SiliconTech

The Science & Technology Magazine

Digital Digest

Our Vision: "To become a center of excellence in the fields of
technical education & research and create responsible citizens"

WORMHOLES

According to the Richard F. Holman ,professor of physics at Carnegie Mellon University, "Wormholes are solutions to the Einstein field equations for gravity that act as "tunnels," connecting points in space-time in such a way that the trip between the points through the wormhole could take much less time than the trip through normal space".

But the question arises do wormholes really exist? Over the years we have seen many sci-fi movies, and have read books where it is possible to travel through these short cuts. Wormholes are possible, according to Einstein's General theory of relativity. Theoretical physicists have hypothesized the existence of such shortcuts through space-time since the 1930s. Wormholes are like tunnels in space through which one can travel galactic distances within a short time. Scientists have shown that wormholes can be created by pulling apart two entangled black holes, where the black holes are the regions with strong gravitational pull that nothing can escape from it.

Recently scientists have found that to create a traversable wormhole, there must be a substance with negative energy called exotic matter a hypothetical particle having negative mass and defying all the currently accepted laws of Physics. It is believed that the solution lies in the cryptic equation "ER = EPR." That is, wormholes between distant points in space-time, the simplest of which are called Einstein-Rosen or "ER" bridges, are equivalent to entangled quantum particles, also known as Einstein-Podolsky-Rosen or "EPR" pairs.

According to some theories, wormholes can exist without exotic matter, if the general relativity is modified. As per the hypotheses wormholes may allow faster than light travel ensuring that the speed of light is not exceeded locally at any time. But while travelling through it lower than light speeds are used. Also if an object travelling faster than light will cross the black hole and will emerge from another end into a different space, time or universe. This will be an inter-universal wormhole, thereby speculating on multiple universes! Moreover, some Physicists believe that these wormholes not only connect two separate regions within the universe but time travel can be possible if one mouth of a wormhole is moved in a specific manner.

These findings may find a way to travel to nearest galaxy in a shortest possible time, through these shortcuts or Wormholes in the not-too-distant future.

Dr. Lopamudra Mitra

Sr. Asst. Professor
Dept. of EEE

Facial Expression Recognition with Age and Gender Detection

Abstract: *This work brings into light the facial expression recognition with age and gender classification which holds innumerable applications but poses to be equally challenging owing to the high amount of intra-class variations involved. Recently many deep learning models have been put forward but still there remains a lot of scope for further improvement. In this work, we focused on the important aspects of a facial image which positively impacts the result obtained, thus resulting in much more efficiency than the pre-proposed models. For this purpose, two separate aspects: expression and the age and gender are considered. For expressions classification the FER13 and RAF-DB dataset are used whereas for age and gender detection the All-Age-Face dataset is used. A user interface is also designed that predicts human expression, age and gender simultaneously by analyzing given images.*

Keywords: Convolutional Neural Networks, Deep Learning, Classifier

I. INTRODUCTION

Expressions are an indispensable part of human life, serving both as a verbal communication enhancer and non-verbal communication medium. The different forms in which they are manifested might sometimes go unnoticed. Therefore, we seek tools which may help us detect and recognize the expressions. Simultaneously, understanding age and gender from the human face plays an essential role in social interaction. To make communication proper and efficient, people subconsciously judge others' age or gender.

Facial expression recognition with age and gender detection has been traditionally associated with the security sector and social media platform but today there is active expansion into other industries including retail, marketing and health. Here is a descriptive explanation of how it helps in the following sectors:

 Security: Companies are training deep learning algorithms to recognize fraud detection, reduce the need for traditional passwords, and to improve the ability to distinguish between a human face and a photograph.

Advertising: The ability to collect and collate masses of personal data has given marketers and advertisers the chance to get closer than ever to their target markets.

Criminal Identification: Artificial Intelligence equipped cameras can be set up in public places prone to theft or crimes. These surveillance cameras can help capture images of criminals which can help identify them.

Mobile Applications: Android applications like AgeBot and Quividi that use this technology to detect age and gender from the photos using facial recognition.

In recent years, there has been an increased rise in the application areas of human expression recognition in various fields including, but not limited to healthcare, security, surveillance, and animation [1]. Also, age and gender detection is helpful in several applications, including re-identification in surveillance videos, intelligent advertising and human-computer interaction and in social media platforms. Nevertheless, accurately and efficiently recognizing expression, age and gender from unconstrained face images is difficult with factors like extreme blur (low-resolution) occlusions, out-of-plane pose variations, expressions and more. Therefore, only with the right tools, classification can be made possible.

Recently, with the use of deep learning and especially Convolutional Neural Networks (CNNs), many features can be extracted and learned for a decent facial expression recognition system [2]. It is, however, noteworthy that in the case of facial expressions, age and gender, much of the clues come from a few parts of the face, e.g. the mouth and eyes, whereas other parts, such as ears and hair, play little part in the output. This means that ideally, the machine learning framework should focus only on the important parts of the face, and be less sensitive to other facial regions.

In this work, a deep learning based framework is proposed for facial expression recognition with age and gender detection. We worked on two separate aspects: one being expression and the other being age and gender. For expressions classification the FER13 and RAF-DB dataset are used whereas for age and gender detection the All-Age-Face dataset is used. For the RAF-

DB dataset, we designed a model not consisting of a localization network and we used the modified model for testing and training for the RAF-DB dataset. We have focused on using CNN for designing our models. A UI (User Interface) is also designed that predicts facial expression, age and gender simultaneously by analyzing the given images.

II. DATASET USED

There are three datasets used for this work: FER-13, RAF-DB, and AAF-DB. We have collected FER13 dataset from kaggle to train and test the model for facial expression recognition. In this dataset images are present in gray-scale pixel form having 48*48 dimension and seven class labels where each class label corresponds to an expression (0-Angry, 1-Disgust, 2-fear, 3-Happy, 4-Sad, 5-surprise, 6-neutral). There are approximate 28,709 data for training and 3,589 data for testing. RaF-DB dataset contains 15,339 RGB, each having a dimension of 100*100 with seven class labels, each corresponding to an expression (Sad, Happy, Disgust, Fear, Surprise, anger and neutral). This dataset consists of two types of dataset. One containing only images of faces and other containing image of faces with noise. We have used a data set that contains only images of faces in RGB form. The All-age- face dataset is being used for age and gender prediction. This data set is collected from Github. This dataset contains 13,322 face images (mostly asian) distributed across all ages (from 2 to 80) including 7381 females and 5941 males and having discrete values for age like [2-9],[10-17], etc. For fair comparison, the images are randomly splited into two sets, one for training and the other for validation.

III. DATA PRE-PROCESSING

In this work, three types of data set is considered to train and test different models. The following different pre-processing techniques are used for different dataset-

A. Pre-Processing in FER-2013:

Initially in this dataset pixels are stored in a list in string form so, at first the string of pixel is converted into integer and then the pixels are resized to 48*48 dimension and stored in a 2-D array.

B. Pre-processing in RAF-DB:

During preprocessing the RGB images is converted into Gray-scale images as computational cost is low using gray scale images. The dimension of each image is resized to 100*100. The pixel values of the image data are subtracted from 255 during normalization, in order to fit the image data into the network to get more accuracy.

C. Pre-processing in AAF(All-Age-Face) dataset:

In this dataset each image is associated with a name that contains person-id, age and gender, so we have extracted age and gender of each image and stored it in an array for further use. During preprocessing the RGB images are converted to Grayscale images and the dimension of each image is resized to 100*100 and later 300*300 to increase the accuracy of the model. Also, for fair classification, the data is reshuffled randomly by using the shuffle function.

IV. THEORITICAL FRAMEWORK

A. Convolution 2 Dimensional Layers

Convolutional layers play a major role in the convolution neural networks. It generally acts as a simple filter for the inputs and results in activation. When we repeatedly apply the same filter to an input then it leads to a map of activation. This indicates strength and location of features in the input (if its an image). Convolutional neural network has the ability to learn a large number of features parallel to specific training data. In a convolution operation multiplication is performed between the two-dimensional input data and the 2-dimensional matrix consisting of weights that is often called as filter or kernel. Generally the multiplication operation applied is a dot product. Filters have the capability to detect features anywhere in the image.

B. Multiple filters in Convolutional Neural Networks

Convolutional Neural networks generally learns multiple filters for a given and that too in parallel. A Convolutional Neural network can easily learn 32-512 filters in parallel. This gives the filters 32 to 512 different ways in reading the patterns in the input or 'seeing' the input data. This leads the CNN to see and learn various specific lines in the data.

C. Edge detection using Convolutional Neural Network

Convolution neural networks act as a heart for most of the computer vision applications. It uses convolution

operation to transform the input images to output. As said already a single step of convolution consists of dot product of pixel values of the image with the filters. A filter is generally a square matrix of size 3*3 or 4*4 and so on. In the process the 2 dimensional filters are processed along the whole matrix containing the pixel values. The resulting matrix after this operation detects edges or transitions between dark and light colours .As we increase the number of filters the network is able to recognize more details in the image.

Creating horizontal edges in the filter helps in detecting horizontal edges in the image and creating vertical edges in the filter helps in detecting vertical edges.

D. Padding

In this process additional pixels are added to the boundary of the matrix containing the image pixel values. This helps in preserving the input image by preventing the image from shrinking because of convolutional filters. Padding can be of two types i.e "valid" and "same". In valid padding we use zero padding and the size of the image shrinks whereas in "same" padding we increase the padding in order to keep the size of the input and output the same. A snapshot of the same is shown in Fig.1.
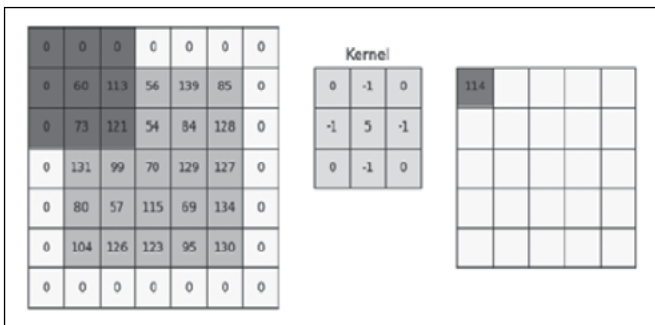

Figure.1: *Example on Padding*

E. Pooling layers

Pooling layers are another building blocks of a Convolutional Neural Network. Its major functionality is to progressively reduce the size of the image pixels representation which in turn reduces the amount of parameters and computation in the network thereby optimizing the network. The most commonly used pooling layer is max-pooling. In max-pooling we generally slice the input matrix and find the maximum value from each slice. The example of the process is shown in fig.2.
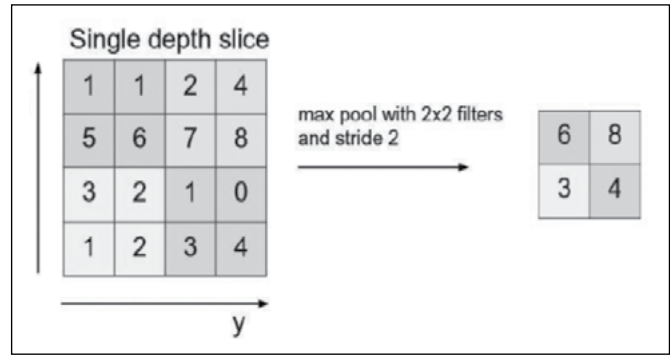

Figure.2: *Pooling layers*

F. Dropout layers

Dropout is a technique used in neural networks where random selection neurons are ignored during training. Due to this the contribution of the neurons gets temporarily removed during forward and backward propagation. If neurons are randomly removed from the network then other neurons have to take part and handle the representation required to make predictions. This results in multiple independent representations learned by the network. Eventually the network becomes less sensitive to specific weights of neurons. This makes the network better capable of generalization and is less likely to overfit the training data. The standard neural network architecture and after applying dropout is shown in fig.3
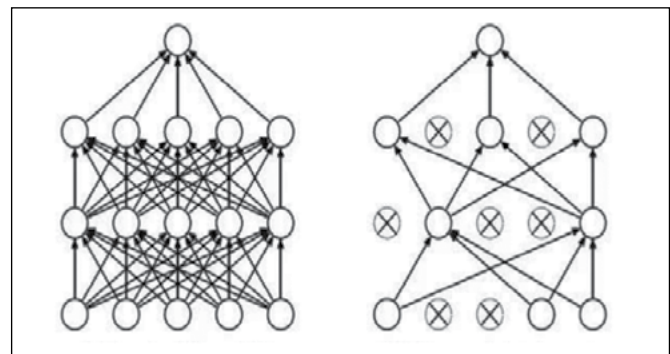

Figure.3: *Neural network architecture*

G. Dense layer

A dense layer is just a regular layer(non-linear layer) of neurons in a neural network. It is a part of a hidden layer of neurons. A Dense layer feed all outputs from the previous layer to all its neurons, each neuron providing one output to the next layer. A dense (10) has 10 neurons. $y=f(w*x+ b)$, the end result is passed through a nonlinear function called Activation function.

Dense implements the operation:-

output= activation(dot(input, kernel)+bias)

H.   Feature Extraction

The feature extraction part consists of four convolutional layers; each two followed by max-pooling layer and rectified linear unit (ReLU) activation function. They are then followed by a dropout layer and two fully-connected layers. The spatial transformer (the localization network) consists of two convolution layers (each followed by max-pooling and ReLU), and two fully-connected layers. After regressing the transformation parameters, the input is transformed to the sampling grid T() producing the warped data. The spatial transformer module essentially tries to focus on the most relevant part of the image, by estimating a sample over the attended region. One can use different transformations to warp the input to the output, here we used an affine transformation which is commonly used for many applications.

V. MODEL ARCHITECTURE

The flowchart for creating different Models is shown in fig. 4. We have taken three types of data sets in this work to train and test the models presented as model-1, model-2 and model-3.
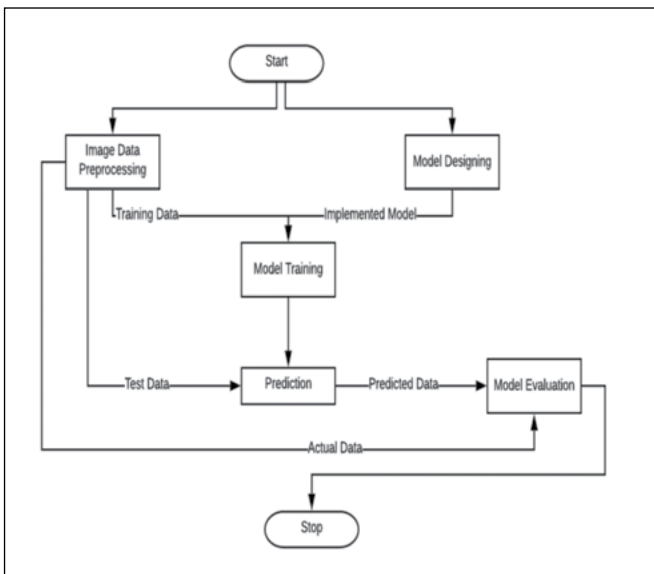

Figure.4: *Flowchart for creating different Models*

The validation mistakes in some cases are quite profound as well. For example, in the case of colored retina images validation, the neural network model is classifying 18 Moderate class as Proliferate_ DRclass. And in total, it is misclassifying 131 images out of 732 validation images. In the case of grayscale validation images, the neural network model is misclassifying 143 images out of 732 validation images.Here, validation results are worse than the colored images validation results. In the case of gaussian filtered validation images, the neural network model is misclassifying 146 images out of 732 validation images. The validation results
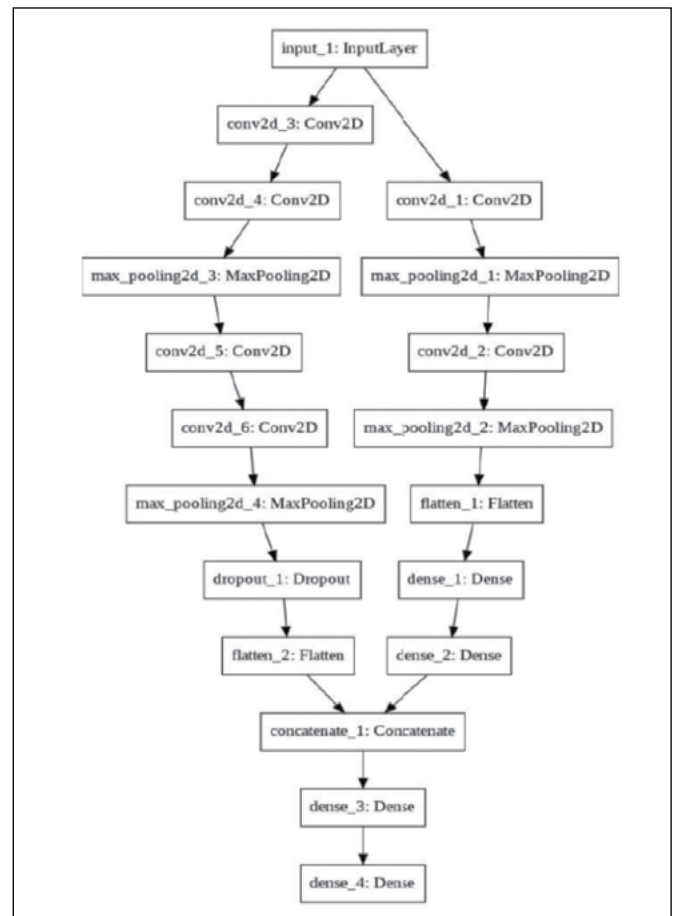
A. Model-1:

Figure.5: *Model Architecture for FER13 dataset*

The model-1 architecture shown in fig.5.is trained and tested with both the datasets i.e. RAF and FER13. The model is equipped with 2 parallel sets of networks, which are then joined together into a

dense layer that, in turn, is connected to the output layer. The upper parallel part of the network is the localization network. It works to localize feature extraction to the part of the image that consists of the facial features useful for expression detection.

B. Model-2:

The model-2 architecture is shown in fig. 6. It consists of 3 sets of feature extractor connected sequentially, each of which consists of 2 convolutional layers, a normalization layer and a dropout layer designed in such a way that unnecessary-feature reduced inputs are provided to the subsequent feature extractor. The snapshot of feature extractor is shown in fig. 7.
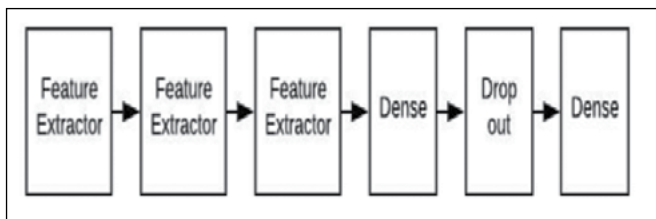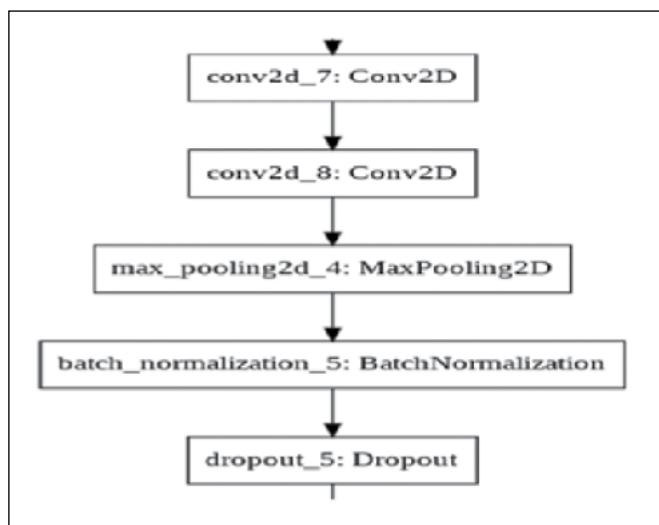

Figure.6: *Model Architecture for RAF dataset*

Figure. 7: *A snapshot of feature extractor*

C. Model-3:

The architecture of model-3 for age and gender detection using AAF datasets is shown in fig. 8.
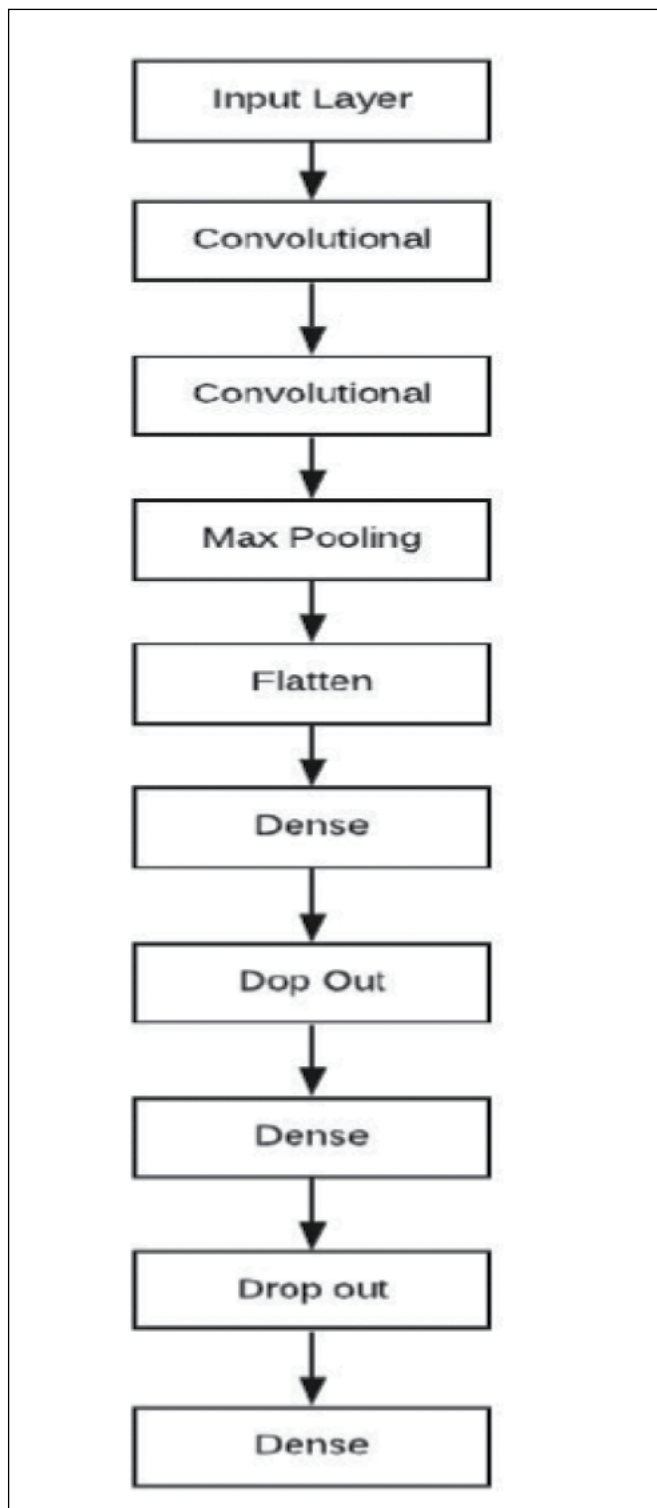

Figure.8: *Model for age and gender detection using AAF datasets*

VI. RESULTS

The model was trained and tested for Facial Expression Recognition along with various other models and was obtained as the one with maximum test accuracy(i.e. 73.65%) amongst them on RAF-DB dataset (used for Facial Expression Recognition). Model 1 was trained and tested on the AAFD dataset for Gender Detection and was observed as a good classifier with a test accuracy of 89.86%. Model 1 was trained and tested on the AAFD dataset for Age Detection and was observed as a bad classifier with a test accuracy of 22.28%. Model 1 is observed to have a degrading performance when the number of classes increases as happens in the case for age detection. Model 2 is a simple approach made towards performing Age and Gender Detection. It was observed that the model is a good classifier when trained and tested for Gender Detection with a test accuracy of 82% but does not perform well in the case of Age Detection. The details of the results obtained are shown in Table-I. The confusion matrix for the three models are presented in Fig. 9, fig. 10 and fig.11 respectively.

