Received: 31 July 2023 | Revised: 20 December 2023 | Accepted: 29 December 2023 | Published online: 2 January 2024

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2024, Vol. 00(00) 1–8
DOI: 10.47852/bonviewJDSIS42021463

BON VIEW PUBLISHING

# Comparative Evaluation of Facial Paradigms for Devanagari Script-Based P300 Speller: Preliminary Findings

**Vibha Bhandari[1], Narendra D. Londhe[1,*] and Ghanahshyam B. Kshirsagar[2]**

[1]*Department of Electrical Engineering, National Institute of Technology Raipur, India*

[2]*Khoury College of Computer Science, Northeastern University, USA*

**Abstract:** Display paradigm design is an important stage in the design of any brain–computer interface device including the P300 speller system. Conventional grey-to-white intensification limits the usability of the existing Devanagari script-based P300 speller (DSP3S) as it suffers from reduced visual contrast, diminished user engagement, limited personalization, minimal cognitive stimulation, and the risk of visual fatigue. The use of a facial paradigm can address the above-mentioned limitations through emotionally resonant stimuli, offering diverse cognitive stimulation, and potentially reducing visual fatigue by introducing dynamic and engaging facial stimuli. In this context, this study is the first attempt to investigate the efficacy of facial paradigms in further improving the usability of DSP3S. The authors investigated three facial paradigms based on the smiley face, famous face, and family faces (FFs) for DSP3S. The three facial paradigms are evaluated based on their effectiveness, efficiency, and satisfaction. Effectiveness was measured using classification performance and amplitude difference waveform, whereas efficiency and satisfaction were measured using NASA-TLX, visual analogue scale-fatigue scale, and system usability scale, respectively. The obtained results showed that using the proposed FF intensification, efficiency, and satisfaction could be improved. However, no significant difference was observed in effectiveness among the three display paradigms. In this study, apart from investigating the usability of three display paradigms, authors also propose brain signal decoding using morphological filtering and extreme gradient boosting. Finally, the study concludes that the design of display paradigms using different intensification patterns affects the usability of the speller system and should be taken into consideration.

**Keywords:** P300 speller, Devanagari script, family face, smiley face, famous face

## 1. Introduction

Brain–computer interface (BCI) establishes a connection between the human brain and its surroundings by converting cognitive activity into computer commands (Wolpaw et al., 2002). In this way, humans can communicate their ideas solely through their brains without really moving (Birbaumer, 2006). Patients with locked-in conditions, such as those with amyotrophic lateral sclerosis, can benefit from BCI. The majority of BCI devices capture brain activity using electroencephalography (EEG). EEG is a quick and reasonably affordable neuroimaging technique that measures the electrophysiological activity of the neurons. Being non-invasive is a huge benefit of EEG, which is crucial for BCI devices.

Speller systems, a prevalent application of BCIs based on EEG, are specifically designed to facilitate communication for individuals suffering with severe motor neuron disorders. These EEG-based BCI spellers leverage event-related potentials (ERPs), suggesting that the cognitive response to a stimulus can be categorized by analyzing voltage deflections in recorded EEG. Among ERP types, visual-evoked ERP is commonly employed in speller systems. The polarity and latency of the voltage deflection serve as indicators for visual evoked ERP components (Sellers & Donchin, 2006). In the context of an oddball paradigm, which involves presenting the subject with repetitive stimuli, ERP components come into play. When the oddball paradigm is used, low-probability target stimuli are combined with high-probability non-target visual stimuli (Jin et al., 2014). The intensification of the desired character serves as the target stimulus in speller systems.

In the traditional P300 speller, the positive voltage peak known as the P300 component, initiating 300 ms after the target stimulus, is extracted to identify the desired symbol through the utilization of the oddball paradigm. The first P300 speller was designed for the English language (Farwell & Donchin, 1988), soon after that researcher identified the need of communication in native languages. Therefore, P300 spellers for different languages including Chinese, Arabic, Japanese, and Devanagari script (DS) were designed by different research groups (Kshirsagar & Londhe, 2019).

A typical P300 speller is accomplished in different stages like design of external stimuli, data recording, feature extraction, ERP classification, and character detection (Kshirsagar & Londhe, 2019). Among different stages of the P300 speller, stimulus design plays a very important role. Conventional grey-to-white intensification in display paradigms is associated with drawbacks

*Corresponding author: Narendra D. Londhe, Department of Electrical Engineering, National Institute of Technology Raipur, India. Email: nlondhe.ele@nitrr.ac.in

such as diminished visual contrast, reduced user engagement, minimal cognitive stimulation, and the possibility of inducing visual fatigue. To overcome these limitations, researchers have explored alternative display paradigms that incorporate more diverse and engaging visual stimuli, such as familiar faces, dynamic images, or personalized content.

