International Information and
Engineering Technology Association
Advancing the World of Information and Engineering

Leveraging UAV Imagery and Deep Learning for Automated Object Detection

Sohan Kanse[1], Vara Prasad Lingam[2], Sahil K. Shah[1], Vidya Kumbahr[1*], T. P. Singh[1], Kumar Karunendra[3]

[1] Symbiosis Institute of Geoinformatics, Symbiosis International (Deemed University), Pune 411016, India
[2] Stantec Resourcenet India Pvt Ltd., Pune 411045, India
[3] Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune 412115, India

Corresponding Author Email: vidya@sig.ac.in

ABSTRACT

India is one of the leading countries in rapid global infrastructure development. Road infrastructure is one of the major contributors to the same. This raises a need for the real-time maintenance of the developed infrastructure. In maintenance, precise identification and management of potholes are important, considering the safety of citizens. The current study presents a geo-intelligent framework for real-time detection of potholes. It uses advanced deep learning techniques such as PSP-Net and U-Net for pothole detection. It employs high-resolution unmanned aerial vehicle (UAV) imagery, digital surface model (DSM), along with training samples identified through annotations for model training and evaluation. Experimental results show that U-Net outperforms PSP-Net with an F1-score of 0.78, demonstrating high precision in pothole determination. This novel framework is further deployed in the form of a toolkit in the ESRI ArcGIS ecosystem. The two tools developed using the Python API were deployed for the determination of pothole volume and fill quantity estimation, respectively. The American Concrete Institute (ACI) approach was used to estimate the amount of repair materials needed for the identified potholes. The study helps in the reduction of man-hour efforts needed for lengthy field surveys for pothole identification. The Geo-Image Analytics toolbox offers a scalable solution for evolving urban infrastructure needs, marking a significant step forward in modernizing pothole management practices and the sustainability of the road infrastructure.

1. INTRODUCTION

Potholes are a significant issue for citizens, with serious economic and safety impacts. Traditional manual inspection methods for pothole detection are time-consuming and comparatively offer less precision, highlighting the need for innovative solutions. High-resolution UAV imagery combined with deep learning techniques and mobile mapping platforms enables automated pothole detection. These technologies can further aid in effectively managing the recognized potholes. Geospatial data provides accurate, location-wise insight, which is essential for mapping, monitoring, and analyzing the road infrastructure. When combined with geospatial data and deep learning, it triggers the automated generation of road damage, traffic patterns, and infrastructure degradation. This integration of both technologies enhances decision-making, supports predictive maintenance, and reduces operational costs by allowing timely interventions, ultimately improving road safety, efficiency, and longevity [1]. The researchers have applied high-resolution geospatial data combined with the road-infrastructure stress indices and/or advanced machine learning techniques for understanding, maintenance, and sustainability of the road infrastructure.

The GIS-based maintenance model developed by Pantha et al. [1], prioritizes pavement and roadside slope stability,

creating maps for maintenance and planning. It enhances decision-making, improving road system serviceability in the Himalayan regions using the International Roughness Index (IRI) and a roadside slope maintenance priority map. Chandimal and Wijesekera [2] addresses transportation issues like traffic congestion and financial burdens, using GIS to identify improvements such as road surface upgrades and alternative paths. It analyzes spatial accessibility in Kaduwela, Colombo, considering road variation, condition, and population. A novel method to monitor road surface conditions of stone pavements using AWZ index in cities is proposed in the study [3]. The results were compared with typical pavement indicators like IRI and PCI. The method was found to be reliable for assessing stone pavement conditions on the entire urban road network. (Hamdi et al. [4] applied Artificial Neural Networks (ANNs) to predict Surface Distress Index (SDI) values with a high correlation factor (R2 score 0.996) and reveal the influence of key input parameters for SDI prediction, such as rutting, crack width, crack area, patching, pothole, and depression. In the study [5], CNN model for road distress mapping using high-resolution remote sensing images was proposed. The model was tweaked to extract finer features, increasing the likelihood of accurate prediction.

Anusree and Rahiman [6] used a stereo vision-based approach for the identification of potholes and estimation of

the volume for identified potholes using disparity maps and 3D point clouds. Kharel and Ahmed [7] used YOLOv5s to identify the potholes in real-time with high precision using high and low resolution images. The study by Lakshmi Kumari et al. [8] explored the use of YOLOv7 and Faster R-CNN with ResNet-50 as a backbone to automate real-time pothole identification. Biçici and Zeybek [9] used three-dimensional point cloud data to automatically detect and measure road distress using UAV images.

The use of CNN models for pothole identification and prediction using biophysical variables was discussed in the study [10]. A pyramid scene parsing network (PSP-Net) to detect the potholes with high-quality results on various datasets such as ImageNet, PASCAL VOC 2012, and Cityscapes benchmarks was discussed in the study [11, 12].