Table I: Model evaluation results

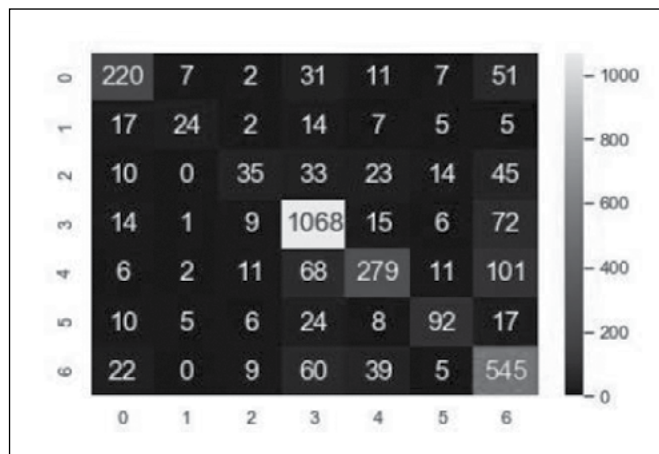| Sl. no	Model	Dataset used	Training accuracy	Testing accuracy
1	Model-1	RAF-DB for FER	83.4%	73.65%
2	Model-2	AAFD for Gender	92%	88%
3	Model-3	AAFD for Age	62%	47%

Figure.9: *Confusion matrix for model-1 on RAF-DB*
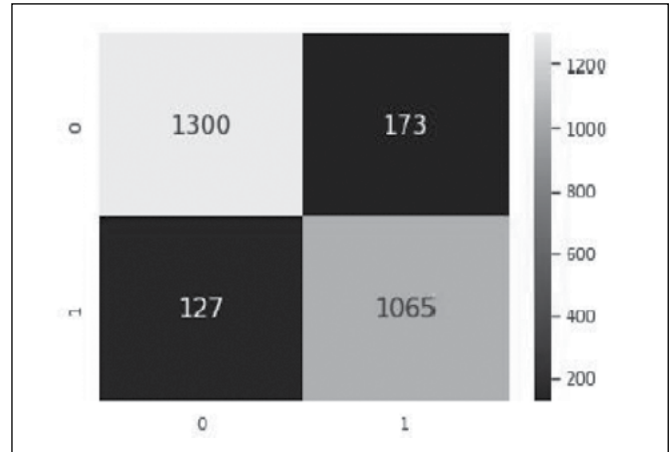
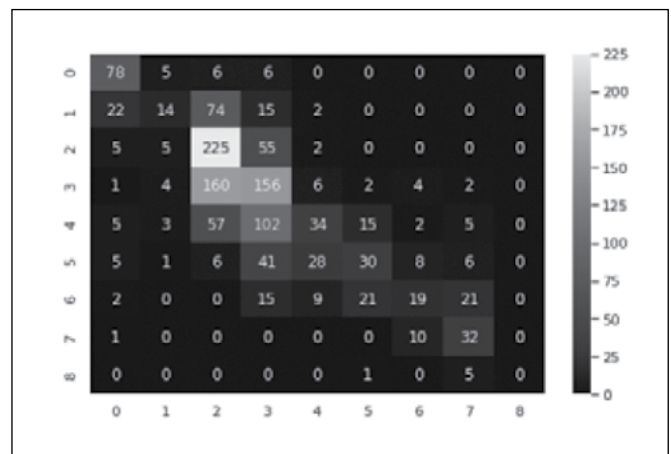***Figure.10:*** *Confusion matrix for model-2 on AAFD for gender*

Figure.11: *Confusion matrix for model-3 on AAFD for age*

VII. CONCLUSIONS

Though many previous methods have addressed the problems of age and gender classification, until recently, much of this work has focused on constrained images taken in lab settings. Such settings do not adequately reflect appearance variations common to the real-world images in social websites and online repositories. Internet images, however, are not simply more challenging: they are also abundant. The easy availability of huge image collections provides modern machine learning based systems with effectively endless training data, though this data is not always suitably labeled for supervised learning. Taking example from the related problem of face expression recognition we explore how well deep

CNN performs these tasks using the gathered data. This work provides results with a deep-learning architecture designed to avoid over fitting due to the limitation of limited labeled data. Two important conclusions can be made from the results obtained. First, CNN can be used to provide improved age and gender classification results, even considering the much smaller size of contemporary unconstrained image sets labeled for age and gender. Second, the simplicity of the model implies that more elaborate systems using more training data may well be capable of substantially improving results beyond those reported here.

REFERENCES

1. Minaee, Shervin, and Amirali Abdolrashidi. "Deep-emotion: Facial expression recognition using attentional convolutional network." arXiv preprint arXiv:1902.01019 (2019).

2. Lee, Jia-Hong, et al. "Joint estimation of age and gender from unconstrained face images using lightweight multi-task cnn for mobile applications." 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, 2018.

**Abhilash Behera, MD Haaris,
Nitu Singh, Prachi Mallick**

Dept. of EEE

NASA's Mars helicopter Ingenuity takes off on historic 1st powered flight on another world

NASA's Ingenuity helicopter lifted off on the Red Planet early this morning (April 19), performing the first-ever powered flight on a world beyond Earth.The 4-lb. (1.8 kilograms) chopper was scheduled to rise from the floor of Mars' Jezero Crater at 12:31 a.m. EDT (0431 GMT) today, get a maximum of 10 feet (3 meters) above the red dirt and land after roughly 40 seconds aloft. At about 6:15 a.m. EDT (1015 GMT), data came down from Ingenuity — via its much larger partner, NASA's Perseverance rover — that the little rotorcraft had hit its marks. The first photo from Ingenuity showed the helicopter's shadow on the Martian surface below, while the Perseverance rover captured stunning video of the historic flight on Mars."Ingenuity has performed its first flight, the first flight of a powered aircraft on another planet!" Ingenuity's chief pilot Håvard Grip said as he confirmed telemetry at NASA's Jet Propulsion Laboratory in Pasadena, California.

Source: ***www.livescience.com, www.space.com***

PV cell fed fuzzy logic controller (mamdani model) based boost converter for 5 level symmetrical cascaded Multi-level inverter

Abstract : *Multi-level inverter technology has emerged recently as a new important alternative in the area of high power medium voltage energy control. A high output power can be obtained from the medium voltage sources such as batteries, solar cells using this multi-level inverter technology. The multi-level inverter is so much on trend because of its advantages such as reduced harmonic distortion in output voltage & lower Electro- Magnetic Interference. While increased number of components as level is increased, voltage balancing method at neutral point, complicated PWM Control method are main disadvantages of this technology. Emerging topologies such as diode- clamped inverter, capacitor clamped inverter and asymmetric hybrid cells have already been used. The new proposed topology uses less number of components as level is increased which lessens the switching losses, subordinate carrier signals for PWM, improved power quality, reduction of harmonics at output voltage. The new converter topology mainly applicable on 5- level & 9- level with the usage of Sinusoidal Pulse Width Modulation (SPWM) technique. The results of this method are also presented & discussed. In this work we have taken solar PV array as the source and used Fuzzy Logic controller for maximizing its power output.*

Keywords: Multi-level inverter (MLI), Pulse width Modulation (PWM), Fuzzy logic controller (FLC),Genetic Algorithm (GA), Harmonics, Photovoltaic (PV), unemployed bees (UB).

I. INTRODUCTION

A power inverter, or inverter, is a power electronic device or circuitry that changes direct current to alternating current. The input voltage, output voltage and frequency, and overall power handling depend on the design of the specific device or circuitry [1]. Typical applications for power inverters include: Portable consumer devices that allow the user to connect a battery, or set of batteries, to the device to produce AC power to run various electrical items such as lights, televisions, kitchen appliances, and power tools. Inverters can be classified according to its waveforms obtained. Broadly it can be classified into single level and multi-level inverter. Multilevel inverters are more advantageous as compared to single level as the Total harmonic distortion will be less. In our project we have simulated a 5 level H bridge type cascaded symmetric inverter. We have taken solar photovoltaic as source and have used fuzzy logic controller to control

II. MULTI-LEVEL INVERTER

The conventional voltage source inverter (VSI) generates a two level output voltage with levels +/ -Vdc/2 and zero, where Vdc is the dc link voltage [2]. The various PWM strategies are applied to high frequency switches to obtain the quality output voltage or current waveform with reduced ripple content. In high-power medium voltage applications, the VSI have limitations such as switching losses at high frequency and voltage/power ratings. Fig.1 shows the multilevel inverter consists of power semiconductors and capacitors voltage sources. The output voltage of the inverter will be stepped waveform which will be nearly equal to sinusoidal waveform. The higher voltage steps leads to reduced dv/dt stress on the load, lesser harmonic content and higher power quality waveforms.
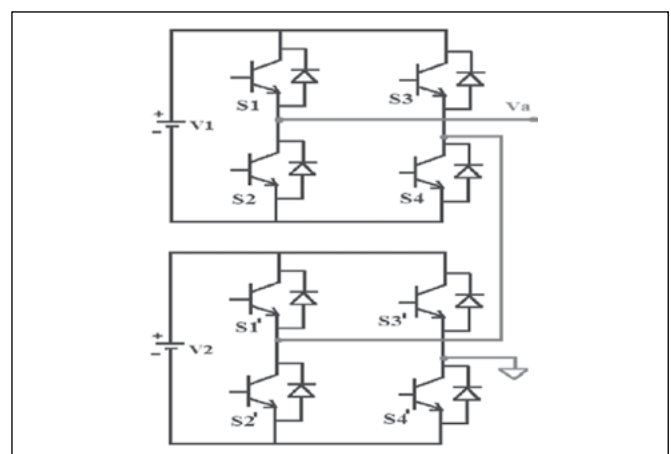

Figure.1: 5-Level Inverter

Pulse width modulation technique:Pulse width modulation (PWM) technology was put forward based on an important conclusion in the sample control theory that when two groups of pulses with the same impulse area but different waveforms are input to an inertial link, the effectiveness of these two groups of impulses are the same. The main principle of PWM technique can be briefly described as shown in fig. 2 [3]. Through ON/OFF control on the semiconductor switching components, a series of pulses with the same amplitude and different width are generated on the output port to replace the sinusoidal wave or other waveforms required. The duty cycle of the output waveform needs to be modulated by a certain rule and as a result both the output voltage and output frequency of the inverter can be regulated.Different types of PWM control technique are single pulse width modulation, multiple pulse width modulation, Sinusoidal pulse width modulation modified& sinusoidal pulse width modulation.
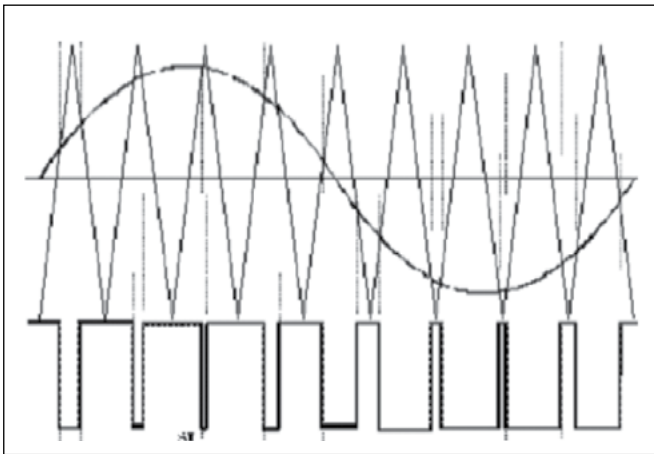

Figure.2: *Generation of Sinusoidal pulse width modulation*

III. SOLAR PHOTOVOLTAIC

A photovoltaic array is a linked collection of photovoltaic modules, one of which is shown in Fig 3. Each photovoltaic (PV) module is made of multiple interconnected PV cells. The cells convert solar energy into direct-current electricity [4]. The MPPT graph is shown in Fig. 4. PV modules are sometimes called solar panels, although that term better applies to solar-thermal water or air heating panels. Photovoltaic modules distinguish themselves from solar cells in that they are conveniently sized and packaged in weather-resistant housings for easy installation and deployment in residential, commercial, and industrial applications. The application and study of photovoltaic devices is known as photovoltaic.
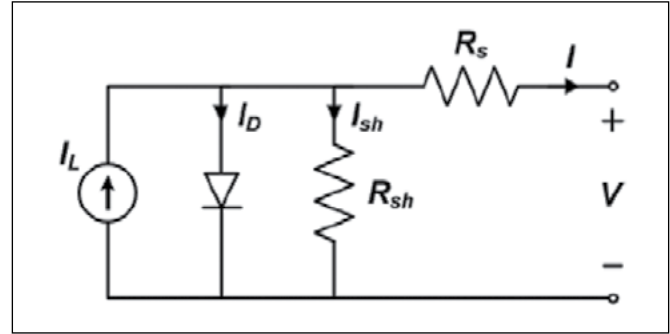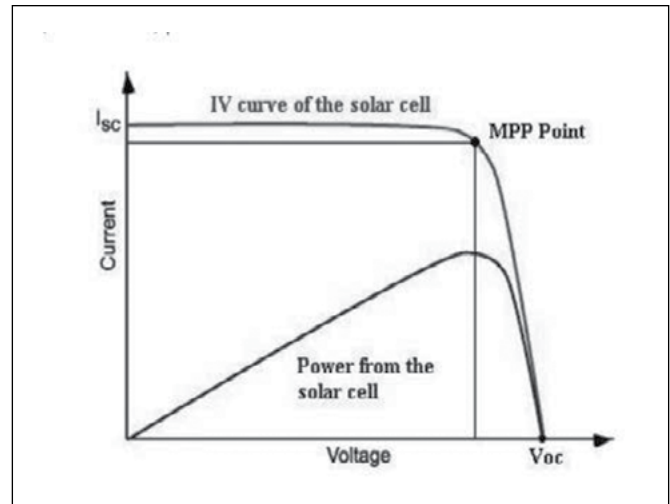

Figure.3: *Equivalent Circuit of a PV Cell*

Figure.4: *Characteristics of a PV Cell*

$$I = I_{ph} - I_s \left(e^{\frac{q(V + IR_s)}{NKT}} - 1 \right)$$

...(1)

I_L is the current generated by solar cell (A)

Io is the diode saturation current (A)

q is the electron charge = 1.6×10-19 Coulomb

n is the ideality factor of diode which varies from 1 to 2

k is Boltzmann constant = 1.38×10-38 Joule/Kelvin

Tr is the rated cell temperature (K)

The cell could be represented by a voltage-current equation:

V= Vsh-RsI ... (2)

where

V is the cell output voltage (V)

I is the load (cell) output current (A)

Vsh is the voltage across shunt resistance (V)

$I_L = G*Isc(Tr)(1+\alpha Isc(T-Tr))/Gr \quad ...(3)$ **...(3)**

where

αIsc is the short circuit temperature coefficient (A/°C)

G is operating solar radiation level (W/m2)

Gr is rated solar radiation level (W/m2)

$$I_o = I_{o(T_r)} \times \left(\frac{T}{T_r}\right)^3 \times e^{\frac{qV_g}{nk} \times \left(\frac{1}{T_r} - \frac{1}{T}\right)} \quad ...(4)$$

$$I_{o(T_r)} = \frac{I_{sc(T_r)}}{e^{\left(\frac{qV_{oc(T_r)}}{nkT}\right)} - 1} \quad ...(5)$$

where,

Isc(Tr) is the short circuit current of the cell at rated temperature (A)

Vg is the band gap voltage (V)

Voc(Tr) is the open circuit voltage of the cell at rated temperature (V)

T is operating temperature of the cell(K)

MPPT and control techniques: In order to increase this efficiency, MPPT controllers are used. Such controllers are becoming an essential element in PV systems shown in Fig. 5. A significant number of MPPT control schemes have be elaborated since the seventies, starting with simple techniques such as voltage and currentfeedback based MPPT to more improved power feedback based MPPT such as the perturbationand observation (P&O) technique or the incremental conductance technique [5]. The photovoltaic module operation depends strongly on the load characteristics, to which it is connected. Indeed, for a load, with an internal resistance Ri, the optimaladaptation occurs only at one particular operating point, called Maximum Power Point (MPP) andnoted in our case Pmax .Maximum power point tracking,frequently referred to as MPPT, is a system used to extract themaximum power of the PV module to deliver it to the loadand the efficiency is increased.
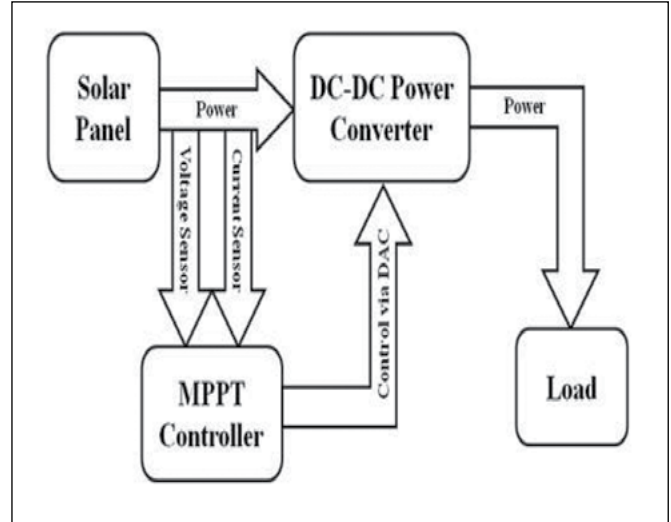

Figure.5: *Block Diagram of PV System*

IV. FUZZYCONTROLLER DESIGN

Fuzzy control is a method that allows the construction of nonlinear controllers from heuristic information that comes from the knowledge of an expert. Figure 8 shows the block diagram of a fuzzy controller [6]. The fuzzification block shown in fig.6 is responsible for processing the input signals and assign them a fuzzy value. The set of rules allows a linguistic description of the variables to be controlled and is based on the knowledge of the process. The inference mechanism is responsible for making an interpretation of the data taking into account the rules and their membership functions. With the defuzzification block, the fuzzy information coming from the inference mechanism is converted into non-fuzzy information that is useful for the process to be controlled. A fuzzy logic controller is shown in fig. 7.
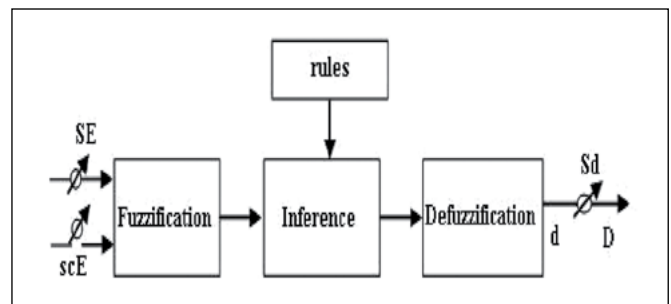

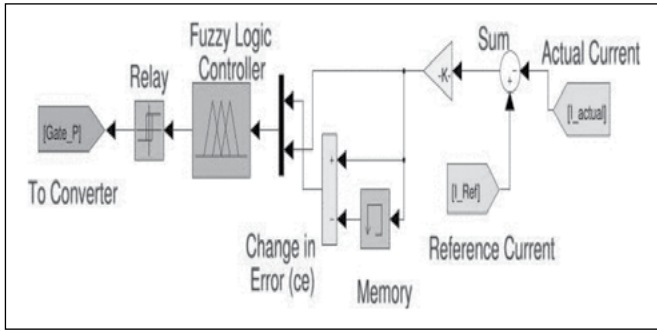**Figure.6:** *Block Diagram of Fuzzy Logic Algorithm*

Figure.7: Fuzzy controller

V. GENETIC ALGORITHM

Gbest-ABC Algorithm: GBest ABC is based on GA problem based on behaviour of honey bee. It incorporates honey bees into employed and unemployed bees (UB). Number of employed bees (EB) is equal to UB. 10 to 15% of UB are scout bees (SB) and rests are onlooker bees (OB). SB search new and abandoned food source (FS) randomly without prior knowledge and gather the information about the quality of FS. As soon as SB finds new FS, it becomes EB and carries nectar to hive. It deposits nectar hive and starts waggle dance to motivate OB [8]. Waggle dance is one of the important activities because updated food information is shared in the dancing area. OB stays in the hive, watches waggle dance and updates food quality information. Food quality generally depends on distance of FS from hive, amount of nectar and food quantity. OB always have option of either becoming SB, stay at hive to watch waggle dance and update information gathered by EB or become EB by motivation of one of the best options among many dancing bees. When EB return to hive with nectar, they also have choice to remain an employed bee, waggle dance and motivate OB to exploit FS or become OB. This process continues until food source is depleted and all employed bees become unemployed [8,9].

