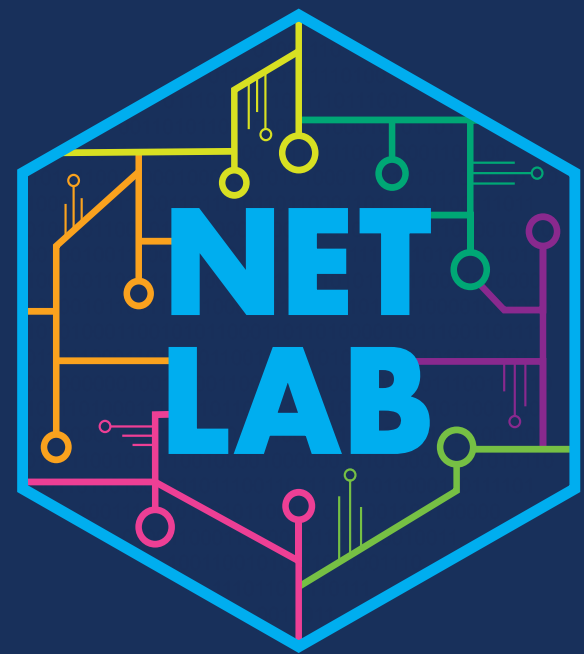


TCAM: Traffic-Camera-Based Air Quality Monitoring Using Multimodal Large Language Models

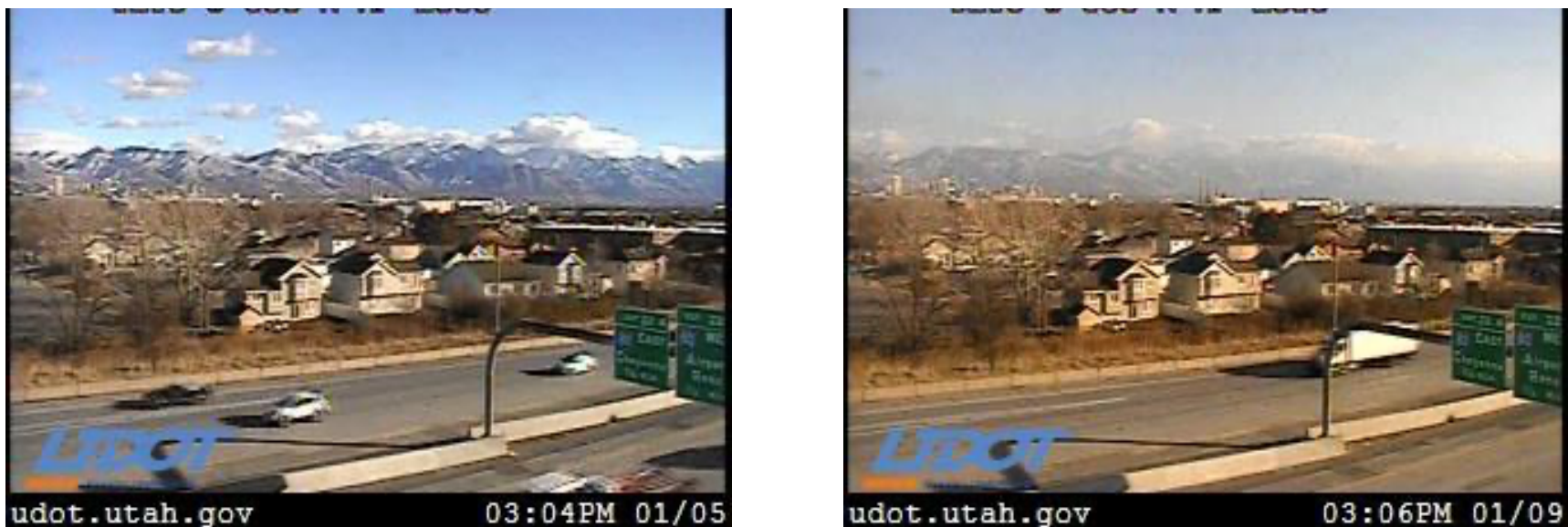
BYU
Electrical & Computer
Engineering

Yujie Rao, Christopher Kitras, Dah-Jye Lee, Philip Lundrigan
Electrical and Computer Engineering Department



