Detection of Exercise and Cooking Scene for Assitance of Visually Impaired People

*Shripad Bhatlawande1, Swati Shilaskar2, Anant Abhyankar3, Mahesh Ahire4, Ankush Chadgal5, Jyoti Madake6

[1]Vishwakarma Institute of Technology, Pune, India

[1*]Shripad.bhatlawande@vit.edu, [2]Swati.shilaskar@vit.edu, [3]anant.abhyankar19@vit.edu, [4]mahesh.ahire19@vit.edu, [5]ankush.chadgal19@vit.edu, [6]Jyoti.madake@vit.edu

Abstract: This paper focuses on scene understanding for a visually impaired person. It classifies an image into two possible scenes namely, i) A scene involving some sort of exercise such as yoga. ii) A scene involving cooking. This system is developed with the aim of assisting the visually impaired people. It will help them to gain knowledge about their surroundings. The existing scene detection systems focus majorly on outdoor activities. This system may be the first system developed to identify indoor activities for visually impaired people. There are considerable challenges in terms of accessibility and usability in the existing systems in terms of lack of useful functions, intuitive feedback, cost, and operation methods. We have developed a system to understand if the detected scene is exercise or cooking. This system uses Scale Invariant Feature Transform to extract the features from the acquired scene. It uses Principal Component Analysis to optimize the dimensions of the feature vector. This optimization is based on maximum explained variance. This optimized feature vector was used to train 5 different classifiers viz- Decision Tree, SVM, K-NN, Random Forest, and Logistic Regression. These classifiers provided recognition accuracy as 75.91%, 80.17%, 79.31%, 80.83%, 79.68% respectively.

Keywords: Scene Detection, Scene Understanding, Image Recognition, Blind, Visual Impairment, SIFT, Random Forest

1 Introduction

Visual Impairment is a medical term used for a person who has any type of vision problem. Whether a person is not able to see anything or just has a problem with not being able to see properly. Many people who are blind cannot see anything but on the other hand, there are people who can see partially and those are called legal blindness.

If we talk about the worldwide statistics of visual impairment people are, In India over an estimated 2.19% of cases of visual impairment are from moderate-severe. And if we talk about the prevalence of pinhole blindness over 0.32% of people are suffering. In short, approximately 18.7 million people are blind in India. One out of every three blind people in the world is Indian.

Age-wise statistics of the people suffering from visual impairment are as 11.6% of people are above the age of 80, 4.1% of people are in between the age of 70-80, 1.6% are from 60-70, and from the age of 50-60 over 0.5% of people are suffering. Their major problems related to mobility are having problems seeing, so they are not able to move freely because they might collide with obstacles on the road or surface.

The visually impairment people have many problems identifying what is in front of them due to which they deal with many problems while moving from one place to another. There are many devices, or we can say Aids are available in the market for the help of visually defected persons some of them are A radio frequency identification (RFID) smart cane project. RFID is used to detect an impediment.[1] Kinesthesia is a high-tech video game designed specifically for visually impaired people. It provides feedback on the surrounding environment using an in-built camera, buzzers, and vibration motors [2]. Device design and functioning - this device is equipped with sensors that detect impediments such as staircases [4] and potholes [5]. Voice commands are also used to transmit information about them. The blind stick navigator is a technology that delivers an SMS notification and position to the appropriate guardian [6]. "Smart gloves for the blind" are wearable devices in the shape of a glove that costs near about INR 4000/-. Its purpose is to detect obstructions so that blind people can navigate independently. It is less expensive since it requires fewer components. Because of its small size, it is flexible, lightweight, and portable. It delivers real-time translation, but it requires extensive training to utilize. Arduino Uno, 1Shield, Ultrasonic Sensors, and a Smartphone were among the hardware components used in the construction. Shield Library and Arduino IDE were used as software components [7]. Be My Third Eye [8] is a type of Smart Electronic Blind Stick with Goggles is a hand-held stick featuring an Ultrasonic and IR sensor, goggles, GSM, GPS module motion sensor, PIR sensor, Arduino Uno, and a price tag of INR 1000/-. A stairwell, a wall, and other things are detected as obstacles. By employing ultrasonic waves to detect nearby barriers and notifying them with a buzzer sound or vibration, the device allows blind people to navigate with speed and confidence. However, accurate obstacle detection [9] is impossible due to limited hardware devices and insufficient data [10] 12-inch braille ruler which is specially designed for binds to draw straight lines and this costs same as the normal ruler [11] E-Z Grabber with Twist Shaft is a tool which costs near about INR 1500/- and is used to reach certain places where it is impossible to reach and is capable of holding a book and has non-slip rubber grip support [12]. Auditory-based Devices for the Blind Perceptual compensation is caused by blindness and manifests as an over-performance of hearing terms of physiology. In truth, a blind person's hearing is significantly more sensitive than a regular person. A highly skilled blind pedestrian can approach an intersection, end to the traffic, and judge the spatial layout of intersecting streets, the width of the street, the number of lanes of traffic in each direction, and the presence of pedestrian islands or medians solely based on auditory information, according to related physiologic research [13]. The vOICe: Meijer, a Dutch physicist, presented a real-time device named voice, which stands for "Oh, I can see." The voice system is a blind navigating system that scans and digitizes an image acquired by a video camera before converting the digital image to sound. It's a wearable gadget. The voice has shown to be highly valuable to the VI (Visually Impaired). The only drawback is that mastering and learning it takes a long time. It's almost as if it's your first language [14]. Sonic Torch and KASPA: Leslie Sy's contributions to solar mobility as a blind aid for the visually handicapped are significant. The majority of the study is focused on binaural, sensor, and Sonic Torch sonar systems. The above uses FM signals to calculate/generate the object's distance based on the pitch of the created sound [15]. Navi (Navigation Assistant for Visually Impaired): R. Nagarajan, Sazali Yaacob, and G. Sainarayanan designed this system. A significant improvement over VOICE. The information in front of the visually impaired was recorded by the vision sensor. The photograph was then analysed to determine what was in front of the person. Object recognition was being done in real-time image processing using fuzzy algorithms at the same time. It was a wearable gadget with much-enhanced image processing algorithms [16]. Blind People's Haptic Devices

