Mesopotamian Journal of Computer Science

Volume 5 | Issue 1 Article 21

1-1-2025

End-to-End License Plate Detection and Recognition in Iraq Using a Detection Transformer and OCR

Younis Al-Arbo
Department of Computer Science, College of Education for Pure Science, University of Mosul, Nineveh, Iraq, younis.bayati@uomsul.edu.iq

Hanaa F. Mahmood
Department of Computer Science, College of Education for Pure Science, University of Mosul, Nineveh, Iraq

Asmaa Alqassab
Department of Computer Science, College of Education for Pure Science, University of Mosul, Nineveh, Iraq

How to Cite This Article
Al-Arbo, Younis; Mahmood, Hanaa F.; and Alqassab, Asmaa (2025) "End-to-End License Plate Detection and Recognition in Iraq Using a Detection Transformer and OCR," *Mesopotamian Journal of Computer Science*: Vol. 5: Iss. 1, Article 21.
DOI: https://doi.org/10.58496/MJCSC/2025/021
Available at: https://map.researchcommons.org/mjcsc/vol5/iss1/21

Mesopotamian journal of Computer Science
Vol. **(2025)**, 2025, **pp**. 329–340
DOI: https://doi.org/10.58496/MJCSC/2025/21 ;ISSN: 2958-6631
https://mesopotamian.press/journals/index.php/cs

Research Article

End-to-End License Plate Detection and Recognition in Iraq Using a Detection Transformer and OCR

Younis Al-Arbo [1,*], Hanaa F Mahmood [2], Asmaa Alqassab[3]

[1,2,3] Department of Computer Science, College of Education for Pure Science, University of Mosul, Mosul, Iraq.

ARTICLEINFO

Article History

Received 27 Jul 2025
Revised 31 Aug 2025
Accepted 19 Sep 2025
Published 30 Sep 2025

Keywords

Automatic License Plate Recognition (ALPR)

DEtection TRansformer (DETR)

Deep Learning for Object Detection

Optical Character Recognition (OCR)

Region-Specific Vehicle Identification

ABSTRACT

Automatic license plate recognition (ALPR) is essential for intelligent transportation systems, traffic monitoring, and law enforcement. Although deep learning has significantly progressed ALPR, the adaptation of these techniques to the specific characteristics of Iraqi plates presents challenges due to the variety of fonts, differing plate dimensions, and intricate real-world driving conditions. This paper presents an end-to-end Automatic License Plate Recognition (ALPR) framework specifically designed for Iraqi license plates. This system integrates a Detection Transformer (DETR) for the identification of plates alongside a Convolutional Recurrent Neural Network (CRNN)-based Optical Character Recognition (OCR) module for the purpose of character recognition. The system is trained and evaluated on a newly developed dataset of 1,000 annotated images representing diverse driving scenarios in Iraq. It achieves a mean Average Precision (mAP@[.5:.95]) of 0.91 for detection and a full-plate OCR accuracy of 93%. The accuracy of DETR surpasses that of YOLOv5, Faster R-CNN, and SSD, as demonstrated by comparative experiments. Conversely, a lightweight transformer (DeiT-Tiny-Det) approaches DETR's performance at faster inference velocities, illustrating a practical trade-off between speed and precision. Ablation studies confirm the importance of robust detection for end-to-end accuracy, while error analysis shows that low-light and character-level confusions remain the main challenges. The results indicate that transformer-based detectors, in conjunction with specialised OCR models, yield dependable region-specific ALPR appropriate for real-world application.

1. INTRODUCTION

Automatic license plate recognition (ALPR) is a critical component in the realisation of smart transportation, traffic regulation, electronic toll collection, and modern urban surveillance [1], [2], [3]. The task involves not only the reliable localization (detection) of license plates in unconstrained images but also the precise recognition of their alphanumeric contents under diverse environmental conditions [4], [5]. Although recent advances in deep learning have transformed ALPR pipelines and led to major breakthroughs in both detection and recognition accuracy, adapting these solutions for region-specific plate standards remains an open challenge [6]–[9].

Iraqi vehicle plates present a particularly compelling case: they use only English uppercase letters (A–Z) and digits (0–9), with the first two digits uniquely denoting the issuing governorate. Despite the apparent simplicity of the character set, practical deployment in Iraq is hindered by significant real-world challenges—including varied plate fonts, aspect ratios, and mounting styles, as well as severe lighting changes, motion blur, occlusion, dirt, and the need for governorate-level recognition.

Traditional ALPR pipelines, originally designed for Western or East Asian plate formats, often require region-specific heuristics and perform suboptimally when directly transferred to the Iraqi context [10], [12]. Most existing systems are not designed to handle the structured mapping between the first two digits and the governorate, which is essential for administrative and law enforcement applications. Moreover, the lack of large, well-annotated, and publicly available datasets for Iraqi plates has limited the development of robust, generalizable models [8], [13].

