Detection of Fishing Activities from Vessel Trajectories

Aida Ashrafi[1(✉)], Bjørnar Tessem[1], and Katja Enberg[2]

[1] Department of Information Science and Media Studies, University of Bergen, Bergen, Norway
aida.ashrafi@uib.no

[2] Department of Biological Sciences, University of Bergen, Bergen, Norway

Abstract. This work is part of a design science project where the aim is to develop Machine Learning (ML) tools for analyzing tracks of fishing vessels. The ML models can potentially be used to automatically analyse Automatic Identification System (AIS) data for ships to identify fishing activity. Creating such technology is dependent on having labeled data, but the vast amounts of AIS data produced every day do not include any labels about the activities. We propose a labeling method based on verified heuristics, where we use an auxiliary source of data to label training data. In an evaluation, a series of tests have been done on the labeled data using deep learning architectures such as Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), 1D Convolutional Neural Network (1D CNN), and Fully Connected Neural Network (FCNN). The data consists of AIS data and daily fishing activity reports from Norwegian waters with a focus on bottom trawlers. Accuracy is higher than or equal to 87% for all deep learning models. Example applications of the trained models show how they can be used in a practical setting to identify likely unreported fishing activities.

Keywords: Fishing Activity Detection · Deep Learning Models · Data Labeling

1 Introduction

One of the Sustainable Development Goals (SDGs) set by the United Nations focuses on life below water (SDG14)[1]. As oceans and marine resources play an important role in different aspects of our life such as economy, food, and ecosystem functioning, stability, and resilience, using them in a sustainable way is a must.

Our effort within this context is a design science project [4] in collaboration with the Norwegian Directorate of Fisheries (NDF). The main goal is to develop deep learning models to help the sustainable use of fish resources. These models can potentially support the surveillance of fishing vessels, by exploiting data about the fishing vessels' movements combined with their reports on fishing activities. The data are from Norwegian waters and are provided by NDF.

[1] https://www.globalgoals.org/goals/14-life-below-water/.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
S. Nurcan et al. (Eds.): RCIS 2023, LNBIP 476, pp. 105–120, 2023.
https://doi.org/10.1007/978-3-031-33080-3_7

1.1 Problem Relevance

There are two approaches towards sustainable fisheries in the real world, and the best results are achieved when both of them are applied simultaneously. These two are: setting rules to regulate exploitation, and providing exploiters with incentives like ownership, for example in the form of quotas. This approach ensures that the long-term benefits of the resource become the priority of the exploiters [3].

In order to prevent Illegal, Unreported, and Unregulated (IUU) fishing, specific quotas are determined and surveillance is done regularly. However, as the high seas are not easily observable, rules are easily and sometimes violated. At the same time, the detection of IUU fishing by fishery inspectors is costly to implement.

Fishery inspectors have access to huge amounts of data about fishing vessels' movements obtained from Satellite-based Automatic Information Systems (S-AIS). However, no satisfactory automated approaches to support the detection of fishing activities are currently in use. While AIS data are mainly used to avoid collision between ships by tracking them, we envision the use of ML models as potential candidates to automate the detection of fishing activities from these data. This enables the inspectors to be aware of when and where fishing has taken or is taking place, and whether it is correctly reported. As a result, they would be able to focus their attention on vessels and locations where there is a higher risk of irregular fishing.

In Fig. 1 AIS data of 27 bottom trawlers on 38 fishing trips are shown. Red points show non-fishing activities and blue ones show fishing activities, based on labeling obtained from the fishing vessels' records.

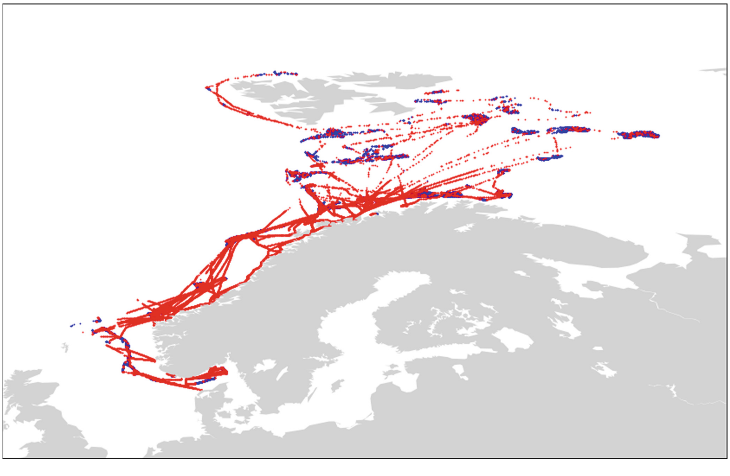


Fig. 1. AIS data for 27 bottom trawlers on 38 fishing trips, red points are non-fishing and blue points are fishing activities according to reports by fishermen used our labeling strategy (Color figure online)

1.2 The Contribution

AIS data include features such as latitude, longitude, speed, and course over ground in addition to the vessel's ID, the so-called Maritime Mobile Service Identity (MMSI). The normal frequency of reporting AIS data is every 5 min (with many exceptions). These discrete data points form the vessels' trajectories.

