Interpretability of Pranayama Effect on Brain Waves Using Explainable Deep Learning

A Project Work Synopsis

Submitted in the partial fulfilment for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE WITH SPECIALIZATION IN

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted by:

Siddhant Gupta(22BAI70088)

Anirudh Sharma(22BAI70109)

Vedansh Maheshwari(22BAI70056)

Shaurya Maan(22BAI70139)

Under the Supervision of:

Dr. Raghav Mehra(E16302)

CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,

PUNJAB

NOV, 2025

Abstract

The project aims to scientifically analyze the impact of *Pranayama*, an ancient yogic breathing practice, on human brain activity using *Electroencephalogram (EEG)* signals. EEG data is collected before, during, and after different Pranayama techniques to study variations in brain wave patterns such as Delta, Theta, Alpha, Beta, and Gamma. The signals are preprocessed to remove noise and artifacts, followed by feature extraction and classification using deep learning models such as *Convolutional Neural Networks (CNNs)* and *Recurrent Neural Networks (RNNs)*. To ensure transparency and trust in the model's predictions, *Explainable Artificial Intelligence (XAI)* techniques like *Grad-CAM, SHAP,* and *LIME* are incorporated, enabling visualization of the most influential brain wave features that drive classification decisions. The study is expected to reveal measurable changes in EEG patterns associated with mental states like relaxation, focus, and meditation during Pranayama. The outcome will be a research-backed, interpretable AI framework and visualization dashboard that bridges traditional wellness practices with modern computational neuroscience.

Table of Contents

Abstract

1. Introduction

1.1 Problem Definition

1.2 Project Overview

1.3 Software Specifications

2. Literature Survey

2.1 Existing System

2.2 Proposed System

2.3 Literature Review Summary

3. Problem Formulation.

4. Research Objective.

5. Methodologies

6. Conclusion.

7. Reference.

1. INTRODUCTION

1.1 Problem Definition

Although *Pranayama*, a yogic breathing technique, has been widely practiced for centuries and is known to promote relaxation, focus, and mental clarity, its **neurophysiological effects** are not fully understood in a quantifiable and interpretable manner. Traditional studies examining brain activity during Pranayama using EEG have primarily relied on conventional statistical or signal processing methods, which provide limited insights and often fail to explain *why* specific changes occur in brain wave patterns.

Furthermore, while *Deep Learning* models can effectively classify complex EEG patterns and identify cognitive states, they typically function as **black-box systems**, offering little transparency into their internal decision-making processes. This lack of interpretability makes it difficult for researchers and practitioners to trust or validate the model's conclusions, particularly in health and neuroscience applications.

Therefore, there is a clear need for a system that can **analyze and interpret EEG signals during Pranayama using transparent, explainable AI techniques**. The challenge lies in developing a framework that not only classifies mental states such as relaxation, focus, and meditative awareness but also provides a clear explanation of which EEG frequency bands or temporal patterns influence these classifications.

This project addresses that problem by combining **EEG-based deep learning models** with **Explainable Artificial Intelligence (XAI)** methods such as Grad-CAM, LIME, and SHAP to uncover the underlying relationships between breathing patterns and brain wave activity. The goal is to create a scientifically grounded, interpretable model that demonstrates the measurable and explainable impact of Pranayama on human brain function.

1.2 Problem Overview

Pranayama, an ancient yogic breathing practice, has been widely recognized for its benefits in enhancing mental clarity, emotional stability, and overall well-being. Scientific studies have suggested that controlled breathing patterns can influence the autonomic nervous system and alter brain wave activity, promoting relaxation and cognitive balance. However, despite these findings, there remains a significant gap in **scientific interpretation and visualization** of how specific Pranayama techniques affect the human brain at the neural level.

