Title: Self-Training with Uncertainty-Aware Style Transfer for Cross-Domain Object Detection

Principal Investigator: [Your Name] **Affiliation:** [Your Institution/Research Group] **Date:** October 15, 2025

Abstract

Modern object detection models achieve remarkable performance but suffer a significant drop in accuracy when deployed in environments (target domains) that differ from their training data (source domain). This problem of **domain shift** is a major obstacle to the real-world application of technologies like autonomous driving, where a vehicle must operate reliably in diverse weather, lighting, and geographic conditions. This proposal outlines a research project to develop a novel framework for unsupervised domain adaptation in object detection. We propose a method that combines generative style transfer with a robust self-training mechanism. Specifically, we will use a Cycle-Consistent Generative Adversarial Network (CycleGAN) to translate images between domains, artificially augmenting the training data. More importantly, we will enhance a self-training pipeline by incorporating **uncertainty estimation**. By using techniques like Monte Carlo Dropout, our model will only leverage pseudo-labels from the target domain in which it has high confidence, preventing the accumulation of errors from incorrect labels. We hypothesize that this uncertainty-aware approach will make the self-training process more stable and effective, leading to a significant improvement in object detection performance in unseen target domains. The proposed research will be evaluated on benchmark datasets like Cityscapes and Foggy Cityscapes, with the goal of creating more robust and reliable perception systems.

Keywords: Research Proposal, Object Detection, Domain Adaptation, Self-Training, Uncertainty Estimation, Style Transfer, Autonomous Vehicles.

Table of Contents

1. Introduction and Problem Statement

1.1. The Success and Brittleness of Modern Detectors

Deep learning-based object detectors like Faster R-CNN and YOLO have become incredibly accurate, forming the perception backbone for many emerging technologies. However, their success is predicated on the assumption that the training and testing data are drawn from the same statistical distribution.

1.2. The Challenge of Domain Shift

In the real world, this assumption is frequently violated. A model trained exclusively on clear, sunny day driving data may fail catastrophically when deployed at night, in the rain, or in a city with different architecture. This phenomenon is known as **domain shift**. Manually annotating data for every possible domain is prohibitively expensive and unscalable. Therefore, **Unsupervised Domain Adaptation (UDA)**, which aims to adapt a model trained on a labeled source domain to an unlabeled target domain, is a critical area of research.

1.3. Research Questions and Objectives

This research aims to answer the following question: *How can we make an object detector robust to domain shift without requiring any labeled data from the new domain?* Our primary objectives are:

1. To design a framework that leverages both image-level style translation and model-level self-training.
2. To develop a novel mechanism to filter noisy pseudo-labels generated during self-training by estimating model uncertainty.
3. To empirically validate the proposed framework and demonstrate its superiority over existing UDA methods.

1.4. Proposed Contribution

The main contribution of this work will be an **uncertainty-aware self-training framework**. While self-training is a common technique, it is often unstable because the model can reinforce its own mistakes by trusting incorrect "pseudo-labels." By introducing a principled mechanism for the model to gauge its own uncertainty, we can select only the most reliable pseudo-labels, leading to more stable and effective adaptation.

2. Literature Review

2.1. State-of-the-Art Object Detection Models

Our work will build upon established object detection architectures. We will consider both a two-stage detector (e.g., Faster R-CNN) and a Transformer-based detector (e.g., DETR) as the base models for our adaptation framework.

2.2. Unsupervised Domain Adaptation (UDA)

UDA is a well-established field in machine learning. The central goal is to leverage a label-rich source domain to learn a task in a label-scarce target domain.

2.3. UDA Techniques in Object Detection

- **Adversarial Training:** These methods use a "domain discriminator" network that tries to distinguish between features from the source and target domains. The main network is then trained to produce features that can "fool" this discriminator, thereby learning domain-invariant features.
- **Style Transfer:** Generative models like GANs are used to translate images from the source style to the target style (e.g., making a sunny image look foggy). This creates a synthetic labeled dataset in the target style.
- **Self-Training:** This involves using an initial model to make predictions on the unlabeled target data. The most confident predictions are then treated as "pseudo-labels" and are used to retrain the model. This approach is powerful but risks error accumulation if the pseudo-labels are noisy. Our proposed work directly addresses this key limitation.

3. Proposed Methodology

3.1. Overall Framework Architecture

The proposed system will consist of three interconnected modules operating on a base object detector. The model will be trained on labeled source data (e.g., sunny images) and unlabeled target data (e.g., rainy images).

