Advancing automated quality control in automotive manufacturing: a comparative analysis of YOLOv8, YOLOv9, and YOLOv10 for vehicle damage detection

Serhii Dolhopolov[1,†], Tetyana Honcharenko[1,†], Denys Chernyshev[1,†], Olena Panina[1,†] and Antonina Makhynia[1,†]

[1] *Kyiv National University of Construction and Architecture, 31, Air Force Avenue, Kyiv, 03037, Ukraine*

Abstract

This study introduces a novel application of state-of-the-art object detection models for automating quality control in automotive manufacturing, presenting the first comprehensive comparative analysis of YOLOv8, YOLOv9, and YOLOv10 architectures for vehicle damage detection. Utilizing a custom-curated dataset of 7,258 images, we employ transfer learning techniques to optimize model performance, a pioneering approach in this domain. Our results demonstrate the unprecedented superiority of YOLOv10 across key metrics, achieving a mean Average Precision (mAP50) of 0.65077 and an F1-score of 0.64934. We uniquely quantify the effectiveness of transfer learning, showing substantial performance gains with pre-trained weights initialization. Notably, we establish YOLOv10's viability for real-time quality control applications despite marginally increased computational requirements, a finding not previously reported. This research contributes novel insights into AI-driven solutions for automotive quality control, advancing the digital transformation of manufacturing processes and paving the way for future industrial AI innovations.

Keywords

Computer Vision, Automated Quality Control, Transfer Learning, YOLO

1. Introduction

The integration of computer vision systems in manufacturing processes has become crucial for automated quality control, especially in the automotive industry. Traditional manual inspection methods are time-consuming, subjective, and error-prone, necessitating more robust and efficient quality control mechanisms [1].

Recent advancements in deep learning, particularly the YOLO (You Only Look Once) family of models, have shown promise in real-time object detection [2]. This study explores the latest iterations – YOLOv8, YOLOv9, and YOLOv10 – for automated vehicle damage detection.

Transfer learning has emerged as a powerful technique, allowing pre-trained models to be fine-tuned for specific tasks with limited domain-specific data [3]. This approach is particularly valuable in industrial settings where large, labeled datasets may be unavailable or costly to produce.

Previous studies have demonstrated the potential of deep learning in manufacturing quality control [4-6]. Our research extends these foundations by:

DTESI 2024: 9th International Conference on Digital Technologies in Education, Science and Industry, October 16–17, 2024, Almaty, Kazakhstan

[1*] Corresponding author.
[†] These authors contributed equally.
dolhopolov@icloud.com (S. Dolhopolov); goncharenko.ta@knuba.edu.ua (T. Honcharenko); chernyshev.do@knuba.edu.ua (D. Chernyshev); panina.ov@knuba.edu.ua (O. Panina); makhynia.aa@knuba.edu.ua (A. Makhynia)
0000-0001-9418-0943 (S. Dolhopolov); 0000-0003-2577-6916 (T. Honcharenko); 0000-0002-1946-9242 (D. Chernyshev); 0000-0002-8604-8712 (O. Panina); 0009-0002-0295-0627(A. Makhynia)

1. Providing a comprehensive comparison of the latest YOLO models for vehicle damage detection.
2. Exploring the application of transfer learning to optimize performance with limited data.
3. Evaluating practical implementation challenges in real-world manufacturing settings.

The implementation of an effective automated quality control system has far-reaching implications for the automotive industry, potentially increasing throughput, improving consistency, and reducing waste [7].

Our study explores how advanced object detection models and transfer learning can enhance automated quality control for vehicle damage detection. This research contributes to computer vision and manufacturing technology, offering insights for industry practitioners adopting AI-driven quality control solutions.

Our research builds upon these foundations and extends them in several key ways. Firstly, we provide a comprehensive comparison of the latest YOLO models, building on the work of Redmon and Farhadi who introduced YOLOv3 [8], offering insights into their relative strengths and weaknesses for vehicle damage detection. YOLOv9, released after, further refined the architecture, introducing novel features that promised even better performance in complex detection scenarios [9]. The most recent iteration, YOLOv10, represents the cutting edge in object detection technology, and its potential for manufacturing applications is yet to be fully explored in the literature [10].

