

Estimação Inteligente de Idade de Telespectadores para Aplicações de Sugestão de Conteúdo em Smart TVs

Trabalho de Conclusão de Curso II

por

Nicoli P. Araújo, Elloá B. Guedes

{npda.eng, ebgcosta}@uea.edu.br

do

Núcleo de Computação
Escola Superior de Tecnologia
Universidade do Estado do Amazonas
Manaus – Amazonas – Brasil

Outline

1. Apresentação

2. Fundamentação Teórica

3. Trabalhos Relacionados

4. Solução Proposta

5. Resultados Parciais

6. Considerações Parciais

Contextualização e Motivação

- ***Smart TVs***
 - Conexão à internet
 - Transmissão de conteúdos advindos de outros dispositivos eletrônicos
- Pesquisa Nacional por Amostra de Domicílios 2015
 - *Smart TVs* representam 68,2% do total de televisores vendidos no primeiro semestre de 2017
- Aumento da popularidade causado por:
 - *Smart TVs versus* televisões convencionais
 - Encerramento da transmissão de sinal analógico da televisão aberta
 - Copa do Mundo 2018
 - Tecnologia 4K

Contextualização e Motivação

- ▶ Sistemas inteligentes de identificação de características:
 - ▶ Dificuldades práticas
 - ▶ Algumas *Smart TVs* possuem câmeras
 - ▶ Aplicações: Recomendação de conteúdo e controle parental

Objetivos

Objetivo Geral

Desenvolvimento de estratégias inteligentes, baseadas na utilização de técnicas de *Deep Learning*, para estimativa da idade de telespectadores a partir de fotografias faciais

Objetivos

Objetivo Geral

Desenvolvimento de estratégias inteligentes, baseadas na utilização de técnicas de *Deep Learning*, para estimativa da idade de telespectadores a partir de fotografias faciais

Objetivos Específicos

1. Referencial teórico sobre redes neurais convolucionais
2. Consolidar uma base de dados
3. Identificar tecnologias adequadas
4. Propor, treinar e testar diferentes estimadores de idade
5. Avaliar comparativamente os estimadores propostos

Justificativas

- Melhor experiência de provimento de conteúdo e de configurações personalizadas
- Prática e a proposição de soluções envolvendo *Machine Learning*
- Laboratório de Sistemas Inteligentes (LSI)

Metodologia

- Fundamentação teórica sobre Machine Learning
- Fluxo de atividades de *machine learning* (Marsland, 2015)
 - Consolidar uma base de dados
 - Propor modelos aplicáveis
 - Treinamento
 - Teste
 - Comparação de métricas de desempenho
- Escrita da proposta e do projeto final do trabalho de conclusão de curso

Cronograma

Tabela 1: Cronograma de atividades

| | 2018 | | | | | | | | | | |
|---|------|----|----|----|----|----|----|----|----|----|----|
| | 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 |
| Escrita da Proposta | X | X | X | X | X | | | | | | |
| Fundamentação Teórica sobre ML | X | X | X | X | | | | | | | |
| Consolidação da Base de Dados | | X | X | | | | | | | | |
| Proposição de Modelos de CNNs | | | | X | X | X | X | X | | | |
| Defesa da Proposta | | | | | X | | | | | | |
| Escrita do Trabalho Final | | | | | | X | X | X | X | X | X |
| Treinamento das CNNs | | | | | X | X | X | X | X | X | |
| Teste das CNNs | | | | | X | X | X | X | X | X | X |
| Comparação de Métricas de Desempenho | | | | | | X | X | X | X | X | X |
| Defesa do Trabalho Final | | | | | | | | | | | X |

