*Article*

# A Public Dataset for Fine-Grained Ship Classification in Optical Remote Sensing Images

Yanghua Di [1,2,3], Zhiguo Jiang [1,2,3] and Haopeng Zhang [1,2,3,*]

1   Department of Aerospace Information Engineering, School of Astronautics, Beihang University,
    Beijing 102206, China; dyhbuaasa@buaa.edu.cn (Y.D.); jiangzg@buaa.edu.cn (Z.J.)
2   Beijing Key Laboratory of Digital Media, Beijing 102206, China
3   Key Laboratory of Spacecraft Design Optimization and Dynamic Simulation Technologies,
    Ministry of Education, Beijing 102206, China
*   Correspondence: zhanghaopeng@buaa.edu.cn; Tel.: +86-10-6171-6978

**Abstract:** Fine-grained visual categorization (FGVC) is an important and challenging problem due to large intra-class differences and small inter-class differences caused by deformation, illumination, angles, etc. Although major advances have been achieved in natural images in the past few years due to the release of popular datasets such as the CUB-200-2011, Stanford Cars and Aircraft datasets, fine-grained ship classification in remote sensing images has been rarely studied because of relative scarcity of publicly available datasets. In this paper, we investigate a large amount of remote sensing image data of sea ships and determine most common 42 categories for fine-grained visual categorization. Based our previous DSCR dataset, a dataset for ship classification in remote sensing images, we collect more remote sensing images containing warships and civilian ships of various scales from Google Earth and other popular remote sensing image datasets including DOTA, HRSC2016, NWPU VHR-10, We call our dataset FGSCR-42, meaning a dataset for Fine-Grained Ship Classification in Remote sensing images with 42 categories. The whole dataset of FGSCR-42 contains 9320 images of most common types of ships. We evaluate popular object classification algorithms and fine-grained visual categorization algorithms to build a benchmark. Our FGSCR-42 dataset is publicly available at our webpages.

**Keywords:** ship classification; remote sensing; fine-grained visual categorization

## 1. Introduction

Fine-grained categorization attracts extensive attention in computer vision field. The task is relatively challenging due to large intra-class differences and small inter-class differences compared with general classification tasks. The fine-grained image recognition should pay more attention to marginal visual differences within subordinate categories. Similar to natural images, ships in remote sensing images are also different from each other in different lighting and imaging conditions. In addition, due to the relatively small inter-class differences, fine-grained ship classification is also challenging and of great significance in understanding remote sensing images.

In recent years, with the rapid development of optical satellites, the outstanding advantages of optical images in ship reconnaissance, especially in the ship classification, have attracted the tremendous attention of marine monitoring departments and scholars. At the same time, driven by the success of deep-learning-based algorithms for extracting deep features, many researchers have pursued approaches based on fine-tuning networks for object detection based on remote sensing image datasets, e.g., DOTA [1] (https://captain-whu.github.io/DOTA/dataset, accessed on 14 December 2020), HRSC2016 [2] (http://www.escience.cn/people/liuzikun/DataSet.html, accessed on 14 December 2020) and NWPU VHR-10 [3] (https://jiong.tea.ac.cn/people/JunweiHan/NWPUVHR10dataset, accessed on 14 December 2020). However, most previous works only distinguish

the ship and background or a few categories like merchant ship, warship, etc. Few efforts have been devoted into fine-grained ship classification in remote sensing images. Since there is no publicly available dataset, evaluating the accuracy of fine-gained ship classification in remote sensing images is also difficult.

Although deep convolutional networks are very effective for feature extraction and object classification, the task of fine-grained ship classification in optical remote sensing images is still filled with difficulties. Firstly, optical remote sensing images containing ships are easily influenced by sea state, weather, illumination, sensor parameters and other factors, causing significant degradation in data quality. Secondly, as an artificial rigid target, the ships are mostly axisymmetric and have different shapes and structures due to their different usage, making the intra-class differences extremely large. In addition, the landing ship is usually parallel to the dock shoreline and ships of different categories often docked close together in many actual scenes. This definitely could cause multiple labels, which significantly influences the accuracy of the classification. Thirdly, methods with more effective capability of learning features are needed for fine-grained ship classification in remote sensing images, since the increasing number of categories to be recognized reduces inter-class differences and makes it difficult to extract discriminative features. Besides the distinct difficulties above, fine-grained ship classification is also challenged by the complexity of background, e.g., the port buildings and shore roads, houses, etc., which makes more difficult to find some robust features to accomplish fine-grained ship classification since these objects may have similar shapes and textures with ships. For example, a inshore ship could own similar appearance and shape as its berth, as is shown in Figure 1.
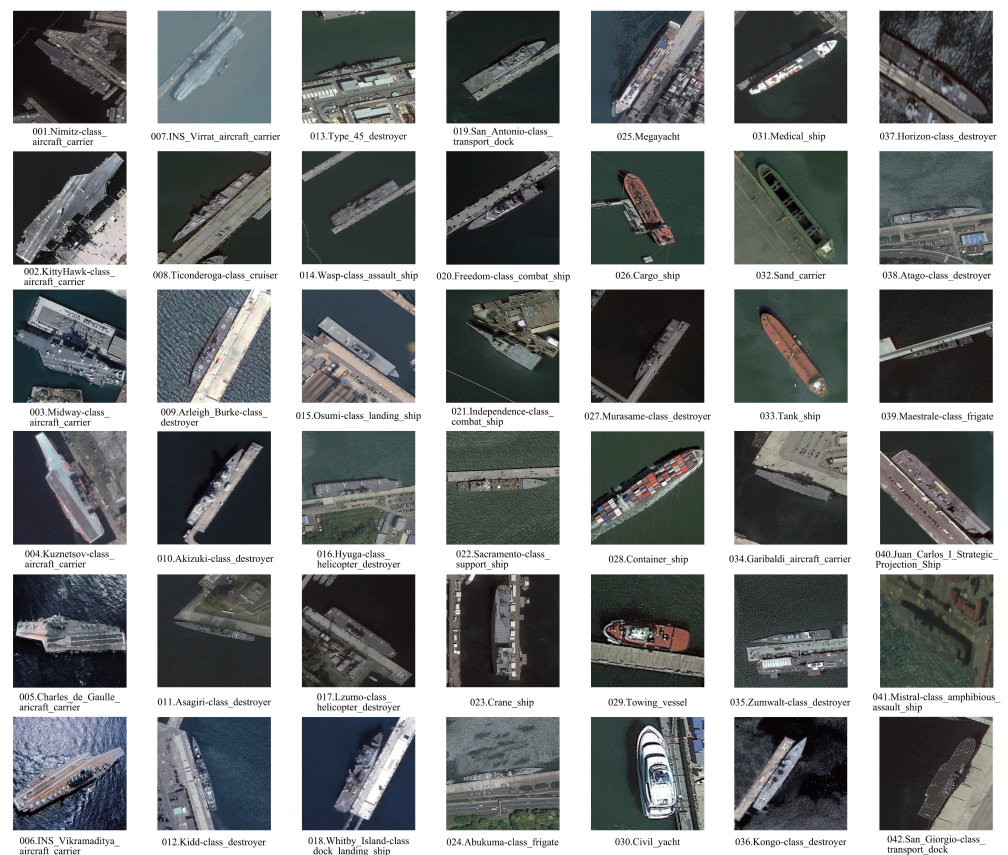


