

## Identification of Vessels on Inland Waters Using Low-Quality Video Streams

Tomasz Hyla  
West Pomeranian University of Technology,  
Szczecin, Poland  
[thyla@zut.edu.pl](mailto:thyla@zut.edu.pl)

Natalia Wawrzyniak  
Marine Technology Ltd., Gdynia, Poland  
[n.wawrzyniak@marinetechonology.pl](mailto:n.wawrzyniak@marinetechonology.pl)

### Abstract

*Video surveillance can be used to monitor recreational and commercial vessels on inland waters. It is difficult to identify vessels on the basis of registration numbers alone, especially when structure and size are not standardized in a way that allows for easy recognition. In this paper, we propose a novel vessel identification method that can work continuously on streams with a high compression rate or visible compression artifacts. This method can identify all types of moving vessels during daylight. It should work with existing video monitoring systems used to monitor inland waters. When a vessel has hull inscriptions, the method can identify that vessel. If hull inscriptions are not present, the method identifies the direction of movement of the vessel. The tests were performed by real-time video playback of prerecorded samples in the test environment. The method correctly identifies 94% of vessels and the performance indicates that it is suitable for practical application.*

### 1. Introduction

There is a need to monitor vessel traffic on both inland and coastal waters. The primary source of information on vessel identification is the automatic identification system (AIS) [1], which is based on radio transponders. However, recreational vessels are not obliged to carry AIS transponders, and therefore the only way to identify them is via video surveillance. On inland waters, especially rivers, River Information Services (RIS) [2] are used to monitor and support ship navigation. In many of these systems, multiple video cameras are placed at different locations and are mainly used to display current traffic to the operator. A useful extension of such systems would be a system that could automatically identify all passing vessels.

It is difficult to identify vessels on the basis of registration numbers alone. The structure and size of vessel registration plates is not standardized in a way that allows for easy recognition. Any commercial vessel must have an International Maritime Organization

(IMO) number [3] written on its hull. However, the IMO number has seven digits, including a control digit, which means that there is a probability of incorrect recognition of 1 in 10. Moreover, on inland waters, many vessels have inscriptions, such as phone numbers and advertisements, that can be mistaken for the vessel name or registration number. In Poland, for example, there are several types of registration numbers for leisure craft that have common prefixes (such as POL for Polish Yachts), but their length usually varies. Additionally, leisure vessels' hull inscriptions are not placed in a standardized way, and they use different fonts, sizes, and colors. Sometimes, a name or registration number is not written on a leisure vessel at all.

However, in Poland and in most of the EU, vessels' registries are public, containing vessel data without information about ownership or personal data. Therefore, vessels' names and registration numbers can be used to match a text string read from a video frame to a specific vessel.

### 1.1 Related works

A number of aspects of automatic vessel identification based on video streams have been investigated over the years alongside general advances in natural scene text recognition [4]. However, the number of documented studies in this narrow field is not impressive. Traditional approaches divide the problem of identifying ships into two parts, the first being the detection and tracking of on-water objects in rapidly varying weather and lighting conditions, and the second the recognition of detected vessel plates or inscriptions to be matched with existing data on ships in the area of observation. Ships detection and tracking problem has been solved in many different ways (e.g. by abovementioned AIS or radar systems), however, the existence of video monitoring systems in ports and other on-shore facilities allowed for computer vision techniques to be used for that purpose. Significant development of methods occurred when high resolution

optical sensors appeared and the computing power of data processing systems increased.

In 2014, a survey of vessel detection and tracking based on maritime surveillance systems was presented [5]. Before the explosion in deep learning techniques, the relevant research was mainly concerned with addressing the problem of background removal in the presence of waves [6,7] and of changing daylight to allow proper extraction of the actual target for further processing. A fusion of Bayesian classification, Kalman filters [8], and particle filters [9] was presented to improve detection and tracking quality.

The use of scene text recognition techniques in the identification process has received attention recently, with the development of methods based on convolutional neural networks (CNNs) [10]. Some of these use CNNs to divide the task into two phases: (1) detecting text areas on the vessel image and (2) recognizing detected inscriptions using optical character recognition (OCR). Zang et al. [11] proposed a method that used a CNN to locate ships and detect plate text lines with a fully convolutional network (FCN). They compared the results with the popular EAST method [12]. Huang et al. [13] proposed a two-step approach with an end-to-end method for naval ship plate number recognition, taking into account differently arranged vessel numbers.