The haptic or touch senses of VI's are heightened in the same way as the auditory senses of the blind are. Various haptic feedback is provided on the screens to offer the visually impaired an idea about the image, similar to reading braille, which is the language for the blind [17]. Prosthetic eye/artificial eye: The artificial eye provides a visual overview of the surroundings in the shortest period of time possible, allowing a vision-impaired individual to interact with normal people. This work proposes an all-optical technique for greatly improving man-made low-light imaging performance. The concept of superposition and the elephant nose, fisheye were the catalysts for this discovery. It boosted scotopic vision with an innovative photosensitivity enhancer, allowing our device to produce high-resolution images in low-light situations. Many applications require this capability [18]. Haptic smartphone: This study suggests the use of smartphone vibration to deliver turn-by-turn wall directions that are given ahead of time. It compares two feedback modalities, Wand and Screen Edge, with a third, Pattern. The prototype is created by performing a poll on eight random people and then walking them on a pre-programmed path using three different types of input, including vibration. The average error rate was found to be 4%. Pattern approach is selected after the experiment, followed by Screen Edge and Wand method [19].

A patient preference management system is explored in [20] is an enhanced Internet of Things (IoT) module with an artificial intelligence structure for efficiently moving data to the cloud and application layer. This system varies from typical IoT systems in that it adds a primary AI screening layer between the network layer and the cloud layer, and a secondary AI screening layer between the application layer and the cloud layer. Design of Deep Learning Algorithm for IoT Application by Image-based Recognition is explored in [21] by using image modification, the Principal Component Analysis (PCA) technique has extracted the different image features, resulting in excellent experimental findings. The high degree of dispersion, or scatter, occurred after projection aided image identification on the IoT. Statistical Segmented Model using Geometrical Features and Gaussian Naïve Bayes is explored in [22] each object component can be separated, and a robust scene model can be trained. The geometrical features that concatenate extreme point features, orientation, and polygon displacement values are then extracted from each component. Object detection is aided by these features, while scene understanding is aided by Gaussian Nave Bayes. The goal of scene understanding is to make machines seem like humans and to grasp visual scenes completely. Cognitive vision influences scene understanding, involving major areas such as computer vision, software engineering, and cognitive engineering [23]. Voice-directed autonomous navigation: This research proposes a voice-directed indoor navigation system for wheelchairs that includes sensors, laser scanners, microphones, and a ring of sonars. It uses the mouse navigation paradigm in an interior setting and can also navigate using voice commands delivered by the user. This greatly aids users with physical and visual impairments in navigating the site [24]. Sightless Literary Devices: There is a paucity of interaction with modern technologies among the world's blind population. This is due to two factors: 1) the restriction of such devices in certain places, and 2) the availability of such devices. 2) Owing to a lack of the necessary software Due to the diversity of the human population, each country has its own set of languages. It is vital to provide effective solutions to such difficulties as a result of the aforementioned elements [25]. Webel Mediatronics Ltd. is the industrial partner for the Bharati Braille Information System, which is a national program managed by IT Kheradpir. For blind individuals, the Braille coding scheme [1-7] is the most common means of reading and writing printed information. Due to differences in script systems, the sightless community in India and the developing globe suffers significant challenges in a) obtaining source printed reading material and b) connecting with the sighted community in writing [26]. Laser Wheelchair is a special wheelchair developed for the physically challenged visually impaired people by the Luca University in Sweden. This wheelchair is expected to be available in the markets as of 2022. Lulea University's laser wheelchair is structurally similar to that of traditional chairs. It comes with sturdy wheels and a durable leather seat which would provide comfort for those visually impaired users who need assistance for mobility [27]. The Finger Reader is a finger reader that uses an ultrasonic holography sensor to read the user's fingerprints. This project has produced excellent results; nonetheless, a study on it is currently ongoing in order to improve it [28]. OrCam My Eye: It is a miniature, portable, lightweight, and portable device the size of a finger. It can be held in the user's hand even magnetically attached to any pair of glasses. This allows users to use both hands for something while wearing the device. Moreover, it does not require an internet connection so that users can use it anywhere and anytime [29]. A system for detecting human posture was introduced. Three-dimensional angles of human arm movement could be determined using parameter data acquired from inertial sensors using a quaternion algorithm for data fusion [30]. 'Human posture recognition' [31] and 'Fascial expression recognition have emerged as a key study subject in both the computer-based intelligent video surveillance system and pattern recognition fields. The study assesses the use of human body posture mechanisms for computer interaction, offering numerous ways for accurate identification and facial expression. In facial recognition, the goal is to recognize faces in any image, extract facial characteristics (eyes and lips), and categorize them into six different moods sadness, fear, neutral, disgust, anger, happy. The training data is subjected to a number of filters and procedures before being classified using a Support Vector Machine (SVM) [32].