Electroencephalography (EEG) provides a powerful non-invasive method to record brain activity and study variations in frequency bands such as Delta, Theta, Alpha, Beta, and Gamma. These frequency bands correspond to different mental and physiological states—

ranging from deep relaxation to alert focus. While traditional analysis methods, such as spectral analysis or statistical comparisons, can capture general changes, they lack the capacity to automatically detect complex patterns or provide deeper interpretability.

The recent advancement of *Deep Learning* models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), offers new possibilities for analyzing EEG data with high accuracy. However, these models often act as **black boxes**, providing limited understanding of how they reach their conclusions. This lack of transparency poses a major challenge in fields related to neuroscience, cognitive science, and health informatics.

To overcome this limitation, the integration of *Explainable Artificial Intelligence (XAI)* techniques—such as Grad-CAM, LIME, and SHAP—can bring interpretability and trust to AI models by identifying which brain wave features most significantly contribute to their predictions. By combining EEG analysis, deep learning, and XAI, this project seeks to **scientifically decode and visualize the impact of Pranayama on brain activity**, enabling both quantitative and explainable insights into how breathing influences mental states like relaxation, focus, and meditation.

1.3 Software Specification

1. **Operating System**

- **Primary OS:** Windows 10/11 or Ubuntu 20.04+ (Linux recommended for better performance in deep learning tasks)
- **Compatibility:** Supports Python environments, GPU acceleration, and open-source EEG processing libraries.

2. Programming Language

- **Python 3.8 or higher**
  - Chosen for its simplicity, extensive scientific libraries, and strong community support in AI, machine learning, and biomedical signal processing.

3. Development Environment

- **Jupyter Notebook / Google Colab:** For coding, testing, and visualizing intermediate results interactively.
- **PyCharm / VS Code:** For integrated project development and model deployment.

4. Libraries and Frameworks

a. Deep Learning & Machine Learning

- **TensorFlow / Keras:** For building and training CNN and RNN models on EEG data.

- **PyTorch:** Alternative framework for model experimentation and explainability research.
- **Scikit-learn:** For preprocessing, feature scaling, and classical machine learning models (SVM, Random Forest).

b. EEG Signal Processing

- **MNE-Python:** For EEG data import, filtering, artifact removal, and band-power extraction.
- **SciPy & NumPy:** For mathematical computations, Fourier transforms, and digital signal processing.
- **PyWavelets:** For wavelet-based time–frequency analysis.

c. Explainable AI (XAI)

- **SHAP (SHapley Additive Explanations):** For global and local feature importance interpretation.
- **LIME (Local Interpretable Model-Agnostic Explanations):** For local instance-level model explanation.
- **Grad-CAM:** For visualizing important EEG spectrogram regions influencing CNN predictions.

d. Visualization Tools

- **Matplotlib / Seaborn:** For plotting EEG signals, spectrograms, and statistical analysis.
- **Plotly:** For interactive charts in the dashboard.
- **Streamlit / Flask:** For developing a user-friendly web dashboard to visualize EEG changes and model explanations.

5. Database / Data Management Tools

- **Pandas:** For structured storage and manipulation of EEG data and metadata.
- **NumPy arrays / CSV files:** For storing intermediate processed datasets.
- **Cloud Storage (Google Drive / AWS / GCP):** For data backup and collaboration.

6. Version Control and Deployment

- **Git / GitHub:** For version tracking of source code and documentation.
- **Docker (optional):** For containerized deployment of the web application.

7. Hardware Compatibility

- Compatible with **CPU and GPU environments**.
- Supports **CUDA (NVIDIA)** for accelerated training if available.

2. LITERATURE SURVEY

2.1 Existing System

In recent years, several research studies have explored the relationship between yogic breathing practices, meditation, and brain activity using *Electroencephalography (EEG)*. These studies have demonstrated that breathing-based relaxation techniques can alter brain wave patterns, improving focus, reducing stress, and promoting relaxation. However, most of these existing systems rely on traditional signal processing and statistical approaches, which have significant limitations in terms of automation, accuracy, and interpretability.