The comprehensive literature review shows that although geospatial techniques that use geospatial data and remote sensing are widely used for pothole identification, the studies that use high-resolution UAV imagery data are limited [13-20]. The review also shows that the estimation of identified pothole volumes has not been considered by researchers. Moreover, the integration of pothole identification systems with GIS software such as ArcGIS for development of integrated tool is observed rarely [21-23]. This lays a strong foundation to amalgamate advanced deep learning techniques with UAV imagery using GIS techniques for precise pothole identification and management. ESRI ArcGIS provides the facility of accessing and managing geospatial data using various tools [24].

From the comprehensive analysis of contemporary studies in pothole identification, it is observed that studies that integrate pothole identification, pothole volume and fill quantity estimation are not available in the easily accessible and deployable formats. Moreover, it is perceived that surveying such potholes manually is slower and inaccurate since it is difficult to measure the volume of the pothole due to its irregular shape. This necessitates the need for a system that automates the detection and maintenance of potholes through an integrated framework.

Considering the limitations of existing works, this study attempts to design and develop a toolbox to detect potholes in real-time settings with the aid of segmentation architectures: PSP-Net and U-Net [12]. Two novel tools, namely 1) Pothole Volume Estimation and 2) Pothole Fill Quantity Estimation, were developed. This affirms integration of identification and fill quantity estimation of potholes in a seamless manner.

The following sections discuss methodology, followed by results and discussion. The main findings and limitations of the proposed method, along with future scope, are discussed in the conclusion.

2. METHODOLOGY

The methodology of the proposed study involved two steps. The first step involved gathering DSMs and high-resolution orthomosaic UAV imagery of 1cm spatial resolution captured from Florida, USA, obtained using OpenAerial Map [13], preprocessing and CNN model development. In the second step the development of tool for pothole volume and fill quantity estimation.

2.1 Dataset preparation

The high-resolution orthomosaic UAV imagery was obtained using OpenAerial Map [25] with a ground sampling distance (GSD) of 1cm/pixel. Table 1 presents the details about imagery acquired. The obtained image was preprocessed to reduce noise, normalize, and align the data for consistency. To prepare the dataset, road images with potholes were identified through manual inspections and annotated. These annotations were used for training and evaluating the segmentation models.

The dataset of annotated bounding boxes of size 6000 was created from the UAV orthomosaic imagery. It was divided into training (80%) and testing (20%) datasets for model training and validation. The detailed methodology of the proposed work is shown in Figure 1.
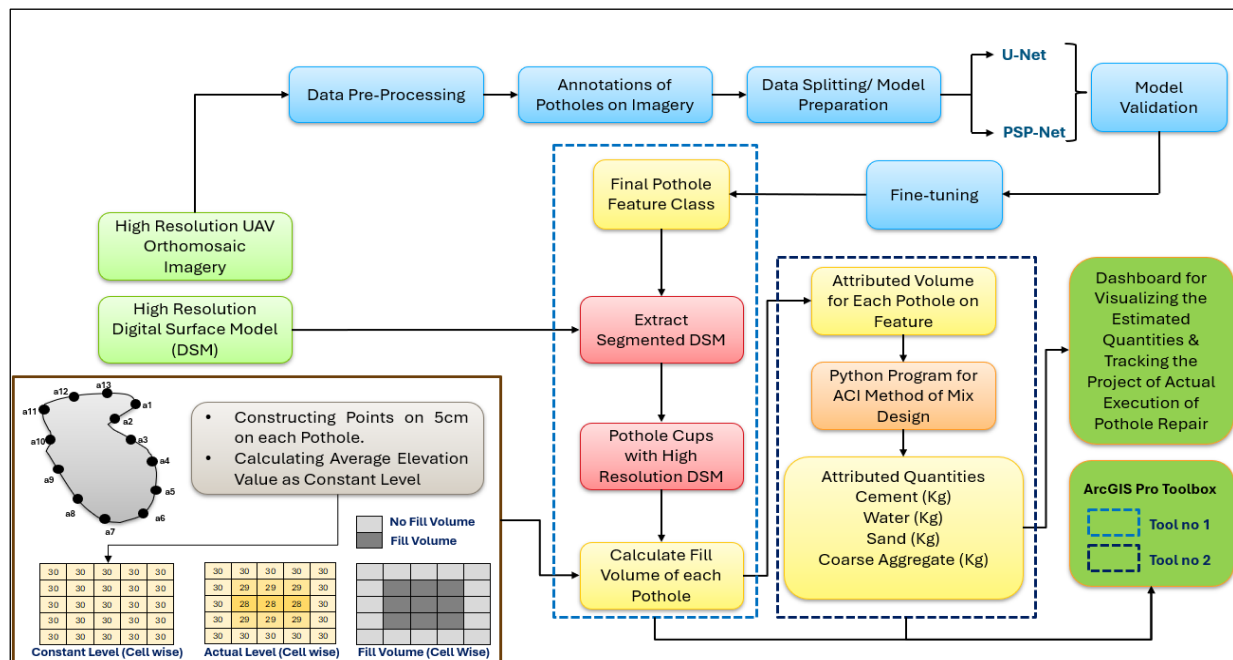

Figure 1. Methodology

Table 1. Dataset details

Parameter	Particulars
Acquisition Date	07-11-2018
Location	Florida, USA
Flight Altitude	60m
Imaging Angle	0^0 (nadir)
Ground Sampling Distance (GSD)	1cm/pixel

2.2 Model design and validation

2.2.1 Pothole identification

For the pothole identification model development and validation, two segmentation architectures U-Net and PSP-Net were employed using the ArcGIS Python API.