In our analysis, we segment these trajectories and label each segment as a fishing or non-fishing activity. Each segment is thus a sequence (time series) of AIS data and the task is to build a model that is able to classify these segments into fishing or non-fishing.

A main contribution of our work is a method to provide more labeled training data for the ML process. This is a significant problem, as it is too time-consuming for an expert to do the labeling. As an alternative, we label the data by using the records from fishing vessels' daily reports on fishing activities. Although this method is based on the reports by the fishermen, domain experts from NDF have verified that the obtained labels are accurate on examples, and that the accuracy is at a level that is sufficient to train a model. The ML models are the second contribution of this work. They should be seen as specific to the data from this region and for a particular fishing gear, as the experts believe that the movement patterns are highly related to these aspects. We focus here on bottom trawlers. However, the general approach should have validity also in other geographical areas and other types of fisheries.

In the next section, the research method is discussed. In Sect. 3 we describe the background for the problem and look into related work on fishery activities. We continue by describing the original data set used, and then the methods used to extract labels of the trajectory segments from the auxiliary data source, ending with a final selection of data (Sect. 4). In Sect. 5 we describe our evaluation approaches and show the models' results, as well as give an example of how the model performs on some interesting cases from the data. Some discussion and conclusions are presented in Sect. 6.

2 Research Method

The research method we have followed is standard design science as described by Hevner et al. [4]. We have created methods for labeling vessel tracking data, used for the further purpose of developing the ML models used to analyse vessel tracks. Two of the most important guidelines in design science are design as an artifact and design evaluation. The most significant design choices we have made for our tools and the sections in which they are explained are indicated in Table 1.

In the evaluations described in Sect. 5, different tests are done on unseen data using different ML architectures. Results are reported in the form of traditional performance metrics from the ML field. As a descriptive evaluation (in the terms of Hevner et al [4]), we provide visualization of speed and tracks of vessels combined with the output of the ML models.

Table 1. Overview of design choices

Design Choice	Related Section
Selection of data from bottom trawlers	Sect. 4.1
Obtaining labels for AIS data using daily catches	Sect. 4.2
Identifying and removing inaccurate data	Sect. 4.2
Selecting features for ML	Sects. 4.3 and 4.4
Representation of the data as track segments and labeling of segments	Sect. 4.5
ML architectures and performance metrics	Sect. 5.1

3 Background and Related Work

Commercial fishery is a complex activity dependent on many factors such as the target species, area of the sea, vessel size, and gear type. In Norwegian waters cod, haddock, pollock (also called saithe), mackerel, and herring are considered the most valuable species.

There are mainly two groups of fishing gears: passive gears such as long line and purse seine, and active gears such as trawl. The species most often targeted with trawl are cod, haddock, and pollock. Our focus in this paper is on recognizing fishing activities for vessels doing bottom trawling, which basically is performed by pulling a bag-formed net along the seafloor to catch the target fish.

There have been several attempts to classify fishing activities using ML. de Souza et al. [9] used ML models, i.e., the Hidden Markov Model (HMM) and data mining to identify fishing activities from AIS data. Three types of fishing gear were analyzed: trawl, longline and purse seine.

Jiang et al. [6] published the first work applying deep learning models for detecting the fishing activities of trawlers. A sliding window technique was used to divide the trajectory into shorter segments. They labeled each window with the same label as that of the middle point. To reduce the noise, undersampling was applied. Linear interpolation was utilized to recapture the trajectory which was then converted to an image matrix. Lastly, an autoencoder was used to detect fishing activities. According to Jiang et al. [5], AIS data is low-dimensional and heterogeneous making it hard to work with deep learning models on these data. They proposed Partition-wise Recurrent Neural Networks (pRNNs) to solve this issue. Their focused fishing gears were long-liners.

Global Fishing Watch (GFW)[2] has provided commercial fisheries datasets that are used in many studies. Kroodsma et al. [7] developed CNN models for the recognition of different vessels' features and also the detection of fishing

[2] https://globalfishingwatch.org/.

activities. In a more recent paper [1], they segmented the trajectories into smaller intervals and used the majority label as the label of each interval. Their proposed model (FishNET) is based on 1D CNN and new features extracted from main features are used to make the method independent of changes in gear type, vessel type, and location.

Shen et al. [8] used a multi-layer Bidirectional LSTM (BiLSTM) model to test the importance of different features in detecting fishing activities with both active and passive gear types used around Taiwan.