While some prior efforts have explored Iraqi license plate recognition, these works have typically relied on limited experiments and non–end-to-end pipelines, often using handcrafted features or CNN-based detectors with constrained generalization capacity. Such approaches struggle under real-world Iraqi conditions, where occlusion, scale variation, and low-light scenarios are common. To date, no study has systematically investigated the use of transformer-based detection combined with modern sequence-based OCR for Iraqi plates. This gap motivates the present work.

*Corresponding author. Email: younis.bayati@uomsul.edu.iq

Convolutional neural networks (CNNs) and hybrid pipelines that combine human-crafted features with machine learning classifiers have been the mainstays of earlier ALPR work [9]. These methods were accurate enough in a controlled setting, but they were not always able to handle the complexities of city life. In particular, CNN-based detectors struggled with scale variation and cluttered backgrounds and different plate dimension, while hybrid pipelines lacked the adaptability needed to handle font diversity and low-visibility scenarios. Recent surveys further highlight the importance of lightweight detectors and synthetic data augmentation to improve generalization under such challenging conditions[11].By contrast, the present study departs from these limitations by leveraging a transformer-based detector (DETR) that incorporates global scene context and eliminates dependence on anchor heuristics, coupled with a deep OCR model specifically tuned to the alphanumeric structure of Iraqi plates

The swift advancement of transformer-based designs such as the Detection TRansformer (DETR) [14], [15] has marked a new phase in object recognition, enabling models to exploit global scene context, discard anchor boxes, and adapt more effectively to varied visual domains. When paired with advances in optical character recognition (OCR), especially deep neural approaches designed for constrained alphanumeric sets, these models offer a promising pathway toward end-to-end ALPR for Iraq and similar contexts [16], [17].

In this work, we propose a unified pipeline that harnesses DETR for plate detection and a dedicated deep OCR system for robust alphanumeric recognition. Our approach is assessed using a freshly developed, realistic dataset comprising 1,000 annotated Iraqi vehicle photos, constituting the most extensive benchmarking initiative to date for the region. The pipeline achieves a mean Average Precision (mAP) of 0.91 for detection and delivers strong recognition performance, as detailed in subsequent sections. Our contributions encompass:

- A comprehensive, annotated dataset of Iraqi license plates encompassing all principal governorates and complex imaging conditions.
- An advanced detection model based on DETR, refined to accommodate the unique visual characteristics of Iraqi license plates.
- A specialized OCR recognition system focused exclusively on English letters and digits, ensuring precise mapping of governorate codes.
- Extensive experimental results and analysis, including state-of-the-art mAP, precision-recall curves, confusion matrices, and region-specific quantitative metrics.

Our results suggest that transformer-based detectors and modern OCR, when adapted for local standards, can deliver reliable, scalable ALPR performance even in the most challenging real-world settings. Related approaches have shown success in other regional contexts [18]– [21], but our work represents the first rigorous evaluation for Iraqi plates using end-to-end deep learning.

2. RELATED WORK

2.1. License Plate Detection

The emergence of deep learning significantly altered the paradigm. Convolutional Neural Networks (CNNs) based on two-stage detectors like Faster R-CNN [3] provided crucial improvements in accuracy by first generating region proposals and then classifying them. Later one-stage detectors, including YOLO (You Only Look Once) [2] and SSD (Single Shot MultiBox Detector) [6], achieved an effective balance between speed and accuracy by conceptualising detection as a single regression task. More recently, transformer-based architectures have entered the computer vision domain. The Detection Transformer (DETR) [12] introduced an end-to-end approach that eliminates the need for hand-designed components like non-maximum suppression (NMS) and anchor generation, instead using a set-based global loss and bipartite matching to directly output detections. DETR offers streamlined pipelines and notable performance; however, its use in license plate detection, especially with difficult datasets, continues to be a subject of ongoing investigation.

Recent reviews indicate that while deep learning methods such as YOLO and its variants demonstrate effectiveness, notable challenges remain in practical applications. These include a high rate of false positives, imbalanced datasets, and a decline in performance in complex environments characterised by poor lighting and varied road conditions. These limitations underline the need for more robust, generalisable approaches to vehicle and license plate detection [23].

The literature reveals a significant gap in the availability of large-scale datasets for Iraqi license plates that are publicly accessible. Previous research has concentrated on small, privately collected datasets. Altyar et al. (2023), for instance, developed a system for Iraqi plates using a dataset of only 326 images, while other studies used fewer than 100 samples [22] . The limited availability of data constrains the generalisation of models and hinders equitable benchmarking. This lack of standardised data significantly impedes the advancement of comprehensive, nationwide ALPR systems.

2.2. License Plate Recognition (OCR)

Once plates are localised, OCR modules are applied to transcribe the alphanumeric sequence. Conventional OCR approaches often employed character segmentation followed by isolated character classification [5], [11], but segmentation-free, deep learning-based models—especially CRNNs [16], attention-based transformer decoders [17], and recent end-to-end methods—now dominate the field [6], [24]. Segmentation-free approaches offer significant benefits in the Iraqi context, as variations in fonts, mounting styles, dirt, and low-light conditions frequently result in touching or distorted characters. Approaching the entire plate as a continuous sequence, instead of relying on exact segmentation, enables these models to reduce error propagation and improve resilience in practical scenarios. Recent studies have proposed joint detection and recognition pipelines to enhance efficiency [20], [24].

