Study Of Deep Learning Techniques For Differently Abled Applications And Detection Techniques

Anandh N[1], Prabu S[2]

[1]Research Scholar, School of computer science and engineering, Vellore institute of technology
[2]Professor , School of computer science and engineering, Vellore institute of technology,
anandh.n@vit.ac.in, sprabu@vit.ac.in
corresponding author: PRABU S,sprabu@vit.ac.in

Abstract:
In worldwide, Visually Impaired Persons (VIP) are facing several issue related to visual impairment and blindness and they are assisted with technical inventions. Based on the survey from World Health Organization (WHO), around 2.2 billion peoples are suffered from visual impairment and among them 1 million peoples are suffered by blindness. Vision is the major sensing organ of the human and to assist the VIP regarding this, there are various digital vision products are in the market which is based on digital technologies and advanced algorithms. This will transform the VIP's vision world into audio to get to know about their surroundings includes objects, motion, obstacles and spatial locations. The objective of this paper is to provide the detailed survey about the existing object recognition, face recognition and text to voice recognition methods proposed to assist VIP. Due to the increase development of machine learning and deep learning algorithms, the digital image recognition and object recognition are become more efficient. These advanced technologies can make sure to assist the VIP to detect and recognize the people face and objects in front of them in the form of audio that are practiced by them on daily basis.
Keywords: Visually Impaired Person, Blindness, deep learning, image recognition, object recognition, text to audio.

1. INTRODUCTION

The organ which is mainly used to observe and learn about the surrounding is human eye. Our daily activity such as reading, writing, moving, observing and much more etc., are based on the vision through eye. Since, vision is the more powerful sense than other, it plays important role in human perceptions regarding the surrounding environment. The visual impairment defined by WHOis the reduction of vision that cannot be cured by contact lenses / glasses which will reduce the person's ability to work on a task [1] and blindness is the severe sight loss that a person cannot be able to see their fingers. Vision impairment can be classified into two types such as distance and near impairment of vision by International Classification of Disease 11. In near VI, after the correction also the vision is as low as M.08 or N6. Distance VI is classified as mild, moderate, severe and blindness based in the Vision worse than 6/12, 6/18, 6/60 and 3/60 correspondingly [2]. The main causes of Vision impairment are cataract, glaucoma, corneal opacity, eye injuries, uncorrected refractive errors, age related macular regeneration, diabetic retinopathy and trachoma [3]. This will affect the mobility and ability to contact with the surroundings of the VIP.

The technologies such as white cans, guide dogs, magnifiers, glasses and screen reading software are used by the VIP for mobility assistance. The white cane length is equally proportional to the range of obstacles detection. Guide dogs are used as a walking assistance to make aware of obstacles in the stepping up and down of the VIP. It cannot give direction assistance to the VIP at complex cases. GPS assisted devices can help the VIP to assist regarding navigation and movement from specific position. These devices can help the VIP regarding their location movement but it cannot avoid the obstacle detection and avoidance. To detect the silent object in front of the VIP, a new technology called echolocation [4] can be used where the sound echoes are recognized with simple mouth clicks. Braille language is used to read document by the VIP or blind person with the knowledge about it. The issue in this is the characters used in braille language are not practical to install everywhere to convey the

information. The VIP can use tactile marks to identify the currency notes and to access the computer and mobile systems, they can use refreshable braille displays, screen readers and screen magnifiers.

To assist the VIP, real time object recognition [5] with two cameras located in person glasses. This also consists of GPS service with ultrasonic sensors to provide the assistance about surrounding objects to VIP. Object detection technique was used here to detect the objects such as face, chairs, tables, doors and other obstacles. These objects are grouped and identified through GPS and sensor located at medium and long distance. To optimize the performance of the recognition, SURF (Speeded-Up Robust features) method was used. ETA (Electronic Travel Aids) was used by the VIP for the detection of obstacles for safe navigation [6]. They provided with the robotic cane for walking assistance which have omnidirectional wheel with LAM base linearization system. This will reduce the risk of falling of a person and maintain the balance.

Received Strength Signal Detection (RFID) was designed to support the blind person to detect the objects [7]. This is specifically designed for medicine searching in the home cabinet. This have the information about the object distance in the form of acoustic signal to identify the medicine. EMC (Electronic Mobility Cane) was designed to vision restoration to detect obstacles [8] which construct a logical map for this assistance that are conveyed to the VIP in the form of audio, voice or vibration. The live of VIP can be enhanced with these science and engineering technologies to provide the independent live to navigate and detect the objects surrounded by them. The devices proposed for this assistance are based on object detection through computer vision with sensors, GPS and distance sensors etc., the main contribution of this paper is to provide an detailed review about the intelligent assistant techniques for VIP with obstacles aware mobility using deep learning techniques.

The remaining section of this survey is organized as follows. Section II provides the detailed review about the Face Detection techniques based on deep learning algorithms.Section III reviewed about the Object detection methods based on deep learning. Section IV concludes the review with challenges and suggestions.

2.   Review about the Face detection approaches using ML/ deep learning algorithms

Biometrics is a technology use to analyze the biological data statistically based on the behavioral and physiological characteristics of a person for identification. This can be used in many applications such as forensics, prison security etc., to provide the secure access. The biological features such as hand geometry, retina, iris, fingerprint, palm print, face are used for this identification. Face is the important part of the body which can speak through the emotions. This is the main source of organ to interact with the society. Due to the reliable technologies, Face Detection and Recognition are the trend to increasing research. Compare to other biometric traits, FD provides numerous benefits. The technique behind FD is to recognize the face of a human based on the features of the face and compare it with the previously recorded inputs. Approximately there are 80 to 90 nodal points in the face that are unique. The distance between these nodal points to eyes, jawline, and cheekbones shape are considered to be the aspects in FD system. These are created as the facepirnt and used to identify the face from the database.

The FD system are main research area in machine learning (ML) and biometric [9]. Machine learning is the technology that a computer can learn by itself based on the training through the algorithms to perform the functions. It consists of various methods includes reinforcement learning, supervised and unsupervised learning [10]. These Ml algorithms can consume lot of storage and resources. Due to this reason, the ML algorithms can be combined with cloud computing, fog computing and enhanced with deep learning algorithms for better performance.

