Indoor Object Detection

Dr. Shripad Bhatlawande   
shripad.bhatlawande@vit.edu

Atharva Chaudhari  
ORCID: 0000-0001-5443-6123

Chinmay Deotale  
ORCID: 0000-0002-7221-0934

Shubham Damale  
ORCID: 0000-0003-0064-5426

\*Department of Electronics and Telecommunication, Vishwakarma Institute of Technology, Pune, India, 411038

Abstract— One of the most vital senses for living is vision. Tens of millions of people residing in this world deal with visual impairment. Those people find difficulty in recognizing indoor objects independently and adequately. The aids that are currently available specifically for indoor object detection are limited and most of them just detect the objects and doesn’t recognize it. The objective of the proposed work is to alternate the visual world into an audio world by notifying the blind people about the indoor objects such as chair, sofa and wardrobe. This will also help them to know more about their surroundings and to find a place to rest/sit such as chair and sofa. It comprehensively includes a variety of crucial strategies, that includes SIFT feature extraction and machine learning algorithm such as support vector machine, decision tree, random forest and K-Nearest Neighbor. Further these models were evaluated based on numerous parameters such as accuracy, recall, precision, f1-score, sensitivity and specificity. The result of model evaluation showed that decision tree gave maximum accuracy of 77.64%, followed by random forest and SVM with 75.2% and 72.31% respectively.

Keywords— Computer Vision, Aid to visual impaired, Household object detection, SIFT, Dimensionality Reduction, Machine Learning.

# **Introduction**

Visual impairment is a reduced capacity of an eye to see clearly. It causes issues that are no longer fixable with the aid of using typical means, such as spectacles. Vision impairment happens when several parts of the eye and brain that are needed for the clear vision get damaged. It can’t be fully cured with medicine and surgery [1]. There are 2.2 billion people suffering from some kind of visual impairment (worldwide). In age wise statistics it was observed majority people are over the age of 50. According to age wise statistics of India alone, the majority of people suffering from visual impairment are above the age of 80 (11.6%), after this 4.1% people are in the age group of 70-79, followed by 1.6% people in 60-69 and 0.5% are in 50-59 [2]. There are many types of information that is encoded in the surrounding but due to visual impairment, these people are not able to detect the objects [3]. They are not able to inspect spaces that are not known. They are not able to detect the object of interest [4].

They need some assistance from other person. Our model is built with the objective of helping visually impaired people for safe mobility. This will aid them to know more about their indoor surroundings by recognizing nearby objects.

# **Literature Survey**

As stated in the introduction section, a lot of people suffer from visual impairment. To tackle this problem, many sensor and computer vision-based object detection systems have been proposed. This section gives a brief survey of recent work done on object detection. For the detection of obstacle, sonar sensor-based vector field histogram algorithm (VFH) was proposed. This algorithm creates feature vector from the reading of sonar sensor. Based upon that it provides audio output. But the limitation of sonar sensor is that it has limited testing distance [5]. Ultrasonic sensor attached on user’s spectacle was proposed to detect the obstacle hence distance of the obstacle from the person and to convey through vibration motor [6]. Amplitude modulated (AM) signal for the detection of the objects was explored in [7]. The object is attached with receiver. Transmitter is carried by blind person that constantly sends the AM signal. Receiver detects the signal, sound beacon emits the sound. Due to inefficiency in terms of bandwidth it may not be able to transmit enough signal to receiver that is required for detection.

To detect indoor human-object interactions, wrist mounted sensor is used. This sensor contains accelerometer and gyroscope. To detect the interactions, Random Forest, J48, and support vector machine (SVM) classifiers are trained. Microprocessor contains the pre-trained model. Based upon the input values of the sensors, audio output is provided [8].

