C3 Project Report

EEG Signal Processing for Autism Spectrum Disorder

Prepared by

Group Number: 9

AMIT KUMAR	IIT2018131	Section B
ASHUTOSH MISHRA	IIT2018133	Section B
SANCHIT JAISWAL	IIT2018180	Section B
KARTIC CHOUBEY	IIT2018181	Section B

Instructor: Dr. Anupam Agarwal, Gopal Chandra Jana, Trapti Shrivastava

Course: IIVP

Date: 30-08-2020

TABLE OF CONTENTS

1 Abstract

In this report, we are going to show how we can use the Electroencephalogram (EEG) signals for finding whether someone is suffering from autistic or not. We also described the research paper , the tools (hardware & software) that we have used for this project & methodology and Implementation of this project.

2 Introduction & Motivation

PROJECT TITLE: EEG Signal Processing for Autism Spectrum Disorder

Autism spectrum disorder (ASD) is a developmental **disorder** that affects communication and behavior. Although **autism** can be diagnosed at any age, it is said to be a "developmental **disorder**" because symptoms generally appear in the first two years of life.

According to the CDC, one in 59 children is estimated to have autism. Autism spectrum disorder is also three to four times more common in boys than in girls, and many girls with ASD exhibit less obvious signs compared to boys. Autism is a lifelong condition. However, many children diagnosed with ASD go on to live independent, productive, and fulfilling lives.

The **electroencephalogram (EEG)** is a recording of the electrical activity of the brain from the scalp. The recorded waveforms reflect the cortical electrical activity. ... Signal frequency: the main frequencies of the human EEG waves are: Delta: has a frequency of 3 Hz or below.

Both of the techniques can be extensively used to deal with the prediction of the EEG signal, which is an emerging technique in today's world for usage in BCI, which is an extensive research topic. EEG signal processing aims at translating raw EEG signals into the class of these signals, i.e., into the estimated mental state of the user. This translation is usually achieved using a pattern recognition approach.

3 Problem Statement & Objective

Statement 1: EEG signal processing for autism spectrum disorder

OBJECTIVE:

The relation of ASD with EEG signal signifies decline of EEG complexity perceived in autistic children. The essential contrasts were seen between cerebrum districts in the correct side of the equator and the focal cortex spoke to by [4]. The relation of ASD with EEG signal signifies decline of EEG complexity perceived in autistic children. The essential contrasts were seen between cerebrum districts in the correct side of the equator and the focal cortex spoke to by [4].The relation of ASD with EEG signal signifies decline of EEG complexity perceived in autistic children. The essential contrasts were seen between cerebrum districts in the correct side of the equator and the focal cortex spoke to by [4].EEG signals are electrical voltage triggered on the electrodes by brain electromagnetic signals (BEMS) [5]. At the hospital, a specialized technician's measures, marks and puts about 16-25 electrodes on the patients. Many research shows that the EEG signals, which consists of Alpha, Beta, Gamma, Delta and Theta, of ASD children are not the same as compared to non-ASD children [6].

Deep learning combines feature extraction and classification processes. Feature extraction is the process of producing a dense but highly meaningful digital presentation of someone fundamental biometric attribute. This can be used to make grouping of samples according to the provided data, and can be useful in early prediction of Autism Spectrum Disorders.

4 Literature Review

Autism spectrum disorder classification electroencephalogram signal using deep learning algorithm.

In the proposed report the authors look at a novel deep learning approach to solve the problem statement. Electroencephalogram (EEG) is a method used in purpose to find out the brain function regardless of the mental states of the individuals [13]. EEG provides robust parameters where it can examine the brain activity at rest state and active state. EEG can also show which part of the brain is active when doing specific tasks and the effect of those specific tasks. There are five types of frequency bands associated with EEG signals namely the alpha, beta, gamma, theta and delta.

