VISÃO COMPUTACIONAL X APRENDIZAGEM PROFUNDA: OS DIFERENTES TIPOS DE TREINAMENTO NO DESENVOLVIMENTO DE VEÍCULOS AUTÔNOMOS

Luiz Henrique Silva[1]
Marcos Antonio Jeremias Coelho[2]
João Mota Neto[3]
Fernando Guessi Placido[4]
Rafael Bernaldo [5]

Resumo: O desenvolvimento de veículos autônomos depende de técnicas avançadas de processamento de imagem e inteligência artificial, sendo a visão computacional e a aprendizagem profunda duas abordagens fundamentais. Enquanto a visão computacional utiliza algoritmos tradicionais para interpretar imagens e extrair características específicas, a aprendizagem profunda emprega redes neurais para reconhecer padrões complexos e tomar decisões de forma autônoma. Neste estudo, foram comparados diferentes modelos de treinamento aplicados a um carro autônomo: um baseado em visão computacional, programado para seguir uma linha amarela central na pista, e outro em aprendizagem profunda, com 5 tipos de treinamento por clonagem comportamental. Os testes mostraram que a visão computacional teve dificuldades em condições de iluminação variáveis, resultando em erros na navegação, enquanto a aprendizagem profunda demonstrou maior precisão e adaptabilidade, conseguindo completar voltas na pista de forma consistente. Os resultados indicam que, embora a visão computacional seja eficaz para tarefas específicas, sua rigidez a torna menos eficiente em ambientes dinâmicos, enquanto a aprendizagem profunda se mostra uma abordagem mais robusta e promissora para o treinamento de veículos autônomos.

Palavras-chave: Visão computacional, aprendizagem profunda, veículo autônomo, redes neurais.

1 INTRODUÇÃO

O avanço da tecnologia tem impulsionado o desenvolvimento de veículos autônomos, tornando-os cada vez mais viáveis para aplicação no trânsito urbano e rodoviário. A automação desse tipo de veículo depende de sistemas inteligentes capazes de interpretar o ambiente ao redor dele, detectar obstáculos, identificar vias

[1] Graduando em Engenharia Mecatrônica, 2025. E -mail: luiz.henrique@satc.edu.br
[2] Prof. do Centro Universitário UniSATC. e-mail: marcos.coelho@satc.edu.br
[3] Coord. da Eng. Mecatrônica UniSATC e-mail: joao.neto@satc.edu.br
[4] Fernando Guessi Placido Prof. do Centro Universitário UniSATCe-mail: fernando.placido@satc.edu.br
[5] Rafael Bernaldo Prof. do Centro Universitário UniSATC E-mail: rafael.bernaldo@satc.edu.br

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

e sinais de trânsito, e tomar decisões em tempo real. Nesse contexto, duas abordagens principais são amplamente utilizadas: a visão computacional e a aprendizagem profunda. Ambas desempenham papéis essenciais na percepção e navegação dos veículos autônomos, mas apresentam diferenças fundamentais em termos de funcionamento, precisão e adaptabilidade.

A visão computacional é uma área da inteligência artificial que busca desenvolver algoritmos capazes de extrair e interpretar informações visuais de imagens e vídeos. Tradicionalmente, essa abordagem utiliza técnicas de classificação de imagem, detecção de bordas, filtros de cores e transformações geométricas para identificar padrões visuais e converter essas informações em comandos para o veículo. No entanto, sua eficácia pode ser limitada por fatores como variações na iluminação, mudanças no ambiente e a necessidade de ajustes manuais para diferentes cenários.

Por outro lado, a aprendizagem profunda (Deep Learning) revolucionou a percepção computacional ao permitir que redes neurais convolucionais (CNNs) aprendam diretamente a partir de grandes volumes de dados, sem necessidade de regras pré-definidas. Essa abordagem tem mostrado um desempenho superior em tarefas complexas, como reconhecimento de objetos, identificação de regiões e tomada de decisão baseada em padrões aprendidos. No contexto dos veículos autônomos, a aprendizagem profunda é frequentemente aplicada por meio da clonagem comportamental, onde o sistema aprende a dirigir imitando as ações humanas, refinando suas decisões à medida que é exposto a mais dados de treinamento.

Diante dessas diferenças, este artigo tem como objetivo comparar as abordagens de visão computacional e aprendizagem profunda no treinamento de um veículo autônomo, avaliando suas vantagens, desafios e limitações. Para isso, foram conduzidos experimentos utilizando um carro autônomo em uma pista de testes, onde o modelo de visão computacional identificava e seguia uma linha amarela, enquanto outro foi treinado por cinco diferentes tipos de aprendizagem profunda com base em um conjunto de dados coletado previamente. Os resultados obtidos demonstram a influência das condições do ambiente e da metodologia de treinamento no desempenho dos modelos, fornecendo uma análise detalhada sobre a aplicabilidade de cada abordagem na navegação autônoma.

