

Towards Successful Implementation of Automated Raveling Detection: Training Data Size, Illumination Scale, and Image Shift

Xinan Zhang, Ph.D. Student – xzhang979@gatech.edu; Yichang (James) Tsai, Ph.D., P.E., Professor – james.tsai@ce.gatech.edu

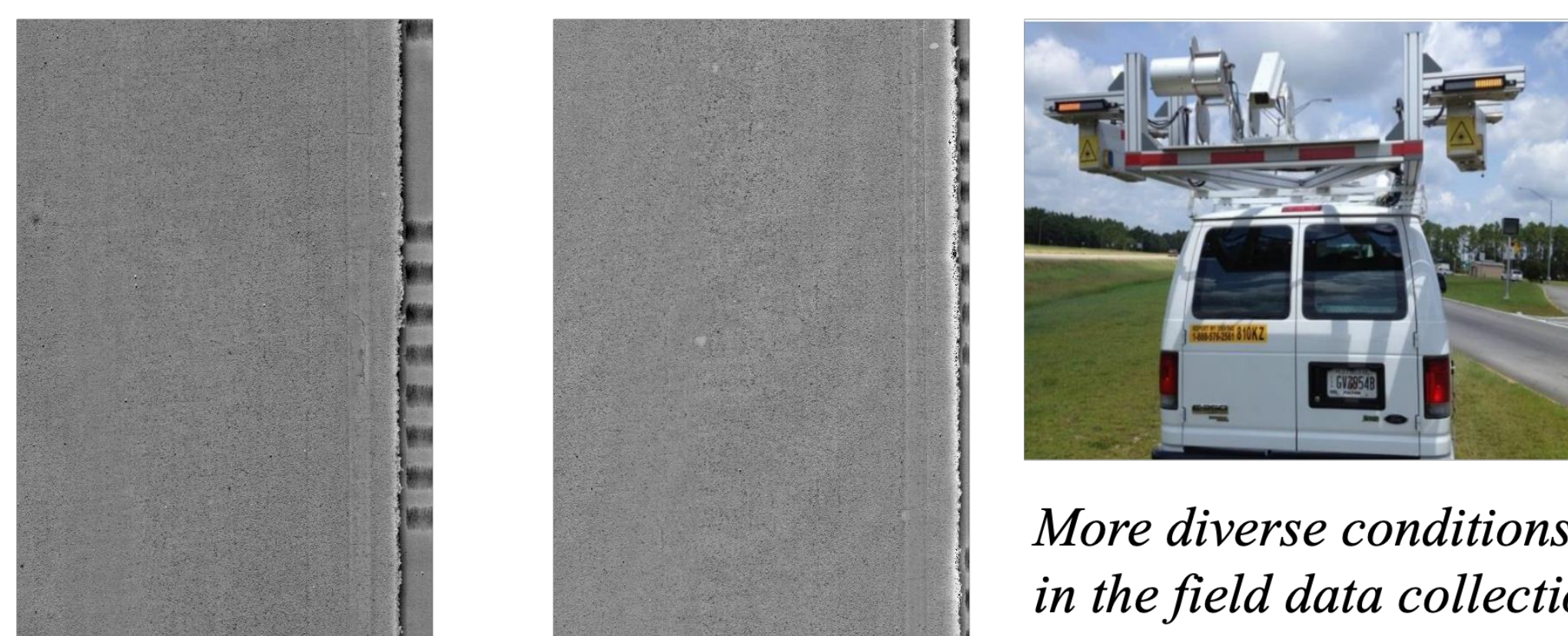
RESEARCH OBJECTIVE

- Raveling, known as loss of aggregates, has been one of major pavement distresses.
- Our objective is to investigate how factors such as **data quantity**, **lighting scales**, and **spatial transformations** influence model robustness.
- We also aims to improve the robustness accordingly.

