

Geração de imagens artificiais e quantização aplicadas a problemas de classificação

Gabriela Salvador Thumé

Orientador: Prof. Dr. Moacir Antonelli Ponti

Co-orientador: Prof. Dr. João do Espírito Santo Batista Neto

Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo

29 de abril de 2016



Conselho Nacional de Desenvolvimento
Científico e Tecnológico



Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses
- Contribuições

2 Quantização de imagens

- Contextualização
- Método

• Experimentos

3 Geração de imagens artificiais

- Contextualização
- Método
- Experimentos

4 Conclusões

- Trabalhos futuros

5 Agradecimentos

Introdução

- ▶ Classificação de imagens: predizer a classe;
- ▶ Extração de características: vetores representativos;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização: classificar novos exemplos;
- ▶ Porém, existem características que não são suficientes para a diferenciação entre as classes;

Introdução

- ▶ Classificação de imagens: predizer a classe;
- ▶ Extração de características: vetores representativos;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização: classificar novos exemplos;
- ▶ Porém, existem características que não são suficientes para a diferenciação entre as classes;

Introdução

- ▶ Classificação de imagens: predizer a classe;
- ▶ Extração de características: vetores representativos;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização: classificar novos exemplos;
- ▶ Porém, existem características que não são suficientes para a diferenciação entre as classes;

Introdução

- ▶ Classificação de imagens: predizer a classe;
- ▶ Extração de características: vetores representativos;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização: classificar novos exemplos;
- ▶ Porém, existem características que não são suficientes para a diferenciação entre as classes;

Introdução

- ▶ Classificação de imagens: predizer a classe;
- ▶ Extração de características: vetores representativos;
- ▶ Algoritmos de aprendizado de máquina: modelo de representação;
- ▶ Generalização: classificar novos exemplos;
- ▶ Porém, existem características que não são suficientes para a diferenciação entre as classes;

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

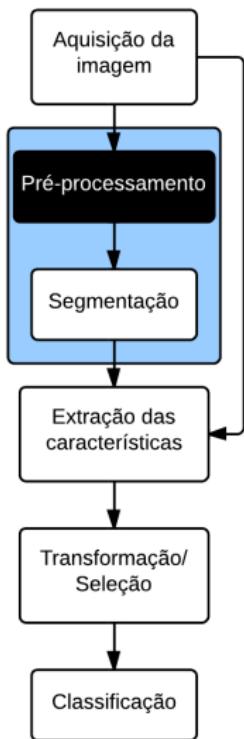


Figura 1: Etapas canônicas do reconhecimento de padrões.

Motivação — Quantização

- ▶ Redução da complexidade no ínicio do reconhecimento;
- ▶ Apesar de fazer parte do pipeline, muitos estudos não descrevem o método de quantização e seus parâmetros;
- ▶ Quantização pode impactar a classificação (Kanan e Cottrell, 2012);
- ▶ Ao negligenciar essa etapa, perde-se a *oportunidade de redução da dimensionalidade do vetor de características e/ou do tempo de execução das etapas posteriores.*

Motivação — Desbalanceamento de classes

- ▶ Diferença entre o número de exemplos disponíveis;
- ▶ Imagens representam eventos importantes mas menos frequentes;
- ▶ Métodos de transformação do espaço e de classificação assumem que a base está balanceada;
- ▶ Preferência à predição da classe majoritária, prejudicando a classificação da minoritária.

Motivação — Rebalanceamento de classes

- ▶ Synthetic Minority Over-sampling Technique (SMOTE) propõe a geração de exemplos a partir dos vetores de características originais das classes minoritárias;
- ▶ Não existem estudos dessas técnicas em dados de informação visual para o rebalanceamento de classes;
- ▶ Proposta: *geração de imagens artificiais a partir do processamento das imagens da classe minoritária.*



Figura 2: Imagem artificialmente gerada.

Hipóteses

Utilizar um número reduzido de cores

- ▶ Juntamente com um método de quantização apropriado;
- ▶ Antes da extração de características;
- ▶ *Obter vetores de características mais compactos e com maior capacidade de discriminação entre classes.*

Geração de imagens artificiais

- ▶ Balancear as classes;
- ▶ Preparação para a extração de características;
- ▶ *Melhorar a acurácia, quando comparada à geração de exemplos artificiais no espaço de atributos.*

Contribuições

- ▶ Demostrar que é possível obter vetores de características compactos e efetivos ao extrair características de imagens quantizadas.
 - ▶ Custo computacional baixo;
 - ▶ Reduzindo a dimensionalidade do vetor após a quantização e posterior extração de características;
 - ▶ Redução do tempo de processamento para os métodos de descrição de textura.

Contribuições

- ▶ Demostrar que a geração de imagens artificiais pode contribuir com o balanceamento entre classes.
 - ▶ Melhorando o *F1-Score* resultante de algoritmos de classificação;
 - ▶ Comparando com a geração de exemplos artificiais no espaço de atributos (SMOTE).

Contribuições em código e reproduzibilidade

- ▶ Código para a quantização:

<http://dx.doi.org/10.5281/zenodo.15932>

- ▶ Código da geração de imagens artificiais:

<https://github.com/GabiThume/msc-src>

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses
- Contribuições

2 Quantização de imagens

- Contextualização
- Método

• Experimentos

3 Geração de imagens artificiais

- Contextualização
- Método
- Experimentos

4 Conclusões

- Trabalhos futuros

5 Agradecimentos

Quantização de imagens

- ▶ Métodos de extração preparados para receber imagens em um canal de cor;
- ▶ Características extraídas para cada canal de cor e posteriormente concatenadas;
- ▶ O pipeline de reconhecimento de imagens inclui a conversão de imagens coloridas em imagens com apenas um canal ($2^3 = 8$ bits);
- ▶ Explorar essa etapa para produzir vetores mais compactos;
- ▶ Diferentes parâmetros de quantização combinados com quatro métodos de extração de cor e um de textura.

Pré-processamento de Imagens — Quantização



(a) Original (b) Intensidade' (c) Gleam



(d) Luminância' (e) MSB

Figura 3: Conversão para a escala de cinza.

Quantização de imagens

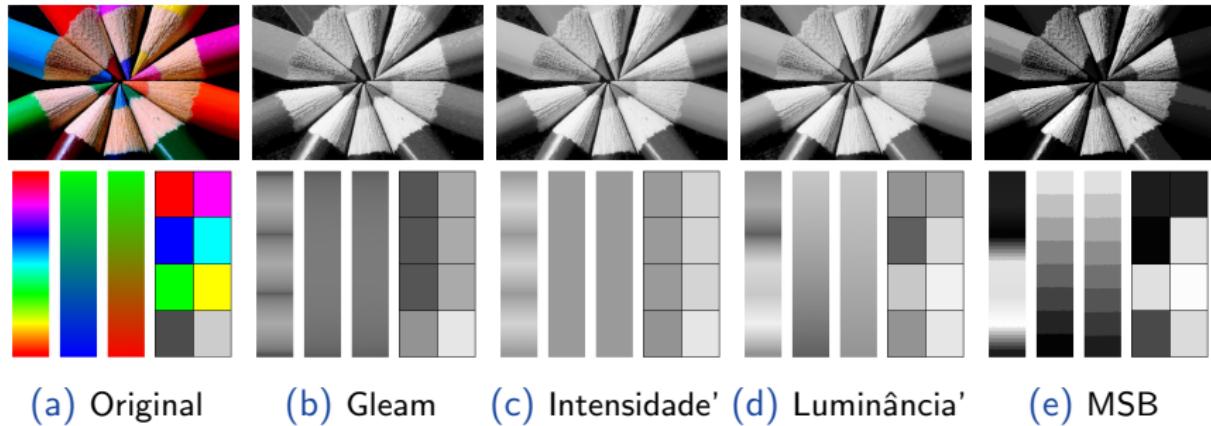


Figura 4: Os métodos *Luminância'* e *MSB* apresentam uma melhor discriminação entre cores. A versão de um canal de cor possui 232 intensidades únicas para o método (e) *MSB* e 184 intensidades para os demais métodos.

