



ISSN: 2454-9940



**INTERNATIONAL JOURNAL OF APPLIED
SCIENCE ENGINEERING AND MANAGEMENT**

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

USE OF DEEP LEARNING FOR GROUND HOLE DETECTION IN SURVEILLANCE

¹YEDLA VISHAL REDDY, ²MADDURI RAKESH, ³NAKKA SAI KIRAN REDDY, ⁴AREPALLY SAI,

⁵KORAPAKA VIJAY KUMAR, ⁶Mr. DHANANJAY, ⁷Dr. M ASHOK KUMAR,

¹²³⁴⁵Student Department of DS, Narsimha Reddy Engineering College, Maisammaguda (V), Kompally, Secunderabad, Telangana-500100.

⁶Assistant Professor, Department of CSE, Narsimha Reddy Engineering College, Maisammaguda (V), Kompally, Secunderabad, Telangana-500100.

⁷Professor, Department of Mechanical Engineering, Narsimha Reddy Engineering College, Maisammaguda (V), Kompally, Secunderabad, Telangana-500100.

Abstract

It is crucial to be vigilant and identify possible dangers for the security and monitoring of border areas. Ground holes may sometimes cause suspicion, even if surveillance drones have improved border area monitoring. A variety of illegal operations may be concealed in ground pits, which are often dug or excavated into the ground. Criminals looking to avoid discovery sometimes target intruders who enter via ground trenches because of how innocuous they are. Automating object recognition using visual data is an area where deep learning has shown encouraging promise. Drones equipped with the deep learning model would provide for an enhanced and all-encompassing surveillance system. For the purpose of training and testing YOLO, this work develops a ground pit image dataset (GPID). Three hundred photographs of ground pits on a variety of surfaces, taken by drone and annotated using web tools, make up this dataset. Compared to other deep-learning models, YOLO's accuracy of over 90% is superior.

Index Terms—Computer Vision, Ground Pit Image Dataset (GPID), YOLO, Deep Learning, CNN.

INTRODUCTION

Pits in the ground may be a security risk near borders because they can hide illegal operations, make it easier to penetrate border defenses, and lead to links across borders. Recognizing and addressing these gaps calls for a comprehensive strategy that integrates cutting-edge monitoring tools, information exchange, and joint operations amongst the

appropriate authorities. Border control authorities may improve their detection and mitigation of ground pit threats by using drones, ground-penetrating radar, and other sensor technologies. Drones can survey large border regions in real time because to their high-resolution cameras, infrared sensors, and sophisticated imaging systems.

Superior to more conventional approaches, they are able to record aerial footage, keep tabs on operations, and spot dangers. Agents on the border can respond quickly to suspicious activity because to real-time data transfer. Drones' ability to automatically detect and identify ground holes may be further enhanced by technological breakthroughs in artificial intelligence (AI) and machine learning (ML). Using a Deep Learning technology called object detection in real-time, it is possible to extract specific items from movies and photos. Automating ground pit identification using visual data analysis is a great way to improve safety during open excavations by recognizing possible dangers and avoiding accidents. Deep learning has shown excellent results in this area. Automated and continuous monitoring of border regions is within its capabilities. Drones free up security guards to concentrate on other high-priority duties by reducing the need for human monitoring and inspections. Automatic analysis of picture or video feeds to detect ground pits using systems based on deep learning would be a vast improvement in detection speed. Take, as an example,



Fig. 1: 150-meter cross-border tunnel detected by BSF along IB in Jammu and Kashmir [2]

The construction of a 150-meter tunnel hole along the Inter National Border (IB) has been done to allow terrorists from Pakistan to infiltrate Jammu and Kashmir, as seen in Figure 1. Drones, in conjunction with ground pit detection systems based on deep learning, may therefore provide a more thorough and reliable monitoring system. We want to present an automated solution that uses a drone's camera to find such ground holes for monitoring. This is driven by our goal. In order to photograph the boundary regions—the area of interest—the drone is equipped with a high-resolution camera, as seen in Figure 2. An object identification system, using Deep Learning methods, automatically identifies the ground pits from the drone's collected photos.

