IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DETECTING STRESS THROUGH 2D ECG IMAGES USING PRETRAINED MODELS, TRANSFER LEARNING AND MODEL COMPRESSION TECHNIQUES.

K.SHIVAKUMAR[1], Dr. K.SANTHISREE [2]

[*1]Student, M.Tech (Data Science), Department of Information Technology, JNTU Hyderabad, Hyderabad,India

[*2]Head of Department, Professor of Information Technology, Department of Information Technology, JNTU Hyderabad, Hyderabad, India

ABSTRACT: Stress is a major part of our everyday life, associated with most activities we perform on a daily basis and if we are not careful about managing stress, it can have a detrimental impact on our health. Despite recent advances in this domain, HRV analysis is still the most common method to detect stress, and although the results that have been produced are admirable, feature extraction is complicated and time consuming. We propose an algorithm to convert 1D (dimensional) ECG data from WESAD (wearable stress and affect detection dataset) into 2D ECG images, which are representative of stress/not stress. It does not require time consuming processes such as feature extraction and filtering. We utilize transfer learning to obtain competitive results. We also demonstrate that model compression techniques can significantly reduce the computational size of the algorithms, without sacrificing much of the performance, as evident from a classification accuracy of 90.62% using the quantization technique. Results substantiate the effectiveness of our proposed method and empirically demonstrates the potential of deep learning algorithms for edge computing and mobile applications, which utilizes low performing hardware.

Keywords – HRV-heart rate value ,ECG- Electrocardiogram,wesad,Stress,Feature Extraction.

1. INTRODUCTION

Whether it is due human activities such as work, relationship or finance, stress is inevitably part of our daily life, and prolonged stress can negatively impact a person's immune system, which can lead to several health problems and aggravate cardiac diseases such as myocardial infarction, hypertension and diabetes .According to a American Psychology Association (APA) report, 75% of Americans experience at least one symptom of stress every month, causing people to miss work and leading to the loss of billions of dollars annually . Observing the on-going trend, it is imperative to detect stress early, in order to prevent chronic outcomes. Heart rate variability (HRV) analysis has been the most common and prevalent method to analyze and detect stress. Analyzed the impact of mental and physical stress associatedwith meditation, presentation, pain, math and exercise. Conducted a study to eliminate the time consuming process associated with feature extraction and utilized transfer learning to predict stress from raw ECG signals, with 90.19% accuracy. Although they did not achieve state of the art results for stress classification, developing an algorithm which was capable of predicting stress through raw ECG's opened the gate for remote monitoring of stress in real-time. Developed an algorithm to convert 1D PPG signals into 2D spatial images, using WESAD dataset. They utilized a CNN algorithm to predict stress with high accuracy, using the 2D spatial images for stress and non stress conditions. Our research was similar theirs, but we converted 1D ECG data into 2D stress images. The aforementioned literatures above, along with other research studies has encouraged us to study the impact of 2D images, which can be used to detect stress using raw ECG data. Most deep learning algorithms utilize enormous number of nodes and hidden layers, which are redundant and do not provide any significant/discriminative information, needed to optimize the error associated with the outcome. This increases model and computational complexity, and increases inference time, while degrading the performance of the algorithms. Utilized pruning to significantly reduce the number of redundant weights, by replacing them with 0, thus notably reducing the inference period and computational complexity. Utilized quantization to convert float32 representations to an int8 representations, resulting in a deep learning algorithm with much lower computational and storage requirements. Developed an algorithm using knowledge distillation, which enabled the transfer of knowledge from a sophisticated model (teacher) to a much smaller student model, where the loss function is optimized to reduce the performance gap between the student and teacher model, resulting in a much lower latency period. Our proposed algorithm resulted in an inference period, which was 10 times faster, without compromising the sophisticated functionality of the algorithm and its performance. We propose an algorithm to convert 1D ECG data from WESAD into 2D images, in order to develop a 2D stress image dataset. Images do not require feature extraction and are easier to automate using cloud applications. Given the recent success of computer vision for real-time detection of fraud, theft and facial recognition, stress images can make a significant difference. It can make automation and real-time detection of stress easier and more effective. We leverage transfer learning and model compression techniques to obtain competitive results, which were better than other research studies conducted to detect stress. Our proposed solution has significantly lower computational and power requirements, than most general deep learning algorithms.

2.LITERATURE REVIEW

REAL-TIME STRESS LEVEL FEEDBACK FROM RAW ECG SIGNALS FOR PERSONALISED, CONTEXT-AWARE APPLICATIONS USING LIGHTWEIGHT CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES:

Human stress is intricately linked with mental processes such as decision making. Public protection practitioners, including Law Enforcement Agents (LEAs), are forced to make difficult decisions during high-pressure operations, under strenuous circumstances. In this respect, systems and applications that assist such practitioners to take decisions are increasingly incorporating user stress level information for their development, adaptation, and evaluation. To that end, our goal is to accurately detect and classify the level of acute, short-term stress, in real time, for the development of personalized, context-aware solutions for LEAs. Deep Neural Networks (DNNs), and in particular Convolutional Neural Networks (CNNs), have been gaining traction in the field of stress analysis, exhibiting promising results.

