Logo detection has gradually become a research hotspot in the field of computer vision and multimedia for its various applications, such as social media monitoring, intelligent transportation, and video advertising recommendation. Recent advances in this area are dominated by deep learning-based solutions, where many datasets, learning strategies, network architectures, and loss functions have been employed. This article reviews the advance in applying deep learning techniques to logo detection. First, we discuss a comprehensive account of public datasets designed to facilitate performance evaluation of logo detection algorithms, which tend to be more diverse, more challenging, and more reflective of real life. Next, we perform an in-depth analysis of the existing logo detection strategies and their strengths and weaknesses of each learning strategy. Subsequently, we summarize the applications of logo detection in various fields, from intelligent transportation and brand monitoring to copyright and trademark compliance. Finally, we analyze the potential challenges and present the future directions for the development of logo detection. This study aims better to inform readers about the current state of logo detection and encourage more researchers to get involved in logo detection.

CCS Concepts: • General and reference → Surveys and overviews; • Computing methodologies → Artificial intelligence;

Additional Key Words and Phrases: Logo detection, computer vision, deep learning, datasets

ACM Reference format:
Sujuan Hou, Jiacheng Li, Weiqing Min, Qiang Hou, Yanna Zhao, Yuanjie Zheng, and Shuqiang Jiang. 2023. Deep Learning for Logo Detection: A Survey. ACM Trans. Multimedia Comput. Commun. Appl. 20, 3, Article 72 (October 2023), 23 pages. https://doi.org/10.1145/3611309

This work is supported by the National Nature Science Foundation of China (Grants No. 62072289, No. 61972378, and No. U19B2040) and CAAI-Huawei MindSpore Open Fund.

Authors’ addresses: S. Hou, J. Li, Q. Hou, Y. Zhao, and Y. Zheng, School of Information Science and Engineering, Shandong Normal University, No. 1 Daxue Road, Shandong, China, 250358; emails: sujuanhou@sdu.edu.cn, [2021317140, 2019309052]@stu.sdu.edu.cn, yannazhao@sdu.edu.cn, zhengyuanjie@gmail.com; W. Min and S. Jiang, Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, No. 6 Kexueyuan South Road, Beijing, 100190, China; emails: {minweiqing, sqjiang}@ict.ac.cn.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
1551-6857/2023/10-ART72 $15.00
https://doi.org/10.1145/3611309
1 INTRODUCTION

With the rapid development of multimedia information, the amount of logo image data on the Internet and social media has shown explosive growth in recent years. Logo images usually consist of texts, shapes, images, or combinations. Logo detection is gaining attention in multimedia due to its wide range of real-world applications, such as intelligent transportation [88, 110], social media monitoring [22, 31], infringement detection [8, 70], and video advertising recommendation [11, 12]. Meanwhile, some competitions have emerged, such as ACM MM2021 Robust Logo Detection Grand Challenge [8, 36, 51, 102] and ICME2022 Few-shot Logo Detection [89]. With the rapid development of powerful tools such as deep learning, logo detection has made great strides, advancing the development of multimedia fields.

The main task of logo detection is to determine the location of a specific logo in images/videos and identify them. Although it may be regarded as a particular object task, logo detection in real-world images can be pretty challenging, since numerous brands may have highly diverse contexts, varied scales, changes in illumination, size, resolution, and even non-rigid deformation (as shown in Figure 1). Many previous works on logo detection employ hand-crafted features (like SIFT [78]) to represent logos and use statistical classifiers for classification. Such methods suffer from complex image preprocessing pipelines and poor robustness when dealing with a much larger number of logos. Recent years have witnessed the rousing success of deep learning since ImageNet Large-scale Visual Recognition Challenge (ILSVRC) [43]. Deep learning-based solutions with expressive feature representation capability offer better robustness, accuracy, and speed and thus attract increasing attention. There are more than 100 papers about logo detection from 1993 to 2022, and a concise milestone of logo detection is shown in Figure 2. We can see that many deep learning-based logo detection strategies have been proposed since 2015. This survey mainly concentrates on deep learning-based solutions specially developed for logo detection.

Even though deep learning has dominated the logo research community, a comprehensive and in-depth survey on deep learning-based solutions is lacking. In this survey, we mainly focus on the advances in recent deep learning for logo detection. We provide in-depth analysis and discussion on existing studies in various aspects, covering datasets, pipelined used, task types, detection strategies, loss functions, and their contributions and limitations. We also try to analyze potential research challenges and future research directions for logo detection. We hope our work could provide a novel perspective to promote the understanding of deep learning-based logo detection, foster research on open challenges, and speed up the development of the logo research field.

The rest of this article is organized as follows. In Section 2, we investigate the public logo detection datasets. In Section 3, we review and organize the currently available work on logo detection. In Section 4, we introduce the applications of logo detection in real-world scenarios. In Section 5, we discuss its challenges and prominent future research directions. Finally, we summarize the whole text in Section 6.

2 LOGO DETECTION DATASETS

Deep learning has brought great success to object detection in recent years, where datasets play a crucial role. Datasets are not only a common basis for comparing and measuring the performance of algorithms but also an essential factor in supporting advanced object detection algorithms. This section provides an overview of the logo datasets for detection. In addition, we also give a brief summary of existing logo classification datasets.

In recent years, many logo datasets intended for detection have been created to solve the problem of large realistic datasets with accurate ground truth. The use of these datasets enables qualitative as well as quantitative comparisons and allows benchmarking of different algorithms. We
conducted statistics on existing datasets commonly used for logo detection and classified them into three types based on their scale: small-scale, medium-scale, and large-scale. Table 1 provides statistics on the available logo datasets, and Figure 3 gives some illustrative examples from these datasets.

The small-scale datasets include BelgaLogos [40], FlickrLogos-32 [77], and so on. As the first benchmark dataset proposed for logo detection, BelgaLogos [40] consists of 10,000 images of...
natural scenes, with 37 different logos and 2,695 instances of logos labeled with bounding boxes. FlickrLogos-32 [77], one of the most popular small-scale datasets for logo detection, comprises 32 different classes with 70 images in each class. The images in this dataset are mainly captured from the real world, and many contain occlusions, appearance changes, and lighting changes, making detecting this dataset very challenging.

Datasets with medium-scale include Logo-Net [28], QMUL-OpenLogo [86], FoodLogoDet-1500 [29], and so on. Logo-Net [28] is built for detecting logos and identifying brands from real-world product images. It consists of two sub-datasets, namely, Logo-18 and Logo-160. QMUL-OpenLogo [86] is an open benchmark for logo detection, constructed by aggregating seven existing datasets and building an open protocol for logo detection evaluation. The QMUL-OpenLogo dataset has a highly imbalanced distribution and significant variation in scale, which is critical to verify the performance of the detection algorithm. FoodLogoDet-1500 [29] is the first high-quality public dataset of food logos with uneven distribution among different food logo classes, which poses a challenge to the small sample food logo detection algorithms. The dataset is composed of 1,500 food logo classes with 99,768 images.

There are also some large-scale datasets, such as LogoDet-3K [96] and PL8K [52]. LogoDet-3K [96] divides all logos into nine super-classes based on the needs of daily life and the main positioning of common enterprises, namely, Clothing, Food, Transportation, Electronics, Necessities, Leisure, Medicine, Sports, and Others. The dataset consists of 3,000 logo classes, 158,652 images, and 194,261 logo objects. The number of logo classes contained in different super-classes varies greatly. For example, the "Food," "Clothes," and "Necessities" contain more images and objects than other super-classes. The imbalanced distribution across different logo classes of LogoDet-3K poses a challenge to effectively detecting logos with few samples. PL8K [52] is a large logo detection dataset constructed semi-automatically. The dataset consists of 7,888 logo brands and 3,017,146 images, and at least 20 images per class.

In addition, there are also some datasets built for logo classification, e.g., WebLogo-2M [82] and Logo-2K+ [97]. Weblogo-2M [82] is obtained by automatic web data acquisition and processing. The dataset excludes images with small widths and/or heights and duplicate images. Compared

| Scale      | #Datasets | #Logos | #Brands | #Images | #Objects | #Public | #Year |
|------------|-----------|--------|---------|---------|----------|---------|-------|
| Small      | BelgaLogos [40] | 37     | 37      | 10,000  | 2,695    | Yes     | 2009  |
|            | FlickrLogos-27 [41] | 27     | 27      | 1,080   | 4,671    | Yes     | 2011  |
|            | FlickrLogos-32 [77] | 32     | 32      | 2,240   | 5,644    | Yes     | 2011  |
|            | MICC-Logos [78] | 13     | 13      | 720     | —        | No      | 2013  |
|            | Logo-18 [28] | 18     | 10      | 8,460   | 16,043   | No      | 2015  |
|            | Logos-32plus [3] | 32     | 32      | 7,830   | 12,302   | No      | 2017  |
|            | TopLogo-10 [85] | 10     | 10      | 700     | —        | No      | 2017  |
|            | Video SportsLogo [57] | 20     | 20      | 2,000   | —        | No      | 2017  |
|            | VLD 1.0 [59] | 66     | 66      | 25,189  | —        | No      | 2017  |
|            | SportLogo [45] | 31     | 31      | 2,836   | —        | Yes     | 2020  |
|            | VLD-45 [115] | 45     | 45      | 45,000  | —        | No      | 2020  |
| Medium     | Logo-160 [28] | 160    | 100     | 73,414  | 130,608  | No      | 2015  |
|            | Logos-in-the-Wild [90] | 871  | 871 | 11,054  | 32,850  | Yes     | 2017  |
|            | QMUL-OpenLogo [86] | 352    | 352     | 27,083  | —        | Yes     | 2018  |
|            | PL2K [20] | 2,000  | 2,000   | 295,814 | —        | No      | 2019  |
|            | FoodLogoDet-1500 [29] | 1,500 | — 99,768 | 145,400 | Yes     | 2021  |
| Large      | Open Brands [38] | 1,216  | 559     | 1,437,812 | 3,113,828 | No      | 2020  |
|            | LogoDet-3K [96] | 3,000  | 2,864   | 158,652 | 194,261  | Yes     | 2020  |
|            | PL8K [52] | 7,888  | 7,888   | 3,017,146 | —        | No      | 2022  |
with other datasets, the Weblogo-2M presents three unique properties inherent to large-scale data exploration for learning scalable logo models: (1) Weak Annotation, (2) Noisy (False Positives), and (3) Class Imbalance. Logo-2K+ [97] is a large-scale high-quality logo dataset. It covers a variety of logo classes from the real world, and different types of logo images have various logo appearances, scales, and backgrounds, since they are collected from different websites. The dataset has high spatial coverage of categories, including Food, Clothes, Institutions, Accessories, Transportation, Electronics, Necessities, Cosmetics, Leisure, and Medical. The number of images is imbalanced among different categories. For instance, "Food" has 769 logo classes, while "Medical" has only 48.

