

NICOLI PINHEIRO DE ARAUJO

**ESTIMAÇÃO INTELIGENTE DE IDADE DE TELESPECTADORES PARA
APLICAÇÕES DE SUGESTÃO DE CONTEÚDO EM *SMART* TVS**

Trabalho de Conclusão de Curso apresentado
à banca avaliadora do Curso de Engenharia
de Computação, da Escola Superior de
Tecnologia, da Universidade do Estado do
Amazonas, como pré-requisito para obtenção
do título de Engenheira de Computação.

Orientador(a): Profa. Dra. Elloá Barreto Guedes da Costa

Manaus – Novembro – 2018

Capítulo 1

Resultados e Discussão

Considerando a abordagem descrita na solução proposta, os resultados da execução das CNNs aplicadas ao problema de estimação de idade a partir de uma imagem de face são apresentados a seguir.

1.0.1 Abordagem 1

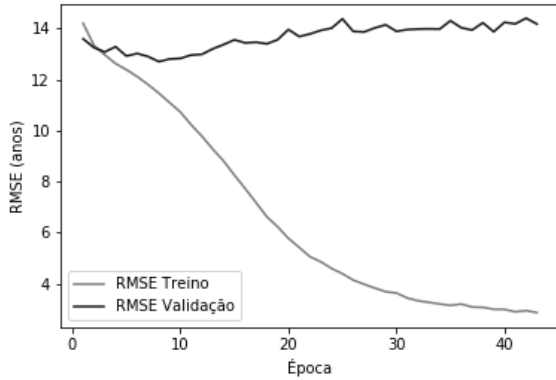
A primeira abordagem de treinamento das CNNs utilizou as imagens da base de dados com equalização por histograma de frequência, normalizadas, e com 50% de chance de estarem rotacionadas horizontalmente. Conforme mencionado na Seção ??, os treinamentos e testes compreenderam as arquiteturas canônicas LeNet e AlexNet com funções de ativação *ReLU* e *Leaky ReLU* nas camadas ocultas e de ativação. É importante ressaltar que neste momento não foram utilizadas técnicas de *transfer learning*.

A CNN que implementa a arquitetura LeNet com função de ativação *ReLU* foi treinada por 43 épocas, obteve MAE de 14.09 e RMSE 17.93. A LeNet com função de ativação *Leaky ReLU* foi treinada por 15 épocas, obteve MAE de 14.44 anos e RMSE de 18.18 anos. Os treinamentos duraram aproximadamente 16 e 12 horas respectivamente, em uma instância do Google Compute Engine com 4 CPUs virtuais e 15 GB de RAM. Os gráficos de treinamento e as retas zero obtidas a partir da apresentação do conjunto de teste aos modelos consolidados podem ser vistos na Figura 1.1. É possível notar que ambas as redes sofreram com overfitting e

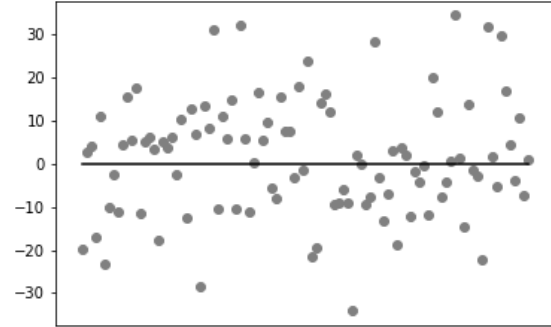
obtiveram grande margem de erro, no entanto a LeNet que utilizou *Leaky ReLU* como função de ativação obteve um desempenho mais satisfatório. .

Figura 1.1: Resultados do treinamento e teste da CNN LeNet.

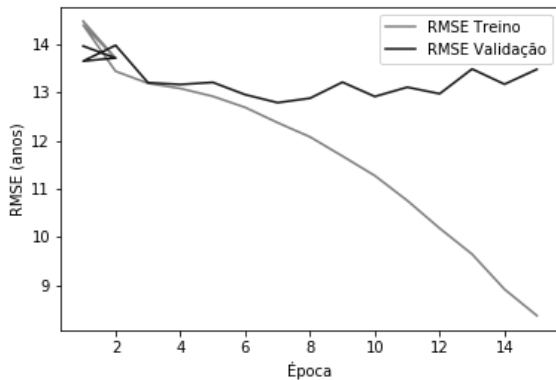
(a) RMSE de treinamento da arquitetura LeNet utilizando funções de ativação *ReLU*.