Due to the development of machine learning and recognition systems, the researches shows interest on pattern recognition with various mining models. The recent research related to face recognition systems are discussed in this section.Salama et al.,[11] proposed a Facial recognition system based on deep computational intelligent approaches. They developed FR systems using cloud and fog computing with deep learning approach called deep convolutional neural networks (DCNN). DCNN used here to extract the features from the input face image in order to compare it with the database. The FR system has

been tested with the ML algorithms called Decision tree, Support vector machine and KNN. They used three datasets for evaluation. the datasets are SDUMLAHMT, 113 and CASIA. Based on the evaluation metrics such as accuracy, precision, sensitivity and specificity and time, the proposed deep face recognition obtains better performance than other algorithms with the accuracy of 99.06%, recall of 99.07% and specificity of 99.10%.

Schiller et al., [12] proposed a relevance based data masking for facial expression recognition to learn automatic emotion recognition. The proposed method can overcome the issues of the traditional FR systems which includes inheritance of the original model structure and restriction to the neural network structure. Based on the relevant features of the input data, the proposed model is trained using the

CNN and validated with the benchmark datasets such as AffectNet [13], FER Dataset [14] and Cohn-Kanade dataset [15]for the detection of facial expressions. They used the CNN architectures such as Xception [16], MobileNetV2 [17], Inception V3 [18] and VGC- Face [19] for training the proposed model. Transfer learning based CNN model has been proposed for face recognition by Prakash et al. [20]. To train the proposed model, the weight of VGG 16 model has been used. The features extracted from the CNN was given as input to the Fully connected layer with the activation function called Softmax for classification. They used the ImageNet database for classification. the evaluation of the proposed model with Yale and AT&T gives better recognition with the accuracy of 96.5%.

Jonnathann et al [21] proposed a comparative study between the profound learning and Artificial intelligence techniques such as ANN, SVM, KNN and deep learning. They used CNN for facial recognition and the datasets such as AR face, Yale, and SDUMLA-HMT [22] have been used for analysis. Ding and Tao [23] proposed a video based face recognition using ensemble CNN. The standard CNN has been improved with truck branch learning. It filter the data from the face images and picked the facial segments. The convolutional layers between the branch and trunk systems are accessing the information shared by the proposed model. They used three video face datasets such as Youtube, PaSC faces and COX face. Shepley [24] analyzed and compared the deep learning based face recognition with the limitations. This paper reviewed and discussed the challenges and research areas for future with novel approaches. They clearly tabled the face detection and recognition datasets that are publically available.

Current research on Face detection and recognition use DCNN for significant performance [25]. Face detection with DCNN can be classified as region based or sliding window based approaches. The region based approaches use selective search [26] to create the collection of regions that includes face. The current methods includes HyperFace [27] and All in once face [28] used this method. The Region based CNN called R-CNN [28] used here to execute the bounding box classification. Sun et al. [29] enhanced the R-CNN with feature concatenation, hard negative mining, multi scale training which reduce the false positive rates and improves the accuracy. Wang et al., [30] proposed a detailed review about the Deep learning based face recognition algorithms, databases, applications and protocols. The loss functions, architectures of Deep face recognition are discussed. Face processing methods are categorized into one to many augmentation and many to one normalization. They reviewed about cross factor, multiple media, heterogeneous and industrial scenes with the challenges and future directions.

Farfade et al [31] addressed the multi view face detection issue with minimally complex deep dense face detector (DFD) approach. This method does not need landmark annotation and the limitation was inadequate to sampling and augmentation. Yang et al., [32] proposed a strategy called Faceness to enhance the face detection where the face images consist of 50% of occlusion. It also accepts the images with different poses and scales through the attributes involved in deep networks. This network was trained with generic objects and part level binary attributes. Viola and Jones [33] proposed a machine learning based approaches for object detection. They used AdaBoost based learning algorithms which is used to select the number of features. This work used cascaded classifiers to remove the background regions and considered only the object oriented regions. This proposed approach obtains high detection rate of 15 frames per second than other algorithms.

Zhang et al.,[34] proposed a multi-task cascaded CNN for face detection and alignment. It combined three stages of DCNN to detect the face and location in the coarse to fine strategy. This will

improve the automatic detection of face than manual sample selection. They analyzed with the datasets such as FDDB, WIDER FACE and AFLW and obtains the improved accuracy of real time performance. Sharma et al., [35] proposed a face detection model with Deep belief Network (DBN) based deep learning algorithm to recognize the face from videos. This model can able to detect the blurred images and also side posed images. The drawback of this method is it cannot detect the eyes with glasses. Mehta et al [36] enhanced the multiviewface detection (FD) with CNN proposed by farfade et al.,[31] enhanced with tagging system. They used Deep Dense Face Detector with DCNN and the detected faces were recognized with LBPH (Local Binary Patterns Histograms). They follow their enhancement process as the following steps, the input image are preprocessed, heat maps are generated to extract the faces, probability of the faces are found and then the image is pass on to tagging algorithm for FD. They evaluated the proposed work with the dataset such as MIT Manipal Farewell 2017, Custom Different Oriented faces, Randomly Selected Celebrities. The evaluation metrics such as Precision, Recall and F measure have been used as the parameters for the algorithm and they obtained 85% accuracy for FD.

The dual layer CNN have been proposed for FD by Osadchy et al., [37]. In this work, the first layer was used to locate the position of the faces and second layer was used to detect the face with the help of the localization. The issues such as rotation and poses are handled by Vaillant et al [37] with the combination of FD and pose estimation. Thus enhancement improved the dual layer CNN with better face detection and estimation. Chi et al.,[38] proposed FD based on End to End CNN trained network. Instead of the prediction of facial landmarks, the geometric transformation matrix was used to align the face of the system. They evaluated with the system with the datasets WIDER FACE [41], CASIA-WebFace [42], FDDB (Face Detection and Benchmark) [43] and LFW (Labelled Face in the Wild) [44] and obtained 89.24% of recall with the accuracy of 98.63%. Fontaine et al.,[39] proposed a computationally efficient real world conditions trained with few samples for training for Face recognition. They used sparse representation for alignment. Zhang and Chi [40] proposed an end to end CNN with Spatial transformation for face recognition. This work is the enhancement of the work [38] with spatial transformation.

