

METODOLOGIA PARA DETECÇÃO E GEORREFERENCIAMENTO DE PLACAS DE SINALIZAÇÃO EM VIAS RURAIS

Marcelo Franco Porto

Lucas Vinicius Ribeiro Alves

Renata Maria Abrantes Baracho Porto

Cleide Maria Martins

Bruna Cristina Beltrão Silva Beleigoli

João Victor de Melo

Universidade Federal de Minas Gerais

RESUMO: O uso da metodologia para detecção e georreferenciamento de placas de sinalização em vias rurais tem como fundamento aprimorar o gerenciamento de ativos rodoviários, especialmente as placas de sinalização. Treinamos o método *YOLO (You Only Look Once)* para detecção e georreferenciamento de sinalização vertical de trânsito em vídeos gerados das vias de estradas rurais. Neste estudo, o objetivo foi identificar a sinalização vertical de trânsito. Para isso, aplicamos através do treinamento de redes neurais convolucionais (*CNN - Convolutional neural network*) as filmagens georreferenciadas. O refinamento dos parâmetros de rede foi efetivamente definido, utilizou-se de um conjunto de dados denominado “*normalbrazilian-traffic-signs_dataset*”; e, a partir disso, conseguiu-se identificar e georreferenciar as placas presentes nos vídeos analisados. É importante ressaltar que a classificação e a precisão dependem de parâmetros, como definição do nível de confiabilidade. Observou-se que a precisão aumenta significativamente ajustando-se o número de épocas de treinamento, entre outros. Isso viabilizou a automatização do processo de identificação das placas a partir dos vídeos das filmagens das vias. Dessa forma, esse estudo contribui para à tomada de decisões relacionadas à gestão de ativos rodoviários no que se refere a sinalização e segurança viária.

Palavras-chave: Automatização. Georreferenciamento. Vias rurais. Gerenciamento.

ABSTRACT: The use of the methodology for detection and georeferencing of road signs in rural roads aims to improve the management of road assets, especially the road signs. We trained the *YOLO (You Only Look Once)* method for detection and georeferencing of vertical traffic signs in videos generated from rural roadways. In this study, the objective was to identify the vertical traffic signs. For that, we applied through the training of convolutional neural networks (*CNN*) the georeferenced footage. The refinement of the network parameters was effectively defined: we used a dataset called “*normal-brazilian-traffic-signs_dataset*”; and, from that, we were able to identify and georeference the signs present in the analyzed videos. It is important to emphasize that the classification and accuracy depend on parameters, such as the definition of the confidence level. We observed that the accuracy increases significantly with the increase we adjusted the number of training epochs; among others. This enabled the automation of the process of identifying the signs from the videos of the filming of the roads. Thus, this study contributes to the decision making related to road asset management regarding signage and road safety.

Keywords: Automation. Georeferencing. Rural Roads. Management.

1. INTRODUÇÃO

As estradas rurais, ou rodovias, são essenciais para o funcionamento do sistema viário, sendo importante que a sinalização esteja adequada para garantir maior segurança aos usuários. Uma sequência de raciocínios, instruções ou operações se torna uma parte importante para alcançar o objetivo de fazer uma melhor gestão de ativos. Através do uso de algoritmos um conjunto de instruções que determinam a sequência de operações a serem realizadas, as redes neurais anotam papel significativo em uma parte específica do método utilizado neste trabalho, cujo intuito é a detecção de placas de trânsito em vias rurais ou rodovias, como maneira de auxiliar na gestão deste ativo para tomada de decisões dos órgãos competentes.

Para tanto, a detecção e a classificação dos sinais de trânsito em ambientes não controlados são um desafio significativo na área de visão computacional devido à grande variabilidade em sua aparência. O reconhecimento desses sinais tem sido objeto de estudo para diversas finalidades, como direção autônoma e assistida HANDMANN *et al.*(1998) (FRANKE *et al.*, 1998). Ao reconhecer os sinais de trânsito, é possível alertar os motoristas sobre ações inadequadas e situações potencialmente perigosas BARÓ *et al.*(2009) e não somente na detecção de objetos.

Observando os dados das condições das rodovias no que se refere a classificação da sinalização da última pesquisa da Confederação Nacional dos Transportes (CNT) de Rodovia (2022); 23,77% das rodovias avaliadas estão em estado péssimo e ruim, sendo que a sinalização está intimamente relacionada a segurança viária.

É essencial conhecer a quantidade e o estado de conservação dos ativos, elaborar um plano de ação e definir as premissas para que as decisões permitam melhorar os níveis de segurança nas vias. Portanto, torna-se indispensável inventariar e georreferenciar esses ativos.