VI. RESULTS

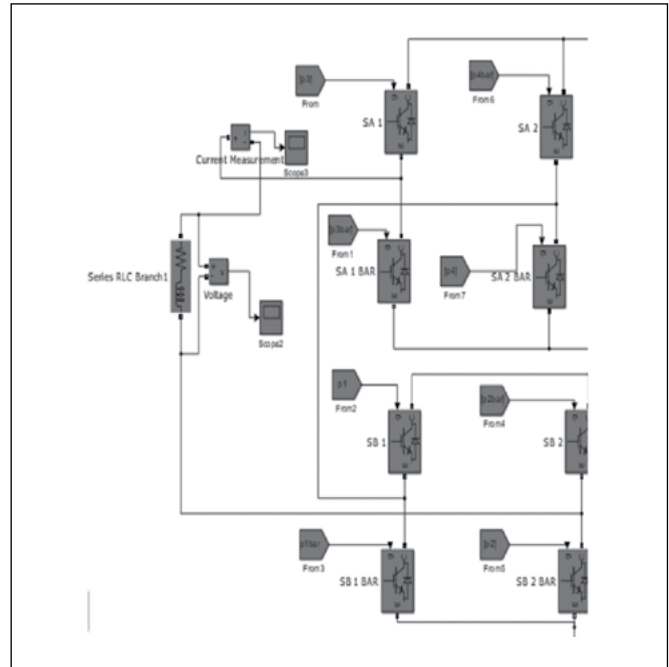Fig. 8 to 14 shows the results of the MATLAB simulation.
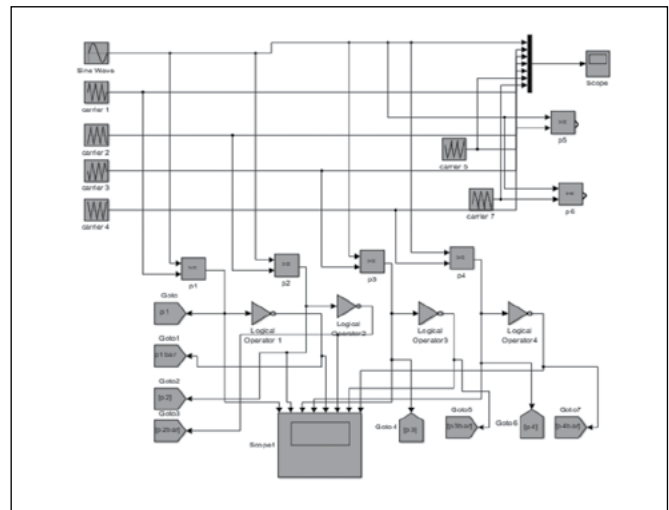

Figure.8: Multilevel Inverter
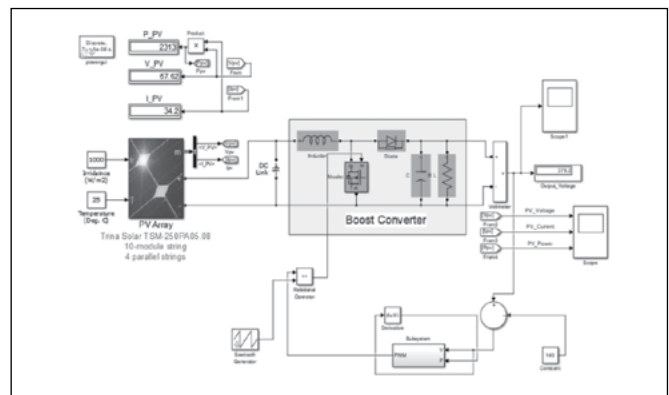

Figure.9: PWM Technique

Figure.10: PV Simulation
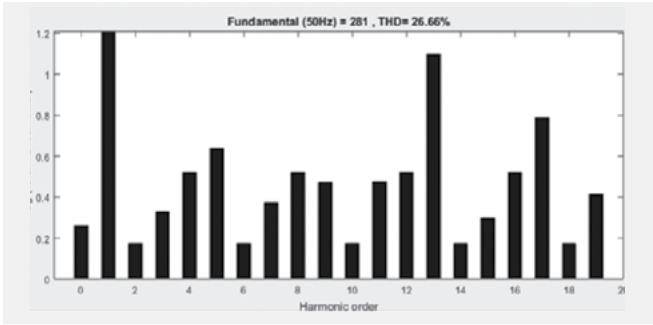
Figure.11: *THD Analysis*

Figure.12: *Waveform Of 5-Level Inverter*
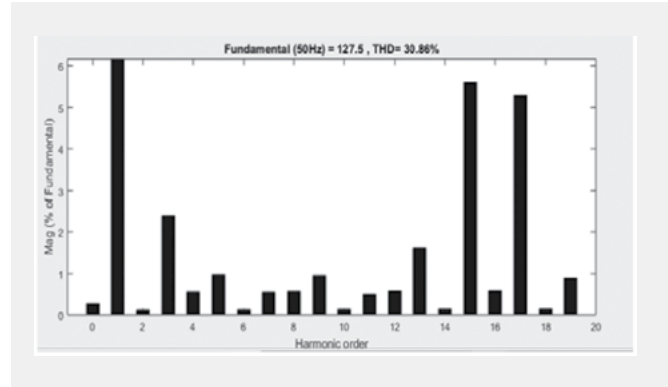
Figure.14: *Harmonic Spectrum*

VII. CONCLUSIONS

In our work we have simulated a H-bridge cascaded multilevel inverter. The THD level reduced to 26.66%. After connecting PV array module it became 26.62 %. For controlling purpose we have used Fuzzy logic controller, in which we got THD as 30.86%. The reason behind adding FLC is to maximize the output of boost converter which is given to the Multilevel Inverter. In future we are planning to integrate our model with Genetic Algorithm with MPPT. Also we will try to implement fuzzy logic in PWM signal given to MLI.

REFERENCES

[1] Keith Corzine, "A New Cascaded Multilevel H-Bridge Drive", Missouri University of Science and Technology, Yakov L. Familiant.

[2] AP. N. S. Chandana, S. Augusti Lindiya, K. Vijayarekha, D. Uma and M. Bharathi, "Analysis Of Various Pulse Width Modulations (Pwm) For Multi-Level Inverter With Reversing Voltage Topology", Department of Electrical Engineering, SASTRA University, India.

[3] G. Bhavanarayana, Chavvakula Swamy, Ganta Mounika, yedlapallinetaji, Sundru Sri durga, G. Bhavanarayana et al., "Analysis and hardware implementation of five level cascaded H Bridge inverter", Int. Journal of Engineering Research and Applications. ISSN: 2248-9622, Vol. 5, Issue 10, (Part - 2) October 2015, pp.43-53.

[5] Different PWM Waveforms Generation for 3-Phase AC Induction Motor with XC164CS, Infineon.

Figure.13: *13 Membership function of the GA*

[6] Photovoltaic cell (PDF). Fraunhofer ISE. 28 July 2014. Archived (PDF) from the original on 31 August 2014. Retrieved 31 August 2014.

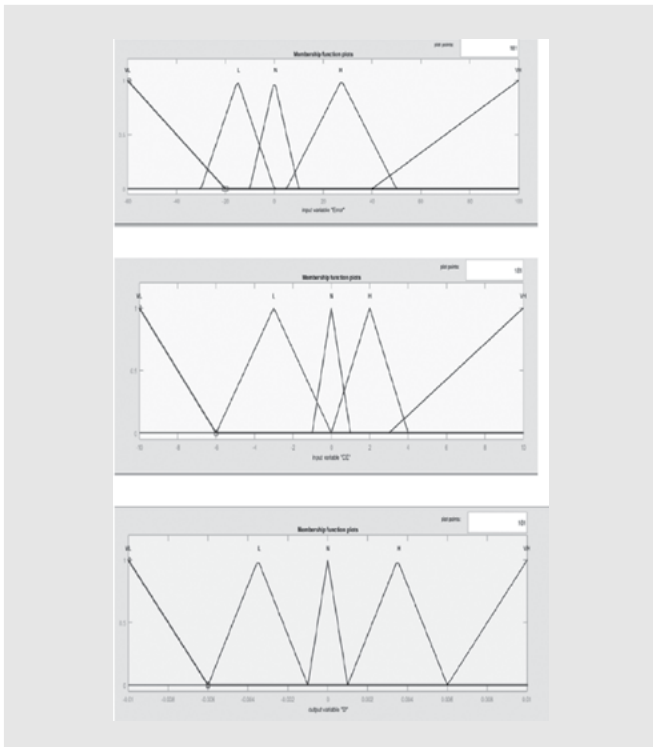[7] John Taborda Giraldo and Omar Rodríguez Álvarez Facultad de Ingeniería, "Fuzzy Logic Based MPPT Controller for a PV System Carlos Robles Algarín, Universidad del Magdalena, Carrera 32 No. 22-08, 470004 Santa Marta, Colombia.

[8] Prakash Kumar Amity University, Gaurav Jain, Rajasthan Technical University, Dheeraj Kumar Palwalia , "Genetic algorithm based maximum power tracking in solar power generation", Rajasthan Technical University.

[9]Sandeep Neupane1, Ajay Kumar2, "Modeling and Simulation of PV array in Matlab / Simulink for comparison of perturb and observe & incremental conductance algorithms using buck converter", Swami Vivekanand Subharti University, Meerut

Smruti Prangya Dash, Devidatta Pattnaik, Nihil Prashant, Nigam Mohanty, Abhijeet Pati, Riya Das

Dept. of Electrical and Electronics Engineering

Can room-temperature superconductors work without extreme pressure?

To find the next big superconductor, it helps to know where to start hunting. Scientists are using computer calculations to theoretically determine the structures and properties of materials and guide the search, theoretical chemist Eva Zurek said March 16 at the meeting. That strategy has paid off in the past. "Theory has played a very important role, in some cases predicting these structures before they were made," said Zurek, of the University at Buffalo in New York. For example, such predictions led researchers to a compound of lanthanum and hydrogen which was found in 2018 to be superconducting at then record-high temperatures up to about –13° Celsius (SN: 9/10/18).Now, predictions have guided scientists to superconductors made of yttrium and hydrogen, Dias reported March 18 at the APS meeting, in work done in collaboration with Zurek. Superconducting at up to about –11° C, Dias's yttrium-hydrogen superconductor is one of the highest-temperature superconductors known. While Dias's carbon, sulfur and hydrogen superconductor is still the temperature record-holder, the new material requires a significantly lower pressure - although at 182 gigapascals, it's still no simple squeeze. Dias and Zurek also reported their results March 19 in Physical Review Letters.

Fingerprint Based Voting Machine with RFID as Alternative

Abstract : The goal of this work is to enhance the overall performance in the Electronic Voting Machine (EVM) by the use of biometric based security system. Biometrics refers to technologies that measure and analyzes human body characteristics for authentication purposes. This technique is used to keep away from fraudulence in balloting systems. Our project "Biometric based EVM" uses the fingerprints as the security system for voting. Thumb effect of each and every man or woman is unique, it helps in minimizing the error. The fingerprint module stores the fingerprint templates and with the help of arduino it checks illegal votes and repetition of votes with correct coding. The system allows only authenticated voters to cast their votes based on his or her fingerprint. Apart from fingerprint security the EVM also supports Radio Frequency Identification (RFID) to help physically disabled peoples. This project reduces human involvement thereby time required for verification, voting and declaration of results becomes more faster and error free, which in turn increases the efficiency of the system. With this fingerprint based Electronic Voting Machine (EVM) the elections ought to be made honest and free from rigging.

Keywords:-Electronics Voting Machine, Biometrics, Fingerprint module, Radio Frequency Identification.

I. INTRODUCTION

Biometrics refers to technologies that measure and analyzes human body characteristics. Biometrics is the automated measurement of physiological or behavioral characteristics of individuals. Physiological characteristics include face, fingerprints, iris and retinal features, hand geometry, and ears.

An EVM consists of two units: the control unit and the balloting unit. These units are joined together by a cable. The control unit of the EVM is kept with the presiding officer or the polling officer. The balloting unit is kept within the voting compartment for electors to cast their votes. This is done to ensure that the polling officer verifies your identity. With the EVM, instead of issuing a ballot paper, the polling officer will press the Ballot Button which enables the voter to cast their vote. A list of candidate's names and/or symbols will be available on the machine with a blue button next to it. The voter can press the button next to the candidate's name they wish to vote for. An Electronic Voting Machine (EVM) is an electronic device used for recording votes and no one can change the program once the controller is manufactured.

The main drawback of this system is that, voter's id checking process is manual hence possibilities of illegal voting by a wrong candidate and also, possibility of multiple votes by same person. Currently the voting system in India is experiencing issues due to pole rigging and duplicity of votes. In rural areas the dominant parties with the help of other government officials try to control the voting and disrespect the unanimity of the voting process resulting in monopoly of majority party. People with physical disabilities also are facing a lot of difficult in casting their votes due to unavailability of any other way of enrollment. An advanced and secure way of vote enrollment and vote casting is of prime importance which could be conducted with minimal man power and interference of any other officials and can be efficient in securing already casted votes if anything goes astray.

To avoid this problem, in our proposed model we are using the fingerprint as an authentication medium for the voters and as an alternative we are using a RFID system in case the voter is not able to enroll his/her fingerprint for authentication. The fingerprint module stores the fingerprint templates and it also helps to check illegal votes and repetition of votes with correct coding. In our proposed model the results are stored in the EVM (controller memory) itself. The results can be viewed from the system itself when the voting process is over.

II. WORKING METHODOLOGY

A. DESIGN

In this paper we proposed a method which uses fingerprint and RFID number as a means of authentication for the enrollment of voters. The voter enrolls their fingerprints for authentication where the RFID is a means of alternative authentication for the voters when the voters are unable to enroll their fingerprint due to some reasons.

Usually the election process starts by enrollment process where the user enrolls their fingerprint or a RFID card as an identity proof. After getting successfully enrolled the voters identity are matched with the previously enrolled identity at the time of casting vote, if there is a match then the voters are allowed to cast their vote. The verification of user identity proof, that helps the voters to know that whether the particular voter can give their

vote in that area or not.

When this verification happens manually it takes too much time. If the voter's data is correctly loaded in the system then the voter can cast their vote. The manual checking can be eliminated totally, as the system will verify authorized RFID number and fingerprints authentication before he/she is going to cast their vote. For the verification and to cast their vote voter needs to provide their fingerprint after selecting an ID for their fingerprint. The system will not proceed to voting process unless the fingerprint is correct and which is found in the database of that system. If not pertaining to that polling booth, the rejection is indicated on the LCD. Once the voter Fingerprint is provided the authentication is done and the user can cast his/her vote. When the voter tries to cast multiple votes the system will check whether the voter has casted his/her vote earlier or not, if yes the voter is not allowed to cast his/her vote again, which is indicated by the LCD. The block diagram of the system is as shown below figure 1.
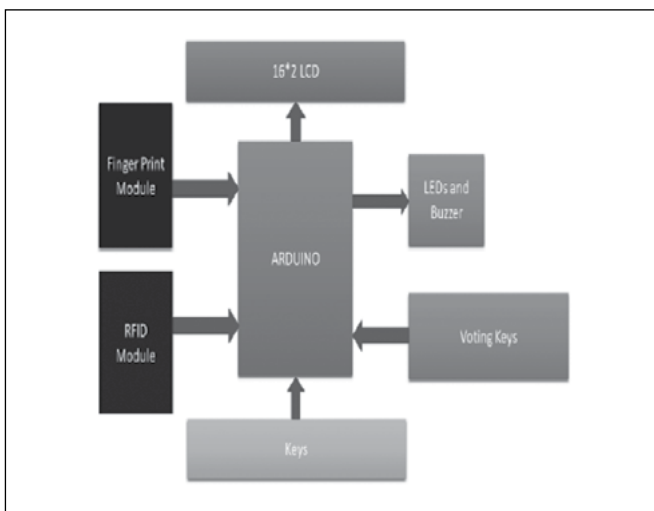

Figure 1: Block diagram of biometric EVM

B. METHODOLOGY

The proposed method uses a finger print module named R305 which can scan and store finger prints up to few hundred and the same can be transferred to database if the storage for finger prints is not enough within the module. The moduleworkswith the concept of optical scanning with the wave length of light in the range 6897 to 25,000 nanometers. The module can be able to do two different ways of searching 1:1 and 1:Ns searching by sending respective commands. The fingerprint module is interfaced with the microcontroller using USART communication. The system will undergo three

different modes of operation. They are a) Enrolling mode, b) Authentication and voting mode and c) Result announcement mode. Abbreviations and Acronyms

1) Enrolling Mode: First of all, user needs to enroll finger or voters with the help of push buttons/keys. To do this user need to press ENROLL key and then LCD asks for entering location/ID where finger will be a store. So now user needs to select an ID (Location) by using UP/DOWN keys. After selecting Location/ID user needs to press an OK key (DEL key). Now LCD will ask to select the means of enrollment which is by fingerprint or RFID card. After selecting the means of authentication the voter provides the fingerprint or RFID card as per the selection. If the voter selects finger print as the method of authentication he/she needs to put their finger over finger print module. Then LCD will ask to remove the finger from finger print module and again ask for placing the finger. Now user needs to put their finger again over finger print module. Now finger print module takes an image and converts it into templates and stores it by selected ID in to the finger print module's memory.

In case of RFID the system asks to scan a RFID card which provides a unique number in HEX format as an authentication number for the voter which is stored in the system. Now voters are be registered and he/she can vote. By same method all the voter can be registered into the system. Now if the user wants to remove or delete any of stored ID then he/she need to press DEL key, after pressing DEL key, LCD will ask to select between fingerprint and RFID, if the user selects fingerprint then the LCD asks to select the location means select ID that to be deleted. Now user needs to select ID and press OK key (same DEL key). Now LCD will let you know that finger has been deleted successfully. In case RFID is selected the LCD asks to scan the particular RFID card to be deleted and after successfully scanning the RFID card the card gets deleted from the system.

2) Authentication and Voting Mode: Now when user wants to vote then he/she needs to press match key and then buzzer will beep and LED will also glow and LCD will ask to place finger over fingerprint module or scan a RFID card for authentication. Now the system provides three attempts to scan your finger. After providing the finger print or the RFID card the system checks the provided identity with the identity

previously enrolled.If finger ID or the RFID card number is detected then LCD will show authorized voter. It means the user is authorized to vote. And then the system moves to next stage for voting. Now Green LED will glow it means now voter can vote for their candidates by pressing a re-elected key. Now if the same voter wants to vote again then it will be indicated by the LCD. Means same voter can't vote again. If any Non-registered user wants to vote then finger print module will not detect its ID into the system and it will be indicated by the LCD.

3) Result Announcement mode: This mode is also used by the election commission to announce the result of a particular area. There no need of a special procedure to enter into this mode. After the voting process ends the results are calculated by the system. The system itself calculated the number of voters enrolled and the votes acquired by the candidates. When the official wants to check the results they press the result key and the system displays the number of votes for each individual candidate and declares the winner of the election.

The flowchart for the above operations which include enrolling voters, authentication, casting of votes and result declaration are shown below in figure 2 & 3.
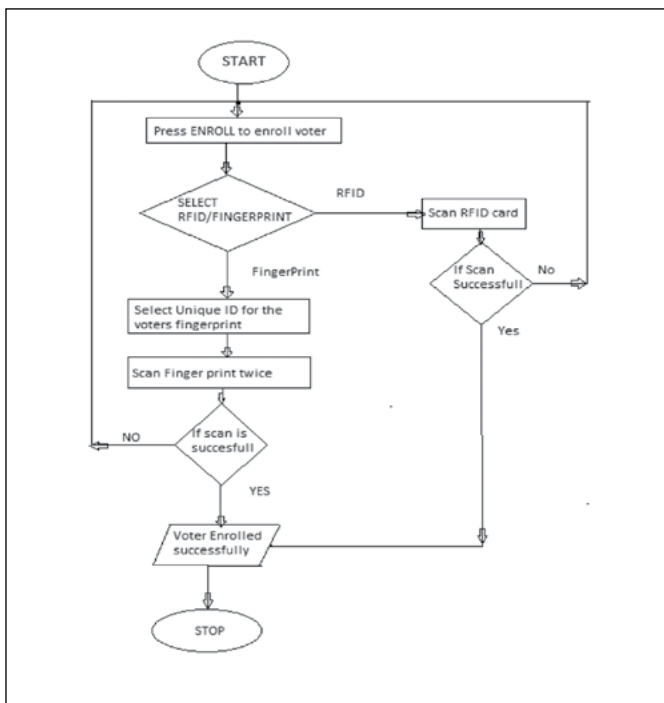

Figure 2: *Flowchart for Enrolment process*
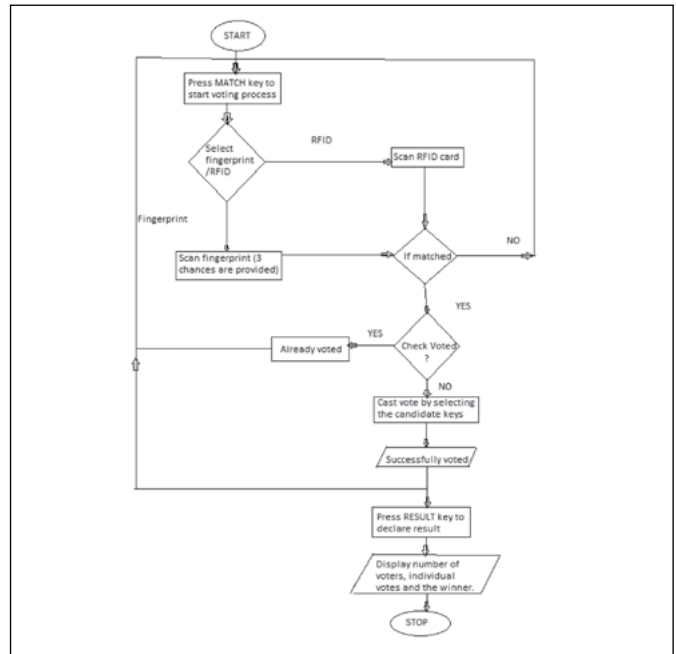
Figure 3: *Flowchart for Authentication process, voting process and result declaration.*

III.RESULT ANALYSIS

Fingerprint verification could also be an honest choice for in e-voting systems, where the system operates in a controlled environment. The voting time is drastically reduced compared to existing methods. The voting time for a voter is reduced to 5 to 10 seconds as the system is able to perform authentication and voting process at the same time.

In this proposed method the voters are enrolled in their respective regions using their finger print or RFID card as a means of authentication. These individual machines are going to serve in their designated polling booths.

Here the enrolled fingerprint or the RFID card number is checked by the system for the eligibility of the voter to cast his/her vote. If the voter is authorized then the voter is allowed to cast his/her voter to their selected candidate by using the re-elected keys. The system doesn't allow multiple votes or invalid votes at the time of elections.

Figure 4: *Prototype of Fingerprint based voting machine*

Figure 5: Enrolment process

Figure 6: Authentication and voting process

Figure 7: Result Declaration

IV. CONCLUSIONS

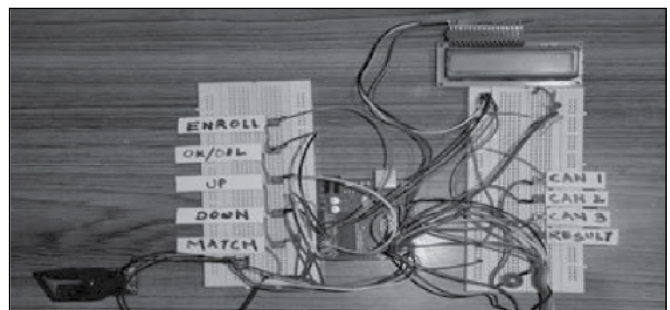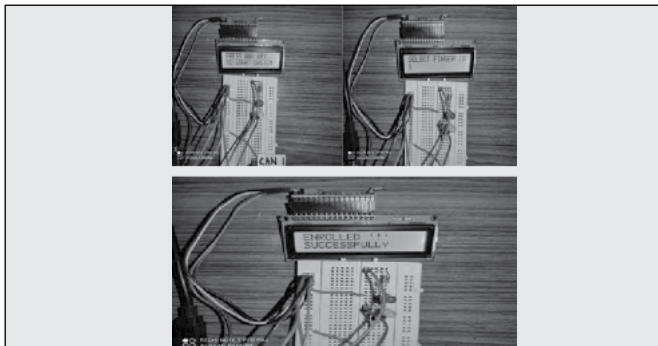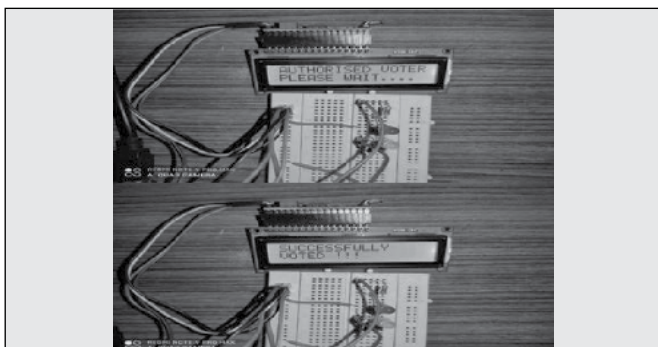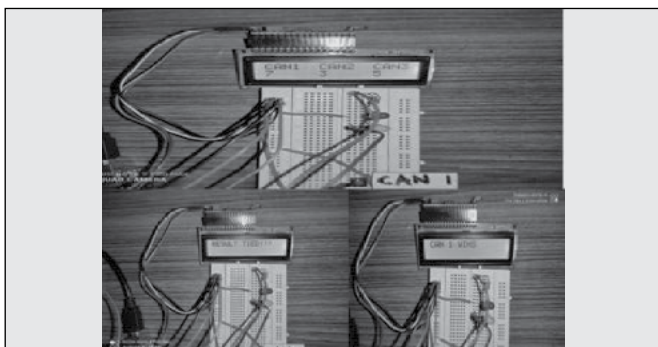The project "Biometric Based Voting Machine" was mainly intended to develop a fingerprint and RFID based advanced Electronic Voting Machine (EVM) which helps in free and fair way of conducting elections. The system allows only authenticated voting than the existing equipment as the person is identified based on his Fingerprint which is unique to each individual. The inclusion of RFID system makes the project more robust by providing an alternate in case a voter is unable to enroll by fingerprint. It is time conscious, less time required for voting & counting and it avoids invalid voting as it prevents unregistered voters from voting. It is ease to transport due to its compact size.

