МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РЕСПУБЛИКИ КАЗАХСТАН
НЕКОММЕРЧЕСКОЕ АКЦИОНЕРНОЕ ОБЩЕСТВО
«АЛМАТИНСКИЙ УНИВЕРСИТЕТ ЭНЕРГЕТИКИ И СВЯЗИ ИМЕНИ ГУМАРБЕКА ДАУКЕЕВА»

ISSN-2790-0886

В Е С Т Н И К

АЛМАТИНСКОГО УНИВЕРСИТЕТА ЭНЕРГЕТИКИ И СВЯЗИ

Учрежден в июне 2008 года

Тематическая направленность: энергетика и энергетическое машиностроение, информационные, телекоммуникационные и космические технологии

3 (74) **2026**

Импакт-фактор - 0.076

Научно-технический журнал

Выходит 4 раза в год

Алматы

ИНФОРМАЦИОННЫЕ, ТЕЛЕКОММУНИКАЦИОННЫЕ И КОСМИЧЕСКИЕ ТЕХНОЛОГИИ

A.A. TEMIRBAYEV, AZATI YAKUFUJIANG
APPLICATION-LEVEL PERFORMANCE EVALUATION OF A MID-SCALE HPC SYSTEM FOR AI WORKLOADS.....139

Y. Nurlanov, A.S. Bissembayev, D.M. Mukhammejanova, A.K. Mukasheva
DISTRIBUTED SYSTEM FOR DATA SHARING AND JOINT PROCESSING BETWEEN RESEARCH CENTERS AND FARMS.....152

D. Milyukov, K. Eslyam, G. Kim, A. Kabdirova
DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR MONITORING MANUFACTURING DEFECTS BASED ON EDGE AI AND YOLO ARCHITECTURE.....163

А.С. Ерболова, Р.С. Орыскалиев, Д.О. Жусипбек, А.С. Шаяхметова
МЕТОД АНАЛИЗА МНОГОМЕРНЫХ ТЕКСТОВЫХ ДАННЫХ НА ОСНОВЕ ПЛОСКОСТИ «ЭНТРОПИЯ–СЛОЖНОСТЬ» 173

А. Құдайберген, Ә. Жора, Р.Қ. Ерғалиев, Е. Нұрғизат
ALATAU СЕЙСМИКАЛЫҚ ЭТАЛОНДЫҚ ЖИНАҒЫ ЖӘНЕ PHASENET НЕЙРОЖЕЛІСІН ЖЕРГІЛІКТІ ДЕРЕКТЕРМЕН ОҚЫТУ 184

Z.B. Mukhtarova, A.T. Zharkimbekova, A.S. Amirova
APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES FOR THE AUTOMATION AND INTELLIGENT ANALYSIS OF INFORMATION SECURITY MANAGEMENT SYSTEM AUDITS.....196

І.М. Самат, Р.К. Ускенбаева, Ж.Б. Кальпеева, Н. Албанбай
ЗАТТАР ИНТЕРНЕТІ ҮШІН ЖЕҢІЛ ГОМОМОРФТЫ ШИФРЛАУ.....207

Zharmakhan K.K., Mukasheva A.K., D.M. Mukhammejanova
HYBRID GAT-AUTOENCODER-BILSTM-ATTENTION ARCHITECTURE FOR MULTI CLASS CYBERATTACK DETECTION IN HETEROGENEOUS IOT NETWORKS 220

K.M. Sagindykov, Zh. Zh. Akhmetova, N. Zh. Mukusheva, A. G. Kerim
RESILIENCE OF CYBERPHYSICAL SYSTEMS TO MULTI-VECTOR ATTACKS IN ENSURING CBF/ECBF SECURITY.....232

A.T. Turarbek, Sh. Temirgaziyeva, A. S. Dokuz
DEVELOPMENT OF A MULTI-SCALE CONVOLUTIONAL ATTENTION ARCHITECTURE (MSCA-NET) FOR AUDIO DEEPFAKE DETECTION USING MELSPECTROGRAMS.....242

Амирханова С.С., Амангелді Ш.С., Дюсенбаева А.А., Садырбаева Д.С.
ИНТЕЛЛЕКТУАЛЬНАЯ IOT-СИСТЕМА МОНИТОРИНГА ДАННЫХ МЕТЕОЗОНДА НА ОСНОВЕ FLATSAT.....257

T.M. Mukanov, A.B. Syzdykov, K.Z. Sakhanov, G.A. Amandykuly.
DEVELOPMENT OF AN ON-BOARD NAVIGATION SYSTEM FOR POCKETQUBE SATELLITE BASED ON THE HIGH-PRECISION MOSAIC-G5 GNSS RECEIVER.....271

ИНФОРМАЦИОННЫЕ, ТЕЛЕКОММУНИКАЦИОННЫЕ И КОСМИЧЕСКИЕ ТЕХНОЛОГИИ

UDC 004.85 DOI 10.51775/2790-0886_2026_74_3_163

DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR MONITORING MANUFACTURING DEFECTS BASED ON EDGE AI AND YOLO ARCHITECTURE

[1]D. Milyukov, [1]K. Eslyam, [1]*G. Kim, [2]A. Kabdirova

[1] NJSC “Manash Kozybayev North Kazakhstan University”, Petropavl, Kazakhstan.
[2] The branch of Joint Stock Company “National Center of Professional Development for Pedagogical Staff ‘Orleu’”, Petropavl, Kazakhstan.