To detect the object, Electromagnetic sensors have also been used [9]. Due to inefficiency of sensor to detect large portion, many infrared sensors mounted on spectacle approach is used. With the help of many sensors, large number of portions for obstacle can be detected. By using receiver, obstacles are detected [10]. To help visually impaired people to detect the objects, blind man sticks and canes have been developed that contains many ultrasonic sensors to detect the obstacles on the bottom, at middle and at high level. Receivers receives the signal, buzzer starts with different types of intensity depending upon the time to receive the signal [11], [12]. Computer Vision and multi-sensor based systems are used for object detection in indoor environment. For classification, support vector machine algorithm is used. Multi-sensor concept is employed by interfacing ultrasonic sensor and Infrared sensor to detect small object near feet [13]. Depth camera-based object detection and localization is implemented using Simultaneous localization and mapping (SLAM) algorithm. By using successive camera frames, it tracks the set points. The limitation of the depth camera is that it can’t detect the transparent objects [14]. Raspberry pi 3, camera, processor, earphone-based system was used for object detection. Convolutional Neural Network model was used for classification. Camera detects the object and based upon the output of convolutional neural network, audio output is provide to blind people. These elements are present on the cap. Its performance depends upon the orientation of the cap [15]. Context-SVM is used in such a way that result of detection of an image is used for classification of an image [16]. Similarly, for object classification, naive bayes algorithm is used. It is useful when there is lot of change in images of particular class. There are many objects in as object bank. Based upon the input image features this algorithm classify the image into pre-defined classes. But the problem with this algorithm is that it follows assumption that, all the features are independent [17]. Due to limitation of dataset, many deep learning concepts are used. Support vector machine (SVM) is trained on this neural network to increase the accuracy of object detection [18]. A system was designed to recognize doors and four types of signs such as exit, Water Closet (WC), disabled exit and confidence zone. Transfer learning technique based approach was used to detect indoor signs and doors and recognition was done using neural network models. [19].

Local kernels are used for structure data that cannot be of fixed length, and fixed features. Kernel provide high value when there is similarity between the images. Support vector machine is used with kernel it provides higher accuracy [20]. To classify the image in real time there is a concept of sliding window. But the limitation of sliding window is that, it takes large time for recognition if feature vectors of image are large. In this method the sub-image is multiplied with only the sub image that has same feature sizes [21]. For the detection of smaller objects various methods are implemented. Lightweight and Stable extraction module reduces the size of an image. Enhance Feature Module achieves more powerful feature extraction by spatial pyramid pooling network. After that accurate detection module determines the scale for more accuracy using the network called as lightweight block [22]. Similarly, a system for recognition and classification of indoor sign was implemented. For feature extraction dense scale invariant feature transform (DSIFT) and histogram of oriented gradients (HOG) was used and for classification, techniques such as the neural network (NN), support vector machine (SVM) and k-nearest neighbors (KNN) was used [23]. A system for indoor navigation using localization technologies such as WiFi, Bluetooth and RFID was explored. It was designed to help a visually impaired person to navigate to precise objects containing RFID tags. For safe mobility Dijkstra’s algorithm and Audio feedback was used [24]. Wearable devices to detect any obstacle on the ground was proposed. This system uses Red, Green, Blue and Depth camera for sensing the obstacles. For object classification, SVM classifier was used [25]. A system for crowd detection using crowd scoring data was implemented in [26]. Camera detects the person and score of people increases and it is compared with the dataset to provide audio output to the bind person. To reduce the collinear features that are extracted from an image, PCA based approach is used. Features of an image is provided to the KNN classifier that classifies the image. But with the high scale data, prediction performance is very less [27] [28].

# **Methodology**

The proposed system aims to detect and classify whether the input image is a household object such as chair, sofa or wardrobe. To achieve this, a Computer Vision and Machine Learning based approach is used. This system is built with the objective of helping visually impaired people for safe mobility. This will help them to know about their surroundings and to find a place to rest/sit such as chair or sofa.

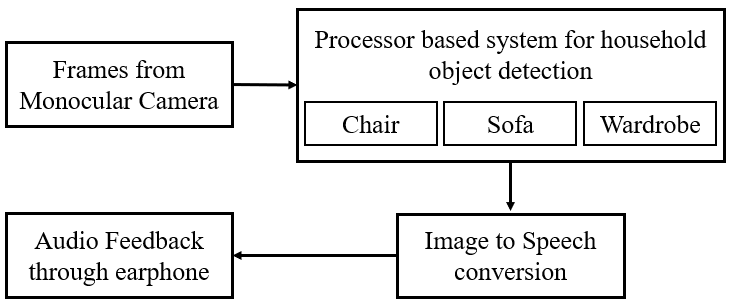


Fig. 1. Block diagram of electronic system for indoor object detection

The electronic system shown in Fig. 1 consists of camera, processor and earphone. Camera captures frames of in-house view. These frames are provided as an input to processor. The processor collects the frames of the surroundings and classifies the frame as a household object such as chair, sofa or wardrobe. The output is provided in form of audio through earphone to the visually impaired person.