Challenge

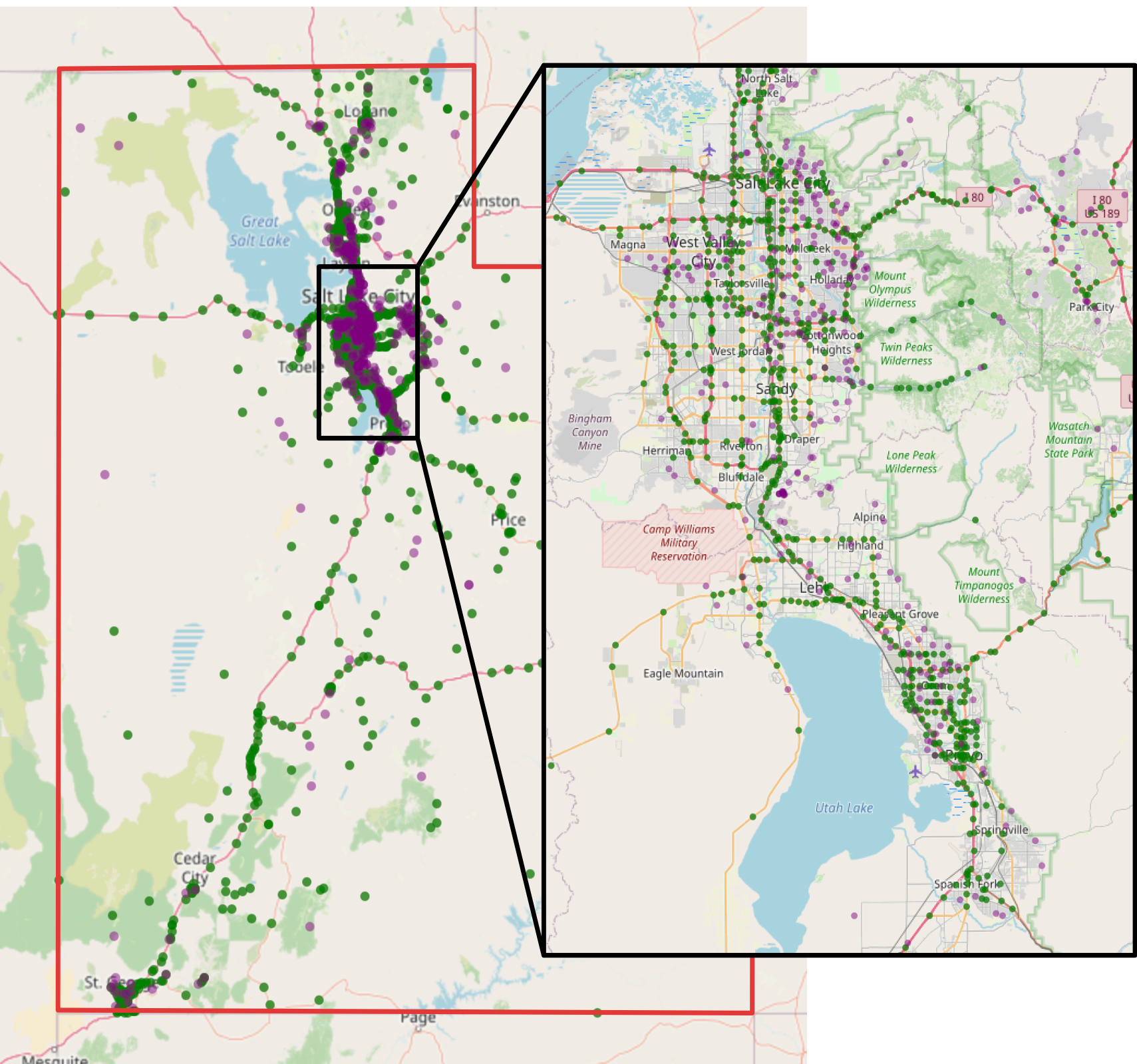
In 2021, air pollution was the second leading risk factor for early death worldwide. To help improve air quality, we must monitor and quantify the air quality problem. Utah has only **26** official state-run air quality monitoring stations, and each station typically represents air quality within a limited local range of about **0.5–4 km**. This leaves many communities without fine-grained air quality information.



Can we reuse existing public infrastructure to estimate air quality?

Public traffic cameras provide a possible way to expand coverage without installing new sensors. Utah already has **1,395** traffic cameras that we can access, many located near roads, neighborhoods, and areas without dense air quality sensor coverage. Traditional CNN models can estimate air quality from images, but they often learn camera-specific patterns. This means a model trained for one camera view may not transfer well to a new location. TCAM addresses this problem by using multimodal large language models that combine traffic camera images with weather data. The goal is not only to estimate PM2.5, but also to improve transferability across different cameras and cities.

