



Augmenting Common Spatial Patterns to deep learning networks for improved alcoholism detection using EEG signals

Neeraj et al.

2026.03.26



Computers in Biology and Medicine (2025)

Overview



01 Author

02 Introduction

03 Proposed Method

04 Experiments

05 Conclusion



Neeraj

Assistant Professor, Ecole Central School of engineering, Mahindra University,
Hyderabad

Verified email at iitp.ac.in

Time series analysis time series forecasting time series classification and generation

<p>Detection of alcoholism using EEG signals and a CNN-LSTM-ATTN network</p> <p>Neeraj, V Singhal, J Mathew, RK Behera Computers in Biology and Medicine 138, 104940</p>	45	2021
<p>A Unified Attentive Cycle-Generative Adversarial Framework for Deriving Electrocardiogram from Seismocardiogram Signal</p> <p>Neeraj, S U, M J, B and R. K IEEE Signal Processing Letters, 1-1</p>	16 *	2022
<p>Power load forecasting based on long short term memory-singular spectrum analysis</p> <p>Neeraj, J Mathew, RK Behera Energy Systems, 1-23</p>	12 *	2021
<p>Fusion of wavelet decomposition and N-BEATS for improved stock market forecasting</p> <p>V Singhal, N Neeraj, J Mathew, M Agarwal</p>	10	2022
<p>A deep learning framework for covid outbreak prediction</p> <p>Neeraj, J Mathew, RK Behera, ZE Panthakkalakath arXiv preprint arXiv:2010.00382</p>	7 *	2020
<p>A Deep Learning Architecture for Spatio-Temporal Feature Extraction and Alcoholism Detection</p> <p>Neeraj, V Singhal, J Mathew 2021 IEEE EMBS International Conference on Biomedical and Health Informatics ...</p>	4	2021

Background study



- One of the main risk factors for numerous health problems is **excessive drinking**.
- Alcoholism can affect a person's **thinking and cognitive abilities**.
- The diagnosis of alcoholism is challenging:
 - blood tests
 - questionnaires

This poses the need for **fast, reliable, automatic, and preferably non-invasive** ways to detect alcoholism.

- Electroencephalogram (EEG) is a non-invasive method that can help detect alcohol use disorders in a subject.
- Proposed: a mechanism to **detect alcoholism** using an **EEG signal** with a **deep learning model** that considers the **spatio-temporal** nature of EEG signals by leveraging **common spatial features (CSP)**.

Data Description

- Machine Learning and Intelligent Systems repository at the University of California, Irvine.
- Multichannel EEG recordings from **alcoholic** and **control patients** are included in this collection.
- 64 electrodes with a frequency of 256 Hz, each trial was recorded for 1 s.
- Each set has 600 samples. There are 300 alcoholic samples and 300 control samples in each of the training and test sets.

Common Spatial Patterns (CSP)

- **High-dimensional** data deteriorates the prediction abilities of the classifier used.
- Another factor that reduces the prediction accuracy is that data from nearby channels are **prone to high correlation**.
- CSP is used to deal with this problem of high dimensionality and correlation.

Common Spatial Patterns (CSP)

1. EEG Spatial Covariance R_i for each sample i

Given EEG signal matrix $A_i \in \mathbb{R}^{C \times T}$

$$R_i = \frac{A_i \cdot A_i^T}{\text{trace}(A_i \cdot A_i^T)}$$

2. Composite Spatial Covariance R_c

$$R_p = \frac{\sum_{A_i \in p} R_i}{N_p}$$

$$R_q = \frac{\sum_{A_i \in q} R_i}{N_q}$$

$$R_c = R_p + R_q$$

Common Spatial Patterns (CSP)

3. Eigenvalue Decomposition

$$R_c = X \cdot \lambda_c \cdot X^T$$

Eigenvector matrix $X \in \mathbb{R}^{C \times C}$. Columns are eigenvectors of R_c .

Each eigenvector represents a principal spatial direction of EEG signal (e.g. [0.2, -0.5, ..., 0.3])

Diagonal matrix λ_c containing eigenvalues which represent variance along each eigenvector.

