

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
- Hipóteses e Objetivos

2 Contextualização

- Pré-processamento
- Redes de Convolução
- Desbalanceamento de classes
- Máquina de Boltzmann restrita
- Extração de características

3 Metodologia

4 Resultados Preliminares

5 Próximos Passos

6 Atividades e cronograma

Introdução

- ▶ Classificação de imagens;
- ▶ Extração de características;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização permite classificar novos exemplos;
- ▶ Características que não são suficientes para a diferenciação entre as classes;
- ▶ Encontrar as características que melhor discriminam as classes.

Introdução

- ▶ Classificação de imagens;
- ▶ Extração de características;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização permite classificar novos exemplos;
- ▶ Características que não são suficientes para a diferenciação entre as classes;
- ▶ Encontrar as características que melhor discriminam as classes.

Introdução

- ▶ Classificação de imagens;
- ▶ Extração de características;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização permite classificar novos exemplos;
- ▶ Características que não são suficientes para a diferenciação entre as classes;
- ▶ Encontrar as características que melhor discriminam as classes.

Introdução

- ▶ Classificação de imagens;
- ▶ Extração de características;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização permite classificar novos exemplos;
- ▶ Características que não são suficientes para a diferenciação entre as classes;
- ▶ Encontrar as características que melhor discriminam as classes.

Motivação

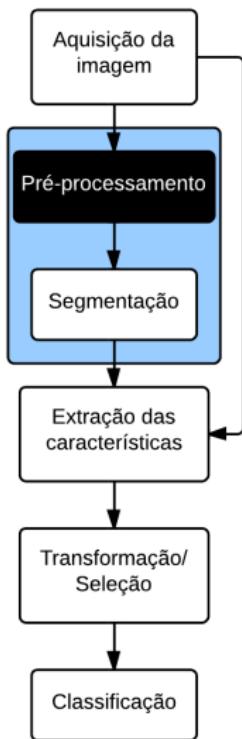


Figura: Etapas canônicas do reconhecimento de padrões.

Motivação

- ▶ Maior esforço ao operar no espaço de características já obtidas;
- ▶ Transformações do espaço ou sistemas complexos de classificação para lidar com as deficiências das características extraídas;
- ▶ Características que podem ser exploradas além dos métodos clássicos;
- ▶ Investigar métodos de processamento e preparação de imagens antes da extração.

Motivação - Características Latentes

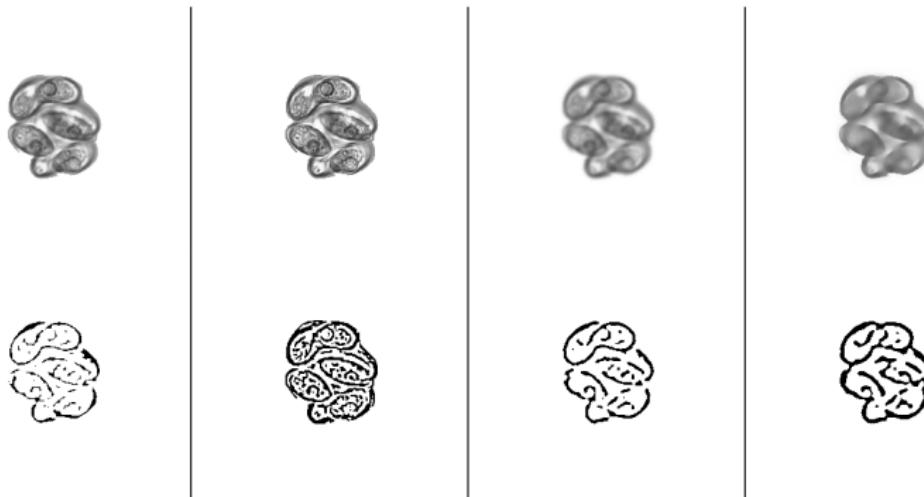


Figura: Características latentes de algas verdes.

Motivação - Características Latentes

- ▶ Justificado o uso de métodos de processamento e preparação de imagens antes da extração;
- ▶ Podem revelar características latentes, não visíveis nas imagens originais;
- ▶ Foco: *revelar características que possam melhor descrever certas classes, utilizando algoritmos sobre as imagens originais.*

Motivação

- ▶ 98% de acurácia após pré-processamento e segmentação (Rocha et al., 2010);
- ▶ Quantização pode impactar a classificação (Kanan e Cottrell, 2012);
- ▶ Quantização permite obter vetores de características mais compactos e com maior capacidade de discriminação entre classes (Ponti et al., 2014);

Continuação.

Motivação - Desbalanceamento de classes

- ▶ Diferença entre o número de exemplos disponíveis;
- ▶ Imagens representam eventos importantes mas menos frequentes;
- ▶ Obstáculo, métodos de transformação do espaço e de classificação assumem que a base está balanceada;
- ▶ Foco: *geração de imagens artificiais a partir do processamento de características das imagens da classe minoritária.*



Figura: Imagem artificialmente gerada.

Proposta da Pesquisa

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

Hipóteses

Métodos de pré-processamento

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

Geração de imagens artificiais

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

Objetivo Geral

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

Espera-se obter características latentes e balancear o número de instâncias de diferentes classes.

Objetivos Específicos

- ▶ Analisar:
 - ▶ impacto de métodos canônicos na classificação;
 - ▶ aprendizado de bases comportadas por redes de convolução.
- ▶ Tornar as características latentes visíveis;
- ▶ Gerar imagens artificiais.
 - ▶ Resultados preliminares;
 - ▶ Matriz de características aprendida por uma máquina de Boltzmann restrita para verificar a relevância das imagens geradas e as imagens originais.

Objetivos Específicos

- ▶ Analisar:
 - ▶ impacto de métodos canônicos na classificação;
 - ▶ aprendizado de bases comportadas por redes de convolução.
- ▶ Tornar as características latentes visíveis;
- ▶ Gerar imagens artificiais.
 - ▶ Resultados preliminares;
 - ▶ Matriz de características aprendida por uma máquina de Boltzmann restrita para verificar a relevância das imagens geradas e as imagens originais.

