J. Integr. Sci. Technol. 2025, 13(1), 1001

Journal of Integrated
SCIENCE & TECHNOLOGY

Chromatic surveillance: Advancing color recognition in security systems

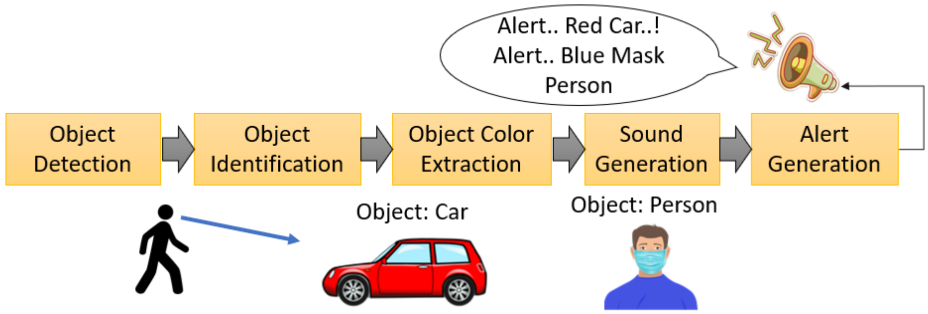
Harshad Lokhande, Sanjay Ganorkar*

Department of Electronics and Telecommunication Engineering, Sinhgad College of Engineering, Savitribai Phule Pune University, Pune. India.

Received on: 14-Feb-2024, Accepted and Published on: 03-Sep-2024

Article

ABSTRACT



In response to safety and security, accurate detection and alert systems for object identification in public spaces are crucial for enhancing surveillance for the visually impaired. Current object recognition algorithms face challenges due to lighting and spatial variations, resulting in inconsistent performance and lack of tailored features like color-based alerts for identification. Our study introduces an innovative algorithm that extracts invariant facial features through a unique segmentation method and K-Means clustering, significantly improving the reliability of face mask detection and enabling precise color recognition for visually impaired individuals through sRGB values and text-to-speech technology. Our research extends to evaluating our algorithm against leading object detection models like MobileNet SSD and Faster R-CNN, achieving a high weighted-average F1-score of 0.9131, showcasing the effectiveness of our approach. Future investigations will focus on enhancing color identification algorithms and expanding auditory signals to cover a wider range of objects, ultimately providing visually impaired individuals with a more comprehensive understanding of their surroundings for safer navigation.

Keywords: K-means, color, color web-space, object color detection, masks, sRGB, MobileNet SSD

INTRODUCTION

The necessity for color detection on identified objects for blind people is underscored by several challenges and opportunities highlighted across various research contexts. Firstly, the global prevalence of visual impairment, affecting approximately 2.2 billion individuals, necessitates innovative solutions to aid in their navigation and interaction with the environment.[1] Blind individuals often encounter difficulties in performing daily activities, including identifying the color of objects, which is crucial for tasks such as recognizing currency and avoiding fraud. This challenge extends to matching the color of clothes, food, and surrounding items, further emphasizing the importance of color detection for enhancing the independence and quality of life of visually impaired people.[2]

Research has shown that object identification technologies have significantly advanced, offering potential for visual substitution through the development of electronic devices that can recognize and provide audio output of objects, including their color.[3] These technologies leverage convolutional neural networks and other algorithms to detect objects and their attributes in real-time, integrating audio devices to communicate this information to the user.[4] Moreover, smart object detection systems have been proposed to ensure a safe living environment for visually impaired individuals by notifying them of detected objects and potential hazards. To address the specific need for color detection, systems have been designed to transform visual information into audio output, including the color of money, thereby preventing undesirable incidents such as fraud.[5] Additionally, color recognizers have been developed to assist visually disabled people in identifying the color of various objects, using sensors to read color values and convey the information through audio output.[6] These innovations highlight the critical role of color detection in improving the mobility, safety, and daily life experiences of blind and visually impaired individuals.[7]

The progress in technological advancements aimed at aiding visually impaired individuals has been notable, particularly in the

*Corresponding Author: Harshad Lokhande, Sanjay Ganorkar
Email: lokhande.harshad@gmail.com, srganorkar.scoe@sinhgad.edu

Cite as: *J. Integr. Sci. Technol., 2025, 13(1), 1001.*
URN:NBN:sciencein.jist.2025.v13.1001
DOI: 10.62110/sciencein.jist.2025.v13.1001

domains of color recognition and object detection. These developments involve the use of voice alerts to provide these individuals with improved awareness of their surroundings. Scholars are primarily focused on the creation of systems that harness sophisticated deep learning algorithms, including the convolutional neural network (CNN) and the You Only Look Once (YOLO) technique, to scrutinize environments and identify objects. Subsequently, this information is conveyed to the user in real-time through text-to-speech (TTS) technology for assistance.[8]

Even all are following preventative measures, still there is chance to get in contact with COVID affected people. So, there is need to develop a system which will alert the person, if mask is there or not. For blind person, knowing the masked faces and alerting if not, will provide the additional preventative measure.