As with common object detection, mAP [19] is the most commonly used evaluation metric to measure logo detectors.

3 LOGO DETECTION METHODS

In logo detection, logo classification is also an essential part. Therefore, we briefly summarize logo classification before introducing logo detection.

3.1 Logo Classification

As one of the most critical tasks in computer vision, logo classification aims to recognize the logo name corresponding to the input image. According to different feature extraction strategies, existing classification methods are divided into two categories: traditional machine learning-based methods and deep learning-based methods. Some representative methods will be described briefly in the following.

3.1.1 Traditional Machine Learning-based Methods. The traditional logo classification methods extract features through manual features, such as SIFT and HOG, and then classify them by a classifier. Support Vector Machine (SVM) [14] is a classifier that performs binary data classification in a supervised learning manner. In traditional logo classification, SVM also shows excellent performance [7, 50, 80, 112]. Carvalho et al. [7] proposed a self-learning and automatic detection method that performs detection without any prior data. The scheme automatically identifies the candidate regions of the localization logo, uses the HOG features extracted from the localization to train the object detector and some sub-detectors and uses the SVM to recognize the logo image.

K-Nearest Neighbor (KNN) is a supervised learning algorithm commonly used in classification [25, 100]. Gopinathan et al. [25] proposed a vehicle logo recognition system in 2018. They used Euclidean distance on the initial training dataset to summarize the pixel intensity distribution by applying each channel’s mean and standard deviation and used the K-means algorithm to cluster the color histogram features to group different logos. Then HOG and KNN algorithms extract logo features and classify logos.

3.1.2 Deep Learning-based Methods. In recent years, with the continuous development of deep learning, deep learning-based solutions have also been successfully applied to logo classification [2, 21, 27, 30, 33, 42, 68, 113]. Karimi et al. [42] used the DCNN logo recognition algorithm, which conducted a pre-trained model for feature extraction and then used SVM for logo classification. They also used transfer learning to improve existing pre-trained models for logo recognition. Finally, the fine-tuned deep model is applied to the parallel structure to obtain a more efficient deep model for logo recognition.

The latest trend in logo classification is to design efficient networks with limited resources [52, 61, 97, 109]. Wang et al. [97] proposed a Discriminative Region Navigation and Augmentation Network (DRNA-Net), which is capable of discovering more informative regions and expanding these regions for logo classification. DRNA-Net is mainly divided into four parts: navigator sub-network, teacher sub-network, region-oriented data augmentation sub-network, and scrutinizer.
sub-network. Specifically, the navigation sub-network first computes the amount of information in all regions generated by pre-defined anchors in the image. To make the navigation sub-network selects the most informative logo-relevant regions, the teacher sub-network then evaluates each region’s confidence that belongs to the ground-truth class. The region-oriented data augmentation sub-network can augment the selected regions to localize more relevant logo regions. Finally, the features of augment regions and the whole image are fused by the scrutinizer sub-network to obtain a unified feature representation for logo prediction. The architecture of DRNA-Net is shown in Figure 4. To distinguish visually similar logos, Li et al. [52] proposed a multi-task learning architecture named SeeTek based on deep learning and scene text recognition. This model combines deep metric learning and text recognition into a single model leading to considerable performance improvements. The architecture of SeeTek is shown in Figure 5.

### 3.2 Logo Detection

Logo detection can be seen as a particular case of general object detection. The purpose of logo detection is to detect logo instances of predefined logo classes in images or videos [28]. In this section, we will present the current state-of-the-art work on logo detection in detail.

Before the development of deep learning, early approaches to logo detection were implemented based on hand-crafted visual features (e.g., SIFT and HOG) and traditional classification models (e.g., SVM) [5, 13, 76, 93]. For example, Sahbi et al. [78] proposed a logo matching system in 2013,
where they used SIFT to conduct logo detection and recognition. The designed system can not only recognize and match multiple reference logos in an image but also is suitable for detecting similar logos. However, such traditional logo detection methods have certain limitations: (1) The region-selective search algorithm based on sliding windows lacks pertinence, especially with high time complexity; (2) the hand-designed features lack robustness to the variation of logo diversity.

In recent years of work, a variety of methods have applied deep learning and have achieved promising results on the task of logo detection. According to different problem scenarios, the existing works are divided into the following aspects, including small-size logo detection, multi-scale logo detection, logo detection in complex scenes, long-tail logo detection, open-set logo detection, and few-shot logo detection, and so on. We will introduce them in the following.

### 3.2.1 Small Logo Detection

Compared to large logos, small logos are difficult to detect for several reasons. First, small logos with the low resolution contain little visual information, so it is difficult to extract discriminative features. Moreover, because of its small size, it is more susceptible to environmental interference, which makes it difficult for the model to accurately locate and identify the small logo. Second, because a small logo covers a small area in the image, its surrounding box is more challenging to locate than a large logo. In the prediction process, the error caused by the deviation of pixels caused by the predicted surrounding box of a small logo is much greater than that caused by a large logo. Finally, there are usually fewer small logo samples. In the training process, when the manually set anchor box is significantly different from the actual boundary box of the small logo, the positive training sample of the small logo will be much smaller than that of the large logo. As a result, the model pays more attention to detecting the large logo, while ignoring the detection of the small logo. Therefore, it is necessary to effectively improve the detection accuracy of the small logos while maintaining the detection accuracy of the large logos.

Vehicle logos are usually considered small logos. In the case of complex backgrounds, a vehicle logo usually occupies only a small part of the image. Therefore, it is challenging to identify vehicle logos in real-world scenarios accurately. In view of the high real-time requirements, the YOLOs are very popular vehicle logo detection. Early YOLO [73] has fast detection speed, but at the expense of accuracy, especially for small objects. The later YOLOs, including YOLOv2 [74], YOLOv3 [34], YOLOv4 [4], YOLOF [9], YOLOR [95], YOLOX [23], YOLOv6 [53], and YOLOv7 [94], improve the overall performance of YOLO, striking a balance between speed and accuracy. At present, YOLO-based methods are widely used in vehicle logo detection [37, 104, 107, 116]. Inspired by the superiority of deep learning in feature extraction, Yin et al. [107] proposed a high-efficiency vehicle logo detection system based on YOLOv2 [74]. Compared with the traditional methods based on manual extraction features, the system has the advantages of self-learning features and direct image input and can realize the dual functions of positioning and recognition of the vehicle logo. Some researchers increase the number of small logo images through data augmentation, so that the model also has the opportunity to focus more on small logo samples during the training process. Yang et al. [104] proposed a data-driven enhanced training method based on YOLOv3 [34]. They combined a feature extraction network with a multi-scale decision scheme to improve the detection accuracy of vehicle logos. However, the detection results of vehicle logos in complex scenes are unsatisfactory. Based on Reference [104], Zhang et al. [116] analyzed the characteristics of vehicle identification in static images and found that both the network for feature extraction and the training strategy for detection has a significant impact on the accuracy of vehicle identification. They proposed a lightweight network structure with separable convolutions to replace traditional methods to improve detection accuracy and speed under the YOLO-based framework. Experiments show that this method improves the real-time performance of vehicle logo detection and also improves the detection accuracy of small-scale objects while maintaining the speed to a
certain extent. By adding peripheral information around the small logo as additional training data, richer contextual information can be obtained, so that the vehicle logo can be detected more easily. Jiang et al. [37] proposed an improved vehicle logo recognition method based on YOLOv4 [4]. They integrated more shallow information by adding a shallow output layer to reduce the loss of small object features in the vehicle logo. In addition, deformable convolution residual blocks are integrated to reconstruct the neck structure to make the model accurately fit the shape and position characteristics of different vehicle logos. Finally, the detection header based on the convolutional transformer block is proposed to obtain richer context information to reduce the impact of complex background on detection. Detection Transformers (DETR) [6] has problems detecting small objects, so it cannot be directly used for multi-scale logo detection. To solve this issue, Velazquez et al. [92] incorporated Feature Pyramid (FP) into DETR architecture. They first split the penultimate layer of the feature pyramid into four patches of the same size as the smallest feature map in the pyramid. They then fed these patches into the DETR pipeline for prediction. This method effectively detects small objects. However, the backward propagation computation of this method increases memory consumption. Although these methods have improved the detection accuracy of small logos to some extent, the detection performance of small logo in specific scenes, such as complex background, is not satisfactory.