In the event that the system identifies any strange diggings, it will flag them as possible threats and notify the relevant security authorities. Consequently, this research employs a deep learning model—specifically, the You only look once (YOLO) model—to detect ground holes in real-time photos acquired by a smart drone. In order to train and evaluate the YOLO model, a dataset called Ground Pit Image Dataset (GPID) was created. This dataset contains photos of ground pits. Around 300 pictures of ground pits on different surfaces, annotated using web tools like Roboflow, make up the GPID dataset [3]. Out of all the deep-learning models we tested, the YOLO model outperformed the competition with an accuracy of over 90%, beating out Faster-RCNN, Fast RCNN, SSD, and others. There are several different versions of YOLO, including the popular little YOLO as well as YOLO v3, YOLO v5, and YOLO v8. We use the GPID dataset to train YOLO-version 8, an existing deep-learning object identification method, with a mean average precision (mAP.50) of 0.759 and a mean absolute precision (mAp) of 0.47 in order to confirm the dataset's practicality. Here is how the remainder of the paper is

structured: Part II reviews the relevant literature. The suggested procedure is detailed in Section III.

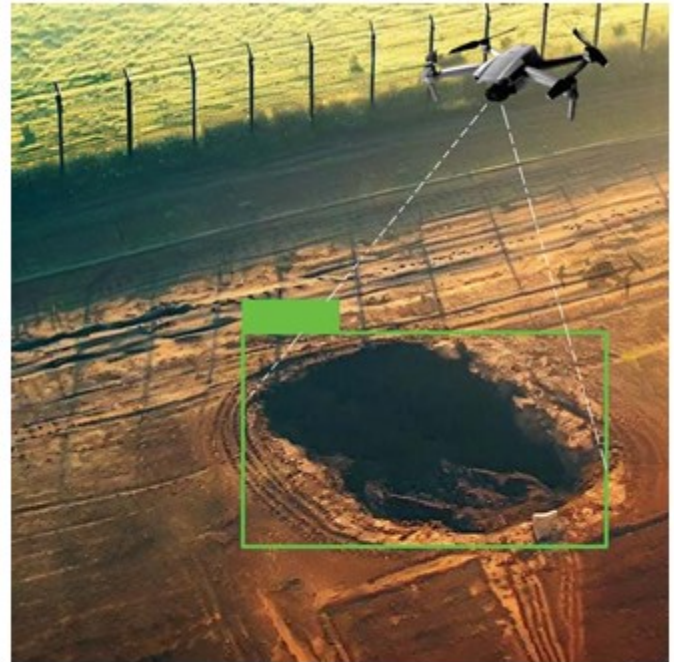


Fig. 2: Ground pit Detection through Drone

with precision. The results of assessing how well the suggested approaches work are presented in Section IV. Section V serves as the paper's last section.

II. RELATED WORKS

Autonomous driving, infrastructure maintenance, security, and surveillance are just a few of the many areas where ground pit detection has recently attracted a lot of interest. Ground hole detection has been the subject of a great deal of research [4]. In order to create an automated machine learning method that permits the gathering of data from crowds, the writers in [5] centered their attention on road pothole identification using data collected from smartphone sensors. The smartphone's accelerometer and GPS have been used to collect data on the road's surface condition and the presence of potholes. To further aid blind people's mobility, the authors of [6] included ultrasonic sensors, buzzers, and a microprocessor inside their sticks to identify impediments. An Arduino IDE was examined by the writers; it used an ultrasonic sensor to detect obstructions at a distance of 100 cm or less. The buzzer will sound an alarm to let the blind person know when it has been detected. Additionally, it is

