

Redes Neurais Convolucionais

Arthur Abrahão Santos Barbosa
Universidade Federal de Pernambuco
Centro de Informática
Pernambuco, Brasil
aasb2@cin.ufpe.br

Filipe Samuel da Silva
Universidade Federal de Pernambuco
Centro de Informática
Pernambuco, Brasil
fss8@cin.ufpe.br

I. Objetivos

A. Objetivo Geral

Desenvolver um Classificador Multiclasse que reconheça as imagens do dataset CIFAR100 [1].

B. Objetivos Específicos

- Compreender a implementação de uma Rede Neural Convolucional
- Demonstrar a Importância do Aprendizado Profundo e suas aplicações
- Demonstrar a eficiência de três arquiteturas importantes para a história do Deep Learning

II. Justificativa

Este projeto foi escolhido com base no fato deste dataset ser bastante usado para testar redes neurais com imagens coloridas, e pelo fato de ter uma divisão bastante equilibrada dos dados. [1].

Sua função é verificar a qual das classes pertence uma imagem de tamanho 32x32.

III. Base de Dados

O dataset usado durante o projeto é o CIFAR-100. Ele tem 100 classes contendo 600 imagens cada, totalizando 60.000 imagens. Das 600 imagens que cada classe possui 100 delas são separadas para teste e 500 delas para treino. As 100 classes do dataset estão agrupadas em 20 super classes do seguinte modo [1]:

- aquatic mammals: beaver, dolphin, otter, seal, whale
- fish: aquarium fish, flatfish, ray, shark, trout
- flowers: orchids, poppies, roses, sunflowers, tulips
- food containers: bottles, bowls, cans, cups, plates
- fruit and vegetables: apples, mushrooms, oranges, pears, sweet peppers
- household electrical devices: clock, computer keyboard, lamp, telephone, television
- household furniture: bed, chair, couch, table, wardrobe
- insects: bee, beetle, butterfly, caterpillar, cockroach
- large carnivores: bear, leopard, lion, tiger, wolf
- large man-made outdoor things: bridge, castle, house, road, skyscraper
- large natural outdoor scenes: cloud, forest, mountain, plain, sea

- large omnivores and herbivores: camel, cattle, chimpanzee, elephant, kangaroo
- medium-sized mammals: fox, porcupine, possum, raccoon, skunk
- non-insect invertebrates: crab, lobster, snail, spider, worm
- people: baby, boy, girl, man, woman
- reptiles: crocodile, dinosaur, lizard, snake, turtle
- small mammals: hamster, mouse, rabbit, shrew, squirrel
- trees: maple, oak, palm, pine, willow
- vehicles 1: bicycle, bus, motorcycle, pickup truck, train
- vehicles 2: lawn-mower, rocket, streetcar, tank, tractor

[1]

IV. Análise Exploratória dos Dados

A. Quantidade de Imagens Por Classes

Tanto o dataset de teste quanto o de treino são bastante equilibrados, possuindo a mesma quantidade de imagens para cada classe:

B. Descrição Estatística dos dados

Porque os dados trabalhados são imagens em vez de tabelas, é mais difícil descrevê-las estatisticamente. Portanto as descrevemos do seguinte modo:

1) Média das Imagens por Classe: Um método possível de analisar as imagens estatisticamente é calcular o valor médio por pixel, e deste modo conseguir uma imagem média que representaria a classe.

- Média da Imagens Não Normalizadas

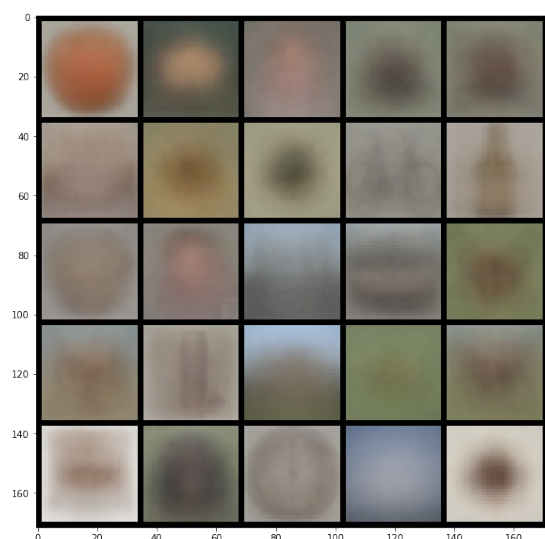


Fig. 1: Da esquerda para a direita, de cima para a baixo: apple, aquarium_fish, baby, bear, beaver, bed, bee, beetle, bicycle, bottle, bowl, boy, bridge, bus, butterfly, camel, can, castle, caterpillar, cattle, chair, chimpanzee, clock, cloud, cockroach