The U-Net model [12, 26] excels at capturing fine details and spatial hierarchies, which makes it suitable for pothole detection in images of road infrastructure. The expanding and contracting paths of the model ensure precise localization and context capture for accurate pothole segmentation. PSP-Net model [11] divides the road surfaces in order to extract the features and train the model using a pyramid pooling module.

With its encoder-decoder architecture, U-Net is an excellent tool for collecting the small details of potholes, and PSP-Net improves the comprehension of the global context of a picture.

The experimentation was undertaken in iterative trails and fine-tuning of hyperparameters until an optimal performance. Table 2 presents the details of hyperparameters used while training the models.

Table 2. Hyperparameter configuration

Hyperparameter	Particulars
Learning Rate	0.001
Batch Size	08
Epochs	100
Optimizer	Adam
Activation	ReLU, Sigmoid
Split Ratio	80:20

The model with the best evaluation performance on the validation dataset was selected for deployment. The final pothole feature class was created using segmentation results obtained from DL techniques, which shows details of the spatial extent and position of identified potholes.

2.2.2 Pothole volume and quantity fill estimation

The segmented DSMs are extracted using the identified feature class and a high-resolution DSM, which represents elevation data essential for calculating the material needed for repairs. A volume computation technique involves creating points within each pothole at 5 cm intervals, comparing actual and constant levels of elevation to determine fill volume. This calculation helps assign a volume property to each pothole, which is critical for resource allocation.

2.3 Geo-Image Analytics toolbox

A crucial component of the proposed work was creating a Python toolbox in ArcGIS Pro. It will automate the operations of estimating pothole volume, quantity fill, and calculating repair quantities. Figure 2 presents the workflow for the development of this toolbox. Two tools, namely 1) Pothole Volume Estimation (Figure 3) and 2) Pothole Fill Quantity Estimation using the ACI method (Figure 4), were developed.

2.3.1 Pothole volume estimation tool

Each Python script from the toolbox analyzed the segmented photos to determine each pothole's measurements and estimate the volume of each pothole. In this stage, feature class (the boundary points) of identified potholes using segmentation model along with the elevation details extracted through DSM are fed as input parameters as a raster layer (Figure 3).

The volume for the identified pothole is calculated using geometric calculations based on the obtained measurements.

2.3.2 Pothole quantity fill estimation tool

Using the ACI approach [27], a Python script was created to estimate the amount of repair materials needed. It calculates the quantity of concrete or asphalt needed for repairs by using the pothole volume and the ACI criteria, as shown in Figure 5. The input parameters required at this stage are the identified pothole boundaries, estimated volume, desirable concrete strength and water to cement ratio for calculating the fill quantity. The output generated is an updated pothole feature class with details such as requirement for cement, sand, water and other aggregate mixture as derived using ACI criteria. To guarantee precise estimations, the script incorporates parameters for material density and other pertinent aspects. The detailed procedure adopted for pothole fill quantity estimate using ACI method [27] is as follows:

The total volume of concrete (V) (Eq. (1)) is assumed to be 1m³, and the sum of the absolute volumes of all ingredients should equal 1:

$$V = \frac{W}{\rho_w} + \frac{C}{\rho_c} + \frac{FA}{\rho_{fa}} + \frac{CA}{\rho_{ca}} + V_{air} \qquad (1)$$

where,

- W: weight of water (kg)
- C: weight of cement (kg)
- FA: weight of fine aggregate (kg)
- CA: weight of coarse aggregate (kg)
- ρ: specific gravity $\times$ 1000 (kg/m³)
- V_{air}: air content (typically 2% for non-air-entrained concrete)
1. Choose target strength based on grade.
   (e.g., M25 $\rightarrow$ 31.6 MPa target).
2. Select water to cement (W/C) ratio (e.g., 0.45).
3. Estimate water content (e.g., 160 kg for 1 cm slump).
4. Calculate cement content using Eq. (2).

$$C = \frac{W}{W/C} \qquad (2)$$

5. Estimate air content (e.g., 2%).
6. Determine coarse aggregate volume from reference ACI tables based on max size and fineness modulus.
7. Compute fine aggregate volume by subtracting all known volumes from 1 m³.
8. Convert volumes to weights using specific gravities as shown in Eq. (3).

