JURNAL PERIKANAN
UNIVERSITAS GADJAH MADA
Terakreditasi Ristekdikti No: 158/E/KPT/2021
ISSN: 2502-5066 (Online)
ISSN: 0853-6384 (Print)
Vol. 26 (1), 35-40
DOI 10.22146/jfs.85043

Machine Learning-Driven Seaweed Genera Identification on a Web Application Using Teachable Machine

Riza Nur Azhar*, Ervia Yudiati & Sri Sedjati
Department of Marine Science, Faculty of Fisheries and Marine Sciences, Diponegoro University, Semarang, Central Java, Indonesia
*Corresponding author, email: rizanur25@gmail.com
Submitted: 08 January 2023; **Revised:** 28 May 2024; **Accepted:** 03 June 2024; **Published:** 30 June 2024

ABSTRACT This study aims to evaluate the use of machine learning technology of Teachable Machine in identifying seaweed genus in Indonesia. The lack of database was the main driver of this study. The research focused on the Panjang Island with three stations. This field data are the foundation for seaweed identification using training machine learning models. In addition to field data, information from the literature on seaweed visual characteristics was also used to support the identification process. The results of machine learning model suggest 99.42% accuracy in identifying 13 classes of 9 seaweed genus. The implementation of the model on the web application showed satisfactory responsive performance, including in the speed test on Google PageSpeed. Overall, the integration of machine learning technology in a web application platform provides a practical solution for accurate and accessible seaweed identification. This invention has the potential in supporting research, conservation, and sustainable utilization of marine resources in Indonesia.

Keywords: Machine learning; seaweed identification; teachable machine

INTRODUCTION

Marine environment plays a pivotal role in the biosphere, housing diverse coastal ecosystems that vital for various life forms (Lobus & Kulikovskiy, 2023). Seaweed, a key macroalgae, contributes significantly to climate adaptation by mitigating wave energy, safeguarding coastlines, and enhancing water pH while providing oxygen, locally countering ocean acidification and deoxygenation (Duarte *et al.*, 2017). The Indonesia's coastlines support seaweed diversity and productivity, housing an estimated 555 known seaweed species worldwide (Sinurat *et al.*, 2023). However, despite the Indonesia's remarkable seaweed diversity, identification methods remain suboptimal due to limited skilled manpower and conventional identification techniques (Rimmer *et al.*, 2021). Inadequate identification impedes conservation efforts and optimal resource utilization. Accurate identification is crucial for biodiversity observation, environmental health assessment, and comprehensive scientific exploration of seaweed ecology and biology (Eggertsen & Halling, 2021).

The remarkable diversity of seaweed in Indonesia presents immense opportunities for scientific exploration, conservation, and sustainable utilization. Effective seaweed identification facilitates biodiversity observation, aids in comprehending marine environmental health, and fosters in-depth scientific research of seaweed ecology and biology (Eggertsen & Halling, 2021). Furthermore, accurate identification is pivotal in formulating precise conservation measures, sustainable marine resource management, and ecosystem restoration (Saleh & Sebastian, 2020). It also enables researchers to monitor climate change impacts and raise awareness about preserving marine and coastal environments.

Existing identification methods based on the morphology of seaweed suffer from environmental variations, impacting its accuracy (Madduppa, 2020). Molecular identification through DNA analysis provides more precise results but requires higher cost (Annisaqois *et al.*, 2018). Hence, the integration of machine learning emerges as a potential, cost-effective, and accurate identification method.

This study aims to apply machine learning, specifically Teachable Machines, for rapid and accurate seaweed genus identification. By developing a web application using TensorFlow's Graphical User Interface (GUI) Object Detection framework, this research pinpoint to contribute to early-stage identification for various applications. Challenges such as data diversity and understanding causal relationships between variables remained, as highlighted by Schölkopf *et al.* (2019).

MATERIALS AND METHODS

Materials

This study utilized a dataset comprising seaweed images obtained from field acquisition and library sources. The dataset served as the primary resource for training and testing the model. The tools employed in the research encompassed various software and hardware components, including a laptop for program execution, Web Browser, Teachable Machine, Virtual Studio Code, Microsoft Word, Microsoft. Excel, Github Pages, and Google PageSpeed. In the field Roll Meter used for demarcating the coastline during seaweed sample collection in the field.