In the context of English language-based P300 spellers, researchers have employed character enhancements based on self-face (Lu et al., 2019), green face (Li et al., 2015), and famous face (FsF) (Kaufmann et al., 2011). Studies on face recognition indicate variations in the amplitude of ERPs between a FsF and a familiar face, such as that of a mother, which holds greater significance for the individual (Lu et al., 2019). Caharel et al. (2005) found that the face of the subject's mother elicits a more pronounced positive response between 300 and 700 ms compared to a FsF. Similar observations have been reported in other face recognition experiments. Sui et al. (2006) noted a higher positivity in ERP amplitude for the face of a classmate or flatmate compared to a celebrity face between 220 and 700 ms. Meijer et al. (2009) and Gunji et al. (2013) demonstrated that familiar faces generate a higher P300 amplitude compared to unfamiliar faces. Based on the aforementioned research on face recognition, the brain processes familiar faces differently, resulting in a waveform with a more substantial positive amplitude. Therefore, in this study, for the first time, we propose a family face (FF) paradigm for DS-based P300 speller (DSP3S). The proposed FF paradigm uses the faces of subjects' family members (mother, father, sibling, cousin, etc.) as familiar faces. In contrast to the previous approach for the English language (Zhaohua et al., 2019) using facial paradigm (famous and familiar) that uses the same face to cover all the characters of flashing rows and columns, we use different faces to cover each character. Therefore, the characters adjacent to the target characters are flashed with different images to avoid possible adjacency distraction. Additionally, we compare the proposed FF paradigm with the FsF (using famous Indian faces) and the smiley face (SF) paradigm. Usability evaluation of the three-display paradigm is presented in this study to select the best paradigm among the three for DSP3S.

The study assesses the three facial paradigms by gauging their effectiveness, efficiency, and user satisfaction. Effectiveness is quantified through classification performance and amplitude difference (AD) waveform analysis, while efficiency and satisfaction are evaluated using NASA-TLX, visual analogue scale-fatigue (VAS-F) scale, and system usability scale (SUS), respectively. The primary aim of the present work is to compare the three facial paradigms. However, the authors also propose to use the morphological feature obtained after morphological filtering for the ERP classification process.

Novel contributions of the proposed work are:

1. Introducing an innovative FF paradigm designed exclusively for DSP3S, providing a novel approach to visual stimuli for enhanced usability.
2. Conducting the inaugural comparative examination of three distinct facial paradigms within the DSP3S framework, offering valuable insights into their respective effectiveness for user interaction and communication.
3. The utilization of mathematical morphology-based ERP classification to quantitatively assess the effectiveness of display paradigms.

The rest of this paper is organized as follows: Section 2 presents the data recording protocol, dataset detail, and proposed methodology for P300 detection and usability evaluation. Experimental results are presented in Section 3, and the work is finally concluded in Section 4.

## 2. Material and Method

### 2.1. Participants

The present study was approved by the Ethical Clearance Committee at NIT Raipur. Four healthy subjects aged 25–35 years, with normal cognitive and normal or corrected-to-normal vision, volunteered for our experiment. Participants were naïve to the BCI, and written informed consent was obtained before their participation.

### 2.2. The spelling paradigm

In this study, we introduce the FF paradigm and conduct a comprehensive comparison with other two facial paradigms namely the SF and FsF paradigms. Each of the three paradigm designs uses a matrix containing 8 rows and 8 columns. The matrix comprises 64 characters, out of which there are 48 Devanagari alphabets, 10 numbers, and 6 symbols. Importantly, it should be noted that the display paradigms were designed using the UMA BCI Speller platform with BCI 2000 in the background.

During the spelling process, rows and columns are flashed randomly. In the FF paradigm, characters in the flashing rows or columns are covered by images depicting the faces of the subject's family members. Notably, to mitigate potential adjacency distraction, each character in the flashing row or column is obscured by a distinct photo of a family member. The design principles of the FsF and SF paradigms align with that of FF. However, in FsF paradigm the characters of the flashing row or column are covered by faces of famous Indian personalities and in the SF paradigm covered by yellow smiley emoji. For a visual representation of these paradigms, refer to Figures 1, 2 and 3.

**Figure 1**
**Display paradigm design with smiley face intensification**


**Figure 2**
**Display paradigm design with famous face intensification**

**Figure 3**
**Display paradigm design with family face intensification**


## 2.3. Experimental setup

The data recording was organized inside a soundproof chamber. Each subject had to spell 10 DS words (comprising a total 41 characters) using the three paradigms. During the spelling process, each row and column was flashed once (total of 16 flashing) in a single trial. There were 15 such trials for each character in a word. Participants were instructed to silently count the number of targeted letter stimulations and then shift their attention to the next letter when the first set was over. During the spelling process, stimulus was provided by covering the characters with face pictures for 120 ms with an 80 ms inter-stimulus difference resulting in total stimulus onset asynchrony (SOA) of 200 ms. One session consists of spelling all 10 words for one paradigm. Each session ended with participants responding to two separate surveys, including the NASA-TLX (Hart & Staveland, 1988) test and VAS-F questionnaires (Kim et al., 2010). A comparative questionnaire for the three paradigms was completed at the end of the three sessions. Sufficient time interval was provided between sessions and between the spelling of different words in a session.

