IJSR CSEIT

International Journal of Scientific Research in Computer Science, Engineering and Information Technology
 | Volume 3 | Issue 3 | ISSN : 2456-3307

Identifying Eye Fixation to Control the Mouse Operations Using Artificial Neural Network

K. N. Priyanka[1], Dr. P. Amudha[2]

[1]PG Scholar, Computer Science and Engineering, Faculty of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India

[2]Associate Professor, Head of the Department Computer Science and Engineering, Faculty of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India

ABSTRACT

Disability is the consequence of an impairment that may be physical, cognitive, mental, sensory, emotional, developmental, or some combination of these. The disability management is a critical task since it is caused by employing a digital system to assist the physically disabled people. This process is completed by applying a digital signal processing system which takes the analog input from the disabled people by using dedicated hardware with software, and then the raw data is converted it into informative data in the form of digital signal. The Iris tracking is the process of measuring either the point of gaze (where one is looking) or the motion of the eye relative to the head. Face and eye detection algorithm is used to identify and recognize the user input using web camera. The predefined pattern matching is initialized for detecting a face from the USB camera feed. Pattern matching algorithm is used to compute the pattern related to the mouse movement.Supervised learning with back propagation is used to identify the input type and click operation on the particular action event. When the position of the user is sufficiently constant, the system for detecting and analyzing blinks and mouse movements is initialized automatically, depending on the involuntary blink of the user. This approach is efficient for disabled people.

I. INTRODUCTION

Image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, a series of images, or a video, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. The word image is also used in the broader sense of any two-dimensional figure such as a map, a graph, a pie chart, or an abstract painting An image is a rectangular grid of pixels. It has a definite height and a definite width counted in pixels Most image-processing techniques involve treating the image as a two-dimensional signal and applying standard signal-processing techniques to it. Images are also processed as three-dimensional signals where the third-dimension being time or the z-axis.

Image processing basically includes the following three steps:

- ✓ Importing the image via image acquisition tools.
- ✓ Analyzing and manipulating the image.
- ✓ Output in which result can be altered image or report that is based on image analysis.

Image processing usually refers to digital image processing, but optical and analog image processing also are possible. The acquisition of images (producing the input image in the first place) is referred to as imaging. Pixel is the term most widely used to denote the elements of a digital image. Each pixel has a color.

CLASSIFICATION OF IMAGES

There are 3 types of images used in Digital Image Processing. They are

1. Binary Image
2. Gray Scale Image
3. Color Image

BINARY IMAGE

A binary image is a digital image that has only two possible values for each pixel. Typically the two colors used for a binary image are black and white though any two colors can be used. The color used for the object(s) in the image is the foreground color while the rest of the image is the background color.

Binary images are also called bi-level or two-level. This means that each pixel is stored as a single bit (0 or 1).This name black and white, monochrome or monochromatic are often used for this concept, but may also designate any images that have only one sample per pixel, such as grayscale images

Binary images often arise in digital image processing as masks or as the result of certain operations such as segmentation, thresholding, and dithering. Some input/output devices, such as laser printers, fax machines, and bi-level computer displays, can only handle bi-level images

GRAY SCALE IMAGE

A grayscale Image is digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray(0-255), varying from black(0) at the weakest intensity to white(255) at the strongest.

Grayscale images are often the result of measuring the intensity of light at each pixel in a single band of the electromagnetic spectrum (e.g. infrared, visible light, ultraviolet, etc.). The conversion of a color image into a grayscale image is converting the RGB values (24 bit) into grayscale value (8 bit).

Figure 1. Gray Scale Image

Grayscale images are distinct from one-bit black-and-white images, which in the context of computer imaging are images with only the two colors, black, and white (also called bi-level or binary images). Grayscale images have many shades of gray in between. Grayscale images are also called monochromatic, denoting the absence of any chromatic variation.But also they can be synthesized from a full color image; see the section about converting to grayscale.

