

Estudo Extensivo de Arquiteturas Neurais para Classificação de Sinais de Trânsito

PSI5886 - Princípios de Neurocomputação
Prof. Dr. Emílio Del Moral Hernandez

Membros:

Alexander Csóka Roque

Aléxis Kenji Uenojo

Fernando Luis Gutiérrez López

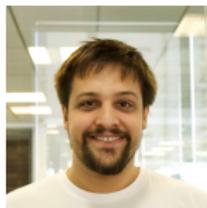
Leonardo Tessarolo Silva

14 de Dezembro de 2018

CONTEÚDO

Introdução
Metodologia
Resultados
Conclusão
Referências

INTRODUÇÃO - EQUIPE



Alexander Csóka
Roque



Aléxis Kenji
Uenojo



Fernando Luis
Gutiérrez López



Leonardo
Tessarolo Silva

DIVISÃO DAS TAREFAS

As tarefas necessárias para o desenvolvimento do projeto foram divididas de modo que, em boa parte delas, todos tiveram alguma participação.

A discussão e especificação das arquiteturas a serem estudadas foram definidas por todos os integrantes, bem como a pesquisa bibliográfica e análise dos resultados.

A implementação e ensaio computacional das redes foi responsabilidade do Alexander.

A elaboração e formatação da documentação do projeto foi responsabilidade do Aléxis, Fernando e Leonardo.

MOTIVAÇÃO

- ▶ Reconhecimento de Sinais de Trânsito (TSR) é essencial em sistemas inteligentes de auxílio ao motorista e piloto automático;
- ▶ Rápido desenvolvimento de veículos autônomos é assunto de muitas pesquisas e interesse de muitas empresas;



Figura 1: *Traffic Sign Recognition System* da Ford. Fonte: [For].

OBJETIVOS

O projeto tem como objetivo:

- ▶ aplicar e consolidar os conhecimentos obtidos na disciplina em um problema real;
- ▶ observar e analisar como a mudança de parâmetros pode influenciar nos resultados de uma RNA;
- ▶ aprofundamento nas técnicas de classificação de imagens por RNAs.

METODOLOGIA - FERRAMENTAS

- ▶ O projeto foi desenvolvido em Python 3.6 com o *framework Tensorflow 1.12.0* com GPU;
- ▶ **Base de dados utilizada:** BelgiumTS Dataset [Tim14].



Figura 2: Amostra do Dataset BelgiumTS. Fonte: [Abd16]

ARQUITETURAS I

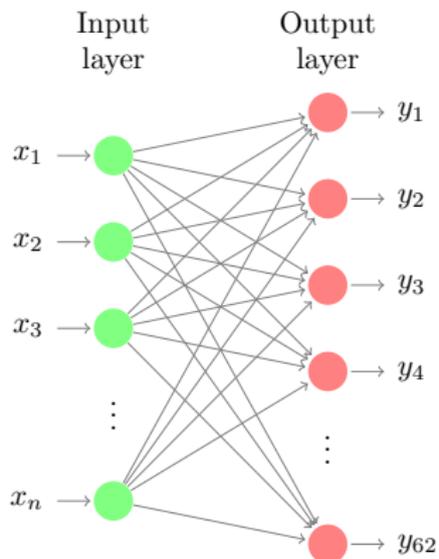
A proposta é realizar diversos ensaios utilizando imagens com dimensões $n = 64 \times 64 \times 3$, com as seguintes arquiteturas:

- ▶ Redes Totalmente Conectadas: camadas escondidas com 62 neurônios variando os seguintes parâmetros
 - ▶ Funções de Ativação;
 - ▶ Taxa de Aprendizado;
 - ▶ Número de camadas totalmente conectadas.
- ▶ Redes Convolucionais: Variação no número de camadas convolucionais.

Cada parâmetro é ensaiado individualmente, ou seja, não é feita a análise de todas as combinações possíveis entre os diversos parâmetros.

