JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718

VOLUME . 4 ISSUE . 1 (2025)

REAL-TIME OBJECT DETECTION FOR AUTOMATED CONSTRUCTION MATERIAL MANAGEMENT

Aqeel Ahmad
School of Computer Science, Minhaj University Lahore, Punjab, Pakistan
***Gulzar Ahmad**
School of Computer Science, Minhaj University Lahore, Punjab, Pakistan
Khalid Masood
School of Computer Science, Minhaj University Lahore, Punjab, Pakistan
Zahid Hasan
Department of Computer Science, National College of Business Administration & Economics, Lahore, Pakistan
Muhammad Mudassar Naveed
School of Computer Science, Minhaj University Lahore, Punjab, Pakistan
Muhammad Sajjad
School of Computer Science, Minhaj University Lahore, Punjab, Pakistan
Adeel Khan
School of Computer Science, Minhaj University Lahore, Punjab, Pakistan
Arslan Ejaz
School of Computer Science, Minhaj University Lahore, Punjab, Pakistan

RECEIVED	ACCEPTED	PUBLISHED
12 Feb 2025	23 Feb 2025	28 March 2025

ABSTRACT
Object detection in construction sites is now replacing traditional methods of quality control and material management. Artificial Intelligence (AI)-driven systems are now being used to detect objects, classify them, and evaluate construction materials using deep learning algorithms. These algorithms enhance the material quality assurance of operational activities. Conventional methods for construction materials are time-consuming, labor-intensive, and prone to error. While the AI-driven approaches are flexible, scalable, and reduce the cost of achieving high accuracy. This study shows the innovative role of object detection in the construction materials industry, emphasizing its benefits, applications, challenges, and future potential. This study proposes the use of a Deep learning -based model, Yolov11, which enhances the capability of real-time object detection by offering a high overall accuracy of 94.3 % precision of 96.3%, thereby enhancing the automated construction site monitoring and construction material management.

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

Introduction

Construction materials are important components used in building and infrastructure, such as dams, roads, and bridges. These materials give us the fundamentals and final elements of a construction project. These materials play a vital role in ensuring structural integrity, cost, sustainability, and quality of a project. Selecting the right materials improves the durability and building safety, reduces building maintenance costs, and helps complete projects within set timelines. There are many types of construction materials, such as natural, man-made, and modern categories. Natural materials contain wood, clay products, and stones, while man-made materials contain cement, concrete, glass, steel, and plastic. Modern materials involve fiber-reinforced polymers, self-healing concrete and bamboo, and fly ash bricks. These materials promote environmental sustainability.

Construction is an important critical sector in world economic development, particularly in urbanization, infrastructure, and industrialization [1]. The well-organized management and detection of construction material are essential for ensuring project quality, time delivery, and cost effectiveness [2]. Traditionally, this procedure has relied on manual work by humans, which results in error-prone and time-consuming, and is subject to variations due to adverse environmental conditions [3]. Therefore, the construction industry needs the integration of automated technologies, particularly with AI, to address the challenges.

Construction sites are more complex environments where various materials, such as steel bars, bricks, pipes, wooden planks, and cement bags, are handled [4]. Misplacement of these materials can affect the cost and increase the delivery time of buildings. Traditional techniques involve human supervision or Radio frequency detection (RFID) systems [5], both have their use limitations. Manual methods involve humans, while sensor-based methods do not work in adverse environments. So, they need a real solution to overcome these limitations. Artificial Intelligence (AI) is playing a significant role in medical [6], road

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

sign detection [7], agriculture [8], and various types of industry.

Recent improvements in AI, particularly in DL, have developed visual recognition tasks. DL models not only detect the objects but also describe their location [6]. Deep learning models, particularly Convolutional Neural Networks (CNNs), have proved outstanding performance in object detection and classification, but have limited use because of the complex environment and lack of real-time object detection [9].

Among the many detection techniques, You Only Look Once(YOLO) [10] stands outclass because of its speed and accuracy, making it suitable for real-time object detection. Yolo has passed through multiple iterations, each enhancing its speed, infrastructure, computational power, and accuracy. Its single-pass capability enables the detection and classification of objects in one shot, making it ideal for environments where quick decisions are required. The Improved version, Yolov11, is efficient for deployment in real-time construction sites because of its lightweight, anchor-free mechanism, and improved features. The goal is to create a robust model that may be deployed on construction sites using edge devices. Furthermore, the implementation of this system at construction sites may track the inventory, reduce the waste, improve the safety compliance minimize human errors.

Related Work:

The construction material detection and monitoring have relied on human work, manual inspections, check-list-based auditing, and sensor-based scanning [11]. Manual inspection can also be affected at start, it is error-prone due to human judgment and inconsistencies in detecting similar-looking materials. The process is labor-intensive and not scalable when applied to large construction sites [12].