3.2.2 Multi-scale Logo Detection. The multi-scale phenomenon exists widely in logos. When detecting multi-scale logos, large-scale logo detection requires strong semantic information for classification, while small-scale logo detection requires fine-grained spatial information to ensure accurate identification and positioning. The general object detectors perform poorly when they are directly used for multi-scale logo detection. The important reason is that the convolutional neural network deepens continuously, leading to the loss of shallow spatial information. To this end, Lin et al. proposed Feature Pyramid Networks (FPN) [58] in 2017, which mainly solves the multi-scale problem in object detection by changing the connectivity of the network. FPN dramatically improves the performance of small object detection without increasing the amount of calculation. In recent years, many works have adopted FPN to solve the multi-scale problem of logo detection [29, 38, 54, 62, 115] and these FPNs have achieved exciting results.

Aiming at the problems of multi-scale and other geometric variations of logos, Meng et al. [62] proposed obtaining sufficient features for logo (OSF-Logo) based on FPN. Specifically, they introduced the Regulated Deformable Convolution (RDC) module in a specific layer of FPN, which allows the sampling point positions of the convolution kernel at different positions to adaptively change according to the content, thereby adapting to the geometric changes of the logo. Meanwhile, they adopt an up-sampling operator in FPN to generate an adaptive kernel that fully aggregates rich semantic information. Therefore, OSF-Logo can quickly perceive a wide range of content and effectively detect multi-scale logos. Similarly, Jin et al. [38] designed a network called “Brand Net.” Brand Net uses FPN to extract multi-scale features. On this basis, it introduces an anchor refinement network in FPN to remove negative anchors, thus greatly reducing the number of anchor boxes. Therefore, it can effectively detect logos with different scales and significantly improve the detection performance of small logos. Recently, Hou et al. [29] proposed a multi-scale feature decoupling network (MFDNet) for logo detection to solve the problem of distinguishing multiple logo categories. MFDNet comprises two components: Balanced Feature Pyramid (BFP) and Feature Offset Module (FOM). The former is designed to fuse multi-scale features using the same deep feature map to integrate balanced semantic features. At the same time, the latter comprises an automatically learned anchor region proposal network for pixel-level offsets to search for the best features for logos. The architecture of MFDNet is shown in Figure 6. Previous works usually detect vehicle logos at the same scale without considering the multi-scale
variation of vehicle logos. However, the vehicle logos captured by the camera are of various sizes. It is essential to design an algorithm with multi-scale vehicle logo detection ability. To this end, Zhang et al. [115] proposed a multi-scale vehicle logo detection (SVLD) based on SSD [60] to assist in the recognition of vehicles by modifying the pretraining strategy, adding a feature extraction layer, and controlling the ratio of positive and negative samples. According to the characteristics of the vehicle, a preset frame and parameter settings are designed to effectively reduce redundant calculations and accelerate the convergence of the model quickly. Compared with SSD, the detection accuracy of SVLD is improved, but the time complexity is slightly increased due to the addition of the feature extraction layer. In addition, although SSD have competitive accuracy compared to other single-stage methods, they are still insufficient in detecting small objects. Li et al. [54] proposed an extended range dependence involutional network (LDI-Net) to solve the multi-scale problem of logos. In LDI-Net, new operators and a self-attention mechanism were combined through LD involution to alleviate the detection difficulties caused by large aspect ratios. The multilevel representation neural architecture search is introduced to construct the topological structure to detect multi-scale logo objects.

3.2.3 Logo Detection in Complex Scenes. Natural visual scenes are usually complex and diverse. Meanwhile, the logo itself has several problems, such as object interference, shape change, and poor perspective effect, which aggravate the difficulty of logo detection. In recent years, academic circles at home and abroad have attached great importance to object detection in complex scenes. When the image is affected by unfavorable factors such as illumination, rotation, occlusion, and so on, the performance of the detector will be significantly reduced. Therefore, it is necessary to design an efficient and robust logo detector [35, 62, 99, 104, 118].

It is challenging to locate the logo in the low-quality image captured under bad weather conditions. Wilms et al. [99] took the airline logos as an example and proposed a logo detection system under adverse conditions. The authors adopt the object proposal generation system AttentionMask [98]. Since the aviation system logo is relatively large, the AttentionMask system removes the scale of small objects to improve execution efficiency and reduce false positives. They also used a data augmentation solution to simulate the effects of severe weather to diversify the training data processing effects. The experimental results show that the proposed strategy can achieve good performance in detecting the logo in complex real-world environments. However, the data obtained under real-world conditions are not as ideal as the data generated by data augmentation, so the performance of the model in detecting images in real-world conditions will be degraded. Considering that locating logos in complex scenarios is challenging due to the large variety of types and appearance of logos, Jain et al. [35] proposed a logo detection framework called LogoNet, which includes a spatial attention module. LogoNet can capture rich spatial information and focus more precisely on the logo object within an image. LogoNet achieves a significant
performance gain within considerable computation time. To detect the same logo with different visual features, Zhao et al. [118] solved this problem by learning a transferable and flexible cross-style relationship under the meta-learning strategy. In addition to the Faster RCNN [75] framework, this method contains a sibling branch to create a pair-wise comparison environment. Meanwhile, the classification head of the detector is remodeled into a matching module that meta-learns how to classify regions through pair-wise matching.

3.2.4 Long Tail Logo Detection. In the traditional logo detection task, the distribution of training data is often subject to artificial checks and balances, that is, there is no significant difference in the number of samples between different categories. However, in the distribution of natural data, the distribution of different logos is often uneven. Using long tail data directly to train the model will lead to over-fitting head data and under-fitting tail data.

The key to solving the long tail distribution is to solve the problem of insufficient attention to the tail category. Most methods achieve sample balance by resampling or setting different weights for different types of losses during training [51, 96, 102]. Wang et al. [96] proposed a robust baseline method Logo-Yolo based on YOLOv3 [34]. They first adopt the K-means clustering algorithm to select the number of candidate anchor boxes and aspect ratio dimensions. They then used Focal loss to address the imbalance of samples in the dataset. Last, they introduced CIoU loss to improve bounding box regression results. Compared with YOLOv3 [34], the performance of this method is significantly improved in predicting small objects and objects in complex backgrounds. However, in some specific cases, the method also shows poor performance, such as detecting similar or occluded logos. In addition, transfer learning, meta-learning, metric learning, and other strategies can solve the problem of few samples. Yohannes et al. [108] designed a framework consisting of a domain adaptation that reduces the loss function between source-target datasets and also represents important source features adopted by the target dataset. The authors add nonlocal blocks and attention mechanisms to achieve a decent performance of logo detection.

3.2.5 Open Set Logo Detection. Open set logo detection relaxes the closed-world assumption that only seen classes appear at inference. It requires the ability to detect unknown logos as well as known ones. In recent years, it has gradually attracted the attention of some researchers [83, 86, 90]. Although there are not many works, it will be a hotspot of research in the future.

The previous logo detection is based on a closed set, assuming that each class has a large amount of training data, as well as fine-grained boundary box annotations. However, this assumption is completely untenable, because logo classes are constantly emerging and it is impossible to annotate every logo image in detail. To this end, Su et al. [86] proposed an open logo detection benchmark and a data augmentation strategy focusing on logo context optimization: Contextual Adversarial Learning (CAL). CAL takes an image with a logo object as input and generates context-consistent synthetic images that can be used as additional training data. However, the distribution of this synthetic image is different from the actual test image. To solve this issue, the same authors proposed a Multi-perspective Cross-class (MPCC) domain adaptation method [83] based on the baseline method CAL [86]. MPCC conducts feature distribution alignment in the data augmentation principle from two perspectives. One is to align the feature distribution between the synthetic logo image of one-shot icon supervised classes and the authentic logo image of the fully supervised class. The other is to align the feature distribution between the logo and non-logo images. This mitigates the domain shift problem between model training and testing on one-shot icon supervised logo classes and reduces the model’s overfitting to the fully labeled logo classes. The final combination of MPCC with Faster R-CNN [75] achieved good results.
3.2.6 Few-shot Logo Detection. Most existing logo detection methods are still data-driven. Few-shot logo detection requires the deep learning model to have a certain generalization ability through a semi-supervised method with less annotated data or a weakly supervised method with incomplete matching annotated data.

Currently, the commonly used methods of few-shot logo detection are mainly the object detectors based on deep learning, and some tricks are used to modify these methods, such as fine-tuning, data enhancement, and sample matching strategy [84, 85, 89, 105]. Su et al. [85] proposed an algorithm for generating Synthetic Context Logo (SCL) training images. When there are only a few available annotation data to train the model, this method can significantly improve the performance of logo detection, avoid a lot of manual labeling costs, and improve the generalization ability of the in-depth learning logo detector model. Later, they developed a deep transfer active learning algorithm named DTAL [84] to select the most valuable samples such that we can label the smallest number of samples to achieve maximum performance improvements in training detection models. Yang et al. [105] proposed an attention-net-based few-shot object detection (AN-FSOD) model for logo detection and recognition, which provides an explicit feature map for detection and classification. They built an end-to-end feature extractor and target detection model while fine-tuning the parameters for the few-shot dataset and achieved satisfactory results.

3.2.7 Others. There are some works for regular logo detection [28, 87]. For example, Hoi et al. [28] proposed a DeepLogo-DRCN scheme for logo detection and brand recognition by exploring several deep region-based convolutional network (DRCN) techniques for object detection. They adopt Selective Search [91] to generate regions of interests (RoIs) and input the RoIs into a fully convolutional network (FCN). Feature vectors are then obtained through fully connected layers (FCs) to train the final classifiers and bounding box regressors. The overall architecture is trained in an end-to-end way. This is the first work on deep learning-based logo detection, which provides a new perspective on logo detection research.

