

NHẬN DẠNG MỘT SỐ LOÀI CÁ BẰNG PHƯƠNG PHÁP HỌC SÂU

FISH SPECIES RECOGNITION USING DEEP LEARNING

TRẦN HỒ ĐẠT^{1a}, TRẦN PHAN AN TRƯỜNG¹,
LÊ THỊ HẠNH HIỀN¹, NGUYỄN HOÀNG ĐÌNH¹

¹Trường Đại học Sư phạm Kỹ thuật Vĩnh Long

^aTác giả liên hệ: datth@vlute.edu.vn

Nhận bài (Received): 7/6/2023; Phản biện (Reviewed): 13/6/2023; Chấp nhận đăng (Accepted): 25/7/2023

TÓM TẮT

Nhận dạng loài cá đóng vai trò cực kỳ quan trọng trong việc đánh giá đa dạng loài cá tại Việt Nam, quản lý sản xuất, và theo dõi hệ sinh thái, đặc biệt là xác định các loài có nguy cơ tuyệt chủng. Điều này làm cho việc nhận dạng loài cá trở thành một yếu tố cần thiết và đòi hỏi sự chính xác và độ tin cậy cao. Tuy nhiên, việc nhận dạng đòi hỏi sự chính xác và độ tin cậy cao thường yêu cầu nhiều thời gian và chi phí. Trong bài báo này, chúng tôi đề xuất kỹ thuật phân loại dựa trên kiến trúc mạng ResNet152V2 và Xception, chúng tôi thu thập bộ dữ liệu gồm 820 hình ảnh của các loài cá và tiến hành phân loại các loài cá dựa trên tập dữ liệu đề xuất. Qua đánh giá thử nghiệm cho thấy độ chính xác của mô hình đề xuất lên đến 97,82%. Cho thấy tiềm năng của phương pháp học sâu trong việc nhận dạng loài cá. Kết quả của nghiên cứu này có thể được ứng dụng trong các lĩnh vực như nuôi trồng thủy sản, nghiên cứu sinh học và giám sát môi trường thủy sản.

Từ khóa: Phân loại cá, học sâu, CNN, ResNet152V2, Xception

ABSTRACT

Fish species identification plays an extremely important role in the assessment of fish species diversity in Vietnam, production management, and ecosystem monitoring, especially in the identification of endangered species. This makes fish species identification a necessity and requires high accuracy and reliability. However, identification that requires high accuracy and reliability often requires a lot of time and cost. In this paper, we propose a classification technique based on ResNet152V2 and Xception network architecture, we collect a data set of 820 images of fish species and conduct classification of fish species based on the data set. propose. Through testing and evaluation, the accuracy of the proposed model is up to 97.82%. Shows the potential of deep learning methods in fish species identification. The results of this study could have applications in areas such as aquaculture, biological research and aquatic environmental monitoring.

Keywords: Fish classification, deep learning, CNN, ResNet152V2, Xception

1. MỞ ĐẦU

1.1. Giới thiệu bài toán

Phát hiện, nhận dạng và phân loại cá tự động là những lĩnh vực nghiên cứu phổ biến và hấp dẫn. Nhiều nhà nghiên cứu đang tham gia vào phát triển nó cho môi trường dưới nước và ngoài nước. Đặc biệt ở Việt Nam một quốc gia có đa dạng địa hình và mạng lưới sông ngòi phong phú tạo ra sự đa dạng về hệ sinh thái cá ở Việt Nam với khoảng 2030 loài cá [9]. Trong những năm gần đây, những tiến bộ trong thị giác máy tính đã cung cấp một cách tiếp cận nhanh chóng và không phá hủy để phân loại cá. Học máy (ML) là một bước đột phá trong trí tuệ nhân tạo (AI) được ứng dụng trong nhiều lĩnh vực khác nhau như robot, y học, bảo mật thông tin và đặc biệt là nuôi trồng thủy sản. Công nghệ học máy được sử dụng rộng rãi trong nhiều ứng dụng nuôi trồng thủy sản, bao gồm phân loại, đếm cá, đo kích thước và xác định hành vi [4]. Học sâu (DL) là một tập hợp con của học máy và ngày càng được áp dụng nhiều hơn trong nuôi trồng thủy sản vì khả năng thể hiện tính năng hiệu quả của nó. DL là một mạng học tập nhiều lớp có thể trích xuất thông tin ngữ nghĩa từ cấp độ pixel, phù hợp để phát hiện hành vi của cá thông qua hình ảnh.