Ding et al [45] proposed deep learning based face detection and assessment with swapped nature. The pairwise comparison of the human subjects are collected in a website. The evaluated results of this work obtains 96% of true positive rate. They created the dataset with swapped still images with the objective of creating the datasets for forensics. Pandey and Sharma [46] reviewed about the face detection and recognition techniques based on neural network techniques. they analyzed the methods used to identify the face from the facial features that are identified, extracted and compared with the template. The facial features were extracted using PCA, LDA, MPCALDA and trained with neural network model called back propagation algorithm.

Ashukumaret al., [47] surveyed about the techniques used for FD. They done a comprehensive review about the challenges, applications of face detection system, feature based and image based FD approaches, Statistical based FD system and the databases used for FD and concluded with the future directions for research on FD. Arunkumar&jain [48] reviewed and compared the FD techniques that are applicable for face detection from images and videos. They also listed the Face recognition models that are frequently used and the databases that are publically available for face recognition.

The Face detection and recognition system using the Machine learning based algorithms are discussed here,Le [49] proposed a FD system based on hybrid model of Artificial neural network with AdaBoost called ABANN. The labelled faces are aligned using active shape and multi layer perceptron. Based on the facial expressions and contours the classifiers can improve accuracy of the face detection. To extract the facial features they used geometric and Independent Component analysis methods. The database used here for evaluation was MIT+CMU [50]. Sharma et al., [51] proposed a FD system with machine learning algorithm and Principal Component analysis. They also evaluated LDA, Multilayer perceptron, Naïve Bayes and SVM and obtained the accuracy of 97% and while using PCA and LDA the accuracy was 100%.

Lahaw et al., [52] used LDA, IDA,PCA and SVM for face recognition. They experiment the FD system with these ML algorithms using AT&T database. The evaluated results of this experiment obtained 96% of accuracy using the hybrid approach of Discrete Wavelet Transform (DWT) and PCA/ICA to reduce the dimension and used SVM as a classifier. Sabri et al., [53] analyzed and compared the ML algorithms such as MLP, NB and SVM for classifying the face using face geometry. They evaluated and concluded that the NB performs better than other algorithms with high precision of 93.16%. Fan et al., [54] proposed a manifold graph based FD called Enhanced Adaptive Locality Preserving Projections (EALPP). This method combines two method such as maximum margin criterion and locality preserving projections (LPP). They used the data sets such as YALE, ORL, UMIST and AR.

Hizem et al., [106][110][111] proposed a camera systems for face recognition with the enhancement of CMOS imaging system [107][112] and synchronized flash camera [108][113]. This paper provides a simplification of software and hardware systems for biometric applications with mobile platform. This will save the computation power and memory. They evaluated the systems with three types of cameras such as DiffCam, FlashCam and normal CCD cam with different illumination. The images captured from the camera was preprocessed using the correlation based method to get a better results. Zhou et al., [109][114] proposed a multi spectrum sensing for face detection with low lighting surrounding. They captured the face with the hybridization of infrared camera and thermal camera. To capture the red eye, they used near Infrared with the LED IR flashes. This bright eye use to localize the eyes and face in 3D position. With the conclusion of this section based on the literature up to 2020, it has been found that the majority of the Face detection and recognition research used the ML algorithm as PCA+SVM and NB gives better performance and deep learning algorithm called CNN and its variants for better detection of face from the images and videos.

3. Review about the Object detection methods using ML/DL algorithms.

Object detection is a crucial part of computer vision for many applications. For our review, the VIP can be able to recognize the objects such as chair, stone, table etc, in front of them through the object detection algorithms. The conventional object detection approaches may use Machine learning methods for recognition with the parameters of object features based trained neural network followed by the classification methods. Masita et al., [55] conducted a review about the recent advancements of deep learning based Object detection (OD) methods. Object detection based on ML algorithms follow the statistical and mathematical equations that are feature based. Neural network are trained and learned based on these features. Liu et al.,[70] provides the detailed review about 300 research papers in the area of object detection using deep learning methods. This paper covered the articles related to object detection, object feature representation, object detection frameworks, context modeling, evaluation metric and training stages with future enhancement of the research.

Zhang et al.[71] discussed about the object detection techniques used before the 2011. Li et al. [72] discussed about the deep learning and handcrafted based feature representation methods for object detection based on statistical learning. Borji et al. [73] surveyed about the salient object detection models. Bengio et al. [74] discussed about the deep learning, unsupervised feature learning, auto encoders, deep networks, manifold learning based methods used for object detection. Litjens et al. [75] reviewed about the research article related to object detection, image classification and segmentation of medical images. Gu et al. [76] provide a detailed comprehensive survey about the Convolutional Neural Network and its applications on computer vision in terms of natural language processing and speech. The machine learning and Deep learning algorithms used for object detection are listed in Table 1.

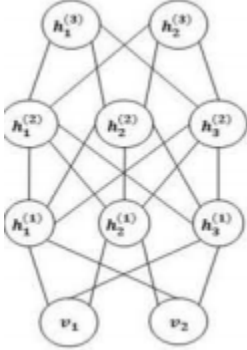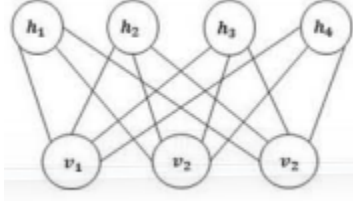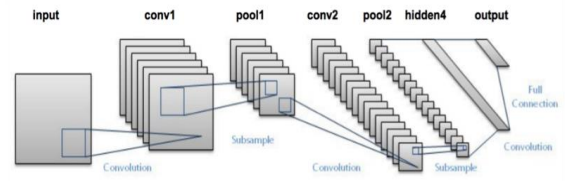Table 1: ML algorithms for Object Detection

