Romanian Journal of Petroleum & Gas Technology
VOL. IV (LXXV) • No. 2/2023

COMPARATIVE ANALYSIS OF OBJECT CLASSIFICATION ALGORITHMS: TRADITIONAL IMAGE PROCESSING VERSUS ARTIFICIAL INTELLIGENCE – BASED APPROACH

Cosmina Mihaela Roșca

Petroleum-Gas University of Ploiești, Romania
email: cosmina.rosca@upg-ploiesti.ro

DOI: 10.51865/JPGT.2023.02.17

ABSTRACT

In the current era of advanced digital technologies, form recognition is integrated into numerous applications, from computer vision to industrial automation. This paper focuses on a comparative analysis of two distinct form recognition algorithms, namely harnessing the power of artificial intelligence (AI) and image processing techniques. The research is motivated by the need to address the trade-off between speed and complexity in form recognition, with a center on real-world applicability.

Traditional image processing-based form recognition approaches often require complex coding, substantial domain expertise, and significant computational resources. This complexity can hinder rapid adaptation to changing requirements and the addition of new forms. The aim is to explore whether AI-powered algorithms can offer a more efficient and versatile alternative, reducing the barriers to entry for form recognition tasks.

The primary goal of the paper is to compare the performance of AI-based form recognition with image processing-based methods in terms of speed and accuracy. The second goal is to assess the ease of adapting AI-based algorithms to new forms without extensive code changes.

Two form recognition algorithms were designed and implemented, one based on artificial intelligence and a second relying on image processing. The AI-powered algorithm uses neural network architecture trained on a predefined dataset of forms. The image processing algorithm employs edge detection and contour analysis techniques.

Keywords: form recognition, artificial intelligence, image processing, machine learning

INTRODUCTION

In nowadays digital era, the field of object detection in images has witnessed a remarkable evolution, offering a myriad of innovative approaches to tackle this complex task. Among the most prominent methods, two distinct paradigms have emerged, each with its own set of strengths and applications, i.e., form recognition using image processing and artificial intelligence (AI) with form recognition and object detection.

Image processing for object detection relies on a rich toolbox of techniques designed to manipulate and analyze visual data. It encompasses a wide range of operations, from basic edge detection and color analysis to more advanced algorithms for feature extraction, pattern recognition, and object tracking. These techniques have proven invaluable in various industries, such as manufacturing, where precise quality control is essential, or in healthcare, where medical image analysis aids in diagnosis and treatment planning.

Secondly, AI-driven form recognition uses deep neural networks and advanced algorithms not only to identify objects but also to understand their context within an image. Through the training of models on vast datasets, AI systems can achieve remarkable accuracy in recognizing objects, even in complex and cluttered visual environments. This has opened up a world of possibilities, from autonomous vehicles that navigate busy city streets to innovative retail experiences that personalize shopping recommendations based on what customers are looking at.

While both paradigms have their unique merits, their synergy is increasingly shaping the future of object identification in images. Integrating the precision of image processing with the contextual understanding of AI promises to revolutionize industries such as autonomous robotics, surveillance, agriculture, and many more. As technology continues to advance, the boundaries of what can be achieved in object identification are continually expanding.

Digital image processing plays an important role in the analysis of remote sensing data. It involves four fundamental steps: image correction/restoration, image enhancement, image transformation, and image classification. Image restoration corrects errors, noise, and geometric distortions introduced during scanning and recording. Image enhancement improves visual impact and information content, aiding computer-based recognition and classification. Information extraction relies on comparisons between images acquired at different times. Unsupervised classification groups reflectance patterns into predefined classes without prior knowledge, while supervised classification uses user guidance to classify features based on spectral characteristics. Change detection tracks seasonal variations and post-classification analysis involves image smoothing and accuracy assessment to yield valuable output in digital image processing [1].

Numerous researches habe explored algorithms that can be implemented in industry to achieve 100% accuracy in object identification. A notable issue encountered in this AI domain is the gradient contribution imbalance during the training process. Paper [2] delves into the root cause of this imbalance, attributing it to variations in example attributes, such as difficulty and shape variation degree. To correct this issue, the paper introduces a novel approach called example attribute-based prediction modulation (EAPM). EAPM operates by defining the attribute of an example based on its prediction and corresponding ground truth. Empirical evidence from experiments conducted on the experimental dataset underscores the substantial improvements that EAPM brings to deep object detectors.