In 2021, Faraj Humaidan developed a custom Arabic character database utilising MATLAB. By applying preprocessing techniques including filtering, edge detection, and greyscale conversion, the system was able to extract 854 accurate OCR recognitions and 846 successful localisations from a dataset of 870 photos.[25] Although early efforts had its limitations due to the computing power and methodology of the time, they were crucial in proving that an ILPR system could be implemented.

2.3. Region-Specific ALPR Systems

The extant literature on ALPR primarily concentrates on standards from Western and Asian contexts; however, a growing body of research is now adapting deep learning methodologies for other regions. For instance, [18] outlines adaptations for Latin American plates, whereas [8], [9], and [21] detail deep learning-based pipelines for Iraqi and Middle Eastern plate formats. Nevertheless, the combination of region-specific OCR and detection transformers for Iraqi plates has been the subject of a limited number of rigorous studies. Larger datasets, more robust models, and a focus on plate-specific structure have been stressed by recent regional research [8], [9], [13]. Our research addresses these requests by offering a thorough, end-to-end pipeline and public benchmarking for Iraq.

3. DATASET

In order to enable rigorous training and evaluation, we constructed a novel dataset of 1,000 high-resolution images of vehicles captured under real-world Iraqi driving conditions, including a wide range of illumination, weather, viewpoints, occlusions, and plate mounting styles [8], [9]. To provide comprehensive coverage of the most popular plate formats and governorates, images were gathered from various locations and environments. The dataset comprises a substantial quantity of images for each major governorate, with Erbil and Baghdad exhibiting the highest representation, thereby mirroring the actual vehicle distribution. Although the dataset offers realistic coverage of regional conditions, its relatively (1,000 images) remains a limitation; we plan to expand it in future work with larger-scale and more balanced data across all governorates. The dataset demonstrates a significant degree of diversity, including both daytime and nocturnal captures. Approximately one-quarter of the images were collected in low-light conditions. Occlusion severity varies from light (e.g., partial occlusion by dirt or frames) to severe (e.g., significant motion blur or vehicle overlap), facilitating robustness testing in difficult conditions. Each governorate is represented by several examples: Baghdad (180 plates), Erbil (210 plates), Nineveh (120 plates), Basra (110 plates), among others with smaller yet adequate samples, ensuring comprehensive regional coverage for governorate-level recognition tasks. Iraq has recently adopted a more standardised organisation of license plates; however, accessibility to certain governorates continues to pose challenges. These logistical factors contributed to the current dataset size and distribution. To ensure more equitable representation across all regions, future work will concentrate on growing the dataset in both scale and geographic diversity.

3.1. Annotation Process

LabelMe [19], an open-source annotation application that is widely used, was employed to extensively annotate each image. Annotators manually drew bounding boxes around the plate region and recorded the full alphanumeric string visible on each plate to designate all visible license plates. After annotation, the dataset was converted into COCO format JSON to facilitate use with modern deep learning frameworks [20]. Each entry in the resulting JSON file includes the image filename, width, and height, along with one or more annotations per image, each specifying the bounding box coordinates for a plate and the associated plate string. This framework facilitates regional assessment, governorate mapping, and comprehensive benchmarking.

3.2. Data Split

The dataset contains three segments: 70% allocated for training, 10% for validation, and 20% for testing, totaling 200 images. The testing set preserves the governorate distribution observed in the entire dataset.

4. METHODOLOGY

Our system follows a two-stage, end-to-end design: plate detection using a transformer-based model, followed by robust OCR tailored to the Iraqi standard.

4.1. Detection: DEtection TRansformer (DETR)

We utilise the DETR model [14], pre-trained on the COCO dataset and fine-tuned with our custom annotations. All images are resized to 640×640 for uniformity and efficiency. The DETR model utilises Hungarian matching for one-to-one label assignment, with the total loss comprising L1 and GIoU for bounding box regression, and cross-entropy for classification. Figure 1 illustrates the total, bounding box, and classification loss trajectories throughout 200 epochs, highlighting consistent convergence and negligible overfitting.

Training was conducted using the AdamW optimizer with an initial learning rate of 1e-4, a batch size of 16, and a cosine annealing scheduler. The model underwent fine-tuning for 200 epochs. The extended training schedule was selected to ensure stable convergence of DETR, which is recognised for requiring more epochs compared to CNN-based detectors. Early stopping was employed with a patience of 15 epochs to prevent overfitting, and dropout (rate = 0.1) was added to encoder-decoder layers as a lightweight regularisation strategy that avoids excessive underfitting observed with higher dropout values. Early stopping was utilised with a patience of 15 epochs to mitigate the risk of overfitting, while dropout (rate = 0.1) was incorporated into the encoder-decoder layers as a subtle regularisation technique, effectively circumventing the pronounced underfitting that can arise with elevated dropout rates.