Methods

Data acquisition and morphological identification

The seaweed image data collection involved processing the data into a format suitable for machine learning models, resulting in a high-quality training dataset. This process aimed to ensure the accurate identification of various seaweed types based on the collected dataset. The data collection method employed was stratified random

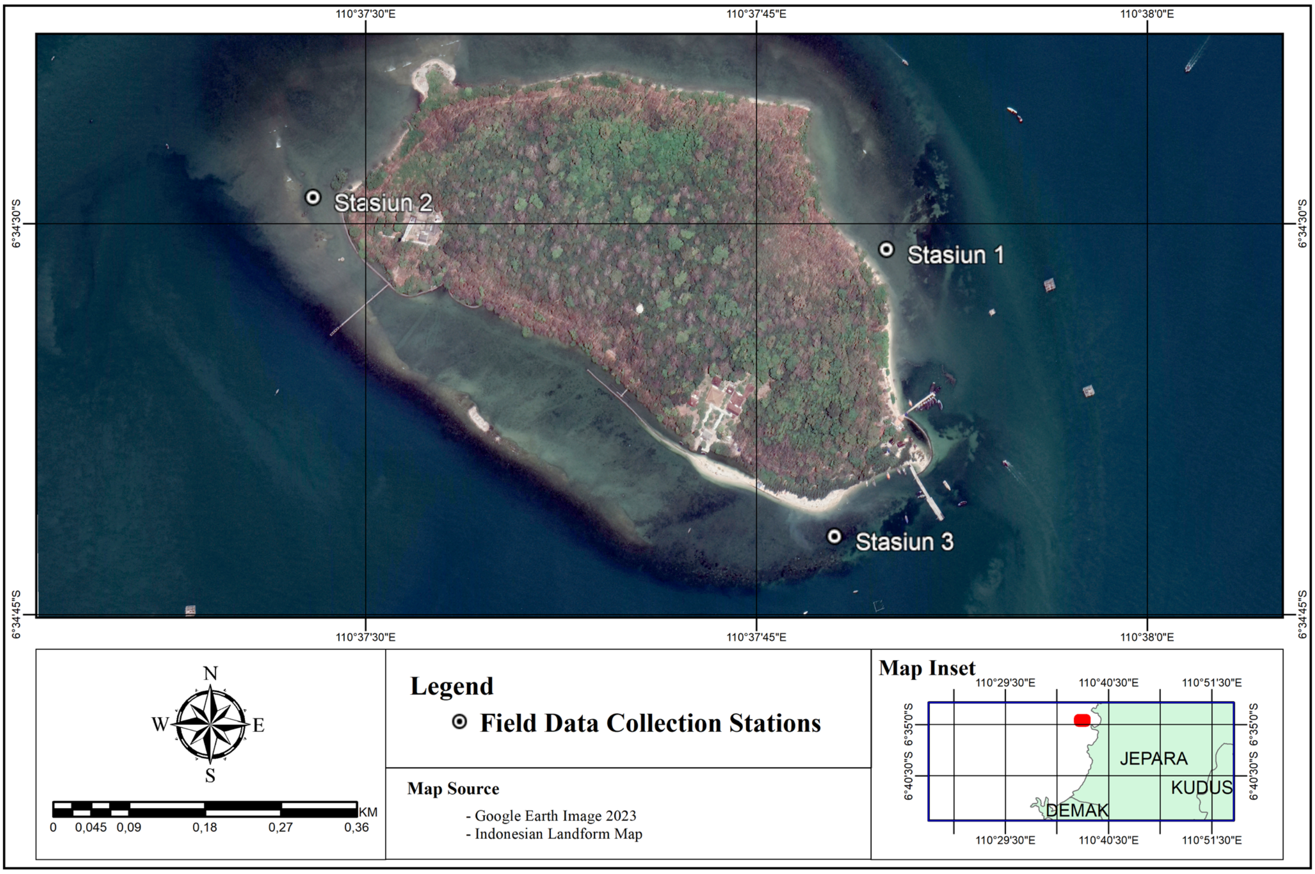

Figure 1. Map of the location of seaweed field image data collection on Panjang Island, Jepara Regency, Central Java conducted in October 2023.