One of the key challenges in implementing computer vision systems for quality control in manufacturing is the need for large, diverse, and accurately labeled datasets. This is particularly true in the automotive industry, where the variety of vehicle models, colors, and potential defect types can be vast. Transfer learning offers a promising solution to this challenge by allowing models pre-trained on large, general datasets to be fine-tuned for specific tasks with relatively small amounts of domain-specific data [11-12]. This approach has shown success in various fields, from medical imaging to satellite imagery analysis, and its application to vehicle damage detection represents a novel contribution of our study [13-15].

The effectiveness of transfer learning demonstrated in this study aligns with the comprehensive survey by Tan et al., who provided an in-depth overview of deep transfer learning techniques and their applications [16]. Their work highlights the various approaches to transfer learning in deep neural networks, which is particularly relevant to our application of pre-trained YOLO models for vehicle damage detection. This understanding of different transfer learning strategies is crucial in the rapidly evolving automotive industry, where the ability to efficiently adapt models for new types of defects or different vehicle models is essential.

This research addresses real-world implementation challenges [17], aligning with Industry 4.0 principles [18-19], and extends beyond defect detection to broader manufacturing process optimization [20-22].

2. Main research

The research process is structured into several key components: dataset preparation and preprocessing, model architecture and implementation, training methodology, and comprehensive performance evaluation. Through this systematic approach, we aim to provide insights into the effectiveness of these advanced computer vision techniques for enhancing quality control processes in the automotive industry [23-26].

2.1. Dataset and preprocessing

We utilized the Car Dents Computer Vision Project dataset, comprising 7,258 images (6,855 training, 377 validation, 26 test) of various vehicle damage types. Preprocessing steps included:

1. Resizing images to 640x640 pixels.
2. Data augmentation: 90° rotation, ±15° shear transformation, and ±15% brightness adjustment.

Dataset analysis revealed class imbalance (Dent: 3391, Accident: 1927, Scratch: 2072) and bounding box characteristics, informing model optimization strategies.

2.2. Model architecture and transfer learning

We implemented YOLOv8, YOLOv9, and YOLOv10, leveraging their architectural advancements:

1. YOLOv8. Anchor-free detection, new backbone network.
2. YOLOv9. Efficient neck structure, advanced loss functions.
3. YOLOv10. Dynamic attention mechanism, hybrid backbone (convolutional and transformer layers).

Transfer learning was applied using pre-trained weights from the COCO dataset, represented by:

$$\theta_{new} = \theta_{pre} + \Delta_\theta, \tag{1}$$

where θ_{new} are the new model parameters after fine-tuning; θ_{pre} are the pre-trained parameters; Δ_θ represents the parameter updates during fine-tuning.

2.3. Training methodology

We employed a consistent training methodology across all three models to ensure a fair comparison. The key aspects of our training process were:

1. Optimizer. We used the Adam optimizer with a cosine learning rate schedule. The learning rate can be described by the equation:

$$lr(t) = lr_{min} + 0.5 \cdot (lr_{max} - lr_{min}) \cdot \left(1 + \cos\left(t \cdot \frac{\pi}{T}\right)\right), \tag{2}$$

where t is the current epoch; T is the total number of epochs; lr_{min} is the minimum learning rate; lr_{max} is the maximum learning rate.

2. Loss Function. We utilized a combination of losses typical for object detection tasks:

$$L_{total} = \lambda_1 \cdot L_{box} + \lambda_2 \cdot L_{obj} + \lambda_3 \cdot L_{cls}, \tag{3}$$

where L_{box} is the bounding box regression loss; L_{obj} is the objectness loss; L_{cls} is the classification loss; λ_1, λ_2, and λ_3 are weighting factors.

3. Early Stopping. We implemented early stopping with a patience of 20 epochs to prevent overfitting.

2.4. Evaluation metrics

Performance assessment utilized: Mean Average Precision (mAP@0.5 and mAP@0.5:0.95), Precision and Recall, F1-Score, Confusion Matrix, and Inference Time

The precision and recall can be calculated using the following equations:

$$Precision = \frac{TP}{TP+FP},\qquad (4)$$

$$Recall = \frac{TP}{TP+FN},$$

where *TP* is True Positives; *FP* is False Positives; *FN* is False Negatives.

