

Sistemas de Visão e Percepção Industrial

6-Complementos sobre Percepção e Representação

Exercícios de Introdução ao Deep Learning

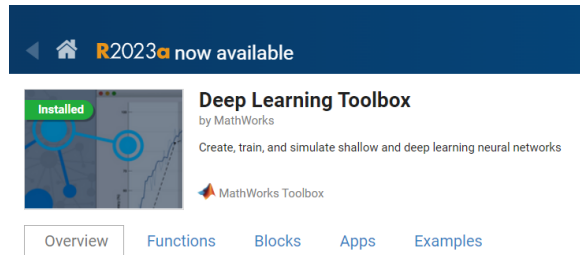
Universidade de Aveiro

Joel Baptista, Vitor Santos, 2022-2023

- 1 Preparação inicial
- 2 Treino da rede neuronal
- 3 Visualização de métricas e filtros

Preparação inicial

- Instalar a Deep Learning Toolbox.
- Esta ferramenta permite:
 - Criar redes neurais.
 - Treinar redes neurais.
 - Utilizar ferramentas de visualização e análise de redes neurais.
- Assume-se que já estejam instalado as outras toolboxes utilizadas nas aulas anteriores.



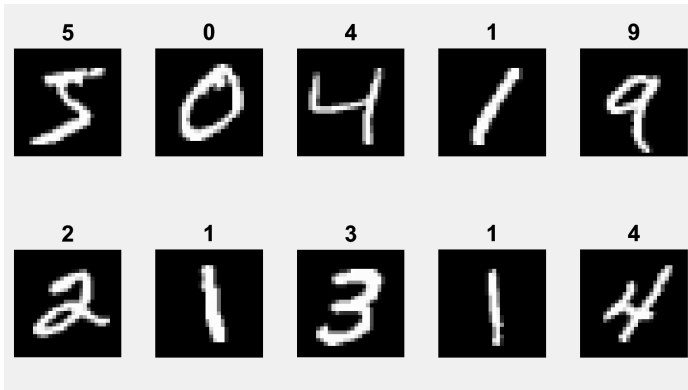
Dataset MNIST

- O dataset MNIST(Modified National Institute of Standards and Technology) é composto por imagens dos algarismos de 0-9 escritos de forma manual.
- As imagens são compostas por apenas um canal de cinzentos de 28x28 pixels.
- O dataset é composto por 60 000 imagens de treino e 10 000 imagens de teste.



Ex 1a) Visualizar dados de treino

- Importar os dados de treino `train.mat`
 - Nota: `train.mat` é uma estrutura com dois campos `XTrain` e `YTrain`.
 - `XTrain` é referente às imagens do dataset.
 - `YTrain` é referente às classificações de cada imagem em `XTrain`.
- Mostrar as primeiras 10 imagens juntamente com as suas respetivas "labels".



Ex. 1b) Preparação para o treino

- Usualmente, são utilizados dados de validação para monitorizar o progresso do treino da rede neuronal.
 - Retirar uma percentagem de amostras (recomendado 30 %) da variável XTrain e colocá-la numa variável XVal.
 - As amostras introduzidas na variável XVal terão de ser retiradas da variável XTrain.
- Criar variável que configura as opções de treino
 - Preencher o campo ValidationData com as variáveis corretas onde está ***

```
options = trainingOptions('adam', ...  
    'MaxEpochs',5, ...  
    'InitialLearnRate',1e-4, ...  
    'ValidationData', {***, ***},...  
    'Plots','training-progress','Verbose',true);
```

