

Seminário de Andamento

Contagem e Classificação de veículos com processamento de vídeo utilizando YOLOv3 e DeepSORT



LSI
Laboratório de
Sistemas Industriais



**DEPARTAMENTO
AUTÔNOMO DE
ESTRADAS DE
RODAGEM**

Anderson Andrei Schwertner

25 de Setembro de 2020

Orientador: Prof. Dr. Valner João Brusamarello

Organização da Apresentação

1. Introdução

- Motivação, justificativa e objetivos da Dissertação de Mestrado.

2. Revisão Bibliográfica

- Redes neurais convolucionais (CNNs), classificação de imagens, detecção de objetos e rastreamento de objetos.

3. Metodologia

- Detector de objetos: YOLOv3
- Rastreador de objetos: DeepSORT
- Contagem e sentido de deslocamento

4. Resultados Parciais

5. Cronograma

Introdução - Motivação e Justificativa

Sistemas de Monitoramento do tráfego tem papel importante no desenvolvimento de cidades inteligentes.

- As contribuições desses sistemas podem tornar a malha de transporte mais eficiente, inteligente e segura.
- Os dados coletados através deles são fundamentais para engenheiros de trânsito em muitas tarefas como:
 - projeto de intersecções;
 - instalação e sincronização de semáforos;
 - construção de lombadas e redutores;
 - instalação ou modificação de sinalização;
- Há também a possibilidade desses sistemas fornecerem a base para o desenvolvimento de novas estratégias para mitigação de congestionamentos.

Introdução - Objetivos

Essa dissertação faz parte do projeto de pesquisa "Sistema de contagem, classificação e direção de fluxo de veículos em estradas do RS".

Objetivo Principal

- Desenvolver um sistema automático para contagem, classificação e determinação da direção tomada pelos veículos em rodovias e intersecções a partir do processamento de vídeo.

Essas informações são atualmente obtidas através do processamento manual de imagens capturadas por uma câmera temporariamente instalada em cada local em que há uma solicitação para o DAER-RS.

Revisão Bibliográfica - Histórico

Deteccção, classificação e rastreamento são tarefas comuns dentro da área de visão computacional mas que observaram grandes avanços a partir de 2012.

- O processo de contagem de veículos já foi implementada com técnicas clássicas de processamento de imagem, como por exemplo filtro de partículas, operações morfológicas, detecção de bordas e análise de histograma.

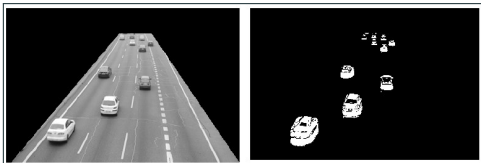


Figura 1: Técnica de subtração de fundo estático.
Adaptado de [1] (2018)

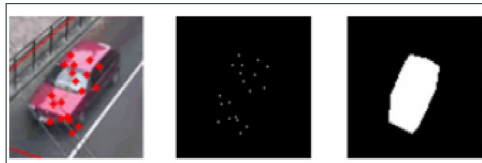


Figura 2: Técnica de filtro de partículas. Adaptado de
[2] (2013)

Revisão Bibliográfica - Histórico

- A classificação de imagens, apesar de intuitiva para nós, é uma tarefa bem complicada para computadores devido ao *gap* entre o conteúdo semântico da imagem e o valor RGB dos pixels.
- Foi somente a partir da popularização das CNNs, originalmente com a Alex-Net em 2012, que ocorreram grandes avanços nessa área.

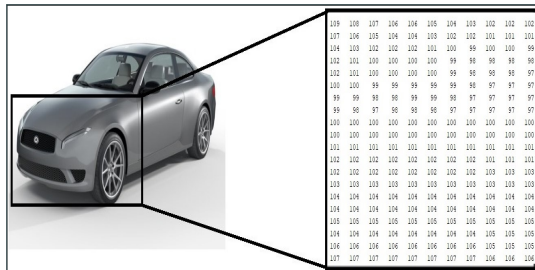


Figura 3: Representação de que, para um computador, uma imagem é apenas uma grade de números entre 0 e 255.

Revisão Bibliográfica - CNNs e a classificação de imagens

Redes Neurais Convolucionais capturam com sucesso as dependências locais e globais das imagens através da convolução de filtros com a imagem.