Ferreira et al. [14] described a procedure for identifying the ships based on matching vessels' plate numbers with information contained in the database of the Vessel Monitoring System (VMS).

## 1.2 Contribution

In this paper, we propose a novel vessel identification method (VIM) designed for inland waters and river information systems. This VIM can work on low-quality streams. By low-quality streams we mean streams with a high compression rate (<10 Mbit/s) and visible compression artifacts, but with the minimum required resolution being full HD (1920 × 1080). This minimum resolution is required because at lower resolutions, inscriptions are usually too small and cannot be correctly recognized.

The method can identify all types of moving vessels during daylight. When a vessel has a hull inscription (name or registration number), we can identify the vessel (from the vessel ID available in a public registry). If hull inscriptions are not present, the method identifies a vessel's directions of movement. The vessel is identified once per passage in front of the camera.

The VIM works in continuous mode, and results are returned on an ongoing basis; i.e., it detects vessels once per second and returns intermediate identification results every 5 s. It returns a final identification decision

up to 15 s after the vessel has passed in front of the camera. Vessels move rather slowly (in comparison with road vehicles), and hence such an identification rate can be considered to be real-time in practical applications. Experimental results show that the method can be applied in practice and can be integrated with a river information system.

The proposed method has as one of its components a vessel detection and tracking method proposed in [15]. The detection method takes as its input video streams and returns a series of pictures with the same tracking ID. The main new contribution compared with [15] is the determination of the vessel ID for the tracking ID.

## 2. Method

The VIM is designed to work with highly compressed video streams (full HD resolution or more) that may have visible compression artifacts. It should work with existing video monitoring systems used to monitor inland waters, where remote locations of some cameras make it impossible to send a high-quality video stream. The video streams must come from a stationary camera located on a bridge or shore near a waterway. To detect every vessel the cameras must be located in narrow passages, where two or more cameras can be placed on both sides of the waterway. This minimizes the chance that a vessel is occluded in views of all cameras at the same time. Such locations are typical for river monitoring systems.

The input for the method is a video stream. The method runs as long as the video stream is available. All types of passing vessels can be identified, but the method is mostly intended for pleasure boats, since they cannot be identified using other systems. The output includes three types of vessel identification:

- 1) Full identification: the vessel inscriptions are matched with a vessel registry without errors; the inscriptions are visible in the streams, they can be read by a human operator, and they are at least 12 pixels high.
- 2) High identification: this is the same as full identification except that a vessel is identified correctly with high probability rather than with certainty, i.e., one or two errors occur in the matching process.
- 3) Movement direction identification: the vessel is not clearly marked, the inscriptions are too small (<12 pixels), or letters are indistinguishable because of compression artifacts.

The method requires the availability of an existing vessel registry containing ships' names and registration numbers. This allows identification of ships' names and registration numbers with no fixed structure.

