

RESEARCH PAPER

Open Access



Generic and attribute-specific deep representations for maritime vessels

Berkan Solmaz^{*†} , Erhan Gundogdu[†], Veysel Yucesoy and Aykut Koc

Abstract

Fine-grained visual categorization has recently received great attention as the volumes of labeled datasets for classification of specific objects, such as cars, bird species, and air-crafts, have been increasing. The availability of large datasets led to significant performance improvements in several vision-based classification tasks. Visual classification of maritime vessels is another important task, assisting naval security and surveillance applications. We introduced, MARVEL, a large-scale image dataset for maritime vessels, consisting of 2 million user-uploaded images and their various attributes, including vessel identity, type, category, year built, length, and tonnage, collected from a community website. The images were categorized into vessel type classes and also into superclasses defined by combining semantically similar classes, following a semi-automatic clustering scheme. For the analysis of the presented dataset, extensive experiments have been performed, involving several potentially useful applications: vessel type classification, identity verification, retrieval, and identity recognition with and without prior vessel type knowledge. Furthermore, we attempted interesting problems of visual marine surveillance such as predicting and classifying maritime vessel attributes such as length, summer deadweight, draught, and gross tonnage by solely interpreting the visual content in the wild, where no additional cues such as scale, orientation, or location are provided. By utilizing generic and attribute-specific deep representations for maritime vessels, we obtained promising results for the aforementioned applications.

Keywords: Fine-grained object categorization, Naval surveillance, Deep representations for maritime vessels, Maritime vessel attributes, Convolutional neural networks, Deep learning

1 Introduction

The coastal and marine surveillance systems are mainly based on sensors such as radar and sonar, which allow detecting marine vessels and taking responsive actions. Vision-based surveillance systems containing electro-optic imaging sensors can also be exploited for developing robust and cost-effective systems. Categorization of maritime vessels is of utmost importance to improve the capabilities of such systems. For a given image of a ship, the goal is to automatically identify it using computer vision and machine learning techniques. Vessel images include important clues regarding different attributes such as vessel type, category, gross tonnage, length and draught. A large-scale dataset would be beneficial for extracting

such clues and learning compelling models from images containing several types of vessels.

Presence of benchmark datasets [1] with large quantities of images and manual labels with meaningful attributes has resulted in a significant increase in visual object categorization performance by allowing the use of convenient machine learning methods such as deep architectures [2]. Later, these powerful deep architectures have been employed in a more challenging problem, fine-grained visual categorization, by either training on datasets from scratch [3], by fine-tuning deep architectures trained on large-scale datasets [4], or by exploiting the previously trained architectures with specific modifications [5].

To classify images with a fine-grained resolution, a considerable amount of training data is necessary for a respectable model generalization. Thus, fine-grained datasets were collected for specific object categories. Some examples are aircraft datasets [6, 7]; Caltech-UCSD bird species dataset [8] consisting of 12 K images, car

*Correspondence: bsolmaz@aselsan.com.tr

[†]Equal contributors

Aselsan Research Center, Ankara, Turkey

make, and model datasets; Stanford cars dataset [9] containing 16 K car images; and CompCars dataset [10] of 130 K images. One work related to marine vessel recognition is [11], where 130,000 random example images from the *Shipspotting* website [12] is utilized and a convolutional neural network [2] is trained for classifying vessel types. In our dataset, 140,000 images are engaged for vessel type classification among 26 superclasses constructed using a semi-supervised clustering approach. Furthermore, constructed vessel superclasses are balanced; the training set is arranged to have an equal number of examples from each superclass, after augmenting data for vessel type classes with lower number of examples. However, there is a significant imbalance of examples among the classes in [11], which may result in a bias in classification towards the dominant classes with more examples. Hence, imbalance makes it more difficult to validate the performance of different classifiers. In this work, for measuring vessel classification performance, we report mean per class accuracies. In addition, we accomplish further important tasks with a vast amount of vessel images and obtain pleasing results, which will be described in details in the following sections.

In order to utilize the state-of-the-art fine-grained visual classification methods for maritime vessel categorization, we collected a dataset consisting of a total of 2 million images downloaded from the *Shipspotting* website [12], where hobby photographers upload images of maritime vessels and corresponding detailed annotations including types, categories, tonnage, draught, length, summer deadweight, year built, and International Maritime Organization (IMO) numbers, which uniquely identify ships. To the best of our knowledge, the collected dataset, MARitime VESseLs (MARVEL) [13, 14], is the largest-scale dataset with meta-data composed of the aforementioned attributes, suited for fine-grained visual categorization, recognition, retrieval, and verification tasks, as well as any possible future applications.

In addition to the introduced large-scale dataset, our other major contributions are presenting generic representations for maritime vessels, as well as targeting visual vessel analysis from five different aspects: (1) *vessel type classification*, (2) *vessel identity verification*, (3) *vessel retrieval*, (4) *vessel identity recognition with and without prior type knowledge*, and (5) *specific vessel attributes (draught, length, gross tonnage, and summer deadweight) prediction and classification*. To verify the practicality of MARVEL and encourage researchers, we present baseline results for these tasks. By providing relevant splits of the dataset for each application and inspecting the consistency of associated labels, we form a comparison basis for visual analysis of maritime vessels. Moreover, we believe our structured dataset will be a benchmark for evaluating approaches designed for fine-grained recognition.

The researchers may also develop several new applications with the help of this dataset in addition to the aforementioned applications.

2 MARVEL dataset properties

MARVEL dataset consists of 2 million marine vessel images collected from *Shipspotting* website [12]. For most of the images in the dataset, the following attributes are available: *beam*, *year built*, *draught*, *flag*, *gross tonnage*, *IMO number*, *name*, *length*, *category*, *summer deadweight*, *MMSI*, *vessel type*.

Among the above attributes, we observe that the most useful and visually relevant categories are as follows: (1) *Vessel type*, (2) *category*, (3) *draught*, (4) *gross tonnage*, (5) *length*, (6) *summer deadweight*, and (7) *IMO number*. *Vessel type* is assigned based on the type of cargo a vessel will be transporting. For instance, if a vessel carries passengers, its type is very likely to be a *Passengers Ship*. The dataset contains 1,607,190 images with valid annotated type labels belonging to one of 197 categories. *Vessel type* histogram, highlighting the major categories, is depicted in Fig. 1c. Another available attribute is *category*, which is another vessel description. Example *categories* with a substantial number of members are *chemical and products tankers*, *containerships built 2001–2010*, and *Tugs* (please see Fig. 1a). All collected images have been assigned a *category* out of 185 categories in MARVEL dataset. *IMO number* is another category, which is an abbreviation for International Maritime Organization number. Similar to the chassis numbers of cars, IMO numbers of vessels uniquely identify the ships registered to IMO regardless of any changes made in their names, flags, or ownerships. Of the collected images, 1,628,056 are annotated with IMO numbers (please refer to Fig. 1b). There are a total of 103,701 unique IMO numbers in MARVEL dataset.

Considering the fact that images which have been assigned identical IMO numbers belong to the same vessels, we are able to check the consistency of other attribute annotations and fill out the missing entries when necessary. First, zero or invalid entries are discarded. Next, we convert all attribute labels to metric unit system to account for the presence of some labels in an imperial system. Finally, we maintain the consistency of labels for each vessel separately by applying median filters on available annotations. Engaging such preprocessing procedures, we obtain very large groups of images that include valid attribute labels. The attributes we focus on are IMO number, vessel type label, draught, gross tonnage, length, and summer deadweight (Fig. 2). For *draught*, an attribute which is defined as the vertical distance between the bottom of vessel hull and waterline, there are 1,067,946 images carrying validated labels. *Gross tonnage* is a unitless index calculated using the internal volume of vessels.

