

Pengenalan Plat Nomor Kendaraan Real-Time pada Kondisi Gelap dan Hujan Menggunakan Deep Learning

Yanuangga Galahartlambang¹, Titik Khotiah², Zahrudin Fanani³, Mohammad bastian Dwiki Rahmat⁴

¹²³⁴Program Studi Teknologi Informasi, Institut Teknologi dan Bisnis Ahmad Dahlan Lamongan

Jl. KH. Ahmad Dahlan No.41, Lamongan 62218

¹yanuangga.id@gmail.com, ²titikaye@gmail.com, ³zahrudin.fanani17@gmail.com,

⁴bastiandwiki22@gmail.com

Abstract

Real-time vehicle license plate recognition is one of the important technologies in intelligent transportation systems, especially in challenging environmental conditions such as darkness and rain. This research aims to develop and test a license plate recognition method using a combination of YOLO v8 for license plate detection and the screen text recognition framework using TPS-ResNet-BiLSTM-Attn for text recognition. The dataset used is 3,000 license plate images divided into 70% for training, 15% for validation, and 15% for testing. This framework is designed to overcome visual challenges such as low lighting and rain interference through the stages of geometric transformation (TPS), feature extraction (ResNet), character sequence modeling (BiLSTM), and attention-based transcription mechanisms (Attention). The test results show different accuracies in various conditions, namely 78.6% in normal conditions, 82.4% in dark conditions, and 80.7% in rainy conditions. This shows that the proposed YOLO v8 object detection and screen text recognition using TPS-ResNet-BiLSTM-Attn text character recognition methods are capable of handling the challenges of license plate recognition under complex environmental conditions, thus having great potential to be applied in more reliable transportation surveillance systems.

Keywords : Number Plate Recognition, YOLO, Text Scene Recognition, OCR,

Abstrak

Pengenalan plat nomor kendaraan secara real-time merupakan salah satu teknologi penting dalam sistem transportasi cerdas, terutama pada kondisi lingkungan yang menantang seperti gelap dan hujan. Penelitian ini bertujuan untuk mengembangkan dan menguji metode pengenalan plat nomor menggunakan kombinasi YOLO v8 untuk melakukan deteksi plat nomor kendaraan dan framework screen text recognition TPS-ResNet-BiLSTM-Attn untuk pengenalan teks. Dataset yang digunakan berjumlah 3.000 gambar plat nomor yang dibagi menjadi 70% untuk pelatihan, 15% untuk validasi, dan 15% untuk pengujian. Framework ini dirancang untuk mengatasi tantangan visual seperti pencahayaan rendah dan gangguan akibat hujan melalui tahapan transformasi geometris (TPS), ekstraksi fitur (ResNet), pemodelan urutan karakter (BiLSTM), dan mekanisme transkripsi berbasis perhatian (Attention). Hasil pengujian menunjukkan akurasi yang berbeda pada berbagai kondisi, yaitu 78,6% pada kondisi normal, 82,4% pada kondisi gelap, dan 80,7% pada kondisi hujan. Hal ini menunjukkan bahwa metode deteksi objek YOLO v8 dan pengenalan karakter teks TPS-ResNet-BiLSTM-Attn yang diusulkan mampu menangani tantangan pengenalan plat nomor dalam kondisi lingkungan yang kompleks, sehingga memiliki potensi besar untuk diterapkan dalam sistem pengawasan transportasi yang lebih andal.

Kata kunci : Pengenalan Plat Nomor, YOLO, Text Scene Recognition, OCR

1. PENDAHULUAN

Pengenalan plat nomor kendaraan telah menjadi komponen penting dalam berbagai aplikasi modern, seperti manajemen lalu lintas, pengawasan keamanan, kontrol akses, dan pengelolaan parkir otomatis. Dengan meningkatnya volume kendaraan dan kebutuhan pengawasan yang efisien, sistem pengenalan otomatis plat nomor (Automatic License Plate Recognition/ALPR) menawarkan solusi yang cepat, akurat, dan andal untuk mengidentifikasi kendaraan tanpa intervensi manusia [1]. Selain itu, dengan perkembangan teknologi berbasis kecerdasan buatan, ALPR juga menjadi pondasi bagi implementasi *smart city* dan sistem transportasi cerdas [2].

