

Colorização de Imagens com Deep Learning

Defesa do Trabalho de Conclusão de Curso I

Giovana de Lucca

Orientadora: Elloá B. Guedes

{gol.eng, elloa.uea}@gmail.com

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

21 de junho de 2018



Agenda

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

Agenda - Seção 1

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

Aprendizado de Máquina

- As técnicas de **Aprendizado de Máquina** têm sido aplicadas com sucesso em um grande número de problemas reais em diversos domínios
- Características: natureza inferencial e a boa capacidade de generalização dos métodos e técnicas desta área
- Algoritmos capazes de aprender padrões por meio de exemplos, baseado-se em dados previamente disponíveis

Aprendizado de Máquina

- As técnicas de **Aprendizado de Máquina** têm sido aplicadas com sucesso em um grande número de problemas reais em diversos domínios
- Características: natureza inferencial e a boa capacidade de generalização dos métodos e técnicas desta área
- Algoritmos capazes de aprender padrões por meio de exemplos, baseado-se em dados previamente disponíveis

Visão Computacional

- A **Visão Computacional** é uma área que procura desenvolver métodos capazes de replicar nos computadores as capacidades da visão humana.
- Procedimentos de extração de características de imagens envolviam um grande esforço devido ao processamento
- Análises de especialistas eram utilizadas para descobrir as regras necessárias para o reconhecimento e extração de certos padrões

Visão Computacional

- A **Visão Computacional** é uma área que procura desenvolver métodos capazes de replicar nos computadores as capacidades da visão humana.
- Procedimentos de extração de características de imagens envolviam um grande esforço devido ao processamento
- Análises de especialistas eram utilizadas para descobrir as regras necessárias para o reconhecimento e extração de certos padrões

Deep Learning

- Algoritmos capazes de lidar com a demanda de cálculos matemáticos complexos automatizados
- Necessidade de ferramentas e algoritmos mais sofisticados
 - Complexidade de problemas computacionais
 - Grande volume de dados constantemente gerados
- Técnicas de Deep Learning: conjunto de camadas de redes neurais artificiais
 - Simplicidade da construção das redes
 - Escalabilidade
 - Transferência de conhecimento

Deep Learning

- Algoritmos capazes de lidar com a demanda de cálculos matemáticos complexos automatizados
- Necessidade de ferramentas e algoritmos mais sofisticados
 - Complexidade de problemas computacionais
 - Grande volume de dados constantemente gerados
- Técnicas de **Deep Learning**: conjunto de camadas de redes neurais artificiais
 - Simplicidade da construção das redes
 - Escalabilidade
 - Transferência de conhecimento

Agenda - Seção 2

- 1 Introdução
- 2 Objetivos**
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

Objetivos

Objetivo Geral

Explorar estratégias para colorização artificial de imagens utilizando técnicas de Deep Learning

Objetivos Específicos

- Consolidar uma base de dados representativa de imagens coloridas para treinamento das redes;
- Descrever o problema da colorização artificial de imagens segundo uma tarefa de Aprendizado de Máquina;
- Explorar a utilização das arquiteturas canônicas de redes neurais convolucionais mediante *Transfer Learning* aplicadas ao problema considerado;
- Propor, treinar e testar diferentes redes neurais convolucionais baseadas nas arquiteturas canônicas elencadas;
- Analisar os resultados obtidos de maneira quantitativa e qualitativa.

Objetivos

Objetivo Geral

Explorar estratégias para colorização artificial de imagens utilizando técnicas de Deep Learning

Objetivos Específicos

- Consolidar uma base de dados representativa de imagens coloridas para treinamento das redes;
- Descrever o problema da colorização artificial de imagens segundo uma tarefa de Aprendizado de Máquina;
- Explorar a utilização das arquiteturas canônicas de redes neurais convolucionais mediante *Transfer Learning* aplicadas ao problema considerado;
- Propor, treinar e testar diferentes redes neurais convolucionais baseadas nas arquiteturas canônicas elencadas;
- Analisar os resultados obtidos de maneira quantitativa e qualitativa.