Table 1. EEG Brain Signal of Normal Human Being [14]

Level	Frequency range	Approximate EEG label
1	64-128 Hz	High Gamma
2	32-64 Hz	Gamma
3	16-32 Hz	Beta
4	8-16 Hz	Alpha
5	4-8 Hz	Theta
6	0-4 Hz	Delta

The authors have used Convolutional Neural Network to extract the required information from the provided data without labeling or segmenting data manually.Deep learning particularly CNN is deployed due to the capability of the algorithm to actually recognize unique features of electrical brainwave patterns from EEG signals. Many studies have shown success in integrating EEG signals into the CNN algorithm. EEG is a non-linear and highly intricate signal that records important data. This data portrays the differences of one human being to another. Since EEG is very complex and non-linear in nature, many linear classifier methods could not accurately detect this signal as EEG.

4.2 Research flowchart

The following research flow was used by the authors.

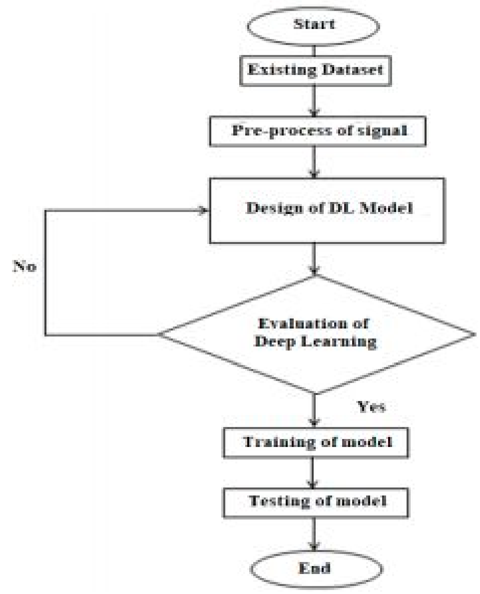


Figure 4. Main development involved

[2][pg 94]

4.2 Preprocessing of EEG signal

All the dataset is pre-processed according to the time taken to record the data. This step is done by using MATLAB software. The intermediate file is a csv file named randomly. Then all the csv files are combined into a single csv file and the samples are labelled according to the following: a normal person is given a label of 1 and an autistic person is given a label of 2. Finally, the data is implemented into a pre-processing algorithm that is used for augmentation and removal of noise using random shuffling and white Gaussian noise.

4.3 Design of deep learning algorithm

The deep learning model is designed to fit the EEG data in 2D matrix form. The model is shown in Figure 5 produced in Python. The proposed CNN architecture has 3 Convolutional layers, each of which is followed by a batch normalization layer, which is then followed by two dense layers of 160 and 20 units. Adam optimizer is used for training of the data on over 200 epochs with a learning rate of 0.0001. The activation function used in all layers except the last layer is ReLU, in the last layer sigmoid activation function is used. The data is distributed as mentioned in Table 3

Table 3. Distribution of dataset

Model	Distribution (%)
Training Set	80%
Validation Set	10%
Test Set	10%

[2][pg 96]