Quantização de imagens

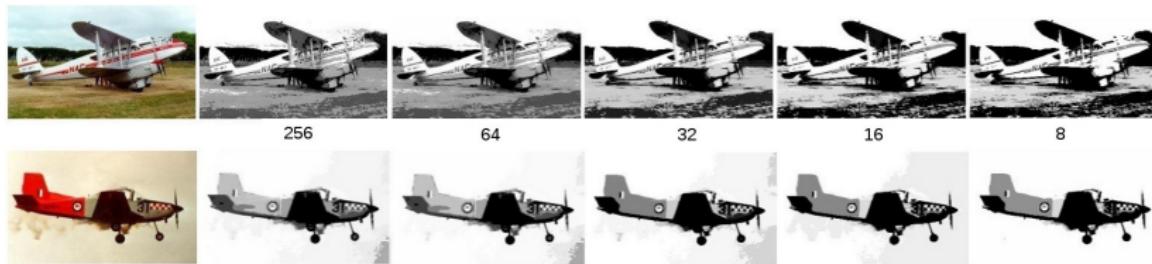


Figura 5: Imagens da base *Caltech101-600* com variações no parâmetro de cor utilizando o método MSB. Com 256 e 64 há preservação das cores. Mas com apenas 32, há perda considerável de informação nas regiões com pouco contraste.

Experimentos

- ▶ Experimentos utilizando um método de extração de características seguido pela classificação (sem posterior seleção de características);
- ▶ Experimentos utilizando o vetor resultante da concatenação de todos os métodos de extração, seguido pela classificação com e sem a seleção de características.

Experimentos

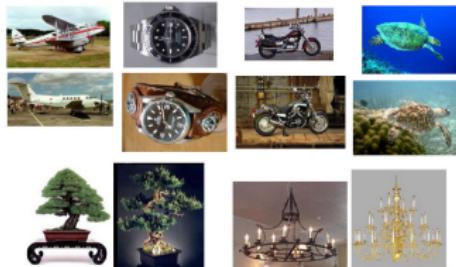


Figura 6: Fluxo das operações e os métodos utilizados nos experimentos.

Experimentos — Bases de Imagens



(a) Corel-1000



(b) Caltech101-600



(c) Produce-1400

Figura 7: Bases de imagens utilizadas nos experimentos de quantização.

Experimentos — Protocolo

1. **Quantização:** *Intensidade', Gleam, Luminância'* e MSB.
2. **Extração de características**
 - ▶ *Auto Color Correlogram* (ACC): quatro distâncias $d = 1, 3, 5$ e 7 ;
 - ▶ *Border-Interior Classification* (BIC): vizinhança de quatro pixels;
 - ▶ *Color Coherence Vector* (CCV): *threshold* = 25;
 - ▶ Haralick-6: pixel vizinho à direita.
3. **Redução da dimensionalidade:** *Locality Preserving Projections* (LPP) com $d = 128, 64, 32$ e 16 dimensões e $k = 10$ vizinhos.
4. **Classificação:** *Support Vector Machines* (SVM) com *grid search* no conjunto de treinamento.

Experimentos — Resultados

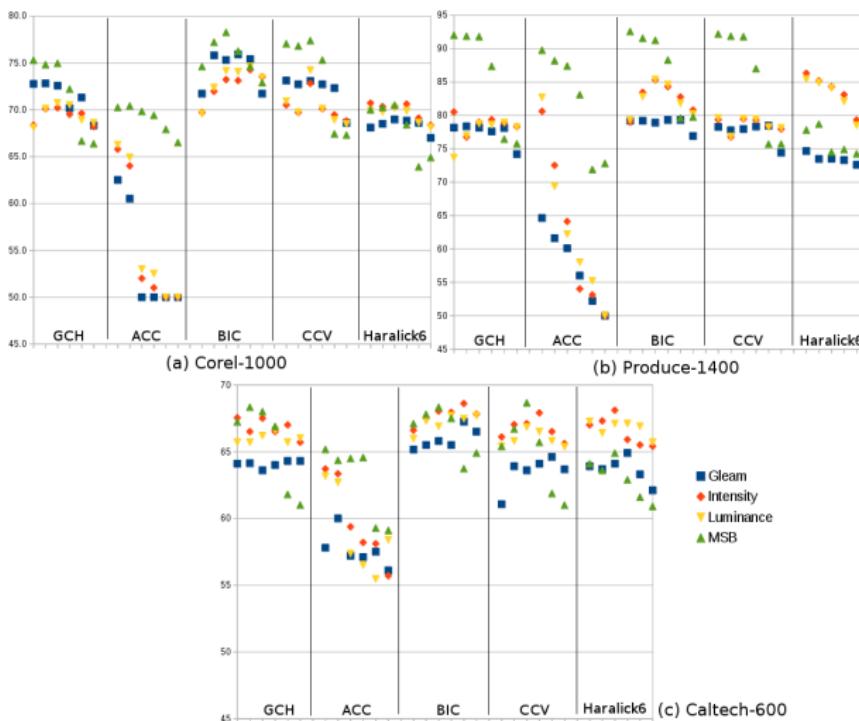


Figura 8: Acurácia média utilizando 256, 128, 64, 32, 16 e 8 cores.

Experimentos — Resultados

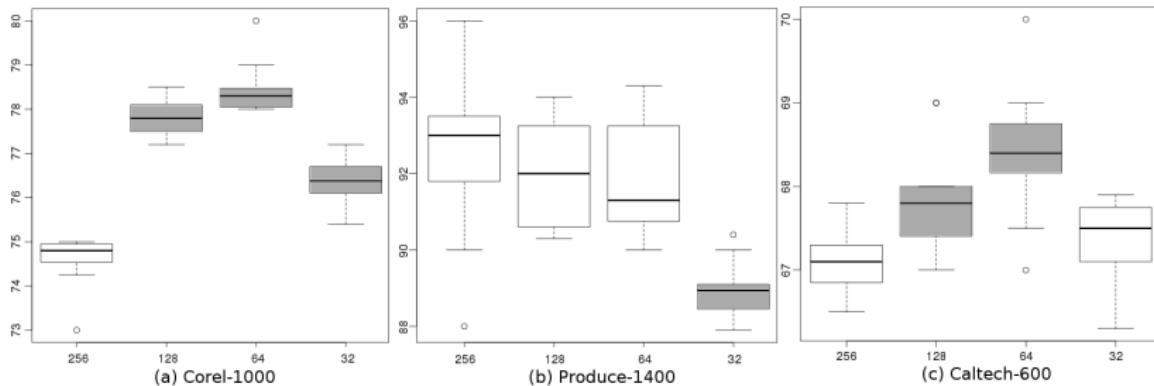


Figura 9: Acurácia média utilizando MSB e BIC ao variar o parâmetro de quantização. Os *boxplots* em cinza correspondem às significâncias estatísticas com $p\text{-value} < 0.01$ quando comparado à 256 cores. Converter para 64 cores prova ser boa escolha de processamento.