The most significant challenge in fishing activity detection tasks is the lack of labeled data [1]. However, the most common open and labeled data sets are limited, even though they are related to many gear types. Annotation of the data by experts is labor intensive and also difficult in reality since even experts are not aware of all fishing patterns.

A work close to ours is by Ferreira et al. [2]. They do not use true labels, but still classify sub-trajectories into fishing and sailing. k-means clustering based on speed and course changes is applied to find the labels in an unsupervised manner. They further use these labels with LSTM and Gated Recurrent Unit (GRU) units to do the classification task. The clustering method involves many tuning details and appears to be rather cumbersome.

4 Datasets and Preprosessing

4.1 Data

The data we use is an AIS dataset that provides fishing vessels' movement data, and DCA (Daily Catch) reports[3] which help to identify fishing activities. The AIS dataset includes information such as vessels' ID (MMSI), message time, latitude, longitude, speed over ground, and course over ground. These data are not labeled. However, DCA reports can be used to extract labels from the fishermen's reporting. DCA data consist of reported fishing intervals (start and stop time, duration, amount, and location of each catch), plus Call-sign for each vessel which itself could be matched to MMSI from the AIS dataset through another data table.

A main difference of our work from earlier studies is our ability to assess the data with the help of domain experts. According to them, the DCA reporting data regarding active gears (like bottom trawl) are cleaner and easier to work with since the patterns are more visible and the duration of fishing activities are recorded more accurately. Also, different types of active gears lead to different patterns. That is why, unlike most of the works in the literature, we have chosen to focus on only one gear type at a time. In this study, we only consider bottom trawlers. With a quick look at our dataset, we could conclude that there are enough data for this gear type to train and test the models. We have used more than 500 different fishing trips of almost 100 vessels between 2015 and 2020.

[3] A part of the electronic reporting by NDF: https://www.fiskeridir.no/Tall-og-analyse/AApne-data/elektronisk-rapportering-ers.

For each vessel, we chose a maximum of 5 fishing trips randomly. This gave us a subset of the data with complete trajectories from whole fishing trips. The vessels have provided departure and landing records (see footnote 3) which can be used to identify fishing trips. Using complete fishing trips allows us to get a representative data set, but visualizing the trips also helps to get a better understanding of the activities in a trip.

4.2 Obtaining Labels

We have used the DCA reports delivered by fishermen to label the AIS data. Some cleaning, however, has been needed. For example, it is immediately meaningless to include fishing activities reported to have zero duration. In addition, vessels are allowed to send correction messages, which are basically duplicates that are sent later. After eliminating these kinds of messages, there would be no overlaps among different records in DCA data.

In the next step, we have labeled the AIS data belonging to the chosen fishing trips. The labels are not certain to be true labels, as they are based on reports by fishermen, who are notoriously being inaccurate. But since the annotation process of AIS data by experts is very time-consuming and the DCA reports can be used to extract labels with a satisfactory level of trust, we chose to use the DCA data to obtain labels. This way we benefit from a larger training set and achieve good performance.

Still, some of the reporting indicated very long or short fishing activities. Such long or short fishing activities are considered to be most likely inaccurate. So, to remove noise from the DCA reports we eliminate the catch activities with a duration of less than 30 min or more than 400 min and define them as irregular messages. The histogram of duration for the most important species is shown in Fig. 2. Finally, we check each AIS data point from desired fishing trips to see if they are located inside any of the fishing intervals. If they are, we label them as fishing, and if not, they are labeled as non-fishing. The ones excluded due to them being outside the acceptable intervals will be used later for evaluation purposes.

4.3 Selecting Features for ML

The normal frequency of sending AIS messages is every 5 min, although this does not happen all the time. Therefore, we decided to add time difference (ΔT) which shows the difference between the time of the current message and the previous one's as a new feature. Messages with very long time-difference (more than 200 min) were also removed. Another feature that has been added is the month of the year that the message has been sent. This is a cyclic attribute and we believe since this feature is related to environmental factors, it can affect the patterns of vessels' movements. The rest of the features are mostly the same as the ones in [8] such as speed (SOG), average speed (S_{avg}), and change in the course (ΔCOG). We also consider changes in the speed (ΔSOG). In our work, instead of distance (distance between the current position and previous position,

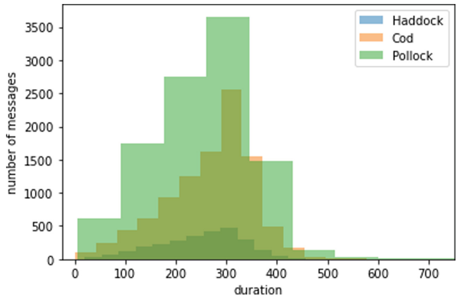


Fig. 2. Duration histogram of catch activities from 2018 for the most important species.