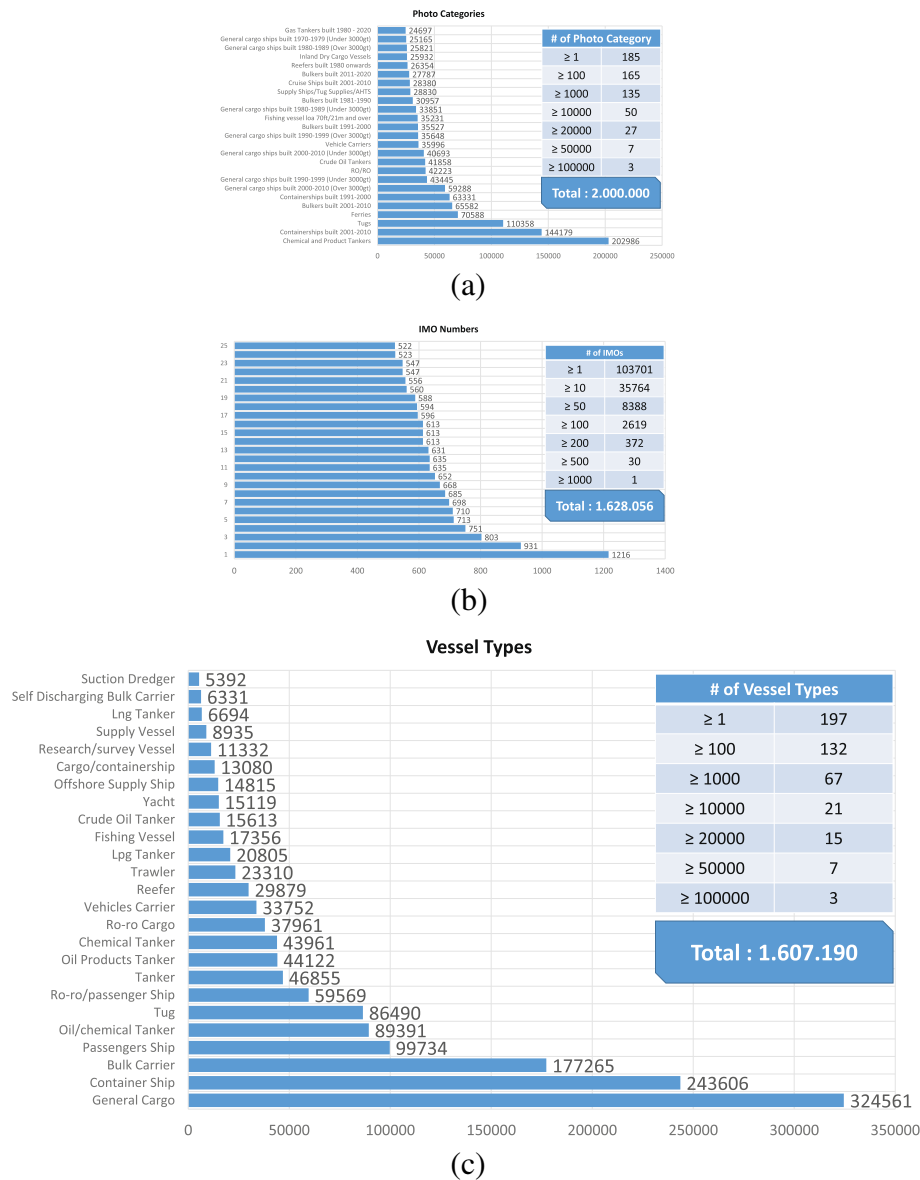
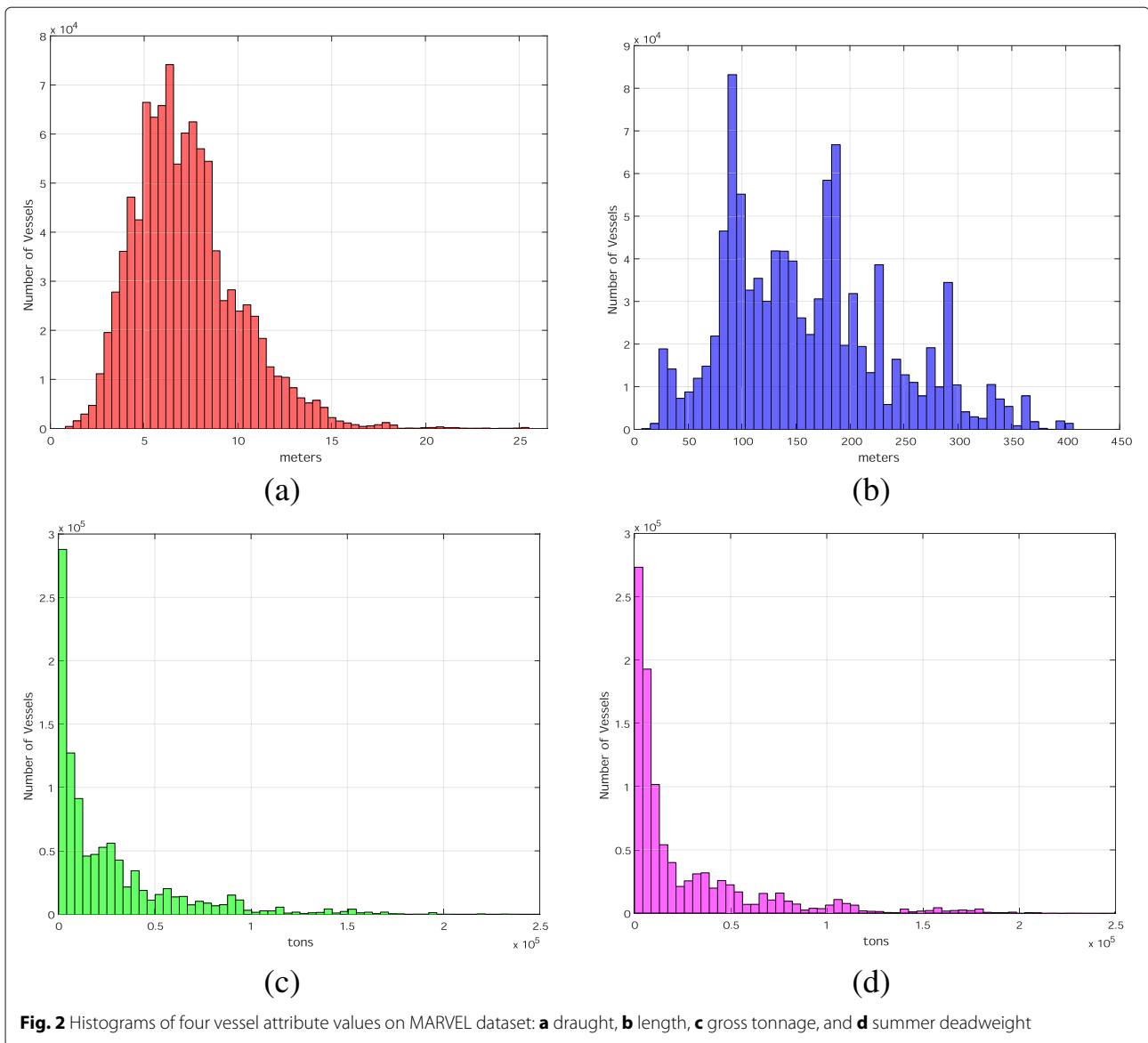


Fig. 1 Distribution of collected vessel images: Number of images belonging to each photo category, individual vessel, and vessel type are depicted in **a**, **b**, and **c**, respectively. The largest group among photo categories is chemical and product tankers. General cargo is the vessel type including highest number of images. Further statistics are provided on the right columns: In **b**, 8388 marine vessels are present containing at least 100 images. In **c**, there are 132 vessel type categories including at least 100 images

There are 1,583,882 images with valid annotated labels for gross tonnage. Validated annotations for *summer deadweight*, a measure of carrying capacity of a ship, are provided for 1,508,974 of all images. *Length* data of the maritime vessels are made available for 1,107,907 images. In summary, when combined, a total of 1,006,868 images retain valid annotated labels for all vessel type, IMO number, draught, length, summer deadweight, and gross tonnage attributes.

3 Potential computer vision tasks on MARVEL dataset

Huge quantity of images and their annotations, existing in MARVEL, makes it applicable to directly employ recent methods utilizing deep architectures such as AlexNet [2] for vessel categorization. One may choose one of the provided vessel attributes such as *vessel type* or *category* and apply classification methods for categorizing images according to the selected attribute.



In MARVEL there are more than 8000 unique vessels (carrying unique IMO numbers) having more than 50 example images as shown in Fig. 1b. It is also feasible to use the dataset for both vessel verification and identity recognition, which could be a vital part of a maritime security system, analogous to a scenario where vehicle make and model recognition is crucial for a traffic security system.

The main foci of this study on MARVEL dataset are five folds: (1) *vessel classification* since content of cargo that a ship carries, specified by its type, is crucial for maritime surveillance, (2) *identity verification* where the ultimate goal is to find out if a pair of images belong to the same vessel with a unique IMO number, (3) *retrieval* where one might desire to query a vessel image and retrieve a certain

number of similar images from a database, (4) *identity recognition* which is a challenging though interesting task which aims at recognizing a specific vessel within vessels of same type or among all other vessels (This might be likened to a facial recognition task.), and finally (5) *specific attribute prediction and classification*, where the objective is to grasp draught, length, gross tonnage, and summer deadweight of a vessel by simply analyzing the 2-D visual content. With an aim to achieve these goals, we design generic and attribute specific representations which are powerful in describing marine vessel images.

For *vessel classification*, one of the most important tasks, we first generate a set of superclasses which may contain vessels of more than one type, since some subsets of *vessel types* are not visually distinguishable even

with human supervision. The sole differences within the subsets arises from the invisible content of cargo rather than the visual appearance of ships. A concrete example of such a case arises for the pair of vessel types: *crude oil tanker* and *oil products tanker*, which is illustrated in Fig. 3. Although the two vessel types have distinct functional differences, their visual characteristics are congruent especially when images are captured by cameras located far away from these vessels; when the vessels occupy a small portion of images and their decks are not visible from such a view point, it is tough to distinguish them. Hence, we merge some of the types to generate superclasses which are semantically correct and visually discriminable. In Section 4, we describe the details for combining *vessel types*. As inspired by [15], the presence of multi-level relevance information and hierarchical grouping of vessels may allow exploitation of MARVEL dataset for a further performance improvement for particular marine vessel recognition tasks in the future.

Vessel verification task serves for deciding whether a pair of vessel images belong to the same vessel or not. This may be beneficial for a naval surveillance scenario, where a specific vessel is required to be tracked using an electro-optic imaging system.

For the task of *vessel retrieval* relating to *vessel classification*, the goal is to retrieve images belonging to providing a query image, several images with similar content are retrieved from the database.

Vessel recognition aims at revealing the accurate identity of a vessel by analyzing an unseen example image of it and finding out the matching vessel within a group of

vessels. This task may be particularly useful for scenarios of marine surveillance and port registration. For this task, first, we performed recognition for vessels considering their type labels, for instance, identifying a passenger ship among other passenger ships. Next, we attempt a more challenging recognition problem, identifying vessels where no additional cues such as *vessel type labels* or *category labels* are given.

Moreover, as novel problems, we attempt tasks of *predicting and classifying vessel attributes*: draught, gross tonnage, length, and summer deadweight. The objective here is to quantify these attributes based on 2-D visual content only, which may ameliorate the practicality of coastal surveillance systems, since that avoids the need for retaining meta-data for optical systems, namely camera parameters, camera position, and distance to the vessel, while estimating physical dimensions of a vessel based on its appearance. Another beneficial use of this task may be for safe marine traffic routing as well as for the calculation of port access and transit fees, when vessel dimensions need to be known. Furthermore, there are studies, proving that presence of attribute-based representations are helpful for several computer vision tasks including object recognition [16], detection [17], and identification [18]. The attribute-based learned representations for marine vessels in this work may be utilized in a similar fashion aiding other visual analysis tasks.

4 Superclasses for vessel types

To generate superclasses from *vessel types*, the first 50 major *vessel types* containing the largest amount of



Fig. 3 Visual comparison of two very similar classes: *crude oil tanker* (top row) and *oil products tanker* (bottom row)

example images are selected and sorted according to their quantity. The *vessel type* with the largest number of images which is employed in our superclass generation, is *general cargo*, consisting of 324,561 example images. The class with the smallest number of images is the *timber carrier*, accommodating only 1837 images. In this work, to investigate the visual similarities among *vessel types*, MatConvNet Toolbox [19] implementation of a pre-trained convolutional neural network (CNN) architecture, VGG-F [20], is adopted. Features are extracted posterior to resizing images to 224×224 . Utilizing the penultimate layer activations of VGG-F [20] as visual representations of images, each image is described by a 4096-dimensional feature vector. Based on these feature vectors, we calculated a dissimilarity matrix for the 50 major vessel classes. To generate superclasses, 1/10 of all collected images belonging to 50 major classes are randomly selected (approximately 130,000 images) and individual class statistics are estimated. Prior to calculating a dissimilarity matrix, we removed outliers following the preprocessing step explained below.