REFERENCES

[1] Meena Chavan, Devanshu Saxena, Mayank Mukherjee, Udhav Varma, "Biometric Based Electronic Voting Machine (EVM)", International Research Journal of Engineering and Technology (IRJET) , Volume: 06, Issue: 06, June 2019

[2] Sanjay Kumar, Manpreet Singh, "Design of a Secure Electronic Voting System using Fingerprint Technique", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 4, No. 1, July 2013

[3] Firas I.Hazzaa, Seifedine Kadry, Oussama Kassem Zein, Web-Based Voting System Using Fingerprint: Design and Implementation, International Journal of computer Applications in Engineering Sciences, VOL II, ISSUE IV, DECEMBER 2012.

[4] João Palas Nogueira and Filipe de Sá-Soares, "Knowledge and Technologies in Innovative Information Systems", Trust in e-Voting Systems, Vol. 129, Pg [51-66], May 2012

[5] Ashok Kumar D.,Ummal Sariba Begum T., A Novel design of Electronic voting system Using Fingerprintdz, International Journal of Innovative Technology & Creative Engineering (ISSN: 2045-8711), Vol.1, No.1.pp:12 19, January 2011.

 [6] Xiao-dan Wu, Yun-feng Wang, Jun-bo Bai, Hai-yan Wang and Chao-Hsien Chu, "RFID application challenges and risk analysis", Industrial Engineering and Engineering Management (IE&EM) 2010 IEEE 17th International Conference, pp. 1086-1090, 29–31 Oct. 2010.

[7] R. Haenni, E. Dubuis, and U. Ultes-Nitsche, "Research on e-voting technologies" Bern University of Applied Sciences, Technical Report 5, 2008.

Aniket Pattnaik
Bibek Anand Moharana
Rachit Pattnaik
Dibyadarshee Pattnaik
Jivitesh Surjit

Dept. of ECE

Twitter Sentiments Analysis using Machine Learning Techniques

Abstract : *Sentiment Analysis is the process of computationally determining whether a piece of writing is positive, negative or neutral. However, human language is complex. Teaching a machine to analyze the various grammatical nuances, cultural variations, and misspellings that occur in social media is a difficult process.Thiswork focuses on designing a tweet classifier based on sentiment analysis using machine learning. It reports on the design of sentiment analysis extracting a required amount of tweets. Results classify customer's perspective via tweets into positive negative or neutral.*

Keywords: Twitter, sentiment analysis, Machine learning, Natural language processing.

I. INTRODUCTION

Micro-blogging today has become a very popular communication tool among Internet users. Millions of messages are appearing daily in popular web-sites that provide services for micro-blogging such as Twitter. Authors of those messages write about their life, share opinions on variety of topics and discuss current issues [1]. Because of a free format of messages and an easy accessibility of micro-blogging platforms, Internet users tendto shift from traditionalcommunication tools (such as traditional blogs or mailing lists)to micro-blogging services [2].Twitter is an American micro-blogging and social networking service on which users can post and interact with messages known as "tweets" [3]. Twitter Sentiment Analysis has been traditionally associated with the security sector but today there is active expansion into other industries including retail, marketing and health [4].

Sentiment Analysis of tweets helps in different sectorsto enhance marketing strategies. Companies consider micro-blogging platform an essential channel for marketing strategy and customer service.Twitter allows business to reach a broad audience and connect with customers without intermediaries[5]. Monitoring Twitter allows companies to understand their audience, keep on top of what's being said about their brand and competitors.It can be used for customer service by spotting dissatisfaction or problem with products also helps in finding people that are happy with your product and can be used to promote the products.

As more and more users post about products and services they use,or express their political and religious views, micro-blogging web-sites become valuable sources of people's opinions and sentiments. Such data can be efficiently used for marketing or social studies. Twitter contains a very large number of very short messages created by the users of this micro-blogging platform.The contents of the messages vary from personal thoughts to public statements. Data from these sources can be used in opinion mining and sentiment analysis tasks.For example, manufacturing companies may be interested in the following questions:

- What do people think about our product (service,company etc.)?

- How positive(or negative)are people about our product?

- What would people prefer our product to be like?

The amount of relevant data is much larger for twitter, as compared to traditional blogging sites. Moreover the response on twitter is more prompt and also more general (since the number of users who tweet is substantially more than those who write web blogs on a daily basis). Sentiment analysis of public is highly critical in macro-scale socioeconomic phenomena like predicting the stock market rate of a particular firm. This could be done by analyzing overall public sentiment towards that firm with respect to time and using economics tools for finding the correlation between public sentiment and the firm's stock market value. Firms can also estimate how well their product is responding in the market, which areas of the market is it having a favorable response and in which a negative response (since twitter allows us to download

stream of geo-tagged tweets for particular locations. If firms can get this information they can analyze the reasons behind geographically differentiated response, and so they can market their product in a more optimized manner by looking for appropriate solutions like creating suitable market segments. Predicting the results of popular political elections and polls is also an emerging application to sentiment analysis. One such study was conducted by N. Azzouza et al. in Germany for predicting the outcome of federal elections in which concluded that twitter is a good reflection of offline sentiment [1].

This work focuses on building a model for Twitter sentiment analysis.Twitter Authentication is an integral part for collecting the data.In order to fetch tweets through twitter API,one needs to register an App through a twitter account. User can fetch any number of tweets and the tweets fetched are in raw format i.e. they are user specific since twitter has no particular format of posting tweets. After the tweets are fetched the job is to process the tweets and remove the terms that are not going to result towards the accuracy or classification in any manner. It could be a challenging task because some users post in their local language that in turn becomes tricky to process. Given a tweet it's pretty clear that not all words in a tweet are useful for classifying it into one of the three categories:positive, negative or neutral [6]. An overview of the Twitter sentiment analysis application is shown in Fig. 1.

Figure.1: *Twitter sentiment analysis*

Given a tweet, the objective of the application is to classify whether the message is of positive, negative, or neutral sentiment. For messages conveying both a positive and negative sentiment, whichever is the stronger sentiment should be chosen.The focus here is on using advanced text mining techniques to analyze the sentiment of the text (in the form of positive, negative and neutral). It is primarily for analyzing conversations, opinions, and sharing of views (all in the form of tweets) for deciding business strategy, political analysis, and also for assessing public actions.

II. LIBRARIES USED

The general libraries used for designing the application are explained below:

A. Tweepy:

Tweepy is a Python library for accessing the Twitter API. It is great for simple automation and creating twitter bots. Tweepy has many featuressuch as : Get tweets from timeline, Creating and deleting Tweets, Follow and unfollow users, etc

B. Nltk:

The Natural Language Toolkit (NLTK) is a suite of libraries and programs for symbolic and statistical natural language processing.It supports classification, tokenization, stemming, tagging, parsing, and semantic reasoning functionalities.

C. Sklearn:

Sklearn is a python module integrating classical machine learning algorithms in the tightly-knit world of scientific Python packages (numpy, scipy, matplotlib). It aims to provide simple and efficient solutions to learning problems that are accessible to everybody and reusable in various contexts: machine-learning as a versatile tool for science and engineering.

D. Numpy:

Numpy is a general-purpose array-processing package. It provides a high-performance multidimensional array object, and tools for working with these arrays. It is the fundamental package for scientific computing with Python. Besides its obvious scientific uses, Numpy can also be used as an efficient multi-dimensional container of generic data.

E. Textblob:

TextBlob is a python library and offers a simple API to access its methods and perform basic NLP task such as part-of-speech tagging, noun, phrase extraction, sentiment analysis, classification, translation and many more.In this work,Textblob is used for sampling the data after pre-processing step.

III. DATA PREPROCESSING:

Raw tweets scraped from twitter generally result in a noisy and obscure dataset. This is due to the casual and ingenious nature of people's usage of social media. Tweets have certain special characteristics such as re-tweets, emoticons, user mentions, etc. which should be suitably extracted. Therefore, raw twitter data must be normalized to create a dataset which can be easily learned by various classifiers. An extensive number of pre-processing steps were applied to standardize the dataset and reduce its size as discussed below.

A. Filtering:

Filtering involves cleaning of raw data. In this step, URL links (E.g. http://twitter.com), special words in twitter (e.g. "RT" which means Re-Tweet), user names in twitter (e.g. @Ron -@ symbol indicating a user name), emoticons are removed[6].

B. Tokenization:

Tokenization focuses on segmentation of sentences. In this step, the text is tokenized or segmented with the help of splitting text by spaces and punctuation marks to form container of words.

C. Removal of Stop words:

Articles such as "a", "an", "the" and other stop words such as "to", "of", "is", "are", "this", "for"are removed in this step.

D. Removing special Character:

In this step special character like '$', '@', '+' etc. which are not going to contribute to the emotions of the text written are removed.

E. Stemminig:

Stemming techniques put word variations like "great", "greatly", "greatest", and "greater" all into one bucket, effectively decreasing entropy and increasing the relevance of the concept of "great". In other words, Stemming allows considering in the same way nouns, verbs and adverbs that have the same radix.

F. Assigning polarity

TextBlob used here works by assigning polarity to the fetched tweets.All the words in each tweet are converted into lower case and compared to predefined 'Stop words' corpus and filtered out.So a data of the form [sentiment, tweet] is prepared and is stored

under a file. Data parsing involves the '@' removal, regex,URL substitution and sentence reduction processes.

IV. CLASSIFIERS USED:

A. SVM Classifier:

SVM was originally designed for binary classification. It can also be extended to multi-class by combining multiples SVMs. It gives an excellent result for text categorization tasks such as sentiment analysis. Here the binary SVM is considered for simplicity. SVM performs classification by finding an optimal hyper-plane that separates two Data Collection Data Preprocessing Feature Extraction Classifier Result classes. The optimal hyper-plane has maximum margin. The distance between nearest data point and hyper-plane is called as margin. The point that lies nearest to hyper-plane is called support vector as shown in fig.2.
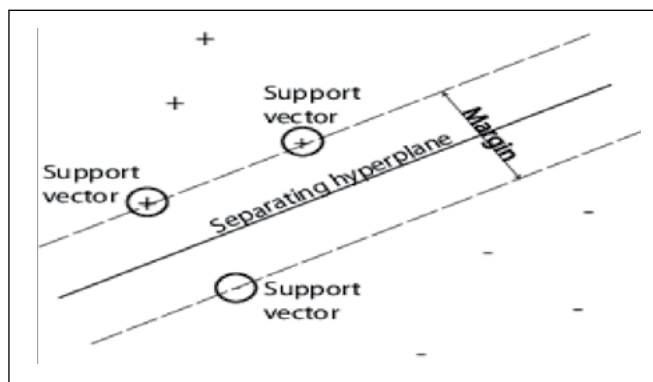
Figure.2: *Support Vector*

IV. CLASSIFIERS USED:

B. K-Nearest Neighbors (KNN) :

The K-Nearest Neighbors (KNN) is also a supervised learning algorithm based on the classes nearest to the point which is to be classified. Based on the values of the K-nearest classes a test set is provided the majority voting class. However, to improve this algorithm weights are assigned to each of the k points according to their distance from the test point. The value of k depends upon the classification problem and the size of dataset.

The prediction probability of SVM and KNN on each test tweet is considered and the class is assigned based on the greater probability.

C. Naïve Bayes Classifier:

After data preparation is done, the next step is data training. A basic Naïve Bayes classifier is built to fit with the data under 'Train.csv'. For the testing purpose, 30% of the dataset is stored under 'Test.csv'. Naive Bayes Classifier assigns the most probable class to the tweet out of all the possible classes.

D. Decision Tree:

The tree represents a test on the attribute of the data set, and its children represent the outcomes. The leaf node represents the final classes of the data points. It is a supervised classifier model which uses data with known labels to form the decision tree and then the model is applied on the test data [4]. For each node in the tree the best test condition or decision has P to be taken. We use the GINI factor to decide the best split.For a given node t, where, p(j | t) is the relative frequency of class 'j' at node 't'.

E. AdaBoost:

Boosting is an ensemble modeling technique which attempts to build a strong classifier from the number of weak classifiers.AdaBoost is short for AdaptiveBoosting and is a very popular boosting technique which combines multiple "weak classifiers" into a single "strong classifier".

Algorithm:

1. Initialize the dataset and assign equal weight to each of the data point.

2. Provide this as input to the model and identify the wrongly classified data points.

3. Increase the weight of the wrongly classified data points.

4. if (got required results)

   Goto step 5

   else

   Goto step 2

5. End.

F. Logistic Regression:

It is used to determine the output or result when there are one or more thanone independent variables. Logistic regression is a regression model. The model builds a regression model to predict the probability that a given data entry belongs to the category numbered as "1". Just like Linear regression assumes that the data follows a linear function, Logistic regression models the data using the sigmoid function shown in fig.3.
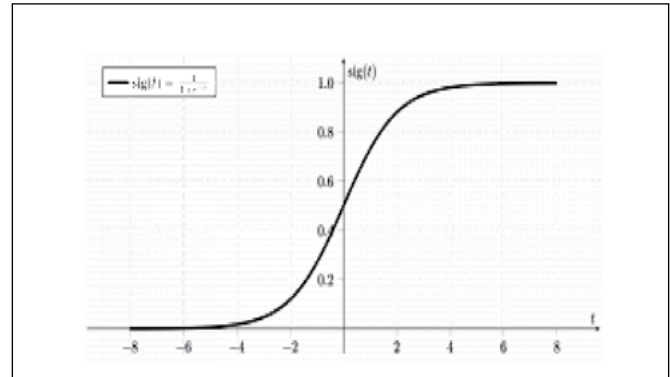

Figure.3. *Sigmoid function*

Logistic regression becomes a classification technique only when a decision threshold is brought into the picture. The setting of the threshold value is a very important aspect of Logistic regression and is dependent on the classification problem itself.The decision for the value of the threshold value is majorly affected by the values of precision and recall. Ideally, we want both precision and recall to be 1, but this seldom is the case. In case of a Precision-Recall trade-off we use the following arguments to decide upon the threshold:

a) Low Precision/High Recall: In applications where we want to reduce the number of false negatives without necessarily reducing the number false positives, we choose a decision value which has a low value of Precision or high value of Recall. For example, in a cancer diagnosis application, we do not want any affected patient to be classified as not affected without giving much heed to if the patient is being wrongfully diagnosed with cancer. This is because, the absence of cancer can be detected by further medical diseases but the presence of the disease cannot be detected in an already rejected candidate.

b) High Precision/Low Recall: In applications where we want to reduce the number of false positives without necessarily reducing the number false negatives, we choose a decision value which has a high value of Precision or low value of Recall. For example, if we are classifying customers whether they will react positively or negatively to a personalised advertisement, we want to be absolutely sure that the customer will react positively to the advertisement because otherwise, a negative reaction can cause a loss potential sale from the customer.

G. Random Forest:

The random forest classifier was chosen due to superior performance over a single decision tree with respect to accuracy. It is essentially an ensemble method based on bagging. The classifier works as follows. The classifier firstly creates k bootstrap samples of Di with each of the samples denoting as Di . A Di has the same number of tuples. By sampling with replacement, it means that some of the original tuples of D may not be included in Di, whereas others may occur more than once.

The classifier then construct a decision tree based on each Di. As a result, a "forest" that consists of k decision tree formed. To classify an unknown tuple, X, each tree returns its class prediction counting as one vote. The final decision of X's class is assigned to the one that has the most votes[2].

V. K-FOLD CROSS VALIDATION AND RESULTS

We split the data set into k parts, hold out one, combine the others and train on them, then validate against the held out portion. We repeat that process k times (each fold), holding out a different portion each time. Then we average the score measured for each fold to get a more accurate estimation of our model's performance.
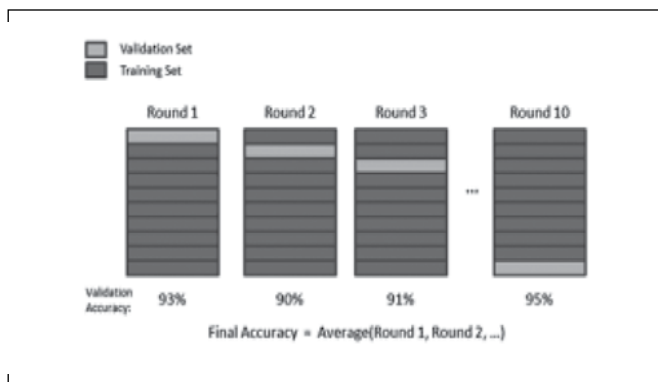

Figure 4: *Fold Cross Validation*

We split the training data into 10 folds and cross validate on them using scikit-learn as shown in the fig. 4. The number of K¬-folds is arbitrary and usually set to 10 it is not a rule. In fact, determine the best K is still an unsolved problem but with lower K: computationally cheaper, less variance, more bias. With large K: computationally expensive, higher variance, lower bias. We can now train the SVM classifier with the training set, validate it using the hold out part of data taken from the training set, the validation set, repeat this 10

times and average the results to get the final accuracy which is about 0.785.

Noticed that to evaluate our classifier, the F1 score and confusion matrix are used. The F1 Scorecan be interpreted as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and worst score at 0. It is a measure of a classifier's accuracy. The F1 score formula is given in equation(1). Here, the precision is the number of true positives (TP) (the number of items correctly labeled as belonging to the positive class) divided by the total number of elements labeled as belonging to the positive class given in equation (2). The formula for recall is the number of true positives divided by the total number of elements that actually belong to the positive class given in equation (3).

$$F1 = \frac{2 * Precision * Recall}{Precison + Recall} \quad \cdots\cdots\cdots\cdots (1)$$

$$Precision = \frac{TP}{TP + FP} \cdots\cdots\cdots\cdots (2)$$

$$Recall = \frac{TP}{TP + FN} \quad \cdots\cdots\cdots\cdots\cdots (3)$$

A precision score of 1.0 means that every result retrieved was relevant (but says nothing about whether all relevant elements were retrieved) whereas a recall score of 1.0 means that all relevant documents were retrieved (but says nothing about how many irrelevant documents were also retrieved).

There is a tradeoff between precision and recallwhere increasing one decrease the other and we usually use measures that combine precision and recall such as F-measure or MCC. A confusion matrixhelps to visualize how the model did during the classification and evaluate its accuracy. In our case we get about 156715 false positive tweets and 139132 false negative tweets. It is "about" because these numbers can vary depending on how we shuffle our data for example. The example of confusion matrix is shown in fig.5. The confusion matrix of the naive SVM can be expressed using a color map where dark colors represent high values and light colors represent lower values as shown in the corresponding color map of the SVMin figure. 6.

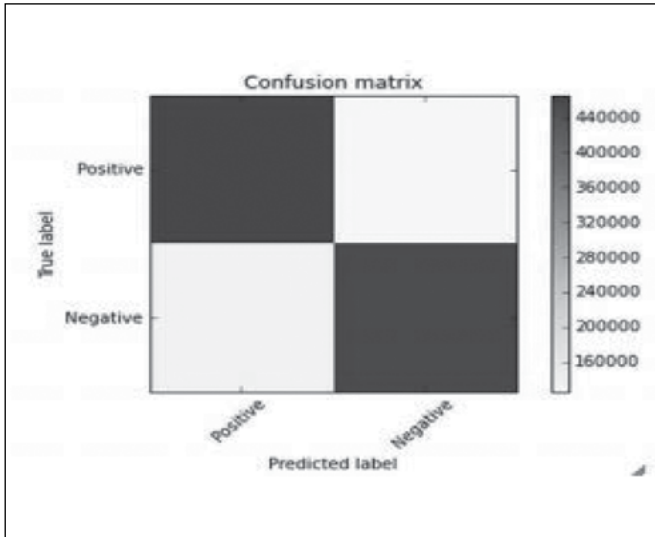**Figure. 5:** *Example of confusion matrix*

Figure. 6: Colormap of the confusion matrix

Hopefully we can distinguish that the number of true positive and true negative classified tweets is higher than the number of false and positive and negative tweets. However from this result we try to improve the accuracy of the classifier by experimenting different techniques and we repeat the same process using the k-fold cross validation to evaluate its averaged accuracy.

Oversampling:

In Machine Learning and Data Science we often come across a term called Imbalanced Data Distribution, generally happens when observations in one of the class are much higher or lower than the other classes. As Machine Learning algorithms tend to increase accuracy by reducing the error, they do not consider the class distribution. This problem is prevalent in examples such as Fraud Detection, Anomaly Detection, Facial recognition etc.SMOTE (synthetic minority oversampling technique) is one of the most commonly used oversampling methods to solve the imbalance problem.It aims to balance class distribution by randomly increasing minority class examples by replicating them. SMOTE synthesis new minority instances between existing minority instances. It generates the virtual training records by linear interpolation for the minority class. These synthetic training records are generated by randomly selecting one or more of the k-nearest neighbors for each example in the minority class. After the oversampling process, the data is reconstructed and several classification models can be applied for the

processed data.Our model classifies the given tweets to either positive, negative or neutral sentiments.To do the sentiment analysis of tweets, the proposed system first extracts the twitter posts from twitter by user. The system can also compute the frequency of each term in tweet. Using machine learning supervised approach help to obtain the results. The screen shorts of the results obtained on four classification algorithms is shown in fig. 7. In this work various types of classifier were explored which are applied on real time tweets.If a topic is provided, it will fetch the real time tweet based on the topic and classify it.Out of all the classifiers that are used for this work, ensemble techniques,Random Forest and adaboost are giving promisingly better accuracy however it is slow in performance compared to the SVM classifier.
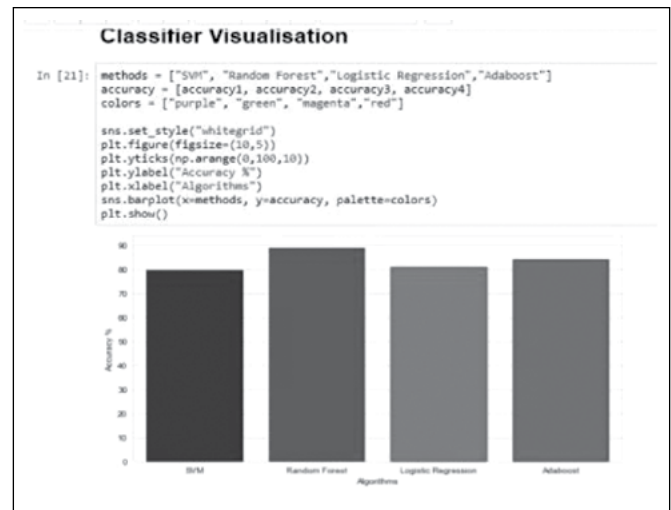
Figure. 7: Classifiers accuracy

VI. CONCLUSIONS

Nowadays, sentiment analysis or opinion mining is a hot topic in machine learning. We are still far from detecting the sentiments of scorpus of texts very accurately because of the complexity in the English language and even more if we consider other languages such as Chinese. In this work we tried to show the basic way of classifying tweets into positive or negative category using different classifier and how language models are related to all the classifiers and can produce better results. We could further improve our classifier by trying to extract more features from the tweets, trying different kinds of features, tuning the parameters of the all the classifiers, or trying another classifier all together.

REFERENCES

1. N. Azzouza, K. Akli-Astouati, A. Ou ssalah, and S. A. Bachir, "A real-time Twitter sentiment analysis using an unsupervised method," in Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics, 2017: ACM, p. 15.

