

Light Actuation Based On Facial Mood Recognition

¹Abhi Kadam *, ²Anupama Mhatre, ³Sayali Redasani, ⁴Amit Nerurkar

^{1,2,3,4}Vidyalankar Institute of Technology, Mumbai University, Mumbai, India

Abstract:

Current lighting technologies extend the options for changing the appearance of rooms and closed spaces, as such creating ambiances with an affective meaning. Using intelligence, these ambiances may instantly be adapted to the needs of the room's occupant(s), possibly improving their well-being. In this paper, we set actuate lighting in our surrounding using mood detection. We analyze the mood of the person by Facial Emotion Recognition using deep learning model such as Convolutional Neural Network (CNN). On recognizing this emotion, we will actuate lighting in our surrounding in accordance with the mood. Based on implementation results, the system needs to be developed further by adding more specific data class and training data.

Keywords: Mood Detection, Convolutional Neural Network, Deep Learning.

1. Introduction

Emotions represent the mood of a person. Emotions constitute an innate and important aspect of human behaviour that colours the way of human communication. They are influenced by personality and temperament and determine the behavior of a person, being closely linked to the nervous system. According to medicine and psychology, emotions are generated by the same central nervous system that controls the entire body, so if a person is experiencing different moods, the system increases or decreases the oxygen levels in the muscles, temperature or pulse. It works like an automatic system.[1] In the current context, the trend in technology is tilted largely towards developing applications that increase the comfort of people and the interaction between humans and computers. The most known example of emotions detection application is based on interpreting facial expressions using image processing algorithms. This area is advanced, leading to facial micro expressions study in order to detect instant mood changes at the level of milliseconds. Studies show that mood states influence our daily life quality and activities, and this is not the only way around.

Research shows that an adequate amount of light improves mood and energy levels, while poor

lighting contributes to depression and other deficiencies in the body. The amount and type of lighting directly affect concentration, appetite, mood, and many other aspects of daily life. Light and color in general have been repeatedly shown to be able to influence people's mood, both at a biological and

Psychological level. From a biological point of view, (short wavelength) light can influence the circadian rhythm, and as such the mood. At a psychological level, (colored) lighting has been shown to carry meaning through associations, thereby influencing affective states[2]. For example, employees working in colorful offices tend to experience more positive moods than employees sitting in neutral or colorless offices. Different hues are also commonly associated with arousal levels: blue, for example, is deemed to be calming, whereas red is considered to be more energizing.

Ambient intelligence research is devoted to designing systems that, embedded in our everyday environments, can help in improving the quality of our lives. Among the various facets of ambient intelligence, a particularly interesting one is the so-called "empathic technology": systems able to detect the affective state of a person and react on it by providing the necessary care. Ambient Lighting can be set using the current mood of the person. One way

to achieve this is to use intelligent lighting solutions to create pleasant ambiances in our rooms. The vast growth in the amount of LED sources, their miniaturization and the possibility to integrate sensors and electronics has created a wealth of options to automatically generate light of any color and intensity at any place or time with very thin, unobtrusive luminaries, making it possible to re-configure light according to a promptly detected negative mood of a room's occupant.

Machine Learning approaches can also be applied to the mood detection system such as deep learning which has the primary concept in studying features of data. There are several models of deep learning. One of them is Convolutional Neural Networks (CNN) which is best suited to work with images. One of CNN research was applied to 3D recognizing sign language using data of hand gestures video taken by using Microsoft Kinect as system and obtained accuracy of 88.5% in gray channel, 94.2% in multi- channel[3]. A character recognition using CNN was also conducted and obtained error rate 4.5% by using 2 hidden layers and 3.05% by using first hidden layer [4]. CNN was used for classifying human emotions from dynamic facial

Expressions in real time which training accuracy of 90.7% and test accuracy of 57.1 % is achieved [5] and another research about facial affect detection using CNN had an accuracy result of 38.1% [6].

2. Convolution Neural Network (Cnn)

CNN is a type of deep neural network which is designed to process 2 dimensional data. CNN can be applied in the image data and sound detection learning process [7]. CNN consists of one or more convolutional and subsampling layer followed by a layer which connects altogether such as a standard neural network. CNN architecture can be seen in Fig. 1.