sampling, dividing the seaweed population into strata based on morphological similarities and sampling from each stratum (Azora, 2021). Field data collection from Panjang Island's waters utilized three stations located at specific coordinates as seen in Figure 1. Using a 100-meter roll meter, three perpendicular lines were drawn toward the sea at 5-meter intervals to create sampling areas. Seaweed samples were taken at these stations, generating image datasets representing each sample. One dataset was allocated for model training, while the other, comprising ten images, was used for Teachable Machine's machine learning data accuracy test. This field data collection aimed to secure valid seaweed data from Panjang Island.

Machine learning model creation

The methodology employed in this study for seaweed identification using machine learning via the Teachable Machine platform. The model training process consisted of two key stages: model training and model conversion. Initially, seaweed images from field collection and library sources were uploaded to Teachable Machine, dividing the data into training and testing sets. The training set facilitated model training, while the testing set evaluated the model's accuracy. Teachable Machine parameters such as batch, epoch, and learning rate were utilized during training to teach the model to identify various seaweed types based on provided images. Post-training, the model was assessed using the testing set to gauge its seaweed identification capability.

Machine learning model evaluation

Subsequently, data evaluation ensured the developed model's accurate and consistent identification performance. Evaluation metrics like Confusion Matrix and Class-wise Accuracy measured the model's performance in identifying seaweed images in the test data, aiming to predict class labels for each input instance. Upon completing training and evaluation, the model was converted to a javascript format, uploaded to the Teachable Machine cloud, and transformed into a functional Application Programming Interface (API) for web application implementation (Pranata *et al.*, 2018).

Machine learning implementation on web application

The implementation of the seaweed identification model into a web application (WebApp) using HyperText Markup Language (HTML). This integration involved linking the Teachable Machine machine learning model with the user interface on a web browser, creating an interactive and user-friendly seaweed identification experience. The integration of the model into the backend using Virtual Studio Code and HTML, enabling users to capture seaweed images via their device cameras. The images are then processed by the model, delivering identification results based on the recognized seaweed type. Additionally, Cascading Style Sheets (CSS) were utilized to construct the user interface, facilitating image capture and seamless presentation of identification outcomes, enhancing user interaction intuitively. Following the design phase, HTML, CSS, and Javascript were uploaded to

Table 1. Morphological genus identification of sample gathered in Panjang Island, Jepara Regency.

No	Identification results	Number of class	Picture gathered	Identification guidelines
1.	*Halimeda* sp.	3	252	(Verbruggen *et al.*, 2004; Hillis *et al.*, 1998)
2.	*Dictyota* sp.	1	132	(Tronholm *et al.* 2010; Lamouroux *et al.*, 1809)
3.	*Padina* sp.	1	43	(Win *et al.*, 2022)
4.	*Sargassum* sp.	2	334	(Mattio *et al.*, 2013)
5.	*Halymenia* sp.	1	100	(Rodríguez-Prieto *et al.*, 2018; Schneider & Wyne, 2007)
6.	*Leathesia* sp.	1	134	(Ailen *et al.*, 2017; Gray, 1821)
7.	*Caulerpa* sp.	2	322	(Estrada *et al.*, 2020)
8.	*Acrosiphonia* sp.	1	143	(Burrows, 1991; Agardh, 1846)
9.	*Turbinaria* sp.	1	98	(Wynne, 2022)

GitHub for broader accessibility.

Web application evaluation

Evaluation of the WebApp's performance was conducted through Google Pagespeed, encompassing performance, accessibility, and Search Engine Optimization (SEO) assessments to ensure optimal functionality (Saputra *et al.*, 2018). Further enhancements and adjustments in response time, result display, memory management, and application performance were undertaken to refine user experience. Seaweed identification accuracy testing on the WebApp, involving image capture, model identification, and accurate display of identification results. The testing utilized a dataset acquired from field data and categorized based on morphological similarities, identifying the genus using the developed WebApp, recording the results, and confidence percentages. The accuracy of the machine learning identification was evaluated by comparing the field dataset with the integrated training model within the TensorFlow machine learning framework (Malahina *et al.*, 2022).