Agenda - Seção 3

- 1 Introdução
- 2 Objetivos
- 3 Justificativa**
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

Justificativa

- Influência da coloração na interpretação das imagens:
 - Valor histórico e artístico de arquivos antigos
 - Imagens de câmeras de segurança com baixa resolução
- Colorizações ajustadas para restauração de algum tipo de perturbação visual
- Prática de conceitos, técnicas e tecnologias de uma área emergente da Computação
- Proposta alinhada com as atividades desenvolvidas pelo Laboratório de Sistemas Inteligentes

Justificativa

- Influência da coloração na interpretação das imagens:
 - Valor histórico e artístico de arquivos antigos
 - Imagens de câmeras de segurança com baixa resolução
- Colorizações ajustadas para restauração de algum tipo de perturbação visual
- Prática de conceitos, técnicas e tecnologias de uma área emergente da Computação
- Proposta alinhada com as atividades desenvolvidas pelo Laboratório de Sistemas Inteligentes

Justificativa

- Influência da coloração na interpretação das imagens:
 - Valor histórico e artístico de arquivos antigos
 - Imagens de câmeras de segurança com baixa resolução
- Colorizações ajustadas para restauração de algum tipo de perturbação visual
- Prática de conceitos, técnicas e tecnologias de uma área emergente da Computação
- Proposta alinhada com as atividades desenvolvidas pelo Laboratório de Sistemas Inteligentes

Justificativa

- Influência da coloração na interpretação das imagens:
 - Valor histórico e artístico de arquivos antigos
 - Imagens de câmeras de segurança com baixa resolução
- Colorizações ajustadas para restauração de algum tipo de perturbação visual
- Prática de conceitos, técnicas e tecnologias de uma área emergente da Computação
- Proposta alinhada com as atividades desenvolvidas pelo **Laboratório de Sistemas Inteligentes**

Agenda - Seção 4

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia**
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

Metodologia

A condução das atividades obedece à metodologia a seguir, composta dos seguintes passos:

1. Estudo dos conceitos relacionados à Aprendizado de Máquina, *Deep Learning* e as principais arquiteturas de redes neurais convolucionais;
2. Estudo do ferramental tecnológico para elaboração e execução de projetos de *Deep Learning*, incluindo Python, Keras, Sci-kit Learn, Google Cloud Platform, dentre outros;
3. Elaborar uma base de dados representativa de imagens coloridas e em escalas de cinza para fins de aprendizado dos padrões de coloração pelas redes neurais convolucionais;
4. Elencar um conjunto de arquiteturas canônicas das redes neurais convolucionais aplicáveis ao problema em questão;

Metodologia

5. Propor modificações nas redes neurais identificadas no passo anterior mediante *Transfer Learning*;
6. Treinar as redes modificadas com os exemplos da base de dados;
7. Testar as redes e coletar métricas de desempenho;
8. Analisar os resultados obtidos identificando as redes mais adequadas ao cenário considerado;
9. Escrita da proposta do Trabalho de Conclusão de Curso;
10. Defesa da proposta do Trabalho de Conclusão de Curso;
11. Escrita do Trabalho de Conclusão de Curso;
12. Defesa do Trabalho de Conclusão de Curso.

Agenda - Seção 5

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma**
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

Cronograma

	2018											
	02	03	04	05	06	07	08	09	10	11	12	
Estudo dos conceitos teóricos relacionados à <i>Machine Learning</i>	x	x	x	x								
Estudo do ferramental tecnológico para elaboração do projeto	x	x	x	x								
Consolidação da base de dados			x	x								
Especificação das arquiteturas canônicas de redes neurais convolucionais			x									
Aplicação das técnicas de <i>Transfer Learning</i> nas redes neurais convolucionais identificadas				x								
Treinamento das redes neurais convolucionais com os exemplos da base de dados				x	x	x	x					
Testes das redes e comparação de métricas de desempenho					x	x	x	x				
Análise dos resultados obtidos									x	x	x	
Escrita da proposta do trabalho	x	x	x	x	x							
Defesa da proposta do trabalho					x							
Escrita do trabalho final						x	x	x	x	x	x	
Defesa do trabalho final											x	

Agenda - Seção 6

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica**
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

Redes Neurais Convolucionais

- Redes *feedforward* com muitas camadas ocultas
- Cada camada possui uma funcionalidade básica específica e produz um **mapa de características**
- Realização de operações de convolução
 - Extração de características por meio de um filtro
 - Parâmetros: extensão espacial, *stride*, *zero-padding*, etc
- Camada *pooling* responsável por produzir um mapa de características condensado
- Camadas completamente conectadas baseadas nos conceitos de uma rede *Multilayer Perceptron*