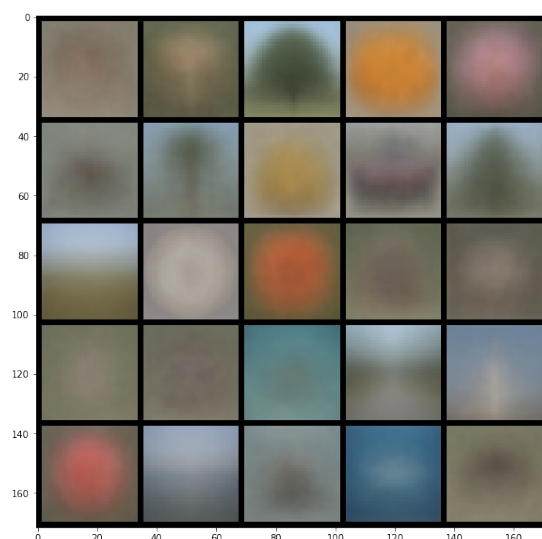


Fig. 3: Da esquerda para a direita, de cima para a baixo: mouse, mushroom, oak_tree, orange, orchid, otter, palm_tree, pear, pickup_truck, pine_tree, plain, plate, poppy, porcupine, possum, rabbit, raccoon, ray, road, rocket, rose, sea, seal, shark, shrew

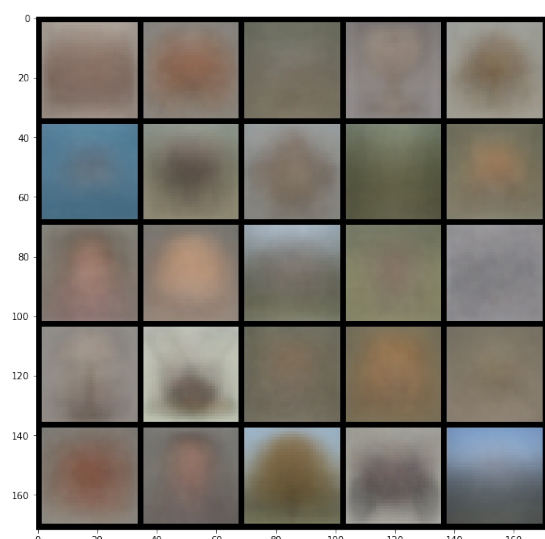


Fig. 2: Da esquerda para a direita, de cima para a baixo: couch, crab, crocodile, cup, dinosaur, dolphin, elephant, flatfish, forest, fox, girl, hamster, house, kangaroo, keyboard, lamp, lawn_mower, leopard, lion, lizard, lobster, man, maple_tree, motorcycle, mountain

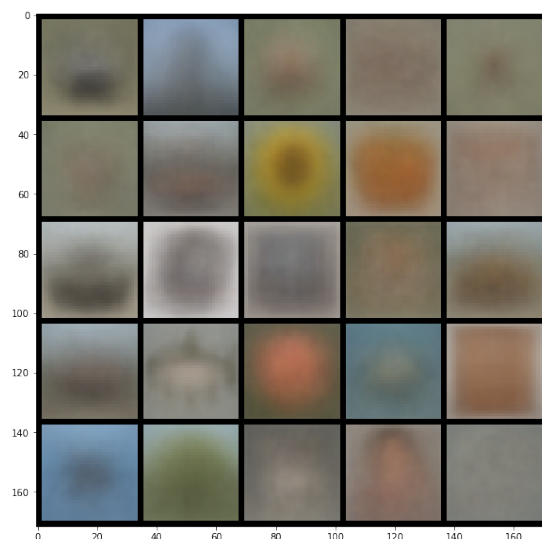


Fig. 4: Da esquerda para a direita, de cima para a baixo: skunk, skyscraper, snail, snake, spider, squirrel, streetcar, sunflower, sweet_pepper, table, tank, telephone, television, tiger, tractor, train, trout, tulip, turtle, wardrobe, whale, willow_tree, wolf, woman, worm

- Média da Imagens Normalizadas

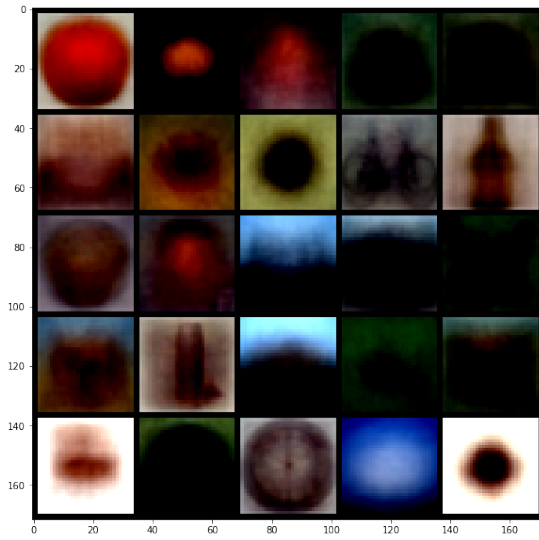


Fig. 5: Da esquerda para a direita, de cima para a baixo: apple, aquarium_fish, baby, bear, beaver, bed, bee, beetle, bicycle, bottle, bowl, boy, bridge, bus, butterfly, camel, can, castle, caterpillar, cattle, chair, chimpanzee, clock, cloud, cockroach

