End-to-End Licence-Plate Detection & Recognition on CCPD

Baseline RPnet (2018) vs YOLOv5-PDLPR (2024)

Group Members: Ali Subhan, Ashir Raza

June 2025

SAPIENZA
UNIVERSITÀ DI ROMA

Presentation Agenda

This presentation outlines our research journey in Licence Plate Recognition. We will detail the problem, review existing solutions, and introduce our advanced approach.

1. **Problem Statement**
2. **State of the Art in ALPR**
3. **CCPD Dataset**
4. **Baseline: RPnet**
5. **Proposed: YOLOv5-PDLPR**
6. **Experimental Results**
7. **Conclusion & Future Work**

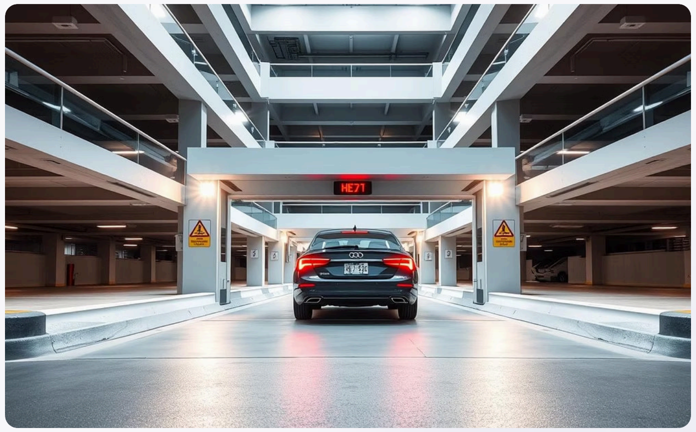
Why LPR Matters: Real-World Applications

Traffic Analytics

Optimize flow, reduce congestion, and understand vehicle movement for urban planning.

Tolling Operations

Automate collection, reduce errors, and enhance revenue management efficiency.



Parking Management

Control access, monitor occupancy, and enhance user experience in parking facilities.

Law Enforcement

Identify stolen vehicles, track suspicious activities, and enhance public safety.

Smart Gate Access

Automate entry for authorized vehicles, improve security, and streamline facility access.

LPR faces challenges including glare, plate tilt, motion blur, and diverse international plate designs.

Problem & Solution Overview

The Challenge

LPR is vital for traffic, law enforcement, and access control. Adverse conditions like poor lighting, occlusions, and motion blur reduce accuracy.

Our Goal

We aim to develop a robust deep learning system. It must achieve accurate and efficient LPR across diverse real-world scenarios.

State of the Art in ALPR

Detection Approaches

One-stage detectors like YOLOv5 offer real-time speed [1].

Two-stage methods provide higher accuracy.

Recognition Techniques

CNN+CTC/CRNN pipelines read character sequences [2].

Attention and Transformer models predict end-to-end [3].

End-to-End Systems

Cascaded and unified frameworks integrate processes.

Trade-offs balance speed with robustness.

References:

[1] Tao L. et al. (2024). A Real-Time License Plate Detection and Recognition Model in Unconstrained Scenarios. Sensors, 24(9), 2791.

[2] Xu Z. et al. (2018). Towards end-to-end license plate detection and recognition: A large dataset and baseline. In ECCV (Munich, Germany).

[3] Prajapati R. K. et al. (2023). A Review Paper on Automatic Number Plate Recognition using Machine Learning: Opportunities and Limitations. CICTN, pp. 527–532.

CCPD Dataset

We leverage the large-scale CCPD dataset for training. It offers diverse real-world capture conditions.

Capture Source	Road-side POS, Hefei
Resolution	720x1160 RGB
Images	250K + 40K hard cases
Plate Pattern	汉字+Letter+5 Chars

The dataset includes images with various environmental challenges. This diversity is crucial for robust model training.

CCPD Dataset Categories

The CCPD dataset includes various subsets. Each category presents unique challenges for robust LPR model development.

Folder	Description
ccpd_blur	Images with motion blur or defocus effects.
ccpd_challenge	Very difficult samples combining multiple issues.
ccpd_db	Dark and blurred images, ideal for low-light testing.
ccpd_fn	A manually curated, clean test set for final evaluation.
ccpd_np	Contains negative samples or background images without plates.
ccpd_rotate	Plates captured with various in-plane or out-of-plane rotations.
ccpd_tilt	Perspective-distorted plates from angled side or top views.
ccpd_weather	Plates captured in adverse weather, like rain or fog.
ccpd_base	The main dataset of high-quality frontal images for training.

