

ANÁLISE DE SENTIMENTOS EM IMAGENS ATRAVÉS DE REDES NEURAIAS CONVOLUCIONAIS

Vinicius Pereira de Oliveira¹; Rogério Bernardes Andrade²; Eduardo Fernandes Saad³

^{1,2,3} Faculdade de Talentos Humanos - FACTHUS, Uberaba (MG), Brasil

viniciuspdoliveira@gmail.com, rbandrade@facthus.edu.br, eduardo.saad@facthus.edu.br

RESUMO: O aprendizado de máquina e a inteligência artificial cada vez mais presentes em ambientes interativos, trazem por sua vez os conceitos de visão computacional e *IoT (Internet of Things)* conectados ao grande fluxo de dados que circulam nas redes e a forma como são tratados, analisados e classificados por dispositivos apropriados. O objetivo do presente trabalho é fazer um comparativo das métricas de precisão e grau de confiança entre duas abordagens diferentes de ferramentas, que apresentam o mesmo conceito de inicial de funcionalidades para a classificação de imagens. Com o conjunto de dados utilizado e o método proposto, obteve-se um melhor resultado através do *Teachable Machine*, entretanto, por meio do uso do *Cloud Annotations* nota-se significativas possibilidades de melhorias por habilitar e fornecer maiores recursos de implementação por meio de codificação.

PALAVRAS CHAVE: Aprendizagem de máquina, Inteligência artificial, Transferência de aprendizado, Visão computacional.

ANALYSIS OF SENTIMENTS IN IMAGES THROUGH CONVOLUTIONAL NEURAL NETWORKS

ABSTRACT: *Machine learning and artificial intelligence increasingly present in interactive environments bring the concepts of computer vision and IoT (Internet of Things) connected to the large flow of data that circulates on networks and the way they are treated, released and classified by appropriated devices. The objective of the present work is to compare the precision and confidence metrics between two different tool approaches, which present the same initial concept of characteristics for an image classification. With the data set used and the proposed method, a better result was obtained through the Teachable Machine, however, with the use of Cloud Annotations, it is possible to notice possible improvements to enable and provide greater implementation resources through coding.*

KEYWORDS: *Machine learning, Artificial intelligence, Transfer learning, Computer vision.*

INTRODUÇÃO

O aumento do uso das redes sociais de forma que vem se tornando cada vez mais popular e acessível o consumo de informações através da internet gera uma quantidade significativa de dados, que são muito valiosos para as empresas que almejam compreender melhor o comportamento dos seus usuários. Desta forma, profissionais que trabalham com a Ciência de Dados desenvolveram ferramentas para gerar informações e conclusões baseados nesses dados. Uma dessas ferramentas é a Análise de Sentimentos (AS) que neste estudo será usada com base na Inteligência Artificial por meio de Redes Neurais Convolucionais.

Becker (2017) propôs o uso de três arquiteturas de Redes Neurais Convolucionais que lidam com dados multilínguas extraídos do *Twitter* contendo quatro línguas. A Análise de Sentimento visa obter conhecimento sobre a polaridade das mensagens postadas, através de diversas técnicas, apoiando-se principalmente nas Redes Neurais Profundas (*Deep Learning*). Os dois primeiros modelos propostos caracterizam-se pelo fato de possuírem um total de parâmetros muito menor que os demais *baselines* considerados, e ainda assim, obtêm resultados superiores

com uma boa margem de diferença. O último modelo proposto é capaz de realizar uma classificação multitarefa, identificando a polaridade das sentenças e também a língua. Com este último modelo obtém-se uma acurácia de 74,43% para AS e 98,40% para Identificação da Língua em um *dataset* com quatro línguas, mostrando-se a melhor escolha entre todos os *baselines* analisados.

Grando (2020) apresenta os conceitos de sala de aula inteligente, visão computacional e redes neurais, com a proposta da construção de uma aplicação capaz de analisar as imagens de estudantes em sala de aula a fim de identificar padrões de emoções. Para o desenvolvimento desse *software* foram estudados os processos de aquisição e processamento de imagens, definição e implementação de uma rede neural, culminando na realização de testes. Testes realizados confirmaram a viabilidade do uso de redes neurais convolucionais para o reconhecimento de emoções nas expressões faciais. Desta forma, torna-se possível integrar o *software* proposto a um sistema *IoT (internet of things)* que constituirá componente relevante a uma sala de aula, adicionando assim um nível preliminar de inteligência a ela.

A pesquisa realizada por Matos (2017), tem como objetivo estabelecer uma metodologia para a criação de uma

rede neural convolucional onde exista um número reduzido de dados para treinamento, usando um cenário de identificação de expressões faciais em imagens e suas respectivas emoções. O autor pretendeu com este trabalho, verificar a possibilidade de leituras em tempo real, tendo como objetivo fornecer dados de situações nas quais os *feedbacks* podem ser fundamentais para a tomada de decisões. Para a criação do *software*, foram utilizados diferentes técnicas de visão computacional e técnicas de aprendizagem de máquina com redes neurais convolucionais, e também utilizados estudos de comportamento humano. Mesmo com um conjunto de dados de treinamento reduzidos, como resultado foi possível uma acurácia de teste final média de 65% com tempo médio de treinamento de dez minutos.