Table 1 shows comparison between different trivial face detection methods. Face detection is a simple issue. Objectives include: detecting faces in images for mask detection, recognizing object colors in a bounding box system, and extracting colored output for sound alerts.

There are various techniques to identify the mask and unmasked faces based on image processing and deep learning. Though recent deep learning methods. are giving promising results in face detection, the deployment of those algorithms on mobile devices is crucial task.

The position of face in video is uncertain as the center of camera may move around the scene. The position of camera is varying like frontal, tilted, or outline. Our motivation is to detect the face covered partially with gadgets, or due to presence of mask. Moreover, the lighting conditions and camera settings have the potential to influence the visual representation of a person's face. In order to address these challenges, scholars have introduced a range of algorithms and methodologies. This work is based upon important aspects to understand the basic idea behind face detection, mask detection and alert generation.

The scope of work further expanded as, review on exisiting systems and technologies. Later the methodlogy exaplins about the face dettection technique and color extraction of detected object. Then the alogirthm explain about voice prompt on identified color. Finally results are compared with state-of-art methods to align the future scope work.

LITERATURE REVIEW

The face detection moves around various challenges like critical background, illumination, skin color, distance, face occlusion, odd expressions and orientation.[9] The distance between facial landmarks with fix length is crucial parameter.[10] So, considering the length and the patterns in landmarks, the face can be detected. But occlusion may change the values in landmarks. In 2001, Voila -Jones method, changed the way to look around face detection problem.[11] The features are used in HAAR Cascade way with AdBoost learning algorithm for overcoming on issues like classic statistical learning approach. Haar Feature Selection is based on analogous characteristics found in human faces, such as the positioning and dimensions of facial elements like the eyes, mouth, nose bridge, and the directional gradients of pixel intensities. A series of 38 cascaded classifiers are utilized to derive a grand total of 6061 features from each frontal face. This technique improved the speed of detection and accuracy. An an algorithm for locally linear regression is utilized for the purpose of detecting facial features in a manner that is not influenced by the orientation or pose of the face.[12] Thus, the issue of variation in appearance of face in front of camera are identified. The rules governing the relationships between facial characteristics are applied in face localization, but their accuracy is compromised by the rules used in face detection.

The face detection method is categorized into four distinct groups.[13] The incorporation of these technologies into a wearable or portable device, like a cap with an embedded camera or a smart stick with a Raspberry Pi, presents a user-friendly interface for individuals with visual impairments. Considering above comparative analysis, the face detection with appearance-based model in combination with feature-based classifier may give good accuracy in detection. The proposed work briefs about face detection, color identification and conversion of alert into sound.

The field of object detection, audio alert, and assistive technologies for the visually impaired has seen significant advancements, leveraging deep learning and machine learning algorithms to enhance the autonomy and safety of visually impaired individuals. Subramanian et al. and another study highlight the application of these technologies in creating visual assistants that use hardware components like Raspberry Pi to detect objects and obstacles, providing audio aid to visually impaired people.[14] A deep learning-based object detection system employing the YOLOV7 algorithm and text-to-speech (TTS) technology for voice-guided object recognition, aims to empower visually challenged individuals to identify objects independently.[15]

Further, an assistive system based on the TensorFlow object detection model and Google Speech’s model has been developed to improve the mobility and safety of visually impaired people in unfamiliar environments by converting object information into speech.[16] A device converting visual information into auditory feedback using a Raspberry Pi and a tactile device has been tested for its effectiveness in scene understanding, demonstrating its utility in real-world applications.[17] The use of the YOLO V4 algorithm and pyttsx3 for detecting objects and converting labels to

