Jurnal Teknik Informatika (JUTIF)
Vol. 6, No. 1, February 2025, pp. 401-410
DOI: https://doi.org/10.52436/1.jutif.2025.6.1.4257
p-ISSN: 2723-3863
e-ISSN: 2723-3871

ENHANCE OBJECT TRACKING ON AUGMENTED REALITY USING HYBRID CONVOLUTIONAL NEURAL NETWORK AND FAST CORNER DETECTION

Nurhadi*[1], Eko Arip Winanto*[2], Saparudin[3]

[1]Department of Informatics, Universitas Dinamika Bangsa, Jambi, Indonesia
[2]Department of Computer Engineering, Universitas Dinamika Bangsa, Jambi, Indonesia
[2]Computer Science, Universiti Teknologi Malaysia, Johor, Malaysia
[3]Department of Informatics Engineering, Telkom University, Indonesia
Email: [1]nurhadi@unama.ac.id, [2]winanto@unama.ac.id, [3]saparudin@telkomuniversity.ac.id

(Article received: December 26, 2024; Revision: February 3, 2025; published: February 19, 2025)

Abstract

Markerless augmented reality (AR) is utilized in applications that do not require anchoring to the real world and do not require the use of physical markers (fiducial markers). Augmented object displays not only float but also allow for the automatic placement of 3D augmented reality objects on flat surfaces to enhance realism in real time. There are two challenges that need to be addressed in Markerless AR systems: object tracking and registration, as well as the influence of light intensity. Therefore, the objective of this research is to propose the use of Convolutional Neural Networks (CNN) and Features from Accelerated Segment Test (FAST) corner detection for tracking or detecting objects in markerless augmented reality systems. Testing was conducted using three epoch schemes: 10, 50, and 100. The test results were measured using several parameters, including the execution time, testing loss, and testing accuracy. The test results indicated an improvement in the performance of the tested object detection. The accuracy testing results of using the CNN and FAST corner detection methods were superior to those of the CNN-only method and FAST corner detection alone, reaching 98%. However, this method increases the processing time for object detection. Thus, the processing time of the CNN without FAST corner detection was faster.

Keywords: *AR, Object tracking, CNN, Fast corner detection, Markerless*

PENINGKATAN PELACAKAN OBJEK PADA AUGMENTED REALITY MENGGUNAKAN HYBRID CONVOLUTIONAL NEURAL NETWORK DAN DETEKSI SUDUT FAST

Abstrak

Augmented Reality (AR) tanpa marker (markerless) digunakan dalam aplikasi yang tidak memerlukan pengikatan pada dunia nyata dan tidak membutuhkan penggunaan marker fisik (fiducial marker). Objek augmentasi tidak hanya dapat ditampilkan secara mengambang, tetapi juga memungkinkan penempatan otomatis objek augmented reality 3D pada permukaan datar untuk meningkatkan realisme secara waktu nyata. Terdapat dua tantangan utama yang perlu diatasi dalam sistem AR tanpa marker, yaitu pelacakan dan registrasi objek, serta pengaruh intensitas cahaya. Oleh karena itu, tujuan penelitian ini adalah untuk mengusulkan penggunaan Convolutional Neural Networks (CNN) dan deteksi sudut dengan Features from Accelerated Segment Test (FAST) untuk melacak atau mendeteksi objek dalam sistem augmented reality tanpa marker. Pengujian dilakukan dengan menggunakan tiga skema epoch, yaitu 10, 50, dan 100. Hasil pengujian diukur menggunakan beberapa parameter, termasuk waktu eksekusi, testing loss, dan akurasi pengujian. Hasil pengujian menunjukkan peningkatan kinerja pada deteksi objek yang diuji. Hasil pengujian akurasi menggunakan metode CNN dan deteksi sudut FAST lebih unggul dibandingkan metode yang hanya menggunakan CNN atau hanya deteksi sudut FAST, dengan mencapai akurasi hingga 98%. Namun, metode ini meningkatkan waktu pemrosesan untuk deteksi objek. Dengan demikian, waktu pemrosesan pada CNN tanpa deteksi sudut FAST lebih cepat.

Kata kunci: AR, Object tracking, CNN, Fast corner detection, Markerless

1. INTRODUCTION

In the development of augmented reality (AR), there are two augmentation techniques: marker-based and markerless [1]. In the context of AR, marker-based refers to what is commonly known as a fiducial markers [2]. Fiducial markers are anchor points that can be placed in a scene to provide a fixed reference point for a position or scale [3][4]. On the other hand, markerless is used in applications that do not require anchoring to the real world or physical markers (fiducial markers). The display of augmented objects not only floats but also allows for the automatic placement of 3D augmented reality objects on flat surfaces to enhance realism in real-time [5].

Regarding feature extraction in AR augmentation, feature extraction in marker-based systems does not represent real-world objects. Marker-based systems require a reference in the form of 2D (flat) barcodes [6]. By contrast, markerless uses real 3D objects for feature extraction in 3-DoF or 6-DoF, contributing to realistic augmentation. This makes markerless techniques superior to marker-based techniques because they do not require physical markers for augmentation, are flexible, can be implemented in various fields, and pose the challenge of proposing reliable tracking methods.