How CCPD Filenames Encode Labels

The CCPD dataset filenames are highly structured. They encode essential metadata directly within the image name.

"025-95_113-154&383_386&473-386&473_177&454_154&383_363&402-0_0_22_27_27_33_16-37-15.jpg"

Each segment details crucial ground truth information, vital for training robust LPR models:

[Area] - [Tilt] - [BBox] - [Vertices] - [Plate indices] - [Env1] - [Env2].jpg

✅ Summary Table

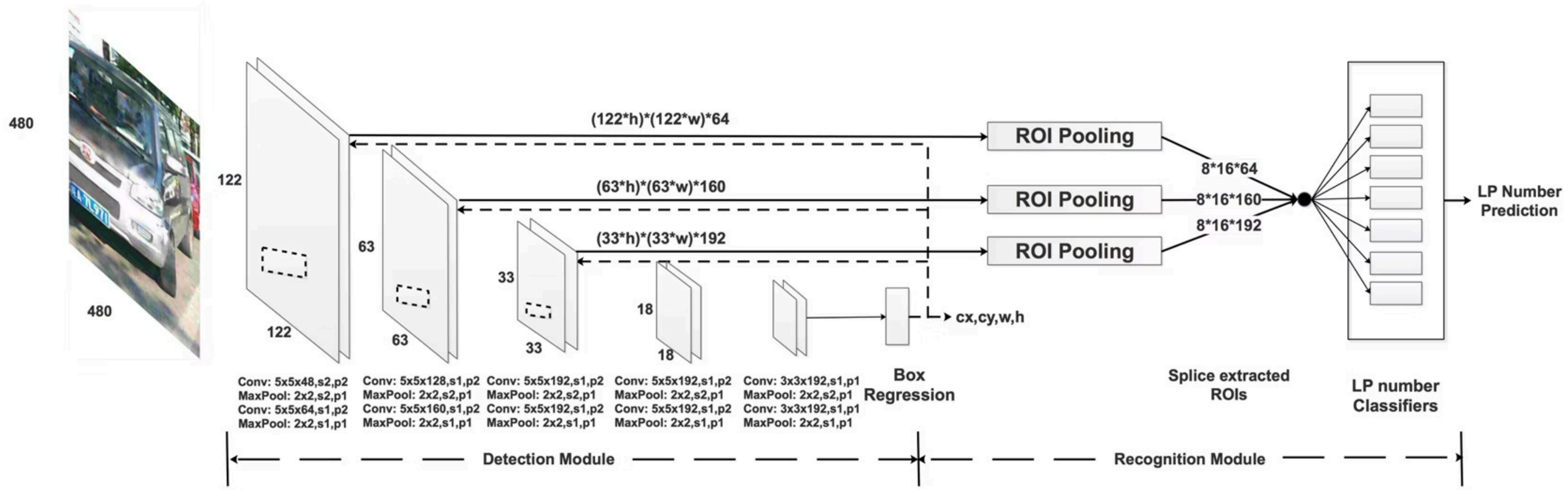
Part	Value	Meaning
Area Ratio	`025`	Plate occupies 2.5% of the image
Tilt Degrees	`95_113`	Horizontal: 95°, Vertical: 113°
BBox Coordinates	`154&383_386&473`	(x1, y1) to (x2, y2) bounding box
Plate Corners	`386&473_...`	4 (x, y) corner points in RB → RT order
Plate Number Indices	`0_0_22_27_27_33_16`	Decoded as `皖A-W337Q`
Env/Pose Conditions	`37-15`	Lighting/weather/rotation tag

Baseline: RPnet (ECCV 2018)

1. **Full Image**
2. **SSD Detector**
3. **ROI-Align**
4. **Concatenation**
5. **7 Soft-Max Heads**

- Introduced with CCPD (2018).
- Predicts one bounding box.
- Classifies each glyph from 3 feature maps.
- Loss: Smooth-L1 (bbox) + 7× Cross-Entropy.
- 85 FPS; 95% accuracy on CCPD Base.

