

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
- Agradecimentos

Motivação

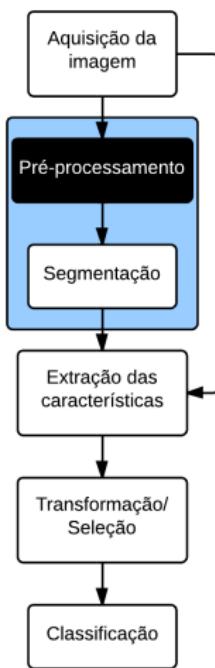
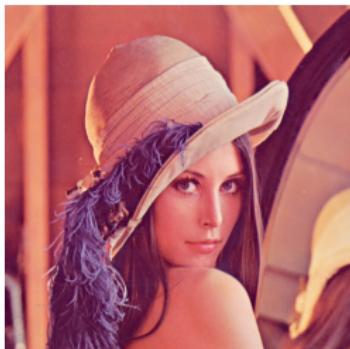


Figura 1: Etapas canônicas do reconhecimento de padrões.

Motivação

- ▶ Maior esforço ao operar no espaço de características já obtidas (extração manual);
- ▶ 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 I — Quantização



(a) Imagem colorida



(b) Imagem quantizada

- ▶ Redução do número de cores (Ex. 24 bits — 8 bits);
- ▶ Simplifica as imagens;
- ▶ Vetores de características mais compactos.

Motivação I — 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 II — 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 II — Rebalanceamento de classes

- ▶ *Synthetic Minority Over-sampling Technique* (SMOTE): geração de exemplos a partir de vetores de características originais;
- ▶ Não existem estudos dessas técnicas em dados de informação visual;
- ▶ 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 o desempenho da classificação, quando comparada à geração de exemplos artificiais no espaço de atributos.*

Contribuições I

- ▶ 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;
 - ▶ Tempo de processamento para os métodos de descrição de textura.

Contribuições II

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

Contribuições em código e reproduzibilidade

- ▶ Código contendo métodos de quantização, extração de características e redução de dimensionalidade:

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

- ▶ Código contendo métodos para a quantização, geração de imagens, SMOTE, visualização e classificação:

<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
- Agradecimentos

Quantização de imagens

- ▶ Métodos de extração estão preparados para receber imagens em um canal de cor;
- ▶ Características podem ser extraídas para cada canal de cor e posteriormente concatenadas;
- ▶ O pipeline inclui a conversão de imagens coloridas em imagens com apenas um canal;
- ▶ **Explorar a quantização para reduzir o tempo de processamento e obter vetores mais compactos.**

Quantização de imagens

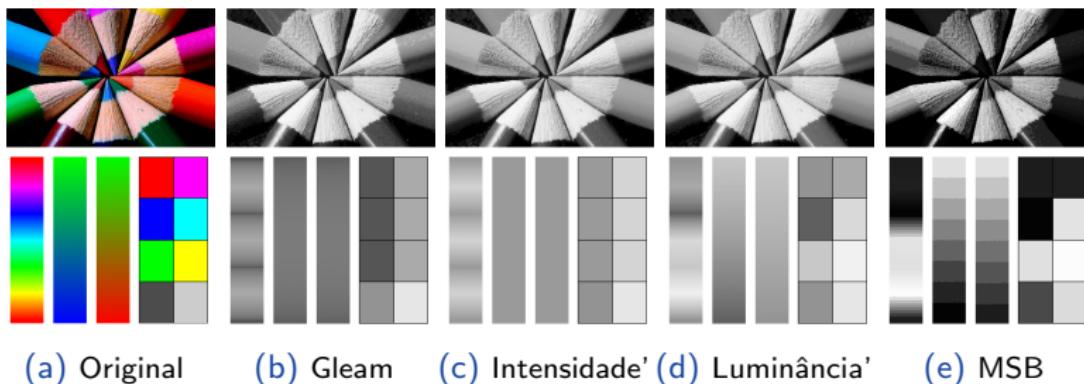


Figura 3: Imagens convertidas em escala de cinza.

- ▶ *Luminância'* e *MSB* (*Most Significant Bits*) apresentam uma melhor discriminação entre cores;
- ▶ A imagem quantizada possui 232 intensidades únicas para o *MSB* e 184 intensidades para os demais métodos.

Quantização de imagens

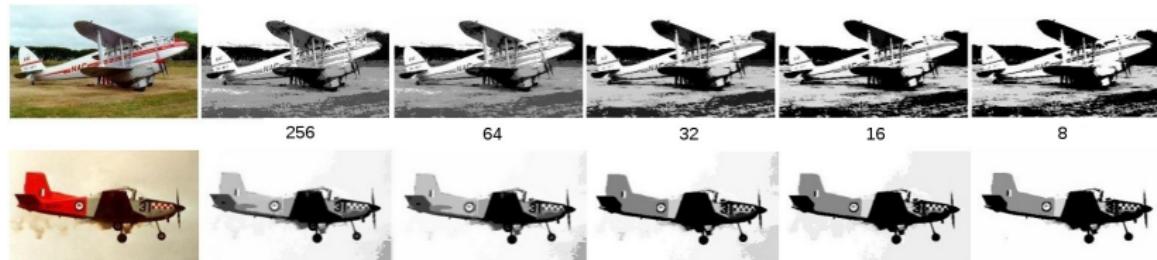


Figura 4: Imagens da base *Caltech101-600* quantizadas com MSB.

- ▶ Com 256 e 64 há preservação da informação;
- ▶ A partir de 32, ainda que seja possível discernir o objeto, há maior degradação, em particular no exemplo da primeira linha.

Experimentos

1. Utilizando um método de extração de características seguido pela classificação (sem posterior seleção de características);
2. Utilizando o vetor resultante da concatenação de todos os métodos de extração, seguido pela classificação com e sem a seleção de características.