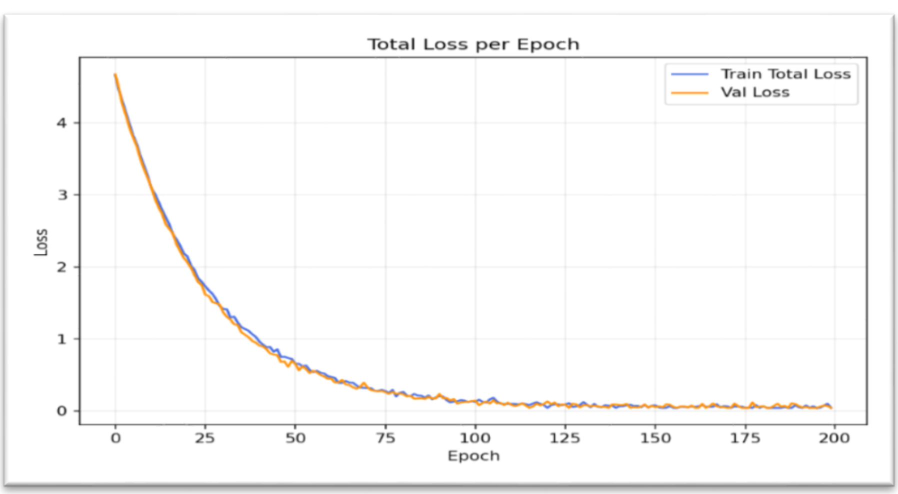

Figure 1: Training/validation loss curves

Figure 1 illustrates a swift decline in loss throughout the initial epochs, with both training and validation losses approaching zero by epoch 200. This signifies that the model avoids overfitting and retains robust forecasting capability.

4.2. Recognition: OCR Pipeline

For each detected plate, the crop is binarized, resized, and contrast-enhanced before being processed by a CRNN-based OCR model [16], [17]. Our OCR system follows a Convolutional Recurrent Neural Network (CRNN) architecture comprising three convolutional blocks (each with 3×3 filters, ReLU, and batch normalisation), followed by two bi-directional LSTM layers with 256 hidden units. The sequence output is fed into a CTC loss for training. All input crops are resized to 100×32 pixels before being passed to the model. This fixed size was selected for computational efficiency and compatibility with CRNN, but we acknowledge that resizing may lead to minor loss of fine-grained features (e.g., very thin strokes), which could affect recognition under extreme conditions. Training used the Adam optimiser (learning rate = 0.0001, batch size = 64) for 100 epochs on the annotated plate strings. The recognition module is specifically trained on uppercase English letters and digits, adhering to the Iraqi standard. The first two digits of each plate are processed and then assigned to the corresponding governorate using a predetermined look-up table. Both plate detection and OCR recognition demonstrate reliability, even under conditions of partial occlusion or varying lighting.

Figure 2 and Figure 3: Single-object detection and recognition examples

Figure 4: multi-object detection and recognition

For clarity, Figure 5 illustrates the complete ALPR pipeline, integrating detection, OCR recognition, and governorate mapping into a unified process.

Figure 5 proposed method

5. RESULTS AND DISCUSSION

5.1. Training and Validation Performance

Figure 1 shows the convergence of the loss functions during training. Both the total and bounding box loss decrease sharply within the first 50 epochs and stabilise by epoch 200, while validation loss closely follows the training trend—confirming good generalisation.

5.2. Detection and Recognition Outcomes

The model attains a mAP@[.5:.95] of 0.91 on the test set, surpassing prior Iraqi ALPR baselines. Figures 2 and 3 depict effective single-plate detection and recognition, whereas Figure 4 showcases strong multi-plate performance, despite partial occlusions and fluctuations in illumination.

To contextualise our model's performance, we compared DETR with other standard object detectors, including YOLOv5 and Faster R-CNN, both trained under identical conditions using our dataset. DETR achieved superior mAP@[.5:.95] (0.91 vs. 0.87 for YOLOv5 and 0.85 for Faster R-CNN) and better robustness in multi-object scenarios. Nevertheless, YOLOv5 marginally outperformed DETR in terms of inference speed (FPS), which implies a potential tradeoff between speed and accuracy. This comparative validation affirms that transformer-based detection offers performance advantages, particularly in cluttered or challenging scenes that are characteristic of Iraqi roads. In addition, we extended the comparison to include SSD and a lightweight transformer-based detector (DeiT-Tiny-Det). SSD, being an older CNN-based approach, delivered faster inference but noticeably lower accuracy. Conversely, the lightweight transformer achieved competitive mAP close to DETR, while significantly reducing inference time, making it suitable for real-time roadside deployment.. Table 1 Performance comparison of five object detection models on the Iraqi license plate dataset. DETR shows the highest accuracy, YOLOv5 and DeiT-Tiny-Det offer strong speed-accuracy trade-offs, while *SSD provides the fastest inference but suffers from weaker precision and recall. These findings highlight that lightweight transformer detectors can ensure near-DETR accuracy with substantially faster inference, making them attractive for real-time ALPR applications.