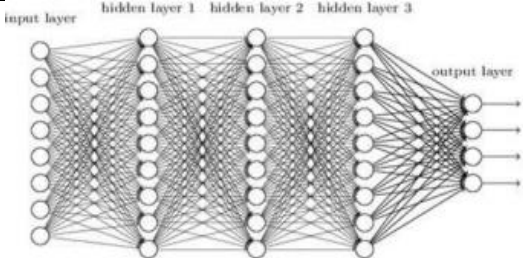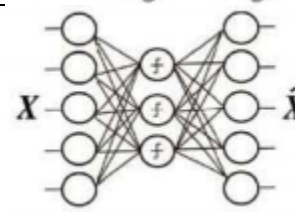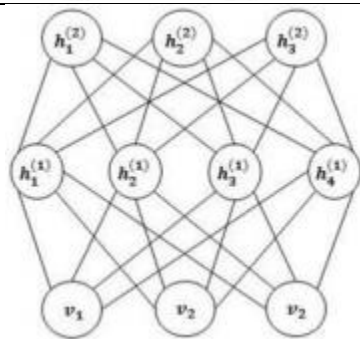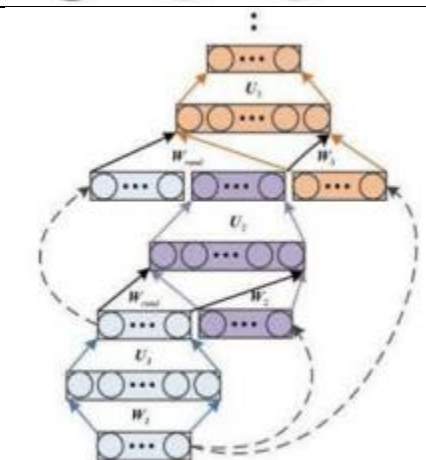| Method	Description
Deformable Models [57]	This is statistical model based on the object actual instance with template deformation. These template information is used to denominate the bonds.
Sparse linear regression [58]	This model used to approximate the regression and noise repeatedly. It used mean residual square for this noise reduction.

Bayesian linear regression [59]	It overcome the regular distribution issue in linear regression through Bayesian inference. In Bayesian inference, the noise variance and prior variance is automatically inferred.
Bayesian logistic regression [60]	This is a binary classification model using the dot product of weight and feature vector of the hyperplane.
Support Vector Machines [61]	Statistical model developed to limit the generalization errors. It is the pattern recognition algorithm developed by Vapnik [62]. This classifies the data point and due to the speed and memory this is highly expensive.

Asim et al [83] discussed about the CNN and its variants in terms of salient, objectness and category specific related methods used to detect the object. Zhao et al.,[84] reviewed about the deep learning based object detection methods. They discussed about the salient object detection, pedestrian detection and face detection. Tong et al.,[85] reviewed about the previous object detection methods with deep learning algorithms which includes data augmentation, context based detection, multi scale feature learning and GAN based detection.

Object detection based on ML algorithms have some drawbacks while the methods are extended to identify the complex objects such as vehicles, peoples and etc. it needs high prior knowledge and related information about it to train the model. Next, the structural similarities are need to be identified based on the image representation. So the knowledge about the object image representation with that features are need to be trained. To overcome these issues, ML algorithms are enhanced with the optimization techniques and deep learning that are computationally intensive. Table 2 shown the List of Deep Learning (DL) algorithms used for object detection.

Table 2: Deep Learning algorithms for Object Detection

Methods	Description	Representation
Deep Boltzmann Machine (DBM) [62]	Feature based learning model which consist of multiple layers with hidden variables. The variables in the layer are considered to be private. The high level features with the sensory inputs were used for this object detection	
Restricted Boltzmann Machine (RBM) [63]	DBM enhanced by limiting the hidden layers count into one layer in RBM for object detection. The activation of one RBM is used to train the next RBM.	
Convolutional Neural Networks (CNN) [65] [66]	It looks like gird structure consist of multiple layers of convolution, pooling input and output layers and activated through the activation function.	

Deep Neural Networks [67]	Neural network consist of multiple hidden layers in DNN. The choice of including the neurons in the layer is complex one. And DNN can fit the data with fewer parameters. This make the DNN as an efficient one.	
Stack Auto-encoder (SAE) [68]	This is the two layer model and trained by itself to remodel the inputs to reduce the reconstruction error.	
Deep Belief Networks (DBN) [69]	This consist of first network as DNN and followed by the stacked RBM which consist of multiple hidden layers and Back propagation algorithm has been used to train it.	
Deep Stacking Neural Networks (DSNs) [70] [71]	Designed to reduce the computational errors during the training of DNN. Classifiers stacking one after another for solving the complex functions.	

Object Detection Dataset:
The most relevant databases used for object detection are PASCAL VOC (2012) [77] consists of 11540 images, ImageNet [78] consist of 14 millions+ images, MS COCO [79] consist of 328000+ images, Places [80] consists of 10 millions+ images and Open images [81] consists of 9 millions+ images, KITTI [86] consists of 15000 images, Caltech [87] consists of 192000 images, FlickrLogos [88] consists of 2500 images,SUN [89]consists of 132000 images, TT100K [90]consists of 100000 images, SOD [91] consists of 4925 images. The sample images from these datasets are shown in Table 3.

Table 3: Object detection dataset name with sample images

Data set name	Sample images

PASCAL VOC	
ILSVRC	
MS COCO	
Open Images	

Tang et al.,[82] proposed an object detection model based on weakly supervised part based learning method based on the size and location. It is based on objectness approach to estimate the salient region. Target objects are classified using the multi-region latent approach. they used PASCAL VOC 2007 AND MS COCO 2014 dataset for evaluation and obtained the result of 4.3 points better than standard model for object detection.