E-mail: kabdirova_a@orleu.edu.kz, gakim@ku.edu.kz, kanatyeslyam@gmail.com, milyukovdo@gmail.com.

Поступила в редакцию 22.07.2026
Получена исправленная версия 30.07.2026
Принята 10.09.2026
Доступна на сайте 30.09.2026

Abstract: *The rapidity of the process in modern manufacturing implies a high speed of continuous monitoring, which often surpasses the potential of manual control. The study presents the adaptive, energy-efficient intelligent platform for monitoring defects with the use of Edge Artificial Intelligence (AI) architecture. In terms of objectives, special attention is paid to solving latent computational problems in real-time processing on the local edge device at "Raduga" LLP.*

At the beginning, the study performs a comparison of One-to-Many and One-to-One detection architectures. The findings show that the Non-Maximum Suppression (NMS) step decreases the speed of local graphics accelerators; accordingly, the system under development uses NMS-free neural network detectors. In order to standardize the process of defining the data flows and algorithms, the following standards for system-structural modeling are used: IDEF0, IDEF3 and DFD. Besides, the system includes a lightweight You Only Look Once (YOLO) 26 Nano detector and ByteTrack object tracking algorithm. Thanks to this software module, it becomes possible to overcome computational difficulties in post-processing of frames and to perform stable object tracking on a moving conveyor. Finally, operational tests show that the system performs with stable real-time speeds on the NVIDIA GeForce RTX 2050 GPU with FPS above 30 for the conveyor speed of 1.5 m/s..

Keywords: *Edge AI, YOLO architecture, defect detection, computer vision, continuous quality control, manufacturing monitoring.*

INTRODUCTION

The issue of continuous quality control has acquired the status of one of the key points of the technological chain in the conditions of modern fast-paced production. Traditional visual quality assurance approaches, largely based on the human operator's actions, show inefficiency due to the operators' mental exhaustion, increasing conveyor speed, and the necessity of stable monitoring around-the-clock. Hence, creation of an intelligent computer vision system has become a top priority for industrial automation.

Despite the existence of ready-made commercial computer vision products, their implementation in SMEs is often restricted by expensive licensing terms and difficulties of integrating into existing legacy systems. This research seeks to implement an adaptive, computationally efficient and cost-effective intelligent platform for defect monitoring, optimized for Edge AI hardware.

The goal of this research is to develop an automated system for defect detection that would provide high precision (98%) of operation in real-time mode on conveyor lines with the speed up to 1.5 m/s. In order to solve this problem, the following multi-stage approach was applied:

• System-Structural Modeling: The business process and information flows were modeled according to IDEF0, IDEF3, and DFD methods, allowing us to identify potential latencies in processing video streams.

• Optimization of Neural Network Architecture: The comparison of "one-to-many" and "one-to-one" architectures was carried out. As a result of the analysis of the influence of the Non-Maximum Suppression (NMS) stage on the GPU performance, the YOLO26 Nano detector was selected.

• The Integration of Tracking Methods: In order to maintain data integrity during continuous movement of conveyor, the ByteTrack object tracking method was used to avoid events' duplication during classification of defects.

The practical significance of this research is confirmed by operational testing of the system at the Raduga LLP production site, revealing its stability during dynamic workload and effectiveness of the data processing on peripheral (Edge) devices. The findings obtained during the research can be scaled for other defect detection problems at similar enterprises, facilitating transition to "smart" production environments without large investments in computing infrastructure.

METHODS

Over the last few years, the drive to automate manufacturing quality control has become a major catalyst for advancing applied computer vision. Reviewing the current literature, research in this domain naturally categorizes into several key methodological areas.

1. Application of Neural Network Architectures for Defect Detection The bulk of contemporary work focuses on adapting the YOLO family of detectors. Foundational studies by Redmon et al. [1] and subsequent advancements by Bochkovskiy et al. [2] clearly established that single-stage detectors are highly capable of handling real-time tasks. In the context of industrial deployments, Liu et al. (2022) [3] proved that YOLO-Nano architectures can maintain the necessary precision (mAP > 90%) even on hardware with constrained computational limits. This directly supports our rationale for selecting the YOLO26 Nano model for our Edge infrastructure.

2. Methods for Enhancing Machine Vision System Performance The scientific community places considerable emphasis on optimizing the post-processing phases of these systems. Specifically, investigations by Ge et al. (2021) into YOLOv6 and YOLOv7 architectures identified the Non-Maximum Suppression (NMS) stage as a major computational bottleneck when processing high-frequency video feeds. A practical workaround, highlighted in recent publications like Wang et al. (2023) [4], is the transition toward NMS-free architectures. This strategy perfectly aligns with our method of optimizing the computational pipeline on the NVIDIA RTX 2050.

3. Integration of Tracking Algorithms in Industrial Environments When monitoring objects continuously moving down a line, preventing event duplication is a major challenge. To address this, researchers frequently integrate specialized tracking algorithms. The work introduced by Zhang et al. (ByteTrack, 2022) [5] demonstrated that track association based directly on detection boxes—without relying on motion approximation, substantially improves object identification, even during partial occlusion. Our system relies heavily on this ByteTrack methodology to maintain stability and tracking accuracy at conveyor speeds up to 1.5 m/s.

