REPUBLIC OF TURKEY
YILDIZ TECHNICAL UNIVERSITY
DEPARTMENT OF COMPUTER ENGINEERING

STRESS DETECTION IN PHYSIOLOGICAL SIGNALS OF NEURODIVERSE PEOPLE USING DEEP LEARNING METHODS

16011044 — İBRAHİM FURKAN ERÇELEBİ

SENIOR PROJECT

Advisor
Assistant Professor Lecturer Hafiza Irem TURKMEN CILINGIR

Aralık, 2024

ACKNOWLEDGEMENTS

I would like to thank my mentor Assistant Professor Lecturer Hafiza Irem TURKMEN CILINGIR and the department members for their support in this project.

İBRAHİM FURKAN ERÇELEBİ

TABLE OF CONTENTS

LIST OF SYMBOLS

μ	Mean
σ^2	Variation
ϵ	Tripple Factor
ω	Angular Frequency
T_n	Chebyshev Polynomial
	Normal Distribution

LIST OF ABBREVIATIONS

ECG	Electrocardioagram , recording hearts electrical activity through cardiac cycles.
TEB	Thoracic(related to upper body) electrical bioimpedance
EEG	Electroencephalograps , recording spontaneous electrical activities of the brain.
EMG	Electromyography , recording electrical activity produced by skeletal muscle
EDA	Electrodermal activity , continues variation of electrical characteristic of the skin.
SCR	Skin Conductance Responses.
SCL	Skin Conductance Level.
PPG	Photoplethysmogram , record blood volume changes in the multivascular bed of tissue.
ITH	Infrared thermopile , measures skin temperature on a scale of 40 to 115° Celsius
ST	Skin temperature
RESP	Respiration
HRV	Heart Rate Variability
BVP	Blood Volume Pulse
ACC	3-axis Accelerometer data
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
Saas	Software as a service
FOSG	Fifth-Order Savitzky-Golay
SOCT	Sixth-Order Chebyshev II

SDNN	Standard deviation of NN intervals
RMSSD	Root mean square of successive differences between normal heartbeats
PNN20	Percentage of successive interbeat intervals that differ by more than 20 ms
PNN50	Percentage of successive interbeat intervals that differ by more than 50 ms
ILSVRC	ImageNet Large-Scale Visual Recognition Challenge
CL	Convolution Layer
MP	Maximum Pooling
AvgPool	Average Pooling
FullConn	Fully Connection Layer
LocResNorm	Local Response Nomalization

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

STRESS DETECTION IN PHYSIOLOGICAL SIGNALS OF NEURODIVERSE PEOPLE USING DEEP LEARNING METHODS

İBRAHİM FURKAN ERÇELEBİ

Department of Computer Engineering
Senior Project

Advisor: Assistant Professor Lecturer Hafiza Irem TURKMEN CILINGIR

Dealing with stress conditions and its relevant changes in the environment is huge task for therapist and psychologist to carry out. Occurrence of stress among neurodivergent people , especially in children , is important field of behavioral sciences due to when and which condition is led to the stress states are hardly recognizable. Preventing these condition beforehand is in scope of recognizing pattern in time series. Stress conditions measurable from body changes like breath rate and sweating. Converting these to numerical data is allow pattern recognition approaches. This approach is done by machine learning or deep learning methods. In this study, stress recognition in normal and neurodivergent children through well known Deep Learning methods called CNN and RNN components and compare their results in success of recognizing stress states.

Keywords: Deep Learning , Stress , Physiology

ÖZET

DERIN ÖĞRENME METODLARIYLA FIZYOLOJIK SINYALLER ÜZERİNDEN NÖROLOJİK FARKLILIKLARI OLAN İNSANLAR ARASINDA STRES TESPİTİ YAPMA

İBRAHİM FURKAN ERÇELEBİ

Bilgisayar Mühendisliği Bölümü
Bitirme Projesi

Danışman: Dr. Öğr. Üyesi Hafiza İrem TÜRKMEN ÇİLİNGİR

Stres koşulları ve çevredeki ilgili değişikliklerle başa çıkmak, terapistler ve psikologlar için gerçekleştirilmesi gereken büyük bir görevdir. Nörodiverjan insanlarda, özellikle çocuklarda stresin ortaya çıkması, stres durumlarının ne zaman ve hangi koşula açığa çıktığının bulunma zorluğu nedeniyle davranış bilimlerinin önemli bir alanıdır. Bu koşulları önceden önlemek, zaman serilerindeki örüntüyü tanıma kapsamındadır. Stres koşulları, nefes alma hızı ve terleme gibi vücut değişikliklerinden ölçülebilir. Bunları sayısal verilere dönüştürmek, örüntü tanıma yaklaşımlarına olanak tanımaktadır. Bu yaklaşım, Makine Öğrenmesi veya Derin Öğrenme yöntemleri ile yapılır. Bu çalışmada, CNN ve RNN bileşenleri olarak adlandırılan iyi bilinen Derin Öğrenme yöntemleri aracılığıyla normal ve nörodiverjan çocuklarda stres tanıma ve stres durumlarını tanıma başarısındaki sonuçlarını karşılaştırdık.

Anahtar Kelimeler: Derin Öğrenme, Stres, Fizyoloji

1 INTRODUCTİON

Metabolic responses to the natural and human environment and happenings can be measurable in a meaningful way. One of these responses can be called 'stress' that stems from non-specific demands [1]. Often this is considered to be any unpleasant case, for example, delaying an appointed task, and leads to physiological stimuli like sweating as a form of stress . Another case with more stimuli generated is the body reaction of automotive drivers going at high speed, which would be an increasing heart rate, rapid changes in skin conductance, and respiration, etc... [2]. These kinds of changes in stress states in specific organs correlate to the sympathetic nervous system. This system generally activates in fight or flight states.

Foundings in handmade recognition of stress states are useful in many human-to-human interaction tasks like therapeutic treatment of neurodiverse peoples. Neurodiverse childrens have detailed procedures of the treatments due to generally giving importance to diagnose and treating in the early stage of mental developments. In these procedures , one of the challenging situations in executing procedures is determining the emotional stability of patients when responding to specific questions. Proposed solutions for optimizing recognition of emotional changes in therapeutic treatments involve making automatic detection by human-computer interaction techniques with machine learning methods. Computer interactions in therapies constitute simulations and detectors that collect input from several signals from specific physiological responses. In the treatment process among children , interactive game systems with specific sensors are favourable for treatments and researches.

Machine and Deep Learning methods are a huge span of usage in every Pattern Recognition areas. These include in medical applications such as detecting lung or brain cancer in early stages. When it comes to recognition of stress , many studies have been done around this subject. Even in this study [3] , a framework model is proposed solely for detecting stress.

1.1 Purpose

Although human emotional situations and stress states can be recognizable by facial expressions and self-report speech recognition , these results can lead to false conclusions for actual results [4]. To obtain more accurate results in the recognition of stress , EDA , BVP, ST, and ACC physiological metrics in time series are processed in this study. The referenced study [5] also used these data, but Machine Learning methods were used. In contrast, Deep Learning Methods (CNN, RNN) are included, and one customized model will be tested as an alternative solution in recognizing stress among neurodiverse children by grouping them as diagnosed names.

2
LITERATURE REVIEW

Stress recognition processes have been studied under numerous topics , tools and methods. Some of them use already collected datasets but different features and different Machine Learning and Deep Learning methods. Summary of information is seen in 2.1.

Generally , data collection process is done by wearable objects attached to the thoracic and wrist and hand areas. The reason for selecting two areas is that significantly specific metrics for recognizing stress are located there. For example , physiological signals which are ECG ,EMG , EDA , RESP , ST from thoracic and wrist in simulated experiments for pilots are collected in this study [6]. In a simulated environment , pilots use different maneuvers and score their experiences by level of stress while sensors were collected for different time-spans . In another experiment only hand skin originated signals which are GSR and PPG are collected from automotive drivers in this study [7]. In this setup drivers for both simulated and real-time environments and different situations labeled as stress level while collecting signals from three numbers of devices connected by bluetooth. These device names are body-worn clip-on Nonin Pulse Oximeter , Abdominal Respiration belt and GSR Velcro electrodes . In this study [8], the experiment was constructed at a job interview of virtual reality and questions responded by autistic adults while sensors in wristbands were collecting physiological signals such as EDA , PPG , ITH and acceleration . Given a time series of each session sample , labeling stress states results are extracted from stress scores in external interview videos.