Redes Neurais Convolucionais

- Redes *feedforward* com muitas camadas ocultas
- Cada camada possui uma funcionalidade básica específica e produz um **mapa de características**
- Realização de operações de **convolução**
 - Extração de características por meio de um filtro
 - Parâmetros: extensão espacial, *stride*, *zero-padding*, etc
- Camada *pooling* responsável por produzir um mapa de características condensado
- Camadas completamente conectadas baseadas nos conceitos de uma rede *Multilayer Perceptron*

Redes Neurais Convolucionais

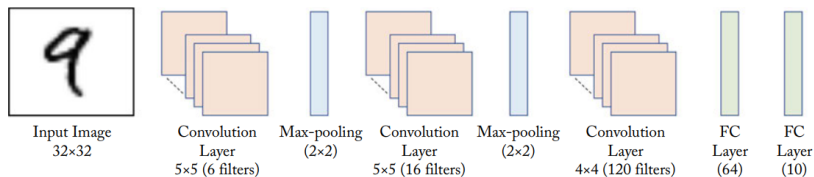
- Redes *feedforward* com muitas camadas ocultas
- Cada camada possui uma funcionalidade básica específica e produz um **mapa de características**
- Realização de operações de **convolução**
 - Extração de características por meio de um filtro
 - Parâmetros: extensão espacial, *stride*, *zero-padding*, etc
- Camada **pooling** responsável por produzir um mapa de características condensado
- Camadas completamente conectadas baseadas nos conceitos de uma rede *Multilayer Perceptron*

Redes Neurais Convolucionais

- Redes *feedforward* com muitas camadas ocultas
- Cada camada possui uma funcionalidade básica específica e produz um **mapa de características**
- Realização de operações de **convolução**
 - Extração de características por meio de um filtro
 - Parâmetros: extensão espacial, *stride*, *zero-padding*, etc
- Camada **pooling** responsável por produzir um mapa de características condensado
- Camadas **completamente conectadas** baseadas nos conceitos de uma rede *Multilayer Perceptron*

Redes Neurais Convolucionais

Figura 1: Arquitetura LeNet.



Arquiteturas Canônicas de Redes Neurais Convolucionais

- Desafio anual desenvolvido pelo projeto **ImageNet**
- *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC)
- Têm promovido algumas arquiteturas de CNNs bem sucedidas em determinados problemas
- ILSVRC 2014
 - 2º lugar: VGGNet - pequenos filtros
 - 1º lugar: GoogLeNet - *Inception Module*
- Vencedora do ILSVRC 2015: ResNet - blocos residuais

Arquiteturas Canônicas de Redes Neurais Convolucionais

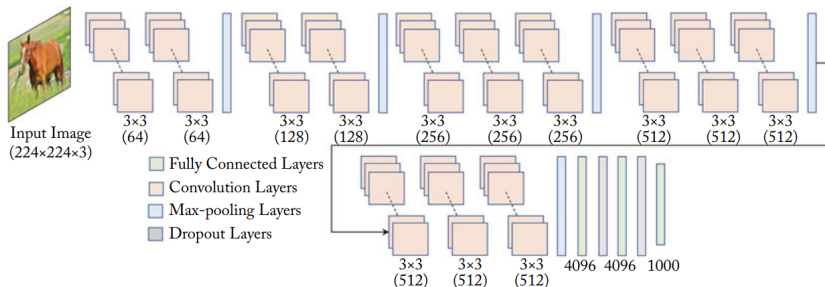
- Desafio anual desenvolvido pelo projeto **ImageNet**
- *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC)
- Têm promovido algumas arquiteturas de CNNs bem sucedidas em determinados problemas
- ILSVRC 2014
 - 2º lugar: VGGNet - pequenos filtros
 - 1º lugar: GoogLeNet - *Inception Module*
- Vencedora do ILSVRC 2015: ResNet - blocos residuais

Arquiteturas Canônicas de Redes Neurais Convolucionais

- Desafio anual desenvolvido pelo projeto **ImageNet**
- *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC)
- Têm promovido algumas arquiteturas de CNNs bem sucedidas em determinados problemas
- ILSVRC 2014
 - 2º lugar: VGGNet - pequenos filtros
 - 1º lugar: GoogLeNet - *Inception Module*
- Vencedora do ILSVRC 2015: ResNet - blocos residuais

Arquiteturas Canônicas de Redes Neurais Convolucionais

Figura 2: Arquitetura VGGNet.



