ABUAD Journal of Engineering Research and Development (AJERD)
ISSN (online): 2645-2685; ISSN (print): 2756-6811

Volume 5, Issue 1, 147-154

Development of a Convolutional Neural Network-Based Object Recognition System for Uncovered Gutters and Bollards

Ibrahim Adepoju ADEYANJU[*1], Muhammed Adekunle AZEEZ[2], Oluwaseyi Olawale BELLO[3], Taofeeq Alabi BADMUS[4] and Mayowa Oyedepo OYEDIRAN[5]

[1,2,4] *Department of Computer Engineering, Federal University, Oye-Ekiti, Nigeria*
ibrahim.adeyanju@fuoye.edu.ng;azmudray1@gmail.com;taofeeq.badmus@fuoye.edu.ng

[3]*Department of Computer Engineering, Ekiti State University, Ado-Ekiti, Ekiti State, Nigeria*
bello.oluwaseyi@eksu.edu.ng

[5]*Department of Computer Engineering, Ajayi Crowther University, Oyo, Oyo State, Nigeria*
mo.oyediran@acu.edu.ng

Corresponding Author: ibrahim.adeyanju@fuoye.edu.ng, *+2348132876689*

Date Submitted: 09/02/2022
Date Accepted: 21/06/2022
Date Published: 30/06/2022

Abstract: *Machine learning and deep learning have advanced considerably over the last few years with machine intelligence transitioning from laboratory to several industrial applications. Among the deep learning techniques, Convolutional Neural Networks (CNN) have been shown to have one of the best performances in image recognition. CNN has been used for the recognition of a lot of outdoor objects such as buildings, potholes, and cars but with little attention to the recognition of uncovered gutters and bollards, typically found in urban areas and higher institution environments of most developing countries. Hence, a CNN-based object recognition system for uncovered gutters and bollards, with high accuracy and low time complexity, was developed in this research. This can be used to aid outdoor navigation for the visually impaired. The images of uncovered gutters and bollards were captured locally with a high-resolution camera. The datasets were pre-processed by resizing the images and annotations carried out to generate the images' textual equivalent as well as define specific object boundaries. CNN was applied for feature extraction and recognition with two convolutional layers, two pooling layers, and a fully connected layer. The system implementation was done with Python programming language, OpenCV libraries, and Yolov4 as the CNN version with a percentage split experimental evaluation methodology. Results from experiments on the uncovered gutter dataset gave accuracy and average computational testing time of 80% and 0.4 s, respectively. Similarly, the bollards dataset with multiple bollards per image gave accuracy and average computational testing time of 72% and 0.47 s, respectively. The output of this research will be useful for outdoor navigation of the visually impaired when integrated into appropriate electronic hardware.*

Keywords: *Deep learning, convolutional neural network, object recognition, gutter, bollard.*

1. INTRODUCTION

The emergence of new technology has tremendously contributed to the way we walk, talk, see and interact in the environment. One of the examples of these technologies is the assistive technology built to support people living with disabilities. Mobility aids such as wheelchairs, scooters, walkers, canes, crutches [1] are available to assist people with physical disabilities [2]. Globally, the number of people of all ages with visual impairments is estimated to be 285 million, of whom 39 million are blind [3]. About 26.3 million people in Africa have a form of visual impairment, 20.4 million have low vision, and 5.9 million are estimated to be blind [4].

Visual impairment can limit people's ability to perform everyday tasks and affect the quality of life and ability to interact with the surrounding world [5]. There are several traditional methods of assistance provided for the visually impaired. These methods include various types of canes that help them in navigation. However, the cane is insufficient because every day routine mobility requires mastery of a regular route, and this requires memorization of the paths for a walk around. Guide dogs are available to lead visually impaired people around obstacles. The dogs have been trained to guide, but they are red-green colour blind and incapable of interpreting street signs [6]. Assistive technologies have been developed to address many of these challenges. Computer software and hardware, such as voice recognition programs, screen readers, and screen enlargement applications, to assist people with movement and sensory impairments use computers and mobile devices [7]. Technological infrastructure aids visually impaired navigation, such as voice announcements on buses and 'talking' crosswalks for outdoor navigation and Braille signs inside buildings. However, the devices are not universally available.