Preprocessing and feature selection methods in these studies are varied. Aside from generally used preprocessing techniques like normalization and filtering , procedures of which algorithmic calculation used depend on specific physiological signals . For GSR and EDA signals, the nature of these physiological signals allow decomposition into tonic and phasic and evaluate peak properties (number of positive and negative peak points , their distances) [7] . For PPG , its preprocess steps are different from GSR due to the nature of it. After preprocessing , a number of statistical features are obtained from signals by different formulations. These formulations are as follows in study [7]: Mean , Signal energy , Time duration , Bandwidth , Time-bandwidth product and Dimensionality . In a previously cited study these formulas are used for both GSR and PPG signals and provide multiple features. In addition, PPG signals have a sub-property called HRV meaningful for evaluating stress state. In the same study , this property was extracted from five numbers of features by lomb periodogram [9]. With the same signals and additional ones as ST and Acceleration in this study [8] , from EDA signal tonic component called skin conductance level (SCL) and phasic component called skin conductance responses (SCRs). Two new properties are extracted from statistical features as mean , peak and standard deviation values. Acceleration is changing in three dimensions so that mean and standard deviation of three axes are extracted as additional features in the same study. On a different side, handcrafted features can be extracted from each ECG, EMG, EDA and RESP signals between sliding windows of time series in this study [6]. There are twenty-five handcrafted features and some of them are named as : normal-to-normal (NN) intervals, the standard deviation of NN intervals (SDNN), root mean square of successive differences(RMSSD) , etc... .

Classification methods in these studies fall into Machine Learning and Deep Learning categories. Although Machine Learning usage outweighs Deep Learning in these studies , it is worth mentioning how Deep Learning is used in classifications. Used Machine Learning methods with generally known ones are : Support Vector Machine (SVM), K-Nearest Neighbourhood (kNN), Random Forest (RF) , Decision Tree (DT) , Logistic Regression (LR), Na¨ıve Bayes (NB), Artificial Neural Networks (ANN) , Gradient Boosting Regression , Elastic Net Regression , Bayesian Ridge Regression , XGBoost Regressor , Support Vector Regression , Feed-Forward Neural Networks(Single-Layer /Mult,-Layer Perception , Cascade Forward Backpropagation , Feed Forward Distributed Time-Delay) and Dynamic or Recurrent Neural Networks (Elman Back-Propagation , Layer Recurrent , Non-linear Autoregressive Networks With Exogenous Inputs) [7] , [8] , [5]. Used Deep Learning methods are : Transformer-Based Convolutional Neural Network (TB-CNN) , Modified AlexNet , Modified ResNet18 and Light RestNet [6].

Table 2.1 Related Literature Analysis

Studies	Data Source	Preprocessing	Feature Extraction	Classification
[7]	Collecting signals from drivers in specific conditions	GSR: Peak Detection Algorithm PPG: Remove Motion Artefact - Heart Rate Extraction	GSR : 8 number of syntactic features PPG : 5 number of syntactic features , derived HRV feaures and five number of HRV spectral features	Gradient Boosting Regression , Elastic Net Regression , Bayesian Ridge Regression , XGBoost Regressor , Support Vector Regression
[6]	Collecting signals from pilot in specific maneuvers	Noise Reduction , High-Low Pass Filtering	handcrafted features for ECG, EMG, EDA, BP in concatenated as 25 number of features in total	Feed-Forward Neural Networks: Single-Layer /Mult,-Layer Perception , Cascade Forward Backpropagation , Feed Forward Distributed Time-Delay Dynamic or Recurrent Neural Networks: Elman Back-Propagation , Layer Recurrent , Non-linear Autoregressive Networks With Exogenous Inputs
[8]	Collecting signals when interview session of autistic adults	EDA: High-Low Pass Filtering	EDA : 6 number of feature PPG : 6 number of features Temperature: 2 number of features Acceleration : 6 number of features	Transformer-Based Convolutional Neural Network (TB-CNN) , Modified AlexNet , Modified ResNet18 , Ligth RestNet
[5]	Collecting signals from autistic children when playing interactive game	BVP : Remove outliers and fill missing values EDA: Filtering and normalization	BVP : 11 number of features EDA :28 number of features ST: 8 number of features	SVM , kNN, RF ,DT , LR , NB , ANN

3
SYSTEM ANALYSIS AND FEASIBILITY

System setup and requirements are raw dataset , proposed and candidate Deep Learning methods , model output performance / comparisons , software and utility modules that run models in efficient hardware. More details are mentioned in this section.

3.1 System Analysis

As in figure 3.1 , the already collected dataset will be used . Experiments of this study constitute outer inputs to an already collected dataset as three different physiological signals as EDA , BVP ,ST and ACC in time series. Dataset formation is described in 3.1. Required features will be extracted in the process of the experiment workflow . Expected input to models is hyper parameters like batch count , number of layers , dropout and name of optimizer and loss functions. Expected output is a set of values that contains classification results as stress / no stress , performance metric of proposed and other candidates models for comparison for giving clear results. Performance metrics are sensitivity , specificity , precision , recall . Also this flow can be recurrent in the classification process until only a specific number of counts for selecting optimized sets of features. These features in this study are the ones used in this study [5].

3.2 Feasibility

Before implementing the experiment , it is crucial to meet resource requirements on the technical side. Similar studies mentioned previously spend financial expenditures for real-time data like . Even though datasets have no samples collected from the setup for experiment of this study , preprocessing and evaluating models require special software and hardware modules. In this section, more details are introduced.

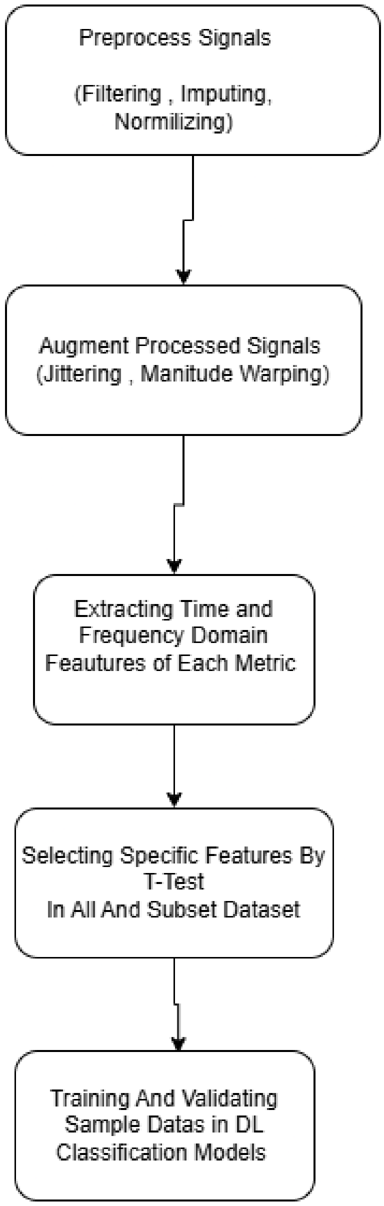

Figure 3.1 Workflow Of Mechanism

3.2.1 Technical Feasibility

The experiment will run on Python language due to support for a wide range of libraries and modules related to data-science. There are other languages like R that are widely used in data science fields .Yet in terms of easy to learn languages , Python outweighs the others. So that selected operating system must support Python interpreter and virtual environment components. Building models with flexible configurations , training / testing models with faster processes and less memory usage are important for customizing models. These criterias are met in PyTorch. For this library , the recommended operating system is Windows with minimum Python version is 3.9 and a package manager such as pip or apt [10]. In the context of this study , image data is not included , so CUDA module is optional for faster processing.

Table 3.1 Dataset Procedures And Abbrevations

Hyper Parameters	Target Balance	
	No Stress // Stress = 1 (BLNC)	No Stress // Stress = 4 (NON $_BLNC$)
Only Original Data ONL $_ORG$	ALL Diagnosed(ALL) , Only Typical Developed(TD)	ALL Diagnosed(ALL) , Only Typical Developed(TD)
Original And Augmented Separated In Test (ORG $_AUG_SEP$)	ALL Diagnosed(ALL) , Only Typical Developed(TD)	ALL Diagnosed(ALL) , Only Typical Developed(TD)
Original And Augmented Combined (ORG $_AUG_MIX$)	ALL Diagnosed(ALL) , Only Typical Developed(TD)	ALL Diagnosed(ALL) , Only Typical Developed(TD)

3.2.1.1 Hardware Feasibility

Mentioned software modules for running experiments are required in specific hardware requirements. List of minimum requirements for effectively running processes are following :

- CPU Unit: with 4 core , base frequency 2300MHz , maximum frequency 1.80 GHz.
- Memory Unit: 16 GB or above.
- Disk Unit : Minimum size of free space is 20 GB for installing and allocating Python and Pythorch. Additional free size must be at least 100MB because dataset size is 78MB.

