

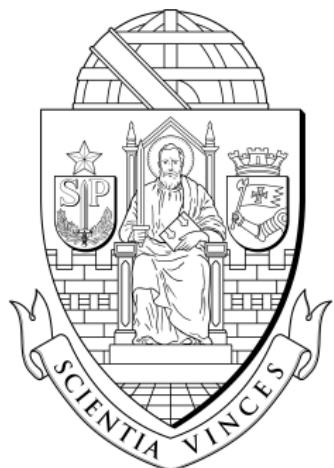
Geração de imagens artificiais e extração de características latentes aplicadas à classificação de imagens

Gabriela Salvador Thumé

Orientador: Prof. Dr. Moacir Pereira Ponti Junior

Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo

27 de março de 2015



Estrutura da Apresentação

1 Introdução

- Motivação
- Contextualização
- Hipóteses e objetivos
- Pré-processamento de imagens
- Extração de características
- Deep learning
- Máquina de Boltzmann restrita

- Desbalanceamento de classes

2 Proposta

- Metodologia
- Resultados esperados
- Atividades e cronograma

3 Resultados Preliminares

- Descrição do experimento
- Resultados
- Artigo

4 Próximos Passos

5 Referências

Estrutura da Apresentação

1 Introdução

- Motivação
- Contextualização
- Hipóteses e objetivos
- Pré-processamento de imagens
- Extração de características
- Deep learning
- Máquina de Boltzmann

restrita

- Desbalanceamento de classes

2 Proposta

3 Resultados Preliminares

4 Próximos Passos

5 Referências



Introdução

- ▶ Classificação de imagens;
- ▶ Algoritmos de aprendizado de máquina;
- ▶ Generalização para classificar novos exemplos;
- ▶ Conjuntos de características que dificultam a diferenciação entre as classes;
- ▶ Encontrar as características que melhor discriminam as classes.

Motivação

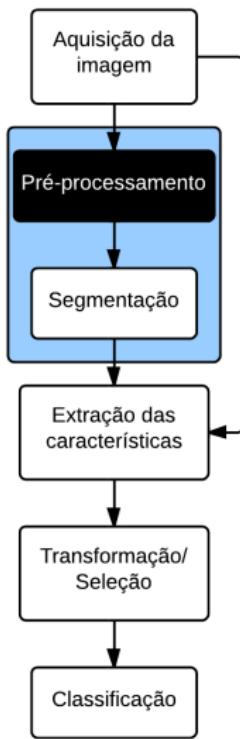


Figura: Etapas Canônicas.

Motivação

- ▶ Maior esforço ao operar no espaço de características já obtidas;
 - ▶ Transformações do espaço ou sistemas de classificação complexos para lidar com as deficiências das características extraídas;
 - ▶ Características que podem ser exploradas além dos métodos clássicos;

Motivação - Características Latentes

- ▶ Métodos de processamento e preparação de imagens antes da extração dessas características;
 - ▶ Podem revelar características latentes, não visíveis nas imagens originais;
 - ▶ Objetivo: encontrar tais características que possam melhor descrever certas classes do problema.

Motivação - Características Latentes

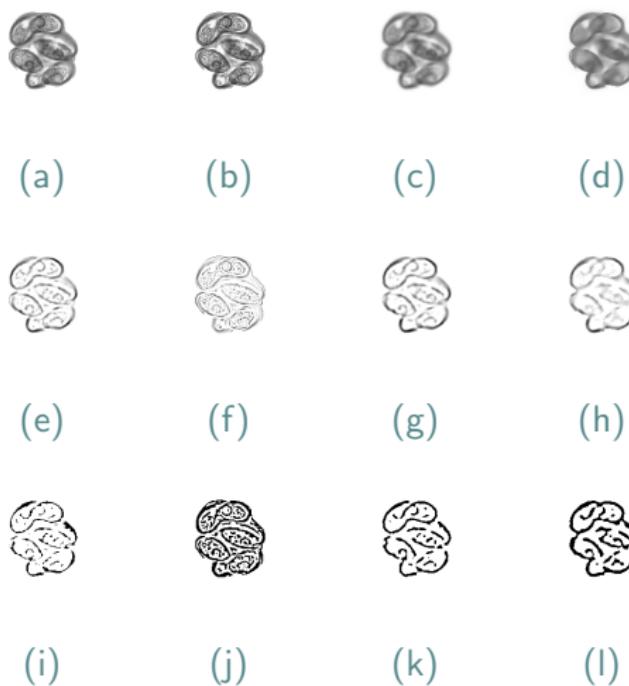


Figura: Características latentes de algas verdes.