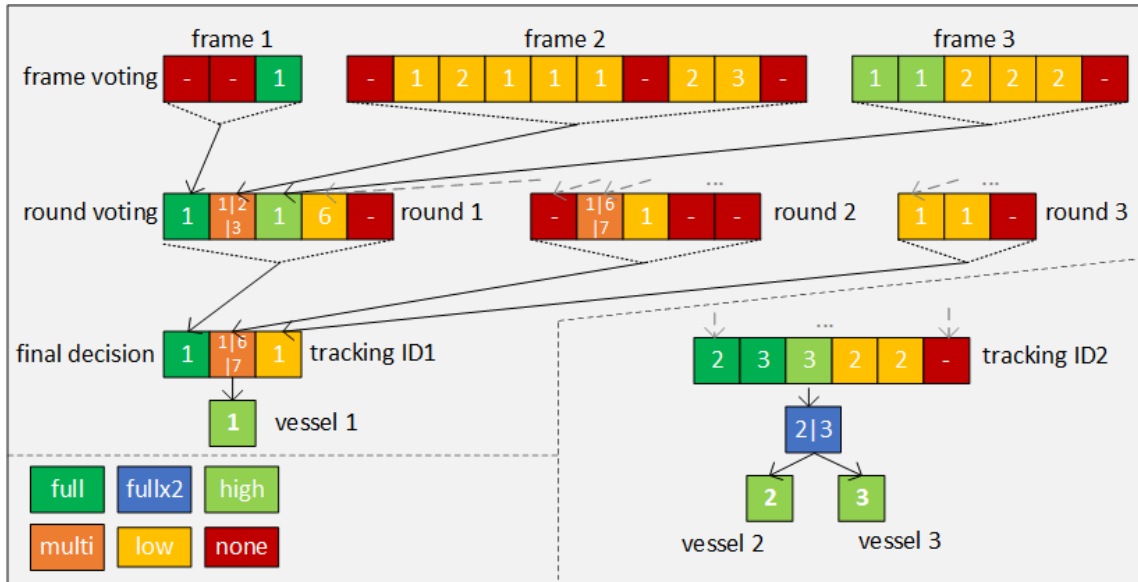


Figure 1. Illustration of the decision process

## 2.1 Informative description

The method consists of three logical layers and uses external components. The first layer is a detection and tracking layer. It uses a method previously proposed in [15]. This detection and tracking layer returns sets of frames every 5 s. Each set of frames has a tracking number. The second layer identifies vessels' inscriptions in each frame set. For increased accuracy, the identification is based on three text detection algorithms: connected component analysis (CCA) [16], maximally stable extremal regions (MSER) [17], and the EAST detector [12], which is based on a deep learning model. In preliminary experiments, different text localization methods were tested on datasets with images of diverse quality. Three selected techniques are based on different approaches to solving this problem. In most cases (79%) EAST detector was sufficient. But in some cases, where it failed, other methods (with a generally low success rate) were able to locate the inscription correctly.

Next, each of the probable text locations is used as an input for the optical recognition algorithm (Tesseract OCR) [18]. The OCR algorithm takes as its input an image containing some inscription and returns the inscription in a text format. The image must contain only the inscription, and the inscription must be horizontal. When the image is of poor quality, the returned text string contains a lot of artifacts. The inscription that has been found is then analyzed, artifacts are removed, and the text string is matched with entries

in the vessel registry. After all frames with the same ID have been analyzed, the method calculates the best identification results and stores these in the vessel status buffer (VSB).

The third layer is responsible for making the final identification decision for each tracking ID. When a vessel passes in front of the camera, i.e. frames with given tracking ID are not present in two consecutive rounds, the layer makes a final identification decision using round decisions stored in the VSB. The identification result is stored in the vessel identification history (VIH) database.

The VIM decides on identification using the maximum available information for a given tracking ID in the set of frames and is designed to correct results in various steps. Internally, the VIM uses six identification categories (namely, double full, full, high, low, multi, and none) that are finally reduced to three categories. Figure 1 illustrates the concept of decision making. The decision process involves three stages:

- Stage 1 (frame voting) uses three different text detection algorithms to increase the probability that all inscriptions are found in the case of low-quality images. The OCR engine is sensitive to the way in which the inscription is clipped from a frame. Even a few pixels' difference could provide better results; for example, one more character could be correctly recognized [4].
- Stage 2 (round voting) is based on the fact that each frame has a different view of a vessel. If identification contains errors, then repeated results could be used to give a better overall result. The

round identification could also be located on another computer.

- Stage 3 (final decision) takes information from all rounds and from the VIH database to produce the final result. If there is ambiguity, then there is a good chance that one of the vessels recently identified is the correct one. Also, this stage is separated from round voting owing to performance requirements.

## 2.2 Method description

The VIM is presented on a block diagram in Figure 2. The descriptions of the numbered blocks in the diagram are as follows:

1. Detection and tracking is based on the method proposed in [15]. In this step, vessels that pass in front of the camera are detected. The detection and tracking method works in intervals (rounds). Every 5 s, it returns information about the detected vessels or nothing. The detection method returns the following: camera ID number, number of detected vessels in the round, vessels' directions of movement with respect to the camera location, and vessel frames (a vessel frame is part of the picture frame clipped based on a vessel's bounding box). All vessel frames have assigned unique tracking IDs. Frames of the same vessel have the same tracking ID.
2. Round processing tries to identify all vessels in one round:
  - 2.1. Each vessel frame is sent to three different text detection algorithms (CCA, MSER, and the EAST detector). Each of these algorithms returns a set of bounding boxes with possible text locations.
  - 2.2. The images of the texts are clipped from the vessel frames based on the bounding boxes and added to one list.
  - 2.3. The OCR engine is used to extract a text string from each image from the list.
  - 2.4. In this step, strings from the OCR are pre-processed:
    - 2.4.1. Strings of length less than four characters are discarded.
    - 2.4.2. All characters except upper- and lower-case letters and numbers are replaced by an asterisk.
    - 2.4.3. If the number of asterisks is greater than 50% of the total length, the string is discarded.
    - 2.4.4. Asterisks at the beginning and end of the string are removed.
    - 2.4.5. Zeros are replaced with letters "O". This is done because some registration

numbers contain both "0" and "O" and in practice they are indistinguishable.

- 2.4.6. The strings are matched to known registration number masks (e.g., IMO number or Polish yacht registry number with "POL" prefix). When there is a match, the remaining strings are rejected.
- 2.5. Each string is categorized into one of the following identification categories:
  - 2.5.1. Full: an exact match is found with a name or registration number from the public registry.
  - 2.5.2. High: the match differs in only one character (including missing and additional letters).
  - 2.5.3. Low: the string has a length greater than five and the match differs in two characters.
  - 2.5.4. None: the string was not matched.
- 2.6. The best identification category (double full, high, low, multi, none) is chosen from Step 2.5. When two or more matches are found in the same category but on different vessels, the matches are integrated into the multi category. The only exception is a situation where two matches in the full category are matched to two different vessels. In such a case, they are integrated into the double full category. This is a borderline case when one vessel passes in front of another in the camera's field of view. The detection method from Step 1 cannot distinguish them and returns them as one.
- 2.7. The results from the previous step (identification category, ship(s) ID) are stored in a temporary buffer.
- 2.8. When all vessel frames from the round with the same tracking ID are processed, then the best result is chosen similarly to Step 2.6.
- 2.9. The identification category for each vessel from the round is stored in the VSB.
3. A final decision is made using the following steps:
  - 3.1. The update vessel status algorithm is triggered once every 5 s. The algorithm verifies the status of all vessels present in the VSB. The buffer contains activity counters for each vessel. When no vessel is present at the time of execution of the algorithm, the counter increases by one. When a vessel is present, it is set to zero.
  - 3.2. When the activity counter equals 3, the vessel is not visible in the camera and the final decision is made.



- 3.3. The final decision algorithm takes the round results for a given vessel and returns the final identification result (i.e., full, high, movement direction) using the following rules:
  - 3.3.1. The round results are sorted in descending order and stored in a list  $L_{rr}$ .
  - 3.3.2. When there is more than one vessel in the same identification category, they are merged into the multi category (see Step 2.6).
  - 3.3.3. When the first element in the list  $L_{rr}$  is double full, full, or high, it is returned as the final decision (double full is divided and returned as two full vessels).
  - 3.3.4. When the first element is in the low or multi category, all vessels from that category are counted; when one of the vessels has more occurrences than others, it is returned using the high identification category. Additionally, if in the previous five rounds a vessel from the list  $L_{rr}$  was identified, its counter is increased by one, since there is a high probability that it is the same vessel.
  - 3.3.5. In other cases, only the movement direction without vessel identification is returned (low and multi categories are only used internally).
- 3.4. The identification result (i.e., ship ID, identification category, time, camera number, movement direction) is stored in the VIH.
- 3.5. The information about the vessel is removed from the VSB.

### 3. Experimental Results

#### 3.1 Implementation

The method is divided into several components to enable scalable and efficient operation (Figure 3). The detection and tracking (Step 1) is implemented as the detection service. The detection service processes one video stream at a time. The text identification (Steps 2.1–2.8) is implemented as the identification service. The rest of the method is implemented in the system core. The system core receives the vessel frames from the detection services, sends them to the identification service, and makes a final decision based on the results from the identification service. It manages all the components. The system core does not contain any computationally intensive operations. In such an architecture, several detection services and identification services can run simultaneously to support

a large number of video streams. All components are implemented using C# and Emgu CV version 4.0.1 (C# wrapper for OpenCV), which includes the Tesseract OCR engine. Additionally, the EAST text detector is used. Moreover, several parts of the code use parallel processing.