There are two main challenges that need to be addressed in Markerless AR systems: object tracking and registration and the influence of light intensity (accuracy) [8]. Previous research [7][9] has proposed markerless techniques for AR, indicating that marker accuracy is lower when augmenting objects influenced by light intensity. This significantly affects the success of the object augmentation. The second challenge is object tracking and registration, which are the essential components of AR systems. Object tracking refers to the process of feature extraction from real objects registered in the AR system as a 3D model [10][11]. Thus, the object registration process or the addition of 3D models to Markerless AR systems becomes more flexible and straightforward. Therefore, this study focuses on addressing the issues of object tracking, registration, and light intensity.

From the methods proposed in previous research [12][13], one method shows potential in handling Markerless AR issues: FAST Corner Detection. The features from the accelerated segment test (FAST) Corner Detection method proposed by [14] claim and prove its superiority in high repeatability under significant aspect changes and different feature types. This makes the method much faster than existing corner detectors (Haris, DoG, SUSAN) [15]. However, it is less effective for detecting objects in environments with many corners or points. There is a current trend in using deep learning to recognize or detect objects deeply in an image [16].

One deep learning method is the Convolutional Neural Network (CNN), which is known for its excellent image classification performance, such as visual object detection and human pose estimation [17],[18],[19]. CNN can be used for sequential data analysis such as voice recognition and search engines. Although physical objects can be detected using deep-learning methods, it is essential to provide an easily understandable form for visualizing information [20],[21],[22]. Thus, deep-learning-based AR enables the augmentation of 3D virtual objects and supports more effective interactions without AR markers. However, it is challenging to detect object information, such as the class, position, and pose of real objects, by scanning the physical object surface using 3D or by identifying key features.

In this study, we propose the use of CNN and FAST Corner Detection for object detection based on deep learning and instant segmentation combined with AR technology to achieve better performance in complex cases. Deep learning can assist users in effectively detecting objects and providing accurate information, considering dynamically changing environments and real-time user situations [23],[24],[25]. Specifically, CNN and FAST Corner Detection are used to efficiently detect physical object instances in object tracking and registration. Additionally, this research makes several contributions, including i) proposing a design architecture of CNN and FAST CD for AR markerless and ii) proposing a method for object detection using CNN and FAST CD for object tracking and registration in AR. This paper is organized into four sections: Section 1 is the introduction; Section 2 provides a brief discussion of the experimental setup; Section 3 provides a detailed description of the experiment and findings; and Section 4 presents the conclusions and suggests potential for future research.

2. METHOD

This section provides a comprehensive account of the experimental intricacies, architectural framework of the proposed methodology, and experimental setup. This study advances a deep learning approach that specifically employs a Convolutional Neural Network (CNN) and FAST Corner Detection for the nuanced tasks of detection and segmentation within augmented reality (AR) systems. The experimental design encompasses various phases, notably the development of the augmented reality interface, implementation of deep learning methodologies, and subsequent visualization of outcomes

2.1. COCO Dataset

In this study, the COCO dataset was employed for the deep-learning training process, specifically for object detection and segmentation in real-world environments. The COCO dataset comprises a collection of images for both training and testing purposes, supplemented with annotated data indicating the outcomes of detected objects. This dataset has been utilized by several researchers in previous studies. For the object detection process in this study, Arduino and other devices were utilized. The COCO dataset from 2017, as detailed in Table 1, consisted of 78,458 images; however, for the training phase, a subset of 42,000 images containing classes of people and objects was utilized for pre-training and testing. Subsequently, in the experimentation phase, Arduino devices were incorporated to assess the integration of physical object data into augmented reality. Examples of images from the COCO dataset and physical objects are as follows:

Table 1. COCO Dataset

2.2. Experiment setup

The present study has a considerable emphasis on the experimental phase for markerless AR using deep learning. The experiment encompasses a comprehensive examination of various aspects, such as the design architecture of AR markerless and object detection techniques, datasets for training and testing purposes, and the configuration of the CNN and FAST CD method and its associated variables. In general, there are three stages of experimentation:

- Proposed architectural design for object detection in AR systems using CNN and FAST corner detection.
- Subsequently, a detailed design of the proposed method is described. This research suggests a deep learning method, namely a Convolutional Neural Network (CNN) and FAST corner detection, for the processes of object detection and segmentation in augmented reality systems.
- Finally, we measured the performance of the proposed system, including the accuracy of the CNN and FAST CD methods.

2.3. Design Architecture CNN and FAST CD

The main objective of this study was to design the architecture of a markerless AR system using a hybrid of CNN and FAST corner detection. The primary distinction in the architecture employing CNN and FAST corner detection, compared to the AR architecture without deep learning methods, lies in the addition of deep learning and FAST corner detection modules and a module comparing the results of CNN and FAST corner detection to enhance detection outcomes. For example, in the object detection process, if the CNN successfully detects four objects in an image, and subsequently, the FAST corner detection detects five objects, the difference of one object is incorporated into the new detection result, thereby improving accuracy. This principle also applies to the reverse processes.

