Smart Surveillance: Real-Time AI-Powered Threat Detection and Response System

Nukala Chaitanya[1], Manjula C[2], Neha Gayathri Devi P[3], Dheerajendra Reddy D[4], Sai Kumar K[5],

Dr. Pogula Sreedevi[6]

[1]Assistant Professor, Dept. of CSE, Rajeev Gandhi Memorial College of Engineering and Technology, Nandyala, Andhra Pradesh, India;

[2,3,4,5] Department of Computer Science and Engineering, Rajeev Gandhi Memorial College of Engineering and Technology, Nandyal- 518501, Andhra Pradesh, India;

[6] Associate Professor, Dept. of CSE, Rajeev Gandhi Memorial College of Engineering and Technology, Nandyala, Andhra Pradesh, India;

Email: [1] chaitu.anju@gmail.com , [2] manjulachavala347@gmail.com , [3] pnehagayathridevi@gmail.com , [4]dheerajendrareddy@gmail.com , [5] venkeysai74@gmail.com , [6] sreedevipogula37@gmail.com .

Abstract

The main objective of the project is to develop an intelligent intelligence analysis using deep learning methods such as neural networks (CNN) and YOLO v8 to improve real-time threat detection. Traditional surveillance systems rely on manual monitoring, which can lead to missing important events due to fatigue or data overload. The project automatically detects potential security threats such as guns, knives, and masked people directly at the security camera level. Existing systems often rely on cloud computing, which causes delays and limits immediate threat protection. The proposed system overcomes this problem by using an intelligent edge model to provide faster and more responsive time. The system also includes motion detection to capture images during abnormal activity, reducing the need for manual monitoring. The system employs YOLO v8 for detecting objects and utilizes CNN for classifying images. Future enhancements will be made to expand the application to cover various types of surveillance and further increase the accuracy of detecting various threats.

Index terms : CNN, YOLO V8, AI-powered surveillance system, real-time threat detection, Traditional surveillance systems

1. Introduction

Imagine: grainy CCTV footage, constantly rewinding, looking for something strange. Now imagine a smart camera analyzing the camera in real time to predict potential threats [1] before they happen. This is the transformative power of artificial intelligence (AI) in analytics. AI is more than just appearances; It focuses on comprehending, anticipating, and ultimately stopping crime. In recent years, there has been an increase in firearm-related violence, which has raised serious concerns about public safety [2]. According to statistics from various law enforcement agencies, there has been an increase in firearm-related incidents in the city, which has led to an urgent need for surveillance and monitoring systems. The emergence of artificial intelligence (AI) and machine learning has opened the way for new solutions to support security protection [3]. Among these, modern security equipment has become an important tool in preventing violence and ensuring the safety of public spaces. The ability to quickly and easily detect weapons can provide valuable information to law enforcement and security personnel that can help reduce threats before they escalate. Traditional security measures such as manual monitoring are often ineffective and do not respond quickly enough to emerging situations. In contrast, AI-powered systems can analyze large amounts of visual information and identify objects with high accuracy [4]. YOLO (Look Alone) is one of the most popular search tools today, known for its speed and accuracy [5]. YOLOv8 is the latest version of the series, which improves on previous versions and improves detection capabilities, especially in difficult areas.

The framework processes images once and is suitable for applications that require real-time analysis. Its architecture is designed to optimize speed and accuracy to provide quality images even in crowded or dynamic environments. The goal of this research is to develop a robust AI model that can detect guns and other weapons in various locations using YOLOv8 and help improve public safety measures. However, the structure of YOLO V8 used in research may vary depending on whether researchers integrate it into their own systems or projects, as shown in

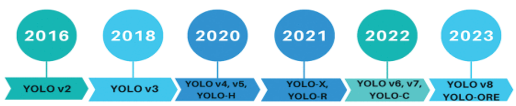

Figure 1. Fig.1.YOLOv8 technology revolutionizes real-time threat detection in video surveillance, providing unparalleled speed and accuracy. Many applications of YOLOv8 go beyond security, helping to create public safety and smarter urban planning.