All the aids which are mentioned above have many limitations but one of the major limitations which we noticed, and which have to be taken care of is cost. All the products which are produced for blinds are made from high technology and hence they are very costly which make these products to be the out of the range. Due to this most people are not able to take this opportunity. And the second one is that most of the tools are highly advanced and need extra care while operating but due to lack of education most of the people are not able to use them and if somehow, they manage to use them they feel very uncomfortable operating the tools without complete knowledge, which also sometimes leads to accidents.

2 Methodology

This paper presents a scene detection system for visually impaired people. It can detect two scenes namely, i) Exercise and ii) Cooking. The block diagram of the system is shown in Fig.1.

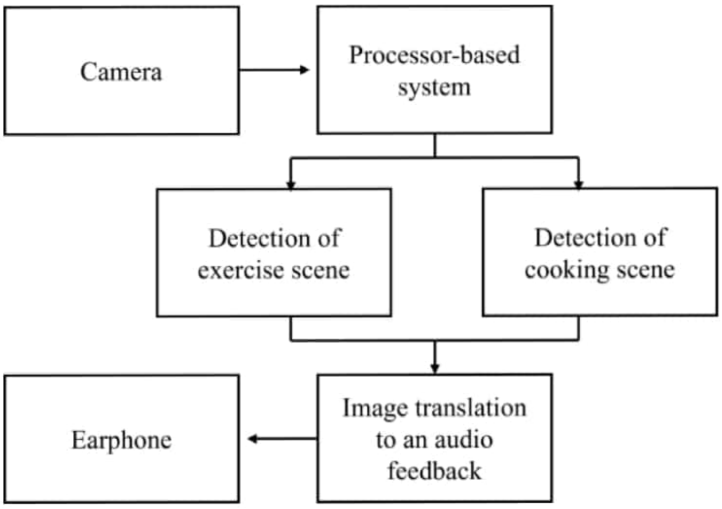

Fig.1. Block diagram of a complete system

It consists of a camera, a processor-based system, and an earphone. The camera captures the scene of the surrounding environment and provides it as an input to a processor-based system. An algorithm has been implemented to detect and interpret exercise and cooking scenes. The system converts the detected scene into audio feedback and conveys it to a visually impaired user via an earphone.

2.1 Dataset and Preprocessing

Total 8220 images were used for implementing this application. The distribution of images in the dataset is shown in Table 1 and Table 2.

Table 1: Details of the images in the dataset

Priority Scenes	No. of Positive Images
Exercise	3550
Cooking	1120

Table 2: Details of the negative images and total images in the dataset

No. of Negative Images	3550
No. of Total Images	8220

The exercise scene was defined as a blend of postures such as stretching of arms, legs, bending position, and yoga. The cooking scene was a typical kitchen scene having objects such as vessels, gas stoves, vegetables, and a pan as shown in Fig.2.

Fig. 2. Sample images of exercise and cooking dataset

The entire dataset was obtained from the internet [33]-[34]. All the images were resized to 200 x 200 pixels and then converted to grayscale. The dataset preprocessing operations are presented in Algorithm 1.

