http://doi.org/10.35784/iapgos.6022

received: 20.03.2024 | revised: 18.05.2024 | accepted: 4.09.2023 | available online: 30.09.2024

REAL-TIME DETECTION AND CLASSIFICATION OF FISH IN UNDERWATER ENVIRONMENT USING YOLOV5: A COMPARATIVE STUDY OF DEEP LEARNING ARCHITECTURES

Rizki Multajam[1], Ahmad Faisal Mohamad Ayob[1], W.S. Mada Sanjaya[2], Aceng Sambas[3,4], Volodymyr Rusyn[5], Andrii Samila[5]

[1]Universiti Malaysia Terengganu, Faculty of Ocean Engineering Technology and Informatics, Kuala Nerus, Malaysia, [2]Universitas Islam Negeri Sunan Gunung Djati, Department of Physics, Bandung, Indonesia, [3]Universiti Sultan Zainal Abidin, Faculty of Informatics and Computing, Campus Besut, Malaysia, [4]Universitas Muhammadiyah Tasikmalaya, Department of Mechanical Engineering, Tasikmalaya, Indonesia, [5]Yuriy Fedkovych Chernivtsi National University, Department of Radio Engineering and Information Security, Chernivtsi, Ukraine

Abstract. *This article explores techniques for the detection and classification of fish as an integral part of underwater environmental monitoring systems. Employing an innovative approach, the study focuses on developing real-time methods for high-precision fish detection and classification. The implementation of cutting-edge technologies, such as YOLO (You Only Look Once) V5, forms the basis for an efficient and responsive system. The study also evaluates various approaches in the context of deep learning to compare the performance and accuracy of fish detection and classification. The results of this research are expected to contribute to the development of more advanced and effective aquatic monitoring systems for understanding underwater ecosystems and conservation efforts.*

Keywords: deep learning, YOLOv5, real-time methods, ONNX, automatic fish detection and classification

WYKRYWANIE I KLASYFIKACJA RYB W CZASIE RZECZYWISTYM W ŚRODOWISKU PODWODNYM PRZY UŻYCIU YOLOV5: BADANIE PORÓWNAWCZE ARCHITEKTUR GŁĘBOKIEGO UCZENIA

Streszczenie. *Niniejszy artykuł bada metody wykrywania i klasyfikacji ryb jako integralną część podwodnych systemów monitorowania środowiska. Wykorzystując innowacyjne podejście, badania koncentrują się na opracowaniu metod w czasie rzeczywistym do bardzo dokładnego wykrywania i klasyfikacji ryb. Wprowadzenie zaawansowanych technologii, takich jak YOLO (You Only Look Once) V5, stanowi podstawę wydajnego i responsywnego systemu. Badanie ocenia również różne podejścia w kontekście głębokiego uczenia się, aby porównać wydajność i dokładność wykrywania i klasyfikacji ryb. Oczekuje się, że wyniki tych badań przyczynią się do rozwoju bardziej zaawansowanych i wydajnych systemów monitorowania zbiorników wodnych w celu zrozumienia podwodnych ekosystemów i wysiłków na rzecz ochrony przyrody.*

Slowa kluczowe: deep learning, YOLOv5, metody czasu rzeczywistego, ONNX, automatyczne wykrywanie i klasyfikacja ryb

Introduction

The advancement of real-time camera technology and artificial intelligence [5] has ushered in new opportunities in the fisheries industry [11, 16]. The combination of real-time cameras with artificial neural networks (ANN) and deep learning techniques [12] promises an efficient and accurate solution for real-time fish detecting [9]. The Artificial Neural Network (ANN) will then indicate whether the object in the camera's image is a fish [22]. The real-time use of cameras combined with deep learning [25] will result in an integrated approach to detect and classify fish in underwater videos [13] thus, in the future, this research can be applied to underwater robot systems to analyze the behaviour of various fish species [3].

This research aims to implement real-time camera technology in conjunction with deep learning for direct fish counting in fisheries environments [16]. By harnessing the power of machine learning algorithms [21], this study not only seeks an efficient alternative to manual counting methods but also aims to pave the way for a more holistic and intelligent fisheries management approach.