**Figure 1.** Typical images of all categories in FGSCR-42 dataset. Examples contain ships in the harbor and sea areas with various backgrounds, different angles, large spatial resolution and different levels of illumination.

There are many famous large natural datasets for fine-grained image classification, such as CUB-200-2011, Stanford Cars and Aircraft datasets, which make significant contribution to the advance in fine-grained classification in natural scenes. Although features learned by deep convolutional networks from natural images can be transferred to remote sensing images [4], considering the difference between natural and remote sensing images, it is not surprising that such features from natural images could not directly be implemented for ship classification in remote sensing images without fine-turning together with a classifier. Although promising progress has been reported for object detection and recognition in remote sensing images using deep learning methods, fine-gained ship classification remains a challenge since there are no public fine-grained ship classification datasets in remote sensing images.

To accelerate fine-grained ship classification research in optical remote sensing images, this paper introduces a wide-variety dataset for Fine-Grained Ship Classification in Remote sensing images (FGSCR-42). We investigate a comprehensive number of ships in remote sensing images and select most common 42 categories for fine-grained visual categorization, as shown in Figure 1. We have collected remote sensing images from Google Earth, as well as popular remote sensing image datasets like DOTA [1], HRSC2016 [2], NWPU VHR-10 [3], containing warships and civilian ships of various scales. The whole dataset contains 9320 images in 42 categories. In addition, compared with previous DSCR [5] (https://github.com/DYH666/DSCR, accessed on 14 December 2020) dataset for ship classification, the number of ship categories would significantly improve the task of fine-grained ship classification in remote sensing images.

The rest of the paper is organized as follows. Section 2 introduces the motivations of this study. Section 3 describes how the FGSCR-42 dataset was created. The details of properties of the dataset are given in Section 4. Evaluation and benchmark results are shown in Section 5. Section 6 concludes the paper.

## 2. Motivations

Over the recent decades, publicly shared datasets have played an important role in data-driven research. Popular datasets like MSCOCO and ImageNet have significantly benefited the research of object detection and image classification, while the CUB-200-2011 [6] (http://www.vision.caltech.edu/visipedia/CUB-200-2011, accessed on 14 December 2020) dataset is instrumental in promoting fine-grained visual categorization research. General fine-grained classification datasets like CUB-200-2011 and Stanford Cars [7] (http://ai.stanford.edu/~jkrause/cars/car_dataset.html, accessed on 14 December 2020) are favored by researchers due to the large number of images, many categories and detailed annotations. To date, little research has been devoted to the task of fine-grained ship classification in remote sensing images. It is very important to propose a fine-grained ship classification dataset for the rapid progress in the field of remote sensing image ship classification, which is of great significance for sea surface and port monitoring. Our goal is to establish a public dataset for fine-grained ship classification in remote sensing images. We put a lot of effort into image collecting, category labeling and dataset organization, which enables FGSCR-42 to make a great contribution to the classification of ships in remote sensing images. Due to the difficulty of acquiring remote sensing images and confirming categories, the number of images in FGSCR-42 is relatively small compared with natural image datasets. However, it has an advantage in the resolution of images, which is more conducive to extracting feature information for fine-grained ship classification in remote sensing images. The comparison of our FGSCR-42 dataset with other fine-gained classification datasets for natural images is illustrated in Table 1, including Stanford Dogs [8] (http://vision.stanford.edu/aditya86/ImageNetDogs, accessed on 14 December 2020), 102 Category Flower Dataset [9] (https://www.robots.ox.ac.uk/~vgg/data/flowers/102, accessed on 14 December 2020), FGVC-Aircraft [10] (https://hal.inria.fr/hal-00842101, accessed on 14 December 2020), etc. Our FGSCR-42 resembles CUB-200-2011 in terms of the image numbers and per class, but CUB-200-2011's

scales are not as many as FGSCR-42 because ships in remote sensing images often have a relatively large scale span.