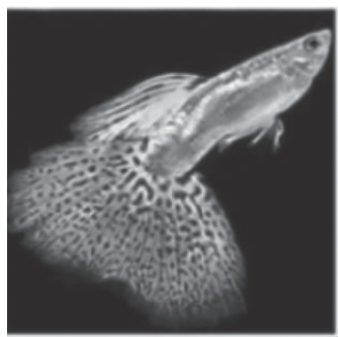
1.2. Những nghiên cứu liên quan

Nour Eldeen M. Khalifa và cộng sự [11] đã đề xuất được xây dựng bằng cách sử dụng mạng thần kinh tích chập sâu (CNN). Nó bao gồm bốn lớp, hai lớp tích chập và hai lớp được kết nối đầy đủ. Độ chính xác mà tác giả đạt được là 85,59%. Alaba SY và cộng sự. [3] đã đề xuất các mạng MobileNetV3-large và VGG16 và một đầu phát hiện SSD. Tác giả tập trung vào thuật toán của Lloyd để phân cụm và CNN để phân loại. Độ chính xác mà tác giả đạt được là 79,7% đã đạt được bằng các phương pháp

được đề xuất. Md Shoaib Ahmed và cộng sự [10] đã đề xuất mô hình dựa trên thuật toán học máy Support Vector Machine (SVM) với chức năng hạt nhân, hệ thống đề xuất đạt được độ chính xác 91,42%. Xiaoling Xu và cộng sự [5] đã đề xuất một phương pháp nhận dạng loài cá dựa trên SE-ResNet152 và tiêu cự cân bằng theo lớp. Phương pháp đề xuất đã được thử nghiệm trên bộ dữ liệu công khai Fish-Pak và đạt được độ chính xác là 91,25%. Ahmad Salman và cộng sự [1] đã đề xuất một phương pháp thống nhất để phát hiện cá di chuyển tự do trong môi trường dưới nước không bị hạn chế bằng cách sử dụng mạng thần kinh chuyển đổi dựa trên khu vực, một kỹ thuật học máy tiên tiến được sử dụng để giải quyết các vấn đề chung về phát hiện và định vị đối tượng. Tác giả sử dụng hai bộ dữ liệu điểm chuẩn được trích xuất từ kho lưu trữ video dưới nước Fish4Knowledge lớn, bộ dữ liệu cảnh phức tạp và bộ dữ liệu cá LifeCLEF 2015 để xác thực tính hiệu quả của phương pháp kết hợp. Đạt được độ chính xác phát hiện (F-Score) lần lượt là 87,44% và 80,02% trên 2 bộ dữ liệu này.

1.3. Đặc điểm của loài cá

Cá một loài thủy sản giàu protein và ít chất béo và là một trong những nguồn thực phẩm chủ yếu. Có hơn 20.000 loài cá trên thế giới, bao gồm khoảng 73 loài Cyclostomata, 800 loài Chondrichthyes và khoảng 20.000 loài Osteichthyes[2]. Trong các vùng nước tự nhiên hoặc trong nuôi trồng thủy sản, các loài cá khác nhau thường sống hoặc sinh sản trong cùng một khu vực, tạo nên sự đa dạng sinh học. Tuy nhiên, việc nhận dạng các loài cá này có thể rất khó khăn do sự tương đồng trong hình dạng và màu sắc. Điều này đòi hỏi sự nghiên cứu và phân tích cẩn thận để xác định chính xác loài cá một cách hiệu quả và đáng tin cậy.



Cá 7 màu



Cá 3 đuôi



Cá lau kiếng

Hình 1. Hình ảnh một số loài cá

1.4. Vai trò của học sâu trong nhận dạng loài cá

Học sâu là một lĩnh vực hiện đang phát triển, khám phá các lĩnh vực trí tuệ nhân tạo và học máy để học các tính năng trực tiếp từ dữ liệu, sử dụng nhiều lớp xử lý phi tuyến tính [6]. Học sâu với Mạng thần kinh chuyển đổi đã nổi lên như một trong những công cụ học máy mạnh mẽ nhất trong phân loại Hình ảnh, vượt qua độ chính xác của hầu hết tất cả các phương pháp phân loại truyền thống khác và thậm chí cả khả năng của con người [11, 19]. Quá trình tích chập có thể đơn giản hóa một hình ảnh chứa hàng triệu pixel thành một tập hợp các bản đồ đặc trưng nhỏ, do đó làm giảm kích thước của dữ liệu đầu vào trong khi vẫn giữ được các đặc trưng khác biệt quan trọng nhất [7].