Table 2. definitions of different features used in our method, i is the index of i^{th} data point in the sequence

feature	definition
ΔT	time difference: $T_{i+1} - T_i$
Month	month of the year coded as $Month_{sin}$, $Month_{cos}$
SOG	Speed over Ground
S_{avg}	Average Speed: $S_{i+1} + S_i/2$
ΔSOG	Speed over ground difference: $SOG_{i+1} - SOG_i$
ΔCOG	Course over ground difference: $COG_{i+1} - COG_i$
P	Position which includes latitude and longitude

ΔP), we have chosen to use latitude and longitude (position, P). Our focus is the Norwegian waters and we get a better result if we are more specific about the location. Furthermore, experts suggest that these patterns are very specific to regions, hence it is potentially better to build and train the models for each area separately. The features we used and their definitions are shown in Table 2.

4.4 Choosing Depth as a Feature

In our later sessions with experts, they suggest using depth as a feature to check for possible improvements plus providing explainability. Depth data is not part of AIS data and we needed to use publicly available bathymetry data sets provided by General Bathymetric Chart of the Oceans (GEBCO)[4]. Considering the position of each data point in AIS dataset as a center, we form a 0.01×0.01 latitude-longitude box around it, find the closest point to it using Haversine distance, and extract the depth value. If there are no points inside the box, we increase the box to $0.1 \times 0.1°$. Adding depth in combination with various other

[4] https://www.gebco.net/.

features did not help in achieving better performance and in some cases cause a slight increase in the loss value. The depth data might thus not be useful for developing models. And we are already considering the location, which obviously is a proxy for depth. But it is still useful for experts to explore the activities on depth maps to achieve a better understanding of the data and to evaluate the model's performance.

4.5 Segmentation of the Vessel Trajectories

The trajectory of a fishing vessel is the track of its movement in the form of a time series. Since it is long and includes different fishing patterns and intervals, it is preferable to divide it into shorter sequences (windows). This process is needed to prepare the input for LSTM, RNN, and 1D CNN. For FCNN, we represent the sequences as non-sequential data points.

It is very common in the literature to use the sliding window technique (Fig. 3) and the window size is mostly between 5 to 15 datapoints [1,6,8]. On the other hand, they have different approaches to label each sequence, either choosing the majority vote of the window or the label of the middle point of a window.

In our work, we tested different alternatives. We trained the model with sequences of length 8 and chose the label of the last point as the sequence's label. We also tried using the middle label and window size 10. No substantial difference in the performance among these choices has been observed, so we decided to keep the first setting.

5 Evaluation and Results

5.1 Using Different Architectures

To compare different deep learning architectures, we have used the same set of hyperparameters. For each of the architectures we trained 10 models, and the average scores of these 10 models are reported as the score. The hyperparameters used were

Number of epochs: 20
Neurons in hidden layers: 128
Optimizer: Adam
Activation functions: ReLU and Softmax
Batch size: 32
Loss function: binary cross entropy loss

The first model we tried was FCNN using the sequence data as a flattened input set. But since the data is in the form of a time series, we assumed that other architectures such as LSTM, RNN, and 1D CNN, which are designed for working with sequences, would be more suitable. Besides segmenting trajectories into sequences to feed into these models, we added the extra hidden specialized

Table 3. Performance on the test set from 2018 using different models

FCNN	Accuracy	Loss	F1-score	Precision	Recall
	90	0.27	89	87	92
LSTM	**92**	**0.22**	**91**	**90**	**94**
BiLSTM	92	0.23	91	90	92
RNN	91	0.23	91	89	93
1D CNN	91	0.25	91	88	93

recurrent or convolutional layer (LSTM, RNN or 1D CNN), as well as a dropout layer with 0.5 dropout rate for those models. We also tried bidirectional LSTM (BiLSTM).

We first tried all the candidate models on the 2018 dataset, with both training and testing data set from that year. All the models except FCNN consider temporal dependency. We, therefore, expected them to perform better than FCNN. The results were in line with that expectation. However, the difference was not substantial. Different metrics such as accuracy, loss, precision, recall, and F-score were calculated. The results are depicted in Table 3.

5.2 Some Observation Related to Overfitting

In our initial tests, the windows were not overlapped. To use the data in the best possible manner, we decided to extract overlapping windows by using the sliding window approach (Fig. 3). This way we can provide more sequences from the same dataset just by shifting the intervals halfway to the right. Adding these overlapped sequences to the original ones caused an overfit, though. We also observed that more complex models (more units per layer and more layers), and more epochs also result in an overfit.

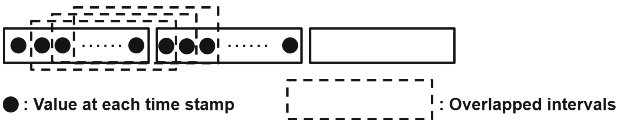


Fig. 3. Overlapped windows

5.3 Using Data from Different Years

To check if there are common fishing patterns among different years and if a model trained on the data from a specific year can perform reasonably on a test set from the following year(s), we did further tests using training and test sets belonging to different years.