**Table 1.** Comparison among FGSCR-42 and recent popular natural image datasets.

| Dataset | Categories | Total Images | Images per Class | Image Width (Pixels) |
|---|---|---|---|---|
| CUB-200-2011 [6] | 200 | 11,788 | 200 | ~500 |
| Stanford Dogs [8] | 120 | 20,580 | 8500 | ~1000 |
| Stanford Cars [7] | 196 | 8144 | 8041 | 128 |
| 102 Category Flower Dataset[9] | 102 | 8189 | 40~258 | 256 |
| FGVC-Aircraft [10] | 102 | 10,200 | 100 | 512 |
| **FGSCR-42** | 42 | 9320 | ~200 | **50~1500** |

## 3. Establishment of FGSCR-42

### 3.1. Images Collection

Considering the accuracy of categories and the precise annotations, we first choose original remote sensing images containing ship instances from popular public remote sensing datasets including DOTA, HRSC2016 and NWPU VHR-10. In addition, we collect optical remote sensing images of ships in 42 categories with wide range of resolutions. To increase the diversity of image data, we review the images of 40 ports around the world for the past twenty years. Obviously, such a large amount of data enables FGSCR-42 more competitive in fine-grained visual categorization. At the same time, we record the exact geographical coordinates of each port and the capture time of the raw remote sensing data to ensure that the data sources do not duplicate.

In order to make our FGSCR-42 dataset suitable for fine-grained ship classification in remote sensing images, we consider the three properties when collecting images, namely, a sufficient number of total images, balanced and enough instances per category, and enough categories to approach practical applications. Due to insufficient efforts are devoted to public fine-grained ship classification dataset in remote sensing images, we show a comparison with remote sensing image scene classification datasets, SAR image classification datasets and remote sensing object detection datasets in Table 2. Noticing that our FGSCR-42 dataset only contains ship images, it can be seen that our FGSCR-42 dataset takes a trade-off between the number of total images and instances per category, and has the most number of ship categories.

**Table 2.** Comparison among FGSCR-42 and some other remote sensing image datasets for classification.

| Dataset | Main Categories | Images | Per Category | Image Width |
|---|---|---|---|---|
| NWPU VHR-10 [11] | 10 | 3896 | ~400 | ~1000 |
| RSOD [12] | 4 | 7400 | ~1800 | ~1000 |
| MSTAR [13] | 3 | 5950 | ~2000 | 128 |
| RSC11 | 11 | 1232 | 100 | 512 |
| RSSCN7 [14] | 7 | 2800 | 400 | 400 |
| WHU-RS19 [15] | 19 | 1005 | 50 | 600 |
| **FGSCR-42** | **42** | **9320** | **~200** | **50~1500** |

### 3.2. Category Selection

In order to achieve better performance in the fine-grained ship classification in remote sensing images, we investigate a large amount of remote sensing image data of sea ships with repeated comparisons and confirmations and eventually determine most common 42 categories for fine-grained visual categorization. As for the rationality of the category selection, the categories are chosen by experts according to the importance of maritime situational awareness and ocean monitoring. It should be noted that through the investi-

gation of the remote sensing image data containing ship targets at this stage, FGSCR-42 basically contains more than 90% of ship targets that can be categorized.

As shown in Table 3, we compare four datasets including ship targets in remote sensing images. DOTA and NWPU VHR-10 treat ship targets as one whole category for object detection task. At the same time, the ship detection dataset HRSC2016 has only 19 categories, which are far from enough for the fine-grained classification of ship targets in remote sensing images. These datasets are mainly concerned with ship detection problems. Based our previous work DSCR [5], FGSCR-42 provides a lot of expansions to ship targets categories. When it comes to the problem of fine-grained classification of specific ship targets in remote sensing images, FGSCR-42 possess the absolute advantage due to the substantial expansion of ship categories, which will do a lot of contributions to ship detection and recognition algorithms research.

**Table 3.** Comparison among ship categories of FGSCR-42 and some other remote sensing datasets.

| Dataset | Ship Categories |
| --- | --- |
| DOTA | 1 |
| NWPU VHR-10 | 1 |
| HRSC2016 | 19 |
| DSCR | 7 |
| **FGSCR-42** | **42** |

### 3.3. Crop Details

Considering that the remote sensing images with ship targets have large spatial size, we adopt two image cropping methods. First, based on the previous research on remote sensing image datasets containing ship instance (e.g., DOTA, HRSC2016, NWPU VHR-10), we use the annotations of the dataset itself due to the high credibility and popularity of the public dataset in remote sensing research. In addition, the more important work we have done is cropping and labeling the large amount of remote sensing image data obtained from different sensors and platforms. As is well known, the ship targets always have high aspect ratio, while the input image size of convolutional neural networks is usually square. It means cropping directly according the bounding boxes will cause unpredictable distortion. In practice, we propose a way to solve the ratio problem. We consider ship targets in remote sensing images into two situations, on the sea surface and in shore. Ships on the sea surface are relatively easier to crop, while ships on shore have some disturbances such as buildings. We first calculate the center of the ship target according to the label, and then choose the larger edge of the bounding box as the side length of the cropped square to ensure that the cropped part includes entire ship target. We make our best to make sure that all ships contained in the same cropped image are in the same category. In our dataset, we mainly consider the case of one ship category per image. We remove the images of multi-label, i.e., the image contains more than two categories of ship targets. Images containing multiple ships will be added in updated version in the future. Figure 2 shows examples of ship target cropping details of FGSCR-42.