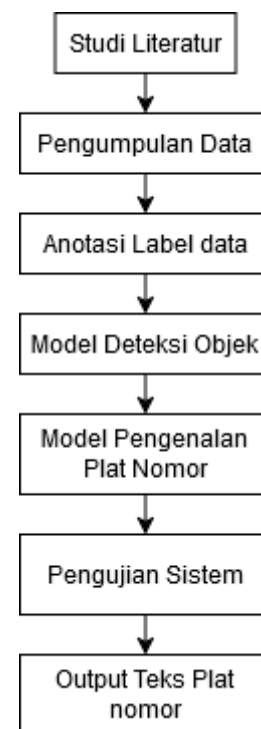
Dalam konteks ini, Komputer Vision telah memainkan peran signifikan dalam pengembangan sistem ALPR. Dengan bantuan algoritma deep learning, sistem pengenalan plat nomor kini dapat dioperasikan secara real-time bahkan dalam skenario yang menantang, seperti kondisi cahaya rendah, hujan deras, atau plat nomor yang terdistorsi. Salah satu model terkemuka dalam deteksi objek, YOLOv8 (You Only Look Once version 8), dikenal sebagai algoritma deep learning dengan akurasi tinggi [3] yang digunakan untuk mendeteksi objek dan efisiensinya dalam mendeteksi area plat nomor secara cepat dan presisi tinggi [4]. Setelah mendeteksi area plat nomor, langkah berikutnya adalah mengenali teks, di mana metode *TPS-ResNet-BiLSTM-Attn* (TRBA) atau *Scene Text Recognition* (STR) memberikan hasil yang andal dalam kondisi kompleks, seperti teks yang miring, buram, atau terdistorsi akibat hujan [5].

Meskipun banyak penelitian sebelumnya telah mengatasi beberapa tantangan teknis dalam ALPR, masih terdapat kesenjangan dalam menangani kondisi lingkungan yang sangat dinamis, seperti hujan deras yang menyebabkan noise optik atau cahaya rendah yang mengurangi visibilitas teks. Penelitian ini berkontribusi dalam pengembangan sistem ALPR dengan memanfaatkan kombinasi YOLOv8 untuk melakukan deteksi objek berupa area plat nomor dan *TPS-ResNet-BiLSTM-Attn* selanjutnya di singkat TRBA untuk pengenalan teks. Kombinasi ini dirancang untuk mengatasi berbagai skenario sulit secara end-to-end dan diimplementasikan secara real-time, menjadikannya solusi yang inovatif dan relevan dengan kebutuhan saat ini dengan menggunakan pendekatan berbasis scene text recognition (STR) untuk meningkatkan akurasi pengenalan teks dalam kondisi lingkungan yang tidak ideal [6].

Penelitian yang dilakukan oleh Vetriselvi et al [1] dan Al-batat et al [2] lebih berfokus pada teknik deteksi objek dan pengenalan karakter dengan menggunakan model seperti YOLO dan OCR, namun tidak secara spesifik menguji kemampuan model di kondisi gelap atau hujan secara bersamaan dalam konteks real-time, ataupun penelitian yang dilakukan oleh galahartlambang et al [7] pengenalan plat nomor kendaraan dilakukan dengan input pengujian menggunakan gambar. Penelitian ini memberikan perhatian lebih pada kedua faktor kondisi lingkungan secara simultan, dan menggunakan kombinasi YOLO v8 dan TPS-ResNet-BiLSTM-Attn, yang dapat mengatasi tantangan visual yang lebih besar dibandingkan dengan model yang hanya mengandalkan satu faktor.