Proposta

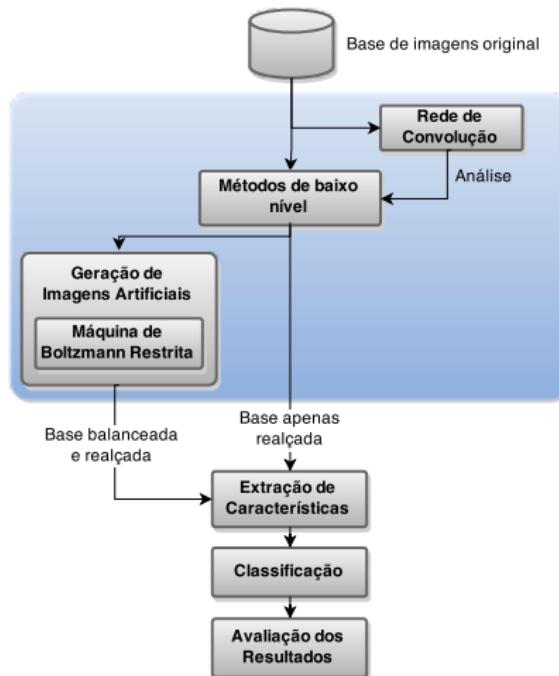


Figura: Estrutura geral desta pesquisa.

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses e Objetivos

2 Contextualização

- Pré-processamento
- Redes de Convolução
- Desbalanceamento de classes
- Máquina de Boltzmann restrita
- Extração de características

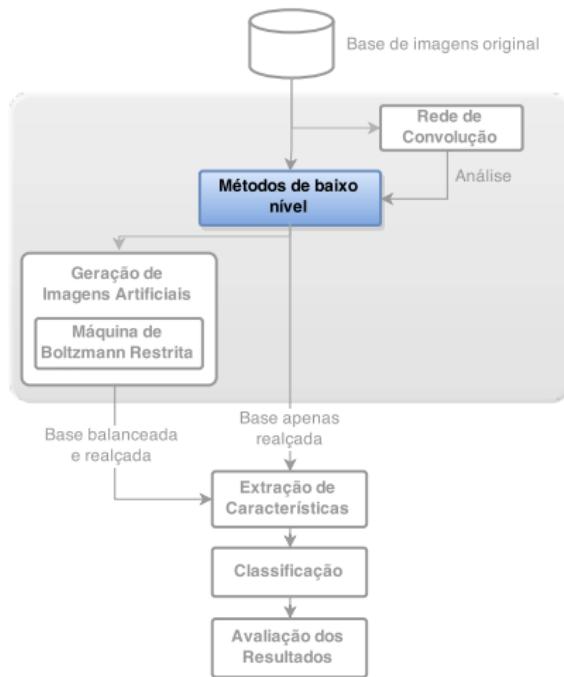
3 Metodologia

4 Resultados Preliminares

5 Próximos Passos

6 Atividades e cronograma

Pré-processamento de Imagens



Pré-processamento de Imagens

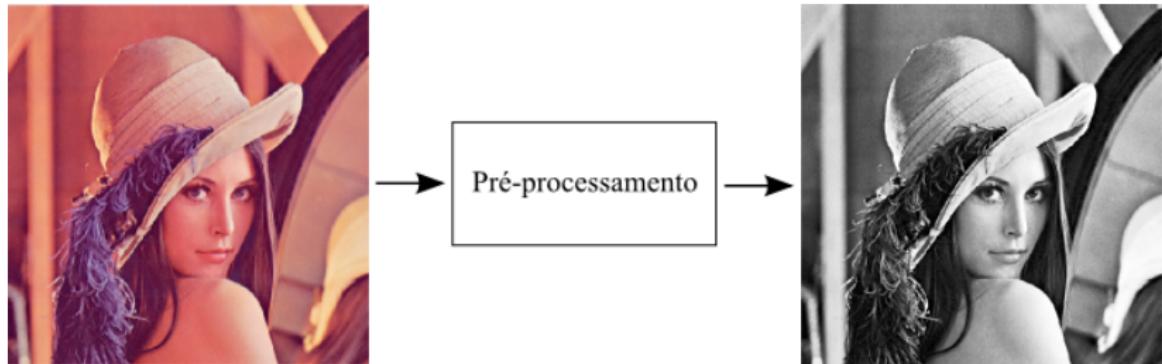


Figura: Conversão em escala de cinza, borramento, realce e de equalização de histograma.

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.

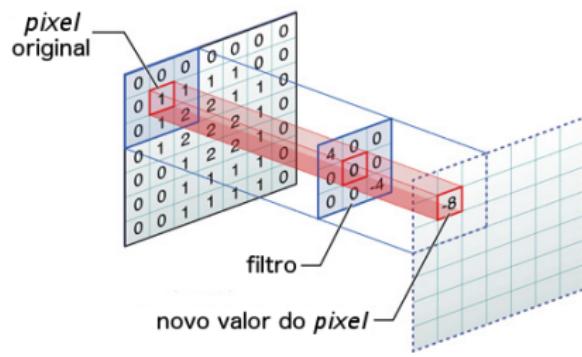
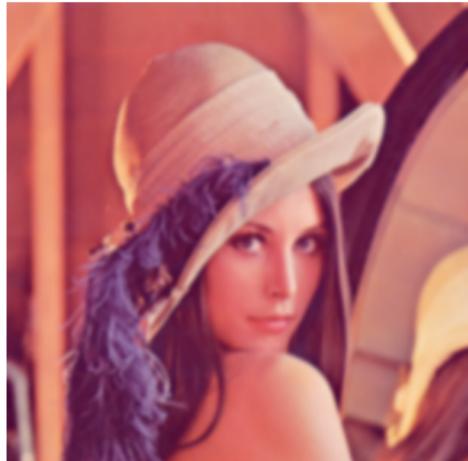


Figura: Convolução com filtro previamente rotacionado.

