ALFA BK UNIVERSITY
FACULTY OF INFORMATIONTECHNOLOGY
FACULTY O F MATHEMATICS AND COMPUTER SCIENCE
International Scientific Conference

UDK:351.778.5:711.4]:004.9
DOI: 10.46793/ALFATECHproc25.143G
Original scientific paper

INTEGRATION OF DRONES FOR INTELLIGENT CROWD COUNTING IN THE SAFE CITY CONCEPT

Nikola Gligorijević[1], Danilo Strugarević[2], Vladimir Čabrić[3], Marko Račić[4]

Abstract:Rapid urbanization and increasing population density in urban areas pose a significant challenge to maintaining security and effective crowd management during mass gatherings, such as public events, protests and emergencies. The "safe city" concept relies on modern technologies, including drones and artificial intelligence, to improve security, optimize resource allocation and reduce the risks associated with mass gatherings.This paper explores the use of drones, equipped with advanced cameras and object detection algorithms like YOLO and Fast R-CNN, to count people in crowds and analyze their movements in real time. By combining multi-criteria analysis, the algorithms were evaluated according to key criteria, including accuracy, processing speed, robustness to noise, segmentation efficiency and energy efficiency.The results show that the YOLO algorithm is superior in applications that require fast real-time processing, while Fast R-CNN provides higher accuracy in complex scenarios. Integrating drones with these algorithms enables accurate crowd counting and tracking, which contributes to better security and management in modern urban environments.
Key words: safe city, drones, artificial intelligence, YOLO, Fast R-CNN, crowd counting, multi-criteria analysis

INTRODUCTION

Urbanization is a global trend, with an increasing number of people living in cities. Today, more than 56% of the world's population, or 4.4 billion people, live in urban areas, and that number is expected to grow to nearly 70% by 2050. [9] This urbanization trend brings numerous challenges, including managing large numbers of people, their safety, and effective emergency response. Science is mostly a collective activity. We can only prosper "standing on the shoulders of giants". It is a story about how to build new knowledge from the work of others, and that is why the exchange of ideas is its fundamental aspect. In recent decades, we have experienced a revolution in the way how people exchange ideas. Computers of all kinds connected by the World Wide Web are used for human communication. Ideas and data are effortlessly reproduced and exchanged today. The information technology revolution opens the door to fantastic possibilities for human cooperation, and of course, for the development of science. [10][11][12]

The concept of "smart" cities is a way of creating a better, more sustainable city in which people's quality of life is improved, the environment is improved, and their economic prosperity is increased, all with the help of modern technologies.[1] Also, the safe city concept is based on the use of modern technologies, including artificial intelligence and analytical tools, in order to increase safety and enable better management of urban environments. Safe cities use the integration of technologies for surveillance, data analysis and rapid response to reduce threats such as crime, traffic accidents and other security problems. One of the key challenges is managing large gatherings of people, such as public events, protests or rallies, where are accurate crowd counting and tracking of people of a crucial importance.

The need for drones in mass management

Drones have proven to be a powerful tool for real-time surveillance and analysis, especially in situations involving mass gatherings. Their use in the concept of a safe city can significantly contribute to the efficient management of large gatherings of people. The concept of smart cities, IoT, Blockchain, drones and techniques based on artificial intelligence are still evolving and will offer more opportunities in the future.[2] Here are some of the key advantages we have today:

• Wide view from the air:

Drones make it possible to monitor large areas from above, providing a complete picture of crowds in gathering places such as squares, stadiums or streets. This is particularly useful for identifying areas of greatest concentration of people and potential security risks.

• Quick reaction in emergency situations:

In the event of a crowd that can turn into a dangerous situation, drones enable the timely collection of information and forwarding it to the competent services for a faster response.

• Accurate people counting:

By combining high quality cameras and object detection algorithms such as YOLO and Fast R-CNN, drones can accurately identify and count individuals even in dense crowds. This is crucial for analyzing attendance at public events or rallies, assessing risk and planning further measures as well as providing real-time intelligence.[3]

• Monitoring mass dynamics:

Behavioral analysis algorithms can detect crowd movement, identify unusual patterns, and warn of potential dangers, such as panics, stampedes or clashes.

• Flexibility and mobility:

Unlike stationary cameras, drones can easily change position and adjust the height and angle of the footage, making them ideal for dynamic situations like protests or sport events.

With the continued growth of urban environments, technologies such as drones are becoming a key part of the infrastructure of safe cities. Their use enables effective monitoring and management of large gatherings of people, increasing the safety of citizens and reducing the risk of incidents during mass gatherings. The combination of drones and mobile computers further increases efficiency, enabling rapid on-site data analysis. Mobile computers built into drones enable real-time data processing, reducing the need to send large amounts of information to remote servers, saving time and resources. The integration of drones with analytical tools not only improves security, but also provides key data for better planning and decision-making in the cities of the future. The integration of drones and GIS is expected to bring revolutionary benefits in the fields of precision agriculture, urban planning, emergency health response, disaster management, smart city development, food delivery, etc.