useful for finding a ground hole in a depression that is 18 cm deep or deeper. Eighty blind people were tested with the stick to see whether it could identify obstructions and holes in the ground. In order to detect and monitor forest fires, a combination of Deep Learning, object recognition, and drones has been studied [7]. In order to monitor and identify forest fires, the authors researched and assessed the latest advancements in drone technology and deep learning technologies. Drones outfitted with customized sensors and cameras allow for efficient and cost-effective real-time surveillance and early detection of fires. The authors analyzed current deep learning object identification breakthroughs, such as YOLO (You Only Look Once) and R-CNN (Region based Convolutional Neural Network), in depth, with an eye toward its possible use in forest fire monitoring. We use a deep learning strategy that is comparable to this one to identify ground holes in a unique dataset. Roadside pothole detection was the primary emphasis of the writers in [8]. Damage to vehicles and a decline in traffic safety are two outcomes of roadside potholes. The You Only Look Once (YOLO) object identification approach was used to identify potholes in real-time by the author. For the purpose of real-time pothole detection in road pictures, they created a YOLO model. A modified version of the VGG16 network is used to identify fractures in [9].

The authors trained their proposed model using their own dataset, CCD1500, and tested it using CFD, DeepCrack, and CrackTree200. Evidence suggests that their suggested model achieves recalls of 90% for CFD, 96% for DeepCrack, and 89.10% for the CrackTree200 dataset, in that order. The authors in [10] used the principal component analysis (PCA) method to classify irregularities into four groups: long bumps, minor bumps, manholes, no anomaly, and others, using data from accelerometers, GPS coordinates, and speeds. This web-based program's main objective is to detect anomalies. In a controlled laboratory environment, anomaly detection has a success rate of 94%, but in real life, it only manages 82%. To find potholes in photos, use the three-step procedure given in [11]. To begin, the shadowy regions around potholes are located and extracted using a histogram and the closure operation of a morphological filter. The second step is to identify possible candidates based on attributes such as pothole density and size. Finally, the pothole is located by considering all of the potential characteristics. With a recall of 73.30%, precision of 80%, and accuracy of 73.50%, the suggested strategy was very effective. An warning system for image-based pothole identification was proposed by Young-

Ro et al. in [20]. Once the photographs have been collected, they are transformed into binary format and then searched for potholes in a database. According to the aforementioned research, drones used for border region surveillance still need the use of deep learning models to identify ground holes or excavations in the vicinity of borders. By locating potential monitoring holes in the ground, the drones may greatly improve border area monitoring. Illicit operations may be concealed in ground pits, which are commonly dug or excavated into the dirt. Criminals looking to avoid discovery sometimes target intruders who enter via ground trenches because of how innocuous they are. Therefore, we tackle the issue of detecting certain types of ground pits in smart-drone photos by using deep learning approaches. Here are the things that we have contributed to this paper: • In order to train our model, we create a collection of ground pit images. In order to guarantee variety in the data set, the development strategy centers on gathering photographs and picking out eligible images. Greater picture datasets may subsequently be generated using the same method. In order to train and evaluate YOLO, we constructed a collection of ground pit images. Drone photos of 300 earth holes on different surfaces make up this dataset. Makesense.AI and Roboflow.com were among the web applications we used to annotate each picture of the earth pit. We validated the viability of deep learning object identification models by evaluating them on this dataset. The border monitoring system is anticipated to be improved with the integration of such a deep learning model into drones.

MATERIALS AND METHODS

Presented below is the suggested method, along with a lengthy account of the dataset's creation and data pre-processing. Figure 2 depicts the steps used to acquire YOLO data and supplement it in order to expand the dataset. In order to assess how well our suggested method works, we detail the experimental setup, data gathering and preparation, and deep learning model in the parts that follow. A. Description of the model Computer vision and object recognition are only two of the many areas that have been profoundly affected by the current revolution in deep learning methods [12]. One field that stands to gain a lot from these innovations is autonomous cars, which are essential since they allow us to travel without physically steering or controlling the vehicle. With the goal of finding effective and reliable

solutions for border area security, this research investigates the possibility of using deep learning models to identify ground holes automatically using visual data. Every thing that came before it makes use of regions.

