

MACHINE LEARNING PARA CLASSIFICAÇÃO DE IMAGENS UTILIZANDO JAVASCRIPT

Debora Pereira de Freitas Gonçalves¹
Ediloy Guilherme Dolinski²
Eduardo Afonso de Cristo³
Higor Pinto Moraes⁴
Pedro Henrique Cavalcante da Rosa⁵
Vinícius Tessele⁶

Resumo: A utilização de *Machine Learning* está cada dia mais presente em nosso cotidiano, estudar a sua criação bem como suas aplicações, são importantes para entender todo o processo que leva os computadores a conseguirem aprender. O objetivo desta pesquisa é avaliar como o algoritmo desempenha seu papel e encontrar formas de aplicar a ferramenta de reconhecimento de imagem no dia a dia, com uma interface web, de fácil manuseio. O processo de implementação da rede neural utilizou recursos de HTML e *JavaScript*, que implementa a biblioteca ML5, Mobile-Net, software de aprendizado de máquina, que utiliza o *TensorFlow*, nas suas camadas internas. O resultado mostra que foi possível realizar o reconhecimento das imagens selecionadas com um grau de precisão elevado, e o desenvolvimento é facilitado pelo uso da biblioteca, também pode se observar que a biblioteca utilizada possui a possibilidade de realizar o treinamento do algoritmo para, sendo assim possível a utilização para os mais diversos fins.

Palavras-chave: Machine Learning. Inteligência artificial. JavaScript. Classificação de imagens.

1 INTRODUÇÃO

Machine learning, conforme é descrito por Soofi e Awan (2017), é uma área da inteligência artificial que utiliza dos preceitos da ciência da computação, estatística, ciência cognitiva, engenharia, teoria da otimização e outras áreas da matemática e ciência. Ainda de acordo com o autor, o aprendizado de máquina pode ser categorizado em duas principais categorias: o *machine learning* supervisionado e o não supervisionado. O modelo não supervisionado retira conclusões do conjunto de dados de entrada, sem possuir uma resposta rotulada, em outras palavras este algoritmo descobre padrões escondidos nos dados sem a

¹Acadêmico do Curso Tecnologia em Análise e Desenvolvimento de Sistemas - TADS da UNIVEL – Centro Universitário UNIVEL.

²Acadêmico do Curso Tecnologia em Análise e Desenvolvimento de Sistemas - TADS da UNIVEL – Centro Universitário UNIVEL.

³Acadêmico do Curso Tecnologia em Análise e Desenvolvimento de Sistemas - TADS da UNIVEL – Centro Universitário UNIVEL.

⁴Acadêmico do Curso Tecnologia em Análise e Desenvolvimento de Sistemas - TADS da UNIVEL – Centro Universitário UNIVEL.

⁵Acadêmico do Curso Tecnologia em Análise e Desenvolvimento de Sistemas - TADS da UNIVEL – Centro Universitário UNIVEL.

⁶ORIENTADOR: Professor Mestre Vinicius Tessele Curso Tecnologia em Análise e Desenvolvimento de Sistemas – TADS da Univel – Centro Universitário UNIVEL.

necessidade de intervenção humana. Já o modelo supervisionado, é uma abordagem que utiliza de grandes quantidades de dados rotulados, estes dados designados ajudam o software a ser treinado, de modo que ele possa classificar dados ou prever resultados com precisão.

Ainda dentro do modelo supervisionado, de acordo com o mesmo autor, existem técnicas, que se dividem em duas principais categorias, a de classificação e a de regressão. Classificação é um modelo que utiliza de *data mining* (mineração de dados), para prever se um dado específico pertence ou não a um grupo de instância de dados, técnica largamente utilizada, quando se trata de *machine learning*. No entanto, a falta de valores em dados no *dataset* pode causar diversos problemas tanto no treinamento, quanto na fase de classificação do *software*, essa falta de valores pode acontecer devido a equívocos, dados reconhecidos como irrelevantes, desvios de dados com outros dados, ou até mal funcionamento do equipamento.

Exemplificando o modelo de classificação, ele pode ser usado desde a diferenciação de maçãs para laranjas, até para a categorização automática da resposta de um cliente em diversas categorias como, críticas construtivas, elogios, críticas destrutivas entre outros.