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

Other devices are electronic canes with recognition of obstacles, sonar vision glasses for obstacle recognition, and GPS navigation devices providing directions.

Over the last years, deep learning methods have been mostly used in computer vision fields more than other machine learning techniques [8]. Deep learning has been used in image classification, object tracking, pose estimation, text detection and recognition, visual saliency detection, action recognition, and scene labelling for people with visual impairment [9]. Deep learning models such as Convolutional Neural Networks (CNNs) are commonly used models. Among different types of models, CNNs have been demonstrated high performance on image classification [10]. One main benefit of CNNs that makes them ideal for object recognition is that they do not rely on inputs of known feature sets, such as the morphological characteristics that taxonomists rely on. Instead, training a CNN only requires a set of labelled images that the model can learn from, developing its own feature set that it uses for identification [11]. CNN has been used for the recognition of a lot of outdoor objects such as buildings, potholes, and cars but with little attention to the recognition of uncovered gutters and bollards, typically found in urban areas and higher institution environments of most developing countries. Hence, a CNN-based object recognition system for uncovered gutters, with high accuracy and low time complexity, was developed in this research to aid outdoor navigation for the visually impaired. Precision, recall, f1-score, accuracy, and average computational testing time were used as performance metrics in a percentage split experimental evaluation methodology with the datasets of uncovered gutters. The system is compared with the existing systems of different intersections over union values.

2. RELATED WORKS

Several research address issues facing the visually impaired and technological approaches to tackle these challenges. Real-time grocery recognition for the visually impaired was designed with a Speeded Up Robust Features (SURF) and a Naive-Bayes [12]. The same SURF descriptor was used for creating a banknote recognition system in the training step of this method is done on banknote ground truth, and features are extracted and stored in a database [13]. A technique to find lost objects was developed. In this work, a Scale Invariant Feature Transform (SIFT) descriptor and colour attributes with sonification were used to guide a person's hand towards the query object. Predefined object patterns were located with SIFT while the colour attribute was used for searching for unknown object patterns [14]. Restroom signage recognition and recognition system were also developed, the shape takes place at the entrance area, and then SIFT features are extracted to recognize the signage bypassing the matching score via thresholding block [15].

A system was proposed for object recognition using a multifunctional classifier interface. The research shows that in difficult situations the system can correctly identify the object [16]. The hybrid CNN and hidden Markov model (HMM) were proposed to recognize the total number of street view buildings [17]. The hybrid model significantly increases the quality of reconnaissance. The CNN refinement model and hybridized optimized segmentation were proposed, and the experiment achieved accurate recognition results [18]. However, the separation of the target object from the background was a challenge, if there are obvious differences between foreground and background, which could affect the results of the recognition.

A deep learning-based pothole recognition system was proposed for the visually impaired safe movement. However, the system is only suitable for a dark environment as the laser patterns used are only visible in a dark situation [19]. Deep learning techniques to develop a system that detects both the potholes, and humps on the road were used, which predict their severity from vibration signals generated by vehicles [20]. The system is effective in detecting the pothole and humps but cannot detect expansion joints and pipeline holes.

Wearable mobility aid uses an RGB-D sensor and a deep learning technique that enables semantic categorization of detected obstacles was also designed. However, these systems still require the use of RGB-D sensors and wearable computers to carry out image processing [21]. A smartphone-based guiding system that uses a real-time object detector was designed to detect obstacles in front of the user [22] [23]. A deep learning approach for floor recognition was also presented. The complexity of identifying the patterns of a floor area in many different scenarios makes the recognition excessively difficult [24].

A convolutional neural network door handle dataset that can be used for model training in the future was designed, the system was trained on a GPU-based deep learning framework called Darknet, the system cannot detect the hand or gives multiple candidates, many of which are just false alarms [25]. Convolutional neural network for other object recognition proposed [26] [27]. The Recognition System has a pre-trained object on the ImageNet dataset to carry out real-time object recognition, the results show that the recognition achieved by the system is fairly accurate, but it fails to identify smaller objects [26] while [27] applied the same technique to potholes recognition system, but benchmarked the system on KITTI ROAD dataset. 20 iterations were carried out for testing, more iterations reduce the performance of the system in real-time object recognition, although the computation time is quite low.

