

Deteccção de Imagens Parecidas

TRABALHO DE VISÃO COMPUTACIONAL

Grupo 02

- Bruno Santos
- Tiago Rodrigues

Sumário da Apresentação

- Introdução
- Revisão Bibliográfica
- Metodologia
- Experimentos
- Resultados
- Conclusões
- Trabalhos Futuros

Introdução

- Imagens parecidas
 - Imagens iguais ou muito próximas
 - Definição subjetiva, dependendo de quais variações fotométricas e geométricas são aceitáveis



Introdução

- Importância de se detectar imagens parecidas
 - Máquinas de Busca
 - Direitos Autorais
 - Spam em Imagens



Introdução

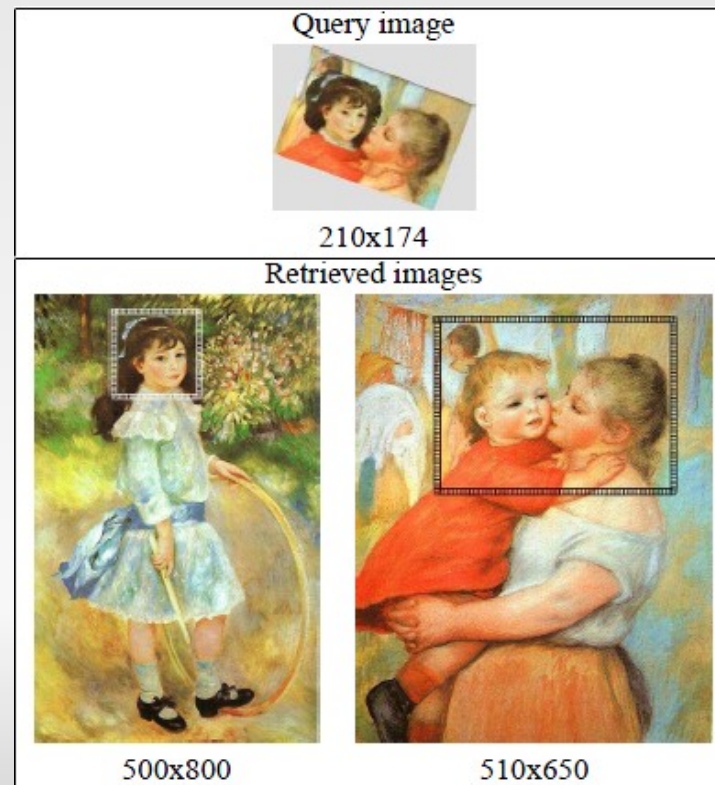
- Agrupadas em 3 categorias na literatura:
 - Cena: diferem na adição, oclusão ou movimento de objetos em primeiro plano
 - Câmera: variações do dispositivo de captura, como ângulo de visão
 - Imagem: variações devido a operações de edição digital, como cor, contraste, alteração da resolução
- Desafios para detecção em larga escala
 - Representação das imagens: eficiente e compacta
 - Alta taxa de acerto na detecção
 - Eficiência
- Problema amplamente estudado na literatura

Trabalhos Relacionados

- Artigos selecionados:
 - *Efficient Near-duplicate Detection and Sub-image Retrieval*
 - *Detection of Near-duplicate Images for Web Search*
 - *Clustering Near-duplicate Images in Large Collections*
 - *Near Duplicate Image Detection: min-Hash and tf-idf Weighting*
 - *Detecting Image Spam using Visual Features and Near Duplicate Detection*

Efficient Near-duplicate Detection and Sub-image Retrieval

- **Autores:** Yan Ke, Rahul Sukthankar, Larry Huston.
- **Publicação:** ACM International Conference on Multimedia, 2004 (MM 04).
- Para comparar os dados da query com as features extraídas da imagem, primeiro escala e rotaciona os pontos de interesse da query de forma que eles fiquem com a mesma escala e rotação da imagem pesquisada de forma a realizar a verificação geométrica das features.