1.5. Các mạng nơ-ron trích xuất đặc trưng

1.5.1. ResNet152V2

ResNet152V2 là một biến thể của ResNet (Residual Network) với 152 layers. ResNet được xem là một trong những kiến trúc quan trọng và hiệu quả nhất trong lĩnh vực học sâu. Một vấn đề thường gặp khi xây dựng các mạng sâu là hiện tượng vanishing gradient (điều khiển gradient mất dần). Điều này có nghĩa là khi thông tin đi qua nhiều lớp, gradient giảm đáng kể và dẫn đến khó khăn trong việc huấn luyện.

ResNet giải quyết vấn đề này bằng cách sử dụng các đường nối (shortcuts) hoặc “residual connections” để cho phép thông tin truyền thẳng qua các lớp. Các residual connections cho phép mạng học các biểu diễn sâu và phức tạp hơn mà không gặp vấn đề vanishing gradient. ResNet152V2 là phiên bản cải tiến của ResNet với 152 layers, sử dụng các kỹ thuật như batch normalization, skip connections và bottleneck architecture để tăng cường khả năng học của mạng.

1.5.2. Xception

Xception (Extreme Inception) là một kiến trúc mạng nơ-ron sâu dựa trên ý tưởng của Inception và có tên gọi là “Extreme” vì nó điều chỉnh cách tính toán trong Inception module. Inception là một kiến trúc mạng nổi tiếng được sử dụng rộng rãi trước đó. Điểm mạnh của Inception là khả năng nắm bắt thông tin ở các mức độ khác nhau bằng cách sử dụng các filter với kích thước khác nhau. Xception đưa ý tưởng này lên một tầm cao mới bằng cách thực hiện các phép tính convolution trực tiếp trên các kênh riêng biệt của dữ liệu đầu vào, thay vì trên toàn bộ không gian. Điều này giúp mạng học các biểu diễn tốt hơn và giảm đáng kể số lượng tham số cần học. Kiến trúc Xception tạo ra một mạng nơ-ron rất mạnh và hiệu quả trong việc nhận dạng hình ảnh.

1.6 Các độ đo đánh giá mô hình

Độ chính xác của các dự đoán tích cực (positive), tính tỉ lệ giữa số lần dự đoán tích cực đúng (true positive – TP) và tổng số dự đoán tích cực (True positive + False positive (FP); Độ bảo phủ (Recall): tỷ lệ phần trăm các trường hợp TP mà mô hình đã phát hiện được, tính bằng tỉ lệ giữa TP và false negatives (FN); Độ chính xác, tính tỉ lệ giữa số lần dự đoán đúng và tổng số dự đoán. Đây là độ đo phổ biến nhất và thường được sử dụng đối với các bài toán phân loại [8].

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

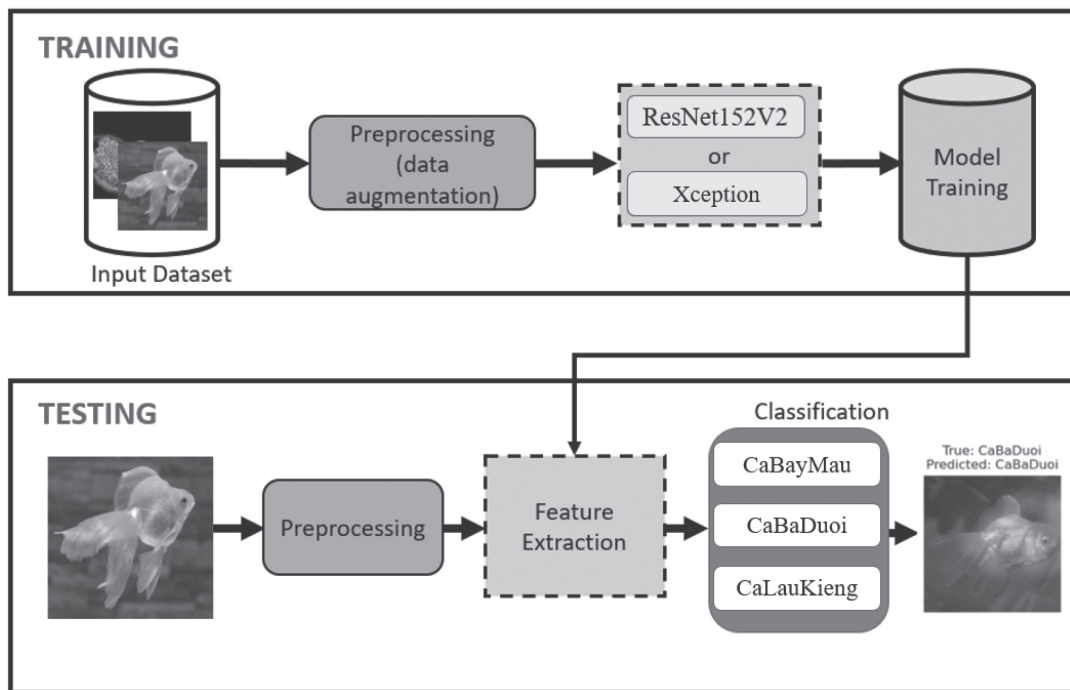
$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{Accuracy} = 2 \times \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