2. METODOLOGI PENELITIAN

Penelitian ini menggunakan metodologi kuantitatif, dengan menggunakan data yang dapat diukur dan menganalisis akurasi dari sistem pengenalan plat nomor. Pada gambar 1 menunjukkan tahapan penelitian yang dilakukan untuk memperjelas alur dari kegiatan yang dilakukan



Gambar 1 Metode Penelitian

2.1. Pengumpulan Data

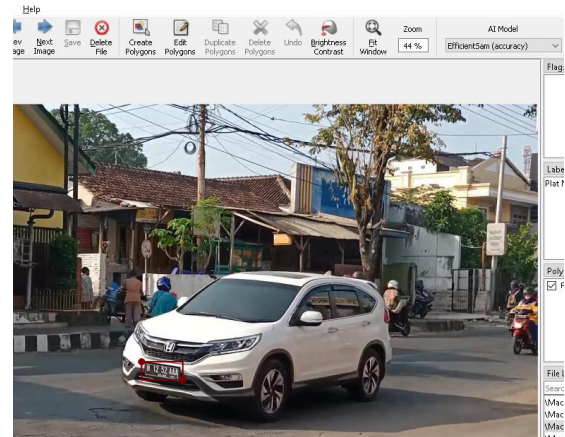
Kegiatan pengumpulan data bertujuan untuk mendapatkan sekumpulan dataset kendaraan beserta plat nomor kendaraan indonesia dalam

berbagai kondisi cahaya dan cuaca. Variasi kondisi perekaman ini bertujuan untuk melatih model deep learning agar mampu mengenali plat nomor secara akurat dalam berbagai skenario, meningkatkan kemampuan generalisasi dan robustness sistem terhadap kondisi dunia nyata[8]. Dataset penelitian ini didapatkan dari rekaman video lalu lintas jalan raya pada siang hari, sore hari dan malam hari. Selain itu dataset penelitian juga didapatkan dari CCTV pintu gerbang parkir dalam perguruan tinggi Institut Teknologi dan Bisnis Ahmad Dahlan Lamongan[7] dan juga video dari Youtube. Jenis kendaraan dalam kumpulan data ini meliputi sepeda motor dan mobil dengan pelat nomor hitam dengan huruf putih, dan pelat nomor putih dengan huruf hitam. Video direkam dengan resolusi full HD (1920 x 1080) dan frame rate 30 fps.

Setelah data terkumpul dan data layak untuk digunakan, dilakukan proses ekstraksi frame video dalam bentuk gambar. Tujuannya agar dapat dilakukan anotasi label gambar. Dataset tersebut selanjutnya dibagi menjadi tiga bagian: training set, validation set, dan test set. Pembagian dataset dilakukan dengan komposisi 70% untuk training set, 15% untuk validation set, dan 15% untuk test set[9].

2.2. Anotasi Label Data

Anotasi label data adalah proses pemberian label atau tag pada gambar untuk mengklasifikasikan atau mengidentifikasi objek yang ada di dalamnya. Proses ini sangat penting dalam berbagai bidang computer vision seperti pengenalan objek, klasifikasi gambar, dan deteksi objek. Proses anotasi label gambar menggunakan labelme. Setiap frame gambar yang terpilih diberi anotasi label menggunakan Labelme. Bounding box digunakan untuk melokalisasi plat nomor pada gambar, dan transkripsi teks plat nomor dicatat sebagai label yang berkorespondensi dengan bounding box tersebut. Anotasi yang akurat sangat penting karena berfungsi sebagai *ground truth* bagi model deep learning selama proses pelatihan, sehingga dapat meningkatkan efisiensi dan akurasi dalam pelabelan data[10][11].



Gambar 2. Anotasi Label Gambar

2.3. Model Deteksi Objek

Pada kegiatan pembuatan model deteksi objek, dilakukan proses training model dengan memuat bobot awal (pre-trained weights) untuk mempercepat konvergensi dan meningkatkan akurasi. Hiperparameter penting yang perlu di konfigurasi meliputi :

- Learning rate : mengontrol kecepatan pembelajaran model
- Batch size : Jumlah sampel yang diproses dalam satu iterasi
- Epoch : Jumlah siklus penuh melalui dataset selama pelatihan

