A. Brömme, N. Damer, M. Gomez-Barrero, K. Raja, C. Rathgeb,
A. Sequeira, M. Todisco, and A. Uhl (Eds.): BIOSIG 2022,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2022

EEG-based biometrics: phase-locking value from gamma band performs well across heterogeneous datasets

Pradeep Kumar G[1], Utsav Dutta[2], Kanishka Sharma[3], Ramakrishnan A G[4]

Abstract: The performance of functional connectivity metrics is investigated for electroencephalogram (EEG)-based biometrics using a support vector machine classifier. Experiments are conducted on a heterogeneous EEG dataset of 184 subjects formed by pooling three distinct datasets recorded with different systems and protocols. The identification accuracy is found to be higher for higher frequency EEG bands, indicating the enhanced uniqueness of the neural signatures in beta and gamma bands. Using all the 56 EEG channels common to the three databases, the best identification accuracy of 97.4% is obtained using phase locking value-based measures extracted from the gamma frequency band. When the number of channels is reduced to 21 from 56, there is a marginal reduction of 2.4% only in the identification accuracy. Additional experiments are conducted to study the effect of the cognitive state of the subject and mismatched train/test conditions on the system performance.

Keywords: biometrics, EEG, functional connectivity, phase locking value, support vector machine.

1 Introduction

Electroencephalogram (EEG)-based biometric systems have been proposed [Ch18]. Recent work on EEG biometrics have primarily been on finding better discriminative features and identification techniques and the use of deep learning frameworks [Ch18]. Commonly used features are Fourier coefficients, autoregressive (AR) model parameters, wavelet coefficients, Shannon entropy, and spectral entropy [YD17]. This study uses functional connectivity (FC)-based metrics which quantify the relation between neural activity observed in different regions of the scalp [BS16].

Chang et al. [Ch20] combined the directed FC measures with signal complexity measures for person recognition systems with the best accuracy of 90.6% using delta band and SVM. Rocca et al. [LR14] implemented spectral coherence measures with Mahalanobis distance-based classifier. They highlighted the higher discriminatory power of modeling the neural activity as interconnected links in the form of connectivity values than the univariate metrics like PSD. Further, they showed that the brain's frontal regions reflected unique, subject-specific neural activity due to the influence of genetic factors. Fraschini et al. [Fr19] used phase-locking value (PLV) and correlation. They reported that the least equal error rate (EER) for person identification of 5.9% is achieved on Dataset-1 considered in the present study using PLV measure in the gamma band rather than the other

[1] Research Student, MILE Laboratory, Department of Electrical Engineering, IISc, India, pradeepkg@iisc.ac.in
[2] Master Student, Carnegie Mellon University, USA, udutta@andrew.cmu.edu
[3] PostDoc, Centre for Neuroscience, IISc, India, kanishkas@iisc.ac.in
[4] Professor, Department of Electrical Engineering, IISc, India, agr@iisc.ac.in

FC measures. Another study [Ko19] reported a mean accuracy of 99% on Dataset-1 with beta and gamma bands using PLV measure. Although the study by Kong et. al. used three datasets, the performance of the algorithm was not tested on the combined dataset.