In case of CUDA is selected , One of Nvidia produced GPU's must be included.

3.2.2 Economic Feasibility

Required all hardware components as whole or part by part have costed in financial term. Depending on the marketplace of available products and currency values in the country , buying these products has a low-medium level of cost. As alternative instead of buying products as whole , third party cloud-based Saas systems for computing resources. One of famous Saas system is Google Colab where the experiment is executed [11]. Colab has other subscription options for more computation resource usage [12]. Optionally in case of process would more slower , other subscription will be use subscriptions. In software side , possible way of access dataset other than institutional permission is purchase the study or subscribe a membership.

3.2.3 Schedule Feasibility

As in figure 3.2 , diagram segmented start and end day of one week . Timeline planned in case one person is the executive of an experiment . Dataset inspection and preprocess steps intersect during one week because it is more time efficient that when at least half of inspection completes , preprocessing starts. But before feature extraction , all preprocessings must complete due to target statistical features being calculated by all sample data. If one sample preprocess is not complete , overall feature values will slightly change. Training classification models and selecting subset of features require a long time so this task is given relatively more weeks.

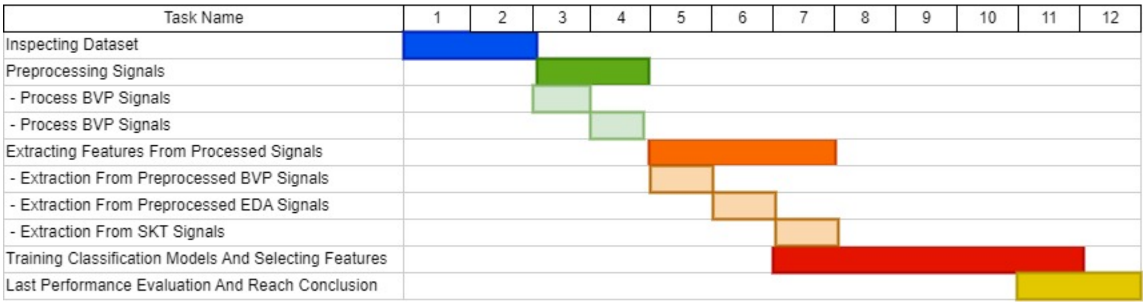

Figure 3.2 Gantt Diagram, Number Indicate Number of Week , Start From October 14th

3.2.4 Legal Feasibility

Dataset that is used in the study required institutional access permission or buy licence for published dataset [13]. We use our university access permission in this study for cost savings. Other than that , no legal restrictions in software and hardware are not found.

4
SYSTEM DESIGN

In the previous chapter , the dataset and preprocess step were defined with less information. In this chapter , the metrics in the dataset , preprocess methods for correct raw data , features from preprocessed data and classification models trained by feature dataset are detailed .

4.1 Material and Dataset

Three signals are collected in a specific procedure. Previously collected dates in the study [5] focused on neurodivergent children. Collection setup in seven minute duration with two separate sessions . In each session , childrens play virtual games called CatchAPet and LeapBall in the scope of E-Therapy. In the CaychAPet game, the user must contact a rabbit that moves in different holes. In the LeapBall game, the user must grab the ball and throw the right basket annotated in the game. When the right actions are done , users gain scores. The hand gestures are important in these games , so movement detectors are attached on the wrists accordingly appointed by therapists. When a child plays these games , measuring physiological metrics are done by another wristband detector called Empatica E4 . This device used many studies related to stress and emotion analysis [14–16]. In that study , physiological metrics are gathered from twenty five children with four different categories , typical development - intellectual disability - dyslexia - obstetric brachial plexus injuries. These are important signals called BVP , EDA and SKT. Generally bio signals are useful in detecting stress state [17]. Specially ECG and EDA signals widely used are promising results in analyzing [18, 19]. More detal about BVP and ECG relationship?. The BVP signal is collected by sampling with 64Hz. The EDA and SKT signals are collected by sampling with 4Hz. At the same time , children's facial expressions that are recorded by camera are evaluated by three experts. These experts diagnose from facial expression in every ten second time stamps as stress / no stress and reaction / no-reaction. Diagnoses about stress labeled as output category for this study.

4.1.1 Preprocess And Feature Extraction

Two signals such as BVP and EDA are preprocceded due to the presence of noisy data and outlier data in the time series of signals. To remove noisy data from signals , SOCT filter for BVP and FOSG filter for EDA are implemented [20]. SOCT filter use below equation 4.1 given angular frequency ω and $s = j\omega$ in n-th order T_n of with static ϵ . The FOSG filter defined below equation 4.2 given i is index at a set of independent variables , j is index at a set of observable variables , C_j is convolution coefficient . In BVP signals , many samples fall into outlier status found by z-score that is bigger than or lower than the negative version. After removing these values , empty spaces in the time series are filled by KNN imputer [21]. Variance of levels between samples of signal time series is very high due to relativity in individual representation of emotions [22]. Elimination to that factor is normalization of BVP and EDA signals in between 0 and 100 values. SKT signals will be used as raw values.

$$H(s) = \frac{1}{\sqrt{1 + \epsilon^2 T_n^2(\omega)}} \tag{4.1}$$

$$Y_j = \frac{1}{2m+1} \sum_{i=-m}^{m} C_i y_{i+j} \tag{4.2}$$

Prepared dataset is relatively small for training and testing performance of DL models in this study. In aim of adding more sample data and improving results , data augmentation methods used in time series datasets are implemented in extracted features [23]. First one is jittering with Gaussian noise . This method uses the Gaussian noise formula 4.3 given x variable , σ_2 parameter in N function . The formula of N function defined in this 4.4 formula with extra μ variable . These noises jitter upon signal waves. Second augmentation method is Slice and shuffle . In this step , signal waves are slicing in specific points of time and replacing random locations in the time series . This method only applied for the features that were invariant in time series . Third augmentation method is Magnitude Warping . This method changes the magnitude of signal waves with a cubic spline curve . The Fourth augmentation method is Window Warping . Like the third one , this changes the window of the signal wave in the fixed duration of time series. As a result , some of them have larger windows and others have smaller window sizes. These augmentations are implemented with the help of the tsgm tool [24].

$$y = x + \mathcal{N}(0, \sigma^2) \tag{4.3}$$

$$f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{4.4}$$

In time series values , many statistical features can be extractable. In three signals , Following features are common among three signals and will be used in the experiment : mean , median , minimum value(min), maximum value(max), kurtness (kurt), skewness (skew), variation (var), quantile 0.05 (Q_05), quantile 0.25 (Q_25), quantile 0.75 (Q_75) and quantile 0.95 (Q_95) . For only EDA signals , two different components are extracted as Tonic EDA (SCL) and Tonic EDA (SCR). The measurement of SCL is usually taken under the processes of multifaceted attentional [25]. And The measurement of SCL is usually taken under the physiological alertness citeboucsein2012electrodermal. Common statistical features are extracted from both of them. For only BVP signals , signals of blood pulses are also related as heart beat signals [26]. So that HRV related features are also extractable from BVP signals and these are : SDNN, RMSSD, PNN20, PNN50. BVP , EDA , SCL and SCR signals have these common features : number of zero crossing (num_zero_cross) , number of positive (num_peak_positive) and negative (num_peak_negative) peak values. The total of 71 features are extracted for this experiment and summary is shown in 4.1 table.

Table 4.1 Table Of Extracted Features From Signals

Signal Groups	Extracted Features
BVP , EDA , SCL , SCR , SKT	mean , median , min , max , kurt , skew, var , Q_05 , Q_25 , Q_75 , Q_95
BVP , EDA , SCL , SCR	num_zero_cross , num_peak_positive , num_peak_negative
BVP	SDNN, RMSSD, PNN20, PNN50

These features are selected by p-values of T-Test function.

4.2 Method

The Deep Learning methods and their models built from them constitute multiple sets of layers which contain a set of nodes with their weights . General process pipeline starts from input nodes as input. Then goes to at least one hidden layer . Finally, the result from nodes of every last hidden layer goes to the output layer with activation function and gives the final value in between 0 and 1. This process is called forward propagation. Another important part of learning is called backward propagation. In this process the loss value is calculated from the final result and expected results. With loss value , the optimizer function will change weights of nodes in every iteration.