2. P.-W. Liang and B.-R. Dai, "Opinion mining on sociamedia data," in Mobile Data Management (MDM), 2013 IEEE 14th International Conference on,2013,vol.2.IEEE, pp. 91-96.

3. N. F. F. da Silva, E. R. Hruschka, and E. R. Hruschka, "Tweet sentiment analysis with classifier ensembles," Decision Support Systems, vol. 66,pp. 170-179, 2014/10/01/ 2014.

4. T. Chalothom and J. Ellman, "Simple Approaches of Sentiment Analysis via Ensemble Learning," Berlin, Heidelberg, 2015: Springer Berlin Heidelberg, pp. 631-639.

5. J. Lin and A. Kolcz, "Large-scale machine learning at twitter," in Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, 2012: ACM, pp. 793-804.

6. E. Haddi, X. Liu, Y. Shi, "The role of text pre-processing in sentiment analysis", Proc. Comp. Sci. 17, pp. 26–32, 2013.

Kumari Madhu
Shalini Kumari
Disha Kumar
Karunesh Kumar

8th Sem CSE

Most of Mars' missing water may lurk in its crust

Planetary scientist Eva Scheller of Caltech and colleagues simulated possible scenarios for water loss on Mars, based on observations of the Red Planet made by rovers and orbiting spacecraft, and lab analyses of Martian meteorites. These simulations accounted for possible water loss to space and into the planet's crust through bodies of water or groundwater interacting with rock.In order for the simulations to match how much water was on Mars 4 billion years ago, how much is left in polar ice caps today and the observed abundance of hydrogen in Mars' atmosphere, 30 to 99 percent of Mars' ancient water must be stashed away inside its crust. The rest was lost to space.Mars' underground water could be mined by future explorers, says Jack Mustard, a planetary geologist at Brown University in Providence, R.I., not involved in the work. The most easily accessible water on Mars may be at its polar ice caps (SN: 9/28/20). But "to get the ice, you've got to go up to [high latitudes] - kind of cold, harder to live there," Mustard says. If water can be extracted from minerals, it could support human colonies at warmer climes closer to the equator.

The Impact of Bioinformatics in the Development of Vaccines Against Covid-19

Severe Acute Respiratory Syndrome Corona virus 2 (SARS-CoV-2),also known as the novel Corona virus is the causative agent for the Corona virus Disease 2019 (COVID-19). Since its first detection in December 2019 the disease has engulfed almost entire world by spreading over more than 200 countries and has taken over 2 million lives. This highly infectious virus spread via respiratory droplets and aerosols when an uninfected person comes in contact with some one who was infected. While masks and proper sanitation practices like constant washing of hands, use of alcohol-based sanitizer and social distancing helped to some extent, the disease could not be battled with without a proper cure or an efficient vaccine. Therefore, researchers around the world have started collaborating and sharing their research data so that with collective efforts, a cure for the Covid-19 could be developed as soon as possible. Bioinformatics surely came out as one of the essential tools to analyse viral data as important information about the genetic makeup of the virus was provided by it and it also assisted directly in the development of drugs or vaccines against the deadly disease. Bioinformatics is one of the reasons why scientists were able to develop a vaccine so quickly.

Bioinformatics significantly reduces the cost of research by narrowing down hundreds of options to a few plausible directions. Scientists upload their findings into a database. A large amount of data can be analysed by bioinformatics to come up with a hypothesis that can be further validated in the laboratory by doing experiments. Instead of doing a bunch of experiments without a clear goal, only a few could be tested as predicted by bioinformatics by analysing the available data. That way a researcher can save a lot of resources and manpower while obtaining results efficiently. In the current COVID-19 crisis, bioinformatics is playing a crucial role by predicting suitable drugs to cure viral targets and subsequently these are now being tested in the laboratory for validation.

In development of a potential vaccine for this deadly disease, bioinformatics helped by predicting possible molecules as inhibitors for the virus that could be tested in the laboratory for efficacy. Biotechnologists were able to use certain computation tools and self-created algorithms to discover certain molecules that could be used as drugs against the disease. From analysing the sequence to the final drug candidate prediction requires several steps to follow which can be broadly divided into five segments:

- Retrieval of the viral sequence from the nucleic acid database
- Analysis of the sequence data by comparison with other viral sequences
- Phylogenetic analysis of the viral sequences to find out how the target virus evolved from others
- Computational modelling of the important viral proteins as drug targets
- Testing several drug molecules against the viral proteins for inhibitory activities.

Immediately after the discovery of the SARS-CoV-2 in Wuhan city of China in December 2019, scientists experimentally determined the DNA sequence of the virus and deposited that in the public database so that others can access it freely and work on it. Biotechnologists downloaded the sequence from the database and started analysing the sequence to extract useful information about the virus as this is the first step in the drug development against COVID-19. This sequence was compared with other viruses, especially other coronaviruses, and found that it has sequence similarity with SARS-CoV which was responsible for the previous SARS outbreak in 2003. Also, this sequence was found very similar to a corona virus found in bats and from that information it was predicted that the virus was probably originated in bats. Subsequently acquiring favourable mutations or changes in its genome the virus jumped to humans from bats thorough a pangolin intermediate – all this information was obtained by analysing viral sequences which again indicates the significance of Bioinformatics.

Once the genetic information of SARS-CoV-2 was analysed to find out genes that it holds and the type of proteins it encodes, scientists generated theoretical models of all the important proteins of the virus. If one can inhibit those essential proteins of the virus, the virus could be stopped from further infecting others. In the absence of any experimental structures of the viral proteins, theoretical models can be built by a process called homology modelling where if structures of similar proteins from other viruses are available those can be used as templates to generate theoretical structures of target proteins. Structures of the proteins are important for designing drugs against those, and therefore, scientists built theoretical models of important

viral proteins that can be targeted by inhibitor molecules.

The final step in the process is to use existing drug molecules or modify the structures of existing drug molecules and 'dock' them against the viral proteins to see whether these molecules are binding to the important sites on the viral proteins or not by computational analysis. If some molecules are found which can bind to the viral proteins with high affinity, then those molecules can be further tested in the laboratory to find out their effectiveness. This way Bioinformatics has helped immensely to design different vaccine candidates prepare a vaccine against the SARS-CoV-2 and stop COVID-19 from further spreading.
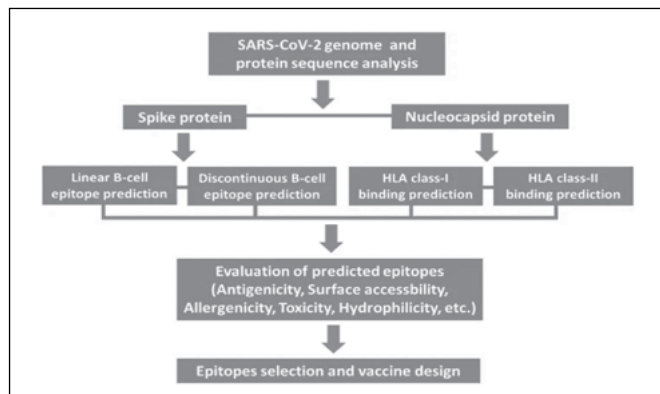
Study workflow. Suitable proteins of SARS-CoV-2 were selected at the first step for epitope prediction. The second step comprised of B- and T-cell epitope analysis with bioinformatics approaches. Epitope evaluation was followed and appropriate ones were chosen for vaccine design

Source: *https://idpjournal.biomedcentral.com/ articles/10.1186/s40249-020-00713-3*

3D structures visualized by UCSF Chimera: Image show the most promising peptides in the envelope protein of COVID-19 (yellow colored) binding to MHC class I

Source: *https://www.hindawi.com/journals/ bmri/2020/2683286/*

Phosphorus for earth's earliest life may have been forged by lightning

The phosphorus that went into building the first DNA and RNA molecules is thought to have come from a mineral called schreibersite, which is typically found in meteorites (SN: 9/7/04). But a new analysis of minerals forged by a lightning strike hints that lightning may have produced enough schreibersite on early Earth to help kickstart life, researchers report online March 16 in Nature Communications.From about 4.5 billion to 3.5 billion years ago, lightning could have forged 110 to 11,000 kilograms of phosphorus-containing compounds each year to help seed life, the researchers conclude.

A Design Approach for FIR Filter using Functionally Linked Cat Swarm Optimization Algorithm

Abstract : *Nowadays, the signal carries all sort of valuable information from one point to another part of the globe. While the signal is transmitted distortion takes place due to the non-ideal nature of the channel. Hence frequent filtering is a basic need in most communication systems. This article analyzes the optimal design approach of modern digital FIR filters; in particular, giving importance to the design of bandstop filters (BSFs) the most difficult one, using artificial swarm intelligence. The linear phase BSF is predominantly measured as a complex, constraint, and highly multimodal engineering problem. The classical optimization procedures fail to solve the exact model of ideal filters, due to their sub-optimal solutions. Hence, the article considers the global optimization technique to avoid premature convergence. Also, it analyzes a novel meta-heuristic algorithm called functionally linked cat swarm optimization, which is proposed to improve the local and global search abilities of conventional cat swarm optimization technique while searching the filter coefficients of linear phase BSF.*

I. INTRODUCTION

A filter is a frequency selecting/rejecting device extensively used for signal restoration or separation. In this digital age, modern filters are the essential blocks in most digital signal processing systems owing to the ease of implementation of higher-order realistic filters at low prices with excellent performance compared to analog types. The easy way of realizing digital filtering is through the convolution operation on the noisy input by the impulse response of the digital filter (DF). Hence DF is the basic building block in most modern communication systems. Also, it has tremendous applications in speech processing, radar signal processing, biomedical signal processing, video signal processing, etc. Typically, two fundamental types of DFs are available, named finite impulse response (FIR) and infinite impulse response types [1,2]. FIR type is generally favored due to its simplicity of design and stability. Again, when the filter weights are symmetrical around the middle position, then the FIR type is confirmed for a precise linear-phase to avoid signal phase distortion totally [3,4]. Hence, this article gives importance to the FIR design aspect of DF. Generally, the filter design approach is measured as a complex, constraint, and highly nonlinear (hence multimodal) optimization problem; hence the objectives function is developed as the sum of error functions requiring minimization to get the best response. Many gradients based traditional minimization tools are available in the literature for exact modeling of linear-phase DFs and these cannot congregate to the global solutions due to erratic behaviors [1-4].

Therefore, modern evolutionary search methods have been suggested for this design due to ease of searching the parameters along with achieving other benefits like the maximum stopband attenuation (SA), etc. Many heuristic optimization methods are available in the literature. Particle swarm optimization (PSO) [5], several variants of PSO such as comprehensive learning PSO (CLPSO) [6], and craziness based PSO (CRPSO) [7] have been used for this filter design approach. But these algorithms are not so effective in finding the universal optima in terms of rate of convergence and solution quality. To relax from these difficulties and to preserve the diversity of the cats, the author has modified the traditional cat swarm optimization (CSO) [8-10] algorithm by developing some suitable nonlinear/trigonometric functions, similar to FLANN/ LFLANN systems [11,12], to improve the performance of the design approach named as the functionally linked CSO (FLCSO) algorithm. Based on the proposed approach, the article presents a high-quality; comprehensive set of results and shows its supremacy.

II. DESIGN FORMULATION

An Nth order band-stop linear-phase filter is modeled as:

$$H(z) = \sum_{n=0}^{N} h(n) z^{-n}$$

Here, the bandstop filter (BSF) has (N+1) number of coefficients in its impulse response, h(n). The elements of h(n) will ascertain the nature of the DF e.g. low/high pass, band pass/stop, etc. This script presents the optimal design of even- order, linear-phase, BSF having the point of symmetry at (N/2+1). Due to the symmetrical nature of the h(n), simply (N/2+1) number of coefficients are required to be searched for the purpose. The optimization algorithm attains the least error among the ideal response and the estimated frequency response by searching the best h(n) within an upper limit of iterations. The optimum h(n) after concatenation represents the final impulse response.

Other supporting DF parameters required for this design are passband ( ) and stopband ( ) cut-off frequencies; passband ( ) and stopband ( ) ripples. These DF parameters are regulated by the searched coefficients. Now, the coefficients of h(n) in Eq. (1) is assumed as a cat vector, { }. The random cat vectors are scattered in an M-dimensional search space within [-1,+1]. Where M = (N/2+1) for an Nth order linear-phase BSF and its sampled frequency response is defined in Eq. (2).

$$H(\omega_k) = \sum_{n=0}^{N} h(n)e^{-j\omega_k n}$$

Where $\omega k = (2\pi k/L)$ and $H(\omega k)$ is the discrete Fourier transform complex vector. The frequency is sampled in $[0, \pi]$ with L (L $\geq$ N) sampling points shown in Eq. (3), where $Hp(\omega)$ is the set of estimated samples of amplitude for the practical DF and $Hi(\omega)$ is the set of ideal samples for the BSF as shown in Eq. (4).

$$H_p(\omega) = \left[H_p(\omega_1), H_p(\omega_2),...,H_p(\omega_L) \right]^T$$

$$H_i(\omega) = \left[H_i(\omega_1), H_i(\omega_2),...,H_i(\omega_L) \right]^T$$

$$H_i(\omega) = \begin{cases} 0 & for \ \omega_{ls} \leq \omega \leq \omega_{us} \\ 1 & otherwise \end{cases}$$

Eq. (5) represents the ideal BSF response with   and   as the upper and the lower stopband edge frequencies respectively. Each position of the cat vector in M-dimensional search space is a solution of h (n). For this design approach, the author has used an efficient error fitness function defined in Eq. (6) including Gibbs phenomena for both pass and stop bands [5].

$$Error = \sum abs \left\{ abs(H_p(\omega)) - abs(H_i(\omega)) \right\}$$

$$J_f = \sum abs[abs(| H_p(\omega)| - 1) - \delta_p] + \sum abs(| H_p(\omega)| - \delta_s)$$

Thus the fitness error function, Jf specified in (7) is the applied fitness function, needs minimization utilizing the bio-inspired techniques PSO, CLPSO, CRPSO, CSO, and FLCSO independently. Each algorithm tries to minimize this error fitness function to achieve the best DF.

III. FLCSO ALGORITHM

This script adopts a novel meta-heuristic evolutionary minimization algorithm named FLCSO, particularly for this DF design. It is modeled on observing a cat's active behavior during chasing which is characterized by two sub-models. In both of the modes, few steps have been modified by the author developing some nonlinear functions to enhance the local and global searching capabilities of conventional CSO [9,10]. A cat has a burly curiosity towards a moving object and possesses excellent hunting skills. Typically, cats spend most of their time sleeping; yet remain alert and travel very slowly. When prey is observed, they run after it speedily spending a huge amount of energy. These 2 features: (i) alert with slow movement and (ii) chasing it rapidly are modeled as seeking and tracing modes respectively, as follows:

A. Seeking mode

This mode is modeled for a local search around each cat. Basic terms used for this mode are: (1) the seeking memory pool (SMP), is the no. of copies of a cat formed for each cat in the seeking mode. (2) The seeking range of the selected dimension (SRD), is the peak difference between the old and new values in the dimension selected for mutation. Here, the SRD fraction is maintained to vary randomly within -1% to +1% (i.e., between -0.002 to 0.002) (non) uniformly. The new position of copies is generated randomly by subtracting or adding an SRD fraction of the current position value by proper selection of trigonometric/algebraic nonlinear functions similar to FLANN/LFLANN algorithms [11,12]. The author has developed the following five new random functions to be subtracted or added to generate new copies of the present position for each seeking cat by creating an automatic change amount within ±0.002 most of the time and found to provide improved local search in the seeking mode comparison to traditional CSO algorithm. Here, random number, is in the range [0, 1]; $0.001 \leq s < 0.004$; q (nearest integer) $< CDC \times M$; in the following expressions (i), (ii), (iii), (iv), and (v).

(i) $s(r_1 - 0.5)$ (ii) $s\{r_2(1,q) - 0.5\}$

(iii) $[\sin\{\min(h(n))\}]^4 - [\sin\{\min(h(n))\}]^5.$

(iv) $s(r_3 - 0.5)[\log_{10}\{abs(sum(h(n)))\} - \{\log_{10}(\max(h(n)))\}].$

(v) $0.5[\{\tan(r_4)\}\{\sin(r_5 - 0.5)\}][\{\tan(r_6)\}\{\sin(r_7 - 0.5)\}][\{\tan(r_8)\}\{\sin(r_9 - 0.5)\}].$

(3) The counts of dimension to change (CDC) are the nos. of dimensions required to be mutated. (4) The term, mixture ratio (MR) is a fraction of the population selected, is used to confirm that each cat spends most of its time sleeping and observing while in the seeking mode. The execution steps for this mode are:

Step 1: Initialize a population, np of M-dimensional, random cat vectors within limits [-1.0, 1.0]. Opt for suitable values of MR, SMP, SRD, and CDC.

Step 2: Compute the fitness value of each cat of the population and select the best cat as the group best (g_{best}) as represented in Eq. (8).

Step 3: Pick up arbitrarily MR fraction of population as seeking cats, others as tracing cats.

Step 4: Create SMP no. of copies for each seeking cat.

Step 5: As per the CDC, remodel the position of each one copy arbitrarily subtracting or adding the SRD fraction of the current position value using any one of the above five described nonlinear functions.

Step 6: Compute the fitness error value of each new copy.

Step 7: Select the best one among all the copies for each seeking cat.

Step 8: Repeat steps 4-7 until all seeking cats are involved.

B. Tracing mode

This mode is modeled for a global search where each cat traces the target with a large step. The author has developed an efficient velocity controlling parameter, to improve the present velocity component. The steps executed in tracing mode are:

Step 1: Define the initial position of each tracing cat in the M-dimensional search space as per Eq. (9).

Step 2: Compute   as per Eq. (10) for Eq. (11) taking three arbitrary cats present in the tracing mode.

Step 3: Update ith tracing cat's velocity as per Eq. (12) for Eq. (13). Also, update each position as per Eq. (14).

Step 4: Evaluate the error fitness values of all updated ith position and compare with its old positions to keep the best one for the ith cat in the tracing mode.

Step 5: Repeat steps 2-4 until all tracing cats are involved.

$$g_{best} = (g_{best1}, g_{best2}, \ldots, g_{bestM})$$

$$P_i = (P_{i1}, P_{i2}, \ldots, P_{iM})$$

$$\delta_{iM} = \frac{(P_{iM} - P_{(i+1)M}) + (P_{(i+2)M} - P_{(i+3)M})}{2}$$

$$\delta_i = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{iM})$$

$$V_{iM} = w * \delta_{iM} + C*r* (g_{bestM} - P_{iM})$$

$$V_i = (V_{i1}, V_{i2}, \ldots, V_{iM})$$

$$P_{iM} = P_{iM} + V_{iM}$$

Here, w is the inertia constant; C is the acceleration factor; r is an arbitrary number [0,1]. The FLCSO algorithm finds its global optimum solution combining the two groups of cats.

C. Application of FLCSO to the BSF Design

The basic steps are:

Step 1: Create   number of cats randomly for the process and fix up the upper limit of iteration cycles. Select suitable values of w and C.

Step 2: Compute the fitness value of each cat (irrespective of the mode) and keep the most excellent one (g_{best}) in the memory to represent the best solution so far.

Step 3: Select arbitrarily some cats into the seeking mode as per MR and the rest as the tracing mode.

Step 4: Update the cat as per its flag, if cat-i is in the seeking mode, involve the cat in the seeking mode process, otherwise involve it in the tracing mode process as described above.

Step 5: Update the position of the   and repeat steps 3-5 until reaching the maximum limit of iterations.

Step 6: Form the full (N+1) number of coefficients of h(n) by copying and concatenating the final   to get the best frequency response of the BSF. The flow chart of the FLCSO algorithm for the purpose is shown in Fig. 1.
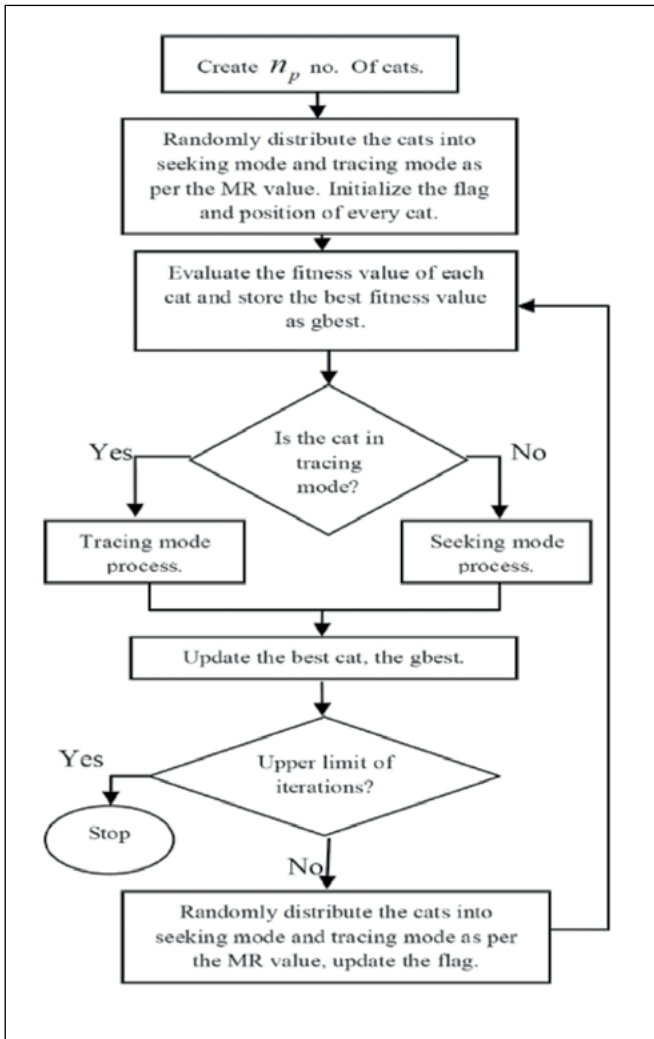

Figure 1: Flow chart of FLCSO for the FIR BSF design approach.

IV. RESULTS AND DISCUSSIONS

The proposed FLCSO algorithm has been applied to solve the BSF design problem for verifying its workability and applicability. Extensive MATLAB simulation was performed to realize the linear-phase BSFs of order

36; hence, the no. of filter coefficients is 37. 256 no. of samples were considered for each simulation and each method was run 30 times to search out the best solution. The design specifications are selected as follows:

- Passband ripple ( ) = 0.02
- Stopband ripple ( ) = 0.004
- Lower passband (normalized) edge frequency,= 0.3
- Lower stopband edge frequency, = 0.35
- Upper stopband edge frequency, = 0.7
- Upper passband edge frequency, = 0.75

Judiciously, the following parameters are selected for getting better results in this BSF design approach considering all the applied methods.

- Population size = 120; No. of iterations = 400;
- SMP,SRD,CDC = 5, 0.2, 0.6;
- w, C = 0.4, 1.8; MR = 0.7.

The best coefficients of the designed linear-phase BSF of order 36 have been searched by the proposed and other methods and are shown in Table I. Table II and Table III summarize the relative findings of important performance parameters using the proposed and other algorithms. The maximum (Max.), minimum (Min.), mean, and standard deviation of SA have been computed and are illustrated in Table II. The proposed FLCSO based approach provides the highest SA of 58.21dB, Min. passband ripple (absolute (Abs.)) of 0.018, and min. stopband ripple (abs.) of 0.005. Also, the transition width and the bandwidth (BW) of the stopband of respective filters using different algorithms are brought for comparison in Table II.