Training a large-scale logo detection model requires a large amount of data, and the logo datasets usually suffer from data scarcity. Data augmentation and transfer learning are common solutions to release the problem of data scarcity. Oliveira et al. [64] proposed an automatic graphic logo detection system based on Fast R-CNN [24], which is robust to unconstrained imaging conditions. They used transfer learning and data augmentation to train a convolutional neural network model and allowed multiple detections of potential regions containing objects. Experimental results demonstrate that this strategy is superior to the traditional methods [77]. However, the adopted SS algorithm still generates redundant calculations, reducing detection speed. In addition, Li et al. [55] built Faster R-CNN for logo detection by using transfer learning, data augmentation, and clustering to guarantee suitable hyperparameters and more precise anchors for RPN. The experimental results show that this improvement could significantly improve the detection accuracy. Montserrat et al. [63] segmented the background image and extracted the depth information to synthesize a new and actual logo image. This method can create many images with only a few manual annotations.

Training a logo detector usually requires a large amount of labeled data, and labeling logo data is time-consuming. Therefore, weakly supervised learning is a vital strategy to solve this issue. Kumar et al. [44] proposed a method for logo detection by using weakly supervised learning of CNN to generate a deep saliency map, which is generated by a single back-propagation of a trained Deep CNN. The authors also used an interactive Grabcut algorithm to calculate the salient segmented regions of the image. However, this method is not sufficiently modeled to detect multiple brand logos in the same image.

To analyze current state-of-the-art deep learning-based methods more comprehensively, the representative models are presented in chronological order in Table 2, and the experimental results...
Table 2. Summary of Representative Deep Learning-based Logo Detection Methods

| Method type | Method | Pipeline | Task | Strategy | Loss function | Year |
|------------|--------|----------|------|----------|---------------|------|
| RCNNs      | DeepLogo-DRCN [28] | RCNN, SPPNet | Regular logo detection | – | – | 2021 |
|            | Improved Faster RCNN [18] | Faster RCNN | Regular logo detection | Transfer learning | – | 2016 |
|            | Deep Region-based Convolutional Networks [30] | Faster RCNN | Small logo detection | Transfer learning | K-means clustering, Data augmentation | – |
|            | Video logo detector [37] | Faster RCNN | Regular logo detection | Transfer learning | – | 2017 |
|            | Faster RCNN+ResNet-50 [3] | Faster RCNN | Long tail logo detection | Data augmentation | Category balance | 2020 |
|            | Cascade detectorRS [36] | Cascade RCNN | Small logo detection | Data augmentation | Multi-scale training | EQLv2 |
|            | Cascade RCNN+Res2Net101 [31] | Cascade RCNN | Long tail logo detection | Data augmentation | Data Sampling | EQLv2 |
|            | Green Hand [102] | DetectorRS | Long tail logo detection | Data augmentation | Resampling, Weighted boxes | EQLv2 |
| YOLOS      | Improved YOLOv3 [184] | YOLOv3 | Small logo detection | Softmax loss | – | 2019 |
|            | Scene recognition CNN [35] | YOLOv3 | Regular logo detection | – | Coordinate error loss, Prediction box loss | 2020 |
|            | Improved YOLOv2 [107] | YOLOv2 | Regular logo detection | Separate convolution | Multi-scale features fusion | – |
|            | Improved YOLOv2 [103] | YOLOv2 | Multi-scale logo detection | Transfer learning | Anchor refinement | – |
|            | Logo-Yolo [96] | YOLOv3 | Regular logo detection | Improved loss | K-means clustering | – |
|            | Scaled YOLOv4 [67] | YOLOv4 | Regular logo detection | – | – | 2021 |
|            | Separable VLD [118] | YOLO | Small logo detection | Separable convolution | – | 2021 |
|            | Improved YOLOv4 [37] | YOLOv4 | Small logo detection | Separable convolution | – | 2022 |
| SSDs       | SVLD [113] | SSD | Multi-scale logo detection | – | – | 2019 |
|            | SSD InceptionV2 [65] | SSD | Regular logo detection | – | – | 2022 |
| FPNs       | Brand Net [38] | – | Small logo detection | Soft mask attention, Weight transfer | Class-aware smooth L1, Cross-entropy loss | 2020 |
|            | OFS-Logo [62] | – | Multi-scale logo detection | Deformable convolution, Up-sampling operator | Class-agnostic smooth L1, Cross-entropy loss | 2021 |
|            | MFNet [29] | – | Multi-scale logo detection | Multi-scale feature fusion, Deformable learning | – | 2021 |
|            | DETR-FP [92] | DETR | Small logo detection | Model fusion | Hungarian loss, Generalized IoU loss | 2021 |
|            | LDI-Net [54] | – | Multi-scale logo detection | Independence, Mutualization, Neural architecture search | Cross-entropy loss, Dynamic smoothL1 loss | 2023 |
| Others     | SLST [52] | Faster RCNN | Weakly supervised logo detection | Incremental learning, Self-training | – | 2017 |
|            | SCL [51] | Faster RCNN | Few-shot logo detection | Synthetic context logo | – | 2018 |
|            | CAL [56] | Faster RCNN, YOLOv2 | Open set logo detection | Context adversarial learning | Softmax cross-entropy loss, Conditional adversarial loss | 2018 |
|            | Weakly supervised learning of CNN [54] | DCNN | Weakly supervised logo detection | Generate saliency map, GradCut segmentation | Softmax loss | 2018 |
|            | Attention logo detection system [99] | AttentionMask | Complex scenes logo detection | Data augmentation | Cross-entropy loss | 2021 |
|            | LogNet [33] | – | Complex scenes logo detection | Spatial attention | – | 2021 |
| Deep attention networks [108] | – | Long tail logo detection | Domain adaptation, Nonlocal block | – | 2021 |
|            | MPCC [83] | Faster RCNN | Open set logo detection | Domain adaptation | Data augmentation | Softmax cross-entropy loss | 2021 |
|            | AN-FSOD [105] | – | Few-shot logo detection | – | cross-entropy loss, MSE loss, MAE loss | 2022 |
|            | DTAL [84] | DCNN | Video logo detection | Active learning, Transfer learning | Meta-class loss | 2021 |

of these models on three popular benchmark datasets, namely, Flickrlogos-32, QMUL-OpenLogo, and Open Brands, are listed in Tables 3, 4, and 5, respectively.

4 APPLICATIONS

Logo detection is widely used in various applications, such as intellectual property protection, autonomous driving, and product branding. In this section, we describe the representative applications of logo detection in detail.

ACM Trans. Multimedia Comput. Commun. Appl., Vol. 20, No. 3, Article 72. Publication date: October 2023.
Table 3. Performance Comparison of Logo Detectors on the FlickrLogos-32 Benchmark

| Method                      | mAP(%) | Year |
|-----------------------------|--------|------|
| FRCN [33]                   | 74.4   | 2015 |
| BD-FRCN-M [64]              | 73.5   | 2016 |
| Video Logo Detector [57]    | 78.6   | 2017 |
| Faster RCNN+VGG16_D2_M3 [55]| 90.3   | 2017 |
| Faster RCNN+CAL [86]        | 74.9   | 2018 |
| Deep Saliency Map [44]      | 75.8   | 2020 |
| Logo-Yolo [96]              | 76.1   | 2020 |
| LogoNet [35]                | 82.2   | 2021 |
| OSF-Logo [62]               | 87.0   | 2021 |
| RetinaNet [79]              | 40.0   | 2021 |
| RCNN [79]                   | 65.0   | 2021 |
| Faster RCNN [79]            | 74.0   | 2021 |
| Scaled YOLOv4 [67]          | 80.4   | 2021 |
| MFDNet [29]                 | 86.2   | 2021 |
| LDI-Net [54]                | 89.8   | 2023 |

Table 4. Performance Comparison of Logo Detectors on the QMUL-OpenLogo Benchmark

| Method                      | mAP(%) | Year |
|-----------------------------|--------|------|
| YOLOv2+CAL [86]             | 49.2   | 2018 |
| Faster RCNN+CAL [86]        | 51.0   | 2018 |
| Logo-Yolo [96]              | 53.2   | 2020 |
| Faster RCNN+CAL+MPCC [83]   | 49.4   | 2021 |
| OSF-Logo [62]               | 53.3   | 2021 |
| Scaled YOLOv4 [67]          | 61.9   | 2021 |
| MFDNet [29]                 | 51.3   | 2021 |
| LDI-Net [54]                | 56.3   | 2023 |

Table 5. Performance Comparison of Logo Detectors on the Open Brands Benchmark

| Method                      | mAP(%) | Year |
|-----------------------------|--------|------|
| Brand Net (SMA) [38]        | 60.4   | 2020 |
| Faster RCNN+RFS+MST [8]     | 64.6   | 2021 |
| Cascade RCNN+Soft NMS [36]  | 65.1   | 2021 |
| Green Hand [102]            | 65.5   | 2021 |
| Cascade RCNN+DCN v2 [51]    | 70.2   | 2021 |

### 4.1 Brand Monitoring

The most obvious use case for logo detection is brand monitoring, which includes brand protection, brand recommendation, and brand identification. Logo plays an essential role in business marketing as a unique brand identity. However, there are many types of brands in real life, and there may be subtle logo differences between the two products. Inter-class similarities and intra-class differences will greatly increase the difficulty of detection. In addition, images may have
highly diverse context, illumination, projection transformation, and resolution. Therefore, how efficiently detecting brands in images is a challenging task. Currently, there is a lot of work on brand identity [31, 38, 64, 65]. For brand logos, the image usually includes some textual description. Textual information is an important source of information for brand identity. Hu et al. [31] proposed a multimodal fusion framework for logo detection. The framework combines image-based logo recognition with contextual features for logo detection using a natural language model. Additional contextual information can alleviate the limitations of detection. The performance of this model for logo detection has been improved, but there are still shortcomings in positioning.