For the rest of the tests, LSTM was chosen. First, the LSTM trained on 2018 data, was tested on data from 2019 and 2020. We also trained an LSTM on 2019 data and tested it on 2020 data. Finally, we trained a third LSTM on the data from 2015, 2016, 2017, 2018 and 2019 and tested it on 2020 data.

The accuracy of all the models was quite high on the test sets, higher than or equal to 87%. The scores are shown in Table 4. It seems that the fishing patterns do not change much over the years as the scores do not vary much. However, they are a bit lower when using a test set from a different year. The reason for this may be that the selection of vessels in the data sets is a bit different, and therefore will give a slightly lower performance.

Table 4. Performance on the test sets from different years using LSTM

train	test	Accuracy	Loss	F1-score	Precision	Recall
2018	2019	91	0.25	91	89	93
2018	2020	88	0.29	87	82	92
2019	2020	89	0.29	88	84	92
2015–2019	2020	87	0.32	86	81	92

5.4 Testing on Outliers

Outliers are, as we use the term, the DCAs which indicate very short or very long catch intervals. We have removed this data in the labeling step so that the model would be trained on accurately labeled fishing activities. However, at test time, we can check this part of the data set to see the difference between the output of the model and the reports by fishermen. Although fishermen labeled all these sequences as fishing, according to experts these reports are most likely incorrect. As expected, our model obtains less accuracy on this part of the data, as the prediction is probably more correct than the ground truth (labels by fishermen). The comparison of the LSTM's performance testing on outliers and non-outliers from 2018 is demonstrated in Table 5. An interpretation of these results could be that about 10–15% of the daily catch reports are incorrect.

Table 5. Comparison of performance on non-outliers vs outliers test set from 2018

	Accuracy	Loss	F1-score
Non-outliers	92	0.22	91
Outliers	85	0.45	91

5.5 Speed Analysis to Test Our Method

According to the experts, vessel speed is one of the most important contributing factors in detecting the fishing activities of bottom trawlers. The results of our tests also confirm this argument, the models can achieve comparable results (all the metrics are less than 1 or 2% below using LSTM) considering speed as the only feature. Therefore, we provide some visualizations of speed analysis which could help to see whether the model is learning and doing something reasonable.

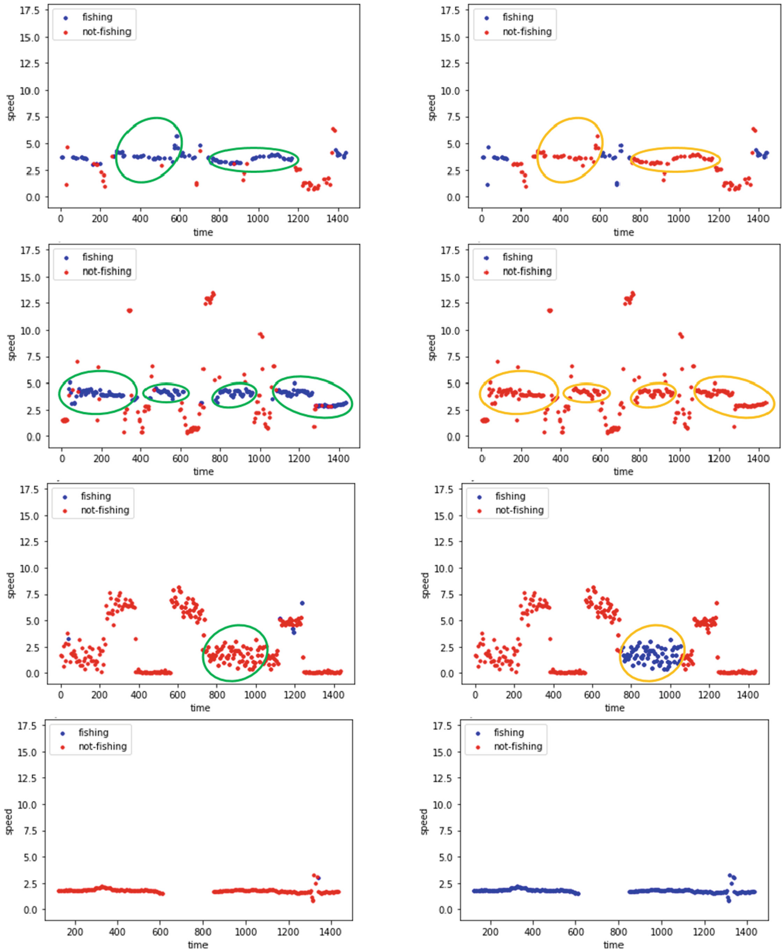


Fig. 4. speed analysis over time for a day: prediction (first column) vs truth or reports (second column)

The model was incorrect in about 10% of the sequences. The visualizations in Fig. 4 show a selection of those and where the discrepancy occurs. We hope that later by combining these graphs with the patterns in the tracks, domain experts could help us to get labels closer to the true labels. These visualizations are for the 2018 test set and predictions are the output of FCNN. The ground truth labels are the labels from using DCA data, not the true labels confirmed by experts.