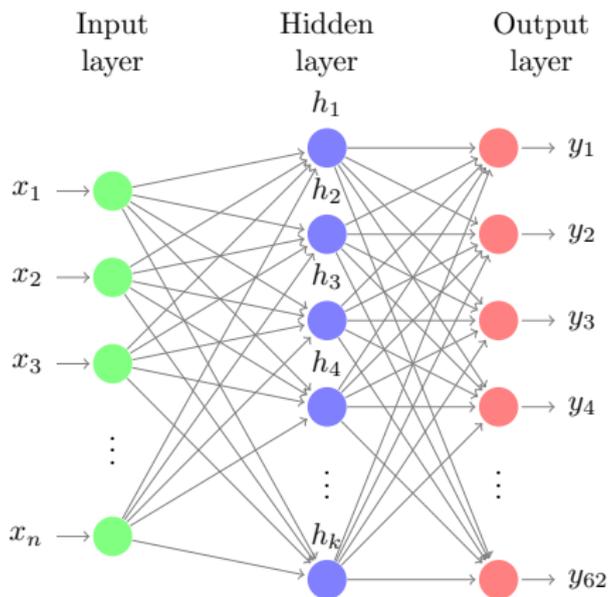
ARQUITETURAS II

- ▶ **Funções de Ativação (taxa de aprendizado fixa em 0.05):** Sigmoid, Tanh e ReLU;
- ▶ **Taxa de Aprendizado (todos utilizando tanh):** 0.0001, 0.0005, 0.001, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1, 0.2, 0.3, 0.4, 0.5, 1.0, 5.0 e 10.0;



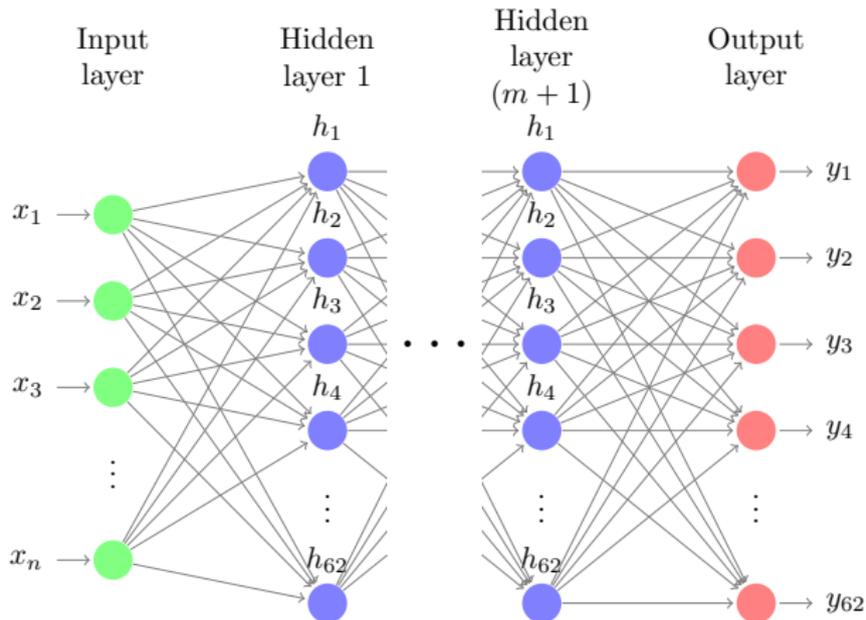
ARQUITETURAS III

- ▶ **Número de Neurônios por camada (k):** 8, 16, 32, 62, 124, 250, 500, 750, 1000, 1500, 2000, 2500, 5000, 7500, 10000, 15000, 20000 e 30000;
- ▶ Foi utilizada tanh e taxa de aprendizado 0.05 em todos os ensaios.



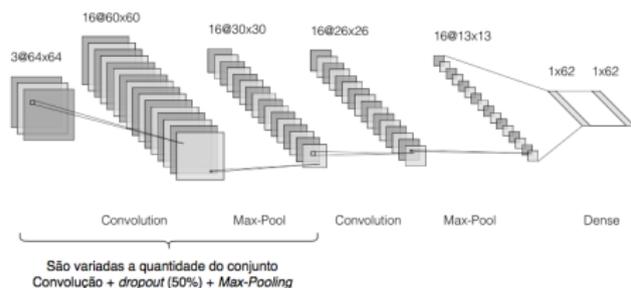
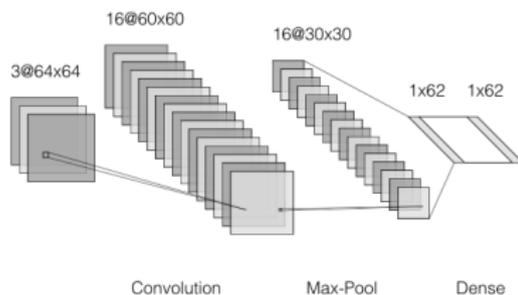
ARQUITETURAS IV

- ▶ **Número de Camadas Escondidas (m):** 1, 2, 3, 4, 5, 7, 10, 15, 20, 25, 30, 40 e 50;
- ▶ Foi utilizada tanh, taxa de aprendizado 0.05 e 62 neurônios em cada camada escondida em todos os ensaios.