Já o modelo regressivo é um método supervisionado que usa um algoritmo, que entende a relação entre variáveis dependentes e independentes. Enquanto o foco da outra técnica é fazer a classificação de dados. O foco do modelo regressivo é fazer previsões de valores numéricos com base em diferentes pontos de dados. Exemplificando o modelo regressivo, ele seria usado para fazer projeções de receitas de vendas para um negócio.

O modelo de construção de *machine learning* escolhido neste artigo, é o modelo supervisionado de classificação, onde foi desenvolvido um site que utiliza *JavaScript*, para que seja possível anexar uma imagem, e o software classificá-la com base em seu treinamento. O objetivo é avaliar como o algoritmo desempenha seu papel e encontrar formas de aplicar a ferramenta de reconhecimento de imagem no dia a dia, por intermédio de uma interface web.

2 REVISÃO BIBLIOGRÁFICA

Tendo exposto o objetivo da pesquisa, é possível perceber diversas utilidades para essa tecnologia, principalmente na descoberta de novos padrões e na classificação automatizada. Como exemplo disso temos a análise de casos de covid-19, com o uso de *machine learning* supervisionada, como descrito no artigo “Corona virus disease (COVID-19) análise de casos usando aplicativos de aprendizado de máquina”, publicado por Kwekha-Rashid et al. (2021).

Ainda na área da saúde temos a classificação de imagens de mamografia, para reconhecimento de câncer de mama, descrito no artigo “Classificação de imagens de mamografia com *Machine Learning* no auxílio de diagnósticos de câncer

de mama” por Wurzel e Martins (2022). Na área dos esportes eletrônicos, temos o treinamento de inteligências artificiais, feitas para descobrir quão rápido um jogo pode ser concluído, assim como descrito por Junior, et al. (2019)

2.1 Machine Learning

De acordo com Chagas (2019), a evolução do *Machine Learning* se deve ao considerável aumento da capacidade de processamento dos computadores atuais, que possibilitam o uso de grandes bancos de dados em seus treinamentos e processamento de sinais, melhorando a acurácia e desempenho das predições. De acordo com Xin e Wang(2019), pelo avanço da conexão de internet móvel, mais e mais imagens são armazenadas na Internet, o que alavanca a capacidade de treinamento do algoritmo avanços estes que trazem um leque maior de oportunidades para o uso do *Machine Learning*.

2.2 Redes Neurais Artificiais

Algo interessante de se notar é que de acordo com Shahida, Rappon e Berta (2019), no reconhecimento de padrões complexos é usado um tipo de processamento que simula um cérebro humano, é dito que os conceitos dentro do campo de estudo relacionado a essas redes neurais artificiais, são análogas à neurociência. RNA, como mencionado, é um ramo da IA que utiliza de técnicas de aprendizado de máquina e que são utilizados em tarefas de classificação.

Essa tecnologia é subdividida em três camadas de neurônios, uma onde acontece o recebimento de dados (input), saída de dados (output), e por último a camada onde acontece o reconhecimento de padrões (hidden), nela também acontece a maior parte do processamento.

2.3 Imagem Digital e Classificação de imagem

Imagem digital é uma matriz onde índices de linhas e colunas indicam um ponto na imagem e este ponto é chamado de *pixel*, então essa imagem é definida como bidimensional, e codificada em números binários (ALVES et al, 2006).

Muitos são os dispositivos ou sensores que são utilizados para capturar imagens, os mais comuns são câmeras fotográficas, scanners, estes dispositivos capturam uma imagem analógica e a transformam em uma representação digital. Imagens já digitalizadas, são submetidas a uma extração de características, por softwares com a capacidade de reconhecer a imagem utilizando algoritmos de RNA.

A Rede Neural Convolucional (RNC), é apropriada para extrair atributos de imagens, através de pesos dos filtros, que são aprendidos durante o processo de treinamento. Os métodos de classificação podem ser divididos em classificação por

pixel ou por regiões, e que por sua vez podem ser classificados como supervisionados e não supervisionados.

