Jurnal Ilmu Kelautan **P-ISSN: 2460-0156**
SPERMONDE (2026) 12(2): 79-88 **E-ISSN: 2614-5049**

SPATIAL ANALYSIS AND POPULATION DENSITY OF GREEN TURTLE (*Chelonia mydas*) IN INDONESIAN WATERS: A SPECIES DISTRIBUTION MODELING APPROACH

Raja Banu Hendriawan[1], Ahmad Rafael Gerald Yosiano[1], Muhammad Ali Husen[1]

Submitted: June 10, 2026 Accepted: September 2, 2026

[1]Student of Marine Information Systems Department, Serang Regional Campus, Indonesian University of Education.
Corresponding Author:
[*]Raja Banu Hendriawan
E-mail: rajabanuhendriawan@upi.edu

ABSTRACT

Indonesia is one of the world's centers of marine biodiversity, serving as an important habitat for the Green turtle (*Chelonia mydas*). Various anthropogenic pressures and environmental changes have led to a decline in habitat quality and turtle populations, necessitating large-scale spatial analysis to support conservation efforts. This study aims to analyze the spatial distribution patterns and population density of Green turtles in Indonesian waters using occurrence data obtained from the Ocean Biodiversity Information System (OBIS). The study applies a combination of Kernel Density Estimation (KDE) and Convolutional Neural Network (CNN) methods to model species distribution and identify potential habitats. Preprocessing stages include removing duplicate data, handling empty values, validating geographic coordinates, and spatial filtering within Indonesian waters. KDE is used to generate continuous density maps and identify population hotspots, while a 1° × 1° spatial grid is used as input for the CNN model. The CNN model is trained using the KDE-derived spatial attributes to classify potential and less potential habitats. The results show that the distribution of Green turtles is uneven and concentrated in several key hotspots, particularly in the Nusa Tenggara, Maluku, Banda Sea, and Papua regions. The CNN model is capable of recognizing spatial habitat patterns well and generating potential habitat predictions consistent with hotspots identified through the KDE. The integration of KDE and CNN has proven effective in mapping species distribution and identifying important Green turtle habitats at a national scale. The results of this study can be used as a scientific basis for planning marine conservation areas, protecting critical habitats, and managing sustainable marine resources in Indonesia.

Keywords: Green turtle, *Chelonia mydas*, Species Distribution Modeling, Kernel Density Estimation, Convolutional Neural Network, Spatial Analysis.

INTRODUCTION

Indonesia is one of the world's centers of marine biodiversity and a primary habitat for the green turtle (*Chelonia mydas*), a species that plays a crucial ecological role in maintaining the balance of coastal and marine ecosystems. Green turtles contribute to the health of seagrass beds and coral reefs through foraging and nesting activities, and serve as indicators of marine environmental health. Globally, green turtles are listed as vulnerable by the IUCN, and in Indonesia they are legally protected because their populations continue to face pressure from human activities and environmental change (Tapilatu et al., 2022).

Green turtle populations in Indonesia face several serious threats that directly impact their spatial distribution and density. Coastal land conversion for development, coastal erosion, loss of natural vegetation, and plastic and light pollution are the main factors reducing the quality of their nesting and feeding habitats (Elfidasari et al., 2022). Anthropogenic pressures such as hunting, fisheries bycatch, and unmanaged tourism also exacerbate population declines. Studies at various locations have shown significant declines in nest numbers and nesting success due to human disturbance and environmental changes (Mujiyanto et al., 2026).

Amidst the complexity of these threats, spatial analysis and species distribution modeling are key approaches to understanding distribution patterns, identifying critical habitats, and supporting evidence-based conservation efforts. The integration of satellite data, UAV surveys, and advanced statistical models such as MaxEnt and GAM enables mapping of potential habitats, predicting distribution changes due to environmental pressures, and evaluating the effectiveness of Marine Protected Areas (MPAs) (Naro-Maciel et al., 2017).