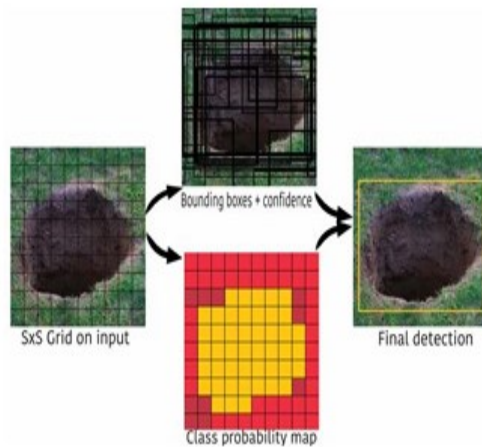


Fig. 3: Working process of YOLO

methods for detecting objects in order to localize them inside the picture. It is not possible for the model to see the whole picture at once. On the contrary, the item is probably included inside certain portions of the photograph. When compared to region-based methods, the YOLO object identification approach couldn't be more different [13]. Bounding boxes and class probabilities for these boxes are predicted by a single neural network model employed in the YOLO [14], [15]. Ground pit detection is addressed in this study by using the capabilities of the YOLO version-8 model. In this section, we will go over the YOLO model's architectural details as seen in Figure 3. An picture is divided into a $S \times S$ grid using this YOLO model. Afterwards, we determine the m th bounding box inside each S^2 grid. Each of the enclosing boxes receives an offset value and a class probability from the model. In order to choose a picture and pinpoint its item inside it, the bounding boxes are evaluated based on whether their class probability is greater than a predetermined threshold number. The YOLO model outperforms competing object identification techniques such as Faster R-CNN and SSD by processing 45 frames per second of the input picture. The reason for this is because the YOLO model's algorithm incorporates geographical restrictions.

Training One improvement over the original YOLO model is YOLO-v8. Our system can identify ground holes efficiently and accurately by combining the capabilities of Deep Learning with the YOLO model. This may aid in border area safety and security monitoring. One of the essential components of the YOLO-v8 model is the Darknet, which is a Deep Neural Network architecture [16]. In order to meet the requirements of YOLO-v8, we adjust the network configuration to include the required amount of layers, filters, and anchor boxes. Depending on the needs of the object detection job, the YOLO-v8 model modifies the architecture.

B. Data Collection and preprocessing

We constructed a set of surface photos with ground holes and called it GPID. Roughly 300 photos of earth pits of varying forms, taken from a variety of weather and angle perspectives, make up the GPID collection. In order to prepare the data for model training, it is first preprocessed by adjusting the picture sizes, adding annotations, and using augmentation methods. In order to train deep learning models to increase performance and minimize overfitting, a big dataset is necessary. This is due to the fact that in order to expand the dataset, data augmentation techniques including cropping, resizing, flipping, rotating, and color modifications are required [12]. Roboflow, an online program, is used to enhance the data, as seen in Figure 4. In order for the model to recognize objects, the picture dataset must be annotated with bounding boxes. We use the online annotation application MakeSense.Ai to add bounding box coordinates and class names to the photos [17]. Annotating data for machine learning applications, particularly in computer vision, has never been easier than with this straightforward and user-friendly tool. The procedure of annotation is shown in Figure 5.

Generating New Version

Prepare your images and data for training by compiling them into a version.
Experiment with different configurations to achieve better training results.

✓ Source Images	Images: 290 Classes: 1 Unannotated: 0
✓ Train/Test Split	Training Set: 203 images Validation Set: 58 images Testing Set: 29 images
✓ Preprocessing	Auto-Orient: Applied Resize: Stretch to 640x640
✓ Augmentation	Flip: Horizontal, Vertical Crop: 0% Minimum Zoom, 10% Maximum Zoom Shear: ±7° Horizontal, ±8° Vertical Brightness: Between -26% and +26% Blur: Up to 1.5px

5 Generate

Review your selections and select a version size to create a moment-in-time snapshot of your dataset with the applied transformations.
Larger versions take longer to train but often result in better model performance. [See how this is calculated](#) »