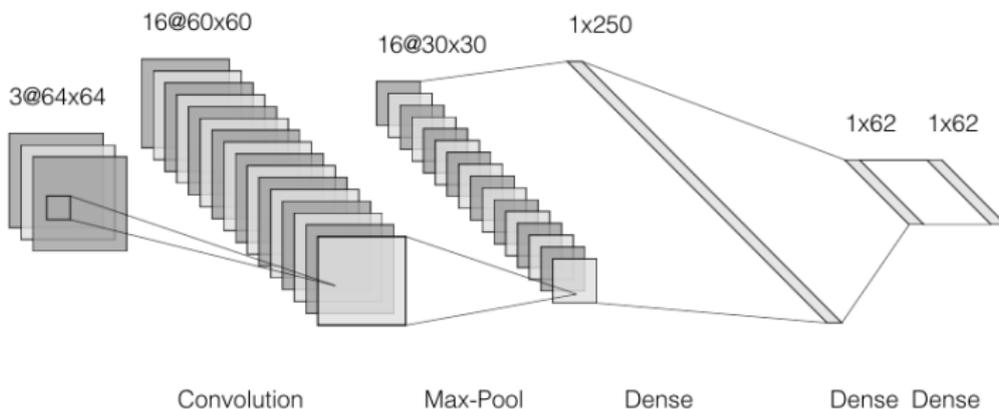
ARQUITETURAS IV

- ▶ **Número de Camadas Convolucionais:** 1, 2, 3 e 4.
- ▶ Todas as redes foram ensaiadas com função de ativação tanh, taxa de aprendizado de 0.05, 16 filtros convolucionais de tamanho 5x5 seguidos de uma camada de *dropout* (50%), *max-pooling* 2x2 com *stride* igual a 2, sem *padding*.



ARQUITETURAS V

- ▶ Dos estudos feitos, foi elaborada uma arquitetura final com uma camada convolucional, *dropout* de 50%, duas camadas totalmente conectadas, sendo a primeira de 250 neurônios e a segunda de 62 neurônios e uma camada de saída também de 62 neurônios. Essa rede foi treinada, então, com uma taxa de aprendizado de 0.05 e função de ativação tanh em um total de 5000 épocas.

