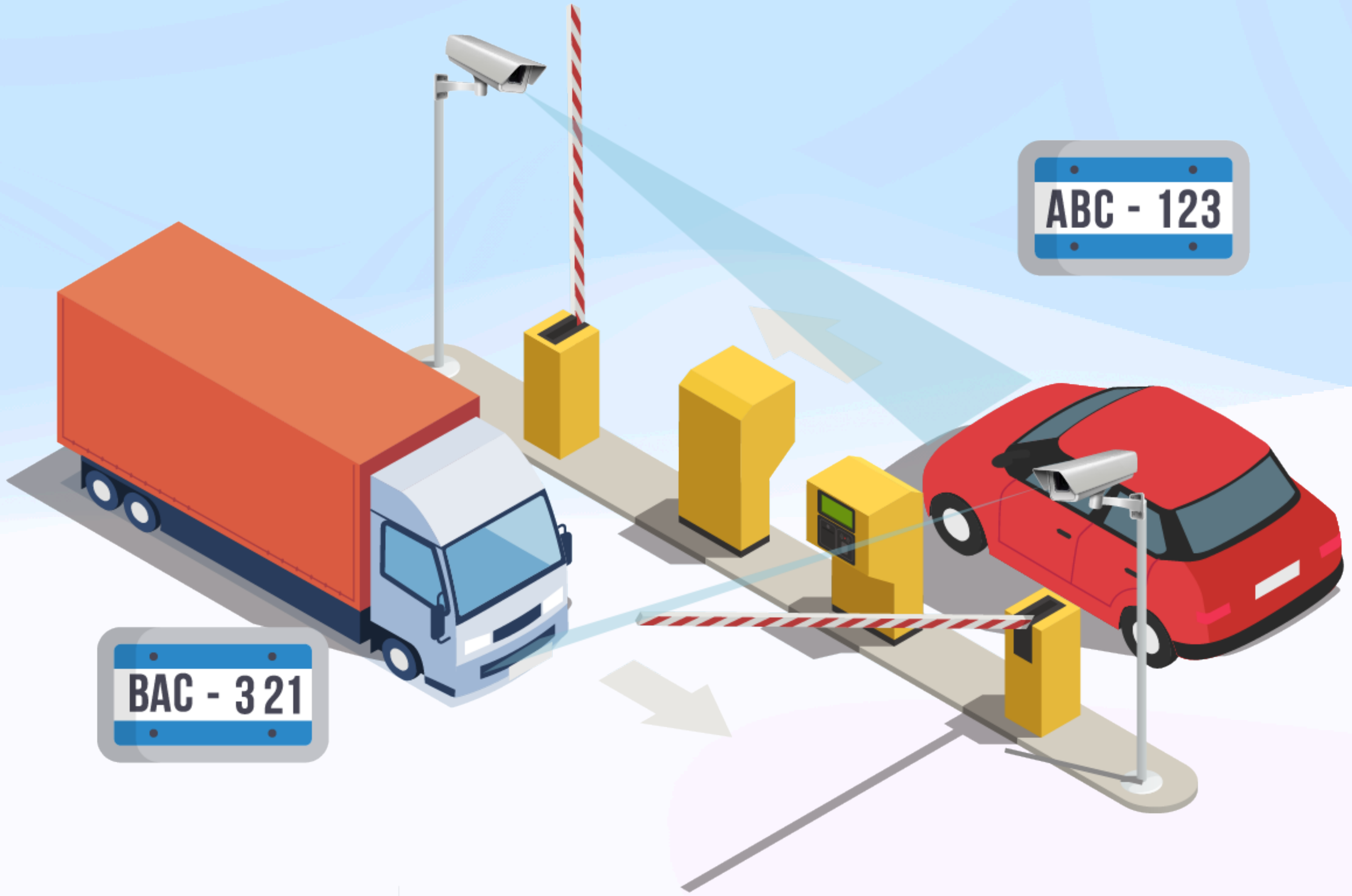
Car Plate Recognition and Reconstruction with Deep Learning

Computer Vision Exam

Esteban Gatein (2222014) - 19/06/2025

Motivation

- License Plate Recognition is essential in:
  - Intelligent traffic systems
  - Law enforcement
  - Automated access control
- Real-world challenges:
  - Motion blur, poor lighting, angle distortion



Source: Automated License Plate Recognition Parking Solution. (s. d.). Axiomtek

Project Scope & Objectives

- **Goal:** implement and evaluate a license plate recognition pipeline
- **Baseline:** build a simple recognition model to decode cropped license plates
- **Paper reproduction:** reproduce the two-stage method proposed by *Tao et al., 2024 - "A Real-Time License Plate Detection and Recognition Model in Unconstrained Scenarios"*
    1. **YOLOv5** for license plate detection
    2. **PDLPR** model for character sequence recognition
- **Objective:** compare the baseline with the reproduced method

Dataset - CCPD

- Real-world plate dataset
- 250k+ images with bounding boxes and ground truth labels
- Covers diverse plate formats, conditions, angles
- Organized in different subsets
  - Trained on *weather* subset
  - Evaluated on *base* subset

Source: Xu, Z., Yang, W., Meng, A., Lu, N., Huang, H., Ying, C., & Huang, L. (2018). Towards End-to-End License Plate Detection and Recognition : A Large Dataset and Baseline.