**A) Image Preprocessing and Feature extraction:**

A custom dataset was used containing a total of 9000 images belonging to three classes. In this, 20% images were obtained from [29] and 80% images were acquired from Internet. The dataset is divided into 3 classes that are, Chair, Sofa and Wardrobe containing 3,000 images each. Images displaying multiple objects were discarded and to maintain region of interest (ROI) 535 images were cropped. Sample images for 3 classes are shown in Fig. 2.

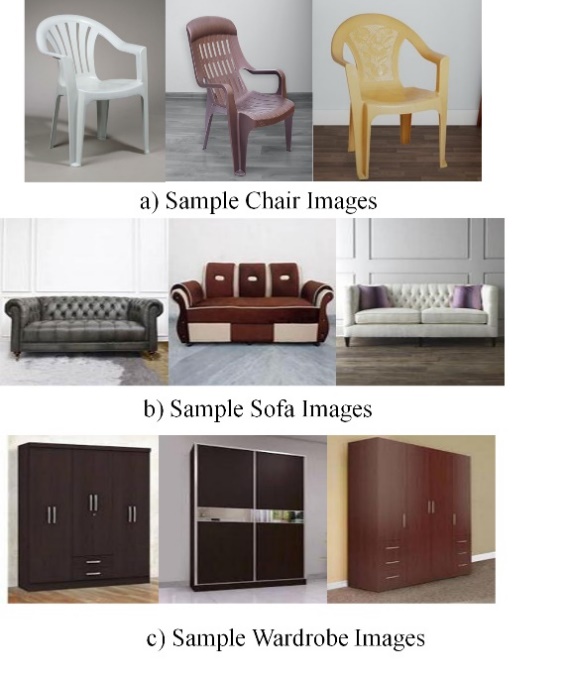


Fig. 2. Sample images of 3 classes

The images were resized to 400x400 pixels. Images were converted from BGR to Grayscale and median filter of kernel size 5 was applied on those images to reduce impulses and excess noise.

Scale Invariant Feature Transform (SIFT) was used for feature detection and description as the features extracted by SIFT are invariant to image scaling and rotation. It can generate large numbers of features that densely cover the image over the full range scales and locations. SIFT algorithm is used for both detecting and describing of keypoints in an image [30]. SIFT algorithm returns keypoints and descriptors of 128 dimensions.

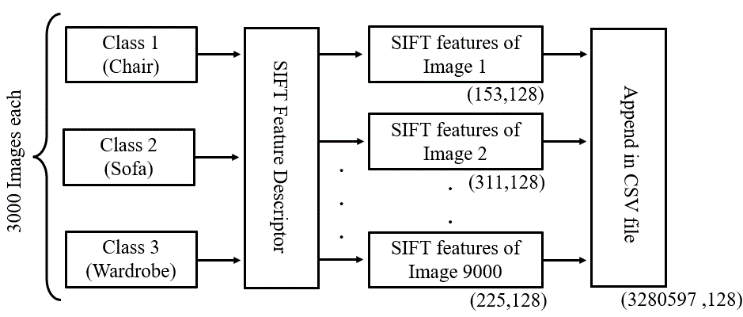


Fig. 3. Feature vector compilation using SIFT

The process of feature vector compilation performed on the images is depicted in Fig. 3. All 9,000 grayscale images were provided to SIFT. The feature vectors obtained from SIFT were appended in CSV file in order to perform further reduction in dimensions. The total number of features extracted from 9000 images were 3,280,597 features and of dimension 128.