$$Absolute\ Volume = \frac{material\ mass}{(material\ relative\ density * water\ mass)} \qquad (3)$$

The python toolbox was developed using ArcGIS Pro 3.1 with the support of ArcGIS API for python, and the python libraries ArcPy and Pandas for processing of geospatial data, model development, integration and deployment.
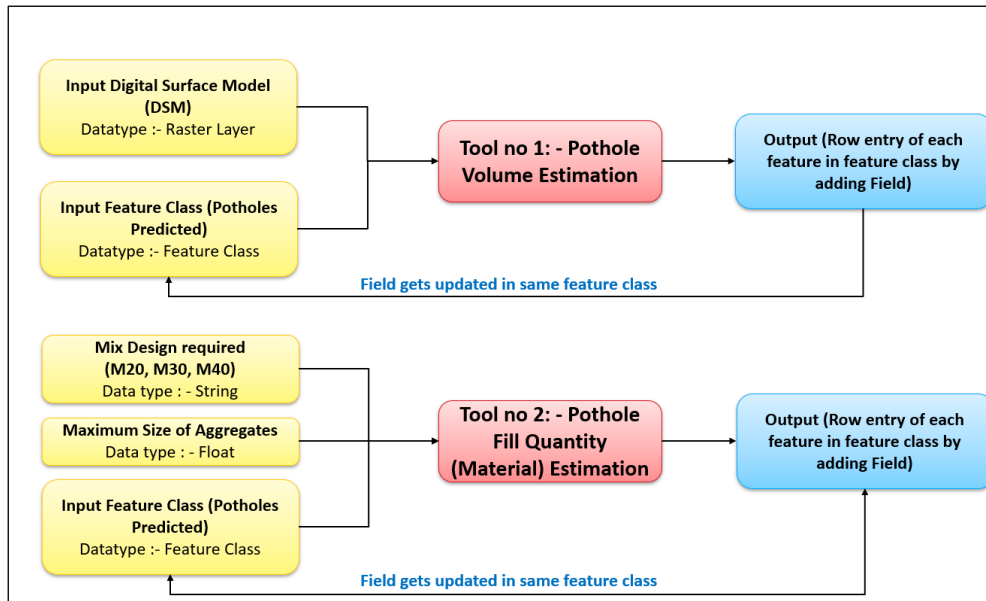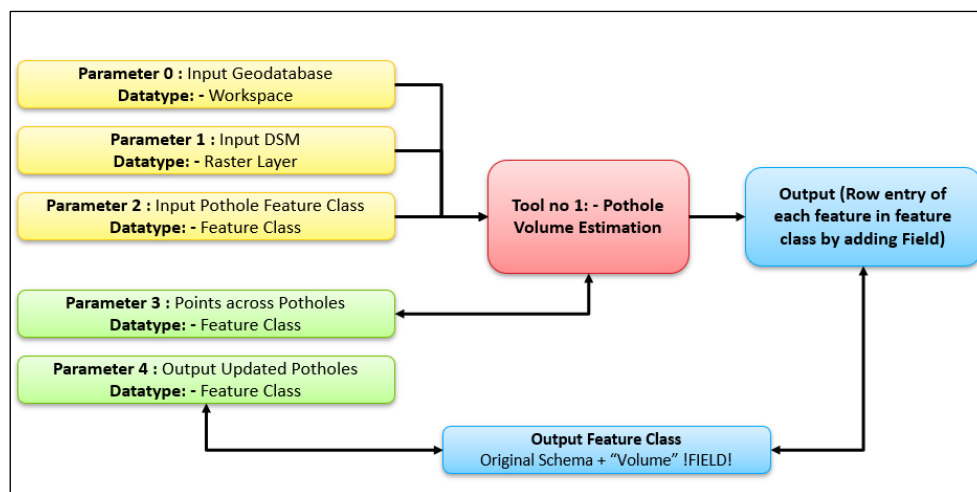
Figure 2. Workflow for toolbox development
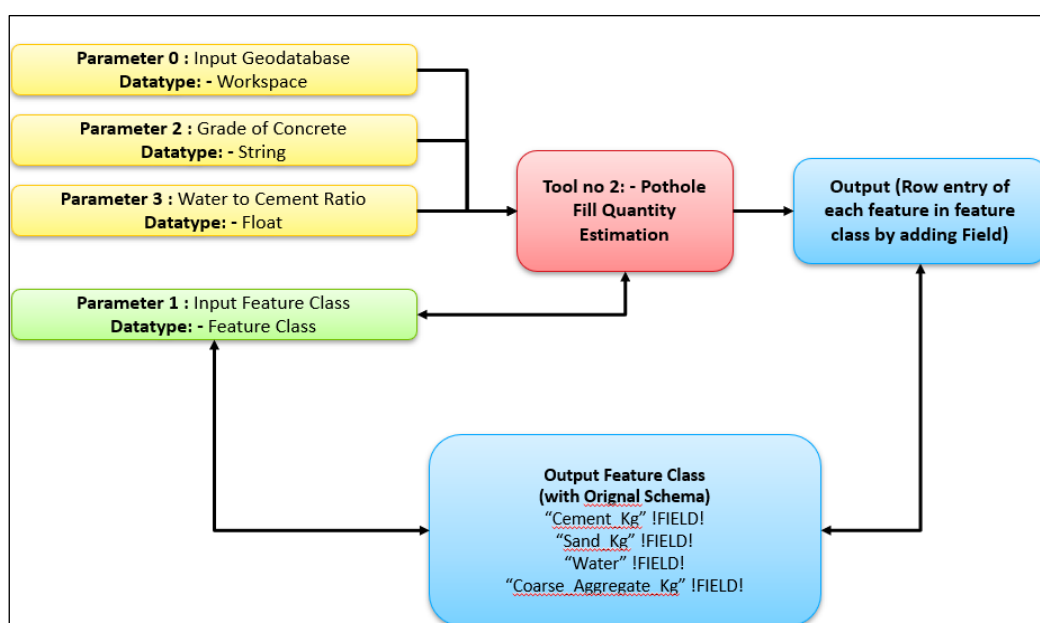

