A Binary Object Detection Pattern Model to Assist the Visually Impaired in detecting Normal and Camouflaged Faces

S. Sajini
Department of Computer and Information Science, Faculty of Science, Annamalai University, India
sajinisheriff@gmail.com (corresponding author)

B. Pushpa
Department of Computer and Information Science, Faculty of Science, Annamalai University, India
pushpasidhu@gmail.com

Received: 15 November 2023 | Revised: 26 November 2023 and 29 November 2023 | Accepted: 4 December 2023

ABSTRACT

This study presents a novel Binary Object Detection Pattern Model (BODPM) to detect objects with face key points and recognize them using the KERAS dataset. The proximity and accuracy of the recognized items were evaluated using computer vision techniques. The object recognition time interval and duration were recorded and stored permanently in a database, and the information was communicated to the visually impaired user as voice output. The normal face, without wearing a mask, was identified using binary patterns with proximity detection. Camouflaged objects were detected in a maximum probability range of 100%. The proposed method was tested, calculating accuracy and score, and compared with existing models, showcasing remarkable performance. The proposed method of normal and camouflage detection is a novel prediction with proximity analysis of objects in a frame.

Keywords-binary object detection pattern model; KERAS dataset; visually impaired; computer vision; camouflaged object detection; parameter assessment

I. INTRODUCTION

Visually Impairment (VI) is a complicated challenge that may lead to unemployment in most areas of work around the world [1]. A native dataset was used in [2] to record and support users to determine objects in front of the visitor office of a visually impaired person. With the use of VI glasses, the visually impaired user was able to recognize visitors to their office and append user information to the native dataset, such as their name, timeslot of arrival, and proximity [3]. This will make it easier for visually impaired people to find employment, as they can record and keep visitor information on hand [4]. The primary objective of this study was to design and construct a framework to identify concealed objects and aid people with VI. This was achieved using the extensively recognized dataset from the KERAS repository [9]. This dataset was used to determine the spatial separation between the objects and to determine their respective designations. The study emphasized on the detection of the proximity of objects to VI individuals, gauging the degree of proximity as a percentage and establishing their significance in predictive analysis. Evaluation metrics, including accuracy and F-score, were considered when subjected to classifier tests. Details concerning user, time allocation, and related information were transmitted as audio input to the speakers. Simultaneously, this information was logged and integrated into the native dataset for update purposes. The scope of this study was limited to the facilitation of people with VI, helping them recognize familiar and unfamiliar individuals or objects and gauging the spatial intervals between them through the application of computer vision techniques. Such an assistive device not only expands the vocational prospects of visually impaired people but also contributes to improving their overall quality of life. In [5], a multitasking CV model was built in hybridization with a segmentation process at training levels. A similar model was also presented in [6], where the computer vision models Random Forest Tree (RFT) (85%), Decision Tree (DT) (80.8%), and K-NN (74.8%) showed good results, but unlike [5], without conducting a separate indoor and outdoor analysis. Table I shows previous studies and proposed models along with their identified gaps.

The reviewed studies were significantly behind in some areas, although they confirmed that more efficient models could not detect objects, faces, and proximity at the same time. These research gaps included the following significant factors: (1) Merged or interlaced objects were difficult to identify [16], (2) could not handle collective problems such as objects, faces, and proximal detections simultaneously [7], and (3) the actual implementation could not be modeled due to the global nature of the COCO dataset [8]. Furthermore, existing models need to accommodate a compatible database such as KERAS [9]. As a result, these gaps must be filled in the proposed model, which should be implemented, tested, and compared with existing models to determine its performance.

TABLE I. EXISTING MODELS AND METHODS

Ref.	Models and Methods	Gaps identified
[10]	Constructed a sequence of actions employing computer vision models	The demand for substantial computational intelligence was noted
[11]	Explored the significance of computer vision in the realm of security and its practical applications.	Stressed the necessity for an object detection-centric safety protocol
[12]	Assessed the effectiveness of distinct classifiers, such as k-means, kNN, DT, and RFT models.	Anticipated enhanced performance from hybrid models when handling complex images
[13]	Investigated challenges and importance of interfacing with smart devices for object tracking, safety, and prediction accuracy	Highlighted limitations of the COCO dataset, prompting the need for updates or hybrid datasets
[14]	Examined a dataset to categorize collected images and segment them into distinct components	Identified the requirement for additional preprocessing for facial object images
[15]	Created an environment and motion-based facial identification system capable of alerting users and detecting concealed objects	Expressed the potential for enhancing proximity and tracking mechanisms in the models
[16]	Employing the You Only Look Once v3 (YOLOv3) model, analogous face detection for animals was conducted, with a specific focus on pigs.	The closeness percentage, however, decreased as the pigs' separation from the source device grew. As a result, the relationship between proximity percentage and separation between the items was inverse
[17]	Used in traffic management and control	The proximity percentage of detection was very low
[18]	Explored techniques in image processing and visual perception to enhance the experience for individuals with VIs, offering improved input at a more affordable expense.	The cost of building the model was comparatively high.