Transfer Learning

- Permite utilizar uma rede pré-treinada e adaptá-la a um novo conjunto de dados
- Modelos treinados com um conjunto genérico de dados conseguem capturar características universais
- Treinamentos com conjuntos de dados bastante conhecidos
 - ImageNet
 - Places50
- Modificação na camada de saída para obter o aprendizado desejado

Transfer Learning

- Permite utilizar uma rede pré-treinada e adaptá-la a um novo conjunto de dados
- Modelos treinados com um conjunto genérico de dados conseguem capturar características universais
- Treinamentos com conjuntos de dados bastante conhecidos
 - ImageNet
 - Places50
- Modificação na camada de saída para obter o aprendizado desejado

Transfer Learning

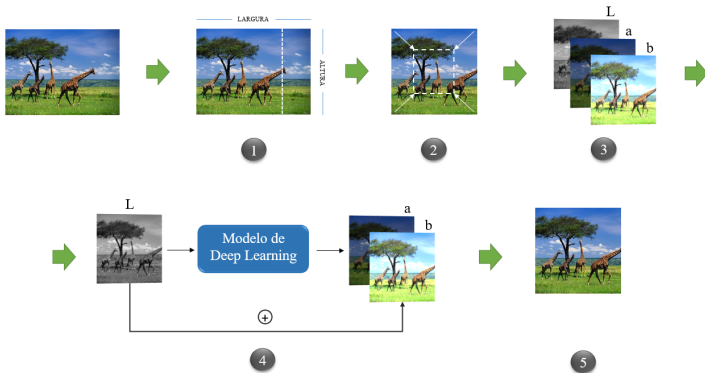
- Permite utilizar uma rede pré-treinada e adaptá-la a um novo conjunto de dados
- Modelos treinados com um conjunto genérico de dados conseguem capturar características universais
- Treinamentos com conjuntos de dados bastante conhecidos
 - ImageNet
 - Places50
- Modificação na camada de saída para obter o aprendizado desejado

Agenda - Seção 7

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta**
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências

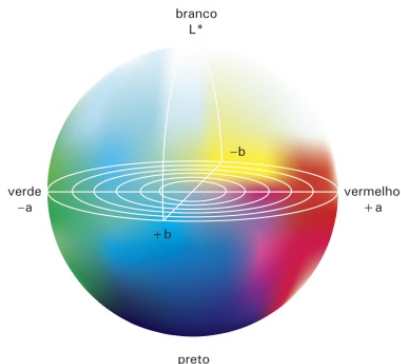
Processo de Aprendizagem de Máquina

- Problema abordado como uma **tarefa de regressão**
- Processo de Aprendizagem de Máquina:



Processo de Aprendizagem de Máquina

- Espaço de cores CIELab
 - Canal L indica luminosidade responsável pela ilustração
 - Canais a e b eixos de cromaticidade responsáveis pela coloração



Processo de Aprendizagem de Máquina

- Treinamentos sucessivos para ajuste de pesos
- Arquiteturas VGGNet e ResNet
- Diferentes parâmetros e hiperparâmetros
- Método *holdout* de validação cruzada
 - 70% dos dados para treinamento
 - 30% dos dados para testes

Métrica de Desempenho

- Métrica **Erro Quadrático Médio** (MSE), definido como:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- y_i é o valor real esperado para o exemplo i
- \hat{y}_i é o valor previsto pelo modelo para o exemplo i
- n é o número de amostras presentes no conjunto de dados

Dados Experimentais

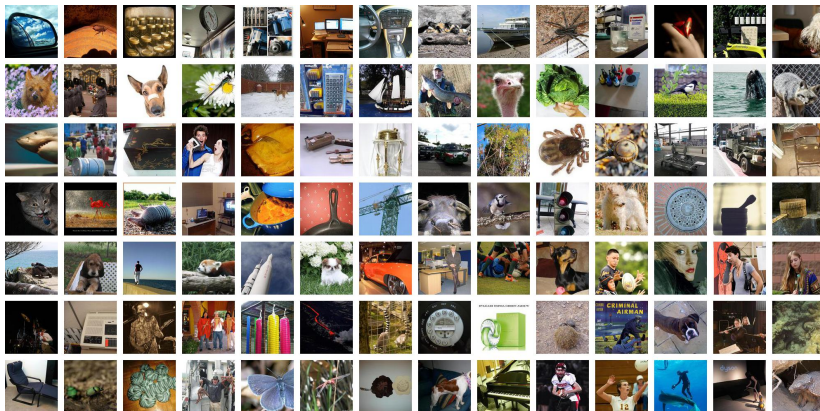
- Conjunto de validação da categoria *Fine-grained classification* do ILSVRC 2012
- Quantidade total: 50 mil imagens
- Imagens coletadas do *Flickr* e outras plataformas similares
- Quantidade de imagens favorável para o treinamento em questão
- Variedade de classes representativa ao cenário de colorização

