

Classificação de imagens utilizando Tensorflow

Tiago Vinicius Lopes Pereira ¹
Victor Hugo Cardoso Mendes ¹

Resumo: Deep learning é uma das bases da inteligência artificial (IA), e o interesse atual em deep learning se deve, em parte, à fascinação por IA. Técnicas de deep learning têm aprimorado a capacidade dos computadores em classificar, reconhecer, detectar e descrever. Este artigo tem como objetivo solucionar problemas de classificação de imagens utilizando o método de deep learning e apresenta o resultado da análise e classificação de imagens de três datasets (fashion MNIST, MNIST Dígitos a mão e Raios-x do peito para pneumonia) com o framework Tensorflow. Os testes foram rodados utilizando uma GPU GTX 660ti

Palavras-chave: Aprendizado profundo. Classificação de imagens. Fashion MNIST. MNIST dígitos manuscritos. Tensorflow.

Abstract: *Terminar o Resumo, e depois traduzir pro inglês*

Keywords: *Deep learning. Fashion MNIST. Image classification. MNIST handwritten digits. Tensorflow.*

1 Introdução

O aprendizado de máquina (machine learning) é um ramo da inteligência artificial baseado na ideia de que sistemas podem aprender com dados, identificar padrões e tomar decisões com o mínimo de intervenção humana.

Aprendizado profundo (deep learning) é um tipo de machine learning que treina computadores para realizar tarefas como seres humanos, o que inclui reconhecimento de fala, identificação de imagem e previsões. Em vez de organizar os dados para serem executados através de equações predefinidas, o deep learning configura parâmetros básicos sobre os dados e treina o computador para aprender sozinho através do reconhecimento de padrões em várias camadas de processamento. [1]

Este artigo surgiu como produto da disciplina optativa Deep Learning para Classificação de Imagem em GPU, do curso de graduação em ciência da computação pela Universidade Estadual do Oeste do Paraná (UNIOESTE).

O objetivo deste artigo é pesquisar o funcionamento dos algoritmos de machine learning com foco em deep learning, e utilizando o framework Tensorflow para a solução de problemas, tornando os alunos capazes de detectar áreas de aplicação de tecnologias baseadas em deep learning para solução de problemas com aceleração de hardware. Foram escolhidos três datasets, MNIST dígitos manuscritos, Fashion MNIST e Chest X-Ray Images.

Os gráficos apresentados são dos logs automaticamente gerados pelo tensorflow e plotados pelo tensorboard (conjunto de ferramentas de visualização embutido na instalação do tensorflow) que facilitam a visualização, debug, e otimização de computações feitas no tensorflow. São apresentados gráficos de acurácia (exatidão mais precisão), de perdas (busca-se o menor número de perdas) e um gráfico de fluxo de dados para representar seu cálculo em termos das dependências entre operações individuais. As épocas (iteração sobre todos os dados de treinamento) podem variar em cada dataset.

¹ Alunos de graduação, UNIOESTE, Campus Foz do Iguaçu (PTI) - Av. Tancredo Neves, 6.731 85856-970 Foz do Iguaçu, Paraná, Brasil

2 Metodologia

O framework escolhido pela dupla foi o Tensorflow devido a sua popularidade e possibilidade de desenvolvimento em Python, ser flexível e poderoso, apesar de ter uma longa curva de aprendizado, é atualmente o principal software para desenvolvimento em deep learning e aplicações de inteligência artificial, com a maior e mais ativa comunidade. Os três datasets foram escolhidos pela turma dentre uma lista de seis para serem utilizados igualmente entre todos, possibilitando a comparação de resultados de cada dataset em diferentes frameworks para solução de problemas com classificação de imagens. O terceiro dataset foi rodado numa maquina com a seguinte especificação: CPU Intel Core i7-2600 @ 3.40GHz x 8, 8GB de memória RAM e uma GPU GeForce GTX 1060 3GB. O sistema operacional utilizado foi o Linux, com a distro Ubuntu 18.04.2 LTS 64bits. As versões do CUDA e do tensorflow são respectivamente, 9.0 e 1.12.0. Apesar de o tensorflow rodar em python2 e python3, neste artigo utilizamos o python3 por ser o mais atualizado.