Faster Regional-Based CNN (Faster R-CNN) was used to simulate a low-cost solar-powered wearable assistive technology was designed. The accuracy rate is 99.35%, but this system was not implemented in a real environment [28]. An integrated AI-Based visual aid with an integrated reading assistant, pre-trained the images with Single Shot Detector (SSD), and video process with Faster R-CNN was designed. The design is compact and works indoors and outdoors, but it lacks advanced functions such as wet-floor and staircase recognition [29]. An Arduino-based monitoring system with the same Faster R-CNN was proposed, the system achieved an accuracy of 91%, but it did not benchmark with an existing system [30].

ISSN (online): 2645-2685; ISSN (print): 2756-6811

3. MATERIALS AND METHODS

3.1 Research Approach

Figure 1 shows the block diagram of the object recognition system that describes the components of the system. The phases are Data acquisition, Image pre-processing, Feature Extraction and Recognition with CNN, Implementation, and Performance Evaluation.

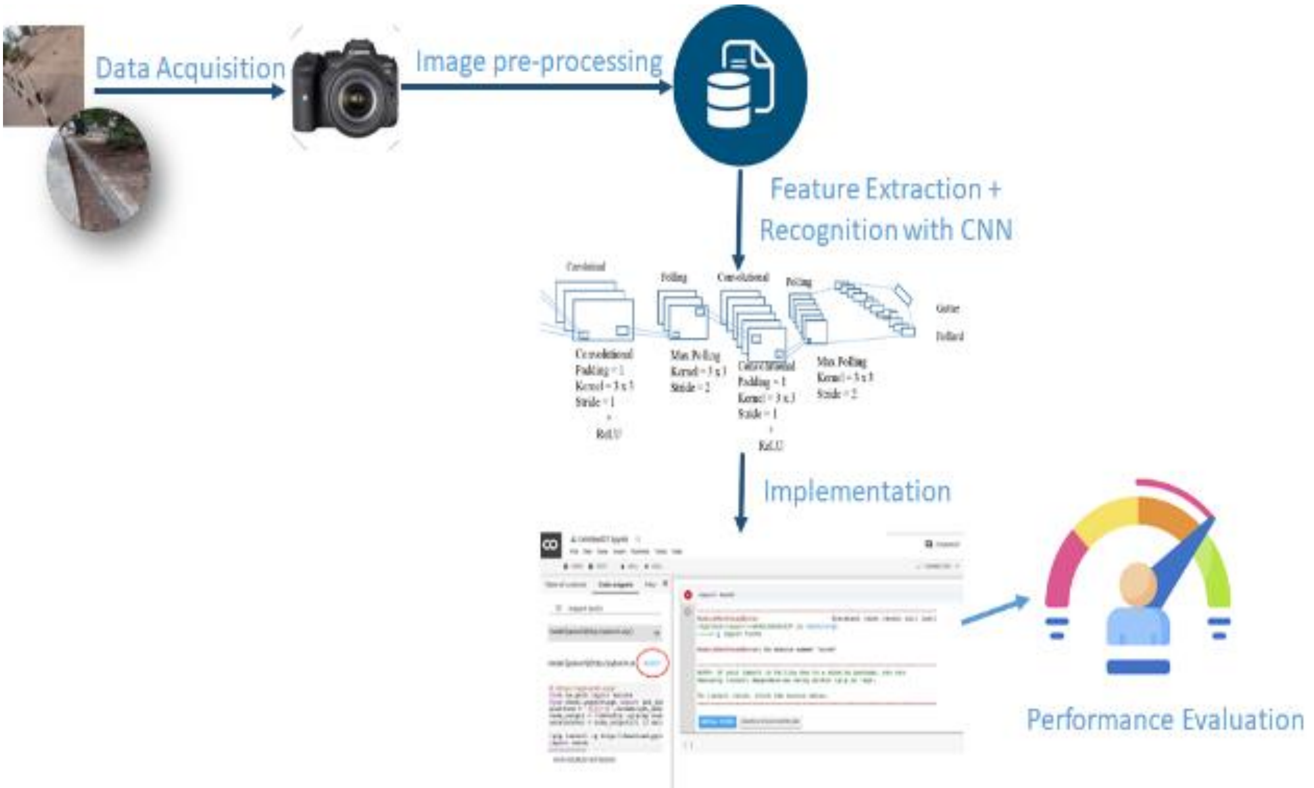

Figure. 1: System Process Flow

3.2 Data Acquisition

1,750 images were captured during the day at the University of Ibadan, Nigeria. Out of the total, 1,200 sets contained gutter objects with a default size of 2448 × 3264 pixels and the rest 550 comprise multiple stationary bollards objects 4000 × 3000 pixels. Table 1 summarized the type, number, and size of the captured images.

Table 1: Data Acquisition Summary

S/N	Objects	Type	Nos of Images	Dimension (Pixels)
1	Gutter	Square	1200	2448 x 3264
2	Bollards	Stationary	550	4000 x 3000

3.3 Image Pre-Processing