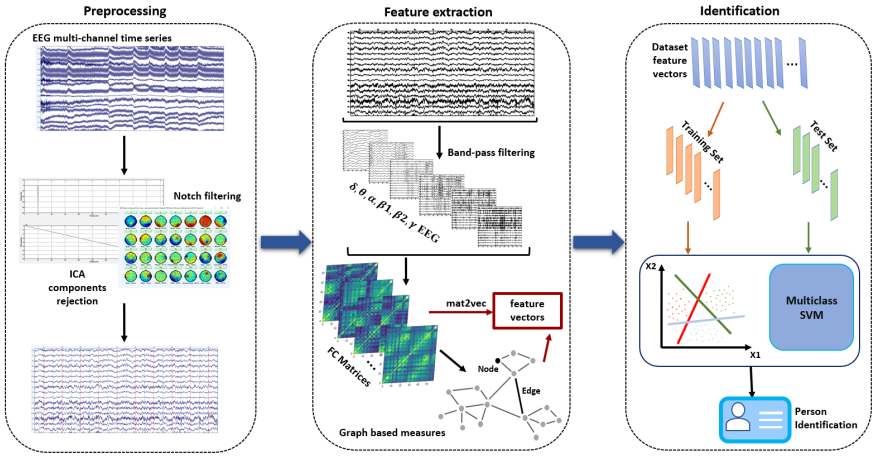

Fig. 1: The proposed EEG-based biometric system. Functional connectivity or graph network-based measures extracted from each frequency band are input to the classifier for training/testing.

1.1 Contribution of the study

We present a biometric framework that explores the potential of FC measures extracted from the individual EEG frequency bands on a heterogeneous dataset from 184 subjects. The data is pooled together from different experimental protocols and recording systems by considering common EEG channels and resampling each dataset to 128 Hz. This increases the generalizability of the technique to data acquired from any EEG system. We use SVM classifier and double k-fold or nested crossvalidation. The study also assesses the performance of two phase-based measures. Fig. 1 presents the proposed methodology.

Tab. 1: Details of the EEG datasets used for the study. In all the cases, only one minute data from the resting state was used. Each database contains 64 channels; however, the combined dataset considers only the 56 channels common to all of them. All the experiments reported use only these 56 channels, or a subset of only 21 channels.

Dataset	EEG data collection setup	# Channels Recorded\Considered	Original sampling frequency (Hz)	# Epochs	# Subjects
Dataset-1	Motor Movement/Imagery	64\56, 21	160	1635	109
Dataset-2	Meditation + Control	64\56, 21	1024	300	14+6
Dataset-3	Meditation	64\56, 21	500	825	55
Combined	-	-\56, 21	-	2760	184

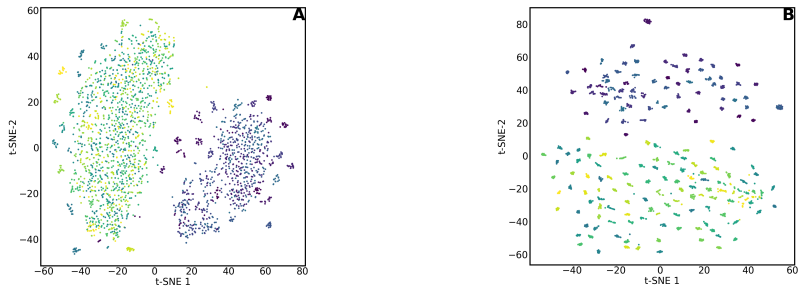

Fig. 2: Lower dimensional visualization using t-SNE: PLV adjacency matrix of combined dataset (184 subjects) **(A)** Alpha band, **(B)** Gamma band. Each subject's samples are identified by a unique shade of colour in the plots. The colours are chosen as distinct hues on a colour palette from the default Python colour set.

2 Materials and Methods

Tab. 1 gives the details of the datasets used for the experiments. Dataset-1 is publicly available. The authors recorded Dataset-2 and Dataset-3, and ethical clearance was obtained from *Institute Ethical Committee with IHEC No: 23-24072019*. Only 1-minute eyes-open baseline data is considered from each dataset. It is split into 4-second long, non-overlapping epochs , resulting in 15 epochs per subject. The Dataset-2 and Dataset-3 [Sh20] will be made public after the possible exploration of the data and other analysis are completed by the group. The preprocessing includes notch filtering to eliminate the line noise at 50 Hz, band-pass filtering from 0.5 to 45 Hz, decomposition by independent component analysis, and manual artifact rejection using EEGLAB [DM04]. To study the effectiveness of different frequency components for the task, the features defined in Sec. 2.2 are extracted for each of the frequency bands delta (0.5 - 4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta1 (12-20 Hz), beta2 (20-30 Hz), and gamma (30-45 Hz).