2.4 JavaScript

O JavaScript, é uma linguagem que foi aberta a publicação em 1995, assim como descrito por Wohlgethan(2018), é uma criação que objetivou facilitar a produção de códigos e sites, sendo mais dinâmica, poderosa e tendo uma sintaxe mais simples.

Hoje essa linguagem é um dos principais pilares da programação, sendo amplamente utilizada no mundo, com uma vasta gama de ferramentas que podem ser utilizadas nesta linguagem, permitindo a programação de inúmeros projetos, inclusive a programação de *machine learning*.

2.2 ML5, Mobile-Net E Tensor Flow

De acordo com Pujara (2020), o modelo mobile-Net se trata de uma arquitetura de rede neural artificial, que diferente de modelos tradicionais, têm seu processamento dividido em processos menores, fazendo dela assim uma rede neural artificial com menor peso para o processamento, e menor complexidade.

Como aponta Lamba (2018), a biblioteca ML5 foi criada na New York University, lançada publicamente em Julho de 2018, com o objetivo de simplificar a abordagem de *machine learning* para entusiastas e iniciantes no assunto, esta biblioteca traz acesso à funções, algoritmos e funcionalidades da Rede Neural Artificial Tensorflow.js.

Criada pelo Google Brain Team em 2015, Tensor Flow é também uma biblioteca de acesso livre, que traz consigo um conjunto de modelos e algoritmos de *Machine Learning* e RNA, sendo útil para fazer o processo de predição de imagens, através do processo das RNAs, sendo possível fazer o treinamento de um algoritmo através dela.

Além disso, traz consigo modelos pré-treinados, que podem ser usados em projetos como o Mobile-Net, modelo este que será utilizado na metodologia do presente trabalho.

3 METODOLOGIA

A metodologia empregada na pesquisa de classificação de imagem é de cunho experimental e de natureza aplicada. Como exposto, o problema que norteia esta pesquisa é avaliar a performance do algoritmo de classificação e reconhecimento de imagens e discutir sobre possibilidades de uso da tecnologia.

Para iniciar a fase de avaliação, foi necessário o desenvolvimento do site utilizando recursos de HTML (Linguagem de Marcação de Hipertexto), e JavaScript que utiliza o ML5 software de aprendizado de máquina, que utiliza o *TensorFlow*, nas suas camadas internas. Essa ferramenta permite que sejam executados testes pelo navegador de internet utilizando a linguagem JavaScript.

Um pre-trained model (modelo pré-treinado) se trata de um modelo previamente treinado, com um grande banco de dados de imagens, procurando por semelhanças entre a imagem, após isso retorna com uma lista de predições sobre a imagem e o nível de confiança sobre a predição.

A aplicação é feita na linguagem JavaScript, onde por meio da biblioteca ml5.js, enviamos uma imagem ao Canvas, com isto, a biblioteca ml5 faz a chamada da função de predição, onde a “Mobile-Net” faz a predição sobre o que ela acha que a imagem se trata, trazendo em forma de texto, o nome da predição feita, como o nível de confiança que ela tem sobre essa predição, indo de 0 a 100%, variando de acordo ao que ela sabe sobre a imagem inserida. Na Figura 1 mostra como deve ser adicionado a biblioteca ao código HTML

Figura 1: Adicionando a biblioteca ao código HTML

```

6 <meta name="viewport" content="width=device-width, initial-scale=1.0">
7 <title>Document</title>
8 <style>
9   #container {
10     border: 5px solid black;
11     padding: 30px;
12     font-size: 30pt;
13   }
14   body {
15     padding: 0;
16     margin: 0;
17     margin-left: 100px;
18   }
19 </style>
20 <script src="./js/p5.js"></script>
21 <script src="./js/sketch.js"></script>
22 <script src="https://unpkg.com/ml5@latest/dist/ml5.min.js"></script>
23 </head>
24 <body>
25   <p id="container">Arraste uma imagem ao canvas</p><br>
26   <p id="prediction"></p>
27 </body>
28 </html>

```

Fonte: O AUTOR, 2022.

Na Figura 2 na função setup, deve ser adicionado a chamada do algoritmo Mobile-Net. Esta chamada de função busca os dados da biblioteca ml5.js, responsável por fazer a predição da imagem.