## 2.4. Data acquisition and pre-processing

The EEG data were recorded using 16-channel ACTi CAP Xpress V-amp EEG recorder with BCI 2000 software. The 16 channels were mounted on the scalp following the international 10–20 system. The impedance of each electrode was kept below 10 kΩ for all the experiments. Left and right mastoids were used as ground and reference, respectively. The sampling frequency was kept 500 Hz. The recorded data were filtered using bandpass filtered with 0.1–12.5 Hz pass band. For ERP extraction, the EEG data 0–600 ms post stimulus were extracted, which give 300 timepoints each ERP. Later all the ERPs are down-sampled to 50 Hz to lessen the computational burden and fast operation. The number of ERPs for 15 and single trial are presented in Table 1.

**Table 1**
**ERP details of 41 Devanagari characters**

| Trial(s) | P300 | Non-P300 | Total |
|---|---|---|---|
| 15 | 78720 | 551040 | 629760 |
| 1 | 5284 | 36736 | 41984 |

## 2.5. Usability evaluation

The usability of the three facial paradigms was evaluated using the methodology outlined by ISO (2000), which included the following three metrics: effectiveness, efficiency, and satisfaction (Medina-Juliá et al., 2020). Effectiveness is the precision with which a user can execute tasks. Different outcomes used to assess the effectiveness in this study are (i) ERP classification performance and (ii) AD between P300 and non-P300 ERP. Efficiency corresponds to the resources used, such as user time and effort, to execute an activity. Two measures were adopted to assess the efficiency: (i) the subjective workload evaluated with NASA-TLX (Hart & Staveland, 1988) which examined the participant's performance, effort, and feelings of frustration as well as the mental, physical, and temporal demands and (ii) the level of weariness felt during the test was measured using the VAS-F scale (0–10) (Kim et al., 2010). The attitude of the users, or their perception of comfort and acceptance with the system, is related to their level of satisfaction. A questionnaire derived from the SUS (Brooke, 1996) and containing six dimensions favorite, complex, comfortable, stressful, controllable, and tiring to assess satisfaction was adopted in this study. The three different intensification patterns were rated among them. For each pattern, three ranks were suggested: rank 1, the least; rank 2, the middle; and rank 3, the most preferred.

## 2.6. ERP classification

*2.6.1. Mathematical morphology*

In this subsection, the authors present a brief explanation of the mathematical morphology-based ERP classification adopted in this study. Morphological operation may be defined as the process of obtaining local information by iteratively transforming the original signal with the morphological operator (MO) and a predefined structural element (SE). This technique may successfully remove signal noise and extract meaningful information. The way the MO and SE are configured has a significant impact on the way morphological filtering works, according to the fundamental theory of MF. Kazakeviciute et al. (2010) identified the four fundamental operators as dilation, erosion, closing, and opening. Let $e(n)$ be the one-dimensional ERP data, and $s(n)$ be the predefined structuring element.

The dilation and erosion operation on $e(n)$ using $s(n)$ are defined as

$$e \oplus s(n) = \max[e(n-m) + s(m) \tag{1}$$

$$e \ominus s(n) = \min[e(n+m) - s(m) \tag{2}$$

where $\oplus$ indicates dilation and $\ominus$ indicates erosion operation. Additionally, the definition of the opening and closing operations is given below (Kazakeviciute et al., 2010):

$$(e \circ s)(n) = ((e \ominus s) \oplus s)(n) \tag{3}$$

$$(e \cdot s)(n) = ((e \oplus s) \ominus s)(n) \tag{4}$$

where $\circ$ indicates opening operation and $\cdot$ indicate closing operation. The convex peak of the signal is smoothed during the opening operation, and the concave peak during the closing operation. To identify the signal's peaks and valleys, opening and closing operations might be used.

Both positive and negative peaks exist in the ERP signal. To detect the bidirectional peak, combination of opening and closing operation can be used. Moreover, the open–close and close–open filters can be expressed as

$$FOC(e(n)) = ((e \circ s) \cdot s)(n) \tag{5}$$

$$FCO(e(n)) = ((e \cdot s) \circ s)(n) \tag{6}$$

In this study, we employ the combination morphological filter hat (CMFH) transform as defined in Equation (7) (Guan et al., 2021). This combination filter is chosen as it better detects the bidirectional peak.

$$CMFH(e(n)) = e(n) - \frac{FOC(e(n)) + FCO(e(n))}{2} \tag{7}$$

After determining a MO in the morphological filter, the next step is to choose an acceptable structuring element. Structure elements come in a variety of shapes, including flat and double tap. Many researchers have demonstrated that there is no discernible difference in feature extraction between flat SE and various types of SE. Furthermore, complex SE forms will increase computational load while decreasing computational efficiency. As a result, the double tap SE is used in the research of this paper. A careful choice of scale of SE is crucial in morphological filtering. If the SE scale is too large, the significant peaks might be eliminated. But when the scale is small, substantial noise cannot be eliminated. The double tap structuring element adopted for this study is presented in Equation (8).

$$SE = \{1\ 0\ 1\} \tag{8}$$

Such structuring elements in the morphological filter reduce background activity and retrieve peak components from the original signal. As an example, the application of CMFH is presented in Figures 4, 5 and 6, where the original ERP with different peaks is shown in Figure 4, the down-sampled ERP is shown in Figure 5, and the result obtained after morphological filtering is shown in Figure 6. As a result of CMFH, the peaks are exposed with higher amplitude and other signal values are reduced greatly in comparison.