Algorithm 1: Image preprocessing & feature extraction

Input: 9000 images of indoor objects

Output: CSV file containing SIFT features of 9000 images

*Initialization:*

**LOOP** process

1. **for** image in directory **do:**

2. read image

3. resize image to 400x400 pixels

4. convert image (BGR to Grayscale)

5. apply median filter

1. perform SIFT feature extraction
2. append feature vectors in CSV file

8. **end for**

9. **return** CSV file (3280597, 128)

For fast convergence of algorithm in real time, dimensionality reduction was applied.

**B) Feature Vector Optimization:**

Unsupervised K-Means clustering algorithm and Principal Component Analysis technique was used for feature vector optimization.

K-Means clustering is a data partitioning technique that is used for clustering. K-Means clustering model was trained on previously saved CSV file containing SIFT feature descriptions of all images. An appropriate value of K=8 (number of clusters to be formed) was decided using elbow method.

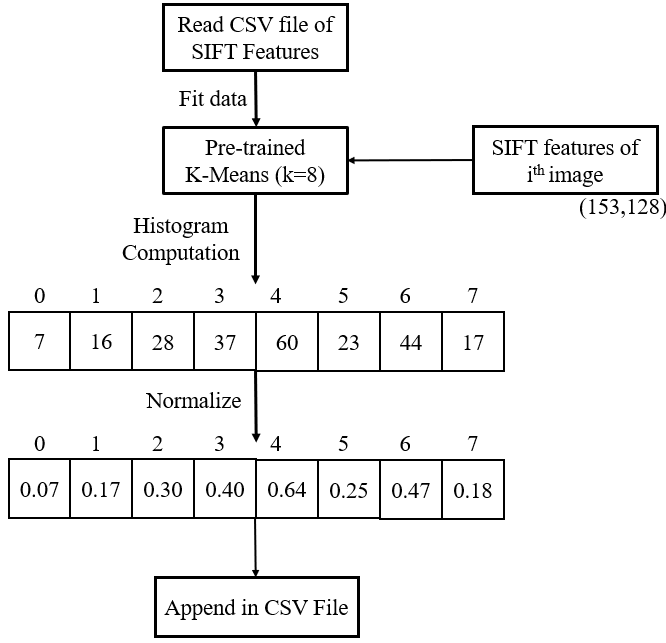


Fig. 4. Dimensionality reduction using K-Means clustering

K-Means model was trained using the data from CSV file. Then SIFT features from an image were extracted and compacted into an 8 bins histogram and hence every image was represented using feature vector of size (1, 8) as shown in Fig. 4. The histogram obtained was then normalized to scale individual samples to have unit norm. This process was carried out for all the images and was saved in new CSV file.

Principal Component Analysis (PCA) was applied on the feature vectors obtained from grounding of SIFT features. It is a dimensionality reduction technique that helps to identify correlated data and patterns in a data set. It reduces data over-fitting issues by decreasing the number of features without losing important information.

PCA was applied on the data. It was observed that first 6 principal components contained 93.48% of the information about the data hence first 6 principal components were considered. The process of dimensionality reduction is described in Algorithm 2.

**Algorithm 2:** Feature vector optimization

Input: 9000 images of objects, Pre-trained K-Means model

Output: Optimized feature vector (9000, 6)

*Initialization:*

**LOOP** process

1. **for** i <=number of images in image folder **do**:

2. read ith image

3. perform SIFT feature extraction (n, 128)

4. ground SIFT vector using pre-trained

K-Means model(1, 8)

5. perform histogram equalization

6. normalize histogram

7. append result in CSV file

8. i=i+1

9. **end for**

10. initialize PCA (n\_components = 6)

11. transform data from CSV file using PCA (9000, 6)

12. **return** final feature vector

Data splitting was performed on final feature vector. Dataset was split in training and testing data in 75% and 25% respectively. The training data was used to train an array of four classifiers.

**C) Classification and Recognition of Household Furniture:**

Machine learning algorithms such as (i) Support Vector Machine (SVM), (ii) Random Forest (RF), (iii) Decision Tree (DT) and (iv) K-Nearest Neighbors (KNN) were used for classification of household furniture. These classifiers were then evaluated to choose parameters that provide better accuracy, less training and testing time, higher precision and recall value, better f1 score and better utility usage.