Sensor-based systems, such as RFID and barcode scanning, have been deployed to track material supply chains [13]. This system requires tags attached to each unit, which can be harsh in a construction environment. Furthermore, the hardware of scanning and human iterations limits them from automation. Image processing methods using rule-based techniques and handcrafted

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

features have been explored in earlier works [14]. But these methods lack robustness in a real-time, unstructured environment. Variations in environment, occlusion, background clutter, and camera angles reduced the detection accuracy of the system [14].

With the invention of Machine learning (ML), researchers moved toward classifiers trained on picture features extracted using algorithms such as Local Binary Pattern, Histogram of Oriented Gradient, Scale-invariant feature extraction [15], Decision Tree, and Support Vector Machine [16].

ML models give better results than rule-based systems, particularly when trained on annotated datasets [17]. The used SVM classifier separates items, reporting over 80% accuracy. Similarly, Random Forest is also being used in industry for differentiating construction materials [18]. Despite improvements, ML models are limited due to manual feature engineering. The Limitations in Traditional and ML Models paved the way for DL models, especially CNNs, which automatically learn features from pixel data. Early models such as VGGNet and AlexNet were adopted for classification tasks on construction material datasets, but failed localization tasks essential for detecting the objects. This gap was filled by the detector Faster-RNN [19], which integrates regional proposal networks (RPN) with CNN-based classifiers. Due to high computational power, they are unsuitable for real-time deployment on edge devices used on construction sites. Single-shot detector SSD[20] and Retina-Net [21] delivered better speed, but sometimes compromise on accuracy. RetinaNet is used to handle class imbalance.

The Yolo family models emerged as an innovation in real-time object detection. Yolov1 [10] offers real-time object detection but struggles with small object detection. Yolov2 and v3 address many of these limitations through multiscale predictions, anchor boxes, and an improved backbone network like Darknet-53 [22]. Yolov4 architecture was best for low-quality images, Yolov5 had a modulator architecture and a Pytorch implementation [23]. Yolov6 to v8 continued this trend with enhancements in feature, transformation attention, and decoupled head signs [24]. This version provides high accuracy, a real-time object detection system, and reduced latency; however, detection of small objects remains challenging. Yolov11 is the latest evaluation of the Yolo family. It has multiple innovations:

- Anchor-free detection
- Transformer-based feature fusion
- Decoupled heads
- Gradient Harmonizing Mechanism [25]

Comparative Studies of Yolo series:

Table 1 below shows the comparative study of the Yolo series.

Table 1: Comparative studies of the Yolo series

YOLO Version	Backbone Network	mAP (%)	Speed (FPS)
YOLOv3	Darknet-53	~33–35	30–45
YOLOv4	CSPDarknet53	~43.5	50+
YOLOv5	CSPDarknet (custom)	~45–50	70–140
YOLOv6	EfficientRep Backbone	~52–55	120+
YOLOv7	E-ELAN	~56–57	150+
YOLOv8	Custom Backbone (CNN-based)	~56–60	180+
YOLOv11	Transformer-CNN hybrid	~60–65	200+

Methodology:

Methodology involves the data acquisition, preprocessing, model selection, model training, evaluation, and deployment in the real world.

Dataset Collection and Annotation:

The dataset used in this study was sourced from Roboflow, a free repository for public use, which provides thousands of computer vision project datasets. The dataset consists of HD-quality images containing various construction materials, such as occlusion, and background conditions. Each images are annotated by annotated tools. These annotated images are then exported to the YOLO format, including the class label, normalized bounding box coordinates, and image dimensions.

Data preprocessing and data distribution:

The dataset passes through many stages before training the model.

Image resizing: All the images are resized to 640*640

Augmentation: This technique is used to increase the picture by rotating and flipping.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

Moto is to increase the robustness and generalizability of the model.

Normalization: Normalization is also done on each image pixel to improve the performance of the model.

The dataset is split into three subsets: 70% for training, 18% for validation, and 12 % for testing, as shown in Figure 1.