Table 1. Comparison between Face Detection Methods

Based on Feature	Based on Knowledge	based on Template	Based on Appearance
• Locate faces by extracting features • classifier differentiates between facial and non-facial regions • This approach gives accuracy upto 94%	• based on set of rules • like face must have nose, etc. • Issue: model is too generalized. • and Unable to find multiple faces in image	• It uses pre-defined template to detect faces by correlation in templates and an input image. • Simple to implement but restricts to template size.	• Relies on statistical analysis. • Eigenface, distribution based and Neural networks-based techniques. • Easy to implement • Accurate

speech has been explored to enable visually impaired users to comprehend their surroundings independently.[18] Moreover, an improved model of YOLOv5, featuring an enhanced decoupled head and improved loss and activation functions, has been proposed for real-time object detection, utilizing a text-to-speech conversion library to inform visually impaired people about detected objects.[19] By employing musical notes to precisely communicate the depth and location of objects, demonstrating high accuracy and effectiveness in outdoor scenes.[20]

Thus, computaional aware approach using sound to communicate through fast object detection can give improved solution with optmization in latency These studies collectively underscore the potential of integrating object detection, audio alerts, and assistive technologies to significantly improve the quality of life for visually impaired individuals.

METHODOLOGY

A. The face detection

Detecting facial landmarks involves a dual-step procedure. The initial step is to pinpoint the location of the face within the image. Subsequently, the process entails identifying the essential facial features on the region of interest of the face.

The best method is to use pre-trained HOG + Linear SVM object detector specifically for the task of face detection. Compute a Histogram of Oriented Gradients (HOG) through the process of calculating the gradient image both in the 'row' and 'col' dimensions, followed by the computation of gradient histograms, subsequent normalization across blocks, and finally, the flattening of the data into a feature vector. This methodological approach facilitates the extraction of relevant information from the image data by capturing the distribution of gradient orientations within localized regions, enabling a robust representation of visual features for various computer vision tasks.

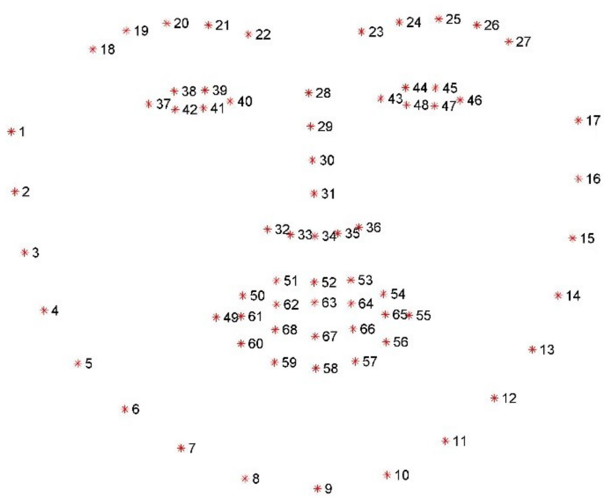


Figure 1. Visualizing the 68 facial landmark coordinates[19]

The face detection is essential step for knowing the ROI (Region of Interset) and to draw the Bounding Box around it as (Cx,Cy,W,H). Once face is detected next step is to find facial landmarks. There are two major approaches to find the facial landmarks with 68 points and 5 points around the face.[21]

Algorithm 1 Face Mask Detection

```
procedure PREPROCESSIMAGE (image)
    resized_image ← RESIZE (image,224,224)
    normalized_image ← NORMALIZE(resized_image, [1,-1])
    blob ← CREATEBLOB (normalized_image)
    return blob
end procedure
procedure LOCATEFACE(image)
    face_box ← DETECTBOUNDINGBOX(image)
    face_roi← image[face_box]
    return face_roi
end procedure
procedure CLASSIFYMASKSTATUS (face_roi,model)
    mask_label← PREDICT(model,face_roi)
    return mask_label
end procedure
pretrainedmodel ← LOADMODEL(…)
image LOADIMAGE(…)
blob ← PREPROCESSIMAGE(image)
faceroi ← LOCATEFACE(blob)
maskstatus←CLASSIFYMASKSTATUS(faceroi,pretrainedmodel)
print maskstatus
```

The visualization of the 68 coordinates' indexes can be observed in Figure 1, providing a clear representation of their spatial distribution. Within the realm of facial analysis, the utilization of the pre-trained facial landmark detector embedded in the dlib library holds significant importance. This detector functions by accurately estimating the precise positioning of 68 (x, y)-coordinates, which correspond to specific facial features and structures present on an individual's face 11. There are a variety of facial landmark detectors, but all methods essentially try to localize and label the following facial regions like Right Eyebrow, Left Eyebrow, Mouth, Nosal part, Jaw, Left Eye, Right Eye.[22]

The algorithm 1, analyzes if a person is wearing a face mask. It involves three main steps: Preprocessing the image, locating the face, and classifying the mask status. Preprocessing includes resizing, normalizing, and transforming the image. Locating the face involves detecting a bounding box and extracting the region of interest. Classifying the mask status predicts if the face is wearing a mask. The algorithm uses a pretrained model to determine and display the mask status of the face in the image, following object detection and classification typical in computer vision tasks.