Each row in Fig. 4 shows the speed analysis of different vessels on different days. Time is in minutes with a day being 1440 min. The blue dots are used to depict non-fishing operations and the red ones are used for fishing activities. The green and yellow circles illustrate the most visible differences in our prediction on the left and the labels by fishermen on the right for the same day.

In the first row, our prediction shows two fishing intervals inside green circles while the yellow circles on the right graph show non-fishing intervals as reported by fishermen. The second row shows the same situation but with no fishing reported by fishermen. In the third row, we observe one dense different interval but this time we predict non-fishing activity for that interval while it is reported as fishing in the reports. The last row shows an example of outlier data and the two figures are different at almost all points. In the reports by fishermen, all the points are reported as fishing which does not seem to be aligned with the rest of the data on which our models are trained. A quick analysis of the patterns seems to suggest that fishing activity with bottom trawls normally is performed at about 5 knots, cresting the discrepancy in row four. Further, there seems to be an expectation by the model that bottom trawling happens with a more stable speed than found in row three.

In general, the model seems to make reasonable decisions regarding fishing activity detection and even has the possibility to correct the reports. We believe that these differences happen when the fishermen do not report their fishing activities in time while they are fishing. Either they report activities carelessly much later, or they report intentionally incorrect intervals. That is why the ML predictions seem more correct in many cases.

5.6 Tracks Analysis to Test Our Method

In our tests, we consider the position (latitude and longitude) as a feature instead of the distance traveled (as common in previous works). This is due to the experts stating that fishing patterns are dependent on the region. This was confirmed as we also obtained better results using the exact positions. In Fig. 5 we see the track of a fishing trip that takes 12 days on a map of the Norwegian water. The trip has been chosen from the test set. The white-colored areas are those with depths higher than 100 or lower than −700 m.

Figure 6 shows the same track as in Fig. 5 labeled by the output of FCNN. On the bottom, we focused on two busy areas from the original map. Green data points are for True Positive (TP), meaning that the model's prediction for that data point was fishing and the label by fisherman's reports was also fishing. Blue points represent True Negative (TN), i.e., where the model output agrees on being non-fishing with the reports. As you can see the colors of the tracks are in line with the 90% accuracy of the model since most of the data points

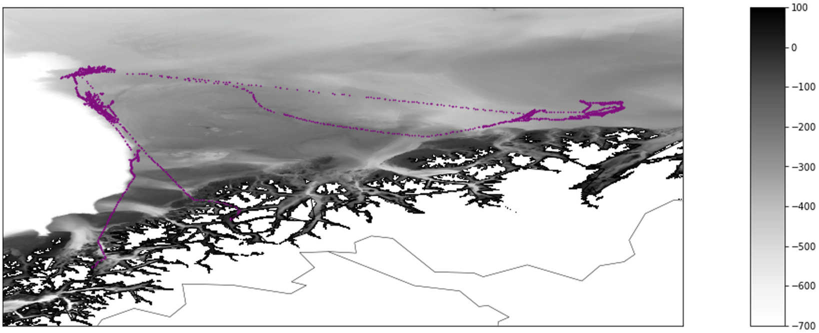


Fig. 5. Track of a fishing vessel from the test set on a map including depth data, white areas have depth lower than -700 or higher than 100 m.