### 3.4. Annotation Method

We consider different ways of classification annotating. Considering the convenience of input, we use image names and labels to generate txt files, which is easily to read and write. We sorted the images into different folders by category. See Figure 3 for example images with annotation. Then, we generate txt files of labeled categories in the corresponding training and testing sets according to the split of FGSCR-42 in Section 3.6. Specifically, we provide the names of all the images and the categories of the random split of training and testing images.
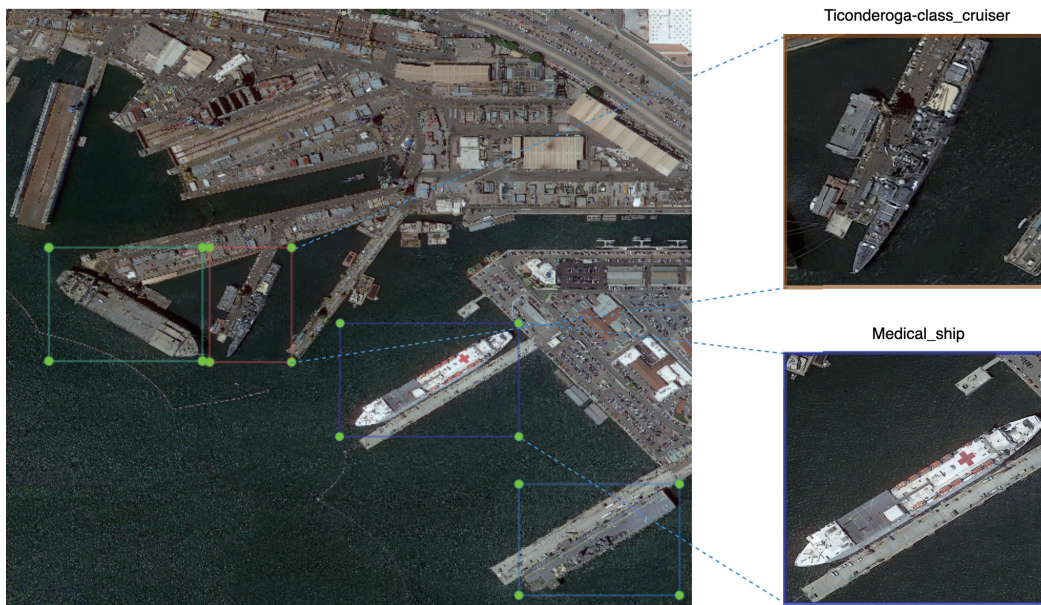
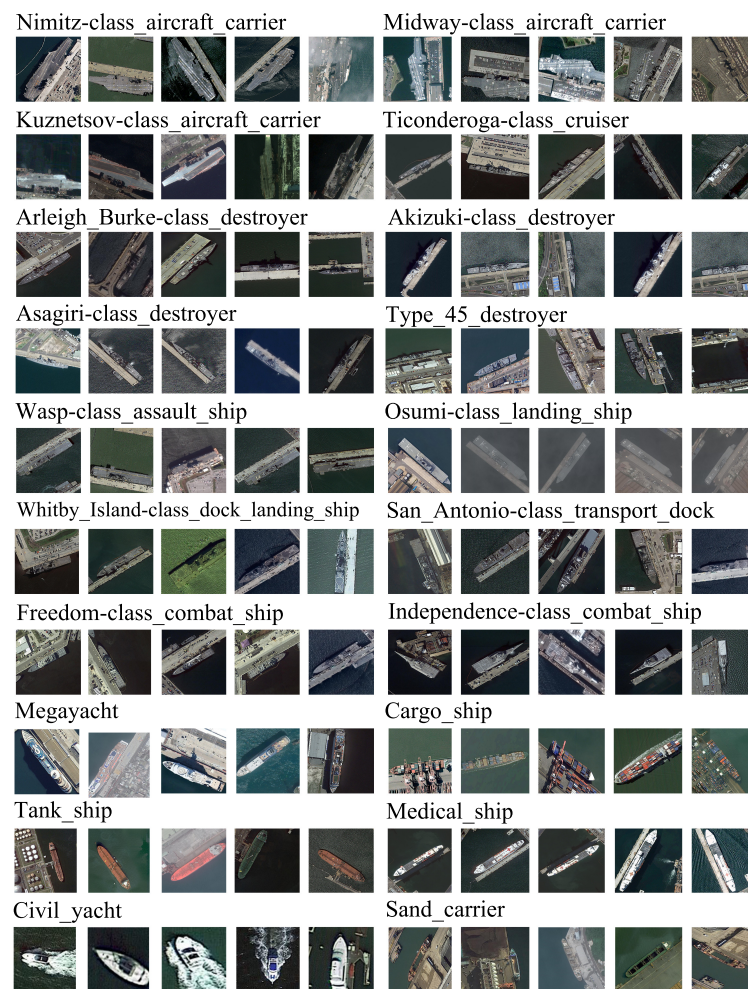**Figure 2.** Examples of ship target cropping details of FGSCR-42.