**Figure 4**
**Extracted ERP**

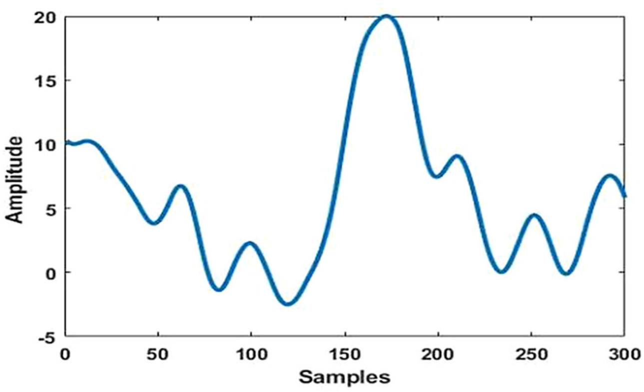



**Figure 5**
**ERP after down sampling**

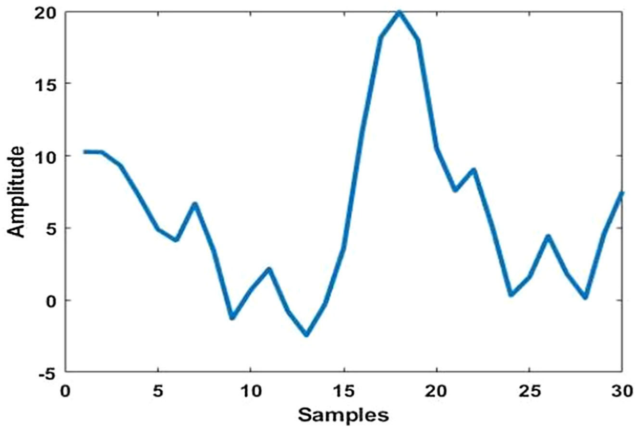
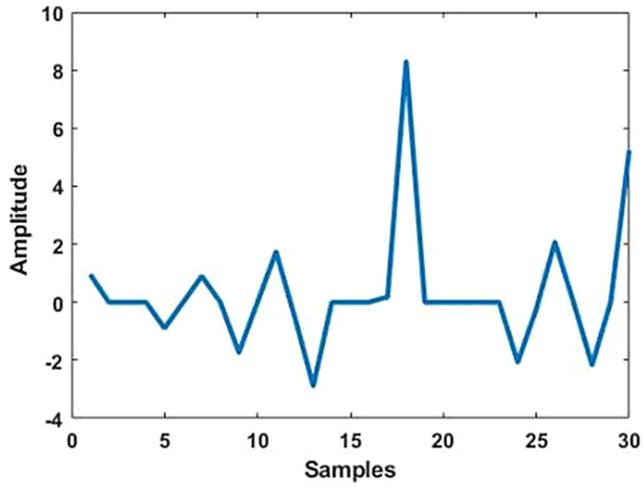



**Figure 6**
**Morphological features obtained after filtering**




Since the dataset for P300 speller consists of target (P300) and non-target (non-P300) ERP, the classification stage aims to distinguish the two. The morphological features obtained after filtering were then fed to extreme gradient boosting (XGBoost) algorithm to perform the binary classification as P300 or non-P300.

### *2.6.2. Extreme gradient boosting*

XGBoost is a machine learning technique that combines the predictions of several weaker models to correctly predict a target variable using a supervised learning strategy. It is a well-known classification tool with fast performance. By adding a new tree in each step to accompany the previously built trees, the additive tree method was used to create the XGBoost model. The accuracy typically increases as more trees are built. In the presented work, XGBoost was used after applying the morphological filtering to the ERPs. XGBoost is used to classify the ERPs to P300 and non-P300. The parameters of XGBoost classifier are specified in Table 2. The classification is evaluated using different metrics as shown in Table 3.