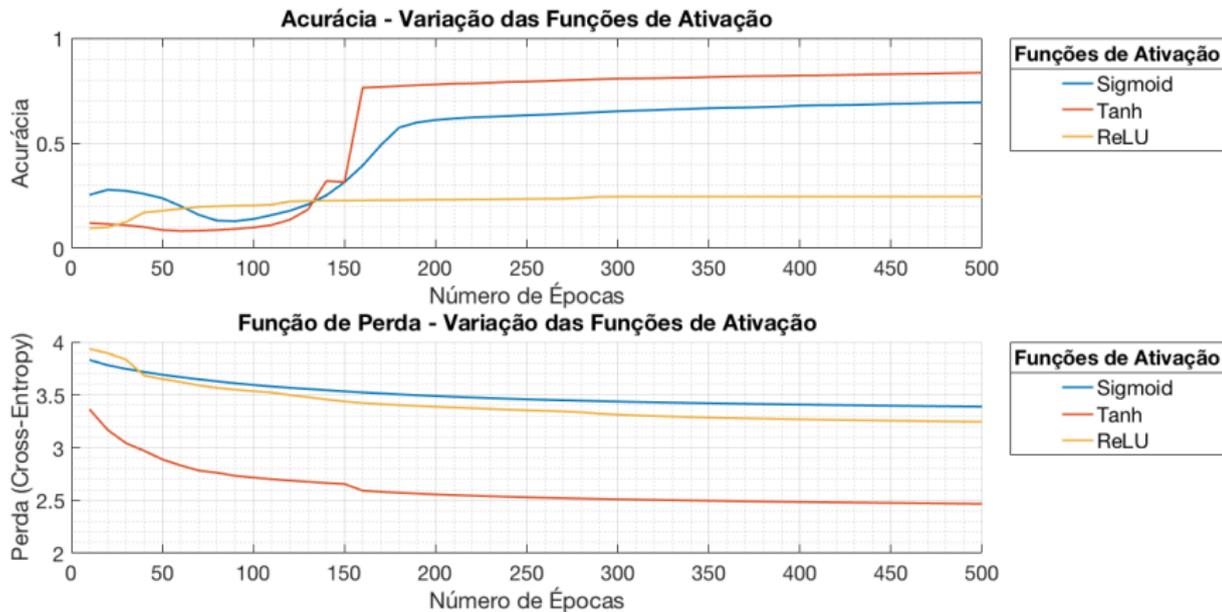


MÉTRICAS

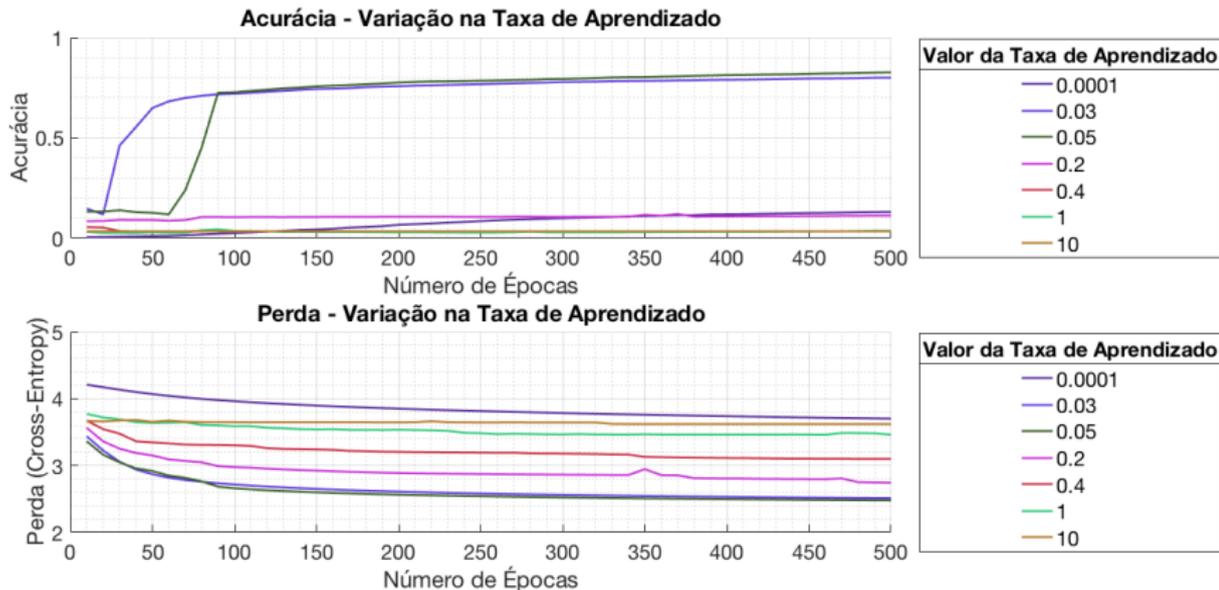
As principais métricas [Tho17] a serem analisadas e comparadas para avaliar a performance das RNAs implementadas são:

- ▶ Perda (*cross-entropy*);
- ▶ Acurácia;
- ▶ Tempo de treinamento;
- ▶ Tempo de processamento.