**Figure 3.** FGSCR-42 Example images with annotation information.

### 3.5. Augmentation

Considering the situation of remote sensing images, ship images are definitely affected by illumination and cloud occlusion, and their spatial resolutions are also different. Consequently, in order to balance the number of each category of images in the dataset, we implement data augmentation for FGSCR-42 by imitating different levels of light, rotation at different angles and cropping of different ratios in the up, down, left and right directions. Particularly, the augmentation ratio of each category is not fixed, and the purpose is to balance the image numbers between categories for better training of classifiers. Taking into account the actual number of ship targets and the requirements of deep learning algorithms, about 200 images per category are confirmed after augmentation. The change of image numbers of FGSCR-42 before and after augmentation can be seen in Figure 4. Owing to the augmentation, FGSCR-42 can better reflect the state of the ship instances in practical situations and have better generalization performance for fine-grained classification task.



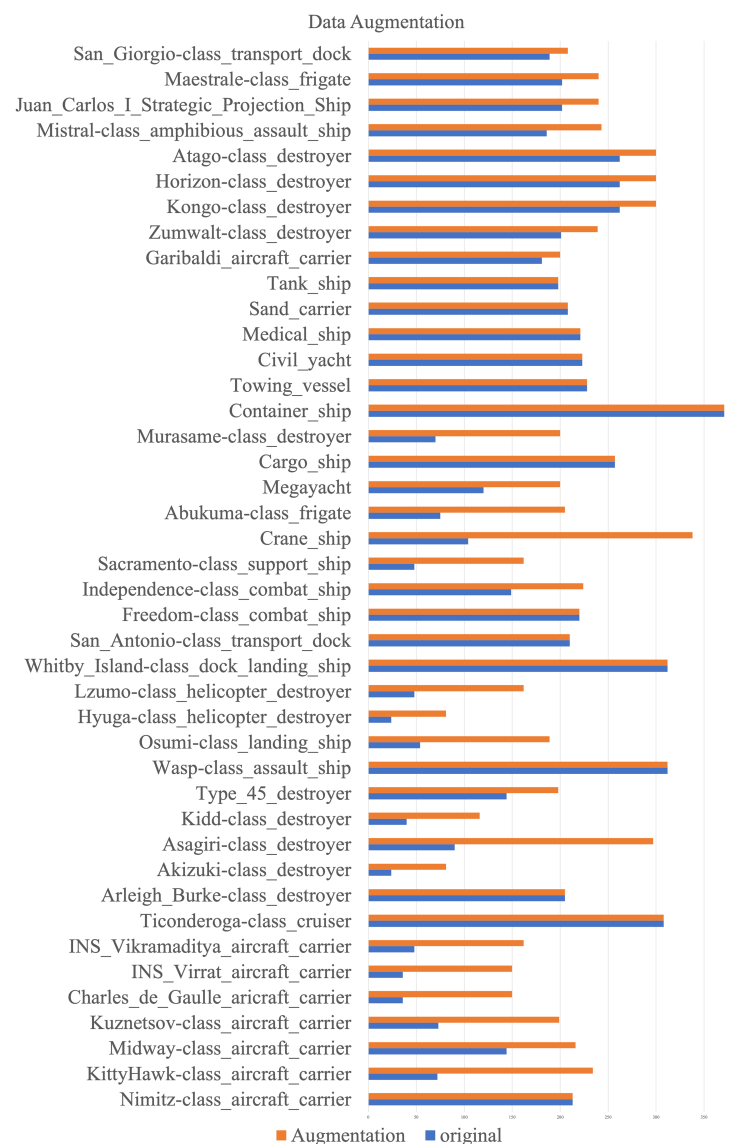**Figure 4.** Comparison of each category of FGSCR-42 between the image numbers after augmentation.

### 3.6. Dataset Splits

In order to make sure that the training data and test data distributions approximately match, referring to the distributions of many public classification datasets, we randomly select half of the original cropped images as the training set, and the rest half as the testing

set. We will publicly provide all the original images with ground truth for training set and testing set.

## 4. Properties of FGSCR-42

### 4.1. Image Size

The size of remote sensing images is generally large compared to those in natural image datasets. The size of images in FGSCR-42 ranges from about $50 \times 50$ to about $1500 \times 1500$ pixels while most images in natural fine-grained datasets (e.g., CUB-200-2011, Stanford Dogs and Stanford Cars) are no more than $500 \times 500$ pixels. It is precisely because of the huge difference in size of ship instances in remote sensing images that higher requirements are put forward for ship fine-grained classification.

### 4.2. Spatial Resolution Information

We also select images with different spatial resolutions in our dataset, which implies the actual size of an instance and plays a significant role in ship classification in remote sensing images. The importance of spatial resolution for fine-grained classification task has two aspects. Firstly, it allows the model to be more adaptive and robust for varieties of objects of the same category. It is known that objects appear smaller when seen from a longer distance. The accuracy of a classification model may reduce when a same object appears in different sizes. This is a common scaling problem for object classification. However, a model can pay more attention to the shape with resolution information provided instead of objects' size. Secondly, due to different spatial resolutions, the same ship instance may occupy different proportions in the image, which is more conducive to enriching the diversity of the dataset.

### 4.3. Various Sizes of Categories

We have previously established a public remote sensing image ship classification dataset, which is named DSCR [5] and contains seven categories of ship targets. The number of categories can meet the basic ship classification task and the recognition task in object detection of remote sensing images. However, the number of categories cannot meet the requirements of fine-grained recognition. Based on our previous work, we investigate more remote sensing images and collect ships of 42 different sub-categories grouped under 10 main categories in FGSCR-42, which has the most classification categories in remote sensing ships as far as we know. The label hierarchy is shown in Figure 5. It is worth mentioned that the 42 categories are formed according to the division of sub-categories containing major categories in DSCR.