RESULTS AND DISCUSSION

Seaweed database morphological identification

Image data of seaweeds were identified based on their genus. Each genus was divided into different classes based on their morphological differences. This identification was done using identification guidelines from research journals as can be seen in Table 1. From this process, image data was obtained for 13 classes belonging to 9 genera. The number of image data and morphological identification guides are grouped by genus and class for each genus, as shown in Table 1.

Machine learning model training

The first step in developing the system into a WebApp is the development of a framework using Teachable Machine. The image data collected is then entered into Teachable Machine according to the class of each image data that will be processed into a convertible model. Model training is carried out using image data from 13 different classes consisting of 9 genera. The training model settings use the original settings, namely Epoch: 50, Batch size: 16 and Learning rate: 0.001. The preview model will appear on the right and is ready for trial before continuing to the advanced test. The settings of the training model depicted in Figure 2.

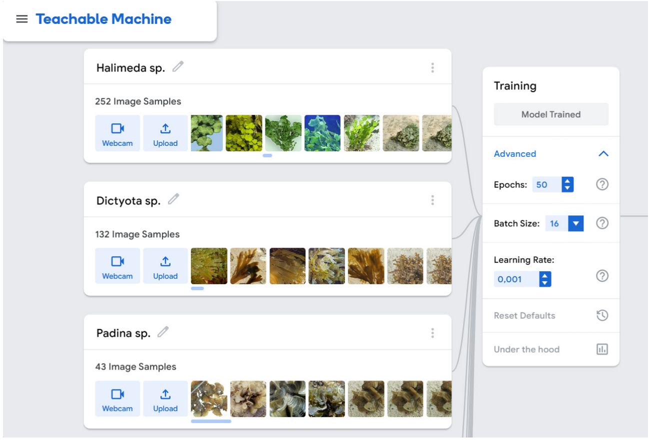

Figure 2. Teachable machine initial display and setup of model training.

Machine learning model evaluation

Accuracy testing per class and confusion matrix were carried out using the built-in test tool from Teachable Machine. A crosscheck system was used with the uploaded database to determine the accuracy of the image data at the model testing stage. The results of the accuracy per class are presented in Table 2.

Table 2. Database crosscheck accuracy test results using Teachable Machine.

No	Database classes	Accuracy	Number of sample
1.	*Halimeda* sp. (1)	1.00	6
2.	*Halimeda* sp. (2)	1.00	15
3.	*Halimeda* sp. (3)	1.00	18
4.	*Dictyota* sp. (1)	1.00	20
5.	*Padina* sp. (1)	1.00	7
6.	*Sargassum* sp. (1)	1.00	27
7.	*Sargassum* sp. (2)	1.00	24
8.	*Halymenia* sp. (1)	1.00	15
9.	*Leathesia* sp. (1)	1.00	21
10.	*Caulerpa* sp. (1)	1.00	23
11.	*Caulerpa* sp. (2)	1.00	26
12.	*Acrosiphonia* sp. (1)	1.00	22
13.	*Turbinaria* sp. (1)	1.00	15

The results for accuracy per class are shown in Table 2 indicating a high accuracy score of 1.00 on a scale of 0-1.00 for all tested classes, demonstrating the ability to accurately recognize each class. The crosscheck accuracy per class was illustrated through a confusion matrix, as depicted in Figure 3.