4. System-Structural Modeling Utilizing formalized modeling techniques (such as IDEF0, IDEF3, and DFD) to map out business processes is an established international standard in systems engineering. As detailed by Satzinger et al. (2015) [6] in their foundational literature on systems analysis, formalizing data flows early in the project lifecycle significantly mitigates the risk of architectural flaws during the actual coding of software modules. This principle strongly justifies our chosen design methodology.

To determine the most suitable neural network architecture for our specific requirements, we performed a comparative analysis of available solutions. This evaluation focused on accuracy (mAP) and operational performance (FPS) metrics, utilizing the NVIDIA RTX 2050 hardware platform. The findings from these tests are detailed in Table 1.

Table 1 — Comparative Analysis of Neural Network Architectures for Defect Detection

Architecture	mAP@0.5 (Accuracy)	FPS (Performance)	NMS Presence	Resource Requirements
RT-DETR	0.96	~15	Да	High
YOLOv8x	0.97	~22	Да	High
YOLOv8n	0.92	~45	Да	Medium
YOLO26 Nano	0.95	51	Нет	Low

As the data shows, the YOLO26 Nano architecture provides an optimal balance of performance [7], demonstrating a speed of 51 frames per second while maintaining high detection accuracy (mAP 0.95), which is critical for real-time production line operations.

Structural analysis techniques were used to describe the system in detail. At the upper level (A0), the "Automatic Defect Detection" process uses the video stream from the assembly line as input. Technical regulations and quality parameters serve as controls, and the Edge Artificial Intelligence (AI) computing module (RTX 2050) and Python-based software serve as mechanisms [8, 9]. The result of the process is data on detected defects and visual reports.

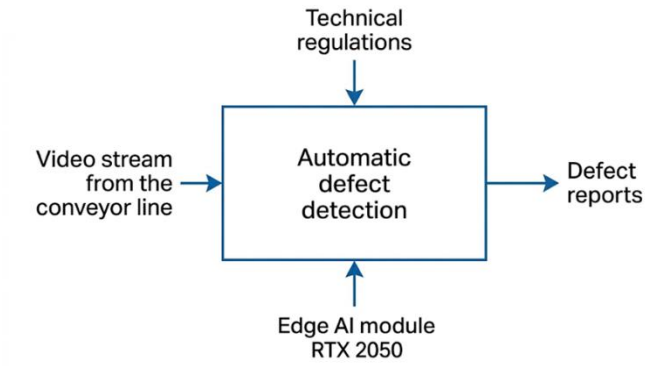


Figure 1 — Integration Definition for Function Modeling (IDEF0) diagram

The system operation scenario includes the following sequence of stages [10]:
Frame capture (OpenCV);
Preprocessing and feeding into You Only Look Once (YOLO) 26 Nano;
Object tracking (ByteTrack) to eliminate double counting;
Decision making on the presence of a defect;
Recording events in the PostgreSQL Database (DB) and notifying the operator

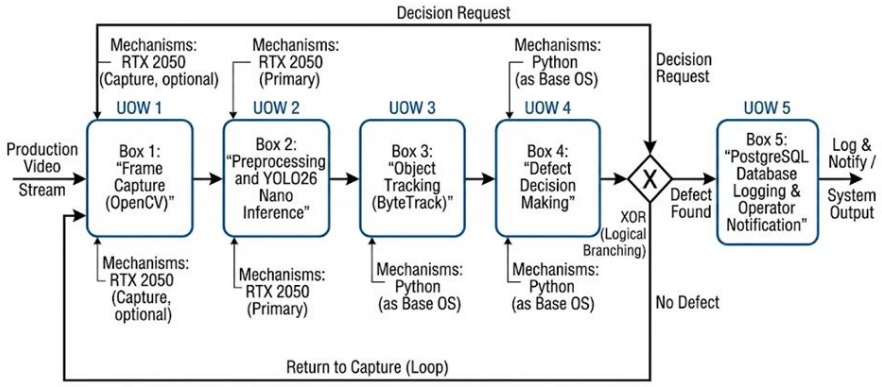


Figure 2 — Integration Definition for Process Description Capture Method (IDEF3) diagram: Intelligent System Workflow

The Data Flow Diagram (DFD) highlights the main data stores (Defect Database) and external objects (operator, pipeline). Data streams show the movement of object metadata from the neural network module to the Streamlit interface through the Plotly visualization layer [11].

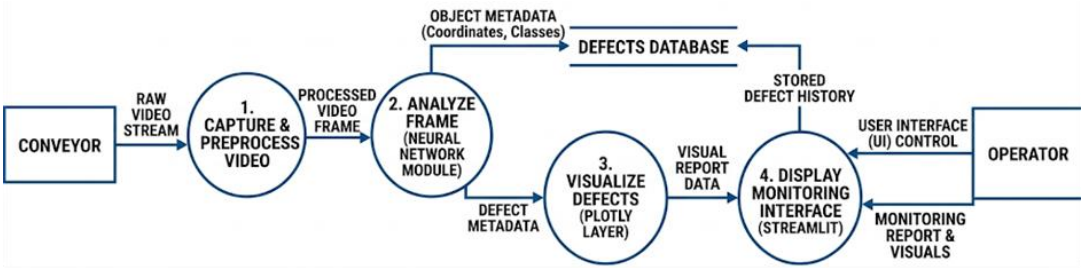


Figure 3 — Data Flow Diagram (DFD) – Defect monitoring system

The developed data flow architecture formed the basis for choosing a specific software and hardware stack. The choice of the technology stack is determined by the strict requirements for an intelligent system for real-time operation directly on local production equipment, which corresponds to the concept of advanced computing (edge AI) [12 ,13]. This approach avoids network delays associated with the transmission of the video stream to external servers, and ensures complete pipeline management autonomy.