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

Este estudo contribui para uma melhor compreensão sobre o uso de visão computacional e aprendizagem profunda no contexto de veículos autônomos, destacando como essas tecnologias podem ser integradas para aprimorar a segurança e a eficiência dos sistemas de direção autônoma.

2 VISÃO COMPUTACIONAL

A visão computacional é uma área da ciência que estuda como as máquinas podem “enxergar” e interpretar o mundo ao seu redor, extraindo informações significativas a partir de imagens capturadas por câmeras e outros dispositivos. Estes estudos surgiram na década de 1950 e se desenvolveram partir de 1970, quando passou a ser associada à Inteligência Artificial. Mesmo com expectativas de uma rápida evolução, os primeiros estudos mostraram que a tarefa era mais complexa do que se imaginava. A visão computacional busca criar sistemas que imitem a capacidade humana de perceber objetos e agir de maneira inteligente, aproximando-se da forma como o cérebro interpreta imagens [1].

Em [2] nos é apresentado um conceito de visão computacional que nos possibilita criar um piloto automático empregando algumas técnicas tradicionais de visão, como algoritmos de detecção de bordas e conversão de espaço de cores, identificando características em imagens capturadas pela câmera, transformando as informações obtidas em comandos de direção e aceleração. Este sistema é vantajoso pois não necessita de intervenção manual para coleta de dados, assim, o usuário pode desenvolver seu próprio algoritmo de visão computacional, ajustando parâmetros que se adaptem a cada situação.

3 APRENDIZAGEM PROFUNDA

As redes neurais de Aprendizagem Profunda (Deep Learning) são estruturas avançadas de redes neurais que se destacam por possuírem várias camadas ocultas, o que as diferencia das redes tradicionais, que geralmente têm uma ou nenhuma camada oculta. Elas são essenciais na construção de sistemas inteligentes devido à sua habilidade superior de aprender e processar grandes volumes de dados, superando as Redes Neurais convencionais. Atualmente, quase todos os sistemas que utilizam inteligência artificial fazem uso do Deep Learning, uma

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

vez que a evolução do poder computacional tornou possível a implementação dessa tecnologia. No passado, a aplicação de Deep Learning era restrita pela capacidade dos computadores, que não conseguiam processar a quantidade de cálculos necessários para a execução, avaliação e ajuste dos pesos nas redes neurais [3].

O piloto automático baseado em aprendizado profundo utiliza uma câmera frontal e uma rede neural convolucional para implementar um sistema de direção autônoma, empregando uma técnica chamada Clonagem Comportamental, também conhecida como Aprendizado por Imitação. O nome "Clonagem Comportamental" reflete o objetivo de criar um piloto automático capaz de replicar as ações humanas. Como o piloto automático baseado em aprendizado profundo, este, depende de imagens capturadas pela câmera, portanto as condições de iluminação desempenham um papel crucial. O modelo de aprendizado profundo se adapta bem a pistas internas, onde é possível controlar as condições de iluminação e os elementos do ambiente. No entanto, pode ser mais desafiador fazer o sistema funcionar corretamente em ambientes externos, onde as condições de iluminação são imprevisíveis e o cenário pode mudar constantemente [2].

4 VISÃO COMPUTACIONAL X APRENDIZAGEM PROFUNDA

A principal diferença entre as duas abordagens está na forma como elas tratam e processam as informações. Enquanto a visão computacional pode envolver abordagens mais convencionais de extração de características específicas e regras programadas, a aprendizagem profunda oferece uma abordagem mais flexível e autônoma, onde os sistemas "aprendem" com exemplos. Em veículos autônomos, isso significa que a visão computacional pode ser usada em conjunto com a aprendizagem profunda para aprimorar a precisão do reconhecimento de objetos e cenários, enquanto a aprendizagem profunda pode lidar com decisões de maior nível, como antecipar movimentos de outros veículos e otimizar rotas. Ambas as tecnologias são essenciais para o treinamento e a operação eficiente de veículos autônomos, mas sua interação permite que o sistema alcance um nível de autonomia e segurança mais avançado [4].

Além disso, a visão computacional no treinamento de veículos autônomos pode ser vista como uma abordagem mais orientada à tarefa específica e direta. Ela

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

é baseada em algoritmos definidos manualmente que procuram identificar objetos específicos e realizar funções bem delineadas, como segmentar imagens e rastrear movimentos. Embora muito eficaz para tarefas simples e bem definidas, como identificar um sinal de trânsito ou um pedestre, ela pode ser menos eficiente em situações mais complexas, que exigem um entendimento mais profundo e generalizado do ambiente. A visão computacional, ao depender de regras rígidas, pode ter dificuldades em lidar com condições imprevisíveis, como mudanças abruptas no cenário ou variações nas condições de iluminação [4].

