N. Zhumakhan[1]

B. Belgibayev[1]

K.Togzhanova[2]

G. Kashaganova[2]

INVESTIGATION OF THE TECHNICAL VISION OF FPV ROBOTS WITH IMAGE PROCESSING AND RECOGNITION

[1]Al Farabi Kazakh National University
[2]Almaty Technological University

Abstract: This review analyzes modern vision systems used in FPV (First-Person View) robots, with a focus on image processing and object recognition technologies. Based on recent scientific literature, the study compares neural network architectures, sensor fusion methods, and implementation approaches under real-time and resource-constrained conditions. A comparative analysis of key technologies (YOLOv8, Squeeze-EnGAN, LiDAR, RGB cameras, and others) is conducted. The findings reveal strengths and trade-offs relevant for autonomous navigation, robotic manipulation, and inspection tasks. Future prospects include edge AI integration, sensor optimization, and robust vision under low light and environmental variability.

Keywords: FPV robots, technical vision, image recognition, neural networks, LiDAR, sensor fusion, computer vision, deep learning

Introduction

Technical vision is a critical area in the context of developing autonomous FPV (first-person) robots. These systems require high accuracy of perception of the environment, optimal response time and processing of visual information in real time. This article provides an overview of current research in the field of image processing and recognition for FPV robots.

Basic research overview

Advanced GAN models, such as Squeeze-EnGAN, can improve image quality in low light conditions, which is critical for FPV drones at night or underground. The authors Haegyo In, Juhum Kweon and Changjoo Moon presented this architecture as an energy-efficient model operating in a non-supervised learning mode [1].

In industry and agriculture, FPV robots can sort fruits or objects on a conveyor belt. In the article by the authors Zhakpher Alihanov, Aidar Moldazhanov, Akmaral Kulmakhambetova and others. It is shown how digital image analysis makes it possible to evaluate the physical parameters of apples with high accuracy, which is important for agricultural robotics [2].

Industrial vision technologies were demonstrated in the work of Haozhan Qu, Jue Wang, and others, where a real-time robotic oyster capture system was implemented based on YOLOv8-OBB with improvements via MobileNetV4. The use of the Kalman filter ensures effective tracking of objects [3].

For spatial perception tasks, it is important to consider sensor comparisons. In the work of Peide Wang [4], a comparative analysis was carried out between LiDAR and cameras, where it was concluded that their fusion was promising to increase the reliability of scene perception by autonomous vehicles and FPV robots.

Additionally, in the classic work by Hiroyuki Kawasaki and co-authors [5], a combination of omnidirectional sonar and omnidirectional vision was proposed, which proved its effectiveness for reliable environmental recognition by mobile platforms.

Finally, an article by Marius Boshoff, Bernd Kulenkampff, and Paul Koslowski [6] explores the use of UAVs as a mobile camera for optimal positioning in visual surveillance and route planning tasks, which is applicable to FPV scenarios with dynamic viewpoint control.

Based on these data, a comparative analysis can be made in Table 1 of the vision technologies for FPV robots.

Table 1. Comparative analysis of technical vision technologies for FPV robots

Technology / Method	**Advantages**	**Limitations**	**Scope of application**
Squeeze-EnGAN	Improves image quality in low light conditions; energy efficient	Limited applicability outside low light conditions	Nighttime FPV systems, underground environments
YOLOv8-OBB + MobileNetV4	High accuracy and speed; shape and orientation detection	Difficulty in learning and adapting to new classes of objects	Industrial manipulators and sorting lines
Digital display image analysis (apples)	High accuracy of measurement of geometrical parameters	Dependence on shooting and calibration conditions	Agricultural LiDAR product quality assessment systems
LiDAR	High accuracy of 3D positioning and range	High cost; weather sensitivity	Spatial navigation, autonomous vehicles
Cameras (RGB)	High resolution, rich visual information	Sensitivity to light and weather conditions	Overall monitoring and control
Merging LiDAR + camera fusion	K Data complementarity, resilience in complex environments	Need for synchronization and calibration	Advanced scene perception, high-level AI tasks
Omnidirectional vision + sonar	Full scene coverage, independent of viewing angle	Limited object detail All-	round mobile platforms
	Flexible view, reduced blind spots	Need for synchronization with FPV tasks	Monitoring, inspection, robot surveillance

The following Figure 1 shows a visual comparison of vision technologies based on five key criteria: Accuracy, Energy Efficiency, Real-Time Suitability, Cost, and Environmental Robustness.

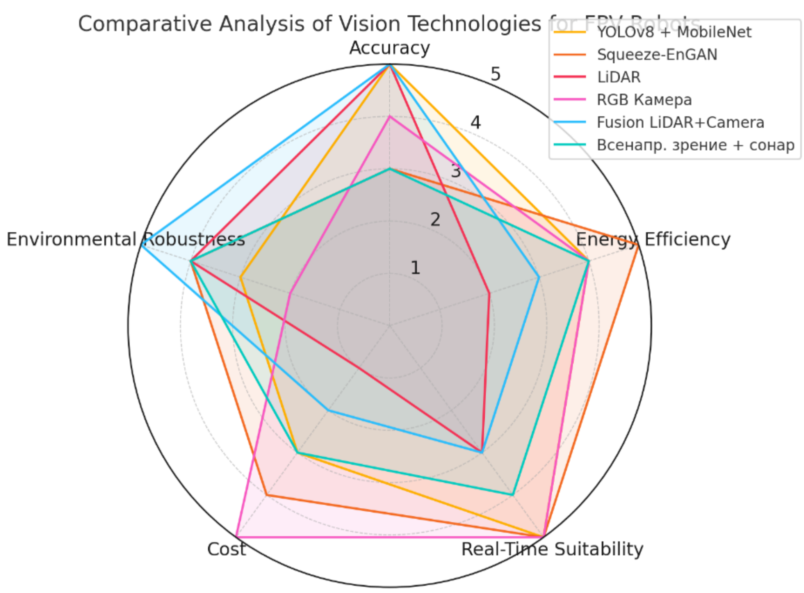

Figure 1. Comparative Analysis of Vision Technologies Based on Key Performance Criteria

The graph allows you to visually assess which technologies are most balanced for use in FPV robots. For example, YOLOv8 with MobileNet demonstrates high accuracy and performance, while Squeeze-

EnGAN provides better energy efficiency. LiDAR and data fusion systems show maximum resilience in a complex environment, but require more resources.

Conclusions

Modern research shows that technical vision for FPV robots is developing in several key areas: improving image quality in difficult lighting conditions, increasing the accuracy of object recognition, as well as integrating multimodal sensors to increase the reliability of perception. Lightweight deep learning models (YOLOv8, MobileNet, Squeeze-EnGAN) demonstrate high applicability in conditions of limited computing resources typical of autonomous platforms.

Future research should focus on: (1) the development of energy-efficient models; (2) Integration of edge AI for on-board processing; (3) more complete fusion of data from visual, ultrasonic and laser sensors. This will improve the reliability and adaptability of FPV systems in real-world applications.

REFERENCES

1. In H., Kweon J., Moon C. Squeeze-EnGAN: Memory Efficient and Unsupervised Low-Light Image Enhancement for Intelligent Vehicles // Sensors. 2025. Vol. 25, No. 6. P. 1825. DOI: 10.3390/s25061825
2. Alihanov Z., Moldazhanov A., Kulmakhambetova A., Zinchenko D., Nurtuleuov A., Shynybai Z., Georgieva T., Daskalov P. Methodology for Determining Basic Physical Parameters of Apples Using Digital Image Analysis // AgriEngineering. 2025. Vol. 7, No. 3. P. 57. DOI: 10.3390/agriengineering7030057
3. Qu H., Wang J., Lei L., Su W. Computer Vision in Robotic Systems: Oyster Identification and Grasping in Real Time // Applied Sciences. 2025. Vol. 15, No. 7. P. 3971. DOI: 10.3390/app15073971
4. Wang P. Comparative study of LiDAR and camera in autonomous driving // Journal of Physics: Conference Series. 2021. Vol. 2093, No. 1. P. 012032. DOI: 10.1088/1742-6596/2093/1/012032
5. Kawasaki H., Furukawa H., Tamura H., Kaneko S. Fusion of omnidirectional sonar and omnidirectional vision for mobile robot environment recognition // Proceedings of the 2000 ICRA. IEEE International Conference on Robotics and Automation. 2000. Vol. 1. P. 894–899. DOI: 10.1109/ROBOT.2000.845343
6. Boshoff M., Kulenkampff B., Koslowski P. Dynamic Camera Planning for Robotic Manufacturing Processes Using UAVs // Robotics. 2025. Vol. 14, No. 3. P. 23. DOI: 10.3390/robotics14030023

Zhumakhan Nurzhan Beibituly, 1st year doctoral student , Al Farabi Kazakh National University, Almaty ,Kazakhstan , nurzahan_14_95@mail.ru

Belgibayev, Baurzhan Abdrahimovich, Doctor of Technical Sciences, Associate Professor, Al Farabi Kazakh National University, Almaty ,Kazakhstan

Kashaganova Gulzhan Bakhytovna, PhD, Associate Professor, Almaty Technological University, Almaty, Kazakhstan

Togzhanova Kulzhan Ondrisovna,PhD, Associate Professor, Almaty Technological University, Almaty, Kazakhstan, togzhanova_kuljan@mail.ru