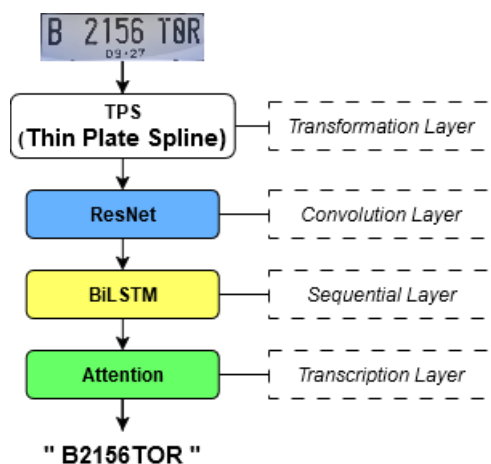
Selama proses training, model belajar mengenali pola dan fitur yang relevan dari dataset. Proses ini dapat dipantau dengan melihat matrix loss, presisi dan recall pada data validasi[12].

Validasi dilakukan selama proses Training untuk memantau kinerja model dan mencegah overfitting. Jika kinerja pada data training jauh lebih baik daripada data validasi, ini bisa menjadi indikasi overfitting. Setelah pelatihan selesai, model dievaluasi menggunakan data uji yang belum pernah dilihat sebelumnya untuk menilai kinerja dari system[13].

2.4. Model Pengenalan Plat Nomor

Dalam pengembangan aplikasi pengenalan plat nomor kendaraan otomatis tantangan seperti distorsi perspektif, pencahayaan buruk, dan noise visual membuat tugas ini menjadi kompleks[14]. Scene Text Recognition (STR) telah muncul sebagai metode yang andal untuk tugas pengenalan teks berbasis gambar[15]. Tahapan STR yang digunakan pada penelitian ini adalah :

- a. Transformation Layer: Thin Plate Spline (TPS)
 TPS digunakan untuk menangani distorsi geometris. Metode ini telah terbukti efektif dalam mengoreksi perspektif gambar. Gambar input yang tidak terstruktur diubah menjadi gambar yang lebih seragam, mempermudah tahap ekstraksi fitur
- b. Convolutional Layer: ResNet
 ResNet digunakan untuk ekstraksi fitur visual dari gambar plat nomor. Residual Network(3), memungkinkan pelatihan jaringan yang lebih dalam tanpa kehilangan informasi. Dalam penelitian ini, ResNet mendeteksi fitur seperti tepi dan pola huruf dari gambar.
- c. Sequential Layer : BiLSTM
 BiLSTM (Bidirectional Long Short-Term Memory) memodelkan hubungan sekuensial antar fitur visual. BiLSTM menangani urutan data dengan cara yang lebih efektif dengan mempertimbangkan hubungan temporal ke depan dan ke belakang[16].
- d. Transcription Layer: Attention
 Attention Layer digunakan untuk memilih informasi relevan dari vektor fitur, sesuai dengan pendekatan yang diusulkan oleh Vaswani et al. [17]. Teknik ini memastikan transkripsi teks yang akurat, meskipun gambar memiliki noise atau huruf yang tumpang tindih[18].



Gambar 3. Framework TPS-ResNet-BiLSTM-Attn

3. HASIL DAN PEMBAHASAN

3.1 Dataset Pelatihan Model Deteksi Objek

Dataset yang dipakai pada penelitian ini terdiri dari 3000 citra dengan berbagai objek. Dataset dikumpulkan dari berbagai sumber dan telah diberi anotasi label dalam bentuk *bounding box* yang menunjukkan lokasi dan kelas objek. Selanjutnya dibagi menjadi proporsi 10% data training, 15% validasi dan 15% testing. Training set yang cukup besar memungkinkan model untuk mempelajari pola dan fitur dari data dengan baik. Validasi set yang memadai membantu dalam memilih hyperparameter yang tepat dan mencegah overfitting. Testing set yang cukup memungkinkan evaluasi performa model pada data yang belum pernah dilihat sebelumnya sehingga menghasilkan estimasi performa yang lebih baik. Tabel 1 merupakan pembagian dataset



(a) Gelap



(b) Hujan

Gambar 4. Pengujian deteksi objek

yang digunakan pada penelitian ini.