2.1 Tensorflow

Tem a capacidade de rapidamente gerar um produto ou serviço a partir do modelo preditivo treinado, sem a necessidade de reimplementar o modelo. Computadores de qualquer natureza podem executar o TensorFlow, desde smartphones a gigantescos clusters de computadores. Oferece a possibilidade de desenvolvimento em Python e C++. O TensorFlow também suporta um ecossistema de poderosas bibliotecas complementares e modelos para experimentar. [2]

2.2 MNIST handwritten digits

O banco de dados MNIST de dígitos manuscritos(MNIST handwritten digits), possui um conjunto de treinamento de 60.000 exemplos e um conjunto de teste de 10.000 exemplos. É um subconjunto de um conjunto maior disponível no NIST. Os dígitos foram normalizados por tamanho e centralizados em uma imagem de tamanho fixo. É um bom banco de dados para pessoas que querem experimentar técnicas de aprendizado e métodos de reconhecimento de padrões em dados do mundo real, enquanto gastam esforços mínimos em pré-processamento e formatação. Normalmente utilizado como "hello world" para algoritmos de classificação de imagens. [3]

Figura 1: Exemplo MNIST Handwritten Digits



2.3 Fashion-MNIST

O Fashion-MNIST é um conjunto de dados de imagens consistindo em um conjunto de treinamento de 60.000 exemplos e um conjunto de teste de 10.000 exemplos. Cada exemplo é uma imagem em escala de cinza 28x28, associada a um rótulo de 10 classes. O Fashion-MNIST serve como um substituto direto do conjunto

de dados MNIST original para algoritmos de benchmarking de aprendizado de máquina. O autor considera o MNIST original muito "fácil" e "comum", visto que redes convolucionais atingem 99.7% e algoritmos clássicos de aprendizagem de máquina atingem 97% facilmente. [4]

Figura 2: Exemplo Fashion-MNIST



2.4 Chest X-Ray Images

Neste artigo usamos o dataset de tomografias do peitoral para identificação de pneumonia em pacientes. Todas as imagens não legíveis foram removidas da coleção e então foram classificadas por dois especialistas antes de serem disponibilizadas para treinamento de inteligência artificial. O conjunto de dados é organizado para cada categoria de imagem (Pneumonia / Normal). Existem 5.863 imagens de raios X (JPEG) e 2 categorias, raios X de pacientes diagnosticados com pneumonia, e raio X de pacientes saudáveis. [5]

Figura 3: Exemplo de tomografias do peitoral



As imagens deste dataset são de tamanhos diferentes, para utiliza-lo, as imagens precisaram ser redimensionadas para um tamanho padrão de 256px de largura e 256px de altura, onde mais que isso excedia a memória de GPU da máquina pela grande quantidade de dados para serem processados por vez.

2.5 TFLearn

Neste artigo utilizamos a biblioteca TFLearn para o treinamento dos modelos em cada dataset, e para a geração de gráficos sobre perda e acurácia em função do tempo.

O TFLearn é uma biblioteca de deep learning modular e transparente construída sobre o Tensorflow. Ele foi projetado para fornecer uma API de nível superior ao TensorFlow para facilitar e agilizar as experimentações, mantendo-se totalmente transparente e compatível com ele. Todas as funções do TFLearn foram construídas sobre tensores e podem ser utilizadas independentemente do TFLearn [6]

3 Resultados

3.1 MNIST handwritten digits

Os seguintes gráficos foram gerados pela biblioteca TFLearn com base nos logs gerados durante o treinamento do modelo, para chegar nesse resultado utilizamos 4 épocas e um tamanho de batch de 128. O treinamento foi feito pelo pacote Deep Neural Network do próprio TFLearn que performa tarefas básicas de classificação utilizando redes neurais.