Figura 2 : Chamando a função ‘Mobile Net’ para fazer a predição.

```
5 ~ function setup() {
7   canvas = createCanvas(400,400);
8   background(128);
9
10  classifier = ml5.imageClassifier('MobileNet', modelReady);
11
12  canvas.drop(loadImgCanvas);
13  pred = select("#prediction");
14
15  container = select("#container");
16
17  container.dragOver(activate);
18  container.dragLeave(deactivate);
19 }
```

Fonte: O AUTOR, 2022.

Já na Figura 3, o algoritmo de classificação realizou as comparações e cálculos necessários, e retornou uma matriz de dados contendo as classes encontradas, mostrando em formato de texto o que ela considera estar contido na imagem, juntamente com o nível de certeza que ela tem sobre a predição.

Figura 3: Resultado do primeiro teste

macaw|| confidence = 0.51%



[sketch.js:681](#)

```
▼ Array(3) ⓘ
  ▶ 0: {label: 'macaw', confidence: 0.5147210359573364}
  ▶ 1: {label: 'toucan', confidence: 0.07742135971784592}
  ▶ 2: {label: 'kite', confidence: 0.015225648880004883}
  length: 3
  ▶ [[Prototype]]: Array(0)
```

Fonte: O AUTOR, 2022.

Pode-se observar que a confiança está entre 0 e 1, sendo que o valor mais próximo de 1 significa maior confiança no caso da Figura 3, a classificação teve 51% de confiança no retorno como Arara (macaw).

Figura 4: Resultado do segundo teste

tiger, Panthera tigris|| confidence = 0.92%



sketch.js:681
▼ (3) [{...}, {...}, {...}] ⓘ
▶ 0: {label: 'tiger, Panthera tigris', confidence: 0.91608822:
▶ 1: {label: 'jaguar, panther, Panthera onca, Felis onca', cor
▶ 2: {label: 'tiger cat', confidence: 0.01476131659001112}
length: 3
▶ [[Prototype]]: Array(0)

Fonte: O AUTOR, 2022.

O segundo teste realizado na Figura 4 apresentou uma confiança de 92%, que na imagem há um tigre, retornando com “tiger, Panthera Tigris”, isso é possível pois o algoritmo conta uma extensa base de dados com milhões de imagens para realizar o seu treinamento.

É importante salientar que o ML5 atende a demandas de reconhecimento de imagens, desde que na sua base de treinamento o algoritmo tenha aprendido a classificar as imagens, sendo assim imagens que não fazem parte do treinamento anterior não sejam classificadas de forma correta, para sanar a falta de treinamento da rede em aplicações específicas o ML5 viabiliza o treinamento de novas imagens agregando o método conhecido como Feature Extractor, com ele é possível que a rede neural seja treinada para classificação de qualquer imagem.

4 RESULTADOS

A possibilidade de utilizar o reconhecimento e classificação de imagens na ciência permite maior eficiência nos processos produtivos, ou em diagnósticos de doenças, fazendo com que tenhamos uma gama imensa de possibilidades. Os resultados dos testes se apresentaram satisfatórios, pois todas as imagens submetidas ao teste retornaram com sucesso o seu conteúdo, fato esse que gera perspectiva positiva, de que o algoritmo possa ser utilizado para os mais diversos fins. A facilidade de implementação da biblioteca e tecnologias já existentes, permite que seja desenvolvido ferramentas de classificação de imagens com acesso a uma grande base de dados já treinada. E por fim, a possibilidade de treinar a rede neural para atingir qualquer objetivo, e nos mais diversos contextos, como agricultura, indústria, saúde e comércio.

5 CONSIDERAÇÕES FINAIS

Levando em consideração o modelo apresentado neste artigo, compreendemos que a área da inteligência artificial e da *Machine Learning* tem

grande potencial justamente por trazer a possibilidade de treiná-la para qualquer objetivo, atendendo as necessidades e agilizando processos repetitivos.