Por outro lado, a aprendizagem profunda oferece uma abordagem mais robusta e adaptativa para o treinamento de carros autônomos. Ela permite que o sistema aprenda a partir de grandes volumes de dados de diferentes cenários, adaptando-se a condições novas e imprevistas sem a necessidade de intervenção humana constante. Isso torna a aprendizagem profunda mais adequada para tarefas de alto nível, como a tomada de decisões complexas e a previsão do comportamento de outros veículos no tráfego. Além disso, a aprendizagem profunda permite que os veículos autônomos se tornem mais precisos à medida que são expostos a mais dados, ajustando-se continuamente para melhorar a segurança e a eficiência na navegação [4]. Em conjunto com a visão computacional, a aprendizagem profunda pode complementar e melhorar os sistemas de percepção, tornando os veículos mais autônomos e capazes de operar de forma eficaz em uma variedade de ambientes dinâmicos.

5 MATERIAIS E MÉTODOS

Os materiais e métodos utilizados neste estudo são descritos a seguir para garantir a reprodutibilidade dos resultados. O modelo de carro utilizado está disponível no site da Donkey Car, e pode ser acessado para melhor entendimento. Para a construção do carro autônomo foram utilizados alguns componentes, sendo eles: base DHK Wolf 2, Raspberry PI 5 8gb, Câmera 5mp wide angle 170°, Módulo PCA 9685, Powerbank 10.000mA/h, Bateria NIMH 6 células 7.2V, acelerômetro e giroscópio MPU 6050, display OLED SSD 1306, roteador Wi-Fi Link One e por fim um joystick (controle) Knup GM034. Para a criação da pista a qual os testes foram executados, foram usadas fitas nas cores branca e amarela, para que pudessem ser feitas as

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

linhas da esquerda e da direita (brancas) e a linha central (amarela), já para ostestes, foi necessária a medição do tamanho da pista com uma trena, e um dois cronômetros, uma para calcular o tempo necessário para o carro concluir o objetivo, e outro para calcular quanto tempo ele esteve fora dos padrões propostos. Além dos componentes físicos, foi utilizado o software Remmina, necessário para criar as redes neurais de treinamento do veículo autônomo, tanto para visão computacional, quanto para aprendizagem profunda.

Fig. 1. Vista dos componentes do carro.

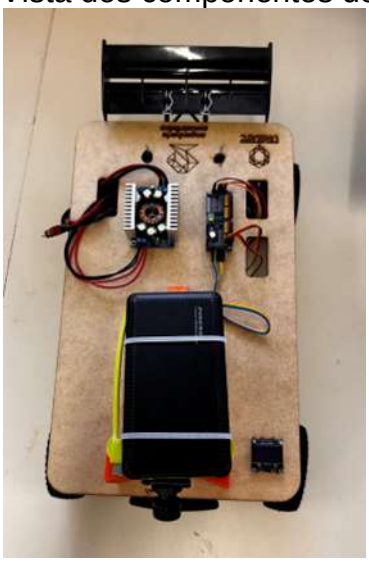

Fig. 2. Modelo de carro utilizado

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

Fig. 3. Modelo da pista utilizada

Como dito anteriormente, os métodos de treinamento utilizados foram visão computacional e aprendizagem profunda. No site Donkey Car é disponibilizado um comando de criação de um carro com características de rede neural de visão computacional, dessa forma, ao conectar o Remmina, introduzimos este comando, onde um algoritmo de dados padrão da Donkey será enviado ao veículo, neste algoritmo, o carro entenderá que deve seguir caminho por uma linha amarela, identificando os limites da pista como uma linha branca, e entendendo a trajetória que deve seguir. No processo de treinamento por visão computacional não há necessidade de um ser humano controlar o carro, deste modo, não é utilizado o joystick, apenas um código de programação é passado ao carro, onde este reúne informações através da câmera, e tem a capacidade de dirigir autonomamente. Para o treinamento por aprendizagem profunda, o processo será parecido, também utilizando o Remmina, porém enviando um outro comando, de criação de um carro de rede neural de aprendizagem profunda, que também está disponível no site da Donkey. Posteriormente, é necessário conectar o joystick ao veículo, e treiná-lo dando voltas na pista durante algum tempo, com ajuda da câmera, ele arrecadará dados que serão utilizados para o seu treinamento, para isso, é necessário arrecadação de 28.000 a 30.000 dados, com esses valores já é possível ter um veículo bem treinado. Após a coleta dos dados, devemos abrir o terminal do computador (nesse caso foi utilizado o sistema operacional Linux) e fazer a sintetização dos dados obtidos, no veículo, dessa forma, o veículo já estará pronto para rodar de forma autônoma.

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

6 RESULTADOS E DISCUSSÕES

Depois de empregar as metodologias descritas acima, foi possível obter certos resultados, alguns satisfatórios, outros nem tanto.

6.1 TESTE DE VISÃO COMPUTACIONAL

O treinamento da rede neural de visão computacional apresentou acertos e erros. Como dito anteriormente, nesse método de treinamento o carro identifica uma linha amarela central, que através da programação passada a ele, este entende que deve seguir caminho pela linha, sendo assim, se faz necessário um traçado vívido e uma boa luminosidade, para que desse jeito, o carro “enxergue” de maneira correta o trajeto que deve seguir. Os erros durante o treinamento ocorreram no momento de colocar o veículo para andar de forma autônoma, justamente por conta desses dois fatores citados, a pouca vivacidade da cor amarela da linha central, e a iluminação ruim no laboratório onde foram feitos os testes, fizeram com que o carro não reconhecesse de forma correta o trajeto ao qual deveria seguir. Muitas vezes, na percepção de sua câmera, foi possível observar alguns “falsos positivos”, ou seja, o veículo identificava amarelo em locais onde não existia, e dessa forma, não seguia o caminho, pois não sabia o local correto por onde andar. Ao desligar as luzes do laboratório percebeu-se uma certa melhora nas condições de trafegabilidade do automóvel, que por sua vez, andou sozinho por alguns metros na pista, antes de parar totalmente.

Fig. 4. Teste de visão computacional com as luzes acesas.

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

Fig. 5. Teste de visão computacional com as luzes apagadas.

Na figura 1 observa-se os falsos positivos citados anteriormente. Nota-se que a câmera identifica muitas áreas em branco, fazendo com que o veículo não entenda o caminho a seguir por identificar que estas áreas são parte da linha amarela.

Já na figura 2 por exemplo, não ocorrem muitos falsos positivos por conta de as luzes estarem apagadas. O carro identifica a linha amarela de melhor forma, e consegue dirigir até certo ponto da pista, porém ao chegar nas curvas, não conseguia fazer a identificação correta, parando de forma completa.

Para experimentos futuros, é notável a necessidade de um bom local para testes de visão computacional, pois, mesmo com as condições ruins, o carro apresentou parte dos resultados satisfatórios ao andar corretamente durante uma parte da pista, mesmo que tenha sido por poucos metros.

6.2 TESTE DE APRENDIZAGEM PROFUNDA

No teste de aprendizagem profunda os resultados encontrados foram satisfatórios dentro do objetivo proposto. Para começar, foi necessária a captação manual de dados para a rede neural, onde foi conectado ao carro um joystick, transformando-o em um carro de controle remoto, dessa forma, durante alguns minutos, foi mostrado ao veículo o que deve ser feito, controlando-o durante todo o trajeto, até arrecadar cerca de 30.000 dados, que posteriormente foram treinados diretamente no terminal do computador. Diferentemente da visão computacional, a

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

aprendizagem profunda não segue uma linha de cor específica, ele segue dados acumulados no treinamento, ou seja, enquanto o carro estava sendo controlado manualmente, a sua câmera captava informações de extremidades da pista, e o principal trajeto por onde deveria andar. Foram adotados dois modelos de captação de dados, no primeiro, o veículo deveria ficar dentro das extremidades das linhas brancas, e seguir o mais próximo possível à linha amarela, chamado de "CarroLinhaAmarela", já no segundo o veículo deveria seguir pela pista da direita, entre as linhas amarela e branca, assim como é feito no trânsito da maioria dos países do mundo, e foi chamado de "CarroEntreLinhas".

No site do Donkey Car são disponibilizados alguns tipos de treinamento para o modelo de aprendizagem profunda, esses modelos são chamados de "Keras", e foram utilizados 5 deles para os testes, sendo estes, Linear, Categorical, Inferred, 3D e RNN. O objetivo principal dos testes para esses modelos de treinamento era que o carro pudesse completar ao menos 10 voltas no percurso para ser considerado um sucesso, mesmo que este tenha cometido erros durante o trajeto. Para estabelecer uma comparação justa entre os veículos foram observadas três estatísticas durante o percurso, sendo elas, o tempo que o carro leva para completar as 10 voltas, a velocidade média do percurso, e a eficiência do automóvel, que é calculada estabelecendo uma relação entre o tempo total em que o veículo esteve dentro dos padrões estabelecidos, sobre o tempo total em que ele esteve no percurso. A partir desses dados, pôde-se inferir uma equação para o cálculo do desempenho de cada um dos modelos estudados.

6.2.1 Modelo de treinamento Linear

O modelo de treinamento Linear é o mais simples dentre os 5 utilizados, ele utiliza uma abordagem simples e eficiente para controlar um carro autônomo, com uma rede neural básica que aprende a partir de imagens capturadas pela câmera do veículo. A rede é treinada para prever dois valores contínuos: direção e aceleração, com base nas imagens de entrada. O modelo usa uma arquitetura com camadas convolucionais seguidas por camadas densas, e a ativação linear permite a previsão direta dos controles. Esse método é vantajoso por ser fácil de implementar e eficiente,

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

funcionando bem em plataformas com recursos limitados, mas pode ter limitações, como dificuldades para aprender a aceleração em alguns casos.

Como expresso anteriormente, foram propostos dois modelos de captação de dados, sendo eles o “CarroLinhaAmarela”, e o “CarroEntreLinhas”.

- CarroLinhaAmarela:

Para este modelo de treinamento os resultados foram satisfatórios, uma vez que, o carro completou as 10 voltas na pista tendo uma taxa de erro muito baixa. Ao considerarmos que o percurso tem uma distância de 13,75m, e o veículo levou 2 minutos e 41 segundos para completar as 10 voltas, podemos estimar que ele estava a uma velocidade média de 0,85m/s. Durante esse tempo, o veículo saiu do traçado da linha amarela apenas 22 vezes, e se manteve sempre dentro dos limites da pista, apresentando um tempo de correção baixíssimo, entre 0,5 e 1 segundo para voltar a seguir a linha. O tempo total em que o veículo esteve fora da linha amarela foi de aproximadamente 16 segundos, o que representa 9,94% do tempo do percurso.

A partir da sexta volta, em uma das retas do circuito, o carro começou a oscilar de forma mais brusca entre as pistas da direita e da esquerda, essas oscilações se fizeram refletir no tempo que o carro demorou para completá-la, levando 16,63 segundos e sendo a mais lenta entre as 10, diferentemente da primeira onde este levou 15,79 segundos, sendo a melhor entre todas. No geral, as cinco primeiras voltas foram mais rápidas do que as cinco últimas, por conta das oscilações que começaram a acontecer.

Os testes para esse modelo foram um sucesso dentro do objetivo imposto, tendo um desempenho dentro do esperado para aquilo que lhe foi proposto.

- CarroEntreLinhas:

O veículo completou as 10 voltas com 100% de eficiência, demonstrando ser um completo sucesso, mantendo-se sempre dentro da pista e sem invadir a faixa da esquerda em nenhum momento. Mostrou grande precisão ao contornar as curvas, garantindo um ótimo desempenho técnico. Além disso, registrou um tempo de 5 minutos e 34 segundos para concluir o percurso, mantendo uma velocidade média de aproximadamente 0,41m/s, fazendo sua melhor volta em 32,18 segundos, e a pior em um tempo de 34,51 segundos, com esses resultados observa-se que o treinamento

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

conseguiu manter um padrão, não havendo grande discrepância de tempo entre as voltas pelo circuito.

6.2.2 Modelo de treinamento Categorical

O modelo Categorical do Donkey Car utiliza uma rede neural para classificar a direção do veículo em categorias discretas, tornando a condução mais estável e menos sensível a ruídos. O treinamento envolve a coleta de imagens e comandos, a conversão dos ângulos em classes e o aprendizado por meio de uma CNN. Esse modelo pode apresentar algumas desvantagens por conta da discretização dos ângulos, reduzindo a precisão, especialmente em curvas suaves, onde a transição entre categorias pode não ser tão fluida quanto em um modelo contínuo.

- CarroLinhaAmarela:

Dentre todos os modelos, o Categorical foi o que apresentou o pior desempenho, uma vez que, o veículo completou apenas duas voltas antes de perder o trajeto na primeira curva da terceira volta, ao seguir reto. Durante essas duas voltas, saiu da linha amarela 16 vezes e ultrapassou os limites das linhas brancas em duas ocasiões antes de sair completamente da pista. O tempo de correção das falhas variou entre 1 e 2,5 segundos, e o veículo conseguiu permanecer no trajeto por um total de 43,33 segundos antes de não conseguir mais seguir corretamente. O carro ficou fora da linha amarela durante um período de aproximadamente 26 segundos, o que representa 60% do tempo do trajeto, sendo assim, o teste para esse modelo não apresentou resultados satisfatórios, pois este se manteve a maior parte do tempo fora da linha amarela, não conseguindo completar as 10 voltas propostas.

- CarroEntreLinhas:

A obtenção de dados para o modelo do CarroEntreLinhas foi feita em uma baixa velocidade, e isso impactou diretamente em todos os treinamentos. Quando operado em modo 100% autônomo, o veículo não teve força de aceleração suficiente para se mover, mesmo tendo sido testado em velocidades equivalentes às dos demais modelos de treinamento, que na ocasião, deram certo. Isto evidencia a falha do

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

treinamento Categorical para este modelo, visto que não se faz possível a captação de estatísticas de tempo, velocidade ou eficiência, já que o veículo foi se quer capaz de se mover no circuito, sendo considerado então um fracasso dentro do objetivo proposto.

6.2.3 Modelo de treinamento RNN

O modelo RNN do Donkey Car utiliza redes neurais recorrentes (RNNs) para prever a direção do veículo com base em sequências de dados temporais, permitindo uma condução mais fluida e adaptativa. Diferente dos modelos baseados apenas em CNNs, o RNN leva em consideração o histórico recente de comandos e imagens, melhorando a estabilidade da direção e antecipando curvas de forma mais eficiente. O treinamento envolve a captura de imagens e comandos ao longo do tempo, processados por camadas recorrentes que aprendem padrões temporais na navegação. Esse modelo é particularmente útil para trajetórias complexas, pois consegue prever ajustes de direção com maior coerência, reduzindo oscilações bruscas e tornando a condução mais natural e estável.

- CarroLinhaAmarela:

O modelo de treinamento RNN demonstrou ser o mais preciso dentre todos os treinamentos, em todos os aspectos, tempo, precisão e correções. O veículo completou 10 voltas mantendo um padrão de qualidade consistente, saindo da linha amarela uma vez por volta, sempre no mesmo ponto. As correções foram rápidas, variando entre 0,5 e 1 segundo, garantindo que não ultrapassasse os limites da pista. O tempo total para concluir as 10 voltas foi de 2 minutos e 33 segundos, demonstrando um desempenho estável e controlado ao longo do percurso. A velocidade média desempenhada pelo veículo foi de aproximadamente 0,90m/s, e pôde-se perceber que o carro melhorou constantemente ao longo do percurso, uma vez que o seu pior desempenho aconteceu na primeira volta, obtendo um tempo de 16,25 segundos, e seu melhor desempenho se deu justamente na última volta, onde fez um tempo de 14,77 segundos. O veículo esteve fora da linha amarela por aproximadamente 7 segundos, o que representa por volta de 4,6% do período em que completou as 10

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

voltas, dessa forma, é possível inferir que o teste proposto foi um sucesso, e o treinamento demonstrou resultados muito satisfatórios.

- CarroEntreLinhas:

O modelo completou as 10 voltas do circuito com poucos erros, apresentando um desempenho estável ao longo do trajeto. Em uma das curvas, invadiu a pista da esquerda em oito ocasiões, porém realizou correções rápidas, com tempo médio de ajuste de aproximadamente meio segundo. O tempo total para a conclusão das 10 voltas foi de 6 minutos e 59 segundos, fazendo um tempo de 39,27 segundos em sua volta mais rápida e 45,36 segundos na pior volta, mantendo uma velocidade média do percurso de aproximadamente 0,33m/s. Durante as curvas, houve uma desaceleração perceptível, influenciada tanto pela baixa velocidade do veículo quanto pelo esterçamento das rodas, que contribuiu para a redução ainda maior da velocidade. No geral o modelo apresentou uma grande eficiência, de 98,81%, ficando apenas 1,19% do tempo fora das extremidades permitidas, desta forma, pode-se inferir que o modelo é um sucesso dentro do objetivo a ele imposto.

6.2.4 Modelo de treinamento 3D

O treinamento 3D do Donkey Car utiliza redes neurais convolucionais tridimensionais (3D CNNs) para processar sequências de frames de vídeo, permitindo uma análise temporal do movimento do veículo em miniatura. Diferente dos modelos tradicionais 2D, que avaliam imagens isoladas, essa abordagem considera múltiplos quadros consecutivos, melhorando a previsibilidade e a suavidade das decisões de direção. O modelo é treinado com dados coletados manualmente, associando imagens capturadas pela câmera do carro aos comandos correspondentes de direção e aceleração. Essa técnica reduz a sensibilidade a ruídos momentâneos e melhora a adaptação a diferentes condições de pista, tornando a navegação autônoma mais estável e eficiente. Apesar disto, a necessidade de um grande volume de dados para um treinamento eficaz pode tornar o processo demorado e complexo, aumentando a dificuldade de ajustes finos no modelo. A latência também pode ser um problema, pois a análise de múltiplos frames exige maior capacidade de processamento, o que pode impactar a resposta do carro em situações dinâmicas. Por fim, a complexidade do

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

modelo pode levar a overfitting(excesso de dados), dificultando a generalização para diferentes condições de pista sem um treinamento mais robusto e diversificado.

- CarroLinhaAmarela:

Este modelo de treinamento apresentou uma grande instabilidade, porém, ainda assim conseguiu completar as 10 voltas propostas, em um tempo total de 3 minutos e 01 segundo. Durante o percurso, saiu da linha amarela 37 vezes, sendo que em 6 dessas ocasiões ultrapassou os limites da pista, sempre no mesmo ponto crítico: a primeira curva. A correção de trajetória apresentou um tempo de resposta elevado, variando entre 4 e 5 segundos, o que impactou a estabilidade e o desempenho geral do carro ao longo das voltas. A velocidade média do carro foi de aproximadamente 0,76m/s, este apresentou instabilidade durante todas as 10 voltas, fazendo seu pior tempo na primeira delas, com 19,32 segundos, e o melhor durante a sétima volta, com 17,57 segundos. O carro autônomo esteve fora da linha amarela por aproximadamente 165 segundos, o que infere uma porcentagem de 91% do tempo fora do trajeto da linha amarela, sendo considerada uma precisão muito ruim para o modelo. Ainda assim, mesmo com a taxa crítica de erro, o método 3D conseguiu completar o trajeto proposto.

- CarroEntreLinhas:

O modelo completou as 10 voltas de forma altamente eficiente, demonstrando um desempenho preciso e estável ao longo do trajeto. Houve apenas uma invasão da pista contrária, que durou um segundo, sendo rapidamente corrigida sem comprometer a trajetória. O tempo total para concluir as 10 voltas foi de 8 minutos e 17 segundos, sendo o maior entre todos os modelos, sua volta mais rápida foi de 44,49 segundos, enquanto a mais lenta foi de 59,68 segundos, demonstrando uma certa discrepância entre as voltas e mantendo uma velocidade média de aproximadamente 0,27m/s. Além disso, apresentou grande precisão nas curvas, com pouca desaceleração, evidenciando um controle refinado e um comportamento consistente durante a navegação no circuito. Apesar do longo tempo para concluir o trajeto, o modelo demonstrou uma grande eficiência, ficando dentro das extremidades propostas durante 99,8%, já que invadiu a pista contrária apenas uma vez, por um

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

segundo, representando 0,2% do tempo. Dentro do que foi proposto, podemos considerar que este modelo é um sucesso, porém lento em comparação aos demais.

6.2.5 Modelo de treinamento Inferred

O modelo Inferred do Donkey Car é um sistema baseado em aprendizado profundo que permite a condução autônoma de veículos utilizando redes neurais convolucionais (CNNs). Ele opera também no paradigma de aprendizado supervisionado, onde a rede é treinada a partir de dados coletados durante a condução manual do veículo. A entrada do modelo consiste em imagens capturadas pela câmera do carro, e a saída são dois valores correspondentes ao ângulo de direção e à aceleração. A arquitetura típica inclui camadas convolucionais, que extraem padrões visuais das imagens, e camadas densas totalmente conectadas, responsáveis por interpretar essas características e gerar os comandos corretos. Para esse tipo de treinamento foi necessária a utilização de um joystick em que o próprio operador estivesse no controle da velocidade, com o carro fazendo apenas o controle da direção, pois o modelo Inferred por padrão mantém uma velocidade elevada quando em estado 100% autônomo, acima da máxima suportada(0,73m/s para o CarroLinhaAmarela, e 0,61m/s para o CarroEntreLinhas) para que o veículo faça o trajeto de forma correta, por conta disto, a captação de estatísticas referentes a esse modelo se tornava inviável, uma vez que, por conta de sua alta velocidade, o veículo acabava por não conseguir fazer curvas, chocando-se contra alguns objetos presentes no laboratório onde os testes foram realizados, e danificando sua estrutura.

- CarroLinhaAmarela:

Neste método de treinamento o veículo completou 10 voltas na pista, saindo da linha amarela 9 vezes apenas, principalmente após curvas, mas sem ultrapassar os limites da pista. O tempo médio de correção foi de 1 a 2 segundos, e, ocasionalmente, o sistema esterçava o volante para o lado errado, porém corrigindo instantaneamente, sem comprometer o percurso. O tempo total para completar as 10 voltas foi de 3 minutos e 08 segundos, mantendo uma velocidade média de 0,73m/s, o que podemos confirmar ao observar os tempos de melhor e pior volta, 18,36 segundos e 20,09 segundos respectivamente, caracterizando a menor das

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

velocidades entre todos os testes para esse modelo, porém com uma grande eficiência ao se manter em cima da linha amarela central, estando fora dela apenas 12,5 segundos aproximadamente, o que representa 6,65% do tempo que levou para completar as 10 voltas, demonstrando ser o segundo método mais eficiente de treinamento.

- CarroEntreLinhas:

O modelo completou as 10 voltas de forma rápida e eficiente, com apenas 8 invasões na pista contrária, cada uma durando um curto período, geralmente cerca de meio segundo, e corrigindo seu curso de maneira ágil. O tempo total para a conclusão das 10 voltas foi de apenas 3 minutos e 53 segundos, demorando 25,09 segundos em sua melhor volta, e 26,42 na pior, o que mostra a constância que o veículo teve durante o circuito, e com uma velocidade média de 0,59m/s, o que o tornou o modelo mais rápido entre os testados. Durante o percurso, executou as curvas com grande precisão e sem apresentar desaceleração, evidenciando um excelente desempenho e controle ao longo do trajeto. Sua eficiência foi ótima, estando 98,28% do tempo dentro dos limites da pista, e apenas 1,72% fora deles. Sendo assim, pode-se inferir que o modelo teve um grande resultado positivo, considerando seu tempo baixo e grande eficiência.

6.2.6 Comparação entre os modelos

A análise comparativa entre os cinco modelos de treinamento aplicados ao modelo da linha amarela permitiu observar diferenças significativas em termos de precisão, estabilidade, tempo de percurso e comportamento em situações críticas.

Para determinar qual dos veículos obteve o melhor desempenho, foi formulada uma equação que relaciona eficiência, tempo total do percurso e velocidade média dos carros no trajeto.

$$Des = 0{,}6 \cdot E/Emax + 0{,}25 \cdot V/Vmax + 0{,}15 \cdot (1 - T/Tmax) \quad (1)$$

O desempenho varia de 0 a 1, sendo 0 o nível mais baixo e 1 o mais alto, para isso, foram definidos pesos para as variáveis de acordo com o grau de

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

importância de cada uma delas, sendo assim, a eficiência recebeu um peso maior, de 0,6 enquanto a velocidade e o tempo receberam tempos menores, de 0,25 e 0,15 respectivamente. Os valores de E e Emax simbolizam a eficiência, e a eficiência máxima atingida por um modelo naquele teste, respectivamente, e os valores de V e Vmax além de T e Tmax seguem o mesmo padrão, porém para velocidade e tempo.

De acordo com a equação do desempenho, a ordem de melhores modelos para o teste da linha amarela ficou da seguinte forma, RNN demonstrando o melhor desempenho, com valor de referência em 0,877, seguido pelos modelos Linear e Inferred, com 0,824 e 0,789 respectivamente. O que mais chama atenção neste treinamento é o valor de referência encontrado para o modelo 3D, que por conta de sua baixa eficiência, acabou tendo o pior desempenho entre todos os treinamentos, 0,273.

O teste entre linhas trouxe alguns resultados diferentes, onde o modelo Inferred acabou por ser o de maior destaque, com desempenho de 0,919, demonstrando uma grande confiabilidade, os demais modelos não tiveram desempenhos ruins, porém ficaram um pouco abaixo deste citado anteriormente, Linear obteve um valor de desempenho de 0,823, enquanto RNN e 3D obtiveram desempenhos de 0,770 e 0,713 respectivamente.

Os resultados indicam que os modelos Inferred e RNN são os mais promissores, destacando-se pelo equilíbrio entre agilidade, precisão e capacidade de correção, o que os torna particularmente adequados para aplicações em ambientes dinâmicos e com exigência de navegação contínua. O modelo Linear por sua vez demonstrou grande eficiência em ambos os casos, e pode muito bem ser utilizado para muitas aplicações no cenário de mobilidade autônoma. Por fim, o modelo 3D apresentou o pior desempenho nos dois testes, porém, ainda demonstra ser um ótimo tipo de treinamento para diferentes tipos de ambientes.

7 CONCLUSÃO

O desenvolvimento de veículos autônomos exige a implementação de tecnologias avançadas de percepção do ambiente, sendo a visão computacional e a aprendizagem profunda duas abordagens essenciais nesse contexto. Ao longo deste estudo, comparou-se o desempenho dessas técnicas no treinamento de um carro

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424

autônomo, avaliando suas vantagens, limitações e aplicabilidade em um certo cenário. Os resultados demonstraram que a visão computacional, apesar de eficiente para tarefas bem definidas, apresenta dificuldades em condições adversas, como variações de iluminação e mudanças inesperadas no ambiente. Em contrapartida, a aprendizagem profunda mostrou-se mais robusta e adaptável, sendo capaz de aprender padrões complexos e operar com maior precisão e consistência.

Embora a aprendizagem profunda tenha se destacado nos testes, a visão computacional ainda desempenha um papel relevante no processamento e interpretação de imagens em sistemas de direção autônoma. A combinação dessas abordagens pode resultar em um sistema híbrido mais eficiente, onde a visão computacional atua na extração inicial de características e a aprendizagem profunda aprimora a tomada de decisão. Assim, este estudo reforça a importância de explorar soluções integradas para maximizar a segurança e a eficiência dos veículos autônomos.

Por fim, esta pesquisa pode servir de base para estudos futuros voltados à otimização das técnicas analisadas, buscando reduzir suas limitações e aprimorar sua aplicação em cenários reais. A evolução dessas tecnologias é essencial para viabilizar a implementação segura e confiável dos veículos autônomos, aproximando-os cada vez mais da realidade do transporte urbano atual.

REFERÊNCIAS

[1] de Milano, Danilo, and Luciano Barrozo Honorato. "Visão computacional." Faculdade de Tecnologia, Universidade Estadual de Campinas (2010).

[2] Donkey Car, "Train autopilot," Donkey Car. [Online]. Disponível em: https://docs.donkeycar.com/guide/train_autopilot/.

[3] Falcão, João Vitor Regis, et al. "Redes neurais deep learning com tensorflow." RE3C-Revista Eletrônica Científica de Ciência da Computação 14.1 (2019).

[4] Janai, J., Güney, F., Behl, A., & Geiger, A. (2020). Computer Vision for Autonomous Vehicles: Problems, Datasets and State-of-the-Art. Foundations and Trends in Computer Graphics and Vision.

DOI: https://doi.org/10.70185/2525-6025.2025.v10.424