4.2 Copyright and Trademark Compliance

With the development of global e-commerce platforms, the logos of major brands have become an important element in the e-commerce market. Some attackers use illegal means to infringe on the original author, so protecting intellectual property rights has also become a focus. Logo detection plays a vital role in preventing the increasing number of counterfeit transaction attempts online [8, 26, 36, 50, 51, 102, 117]. Infringement detection of logos has been studied for images and videos. Chen et al. [8] built a highly optimized and robust detector by using techniques such as data augmentation for the detection of logos (515 categories) in e-commerce images. Patalappa et al. [70] extended the dataset through data augmentation techniques. They applied the SSD algorithm to detect broadcast logos in the context of broadcast and broadband content piracy and proved the effectiveness of the application of SSD in content piracy video analysis monitoring tasks. An efficient way to identify the source of a web video is to detect some specified logos. To identify the source of illegal videos, Ye et al. [106] proposed a logo detection system “GeLoGo” to detect logos in Web-scale videos. The system consists of four main modules: the keyframe module, the proposal module, the spatial verification module, and the temporal verification module. The keyframe module first extracts several frames with short fragment detection, which can significantly reduce the data. Second, the proposal module extracts candidate boxes through a set of pre-trained identity detectors. Then, the spatial validation module detects logos through the ResNet network. Finally, the temporal verification module detects the detection result by detecting the recognition position.

4.3 Intelligent Transportation

In intelligent transportation, vehicle logo detection is significantly growing as it can effectively assist vehicle management. In addition, due to the increasing number of vehicle violations and traffic accidents, the realization of intelligent transportation systems is crucial. However, in the real world, the detection of vehicle logos is seriously disturbed due to occlusion, illumination, and low image resolution. Therefore, an efficient vehicle logo detection system is essential. In recent years, methods for vehicle logo detection have been proposed successively [25, 59, 68, 81, 100, 103, 104, 112]. Lu et al. [61] designed a feature extraction module VLF-Net and category-consistent mask learning module for vehicle logo detection. Yu et al. [110] proposed a cascaded deep convolutional neural network to detect vehicles without relying on the presence of license plates. The network is a two-stage framework including two components: a region proposal network and a convolutional capsule network. The former is responsible for generating region proposals that may contain vehicle logos, while the latter is responsible for classifying the proposals into background and vehicle logos. Surwase et al. [88] proposed a multi-scale multi-stream deep network for vehicle logo detection. The network processes the input image through a multi-scale stream to extract robust features, followed by logo recognition. The network follows a knowledge sharing strategy, where the learned features are shared on each stream of the network. Yu et al. [111] proposed VLDetNet for vehicle logo detection. It introduces the multi-branch block (vBlock), which can be combined with other networks to improve recognition accuracy. They built vMNetV1 as a feature extraction
network by combining vBlock and MobileNetV1. In addition, VLDetNet extracts the spatial attention information from the underlying feature map as the spatial attention weight to facilitate the extraction of the vehicle logo.

4.4 Document Categorization

A logo often appears in commercial documents and can be used as a statement of ownership. With the increasing demand for intelligent document image analysis and retrieval, automatic identity detection has made great progress [15, 17, 46–49, 56, 71]. For example, Ogier et al. [46] proposed a key-point matching method that combines SIFT and BRIEF descriptors for document retrieval. The BRIEF descriptor is used to post-filter the matching keys, thus increasing the number of matching keys with the same precision level. Alaei et al. [1] proposed a complete system for logo detection in document images. They proposed a template-based recognition strategy for under or over-segmentation problems that may arise during detection. To speed up the matching of templates in the recognition phase, a search space reduction strategy that utilizes the geometric properties of logo-patches and template logo-models is proposed to reduce the number of template logo-models. Zhu et al. [119] proposed a multi-scale method for document image logo detection and extraction. The authors used an augmentation strategy across multiple image scales to classify and localize logos accurately.

4.5 Other Applications

Logo detection can offer various benefits for advertisers looking to improve their marketing strategies effectively. As shown in Figure 7, Liao et al. [57] proposed a video logo detection framework that used a mutual-enhanced approach to improve logo detection through the information obtained from other simultaneously occurred logos. Based on the Fast R-CNN model, a uniformity-enhanced reordering method is proposed to improve the accuracy of in-frame region suggestions by analyzing the features of logos in videos. In addition, they proposed a frame propagation enhancement method to assist in the detection of adjacent frames. Finally, the effectiveness of the method is demonstrated on the FlickrLogos-32 dataset and the video dataset. Recently, food logo detection [29] has become more popular, because it has many valuable applications, such as food brand management and displaying nutritional information. Logo detection is also used in other areas, such as video semantic annotation [32, 66, 67] and noise recognition [10, 27, 72, 80]. TV logo is one of the important symbols that distinguish TV stations. Therefore, it is of great significance to effectively identify TV logos, because they have been widely used in video analysis, video management, and other fields [39, 101]. However, it is a challenging task, because TV logos are often varied, have small areas, less information, and complex backgrounds. To improve the
generalization ability of the model, Zhang et al. [114] proposed a two-stage scalable logo detection and recognition method (SLDR), which uses the batch-hard metric learning method to train the matching model to determine the logo category quickly. The detection and recognition mechanism of SLDR can also extend the detection target to unknown categories. Since the TV logo photographed is susceptible to illumination, occlusion, noise, and other factors, the TV logo may only take up a small part of each image. To this end, Pan et al. [69] proposed an efficient CNN to solve this problem. They used the maximum stable extremal region (MSER) algorithm to extract candidate frames. Since this algorithm tends to produce a large number of candidate frames without TV logos, they designed certain geometric constraints to remove non-logo objects. Finally, the images were fed into the CNN network for classification and achieved good results.

5 CHALLENGES AND FUTURE DIRECTIONS

5.1 Challenges

Logo detection has received more attention in the past few years for its wide applications. However, robust and accurate logo detection is still challenging in real-world scenarios, because the logo images have distinctive properties, which may form significant obstacles to current progress. We try to summarize them as following:

1. Small size: Unlike generic objects, logos tend to be small in size, making it difficult to distinguish them from contexts, especially the complex background. Although contextual information is crucial in small object detection, because small objects carry limited information [16], it is not always beneficial for logo detection, especially when it is too complex and unrelated to the target logo, as shown in Figure 8(a). We know that in the deep learning-based detection architecture, different layers of feature maps with various spatial resolutions are produced due to pooling and subsampling processes. Early-layer feature maps represent small reception fields but lack high-level semantic information critical for logo detection. In contrast, high-level feature maps help identify large logos but may not detect smaller ones.

2. Highly diverse backgrounds: A logo is usually associated with other visible entities, like a particular service or product. Thus, a logo may appear in various situations superimposed on objects such as geometrical renderings, shirts of persons, jerseys of players, boards of shops, billboards, or posters on sports playfields. As shown in Figure 8(b), a brand like “Nike” may serve many types of products, such as jackets, shoes, or trousers, resulting in the same logo being attached to a variety of different objects, which poses challenges to designing a robust logo detector when considering image statistics from the entire image.

3. Sub-branding: Sub-branding is often considered when a company releases a new product. Sub-brand detection facilitates brand monitoring but suffers from the same problems as fine-grained classification, such as subtle differences between sub-brands and parent brands and between sub-brands. In real logo detection, some sub-brands are treated as logos different from the parent brand or other sub-brands, because they have customer expectations and personalities that are different from the parent brand. Thus, the parent brand and its sub-brands usually have remarkably similar context information (background, packaging, etc., as shown in Figure 8(c)). When considering image-level feature maps, in this case, there are probably several remarkably similar feature maps extracted from different logo images, which challenges identifying different sub-brand logos accurately when using the deep learning-based detector.

Many approaches to tackle logo detection tasks have been developed based on object detection methods and have achieved promising results, but there is still much room for improvement. For example, Reference [104] proposed a vehicle logo detection system based on YOLOv3. Though the performance of small logo detection has been improved to a certain extent, it fails to accurately detect small logos with complex letter patterns. Reference [18] designed an improved anchor-based
scheme for small logos by utilizing higher resolution feature maps. However, it is too computationally expensive for constrained scenarios and too slow for real-time applications due to adopting a two-phase detection strategy. It is worth mentioning that the best detection performance only reaches 61.9% of mAP on QMUL-OpenLogo [67], in which the logo detector is designed based on YOLOv4 [4]. Therefore, simply borrowing the existing detectors from general object detection is considered insufficient for logo detection. Designing a new detection paradigm that is suitable for logos is essential.

5.2 Future Directions

Some future research directions in logo detection are as follows:

**Lightweight logo detection:** Lightweight networks are designed to reduce model complexity further while maintaining model accuracy. With the rapid development of mobile terminals such as mobile phones and computers, people pay more attention to lightweight detectors. Lightweight logo detection can help consumers quickly identify and accurately search for the products they need and thus can significantly improve the efficiency of online shopping. Although significant efforts have been made recently, the speed gap between the machine and the human eye is still huge. The detection accuracy is also insufficient, especially for small logos. Therefore, improving the accuracy of lightweight logo detectors is one of the future development trends.

**Weakly supervised logo detection:** A fully supervised logo detector must be trained on fully labeled large-scale logo datasets. However, the labeling of existing datasets is mainly done manually, which poses a challenge to the generalization and robustness of existing models due to the problem of data quality. With the rapid growth of logo data volume, data annotation costs are getting higher and higher. Developing weakly supervised logo detection techniques that only use image-level annotations or partial bounding box annotations to train logo detectors significantly
reduces costs and improves detection flexibility. Therefore, developing weakly supervised logo detection methods without fully labeled training data is an important issue for future research.