The software implementation of the algorithms is made in Python 3.11. This made it possible to effectively integrate modern signal processing and computer vision libraries, as well as provide direct work with the computing power of the NVIDIA RTX graphics processor (GPU) in 2050 [14, 15].

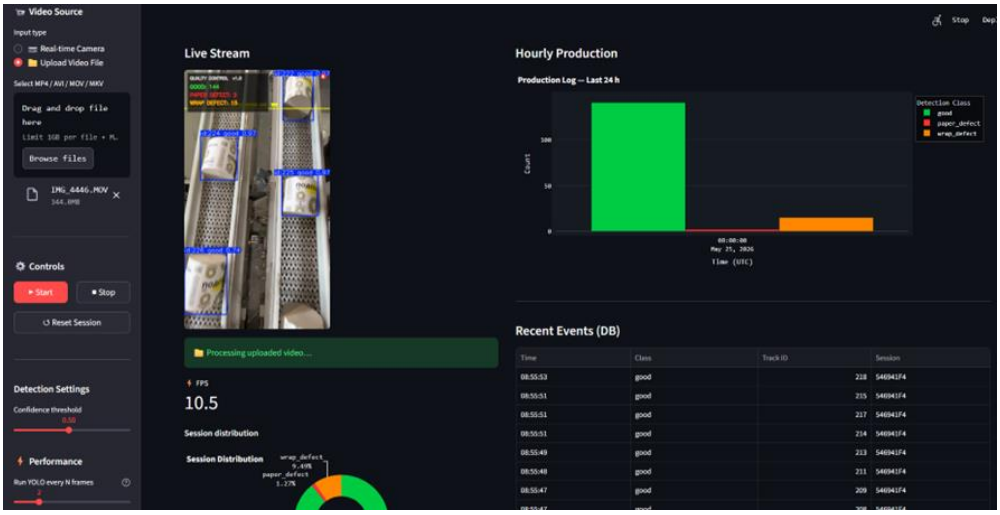


Figure 4 — Wireframe of the dashboard layout

The architecture of the system is based on the seamless interaction of four key components:

Neural Network Detection module: The lightweight YOLO model is used as the visual control core. An important characteristic of the chosen model is the use of architecture without the use of non-maximal suppression (NMS) (without the stage of non-maximal suppression). This solution significantly reduces the overall computing load on the equipment and almost completely eliminates delays during post-processing of frames, which is a critical condition for Edge AI class systems when working with fast-moving objects on the conveyor [16].

Object Tracking (ByteTrack): To avoid double counting of the same defect or product, the ByteTrack algorithm is used. This guarantees stable tracking of the movement of objects in the video stream, preserving their identifiers even in conditions of short-term dimming or sudden changes in industrial lighting [17].

User Interface (Streamlit): A web-based dashboard for monitoring and managing technological processes is implemented using the Streamlit platform. This solution allows operators to access the interface through any standard web browser [18].

Database: The PostgreSQL Relational Database Management System (DBMS) is used to securely store the history of detected incidents, defect metadata, and aggregated statistics. The choice of a relational structure is determined by the need to ensure data integrity when registering high-frequency information flows for their subsequent retrospective analysis [19].

RESULTS

To support operational monitoring and retrospective analysis of production metrics, a dynamic visualization module based on the Plotly library is integrated into the system. This tool equips the operator with interactive charts showing the distribution of detected incidents by defect type, hourly intervals, and production shifts. Introducing this analytical layer does more than just help eliminate defects in real-time; it also pinpoints systemic anomalies in the conveyor equipment's operation, which is essential for scheduling preventive maintenance.

The viability of the proposed hardware-software complex was proven during experimental trials using actual production data. The trained neural network model displayed high precision in both localization and classification, with the mean average precision (mAP@0.5) reaching 0.95. Throughout the evaluation, overall recognition accuracy stood at 98%, and the completeness of defect detection (recall rate) reached 92.9%. Performance-wise, the system running on an Edge module (NVIDIA RTX 2050) maintained 51 frames per second (FPS) with a low output latency of just 19 ms.

.

The training curves are shown in accordance with Figure 5.