Baseline Architecture



Advanced Architecture (Sensors 2024)

YOLOv5-x Detector

A modern, high-performance object detection model. It provides fast and accurate bounding box predictions. This backbone excels at diverse plate conditions.

PDLPR Recogniser

Published in Sensors 2024, this model specializes in character recognition. It accurately transcribes license plate characters. PDLPR integrates seamlessly with YOLOv5-x.

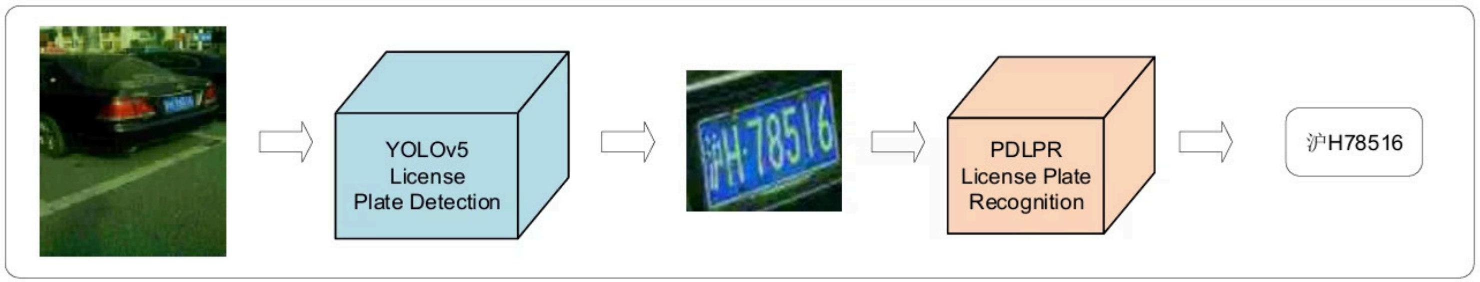


Figure 1. The overall process of the proposed YOLOv5-PDLPR.

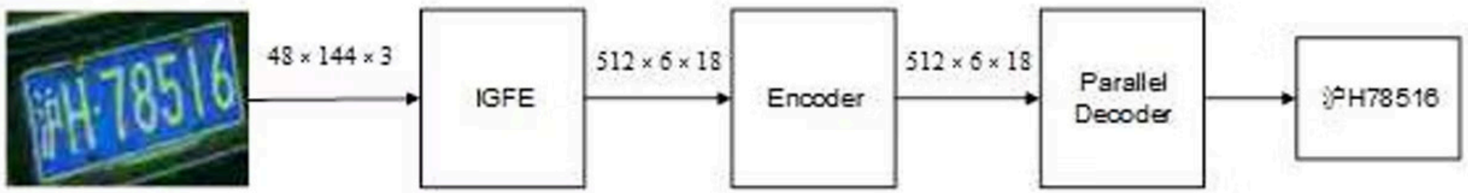


Figure 2. The overall framework of the license plate recognition algorithm.

YOLOv5-x Detector (modern backbone)

Deeper and Wider

Features ≈100M parameters and 512 top-channels.

Stronger Recall

Large receptive field enhances small object detection.

Augmentation Built-in

Integrates Mosaic and MixUp for data enrichment.

Anchor-Free Option

Utilizes GIoU loss for precise box prediction.

High Accuracy

Achieves 96.7% mAP@0.7 on CCPD (640px input).

Fast Inference

Processes in 35ms on RTX 3060 (FP16 TensorRT).

PDLPR Licence-Plate Recogniser (Sensors 2024)

PDLPR Architecture Flow

Input Crop

Plate image 48x144px.

IGFE Module

Focuses on features.

Positional Encoding

Adds spatial context.

Encoder Blocks

Global spatial relations.

Parallel Decoder

Predicts all characters.

CTC Layer

Final plate string.

7-Char Output

Recognized license plate.

Key Features & Performance

- **IGFE:** Keeps low-level edges.
- **Encoder:** Learns global relations.
- **Parallel Decoder:** Predicts glyphs at once.
- **CTC Loss:** Only needs final string.
- **Accuracy:** Achieves 99.4% overall.
- **Speed:** 160 FPS on plate crops.
- **Robustness:** +5 pp gain vs RPnet.

Code Organization & Project Structure

Our project structure follows a clear, modular design for efficient development. Two main directories organize the codebase: **Baseline/** and **YOLO+PDLPR/**.