Dados Experimentais

- Conjunto de validação da categoria *Fine-grained classification* do ILSVRC 2012
- Quantidade total: 50 mil imagens
- Imagens coletadas do *Flickr* e outras plataformas similares
- Quantidade de imagens favorável para o treinamento em questão
- Variedade de classes representativa ao cenário de colorização

Dados Experimentais

Figura 3: Visão geral do conjunto de dados.



Limpeza e Pré-Processamento da Base de Dados

- Procedimento necessário para estruturação da base de dados sem que haja sobrecarga computacional
- Redimensionamento das imagens
 - Tamanhos iguais de altura e largura
 - Dimensão de 128×128 pixels
- Selecionar imagens do espaço de cores RGB
 - Eliminar imagens com outros espaços de cores
 - Eliminar imagens em tons de cinza
- Aproximadamente 900 imagens foram descartadas
- Imagens convertidas para o espaço de cores CIELab

Limpeza e Pré-Processamento da Base de Dados

- Procedimento necessário para estruturação da base de dados sem que haja sobrecarga computacional
- Redimensionamento das imagens
 - Tamanhos iguais de altura e largura
 - Dimensão de 128×128 pixels
- Selecionar imagens do espaço de cores RGB
 - Eliminar imagens com outros espaços de cores
 - Eliminar imagens em tons de cinza
- Aproximadamente 900 imagens foram descartadas
- Imagens convertidas para o espaço de cores CIELab

Limpeza e Pré-Processamento da Base de Dados

- Procedimento necessário para estruturação da base de dados sem que haja sobrecarga computacional
- Redimensionamento das imagens
 - Tamanhos iguais de altura e largura
 - Dimensão de 128×128 pixels
- Selecionar imagens do espaço de cores RGB
 - Eliminar imagens com outros espaços de cores
 - Eliminar imagens em tons de cinza
- Aproximadamente 900 imagens foram descartadas
- Imagens convertidas para o espaço de cores CIELab

Limpeza e Pré-Processamento da Base de Dados

- Procedimento necessário para estruturação da base de dados sem que haja sobrecarga computacional
- Redimensionamento das imagens
 - Tamanhos iguais de altura e largura
 - Dimensão de 128×128 pixels
- Selecionar imagens do espaço de cores RGB
 - Eliminar imagens com outros espaços de cores
 - Eliminar imagens em tons de cinza
- Aproximadamente 900 imagens foram descartadas
- Imagens convertidas para o espaço de cores CIELab

Modelos e Parâmetros Considerados

- Abordagem *transfer learning*
- Utilização do *framework Keras*
 - Arquiteturas canônicas disponíveis
 - Pré-treinamento com base de dados do ImageNet
- Utilização das redes **VGGNet16** e **ResNet50**
- Últimas camadas substituídas por camadas completamente conectadas adaptadas ao problema
- Eventuais re-treinos para ajustes
- 100 épocas
- Demais parâmetros preservados

Modelos e Parâmetros Considerados

- Abordagem *transfer learning*
- Utilização do *framework Keras*
 - Arquiteturas canônicas disponíveis
 - Pré-treinamento com base de dados do ImageNet
- Utilização das redes **VGGNet16** e **ResNet50**
- Últimas camadas substituídas por camadas completamente conectadas adaptadas ao problema
- Eventuais re-treinos para ajustes
- 100 épocas
- Demais parâmetros preservados

Agenda - Seção 8

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais**
- 9 Considerações Parciais
- 10 Referências

Resultados Parciais

- Atividades propostas executadas conforme o cronograma
 - Estudos das tecnologias e conceitos teóricos
 - Consolidação da base de dados
 - Especificação das arquiteturas para *transfer learning*
- Treinamento de um modelo para fins de teste
 - 1% da base de dados: 500 exemplos
 - Método *holdout* de validação cruzada 70/30
 - 3 épocas e taxa de aprendizado de 10^{-4}
 - Outros parâmetros preservados
 - Métrica MSE igual 261,175
 - 6 minutos de tempo de execução