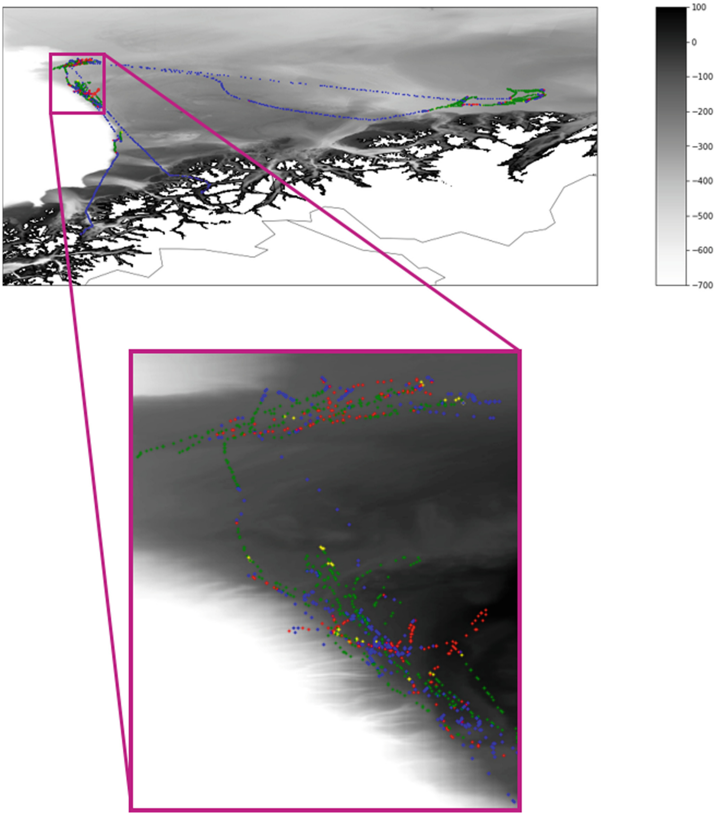
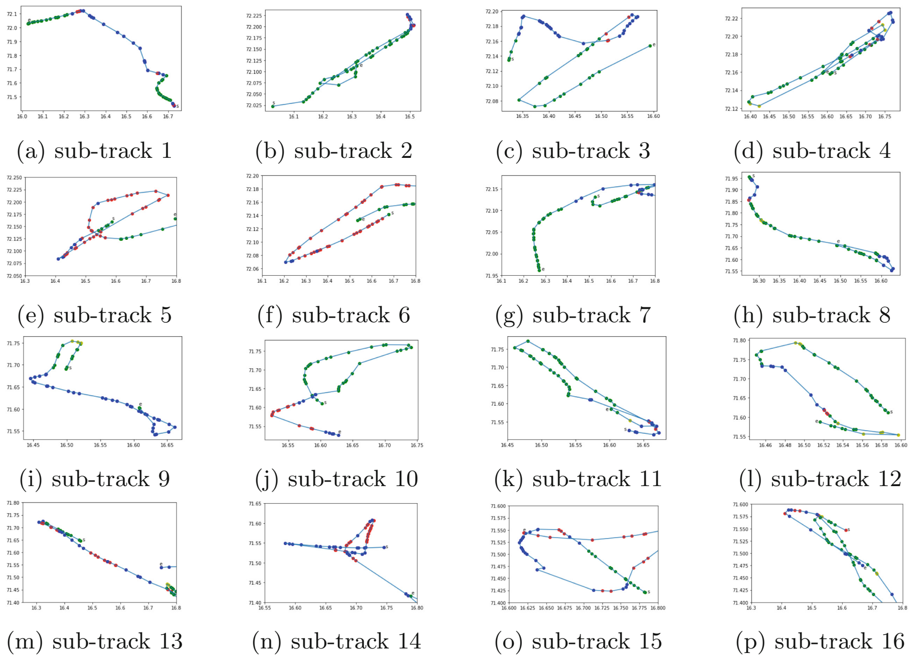


Fig. 6. Track of a fishing vessel with the labels by the model on the top, focus on the busy area on the bottom. Green for TP, blue for TN, red for FP, yellow for FN. (Color figure online)



(a) sub-track 1 (b) sub-track 2 (c) sub-track 3 (d) sub-track 4

(e) sub-track 5 (f) sub-track 6 (g) sub-track 7 (h) sub-track 8

(i) sub-track 9 (j) sub-track 10 (k) sub-track 11 (l) sub-track 12

(m) sub-track 13 (n) sub-track 14 (o) sub-track 15 (p) sub-track 16

Fig. 7. 16 sub-tracks from the focused area of Fig. 6 with labels, green: TP, blue: TN, red: FP, yellow: FN. The horizontal axis shows longitude and the vertical one is latitude. Starting point is shown with 's' and the ending point with 'e'. (Color figure online)

are in green and blue. Red points represent False Positive (FP) where the model outputs fishing while the fisherman reported non-fishing activity. Finally, yellow points are the ones used for False Negative (FN) for which the model suggests non-fishing whereas the reports show fishing activity. We can observe that most false predictions belong to FP.

We again split the tracks in this focused area into 16 sub-trajectories of lengths 53 and 54 to see the patterns better. In Fig. 7, these sub-parts with their starting (s) and ending (e) points are shown. Some of the sub-tracks are not placed completely inside the focused area. Figures 7e, 7f, 7g, 7m, 7n, 7o, and 7p are examples in which this situation happens, therefore, the sub-tracks are cut reaching the borders and come back to the frame again when the position is inside the focused area again. In some of the sub-tracks such as Figs. 7e, 7f, and 7n some larger red intervals can be observed. This is where the model predicts fishing activities while they are reported as non-fishing by fishermen. These intervals most likely include unreported operations and finding them is the main purpose of our project. In some of the sub-tracks such as Figs. 7a, 7b, and 7c there are some red data points spread inside blue intervals. These data points are predicted as fishing by the model while the data points in their neighborhood are labeled as non-fishing by both the model and the fishermen. These points are most likely mistakes by the model but even in that case, they are very rare.