One way to manually count the number of fish is by placing several fish on a container placed above a water tank or pond, then pushing some fish into a channel and counting them one by one [17]. Meanwhile, fish counting using artificial intelligence technology, one of which involves the implementation of Artificial Neural Networks (ANN) [28]. Deep learning methods [8, 18] are employed to classify fish species [3, 8, 10, 13, 17, 18, 21, 28]. Therefore, the use of real-time cameras combined with deep learning [26] can result in accurate and real-time fish counting.

The results of fish classification can provide significant benefits in various aspects, including faster and more accurate identification of fish species [27], monitoring fish populations, and supporting fisheries resource management to maintain environmental sustainability and ensure an adequate supply of fish for communities [4].

1. Theoretical background

The current technological advancements have significant impacts on various sectors [15], including the cultivation of flora and fauna. The use of artificial neural networks has achieved great success, particularly in handling text, images, videos, and so forth [7]. Furthermore, the use of artificial neural networks can be employed in sorting machines and the classification of fruit ripeness [1]. Furthermore, the advancement of artificial intelligence can be applied in robotics systems to recognize human faces in real-time using cameras [23], by embedding a database into the system, enabling the robot's camera to recognize these facial patterns.

1.1. You Look Only Once (YOLO) v5

YOLO (You Only Look Once) [20] is a renowned object detection model known for its high speed and efficient approach to address issues present in traditional architectures [29].

In YOLO, object detection and class probability predictions are performed directly on the image with a single evaluation [14], avoiding the complexity of conventional detection pipelines. YOLO v5, as the latest iteration, continues to improve performance and accuracy by utilizing regression as the detection approach and optimizing the network architecture. With 24 convolutional layers followed by 2 fully connected layers [6], the model is capable of making global predictions about the image, resulting in fewer false positive predictions. As shown in Fig. 1.

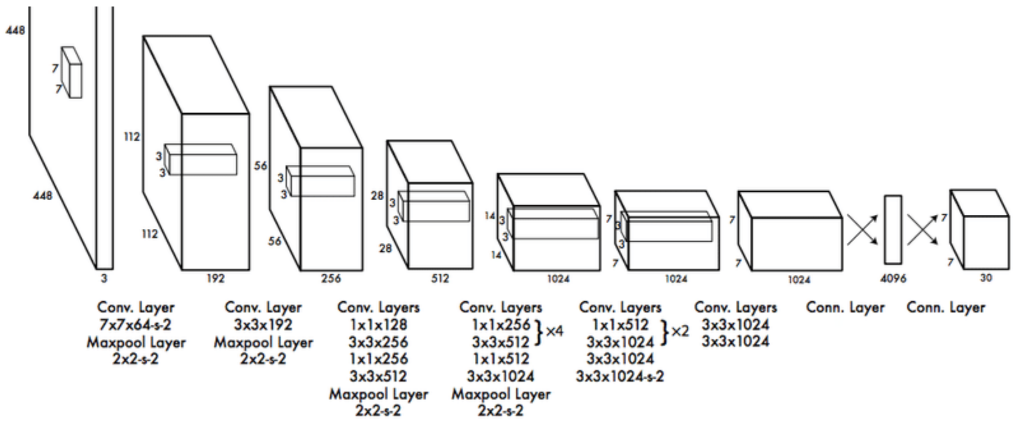


Fig. 1. The Architecture of YOLO [20]

In Fig. 1, it can be observed that the alternating 1×1 convolutional layers are used to reduce the feature space from the previous layers [20]. Pretraining is conducted on the convolutional layers for the ImageNet classification task with half resolution (224×224 input images), and then the resolution is doubled for the detection process [20]. In addition, YOLO can also be applied to detect objects in water [2] and doing it in real-time object detection in videos on embedded devices [24].

1.2. Evaluation of model

The commonly used evaluation metric is accuracy, which measures how often the model provides correct predictions in classifying images of an object. This accuracy approach is calculated by dividing the number of correct predictions by the total number of samples [19]. The commonly used metrics in cases of image classification when the number of positive and negative samples is imbalanced include precision, recall, and F1-Score.