4.1 Outlier removal

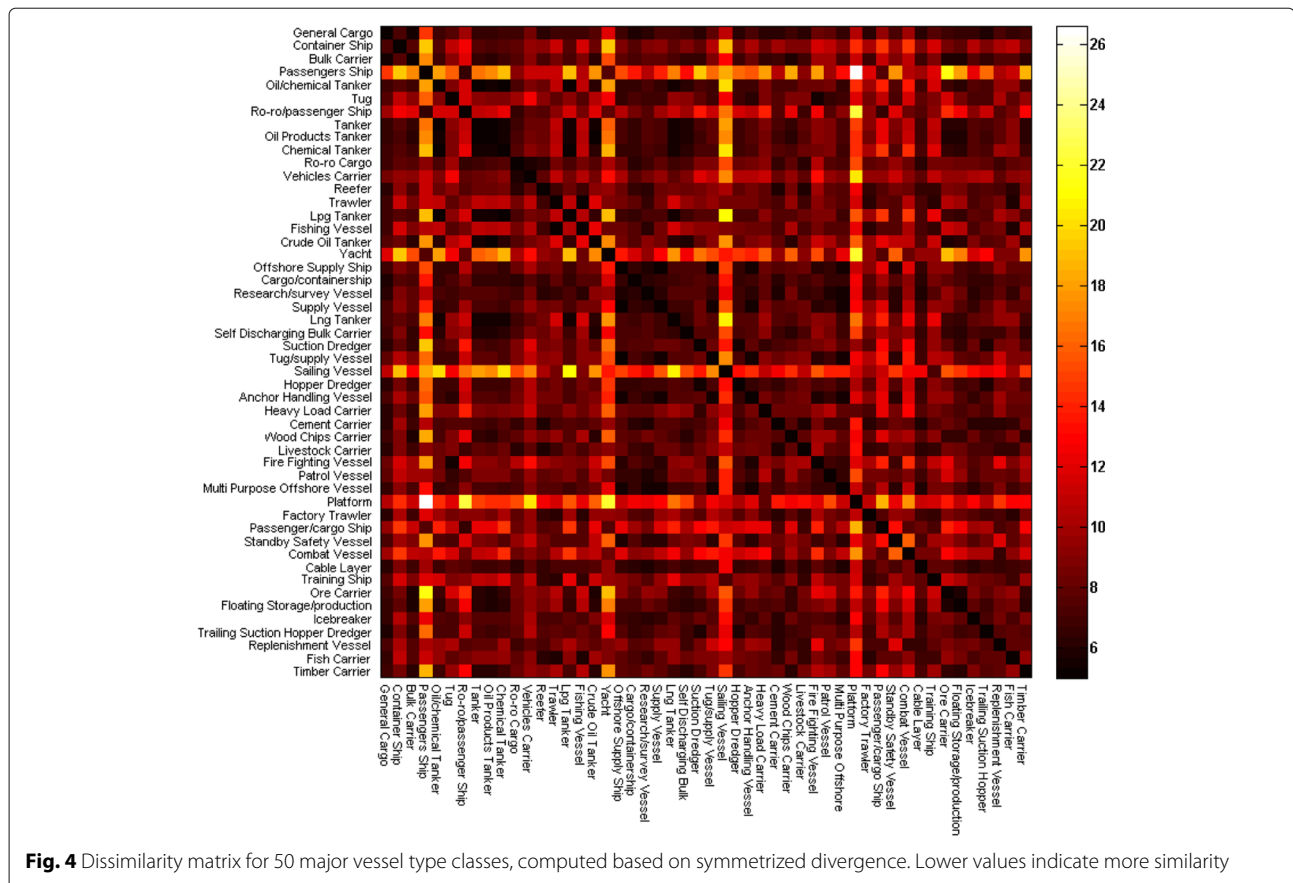
Although image annotations for most categories are valid and correct, interior images of vessels are also present in MARVEL dataset. Thus, we prune outliers within individual *vessel types* and avoid them while computing the dissimilarity matrix. First, feature vector dimensionality is reduced to 10 by principal component analysis (PCA) using all examples of 50 major vessel type classes, since Kullback-Leibler divergence is utilized in dissimilarity computation and determinants of very high dimensional matrices become unbounded. After dimensionality reduction, each vessel type class is processed independently and Gaussian distributions are fitted; means and covariances of each distribution are estimated. The feature vectors of corresponding classes are whitened to obtain unit variance within each class. We intent to filter out unlikely examples in the dataset to obtain a clear dissimilarity matrix. Next, we utilize χ^2 distribution since the dataset is already whitened. For each example in individual classes, the sum of the square values of the 10-dimensional feature vectors are used as samples drawn from the χ^2 distribution with 10° of freedom. Cumulative distribution function (cdf) value for each sample is calculated and removed from the class set if the cdf value is greater than 0.95, which corresponds to the samples drawn from the 5% tail of the χ^2 distribution.

4.2 Dissimilarity matrix and superclass generation

Once outliers are removed from each vessel type class by the above procedure, the remaining examples are used to

compute a dissimilarity matrix. We compute symmetrized divergence as the dissimilarity index. Symmetrized divergence $D_S(P, Q)$ of two classes, namely P and Q , is defined as $D_S(P, Q) = \frac{1}{2}D_{KL}(P||Q) + \frac{1}{2}D_{KL}(Q||P)$, where $D_{KL}(\cdot||\cdot)$ stands for Kullback-Liebler divergence of two multivariate Gaussian distributions. The computed dissimilarity matrix is depicted in Fig. 4.

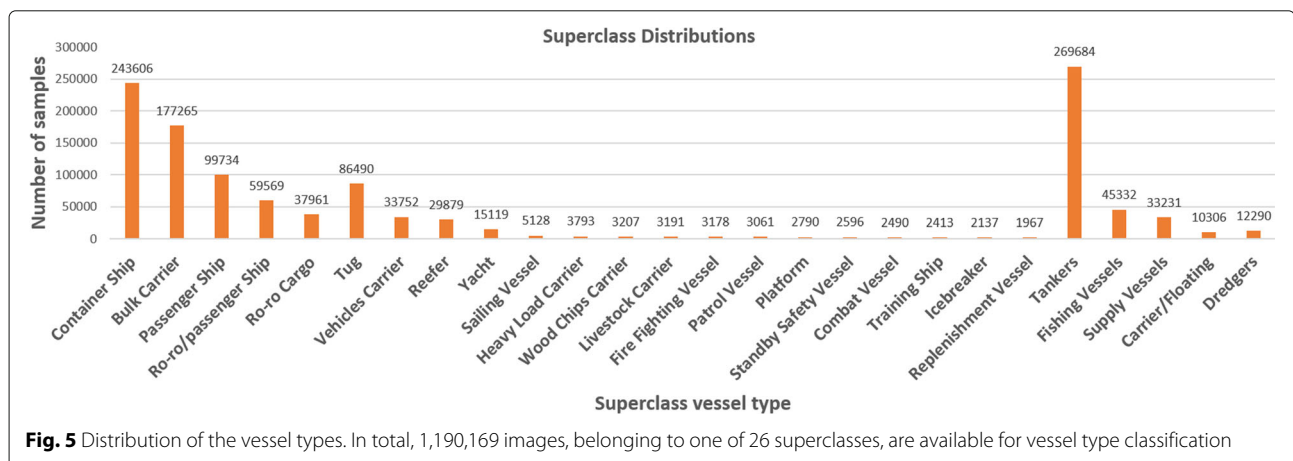
By exploiting the dissimilarity matrix, we merge similar vessel type classes using a threshold. Prior to thresholding, we applied spectral clustering methods with the help of the dissimilarity matrix. Nevertheless, the resulting groups were not semantically meaningful. Hence, we opt to continue by increasing the threshold for the similarities of the pairs of classes (i.e., this corresponds to each entry of the dissimilarity matrix). If dissimilarity index of a pair of classes is below a threshold, the pair is assigned to the same superclass. We keep increasing the threshold before it reaches to a point where semantically irrelevant classes (human supervision is adopted here) start to merge, and we define it as the final threshold for clustering. The majority of the resulting superclasses contain reasonable classes. The generated vessel type superclasses with more than one *vessel type* are (1) *tankers* (consisting of *oil products tanker*, *oil/chemical tanker*, *tanker*, *chemical tanker*, *crude oil tanker*, *lpg tanker*, *lng tanker*, *ore carrier*), (2) *carrier/floating* (consisting of *timber carrier*, *floating storage production*, *self discharging bulk carrier*), (3) *supply vessels* (which contain *offshore supply ship*, *supply vessel*, *tug/supply vessel*, *anchor handling vessel*, *multi purpose offshore vessel*), (4) *fishing vessels* (which include *trawler*, *fishing vessel*, *factory trawler*, *fish carrier*), and (5) *dredgers* (which contain *suction dredger*, *hopper dredger*). Finally, marginal adjustments are done manually to make all superclasses as meaningful as possible. These adjustments include merging superclass containing only *trailing suction hopper dredger* with superclass consisting of *Suction Dredger* and *Hopper Dredger*. In addition, seven *vessel types* are removed entirely from the set of superclasses. The classes to be eliminated are decided according to the average dissimilarity of the classes to the rest. The salient overall dissimilarity scores are detected manually. The removed classes are, namely (1) *general cargo* (it is significantly confusing with the *container ship* and *ro-ro cargo*), (2) *cargo/containership*, (3) *research/survey vessel*, (4) *cement carrier*, (5) *multi purpose offshore vessel*, (6) *passenger/cargo ship*, and (7) *cable layer*. The removed classes both visually and functionally contain more than at least two separate classes, i.e., *passenger/cargo ship* involve both passenger vessels and general cargo vessels. The merged classes with thresholding also contain visually very meaningful *vessel types*, i.e., all of the fish-related vessels are clustered within the same superclass. The distribution of final 26 superclasses can be viewed in Fig. 5.