Resultados Parciais

- Atividades propostas executadas conforme o cronograma
 - Estudos das tecnologias e conceitos teóricos
 - Consolidação da base de dados
 - Especificação das arquiteturas para *transfer learning*
- Treinamento de um modelo para fins de teste
 - 1% da base de dados: 500 exemplos
 - Método *holdout* de validação cruzada 70/30
 - 3 épocas e taxa de aprendizado de 10^{-4}
 - Outros parâmetros preservados
 - Métrica MSE igual 261,175
 - 6 minutos de tempo de execução

Agenda - Seção 9

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais**
- 10 Referências

Considerações Parciais

- Colorização de imagens utilizando modelos de redes neurais convolucionais e técnicas de *Deep Learning*
- Imagens no espaço de cores CIELab
 - Canal L como entrada das redes
 - Canais a e b como saída das redes
- Rede VGGNet com 1% da base de dados e MSE de 261,175
- Continuidade: treinamentos de outros modelos
 - Base de dados completa
 - Arquiteturas VGGNet e ResNet
 - 100 épocas
 - Métrica MSE

Considerações Parciais

- Colorização de imagens utilizando modelos de redes neurais convolucionais e técnicas de *Deep Learning*
- Imagens no espaço de cores CIELab
 - Canal L como entrada das redes
 - Canais a e b como saída das redes
- Rede VGGNet com 1% da base de dados e MSE de 261,175
- Continuidade: treinamentos de outros modelos
 - Base de dados completa
 - Arquiteturas VGGNet e ResNet
 - 100 épocas
 - Métrica MSE

Considerações Parciais

- Colorização de imagens utilizando modelos de redes neurais convolucionais e técnicas de *Deep Learning*
- Imagens no espaço de cores CIELab
 - Canal L como entrada das redes
 - Canais a e b como saída das redes
- Rede VGGNet com 1% da base de dados e MSE de 261,175
- **Continuidade:** treinamentos de outros modelos
 - Base de dados completa
 - Arquiteturas VGGNet e ResNet
 - 100 épocas
 - Métrica MSE

Cronograma Parcial

	2018											
	02	03	04	05	06	07	08	09	10	11	12	
Estudo dos conceitos teóricos relacionados à <i>Machine Learning</i>	x	x	x	x								
Estudo do ferramental tecnológico para elaboração do projeto	x	x	x	x								
Consolidação da base de dados			x	x								
Especificação das arquiteturas canônicas de redes neurais convolucionais			x									
Aplicação das técnicas de <i>Transfer Learning</i> nas redes neurais convolucionais identificadas				x								
Treinamento das redes neurais convolucionais com os exemplos da base de dados				x	x	x	x					
Testes das redes e comparação de métricas de desempenho					x	x	x	x				
Análise dos resultados obtidos									x	x	x	
Escrita da proposta do trabalho	x	x	x	x	x							
Defesa da proposta do trabalho					x							
Escrita do trabalho final						x	x	x	x	x	x	
Defesa do trabalho final											x	

Agenda - Seção 10

- 1 Introdução
- 2 Objetivos
- 3 Justificativa
- 4 Metodologia
- 5 Cronograma
- 6 Fundamentação Teórica
- 7 Solução Proposta
- 8 Resultados Parciais
- 9 Considerações Parciais
- 10 Referências**

Referências

- ZHANG, R.; ISOLA, P.; EFROS, A. A. Colorful image colorization. In: *ECCV*. Bekerley, California: Cornell University Library, 2016.
- CHOLLET, F. *Deep Learning with Python*. 1. ed. Shelter Island, New York: Manning, 2017.
- KHAN, S. et al. *A Guide to Convolutional Neural Networks for Computer Vision*. 1 ed. Australia: Morgan and Claypool, 2018. v. 1. (Synthesis Lectures on Computer Vision, v. 1).
- RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, v. 115, n. 3, p. 211-252, 2015.

Colorização de Imagens com Deep Learning

Defesa do Trabalho de Conclusão de Curso I

Giovana de Lucca

Orientadora: Elloá B. Guedes

{gol.eng, elloa.uea}@gmail.com

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

21 de junho de 2018