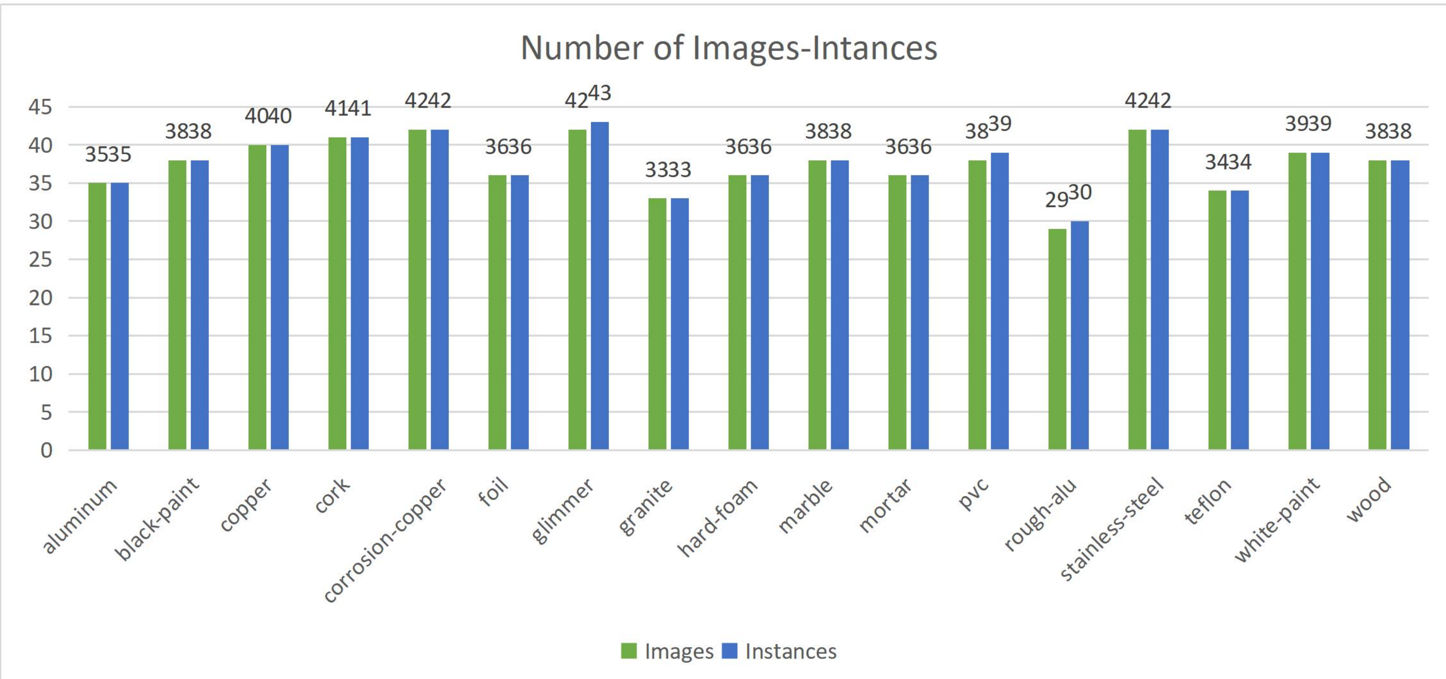


Figure 1: Data distribution

Yolov11 model architecture:

The advancements in YOLOv11 from the previous version are given below.

- Anchor-free detection
- Transformer-based feature fusion
- Decoupled detection heads

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718

VOLUME . 4 ISSUE . 1 (2025)

Model Framework:

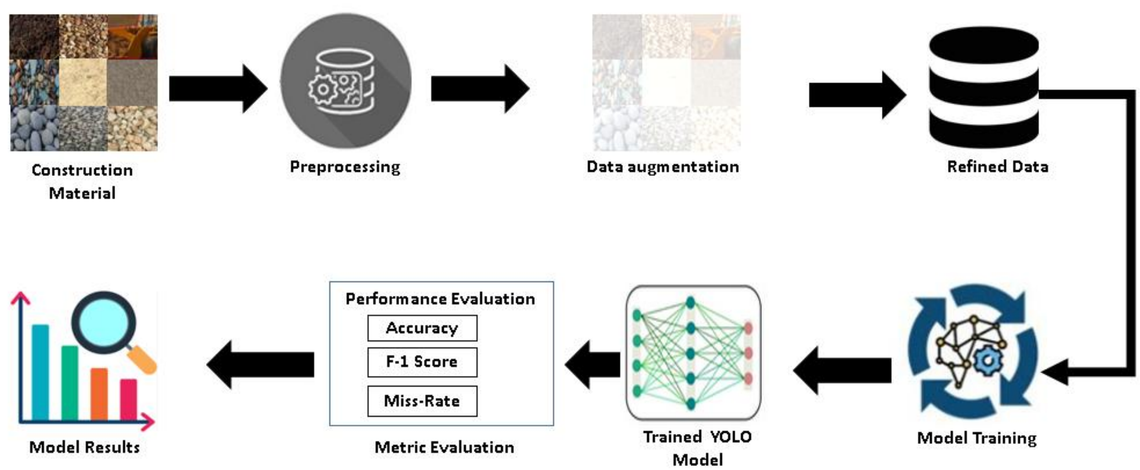


Figure 2: Proposed System

Figure 2 represents the proposed construction material works using a Yolo-based DL pipeline to detect the material automatically, such as Aluminum, rubber, wood, and pipes etc. The workflow of the models starts with image data acquisition, which consists of raw images. The images pass through different preprocessing steps such as resizing, augmentation. The main goal is to make the model perfect to improve the robustness and generalizability of the model. Integration of original and augmented images continues to refine the data, which is used to train the Yolo model to detect the objects. During the training phase, it learns the classify the objects in an adverse environment and complex background. After the completion of training, performance was evaluated with Accuracy, F-1 score, and Miss Rate. These metrics detect the reliability, precision, and error tendencies. Finally, the model results are visualized, enabling the model to be deployed in real-time construction monitoring.