Experimentos — Resultados

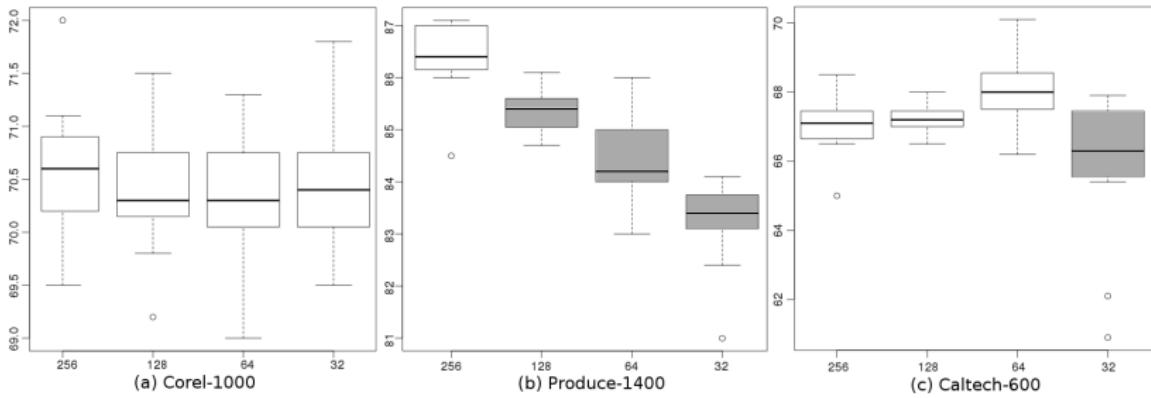


Figura 10: Acurácia média utilizando *Luminância*' e Haralick ao variar o parâmetro de quantização. Dependendo da base, a redução pode manter ou degradar a acurácia.

Experimentos — Quantização versus LPP

- ▶ Imagem convertida para escala de cinza com MSB em 256 cores;
- ▶ Extração de características com o método BIC;
- ▶ Vetor de entrada para LPP:
 - ▶ Versões reduzidas de 256, 128 e 64 dimensões.
- ▶ Comparação da acurácia dos vetores reduzidos pela quantização.

Experimentos — Resultados da quantização versus LPP

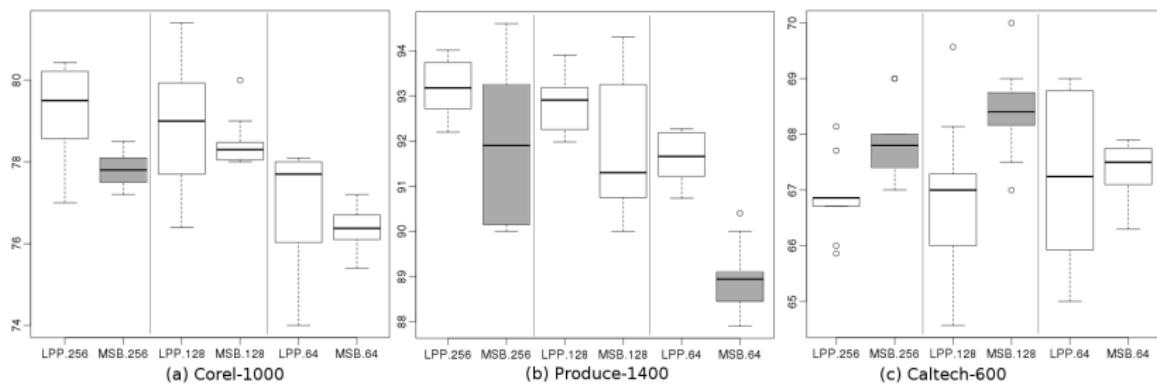


Figura 11: Acurácia para os métodos MSB e LPP. A comparação foi realizada com a mesma dimensionalidade. Se utilizado um número de cores correto, é possível manter ou melhorar a acurácia.

Experimentos — Concatenação

- ▶ Número de dimensões de um único descritor pode ser baixo;
- ▶ Imagens convertidas em escala de cinza com 256 cores;
- ▶ Descritas por todos os métodos e suas características concatenadas em um vetor $D = 2310$;
- ▶ Redução de dimensionalidade com LPP para $d = 1160, 582, 294$ e 150. Mesmas dimensões de utilizar quantização como redução.

Experimentos — Resultados da concatenação

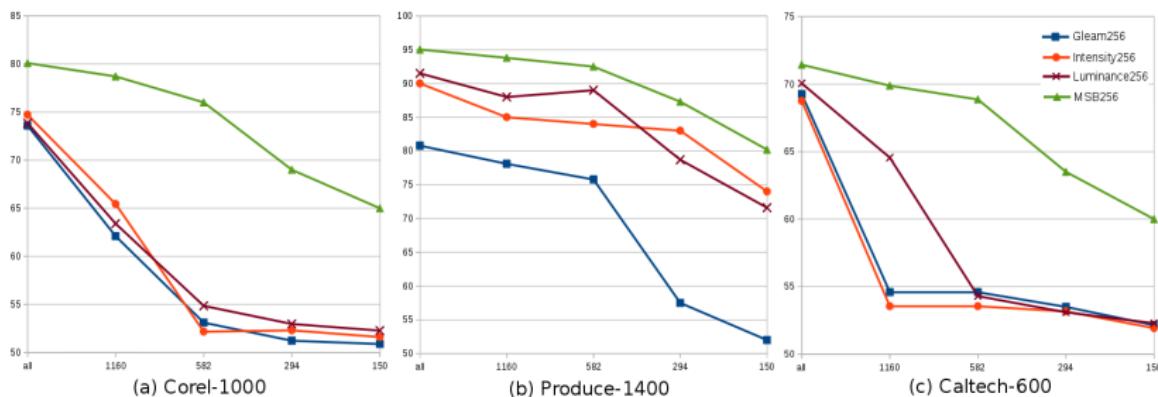


Figura 12: Acurácia para *Gleam*, *Intensidade*', *Luminância*' e *MSB*. O vetor de características com $D = 2310$ sofreu redução da dimensionalidade com o LPP para $d = 1160, 582, 294$ e 150 .