**Table 2**
**Parameters of XGBoost classifier**

| Parameter | Value |
|---|---|
| Learning rate | 0.1 |
| Number of estimators | 250 |
| Maximum tree depth | 10 |
| Minimum child weight | 1 |
| Gamma | 0.2 |
| Subsample | 0.9 |
| Col sample by tree | 0.9 |
| Objective | Binary logistic |
| Number of threads | 6 |

**Table 3**
**Evaluation metrics**

| Measure | Formula |
|---|---|
| Accuracy (ACC) | $\frac{TN+TP}{TN+TP+FN+FP}$ |
| Precision or positive predictive value (PPV) | $\frac{TP}{TP+FP}$ |
| Recall or true positive rate (TPR) | $\frac{TP}{TP+FN}$ |
| F1 score | $2 \times \left[\frac{TPR \times PPV}{PPV+TPR}\right]$ |

*Note: TP – true positive, FP – false positive, TN – true negative, FN – false negative*

## 3. Result and Discussion

### 3.1. Effectiveness

Figure 7 presents the single trial offline classification performance of the three paradigms on the test dataset in terms of accuracy, true positive rate, true negative rate, and F1 score. The classification was performed using the leave-one-subject-out approach and the presented result is the average of the 4-fold cross-validation performance. Adopting the FFs paradigm improves the classification accuracy by 0.8% and 2.52% compared to SF and FsF; however, the difference is not statistically significant. The better performance in the case of the FF paradigm may be attributed to the larger P300 and N170 components in the recorded EEG. Nonetheless, there is a minor overfitting concern observed in the results, where there is a slight variance between the outcomes on the training and test datasets. It is worth noting that the disparity is not statistically significant for any of the display paradigms ($p$-value = 0.01, $t$-test). The AD waveform in contrast to target or non-target ERP waveforms is more interesting to analyze because it demonstrates how the amplitudes of the ERP target and non-target signals differ, making it simpler for the classifier to distinguish between the two signals. The grand averaged AD waveform for the Pz channel is shown in Figure 8. It can be observed from Figure 8 that FF elicits slightly higher P300 compared to other two. The AD signal and classification accuracy do not, on average, differ significantly between the three facial paradigms.

**Figure 7**
**Classification performance of the three facial paradigms on the test data**

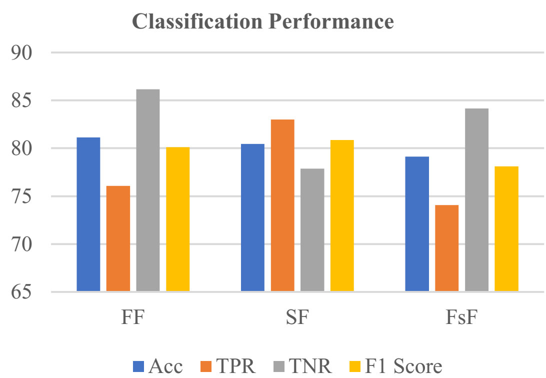
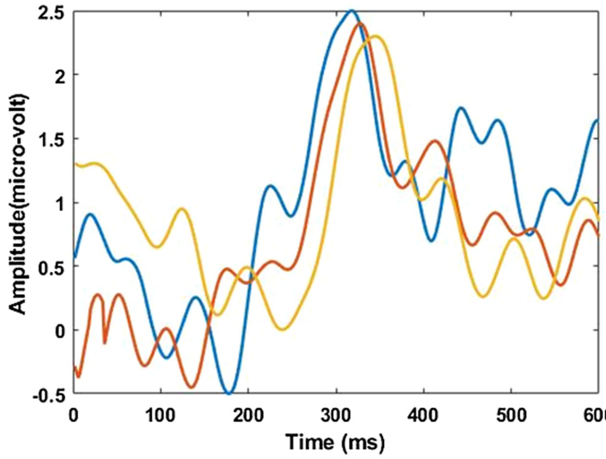



**Figure 8**
**Grand averaged amplitude difference waveform for channel Pz yellow (SF), red (FsF), and blue (FF)**




### 3.2. Efficiency

The effects of facial paradigms on VAS-F, overall workload (NASA-TLX), and subjective workload dimensions (mental demand, physical demand, temporal demand, effort, performance, and frustration) are reported in Table 4. The results are in accordance with the average score of all the participants. Clearly, SF and FsF incur 15.16% ($p$-value = 0.03, $t$-test) and 19.6% ($p$-value = 0.01, $t$-test) more total workload compared to FF paradigm. Also, the fatigue (VAS-F) for SF and FsF is 56.86% ($p$-value = 0.01, $t$-test) and 52.06% ($p$-value = 0.04, $t$-test) higher compared to FF. However, no significant difference was found between SF and FsF.

**Table 4**
**Efficiency grade based on NASA TLX score and VAS-F scale**

| Parameter | SF | FsF | FF |
|---|---|---|---|
| Mental demand (0–12) | 9.75±0.5 | 9.25±0.95 | 7.75±0.95 |
| Physical demand (0–12) | 7.00±1.41 | 7.75±0.5 | 5.5±1.00 |
| Temporal demand (0–12) | 5.75±2.21 | 8.5±1.73 | 6.25±0.95 |
| Performance (0–12) | 9.25±0.95 | 10.25±0.5 | 11.0±0.00 |
| Effort (0–12) | 9.00±1.41 | 10.75±0.5 | 8.0±0.16 |
| Frustration (0–12) | 8.00±0.81 | 8.5±1.2 | 5.75±0.95 |
| Total workload (0–100) | 70.25±6.34 | 73.00±6.97 | 61.0±4.69 |
| VAS-F (0–10) | 7.2±0.72 | 6.98±1.17 | 4.59±1.85 |