Desbalanceamento de classes

- ▶ Diferença entre o número de exemplos disponíveis;
- ▶ Obstáculo;
- ▶ Métodos de transformação do espaço de características e de classificação assumem que as classes da base estão平衡ados;
- ▶ Objetivo: geração de imagens artificiais.

Proposta da Pesquisa

Melhorar a classificação de imagens, utilizando métodos de processamento com foco na **extração de características latentes** e no **rebalanceamento de classes**.

Contextualização

- ▶ Grupo de pesquisa em Visualização, Imagens e Computação Gráfica (VICG)
 - ▶ Visualização de informação com projeções multidimensionais e árvores;
 - ▶ Extração e classificação de imagens.
- ▶ (Rocha et al., 2010) 98% de acurácia após aquisição, pré-processamento e segmentação;
- ▶ (Kanan e Cottrell, 2012) Quantização pode impactar a classificação;
- ▶ (Ponti et al., 2014) Quantização permite obter vetores de características mais compactos e com maior capacidade de discriminação entre classes;
- ▶ Continuidade a essa pesquisa: analisar deep learning (aprendem quais operações geram as características).

Hipóteses

Geração de imagens artificiais

- ▶ Balancear as classes;
- ▶ Melhorar a acurácia de algoritmos de classificação, versus geração de exemplos artificiais no espaço de atributos

Métodos de pré-processamento

- ▶ Extração de características latentes que aumentem a variância entre as classes, sem aumentar a variância intra-classe;
- ▶ Melhorar a classificação.

Objetivos

Geral

Investigar os métodos de pré-processamento para preparar uma coleção de imagens para a extração de características.

Obter características latentes e balancear o número de instâncias de diferentes classes.

Objetivos

Específicos

- ▶ Analisar:
 - ▶ impacto de métodos canônicos na classificação;
 - ▶ aprendizado de bases bem discriminadas por CNN.
- ▶ Tornar as características latentes visíveis. Aumentar a variância entre as classes;
- ▶ Gerar imagens artificiais a partir das imagens pertencentes às classes minoritárias, compensando o desbalanceamento.
 - ▶ Resultados preliminares
 - ▶ Matriz de características aprendida por RBM para verificar se as imagens geradas são relevantes para o aprendizado e escolher as imagens originais.

Contextualização

- ▶ Métodos canônicos de pré-processamento
- ▶ Redes de convolução e máquinas de Boltzmann restritas;
- ▶ Desbalanceamento de classes.

Pré-processamento de Imagens

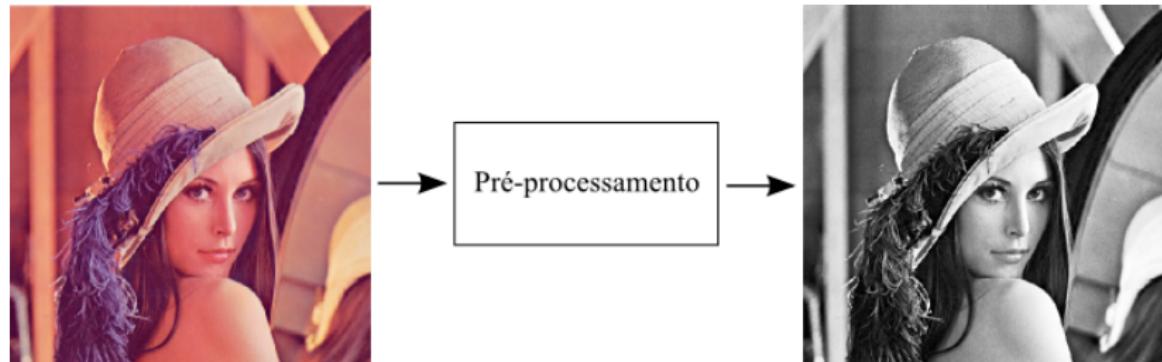


Figura: Borramento, realce e de equalização de histograma.