Pré-processamento de Imagens - Convolução



(a) Original



(b) Filtragem Gaussiana

Pré-processamento de Imagens - Realce



(a) Original

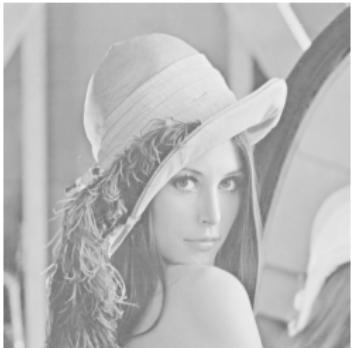


(b) *Unsharp masking*

Pré-processamento de Imagens - Quantização



(a) Original

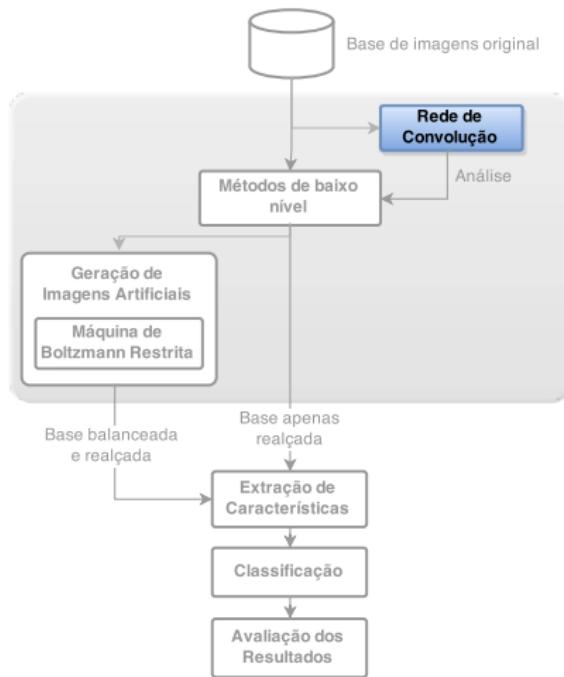


(b) Intensidade



(c) MSB

Redes de Convolução



Redes Neurais

- ▶ Imagens como treinamento para inferir as regras para a classificação;
- ▶ Conhecimento através da experiência: ao tentar uma solução e errar, aprendem e podem tentar novamente;
- ▶ Aprendizado: ajuste dos pesos entre a saída esperada e a produzida.

Redes de Convolução - Deep Learning

- ▶ Reconhecimento humano de novos padrões - capacidade de generalização (hierarquias);
- ▶ Simular o funcionamento do cérebro humano por meio de camadas;
- ▶ Redes profundas possuem duas ou mais camadas ocultas;
- ▶ Subdividem em representações mais simples;
- ▶ Representam o estado da arte em visão computacional
 - ▶ Há falta de estudo sobre suas propriedades;

Redes de Convolução

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

- ▶ Multiplicações de um filtro espacial pela imagem de entrada, resultando em um *mapa de características ativadas*;
- ▶ Vetor de parâmetros capaz de aprender;
- ▶ Um filtro é capaz de extrair apenas um tipo de característica;
- ▶ Camada de convolução = muitas convoluções em paralelo.

Redes de Convolução

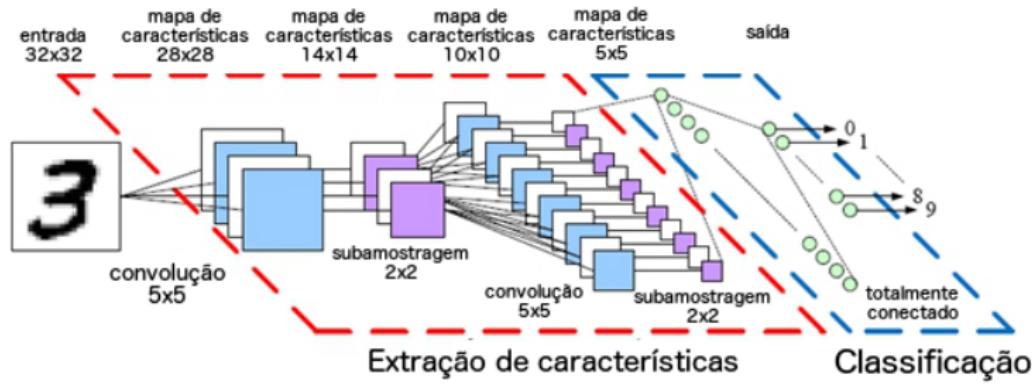


Figura: Arquitetura de uma CNN. Fonte: <http://parse.ele.tue.nl/education/cluster2>

Redes de Convolução

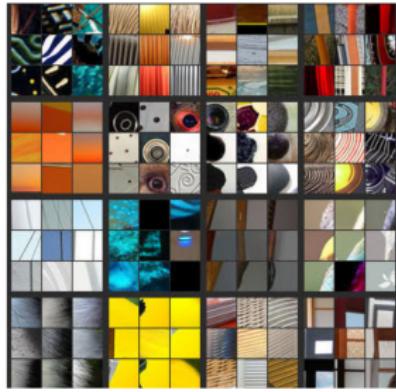
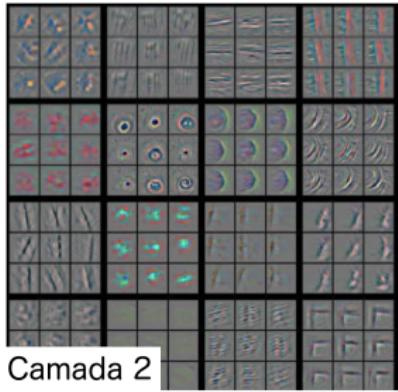
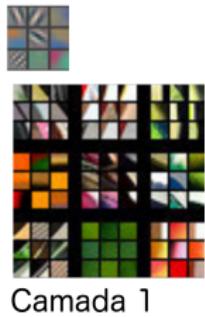


Figura: Primeira: ativar características de borda. Segunda camada: formas simples e texturas similares. Fonte: [21]

