

Identificando Modelos de Aviões a Partir de Imagens Utilizando Redes Neurais Artificiais

Cândido Signorini Moraes¹; Ricardo Matsumura Araujo²

¹Universidade Federal de Pelotas (UFPEL) – csmoraes@inf.ufpel.edu.br

² Universidade Federal de Pelotas (UFPEL) – ricardo@inf.ufpel.edu.br

1. INTRODUÇÃO

Modelos de aviões, apesar de terem um formato parecido, apresentam certas características que os diferem dos outros, mesmo que pertençam ao mesmo fabricante. Identificar um modelo de avião por modelo ou fabricante pode não ser uma tarefa trivial para a maioria das pessoas. Ciente desse fato, esse artigo apresenta um classificador de aviões capaz de categorizar uma imagem entre oito tipos (classes) diferentes com 91.88% de acurácia confeccionado a partir da rede neural convolucional Alexnet. O classificador obtido poderia ser facilmente implementado em um aplicativo para celulares, que retornaria o modelo da aeronave fotografada. Tal aplicativo seria útil em uma situação de espera em um aeroporto para identificar o modelo de um avião estacionado ou em procedimento de pouso e decolagem, por exemplo.

Redes Neurais são modelos de Inteligência Artificial que podem ser utilizadas, dentre outras tarefas, para classificar algum tipo de informação em categorias pré-definidas. O princípio básico das redes neurais convolucionais parte dos experimentos de Hubel e Wiesel, realizados em 1962, que mostraram que neurônios são ativados quando expostos a algumas linhas e curvas. Convoluções atuam como filtros que, quando aplicados a pedaços de uma imagem, ressaltam suas características.

Em 1998 Yann Lecun aplicou com sucesso a primeira rede neural convolucional, denominada Lenet. No entanto, foi apenas em 2012 que as redes neurais convolucionais se tornaram populares, quando a rede Alexnet, proposta por Alex Krizhevsky, venceu o campeonato de classificação de imagens Imagenet.

Os aviões selecionados para fazer parte desse trabalho foram os modelos A320, A330 e A380 da fabricante francesa Airbus; 72 da também francesa ATR; 737, 747 e 777 da americana Boeing e o E-190 da nossa Brasileira Embraer.

As aeronaves bimotoras A320, 737 e E-190 competem no mesmo segmento de fuselagem estreita (*Narrowbody*), apresentando um corredor único, para distâncias e quantidade de passageiros médias. Já, os bimotores 777 e A330 são aviões de fuselagem larga (*Widebody*), de grande alcance utilizados em voos intercontinentais.

Os aviões A380 e 747 se diferem dos demais por serem quadrimotores e possuírem dois andares. Enquanto o A380 possui um segundo andar em toda sua fuselagem, o 747 o apresenta apenas na parte frontal, formando dessa forma sua icônica “corcunda”. O modelo da Airbus é considerado o maior avião do mundo em termos de capacidade de passageiros, sendo capaz de transportar 525 passageiros na configuração de 3 classes e até 853 quando disponibilizado apenas classe econômica.

O ATR-72 é o único, dentre os oito modelos utilizados nesse projeto que não é impulsionado por um motor a jato. O fato de ser um bimotor turbo-hélice, permite sua utilização em aeroportos com menor infraestrutura, que possuam uma pista

curta ou que não tenham pavimentação. É utilizado, portanto, para rotas curtas com uma capacidade de transportar 72 passageiros.

Na literatura são encontrados trabalhos relacionados como KARACOR (2011) que demonstra um classificador de aeronaves feito com base em uma rede neural artificial não convolucional, capaz de identificar helicópteros e aviões militares, diferentemente da proposta desse artigo que foca apenas em aviões comerciais

2. METODOLOGIA

A arquitetura de rede neural convolucional utilizada para o desenvolvimento desse trabalho foi a Alexnet (KRIZHEVSKY, 2012). A Alexnet possui 650 mil neurônios, 60 milhões de parâmetros dispostos em 8 camadas. As cinco primeiras são do tipo convolucionais enquanto as três últimas são do tipo *Fully Connected*.

Esse trabalho foi desenvolvido usando a biblioteca PyTorch na plataforma Google Colaboratory, que disponibiliza Graphic Processing Units (GPUs), de forma gratuita por um certo tempo, de modo a agilizar o processo de treinamento e validação.

O conjunto de imagens (Dataset) utilizado para treinamento e validação foi retirado do site JetPhotos. Ao todo foram coletadas 10800 imagens, das quais 10000 foram destinadas ao treinamento e 800 separadas para teste da proposta, distribuídas de forma igualitária entre todos os modelos de aeronaves, isto é, 250 imagens para treinamento e 100 para teste de cada modelo. O dataset inclui imagens dos aviões de perfil, com o nariz tanto para esquerda quanto para direita, assim como o avião no solo ou em processo de decolagem ou pouso.