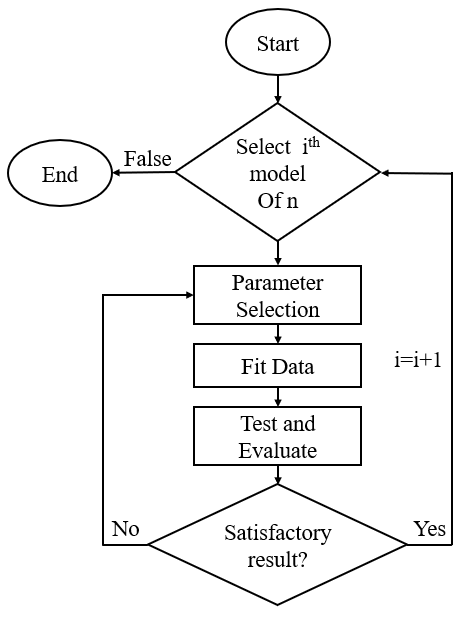


Fig. 5. Parameter selection for classifiers

The best parameters were chosen for the model using process shown in Fig. 5. A model was initialized with its default parameters. The training data was fitted to the model and further used for prediction on testing data. The model was further evaluated based on accuracy, recall, precision, f1-score, sensitivity and specificity. If the results were not satisfactory, same model was initialized using appropriate parameters. This process was repeated till the best possible parameters were obtained. Next model was chosen when satisfactory results of previous model was obtained.

First classifier used was Support Vector Machine (SVM). It constructs a hyper-plane in a higher dimensional space that can be used for classification [31]. SVM can be expressed using equation:

(1)

Where ‘k’ and ‘a’ are parameters defined as SVM parameters as signifying weighted vector and bias respectively. The optimization problem is to reduce the square of the weighted vector as:

(2)

Linear, Polynomial and Radial Basis Function (RBF) kernel were used to train the model.

Second classifier used was Decision Tree. In a decision tree, features of the dataset are represented by internal nodes, branches of the tree represent the conditions depending on that the decision is taken and leaf nodes represent the output. Information Gain is used for splitting the nodes if the target variable is categorical. It works on the concept of the entropy. Entropy is used for calculating the purity of a node.

(3)

Where, is probability of element/class i.

Mathematically, Information Gain (G) is represented as:

(4)

Where, “b” is the dataset before the split, k is the number of subsets generated by the split, and (j, a) is subset j after the split.

The third classifier used was Random Forest. The algorithm contains many decision trees and group the models that perform better than random guessing to build robust learner. Finally, accuracy is obtained by averaging the result from the decision trees that is signified in equation (5) [32] [33]:

(5)

Where, M is prediction obtained, ith is prediction of that particular decision tree and ‘i’ is number of trees.

Final classifier used was K-Nearest Neighbor (KNN). It works by considering K- Nearest data points to predict the class. Learning of this algorithm is instance based.

The equation for KNN is given in equation (6).

(6)

The detailed analysis of performance of all classifiers is discussed in next section.

# **Result And Discussion**

Features were extracted using SIFT and after Dimensionality reduction, dataset was divided into training and the testing sets. The system required 17.33 minutes for feature extraction using SIFT and 19.12 minutes for Dimensionality reduction (K-Means clustering and PCA) when tested on a laptop with processor specification of intel core i5 8th Generation with CPU clock speed of 1.8 GHz and 8 GB RAM.

Table 1: Accuracy comparison between classifiers

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model  Name | Test  Acc. | Train  Acc. | Training Time  MM:SS:MS | Testing Time  MM:SS:MS |
| SVM | 72.31 | 93.44 | 0:07.67 | 0:02.67 |
| DT | 77.64 | 99.42 | 0:00.047 | 0:00.001 |
| RF | 75.2 | 94.11 | 0:01.92 | 0:00.15 |
| KNN | 53.29 | 56.52 | 0:00.40 | 0:00.26 |

*\*Note: Acc.-Accuracy, SVM- Support Vector Machine, DT- Decision Tree, RF- Random Forest, KNN- K Nearest Neighbors, MM: SS: MS - Minutes: Seconds: Milliseconds*

The evaluation of models showed that Decision Tree provided maximum accuracy of 77.64% and KNN had minimum accuracy of 53.29%. KNN is a distance-based algorithm. While working on dataset of large size and high dimension, the performance of the algorithm degrades. This is the reason that KNN gave least accuracy. Decision tree performed slightly better than SVM as it deals with collinearity in categorical data better. Random Forest provides the high performance although the training accuracy is drastically high when compared to the testing accuracy as described in the Table 1. This implies that the Random Forest is over-fitting the data that will affect its accuracy when used in real life scenarios. So, considering the over-fitting anomaly it can be analyzed that SVM and Decision tree classifier will be an ideal choice with real world implementation.