2. Literature survey

2.1 Importance of Instant Weapon Detection
The success of instant weapon detection using YOLOv8 is critical for ensuring security in public spaces. In an era where crises can occur unexpectedly, the ability to detect threats swiftly is vital. Integrating advanced AI models into surveillance systems for schools, airports, shopping malls, and other high-risk locations enhances security measures and fosters public confidence. The mere presence of such systems acts as a deterrent to criminal activities. In schools, for instance, immediate alerts to security personnel or law enforcement upon detecting a weapon can significantly reduce response times and prevent casualties. Airports, particularly high-risk zones, can leverage real-time alerts to

initiate rapid evacuations or lockdowns, ensuring adherence to security protocols while minimizing panic.

2.2 Transforming Traditional Surveillance
This capability is redefining conventional surveillance by integrating AI-driven security measures. The technology seamlessly integrates with existing security infrastructures, including drone-based aerial surveillance for large-scale events, offering a multi-perspective approach to threat detection. Furthermore, its synergy with AI-powered technologies, such as facial recognition, enhances security management. Combining weapon detection with personal identification improves situational awareness for law enforcement, enabling immediate response to potential threats through multi-source intelligence.

2.3 Data-Driven Security and Policy Implications
By collecting and analyzing data on weapon detection incidents, law enforcement and policymakers can better identify crime patterns, assess security risks, and optimize resource allocation. This data-driven approach facilitates the development of evidence-based security policies aimed at crime prevention and public safety enhancement.

2.4 Need for AI Investment in Security
Investing in AI-driven technologies is imperative for advancing public safety measures. AI-powered security systems represent a forward-thinking approach to addressing modern security challenges, reinforcing the necessity for continued research and implementation in this field.

3.Proposed System

The proposed system will overcome this problem by using edge intelligence model to provide faster detection and response time. The system also includes motion detection to capture images during abnormal activity, reducing the need for manual monitoring. The system employs YOLO v8 for detecting objects and uses CNN for classifying images Future improvements will be made to extend the application to various types of surveillance and further increase the accuracy of detecting various threats. Machines that use images to represent objects now rely on data from a single frame. The camera in the focal area captures the image and tracks changes in the frame in detail. The You Look Alone (YOLO) 8 algorithm is the standard for detecting objects in images. It is now known for its accuracy in acquiring information and its speed of operation.

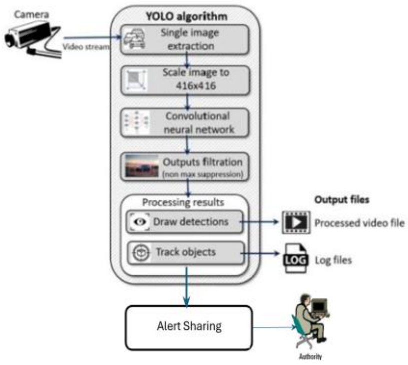

Fig.2. Proposed system architecture

This provides an overview of the "look at once" rule algorithm (YOLOv8) and its variants, comparing different YOLO versions and their performance relative to convolutional neural networks (CNNs). It highlights ongoing developments in the YOLOv8 algorithm. The paper details the development of intelligence models for real-time weapon detection, aimed at improving security in public spaces like schools, airports, and transportation systems. With global violence on the rise, the need for effective surveillance tools is critical. The model was trained on a large dataset of firearms and edged weapons, ensuring a robust learning process. Performance evaluation was done using key metrics such as precision, recall, F1 score, and mean accuracy (mAP), as well as Intersection-over-Union (IoU) thresholds. The results show that the YOLOv8-based weapon detection model performs effectively in distinguishing weapons from non-weapons with minimal errors and is capable of real-time processing. This research not only contributes to computer vision but also addresses the societal need for improved security measures, enhancing the ability of law enforcement and public safety organizations to respond to threats.

4. Working methodology

The YOLO algorithm flowchart consists of multiple key elements that work together to detect objects in images. It starts with a still image as input, which may undergo preprocessing like resizing or normalization depending on the YOLO version. The image is then processed through a convolutional neural network (CNN) backbone, such as Darknet or ResNet, which extracts hierarchical features at different scales. This allows the model to capture both low-level details and high-level semantic information.

The algorithm's head component applies convolutional techniques to estimate bounding boxes, confidence scores, and object quality based on the extracted feature maps. Bounding boxes are defined using coordinates (x, y, width, height) to locate potential objects. This technology represents a major step toward smarter, AI-driven security systems. By leveraging AI for public safety, the research emphasizes the need for innovation, collaboration among technology experts, and proactive legislation to protect communities from emerging threats.

