

Classificador de cachorros e gatos

Projeto de Machine Learning

Matheus Andrade de Souza

Prof. André Gustavo Maletzke

Dataset

- Possui um total de 37461 imagens distintas de cachorros e gatos previamente separadas em três subconjuntos:
 - Treino: contém 20000 imagens no total, sendo 10000 de cachorros e 10000 de gatos.
 - Teste: contém 12461 imagens, sendo 6219 de cachorros e 6242 de gatos.
 - Validação: contém 5000 imagens, sendo 2500 cachorros e 2500 gatos.
- Todas as imagens estão em formato JPG.
- Imagens tem um tamanho médio de aproximadamente 25 KB.
- Imagens possuem um tamanho de resolução diferentes umas das outras.

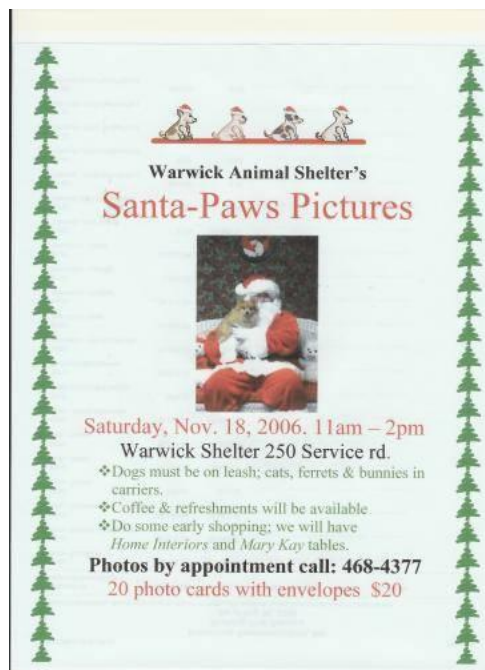
Dificuldades

- Conjunto de dados possui elementos inconsistentes.
- Necessário realizar uma pré-seleção sobre os dados manualmente.



Pré-processamento

Imagens foram selecionadas e removidas manualmente.



YAHOO! MAIL



Pré-processamento

- Ainda existem dados inconsistentes após a remoção manual devido ao grande volume de imagens.
- Conjuntos de treino, teste e validação agrupados em um único conjunto.
- Nova separação em subconjuntos de treino e teste com uma proporção de 20% de dados para o teste.
- Novo *dataset* possui 37240 imagens no total (37461 originalmente):
 - Treino: 29791 imagens no total, sendo 14913 cachorros e 14878 gatos.
 - Teste: 7449 imagens no total, sendo 3729 cachorros e 3720 gatos.
- Todas as imagens foram redimensionadas para 224x224 *pixels* durante o treino.

Distribuição total

Distribuição de Dogs e Cats

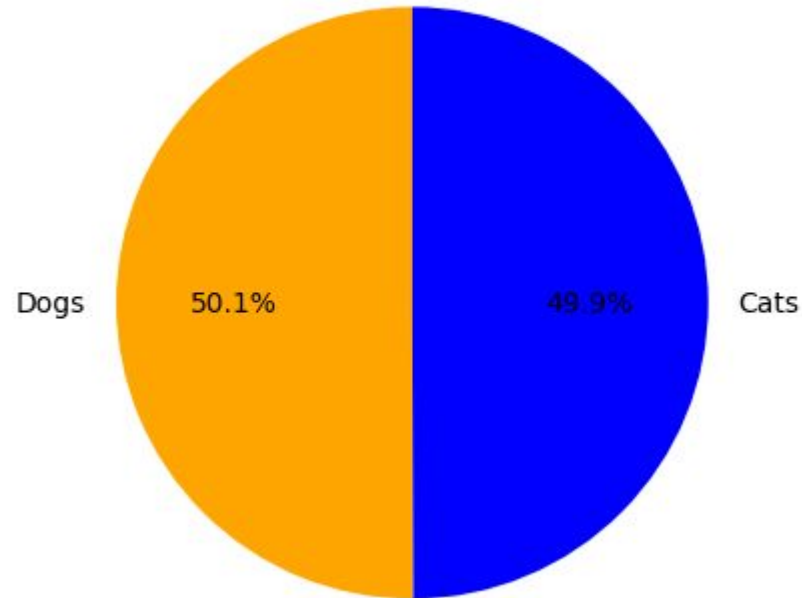


Figura 1: Distribuição de classes
(Autoria própria)

CNN (*Convolutional Neural Network*)

- CNN significa Rede Neural Convolucional.
- Técnica *Machine Learning* inspirada no funcionamento do cérebro humano.
- Utiliza camadas convolucionais para extrair características em dados.
- No final utiliza-se uma rede neural com camadas totalmente conectadas para determinar a probabilidade das classes.

Convolution Neural Network (CNN)

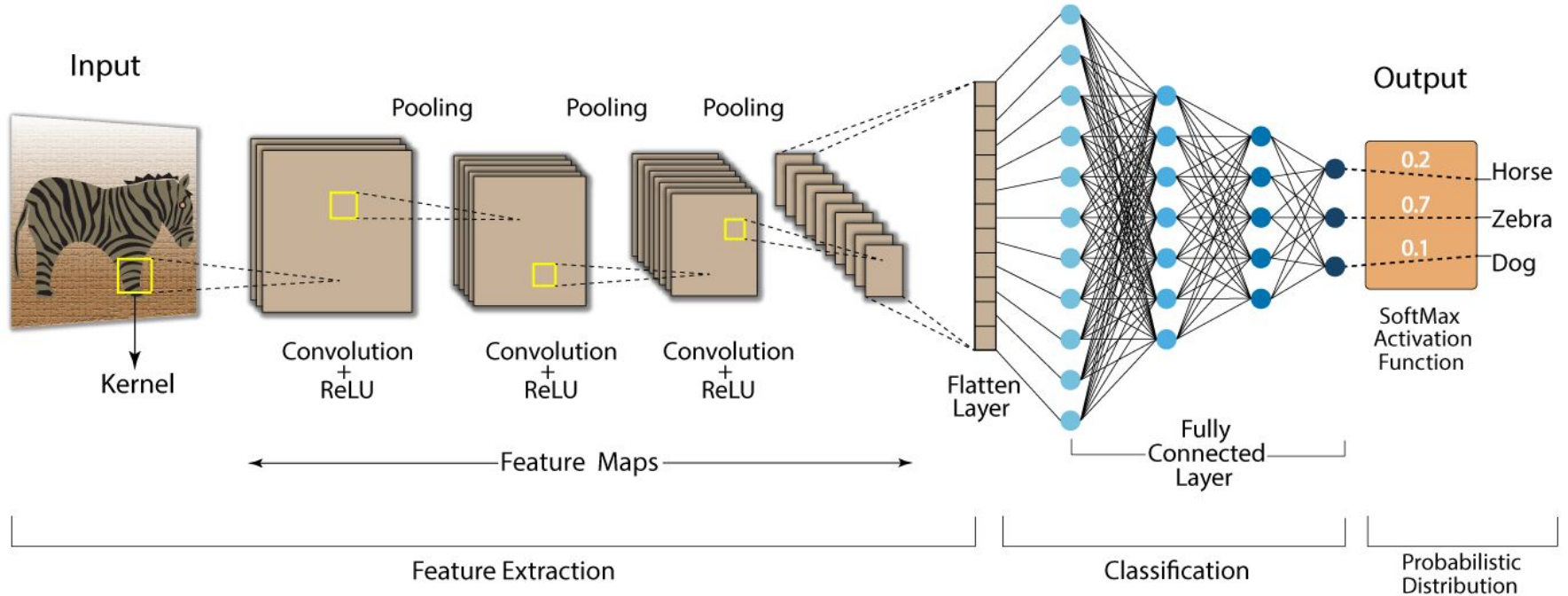


Figura 2: Estrutura CNN
(Basics of CNN in Deep Learning, 2023)