Algorithm 1: Pre-processing of Images

Input: Input image
Output: Pre-processed images
Initialization:
LOOP Process
1: **for** every image in the directory **do**
2: Resizing the images to 200 x 200 pixels
3: Gray scaling the image
4: **end for**
5: **return** pre-processed image

2.2 Feature Vector Compilation

Scale-Invariant Feature Transform (SIFT) was used to extract features from all images in the dataset as shown in Fig.3. SIFT is a computer vision algorithm that recognizes and matches local features in images. It finds critical locations and then provides descriptors that are utilized for object recognition, SIFT has the capability of uniform scaling and orientation. SIFT is invariant to scale, illumination changes, rotations, occlusions whereas other algorithms fail if they have scale variance, rotation, non-uniform illumination. Therefore, SIFT is superior to other machine learning feature extraction to get better accuracy. Features of every image were individually extracted by SIFT and appended into a CSV file. An Mx128 feature vector array is generated for every image. Where 'M' represents a number of key points extracted from an image. The size of the SIFT feature vector for the entire dataset was 2019256 x 128.

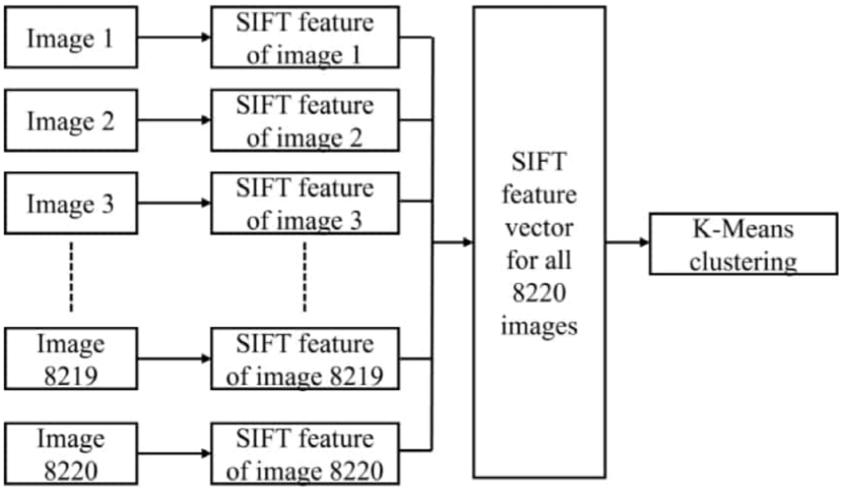

Fig.3. Training of large feature vector

The large size feature vector was divided into 5 clusters by using the K-Means clustering algorithm. The number of clusters was decided based on the elbow method. SIFT features of all images were predicted with the pre-trained K-Means model. Then the output histogram of each image was normalized and appended into a CSV file. The feature vector was reduced to 8220 x 5 after applying the K-Means clustering algorithm. The principal vector component analysis (PCA) was used to further reduce the dimensions of the feature vector to 8220 x 4. We measured the information contained in each bin, by comparing the maximum explained variance that was accommodated in the first 4 principal components. The information table in each principal component is shown in Table 3. After adding the first 4 principal components 99% of information was obtained. Therefore, 4 PCA components were selected based on maximum explained variance.

Table 3: The percentage of variance explained by principal components (PC) generated by PCA

PC1	PC2	PC3	PC4	PC5
78.16%	13.14%	4.77%	2.93%	0.98%

The overall procedure for dimensionality reduction and feature vector compilation is shown in Fig.4. and presented in Algorithm 2. The optimized feature vector was then provided as an input to an array of classifiers.

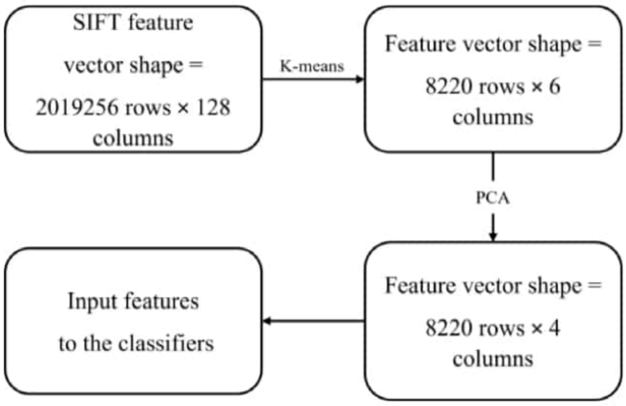

Fig.4. SIFT feature compilation and dimensionality reduction

Algorithm 2: Dimensionality Reduction

Input: Feature vector (2019256 rows × 128 columns)
Output: Optimized feature vector (8220 rows × 4 columns)