Figure 3. Tool for pothole volume estimation

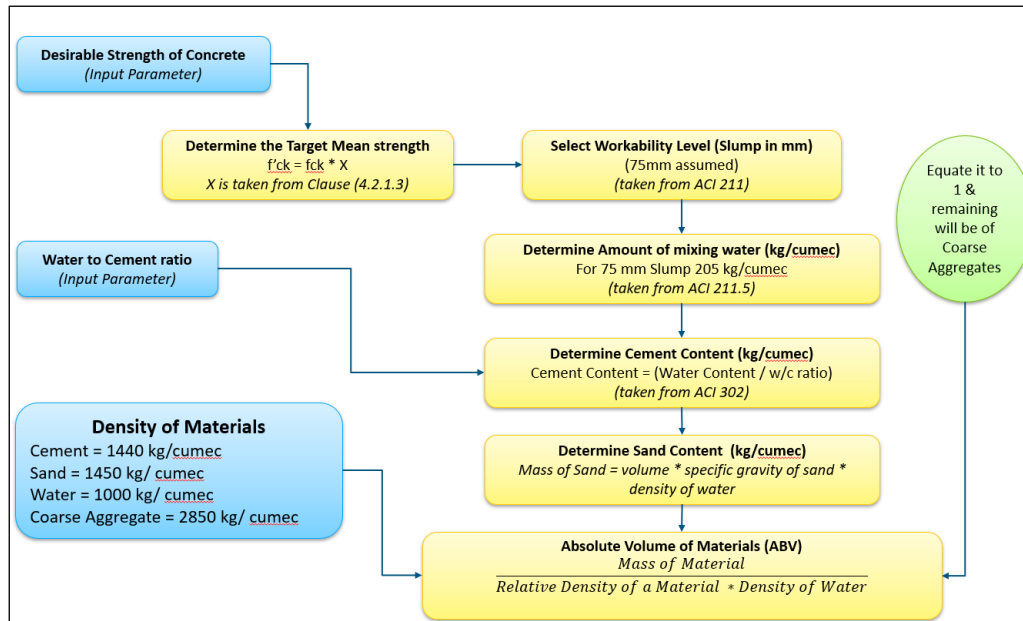**Figure 4.** Tool for quantity estimation using ACI method

Figure 5. ACI method of mix-concrete design for pothole repair

3. RESULTS AND DISCUSSIONS

The performance of CNN models was evaluated using metrics like pixel accuracy, precision, recall, F1-score, and intersection over union (IoU). It is observed that, with a recall of 0.76, precision of 0.82, accuracy of 0.72, IoU of 0.68, and an F1 score of 0.78, the U-Net model performed well. The PSP-Net model provided an F1 score of 0.61, an IoU of 0.76, an accuracy of 0.85, a precision of 0.58, and a recall of 0.72. The above findings demonstrated the trade-offs between various model properties, indicating the U-Net model performs better in terms of precision and F1-score, and the PSP-Net model demonstrates higher accuracy and IoU (Table 3).

While PSP-Net achieves a higher IoU than U-Net, it records a lower F1-score, pointing to an important trade-off between region-level accuracy and pixel-wise recall. The architecture of PSP-Net, with its multi-scale context aggregation through pyramid pooling, tends to produce smoother and more conservative segmentations. This often leads to better overlap with the ground truth as measured by IoU but can miss the finer details/smaller structures, thus reduce the recall and lower the F1-score. On the contrary, U-Net benefits from direct skip connections that help retain spatial details, making it more effective at capturing object boundaries and smaller regions. As a result, it tends to achieve higher recall, which contributes to an improved F1-score. Thus the selection of model in cases where detecting small or minor regions like potholes drives the overall efficacy of the developed system.