Image processing is applied in multiple research areas using different algorithms, but having the same image methods for object identification. The healthcare field is among the most prominent domains for object identification in images due to its widespread applications. Reference [3] compiles experimental findings and analyses conducted on medical images about diabetic-related issues. The research spans from fundamental image processing techniques, such as image enhancement, to advanced image segmentation methods. Its primary purpose is to raise awareness about diabetes and its associated factors, presenting image-processing approaches for detection and prediction straightforwardly. This resource proves valuable for researchers, engineers, medical professionals, and bioinformatics experts.

In the realm of computer vision, object detection is a critical challenge, encompassing both object classification and precise localization within images or video frames. Identifying objects in images faces three major challenges:

- The ability to recognize objects in noisy images;
- The speed at which identification can be performed;
- Limited memory resources placed on the system running the algorithm within the application.

Article [4] introduces the pervasive issue of image noise, proposing a low-cost technique for detecting objects in noisy images. Utilizing the Single Shot MultiBox Detector (SSD) and retraining conventional detectors on noisy images, the paper achieves satisfactory performance, particularly beneficial for surveillance applications with challenging lighting conditions.

Object identification in videos is the most challenging task in image processing. Paper [5] proposes a robust video object segmentation algorithm that leverages wavelet transforms and moving edge detection. Through a sequence of operations, including wavelet domain change detection and edge mapping, it effectively identifies and tracks moving objects in videos. The method's resilience to object motion and local deformations is substantiated through experiments and object evaluation. Research [6] explores the use of Light Detection and Ranging (LiDAR) and streaming video to enhance situational awareness and detection sensitivity in urban environments, particularly for radiological and nuclear material detection. The study demonstrates real-time object detection, tracking, and source-object attribution analysis, showcasing the fusion of radiological data with tracking information.

In the article [7], the focus is on anchor-based object detection models with sparsely annotated datasets. The paper introduces innovative solutions, namely the anchor-less object detector and the single-object tracker for semi-supervised learning-based object detection. The approach leverages bidirectional single-object tracking to obtain dense annotations, leading to remarkable performance in object detection challenges.

Research [8] comprehensively explores various Object Detection Algorithms, including face detection, skin detection, color detection, shape detection, and target detection. These algorithms are implemented in MATLAB 2017b for video surveillance applications, offering improved accuracy. The paper also discusses the challenges and diverse applications of Object Detection methods across fields like security and healthcare.

Paper [9] addresses the scarcity of datasets for foreign object intrusion detection on railway tracks. It conducts a comparative study between YOLOv5 and Fast RCNN in this context. Utilizing UAV aerial images and edge detection techniques, the study establishes the superiority of YOLOv5 in terms of detection rate and speed for real-time tracking foreign object intrusion detection. Paper [10] delves into the challenge of deploying object detection networks on embedded devices with limited computational and memory resources. It introduces Tiny SSD, a highly optimized deep convolutional neural network designed for real-time embedded object detection. With a significantly reduced model size, Tiny SSD maintains impressive object detection performance, making it a viable choice for embedded scenarios.

Comprehensive insights into the domain of digital image processing, covering various aspects of this multidisciplinary field are presented in papers [11 - 13]. Paper [11] presents the significance and applications of digital image processing, emphasizing its exponential growth across diverse sectors. The paper delves into the fundamental principles of image sampling, quantization, enhancement, restoration, compression, and analysis. It underscores the importance of efficient data representation and reduction to handle the immense information contained in digital images. Article [12] focuses on image processing within mechanical and electrical engineering contexts. It distinguishes between analog and digital image processing, detailing the phases of preprocessing, augmentation, display, and information extraction. The paper highlights the application of machine learning and deep learning in detecting and measuring defects on metal surfaces, particularly in the context of pantograph inspection for electric rail cars. Research [13] explores the manipulation and analysis of images by computers, distinguishing between image processing and image analysis. It covers various aspects of image processing, such as digitization, image models, compression, enhancement, restoration, reconstruction, matching, segmentation, and description. The paper also touches upon emerging architectures and technologies for efficient image processing.

The extensions of image processing concepts, particularly in the context of error diffusion neural networks are presented in this paper [14]. It explores applications like color halftoning and discusses how different filter designs with varying spectral responses can lead to features such as edge enhancement, offering potential applications in feature extraction and automatic target recognition.