Paper ref	Object Detection Model	Description	Dataset used
Tan et al., [92]	EfficientDet	Initially they developed BiFPN (weighted bi-directional feature pyramid network", then they proposed compound scaling method to scale the resolution, feature backbone.	COCO
He et al., [93]	Mask RCNN	Efficient Object Detection with segmentation mask. This is the extension of faster RCNN. Training this method is simple.	COCO 2016
Girshick et al., [94]	Region-based CNN (R-CNN)	Bounding box object detection with the evolution of CNN that can run independently with each ROI.	PASCAL VOC

Ren et al., [95]	Faster RCNN	Enhancement of Region Proposal Network (RPN). This is flexible and robust.	PASCAL VOC 2007, 2012, and MS COCO,I LSVRC
Ghaisi et al., [96]	NAS-FPN (Neural Architecture Search- feature pyramid network)	NAS is enhanced with the novel search space that are scalable and cover all cross scale connections. It consists of top down and bottom up connections.	COCO
Liu et al., [97]	PANET (Path Aggregation Network)	PANet has been used for instance segmentation with the enhancement of localization signals and bottom up path augmentation. They enhanced the feature level detection with feature pooling concept	COCO 2017
Dai et al., [98]	R-FCN (Region based Fully Convolutional Network)	To address the issues in translation variance of image classification and object detection, position sensitive score maps has been proposed.	PASCAL VOC
Liu et al., [99]	SSD: Single Shot MultiBox Detector	Output bounding box has been divided into default boxes based on aspect ratio and scales per location of the feature map. It reduces the frequent feature resampling stages.	PASCAL VOC, COCO, and ILSVRC
Redmon et al., [100]	YOLO (You Only Look Once)	This separates the bounding boxes with the respective class labels. Single neural network has been use to predict the class labels and bounding boxes from the full images.	PASCAL VOC 2007

The recent research works that helps to the VIP to recognize the human face, objects in front of them using the deep learning models. Jayashree and kalpana [105] reviewed about the techniques involved in face recognition, text detection and object detention for VIP. Mareeswari et al., [101] proposed a series of procedure to detect the objects in front of the VIP through the camera attached to the eye glasses. They used the machine learning algorithms such as DWT (Discrete Warier Transform), SURF (Speeded-up Robust Features) and OCR (Optical Characteristics Recognition). Joshi et al., [102] proposed a Multi object detection method using Artificial intelligence for VIP. It was fully automotive to assist the VIP about their surrounding environment. They discussed about the deep learning based object recognition and distance measuring sensor. The evaluation results obtained from this work was 99.69% of object detection. Yao et al., [103] proposed a smart wearable image recognition for VIP using cloud and local processing. Cloud server performs the image recognition and image upload and information transfer are processing by the local processor. They used PASCAL VOC and LFW datasets for evaluation and obtained the high level of accuracy. Shaik et al., [104] proposed an object recognition system for VIP using the YOLO v3 machine learning algorithm. the evaluation database was COCO. The experimental results of this work obtained 95% accuracy as a overall performance and 100% accuracy on detecting the objects such as chair, clock an cell phone.

4. CONCLUSION

Vision are the important aspect of human to survive to access the information surrounded by the person. For the peoples those who are poor in vision can be able to interact with the surroundings through

the advanced technologies such as computer vision services and products. This paper discusses the comprehensive review about the face recognition, object detection techniques using the deep learning algorithms that can improve the exiting recognition methods. This review based on the techniques and datasets can help the researchers to direct the further improvement of the recognition of surrounding environment of the VIP.

References:

1. World Health Organization, www.who.com .
2. Bourne, R.R.A.; Flaxman, S.R.; Braithwaite, T.; Cicinelli, M.V.; Das, A.; Jonas, J.B.; Keeffe, J.; Kempen, J.H.; Leasher, J.; Limburg, H.; et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis. Lancet Glob. Health 2017, 5, e888–e897
3. Blindness and Vision Impairment. Available online: https://www.who.int/news-room/fact-sheets/detail/ blindness-andvisual-impairment (accessed on 12 June 2020).
4. Thaler, L.; Arnott, S.R.; Goodale, M.A. Neural Correlates of Natural Human Echolocation in Early and Late Blind Echolocation Experts. PLoS ONE 2011, 6, e20162
5. Zraqou, Jamal &Alkhadour, Wissam& Siam, Mohammad. (2017). "Real-Time Objects Recognition Approach for Assisting Blind People." Multimedia Systems Department, Electrical Engineering Department, Isra University, Amman-Jordan Accepted 30 Jan 2017, Available online 31 Jan 2017, Vol.7, No.1
6. Van Lam, P.; Fujimoto, Y.; Van Phi, L. A Robotic Cane for Balance Maintenance Assistance. IEEE Trans. Ind. Inform. 2019, 15, 3998–4009
7. A. Dionisi, E. Sardini and M. Serpelloni, "Wearable object detection system for the blind," 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Graz, 2012, pp. 1255-1258, doi: 10.1109/I2MTC.2012.6229180
8. Bhatlawande, S.; Mahadevappa, M.; Mukherjee, J.; Biswas, M.; Das, D.; Gupta, S. Design, Development, and Clinical Evaluation of the Electronic Mobility Cane for Vision Rehabilitation. IEEE Trans. Neural Syst. Rehabil. Eng. 2014, 22, 1148–1159
9. Sareen P. Biometrics—introduction, characteristics, basic technique, its types and various performance measures. Int J Emerg Res Manag Technol. 2014; 3: 109–119
10. Haffner P. What is machine learning–and why is it important. Interactions. 2016; 7
11. SalamaAbdELminaam D, Almansori AM, Taha M, Badr E (2020) A deep facial recognition system using computational intelligent algorithms. PLoS ONE 15(12): e0242269. https://doi.org/ 10.1371/journal.pone.0242269.
12. Schiller Dominik, Huber Tobias, Dietz Michael, and Andre´ Elisabeth. "Relevance-based data masking: a model-agnostic transfer learning approach for facial expression recognition." Front. Comput. Sci., 10 March 2020 | https://doi.org/10.3389/fcomp.2020.00006.
13. Mollahosseini, A., Hasani, B., and Mahoor, M. H. (2017). Affectnet: a database for facial expression, valence, and arousal computing in the wild. *IEEE Trans. Affect. Comput.* 10, 18–31. doi: 10.1109/TAFFC.2017.2740923
14. Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., et al. (2013). "Challenges in representation learning: a report on three machine learning contests," in *International Conference on Neural Information Processing* (Daegu: Springer), 117–124
15. Kanade, T., Cohn, J. F., and Tian, Y. (2000). "Comprehensive database for facial expression analysis," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)* (Grenoble: IEEE), 46–53
16. Chollet, F. (2017). "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Honolulu, HI), 1251–1258.

17. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2018). "Mobilenetv2: inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Salt Lake City, UT), 4510–4520.

18. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV), 2818–2826

19. Parkhi, O. M., Vedaldi, A., Zisserman, A., et al. (2015). "Deep face recognition," in *bmvc*, Vol. 1 (Swansea), 6.

20. Prakash, R. Meena, N. Thenmoezhi, and M. Gayathri. "Face Recognition with Convolutional Neural Network and Transfer Learning." In 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 861–864. IEEE, 2019.

21. Power Jonathan D., Plitt Mark, Gotts Stephen J., KunduPrantik, Voon Valerie, Bandettini Peter A., and Martin Alex. "Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data." Proceedings of the National Academy of Sciences 115, no. 9 (2018): E2105–E2114. https://doi.org/10.1073/pnas.1720985115 PMID: 29440410

22. Yin Y, Liu L, Sun X, SDUMLA-HMT: A multimodal biometric database. In: Chinese conference on biometric recognition. Beijing, China: Springer; 2011. pp. 260–268.

23. Ding C, Tao D. Trunk-branch ensemble convolutional neural networks for video-based face recognition. IEEE Trans Pattern Anal Mach Intell. 2017; 40: 1002–1014. https://doi.org/10.1109/TPAMI.2017. 2700390 PMID: 2847504

24. ANDREW JASON SHEPLEY, "Deep Learning For Face Recognition: A Critical Analysis",Cornell University, **Computer Vision and Pattern Recognition,** arXiv:1907.12739**, 2019.**

25. Krizhevsky, A., I. Sutskever, and G. Hinton, ImageNet classification with deep convolutional neural networks. Communications of the ACM, 2017. 60(6): p. 84-90.

26. Uijlings, J., et al., Selective Search for Object Recognition. International Journal of Computer Vision, 2013. 104(2): p. 154-171

27. Ranjan, R., V.M. Patel, and R. Chellappa, HyperFace: A Deep Multi-task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017: p. 1-1.

28. Ranjan, R., et al., An All-In-One Convolutional Neural Network for Face Analysis. 2017. 17-24

29. Sun, X., P. Wu, and S.C.H. Hoi, Face detection using deep learning: An improved faster RCNN approach. Neurocomputing, 2018. 299: p. 42-50.

30. Mei Wang, Weihong Deng, "Deep Face Recognition: A Survey", arXiv:1804.06655v9 [cs.CV] 1 Aug 2020

31. Farfade, S., M. J. Saberian, and L.-J. Li, Multi-view Face Detection Using Deep Convolutional Neural Networks. 2015

32. Yang, S., et al. From Facial Parts Responses to Face Detection: A Deep Learning Approach. in 2015 IEEE International Conference on Computer Vision (ICCV). 2015.

33. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Kauai, HI, USA, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517.

34. K. Zhang, Z. Zhang, Z. Li and Y. Qiao, "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," in *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499-1503, Oct. 2016, doi: 10.1109/LSP.2016.2603342.

35. Manik Sharma, J Anuradha, H KManne and G S CKashyap, "Facial detection using deep learning", IOP Conf. Series: Materials Science and Engineering 263 (2017) 042092 doi:10.1088/1757-899X/263/4/042092

36. J. Mehta, E. Ramnani and S. Singh, "Face Detection and Tagging Using Deep Learning," *2018 International Conference on Computer, Communication, and Signal Processing (ICCCSP)*, Chennai, 2018, pp. 1-6, doi: 10.1109/ICCCSP.2018.8452853.

37. M. Osadchy, Y. L. Cun, and M. L. Miller, "Synergistic face detection and pose estimation with energy-based models," Journal of Machine Learning Research, vol. 8, no. May, pp. 1197–1215, 2007

38. L. Chi, H. Zhang, and M. Chen, "End-to-end face detection and recognition," arXiv preprint arXiv:1703.10818, 2017.

39. X. Fontaine, R. Achanta, and S. Susstrunk, "Face recognition in real- ¨ world images," in 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), March 2017, pp. 1482–1486

40. Hongxin Zhang, Liying Chi, "End-to-End Spatial Transform Face Detection and Recognition", Virtual Reality & Intelligent Hardware Volume 2, Issue 2, April 2020, Pages 119-131

41. S. Yang, P. Luo, C. C. Loy, and X. Tang. Wider face: A face detection benchmark. In Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pages 5525–5533, June 2016. 1, 2, 5

42. D. Yi, Z. Lei, S. Liao, and S. Z. Li. Learning face representation from scratch. arXiv, 2014. 1, 2, 5

43. V. Jain and E. G. Learned-Miller. Fddb: A benchmark for face detection in unconstrained settings. UMass Amherst Technical Report, 2010. 1, 2, 5

44. G. B. Huang, M. Ramesh, T. Berg, and E. Learned Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments Technical report, Technical Report 07-49, Universityof Massachusetts, Amherst, 2007. 1, 2, 5

45. Xinyi Ding , ZohrehRaziei, Eric C. Larson, Eli V. Olinick, Paul Krueger3 and Michael Hahsler, "Swapped face detection using deep learning and subjective assessment", EURASIP Journal on Information Security, (2020) 2020:6 https://doi.org/10.1186/s13635-020-00109-8

46. Shailypandey and Sandeep Sharma, "Review: Face Detection and Recognition Techniques", International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 4111-4117

47. Ashu Kumar, · Amandeep Kaur,· MunishKuma, "Face detection techniques: a review", Artificial Intelligence Review, Springer, 2018, https://doi.org/10.1007/s10462-018-9650-2

48. Arun Kumar Dubey&Vanita Jain (2019) A review of face recognition methods using deep learning network, Journal of Information and Optimization Sciences, 40:2, 547-558, DOI: 10.1080/02522667.2019.1582875

49. Thai Hoang Le, "Applying Artificial Neural Networks for Face Recognition", Hindawi Publishing Corporation Advances in Artificial Neural Systems Volume 2011, Article ID 673016, 16 pages doi:10.1155/2011/673016

50. CBCL Database #1, Center for Biological and Computational Learning at MIT and MIT, http://cbcl.mit.edu/software-datasets/FaceData2.html

51. S. Sharma, M. Bhatt and P. Sharma, "Face Recognition System Using Machine Learning Algorithm," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 1162-1168, doi: 10.1109/ICCES48766.2020.9137850.