Experimentos — Resultados da concatenação

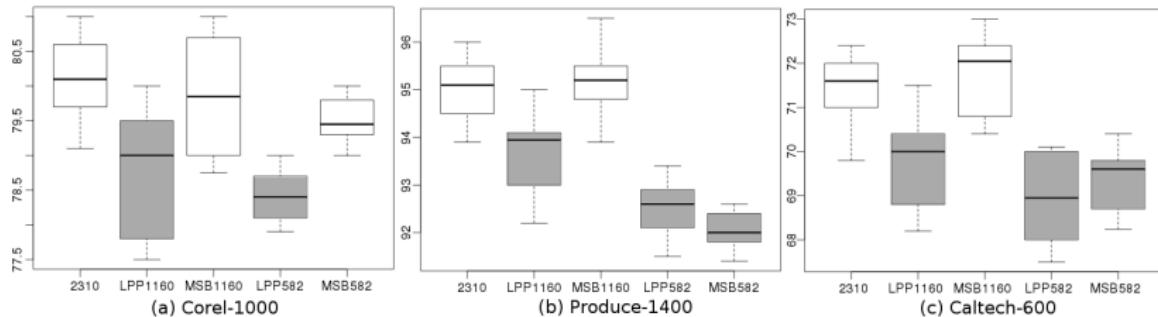


Figura 13: Comparação da redução de dimensionalidade obtida com LPP e MSB. Utilizar 128 ou 64 cores é indicado como uma boa escolha do parâmetro de quantização.

Experimentos — Resultados da concatenação

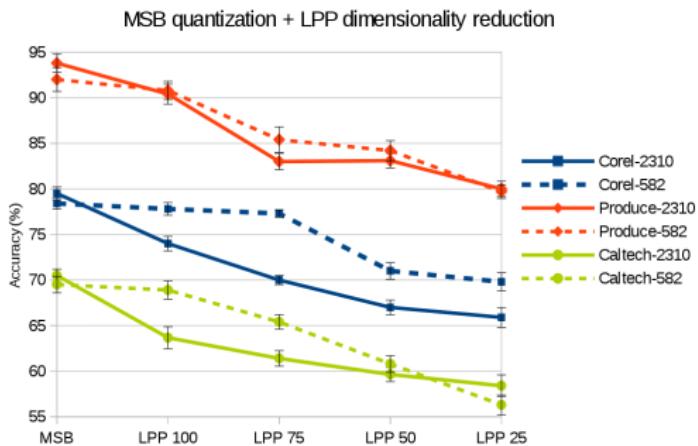


Figura 14: Projeção LPP sobre o espaço produzido pelo método MSB com 256 ($d = 2310$) e 64 cores ($d = 582$). Em geral, as projeções com as imagens quantizadas em 64 cores foram melhores.

Experimentos — Discussão

- ▶ Aplicar a quantização na etapa de pré-processamento pode reduzir significativamente a dimensionalidade, enquanto melhora ou mantém a classificação do sistema;
- ▶ Inclusive quando comparado com métodos mais complexos para a redução da dimensionalidade;
- ▶ O vetor concatenado possui $9C + 6$ dimensões e o tempo de extração de todas as características é $f(N) = 42N + 6C^2$. Utilizar 64 cores ao invés de 256 para 100 imagens corresponde a uma redução de 74,6% do número de instruções.

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses
- Contribuições

2 Quantização de imagens

- Contextualização
- Método

● Experimentos

3 Geração de imagens artificiais

- Contextualização
- Método
- Experimentos

4 Conclusões

- Trabalhos futuros

5 Agradecimentos

Desbalanceamento de classes

- ▶ Número desbalanceado de exemplos: majoritárias x minoritárias;
- ▶ Abordagens:
 - ▶ *Pré-processamento ao reamostrar os dados:*
 - ▶ Aumentar a minoritária (Em geral, melhores resultados);
 - ▶ Diminuir a majoritária.
 - ▶ *Modificar métodos de aprendizagem:* adicionar funções de custo na classificação;

Desbalanceamento de classes — Subamostragem

- ▶ Diminuir o número de elementos do conjunto;
- ▶ Eliminar elementos distantes da fronteira de decisão (menos relevantes);
- ▶ Normalmente apresentam resultados piores;
- ▶ Pode remover informações essenciais dos dados originais;
- ▶ Não há melhor para todos os cenários.

Desbalanceamento de classes — Sobreamostragem

- ▶ Aumentar o número de elementos;

SMOTE (Chawla et al., 2002)

- ▶ 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;

Desbalanceamento de classes — SMOTE

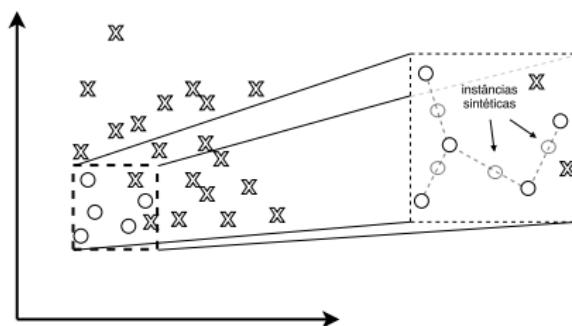


Figura 15: SMOTE

- ▶ Força uma região de decisão maior e mais geral;
- ▶ Rebalancear ao gerar novos elementos, ao invés de replicá-los;
- ▶ **Diferentes estratégias para criar exemplos sintéticos podem melhorar a performance da classificação;**
- ▶ Utilizado para comparação.

Geração de imagens artificiais

- ▶ Compensar a baixa disponibilidade de exemplos de uma determinada classe;
- ▶ Permitir a extração de informações antes não disponíveis nas imagens originais por meio da combinação ou perturbação das imagens de entrada.

Geração de imagens artificiais

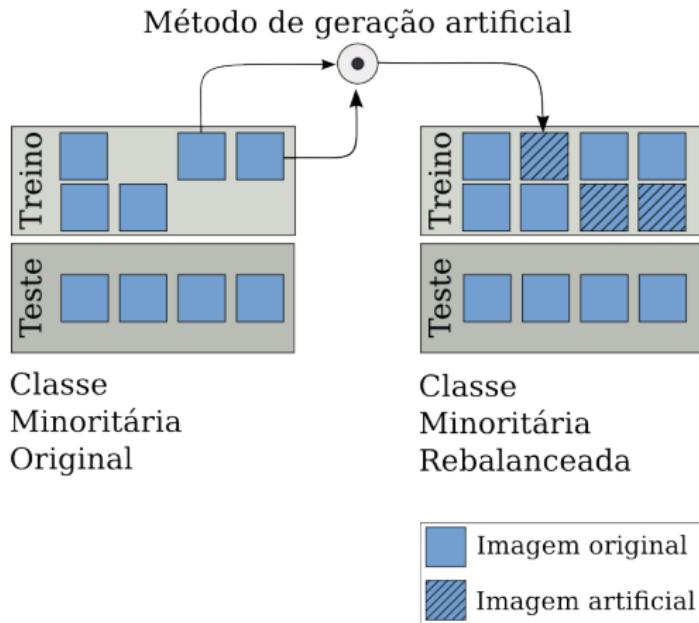


Figura 16: Geração de imagens artificiais da classe minoritária para rebalancear a base de imagens.

Geração de imagens artificiais — Borramento



(a) Original



(b) Imagem artificial

Figura 17: Filtro bilateral. A imagem (b) possui detalhes borradados, porém preservando as bordas.

Geração de imagens artificiais — Aguçamento



(a) Original



(b) Imagem artificial

Figura 18: *Unsharp mask.* A imagem resultante (b) apresenta saliência nas transições de intensidade.

Geração de imagens artificiais — Adição de ruído



(a) Original



(b) Imagem artificial

Figura 19: Ruído de Poisson. Regiões claras de (b) apresentam mais ruído que as regiões escuras.

Geração de imagens artificiais — SMOTE visual



(a) Original



(b) Original



(c) Imagem artificial

Figura 20: Smote ao nível de pixels. É possível notar a sobreposição de uma “sombra” da Figura (b) em (a).

