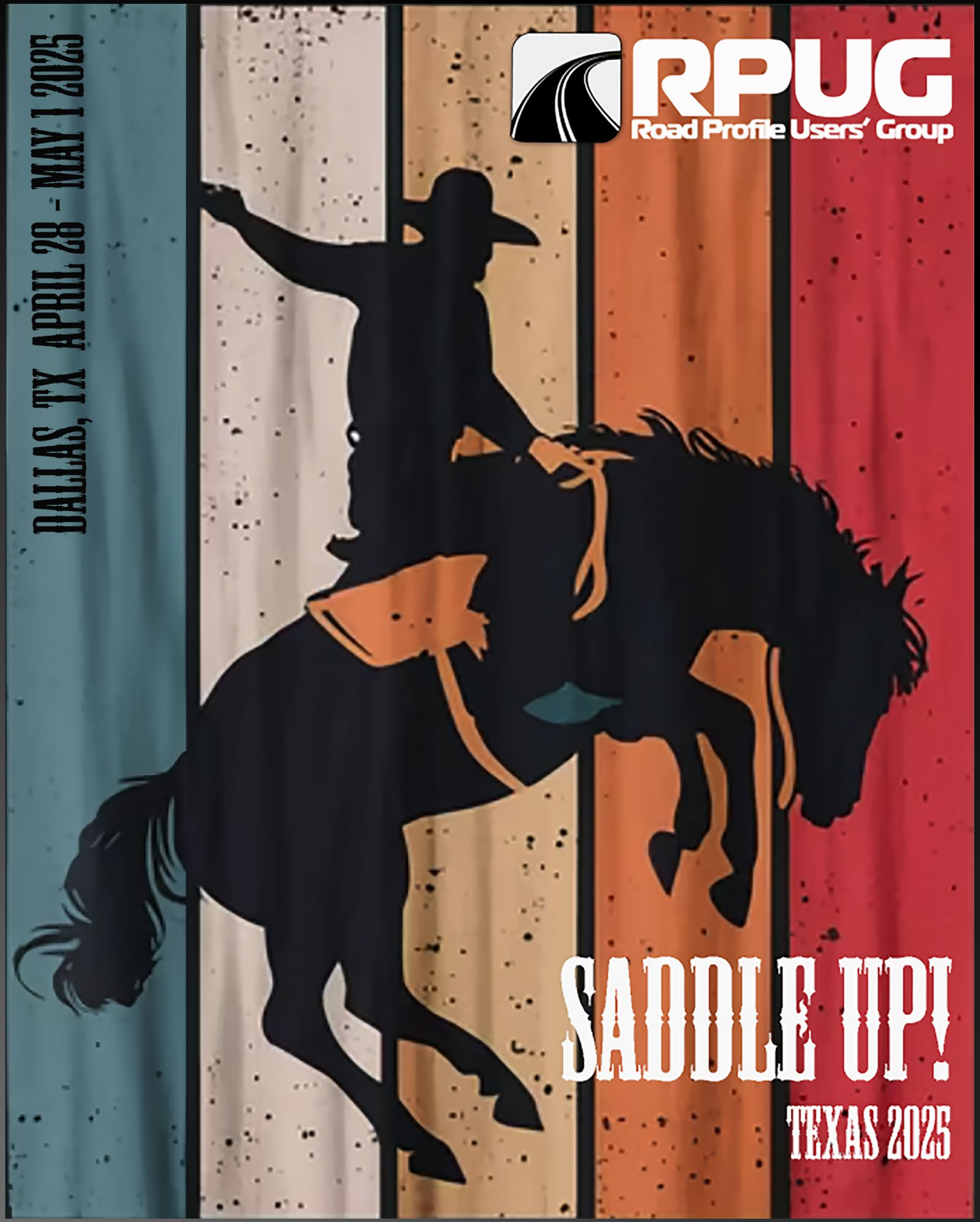


IMPLEMENTING AI/ML IN AUTOMATED PAVEMENT DISTRESS DETECTION AT THE STATE LEVEL -LESSONS LEARNED IN TEXAS

Haitao Gong (presenter), TXST

Feng Wang, TXST

Jelena Tesic, TXST

Andrew Scouten, TXST

4/30/2025

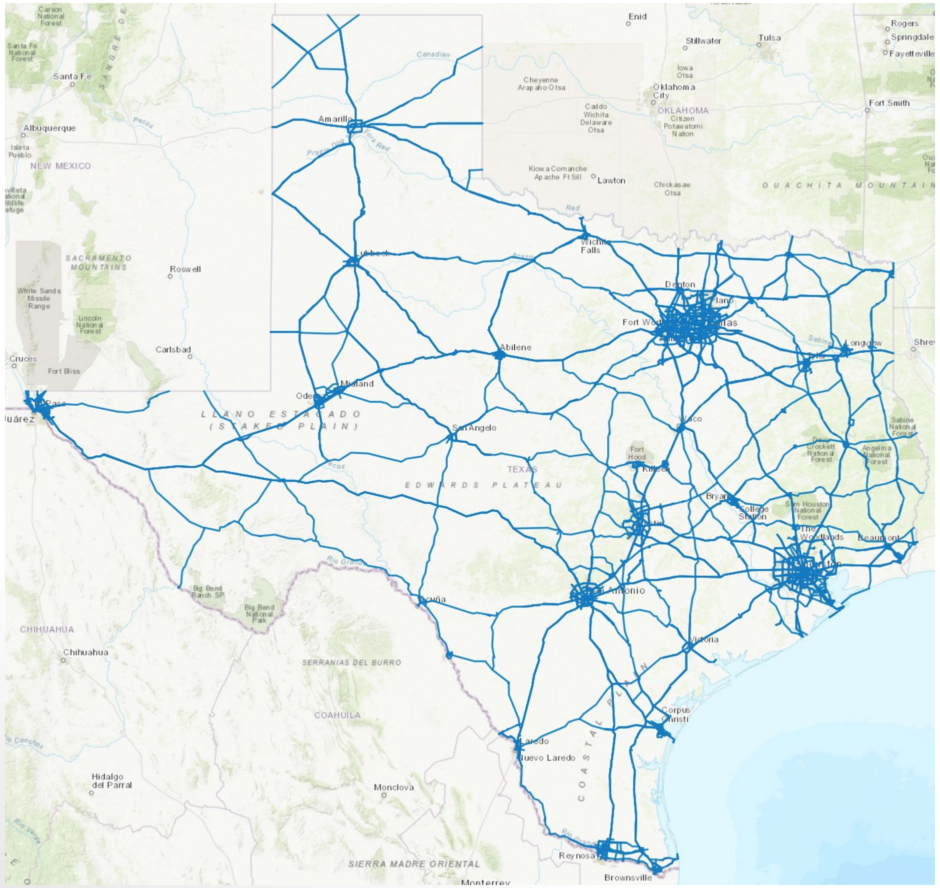
RPUG
Road Profile Users' Group

OUTLINES

- Introduction
- Project Overview
- Lesson 1: know your data
- Lesson 2: know when your model fails
- Future directions

INTRODUCTION

- Distress detection model must operate reliably across an entire state-level highway network
- Errors at scale have real consequences
- Most existing research focuses on localized or Small-scale dataset
- The work addresses a critical gap



INTRODUCTION

- Characteristics varies depending on :
  - Pavement condition
  - Material type
  - Pavement age
  - Traffic pattern
  - Environmental exposure
  - …