2. KẾT QUẢ NGHIÊN CỨU

2.1. Phương pháp đề xuất

Để giải quyết bài toán, chúng tôi đề xuất sử dụng mô hình tổng quát như **Hình 2**. Mô hình tổng quát gồm có hai giai đoạn: Giai đoạn huấn luyện và kiểm thử được mô tả cụ thể như hình sau:



Hình 2. Mô hình đề xuất phát hiện và phân loại các loài cá ở Việt Nam.

2.1.1. Đào tạo mô hình

Ở giai đoạn tiền xử lý, chúng tôi thực hiện tăng cường dữ liệu bằng cách xoay ảnh, lật ảnh ngẫu nhiên từ 0 đến 180 độ và thực hiện chuẩn hóa kích thước ảnh thành 512 x 512. Sau khi quá trình tiền xử lý dữ liệu, chúng tôi sẽ tiến hành rút trích

đặc trưng. Chúng tôi đề xuất phương pháp rút trích đặc trưng với hai mô hình mạng là ResNet152V2 và Xception như đã trình bày ở phần 2.1 trong bài báo này và sau đó thực hiện đào tạo dữ liệu theo mô hình.

2.1.2. Kiểm thử

Ở giai đoạn kiểm thử, dữ liệu đầu vào

sẽ là tập dữ liệu hình ảnh được mô tả trong phần 2.2 để nhận dạng và phân loại cá. Ở quá trình tiền xử lý, ảnh đầu vào sẽ được chuẩn hóa về kích thước thành 512 x 512, sau đó đưa qua các mô hình mạng được đề xuất từ giai đoạn thứ nhất để dự đoán.

Bảng 1. Các kịch bản được đề xuất và các tham số huấn luyện

Kịch bản	Kiến trúc mạng	Epochs	Learning rate	Input shape	Activation Function
1	ResNet152V2	25	0.001	512 x 512	Softmax
2	Xception	25	0.001	512 x 512	Softmax

2.2.2. Môi trường cài đặt và tập dữ liệu thực nghiệm

a. *Môi trường cài đặt:* Hệ thống được cài đặt bằng ngôn ngữ Python và chạy trên cùng một môi trường Kaggle (GPU P100). Thư viện hỗ trợ đào tạo mô hình mạng sử dụng là Tensorflow và Keras.

b. *Tập dữ liệu thực nghiệm:* Dữ liệu được chúng tôi thu thập từ một số loài cá phổ biến tại Việt Nam. Bộ dữ liệu này được chúng tôi thu thập từ năm 2023. Bộ dữ liệu bao gồm 820 hình ảnh với kích thước hình ảnh trung bình là 500*500 pixel. Các hình ảnh ở định dạng PNG và JPG. Các hình