Figura 4: Acurácia em função do tempo para dígitos numéricos

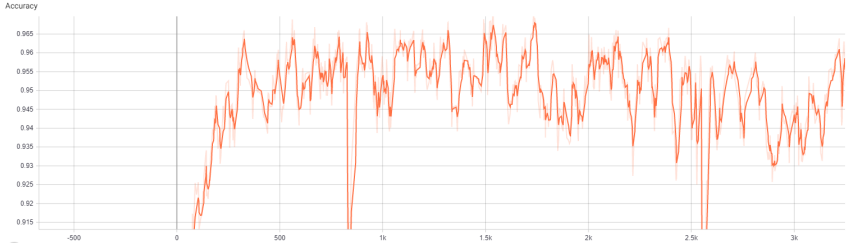
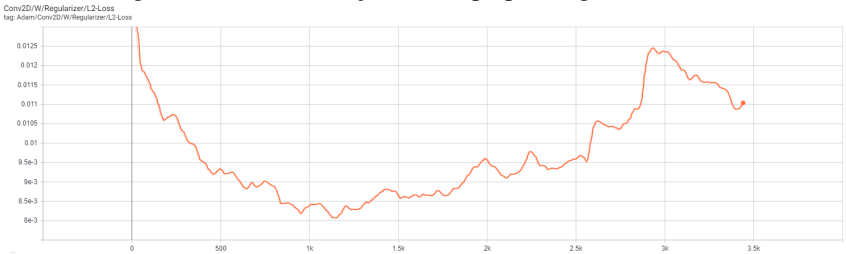


Figura 5: Perda em função do tempo para dígitos numéricos



Com o modelo treinado, consideramos sete imagens de teste, sendo respectivamente os números: 7, 2, 1, 0, 4, 2, 4.

Tabela 1: Classificação

Imagem	0 (%)	1 (%)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	9 (%)
(7)	0.0001	0.0001	0.0001	0.0003	0.0003	0.0001	0.0001	99.9904	0.0002	0.0089
(2)	0.0005	16.6347	99.9805	0.0002	0.0002	0.0004	0.0002	0.0007	2.7425	0.0008
(1)	0.0006	99.9998	0.0001	0.0007	0.0008	0.0002	0.0001	0.0008	0.0002	0.0001
(0)	99.9950	0.0002	0.0002	0.0005	0.0002	0.0007	0.0001	0.0007	0.0002	0.0044
(4)	0.0019	0.0003	0.0002	0.0002	99.9990	0.0001	0.0006	0.0008	0.0005	0.0009
(2)	0.0005	99.9995	0.0001	0.0001	0.0060	0.0002	0.0007	0.0038	0.0002	0.0004
(4)	0.0003	0.0023	0.0002	0.0009	99.7840	0.0017	0.0012	0.0013	0.1291	8.2812

3.2 Fashion-MNIST

Os seguintes gráficos foram gerados pela biblioteca TFLearn com base nos logs gerados durante o treinamento do modelo, para chegar nesse resultado utilizamos 10 épocas e um tamanho de batch de 128. O treinamento foi feito pelo pacote Deep Neural Network do próprio TFLearn que performa tarefas básicas de classificação utilizando redes neurais.

Figura 6: Acurácia em função do tempo para imagens de roupas

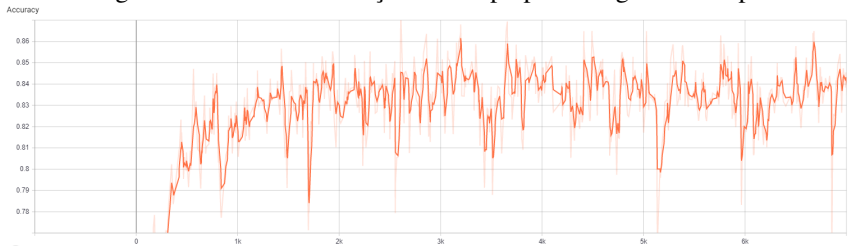
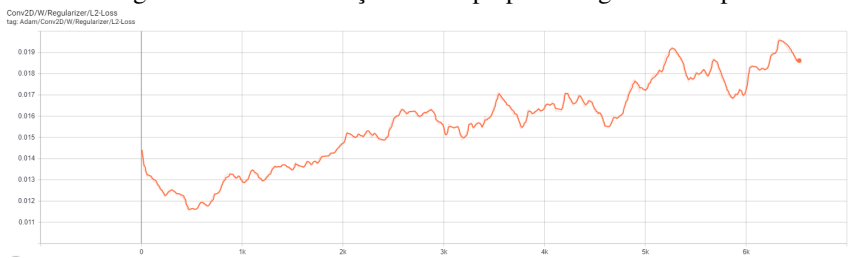