**Video logo detection:** Logo promotion exists in most video advertisements. Accurate and reasonable advertising is targets businesses, so it is essential to identify the target logo effectively. Current logo detectors are commonly used to detect logos of individual images, which may lead to a lack of correlation between consecutive images. In addition, there are problems such as motion blur and target occlusion in logo detection of videos, which can highly affect the performance of the detector. There are already methods trying to detect video logos. Currently, there are methods [57, 84] trying to detect video logos, but they have not achieved satisfactory results. Therefore, it is an important research direction to improve logo detection by utilizing temporal and spatial correlations.

**Tiny logo detection:** Tiny logos contain insufficient information, making it more challenging to extract discriminative features. Meanwhile, tiny logos contain relatively few samples and thus are easily disturbed by environmental factors. Numerous real-world tiny logos may have highly diverse contexts, which can cause the same logo to appear very different in different real scenarios. Observing the results of existing detection methods on tiny logos, we can find that the performance of this methods [92, 104, 116] is unsatisfactory in real scenarios.

**Long tail logo detection:** With the rapid development of society, some enterprises are currently leading the way, and their enterprise logos naturally often appear in front of the public. In contrast, the logos of most small enterprises are not paid much attention to, because they appear less. The long-tail distribution of logos exists all the time in reality. Some logo classes contain a large number of samples, while others contain few samples, making it difficult for logo detectors to detect logo classes with few samples. In previous logo detection methods, there is little focus on the long-tailed distribution in logo images. Therefore, it is one of the future tasks to utilize the unbalanced data to train efficient and accurate detectors.

**Incremental logo detection:** As a company’s signature or symbol, the logo will continue to increase as the number of companies increases. Therefore, the logo detector must learn new logos. However, retraining the logo detector is time-consuming and resource-intensive by discarding the previously learned knowledge. Incremental learning enables a system to continuously learn new knowledge from new samples while retaining most previously learned knowledge without catastrophic forgetting. It can reduce the complexity of the model to a certain extent, thereby reducing the training cost of the model. However, no incremental learning method currently can show good performance under all conditions. Therefore, incremental learning for logo detection is one of the future research directions.

In fact, existing studies generally assume close-environment scenarios. For example, the data distribution of logo datasets is invariable, and all the training data have known classes and fine-grained bounding box annotations in advance. Although Su et al. [86] proposed an assumption to simulate the open deployment, they mainly focus on the problem of limited logo data by presenting a context adversarial learning approach to generate context-consistent synthetic training data automatically. How to enable the trained model to be updated according to emerging new logo classes or data distribution change, or even we hardly know what changes is an essential requirement in open-environment.

### 6 CONCLUSION

In this work, we offer a comprehensive survey of previous and current approaches that helps logo detection. We mainly review the recent advances in deep learning-based logo detection by summarizing classical solutions. In addition, we comprehensively review the commonly used datasets, summarize the related applications of logo detection, and predict future research directions.
Although the achievement of logo detection has been effective recently, there is still much room for further development. We hope this article will better inform readers about the current development of logo detection and inspire more people to get involved in logo detection.

REFERENCES

[1] Alireza Alaei and Mathieu Delalandre. 2014. A complete logo detection/recognition system for document images. In Proceedings of the 11th IAPR International Workshop on Document Analysis Systems. 324–328.
[2] Marisa Bernabeu, Antonio Javier Gallego, and A. Pertusa. 2022. Multi-label logo recognition and retrieval based on weighted fusion of neural features. Retrieved from https://arXiv:2205.05419
[3] Simone Bianco, Marco Buzzelli, Davide Mazzini, and Raimondo Schettini. 2017. Deep learning for logo recognition. Neurocomputing 245 (2017), 23–30.
[4] Alexey Bochkovskiy, Chien Yao Wang, and H. Liao. 2020. YOLOv4: Optimal speed and accuracy of object detection. Retrieved from https://arXiv:2004.10934
[5] Raluca Boia, Corneliu Florea, and Laura Florea. 2015. Elliptical ASIFT agglomeration in class prototype for logo detection. In Proceedings of the British Machine Vision Conference. 115.1–115.12.
[6] Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, and Sergey Zagoruyko. 2020. End-to-end object detection with transformers. In Proceedings of the European Conference on Computer Vision. 213–229.
[7] Pedro Carvalho, Américo Pereira, and Paula Viana. 2021. Automatic TV logo identification for advertisement detection without prior data. Appl. Sci. 11, 16 (2021), 74–94.
[8] Hang Chen, Xiao Li, Zefan Wang, and Xiaolin Hu. 2021. Robust logo detection in e-commerce images by data augmentation. In Proceedings of the 29th ACM International Conference on Multimedia. 4789–4793.
[9] Qiang Chen, Yingming Wang, Tong Yang, X. Zhang, Jian Cheng, and Jian Sun. 2021. You only look one-level feature. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 13034–13043.
[10] Yu Chen and Vrizlynn L. L. Thing. 2011. A noise-tolerant enhanced classification method for logo detection and brand classification. In Proceedings of the FGIT-SCiTech. 31–42.
[11] ZhiQi Cheng, Yang Liu, Xiao Wu, and XianSheng Hua. 2016. Video ecommerce: Towards online video advertising. In Proceedings of the 24th ACM International Conference on Multimedia. 1365–1374.
[12] C. Constantinopoulos, E. Meinhardt, Y. Liu, and V. Caselles. 2011. A robust pipeline for logo detection. In Proceedings of the IEEE International Conference on Multimedia and Expo. 1–6.
[13] Corinna Cortes and Vladimir Naumovich Vapnik. 1995. Support-vector networks. Mach. Learn. 20 (1995), 273–297.
[14] Mickael Coustaty, Benjamin Duthil, Vincent Courboulay, and Jean Marc Ogier. 2013. Adaptive saliency for logo spotting. In Proceedings of the IAPR International Workshop on Graphics Recognition.
[15] Santosh Kumar Divvala, Derek Hoiem, James Hays, Alexei A. Efros, and Martial Hebert. 2009. An empirical study of context in object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 1271–1278.
[16] Umesh D. Dixit, M. S. Shirdhonkar, and G. R. Sinha. 2022. Automatic logo detection from document image using HOG features. Multimedia Tools Appl. 82 (2022), 863–878.
[17] Christian Eggert, Dan Zecha, Stephan Brehm, and Rainer Lienhart. 2017. Improving small object proposals for company logo detection. In Proceedings of the ACM International Conference on Multimedia Retrieval. 167–174.
[18] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John M. Winn, and Andrew Zisserman. 2009. The pascal visual object classes (VOC) challenge. Int. J. Comput. Vision 88, 2 (2009), 303–338.
[19] István Fehérvári and Srikar Appalaraju. 2019. Scalable logo recognition using proxies. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision. 715–725.
[20] Antonio Javier Gallego, A. Pertusa, and Marisa Bernabeu. 2019. Multi-label logo classification using convolutional neural networks. In Pattern Recognition and Image Analysis. Springer, 485–497.
[21] Yue Gao, Yi Zhen, Haojie Li, and Tat Seng Chua. 2016. Filtering of brand-related microblogs using social-smooth multiview embedding. IEEE Trans. Multimedia 18 (2016), 2115–2126.
[22] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun. 2021. YOLOX: Exceeding YOLO series in 2021. Retrieved from https://arXiv:2107.08430
[23] Ross B. Girshick. 2015. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision. 1440–1448.
[24] S. Gopinathan and Giddaluru Lalitha. 2018. Vehicle logo recognition using enhanced learning: VLR recognition. In Proceedings of the 2nd International Conference on I-SMAC. 707–712.
[26] D. S. Guru and N. Vinay Kumar. 2017. Symbolic representation and classification of logos. In Proceedings of the International Conference on Computer Vision and Image Processing. 535–569.

[27] Hendrick, Chih Min Wang, Ariprihartta, Ciou Guo Jhe, Ping Chong Tsu, and Gwo Jia Jong. 2018. The halal logo classification by using NVIDIA DIGITS. In Proceedings of the International Conference on Applied Information Technology and Innovation. 162–165.

[28] Steven C. H. Hoi, Xiongwei Wu, Hantang Liu, Yue Wu, Huiqiong Wang, Hui Xue, and Qiang Wu. 2015. LOGO-Net: Large-scale deep logo detection and brand recognition with deep region-based convolutional networks. Retrieved from https://arXiv:1511.02462

[29] Qiang Hou, Weiqing Min, Jiaoyan Zheng, and Shuqiang Chen. 2021. FoodLogoDet-1500: A dataset for large-scale food logo detection via multi-scale feature decoupling network. In Proceedings of the 29th ACM International Conference on Multimedia. 4670–4679.

[30] Sujuan Hou, Jianwei Lin, Shangbo Zhou, Maolong Qin, Weikuan Jia, and Yuanjie Zheng. 2017. Deep hierarchical representation from classifying logo-405. Complexity 29, 6 (2017), 1–12.

[31] Changbo Hu, Qun Li, Zhen Zhang, Kenghao Chang, and Ruofei Zhang. 2020. A multimodal fusion framework for brand recognition from product image and context. In Proceedings of the IEEE International Conference on Multimedia & Expo Workshops. 1–4.

[32] ChungLin Huang, HuangChia Shih, and ChuangYuan Chao. 2006. Semantic analysis of soccer video using dynamic bayesian network. IEEE Trans. Multimedia 8 (2006), 749–760.