II. METHODOLOGY

Computer vision is an advanced technology that offers users valuable information and experience [20]. Its diverse applications include tasks such as drone image processing analysis and the development of educational visual aids. In the context of this study, the potential of this technology was harnessed to assist people with VI in object recognition and information retrieval. This was achieved using the KERAS dataset [21] in conjunction with information from other global datasets. In [22-23], a mobile platform was designed to both detect the proximity of objects and accurately identify them. This platform facilitated real-time object recognition and determination of source distance. However, it faced limitations in detecting camouflaged or masked faces directly in front of the user. The user would receive an auditory message detailing the identified and recorded object, excluding those that were camouflaged. This study addressed this limitation by devising a multiphase framework model, as shown in Figure 1.

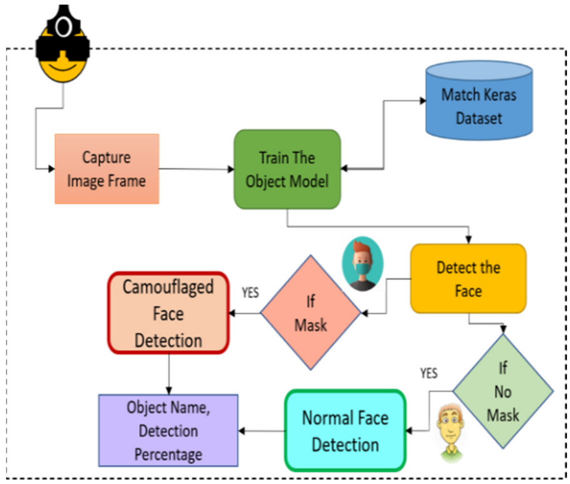

Fig. 1. Model structure for implementing camouflaged object detection.

The system incorporates a webcam or pinhole camera into smart blind glasses, which the visually impaired user employs to capture images. Alternatively, in the scenario of IoT-enabled glasses, a pi camera can be used [24]. This camera processes the image and employs box-shaped face detection techniques to identify blind spots on the object. The KERAS dataset is used to associate the recognized box face with the identification and labeling of the objects. The identified object is transmitted to the processing unit of the smart system, where it is stored as an output along with contextual information, timestamp, and date of detection. Subsequently, the text-to-speech conversion tool [25] converts the image data into textual form and the result is audibly relayed through internal or external speakers. BODPM introduces an innovative concept comprising three primary phases:

- Training the model for object detection
- Detecting normal or camouflaged objects
- Evaluating parameters and providing speech-based assistance

Figure 2 shows the framework for these three phases. In the initial phase of object recognition, the visually impaired individual captures an image using his smart blind glasses placed within his usual field of vision. The acquired image undergoes additional analysis to identify the specific boundary that outlines the face of the detected object. This identification process involves a comparison with the KERAS dataset used in TensorFlow-based object detection methods [26], particularly within models grounded in computer vision. The result of this identification becomes an integral part of the output from the first phase, serving as input for the subsequent phase focused on detecting either normal or camouflaged objects. Using the visual patterns established in the second phase, a novel algorithm called the Binary Object Detection Pattern Model (BODPM) is employed. This algorithm trains the models and differentiates between normal and camouflaged objects. The findings of the object detection phase, along with associated information, are subsequently forwarded to the final phase of the model.

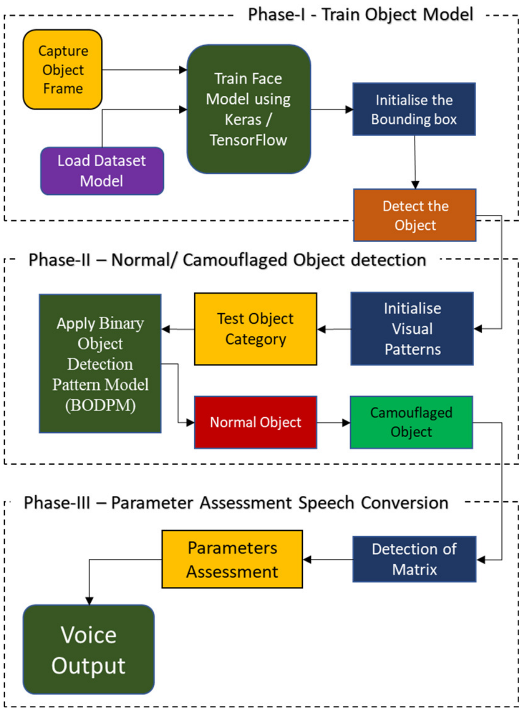

Fig. 2. The Binary Object Detection Pattern Model (BODPM) framework.

Within this phase, an integrated proximity detection algorithm [27-28] calculates the distance between recognized objects and the visually impaired user. This model also records nearby objects, along with essential details such as the date and time of identification, during the transition from textual information to voice-based output in the text-to-voice conversion phase [25]. Here, the conversion engine processes the textual information and delivers the result audibly to the user. In a broader context, BODPM comprises two key phases, each underpinned by the following algorithms.

Figure 3 shows the initialization of the learning rate, epochs for time measurement, and batch size with two categories of detections, including with and without mask, respectively. The image is captured from KERAS to preprocess and training. The image is converted into a matrix form with textual components. The textual components are further segregated into binary data to train the models on the x and y-axis of the image to be captured. Thus, the binarization associates the best features to be captured for a better augmentation of the data captured from the image. Then, the models from the TensorFlow and Keras datasets are used to evaluate the parameters *AveragePooling2D*, *Dense*, and *Dropout*, respectively. The initial base and face objects are loaded and trained with the models. Training is a continuous process, extended until further training is not required. If the model is trained, the model is compiled with loss and optimizer analysis. The model is tested and evaluated using the accuracy measure. Then the fitness function is applied and the model prediction is initiated. Finally, the trained models are evaluated and the efficiency was calculated. After initiating the fitness function to train the model, the object detection model was initiated as shown in Figure 3.

```
Algorithm Camouflaged_Model_Train (CMT)
Step 1: Initialize the Learning Rate, Epochs for
Time Measurement, and Batch Size.
Step 2: Create a list with two categories 'Normal'
and 'Camouflaged'.
Step 3: Capture the images from the KERAS dataset
and preprocess them.
Step 4: Identify the image matrix in text form.
Step 5: Convert the text images to Binary data
with training models in the x and y axes.
Step 6: Generate the image data for data
augmentation including the following: =
{rotation_range, zoom_range, width_shift_range,
height_shift_range, shear_range, horizontal_flip,
fill_mode}.
Step 7: Load the TensorFlow-based KERAS models
with input shapes and size.
Step 8: Load the head model with evaluations
including AveragePooling2D, Dense, Dropout.
Step 9: Load the initial image parameters for base
objects.
Step 10: Test if the objects were trained. If Yes,
no further training, else training is required.
Step 11: If the model is trained, compile the
model with loss and optimizer analysis.
Step 12: Evaluate the model testing for Accuracy.
Step 13: Apply the fitness function and train the
model using validation data.
Step 14: Test epochs until the end of the batch
size.
Step 15: Apply the model and predict the
arguments.
Step 16: Train the model in the background until
detection is completed.
Step 17: Evaluate accuracy and plot the loss in
training.
Step 18: Calculate Train_efficiency : = Loss /
Accuracy.
Step 19: Plot the image and display the results.
Step 20: Terminate when object detection is
completed.
End CMT
```

Fig. 3. Camouflaged Model Train (CMT) algorithm for initiating object recognition.

As shown in Figure 4, the initial image is captured from the video stream of a local web camera or any external device. FaceNet is used to detect the shape of the image. The confidence level is calculated from the image matrix. The bounding box for the object is measured using the product of *shape_detections* × *array_matrix*. The size of the object is then calculated from the selection points on the image face. The preprocessed image is converted from BGR to RGB using color and resizing options. Later, the detected faces are identified based on the user list and the use model algorithm to detect protective items. The path was assigned to calculate the weight to find key facial points. In the follow-up process, when the object comes inside the frame, the mask detection model is loaded to identify the video device and load the stream of bytes. The individual frame of the image is loaded and resized to the frame of the image based on visual patterns. The camouflaged_object_detection fitness function is applied to the captured frame to identify the parameters in the start and end points of the x and y axes (*xstart*, *ystart*, *xend*, *yend*, respectively) to identify objects and draw bounding boxes. The face is tested if it is normal or camouflaged. If normal, then the object is indicated with a bounding box in red color, else the green color is used to indicate the camouflaged image. The *probability_of_mask* is calculated by finding the maximum of *mask* and *without_mask* multiplied by 100. The output bounding box is adjusted based on the detected object. Then, the color changes are modified based on the camouflaged or normal face. Finally, the probability percentage and the accuracy are displayed based on the detected object. The result is also saved in the native dataset to be used for future reference and verification by the visually impaired person, and is also a source of visitors at a particular point in time.

```
Algorithm Camouflaged_Object_Detection (COD)
Step 1: Capture an image from an external camera.
Step 2: Detect the face as shape using facenet.
Step 3: The confidence level is calculated from the
image matrix.
Step 4: Calculate bounding_box: = shape_detections ×
array_matrix.
Step 5: Calculate the size of the object based on the
selection points of the image face
Step 6: Preprocess image from BGR to RGB using color
and resizing options.
Step 7: Detect faces based on the list of users and
use the model algorithm to detect protective items.
Step 8: Assign path and weight to calculate
face_keypoints.
Step 9: Load mask detection model.
Step 10: Identify the video device and load the
stream of bytes.
Step 11: Read an individual frame from the image.
Step 12: Resize the frame of the image based on
visual patterns.
Step 13: Apply the Camouflaged_Object_Detection
fitness function on the captured frame.
Step 14: Identify the parameters in the start and the
end points of the x and y axes.
Step 15: Identify the object and draw the bounding
box.
Step 16: Test if the face is normal or camouflaged.
Step 17: If normal, indicate the bounding box in red
color, else in green.
Step 18: Calculate Probability_of_Mask: = Max(mask,
without_mask)×100.
Step 19: Adjust the output bounding box based on the
detected object.
Step 20: Modify color changes based on camouflaged or
normal face.
Step 21: Display the probability percentage and
accuracy based on the detected object.
End COD
```

Fig. 4. Camouflaged Object Detection (COD) algorithm for object recognition.

III. RESULT AND DISCUSSIONS

The proposed model was compared with existing models. Previous studies showed that the RFT, DT, and kNN models yielded accuracies of 85, 80.8, and 74.8%, respectively. The proposed model was tested using different image parameters. The model was implemented in Python coupled with TensorFlow. Figure 5 shows a preliminary object detection stage.

As shown in Figure 5, the proposed model demonstrated its efficiency in recognizing specific individuals in two distinct locations. Both the timestamp of the visitor's presence and the proximity of the human subject were logged, exhibiting an improved performance rate of 82%. Additionally, all relevant data were systematically stored in a dedicated database. The proposed model surpassed the DT (80.8%) and kNN (74.8%) models. Furthermore, even the RFT model's 85% accuracy was exceeded in outdoor detections, achieving an impressive peak accuracy level of 98%. Figure 6 shows that the proximity performance was higher. The selected entities were the potted plant and humans, with a peak performance of 98%, surpassing existing models. The proposed model also recorded the object's date and time. Other tests explored the effect of merging two objects, one of which was a person. While the individual was identified with a proximity efficiency of 65%, the potted plant was detected with a proximity efficiency of 77%. Consequently, as shown in Figure 7, efficiency decreases when two items are combined.

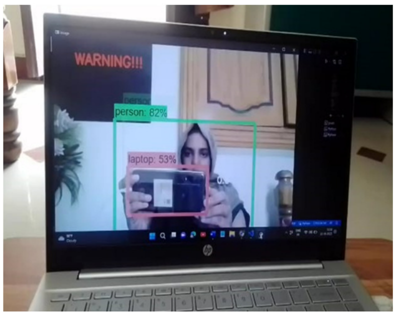

Fig. 5. Object detection using the proposed model. (one of the authors captured during the detection process).

Fig. 6. Object detection in an outside environment (the participants gave permission to show their faces).