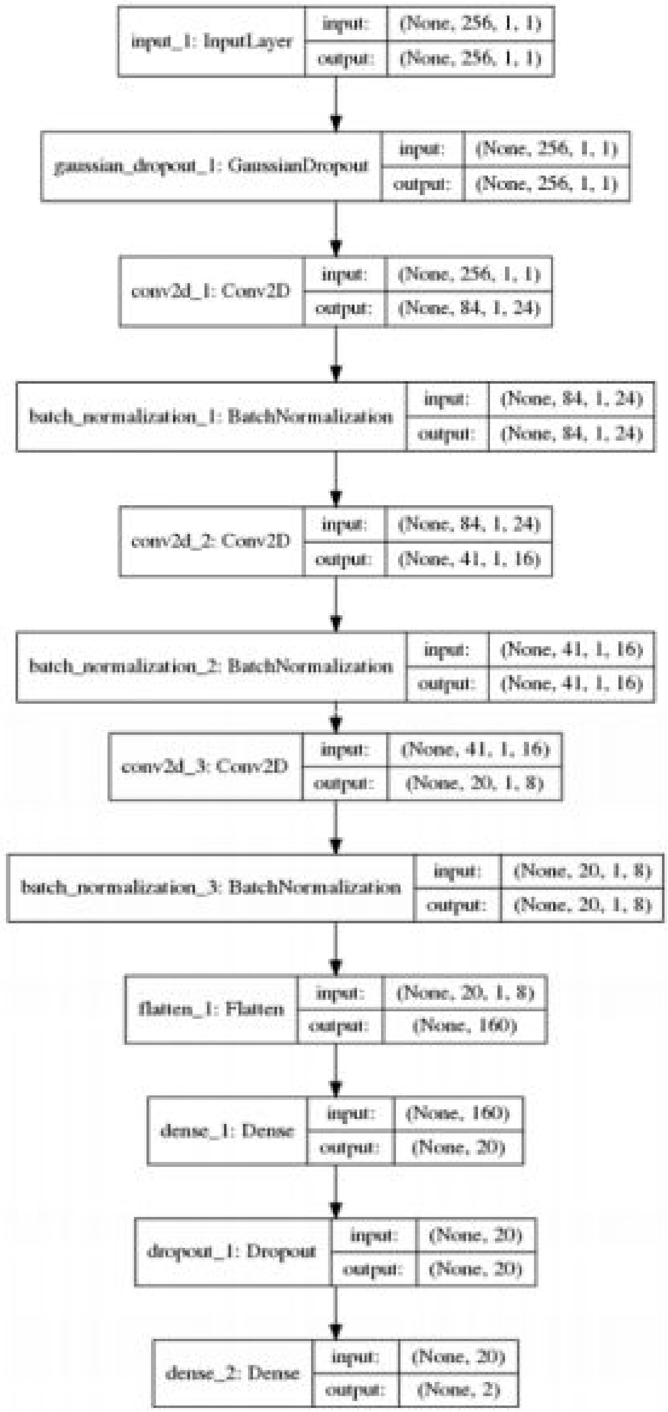


Figure 5. Design of Deep Learning Model

[2][pg 97][2][pg 94]

5 Methodology

5.1 Data Analysis and feature extraction:

The data contains eeg reports from different subjects recorded in mostly a calm and stable environment.EEG provides an image of electrical activity in the brain represented as waves of varying frequency, amplitude, and shape. It can be used to measure brain activity that occurs during an event--like the completion of a task or the presentation of a stimulus--or to measure spontaneous brain activity that happens in the absence of a specific event.

There are five types of frequency bands associated with EEG signals namely the alpha, beta, gamma, theta and delta. As aforementioned each of these brain waves has their own frequency range and each is associated with its own physiological characteristics which govern its own cognitive (the mental action or process of acquiring knowledge)properties. During EEG recording, patients will be shown a series of simulations while sitting down.

Wavelet level	Wavelet frequency range	Approximate EEG label
1	**32–64Hz**	**gamma**
2	**16–32Hz**	**beta**
3	**8–16Hz**	**alpha**
4	**4–8Hz**	**theta**
5	**0–4Hz**	**delta**