Experimentos

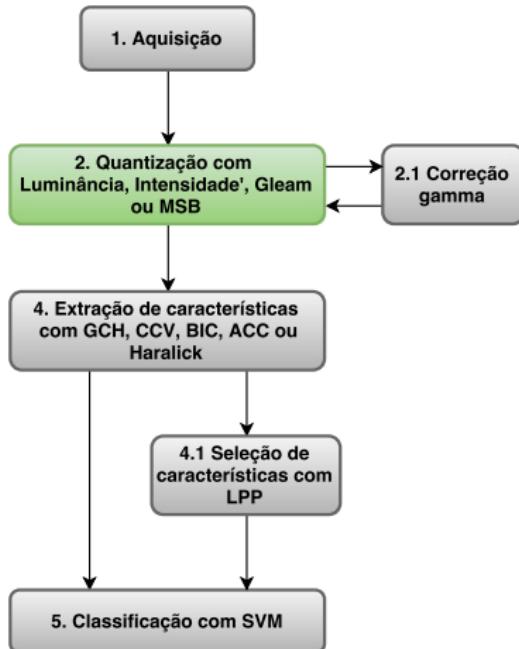
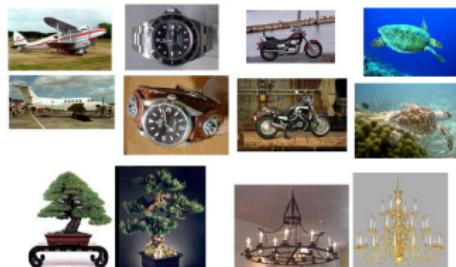


Figura 5: Fluxo das operações e os métodos utilizados nos experimentos.

Experimentos — Bases de Imagens



(a) Corel-1000



(b) Caltech101-600



(c) Produce-1400

Figura 6: Bases de imagens utilizadas nos experimentos de quantização.

Experimentos — Protocolo

1. **Quantização^a:** *Intensidade', Gleam, Luminância'* e MSB;
2. **Extração de características^b:**
 - ▶ *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^c:** *Locality Preserving Projections* (LPP) com $k = 10$ vizinhos;
4. **Classificação:** *Support Vector Machines* (SVM) com *grid search* no conjunto de treinamento.

^aKanan e Cottrell, 2012; Ponti e Escobar, 2013

^bPenatti et al., 2012

^cZhuo et al., 2014

Experimentos — Geral

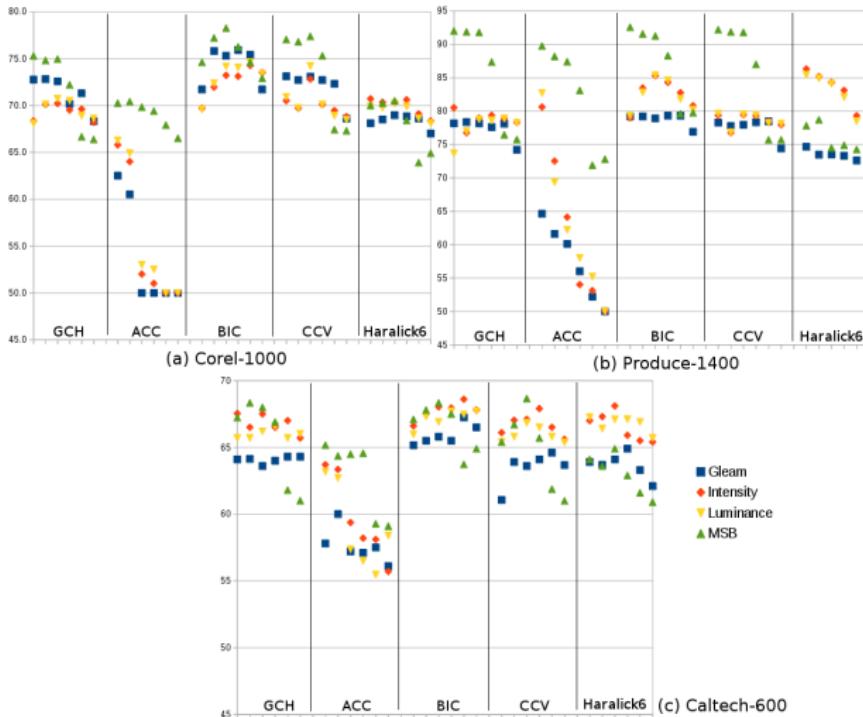


Figura 7: Acurácia média utilizando 256, 128, 64, 32, 16 e 8 cores.

Experimentos — Características de cor

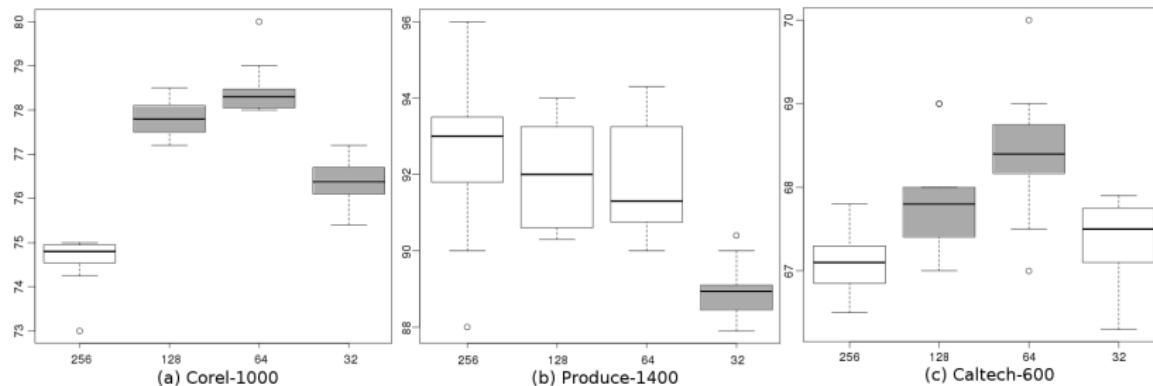


Figura 8: Acurácia média utilizando MSB e BIC.

- ▶ Boxplots em cinza correspondem às significâncias estatísticas com $p\text{-value} < 0.01$ quando comparado à 256 cores;
- ▶ Quantizar para 64 cores prova ser uma escolha interessante de pré-processamento.