Furthermore, the electrocardiogram (ECG) signals have also been widely adopted for estimating levels of stress. In this work, we propose two CNN architectures for the stress detection and 3- level (low, moderate, high) stress classification tasks, using ultra short-term raw ECG signals (3 s). One architecture is simple and with a low memory footprint, suitable for running in wearable edge-computing nodes, and the other is able to learn more complex features, having more trainable parameters. The models were trained on the two publicly available stress classification datasets, after applying pre-processing techniques, such as data pruning, down-sampling, and data augmentation, using a sliding window approach. After hyper parameter tuning, using 4-fold cross-validation, the evaluation on the test set demonstrated state-of-the-art accuracy both on the 3- and 2-level stress classification task using the DriveDB dataset, reporting an accuracy of 83.55% and 98.77% respectively.

3.SYSTEMANALYSIS

Existing Work:

Most deep learning algorithms utilize enormous number of nodes and hidden layers, which are redundant and does not provide any significant/discriminative information, needed to optimize the error associated with the outcome. This increases model and computational complexity, and increases inference time, while degrading the performance of the algorithms (He, Zhang, & Sun, 2017). Zhou, MoosaviDezfooli, Cheung, and Frossard (2018) utilized pruning to significantly reduce the number of redundant weights, by replacing them with 0, thus notably reducing the inference period and computational complexity. Zhou et al. (2018) utilized quantization to convert float32 representations to an int8 representations, resulting in a deep learning algorithm with much lower computational and storage requirements.

Disadvantages of Existing System:

1. Does not provide any significant/discriminative information, needed to optimize the error associated with the outcome

2.They did not achieve state of the art results for stress classification

Proposed System:

We propose an algorithm to convert 1D (dimensional) ECG data from WESAD (wearable stress and affect detection dataset) into 2D ECG images, which are representative of stress/not stress. It does not require time consuming processes such as feature extraction and filtering. We utilize transfer learning to obtain competitive results. We also demonstrate that model compression techniques can significantly reduce the computational size of the algorithms, without sacrificing much of the performance, as evident from a classification accuracy of 90.62% using the quantization technique. Results substantiate the effectiveness of our proposed method and empirically demonstrates the potential of deep learning algorithms for edge computing and mobile applications, which utilizes low performing hardware.

Advantages of proposed system:

1.It does not require time consuming processes such as feature extraction and filtering.

2.Results substantiate the effectiveness of our proposed method and empirically demonstrates the potential of deep learning algorithms for edge computing and mobile applications, which utilizes low performing hardwar

4.ALGORITHMS

1D TO 2D DATA TRANSFORMATION ALGORITHM:

We propose an algorithm which transforms the 1D data from WESAD and segments them into 2D stress images. We also discuss the algorithms utilized to classify stress from 2D images through transfer learning and model compression techniques. This is the only literature which examines the impact of pruning, quantization and knowledge distillation for stress prediction, using 2D ECG images.

TRANSFORMATION OF 1D SIGNALS TO 2D IMAGES AND CLASSIFICATION

Which proposed the theory that 2D images were more effective and produced better results in comparison to 1D data we propose an algorithm to convert 1D ECG data from WESAD into 2D grayscale images? stress in real-time using ECG signals for stressed and non stressed conditions from WESAD, through a on-chip reservoir computer (RC) digital classifier stress in real-time using ECG signals for stressed and non stressed conditions from WESAD, through a on-chip reservoir computer (RC) digital classifier the difference between spatial images for stressed and non stressed condition using WESAD dataset. Unlike us, the latter classified segmented PPG images and discriminated between them for stressed and non-stressed conditions. Real-time analysis of stress is an on going problem, extremely low epochs can significantly reduce the amount of information from 1D data, but cloud applications have utilized computer vision and 2D images to provide smart security, detect fraud and faces in real-time The aforementioned literature's and success of computer vision, only encourages the idea that a stressimage dataset would allow for effective detection of stress in real-time. Algorithm 1 details the steps required to detect the R-peaks from an ECG signal and segment them, prior to using OpenCV to convert each segmented data into ECG images. The images were resized to size128 × 128. Since the color does not provide any significant information about an ECG, we used grayscale images to reduce the input dimension, thus removing any redundant information, making it easier and faster to train the algorithm.

PRE-TRAINED ALGORITHMS:

Pre-trained Models:
A pre-trained model has been previously trained on a dataset and contains the weights and biases that represent the features of whichever dataset it was trained on. Learned features are often transferable to different data. For example, a model trained on a large dataset of bird images will contain learned features like edges or horizontal lines that you would be transferable your dataset.

VGG-16:

VGG-16 is a CNN Architecture that won the 2014 ImageNet Large Scale Visual Recognition Chal-lenge (ILSVRC). The VGG-16 architecture is an enhancement to the AlexNet design. Instead of directly employing the VGG16, we made some slight modification, but kept the feature extraction part of the model which consists of consecutive convolutional and pooling layers. Instead of directly transitioning to fully connected layers, two intermediate layers, dropout and global average pooling are used.