RESEARCH BACKGROUND



Path Offset Lighting Difference More Conditions

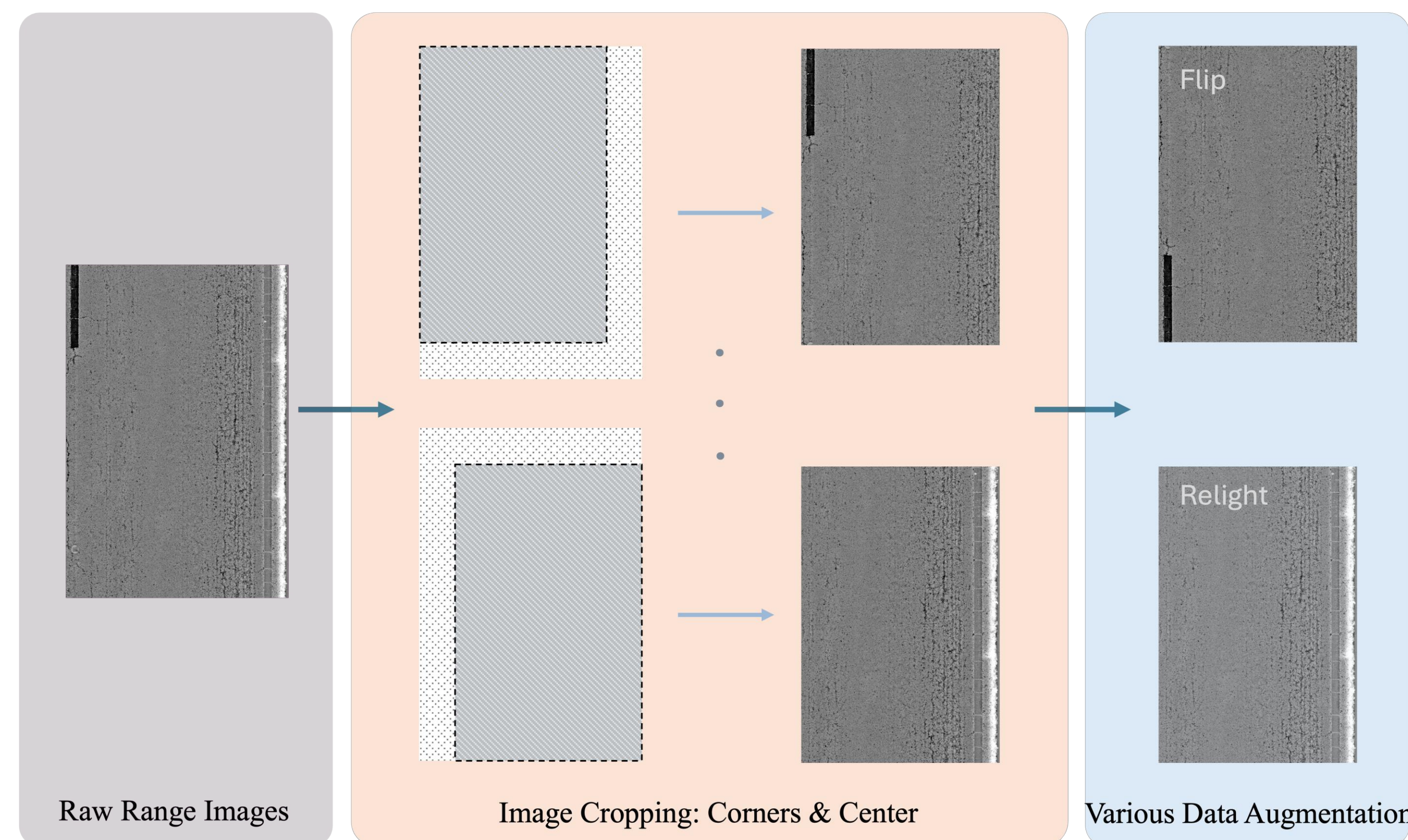


More diverse conditions in the field data collection

- Previous studies provide models that achieve good performance on testing datasets;
- however, in **real-world large-scale implementation**, their performance sometimes drops unexpectedly.
- A reliable evaluation & training framework is needed.