6 Conclusion and Discussion

Sustainable use of fish resources is highly important for the development of humanity's future, and the long-term goal is to avoid overfishing. While many fishing nations are active in the surveillance of fishing vessels' behavior on the ocean, there are still unreported activities going on in many areas. In this work, we tried to automate fishing activity detection.

The most common problem for tools like this is the limited number of labeled data and the expensive annotation process. We have proposed a method to extract labels using an auxiliary data source, namely the publicly available daily catch reports from NDF. Our alternative approach achieves accuracy at par or better than the works where they use expert labels. The labels are based on fishermens' reports and may be incorrect, but by removing outliers, they can be trusted to be at a sufficiently good level to be used in ML. Thus we can have significantly more data to train our model than in previous work.

Since the data is spatiotemporal we opted to use different models such as LSTM, RNN, and 1D CNN, and for that, we segmented data into smaller windows. We also explored the performance of FCNN. The difference was small. The 1 to 2% difference in accuracy might be due to the importance of temporal dependency. All the models are fairly simple and using more complex models or overlapped sequences ended in over-fitting. Testing the models on different years still gave a high performance. Those parts of the data which were left out during the training were also used to test the model. In this case, the model showed fairly good performance and seems to be able to identify irregular reporting.

We have developed models by using different features which were partly different from previous works. As experts suggested the impact of region on fishing patterns, we picked the exact position rather than the distance. Depth was omitted from the feature set, but can still be useful in order to help the expert get a better interpretation of the model output. Visualizing tracks and patterns, complemented with depth maps could help experts in the process of providing labels closer to true labels.

There are some limitations to our work. First, the models we used could only obtain around 90% accuracy by testing different sets of features, architectures, and hyperparameters. The various tests indicate no promise for substantial improvements in the performance. We believe this comes from the heuristics for removing inaccurate data being imperfect, which again created noise in the data labels. Second, our work is not totally comparable with the previous studies on the same topic since our focus is narrowed down to Norwegian waters and bottom trawlers. Although the method could be generalized to other regions and gears provided there are data, we cannot guarantee generalizability of the results.

In the future, we are going to extend our work to other active gear types and also to passive gear types. Further, we believe there are more major features such as seabed surface which can be included to increase the performance. We are also going to develop an application that presents results of the model, speed, and track analyses to domain experts. They will be invited to correct the labeling in an easy way and thus help us to obtain more true labels.

References

1. Arasteh, S., et al.: Fishing vessels activity detection from longitudinal ais data. In: Lu, C.T., Wang, F., Trajcevski, G., Huang, Y., Newsam, S., Xiong, L. (eds.) Proceedings of the 28th International Conference on Advances in Geographic Information Systems, pp. 347–356 (2020). https://doi.org/10.1145/3397536.3422267
2. Ferreira, M.D., Spadon, G., Soares, A., Matwin, S.: A semi-supervised methodology for fishing activity detection using the geometry behind the trajectory of multiple vessels. Sensors **22**(16), 6063 (2022). https://doi.org/10.3390/s22166063
3. Heino, M., Enberg, K.: Sustainable use of populations and overexploitation. In: Encyclopedia of Life Sciences (eLS). Wiley, Chichester (2008). https://doi.org/10.1002/9780470015902.a0020476
4. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. Manag. Inf. Syst. Q. **28**, 75–106 (2004)
5. Jiang, X., Liu, X., de Souza, E.N., Hu, B., Silver, D.L., Matwin, S.: Improving point-based AIS trajectory classification with partition-wise gated recurrent units. In: 2017 International Joint Conference on Neural Networks (IJCNN), pp. 4044–4051 (2017). https://doi.org/10.1109/IJCNN.2017.7966366
6. Jiang, X., Silver, D.L., Hu, B., de Souza, E.N., Matwin, S.: Fishing activity detection from AIS data using autoencoders. In: Khoury, R., Drummond, C. (eds.) AI 2016. LNCS (LNAI), vol. 9673, pp. 33–39. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-34111-8_4
7. Kroodsma, D., et al.: Tracking the global footprint of fisheries. Science **359**, 904–908 (2018). https://doi.org/10.1126/science.aao5646
8. Shen, K., Chu, Y., Chang, S.J., Chang, S.: A study of correlation between fishing activity and AIS data by deep learning. TransNav Int. J. Mar. Navig. Saf. Sea Transp. **14**, 527–531 (2020). https://doi.org/10.12716/1001.14.03.01
9. de Souza, E.N., Boerder, K., Matwin, S., Worm, B.: Improving fishing pattern detection from satellite AIS using data mining and machine learning. PLoS ONE **11**(9), 1–2 (2016). https://doi.org/10.1371/journal.pone.0163760