Paper [15] examines the landscape of algorithm development environments, function libraries, source code repositories, and specialized data processing packages for image and video processing. It delineates the suitability of these tools for different tasks, emphasizing the importance of choosing the right tool based on the specific requirements of the task at hand.

These papers contribute in developing of the object detection domanin, addressing key challenges, and offering practical solutions applicable to a wide range of fields and scenarios. Their findings collectively enhance our understanding and capabilities in this critical area of computer vision.

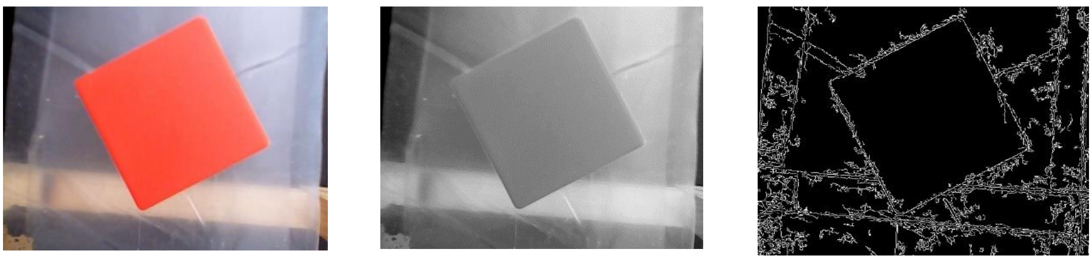
CLASSIC IMAGE PROCESSING ALGORITHM

In this section, I will present two algorithms designed to perform object identification and classify objects into specific categories: squares or triangles. Subsequently, these two algorithms will undergo testing to assess their efficiency. The implementation of the two algorithms was conducted using the Visual Studio development environment and the C# programming language. To facilitate a meaningful comparison between the two algorithms, the background remained unaltered, with only the objects to be identified being modified in the test images. In other words, the position of the object or the object itself was modified within the images. It is important to note that the algorithm does not consider the color or position of the object in the image but only focuses on the object's shape.

The classic algorithm that uses image processing is based on approximating the areas and the number of edges of the object for identification purposes. Specifically, for the square class, it calculates four sides, and for the triangle class, it calculates three sides. Figure 1 illustrates an example of an object to be identified. In this example, a triangle is presented. In order to perform object detection within an image, the algorithm undergoes the following stages:

- **Grayscale Conversion**: The image is converted to grayscale, simplifying the subsequent contour detection;
- **Edge Detection**: This step involves detecting edges within the image using the Canny edge detector. This process highlights regions of significant contrast within the image;
- **Contour Extraction**: Following edge detection, the algorithm extracts contours from the image. Contours represent the boundaries of objects in the image;
- **Contour Approximation**: The algorithm approximates the contours to simplify their representation. This step reduces the complexity of the contour while retaining its essential shape characteristics;
- **Identification of Shapes**: Each approximated contour is examined to determine the number of sides or edges it possesses. This count is critical for distinguishing between square and triangular shapes. For example, if the contour has three sides, it is identified as a triangle, while a count of four sides signifies a square;
- **Object Classification**: Based on the identified number of sides, the algorithm classifies the object as either a square or a triangle.

Figure 1a exemplifies this process by showcasing an object, in this case, a square, which serves as a representation of the object identification process. This algorithm systematically evaluates the edges, contours, and number of sides in order to accurately classify objects within an image. In Figure 1a, it is presented the original image, which has been acquired using a Webcam positioned on a conveyor belt. It is important to emphasize that all images subjected to these rigorous testing procedures were acquired under standardized conditions, employing the identical webcam system. Figure 1b presents the same image from Figure 1a transformed into a grayscale representation. Finally, Figure 1c provides an illustration of the image following the identification and accentuation of contours that are pivotal to the analytical objectives.



(a) (b) (c)

Figure 1. *Object example: a) original image; b) grayscale image; c) contours image*

For the image depicted in Figure 1, the algorithm successfully identified the object as a square. Nevertheless, upon closer examination of Figure 1c, several limitations might be identified:

- The marked contours in the image may potentially describe other geometric elements, such as triangles or additional squares within the same image;
- In an effort to address the previous limitation, area thresholds have been empirically established. However, these thresholds do not guarantee the accurate identification of objects within the image;
- In scenarios where the object is situated on a moving conveyor belt, and the Webcam captures only a partial view of the object, the identification process becomes inaccurate. Consequently, the algorithm's applicability is restricted to instances where the object is fully captured within the frame.