In the context of tropical marine conservation management, understanding spatial distribution patterns and population density levels is fundamental information that is absolutely necessary for formulating protected area zoning policies (Mazaris et al., 2017). Through macro-scale spatial mapping, areas that are centers of turtle activity, both as foraging grounds and critical migration corridors, can be identified. However, the

main challenges to marine megafauna research in Indonesia have been the limited availability of continuous field data and logistical constraints, resulting in most previous studies being local and spatially limited to just one or two nesting beaches.

Advances in computational data science technology and information transparency now provide an alternative solution to overcome these data limitations through the use of global biodiversity repositories, such as the Ocean Biodiversity Information System (OBIS). OBIS provides integrated marine biodiversity occurrence records that have been scientifically validated by researchers worldwide (Appeltans et al., 2024). Large-scale geographic coordinate data from OBIS can be extracted and processed quantitatively using geographic information system-based methods, one of which is Kernel Density Estimation (KDE). The KDE method has proven reliable in modeling the intensity of random point distributions (point patterns) into a visual representation of spatially continuous density areas (hotspots) (Tapilatu et al., 2022).

Given the importance of protecting endangered species in Indonesian waters and the potential use of computationally based secondary data, this study was conducted to fill the gap in large-scale spatial information. This study aims to analyze spatial distribution patterns at a macro level, map population concentration points (hotspots), and identify critical distribution areas of green turtles (*Chelonia mydas*) throughout Indonesian waters based on OBIS data. The results of this distribution modeling are expected to provide objective scientific contributions and serve as recommendations in strengthening the zoning strategy for sustainable Marine Protected Areas (MPAs) in Indonesia, in line with the spirit of marine space management which is the main focus of marine science studies.

MATERIALS AND METHODS

Study Area

The research was conducted in Indonesian waters located at astronomical coordinates 95° – 141° East Longitude and 11° South Latitude – 6° North Latitude. This study area encompasses the dynamics of various complex tropical marine ecosystems that serve as critical habitats for Green turtle (*Chelonia mydas*), including seagrass beds, coral reefs, and coastal zones. Indonesian waters have very high ecological significance at the global level due to their position within the Coral Triangle region, which acts as a major migration corridor and foraging ground for various marine megafauna species (Block et al., 2011).



Figure 1. Map of Indonesian Waters (Research Geographical Boundaries)

Determining and mapping the boundaries of the study area precisely is a crucial step in limiting the scope of spatial analysis, so that the results of the ecological interpretation obtained remain within valid biogeographic boundaries (Read et al., 2009). A comprehensive computational analysis was carried out on all recorded points of Green turtle emergence within the territorial boundaries of these waters after going through a spatial filtering stage.

Data Acquisition and Preprocessing

A secondary dataset recording large-scale green turtle (*Chelonia mydas*) occurrences was obtained through data mining on the global Ocean Biodiversity Information System (OBIS) platform in CSV format. Processing large-scale datasets sourced from public repositories requires a structured and rigorous cleaning process to evaluate quality and address limitations and biases before integration into a distribution mapping

model (Hughes et al., 2024). The main data attributes extracted include geographic coordinates of latitude, longitude, observation date (event date), specimen identity (organism ID), and number of individuals (individual count) which are aligned to the universal coordinate system World Geodetic System 1984.

The data preprocessing stage is carried out through several stages, including removing duplicate data, eliminating missing values, and filtering latitude-longitude boundaries. Next, a land-sea masking process is performed using Indonesian administrative boundaries as a reference to separate and remove anomalous observation points recorded on land. This step is essential in marine wildlife ecology studies, eliminating bias due to coordinate recording errors and ensuring a true representation of ocean movements (Sequeira et al., 2018). In the final stage of cleaning, all numeric variables are transformed using the Min-Max Scaling method into the range 0 to 1. Standardizing the data scale through this normalization feature has been empirically proven to accelerate the convergence of machine learning algorithms when processing numeric features (Aksoy & Haralick, 2001).