COLOR IMAGE

A (digital) color image is a digital image that includes color information for each pixel. Each pixel has a particular value which determines its appearing color. This value is qualified by three numbers giving the decomposition of the color in the three primary colors Red, Green and Blue. Any color visible to human eye can be represented this way. The decomposition of a color in the three primary colors is quantified by a number between 0 and 255. For example, white will be coded as R = 255, G = 255, B = 255; black will be known as (R,G,B) = (0,0,0); and bright pink will be : (255,0,255).

In other words, an image is an enormous two-dimensional array of color values, pixels, each of them coded on 3 bytes, representing the three primary colors. This allows the image to contain a total of 256x256x256 = 16.8 million different colors. This technique is also known as RGB encoding, and is specifically adapted to human vision.

1.3 STEPS INVOLVED IN DIGITAL IMAGE PROCESSING

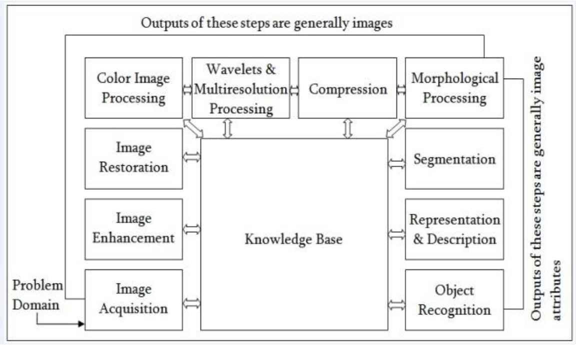

Figure 2. Steps involved in Digital Image Processing

Digital image processing techniques help in manipulation of the digital images by using computers The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction. Image acquisition is the first process acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves pre-processing, such as scaling Closely related to image processing are computer graphics and computer vision. In computer graphics, images are manually made from physical models of objects, environments, and lighting, instead of being acquired (via imaging devices such as cameras) from natural scenes, as in most animated movies. Computer vision, on the other hand, is often considered high-level image processing out of which a machine/computer/software intends to decipher the physical contents of an image or a sequence of images (e.g., videos or 3D full-body magnetic resonance scans).

II. LITERATURE SURVEY

Saleh Albelwi et.al[1] proposed a new optimization objective function that pools the error rate and the information cultured by a set of feature maps by means of deconvolutional networks. Compare to Nelder-Mead Method (NMM) the CNN architecture improves performance, reduces error rate and it is more effective. CNN uses MNIST dataset which performs better than CIFAR-10 and CIFAR-100 datasets. This approach reduces the size of strides, provides high recognition performance.

Srinivas S S Kruthiventi et.al[2] proposed a deep convolutional neural network(CNN) which is able to predict eye fixations and segments salient objects in an integrated framework. Using number of datasets our CNN network provides better improvement in eye fixation prediction and salient object segmentation. This proposed network handles multi-scale aspects and it has branched architecture so it captures both the high-level and low-level Semantics required for salient object segmentation.

Ejaz Ahmed et.al[3] proposed a method for simultaneously learning features and similarity metric for person re-identification. To address the problem of re-identification a deep convolutional architecture is used. The proposed provides better performance eve using small data set (VIPeR). Large dataset (CUHK03) and a medium-sized data set (CUHK01) is also been used. cross-input neighbourhood differences layer, and a subsequent layer provides efficient usage of network.

Muhammad Imran Razzak et.al[4] discussed about deep learning architecture and its usage in Medical image segmentation and classification. Compare to other real world problem deep learning in health care especially in medical image processing is quite slow. One of the big obstacles is unattainability of annotated dataset. Bigger the data better the results. It deals with complex data and gives positive feedback.

Dennis Sng et.al[5] proposed the deep learning algorithms applied to video analytics of smart city for image classification, object detection, object tracking,

face recognition, and scene labelling. Deep learning increases the performance of human capabilities. Big training data are available for building large neural networks and it is used in wide range of applications.
Yingjie Xia et.al[6] proposed that a deep learning approach for the effective recognition of 3DObjects with blocking. Using encode–decode deep learning network this approach constructs multi-view shape model based on 3D. Converting deep learning to RGB-D data is more benefit for recuperating missing information. This approach provides accuracy and improves the efficiency of feature representation and increase the performance of object.