Pré-processamento de Imagens - Convolução

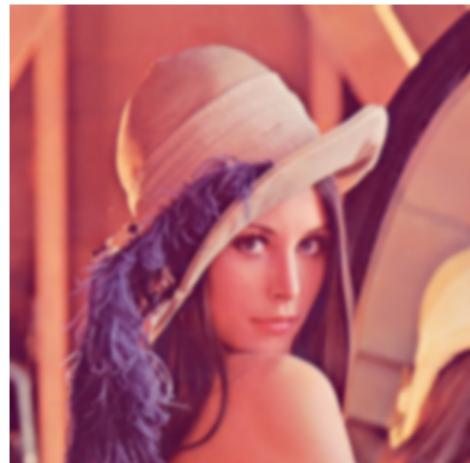
- ▶ Percorre a imagem com um filtro espacial rotacionado em 180°;
- ▶ Cria cada novo pixel com as mesmas coordenadas do centro da vizinhança contendo o valor resultante da filtragem.

$$\text{Mapa de características} = \text{imagem de entrada} * \text{filtro} \quad (1)$$

Pré-processamento de Imagens - Convolução



(a) Original



(b) Filtragem Gaussiana

Pré-processamento de Imagens - Realce

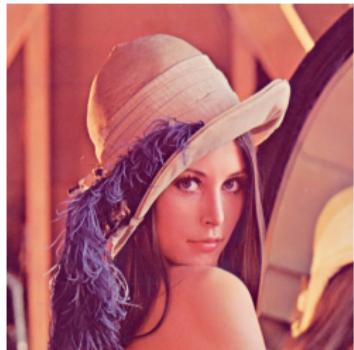


(a) Original

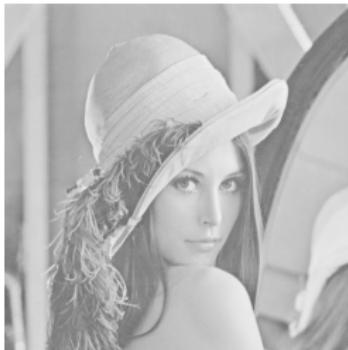


(b) *Unsharp masking*

Pré-processamento de Imagens - Quantização



(a) Original



(b) Intensidade



(c) MSB

Extração de Características

- ▶ Descrever as informações visuais relevantes em um vetor de características;
- ▶ Entrada para o classificador de padrões;
- ▶ Salientar as diferenças entre imagens de classes distintas e suavizar possíveis diferenças de imagens da mesma classe.

Textura: suavidade, aspereza e uniformidade. Ex. entropia;

Forma: características externas. Ex. da curvatura;

Cor: distribuição espacial de cores na imagem. Ex. histograma.

Extração de Características

- GCH** Histograma global de cor. N dimensões, sendo N o número de intensidades.
- CCV** Vetor de coerência de cor. Classifica os pixels da imagem em coerentes e incoerentes, computa os respectivos histogramas de acordo com um *threshold* e os concatena. 2ND.
- BIC** Classificação de pixels de borda e interior. Mesma cor que seus vizinhos, é pixel de interior. Computa dois histogramas 2ND.
- ACC** Auto-correlograma de cor: captura a correlação espacial entre cores idênticas. Probabilidade de encontrar dois pixels com a mesma cor, a uma distância d um do outro. Para 1, 3, 5 e 7; $4N$ características.
- Haralick** entropia, homogeneidade, contraste, correlação, probabilidade máxima e uniformidade. 6D.

Contextualização - Redes Neurais

- ▶ Redes neurais de convolução (CNN)
 - ▶ Camadas de neurônios para aprender as melhores características que diferenciam as classes;
 - ▶ Aprendizado: ajuste dos parâmetros entre a saída esperada e a produzida.
- ▶ Máquinas de Boltzmann restritas (RBM)
 - ▶ Aprende a representação das imagens de entrada;
 - ▶ Definir quais imagens são relevantes para o aprendizado.

Aprendem versões processadas das imagens de entrada, indicando que os filtros aprendidos são os que melhor diferenciam as classes.

Deep Learning

- ▶ Reconhecimento humano de novos padrões está relacionada com a capacidade de generalização, através de hierarquias.
- ▶ Simular o funcionamento do cérebro humano por meio de múltiplas camadas.
- ▶ Imagens como treinamento para automaticamente inferir quais são as regras para o reconhecimento;
- ▶ Conhecimento através da experiência: ao tentar uma solução e errar, aprendem e podem tentar novamente. Neurônios de entrada, ocultos e de saída. Aprender = encontrar os pesos para exibir o comportamento esperado
- ▶ Redes neurais profundas possuem uma estrutura de muitas camadas – duas ou mais ocultas
- ▶ Subdividem em problemas mais simples de serem resolvidos, podendo chegar ao nível de pixels isolados.