4.3 Superclass classification

As demonstrated in Fig. 5, there exists an imbalance between superclasses. Nevertheless, even the superclass with the least amount of examples has a large quantity of examples. Therefore, to classify superclasses of vessels, it is feasible to train a deep CNN architecture AlexNet [2]. To avoid the imbalance between superclasses, we acquire equal numbers of samples from each class for both training and testing, as 8192 and 1024 images,

respectively. For superclasses with examples less than the required amount, we generate additional examples by data augmentation (using different croppings of images). Consequently, our training and test sets contain 212,992 and 26,624 examples, respectively, although we have 140,000 unique examples. We should also note that no images of the same vessels appear in both training and test sets. The classification performance is quantified by the help of a normalized confusion matrix [7]. The practical +



metric for a fine-grained classification task can be the class-normalized average classification accuracy, which is calculated as the average of diagonal elements of a normalized confusion matrix, C , entries of which are defined as follows [6]:

$$C_{pq} = \frac{|\{i : \hat{y}_i = q \wedge y_i = p\}|}{|\{i : y_i = p\}|}, \quad (1)$$

where $|\cdot|$ denotes the cardinality of the set, \hat{y}_i indicates the estimated class label, and y_i is the actual label for the i^{th} training example. The final performance measure is the mean of the diagonal elements of the matrix C . This value for 26 superclasses is 73.14% for the normalized confusion matrix depicted in Fig. 6. To emphasize the validity and efficacy of the learned network, we also compare it with another method utilizing multi-class support vector machine (SVM) with the Crammer and Singer multi-class SVM [21] implementation of [22] in LIBLINEAR [23] library. The feature vectors for training SVM are extracted from the VGG-F network of [20], their dimensionality is reduced to 256, and PCA whitening is applied. Due to memory requirements and computational complexity in optimization, we use half of the training set. We report the class-normalized average classification accuracy in testing as 53.89%. Compared to the use of pre-learned VGG-F weights with an SVM classifier, AlexNet trained from scratch has 35% improvement in accuracy.

5 Experiments on potential applications

In this section, we make use of our dataset, MARVEL, for potential maritime applications and vessel verification, retrieval, identity recognition, and attribute prediction and classification. In the following subsections, these applications and necessary experimental settings are explained.

During all experiments, we follow training and testing strategies similar to [10]. First, 8000 vessels with unique IMO numbers are selected such that each vessel will have 50 example images, resulting in a total of 400,000 images. This data is divided into two splits: training and testing. The training set consists of 4035 vessels (201,750 example images in total), and the test set contains 3965 vessels (198,250 example images in total). There exist 109 vessel type labels among 400,000 examples, and training and test sets are split in a way that the number of vessel types are identical in both sets. In the rest of the paper, we call the training split of this subset as IMO training set, and the test split as IMO test set.

We propose three deep CNN-based generic representations for marine vessels on IMO training set by making use of *vessel type* and/or *vessel IMO* labels. Hence, we train the same architecture of [2] as in *vessel classification* task and modify it accordingly with an aim to capture more details in vessel images: For the last layer, rather than 26 label classes, we use 109, 4035, and 4144 label classes.

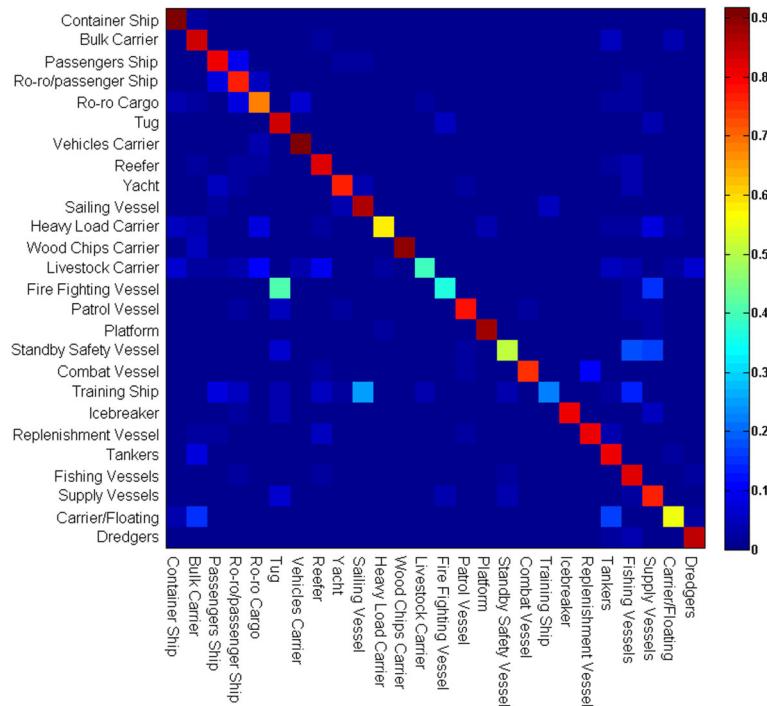


Fig. 6 Normalized confusion matrix for categorization of 26 superclasses representing vessel types. Accuracy, computed by averaging diagonal entries, is 73.14%

These three different classifiers focus on discriminating vessel types, vessel IMO numbers (classifying individual vessels on IMO training set), and both vessel types and IMO numbers (jointly classifying type and IMO numbers of vessels on IMO training set), respectively. We compare the performances of these three representations over computer vision tasks, which are described below in details.

Deep representations for example images are extracted as the penultimate layer activations of the trained networks (as in the superclass generation part in Section 4) with 4096 dimensions. More discriminative features being desired, we extract the penultimate layer activations prior to the rectified linear unit (ReLU) layer, which carry more information than the layer after ReLU since the negative values are cast to zero after ReLU. This choice makes our vessel verification performance better than the case with the deep representations after ReLU case.

During all experiments utilizing convolutional neural networks, we select batch sizes as 256 without normalization and decaying learning rates, consisting of logarithmically equally spaced values between 0.01 and 0.0001. For superclass classification, we train the networks for 60 epochs and for attribute classification and prediction, we train the networks for 50 epochs, since we notice that the training error does not decrease with further training. The implementation of the networks are based on the MatConvNet Toolbox [19].

5.1 Vessel verification, retrieval, and recognition

5.1.1 Vessel verification

Akin to face verification [24], car model verification is applied in CompCars dataset [10] to serve for conceivable purposes in transportation systems. That kind of task is claimed to be more complicated compared to face verification, since car model verification is performed on images with unconstrained viewpoints. On MARVEL dataset, we perform maritime *vessel verification* where the attribute to be verified is the vessel identity. *Please note that our task is more challenging compared to identifying other attributes such as category or vessel type.* Furthermore, this problem is more challenging than both car model and face verification tasks, since it is desired to identify/verify pairs of individual vessels by looking only at their appearances which have more diversity.

After extracting the generic deep representations (109 and 4144-dimensional output based), 50,000 positive pairs (belonging to same vessels) and 50,000 negative pairs (belonging to different vessels) are picked randomly from both training and test splits out of 201,750 training examples and 198,250 test examples, respectively¹. For all 400,000 training and testing examples, feature vector dimensionality is reduced to 100 by PCA exploited with only training examples. Moreover, all 100-dimensional

examples are PCA whitened since whitening increases performance of SVM classifier. Concatenating two 100-dimensional vectors, we describe each pair of vessel during verification experiments. Finally, for each generic representation, we train a binary SVM classifier with a radial basis function kernel on the generated training set by using the implementation of LIBSVM library [25]. Additionally, we attempt end-to-end learning for verification. For this experiment, we construct a Siamese neural network, based on AlexNet architecture, with shared weights, and added a contrastive loss layer after the last fully connected layers. Contrastive loss [26], incurring for similar and dissimilar pairs of images is defined as,

$$L = (1 - Y) \frac{1}{2} (D_W)^2 + Y \frac{1}{2} \{ \max(0, m - D_W) \}^2 \quad (2)$$

where Y is a binary label, assigned to 1 for similar images, otherwise set to 0. $m > 0$ is a margin set for dissimilar pairs, and D_W is the distance to be learned for pairs of images, \vec{X}_1 and \vec{X}_2 . D_W is calculated as the Euclidean distance between outputs of parametrized function G_W .

$$D_W(\vec{X}_1, \vec{X}_2) = \|G_W(\vec{X}_1) - G_W(\vec{X}_2)\|_2 \quad (3)$$

The precision recall curves for the two generic representations and the Siamese network-based representation, obtained by varying the classification thresholds, are plotted in Fig. 7. We also compare the performance of SVMs with the nearest neighbor (NN) classifiers. For NN classifier, each test pair is assigned the label of the training pair for which the Euclidean descriptor distances are the smallest. The resulting precision and recall values of SVM and NN classifier are presented in Table 1. All classifiers

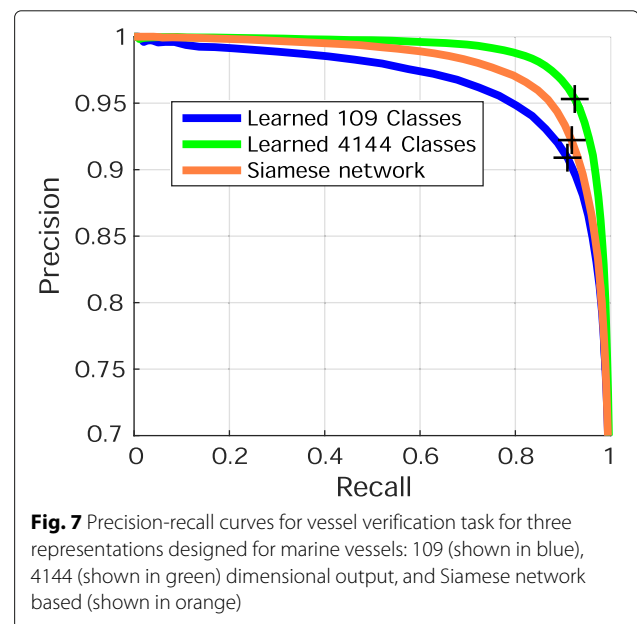


Table 1 Vessel verification results on 50,000 positive pairs and 50,000 negative pairs of vessels for the nearest neighbor and SVM classifiers by utilizing the generic and end-to-end learning-based vessel representations learned in IMO training set, which does not contain any images of the vessels in IMO test set

	Representation	True positives	True negatives	False positives	False negatives	Accuracy	Precision	Recall
NN	109-dimensional output based	44,978	40,198	9,802	5,022	85.18%	0.82	0.90
SVM	109-dimensional output based	45,503	45,422	4,578	4,497	90.93%	0.91	0.91
NN	4144-dimensional output based	47,305	41,148	8,852	2,695	88.45%	0.84	0.95
SVM	4144-dimensional output based	46,225	47,744	2,256	3,775	93.97%	0.95	0.92
NN	Siamese network based	44,459	40,390	9,610	5,541	84.85%	0.82	0.89
SVM	Siamese network based	45,869	46,150	3,850	4,131	92.02%	0.92	0.92

are quite satisfactory, which is very promising for a real-world verification application. SVM performs better than NN for all tested representations, since it generalizes better, making use of all training data while learning support vectors. The 4144-dimensional output-based generic representation, carrying finer details for the vessels performs the best for both classifiers. Verification performance is slightly lower for end-to-end learning -based representation compared to the 4144-dimensional output-based vessel representation. One reason may be the limitation in random and insufficient sampling of image pairs out of 4035 different vessels during training.