Treinamento

- Foram treinados quatro modelos diferentes:
 - Modelo A1: 50-layer Residual Network(resnet50) pré-treinada com 14 camadas congeladas.
 - Modelo A2: 50-layer Residual Network(resnet50) pré-treinada com 88 camadas congeladas.
 - Modelo A3: 50-layer Residual Network(resnet50) pré-treinada com 174 camadas congeladas.
 - Modelo B: MobileNetV3-Small pré-treinada.
- Todos os modelos foram treinados com um tamanho de lotes igual a 32.
- Todos os modelos foram treinados com 15 épocas.
- Algumas imagens serão aleatoriamente espelhadas horizontalmente.
- Redes pré-treinadas, ou seja, os pesos de suas camadas já estão ajustados com base no *dataset* utilizado durante o pré-treino.
- Técnica de treinamento “*One Cycle Policy*” (Política de Um Ciclo).
- Treinados com uma taxa de aprendizado de 0.00012.

Modelo A1 (resnet50)

- Possui um total de 176 camadas.
- As 14 primeiras camadas foram congeladas, ou seja, seus pesos não foram alterados durante o treinamento.
- Duração do treinamento de aproximadamente 60 minutos.

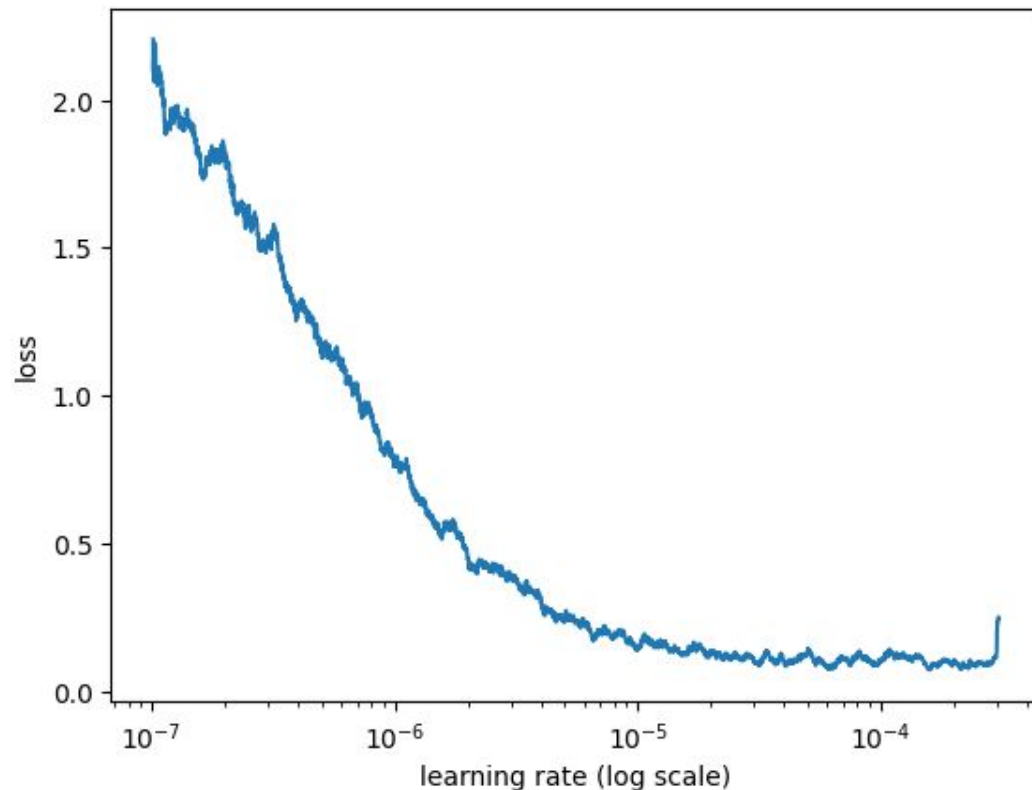
Estimando a taxa de aprendizado

Foi utilizada a função

`lr.find()` da biblioteca `ktrain`
para estimar a taxa.

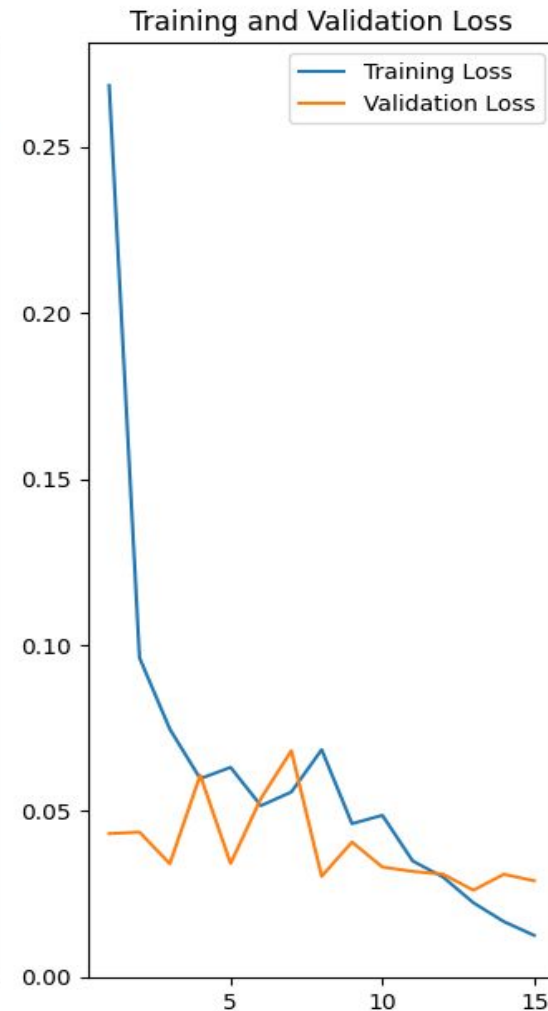
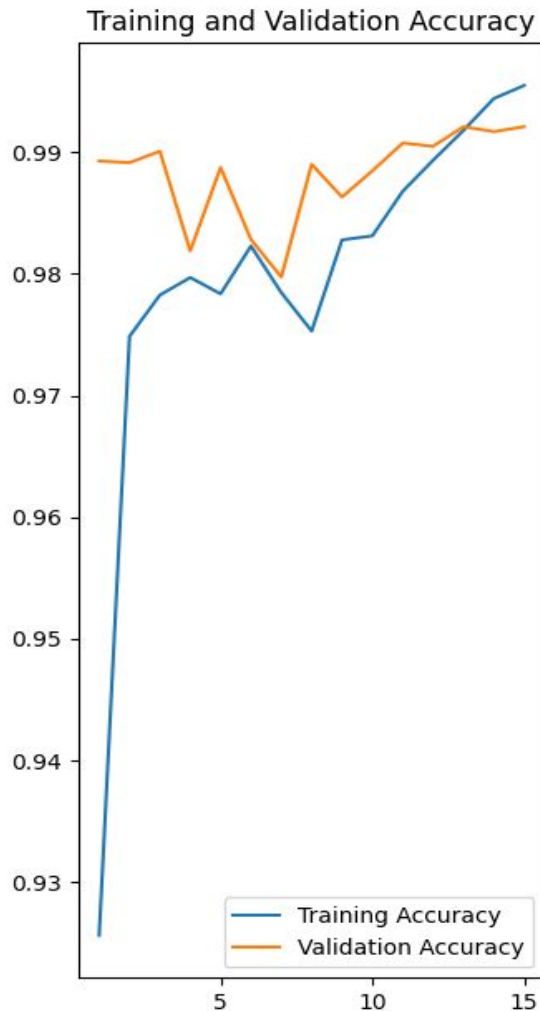
Taxa escolhida de 0.00012 para
todos os modelos.