Initialization:
LOOP Process

```
for every image in a specified path do
    Perform prediction on pre-trained k-means clustering
    Normalized the data
    Append it to feature vector
end for
Standardization using standard scalar
Perform PCA with n_components=4
Perform PCA transform
return optimized feature vector (8220 rows × 4 columns)
```

2.3 Classification and Detection of Scene

This system uses an array of five classification algorithms namely, i) Decision Tree (DT) ii) Support Vector Machine (SVM) iii) K-Nearest Neighbour (K-NN) iv) Random Forest (RF) and v) Logistic Regression (LR) for classification and detection of desired scenes. Performance parameters such as training accuracy, testing accuracy, precision score, recall score, F1 score, were measured for the assessment of the detection.

The first classifier used was DT. It generates rules to forecast any unseen data.

$$E(s) = \sum_{i=1}^{c} - p_i \log_2 p_i \qquad (1)$$

Here p_i, stands for the probability of class i.

The second classifier used was SVM. It describes the best decision boundary for categorizing n-dimensional space into classes. In SVM for multiclass classification, we used two methods namely i) One-vs-rest (OVR), ii) One-vs-one (OVO). The equation for SVM is

$$\mathrm{L}(w) = \sum_{i=1} \max\big(0, 1 - b_i (w^s m_i + z)\big) + \lambda \parallel w \parallel_2^2 \qquad (2)$$

where w is the input and m_i, is the support vector. Coefficients b_i , z and λ are estimated from training data.

The K-NN was the third classifier used. It predicts the values of new data points based on feature similarity. The new data point will be assigned a value depending on how closely it matches the points in the training set. It is based on Euclidean distance measure and its equation is,

$$\mathrm{D} = \sqrt{(m_1 - m_2)^2 + (n_1 - n_2)^2} \qquad (3)$$

Where D is the calculated distance, m_1 and n_1, are x coordinates and y coordinates of the feature vector.

The Random Forest algorithm is the fourth classifier used. The dataset is divided into subsets by a random forest classifier. Every decision tree receives these subsets, and each one generates its own output.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (fi - bi)^2 \qquad (4)$$

Where N is the number of data points, fi is the value returned by the model, and bi is the actual value for data point i.

The last classifier used was LR. It is the supervised machine learning algorithm.

$$p = \frac{e^{a+bx}}{1+e^{a+bx}} \quad (5)$$

The probabilistic value ranges in the form 0 and 1. The outcomes are categorized in the form of 1 and 0.