Experimentos — Características de textura

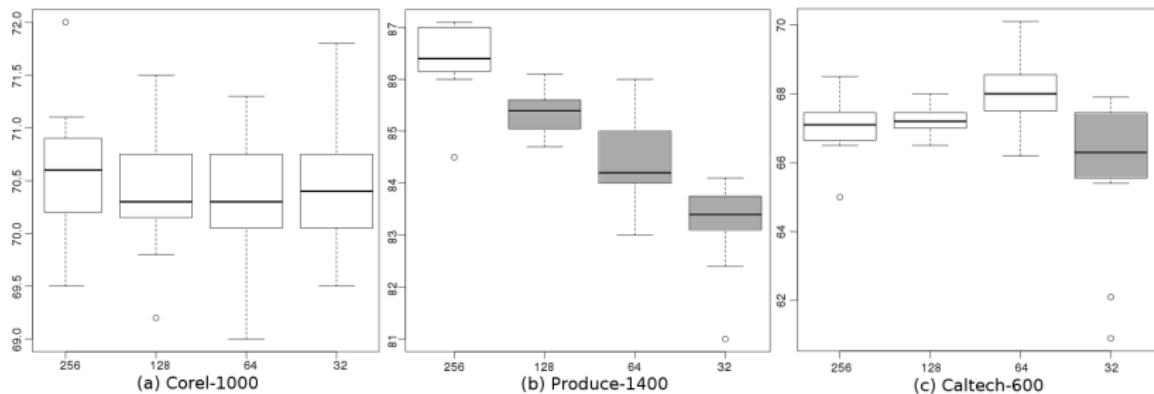


Figura 9: Acurácia média utilizando *Luminância'* e Haralick.

- ▶ Dependendo da base, a redução pode manter ou degradar a acurácia.

Experimentos — Quantização versus LPP

Análise do desempenho da quantização versus LPP:

1. Converter a imagem para escala de cinza com MSB em 256 cores;
2. Extrair as características com o método BIC;
3. Reduzir a dimensionalidade com o método LPP:
 - ▶ Versões reduzidas de 256, 128 e 64 dimensões.
4. Comparar a acurácia dos vetores com dimensionalidade reduzida.

Experimentos — Quantização versus LPP

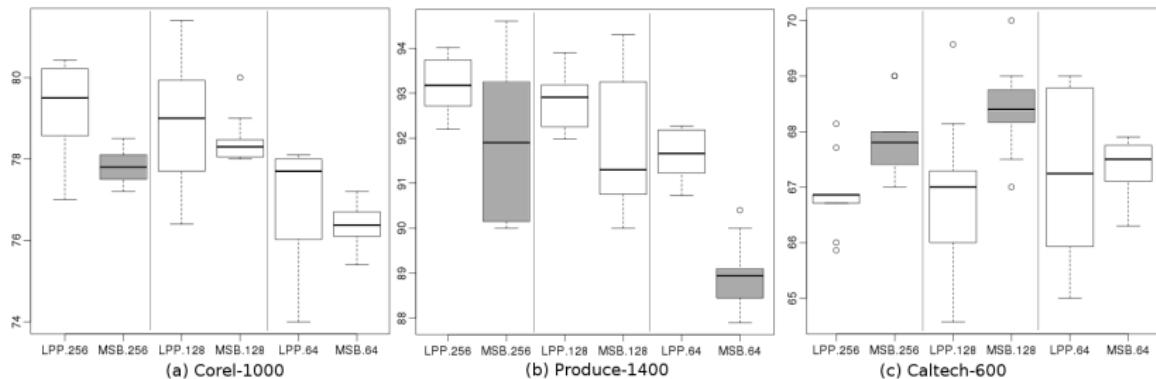


Figura 10: Acurácia comparando os métodos MSB e LPP.

- ▶ Comparação realizada com a mesma dimensionalidade;
- ▶ Quantizar imagens pode produzir resultados similares a projetar as características em um espaço de menor dimensão;
- ▶ Importante explorá-la para melhorar a discriminação entre as classes.

Experimentos — Concatenação dos vetores

1. Converter a imagem em escala de cinza com 256 cores;
2. Extrair as características com **todos os métodos**;
3. Concatenar as características em um vetor com $D = 2310$;
4. Analisar:
 - ▶ Redução da dimensionalidade com LPP para $d = 1160, 582, 294$ e 150 ;
 - ▶ A redução da quantização versus LPP para $d = 1160$ e 582 ;
 - ▶ Uso do LPP sobre os vetores quantizados para $d = 582$;

Experimentos — LPP

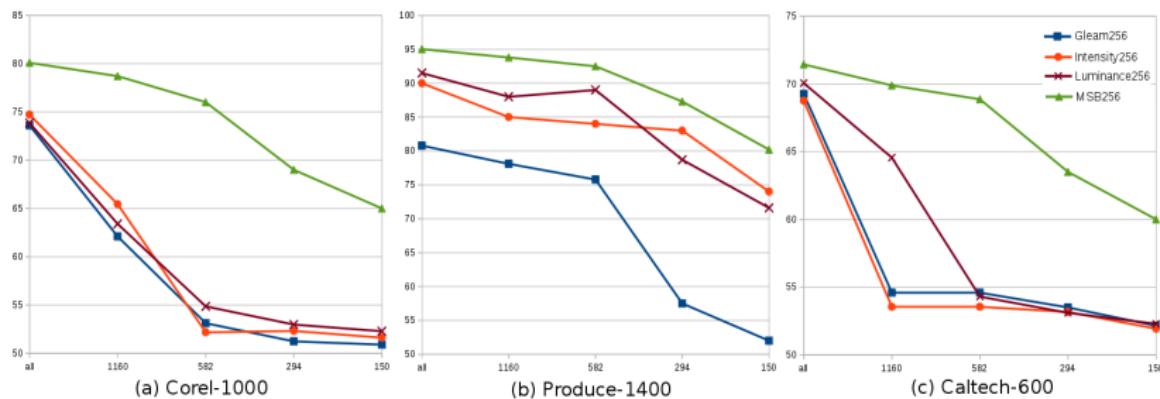


Figura 11: Acurácia para *Gleam*, *Intensidade*', *Luminância*' e *MSB*.

- ▶ O método de quantização MSB obteve melhor desempenho;
- ▶ Utilizar todos os vetores melhorou a acurácia em relação ao melhor descriptor individual.