2.1 Features based on functional connectivity (FC)

Scalar bivariate FC metrics computed between pairs of electrodes over a pre-defined epoch assess the relationship between EEG data from different scalp regions [BS16]. Since the connectivity matrix is symmetric, only its upper triangular elements are concatenated into a row vector of dimension $N_{ch}(N_{ch}-1)/2$, where N_{ch} is the No. of channels used. Pearson's correlation, phase locking value, and phase lag index [BS16] are used.

Discriminability of the different features In order to visualize the ability of the different features defined in Sec. 2.2 to discriminate between the subjects, t-distributed stochastic neighbor embedding (t-SNE) [VdMH08] is used to reduce the feature representation to two dimensions. t-SNE is a non-deterministic method that reduces the dimensionality of any data while retaining the relative 'distance' between points across dimensions. It minimizes the KL-divergence between the probability distributions of the data in the high and low dimensional spaces. The t-SNE plots in Fig. 2 show that the data from different subjects cluster more robustly in the gamma than the alpha band. Since t-SNE uses Gaussian kernels for building probability density estimates, we use the SVM classifier with the radial basis function (RBF) kernel.

2.2 Subject Identification Experiments

Each feature defined in Sec. 2.2 in its vectorized form of connectivity matrix is assessed individually for each of the four datasets. We use multi-class SVMs [MLH03] capable of dealing with high dimensional data with minimal tuning. The SVM employs a 'one-vs-rest' approach where a different binary classification model is trained for each target class. RBF kernel is used. A double k-fold or nested cross-validation approach [St74] is used to assess the algorithm's performance. The dataset is split into k1 = 10 folds, of which (k1-1) folds form the training set, and one fold is held out as the unseen test set. The training set is then divided into k2 = 3 folds, of which one fold is held out as a validation set while the others are combined to train the model. For each of the k1 folds, the best performing hyperparameter set is used to estimate the performance on the unseen test set. For each set of features, the SVM model is optimized using grid search of the hyperparameters C (0.1, 1, 10, 100) and gamma (1, 0.1, 0.01, 0.001).

Experiments using lower number of EEG channels: To study the effect of the number of channels employed on the biometric identification accuracy, additional experiments are conducted using only the subset of 21 EEG channels corresponding to the 10-20 electrode system. If very good performance can be obtained with a few channels, it will facilitate real-life deployment of EEG-based biometrics.

Experiments to study the effect of cognitive state: Different mental activities involve distinct cortical networks, which result in distinct functional connectivity configurations of the EEG. To test whether the biometric recognition performance depends on the cognitive state of the subject, we conducted experiments by training and testing our system using data corresponding to the task, rather than the resting baseline. In our case, the tasks vary across the datasets. Dataset 1 corresponds to motor imagery; datasets 2 and 3 involve EEG from distinct meditative states. We conducted these experiments for both 56 and 21-channel data and compared their performance. Only the PLV feature is used to study the effect of cognitive states.

Experiments with mismatched train/test conditions: The performance of any biometric system degrades when the testing conditions are different from those of enrolment. In our case, all the original experiments were conducted under matching conditions of resting state. However, to study the robustness of the system under mismatched train/test conditions, we also performed experiments where the enrolment uses the resting state EEG, whereas the testing uses task-based EEG and vice versa.

3 Results

3.1 Performance using different features from each of the bands

Performance on each dataset using different FC metrics and 56 channels are the left side figures in Tab. 2. The values are average double k-fold cross-validated identification accuracies with the standard error. The best accuracies obtained are 99.4%, 93%, 97% and 97.4% for the datasets 1, 2, 3 and the combined one, respectively. For all the datasets,

Tab. 2: System performance on resting state EEG data from all the 56\21 channels common to the three databases, for each frequency band. Identification accuracies (in %) on all the datasets using different FC metrics. The best accuracy obtained for each dataset is shown in bold. PLV feature extracted from the gamma band performs the best on all the datasets.