A rede neural utilizada nesse contexto é a *YOLO* ("You Only Look Once", Você Só Olha Uma Vez), uma arquitetura amplamente conhecida que, recentemente, chegou a sua oitava versão. Esta rede é empregada para a detecção e reconhecimento de objetos em tempo real, em imagens e vídeos. A rede que foi introduzida pela primeira vez por Joseph Redmon, Santosh Divvala, Ross Girshick e Ali Farhadi em um artigo publicado em 2016. A principal inovação desta rede é sua capacidade de prever, simultaneamente, caixas delimitadoras e probabilidades de classe para múltiplos objetos em uma única passagem pela rede neural. Isso é diferente de muitos outros métodos de detecção de objetos que envolvem múltiplas etapas ou abordagens de janela deslizante, que podem ser computacionalmente caras e mais lentas.

De acordo com Redmon *et al.*(2016) o uso do algoritmo *YOLO* divide a imagem em uma grade e atribui a cada célula da grade outras várias caixas delimitadoras (*bounding boxes*) e pontuações de confiança que representam a probabilidade de conter objetos. Em seguida, ele usa uma regressão para ajustar essas caixas delimitadoras para melhor se ajustarem aos objetos reais. Uma característica importante do *YOLO* é sua capacidade de detectar múltiplos objetos em uma única imagem, independentemente de sua sobreposição. Além disso, o algoritmo é capaz de reconhecer uma ampla variedade de classes de objetos, desde pessoas, carros até animais e objetos domésticos.

Este artigo visa aprimorar a detecção de sinalização vertical de trânsito, utilizando o processamento de redes neurais. O objetivo é automatizar a coleta de informações a partir de filmagens e dados de georreferenciamento.

2. REVISÃO DE LITERATURA

Na literatura, o uso de redes neurais é amplamente conhecido para classificação de imagens, e o método usado para o presente trabalho, como podemos verificar na parte de metodologia. Por exemplo, Njyou (2022), para 43 classes, usou o sistema alemão de

classificação de imagens, o GTSRB (*German Traffic Sign Recognition*), com um conjunto de dados composto de 39186 imagens para treinamento e 12630 imagens para teste, em que uma rede neural de 10 camadas foi usada (sendo 4 convoluções e 2 max pooling), obtendo-se nesse trabalho uma acurácia de aproximadamente 99,20%, ainda, o autor faz a distinção entre *traffic sign recognition* em duas categorias: *traffic sign detection* e *traffic sign classification*. Ainda, o autor mostra que houve evolução das *Convolutional Neural Networks (CNN)*, ou seja, Redes Neurais Convolucionais, até a que é usada hoje em dia, por exemplo, não se tanto a função de ativação tanh, e sim, a função de ativação Relu.

No trabalho, *Chen (2017)* usa uma CCNN (combined CNN) para que a partir de um processo, descobre-se a classe associada com sua respectiva probabilidade, ou seja, em linhas gerais, um número que associa quanto é acurado a detecção de uma certa classe. Para isso, foram usadas duas arquiteturas distintas para cada processo de reconhecimento. Foi usado novamente, o conjunto de dados do sistema alemão de placas (GTSRB) com 26640 para treinamento, e 12569 imagens para teste, obtendo-se uma acurácia separada para as arquiteturas A e B, ambas superiores ou igual a 95 % aproximadamente. Por fim, o autor cita que o uso do método CCNN é promissor para aplicações em tempo real, com um *time cost* de 2,7 ms e uma acurácia de 98,26 % com um *dataset* aumentado.

De acordo com *Naftali et al. (2022)*, temos uma comparação entre diversos tipos objetos, comparando algoritmos de detecção como o SSD MobileNetv2 FPN-lite 320x320, o YOLOv3, o YOLOv4, o YOLOv5l, e o YOLOv5s em tempo real, além disso, os autores usam técnicas de *augmentation* como *resizing*, *hue shifting*, e ruído, em seguida, há uma análise comparativa dos parâmetros conhecidos retornados. Nesse contexto, o YOLO, inicialmente redefine o propósito dos classificadores para fazer detecção de objetos, ainda conforme *REDMON et al. (2016)*, usando *bounding boxes* com avaliação probabilística para cada classe. Nesse modelo proposto, temos 24 redes de convolução, além de 2 camadas de neurônios totalmente conectados, e são usadas funções de ativação *Leaky RELU* para todas as camadas menos na última.

O YOLO v8 é uma rede que possui características de melhoria no Backbone, dado que o YOLOv8 utiliza uma poderosa rede de base, como o *DarkNet*, para extrair características de alto nível das imagens de entrada. Essas características são então utilizadas para a detecção de objetos, garantindo uma melhor representação dos objetos em cenas complexas. Além disso, temos nesse modelo, a *Feature Pyramid Network (FPN)* que concerne lidar com o desafio de detectar objetos em diferentes escalas, o YOLOv8 incorpora o FPN, que combina características de diferentes níveis da rede.