In more complex input data such as matrix representing multiple dimensions, processing reducing volume of data without losing context is allowed in CNN models [27]. These models have two common components called Convolution Layer(CL) and Pooling Layer. Convolution Layers get a specific part of matrix value with kernels and convolve as one data segment to the next data matrix. Pooling Layer reduces convoluted matrix values into smaller and abstracted data matrix. CNN models have a wide range of applications and one of them is time series forecasting **??**. Analyzing signals in time series data in CNN models is the main purpose of the experiment . Their input is an augmented set of statistical and physiological features. Input shape will be in (71, 1) and output size will be reshaped to give two channel response for possibilities of stress/no-stress. 80 percent of the whole augmented dataset is separated for training models and the rest of the dataset samples will be used as testing model performance.The hyper parameters for the employment of these models are separated into three scenarios as seen in 4.2. Each scenario is only for that parameter , not strict to all . It means that in testing , all three hyper parameters use different combinations of scenarios.

Table 4.2 Hyper Parameters and Scenarios

	Scenario No		
Hyper Parameters	No 1	No 2	No 3
Batch Size	5	10	10
Number Of Epochs	100	200	400

The one known models will be employed. That model name is AlexNet [28]. This model is designed for competition of ILSVRC-2010 . The model architecture is designed in eight batch of layers. The first batch is CL which takes input with 11x11 kernel size , and then Response Normalization Layer (RNL) gives output to the second Layer where MP with 3 pixel kernel size is in between. The second Layer has a CL with 5x5 kernel size and then Response Normalization Layer (RNL) gives output to the second batch where MP with 3 pixel kernel size is in between. The third , fourth and fifth batch has only CL with 3x3 kernel size and there is no pooling mechanism , only direct connection. Output of fifth batch under MP with 3 kernel size connected to the dense layer . The sixth and seventh batches contain dense layers where connection has been cut out. For both layers , half of the connections drop out and have 4096 neurons . In the eighth batch, the activation layer filters out and gives output with the desired number of classes. In this study , our modified AlexNet model for training is seen in image 4.1

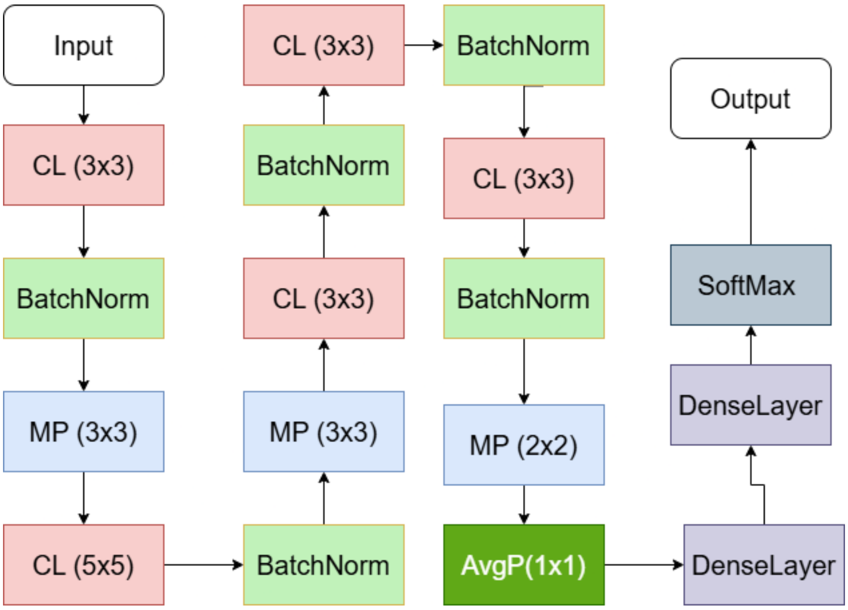

Figure 4.1 Modified AlexNet Model Diagram

From this far , all models have forward propagation and backpropagation. This type of flow has one key issue: input depends on strictly constrained features in one-way process flow without capturing previous flows. In case of processing time or space dependent features like strings of paragraphs or time series of features , it is crucial to capture previous results of flows. Recurrent Neural Networks(RNN) give a solution by capturing previous states of outputs in specific structures called hidden states. After each output from the RNN layer, hidden state captures and concatenates activation function with parameters of previous weights and bias values in the next training flow. Different from the rest of the DN flows, backpropagation is implemented through all layers and its hidden states. Other modern version of RNN is LSTM which means Long-Short Term Memory . Our study is conducted upon this model Tell details of RNN and the position of LSTM among them. [29]

In this study. Four types of LSTM architecture will be tested. The first one is a simple LSTM structure with a number of 2 layers, and 0.4 dropout and hidden size is half of input. The second one is the same with the difference of hidden size is two times larger than input. The third one is same number ber of hidden size as first one but with 4 layers. The fourth one is same number of hidden size as second one but with 8 layers. The output of the LSTM goes through the activation layer with two output.

5
EXPERİMENT RESULTS

In this episode , all process flows start from the collecting data to training have ended to building a model for detecting stress status from children in variations of processed datasets. At the same time , collecting train results gives insight from model trainings as seen in images 5.1 , 5.2, 5.3, 5.4 , 5.5 , 5.6 , 5.7, 5.8, 5.9, 5.10 , 5.11 , 5.12 .

5.1 Performance Analysis

In below figures, general tendency in the train process of the dataset with only original samples or separated from augmented ones, is very disruptive and constrained in specific range of rate. One of the possible reasons is that extracting features in these signals gives weak indications of the stress states. Although in this study, raw signals are not used to train models, [30] the study compares the raw signal dataset to the feature extracted dataset, and the second one gives a higher rate of accuracy.

When comparing balanced and non-balanced samples among original samples , in accord with the existence of augmented data, the accuracy and loss trend is changing from a more unstable to a stable development. This development also reflects on the confusion matrix .

5.2 Comparative Analysis

In comparison of AlexNet and LSTM , changes of accuracy and other metrics are dratisitcally different lines. Both of them give low level of accuracy development in original and augmentation separated datasets , LSTM increases its performance in non-balanced datasets relatively. This comes from being able to keep previous states and evaluate next states. That advantage gives LSTM the upper hand in contextual recognition. But it requires larger epoch size to attain to the point where AlexNet arrives more quickly.