Table 4: Segmentation of various EEG signals

5.2 Data Preprocessing:

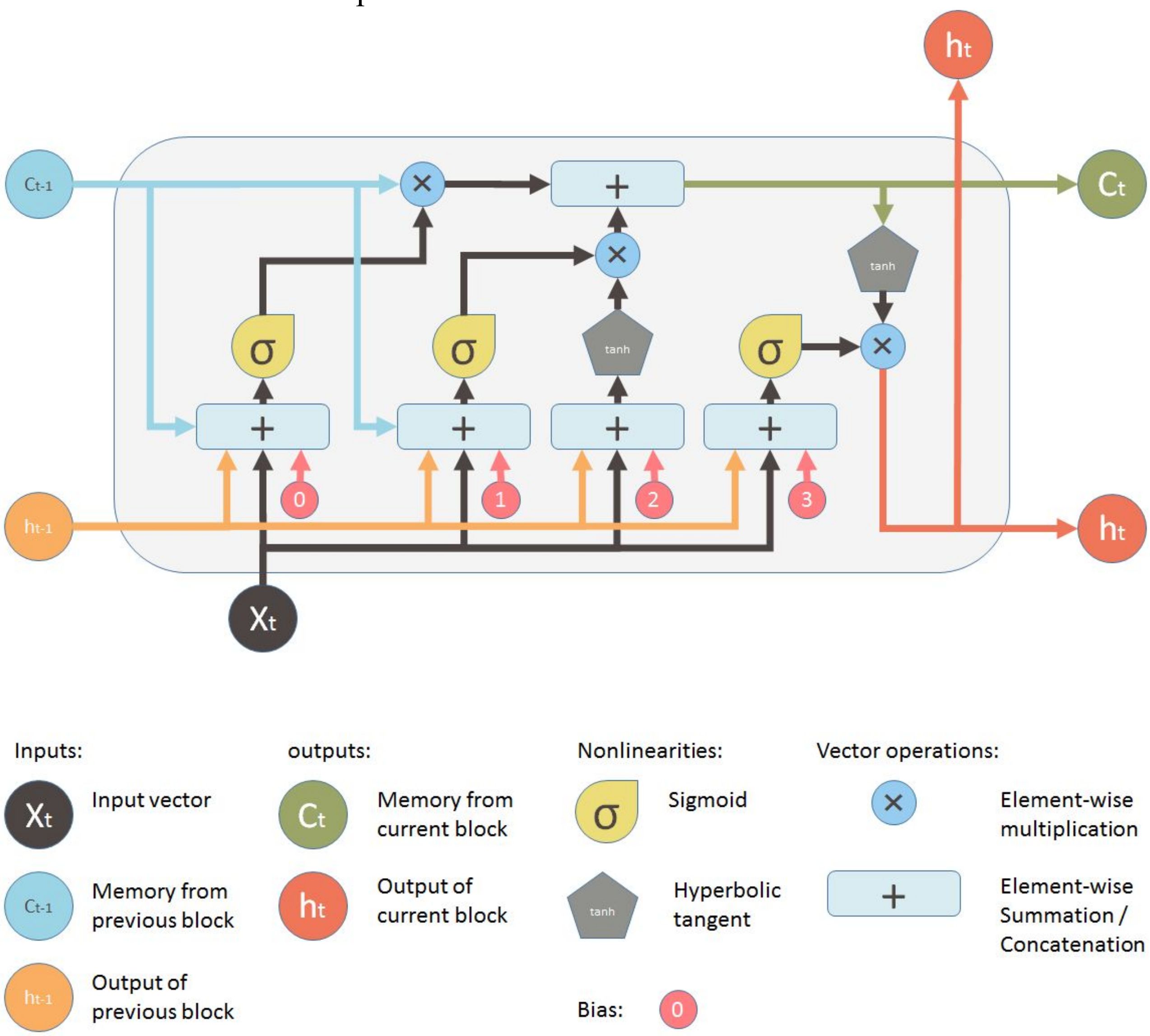
The EEG data of both datasets are preprocessed and the DE features are extracted.

- For SAFARI dataset, DE features are extracted per second from five frequency bands: δ : 1–3 Hz, θ :4–7 Hz, α : 8–13 Hz, β : 14–30 Hz, and γ : 31–50 Hz .
- The input data of each subject thus have 310 dimensions (62 channels ×5 frequency bands). The number of samples with labels for each subject is 3394.
- For Australian EEG dataset, DE features are extracted per second except for δ frequency due to the fact that the low frequency band is filtered in this dataset. The input data of each subject thus have 128 dimensions (32 channels ×4 frequency bands). The number of samples with labels for each subject is 2400.

We are proposing use of data preprocessing techniques in the likes of discrete wavelet transform on eeg data, to remove major artifact noise. If required we will also use a non-linear method, namely Shannon entropy.

5.3 Proposed Architecture:

The proposed architecture of neural network is RNN with 2 LSTM layers with Dropouts for regularization and Fully Connected layer with 3 units and softmax as activation function for Output.



6 Software & Hardware Requirements

The Software Requirements for this project are:

1. Miniconda
2. TensorFlow
3. Scipy
4. Jupyter

7 Implementation

The current implementation architecture is using an RNN with 2 layers of LSTM with 64 and 32 memory cells. The activation functions used are ReLU. The last layer is a fully connected layer(Dense) with 3 units and softmax as activation. We have chosen softmax activation as we are doing a multiclass classification and the softmax is useful in limiting the total prediction probability to 1. For preprocessing we have used Fast Fourier Transform to convert the EEG signal to preprocessed context data.

