing interests” statement if appropriate. State whether the experts were blinded to each other and to participant information.
- About 48% of a general volunteer sample is labelled “overthinking”. Please discuss whether this prevalence is plausible for the construct.

Should address
- State whether augmentation was restricted to the training folds in cross-validation, and whether the test set influenced model selection.
- Describe the baselines fully, and include a plain CNN without attention and a pretrained backbone.
- Justify vertical-flip augmentation for anatomically oriented images (Fig. 5), or test without it.
- Confirm that participant consent covers labelling individuals with a psychological state from their images, and discuss the risk of over-interpretation.

3. Replicability of methods

Must address
- Provide an operational labelling protocol: colour thresholds, area fractions and example images for each class.
- Describe the acquisition-to-colour-map pipeline as far as the patent allows (camera model and settings, lighting). If it cannot be disclosed, state this clearly as a limitation.
- Specify the ROI rule (size and offset relative to the face box, mask construction) and report Haar-cascade failure rates. The statement in Section 3.2 that no segmentation was applied because ROIs were “already extracted” is confusing and should be clarified.

Should address
- Report random seeds, input normalisation, weight initialisation, the model-selection criterion, software versions, and the SVM features and kernel.
- Confirm that the repository contains split indices, seeds and baseline code.

4. Statistical analysis

Must address

- The chi-square test in Section 4.9 should be removed or redone. Table 8 compares total counts of predicted and labelled positives, not per-image agreement, so it cannot show agreement. The hypotheses are also reversed, and failing to reject a null hypothesis is not evidence of agreement. Please replace it with the model-versus-expert confusion matrix, Cohen's κ and its confidence interval.
- The McNemar p-values (0.007 to 0.031) look implausible for a test set of about 130 images with differences of 2 to 3 points. Please report the discordant-pair counts, use exact tests and correct for multiple comparisons. If the p-values cannot be reproduced, the significance claims should be withdrawn.
- Please add 95% confidence intervals (for example by bootstrap) for all accuracies, F1 scores and AUCs. With so few test images, most differences between models are probably within noise.
- Please report mean ± SD over multiple seeds and splits for every model before claiming that SE attention is superior.

Should address

- Acknowledge that the two expert datasets agree on about 98% of images, so they are near-duplicates and not independent replications.
- Within-site cross-validation SD should not be described as evidence of "robust generalisation".

5. Data availability

Must address

- Please deposit in a public repository the de-identified ROI images, both sets of expert labels, the split and fold assignments, and the per-image model predictions. The ROI crops show an abstract colour map above the head and appear to carry little identification risk, although the ethics committee should confirm this. The privacy concern applies mainly to the full-body scans.
- If any item truly cannot be shared, please state exactly which and why, and describe a concrete access route in place of "on reasonable request".

6. Conclusions

Must address

- Reword the title, contributions list and results text to say that the model reproduces expert-assigned labels, and does not "detect overthinking".
- Remove or support the claims of "real-time screening", "a safer and more accessible alternative to EEG/fMRI", "reducing reliance on subjective assessments" (circular, since the labels are themselves subjective) and "robust generalisation".
- Before augmentation, the text says SE-BioNet performed best, but Table 5 shows Multi-Head Attention ahead (82.70% vs 81.56%). Please correct this.
- Treat the t-SNE plots as illustrative only, not as evidence of discriminative features.

Points that must be addressed for the article to be scientifically sound

1. Reframe the claim (title, abstract, contributions) as reproduction of expert labels, or add independent validation.
2. Run confound and shortcut-learning controls.
3. Replace the chi-square analysis with a proper model-versus-expert agreement analysis.
4. Reproduce or withdraw the McNemar results, correct for multiple testing, and add confidence intervals and multi-seed results.
5. Reconcile all tables and correct the parameter count.
6. Release the ROI images, both label sets, splits and predictions, or justify non-release

specifically.
7. Provide an operational labelling protocol, the ROI rule and acquisition details as far as possible.
8. Clarify the annotators' independence and update the competing-interests statement.
9. Remove unsupported claims about real-time screening, superiority over EEG and fMRI, and generalisation.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
No

Are all the source data underlying the results available to ensure full reproducibility?
No

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: My area of research includes AI, ML, Image processing, Human Biofield, Human -centric AI.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research