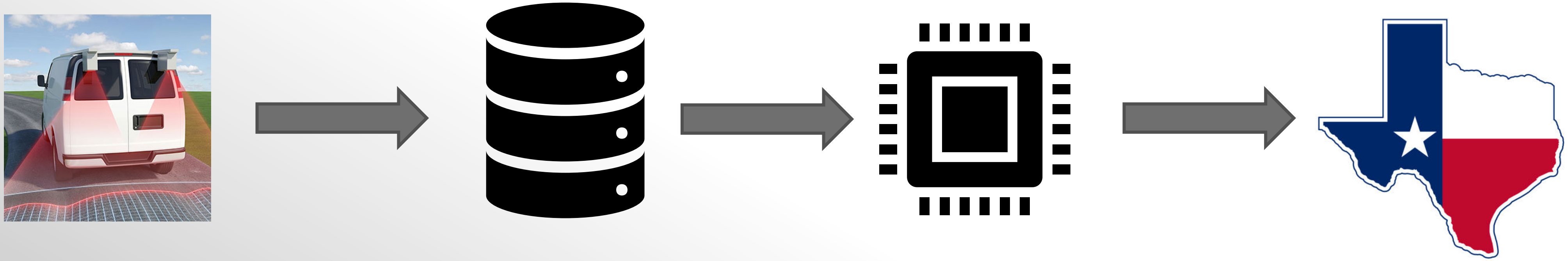
PROJECT OVERVIEW

TxDOT #7150 "*Artificial Intelligence for Pavement Condition Assessment from 2D/3D Surface Images*"

- A real-world Stress Test
  - Texas manages over 200,000 lane miles of highway
  - Multiple pavement types are included: CRCP, JCP, ACP
  - Statewide image data collected by vendor-operated vans
  - a unique opportunity to study model generalization across real-world diversity

PROJECT OVERVIEW

- Technical framework



2D/3D image preparation
09/2022-02/2023

Image library build up
03/2023-12/2023

Model development
01/2024-02/2025

Pilot study
02/2025 - present

PROJECT OVERVIEW

- Primary goal: explore the capabilities and limitations of AI
  - Focused on evaluating feasibility, reliability, and scope.
- Early observations reveal practical challenges
  - Variability in data quality, surface conditions, and distress interpretation affect detection outcomes.
- key focus: where AI succeeds and where it fails
  - Results will guide future integration strategies and model refinement.