Baseado nestes trabalhos, o presente estudo tem por objetivo fazer um comparativo entre as ferramentas *Teachable Machine* e *Annotations*, com *Python* e *TensorFlow*, utilizando como base as métricas de precisão e grau de confiança, comum às duas plataformas, para verificar em mesmas condições de experimento aquela que apresenta melhores resultados.

FUNDAMENTAÇÃO TEÓRICA

Análise de Sentimento. A mineração de opiniões (MO), comumente conhecida como análise de sentimento (AS), é o campo de estudo que analisa as opiniões de pessoas, sentimentos, avaliações, atitudes e emoções em relação a entidades como serviços, organizações, eventos, produtos e indivíduos. O termo análise de sentimento surgiu pela primeira por Pang, Lee e Vaithyanathan (2002). Consiste de uma área multidisciplinar que abrange o processamento de linguagem natural (NLP), linguística e inteligência artificial para identificar, extrair e quantificar informações subjetivas de forma sistemática. A AS é amplamente aplicada em marketing, relações públicas, investimento, campanhas políticas e em muitas outras áreas (SANTOS, 2017).

As reações humanas a determinadas situações variam de acordo com o contexto social em que se depara mediante ao momento e cenário estabelecidos. Entretanto, algumas expressões faciais puderam ser catalogadas como sendo universais (EKMAN, 2003), apesar disso, existe uma vasta variância quando estas expressões são demonstradas, abrangendo desde reações simples e totalmente compreensíveis por quem se observa, até movimentos sutis de músculos da face não tão evidentes e notáveis. Estas expressões minúsculas, quase imperceptíveis, nem sempre são facilmente captadas por humanos, uma vez que podem durar apenas uma fração de segundos. Embora este tipo de limitação não ocorra em análises feitas por computadores, a questão passa a ser como um *software* pode fazer medições e interpretações com relação a comportamentos humanos (MATOS, 2017).

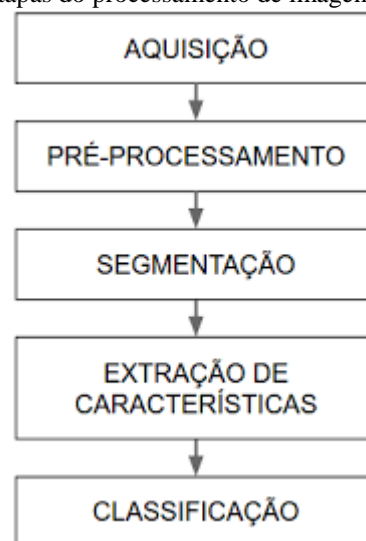
Visão Computacional. No dia a dia a visão humana é utilizada para observar o ambiente, as ações, os objetos ao nosso redor e, obter informações importantes

sobre eles, porém apesar de ser uma habilidade básica para os seres humanos a mesma tarefa torna-se um tanto quanto complexa quando realizada por uma máquina (RAUTARAY e AGRAWAL, 2012).

A visão computacional procura auxiliar na resolução de problemas abstratos, simulando a visão humana, capturando, processando e analisando imagens tentando imitar a habilidade do ser humano de tomar decisões de acordo com as informações obtidas. A forma com que o computador consegue interpretar o ambiente que está ao seu redor, extraindo informações através de câmeras de vídeos e sensores é responsabilidade da visão computacional, o que permitirá ao computador reconhecer e manipular os objetos que compõem as imagens capturadas (LI e SHI, 2018). Conforme as tecnologias vão evoluindo, novas áreas acabam se beneficiando da integração de processamento de imagens com a inteligência artificial (GRANDO, 2020).

ETAPAS DA VISÃO COMPUTACIONAL E O PROCESSAMENTO DIGITAL DE IMAGENS

Figura 1: Etapas do processamento de imagens.



Fonte: Os autores (2021).

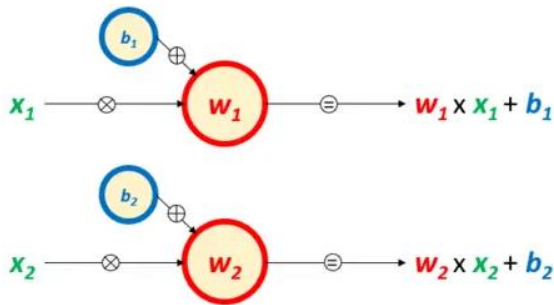
Redes Neurais. Uma das principais vantagens de se utilizar redes neurais é a sua capacidade de aprender através de seu próprio uso e melhorar o seu desempenho através dessa aprendizagem. Aprender significa adaptar-se ao ambiente, modificando o seu comportamento ao longo do tempo e de acordo com as regras, melhorando sua capacidade na resolução de problemas (AMBRÓSIO, 2002).

A partir deste modelo de rede neural surgiram dois novos paradigmas na área de inteligência artificial: simbólica e conexionista. A inteligência artificial simbólica busca simular o comportamento de inteligência humana sem uma inspiração biológica para realizá-la. A inteligência artificial conexionista busca simular a inteligência através de um sistema que possua uma estrutura parecida com a do cérebro humano, onde este sistema seria capaz de apresentar

inteligência, por meio da aprendizagem e do erro (FURTADO, 2019).

Funções de Ativação. As redes neurais são estruturas que promovem transformações matemáticas (Fig. 2) nos dados que recebem para processar. Em cada neurônio de cada camada, elas multiplicam o valor de entrada pelo peso do neurônio correspondente, somam com o *bias* associado ao neurônio, e passam esse valor adiante.