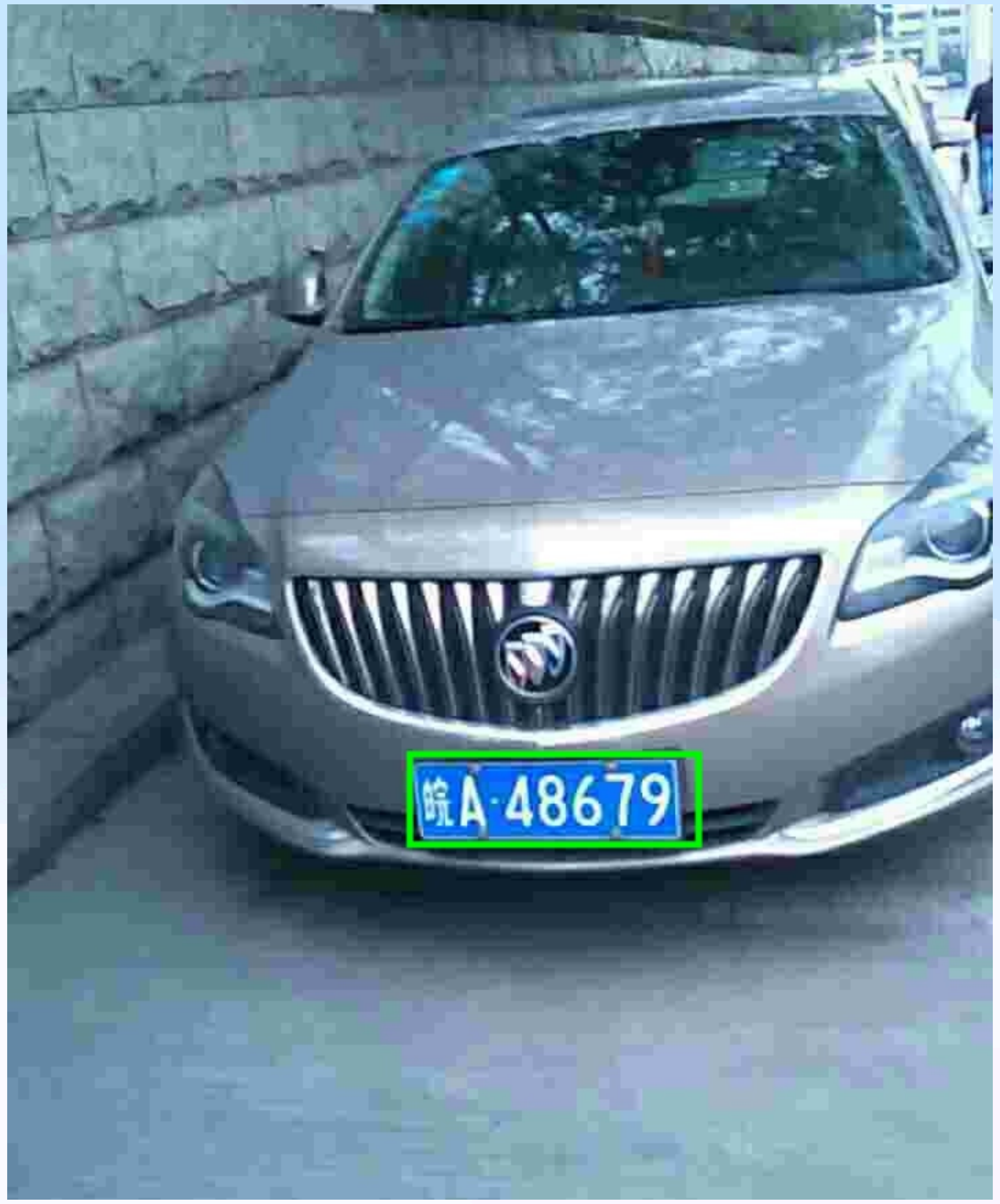
Baseline Model

- **Bounding box model:**
  - Simple CNN to predict the coordinates, uses MSE as loss
- **OCR model:**
  - Sequence-to-sequence architecture with attention
  - Decodes the text from the cropped images, uses cross-entropy as loss
- **Limitations:**
  - OCR model sensitive to misaligned or bad crops
  - Character predictions deteriorate with input noise

Paper Reproduction

YOLOv5 Detection Results

- Trained with plate bounding boxes from CCPD
- Accurate and fast plate detection:
    - Reached accuracy (IoU > 0.7): 93.4%