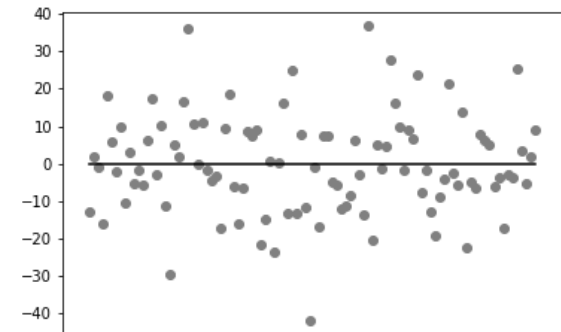
(b) Reta-0 LeNet *ReLU*.



(c) RMSE de treinamento da arquitetura LeNet utilizando funções de ativação *Leaky ReLU*.



(d) Reta-0 LeNet *Leaky ReLU*.

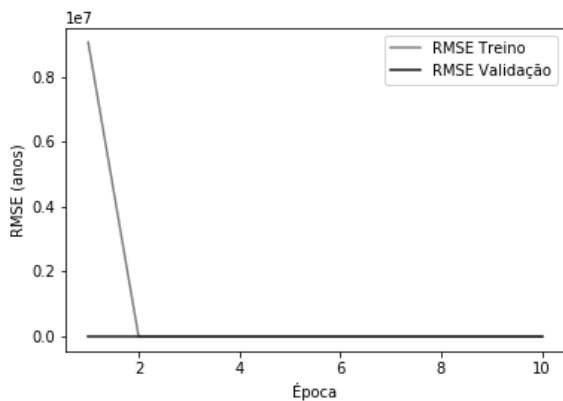


A CNN que implementa a arquitetura AlexNet com função de ativação *ReLU* foi treinada por 10 épocas, obteve MAE de 38.63 anos e RMSE de 41.22 anos. A AlexNet com função de ativação *Leaky ReLU* foi treinada por 30 épocas, obteve MAE de 14.44 anos e RMSE de 15.33 anos. Os treinamentos duraram aproximadamente 15 e 38 horas respectivamente, na mesma instância do Google Compute Engine utilizada para o treinamento das redes LeNet. Os gráficos de treinamento e as retas zero obtidas a partir da apresentação do conjunto de teste aos modelos consolidados podem ser vistos na Figura 1.2. É possível notar que a AlexNet que utiliza *ReLU* sofreu de *dying ReLU problem*, o que culminou em previsões iguais a zero para

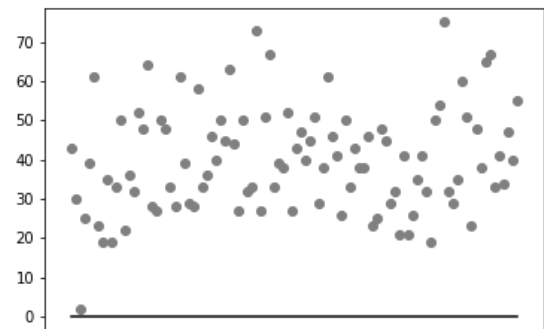
todos os exemplos do conjunto de teste, o que pode ser evidenciado pela ausência total de acertos conforme mostra a Figura 1.2b. Para contornar este problema, a AlexNet que utilizou *Leaky ReLU* como função de ativação foi capaz de convergir para uma solução mais adequada, prevendo idades mais próximas às reais.

Figura 1.2: Resultados do treinamento e teste da CNN AlexNet.

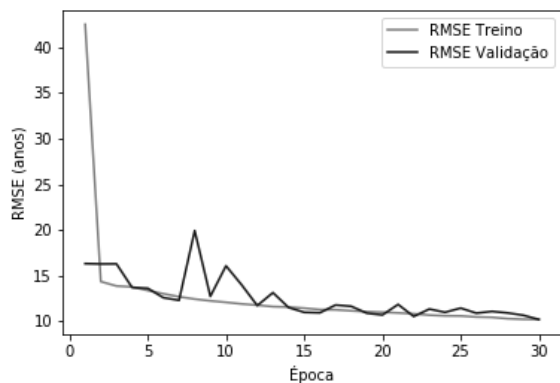
(a) Treinamento AlexNet *ReLU*.



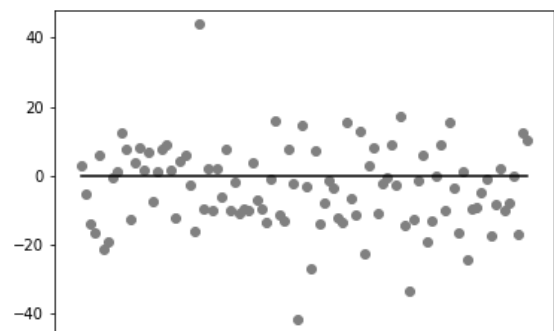
(b) Reta-0 AlexNet *ReLU*.