In our experiments, the following computers were used:

1. Streaming device: Intel core i5-7440HQ, 16 GB RAM, SSD 500 GB, NVIDIA GeForce 930MX, Windows 10 Pro (laptop).
2. Detection service: Intel Core i7-8750H, 16 GB RAM, SSD 256 GB, NVIDIA GeForce GTX 1050Ti, Windows 10 Pro (laptop).
3. Identification service: Intel Core i7-8700K, 32 GB RAM, SSD 1 TB, NVIDIA Quadro P4000, Windows 10 Pro (workstation).
4. System core: laptop with a specification like that of the streaming device.

The network cameras were simulated by the VLC media player, which was used to stream video files over the network. This approach enabled us to repeat tests when necessary.

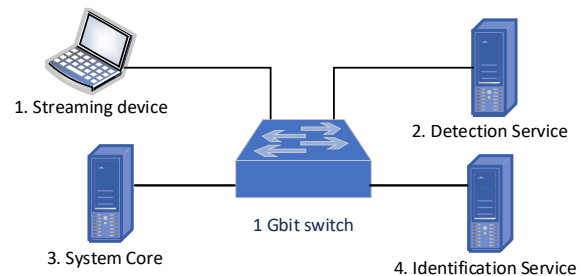


Figure 3. Method deployment

#### 3.2 Data Set

The dataset contained 30 video samples recorded near places where river monitoring system cameras were mounted. The samples were recorded using three different cameras (a GoPro Hero 6 camera (1), an AXIS IP camera Q1765-LE (2), and a Dahua IP camera IPC-HFW81230E-ZEH 12Mpx (3)) with different streaming qualities (10 samples for each camera). Figure 4 shows sample frames of video samples from each camera. It also shows the quality and size of the hull inscriptions used to identify sample frames. The qualities of the samples are as follows:

1. Full high-definition  $1920 \times 1080$ , 30 frames/s, bitrate 20 Mb/s, H.264 AVC Baseline@L4.
2.  $4000 \times 3000$ , 20 frames/s, bitrate 8 Mb/s, H.264 AVC High@L5.1.

3. Full high-definition 1920 × 1080, 25 frames/s, bitrate 2–7 Mb/s, H.264 AVC Main@L4.1.

The samples from cameras (2) and (3) were heavily compressed and contained a large number of visible compression artifacts. In some video samples, there was

more than one vessel. The number of vessels depended on the traffic on the waterway. The length of the video samples varied between 44 and 434 s.