Deep Learning

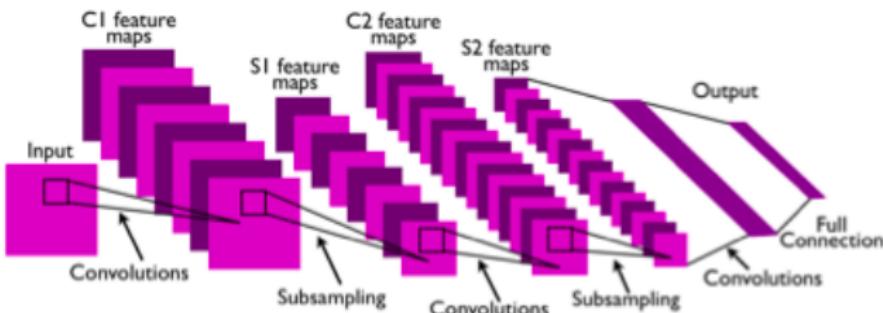


Fig. 1. A typical ConvNet architecture with two feature stages

Figura

Máquina de Boltzmann restrita

- ▶ Rede neural estocástica que treina um modelo a partir dos vetores de entrada;
- ▶ Os pixels correspondem às unidades visíveis e os detectores de características às unidades ocultas;
- ▶ Começa em um estado aleatório e atualiza seus pesos até encontrar uma distribuição que esteja em equilíbrio;
- ▶ Identificar se uma imagem é relevante para o aprendizado.

Desbalanceamento de classes

- ▶ Número desbalanceado de exemplos. Majoritárias x minoritárias.
- ▶ Abordagens
 - ▶ Reamostrar os dados;
 - ▶ Adicionar funções de custo na classificação.

Desbalanceamento de classes - Sobreamostragem

- ▶ Aumentar, utilizando alguma estratégia, o número de elementos do conjunto;
- ▶ Replicação não reporta melhorias;

SMOTE

- ▶ A diferença entre o vetor de características de um elemento e do seu vizinho mais próximo é multiplicada por um número entre zero e um;
- ▶ Valor adicionado ao vetor original, criando um novo elemento entre os dois vetores originais;
- ▶ Força uma região de decisão maior e mais geral para serem aprendidas como exemplos da classe minoritária;

Contextualização - SMOTE

- ▶ Rebalancear classes ao gerar artificialmente novos elementos, ao invés de apenas replicá-los;
- ▶ Sobre os vetores de características previamente extraídos;
- ▶ (Chawla et al., 2002) **Diferentes estratégias para criar exemplos sintéticos podem melhorar a performance da classificação.**

Desbalanceamento de classes - Subamostragem

- ▶ Diminuir, utilizando alguma estratégia, o número de elementos do conjunto;
- ▶ Eliminar elementos da classe majoritária que estão distantes da fronteira de decisão;
- ▶ Normalmente apresentam resultados piores;
- ▶ Pode remover informações essenciais dos dados originais;
- ▶ Não há melhor para todos os cenários.

Estrutura da Apresentação

- Atividades e cronograma

1 Introdução

2 Proposta

- Metodologia
- Resultados esperados

3 Resultados Preliminares

4 Próximos Passos

5 Referências

Metodologia

Metodologia

Características latentes

- ▶ Métodos de pré-processamento de imagens com o objetivo de obter imagens processadas que sejam melhor caracterizadas para a etapa de classificação.
- ▶ O enfoque em realçar determinadas características, utilizando diversos algoritmos sobre as imagens originais.

Metodologia

Redes neurais

- ▶ Estado da arte da classificação, reconhecimento e localização de objetos;
- ▶ Identificar as características relevantes em imagens;
- ▶ RBM são redes neurais mais simples, convenientes para a verificação da relevância de uma imagem para o aprendizado.

Metodologia

Desbalanceamento de classes

Pesquisar métodos que visam equilibrar o número de imagens em cada classe.

Metodologia

Descritores

Investigar métodos capazes de descrever as propriedades das imagens.