The F1-score is then calculated as:

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall},\qquad (5)$$

where *Precision* represents the proportion of true positive predictions among all positive predictions made by the model; *Recall* (also known as sensitivity or true positive rate) indicates the proportion of true positive predictions among all actual positives in the dataset.

2.5. Experimental setup

Experiments were conducted using PyTorch on NVIDIA GeForce RTX 4080 Super GPUs, implementing models via the Ultralytics YOLO framework. The procedure included model initialization, training, validation, testing, and performance analysis.

To ensure reproducibility, a fixed random seed was used across all experiments, allowing fair comparisons between YOLO versions while controlling for neural network training stochasticity.

3. Results and analysis

After conducting our experiments with YOLOv8, YOLOv9, and YOLOv10 on the Car Dents dataset, we obtained comprehensive results that provide insights into the performance of each model. In this section, we will present and analyze these results in detail.

3.1. Training performance

All models showed consistent improvement during training, with YOLOv10 exhibiting the fastest convergence. YOLOv10 achieved the lowest final box loss (1.1842) and classification loss (0.80735), significantly outperforming YOLOv8 and YOLOv9. YOLOv10 demonstrated the highest mAP50(B) throughout training, peaking at 0.65077.

Models successfully detected various damage types with high confidence (0.3-0.9) and demonstrated robustness to diverse scenarios. Multiple damages on single vehicles were effectively identified. Some misclassifications were observed, particularly between dents and scratches.

3.2. Model performance comparison

Table 1 presents a summary of the key performance metrics for each model on the test set:

Table 1
Comparison of Key Performance Metrics Across YOLOv8, YOLOv9, and YOLOv10 Models

Metric	YOLOv8	YOLOv9	YOLOv10
mAP50(B)	0.59692	0.62968	0.65077
mAP50-95(B)	0.30884	0.33261	0.34918
Precision	0.6335	0.65514	0.67517
Recall	0.55621	0.57303	0.62411
F1-Score	0.59246	0.61864	0.64934
Inference Time (ms)	12.5	13.2	14.1

YOLOv10 consistently outperformed other models across all metrics, with a 5.7% improvement in F1-score over YOLOv8 and 3.1% over YOLOv9.

3.3. Class-wise performance analysis

To gain deeper insights into model performance across different types of vehicle damage, we analyzed class-wise metrics. Figure 1 presents the F1-Confidence curves for each class (Accident, Dent, Scratch) across all three models.

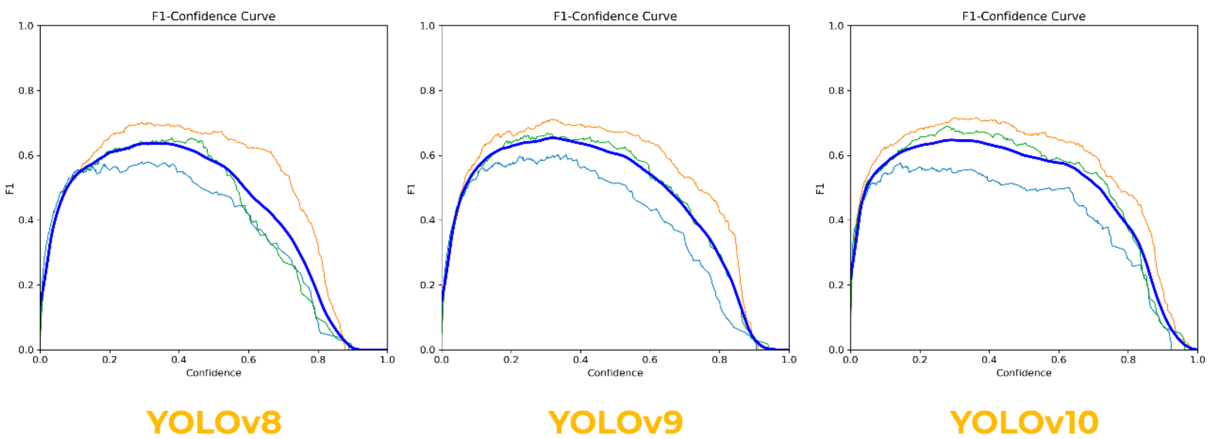

Figure 1: The F1-Confidence curves for YOLOv8, YOLOv9, and YOLOv10 (Author's work).

