Denoising EEG Signals Using Deep Learning

Shubh Goel

Supervised By: Prof. Lalan Kumar

Introduction

- Electroencephalography (EEG) - electrical response of brain cells in the cerebral cortex
- Analysis of EEG - physiological, psychological and pathological information.
- The recorded electroencephalography (EEG) signals are usually contaminated by many artifacts
- These artifacts exist in almost the entire acquisition process and often mask the waveform characteristics of EEG, which makes the reading of EEG signals more difficult and brings great difficulties to the subsequent research and application of EEG signal

Data Preprocessing

- WAY-EEG-GAL Dataset*
  - Twelve participants performed lifting series in which the object’s weight, surface friction , or both, were changed unpredictably between trials, thus enforcing changes in fingertip force coordination.
  - 32-channel EEG was recorded for each participant.
- Removed various artifacts from each of the participant’s EEG Signal by performing Independent Component Analysis using EEGLAB toolbox. This formed the (Raw EEG, Denoised EEG) pairs that are then fed to the model for training.
- No. of channels used for training – 18 (around motor-cortex region)

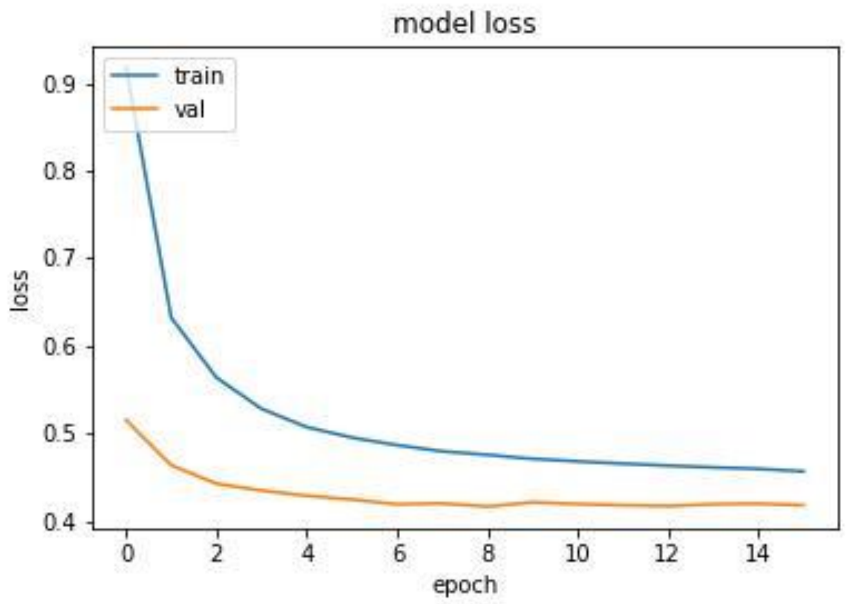
*Luciw, Matthew D., Ewa Jarocka, and Benoni B. Edin. "Multi-channel EEG recordings during 3,936 grasp and lift trials with varying weight and friction." Scientific data 1, no. 1 (2014): 1-11.

Subjective Analysis

- The models are trained on each of the subject’s data separately and then the results are analyzed.

MLP Model

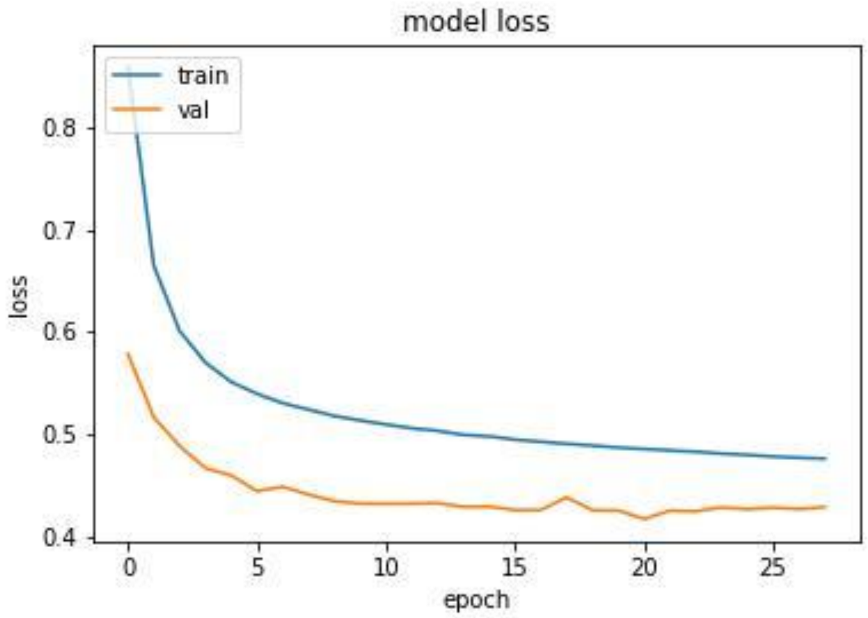
Input Layer(18) → Dense(64) (•BatchNorm •ReLU •DropOut) → Dense(128) (•BatchNorm •ReLU •DropOut) → Dense(128) (•BatchNorm •ReLU •DropOut) → Dense(64) (•BatchNorm •ReLU •DropOut) → Output Layer(18)