Figura 3: Taxa de aprendizado
(Autoria própria).

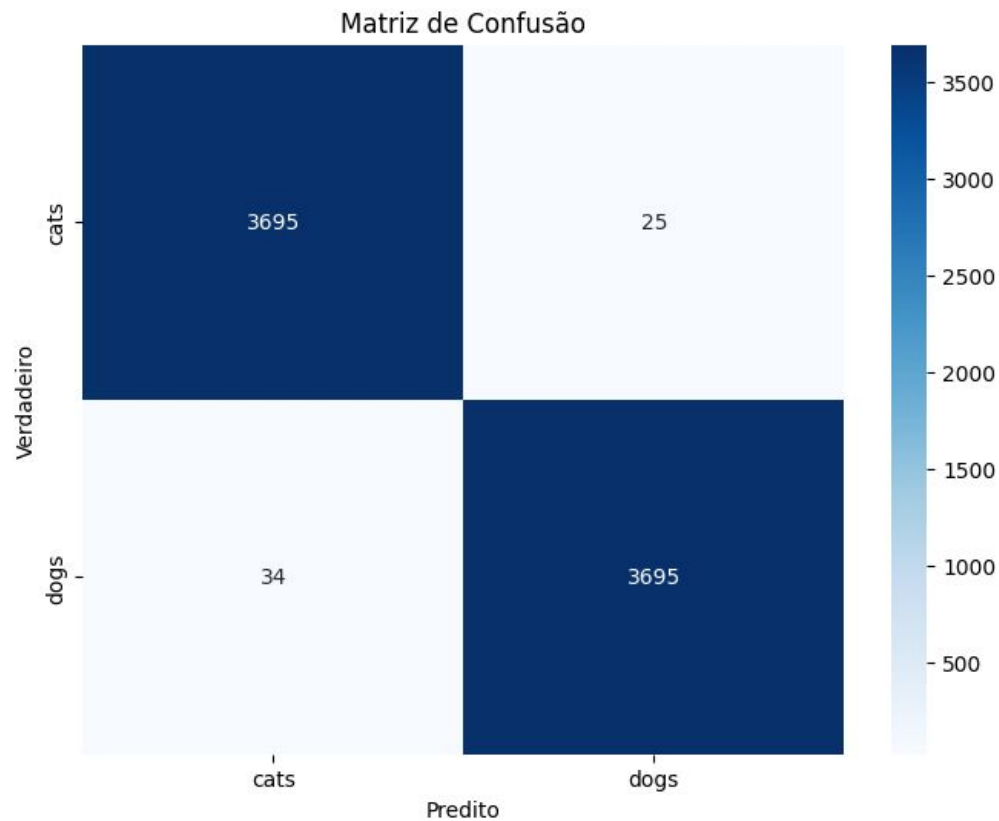


Modelo A1 (resnet50) Resultados

- Acurácia final de
 $0.9920795 \approx 99.21\%$
- Loss final de
 $0.029000513223961362 \approx 2.9\%$



Modelo A1 (resnet50) - Resultados

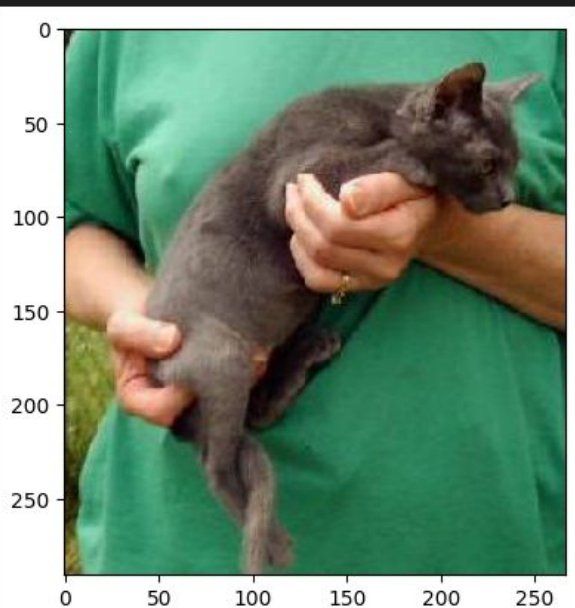


Fonte: Autoria própria.

Analizando erros (Cachorros)

Dataset possui erros nos labels de seus dados.

File classified 'dogs\dog (4647).jpg' as 'CAT' [0.51531357 0.4846864]



File classified 'dogs\dog (5189).jpg' as 'CAT' [9.9967718e-01 3.2280004e-04]



Analizando erros (Cachorros)

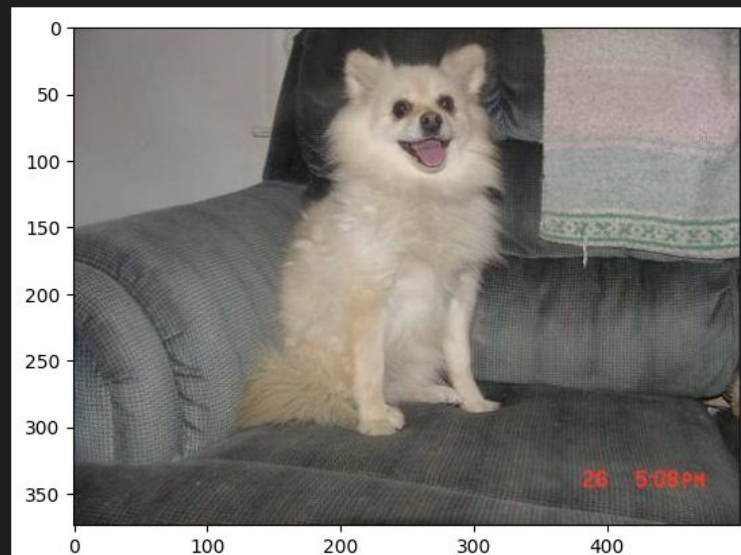
File classified 'dogs\dog.9683.jpg' as 'CAT' [0.9411265 0.0588735]