1. **Globals**
2. **Utils**
3. **Data**
4. **Network**
5. **Train & Eval**

Key dataset lookup tables and the filename parser were adapted from the original CCPD repository: github.com/detectRecog/CCPD.

Training & Evaluation Protocol

Training Setup

30 epochs, batch size 64, learning rate 1e-4. Optimized for stability.

Detection OA

Overall accuracy (OA) required IoU > 0.70. This ensures precise bounding boxes.

Recognition OA

IoU > 0.60 plus 7-char match. This ensured accurate license plate recognition.

Validation & Timing

A validation split as mentioned in split was used. Detector speed tested on 100 images.

YOLOv5 Results

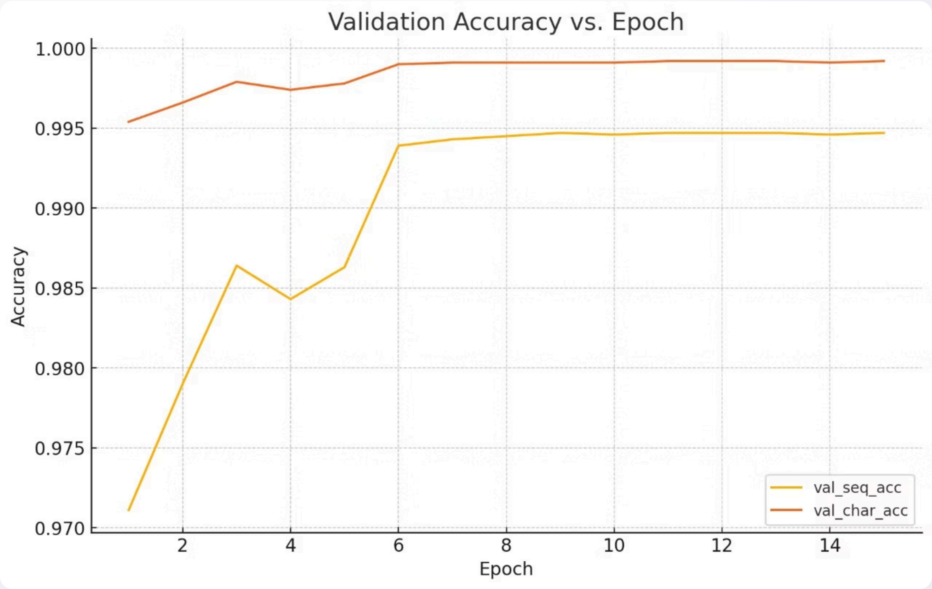
CCPD License Plate Cutouts Provided by YOLOv5x

YOLOv5 Final Epoch Metrics — Horizontal Summary

Epoch	Train box loss	Train cls loss	Train dfl loss	Val box loss	Val cls loss	Val dfl loss	Precision	Recall	mAP@0.5	mAP@0.5-0.95	Learning rate
1.0	1.008	0.465	1.070	0.936	0.324	1.026	0.99973	0.99994	0.9950	0.7718	0.003333