Precision is one of the essential evaluation metrics in assessing the performance of a classification system. Precision provides an insight into how accurately the model identifies positive outcomes. The precision formula is expressed as:

$$\text{Precision} = \frac{TruePositive}{TruePositive + FalsePositive} \quad (1)$$

Precision measures, among all the model's positive predictions, what percentage is truly relevant or correct. High precision indicates that the model tends to provide accurate positive predictions and minimizes false positive predictions. Recall, also known as Sensitivity or True Positive Rate, is a crucial evaluation metric in assessing the performance of a classification system. Recall provides an insight into how well the model can recognize all positive outcomes that should be identified. The recall formula is expressed as:

$$\text{Recall} = \frac{TruePositive}{TruePositive + FalseNegative} \quad (2)$$

Recall measures, among all the actual positive instances, what percentage is successfully identified by the model. High recall indicates that the model can detect most of the actual positive instances.

$$\text{Specificity} = \frac{TrueNegative}{FalseNegative + TrueNegative} \quad (3)$$

Specificity indicates the model's ability to correctly classify all negative outcomes. Meanwhile, F1-Score provides a balanced measurement between precision and recall by combining. F1-Score provides a holistic overview of the model's performance by considering both aspects. F1-Score is calculated using the harmonic mean of precision and recall:

$$\text{F1-Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (4)$$

F1-Score has a range of values between 0 and 1, where a value of 1 indicates perfect performance with optimal precision and recall. F1-Score is particularly valuable when there is an imbalance between positive and negative classes in the dataset, as it can provide a more accurate representation of the overall model performance.

In its application context, F1-Score is commonly used in various fields, including image processing, object detection, and evaluating the performance of classification models in imbalanced class scenarios.

2. Experimental method

2.1. The method and system for fish classification

In this experimental method, we attempted to collect samples of goldfish (Carassius auratus) and koi fish (Cyprinus rubrofuscus), totalling 44 samples for each type of fish. Subsequently, the fish samples were classified using Roboflow and YOLO to generate an .onnx file, which is a deep learning model file. Next, to test the samples, a testing program was developed using Python.

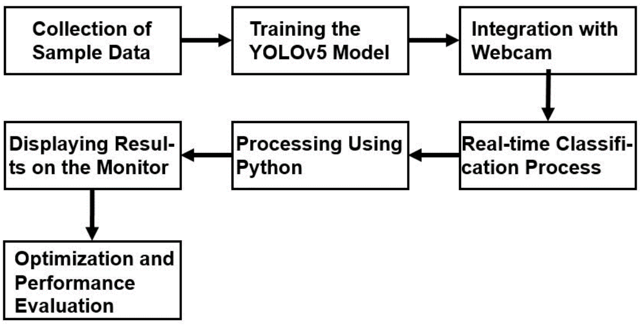


Fig. 2. Illustration of a classification system

Fig. 2 illustrates the flow conducted during the research.

Firstly, collecting samples of goldfish and koi fish to be used as a database. Subsequently, the fish samples will be classified using Roboflow and YOLOv5, to produce a fish classification model created in the ONNX format. Next, the model is tested using fish and non-fish objects.

2.2. Method for real-time fish observation

In the experimental process of real-time fish classification using YOLOv5, we commenced by gathering diverse samples of goldfish (Carassius auratus) and koi fish (Cyprinus rubrofuscus), totalling 44 samples for each species. Subsequently, these fish samples underwent classification employing Roboflow and YOLOv5, resulting in the creation of an ONNX file – a deep learning model file. The webcam was then utilized to capture real-time images of fish objects, with the Python programming language being employed to process the webcam recordings. This processing incorporated the previously generated ONNX file as the learning model.

The devices used for fish classification include a 1080-pixel HD camera, an Intel Core i3 processor, samples of koi fish (Cyprinus rubrofuscus) and goldfish (Carassius auratus), and Python 3.9. Each device serves its respective function, collectively forming a system for fish classification as depicted in Fig. 3.

