Queensland University of Technology

LA TROBE UNIVERSITY

SENTIENT

SMARTSAT Cooperative Research Centre

Diversity is Definitely Needed: Improving Model-Agnostic Zero-shot Classification via Stable Diffusion

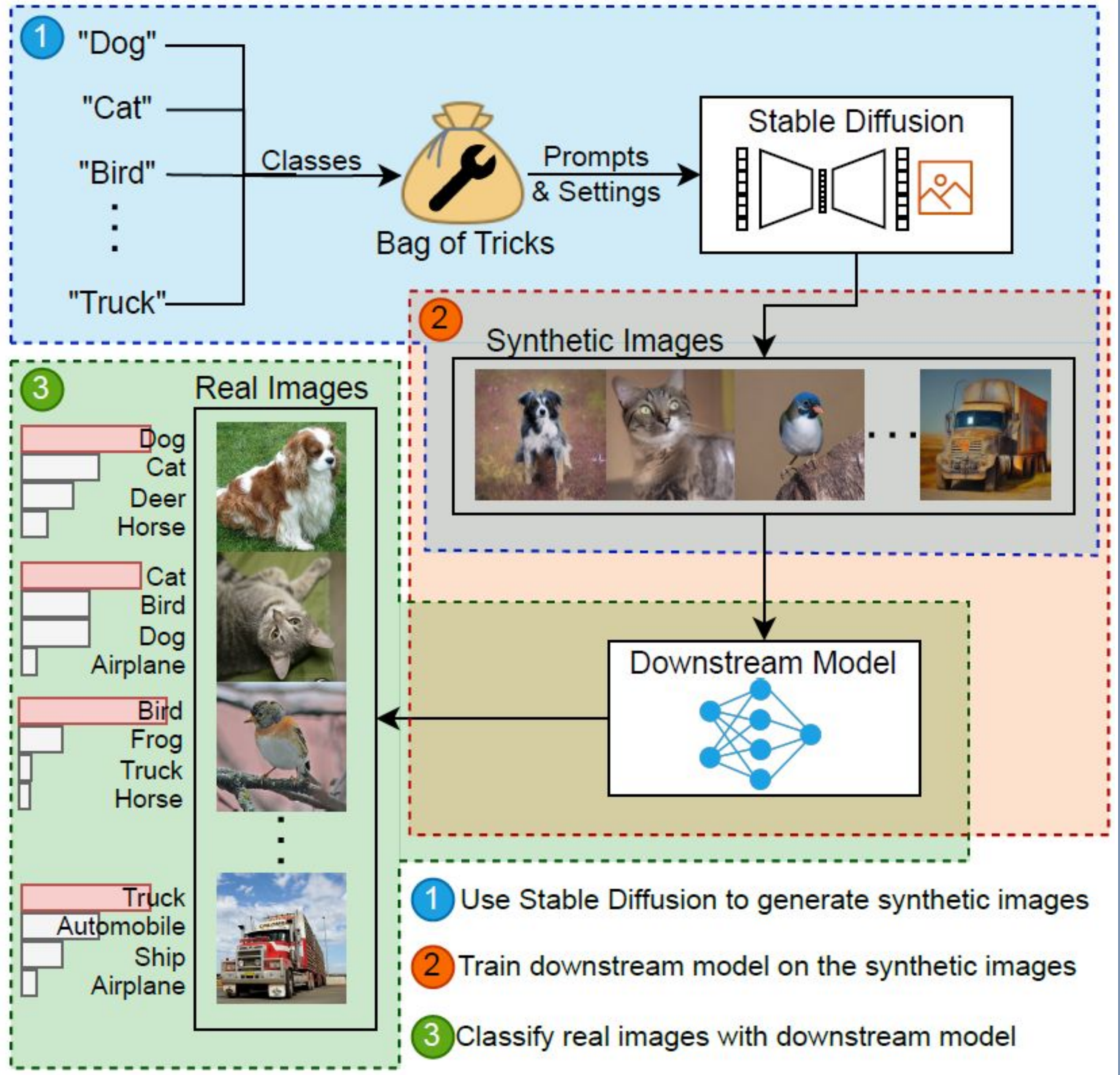
Jordan Shipard, Arnold Wiliem, Kien Nguyen Thanh, Wei Xiang, Clinton Fookes

Paper

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Overview of Method

- How do we solve **M**odel **A**gnostic **Z**ero-**S**hot **C**lassification?
- Use **stable diffusion** to generate **synthetic training data** to train **downstream models**.
- The performance of real data is dependent on the **diversity** of the synthetic images used for training.
- **Enhance** the **diversity** of the synthetic training data with **adjustments** to the **prompts** and generation **settings** used.