**Table 5**
**Comparative analysis between existing P300 speller paradigm and proposed facial paradigm**

| S. no. | Paradigm | Ref | Subjects | Trial | Classifier | % Acc |
|---|---|---|---|---|---|---|
| 1 | RC | (Rakotomamonjy & Guigue, 2008) | 2 | 5 | SWLDA | 74.5 |
| 2 | RC | (Cecotti & Graser, 2011) | 24 | 5 | CNN | 59.44 |
| 3 | RC | (Minett et al., 2012) | 9 | 5 | WE-SVM | 79.2 |
| 4 | RC | (Chaurasiya et al., 2016) | 2 | 5 | ESVM | 73.5 |
| 5 | RC | (Liu et al., 2018) | 2 | 5 | CNN | 69 |
| 6 | RC | (Xiao et al., 2019) | 12 | 1 | DCPM | 82.17 AUC |
| 7 | RC | (Kshirsagar & Londhe, 2019) | 10 | 3 | DCNN | 88.22 |
| 8 | RC | (Kshirsagar & Londhe, 2019) | 10 | 3 | SAE | 84.85 |
| 9 | RC | (Kshirsagar & Londhe, 2022) | 10 | 1 | DS-P3SNet | 85.85 |
| 10 | Familiar face | (Li et al., 2015) | 12 | 2 | BLDA | 75.6 |
| 11 | Green familiar face | (Li et al., 2015) | 12 | 2 | BLDA | 86.1 |
| 12 | Famous face | (Lu et al., 2020) | 10 | 2 | BLDA | 80 |
| 13 | Self-face | (Lu et al., 2020) | 10 | 2 | BLDA | 85.3 |
| 14 | **SF** | **This work** | **4** | **1** | **MF-XGBoost** | **79.12** |
| 15 | **FsF** | **This work** | **4** | **1** | **MF-XGBoost** | **80.45** |
| 16 | **FF** | **This work** | **4** | **1** | **MF-XGBoost** | **81.12** |

## 3.3. Satisfaction

Each variable in the satisfaction construct was considered as either positive (*controllable*, *comfortable*, and *favorite*) or negative (*complex*, *stressful*, and *tiring*). It can be observed from Figure 9 that SF and FsF paradigm presented more negative than positive ranks as they were considered more tiring and less comfortable. In contrast, most dimensions for FF paradigm presented positive values. So, in terms of satisfaction, users found the FF paradigm to be significantly ($p < 0.01$, *t*-test) adequate.

**Figure 9**
**Satisfaction index for each of three paradigms. Ranks 1, 2, and 3 were transformed into positive or negative values, *controllable*, *comfortable*, and *favorite* dimensions received positive ratings. *Complex*, *stressful*, and *tiring* aspects were given negative rating**


## 3.4. Comparative analysis

We present a comparative analysis of the proposed approach with recent studies incorporating a reduced number of trials from the literature in Table 5. First, we compare the proposed approach with existing studies on DS-based P300 spellers. Subsequently, we conducted a comparison of P300 classification accuracy on different languages, considering various approaches that involve a reduced number of trials or utilize facial paradigms.

(i) The existing DS-P3S demonstrates a P300 detection accuracy ranging from 92 to 95% with 15 trials (Kshirsagar & Londhe, 2019; Kshirsagar & Londhe, 2022). However, this accuracy diminishes when the number of trials is reduced. Specifically, for 3–5 trials, the P300 detection accuracy falls within the range of 73–88% (Chaurasiya et al., 2016; Kshirsagar & Londhe, 2019). In comparison to the above-mentioned studies, the proposed approach reports a comparable performance of 81.12% accuracy using a simple classifier on a single trial. Although Kshirsagar and Londhe (2022) report a better accuracy of 85% for a single trial, our suggested method is comparatively less complex because it does not require training several models.

(ii) P300 spellers for the English language(Cecotti & Graser, 2011; Liu et al., 2018; Minett et al., 2012; Rakotomamonjy & Guigue, 2008) have accuracy in the range of 69–79% with 5 trials. A recent study (Xiao et al., 2019) achieved an area under the receiver operating characteristic curve score of 82.17 for single-trial P300 detection. The proposed approach achieves a comparable classification accuracy of 81.12% with only a single trial. Moreover, in contrast to other facial paradigm-based P300 spellers (Li et al., 2015; Lu et al., 2020) which typically achieve P300 classification accuracy ranging from 75.6 to 86.1% with two trials, our proposed approach stands out by achieving an 81.12% classification accuracy with just a single trial. This signifies a notable improvement, as it reduces the time required for spelling by half compared to existing facial paradigms.

Based on the aforementioned comparison, it is evident that the proposed approach demonstrates performance on par with existing methods, especially with a straightforward classifier design, a diminished number of trials, and training involving fewer subjects. There is potential for further enhancement in the performance of the proposed approach through training on a larger dataset, thereby encompassing a greater number of subjects.