5.1.2 Vessel retrieval

Compelling amount of research efforts [27–30] have been put on content-based image retrieval (CBIR) as volumes of image databases are dramatically growing. Particularly, vessel retrieval is another promising application, potentially required in a maritime security system, where a user would like to query a database with a vessel image and retrieve similar images. It may also help annotating vessel images uploaded to a database when no meta-data is present. In our application, the retrieved content is not chosen as either the superclasses of *vessel types* that we

constructed as the coarse attribute in Section 4.3, or the IMO number (aiming to identify the exact vessel), which is too fine for a retrieval task (This is studied as a recognition problem in Section 5.1.3.). Instead, we use 109 *vessel types* of the 8000 unique vessels with 50 example images, as the content for the retrieval task. We perform content based vessel retrieval (CBIR), using Euclidean (L^2) and chi-squared (χ^2) distances as the similarity metric for four different vessel representations.

The first representation is one of the presented generic descriptions for marine vessels, a 109-dimensional classifier output of the network, trained on IMO training set. The second representation is the 4144 dimensional output-based generic description designed for distinguishing both vessel types and identities. Third representation is based on a Siamese network similar to the one, end-to-end trained in Section 5.1.1. However, this network focuses on matching vessel types. On the other hand, we also compare these learned deep representations (employing the content information) with another effective representation, designed for object classification. Hence, we use pre-learned VGG-F weights to extract 4096-dimensional features. We train a multi-class SVM to train a classifier for 109 vessel types on the IMO training

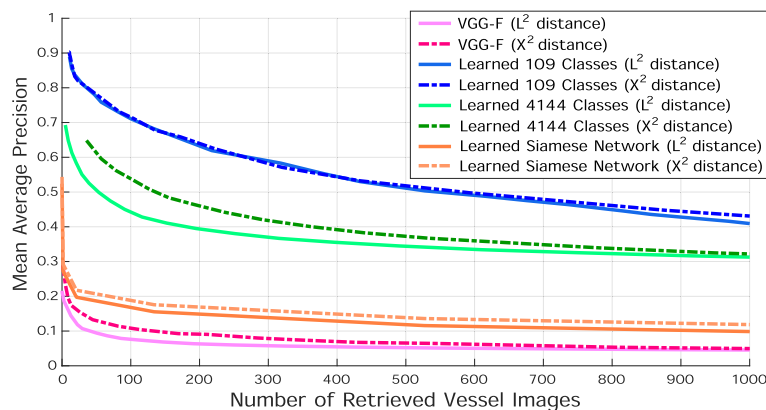


Fig. 8 Vessel retrieval results for four representations: the feature vectors of pre-trained VGG-F network (shown in magenta), AlexNet network based 109 (shown in blue), 4144 (shown in green) dimensional output based, and Siamese network (shown in orange) representations

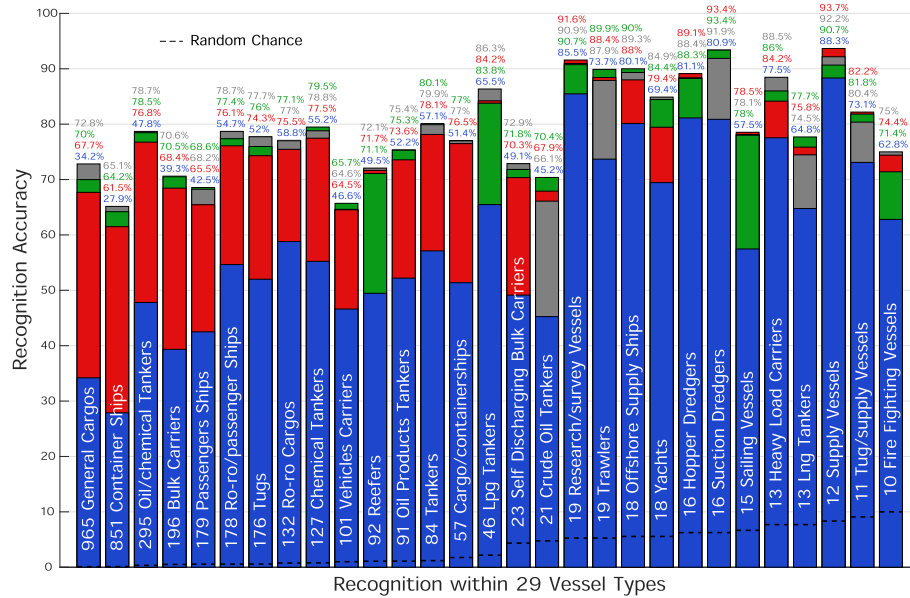


Fig. 9 Vessel type specific recognition: Average recognition accuracies computed within each of the 29 vessel types on IMO testing set are depicted for extracted 109- (blue), 4035- (red), and 4144- (green) dimensional output-based representations and VGG-VD-19-based 4144-dimensional output-based representation (gray) learned in IMO training set

set. For each example, classifier responses of dual combinations of 109 classes (generated during the multi-class SVM phase) are utilized as $\binom{109}{2}$ dimensional feature vectors. By utilizing these four representations, various numbers of images are retrieved and mean average precision curves are generated, as depicted in Fig. 8.

Here, the deep representations learned specifically for maritime vessels significantly outperform the deep representation (VGG-F) learned for general object categorization for 1000 classes [2, 20] for both distance metrics. In addition, χ^2 distance is superior in CBIR than L^2 distance, for the tested representations. A 109-dimensional output-based generic representation performs the best in this experiment, since it is specifically designed for learning vessel types. The retrieval performance of Siamese network, utilizing end-to-end learning, is lower, compared to 109 and 4144-dimensional representations.

5.1.3 Vessel recognition

Visual object recognition is one of the most crucial topics of computer vision. Especially, face recognition has been studied extensively, and state-of-the-art methods [31, 32],

which perform effectively on the benchmark datasets [33–35], have been proposed. Since encouraging performance results are obtained with recent methods, another application performed, utilizing MARVEL, is *vessel recognition* task, where the ultimate goal is to perceive a vessel's identity by its visual appearance. It might not be meaningful for object types, other than maritime vessels or faces, such as cars, since same car models with same color have no visual differences and technically are not distinguishable. Nevertheless, individual vessels generally carry distinctive features, as shapes of vessels belonging to the same *vessel type category* may vary significantly due to their customized construction processes. Here, we utilize the learned generic vessel representations as feature vectors for vessels.

We perform *identification* for two scenarios. First, we assume the vessel type labels are provided. Hence, recognition is performed among individual classes separately, e.g., vessels belonging to the *passenger ships* class are learned and recognized. Multi-class SVMs are trained for images belonging to each vessel type and classification is done. Among the 3965 vessels in IMO test set, there exist 29 *vessel types* that have at least 10

Table 2 Vessel recognition performance on IMO testing set, composed of 3965 marine vessels, by utilizing nearest neighbor search on 109-, 4035-, and 4144-dimensional output-based representations learned in IMO training set

	109-dim. output based representation	4035-dim. output based representation	4144-dim. output based representation	4144-dim. output based representation (VGG-VD-19)
Recognition accuracy	23.87%	59.25%	65.13%	65.78%

Table 3 Vessel attribute prediction performance, measured as correlation of manual truth and predicted labels for 158,850 images in IMO testing set

	Draught	Gross tonnage	Length	Summer deadweight
SVM	0.7556	0.8301	0.8696	0.7930
CNN	0.7911	0.2699	0.9042	0.0830

unique vessels, and each unique vessel has 50 example images. For recognition, we first divide the examples of each vessel into fivefolds where each fold has 10 examples per vessel. The training and testing sets contain fourfolds (40 examples) and onefold (10 examples) per vessel, respectively. We perform fivefold cross-validation for classifying all 50 example images of each vessel. For each multi-class SVM, the number of classes equals the number of unique vessels of that particular *vessel type*. In Fig. 9, the recognition performances are illustrated for each *vessel type* and by using each generic vessel representation as feature vectors. Representations trained over 4035- and 4144-dimensional output labels, which aim to learn specific vessels in IMO training set, perform significantly better than the representation trained on 109-dimensional output labels which only learns vessel types on IMO training set. Being able to learn both, hence extracting both coarse and fine details, 4144 dimensional output-based representation is the best of three for generic vessel description. Random chance for recognition is also depicted in the figure in order to prove the success of the presented generic marine vessel representations. Additionally, we tested the performance of 4144 dimensional representation when employing a deeper neural network VGG-VD-19 [20], and we obtain high performance similarly.

Vessels belonging to *research survey vessels*, *suction dredgers*, and *supply vessels* type classes of are the most distinguishable ones with recognition accuracies above 90%. On the other hand, vessels of *crude oil tankers*, *vehicle carriers*, and *containership* classes have less distinct differences and a slightly lower recognition performances are achieved, compared to the rest of the classes. Please note that, as number of unique vessels increase in a vessel type group, the random chance and recognition rates slightly decrease as expected, since it becomes a more challenging recognition problem. Yet, recognition accuracies over 77% can be obtained even though the number of unique vessels exceeds a hundred, such as in *ro-ro cargo* and *chemical tanker* vessel types.