Spatial Analysis and Grid-Based Input Formulation

Initial spatial analysis was applied to transform discrete point data into continuous spatial representation using the Kernel Density Estimation (KDE) method. KDE is a non-parametric statistical approach used to estimate the probability density function of a population based on the distribution of sample points in the field (Silverman, 2002). Kernel density calculations in two-dimensional space use Gaussian basis functions to estimate density based on the total selected observation data, smoothing parameters (bandwidth), Gaussian kernel functions, and the coordinates of the sample points.

To create an input data format suitable for artificial intelligence models, Indonesian waters are divided into a two-dimensional spatial grid structure (2D Grid-Map) with a grid resolution of 1°×1°. Each grid unit represents a habitat unit, storing a density index value based on the accumulated average KDE value within its coverage area. This transformation from discrete point coordinates to a grid matrix structure aims to transform conventional spatial data into structured images ready for extraction by the convolutional layer in a deep learning model.

Grids with high KDE index values are classified as potential habitat areas (labeled binary 1), while grids with low KDE index values are categorized as less potential habitat (labeled binary 0). This grid-based discrete class division is the main basis for preparing labeled data for the purposes of modeling distribution estimates in artificial intelligence and managing uncertainty in conservation decision making (Sinclair et al., 2010).

Species Distribution Modeling Using Convolutional Neural Network

The green turtle spatial distribution estimation model was built using a deep learning approach based on a Convolutional Neural Network (CNN) architecture. This convolutional neural network was implemented due to its superior performance in extracting spatial features, recognizing visual hierarchies, and detecting geometric representation patterns in complex environmental matrices (Brodrick et al., 2019). The input variables fed into the network are grid image matrices containing latitude, longitude, and density index values calculated from the KDE. The output variable is a predicted probability of a binary habitat classification (potential vs. less potential).

The CNN architecture designed in this study consists of an input layer, a convolutional layer that extracts local spatial features from density images, a max pooling layer that reduces the matrix dimensionality for computational efficiency, a flatten layer that converts the matrix into a one-dimensional vector, a fully connected layer (dense), and an output layer. The final layer uses a sigmoid activation function to generate binary classification probability values, while the hidden layers utilize Rectified Linear Unit (ReLU) activation.

To ensure the system's ability to automatically recognize patterns, the architecture is also configured to adapt to the spatial patterns of marine megafauna. This spatial dataset was randomly divided, with 80% allocated as training data and 20% as testing data using a hold-out validation method.

Throughout the learning iteration process, the network weights are optimized using the Adam Optimizer algorithm with a binary cross-entropy loss function to minimize the prediction error rate (Kingma & Ba, 2015). The performance and validity of the CNN model are then evaluated using a series of standard evaluation metrics to measure the model's generalization ability to new data and to ensure the model avoids

the phenomenon of overfitting (Hinton, 2014). The evaluation metrics include testing the Precision, Recall, F1-Score, Confusion Matrix, and overall accuracy levels (Accuracy).