1. Traditional EEG-Based Studies

Earlier research focused primarily on analyzing EEG data using basic frequency-domain methods such as *Fast Fourier Transform (FFT)* or *Power Spectral Density (PSD)* to study the effects of Pranayama or meditation.

- Studies have shown that Alpha and Theta band power increases during slow and rhythmic breathing, indicating relaxation and meditative states.
- Techniques like *Anulom-Vilom* and *Bhramari Pranayama* were found to enhance calmness by modulating the autonomic nervous system.
- However, these methods were largely manual and limited to simple feature extraction, unable to capture complex temporal or spatial relationships in EEG data.

2. Machine Learning-Based EEG Classification

Recent works have applied *machine learning algorithms* such as Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbors (k-NN) to classify EEG data for tasks like stress detection, emotion recognition, and meditation state identification.

- While these models improved accuracy compared to traditional methods, they still depend heavily on handcrafted features extracted by experts.
- They lack generalization when applied across subjects or different breathing techniques.
- Moreover, these models provide little insight into why a particular prediction was made, which limits their scientific interpretability.

3. Deep Learning Approaches

With the rise of deep learning, models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been employed to automatically learn patterns from raw EEG signals.

- CNNs can extract spatial–temporal features from EEG spectrograms, while RNNs capture sequential dependencies.
- These approaches have shown significant improvements in classification accuracy for mental state detection and meditation analysis.
- However, most of these systems act as black-box models, offering predictions without any interpretability or transparency.

3. Gaps in the Existing Systems

Despite advancements, existing EEG analysis systems for Pranayama and meditation suffer from the following limitations:

- Lack of explainability in AI models — users and researchers cannot understand how or why a model arrives at a decision.
- Limited datasets and participant diversity restrict the generalization of results.
- Absence of integrated visualization tools to display changes in brain activity in an intuitive manner.
- Dependence on complex manual preprocessing and feature engineering.

2.2 Proposed System

The proposed system, titled **"Interpretability of Pranayama Effect on Brain Waves Using Explainable Deep Learning,"** aims to scientifically analyze and interpret how different Pranayama techniques influence brain activity using EEG (Electroencephalogram) signals. The system integrates **Deep Learning** and **Explainable Artificial Intelligence (XAI)** methods to classify mental states such as relaxation, focus, and meditative awareness while providing clear explanations of the model's reasoning.

2.3 Literature Review Summary

Year	Citation (Author)	Tools / Software	Technique	Evaluation Parameter
2011	Telles, S., et al.	EEG Device, MATLAB	Alternate Nostril Breathing and Breath Awareness	Change in Alpha and Theta band power
2012	Bhavanani, A. B., et al.	Physiological Sensors, Statistical Analysis Tools	Chandra Nadi Pranayama (Left Nostril Breathing)	Heart Rate, Respiratory Rate, Autonomic Balance
2020	Kaur, M., & Singh, A.	MATLAB, Python (Scikit-learn)	EEG Signal Analysis using SVM and Random Forest	Classification Accuracy, Confusion Matrix
2022	Upadhyay, D., & Dutta, A.	Python, TensorFlow, EEG Dataset	Deep Learning-Based EEG Classification for Meditation States	Accuracy, Precision, F1-Score
2017	Samek, W., Wiegand, T., & Müller, K.R.	TensorFlow / PyTorch	Explainable AI Techniques (Grad-CAM, Layer-wise Relevance Propagation)	Visualization and Model Interpretability
2017	Lundberg, S. M., & Lee, S.-I.	Python (SHAP Library)	SHAP – Shapley Additive Explanations	Feature Importance, Global and Local Interpretability
2016	Ribeiro, M. T., Singh, S., & Guestrin, C.	Python (LIME Library)	LIME – Local Interpretable Model-Agnostic Explanations	Local Explanation, Model Transparency
2016	Bashivan, P., Rish, I., Yeasin, M., & Codella, N.	PyTorch, EEG Dataset	Deep Recurrent-Convolutional Neural Network for EEG	Accuracy, Temporal-Spatial Representation Performance