As a second scenario for recognition, we attempt recognition of vessels when there is no prior information, namely, when type labels are not present. Here, the goal is

to classify images of 3965 vessels in IMO testing set by the use of generic vessel representations learned on images of IMO training set. Large number of classes makes it computationally infeasible to train models with a SVM; thus, we employ a nearest neighbor classifier for this experiment. In a similar setting, we split images of individual vessels in IMO testing set into five non-overlapping folds (fourfolds as a training and onefold as a testing split), and we perform fivefold cross-validation for and classify all 50 example images of each vessel. For each image in a testing fold, we find the best matching image among training images and assign its label for the test image. Repeating the same experiment for four generic representations, we conclude that 4144-dimensional output-based representations (AlexNet based and VGG-VD-19 based) perform better than the other two. The recognition rates are listed in Table 2.

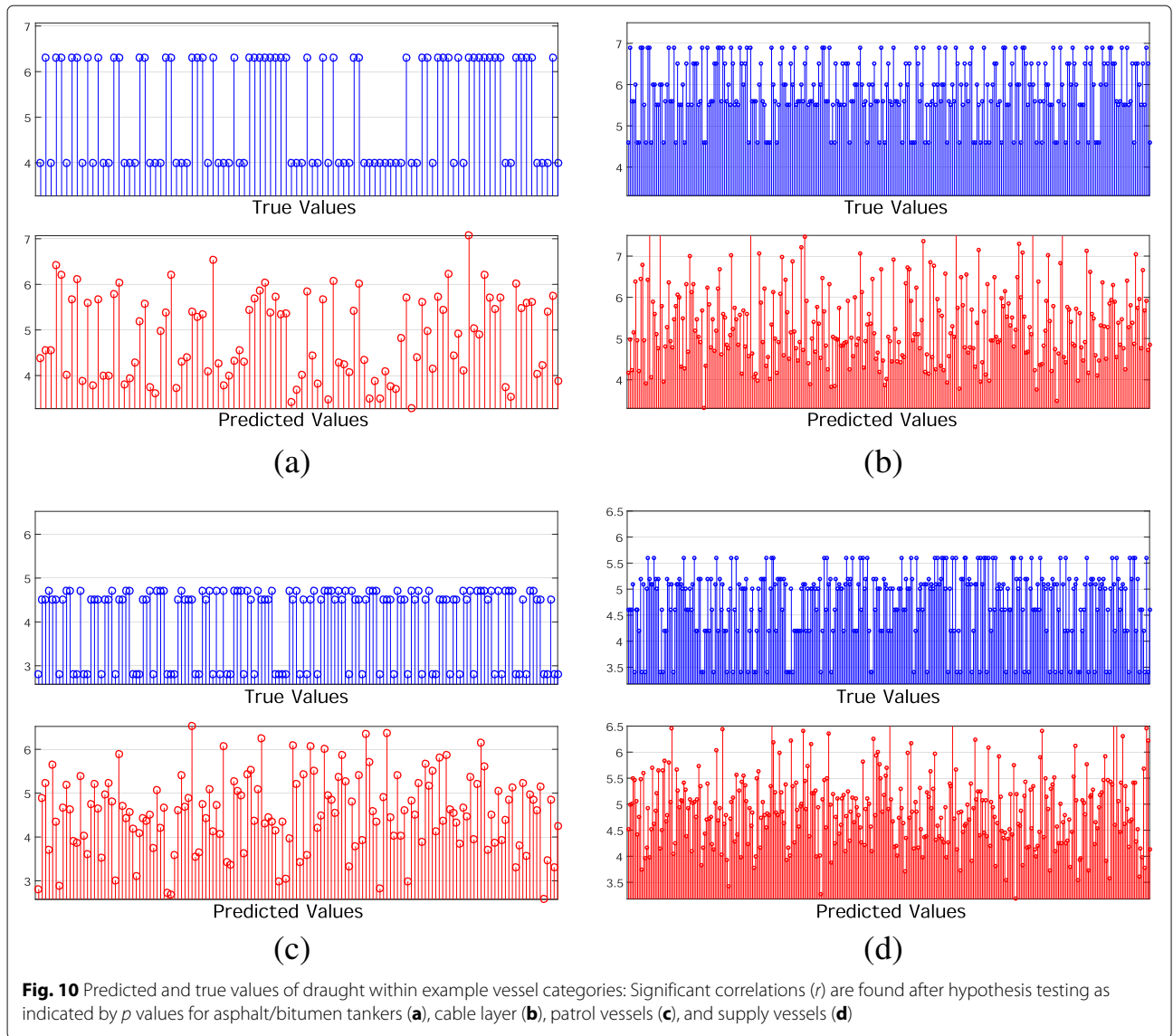
5.2 Vessel attribute prediction and classification

MARVEL dataset includes several labeled vessel attributes some of which relate to the visual content. Here, as interesting applications, by studying only the visual content, we targeted predicting and classifying four important attributes: draught, gross tonnage, length, and summer deadweight.

The draught of a vessel is a measure describing the vertical distance between the waterline and the bottom of vessel hull. Draught, defining the minimum depth of water a vessel can operate, is an important factor for navigating and routing vessels while avoiding shallow water pathways. Length of a vessel does matter for navigation and marine traffic routing, as well as for calculating fees during vessel registration. Consequently, estimating length of a vessel effectively from a single image may be very beneficial for maritime applications. Gross tonnage is a nonlinear measure calculated based on overall interior volume (from keel to funnel) of a vessel. It is important in determining the number of staff, safety rules, registration fees, and port dues. Summer deadweight defines how much mass a ship can safely carry. It excludes the weight of the ship and includes the sum of the weights of cargo, fuel, fresh water, ballast water, provisions, passengers, and crew [36].

Table 4 Vessel attribute prediction performance, measured as coefficient of determination between manual truth and predicted labels for 158,850 images in IMO testing set

	Draught	Gross tonnage	Length	Summer deadweight
SVM	0.598	0.554	0.743	0.481
CNN	0.770	0.419	0.863	0.466



Such efforts of attribute estimation is especially valuable for coastal guarding and surveillance, since it allows grasping the physical specifications of a vessel remotely and only by a captured image. In order to achieve these objectives, we both test the use of our powerful 4144-dimensional output-based generic vessel representation and also employ specific attribute-based deep representations. *Please note that estimating these attributes are very challenging due to the lack of notion of scale, pose, perspective, camera parameters, etc. The only available information is the appearance of a vessel.* For all experiments of attribute prediction, we learn models in IMO training set and evaluate performances of the learned models in IMO testing set. Images missing valid attribute labels were not used in these experiments. Attribute labels, as opposed to being discrete numbers as in vessel type labels or IMO

number labels, are continuous and might be unique for each vessel.

We design two sets of experiments: regression and classification. Approaching the problem as a regression task, we represent vessel images by either generic deep models we designed for marine vessels or deep models trained for estimating specific attributes. As in the previous experiments, we extract the penultimate layer activations of the trained networks as feature vectors and utilize a support vector regressor [25, 37] for prediction. For learning attribute-specific deep models, we use AlexNet as a base CNN architecture and modify the last loss layer with an objective to minimize an $L2$ -norm loss, approaching the problem as a least squares regression. For performance evaluation, we compute two measures.

Table 5 Vessel attribute classification performance of generic and attribute-specific representations, calculated for four attributes on 158,850 images of IMO testing set

Classified attribute	Employed representation	Top 1 accuracy	Top 2 accuracy	Top 3 accuracy	Top 4 accuracy	Top 5 accuracy
Draught	Generic model + SVM	0.1302	0.3104	0.4432	0.5506	0.6320
Gross tonnage	Generic model + SVM	0.4755	0.6393	0.7418	0.8178	0.8678
Length	Generic model + SVM	0.4539	0.6345	0.7317	0.8019	0.8510
Summer deadweight	Generic model + SVM	0.4304	0.6209	0.7310	0.7998	0.8525
Draught	Attribute-specific trained CNN	0.1834	0.4159	0.5761	0.6884	0.7774
Gross tonnage	Attribute-specific trained CNN	0.5515	0.7492	0.8556	0.9131	0.9454
Length	Attribute-specific trained CNN	0.5289	0.7266	0.8257	0.8896	0.9328
Summer deadweight	Attribute-specific trained CNN	0.5155	0.7364	0.8317	0.8938	0.9288

The first measure is Pearson correlation coefficient between predicted labels and manual truth. It is defined as,

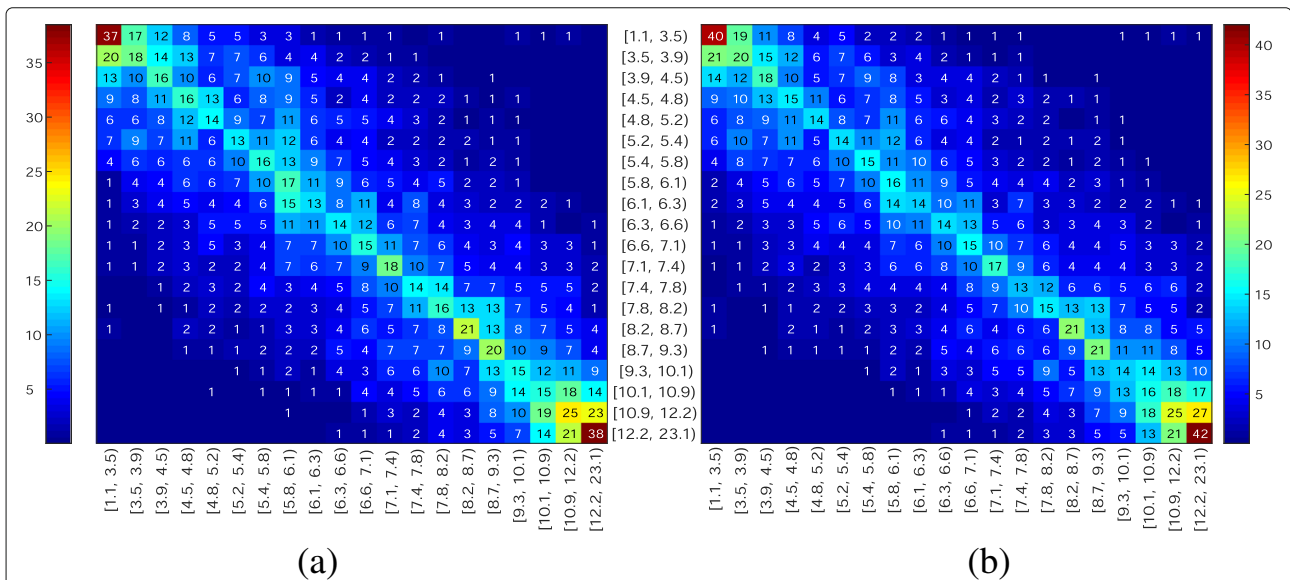
$$r = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}, \quad (4)$$

where \hat{y}_i and y_i are single indexed samples of predicted labels and true labels, respectively. N is the sample size, which is 158,850, corresponding to all test images with

valid attribute labels. These results are given in Table 3. The highest correlations obtained are 0.9042 for length, 0.7911 for draught, 0.8301 for gross tonnage, and 0.7930 for summer deadweight.