O algoritmo é capaz de detectar objetos, independentemente do tamanho, com precisão, o que melhora significativamente o desempenho geral de detecção. Uma característica distintiva é a sua capacidade de fazer previsões em múltiplas escalas. Em resumo, o modelo é projetado para identificar objetos em várias escalas utilizando caixas âncora (anchor boxes). Essa funcionalidade permite ao algoritmo lidar de maneira mais eficiente com objetos de tamanhos variados.

Ao incorporar técnicas de regressão de caixa delimitadora (*bounding box*), o algoritmo prevê com precisão, as coordenadas das caixas delimitadoras dos objetos, reduzindo tanto os falsos positivos quanto os falsos negativos, e por fim, a rede utiliza a supressão de não-máximo (*non-maximum suppression - NMS*) para remover previsões redundantes de caixas delimitadoras. Nessa etapa de pós-processamento, temos a garantia de que cada objeto será detectado apenas uma vez, eliminando detecções duplicadas e aprimorando os resultados finais.

3. GEORREFERENCIAMENTO

Segundo Yao (2020), o georreferenciamento é o processo de associar locais a objetos geográficos dentro de um quadro geográfico de referência. Além disso, de acordo com a Enciclopédia Internacional da Geografia Humana, existem dois mecanismos distintos de georreferenciamento: georreferenciamento métrico e georreferenciamento indireto. O primeiro é baseado em coordenadas, já o segundo é baseado em bancos de dados GIS (*geographic information system*), e esses possuem coleções de características espaciais referenciadas por coordenadas. Além disso, podemos georreferenciar objetos de modo indireto, ou seja, por meio de atributos, como índices ou nomes associados aos objetos georreferenciados metricamente. Usaremos isso de forma sistemática neste trabalho, em consonância com o uso de redes neurais, para detecção de placas em vias rurais.

Outro conceito importante é o de geocódigo, sendo este uma importante técnica de georreferenciamento, que segundo Goldberg (2011) refere-se à associação entre endereços de ruas e as coordenadas geográficas, sendo que uma definição melhor é dada, considerando-se a abrangência de vários tipos de características geográficas. Além disso, o mesmo autor menciona o processo de geocodificação como sendo composto por três partes: a tokenização, a padronização e a associação de endereço, conforme vemos na Figura. 1:

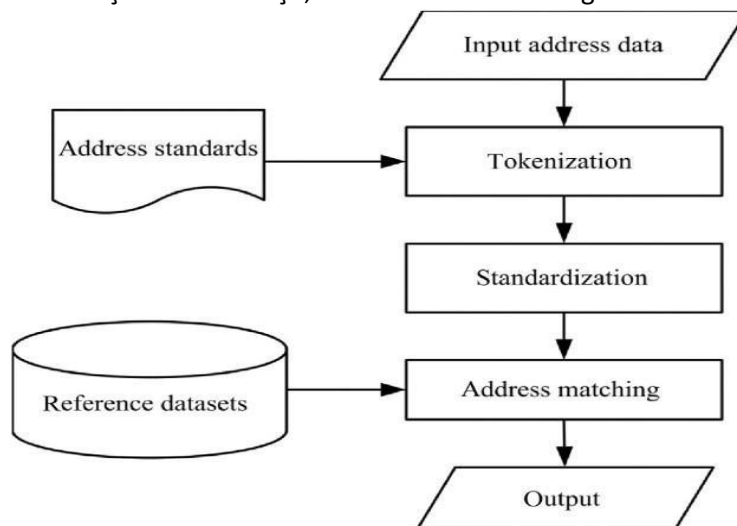
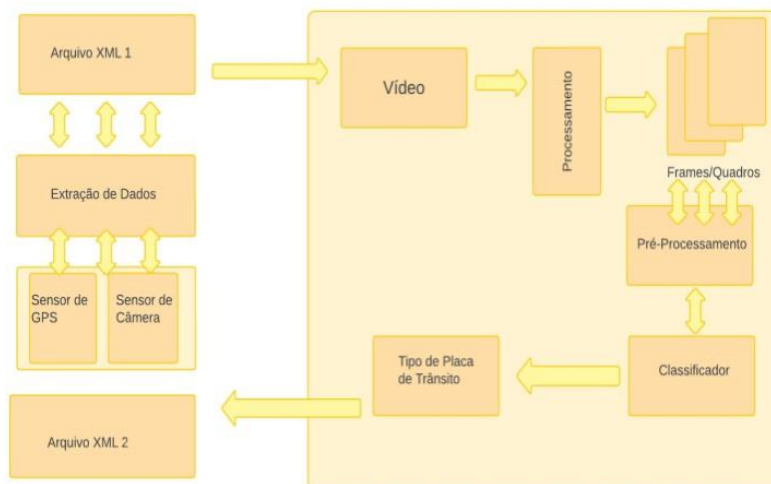


Figura 1 - Fluxograma da Geocodificação. Fonte: Goldenberg (2011).