[33] Forrest N. Iandola, Anting Shen, Peter Gao, and Kurt Keutzer. 2015. DeepLogo: Hitting logo recognition with the deep neural network framework. Retrieved from https://arXiv:1510.02131

[34] J. Redmon and A. Farhadi. 2018. YOLOv3: An incremental improvement. Retrieved from https://arXiv:1804.02767

[35] Rahul Kumar Jain, Taro Watasue, Tomohiro Nakagawa, SATO Takahiro, Yutaro Iwamoto, RuAN Xiang, and Yen Wei Chen. 2021. LogoNet: Layer-aggregated attention centernet for logo detection. In Proceedings of the IEEE International Conference on Consumer Electronics. 1–6.

[36] Xiaojun Jia, Huanqian Yan, Yonglin Wu, Xingxing Wei, Xiaochnan Cao, and Yong Zhang. 2021. An effective and robust detector for logo detection. Retrieved from https://arXiv:2108.00422

[37] Xiaoli Jiang, Kai Sun, Liqun Ma, Zhijian Qu, and Chongguang Ren. 2022. Vehicle logo detection method based on improved YOLOv4. Electronics 11, 20 (2022).

[38] Xuan Jin, Wei Su, Rong Zhang, Yuan He, and Hui Xue. 2020. The open brands dataset: Unified brand detection and recognition at scale. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. 4387–4391.

[39] He Jingmeng, Xie Yuxiang, Luan Xiaodao, Niu Xiaodao, and Zhang Xiaodao. 2016. A TV logo detection and recognition method based on SURF feature and bag-of-words model. In Proceedings of the 2nd IEEE International Conference on Computer and Communications. 370–374.

[40] Alexs Joly and Olivier Buisson. 2009. Logo retrieval with a contrario visual query expansion. In Proceedings of the 17th ACM International Conference on Multimedia. 581–584.

[41] Yannis Kalantidis, Lluis Garcia Pueyo, Michele Trevisiol, Van Zwol Roelof, and Yannis Avrithis. 2011. Scalable triangulation-based logo recognition. In Proceedings of the 1st ACM International Conference on Multimedia Retrieval. 1–7.

[42] Mohammad Karimi and Alireza Behrad. 2019. Logo recognition by combining deep convolutional models in a parallel structure. In Proceedings of the 4th International Conference on Pattern Recognition and Image Analysis. 216–221.

[43] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 6 (2012), 84–90.

[44] G. Kumar. P. Keserwani, P. P. Roy, and D. P. Dogra. 2021. Logo detection using weakly supervised saliency map. Multimedia Tools Appl. 80 (2021), 4341–4365.

[45] Andrey Kuznetsov and Andrey V. Savchenko. 2020. A new sport teams logo dataset for detection tasks. In Proceedings of the International Conference on Computer Vision and Graphics. 87–97.

[46] Viet Phuong Le, Nibal Nayef, Muriel Visani, Jean Marc Ogier, and Cao De Tran. 2014. Document retrieval based on logo spotting using key-point matching. In Proceedings of the 22nd International Conference on Pattern Recognition. 3056–3061.

[47] Viet Phuong Le, Nibal Nayef, Muriel Visani, Jean Marc Ogier, and Cao De Tran. 2016. Time-efficient logo spotting using text/non-text separation as preprocessing and approximate nearest neighbor search. In Proceedings of the Conférence en Recherche d’Informations et Applications French Information Retrieval Conference and Colloque International Francophone sur l’Ecrit et le Document (CORIA-CIFED’16).

[48] Viet Phuong Le, Muriel Visani, Cao De Tran, and Jean Marc Ogier. 2012. Logo spotting for document categorization. In Proceedings of the 21st International Conference on Pattern Recognition. 3484–3487.
Deep Learning for Logo Detection: A Survey

[49] Viet Phuong Le, Muriel Visani, De Cao Tran, and Jean Marc Ogier. 2013. Improving logo spotting and matching for document categorization by a post-filter based on homography. In Proceedings of the 12th International Conference on Document Analysis and Recognition. 270–274.

[50] Baiying Lei, Vrizlynn L. L. Thing, Yu Chen, and Wee Yong Lim. 2012. Logo classification with edge-based daisy descriptor. In Proceedings of the IEEE International Symposium on Multimedia. 222–228.

[51] Fuxing Leng. 2021. A gradient balancing approach for robust logo detection. In Proceedings of the 29th ACM International Conference on Multimedia. 4765–4769.

[52] Chenge Li, István Fehér-vári, Xiaonan Zhao, Ives Macédo, and Srikar Appalaraju. 2022. SeeTek: Very large-scale open-set logo recognition with text-aware metric learning. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 587–596.

[53] Chuyin Li, Lu Li, Hongliang Jiang, Kaiheng Weng, Yifei Geng, L. Li, Zaidan Ke, Qingyuan Li, Meng Cheng, Weiqiang Nie, Yiduo Li, Bo Zhang, Yufei Liang, Linyuan Zhou, Xiaoming Xu, Xiangxiang Chu, Xiaoming Wei, and Xiaolin Wei. 2022. YOLOv6: A single-stage object detection framework for industrial applications. Retrieved from https://arXiv:2209.02976.

[54] Xingzhuo Li, Sujuan Hou, Baisong Zhang, Jing Wang, Weikuan Jia, and Yuanjie Zheng. 2023. Long-range dependence involutional network for logo detection. Entropy 25, 1 (2023).

[55] Yuanyuan Li, Qiuyue Shi, Jianguan Deng, and Fei Su. 2017. Graphic logo detection with deep region-based convolutional networks. In IEEE Visual Communications and Image Processing. Springer, 1–4.

[56] Zhe Li, Schulte Austum Matthias, and Martin Neschen. 2010. Fast logo detection and recognition in document images. In Proceedings of the 20th International Conference on Pattern Recognition. 2716–2719.

[57] Yuan Liao, Xiaoming Hou, Baisong Zhang, Jing Wang, Weikuan Jia, Yuanjie Zheng, and Yuanjie Zhang. 2023. Long-range dependence pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 936–944.

[58] Jiajun Liu, Fei Shen, Mengwan Wei, Yuzhao Zhang, H. Zeng, Jianqiang Zhu, and Canhui Cai. 2019. A large-scale benchmark for vehicle logo recognition. In Proceedings of the 4th International Conference on Image, Vision, and Computing. 479–483.

[60] W. Liu, Dragomir Anguelov, D. Erhan, Christian Szegedy, Scott E. Reed, Cheng Yang Fu, and Alexander C. Berg. 2016. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision. 21–37.

[61] Wanglong Lu, Hanli Zhao, Qi He, Hui Huang, and Xiaogang Jin. 2021. Category-consistent deep network learning for accurate vehicle logo recognition. Neurocomputing 463 (2021), 623–636.

[62] Ye Meng, Sujuan Hou, Jing Wang, Weikuan Jia, Yuanjie Zheng, and Awudu Karim. 2021. An adaptive representation algorithm for multi-scale logo detection. Displays 70 (2021), 102090.

[63] Daniel Mas Montserrat, Qian Lin, Jan P. Allebach, and Edward J. Delp. 2018. Logo detection and recognition with synthetic images. Electron. Imag. 2018, 10 (2018), 337–1–337–7.

[64] Gonçalo Oliveira, Xavier Frazão, André Pimentel, and Bernadete Ribeiro. 2016. Automatic graphic logo detection via fast region-based convolutional networks. In Proceedings of the International Joint Conference on Neural Networks. 985–991.

[65] Oscar Orti, Rubén Tous, Mauro Gomez, Jonatan Poveda, L. Cruz, and Otto Wüst. 2019. Real-time logo detection in brand-related social media images. In Advances in Computational Intelligence. Springer, 125–136.

[66] Nedere Ozay and Bülent Sankur. 2009. Automatic TV logo detection and classification in broadcast videos. In Proceedings of the 17th European Signal Processing Conference. 839–843.

[67] Karel Paleček and Josef Chaloupka. 2021. Logo detection and identification in system for audio-visual broadcast makeover in the context of content piracy. Global Trans. Proc. 2, 2 (2021), 421–428.

[68] Chun Pan, Zhiyuan Yan, Xiaoming Xu, Mingxia Sun, Jie Shao, and Di Wu. 2013. Vehicle logo recognition based on deep learning architecture in video surveillance for intelligent traffic system. In Proceedings of the IET International Conference on Smart and Sustainable City. 123–126.

[69] Da Pan, Ping Shi, Zhe Qiu, Yuan Sha, Xiu Zhongdi, and Jing Zhoushao. 2016. TV logo classification based on convolutional neural network. In Proceedings of the IEEE International Conference on Information and Automation. 1793–1796.

[70] Kiran Kumar Jakkur Patalappa and Supriya Maganahalli Chandramouli. 2021. Robust recognition of logo under makeovers in the context of content piracy. Global Trans. Proc. 2, 2 (2021), 421–428.

[71] The Anh Pham, Mathieu Delalandre, and Sabine Barrant. 2011. A contour-based method for logo detection. In Proceedings of the International Conference on Document Analysis and Recognition. 718–722.

[72] Pichitchai Pimkote and Thanapat Kangkachit. 2018. Classification of alcohol brand logos using convolutional neural networks. In Proceedings of the International Conference on Digital Arts, Media and Technology. 135–138.
[73] Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779–788.

[74] Joseph Redmon and Ali Farhadi. 2017. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 6571–6575.

[75] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. 39, 6 (2015), 1137–1149.

[76] Jerome Revaud, Matthijs Douze, and Cordelia Schmid. 2012. Correlation-based burstiness for logo retrieval. In Proceedings of the 20th ACM International Conference on Multimedia. 965–968.

[77] Stefan Romberg, Lluís Garcia Pueyo, Rainer Lienhart, and Van Zwol Roelof. 2011. Scalable logo recognition in real-world images. In Proceedings of the 1st ACM International Conference on Multimedia Retrieval. 1–8.