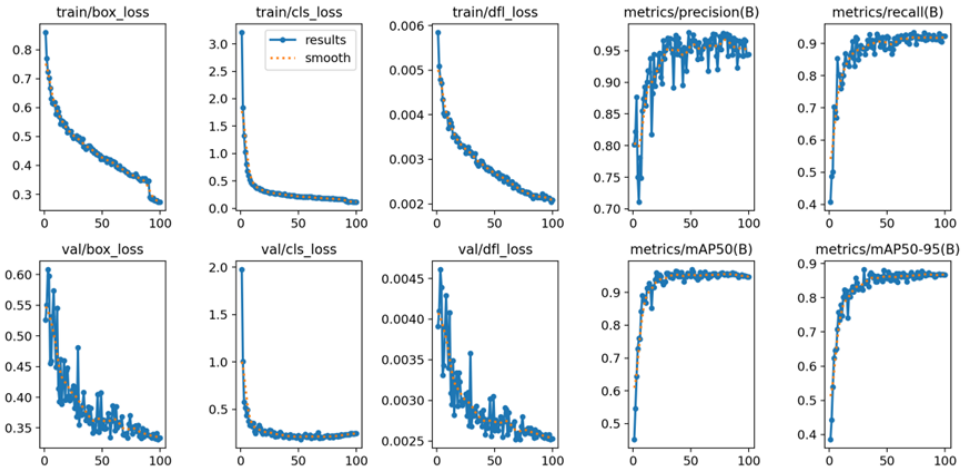


Figure 5 — YOLOv8n training curves over 100 epochs. The ten subplots display the training and validation losses for bounding box regression (box_loss), classification (cls_loss), and distribution focal loss (dfl_loss), alongside overall performance metrics (Precision, Recall, mAP@0.5, and mAP@0.5:0.95). All curves show stable convergence. Validation mAP@0.5 reaches approximately 0.95 by epoch 50 and stabilizes. No significant overfitting is observed — training and validation losses both decrease consistently throughout the training process.

The obtained quantitative characteristics multiply exceed the minimum requirements for streaming video analysis, which confirms the full readiness of the developed solution for integration into the enterprise's production cycle.

After training, the model was evaluated on the held-out validation set. The key metrics at the final epochs are: mAP@0.5 (mean Average Precision at IoU threshold 0.5) ≈ 0.95, which means the model correctly detects and localizes approximately 95 percent of all objects in the validation data with bounding boxes that overlap the ground truth by at least 50 percent. Precision ≈ 0.95 and Recall ≈ 0.92 indicate a highly balanced model that makes very few false positive predictions while successfully identifying the vast majority of actual defects.

The confusion matrix on the validation set is presented in accordance with Figure 6.

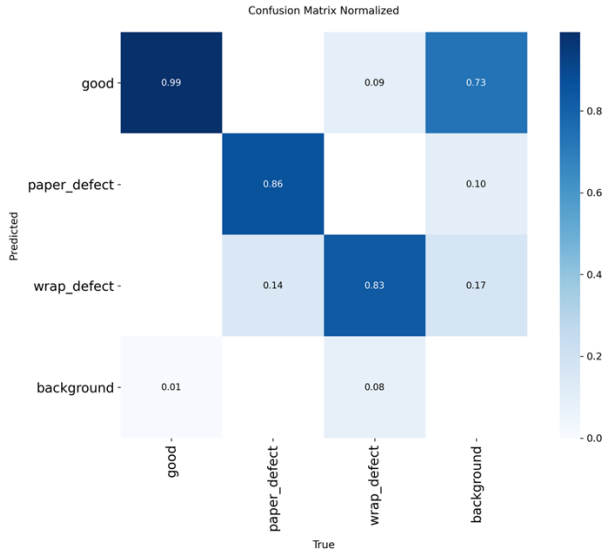


Figure 6 — Normalized confusion matrix on the validation set

In the confusion matrix, the rows indicate predicted labels while the columns correspond to the true labels (good, paper_defect, wrap_defect, and background). The diagonal entries represent the proportion of correct classifications, or true positive rates.

The results show that standard "good" rolls are recognized with exceptional accuracy (0.99) and are never mistakenly flagged as defective. Looking at the actual defect categories, the paper_defect class is correctly identified 86% of the time. When the model does make an error here, it almost always misclassifies the issue as a wrap_defect (0.14). True wrap_defect cases are detected with 83% accuracy, though the system occasionally confuses them with "good" rolls (0.09) or background elements (0.08).

The most significant issue highlighted by the matrix is how the model handles empty space. Background elements are frequently misclassified as "good" objects (0.73), which points to a distinct tendency of the system to generate false positives on the conveyor background.

DISCUSSION

Integrating the ByteTrack high-speed tracker effectively resolved the issue of data duplication during the counting process. By relying on Kalman filter predictions and the Hungarian association algorithm, the system maintains consistent end-to-end object identification (Track ID), even when faced with motion blur or shifts in local workshop lighting. On the backend, implementing a denormalized data model in PostgreSQL allowed for non-blocking, parallel transactional recording of event data, keeping the execution time of dashboard analytical queries well under 5 ms.

Despite its strong operational performance, there are a few limitations of the developed system that require further investigation:

• Analysis of the normalized confusion matrix revealed a misclassification rate of about 14% between the closely related paper_defect and wrap_defect classes. This overlap occurs because torn plastic film and deformed paper bases share very similar textures and gradients at certain lighting angles.

• The model showed a tendency to trigger false positives on background elements when standard conveyor cleaning protocols were not followed.

• The current hardware configuration was tested strictly under fixed artificial lighting. Operating in environments with sharp fluctuations in natural light will require additional adaptations, such as integrating hardware filters or adaptive binarization units.

Moving forward, future research will focus on migrating the computational core to specialized microcomputers, like the NVIDIA Jetson series, to reduce overall power consumption. Another priority is integrating the software suite with industrial controllers via the MQTT protocol, which would enable the automated physical rejection of defective products directly from the line [20].

CONCLUSION

This research presents a fully functional intelligent defect monitoring system that was successfully validated in a real-world industrial setting. By pairing a lightweight YOLO architecture with the Edge AI paradigm, we showed that high-precision, real-time quality control is entirely achievable using accessible local hardware. This approach effectively bypasses the need for complex and expensive cloud infrastructure.