Figure 1 depicts the architecture of the markerless AR system using CNN and FAST corner detection proposed in this study. Generally, there are four main stages in the experimental setup, as illustrated in Figure 1: (1) AR interface, (2) deep learning module, and (3) visual augmentation module. This can be described as follows.

- AR Interface Module: The AR interface module allows AR devices to capture images for physical object recognition and send them to the CNN and FAST corner detection modules.
- Deep Learning (CNN) and FAST CD Module: The deep learning and FAST CD modules detect physical objects in captured images, perform object segmentation, and then send the findings to the AR interface module. This system operates on the server.
- Visual Augmentation Module: The Visual augmentation scans the real environment using AR devices and constructs a 3D map as the users move in the real environment. Simultaneously, visual augmentation matches the location of the real object in the 3D map immediately after the deep learning module determines the real object. Unity3D engine was utilized in this study to build an AR environment for AR devices

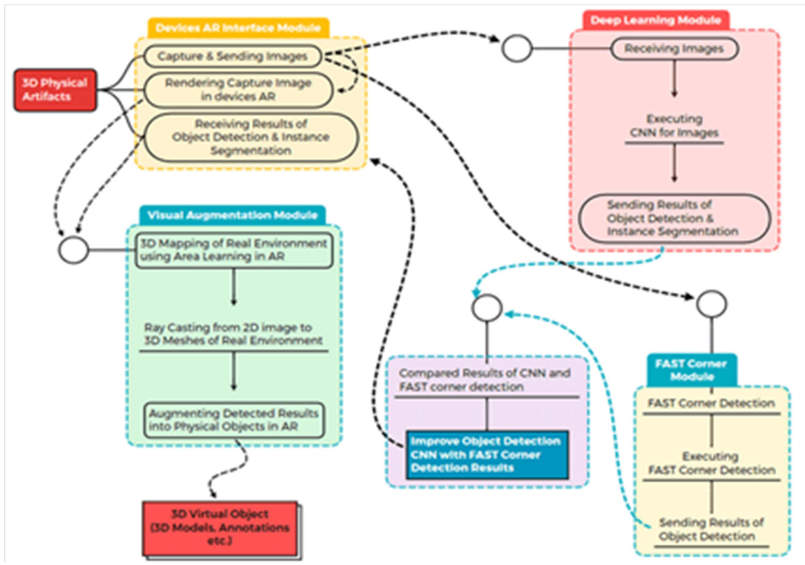


Figure 1. Proposed design architecture AR Markerless

2.4. Proposed Method

This section elaborates the detailed design of the proposed method. This research suggests a deep learning method, specifically utilizing convolutional neural networks (CNN) and FAST corner detection, for the processes of object detection and segmentation in augmented reality systems. Several stages must be traversed to complete the CNN and FAST CD methods for markerless AR. The design of the method is divided into three components: FAST corner detection, CNN, hybrid CNN and FAST corner detection

2.5. Fast Corner Detection

In the initial section, the FAST corner detection method is described as a proposal for the 3D object tracking process in markerless AR systems. The FAST corner detection method was designed to identify objects in an image captured by devices in the AR architecture. Figure 2 illustrates the outcomes of the corner detection process using FAST corner detection. In this image, the corners are derived from the original image, resulting in the detection of objects. However, the effectiveness of object detection is compromised because of the abundance of points on the grass obtained and variations in color gradients within the object image.

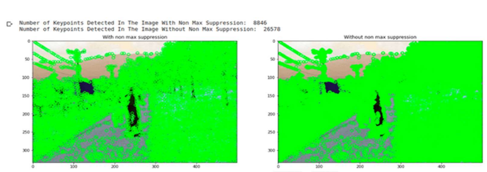


Figure 2. Result of FAST CD

2.6. CNN

Convolutional neural networks (CNNs) are popular deep learning techniques for object detection. In a CNN, an image is divided into multiple parts, with each element serving as the input for the network, ultimately producing object classes through the convolution and pooling layers. TensorFlow was employed for both the training and testing processes. The COCO dataset was used to develop the CNN model by leveraging the pre-trained ResNet architecture. In the initial phase of the study, the CNN was trained with ten classes, followed by a second training phase with 30 classes. The CNN structure differs from other methods, with convolutional layers stacked atop each other, followed by multiple layers, culminating in a fully connected layer. The convolutional layers are pivotal components of a CNN that extract local features and employ kernel weight-sharing mechanisms.

Table 2 Variable CNN

Number of Layer	4 (1 input, 2 hidden, 1 output)
Node	12 node, 8 node, 8 node, 8 node
Activation	Relu, relu, relu, sigmoid
Input dimension	dimConvol
Epoch	10, 50, 100
batch_size	10

The parameters for the CNN in table 2, obtained from training, include object recognition and instance segmentation, which are utilized in the user study. The configuration variables for the constructed CNN are determined based on the input dimensions derived from the dataset.

2.7. Design method CNN and FAST CD

The aim of this study is to design a method that employs a CNN and FAST corner detection for the detection system in markerless AR. The objective of integrating the hybrid method of CNN and FAST corner detection is to rectify the detection errors or complement the detection outcomes of each method. This research proposes a parallel execution of the hybrid CNN and FAST CD methods, where the results of each method are compared to achieve improved outcomes.