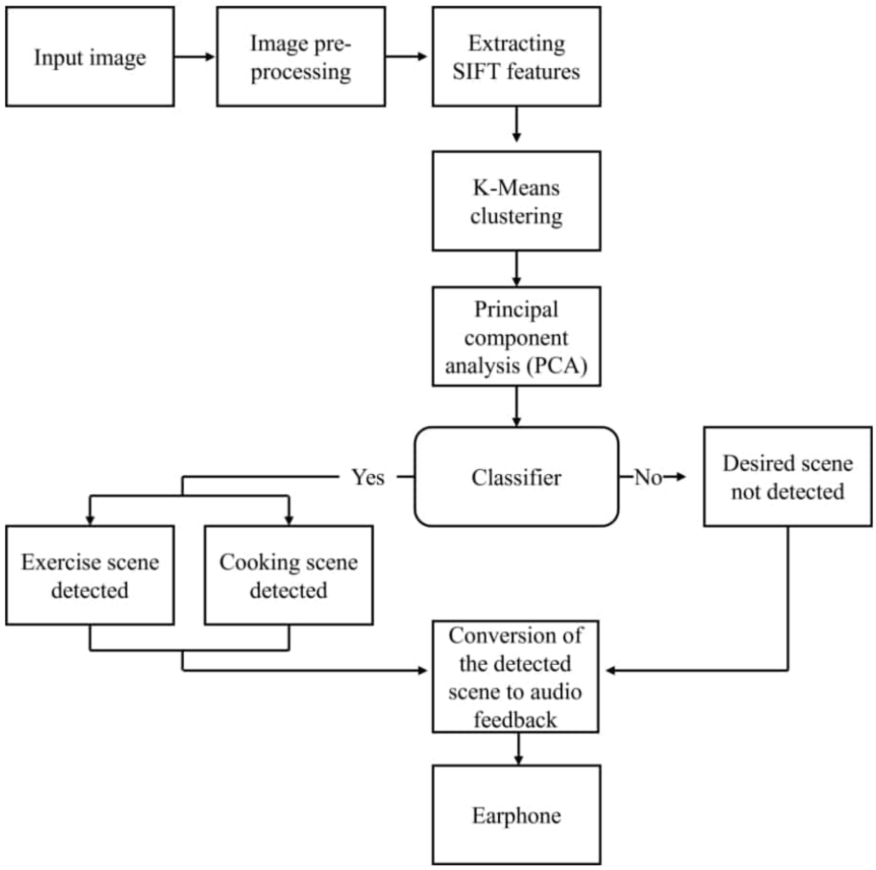

Fig.5. Overall flow diagram for scene detection and communication

The overall process for classification and detection of the exercise and cooking scene is described in Algorithm 3. The overall flow diagram of the proposed system is shown in Fig .5. The system classifies the detected scene and label it 0 if it is exercise scene. It assigns label 1, if the detected scene is cooking scene. The system generates spoken message "Exercise" or "Cooking" based on label 0 and 1 respectively.

Algorithm 3: Classification and Detection of the Scene

Input: Optimized feature vector (8220 rows × 4 columns)
Output: Classified Exercise and Cooking Scene
Initialize: min_accuracy = 75.91%
Procedure:
1: Fit the model with training data for the Random Forest classifier
2: Predict the test data using this pre-trained model
3: **if** (predicted class label ==0 && accuracy >= min_accuracy)
4: **return** Exercise Scene Detected
5: **else if** (predicted class label ==1 && accuracy >= min_accuracy)
6: **return** Cooking Scene Detected
7: **else if** (predicted class label ==2 && accuracy >= min_accuracy)
8: **return** Neither Exercise nor Cooking Scene
9: **else**
10: **return** Not able to classify the scene

3. Result

An array of five classifiers was used for the assessment of scene detection. The detailed performance analysis of the five classifiers is shown in Table 4. and Table 5.

Table 4: Training accuracy and Testing accuracy analysis for detection of exercise and cooking scene

Priority Classifiers	Training accuracy	Testing accuracy
Decision Tree	75.71%	75.91%
SVM – OVR Linear	78.68%	79.50%
SVM – OVR Polynomial	76.58%	77.98%
SVM – OVR RBF	79.80%	80.17%
SVM – OVO Linear	78.68%	79.50%
SVM – OVO Polynomial	76.58%	77.98%
SVM – OVO RBF	79.80%	80.17%
K-NN	84.48%	79.31%
Random Forest	92.92%	80.83%
Logistic Regression	78.90%	79.68%

Table 5: Precision score, Recall score, F1 score for exercise and cooking scene detection

Priority Classifiers	Precision score	Recall score	F1 score
Decision Tree	50.71%	58.29%	54.22%
SVM – OVR Linear	74.64%	67.13%	68.22%
SVM – OVR Polynomial	76.09%	63.85%	64.25%
SVM – OVR RBF	75.25%	68.18%	69.40%
SVM – OVO Linear	74.50%	67.13%	68.22%
SVM – OVO Polynomial	76.09%	63.85%	64.25%
SVM – OVO RBF	75.25%	68.18%	69.14%
K-NN	72.43%	70.09%	70.88%
Random Forest	75.36%	71.58%	72.80%
Logistic Regression	73.86%	68.34%	69.50%

Total 1000 test images (exercise and cooking) were used for the performance evaluation of classifiers. All classifiers accurately classified majority images. Among all classifiers, the Random Forest classifier provided the highest training and testing accuracy of 92.92% and 80.83% respectively. This model uses a Random Forest classifier for predicting the exercise and cooking scene.

4 Conclusion

This system helped in classifying an image into two possible scenes. The two distinct scenes are i) A scene involving some sort of exercise such as yoga, ii) A scene involving cooking. It is very important for everyone to recognize surroundings while walking or moving from one place to another. This project will help visually disabled people to sense what's happening near them.

The system accuracy reduces if there is more than one person at a time in the frame. The system performance is reduced when there is non-uniform illumination in the captured image.

This system can be extended to include activities such as swimming, playing cricket, and many other things if trained accordingly. This can be further improved with a Graphical user interface (GUI) for better understanding and can also be deployed in a kind of hardware such as a blind stick that will alert the user that there are people in front of you performing the following tasks.

