

Análise do desempenho de modelos de *optical character recognition* de propósito geral em placas veiculares

Luís Gustavo Werle Tozevich¹, Gabriel Machado Lunardi¹,
Thiago Lopes Trugillo da Silveira¹, Adriano Quilião de Oliveira²

¹Centro de Tecnologia, Universidade Federal de Santa Maria

²Coordenadoria Acadêmica de Cachoeira do Sul, Universidade Federal de Santa Maria

lgtozevich@inf.ufsm.br,

{gabriel.lunardi, thiago.silveira, adriano.q.oliveira}@ufsm.br

Abstract. *Automatic License Plate Recognition (ALPR) is a crucial technology for various fields, such as transportation systems, where specialized development can entail significant costs and complexity. In this context, general-purpose Optical Character Recognition (OCR) engines emerge as a potential alternative. This work conducted a quantitative evaluation of three generic OCRs (EasyOCR, Pytesseract, Keras-OCR). The results reveal a performance discrepancy: state-of-the-art models reach 95.6% accuracy, while the best-performing generic OCR achieved only 16.5%. Such findings demonstrate that non-specialized OCR tools lack suitability for real-world ALPR applications.*

Resumo. *O Reconhecimento Automático de Placas Veiculares (ALPR) é uma tecnologia importante para diversas áreas, como sistemas de transporte, cujo desenvolvimento pode apresentar altos custos e complexidade. Nesse contexto, modelos de Optical Character Recognition (OCR) de propósito geral emergem como uma potencial alternativa. Este trabalho realizou uma avaliação quantitativa de três OCRs genéricos (EasyOCR, Pytesseract, Keras-OCR). Os resultados revelam uma discrepância de desempenho: modelos estado da arte atingem 95,6% de acurácia, enquanto o OCR genérico de melhor performance alcançou somente 16,5%. Tais achados demonstram que ferramentas de OCR não especializadas carecem de adequação para aplicações reais em ALPR.*

1. Introdução

O Reconhecimento Automático de Placas Veiculares (ALPR, *Automatic License Plate Recognition*) constitui um componente tecnológico essencial para sistemas de transporte inteligentes e aplicações de segurança pública [Du et al. 2012]. Trabalhos recentes demonstram que arquiteturas de *deep learning*, customizadas para esta tarefa, alcançam acurácia próxima ao estado da arte (SOTA). Modelos como Redes Neurais Convolucionais Recorrentes (CRNN) [Shi et al. 2016], framework TRBA [Baek et al. 2019] e, mais recentemente, arquiteturas baseadas em *Transformers* como o ViTSTR [Atienza 2021], reportam taxas de acerto superiores a 95% em *benchmarks*. Contudo, o desenvolvimento, treinamento e implementação de tais sistemas especializados demandam elevado custo computacional, conhecimento técnico e grandes conjuntos de dados, representando barreiras para sua adoção em larga escala [Bulan et al. 2015].

Todavia, modelos de *Optical Character Recognition* (OCR) de propósito geral oferecem uma alternativa de baixo custo e rápida implementação, devido à sua natureza *off-the-shelf*. Contudo, a viabilidade de sua aplicação direta em ALPR é incerta, uma vez que placas veiculares impõem desafios como a coexistência de múltiplos padrões, formatos, variações acentuadas de iluminação e ângulo de captura.

Neste trabalho, busca-se analisar o potencial e as limitações de modelos OCR de propósito geral para o reconhecimento de placas veiculares. Para tanto, analisa-se três motores OCR: EasyOCR [Jaided AI 2024], Keras-OCR [Morales 2020] e PyTesseract [Tesseract-OCR. 2024], no *dataset* brasileiro RodoSol-ALPR [Laroca et al. 2022].

2. Metodologia

A avaliação foi realizada no conjunto de dados RodoSol-ALPR, composto por 20.000 imagens capturadas em praças de pedágio da concessionária Rodovia do Sol na rodovia ES-060 [Laroca et al. 2022]. Esse *dataset* abrange diferentes tipos de veículos, bem como dois padrões de placas: o padrão brasileiro e o Mercosul. As condições de captura incluem cenários diurnos, noturnos, em dias chuvosos e em diferentes distâncias da câmera. Na avaliação dos métodos, seguiu-se a divisão proposta pelos autores, utilizando 8000 imagens no conjunto de teste. A Figura 1 apresenta exemplos de imagens do *dataset*.