The proposed design of the CNN and FAST corner detection method for this study is depicted in Figure 3. There are three main components: the CNN module, the FAST corner detection module, and a comparative module to complement the detection outcomes.

2.8. Analysis tools

The deep learning server was run on a computer with a CPU i7 processor and 16 GB RAM. For develop Unity AR using CPU i7 and 12 GB RAM and Octa-core Max 2.2Hhz, ram 4.00+2.00 GB, android 12 for AR markerless.

3. RESULTS AND DISCUSSION

This comprehensive examination outlines the preparation of the dataset, the methodology employed for feature selection and extraction, the implementation of the CNN-FCD, and a thorough analysis of the experimental outcomes, including a discussion of the results. This section presents in more detail the results obtained from the experiment. In addition, it also analyzes experimental results and compares experimental results with other previous studies.

3.1. Result FAST Corner detection

This comprehensive examination outlines the preparation of the dataset, methodology employed for feature extraction, implementation of CNN and FAST CD, and a thorough analysis of the experimental outcomes, including a discussion of the results. This section presents the results of the experiments in more detail. In addition, the experimental results are analyzed and compared with those of previous studies. Table 3 provides examples of the test results for object detection using FAST corner detection. The testing outcomes were less satisfactory, as observed in instances 2, 5, and 6. This can be attributed to the images containing numerous points or corners that closely resemble the main object

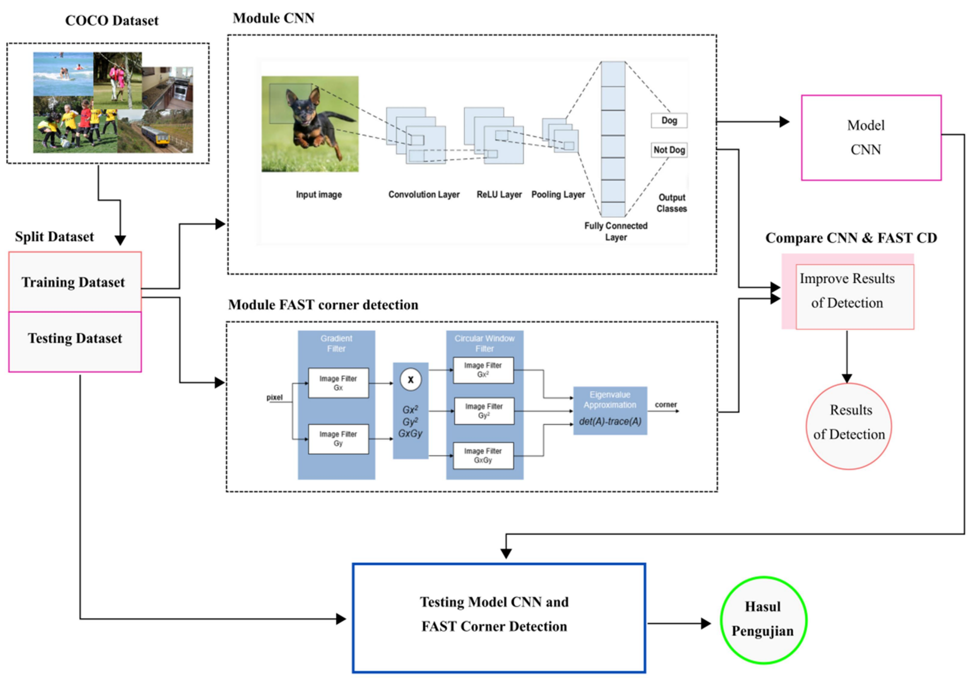
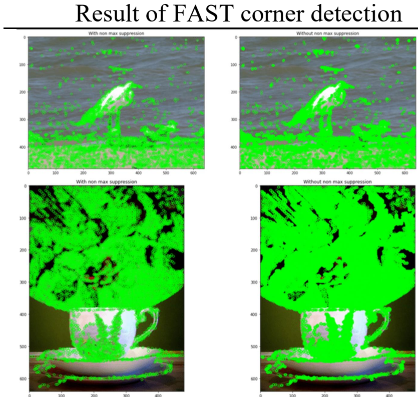
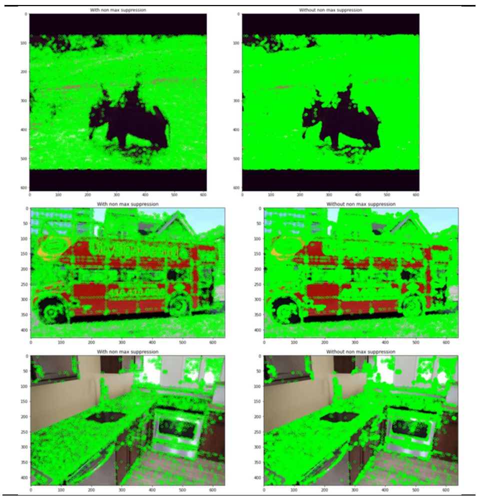


Figure 3. Hybrid CNN and FAST CD

Table 3 Result of FAST CD

Result of FAST corner detection




3.2. Results of CNN

The COCO dataset was used to evaluate the Convolutional Neural Network (CNN) algorithm. This test aims to assess the performance of the CNN in its ability to detect objects and segment images, which will subsequently be utilized in the tracking process within an augmented reality (AR) system. The training and testing processes involved a subset of images from the COCO dataset, specifically 78,458 images for training and 42,100 for testing, which were selected randomly. The evaluation of the testing results included an examination of the confusion matrix values to identify correctly and incorrectly classified images. In this testing process, an image is input into the AR system and the detected object results are displayed as the testing dataset outcomes.