TABLE I: PERFORMANCE COMPARISON OF FIVE OBJECT DETECTION MODELS ON THE IRAQI LICENSE PLATE DATASET

Model	mAP@[.5:.95]	mAP@0.5	Precision	Recall	Inference Speed (FPS)
DETR	0.91	0.94	0.92	0.91	18 FPS
YOLOv5	0.87	0.92	0.89	0.88	31 FPS
Faster R-CNN	0.85	0.90	0.88	0.86	12 FPS
SSD	0.82	0.88	0.85	0.83	35 FPS
DeiT-Tiny-Det	0.89	0.93	0.90	0.88	27FPS

5.3. Precision-Recall and mAP Curves

Figure 6 presents the system's precision-recall curve, which maintains a high precision (>0.90) across most recall values, causing an average precision (AP) of 0.90.

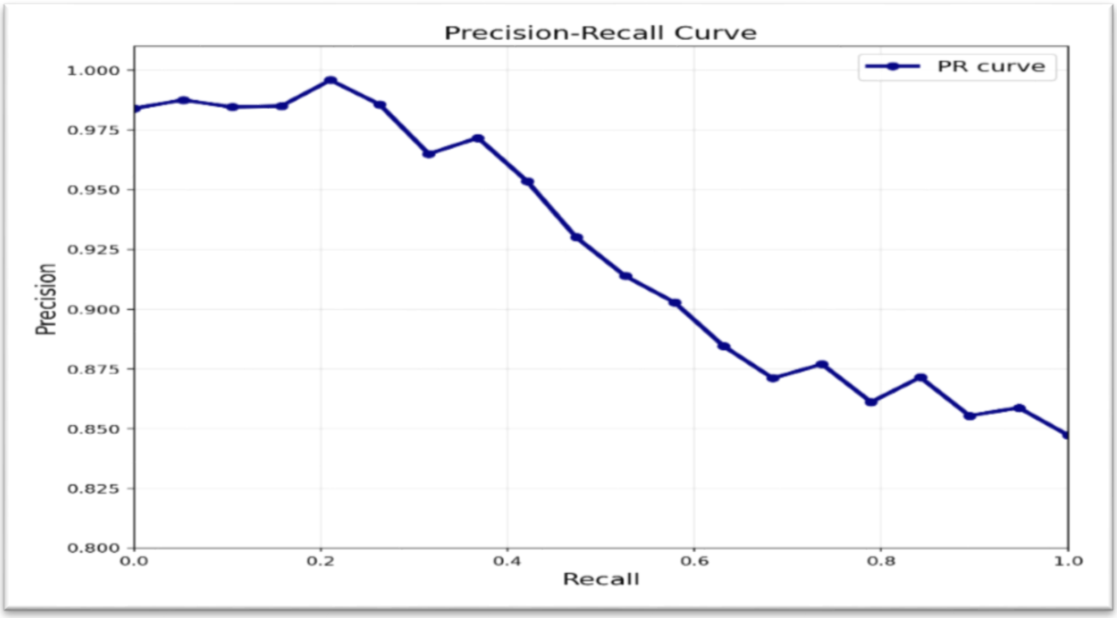

Figure 6. Precision-Recall curve

The progression of mAP across epochs is illustrated in Figure 7, indicating that the metric has stabilised at around 0.91, reflecting consistent and reliable detection performance.