Yoshua Bengio et.al[7] proposed deep architecture and explains motivations and principles regarding the learning. This approach use unsupervised learning for single-layer models such as Restricted Boltzmann Machines and deeper models such as Deep Belief Networks. This paper motivates in learning AI approach, then representation and classification and theoretical architecture on computational architecture.

Zhangyang Wang et.al[8] proposed the context of deep learning to follow the sparse coding-based clustering Pipeline. Based on the sparse coding algorithm, TAGnet a feed-forward network structure is constructed trained with task specific loss functions. To intermediate feature hierarchy an auxiliary clustering task has been introduced, DTAGnet is used to boost sparse coding-based clustering and provides high effectiveness and robustness.

Bolei Zhou et.al[9] proposed scene-centric database called Places which is used to compare density and diversity of image for different data sets. To represent places and scenes in the real world Places database is intended. To establish and to learn deep features of scene recognition CNN been used and it provides better results. Visualization in CNN provides internal representation difference between Scene-centric networks and object-centric. This performance is measured using deep features on current scene benchmarks.

III. PROPOSED WORK

System description explains about the input and output design of the work. The accurate measurement of three-dimensional eye movements is desirable in many areas, such as in oculomotor and vestibular research, medical diagnostics, and photo-refractive surgery. The three main ways to measure three-dimensional eye movements are to use scleral search coils, electro-oculography, or video-oculography. Video-oculography is the only one of these options that is suited for clinical practice, since scleral search coils can be uncomfortable and electro-oculography has low spatial resolution.

Using video-oculography, horizontal and vertical eye movements tend to be easy to characterise, because they can be directly deduced from the position of the pupil. Torsional movements, which are rotational movements about the line of sight, are rather more difficult to measure; they cannot be directly deduced from the pupil, since the pupil is normally almost round and thus rotationally invariant. One effective way to measure torsion is to add artificial markers (physical markers, corneal tattoos, scleral markings, etc.) to the eye and then track these markers. However, the invasive nature of this approach tends to rule it out for many applications. Non-invasive methods instead attempt to measure the rotation of the iris by tracking the movement of visible iris structures.

To measure a torsional movement of the iris, the image of the iris is typically transformed into polar co-ordinates about the centre of the pupil; in this co-ordinate system, a rotation of the iris is visible as a simple translation of the polar image along the angle axis. Then, this translation is measured in one of three ways: visually, by using cross-correlation or template matching, or by tracking the movement of iris features.Methods based on visual inspection provide reliable estimates of the amount of torsion, but they are labour intensive and slow, especially when high accuracy is required. It can also be difficult to do visual matching when one of the pictures has an image of an eye in an eccentric gaze position.

If instead one uses a method based on cross-correlation or template matching, then the method will have difficulty coping with imperfect pupil tracking, eccentric gaze positions, changes in pupil size, and nonuniform lighting. There have been some attempts to deal with these difficulties but even after the corrections have been applied, there is no guarantee that accurate tracking can be maintained. Indeed, each of the corrections can bias the results.

The remaining approach, tracking features in the iris image, can also be problematic. Features can be marked manually, but this process is time intensive, operator dependent, and can be difficult when the image contrast is low. Alternatively, one can use small local features like edges and corners. However, such features can disappear or shift when the lighting and shadowing on the iris changes, for example, during an eye movement or a change in ambient lighting. This means that it is necessary to compensate for the lighting in the image before calculating the amount of movement of each local feature.

To validate the correctness of our methodology, we check whether it works on a video of a human eye by comparing its torsional estimates with estimates obtained through visual matching. Then, we check to see whether the three-dimensional eye positions from a standard nine-point calibration obey Listing's law.

IV. SYSTEM IMPLEMENTATION

The proposed algorithm is initialized on detecting a face from the USB camera feed, under satisfactory illumination. When the position of the user is sufficiently constant, the system for detecting and analyzing blinks and mouse movements is initialized automatically, depending on the involuntary blink of the user.