Fully connected layers are vulnerable to overfitting, limiting the entire network's generalization capabilities. To avoid this, dropout is added as an intermediate layer for regularization. The reason behind using the global average pooling layer is it enforces a strong correspondence between the feature maps and categories. Finally, three fully connected dense layers are used, where the final layer is comprised of three perceptions for three categories.

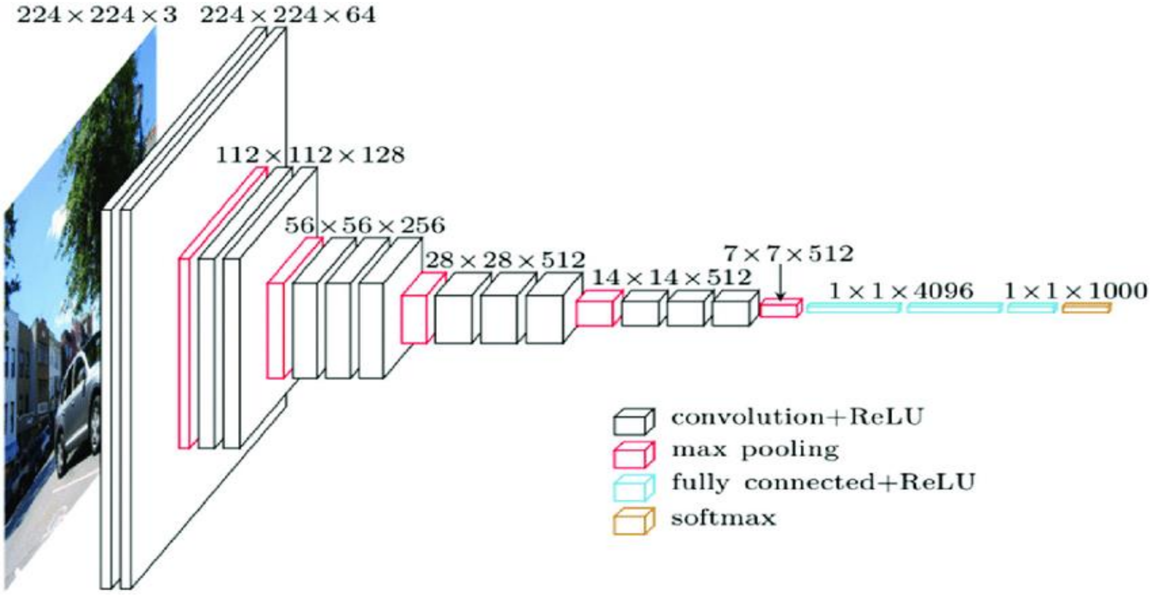


fig. Architecture of VGG-16

VGG-19:

The VGG-19 Neural Network consists of 19 layers of deep neural network and has more weight. The size of "VGG-19"network in terms of fully connected nodes is574 MB. As the number oflayer increases, accuracy of DNN is improved. The VGG-19 model comprised of 19 deeptrainable layers performing convolution, which is fully connected with max pooling and dropoutlayers.

Convolutional layer is trained to perform customized classification role that involved densely connected classifier and to regularize a dropout layer was used. TheVGG-19 architecture is so beneficial and it simply uses 3 × 3 convnet arranged as above to extend the depth.

VGG-19 Net extracted the low-level and high-level features of images layer by layer, and finally realized image classification, so as to achieve the recognition accuracy requirements. The main identification process in this paper is divided into data collection, data preprocessing, model training, model test and model test.The VGG-19 Neural Network consists of 19 layers of deep neural network and has more weight. The size of "VGG-19" network in terms of fully connected nodes is574 MB. As the number of layer increases, accuracy of DNN is improved. The VGG-19 model comprised of 19 deep trainable layers performing convolution, which is fully connected with max pooling and dropout layers aforementioned literature's and success of computer vision, only encourages the idea that a stressimage dataset would allow for effective detection of stress in real-time. Algorithm 1 details the steps required to detect the R-peaks from an ECG signal and segment them, prior to using OpenCV to convert each segmented data into ECG images. The images were resized to size128 × 128. Since the color does not provide any significant information about an ECG, we used grayscale images to reduce the input dimension, thus removing any redundant information, making it easier and faster to train the algorithm.

aforementioned literature's and success of computer vision, only encourages the idea that a stressimage dataset would allow for effective detection of stress in real-time. Algorithm 1 details the steps required to detect the R-peaks from an ECG signal and segment them, prior to using OpenCV to convert each segmented data into ECG images. The images were resized to size128 × 128. Since the color does not provide any significant information about an ECG, we used grayscale images to reduce the input dimension, thus removing any redundant information, making it easier and faster to train the algorithm.