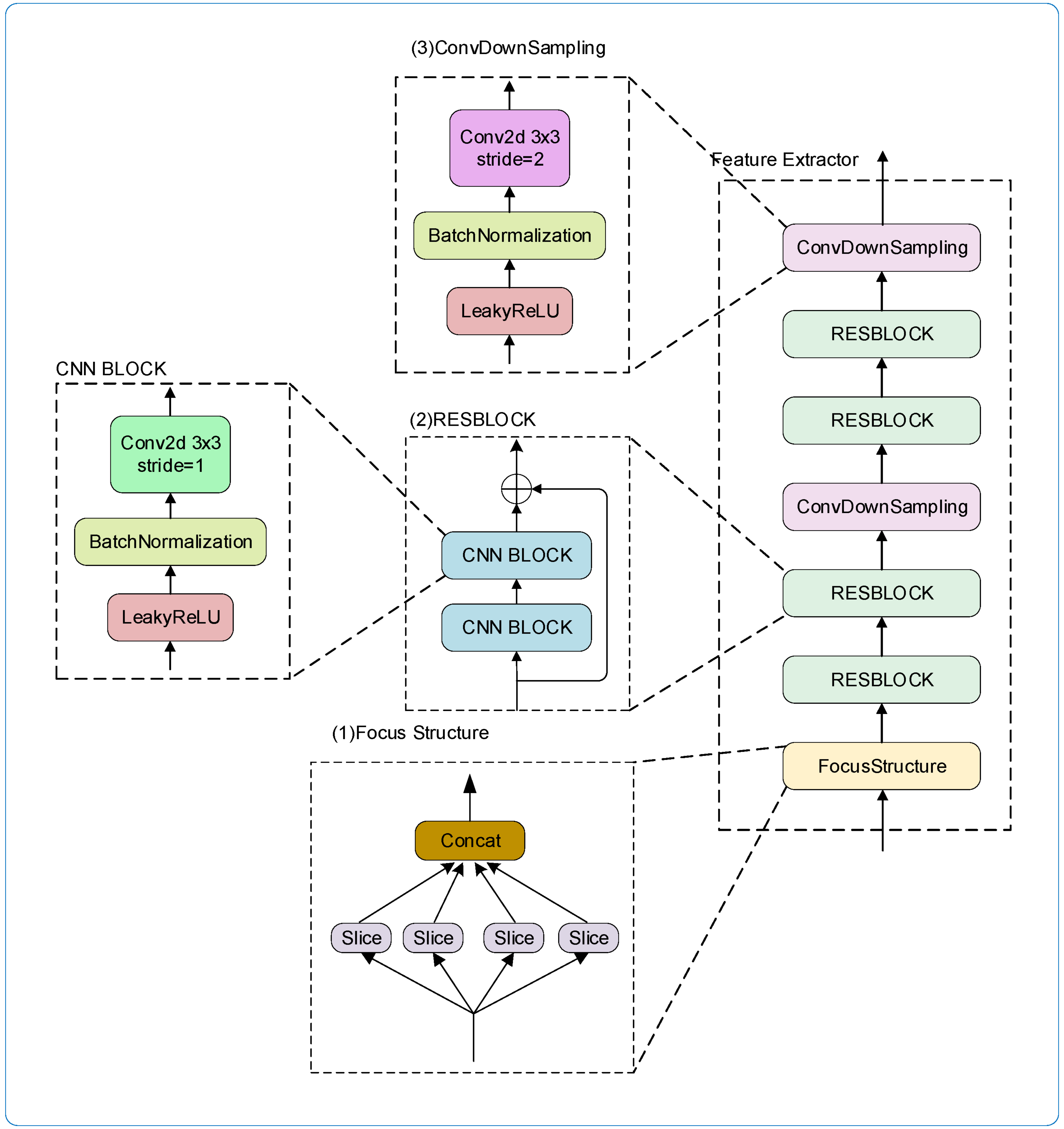


Inference example

Paper Reproduction

PDLPR

- Parallel Decoding License Plate Recognition (PLDPR)
- Transofrmer-based with CTC loss
- Training issues:
    - Not enough ressources
    - Model colapse
- Not a structural issue



Source: Tao, L., Hong, S., Lin, Y., Chen, Y., He, P., & Tie, Z. (2024). A Real-Time License Plate Detection and Recognition Model in Unconstrained Scenarios

Results and observations

- For the detection plate model
  - The classic CNN with MSE could not learn the bboxes (acc. 5.9%)
  - YOLOv5 solved the task (acc. 93.4%)
- For the text recognition model
  - Baseline not complex enough, but started learning
    - Normalized Levenshtein distance < 1
    - Could predict the first two characters
  - Issues with PLDPR, no structural problems (lack of resources)

Why the Paper Works Better

Theoretically

- **YOLOv5**
  - Trained for object detection → accurate bounding boxes
- **PDLPR**
  - Uses self-attention and parallel decoding
  - Robust to variable-length (not necessary) and noisy inputs
- **Baseline limits**
  - No detection logic for the bounding box & OCR model not complex enough

Final Evaluation

- **Positive outcomes**
  - **YOLOv5** detector performed well on CCPD
  - **CNN** baseline provided some useful bounding boxes
  - **Seq2Seq** provided some accurate characters (measured with Levenshtein)
- **Remaining challenges**
  - CNN accuracy far from YOLOv5
  - Could not get results from PDLPR