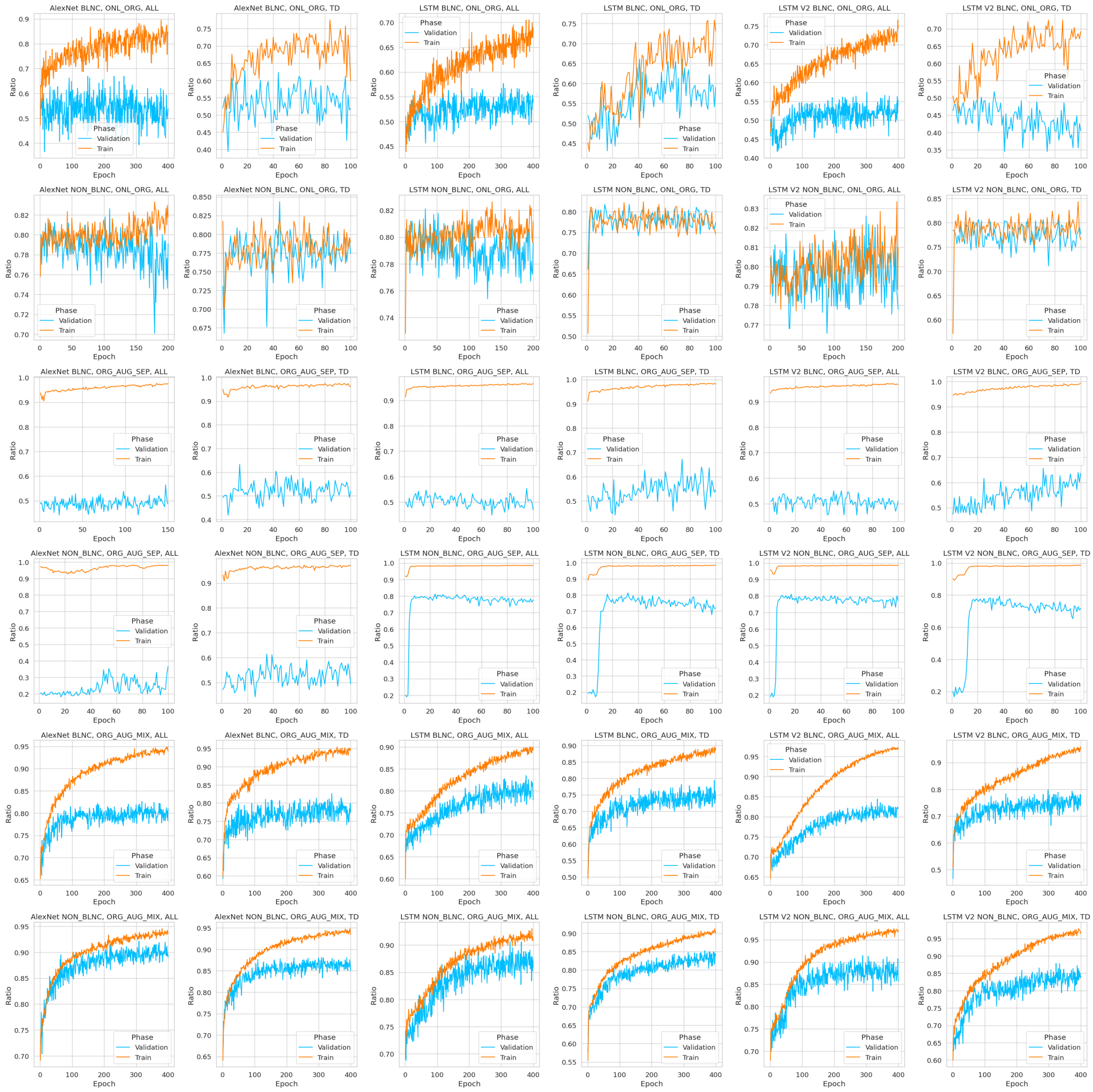

Figure 5.1 Train Accuracy Process Results By Model And Procedure Codes

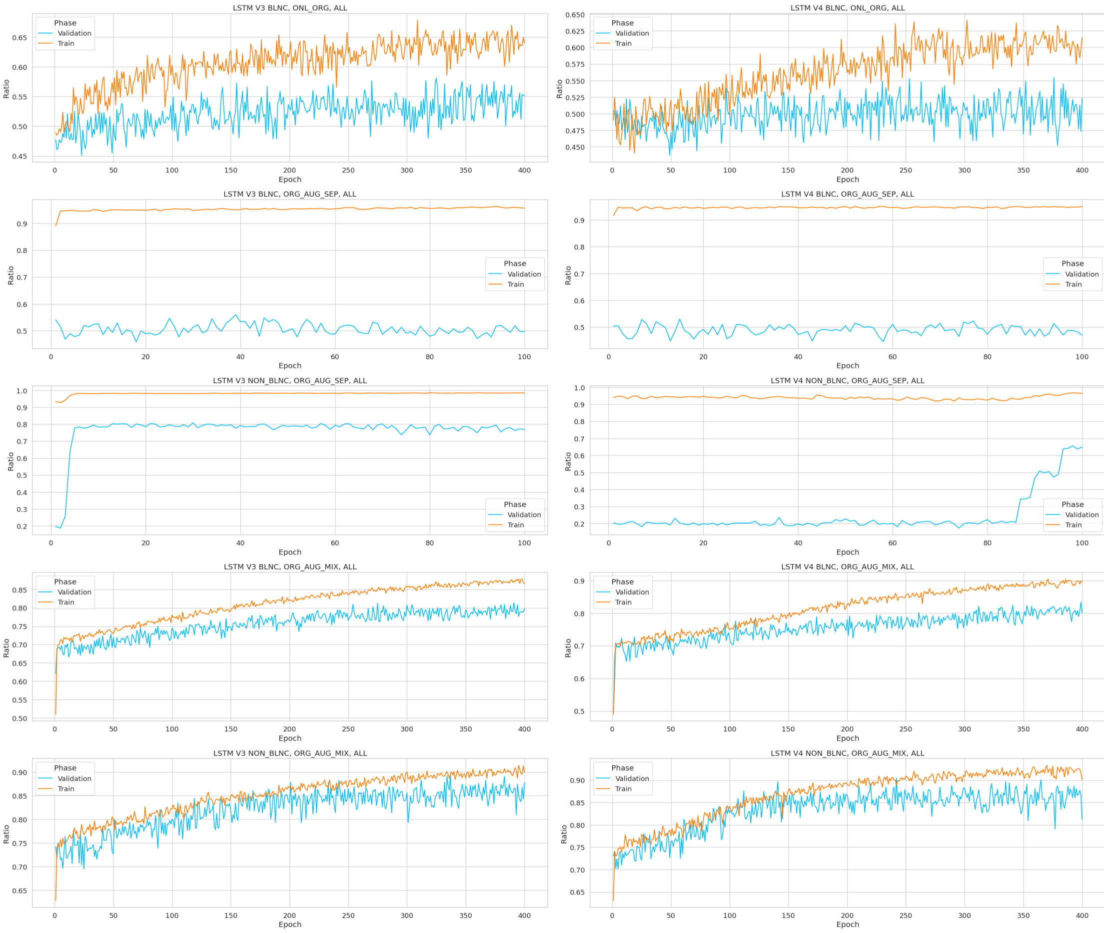

Figure 5.2 Train Accuracy Process Results Of LSTM V3 and V4 Model With Procedure Codes