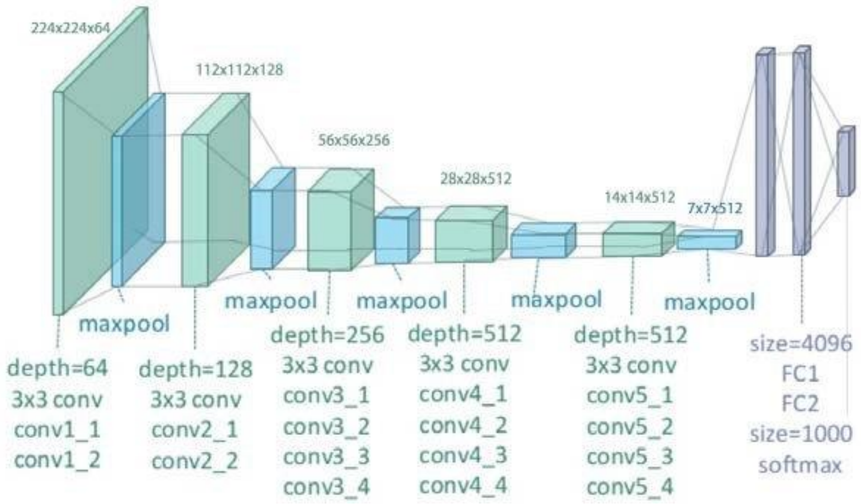


fig. Architecture of VGG-19

3.RESNET 50:

ResNet stands for Residual Network and is a specific type of convolutional neural network (CNN).ResNet-50 is a 50-layer convolutional neural network (48 convolutional layers, one MaxPool layer, and one average pool layer). Residual neural networks are a type of artificial neural network (ANN) that forms networks by stacking residual blocks.

ResNet-50 Architecture:

The 50-layer ResNet uses a bottleneck design for the building block. A bottleneck residual block uses 1×1 convolutions, known as a “bottleneck”, which reduces the number of parameters and matrix multiplications. This enables much faster training of each layer. It uses a stack of three layers rather than two layers. A 7×7 kernel convolution alongside 64 other kernels with a 2-sized stride.
A max pooling layer with a 2-sized stride.
9 more layers—3×3,64 kernel convolution, another with 1×1,64 kernels, and a third with 1×1,256 kernels. These 3 layers are repeated 3 times.
12 more layers with 1×1,128 kernels, 3×3,128 kernels, and 1×1,512 kernels, iterated 4 times.
18 more layers with 1×1,256 cores, and 2 cores 3×3,256 and 1×1,1024, iterated 6 times.
9 more layers with 1×1,512 cores, 3×3,512 cores, and 1×1,2048 cores iterated 3 times.
(up to this point the network has 50 layers)
Average pooling, followed by a fully connected layer with 1000 nodes, using the softmax activation function.

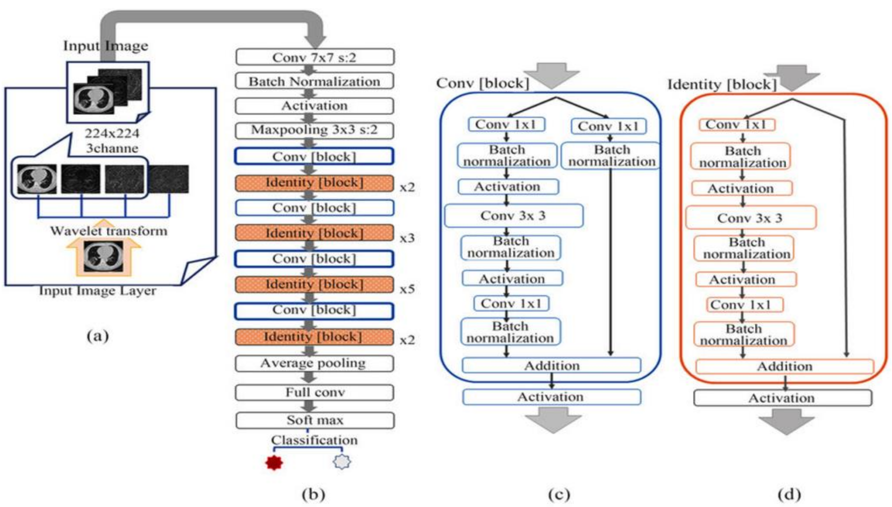


Fig: Architecture of Resnet-50

4.Inception v3:

Inception v3 mainly focuses on burning less computational power by modifying the previous Inception architectures. This idea was proposed in the paper Rethinking the Inception Architecture for Computer Vision, published in 2015. It was co-authored by Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, and Jonathon Shlens.
In comparison to VGGNet, Inception Networks (GoogLeNet/Inception v1) have proved to be more computationally efficient, both in terms of the number of parameters generated by the network and the economical cost incurred (memory and other resources). If any changes are to be made to an Inception Network, care needs to be taken to make sure that the computational advantages aren't lost. Thus, the adaptation of an Inception network for different use cases turns out to be a problem due to the uncertainty of the new network's efficiency. In an Inception v3 model, several techniques for optimizing the network have been put suggested to loosen the constraints for easier model adaptation. The techniques include factorized convolutions, regularization, dimension reduction, and parallelized computations.