4.1. Leveraging CNNs and YOLOv8 for Real-time Threat Detection

This system aims to utilize the power of Convolutional Neural Networks (CNNs) and the YOLOv8 object detection framework to automatically detect potential security threats like guns, knives, and masked individuals directly from security camera feeds in real-time.

High-resolution security camera: Captures high-quality images with sufficient resolution for accurate object detection.

4.2. YOLOv8 Object Detection

Model Training: Train the YOLOv8 model on large and diverse image and video data containing various threatening objects (guns, knives, masks, etc.). The YOLOv8 learning model can be deployed on an end device (e.g., a GPU-accelerated computer) or the cloud to provide real-time processing of video streams, identify threats using bounding boxes and class capabilities.

4.3.CNN-Based Threat Classification

Fine-tuning the CNN: Teach the CNN model to further classify detected objects into specific threat types (e.g., gun, rifle, knife, mask). This increases accuracy and reduces distortion. Alarm System:

Alarm System: When a threat is detected (e.g., gun, knife, masked person), the system will sound an alarm. These can include:

- **Visual alerts:** Display alert messages or highlight items on security camera feeds.
- **Data Collection:** Collect large and diverse image and video files that include:
- **Objects of Interest**: Guns, knives, other weapons, masked person, bad behavior (e.g., running)
- **Images:** Images taken from various locations (airports, public places, etc.) with lighting contrast, camera angle, and crowd level. This backbone consists of a set of convolutional algorithms, segmentation, and other techniques to achieve good results. For every object in the picture, estimate the bounding boxes and class probability.
- **Anchor boxes**: Predefined boxes with different sizes and aspect ratios that are used to estimate object locations
- **Object scores:** Estimate the probability that an object appears in an anchor box.
- **Probability**: Calculate the likelihood that an item will be present in each class (e.g. gun, knife, masked person, background). Training Loss function: A combination of loss functions is often used: o Intersection of box return loss: Determines how much the actual bounding box coordinates differ from the projected ones. (e.g. Intersection Over Unit (IoU) loss).
- **Objectivity Loss:** Calculates the discrepancy between actual and expected scores.
- **Optimizer:** Use an optimizer like Adam or SGD to adjust the model weights during training. Convergence. Inference and post-processing.
- **Inference:** Feed the input images to the YOLOv8 model to get the predictions. And it overlaps a lot. Instant Deployment:
- **Optimization**: Optimize the speed of the model (e.g. using techniques such as quantization, pruning, etc.). Action must be taken immediately. Be responsible and ethical, avoid injustice and abuse. And updated with new data and improved algorithms.

By leveraging the power of CNN’s and YOLOv8, combined with robust data collection and modeling, it is possible to produce real-time, high-quality results that can improve public safety and security. The exact details of system implementation and design will depend on many factors, including the specific needs, available resources, and the level of accuracy and performance required.

4.4. Data Collection and Preparation:

The input photos are divided into SXS grids by the YOLO V8 architecture. Each cell is responsible for estimating a fixed number of boxes defined by four coordinates (x, y, w, h) and a confidence interval C.

The confidence score is calculated using the equation:

$$C = P(\text{Object}) \times \text{IoU} \tag{1}$$

Where P(object) is the probability that the bounding box contains the object, and IoU (Intersection of Union) measures the overlap of the estimated bounding box and the box’s true location on the ground. This approach increases speed and efficiency by allowing YOLO to issue predictions for multiple objects in a single pass. Subsequent iterations of YOLO introduced many improvements to increase accuracy, speed, and robustness. YOLOv8 is still pushing the limits of real-time detection. YOLOv8 was designed by its leaders by incorporating modern technologies such as high-performance computing into the framework to improve performance and efficiency. The introduction of new training methods such as machine learning and self-monitoring makes YOLOv8 adaptable to a wide range of scenarios and datasets. Many products are used to optimize the performance of the model. The final loss function is expressed as:

Algorithm 1 Object Detection Algorithm
1.Input: Image I
2. Preprocess image I to fixed size W x H
3. Divide image into S x S grid cells
4. for each grid cell g do
5. for each bounding box b in cell g do
6. Predict coordinates (x, y, w, h) and confidence score C
7. Predict class probabilities P(Class|g)
8. end for
9. end for
10. Apply Non-Maximum Suppression (NMS) to filter overlapping boxes
11. Output: Detected objects with bounding boxes and class labels