A local template of the open eye is used for the subsequent tracking of the eye. On performing the training procedure required for each session, scores based on variance projection as well as the relative positions of the iris are analyzed and interpreted to perform the various mouse functions accordingly Face and eye detection algorithm is used to identify and recognize the user input

1. Camera is activated to capture the live feed provided from the user.
2. From the captured input, Region Of Interest(ROI) is identified by checking face pattern.
3. Captured input is stored as image and the pixel point values are normalized to recognize the face input and eye input.
4. By using edge detection and shape detection, the eye pattern is recognized and the input is filtered from the image.

This analysis is performed at each frame. It is assumed that a frontal face is initially detected and tracked using an appropriate procedure. Once the face is detected and stabilized, it is necessary to locate the eyes in order to track the iris and analyze blinks.

An involuntary eye blink triggers the eye localization process.To accomplish this, the deference image of the head region of consecutive frames is created and then thresholded with a suitable threshold value. Some room for head movement should be accounted for without resulting in errors in mouse cursor movement. It is necessary to distinguish head movements from iris movements. Thus tracking of the head to a small extent is necessary.

The predefined pattern is commonly known to be as the default pattern. The default pattern technique is normally applied to all experimental purpose and this is to ensure that the absolute or the required data is obtained at the user end. Generally, the predefined pattern technique is carried out for comparative study, where before an execution of any task the data's, more precisely the current data are compared with the predefined pattern and are then subjected to implementation.

Gaussian and wiener filter is used to apply the thresholding by calculating the probabilistic distribution value for estimating the lower bound and upper bound to applying the noise removal process.

1. Each image is sperated in rgb color vector
2. In each color vector, row and columns are processed separately
3. For each row, from the list of pixel values, summation and average value is estimated.
4. From the computed average value, standard deviation and the variance value with the range of expected value is estimated
5. From he estimated std dev and mean value, probability of distribution is computed to apply the filtering operation of the image pixel
6. The upper bound and lower bound is estimated for the distribution of the pixel valued.
7. The boundaries are fixed as the threshold and the pixel piints ranges between the boundaries are filtered using the probability filters and the remaining pixel points are ignored.

The predefined pattern is commonly known to be as the default pattern. The default pattern technique is normally applied to all experimental purpose and this is to ensure that the absolute or the required data is obtained at the user end.Generally, the predefined pattern technique is carried out for comparative study, where before an execution of any task the data's, more precisely the current data are compared with the predefined pattern

The way of obtaining the predefined pattern can vary from different manufactures. In this page processing technique the predefined pattern is obtained by making use of the angle calculation. To be more detailed, the user who has been sensed by the camera is asked to focus at four different points from the page provided.

The pattern matching algorithm involves the following steps:

- ✓ The input video frame and the template are reduced in size to minimize the amount of computation required by the matching algorithm.
- ✓ Normalized cross correlation, in the frequency domain, is used to find a template in the video frame.
- ✓ The location of the pattern is determined by finding the maximum cross correlation value.

The points that are to be focused are namely, the center point, the top of the page, the bottom of the page, the left and right corner. The system is trained in such a way that , it can calculate the angle based on the obtained reference points ($\theta1,\theta2,\theta3,\theta4,\theta5$).

The range is varied from 1024*768 and 1344*780, which are normally considered as the horizontal and vertical axis. The angle that is made use of, in this technique is the inverse tangent. The inverse tangent or the tan' is taken to be as the reference parameter and is calculated as,

tan'(b/a) => tan' (y dist / x dist)(1)

The angle we make use of relies on degree calculation . Hence all the radian values are converted to degree before they are taken for the predefined pattern recognition. The angle that is calculated for the purpose of predefined pattern recognition varies from user to user. And this is because of the varying physical characteristics of every user. Depending upon the area over which the users place their line of sight, the predefined pattern calculation is varied.

Once the pattern is recorded for one single user, it remains the same throughout the page. And every time the user reads a line the current angle is compared with the predefined angle and then the line is read. The angles that are calculated are placed as an array matrix in the memory that are then used for comparative study. The lines in the page can be read with or without the use of the mouse pointer. The case in which there is a demand for the cursor is that when a random pattern is recorded. In this case the user is advised to make use of the mouse pointer. The same may be applied for closing a tab. In all other cases the use of mouse pointer is not a demand.