GitHub Code

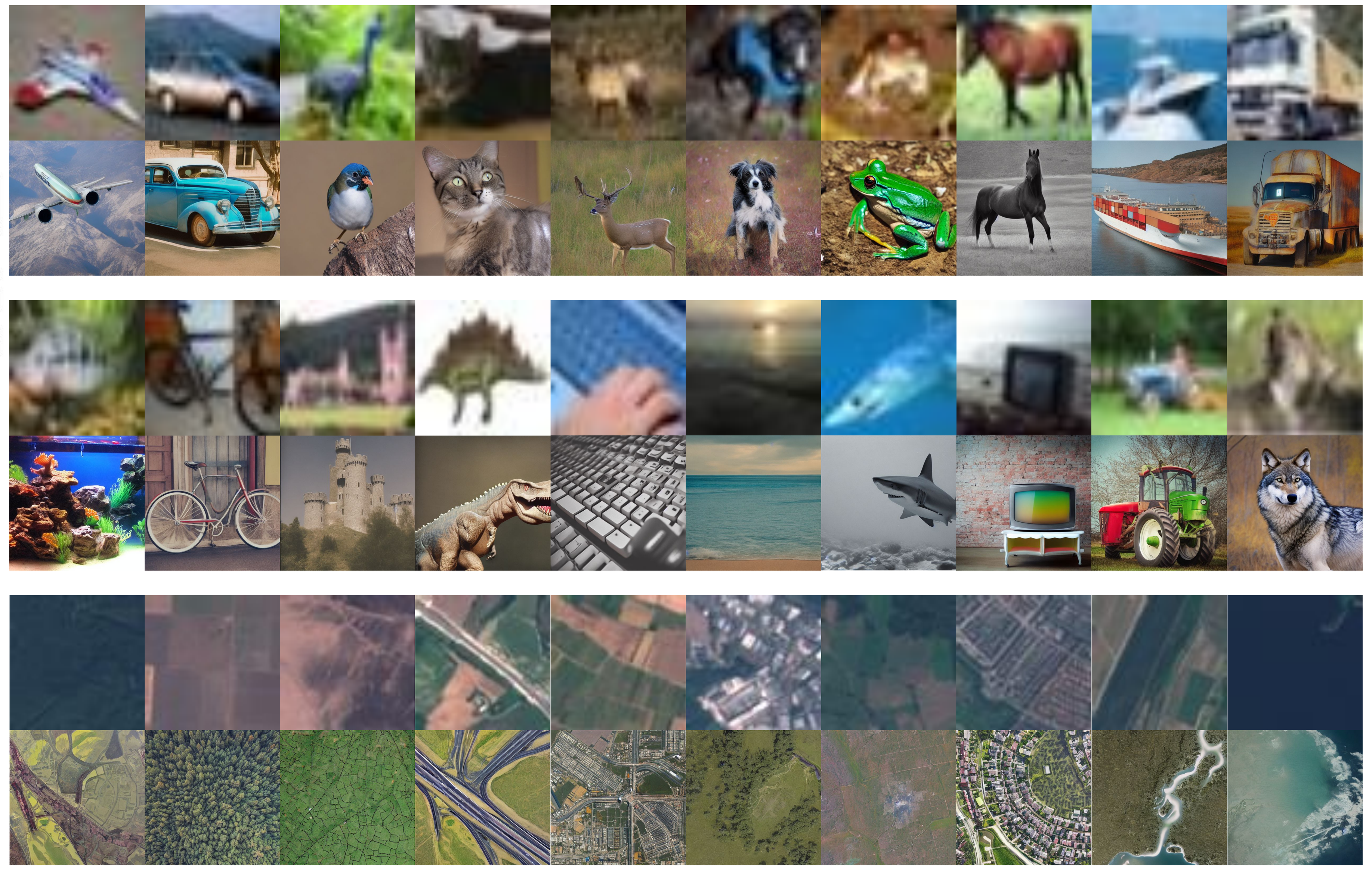
Real (top) vs Synthetic (bottom) Datasets