2.2 Kết quả thực nghiệm

2.2.1. Các kịch bản áp dụng

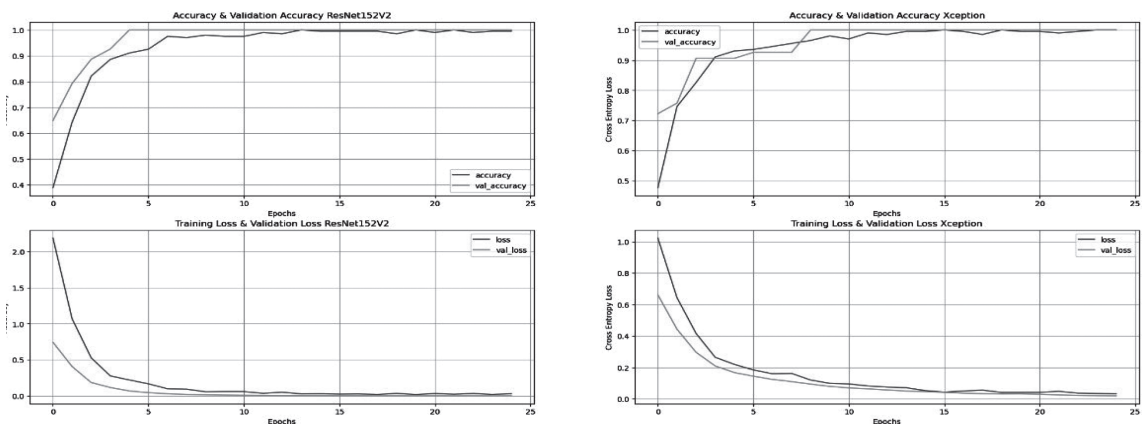
Để tiến hành thực nghiệm cho mô hình đề xuất, chúng tôi thực hiện hai kịch bản với các tham số huấn luyện như sau:

ảnh được phân loại thành ba loại: cá 7 màu, cá 3 đuôi và cá lau kiếng.

2.2.3. Các kết quả thực nghiệm

a. Độ đo Loss và Accuracy

Trong quá trình huấn luyện tôi chọn giá trị 512 x 512 cho kích thước ảnh đầu vào. Hình 3 thể hiện giá trị Loss và Accuracy của các kịch bản 1 và kịch bản 2. Độ chính xác của các kịch bản lần lượt là 95,12% và 97,82%. Kết quả thực nghiệm cho thấy kịch bản 2 có giá trị Loss thấp hơn và độ chính xác cao hơn trên tập dữ liệu đề xuất phù hợp cho nhận dạng các loài cá trong thực tế.

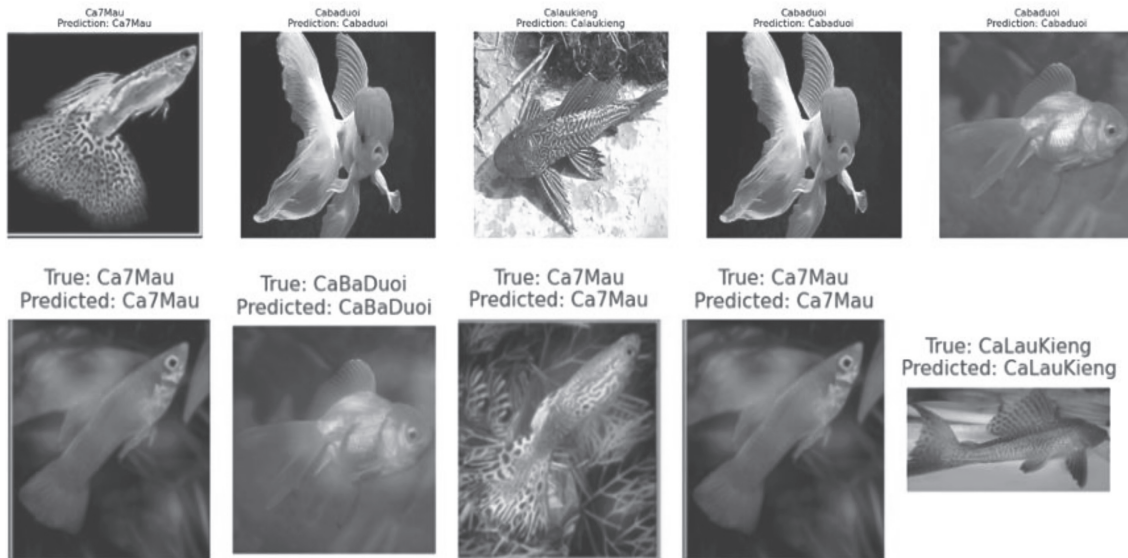


Hình 3. Biểu đồ về độ đo Accuracy và Loss của kịch bản 1 và 2

b. Một số kết quả thực nghiệm

Hình 4 trình bày một số kết quả thực nghiệm của quá trình nhận dạng và phân

loại một số loài cá tại Việt Nam dựa trên mô hình mạng đề xuất.