3. PROBLEM FORMULATION

Pranayama, a fundamental aspect of yogic science, is known to influence physiological and psychological functions through controlled breathing techniques. Various studies have indicated that Pranayama practices can alter **brain wave activity**, enhance focus, and reduce stress. However, the **quantitative and interpretable analysis** of how these breathing techniques impact the brain remains limited. Traditional EEG-based studies rely on manual signal processing methods that are insufficient to capture complex temporal and spectral patterns, while modern deep learning approaches often function as **black-box systems** with little or no interpretability.

To address this gap, the problem can be formulated as follows:

"To develop an explainable deep learning framework capable of classifying and interpreting EEG signals recorded before, during, and after different Pranayama techniques, thereby identifying and explaining the neural patterns associated with relaxation, focus, and meditative states."

Mathematical / Conceptual Formulation

Let the EEG dataset be represented as:

D={(Xi,Yi)}i=1nD = \{(X_i, Y_i)\}_{i=1}^{n}D={(Xi,Yi)}i=1n

where:

- XiX_iXi = EEG signal segments (multi-channel time series or spectrograms)
- YiY_iYi = corresponding mental state labels (Relaxed, Focused, Meditative, etc.)

The objective of the model fθf_\thetafθ parameterized by θ\thetaθ is to learn a mapping:

fθ:Xi→Yif_\theta : X_i \rightarrow Y_ifθ:Xi→Yi

such that classification accuracy is maximized while maintaining interpretability through explainability functions E(fθ,Xi)E(f_\theta, X_i)E(fθ,Xi), where EEE represents the XAI explanation method (e.g., SHAP, LIME, or Grad-CAM).

Thus, the overall problem is an optimization task:

maxθ Accuracy(fθ(X),Y)subject to interpretability constraints defined by E(fθ,X)\max_\theta \, \text{Accuracy}(f_\theta(X), Y) \quad \text{subject to interpretability constraints defined by } E(f_\theta, X)θmaxAccuracy(fθ(X),Y)subject to interpretability constraints defined by E(fθ,X)

Problem Objective

To build a system that:

- Automatically analyzes EEG signals corresponding to different phases of Pranayama.
- Accurately classifies the mental states induced by the practices.
- Provides clear and visual explanations of how the model reaches its conclusions, using XAI techniques.

4. OBJECTIVES

The primary objective of the project **“Interpretability of Pranayama Effect on Brain Waves Using Explainable Deep Learning”** is to scientifically analyze and interpret how different Pranayama techniques influence human brain activity using EEG signals. The project combines deep learning with explainable AI to ensure that model predictions are transparent and interpretable.

Main Objective

To develop an **explainable deep learning framework** that classifies EEG signals recorded before, during, and after Pranayama into distinct mental states (such as relaxation, focus, or meditative state) and provides **interpretable visual explanations** of the underlying neural patterns.

Specific Objectives:

1. **To collect and preprocess EEG data** recorded during different Pranayama techniques (Anulom-Vilom, Bhramari, Kapalbhati, etc.) to ensure clean and reliable signals for analysis.
2. **To extract meaningful features** from EEG signals by analyzing frequency bands (Delta, Theta, Alpha, Beta, Gamma) and generating time–frequency representations such as spectrograms.
3. **To design and train deep learning models** (CNNs and RNNs/LSTMs) capable of classifying EEG data into mental states induced by Pranayama.
4. **To integrate Explainable AI (XAI) techniques** such as Grad-CAM, SHAP, and LIME to interpret and visualize the model’s decision-making process.
5. **To build an interactive dashboard** that displays EEG patterns, classification outputs,

and interpretability visualizations for better understanding by researchers and practitioners.

6. **To evaluate and validate** the model's performance and interpretability using appropriate metrics like accuracy, precision, recall, and feature importance visualization.