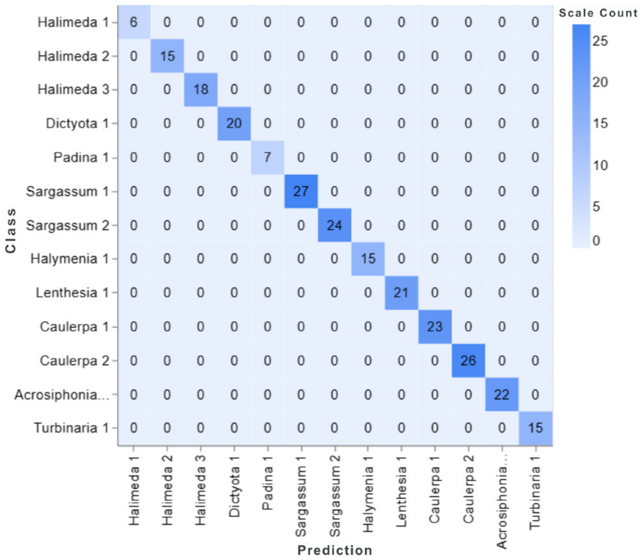

Figure 3. Database crosscheck result in confusion matrix database machine learning. The scale count represents the images being tested. The y-axis represents the actual classes of the tested images, while the x-axis shows the predicted classes, indicating the machine learning model's predictions.

The training dataset achieved perfect accuracy for all classes, as shown in Figure 3 The confusion matrix in Figure 3 indicates alignment between the y and x axes, signifying no misclassifications or identification errors within the database. These results are consistent with Baihaqi *et al.* (2022) which demonstrated similar results through analogous Teachable Machine testing. However, it is important to note that these results are crosschecks from the model database. Accuracy after implementation in web application or other applications may vary due to data quality, lighting, image capture angle, and camera types, which could potentially introduce distortions or alter the quality of the testing images (Dodge & Karam, 2016).

Model conversion and implementation

Databases that have undergone model training and achieved high accuracy rates in the per-class accuracy test and confusion matrix are uploaded to a Google API. The API is included in the javascript available after model conversion and can be implemented in HTML. The HTML can then be uploaded to cloud-based hosting using GitHub Pages to create a web application. The result of converting the model into javascript can be seen in Figure 4.

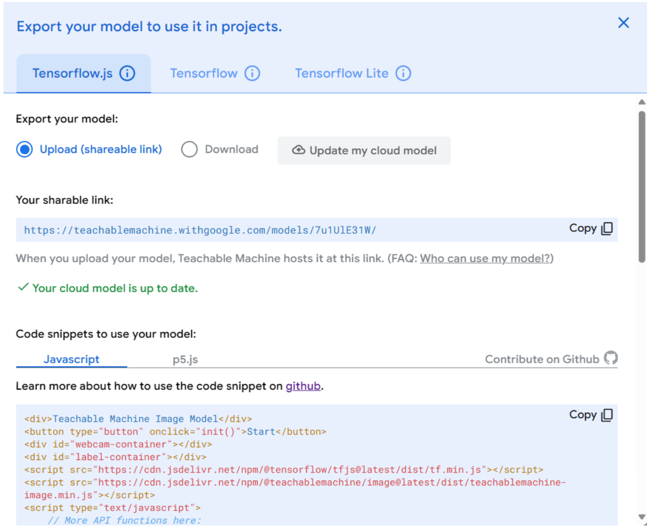

Figure 4. Model to javascript conversion on Teachablemachine.

HyperText Markup Language with javascript, which is then uploaded to the Github Pages repository so that the website can be accessed by devices that have a browser and a camera. The results of the implementation can be seen on a local computer, as shown in Figure 5.

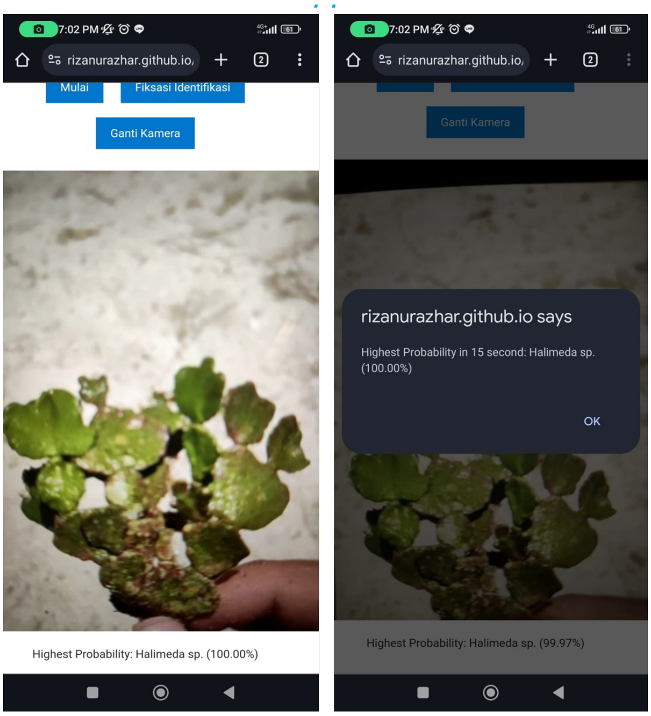

Figure 5. Web application display and identification pop up on seaweed genus identification result.

