



Reconhecimento de Doodles

Trabalho Final - Deep Learning

Daniel dos Santos

19 de dezembro de 2022

Objetivo

Utilizar os conhecimentos de Deep Learning adquiridos em aula, para resolver um problema de visão computacional. O trabalho deve seguir algumas diretrizes:

Objetivo

Utilizar os conhecimentos de Deep Learning adquiridos em aula, para resolver um problema de visão computacional. O trabalho deve seguir algumas diretrizes:

- Não deve ser muito simples;

Objetivo

Utilizar os conhecimentos de Deep Learning adquiridos em aula, para resolver um problema de visão computacional. O trabalho deve seguir algumas diretrizes:

- Não deve ser muito simples;
- Não deve ser muito complicado;

Objetivo

Utilizar os conhecimentos de Deep Learning adquiridos em aula, para resolver um problema de visão computacional. O trabalho deve seguir algumas diretrizes:

- Não deve ser muito simples;
- Não deve ser muito complicado;
- Ter um benchmark para o problema.

Tema proposto

Classificação de Doodles



Tema proposto

Classificação de Doodles



Será utilizado o conjunto de dados "Quick, Draw!" da Google, que contém 365 categorias de doodles com cerca de 120000 exemplos para cada categoria.

Exemplo do conjunto de dados i

Cada elemento possui a seguinte estrutura:

```
{  
  "key_id": "5891796615823360",  
  "word": "nose",  
  "countrycode": "AE",  
  "timestamp": "2017-03-01 20:41:36.70725 UTC",  
  "recognized": true,  
  "drawing": [[[129, 128, 129, 129, 130, ...]]]  
}
```


Exemplo do conjunto de dados ii

Formato do desenho é dado por traços, da seguinte forma:

```
[  
  [ // Primeiro traço  
  
    [x0, x1, x2, x3, ...],  
    [y0, y1, y2, y3, ...],  
    [t0, t1, t2, t3, ...]  
  ],  
  ... // Demais traços  
]
```

Exemplo do conjunto de dados iii



Figura 1: Exemplo de imagem já processada, em seu tamanho original (256x256).

Exemplo do conjunto de dados iv

Devido a uma limitação de tempo, foi utilizado um subconjunto sorteado aleatoriamente de apenas 20 categorias do total de 365, sendo ele:

- Angel;
- Bicycle;
- Cat;
- Alarm Clock;
- Cup;
- Diamond;
- Dolphin;
- Door;
- Eye;
- Eyeglasses;
- Hat;
- Headphones;
- Helicopter;
- House;
- House Plant;
- Octopus;
- Popsicle;
- Ship;
- Umbrella;
- Windmill;

Ferramentas i

C++ Para processamento das imagens com OpenCV;
Python Treinamento da rede neural utilizando Keras;
Treinado localmente em um MacBook Air M1, 8Gb utilizando Metal.

Pré-processamento i

Foi utilizado o OpenCV para traçar um linha entre os pontos dos traços e gerar as imagens, de forma que ficassem alinhadas ao canto superior esquerdo.

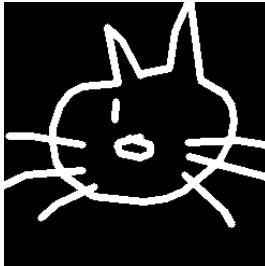
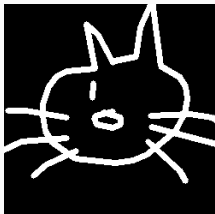


Figura 2: Imagem gerada com o OpenCV

Pré-processamento ii

Aplicar Anti-Aliasing e diminuindo a imagem para 32x32.



(a) Original



(b) Anti-Aliasing



(c) Redimensionado

Figura 3: Pré-processamento

Visualização i

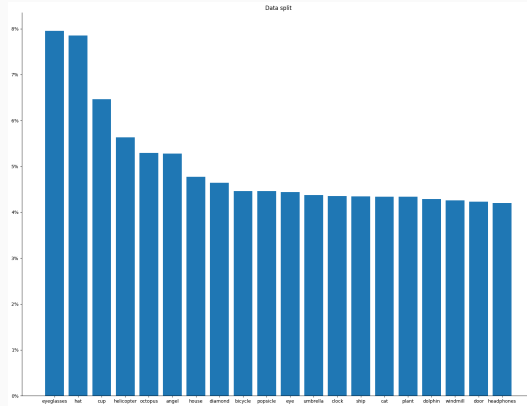


Figura 4: Distribuição das categorias no conjunto de treino é balanceada.

