

Advanced topics in signal and image processing – Final project

Research Proposal: A Solution to Domain Shift in Object Recognition Using Diffusion Models.

1. Background This research is based on the paper "**DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation**¹" by Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, Kfir Aberman, Google Research & Boston University from CVPR 2023. With Publicly available code.²

The paper introduces the DreamBooth method, which allows a user to fine-tune a diffusion model to learn a new concept (such as a specific object) from just a few images. This method enables the generation of new and unique images of that object in various contexts and environments while maintaining high fidelity. This application provides an effective solution to problems of data scarcity and domain shift.

2. Motivation The domain shift problem in object recognition is a key challenge, as a model trained in one environment often struggles to recognize objects when they appear against a different background. This is particularly critical in cases where there is a limited number of images of a specific object. The goal of this research is to use DreamBooth as a generative tool to create a rich synthetic dataset that combines the object with a new, specific background. This approach will allow for the training of a stable and robust object recognition model.

3. Research Method The research will proceed with the following steps, adapted for the constraints of limited home computing power:

1. **LoRA Model Training:** Training a small LoRA model using a limited number of images of the new object (e.g., a warship) with the DreamBooth method.
2. **Synthetic Dataset Generation:** Combining the new background images with the trained LoRA model and using the SDEdit technique (implemented in the Stable Diffusion GUI³) to create new images where the object is realistically integrated into the background. Image + Text to Image algorithm.
3. **Automatic Data Annotation:** Create bounding boxes around the objects in the newly generated synthetic images Knowing g.t. bounding box label/Using a pre-trained YOLO model.
4. **Object Detection Model Training and Evaluation:** Training a new object detection model (such as YOLO) on the synthetic dataset. Afterward, comparing its performance to a model trained on real images of the object in the same new background to test the effectiveness of the method.

4. Conclusion This research proposes an innovative and effective method for solving the problems of domain shift and data scarcity by creating high-quality synthetic data using a DreamBooth model.

¹ <https://arxiv.org/pdf/2208.12242>

² <https://github.com/google/dreambooth>

³ <https://github.com/AUTOMATIC111/stable-diffusion-webui>