The various transformations applied to the input image can be summarized as follows:

$$\text{Loss} = \lambda_{coord} \sum_i \sum_j \text{Loss}_{coord} + \lambda_{noobj} \sum_i \sum_j \text{Loss}_{noobj} + \sum_i \sum_j \text{Loss}_{class} \tag{2}$$

The loss function used throughout the study is important. has good performance standards. The regional loss can be expressed as a combination of confidence loss and distribution loss:

$$\text{Loss}=\lambda_{coord} + \text{Loss}_{coord} + \lambda_{noobj} \cdot \text{Loss}_{noobj} + \text{Loss}_{class} \tag{3}$$

Where:

$$\text{Loss}_{coord} = \sum_{i=0}^{N} \sum_{j=0}^{B} ((X_{ij} - \hat{X}_{ij})^2 + (Y_{ij} - \hat{Y}_{ij})^2 + (W_{ij} - \hat{W}_{ij})^2 + (H_{ij} - \hat{H}_{ij})^2) \tag{4}$$

Here, N is the number of samples, and B is the number of bounding boxes predicted per grid cell.

The λ_{cord} and λ_{noobj} hyperparameters control the weight of the localization loss and no-object loss, respectively. These parameters can be adjusted to emphasize different aspects of the training process.

4.5. Fast search using YOLOv8:

For immediate applications, fast detection of the model is essential. The average inference time per frame of different solutions shows how well the model performs with large inputs. The model developed from the YOLOv8 model demonstrates its ability to detect weapons in a variety of real-world scenarios.

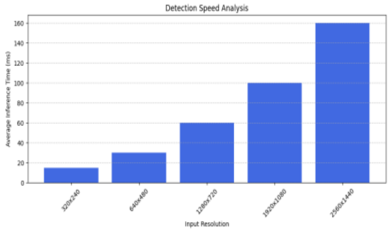

Fig.3. Detection of speed analysis

The image shows how the model creates boxes around the test objects and assigns labels to the class and confidence scores to indicate the accuracy of each test. The containers are color-coded according to confidence, allowing for quick assessment of the reliability of findings. For example, higher scores are typically represented by brighter colors, while lower scores may use darker tones. This visual feedback helps users quickly determine a model's performance in different situations, such as crowded spaces or changing lighting conditions. It's a significant advancement at the intersection of AI and public safety.

In the next step, the algorithm uses maximum limit (NMS) to eliminate inconsistent or overlapping detections. NMS ensures that only the most reliable, non-overlapping data is stored, thus eliminating the occurrence of the same product. The output of the YOLO algorithm is a series of bounding boxes, each associated with a confidence score and class probability. These checkboxes represent the objects detected in the input image, along with their class labels and confidence scores. It has become a technique for identifying objects in the field of computer vision. In the past, people used techniques such as sliding window, RCNN, Fast RCNN, and Faster RCNN for object detection. It outperforms all other algorithms. This version leverages the latest advances in deep learning and computer vision to deliver exceptional speed and accuracy. Its design meets a variety of application needs, and since it is implemented in the easy-to-use Ultralytics Python package, it can be adapted to a variety of hardware platforms, from edge products to cloud APIs. State-of-the-art (SOTA) object detection algorithms are so fast that they have become a standard way to identify objects in computer vision. In the past, sliding windows were the most effective method for target detection. Later improvements were made and faster versions of detection products were released such as CNN, R-CNN, Fast RCNN, etc. Some ideas. We will delve into the potential of YOLOv8 and understand its progress, how to use it seamlessly on a specific dataset, and try to understand the evolution of YOLO and the challenges and limitations of previous YOLO versions.

5. YOLOv8 technology

YOLOv8 technology has brought about significant changes in the way we monitor areas in real time. It is the eighth version of a series of product detection models that promise faster, more reliable and more accurate threat detection. This is not just a new change, but an overhaul that will have a major impact on global security surveillance. However, their use in criminal activities has also increased. This situation reveals the urgent need for advanced counter-drone systems. YOLOv8 overcomes this challenge by knowing the disadvantages of harmless drones. Combining high-tech technologies with machine tools such as TensorFlow and PyTorch, YOLOv8 is more than just an improvement. It is a complete revolution in the field of video surveillance. Intersection of Unity (IoU) is the key to verifying the truth. It shows how close YOLOv8's predictions are to the truth.