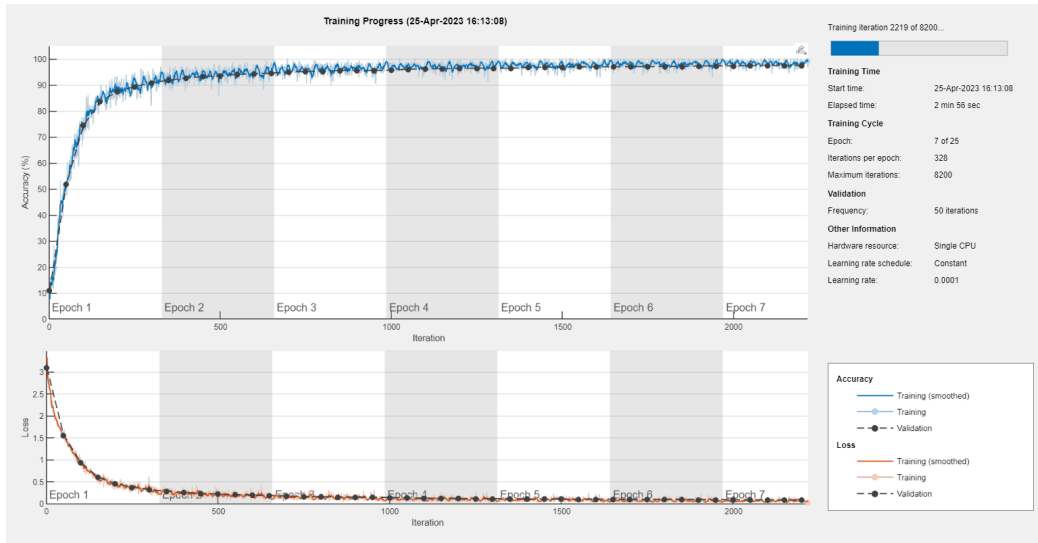
Treino da rede neuronal

Ex. 2) Treinar a rede neuronal convolucional

- A arquitetura da rede é configurada com um "array de layers".
- O fluxo de informação percorre o array, sendo que a imagem entra no `imageInputLayer` e a classificação sai no `classificationLayer`.
- Preencher os espaços assinalados com `***` com números relativos ao problema em questão.
- Utilizar a função `trainNetwork` para treinar a rede neuronal convolucional.

```
layers=[  
    imageInputLayer([ *** ],'Name','input')  
    convolution2dLayer(3,6,'Padding','same')  
    reluLayer  
    batchNormalizationLayer  
    maxPooling2dLayer(2,'Stride',2)  
    convolution2dLayer(3,16)  
    reluLayer  
    batchNormalizationLayer  
    maxPooling2dLayer(2,'Stride',2)  
    fullyConnectedLayer(120,'name','f1')  
    reluLayer  
    fullyConnectedLayer(84,'name','f2')  
    reluLayer  
    fullyConnectedLayer(***,'name','f3')  
    softmaxLayer  
    classificationLayer  
];
```

Ilustração do treino da rede neuronal convolucional



Visualização de métricas e filtros

Ex. 3) Visualização de métricas de treino

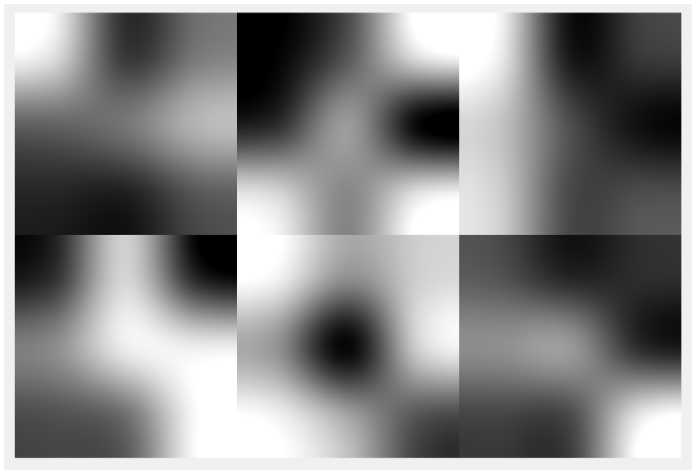
- Importar a estrutura `test.mat` e que tem as imagens para testar o modelo treino.
- Calcular a percentagem de acerto (accuracy) usando as imagens de acerto.
- Questões:
 - Porque estamos a testar em imagens não utilizadas no treino?
 - A percentagem de acerto do teste está perto do treino e da validação?
 - O que significa estarem perto?
 - O que significa estarem longe?

Ex. 4) Visualização de filtros de convolução

- Analisar os filtros de convolução otimizados pela primeira camada convolucional.
- Utilizar a função `montage` para fazer a visualização.
 - Nota: utilizar a função `rescale` antes para obter uma melhor visualização.
- Depois do treino, cada `convolution2dLayer` aprendeu um número fixo de filtros de convolução que otimiza a classificação das imagens do dataset MNIST.
 - Quantos filtros de convolução aparecem? Porquê esse número?
 - Que informação é possível tirar destes filtros?
 - Os filtros obtidos são iguais aos apresentados no slide seguinte? Porquê?

```
% Extract convolution filters  
w1 = net.Layers(2).Weights;
```

Imagem com os filtros de convolução



- Nota: O efeito "blurred" é produzido pela função montage. Na verdade, os filtros são matrizes 3x3.

Ex. 5) Visualização da matriz de confusão

- Analisar os resultados do teste com uma matriz de confusão.
- Utilizar a função `confusionchart` para fazer a visualização.
- A matriz de confusão permite rapidamente adquirir uma ideia geral da performance do modelo de forma visual.
- Esta ferramenta também permite analisar ao detalhe os erros que o modelo produz, especificando quais as classes que causam mais confusão no modelo treinado.
 - O que significam as colunas da matriz?
 - O que significam as linhas da matriz?
 - O que é que significa a diagonal da matriz?

Ilustração da matriz de confusão

0	962		2	1		2	7	1	3	2
1		1121	4			1	3	1	5	
2	7	1	984	8	1		2	13	16	
3			6	976		9		7	8	4
4		1	7	1	941		12	3	6	11
5	2	1	1	5		864	8	1	8	2
6	7	2	2		5	6	932		4	
7		5	18	2	2			983	3	15
8	3		5	5	6	5		4	941	5
9	6	5	3	6	6	7		7	10	959
	0	1	2	3	4	5	6	7	8	9

True Class

Predicted Class