5.Dense net 169:

DenseNet is a network architecture where each layer is directly connected to every other layer in a feed-forward fashion (within each *dense block*). For each layer, the feature maps of all preceding layers are treated as separate inputs whereas its own feature maps are passed on as inputs to all subsequent layers. This connectivity pattern yields state-of-the-art accuracies on CIFAR10/100 (with or without data augmentation) and SVHN. On the large scale ILSVRC 2012 (ImageNet) dataset, DenseNet achieves a similar accuracy as ResNet, but using less than half the amount of parameters and roughly half the number of FLOPs.
Whereas traditional convolutional networks with L layers have L connections - one between each layer and its subsequent layer - our network has L(L+1)/2 direct connections. For each layer, the feature-maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs into all subsequent layers. DenseNets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters. We evaluate our proposed architecture on four highly competitive object recognition benchmark tasks (CIFAR-10, CIFAR-100, SVHN, and ImageNet). DenseNets obtain significant improvements over the state-of-the-art on most of them, whilst requiring less memory and computation to achieve high performance.

TRANSFER LEARNING MODELS:

Transfer learning is a method for feature representation from a pre-trained model facilitating us that we don't need to train a new model from scratch. A pre-trained model is usually trained on a huge dataset such as ImageNet and the weights obtained from the trained model can be used for any other related application with your custom neural network. These newly built models can directly be used for predictions on relatively new tasks or can be used in training processes for related applications. This approach not only reduces the training time but also lowers the generalization error.

1.Pruning:

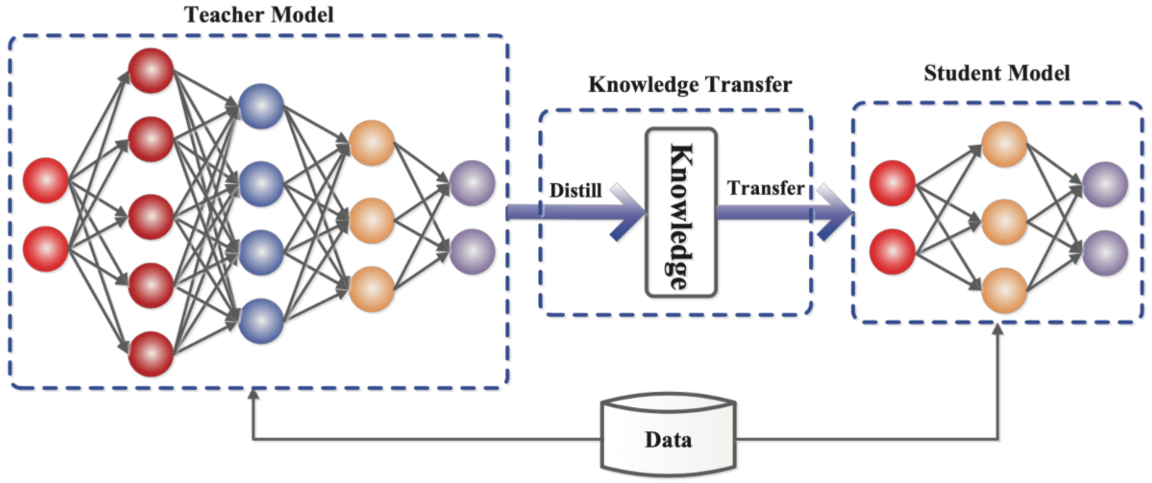
Pruning in deep learning basically used so that we can develop a neural network model that is smaller and more efficient. The goal of this technique is to optimize the model by eliminating the values of the weight tensors. The aim is to get a computationally cost-efficient model that takes a less amount of time in training. The necessity of pruning on one hand is that it saves time and resources while on the other hand is essential for the execution of the model in low-end devices such as mobile and other edge devices. Deep Learning Algorithms like Convolutional Neural Networks (CNNs) suffer from different issues, such as computational complexity and the number of parameters. Hence the goal here is to test different methodologies for pruning a deep learning model and suggest a different approach for Pruning.

2.Quantization:

Quantization is a method to bring the neural network to a reasonable size, while also achieving high performance accuracy. This is especially important for on-device applications, where the memory size and number of computations are necessarily limited. Quantization is the process of approximating a neural network that uses floating-point numbers by a neural network of low bit width numbers. This dramatically reduces both the memory requirement and computational cost of using neural networks.

3.Knowledge distillation:

Knowledge Distillation is a process of condensing knowledge from a complex model into a simpler one. It originates from Machine Learning, where the goal is to create models that can learn from data and make predictions. Early applications of Knowledge Distillation focused on creating smaller, more efficient models that could be deployed on devices with limited resources.
Distillation of knowledge means that knowledge is transferred from the teacher network to the student network through a loss function where the optimization target is to match the class-wise probability distribution of the student network to the probability output by the teacher. This concept of model compression was generalized, and the concept of distillation was formulated in 2015 by Hinton et al



EXPERIMENTAL RESULTS:

1.LOOCV(Leave One Out Cross-Validation)

It is a type of cross-validation approach in which each observation is considered as the validation set and the rest (N-1) observations are considered as the training set. In LOOCV, fitting of the model is done and predicting using one observation validation set. Furthermore, repeating this for N times for each observation as the validation set. Model is fitted and the model is used to predict a value for observation. This is a special case of K-fold cross-validation in which the number of folds is the same as the number of observations(K = N). This method helps to reduce Bias and Randomness**.** The method aims at reducing the Mean-Squared error rate and prevent over fitting.