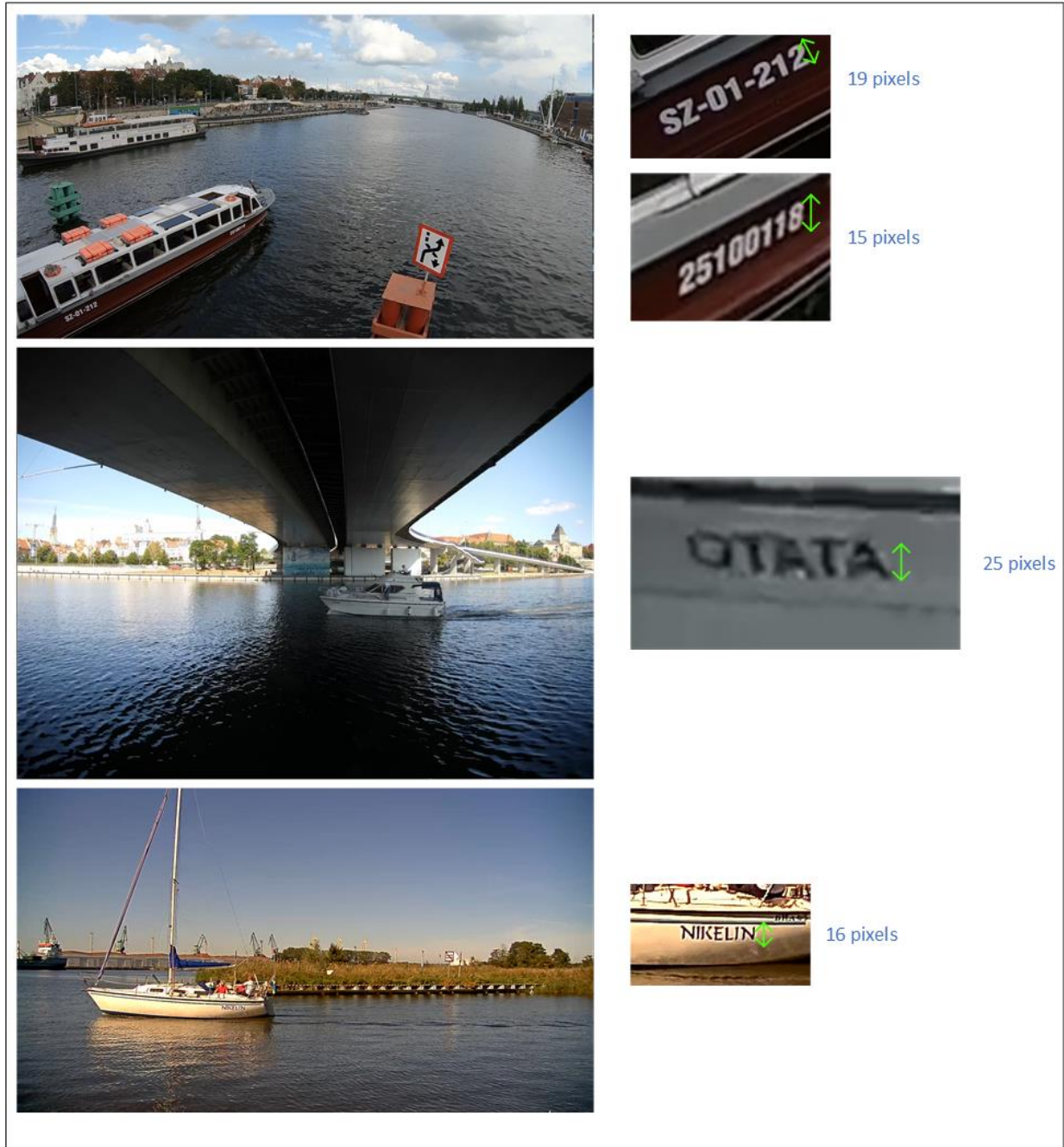


Figure 4. Frames from the dataset samples

### 3.3 Results

The dataset contained 53 vessels, from which the name or registration number of 36 vessels could be identified. Vessels could be identified once per passage in front of a camera. The movement direction could be identified for all 53 vessels.

The dataset results are presented in Table 1. The  $F_1$  score for the vessel identification was 0.971 and for the identification of the direction of movement was 0.972. The method identified 34 of 36 vessels (94%). In total, 25 (69%) were identified as full identification and 9 (25%) were identified as high identification. Two vessels were not identified. There were no incorrect full or high results. The movement direction was correctly identified for 52 vessels (98%). One vessel was not recognized (2%): it merged with another vessel during the detection and tracking stage. Additionally, there were two erroneous movement detections. Two vessels were detected twice and had two tracking IDs assigned.

**Table 1. Identification results for the vessel identification method**

Identification result	Number of vessels	Percentage of vessels
1) Vessel ID - correct	34	94%
a) Full - correct	25	69%
b) Full - incorrect	0	0%
c) High - correct	9	25%
d) High - incorrect	0	0%
e) Missing - incorrect	2	6%
2) Movement - correct	52	98%
a) Incorrect-additional	2	4%
b) Incorrect-missing	1	2%

### 4. Discussion

The method is designed in such a way that computationally intensive operations can be implemented in separate services that run independently. This allows scaling of the solution to a large number of video streams (20–100) by using more computational resources. The tests were performed by real-time video playback of prerecorded samples from a streaming device that emulated a surveillance camera.

The most time-consuming operations are text detection and text optical recognition, since these both use deep learning models. In our method, we do not use a complete frame, but only a clipped vessel frame,

which significantly reduces the text detection time. However, the OCR needed during identification remains the main bottleneck. The average identification time (using computer No. 3) is around 500 ms, or 900 ms when the frame does not contain a vessel (this time is longer because the method stops when a vessel is identified). Another bottleneck is caused by the need for background model updating during the detection process. The detection time (for a 1 s video period) is between 400 and 800 ms (using computer No. 2; the 800 ms is required for 4K resolution), when the background model is updated three times per second and the detection rate is one per second. The system core component processes each require less than 1 ms, because they do not involve any computationally intensive operations.