Metodologia

Classificador de padrões

- ▶ Naive Bayes, Optimum-Path Forest, KNN (*K-Nearest Neighbor*) e Support Vector Machines.
- ▶ Não é o foco deste estudo, após experimentos um destes será escolhido.

Metodologia

Implementação

- ▶ Biblioteca OpenCV;
- ▶ Linguagens de programação C++ e Python;
- ▶ Código disponível em <https://bitbucket.org/moacirponti/imagefeatureextraction/overview>.

Metodologia

Bases de imagens

- ▶ Objetivos com viés genérico;
- ▶ Diversas coleções de imagens com o objetivo de estabelecer ou refutar as hipóteses levantadas;
- ▶ Resultados preliminares com a base de imagens COREL^a;

^aDisponível em <http://wang.ist.psu.edu/docs/related/>



Figura: Base de imagens COREL-1000

Metodologia

Bases de imagens

- ▶ Cor: COREL-1000;
- ▶ Textura: Describable Textures Dataset^a (DTD);
- ▶ Forma: Leafsnap^b;

^a<http://www.robots.ox.ac.uk/~vgg/data/dtd/>

^b<http://leafsnap.com/dataset/>



Figura: Base de imagens Leafsnap.



Figura: Base de imagens DTD.

Metodologia

Experimentos

- ▶ Experimentos direcionados a explorar as etapas de pré-processamento, para melhorar a acurácia da classificação de bases de imagens;
- ▶ Entrada: imagens originais provenientes de diversas coleções disponíveis na literatura;
- ▶ Resultado: medidas estatísticas da classificação após a alteração das imagens com os métodos de realce de características relevantes.

Metodologia

Análise dos dados

- ▶ Comparar a classificação das imagens originais com aquelas tratadas pelo método proposto;
- ▶ Ainda, o método de rebalanceamento de classes será comparado com técnicas disponíveis na literatura, como o SMOTE.

Metodologia

Avaliação - F1Score

Efetividade da classificação em cenários desbalanceados.

A precisão (medida da exatidão): dos exemplos classificados como positivos, quantos realmente são.

$$P = \frac{VP}{VP + FP}, \quad (2)$$

A revocação é a medida de completude: quantos exemplos positivos foram corretamente classificados como tal.

$$R = \frac{VP}{VP + FN} \quad (3)$$

$$F1 = 2 \frac{PR}{P + R} \quad (4)$$

Metodologia

Avaliação - Friedman

- ▶ Determinar se há diferença significante em uma amostra de resultados gerados;
- ▶ As performances dos algoritmos são analisados e um *rank* é atribuído para cada conjunto de dados;
- ▶ A hipótese nula a ser testada é que não há diferença estatística relevante entre as observações;
- ▶ P-valor indica o quão estatisticamente significante o resultado é: quanto menor o seu valor, maior a evidência contra a hipótese nula (limiar de 0,05).

Avaliação - Mahalanobis

Distância entre a média das classes e a variância dentro das classes.

Baseia na correlação entre as variáveis e pode ser definida por

$$D_m(x_i) = \sqrt{(x_i - \mu) C^{-1} (x_i - \mu)^T},$$

onde x_i é um vetor de valores, μ a média e C a matriz de covariância.

Resultados Esperados

- ▶ Melhorar a classificação, validando-a através do cálculo da medida-F1.
- ▶ A análise das características aprendidas por uma rede neural de convolução será realizada ao executar o treinamento com bases específicas que destaqueem propriedades como cor, textura e forma.
- ▶ Os resultados serão obtidos a partir da escolha de quais imagens adicionam informação ao conjunto de treino. As redes RBM serão utilizadas para este fim.
- ▶ Bases naturalmente não balanceadas serão testadas e seus resultados avaliados.

Atividades e Cronograma

Tabela: Duração de cada atividade a partir de 24/02/2014.

Atividade	2014		2015		2016
	1º Sem.	2º Sem.	1º Sem.	2º Sem.	1º Sem.
1					
2					
3					
4					
5					
6					

Estrutura da Apresentação

1 Introdução

- Descrição do experimento
- Resultados
- Artigo

2 Proposta

4 Próximos Passos

3 Resultados Preliminares

5 Referências

Resultados Preliminares



Figura: Geração de imagens artificiais para o rebalanceamento de classes por meio de: borramento, adição de ruído, mistura e combinação.

