# AIL862: Assignment 3
# 2024AIB2289

Kashish Srivastava
MTech ScAI

***Abstract*—Deep learning models trained on synthetic data usually experience performance degradation when testing on some real-world data due to domain shifts. This study investigates whether increasing the size of synthetic datasets or incorporating domain adaptation techniques can reduce this gap. We trained a MobileNetV2 classifier on synthetic image dataset and evaluated its performance on some real-world images, comparing the baseline results with models trained using domain adaptation. Results we observed demonstrate that domain adaptation increases classification performance, achieving a test accuracy improvement from 84.55% to 91.62%.**

## I. Introduction

Deep learning models usually struggle when they are trained on synthetic images and evaluated on real-world image data. This performance drop is due to domain shifts, where differences in texture, lighting, and object representation make models to generalize poorly. Addressing this issue is crucial for real-world deployment, where labeled real-world data is often scarce or expensive to obtain.

This study tries to bridge the synthetic to real performance gap by:

- Evaluating how increasing training dataset size of synthetic images affects the model performance.
- Implementing a domain adaptation approach using pseudo-labeling to increase real-world classification accuracy.

We train and validate models on synthetic data and assess their performance on a real-world test set, comparing baseline results to those obtained with domain adaptation.

## II. Methodology

### A. Dataset

We generated synthetic images for three object classes using a text-to-image model. We have used the stable diffusion model for the text-to-image model in this study. The dataset consists of:

- **Synthetic Data:** Used for training, with sample sizes of 500 and 1000 images per class. Classes include 'dog','cat' and 'bird'.
- **Real Data:** Used for testing, containing real-world images corresponding to the three classes. Classes include 'dog','cat' and 'bird'.
- **Unlabeled Real Data:** Used in domain adaptation for pseudo-labeling. Images from real world dataset is used to achieve 90% in-domain classes and 10% of images from other classes (eg. 'beard','koala' etc.)

The training data is divided into 80% for training and 20% for validation.

### B. Model Architecture and Training

We use MobileNetV2, a lightweight convolutional neural network designed for efficient image classification. The model's feature extraction layers are frozen, and only the final classification layer is trained. Training is performed using the Adam optimizer with an initial learning rate of 0.001, and a step-based learning rate scheduler reduces the learning rate by a factor of 0.1 every seven epochs. The loss function is cross-entropy loss.

### C. Domain Adaptation Approach

In order to improve generalization, we implemented domain adaptation through pseudo-labeling:

1) Training a model on synthetic dataset.
2) Using this model to generate pseudo-labels for the unlabeled real-world dataset.
3) Retraining the model using both synthetic data and pseudo-labeled real data.

Pseudo-labels with confidence scores above a threshold (0.51) are selected for training, reinforcing the model's ability to adapt to real-world images.

## III. Results and Discussion

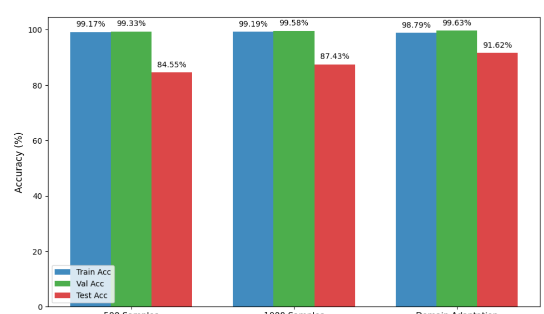


Fig. 1. Comparison of Accuracy Across Different Methods

Figure 1 visualizes the performance comparison across different methods, clearly showing the superiority of the domain adaptation approach in terms of test accuracy.

Increasing the synthetic dataset size from 500 to 1000 images per class improved test accuracy from 84.55% to 87.43%. However, domain adaptation significantly outperformed dataset scaling alone, boosting test accuracy to

TABLE I
PERFORMANCE COMPARISON OF DIFFERENT TRAINING STRATEGIES

| | **Train Acc** | **Val Acc** | **Test Acc** |
|---|---|---|---|
| 500 Samples | 99.17% | 99.33% | 84.55% |
| 1000 Samples | 99.19% | 99.58% | 87.43% |
| Domain Adaptation | 98.79% | 99.63% | 91.62% |

91.62%. The improvement is attributed to the inclusion of real-world pseudo-labeled data, which helps the model better generalize to the target domain.

Additionally, while synthetic-only training resulted in overfitting (with near-perfect training accuracy but lower test accuracy), domain adaptation balanced performance across domains, demonstrating its effectiveness.

## IV. CONCLUSION

Models trained on synthetic image dataset leads to lower performance when transitioning to real-world data. While increasing the dataset size may help improve generalization, domain adaptation through pseudo-labeling provides a much more substantial accuracy increase.