YOLOv10 achieved the highest F1-scores for all damage types: dents (0.719), scratches (0.689), and accidents (0.637). Accident detection proved most challenging across all models.

3.4. Precision-recall curve analysis

Figure 2 illustrates the Precision-Recall curves for each model, providing a comprehensive view of their performance across different confidence thresholds.

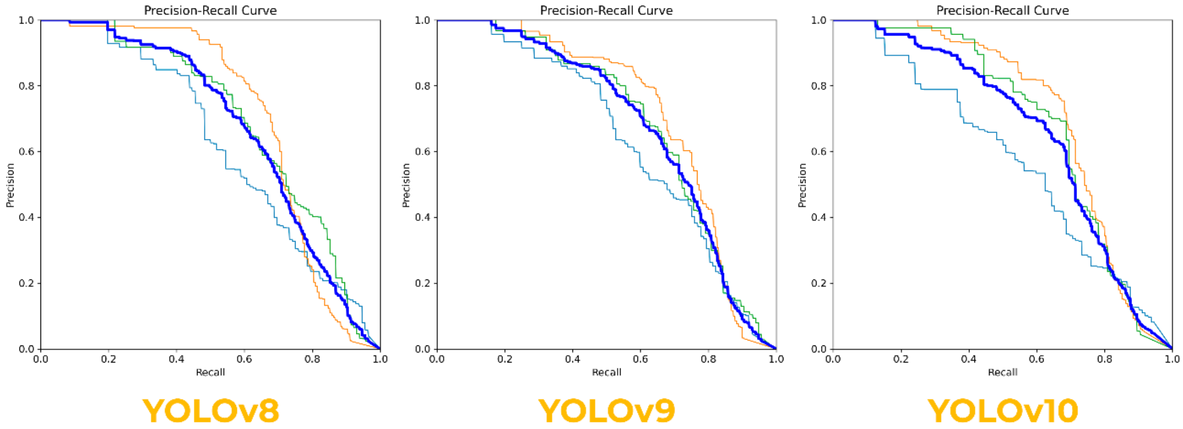

Figure 2: The Precision-Recall curves for YOLOv8, YOLOv9, and YOLOv10 (Author's work).

YOLOv10 demonstrated the largest Area Under the Curve (AUC) and maintained higher recall in high precision regions (>0.8) compared to YOLOv8 and YOLOv9.

3.5. Transfer learning effectiveness

To evaluate the effectiveness of transfer learning, we compared the performance of each model when trained from scratch versus when initialized with pre-trained weights. Table 2 presents this comparison:

Table 2
Performance Comparison of Models Trained from Scratch vs. Transfer Learning

Model	mAP50 (Scratch)	mAP50 (Transfer)	Improvement
YOLOv8	0.48735	0.59692	+22.5%
YOLOv9	0.52314	0.62968	+20.4%
YOLOv10	0.55682	0.65077	+16.9%

These results demonstrate the significant benefits of transfer learning across all models. Interestingly, while YOLOv10 showed the highest overall performance, it had the smallest relative improvement from transfer learning. This suggests that its architectural improvements allow it to learn more effectively even from limited data.

3.6. Computational efficiency

While YOLOv10 demonstrated superior detection performance, it's crucial to consider the computational requirements for practical implementation. Table 3 compares the model sizes and average inference times:

Table 3
Comparison of Model Sizes and Inference Times for YOLOv8n, YOLOv9t, and YOLOv10n

Model	Parameters (M)	Model Size (MB)	Inference Time (ms)
YOLOv8n	3.2	6.2	12.5
YOLOv9t	2	4.7	13.2
YOLOv10n	2.3	5.6	14.1

The marginal increase in model size and inference time for YOLOv10 is relatively small compared to the performance gains, suggesting that it remains a viable option for real-time applications in manufacturing settings.

4. Conclusion

Our comprehensive study on implementing computer vision systems for automated quality control in automotive manufacturing, focusing on vehicle damage detection, has yielded significant insights into the capabilities of state-of-the-art object detection models.