Web application performance test

Web application development is carried out by testing WebApp performance with Google PageSpeed which will produce a performance test. The results of the Google PageSpeed analysis are presented in Figure 6.

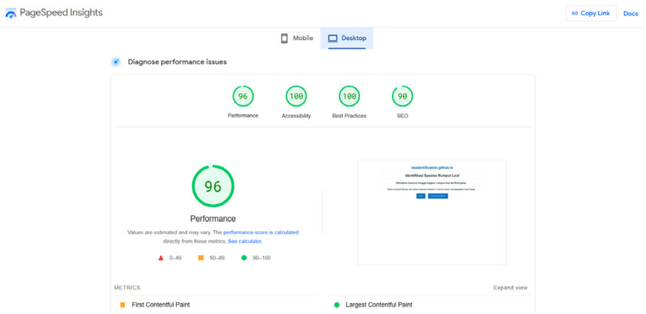

Figure 6. Google PageSpeed performance test result.

The comprehensive Google PageSpeed evaluation revealed the site's commendable performance, as shown in Figure 6. This evaluation included several key components scored on a scale of 0-100. Performance scored 96, indicating exceptional functionality, which may be influenced by content weight, server responsiveness, or JavaScript rendering. Accessibility scored a perfect 100, indicating universal accessibility, even for users who rely on assistive technologies. Best Practices also scored 100, reflecting adherence to modern web standards, including secure HTTPS usage and efficient script coding. SEO (Search Engine Optimization) scored 90, demonstrating strong optimization for search engines, including factors such as meta descriptions and proper image annotation. Taken together, these results demonstrate a well-designed site in terms of performance, accessibility, best practices, and SEO considerations. However, there's room for further refinement and improvement, as identified by Google PageSpeed Insights, which may indicate potential areas for future research or development.

The results of the web application accuracy test of 10 test image data from each class totaling 13 classes that have not been identified produce data with varying confidence levels for each class. The identification results can be seen in Table 3.

Table 3. Identification accuracy results using machine learning integrated in the web application on a scale of 0-100.

No	Database classes	Number of samples	Accuracy (%)
1	*Halimeda* sp. (1)	38	99.91
2	*Halimeda* sp. (2)	97	99.93
3	*Halimeda* sp. (3)	117	99.80
4	*Dictyota* sp. (1)	132	98.97
5	*Padina* sp. (1)	43	99.62
6	*Sargassums* sp. (1)	178	98.95
7	*Sargassum* sp. (2)	156	99.33
8	*Halymenia* sp. (1)	100	99.89
9	*Leathesia* sp. (1)	134	99.97
10	*Caulerpa* sp. (1)	150	98.71
11	*Caulerpa* sp. (2)	172	98.91
12	*Acrosiphonia* sp. (1)	143	99.32
13	*Turbinaria* (1)	98	99.19
Average accuracy			99.42

The accuracy tests performed for the integrated machine learning in the developed web application are shown in Table 3, which shows 13 database classes tested with ten image samples, resulting in an average accuracy of 99.42%. This rate, although not a universal standard, is considered good. The accuracy achieved may vary depending on the specific application, the consequences of misclassification errors, and the state of the art in the respective field (Rudzicz *et al*., 2019). For example, in medical imaging, machine learning-based skin cancer classification studies achieved classification accuracies of 92.1% (Javaid *et al*., 2021). Similar studies in agriculture reported weed detection accuracies of 96% and 94% using different machine learning algorithms (Islam *et al*., 2021). Other image classification studies using deep learning reported test accuracies of 100% for specific flowers and 99.30% for rose classification (Pratiwi *et al*., 2021). These examples highlight that acceptable machine learning image identification accuracy depends on the context, algorithms, machine learning techniques, and data used.