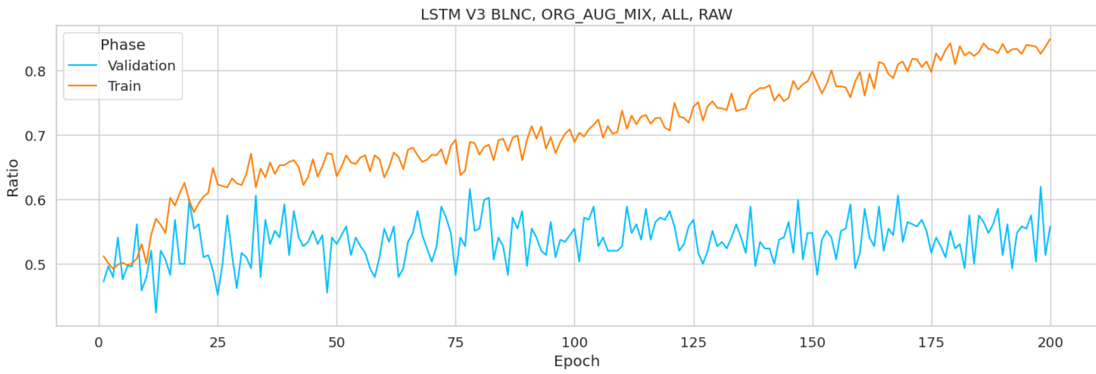

Figure 5.3 Train Accuracy Process Results Of LSTM V3 With Raw Signal Dataset

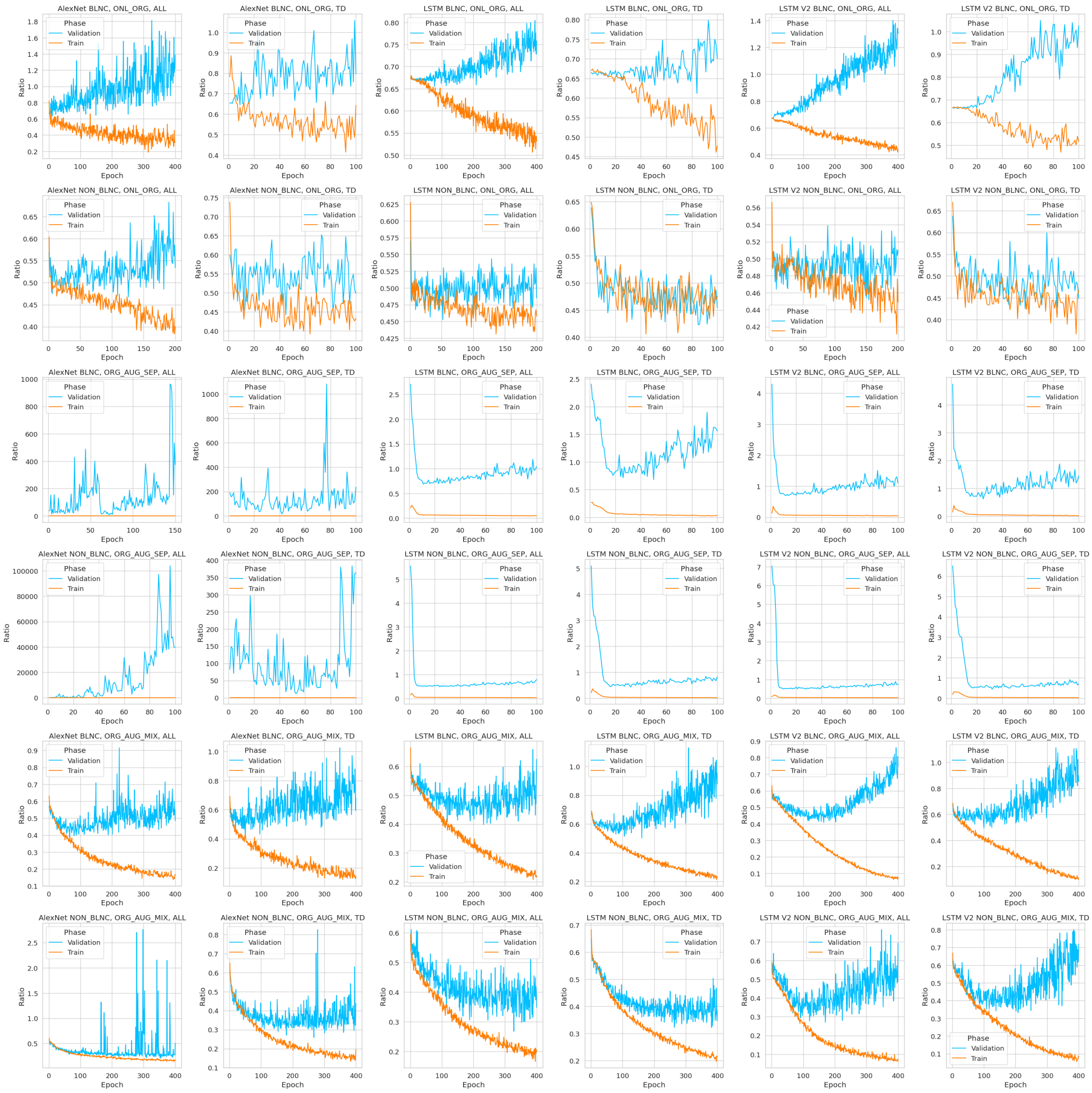

Figure 5.4 Train Loss Process Results By Model And Procedure Codes

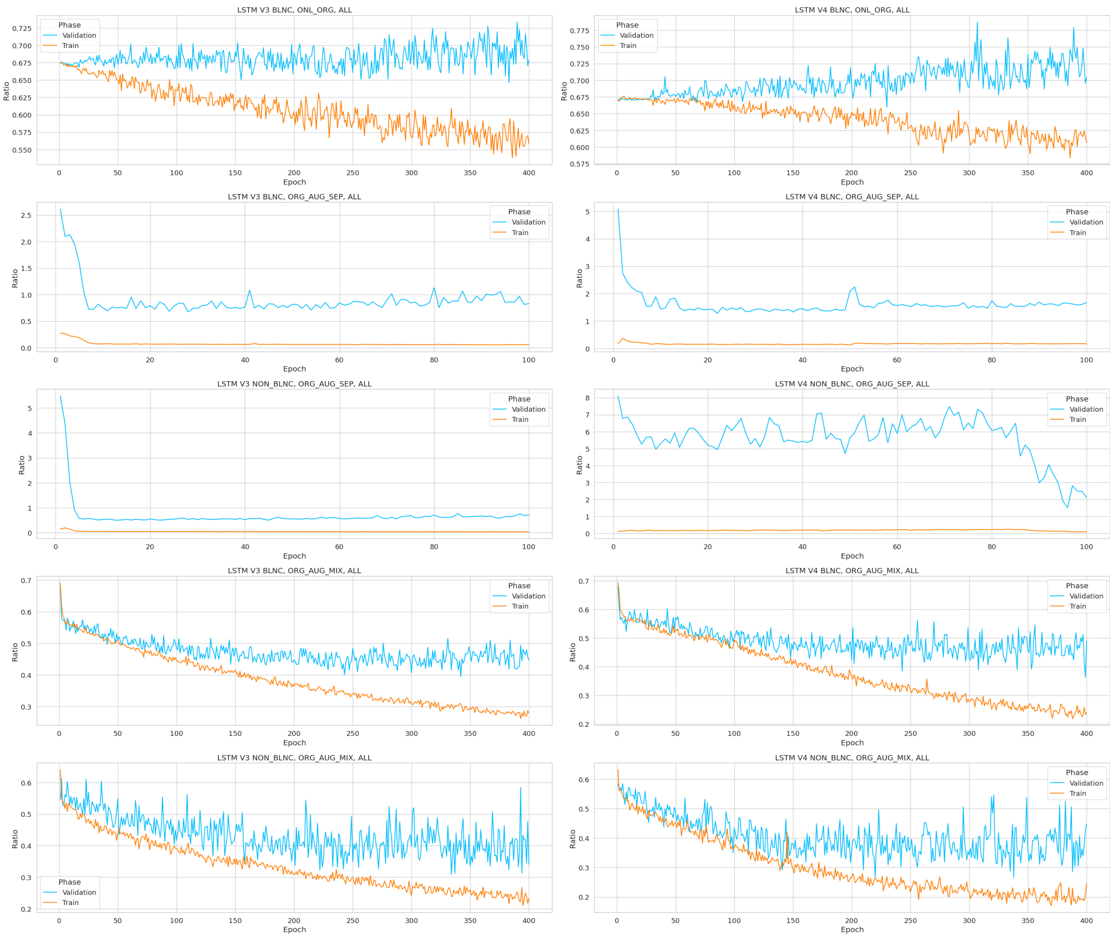

Figure 5.5 Train Loss Process Results Of LSTM V3 and V4 Model With Procedure Codes