Figura 1. Exemplos do *dataset* RodoSol

A escolha dos modelos OCR considerou a ampla utilização na comunidade e facilidade de uso. O PyTesseract é um wrapper para o motor Tesseract, que utiliza redes neurais recorrentes do tipo LSTM (*Long Short-Term Memory*); o EasyOCR combina um detector de texto (como o CRAFT) com um reconhecedor baseado em CRNN (*Convolutional Recurrent Neural Network*); e o Keras-OCR oferece um pipeline semelhante, também fundamentado em arquiteturas de CRNNs.

A fim de isolar o desempenho da etapa de reconhecimento, os três motores foram aplicados diretamente sobre as imagens das placas previamente recortadas a partir das coordenadas do *ground truth* presente no *dataset*, evitando que falhas na detecção influenciassem os resultados. Todos os métodos foram testados em sua configuração padrão, sem treinamento adicional ou *fine tuning*, e avaliados por duas métricas principais: acurácia de placa completa (APC) e taxa de reconhecimento de caracteres (CRR).

3. Resultados

Conforme detalhado na Tabela 1, há uma disparidade de desempenho entre os motores OCR de propósito geral e as abordagens especializadas. Os resultados dos OCRs (EasyOCR, Keras-OCR e PyTesseract) foram computados neste trabalho, enquanto os dos modelos SOTA (ViTSTR-Base [Atienza 2021], TRBA [Baek et al. 2019] e CRNN [Shi et al. 2016]) foram extraídos de [Laroca et al. 2023], obtidos sob o protocolo

traditional-split com avaliação no conjunto RodoSol-ALPR, no qual os modelos são treinados na união dos conjuntos de treinamento e avaliados nos dados do conjunto de teste do modelo analisado. O melhor OCR genérico testado (EasyOCR) atingiu 16,5% de acurácia, contrastando com os 95,6% do modelo ViTSTR-Base. Essa diferença evidencia a inadequação de ferramentas não especializadas para a tarefa de ALPR quando comparadas a modelos treinados especificamente para este domínio.

Tabela 1. Métricas de desempenho por modelo.

Métrica	ViTSTR-Base	TRBA	CRNN	EasyOCR	Keras-OCR	PyTesseract
APC (%)	95.6	80.6	73.6	16.5	13.4	14.7
CRR (%)	—	—	—	66.6	63.0	27.8

Analisando os resultados de CRR, percebe-se que os modelos de propósito geral apresentam limitações no reconhecimento de placas completas. Embora o EasyOCR e o Keras-OCR apresentem um CRR de 66,6% e 63,0%, respectivamente. No entanto, como detalhado na Figura 2, esse desempenho não é uniforme, sendo inferior para placas de motocicleta, que possuem variações geométricas e duas linhas. Além da dificuldade com a estrutura das placas, a análise dos erros revela que as falhas mais comuns ocorrem na troca de caracteres visualmente similares. No modelo EasyOCR, a letra 'O' foi incorretamente classificada como o número '0' em 346 ocasiões, enquanto no Keras-OCR, a letra 'B' foi confundida com o número '8' 196 vezes. De forma ainda mais acentuada, o PyTesseract classificou erroneamente a letra 'I' como o número '1' em 146 instâncias, superando o número de acertos para o próprio caractere 'I' (91). Tais trocas poderiam ser corrigidas com heurísticas [Montazzolli and Jung 2017], mas optou-se por não aplicá-las para avaliar o desempenho bruto dos modelos OCR. Essas confusões sistemáticas comprometem a integridade da leitura e justificam o baixo desempenho geral.

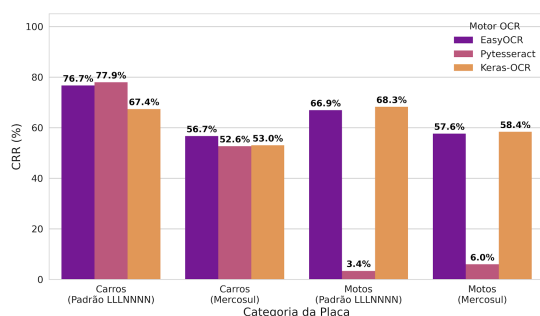


Figura 2. CRR por categoria e tipo de placa.