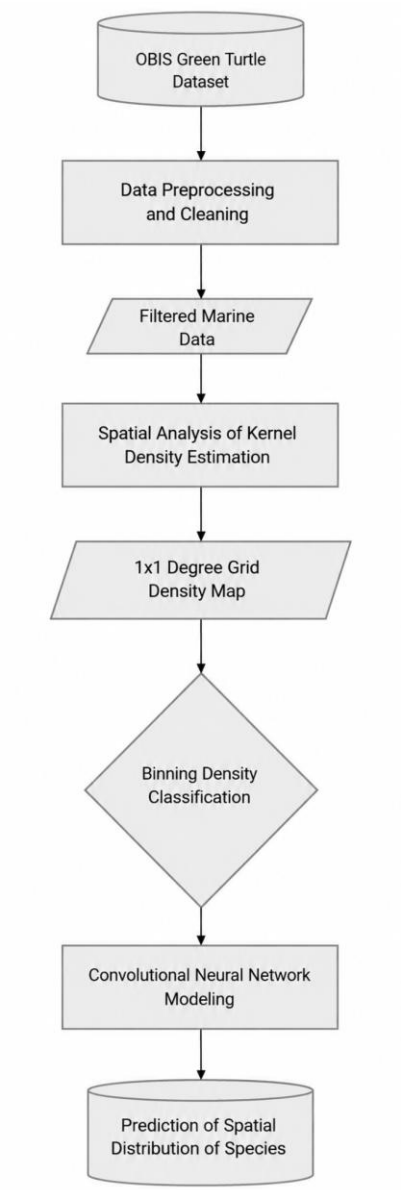


Figure 2. Flowchart of CNN-Based Green turtle Spatial Distribution Modeling Methodology.

RESULTS AND DISCUSSION

Characteristics of Green turtle Occurrence Data

In the initial stage, data cleaning and validation of green turtle (*Chelonia mydas*) occurrences were conducted, obtained from the Ocean Biodiversity Information System (OBIS). This process included removing duplicate data, blank values, and coordinates outside Indonesian territorial boundaries. After the spatial validation process, a number of observation points were obtained that represent the distribution of green turtles across various Indonesian waters. The data distribution indicates that green turtle occurrences are not evenly distributed, but rather concentrated in certain areas known to have habitat characteristics that support foraging and migration activities.

Table 1. Green turtle Dataset Statistics

Parameter	Value
Total Data	791.0000
Minimum Latitude	-10.9833
Maximum Latitude	-6.2146
Minimum Longitude	105.6000
Maximum Longitude	132.8167
Average Latitude	-10.0995
Average Longitude	120.3031

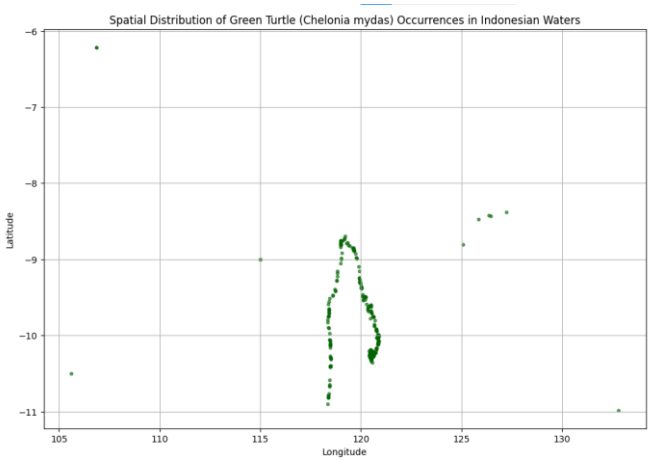


Figure 3. Distribution of green turtle emergence points in Indonesia

Based on the distribution map, observation points tend to be concentrated in eastern Indonesia, particularly the Nusa Tenggara region, the Sawu Sea, Maluku, and parts of West Papua. This pattern indicates that these areas have suitable oceanographic conditions to support green turtle activity.

Spatial Density Analysis Using KDE

Spatial density analysis using Kernel Density Estimation (KDE) produces a continuous density surface that describes the concentration level of Green turtle occurrences in the research area.