Redes de Convolução

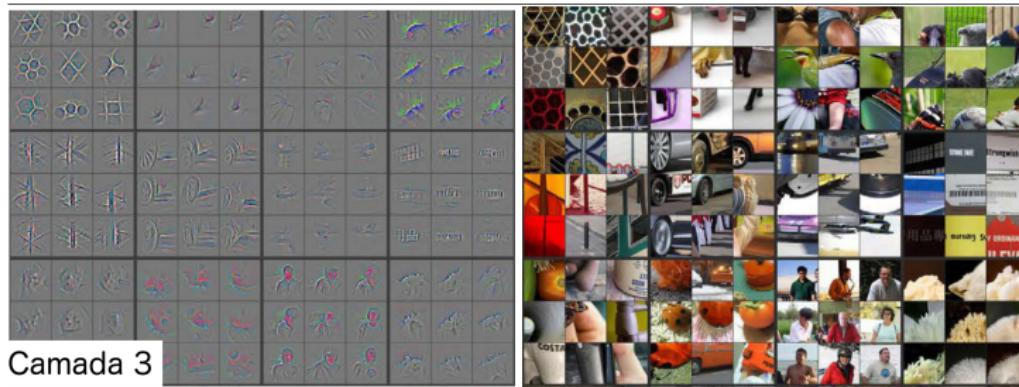


Figura: Terceira camada. Fonte: [21]

Redes de Convolução



Figura: Quarta e quinta camada. Complexos pedaços das imagens ativados. Fonte: [21]

Redes de Convolução

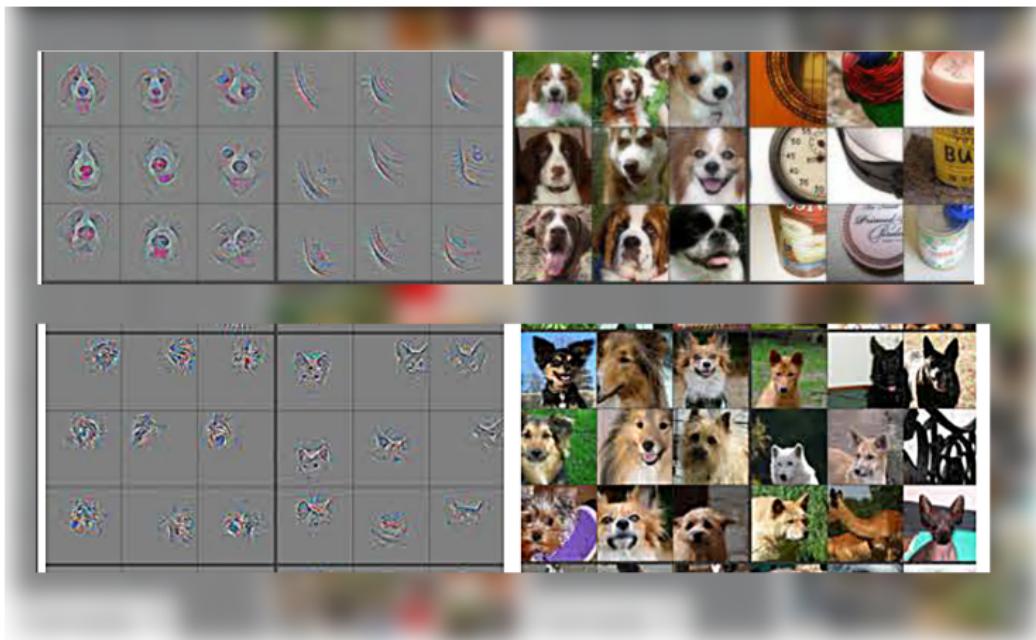
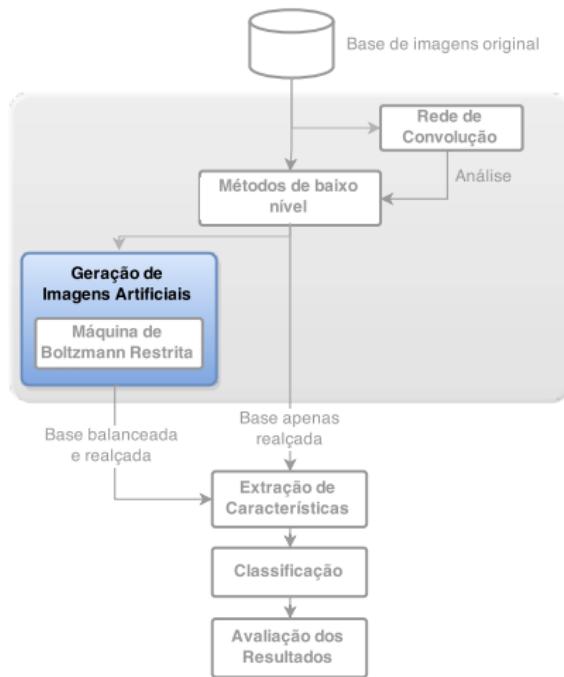


Figura: Quarta e quinta camada. Insere abstração e invariância.
Estruturas discriminantes. Fonte: [21]

Desbalanceamento de classes



Desbalanceamento de classes

- ▶ Número desbalanceado de exemplos. Majoritárias x minoritárias.
- ▶ Abordagens:
 - ▶ *Modificar métodos de aprendizagem:* adicionar funções de custo na classificação;
 - ▶ *Pré-processamento ao reamostrar os dados:*
 - ▶ Aumentar a minoritária;
 - ▶ Diminuir a majoritária.

Desbalanceamento de classes - Subamostragem

- ▶ Diminuir o número de elementos do conjunto;
- ▶ Podem remover informações essenciais dos dados originais;
- ▶ Eliminar elementos distantes da fronteira de decisão;
- ▶ Normalmente apresentam resultados piores.

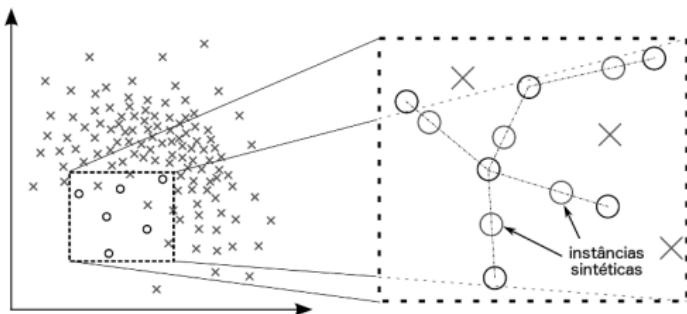
Desbalanceamento de classes - Sobreamostragem

- ▶ Aumentar o número de elementos;