This is a process of removing low-frequency background noise, normalizing the intensity of images, the image resizes, removing or enhancing data images before computational processing. The pre-processing techniques applied to the images captured are image resizing with FastStone Image viewer software. Faststone is proprietary software but it is free for non-commercial use [31].

1) Image Resizing: The sizes of gutter object images were manually reduced to 245 × 326 pixels from the original dimension of 2448 × 3264 pixels. The image of the bollard objects dimension was reduced to 400 × 300 pixels from the initial 4000 × 3000 pixels as shown in Table 2. This was done to reduce the training time, and at the same time retain the qualities of the images.

ISSN (online): 2645-2685; ISSN (print): 2756-6811

Table 1: Image resize details

S/N	Image	Initial Size (Pixels)	New Size (Pixels)
1	Gutter	2448 x 3264	245 x 326
2	Bollards	4000 x 3000	400 x 300

2) Image Annotations: The annotation of an object in the pictures creates metadata for the object in the dataset. Each image was annotated with LabelImg software [32] to generate the images' textual equivalent and define specific object boundaries. The directory that consists of the images was used as Open Directory, while the one contains the equivalent Yolo text files saved in the Change Save Directory. From the annotation, the text files of the objects annotated in the images were generated with the coordinate values (<class id> <Xo/X> <Yo/Y> <W/X> <H/Y>). The first field is the class that is set to 0, while the others are x, y, width, and height of the object.

3.4 Feature Extraction and Recognition with CNN

The CNN model used for this research has two convolutional layers, followed by two pooling layers, and fully connected layers as shown in Figure 2. The first hidden layer is the convolutional layer and Rectified Linear Units (ReLU) as an activation function. Each of these components is discussed as applied to the images subsequently before the final prediction.