Training time as well as testing time was maximum for SVM followed by Random Forest, KNN and Decision Tree. Decision tree was fastest as it generalizes the data in advance whereas classifiers such as KNN tends to be slower with large datasets because they scan the whole dataset for prediction.

Fig. 6. Evaluation parameter comparison of classes with SVM (RBF kernel)

Accuracy obtained by SVM was 72.31% on testing data whereas it was 93.44% on training data. Chair had higher values in every parameter when compared with other class of Sofa and Wardrobe when trained on SVM with RBF kernel as observed from Fig. 6.

Table 2: Evaluation parameters for Decision Tree

|  |  |  |  |
| --- | --- | --- | --- |
| Decision Tree | Chair | Sofa | Wardrobe |
| Precision | 80% | 77% | 76% |
| Recall | 80% | 72% | 81% |
| F1-score | 80% | 74% | 79% |
| Sensitivity | 92.18% | 71.91% | 81.01% |
| Specificity | 89.98% | 89.11% | 87.37% |

Accuracy obtained by Decision tree classifier was highest with 77.64. Value of Precision was highest for Chair as shown in table 2. It shows that the proportion of positive identifications that were actually correct were highest for chair and lowest for Wardrobe. Conversely, the Value of Recall was highest for Wardrobe. It shows that proportion of actual positives identified correctly was highest for Wardrobe followed by Chair and Sofa. Other parameters such as F1-score, Sensitivity and Specificity was highest for Chair class.

SVM with RBF kernel was evaluated on basis of utility by taking 1000 house-hold images of each class to predict through the model. Out of the 1000 images of chair, 735 images were correctly classified whereas 265 were misclassified. In this, 124 chair images were misclassified as Sofa and 141 were misclassified as Wardrobe. In case of Sofa, 815 out of 1000 were correctly classified and for Wardrobe 824 gave correct prediction. It was observed that Wardrobe and Sofa performed better than Chair as the number of correct classifiers were higher as compared to Chair class.

# **Conclusion**

The paper proposes a novel approach that will be effective in providing assistance to visually impaired people by detecting and recognizing indoor objects. The system will also provide an audio alert after the object is detected. It will help the visually impaired person to provide a comprehensive idea about the surrounding indoor environment and also find a place to rest i.e., chair or sofa. Our system specifically targets indoor object recognition as not much work is done in this area. The camera can be mounted on blind glasses and so that it will be comfortable for them and can be integrated as a part of their lives. Model was trained on features that were extracted from 3000 images from each of 3 classes (Chair, Sofa and Wardrobe) and SIFT was used for feature detection and description and multiple classifiers were used. Out of all classifiers, Decision Tree gave the best accuracy i.e. 77.64%.

Since a large population is suffering from some kind of visual impairment, this system can be of a great use to them. It can also be used by senior citizens who have blurred vision. The proposed system can be improved by adding more categories of indoor objects such as table, bed etc. The model with some modification can also be used for obstacle detection to ensure safety of blind person once few more categories of indoor objects are added and the model can be trained using a dataset with a greater number of images to improve the accuracy. The proposed model provides high accuracy with testing data as well as in further utility evaluation. Further a lidar sensor can also be added to estimate the distance of the object from the person.