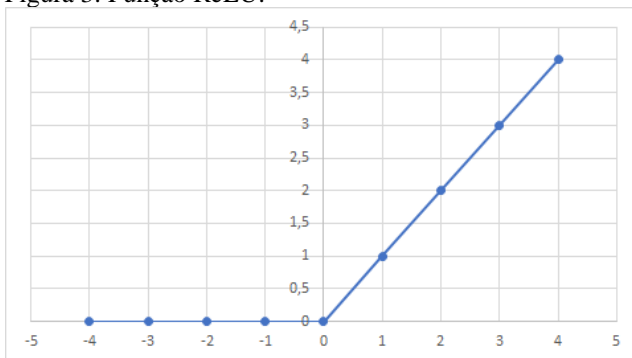
Figura 2: Operações matemáticas com dois neurônios.



Fonte: (IA EXPERT ACADEMY, 2020).

Função ReLU. A função ReLU (do inglês *rectified linear unit*) é uma função de ativação não linear amplamente utilizada para extração de características (Fig. 3). A função ReLU pode ser utilizada tanto para pré-treinamento como para classificação e possui uma alta velocidade de convergência (HARA; SAITO; SHOUNO, 2015).

Figura 3: Função ReLU.



Fonte: (GRANDO, 2020).

Arquitetura das Redes Neurais. Uma das primeiras e mais simples formas de redes neurais artificiais foram os *perceptrons*, introduzido por Rosenblatt no ano de 1958. Os *perceptrons* de uma camada são úteis para a solução de problemas linearmente separáveis, podendo classificar padrões somente entre duas classes (AMBRÓSIO, 2002). Quando um problema não admite uma separação linear exata, torna-se necessário utilizar os *perceptrons* multicamadas (MLP, do inglês *Multi Layer Perceptrons*), onde os neurônios estão agrupados em várias camadas que podem ser basicamente divididos em três

partes: camada de entrada, onde são originados os dados; camadas intermediárias, também chamadas de camadas ocultas, onde são realizadas as extrações de características e inferência; camada de saída, responsável pela conclusão e apresentação do processamento. Ainda segundo Furtado (2019), normalmente as arquiteturas fundamentais podem ser subdivididas em três diferentes classes: Rede *feedforward* de uma camada; Rede *feedforward* de múltiplas camadas; Rede recorrente ou realimentada.

Treinamento e Aprendizagem. Temos como uma das principais vantagens de se utilizar redes neurais é a sua capacidade de aprender através de seu próprio uso e aprimorar o seu desempenho através dessa aprendizagem. Aprender significa adaptar-se ao ambiente, modificando o seu comportamento ao longo do tempo e de acordo com as regras, melhorando sua capacidade na resolução de problemas. O processo pelo qual a rede neural artificial precisa ser submetida para que atinja seu destino de resultados denomina-se etapa de treinamento, onde seus pesos e vieses devem passar pelo processo de refinamento, de maneira que uma entrada futura seja corretamente classificada. Somente são permitidas alterações dos parâmetros de uma rede durante a etapa de treinamento, devendo permanecer estáticos durante as fases de teste e de execução (AMBRÓSIO, 2002).

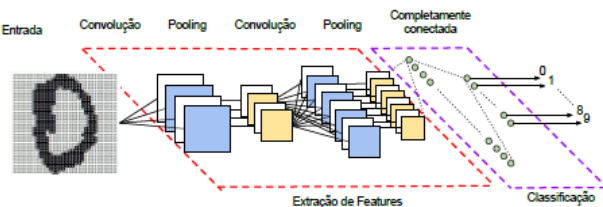
Em uma rede neural o conceito de aprendizado distingue-se de treinamento, pois o aprendizado está relacionado a uma tarefa que a rede está executando em função do treinamento, da topologia e da arquitetura da rede, já o treinamento é o processo de ensinar a rede neural (FURTADO, 2019). Para realizar o treinamento são apresentadas amostras de dados na entrada da rede e a saída é redirecionada para uma resposta esperada, forçando assim que seus pesos e parâmetros sejam modificados até atingirem uma validação desejada. Ao finalizar a etapa de treinamento, a rede terá então o conhecimento necessário sobre o ambiente em que está atuando (FLECK et al., 2016). Etapas de treinamento da rede: Treinamento Supervisionado; Treinamento Não-Supervisionado; Treinamento por Reforço.

Redes Neurais Convolucionais. Uma Rede Neural Convolutiva (ConvNet / *Convolutional Neural Network* / CNN) é um algoritmo de Aprendizado Profundo que pode captar uma imagem de entrada, atribuir importância (pesos e vieses que podem ser aprendidos) a vários aspectos / objetos da imagem e ser capaz de diferenciar um do outro (Fig. 4). O pré-processamento exigido em uma ConvNet é muito menor em comparação com outros algoritmos de classificação. Enquanto nos métodos primitivos os filtros são feitos à mão, com treinamento suficiente, as ConvNets têm a capacidade de aprender esses filtros / características (DEEP LEARNING BOOK, 2019).