Descrição do Experimento

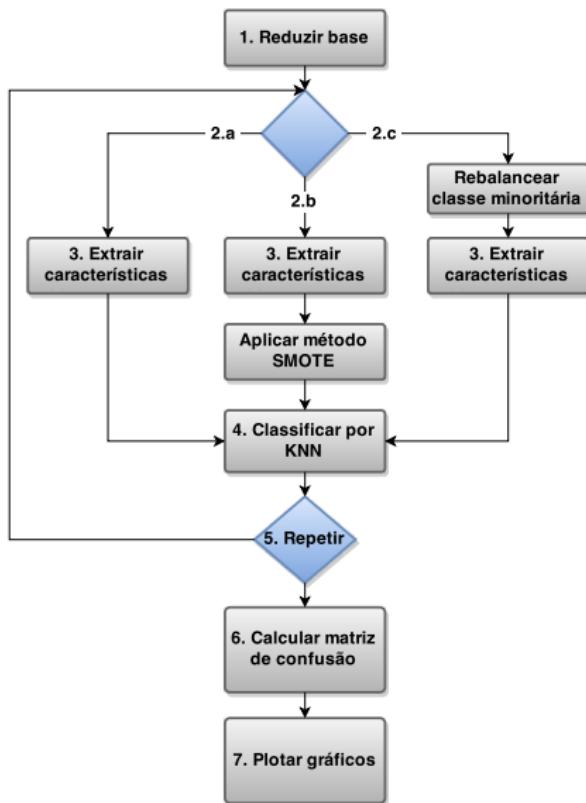


Figura: Fluxo dos resultados preliminares.

Descrição do Experimento

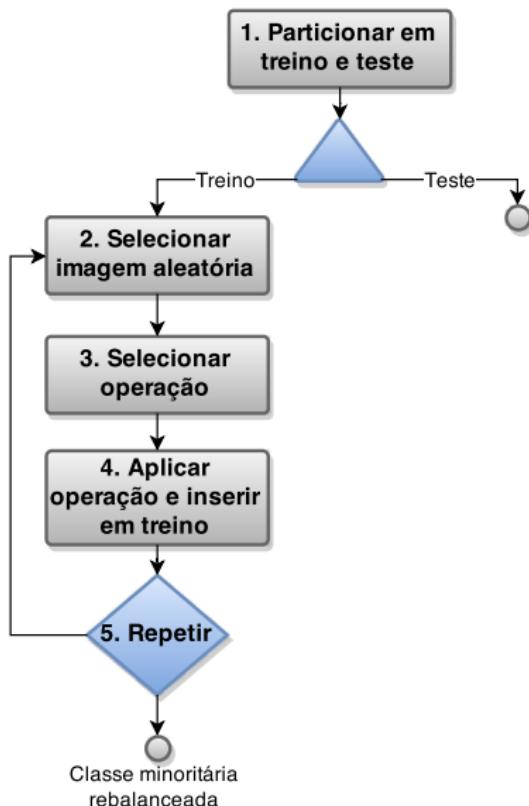


Figura: Fluxo da geração artificial.

Resultados Preliminares

Classes com maior dificuldade de diferenciação, havendo alta taxa de sobreposição de intensidades de cores e texturas.



Figura: Classes “praia” e “montanha” da base de imagens COREL-1000.

Resultados Preliminares - Melhor

- ▶ Ganho estatístico da medida-F, quando comparado à geração de exemplos artificiais no espaço de atributos;
- ▶ Descriptor ACC, conversão MSB e operação de pré-processamento combinação.

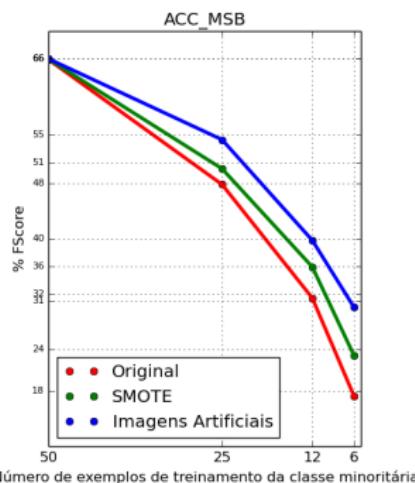


Figura: Resultado obtido com a operação de combinação.