Outline

1. Apresentação

2. Fundamentação Teórica

3. Trabalhos Relacionados

4. Solução Proposta

5. Resultados Parciais

6. Considerações Parciais

Redes Neurais Artificiais

McCulloch e Pitts, 1943

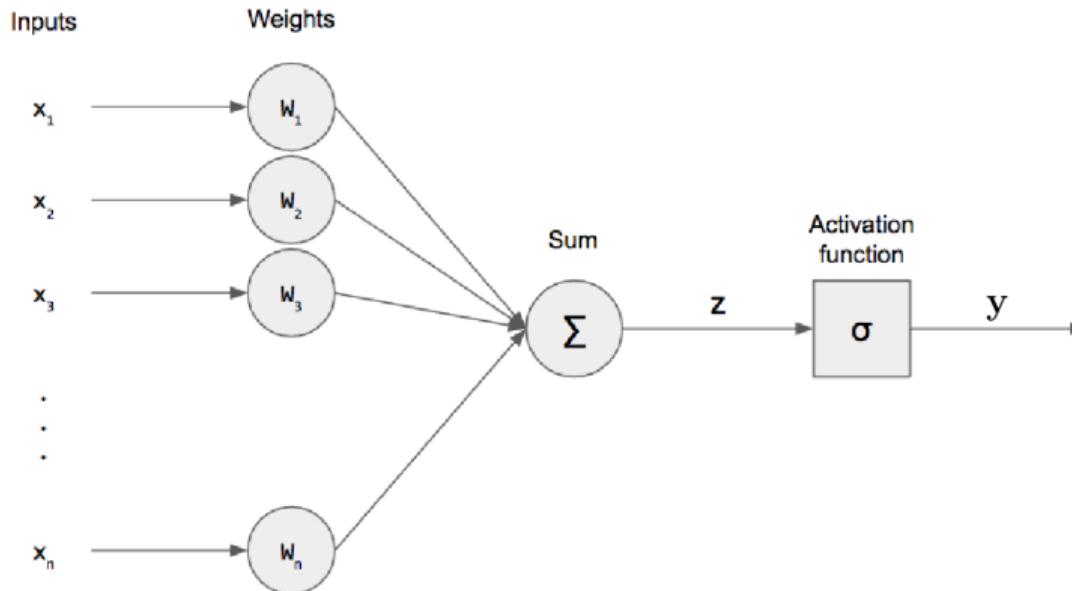


Figura 1: Representação de um neurônio artificial

Redes Neurais Convolucionais

- Topologia bem definida e estrutura em grid
- Destaca-se no reconhecimento de padrões em dados de alta dimensionalidade
- Diferentes tipos de camadas:
 - Camada convolucional
 - Camada de ativação
 - Camada de pooling

Redes Neurais Convolucionais

▪ Convolução

$$S(i, j) = I(i, j) * K(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n) \quad (1)$$

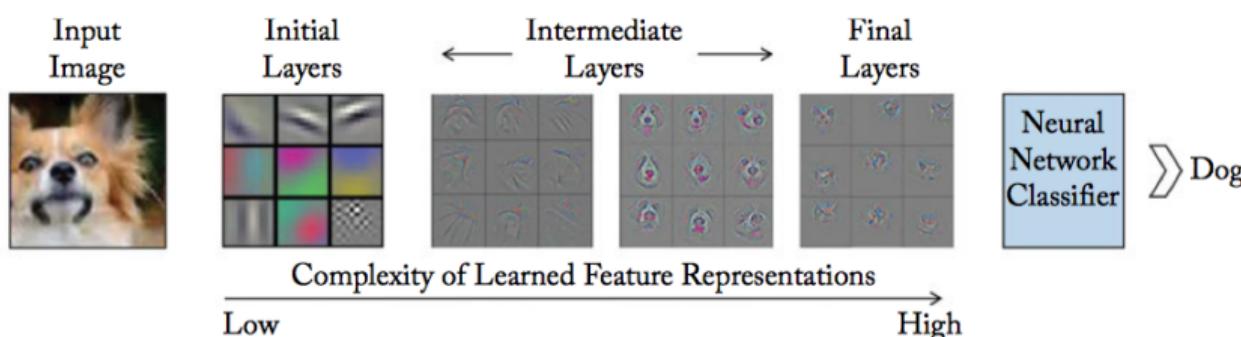


Figura 2: Papel das camadas convolucionais e *feature maps* nas CNNs.

Modelos Canônicos de Redes Neurais Convolucionais

- Arquiteturas que trouxeram contribuições importantes
- Comuns ainda hoje no cenário de DL
- LeNet (1998)
- AlexNet (2012)
- VGG (2014)
- Inception (2014)
- ResNet (2015)
- *Transfer Learning*: Aproveitamento de parâmetros treinados

Outline

1. Apresentação

2. Fundamentação Teórica

3. Trabalhos Relacionados

4. Solução Proposta

5. Resultados Parciais

6. Considerações Parciais

Trabalhos Relacionados

‣ DEX (Rothe,2015)

- Problema de classificação: idade entre 0 e 100
- Datasets: *IMDb*, *Wikipedia* e *Looking At People*
- Ensemble de 20 VGG-16, *Transfer Learning*
- MAE (*Mean Average Error*) de 3.221 na fase de testes

Trabalhos Relacionados

‣ **DEX** (Rothe,2015)

- Problema de classificação: idade entre 0 e 100
- Datasets: *IMDb*, *Wikipedia* e *Looking At People*
- Ensemble de 20 VGG-16, *Transfer Learning*
- MAE (*Mean Average Error*) de 3.221 na fase de testes

‣ **AgeNet** (Liu et al., 2015)

- Ensemble de modelos regressores e classificadores com diferentes métodos de normalização de face
- CNNs inspiradas na *GoogLeNet*
- MAE de 3.3345

Trabalhos Relacionados

➢ Levi e Hassner, 2015:

- Idade e gênero
- Problema de classificação: 8 classes de idade
- Três camadas convolucionais com *dropout* e funções de ativação *ReLU*, seguidas por três camadas totalmente conectadas
- Acurácia de $84.7\% \pm 2.2$

Outline

1. Apresentação

2. Fundamentação Teórica

3. Trabalhos Relacionados

4. Solução Proposta

5. Resultados Parciais

6. Considerações Parciais

Tarefa de Aprendizado

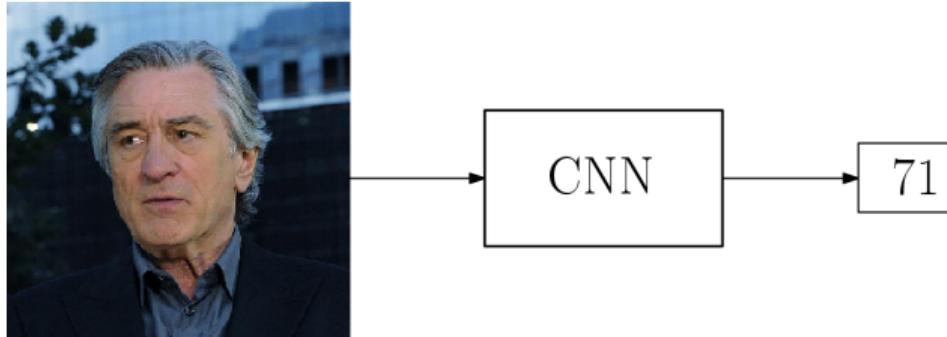
- Tarefa de Regressão

- Entrada:

- Imagem em cores RGB de dimensões 224×224 pixels contendo uma face humana centralizada

- Saída:

- Estimativa de idade, em anos, da pessoa correspondente



Tarefa de Aprendizado

- Métricas de desempenho:

- *Root Mean Squared Error (RMSE)*

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2}. \quad (2)$$

- *Mean Absolute Error (MAE)*

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}| \quad (3)$$

Conjunto de Dados

- Base de dados experimentais IMDb
 - 452.132 exemplos
 - 20.284 dos atores mais populares listados no site IMDb
 - Organizada por Rothe et al., 2015
 - Imagens e meta-dados

Conjunto de Dados

| Meta-dado | Valor |
|---------------------------|---|
| ID Celebridade | 16349 |
| Nome | Robert De Niro |
| Endereço da imagem | imdb/34/nm0000134_rm334009 0368_1943-8-17_2011.jpg |
| Pontuação da Face | 5.21396 |
| Pontuação da Segunda Face | NaN |
| Localização da Face | (663.65, 992.475, 590.134, 918.959) |
| Data de Nascimento | 1943 – 08 – 17 |
| Ano da Foto | 2011 |
| Gênero | Masculino |

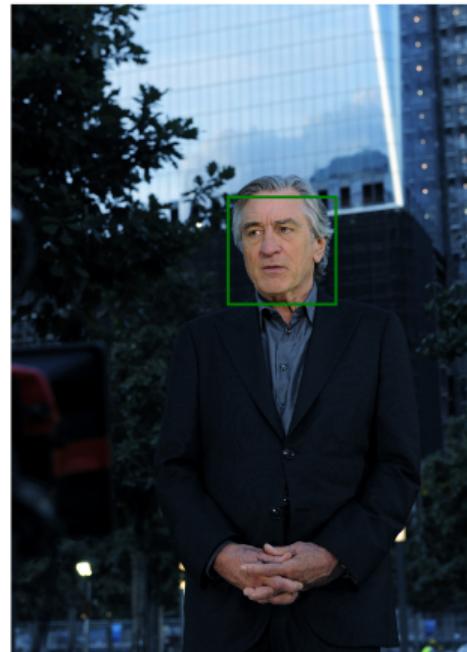


Figura 3: Exemplo de imagem do conjunto de dados contendo apenas um rosto.

Conjunto de Dados

| Meta-dado | Valor |
|---------------------------|---|
| ID Celebridade | 16349 |
| Nome | Robert De Niro |
| Endereço da imagem | imdb/34/nm0000134_rm14800 44288_1943-8-17_2012.jpg |
| Pontuação da Face | 5.51656 |
| Pontuação da Segunda Face | 4.55379 |
| Localização da Face | (1392.72, 1614.18, 225.55, 447.003) |
| Data de Nascimento | 1943 – 08 – 17 |
| Ano da Foto | 2012 |
| Gênero | Masculino |



Figura 4: Exemplo de imagem do conjunto de dados contendo mais de um rosto com a classificação errônea.