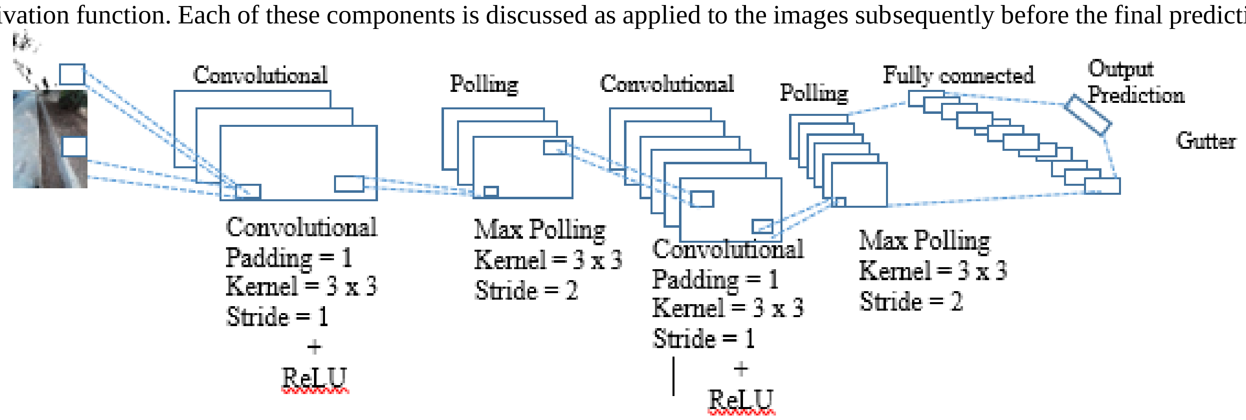

Figure 2: Proposed CNN Architecture for object recognition

3.4.1 Convolutional layer feature extractions

The convolutional layer was used for feature extractions. The network tried every possible matching feature from the input image that contains either gutter or bollards and compared it to the object to be recognized i.e. gutter, bollards. The input image of 245 × 326 pixels, 400 × 300 pixels, and a 3×3 size 2 Dimension convolution kernel. Multiple convolution kernels were used which run over the image and compute a dot product.

3.4.2 ReLU activation

The convolutional layer contains ReLU activation to make all negative values zero. ReLU activation function was used to introduce nonlinearity. The negative pixel value from the previous convolved feature map matrix was replaced by 0 after ReLU operations. ReLU is defined by the following function:

$$ReLU(x) = \max\,(0, x) \ldots\ldots\ldots\ldots\ldots\ldots. (1)$$

From equation 1, the output of ReLU is the maximum value between zero and the input value. Where x is an input value

3.4.1 Convolutional layer feature extractions

The convolutional layer was used for feature extractions. The network tried every possible matching feature from the input image that contains either gutter or bollards and compared it to the object to be recognized i.e., gutter, bollards. The input image of 245 × 326 pixels, 400 × 300 pixels, and a 3×3 size 2 Dimension convolution kernel. Multiple convolution kernels were used which run over the image and compute a dot product.

3.4.4 Fully Connected Layer

The fully connected layer 2 contains one neuron where softmax classifier activation was used to predict the output of the model, and this represents a class of gutters. The softmax function is used as an activation function for the gutter prediction. Here is the equation for the softmax activation function.

$$softmax(Z_i) = \frac{e^{Z_i}}{\sum_{j=1}^{k} e^{Zj}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

Where Z_i = input vector, e^{Z_i} = standard exponential function for input vector, k = number of classes in the multi-class classifier, e^{Z_j} = standard exponential function for output vector.

4. IMPLEMENTATION

The implementation of the models was performed in the Google Colab (cloud) with Python 3 Google Compute Engine backend (GPU) 2.50G with 2.16G of memory assigned for training. OpenCV version 3.2.0 and GPU: Tesla T4 used to minimize the computational time. The Yolov4 model was implemented by cloning Darknet respiratory in colab VM. Yolov4 mainly supports end-to-end training and testing and can detect and recognize multiple objects in images with better accuracy and faster speed [33]. At the first stage, the runtime type was changed to GPU, then the Darknet git repository was cloned onto the colab VM. Thereafter, yolov4(for gutters and bollards) folders were created in google drive. The files (i.e. "process.py", "yolov4-custom.cfg", "obj.zip" "obj.data", "obj.names" and) needed for training were created. Adjustments were made in the .cfg file to reflect the class which was set to 0 for one class of object, and batch number 64, subdivision 16, width (416) x height (416), max_batches 6000, steps (4800,5400), filters 18 and classes 0. The google drive was mounted and linked to the parent folder. The OpenCV and GPU were enabled by making changes in the "Make file" and the run command was executed to build the Darknet files created were transferred from google drive to the Darknet directory in Collaboration (Colab) Virtual Machine. The process.py python script was executed to create the train.txt and test.txt files. The pre-trained yolov4(for gutters) weights were downloaded. Then the detector training was executed at this stage, and the performance of the system was checked.

5. PERFORMANCE METRICS

The Intersection Over Union (IoU) is a number from 0 to 1 that specifies the amount of overlap between the predicted and ground truth bounding box. The prediction is considered to be True Positive if IoU > threshold and False Positive if IoU < threshold. Poor prediction means the IoU value is less than the threshold value, good prediction is IoU greater than or equal to the threshold value, while prediction is classified to be perfect when IoU is greater than 0.9(90%).