Performance Metrics:

These metrics are used to evaluate the model's performance.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718

VOLUME . 4 ISSUE . 1 (2025)

Table 2: Performance metrics

Metrics	**Formula**
Accuracy	tp+fp/tp+fp+tn+fn
Precision	tp/tp+fp
Recall	tp/tp+fn
F-1 Score	2*pre*Rec/pre+Rec

Confusion Metric: A metric that is used for the overall classification of the classes. The confusion metric of the proposed model is shown below in Table 3.

Table 3: Confusion Metric

Predicted \ True	aluminum	black-paint	copper	cork	corrosion-copper	foil	glimmer	granite	hard-foam	marble	mortar	pvc	rough-alu	stainless-steel	teflon	white paint	wood
aluminum	33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
black-paint	1	32	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
copper	0	1	35	0	0	0	0	0	0	0	0	1	0	0	0	1	0
cork	0	0	0	39	0	0	0	0	0	0	0	0	0	0	0	0	0
corrosion-copper	0	0	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0
foil	0	0	0	0	0	39	0	0	0	0	0	0	0	0	0	0	0
glimmer	0	0	0	0	0	0	38	0	0	0	0	0	0	0	0	0	0
granite	0	0	0	0	0	0	0	38	0	0	0	0	0	0	0	0	0
hard-foam	0	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0
marble	0	0	0	0	0	0	0	0	0	36	0	0	0	0	0	1	0
mortar	0	0	0	0	0	0	0	0	0	0	36	0	0	0	0	0	0
pvc	0	0	0	0	0	0	0	0	0	0	0	35	0	0	0	0	0
rough-alu	0	0	0	0	0	0	0	0	0	0	0	1	35	0	0	0	0
stainless-steel	0	0	0	0	0	0	0	0	0	0	0	0	0	39	0	0	0
teflon	0	0	0	0	0	0	0	0	0	0	0	0	0	0	33	0	0
white-	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	35	0

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

paint																	
wood	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	35

Accuracy Graph:

The accuracy curve shows a consistent performance of the model throughout the training. Initially, the model quickly learns the basic features. An abrupt change occurs from 0.02 to 0.8 mAP in early epochs. Mid-training shows the model refines its accuracy and learn slowly. After 50 epochs, the model gains an accuracy of 94%, indicating no room for further gains. Minor fluctuations occur near the end of the training, reflecting high model performance. Overall, the curve shows that the model is perfect and well-trained for object detection tasks, as shown in Figure 3.

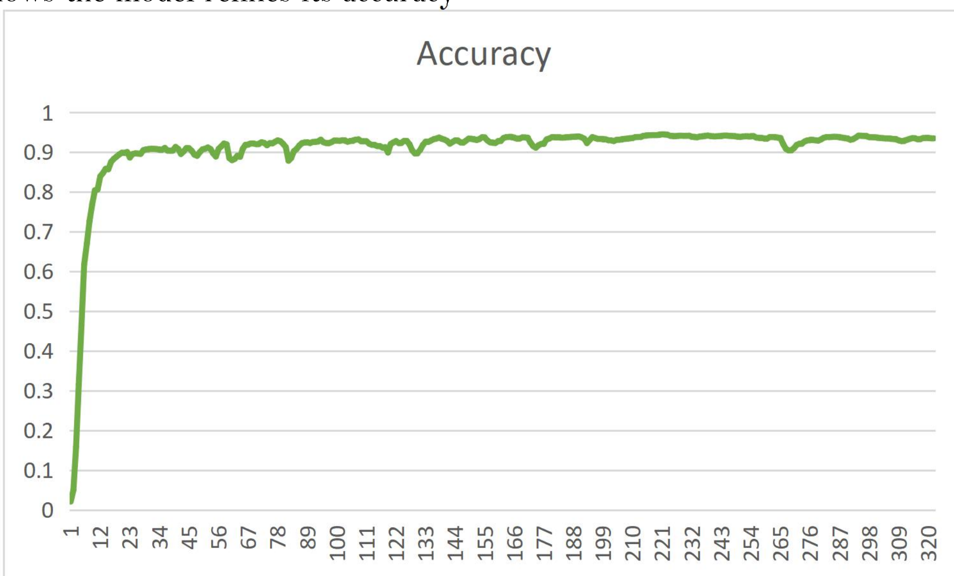


Figure 3: Accuracy Graph:

Precision & Recall Graph:

The material classification model shows high overall performance, with a recall of 0.90, a precision of 0.963, and an F-1 score of 0.93. Most of the classes show excellent performance, showing minimal loss of misclassifications. Many classes show balanced precision and recall, offering stable and reliable predictions. However, a few classes, such as PVC and rough-alu, show lower F-1 scores, indicating the model

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

struggles. This is due to imbalanced data. Figure 4 shows the precision and recall curve.