### 4.4. Various Aspect Ratios of Instances

FGSCR-42 has various aspect ratios for ship targets. For example, large warships generally have a larger aspect ratio as small yachts generally have a smaller aspect ratio. In addition, the aspect ratio between different sub-categories of the same main-category may also vary greatly, which makes FGSCR-42 suitable for ship fine-grained classification with large intra-class differences and small inter-class differences in remote sensing images.
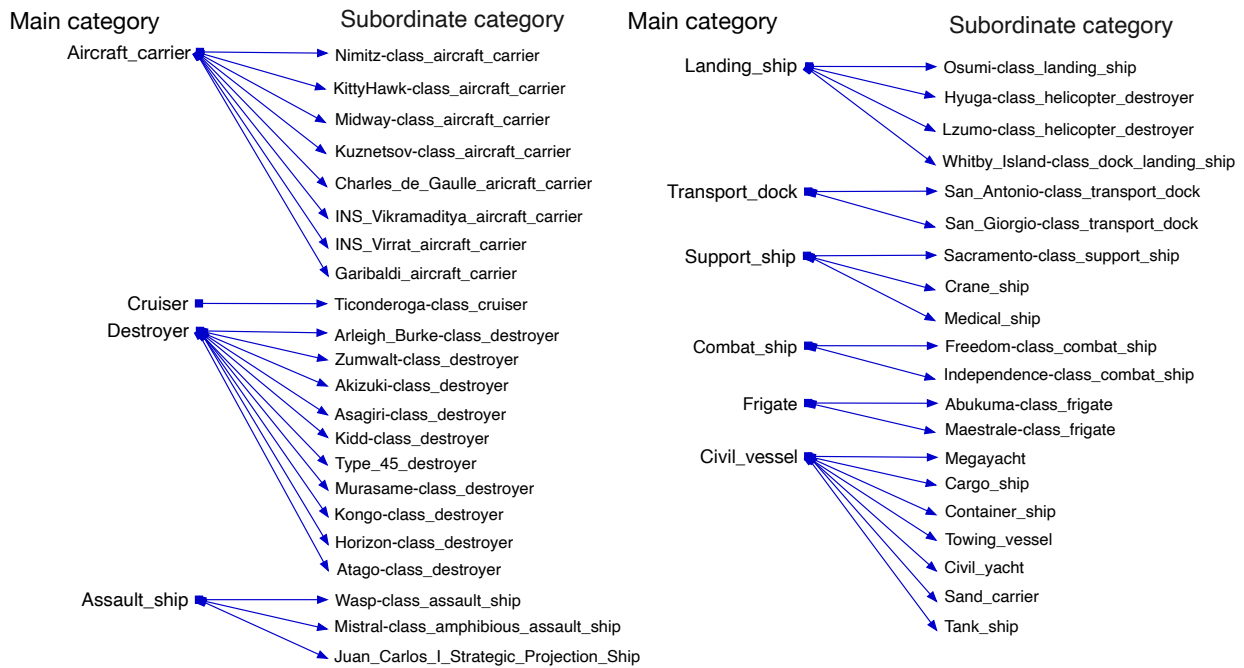
**Main category**      **Subordinate category**      **Main category**      **Subordinate category**

Aircraft_carrier → Nimitz-class_aircraft_carrier
→ KittyHawk-class_aircraft_carrier
→ Midway-class_aircraft_carrier
→ Kuznetsov-class_aircraft_carrier
→ Charles_de_Gaulle_aricraft_carrier
→ INS_Vikramaditya_aircraft_carrier
→ INS_Virrat_aircraft_carrier
→ Garibaldi_aircraft_carrier

Cruiser → Ticonderoga-class_cruiser

Destroyer → Arleigh_Burke-class_destroyer
→ Zumwalt-class_destroyer
→ Akizuki-class_destroyer
→ Asagiri-class_destroyer
→ Kidd-class_destroyer
→ Type_45_destroyer
→ Murasame-class_destroyer
→ Kongo-class_destroyer
→ Horizon-class_destroyer
→ Atago-class_destroyer

Assault_ship → Wasp-class_assault_ship
→ Mistral-class_amphibious_assault_ship
→ Juan_Carlos_I_Strategic_Projection_Ship

Landing_ship → Osumi-class_landing_ship
→ Hyuga-class_helicopter_destroyer
→ Lzumo-class_helicopter_destroyer
→ Whitby_Island-class_dock_landing_ship

Transport_dock → San_Antonio-class_transport_dock
→ San_Giorgio-class_transport_dock

Support_ship → Sacramento-class_support_ship
→ Crane_ship
→ Medical_ship

Combat_ship → Freedom-class_combat_ship
→ Independence-class_combat_ship

Frigate → Abukuma-class_frigate
→ Maestrale-class_frigate

Civil_vessel → Megayacht
→ Cargo_ship
→ Container_ship
→ Towing_vessel
→ Civil_yacht
→ Sand_carrier
→ Tank_ship

**Figure 5.** Label hierarchy of all categories in FGSCR-42.

## 5. Evaluations

### 5.1. Baseline Models

Common classification networks can be used for fine-grained ship classification in remote sensing images, but the results need to be improved. For the comprehensiveness of the results, we eventually selected VGG [16], ResNet [17], ResNext [18], and DenseNet [19], as our testing algorithms for building a benchmark on FGSCR-42. Furthermore, in order to apply FGSCR-42 to fine-grained ship classification task in remote sensing images, we selected four popular fine-grained recognition algorithms, namely B-CNN [20], RA-CNN [21], DCL [22], and TASN [23], for better verification. These baseline models cover the popular deep learning classification networks and the state-of-the-art fine-grained recognition networks, thus can offer valuable benchmark results for dataset evaluation.

