**Problems that want to be solved:**

increasing the number of dimensions leads to growth ratio of saddle points to local minimums. This phenomenon results in slow convergence in deep neural network.

hyperparameters seriously influence the convergence and the performance of the deep neural network.

Hyperparameter tuning is one of the approaches in deep learning, which leads to fast convergence

because of its ability to find better local minimums. In this paper, a new adaptive hyperparameter

tuning method is proposed to improve training of Convolutional Neural Networks (CNNs)

Solution:

The proposed method minimizes nonconvex error function in high‑dimensional space. For this purpose, the objective function is transferred to Lagrange paradigm as a two‑constrained optimization problem including either quality or inequality constraints. Then, Karush–Kuhn–Tucker (KKT) system is used to translate the problem into a standard nonlinear framework. Finally, the hyperparameter tuning is achieved using iterative Newton and dual active set techniques. The proposed method is applied on Adadelta module of a CNN to obtain a well‑fit model for distinguishing P300 and non‑P300 signals.

Data Set: EPFL BCI dataset

Table

Description automatically generatedEPFL BCI dataset was applied. It has been captured by the Biosemi system with 32 electrodes located according to standard 10‑20 international system position at 2048 HZ. The EPFL BCI dataset is composed of eight available subjects

Flashing Details:

The data of each subject is composed of four sessions. Each of the sessions consisted of six runs; they corresponded to six images which had displayed in a six‑cell paradigm. The images were flashed in random order. Each flash of an image lasted for 100 MS and then during 300 MS none of them were flashed. In a six‑cell paradigm

Preprocessing:

1-Refenceing:

The EEG voltages have been recorded using 32 electrodes which have depended on each other. Hence, activities in the reference electrode may be reflected in the recorded signal of others. Referencing method used to remove such anomaly effects in the activity of electrodes

2-Filtering:

A sixth‑order forward–backward Butterworth bandpass filter was used at 1.0 and 12.0 Hz cutoff signal frequency,[30] to remove additional frequencies and noise.

3-Downsampling:

The EEG was downsampled from 2048 to 32 Hz by selecting each 64th sample from the bandpass‑filtered data

4-Signal trial extraction:

Single trials of duration 1000 MS were extracted from the data. These trials were started at stimulus onset, that is, at the beginning of the intensification of an image and were ended 1000 MS after stimulus onset. Due to the ISI of 400 MS, the last 600 MS of each trial overlapped with the last 600 MS of the following trial

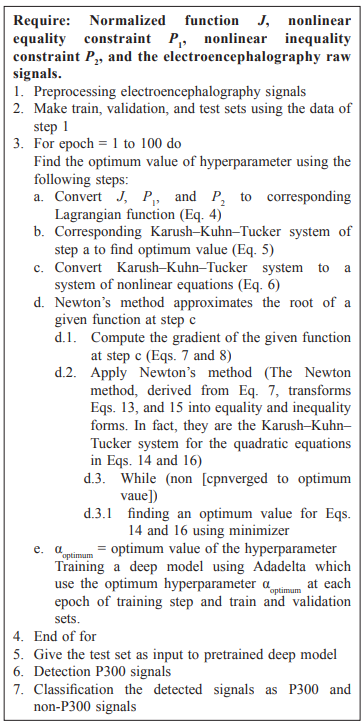
5-Windsorzing:

To reduce the effects of outliers in EEG signal such as muscles activity and eye blinks, the 10th percentile and the 90th percentile of data of each electrode were computed. Amplitude values lying out of this range were replaced by start and end limits of this range, respectively.

6-Normalizing:

The last stage of preprocessing is mapping the EEG signal to the range of (0, 1) to set data onto the same range, so the computational complexity has a sharp decrease

Modeling Algorithm:



Accuracy: average accuracy is 91%

Extra Informations:

A pretrained CNN is applied to detect P300 signals. The CNN builds up a model to map between the features and EEG signals categories, which are known as P300 and non‑P300 with binary labels

Electrods Position: 32 electrodes positioned at:

