



**UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA**

Relatório Final
Trabalho de Conclusão de Curso

VISÃO COMPUTACIONAL APLICADA A VEÍCULOS AUTÔNOMOS

Autor: Pedro Gabriel

Orientadora: Profa. Dra. Ludmila Corrêa de Alkmin e Silva

Campinas
2022

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA MECÂNICA

Relatório Final
Trabalho de Conclusão de Curso

VISÃO COMPUTACIONAL APLICADA A VEÍCULOS AUTÔNOMOS

Autor: Pedro Gabriel

Orientadora: Profa. Dra. Ludmila Corrêa de Alkmin e Silva

Curso: Engenharia de Controle e Automação

Trabalho de conclusão de curso apresentado à Comissão de Graduação da Faculdade de Engenharia Mecânica como requisito para a obtenção do título de Engenheiro de Controle e Automação.

Campinas
2022

Agradecimentos

Este trabalho não poderia ser terminado sem a ajuda de diversas pessoas às quais presto minha homenagem:

Primeiramente, gostaria de agradecer à professora Ludmila por me acompanhar nesse trabalho e por toda sua dedicação ao longo dos projetos que fizemos juntos ao longo de minha graduação.

Para os membros do laboratório de sistemas integrados da FEM (Labsin), que me auxiliaram durante os anos que estivemos juntos e contribuíram de maneira indispensável para a minha formação acadêmica.

Aos coordenadores de Graduação do curso de Engenharia de controle e automação da FEM/Unicamp, Prof.Dr. Tiago Henrique Machado e Prof.Dr. Gregory Bregion Danieli.

À minha família cujo apoio e suporte me possibilitou chegar até aqui.

Para todos meus colegas que me acompanharam durante a graduação, cuja companhia foi extremamente importante para meu crescimento pessoal e profissional.

Para a minha namorada Alessandra, que me deu uma grande suporte durante grande parte dessa jornada sempre me incentivando e me animando nesse percurso.

Resumo

Gabriel, Pedro, Visão computacional aplicada a veículos autônomos, Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Trabalho de Conclusão de Curso, (2022).

A aplicação de veículos autônomos à mobilidade urbana é uma mudança promissora, capaz de trazer mais eficiência e segurança à locomoção. Porém que ainda enfrenta diversos desafios no processo para se tornar verdadeiramente autônomo. A utilização de ferramentas de visão computacional vem sendo fundamental nesse processo e cada vez mais utilizada, com a evolução de técnicas de aprendizado de máquinas e a disponibilidade de maiores bancos de dados, permitindo a construção de sistemas cada vez mais complexos. Dentro desse contexto, esse trabalho tem por objetivo fazer a construção de um algoritmo de visão computacional em tempo real, capaz de processar tanto a detecção de objetos dentro do campo de visão do veículo quanto a identificação de uma trajetória para o mesmo. Esse algoritmo é composto por duas etapas distintas: A primeira sendo uma implementação de aprendizado de máquinas de modo a utilizar a rede You Only Look Once (YOLO) para identificação de uma série de objetos. Com relação a segunda parte, tem-se a utilização de um método clássico de identificação de trajetória a partir de marcações da estrada, fazendo uso de diversos tratamentos para extração de informação da imagem, chamado *B-Snake*. Assim, a análise irá abordar a utilização desse algoritmo em tempo real aplicado a condições próximas ao esperado para a aplicação em um veículo urbano padrão, observando tanto tempo de execução como robustez de cada método.

Palavras Chave: **Visão Computacional, Veículo autônomo, Aprendizado de máquinas**

Lista de Figuras

| | | |
|------|---|----|
| 1.1 | Estrutura de uma ANN. Fonte: (O'SHEA; NASH, 2015) | 14 |
| 1.2 | Funcionamento de uma camada convolucional em uma CNN. Fonte: (O'SHEA; NASH, 2015) | |
| 1.3 | Imagens do banco de dados <i>BDD100K</i> . Fonte: (YU et al., 2018) | 17 |
| 1.4 | (a) Modelo 2D da estrada com linhas paralelas. (b) Modelo da estrada com linhas paralelas e | |
| 1.5 | Estrutura da rede neural YOLO. Fonte: (REDMON; DIVVALA et al., 2016) | 27 |
| 1.6 | Estrutura da rede neural YOLO V3. Fonte: (LIU; ZHANG, 2020) | 28 |
| 1.7 | Saída de um célula da rede YOLO. Fonte: (PYLESSONS.COM, 2022) . . . | 30 |
| 1.8 | Processo de redimensionamento de uma caixa ancora. Fonte: (REDMON; FARHADI, 2018) | |
| 2.1 | Fluxo de trabalho na etapa de detecção de objetos. Fonte: Autor | 37 |
| 2.2 | Fluxo de trabalho na etapa de navegação. Fonte: Autor | 48 |
| 2.3 | Processo de identificação dos pontos de trajetória da navegação. Fonte: (WANG; TEOH; SHI) | |
| 2.4 | Pontos de referencia para a trajetória do veiculo. Fonte: (WANG; TEOH; SHEN, 2004) | 53 |
| 2.5 | Representação gráfica do gradiente obtido por Canny Edge Detection, separado em ambos eixos | |
| 2.6 | Representação gráfica do GVF, separado em ambos eixos. Fonte: Autor . . | 57 |
| 2.7 | Representação gráfica do gradiente obtido por Canny Edge Detection. Fonte: Autor | 58 |
| 2.8 | Representação gráfica do gvf. Fonte: Autor | 59 |
| 3.1 | Resultados obtidos em ambiente diurna após 1200 lotes. Fonte: Autor . . . | 64 |
| 3.2 | Resultados obtidos em ambiente noturno após 1200 lotes. Fonte: Autor . . | 65 |
| 3.3 | Ambiente contendo muitos objetos para detecção. Fonte: Autor | 66 |
| 3.4 | Erro de detecção de veiculo próximo. Fonte: Autor | 67 |
| 3.5 | Influencia do número de seções na detecção da trajetória. Fonte: Autor . . | 69 |
| 3.6 | Desempenho do algoritmo para retirada de linhas indesejadas. Fonte: Autor | 70 |
| 3.7 | Processamento do trajeto em uma rua com múltiplas marcações. Fonte: Autor | 71 |
| 3.8 | Preparo para a construção da trajetória. Fonte: Autor | 72 |
| 3.9 | <i>Spline</i> da trajetória. Fonte: Autor | 73 |
| 3.10 | Curva final após resultado da minimização. Fonte: Autor | 74 |
| 3.11 | <i>B-Snake</i> executado em uma imagem com visão comprometida. Fonte: Autor | 75 |

Lista de Tabelas

| | | |
|-----|---|----|
| 3.1 | Resultados de validação da rede YOLO. Fonte: Autor | 68 |
| 3.2 | Configurações do computador utilizado para simular o processamento do algoritmo. Fonte: A | |
| 3.3 | Tempo de execução da etapa de configuração. Fonte: Autor | 73 |
| 3.4 | Tempo de execução da etapa por quadro. Fonte: Autor | 74 |

Glossário

Canny Edge Detection Algoritmo utilizado para detecção de bordas em uma imagem, por meio da variação de intensidade entre pixels em uma vizinhança. 5, 19, 22, 23, 36, 47, 48, 56, 58, 60

Deep Learning Nomenclatura na qual é conhecido estrutura de rede neurais com grande quantidade de camadas escondidas. 24

Filtro de Kalman Algoritmo construído para identificar os estados de um sistema, através de uma observação do mesmo que pode conter ruído. 19

Framework Conjunto de ferramentas confeccionado para auxiliar no desenvolvimento de projetos em uma determinada área. 12, 24, 36, 39, 41, 43, 45–47, 66, 82

Json (JavaScript Object Notation) É uma formatação para arquivo de dados com formatação similar a objetos escritos na Linguagem Json, sendo de fácil leitura para usuários e usando estrutura familiar a diversas linguagens de programação.. 43

Limiar de cores Técnica na qual é utilizada a detecção de cores, por meio de limiares definindo a range de valores correspondentes ao range de cores desejados. 19

OpenCV Open Computer Vision (OpenCV) é um framework que permite acesso a diversas API para trabalhos com visão computacional. 35, 36, 48

Overfitting Problema causado durante o treinamento de aprendizado de máquinas onde o modelo se torna especializado nos dados apresentados no seu treinamento e tem um resultado insatisfatório na sua aplicação com outros dados. 15, 38

Spline Função matemática que define polinômios para representar trechos de uma curva complexa, separado por pontos de observação da curva conhecidos como nós. Tem como objetivo a modelagem de uma curva complexa pelo uso de polinômios mais simples para cada trecho. 23

Transformada de Hough Técnica de extração de características utilizada no processamento de imagens. O propósito dessa técnica é a identificação de objetos de uma determinada forma, normalmente aplicado para detecção de retas. 19, 22, 36, 48, 49, 69, 70, 76

Waymo Waymo é a atual empresa que trabalha com veículos autônomos, dando continuidade ao projeto de carros autônomos da Google. 11

Siglas

ADAS Advanced Driver-Assistance System. 10

ANN Artificial Neural Network. 5, 13, 14

CHEVP Canny/Hough Estimation of Vanishing Points. 22, 47, 60

CNN Convolutional Neural Network. 5, 14–16, 19, 26

GPU Graphics Processing Unit. 24

GVF Gradient Vector Flow. 5, 55, 57

ML Machine Learning. 12, 13, 15

OpenCV Open Computer Vision. 7, 24

RCNN Regional Convoluted Neural Network. 24

SGD Stochastic Gradient Descent. 32

SSD Single Shot Detector. 24

YOLO You Only Look Once. 4–6, 24, 26–31, 35, 36, 39–41, 62, 63, 67, 68, 72, 76, 82

Sumário

| | |
|--|-----------|
| Introdução | 10 |
| 1 Revisão Bibliográfica | 12 |
| 1.1 Aprendizado de máquinas | 12 |
| 1.1.1 Redes Neurais Convolucionais (CNN) | 14 |
| 1.1.2 Banco de dados | 15 |
| 1.2 Visão computacional | 17 |
| 1.2.1 Navegação | 18 |
| 1.2.2 Detecção de objetos | 24 |
| 1.3 YOLO | 26 |
| 1.3.1 Funcionamento | 28 |
| 1.3.2 Função de custo | 31 |
| 2 Metodologia | 34 |
| 2.1 Técnicas utilizadas | 34 |
| 2.2 Detecção de objetos | 36 |
| 2.2.1 Escolha do banco de dados | 37 |
| 2.2.2 Adaptação do modelo YOLO | 39 |
| 2.2.3 Treino do modelo | 41 |
| 2.2.4 Teste e validação | 46 |
| 2.3 Navegação | 47 |
| 2.3.1 Estimativa base das marcações | 47 |
| 2.3.2 Ajuste dos parâmetros do modelo | 55 |
| 2.3.3 Aplicação em tempo real | 59 |
| 2.4 Aplicação de ambos os processos | 60 |
| 3 Resultados | 62 |
| 3.1 Detecção de objetos | 62 |
| 3.1.1 Validação do rede neural | 66 |
| 3.2 Navegação | 68 |
| 3.3 Detecção em tempo real | 72 |
| Conclusões | 76 |
| Referências bibliográficas | 78 |
| A Implementação da rede YOLO usando Keras | 82 |
| B Implementação do algoritmo de navegação B-Snake | 95 |

Introdução

A utilização de veículos autônomos é uma das maiores promessas para mudar a maneira como lidamos com a mobilidade atualmente, tendo a capacidade de melhorar a segurança, rapidez e eficiência da locomoção. Porém a aplicação dentro de um ambiente urbano é extremamente complexo, a grande quantidade de variáveis e situações imprevisíveis faz com que o projeto de um modelo capaz de prever e atuar dentro dessas condições seja um verdadeiro desafio.(JANAI et al., 2020) Assim a construção de um veículo totalmente autônomo enfrenta dificuldades em alguns aspectos para poder ser utilizado na vida cotidiana, sendo elas: Robustez, precisão, execução em tempo real e preço.(LÓPEZ et al., 2017) Para que um sistema como esse que possui um alto nível de criticidade, uma falha no sistema poderia ser fatal, é extremamente importante que a confiabilidade do mesmo seja alta. Além da necessidade de que o sistema seja sustentável em termos de custo e energia de modo a ser comercializável.

Por outro lado, mesmo que um sistema totalmente autônomo ainda não esteja ao alcance da evolução dos campos de estudos em visão computacional e aprendizado de máquinas vem trazendo essa tecnologia cada vez mais perto de se concretizar, sendo um campo que recebe atualmente grande investimentos e incentivos.(AGARWAL; CHIANG; SHARMA, 2018)

Atualmente já é possível ver avanços significativos com veículos capazes de lidar com situações específicas onde o ambiente é mais previsível que as rodovias (LÓPEZ et al., 2017), ou com tecnologias que permitem uma assistência ao motorista.

Dentre esses casos, têm-se os veículos que atualmente são o estado da arte no quesito de direção autônoma. Como é o caso de alguns veículos Tesla, possuindo um avançado sistema Advanced Driver-Assistance System (ADAS), integrado com diversas funcionali-

dades que aumentam o conforto e segurança. Esse sistema é capaz de atuar de maneira autônoma em diversas circunstâncias, podendo realizar manobras em estacionamento, dirigir com conhecimento de tráfego, dentre outros.(TESLA, INC., s.d.) Porém ainda sendo necessário a presença de um motorista atento e disponível para intervir no controle do veículo.

Outro exemplo é o Waymo, inicialmente sendo o projeto de veículos autônomos da Google, é um dos primeiros veículos a receberem licença para atuar de maneira autônoma em Phoenix e São Francisco nos Estados Unidos. Esse sistema foi um dos mais treinados até os dias de hoje, passando por milhões de milhas em percursos reais e bilhões de milhas em simulação. Esse sistema pode dirigir de maneira autônoma porém atuando em ambientes previamente conhecidos e mapeados.(WAYMO LLC, s.d.) Ambos veículos mencionados acima usam diversos sistemas para a coleta de informação do meio ao seu redor, desde lidars a câmeras, para não só identificar coisas ao seu redor mas também para fazer a análise de como será o comportamento das mesmas. Isso é fundamental para o desenvolvimento de um sistema seguro que irá atuar em um ambiente complexo como o do tráfego.

Dentro desse trabalho serão abordados estudos e simulações com intuito de desenvolver um sistema capaz de, por meio de técnicas de aprendizado de máquinas e visão computacional para identificação de trajeto e objetos ao entorno do veículo, fazer uma análise básica de seu estado de navegação.

Capítulo 1

Revisão Bibliográfica

Durante esse capítulo do trabalho será feita uma revisão com relação as técnicas utilizadas, de modo a verificar o estado da arte e principais técnicas dentro do campo de interesse desse trabalho. Essa revisão será feita abordando diversos conceitos aplicados a veículos autônomos com um foco principalmente dentro do campo de visão computacional, como aprendizado de máquina, navegação de veículos autônomos por meio de visão.

Essa revisão também irá tratar sobre as ferramentas usualmente aplicadas a desafios envolvendo veículos autônomos, como principais Frameworks de visão computacionais e os principais bancos de dados para aplicação em aprendizado de máquina, tanto na questão de imagens como de outros sensores utilizados. Assim revisão busca obter as ferramentas e técnicas necessárias para fazer o desenvolvimento futuro do trabalho.

1.1 Aprendizado de máquinas

O aprendizado de máquinas (Machine Learning (ML)) é uma área de estudos em crescente desenvolvimento em diversas áreas de aplicação, desde análise de grandes quantidades de dados para levantamento estatísticos até o tratamento de imagens. Esse crescimento se deve principalmente a um avanço na disponibilidade de grandes quantidades de dados e crescimento em poder computacional visto nos últimos anos, se tornando também uma área de estudo em destaque.

Porém esse conceito já vem sendo construído desde a década de 50, começando com a ideia de replicar o processo de aprendizado de conceitos abstratas como é feito pelos humanos. (BADILLO et al., 2020) O aprendizado consistiria assim na capacidade da máquina de perceber novos padrões, valores ou aprender maneiras de executar uma desejada tarefa. Assim em 1959 Arthur Samuel define aprendizado de máquina como o campo de estudos que consiste em dar a máquina a capacidade de executar uma tarefa sem uma programação explícita da mesma. (SAMUEL, 1967)

Um dos aspectos centrais do trabalho com ML consiste em seu treinamento, processo no qual são apresentados dados ao algoritmo com o intuito de obter um modelo capaz de identificar um padrão nos dados fornecidos. Dentro desse contexto pode se classificar a estrutura de ML em duas categorias principais, de acordo com (BADILLO et al., 2020).

- Aprendizado Supervisionado
 - Essa estrutura de aprendizado é composta pelo treinamento do sistema por meio de dados categorizados e com resultados esperados para o processamento dessas entradas, de modo que o sistema seja capaz de aprender modelos que relacionam entradas a determinadas saídas e possa futuramente obter resultados para novos dados desconhecidos.
 - Exemplos: Problemas de classificação e regressão.
- Aprendizado Não Supervisionado
 - Nesse aprendizado não se utiliza de dados já categorizado, com intuito de que o sistema seja capaz de encontrar padrões nesse conjunto de dados. Normalmente esse tipo de treinamento é utilizado quando não se tem essa categorização de dados definida antecipadamente ou se deseja não haver uma influencia prévia no padrão que a máquina irá encontrar.
 - Exemplos: Clusterização e redução de dimensão.

Com o desenvolvimento dos estudos no campo de ML, foram desenvolvidas sistemas computacionais baseados na estrutura biológica do cérebro humano, baseada em camadas de nódulos computacionais e com uma alta quantidade de interconexões entre eles, como mostrado na figura 1.1, esses sistemas chamados de Artificial Neural Network (ANN).