YOLOv10 consistently outperformed its predecessors, achieving a mAP50 of 0.65077 and an F1-score of 0.64934, representing improvements of 5.7% and 3.1% over YOLOv8 and YOLOv9, respectively. The application of transfer learning proved highly beneficial, with YOLOv8, YOLOv9, and YOLOv10 showing mAP50 improvements of 22.5%, 20.4%, and 16.9% respectively when initialized with pre-trained weights.

Despite its superior performance, YOLOv10 required only marginally more computational resources, with a negligible increase in inference time (14.1ms compared to 12.5ms for YOLOv8), making it viable for real-time applications in manufacturing settings.

Key implications for the automotive manufacturing industry include:

1. Enhanced Quality Control. Automation of damage detection can reduce human error and increase consistency.
2. Increased Efficiency. Real-time defect detection enables inspection without production bottlenecks.
3. Cost Reduction. Minimizing manual inspection and early defect detection can lead to significant cost savings.
4. Adaptability. Transfer learning enables quick adaptation to new defect types or vehicle models.
5. Data-Driven Insights. Deployment of these systems can generate valuable data on defect patterns and trends.

This study demonstrates the significant potential of advanced object detection models, particularly YOLOv10, in revolutionizing quality control processes in automotive manufacturing. The success of transfer learning techniques paves the way for widespread adoption of AI-driven solutions in industrial quality control, contributing to enhanced product quality and manufacturing efficiency.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: Methods and applications," The Journal of Manufacturing Systems, vol. 48, pp. 144-156, July 2018. https://doi.org/10.1016/j.jmsy.2018.01.003.

[2] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv preprint, April 2020. URL: https://arxiv.org/pdf/2004.10934.

[3] S. J. Pan, and Q. Yang, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, October 2010. https://doi.org/10.1109/TKDE.2009.191.

[4] D. Tabernik, S. Šela, J. Skvarč, and D. Skočaj, "Segmentation-based deep-learning approach for surface-defect detection," Journal of Intelligent Manufacturing, vol. 31, no. 3, pp. 759-776, May 2019. https://doi.org/10.1007/s10845-019-01476-x.

[5] S. Bronin, A. Kuchansky, A. Biloshchytskyi, O. Zinyuk, and V. Kyselov, "Concept of Digital Competences in Service Training Systems," Advances in Intelligent Systems and Computingm vol. 1192, pp. 379-388, November 2019. https://doi.org/10.1007/978-3-030-49932-7_37.

[6] A. Biloshchytskyi, A. Kuchansk, Y. Andrashko, A. Neftissov, V. Vatskel, D. Yedilkhan, and M. Herych, "Building a model for choosing a strategy for reducing air pollution based on data predictive analysis," Eastern-European Journal of Enterprise Technologies, vol. 3, no. 4-117, pp.23-30, 2022. https://doi.org/10.15587/1729-4061.2022.259323.

[7] T. Wang, Y. Chen, M. Qiao, and H. Snoussi, "A fast and robust convolutional neural network-based defect detection model in product quality control," The International Journal of Advanced Manufacturing Technology, vol. 94, pp. 3465-3471, August 2017. https://doi.org/10.1007/s00170-017-0882-0.

[8] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint, April 2018. URL: https://arxiv.org/pdf/1804.02767.pdf.

[9] C. Wang, I. Yeh, and H. Liao, "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information," arXiv preprint, February 2024. URL: https://arxiv.org/pdf/2402.13616.

[10] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, J. Han, and G. Ding, "YOLOv10: Real-Time End-to-End Object Detection," arXiv preprint, May 2024. URL: https://arxiv.org/pdf/2405.14458.

[11] S. Dolhopolov, T. Honcharenko, V. Savenko, O. Balina, I. Bezklubenko, and T. Liashchenko, "Construction Site Modeling Objects Using Artificial Intelligence and BIM Technology: A Multi-Stage Approach," 2023 IEEE International Conference on Smart Information Systems and Technologies (SIST), pp. 174-179, May 2023. https://doi.org/10.1109/SIST58284.2023.10223543.