Geração de imagens artificiais — Mistura ponderada



(a) Original



(b) Original



(c) Imagem artificial

Figura 21: Soma ponderada de duas imagens. A imagem (c) representa a mistura de (a) e (b).

Geração de imagens artificiais — Mistura limiarizada



(a) Original



(b) Original



(c) Imagem artificial

Figura 22: Mistura de *thresholds*. A imagem resultante (c) é uma composição do *foreground* da primeira imagem sobre a segunda.

Geração de imagens artificiais — Mistura saliente



(a) Original



(b) Original



(c) Imagem artificial

Figura 23: Combinação de saliência. A imagem resultante (c) apresenta a região saliente de (b) sobreposta em (a).

Geração de imagens artificiais — Composição



Figura 24: Várias imagens, dispostas em um mosaico, formam a imagem resultante. Cada célula do mosaico sofre uma operação, sorteada no momento da geração da imagem.

Experimentos

Comparar a classificação:

- ▶ Base original;
- ▶ Geração de imagens artificiais com os métodos anteriores;
- ▶ Técnica de sobreamostragem SMOTE.

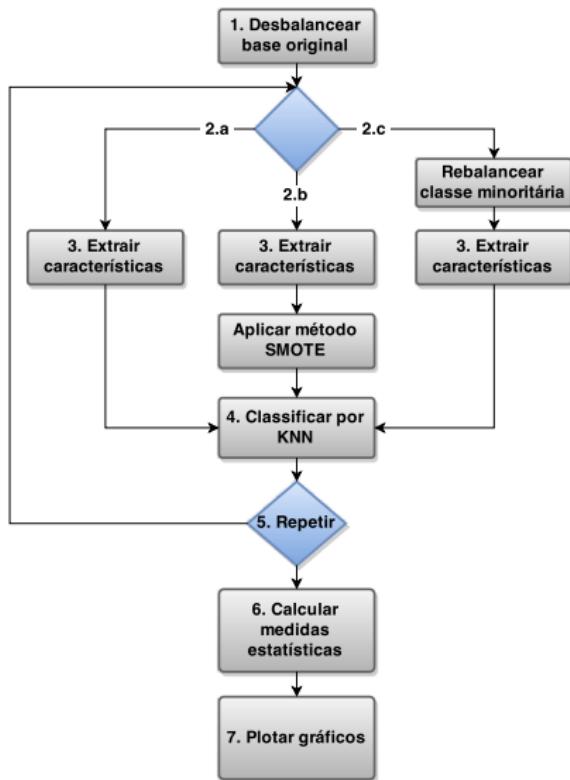


Figura 25: Fluxo de operações para obtenção dos resultados do rebalanceamento de classes.

Experimento — Visualização

- ▶ Visualização do espaço obtido após a geração de imagens;
- ▶ Verificar a definição da classe minoritária em relação ao espaço original;
- ▶ Analisar qual método mais se assemelha à distribuição original dos dados.

Protocolo — Visualização I

1. **Imagens originais:** *Horse* e *Elephant* com 100 imagens cada;



(a) Elephant



(b) Horse

Figura 26: Originalmente da base de imagens Corel-1000, a principal característica dessas imagens é a diferença de cores, contendo pequeno grau de sobreposição.

Protocolo — Visualização II

2. **Desbalanceamento:** cada classe foi dividida em 50% para treino e 50% para teste, de maneira aleatória. Após, a classe *Horse* sofreu remoção de 50% do seu conjunto de treino;

Protocolo — Visualização III

3. Método para geração de imagens: *mistura ponderada*;



(a) Original



(b) Original



(c) Imagem artificial

Figura 27: Mistura ponderada. A imagem resultante (c) é composta pela mistura de (a) e (b).

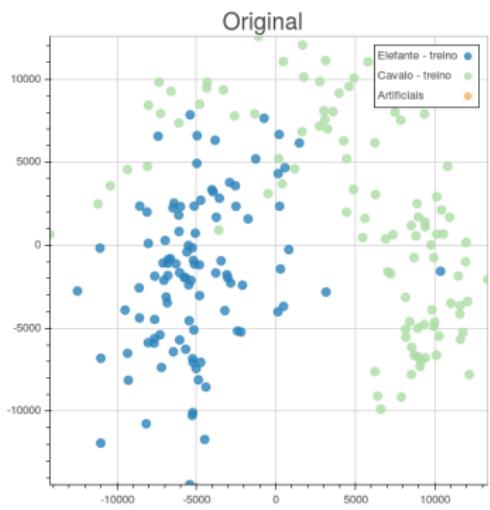
4. Conversão em escala de cinza: *Intensidade*;

5. Extração de características: BIC;

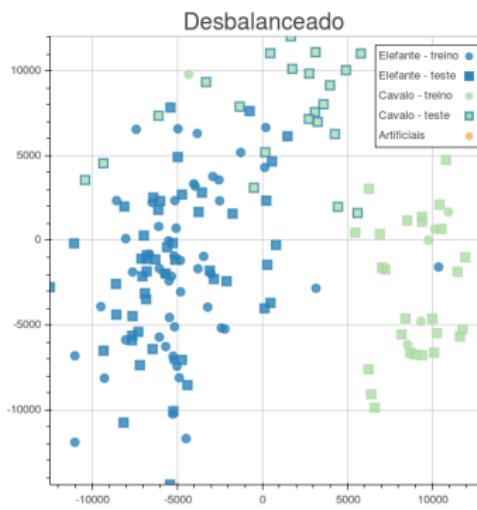
Protocolo — Visualização IV

6. **Classificação:** KNN com $K = 1$;
7. **Projeção multidimensional para visualização:** projetados os dois componentes principais encontrados ao aplicar PCA nos vetores de características.

Experimento — Visualização



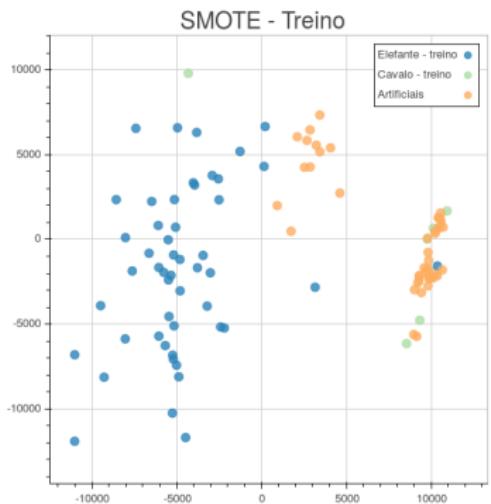
(a) Original



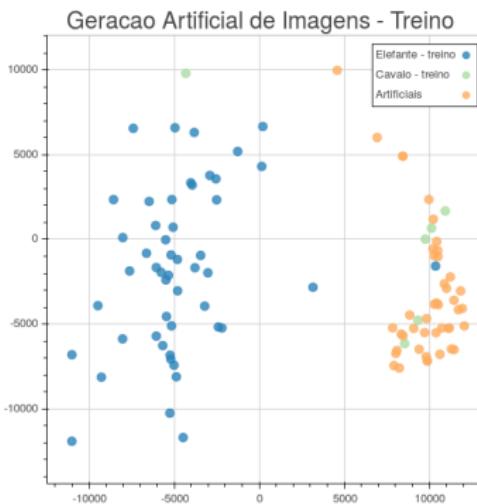
(b) Desbalanceado

Figura 28: Projeção dos dois componentes principais obtidos com o PCA. À direita, as mesmas classes após a remoção de 50% do treino da classe Horse.