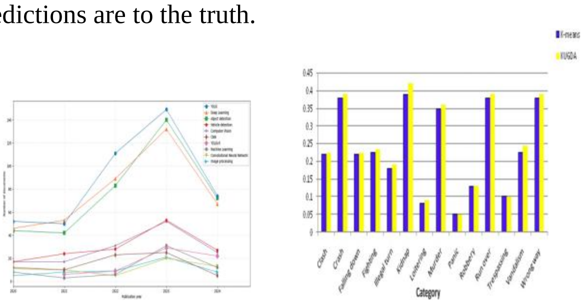

Fig.4. CNN, YOLOv8's predictions of truth and Yolov8, CNN multiple groups testing

Average Precision (AP) and Mean Average Precision (MAP) show the accuracy for different product types. This is very important for use in many cases. Precision and memory help to avoid errors when searching for the right product. This is important when both types of errors are serious. CNN, YOLOv8's model.val() function checks all of these parameters in multiple groups during testing..

5.1. Comparative Speed Analysis

YOLOv8 is known for its speed and accuracy. It works well using the COCO dataset which shows its accuracy and speed. This is crucial when quick decisions are necessary. Tools such as F1 Score Curves and Precision-Recall Curves illustrate its effectiveness.

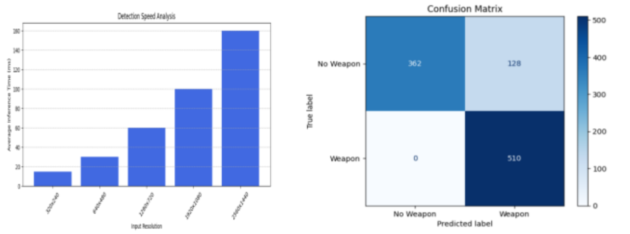

Fig.5. Detection of speed analysis and Detection matrix of weapon detection

These help the model to be better. It is highly reliable in the real world and improves performance at all scales. and tracking. Advanced cameras and sensors capture high-resolution images, which are then identified and tracked for movement of people, threats, and other objects of interest in the observation area. By quickly processing visual data and using intelligent tracking devices, these systems can track the contours of multiple targets simultaneously and provide information and instant alerts to children about safety risks or security issues. Each area is analyzed in detail, providing a wealth of information on current usage. Among the various YOLO architectures evaluated (v2-v8, H, X, R, C), YOLO v8 performs well with an average accuracy (mAP) of 0.99.

YOLOv8 is replacing video surveillance technology. It has revolutionized the way we use security cameras in the real world. Public spaces can now monitor and manage traffic better than ever before. It scores well in detecting when items are removed from the site. With YOLOv8, video forensics technology has moved beyond the old methods. Drones can now be effectively identified and tampered with. This is thanks to research that combines YOLOv8 with a camera to prevent accidents. This shows interest in using YOLOv8 to analyze drone data. It has also developed a number of methods to solve the challenge of discovering hidden objects in video. This development is important for the future of security analysis.

Table 1: Y**OLOv8 for security surveillance**

Capability	**Statistic**	**Real-world Impact**
Forgery Localization F1-Score	0.99	Enhances reliability in video surveillance
Obstacle Detection for UAVs	F1 score of 96% in 200 epochs	Improves navigational safety for drones
Generative models in occlusion	Effective recognition with occlusions	Advances object detection despite visual interferences

Classification is the simplest of the other tasks and involves dividing the entire image into one of a set of predefined categories. The image classifier generates a label with a confidence score. Important issues that need to be carefully planned and coordinated.

Privacy and security:

Privacy and security are important to address public concerns and ensure compliance with laws governing the use of personal information.

1.Training and supervision: Training is not required among employees and workers. Regular performance monitoring is essential to ensure performance and to keep up with changing threats and technological advances. urine.

2. Ethical issues: Considerations regarding the use of AI-based surveillance should balance public safety and individual privacy rights, and should be guided by clear instructions and stakeholder participation.