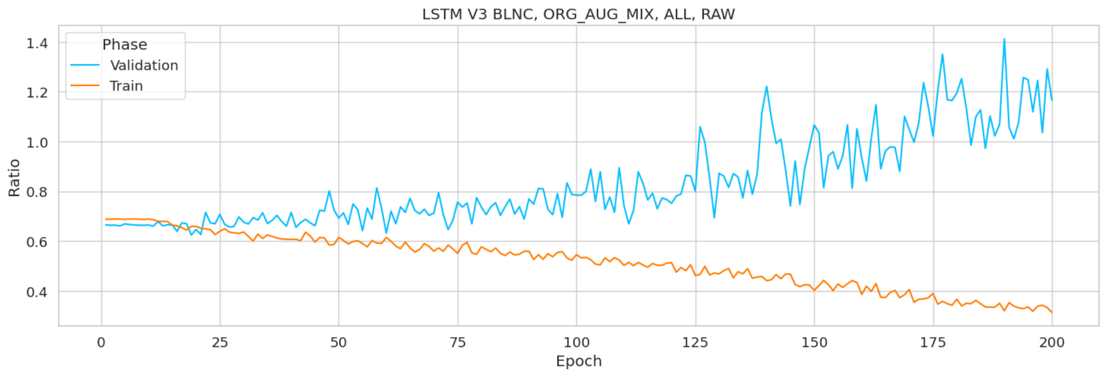

Figure 5.6 Train Loss Process Results Of LSTM V3 With Raw Signal Dataset

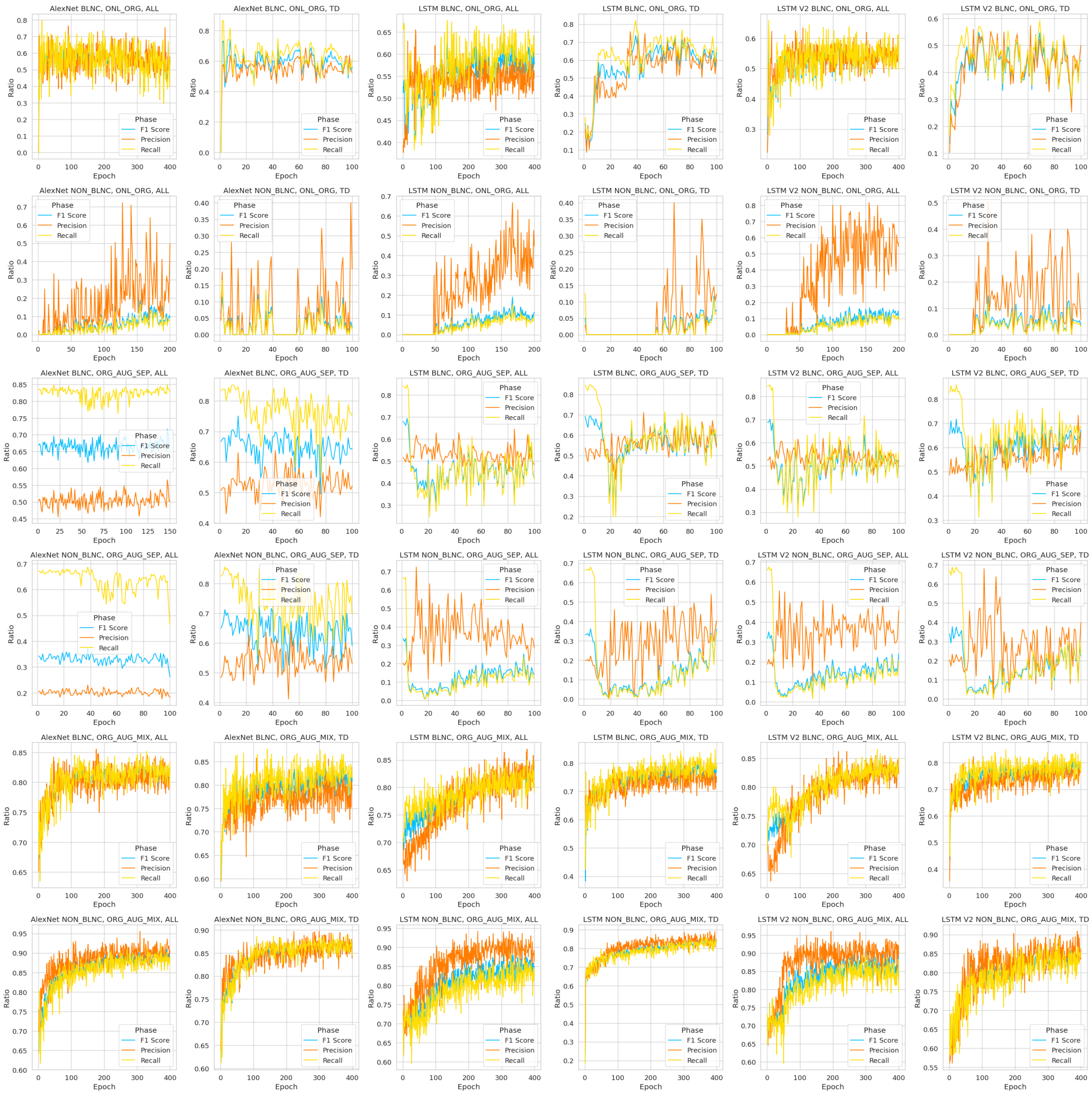

Figure 5.7 Validation Precision , F1 Score and Recall Process Results By Model And Procedure Codes

Figure 5.8 Validation Precision , F1 Score and Recall Process Results Of LSTM V3 and V4 Model With Procedure Codes

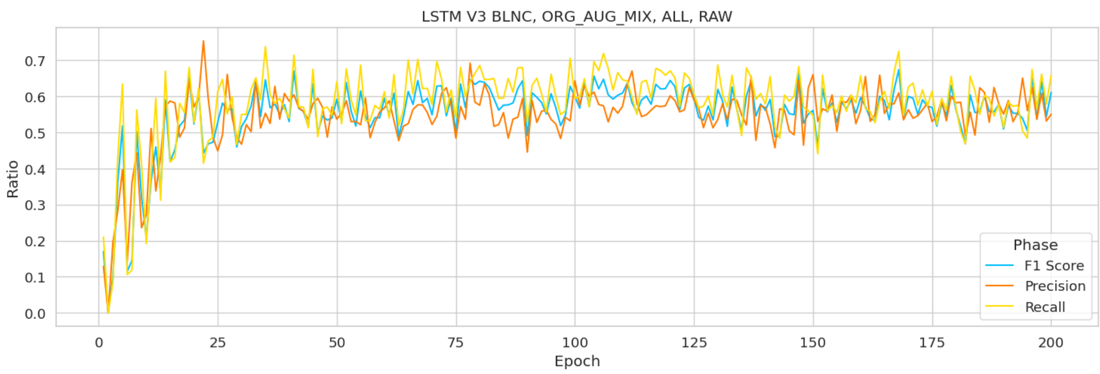

Figure 5.9 Validation Precision , F1 Score and Recall Process Results Of LSTM V3 With Raw Signal Dataset

Figure 5.10 Confusion Matrix Results By Model And Procedure Codes

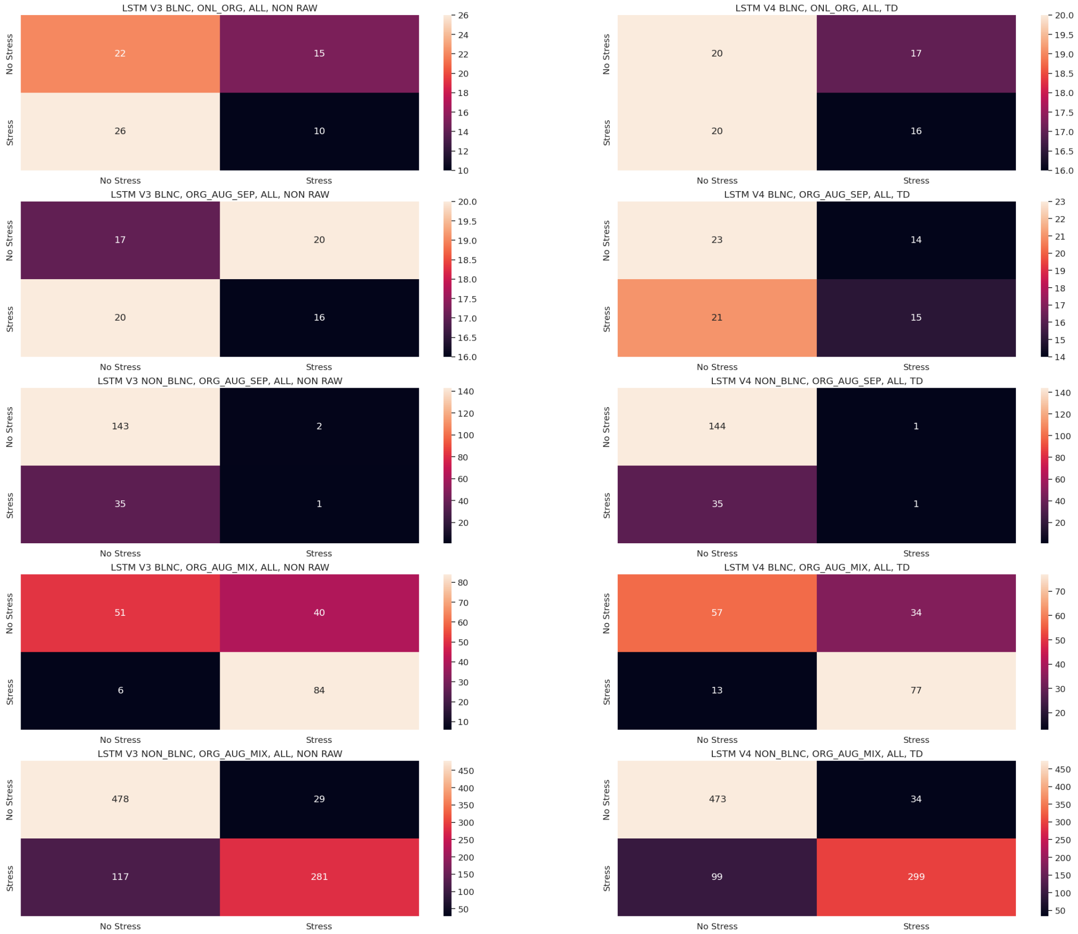

Figure 5.11 Confusion Matrix Results Of LSTM V3 and V4 Model With Procedure Codes

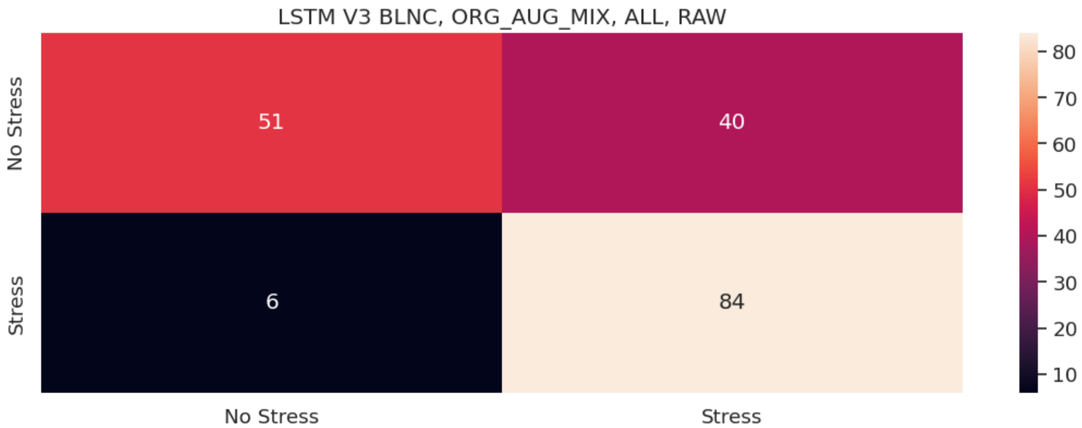

Figure 5.12 Confusion Matrix Results Of LSTM V3 With Raw Signal Dataset

6
DISCUSSION AND CONCLUSION

Goal of this study is detecting stress from physiological signals metrics among normal and neurodivergent children. After preprocessing raw signals by filtering and imputing removed data, different augmentations like Jittering and Warping implemented. Number of time domain features are extracted from both original and augmented samples. Execute models in with different parameters and different versions of processed datasets and obtained variety of train and test results.