	mse	pcc
0	0.574838	0.678749
1	0.537207	0.765521
2	0.411375	0.677787
3	0.382918	0.761234
4	0.493031	0.792084
5	0.465836	0.777543
6	0.446246	0.780797
7	0.311796	0.872219
8	0.398787	0.815311
9	0.193352	0.897180
10	0.325686	0.769134
11	0.340887	0.815903
12	0.383102	0.797253
13	0.285175	0.791208
14	0.271698	0.877171
15	0.500641	0.726486
16	0.517617	0.649403
17	0.381402	0.773079

Participant 8’s data gave the best results with test set loss of 0.4012

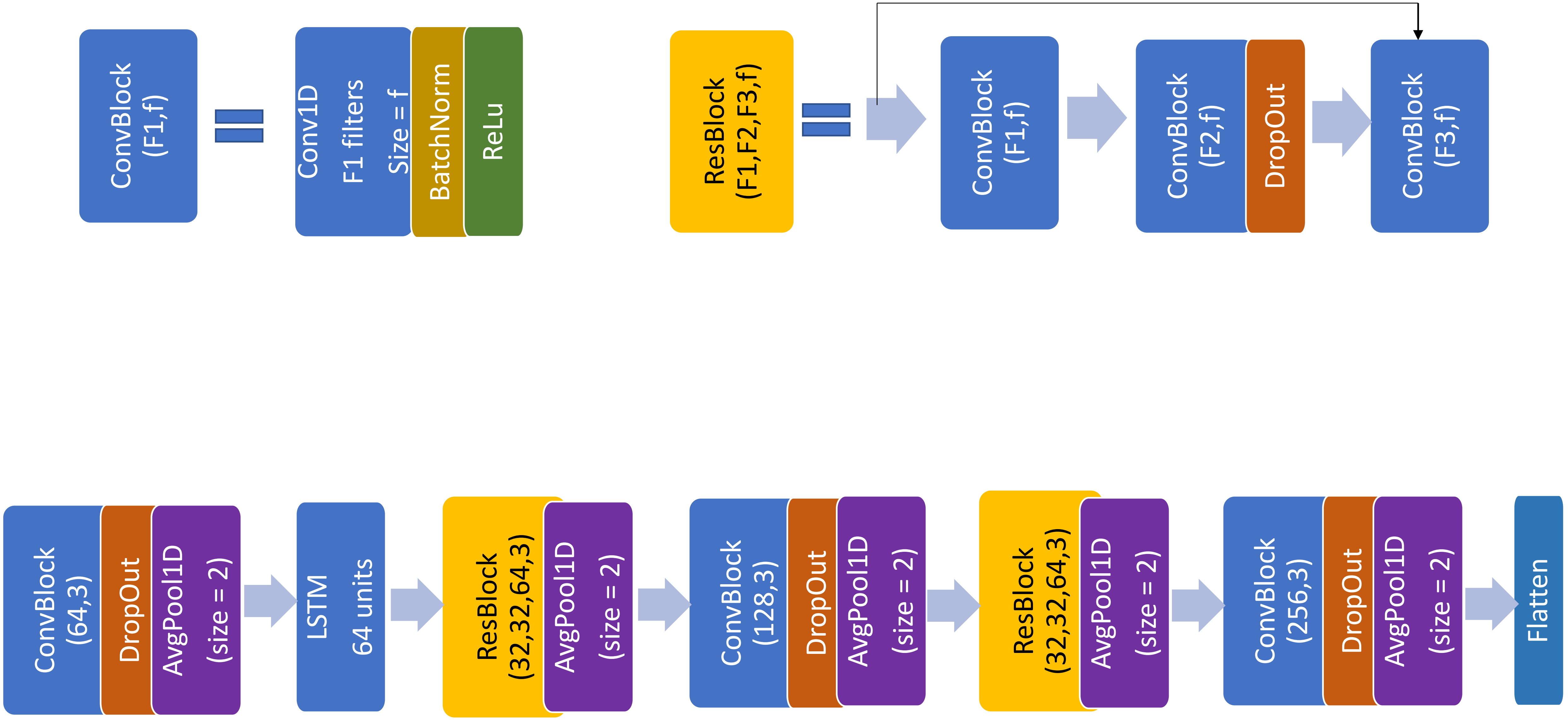
CNN-LSTM Model



	mse	pcc
0	0.549174	0.683907
1	0.484211	0.774394
2	0.400108	0.688294
3	0.374169	0.768954
4	0.490898	0.774240
5	0.444602	0.781117