Table 4 Result CNN

Results of CNN

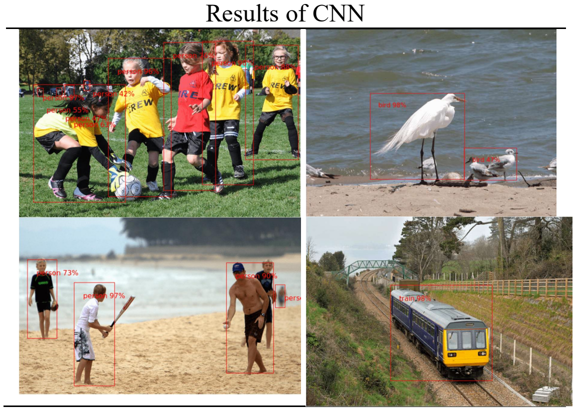


Table 4 presented as a visual outcome, illustrates the iterative trial-and-error steps employed to assess the performance of the Convolutional Neural Network (CNN) algorithm in this study. Images encompass various objects, including humans, animals, and vehicles, thereby increasing the complexity of recognition. The initial testing results indicate satisfactory classification performance by the CNN, as depicted in table 3, despite some minor errors attributed to unclear or excessively small objects. The detailed results of this testing, utilizing a ((CNN), can be observed in the following table 5.

Table 5. Example of testing CNN

Images	Name of Object	Object Detection	CNN	Accuracy
1	Person	Person	Person	P1 : 88%, P2 : 98%, P3 : 78%, P4 : 96%, P5 : 97%, P6 : 42%, P7 : 63%, P9 : 47%, P10 : 55%
2	Person	Person	Person	P1 : 90%, P2 : 97%, P3 : 73% P4 : not detected
3	Animal	Bird	Bird	B1 : 98%, B2 : 47%
4	Train	Vehicle	Train	K1 : 98%

CNN testing was conducted in three testing scenarios involving training processes with 10 epochs, 50 epochs, and 100 epochs. Each testing scenario employed the COCO 2017 dataset consisting of 40,670 randomly selected images in the training process

3.3. Result of Proposed Method

Testing in this research involves evaluating the proposed architecture, method and calculating the performance of the AR system using CNN and FAST corner detection. The testing of the architecture and method was conducted simultaneously because the architecture and method form an integrated unit, where the CNN and FAST corner detection methods are encompassed within the proposed AR architecture. Consequently, the testing results were consolidated into a unified performance metric, including the accuracy, measurement time, and testing loss. The testing analysis was conducted in two parts: testing the CNN without FAST Corner Detection and testing the CNN with FAST Corner Detection

3.3.1. Result of CNN without FAST Corner

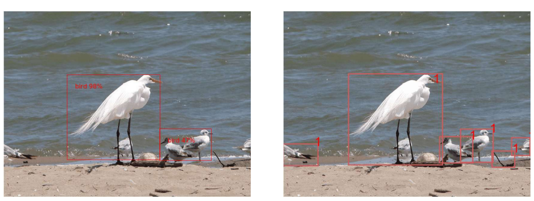
In this section, we discuss the results of epoch testing using the Convolutional Neural Network (CNN) method for object detection and segmentation in the AR system. The number of epochs used in this study ranged between 10, 50, and 100. The comparative experimentation with different epoch values for the CNN algorithm in this study was conducted three times. The experimental results presented in Table 6 demonstrate that the values for each testing epoch yield varying outcomes, although not significantly different. In terms of execution time, testing loss, and testing accuracy, Experiment 1 produced favorable results compared with the other experiments, although the difference was not highly significant.

Table 6. Testing CNN without FAST CD

Test Parameters	Testing		
	1	**2**	**3**
	Epoch 10		
Execution Time	58.01 ms	60,02 ms	58.08 ms
Loss Testing	2.522284	2.732432	2.62424
Testing Accuration	0.8378368	0.834123	0.824123
	Epoch 50		
Execution Time	98.15 ms	99,21 ms	98.01 ms
Loss Testing	2.123234	2.132631	2.25322
Testing Accuration	0.89783823	0.884122	0.87213
	Epoch 100		
Execution Time	158.14 ms	149,24 ms	154.81 ms
Loss Testing	2.25364	2.14264	2.34421
Testing Accuration	0.90723	0.89439	0.91214

3.3.1. Result of CNN with FAST Corner

This study proposes the utilization of the CNN method with FAST corner detection to enhance the performance of the CNN in detecting an object within an image in an AR system. The objective of employing CNN and FAST corner detection is to compare the results of detection and rectify inaccuracies in the detection process when using CNN without FAST corner detection, as illustrated in Figure 4(a), where the object cannot be detected entirely.