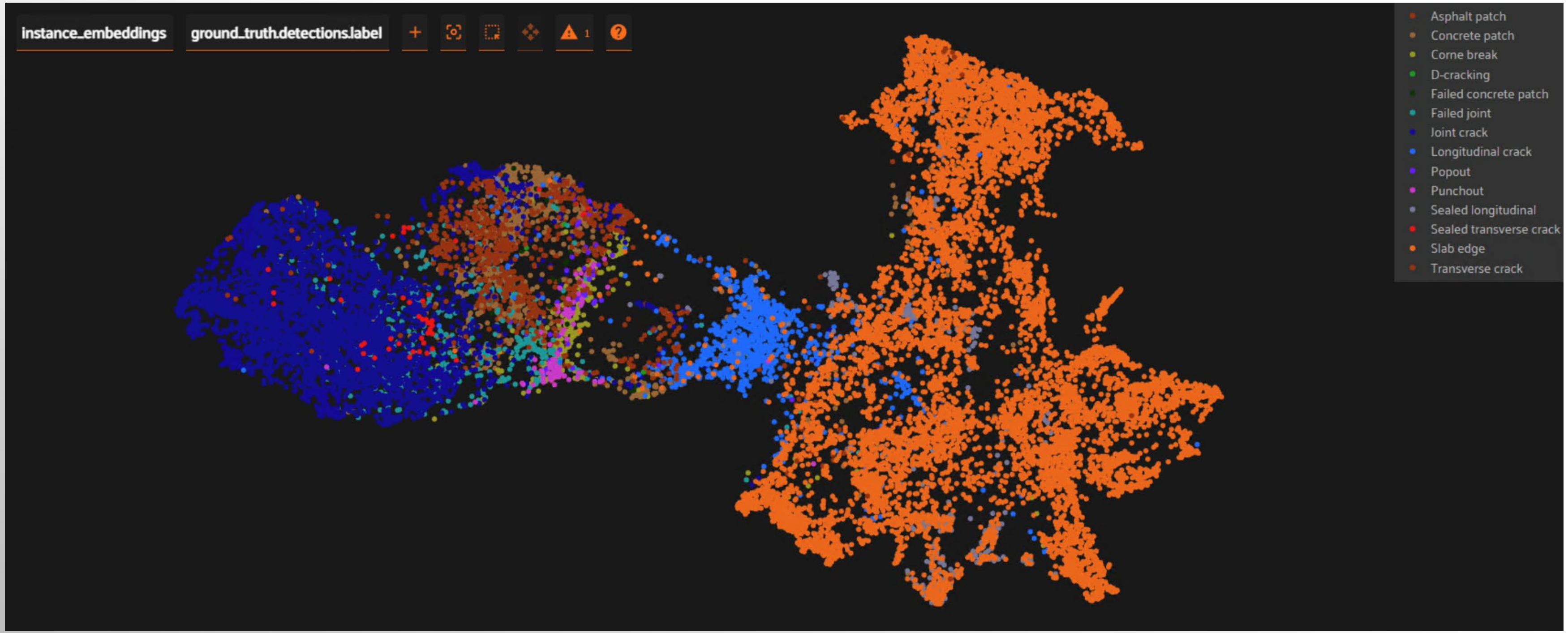
LESSON 1: KNOW YOUR DATA

- What is the image library
  - Intensity & range image
  - Pavement types
  - Annotations

```
Image Library
├── CRCP
│   ├── Image_0001.psi
│   └── Image_0001.txt
├── JCP
│   ├── Image_0001.psi
│   └── Image_0001.txt
└── ACP
    ├── Image_0001.psi
    └── Image_0001.txt
```