6. Benefits

Surveillance Video Analytics: Its Importance in Today's World: The main objectives that highlight the importance of this topic are listed below. It is difficult and frustrating for people to watch movies continuously. Intelligent surveillance video analysis is the solution for complex tasks. Wisdom must be found in every real situation. Knowledge of products and knowledge of operations required the highest accuracy. Functions such as crowd analysis still need a lot of development. The time required to generate a response in real life is very important. Predicting certain movements, actions, or actions can be helpful in emergencies such as emergencies. A wealth of information is available in video format. Some data uses a technique similar to binary classification to determine which behavior is different. There are various methods for criminal detection and crowd analysis. The next section discusses the designated application areas. A significant portion of existing work offers context-specific solutions.

- Traffic and major intersections
- Address
- Discussion board
- A festival as part of a religious organization
- Inside the office
- The crowd is the most difficult to identify in the situations listed. All actions, behaviors and movements must be recognized.

7.Conclusion

Our results from extensive testing demonstrate the model's effectiveness by striking a balance between accuracy and recall that is essential to minimizing bias and false positives. This is particularly true when the stakes are extremely high. Additional insights, including regression of accuracy and uncertainty matrices, provide a better understanding of the model's performance, highlighting its benefits across different weapons, not just one. Furthermore, the analysis of the detection speeds of different solutions demonstrates the effectiveness of this model, indicating its suitability for integration into the surveillance plan. With urban security issues becoming increasingly important, the implementation of such solutions can play an important role in measuring security. Future work will focus on further developing the model, expanding the data to cover different locations and weapon types, optimizing it to meet faster requirements, and exploring different ways to integrate audio and video data to improve discovery. Through these efforts, we focus on using technology to solve public safety issues and contribute to a safer society.

8.Future work

Some of the performance improvement of the threat detection system including but not limiting to further research:

- Ensemble Methods: The output of several models can probably be combined to improve the ability to detect objects. For example, an ensemble made of YOLOv8 and other lightweight model could provide optimal balance between speed and accuracy performance in challenging environments.

9. Acknowledgement

The authors sincerely acknowledge the students, staff, and authority of Rajeev Gandhi Memorial College of Engineering and Technology for their cooperation in research.

10.References

[1] Ahmed Abdelmoamen Ahmed ., Mathias Echi (2021). "Hawk-Eye: An AI-Powered Threat Detector for Intelligent Surveillance Cameras."
[2] Petrie, C., et al. (2005). "Firearm-related violence and public safety." *Journal of Public Safety*.
[3] Fowler, J., et al. (2015). "Artificial Intelligence in Security: New Solutions for Old Problems." *International Journal of Security Studies*.
[4] Shahnoor, M., et al. (2022). "AI-Powered Surveillance: Analysing Visual Information for Security." *Journal of Computer Vision and Image Processing*.
[5] Redmon, J. (2016). "You Only Look Once: Unified Real-Time Object Detection." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
[6] Deshpande, A., et al. (2023). "Data Collection and Training for Weapon Detection Models." *Journal of Machine Learning Research*.
[7] Dong, Y., et al. (2024). "Challenges in Real-Time Threat Detection: The Limitations of Cloud Computing." *International Journal of Security Technology*.
[8] Brodie, J., et al. (2005). "Data-Driven Approaches in Law Enforcement." *Journal of Criminal Justice Research*
[9] Manasa, M. & Sowmya, D. & Reddy, Y. & Sreedevi, Pogula. (2024). Attention-Based Image Caption Generation. 10.1007/978-981-97-8031-0_38.
[10] Sunil, N., & Narsimha, G. (2024). Image-based random rotation for preserving the data in data mining process. Signal Image and Video Processing, 18(4), 3893–3902. https://doi.org/10.1007/s11760-024-03050-2
[11] Lalitha, R. V. S., Jalligampala, D. L. S., Kavitha, K., Vahida, S., & Rajasekhar, G. (2021). New Directions in Traffic Control Analysis through Video Surveillance. E3S Web of Conferences, 309, 01099. https://doi.org/10.1051/e3sconf/202130901099
[12] Annapurna, B., Reddy, T. R., Raghavendran, C. V., Singh, R. K., & Prasad, V. V. (2020). Coordinate Access System for live video acquisition. Journal of Physics Conference Series, 1712(1), 012034. https://doi.org/10.1088/1742-6596/1712/1/012034