Our best version, **PGML-TCAM**, adds physics-guided weather indicators to help the model distinguish pollution-related haze from fog, rain, clouds, and low-light conditions.



Methodology

TCAM is an end-to-end pipeline that uses public traffic cameras and weather data to estimate PM2.5. The system collects traffic camera images, matches them with weather observations and nearby PM2.5 measurements, and uses the aligned records to train multimodal models.

Data Collection

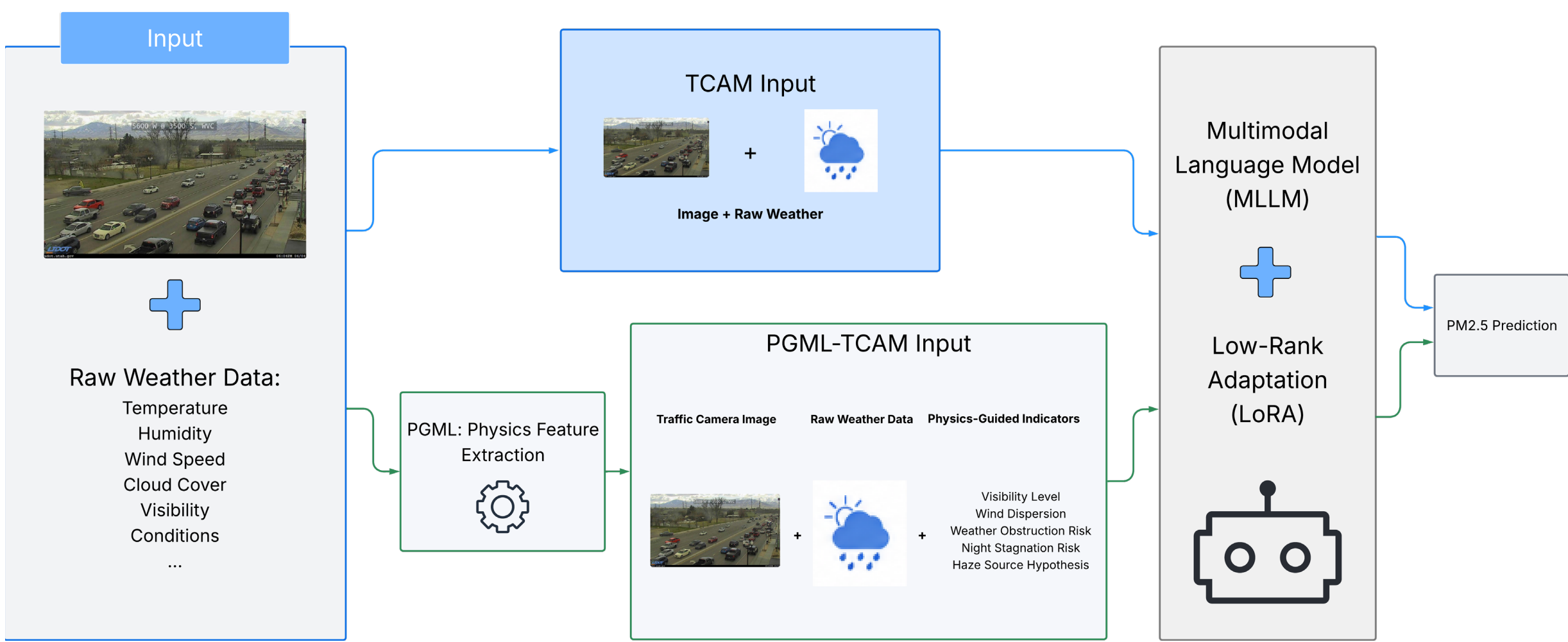
We collect public traffic camera images across Utah using Python scripts. The system also gathers weather observations and PM2.5 measurements from nearby monitoring stations. These data sources allow TCAM to connect what the camera sees with local weather conditions and measured air quality.

Instruction Fine-Tuning

We fine-tune multimodal large language models using instruction fine-tuning with LoRA. Instead of retraining the full model, LoRA updates a small set of adapter parameters, making training more efficient. The model learns to use both visual cues, such as haze and visibility, and weather information, such as wind, humidity, and cloud cover, to predict PM2.5.