Experimento — Visualização



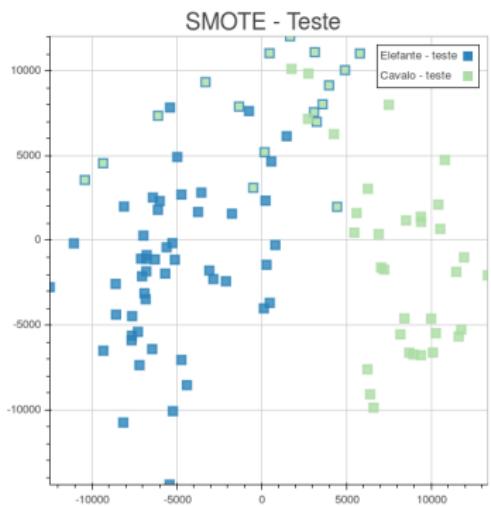
(a) SMOTE



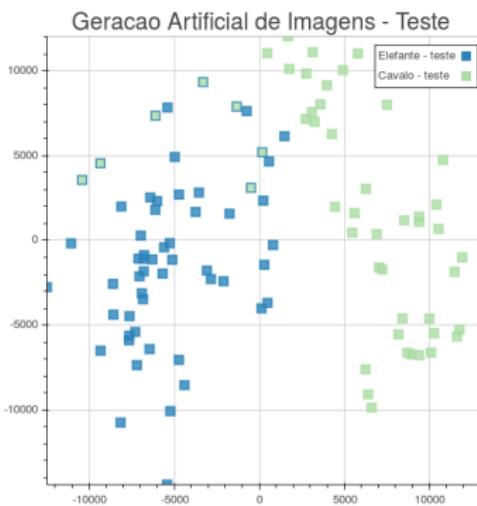
(b) Geração de imagens artificiais

Figura 29: Novos exemplos de treinamento da geração com SMOTE e no campo visual com o método *mistura* (em laranja). Ganho de mais de 10% de acurácia.

Experimento — Visualização



(a) Smote



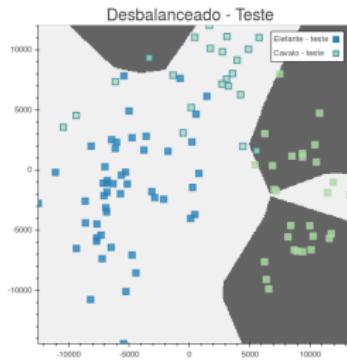
(b) Geração de imagens

Figura 30: A cor da borda dos marcadores representa a classe predita pelo classificador K-NN com $K = 1$ após o treinamento realizado com as bases rebalanceadas.

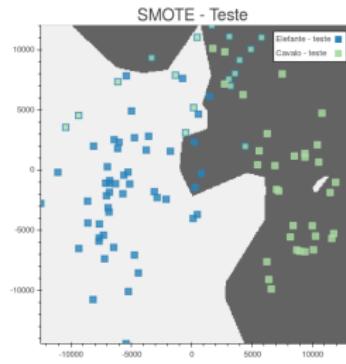
Experimento — Visualização

- ▶ A geração de imagens melhorou a definição da classe minoritária;
- ▶ Se assemelhou à distribuição dos dados originais;
- ▶ Ao interpolar os vetores originais o SMOTE pode criar exemplos em regiões que pertencem a outra classe;
- ▶ Além disso, não extrapolou a sua região.

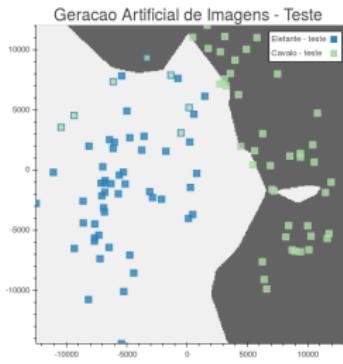
Experimento — Visualização



(a) Desbalanceado



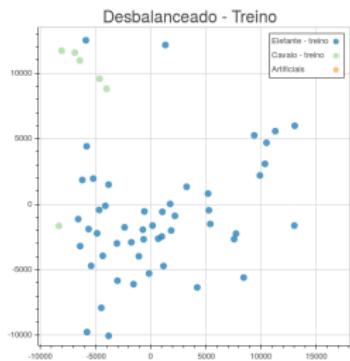
(b) Smote



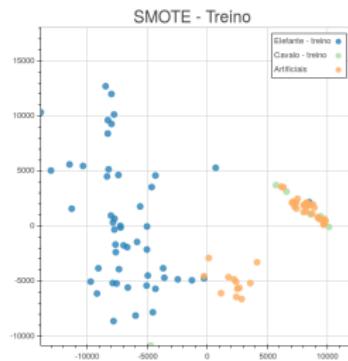
(c) Geração de imagens

Figura 31: Região de decisão com 1-NN. A classe minoritária apresenta-se melhor representada e o SMOTE ocasionou uma certa invasão do espaço de características da classe majoritária.

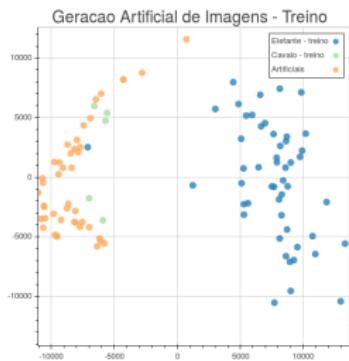
Experimento — Visualização



(a) Desbalanceado



(b) Smote



(c) Geração de imagens

Figura 32: Melhores subespaços encontrados (componentes recalculadas). A geração de imagens artificiais proporciona a criação de um subespaço que melhor discretiza as classes.

Experimento — Visualização

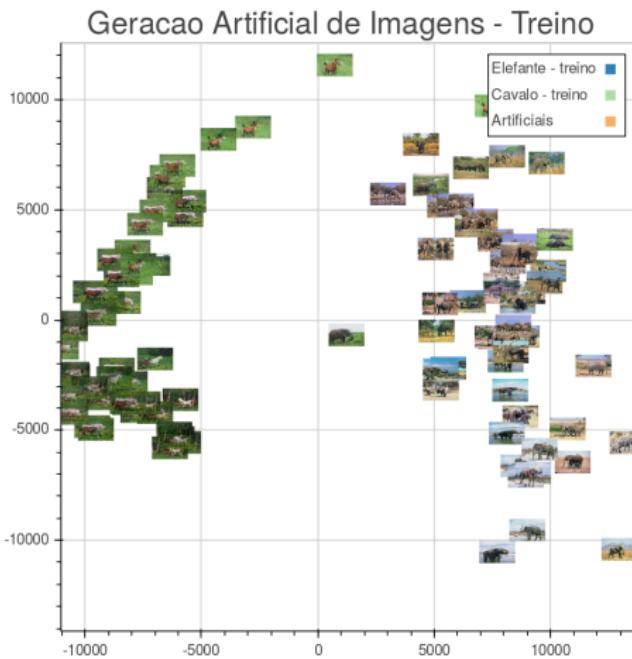


Figura 33: Impacto do método de extração de características BIC na separação entre classes.