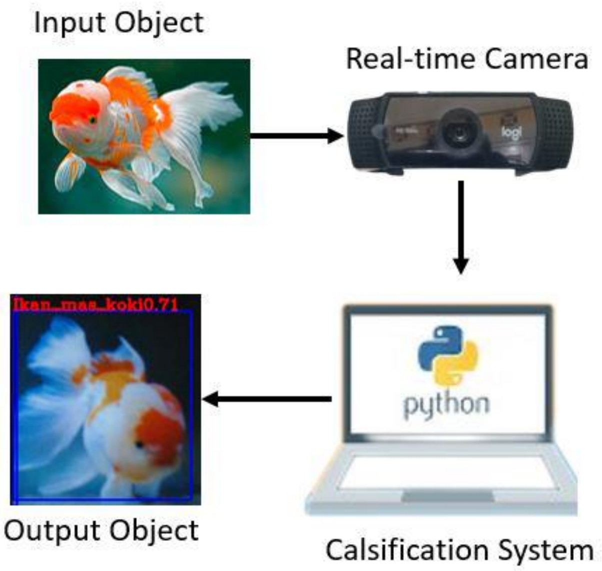


Fig. 3. Implementation in real-time fish detection

Fig. 3 illustrates the implementation of the YOLOv5 classification results. The webcam will capture real-time images of fish objects, and the output from the webcam recording will be processed in a Python program, where the ONNX file has been incorporated as the learning model. Subsequently, the system will display images of fish on the monitor, complete with the type of fish captured by the camera. The system successfully displayed real-time images of detected fish, accompanied by their respective species, on the monitor. Additionally, we optimized the model for improved speed and accuracy, evaluating performance metrics such as response time and accuracy rate. The real-time fish classification system demonstrates potential applications in fisheries monitoring, environmental research, and the aquaculture industry, paving the way for ongoing advancements in accuracy and technological integration.