6	0.435493	0.781734
7	0.302878	0.876562
8	0.389249	0.811479
9	0.198923	0.883389
10	0.332382	0.776678
11	0.349075	0.813089
12	0.394385	0.786514
13	0.272439	0.803187
14	0.297673	0.854438
15	0.523166	0.711350
16	0.538905	0.615994
17	0.425469	0.741683

Participant 8’s data gave the best results with test set loss of 0.4000

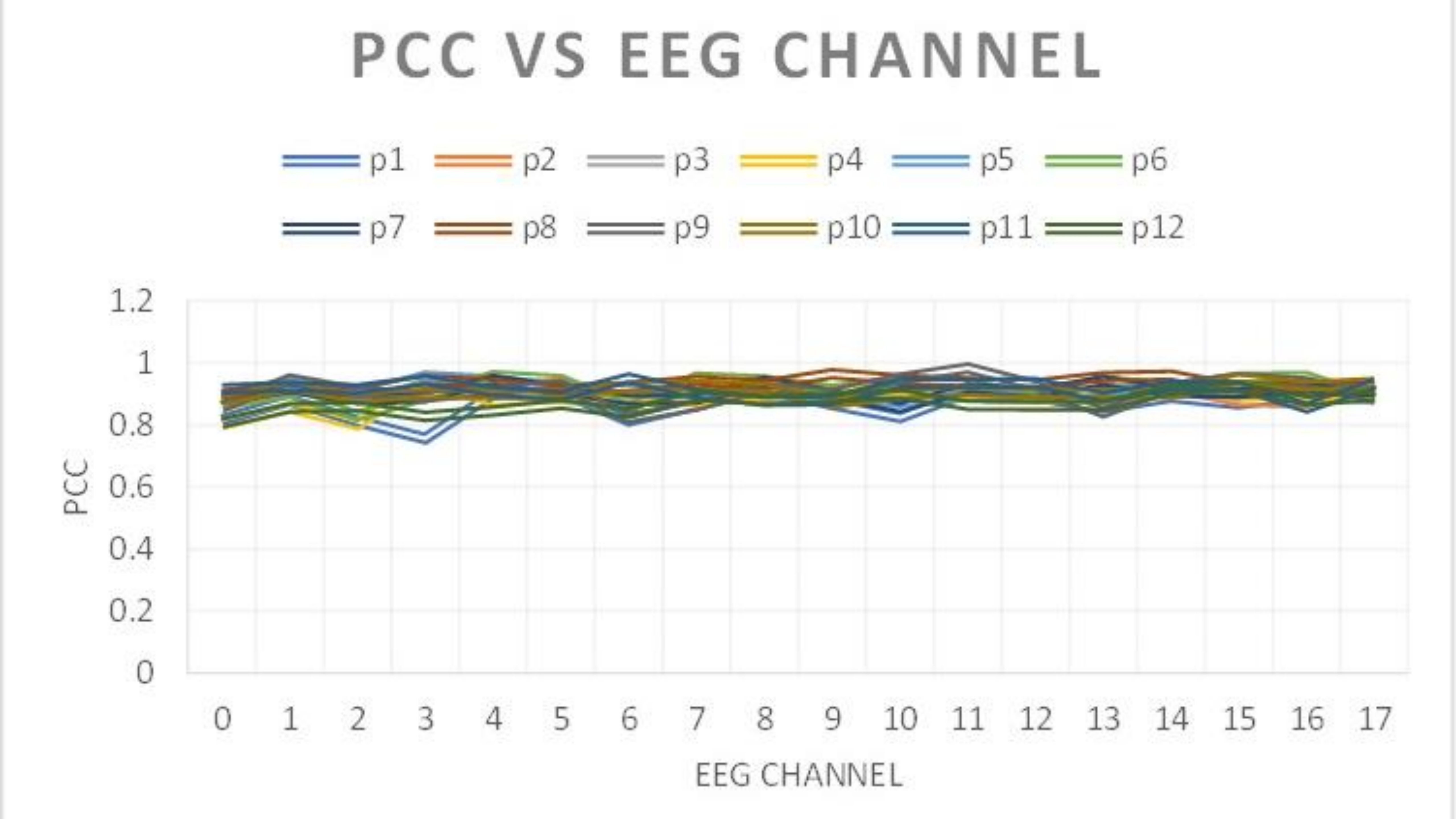
LSTM-ResNet Model



Results obtained after running the LSTM-Resnet Model on pre-processed data

- After pre-processing the data correctly, the LSTM-Resnet Model was trained on the data. I obtained the following PCC after running the Model on the test set of each participant.