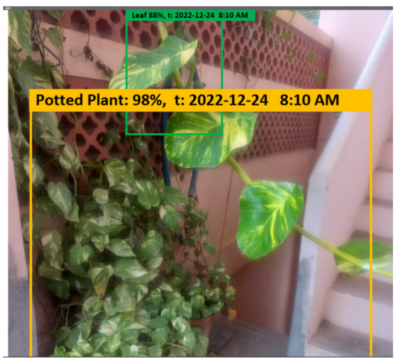

Fig. 7. Object detection of mixed objects.

The proposed model outperformed the existing models with an average difference of 1%, which is a considerable improvement. The experiments demonstrated that the entities could be recognized and locally detected without any systemic mistakes. As shown in Figure 8, the camouflaged detection was completed with a normal face indicated by the red bounding box, and the proximity detection accuracy was 100% for direct faces and 99.51% for faces that were turned. Figure 9 shows that face detection in masked form is indicated with a green bounding box, achieving 100% proximity efficiency. The proposed model was examined and implemented to identify that it could fit into any computer vision model. The proposed hybrid model was able to identify camouflaged faces and objects, outperforming existing models. The model was capable of identifying the objects and faces in both inside and outside environments with remarkable efficiency.

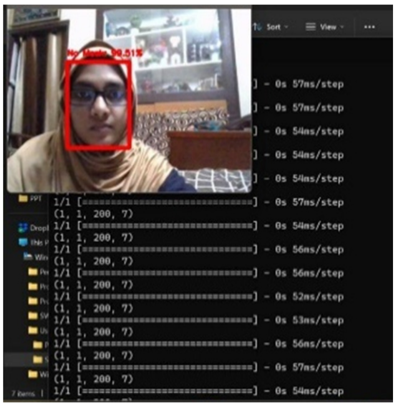

Fig. 8. Object detection in the normal face with red bounding box (one of the authors captured during the detection process).

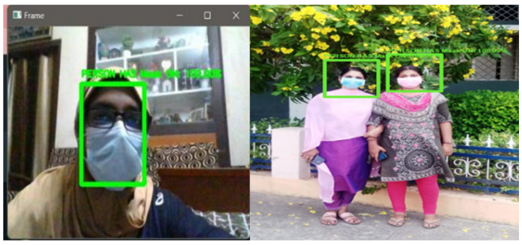

Fig. 9. Object and face recognition after masking the face (the participants gave permission to show their faces).

This study designed and implemented a novel framework called BODPM, which works as a hybrid of existing models. The proposed framework covered three aspects, including image capture, recognition, and proximity detection at the same time. In addition, this model also recorded the time slots in conjunction with the system information. Another noteworthy result of this study was the detection of camouflaged faces. Table II shows the overall comparative analytics of the proposed object detection process with the existing models. The proposed BODPM model outperformed RFT, DT, and kNN, with increased performance in positive measures such as accuracy and F-score, and reduced error rate.

TABLE II. COMPARATIVE ANALYTICS OF BODPM AND EXISTING MODELS IN OBJECT DETECTION

Model	Accuracy	Error Rate	F-score
RFT	85%	0.67	0.83
DT	80.8%	0.76	0.78
kNN	74.8%	0.91	0.65
Proposed BODPM	98%	0.43	0.98

IV. CONCLUSION

Based on the experimental results, the main objective of introducing an enhanced model was successfully realized, exhibiting superior performance over existing models. The proposed BODPM model demonstrated an outstanding peak performance of 98%, which has the potential to serve as a benchmark for subsequent research endeavors. In particular, both indoor and outdoor object detections achieved significantly high levels of reliability. Remarkably, the proposed framework effectively addressed all the shortcomings of previous models, aligning with the intended objectives. The integration of IoT-enabled devices would facilitate the proficient management of substantial quantities of image and video data, addressing a current limitation of the model. Consequently, this inventive framework has proven highly successful in providing a pertinent solution for individuals with VI, while also holding the potential to advance object recognition alongside proximity detection in the future.

REFERENCES

[1] D. L. Hall, J. Allen-Collinson, and P. C. Jackman, "'The agenda is to have fun': exploring experiences of guided running in visually impaired and guide runners," *Qualitative Research in Sport, Exercise and Health*, vol. 15, no. 1, pp. 89–103, Jan. 2023, https://doi.org/10.1080/2159676X.2022.2092200.