(a) (b)

Figure 4. (a) result of CNN without FAST CD, and (b) result CNN with FAST CD

Figure 4(a) shows an example of object detection testing using only the CNN method. In this test, only two objects were detected, one of which had a low accuracy. Figure 4(b) demonstrates the utilization of the CNN method with FAST corner detection in object detection. The results obtained by combining CNN with FAST corner detection exhibited superior detection outcomes compared with using CNN alone. The next test involves the proposed method, namely CNN with FAST corner detection for the AR system, with three epoch tests, similar to CNN without FAST corner detection.

The comparative results of each epoch experiment for the CNN and FAST corner detection algorithms in this study were obtained three times. The experimental results presented in Table 7 show that the values for each epoch yielded varying outcomes, although they were not significantly distant. In terms of execution time, testing loss, and testing accuracy, with an iteration of epoch = 100, Experiment 1 produced favorable results, achieving an accuracy of 0.98 compared to the other experiments.

Table 7. Testing CNN with FAST CD

Parameters	Testing		
	1	2	3
	Epoch 10		
Execution Time	78.09 ms	70,32 ms	75.34 ms
Loss Testing	1.824	2.1322	2.034
Testing Accuration	0.9373	0.92302	0.924123
	Epoch 50		
Execution Time	173.73 ms	161,21 ms	147.34 ms
Loss Testing	1.524	1.627	1.833
Testing Accuration	0.9454	0.9575	0.9538
	Epoch 100		
Execution Time	213.14 ms	219,27 ms	221.13 ms
Loss Testing	1.3441	1.223	1.3242
Testing Accuration	0.9801	0.9741	0.9823

4. DISCUSSION

In this section, a comparison is made between the results of the CNN method for the AR system and the proposed method, namely, CNN and FAST corner detection. The objective of this section was to assess the performance of the proposed method. Table 8 presents the comparison results of the execution times from the experiments conducted in this study. From Table 8, it can be observed that testing for the execution time parameter increases as the number of iterations increases. This holds true for both the CNN with FAST corner detection and the CNN without FAST corner detection testing processes. In the testing process, the CNN and FAST corner detection methods require more time than the CNN method without FAST corner detection.

The next comparison is the accuracy results obtained from the testing of the CNN with FAST corner detection and CNN without FAST corner detection. Table 9 displays the accuracy comparison results for each epoch testing, including the experiments with three epochs: 10, 50, and 100. The results show that the accuracy of using the CNN and FAST corner detection methods for object detection in the AR system is higher than that of using CNN alone. The improvement in accuracy is not only observed in a single epoch but also across all epoch tests (10, 50, and 100), demonstrating a significant enhancement compared to the CNN method without FAST corner detection.

Table 8. Comparison of execution time of CNN results and FAST corner detection and CNN without FAST corner detection

Epoch	Testing 1	Testing 2	Testing 3
		CNN	
10	58.01	60.02	58.08
50	98.15	99.21	98.01
100	158.14	149.24	154.81
		CNN + FAST corner detction	
10	78.09	70.32	75.34
50	173.73	161.21	147.34
100	213.14	219.27	221.13

Table 9. Comparison of the accuracy of CNN and FAST corner detection results and CNN without FAST corner detection

Epoch	Testing 1	Testing 2	Testing 3
		CNN	
10	0.8378368	0.834123	0.824123
50	0.89783823	0.884122	0.87213
100	0.90723	0.89439	0.91214
		CNN + FAST corner detction	
10	0.9373	0.92302	0.924123
50	0.9454	0.9575	0.9538
100	0.9801	0.9741	0.9823

Subsequently, object augmentation testing was conducted in an AR system. The AR system designed in the initial system must determine the visual field, that is, the physical environment of the AR. This testing involves various experiments, starting with testing on textured surfaces, portraits, and landscape orientation alongside real objects of varying sizes, and testing with detailed annotations of augmented objects. Figure 5 illustrates the testing of AR objects on textured surfaces with dark colors. The testing is successful, but it requires more time because the augmentation time is influenced by the light intensity. Figure 6 depicts the augmentation testing on flat surfaces above other real objects. Experiment 6 is an example of testing with three object models simultaneously and detailed annotations of these objects. Testing is also performed by inserting real objects in the midst of virtual objects in an AR system constructed using CNN and FAST corner detection for object detection and segmentation processes. This test indicated that augmentation can also be applied to objects other than the initial visual field.

The testing shown in Figure 7 involves the augmentation of virtual objects between large real objects. The objective of this test is to determine whether the AR system can augment virtual objects in large real objects. Subsequent testing involved the use of the AR system for augmentation with large-sized objects. In this test, an attempt was made to augment a chair object in the AR field, which contains other large objects in the AR environment, as shown in Figure 8.