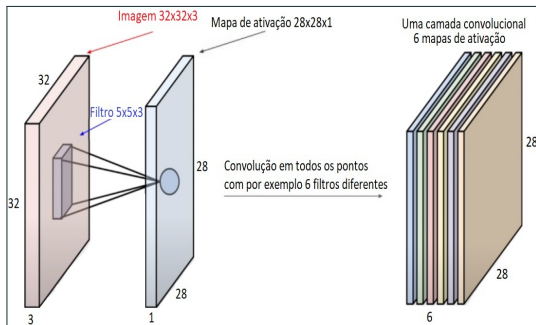


Figura 4: Representação do processo de convolução e de uma camada convolucional, adaptado de [3].

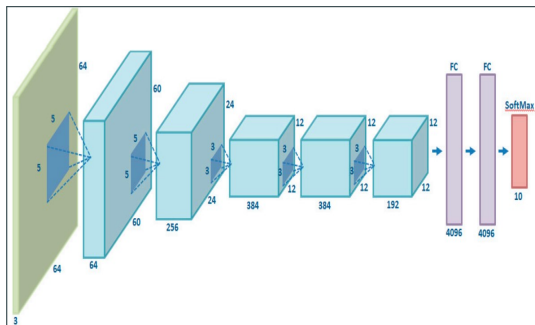


Figura 5: Visualização da arquitetura da AlexNet [4].

Revisão Bibliográfica - CNNs e a classificação de imagens

- As metodologias baseadas em CNNs podem ser diretamente associadas a maior parte do progresso obtido na área nos últimos anos.
- Devido ao seu desempenho passaram também a ser utilizadas com êxito como *backbone* em algoritmos de detecção de objetos.

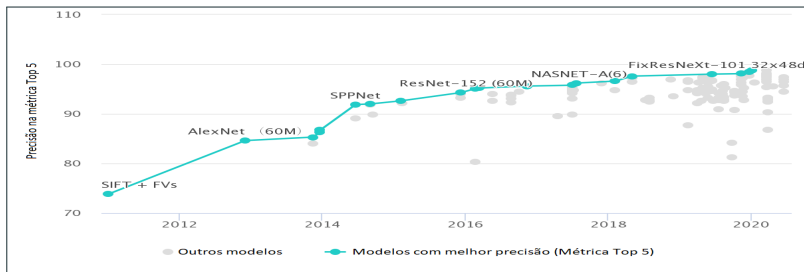


Figura 6: Estado da arte (Top 5) na tarefa de classificação de imagem na base de dados ImageNet [5].

Revisão Bibliográfica - Detecção de objetos

A tarefa de detecção de múltiplos objetos é mais complexa e exige mais etapas que a de classificação de imagens.

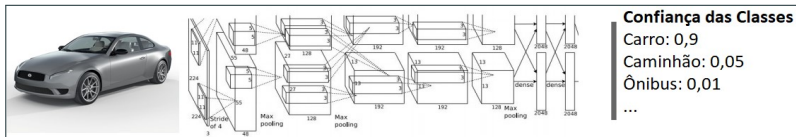


Figura 7: Exemplo de classificação de imagem. Adaptado de [3].



Figura 8: Detecção de múltiplos objetos em uma imagem. Adaptado de [6].

Revisão Bibliográfica - Detecção de objetos

Os algoritmos de detecção de objetos podem ser divididos em dois grupos principais:

- Métodos de estágio único
 - Estimam as *bouding boxes* e classificam os objetos simultaneamente. Alguns exemplos são:
 1. Single Shot Detector (SSD)
 2. You Only Look Once (YOLO)
- Métodos de estágio duplo
 - Primeiramente encontram regiões de interesse para em seguida classificá-las separadamente. Alguns exemplos são:
 1. R-CNN
 2. Faster R-CNN

Detectores de estágio único são mais rápidos ao custo de precisão se comparados aos de estágio duplo

Revisão Bibliográfica - Detecção de objetos

Assim como para classificação de imagens, as CNNs foram responsáveis pela recente melhoria na precisão dos métodos de detecção de objetos.

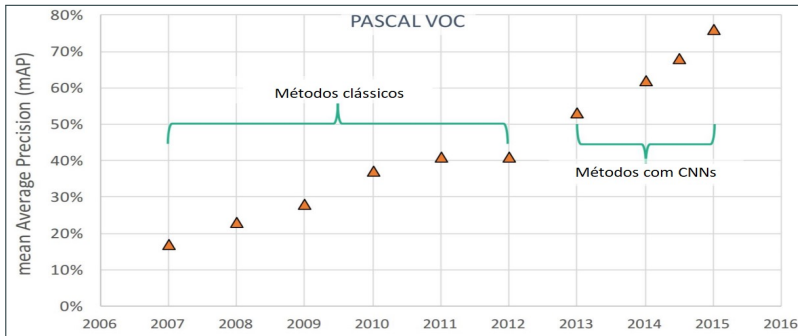


Figura 9: Histórico da precisão de algoritmos de detecção de objetos na base PASCAL VOC. Adaptado de [3].