REFERÊNCIAS

ALVES, Wonder Alexandre Luz; DE ARAÚJO, Sidnei Alves; LIBRANTZ, **André Felipe Henriquez. Reconhecimento de padrões de texturas em imagens digitais usando uma rede neural artificial híbrida.** *Exacta*, v. 4, n. 2, p. 325-332, 2006.

CHAGAS, Edgar Thiago De Oliveira. **Deep Learning e suas aplicações na atualidade.** *Revista Científica Multidisciplinar Núcleo do Conhecimento*. (2019). Disponível em: <https://www.nucleodoconhecimento.com.br/administracao/deep-learning>

CLEMENTE, Bianca Cristina et al. Caça objetos: reconhecimento de imagem como um jogo. 2021.

JUNIOR, Emanuel Sales dos Santos et al. **Reinforcement Learning for Pac-Man Speedrun Training.** *Brazilian Journal of Development*, 21 nov. 2019. Disponível em: <https://brazilianjournals.com/ojs/index.php/BRJD/article/view/4755/4384>. Acesso em: 20 out. 2022.

KWEKHA-RASHID, A.S., Abduljabbar, H.N. & Alhayani, B. **Coronavirus disease (COVID-19) cases analysis using machine-learning applications.** *Appl Nanosci* (2021). Disponível em: <https://doi.org/10.1007/s13204-021-01868-7> . Acesso em: 20 out. 2022.

OSISANWO, F.Y. et al. **Supervised Machine Learning Algorithms: Classification and Comparison.** *International Journal of Computer Trends and Technology (IJCTT)*, v. 48, jun. 2017. Disponível em: https://www.researchgate.net/profile/J-E-T-Akinsola/publication/318338750_Supervised_Machine_Learning_Algorithms_Classification_and_Comparison/links/596481dd0f7e9b819497e265/Supervised-Machine-Learning-Algorithms-Classification-and-Comparison.pdf. Acesso em: 19 out. 2022.

PUJARA, Abhijeet. **Image Classification With MobileNet**(2020). Disponível em: <https://medium.com/analytics-vidhya/image-classification-with-mobilenet-cc6fbb2cd470>. Acesso em 23 out. 2022.

SARKER, I.H. **Machine Learning: Algorithms, Real-World Applications and Research Directions.** SN COMPUT. SCI. 2, 160 (2021). Disponível em: <https://doi.org/10.1007/s42979-021-00592-x> acesso em: 19 out. 2022.

SHAHIDA, Nida; RAPPON, Tim; BERTA, Whitney. Applications of artificial neural networks in health care organizational decision-making. Applications of artificial neural networks in health care organizational decision-making: A scoping review, Plos One, 19 fev 2019. acesso em: 22 de out. 2022, disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0212356>

SOOFI, Aized Amin; AWAN, Arshad. **Classification Techniques in Machine Learning: Applications and Issues.** Journal of Basic & Applied Sciences, 29 ago. 2017. Disponível em: <https://setpublisher.com/pms/index.php/jbas/article/view/1715/1579>. Acesso em: 19 out. 2022.

WOHLGETHAN, Eric. Supporting Web Development Decisions by Comparing Three Major JavaScript Frameworks: Angular, React and Vue.js, Hamburg university of Applied sciences, 15 maio 2018. Acesso em 23 de out. 2022, disponível em: https://reposit.haw-hamburg.de/bitstream/20.500.12738/8417/1/BA_Wohlgethan_2176410.pdf

WURZEL, Pedro; MARTINS, Mirkos Ortiz. **Classificação de imagens de mamografia com Machine Learning no auxílio de diagnósticos de câncer de mama.** Disciplinarum Scientia| Naturais e Tecnológicas, v. 23, n. 2, p. 1-17, 2022.

XIN, Mingyuan X; WANG, Yong W. **Research on image classification model based on deep convolution neural network**(2019). Disponível em: <https://jivp-urasipjournals.springeropen.com/articles/10.1186/s13640-019-0417-8> Acesso em 23 out. 2022.

YEGULAP, Serdar.**What is TensorFlow? The machine learning library explained**(2022). Disponível em: <https://www.infoworld.com/article/3278008/what-is-tensorflow-the-machine-learning-library-explained.html>. Acesso em 23 out. 2022.