PROPOSED FRAMEWORK

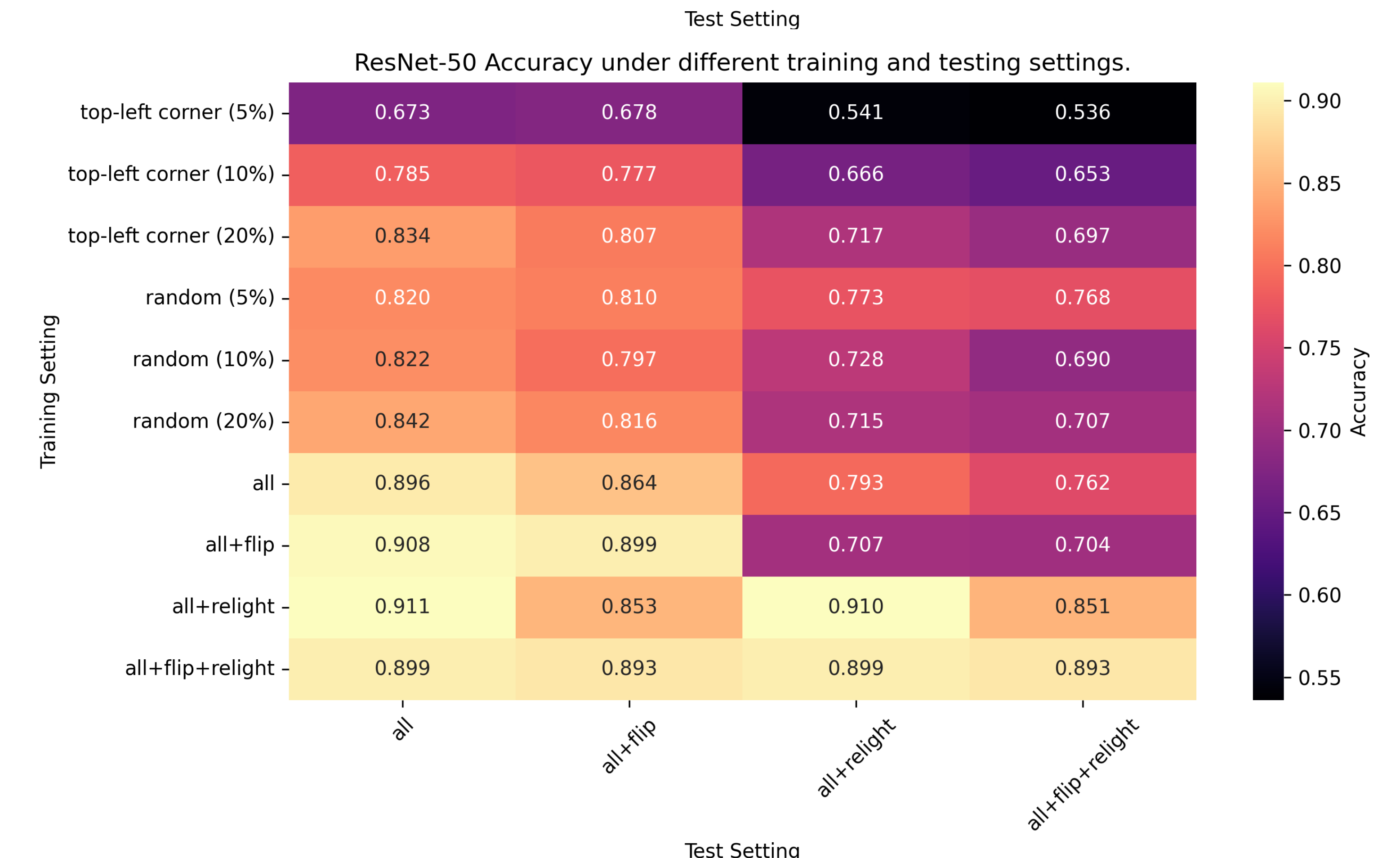
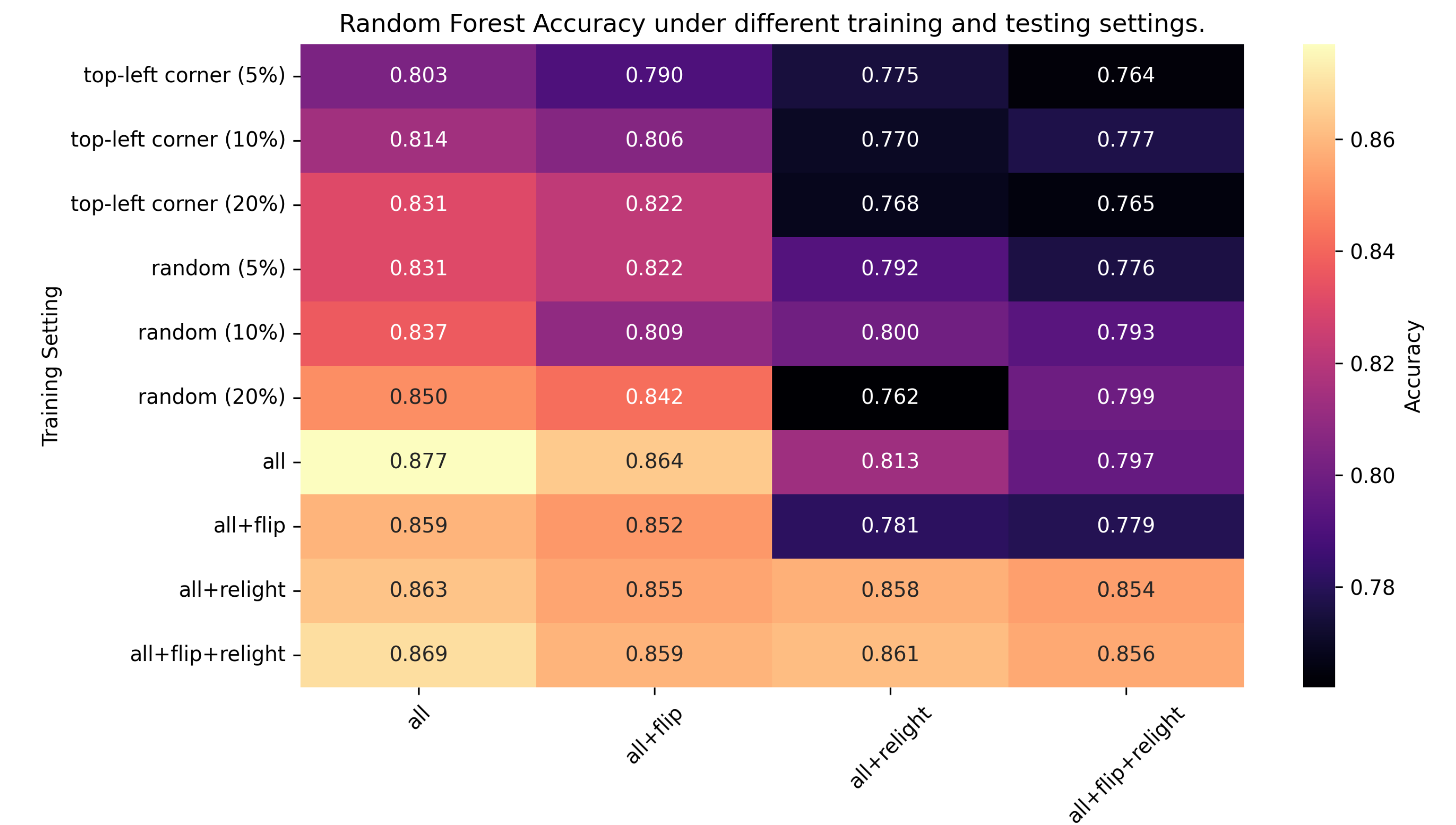
- To accurately evaluate model performance, an annotated dataset that reflects a spectrum of variations is essential. However, this presents several challenges: **Limited variation in existing datasets**, **Cost and labor of data annotation**, **Dynamic nature of variations**.



- The proposed benchmark consists of three stages:
 - Augmented Dataset Generation
 - Controlled Experimental Design
 - Consistent Training and Evaluation Protocol

EXPERIMENTS

- For our experiments, we selected **Random Forest** (ML) and **ResNet50** (DL) as representative models due to their strong performance in prior research.



FINDINGS & SUMMARY

- Extensive experiments demonstrate that **increasing data quantity and diversity** is crucial for model performance.
- By incorporating data augmentation techniques, both ML and DL models become **more robust** to variations conditions.
- **More variations** can be included in this framework to make it more comprehensive.