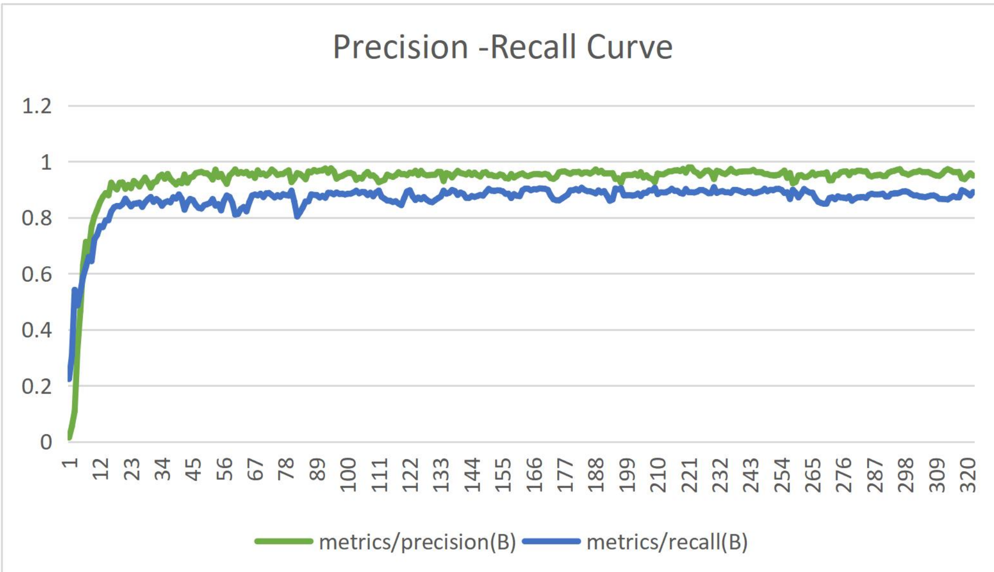


Figure 4: Precision and Recall Graph

F-1 Score Graph:

F-1 score curve values, as shown in Figure 5, show the trend of model training. The value reached 80% after 50 epochs, and 90% after 70 epochs, indicating the best gains in classification tasks. The model maintains a stable F-1 score, achieving at 93% value. The curve shows the effective learning of the model, consistent enhancements, and highest accuracy by the end of the training.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

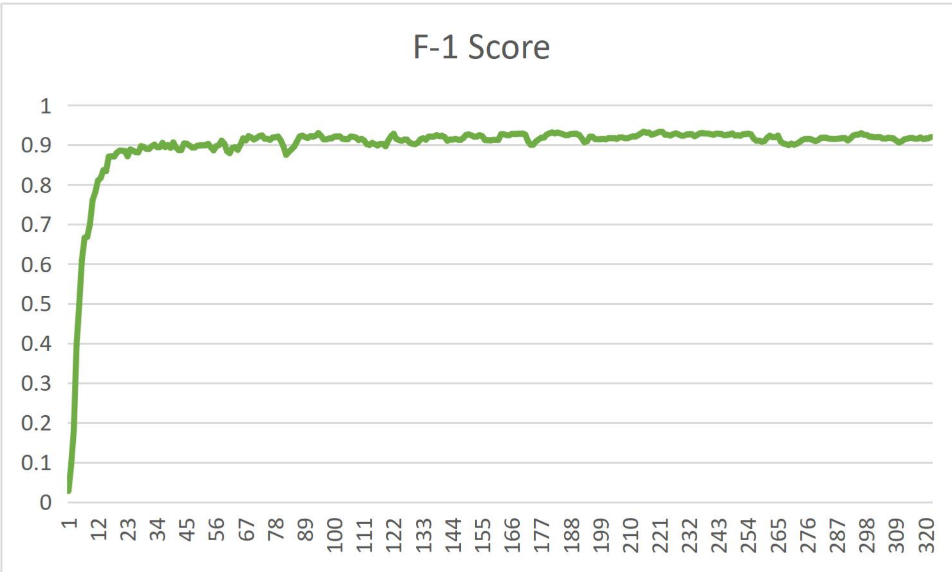


Figure 5: F-1 Score Curve Graph

Training Loss & Validation Loss Curve:

The training and validation class loss curve shown in Figure 6 represents generalization and effective learning of the model. Initially, both losses are very high but decrease steadily over time. The training and validation losses dropped from 4.78 to 0.25 and from 7.5 to 0.34, respectively, indicating the best model learning rate and avoiding overfitting of the model. This trend shows that the model performs well on the unseen dataset with monitored losses.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