Figura 7: Perda em função do tempo para imagens de roupas



Com o modelo treinado, consideramos sete imagens de teste, sendo respectivamente as peças de roupa: Bota, Agasalho, Calça, Calça, Camisa, Calça e Casaco.

Tabela 2: Legenda para as classificações

Classe	Número
Camiseta	(0)
Calça	(1)
Agasalho	(2)
Suéter	(3)
Casaco	(4)
Sandália	(5)
Camisa	(6)
Tênis	(7)
Bolsa	(8)
Bota	(9)

Tabela 3: Classificação

Imagem	(0) (%)	(1) (%)	(2) (%)	(3) (%)	(4) (%)	(5) (%)	(6) (%)	(7) (%)	(8) (%)	(9) (%)
(9)	0.0003	0.0009	0.0001	0.0001	0.0001	0.0002	0.0002	0.0529	0.0006	99.9467
(2)	0.0382	0.0009	99.8950	0.0180	0.0301	0.0004	0.0148	0.0014	0.0008	0.0001
(1)	0.0004	99.9953	0.0008	0.0024	0.0014	0.0005	0.0002	0.0001	0.0003	0.0004
(1)	0.0029	99.9950	0.0006	0.0012	0.0005	0.0003	0.0007	0.0002	0.0004	0.0003
(6)	23.9528	0.0014	23.3576	11.7867	1.0761	0.0008	39.6242	0.0151	0.1822	0.0046
(1)	0.0989	99.4237	0.1294	0.0285	0.3107	0.0001	0.0014	0.0003	0.0066	0.0006
(4)	2.1062	0.2311	43.2329	1.7401	51.9477	0.0001	0.5625	0.0480	0.1308	0.0007

3.3 Chest X-Ray Images

Os seguintes gráficos foram gerados pela biblioteca TFLearn com base nos logs gerados durante o treinamento do modelo, para chegar nesse resultado utilizamos 6 épocas e um tamanho de batch de 128. O treinamento foi feito pelo pacote Deep Neural Network do próprio TFLearn que performa tarefas básicas de classificação utilizando redes neurais.

Figura 8: Acurácia em função do tempo para tomografias do peitoral

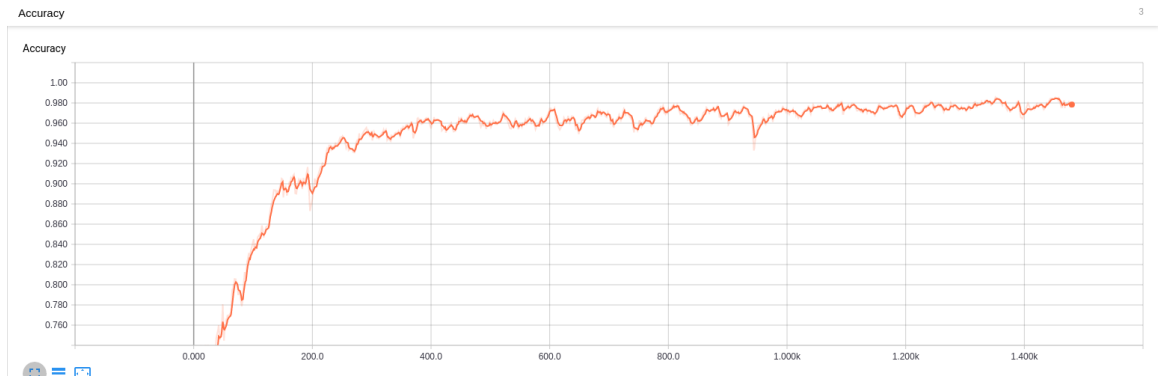
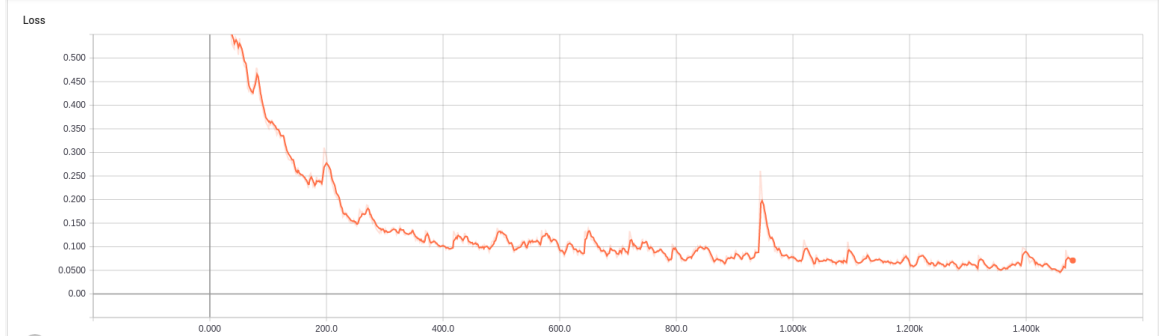