##### References

1. Web resource URL: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>(Last accessed on: 31/12/2021)
2. Web resource URL: <https://www.livemint.com/news/india/estimates-of-blindness-reduced-by-47-in-12-years-govt-survey-11570733865393.html> (Last accessed on: 31/12/2021)
3. J. P. Bigham, C. Jayant, A. Miller, B. White and T. Yeh, "VizWiz::LocateIt - enabling blind people to locate objects in their environment," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2010, pp. 65-72, doi: 10.1109/CVPRW.2010.5543821.
4. Khan, Akif; Khusro, Shah (2020). An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities. Universal Access in the Information Society, (), –. doi:10.1007/s10209-020-00733-8
5. B. Li et al., "Vision-Based Mobile Indoor Assistive Navigation Aid for Blind People," in IEEE Transactions on Mobile Computing, vol. 18, no. 3, pp. 702-714, 1 March 2019, doi: 10.1109/TMC.2018.2842751.
6. K. Patil, Q. Jawadwala and F. C. Shu, "Design and Construction of Electronic Aid for Visually Impaired People," in IEEE Transactions on Human-Machine Systems, vol. 48, no. 2, pp. 172-182, April 2018, doi: 10.1109/THMS.2018.2799588
7. P. Blenkhorn and D. G. Evans, "A system for enabling blind people to identify landmarks: the sound buoy," in IEEE Transactions on Rehabilitation Engineering, vol. 5, no. 3, pp. 276-278, Sept. 1997, doi: 10.1109/86.623019.
8. S. Tahir, A. Raheel, M. Ehatisham-Ul-Haq and A. Arsalan, "Recognizing Human-Object Interaction (HOI) Using Wrist-Mounted Inertial Sensors," in IEEE Sensors Journal, vol. 21, no. 6, pp. 7899-7907, 15 March15, 2021, doi: 10.1109/JSEN.2020.3044315.
9. E. Cardillo et al., "An Electromagnetic Sensor Prototype to Assist Visually Impaired and Blind People in Autonomous Walking," in IEEE Sensors Journal, vol. 18, no. 6, pp. 2568-2576, 15 March15, 2018, doi: 10.1109/JSEN.2018.2795046.
10. B. Ando and S. Graziani, "Multisensor Strategies to Assist Blind People: A Clear-Path Indicator," in IEEE Transactions on Instrumentation and Measurement, vol. 58, no. 8, pp. 2488-2494, Aug. 2009, doi: 10.1109/TIM.2009.2014616.
11. K. Kumar, B. Champaty, K. Uvanesh, R. Chachan, K. Pal and A. Anis, "Development of an ultrasonic cane as a navigation aid for the blind people," 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2014, pp. 475-479, doi: 10.1109/ICCICCT.2014.6993009.
12. A. Krishnan, G. Deepakraj, N. Nishanth and K. M. Anandkumar, "Autonomous walking stick for the blind using echolocation and image processing," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), 2016, pp. 13-16, doi: 10.1109/IC3I.2016.7917927.
13. C. T. Patel, V. J. Mistry, L. S. Desai and Y. K. Meghrajani, "Multisensor - Based Object Detection in Indoor Environment for Visually Impaired People," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, pp. 1-4, doi: 10.1109/ICCONS.2018.8663016.
14. J. Bai, S. Lian, Z. Liu, K. Wang and D. Liu, "Virtual-Blind-Road Following-Based Wearable Navigation Device for Blind People," in IEEE Transactions on Consumer Electronics, vol. 64, no. 1, pp. 136-143, Feb. 2018, doi: 10.1109/TCE.2018.2812498.
15. A. Nishajith, J. Nivedha, S. S. Nair and J. Mohammed Shaffi, "Smart Cap - Wearable Visual Guidance System for Blind," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), 2018, pp. 275-278, doi: 10.1109/ICIRCA.2018.8597327.
16. Q. Chen, Z. Song, J. Dong, Z. Huang, Y. Hua and S. Yan, "Contextualizing Object Detection and Classification," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 1, pp. 13-27, 1 Jan. 2015, doi: 10.1109/TPAMI.2014.2343217.
17. L. Zhang, X. Zhen and L. Shao, "Learning Object-to-Class Kernels for Scene Classification," in IEEE Transactions on Image Processing, vol. 23, no. 8, pp. 3241-3253, Aug. 2014, doi: 10.1109/TIP.2014.2328894.