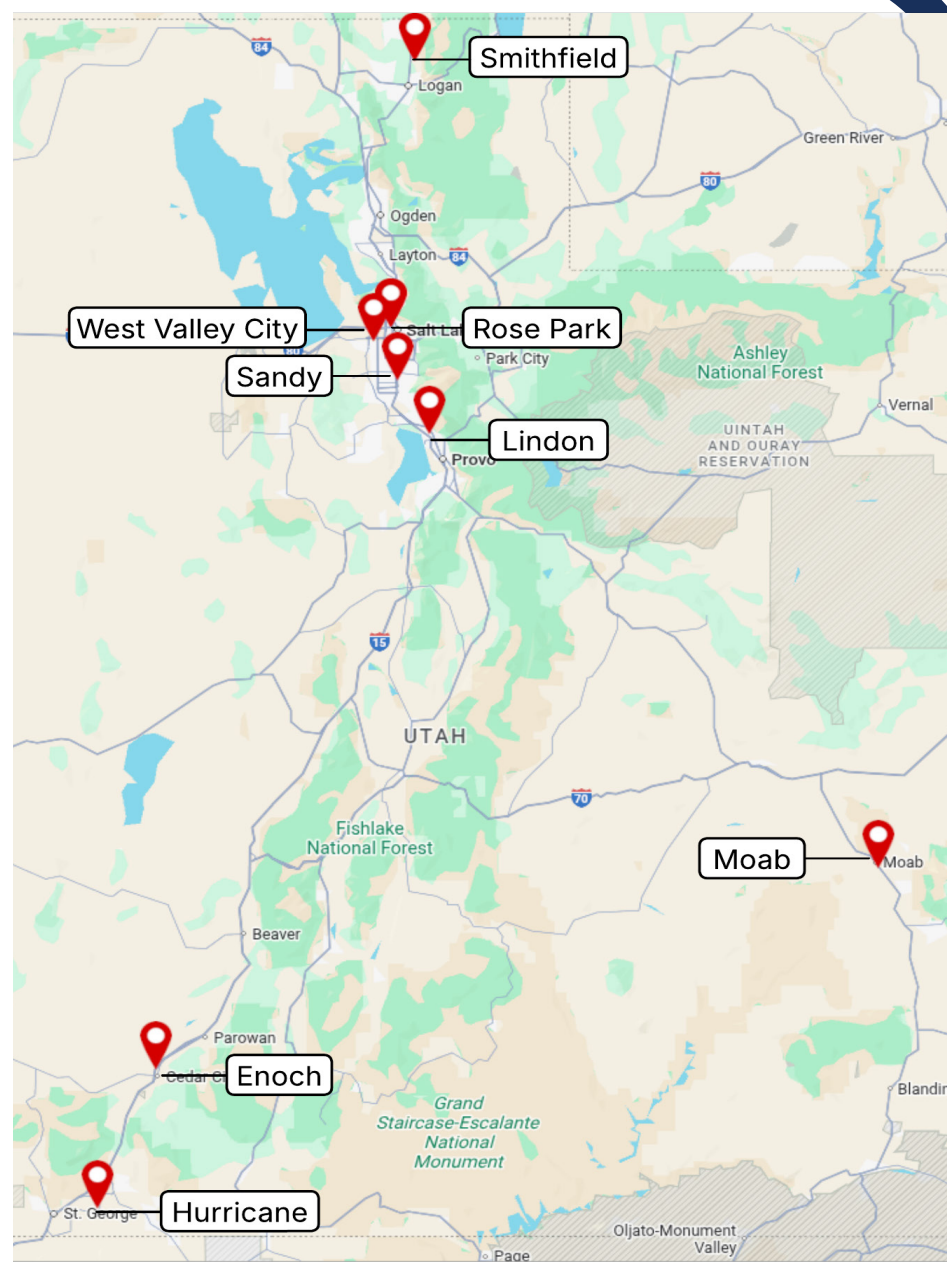
PGML-TCAM

PGML-TCAM is the physics-guided version of TCAM. It adds weather-based indicators such as visibility level, wind dispersion, weather obstruction risk, nighttime stagnation risk, and haze source hypotheses. These indicators help the model separate pollution-related haze from fog, rain, clouds, and low-light conditions, improving prediction consistency and transfer to unseen locations.