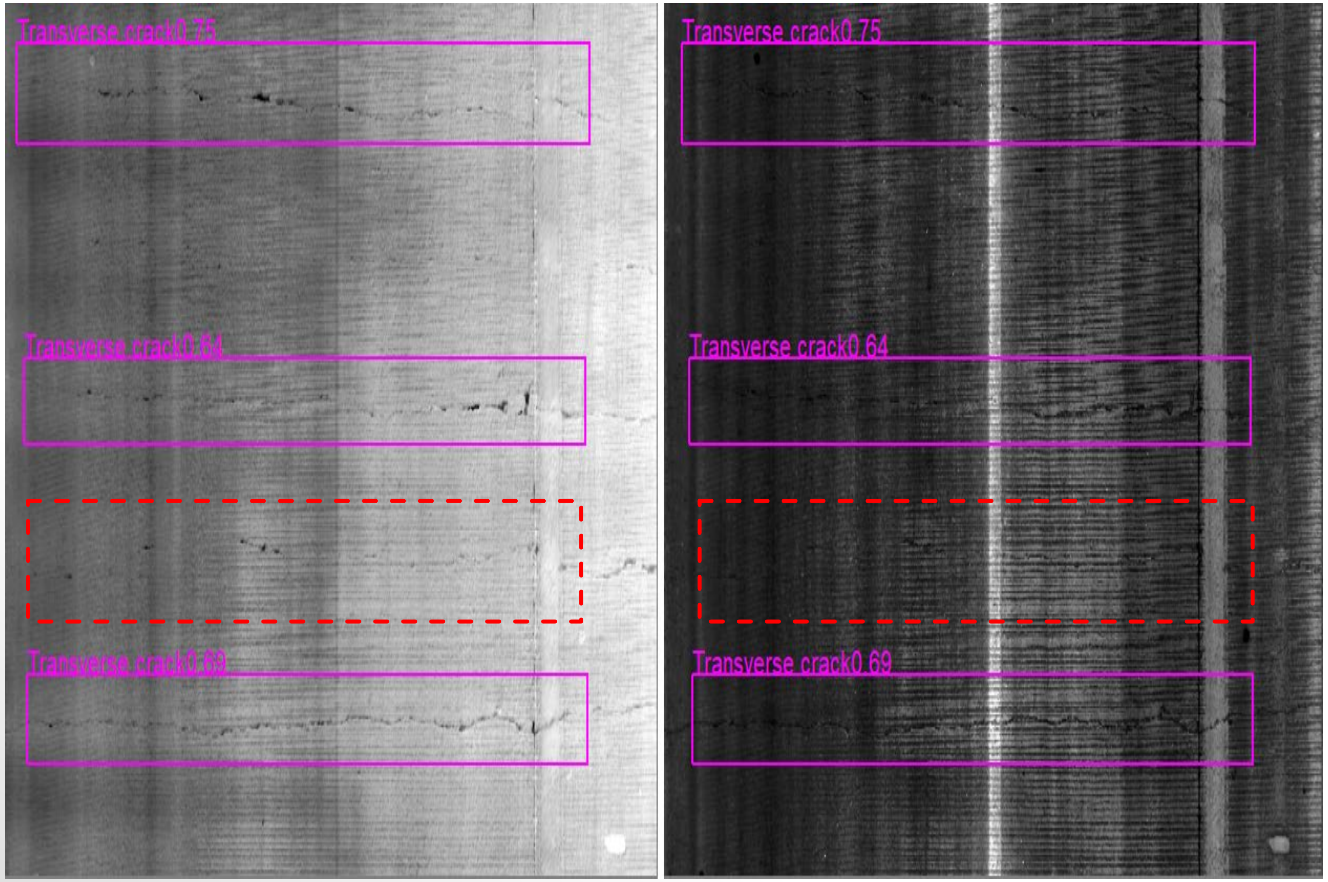
LESSON 1: KNOW YOUR DATA

- Data imbalance:
  - Sample imbalance across distress types
  - Further investigation