Hình 4. Một số kết quả thực nghiệm

Kết quả thực nghiệm cho thấy mô hình đạt độ chính xác cao trong nhận dạng 3 loại cá đề xuất. Ở các loài cá có màu sắc và kích thước khác nhau, mô hình kịch bản 2 cho kết quả nhận dạng chính xác trong các trường hợp khác nhau của tập dữ liệu kiểm thử.

3. KẾT LUẬN

Trong bài báo này, chúng tôi đã tiến hành nghiên cứu về việc nhận dạng loài cá tại Việt Nam bằng phương pháp học sâu. Việc nhận dạng loài cá tại Việt Nam mang ý nghĩa quan trọng trong việc xác định mức độ phong phú và quản lý hiệu quả các loài cá trong hệ thống đa dạng sinh học của Việt Nam. Chúng tôi đã sử dụng các kiến trúc

mạng học sâu bao gồm ResNet152V2 và Xception đồng thời xây dựng bộ dữ liệu để nhận dạng các loài cá dựa trên hình ảnh của chúng. Kết quả đánh giá thử nghiệm cho thấy mô hình Xception đạt được độ chính xác cao khoảng 97,82%. Điều này chứng tỏ tiềm năng lớn của phương pháp học sâu trong việc nhận dạng loài cá tại Việt Nam, đồng thời khẳng định tính chính xác và độ tin cậy cao của mô hình. Trong thời gian tới, chúng tôi tiếp tục thực hiện nhận dạng trên nhiều mô hình mạng khác nhau và nhiều loài cá phổ biến đồng thời cải tiến mô hình để nhận dạng bệnh trên các loài cá góp phần nâng cao sản lượng thủy sản tại Việt Nam.

TÀI LIỆU THAM KHẢO

- [1] Ahmad Salman, Shoaib Ahmad Siddiqui, Faisal Shafait, Ajmal Mian, Mark R Shortis, Khawar Khurshid, Adrian Ulges, Ulrich Schwanecke; Automatic fish detection in underwater videos by a deep neural network-based hybrid motion learning system. ICES Journal of Marine Science, Volume 77, Issue 4, July-August 2020, Pages 1295–1307,
- [2] Ashish Issac, Malay Kishore Dutta, Biplab Sarkar; Computer vision based method for quality and freshness check for fish from segmented gills. Computer vision based method for quality and freshness check for fish from segmented gills.

- [3] Alaba, S.Y.; Nabi, M.M.; Shah, C.; Prior, J.; Campbell, M.D.; Wallace, F.; Ball, J.E.; Moorhead, R. Class-Aware Fish Species Recognition Using Deep Learning for an Imbalanced Dataset. *Sensors* 2022, 22, 8268.
- [4] Bradley, D.; Merrifield, M.; Miller, K.M.; Lomonico, S.; Wilson, J.R.; Gleason, M.G. Opportunities to Improve Fisheries Management through Innovative Technology and Advanced Data Systems. *Fish Fish.* 2019, 20, 564–583. [Google Scholar] [CrossRef]
- [5] Xiaoling Xu, Wensheng Li, Qingling Duan; Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification. Version of Record 28 December 2020.
- [6] Deng, L., Yu, D. (2014). *Deep Learning: Methods and Applications*. Boston.
- [7] Hua, K.-L., et al. (2015). Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and therapy*.
- [8] Meysam Vakili, Mohammad Ghamsari and Masoumeh Rezaei (2020). Performance Analysis and Comparison of Machine and Deep Learning Algorithms for IoT Data Classification. arXiv:2001.09636 or arXiv:2001.09636v1.
- [9] Image processing based technique for classification of fish quality after cypermethrine exposure LWT - Food Sci. Technol. (2016)
- [10] Md Shoaib Ahmed, Tanjim Taharat Aurpa, Md. Abul Kalam Azad; Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture. Version of Record 9 August 2022.
- [11] Khalifa, N.E.M., Taha, M.H.N., Hassanien, A.E. (2019). Aquarium Family Fish Species Identification System Using Deep Neural Networks. In: Hassanien, A., Tolba, M., Shaalan, K., Azar, A. (eds) *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2018. AISI 2018. Advances in Intelligent Systems and Computing*, vol 845.