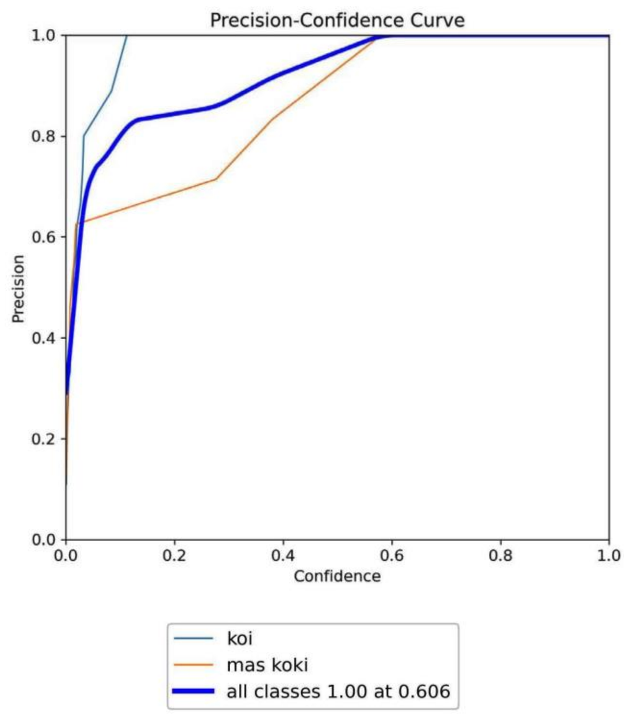
3. Result and discussion

3.1. The YOLOv5 classification process

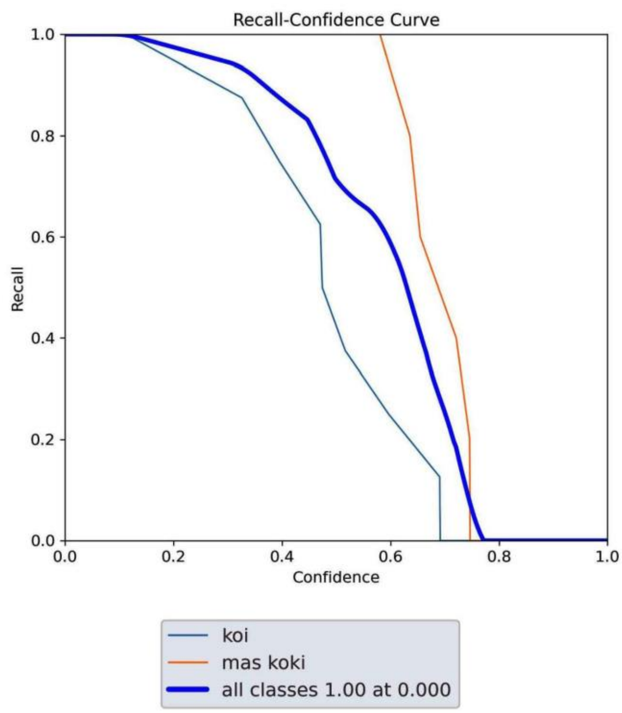
The Precision and Recall curve is a powerful visual tool for evaluating the performance of object detection models like YOLOv5. Precision and Recall are two key metrics used to measure how well the model can correctly identify objects and how many positive objects are overlooked. The precision and recall graph of goldfish and koi, as indicated in Fig. 4.

In conclusion, the Precision and Recall curve in Fig. 4 serves as a crucial tool for evaluating and optimizing the YOLOv5 model's performance in detecting goldfish and koi fish, providing insights for fine-tuning the model based on specific application requirements.

The next step to optimize the model balance for goldfish and koi fish detection utilizes the F1-Score, calculated with the formulation that integrates precision and recall. This process results in a graph, as seen in Fig. 5.



a)



b)

Fig. 4. The precision and recall graph of goldfish and koi: a – precision curve, b – recall curve

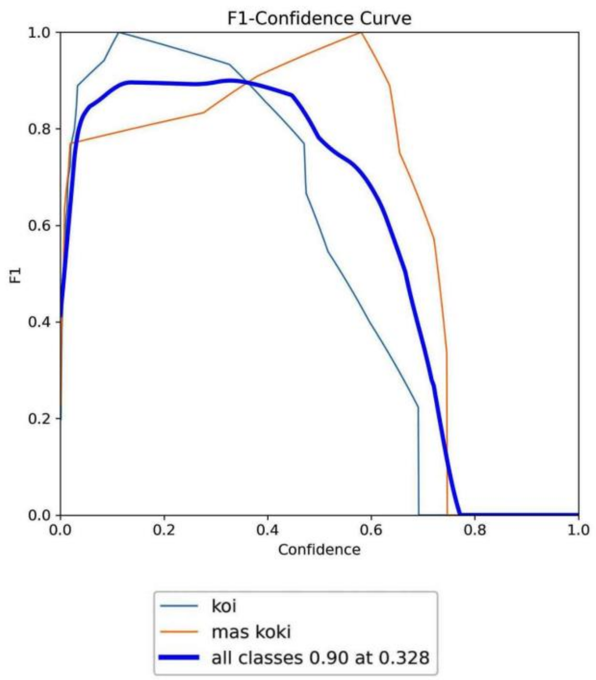


Fig. 5. The F1-Score curve for the goldfish and koi fish model

Fig. 5 the curve visually represents the outcomes of this calculation, offering insights into the model's performance in achieving an optimal balance between precision and recall. It aids in determining the optimal detection threshold for this fish detection application.

3.2. Application in the fish detection system

The YOLOv5 fish detection application utilizes a deep neural network to rapidly and accurately detect fish in images, involving training the model with a properly annotated dataset and evaluating performance using metrics such as precision, recall, and mAP.

In this project, only two types of fish are classified: koi fish (Cyprinus rubrofuscus) and goldfish (Carassius auratus). The webcam will capture real-time images of fish in the aquarium. Subsequently, the captured images will be processed on the computer through Python programming, distinguishing between fish and other objects in their surroundings. The results of fish classification can be observed in Fig. 6.