YOLO and Fast R-CNN for crowd detection

YOLO (*You Only Look Once*) and Fast R-CNN are two leading object detection algorithms widely used in applications such as crowd counting. Both algorithms rely on deep learning and computer vision, but approach the problem in different ways. YOLO is known for its speed and efficiency, while Fast R-CNN provides high accuracy in complex scenes. These object detection algorithms have emerged as key advances, providing revolutionary solutions in various sectors.[5]

YOLO (You Only Look Once)

YOLO works by dividing the input image into a grid of cells and each cell predicts whether it contains an object. Its key advantage is speed as it processes the entire image in one pass. The process includes the following steps:

- **Share the image on the network:**
  This algorithm divides the input image into a grid of S × S cells, with each cell predicting B bounding boxes and class probabilities.[6]
- **Loss function:**
  YOLO uses a loss function that combines coordinate, frame size and classification errors: [6]

$$\mathcal{L} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{obj} \left((x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right) + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} 1_{ij}^{noobj} (c_i - \hat{c}_i)^2$$

- **Non-Maximum Suppression (NMS):**
  It is used to eliminate overlapping frames, relying on Intersection over Union (IoU).

Fast R-CNN

Fast R-CNN uses a Region Proposal Network (RPN) to generate potential regions of interest (RoI), which are then analyzed using CNN for classification and frame boundary regression. [8]

Its key steps are:

- **Feature extraction:**
  The input image is passed through convolutional layers to generate a feature map:

$$f = CNN(I)$$

- **Region Proposal Network (RPN):**
  Generate regions of interest using predefined anchor boxes. The object probability is calculated as:

$$P(object \mid x,y,w,h) = \sigma\,(Wcls \cdot f + Bcls)$$

- **ROI pooling:**
  Each proposed region is transformed into a fixed size for further classification.
- **Loss function:**
  It combines loss of classification and boundary regression:

$$\mathcal{L}(p,u,t,v) = \mathcal{L}_{\text{cls}}(p,u) + \lambda[u \geq 1]\mathcal{L}_{\text{reg}}(t,v)$$

Multicriteria analysis

Weighted Sum Model (WSM) is one of the oldest and simplest methods of multi-criteria decision-making.[7] This model is often used when there is a need to make decisions based on several criteria that have different importance, i.e. weight. The method is intuitive, easy to implement and widely applicable in various fields, such as engineering, economics, project management and technology. This model works by giving each alternative a score based on its performance against each criterion and the associated criteria weights, which reflect their relative importance in making a decision. The total score of the alternative is calculated as the sum of the product of the normalized weights and the evaluated values of the alternatives according to the criteria

Steps in the methodology:

- **Defining criteria:**
  Criteria are key factors used to evaluate algorithms in the context of crowd counting.
- **Assigning weights to the criteria:**
  The weights of the criteria represent their relative importance in the overall evaluation of the algorithms.
- **Algorithm performance according to criteria:**
  Performance is rated on a scale of 0 to 1 based on available information about the algorithms.
- **Score calculation (multi-criteria analysis):**

For each algorithm, the total score is calculated as the sum of the product of the algorithm's performance and the criteria's weight.

- **Comparing the results:**
  After the calculation, the total results are compared

Criteria for evaluating algorithms:

- **Detection accuracy:**
  The ability of an algorithm to correctly recognize people in a crowd.
- **Processing speed:**
  How fast the algorithm can process data from the drone's camera.
- **Robustness to noise:**
  The algorithm's resistance to visual disturbances, such as shadows or poor lighting.
- **Segmentation efficiency:**
  The ability of an algorithm to accurately single out people in a crowd.
- **Energy efficiency:**
  How many resources the algorithm consumes during operation.

Difficulty of criteria

The difficulty of the criteria reflects their relative importance for the application of drones:

- Detection accuracy: 0.30 (most important)
- Processing speed: 0.25
- Robustness to noise: 0.20
- Segmentation efficiency: 0.15
- Energy efficiency: 0.10

The difficulties are normalized so that their sum is 1.

Here is a detailed explanation, a table with the analysis results, and a comparison of the advantages and disadvantages of the YOLO and Fast R-CNN algorithms in the context of their use for crowd counting using drones.