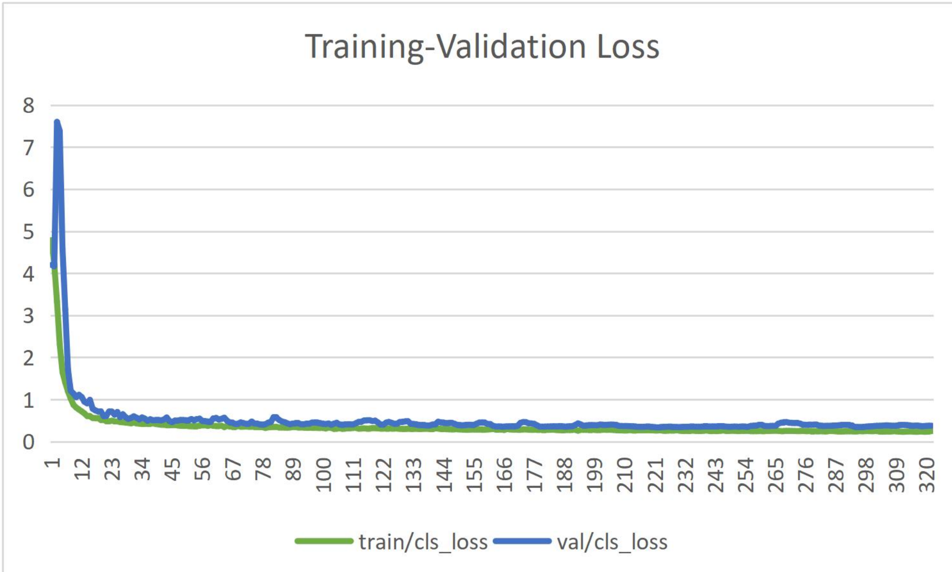


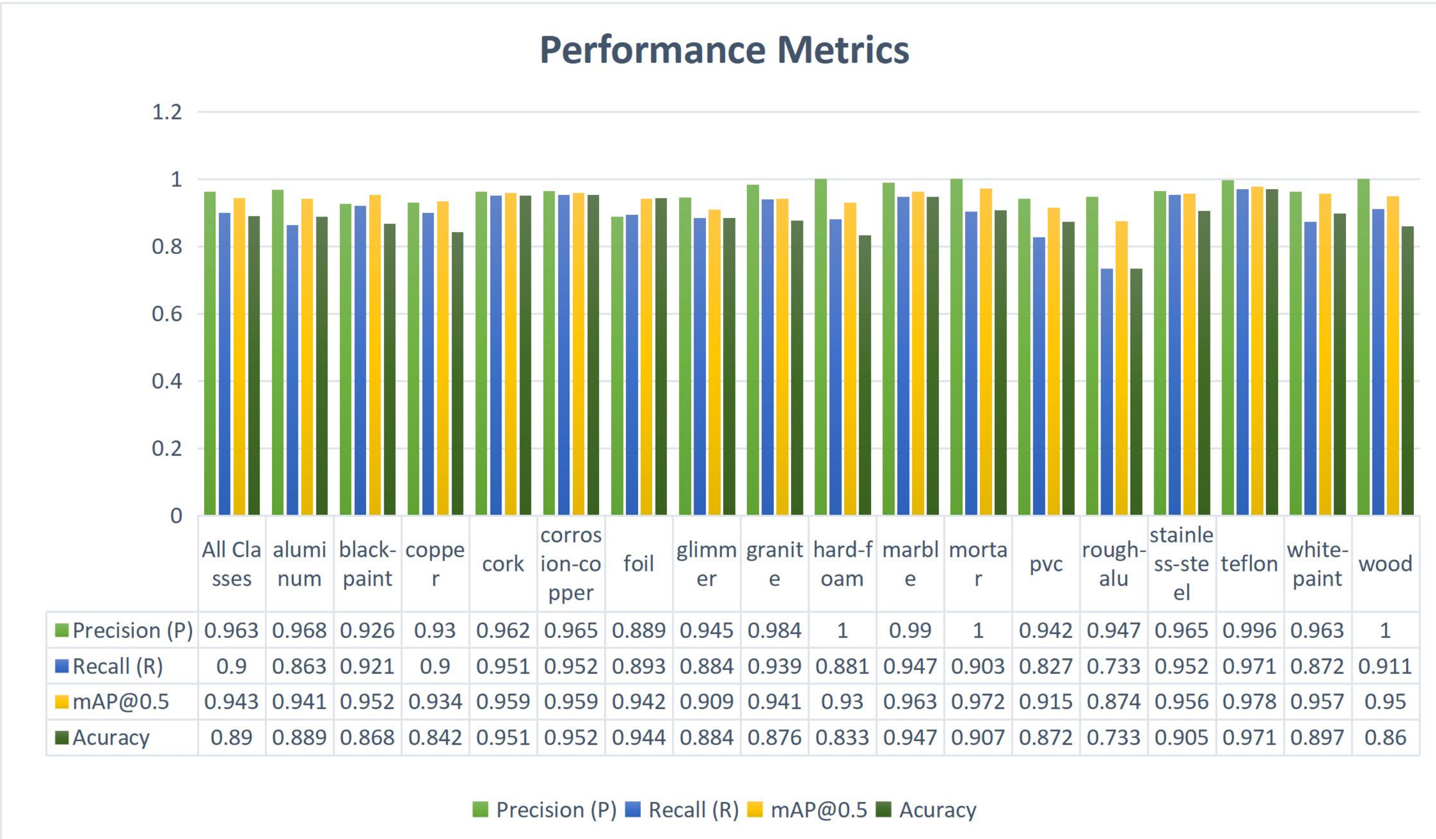
Figure 6: Training and Validation Loss Curves

Results of Yolov11 Model:

The proposed Yolov11 model performs strongly in the classification of construction materials, with high recall, precision, and accuracy as shown in Figure 7. Overall, the model is suitable with minor improvements needed for certain materials.