Data	FC metric	Delta	Theta	Alpha	Beta1	Beta2	Gamma
Dataset-1 (109 subjects)	PLI	12\9	21\7	47\14	79\38	95.4 \65	99.1\89
	PLV	76\47	76\46	76\46	93\77	96\87	**99.4\96**
	COR	70\45	69\43	69\44	91.4\75	97\86	97.8\94
Dataset-2 (20 subjects)	PLI	9\10	18\16	28\24	19\15	32\19	46\30
	PLV	89\83	91\89	88\86	**93**\92	92**93**	**93\93**
	COR	88\85	88\87	87\83	91\88	92**93**	**93\93**
Dataset-3 (55 subjects)	PLI	5\4	10\6	5\1	10\8	8\6	18\12
	PLV	73\55	68\58	70\59	89\87	96\95	**97\96**
	COR	73\57	66\61	68\62	87\83	95\94	96\96
Combined (184 subjects)	PLI	8\5	15\6	27\13	53\26	61\43	68\59
	PLV	73\51	71\53	68\54	91\81	96\89	**97.4\95**
	COR	67\50	60\49	60\49	86\76	94 \88	95.7\93

the feature resulting in highest accuracy is the PLV derived from the gamma band. The accuracy is lower when the FC measures are extracted from the complete EEG signal (full bandwidth). Hence, those results are not reported. The optimal FC features are from higher frequency bands beta1, beta2, and gamma, indicating higher discriminatory neural activity at these frequencies. Across all datasets, PLV-gamma displays strong predictive capability across different sampling frequencies, subjects, data acquisition protocols, and EEG systems, thus highlighting its potential for use with different EEG datasets. In the literature, there is only one other study by [Ch19], which deals with a large number of subjects. The results of the above work are compared with those of our study in Tab. 3. Even with a deep neural network, they could achieve a performance of only 96.3%, whereas our method has achieved a recognition rate 1.1% above this value on a database with 27 more subjects.

The effect of epoch length on the identification accuracy is analyzed using the best performing feature of PLV-gamma on the combined dataset. The performance for epoch lengths of 2 to 6 sec in steps of 1 sec is evaluated. Epoch length of 4 seconds resulted in the highest accuracy, and hence the analysis has been carried out with 4-sec epochs.

3.2 Results of additional experiments

The right side entries in Tab. 2 list the performance of the FC metrics PLI, PLV and COR derived from only the classical 21-channel subset of the data. We see that the performance with 21 channels degrades heavily for all the features derived from the lower frequency bands. Gamma band gives the best accuracy of 95% for the PLV feature, which is only 2.4% less than the corresponding figure for the complete 56-channel data. Thus there is promise for EEG biometrics even with lower number of channels.

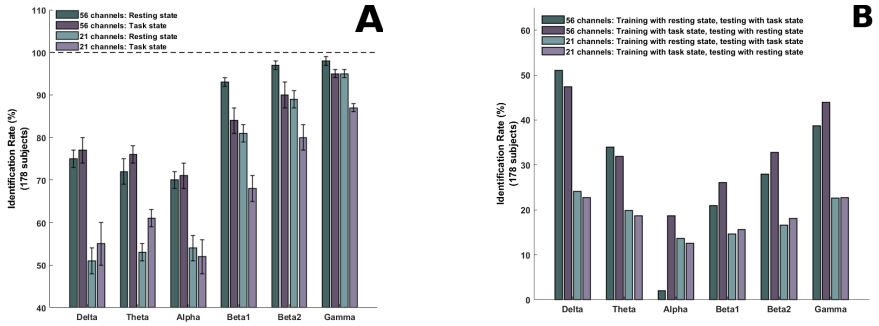

Fig. 3: **(A)** Crossvalidation accuracies for each frequency band from 56 or 21 channels in resting and task state conditions. Since there is no meditation data for the 6 control subjects in dataset2, the results are shown only for the remaining 178 subjects, for the best feature of PLV. The best accuracy, which occurs for the gamma band, degrades only by about 2% when task-based data is used. **(B)** Performance with mismatched train/test conditions for 56 and 21 channels, for each frequency band.The train and test data belong to different states (resting versus task or vice versa). Performance significantly reduces when test conditions differ from the training one.