The actual high-resolution UAV image with ground truth pothole annotations recorded through manual field visits, and predicted results are shown in Figures 6 and 7, respectively.

The model was trained using the annotated samples of potholes using the high-resolution UAV imagery. From the results, it is apparent that model learned the minute features for correct discrimination of potholes. The accuracy of model is driven by the spatial resolution of the aerial images captured and correctness of the annotations. It can be used across different geographical locations in similar experimental settings and resolution of input images. Further, the calculations can be used for estimation of pothole volume and fill quantity. The factors such as the effects of shadow, lighting and angle of image capture play a vital role in models learnability and robustness [28].

Table 3. Comparative analysis of deep learning models for pothole identification

Model	Recall	Precision	Accuracy	Intersection over Union (IoU)	F1-Score
U-Net	0.76	0.82	0.72	0.68	0.78
PSP-Net	0.72	0.58	0.85	0.76	0.61

An integrated dashboard using ArcGIS Pro was developed to monitor the progress of pothole repair through field updates in real-time. A dedicated data pipeline was designed to fetch the data of roads, edges, and pothole polygons. Figure 8 shows the workflow adapted for designing this integrated dashboard and the obtained results. The dashboard was developed using ArcGIS Dashboard hosted over ArcGIS Online (AGOL) platform. This dasboard hosted a web-map with layers such as: road center line (RDC) viewer, road edges (RE) viewer, pothole points (PEP) viewer etc. Each layer can be fetched separately from the dasboard.Its integration with web-map allows real-time pothole identification and monitoring of repairs.

Field personnel were able to update the status of pothole repairs and material procurement directly from the site. Stakeholders were able to promptly resolve any concerns that developed during the repair work execution by using the dashboards to make educated decisions. The performance of proposed system is constrained due to the effects of lower resolution imagery, variations in the lighting or angle of image capture. In future, this study can be extended with more generalizability and robustness by training the models with more diverse datasets.

Figure 6. High-resolution UAV orthomosaic imagery and ground truth annotations for manually inspected potholes(annoted in yellow boxes)
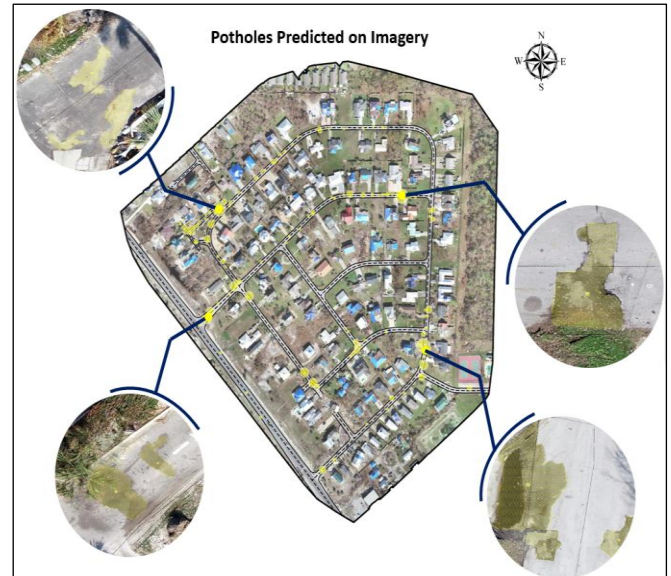
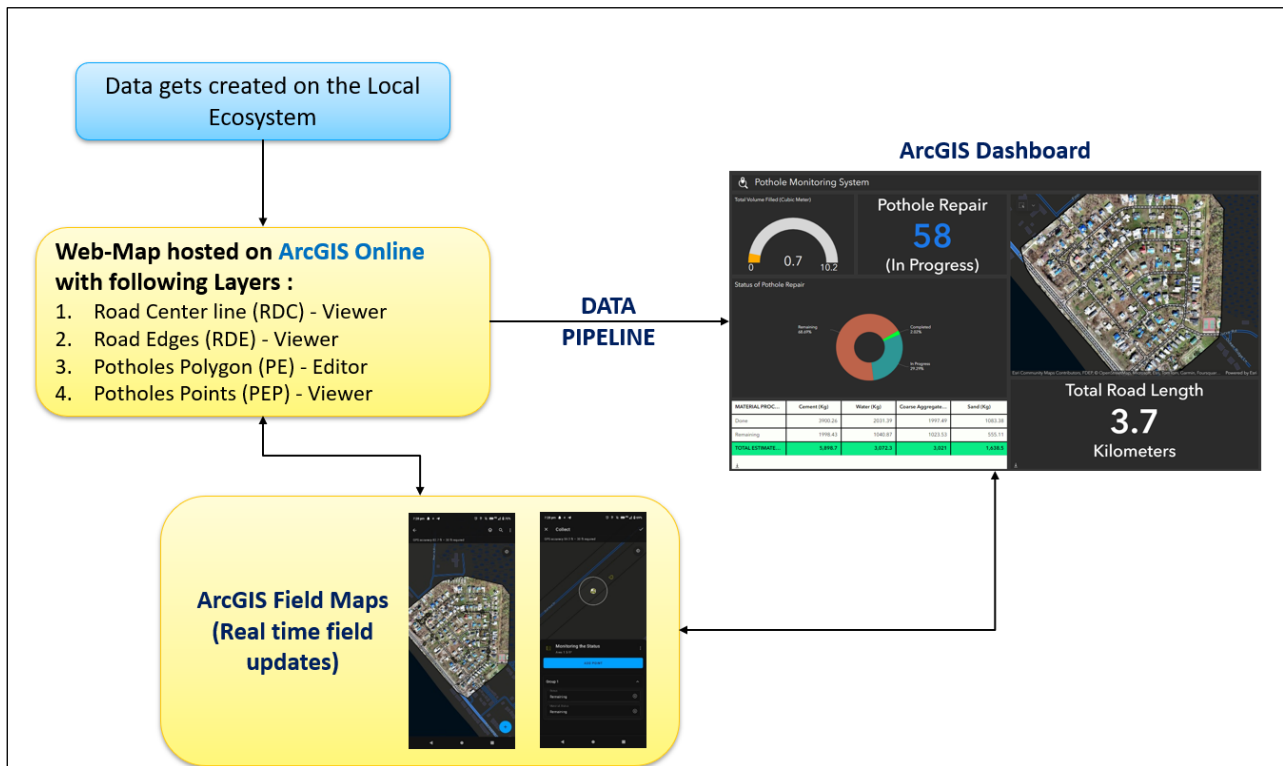