SMOTE

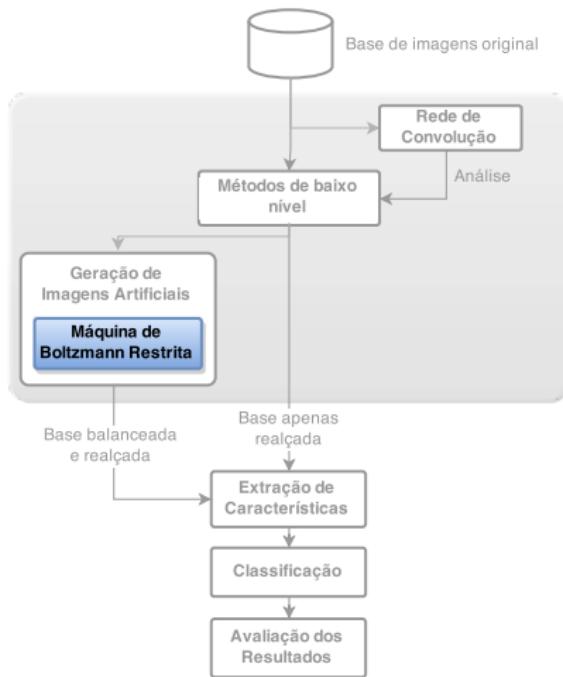
- ▶ Multiplica a diferença entre o vetor de características de um elemento e do seu vizinho mais próximo por um número $0 \leq x \leq 1$;
- ▶ Adiciona ao vetor original, criando um novo elemento entre os dois vetores originais;
- ▶ Aprendido como exemplo da classe minoritária;
- ▶ Força uma região de decisão maior e mais geral;

Desbalanceamento de classes - SMOTE



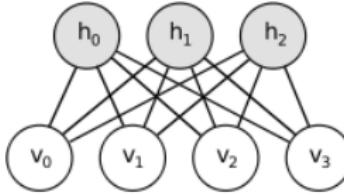
- ▶ Rebalancear ao gerar novos elementos, ao invés de replicá-los;
- ▶ Sobre os vetores de características previamente extraídos;
- ▶ (Chawla et al., 2002) **Diferentes estratégias para criar exemplos sintéticos podem melhorar a performance da classificação;**
- ▶ Utilizado para comparação.

Máquina de Boltzmann Restrita



Máquina de Boltzmann restrita

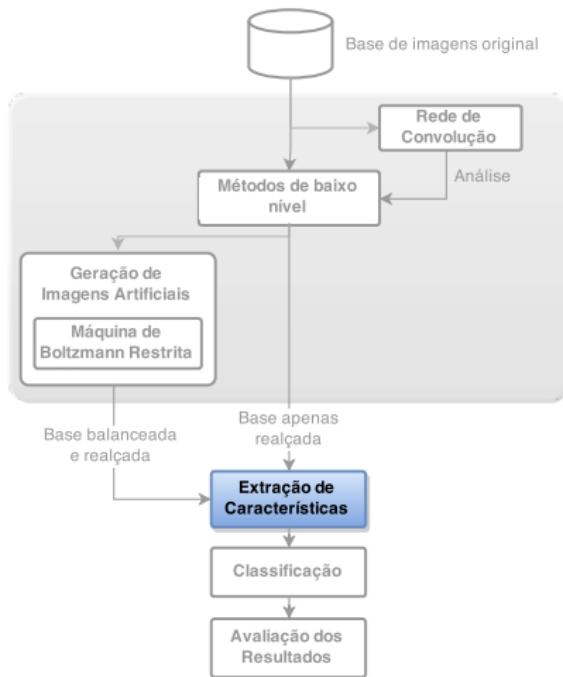
- ▶ Pixels: unidades visíveis. Camada oculta: correlação entre os pixels;
- ▶ Aprendem representações das imagens;
- ▶ Verificação da relevância de uma imagem para o aprendizado:
 - ▶ Utilizar apenas um neurônio oculto, com a matriz de pesos como memória associativa;
 - ▶ Esse memória não se modifica se for dado como entrada uma imagem cujos padrões já foram aprendidos.



Redes Neurais

- ▶ Redes neurais de convolução (CNN)
 - ▶ Compreendem todo o pipeline com camadas de neurônios;
 - ▶ Aprendem as melhores características que diferenciam as classes, utilizando convolução;
- ▶ Máquinas de Boltzmann restritas (RBM)
 - ▶ Aprende a representação das imagens de entrada;
 - ▶ Definir quais imagens são relevantes para o aprendizado.

Extração de Características



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 (Ex. algas - forma).

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

Forma: características externas. Ex. curvatura;

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

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses e Objetivos

2 Contextualização

- Pré-processamento
- Redes de Convolução
- Desbalanceamento de classes
- Máquina de Boltzmann restrita
- Extração de características

3 Metodologia

4 Resultados Preliminares

5 Próximos Passos

6 Atividades e cronograma

Metodologia - Pesquisa Bibliográfica

- ▶ Características latentes;
- ▶ Redes neurais CNN e RBM;
- ▶ Desbalanceamento de classes;
- ▶ Descritores de características;
- ▶ Classificador de padrões.

Metodologia - Implementação

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

Metodologia - Bases de Imagens

- ▶ 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

Bem comportadas

- ▶ 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: DTD.



Figura: Leafsnap.

Metodologia - Experimentos

- ▶ Explorar algoritmos para evidenciar características relevantes:
 - ▶ Melhorar discriminação entre as classes;
 - ▶ Auxiliar no rebalanceamento.
- ▶ Entrada: imagens originais das coleções disponíveis na literatura;
- ▶ Resultado: medidas estatísticas da classificação.

Metodologia - Análise dos resultados

- ▶ Comparar a classificação:
 - ▶ Base original;
 - ▶ Base realçada pelo método proposto.
- ▶ Comparar os métodos:
 - ▶ Geração artificial;
 - ▶ Técnicas de sobreamostragem disponíveis na literatura, como o SMOTE.

Metodologia - Avaliação da Classificação

Medida F1

Problema da acurácia: minoritária sem resultados corretos.

- ▶ Precisão (exatidão): dos exemplos classificados como positivos, quantos realmente são.

$$P = \frac{VP}{VP + FP}$$

- ▶ Revocação (completude): exemplos positivos corretamente classificados como tal.

$$R = \frac{VP}{VP + FN}$$

$$F1 = 2 \frac{P \cdot R}{P + R}$$

Metodologia - Avaliação da Classificação

Medida F1

Problema da acurácia: minoritária sem resultados corretos.