52. Z. B. Lahaw, D. Essaidani and H. Seddik, "Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM Algorithms," 2018 41st International Conference on Telecommunications and Signal Processing (TSP), Athens, 2018, pp. 1- 5. doi: 10.1109/TSP.2018.8441452

53. N. Sabri et al., "A Comparison of Face Detection Classifier using Facial Geometry Distance Measure," 2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 2018, pp. 116-120.doi: 10.1109/ICSGRC.2018.8657592

54. J. Fan, Q. Ye and N. Ye, "Enhanced Adaptive Locality Preserving Projections for Face Recognition," 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nanjing, 2017, pp. 594-598. doi: 10.1109/ACPR.2017.123

55. K. L. Masita, A. N. Hasan and T. Shongwe, "Deep Learning in Object Detection: a Review," 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD), Durban, South Africa, 2020, pp. 1-11, doi: 10.1109/icABCD49160.2020.9183866.

56. Liu, L., Ouyang, W., Wang, X. *et al.* Deep Learning for Generic Object Detection: A Survey. *Int J Comput Vis* **128,** 261–318 (2020). https://doi.org/10.1007/s11263-019-01247-4

57. T. Albrecht, M. Luthi and T. Vetter, Deformabl e Models, University of Basel, (Switzerland 2015)

58. L. Perrinet, Sparse Models for Computer Vision, Biologically inspired computer vision, (2015)

59. B.A Hoadley, Bayesian Look at Inverse Linear Regression, Journal of the American Statistical Association , Vol. 65, issue 329 (2012), pp. 356-369

60. Z. Xu and R. Akella, A bayesian logistic regression model for active relevance feedback, Proceedings ofthe 31st annual international ACM SIGIR conference on Research and development in information retrieval (Singapore, 2008), pp. 227-234

61. A.M. Farayola , AN. Hasan and Ali. Ahmed, Efficient Photovoltai c MPPT System Using Coarse Gaussian Support Vector Machine and Artificial Neural Network Techniques, in International Journal of Innovative Computing Information and Control (UICIC),Vol. 14, No. 1 (Feb. 2018)

62. V. Vapnik, Statistical Learning Theory, Wiley New York, Inc. (1998).

63. L. Deng and D. Yu, Deep Learning: Methods and Applications, Foundations and Trends in Signal Processing, Vol. 7: No. 3-4 (2014), pp 197-387.

64. B. Cyganek, Object detection and Recognition in Digital Images, Theory and Practice, (May, 2016)

65. 1. Goodfellow, Y. Bengio and A Courville, Deep Learning (2013), pages 654-720

66. A. Krizhevsky, I Sutskever. and G Hinton, ImageNet classification with deep convolutional neural Networks, Advances in Neural Information Processing Systems, (2012), pp. 1106-1114

67. A. Michael Nielsen, Neural Networks and Deep Learning, Determination Press, 2015

68. P. Vincent, H. Larochelle, Y. Bengio and P.A. Manzagol, Extracting and composing robust features with denoisingautoencoders, In ICML 2008, pg. 241.

69. N. Le Roux, and Y. Bengio, Representational power of restricted Boltzmann machines and deep belief networks, Neural Computation, Vol. 20, No.6 (2008), pp. 1631-1649.

70. Liu, L., Ouyang, W., Wang, X. *et al.* Deep Learning for Generic Object Detection: A Survey. *Int J Comput Vis* **128,** 261–318 (2020). https://doi.org/10.1007/s11263-019-01247-4

71. Zhang, X., Yang, Y., Han, Z., Wang, H., &Gao, C. (2013). Object class detection: A survey. ACM Computing Surveys, 46(1), 10:1–10:53

72. Li, Y., Wang, S., Tian, Q., & Ding, X. (2015b). Feature representation for statistical learning based object detection: A review. Pattern Recognition, 48(11), 3542–3559

73. Borji, A., Cheng, M., Jiang, H., & Li, J. (2014). Salient object detection: A survey, 1, 1–26. arXiv:1411.5878v1.

74. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE TPAMI, 35(8), 1798–1828

75. Litjens, G., Kooi, T., Bejnordi, B., Setio, A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60–88.

76. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., et al. (2018). Recent advances in convolutional neural networks. Pattern Recognition, 77, 354–377

77. Everingham, M., Eslami, S., Gool, L. V., Williams, C., Winn, J., &Zisserman, A. (2015). The pascal visual object classes challenge: A retrospective. IJCV, 111(1), 98–136

78. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet large scale visual recognition challenge. IJCV, 115(3), 211–252

79. Lin, T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., &Zitnick, L. (2014). Microsoft COCO: Common objects in context. In ECCV (pp. 740–755)

80. Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., &Torralba, A. (2017a). Places: A 10 million image database for scene recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(6), 1452–1464

81. Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., PontTuset, J., et al. (2018). The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. arXiv:1811.00982

82. Y. Tang, X. Wang, E. Dellandréa, and L. Chen, "Weakly Supervised Learning of Deformable Part-Based Models for Object Detection via Region Proposals," IEEE Trans. Multimed., vol. 19, no. 2, pp. 393–407, 2017

83. AsimSuhail, ManojJayabalan, VineshThiruchelvam, "CONVOLUTIONAL NEURAL NETWORK BASED OBJECT DETECTION: A REVIEW", Journal of Critical Reviews, Vol 7, Issue 11, 2020

84. Z. Zhao, P. Zheng, S. Xu and X. Wu, "Object Detection With Deep Learning: A Review," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, pp. 3212-3232, Nov. 2019, doi: 10.1109/TNNLS.2018.2876865.

85. Kang Tong, Yiquan Wu ∗, Fei Zhou, "Recent advances in small object detection based on deep learning: A review",Elsevier, Image and Vision Computing, Volume 97, May 2020, 103910

86. A. Geiger, P. Lenz, R. Urtasun, Are we ready for autonomous driving? The KITTI vision benchmark suite, Computer Vision and Pattern Recognition 2012, pp. 3354–3361.