Table I. of Optimized Coefficients BSF Searched

H(36-N) = H(N)	Pso	Crpso	Clpso	Cso	Flcso
h(0)	-0.0012	-0.0051	-0.0047	0.0013	0.0003
h(1)	-0.0262	-0.0250	-0.0311	-0.0168	-0.0113
h(2)	0.0057	0.0015	0.0047	0.0028	0.0035
h(3)	-0.0025	0.0017	-0.0080	0.0023	0.0041
h(4)	0.0103	0.0074	0.0061	0.0042	0.0039
h(5)	0.0333	0.0312	0.0278	0.0256	0.0221
h(6)	-0.0202	-0.0194	-0.0256	-0.0212	-0.0200
h(7)	-0.0246	-0.0272	-0.0268	-0.0203	-0.0213
h(8)	-0.0004	-0.0034	-0.0021	0.0005	0.0012
h(9)	-0.0267	-0.0232	-0.0259	-0.0217	-0.0213

h(10)	0.0568	0.0485	0.0533	0.0515	0.0510
h(11)	0.0413	0.0466	0.0341	0.0418	0.0411
h(12)	-0.0527	-0.0479	-0.0549	-0.0515	-0.0503
h(13)	-0.0047	-0.0001	-0.0013	0.0000	- 0.0005
h(14)	-0.0861	-0.0803	-0.0897	-0.0851	-0.0854
h(15)	-0.0526	-0.0502	-0.0515	-0.0464	-0.0467
h(16)	0.2993	0.2978	0.2934	0.2974	0.2955
h(17)	0.0328	0.0328	0.0271	0.0291	0.0299
h(18)	0.5991	0.5992	0.5938	0.5999	0.6002

Table II. Performance Parameters of the BSFs

Technique	SA (dB)			Standard deviation	Normalized Transition	
	Max.	Min.	Mean		width	BW
PSO	71.2	26.0	38.91	8.22	0.065	0.42
CRPSO	59.8	29.9	42.14	4.64	0.085	0.41
CLPSO	74.9	27.0	39.53	10.98	0.072	0.42
CSO	66.2	31.5	49.25	2.82	0.085	0.42
FLCSO	103.2	44.4	58.21	3.14	0.085	0.42

The comparative magnitude plots are illustrated both in Figs. 2 and 3. The simulation findings confirm that the proposed FLCSO based approach of BSF design results in Min. passband ripple and Min. stopband ripple which is observed in Figs. 4 and 5 respectively and are listed in Table III. The proposed FLCSO based design shows 8.96 dB, 18.68 dB, 16.07 dB, and 19.3 dB improvements for SA compared to CSO, CLPSO, CRPSO, and PSO BSFs respectively. Also, the proposed design shows 0.008 (Abs.), 0.031, 0.024, and 0.038 improvements for passband ripple compared to CSO, CLPSO, CRPSO, and PSO BSFs respectively. Similarly, the proposed FLCSO based design shows 0.003 (Abs.), 0.017, 0.013, and 0.027 improvements for stopband ripple compared to CSO, CLPSO, CRPSO, and PSO linear-phase BSFs respectively. Also, it is marked that with almost the same level of transition band, the proposed BSF design provides the Max. SA(dB) and the least pass/stop band ripples (Abs.) without changing its BW compared to other reported approaches.

Table III. Stop/Pass Band Ripple of the BSFs

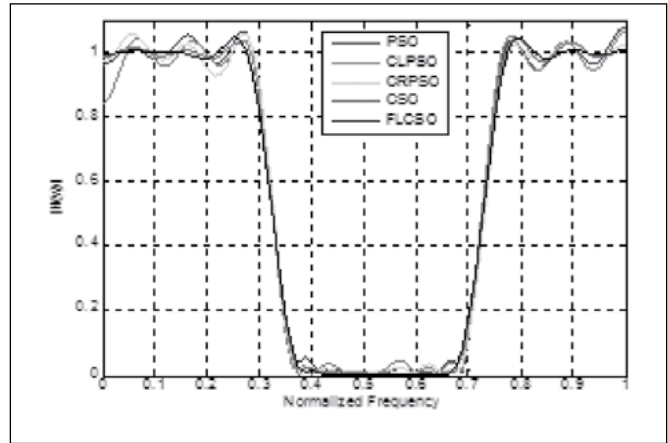| Technique	Stopband Ripple (Abs.)			Passband Ripple (Abs.)		
	Min.	Max.	Mean	Min.	Max.	Mean
PSO	0.001	0.056	0.032	0.022	0.084	0.065
CRPSO	0.003	0.034	0.018	0.02	0.06	0.042
CLPSO	0.001	0.043	0.022	0.03	0.066	0.049
CSO	0.001	0.024	0.008	0.012	0.05	0.026
FLCSO	0.000	0.014	0.005	0.001	0.04	0.018

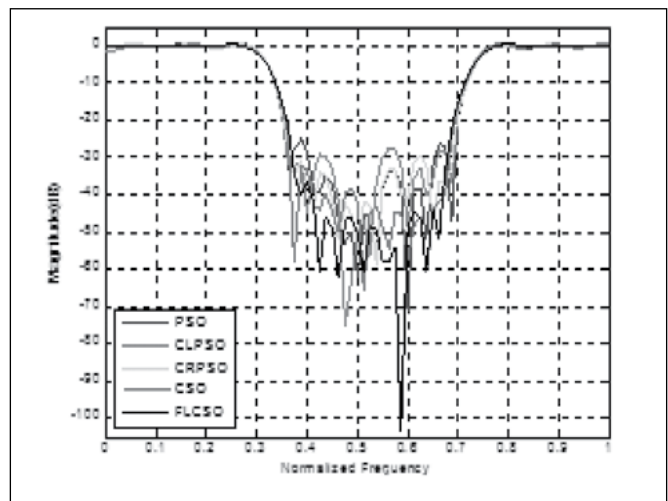*Figure. 2: The magnitude (Abs.) response of the BSF.*

Figure. 3: Magnitude (dB) response of the BSF

The fitness error curve of the respective filter is compared in Fig. 6. The FLCSO converges to a much lower error fitness value compared to PSO, CLPSO, CRPSO, and CSO techniques. From Fig. 6, it is found that the conventional PSO converges to the least possible fitness error value of 6.74 in 309 iteration cycles; the CLPSO converges to the least fitness error value of 6.07 in 322 iterations; the CRPSO converges to the lowest error fitness value of 5.55 in 292 iterations; the CSO converges to the lowest fitness error value of 4.27 in 255 iterations; whereas, the proposed FLCSO converges to the smallest error fitness value of 3.36 in 228 iterations while searching the optimal DF coefficients. With scrutiny to the above reality, it is ultimately concluded that the performance of the FLCSO technique is the best amid all these techniques in solving this design approach.

Step 4: Update the cat as per its flag, if cat-i is in the seeking mode, involve the cat in the seeking mode process, otherwise involve it in the tracing mode process as described above.

Step 5: Update the position of the and repeat steps 3-5 until reaching the maximum limit of iterations.

Step 6: Form the full (N+1) number of coefficients of h(n) by copying and concatenating the final to get the best frequency response of the BSF. The flow chart of the FLCSO algorithm for the purpose is shown in Fig. 1.

V. CONCLUSIONS

In this work, an improved version of the CSO algorithm named the FLCSO algorithm has been analyzed while designing the linear-phase BSF. It is evident from the simulation analysis and the convergence graphs, that FLCSO can provide much accurate optimal impulse response characteristics in terms of Max. SA and least stop/pass band ripples compared to other global search methods reported in this work. Also, simulation results demonstrate that the proposed FLCSO can provide a better quality solution compared to the conventional CSO while designing the linear-phase BSF. Thus, it is concluded that the novel FLCSO can be considered as a global optimizer for this design approach of modern filters in the practical signal processing framework.
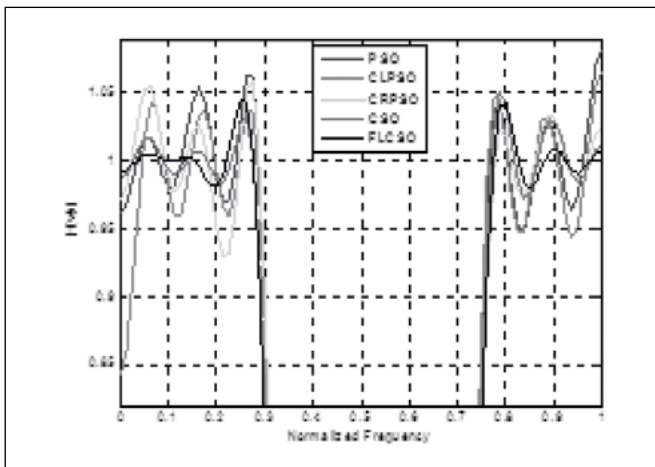

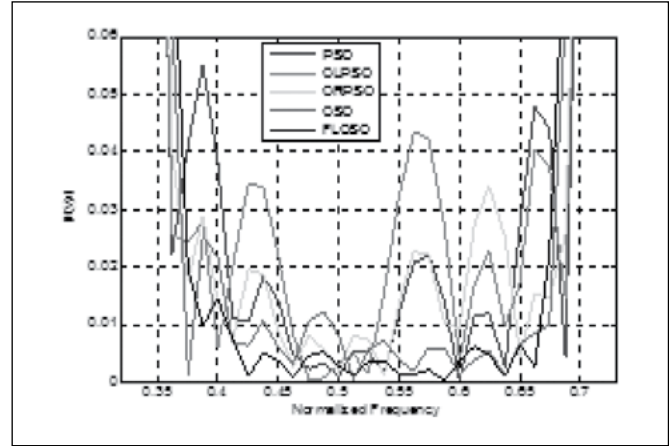*Figure. 4: Passband ripple plots of the BSFs of order 36.*

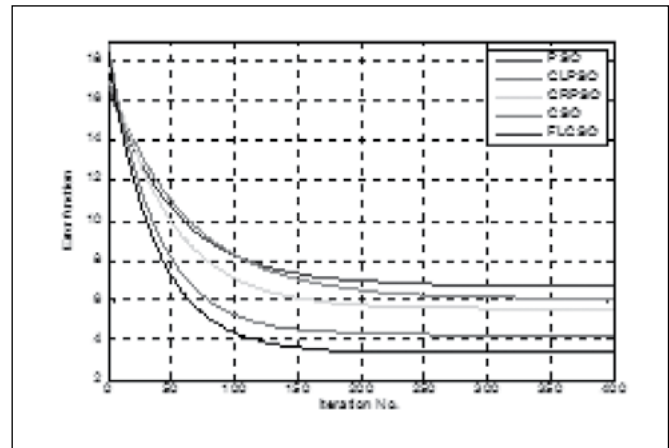*Figure. 5: Stopband ripple plots of the BSFs of order 36.*

Figure. 6: Fitness curve of BSF for order 36.

REFERENCES

1. J. I. Ababneh, and M. H. Bataineh, "Linear phase FIR filter design using particle swarm optimization and genetic algorithms," Digital Signal Processing, vol. 18(4), pp. 657–668, Jul. 2008.

2. B. Li, Y. Wang, T. Weise, and L. Long, "Fixed-point digital IIR filter design using two-stage ensemble evolutionary algorithm," Applied Soft Computing, vol. 13, No.1, pp. 329-338, Jan. 2013.

3. D. Schlichtharle, Digital Filters, Basics and Design, 2nd ed., Springer: Berlin Heidelberg, 2011, pp.329-371.

4. Z. Ye and C.-H. Chang, "Local search method for FIR filter coefficients synthesis," 2nd IEEE Int. Workshop on Electronic Design, Test and Applications, Proc. DELTA, pp. 255–260, Jan. 2004.

5. M. Kumar, and T. N. Sasamal, "Design of FIR filter using PSO with CFA and inertia weight approach," in: IEEE International Conference on Computing, Communication, and Automation, ICCCA, pp. 1331-1334, May 2015.

6. H. Ali, and F. A. Khan "Attributed multi-objective comprehensive learning particle swarm optimization for optimal security of networks," Applied Soft Computing, vol. 13, No. 9, pp. 3903–3921, Sept. 2013.

7. N. A. Firdausanti, and Irhamah "On the Comparison of Crazy Particle Swarm Optimization and Advanced Binary Ant Colony Optimization for Feature Selection on High-Dimensional Data," The Fifth Information Systems International Conference, Procedia Computer Science, vol. 161, pp. 638–64, 2019.

8. B. Santosa, and M. K. Nigrum, "Cat Swarm Optimization for clustering," in: International Conference of Soft Computing and pattern recognition, SOCPAR'09, pp. 54-59, Dec. 2009.

9. J. So, and W. K. Jenkins, "Comparison of Cat Swarm Optimization with Particle Swarm Optimization for IIR System Identification," IEEE Asilomar Conference on Signals, Systems and Computers, pp. 903-910, 2013.

10. G. Panda, P. M. Pradhan, and B. Majhi, "IIR System Identification using Cat Swarm Optimization," Expert System with Application, vol. 38, No.10, pp. 12671–83, Sept. 2011.

11. S. Dehuri, R. Roy, Sung-Bae Cho, and A. Ghosh, "An improved swarm optimized functional link artificial neural network (ISO-FLANN) for classification," The Journal of Systems and Software, vol. 85, pp. 1333– 1345, June 2012.

12. G. Das, P. K. Pattnaik, and S. K. Padhy, "Artificial Neural Network trained by Particle Swarm Optimization for non-linear channel equalization," Expert Systems with Applications, vol. 41, pp. 3491–3496, June 2014.

Dr Judhisthir Dash

ECE Dept.

Quotes by famous female scientists

1. Don't let anyone rob you of your imagination, your creativity, or your curiosity. It's your place in the world; it's your life. Go on and do all you can with it, and make it the life you want to live."
- Mae Jemison, first African American woman astronaut in space.

2. "Certain people – men, of course – discouraged me, saying [science] was not a good career for women. That pushed me even more to persevere."
- Francoise Barre, virologist who won the 2008 Nobel Prize in Physiology and Medicine.

3. "If you're beautiful, you're led to believe that you can't also be smart. But you can be fun and fit and social and be really smart. And the smarter you are, the more capable you'll be to handle whatever challenges come up in life." **- Danica McKellar, Mathematician.**
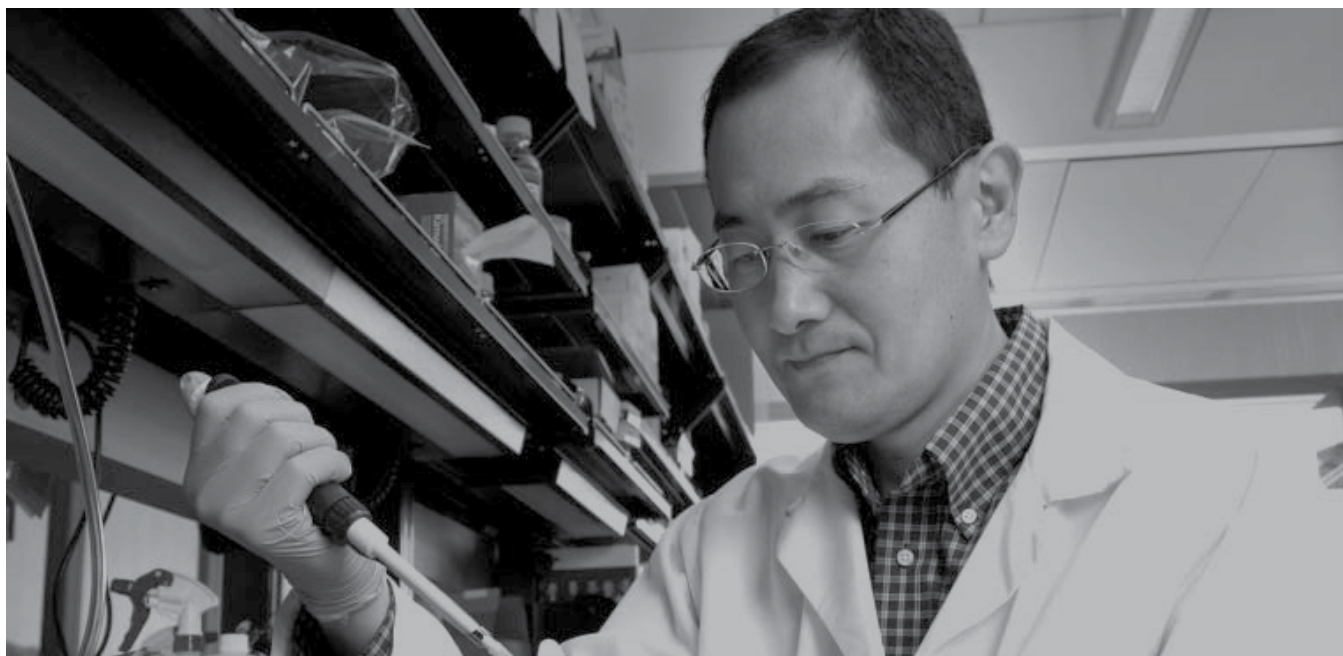
SHINYA YAMANAKA

Born in a town of Japan known as Osaka and famous as "Rugby football town", this great Scientist and Doctor is famous for developing a revolutionary method for generating stem cells from existing cells of the body, which fetched him Nobel Prize for Physiology or Medicine in 2012 for his discovery.

Yamanaka received an M.D. from Kobe University in 1987 and a Ph.D. in pharmacology from the Osaka City University Graduate School in 1993. From 1993 to 1996, he was at the Gladstone Institute of Cardiovascular Disease, San Francisco, where he began investigating the c-Myc gene in different strains of knockout mice (mice in which a specific gene has been rendered nonfunctional in order to investigate the gene's function). In 1996 Yamanaka returned to Osaka City University, where he remained until 1999, but found himself mostly looking after mice in the laboratory, not doing actual research, then decided to focus on the research and took a position at Nara Institute of Science and Technology. In 2006, he and his team generated induced pluripotent stem cells (iPS cells) from adult mouse fibroblasts (connective-tissue cells). The following year Yamanaka reported that he had derived iPS cells from human adult fibroblasts—the first successful attempt at generating human versions of these cells. This discovery marked a turning point in stem-cell research, because it offered a way of obtaining human stem cells without the controversial use of human embryos.

When asked about the inspiration for the research the scientist answered-"When I saw the embryo, I suddenly realized there was such a small difference between it and my daughters. I thought, we can't keep destroying embryos for our research. There must be another way." In order to achieve this goal he dedicated himself for many years, faced many hurdles but never gave up. He possesses a can-do attitude.

Now, let's know more about the discovery, this method involved inserting specific genes into the nuclei of adult cells (e.g., connective-tissue cells), a process that resulted in the reversion of cells from an adult state to a pluripotent state. As pluripotent cells, they had regained the capacity to differentiate into any cell type of the body. Thus, the reverted cells became known as induced pluripotent stem (iPS) cells.

Yamanaka received multiple awards for his contributions to stem-cell research, including the Robert Koch Prize (2008), the Shaw Prize in Life Science and Medicine (2008), the Gairdner Foundation International Award (2009), the Albert Lasker Basic Medical Research Award (2009), and the Millennium Technology Prize (2012).

STRESS METER USING ARDUINO

Abstract : - In India, many people die not because of lack of facilities but because of lack of knowledge among the people. Using Arduino, we designed a device to diagnose stress. It is very cheap, economical and run on a 9 volt battery. Stress meter is a device which used to measure stress of the body depending upon the current flow through our body. Our body resistance level varies with stress of our body and mind. A highly stressed body results in increase of blood flow through our body mechanism which result in increase in ion concentration in our body and thus allow more current to flow across our body. While a stress-free human being has high resistance due to less blood flow and thus result in less ion concentration all our body.

Keywords: Light emitting diode (LED), Integrated Development Environment (IDE), heart rate (HR), various indices of heart rate variability (HRV), blood pressure (BP), Wavelet-Independant Component Analysis (WICA).

I. INTRODUCTION

Stress is a term that refers to the sum of the physical, mental, and emotional strains or tensions on a person. Feelings of stress in humans result from interactions between persons and their environment that are perceived as straining or exceeding their adaptive capacities and threatening their well-being. The element of perception indicates that human stress responses reflect differences in personality as well as differences in physical strength or health. This stress meter allows assessing one's emotional pain. If the stress is very high, it gives visual indication through LED display along with a warning red light and buzzer sound.

II. PRINCIPLE OF STRESS METER

The gadget is based on the principle that the resistance of the skin varies in accordance with your emotional states. If the stress level is high the skin offers less resistance, and if the body is relaxed the skin resistance is high. The low resistance of the skin during high stress is due to an increase in the blood supply to the skin which increases the permeability of the skin. High stress causes sweating and leakage of water from the blood vessels in the skin and this makes the skin moist and electrical conductivity increases.

III. RESEARCH BACKGROUND

Stress is a term that refers to the sum of the physical, mental, and emotional strains or tensions on a person. Feelings of stress in humans result from interactions between persons and their environment that are perceived as straining or exceeding their adaptivecapacitiesand

threateningtheirwell-being.Theelementofperception indicates that human stress responses reflect differences in personality as well as differences in physical strength or health. This Stress meter allows to assess one's emotional pain. If the stress is very high, it gives visual indication through LED display along with a warning yellow light. Stress meter is based on the principle that the resistance of the skin varies in accordance with your emotional states. Resistance varies inversely proportional to the stress. If the stresslevel ishighthe skin offers less resistance, and if relaxed, resistance is high.In an article "Stress and Mind Control", 2008, Roberto Bonomi stated that "When we speak of the fabulous relaxation capacity that mind control gives us, the first thing that comes to our mind, is that we will be able to take off, the excesses of nervous tension, the stress; and this is a great benefit. Because suppose that you could measure stress in inches, and that you have stress zero when the meter is located in zero." Based on this our project is aimed to give a visual indication of one's stress through a light-emitting diode display along with a warning light [1].

A number of physiological markers of stress have been identified, including electro dermal activity, heart rate (HR), various indices of heart rate variability (HRV), blood pressure (BP), muscle tension, and respiration [2]. Other than that, some biological markers that are known to change with changes in the stress level of the person, such as the blood composition, respiration and hormonal changes.

Nowadays, there are many researches related to the stress have been conducted. Satoshi Suzuki and colleagues

proposed measurement of heart rate variability using compact microwave radar for stress monitoring. A novel method for non-contact monitoring of stress-induced autonomic activation through the back of a chair, using compact 24 GHz microwave radar had developed. As a reference, pectoral ECG was monitored along the noncontact monitoring, the output signals from both of the contact and noncontact system were sampled through A/D converter with the same sampling rate of 100 Hz. A non-contact autonomic activation measurement system for noncontact measurement of HRV is designed. The system consists of prototype compact microwave radar, control unit for microwave radar, an A/D converter and a personal computer [3].

Syahrull Hi-Fi Syam B Ahmad Jamil and colleagues is conducted a research in determination of stress index. Four differents types of simulation will undergo constantly in order to make sure the subjects are really under stress during the measurement processes which are hearing test, cognitive test, emotion test and environment test. The collected data from ECG are analysis using Power Spectral technique, Wavelet Transform (WT) is technique for EEG diagnosis while Wavelet-Independant Component Analysis (WICA) is technique for EMG diagnosis. Then Artificial Intelligence is use to determine the stress index which there are two types of approaches are use in order to interpret and elaborate stress index. The approaches are qualitative approach and quantitative approach.

A research about detecting mental stress using unobtrusive wearable sensors of heart rate is performed by Jongyoon Choi and Ricardo utierrez-Osuna. The approach relies on estimating the state of the autonomic nervous system from an analysis of heart rate variability. Namely, we use a nonlinear system identification technique known as principal dynamic modes (PDM) to predict the activation level of the two autonomic branches: sympathetic (i.e. stress-inducing) and parasympathetic (i.e. relaxation related) [4]. We validate the method on a discrimination problem with two psycho-physiological conditions, one associated with mental tasks and one induced by relaxation exercises.