Detection of Near-duplicate Images for Web Search

- **Autores:** Jun Jie Foo, Justin Zobel, Ranjan Sinha, S. M. M. Tahaghoghi.
- **Publicação:** ACM International Conference on Image and Video Retrieval, 2007 (CIVR 07)
- Comparação entre PCA-SIFT e Dynamic Partial Functions contra métodos que utilizam uma combinação de descritores locais e "hash-based counting". Além da comparação é feita uma pesquisa a cerca das variações que podem ser encontradas em imagens parecidas na web.

Alteration Type	Celebrities	Kids-related	Cars	Heroes	Deceased	Total instances
original	100	96	99	97	118	510
combination	62	59	8	110	152	391
scale	72	42	16	95	143	368
crop/zoom	87	24	21	69	153	354
P.I.P.	15	13	6	43	57	134
contrast (+/-)	13	4	3	18	54	92
border/frame	6	2	3	18	56	85
grayscale	2	2	1	1	27	33
recolouring	1	0	2	7	19	29
mirror	0	1	1	1	7	10
rotate	0	1	5	0	0	6

Clustering Near-duplicate Images in Large Collections

- **Autores:** Jun Jie Foo, Justin Zobel, Ranjan Sinha.
- **Publicação:** ACM SIGMM International Workshop on Multimedia Information Retrieval, 2007 (MIR 07)
- Utiliza técnicas de agrupamento combinado com a técnica do SIFT e uma adaptação de técnicas de agrupamento de textos quase-duplicados. Escolheram o PCA-SIFT pois esse se mostrou mais compacto e com acurácia mais alta. A metodologia sugerida no trabalho é bem interessante, combinando técnicas de visão computacional com métodos de agrupamento de textos.

Near Duplicate Image Detection: min-Hash and tf-id weighting

- **Autores:** Ondřej Chum, James Philbin, Andrew Zisserman.
- **Publicação:** British Machine Vision Conference, 2008 (BMVC 08).
- Representa a imagem por um conjunto esparsos de “visual words”. A semelhança é medida pela interseção dos conjuntos. A vantagem desta escolha de representação de imagem e medida de semelhança é que é possível obter um resultado bem eficiente. Outras medidas de similaridade utilizadas são:
 - Representar a imagem por um conjunto de “visual words” com pesos. Cada “visual words” tem um peso.
 - Extender a medida de similaridade para computar um histograma ponderado de interseção, o qual é capaz de levar em consideração a frequência do termo.

Query image:



Results

Set overlap, weighted set overlap, weighted histogram intersection

Detecting Image Spam using Visual Features and Near Duplicate Detection

- **Autores:** Bhaskar Mehta, Saurabh Nangia, Manish Gupta.
- **Publicação:** International World Wide Conference, 2008 (WWW 08)
- Algumas características extraídas:
 - Média RGB;
 - Histograma de cor;
 - “Color moment” : distribuição de cores na imagem é tratada como uma probabilidade de distribuição;
 - Frequência de vértices;
- Para comparação utiliza o classificador Support Vector Machine (SVM).

Metodologia

- Nova proposta baseada nos trabalhos relacionados lidos
- Criação da base de dados
- Extração de características das imagens
- Uso de classificação como mecanismo de detecção

- Imagens de uma base criada para o *The 2006 PASCAL Visual Object Classes Challenge* (VOC2006), disponível na Web.
- Composta por 5.304 imagens coletadas de fotógrafos profissionais, do Flickr e da base de dados da Microsoft Research Cambridge.
- Images de: bicicletas, ônibus, gatos, carros, vacas, cachorros, cavalos, motocicletas, pessoas e ovelhas.
- Formato: PNG