Existem atualmente modelos de redes neurais convolucionais (CNN) pré-treinadas, para uso variado, que aceleram o desenvolvimento de produtos de software nesta área. Diversos desenvolvedores (*Google, Microsoft, IBM*, entre outros) disponibilizam suas implementações de redes

neurais para que sejam adaptadas e reutilizadas em outros cenários, na forma de pacotes de software. Portanto, nesta etapa pode ser utilizado o processo de *Transfer Learning* (do inglês, transferência de aprendizado), que possibilita que uma CNN já treinada seja reutilizada com um novo conjunto de dados. Desta forma ela aprende incrementalmente novas classificações (PAN; YANG, 2010).

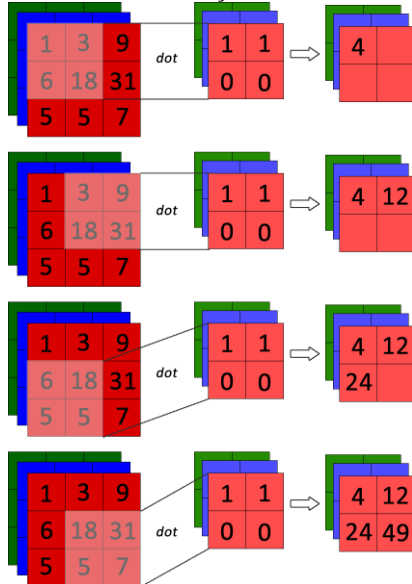
Figura 4: Rede Neural Convolucional.



Fonte: (VARGAS; PAES; VASCONCELOS, 2016).

Camada de Convolução. A camada de convolução é composta por diversos neurônios, onde cada um é responsável por aplicar um filtro em um pedaço específico da imagem (Fig. 5). Os pesos atribuídos às conexões de um neurônio podem ser representados como uma matriz que identifica os filtros de convolução (também conhecidos como *kernel* ou máscara), possuindo valores reais e sendo capazes de aprender com o processo de treinamento (VARGAS; PAES; VASCONCELOS, 2016).

Figura 5: Camada de Convolução.

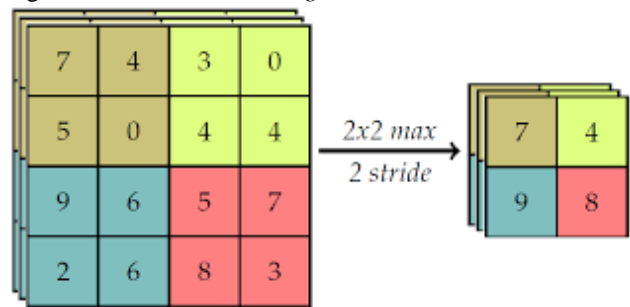


Fonte: (PACHECO, 2019).

Camada de Pooling. A camada de *pooling* serve para simplificar os dados da camada anterior, reduzindo a dimensionalidade das informações. A redução das informações é necessária para que o treinamento possa ser realizado com maior agilidade, reduzindo os sobre ajustes e a invariância de transições (LEMLEY; BAZRAFKAN; CORCORAN, 2017). A camada de *pooling* funciona

agrupando um conjunto de dados, utilizando-se dos conceitos de filtros e *strides* (quantidade de passos que podem ser dados de uma única vez), como, por exemplo em uma matriz de 4 x 4 sendo aplicado um filtro com configuração de 2 x 2 que irá selecionar um valor para representar cada área dessa matriz através de uma função para produzir o resultado final (GRANDO, 2020).

Figura 6: Camada de *Pooling*.



Fonte: (PACHECO, 2019).

Softmax. Softmax é uma generalização da função sigmóide para casos não-binários. Ela não costuma ser aplicada às camadas escondidas da rede neural, mas sim na camada de saída de problemas de classificação multiclasse, já que sua característica é produzir valores no intervalo [0, 1] onde sua soma é igual a 1. Ou seja, num problema com 4 classes, por exemplo, a função softmax vai produzir 4 valores, que somam 1, onde cada valor representa a probabilidade da instância pertencer a uma das 4 classes possíveis (IA EXPERT ACADEMY, 2020).

Imagenet. Competição *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) em 2012. A competição *ImageNet* consiste em avaliar diversos algoritmos para detecção de objetos e classificação de imagens em larga escala. A motivação desta competição é permitir que os pesquisadores façam a comparação e o progresso na detecção em uma ampla variedade de objetos e medir o progresso da área de visão computacional (SILVA, 2018).

TensorFlow. Para realizar implementações baseadas em técnicas de visão computacional e redes neurais convolucionais é necessário utilizar ferramentas que facilitam e auxiliam no desenvolvimento de tais tarefas. O *TensorFlow* é uma biblioteca de código aberto de *machine learning* para criação e treinamento de sistemas de redes neurais e detecção de padrões, tendo sido desenvolvido originalmente pela *Google Brain Team* e lançado em 2015, oferece APIs para pesquisa e desenvolvimento desktop, mobile, web e cloud (TENSORFLOW, 2019).

A principal característica do *TensorFlow* é a possibilidade de realizar operações em matrizes com grande volume de dados de forma eficiente.

As imagens têm dimensões 48 x 48 pixels. As características sob análise são os valores de cada pixel. A rede neural cria um modelo interno das relações entre pixels. Para tanto, os dados de entrada são enviados para a primeira

camada escondida (*hidden layer*) com seus respectivos pesos. Nesta camada os dados passam pela função ativação. Com base no resultado desta função de ativação, o neurônio decide se deve ou não enviar o dado para uma próxima camada escondida ou para a saída final da rede. Uma vez que o resultado tenha sido obtido, comparam-se os valores com os valores esperados. Para determinar quão diferente os resultados se apresentam, é aplicada uma função de custo (*loss function*).