[2] H. Xia, C. Yao, Y. Tan, and S. Song, "A dataset for the visually impaired walk on the road," *Displays*, vol. 79, Sep. 2023, Art. no. 102486, https://doi.org/10.1016/j.displa.2023.102486.

[3] A. Ben Atitallah *et al.*, "Embedded implementation of an obstacle detection system for blind and visually impaired persons' assistance navigation," *Computers and Electrical Engineering*, vol. 108, May 2023, Art. no. 108714, https://doi.org/10.1016/j.compeleceng.2023.108714.

[4] M. Shoaib, S. Khan, D. Fitzpatrick, and I. Pitt, "A mobile e-learning application for enhancement of basic mathematical skills in visually impaired children," *Universal Access in the Information Society*, Mar. 2023, https://doi.org/10.1007/s10209-023-00990-3.

[5] L. Beyer *et al.*, "A Study of Autoregressive Decoders for Multi-Tasking in Computer Vision." arXiv, Mar. 30, 2023, https://doi.org/10.48550/arXiv.2303.17376.

[6] A. Bhattacharya, P. Shukla, A. Banerjee, S. Jaipuria, N. C. Narendra, and D. Garg, "Multitask Scheduling of Computer Vision Workload on Edge Graphical Processing Units," in *2023 15th International Conference on COMmunication Systems & NETworkS (COMSNETS)*, Bangalore, India, Jan. 2023, pp. 588–593, https://doi.org/10.1109/COMSNETS56262.2023.10041358.

[7] Q. Zheng, Z. Peng, P. Zhu, Y. Zhao, R. Zhai, and W. Ma, "An Object Recognition Grasping Approach Using Proximal Policy Optimization With YOLOv5," *IEEE Access*, vol. 11, pp. 87330–87343, 2023, https://doi.org/10.1109/ACCESS.2023.3305339.

[8] K. Tong and Y. Wu, "Rethinking PASCAL-VOC and MS-COCO dataset for small object detection," *Journal of Visual Communication and Image Representation*, vol. 93, May 2023, Art. no. 103830, https://doi.org/10.1016/j.jvcir.2023.103830.

[9] A. D. Logronio, R. Christine Reyes, and N. B. Linsangan, "Age Range Classification Through Facial Recognition Using Keras Model," in *2023*

15th International Conference on Computer and Automation Engineering (ICCAE), Sydney, Australia, Mar. 2023, pp. 343–349, https://doi.org/10.1109/ICCAE56788.2023.10111149.

[10] U. Iqbal, M. Z. Bin Riaz, J. Barthelemy, and P. Perez, "Quantification of visual blockage at culverts using deep learning based computer vision models," *Urban Water Journal*, vol. 20, no. 1, pp. 26–38, Jan. 2023, https://doi.org/10.1080/1573062X.2022.2134041.

[11] N. Khan, S. F. A. Zaidi, J. Yang, C. Park, and D. Lee, "Construction Work-Stage-Based Rule Compliance Monitoring Framework Using Computer Vision (CV) Technology," *Buildings*, vol. 13, no. 8, 2023, https://doi.org/10.3390/buildings13082093.

[12] H. H. Htun, M. Biehl, and N. Petkov, "Survey of feature selection and extraction techniques for stock market prediction," *Financial Innovation*, vol. 9, no. 1, Jan. 2023, Art. no. 26, https://doi.org/10.1186/s40854-022-00441-7.

[13] S. Arshad, O. Akinade, S. Bello, and M. Bilal, "Computer vision and IoT research landscape for health and safety management on construction sites," *Journal of Building Engineering*, vol. 76, Oct. 2023, Art. no. 107049, https://doi.org/10.1016/j.jobe.2023.107049.

[14] B. Guzder-Williams, E. Mackres, S. Angel, A. M. Blei, and P. Lamson-Hall, "Intra-urban land use maps for a global sample of cities from Sentinel-2 satellite imagery and computer vision," *Computers, Environment and Urban Systems*, vol. 100, Mar. 2023, Art. no. 101917, https://doi.org/10.1016/j.compenvurbsys.2022.101917.

[15] R. Kalshetty and A. Parveen, "Abnormal event detection model using an improved ResNet101 in context aware surveillance system," *Cognitive Computation and Systems*, vol. 5, no. 2, pp. 153–167, 2023, https://doi.org/10.1049/ccs2.12084.