Mathematical Expression
LOOCV involves one fold per observation i.e each observation by itself plays the role of the validation set. The (N-1) observations play the role of the training set. With least-squares linear, a single model performance cost is the same as a single model. In LOOCV, refitting of the model can be avoided while implementing the LOOCV method. **MSE (**Mean squared error**)** is calculated by fitting on the complete dataset.

$$CV_{(n)} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i - y_i}{1 - h_i} \right)^2$$

In the above formula, $\mathbf{h_i}$ represents how much influence an observation has on its own fit i.e between 0 and 1 that punishes the residual, as it divides by a small number. It inflates the residual.

2.10-k fold cross validation:

K-fold cross-validation is a technique for evaluating predictive models. The dataset is divided into k subsets or folds. The model is trained and evaluated k times, using a different fold as the validation set each time. Performance metrics from each fold are averaged to estimate the model's generalization performance. This method aids in model assessment, selection, and hyperparameter tuning, providing a more reliable measure of a model's effectiveness.

In each set (fold) training and the test would be performed precisely once during this entire process. It helps us to avoid overfitting. As we know when a model is trained using all of the data in a single short and give the best performance accuracy. To resist this k-fold cross-validation helps us to build the model is a generalized one.

To achieve this K-Fold Cross Validation, we have to split the data set into three sets, Training, Testing, and Validation, with the challenge of the volume of the data.

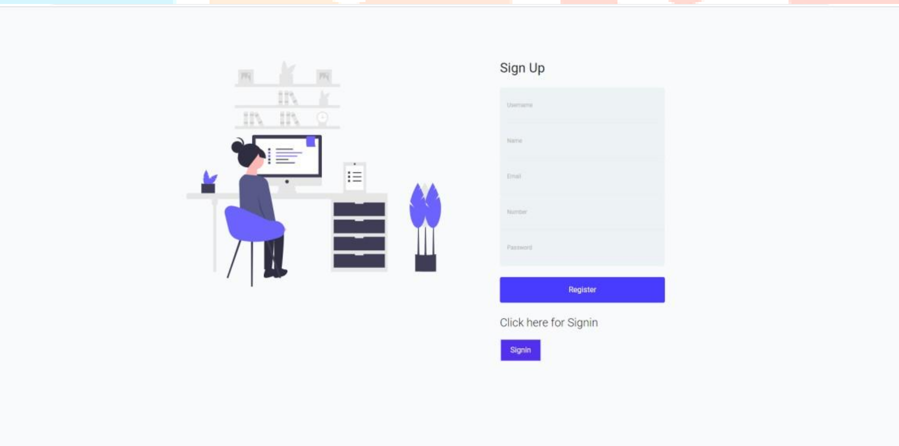
Here Test and Train data set will support building model and hyperparameter assessments.

In which the model has been validated multiple times based on the value assigned as a parameter and which is called K and it should be an INTEGER.

Make it simple, based on the K value, the data set would be divided, and train/testing will be conducted in a sequence way equal to K time.

5.EXPERIMENTAL RESULTS:

A registration page enables the users and organizations to independently register and gain access toyour system.