Visualização ii

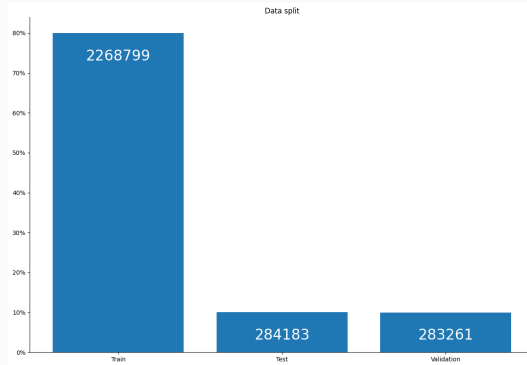


Figura 5: Distribuição entre treino e teste.

Topologia i

Elshenaway e Guirguis (2021) sugerem a seguinte topologia para rede:

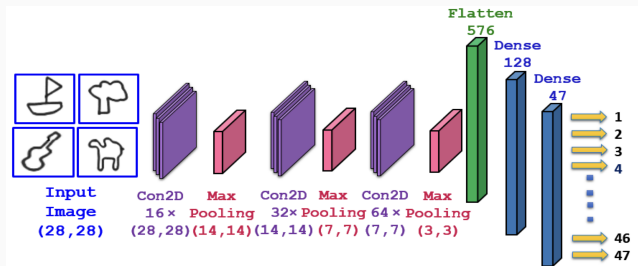


Figura 6: Topologia sugerida por Elshenaway e Guirguis (2021)

Topologia ii

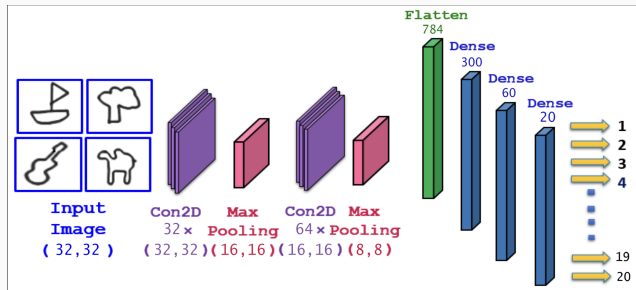


Figura 7: Topologia utilizada neste trabalho.

Validação Cruzada i

Método: **Kfold** K=5, mantendo a proporção das classes;

Validação Cruzada ii

Tabela 1: Resultados da Validação Cruzada

Fold	Acurácia	Tempo de Treinamento (s)
I	0.949	2311
II	0.948	1855
III	0.949	2115
IV	0.949	2776
V	0.950	1610
Média	0.949	2133.4

Resultados i

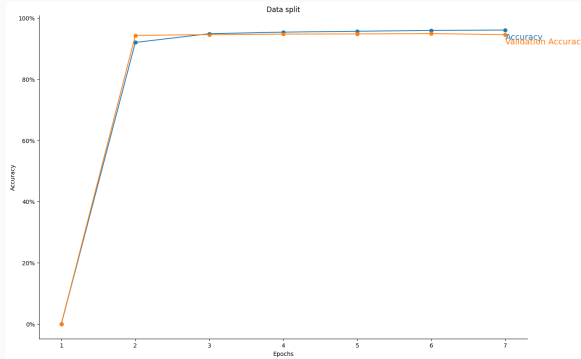


Figura 8: Convergência do modelo por epoch

Resultados ii

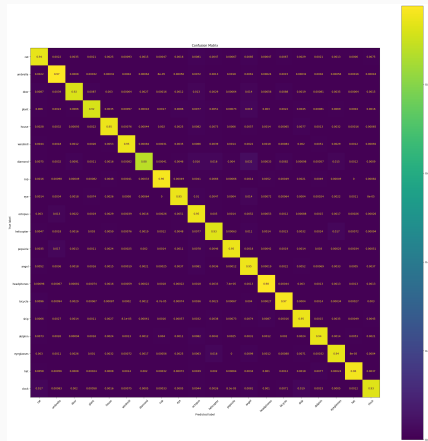


Figura 9: Matriz de confusão do melhor modelos.

Resultados iii

Tabela 2: Resultados no conjunto de teste

Acurácia	F1-Score	Tempo de Treinamento (s)
0.945	0.945	7063

Trabalhos Futuros

- Todas as 365 categorias;
- Mais pré-processamento (deskewing, cropping, etc);
- Testar outras topologias de rede;

Extra



Referências



Elshenaway, A. R. and Guirguis, S. K. (2021).

On-air hand-drawn doodles for iot devices authentication during covid-19.

IEEE Access, 9:161723–161744.



Monica, E. D., Davu, P., Caroline, C. P., and Jagannath, D. (2019).

Doodle recognition using machine learning for hearing and speech-impaired people.

In 2019 2nd International Conference on Signal Processing and Communication (ICSPC), pages 109–112.