Experimentos — Quantização versus LPP

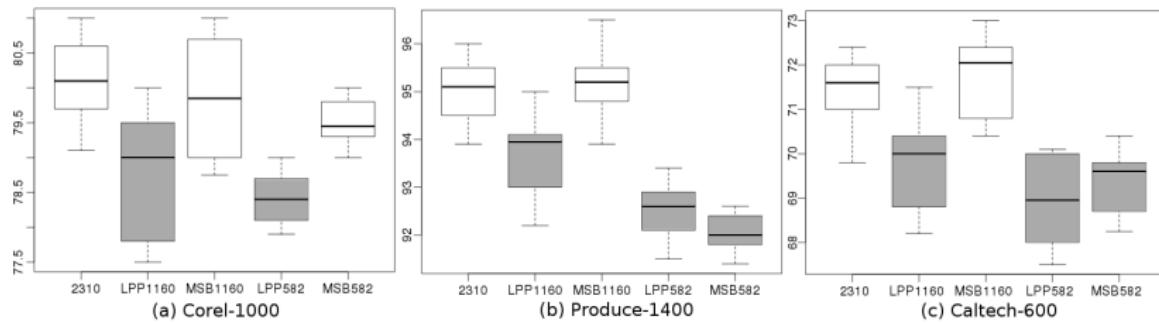


Figura 12: Comparação da redução de dimensionalidade com LPP e MSB.

- ▶ Utilizar 128 ou 64 cores é indicado como uma boa escolha do parâmetro de quantização.

Experimentos — Quantização + LPP

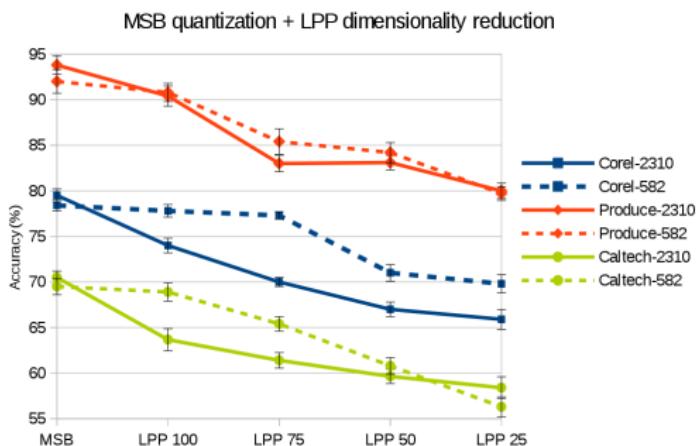


Figura 13: Redução da dimensionalidade com LPP sobre vetores quantizados.

- ▶ Usar apenas $d = 100$ dimensões provou manter as acuráciais usando 64 cores;
- ▶ Em geral, as projeções com LPP utilizando imagens quantizadas em 64 cores foram melhores (remoção de informações confusas).

Quantização — Discussão

1. Redução significativa da dimensionalidade:

- ▶ Enquanto melhora ou mantém a classificação do sistema;
- ▶ Inclusive quando comparado com métodos mais complexos.

2. Redução do tempo de execução:

- ▶ Vetor concatenado: $9C + 6$ dimensões;
- ▶ Tempo de extração das características: $f(N) = 42N + 6C^2$;
- ▶ Ex. 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
- Agradecimentos

Desbalanceamento de classes

- ▶ Número desbalanceado de exemplos:
 - ▶ Majoritárias x minoritárias.
- ▶ Abordagens comuns:
 1. *Pré-processamento ao reamostrar os dados:*
 - ▶ Aumentar a minoritária (em geral, melhores resultados);
 - ▶ Diminuir a majoritária.
 2. *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);
- ▶ Apresentam resultados piores, em geral:
 - ▶ Pode remover informações essenciais dos dados originais;
- ▶ Não há melhor método para todos os cenários.

Desbalanceamento de classes — Sobreamostragem

- ▶ Aumentar o número de elementos da classe minoritária.

SMOTE (Chawla et al., 2002)

1. Calcula a diferença entre o vetor de características de um elemento e do seu vizinho mais próximo;
2. Multiplica essa diferença por um número $0 \leq x \leq 1$;
3. Adiciona ao vetor original, criando um novo elemento entre os dois vetores originais;
4. Aprendido como exemplo da classe minoritária.

Desbalanceamento de classes — Sobreamostragem

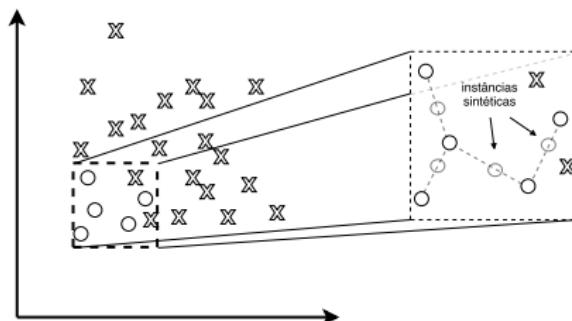


Figura 14: SMOTE

- ▶ Rebalancear ao gerar novos elementos, ao invés de replicá-los;
- ▶ **Diferentes estratégias para criar exemplos sintéticos podem melhorar a performance da classificação;**
- ▶ Utilizado como comparação.

Geração de imagens artificiais

- ▶ Compensar a baixa disponibilidade de exemplos de uma classe;
- ▶ Permitir a extração de informações antes não disponíveis nas imagens originais:
 - ▶ Combinação ou perturbação das imagens de entrada.
- ▶ Utilizado como entrada para métodos de extração ou CNNs (*Data Augmentation*):
 - ▶ SMOTE opera no espaço de características já extraídas.