Entende-se por *backpropagation* ou retro propagação os ajustes dos pesos e denomina-se *feedforward* o envio dos dados por meio da rede. O ciclo entre o processamento dos dados pela rede e a retro propagação para ajustes dos pesos recebe o nome de época (*epoch*). É necessário também encontrar um valor para que a quantidade de épocas processadas não leve a uma situação de *overfitting*, cenário em que dados irrelevantes de uma imagem podem ser utilizados para realizar a classificação (MATOS, 2017).

Com a transferência de aprendizado (*Transfer Learning*) e realizando um novo treinamento sobre os pesos desta rede, é possível adequá-la para determinação de novas categorias, portanto, realizar um treinamento apenas sobre a última camada de uma rede, obtendo-se resultados satisfatórios, com tempo de treinamento baixo e sem a necessidade de hardwares específicos como GPUs (*Graphics Processing Units*, unidades de processamento gráfico). Uma vez que o tempo de treinamento é relativamente curto, é possível realizar uma maior quantidade de ajustes durante as etapas de treinamento. A maior vantagem da estratégia de transferência de aprendizado é que várias das camadas já treinadas para distinguir objetos específicos, podem ser reutilizadas na classificação de novas categorias sem que sejam necessárias alterações em sua estrutura base.

MÉTRICAS

Precisão. É obtida dividindo-se a quantidade de verdadeiro positivos pelo total de amostras classificadas como positivas (verdadeiro positivos + falso positivos).

Verdadeiro Positivo (VP): segmentos que pertencem a classe positiva e foram classificados como positivos.

Verdadeiro Negativo (VN): segmentos que pertencem a classe negativa e foram corretamente classificados como negativos.

Falso Positivo (FP): segmentos que pertencem a classe negativa e foram classificados como positivos.

Falso Negativo (FN): segmentos que pertencem a classe positiva e foram classificados como negativos.

Grau de Confiança. O grau de confiança refere-se à taxa de sucesso dentro de um número de amostras, ou seja, com que frequência esse tipo de intervalo vai capturar o parâmetro de interesse. Um intervalo de confiança específico fornece uma amplitude de valores plausíveis para o parâmetro de interesse.

MATERIAIS E MÉTODOS

Para a realização deste trabalho foi utilizado um *dataset* público disponibilizado pela plataforma *Kaggle*. O mesmo possui imagens no padrão 48 x 48 em escala de cinza. A base de dados disponível é desbalanceada, sendo assim, foi feita a curadoria das imagens para serem acrescentadas na implementação, obtendo-se, desta forma, um novo *dataset* de 12.000 imagens divididas nas seguintes classificações: Feliz (3.000 imagens), Triste (3.000 imagens), Surpreso (3.000 imagens) e Neutro (3.000 imagens).

Para a etapa de testes temos um *dataset* com 400 imagens, dividido em 100 imagens para cada uma das classes de cada sentimento.

Sentimentos Em Expressões Faciais. As emoções quando tratadas no âmbito do estímulo causado por determinadas interações, na maioria das vezes resulta em reações que são manifestações voluntárias e involuntárias acionadas no nosso comportamento, ocasionando por sua vez estados de alterações da expressão facial nos seres humanos. O cérebro de cada pessoa que é exposta a determinada situação reage de uma forma diferente, sendo assim, expressamos em nossas faces por um determinado período de tempo sentimentos e emoções geradas por influências internas ou externas que podem ser perceptíveis para aqueles que estão ao nosso redor. Para o desenvolvimento deste trabalho foram escolhidos quatro sentimentos: alegria, tristeza, surpresa e neutralidade; demonstrados respectivamente em cada uma das linhas da Figura 7.

Figura 7: Amostras do *dataset* utilizado pelo autor.



Fonte: (KAGGLE, 2020).

Para o desenvolvimento deste projeto foi utilizado um Notebook (HP Intel Core i3, 8GB de RAM, HD de 120GB), linguagem *Python* para inferências no código e o *TensorFlow* para a criação da rede neural convolucional.

Teachable Machine. O *Teachable Machine* foi configurado da seguinte forma para receber os dados e imagens para o treinamento e teste dos dados:

- Épocas (*epochs*): 50
- Tamanho do lote (*batch size*): 16
- Taxa de aprendizado (*learning rate*): 0,001

Annotations com Python e TensorFlow. Para a *Cloud Annotations* juntamente com a linguagem de programação *Python* e utilizando *Tensorflow*, temos a seguinte configuração para o treinamento e teste:

- Camadas: 4, sendo que, 2 camadas são ocultas
- Épocas (*epochs*): 50
- Tamanho do lote (*batch size*): 16
- Taxa de aprendizado (*learning rate*): 0,001

Desenvolvimento das Redes Neurais Convolucionais. Por meio do estudo realizado abrangendo as informações descritas nos tópicos anteriores, foi verificado e analisado as etapas e processos para a criação dos ambientes de implementação.

No ambiente de desenvolvimento *Teachable Machine*, plataforma *web* pertencente ao grupo *Google*, foi inserido o *dataset* de treino com as 4 classes de sentimentos trabalhadas neste projeto: feliz, triste, surpreso e neutro, contendo 3.000 imagens em cada uma destas classes. Foi realizado o treinamento do modelo. Posteriormente, foram utilizadas 100 imagens de cada sentimento (feliz, triste, surpreso e neutro) para fazer a validação e o teste do modelo criado, de modo que pudessem ser constatadas as métricas estabelecidas para este trabalho: precisão e grau de confiança.