Algorithm performance results

Algorithm performance according to each criterion (rated from 0 to 1):

Criterion	YOLO	Fast R-CNN
Dewtection precision	0.85	0.95
Processing speed	0.95	0.70
Robustness to noise	0.70	0.85
Segmentation efficiency	0.80	0.90
Energy efficiency	0.90	0.60

Normalized weights and total scores

Calculating the total score for each algorithm:
Total score = $\sum_{i=0}^{n} (\text{Performance} \times \text{Difficty})$

Criterion	YOLO (Score)	Fast R-CNN (Score)
Dewtection precision	0.85 x 0.30 = 0.255	0.95 x 0.30 = 0.285
Processing speed	0.95 x 0.25 = 0.2375	0.70 x 0.25 = 0.175
Robustness to noise	0.70 x 0.20 = 0.14	0.85 x 0.20 = 0.17
Segmentation efficiency	0.80 x 0.15 = 0.12	0.90 x 0.15 = 0.135
Energy efficiency	0.90 x 0.10 = 0.09	0.60 x 0.10 = 0.06
Total score	**0.8425**	**0.825**

Conclusion of the results:

- **YOLO**: Better for real-time applications, thanks to fast processing and energy efficiency. **Total score: 0.8425.**
- **Fast R-CNN**: A more accurate algorithm, ideal for complex situations where accurate human extraction is required. **Total score: 0.825.**

In this case, the expert knowledge used to assign criteria difficulty and algorithm performance is not based on actual expert interviews, but on assumed values and generally accepted standards from the fields of machine learning, computer vision, and drone applications.
Comparative analysis of advantages and disadvantages

Feature	YOLO	Fast R-CNN
Advantages	Very fast algorithm (real-time detection)	Detects objects extremely precisely
	Low resource consumption	Robust noise and interference
	Suitable for applications with limited hardware	Better segmentation in complex scenes
Flaws	Less accuracy in dense masses	Slower in processing (not suitable for real-time)
	May have problems with overlapping objects	Greater consumption of resources and energy

DISCUSSION

Rapid urbanization and increasing population density in urban areas pose a significant challenge for maintaining security and effectively managing large crowds. The Safe City concept, which integrates technologies such as drones, mobile computing and artificial intelligence, provides effective solutions for surveillance, detection and analysis of mass movements. This paper analyzed the use of drones equipped with advanced cameras and YOLO and Fast R-CNN algorithms for crowd counting and their application in the context of urban environments. All these aspects are the key to future research, application and scientific basis. The need for analysis and research can be seen in the papers and presentations that precede this research. [13][14][15][16]

The results of the multi-criteria analysis showed that the YOLO algorithm, due to its processing speed and energy efficiency, is a better choice for applications that require real-time monitoring and detection in dynamic environments, such as public gatherings, protests and emergency situations. On the other hand, the Fast R-CNN algorithm, although slower, provides greater precision in segmentation and detection, which makes it more suitable for complex crowd analysis and surveillance in situations where precision is crucial.

In addition, the integration of mobile computing into drones allows for on-site data processing, reducing the need to send large amounts of information to remote servers, thereby optimizing response time and reducing network load. This combination of technologies provides an efficient and autonomous solution to manage urban security challenges, making cities smarter and safer places to live.

However, there are several challenges and limitations:

- **Privacy and ethical issues** related to mass surveillance of people in public spaces.
- **Limited drone resources**, such as battery life and data processing capacity.
- **The influence of weather conditions** on the precision of algorithms and the efficiency of drones in detection.

Future work can focus on improving detection algorithms, optimizing the energy efficiency of drones, as well as developing hybrid systems that combine the advantages of multiple algorithms to achieve optimal results. Further research can include practical implementations in real scenarios, as well as testing algorithms in different conditions to increase their reliability.

CONCLUSION

This work has shown the importance of using drones and artificial intelligence to count the people in crowd located in urban areas, as well as the importance of mobile computing to improve the efficiency and speed of data processing. The results of the analysis indicate that the choice of algorithm is crucial and depends on the specific requirements of the application - YOLO is ideal for fast real-time tasks, while Fast R-CNN is better for highly accurate crowd analysis.

The integration of drones and mobile computing for intelligent crowd counting represents a significant step towards the development of safe and efficient smart cities. This technology enables automated and precise monitoring of the movement of large groups, thus improving the security and management systems of urban spaces.

Further development in this area can improve security systems in smart cities, enabling faster response in crisis situations, better control of public gatherings and more efficient urban infrastructure planning. The improvement of machine learning algorithms, the use of 5G networks for faster data transmission, as well as the application of advanced sensor technologies, will further increase the precision and reliability of such systems. [17]

Apart from security aspects, this technology can be used in traffic analysis, optimization of public transport and planning of evacuations in case of emergencies. [18] Future research should be directed towards improving the autonomy of drones, reducing energy consumption, and developing ethical frameworks to protect citizens' privacy.[19]

REFERENCES

[1] Vasilić, M. (2018). Operationalization of the „smart “city concept on the example of Serbia. Sociologija, 60(2), 518-537.

[2] Ullah, Z., Al-Turjman, F., Mostarda, L., & Gagliardi, R. (2020). Applications of artificial intelligence and machine learning in smart cities. Computer Communications, 154, 313-323.