A login page is a set of credentials used to authenticate a registered user

Fig:Login page

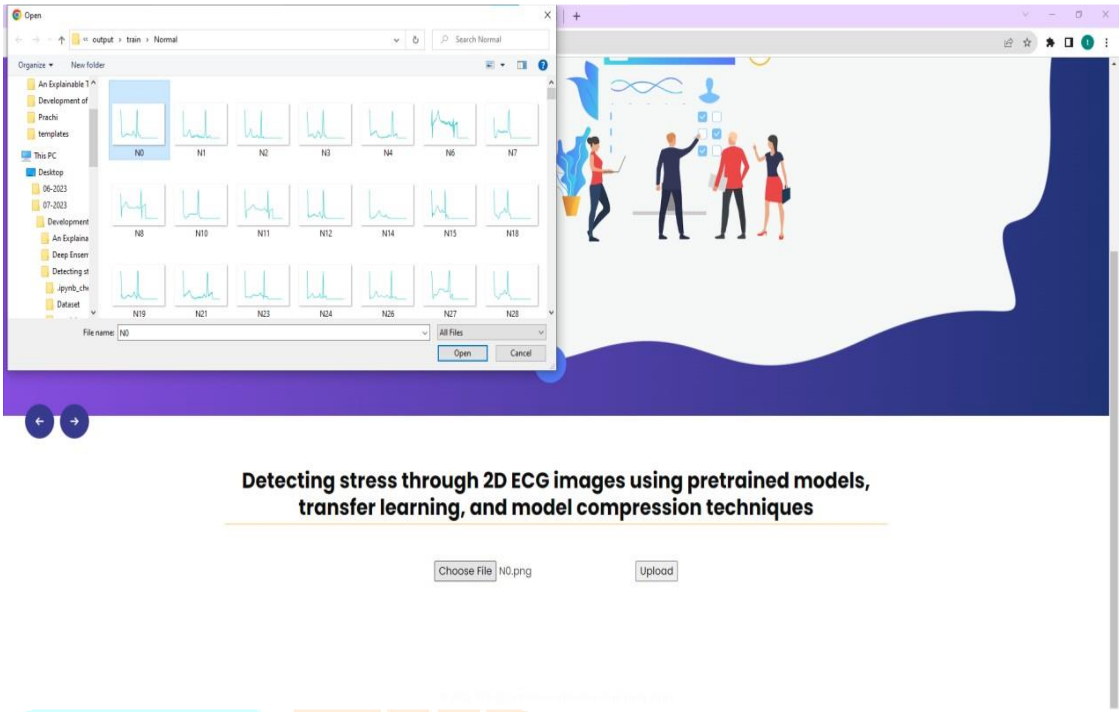
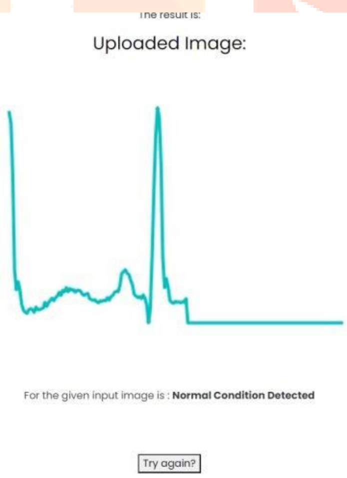


FIG:Web page showing the test cases for detecting the ECG images whether the obtainedimages are related to Normal image or Stress image.



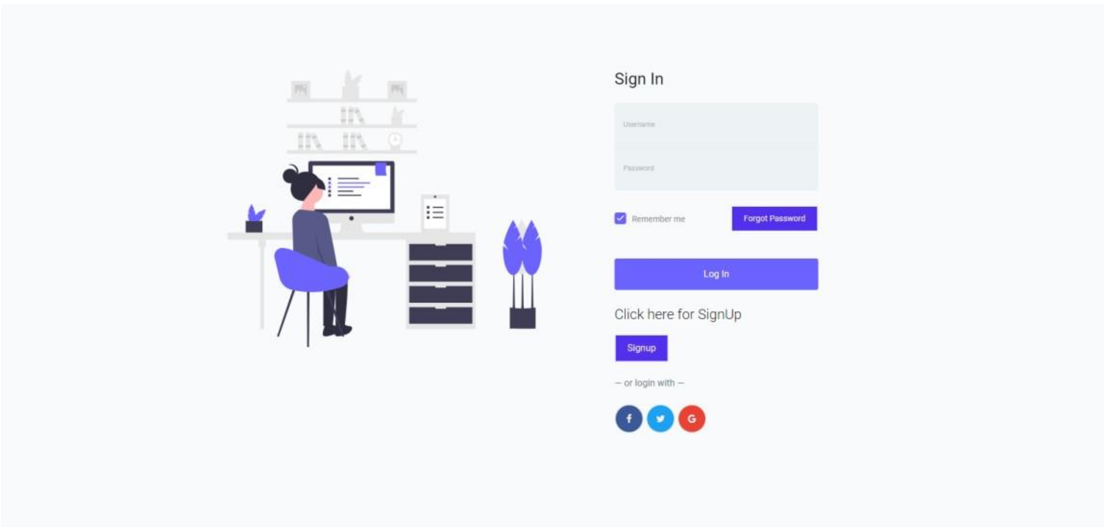


FIG:The above image shows that the ECG image is Normal.