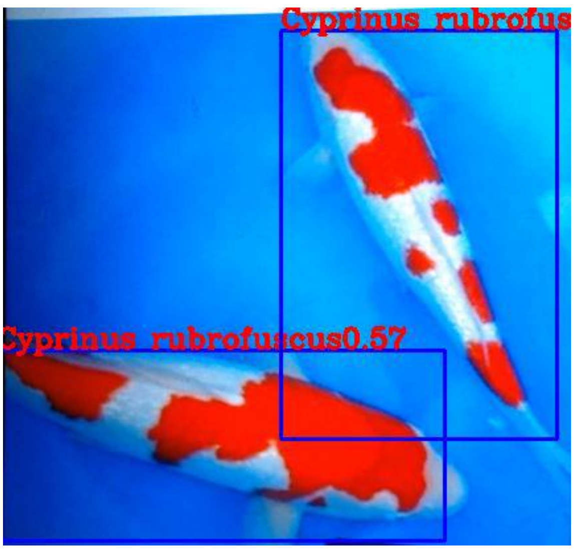


Fig. 6. Classification results of Koi fish Cyprinus rubrofuscus)

Fig. 6 represents the results of fish detection using YOLOv5 by classifying fish types. From the image, koi fish (Cyprinus rubrofuscus) is detected with an accuracy of 57%. Subsequently, testing is also conducted on the goldfish (Carassius auratus) type, as shown in Fig. 7.

Fig. 7. Classification results of Golden fish (Carassius auratus)

Fig. 7 shows the results obtained after detecting goldfish. The processing of goldfish figure yielded a result of 64%. The accuracy percentage of fish is influenced by the visibility range and 5 water clarity, which significantly affect the camera. Water turbidity will consistently reduce the capture results of objects by the camera.

In this research, there is a constraint in observing fish because the displayed images on the monitor always experience delays due to insufficient computer processor specifications. To maximize the fish observation process, supporting devices such as a high-resolution camera and a computer with a fast processor specification are required.

4. Conclusion

Classification of fish underwater holds great potential for providing new insights in the development of underwater sciences. The use of YOLOv5 in fish classification, with the evaluation of metrics such as Precision, Recall, and F1-Score, demonstrates satisfactory object detection capabilities, enabling model optimization to achieve a balance between accuracy and detection completeness. To obtain optimal results in fish observation underwater, a high-speed processor is essential to ensure uninterrupted display of images.

References

[1] Abdul Aziz M. F. et al.: Development of Smart Sorting Machine using artificial intelligence for Chili Fertigation Industries. Journal of Automation, Mobile Robotics and Intelligent Systems 28, 2022, 44–52 [https://doi.org/10.14313/jamris/4-2021/26].

[2] Ayob A. et al.: Analysis of pruned neural networks (mobilenetv2-yolo v2) for underwater object detection. 11th National Technical Seminar on Unmanned System Technology 2019 NUSYS'19, Springer Singapore, Singapore, 2021, 87–98.

[3] Boudhane M., Benayad N.: Underwater Image Processing Method for Fish Localization and Detection in Submarine Environment. Journal of Visual Communication and Image Representation 39, 2016, 226–238 [https://doi.org/10.1016/j.jvcir.2016.05.017].

[4] Brownscombe J. W. et al.: The Future of Recreational Fisheries: Advances in Science, Monitoring, Management, and Practice. Fisheries Research 211, 2019, 247–255 [https://doi.org/10.1016/j.fishres.2018.10.019].

[5] Chen PH. C. et al.: An Augmented Reality Microscope with Real-time Artificial Intelligence Integration for Cancer Diagnosis. Nature Medicine 25(9), 2019, 1453–1457 [https://doi.org/10.1038/s41591-019-0539-7].

[6] Du J.: Understanding of Object Detection Based on CNN Family and YOLO. Journal of Physics: Conference Series 1004, 2018, 012029 [https://doi.org/10.1088/1742-6596/1004/1/012029].

[7] Fan F.-L. et al.: On Interpretability of Artificial Neural Networks: A Survey. IEEE Transactions on Radiation and Plasma Medical Sciences 5(6), 2021, 741–760 [https://doi.org/10.1109/trpms.2021.3066428].