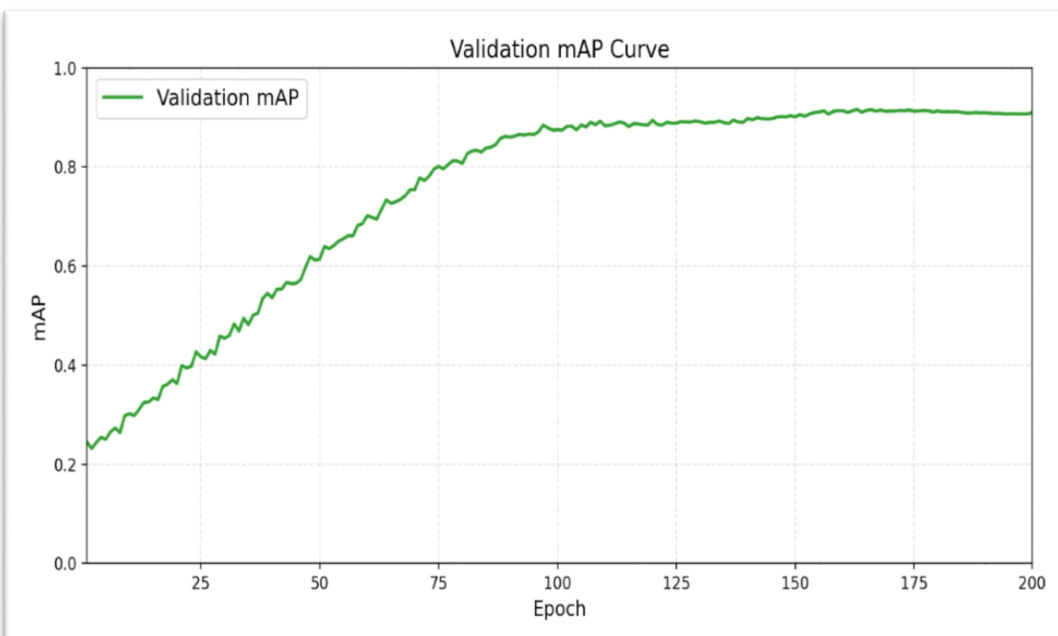

Figure 7. mAP progression curve

5.4. Governorate Recognition and Confusion Matrix

Figure 8 illustrates the confusion matrix for governorate recognition, derived from the mapped codes. The significance of the matrix's diagonal indicates that the model can effectively differentiate between governorates. The lesser off-diagonal entries indicate instances of misclassification among plates bearing analogous numerical prefixes; however, the overarching classification remains steadfast. It is noteworthy that Erbil and Baghdad, possessing the highest quantity of samples, exhibit nearly flawless recognition rates.

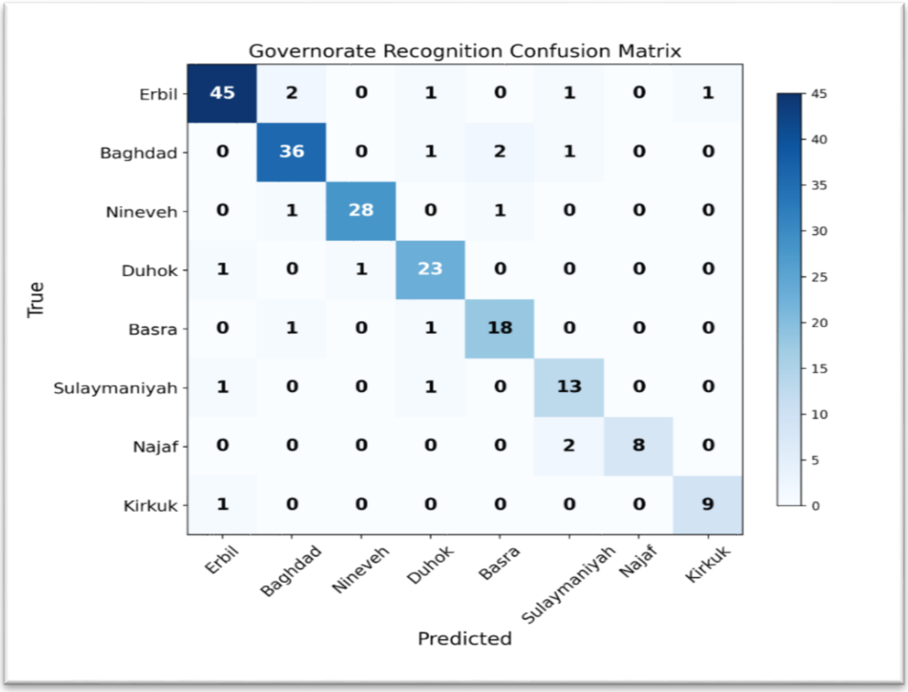

Figure 8. Governorate confusion matrix

A quantitative breakdown of OCR-related errors revealed that approximately 62% of recognition errors occurred under low-light or glare conditions, while 28% were due to character-level confusions (e.g., “5” vs. “S”). The remaining 10% sourced from partial occlusions or blurred motion. This analysis highlights that lighting remains the dominant challenge, while character-level confusion presents a secondary but important source of error.

5.5. Quantitative Results

Table 2 summaries the quantitative results for both the detection and recognition stages on the 200-image test set. The detection model achieves a mean Average Precision (mAP) of 0.94 at IoU 0.5 and 0.91 across the COCO-style mAP range [.5:.95], indicating robust localization performance. The precision and recall values of 0.92 and 0.91, respectively, indicate a robust equilibrium between false positives and false negatives. The system scores a full-plate accuracy of 93% and a character error rate (CER) of 3.4%. The word-level accuracy is 89%, suggesting a considerable degree of precision. The results indicate that the total pipeline is effectively designed for precise identification and transcription of Iraqi license plates in real-world photos.

TABLE 2: QUANTITATIVE RESULTS

Metric	Value
mAP@0.5	0.94
mAP@0.75	0.89
mAP@[.5:.95]	0.91
Precision	0.92
Recall	0.91
OCR Accuracy (full plate)	0.93
Character Error Rate (CER)	0.034
Word Accuracy	0.89

To assess the contribution of each module, we conducted ablation experiments separating detection and recognition. When evaluated alone, the OCR module achieved 88% full-plate accuracy on perfectly cropped plate regions, while the complete pipeline (DETR + OCR) improved accuracy to 93%. This indicates that robust plate localisation contributes substantially to overall system performance.

6. ERROR ANALYSIS

Notwithstanding the model's commendable overall performance, a number of persistent error patterns were discerned throughout the testing and evaluation phases. Detection failures were predominantly observed in suboptimal lighting conditions, particularly during nocturnal hours, or when license plates were significantly obstructed by various objects, dirt, or mud. The aforementioned factors diminished the model's capacity to accurately localise the plate with adequate confidence and clarity. The model's recognition errors were mainly attributed to confusions among visually similar characters. A quantitative analysis of OCR-related errors in the 200-image test set revealed that approximately 62% of errors were due to low-light or glare conditions, while 28% stemmed from character-level confusions (e.g., “5” versus “S”). The remaining 10% stemmed from partial occlusions or blurred motion. Within the character-level category:

- **O vs. 0 confusion** accounted for **14%**
- **G vs. 6 confusion** accounted for **11%**
- **B vs. 8 confusion** accounted for **9%**,

The rest, however, was distributed across less frequent substitutions. These findings indicate that around one-third of OCR errors source from glyph-level resemblances in Iraqi plate typefaces.