For the task of fine-grained ship classification, the performance is evaluated as the accuracy of Top-1 classification. Formally, the Top-1 accuracy of *Class i* is defined as

$$P_i = \frac{N_i^c}{N_i} \tag{1}$$

where $N_i^c$ represents the correct number of each category in the test set and $N_i$ represents the number of each category in the test set. Then, overall precision $P$ can be calculated as

$$P = \frac{\sum_{i=1}^{C} N_i^c}{\sum_{i=1}^{C} N_i} \tag{2}$$

where $C$ represents the number of image categories.

### 5.2. Benchmark Results

Benchmark results of selected baseline models are shown in Table 4. Fine-grained classification network is different from image classification network. We selected four general CNN classification networks (VGG [16], ResNet [17], ResNext [18], and DenseNet [19]) and four fine-grained classification networks (B-CNN [20], RA-CNN [21], DCL [22], and TASN [23]) to build a benchmark on FGSCR-42. All of the general classification networks can be able to perform fine-grained classification tasks, but their results are not satisfactory.

Especially for the VGG series network, there is no advantage in training time and model size. Comparatively, the fine-grained image classification algorithms have better performance in fine-grained ship classification task than general CNN classification networks. In particular, we list the accuracy of each category when using the B-CNN algorithm for fine-grained classification on FGSCR-42 in Table 5. Calculating the accuracy of each category allows us to get the characteristics of the dataset. Aircraft carriers, for example, are more accurate than civil yachts. Similar characteristics of the dataset can be obtained by the B-CNN algorithm. The classification accuracy of some smaller ships on the spatial scale is not ideal, probably because the feature information of the lower accuracy categories is not rich and the intra-class difference is relatively large.

**Table 4.** Fine-gained classification results of baseline networks. (Top-1 Accuracy).

| Model | Accuracy | Model | Accuracy |
|---|---|---|---|
| VGG19 | 77.36 | B-CNN | 89.53 |
| ResNet-50 | 87.24 | RA-CNN | 91.63 |
| DenseNet | 88.69 | DCL | 93.03 |
| ResNext-50 | **89.16** | TASN | **93.51** |

**Table 5.** Accuracy of each category when using the B-CNN algorithm for fine-grained classification on FGSCR-42.

| Category | Accuracy | Category | Accuracy |
|---|---|---|---|
| 001.Nimitz-class_aircraft_carrier | 89.77 | 022.Sacramento-class_support_ship | 80.25 |
| 002.KittyHawk-class_aircraft_carrier | 94.05 | 023.Crane_ship | 95.83 |
| 003.Midway-class_aircraft_carrier | 98.35 | 024.Abukuma-class_frigate | 88.00 |
| 004.Kuznetsov-class_aircraft_carrier | 93.65 | 025.Megayacht | 90.63 |
| 005.Charles_de_Gaulle_aricraft_carrier | 95.29 | 026.Cargo_ship | 56.84 |
| 006.INS_Vikramaditya_aircraft_carrier | 95.83 | 027.Murasame-class_destroyer | 88.46 |
| 007.INS_Virrat_aircraft_carrier | 98.06 | 028.Container_ship | 91.50 |
| 008.Ticonderoga-class_cruiser | 98.84 | 029.Towing_vessel | 61.70 |
| 009.Arleigh_Burke-class_destroyer | 93.04 | 030.Civil_yacht | 40.00 |
| 010.Akizuki-class_destroyer | 89.29 | 031.Medical_ship | 88.61 |
| 011.Asagiri-class_destroyer | 92.76 | 032.Sand_carrier | 60.95 |
| 012.Kidd-class_destroyer | 89.74 | 033.Tank_ship | 96.43 |
| 013.Type_45_destroyer | 95.60 | 034.Garibaldi_aircraft_carrier | 93.75 |
| 014.Wasp-class_assault_ship | 97.88 | 035.Zumwalt-class_destroyer | 91.92 |
| 015.Osumi-class_landing_ship | 84.50 | 036.Kongo-class_destroyer | 90.18 |
| 016.Hyuga-class_helicopter_destroyer | 92.59 | 037.Horizon-class_destroyer | 93.75 |
| 017.Lzumo-class_helicopter_destroyer | 97.96 | 038.Atago-class_destroyer | 96.67 |
| 018.Whitby_Island-class_dock_landing_ship | 92.50 | 039.Maestrale-class_frigate | 87.13 |
| 019.San_Antonio-class_transport_dock | 97.48 | 040.Juan_Carlos_I_Strategic_Projection_Ship | 88.24 |
| 020.Freedom-class_combat_ship | 95.54 | 041.Mistral-class_amphibious_assault_ship | 94.25 |
| 021.Independence-class_combat_ship | 91.67 | 042.San_Giorgio-class_transport_dock | 98.59 |

## 6. Conclusions

In this paper, we presented FGSCR-42, a new large public dataset for fine-grained ship classification, which has a wide variety of categories of most warships and civil vessels in remote sensing images. The dataset contains 9320 ship images of 42 categories, about 200 images in each category. As the first public dataset specially released for fine-grained ship classification, we believe that FGSCR-42 has the potential of introducing ship recognition as a novel domain in FGVC to the wider computer vision community. It could be used to evaluate fine-grained ship classification methods and train ship detection methods as well. Considering the application of the dataset, we have also established a benchmark for fine-grained ship classification. This could be useful to future development of fine-grained classification algorithms in remote sensing images. In the future work, we

will further expand the amount of categories and the size of the dataset. If possible, we will invite more experts to make more precise annotations of the dataset.