The example presented in Figure 1 represents an ideal case where the Webcam has captured the entire object. Figure 2 illustrates all 15 test images used to assess the accuracy of the classical image processing algorithm.

For the square class, the classical image processing algorithm was assessed using the 15 images presented in Figure 2. Following the algorithm's execution, only two images were misclassified as containing a triangle class object. Figure 3 illustrates the outcome of the algorithm's execution, wherein it can be observed that images "i.jpg" and "k.jpg" from Figure 2 were incorrectly classified.

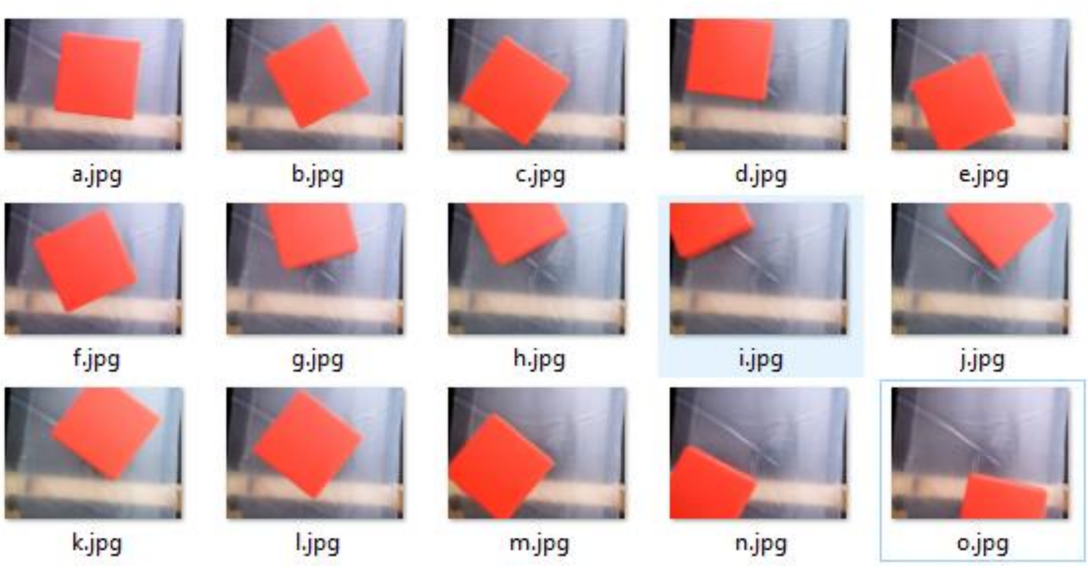


Figure 2. *Test images for the square class*

Although some images contain objects with more than 4 sides (for instance "n.jpg"), they were classified as belonging to the square class due to the area condition being met. In this way, the algorithm's accuracy was significantly improved.

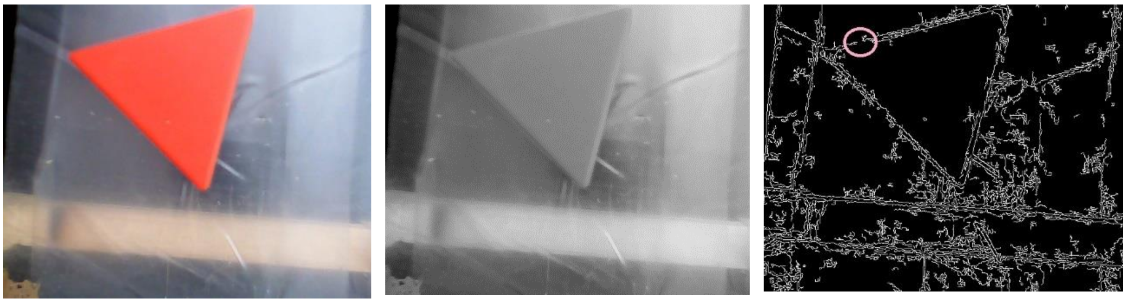
```
C:\Windows\system32\cmd.exe
Image  a.jpg - Square
Image  b.jpg - Square
Image  c.jpg - Square
Image  d.jpg - Square
Image  e.jpg - Square
Image  f.jpg - Square
Image  g.jpg - Square
Image  h.jpg - Square
Image  i.jpg - Triangle
Image  j.jpg - Square
Image  k.jpg - Triangle
Image  l.jpg - Square
Image  m.jpg - Square
Image  n.jpg - Square
Image  o.jpg - Square
```

Figure 3. *Classic image processing algorithm execution for square class tests*

Next, I will present the analysis conducted for the triangle class. For this test, 43 images were used, with the triangular object being marked differently in the experiments. Figure 4 displays the test images. The execution results indicate that only 2 images were correctly classified, namely "a.jpg" and "zz7.jpg".