In the case of high-quality video streams, when hull inscriptions are clearly visible and contrasted with the background, it is possible to reduce the number of operations by using only one text detection algorithm. In our implementation, we also execute several parts of the algorithm in parallel, including text detection, OCR, and text parsing.

The method's robustness results mainly from the construction of the decision process. The ship passes in front of the camera in a few to tens of seconds. Only one frame from single passage with readable inscription located and extracted by any of three text location methods can be sufficient to identify a vessel. The VIM worked well when identifying commercial vessels, since their inscriptions are placed according to binding rules. The inscriptions are easy to recognize. In the tests, the method worked on most of the video samples, because their height exceeded 12 pixels. This is the threshold at which Tesseract starts to return good results.

One part of our data set consisted of video samples with a resolution of  $4000 \times 3000$  pixels and with a high compression rate. The test showed that such streams did not produce good results, since vessels' inscriptions were blurred by compression (sometimes even when they were 20 pixels in height) and it was not possible to recognize them. Such vessels are assigned to the movement direction identification category. The structure of the method involves several text detection algorithms and multiple frames to provide correct reading of inscriptions. The voting algorithm recognizes a vessel correctly when some text string is repeated in multiple vessel frames.

Additionally, the identification part of the method corrects two kind of errors made by the detection and tracking algorithms: (1) when two ships are visible in the vessel frame (this is the situation where both vessels are in the field of view of the camera); (2) when a vessel is split into two tracking objects (this is the situation, for



example, where a long barge slowly moves under a bridge and the background model incorrectly merges it with its background).

The proposed method differs significantly from existing methods of vessel identification mainly due to its applicability to recognize wide range of vessels of many types and therefore using different hull marking rules. Ferreira et al. [14] proposed a solution for fishing vessels, which first uses a classifier to recognize the bow of the vessel and then finds text regions using MSER algorithm. In our case, there is no easily distinguishable feature that would differentiated the process of text recognition for each type of a vessel. Method presented by Zhang et al. [11] also focused only on commercial ships in marine port of Dongying, China which have large and distinct IMO hull markings, locally often accompanied by Chinese characters. Therefore the authors enhanced the use of EAST detector by adding FCN to properly detect horizontal text lines for Chinese characters. Huang et al. [13] presented interesting end-to-end solution that includes solving arbitrary-oriented text issue, but the method is designed to work with clearly marked ships (naval warships) whose plates consist only of numbers and use large fonts.

## 5. Conclusion and future works

In the proposed method, some components (e.g., text detection and OCR) could be replaced with better approaches and some parameters could be adjusted to provide better output, e.g. when a new OCR engine that better recognizes low quality (small) characters will be available. In the future, we plan to do more testing with streams of different quality to verify which parameters produce the best results.

We expect to encounter two main challenges during the system deployment. In the experiments, the monitoring camera was emulated by using a streaming device that streamed video samples from the test data set. The system must work simultaneously with approximately 20 cameras. The system will be able to process that number of video streams, because it consists of several components that can have multiple instances. Only the system core must be run in a single instance, but it does not contain computationally intensive operations. The challenge is to minimize the total performance requirements, so the system will be more economically viable.

The second challenge is related to the camera location in relation to the sun position as direct sunlight into the camera greatly reduces video quality which influences identification quality. Because usually two or more cameras are in one place (e.g., on one bridge) pointing in different directions, we will probably create a mechanism that selects the camera providing the best

video quality. Optionally, we will create voting mechanism that will take as input identification results from a few cameras that can see the same vessel in a short period of time.

Moreover, we seek to conduct further testing to include integration with RIS systems. Additionally, we plan to extend the VIM method and add a vessel classification module. The vessels will be classified into one of the categories if their identification is negative. In such case, the system will gather statistical information about a number of vessels from a given category (e.g. motorboat, barge, yacht, etc.) that passed in front of the camera and also about their direction of movement.

## 6. Acknowledgement

This scientific research work was supported by the National Centre for Research and Development (NCBR) of Poland under grant No. LIDER/17/0098/L-8/16/NCBR/2017.