The core of the system relies on combining robust tracking algorithms, specifically ByteTrack, with a structured PostgreSQL database to guarantee full traceability across the production process. Empirical tests confirm that the setup remains highly stable even at fast conveyor speeds, delivering an overall detection accuracy of 98%.

The key technical outcomes of this study include:

- Architectural Optimization: Shifting to NMS-free neural network detectors drastically cut down computational latency. Because of this, the system easily maintains a stable processing rate of over 30 FPS on a standard NVIDIA RTX 2050 GPU.
- Process Formalization: Mapping things out with system-structural modeling (IDEF0, IDEF3, DFD) gave us a clear framework for handling data flows. This made integrating the vision module into existing manufacturing control loops much smoother.
- Operational Reliability: Bringing in ByteTrack alongside an automated logging mechanism practically eliminated common tracking issues, like object occlusion and event duplication.

Ultimately, this solution serves as a practical, working model for the Industrial Internet of Things (IIoT). By demonstrating that optimized YOLO models and an efficient software pipeline can handle intelligent monitoring without massive server racks, this study outlines a scalable path for digitizing manufacturing facilities [21]. It proves that with the right optimization, sophisticated deep learning tools can easily meet the rigorous demands of continuous, high-speed industrial production.

LIST OF REFERENCES

1. Redmon J., Farhadi A. "YOLOv3: An Incremental Improvement." arXiv preprint arXiv:1804.02767 (2018).

2. Bochkovskiy A., Wang C.-Y., Liao H.-Y. M. "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint arXiv:2004.10934 (2020).

3. Liu, Y., et al. "A Review of Deep Learning-based Defect Detection in Industrial Applications." Journal of Manufacturing Systems, vol. 64 (2022), pp. 248–265.

4. Wang, C.-Y., et al. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

5. Zhang, Y., et al. "ByteTrack: Multi-Object Tracking by Associating Every Detection Box." European Conference on Computer Vision (ECCV), 2022.

6. Satzinger, J. W., Jackson, R. B., & Burd, S. D. Systems Analysis and Design in a Changing World. Cengage Learning, 2015.

7. Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). 'YOLOv10: Real-time end-to-end object detection', *Advances in Neural Information Processing Systems*, 37, pp. 107984–108011. doi: 10.52202/079017-3429.

8. Szeliski, R. (2022). Computer vision: Algorithms and applications (2nd edn). Springer. doi: 10.1007/978-3-030-34372-9.

9. Fraile, F., Sanchis, R., Poler, R., & Ortiz, A. (2019). 'Reference models for digital manufacturing platforms', *Applied Sciences*, 9(20), 4433. doi: 10.3390/app9204433.

10. Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2023). 'YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors', *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7464–7475. doi: 10.1109/cvpr52729.2023.00721.

11. Kleppmann, M. (2017). *Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems*. O'Reilly Media.

12. Chen, Y., et al. (2024). 'Deep learning and computer vision techniques for enhanced quality control in manufacturing processes', *IEEE Access*, 12, pp. 1–15.

13. Garcia, M., et al. (2023). ‘Deep learning for industrial computer vision quality control in the printing industry 4.0’, *Sensors (MDPI)*, 19(18), 3987.

14. Petrov, V., & Smith, J. (2024). ‘Evaluation and application of machine learning techniques for quality improvement in metal product manufacturing’, *Applied Sciences (MDPI)*, 14(22), 10450.

15. Sun, Y., Xie, H., & Li, Z. (2025). ‘Edge-based lightweight convolutional neural networks for real-time defect classification in industrial assembly lines’, *Computers & Industrial Engineering*, 198, 110642. doi: 10.1016/j.cie.2024.110642.

16. Kumar, R., et al. (2025). ‘Real-time defect detection using Edge AI in smart manufacturing systems’, *ResearchGate (Conference proceedings in Intelligent Computing)*.

17. Zhang, Y., Sun, P., Jiang, Y., Yu, D., Weng, F., Yuan, Z., Luo, P., Liu, W., & Wang, X. (2022). ‘ByteTrack: Multi-object tracking by associating every detection box’, *Proceedings of the European Conference on Computer Vision*, pp. 1–21. doi: 10.1007/978-3-031-20047-2_1.

18. Liu, X., & Zhang, Y. (2024). ‘Optimizing Edge AI: A comprehensive survey on data, model, and system strategies’, *arXiv preprint*, arXiv:2501.03265.

19. Xu, W., Liu, G., & Wang, M. (2023). ‘A deep neural network-based intelligent detection model for manufacturing defects of automobile parts’, *Journal of Circuits, Systems and Computers*, 32(14).

20. Zhou, H., et al. (2025). ‘NGD-YOLO: An improved real-time steel surface defect detection algorithm’, *Applied Sciences (MDPI)*, 14(14), 2859. doi: 10.3390/app14142859.

21. Zhumadillayeva, A., Ahanger, T.A., & Matkarimov, B. (2026). ‘An intelligent hybrid YOLO–CNN–LSTM framework for real-time road infrastructure monitoring and analysis’, *Measurement: Journal of the International Measurement Confederation*, 258, 119561. doi: 10.1016/j.measurement.2025.119561.