Tabel 1. Dataset Model Deteksi Objek

Nama Dataset	Cuaca/Cahaya	Jumlah
Training	Normal	850
	Gelap	625
	Hujan	625
Validasi	Normal	200
	Gelap	125
	Hujan	125
Testing	Normal	200
	Gelap	125
	Hujan	125

3.2 Hasil Pelatihan Pengujian Model Deteksi Plat Nomor

Hasil training dataset menggunakan YOLO v8 dapat dilihat pada tabel 2.

Tabel 2. Hasil Pelatihan model YOLO v8

Nama Dataset	Precision	Recall	F1-Score	mAP
Normal	0,948	0,936	0,937	0,961
Gelap	0,963	0,941	0,953	0,973
Hujan	0,957	0,926	0,936	0,968

Pada table 1 dan table 2 menampilkan hasil evaluasi performa pelatihan dan pengujian model menggunakan Yolo v8. Kolom precision menunjukkan tingkat ketepatan pada model dalam melakukan mendeteksi objek, recall menunjukkan kemampuan model untuk mendeteksi semua objek yang ada, F1-Score adalah rata-rata harmonic antara precision dengan recall, sedangkan mAP (Mean Average Precision) mengukur akurasi model dalam mendeteksi objek di semua kelas. Semakin tinggi nilai-nilai pada kolom tersebut semakin baik pula performa model dalam mendeteksi objek.

Hasil dari pengujian model untuk deteksi objek dapat dilihat pada tabel 2.

Tabel 2. Hasil Pengujian Model YOLO v8

Nama Dataset	Precision	Recall	F1-Score	mAP
Normal	0,918	0,813	0,842	0,846
Gelap	0,937	0,837	0,825	0,837
Hujan	0,914	0,782	0,841	0,807

Hasil pengujian deteksi objek dapat dilihat pada gambar 4. Gambar 4a deteksi plat nomor pada saat kondisi gelap, gambar 4b deteksi plat nomor pada saat kondisi hujan.

3.3 Hasil Pelatihan Pengujian Pengenalan Teks Karakter

Model training dataset plat nomor menggunakan framework TRBA berguna untuk mengenali karakter teks dari potongan (bounding box) plat nomor. Potongan plat nomor tersebut merupakan sekumpulan dataset yang dipotong berdasarkan hasil prediksi inference dalam bentuk pemberian bounding box pada YOLO v8. Tabel 3 hasil pelatihan TRBA

Tabel 3. Hasil Training framework TRBA

Nama Dataset	Epoch	Accuracy
Normal	100	91,4 %
Gelap	100	87,2 %
Hujan	100	85,3 %

Tabel 4 merupakan hasil testing model menggunakan framework TRBA, pada table 4 akurasi testing sedikit lebih rendah dibandingkan akurasi training framework TRBA karena tantangan pada data uji, seperti noise atau variasi lain yang tidak ada dalam data training.

Tabel 4. Hasil Testing framework TRBA

Nama Dataset	Epoch	Accuracy
Normal	100	78,6 %
Gelap	100	82,4 %
Hujan	100	80,7 %

3.4 Hasil Pengujian Sistem

Hasil pengujian system dapat dilihat pada gambar 5, yang menguraikan tentang tahapan framework TRBA dalam melakukan ekstraksi karakter teks dari plat nomor yang telah dideteksi sebelumnya.

Gambar 5a. gambar hasil deteksi plat nomor kendaraan, gambar 5b hasil potongan gambar plat nomor yang akan diproses TRBA, gambar 5c merupakan tahapan TPS (*Transformation*) untuk memperbaiki distorsi geometris, gambar 5d merupakan tahapan ResNet(*Feature Extraction*) ekstraksi fitur visual utama yang ditampilkan dalam bentuk deteksi tepi untuk mengenali pola karakter, gambar 5e adalah tahapan BiLSTM (*sequential layer*) proses pemodelan urutan karakter untuk memahami konteks teks, dan gambar 5f adalah tahapan Attention(*Transcription Layer*) yang memberikan prediksi akhir pada area penting menggunakan Attention, memberikan hasil lebih jelas untuk pengenalan teks. Output teks terdapat pada gambar 5g.