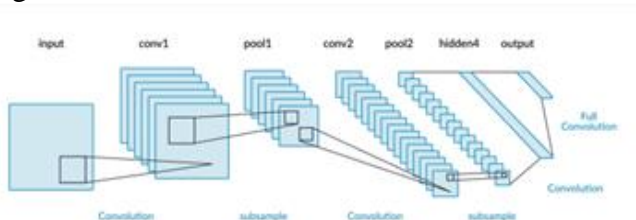


Fig.1 Architecture of Convolution Neural Network [5]

Convolutional layer conducts convolutional operation on the output resulted from the previous layer. The layer is a main process for a basic CNN. Within image processing, convolution means to apply a kernel into overall image of enabled offset. Kernel moves from the upper left corner into the bottom right corner. In each movement, Kernel conducts addition from multiplication of each point within the kernel with each point within image entered. The result is a convolution of pixel value output in the form of new matrix [7].

Subsampling layer is a process of reducing sizes of image data. Within image processing, subsampling aims to increase position invariance of features. In most CNN methods, the used subsampling is max pooling. Max pooling divides output from convolutional layer into small grids and extracts the maximal value of each grid to compile matrix of reduced image [7]. The size of images reduces as it goes deeper into the CNN and the number of kernels increases as we go deeper into CNN.

3. Methodology

There are various stages of implementation such as literature study, data collection and preparation, system design and development, analysis and evaluation as well as conclusion. The design of implementation stages is seen in Fig. 2.

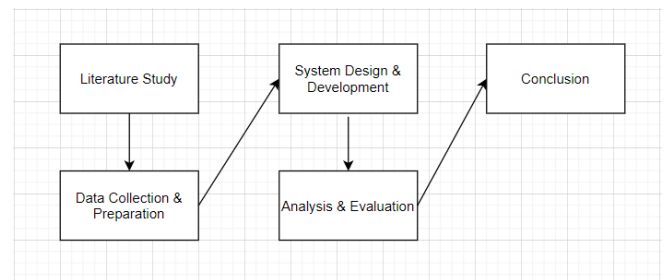


Fig. 2 Implementation Stages

1) *Literature Study:* The literature study was a process for finding out sources and references related to the research. This stage includes study about image processing, computer vision, face detection, artificial neural network, deep learning, and convolutional neural network. It also consisted of research and study of hardware required for light actuation.

2) *Data Collection and Preparation:* The dataset used to validate the system is the FER2013 dataset. This dataset consists of 48x48 grayscale

images of various emotions such as Angry, Happy, Sad, Surprise as well as Neutral. CNN method was used to create a mood detection system of face image and its label. Then Neural Network is implemented to classify the image according to the available label.

3) *System Design and Development:* Waterfall method was used for developing the system. The stages are analysis, design, coding and testing.

4) *Analysis and Evaluation:* The system's output data were analysed and evaluated at this stage. The analysis and evaluation process addressed to find out the level of accuracy and losses.

5) *Conclusion:* At this stage, necessary conclusions were drawn and limitations as well as future scope was discussed.

A. Mood Detection System

The Mood Detection System is a Deep Learning model called Convolution Neural Network. This system classifies the images in one of the five standard emotions such as Angry, Happy, Neutral, Sad or Surprise. The first step was to capture the image of the person. After the image is captured, the next step was to detect the face of the person. For that purpose, Haar Cascade Classifier for frontal face detection was used. The image captured requires some preprocessing to be done so that they can be prepared for training and testing and extract necessary information which can be used for light actuation.

The following preprocessing steps were used:

1. *Scaling:* Image scaling refers to the resizing of a digital image. Our model was trained on grayscale images of 48x48 resolution. Each frame was scaled by resizing frame resolution at first 264x264 pixels into 48x48 pixels.

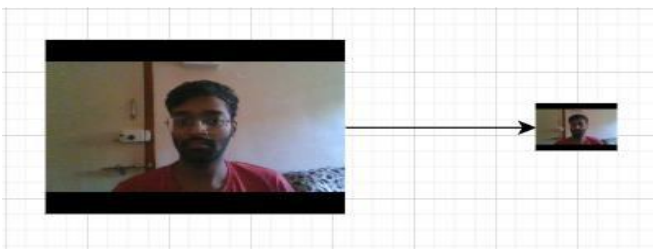


Fig. 3: Scaling Process

2. *Grayscale:* Image is a representation of matrix values of pixels. The color image is converted into grayscale color space. The left image is an original one and the right image is a converted one. It is done in order the process of calculation will not be larger, hard, too long and simple.