Revisão Bibliográfica - Rastreamento de objetos

Esses recentes avanços tornaram o rastreamento por detecção em uma das abordagens mais populares para o rastreamento de objetos em vídeos.

- Rastreamento por detecção é um processo de duas etapas:
 1. Objetos são detectados em cada *frame* utilizando um algoritmo como o SSD;
 2. Objetos são rastreados pela associação das detecções em frames subsequentes combinadas com a previsão de posições futuras.

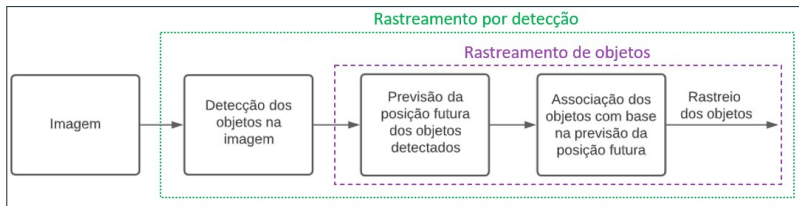


Figura 10: Fluxograma genérico de um algoritmo de rastreamento por detecção

Metodologia - Detector de objetos: YOLOv3

Utiliza uma combinação da DarkNet-53 (backbone) anexada a mais 53 camadas responsáveis pela detecção, totalizando 106 camadas.

- Realiza a detecção em 3 "resoluções" diferentes para encontrar objetos de diferentes tamanhos.

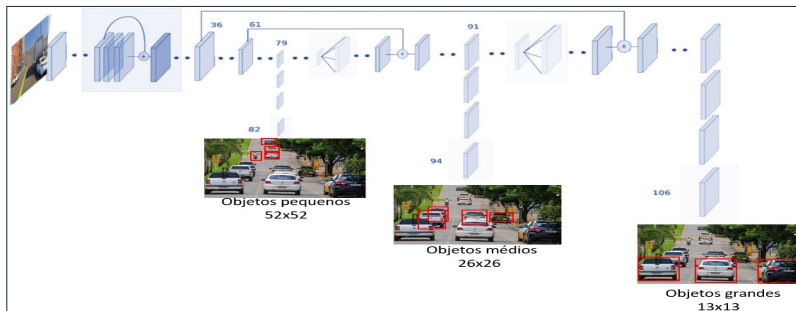


Figura 11: Demonstração da detecção de objetos em 3 estágios do YOLOv3. Adaptado de [7]

Metodologia - Detector de objetos: YOLOv3

- Na última camada da DarkNet-53 as características da imagem estão contidas em um mapa de ativação de 13x13.
- Cada célula é o centro de uma âncora onde são estimados:
 1. Os offsets ($\Delta x, \Delta y, \Delta h, \Delta w$) para as 3 bounding boxes padronizadas;
 2. A probabilidade da célula ser o centro de um objeto;
 3. O score de cada classe de objeto.
- O tensor de saída para cada célula é então $3 \times (5 + \text{Classes de objeto})$.



Figura 12: Visualização das células

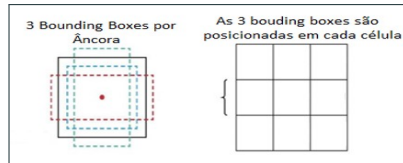


Figura 13: Exemplo de Âncoras e bounding boxes

Metodologia - Detector de objetos: YOLOv3

- O mesmo processo é realizado nos mapas 26x26 e 52x52 totalizando 10647 bounding boxes armazenados em 3 tensores.



Figura 14: Visualização das células nos 3 estágios de detecção

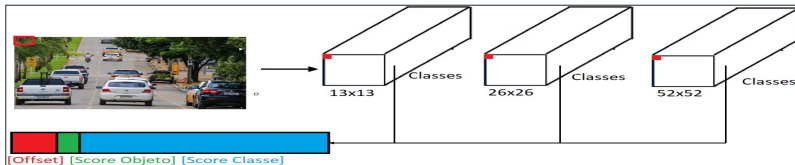


Figura 15: Tensores de saída contendo as informações das detecções nos 3 níveis

Metodologia - Detector de objetos: YOLOv3

- Detecções com confiança abaixo de 0,5 são descartadas.
- Detecções repetidas são eliminadas através da supressão de não máximos se a $IoU \geq 0.5$ considerando a melhor detecção.