Figura 9: Perda em função do tempo para tomografias do peitoral



Com o modelo treinado, consideramos sete imagens de teste, sendo elas: 4 classificadas com pneumonia(1) e 3 classificadas como normais(0).

Tabela 4: Classificação

Imagem	Normal (%)	Pneumonia (%)
(0) Imagem1	12.5274	87.4726
(0) Imagem2	0.0968	99.9032
(0) Imagem3	0.6776	99.3224
(1) Imagem99	0.006	99.9940
(1) Imagem1	0.0084	99.9916
(1) Imagem3	0.0021	99.9979
(1) Imagem4	0.0084	99.9916

Figura 10: Modelo de matriz de confusão

	NORMAL	PNEUMONIA
NORMAL	VERDADEIRO POSITIVO	FALSO NEGATIVO
PNEUMONIA	FALSO POSITIVO	VERDADEIRO NEGATIVO

Figura 11: Matriz de confusão pneumonia

		Classes Saída	
		NORMAL	PNEUMONIA
Classes Entrada	NORMAL	0	7
	PNEUMONIA	0	7

4 Discussão

5 Conclusão

Agradecimentos

O presente trabalho foi realizado com apoio do docente de graduação em ciência da computação da Universidade Estadual do Oeste do Paraná, Cláudio Roberto Marquetto Maurício, e dos alunos da disciplina optativa Deep Learnig para Classificação de Imagem em GPU (2019).

Referências

[1] SAS. *Deep Learning O que é e qual sua importância?* Disponível em: <https://www.sas.com/pt_br/insights/analytics/deep-learning.html>. Acesso em: 07 mai. 2019.

[2] DATASCIENCEACADEMY. *O QUE É O TENSORFLOW MACHINE INTELLIGENCE PLATFORM?* Disponível em: <<http://datascienceacademy.com.br/blog/o-que-e-o-tensorflow-machine-intelligence-platform/>>. Acesso em: 07 mai. 2019.

[3] LECUN CORINNA CORTES, C. J. B. Y. *THE MNIST DATABASE of handwritten digits*. Disponível em: <<http://yann.lecun.com/exdb/mnist/index.html>>. Acesso em: 07 mai. 2019.

[4] XIAO, H.; RASUL, K.; VOLLGRAF, R. *Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms*. 2017.

[5] KERMANY, D. S. et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, v. 172, n. 5, p. 1122–1131.e9, feb 2018. ISSN 00928674. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0092867418301545>>.

[6] DAMIEN, A. *TFLearn: Deep learning library featuring a higher-level API for TensorFlow*. Disponível em: <<http://tflearn.org/>>. Acesso em: 21 jun. 2019.

Figura 12: Fluxo de dados para tomografias do peitoral