A tokenização se refere à prática, como citado pelo autor, de dividir um endereço completo em componentes distintos. A partir de então, a padronização/*standardization* "pega" cada componente e o transforma numa forma padronizada de tal modo que possa ser usada na associação de endereço/*address matching*. Nesta última fase, componentes de uma entrada são usados para registrar nos conjuntos de dados de referência.

4. METODOLOGIA

Nesta etapa, enumeramos diversos passos que formam o conjunto de nossa metodologia: 1) obtenção a partir de filmagem em vias rurais (em arquivo do tipo .avi) e dados de georreferenciamento (em arquivo do tipo .xml); 2) transformação da filmagem em *frames* (quadros); 3) classificação de placas de sinalização encontradas dentro dos quadros; 4) união dos dados de georreferenciamento (presente inicialmente no primeiro arquivo que é enviado a um segundo). No primeiro passo, a partir de sensores GPS e câmera, obtemos os dados de georreferenciamento e filmagem do trecho correspondente, de forma que teremos todos os dados para fazermos a automação de ativos que este trabalho propõe. Os passos estão apresentados abaixo:



F
i
r
a
2
-

g u

Fluxograma do método. Fonte:
Os autores.

Assim, focando inicialmente no passo (3), no que concerne à rede neural. Foram realizados, nesta etapa, três experimentos para 232 classes, sendo as classes referentes às placas de sinalização. Foram utilizados o conjunto de dados normal-brazilian-traffic-signs_dataset, com 2816 imagens para treinamento e 402 imagens para teste com um total de 3218 imagens. Além disso, foi utilizado um batch = 16 e 50 épocas para cada um dos experimentos, que foram designados por experimentos 1, 2 e 3, designado se realizado pelo framework de YOLOv5, YOLOv7 e YOLOv8 respectivamente. Para verificar a acurácia das redes, usaremos a seguinte métrica:

$$mAP = \frac{1}{|C|} \sum_{c \in C} \frac{|TP(c)|}{|TP(c)| + |FP(c)|} \quad (1)$$

Acima, observamos a métrica utilizada neste trabalho para avaliarmos a acurácia de nosso modelo de rede neural, ou seja, usaremos a *mean average precision* (média das precisões médias) dada pelo seguinte: cV representa o número de 232 classes consideradas. Além disso, $TP(c)$ simboliza um verdadeiro positivo, ou seja, supondo que a classe c apontou um certo valor, verifica-se que de fato o resultado foi aquele, e $FP(c)$ simboliza um falso positivo, ou seja, corresponde aos erros de detecção de classes de placas de sinalização. Desse modo, $FP(c) + TP(c)$ corresponde ao número de placas total para uma dada classe, por exemplo, a classe de “pare”, e então se calcula os acertos $TP(c)$ sobre o total (uma probabilidade associada), e soma-se para cada classe (dentro das 232), sendo que por fim, toma-se a média (obtendo-se uma probabilidade ou acurácia global).

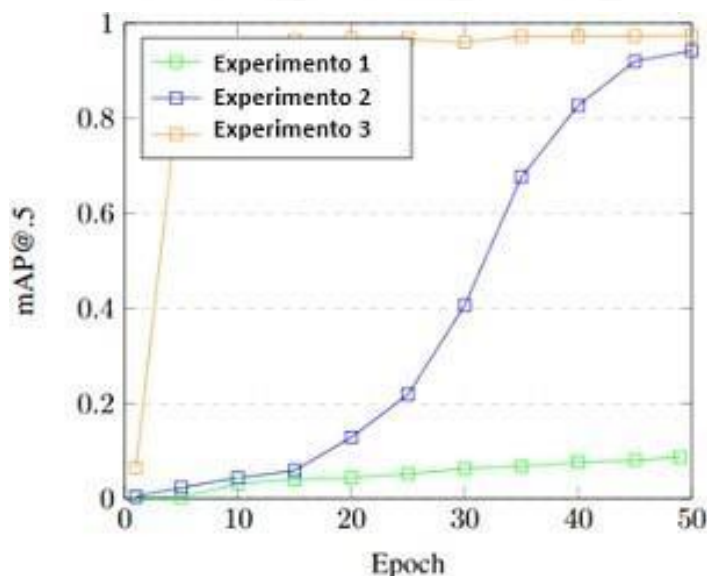


Figura 3 - mAP@.5 para diferentes épocas.