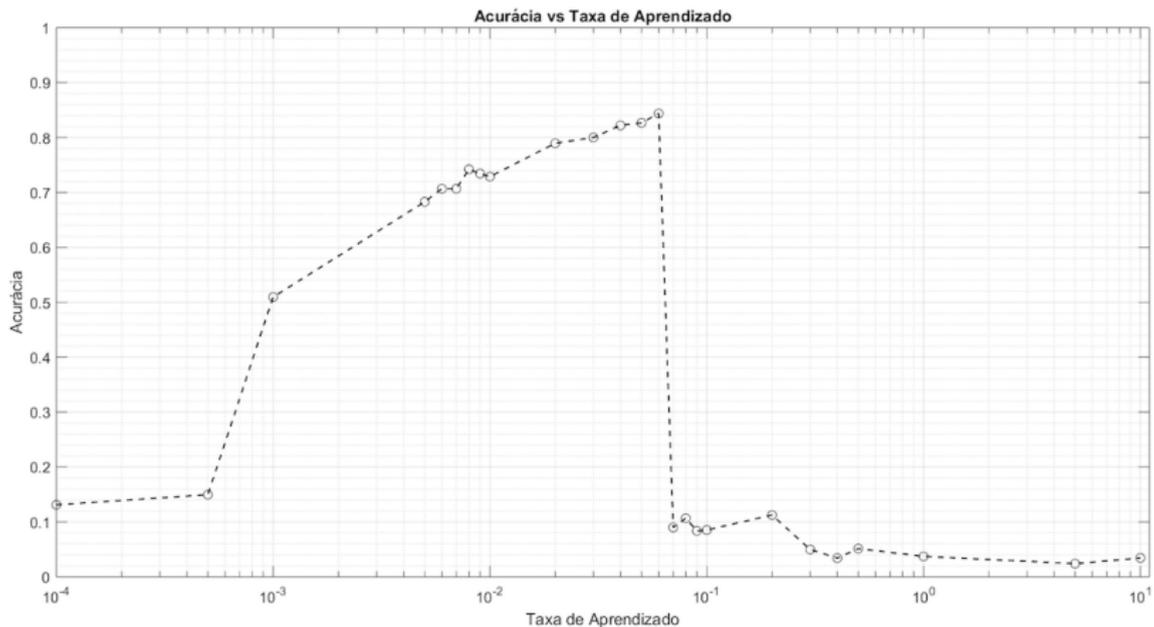
RESULTADOS - FUNÇÃO DE ATIVAÇÃO



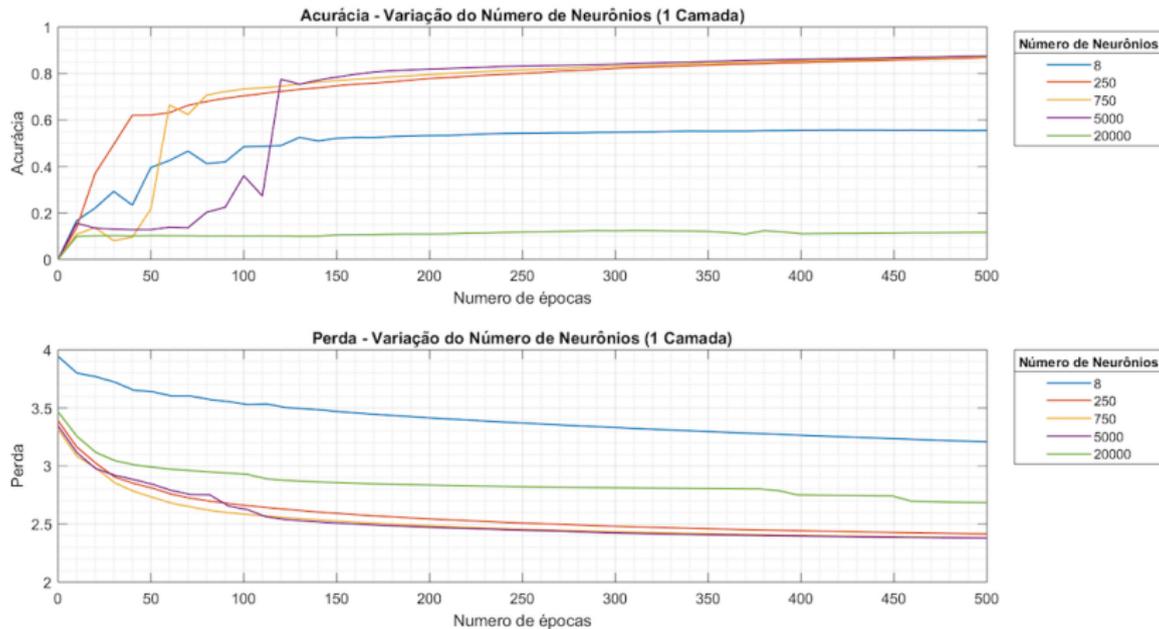
RESULTADOS - TAXA DE APRENDIZADO



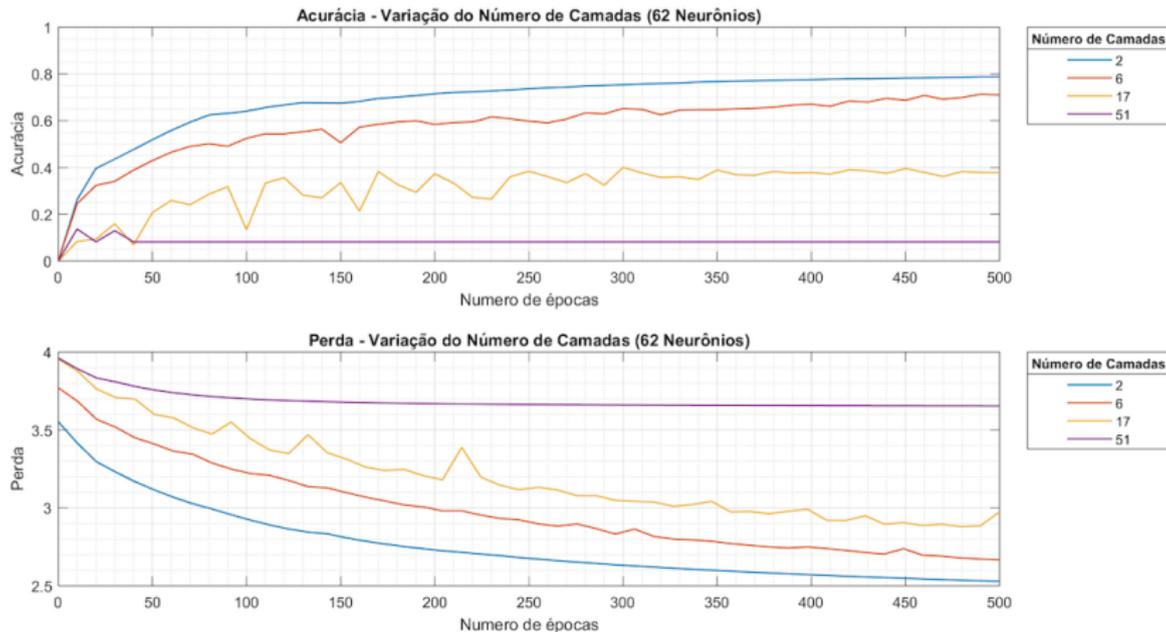
RESULTADOS - TAXA DE APRENDIZADO



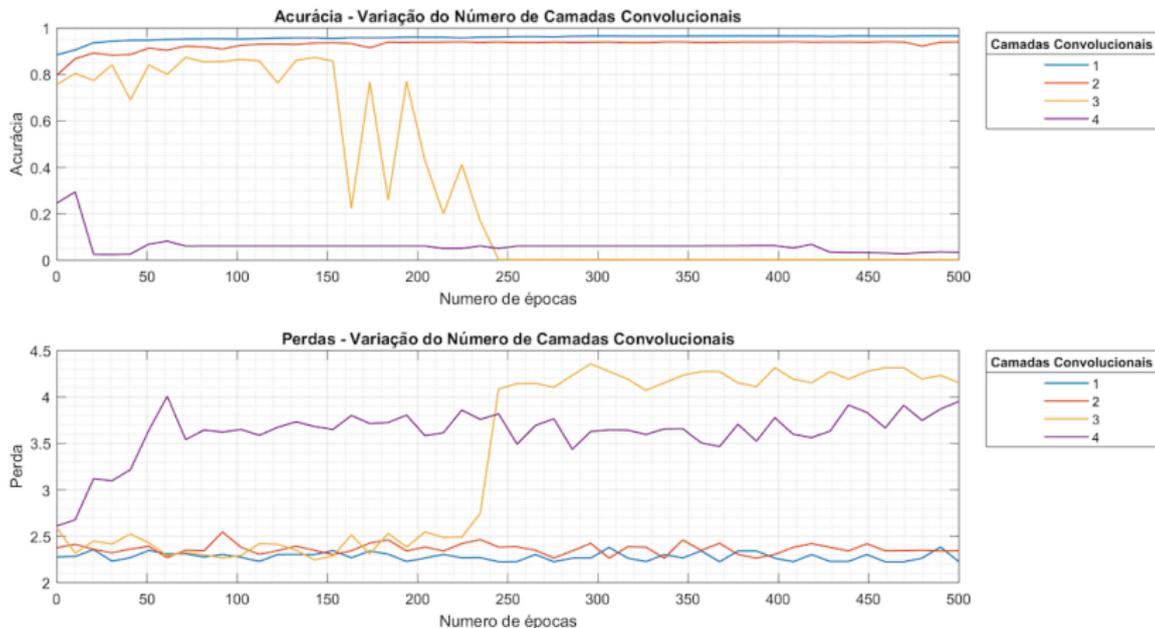
RESULTADOS - NÚMERO DE NEURÔNIOS



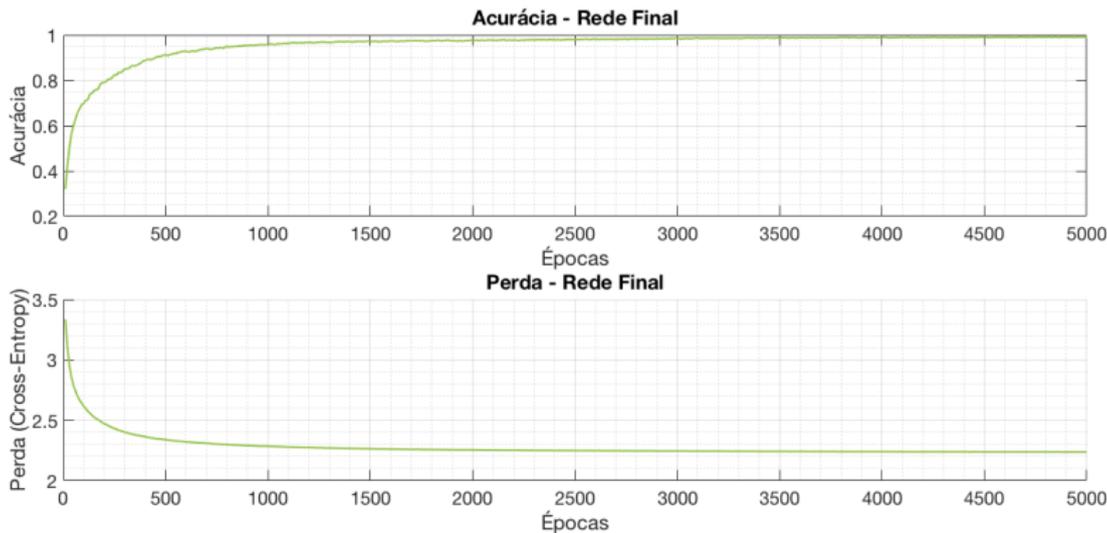
RESULTADOS - NÚMERO DE CAMADAS



RESULTADOS - NÚMERO DE CAMADAS CONVOLUCIONAIS



RESULTADOS - REDE FINAL



RESULTADOS - TEMPOS