The COVID19 Asian Face Mask Dataset (CAFMD) is used to dtect the masked faces with diverse array of image types showcasing various people in different scenarios, ranging from individuals in isolation to various scenes.[23] In light of the literature reviewed, the face mask detection methods MobileNet SSD and Faster RCNN are utilized. MobileNet SSD, renowned for its lightweight design, utilizes depth-wise separable convolutions to reduce computational costs and model size, making it suitable for mobile and edge devices.[24] Conversely, Faster R-CNN excels in accuracy due to its region proposal network (RPN) but demands higher computational resources, making it more suitable for applications prioritizing precision over speed.[25] So, MobileNet SSD is considered for high-speed detection with acceptable accuracy, while Faster R-CNN is preferred for scenarios emphasizing precise detection and possessing sufficient computational resources.

B. Identifying the color

The color identification is critical task. From an image masked person is detected and the ROI is marked as Bounding Box. The bounding box decides the presence of object if IoU (Intersection over Union) is greater than 0.65. The color identification works in 2 stages as shown in Figure 4.

1. Detecting Dominant Color
2. Extracting the Color Name

Algorithm 2 Color Detection and Conversion

procedure DetectAndConvertColor*(input_cropped_ROI)*
    dominant_color ← FindDominantColor *(input_cropped_ROI)*
if sRGBValueMatches(dominant_color) **then**
    css_color_value ← ConvertToCSSColorSpace *(dominant_color)*
color_name ← MatchWithCSSKey*(css_color_value)*
    return *color_name*
else
    return "No Match"
end if
end procedure

The algorithm 2, shows all the steps involved in detecting and extracting the color of an identified object from an image

Detecting Dominant Color

Once the detected object id located as ROI, the image is cropped in size of from a two-dimensional to one-dimensional data set. All calculations are done in floating point 64-bit values. The color space is mapped into 25 distinct values. Hence the K-Means cluster is used to segment the different regions from the cropped image. Iteration parameters defined for stopping condition of K-Means. If an error is less than 0.5 EPS or the number of iterations exceeds MAX 10, then K-Means is converged and it will stop to allocate the segmented regions. The detected value of the dominant color is stored in the form of a tuple. For eg. [125,125,0] = [B, G, R].

Extracting Color Name

The extracted color values are treated via web color space as sRGB format. sRGB values for illuminated objects of scene are simply combinations of CIE XYZ valuee. And these values are computed using derived relationship as follows given by equation 1.

$$\begin{bmatrix} Rs\ RGB \\ Gs\ RGB \\ Bs\ RGB \end{bmatrix} = \begin{pmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.09692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{pmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \qquad (1)$$

These values are used as "color space in web" mentioned in Figure 2, for representing color numerically in terms of three or more coordinates. Later this sRGB is converted into HSL format (Hue – Saturation- Lightness) using hsl() method. The colors are adopted from newest CSS-3 color module. The CSS-3 supports variable opacity of colors by allowing specific rgba() and hsla() constructs. If a (opacity parameter) = 0 image is completely transparent, A = 1 is is completely visible. In CSS-3 color scheme, the color value is used as keyword. The following image is showing basic color formats in CSS-3.

Color names and sRGB values

Named	Numeric	Color name	Hex rgb	Decimal
		black	#000000	0,0,0
		silver	#C0C0C0	192,192,192
		gray	#808080	128,128,128
		white	#FFFFFF	255,255,255
		maroon	#800000	128,0,0
		red	#FF0000	255,0,0
		purple	#800080	128,0,128
		fuchsia	#FF00FF	255,0,255
		green	#008000	0,128,0
		lime	#00FF00	0,255,0
		olive	#808000	128,128,0
		yellow	#FFFF00	255,255,0