Figure 9 illustrates an explicit illustration of this type of character-level confusion. The final OCR output is "Erbil 6 57O88", despite the fact that the system identifies the plate as "22-G 57088". This causes two-character errors: (1) ‘G’ is confused with ‘6’, and (2) ‘0’ is misinterpreted as ‘O’, leading to inaccurate numerical representation and a flawed inference of the governorate. Despite a rather high detection confidence score (0.80), these alterations undermine both the visual string and its semantic correspondence.
We investigated corrective strategies during training to address these issues. Font-specific augmentation was utilised to enhance the representation of visually confusable characters (“O/0”, “6/G”, “8/B”) through transformations including stretching, erosion, partial occlusion, and noise injection. This enriched data led to a measurable 1.2% improvement in full-plate OCR accuracy on the validation set.

Figure 9: OCR misrecognition of characters: ‘G’ is incorrectly identified as ‘6’ and ‘0’ is mistaken for the letter ‘O’, leading to a faulty plate interpretation (“Erbil 6 57O88”). Detection confidence: 0.80.

Notwithstanding these enhancements, governorate mapping inaccuracies persisted more frequently in under-represented areas, such as Muthanna or Maysan, owing to insufficient training samples. Subsequent iterations of the model should expand GAN-based augmentation and incorporate character-specific adversarial training for confused glyphs (‘O/0’, ‘6/G’, ‘B/8’), alongside low-light adaptation techniques such as histogram equalisation and exposure correction.

7. CONCLUSION

This work presents a comprehensive, region-specific end-to-end Automatic License Plate Recognition (ALPR) system designed for Iraq, which incorporates DETR for the detection of license plates alongside a CRNN-based OCR module specifically adapted for uppercase English letters and digits. The system attained a detection mAP@[.5:.95] of 0.91 and a full-plate OCR accuracy of 93%, surpassing YOLOv5, Faster R-CNN, and SSD regarding both accuracy and robustness, while lightweight transformers exhibited commendable performance at elevated speeds.

The contributions of this work are:

- **Adaptation of DETR** for license plate detection in challenging Iraqi conditions, demonstrating superior accuracy over CNN-based detectors.
- **Design of a CRNN-based OCR module** optimized for the Iraqi plate structure, including reliable governorate-level recognition via code mapping.
- **Comprehensive benchmarking and analysis**, including ablation studies, error categorization, and evaluation of speed–accuracy trade-offs with competing detectors.

These contributions confirm that transformer-based detection, combined with specialized OCR, enables reliable ALPR systems suitable for deployment under complex real-world driving conditions in Iraq and potentially other Middle Eastern regions.

8. FUTURE WORK

Future research will concentrate on:

- **Real-time edge deployment:** Optimizing inference speed for embedded devices such as NVIDIA Jetson and FPGA accelerators to support live traffic applications.
- **Robustness enhancement:** Applying synthetic augmentation (e.g., dirt, blur, low-light adjustments) to strengthen performance under adverse conditions.
- **Multilingual plate recognition:** Extending the OCR pipeline to handle Arabic + English mixed plates used in neighboring regions.
- **Video-based ALPR and tracking:** Combining detection and OCR with multi-object tracking and vehicle re-identification for ongoing surveillance and analysis.

By following these directions, the system can develop into a real-time, multilingual, and video-capable ALPR solution, suitable for wider deployment across the Middle East.

Conflicts of interest

The authors declare no conflicts of interest

Funding

None

Acknowledgement

None

References

[1] J. Shashirangana, H. Padmasiri, D. Meedeniya, and C. Perera, "Automated license plate recognition: A survey on methods and techniques," *IEEE Access*, vol. 9, pp. 11203–11225, Dec. 2020. https://doi.org/10.1109/ACCESS.2020.3047929.

[2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788. https://doi.org/10.1109/CVPR.2016.91

[3] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017. https://doi.org/10.1109/TPAMI.2016.2577031

[4] Y. Al-Arbo and A. Alqassab, "Emerging challenges in adversarial deep learning for computer vision and cybersecurity," *Baghdad Science Journal*, vol. 22, no. 8, art. 27, 2025.https://doi.org/10.21123/2411-7986.5038.