## 4. Conclusion

This work is the first study to compare the facial paradigms for DSP3S. In this study, it has been demonstrated that when designing the P300 speller, the intensification pattern is important and should be taken into consideration. Summarizing the present work, the FF paradigm is most preferred in terms of satisfaction. The FsF paradigm is associated with the highest workload according to the NASA-TLX scores; it suggests that participants perceive this task as requiring more mental and cognitive effort. On the other hand, the FF paradigm is associated with the least workload; hence, it may be concluded that participants find it easier and less mentally demanding, possibly due to the familiarity and personal nature of family members' faces. However, considering the effectiveness dimension, no significant difference is observed across three paradigms concerning both classification accuracy and AD waveform. This suggests that, in terms of effectiveness, the cognitive and neural processes involved in processing these faces might be more similar than different. It is likely that with a larger sample size in the current study, the FF paradigm could have demonstrated statistical significance across all usability dimensions as the most appropriate intensification pattern. Finally, it will be interesting in future studies to investigate the proposed approach with a larger database. Also, in future work, we would compare morphological filtering-based classification with other state-of-the-art classifications.

## Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

## Conflicts of Interest

Narendra D. Londhe is an editorial board member for *Journal of Data Science and Intelligent Systems* and was not involved in the editorial review or the decision to publish this article. The authors declare that they have no conflicts of interest to this work.

## Data Availability Statement

The data that support the findings of this study are not publicly available due to privacy concerns. However, anonymous data are available on reasonable request. Requests should be made to Narendra D. Londhe (nlondhe.ele@nitrr.ac.in) and should include a brief description of the intended use of the data.

## References


Birbaumer, N. (2006). Breaking the silence: Brain–computer interfaces (BCI) for communication and motor control. *Psychophysiology*, *43*(6), 517–532. https://doi.org/10.1111/j.1469-8986.2006.00456.x

Brooke, J. (1996). SUS—A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

Caharel, S., Courtay, N., Bernard, C., Lalonde, R., & Rebaï, M. (2005). Familiarity and emotional expression influence an early stage of face processing: An electrophysiological study. *Brain and Cognition*, *59*(1), 96–100. https://doi.org/10.1016/j.bandc.2005.05.005

Cecotti, H., & Graser, A. (2011). Convolutional neural networks for P300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(3), 433–445.

Chaurasiya, R. K., Londhe, N. D., & Ghosh, S. (2016). Binary DE-based channel selection and weighted ensemble of SVM classification for novel brain–computer interface using Devanagari script-based P300 speller paradigm. *International Journal of Human–Computer Interaction*, *32*(11), 861–877.

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, *70*(6), 510–523. https://doi.org/10.1016/0013-4694(88)90149-6

Guan, R., Liao, Y., & Wang, S. (2021). Adaptive morphological analysis method and its application for bearing fault diagnosis. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1–10. https://doi.org/10.1109/TIM.2021.3072116

Gunji, A., Goto, T., Kita, Y., Sakuma, R., Kokubo, N., Koike, T., . . . , & Inagaki, M. (2013). Facial identity recognition in children with autism spectrum disorders revealed by P300 analysis: A preliminary study. *Brain & Development*, *35*(4), 293–298. https://doi.org/10.1016/j.braindev.2012.12.008

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Jin, J., Allison, B. Z., Zhang, Y., Wang, X., & Cichocki, A. (2014). An ERP-based BCI using an oddball paradigm with different faces and reduced errors in critical functions. *International Journal of Neural Systems*, *24*(08), 1450027.

Kaufmann, T., Schulz, S. M., Grünzinger, C., & Kübler, A. (2011). Flashing characters with famous faces improves ERP-based brain–computer interface performance. *Journal of Neural Engineering*, *8*(5), 056016. https://doi.org/10.1088/1741-2560/8/5/056016

Kazakeviciute, M., Juozapavičius, A., & Samaitiene, R. (2010). Morphological filtering of EEG. *Materials Physics and Mechanics*, *9*(3), 185–193.

Kim, E., Lovera, J., Schaben, L., Melara, J., Bourdette, D., & Whitham, R. (2010). Novel method for measurement of fatigue in multiple sclerosis: Real-time digital fatigue score. *Journal of Rehabilitation Research and Development*, *47*(5), 477–484. https://doi.org/10.1682/jrrd.2009.09.0151

Kshirsagar, G. B., & Londhe, N. D. (2019). Improving performance of Devanagari script input-based P300 speller using deep learning. *IEEE Transactions on Biomedical Engineering*, *66*(11), 2992–3005. https://doi.org/10.1109/TBME.2018.2875024

Kshirsagar, G. B., & Londhe, N. D. (2022). DS-P3SNet: An efficient classification approach for Devanagari scriptbased P300 speller using compact channelwise convolution and knowledge distillation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *52*(12), 7431–7443.