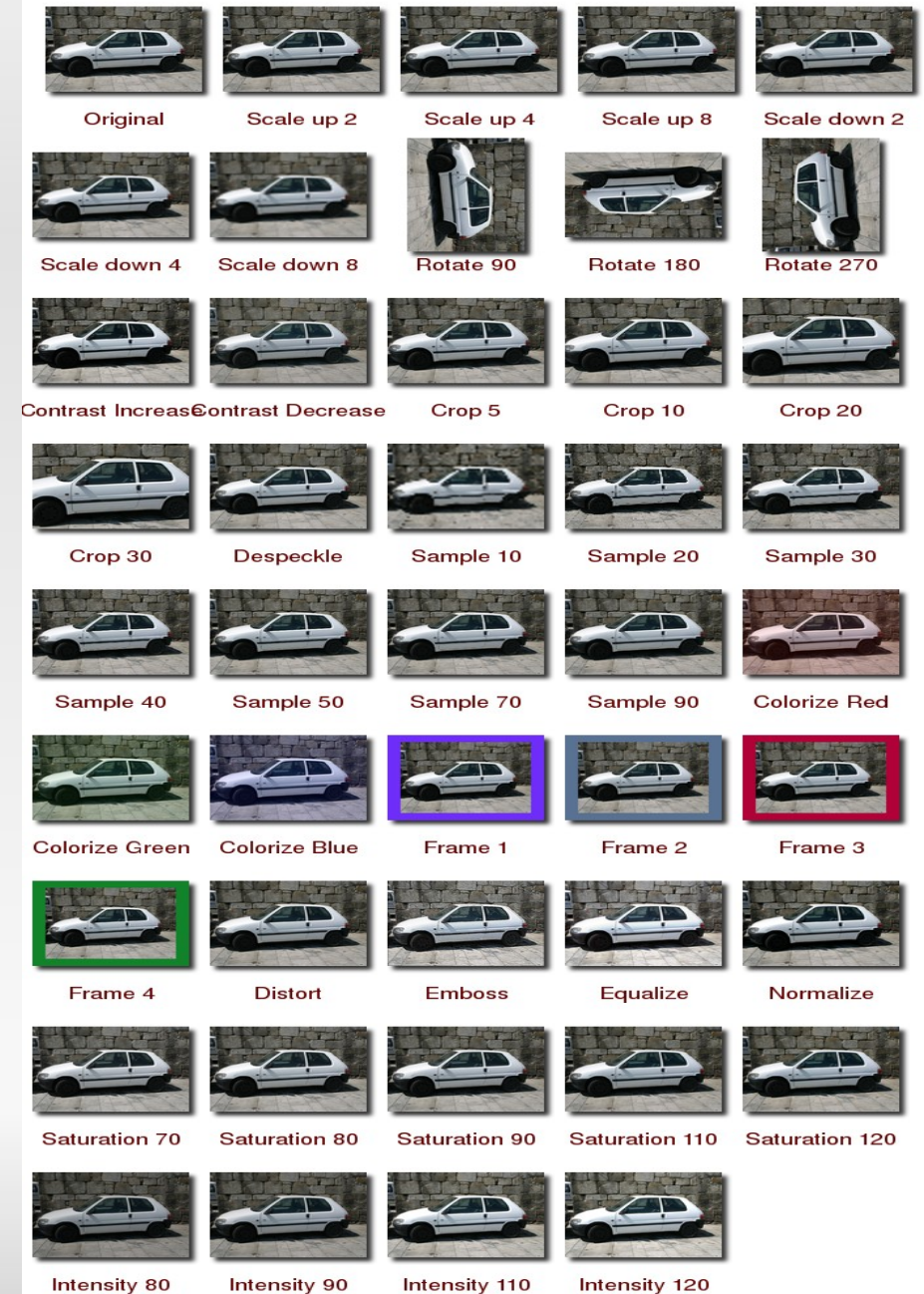
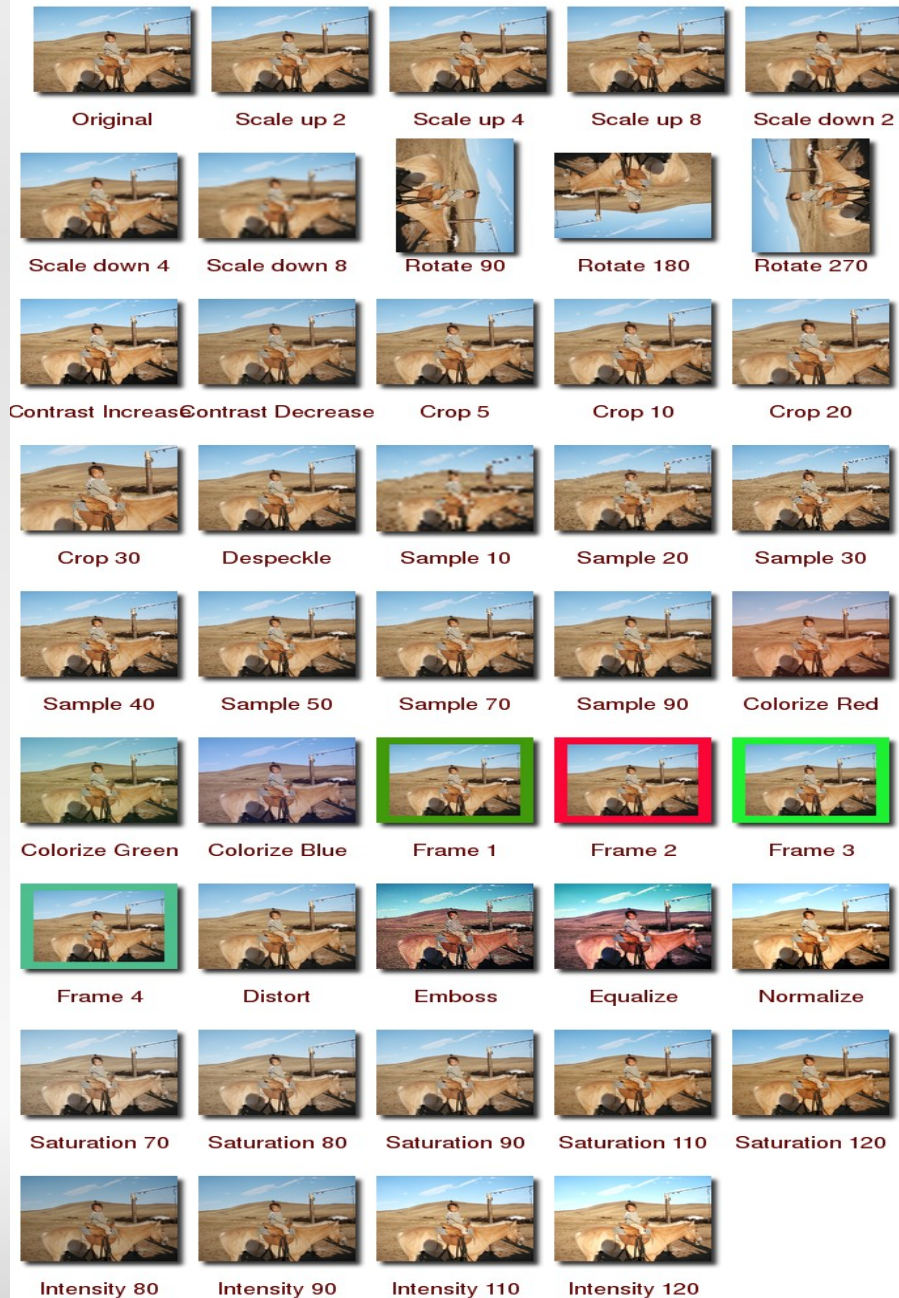
Metodologia