4. Whitening Transformation

Whitening transformer/whitening matrix $Y = \sqrt{\lambda_c^{-1}} \cdot X^T$

$$Z_p = Y \cdot R_p \cdot Y^T$$

$$Z_q = Y \cdot R_q \cdot Y^T$$

$$Z_p + Z_q = I$$

Common Spatial Patterns (CSP)

Why **whitening**?

Whitening transforms data so that the covariance becomes identity: $YR_cY^T = I$

Meaning:

- all channels become **uncorrelated**
- variance becomes **equalized** (= 1)

Common Spatial Patterns (CSP)

5. Eigen Decomposition of Whitened Matrix

$$Z_p = Q \cdot \lambda_p \cdot Q^T$$

These eigenvectors maximize the variance difference between classes.

6. Final Projection Matrix (Spatial Filters)

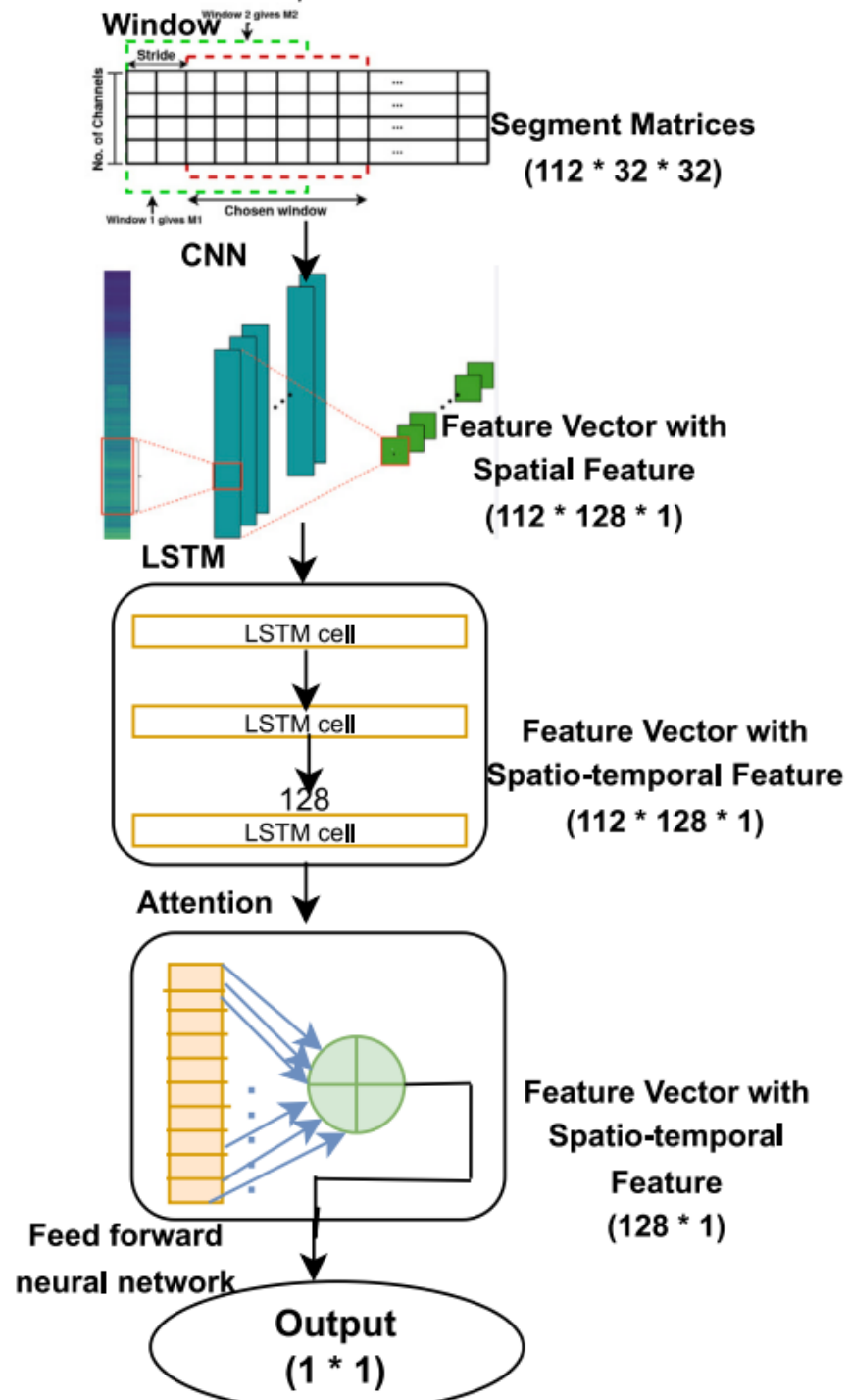
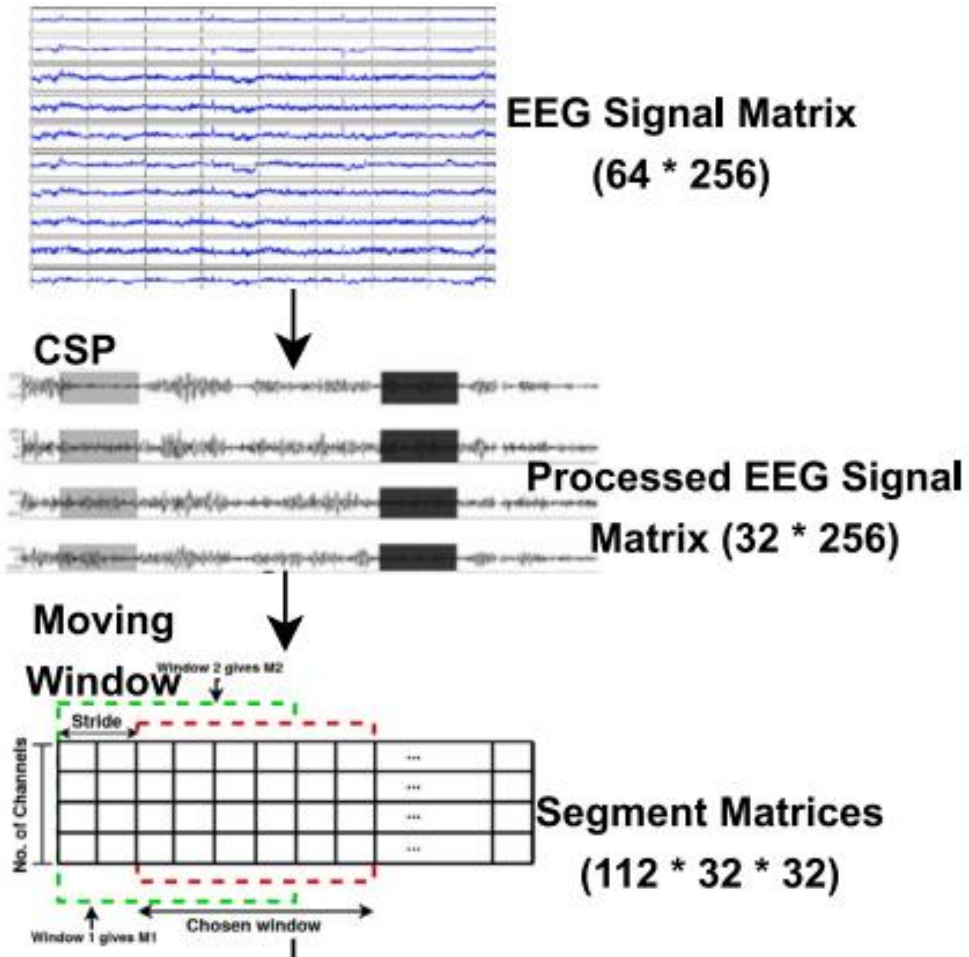
$$W = Q^T \cdot Y$$

This matrix projects EEG signals into **discriminative spatial components**.