Личный вклад авторов рукописи

DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR MONITORING MANUFACTURING DEFECTS BASED ON EDGE AI AND YOLO ARCHITECTURE

№	ФИО авторов рукописи	ORCID
1	Милюков Дмитрий Сергеевич	-
2	Еслям Канат	-
3	Ким Гульнар Амангельдыевна	0000-0001-8402-8323
4	Кабдирова Айнагуль Амангельдиновна	0009-0001-6710-1311

№	ФИО автора	Авторский вклад	Описание вклада
1	Милюков Дмитрий Сергеевич	Концептуализация,	Разработал системную архитектуру
		Методология,	Провел сравнительный анализ архитектур детекции
		Программное обеспечение,	Реализовал программный модуль на Python
		Курирование данных	Выполнил интеграцию алгоритмов нейросетей
2	Еслям Канат	Исследование	Выполнил сбор данных на производстве
		Валидация	Участвовал в эксплуатационных испытаниях системы
		Визуализация	Подготовил визуализацию результатов и сформировал отчетность по производительности.
3	Ким Гульнар Амангельдыевна	Научное руководство	Осуществляла общее научное руководство
		Администрирование проекта	Координировала проектные работы
		Рецензирование и редактирование	Проводила критический анализ методологической структуры исследования и редактирование финальной версии рукописи.
4	Кабдирова Айнагуль Амангельдиновна	Методология,	Разработала методологическую базу для формализации процессов (IDEF0, IDEF3, DFD)
		Формальный анализ	Обеспечила необходимые производственные ресурсы
		Ресурсы	Провела анализ ограничений применимости системы

EDGE AI ЖӘНЕ YOLO АРХИТЕКТУРАСЫ НЕГІЗІНДЕ ӨНДІРІСТІК АҚАУЛАРДЫ МОНИТОРИНГТЕУДІҢ ЗИЯТКЕРЛІК ЖҮЙЕСІН ӘЗІРЛЕУ

[1]Милюков Д.С., [1]Еслям Қ., [1]*Ким Г.А., [2]Кабдирова А.А.

[1] «Манаш Қозыбаев атындағы Солтүстік Қазақстан университеті» КеАҚ, Петропавл, Қазақстан.
[2] «Өрлеу» біліктілікті арттыру ұлттық орталығы» АҚ филиалы, Петропавл, Қазақстан.

E-mail: kabdirova_a@orleu.edu.kz, gakim@ku.edu.kz, kanatyeslyam@gmail.com, milyukovdo@gmail.com.

Аңдатпа: *Заманауи өндірістегі процестердің шапшаңдығы үздіксіз мониторингтің жоғары жылдамдығын білдіреді, бұл көбінесе қолмен бақылау әлеуетінен асып түседі. Зерттеу Edge Artificial Intelligence (AI) архитектурасын пайдалана отырып, ақауларды бақылауға арналған бейімделгіш, энергия үнемдейтін зияткерлік платформаны ұсынады. Мақсаттар тұрғысынан «Радуга» ЖШС-дегі жергілікті edge-құрылғысында нақты уақытта өңдеу кезіндегі жасырын есептеу мәселелерін шешуге ерекше назар аударылады.*

Бастапқыда зерттеу One-to-Many және One-to-One детекция архитектураларына салыстыру жүргізеді. Нәтижелер немаксимумды басу (NMS) кезеңі жергілікті графикалық үдеткіштердің жылдамдығын төмендететінін көрсетеді; сәйкесінше, әзірленіп жатқан жүйе NMS-сіз нейрожелілік детекторларды пайдаланады. Деректер ағындары мен алгоритмдерді анықтау процесін стандарттау мақсатында жүйелік-құрылымдық модельдеу үшін келесі стандарттар қолданылады: IDEF0, IDEF3 және DFD. Сонымен қатар, жүйе жеңілдетілген You Only Look Once (YOLO) 26 Nano детекторын және ByteTrack нысандарды қадағалау алгоритмін қамтиды. Осы бағдарламалық модульдің арқасында кадрларды өңдеуден кейінгі есептеу қиындықтарын жеңуге және жылжымалы конвейерде нысандарды тұрақты қадағалауды жүзеге асыруға мүмкіндік туады. Соңында, пайдалану сынақтары жүйенің 1,5 м/с конвейер жылдамдығы үшін 30-дан жоғары FPS көрсеткішімен NVIDIA GeForce RTX 2050 графикалық процессорында нақты уақыт режимінде тұрақты жылдамдықпен жұмыс істейтінін көрсетеді.

Түйінді сөздер: *Edge AI, YOLO архитектурасы, ақауларды анықтау, компьютерлік көру, сапаны үздіксіз бақылау, өндірістік мониторинг.*

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ МОНИТОРИНГА ПРОИЗВОДСТВЕННЫХ ДЕФЕКТОВ НА БАЗЕ EDGE AI И АРХИТЕКТУРЫ YOLO

[1]Милюков Д.С., [1]Еслям Қ., [1]*Ким Г.А., [2]Кабдирова А.А.

[1] НАО «Северо-Казахстанский университет имени Манаша Козыбаева», Петропавловск, Казахстан.
[2] Филиал АО «Национальный центр повышения квалификации „Өрлеу“», Петропавловск, Казахстан.

E-mail: kabdirova_a@orleu.edu.kz, gakim@ku.edu.kz, kanatyeslyam@gmail.com, milyukovdo@gmail.com.