Limpeza e Pré-processamento dos dados

- Versão original: 267 GB
- Faces recortadas: 7,1 GB
- Cálculo do atributo alvo: **Idade**
- Itens descartados:
 - Idade e gênero apresentando valores nulos, inválidos ou negativos
 - Múltiplos exemplos referentes à mesma pessoa com a mesma idade
 - Idade maior que 100 anos ou não compatível com os dados da celebridade referida nos meta-dados
 - Ausência de rosto
 - Presença de mais de uma face na imagem

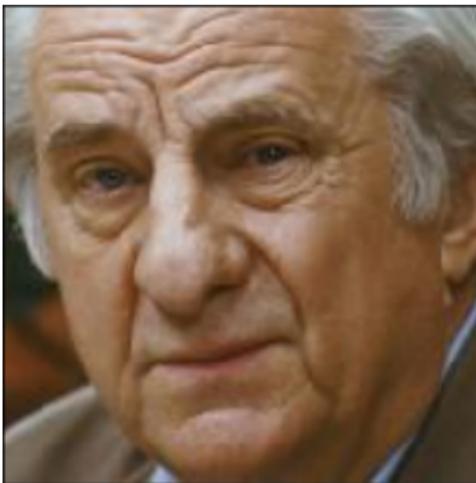
Limpeza e Pré-processamento dos dados

- Padronização das dimensões das imagens
 - 224×224 pixels
 - RGB
- Descarte de meta-dados irrelevantes para a tarefa de aprendizado

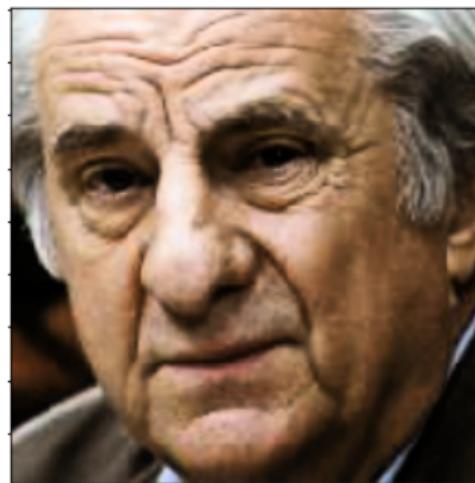
Limpeza e Pré-processamento dos dados

‣ Equalização por histograma

(a) Imagem sem equalização de histograma.



(b) Imagem após equalização por histograma.



Limpeza e Pré-processamento dos dados

- Conjunto de dados consolidado:
 - 47.950 exemplos
 - 14.607 celebridades
 - 1,2GB em disco

- Divisão obedecendo o método *Holdout*
 - Treinamento–Validação–Teste
 - 70% – 10% – 20%
 - 33.565 – 4.795 – 9.590

Modelos de CNN Considerados

▣ Arquiteturas

- LeNet
- AlexNet
- VGG-16
- SqueezeNet

Modelos de CNN Considerados

- Funções de ativação: *ReLU* ou *Leaky ReLU*
- Método de otimização do gradiente descendente *Adam*
- Camadas de saída com apenas um neurônio
- *batch size* igual a 64
- Número de épocas obtida de maneira experimental

Outline

1. Apresentação

2. Fundamentação Teórica

3. Trabalhos Relacionados

4. Solução Proposta

5. Resultados Parciais

6. Considerações Parciais

Resultados Parciais

- Estimação de idade a partir de uma imagem de face
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 1
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 2
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 3
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 4
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 5
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 6
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 7
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 8
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Resultados Parciais

- Abordagem 9
- LeNet, AlexNet, VGG-16, SqueezeNet
- **9 Abordagens**

Outline

1. Apresentação

2. Fundamentação Teórica

3. Trabalhos Relacionados

4. Solução Proposta

5. Resultados Parciais

6. Considerações Parciais

Considerações Parciais

Estimação Inteligente de Idade de Telespectadores para Aplicações de Sugestão de Conteúdo em Smart TVs

Elaborar estratégias inteligentes para estimação de idade de telespectadores de *Smart TVs* a partir de suas respectivas fotografias faciais

Modelos de CNNs baseados na LeNet e AlexNet

Trabalho de Conclusão de Curso II

- Dois perfis de hiperparâmetros considerados
 - por
- Melhoria de performance verificada na AlexNet
- Em suma:
 - {npda.eng, ebgcosta}@uea.edu.br
 - LeNet exibiu valores positivos e negativos próximos de zero
 - AlexNet forneceu previsões que giravam em torno da média dos dados

Núcleo de Computação

Escola Superior de Tecnologia

Universidade do Estado do Amazonas

Manaus – Amazonas – Brasil

Considerações Parciais

Próximos Passos

- ▶ Pesquisar e adotar estratégias que possam minimizar os problemas identificados
 - ▶ Considerar a tarefa como um problema de classificação?
- ▶ Proposição, treinamento e teste de outras redes inspiradas em outros modelos canônicos