File classified 'dogs\dog.2458.jpg' as 'CAT' [0.9822326 0.0177674]



File classified 'dogs\dog.4616.jpg' as 'CAT' [0.9978435 0.00215652]

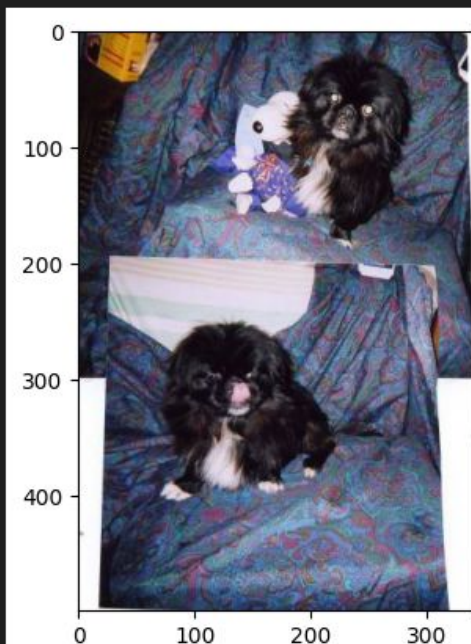


Analizando erros (Cachorros)

File classified 'dogs\dog (926).jpg' as 'CAT' [0.72789776 0.2721023]



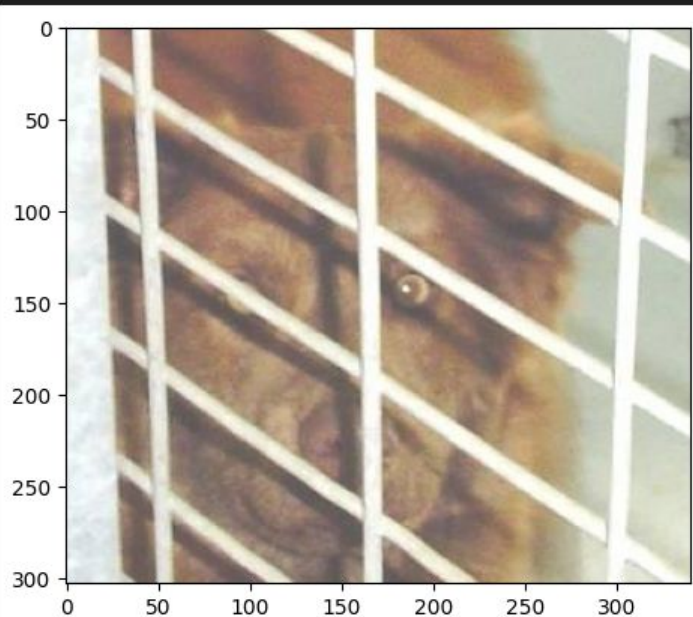
File classified 'dogs\dog.1069.jpg' as 'CAT' [0.74389815 0.25610182]



Analizando erros (Gatos)

Dataset possui mais erros nos *labels* de seus dados.

File classified 'cats\cat (3121).jpg' as 'DOG' [0.00111568 0.9988844]



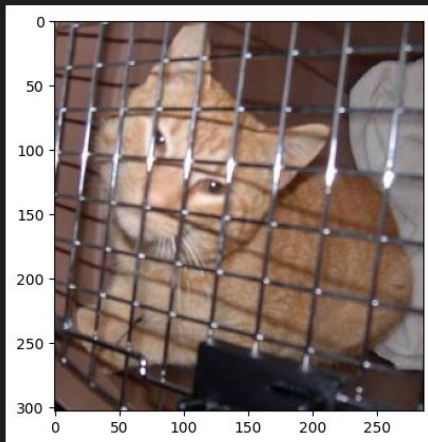
File classified 'cats\cat.11222.jpg' as 'DOG' [9.603498e-14 1.000000e+00]



Analizando erros (Gatos)

Detector de grades ?

File classified 'cats\cat.11281.jpg' as 'DOG' [0.08817233 0.91182774]



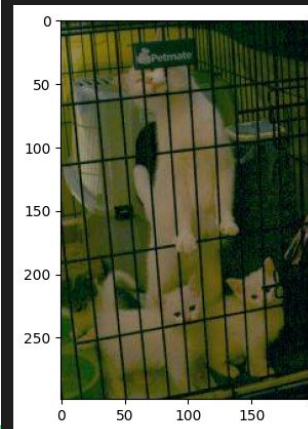
File classified 'cats\cat.5173.jpg' as 'DOG' [0.03278645 0.9672135]



File classified 'cats\cat.2267.jpg' as 'DOG' [0.00806716 0.9919328]



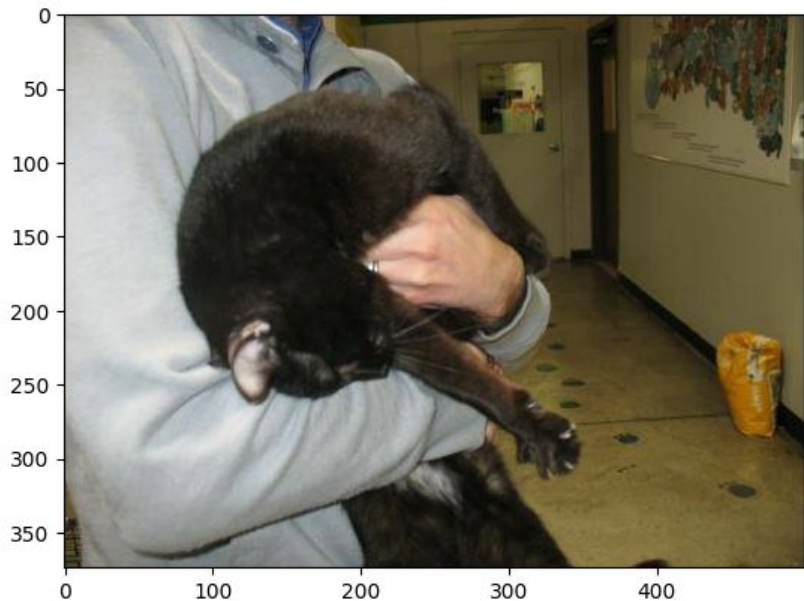
File classified 'cats\cat.12378.jpg' as 'DOG' [0.04470102 0.95529896]



Analizando erros (Gatos)

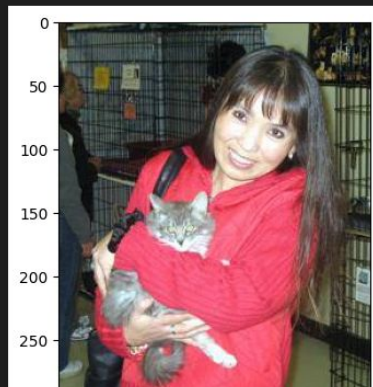
Detector de pessoas ?

File classified 'cats\cat.2748.jpg' as 'DOG' [0.0844738 0.9155262]



Fonte: Autoria própria.