Geração de imagens artificiais

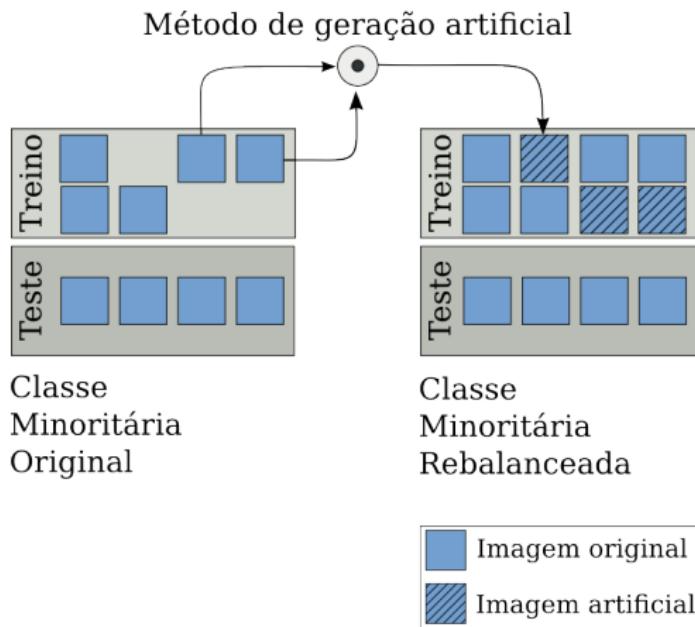


Figura 15: Geração de imagens artificiais da classe minoritária para rebalancear a base de imagens.

Geração de imagens artificiais — Borramento



(a) Original

(b) Imagem artificial

Figura 16: Filtro bilateral.

- ▶ A nova imagem possui detalhes borradados, porém com bordas preservadas.

Geração de imagens artificiais — Aguçamento



(a) Original



(b) Imagem artificial

Figura 17: *Unsharp mask*.

- ▶ A imagem resultante apresenta saliência nas transições de intensidade.

Geração de imagens artificiais — Adição de ruído



(a) Original



(b) Imagem artificial

Figura 18: 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 19: 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 20: 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 21: Mistura de *thresholds* de OTSU.

- ▶ A imagem resultante 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 22: Combinação de saliência obtida com GMR (*Graph-Based Manifold Ranking*).

- ▶ A imagem resultante (c) apresenta a região saliente de (b) sobreposta

Geração de imagens artificiais — Composição



(a) Imagem artificial



(b) Imagem artificial

Figura 23: Várias imagens, dispostas em um mosaico.

- ▶ Mistura de informações de várias imagens;
- ▶ Cada célula do mosaico sofre uma operação, sorteada no momento da geração da imagem.

Experimentos

Comparar o desempenho da classificação:

1. Base original (desbalanceada);
2. Geração de imagens artificiais com os métodos anteriores;
3. Técnica de sobreamostragem SMOTE.

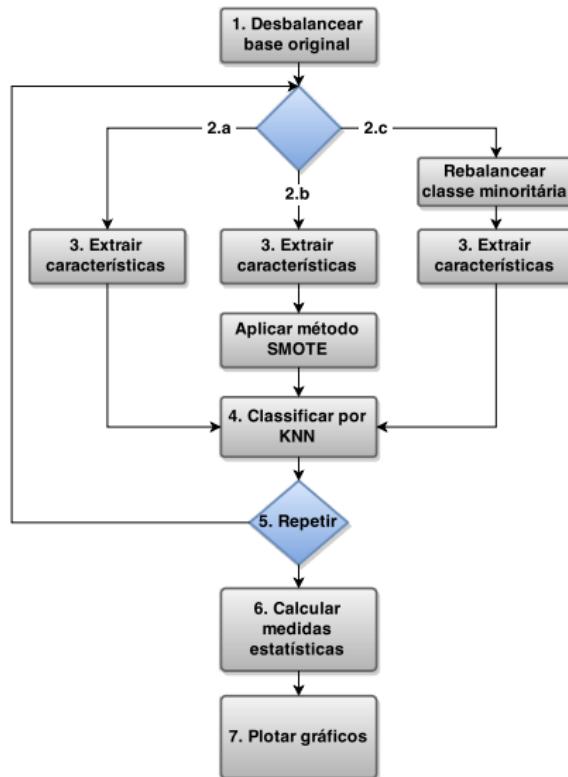


Figura 24: Fluxo de operações para obtenção dos resultados do rebalanceamento de classes.

Experimento — Visualização

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

Experimento de Visualização — Protocolo

1. **Imagens originais:** *Horse* e *Elephant* com 100 imagens cada;
2. **Desbalanceamento:** cada classe foi dividida em 50% para treino e 50% para teste, de maneira aleatória. Após, a classe *Horse* é desbalanceada, removendo 88% do conjunto de treino aleatoriamente.
3. **Método para geração de imagens:** *mistura ponderada*;
4. **Conversão em escala de cinza:** *Intensidade'*;
5. **Extração de características:** BIC;
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

Imagens originais



(a) Elephant



(b) Horse

Figura 25: Originalmente da base de imagens Corel-1000.

- Diferença de cores, contendo pequeno grau de sobreposição.

Experimento – Visualização

Método para geração de imagens



(a) Original



(b) Original



(c) Imagem artificial

Figura 26: Mistura ponderada.

Experimento — Visualização

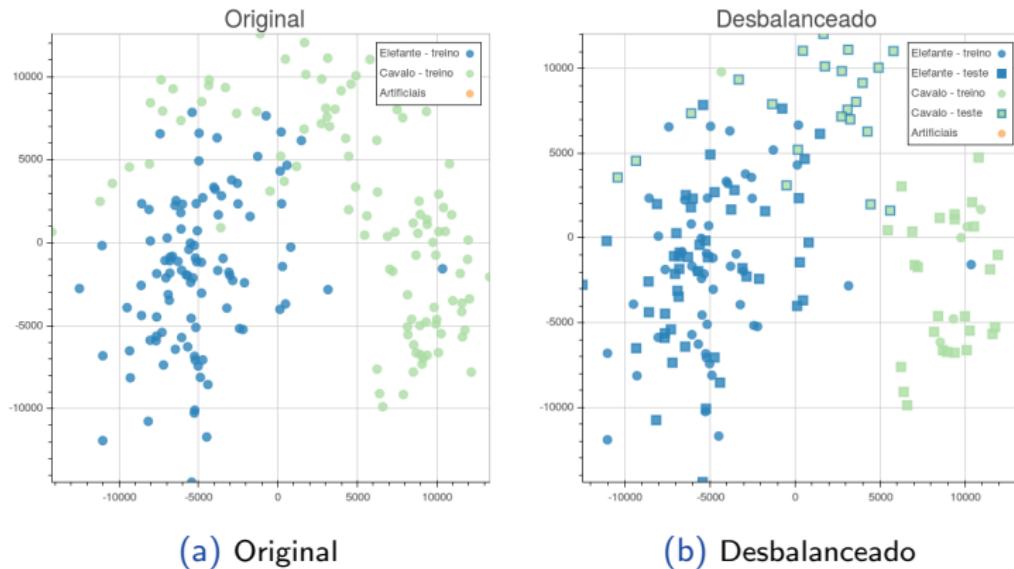


Figura 27: Projeção dos dois componentes principais obtidos com o PCA.