Figure 2. The basic Color Table in sRGB – CSS3

The extracted dominant pixels are used as input for getting RGB value (in Hex value). This HEX value is converted to RGB triplet and passed as key to extract the color from web-color space. To identify a color that is not available in the list of dominant colors detected using the logic provided, you can follow these steps mentioned in below Figure 3,

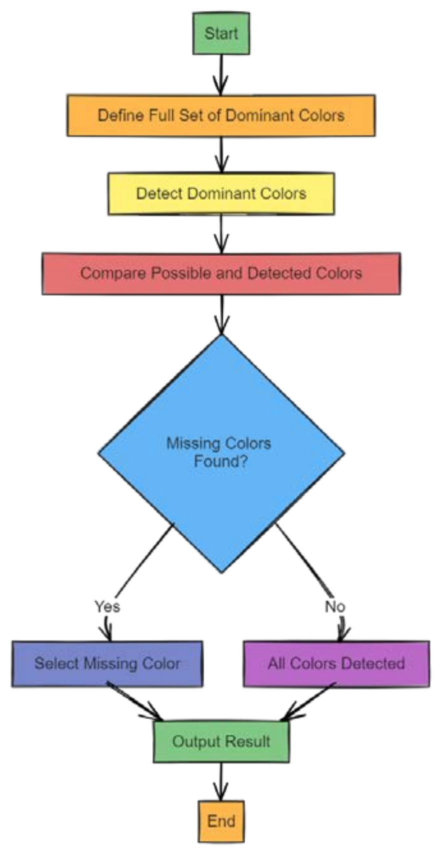


Figure 3. The graph diagram representing the algorithm to identify a missing color

C. The sound alert mechanism

Generating an alert if a person is masked or unmasked is accomplished with the “winsound” as sound alert generation module.[14]

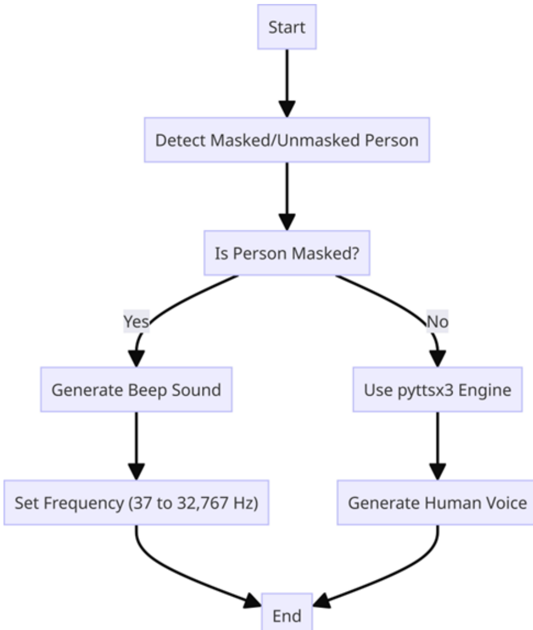


Figure 4. Flowchart of Sound Alert Mechanism

This module generates sound in the form of “beep” or also add human voice using the pyTTS engine.[26] The frequency must be in the range 37 through 32,767 hertz in order to have clear audible signal. The figure 4, shows flowchart for sound alert.

The “pyttsX3” engine is applied for human based voice annotations. The pyTTS engine is used to generate a sound instance. The driver is responsible for loading speech engine. This generates sound alert in either “beep” or in voice based given text message.[27] The alert will run for the time allotted in “RUN Instance”.

Results and Discussions

The training process relies on Stochastic Gradient Descent (SGD) with specific parameters for learning rate (α = 10-3), momentum (β = 0.7), and number of epochs (250). Utilizing the NVIDIA RTX 2060, the training involved a dataset split into train and test sets containing 945 masked and 690 unmasked labels across 700 images, with pretrained weights used for transfer learning and feature extraction using VGG16 architecture to obtain a feature vector of size 25,088 from the output of 512 x 7 x 7 before the final POOLING layer, necessitating setting N to the total number of images in the training dataset.[28]

Figure 5. Detecting K-Means Clusters (From left a. original image, b. K=2 , c. K =3)

Above figure 5 shows, detected the ROI and finding the clusters to identify the dominant colour. In the example shown below, the RGB images is converted to grey scale. Then K- means is applied with K=2 and K= 3. Once the regions are clustered, the extracted region converted to RGB and its dominant color value is extracted. It shows the steps of a K-means clustering algorithm applied to an image for segmentation purposes. The image on the left is the original grayscale input, while the subsequent images represent various stages of K-means clustering.