Outcome of Objectives.

By achieving these objectives, the project will provide:

- A **scientific and interpretable explanation** of how Pranayama alters brain activity.
- A **transparent AI-based system** that bridges the gap between deep learning accuracy and human understanding.
- A **foundation for future wellness and neurocognitive applications**, integrating yoga, neuroscience, and artificial intelligence.

5. METHODOLOGY

The methodology for the project **“Interpretability of Pranayama Effect on Brain Waves Using Explainable Deep Learning”** outlines the systematic process used to collect, process, analyze, and interpret EEG data to study the effect of Pranayama on brain activity. The approach integrates traditional EEG signal analysis with deep learning and explainable AI (XAI) techniques to ensure both accuracy and interpretability of results.

1. Data Collection

- **Participants:** Healthy adult volunteers aged 18–50 years are selected for EEG recording.
- **Session Design:** Each session consists of three phases —
    1. **Pre-Pranayama (Baseline):** 5 minutes of normal breathing.
    2. **During Pranayama:** 5 minutes of guided breathing (Anulom-Vilom, Bhramari, or Kapalbhati).
    3. **Post-Pranayama:** 5 minutes of rest after the breathing exercise.
- **EEG Device:** EEG signals are recorded using consumer-grade or research-grade headsets (e.g., Muse, Emotiv, or OpenBCI) with a minimum of 8 channels and a sampling rate of 250–500 Hz.
- **Environment:** Recordings are done in a quiet, well-lit, and controlled environment to minimize external noise and movement artifacts.

2. Data Preprocessing

EEG signals are raw and prone to noise; hence, preprocessing is a crucial step.

- **Filtering:**
    - Apply a *band-pass filter* (0.5–45 Hz) to retain brain wave frequencies of interest.
    - Use a *notch filter* (50/60 Hz) to remove electrical interference.
- **Artifact Removal:**
    - Eliminate noise from eye blinks, muscle movement, or head motion using *Independent Component Analysis (ICA)* or automated artifact rejection algorithms.
- **Segmentation:**
    - Divide EEG data into small time windows (e.g., 2–5 seconds).
    - Label each segment according to the Pranayama phase (Before, During, After).
- **Normalization:**
    - Normalize signals to ensure uniform amplitude range across participants and sessions.

3. Feature Extraction

To identify patterns in brain activity, features are extracted in both the **frequency** and **time–frequency** domains.

- **Frequency Domain Features:**
  - Calculate power spectral density (PSD) using *Fast Fourier Transform (FFT)*.
  - Extract band powers for **Delta (0.5–4 Hz)**, **Theta (4–8 Hz)**, **Alpha (8–12 Hz)**, **Beta (12–30 Hz)**, and **Gamma (30–45 Hz)**.
- **Time–Frequency Features:**
  - Generate spectrograms or wavelet transforms to visualize brain activity over time.
  - Use these as input images for CNN-based classification.

4. Model Development

Two deep learning architectures are employed:

- **Convolutional Neural Networks (CNNs):**
  - Used for classifying spectrogram images derived from EEG signals.
  - Automatically learns spatial and temporal patterns associated with mental states.
- **Recurrent Neural Networks (RNNs) / LSTMs:**
  - Applied to sequential EEG data to model temporal dependencies.
- **Training and Evaluation:**
  - Dataset split into *Training (70%)*, *Validation (15%)*, and *Testing (15%)* sets.
  - Performance measured using metrics such as *Accuracy, Precision, Recall, and F1-score.*

5. Explainable AI (XAI) Integration

To ensure transparency and interpretability in model predictions, multiple XAI techniques are applied:

- **Grad-CAM:** Highlights which parts of EEG spectrograms most influenced the CNN's prediction.
- **SHAP (Shapley Additive Explanations):** Quantifies the contribution of each frequency band to the final output.
- **LIME (Local Interpretable Model-Agnostic Explanations):** Provides localized explanations for individual EEG samples.
  These methods make the deep learning model's reasoning process transparent, enabling researchers to identify *why* specific mental states were classified.