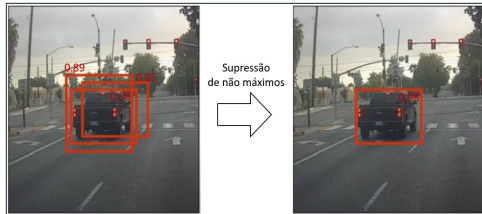


Figura 16: Exemplo de supressão de não máximos



Figura 17: Métrica Intersection over Union

Metodologia - Rastreador de objetos: DeepSORT

- Os objetos já rastreados têm suas novas posições estimadas usando um Filtro de Kalman com um modelo linear de velocidade constante.
- A associação entre as novas detecções e as posições estimadas é resolvida através do método húngaro.
- São utilizadas duas métricas nessa associação, uma representando informações de movimento e outra de aparência, esta obtida através de uma CNN. As duas métricas são combinadas através de uma soma ponderada.

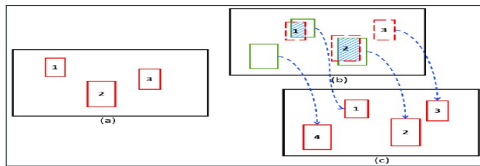


Figura 18: Exemplo do problema de associação

Metodologia - Contagem e sentido de deslocamento

As coordenadas iniciais e finais de cada veículo rastreado são armazenadas para a determinação de regiões de contagem na imagem.

- O conjunto de coordenadas é analisado como um problema de clusterização que quando resolvido retorna regiões da imagem em que carros surgem e desaparecem.
- Nessas regiões é analisada a quantidade de entradas e saídas resultando na contagem e sentido de deslocamento dos veículos.



Figura 19: Exemplo de determinação das regiões para contagem de veículos na imagem

Demonstração da implementação

Demonstração da implementação

	2020				2021
	1º Tri	2º Tri	3º Tri	4º Tri	1º Tri
Revisão bibliográfica	✓	✓	X		
Filtragem e preparação do Banco de Dados	✓	✓			
Estudo de Python e Deep Learning	✓	✓			
Detector de veículos		✓			
Rastreador de veículos, Contagem e Direção			X	X	
Obtenção e Análise dos resultados			X	X	
Escrita de artigo para publicação			X		
Escrita da dissertação		✓	X	X	X
Defesa					X

Figura 20: Cronograma com o planejamento da dissertação

- 20 créditos cursados em 2019;
- Prova de proficiência em lingua inglesa já realizada;
- Estágio Docência em andamento.

Seminário de Andamento

Contagem e Classificação de veículos com processamento de vídeo utilizando YOLOv3 e DeepSORT



LSI
Laboratório de
Sistemas Industriais










**DEPARTAMENTO
AUTÔNOMO DE
ESTRADAS DE
RODAGEM**

Anderson Andrei Schwertner

25 de Setembro de 2020

Orientador: Prof. Dr. Valner João Brusamarello

-  P. N. Chowdhury, T. C. Ray, and J. Uddin, “A vehicle detection technique for traffic management using image processing,” in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, IEEE, feb 2018.
-  C. Bouvié, J. Scharcanski, P. Barcellos, and F. L. Escouto, “Tracking and counting vehicles in traffic video sequences using particle filtering,” pp. 812–815, IEEE, 2013.
-  F.-F. Li, J. Johnson, and S. Young, “Lecture notes in convolutional neural networks,” April 2017.
-  J. Llamas, P. M. Lerones, R. Medina, E. Zalama, and J. Gomez-Garcia-Bermejo, “Classification of architectural heritage images using deep learning techniques,” *Applied Sciences*, vol. 7, p. 992, sep 2017.
-  P. with Code, “Image classification on imagenet,” 2020.

-  Z. Tang, G. Wang, H. Xiao, A. Zheng, and J.-N. Hwang, "Single-camera and inter-camera vehicle tracking and 3d speed estimation based on fusion of visual and semantic features," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, jun 2018.
-  S. Dulepet, P. Maji, M. Harsh, and K. Burke, "Deploying a scalable object detection inference pipeline," 2020.