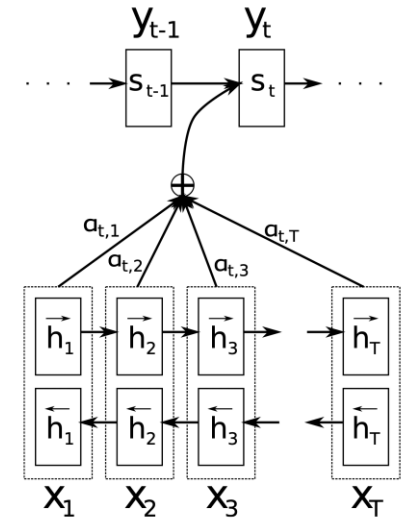
CSP finds spatial patterns that separate alcoholic vs control EEG.

03 Proposed Method

CSP + Deep Learning



Bahdanau Attention



$$c_t = \sum_{i=0}^N a_t \cdot i_t$$

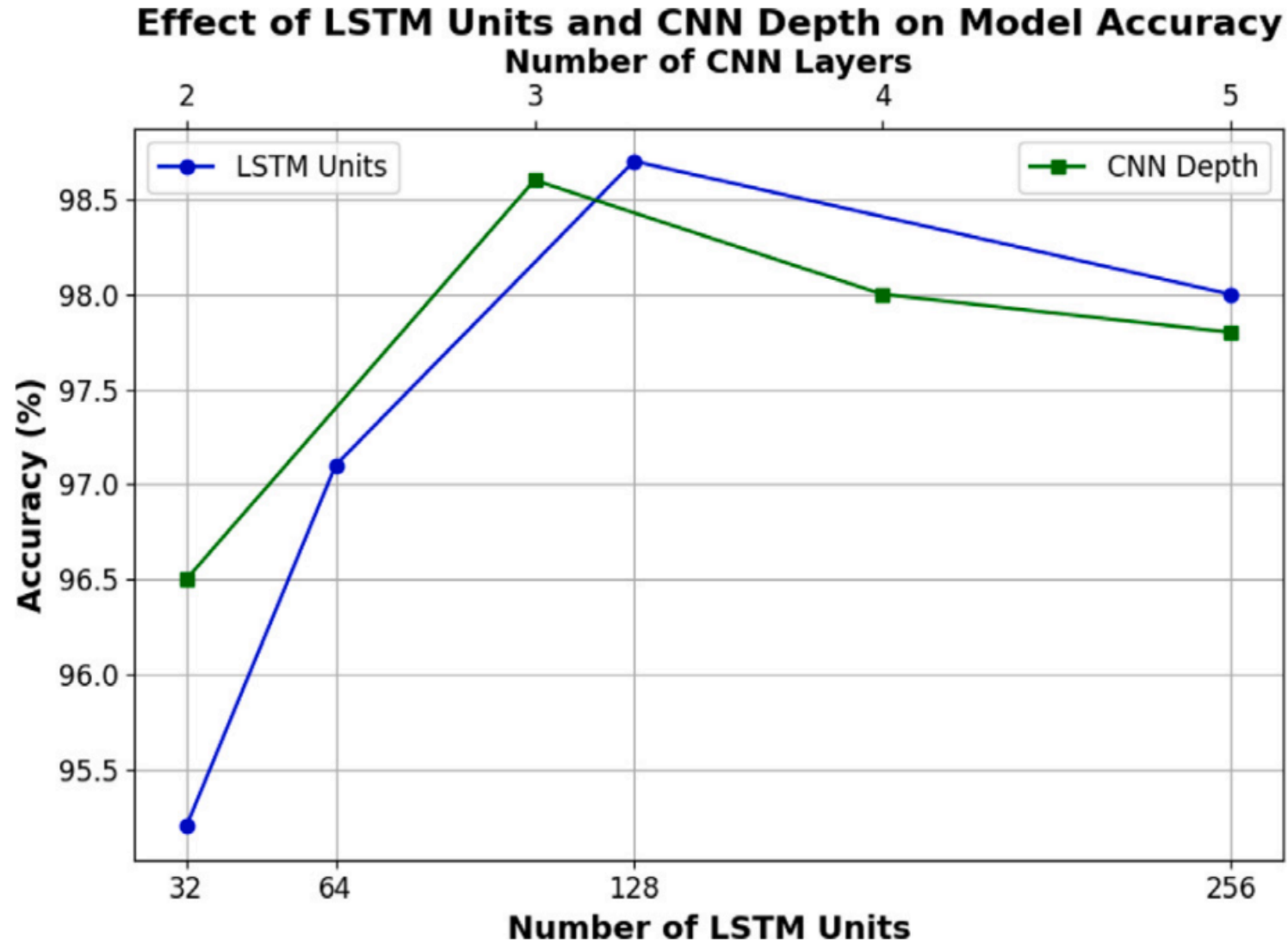
$$\text{softmax}(a_t) = \frac{\exp(q_t)}{\sum_{i=0}^N \exp(q_t)}$$