Figure 5. Augmentation in less intencity

Figure 6 Augmentation with real object

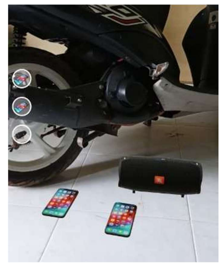


Figure 7. Augmentation with big real object

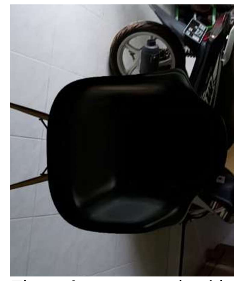


Figure 8. Augmentation big object and real object

4.1. Comparison of Experimental Results with Previous Research

This section presents the results of the experiments and compares them with the findings of prior research. The objective is to evaluate the performance of the CNN in comparison with other methods in the context of augmented reality (AR). While some previous studies have utilized different datasets, this is due to the AR system's reliance on real-world testing data, making it challenging to compare the same datasets and features. The results of this research indicate enhanced AR performance through the application of a CNN. Table 10 provides a comparison of the accuracies achieved using deep learning in this study. The experimental results demonstrate an improvement in the accuracy of the proposed method. Consequently, it can be concluded that the proposed method is successful and exhibits a satisfactory performance. Although some studies did not specify accuracy figures, they asserted that the proposed methods achieved satisfactory accuracy levels.

Table 10. Comparison with previous method

Ref	Methods	Accuracy (%)	Successful Recognize object
D. Conference, P. Stief, et al (2022)	Machine Learning	82.19	√
A. Rahman et al., (2021)	DNN	89.20	√
K. Park, S. Ho, M. Kim, and J. Yeol (2020)	CNN-RGB-D	n/	√
K. Park, M. Kim, S. H. Choi, and J. Y. Lee (2020)	CNN	n/	√
This propose study	CNN + FAST CD	98,98	√

5. CONCLUSION

This study focuses on addressing issues related to object tracking and registration in marker-less augmented reality (AR) systems. To address these challenges, this study proposes an architecture and method for an AR system using a hybrid of a Convolutional Neural Network (CNN) and FAST corner detection. Based on tests employing the proposed CNN and FAST corner detection method for object detection and tracking in markerless AR systems, experiments were conducted with three epoch schemes: 10, 50, and 100. The results of the experiments were measured against several parameters, including execution time, loss testing, and testing accuracy. The findings from the experiments indicate an improvement in the performance of object detection. The accuracy testing results obtained using the CNN and FAST corner detection methods surpassed those of the CNN-only method, reaching 98%. However, this method increases the processing time for object detection. Consequently, the processing time for CNN without FAST corner detection was faster. Additionally, the testing included the successful implementation of 3D object augmentation in markerless AR fields. Testing involves various schemes for determining the augmentation field, placing single and multiple objects for augmentation, and augmenting other physical objects. Future research endeavors will propose a specialized method for image segmentation to enhance the performance of the detection system in markerless AR module architecture, and methods to compare CNN and FAST CD modules to determine improvements in detection results.

UCAPAN TERIMA KASIH

This research was supported by the University of Dinamika Bangsa through the development of human resource programs.

DAFTAR PUSTAKAA

[1] C. Gsaxner *et al.*, “Augmented Reality for Head and Neck Carcinoma Imaging: Description and Feasibility of an Instant Calibration, Markerless Approach,” *Comput. Methods Programs Biomed.*, vol. 200, p. 105854, 2021, doi: 10.1016/j.cmpb.2020.105854.

[2] W. Li, J. Wang, S. Jiao, M. Wang, and S. Li, “Fully convolutional network-based registration for augmented assembly systems,” *J. Manuf. Syst.*, no. December, 2021, doi: 10.1016/j.jmsy.2020.12.017.

[3] A. Brunzini, M. Ciccarelli, M. Sartini, G. Menchi, A. Papetti, and M. Germani, "A novel approach to use marker-less mixed reality applications with in-motion systems," in *Advances on Mechanics, Design Engineering and Manufacturing IV*, S. Gerbino, A. Lanzotti, M. Martorelli, R. M. Buil, C. Rizzi, and L. Roucoules, Eds. Cham: Springer International Publishing, 2023, pp. 1401–1412. doi: 10.1007/978-3-031-15928-2_122.

[4] S. Boonbrahm, P. Boonbrahm, and C. Kaewrat, “The use of marker-based augmented reality in space measurement,” *Procedia Manuf.*, vol. 42, no. 2019, pp. 337–343, 2020, doi: 10.1016/j.promfg.2020.02.081.

[5] X. Hu, F. R. y Baena, and F. Cutolo, "Head-mounted augmented reality platform for markerless orthopaedic navigation," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 2, pp. 910–921, 2022. doi: 10.1109/JBHI.2021.3088442

[6] T. Jia, Z. A. Taylor, and X. Chen, "Long term and robust 6DoF motion tracking for highly dynamic stereo endoscopy videos," *Computerized Medical Imaging and Graphics*, vol. 94, p. 101995, 2021. doi: 10.1016/j.compmedimag.2021.101995.

[7] C.-H. Chu, C.-Y. Weng, and Y.-T. Chen, "Enhancing manual inspection in semiconductor manufacturing with integrated augmented reality solutions," Journal of Manufacturing Systems, vol. 77, pp. 933–945, 2024. doi: 10.1016/j.jmsy.2024.10.028.

[8] M. Venkatesan *et al.*, “Virtual and augmented reality for biomedical applications,” *Cell Reports Med.*, vol. 2, no. 7, pp. 1–13, 2021, doi: 10.1016/j.xcrm.2021.100348.