IRIS MOVEMENT DETECTION

The user input is captured by camera. The input image undergoes image enhancement technique to extract the iris movement. The enhancement technique such as sharpening and segmentation. Sharpening is done to neglect the background of the image this in turn gives the accurate pixel value of iris. Second technique is segmentation this is done for grouping similar pixels this in turn helps to split the iris from the input image.

Then the image is given to the motion sensor. Motion sensor is used to detect the movement of iris. It will detect only pixel point and its coordinate. Using this coordinates the movement of iris is calculated.

The pixel coordinates are generally in (x,y)form. By considering the previous and current y point position, whether the user is reading current or next or previous line can be determined. If the user is reading the same line then by comparing the current and previous x point position, whether the user is reading from right to left or from left to right can be determined. In general, human cannot view a line with exact straight line. Hence, we considering the absolute difference between the y points.

The absolute difference is less than difference coordinate than the user is reading the same line. If the absolute difference is greater than difference coordinate then scroll operation has to be performed. The obtained pattern has to be matched with pre defined pattern. Then the required action has to be performed. Then the task has to be executed. The output is obtained. After that the feedback has to be collected from the user. if the feedback is positive, the process is continued.If it is negative, the entire system has to be remodelled. If it is neutral, additional features has to be added to the system.

CCD sensor / Web Camera is used capture the live feed from the user. Supervised learning with back propagation is used to identify the input type and click operation on the particular action event.Pattern matching algorithm is used to compute the pattern related to the mouse movement

1. From the continuous images, the ROI is identified and filtered from the images
2. The pixel points of the ROI is compared in each images
3. The movement pattern is identified by checking the same pixel of the detected shapes.
4. From the movement of The pixel values, pattern of mouse movement is detected
5. Mouse control is selected and the move operation is executed.

The module under the test is seen as a single unit. The correctness of the module are evaluated by comparing the obtained output with the expected output.

V. RESULTS AND DISCUSSION

Detecting movement of the eye and calculating the eye detection delay using timestamp. In existing approach eye detection delay is higher compared to proposed approach.

Eye Detection Delay = Eye Detection Timestamp - Program Initiated Timestamp (2)

Click operation based on the eye tab using iris movement either single tab or double tab operation. Click operation delay is calculated by using click timestamp and eye tab timestamp.

Click Operation Delay = Click Timestamp - Eye Tab Timestamp (3)

Task execution delay is calculated by eye movement, click operation, double click, thrice eye tab timestamp.

Task Execution Delay = (Eye Movement Timstamp - Eye Movement Initiated Timestamp) + (Click Operation Delay) + (Double Click Timestamp - thrice Eye tab Timestamp) (4)

Accuracy is used to check the accuracy of the work. Accuracy can be calculated by successful task execution by number of attempts.

Accuracy = Successful Task Execution / Number of Attempts (5)