CIFAR10

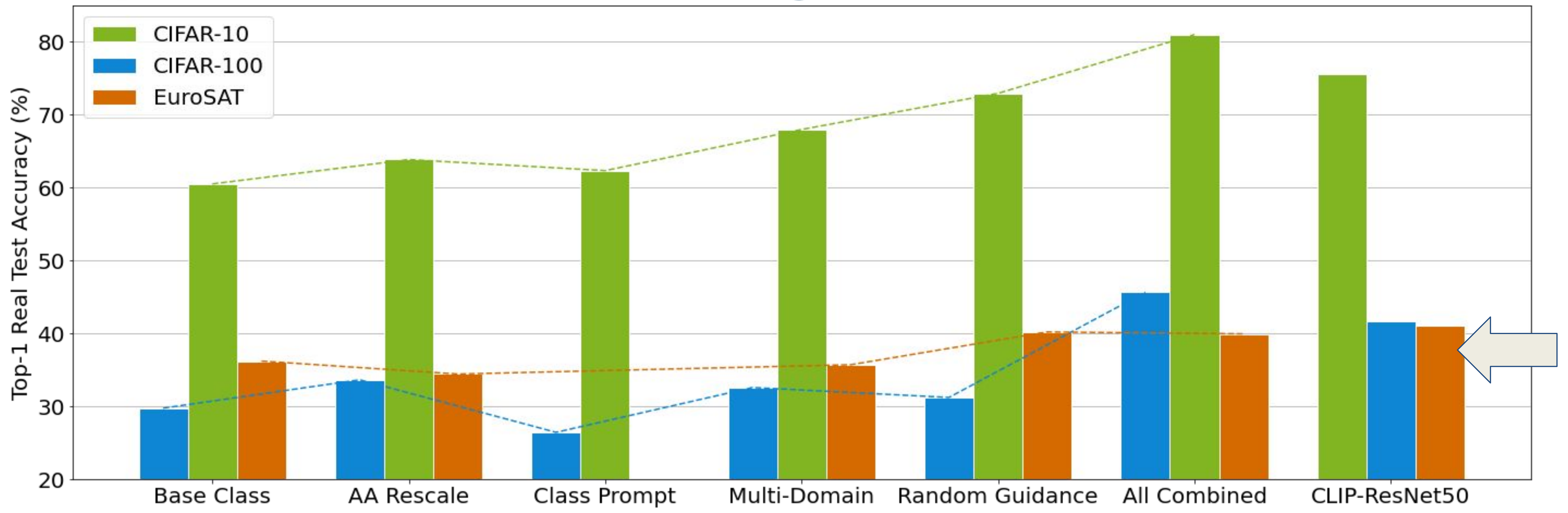
Download

CIFAR100

EuroSAT

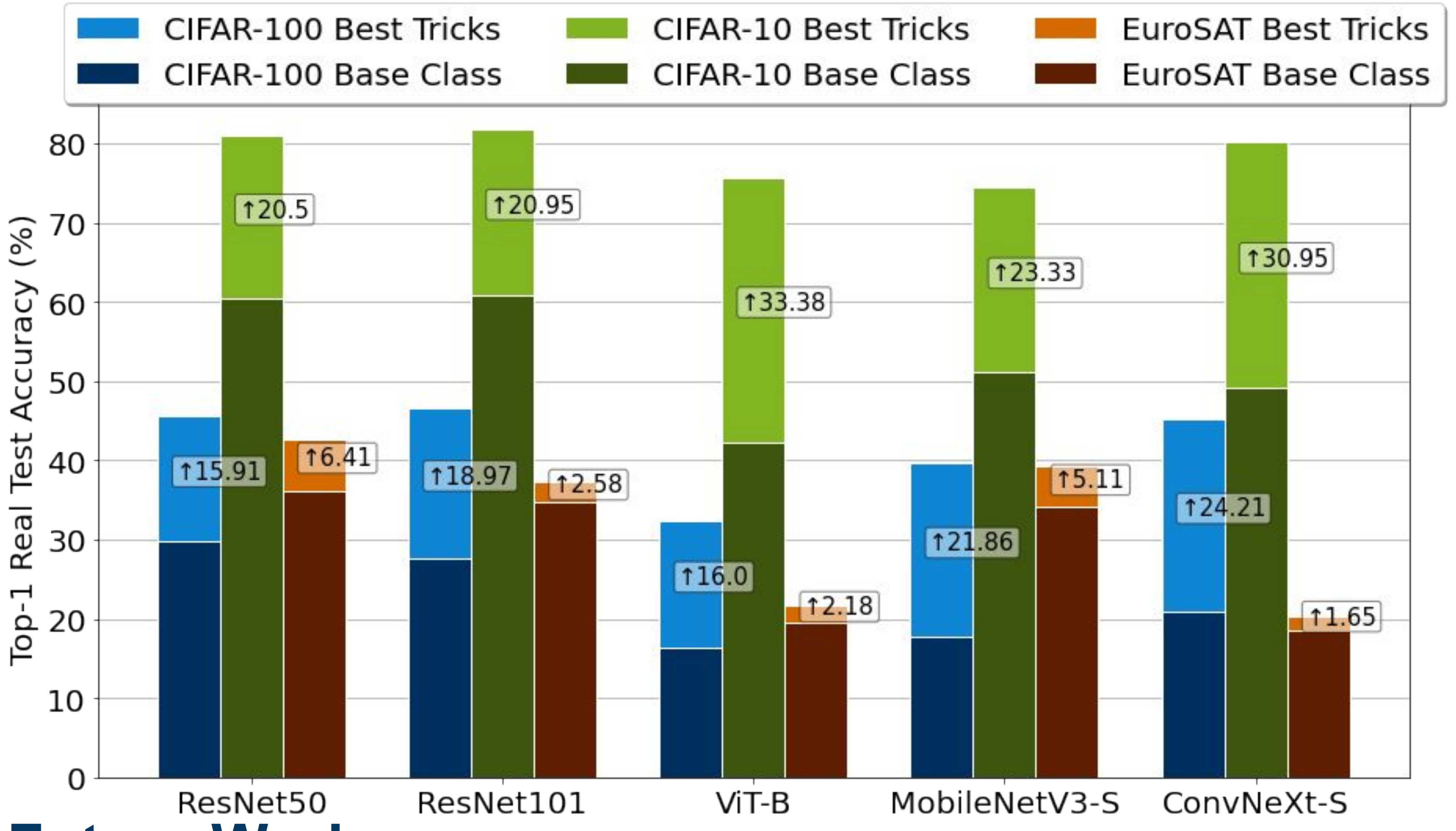


Improvement from Bag of Tricks



CLIP Zero-shot Baseline [2]

Improvement with Different Models



Future Work

- More datasets, such as ImageNet and domain specific datasets, such as additional remote sensing datasets/medical datasets.
- Making the method more efficient through reducing the number of training images needed.
- Expanding to other computer vision tasks, such as semantic segmentation and object detection.

Bag of Tricks

Base Class: "An image of a {class}"
Class Prompt: "{class}"
Multi-Domain: "a {domain} of a {class}"
Random Guidance: "An image of a {class}" + random unconditional guidance [1] value set between 1-5

Base Class

Class Prompt

Multi-Domain

Random Guidance

Questions?

Join our Zoom call during the poster session or email us a question.

Meeting ID: 926 453 9384
Email: jordan.shipard@hdr.qut.edu.au

Zoom link

Acknowledgement

This work has been supported by the SmartSat CRC, whose activities are funded by the Australian Government's CRC Program; and partly supported by Sentient Vision Systems. Sentient Vision Systems is one of the leading Australian developers of computer vision and artificial intelligence software solutions for defence and civilian applications.

Citations

[1] "Classifier-free diffusion guidance", Jonathan Ho and Tim Salimans, NIPS 2021
[2] "Learning Transferable Visual Models From Natural Language Supervision", Radford et al., ICML 2021