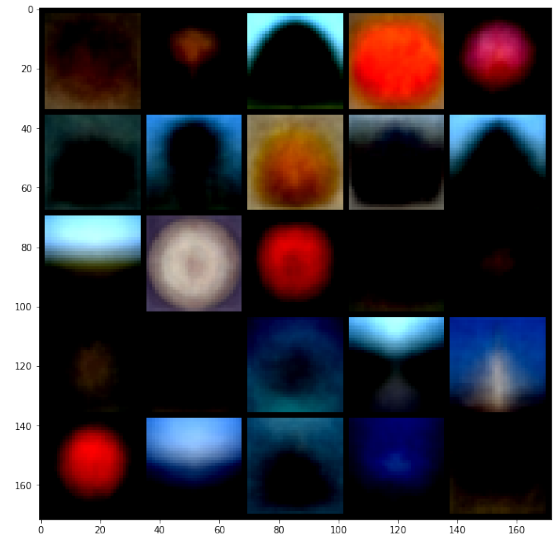


Fig. 7: Da esquerda para a direita, de cima para a baixo: mouse, mushroom, oak_tree, orange, orchid, otter, palm_tree, pear, pickup_truck, pine_tree, plain, plate, poppy, porcupine, possum, rabbit, raccoon, ray, road, rocket, rose, sea, seal, shark, shrew

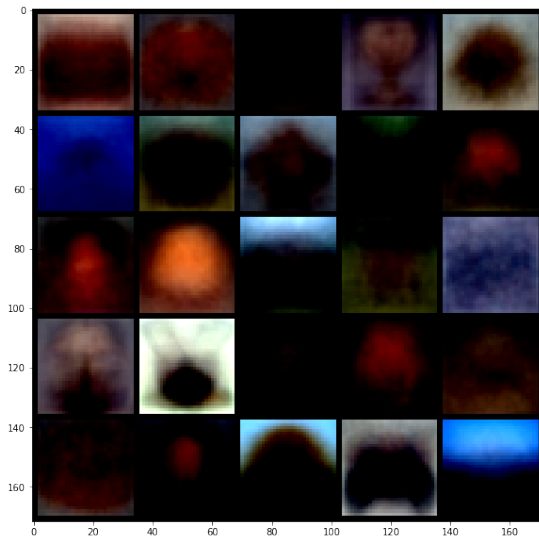


Fig. 6: Da esquerda para a direita, de cima para a baixo: couch, crab, crocodile, cup, dinosaur, dolphin, elephant, flatfish, forest, fox, girl, hamster, house, kangaroo, keyboard, lamp, lawn_mower, leopard, lion, lizard, lobster, man, maple_tree, motorcycle, mountain

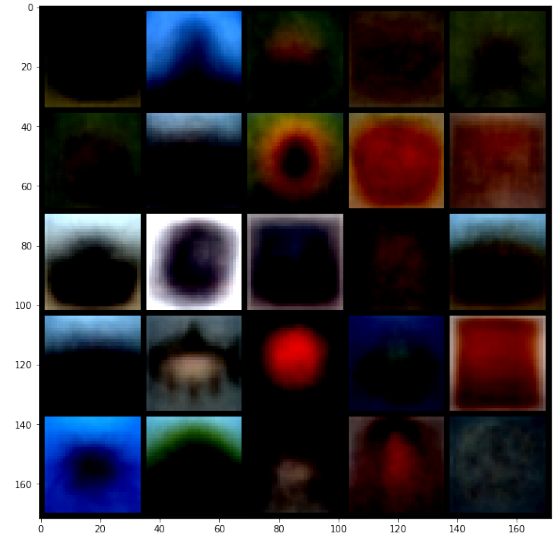


Fig. 8: Da esquerda para a direita, de cima para a baixo: skunk, skyscraper, snail, snake, spider, squirrel, streetcar, sunflower, sweet_pepper, table, tank, telephone, television, tiger, tractor, train, trout, tulip, turtle, wardrobe, whale, willow_tree, wolf, woman, worm

V. Sobre as Métricas Utilizadas

A. Precision

Precision é a razão

$$\frac{A_c}{A_c + A_e} \quad (1)$$

onde:

- A_c é o número de amostras corretamente classificadas de uma determinada classe.
- A_e é o número de amostras erroneamente classificadas como sendo desta determinada classe.