JOURNAL OF EMERGING TECHNOLOGY AND DIGITAL TRANSFORMATION
ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718

VOLUME . 4 ISSUE . 1 (2025)



	All Classes	aluminum	black-paint	copper	cork	corrosion-copper	foil	glimmer	granite	hard-foam	marble	mortar	pvc	rough-alu	stainless-steel	teflon	white-paint	wood
Precision (P)	0.963	0.968	0.926	0.93	0.962	0.965	0.889	0.945	0.984	1	0.99	1	0.942	0.947	0.965	0.996	0.963	1
Recall (R)	0.9	0.863	0.921	0.9	0.951	0.952	0.893	0.884	0.939	0.881	0.947	0.903	0.827	0.733	0.952	0.971	0.872	0.911
mAP@0.5	0.943	0.941	0.952	0.934	0.959	0.959	0.942	0.909	0.941	0.93	0.963	0.972	0.915	0.874	0.956	0.978	0.957	0.95
Acuracy	0.89	0.889	0.868	0.842	0.951	0.952	0.944	0.884	0.876	0.833	0.947	0.907	0.872	0.733	0.905	0.971	0.897	0.86

Figure 7: Results of the Model

ONLINE ISSN
3006-9726
PRINT ISSN
3006-9718
VOLUME . 4 ISSUE . 1 (2025)

Conclusion:

The proper management of construction materials is a vital component in ensuring the safety, enhancing productivity, and reducing waste on construction sites. The traditional methods and ML methods have been used, offering the best results, but are limited in complex environmental conditions. This study introduces the DL-based model YOLOv11 to detect the construction materials in real-time. By using a labeled dataset, the model offers high accuracy in detecting multiple types of materials in adverse conditions, achieving a high **overall precision of 96.3% and accuracy of 94.3%**. The system delivers a robust, real-time interface and scalable capabilities that can be integrated into edge devices, thereby supporting smart construction initiatives. Future work will focus on expanding the datasets, incorporating Explainable Artificial Intelligence (XAI) methods for understanding, and incorporating the object detection system into a wider construction management system.

References:

[1] S. Durdyev and S. Ismail, “Role of the construction industry in economic development of Turkmenistan,” *Energy Educ. Sci. Technol. Part A Energy Sci. Res.*, vol. 29, pp. 883–890, Jul. 2012.

[2] J. Pipaliya, *REVIEW PAPER ON QUALITY MANAGEMENT PRACTICES FOR CONSTRUCTION PROJECT DELIVERY*. 2024. doi: 10.13140/RG.2.2.20760.51203.

[3] A. Setayesh, E. Grosse, C. Glock, and W. Neumann, “Determining the source of human-system errors in manual order picking concerning human factors,” *Int. J. Prod. Res.*, vol. 60, Nov. 2021, doi: 10.1080/00207543.2021.1991022.

[4] A. T. Lima *et al.*, “Mapping circular economy practices for steel, cement, glass, brick, insulation, and wood – A review for climate mitigation modeling,” *Renew. Sustain. Energy Rev.*, vol. 202, p. 114697, 2024, doi: https://doi.org/10.1016/j.rser.2024.114697.

[5] A. Akbari, “The application of radio-frequency identification (RFID) technology in the petroleum engineering industry: Mixed review,” *Pet. Res.*, 2025,

doi: https://doi.org/10.1016/j.ptlrs.2025.05.001.

[6] Muhammad Sajjad *et al.*, "AI-DRIVEN APPROACH FOR EARLY PROSTATE CANCER DETECTION AND DIAGNOSIS," *Spectr. Eng. Sci.*, vol. 3, no. 7 SE-Articles, pp. 1141–1151, Jul. 2025, [Online]. Available: https://www.sesjournal.com/index.php/1/article/view/700

[7] Muhammad Sajjad, Aqeel Ahmad, Ayesha Batool, Muhammad Mudassar Naveed, Arslan Ejaz, and Usman Aftab Butt, "AUTONOMOUS ROAD SIGN DETECTION USING DEEP LEARNING MODELS," *Spectr. Eng. Sci.*, vol. 3, no. 7 SE-Articles, pp. 609–620, Jul. 2025, [Online]. Available: https://sesjournal.com/index.php/1/article/view/625

[8] G. Ahmad, A. Khan, M. Sajjad, A. Saddiqa, M. M. Naveed, and Z. Riaz, "TRANSFER LEARNING-BASED SMART CROP PROTECTION SYSTEM FOR ANIMAL DETECTION AND DETERRENCE," *Spectr. Eng. Sci.*, vol. 3, no. 7, pp. 810–899, 2025.