87. C. Wojek, P. Dollar, B. Schiele, P. Perona, Pedestrian detection: an evaluation of the state of the art, IEEE Trans. Pattern Anal. Mach. Intell. 34 (4) (2012) 743–761

88. S. Romberg, L.G. Pueyo, R. Lienhart, R.V. Zwol, Scalable logo recognition in realworld images, International Conference on Multimedia Retrieval 2011, pp. 25–33.

89. J. Xiao, K.A. Ehinger, J. Hays, A. Torralba, A. Oliva, SUN database: exploring a large collection of scene categories, Int. J. Comput. Vis. 119 (1) (2010) 3–22

90. Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, S. Hu, Traffic-sign detection and classification in the wild, Computer Vision and Pattern Recognition, 2016

91. C. Chen, M.-Y. Liu, O. Tuzel, J. Xiao, R-CNN for small object detection, Asian Conference on Computer Vision 2016, pp. 214–230.

92. Mingxing Tan Ruoming Pang Quoc V. Le, "EfficientDet: Scalable and Efficient Object Detection", arXiv:1911.09070v7 [cs.CV] 27 Jul 2020

93. Kaiming He Georgia GkioxariPiotr Dollar Ross Girshick, "Mask R-CNN", arXiv:1703.06870v3 [cs.CV] 24 Jan 2018

94. R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, 2014

95. S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In NIPS, 2015

96. GolnazGhaisiTsung-Yi Lin Ruoming Pang Quoc V. Le Google Brain, "NAS-FPN: Learning Scalable Feature Pyramid Architecture for Object Detection", arXiv:1904.07392v1 [cs.CV] 16 Apr 2019

97. Shu Liu, Lu Qi† Haifang Qin, Jianping Shi, JiayaJia, "Path Aggregation Network for Instance Segmentation", arXiv:1803.01534v4 [cs.CV] 18 Sep 2018

98. Jifeng Dai, Yi Li, Kaiming He, Jian Sun, "R-FCN: Object Detection via Region-based Fully Convolutional Networks",

99. Wei Liu1 , Dragomir Anguelov2 , Dumitru Erhan3 , Christian Szegedy3 , Scott Reed4 , Cheng-Yang Fu1 , Alexander C. Berg1, "SSD: Single Shot MultiBox Detector", arXiv:1512.02325v5 [cs.CV] 29 Dec 2016

100. Joseph Redmon∗ , Santosh Divvala∗†, Ross Girshick¶ , Ali Farhadi∗, "You Only Look Once: Unified, Real-Time Object Detection", arXiv:1506.02640v5 [cs.CV] 9 May 2016.

101. V. Mareeswari, VijayanRamaraj, G. Uma Maheswari, R. Sujatha, Preethi. E, "Accessibility Using Human Face, Object and Text Recognition for Visually Impaired People", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-6, April 2019

102. Rakesh Chandra Joshi 1 , SaumyaYadav 1 , Malay Kishore Dutta 1,* and Carlos M. Travieso-Gonzalez , "Efficient Multi-Object Detection and Smart Navigation Using Artificial Intelligence for Visually Impaired People",entropy, MDPI, Entropy 2020, 22, 941; doi:10.3390/e22090941

103. Shiwei Chen, Dayue Yao, Huiliang Cao * and Chong Shen , "A Novel Approach to Wearable Image Recognition Systems to Aid Visually Impaired People", MDPI, Appl. Sci. 2019, 9, 3350; doi:10.3390/app9163350

104. Shifa Shaikh, VrushaliKarale and Gaurav Tawde "Assistive Object Recognition System for Visually Impaired", International Journal of Engineering Research & Technology (IJERT), Vol. 9 Issue 09, September-2020.

105. N.Jayashree, Dr.Y.Kalpana, "Survey on Face Recognition, Object and Text Detection For Visually Challenged People", International Journal of Pure and Applied Mathematics Volume 119 No. 10 2108, 161-168 ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version).

106. Hizem W., Krichen E., Ni Y., Dorizzi B., Garcia-Salicetti S. (2005) Specific Sensors for Face Recognition. In: Zhang D., Jain A.K. (eds) Advances in Biometrics. ICB 2006. Lecture Notes in Computer Science, vol 3832. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11608288_7

107. Y. Ni, X.L. Yan, "CMOS Active Differential Imaging Device with Single in-pixel Analog Memory", Proceedings of IEEE European Solid-State Circuits Conference (ESSCIRC'02), pp. 359-362, Florence, Italy, Sept. 2002.

108. W. Hizem, Y. NI and E. Krichen, "Ambient light suppression camera for human face recognition" CSIST Pekin 2005

109. Mingyuan Zhou, Haiting Lin, Jingyi Yu and S. S. Young, "Hybrid sensing face detection and recognition," 2015 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington, DC, 2015, pp. 1-9, doi: 10.1109/AIPR.2015.7444542.

110. K.Venkatachalam, A.Devipriya, J.Maniraj, M.Sivaram, A.Ambikapathy, Iraj S Amiri, "A Novel Method of motor imagery classification using eeg signal", Journal Artificial Intelligence in Medicine Elsevier, Volume 103, March 2020, 101787

111. Yasoda, K., Ponmagal, R.S., Bhuvaneshwari, K.S. K Venkatachalam, " Automatic detection and classification of EEG artifacts using fuzzy kernel SVM and wavelet ICA (WICA)" Soft Computing Journal (2020).

112. P. Prabu, Ahmed Najat Ahmed, K. Venkatachalam, S. Nalini, R. Manikandan,Energy efficient data collection in sparse sensor networks using multiple Mobile Data Patrons, Computers & Electrical Engineering,Volume 87,2020.

113. V.R. Balaji, Maheswaran S, M. Rajesh Babu, M. Kowsigan, Prabhu E., Venkatachalam K,Combining statistical models using modified spectral subtraction method for embedded system,Microprocessors and Microsystems, Volume 73,2020.

114. Malar, A.C.J., Kowsigan, M., Krishnamoorthy, N. S. Karthick, E. Prabhu & K. Venkatachalam (2020). Multi constraints applied energy efficient routing technique based on ant colony optimization used for disaster resilient location detection in mobile ad-hoc network. Journal of Ambient Intelligence and Humanized Computing, 01767-9.