8 Comparison

➔ <u>Base Paper Architecture vs Our Architecture</u> :

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv1d_26 (Conv1D)           (None, 1, 32)             244640
_________________________________________________________________
max_pooling1d_21 (MaxPooling (None, 1, 32)             0
_________________________________________________________________
lstm_9 (LSTM)                (None, 2)                 280
_________________________________________________________________
dense_16 (Dense)             (None, 3)                 9
=================================================================
Total params: 244,929
Trainable params: 244,929
Non-trainable params: 0
_________________________________________________________________
```

BASE PAPER

```
Model: "sequential_3"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_8 (LSTM)                (None, 1, 64)             668928
_________________________________________________________________
dropout_8 (Dropout)          (None, 1, 64)             0
_________________________________________________________________
lstm_9 (LSTM)                (None, 1, 32)             12416
_________________________________________________________________
dropout_9 (Dropout)          (None, 1, 32)             0
_________________________________________________________________
lstm_10 (LSTM)               (None, 16)                3136
_________________________________________________________________
dropout_10 (Dropout)         (None, 16)                0
_________________________________________________________________
dense_3 (Dense)              (None, 3)                 51
=================================================================
Total params: 684,531
Trainable params: 684,531
Non-trainable params: 0
_________________________________________________________________
```

➔ Evaluation Metric:

BASE

	precision	recall	f1-score	support
0	0.99	0.96	0.98	148
1	0.95	0.86	0.90	136
2	0.88	0.99	0.93	143
accuracy			0.94	427
macro avg	0.94	0.94	0.94	427
weighted avg	0.94	0.94	0.94	427

---

	precision	recall	f1-score	support
0	1.00	1.00	1.00	148
1	0.98	0.94	0.96	136
2	0.95	0.99	0.97	143
accuracy			0.98	427
macro avg	0.98	0.98	0.98	427
weighted avg	0.98	0.98	0.98	427

➔ Accuracy:

BASE