Figure 4. *Test images for the triangle class*

The explanation for not classifying correctly the other images is that the transformation to grayscale destroyed the image's contour, as shown in Figure 5b. In the contour marking step, the triangle object could no longer be identified, as seen in Figure 5c. In this figure, the pink lens marks the interruption of the contour, which will prevent its identification as belonging to the triangle class. These situations can be prevented by adjusting the parameters for grayscale conversion, but changing them will affect the identification of the square class.



(a) (b) (c)

Figure 5. *Object example: a) original image; b) grayscale image; c) contours image*

As a result, it is obvious that the proposed implementation of the classical image processing algorithm is not efficient for real-time object classification. Empirical adjustments to grayscale transformation parameters or object area evaluation are time-consuming, and altering the background or adding a new object to the classification would require substantial code modifications.

In light of these considerations, the following section will analyze the proposal for integrating a specific artificial intelligence algorithm using the form recognition method within the machine learning component.

ARTIFICIAL INTELLIGENCE ALGORITHM

The implementation of the artificial intelligence algorithm is based on the machine learning component. This component focuses on the development of algorithms enabling computers to learn from data and make predictions or decisions based on different features identified. It involves the automatic improvement of computer systems through experience rather than explicit programming. The algorithm was developed using Visual Studio environment with C# programming language and ML.NET tool. This tool is an open-source, cross-platform machine learning framework developed by Microsoft. It is designed to provide developers with an easy way to integrate machine learning models into their applications. ML.NET supports a wide range of machine learning tasks, including object detection, classification, regression, clustering, recommendation, and anomaly detection.

The Image Classification component in ML.NET is a machine learning module designed for image analysis tasks. It enables the development of models for classifying images into predefined categories or labels. This component is particularly useful in scenarios where you want to automatically assign labels or tags to images based on their content. The proposed algorithm uses the Image Classification component of the ML.NET tool.

The development of the application involves the following stages:

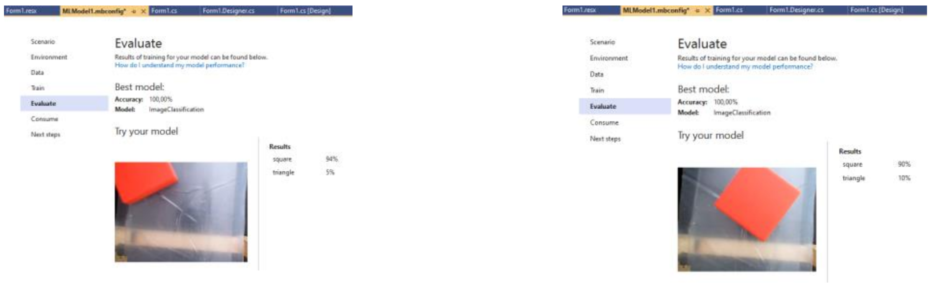
- **Model Construction Based on Test Images**: For each category, there were 200 training images, distinct from the images used for testing;
- **Model Training**: Model training is a crucial step in developing an image classification application. It involves using a labeled dataset comprising images and their corresponding labels. ML.NET uses this data to adjust the model's parameters, enabling it to accurately predict labels for new, unseen images. During this phase, the AI component identifies common features among images

belonging to the same category while distinguishing them from features common to other categories;

- **Model Evaluation**: Following training, it is essential to evaluate the model's performance using a separate validation dataset. ML.NET offers various metrics such as accuracy, precision, recall, and F1-score to assess how effectively the model classifies images;
- **Inference integration**: Once the model is trained, it is ready for inference, which entails predicting labels for new, unlabeled images. The model takes an image as input and provides predicted labels along with associated confidence scores. ML.NET's Image Classification component seamlessly integrates with the broader ML.NET ecosystem, facilitating the incorporation of image classification capabilities into your .NET applications.