3.2. Module 1: Cross-Domain Style Transfer

We will first train a CycleGAN model to learn the mappings between the source and target domains. This will allow us to translate a source image into a "fake" target image (e.g., sunny -> rainy) and vice-versa. This provides a basic form of data augmentation, allowing the detector to see labeled images that look like they are from the target domain.

3.3. Module 2: Self-Training with Pseudo-Labeling

In parallel, we will use the model trained on the source data to generate predictions (bounding boxes and classes) for the unlabeled target domain images. These predictions will serve as initial pseudo-labels.

3.4. The Core Innovation: Uncertainty-Aware Label Filtering

This is the central component of our proposal. Instead of naively trusting all pseudo-labels above a simple confidence threshold, we will estimate the model's **uncertainty** for each prediction. We will use **Monte Carlo Dropout**, a technique where dropout is applied at inference time over multiple forward passes. The variance in the resulting predictions serves as a strong indicator of model uncertainty. We will then filter the pseudo-labels using a combined score of confidence and low uncertainty. Only the most certain and confident pseudo-labels will be added to a replay buffer used to fine-tune the detector, making the adaptation process robust to noise.

4. Experimental Setup and Evaluation

4.1. Datasets and Benchmarks

We will focus on autonomous driving scenarios. The primary experiment will be adapting from the **Cityscapes** dataset (clear weather) to the **Foggy Cityscapes** dataset. We will also evaluate on other common shifts, such as adapting from synthetic data (**Sim10k**) to real-world data (**KITTI**).

4.2. Baseline Models for Comparison

We will compare our method against three baselines:

1. A "Lower Bound" model trained only on source data.
2. A state-of-the-art adversarial training method for UDA.
3. A standard self-training method without uncertainty awareness.

4.3. Evaluation Metrics

Performance will be measured using the standard object detection metric, **mean Average Precision (mAP)**, calculated on the labeled validation set of the target domain.

5. Expected Results and Broader Impact

5.1. Hypothesized Performance Gains

We expect our uncertainty-aware framework to significantly outperform the baselines. We hypothesize that by reducing the noise in the pseudo-labeling process, our model will adapt more effectively, resulting in a 5-10% absolute improvement in mAP on the target domain compared to standard self-training methods.

5.2. Impact on Autonomous Systems and Robotics

A more robust perception system directly translates to increased safety and reliability for autonomous vehicles, drones, and industrial robots. This research could help bridge the gap between development and real-world deployment of these technologies.

6. Plan of Work and Timeline

The project is planned for a 12-month period.

- **Phase 1 (Months 1-2):** In-depth literature review; setting up the computational environment and baseline models.

- **Phase 2 (Months 3-6):** Implementation of the style transfer module and the uncertainty-aware self-training loop.
- **Phase 3 (Months 7-10):** Conducting extensive experiments on benchmark datasets, analyzing results, and performing ablation studies.
- **Phase 4 (Months 11-12):** Writing a research paper for submission to a top-tier computer vision conference (e.g., CVPR, ICCV) and finalizing the project report.

7. Ethical Considerations

The primary application of this research is to enhance safety in autonomous systems. However, object detection technology can also be used for surveillance. Our research will be conducted transparently, and we will focus our evaluation on publicly available datasets related to driving. We will not use private or personally identifiable data. The code and models will be made publicly available to ensure reproducibility and encourage positive use.

8. Conclusion

This research proposal addresses the critical problem of domain shift in object detection. By proposing a novel framework that integrates style transfer with a more robust, uncertainty-aware self-training mechanism, we aim to significantly advance the state of the art in unsupervised domain adaptation. The successful completion of this project will produce more reliable perception models, thereby accelerating the safe and responsible deployment of AI in real-world applications.

9. References

[Hoffman et al., 2018] Hoffman, J., Tzeng, E., Park, T., Zhu, J. Y., Isola, P., Saenko, K., ... & Darrell, T. (2018). CyCADA: Cycle-consistent adversarial domain adaptation. *International conference on machine learning*.

[Saito et al., 2017] Saito, K., Watanabe, K., Ushiku, Y., & Harada, T. (2017). Asymmetric tri-training for unsupervised domain adaptation. *International conference on machine learning*.

[Gal & Ghahramani, 2016] Gal, Y., & Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. *International conference on machine learning*.

[Zou et al., 2018] Zou, Y., Yu, Z., Kumar, B., & Wang, J. (2018). Unsupervised domain adaptation for semantic segmentation via class-balanced self-training. *European conference on computer vision*.