Puderem ser estabelecidos como parâmetros de comparação entre as duas ferramentas a precisão e o grau de confiança, devido ser a única métrica em comum presente no *Teachable Machine* e na *Cloud Annotations*.

O *Teachable Machine* é uma plataforma livre e aberta para qualquer tipo de usuário, não sendo necessário ter conhecimentos técnicos, específicos ou de codificação para que se possa desenvolver modelos de aprendizado de máquina, tornando assim a ferramenta mais abrangente e acessível, obtendo-se cada vez mais usuários usufruindo deste tipo de tecnologia em atividades cotidianas ou até mesmo para fins profissionais.

Na outra arquitetura utilizada, o *Cloud Annotations* da IBM, foi imputado o mesmo *dataset* de treino para a realização desta etapa do desenvolvimento do projeto. Neste ambiente, por sua vez, também foram usadas as 12.000 imagens subdivididas nas 4 classes de sentimentos abordadas neste trabalho, realizando assim o treinamento e validação do *dataset*. Concluído o pré-processamento, foi realizada a etapa de testes, utilizando nesta fase 400 imagens, sendo: 100 imagens para a classe "feliz", 100 imagens para a classe "triste", 100 imagens para a classe "surpreso" e 100 imagens para a classe "neutro". O *Cloud Annotations* conta com a linguagem *Python*, através do ambiente de programação do *Jupyter Notebook*. O *Python* permite o uso de bibliotecas como o *OpenCV*, o *Keras*, além de outras como o *TensorFlow* que foi utilizado durante o andamento deste trabalho.

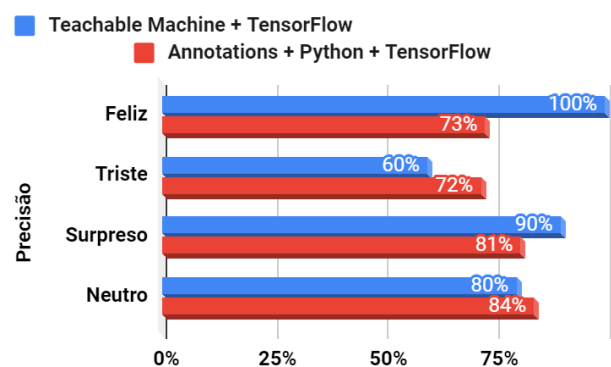
RESULTADOS E DISCUSSÃO

De acordo com análise obtida após o treinamento de 12.000 imagens do *dataset*, validação e teste para cada uma das classes de sentimentos do *dataset* no modelo de

redes neurais convolucionais gerado e implementado em cada uma das ferramentas, apresenta-se abaixo os dados colhidos.

No Gráfico 1, notamos que obtivemos um melhor desempenho de precisão no *Teachable Machine* para as classes "feliz" e "surpreso", tendo em cada uma destas classes os respectivos valores 100% e 90%. Enquanto no *Annotations* temos para estas mesmas classes os valores de 73% e 81%. Em contrapartida, para as classes "triste" e "neutro", os resultados referentes a precisão são melhores desta vez no *Annotations*, obtendo-se respectivamente 72% e 84%, para 60% e 80% no *Teachable Machine*.

Gráfico 1: Precisão.

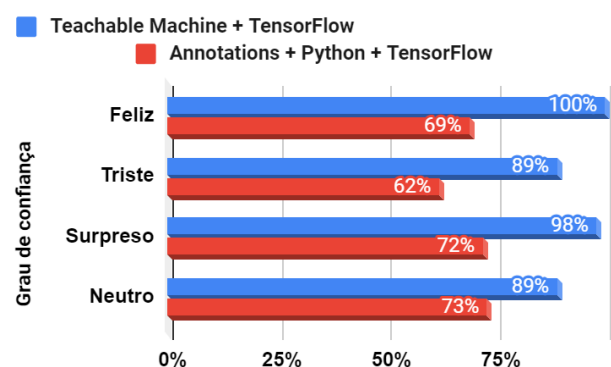


Fonte: Os autores (2021).

Durante a etapa de análise de desempenho das CNNs para avaliação das representações, mesmo que executados com sob as mesmas condições de base de dados, nota-se que o método de amostragem utilizado por cada uma das plataformas corresponde de uma maneira e as tratativas para com o *dataset* resulta de formas distintas, inclusive no tempo de processamento.

No Gráfico 2, temos os resultados obtidos para os sentimentos analisados neste trabalho para os modelos de redes neurais convolucionais em um estudo comparativo entre as ferramentas *Teachable Machine* e *Cloud Annotations* em relação ao grau de confiança, ou também podendo ser considerado como a *SoftMax*.

Gráfico 2: Grau de confiança.



Fonte: Os autores (2021).