LESSON 1: KNOW YOUR DATA

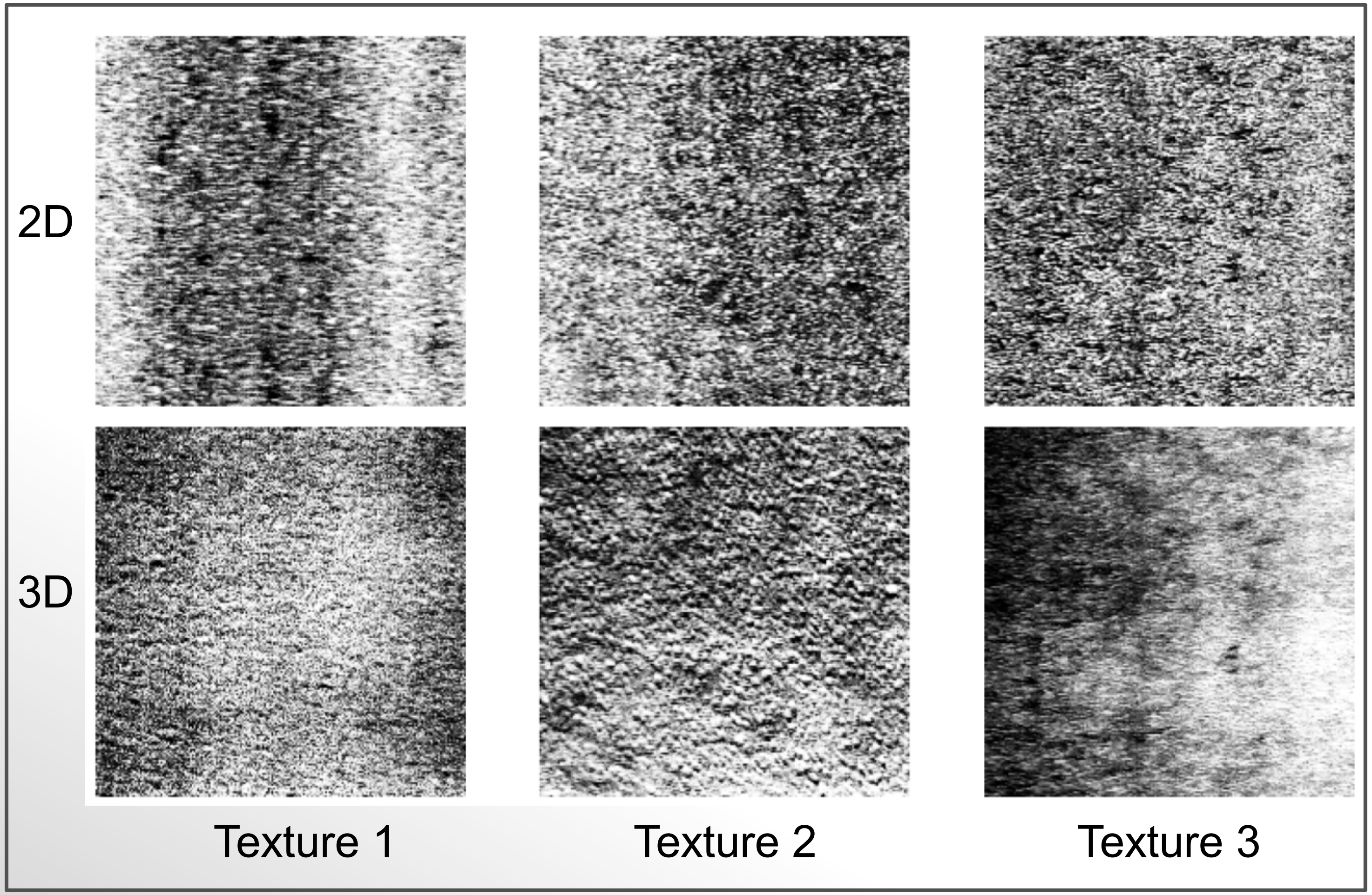
- Annotation consistency:
  - Subjectivity
  - Annotation Standard
- Impact on model training



Crack or not crack?

LESSON 1: KNOW YOUR DATA

- Diversity coverage:
  - Impossible to cover all scenarios
- Impact on model training
  - Uncertainty over new/unseen data



WHAT DEFINES A HIGH-QUALITY DATASET?

- Representative coverage of real-world condition
- Balanced or intentionally curated class distribution
- Consistent and reliable annotation
- Sufficient quantity with verified quality

LESSON 2: KNOW WHEN YOUR MODEL FAILS

- Common model Failures:
  - Misclassification
  - Missed detection
  - Hallucination

LESSON 2: KNOW WHEN YOUR MODEL FAILS

- Misclassification:
  - Distress detected, but assigned wrong class
  - Label inconsistency
  - Domain shift
- How to mitigate:
  - Annotation check and model retrain
  - Complementary data collection

LESSON 2: KNOW WHEN YOUR MODEL FAILS

- Missed Detection:
  - Rare, subtle, or degraded features
  - Model's limited exposure to edge cases
  - Domain shift
- How to mitigate:
  - Complementary data collection