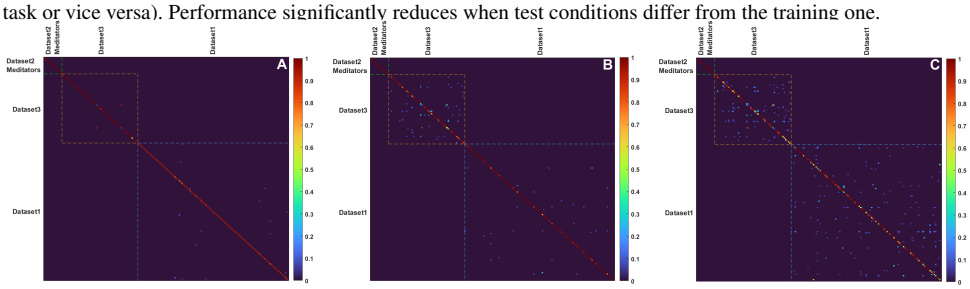

Fig. 4: Confusion matrices (shown as images) using PLV-gamma for 178 subjects **(A)** resting-state, **(B)** task-data with 56 channels; **(C)** task data of 21 channels. The dotted lines with different colors represent subjects from different datasets. Confusions increase if lower number of channels and/or the task-based data are used.

The bar chart in Fig. 3A compares the performance of the task-based data with that using resting data - both for 56 and 21-channels. As there was no meditation data for the six control subjects in dataset2, for fair comparison, the results of all the experiments are shown only for the remaining 178 subjects. For lower frequency bands, the performance with task-based data is superior by a few percentage. With beta, and gamma bands, the resting state data entails better accuracies. Using gamma band, accuracies of 97 and 94% are obtained with resting and task-based data, respectively. Using only 21 channels, the performance is dismally low for low frequency bands, but gradually improves with increasing frequencies. Using 21-channel, resting state, gamma-band data, the performance is lower than that for 56-channels by only 2.5%. However, with the task-based data, the performance comes down by nearly 9% with the use of only the 10-20 channels.

Figure 4 shows the confusion matrices for three distinct cases as color-coded images, where different hues represent different levels of confusion. Figure 4B shows that confusion increases with task-based data. Figure 4C shows that the confusion significantly increases further with task-based data of only 21 channels. The results with mismatched

train/test conditions are shown in Fig. 3B. Compared to the matched conditions, the accuracy reduces significantly when the enrolment and the test conditions differ. For the high frequency bands, there is marginally less reduction in the accuracy when enrolment data is task-based and test data is from resting state. This trend reverses for low frequency bands. Accuracy is less than 25% for all bands with 21-channel data under mismatched conditions.

4 Discussion

Our objective is to assess the utility of EEG for person identification on heterogenous datasets using FC. FC measures perform better than AR and PSD features (see Tab. 3). PLV performs robustly for the combined dataset comprising three different datasets. Accuracy is higher with higher frequency bands due to better clustering of subjects (see Fig. 2), with competitive accuracies obtained from beta and gamma bands. Kavitha et. al [TV18] reported high accuracy using resting state gamma power values of all the channels. Crobe et. al [Cr16] obtained the least equal error rate using the gamma band. Thus gamma band features are emerging as distinctive markers for individual identification.

Tab. 3: Performance comparison with a similar study [Chen et al., 2019] for individual identification on a heterogeneous EEG dataset. Our results [PLV-gamma] on a dataset of higher number of subjects are comparable or better than those of Chen et al., 2019 for any feature and/or classifier.