Criação da base de dados

- 10 imagens selecionadas aleatoriamente, aplicando 43 transformações, seguindo metodologia do artigo "*Efficient Near-duplicate Detection and Sub-image Retrieval*":
- Programa utilizado: **Image Magick**
- Colorizing: aumentar em 10% os canais RGB
- Contrast: aumentar e diminuir
- Cropping: 5%, 10%, 20% e 30% preservando a região central e voltando para o tamanho original
- Downsampling: 10%, 20%, 30%, 40%, 50%, 70% e 90%
- Framing: adicionar uma moldura de espessura 10% do tamanho com cores aleatórias
- Rotation: 90°, 180° e 270°
- Saturation: 70%, 80%, 90%, 110% e 120%
- Scaling: aumenta e diminui a imagem em 2, 4 e 8 vezes, voltando ao tamanho original
- Intensity: 80%, 90%, 110% e 120%
- Despeckle
- Distort
- Emboss
- Equalize
- Normalize

Metodologia

Criação da base de dados



Metodologia

Extração de características das imagens

- Objetivo: criar vetor que representa a imagem (impressão digital)
- Algumas características extraídas seguindo metodologia do artigo *"Detecting Image Spam using Visual Features and Near Duplicate Detection"*
- MATLAB
 - Número de Cores Únicas
 - RGB: média em cada canal
 - Color Histogram: 64 bins
 - Color Moment RGB: mean, standard deviation e skewness para cada canal
 - HSV: número médio em cada canal
 - Edge: quantidade de pixels nas bordas detectadas usando Canny
 - Corners: quantidade de corners detectados usando Harris
 - Texture First Order Statistics: mean, variance, skewness, kurtosis, median e mode