O *Teachable Machine* resultou nesta etapa, para as classes “feliz”, “triste”, “surpreso” e “neutro”, os seguintes e respectivos valores de grau de confiança: 100%, 89%, 98% e 89%. Por outro lado, o *Annotations* obteve para a classe “feliz” 69%, “triste” 62%, “surpreso” 72% e para a classe “neutro” 73% de grau de confiança. Apesar de o *Teachable Machine* ter alcançado melhores resultados nestas comparações apresentadas, mudanças no número de épocas poderiam ocasionar *overfitting*, ou alterações para remodelagem, por exemplo, gerar *backpropagation* (retropropagação), ou ajustes de pesos e *biases* (vieses) desequilibrando outros métodos do modelo de treinamento em questão e desta maneira necessitaria controles e acertos mais específicos do projeto, o que não seria possível de se realizar através do *Teachable Machine* por ser uma ferramenta visual.

O *Annotations* por sua vez, permite uma grande flexibilidade de recursos que podem ser implementados, permitindo assim ir muito além dos resultados esperados e obtidos no *Teachable Machine*.

O *Annotations*, por ser uma ferramenta onde há espaço específico destinado à codificação através de linguagens de programação, possibilita desta forma customizar a plataforma de diversas maneiras, inserindo complementos e recursos aprimorados a fim de melhorar a implementação de um modelo de aprendizado de máquina, como por exemplo: mudar o otimizador, elaborar o algoritmo, utilizar *PyThorc* ao invés de *TensorFlow*, usar diversas camadas intercaladas de convolução e *pooling*, chamados de filtros, aperfeiçoar a modelagem de dados por meio de outros vetores e matrizes, podendo desta forma aprimorar o processamento e os resultados, além de possibilitar o uso de diversas bibliotecas existentes que aumentam os graus de precisão e otimização de modelos de *machine learning*, *deep learning*, *transfer learning*, dentre outros.

O *Teachable Machine*, uma solução “*black box*”, que apesar de ser viável dependendo e de acordo com o que será desenvolvido, muita das vezes tornará o projeto limitado, notando-se como ponto importante a escassez de métricas de análise e comparação.

CONCLUSÃO

Este trabalho teve como objetivo apresentar a análise de sentimentos em imagens em um estudo comparativo entre as ferramentas *Teachable Machine* e *Annotations*, com *Python* e *TensorFlow*, através de redes neurais convolucionais. Conclui-se que para este experimento, dentro das condições dispostas, o *Teachable Machine* foi a ferramenta que apresentou melhores resultados. Deve-se ressaltar que as métricas estabelecidas para este comparativo foram precisão e grau de confiança, por serem as que constam nas duas ferramentas utilizadas para os estudos.

Os resultados também mostram que o uso de determinadas ferramentas deve atender o maior número de recursos possíveis para uma implementação, visando

solucionar as necessidades de um projeto em sua magnitude. Por esse aspecto, nota-se a possibilidade de melhorias que podem ser realizadas através do modelo programado. Tendo em vista a era do *Big Data* (grande volume de dados), as arquiteturas que melhor se adaptarem em satisfazer desde usuários comuns das crescentes redes sociais até os profissionais que atuam com IA, Cientistas de Dados e outras diversas áreas, estarão um passo à frente das demais.

Como sugestão de trabalhos futuros seria a realização de testes utilizando outras CNNs. Estudos voltados para *fairness* buscando verificar o comportamento de um *dataset* sem viés visando o equilíbrio dos dados com atenção direcionada para a pele negra. Implementação da análise de sentimentos em tempo real juntamente com *IoT*, usando recursos para dispositivos móveis.

REFERÊNCIAS

AMBRÓSIO, P. E. **Redes neurais artificiais no apoio ao diagnóstico diferencial de lesões intersticiais pulmonares**. Dissertação (Mestrado)—Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2002. Disponível em: <<https://teses.usp.br/teses/disponiveis/59/59135/tde-26102002-155559/pt-br.php>> Acesso em: 4 de Maio de 2021.

BECKER, Willian Eduardo. **Uma abordagem de redes neurais convolucionais para análise de sentimento multi-lingual**. 2017. 85 p. Dissertação (Mestrado) - Programa de Pós-Graduação em Ciência da Computação, PUCRS. Disponível em: <<https://l1library.org/document/q2g572jy-abordagem-redes-neurais-convolucionais-analise-sentimento-multi-lingual.html>> Acesso em: 10 de Julho de 2020.

DATA SCIENCE ACADEMY. **Deep Learning Book, 2019**. Disponível em: <<https://www.deeplearningbook.com.br/introducao-as-redes-neurais-convolucionais/>>. Acesso em: 7 de Abril de 2021.

DIDÁTICA TECH INTELIGÊNCIA ARTIFICIAL E DATA SCIENCE. 2020. **Visão Computacional e Processamento de Imagens**. Disponível em: <<https://didatica.tech/visao-computacional-e-processamento-de-imagens/>> Acesso em: 1 de Junho de 2021.

EKMAN, P. **Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life**. Henry Holt and Company. 2003. Disponível em: <<https://books.google.com.br/books?id=JJ9xHXo7hPgC>> Acesso em: 5 de Maio de 2021.

FLECK, L. et al. Redes neurais artificiais: princípios básicos. 2016. **Revista Eletrônica Científica Inovação e Tecnologia**, v. 7, n. 15, p. 47–57. Disponível em:

<<https://periodicos.utfpr.edu.br/recit/article/view/4330>>
Acesso em: 13 de Abril de 2021.

FURTADO, M. I. V. **Redes neurais artificiais: uma abordagem para sala de aula.** Ponta Grossa, PR. Atena Editora, 2019. Disponível em: <<https://pdfcoffee.com/e-book-redes-neurais-artificiais-uma-abordagem-para-sala-de-aula-pdf-free.html>> Acesso em: 5 de Maio de 2021.