- ▶ Precisão (exatidão): dos exemplos classificados como positivos, quantos realmente são.

$$P = \frac{VP}{VP + FP}$$

- ▶ Revocação (completude): exemplos positivos corretamente classificados como tal.

$$R = \frac{VP}{VP + FN}$$

$$F1 = 2 \frac{P \cdot R}{P + R}$$

Metodologia - Avaliação dos Vetores

Mahalanobis

- ▶ Distância entre a média e a variância das classes;
- ▶ Quanto maior a distância entre as distribuições das classes melhor;
- ▶ Antes e depois dos processamentos a serem realizados.

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.

Metodologia - Avaliação

Teste de Friedman

- ▶ Determinar se há diferença significante entre os resultados gerados;
- ▶ Analisa as performances dos algoritmos e atribui um *rank*;
- ▶ A hipótese nula a ser testada é que não há diferença estatística relevante entre as observações;
- ▶ P-valor indica essa significância (limiar de 0,05).

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses e Objetivos

2 Contextualização

- Pré-processamento
- Redes de Convolução
- Desbalanceamento de classes
- Máquina de Boltzmann restrita
- Extração de características

3 Metodologia

4 Resultados Preliminares

5 Próximos Passos

6 Atividades e cronograma

Descrição do Experimento

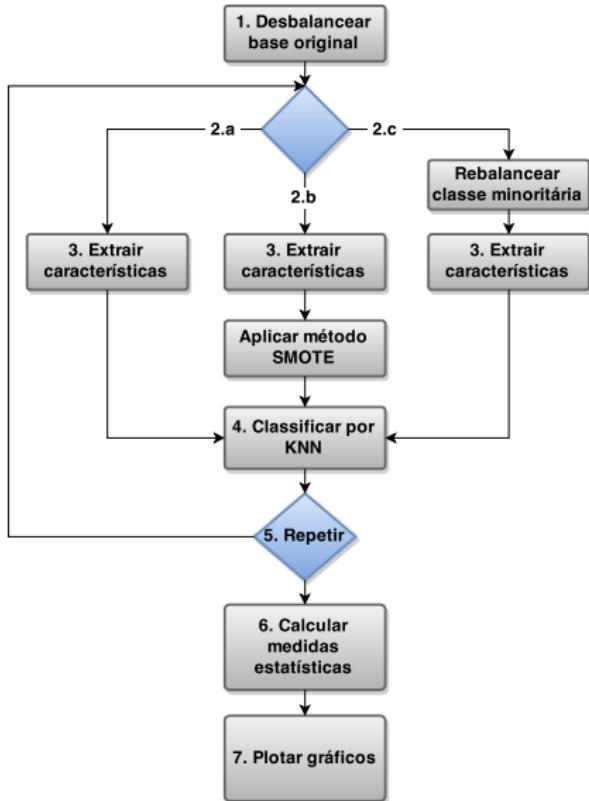


Figura: Fluxo dos resultados preliminares.

Descrição do Experimento

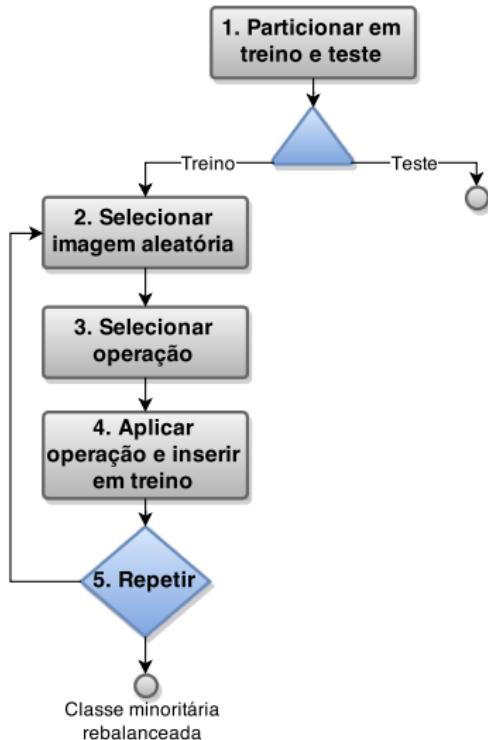


Figura: Fluxo da geração artificial.

Descrição do Experimento

- ▶ Classes praia e montanha da COREL-1000;
- ▶ 50% de teste/treino, com desbalanceamento da classe minoritária em 50, 25, 12 e 6;
- ▶ **Geração artificial:** borramento, adição de ruído, realce, mistura e combinação;
- ▶ **Quantização:** Intensidade para Haralick e MSB para os outros;
- ▶ **Extração de características:** ACC, CCV, BIC, GCH e Haralick;
- ▶ **Classificação** com KNN ($K=1$).

Descrição do Experimento

Classes com alta sobreposição nas características de cor e textura.



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

Descrição do Experimento - Geração Artificial



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

Descrição do Experimento - Extração de Características

GCH *Histograma global de cor - N* (intensidades).

CCV *Vetor de coerência de cor.* Classifica os pixels da imagem em coerentes e incoerentes de acordo com um *threshold*, calcula e concatena os histogramas - $2N$.

BIC *Classificação de pixels de borda e interior.* Mesma cor que seus vizinhos, é pixel de interior. Calcula dois histogramas - $2N$.

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. 1, 3, 5 e 7 - $4N$.

Haralick Entropia, homogeneidade, contraste, correlação, probabilidade máxima e uniformidade - 6.

Descrição do Experimento - Extração de Características

GCH *Histograma global de cor - N* (intensidades).

CCV *Vetor de coerência de cor.* Classifica os pixels da imagem em coerentes e incoerentes de acordo com um *threshold*, calcula e concatena os histogramas - $2N$.

BIC *Classificação de pixels de borda e interior.* Mesma cor que seus vizinhos, é pixel de interior. Calcula dois histogramas - $2N$.

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. 1, 3, 5 e 7 - $4N$.

Haralick Entropia, homogeneidade, contraste, correlação, probabilidade máxima e uniformidade - 6.

Descrição do Experimento - Extração de Características

GCH *Histograma global de cor - N* (intensidades).