The results of the seaweed identification web application identification accuracy testing, as shown in Table 3, align with the crosscheck database, indicating high-quality outcomes. These results were achieved due to significant differences of the morphological properties among the seaweed genera. Similar observations were made by Malahina *et al*. (2022), suggesting that the accuracy of identification is notably affected by the similarity of data among classes. Other factors that can affect accuracy include lighting, image capture angle, and image quality. To facilitate identification and achieve accuracy ranging from 90% to 100%, it is recommended to have clear imaging of different parts of the seaweed with contrasting backgrounds.

An accuracy rate averaging 99.42% is considered good, considering the complexity of similar seaweed shapes and colors within the same genus and variations in the seaweed dataset used (Rudzicz *et al*., 2019). Comparatively, Malahina *et al*., (2022) achieved an average accuracy of 91.8% in real-time student face identification using Teachable Machine. This rate might be influenced by suboptimal backgrounds for facial identification. Another study using a YOLO (You Only Look Once) deep learning algorithm for reef fish identification reached an 82.82% accuracy, possibly influenced by the focus on fish identification and different machine learning algorithms (Yusup *et al*., 2020). Additionally, Chazar & Rafsanjani (2022) attained 100% accuracy in identifying varied seedlings, highlighting the significant impact of the research object and utilized algorithms on identification accuracy. Nevertheless, the seaweed identification research achieved remarkably high accuracy, signifying the substantial potential of Machine Learning in seaweed identification.

CONCLUSION AND RECOMMENDATION

Conclusion

The study on seaweed identification using machine learning through Teachable Machine, implemented in a web application, concludes that the successful integration of the machine learning model into a web application connected to the Google API enables practical and accessible seaweed identification. Functional testing revealed commendable responsiveness and performance, as evidenced by the positive results obtained from Google PageSpeed assessments. Field data collection yielded 13 classes of seaweed from 9 genera and 13 species. The machine learning model was tested and demonstrated high accuracy levels, 99.42%. The confusion matrix confirmed precise identification without significant errors during crosscheck database testing.

Recommendation

Collaborative research involving teams from various fields, such as computer science, biology, and user interface design, can be conducted to gain broader and more comprehensive insights. In this application, identification is limited to the genus level because species identification requires molecular analysis, given the similar morphological traits between species and high level of morphological plasticity in seaweed. Further research using molecular identification techniques is necessary to validate species identification more thoroughly and increase accuracy.

AUTHORS' CONTRIBUTIONS

RNA is doing research ideas, data generation, sample image collection and data analysis, EY is written manuscript, SS is doing manuscript preparation and translation.

ACKNOWLEDGEMENT

The author expresses gratitude to Dr. Ir. Ervia Yudiati, M.Sc. and Dr. Ir. Sri Sedjati, M.Si. for their invaluable guidance and support throughout the course of this research. The author also thanks the Faculty of Fisheries and Marine Science at Diponegoro University, and the Coastal community at Teluk Awur and Panjang Island Jepara.

REFERENCES

Annisaqois, M., G. Gerung, S. Wullur, B. Sumilat, W. Wagey & S. Mandagi. 2018. Analisis molekuler DNA alga merah (*Rhodophyta*) *Kappaphycus sp*. J. Pesisir dan Laut Tropis. 6 (1): 107-112. https://doi.org/10.35800/jplt.6.1.2018.20589

Azora, P. 2021. Analisis quick count dengan menggunakan metode stratified random sampling studi kasus pemilu gubernur Kalimantan Barat 2018. Bimaster. 10 (1): 43–50.

Baihaqi, M.B., Y. Litanianda & A. Triyanto, 2022. Implementasi tensor flow lite pada teachable untuk identifikasi tanaman aglonema berbasis android. KOMPUTEK. 6 (1): 70-80.

Chazar, C & M.H. Rafsanjani. 2022. Penerapan Teachable Machine Pada Klasifikasi Machine Learning Untuk Identifikasi Bibit Tanaman. INOTEK. 2 (1): 32-40.

Dodge, S & L. Karam. 2016. Understanding how image quality affects deep neural networks. QoMEX, 1-6.