Figure 7. Pothole prediction results

Figure 8. Workflow of pothole monitoring dashboard

4. CONCLUSIONS

The study presented the Geo-Image Analytics toolbox, an innovative approach that integrates deep learning models, geospatial data, and the ArcGIS Python API for pothole identification and maintenance. The proposed system explores the potential of advanced deep learning techniques such as PSP-Net and U-Net for the identification of potholes using high-resolution UAV imagery and DSM. Further, the U-Net model is deployed in the form of a toolbox for ready use by the stakeholders. The proposed method also provides the facility for the calculation of the fill quantity of detected potholes with the help of the ACI method. The DSM model with high-resolution imagery helped to detect small and shallow potholes with improved accuracy. An integrated dashboard is deployed for real-time monitoring of pothole repairs and field maps. With this framework, stakeholders can track the progress of the repair and maintenance of identified potholes, contributing to sustainable road infrastructure management and making informed decisions. It also ensures road safety, infrastructure management, operational efficiency, and improved livelihood of the citizens. This system acts as a prototype for further research that will use technology to manage infrastructure, making community road networks safer and more dependable in the end.

REFERENCES

[1] Pantha, B.R., Yatabe, R., Bhandary, N.P. (2010). GIS-based highway maintenance prioritization model: An integrated approach for highway maintenance in Nepal mountains. Journal of Transport Geography, 18(3): 426-433. https://doi.org/10.1016/j.jtrangeo.2009.06.016

[2] Chandimal, W.P.A.I.M., Wijesekera, N.T.S. (2008). A GIS approach to identify road network improvement needs: Case study of Kaduwela, Sri Lanka. Engineer: Journal of the Institution of Engineers, Sri Lanka, 41(5): 116-125. https://doi.org/10.4038/engineer.v41i5.7111

[3] Bruno, S., Vita, L., Loprencipe, G. (2022). Development of a GIS-based methodology for the management of stone pavements using low-cost sensors. Sensors, 22(17): 6560. https://doi.org/10.3390/s22176560

[4] Hamdi, Hadiwardoyo, S. P., Correia, A. G., Pereira, P., Cortez, P. (2017). Prediction of surface distress using neural networks. AIP Conference Proceedings, 1855(1): 040006. https://doi.org/10.1063/1.4985502

[5] Madhumita, D., Bharath, H.A. (2023). Deep Learning based approach for road distress mapping using VHR images. In IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, pp. 1549-1552. https://doi.org/10.1109/igarss52108.2023.10283229

[6] Anusree, B., Rahiman, V.A. (2021). Pothole detection and volume estimation based on disparity transformation with histogram thresholding. Proceedings of the Yukthi 2021- The International Conference on Emerging Trends in Engineering – GEC Kozhikode, Kerala, India. http://dx.doi.org/10.2139/ssrn.3973013

[7] Kharel, S., Ahmed, K.R. (2022). Potholes detection using deep learning and area estimation using image processing. In Proceedings of SAI Intelligent Systems Conference, pp. 373-388. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-82199-9_24

[8] Lakshmi Kumari, P.D.S.S., Ramacharanteja, G.S.S., Suresh Kumar, S., Bhuvana Sri, G., Jyotsna, G.S.N., Safalya, A.H.K.N. (2023). Developing an automated system for pothole detection and management using Deep Learning. In International Conference on Advanced Communication and Intelligent Systems, pp. 12-22. Cham: Springer Nature, Switzerland. https://doi.org/10.1007/978-3-031-45124-9_2

[9] Biçici, S., Zeybek, M. (2021). An approach for the automated extraction of road surface distress from a UAV-derived point cloud. Automation in Construction, 122: 103475. https://doi.org/10.1016/j.autcon.2020.103475

[10] Lee, S.Y., Le, T.H.M., Kim, Y.M. (2023). Prediction and detection of potholes in urban roads: Machine learning and deep learning based image segmentation approaches.