[9] O. Olabanjo *et al.*, "Application of Machine Learning and Deep Learning Models in Prostate Cancer Diagnosis Using Medical Images: A Systematic Review," 2023. doi: 10.3390/analytics2030039.

[10 J. Redmon and A. Farhadi, "Yolo9000: better, faster, stronger arXiv preprint," *arXiv Prepr. arXiv*, vol. 1612, 2017.

[11] Y. Nakanishi, T. Kaneta, and S. Nishino, "A Review of Monitoring Construction Equipment in Support of Construction Project Management," *Front. Built Environ.*, vol. 7, Jan. 2022, doi: 10.3389/fbuil 2021.632593.

[12] H. Agency, *Inspection manual for highway structures: Vol. 1: Reference manual*, vol. 1. The Stationery Office, 2007.

[13] X. Jia, Q. Feng, T. Fan, and Q. Lei, "RFID technology and its applications in Internet of Things (IoT)," *2012 2nd Int. Conf. Consum. Electron. Commun. Networks, CECNet 2012 - Proc.*, Apr. 2012, doi: 10.1109/CECNet

2012.6201508.

[14] A. M. Darwish and A. K. Jain, "A rule-based approach for visual pattern inspection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 1, pp. 56–68, 1988.

[15] T. Lindeberg, "Scale invariant feature transform," 2012.

[16] M. N. Khan, A. Das, and S. Wulff, "Multilevel weather detection based on images: a machine learning approach with histogram of oriented gradient and local binary pattern-based features," *J. Intell. Transp. Syst.*, vol. 25, pp. 1–20, Jul. 2021, doi: 10.1080/15472450.2021.1944860.

[17] W. Olabiyi, D. Akinyele, and E. Joel, "The Evolution of AI: From Rule-Based Systems to Data-Driven Intelligence," Jan. 2025.

[18] Z. Zheng, L. Zhou, and H. Wu, "Construction cost prediction system based on Random Forest optimized by the Bird Swarm Algorithm," *Math. Biosci. Eng.*, vol. 20, pp. 15044–15074, Jul. 2023, doi: 10.3934/mbe 2023674.

[19] S. A. Khan, H. J. Lee, and H. Lim, "Enhancing Object Detection in Self-Driving Cars Using a Hybrid Approach," 2023. doi: 10.3390/electronics12132768.

[20] J. P. Vasconez, J. Delpiano, S. Vougioukas, and F. Auat Cheein, "Comparison of convolutional neural networks in fruit detection and counting: A comprehensive evaluation," *Comput. Electron. Agric.*, vol. 173, p. 105348, 2020, doi: https://doi.org/10.1016/j.compag.2020.105348.

[21] M. Carranza-García, J. Torres-Mateo, P. Lara-Benítez, and J. García-Gutiérrez, "On the Performance of One-Stage and Two-Stage Object Detectors in Autonomous Vehicles Using Camera Data," 2021. doi: 10.3390/rs13010089.

[22] S. Kang, Z. Hu, L. Liu, K. Zhang, and Z. Cao, "Object Detection YOLO Algorithms and Their Industrial Applications: Overview and Comparative Analysis," 2025. doi: 10.3390/electronics14061104.

[23] M. A. Alhassan and E. Yılmaz, "Evaluating YOLOv4 and YOLOv5 for Enhanced Object Detection in UAV-Based Surveillance," 2025. doi: 10.3390/pr13010254.

[24] J. Terven and D.-M. Cordova-Esparza, *A Comprehensive Review*

of YOLO: From YOLOv1 to YOLOv8 and Beyond. 2023. doi: 10.48550/arXiv.2304.00501.

[25] B. Li, Y. Liu, and X. Wang, "Gradient harmonized single-stage detector," in *Proceedings of the AAAI conference on artificial intelligence*, 2019, pp. 8577–8584.

[26] A. Suleiman and V. Sze, *Energy-Efficient HOG-based Object Detection at 1080HD 60 fps with Multi-Scale Support.* 2014. doi: 10.13140/2.1.4752.0964.

[27] M. Wasala and T. Kryjak, *Real-time HOG+SVM based object detection using SoC FPGA for a UHD video stream.* 2022. doi: 10.36227/techrxiv 19635429.v1.

[28] W. Fang, L. Ding, B. Zhong, P. E. D. Love, and H. Luo, "Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach," *Adv. Eng. Informatics*, vol. 37, pp. 139–149, 2018, doi: https://doi.org/10.1016/j.aei.2018.05.003.

[29] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.

[30] J. Ryu, D. Kwak, and S. Choi, "YOLOv8 with Post-Processing for Small Object Detection Enhancement," 2025. doi: 10.3390/app15137275.

[31] M. Tan, R. Pang, and Q. V Le, "Efficientdet: Scalable and efficient object detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10781–10790.

[32] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2016.

*Gulzar Ahmad