Maximum Version Size

696 images (3x)

Generate

Fig. 4: Process of automatic data splitting, data pre processing, and data augmentation

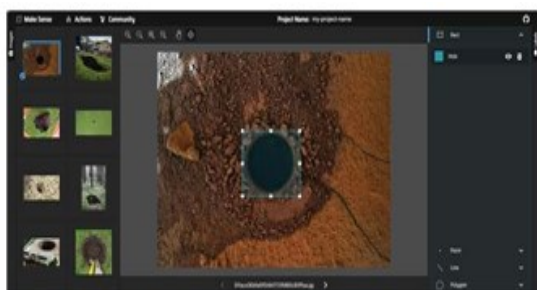


Fig. 5: MakeSense.AI Interface

C. Experimental setup

On two separate platforms, this experimental arrangement is carried out. The model is trained on the Google Colab platform and then evaluated on the local PC. First, preparing the model: For this experiment to run, we need Python, CUDA, cuDNN, and a deep learning framework like Darknet, PyTorch, or TensorFlow, among others, in order to train YOLO-v8. In terms of hyperparameters and training possibilities, we have set up YOLO-v8. For this implementation, we used Google Colab notebook to train our model. Colab is a great environment for

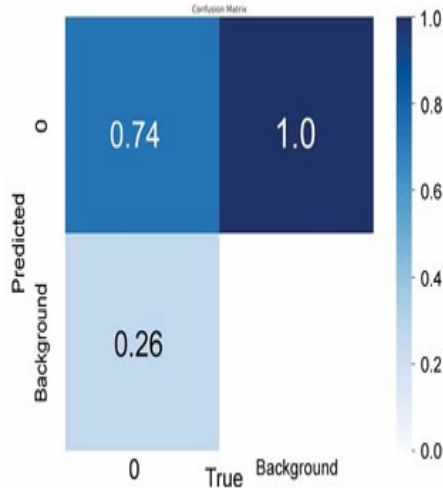
executing Python code, especially for data analysis, machine learning, and research projects, and it comes with a powerful 15Gb GPU Tesla t4. Its accessibility from any device with an internet connection is made possible by its cloud-based nature, which removes the requirement for local installation and setup. 2) The model is tested on a computer system with a 12th-Gen Intel(R)-Core(TM)-i5-1240P (12 core) 4.40GHz CPU and an NVIDIA GeForce GTX 1650 graphics card with 4GB of RAM. This model is called YOLO-v8. Here, we divided our GPID dataset in half. One half was utilized for training, the other for testing, and the last for validation. A total of 88% of the data went into the training set. Hyperparameters like as learning rate, batch size, and data augmentation approaches are part of the training setup. To begin training, we input the YOLO-v8 model the enhanced training data. While training, the model predicts the class probabilities of each item in the pictures and the coordinates of each bounding box [18]. Observing the loss graphs in 8 allows one to keep tabs on the training process. The YOLO-v8 model is trained using many hyper-parameters, including the learning rate, numerous epochs, and regular intervals to preserve the model weights for inference and assessment later on. After the YOLO-v8 model has been trained to perform well, it is deployed on the local system according to the requirements specified before. Drones may use this trained model to identify objects in unseen photos or videos in real-time by integrating it into their systems.

RESULTS AND DISCUSSION

To evaluate the performance of the proposed model on the separate test dataset, different performance metrics have been used, including precision, recall, average precision (AP), and mean average precision (mAP) to measure the model's accuracy and detection quality [19]. Here details of all evaluation metrics are as follows:

Confusion Matrix:

A summary table detailing how well a categorization model performed. It breaks down the statistics for each class, showing the total number of TP, TN, FP, and FN. In doing so, it aids in evaluating the model's class prediction and misclassification capabilities. The most common format for its representation is a square matrix. The ground pit is represented by Class 0 in this (Matrix.6).



Accuracy: It evaluates how well the model can distinguish between actual and expected positive occurrences. The following equation is used to compute it, and it is based on the accuracy of positive predictions; it is also the ratio of true positives to the total of true positives and false positives (see Curve.7b).