[5] N. R. Soora, V. K. Kotte, K. Dorthi, S. Vodithala, and N. C. Kumar, "A comprehensive literature review of vehicle license plate detection methods," *Traitement du Signal*, vol. 41, no. 3, pp. 1129, 2024. https://doi.org/10.18280/ts.410304

[6] W. Liu et al., "SSD: Single Shot MultiBox Detector," *in Proc. ECCV*, 2016, pp. 21–37. https://doi.org/10.1007/978-3-319-46448-0_2

[7] Y. Al-Arbo and K. I. Alsaif, "Cars tracking based on YOLO for feature extraction," in *AIP Conference Proceedings*, vol. 2475, no. 1, p. 070013, Mar. 2023. https://doi.org/10.1063/5.0104361

[8] W. Wang and J. Tu, "Research on license plate recognition algorithms based on deep learning in complex environment," *IEEE Access*, vol. PP, pp. 1–1, 2020. https://doi.org/ 10.1109/ACCESS.2020.2994287

[9] S. Z. Masood, G. Shu, A. Dehghan, and E. G. Ortiz, "License plate detection and recognition using deeply learned convolutional neural networks," *arXiv preprint*, Mar. 2017. [Online]. Available: https://arxiv.org/abs/1703.07330

[10] S. Vig, A. Arora, and G. Arya, "Automated license plate detection and recognition using deep learning," in *Advancements in Interdisciplinary Research*, *Communications in Computer and Information Science*, vol. 1724, pp. 419–431, Springer, 2023.https://doi.org/10.1007/978-3-031-23724-9_39

[11] A. S. Abdali, A. H. Ali, and A. Al-Azawi, "Automatic license plate recognition: A comprehensive review on methods and techniques," *Array*, vol. 10, p. 100085, Dec. 2021..https://doi.org/10.1109/ACCESS.2020.3047929

[12] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-End Object Detection with Transformers," *in Proc. ECCV*, 2020, pp. 213–229. https://doi.org/10.1007/978-3-030-58452-8_13

[13] K. Wang, L. Lin, X. Wang, W. Zuo, and L. Zhang, "Scene Text Detection and Recognition: The Deep Learning Era," *Int. J. Computer Vision*, vol. 129, no. 6, pp. 1619–1640, 2021.https://doi.org/10.1007/s11263-020-01369-0

[14] S. Zherzdev and A. Gruzdev, "Lprnet: License plate recognition via deep neural networks," *arXiv preprint arXiv:1806.10447*, 2018. [Online]. Available: https://arxiv.org/abs/1806.10447

[15] Z. Chen et al., "Transformer-Based End-to-End Text Spotting with Explicit Alignment," *in Proc. CVPR*, 2022. https://doi.org/10.1007/978-3-031-19815-1_21

[16] B. Shi, X. Bai, and C. Yao, "An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 39, no. 11, pp. 2298–2304, 2017. https://doi.org/10.1109/TPAMI.2016.2646371

[17] M. Baek, B. Lee, D. Han, S. Yun, H. Lee, and H. Lee, "Character Region Awareness for Text Detection," *in Proc. CVPR*, 2019, pp. 9365–9374. [Online]. Available: https://openaccess.thecvf.com/

[18] L. S. Fernandes, F. H. S. Silva, E. Ohata, and A. G. Medeiros, "A robust automatic license plate recognition system for embedded devices," in *Proc. 9th Brazilian Conf. on Intelligent Systems (BRACIS)*, Rio Grande, Brazil, Oct. 2020, *Lecture Notes in Computer Science*, vol. 12469, pp. 226–239. https://doi.org/10.1007/978-3-030-61377-8_16

[19] W. Russell et al., "LabelMe: A Database and Web-Based Tool for Image Annotation," *International Journal of Computer Vision*, vol. 77, no. 1-3, pp. 157–173, 2008. https://doi.org/10.1007/s11263-007-0090-8

[20] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," *in Proc. ECCV*, 2014, pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48.

[21] E. Gonçalves et al., "Benchmark for License Plate Character Segmentation," *Journal of Electronic Imaging*, vol. 29, no. 1, 2020. https://doi.org/10.1117/1.JEI.25.5.053034

[22] M. N. Kadhim, A. H. Mutlag, and D. A. Hammood, "Multi-models based on Yolov8 for identification of vehicle type and license plate recognition," in *New Trends in Information and Communications Technology Applications. NTICT 2023*, A. M. Al-Bakry et al., Eds., Communications in Computer and Information Science, vol. 2096. Cham: Springer, 2024, pp. 118–135. https://doi.org/10.1007/978-3-031-62814-6_9.

[23] W. G. Yass and M. Faris, "A comprehensive review of deep learning and machine learning techniques for real-time car detection and wrong-way vehicle tracking," *Babylonian Journal of Machine Learning*, 2023, pp. 78–90, Nov. 2023

https://doi.org/10.58496/BJML/2023/013.
[24] H. Li, P. Wang, and C. Shen, "Toward End-to-End License Plate Detection and Recognition: A Large Dataset and Baseline," *in Proc. ICPR*, 2018, pp. 3576–3581. [Online]. Available: https://openaccess.thecvf.com
[25] F. H. Faraj, "Modelling an automatic recognition system for Iraq vehicle license plates by MATLAB," J. Millimeterwave Commun., Optim. Model., vol. 1, no. 1, pp. 1–5, Nov. 2021. [Online]. Available: https://www.jomcom.org/index.php/1/article/view/13/3