Next, all the images used for the classical image processing algorithm were tested. The algorithm successfully classified all the images correctly. In Figure 6, two tests that were misclassified by the classical algorithm are presented. Figures 6a and 6b correspond to the images "i.jpg" and "k.jpg" associated with the square category. For these images, the confidence level is 94% and 90%, respectively.

Figure 8 presents two examples associated with the triangle category. None of these images were correctly classified by the classical algorithm. In Figure 8a, the image "p.jpg" contains a triangle object of the same color as the square objects, which could potentially confuse the algorithm. However, in Figure 8a, a triangle was identified with a confidence level of 99%. In Figure 8b, the image "zz1.jpg" contains a partial triangle object. In this case as well, successful identification was achieved, with a confidence coefficient of 100% due to the recognition of the three sides.



(a) (b)

Figure 7. *Square object identified with: a) 94% confidence level; b) 90% confidence level*

(a) (b)

Figure 8. *Triangle object identified with: a) 99% confidence level; b) 100% confidence level*

The integration of the image classification component into the algorithm has yielded significant improvements in object recognition. This approach uses machine learning and deep learning techniques to enhance the accuracy of object identification. Unlike the classical algorithm, which relied on heuristics and predefined rules, the machine learning-based approach learns from the two datasets of labeled images.

The algorithms underwent rigorous testing within a real-time sorting conveyor system, effectively replicating an industrial sorting environment. This dynamic setup was the basis of assessing their performance under conditions closely mirroring actual operational scenarios. As pieces flowed along the sorting conveyor, the algorithms swiftly and accurately recognized and classified them based on their shapes. The testing prototype provided valuable information about their speed and accuracy. The results confirmed their practical viability in an automated and time-sensitive context. The results obtained from the real-time tests delivered crucial insights into the algorithms' effectiveness and robustness in sorting applications.

This study's originality relies on its comprehensive comparison of two distinct form recognition approaches, highlighting the efficiency and adaptability of AI-based methods. Additionally, the research underscores the ease of knowledge base expansion in the AI algorithm, enabling rapid integration of new forms without the need for extensive code changes. This unique combination of speed, accuracy, and adaptability highlights the advantages of using AI algorithms in real-time applications.

The research showcases the clear advantages of leveraging artificial intelligence for form recognition tasks. The AI-powered algorithm outperforms its image-processing counterpart in terms of speed and accuracy and offers a user-friendly knowledge base expansion mechanism. This ease of adaptability positions AI-powered form recognition as a superior choice for applications where speed, precision, and flexibility are paramount. As we move towards an era of increased automation and evolving form recognition requirements, AI emerges as a transformative force in simplifying complex tasks.

As a result, the algorithm can adapt to various lighting conditions and object orientations. This adaptability has proven crucial in real-world scenarios where objects may appear differently due to changing environmental factors. The model's ability to generalize from its training data allows it to correctly classify objects even when presented with challenging variations.

CONCLUSIONS

In this paper, two algorithms were introduced, in particular an algorithm that uses classical image processing techniques and another algorithm that leverages specific artificial intelligence methods. Both algorithms aimed to classify an object identified within an image. The two algorithms were tested using the same set of images.

The algorithm designed with AI image classification demonstrates superior performance due to its machine learning-based approach. By training on a diverse dataset, it can adapt to a wide range of object variations, lighting conditions, and backgrounds. This adaptability is a significant advantage over the classical algorithm, which relies on predefined rules and heuristics.

The main advantage of the AI image classification-based algorithm is its ability to generalize, providing reliable results in real-world scenarios where objects may appear differently from those encountered during training.

However, it is crucial to acknowledge some of the limitations of the image classification-based algorithm. One notable drawback is the "black box" nature of machine learning models. While the algorithm excels at object recognition, the programmer may not have a deep understanding of the specific features or criteria driving its decisions. This lack of interpretability can be a challenge when fine-tuning or debugging the algorithm is required.

Additionally, there are scenarios in which the algorithm may not correctly identify an object. It is important to recognize that machine learning models operate based on patterns and probabilities, and there will always be cases that fall outside their training data distribution. As a result, when integrating this algorithm into production systems, a human operator should be on hand to ensure the quality of object identification and intervene when necessary.