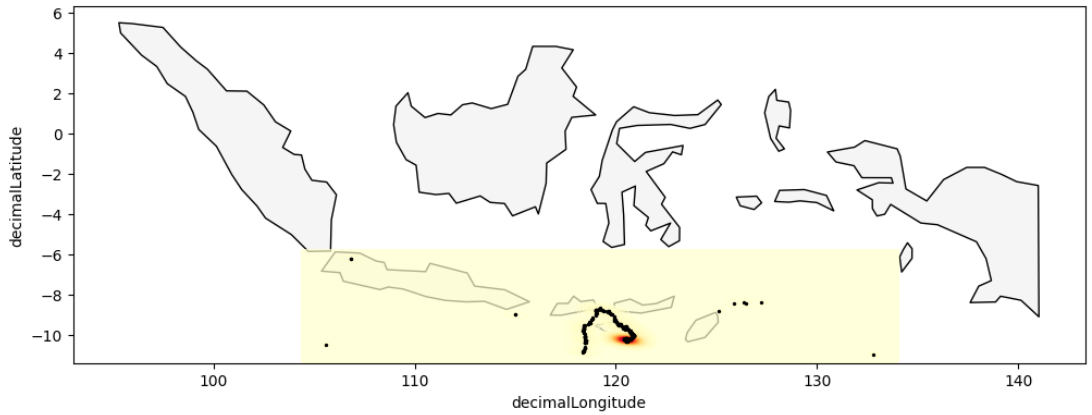


Figure 4. Kernel Density Estimation (KDE) Map

The KDE results indicate several key hotspots with higher densities than other areas. Areas with the highest densities are generally located in waters adjacent to seagrass beds and coral reef ecosystems.

High densities indicate that these areas have the potential to be important habitats for green turtles due to their abundant food sources and relatively stable environmental conditions.

Table 2. Density Classification Based on KDE Values

Parameter	Value
Minimum Density	0.003633
Maximum Density	2.162854
Mean Density	1.646825
Median Density	2.151897
Standard Deviation	0.870535

This finding is in line with previous research which stated that the distribution of green turtles is greatly influenced by the availability of seagrass beds and the quality of coastal habitats.

Grid-Based Habitat Representation

The KDE results are then transformed into a 1° × 1° spatial grid to produce a more structured habitat representation that is compatible with the needs of the CNN model.

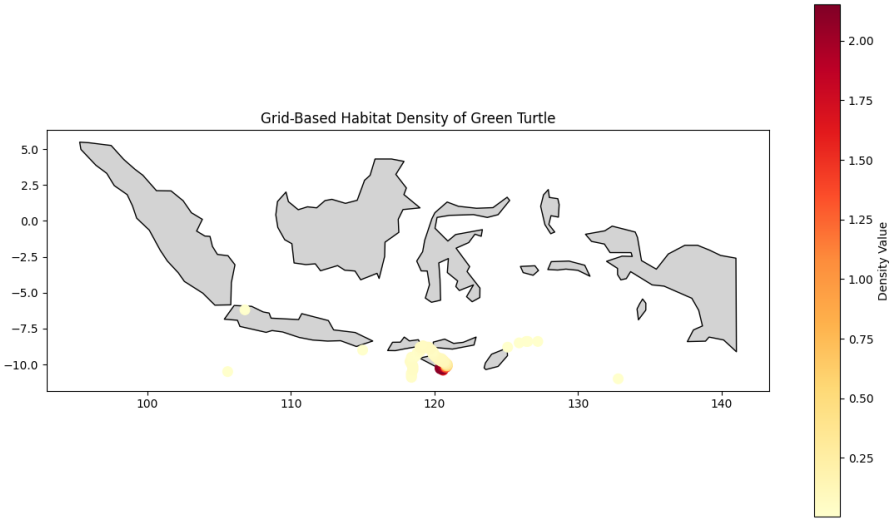


Figure 5. Green turtle Habitat Density Grid Map

Table 3. Ten Grids with the Highest Density Values

lat_grid	lon_grid	kde_density
-10.3	120.5	2.152605
-10.3	120.6	2.128150
-10.2	120.5	2.125915
-10.3	120.4	2.080826
-10.2	120.4	2.067159
-10.2	120.6	2.010163
-10.2	120.7	1.858567
-10.4	120.6	1.683094
-10.2	120.8	1.339193
-10.1	120.8	0.764294

Grid-based representation allows for a more systematic identification of potential habitats than using individual observation points. Grids with high KDE values indicate areas with consistently higher concentrations of Green turtle occurrences.

Species Distribution Modeling Using CNN

The CNN model was built using gridded data labeled by habitat density category. The training process was conducted using 80% training data and 20% testing data.