**Author Contributions:** Conceptualization, Z.J. and H.Z.; methodology, Y.D. and H.Z.; software, Y.D.; validation, Y.D.; formal analysis, Y.D. and H.Z.; investigation, Y.D. and H.Z.; resources, Z.J. and H.Z.; data curation, Y.D.; writing—original draft preparation, Y.D. and H.Z.; writing—review and editing, H.Z.; visualization, Y.D.; supervision, Z.J. and H.Z.; project administration, Z.J. and H.Z.; and funding acquisition, Z.J. and H.Z. All authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1.  Xia, G.S.; Bai, X.; Ding, J.; Zhu, Z.; Belongie, S.; Luo, J.; Datcu, M.; Pelillo, M.; Zhang, L. DOTA: A large-scale dataset for object detection in aerial images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 23–26 October 2018; pp. 3974–3983.
2.  Liu, Z.; Wang, H.; Weng, L.; Yang, Y. Ship rotated bounding box space for ship extraction from high-resolution optical satellite images with complex backgrounds. *IEEE Geosci. Remote. Sens. Lett.* **2016**, *13*, 1074–1078. [CrossRef]
3.  Cheng, G.; Zhou, P.; Han, J. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Trans. Geosci. Remote. Sens.* **2016**, *54*, 7405–7415. [CrossRef]
4.  Fan, H.; Gui-Song, X.; Jingwen, H.; Liangpei, Z. Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery. *Remote Sens.* **2015**, *7*, 14680–14707.
5.  Di, Y.; Jiang, Z.; Zhang, H.; Meng, G. A public dataset for ship classification in remote sensing images. In Proceedings of the SPIE 11155, Image and Signal Processing for Remote Sensing XXV, Strasbourg, France, 9–12 September 2019; pp. 111551M–1–111551M–7.
6.  Wah, C.; Branson, S.; Welinder, P.; Perona, P.; Belongie, S. The Caltech-UCSD Birds-200-2011 Dataset. Number CNS-TR-2011-001. Available online: http://www.vision.caltech.edu/visipedia/CUB-200-2011.html (accessed on 26 October 2011).
7.  Krause, J.; Stark, M.; Deng, J.; Fei-Fei, L. 3D Object Representations for Fine-Grained Categorization. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Sydney, Australia, 2–8 December 2013; pp. 554–561.
8.  Khosla, A.; Jayadevaprakash, N.; Yao, B.; Fei-Fei, L. Novel Dataset for Fine-Grained Image Categorization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 554–561.
9.  Nilsback, M.; Zisserman, A. Automated Flower Classification over a Large Number of Classes. In Proceedings of the 2008 Sixth Indian Conference on Computer Vision, Graphics Image Processing, Bhubaneswar, India, 16–19 December 2008; pp. 722–729.
10. Maji, S.; Rahtu, E.; Kannala, J.; Blaschko, M.; Vedaldi, A. Fine-Grained Visual Classification of Aircraft. *arXiv* **2013**, arXiv:1306.5151.
11. Cheng, G.; Han, J.; Zhou, P.; Guo, L. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS J. Photogramm. Remote Sens.* **2014**, *98*, 119–132. [CrossRef]
12. Long, Y.; Gong, Y.; Xiao, Z.; Liu, Q. Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2486–2498. [CrossRef]
13. Ross, T.D.; Worrell, S.W.; Velten, V.J.; Mossing, J.C.; Bryant, M.L. Standard SAR ATR evaluation experiments using the MSTAR public release data set. In *Algorithms for Synthetic Aperture Radar Imagery V*; International Society for Optics and Photonics: Washington, WA, USA, 1998; Volume 3370, pp. 566–574.
14. Zou, Q.; Ni, L.; Zhang, T.; Wang, Q. Deep learning based feature selection for remote sensing scene classification. *IEEE Geosci. Remote. Sens. Lett.* **2015**, *12*, 2321–2325. [CrossRef]
15. Cheng, G.; Han, J.; Lu, X. Remote sensing image scene classification: Benchmark and state of the art. *Proc. IEEE* **2017**, *105*, 1865–1883. [CrossRef]
16. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
17. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
18. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated residual transformations for deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1492–1500.
19. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
20. Lin, T.Y.; RoyChowdhury, A.; Maji, S. Bilinear CNN Models for Fine-grained Visual Recognition. In Proceedings of the International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.

21. Fu, J.; Zheng, H.; Mei, T. Look Closer to See Better: Recurrent Attention Convolutional Neural Network for Fine-grained Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4476–4484.

22. Chen, Y.; Bai, Y.; Zhang, W.; Mei, T. Destruction and Construction Learning for Fine-Grained Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 5152–5161.

23. Zheng, H.; Fu, J.; Zha, Z.J.; Luo, J. Looking for the Devil in the Details: Learning Trilinear Attention Sampling Network for Fine-grained Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 5012–5021.