CCV *Vetor de coerência de cor.* Classifica os pixels da imagem em coerentes e incoerentes de acordo com um *threshold*, calcula e concatena os histogramas - $2N$.

BIC *Classificação de pixels de borda e interior.* Mesma cor que seus vizinhos, é pixel de interior. Calcula dois histogramas - $2N$.

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. 1, 3, 5 e 7 - $4N$.

Haralick Entropia, homogeneidade, contraste, correlação, probabilidade máxima e uniformidade - 6.

Descrição do Experimento - Extração de Características

GCH *Histograma global de cor - N* (intensidades).

CCV *Vetor de coerência de cor.* Classifica os pixels da imagem em coerentes e incoerentes de acordo com um *threshold*, calcula e concatena os histogramas - $2N$.

BIC *Classificação de pixels de borda e interior.* Mesma cor que seus vizinhos, é pixel de interior. Calcula dois histogramas - $2N$.

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. 1, 3, 5 e 7 - $4N$.

Haralick Entropia, homogeneidade, contraste, correlação, probabilidade máxima e uniformidade - 6.

Descrição do Experimento - Extração de Características

GCH *Histograma global de cor - N* (intensidades).

CCV *Vetor de coerência de cor.* Classifica os pixels da imagem em coerentes e incoerentes de acordo com um *threshold*, calcula e concatena os histogramas - $2N$.

BIC *Classificação de pixels de borda e interior.* Mesma cor que seus vizinhos, é pixel de interior. Calcula dois histogramas - $2N$.

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. 1, 3, 5 e 7 - $4N$.

Haralick Entropia, homogeneidade, contraste, correlação, probabilidade máxima e uniformidade - 6.

Resultados Preliminares - Melhor Rebalanceamento

- ▶ Ganho estatístico da medida-F, quando comparado à geração no espaço de atributos;

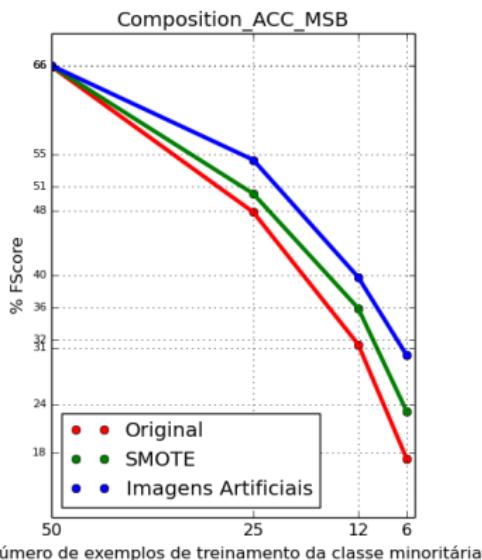


Figura: Resultado obtido com a operação de combinação.

Resultados Preliminares - Pior Rebalanceamento

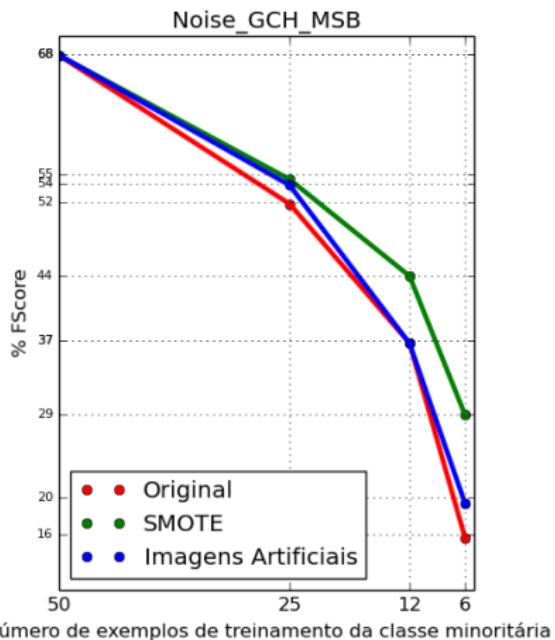


Figura: Piores resultados, obtidos com a adição de ruído.

Resultados Preliminares

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

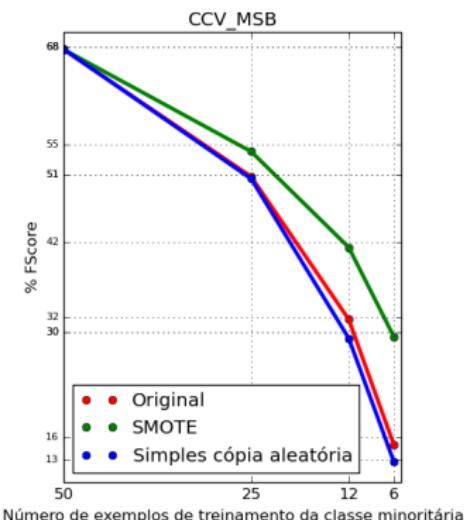


Figura: Simples replicação de exemplos sem pré-processamento.

Resultados Preliminares

- ▶ Melhores operações: todas, apenas mistura e apenas composição;
- ▶ Piores: apenas borramento, ruído e *unsharp masking* (filtros comuns);
- ▶ Descritor que melhor se beneficiou: ACC;
- ▶ Descritores que não se beneficiaram: 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
Artificial	1.3863
Smote	1.6136
Original	3.0000

- ▶ Em algumas execuções: Artificial (1), SMOTE (2) e Original (3).

Resultados Preliminares - Rede de Convolução

<http://caffe.berkeleyvision.org/>

Tabela: Treinamento das classes praia e montanha da base COREL-1000.

Bases	Medida F1
Original não balanceada	0.708
Desbalanceada em 50%	0.577
Rebalanceada	0.677

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses e Objetivos

2 Contextualização

- Pré-processamento
- Redes de Convolução
- Desbalanceamento de classes
- Máquina de Boltzmann restrita
- Extração de características

3 Metodologia

4 Resultados Preliminares

5 Próximos Passos

6 Atividades e cronograma

Resultados Esperados

- ▶ Melhorar a classificação, validando-a com a medida-F1.
 - ▶ *Pré-processamento* de imagens que caracterizem melhor aspectos de suas classes;
 - ▶ *Geração artificial de imagens* de classes minoritárias.
- ▶ Analisar as características aprendidas com a CNN em bases específicas;
- ▶ Escolher imagens que adicionam informações (RBM);
- ▶ Testar bases naturalmente não balanceadas.