Precision é intuitivamente a habilidade do classificador não marcar como pertencente a uma classe uma amostra que não pertence a esta. O melhor valor de Precision é 1 e o pior é zero. [2]

B. Accuracy

Accuracy é a fração de amostras preditas corretamente, e é dada pela seguinte fórmula:

$$\frac{\sum_1^n A_c}{\sum_1^n A_t} \quad (2)$$

onde:

- n é o número de classes
- A_c é o número de amostras corretamente classificadas de uma determinada classe.
- A_t é o número de amostras que pertencem a uma determinada classe

[3]

C. Recall-Score

O Recall Score é a razão:

$$\frac{A_c}{A_t} \quad (3)$$

onde:

- A_c é o número de amostras classificadas corretamente de uma determinada classe
- A_t é o número de amostras que pertencem a esta classe

O Recall Score é intuitivamente a habilidade do classificador de encontrar todas as amostras pertencentes a uma classe específica. O melhor valor do Recall Score é 1 e o pior valor é 0. [4]

D. F1-Score

O F1 Score pode ser interpretado como a média ponderada da precisão e recall. O melhor valor que o F1 score pode alcançar é 1, o pior é 0. A contribuição relativa da precisão e recall para o F1 score são iguais. A fórmula para o F1 score é:

$$F1 = \frac{2 \cdot (precision \cdot recall)}{precision + recall} \quad (4)$$

[5]

E. Confusion Matrix

No caso de classificação multiclasse, uma confusion matrix é dividida em NxN categorias (onde N é o número de classes do problema), cada uma apresentando a quantidade de amostras que se encaixam nesta. A diagonal do meio representa a quantidade de amostras classificadas corretamente e as demais seções da matriz demonstram o número de amostras classificados erroneamente, a quais classes eles pertencem e em quais classes eles foram classificados.

[6]

VI. Redes Neurais Convolucionais

A. O Nosso Modelo

B. AlexNet

C. GoogLeNet

VII. Experimentos

A. CIFAR10

B. Similaridade de Pixel

Baseado na abordagem do fastai [7] iremos criar um modelo básico que não utiliza aprendizado de máquina, para ter uma precisão como base para verificar o desempenho dos próximos modelos. Esse método basicamente calcula uma imagem média para cada classe do conjunto de treino, esta imagem média é basicamente uma imagem formada pela média de cada pixel das imagens de uma classe.

Para fazer a predição esta arquitetura basicamente calcula a distância de uma imagem a imagem média, e escolhe a classe com a menor distância. Podemos usar o erro quadrático médio ou o valor absoluto das diferenças como valor da distância total entre uma imagem e a média da sua classe.

Porém, ao contrário da abordagem original do fastai para o dataset mnist com apenas duas classes com imagens em preto e branco, que teve uma 'accuracy' razoável com este método, temos que a 'accuracy' que tivemos com o CIFAR 100 que possui muito mais classes, e imagens coloridas, não é melhor do que selecionarmos uma classe ao acaso.

C. Treinando os Modelos

D. Apenas um Epoch

E. Imagens em Grayscale

F. Limitando os Dados

G. Sem Normalização e sem Data Augmentation

VIII. Análise dos Resultados

IX. Conclusões e Discussões

References

- [1] A. Krizhevsky. (2009) Cifar-10 and cifar-100 datasets. [Online]. Available: <https://www.cs.toronto.edu/~kriz/cifar.html>
- [2] "sklearn.metrics.precision_score." [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.precision_score.html

- [3] A. Long, “Understanding data science classification metrics in scikit-learn in python,” Aug. 2018. [Online]. Available: <https://towardsdatascience.com/understanding-data-science-classification-metrics-in-scikit-learn-in-python-3bc336865019>
- [4] “sklearn.metrics.recall_score.” [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.recall_score.html
- [5] sklearn.metrics.f1_score. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html
- [6] B. Shmueli, “Multi-class metrics made simple, part i: Precision and recall,” Jul. 2019. [Online]. Available: <https://towardsdatascience.com/multi-class-metrics-made-simple-part-i-precision-and-recall-9250280bddc2>
- [7] J. Howard and S. Gugger, Deep Learning for Coders with Fastai and Pytorch: AI Applications Without a PhD. O’Reilly Media, Incorporated, 2020. [Online]. Available: <https://books.google.no/books?id=xd6LxgEACAAJ>
- [8] A. Krizhevsky, “Learning multiple layers of features from tiny images,” University of Toronto, Toronto, Ontario, Tech. Rep. 0, 2009.
- [9] Muthu, “Understanding the classification report through sklearn,” Jul. 2018. [Online]. Available: <https://muthu.co/understanding-the-classification-report-in-sklearn/>