IV. ARDUINO

Arduino shown in Fig. 1 is an open-source electronics platform based on easy-to-use hardware and software. Arduino consists of both a physical programmable circuit board (often referred to as a microcontroller) and a piece of software, or IDE (Integrated Development Environment) that runs on your computer, used to write

and upload computer code to the physical board. Arduino boards are able to read inputs e.g. light on a sensor, a finger on a button, or a twitter message-and turn into an output activating a motor, turning on an LED, publishing something online.
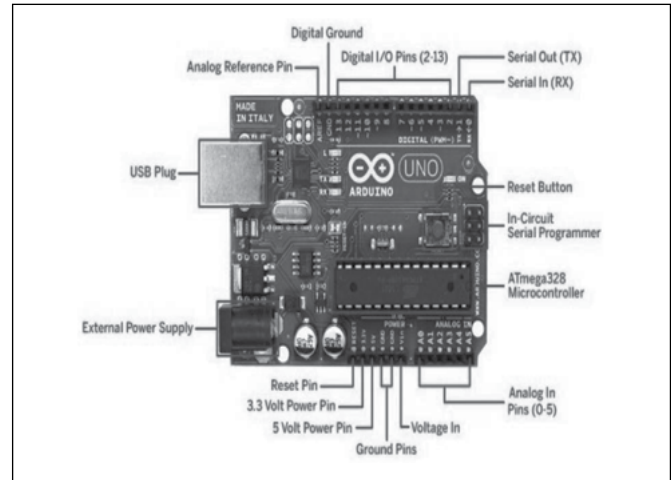

Figure. 1. *Pin configuration of Arduino Uno*

V. PIEZO-BUZZER

A piezo-buzzer is a sound producing device. The main working principle is based on the theory that, whenever an electric potential is applied across a piezoelectric material, a pressure variation is generated. A piezo-buzzer consists of piezo crystals in between two conductors as shown in Fig. 2 and 3. When a potential difference is applied across these crystals, they push one conductor and pull the other conductor by their internal property. The continuous pull and push action generates a sharp sound wave. Piezo-buzzers generate a loud & sharp sound. So, they are typically used as an alarm circuits. Also they are used to make an alert of an event, signal or sensor input. A special characteristic of piezo-buzzer is, the sound pitch or level is not depended on the voltage level that is, it works only in a specific voltage range. Typically, a piezo-buzzer can generate a sound in the range of 2 to 4 kHz.
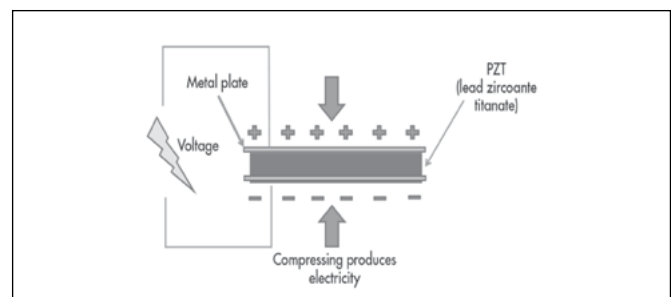

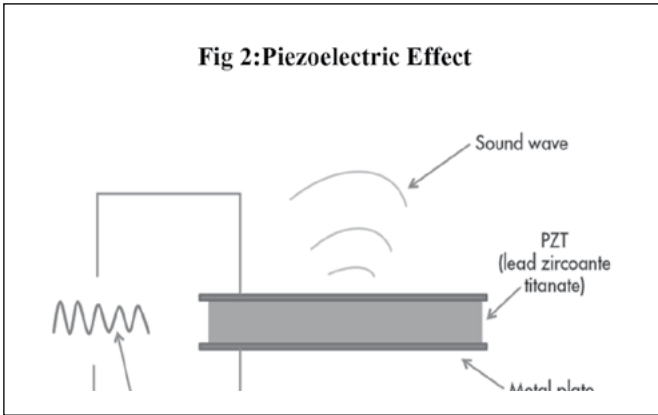**Figure. 2.** *Piezoelectric Effect*

Fig 2:Piezoelectric Effect

Figure. 3. *Inverse Piezoelectric Effect*

The piezoelectric effect is very useful within many applications that involve the production and detection of sound, generation of high voltages, electronic frequency generation, microbalances, and ultra fine focusing of optical assemblies. It is also the basis of a number of scientific instrumental techniques with atomic resolution, such as scanning probe microscopes (STM, AFM, etc).

TOUCHPAD

A touchpad shown in Fig. 4 is a device for pointing (controlling input positioning) on a computer display screen. While touchpad's, like touch screens, by their design are able to sense absolute position, resolution is limited by their size.
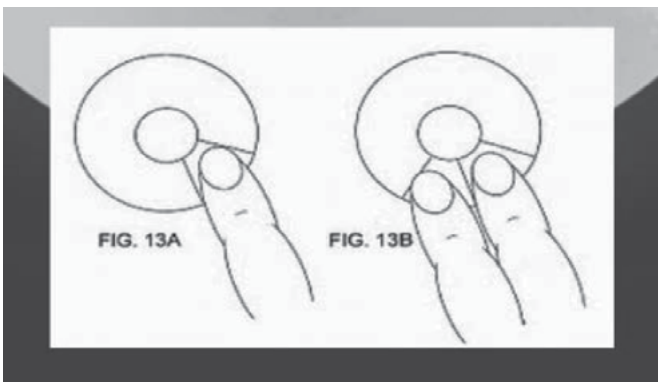

Figure. 4. *Touchpad*

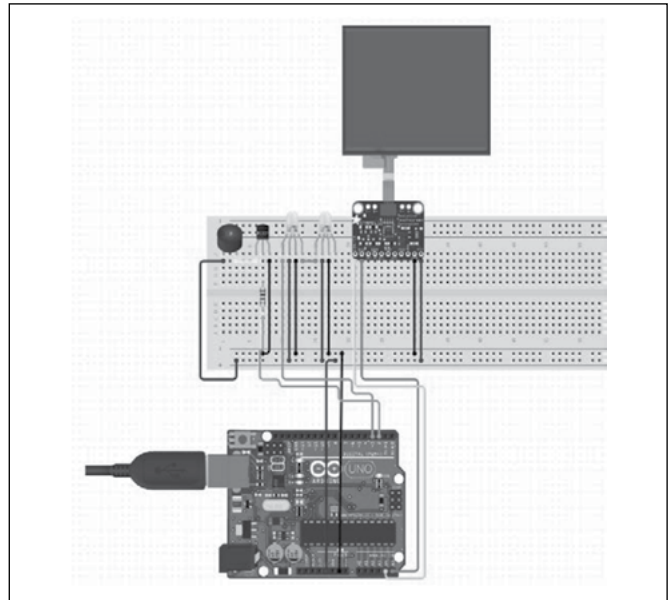Figure 5 shows the complete circuit diagram of the stress meter. The operational flow chart is shown in Fig. 6.
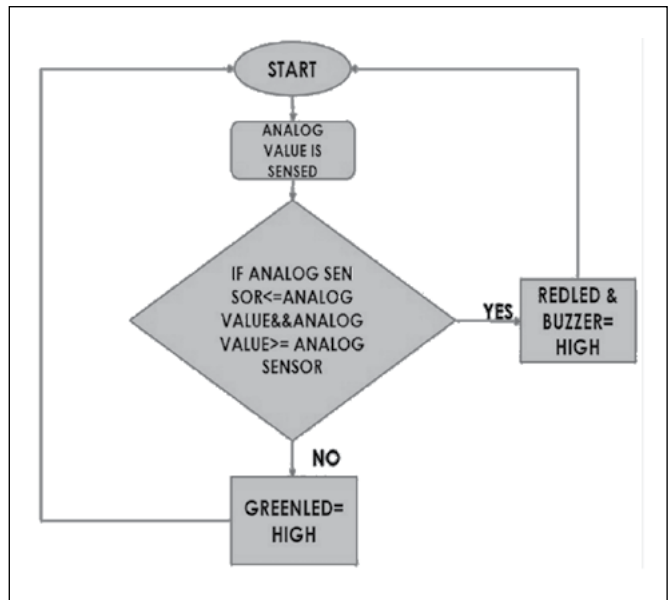

Figure. 5. *Circuit diagram of the stress meter*

Figure. 6. *Operational Flow Chart*
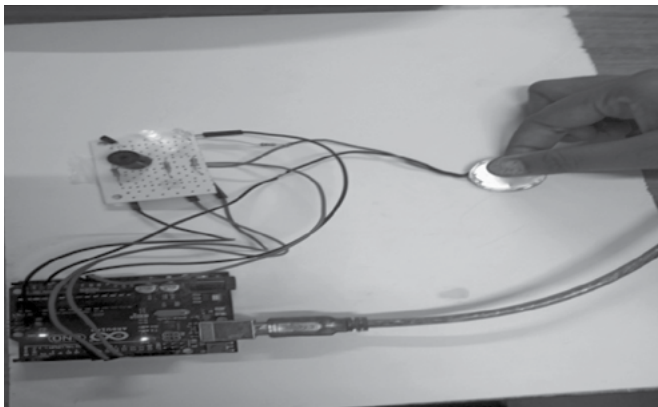
VI. RESULTS

The results are shown in Fig. 7 and 8.

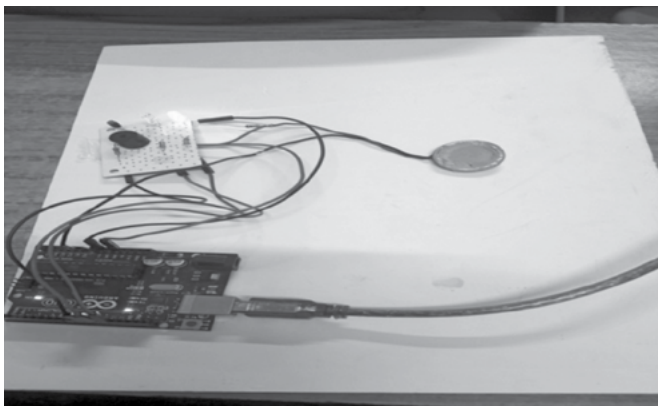*Figure. 7: Low stress condition (the green led glows)*

Figure. 8: High stress condition(the red led glows and the peizo-buzzer starts buzzing)

VII. CONCLUSIONS

As for conclusion, to measure human stress level based on skin resistance, the construction of stress measurement circuit should be done carefully in order to get the precise value. Better understanding on the sensor is necessary due to the important of skin resistance measurement. The skin resistance value is always varying time by time. It needs a protocol to measure the stress level. The stress test using stroop color and arithmetic test are simple and easy to implement to various groups of subject. However, the is no ground truth has been prove by earlier researchers because the level to inducing the stress are also depending on age, gender, culture, environmental, experience, IQ level and many other factors.

REFERENCE:

1. D. Mohan Kumar, "Stress Meter", Electronics For You, September 2005.

2. Jongyoon Choi and Ricardo Gutierrez-Osuna, "Using Heart Rate Monitors to Detect Mental Stress", Body Sensor Networks, 2009.

3. D. Mohan Kumar, "Skin Response Meter", Electronics for You Magazine, Vol 41, no 6, pp. 88, June 2009.

4. Bandana Mallick, Ajit Kumar Patro," Heart Rate Monitoring System Using Finger Tip Through Arduino And Processing Software", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 5, Issue 1, January 2016.

Rakesh Kumar Rout, GouravPattnaik, Arpit Beuria, MitrabindaParida,Manisha Das, Aparna Priyadarsini, Department of Electrical and Electronics Engineering

LOAD FREQUENCY CONTROL OF A MULTI SOURCE POWER SYSTEM WITH HVDC LINK

Abstract : *- In this work, a hybrid system based on Proportional-Integral-Derivative controller is proposed for load frequency control of a two-area interconnected multi-source power system with and without HVDC link. Frequency control is one of the most important services in power system and is achieved through governor response and load frequency control. Each area of the multi-source power system consists of a reheat thermal, a hydro and a gas unit and three PID controllers are used in this study, one for each generating unit. The conventional PID controllers are used to study the dynamic performance of the power system by applying a step load change of 0.01 p.u. in area1.*

Keywords: High voltage DC (HVDC), Load Frequency control (LFC), automatic voltage regulator (AVR), Extra high voltage (EHV), Proportional-Integral-Derivative controller (PID)

I. INTRODUCTION

Electric energy is the most popular form of energy, because it can be transported easily at high efficiency and reasonable cost. The first demonstration of electric light in Calcutta (now Kolkata) was conducted on 24th July 1879 by P.W. Fleury & Co. The introduction of electricity in Calcutta was a success, and power was next introduced in Bombay (now Mumbai) in 1882[1]. The Bombay Electric Supply & Tramways Company set up a generating station in 1905 to provide electricity for tram way[2]. The first hydroelectric installation in India was installed near a tea estate for Darjeeling Municipality in 1897[3]. The first electric street light in Asia was lit on 5th August 1905 in Bangalore [4]. The first electric train in country ran between Victoria Terminus and Kurla on 3rd February 1925[5].

As the need for parallel operation and interconnection became evident, a standard frequency of 50 Hz was adopted throughout the country. The transmission voltage has risen steadily, and extra high voltage (EHV), in commercial use is 765kV, was put into operation. For transmitting power over long distances it was more economical to convert EHV ac to EHV dc.

India began using grid management on a regional basis in 1960s. Individual state grids were interconnected to form 5 regional grids covering mainland India, the Northern, Eastern, Western, Southern and North-Eastern Grids. These regional links were established to enable transmission of surplus electricity between states in each region. Regional Grids were initially interconnected by asynchronous high-voltage direct current (HVDC) back-to-back links facilitating limited exchange of regulated power. Later on these links were upgraded to high capacity synchronous links [6].

In 1990, the Indian government began planning for a national grid. The first interconnection of regional grids was established in October 1991 between North-Eastern and Eastern grid. The Western and Northern grids were interconnected with North-Eastern and Eastern grid in March 2003 and August 2006, forming a Central Grid that was synchronously connected and operating at one frequency[6]. The sole remaining regional grid, the Southern grid, was synchronously interconnected to the Central Grid in December 2013 with commissioning of the 765kV Raichur-Solapur transmission line, establishing the National Grid [6, 7].

In order to keep the interconnected power system in the steady-state, the active and reactive power generation and demand must be equal. The objective of the control strategy is to maintain the power balance in a manner which is both reliable and economic while maintaining voltage and frequency within permissible limits. Changes in real power mainly affects the system frequency, while reactive power is less sensitive to changes in frequency and is mainly dependent on changes in voltage magnitude. Thus, the real and reactive powers are controlled separately. The load frequency control (LFC) loop controls the real

power and frequency whereas the automatic voltage regulator (AVR) loop regulates the reactive power and voltage magnitude. LFC has gained in importance with the growth of interconnected systems and has made the operation of interconnected systems possible. Today, it is still the basis of many advanced concepts for the control of large systems [8].

II. A GLANCE TO THE CONCEPTS

1. LFC

Load-frequency control (LFC) is a type of integral control that restores the system frequency and power flows to adjacent areas back to their values before a change in load. The power transfer between different areas of a system is known as "net tie-line power" [9]. The active and reactive power demands are never steady and they continuously changes with the rising or falling trend of load demand. There is a change in frequency with the change in load which causes problems such as

1. Most AC motors run at speeds that are directly related to frequency. The speed and   induced electro motive force (e.m.f) may vary because of the change of frequency of the power circuit.

2. When operating at frequencies below 49.5 Hz; some types of steam turbines, certain rotor states undergo excessive vibration.

3. The change in frequency can cause mal operation of power converters by producing harmonics.

4. For power stations running in parallel it is necessary that frequency of the network must

2. THERMAL POWER PLANT

Fig. 1 shows the power plant automatic load frequency control block diagram and the followings are the equations:

- Speed Governor

$$\Delta P_g(s) = \Delta P_{ref_0} - \frac{1}{R}\Delta f_0$$

- Hydraulic Valve Actuator

$$\Delta P_V(s) = \frac{1}{1+sT_H}\Delta P_G(s)$$

- Turbine-Generator

$$\Delta P_T(s) = \frac{1}{1+sT_T}\Delta P_V(s)$$

- Closing ALFC loop

$$\Delta P_T - \Delta P_D = \frac{2H}{f_0}\frac{d}{dt}(\Delta f) + D\,\Delta f$$

$$\Delta f(s) = \frac{1}{D+s\frac{2H}{f_0}}[\Delta P_T(s) - \Delta P_D(s)]$$

- Signal fed to integrator is ACE = Δf

- Thus,

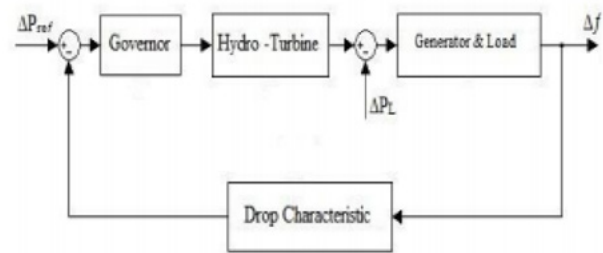$$\Delta P_{ref}(s) = -\frac{K_i}{s}\Delta f(s)$$

Figure. 1: *Low stress condition (the green led glows)*

3. HYDROELECTRIC POWER PLANT

Fig. 2 shows the hydro-electric power plant block diagram and the followings are the equations:

Governor modelling: $\frac{1}{1+ST_g}$

Hydro-turbine modelling: $\frac{1-ST_W}{1+0.5ST_W}$

Generator and load modelling: $\frac{1}{1+ST_p}$
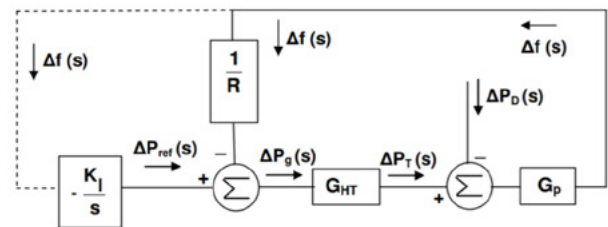
Droop characteristics modelling: $\frac{1}{R_p}$

Figure. 2: *Hydro turbine, its Governing and Generator-Load system model*

4. GAS TURBINE POWER PLANT

Fig. 3 shows the Gas Turbine unit with ALFC the followings are the equation:

- Valve positioner: $\frac{1}{c_g + sb_g}$

- Speed governor: $\frac{1+sX_G}{1+sY_G}$

- Fuel system and Combuster: $\frac{1+sT_{CR}}{1+sT_F}$
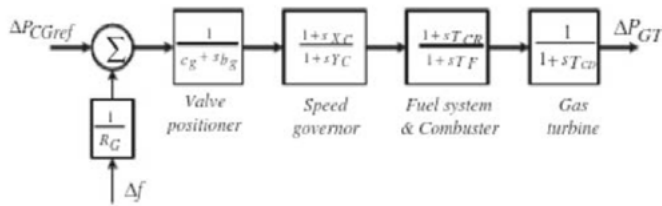
- Gas Turbine: $\frac{1}{1+sT_{CD}}$

Figure. 3: *Block Diagram of a Gas Turbine unit with ALFC*

5. TIE-LINE

The transmission lines that connect an area to its neighboring area are called Tie line shown in Fig. 4. Power sharing between two areas occurs through these tie-lines. Power flow between different areas is regulated by holding frequency constant. Which unlike transmission lines facilitates bidirectional power flow [10].
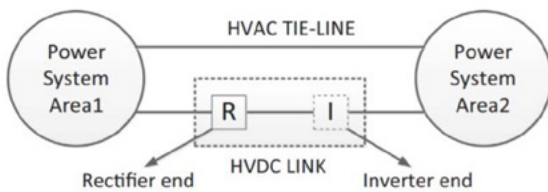

Figure. 3: *Two area system with HVAC and HVDC tie-lines in parallel*

a. AC TIE-LINE

In normal operating conditions power flowing through tie-line is given by,

$$P_{12} = \frac{|V_1^0| - |V_2^0|}{X} \sin(\delta_1^0 - \delta_2^0)$$

Where δ_1^0 and δ_2^0 are angles of end voltages V_1 and V_2 respectively as shown in Fig. 5. The order of subscripts indicate that the power is defined in direction 1 to 2 [11]

But when power demand changes, the change in tie-line power ΔP_{12} is given as,

$$\Delta P_{12} = \frac{|V_1^0| - |V_2^0|}{X} \cos(\delta_1^0 - \delta_2^0)(\Delta\delta_1 - \Delta\delta_2),$$

Now we can define synchronizing coefficient T^0 is given as,

$$T^0 = \frac{|V_1^0| - |V_2^0|}{X} \cos(\delta_1^0 - \delta_2^0)$$

In terms of frequency, $\Delta\delta = 2\pi \int \Delta f \, dt$

Thus,

$$\Delta P_{12} = \frac{2\pi T^0}{s}[\Delta f_1 - \Delta f_2]$$
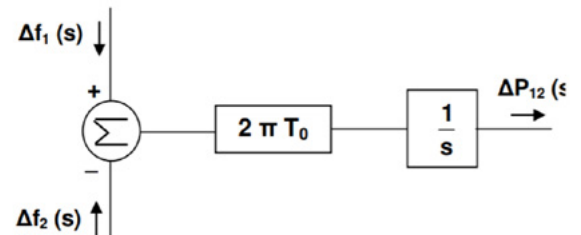

Figure. 5: *AC tie-line*

b. DC TIE-LINE

The DC link is modeled in constant current mode and the incremental power flow through DC link is modeled with incremental change in frequency at rectifier end. The transfer function is for the HVDC link is given by,

$$\frac{SK_{DC}}{1 + ST_{DC}}$$

Where, K_DC and T_DC represents gain of HVDC link and time constant of HVDC link respectively.[12]

6. CONTROLLER

A controller is the most important component of the control system. It is responsible for the performance of the control system. It is a device or an algorithm that works to maintain the value of the controlled variable at set point as shown in fig.6. The controller receives the difference between the reference set point and the measured output and generates a control action to make error to zero.
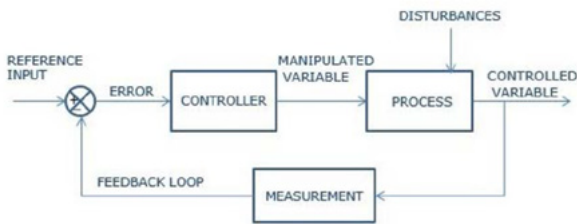

Figure. 6: *Structure of closed loop control system*

PID controller is shown in fig.7 has the optimum control dynamics including zero steady state error, fast response (short rise time), no oscillations and higher stability. The necessity of using a derivative gain component in addition to the PI controller is to eliminate the overshoot and the oscillations occurring in the output response of the system. One of the main advantages of the P-I-D controller is that it can be used with higher order processes including more than single energy storage.
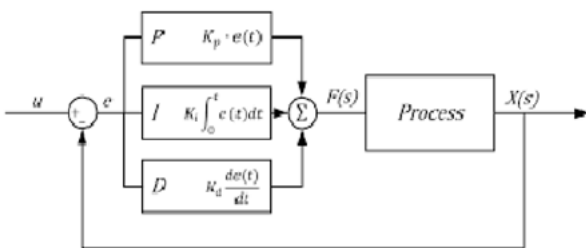

Figure. 7: *PID controller*

TUNING OF PID CONTROLLER

1. Set all gains to zero.

2. Increase the P gain until the response to a disturbance is steady oscillation.

3. Increase the D gain until the oscillation go away (i.e. it's critically damped).

4. Repeat steps 2 and 3 until increasing the D gain does not stop the oscillation.

5. Set P and D to the last stable values.

6. Increase the I gain until it brings you to the set point with the number of oscillation desired (normally zero but a quicker response can be had if you don't mind a coupled oscillations of overshoot).

   a. ZEIGLERS-NICHOLS METHOD

    It is another popular method for tuning PID controllers. Ziegler and Nichols presented two classical methods for determining values of proportional gain, integral time and derivative time based on transient response characteristics of a given plant or system [13].