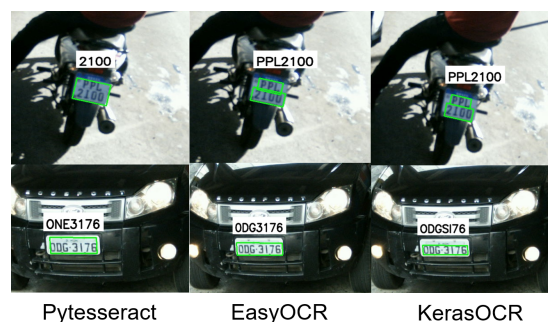


Figura 3. Exemplos das previsões.

As falhas mencionadas, ilustradas na Figura 3, reforçam que a ausência de um treinamento focado resulta em baixa capacidade de generalização. Conclui-se, portanto, que os OCRs de propósito geral analisados são inadequados para aplicações reais de ALPR.

4. Conclusão

Os resultados obtidos evidenciam as limitações do uso de sistemas OCR de propósito geral no contexto de ALPR. Apesar de sua facilidade de implementação e caráter *plug-and-play*, os modelos avaliados apresentaram desempenho inferior às soluções especiali-

zadas. Desta forma, os resultados corroboram com a necessidade de arquiteturas e procedimentos específicos para o domínio de placas veiculares, incluindo conjuntos anotados representativos e estratégias de adaptação.

Como trabalhos futuros, propõe-se: (i) avaliar o impacto de técnicas de pré-processamento de imagem sobre a acurácia de reconhecimento; (ii) empregar *oriented bounding boxes* para melhorar a localização de placas rotacionadas; (iii) medir o desempenho *end-to-end*, integrando um detector de placas ao *pipeline* de OCR; e (iv) investigar o desempenho de modelos OCR mais recentes, a fim de verificar se avanços recentes podem reduzir a diferença de acurácia observada. Essas estratégias podem aumentar as taxas de acerto e viabilizar o uso destas soluções de baixo custo em aplicações reais.

Referências

- Atienza, R. (2021). Vision transformer for fast and efficient scene text recognition. In *International conference on document analysis and recognition*, pages 319–334. Springer.
- Baek, J., Kim, G., Lee, J., Park, S., Han, D., Yun, S., Oh, S. J., and Lee, H. (2019). What is wrong with scene text recognition model comparisons? dataset and model analysis. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 4715–4723.
- Bulan, O., Kozitsky, V., and Burry, A. (2015). Towards annotation free license plate recognition. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, pages 1495–1499. IEEE.
- Du, S., Ibrahim, M., Shehata, M., and Badawy, W. (2012). Automatic license plate recognition (alpr): A state-of-the-art review. *IEEE Transactions on circuits and systems for video technology*, 23(2):311–325.
- Jaied AI (2024). Easyocr: Ready-to-use ocr with 80+ supported languages. <https://github.com/JaiedAI/EasyOCR>. Acessado em 24 de agosto de 2025.
- Laroca, R., Cardoso, E. V., Lucio, D. R., Estevam, V., and Menotti, D. (2022). On the cross-dataset generalization in license plate recognition. In *International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 166–178.
- Laroca, R., Zanlorensi, L. A., Estevam, V., Minetto, R., and Menotti, D. (2023). Leveraging model fusion for improved license plate recognition. In *Iberoamerican Congress on Pattern Recognition (CIARP)*, pages 60–75.
- Montazzolli, S. and Jung, C. (2017). Real-time brazilian license plate detection and recognition using deep convolutional neural networks. In *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 55–62.
- Morales, F. (2020). Keras-ocr: End-to-end trainable ocr pipeline. <https://github.com/faustomorales/keras-ocr>. Acessado em 24 de agosto de 2025.
- Shi, B., Bai, X., and Yao, C. (2016). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE transactions on pattern analysis and machine intelligence*, 39(11):2298–2304.
- Tesseract-OCR. (2024). Tesseract open source ocr engine. <https://github.com/tesseract-ocr/tesseract>. Acessado em 24 de agosto de 2025.