- À direita, as classes após a remoção de 88% do treino da classe *Horse*.

Experimento — Visualização

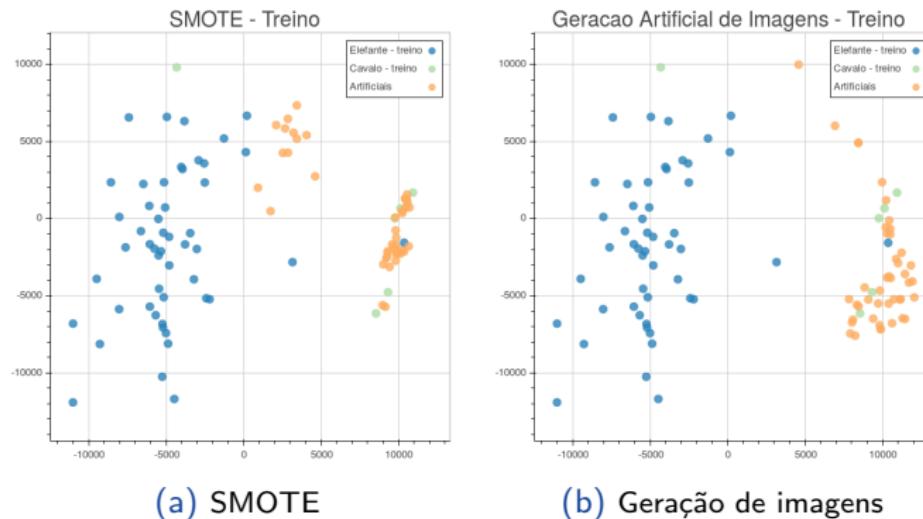


Figura 28: Novos exemplos de treinamento.

- ▶ Projeção no plano da base original desbalanceada;
- ▶ O método SMOTE gerou exemplos próximos a uma linha reta;
- ▶ A geração de imagens proporcionou mais dispersão.

Experimento — Visualização

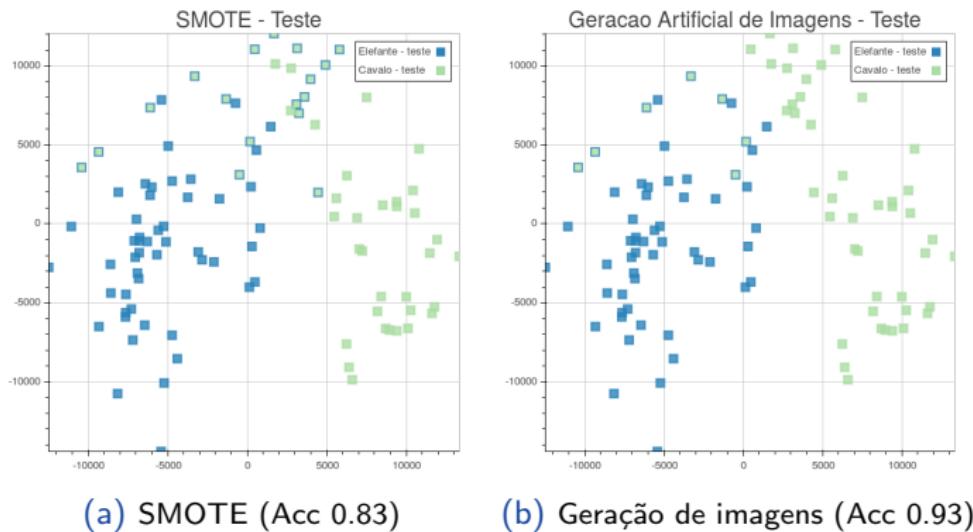
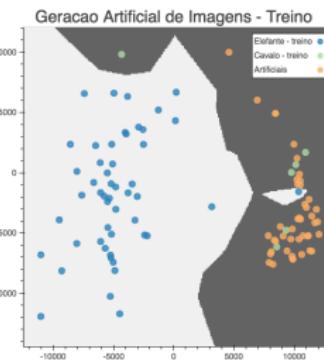
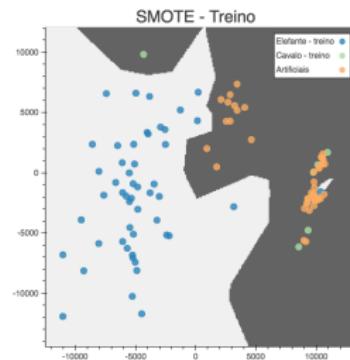
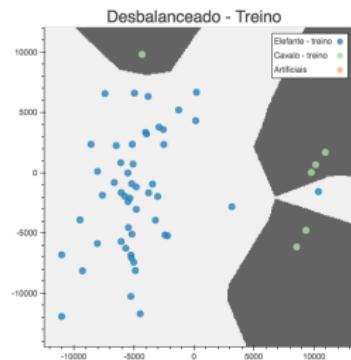


Figura 29: Resultado do treinamento com a base rebalanceada.

- ▶ A cor da borda dos marcadores representa a classe predita pelo classificador;
- ▶ A melhoria corresponde ao aumento da acurácia.

Experimento — Visualização



(a) Desbalanceado

(b) SMOTE

(c) Geração de imagens

Figura 30: Região de decisão com 1-NN.

- ▶ A classe minoritária apresenta-se melhor representada;
- ▶ SMOTE ocasionou uma certa invasão do espaço de características da classe majoritária.