[9] J. M. Runji, Y.-J. Lee, and C.-H. Chu, "Systematic literature review on augmented reality-based maintenance applications in manufacturing centered on operator needs," *International Journal of Precision Engineering and Manufacturing-Green Technology*, vol. 10, no. 2, pp. 567–585, Mar. 2023. doi: 10.1007/s40684-022-00444-w.

[10] J. Angelopoulos and D. Mourtzis, "An intelligent product service system for adaptive maintenance of engineered-to-order manufacturing equipment assisted by augmented reality," *Applied Sciences*, vol. 12, no. 11, Article no. 5349, 2022. doi: 10.3390/app12115349.

[11] J. Song and J. Kook, "Visual SLAM based spatial recognition and visualization method for mobile AR systems," *Applied System Innovation*, vol. 5, no. 1, Article no. 11, 2022. doi: 10.3390/asi5010011.

[12] M. Wang, C. Sun, and A. Sowmya, "Efficient corner detection based on corner enhancement filters," *Digital Signal*

Processing, vol. 122, p. 103364, 2022. doi: 10.1016/j.dsp.2021.103364.(E. Rosten and T. Drummond, “Machine Learning for High-Speed Corner Detection,” in *Computer Vision -- ECCV 2006*, 2006, pp. 430–443.)

[13] N. Nurhadi, E. A. Winanto, and S. Saparudin, “Enhance Object Tracking on Augmented Reality Markerless using FAST Corner Detection,” *2021 IEEE 5th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng.*, pp. 66–70, 2022.

[14] V. Arulkumar, S. Jaya Prakash, E. K. Subramanian, and N. Thangadurai, "An intelligent face detection by corner detection using special morphological masking system and fast algorithm," in *Proc. 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, Trichy, India, 2021, pp. 1556-1561. doi: 10.1109/ICOSEC51865.2021.9591857.

[15] S. Chakraborty and L. Dey, "Virtual and Augmented Reality with Embedded Systems," in *Computing for Data Analysis: Theory and Practices*, Singapore: Springer Nature Singapore, 2023, pp. 75-96. doi: 10.1007/978-981-19-8004-6_4.

[16] X. Shi, W. Xu, and J. Wang, "An improved feature point-based LK optical flow algorithm for robotic arm control," in *Eighth International Symposium on Advances in Electrical, Electronics, and Computer Engineering (ISAEECE 2023)*, Shubin Yan and Pavel Loskot, Eds., SPIE, 2023, vol. 12704, pp. 127043G. doi: 10.1117/12.2680052.

~~[17]~~ J. Li, B. Wang, S. Zhu, X. Cao, F. Zhong, W. Chen, T. Li, J. Gu, and X. Qin, "BCOT: A Markerless High-Precision 3D Object Tracking Benchmark," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 6697-6706.

[18] Y. Ghasemi, H. Jeong, S. H. Choi, K.-B. Park, and J. Y. Lee, "Deep learning-based object detection in augmented reality: A systematic review," *Computers in Industry*, vol. 139, 2022, p. 103661. doi.org/10.1016/j.compind.2022.103661.

[19] K. Mahdikhanlou and H. Ebrahimnezhad, "3D object deforming and manipulating through dynamic hand gestures," *Entertainment Computing*, vol. 43, 2022, p. 100507. [Online]. Available: https://doi.org/10.1016/j.entcom.2022.100507.

[20] A. Gaurav, B. B. Gupta, and P. K. Panigrahi, "A comprehensive survey on machine learning approaches for malware detection in IoT-based enterprise information system," *Enterprise Information Systems*, vol. 17, no. 3, p. 2023764, 2023. [Online]. Available: https://doi.org/10.1080/17517575.2021.2023764.

[21] M. M. Srikantamurthy, V. P. S. Rallabandi, D. B. Dudekula, S. Natarajan, and J. Park, "Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning," *BMC Medical Imaging*, vol. 23, no. 1, p. 19, 2023., doi.org/10.1186/s12880-023-00964-0.

[22] G. Apostolopoulos, D. Andronas, N. Fourtakas, and S. Makris, “Operator training framework for hybrid environments: An Augmented Reality module using machine learning object recognition,” *Procedia CIRP*, vol. 106, pp. 102–107, 2022, doi: 10.1016/j.procir.2022.02.162.

[23] A. Rahman *et al.*, “An integrated framework of sensing, machine learning, and augmented reality for aquaculture prawn farm management,” *Aquac. Eng.*, vol. 95, no. July, p. 102192, 2021, doi: 10.1016/j.aquaeng.2021.102192.

[24] K.-B. Park, S. H. Choi, and J. Y. Lee, "Self-training based augmented reality for robust 3D object registration and task assistance," *Expert Systems with Applications*, vol. 238, Part F, p. 122331, 2024.

[25] S. Raj, L. R. D. Murthy, T. A. Shanmugam, G. Kumar, A. Chakrabarti, and P. Biswas, "Augmented reality and deep learning based system for assisting assembly process," *Journal on Multimodal User Interfaces*, vol. 18, no. 1, pp. 119-133, 2024. [Online]. Available: https://doi.org/10.1007/s12193-023-00428-3.