File classified 'cats\cat.4190.jpg' as 'DOG' [0.00267871 0.9973213]

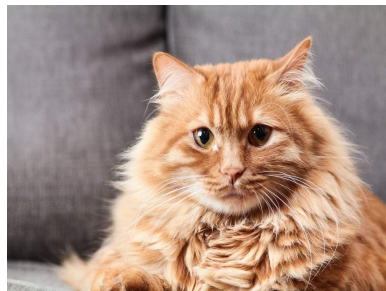


File classified 'cats\cat.11297.jpg' as 'DOG' [0.35174605 0.6482539]



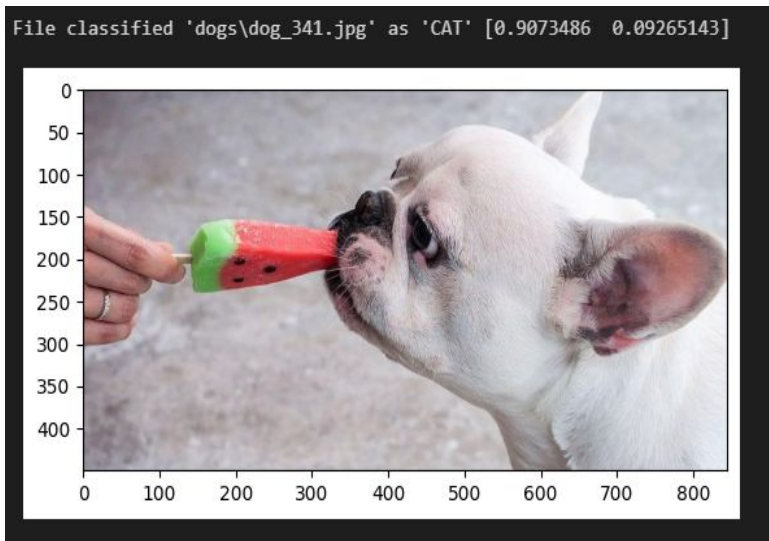
Testes com *dataset* diferente.

- Dataset possui um total de 676 imagens, sendo 336 gatos e 340 cachorros.



Fonte: *Dataset* de teste.

Analizando erros do novo *dataset* (Cachorros)



Analizando erros do novo *dataset* (Gatos)

File classified 'cats\cat_280.jpg' as 'DOG' [0.3227242 0.6772758]

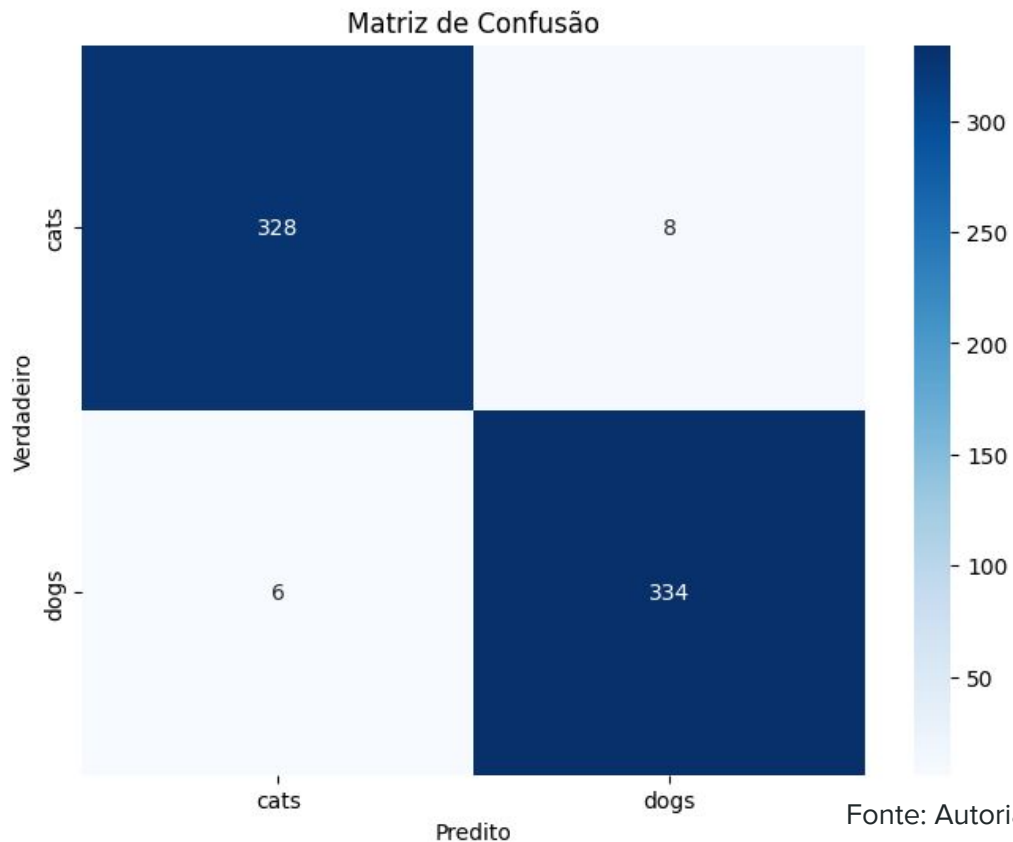


File classified 'cats\cat_548.jpg' as 'DOG' [0.413508 0.586492]



Resultados teste como *dataset* diferente.