Tabela 1: Medidas de Tempo - Variação das Funções de Ativação

Função de Ativação	Tempo de Treinamento (s)	Tempo de Processamento (s)
Sigmoid	105,34	0,19
Tanh	105,42	0,20
ReLU	107,48	0,20

RESULTADOS - TEMPOS

Tabela 2: Medidas de Tempo - Variação do Número de Camadas Convolucionais.

No. de Camadas	Tempo de Treinamento (s)	Tempo de Processamento (s)
1	335,89	0,278
2	478,64	0,285
3	629,87	0,495
4	766,97	0,598

RESULTADOS - TEMPOS

Tabela 3: Medidas de Tempo - Variação do Número de Camadas Fully Connected (com 62 Neurônios em cada camada).

No. de Camadas	Tempo de Treinamento (s)	Tempo de Processamento (s)
3	107,29	0,196
4	106,71	0,193
5	106,51	0,192
6	105,84	0,194
7	106,12	0,196
9	106,21	0,189
12	106,09	0,194
17	107,17	0,196
22	107,96	0,191
27	108,59	0,197
32	108,85	0,199
42	109,63	0,195
52	110,51	0,2015

RESULTADOS - TEMPOS

Tabela 4: Medidas de Tempo - Variação do Número de Neurônios em Uma Camada Fully Connected.