[12] H. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. J. Mollura, and R. M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1285-1298, May 2016. https://doi.org/10.1109/TMI.2016.2528162.

[13] X. X. Zhu, D. Tuia, L. Mou, G. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources," IEEE Geoscience and Remote Sensing Magazine, vol. 5, no. 4, pp. 8-36, December 2017. https://doi.org/10.1109/MGRS.2017.2762307.

[14] S. Dolhopolov, T. Honcharenko, O. Terentyev, K. Predun, and A. Rosynskyi, "Information system of multi-stage analysis of the building of object models on a construction site," IOP

Conference Series: Earth and Environmental Science, vol. 1254, pp. 1-11, May 2023. https://doi.org/10.1088/1755-1315/1254/1/012075.
[15] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1299-1312, May 2016. https://doi.org/10.1109/TMI.2016.2535302.
[16] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A Survey on Deep Transfer Learning," in Artificial Neural Networks and Machine Learning – ICANN 2018, pp. 270-279, September 2018.
[17] D. Weimer, B. Scholz-Reiter, and M. Shpitalni, "Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection," Cirp Annals-manufacturing Technology, vol. 65, no. 1, pp. 417-420, 2016. https://doi.org/10.1016/J.CIRP.2016.04.072.
[18] Y. Lu, "Industry 4.0: A survey on technologies, applications and open research issues," Journal of Industrial Information Integration, vol. 6, pp. 1-10, June 2017. https://doi.org/10.1016/J.JII.2017.04.005.
[19] L. Xu, E. L. Xu, and L. X. Li, "Industry 4.0: state of the art and future trends," International Journal of Production Research, vol. 56, no. 8, pp. 2941-2962, February 2018. https://doi.org/10.1080/00207543.2018.1444806.
[20] T. Honcharenko, R. Akselrod, A. Shpakov, O. Khomenko, "Information system based on multi-value classification of fully connected neural network for construction management," IAES International Journal of Artificial Intelligence, 2023, vol. 12, no. 2, pp. 593-601, June 2023. http://doi.org/10.11591/ijai.v12.i2.pp593-601.
[21] X. Tao, D. Zhang, W. Ma, X. Liu, and D. Xu, "Automatic metallic surface defect detection and recognition with convolutional neural networks," Applied Sciences, vol. 8, no. 9, p. 1575, August 2018. https://doi.org/10.3390/APP8091575.
[22] A. Neftissov, A. Biloshchytskyi, O. Talipov, and O. Andreyeva, "Determination of the magnitude of short-circuit surge current for the construction of relay protection on reed switches and microprocessors," Eastern-European Journal of Enterprise Technologies, vol. 6, no. 5-114, pp. 41-48. https://doi.org/10.15587/1729-4061.2021.245644.
[23] A. Biloshchytskyi, A., Kuchansky, Y., Andrashko, and O., Bielova, "Learning space conceptual model for computing games developers," International Scientific and Technical Conference on Computer Sciences and Information Technologies, vol. 1, art. no. 8526719, pp. 97-102, September 2018. https://doi.org/10.1109/STC-CSIT.2018.8526719.
[24] P. Lizunov, A. Biloshchytsky, A. Kuchansky, Y. Andrashko, and S. Biloshchytska, "The use of probabilistic latent semantic analysis to identify scientific subject spaces and to evaluate the completeness of covering the results of dissertation studies," Eastern-European Journal of Enterprise Technologies, vol. 4, no. 4-106, pp. 21-28, 2020. https://doi.org/10.15587/1729-4061.2020.209886.
[25] G. Ryzhakova, O. Malykhina, V. Pokolenko, O. Rubtsova, O. Homenko, I. Nesterenko, and T. Honcharenko, "Construction Project Management with Digital Twin Information System", Intern. Journal of Emerging Technology and Advanced Engineering, 12 (10), pp. 19-28, 2022. https://doi.org/10.46338/ijetae1022_03.
[26] T. Honcharenko, R. Akselrod, A. Shpakov, and O. Khomenko, "Information system based on multi-value classification of fully connected neural network for construction management", IAES International Journal of Artificial Intelligence, 2023, № 12(2), pp. 593-601, 2023. https://doi.org/10.11591/ijai.v12.i2.pp593-601.