



Universidade Estadual de Londrina  
Departamento de Computação - CCE

# **Trabalho - Computação Gráfica: Classificação de imagens utilizando o dataset MNIST**

Eduardo Henrique Giroto

Londrina, 3 de novembro de 2019



## Sumário

<b>1</b>	<b>Introdução</b>	<b>2</b>
<b>2</b>	<b>Pré-Processamento</b>	<b>2</b>
<b>3</b>	<b>Extração de características</b>	<b>2</b>
3.1	<i>Histogram of Oriented Gradients</i> (HoG) . . . . .	2
<b>4</b>	<b>Modelo de Aprendizado de Máquina</b>	<b>3</b>
<b>5</b>	<b>Métricas de Avaliação</b>	<b>3</b>
5.1	Recall, Precision, Accuracy, F1-Score . . . . .	3
<b>6</b>	<b>Resultados obtidos</b>	<b>3</b>





## 4 Modelo de Aprendizado de Máquina

O modelo utilizado para a classificação de imagens foi uma Rede Neural Multi-Camadas, também chamado de MLP (*Multi-Layer Perceptron*). O modelo foi criado com o auxílio das bibliotecas Keras e TensorFlow. A Rede Neural implementada é organizada em camadas distintas em que cada uma possui um conjunto de neurônios.

A rede neural utilizada neste trabalho contém 4 camadas (*layers*), com cada uma tendo as seguintes características:

- O primeiro layer receberá como input o vetor de características dado pelo HoG. Ou seja, como o vetor possui tamanho igual a 441, nossa primeira camada terá 441 neurônios.
- As camadas 1 e 2 terão arbitrariamente 128 e 64 neurônios.
- Por fim, a última camada, terá a quantidade de neurônios correspondente à quantidade de classes que o nosso problema de classificação possui: 10

Para as camadas ocultas, foi utilizada a função de ativação sigmoid e na última camada foi usada a função softmax. Para treinar o modelo, foi escolhido o algoritmo *Stochastic Gradient Descent* (SGD). A learning rate do nosso algoritmo será igual a 0.01, e a loss function será a `categorical_crossentropy`, pois nosso problema é multi-classe.

## 5 Métricas de Avaliação

### 5.1 Recall, Precision, Accuracy, F1-Score

Em termos simples, *recall* (revocação) diz o quanto o resultado é completo ou não, e *precision* (precisão) indica o quanto os resultados são úteis, uma alta precisão significa que um algoritmo retornou mais resultados relevantes do que irrelevantes.

Acurácia indica uma performance geral do modelo. Dentre todas as classificações, quantas o modelo classificou corretamente. F1-Score representa a média harmonica entre precision e recall.

## 6 Resultados obtidos

Durante a fase de treinamento, o modelo foi ajustado com diferentes valores de época e lotes para a análise do melhor desempenho de nossa Rede Neural.

- Época: Uma passagem por todas as linhas no conjunto de dados de treinamento.
- Lote: Uma ou mais amostras consideradas pelo modelo em uma época antes da atualização dos pesos.

Ao final, dos experimentos, escolhemos 200 para o número de épocas e 10 lotes. Isso significa que cada época envolverá (200/10) 20 atualizações nos pesos do modelo. A tabela 1 contém a matriz de confusão com os valores das métricas de avaliação e a figura 2 mostra o gráfico com as curvas referentes aos datasets de treino e validação. Quanto mais próximas, maior o sinal de que não houve overfitting durante a fase de treinamento.



Classe	Precision	Recall	F1-Score
0	100%	100%	100%
1	83%	100%	91%
2	92%	100%	96%
3	78%	100%	88%
4	100%	82%	90%
5	100%	100%	100%
6	92%	92%	92%
7	83%	100%	91%
8	100%	64%	78%
9	100%	100%	100%

Tabela 1: Matriz de Confusão

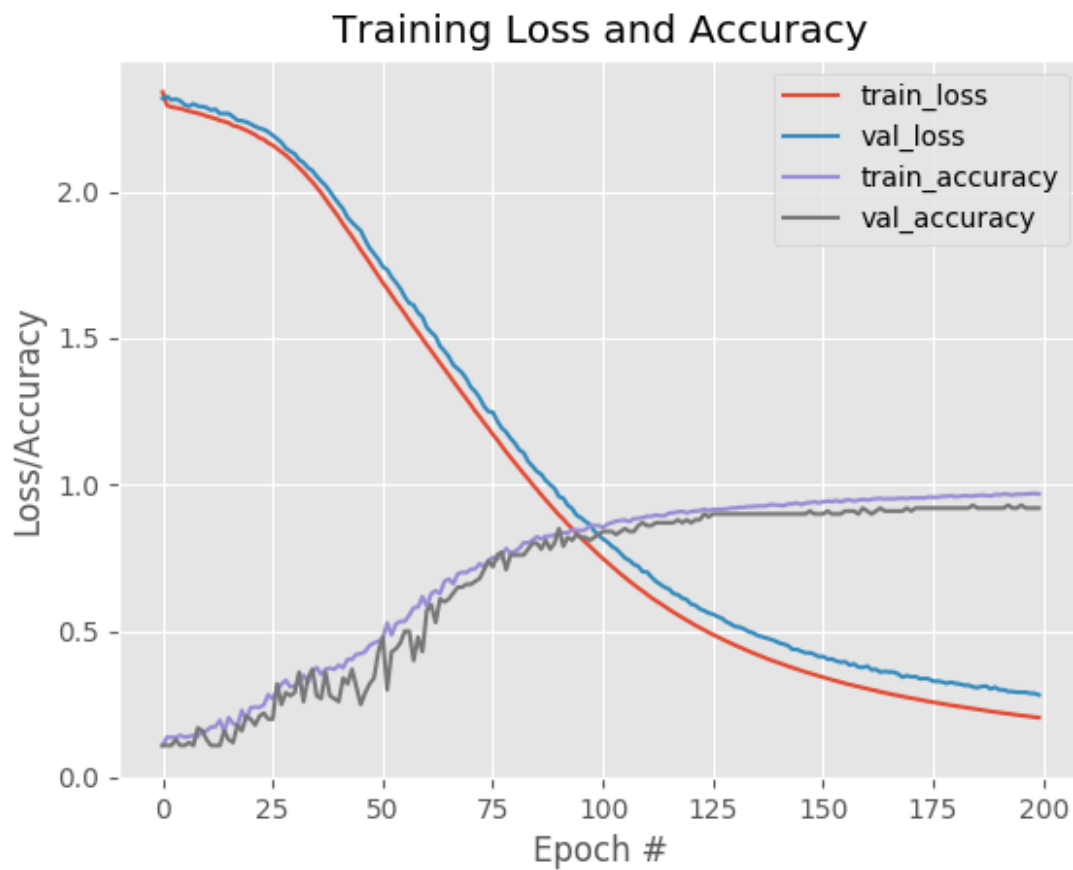


Figura 2: Curvas dos datasets de treino e validação

Ao final do treinamento e com a validação, a rede conseguiu atingir uma acurácia de 93%.