LESSON 2: KNOW WHEN YOUR MODEL FAILS

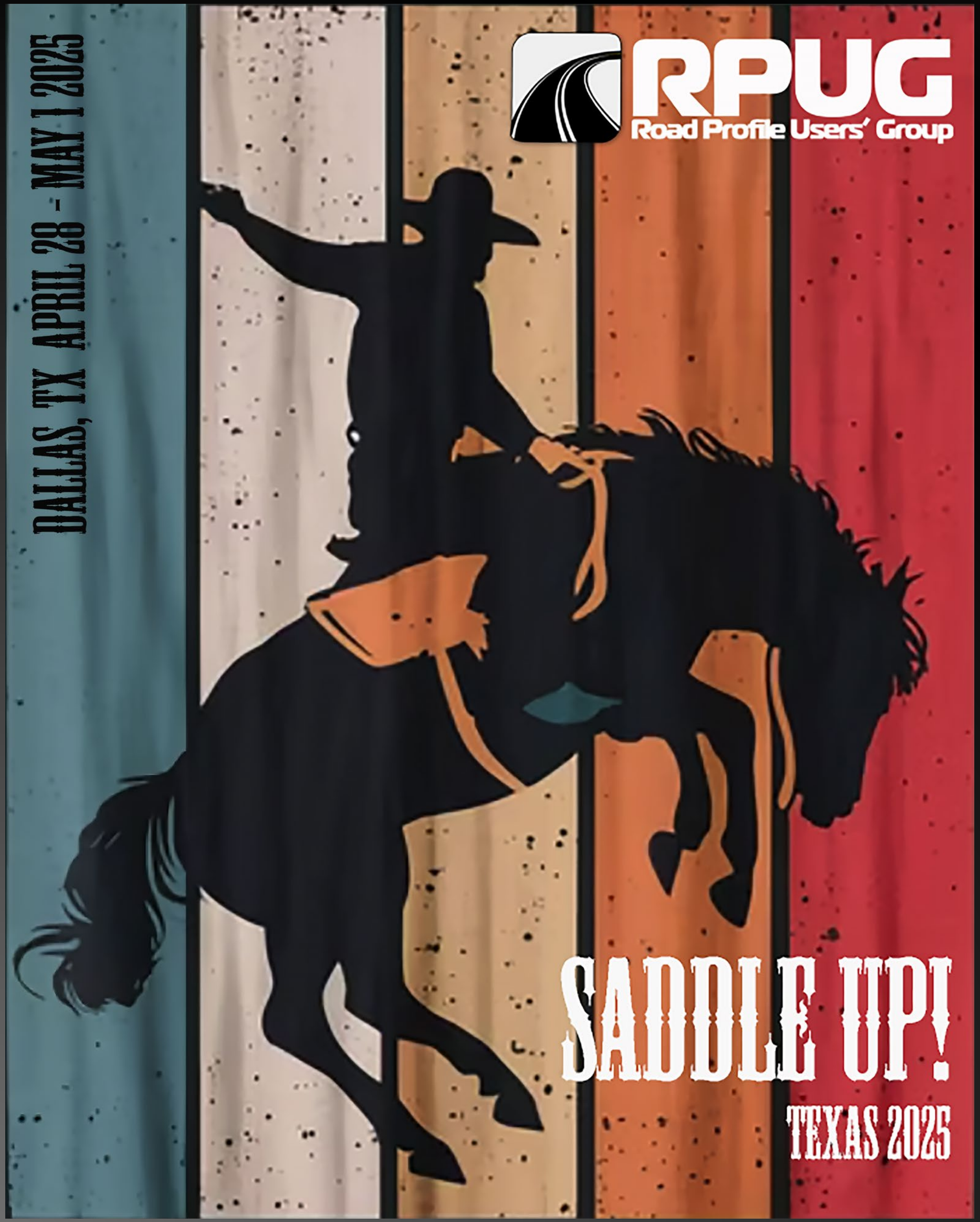
- Hallucination:
  - Mode predicts distress none exists
  - Noise, low-contrast regions
  - Domain shift
- How to mitigate:
  - Negative sample
  - Complementary data collection

FUTURE DIRECTION

- Continuous Improvement Loop
  - Complimentary data collection
  - Field feedback → retraining → better generalization
  - Active learning

ACKNOWLEDGEMENT

- TxDOT
  - For funding, data access, and statewide collaboration
  - Special thanks to TxDOT's pavement division and Andy Mergenmeier
- Pathway Services Inc.
  - For providing high-resolution pavement imaging data



THANK YOU!

RPUG
Road Profile Users' Group