Experimento — Visualização

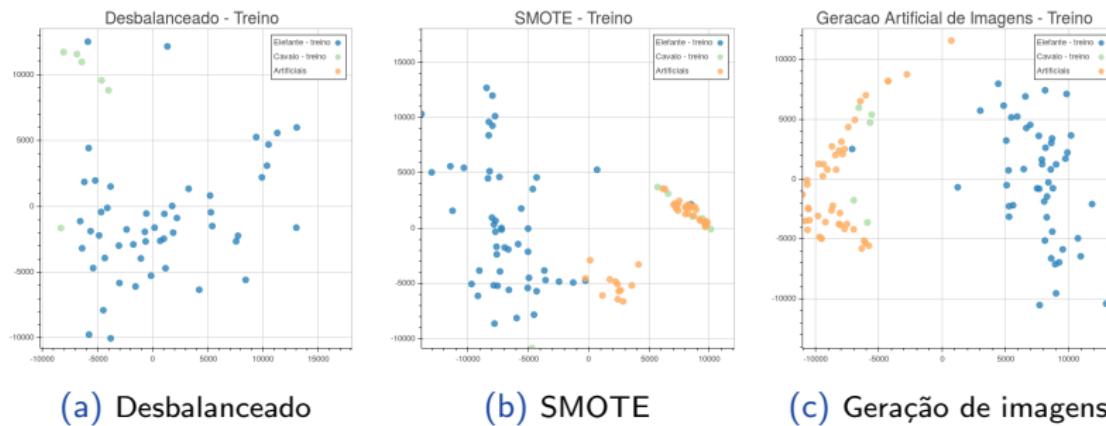


Figura 31: Melhores subespaços encontrados (componentes recalculadas).

- ▶ A geração de imagens artificiais proporcionou a criação de um subespaço que melhor discretiza as classes.

Experimento — Visualização

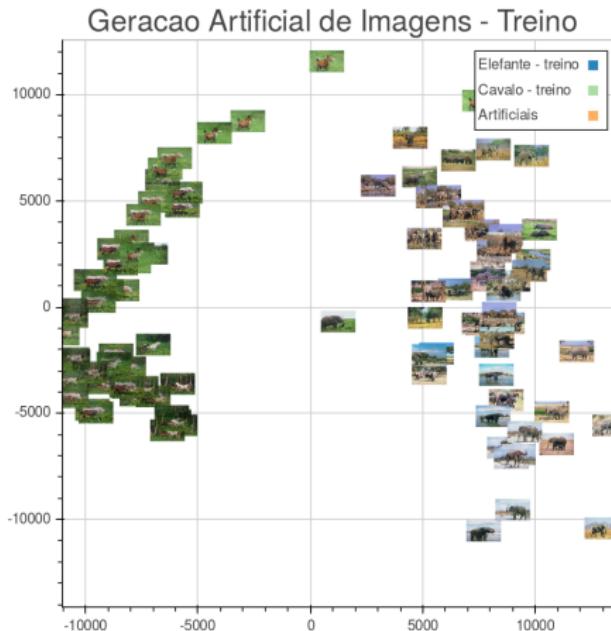


Figura 32: Impacto do método de extração de características BIC na separação entre classes.

Experimento — Visualização

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

Experimentos — Validação

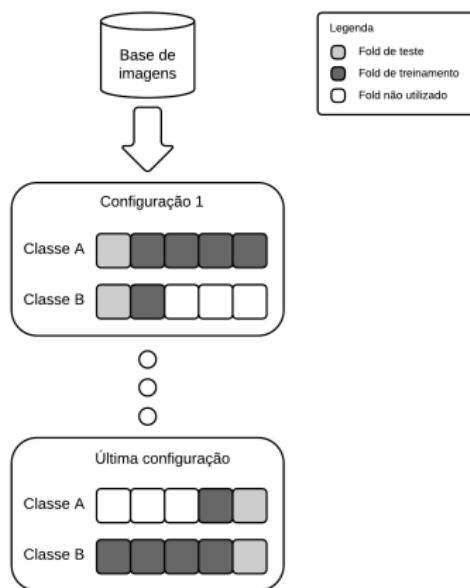


Figura 33: Validação k -fold com o objetivo de prover mais robustez ao sistema.

Experimentos — Protocolo

1. **Quantização^a:** *Intensidade', Gleam, Luminância', Luma e MSB;*
2. **Extração de características^b:**
 - ▶ *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;
 - ▶ *Histogram of Oriented Gradients^c* (HOG): 16x16 células;
 - ▶ *Local Binary Patterns* (LBP) com padrões uniformes^d: janelas 2x2.
3. **Classificação:** KNN com $K = 1$;
4. **Medida estatística:** *F1-Score.*

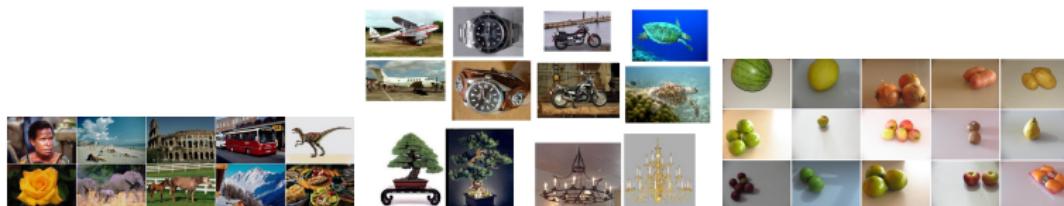
^aKanan e Cottrell, 2012; Ponti e Escobar, 2013

^bPenatti et al., 2012

^cDalal e Triggs, 2005

^dOjala et al, 2002

Experimentos — Bases de Imagens



(a) Corel-1000

(b) Caltech101-600

(c) Produce-1400



(d) Cifar-10

(g) Flickr 15k

(h) ImageNet

Figura 34: Bases de imagens utilizadas nos experimentos de geração.