Abaixo temos uma tabela referente aos três experimentos: conforme o número de épocas aumenta, temos um aumento de nossa acurácia, dada em mAP, ou seja, na métrica definida em 3.1.

Tabela 1: mAP@.5 para diferentes épocas.

| Experimento 1 | | Experimento 2 | | Experimento 3 | |
|---------------|---------|---------------|---------|---------------|--------|
| Épocas | mAP | Épocas | mAP | Épocas | mAP |
| 1 | 0,00259 | 1 | 0,00446 | 1 | 0,0651 |
| 5 | 0,00259 | 5 | 0,0231 | 5 | 0,89 |
| 10 | 0,0312 | 10 | 0,0435 | 10 | 0,952 |
| 15 | 0,0409 | 15 | 0,059 | 15 | 0,965 |
| 20 | 0,0437 | 20 | 0,129 | 20 | 0,968 |
| 25 | 0,0518 | 25 | 0,219 | 25 | 0,967 |
| 30 | 0,0629 | 30 | 0,407 | 30 | 0,96 |
| 35 | 0,0677 | 35 | 0,677 | 35 | 0,973 |
| 40 | 0,0765 | 40 | 0,827 | 40 | 0,973 |
| 45 | 0,0808 | 45 | 0,92 | 45 | 0,973 |
| 50 | 0,0868 | 50 | 0,042 | 50 | 0,974 |

5. ESTUDO DE CASO

Apresenta-se a seguir a aplicação prática do método.

5.1 Do Treinamento do Modelo YOLO v8

O modelo YOLO v8 apresentou melhor resultado e por essa razão foi utilizado para o estudo de caso. Inicialmente, para os casos de *validation*, obtivemos o seguinte dentre cada um dos três casos distintos com as porcentagens de predição da Figura 3.

Na Figura 5, o número 1 refere-se a um acerto da placa detectada. Por exemplo, é certo que uma das placas da sexta figura (da esquerda para a direita, de cima para baixo) representa a placa A-12, além disso na mesma imagem a placa R-2 representa uma placa com 90% de correspondência. Outro exemplo se refere a primeira imagem na qual podemos ver próximo a

uma girafa a placa I-SA19 com um percentual de 100% de correspondência, indicando que a predição correspondeu aquilo que verdadeiramente é, a girafa relacionada com seu *label*.

Na figura 4, podemos ver a evolução da precisão ao longo da confiança (representado pela linha azul). Essa curva chamada de Precisão-Confiança oferece uma ferramenta valiosa para analisar a relação entre a confiança do modelo e a precisão de suas previsões. Observando que a precisão representa, na métrica utilizada, a acurácia de predições corretas feitas pelo modelo em diferentes margens de confiança. Por exemplo, na curva o valor para a precisão de 0,9 está projetado sob o ponto de confiança de cerca de 0,65. Isso indica que, para um determinado limiar de confiança (neste caso 0,65) o modelo está fornecendo uma precisão de 0,9. Em termos práticos, isso indica que, quando o modelo faz previsões com uma confiança acima de 0,65 há uma alta probabilidade de que 90 por cento dessas previsões sejam corretas, sugerindo que o modelo tem um bom desempenho em termos de precisão nessas situações específicas. Outro exemplo: tome o ponto de precisão 1,0 e, ao projetá-lo, temos uma confiança de 1,0, indicando que o modelo está extremamente seguro em suas previsões.

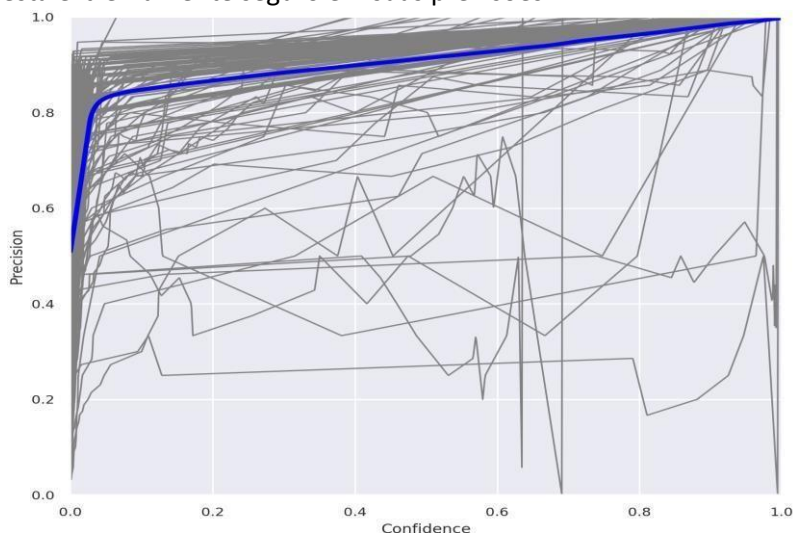


Figura 4 - Curva P da rede YOLO v8 precisão e confiabilidade.