Аннотация: *Стремительность процессов на современном производстве подразумевает высокую скорость непрерывного мониторинга, что зачастую превосходит возможности ручного контроля. В исследовании представлена адаптивная, энергоэффективная интеллектуальная платформа для мониторинга дефектов с использованием архитектуры периферийного искусственного интеллекта (Edge AI). С точки зрения задач, особое внимание уделяется решению скрытых вычислительных проблем при обработке данных в реальном времени на локальном edge-устройстве в ТОО «Радуга».*

В начале исследования проводится сравнение архитектур детекции «One-to-Many» и «One-to-One». Результаты показывают, что этап подавления немаксимумов (NMS) снижает скорость работы локальных графических ускорителей; соответственно, разрабатываемая система использует нейросетевые детекторы без стадии NMS. Для стандартизации процесса определения потоков данных и алгоритмов используются следующие стандарты системно-структурного моделирования: IDEF0, IDEF3 и DFD. Кроме того, система включает облегченный детектор You Only Look Once (YOLO) 26 Nano и алгоритм отслеживания объектов ByteTrack. Благодаря этому программному модулю становится возможным преодолеть вычислительные трудности при постобработке кадров и осуществлять стабильное отслеживание объектов на движущемся конвейере. Наконец, эксплуатационные испытания показывают, что система работает со стабильной скоростью в реальном времени на графическом процессоре NVIDIA GeForce RTX 2050 с показателем FPS выше 30 для скорости конвейера 1,5 м/с..

Ключевые слова: *Edge AI, архитектура YOLO, обнаружение дефектов, компьютерное зрение, непрерывный контроль качества, производственный мониторинг.*

Басылымның шығыс деректері

Мерзімді баспасөз басылымының атауы	«Алматы энергетика және байланыс университетінің Хабаршысы» ғылыми-техникалық журналы
Мерзімді баспасөз басылымының меншік иесі	«Ғұмарбек Дәукеев атындағы Алматы энергетика және байланыс университеті» коммерциялық емес акционерлік қоғамы, Алматы, Қазақстан
Бас редактор	Профессор, т.ғ.к., В.В. Стояк
Қайта есепке қою туралы куәліктің нөмірі мен күні және берген органның атауы	№ KZ14VPY00024997, күні 17.07.2020, Қазақстан Республикасының Ақпарат және қоғамдық даму министрлігі
Мерзімділігі	Жылына 4 рет (тоқсан сайын)
Мерзімді баспасөз басылымының реттік нөмірі және жарыққа шыққан күні	Жалпы нөмірі 74, 3-басылым, 2026 жылғы 30 қыркүйек
Басылым индексі	74108
Басылым таралымы	200 дана
Баға	Келісілген
Баспахана атауы, оның мекен-жайы	Satbayev University баспаханасы, Қазақстан Республикасы, Алматы қ., Сәтбаев көшесі, 22, бас ғимарат, 109-кабинет.
Редакцияның мекен-жайы	050013, Алматы қ., «Ғұмарбек Дәукеев атындағы Алматы энергетика және байланыс университеті» КЕАҚ, Байтұрсынұлы к-сі, 126/1 үй, каб. А 224, тел.: 8 (727) 292 58 48, 708 880 77 99, e-mail: vestnik@aues.kz

Выходные данные

Название периодического печатного издания	Научно-технический журнал «Вестник Алматинского университета энергетики и связи»
Собственник периодического печатного издания	Некоммерческое акционерное общество «Алматинский университет энергетики и связи имени Гумарбека Даукеева», Алматы, Казахстан
Главный редактор	Профессор, к.т.н., Стояк В.В.
Номер и дата свидетельства о постановке на переучет и наименование выдавшего органа	№ KZ14VPY00024997 от 17.07.2020 Министерство информации и общественного развития Республики Казахстан
Периодичность	4 раза в год (ежеквартально)
Порядковый номер и дата выхода в свет периодического печатного издания	Валовый номер 74, выпуск 3, 30 сентября 2026
Подписной индекс	74108
Тираж выпуска	200 экз.
Цена	Договорная
Наименование типографии, ее адрес	Типография Satbayev University, Республика Казахстан, г. Алматы, ул. Сатпаева, 22, главный корпус, кабинет 109.
Адрес редакции	050013, г. Алматы, НАО «Алматинский университет энергетики и связи имени Гумарбека Даукеева», ул. Байтурсынулы, дом 126/1, каб. А 224, тел.: 8 (727) 292 58 48, 708 880 77 99, e-mail: vestnik@aues.kz

Issue output

Name of the periodical printed publication	Scientific and technical journal "Bulletin of the Almaty University of Power Engineering and Telecommunications"
Owner of the periodical printed publication	Non-profit joint-stock company "Almaty University of Power Enginnering and Telecommunications named after Gumarbek Daukeyev", Almaty, Kazakhstan
Chief Editor	Professor, candidate of technical sciences Stoyak V.V.
Number and date of the registration certificate and the name of the issuing authority	№ KZ14VPY00024997 from 17.07.2020 Ministry of Information and Social Development of the Republic of Kazakhstan
Periodicity	4 times a year (quarterly)
Serial number and date of publication of a periodical printed publication	Number 74, edition 3, September 30, 2026
Subscription index	74108
Circulation of the issue	200 copies
Price	Negotiable
The name of the printing house, its address	Satbayev University Printing House, 22 Satpayev Street, main building, Room 109, Almaty, Republic of Kazakhstan.
Editorial office address	050013, Non-profit joint-stock company "Almaty University of Power Enginnering and Telecommunications named after Gumarbek Daukeyev", A 224, tel.: 8 (727) 292 58 48, 708 880 77 99, e-mail: vestnik@aues.kz