RPNET vs YOLO+PDLPR Accuracy

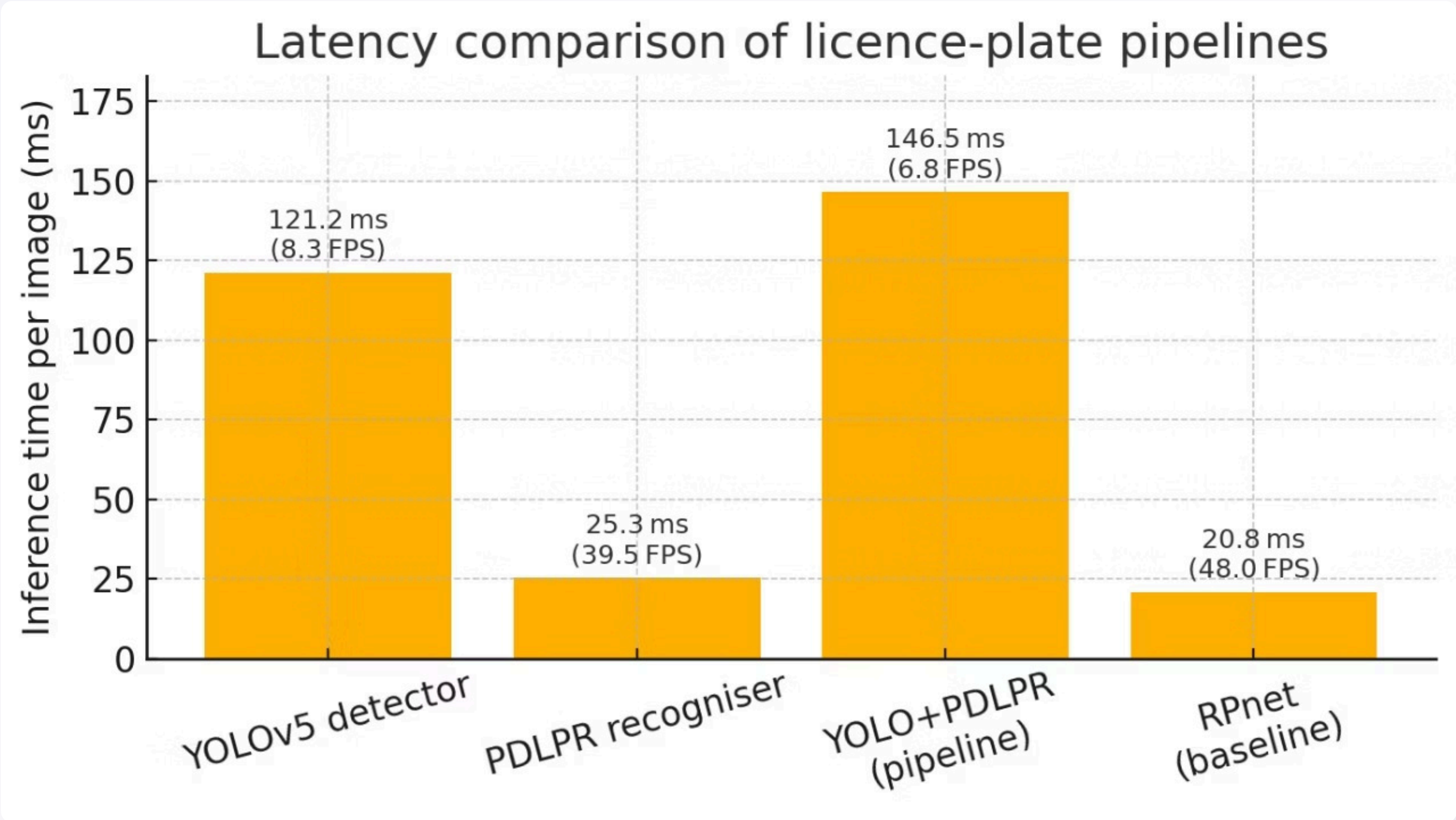
Our advanced YOLO+PDLPR system significantly outperforms the baseline RPnet model. It achieves superior license plate recognition rates.