Initial Image (A): A grayscale image where intensity values range from 0 (black) to 255 (white).

- **Initial Centroids:** The algorithm chooses k initial centroids (either randomly or based on some heuristic).
- **Assignment Step (B):** Each pixel in the image is assigned to the nearest centroid. This step can produce an image where each pixel is colored according to the centroid it's closest to, often resulting in a segmented image where each segment corresponds to one of the k clusters.[29]
- **Update Step (C and onwards):** The centroids are recalculated as the mean of all pixels belonging to their cluster. The assignment step is then repeated with the new centroids, and this process iterates until convergence — that is, until the centroids no longer change significantly between iterations, or the assignments of pixels to centroids remain stable.

Figure 6. Multiple face mask detection on Faster RCN

The results from Faster RCNN show high confidence scores (95% to 99%) for "masked" classifications, indicating the model's robust ability to detect masks on faces as shown in figure 6. However, one instance labeled as "unmasked" received a lower confidence score of 68%, suggesting uncertainty possibly due to factors like partial visibility of the face or ambiguous features.[30]

The MobileNet SSD model demonstrates strong performance in detecting face masks in this particular sample. It exhibits resilience when faced with various types of masks and different orientations of faces, consistently providing confident predictions. This quality makes it suitable for applications requiring real-time processing, where both speed and accuracy are crucial.[31]

Figure 7. Multiple face mask detection on MobileNet SSD

As shown in Figure 7, the detection system with Mobilenet SSD accurately identifies various mask styles, such as surgical, N95, cloth, and patterned masks, even when designs are intricate or resemble faces like skulls, with confidence scores generally above 80% and often reaching 99% or 100%. Factors like visible mask parts, color contrast with skin, or patterns on masks may influence the confidence score variations, with some cases at 71% possibly due to mask positioning, unique patterns, or facial accessories affecting accuracy.[32]

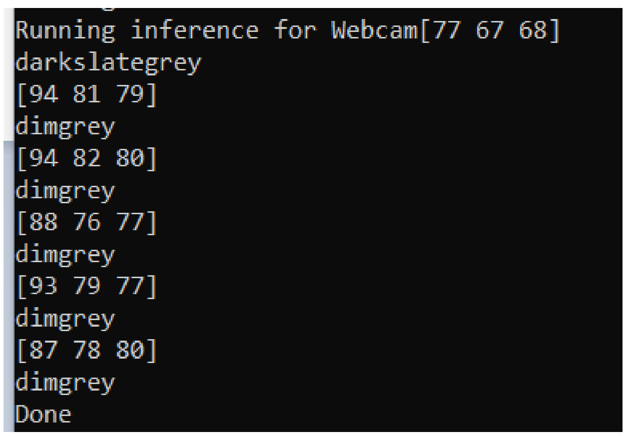


Figure 8. Extraction of dominant color value

The algorithm takes input data from a webcam, likely in the form of pixel values within an image. The figure 8 shows dominant colors from frame captured. Here each pixel has an RGB value associated with it, which denotes its color. The algorithm analyzes the RGB values and may compare them with a predefined set of colors to find the closest match.[33] This could be done using various methods such as nearest neighbor search in the RGB color space. Once the closest match is found, the algorithm assigns the predefined name of that color (like "darkslategrey" or "dimgrey") to the detected color.

Based on experimental findings, the detailed prediction outcomes are compared with the ground truth labels to compute the confusion matrix and F1 score. This process facilitates the assessment of raw prediction results against actual data, thus aiding in enhancing accuracy and expanding the scope of future work.

Table 2. Confusion Matrix evaluation chart (T: Target, O: Output)

T / O	Masked	Unmasked	Unknown	Sum
Masked	375	5	10	390
	48.51 %	0.65 %	1.29 %	96.15 %
Unmasked	4	310	7	321
	0.52 %	40.10 %	0.91 %	96.57 %
Unknown	50	8	4	62
	6.47 %	1.03 %	0.52 %	6.45 %
Sum	429	323	21	689 / 773
	87.41 %	95.98 %	19.05 %	89.13 %

The model demonstrates high precision in both Masked and Unmasked classes, with low misclassification rates. However as shown in Table 2 the confusion matrix, it struggles with correctly identifying the Unknown class, as evidenced by 50 instances misclassified as Masked and 8 as Unmasked.