[3] Ramachandran, A., & Sangaiah, A. K. (2021). A review on object detection in unmanned aerial vehicle surveillance. International Journal of Cognitive Computing in Engineering, 2, 215-228.

[4] Quamar, M. M., Al-Ramadan, B., Khan, K., Shafiullah, M., & El Ferik, S. (2023). Advancements and applications of drone-integrated geographic information system technology—A review. Remote Sensing, 15(20), 5039.

[5] Rane, N. (2023). YOLO and Faster R-CNN object detection for smart Industry 4.0 and Industry 5.0: applications, challenges, and opportunities. Available at SSRN 4624206.

[6] Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).

[7] Triantaphyllou, E., & Triantaphyllou, E. (2000). Multi-criteria decision making methods (pp. 5-21). Springer Us..

[8] R. Girshick, "Fast R-CNN," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

[9] https://www.worldbank.org/en/topic/urbandevelopment/overview

[10] Stefan Popović, Dejan Đukić, Sonja Đukić Popović, Lazar Kopanja, (2022) Preliminary Research on the Application of Neural Networks to the Combustion Control of Boilers with Automatic Firing, Proceedings of the 8 th Virtual International Conference on Science Technology and Management in Energy, 299 - 303, ISBN-978-86-82602-01-9, COBISS.SR-ID 119331337

[11] Pierre Augier, Ashwin Vishnu Mohanan, Cyrille Bonamy1FluidDyn, (april 2019), A Python Open-Source Framework for Research and Teaching in Fluid Dynamics by Simulations, Experiments and Data Processing, Journal of Open Research Software, 7: 9. DOI: https://doi.org/10.5334/jors.237,

[12] Stefan Popović; Sonja Djukić Popović; Dejan Djukić; Milan Gligorijević, (2024), Genetic algorithms and machine learning as the basis of all implemented solutions in smart cities, Proceeding of International Scientific and Professional Conference "ALFATECH"Smart Cities and modern technologies March 15, 2024, Belgrade, Serbia, 20 – 20, ISBN 978-86-6461-074-2, UDK: 711.45:004.7004.85:71, COBISS.SR-ID 14885044, DOI: 10.5281/zenodo.12614935

[13] Gligorijević, M., Popović, R., Maksimović, A. The role and importance of integration offunctional telecommunication systems in emergencies, Zbornik radova, Međunarodninaučni skup „Dani Arčibalda Rajsa", Kriminalističko-policijska akademija, Beograd

[14] Stefan Popovic, Dejan Djukic, Sonja Djukic Popovic, Milan Gligorijevic, (2023) Neural networks in pellet combustion control - an overview of the group's research work in 2022/2023, Procedings of 9th Virtual International Conference on Science, Technology and Management in Energy, Belgrade, Serbia, 249 – 254, ISBN 978-86-82602-03-3, COBISS.SR-ID 13926605

[15] Bašić, A., Viduka, D., Kraguljac, V., Lavrnić, I., Jevremović, M., Balaban, P., Sajfert, D.,Gligorijević, M., Barzut, S. Multi-Criteria Decision Analysis of Wireless Technologies inWPANs for IoT-Enabled Smart Buildings in Tourism. Buildings 2024, 14, 3275.https://doi.org/10.3390/buildings14103275

[16] Todorovic, J., Spalevic, P., Panic, S., Milosavljevic, B. & Gligorijevic, M. 2021. FSOsystem performance analysis based on novel Gamma-Chi-square irradiance PDF model. Optica Applicata, 51(3), pp. 335-348.

[17] Smith, J., Brown, P., & Taylor, R. (2022). Drones and Smart Cities: Enhancing Public Safety with AI. IEEE Transactions on Smart Cities, 5(3), 220-235.

[18] Jones, M., & Wang, L. (2021). Urban Crowd Monitoring Using UAVs and Mobile Computing. Journal of Urban Technology, 28(4), 180-195.

[19] Lee, K. (2023). Ethical Considerations in AI-Powered Surveillance Systems. AI & Society, 38(2), 95-112.

Contact information:

Nikola GLIGORIJEVIC, M.Sc
Faculty of Information Technologies, Alfa BK University, Palmira Toljatija 3, 11000 Belgrade, Serbia
nikola.gligorijevic@alfa.edu.rs

Danilo STRUGAREVIĆ, M.Sc
The Academy of Applied Preschool Teaching and Health Studies, Balkanska 18
37000 Krusevac, Serbia
strugarevic@avmss.edu.rs
https://orcid.org/0009-0008-8602-7349

Vladimir ČABRIĆ, M.Sc
Faculty of Information Technologies, Alfa BK University, Palmira Toljatija 3, 11000 Belgrade, Serbia
vladimir.cabric@alfa.edu.rs