Figura 1 – Imagens de aviões pertencentes ao dataset. (A) Embraer E-190, (B) ATR-72, (C) Airbus A320 e (D) Boeing 747

As imagens obtidas do site JetPhotos possuem resolução 400 x 225. Como a resolução de entrada da rede Alexnet é de 256 x 256, as imagens tiveram de ser redimensionadas e para manter a proporção (aspect ratio), um *padding* - inserção de barras pretas - foi aplicado.

Com o objetivo de acelerar o processo de treinamento, foi utilizado *Transfer Learning*, que é o processo de aproveitar parte de uma rede já treinada, capaz de generalizar formas e padrões para uma nova aplicação. No caso desse trabalho,

foram reaproveitados os pesos da parte convolucional da Alexnet utilizada na competição Imagenet, enquanto a parte *Fully Connected* foi totalmente treinada para este problema em específico.

Como já citado anteriormente, devido as características do dataset, que possui aviões com o nariz tanto para esquerda como para direita, além de estar inclinado em alguns casos, foi aplicado *RandomHorizontalFlip* no contexto de *Data Augmentation*. Esse procedimento espelha de forma aleatória, o eixo horizontal da imagem, com o objetivo de aumentar a variabilidade das amostras. Houve a tentativa de se utilizar também o *RandomRotation*, que aplica uma inclinação de 15°, uma vez que em algumas fotos, os aviões aparecem inclinados durante o procedimento de pouso ou decolagem. No entanto, seu uso piorou os resultados da rede e por isso não foi utilizado na versão final.

Uma preocupação que se teve, foi de evitar que a rede aprendesse a diferenciar aviões pelo logo da companhia aérea, uma vez que não há como garantir que não existam mais fotos de uma companhia aérea do que outra no dataset, ou em um determinado tipo de avião. Por isso, as imagens foram transformadas para escala de cinza.

3. RESULTADOS E DISCUSSÃO

Inicialmente, a transformação para escala de cinza não foi bem-sucedida, tendo a rede obtido péssimos resultados. Em uma segunda tentativa, após testes de diferentes valores de *Learning Rate* e *Momentum*, além da retirada do *RandomHorizontalFlip* conseguiu-se obter um resultado bastante satisfatório.

O melhor resultado foi obtido utilizando um otimizador do tipo Descida do Gradiente Estocástico (SGD) com *Learning Rate* igual a 0,001 e *Momentum* igual a 0,7. Como critério de Loss foi utilizado o CrossEntropy. Devido às restrições de tempo de execução da plataforma Google Colaboratory, o treinamento ocorreu em apenas 10 épocas, o que foi suficiente para atingir uma acurácia de 91.88%.

Durante o treinamento o Loss variou de 1,34 no início do treinamento até 0,16 ao final. No conjunto de testes, a acurácia TOP1, ou seja, a porcentagem de acertos obtida pela rede foi de 91.88 %.

A rede neural calcula a probabilidade de uma dada foto pertencer a cada uma das oito classes. A acurácia TOP5 leva em consideração o número de vezes em que a classe correta esteve entre as 5 classes mais prováveis de acordo com a saída da rede. A rede em questão obteve acurácia TOP5 foi de 99.5%.

A Tabela 1 mostra a acurácia obtida pela rede em cada classe.

Aeronave	Acurácia (%)
Airbus A320	97
Airbus A330	89
Airbus A380	82
ATR-72	99
Boeing 737	83
Boeing 747	94
Boeing 777	84
Embraer E-190	96

Tabela 1 – Acurácia obtida por cada classe

Analisando a Tabela 1 é possível notar que a classe em que a rede se saiu melhor é aquela que apresenta o modelo 72 da fabricante francesa ATR. Esse resultado já era esperado, uma vez que esse modelo possui características únicas

em relação aos demais, como a presença de um motor a hélice, asas e estabilizadores horizontais acima das janelas, e tamanho reduzido em comparação aos demais.

4. CONCLUSÕES

Esse trabalho apresentou um classificador de modelos de aeronaves utilizando a arquitetura de redes neurais convolucionais Alexnet. A acurácia obtida foi de 91.88 %, o que mostra que a solução funcionou de maneira adequada e de acordo com o esperado.

Os valores de acurácia obtidos, demonstraram que a rede foi capaz de abstrair e aprender de forma bastante satisfatória a identificar modelos de aeronaves, conseguindo classificar com elevada precisão um avião entre oito possibilidades diferentes.

5. REFERÊNCIAS BIBLIOGRÁFICAS

LECUN, Yann et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, v. 86, n. 11, p. 2278-2324, 1998.

KRIZHEVSKY, Alex; SUTSKEVER, Ilya. Hinton. G., ImageNet Classification with Deep. Convolutional Neural Networks, **NIPS**, 2012.

JETPHOTOS. Disponível em: <https://www.jetphotos.com/>

KARACOR, Adil Gursel; TORUN, Erdal; ABAY, Rasit. Aircraft classification using image processing techniques and artificial neural networks. **International Journal of Pattern Recognition and Artificial Intelligence**, v. 25, n. 08, p. 1321-1335, 2011.

TAN, Chuanqi et al. A survey on deep transfer learning. In: **International Conference on Artificial Neural Networks**. Springer, Cham, 2018. p. 270-279.