Metodologia

Extração de características das imagens

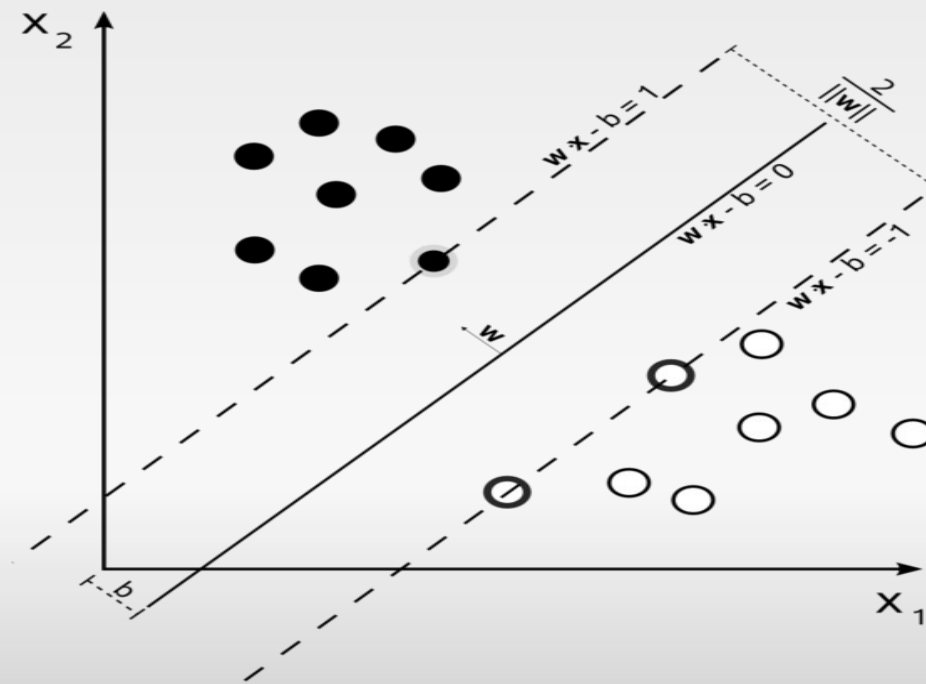
- Total: 88 características de cada imagem

```
1 1:22209 2:28432.1039041916 3:28678.2268023952 4:26316.5353053892 5:1646.0000000000 6:1136.0000000000 7:946.0000000000 8:1013.0000000000 9:1064.0000000000
6 1:21692 2:30712.2395605469 3:31202.1388802083 4:30148.2991829427 5:238.0000000000 6:2406.0000000000 7:7073.0000000000 8:6660.0000000000 9:6726.0000000000
4 1:187401 2:35795.9391093333 3:36668.3832480000 4:32218.2303413333 5:0.0000000000 6:0.0000000000 7:0.0000000000 8:0.0000000000 9:0.0000000000 10:0.0000000000
3 1:306070 2:26691.3554524740 3:26205.6380143229 4:17565.7398470052 5:0.0000000000 6:3.0000000000 7:42.0000000000 8:134.0000000000 9:378.0000000000 10:628.0000000000
8 1:162120 2:34177.4668708709 3:34297.8577237237 4:33979.2649969970 5:0.0000000000 6:0.0000000000 7:0.0000000000 8:2.0000000000 9:687.0000000000 10:4476.0000000000
5 1:57525 2:25942.5075000000 3:25145.3741471354 4:9959.7404947917 5:6.0000000000 6:84.0000000000 7:464.0000000000 8:1060.0000000000 9:1161.0000000000 10:1.0000000000
8 1:31199 2:27341.9225585586 3:27438.2351351351 4:27183.3528288288 5:379.0000000000 6:200.0000000000 7:306.0000000000 8:1047.0000000000 9:7752.0000000000 10:1.0000000000
7 1:108250 2:22782.2556733871 3:23962.2415403226 4:22055.0965766129 5:179.0000000000 6:408.0000000000 7:649.0000000000 8:1430.0000000000 9:2868.0000000000 10:1.0000000000
7 1:108734 2:24049.0178830645 3:24907.7281532258 4:23504.0402096774 5:84.0000000000 6:273.0000000000 7:455.0000000000 8:1031.0000000000 9:2089.0000000000 10:1.0000000000
10 1:72440 2:33013.7355293843 3:33029.7485657649 4:33067.1157921331 5:1498.0000000000 6:2962.0000000000 7:4928.0000000000 8:4035.0000000000 9:4028.0000000000 10:1.0000000000
1 1:47932 2:32585.7570750751 3:32685.7825885886 4:30036.6698558559 5:5758.0000000000 6:4437.0000000000 7:3143.0000000000 8:3139.0000000000 9:3767.0000000000 10:1.0000000000
10 1:72440 2:29552.7557719216 3:29729.2730449316 4:28599.0374455846 5:0.0000000000 6:837.0000000000 7:5760.0000000000 8:8286.0000000000 9:5182.0000000000 10:1.0000000000