$$q_t = F_{fcn}(i_t, p_{t-1})$$

Grid search using accuracy metric.

CSP subspace size (C')	4	8	16	32
Window size				
4	81.83	93.50	95.16	98.00
8	84.83	90.00	95.33	97.83
16	84.16	93.00	96.33	98.00
32	83.50	95.16	96.83	98.66
64	85.83	94.66	95.83	97.83

04 Experiment Results



Performance evaluation of the proposed model based on 30 trials.

Metric	Mean	Median	Standard deviation	Min	Max
Accuracy	98.60	98.66	0.24	98.00	99.00
F1-score	98.59	98.65	0.25	97.97	99.00
Recall	98.24	98.33	0.57	96.66	99.33
Precision	98.95	98.99	0.36	98.01	99.32
MCC	97.20	97.33	0.49	96.03	98.00

Confusion matrix.

		Actual class	
		Alcoholic	Control
Predicted class	Alcoholic	297	3
	Control	5	295

04 Experiment Results

Comparison with existing methods.

Model	Classification technique	Features used	Model accuracy (%)
Kumar et al [17]	SVM and fuzzy C-mean clustering	Absolute power, relative power and peak power frequency	88.00
Ehlers et al. [18]	Discriminant analysis classifier	Hurst exponent, correlation dimension, LLE and entropy	88.00
Shri et al. [5]	SVM and BPNN	ApEn ranked using ANOVA test scores	90.00
Fayyaz et al. [8]	LSTM	Features obtained from peak visualization method	90.00
Kannathal et al. [19]	Discriminant analysis classifier	Correlation dimensions	90.00
Acharya et al. [4]	SVM	LLE, SampEn, and some higher order spectra features	91.70
Faust et al. [6]	GMM, KNN, PNN, DT, FSC and NBC	Higher order spectra features	92.40
Farsi et al. [9]	ANN and LSTM	PCA with ANN and temporal features using LSTM.	93.00
Zhu et al. [7]	KNN and SVM	HVGE and SampEn	95.80
Proposed model	CNN-LSTM-ATTN	CSP and moving window	98.66

Performance of various variations of the proposed model(in %).

Model	Accuracy	F1-score	Recall	Precision	MCC	AUC
CSP-LSTM	84.83	84.95	85.66	84.26	69.67	92.83
CSP-LSTM-ATTN	93.66	93.62	93.00	94.25	87.34	98.12
CSP-CNN-LSTM	95.50	95.60	98.00	93.33	91.11	98.84
CSP-CNN-LSTM-ATTN	98.60	98.59	98.24	98.95	97.20	99.64

Common Spatial Patterns + Deep Learning

- The application of CSP not only linearly transforms a data matrix into a lower-dimensional subspace but also maximizes the discriminability of two classes.
- This work shows how deep learning can extract spatio-temporal features from multi-channel EEG signals and use these features for accurate classification.

Future Work

- Examine using sophisticated deep learning architectures, including Transformers and Graph Neural Networks (GNNs), to increase the model's efficacy.
- Try to incorporate multi-modal data (EEG + MRI), using explainable AI techniques to enhance model interpretability, or validating results on larger, more diverse datasets.



Thank You



2026.03.26

BrainLAB Journal Club