(a) Deteksi kondisi gelap



(a) Deteksi kondisi hujan



(b) Potongan kondisi gelap



(b) Potongan kondisi hujan



(c) TPS kondisi gelap



(c) TPS kondisi hujan



Gambar 5. Hasil Pengujian Sistem

4. Kesimpulan dan Saran

Berdasarkan hasil penelitian dan pengujian menggunakan framework TPS-ResNet-BiLSTM-Attn, diperoleh kesimpulan:

1. Akurasi Pengenalan:

Pada kondisi normal, akurasi pengenalan mencapai 78,6%, yang menunjukkan bahwa framework dapat mengenali teks plat nomor dengan cukup baik. Pada kondisi gelap, akurasi meningkat menjadi 82,4%, menandakan bahwa mekanisme Attention bekerja optimal dalam fokus pada area penting meskipun dalam pencahayaan rendah. Pada kondisi hujan, akurasi adalah 80,7%, menunjukkan bahwa TPS efektif dalam mengurangi distorsi akibat tetesan air, namun masih ada tantangan seperti noise visual

2. Efektivitas Framework:

TPS membantu mengoreksi distorsi geometris, sehingga model dapat membaca karakter dengan lebih presisi. ResNet berhasil mengekstraksi fitur utama meskipun dalam kondisi lingkungan yang sulit. BiLSTM

memahami urutan karakter, sehingga mampu mengenali teks meskipun ada noise atau ketidakjelasan sebagian karakter. Attention Mechanism memastikan fokus model pada karakter penting di gambar, yang berkontribusi besar pada akurasi pengenalan.

Saran pada penelitian ini yaitu:

1. Peningkatan Akurasi:

Perluasan dataset dengan variasi lebih banyak pada kondisi lingkungan seperti hujan deras, kabut, dan sudut pengambilan gambar yang ekstrem untuk melatih model lebih robust. Gunakan teknik data augmentation seperti rotasi, blur, atau penambahan noise untuk meningkatkan kemampuan generalisasi model.

2. Optimalisasi Framework:

Lakukan penyederhanaan arsitektur ResNet untuk mengurangi beban komputasi, terutama jika model akan diterapkan pada perangkat dengan sumber daya terbatas. Terapkan teknik transfer learning dengan memanfaatkan model pretrained yang lebih spesifik untuk tugas pengenalan teks dalam kondisi lingkungan ekstrem.

3. Evaluasi Lebih Lanjut:

Uji model pada dataset dengan format plat nomor dari negara atau wilayah lain untuk memastikan generalisasi. Analisis kesalahan pengenalan untuk memahami karakter atau kondisi yang sering menyebabkan prediksi salah, kemudian lakukan fine-tuning model.

5. UCAPAN TERIMA KASIH

Penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada Ristek Dikti dari Direktorat Riset dan Pengabdian kepada Masyarakat (DRPM) atas dana hibah yang diberikan untuk mendukung penelitian ini. Bantuan pendanaan ini sangat membantu dalam pelaksanaan dan pengembangan penelitian yang bertujuan untuk memberikan kontribusi signifikan dalam bidang teknologi pengenalan plat nomor kendaraan. Kami berharap hasil penelitian ini dapat memberikan manfaat yang luas bagi perkembangan ilmu pengetahuan dan teknologi di Indonesia. Serta kampus ITBAD Lamongan bidang Lembaga Penelitian dan Pengabdian kepada Masyarakat. Dalam hal memberikan kesempatan dan dukungan agar kegiatan penelitian ini dapat

berjalan dengan lancar. Dan saya ucapkan terima kasih kepada seluruh rekan-rekan dan mahasiswa atas kerja sama dan kerja samanya dalam melakukan proses penelitian ini..

