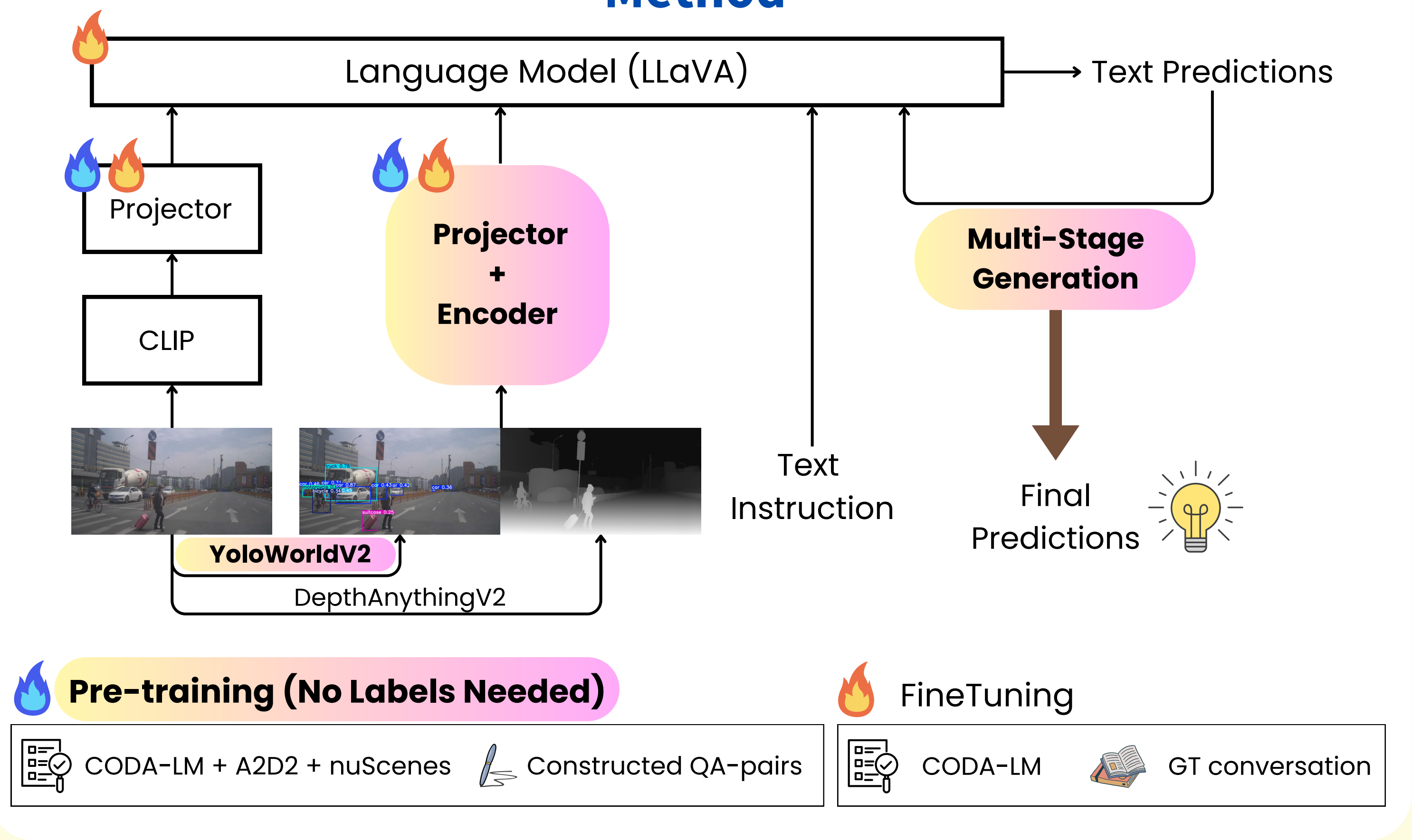# PREVISION

Pre-training Enhanced Versatile Integration of Semantics, Images, and Object Detection for Novel Corner Case Analysis

許景淯(B10901036) 傅啟恩(B11901174) 李冠儀(B10901091) 劉庭均(B10901039)

NTU DLCV-FALL 2024 Analysis of Corner Cases in Autonomous Driving

Method

Language Model (LLaVA) → Text Predictions

Projector

CLIP

Projector + Encoder

Multi-Stage Generation

Text Instruction

Final Predictions

YoloWorldV2

DepthAnythingV2

🔵 **Pre-training (No Labels Needed)**

CODA-LM + A2D2 + nuScenes   Constructed QA-pairs

🔥 FineTuning

CODA-LM   GT conversation

Ablation Study

Conclusion