Fig. 4: Grayscale Process

3. *Histogram normalization:* Histogram Equalization is a computer image processing technique used to improve contrast in images. It accomplishes this by effectively spreading out the most frequent intensity values, i.e. stretching out the intensity range of the image. This method usually increases the global contrast of images when its usable data is represented by close contrast values. This allows for areas of lower local contrast to gain a higher contrast.

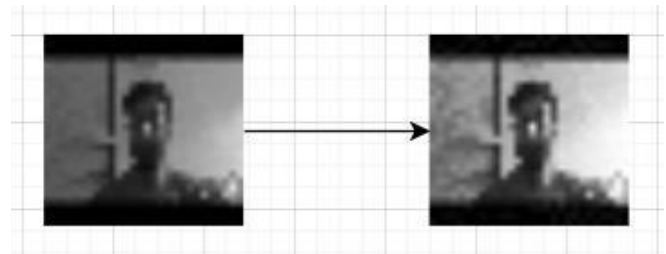


Fig. 5: Histogram Normalization Process

Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. A characteristic of these large data sets is a large number of variables that require a lot of computing resources to process. Feature extraction is the name for methods that select and /or combine variables into features, effectively reducing the amount of data that must be processed, while still accurately and completely describing the original data set. The Convolution Neural Network is used to learn and extract features from input data then be classified to respective class. The features extraction and

classification process use layers of CNN model. There are 4 blocks of layers consisting of 2 convolution layers followed by max pooling layer. Max pooling by the scale of 2x2 was used within subsampling layer. The number of epochs used are 25 and the batch size is 32. Training process used backpropagation network, while testing process used feedforward network. The initial kernel value within the training process was random, then it was renewed to determine classification pattern.

Based on the scenario testing result, the highest accuracy of 60% is obtained by using epoch 25 and batch 32.

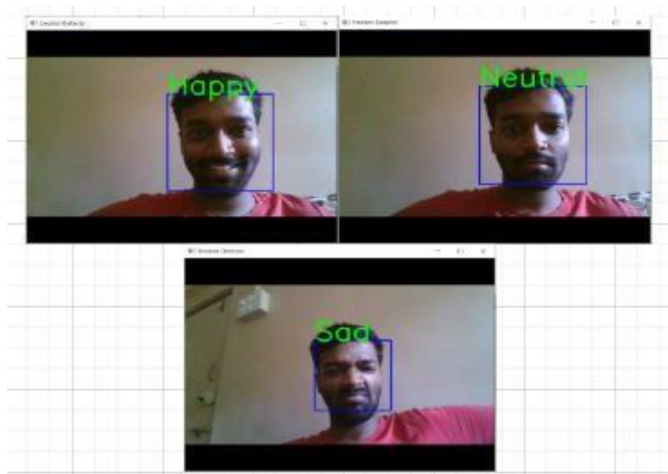


Fig. 6: Emotion Detector Output

B. Light Actuation

Light is one of the principal environmental factors that influence on a person. Colour is another factor that is constantly present in any environment. Many authors highlight that certain colours are highly influencing for emotion elicitation and regulation. The four possible explanations to colour and emotion association, which include references to metonymic and metaphoric thinking, formation of specific emotional reactions for colour perception, and sharing connotative structure in the language for colour and emotion terms, have been validated [8].

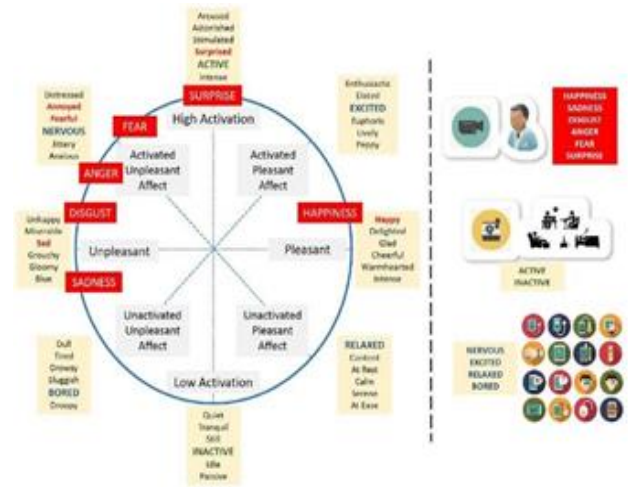


Fig. 7: Circumplex model for emotion detection [2].