Developments in the Built Environment, 13: 100109. https://doi.org/10.1016/j.dibe.2022.100109

[11] Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J. (2017). Pyramid scene parsing network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, pp. 2881-2890. https://doi.org/10.1109/cvpr.2017.660

[12] J Ghosh, R., Shah, S.K., Kumbhar, V. (2023). Deep learning-based liver tumor segmentation: A comparative study of U-Net variants for medical imaging analysis. In 2023 Global Conference on Information Technologies and Communications (GCITC), Bangalore, India, pp. 1-7. https://doi.org/10.1109/GCITC60406.2023.10426259

[13] Ling, S., Pan, Y., Chen, W., Zhao, Y., Sun, J. (2024). Pothole detection based on superpixel features of unmanned aerial vehicle images. International Journal of Pavement Research and Technology, 1-11. https://doi.org/10.1007/s42947-024-00436-w

[14] Ruseruka, C., Mwakalonge, J., Comert, G., Siuhi, S., Ngeni, F., Anderson, Q. (2024). Augmenting roadway safety with machine learning and deep learning: Pothole detection and dimension estimation using in-vehicle technologies. Machine Learning with Applications, 16: 100547. https://doi.org/10.1016/j.mlwa.2024.100547

[15] Pan, Y., Zhang, X., Cervone, G., Yang, L. (2018). Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(10): 3701-3712. https://doi.org/10.1109/JSTARS.2018.2865528

[16] Ping, P., Yang, X., Gao, Z. (2020). A deep learning approach for street pothole detection. In 2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, UK, pp. 198-204. https://doi.org/10.1109/BigDataService49289.2020.00039

[17] Ahmed, K.R. (2021). Smart pothole detection using deep learning based on dilated convolution. Sensors, 21(24): 8406. https://doi.org/10.3390/s21248406

[18] Varona, B., Monteserin, A., Teyseyre, A. (2020). A deep learning approach to automatic road surface monitoring and pothole detection. Personal and Ubiquitous Computing, 24(4): 519-534. https://doi.org/10.1007/s00779-019-01234-z

[19] Rastogi, R., Kumar, U., Kashyap, A., Jindal, S., Pahwa, S. (2020). A comparative evaluation of the deep learning algorithms for pothole detection. In 2020 IEEE 17th India council international conference (INDICON), New Delhi, India, pp. 1-6. https://doi.org/10.1109/INDICON49873.2020.9342558

[20] Becker, Y.V.F., Siqueira, H.L., Matsubara, E.T., Gonçalves, W.N., Marcato, J.M. (2019). Asphalt pothole detection in UAV images using convolutional neural networks. In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, pp. 56-58. https://doi.org/10.1109/IGARSS.2019.8900621

[21] Faisal, A., Gargoum, S. (2025). Cost-effective LiDAR for pothole detection and quantification using a low-point-density approach. Automation in Construction, 172: 106006. https://doi.org/10.1016/j.autcon.2025.106006

[22] Meier, J., Welborn, E., Diamantas, S. (2025). Pothole segmentation and area estimation with thermal imaging

using deep neural networks and unmanned aerial vehicles. Machine Vision and Applications, 36(1): 17. https://doi.org/10.1007/s00138-024-01637-w

[23] Pan, Z., Guan, J., Yang, X., Guo, A., Wang, X. (2024). 3D profile-based pothole segmentation and quantification. International Journal of Hydromechatronics, 7(1): 16-30. https://doi.org/10.1504/IJHM.2024.135980

[24] Johnston, K., Ver Hoef, J.M., Krivoruchko, K., Lucas, N. (2001). Using ArcGIS geostatistical analyst (Vol. 380). Redlands: Esri.

[25] OpenAerialMap. OpenAerialMap. https://openaerialmap.org/.

[26] Ronneberger, O., Fischer, P., Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, pp. 234-241. Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

[27] Chhachhia, A. (2021). Concrete mix design by IS, ACI and BS methods: A comparative analysis. Journal of Building Material Science, 2(1): 30-33. https://doi.org/10.30564/jbms.v2i1.2636

[28] Sun, Q., Qiao, L., Shen, Y. (2025). Pavement potholes quantification: A study based on 3D point cloud analysis. IEEE Access, 13: 12945-12955 https://doi.org/10.1109/ACCESS.2025.3531766