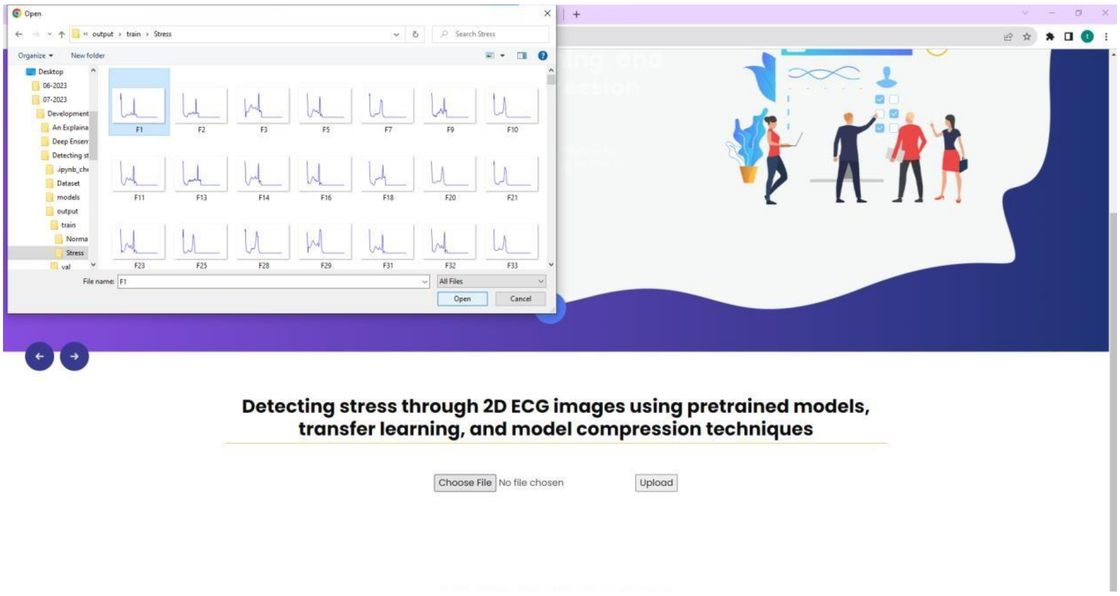


FIG:Web page showing the test cases for detecting the ECG images whether the obtainedimages are related to Normal image or Stress image.

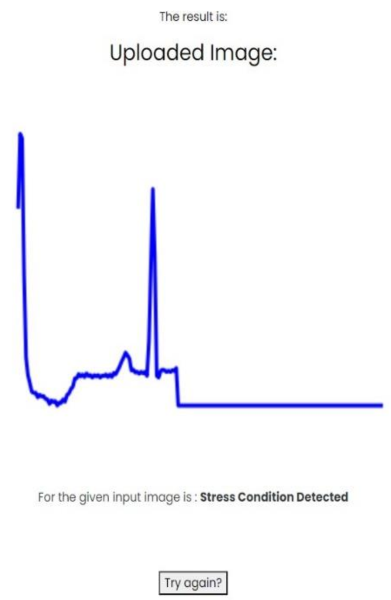


FIG:The above image shows that the ECG image is Stress.

6.CONCLUSION:

The paper introduces how to detect stress through 2D ECG images using pre-trained models, transfer learning and model compression techniques. The proposed method addresses the drawbacks of basic methods like does not provide any significant/discriminative information, needed to optimize the error associated with the outcome. They did not achieve state of the art results for stress classification.

We developed a stress image dataset, which was used to predict stress with 90.62% accuracy, throughmodel compression techniques. Given the success of computer vision for real-time detection, we are optimistic that stress images and lightweight models, can make a significant impact and make it easierto automate algorithms to predict stress in real-time.

In the near future, we want to use statistical testing to test for bias, variance associated with the aforementioned methods.We also want to classify 2D spectrograms using LTSM and classify 2D datathrough cloud applications.

Experimental results prove that we beat the previous state-of-the-art in terms of classification accuracy, precision and recall. The important finding of this study is that the ECG stress or unstress ofdifferent dataset increases the performance of the machine learning task as compare to the other optimizers to find the greater accuracy individually.

7.REFERENCES:

[1] Ahmad, Z.Tabassum, A. Guan, L. & Khan, "ECG heart-beat classification using multimodal imagefusion" In ICASSP 2021-2021, IEEE international conference on acoustics, speech and signal processing ,pp.1330–1334.

[2] Bong, S. Z. Murugappan, M. & Yaacob, "Analysis of electrocardiogram (ECG) signals for human emotional stress classification" in International conference on intelligent robotics, automation, and manufacturing ,pp.198–205.

[3] Chandrasekaran, S. T. Bhanushali, S. P. Banerjee, I. & Sanyal, " A bioinspired reservoir-computer for real-time stress detection from ECG signal" IEEE Solid-State Circuits Letters, 3, 290–293.

[4] Cho, H.M. Park, H.Dong, S.Y.Youn, "Ambulatory and laboratory stress detection based on raw electrocardiogram signals using a convolutional neural network Sensors" 19(20), 4408.

[5] Elzeiny, S.Qaraqe, "Stress classification using photoplethysmogram-based spatial and frequency domain images Sensors" 20(18), 5312, Frankle, J.Carbin, M. (2018), The lottery ticket hypothesis: Finding sparse, trainable neural networks, arXiv preprint arXiv:1803.03635.

[6]K.Zhang, X. Ren, S. Sun, " Deep residual learning for image recognition .In Proceedings of the IEEE conference on computer vision and pattern recognition" ,(pp. 770–778).

[7] Hinton, G. Vinyals, O. Dean, J.et al, (2015) "Distilling the knowledge in a neural network" ,2, (7), arXiv, preprint arXiv:1503.02531.

[8] Huang, J. Chen, B. Yao, B. He, W. (2019) "ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network" , IEEE Access, 7, 92871–92880.

[9] Huang, G. Liu, Z. Van Der Maaten, L. Weinberger, K. Q. (2017) "Densely connected convolutional networks" In Proceedings of the IEEE conference on computer vision and pattern recognition ,(pp. 4700–4708).