6 1:14178 2:25809.8005729167 3:26268.2343619792 4:25413.1270963542 5:77.0000000000 6:1104.0000000000 7:2022.0000000000 8:2061.0000000000 9:2140.0000000000 10:1.0000000000
9 1:53081 2:25962.0504850261 3:22945.3272623698 4:20766.3571126302 5:253.0000000000 6:1722.0000000000 7:1044.0000000000 8:1319.0000000000 9:1658.0000000000 10:1.0000000000
5 1:55758 2:32139.1231250000 3:31043.0284049479 4:10207.2458398437 5:3.0000000000 6:51.0000000000 7:200.0000000000 8:688.0000000000 9:985.0000000000 10:95.0000000000
4 1:43987 2:35795.9249493333 3:36668.4433760000 4:32218.2724960000 5:0.0000000000 6:0.0000000000 7:0.0000000000 8:0.0000000000 9:0.0000000000 10:0.0000000000
1 1:52996 2:36386.5870969089 3:25057.8413784461 4:31320.9560693400 5:6531.0000000000 6:4606.0000000000 7:3868.0000000000 8:4160.0000000000 9:4404.0000000000 10:1.0000000000
6 1:23749 2:25802.7590071614 3:26244.0184700521 4:25416.3136621094 5:307.0000000000 6:4434.0000000000 7:7960.0000000000 8:8313.0000000000 9:8469.0000000000 10:1.0000000000
4 1:43987 2:35630.2632853334 3:36590.0056000000 4:31694.8432320000 5:0.0000000000 6:0.0000000000 7:0.0000000000 8:0.0000000000 9:0.0000000000 10:0.0000000000
5 1:57525 2:33066.0054329427 3:33135.0366210937 4:32695.4811360677 5:1294.0000000000 6:2816.0000000000 7:4243.0000000000 8:4843.0000000000 9:4938.0000000000 10:1.0000000000
1 1:165356 2:28447.8295435435 3:28728.4751111111 4:26483.3477657658 5:825.0000000000 6:1958.0000000000 7:3058.0000000000 8:4001.0000000000 9:4805.0000000000 10:1.0000000000
7 1:108734 2:23203.8066693549 3:24277.1764959677 4:22522.5815362903 5:152.0000000000 6:367.0000000000 7:572.0000000000 8:1327.0000000000 9:2658.0000000000 10:1.0000000000
3 1:31705 2:26168.3154720052 3:25915.1823144532 4:17872.0722623698 5:1.0000000000 6:11.0000000000 7:45.0000000000 8:134.0000000000 9:329.0000000000 10:540.0000000000
5 1:57526 2:24011.5662141023 3:40438.6654437030 4:19875.0797322591 5:2.0000000000 6:41.0000000000 7:186.0000000000 8:640.0000000000 9:933.0000000000 10:90.0000000000
10 1:72440 2:29760.7887943097 3:29907.8775419776 4:28965.9997123756 5:0.0000000000 6:835.0000000000 7:5755.0000000000 8:8095.0000000000 9:4968.0000000000 10:1.0000000000
```