## 7. References

- [1] B. Tetreault, "Use of the Automatic Identification System (AIS) for Maritime Domain Awareness", Proceedings of OCEANS 2005 MTS/IEEE, Washington D.C., USA, vol. 2, September 2005, pp. 1590-1594.
- [2] G. Schilk, L. Seemann, "Use of ITS Technologies for Multimodal Transport Operations – River Information Services (RIS) Transport Logistics Services", *Procedia - Social and Behavioral Sciences*, vol. 48, 2012, pp. 622-631.
- [3] International Maritime Organization, "Resolution A.600(15): IMO Ship Identification Number Scheme". Resolutions and Other Decisions (Resolutions 596 - 635): Assembly, Fifteenth Session, 9-20 November 1987, ISBN 92-801-1234-1, 1988, pp. 17–20.
- [4] S. Long, X. He, and C. Ya, "Scene text detection and recognition: The deep learning era", *CoRR*, abs/1811.04256, 2018, pp. 1-26.
- [5] R. da Silva Moreira, N. F. F. Ebecken, A.S. Alves, F. Livernet, and A. Campillo-Navetti, "A survey on video detection and tracking of maritime vessels", *Int. Jour. of Res. and Rev. in App. Sci.*, vol. 20 (1), 2014, pp. 37–50.
- [6] W.-C. Hu, C.-Y. Yang, and D.-Y. Huang, "Robust real-time ship detection and tracking for visual surveillance of cage aquaculture", *Jour. of Vis. Comm. and Image Repr.*, vol. 22 (6), 2011, pp. 543–556.
- [7] Z. L. Szapak, and J. R. Tapamo, "Maritime surveillance: Tracking ships inside a dynamic background using a fast level-set", *Expert Systems with Applications*, vol. 38 (6), 2010, pp. 6669–6680.
- [8] Y. J. Kim, Y. K., Chung, and B.G. Lee, "Vessel tracking vision system using a combination of Kalman filter, Bayesian classification, and adaptive tracking algorithm", *Proceedings of 16th International*

- Conference on Advanced Communication Technology, 2014, pp. 196–201.
- [9] N. Kaido, S. Yamamoto, and T. Hashimoto, “Examination of automatic detection and tracking of ships on camera image in marine environment”, In Proc. Of 2016 Techno-Ocean, 2016, pp. 58–63.
- [10] Z. Liu, J. Hu, L. Weng, and Y. Yang, “Rotated region based CNN for ship detection”, Proc. IEEE Int. Conf. Image Process., Sep. 2017, pp. 1-5.
- [11] W. Zhang, H. Sun, J. Zhou, X. Liu, Z. Zhang, and G. Min, “Fully Convolutional Network Based Ship Plate Recognition”, 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 1803-1808.
- [12] X. Zhou, C. Yao, H. Wen, Y. Wang, S. Zhou, W. He, and J. Liang, “EAST: An Efficient and Accurate Scene Text Detector”, In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017, pp. 2642-2651.
- [13] H. Huang, X. Xu, X. Xia, and Y. Zhang, “End-to-End Vessel Plate Number Detection and Recognition Using Deep Convolutional Neural Networks and LSTMs”, 2018 11th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 2018, pp. 195-199.
- [14] J. C. Ferreira, J. Branquinho, P.C. Ferreira, and F. Piedade, “Computer vision algorithms fishing vessel monitoring—identification of vessel plate number”, Ambient Intelligence– Software and Applications – 8th International Symposium on Ambient Intelligence (ISAmI 2017), Advances in Intelligent Systems and Computing, De Paz, J. F, et al. (eds.), Springer International Publishing, vol. 615, 2017, pp. 9-17.
- [15] N. Wawrzyniak, T. Hyla, and A. Popik, “A Vessel Detection and Tracking Method Based on Video Surveillance”, Sensors, vol. 19 (23), 2019, art. no. 5230.
- [16] S. Mule, S. N. Holambe, “Detecting Text in Natural Scenes with Connected Component Clustering and Nontext Filtering”, IRJET, vol. 03(09), 2016, pp 625-629.
- [17] J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions”, Image and Vision Computing, vol. 22, 2004, pp. 761-767.
- [18] C. Patel, C.; A. Patel, and D. Patel, “Optical character recognition by open source OCR tool tesseract: A case study”, Int. J. Comput. Appl., vol. 55, 2012, pp. 50–56.