Study	**# Subj**	**# Channels**	**Frequency band**	**Classifier**	**Feature set**	**Accuracy**
Chen 2019	157	28	0.1-32 Hz	SVM	AR coefficients	86.3%
Chen 2019	157	28	0.1-32 Hz	SVM	PSD	92.9%
Chen 2019	157	28	0.1-32 Hz	GSLT-CNN	Raw EEG data	96.3%
PLV-gamma	184	21	Gamma (30-45 Hz)	SVM	FC measures	**95%**
PLV-gamma	184	56	Gamma (30-45 Hz)	SVM	FC measures	**97.4%**

5 Conclusion

Our results on heterogenous datasets are competitive to recent studies on EEG biometrics, although such studies employ a single dataset recorded under identical conditions. Our dataset1 from Physionet was recorded with a system and for tasks distinct from the ones we used to record the other datasets. The results are only marginally lower when the channels are reduced from 56 to 21. Our results show that PLV-gamma feature is robust for biometrics and works independent of the system or the recording conditions, provided the enrolment and testing data are matched. The use of end-to-end deep neural networks with automated feature extraction from the EEG have resulted in better performance recently [Pa22]. However, [Ch19] achieved only 96.3% accuracy using their GSLT-CNN framework. Thus, the performance of our system is comparable to the best in the literature on large datasets. The code to compute the metrics and test the classifier used in this article are publicly available at `www.github.com/MILE-IISc/Neuroscience/EEG-Biometrics`.

Acknowledgement

The authors thank Schalk and colleagues for making the motor-imagery EEG dataset public. The authors are immensely thankful to Brahma Kumaris Organization for lending their EEG system and encouraging their meditators to participate in our experiments.

References

[BS16] Bastos, A.; Schoffelen, J-M: A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. Frontiers in Sys. Neurosci., 9:175, 2016.

[Ch18] Chan, H-L et al.: Challenges and future perspectives on electroencephalogram-based biometrics in person recognition. Frontiers in Neuroinf., 12:66, 2018.

[Ch19] Chen, JX; Mao, ZJ; Yao, WX; Huang, YF: EEG-based biometric identification with convolutional neural network. Multimedia Tools and Applications, pp. 1–21, 2019.

[Ch20] Chang, W. et al.: An EEG based familiar and unfamiliar person identification and classification system using feature extraction and directed functional brain network. Expert Sys. with Applications, 158:113448, 2020.

[Cr16] Crobe, A. et al.: Minimum spanning tree and k-core decomposition as measure of subject-specific EEG traits. Biomed. Physics & Eng. Exp., 2(1):017001, 2016.

[DM04] Delorme, A.; Makeig, S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including ICA. J. Neurosci. Methods, 134(1):9–21, 2004.

[Fr19] Fraschini, M. et al.: Robustness of functional connectivity metrics for EEG-based personal identification over task-induced intra-class and inter-class variations. Pat. Recog. Let., 125:49–54, 2019.

[Ko19] Kong, W. et al.: EEG fingerprints: phase synchronization of EEG signals as biomarker for subject identification. IEEE Access, 7:121165–121173, 2019.

[LR14] La Rocca, D. et al.: Human brain distinctiveness based on EEG spectral coherence connectivity. IEEE Trans. Biomed. Eng., 61(9):2406–2412, 2014.

[MLH03] Meyer, D.; Leisch, F.; Hornik, K.: The support vector machine under test. Neurocomput., 55(1-2):169–186, 2003.

[Pa22] Panzino, Andrea; Orrù, Giulia; Marcialis, Gian Luca; Roli, Fabio: EEG personal recognition based on qualified majority over signal patches. IET Biomet., 11(1):63–78, 2022.

[Sh20] Sharma, K. et al.: Brain-electric activity during eyes open Brahma Kumaris Rajayoga meditation. MindRxiv, doi:10.31231/osf.io/e48bv, 2020.

[St74] Stone, M.: Cross-validatory choice and assessment of statistical predictions. J. Royal Stat. Soc.: Series B (Methodological), 36(2):111–133, 1974.

[TV18] Thomas, K. P.; Vinod, A. P.: EEG-based biometric authentication using gamma band power during rest state. Circ. Sys. Sig. Process., 37(1):277–289, 2018.

[VdMH08] Van der Maaten, L.; Hinton, G.: Visualizing data using t-SNE. J. Mach. Learn. Res., 9(11), 2008.

[YD17] Yang, S.; Deravi, F.: On the usability of electroencephalographic signals for biometric recognition: A survey. IEEE Trans Human-Machine Systems, 47(6):958–969, 2017.