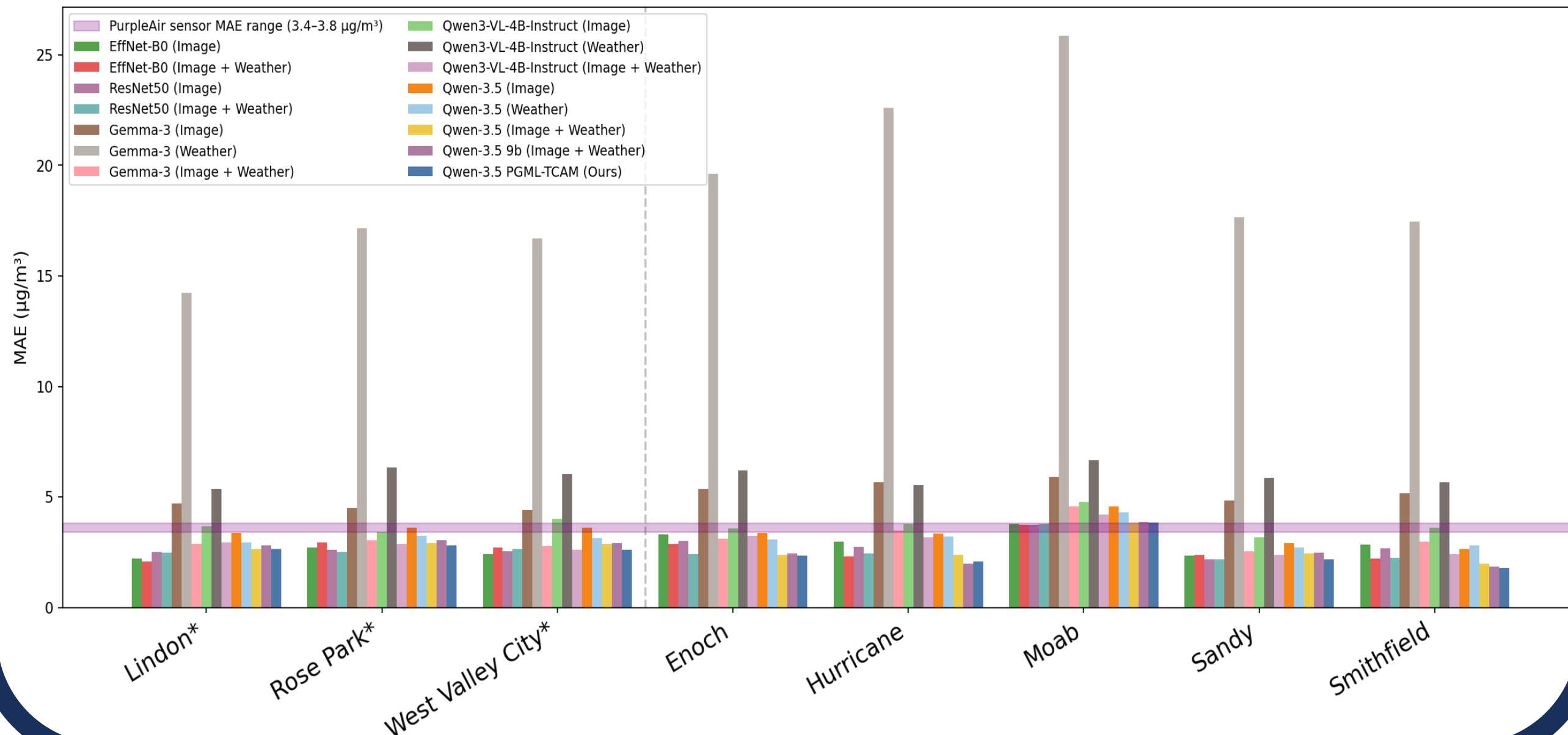


Results

We evaluated TCAM on a deployment-style test set from **April 2025**, using **22,387** traffic camera images from eight Utah cities. Five of these cities were unseen during training, allowing us to test whether the model can transfer to new locations. We compared PGML-TCAM with CNN baselines, CNN image-plus-weather models, and multiple MLLMs, including Gemma-3, Qwen3-VL-4B-Instruct, Qwen-3.5 4B, and Qwen-3.5 9B.



PGML-TCAM achieved the best overall balance among the tested models. Across all eight cities, it reached the highest correlation of **0.401**, the lowest MAE of **2.47 $\mu\text{g}/\text{m}^3$** , and the lowest MAPE of **50.0%**, while matching the lowest rounded RMSE of **3.69 $\mu\text{g}/\text{m}^3$** . On the five unseen cities, PGML-TCAM also performed best, with an RMSE of **3.62 $\mu\text{g}/\text{m}^3$** , an MAE of **2.39 $\mu\text{g}/\text{m}^3$** , and the strongest transferability score of **0.926**. These results show that physics-guided indicators help TCAM generalize better to new locations.



Future Work

Future work will test TCAM outside Utah and across different camera networks to better measure generalization. We also plan to improve PGML-TCAM with stronger temporal context, smoke-event indicators, inversion indicators, and better calibration methods so the model can respond more accurately to sudden PM2.5 changes while avoiding errors from fog, rain, and low-light conditions.