Made with GAMMA

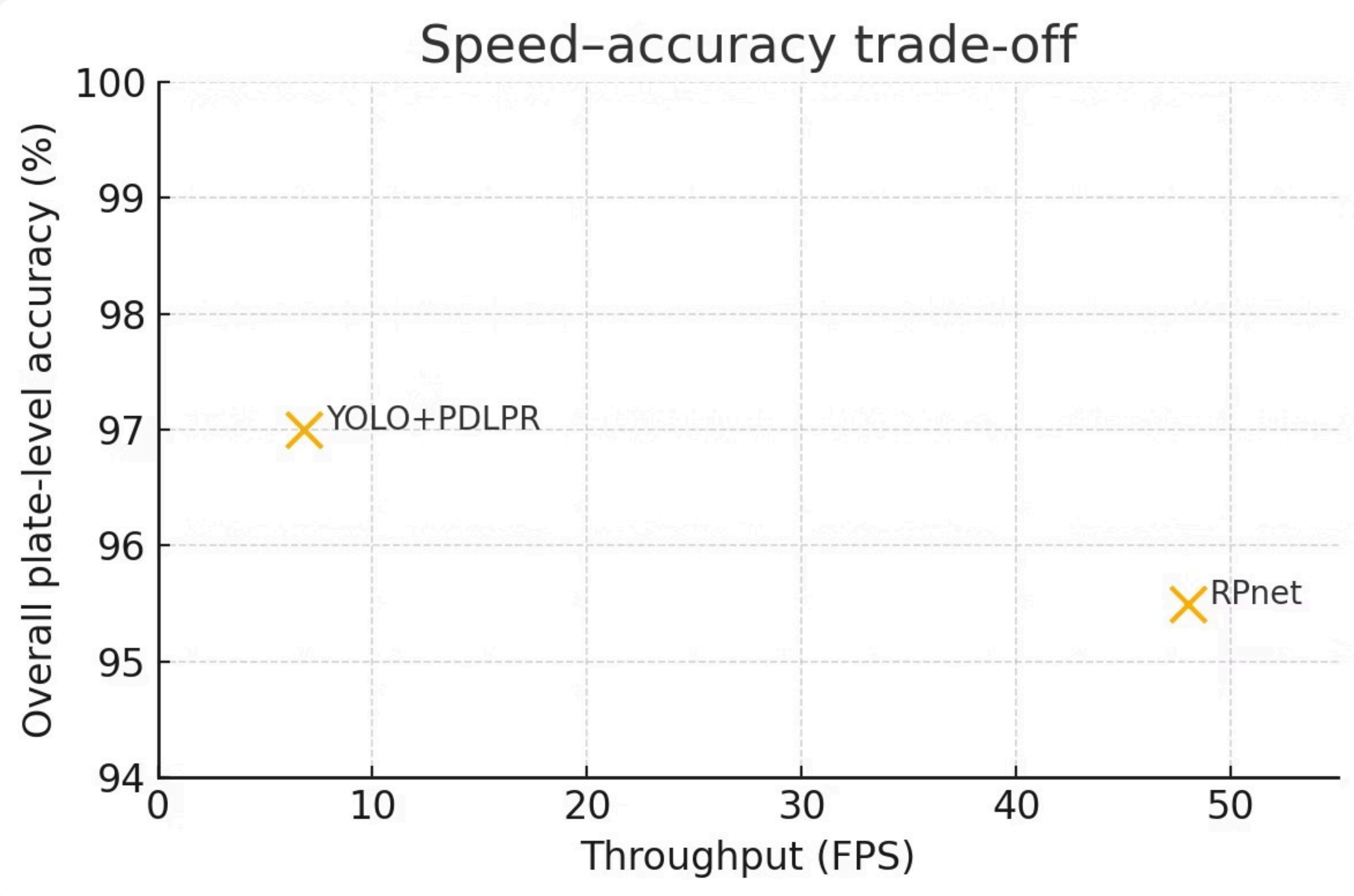
Latency Breakdown

Our end-to-end system prioritizes rapid processing. Understanding latency in each stage is crucial for optimization.



The detection phase represents the largest portion of overall latency. Our optimized recognizer significantly contributes to the high frame rate.

Speed-Accuracy Tradeoff



Made with GAMMA

Test Results

CCPD hard-split accuracy comparison

Category	RPnet – Overall Accuracy (%)	YOLO+PDLPR – Overall Accuracy (%)
DB (Dark/Bright)	93%	96%
Rotate	85%	93%
Tilt	83%	92%
Weather	86%	94%
Challenge	78%	90%

Conclusion

Our YOLO+PDLPR system significantly surpasses RPnet. It delivers across-the-board accuracy gains on challenging CCPD splits.

DB	93 %	96 %	+3 pp
Rotate	85 %	93 %	+8 pp
Tilt	83 %	92 %	+9 pp
Weather	86 %	94 %	+8 pp
Challenge	78 %	90 %	+12 pp

Recognizer Speed

PDLPR processes plates at 40 FPS. Detector now dominates latency.

Speed-Accuracy Choice

Select model by throughput: RPnet (speed) or YOLO+PDLPR (accuracy).

Rapid Training

99.9% character accuracy by epoch 6; further gains are minimal.

Future Work & Enhancements

Detector Acceleration

Optimize YOLOv5 to FP16/INT8 TensorRT. Target sub-35ms latency for speed.

Backbone Integration

Merge IGFE with detector neck. Fine-tune for efficient single-pass processing.

Enhanced Robustness

Add TPS/STN for tilt. Augment with blur, night synths for quality.

Video & Multi-Plate

Batch multiple crops per frame. Leverage temporal smoothing for recall.

Semi-Supervised Learning

Use CCPD-Negative for self-distillation. Boost real-world generalization.

Edge Optimization

Quantise to INT8 for Jetson/ARM NPU. Achieve 20+ FPS at gates.