(c) Treinamento AlexNet *Leaky ReLU*.



(d) Reta-0 AlexNet *Leaky ReLU*.



Obedecendo ao método de validação cruzada *holdout* previamente mencionado, os resultados desta abordagem encontram-se sintetizados na Tabela 1.1.

Tabela 1.1: Resultados do treino e teste dos modelos propostos na Abordagem 1.

Rede	Função de ativação	Parâmetros	Épocas	Tempo de treinamento	MAE Teste	RMSE Teste
LeNet	<i>Leaky ReLU</i>	params	15	12 h	14.44	18.18
LeNet	<i>ReLU</i>	params	43	16 h	14.09	17.93
AlexNet	<i>ReLU</i>	58.286.145	10	15 h	38.63	41.22
AlexNet	<i>Leaky ReLU</i>	params	30	40 h	15.33	18.58

1.0.2 Abordagem 2

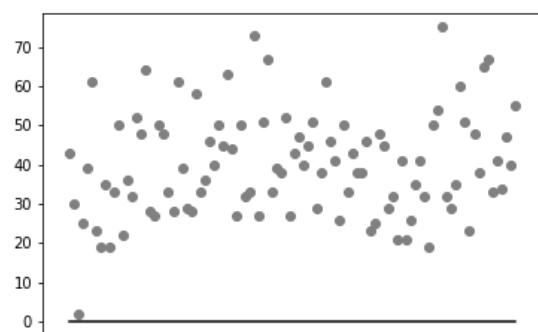
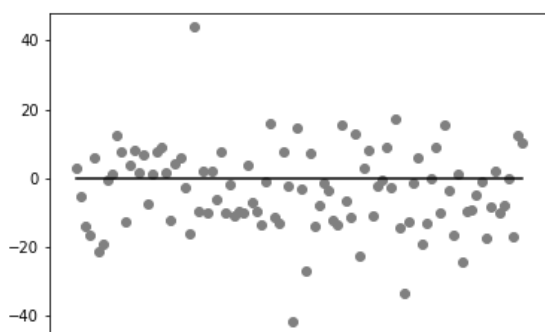
- Mesmas redes - Normalização das imagens, equalização por histograma -> o que é - data augmentation -> mais técnicas de data augmentation

1.0.3 Abordagem 3

Outras arquiteturas VGG com transfer learning 1. Retirar última camada (softmax) e adicionar leaky relu 2. Retirar duas últimas camadas (dense e softmax) e adicionar leaky relu

Figura 1.3: Redes neurais biológicas.

(a) Reta-0 Alexnet LReLU com imagens normaliza- (b) Reta-0 Alexnet ReLU com imagens normaliza-
das e equalizadas das e equalizadas



Capítulo 2

Considerações Finais

O objetivo deste trabalho consiste em elaborar estratégias inteligentes para estimação de idade de telespectadores de *Smart* TVs a partir de suas respectivas fotografias faciais. Para este fim, foram propostos, treinados e testados em caráter preliminar dois modelos de CNNs já bem estabelecidos na literatura, a LeNet e AlexNet, com dois perfis de hiperparâmetros cada um.

Com isto, observou-se uma melhora significativa na performance da AlexNet, enquanto o RMSE da LeNet não sofreu grandes mudanças. Quanto às saídas das redes, a LeNet exibiu valores positivos e negativos próximos de zero, e a AlexNet forneceu previsões que giravam em torno da média dos dados. Estes resultados são preliminares e certamente outros modelos e parâmetros serão investigados conforme previsto na metodologia e cronograma deste trabalho de conclusão de curso.

Nos próximos meses, os esforços estarão concentrados em pesquisar e adotar estratégias que possam minimizar os problemas identificados, como substituir as funções de ativação das camadas ocultas por outras variantes da *ReLU*, adotar métodos específicos de inicialização de pesos, normalização de *batch*, entre outros. Planeja-se também a proposição, o treinamento e teste de outras redes inspiradas em outros modelos canônicos.

O problema em questão é importante do ponto de vista prático para o desenvolvimento de diversas soluções de recomendação de conteúdo e controle parental em *Smart* TVs, auxiliando no desenvolvimento destas soluções tecnológicas. Considerando a formação de uma bacharelada em Engenharia de Computação, endereçar este problema permite a prática de diversos conceitos

vistos ao longo do curso, em especial relacionados às disciplinas de Inteligência Artificial, Redes Neurais, Processamento Digital de Imagens, *Machine Learning* e Sinais e Sistemas.

Referências Bibliográficas