[8] Hong S. et al.: Opportunities and Challenges of Deep Learning Methods for Electrocardiogram Data: A Systematic Review. Computers in Biology and Medicine 122, 2020, 103801 [https://doi.org/10.1016/j.compbiomed.2020.103801].
[9] Hu J. et al.: Real-time Nondestructive Fish Behavior Detecting in Mixed Polyculture System Using Deep-learning and Low-cost Devices. Expert Systems With Applications 178, 2021, 115051 [https://doi.org/10.1016/j.eswa.2021.115051].
[10] Iqbal M. A. et al.: Automatic Fish Species Classification Using Deep Convolutional Neural Networks. Wireless Personal Communications 116(2), 2019, 1043–1053 [https://doi.org/10.1007/s11277-019-06634-1].
[11] Isabelle D. A., Westerlund M.: A Review and Categorization of Artificial Intelligence-Based Opportunities in Wildlife, Ocean and Land Conservation. Sustainability 14(4), 2022, 1979 [https://doi.org/10.3390/su14041979].
[12] Ismail N., Owais A. M.: Real-time Visual Inspection System for Grading Fruits Using Computer Vision and Deep Learning Techniques. Information Processing in Agriculture 9(1), 2022, 24–37 [https://doi.org/10.1016/j.inpa.2021.01.005].
[13] Jalal A. et al.: Fish Detection and Species Classification in Underwater Environments Using Deep Learning with Temporal Information. Ecological Informatics 57, 2020, 101088 [https://doi.org/10.1016/j.ecoinf.2020.101088].
[14] Jing L. et al.: Video You Only Look Once: Overall Temporal Convolutions for Action Recognition. Journal of Visual Communication and Image Representation 52, 2018, 58–65 [https://doi.org/10.1016/j.jvcir.2018.01.016].
[15] Khan A. N. et al.: Sectorial Study of Technological Progress and CO2 Emission: Insights From a Developing Economy. Technological Forecasting and Social Change 151, 2020, 119862 [https://doi.org/10.1016/j.techfore.2019.119862].
[16] Khokher M. R. et al.: Early Lessons in Deploying Cameras and Artificial Intelligence Technology for Fisheries Catch Monitoring: Where Machine Learning Meets Commercial Fishing. Canadian Journal of Fisheries and Aquatic Sciences 79(2), 2022, 257–266 [https://doi.org/10.1139/cjfas-2020-0446].
[17] Klapp I. et al.: Ornamental Fish Counting by Non-imaging Optical System for Real-time Applications. Computers and Electronics in Agriculture 153, 2018, 126–133 [https://doi.org/10.1016/j.compag.2018.08.007].
[18] Liu H., Lang B.: Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey. Applied Sciences 9(20), 2019, 4396 [https://doi.org/10.3390/app9204396].
[19] Mada Sanjaya W. S.: Deep Learning Citra Medis Berbasis Pemrograman Python. Bolabot, 2023.
[20] Redmon J. et al.: You Only Look Once: Unified, Real-Time Object Detection. arXiv.org, 8 June 2015, arxiv.org/abs/1506.02640.
[21] Reynard D., Shirgaokar M.: Harnessing the Power of Machine Learning: Can Twitter Data Be Useful in Guiding Resource Allocation Decisions During a Natural Disaster? Transportation Research Part D: Transport and Environment 77, 2019, 449–463 [https://doi.org/10.1016/j.trd.2019.03.002].
[22] Rico-Díaz Á. J. et al.: An Application of Fish Detection Based on Eye Search With Artificial Vision and Artificial Neural Networks. Water 12(11), 2020, 3013 [https://doi.org/10.3390/w12113013].
[23] Sanjaya W. S. et al.: The Design of Face Recognition and Tracking for Human-robot Interaction. 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering – ICITISEE). IEEE, 2017 [https://doi.org/10.1109/icitisee.2017.8285519].
[24] Shafiee M. J. et al.: Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video. arXiv.org, 18 Sept. 2017, arxiv.org/abs/1709.05943.
[25] Unlu E. et al.: Deep Learning-based Strategies for the Detection and Tracking of Drones Using Several Cameras. IPSJ Transactions on Computer Vision and Applications 11(1), 2019 [https://doi.org/10.1186/s41074-019-0059-x].
[26] Wang D. et al.: UAV Environmental Perception and Autonomous Obstacle Avoidance: A Deep Learning and Depth Camera Combined Solution. Computers and Electronics in Agriculture 175, 2020, 105523 [https://doi.org/10.1016/j.compag.2020.105523].
[27] Xiu L. et al.: Fast Accurate Fish Detection and Recognition of Underwater Images With Fast R-CNN. OCEANS 2015 – MTS/IEEE Washington. IEEE, 2015 [https://doi.org/10.23919/oceans.2015.7404464].
[28] Zhang L. et al.: Automatic Fish Counting Method Using Image Density Grading and Local Regression. Computers and Electronics in Agriculture 179, 2020, 105844 [https://doi.org/10.1016/j.compag.2020.105844].
[29] Zhao Zhong-Qiu et al.: Object Detection With Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems 30(11), 2019, 3212–3232 [https://doi.org/10.1109/tnnls.2018.2876865].