Resultados Preliminares - Pior

- ▶ Adição de ruído, extração com CCV e a quantização por MSB;
- ▶ Outros casos que não obtiveram o resultado esperado envolveram as operações de borramento e de *unsharp masking*.

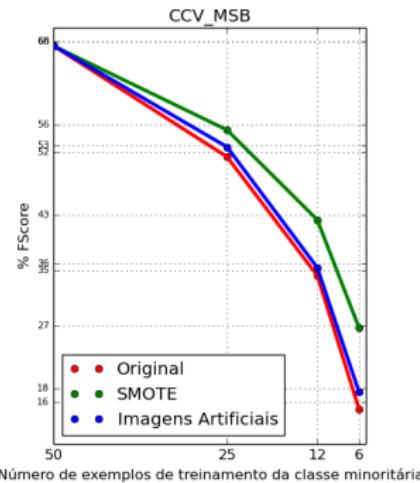


Figura: Piores resultados, obtidos com a adição de ruído.

Resultados Preliminares

- ▶ Melhores operações: todas, apenas mistura e apenas composição;
- ▶ Piores: apenas borramento, ruído ou *unsharp masking*;
- ▶ Melhor extrator: ACC;
- ▶ Piores: CCV e GCH;

Resultados Preliminares

- ▶ Teste de Friedman para todas as execuções das melhores operações.
- ▶ P-valor = $4.24E^{-11}$. Hipótese nula rejeitada;

Tabela: Posição média dos algoritmos utilizando Friedman

Algoritmos	Posição
Original	3.0000
Smote	1.6136
Artificial	1.3863

- ▶ Em algumas execuções: geração artificial (1), SMOTE (2) e imagens originais (3).

Resultados Preliminares

- ▶ Replicação: SRS - *Simple Random Sampling*;
- ▶ Não adiciona informações novas para o aprendizado.

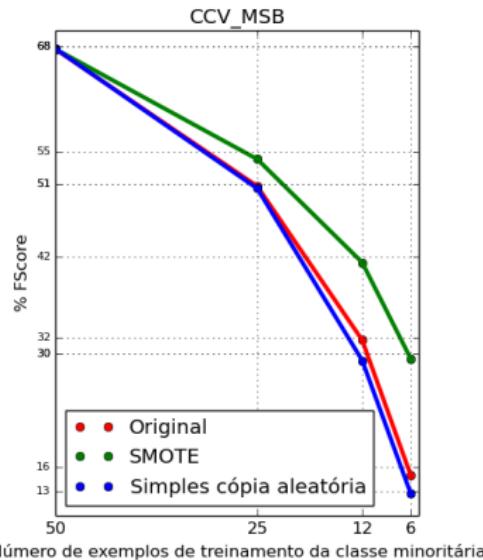


Figura: Simples replicação de exemplos sem realizar nenhuma operação

Resultados Preliminares

Adicionar aqui os novos resultados

Artigo na Neurocomputing

Ponti, M.; Nazaré, T; Thumé, G. Image quantization as a dimensionality reduction procedure in color and texture feature extraction, submitted to Neurocomputing, 2014.

Estrutura da Apresentação

1 Introdução

2 Proposta

3 Resultados Preliminares

4 Próximos Passos

5 Referências

Próximos Passos

- ▶ Analisar as características latentes que as redes neurais de convolução conseguem extrair
 - ▶ Bases bem discriminadas quanto às propriedades de textura, cor e forma.
- ▶ Analisar a memória associativa de uma rede de Boltzmann
 - ▶ As imagens geradas foram adicionadas no conjunto de treino sem verificação da sua relevância;
 - ▶ Escolha para qual imagem original utilizar, ao invés do método aleatório utilizado nos resultados preliminares.

Próximos Passos

- ▶ Desbalancear as 10 classes da COREL-1000 em escada;
- ▶ Utilizar todas as classes com apenas 1 desbalanceada;
- ▶ Analisar o treinamento da CNN para as 3 bases bem discriminadas.

Agradecimentos?

Estrutura da Apresentação

1 Introdução

2 Proposta

3 Resultados Preliminares

4 Próximos Passos

5 Referências

Referências

Geração de imagens artificiais e extração de características latentes aplicadas à classificação de imagens

Gabriela Salvador Thumé

Orientador: Prof. Dr. Moacir Pereira Ponti Junior

Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo

27 de março de 2015