For Light Actuation, the hardware set up has LED strips with the Arduino due to its simplicity and versatility, its strong executive power and reliability. The aim is to change the colour of LEDs as per mood detected. The LED are connected to breadboard with 220 ohms resistors and jumping wires. Arduino receives the information about the emotion from the computer through the USB serial port. The Arduino is coded to actuate light according to that mood.

Based on Circumplex Model for emotion detection [2], the lights are actuated in the following ways:

1. Activated Unpleasant Affect has to transit to Unactivated Pleasant Affect. This is achieved in two steps. Firstly, colour/light LED strips are changed to Blue, Light blue or Purple, and, then to Pink, Green or Yellow. The order in these lists correspond to the velocity at which these colours provide the desired effect. Obviously, this has also to be personalised in accordance with the person’s reactions[2].
2. Unactivated Unpleasant Affect transits directly to Unactivated Pleasant Affect with colour/light Pink, Green or Yellow. Remember that Unactivated Unpleasant Affect is a health affective state, but try to maintain the person as “Pleasant” as possible. This is why, it is foreseen that the affective state will balance between Unactivated Unpleasant Affect and Activated Pleasant Affect. In order to transit to Activated Pleasant Affect newly two steps are necessary. In the first one, colour/light is changed to Orange, and in the second one colour/light is changed to Red[2].

3. The Activated Pleasant Affect case is similar to the previous one. Here, firstly colour/light is changed to Blue, and in the second one colour/light is changed to Yellow.[2]

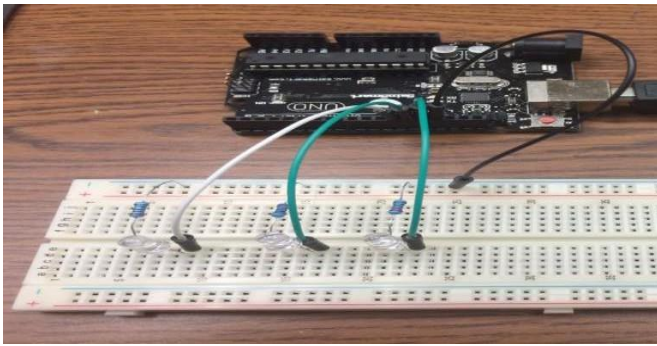


Fig. 8: Light Actuation Setup

Conclusion

The mood detection system developed uses the Convolutional Neural Network method which gives a good accuracy, though the data during the training process is limited. The system can be further enhanced by increasing its accuracy. Many other parameters such as sound, pulses, etc. can be used to detect the mood. Based on detected mood, light actuation of surrounding was also achieved.

Acknowledgment

We take this opportunity to express our deepest gratitude towards our project guide Prof. Amit Nerurkar, who has been the driving force behind this project and whose guidance and co-operation has been a source of inspiration for us.

References

- [1.] Alexandra Cernian, Adriana Olteanu, Dorin Carstoiu, Cristina Mares, "Mood Detector – On Using Machine Learning to Identify moods and Emotions", University Politehnica of Bucharest, 2017
- [2.] Antonio Fernández-Caballero, Arturo Martínez-Rodrigo, José Manuel Pastor, José Carlos Castillo, Elena Lozano-Monator, María T. López, Roberto Zangróniz, José Miguel Latorre, Alicia Fernández-Sotos, "Smart environment architecture for emotion detection and regulation", Journal of Biomedical Informatics 64 (2016) 55–73
- [3.] J. Huang, W. Zhou, H. Li and W. Li, "Sign Language Recognition Using 3D Convolution Neural Networks", Ghent University, ELIS, Belgium, 2014.
- [4.] L. Deng, "The MNIST Database of Handwritten", IEEE Signal Processing Magazine, pp. 141-142, 2012.
- [5.] D. Duncan, G. Shine, and C. English. "Facial Emotion Recognition in Real Time", Stanford University, 2016.
- [6.] S. Wang. "Facial Affect Detection Using Convolutional Neural Networks", Stanford University, 2016.
- [7.] S. Anwar, Hwang, K., and Sung, W. "Fixed Point Optimization of Deep Convolutional Neural Networks for Object Recognition", ICASSP, 1131-1135, 2015.
- [8.] C. Soriano, J. Valenzuela, Emotion and colour across languages: implicit associations in Spanish colour terms, Soc. Sci. Inform. 48 (3) (2009) 421–445.