Duarte, C. M., J. Wu , X. Xiao, A. Bruhn & D. Krause-Jensen. 2017. Can seaweed farming play a role in climate change mitigation and adaptation?. Front. Mar. Sci. 4 (100): 1-8.

Eggertsen, M & C. Halling. 2021. Knowledge gaps and management recommendations for future paths of sustainable seaweed farming in the Western Indian Ocean. Ambio. 50 (1): 60-73.

Islam, N., M.M. Rashid, S. Wibowo, C.Y. Xu, A. Morshed, SA. Wasimi & S.M. Rahman. 2021. Early weed detection using image processing and machine learning techniques in an Australian chilli farm. Agriculture. 11 (5): 387.

Javaid, A., M. Sadiq & F. Akram. 2021. Skin cancer classification using image processing and machine learning. IBCAST. 439-444.

Lobus, N.V & M.S. Kulikovskiy. 2023. The Co-Evolution Aspects of the Biogeochemical Role of Phytoplankton in Aquatic Ecosystems: A Review. Biology. 12 (1): 92.

Madduppa, H. 2020. Perbandingan hasil metode identifikasi spesies: Morfologi dan molekuler pada ikan julung-julung di TPI (Tempat Pelelangan Ikan) Muara Angke, DKI Jakarta. Jurnal Kelautan: Indonesian Journal of Marine Science and Technology. 13 (3): 168-175.

Malahina, E.A.U., R.P. Hadjon & F.Y. Bisilisin. 2022. Teachable Machine: Real-Time Attendance of Students Based on Open Source System. IJICS. 6 (3): 140-146.

Pranata, B.A., A. Hijriani & A. Junaidi. 2018. Perancangan Application Programming Interface (API) Berbasis Web Menggunakan Gaya Arsitektur Representational State Transfer (Rest) Untuk Pengembangan Sistem Informasi Administrasi Pasien Klinik Perawatan Kulit. Jurnal Komputasi. 6 (1): 33-89.

Pratiwi, H.A., M. Cahyanti & M. Lamsani. 2021. Implementasi Deep Learning Flower Scanner Menggunakan Metode Convolutional Neural Network. Sebatik. 25 (1): 124-130.

Rimmer, M.A., S. Larson, I. Lapong, A.H. Purnomo, P.R. Pong-Masak, L. Swanepoel & N.A. Paul. 2021. Seaweed aquaculture in Indonesia contributes to social and economic aspects of livelihoods and community wellbeing. Sustainability. 13 (19): 10946. https://doi.org/10.3390/su131910946

Rudzicz, F., P.A. Paprica & M. Janczarski. 2019. Towards international standards for evaluating machine learning. In SafeAI@ AAAI. 2301 (10).

Saleh, H & E. Sebastian. 2020. Seaweed Nation: Indonesia's New Growth Sector; Backgrounder, No. 02/2020. VIC. Australia–Indonesia Centre: Caulfield East, Australia. 18 pp.

Saputra, D.D., I. Ilhamsyah & D. Prawira. 2020. Implementasi Framework Accelerated Mobile Pages Pada Pengembangan Website Program Studi Sistem Informasi. Jurnal Komputer dan Aplikasi. 8 (2): 67-78.

Schölkopf, B., F. Locatello, S. Bauer, N.R. Ke, N. Kalchbrenner, A. Goyal & Y. Bengio. 2021. Toward causal representation learning. IEEE. 109 (5): 612-634.

Sinurat, E., D. Fransiska, B.S.B. Utomo, S. Subaryono & N. Nurhayati. 2023. Characteristics of Powder Agar Extracted from Different Seaweeds Species and Locations. J Appl Phycol.

Yusup, I.M., M. Iqbal & I. Jaya. 2020. Real-time reef fishes identification using deep learning. IOP Conference Series: Earth and Environmental Science. 429 (1): 012046.

How to cite this article Azhar, R.N., E. Yudiati & S. Sedjati. 2024. Machine learning-driven seaweed genera identification on a web application using teachable machine. Jurnal Perikanan Universitas Gadjah Mada. 26 (1): 35-40 https://doi.org/10.22146/jfs.92955