Podemos verificar que para cada placa é atribuída uma probabilidade de termos acertado na sua identificação, conforme demonstrado na figura 5.

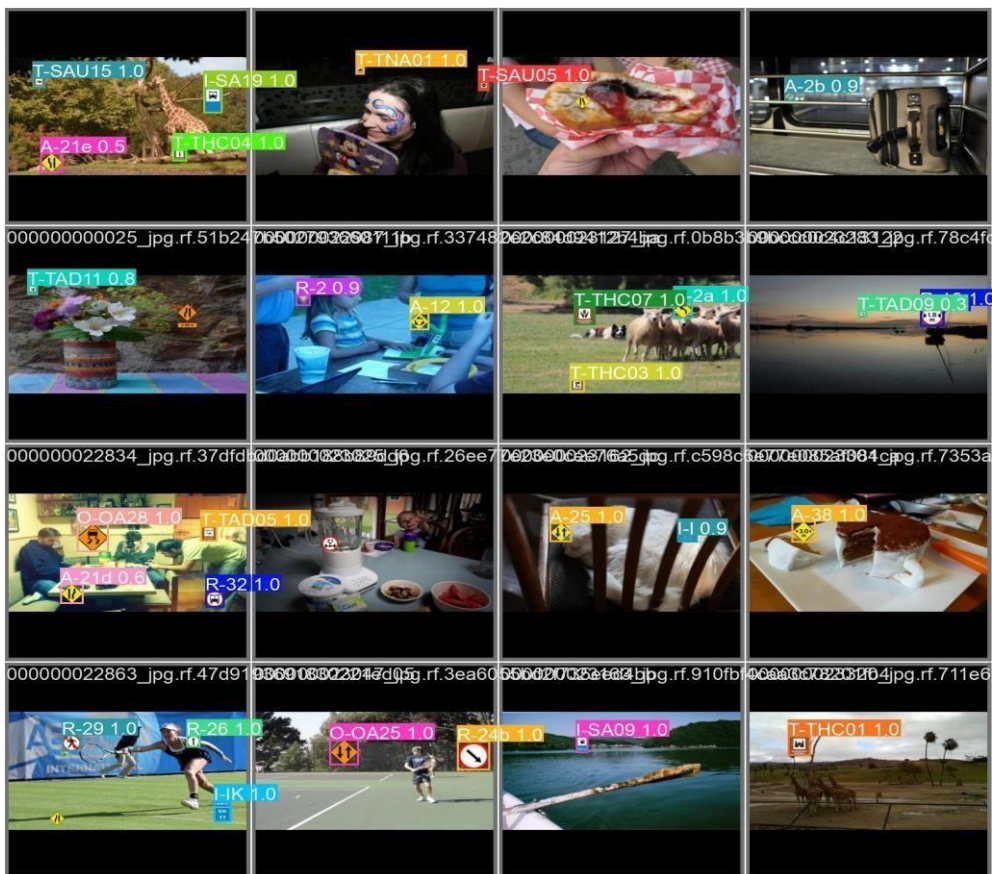


Figura 5 - Validação para um conjunto de placas.

O *dataframe* foi desenvolvido através de um vídeo específico com duração de 23 minutos e 17 segundos, para a detecção de placas. Para isso, utilizou-se um número de 10 frames por segundo. Após a conversão em frames, foram transformados em um conjunto de dados para input na rede neural, conforme o *framework* adotado. Além disso, utilizou-se os seguintes parâmetros de georreferenciamento: coordenadas, nome da via e características geométricas da via. Analisou-se o lado direito da via, com uma câmera apropriada.

Inicialmente (numa primeira abordagem), a partir frames, de um carro em movimento, durante um trecho da BR 262, entre os km 550 e 570 obteve-se dados de georreferenciamento que permitiu localizar dado pelo tempo de captura de cada frame, a localização de cada placa, bem como seu tipo, além de características geométricas da via. No entanto, neste último caso, é necessário o uso de redes específicas como visto em *Xin Chen et al. (2022)*. Nesse estudo, o objetivo não é o uso dessas técnicas de características geométricas. Foi utilizado unicamente a característica do ângulo da curva em avaliação, conforme pode ser visto abaixo nos resultados iniciais do estudo de caso:



Figura 6 - Detecção de placas (BR 262) da rede YOLO v8 para 10 épocas.