```
Accuracy after Fine tuning Model 0.9391100406646729
```

```
Accuracy after Fine tuning Model 0.976580798625946
```

9 Results

In our model, we are predicting whether a child is autistic or not with the help of signals produced by a child's brain. By using the preprocessed dataset of Brain Signals, The Accuracy of our model comes out to be 92%. That means for every 100 brain signals, our model will predict 92 times right.

10 Conclusion

By this, we can conclude that by using RNN with 2 layers of LSTM with 64 and 34 memory cells which is a deep learning algorithm, we can find out whether a child is autistic or not by using our algorithm on their brain signals. EEG signal processing can be used in Neuromarketing, Social Interactions, Psychology and Neuroscience, Clinical and Psychiatric Studies and Brain Computer Interfaces(BCI).

11 References

[1] Sutrisno Ibrahim, Ridha Djemal , Abdullah Alsuwailem. Electroencephalography (EEG) signal processing for epilepsy and autism spectrum disorder diagnosis.Electrical Engineering Department, King Saud University, Riyadh, Saudi Arabia.

[2] N. A Ali, A.R Syafeeza, A. S Jaafar, M.K Mohd Fitri Alif . Autism spektrum disorder classification on electroencephalogram signal using deep learning algorithm.Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia. (BASE PAPER)

[3] Lonnie Zwaigenbaum, Jessica A. Brian, Angie Ip. Early detection for autism spectrum disorder in young children . Canadian Paediatric Society, Autism Spectrum Disorder Guidelines Task Force, Ottawa, Ontario.

[4] William J. Bosl, Helen Tager-Flusberg & Charles A. Nelson.EEG Analytics for Early Detection of Autism Spectrum Disorder: A data-driven approach. Published on 1 May 2018.

[5] Ridha Djemal, Khalil AlSharabi, Sutrisno Ibrahim, and Abdullah Alsuwailem. EEG-Based Computer Aided Diagnosis of Autism Spectrum Disorder Using Wavelet, Entropy, and ANN. Electrical Engineering Department, College of Engineering, King Saud University, Box 800, Riyadh 11421, Saudi Arabia.

[6] The-Hanh Pham, Jahmunah Vicnesh, Joel Koh En Wei, Shu Lih Oh, N. Arunkumar, Enas. W. Abdulhay, Edward J. Ciaccio and U. Rajendra Acharya. Autism Spectrum Disorder Diagnostics System Using HOS Bispectrum with EEG Signals. Published: 4 February 2020

[7] T. Heunis, C.Aldrich, J. M. Peters, S. S. Jeste, M. Sahin, C. Scheffer and P. J. de Vries. Recurrence quantification analysis of resting state EEG signals in autism spectrum disorder - a systematic methodological exploration of technical and demographic confounders in the search for biomarkers. https://doi.org/10.1186/s12916-018-1086-7.

[8] N. Fauzan And N. H. Amran, “Brain Waves And Connectivity Of Autism Spectrum Disorders,” Procedia - Soc.
Behav. Sci., Vol. 171, Pp. 882–890, 2015.

[9] L. Bote-Curiel, S. Muñoz-Romero, A. Guerrero-Curieses, And J. L. Rojo-Álvarez, “Deep Learning And Big Data In
Healthcare: A Double Review For Critical Beginners,” Appl. Sci., Vol. 9, No. 11, 2019.

[10] F. C. Morabito Et Al., “Deep Convolutional Neural Networks For Classification Of Mild Cognitive Impaired And
Alzheimer’s Disease Patients From Scalp Eeg Recordings,” 2016 Ieee 2nd Int. Forum Res. Technol. Soc. Ind.
Leveraging A Better Tomorrow, Rtsi 2016, 2016.

[11] W. H. L. Pinaya, A. Mechelli, And J. R. Sato, “Using Deep Autoencoders To Identify Abnormal Brain Structural
Patterns In Neuropsychiatric Disorders: A Large-Scale Multi-Sample Study,” Hum. Brain Mapp., Vol. 40, No. 3,
Pp. 944–954, 2019.

[12] G. Cho, J. Yim, Y. Choi, J. Ko, And S. H. Lee, “Review Of Machine Learning Algorithms For Diagnosing Mental
Illness,” Psychiatry Investig., Vol. 16, No. 4, Pp. 262–269, 2019.

[13] W. Bosl, A. Tierney, H. Tager-Flusberg, And C. Nelson, “Eeg Complexity As A Biomarker For Autism Spectrum
Disorder Risk,” Bmc Med., Vol. 9, 2011.

[14] M. S. Mythili And A. R. M. Shanavas, “A Study On Autism Spectrum Disorders Using Classification Techniques,”
Ijcsit, Vol. 5, No. 6, Pp. 7288–7291, 2014.

[15] J. B. Ewen Et Al., “Decreased Modulation Of Eeg Oscillations In High-Functioning Autism During A Motor
Control Task,” Vol. 10, No. May, Pp. 1–11, 2016.

[16] T. Heunis Et Al., “Recurrence Quantification Analysis Of Resting State Eeg Signals In Autism Spectrum Disorder -
A Systematic Methodological Exploration Of Technical And Demographic Confounders In The Search For
Biomarkers,” Bmc Med., Vol. 16, No. 1, Pp. 1–17, 2018.

[17] I. Vasilev, D. Slater, G. Spacagna, P. Roelants, And V. Zocca, Python Deep Learning: Exploring Deep Learning
Techniques And Neural Network Architectures With Pytorch, Keras And Tensorflow, 2nd Edition. Birmingham, Uk:
Packt Publishing, 2018.

[18] M. J. Alhaddad Et Al., “Diagnosis Autism By Fisher Linear Discriminant Analysis Flda Via Eeg,” Vol. 4, No. 2, Pp. 45–54, 2012.

[19] Gavin Hackeling, Mastering Machine Learning With Scikit-Learn Second, 2nd Edition. Packt Publishing, 2017.

[20] P. Marzuki, A. R. Syafeeza, A. N. Alisa, And M. K. M. F. Alif, “An Improved Of Malaysian License Plates Detection Using Deep Convolutional Neural Networks,” Symp. Electr. Mechatronics Appl. Sci. 2018, Vol. 2018, No. November, Pp. 75–76, 2018.

[21] N. A. Ali, A. R. Syafeeza, L. J. Geok, Y. C. Wong, N. A. Hamid, And A. S. Jaafar, “Design Of Automated Computer-Aided Classification Of Brain Tumor Using Deep Learning,” In Intelligent And Interactive Computing, 2019, Pp. 285–291.

[22] AlexNet Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks", Advances in neural information processing systems, 2012