GRANDO, Aldoir. **Análise facial de emoções utilizando redes neurais no contexto de uma sala de aula inteligente.** 2020. Disponível em: <<https://repositorio.ucs.br/xmlui/handle/11338/6360>> Acesso em: 10 Julho de 2020.

HARA, K.; SAITO, D.; SHOUNO, H. Analysis of function of rectified linear unit used in deep learning. 2015. **International Joint Conference on Neural Networks (IJCNN)**. [S.l.: s.n.], 2015. p. 1–8. Disponível em: <https://www.researchgate.net/publication/308834365_Analysis_of_function_of_rectified_linear_unit_used_in_deep_learning> Acesso em: 2 de Junho de 2021.

IA EXPERT ACADEMY. **Funções de ativação: definição, características, e quando usar cada uma.** 2020. Disponível em: <<https://iaexpert.academy/2020/05/25/funcoes-de-ativacao-definicao-caracteristicas-e-quando-usar-cada-uma/>> Acesso em: 2 de Junho de 2021.

IMAGENET. **ImageNet classification with Python and Keras.** Disponível em: <<https://www.pyimagesearch.com/2016/08/10/imagenet-classification-with-python-and-keras/>> Acesso em: 19 de Maio de 2021.

KERAS. **Image data preprocessing.** Disponível em: <<https://keras.io/api/preprocessing/image/>> Acesso em: 19 de Maio de 2021.

LEMLEY, J.; BAZRAFKAN, S.; CORCORAN, P. Deep learning for consumer devices and services: Pushing the limits for machine learning, artificial intelligence, and computer vision. **IEEE Consumer Electronics Magazine**, v. 6, n. 2, p. 48–56, April 2017. Disponível em: <<https://ieeexplore.ieee.org/document/7879402>> Acesso em: 8 de Junho de 2021.

LI, X.; SHI, Y. Computer vision imaging based on artificial intelligence. 2018. **International Conference on Virtual Reality and Intelligent Systems (ICVRIS)**. [S.l.: s.n.], p. 22–25. Disponível em: <<https://ieeexplore.ieee.org/document/8531341>> Acesso em: 8 de Junho de 2021.

MATOS, Hamilton de. **Desenvolvimento e validação de um sistema de identificação de emoções por visão**

computacional e redes neurais convolucionais com transferência de aprendizado. 2017. 50 f. Dissertação (Engenharia Elétrica) - Universidade Presbiteriana Mackenzie, São Paulo. Disponível em: <<http://tede.mackenzie.br/jspui/bitstream/tede/3553/5/HAMILTON%20DE%20MATOS.pdf>> Acesso em: 13 de Agosto de 2020.

PACHECO, A. G. C. **Classificação de espécies de peixe utilizando redes neurais convolucionais.** CORR, abs/1905.03642, 2019. Disponível em: <<https://paperswithcode.com/paper/190503642>> Acesso em: 8 de Maio de 2021.

PAN, S. J.; YANG, Q. A survey on transfer learning. **IEEE Transactions on Knowledge and Data Engineering**, v. 22, n. 10, p. 1345–1359, Oct 2010. Disponível em: <<https://ieeexplore.ieee.org/document/5288526>> Acesso em: 8 de Maio de 2021.

PANG, B.; LEE, L.; VAITHYANATHAN, S. Thumbs up: Sentiment classification using machine learning techniques. In: **Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10**. Stroudsburg, PA, USA: Association for Computational Linguistics, 2002. (EMNLP '02), p. 79–86. Disponível em: <<https://www.aclweb.org/anthology/W02-1011/>> Acesso em: 10 de Abril de 2021.

RAUTARAY, S.; AGRAWAL, A. Vision based hand gesture recognition for human computer interaction: A survey. **Artificial Intelligence Review**, v. 43, 01 2012. Disponível em: <https://cgvr.informatik.uni-bremen.de/teaching/studentprojects/nui4cars/wp-content/uploads/2013/06/survey_Agrawal_AI2012_handR.ecod.pdf> Acesso em: 11 de Abril de 2021.

SANTOS, Igor Pedro Pinto dos. **Análise de sentimento usando redes neurais de convolução.** 2017. 132f. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2017. Disponível em: <https://www.pel.uerj.br/wp-content/uploads/uploads-manuais/sistemas-inteligentes/Dissertacao_Igor_dos_Santos.pdf> Acesso em: 1 de Junho de 2021.

SILVA, R. E. V. **Um estudo comparativo entre redes neurais convolucionais para a classificação de imagens.** 2018. Disponível em: <http://www.repositorio.ufc.br/bitstream/riufc/39475/1/2018_tcc_revsilva.pdf> Acesso em: 2 de Junho de 2021.

TENSORFLOW. **Transfer learning for image classification.** 2015. Disponível em:

https://www.tensorflow.org/lite/tutorials/model_maker_image_classification> Acesso em: 1 de Maio de 2021.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N.
Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: ALIAGA, D. G. et al. (Ed.). Proceedings... Porto Alegre: **Sociedade Brasileira de Computação**, 2016. Disponível em: <<http://gibis.unifesp.br/sibgrapi16/e proceedings/wuw/7.pdf>> Acesso em: 21 de Abril de 2021.