[78] Hichem Sahbi, Lamberto Ballan, Giuseppe Serra, and A. Bimbo. 2013. Context-dependent logo matching and recognition. IEEE Trans. Image Process. 22, 3 (2013), 1018–1031.

[79] Salma Sahel, Mashael Alsahafi, Manal Alghamdi, and Tahani Alsubait. 2021. Logo detection using deep learning with pretrained CNN models. Eng., Technol. Appl. Sci. Res. 11, 1 (2021), 6724–6729.

[80] Mukhil Azhagan Mallaiyan Sathiaseelan, Olivia P. Paradis, Rajat Rai, Suryaprakash Vasudev Pandurangi, Manoj Yasaswi Vutukuru, Shayan Taheri, and Navid Asadizanjani. 2021. Logo classification and data augmentation techniques for PCB assurance and counterfeit detection. In Proceedings of the 47th International Symposium for Testing and Failure Analysis. 12–19.

[81] Sittampalam Sotheswaran and Amirthalingam Ramanan. 2014. A classifier-free codebook-based image classification of vehicle logos. In Proceedings of the 9th International Conference on Industrial and Information Systems. 1–6.

[82] Hang Su, Shaogang Gong, and Xiatian Zhu. 2017. Weblogo-2M: Scalable logo detection by deep learning from the web. In Proceedings of the IEEE International Conference on Computer Vision Workshops. 270–279.

[83] Hang Su, Shaogang Gong, and Xiatian Zhu. 2021. Multi-perspective cross-class domain adaptation for open logo detection. Comput. Vision Image Understand. 204 (2021), 103156.

[84] Hang Su and Guoping Qiu. 2022. Few-shot transfer active learning for logo detection in sports video. SSRN Electronic J. (Apr. 2022).

[85] Hang Su, Xiatian Zhu, and Shaogang Gong. 2017. Deep learning logo detection with data expansion by synthesising context. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision. 530–539.

[86] Hang Su, Xiatian Zhu, and Shaogang Gong. 2018. Open logo detection challenge. In Proceedings of the IEEE International Conference on Computer Vision. 111–119.

[87] G. Naga Sujini. 2022. Automatic logo detection with deep region-based convolutional networks. Int. J. Res. Appl. Sci. Eng. Technol. 10, 6 (2022), 3267–3274.

[88] Snehal Surwase and Meenakshi M. Pawar. 2021. Multi-scale multi-stream deep network for car logo recognition. Evolutionary Intelligence 16 (2021), 485–492.

[89] Kang Tong, Ka Wai Cheung, and Xiaochuan Yu. 2022. ICME 2022 few-shot logo detection top 9 solution. Retrieved from https://arXiv:2206.11462

[90] Andras Túzkö, Christian Herrmann, Daniel Manger, and Jürgen Beyerer. 2018. Open set logo detection and retrieval. In Proceedings of the International Conference on Computer Vision Theory and Applications. 284–292.

[91] Jasper R. R. Uijlings, Koen E. A. van de Sande, Theo Gevers, and Arnold W. M. Smeulders. 2013. Selective search for object recognition. Int. J. Comput. Vision 104 (2013), 154–171.

[92] Diego A. Velazquez, Josep M. Gonfaus, Pau Rodríguez, F. Xavier Roca, Seiichi Ozawa, and Jordi González. 2021. Logo detection with no priors. IEEE Access 9 (2021), 106998–107011.

[93] Chengde Wan, Zhicheng Zhao, Xin Guo, and Zhiyi Cai. 2013. Tree-based Shape Descriptor for scalable logo detection. In Visual Communications and Image Processing. Springer. 1–6.

[94] Chien Yao Wang, Alexey Bochkovskiy, and X. Liao. 2022. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Retrieved from https://arXiv:2207.02696

[95] Chien Yao Wang, I. Hau Yeh, and Hongpeng Liao. 2021. You only learn one representation: unified network for multiple tasks. Retrieved from https://arXiv:2105.04206

[96] Jing Wang, Weiqing Min, Sujuan Hou, Shengnan Ma, Yuanjie Zheng, and Shuqiang Jiang. 2022. LogoDet-3K: A large-scale image dataset for logo detection. ACM Trans. Multimedia Comput., Commun. Appl. 18, 1 (2022), 1–19.

[97] Jing Wang, Weiqing Min, Sujuan Hou, Shengnan Ma, Yuanjie Zheng, Haishuai Wang, and Shuqiang Jiang. 2020. Logo2K+: A large-scale logo dataset for scalable logo classification. In Proceedings of the AAAI Conference on Artificial Intelligence. 6194–6201.

[98] Christian Wilms and Simone Frintrop. 2018. AttentionMask: Attentive, efficient object proposal generation focusing on small objects. In Proceedings of the Asian Conference on Computer Vision. 678–694.

[99] Christian Wilms, Rafael Heid, Mohammad Araf Sadeghi, Andreas Ribbrock, and Simone Frintrop. 2021. Which airline is this? Airline logo detection in real-world weather conditions. In Proceedings of the 25th International Conference on Pattern Recognition. 4996–5003.
Zhiqiang Xiang, Yuexian Zou, Xiaoqun Zhou, and Xiaolin Huang. 2016. Robust vehicle logo recognition based on locally collaborative representation with principal components. In Proceedings of the 6th International Conference on Information Science and Technology. 487–491.

Jiayu Xu, Dongming Zhang, Guoqing Jin, Xiuguo Bao, Qingsheng Yuan, and Yongdong Zhang. 2018. PNET: Pixel-wise TV logo recognition network. J. Comput.-aided Design Comput. Graph. 30 (2018), 1878–1889.

Weipeng Xu, Ye Liu, and Daquan Lin. 2021. A simple and effective baseline for robust logo detection. In Proceedings of the 29th ACM International Conference on Multimedia. 4784–4788.

Shuo Yang, Chunjuan Bo, Junxung Zhang, and Meng Wang. 2020. Vehicle logo detection based on modified YOLOv2. In Proceedings of the 2nd EAI International Conference on Robotic Sensor Networks. 75–86.

Shuo Yang, Junxung Zhang, Chunjuan Bo, Meng Wang, and Lijun Chen. 2019. Fast vehicle logo detection in complex scenes. Optics Laser Technol. 110 (2019), 196–201.

Zhixiong Yang, Huaizhang Liao, Haoyu Zhang, Weijie Li, and Jingyuan Xia. 2022. Representation based few-shot learning for brand-logo detection. In Proceedings of the IEEE 4th International Conference on Power, Intelligent Computing and Systems. 350–354.

Q. Ye, Z. Luo, X. Xiao, and S. Ge. 2017. GeLoGo: Detecting TV logos from web-scale videos. In Proceedings of the IEEE 3rd International Conference on Multimedia Big Data. 250–251.

Kangning Yin, Shaoqi Hou, Ye Li, Chao Li, and Guangqiang Yin. 2020. A real-time vehicle logo detection method based on improved YOLOv2. In Proceedings of the International Conference on Wireless Algorithms, Systems, and Applications. 666–677.

Ervin Yohannes, Chih Yang Lin, Timothy K. Shih, Chen Ya Hong, Avirmed Enkhbat, and Fitri Utaminingrum. 2021. Domain adaptation deep attention network for automatic logo detection and recognition in google street view. IEEE Access 9 (2021), 102623–102635.

W. Yousaf, A. Umar, S. H. Shirazi, Z. Khan, and M. Zaka. 2021. Patch-CNN: Deep learning for logo detection and brand recognition. J. Intell. Fuzzy Syst. 40, 2 (2021), 1–14.

Yongtao Yu, Haiyan Guan, Dilong Li, and Changhui Yu. 2021. A cascaded deep convolutional network for vehicle logo recognition from frontal and rear images of vehicles. IEEE Trans. Intell. Transport. Syst. 22, 2 (2021), 758–771.

YeonJu Yu, Daizong Shi, Wenjia Zhu, Zhen Liang, Hua Li, Wei Jia, and Qiang Lu. 2022. VLDetNet: An improved vehicle logo detection method based on vblock. J. Electron. Imag. 31, 6 (2022), 063042.

Bailing Zhang and Hao Pan. 2013. Reliable classification of vehicle logos by an improved local-mean based classifier. In Proceedings of the 6th International Congress on Image and Signal Processing. 176–180.

Fan Zhang, Lin Cao, and Dong Ming Zhang. 2017. TV logo dataset and annotations for the convolution neural network. In Proceedings of the 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics. 1–6.

Guangpeng Zhang, Dongming Zhang, Jing Zhang, Chuanning Wang, Lidong Wang, and Xueqiang Zou. 2022. TV logo detection and recognition based on data synthesis and metric learning. J. Softw. 9 (2022), 3180–3194.

Junxung Zhang, Lijun Chen, Chunjuan Bo, and Shuo Yang. 2021. Multi-scale vehicle logo detector. Mobile Netw. Appl. 26, 1 (2021), 67–76.

Junxung Zhang, Shuo Yang, Chunjuan Bo, and Zhiyuan Zhang. 2021. Vehicle logo detection based on deep convolutional networks. Comput. Electric. Eng. 90 (2021), 107004.

Yifei Zhang, Mingming Zhu, Daling Wang, and Shi Feng. 2014. Logo detection and recognition based on classification. In Web-Age Information Management. Springer, 805–816.

Zhiyuan Zhao and Qingjie Liu. 2022. Robust logo detection across large style variations. In Proceedings of the International Conference on Artificial Neural Networks. 624–634.

Guangyu Zhu and David S. Doermann. 2007. Automatic document logo detection. In Proceedings of the 9th International Conference on Document Analysis and Recognition. 864–868.

Received 21 October 2022; revised 6 March 2023; accepted 25 July 2023