An analysis of classification metrics reveals insights into the model's predictive accuracy and robustness. Both Masked and Unmasked classes exhibit high precision (0.9615 and 0.9657) and recall (0.8741 and 0.9598), indicating strong performance with minimal false positives or negatives. In contrast, the Unknown class shows significantly lower precision (0.0645) and recall (0.1905), leading to a low F1-score (0.0964), highlighting the need for improvement. Despite a high overall accuracy (0.8913) and relatively low misclassification rate (0.1087), the macro-average F1-score (0.6583) is moderate, primarily impacted by the Unknown class. However, the weighted-average F1-score (0.9131), which considers support, remains high.

Conclusion

The identification of object and alerting audio-visually within a complex environment to blind person is a pivotal challenge. This process is operationalized through a bifurcated approach, initially recognizing the object and extracting the color parameters from located Region of Interest (ROI). The algorithm proposed herein employs a feature extraction robust to changes in scale and luminance, compared over MobileNet SSD and Faster-RCNN. Moreover, it incorporates a clustering mechanism via K-Means to amalgamate segmented regions and extracting color of detected object with CIE values. The preeminent hue of the isolated segment is transposed to standard RGB (sRGB) color space, enabling the determination of the mask's color within a defined color spectrum with 98.13 % accuracy at 0.9131 F1-score. Subsequently, this chromatic data is aurally conveyed to visually impaired individuals as an auditory cue with tts audio engines, thereby signaling the presence of individuals wearing masks. The work delimits scope to only bright images only, the color detection can be increased through preprocessing techniques with Histogram Equalization. Further detection accuracy can be increased by complex models like YOLO and multimodal detection.

Acknowledgments

The Authors thank the Research Center at Sinhgad College of Engineering, Savitribai Phule Pune University and School of Computing, MIT ADT University, for their essential hardware,

software, and data collection support, which was crucial to the success of our research.

Conflict of Interest

The authors do not have conflict of interest for publication of this work as this study was not supported by any funding agency.