Próximos Passos - Curto Prazo

- ▶ Desbalancear as 10 classes da COREL-1000 em escada;
- ▶ Utilizar todas as classes com apenas 1 desbalanceada;
- ▶ Analisar as características latentes que as redes neurais de convolução conseguem extrair.
 - ▶ Bases comportadas quanto às propriedades de textura, cor e forma.

Próximos Passos - Médio Prazo

- ▶ Analisar a memória associativa aprendida com uma máquina de Boltzmann restrita.
 - ▶ Escolher para qual imagem original utilizar, ao invés do método aleatório utilizado nos resultados preliminares;
 - ▶ Verificação da relevância das imagens geradas.

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses e Objetivos

2 Contextualização

- Pré-processamento
- Redes de Convolução
- Desbalanceamento de classes
- Máquina de Boltzmann restrita
- Extração de características

3 Metodologia

4 Resultados Preliminares

5 Próximos Passos

6 Atividades e cronograma

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.
Disciplinas	[REDACTED]	[REDACTED]			
Revisão		[REDACTED]	[REDACTED]	[REDACTED]	
Geração artificial		[REDACTED]	[REDACTED]	[REDACTED]	
Características latentes			[REDACTED]	[REDACTED]	
Experimentos		[REDACTED]	[REDACTED]	[REDACTED]	
Escrita científica		[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

Artigo aceito na Neurocomputing

Image quantization as a dimensionality reduction procedure in color and texture feature extraction

Monair Ponti, Tiago S. Nazaré, Gabriela S. Thumé

Instituto de Ciências Matemáticas e de Computação

Universidade de São Paulo

São Carlos, SP — Brazil — 13566-590

Abstract

The image-based visual recognition pipeline includes a step that converts color images into images with a single channel, obtaining a color-quantized image that can be processed by feature extraction methods. In this paper we explore this step in order to produce compact features that can be used in retrieval and classification systems. We show that different quantization methods produce very different results in terms of accuracy. While compared with more complex methods, this procedure allows the feature extraction in order to achieve a significant dimensionality reduction, while preserving or improving system accuracy. The results indicate that quantization simplify images before feature extraction and dimensionality reduction, producing more compact vectors and reducing system complexity.

Keywords: Image quantization, LPP, compact features extraction

Email addresses: ponti@usp.br (Monair Ponti), tiagossn@usp.br (Tiago S. Nazaré), gabithume@usp.br (Gabriela S. Thumé)

Agradecimentos

Moacir Pereira Ponti Junior



Referências I

- [1] GE E Batista, RC C Prati, and MC C Monard. A study of the behavior of several methods for balancing machine learning training data. *ACM Sigkdd Explorations Newsletter*, 6(1):20–29, 2004.
- [2] Yoshua Bengio. Learning Deep Architectures for AI. *Foundations and Trends in Machine Learning*, 2(1):1–127, January 2009.
- [3] Yoshua Bengio, Ian J. Goodfellow, and Aaron Courville. Deep learning. Book in preparation for MIT Press, 2014.
- [4] NV V Chawla, LO O Hall, and KW W Bowyer. SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16:321–357, 2002.
- [5] R. De Maesschalck, D. Jouan-Rimbaud, and D.L. Massart. The Mahalanobis distance. *Chemometrics and Intelligent Laboratory Systems*, 50(1):1–18, January 2000.
- [6] Asja Fischer and Christian Igel. Training restricted Boltzmann machines: An introduction. *Pattern Recognition*, 47(1):25–39, January 2014.
- [7] Rafael C Gonzalez and Richard E Woods. *Digital Image Processing*. Prentice-Hall, 3 edition, 2007.
- [8] Robert M. Haralick, K. Shanmugam, and Its'Hak Dinstein. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6):610–621, November 1973.
- [9] Jing Huang, S Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih. Image indexing using color correlograms. In *Computer Society Conference on Computer Vision and Pattern Recognition*, pages 762–768. IEEE, 1997.
- [10] Nathalie Japkowicz and Shaju Stephen. The class imbalance problem : A systematic study. *Intelligent data analysis*, 6(5):429–449, 2002.

Referências II

- [11] Christopher Kanan and Garrison W Cottrell. Color-to-grayscale: does the method matter in image recognition? *PLoS one*, 7(1):e29740, January 2012.
- [12] Yann LeCun, Koray Kavukcuoglu, and Clement Farabet. Convolutional networks and applications in vision. In *International Symposium on Circuits and Systems*, pages 253–256. IEEE, May 2010.
- [13] Aravindh Mahendran and Andrea Vedaldi. Understanding deep image representations by inverting them. arXiv preprint arXiv:1412.0035, 2014.
- [14] Michael A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.
- [15] Greg Pass, Ramin Zabih, and Justin Miller. Comparing images using color coherence vectors. In *Proceedings of the fourth ACM international conference on Multimedia*, pages 65–73, New York, USA, February 1996. ACM Press.
- [16] CT T Picon, Isadora Rossi, and MP Ponti Ponti Jr. Análise da classificação de imagens por descritores de cor utilizando várias resoluções. In *Workshop of Undergraduate Works*. SIBGRAPI, 2011.
- [17] Moacir Ponti, Tiago Nazaré, and Gabriela Thumé. Image quantization as a dimensionality reduction procedure in color and texture feature extraction. Submitted to Neurocomputing, 2014.
- [18] Anderson Rocha, Daniel C. Hauagge, Jacques Wainer, and Siome Goldenstein. Automatic fruit and vegetable classification from images. *Computers and Electronics in Agriculture*, 70(1):96–104, January 2010.
- [19] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. September 2014.
- [20] Renato O. Stehling, Mario A. Nascimento, and Alexandre X. Falcão. A compact and efficient image retrieval approach based on border/interior pixel classification. In *Proceedings of the eleventh international conference on Information and knowledge management*, pages 102–109, New York, USA, November 2002. ACM Press.
- [21] Matthew D Zeiler and Rob Fergus. Visualizing and Understanding Convolutional Networks. *Computer Vision–ECCV 2014*, pages 818–833, November 2013.

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