[16] I. Bakhshayeshi, E. Erfani, F. R. Taghikhah, S. Elbourn, A. Beheshti, and M. Asadnia, "An Intelligence Cattle Re-Identification System over Transport by Siamese Neural Networks and YOLO," *IEEE Internet of Things Journal*, pp. 1–1, 2023, https://doi.org/10.1109/JIOT.2023.3294944.

[17] M. I. Basheer Ahmed *et al.*, "A Real-Time Computer Vision Based Approach to Detection and Classification of Traffic Incidents," *Big Data and Cognitive Computing*, vol. 7, no. 1, Mar. 2023, Art. no. 22, https://doi.org/10.3390/bdcc7010022.

[18] U. Akshaya Devi and N. Arulanand, "Detection of Road Potholes Using Computer Vision and Machine Learning Approaches to Assist the Visually Challenged," in *Smart Computer Vision*, B. V. Kumar, P. Sivakumar, B. Surendiran, and J. Ding, Eds. Cham, Switzerland: Springer International Publishing, 2023, pp. 61–79.

[19] K. Malik, C. Robertson, S. A. Roberts, T. K. Remmel, and J. A. Long, "Computer vision models for comparing spatial patterns: understanding spatial scale," *International Journal of Geographical Information Science*, vol. 37, no. 1, pp. 1–35, Jan. 2023, https://doi.org/10.1080/13658816.2022.2103562.

[20] M. Saini and S. Susan, "Tackling class imbalance in computer vision: a contemporary review," *Artificial Intelligence Review*, vol. 56, no. 1, pp. 1279–1335, Oct. 2023, https://doi.org/10.1007/s10462-023-10557-6.

[21] D. Gupta, "Image Segmentation Keras : Implementation of Segnet, FCN, UNet, PSPNet and other models in Keras." arXiv, Jul. 24, 2023, https://doi.org/10.48550/arXiv.2307.13215.

[22] A. I. H. Holguín, L. C. Méndez-González, L. A. Rodríguez-Picón, I. J. C. P. Olguin, A. E. Q. Carreón, and L. G. G. Anaya, "Assistive Device for the Visually Impaired Based on Computer Vision," in *Innovation and Competitiveness in Industry 4.0 Based on Intelligent Systems*, L. C. Méndez-González, L. A. Rodríguez-Picón, and I. J. C. Pérez Olguín, Eds. Cham, Switzerland: Springer International Publishing, 2023, pp. 71–97.

[23] W. Ali, G. Wang, K. Ullah, M. Salman, and S. Ali, "Substation Danger Sign Detection and Recognition using Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 13, no. 1, pp. 10051–10059, Feb. 2023, https://doi.org/10.48084/etasr.5476.

[24] N. A. Alsharif, S. Mishra, and M. Alshehri, "IDS in IoT using Machine Learning and Blockchain," *Engineering, Technology & Applied Science Research*, vol. 13, no. 4, pp. 11197–11203, Aug. 2023, https://doi.org/10.48084/etasr.5992.

[25] K. Guravaiah, Y. S. Bhavadeesh, P. Shwejan, A. H. Vardhan, and S. Lavanya, "Third Eye: Object Recognition and Speech Generation for Visually Impaired," *Procedia Computer Science*, vol. 218, pp. 1144–1155, Jan. 2023, https://doi.org/10.1016/j.procs.2023.01.093.

[26] R. Doyaoen, S. Aromin, S. Arellano, and T. Palaoag, "Real-time building instance detection using tensorflow based on facade images for urban management," *AIP Conference Proceedings*, vol. 2602, no. 1, May 2023, Art. no. 030023, https://doi.org/10.1063/5.0146699.

[27] H. Chen, L. Hou, G. (Kevin) Zhang, and S. Wu, "Using Context-Guided data Augmentation, lightweight CNN, and proximity detection techniques to improve site safety monitoring under occlusion conditions," *Safety Science*, vol. 158, Feb. 2023, Art. no. 105958, https://doi.org/10.1016/j.ssci.2022.105958.

[28] M. M. H. Milu, M. A. Rahman, M. A. Rashid, A. Kuwana, and H. Kobayashi, "Improvement of Classification Accuracy of Four-Class Voluntary-Imagery fNIRS Signals using Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 13, no. 2, pp. 10425–10431, Apr. 2023, https://doi.org/10.48084/etasr.5703.