References

1. F. Sandid, A. Douik. Robust color texture descriptor for material recognition. *Pattern Recognit Lett* **2016**, 80, 15–23.
2. I. Abu Doush, S. AL-Btoush. Currency recognition using a smartphone: Comparison between color SIFT and gray scale SIFT algorithms. *Journal of King Saud University* - Computer and Information Sciences **2017**, 29 (4), 484–492.
3. H.N. Lokhande, S.R. Ganorkar. Challenges in Scene Interpretation for Video Surveillance in *Test-Engineering and Management*, **2020**, vol. 83, no. 2, pp. 2168–2276
4. R. Jadhav, D.S. Anand, A.K. Gupta, et al. Real-Time Object Detection for Visually Challenged. In *Machine Learning and Information Processing*; Springer Singapore, **2021**; pp 281–296.
5. P. Zhao. Robust wood species recognition using variable color information. *Optik - International Journal for Light and Electron Optics* **2013**, 124 (17), 2833–2836.
6. J. Zhu, J. Yu, C. Wang, F.Z. Li. Object recognition via contextual color attention. *J Vis Commun Image Represent* **2015**, 27, 44–56.
7. R. Gong, Q. Wang, X. Shao, J. Liu. A color calibration method between different digital cameras. *Optik (Stuttg)* **2016**, 127 (6), 3281–3285.
8. A. Nadeem, M. Ashraf, K. Rizwan, et al. A Novel Integration of Face-Recognition Algorithms with a Soft Voting Scheme for Efficiently Tracking Missing Person in Challenging Large-Gathering Scenarios. *Sensors* **2022**, 22 (3).
9. B. Shah, S. Shah, P. Shah, A. Shah. Survey on Object Detection, Distance Estimation and Navigation Systems for Blind People. In *Machine Learning for Predictive Analysis*; Springer Singapore, **2020**; pp 463–472.
10. Y. Liu, D. Zhang, G. Lu, W.Y. Ma. A survey of content-based image retrieval with high-level semantics. *Pattern Recognit* **2007**, 40 (1), 262–282.
11. E. Rachmawati, I. Supriana, M.L. Khodra, F. Firdaus. Integrating semantic features in fruit recognition based on perceptual color and semantic template. *Information Processing in* Agriculture **2022**, 9 (2), 316–334.
12. X. Chai, S. Shan, X. Chen, W. Gao. Locally Linear Regression for Pose-Invariant Face Recognition. *IEEE Transactions on Image Processing* **2007**, 16 (7), 1716–1725.
13. M. Alduhayyem, M. M. Alnfiai, N. Almalki, et al. IoT-Driven Optimal Lightweight RetinaNet-Based Object Detection for Visually Impaired People. *Computer Systems Science and Engineering* **2023**, 46 (1), 475–489.
14. H. Takizawa, S. Yamaguchi, M. Aoyagi, N. Ezaki, S. Mizuno. Kinect cane: an assistive system for the visually impaired based on the concept of object recognition aid. *Pers Ubiquitous Comput* **2015**, 19 (5–6), 955–965.
15. B. Uttl, P. Graf, P. Santacruz. Object color affects identification and repetition priming. *Scand J Psychol* **2006**, 47 (5), 313–325.
16. J. Shen, Y. Chen, H. Sawada. A Wearable Assistive Device for Blind Pedestrians Using Real-Time Object Detection and Tactile Presentation. *Sensors* **2022**, 22 (12), 4537.
17. A. Jiang, W. Fang, J. Liu, et al. The effect of colour environments on visual tracking and visual strain during short-term simulation of three gravity states. *Appl Ergon* **2023**, 110.
18. F. Andika, J. Kustija. Nominal of Money and Colour Detector for the Blind People. *IOP Conf Ser Mater Sci Eng* **2018**, 384, 12023.
19. P. Xia, S. Wang, F. Gong, W. Cao, Y. Zhao. Rapid recognition method of red brick content in recycled brick-concrete aggregates and powder based on color segmentation. *Journal of Building Engineering* **2024**, 84, 108633.
20. Y.C. Wong, J.A. Lai, S.S.S. Ranjit, A.R. Syafeeza, N.A. Hamid. Convolutional Neural Network for Object Detection System for Blind People. **2019**, 11, 1–6.
21. J. Madake, S. Bhatlawande, A. Solanke, S. Shilaskar. A Qualitative and Quantitative Analysis of Research in Mobility Technologies for Visually Impaired People. *IEEE Access* **2023**, 11, 82496–82520.
22. G. Burel, D. Carel. Detection and localization of faces on digital images. *Pattern Recognit Lett* **1994**, 15 (10), 963–967.
23. B. Mocanu, R. Tapu, T. Zaharia. When Ultrasonic Sensors and Computer Vision Join Forces for Efficient Obstacle Detection and Recognition. *Sensors* **2016**, 16 (11), 1807.
24. X. Chen, J. Xu, Z. Yu. A 68-mw 2.2 Tops/w Low Bit Width and Multiplierless DCNN Object Detection Processor for Visually Impaired People. *IEEE Transactions on Circuits and Systems for Video Technology* **2019**, 29 (11), 3444–3453.
25. H. Fu, H. Ma, G. Wang, X. Zhang, Y. Zhang. MCFF-CNN: Multiscale comprehensive feature fusion convolutional neural network for vehicle color recognition based on residual learning. *Neurocomputing* **2020**, 395, 178–187.
26. C. Singh, J. Singh. Geometrically invariant color, shape and texture features for object recognition using multiple kernel learning classification approach. *Inf Sci (N Y)* **2019**, 484, 135–152.
27. E. Wnuk, A. Verkerk, S.C. Levinson, A. Majid. Color technology is not necessary for rich and efficient color language. *Cognition* **2022**, 229, 105223.
28. A.B. Gunduz, B. Taskin, A.G. Yavuz, M.E. Karsligil. A better way of extracting dominant colors using salient objects with semantic segmentation. *Eng Appl Artif Intell* **2021**, 100, 104204.
29. A. Amelio, G. Bonifazi, E. Corradini, et al. Defining a deep neural network ensemble for identifying fabric colors. *Appl Soft Comput* **2022**, 130, 109687.
30. Y. Oishi, H. Nagasawa, K. Hirayama, K. Suzuki. Neural bases of color-specific semantic loss: Two cases of object-color knowledge impairment. *Cortex* **2021**, 141, 211–223.
31. H. Karimipour, C. Witzel. Colour expectations across illumination changes. *Vision Res* **2024**, 222, 108451.
32. P.E. Pavlova, K.P. Cyrrilov, I.N. Moumdjiev. Application of HSV colour system in identification by colour of biological objects on the basis of microscopic images. *Computerized Medical Imaging and Graphics* **1996**, 20 (5), 357–364.
33. Q. Zaidi, M. Bostic. Color strategies for object identification. *Vision Res* **2008**, 48 (26), 2673–2681.