Table 4. CNN Parameters

Parameter	Value
Input Features	3
Training Data	80%
Testing Data	20%
Epoch	50
Batch Size	16
Optimizer	Adam
Loss Function	Binary Crossentropy
Activation Hidden Layer	ReLU
Activation Output Layer	Sigmoid

Model Performance Evaluation

The evaluation was performed using the Accuracy, Precision, Recall, and F1-Score metrics.

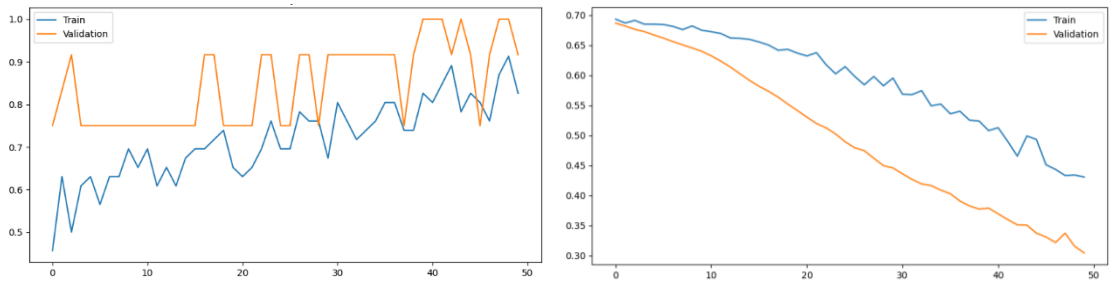


Figure 6. Accuracy Curve and Loss Curve

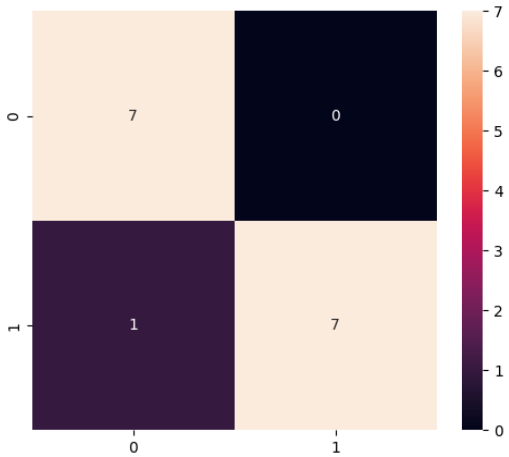


Figure 7. Confusion Matrix

Table 5. CNN Evaluation Results

Metric	Mean	Std
Accuracy	0.89	0.04
Precision	0.91	0.05
Recall	0.86	0.07
F1 Score	0.88	0.04

The training curve shows a steady increase in accuracy as the number of epochs increases, while the loss value consistently decreases. This indicates that the model is capable of learning the spatial patterns of Green turtle habitat well.

If the accuracy value is above 80%, the model can be considered to have good classification ability in distinguishing potential habitat from less potential habitat.

Predicted Distribution of Green turtle Habitat

The final stage of the research produced a distribution map of potential Green turtle habitat based on the CNN prediction results.