Nonetheless, with proper monitoring and validation, the AI image classification algorithm can be successfully integrated into real-time systems. Its flexibility, adaptability, and enhanced accuracy make it a valuable tool for automating object recognition tasks, ultimately improving efficiency and reliability across a range of applications.

In conclusion, the algorithm using image classification represents a significant advancement in object recognition, surpassing the capabilities of the classical image processing approach. While it may introduce challenges related to interpretability and occasional misclassifications, its overall performance and adaptability make it a powerful tool for modern applications. When integrated into production systems, it can operate effectively with human oversight, ensuring both accuracy and efficiency in object identification tasks. The presentation of the automated system utilizing the AI algorithm for sorting parts on a conveyor belt will be conducted as part of future work.

REFERENCES

[1] Reddy G.P.O., Digital Image Processing: Principles and Applications. Geospatial Technologies in Land Resources Mapping, Monitoring and Management, Springer, 2018. doi: 10.1007/978-3-319-78711-4_6

[2] Wu Z., Liu C., Huang C., Wen J. Xu Y., Deep Object Detection with Example Attribute-Based Prediction Modulation, ICASSP 2022 – 202,2 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022. doi: 10.1109/ICASSP43922.2022.9746194

[3] Kumar A., Shaik F., Image Processing in Diabetic Related Causes, Springer Nature, 2016.

[4] Nayan A.-A., Saha J., Mahmud K.R., Al Azad A.K., Kibria M.G., Detection of Objects from Noisy Images, 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), IEEE, 2020. doi: 10.1109/STI50764.2020.9350521

[5] Zhang X.-Y., Zhao R.-C., Automatic Video Object Segmentation using Wavelet Transform and Moving Edge Detection, 2006 International Conference on Machine Learning and Cybernetics, IEEE, 2006. doi: 10.1109/ICMLC.2006.258748

[6] Marshall M.R., Hellfeld D., Joshi T.H.Y., Salathe M., Bandstra M.S., Bilton K.J., Vetter K., 3-D Object Tracking in Panoramic Video and LiDAR for Radiological Source–Object Attribution and Improved Source Detection. IEEE Trans. Nucl. Sci., 68(2), 189–202, 2020. doi: 10.1109/TNS.2020.3047646

[7] Yoon J., Hong S., Choi, M.-K., Semi-Supervised Object Detection With Sparsely Annotated Dataset, 2021 IEEE International Conference on Image Processing (ICIP), IEEE, 2021. doi: 10.1109/ICIP42928.2021.9506790

[8] Raghunandan A., Mohana, Raghav P., Aradhya H.V.R., Object Detection Algorithms for Video Surveillance Applications, 2018 International Conference on Communication and Signal Processing (ICCSP), IEEE, 2018. doi: 10.1109/ICCSP.2018.8524461

[9] Cai X., Ding X., A comparative study of machine vision-based rail foreign object intrusion detection models, 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), IEEE, 2023. doi: 10.1109/ICPECA56706.2023.10075769

[10] Womg A., Shafiee M.J., Li F., Chwyl B., Tiny SSD: A Tiny Single-Shot Detection Deep Convolutional Neural Network for Real-Time Embedded Object Detection, 15th Conference on Computer and Robot Vision (CRV), IEEE, 2018. doi: 10.1109/CRV.2018.00023

[11] Da Silva E.A.B., Mendonça G.V., Digital Image Processing, The Electrical Engineering Handbook. Academic Press, 2005. doi: 10.1016/B978-012170960-0/50064-5

[12] Arora V., Yin-Kwee Ng E., Singh A., Machine learning and its applications. Smart Electrical and Mechanical Systems, Academic Press, 2022. doi: 10.1016/B978-0-323-90789-7.00002-6

[13] Chellappa R., Rosenfeld A., Image Processing. Encyclopedia of Physical Science and Technology (Third Edition), Academic Press, 2003, doi: 10.1016/B0-12-227410-5/00841-3

[14] Shoop B.L., Information processing | Optical Digital Image Processing. Encyclopedia of Modern Optics, Elsevier, 2005. doi: 10.1016/B0-12-369395-0/00811-3

[15] Slatton K.C., Evans B.L., Software for Image and Video Processing. Handbook of Image and Video Processing (Second Edition), Academic Press, 2005. doi: 10.1016/B978-012119792-6/50100-5

Received: November 2023; Accepted: November 2023; Published: November 2023