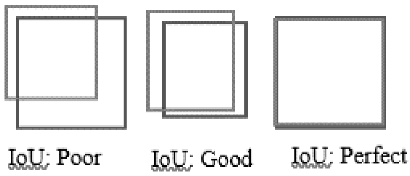

Figure 3: IoU Prediction Bounding Box Classification

Precision is related to the exactness and is expressed by

$$Precision = \frac{TP}{TP + FP} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3)$$

Recall means the recognition of completeness and is expressed by

$$Recall = \frac{TP}{TP + FN} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (4)$$

Accuracy means to the average correctness of a classification process is expressed by

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \dots\dots\dots\dots\dots\dots\dots (5)$$

F1 Score is the weighted average of Precision and Recall.

$$F1\ score = 2 * \frac{precision\ x\ recall}{precision + recall} \dots\dots\dots\dots\dots\dots\dots (6)$$

Table 2: Evaluation parameters

Value	Description
TP(True Positive)	Labelled Object Correctly Identified as Gutter
FP(False Positive)	Labelled Object Incorrectly Identified as Gutter
FN(False Negative)	Unlabelled Object Incorrectly Identified as Gutter
TN(True Negative)	Unlabelled Object Correctly Identified as unlabelled Gutter

 ISSN (online): 2645-2685; ISSN (print): 2756-6811

Table 3 shows the description of the metrics used for object classification which designates whether an object is predicted appropriately or not. The labelled objects would be classified as True Positive if the prediction is True and as False Positive if is False. Unlabelled objects are classified as False Negative when the prediction is False while as True Negative when is True.

6. EXPERIMENTAL SETUP

A total number of 110 pre-processed images were used for the experiment. 85 of the images were for training, 10 for validation, and 15 for testing. The results of validation showed the best models as weight. The models that showed the best predictions were used to analyze 15 images for gutter objects testing. To recognize the objects changes were made to the custom configuration file to set it to test mode. The batch and division lines were changed to 1 automatically with python code. Figure 4 shows the objects recognized with the percentage of prediction in the testing images.

Figure 4: Gutter and Bollards Prediction Outputs

7. RESULTS AND DISCUSSION

7.1 Result for Gutter Recognition

A total of 15 gutter and non-gutter images were used to test the developed system. The images were sorted in descending order of confidence score where 93% is the highest score, and 78% is the lowest score. Images 1, 2, 5, and 6 have 93% confidence, Images 4, 7, 8, and 9 have 81% confidence, Image 3 and 10 have 78% confidence and no object was recognized in images 11 to 15. When prediction unsuccessfully recognized existing gutter objects it was classified as FN and the images without the gutter objects were classified as TN.

7.2 Result for Bollards Recognition

A total of 15 bollard and non-bollard images were used to test the developed system. The images were sorted and arranged based on the actual numbers of objects present and some predicted ones (images with bounding boxes). A confidence score of 100% is the highest score, and 41% is the lowest score. Out of 39 objects present, 38 were recognized as bollard objects, the confidence level considered is 80% (IoU = 0.8), the ones predicted to be TP have 80% or more confidence while those with less than 80% confidence (IoU < 0.8) are FP. When prediction fails to recognize the existing bollard object, it is FN, and the images without a bollard object are TN.

ISSN (online): 2645-2685; ISSN (print): 2756-6811

7.3 Result Discussion

From the result presented in Table 4, the model of gutter objects performed better with a precision of 0.80 and recall of 0.89, and the result shows how efficient and effective the models performed. The system accuracy is 80% and f1-score of 84% and the recognition time for an image is 0.4 seconds. The model of bollard objects produced a precision of 0.71 and a recall of 0.96. Its accuracy is 72%, and f1-score is 82%. It showed that the model is better for recognizing bollard objects even at an upper confidence level of 80% (IoU = 0.8). However, the recognition time is 0.47 seconds compared with 0.4 seconds for the gutter object because bollard images consist of multiple objects and each gutter image has a single object recognized.