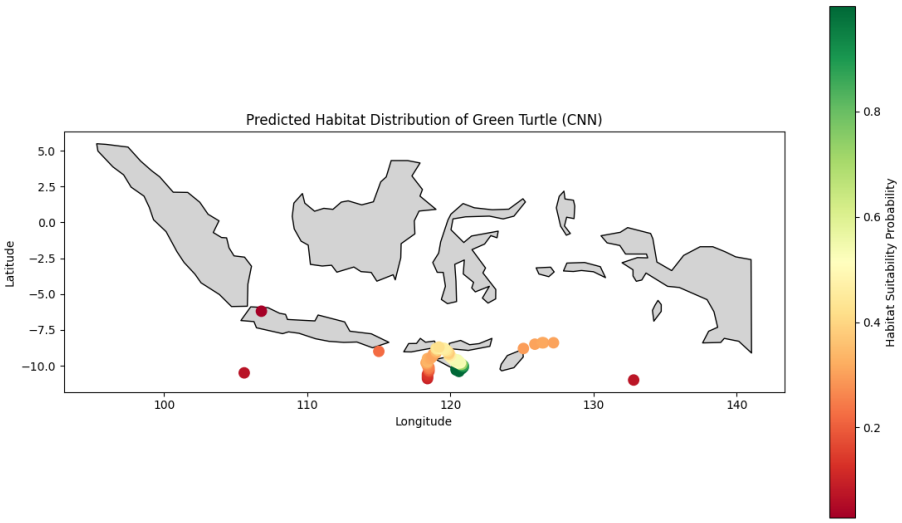


Figure 8. Predicted Map of Potential Green turtle Habitat Distribution

The prediction map shows that potential green turtle habitat is concentrated in several Indonesian waters previously identified as hotspots through the KDE analysis. The agreement between the KDE results and CNN predictions indicates that the model successfully captures the key spatial patterns found in the observational data.

Areas with high habitat probability are generally located in waters with favorable ecological characteristics, such as seagrass beds, coral reefs, and high primary productivity. Conversely, areas with low habitat probability tend to be located in open ocean areas with limited food sources and protective habitat.

CONCLUSION

This study successfully applied a spatial analysis approach based on Kernel Density Estimation (KDE) and Convolutional Neural Networks (CNN) to model the distribution and population density of Green turtles (*Chelonia mydas*) in Indonesian waters using occurrence data obtained from the Ocean Biodiversity Information System (OBIS). Data preprocessing, including data cleaning, geographic coordinate validation, Indonesian region filtering, and data transformation into a spatial grid, resulted in a representative dataset for species distribution analysis at a national scale.

The KDE analysis results indicate that the distribution of Green turtles in Indonesia is not homogeneous, but rather forms several hotspots with high densities concentrated in certain regions. Eastern Indonesia, particularly the waters of Nusa Tenggara, Maluku, the Banda Sea, and Papua, was identified as an area with a higher occurrence intensity than other regions. This finding indicates that these regions possess habitat characteristics that support foraging, migration, and habitat use by Green turtles.

Furthermore, the species distribution modeling using CNN was able to recognize spatial patterns from the KDE density-based grid data and generate predictions of potential habitats with good performance based on model evaluation metrics. The agreement between the hotspots generated by the KDE and the potential habitat areas predicted by the CNN indicates that the combination of the two methods effectively captures the spatial distribution patterns of Green turtles. This approach offers the advantage of integrating spatial density analysis and artificial intelligence to support the identification of critical habitats over large areas.

Overall, this study demonstrates that utilizing open biodiversity data combined with spatial analysis and deep learning methods can be an efficient alternative to support the conservation of protected marine

species. The resulting information on distribution hotspots and potential habitats can be used as a scientific basis for zoning planning of Marine Protected Areas (MPAs), critical habitat management, and more adaptive and sustainable Green turtle conservation decision-making in Indonesia.

ACKNOWLEDGMENT

The author would like to thank the Study Program and Faculty for their academic support during this research. He also thanks the supervisor and all parties who provided guidance, input, and support throughout the development of this research.

The author appreciates the Ocean Biodiversity Information System (OBIS) for providing open ocean biodiversity data that made this research possible. He also thanks the developers of the open-source software and scientific libraries used in this research, including Python, TensorFlow, Scikit-learn, GeoPandas, and various other spatial analysis tools.

We hope that the results of this research will contribute scientifically to the development of marine science, the conservation of green turtles (*Chelonia mydas*), and the sustainable management of marine protected areas in Indonesia.

REFERENCES

Aksoy, S., & Haralick, R. M. (2001). Feature normalization and likelihood-based similarity measures for image retrieval, *22*, 563–582.