In general , dataset class distribution in datasets gives a clear position in the development of models. Models unexpectedly give stable scores in imbalanced datasets and maybe accuracy increases. But testing from original or mixed metrics gives different solutions. Also, removing specific groups like neurodivergent samples changes the calculation of scores.

Recognition studies around stress and emotional states must be carried out with a large collection of data. The importance of sample size and distribution of target state such as stress-no stress are crucial data science analyses. In more advanced analysis , stress studies around neurodivergent children require more data acquisition on physiological metrics.

References

[1] T. A. I. of Strees. "What is stress?" (), [Online]. Available: https://www.stress.org/what-is-stress.

[2] M. Singh and A. B. Queyam, "Correlation between physiological parameters of automobile drivers and traffic conditions," *International Journal of Electronics Engineering*, vol. 5, no. 2, pp. 6–12, 2013.

[3] B. Hwang, J. You, T. Vaessen, I. Myin-Germeys, C. Park, and B.-T. Zhang, "Deep ecgnet: An optimal deep learning framework for monitoring mental stress using ultra short-term ecg signals," *Telemedicine and e-Health*, vol. 24, no. 10, pp. 753–772, 2018, PMID: 29420125. DOI: 10.1089/tmj.2017.0250. eprint: https://doi.org/10.1089/tmj.2017.0250. [Online]. Available: https://doi.org/10.1089/tmj.2017.0250.

[4] S. K. Khare and V. Bajaj, "Time–frequency representation and convolutional neural network-based emotion recognition," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 2901–2909, 2021. DOI: 10.1109/TNNLS.2020.3008938.

[5] B. Coşkun, D. E. Barkana, İ. Uzun, H. Bostancı, and D. Tarakçı, "Classification of stress and participation using physiological signals of children during serious game-based therapy," in *2023 Medical Technologies Congress (TIPTEKNO)*, IEEE, 2023, pp. 1–4.

[6] Y. Li, K. Li, J. Chen, S. Wang, H. Lu, and D. Wen, "Pilot stress detection through physiological signals using a transformer-based deep learning model," *IEEE Sensors Journal*, vol. 23, no. 11, pp. 11 774–11 784, 2023.

[7] R. R. Singh, S. Conjeti, and R. Banerjee, "A comparative evaluation of neural network classifiers for stress level analysis of automotive drivers using physiological signals," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 740–754, 2013.

[8] M. Migovich, D. Adiani, M. Breen, A. Swanson, T. J. Vogus, and N. Sarkar, "Stress detection of autistic adults during simulated job interviews using a novel physiological dataset and machine learning," *ACM Trans. Access. Comput.*, vol. 17, no. 1, Mar. 2024, ISSN: 1936-7228. DOI: 10.1145/3639709. [Online]. Available: https://doi.org/10.1145/3639709.

[9] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, 2005, Cited by: 1601. DOI: 10.1109/TITS.2005.848368. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-21644474051&doi=10.1109%2fTITS.2005.848368&partnerID=40&md5=9f545000bca7919f87cb1f7e15c0ce1a.

[10] "Start locally | pytorch." (), [Online]. Available: https://pytorch.org/get-started/locally/#windows-prerequisites.

[11] "Colab.google." (), [Online]. Available: https://colab.google.com/.

[12] "Colab subscription services." (), [Online]. Available: https://colab.research.google.com/signup.

[13] "Get institutional access to ieee information." (), [Online]. Available: https://www.ieee.org/publications/subscriptions/subterms/get-access.html.

[14] J. G. Negrão *et al.*, "The child emotion facial expression set: A database for emotion recognition in children," *Frontiers in psychology*, vol. 12, p. 666 245, 2021.

[15] V. Chandra, A. Priyarup, and D. Sethia, "Comparative study of physiological signals from empatica e4 wristband for stress classification," in *Advances in Computing and Data Sciences: 5th International Conference, ICACDS 2021, Nashik, India, April 23–24, 2021, Revised Selected Papers, Part II 5*, Springer, 2021, pp. 218–229.

[16] G. Cosoli, A. Poli, L. Scalise, and S. Spinsante, "Measurement of multimodal physiological signals for stimulation detection by wearable devices," *Measurement*, vol. 184, p. 109 966, 2021.

[17] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on psychological stress detection using biosignals," *IEEE transactions on affective computing*, vol. 13, no. 1, pp. 440–460, 2019.

[18] K. Nkurikiyeyezu, A. Yokokubo, and G. Lopez, "The effect of person-specific biometrics in improving generic stress predictive models. arxiv 2019," *arXiv preprint arXiv:1910.01770*, 2019.

[19] S. Sriramprakash, V. D. Prasanna, and O. R. Murthy, "Stress detection in working people," *Procedia computer science*, vol. 115, pp. 359–366, 2017.

[20] M. Sevil *et al.*, "Discrimination of simultaneous psychological and physical stressors using wristband biosignals," *Computer methods and programs in biomedicine*, vol. 199, p. 105 898, 2021.

[21] M. Kuhn, *Applied predictive modeling*, 2013. [Online]. Available: https://books.google.com.tr/books?id=3ZqtzQEACAAJ.

[22] B. Zhao, Z. Wang, Z. Yu, and B. Guo, *Emotionsense: Emotion recognition based on wearable wristband. in 2018 ieee smartworld, ubiquitous intelligence & computing, advanced & trusted computing, scalable computing & communications, cloud & big data computing, internet of people and smart city innovation (smartworld/scalcom/uic/atc/cbdcom/iop/sci). 346–355*, 2018.

[23] G. Iglesias, E. Talavera, Á. González-Prieto, A. Mozo, and S. Gómez-Canaval, "Data augmentation techniques in time series domain: A survey and taxonomy," *Neural Computing and Applications*, vol. 35, no. 14, pp. 10 123–10 145, 2023.

[24] A. Nikitin, L. Iannucci, and S. Kaski, "Tsgm: A flexible framework for generative modeling of synthetic time series," *arXiv preprint arXiv:2305.11567*, 2023.

[25] J. T. Cacioppo, L. G. Tassinary, and G. Berntson, *Handbook of psychophysiology*. Cambridge university press, 2007. [Online]. Available: https://books.google.com.tr/books?id=E7hRKwVBXb4C&lpg=PR9&ots=VJeCOUmCCV&dq=Handbook%20of%20psychophysiology&lr&hl=tr&pg=PR9#v=onepage&q=Handbook%20of%20psychophysiology&f=false.

[26] M. Nitzan, S. Turivnenko, A. Milston, A. Babchenko, and Y. Mahler, “Low-frequency variability in the blood volume and in the blood volume pulse measured by photoplethysmography,” *Journal of biomedical optics*, vol. 1, no. 2, pp. 223–229, 1996.

[27] J. Wu, “Introduction to convolutional neural networks,” *National Key Lab for Novel Software Technology. Nanjing University. China*, vol. 5, no. 23, p. 495, 2017.

[28] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25, Curran Associates, Inc., 2012. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

[29] S. Zargar, “Introduction to sequence learning models: Rnn, lstm, gru,” *Department of Mechanical and Aerospace Engineering, North Carolina State University*, 2021.

[30] K. Motaman, K. Alipour, B. Tarvirdizadeh, and M. Ghamari, “A stress detection model based on lstm network using solely raw ppg signals,” in *2022 10th RSI International Conference on Robotics and Mechatronics (ICRoM)*, IEEE, 2022, pp. 485–490.

Curriculum Vitae

FIRST MEMBER

Name-Surname: İBRAHİM FURKAN ERÇELEBİ
Birthdate and Place of Birth: 10.12.1997, İstanbul
E-mail: furkan.ercelebi@std.yildiz.edu.tr
Phone: 0541 513 74 60
Practical Training:

Project System Informations

System and Software: Windows, Python
Required RAM: 2GB
Required Disk: 256MB