18. S. Akcay, M. E. Kundegorski, C. G. Willcocks and T. P. Breckon, "Using Deep Convolutional Neural Network Architectures for Object Classification and Detection Within X-Ray Baggage Security Imagery," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 9, pp. 2203-2215, Sept. 2018, doi: 10.1109/TIFS.2018.2812196.
19. M. Afif, R. ayachi, Y. Said, E. Pissaloux and M. Atri, "Recognizing signs and doors for Indoor Wayfinding for Blind and Visually Impaired Persons," 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 2020, pp. 1-4, doi: 10.1109/ATSIP49331.2020.9231933.
20. H. Sahbi, J. -Y. Audibert and R. Keriven, "Context-Dependent Kernels for Object Classification," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 4, pp. 699-708, April 2011, doi: 10.1109/TPAMI.2010.198.
21. Pang, Yanwei; Zhang, Kun; Yuan, Yuan; Wang, Kongqiao (2014). Distributed Object Detection With Linear SVMs. IEEE Transactions on Cybernetics, 44(11), 2122–2133. doi:10.1109/TCYB.2014.2301453
22. Tao, Ye, Zhao Zongyang, Zhang Jun, Chai Xinghua, and Zhou Fuqiang. "Low-altitude small-sized object detection using lightweight feature-enhanced convolutional neural network." Journal of Systems Engineering and Electronics 32, no. 4 (2021): 841-853.
23. Z. Ni, S. Fu, B. Tang, H. He and X. Huang, "Experimental studies on indoor sign recognition and classification," 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), 2014, pp. 489-494, doi: 10.1109/CIDM.2014.7008707.
24. Abu Doush, Iyad, Sawsan Alshattnawi, Abdel Karim Al Tamimi, Bushra Alhasan and Safaa Hamasha. “ISAB: Integrated Indoor Navigation System for the Blind.” Interact. Comput. 29 (2017): 181-202.
25. Li, Zhongen; Song, Fanghao; Clark, Brian C.; Grooms, Dustin R.; Liu, Chang (2020). A Wearable Device for Indoor Imminent Danger Detection and Avoidance With Region-Based Ground Segmentation. IEEE Access, 8(), 184808–184821. doi:10.1109/access.2020.3028527
26. Y. Sun, W. Meng, C. Li, N. Zhao, K. Zhao and N. Zhang, "Human Localization Using Multi-Source Heterogeneous Data in Indoor Environments," in IEEE Access, vol. 5, pp. 812-822, 2017, doi: 10.1109/ACCESS.2017.2650953.
27. Zhou, Yanhong, Shukai Cao, Dong Wen, Huiyang Zhang, and Liqiang Zhao. "The study of face recognition based on hybrid principal components analysis and independent component analysis." In 2011 International Conference on Electronics, Communications and Control (ICECC), pp. 2964-2966. IEEE, 2011.
28. Xiaodan Wang, ; Zhaohui Shi, ; Chongming Wu, ; Wei Wang, (2006). [IEEE 2006 6th World Congress on Intelligent Control and Automation - Dalian, China ()] 2006 6th World Congress on Intelligent Control and Automation - An Improved Algorithm for Decision-Tree-Based SVM. , (0), 4234–4238. doi:10.1109/wcica.2006.1713173
29. Web resource URL: https://www.kaggle.com/akkithetechie/furniture-detector (Last accessed on: 31/12/2021)
30. Zhu Daixian, "SIFT algorithm analysis and optimization," 2010 International Conference on Image Analysis and Signal Processing, 2010, pp. 415-419, doi: 10.1109/IASP.2010.5476084.
31. C.-C. Chang and C.-J. Lin, “Libsvm: a library for support vector machines,” ACM Transactions on Intelligent Systems and Technology (TIST), vol. 2, no. 3, p. 27, 2011.
32. Vrushali Y Kulkarni and Pradeep K Sinha, “Efficient Learning of Random Forest Classifier using Disjoint Partitioning Approach”, in Proc. of the World Congress on Engineering, London, U.K, July 3-5, vol. 2, 2013.
33. Anna Bosch, Andrew Zisserman and Xavier Munoz, “Image Classification using Random Forests and Ferns,” in IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, October 14-20, pp. 1–8, 2007, IEEE.