

Abstract of Doctoral Dissertation

Title: Robust and Accurate Logo Detection Models with Dual-Attention and Adversarial Domain Adaptation

Doctoral Program in Advanced Information Science and Engineering
Graduate School of Information Science and Engineering
Ritsumeikan University

ジャイン ラフル クマール
JAIN Rahul Kumar

Logo is the most common and important entity for every brand in the modern era. A logo is a unique symbol which identifies product and services of any company. It is a mean of identification of a company and distinguishes a brand from others. Automatic detection of logos in real-world images is a crucial problem. It has significant importance in several real-world applications such as brand promotion, brand visibility management, copyright infringement detection, social media monitoring, analyzing advertisements and sponsorships on various platforms. To date, various Logo detection methods have been proposed. Most of them need enormous annotated data for training. In addition, there are domain shifts between the training data (source domain) and test data (target data) resulting in reduction of detection accuracy. In order to overcome these problems, I proposed three novel methods: (1) A weakly supervised logo detection method using a dual-attention architecture; (2) A LogoNet based on dual-attention and domain adaptation for robust and accurate logo detection; (3) An unsupervised logo detection method using adversarial learning from synthetic to real images. The main achievements are summarized as below:

- (1) Most existing logo detection methods need enormous annotated data for training, which are based on strong object level manually annotated bounding boxes. In this study, I proposed a weakly supervised logo detection method using a dual-attention architecture. By using the dual-attention (channel and spatial dimensions), we can focus on important features and areas and automatically find the location of logo in the image. Thus, the proposed method just needs image-level annotations for training and do not need object-level annotations (bounding box), that is called as weakly supervised logo detection. In addition, by using dual-attention architecture, we can also improve the detection accuracy for images with complex background.
- (2) To perform precise and efficient Logo detection, I proposed a framework which consist of a robust feature extraction network with spatial and channel attention modules and an anchorfree detection head. We also proposed a lightweight CNNs module architecture for fast detection and practical applications. In most existing methods, there are domain shifts between the training data (source domain) and test data (target data) resulting in reduction of detection accuracy. To address this problem, in this study, I introduced domain adaptation-based approaches to train the detection frameworks to align networks between datasets from different domains. In the proposed method (called as LogoNet), the unannotated target domain data are also used for training together with the annotated source domain data. By using the LogoNet, we realized a robust and accurate logo detection.
- (3) Further, I proposed a novel domain adaptation method using mid-level feature maps entropy minimization to improve the accuracy of domain adaptation. By using the proposed domain adaptation method, I realized an unsupervised logo detection method from synthetic to real images. We use labeled synthetic images (1-shot logo icon images) and only unlabeled real images for training. The domain shifts between the synthetic images and real images are significantly aligned using mid-level feature maps entropy minimization for adversarial learning. The proposed framework addresses the domain gap problem and is also appropriate for practical applications.

博士論文要旨

論文題名：デュアルアテンションとドメイン適応を利用した 頑健で高精度なロゴ検出法

立命館大学大学院情報理工学研究科
情報理工学専攻博士課程後期課程

ジャイン ラフル クマール
JAIN Rahul Kumar

ロゴは企業や製品のシンボルマークである。ロゴの自動検出は、ブランディング戦略、ブランド認知度の管理、著作権の侵害防止、及びソーシャルメディアでの製品ブランドのモニタリングなどにおいて重要な役割を果たしている。これまで、様々なロゴ検出法が提案されているが、いずれも膨大な学習データが必要である。また、学習に用いたソースデータと検出対象となるターゲットデータ間にドメインシフトが存在するため、検出精度が低下する。本研究では、上記の問題点を克服するために、デュアルアテンションを用いた弱教師学習型ロゴ検出法を開発した。また、ドメイン適応法を開発し、ドメインシフトに対応できる頑健なロゴ検出法 (LogoNet) を開発した。さらに、人工的に生成したロゴ画像を用いた教師なし学習型ロゴ検出法を開発した。主な研究成果を以下に示す。

- (1) 深層学習を用いた高精度なロゴ検出法が提案されているが、膨大な学習データが必要である。本研究では、チャンネル方向と空間方向のデュアルアテンションを用いた弱ロゴ検出ネットワークを提案した。注目すべき場所を自動的に強調させることによってイメージレベルでのアノテーションだけでロゴ検出（弱教師型ロゴ検出）ができるようになった。また、複雑な背景を持つ画像におけるロゴ検出精度も向上した。
- (2) これまで提案された手法において、学習に用いたソースデータと検出対象となるターゲットデータ間にドメインシフトが存在するため、検出精度が低下するという問題点があった。本研究では、デュアルアテンションとドメイン適応を用いた LogoNet を提案した。アノテーションされているソースデータに加え、アノテーションされていないターゲットドメインデータも学習に加えるドメイン適応法は、異なるドメインデータに対しても頑健なロゴ検出ができた。
- (3) 中間層特徴マップを利用したドメイン適応法を新たに提案し、頑健で高精度なロゴ検出を実現した。また、提案したドメイン適応法を利用し、人工的に生成したロゴ画像を用いた教師なし学習型ロゴ検出法を開発した。提案法は、生成ロゴ画像のみを学習に用いる。アノテーションされた実画像は不要である。生成画像と実画像間のドメインシフトはドメイン適応法によってアライメントされる。