Acurácia: 97.93%



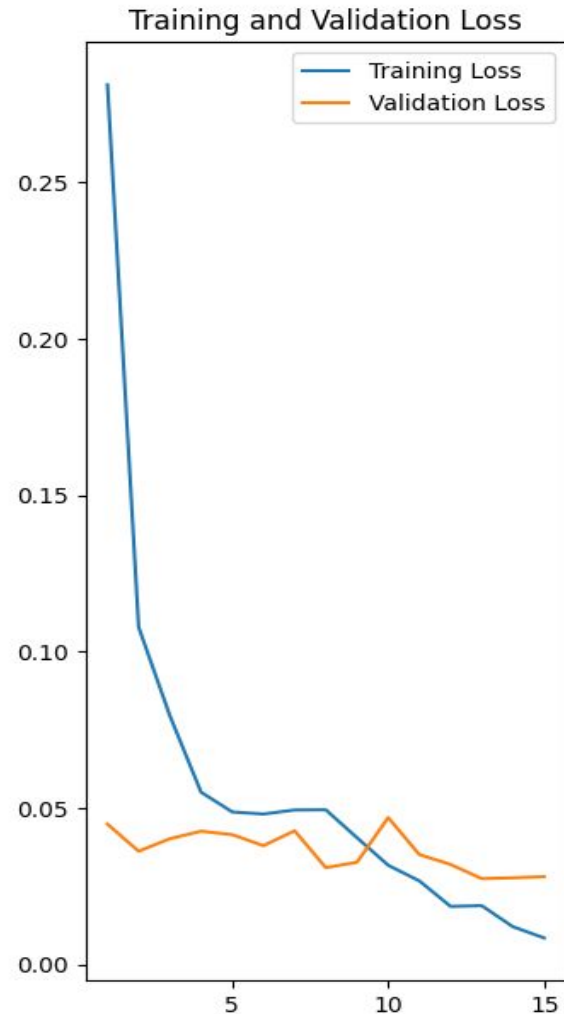
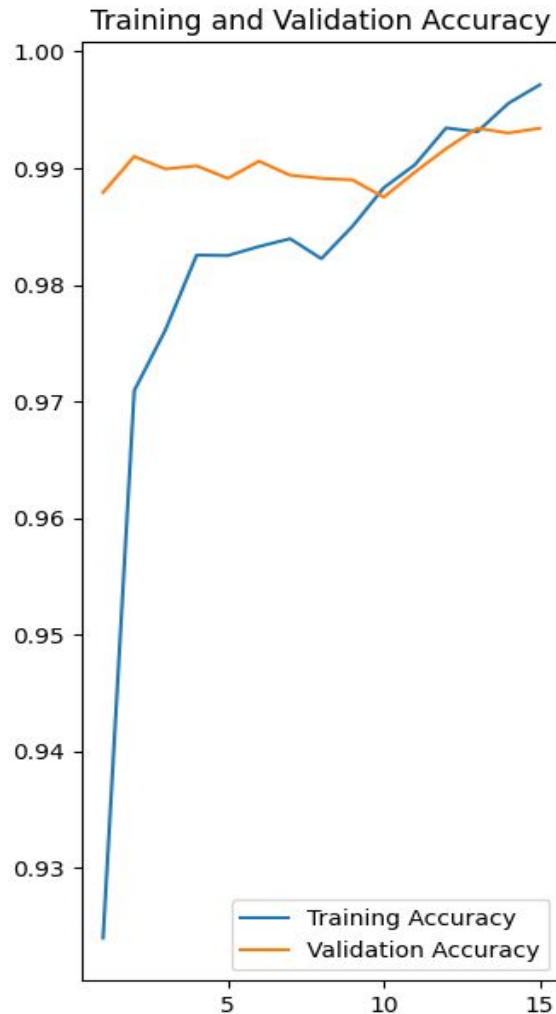
Fonte: Autoria própria.

Modelo A2 (resnet50)

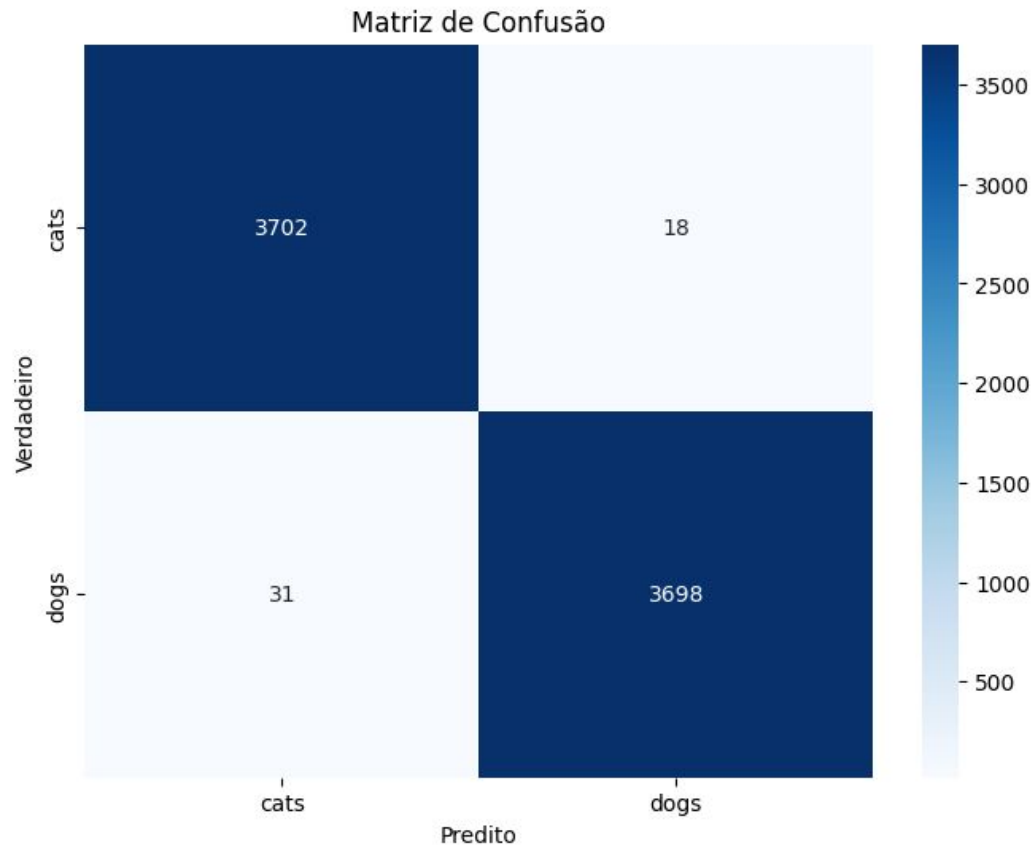
- Possui um total de 176 camadas.
- As 88 primeiras camadas foram congeladas.
- Duração do treinamento de aproximadamente 34 minutos.

Modelo A2 (resnet50) Resultados

- Acurácia final de $0.9934219 \approx 99.34\%$
- Loss final de $0.02815775488966852 \approx 2.81\%$