Table 4: Comparison of Gutter and Bollard Results

Performance Metrics	Gutter Prediction	Bollards Prediction
Precision	0.80	0.71
Recall	0.89	0.96
Accuracy	0.80	0.72
f1-score	0.84	0.82
Total Recognition Time	0.4s	0.47s

8. CONCLUSION

The research work has developed CNN-based systems for uncovered gutter object recognition. The process started with capturing uncovered gutter images with a high-resolution camera. The image datasets were pre-processed, recognized with CNN (Yolov4) algorithm, and implemented with OpenCV library. Results from experiments on the uncovered gutters dataset gave precision, recall, f1-score, accuracy, and average computational testing time of 0.8, 0.89, 0.84, 80%, and 0.4s, respectively. Also, the bollards dataset with multiple bollards per image gave precision, recall, f1-score, accuracy, and average computational testing time of 0.71, 0.96, 0.82, 72%, and 0.47s, respectively. The research work has developed CNN-based systems for uncovered gutter and bollard object recognition and introduced a locally available annotated dataset for public use. This research will be useful for outdoor navigation of the visually impaired when integrated into appropriate electronic hardware.

REFERENCES

[1] Prajapati, G. & Sharmila, K. (2021). Role of Assistive Devices in Wellbeing of Elderly. *A Review. Indian Journal of Gerontology*, 34(3), 394-402.

[2] Mann, W. C. & Helal, S. (2017). *Technology and chronic conditions in later years: reasons for new hope*. In New Dynamics in Old Age Individual, Environmental, and Societal Perspectives, 271-290.

[3] WHO & WorldBank (2021). “Disability Inclusion Overview,”. [Online]. Available: https://www.worldbank.org/en/topic/disability.

[4] WHO (2019). “Blindness and vision impairment. Universal Eye Health.,”. [Online]. Available: https://www.vision2020australia.org.au/resources/universal-eye-health-a-global-action-plan-2014-2019/.

[5] WHO (2014). Visual impairment and blindness. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs282/en/.

[6] Giudice, N. A. (2018). Navigating without Vision in Principles of Blind Spatial Cognition. *In Handbook of behavioural and cognitive geography*, Edward Elgar Publishing.

[7] Sekharaiah, K.C., Babu, J. R. & Upakara, R. (2011). Towards ICT Applications for Differently Abled People. *IUP Journal of Information Technology*, 7(2)

[8] Voulodimos, A., Doulamis, N., Doulamis, A. & Protopapadakis, E. (2018) “Deep Learning for Computer Vision,” *A Brief Review. Computational Intelligence and Neuroscience*.

[9] Guo, T., Dong, J., Li, H. & Gao, Y. (2017).Simple convolutional neural network on image classification. *2017 IEEE 2nd International* Conference *on Big Data Analysis (ICBDA*), IEEE, 721-724. doi: 10.1109/ICBDA.2017.8078730

[10] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. & Lew, M. (2016). Deep learning for visual understanding: A review. *Neurocomputing*, 187, 27-48.

[11] Spiesman, B., Gratton, C., Hatfield, R., Hsu, W., Jepsen, S., McCornack, B. & Wang, G. (2021). Assessing the Potential for Deep Learning and Computer Vision to Identify Bumble Bee Species from Images. *Scientific Reports*, 11(1), 1-10.

[12] Winlock, T., Christiansen, E. & Belongie, S. (2010). Toward Real-Time Grocery Detection for the Visually Impaired. *In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, 49-56.

[13] Hasanuzzaman, F., Yang, X. & Tian, Y. (2012). Robust and Effective Component-Based Banknote Recognition for the Blind. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6), 1021-1030.

[14] Schauerte, B., Martinez, M., Constantinescu, A. & Stiefelhagen, R. (2012). An Assistive Vision System for the Blind that Helps Find Lost Things. *In International Conference on Computers for Handicapped Persons*, 566-572.

[15] Zhou, W., Li, H., Wang, M., Lu, Y. & Tian, Q. (2012). Binary SIFT: towards efficient feature matching verification for image search. *In Proceedings of the 4th International Conference on Internet Multimedia Computing and Service*, 1-6.