Eng. Rizki Multajam
e-mail: multajamrizki85@gmail.com

Rizki Multajam received his Bachelor's degree in Physics in 2016 from the State Islamic University of Sunan Gunung Djati Bandung, Indonesia. In 2023, he pursued a Master's degree in Mechanical Engineering specializing in Maritime Engineering and Technology at Universiti Malaysia Terengganu, Malaysia. His current research focus includes instrumentation physics, robotics, control systems, automation systems, mechatronics, signal processing, computer vision, underwater vehicles, artificial intelligence, and machine learning.

https://orcid.org/0009-0005-3657-4588

Ph.D. Ahmad Faisal Mohamad Ayob
e-mail: ahmad.faisal@umt.edu.my

Associate professor in Maritime Technology at the Universiti Malaysia Terengganu. Obtained his Ph.D. in Mechanical Engineering from the University of New South Wales at Australian Defense Force Academy in 2011. His focus area is hull performance design, specifically high-speed boat. Dr. Faisal is also focusing on software-hardware development, specifically ship simulators with the marine engineering and marine industry. Subject taught by Dr. Faisal are Engineering Dynamics and Marine Engineering Design. He served as an Executive Committee member for IEEE-Ocean.

https://orcid.org/0000-0001-7123-4662

Ph.D. W.S. Mada Sanjaya
e-mail: madasws@gmail.com

Mada Sanjaya W.S. received his Ph.D in Mathematics from the University Malaysia Terengganu, Malaysia in 2012. He was first appointed as a Lecturer at the UIN Sunan Gunung Djati Bandung, Indonesia in 2009. His research interests include nonlinear dynamical systems, chaotic systems, artificial intelligence, soft computing and robotic systems.

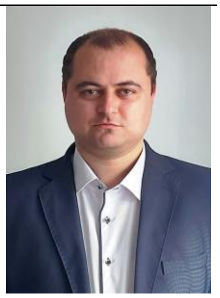
https://orcid.org/0000-0003-1963-193X

Ph.D. Aceng Sambas
e-mail: acengsambas@unisza.edu.my

Aceng Sambas received his Ph.D in Mathematics from the Universiti Sultan Zainal Abidin (UniSZA), Malaysia in 2020. He has assistant professor in the Universiti Sultan Zainal Abidin, Malaysia and Muhammadiyah University of Tasikmalaya, Indonesia since 2022. His current research focuses on dynamical systems, chaotic signals, electrical engineering, computational science, signal processing, robotics, embedded systems and artificial intelligence.

https://orcid.org/0000-0002-1623-0770

Ph.D. Volodymyr Rusyn
e-mail: rusyn_v@ukr.net



Yuriy Fedkovych Chernivtsi National University Department of Radio Engineering and Information Security. Ph.D. (Engineering Sciences) assistant professor of Department of Radio Engineering and Information Security.
Research interests: modeling of nonlinear equations and chaotic generators; control of chaotic oscillations. Author of nearly 140 publications.

https://orcid.org/0000-0001-6219-1031

Prof. Andrii Samila
e-mail: a.samila@chnu.edu.ua

Yuriy Fedkovych Chernivtsi National University. D.Sc. (Engineering), Full Professor, Vice Rector for Scientific Research.
Research interests: IoT, Microelectronics & Electronic Packaging, Signal Processing, Computer Hardware Design, Robotics, High Energy & Nuclear Physics. Author of nearly 200 publications in this research area.

https://orcid.org/0000-0001-8279-9116