É importante destacar que nem todas as placas tiveram detecção correta. Isso se deve ao fato de que, para o caso acima, usamos apenas 10 épocas. Outro fator que influenciou foi o fato de o *data augmentation* não ser tão grande (além disso, nosso *dataset* utilizado não possui a placa R-7, de não ultrapasse).

Nesse contexto, para determinarmos o ângulo da via θ , devemos resolver a seguinte equação transcendente:

$$\sin\theta = L/L' \cdot \theta, \quad (2)$$

onde L é dado pela distância que se devia ter percorrido em linha reta, e L' é a distância que foi percorrida de fato. Usando o método de Newton Raphson para resolver (2), obteremos o seguinte:

Tabela 2: Estudo de caso da BR 262.

| Tipo de placas | Horário | Coordenada |
|----------------|----------|------------|
| A-2b | 14:59:00 | Km 549 |
| R-7 | 15:05:00 | Km 553 |
| R-19 | 15:06:00 | Km 554 |
| A-2a | 15:11:00 | Km 558 |

Estudo de Caso 2

Apresenta-se o desenvolvimento de um segundo estudo, MG-353:



Figura 7- Detecção de placas da rede YOLO v8 para 10 épocas.

Analisando a aplicação da rede, na Figura 7 podemos notar que o modelo acertou em duas das placas: a placa R-19 (limite de velocidade) e a placa A-18 (lombada), apesar de que nesta última o índice de confiança foi pequeno.

Tabela 3: Tabela de resultados testados.

| Tipo de Placa | Horário | Coordenada (X,Y) | Via (Ângulo em graus)x 10 |
|---------------|----------|--------------------------------------|---------------------------|
| R-19 | 08:41:38 | -21,6780087066667;-43,3444093666667 | 7,6491103897046 |
| A-18 | 08:42:39 | -21,6779762133333;-43,3443637916667 | 5,70865879285988 |
| A-18 | 08:42:59 | -21,6779971216667;-43,3443928783333 | 6,67509502560549 |
| A-1b | 08:45:24 | -21,6778994483333; -43,3442538900000 | Não Avaliado |



O vídeo foi dividido em *frames*, sendo que cada um foi nomeado pelo tempo em relação ao vídeo total (de cerca de 23min e 17s). O correspondente tempo do *frame* é associado a uma informação de georreferenciamento de tempo no arquivo .XML (sincronismo), o qual é um arquivo de descrição de georreferenciamento, e a partir do tempo do frame, obtemos os valores de latitude e longitude, respectivamente X e Y, que são incorporados ao nome do arquivo, além do horário em que a imagem foi feita. Em seguida esses dados são extraídos e colocados na tabela, sendo que o ângulo é calculado conforme a equação acima descrita (ou seja, temos uma função que é chamada para o cálculo). Pode-se notar que, usando uma velocidade constante para o carro de 20 km/h (aproximadamente 5,5 m/s), o ângulo é obtido pela Equação 2, tendo como referência o L dado como distância entre pares de coordenadas sucessivas, e L' dado pela $v \cdot t$, onde t é o tempo entre dois horários sucessivos.

6. CONCLUSÃO E DISCUSSÃO

A rede *YOLOv8*, apresenta algumas limitações, como a necessidade de um treinamento extenso para alcançar resultados ótimos e depois de um treinamento relativamente longo, além de que ela depende de muitos parâmetros para dar uma classificação correta, como o *bounding box* com dimensões adequadas em relação ao restante da imagem, bem como parâmetros relacionados à escala de cor. Apesar disso, o seu uso se demonstrou promissor nos dois estudos de caso realizados, uma vez que a acurácia aumenta rapidamente com o número de épocas de treinamento. Observou-se que, com um *dataset augmentation* gerando muitas imagens deveriam-se obter melhores resultados, fato que não ocorreu nos estudos realizados. Durante a análise, notou-se que a qualidade da câmera desempenha um papel significativo no processo de tomada de decisão. Isso sugere que uma câmera de qualidade superior poderia, de fato, contribuir para a obtenção de resultados mais precisos e confiáveis. O estudo também aponta oportunidades para trabalhos futuros como a integração desse sistema de detecção por imagens acoplado e integrado a um modelo 3D BIM (*Building Information Moldelling*).

Agradecimentos

Esse trabalho foi apoiado de forma colaborativa pelo projeto Transcolar Rural (EE/UFMG), Fundação de Amparo à pesquisa do Estado de Minas Gerais (FAPEMIG) e pelo Departamento de Estradas de Rodagem de Minas Gerais (DER-MG).