FIRST METHOD

• Obtain a unit step response of the plant experimentally and it may look's' shaped curve as shown in figure below. This method applies, if obtained response exhibit s-shaped curve for unit step input otherwise it cannot be applied. This curve can also be obtained by dynamic simulation of the plant as shown in Fig. 8
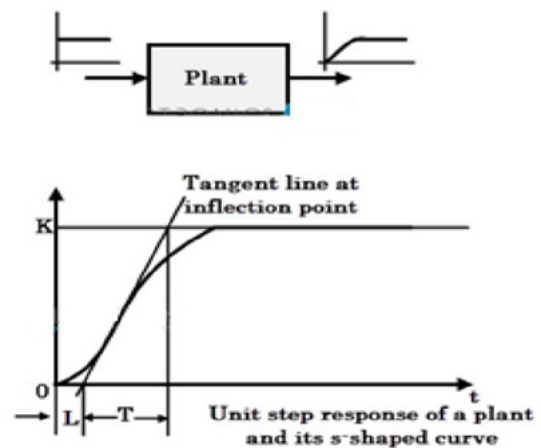

Figure. 8: *Tuning of PID controller with system producings shaped curve*

- Obtain two constants, delay time L and time constant T by drawing a tangent line at the inflection point of the s-shaped curve.

- Set the parameters of Kp, Ti, and Td values from the table given below for three types of controllers.

SECOND METHOD

- It is very similar to the trial and error method where integral and derivative terms are set to the zero, i.e., making Ti infinity and Td zero [14].

- Increase the proportional gain such that the output exhibits sustained oscillations as shown in Fig. 9. If the system does not produce sustained oscillations then this method cannot be applied. The gain at which sustained oscillations produced is called as critical gain.
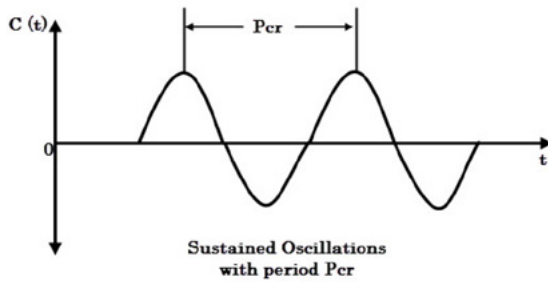

Figure. 9: *Tuning of PID controller with system producing sustained oscillations*

Once the sustain oscillations are produced, set the values of Ti and Td as per the given table for P, PI and PID controllers based on critical gain and critical period.

III. HE PROPOSED SYSTEM

In this paper LFC issue in a two area interconnected multi source power system is presented. Each area of the power system comprises of a reheat thermal unit, a hydro unit and a gas unit. A step load change of 0.01 pu (1%) is applied in area1 to study the dynamic behavior of both the power systems (with and without HVDC link). Each area is equipped with a PID controller to control the frequency and tie-line power deviations. The linearized transfer function model of two area interconnected power system is shown in Fig.10. The parameters of the power system and their nominal values are given in Table 1. There are three generating units in each area and each unit has its own speed governing system, turbine and generator.
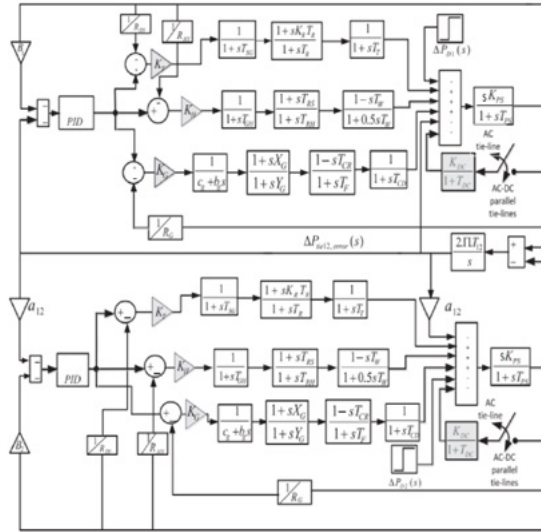

Figure. 10: *The Proposed System*

Table- 1: Operating parameters of proposed system and their nominal value

Parameters	Symbol	Nominal value
Frequency Bias Constants	B_i	0.4312 MW/Hz
Tie Line Power Coefficient	T_{12}	0.0433
Regulation Time Constant	R_i	2.4 Hz/MW
Control Area Capacity Ratio	a_{12}	−1
Participation factors for Thermal	K_T	0.543478
Participation factors for Hydro	K_H	0.326084
Participation factors for Gas	K_G	0.130438
Control Area Gain	K_{PS}	68.9566
Control Area Time Constant	T_{PS}	11.49 s
Governor Time Constant	T_{SG}	0.08 s
Turbine Time Constant	T_T	0.3 s
Reheat Gain	K_R	0.3
Reheat Time Constant	T_R	10 s
Hydro Turbine Speed Governor Time Constant	T_{GH}	0.2 s
Hydro Turbine Speed Governor Reset Time	T_{RS}	5 s
Hydro Turbine Speed Governor Transient Droop Time Constant	T_{RH}	28.75 s
Nominal Starting Time Of Water in penstock	T_W	1 s
Gas Turbine Constant Of Valve Positioner	b_g	0.05
Gas Turbine Valve Positioner	c_g	1
Lag Time Constant Of Gas Turbine Speed Governor	Y_G	1 s
Lead Time Constant Of Gas Turbine Speed Governor	X_G	0.6 s
Gas Turbine Combustion Reaction Time Delay	T_{CR}	0.01 s
Gas Turbine Fuel Time Constant	T_F	0.23 s
Gas Turbine Compressor Discharge Volume– Time Constant	T_{CD}	0.2 s
Gain of HVDC Link	K_{DC}	1.0
Time Constant of HVDC Link	T_{DC}	0.2 s

Input to PID controllers are their respective area control errors (ACEs) which are given by,

$$ACE_1 = \Delta P_{12} + B_1\,\Delta f_1$$

$$ACE_2 = \Delta P_{21} + B_2\,\Delta f_2$$

where ΔP_12 is the change in tie-line power in area1 and ΔP_21 is that in area2, B_1 and B_2 are frequency bias factors of area1 and area2 respectively. Δf_1 is the frequency deviation in area1 and Δf_2 is that in area2.[24]

Table- 2: PARAMETERS

Parameters	Symbol	Nominal value
Frequency Bias Constants	B_i	0.4312 MW/Hz
Tie Line Power Coefficient	T_{12}	0.0433
Regulation Time Constant	R_i	2.4 Hz/MW
Control Area Capacity Ratio	a_{12}	−1
Participation factors for Thermal	K_T	0.543478
Participation factors for Hydro	K_H	0.326084
Participation factors for Gas	K_G	0.130438
Control Area Gain	K_{PS}	68.9566
Control Area Time Constant	T_{PS}	11.49 s
Governor Time Constant	T_{SG}	0.08 s
Turbine Time Constant	T_T	0.3 s
Reheat Gain	K_R	0.3
Reheat Time Constant	T_R	10 s
Hydro Turbine Speed Governor Time Constant	T_{GH}	0.2 s
Hydro Turbine Speed Governor Reset Time	T_{RS}	5 s
Hydro Turbine Speed Governor Transient Droop Time Constant	T_{RH}	28.75 s
Nominal Starting Time Of Water in penstock	T_W	1 s
Gas Turbine Constant Of Valve Positioner	b_g	0.05
Gas Turbine Valve Positioner	c_g	1
Lag Time Constant Of Gas Turbine Speed Governor	Y_G	1 s
Lead Time Constant Of Gas Turbine Speed Governor	X_G	0.6 s
Gas Turbine Combustion Reaction Time Delay	T_{CR}	0.01 s
Gas Turbine Fuel Time Constant	T_F	0.23 s
Gas Turbine Compressor Discharge Volume– Time Constant	T_{CD}	0.2 s
Gain of HVDC Link	K_{DC}	1.0
Time Constant of HVDC Link	T_{DC}	0.2 s

SIMULATION RESULTS & ANALYSIS

To study the dynamic performance of the conventional multi-source power system with HVAC tie-lines only and with both HVAC and HVDC tie-lines, a step load perturbation of 1% i.e. 0.01 p.u. is applied in area1. The gains of the conventional PID controllers are optimized to KP = 0.6227 KI = 1.787 KD = 0.8269.

a. WITHOUT HVDC LINK

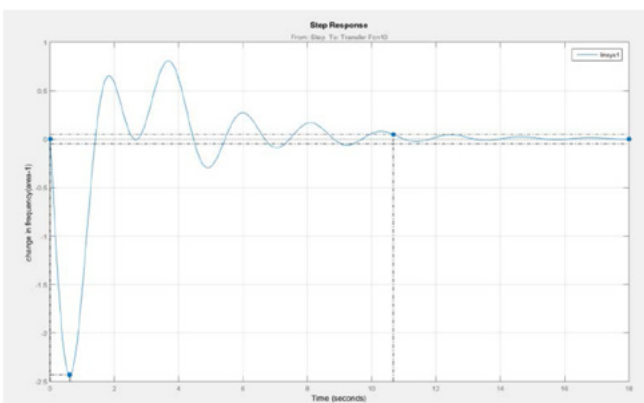The results without HVDC link is shown in Fig. 11 to 13 and Table 2 specifies the values.

Figure. 11: *Linearized Model of change in frequency of area-1 of ALFC system without HVDC link*

Figure. 12: *Linearized Model of change in frequency of area-2 of ALFC system without HVDC link*

Figure. 13: *Linearized Model of tie-line power of ALFC system without HVDC link*

Table- 2: Performance parameters without HVDC link

	PEAK RESPONSE	SETTLING TIME	RISE TIME	STEADY STATE
AREA 1	-2.43	10.7	0	0
AREA 2	-1.73	13.8	0	0
TIE-LINE POWER	-2.13	12.3	0	0

b) WITH HVDC LINK

The results with HVDC link is shown in Fig. 14 to 16 and Table 3 specifies the values.

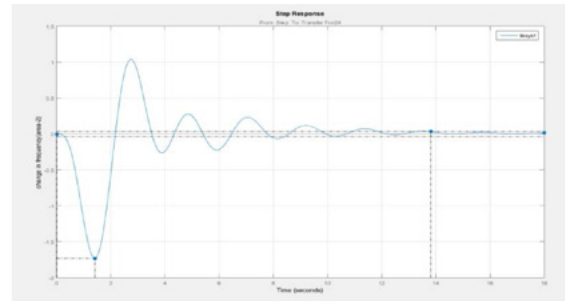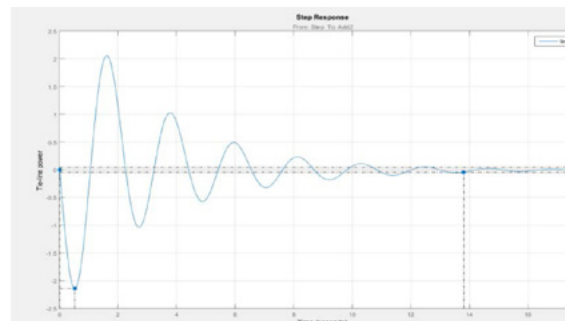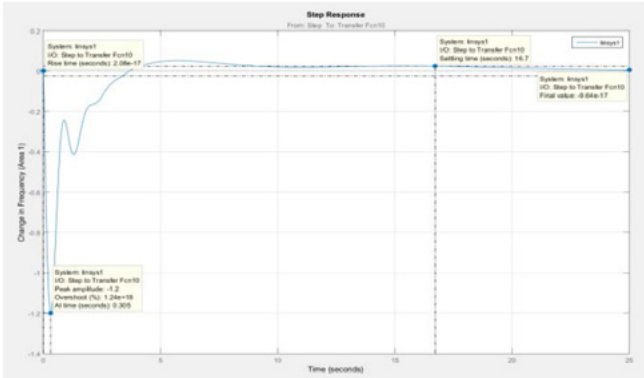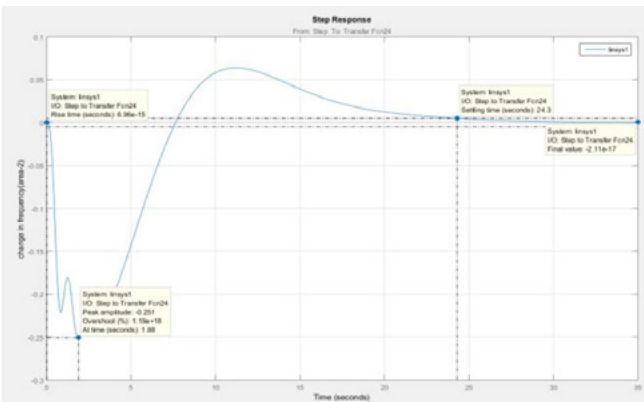*Figure. 14: Linearized Model of change in frequency of area-1 of ALFC system with HVDC link*

Figure. 15: Linearized Model of change in frequency of area-2 of ALFC system with HVDC link
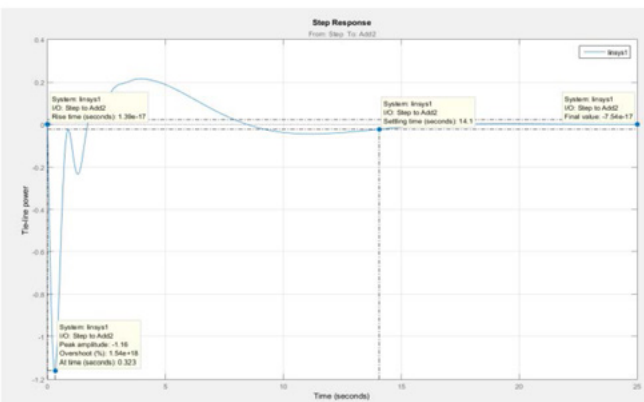

Figure. 16: Linearized Model of tie-line power of ALFC system with HVDC link

Table- 3: Performance parameters with HVDC link

	PEAK RESPONSE	SETTLING TIME	RISE TIME	STEADY STATE
AREA 1	-1.2	16.7	2.08e-1	-9.64e-17
AREA 2	-1.2	16.7	0	0
TIE-LINE POWER	-1.23	14.1	1.39e-17	-7.22e-17

Due to the remarkable growth of High Voltage Direct Current (HVDC) transmission system in the past decade, HVDC links are now being used in parallel with the existing EHVAC (extra high voltage AC) for interconnecting two control areas. HVDC link in parallel with HVAC link not only offers environmental and economic benefits but also ensures good quality of electrical power supply. It also improves the dynamic performance of the system when the system is subjected to small disturbances.

V. CONCLUSIONS

In this paper the gains of the conventional PID controller is optimized, by tuning of the PID controller, in order to tackle the load frequency control problem of a two-area interconnected multi-source power system. Each area of the power system has a reheat thermal unit, a hydro unit and a gas unit and two PID controllers are used in this study, one for each area. First, a step load change of 0.01 pu (1%) is applied in area1, and then, step load change of 0.01 pu (1%) is applied in both the areas to study the dynamic behavior of both the power systems (with and without HVDC link). From the results obtained it is concluded that HVDC link along with HVAC link i.e., hybrid tie-lines yielded superior results than the conventional HVAC tie-lines.

REFERENCES

1. Cohn N., "Some aspects of tie-line bias control on interconnected power systems", Am Inst Elect Eng Trans 1957, Vol. 75, p.p.1415–36.

2. Khuntia SR, Panda S., " Simulation study for automatic generation control of a multi-area power system by ANFIS approach", Appl Soft Comput 2012, Vol.12, p.p.333–41.

3. Abraham RJ, Das D, Patra A. "Automatic generation control of an interconnected hydrothermal power system considering superconducting magnetic energy storage", Electr Power Energy Syst 2007, Vol.29, p.p.571–9.

4. Ali ES, Abd-Elazim SM. "Bacteria foraging optimization algorithm based load frequency controller for interconnected power system.", Electr Power Energy Syst 2011, Vol. 33, p.p.633–8.

5. Sabahi K, Sharifi A, Aliyari M, Teshnehlab M, Aliasghary M. "Load frequency controller in interconnected power system using multi-objective PID controller", Journal of Appl Sci 2008 Vol. 8, p.p.3676–82

6. Lee KA, Yee H, Teo CY. "Self-tuning algorithm for automatic generation control in an interconnected power system.", Electr Power Syst Res 1991, Vol.20 p.p.157–65.

7. Nanda J, Mishra S, Saikia LC. "Maiden application of bacterial foraging based optimization technique in multi-area automatic generation control.", IEEE Trans Power Syst 2009, Vol.24(2), p.p.602–9.

8. Lu CF, Liu CC, Wu C. "Effect of battery energy storage system on load frequency control considering governor dead-band and generation rate constraints", IEEE Trans Power Syst 1995, Vol.10(3), p.p.555–61

9. Chaturvedi DK, Satsangi PS, Kalra PK. "Load frequency control: a generalized neural network approach.", Electr Power Energy Syst 1999, Vol.21(6), p.p.405–15.

10. Ghosal SP. "Optimization of PID gains by particle swarm optimization in fuzzy based automatic generation control", Electr Power Syst Res 2004, Vol.22(3), p.p. 203–12.

11. Mohanty B, Panda S, Hota PK. "Controller parameters tuning of differential evolution algorithm and its application to load frequency control of multisource power system", Electr Power Energy Syst 2014, Vol.54,p.p.77–85.

12. https://www.electricaleasy.com/2015/08/thermal-power-plant.html

13. http://www.srmuniv.ac.in/sites/default/files/downloads/chapter2_power_system_operation.pdf

14. https://www.electronicshub.org/pid-controller-working-and-tuning-methods/

Sourav Kumar Bisoy, Biswajit Sahoo, Ujjwal Sinha, Sidhant Patra and Bikash Kumar Parida

Department of Electrical and Electronics Engineering.

ENERGY-AWARE ROUTING ALGORITHMS IN SOFTWARE DEFINED NETWORKS

Now-a-days, networking industries are developing very fast. For flawless operation, it demands of intelligent network programs. This has encouraged the expansion of Software Defined Networking (SDN). In addition to network programmability, SDN decouples data plane and control plane. The central controller presents in control plane takes the overall decision of the network without interfering the backbone network architecture. The result of this, SDN is widely employed in data centres, rural connection like many more applications. However, there are certain issues inherent to SDN controller placement, correctness of flow table entries, traffic engineering, load balancing, security and energy efficient, etc. These issues have to be handled in a careful manner for successful operations of SDN.

In SDN, major parts of energy consumption are done by network components like servers, switches, routers and links. The maximum amount of energy consumption is associated with lighting equipment's, UPS power supply, air conditioning, refrigeration, and etc. The rapid growth in energy consumption in Information and Communication Technologies (ICT) and CO2 emissions have an adverse effect on society and the environment. This causes a decrease in the performance of SDN. Energy consumption puts an important challenge in the operation of SDN. Therefore, to increase the efficiency of SDN these issues need to be addressed properly. In this thesis, analytical models were developed to compare the performance of the network and energy savings. These models help the network administrator to adjust the other parameters of the system to achieve high throughput with low latency to meet the network performance.

The network performance is directly proportional to energy savings of the network. To achieve the major amount of energy saved, some of the techniques are based on adaptive link rates, idle to sleep power mode, monitoring of flow table, etc. To optimize the memory size of the flow tables, Ternary Content Addressable Memory (TCAM) is monitored effectively. Frequent access to TCAM may cause a major amount of energy consumption. TCAM is very expensive and power hungry. To prevent this in heavy traffic load the controller installs the flow table in switches under some constraints. Network intelligence along with virtualization plays an important role to save a major amount of energy consumption.

Unutilized link operated in high data rate is the major cause of energy consumption in wired networks. In this thesis, we have proposed energy efficient routing algorithm that uses discrete data link rate. Energy consumption can be reduced to a certain level if all the links start operating as per their utilization factor. The output buffer queue with two levels of threshold values is used to monitor the behaviour of the network. In another approach we have considered end host awareness to decide which idle link or switch to put into a sleep state. The controller takes this decision. At the same time, it cares about network partition.

Energy consumption in SDN is affected by the load of the network. In high traffic condition energy consumption is increased exponentially with load of network. In this thesis, we have proposed some energy efficient techniques by carefully monitoring the data rate, number of idle device and links in SDN, etc. These methods help to increase the performance characteristics of SDN. In one of the methods, the load of the network is shared by distributing the load in the alternative path. In another approach, the two layers of controller architecture are used to balance the load of the network. Logically, two kinds of controllers are used. One form is super controller and another form is sub controller. In high traffic the flow table entries are frequently changing. Frequent access of TCAM may cause a major amount of energy loss. To prevent this frequent monitoring of TCAM is required. Wild card entry and default port can be used to minimize the number of entries for flow table TCAM. An analytical model has been derived to minimize energy consumption in SDN using flow table monitoring.

Plastic Waste

The Novel Coronavirus has had severe impact on global society, and debilitating effects in India. The obvious are loss of jobs and income, decline of the general health of the economy, and the ill-effects on physical and mental health. However, there are other tangible effects that COVID-19 has brought with it - generation of a tremendous amount of plastic waste from the healthcare industry. Personal protective equipment used by frontline healthcare workers include gowns, gloves and masks which are disposed of. This, along with the regularly created disposable plastic produces a significant volume of waste. How does one tackle this problem?

Recycling is one answer. According to sources, India with an approximately 60% recycle rate, is actually a relatively innocuous producer of plastic waste when compared with the fact that only about 9% of all plastic ever produced has been recycled[1]. The ideal pathway for the recycle process is to carefully segregate, collect, transport and treat. This, however, is easier said than done.

One such effort launched in India by Dow Chemical is called Rethink+ [2]. This program connects the plastic waste generators, aggregators, processors and recyclers. A doorstep plastic waste pickup is scheduled via an app. Also on-board are IT companies providing cloud-based solutions for waste recycling and NGOs for creating awareness. The program was a resounding success in Pune, and will hopefully permeate to other parts of the country. A lot of recycled plastic also ends up in the construction business as portion of an aggregate mixture for making roads and concrete blocks. Innovative ideas and solutions are the need of the hour to tackle the plastic waste problem - otherwise our oceans and land are rapidly becoming dumping grounds.

Sources:

[1] United Nations data
[2] Dow India launches Rethink+, India's first digital waste management program to recycle plastic waste

The Science & Technology Magazine

Digital
Digest

Silicon Institute of Technology
सिलिकन प्रौद्योगिकी संस्थान

Bhubaneswar Campus
An Autonomous Institute
Silicon Hills, Patia
Bhubaneswar - 751024

Sambalpur Campus
An Affiliated Institute
Silicon West, Sason
Sambalpur - 768200

Contents

Editorial Team
Dr. Jaideep Talukdar
Dr. Pamela Chaudhury
Dr. Lopamudra Mitra

Members
Dr. Bhagyalaxmi Jena
Nalini Singh
Dr. Soumya Priyadarsini Panda
Dr. Priyanka Kar

Student Members
Tanmaya Bal
Rohit Kumar Nayak

Media Services
G. Madhusudan

Circulation
Sujit Kumar Jena

Make your submissions to:
publication@silicon.ac.in