6. Visualization and Dashboard Development

A user-friendly interface is developed using **Streamlit or Flask** to present:

- EEG wave plots (before, during, and after Pranayama).
- Predicted mental state (Relaxed, Focused, Meditative).
- XAI visualizations (heatmaps, feature-importance graphs).
  This helps researchers and practitioners easily interpret the AI's findings.

7. Result Analysis

- Compare EEG changes across different Pranayama techniques.
- Evaluate which frequency bands show the most significant variations post-Pranayama (e.g., increased Alpha and Theta power).
- Analyze the interpretability outputs from XAI to validate physiological relevance.
- Perform statistical analysis to confirm consistency of results across participants.

8. Deployment and Documentation

- The trained model and dashboard are integrated into a deployable system for demonstration.
- Documentation includes dataset description, preprocessing pipeline, model parameters, and explainability results.

6.CONCLUSION

The project **"Interpretability of Pranayama Effect on Brain Waves Using Explainable Deep Learning"** successfully integrates the fields of *neuroscience, artificial intelligence,* and *yogic science* to provide a scientific understanding of how Pranayama influences brain activity. Through the use of EEG signal analysis, deep learning models, and explainable AI (XAI) techniques, the system is able to classify and interpret the mental states induced by different Pranayama practices such as *Anulom-Vilom*, *Bhramari*, and *Kapalbhati*.

The results of this approach are twofold — it not only quantifies the measurable changes in brain wave patterns (such as increases in Alpha and Theta activity indicating relaxation and focus) but also explains *why* and *how* the AI model reaches its conclusions. The integration of **Grad-CAM, SHAP,** and **LIME** provides interpretability and transparency, overcoming the limitations of traditional black-box deep learning models.

The developed system serves as a foundation for **data-driven, interpretable wellness research**, offering insights that can benefit both neuroscience and yoga therapy. Furthermore, the creation of an interactive dashboard enhances user understanding by visually demonstrating the neural impact of Pranayama.

In conclusion, the project demonstrates that the combination of *Deep Learning* and *Explainable AI* can be a powerful tool for exploring the connection between ancient yogic techniques and modern brain science. It opens new possibilities for future research in **mental health monitoring, stress reduction, cognitive enhancement,** and **personalized wellness technologies**.

7.REFERENCES

1. Telles, S., Singh, N., & Balkrishna, A. (2011). "EEG Changes during Alternate Nostril Breathing and Breath Awareness." *International Journal of Yoga*, 4(1), 22–27.
2. Bhavanani, A. B., Madanmohan, & Sanjay, Z. (2012). "Immediate Effect of Chandra Nadi Pranayama on Autonomic and Respiratory Variables." *Indian Journal of Physiology and Pharmacology*, 56(1), 80–87.
3. Kaur, M., & Singh, A. (2020). "Analysis of Meditation EEG Signals Using Machine Learning Techniques." *International Journal of Scientific & Technology Research*, 9(2), 1303–1308.
4. Upadhyay, D., & Dutta, A. (2022). "Deep Learning-Based EEG Signal Classification for Meditation States." *Biomedical Signal Processing and Control*, 71, 103158.
5. Samek, W., Wiegand, T., & Müller, K. R. (2017). "Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models." *arXiv preprint arXiv:1708.08296.*
6. Lundberg, S. M., & Lee, S.-I. (2017). "A Unified Approach to Interpreting Model Predictions." *Advances in Neural Information Processing Systems (NeurIPS)*, 30, 4765–4774.
7. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You? Explaining the Predictions of Any Classifier." *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'16)*, 1135–1144.
8. Bashivan, P., Rish, I., Yeasin, M., & Codella, N. (2016). "Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks." *International Conference on Learning Representations (ICLR)*.