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

First, there's recall, sometimes called sensitivity or the true positive rate, which is a measure of the model's capacity to pick out positive examples from among the real positive examples. The ratio of genuine positives to the total of true positives and false negatives is mathematically represented by curve7d. It alludes to the comprehensiveness of optimistic forecasts determined by

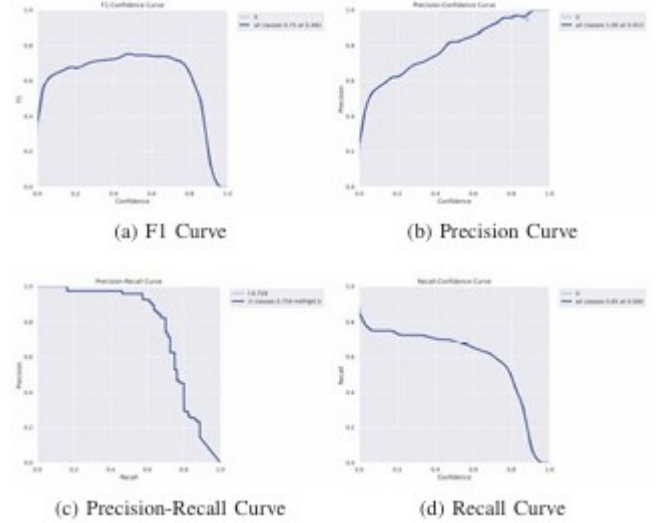


Fig. 7: Evaluation metrics

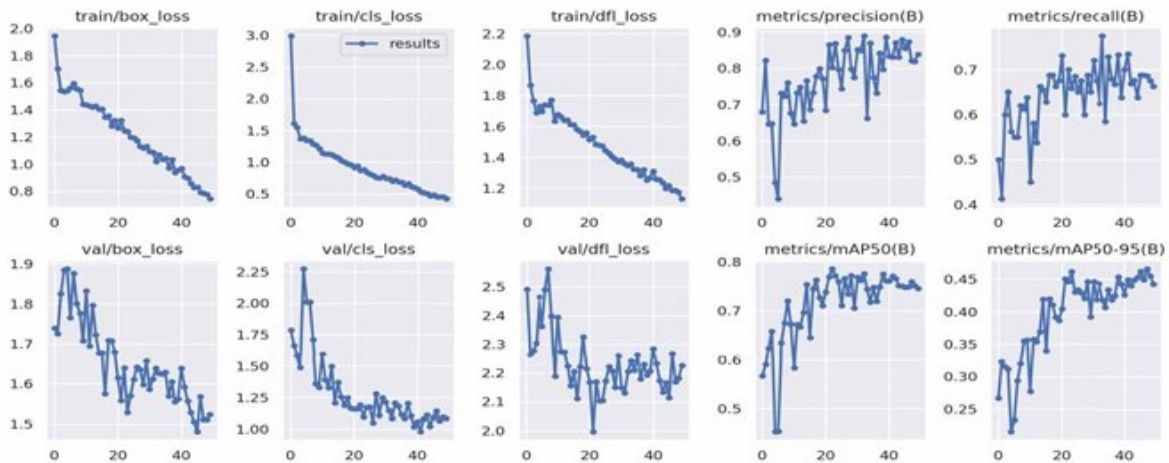


Fig. 8: Overall Results

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

F1 Score: A measure that integrates recall and accuracy into one number is the F1 Score, and it may be used to assess the model's performance. This statistic is the geometric mean of recall and accuracy. This helps in determining the model's efficacy (refer

to Fig.7a(a)), and the following equation assesses that.