	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	p11	p12
0	0.817587	0.859926	0.894835	0.803631	0.896244	0.870616	0.883873	0.900398	0.859847	0.879004	0.91736	0.807091
1	0.879028	0.92364	0.915608	0.85616	0.937544	0.902751	0.91815	0.923624	0.948275	0.92577	0.924278	0.855459
2	0.813242	0.894681	0.8966	0.801986	0.897171	0.828911	0.882464	0.906544	0.908662	0.887809	0.916671	0.859708
3	0.756481	0.893652	0.89694	0.945348	0.957413	0.892705	0.889201	0.946793	0.89032	0.902177	0.945411	0.827603
4	0.901849	0.920719	0.906908	0.876998	0.945184	0.958887	0.946419	0.934533	0.911556	0.897426	0.908853	0.845939
5	0.907936	0.935837	0.884614	0.888594	0.921433	0.945811	0.913063	0.912809	0.929677	0.902567	0.89538	0.867336
6	0.814231	0.881181	0.876035	0.896771	0.915551	0.855453	0.881144	0.924144	0.823122	0.921719	0.951132	0.838996
7	0.863572	0.876592	0.894177	0.875771	0.902897	0.953736	0.883849	0.942903	0.862753	0.929469	0.901835	0.894068
8	0.915681	0.932734	0.884877	0.895867	0.924713	0.944685	0.942077	0.931236	0.921983	0.920441	0.876278	0.881098
9	0.867768	0.91028	0.889342	0.914977	0.887874	0.915839	0.899242	0.964702	0.901236	0.869692	0.879188	0.889122
10	0.825174	0.921581	0.889511	0.909879	0.873896	0.941482	0.856356	0.94745	0.950631	0.898447	0.935952	0.900118
11	0.905452	0.892117	0.913047	0.897667	0.892961	0.93755	0.920268	0.938763	0.984071	0.911237	0.932019	0.863692
12	0.892457	0.886218	0.914325	0.887106	0.889813	0.885089	0.903917	0.932564	0.927376	0.903133	0.937005	0.862795
13	0.855735	0.887959	0.870908	0.898801	0.909608	0.945458	0.943039	0.955814	0.840483	0.902485	0.889689	0.862156
14	0.890474	0.922943	0.913126	0.915154	0.926449	0.920378	0.905398	0.960305	0.923439	0.89924	0.927011	0.913831
15	0.868761	0.878257	0.912991	0.895099	0.927027	0.95339	0.902602	0.920167	0.952728	0.948735	0.925489	0.920692
16	0.897709	0.871546	0.904994	0.895793	0.909791	0.955727	0.920589	0.927226	0.938588	0.926607	0.856048	0.877879
17	0.886341	0.93613	0.898416	0.903243	0.912638	0.881316	0.907613	0.92403	0.888521	0.93654	0.934586	0.894157



PCC VS EEG CHANNEL
p1 p2 p3 p4 p5 p6
p7 p8 p9 p10 p11 p12
1.2
1
0.8
0.6
0.4
0.2
0
PCC
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
EEG CHANNEL

Potential Applications

- **Clinical diagnosis:** EEG is a commonly used tool for diagnosing neurological disorders such as epilepsy, sleep disorders, and dementia. However, artefacts in EEG data can make diagnosis challenging. The developed model can improve the accuracy of EEG-based diagnoses and help clinicians make more informed treatment decisions.
- **Neuroscience research:** EEG is also widely used in cognitive neuroscience research to investigate brain function and cognition. The developed model can help researchers obtain cleaner EEG signals and thus more accurately measure brain activity.
- **Neuroengineering:** EEG-based brain-computer interfaces (BCIs) are used to control prosthetic devices, communicate with locked-in patients, and even play video games. However, artefacts in EEG data can degrade the performance of BCIs. The developed model can improve the accuracy and reliability of BCIs, making them more useful in practical applications.

Future Work

- **Generalization:** I trained my model on 12 different subjects. To make the model more useful in practice, I could try to generalize it to new subjects. This could involve collecting more EEG data from different subjects and training the model on a larger and more diverse dataset.
- **Real-time processing:** EEG-based applications often require real-time processing, which can be challenging for complex neural network models. I could investigate ways to optimize my model for real-time processing, such as reducing the number of layers or using a more efficient architecture.