The second measure we report is the coefficient of determination, namely R^2 , which quantifies how well regression model fits the data. It is calculated as,

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2}. \quad (5)$$

**Fig. 11** Confusion matrices for classifying draught: **a** generic vessel features combined with a support vector machine classifier and **b** learned draught-specific representation combined with a softmax classifier

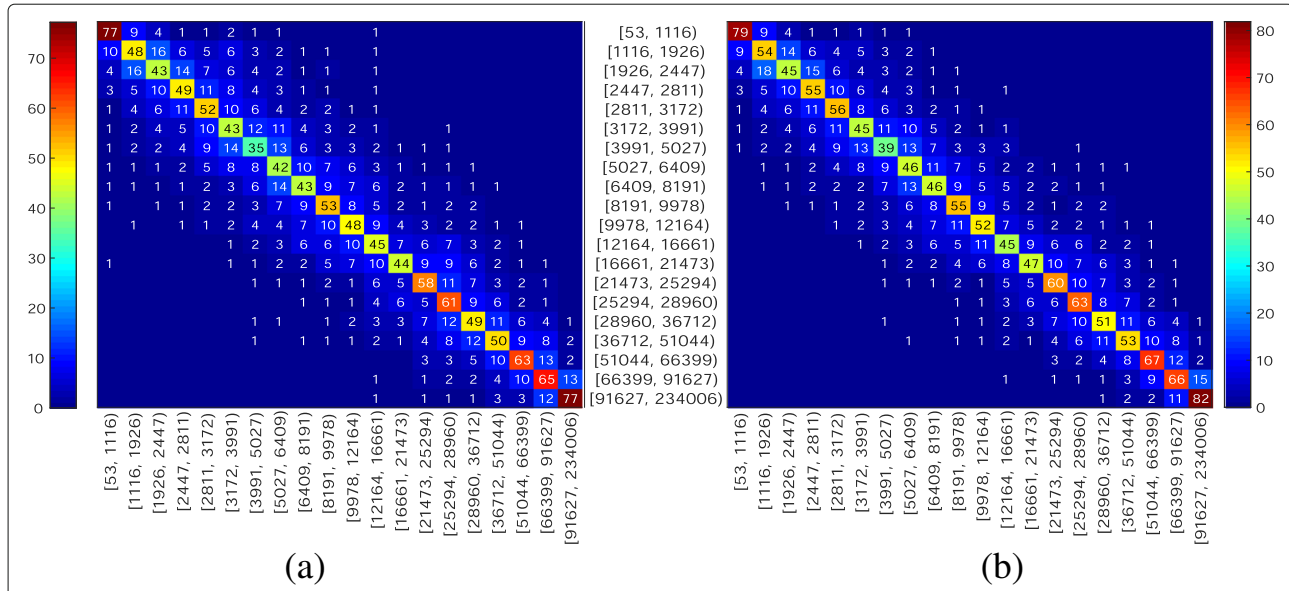


Fig. 12 Confusion matrices for classifying gross tonnage: **a** generic vessel features combined with a support vector machine classifier and **b** learned gross tonnage-specific representation combined with a softmax classifier

Table 4 shows the R^2 values when predicting four attributes. SVM classifier employs the generic representation learnt for vessel type classification, whereas CNN employs a representation specifically learnt for predicting attributes. Table 4 shows that attribute-based representation performs better for predicting length and draught; nevertheless, it performs slightly worse for gross tonnage and summer deadweight. Thus, we may conclude that for predicting physical attributes, values of which are visually explicit, specific representations are more effective. For predicting attributes such as weight, our method relies on vessel type classification.

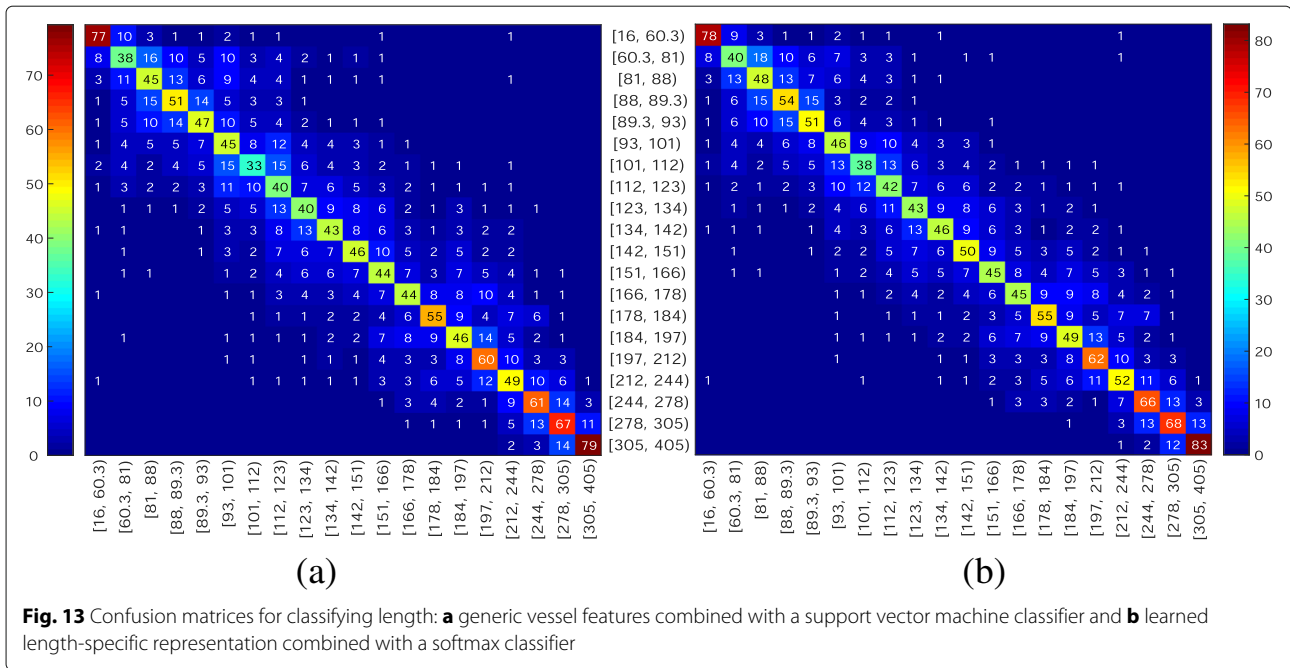
For further analysis, we plot predicted draught values for four example vessel categories separately in Fig. 10. The annotated attributes differ for individual vessels within specific vessel categories. However, the significant correlations, between the true values and predicted values for vessels belonging to the same types, show that learnt representations, capturing visual cues, are effective in attribute prediction. The trained neural networks simply try to estimate vessel attributes similar to how human can do, based on clues such as vessel type and also appearance (visible parts of a vessel).

As another experiment, we quantize the attribute labels and relabel and assign the images in IMO training set accordingly to 20 distinct classes such that each class has equivalent number of examples for a balanced training. Next, we train a multi-class classifier, using both the generic vessel representation (combined with a nonlinear SVM) and also specific deep representations (softmax classifier) for each attribute. For instance, in training, we

use a total of 134,000 images for draught, 142,000 images for gross tonnage, 140,000 images for length and 148,000 images for summer deadweight. For testing, we use all 158,850 images of IMO test set for which all attribute annotations are present. Top five classification accuracies for the attributes and employed representations are summarized in Table 5. Though generic vessel representation performs reasonably well, trained deep models which focus on specific attributes are significantly better in attribute categorization. The classification results are also depicted as normalized confusion matrices in Figs. 11, 12, 13, and 14. The imbalance of the training set results in coarser ranges for classes around the extrema values and very fine classes otherwise. The entries of the confusion matrices are high valued along the diagonal entries, which shows that the learned models are effective in capturing the desired attribute information.

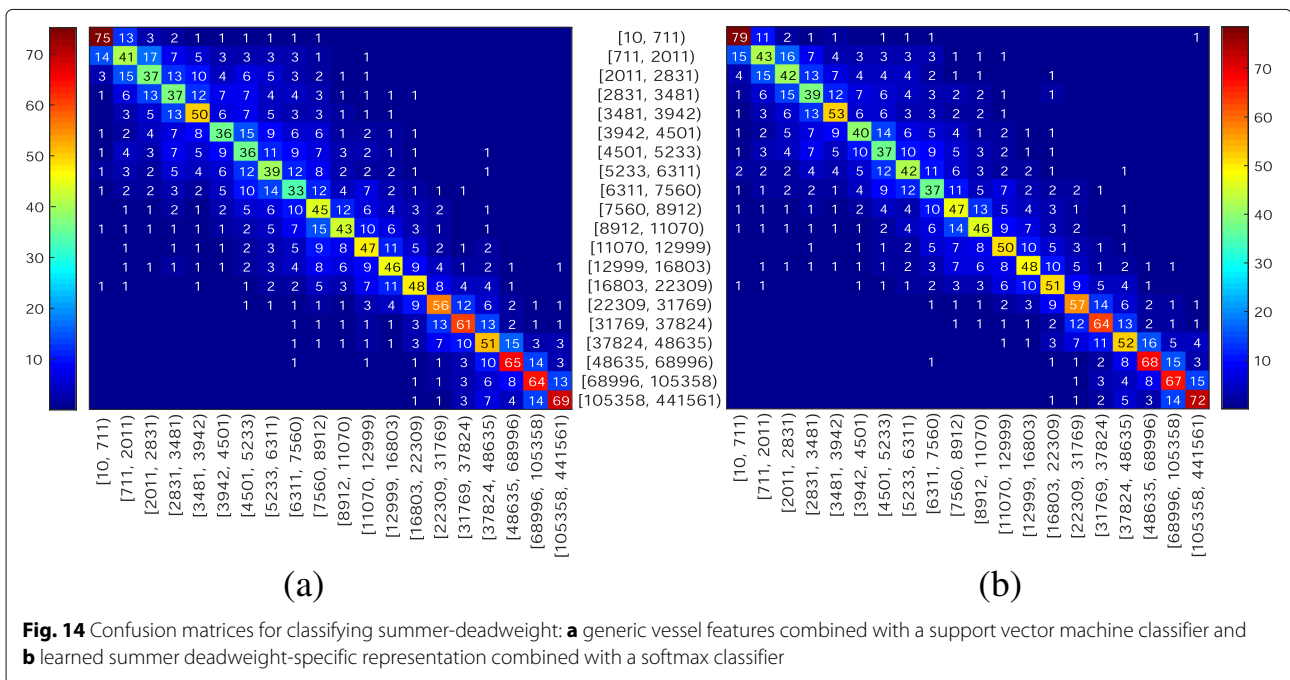
6 Discussions

Introducing MARVEL, a large-scale dataset for maritime vessels, our goal is to point out several research problems and applications for maritime images. MARVEL dataset, composed of a massive number of images and their meta-data, carries interesting attributes to be considered for visual analysis tasks. In this work, we presented our efforts for visual classification of maritime vessel types, retrieval, identity verification, identity recognition, and estimation of physical attributes such as draught, length, and tonnage of vessels. For each of these tasks, we provide the details (experimental settings, labels, training and testing splits) to make results reproducible.