Li, Q., Liu, S., Li, J., & Bai, O. (2015). Use of a green familiar faces paradigm improves P300-speller brain–computer interface performance. *PLoS ONE*, *10*(6), e0130325. https://doi.org/10.1371/journal.pone.0130325

Liu, M., Wu, W., Gu, Z., Yu, Z., Qi, F., & Li, Y. (2018). Deep learning based on batch normalization for P300 signal detection. *Neurocomputing*, *275*, 288–297.

Lu, Z., Gao, N., Zhou, W., Yang, J., Wu, J., & Li, Q. (2019). A comparison of facial p300-speller paradigm based on famous face and the familiar face. In *12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics,* 1–6. https://doi.org/10.1109/CISP-BMEI48845.2019.8965892

Lu, Z., Li, Q., Gao, N., & Yang, J. (2020). The self-face paradigm improves the performance of the P300-speller system. *Frontiers in Computational Neuroscience*, *13*, 93. https://doi.org/10.3389/fncom.2019.00093

Medina-Juliá, M. T., Fernández-Rodríguez, Á., Velasco-Álvarez, F., & Ron-Angevin, R. (2020). P300-based brain–computer interface speller: Usability evaluation of three speller sizes by severely motor-disabled patients. *Frontiers in Human

*Neuroscience*, *14*, 583358. https://doi.org/10.3389/fnhum.2020.583358

Meijer, E. H., Smulders, F. T. Y., & Wolf, A. (2009). The contribution of mere recognition to the P300 effect in a concealed information test. *Applied Psychophysiology & Biofeedback*, *34*(3), 221–226. https://doi.org/10.1007/s10484-009-9099-9

Minett, J. W., Zheng, H. Y., Fong, M. C., Zhou, L., Peng, G., & Wang, W. S. (2012). A Chinese text input brain–computer interface based on the P300 speller. *International Journal of Human-Computer Interaction*, *28*(7), 472–483.

Rakotomamonjy, A., & Guigue, V. (2008). BCI competition III: Dataset II-ensemble of SVMs for BCI P300 speller. *IEEE Transactions on Biomedical Engineering*, *55*(3), 1147–1154.

Sellers, E. W., & Donchin, E. (2006). A P300-based brain–computer interface: Initial tests by ALS patients. *Clinical Neurophysiology*, *117*(3), 538–548. https://doi.org/10.1016/j.clinph.2005.06.027

Sui, J., Zhu, Y., & Han, S. (2006). Self-face recognition in attended and unattended conditions: An event-related brain potential study. *NeuroReport*, *17*(4), 423–427. https://doi.org/10.1097/01.wnr.0000203357.65190.61

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, *113*(6), 767–791. https://doi.org/10.1016/s1388-2457(02)00057-3

Xiao, X., Xu, M., Jin, J., Wang, Y., Jung, T. P., & Ming, D. (2019). Discriminative canonical pattern matching for single-trial classification of ERP components. *IEEE Transactions on Biomedical Engineering*, *67*(8), 2266–2275.

**How to Cite:** Bhandari, V., Londhe, N. D., & Kshirsagar, G. B. (2024). Comparative Evaluation of Facial Paradigms for Devanagari Script-Based P300 Speller: Preliminary Findings. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS42021463

## Appendix

**P300 speller:** The P300 speller is a brain–computer interface system designed to enable communication for individuals with severe motor disabilities. It relies on the P300 event-related potential, a distinctive brainwave that occurs approximately 300 milliseconds after a person recognizes a relevant stimulus. By presenting a matrix of characters or symbols on a computer screen and analyzing the user's brain responses to focus on specific items, the P300 speller allows individuals to spell out words or convey messages using their brain activity.

**Devanagari script:** The ancient Indian script known as Devanagari is mostly used to write Sanskrit, Hindi, Marathi, and a few other Indian languages. In a Devanagari script-based P300 speller, the character matrix consists of Devanagari script characters and the user can choose the target characters by concentrating on the target character during the spelling process. The design of the Devanagari script-based P300 speller is mainly driven by the idea of providing a communication facility in the native language to patients suffering from motor neuron diseases.

**Family face paradigm:** Family face paradigm is a newer approach in P300 speller design. In the family face paradigm, family member faces are used as visual stimuli in the traditional row–column matrix speller design. During the spelling process, as the row and column of the matrix are flashed, the characters in the flashing rows and columns are covered with the faces of the family members. The objective of this paradigm is to introduce emotional connection and familiarity to the design of the display paradigm.

**Smiley face paradigm:** In this paradigm, the conventional character matrix incorporates visual stimuli represented by yellow smiley faces. Throughout the spelling process, as the row and column of the matrix are flashed, they are overlaid with yellow smiley faces. The objective of this paradigm is to introduce positivity and cheerful emotion to the design of the display paradigm.

**Famous face paradigm:** In the famous face paradigm, the characters in the flashing row and column are covered with images of famous personalities during the P300 spelling operation. This design choice adds a sense of familiarity to the conventional row–column matrix to improve the user experience during the spelling process.