Experimentos — Validação

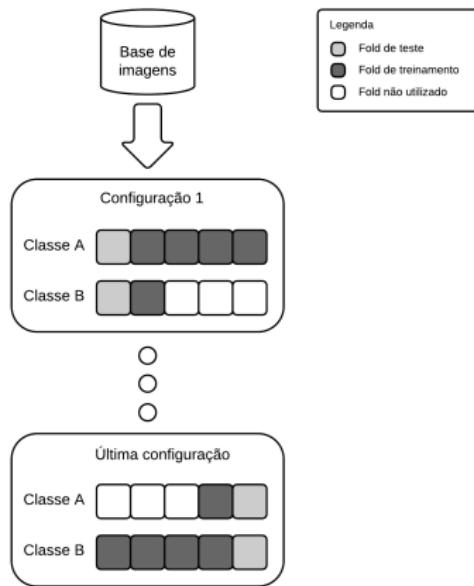


Figura 34: Validação k -fold com o objetivo de prover mais robustez ao sistema.

Experimento

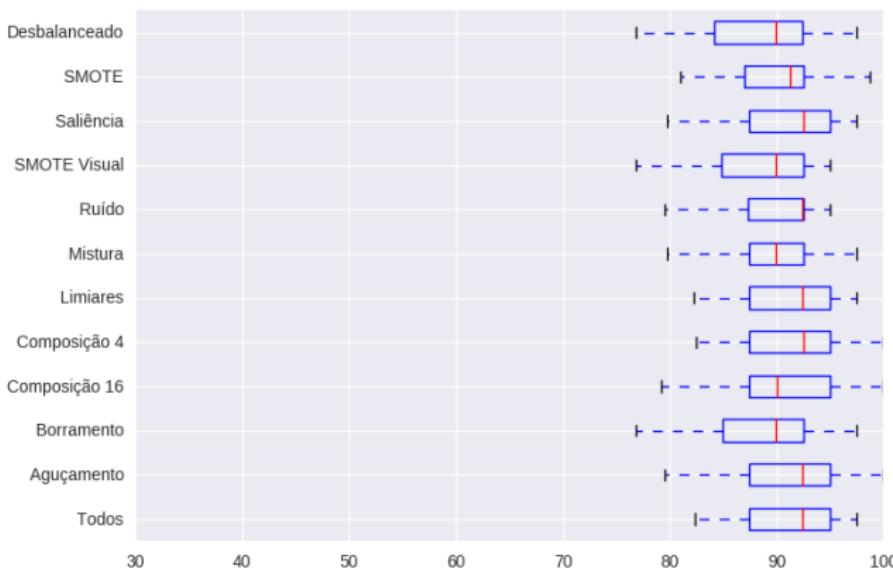


Figura 35: Boxplot para os métodos *Gleam* e ACC, combinação com melhor F1-Score para *Horse* e *Elephant*.

Experimento

Tabela 1: A geração com *Composição 4* obteve maior valor de *F1-Score*. Difere do resultado utilizando BIC e Intensidade'.

Gleam & ACC	Média	Desvio Padrão
Todos	91.090913	4.559066
Aguçamento	91.002678	4.907016
Borramento	89.394500	5.103498
Composição 16	90.934305	4.399334
Composição 4	91.773528	4.909852
Limiares	90.893133	5.285833
Mistura	90.177055	4.409787
Ruído	89.337770	5.169757
SMOTE Visual	88.616535	5.567976
Saliência	91.282655	4.230281
SMOTE	90.173808	4.566863
Desbalanceado	88.258567	5.538461

Experimento



(a) Imagem gerada

Figura 36: A imagem gerada apresenta uma *composição* de quatro imagens da classe *Elephant*.

Teste *post-hoc* HSD de Tukey

- ▶ Não há diferença estatística relevante entre a base desbalanceada e o SMOTE ($p\text{-value} = 0.2073$);
- ▶ Significância entre o desbalanceamento e a geração de imagens ($p\text{-value} = 0.0062$);
- ▶ O melhor método para rebalancear essas classes é a geração de imagens utilizando o método de composição de quatro imagens.

Ranking dos métodos de rebalanceamento

Tabela 2: Resultados acumulados de todos os experimentos. Diferenças não significativas de F1-Score possuem o mesmo peso que as significantes.

Cenários de duas classes	Cenário Multiclasses	Todos
Aguçamento (16)	SMOTE (13)	Limiares (35)
Limiares (17)	Mistura (14)	Mistura (39)
Saliência (22)	Limiares (18)	SMOTE (42)
Todos (25)	Todos (22)	Aguçamento (43)
Mistura (25)	Saliência (24)	Saliência (46)
Composição 4 (29)	Aguçamento (27)	Todos (47)
SMOTE (29)	Composição 4 (28)	Composição 4 (57)
Composição 16 (37)	Ruído (28)	Composição 16 (69)
Borramento (44)	Borramento (29)	Borramento (73)
Ruído (47)	Composição 16 (32)	Ruído (75)
SMOTE Visual (47)	Desbalanceado (34)	Desbalanceado (86)
Desbalanceado (52)	SMOTE Visual (42)	SMOTE Visual (89)

F1-Scores de cada método

Tabela 3: Média ordenada pela coluna de todos os experimentos.

Métodos	Cenários de duas classes	Cenário Multiclasses	Todos
Aguçamento	84,473575	74,817654	80,182055
Limiares	84,332408	74,472474	79,950215
Saliência	84,172238	74,248645	79,761752
Composição 4	83,123738	74,667785	79,365537
Composição 16	82,977850	74,581912	79,246322
Mistura	83,124582	74,045887	79,089606
Borramento	82,164793	74,555320	78,782805
Desbalanceado	81,314335	74,551506	78,308633
Todos	82,358089	71,943914	77,729567
Ruído	81,179247	72,609639	77,370532
SMOTE	79,085501	73,329079	76,527091
SMOTE Visual	78,811666	70,439604	75,090749

Resultados — Fine tuning

- ▶ Modificar parâmetros dos métodos e verificar a diferença da acurácia com os parâmetros padrões;
- ▶ Ligeira melhora: entre 1% e 3%;
- ▶ No geral, possível superar os resultados do SMOTE.

Resultados — Rede de Convolução

[TODO]

Tabela 4: Treinamento de uma Rede Neural de Convolução das classes *Horse* e *Elephant* da COREL-1000.

Base	Medida F1
Original balanceada	
Desbalanceada	
Rebalanceada com geração de imagens artificiais	

Discussão

- ▶ Pode haver ganho estatístico do F1-Score ao gerar imagens (versus a geração de exemplos no espaço de atributos);
- ▶ Na maioria dos experimentos, a geração obteve significância estatística relevante quando comparada à base desbalanceada;
- ▶ Pode gerar novas informações relevantes para a classificação de imagens;
- ▶ Um estudo aprofundado de cada contexto pode indicar quais operações aplicar.