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3)$$

One assessment measure for object identification is mAP, which stands for "mean average precision." It's used to calculate the average accuracy across various degrees of recall. By adjusting the confidence threshold for object detection and computing precision (7b) and recall (7d) at each threshold, the precision recall (7c) curve is produced. A class's mAP is its average area under the precision-recall curve (AP). This metric gives object detection models a general idea of how well they're doing. Figure 8 shows that the mean average precision (mAP.50) is 0.759 and the mAP(0.5-0.95) is 0.47. Classification and object identification models often use these metrics for evaluation because they provide light on several facets of model performance, including precision, completeness, and accuracy. In doing so, they aid in determining how well the model predicts outcomes and where it may be improved.

TABLE I: Comparative analysis of YOLO with other methods

Method	mAP
YOLO-v8	90.43%
Faster R-CNN	85.21%
Fast R-CNN	83.21%
R-CNN	78.02%
SSD	76.45%

Results from testing YOLO-v8 on the coverage hole/dig detection challenge are shown below. Evaluation of the model's accuracy, speed, and resilience. The number of epochs is usually shown on the horizontal axis of a standard YOLO training plot, as seen in Figure 8. In these graphs, the loss value is shown along the vertical axis. The training performance of the model may be evaluated by looking at its loss. The model's performance is improved with lower loss levels and worsened with greater ones. Training should focus on reducing the lost value as much as possible. Figure 8 shows that after 40 epochs, losses are almost nonexistent. The terms "box loss," "cls loss," and

"dfloss" are used to assess the effectiveness of the loss in the YOLOv8 framework. All of these things go into the final loss figures. As a rule, the total loss amount is just the weighted sum of all of these losses. It is implementation-dependent as to which units are used for the vertical axis. But in most cases, they show how much of a mistake there was or how much the actual numbers differed from the predictions. In Fig. 9, we can see the outcome of the working model's confidence-based predictions.

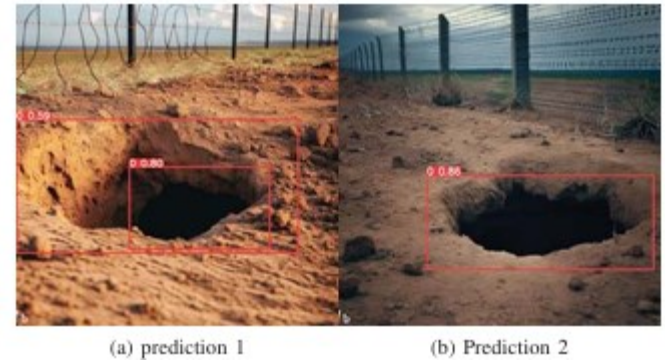


Fig. 9: Result of the working model with Confidence

In order to demonstrate the approach's potential in relation to the mAP, we conclude by comparing our suggested method with other current models, as given in Table I. When it comes to finding flaws in current models like R-CNN, Fast R-CNN, Fast R-CNN, and SSD, our YOLO-v8-based approach achieves an accuracy of over 90%.

CONCLUSION

Illicit operations may be concealed in ground pits, which are commonly dug or excavated into the dirt. Ground trenches are an enticing target for would-be invaders because of how undetectable they are. Therefore, by locating such ground pits for monitoring, drones may significantly improve border area monitoring. Using deep learning approaches, this article addressed the recognition of such ground holes in smart drone photos. Ground Pit Image collection (GPID) is an image collection that we created for the purpose of recognizing earth holes. It contains 300 annotated photos of various sorts of earth holes on different surfaces. A more thorough and reliable monitoring system is provided by our suggested

method, which identifies ground pits with an accuracy of over 90%.