Modelo A2 (resnet50) - Resultados



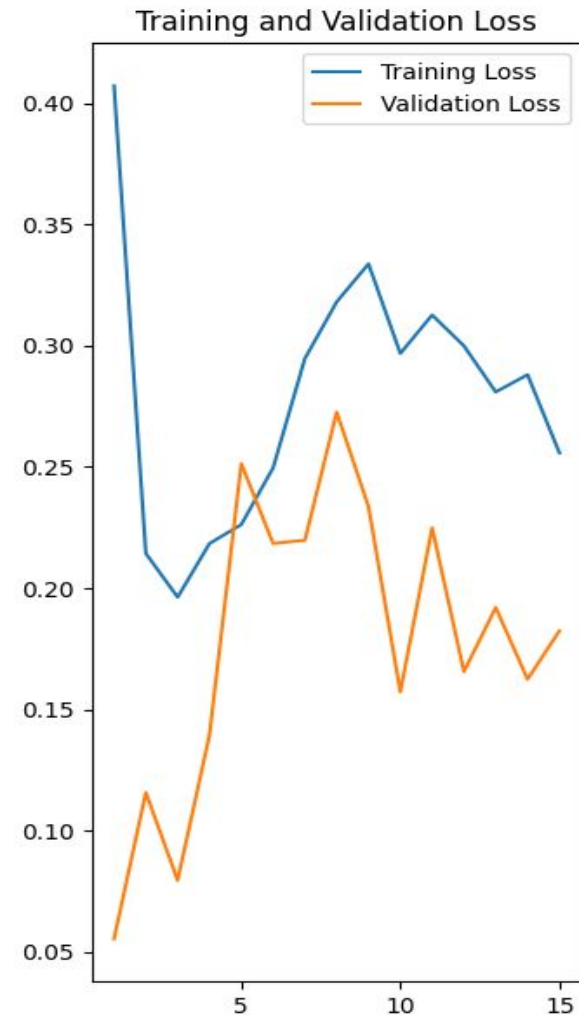
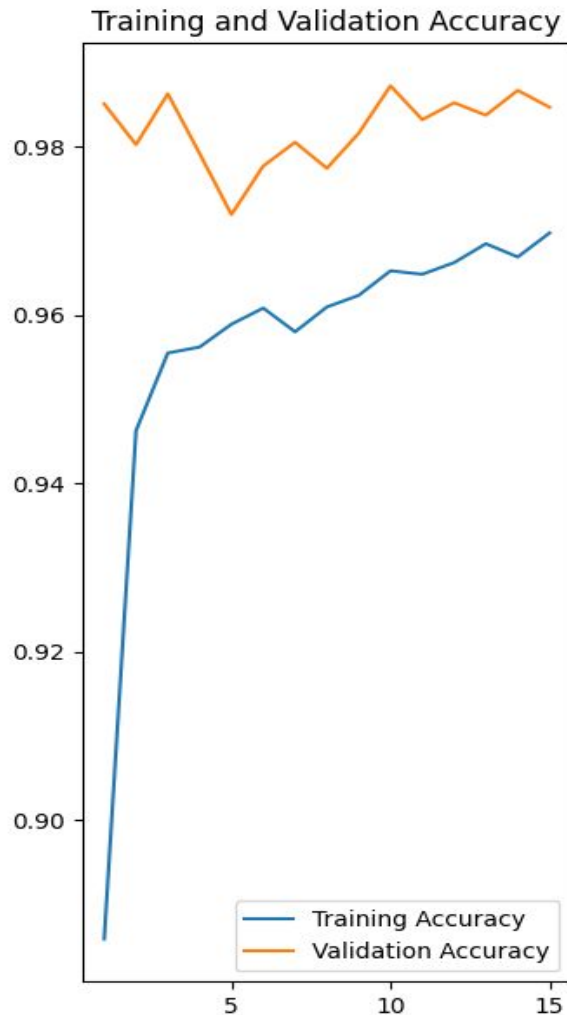
Fonte: Autoria própria.

Modelo A3 (resnet50)

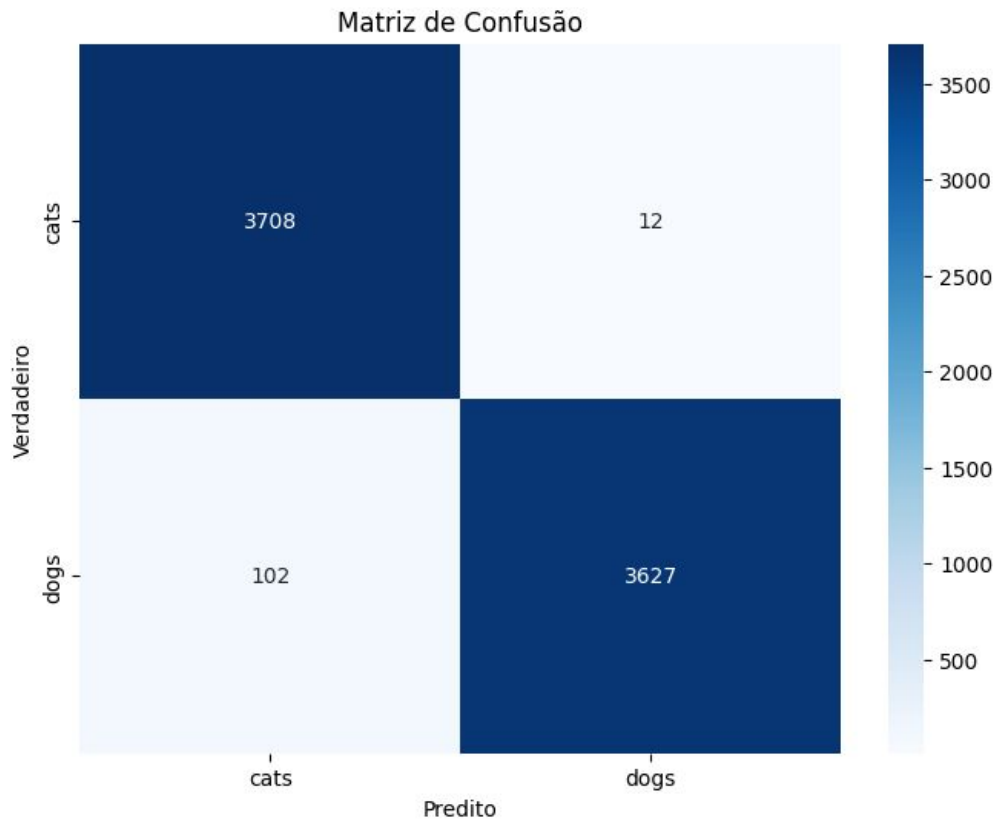
- Possui um total de 176 camadas.
- As 174 primeiras camadas são congeladas por padrão, logo apenas as duas últimas camadas tem seus pesos alterados.
- Duração do treinamento de aproximadamente 22 minutos.

Modelo A3 (resnet50) Resultados

- Acurácia final de
0.9846959 \approx 98.47%
- Loss final de
0.18231066803001056 \approx
18.23%