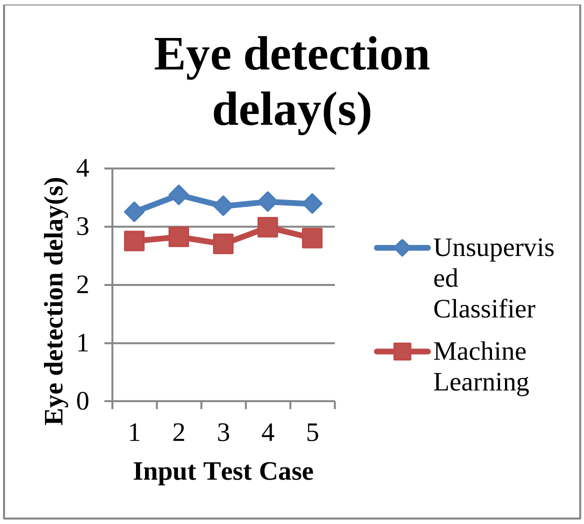

Figure 3. Eyedetection Delay

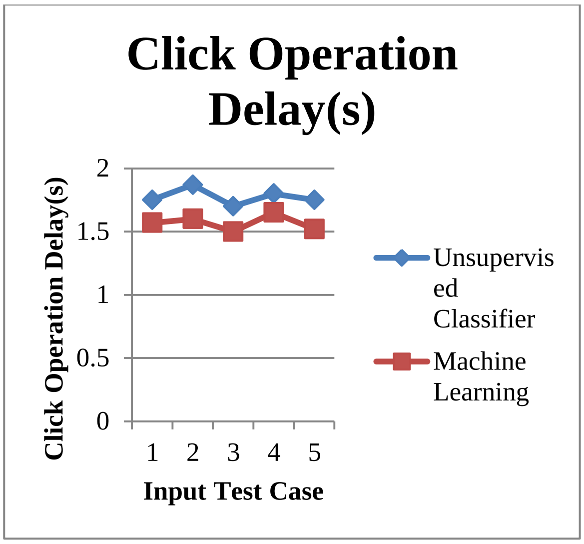

Figure 4. Click Operation Delay

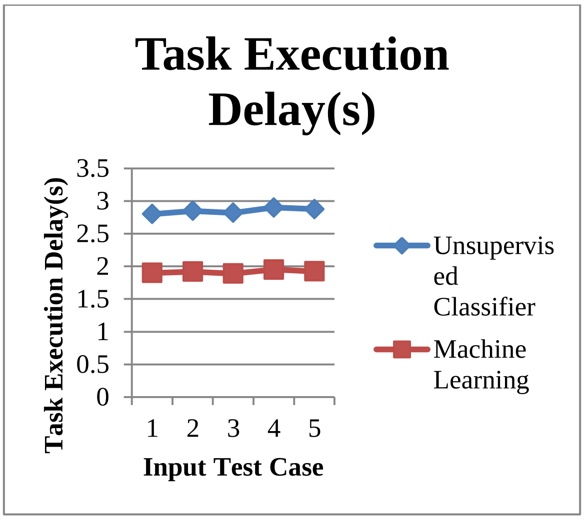

Figure 5. Task Execution Delay

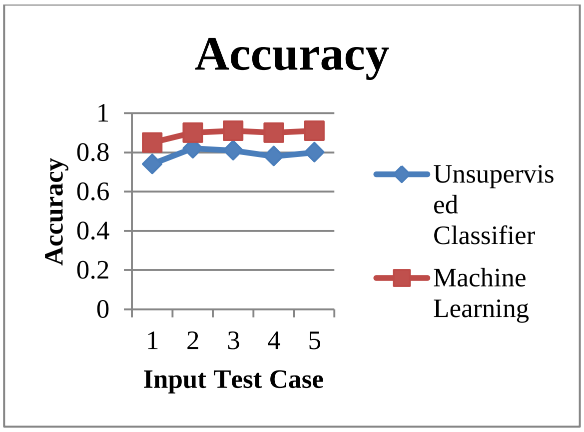

Figure 6. Accuracy

VI. CONCLUSION

The Artificial Neural Network is primarily employed to perform the learning process to improve the accuracy of the pattern recognition and matching process. Using the ability of automatic learning and improve the experience is evolved to create the required patterns with the machine learning. The machine learning algorithm is mainly employed to produce the perceptive and functional model to handle the disability management. Providing the ease of access to operate the computer with the help of iris movement is developed in the proposed work by applying the digital signal processing system. The prediction of task to be performed is based on the input operation, Task execution and output production with the feedback collection are used to improve the ANN. The performance evalution showed that the proposed ANN is achieved better performance compare to the existing CNN in terms of eye detection delay, task identification delay and task execution delay. The eye fixation can be furtherly enhanced by employing the secure authentication based on the iris biometric identification that uses mathematical pattern-recognition techniques on video images of one or both of the irises of an individual's eyes, whose complex patterns are unique, stable, and can be seen from some distance. The iris biometric and its movement in various direction can be employed as credential verification system instead of pattern, pin and password to access the devices such as mobile, laptop, computer and PDA. This requires the speed of matching and its extreme resistance to false matches