Acknowledgemments

We express our sincere gratitude to the visually impaired participants in this study, orientation and mobility (O&M) experts and authorities at The Poona Blind Men's Association, Pune. The authors thank the La Fondation Dassault Systemes, India and Vishwakarma Institute of Technology Pune for providing support (IN-2019-056) to carry out this research work

References

1. Dey, Naiwrita; Paul, Ankita; Ghosh, Pritha; Mukherjee, Chandrama; De, Rahul; Dey, Sohini (2018). *[IEEE 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) - Coimbatore, India (2018.3.1-2018.3.3)] 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) - Ultrasonic Sensor Based Smart Blind Stick.*
2. Mennens, J.; Van Tichelen, L.; Francois, G.; Engelen, J.J. (1994). *Optical recognition of Braille writing using standard equipment. IEEE Transactions on Rehabilitation Engineering, 2(4), 207–212.*
3. Yokokohji, Y.; Yoshikawa, T. (1994). Bilateral control of master-slave manipulators for ideal kinesthetic coupling-formulation and experiment. IEEE Transactions on Robotics and Automation, 10(5), 605–620.
4. Munoz, Rai; Xuejian Rong, ; Tian, Yingli (2016). *[IEEE 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW) - Seattle, WA, USA (2016.7.11-2016.7.15)] 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW) - Depth-aware indoor staircase detection and recognition for the visually impaired.*
5. Reddy, E. J., Reddy, P. N., Maithreyi, G., Balaji, M. B. C., Dash, S. K., & Kumari, K. A. (2020). *Development and Analysis of Pothole detection and Alert based on NodeMCU. 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE).*
6. Saaid, M. F., Ismail, I., & Mohd Zikrul Hakim Noor. (2009). *Radio Frequency Identification Walking Stick (RFIWS): A device for the blind. 2009 5th International Colloquium on Signal Processing & Its Applications.*
7. Jain, Sambhav; Varsha, Sushanth D; Bhat, Vijetha N; Alamelu, J V (2019). *[IEEE 2019 International Conference on Communication and Signal Processing (ICCSP) - Chennai, India (2019.4.4-2019.4.6)] 2019 International Conference on Communication and Signal Processing (ICCSP) - Design and Implementation of the Smart Glove to Aid the Visually Impaired.*
8. Arvitha, M. V. S., Biradar, A. G., & Chandana, M. (2021). *Third Eye for Visually Challenged Using Echolocation Technology. 2021 International Conference on Emerging Smart Computing and Informatics (ESCI).*
9. Long Chen, ; Bao-long Guo, ; Wei Sun, (2010). *[IEEE 2010 Fourth International Conference on Genetic and Evolutionary Computing (ICGEC 2010) - Shenzhen (2010.12.13-2010.12.15)] 2010 Fourth International Conference on Genetic and Evolutionary Computing - Obstacle Detection System for Visually Impaired People Based on Stereo Vision.*
10. *Joe Louis Paul, I; Sasirekha, S; Mohanavalli, S; Jayashree, C; Moohana Priya, P; Monika, K (2019). [IEEE 2019 International Conference on Computational Intelligence in Data Science (ICCIDS) - Chennai, India (2019.2.21-2019.2.23)] 2019 International Conference on Computational Intelligence in Data Science (ICCIDS) - Smart Eye for Visually Impaired-An aid to help the blind people.*
11. Mennens, J.; Van Tichelen, L.; Francois, G.; Engelen, J.J. (1994). *Optical recognition of Braille writing using standard equipment. IEEE Transactions on Rehabilitation Engineering,*
12. Vaccarella, A.; Comparetti, M. D.; Enquobahrie, A.; Ferrigno, G.; De Momi, E. (2011). *[IEEE 2011 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society - Boston, MA (2011.08.30-2011.09.3)] 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society - Sensors management in robotic neurosurgery: The ROBOCAST project.*