No. de Neurônios	Tempo de Treinamento (s)	Tempo de Processamento (s)
8	103,24	0,19
16	102,77	0,19
32	102,66	0,19
62	104,08	0,19
124	106,3	0,19
250	109,46	0,2
500	115,47	0,21
750	126,74	0,31
1000	133,52	0,21
1500	142,13	0,27
2000	155,03	0,24
2500	168,07	0,25
5000	230,43	0,31
7500	291,67	0,36
10000	353,03	0,42
15000	482,25	0,54
20000	604,43	0,64
30000	855,07	0,86

RESULTADOS - CONSOLIDADOS

Tabela 5: Resultados consolidados dos ensaios.

Parâmetro	Tempo de Treinamento (s)	Tempo de Processamento (s)	Acurácia (%)	Perda
Função de Ativação				
Sigmoid	105,34	0,19	69,4	3,39
lanh	105,42	0,20	83,6	2,47
ReLU	107,48	0,20	24,6	3,23
Camadas Fully Connected				
3	107,29	0,196	78,8	2,53
4	106,71	0,193	76,9	2,61
5	106,31	0,192	75,6	2,63
6	105,84	0,194	71,1	2,67
7	106,12	0,196	59,2	2,73
9	106,21	0,189	53,3	2,74
12	106,09	0,194	55,2	2,71
17	107,17	0,196	37,8	2,97
22	107,96	0,191	32,5	2,89
27	108,59	0,197	19,3	3,17
32	108,85	0,199	16,6	3,09
42	109,63	0,195	22,2	2,99
52	110,51	0,2015	8,2	3,65
No. Neurônios				
8	103,24	0,19	55,5	3,21
16	102,77	0,19	64,4	2,96
32	102,66	0,19	80,4	2,69
62	104,08	0,19	78,2	2,54
124	106,3	0,19	83,3	2,44
250	109,46	0,2	86,9	2,41
500	115,47	0,21	86,8	2,40
750	126,74	0,31	87,2	2,38
1000	133,52	0,21	86,0	2,39
1500	142,13	0,27	86,5	2,38
2000	135,03	0,24	86,5	2,38
2500	168,07	0,25	85,8,0	2,40
3000	230,43	0,31	87,5	2,38
7500	291,67	0,36	28,0	2,47
10000	333,03	0,42	10,3	2,55
15000	482,25	0,54	8,6	2,58
20000	604,43	0,64	11,6	2,68
30000	855,07	0,86	8,4	2,73
No. Camadas Convolucionais				
1	335,89	0,278	96,6	2,23
2	478,64	0,285	94,0	2,34
3	629,87	0,495	0,2	4,15
4	766,97	0,598	3,3	3,95
Rede Final				
	3.345,4	0,296	99,0	2,24

CONCLUSÕES

- ▶ A definição de parâmetros de uma RNA é, de fato, um assunto complexo e de grande relevância, sendo observado no trabalho que existem valores ótimos para os hiperparâmetros, não bastando aumentar a capacidade da rede para obter melhores resultados;
- ▶ Devido ao caráter estocástico de uma realização de treinamento, é uma boa prática tomar médias de várias realizações de treinamentos para inferir comportamentos de maneira mais precisa;
- ▶ Seria interessante observar parâmetros como peso e gradiente [GB10], o que poderia ajudar na explicação mais precisa da causa de alguns fenômenos observados nos ensaios;

REFERÊNCIAS I



Waleed Abdulla, *Traffic sign recognition with tensorflow*, Disponível em: <https://medium.com/@waleedka/traffic-sign-recognition-with-tensorflow-629dfc391a6>, 2016, Acesso em: 2018-11-25.



Haynes Ford, *Ford sign recognition system*, Disponível em: <https://www.haynesford.co.uk/Tech-Ford-Traffic-sign-recognition-system>, Acesso em: 2018-11-25.



Xavier Glorot and Yoshua Bengio, *Understanding the difficulty of training deep feedforward neural networks*, Proceedings of the thirteenth international conference on artificial intelligence and statistics, 2010, pp. 249–256.



Martin Thoma, *Analysis and optimization of convolutional neural network architectures*, Masters's thesis, Karlsruhe Institute of Technology, Karlsruhe, Germany, June 2017.



R. Timofte, *Belgiumts dataset*, Disponível em: <https://btsd.ethz.ch/shareddata/>, 2014, Acesso em: 2018-11-07.

DÚVIDAS

Perguntas?
Obrigado!