Modelo A3 (resnet50) - Resultados

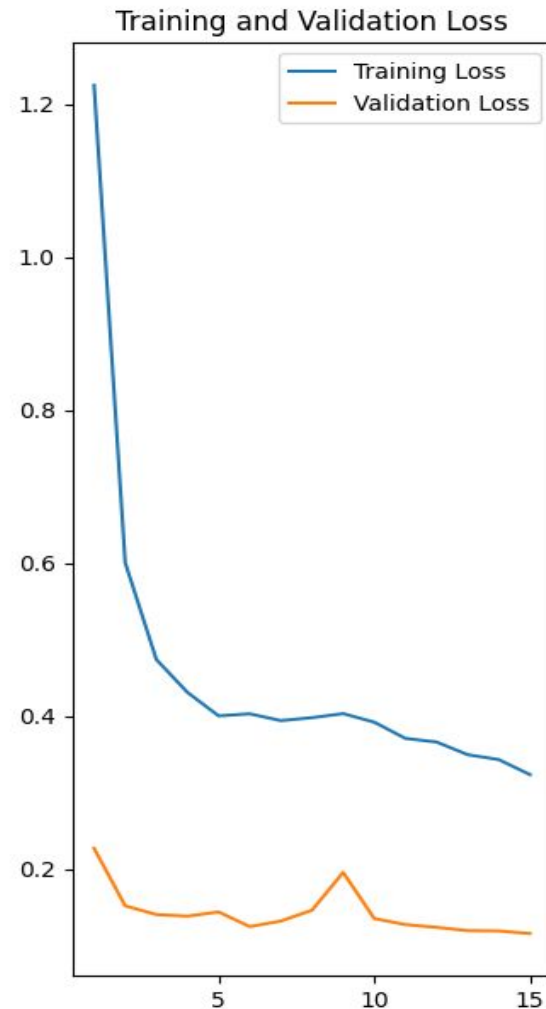
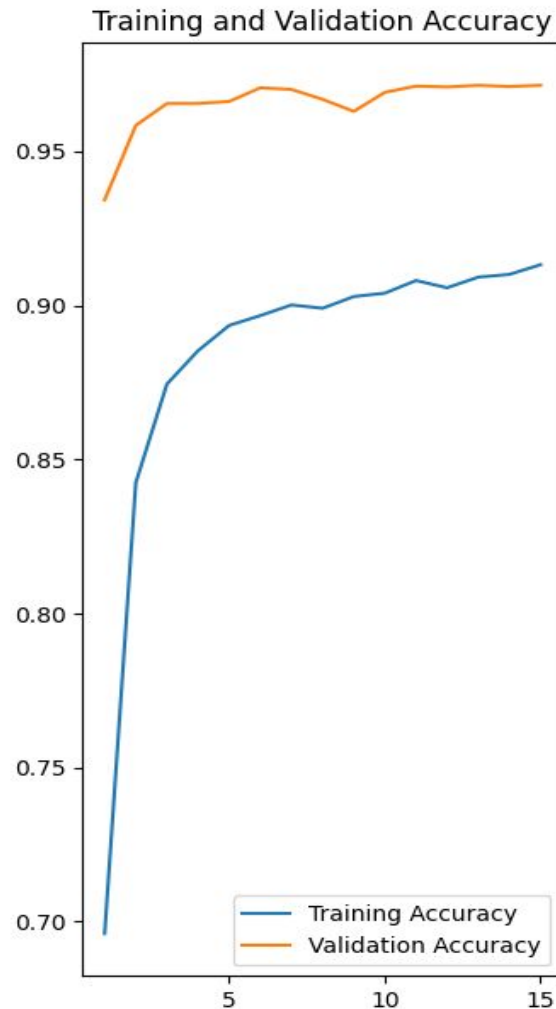


Modelo B (mobilenetv3)

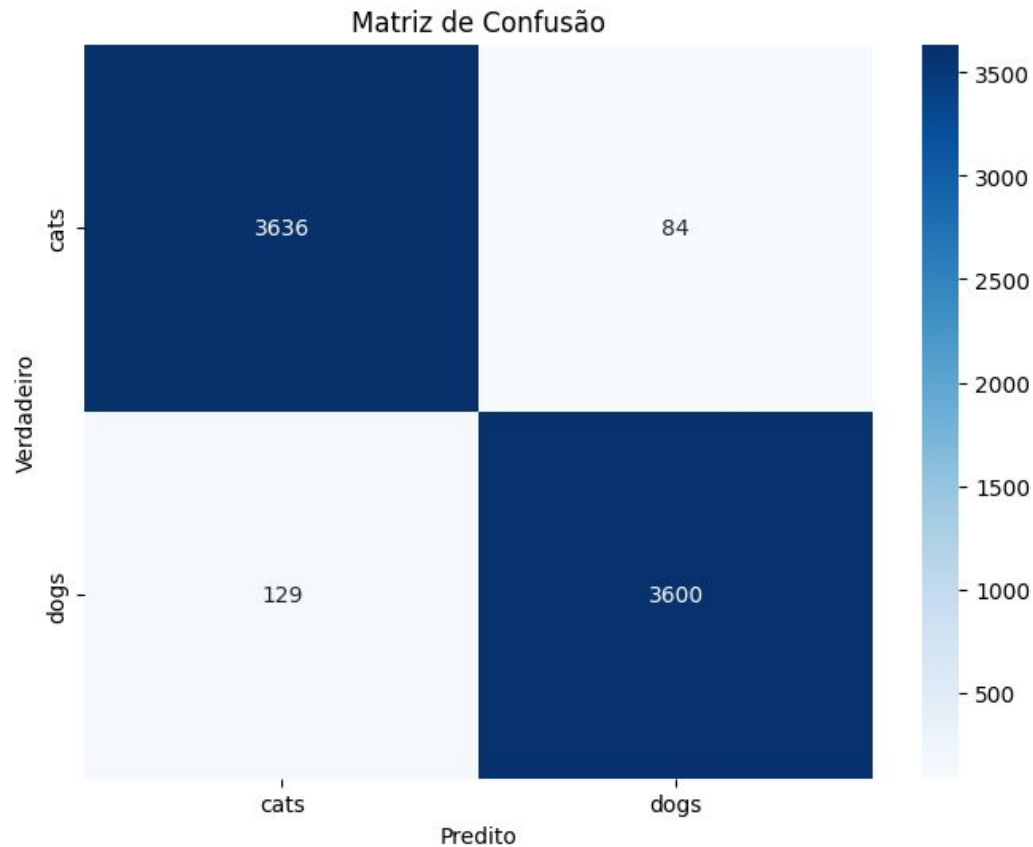
- Rede estruturada com o objetivo de ser mais leve durante o treinamento e previsões para ser utilizada em dispositivos móveis ou embarcados.
- Possui um total de 231 camadas.
- As primeiras camadas 228 são congeladas por padrão, logo apenas as três últimas camadas tem seus pesos alterados.
- Duração do treinamento de aproximadamente 21 minutos.

Modelo B (mobilenetv3) Resultados

- Acurácia final de $0.97140557 \approx 97.14\%$
- Loss final de $0.11548270444097462 \approx 11.54\%$



Modelo B (mobilenetv3) - Resultados



Fonte: Autoria própria.

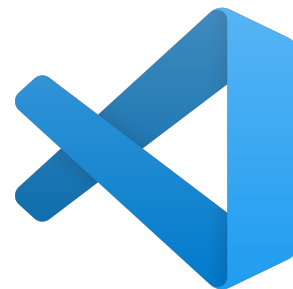
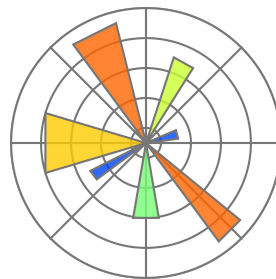
Conclusão

- *Dataset* de treino possui viés, incoerência na rotulação e dados inconsistentes.
- Modelos provavelmente super ajustados ao *dataset*(*overfitting*).

Modelo	Acurácia	Loss	Tempo de treino aproximado
A1	99.21%	2.9%	60 min
A2	99.34%	2.81%	34 min
A3	98.47%	18.23%	22 min
B	97.14%	11.54%	21 min

Tabela 1: Medição dos modelos
(Autoria própria).

Tecnologias utilizadas



Referências

Dataset de treino. dogs vs cats. Disponível em: <<https://www.kaggle.com/datasets/moazeldsokyx/dogs-vs-cats>>.

Figura 2: Estrutura CNN. KALITA, D. Basics of CNN in Deep Learning. Disponível em:
<<https://www.analyticsvidhya.com/blog/2022/03/basics-of-cnn-in-deep-learning/>>.

MAIYA, A. S. amaiya/ktrain. Disponível em: <<https://github.com/amaiya/ktrain>>.

Dataset de testes. Cats and Dogs image classification. Disponível em:
<<https://www.kaggle.com/datasets/samuelcortinhas/cats-and-dogs-image-classification/>>.