REFERÊNCIAS BIBLIOGRÁFICAS

- Anjan Gudigar, Shreesha Chokkadi, and Rajendra U. "A review on automatic detection and recognition of traffic sign." *Multimedia Tools and Applications*, vol. 75, 2016, pp. 333-364. doi: 10.1007/s11042-014-2293-7.
- Baró, X.; Escalera, S.; Radeva, P.; Vitrià, J. Generic Object Recognition in Urban Image Databases. In: 12th INTERNATIONAL CONFERENCE OF THE CATALAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE. Proceedings... 2009a. p. 27-34.186
- CNT. Somente 12,4% da malha viária rodoviária brasileira é pavimentada Confederação Nacional do Transporte - CNT Disponível em: Acesso em: junho 2023.
- Daniel W. Goldberg, John P. Wilson, and Craig A. Knoblock. "From text to geographic coordinates: The current state of geocoding." *URISA Journal*, vol. 19, no. 1, 2007, pp. 3346. Advances in geocoding research and practice." *Transactions in GIS*, vol. 15, no. 6, 2011, pp. 727-733. doi: 10.1111/j.1467-9671.2011.01298.x. Available at: <http://dx.doi.org/10.1111/j.1467-9671.2011.01298.x>.
- Franke, U.; Gavrilu, D.; Gorzig, S.; Lindner, F.; Paetzold, F. and Wohler, C. Autonomous driving goes downtown. *IEEE Intelligent Systems*, v. 13, n. 6, p. 40-48. 1998.
- Handmann, U.; Kalinke, T.; Tzomakas, C.; Werner, M.; Seelen W. V. An image processing system for driver assistance. In *IEEE International Conference on Intelligent Vehicles*, Stuttgart, Germany. Proceedings... 1998. p. 481-486.
- IBGE. Censo demográfico 2019. cidades.ibge.gov.br Disponível em: Acesso em: junho/ 2023.
- L. Chen, G. Zhao, J. Zhou, and L. Kuang. "Real-Time Traffic Sign Classification Using Combined Convolutional Neural Networks." In *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*, Nanjing, China, 2017, pp. 399-404. doi: 10.1109/ACPR.2017.12.
- Li, Ying, Saldanha, Carlos, Lalonde, Marc Geomodeling: Georeferencing real world objects (1996) *Proceedings of the Vision Interface Conference*, pp. 71-74. Cited 2 times
- Naftali, M.G., Sulistyawan, J.S., & Julian, K. (2022). "Comparison of Object Detection Algorithms for Street-level Objects." *arXiv preprint arXiv:2208.11315*.
- Nascimento Luciana da Silva; Espíndola Aline Calheiros; Oliveira Ricardo Venescau de Almeida; Júnior Ernesto Ferreira Nobre Aplicação de aprendizagem profunda para detecção automática de sinalização vertical em rodovias 35 Congresso de Pesquisa e ensino de transportes ANPET, 2021.
- Oliveira,m.m.; OLIVEIRA,; , OLIVEIRA, C. Y. Localização e reconhecimento de sinais de trânsito para sistemas automatizados de controle de veiculos (1998) *ANAIS DA SPIE- Sociedade Internacional de Engenharia Óptica*.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). "You Only Look Once: Unified, Real-Time Object Detection." *arXiv preprint arXiv:1506.02640*.
- Youssef Njayou. "Traffic sign classification using CNN and detection using Faster-RCNN and YOLOV4."
- Xiaobai A. Yao. "Georeferencing and Geocoding." In Audrey Kobayashi (Ed.), *International Encyclopedia of Human Geography (Second Edition)*, Elsevier, 2020, pp. 111-117. ISBN: 9780081022962. Available at: <https://doi.org/10.1016/B978-0-08-102295-5.10548-7>.

Xin Chen, Qun Sun, Wenyue Guo, Chungping Qiu, Anzhu Yu, GA-Net: A geometry prior assisted neural network for road extraction, International Journal of Applied Earth Observation and Geoinformation, Volume 114, 2022, 103004, ISSN 1569-8432,



<https://doi.org/10.1016/j.jag.2022.103004>.

Njayou, Y. (2022). Traffic sign classification using CNN and detection using Faster-RCN and YOLOV4. Heliyon, 8(12), e11792. <https://doi.org/10.1016/j.heliyon.2022.e11792>

YOLO. "normal-brazilian-traffic-signs Dataset." Open Source Dataset. Roboflow Universe, Roboflow, Oct 2022. Available online at: <https://universe.roboflow.com/yolo-uiruj/normal-brazilian-traffic-signs>. (Visited on May 12, 2023)