VII. REFERENCES

[1]. Srinivas S S Kruthiventi, Kumar Ayush, and R. Venkatesh Babu "DeepFix: A Fully Convolutional Neural Network for predicting Human Eye Fixations" DOI 10.1109/TIP.2017.2710620, 2017

[2]. J. Pan, K. McGuinness, E. Sayrol, N. O'Connor, and X. Giro-i Nieto, "Shallow and deep convolutional networks for saliency prediction," arXiv preprint arXiv:1603.00845, 2016

[3]. C. Sheth and R. Venkatesh Babu, "Object saliency using a background prior," in IEEE International Conference on Acoustics, Speech and Signal Processing, 2016.

[4]. S. S. Kruthiventi, V. Gudisa, J. H. Dholakiya, and R. Venkatesh Babu, "Saliency unified: A deep architecture for simultaneous eye fixation prediction and salient object segmentation," in IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[5]. H.HadizadehandI.V.Bajic,"Saliency-awarevideocompression,"IEEE Transactions on Image Processing, vol. 23, no. 1, pp. 19–33, 2014.

[6]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems, 2012.

[7]. L. Bottou, "Large-scale machine learning with stochastic gradient descent," in Proceedings of COMPSTAT'2010, 2010, pp. 177–186.

[8]. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, pp. 1–42, April 2015.

[9]. B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva, "Learning deep features for scene recognition using places database," in Advances in Neural Information Processing Systems, 2014.

[10]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[11]. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," arXiv preprint arXiv:1409.4842, 2014.

[12]. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[13]. T. Judd, F. Durand, and A. Torralba, "A benchmark of computational models of saliency to predict human fixations," in MIT Technical Report, 2012.

[14]. A. Borji and L. Itti, "Cat2000: A large scale fixation dataset for boosting saliency research," CVPR workshop on "Future of Datasets", 2015, arXiv preprint arXiv:1505.03581.

[15]. S. Hossein Khatoonabadi, N. Vasconcelos, I. V. Bajic, and Y. Shan, "How many bits does it take for a stimulus to be salient?" in IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[16]. M. Kummerer, L. Theis, and M. Bethge, "Deep gaze i: Boosting saliency prediction with feature maps trained on imagenet," arXiv preprint arXiv:1411.1045, 2014.

[17]. C. Shen and Q. Zhao, "Learning to predict eye fixations for semantic contents using multi-layer sparse network," Neurocomputing, vol. 138, pp. 61–68, 2014.

[18]. E. Vig, M. Dorr, and D. Cox, "Large-scale optimization of hierarchical features for saliency prediction in natural images," in IEEE Conference on Computer Vision and Pattern Recognition, 2014.

[19]. N. Liu, J. Han, D. Zhang, S. Wen, and T. Liu, "Predicting eye fixations using convolutional neural networks," in IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[20]. A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: an astounding baseline for recognition," in Computer Vision and Pattern Recognition Workshops, 2014.

[21]. X. Huang, C. Shen, X. Boix, and Q. Zhao, "SALICON: Reducing the semantic gap in saliency prediction by adapting deep neural networks," in IEEE International Conference on Computer Vision, 2015.

[22]. R. Zhao, W. Ouyang, H. Li, and X. Wang, "Saliency detection by multicontext deep learning," in IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[23]. S. Srivatsa R and R. Venkatesh Babu, "Salient object detection via objectness measure," in IEEE International Conference on Image Processing, 2015.

[24]. Deeplearning.net," http://deeplearning.net/tutorial/.

[25]. Z. Bylinskii, P. Isola, C. Bainbridge, A. Torralba, and A. Oliva, "Intrinsic and extrinsic effects on image memorability," Vision research, vol. 116, pp. 165–178, 2015.

[26]. Ms. Sonali. B. Maind, Ms. Priyanka Wankar," Research Paper on Basic of Artificial Neural Network"International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Vol: 2 Issue: 1.