

Criação de banco de dados de imagens de satélite para detecção de veículos

Luan André Contel

Prof. Dr. André Luiz Cunha

Universidade de São Paulo

luanppp10@usp.br

Objetivos

Essa pesquisa buscou avaliar técnicas de Visão Computacional voltadas para o monitoramento de veículos, como a detecção por veículos por imagens de satélite (Chen et al, 2014; Jiang et al, 2015). O objetivo desta pesquisa é demonstrar a viabilidade do uso de imagens de satélite para detecção de veículos, o que pode auxiliar os estudos de tráfego em qualquer região. Os desafios deste trabalho incluem a calibração de um sistema de detecção veicular automático, além da criação e disponibilização do banco de dados de veículos rotulados em imagens de satélite nas rodovias brasileiras.

Métodos e Procedimentos

A construção do banco de dados envolveu a coleta de imagens de satélite utilizando o *software* Google Earth. A rotulação, que é a criação de caixas de contorno em torno dos objetos detectados, que no caso são veículos leves e pesados, foi realizada por meio do *software* Label Studio e com o auxílio das ferramentas SAHI (*Slicing Aided Hyper Inference*) e SAM (*Segment Anything Model*). O treinamento dos modelos de detecção sugeridos foram realizados pelo YOLO v8 e na variação medium (yolov8m). Além disso, ao todo foram treinadas 9 combinações distintas de imagens, variando a proporção entre treino, teste e validação, como detalhado na Tabela 1.

Tabela 1 - Divisão das imagens em treino, teste e validação

Combinação	treino (%)	validação (%)	teste (%)
1	64	16	20
2	48	32	20
3	32	48	20
4	48	12	40
5	36	24	40
6	24	36	40
7	32	8	60
8	24	16	60
9	16	24	60

Dessa forma, o desempenho dos modelos treinados foi avaliado com base na métrica de *precision*. Por fim, o banco de dados foi padronizado e disponibilizado para a comunidade científica.

Resultados

Os resultados foram divididos em 3 partes: criação do banco de dados, treinamento dos modelos de detecção e avaliação dos resultados obtidos. Inicialmente, a criação do banco de dados foi condicionada por 2 processos: coleta de dados e rotulagem dos veículos (leves e pesados). O processo da coleta de dados gerou ao todo 701 imagens. A rotulagem dos veículos foi realizada pelo

software Label Studio e esse processo foi realizado tanto manualmente (média de amostragem de 4 imagens/h) quanto com o auxílio de ferramentas como o SAHI e SAM (média de 8 imagens/h) - melhorando a taxa de produtividade de anotação dos veículos. O treinamento dos modelos de detecção foi realizado em YOLO e com base nos conjuntos de imagens treino e teste. Após o treinamento dos modelos de detecção foi realizada a inferência de todas as imagens do conjunto de validação e, a partir das caixas de contorno previstas pelos modelos, foi possível calcular o IoU (*Intersection over Union*) e a precisão para todas as combinações. Dessa forma, foram avaliadas as precisões dos modelos utilizando

diferentes limiares de IoU (0,5; 0,6 e 0,7). Os resultados indicaram que, à medida que o limiar de IoU aumenta, a precisão média diminui, devido ao maior rigor nas detecções verdadeiras. As distribuições das precisões dos modelos foram comparadas usando o teste Qui-Quadrado de Pearson, o que revelou que a proporção de imagens de treino, validação e teste influencia significativamente na qualidade dos modelos. A combinação 4 se destacou como a melhor entre as testadas, com a maior precisão média para IoU de 0,5 e 0,6, e o menor desvio padrão para IoU de 0,7, como pode ser visualizado na Tabela 2.

Tabela 2 - Média e desvio padrão para cada combinação

		Comb1	Comb2	Comb3	Comb4	Comb5	Comb6	Comb7	Comb8	Comb9
IoU50	Média	0,835	0,809	0,826	0,853	0,853	0,825	0,820	0,840	0,838
	Desvio Padrão	0,124	0,138	0,142	0,118	0,128	0,146	0,158	0,116	0,111
IoU60	Média	0,777	0,743	0,765	0,794	0,777	0,760	0,747	0,759	0,759
	Desvio Padrão	0,149	0,148	0,146	0,132	0,145	0,163	0,166	0,133	0,124
IoU70	Média	0,609	0,570	0,582	0,605	0,588	0,598	0,555	0,548	0,571
	Desvio Padrão	0,192	0,183	0,176	0,157	0,184	0,192	0,170	0,158	0,167

Conclusões

Portanto, nota-se que o uso dos modelos SAM e SAHI foram eficientes para aumentar a produtividade das anotações das imagens. Além disso, observou-se que as precisões dos modelos diminuíram conforme o limiar de IoU aumentou, uma vez que limiares mais altos são mais rigorosos e exigem maior correspondência entre caixas de detecção previstas e reais. Além disso, dentre todas as combinações testadas, a combinação 4 foi a que obteve os melhores resultados de precisão. Por fim, pode-se concluir que é possível detectar veículos por imagens de satélite, pois existem bancos de dados e ferramentas acessíveis para esta tarefa.

Agradecimentos

Gostaria de agradecer o professor André Luiz Cunha pela oportunidade de realizar essa pesquisa e a Universidade de São Paulo pelo auxílio fornecido, que foi de suma importância para a minha continuação na graduação.

Referências

- CHEN, Xueyun et al. Vehicle detection in satellite images by hybrid deep convolutional neural networks. **IEEE Geoscience and remote sensing letters**, v. 11, n. 10, p. 1797-1801, 2014.
- JIANG, Qiling et al. Deep neural networks-based vehicle detection in satellite images. In: **2015 international symposium on bioelectronics and bioinformatics (ISBB)**. IEEE, 2015. p. 184-187.