Daftar Pustaka:

- [1] T. Vetrisevi et al., "Deep learning based license plate number recognition for smart cities," *Comput. Mater. Contin.*, vol. 70, no. 1, pp. 2049–2064, 2021, doi: 10.32604/cmc.2022.020110.
- [2] R. Al-batat, A. Angelopoulou, S. Premkumar, J. Hemanth, and E. Kapetanios, "An End-to-End Automated License Plate Recognition System Using YOLO Based Vehicle and License Plate Detection with Vehicle Classification," *Sensors*, vol. 22, no. 23, 2022, doi: 10.3390/s22239477.
- [3] N. Giarsyani, "Komparasi Algoritma Machine Learning dan Deep Learning untuk Named Entity Recognition: Studi Kasus Data Kebencanaan," *Indones. J. Appl. Informatics*, vol. 4, no. 2, p. 138, 2020, doi: 10.20961/ijai.v4i2.41317.
- [4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 779–788, 2016, doi: 10.1109/CVPR.2016.91.
- [5] R. Ronen, S. Tsiper, O. Anschel, I. Lavi, A. Markovitz, and R. Manmatha, "GLASS: Global to Local Attention for Scene-Text Spotting," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13688 LNCS, pp. 249–266, 2022, doi: 10.1007/978-3-031-19815-1_15.
- [6] S. Rajebi, S. Pedrammehr, and R. Mohajerpoor, "A License Plate Recognition System with Robustness against Adverse Environmental Conditions Using Hopfield's Neural Network," *Axioms*, vol. 12, no. 5, 2023, doi: 10.3390/axioms12050424.
- [7] Y. Galahartlambang, T. Khotiah, Zahrudin Fanani, and Afifatul Aprilia Yani Solekhah, "Deteksi Plat Nomor Kendaraan Otomatis Dengan Convolutional Neural Network Dan Ocr Pada Tempat Parkir Itb Ahmad Dahlan Lamongan," *J. Manaj. Inform. dan Sist. Inf.*, vol. 6, no. 2, pp. 114–122, 2023, doi: 10.36595/misi.v6i2.754.
- [8] R. Laroca et al., "A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector," *Proc. Int. Jt. Conf. Neural Networks*, vol. 2018-July, 2018, doi: 10.1109/IJCNN.2018.8489629.
- [9] Y. Liu et al., "An Improved Tuna-YOLO Model Based on YOLO v3 for Real-Time Tuna Detection Considering Lightweight Deployment," *J. Mar. Sci. Eng.*, vol. 11, no. 3, 2023, doi: 10.3390/jmse11030542.
- [10] F. Demrozi, C. Turetta, F. Al Machot, G. Pravadelli, and P. H. Kindt, "A Comprehensive Review of Automated Data Annotation Techniques in Human Activity Recognition," pp. 1–37, 2023, [Online]. Available: <http://arxiv.org/abs/2307.05988>
- [11] L. Lakhdar and M. S. Allili, A survey on machine learning techniques for semantic image and video annotations, no. Iceis. Atlantis Press International BV, 2024. doi: 10.2991/978-94-6463-496-9_14.
- [12] M. Yaseen, "What is YOLOv8: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector," vol. 8, pp. 1–10, 2024, [Online]. Available: <http://arxiv.org/abs/2408.15857>
- [13] A.-R. A. Gamani, I. Arhin, and A. K. Asamoah, "Performance Evaluation of YOLOv8 Model Configurations, for Instance Segmentation of Strawberry Fruit Development Stages in an Open Field Environment," pp. 1–15, 2024.
- [14] J. Baek et al., "What is wrong with scene text recognition model comparisons? dataset and model analysis," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2019-Octob, pp. 4714–4722, 2019, doi: 10.1109/ICCV.2019.00481.
- [15] 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, doi: 10.1109/TPAMI.2016.2646371.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
- [17] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks," *ACM Int. Conf. Proceeding Ser.*, vol. 148, pp. 369–376, 2006, doi: 10.1145/1143844.1143891.
- [18] A. Vaswani et al., "Attention is all you need," *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Nips, pp. 5999–6009, 2017.