For organizing the dataset, first, we performed semantic analysis and combined vessel type classes which are visually indistinguishable. Next, we pruned annotations for attributes semi-automatically, converting them to certain metric units, filtering out the missing and wrong entries and ensured reliability of the labels. We also present baseline results for several computer vision tasks to inspire future applications on MARVEL. Moreover, we provide generic deep representations for maritime vessels and prove their success in aforementioned tasks

by performing extensive experiments. We achieve promising performance in vessel classification, recognition, and retrieval. Moreover, we observe that attributes are predictable as long as they are visually distinguishable. Hence, attributes such as length and draught can be estimated accurately and by solely exploiting visual data. What remains of key interest for future work is the enhancement of performance for the aforesaid tasks, which can be fulfilled by utilizing more powerful visual representations, developing sophisticated methods.



Endnote

¹A negative pair indicates a pair of different vessel images, whereas a positive pair corresponds to a pair of vessel images belonging to a unique vessel.

Acknowledgements

We would like to thank to Koray Akçay for his invaluable support and special consultancy for maritime vessels.

Authors' contributions

VY took charge of data collection and organization. VY and EH generated the statistics of the collected dataset. EH implemented and performed the representation learning for marine vessels and carried out vessel type classification experiments. BS designed the marine vessel applications (verification, retrieval, recognition, attribute estimation) and implemented and carried out the related experiments. BS proposed superclass generation and EH implemented and executed the task. EH organized the initial manuscript and BS created the supplemental. BS later revised and extended the work and writing. AK coordinated the work during the study and did English revising. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 21 June 2017 Accepted: 7 November 2017

Published online: 11 December 2017

References

- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet large scale visual recognition challenge. *Int J Comput Vis (IJCV)* 115(3):211–252. doi:10.1007/s11263-015-0816-y
- Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*. Curran Associates Inc., Lake Tahoe, Nevada. pp 1097–1105. <http://dl.acm.org/citation.cfm?id=2999134.2999257>
- Lin D, Shen X, Lu C, Jia J (2015) Deep lac: Deep localization, alignment and classification for fine-grained recognition. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp 1666–1674. doi:10.1109/CVPR.2015.7298775
- Xie S, Yang T, Wang X, Lin Y (2015) Hyper-class augmented and regularized deep learning for fine-grained image classification. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp 2645–2654. doi:10.1109/CVPR.2015.7298880
- Liu L, Shen C, van den Hengel A (2015) The treasure beneath convolutional layers: Cross-convolutional-layer pooling for image classification. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp 4749–4757. doi:10.1109/CVPR.2015.7299107
- Maji S, Rahtu E, Kannala J, Blaschko M, Vedaldi A (2013) Fine-grained visual classification of aircraft. *arXiv preprint arXiv:1306.5151*
- Vedaldi A, Mahendran S, Tsogkas S, Maji S, Girshick R, Kannala J, Rahtu E, Kokkinos I, Blaschko MB, Weiss D, Taskar B, Simonyan K, Saphra N, Mohamed S (2014) Understanding objects in detail with fine-grained attributes. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Institute of Electrical and Electronics Engineers, USA. pp 3622–3629
- Wah C, Branson S, Welinder P, Perona P, Belongie S (2011) The Caltech-UCSD Birds-200-2011 Dataset. Technical Report CNS-TR-2011-001. California Institute of Technology
- Krause J, Stark M, Deng J, Fei-Fei L (2013) 3d object representations for fine-grained categorization. In: *Computer Vision Workshops (ICCVW)*, 2013 IEEE International Conference On. pp 554–561. doi:10.1109/ICCVW.2013.77
- Yang L, Luo P, Loy CC, Tang X (2015) A large-scale car dataset for fine-grained categorization and verification. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp 3973–3981. doi:10.1109/CVPR.2015.7299023
- Dao-Duc C, Xiaohui H, Morère O (2015) Maritime vessel images classification using deep convolutional neural networks. In: *Proceedings of the Sixth International Symposium on Information and Communication Technology*. SolICT 2015. ACM, New York. pp 276–281. doi:10.1145/2833258.2833266. <http://doi.acm.org/10.1145/2833258.2833266>
- Ship Photos and Ship Tracker. www.shipspotting.com. Accessed 1 May 2017
- Gundogdu E, Solmaz B, Yücesoy V, Koç A (2016) MARVEL: a large-scale image dataset for maritime vessels. In: *Asian Conference on Computer Vision*. Springer International Publishing, Cham. pp 165–180
- Solmaz B, Gundogdu E, Karaman K, Koç A, et al (2017) Fine-grained visual marine vessel classification for coastal surveillance and defense applications. In: *Electro-Optical Remote Sensing XI*. vol. 10434. International Society for Optics and Photonics, USA. p 104340A
- Zhang X, Zhou F, Lin Y, Zhang S (2016) Embedding label structures for fine-grained feature representation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Institute of Electrical and Electronics Engineers, USA. pp 1114–1123
- Farhadi A, Endres I, Hoiem D, Forsyth D (2009) Describing objects by their attributes. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Institute of Electrical and Electronics Engineers, USA. pp 1778–1785. doi:10.1109/CVPR.2009.5206772
- Lampert CH, Nickisch H, Harmeling S (2009) Learning to detect unseen object classes by between-class attribute transfer. In: *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On*. pp 951–958. doi:10.1109/CVPR.2009.5206594
- Sun Y, Bo L, Fox D (2013) Attribute based object identification. In: *2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany, May 6-10, 2013*. pp 2096–2103. doi:10.1109/ICRA.2013.6630858
- Vedaldi A, Lenc K (2015). In: *Proceedings of the 23rd ACM international conference on Multimedia*. ACM. pp 689–692
- Chatfield K, Simonyan K, Vedaldi A, Zisserman A (2014) Return of the devil in the details: delving deep into convolutional nets. *arXiv preprint arXiv:1405.3531*
- Crammer K, Singer Y (2002) On the learnability and design of output codes for multiclass problems. *Mach Learn* 47(2):201–233. doi:10.1023/A:1013637720281
- Keerthi SS, Sundararajan S, Chang KW, Hsieh CJ, Lin CJ (2008) A sequential dual method for large scale multi-class linear svms. In: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '08. ACM, New York. pp 408–416. doi:10.1145/1401890.1401942. <http://doi.acm.org/10.1145/1401890.1401942>
- Fan RE, Chang KW, Hsieh CJ, Wang XR, Lin CJ (2008) LIBLINEAR: a library for large linear classification. *J Mach Learn Res* 9:1871–1874
- Sun Y, Wang X, Tang X (2014) Deep learning face representation from predicting 10,000 classes. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. Institute of Electrical and Electronics Engineers, USA. pp 1891–1898. doi:10.1109/CVPR.2014.244
- Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2:27–12727. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Hadsell R, Chopra S, LeCun Y (2006) Dimensionality reduction by learning an invariant mapping. In: *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference On*. vol. 2. IEEE, USA. pp 1735–1742
- Guo JM, Prasetyo H (2015) Content-based image retrieval using features extracted from halftoning-based block truncation coding. *IEEE Trans Image Process* 24(3):1010–1024. doi:10.1109/TIP.2014.2372619
- Qiu G (2003) Color image indexing using btc. *IEEE Trans Image Process* 12(1):93–101
- Lai CC, Chen YC (2011) A user-oriented image retrieval system based on interactive genetic algorithm. *IEEE Trans Instrum Meas* 60(10):3318–3325. doi:10.1109/TIM.2011.2135010
- Gordo A, Almazan J, Revaud J, Larlus D (2017) End-to-end learning of deep visual representations for image retrieval. *Int J Comput Vis* 124(2):237–254
- Lai J, Jiang X (2016) Classwise sparse and collaborative patch representation for face recognition. *IEEE Trans Image Process* 25(7):3261–3272. doi:10.1109/TIP.2016.2545249

32. Gong D, Li Z, Tao D, Liu J, Li X (2015) A maximum entropy feature descriptor for age invariant face recognition. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Institute of Electrical and Electronics Engineers, USA. pp 5289–5297. doi:10.1109/CVPR.2015.7299166
33. Lee KC, Ho J, Kriegman DJ (2005) Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans Pattern Anal Mach Intell* 27(5):684–698. doi:10.1109/TPAMI.2005.92
34. Sim T, Baker S, Bsat M (2003) The cmu pose, illumination, and expression database. *IEEE Trans Pattern Anal Mach Intell* 25(12):1615–1618. doi:10.1109/TPAMI.2003.1251154
35. Ricanek K, Tesafaye T (2006) Morph: a longitudinal image database of normal adult age-progression. In: 7th International Conference on Automatic Face and Gesture Recognition (FGR06). pp 341–345. doi:10.1109/FGR.2006.78
36. Turpin EA, McEwen WA (1980) *Merchant Marine Officers' Handbook*. 4th edn.. Cornell Maritime Press, Centreville, Maryland
37. Schölkopf B, Smola AJ, Williamson RC, Bartlett PL (2000) New support vector algorithms. *Neural computation* 12(5):1207–1245

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)