Experimentos — Resultados

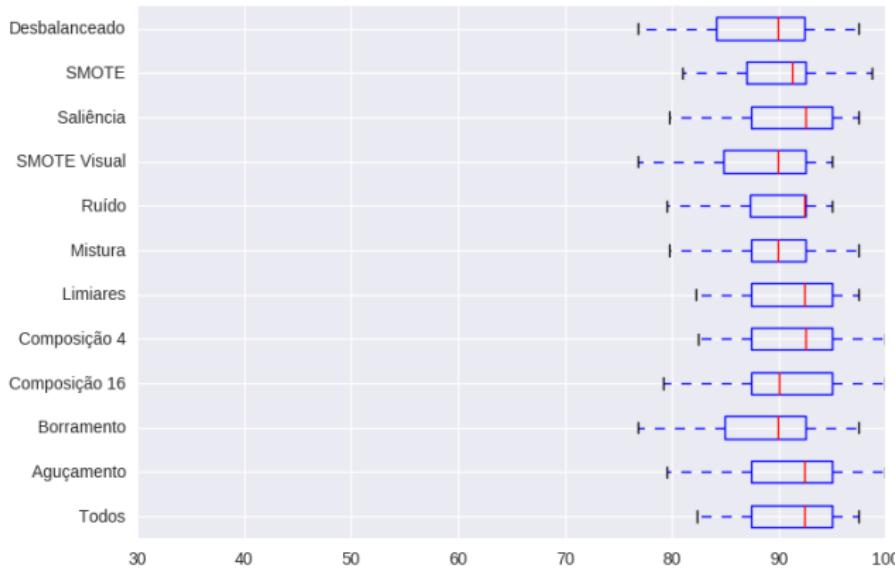


Figura 35: Combinação com melhor F1-Score para *Horse* e *Elephant*: Gleam e ACC.

Experimentos — Resultados

Tabela 1: A geração com *Composição 4* obteve maior valor de *F1-Score*.

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

Experimentos — Resultados



(a) Imagem gerada

Figura 36: Composição de quatro imagens da classe *Elephant*.

Experimentos — Resultados

Teste post-hoc HSD (*Honest Significant Difference*) 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.

Experimentos — *F1-Scores* de cada método

Tabela 2: 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

Experimentos — Média da melhoria para cada método

Tabela 3: Média da diferença entre o método de rebalanceamento e a base desbalanceada, considerando todos os experimentos.

Métodos	Diferença média da melhoria	Desvio Padrão
Aguçamento	16,860795	2,206283
Limiares	14,774237	2,190548
Saliência	13,078076	2,544774
Composição 4	9,512135	1,525219
Composição 16	8,439199	1,351292
Mistura	7,028758	2,306720
Borramento	4,267550	0,819487
Desbalanceado	0,000000	0,000000
Todos	-5,211596	5,922549
Ruído	-8,442906	4,059528
SMOTE	-16,033877	9,002886
SMOTE Visual	-28,960951	6,581923

Experimentos — Fine tuning

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

Experimentos — Rede de Convolução

Tabela 4: Treinamento de uma Rede Neural de Convolução.

Base	Acurácia média	<i>Elephant</i>	<i>Horse</i>
Original balanceada	0.96 ± 0.011	20 Teste — 80 Treino	20 Teste — 80 Treino
Desbalanceada	0.91 ± 0.009	20 Teste — 20 Treino	20 Teste — 80 Treino
Geração de imagens	0.97 ± 0.020	20 Teste — 80 Treino	20 Teste — 80 Treino

Geração de imagens — 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;
- ▶ Pode ser utilizado como *Data Augmentation* em CNNs;
- ▶ 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
- Agradecimentos

Conclusões — Quantização

- ▶ Alternativa/complemento à seleção de características;
- ▶ Redução de 256 para 64 cores produziu resultados significativos;
- ▶ Aplicar o LPP em imagens quantizadas com 64 cores obteve melhores resultados do que em 256;
- ▶ 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 mostrou que a geração é capaz de ocupar uma região mais abrangente do espaço.

Artigo publicado

Ponti, Moacir, Tiago S. Nazaré, and Gabriela S. Thumé. "Image quantization as a dimensionality reduction procedure in color and texture feature extraction." *Neurocomputing* 173 (2016): 385-396.

Neurocomputing 173 (2016) 385–396



Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom



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):
 - ▶ Verificação da relevância das imagens geradas.

Trabalhos futuros — Geração de imagens

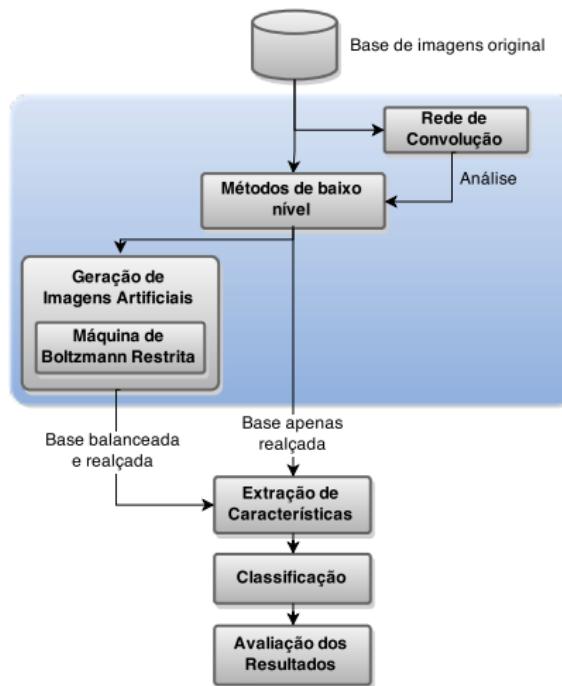


Figura 37: Sugestão de extensão do *pipeline* atual para trabalhos futuros.

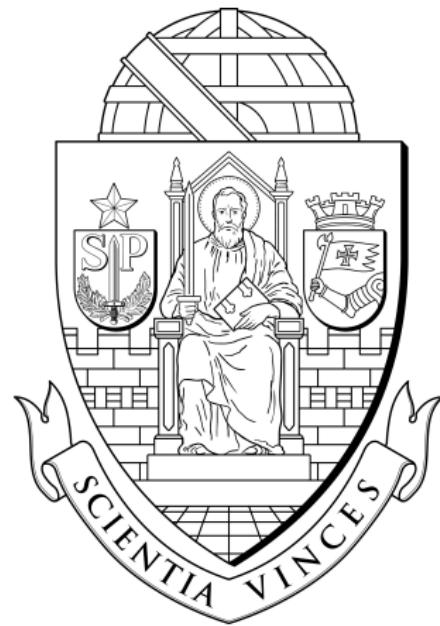
Agradecimentos

Moacir Antonelli Ponti

João do Espírito Santo Batista Neto



*Conselho Nacional de Desenvolvimento
Científico e Tecnológico*



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.
- [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.

Referências II

- [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