13. Leporini, Barbara; Rosellini, Michele; Forgione, Nicola (2020). *Designing assistive technology for getting more independence for blind people when performing everyday tasks: an auditory-based tool as a case study. Journal of Ambient Intelligence and Humanized Computing,*
14. KARTHIK, A.; RAJA, V. KAARTHICK; PRABAKARAN, S. (2018). *[IEEE 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT) - Chennai, India (2018.2.15-2018.2.17)] 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT) - Voice Assistance for Visually Impaired People*
15. *Ng, S.S. (1991). [IEEE Annual International Conference of the IEEE Engineering in Medicine and Biology Society Volume 13: 1991 - Orlando, FL, USA (31 Oct.-3 Nov. 1991)] Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society Volume 13: 1991 - Sonic Electronic Guide For The Blind*
16. Khan, S., Nazir, S., & Khan, H. U. (2021). Analysis of Navigation Assistants for Blind and Visually Impaired People: A Systematic Review. IEEE Access, 9, 26712–26734.
17. Menikdiwela, M. P.; Dharmasena, K.M.I.S.; Abeykoon, A.M. Harsha S. (2013). *[IEEE 2013 International conference on Circuits, Controls and Communications (CCUBE) - BENGALURU, India (2013.12.27-2013.12.28)] 2013 International conference on Circuits, Controls and Communications (CCUBE) - Haptic based walking stick for visually impaired people.*
18. M. Ponnavaillo and V. P. Kumar, "The artificial eye," in IEEE Potentials, vol. 18, no. 5, pp. 33-35, Dec. 1999-Jan. 2000
19. *Akhter, S.; Mirsalahuddin, J.; Marquina, F.B.; Islam, S.; Sareen, S. (2011). [IEEE 2011 37th Annual Northeast Bioengineering Conference (NEBEC) - Troy, NY, USA (2011.04.1-2011.04.3)] 2011 IEEE 37th Annual Northeast Bioengineering Conference (NEBEC) - A Smartphone-based Haptic Vision Substitution system for the blind.*
20. Sathesh, A. "Computer Vision on IOT Based Patient Preference Management System." Journal of Trends in Computer Science and Smart Technology 2, no. 2 (2020): 68-77.
21. Jacob, I. Jeena, and P. Ebby Darney. "Design of Deep Learning Algorithm for IoT Application by Image based Recognition." Journal of ISMAC 3, no. 03 (2021): 276-290.
22. A. A. Rafique, A. Jalal and A. Ahmed, "Scene Understanding and Recognition: Statistical Segmented Model using Geometrical Features and Gaussian Naïve Bayes," *2019 International Conference on Applied and Engineering Mathematics (ICAEM)*, 2019, pp. 225-230
23. S. Aarthi and S. Chitrakala, "Scene understanding — A survey," *2017 International Conference on Computer, Communication and Signal Processing (ICCCSP)*, 2017, pp. 1-4, doi: 10.1109/ICCCSP.2017.7944094.
24. *Megalingam, Rajesh Kannan; Nair, Ramesh Nammily; Prakhya, Sai Manoj (2011). [IEEE Electronic Systems Technology (Wireless VITAE) - Chennai, India (2011.02.28-2011.03.3)] 2011 2nd International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE) - Automated voice based home navigation system for the elderly and the physically challenged.*
25. *Basu, A.; Dutta, P.; Roy, S.; Banerjee, S. (1998). A PC-based Braille library system for the sightless. , 6(1), 60–65. doi:10.1109/86.662621*
26. *Chaves, D. R.; Peixoto, I. L.; Lima, A.C.O.; Vieira, M. F.; de Araujo, C. J. (2009). [IEEE 2009 IEEE International Workshop on Medical Measurements and Applications (MeMeA) - Cetraro, Italy (2009.05.29-2009.05.30)] 2009 IEEE International Workshop on Medical Measurements and Applications - Microtuators of SMA for Braille display system.*
27. *Trieu, Hoang T.; Nguyen, Hung T.; Willey, Keith (2008). [IEEE 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - Vancouver, BC (2008.08.20-2008.08.25)] 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - Shared control strategies for obstacle avoidance tasks in an intelligent wheelchair. , (), 4254–4257. doi:10.1109/iembs.2008.4650149.*
28. *Kowshik, S.; Gautam, V.R; Suganthi, K. (2019). [IEEE 2019 11th International Conference on Advanced Computing (ICoAC) - Chennai, India (2019.12.18-2019.12.20)] 2019 11th International Conference on Advanced Computing (ICoAC) - Assistance For Visually Impaired Using Finger-Tip Text Reader Using Machine Learning.*
29. *O. Younis, W. Al-Nuaimy, F. Rowe and M. H. Alomari, "Real-time Detection of Wearable Camera Motion Using Optical Flow," 2018 IEEE Congress on Evolutionary Computation (CEC), 2018, pp. 1-6*

30. W. Ni et al., "Human posture detection based on human body communication with muti-carriers modulation," 2016 39th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2016, pp. 273-276
31. F. -. C. Adochiei, I. R. Adochiei, R. Ciucu, G. Pietroiu-Andruseac, F. C. Argatu and N. Jula, "Design and Implementation of a Body Posture Detection System," 2019 E-Health and Bioengineering Conference (EHB), 2019, pp. 1-4.
32. C. Jain, K. Sawant, M. Rehman and R. Kumar, "Emotion Detection and Characterization using Facial Features," 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE), 2018, pp. 1-6.
33. Shruti Saxena, *Yoga Pose Image Classification Dataset,* Apr. 3, 2021. Accessed on: Nov. 25, 2021. [Online]. Available: https://www.kaggle.com/shrutisaxena/yoga-pose-image-classification-dataset
34. *"Cooking using pan,"* Oct. 20,2021. Accessed on: Nov. 25, 2021.[Online].Available: https://www.google.co.in/search?q=cooking+using+pan&hl=en&tbm=isch&source