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses
- Contribuições

2 Quantização de imagens

- Contextualização
- Método

- Experimentos

3 Geração de imagens artificiais

- Contextualização
- Método
- Experimentos

4 Conclusões

- Trabalhos futuros

5 Agradecimentos

Conclusões — Quantização

- ▶ Alternativa/complemento à seleção de características;
- ▶ Redução de 256 para 64 produziu bons resultados;
- ▶ Utilizar quantização como um primeiro passo e então o LPP;
- ▶ Faz parte do pipeline: não aumenta o custo computacional do sistema e simplifica os passos subsequentes;
- ▶ Reduz a dimensão dos vetores de características de cor;
- ▶ Reduz o tempo de computação para os descritores de textura.

Conclusões — Geração de imagens artificiais

- ▶ Novas informações para a classificação de imagens;
- ▶ Melhorou a acurácia dos algoritmos de classificação, inclusive quando comparada à geração de exemplos no espaço de atributos;
- ▶ Provou contribuir com o balanceamento entre classes;
- ▶ A visualização do espaço (características das novas imagens e as resultantes da interpolação dos vetores) mostrou que a geração é capaz de ocupar uma região mais abrangente do espaço.

Artigo publicado

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

Neurocomputing 173 (2016) 385–396

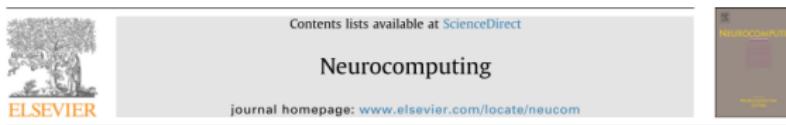


Image quantization as a dimensionality reduction procedure in color and texture feature extraction



Moacir Ponti*, Tiago S. Nazaré, Gabriela S. Thumé

Instituto de Ciências Matemáticas e de Computação Universidade de São Paulo São Carlos, 13566-590 SP Brazil

ARTICLE INFO

Article history:
Received 2 July 2014
Received in revised form
9 April 2015
Accepted 16 April 2015
Available online 1 September 2015

Keywords:
Image quantization
LPP
Compact features extraction

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.

© 2015 Elsevier B.V. All rights reserved.

Trabalhos futuros — Quantização

- ▶ Estudo da influência da quantização em outros métodos de extração de características;
- ▶ MSB se sobressaiu nos resultados do uso da quantização:
 - ▶ Variações desse método podem produzir melhores mapas de cor para a etapa de reconhecimento.

Trabalhos futuros — Geração de imagens

- ▶ Análise dos espaços encontrados para os diferentes métodos de geração;
- ▶ O impacto em extractores de características pode sugerir quais são as características latentes;
- ▶ Outros métodos para geração podem ser sugeridos;
- ▶ SMOTE visual pode ser realizado com imagens próximas ao invés de aleatórias.

Trabalhos futuros — Geração de imagens

- ▶ Estado da arte de extração e classificação de imagens corresponde ao uso de Rede Neural Convolucional (SCHMIDHUBER, 2014):
 - ▶ Aprender quais são as melhores características que diferenciam as classes de imagens;
 - ▶ Podem indicar possíveis operações para a geração de imagens artificiais.
- ▶ Analisar a memória associativa aprendida com uma máquina de Boltzmann restrita (FISCHER; IGEL, 2014):
 - ▶ Escolher para qual imagem original utilizar, ao invés do método aleatório utilizado nos resultados preliminares;
 - ▶ Verificação da relevância das imagens geradas.

Trabalhos futuros — Geração de imagens

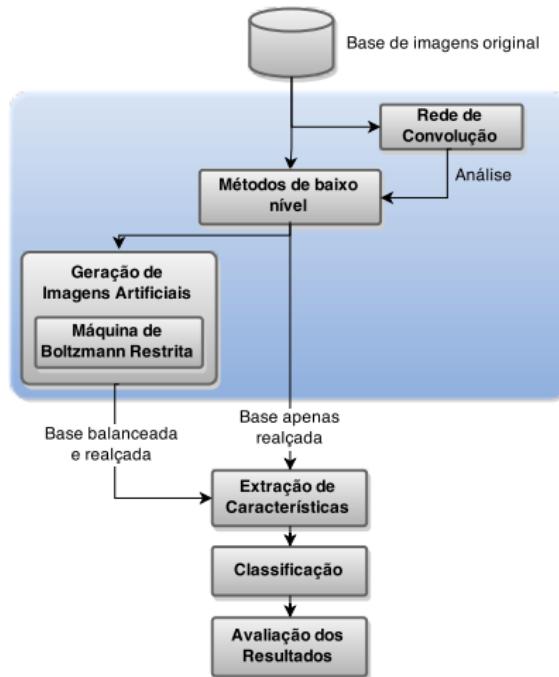


Figura 37: Trabalhos futuros.

Estrutura da Apresentação

1 Introdução

- Motivação
- Hipóteses
- Contribuições

2 Quantização de imagens

- Contextualização
- Método

● Experimentos

3 Geração de imagens artificiais

- Contextualização
- Método
- Experimentos

4 Conclusões

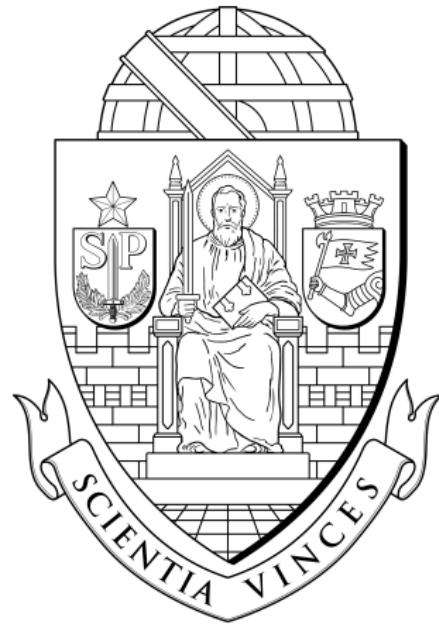
- Trabalhos futuros

5 Agradecimentos

Agradecimentos

Moacir Antonelli Ponti

João do Espírito Santo Batista Neto



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] Asja Fischer and Christian Igel. Training restricted Boltzmann machines: An introduction. *Pattern Recognition*, 47(1):25–39, January 2014.
- [6] Rafael C Gonzalez and Richard E Woods. *Digital Image Processing*. Prentice-Hall, 3 edition, 2007.
- [7] 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.
- [8] 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.
- [9] Nathalie Japkowicz and Shaju Stephen. The class imbalance problem : A systematic study. *Intelligent data analysis*, 6(5):429–449, 2002.
- [10] Christopher Kanan and Garrison W Cottrell. Color-to-grayscale: does the method matter in image recognition? *PloS one*, 7(1):e29740, January 2012.

Referências II

- [11] 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.
- [12] Aravindh Mahendran and Andrea Vedaldi. Understanding deep image representations by inverting them. arXiv preprint arXiv:1412.0035, 2014.
- [13] Michael A. Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.
- [14] 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.
- [15] 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.
- [16] Moacir Ponti, Tiago Nazaré, and Gabriela Thumé. Image quantization as a dimensionality reduction procedure in color and texture feature extraction. Submitted to Neurocomputing, 2014.
- [17] 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.
- [18] Karen Simonyan and Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. September 2014.
- [19] 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.
- [20] 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 quantização aplicadas a problemas de classificação

Gabriela Salvador Thumé

Orientador: Prof. Dr. Moacir Antonelli Ponti

Co-orientador: Prof. Dr. João do Espírito Santo Batista Neto

Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo

29 de abril de 2016