Appeltans, W., Provoost, P., Principe, S. C., Driedger, A., Webb, T., & Costello, M. J. (2024). Biodiversity knowledge and threats on marine life: assessing no-take zones as a refuge for marine species. *STATE OF THE OCEAN REPORT*, 31.

Block, B. A., Jonsen, I. D., Jorgensen, S. J., Winship, A. J., Shaffer, S. A., Bograd, S. J., … Costa, D. P. (2011). Tracking apex marine predator movements in a dynamic ocean. *Nature*, 1–5. http://doi.org/10.1038/nature10082

Brodrick, P. G., Davies, A. B., & Asner, G. P. (2019). Uncovering Ecological Patterns with Convolutional Neural Networks. *Trends in Ecology & Evolution*, *34*(8), 734–745. http://doi.org/10.1016/j.tree.2019.03.006

Elfidasari, D., Sabil, M. Q. T., Sasaerilla, Y., & Sugoro, I. (2022). Distribution and nesting habitat of green turtle (*Chelonia mydas*) in Pangumbahan Turtle Conservation Area, Sukabumi, Indonesia. *AACL Bioflux*, *15*(5), 2329–2338. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85139153965&partnerID=40&md5=a82cabd86d119f894dfc5126022a5915

Hinton, G. (2014). Dropout : A Simple Way to Prevent Neural Networks from Overfitting, *15*, 1929–1958.

Hughes, A. C., Dorey, J. B., Bossert, S., & Orr, M. C. (2024). Big data, big problems? How to circumvent problems in biodiversity mapping and ensure meaningful results, 1–15. http://doi.org/10.1111/ecog.07115

Kingma, D. P., & Ba, J. L. (2015). A : a m s o, 1–15.

Mazaris, A. D., Schofield, G., Gkazinou, C., Almpanidou, V., & Hays, G. C. (2017). Global sea turtle conservation successes. Sci. Adv.3,e1600730(2017).DOI:10.1126/sciadv.1600730

Mujiyanto, M., Nastiti, A. S., Sugianti, Y., Fitriyana, E. H., Permana, T. D., Yuneni, R. R., Wiadnyana, N. N. (2026). Assessment of green turtle nesting under environmental pressure at Sindangkerta Beach, West Java, Indonesia. *Regional Studies in Marine Science*, *93*. http://doi.org/10.1016/j.rsma.2025.104721

Naro-Maciel, E., Hart, K. M., Cruciata, R., & Putman, N. F. (2017). DNA and dispersal models highlight constrained connectivity in a migratory marine megavertebrate. *Ecography*, *40*(5), 586–597. http://doi.org/10.1111/ecog.02056

Read, A. J., Fujioka, E., Best, B. D., Donnelly, B., Hazen, L. J., Kot, C., … Crowder, L. B. (2009). The Official Magazine of the Oceanography Society. http://doi.org/10.5670/oceanog.2009.42.COPYRIGHT

Sequeira, A. M. M., Rodríguez, J. P., Eguíluz, V. M., Harcourt, R., Hindell, M., Sims, D. W., & Duarte, C. M. (2018). Convergence of marine megafauna movement patterns in coastal and open oceans. http://doi.org/10.1073/pnas.1716137115

Silverman, B. W. (2002). DENSITY ESTIMATION FOR STATISTICS AND DATA ANALYSIS, (1951), 1–22.

Sinclair, S. J., White, M. D., & Newell, G. R. (2010). How useful are species distribution models for managing biodiversity under future climates?. *Ecology and Society*, *15*(1).

Tapilatu, R. F., Wona, H., Mofu, B., Kolibongso, D., Alzair, N., Erdmann, M., & Maruanaya, B. (2022). Foraging habitat characterization of green turtle, *Chelonia mydas*, in the Cenderawasih Bay, Papua, Indonesia: Insights from satellite tag tracking and seagrass survey. *Biodiversitas*, *23*(6), 2783–2789. http://doi.org/10.13057/biodiv/d230601