[16] Park, J., Li, S., Wen, W., Tang, P. T. P., Li, H., Chen, Y & Dubey, P. (2016). Faster CNNs with Direct Sparse Convolutions and Guided Pruning. *arXiv preprint*, https://doi.org/10.48550/arXiv.1608.01409

[17] Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. & Lew, M. S. (2016).“Deep learning for visual understanding: A review,” Neurocomputing, 187, 27-48.

[18] Wu, Z., Rengin, C., Shuyan, X., Xiaosi, W. & Yuli, F. (2017). Research and Improvement of WiFi Positioning Based on k Nearest Neighbor Method. *Comput. Eng.*, 43(3), 289-293.

[19] Chanana, P. A., Paul, R. A., Balakrishnan, M. & Rao, P. (2017). Assistive Technology Solutions for Aiding Travel of Pedestrians with Visual Impairment. *Journal of Rehabilitation and Assistive Technologies Engineering,* 4. https://doi.org/10.1177/2055668317725993

[20] Thomas, A., Harikrishnan, P. M., Nisha, J. S., Gopi V. P., & Palanisamy, P. (2021). Pothole and Speed Bump Classification using a Five-Layer Simple Convolutional Neural Network. *In Proceedings of International Conference on Recent Trends in Machine Learning, IoT, Smart Cities and Applications*, 17(16), 491-499.

[21] Poggi, M. & Mattoccia, S. (2016). A wearable mobility Aid for the Visually Impaired Based on Embedded 3D Vision and Deep Learning, *2016 IEEE Symposium on Computers and Communication (ISCC),* 208-213.

[22] Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Recognition. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 779-788.

[23] Lin, B. S., Lee, C. C. & Chiang, P. Y. (2017). “Simple Smartphone-Based Guiding System for Visually Impaired People, *Sensors*, 17(6), 1371. doi: 10.3390/s17061371

[24] Delahoz, Y. & Labrador, M. A. (2017). A Deep Learning-Based Floor Detection System for the Visually Impaired. *In 2017 IEEE 15th Intl Conference on Dependable, Autonomic and Secure Computing, 15th Intl Conference on Pervasive Intelligence and Computing, 3rd Intl Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress*, 883-888.

[25] Niu, L., Qian, C., Rizzo, J. R., Hudson, T., Li, Z., Enright, S. & Fang, Y. (2017). A Wearable Assistive Technology for the Visually Impaired with Door Knob Recognition and Real-Time Feedback for Hand-to-Handle Manipulation. *In Proceedings of the IEEE International Conference on Computer Vision Workshops*, 1500-1508.

[26] Potdar, K., Pai, C. D., & Akolkar, S. (2018). A convolutional neural network based live object recognition system as blind aid, *arXiv preprint,* arXiv:1811.10399.

[27] Islam, M. M., & Sad, M. S. (2018). Path Hole Recognition to Assist the Visually Impaired People in Navigation. *4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT)*, 268-273. doi: 10.1109/CEEICT.2018.8628134

[28] Calabrese, B., Velázquez, R., Del-Valle-Soto, C., De Fazio, R., Giannoccaro, N. I. & Visconti, P. (2020). Solar-Powered Deep Learning-Based Recognition System of Daily Used Objects and Human Faces for Assistance of the Visually Impaired. *Energies,* 13(22), 6104. https://doi.org/10.3390/en13226104

[29] Khan, M. A., Paul, P., Rashid, M., Hossain, M. & Ahad, M. A. R. (2020). An AI-Based Visual Aid with Integrated Reading Assistant for the Completely Blind. *IEEE Transactions on Human-Machine Systems,* 50(6), 507-517.

[30] Ding, N., Dai, Q., & Fan, Y. (2021). Design of Arduino PM2.5 Air Quality Monitoring System. *Proceedings of 2021 IEEE International Conference on Power Electronics Computer Applications (ICPECA)*, 798-800.

[31] The FastStone website. [Online]. (2021). Available https://www.faststone.org/

[32] The labelimg website. [Online]. (2021). Available: https://sourceforge.net/projects/labelimg.mirror/

[33] Du, J. (2018). Understanding of object detection based on CNN family and YOLO. Journal of Physics: Conference Series. 1004(1), IOP Publishing.