Gradual Domain Adaptation in Scene Text Recognition

Lê Nhật Kha

University of Information Technology, VNU-HCM
HCMC, Vietnam

What ?

We introduce a framework address domain gap problem in Scene Text Recognition:

- Proposed a novel method to improve traditional Unsupervised Domain Adaptation

- Boosted performance of STR models

Why ?

- Scene Text Recognition (STR) encounters significant challenges when the training data (synthetic) and testing data (real-world) belong to two different distribution domains, known as the domain gap.
- Domain adaptation techniques applied to the STR problem often prove ineffective when faced with a substantial domain gap.

Overview
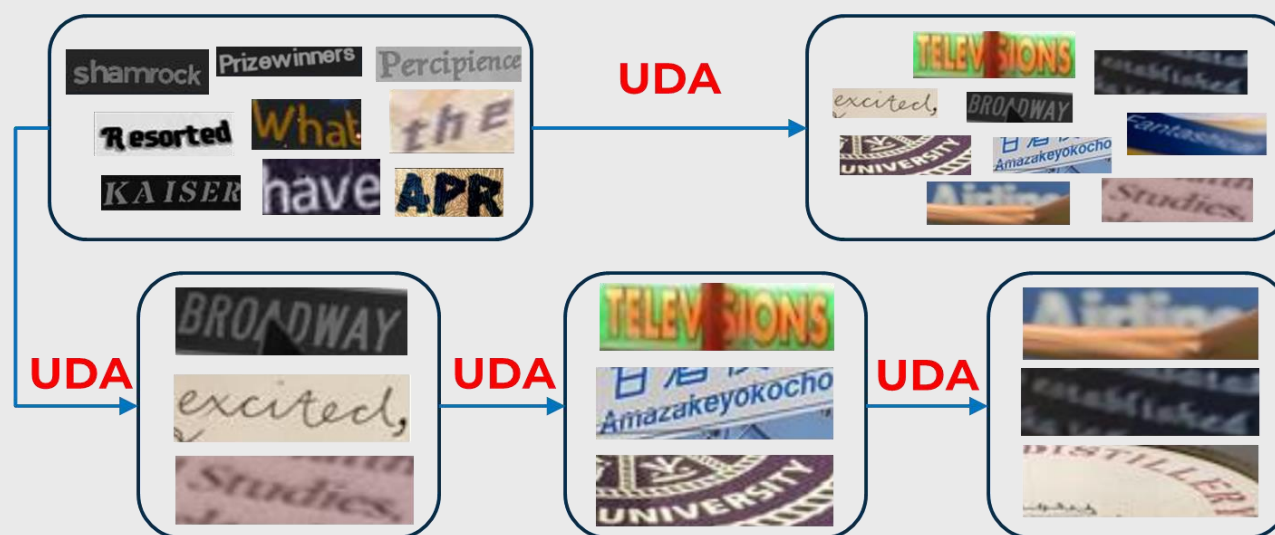
Description

1. Introduction:

- Natural Scene Text Recognition (STR) is a crucial problem in the field of Computer Vision.
- State-of-the-art models (SOTA) prioritize using synthetic data (automatically generated data) for model training.
- Baek and colleagues [4] have pointed out a substantial difference between real-world and synthetic data, referred to as the domain gap.
- To address this issue, some authors have proposed methods using Unsupervised Domain Adaptation (UDA) techniques, aiming to enable models trained on the source domain (synthetic data) to perform well on the target domain (real-world data).
- However, the limitation of this method is that as the domain gap between the source and target domains increases, the effectiveness of UDA diminishes.
- To address this challenge, the team decided to explore the Gradual Domain Adaptation (GDA) method, with the main idea of breaking down the domain gap and implementing progressive self-training.

2. Research content:

- Gradual domain adaptation (GDA) is a method that combines multiple steps of Unsupervised Domain Adaptation (UDA).
- Instead of performing self-training only once for the entire target domain data, GDA divides the data into sub-domains, where the domain gap of each sub-domain with respect to the source domain tends to increase gradually.
- With this approach, the model adapts gradually from the source domain to the target domain, progressively improving its performance.
- This method can be seen as a form of progressive self-training, starting from easy to hard instances.

3. Scientific research method:

- With the hypothesis that "breaking down the domain gap will enhance the effectiveness of domain adaptation," the team conducted experiments by subdividing the data for self-training.

- Initially, the data was randomly partitioned, and subsequently, it was subdivided with an ordered approach using proposed sorting methods. The experimental results were compared to validate the hypothesis and the proposed methods.

4. Expectation:

- The proposed method aims to enhance the effectiveness of Unsupervised Domain Adaptation (UDA) techniques.
- It achieves improved performance for baseline STR models without incurring additional labeling costs. Through evaluations on various scene-text benchmark datasets, the proposed method demonstrates superior effectiveness compared to other UDA methods, achieving state-of-the-art domain adaptation results for Scene Text Recognition (STR).
- The synthesis of these results has been presented in a paper, and it has been accepted for publication at major conferences such as CVPR2024.

NII

Le Nhat Kha – University of Information Technology, VNU-HCM
TEL : 0913078252   Email : 20520208@gm.uit.edu.vn