REFERENCES

- [1]. [1] S. P. Kapur, "India and pakistan's unstable peace: Why nuclear south asia is not like cold war europe," *International Security*, vol. 30, no. 2, pp. 127–152, 2005.
- [2]. [2] Wikipedia. (2023) 150-metercross-border tunnel detected by bsf along ib in jk. [Online]. Available: <https://www.orissapost.com/another-150m-cross-border-tunnel-detected-by-bsf-along-ib-in-jk/>
- [3]. [3] roboflow. (2023) Everything you need to build and deploy computer vision models. [Online]. Available: <https://roboflow.com/>
- [4]. [4] Y.-M. Kim, Y.-G. Kim, S.-Y. Son, S.-Y. Lim, B.-Y. Choi, and D. H. Choi, "Review of recent automated pothole-detection methods," *Applied Sciences*, vol. 12, no. 11, p. 5320, 2022.
- [5]. [5] C. Wu, Z. Wang, S. Hu, J. Lepine, X. Na, D. Ainalis, and M. Stettler, "An automated machine-learning approach for road pothole detection using smartphone sensor data," *Sensors*, vol. 20, no. 19, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/19/5564>
- [6]. [6] S. A. Yahaya, L. J. Jilantikiri, G. S. Oyinloye, E. J. Zaccheus, J. O. Ajiboye, and K. A. Akande, "Development of obstacle and pit detecting ultrasonic walking stick for the blind," *FUOYE Journal of Engineering and Technology*, vol. 4, no. 2, 2019.
- [7]. [7] M. Yandouzi, M. Grari, B. Mohammed, I. Idrissi, O. Moussaoui, M. Azizi, K. Ghomid, and K. E. Aissa, "Investigation of combining deep learning object recognition with drones for forest fire detection and monitoring," *International Journal of Advanced Computer Science and Applications*, vol. 14, pp. 377–384, 03 2023.
- [8]. [8] N. Patel, V. Dabhi, and R. Adhvaryu, "Identify road potholes using image semanticsegmentationforadvance driver assistant system," *Journal of Data Acquisition and Processing*, vol. 38, no. 2, p. 2307, 2023.
- [9]. [9] Z. Qu, J. Mei, L. Liu, and D.-Y. Zhou, "Crack detection of concrete pavement with cross-entropy loss function and improved vgg16 net work model," *Ieee Access*, vol. 8, pp. 54564–54573, 2020.
- [10]. [10] N. Silva, V. Shah, J. Soares, and H. Rodrigues, "Road anomalies detection system evaluation," *Sensors*, vol. 18, no. 7, p. 1984, 2018.
- [11]. [11] S.-K. Ryu, T. Kim, and Y.-R. Kim, "Image-based pothole detection system for its service and road management system," *Mathematical Problems in Engineering*, vol. 2015, pp. 1–10, 2015.
- [12]. [12] N. Ahmad, S. Arya, and D. Singh, "Predicting risky environment for child inside house using deep learning," in *2023 International Conference on Emerging Smart Computing and Informatics (ESCI)*. IEEE, 2023, pp. 1–6.
- [13]. [13] T. Diwan, G. Anirudh, and J. V. Tembhurne, "Object detection using yolo: Challenges, architectural successors, datasets and applications," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243–9275, 2023.
- [14]. [14] J.-H. Kim, N. Kim, and C. S. Won, "High-speed drone detection based on yolo-v8," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–2.
- [15]. [15] A. Melino-Carrero, ' A. N. Su' arez, C. Losada-Gutierrez, M. Marron Romera, I. G. Luna, and J. Baeza-Mas, "Object detection for functional assessment applications," in *International Conference on Engineering Applications of Neural Networks*. Springer, 2023, pp. 328–339.
- [16]. [16] N. Ahmad, Z. Khan, and D. Singh, "Student engagement prediction in moocs using deep learning," in *2023 International Conference on Emerging Smart Computing and Informatics (ESCI)*. IEEE, 2023, pp. 1–6.
- [17]. [17] (2023) makesense.ai. [Online]. Available: <https://www.makesense.ai/>
- [18]. [18] R. Alam, N. Ahmad, S. Shahab, and M. Anjum, "Prediction of dropout students in massive open online courses using ensemble learning: A pilot study in post-covid academic session," in *Mobile Computing and Sustainable Informatics: Proceedings of ICMCSI 2023*. Springer, 2023, pp. 549–565.
- [19]. [19] N. Ahmad, A. Gupta, and D. Singh, "Using deep transfer learning to predict student engagement in online courses," in *International Conference on Machine Learning, Image Processing, Network Security and Data Sciences*. Springer, 2022, pp. 27–36.