Metodologia

Uso de classificação como mecanismo de detecção

- Classificação
 - Agrupar determinados objetos em classes pré-definidas
 - Duas fases: treino e teste
- Classificador: Support Vector Machine (SVM)
 - [LibSVM](#) 2.86 – escolhe melhores parâmetros
 - Kernel RBF (Radial Basis Function): não-linear



Metodologia

Uso de classificação como mecanismo de detecção

- Validação
 - 10-fold cross-validation
 - Quebra a base em 10 conjuntos de tamanho 1/10
 - Usa 9 conjuntos para formar o treino e 1 como teste
 - Repete 10 vezes, cada hora mudando o teste
- Métricas utilizadas
 - Precisão = $tp/(tp+fp)$
 - Revocação = $tp/(tp+fn)$

Metodologia

Uso de classificação como mecanismo de detecção

- Abordagens utilizadas
 - Classificação Multi-classe
 - 10 conjuntos de 44 imagens parecidas
 - 10 classes
 - Classificação Binária
 - Compara vetor de características de pares de imagens
 - Valor absoluto da diferença de cada característica
 - 2 Classes: Parecidas ou Diferentes

Experimentos

Classificação Multi-classe

- 30 execuções
- Métricas:
- Matriz de confusão

Métrica	Valor
Precisão	88,676
Revocação	87,906

	Previsão									
	1	2	3	4	5	6	7	8	9	10
1	112	0	3	2	2	4	3	5	0	1
2	0	121	0	8	0	0	0	2	0	1
3	0	0	119	0	1	0	0	2	10	0
4	5	3	0	115	0	0	1	3	4	1
5	0	0	7	0	114	0	3	5	3	0
6	0	0	1	0	0	120	0	2	3	6
7	3	0	3	0	3	0	112	7	4	0
8	1	0	2	3	0	2	2	116	1	5
9	0	0	11	0	0	3	3	3	112	0
10	0	0	2	0	0	6	0	7	4	113

Experimentos

Classificação Multi-classe – Análise dos Casos de Erro

- Transformações

- Framing
- Original
- Emboss
- Equalize
- Intensity
- Normalize
- Colorize

- Matriz de Confusão

- Características de cores ficaram próximas



Experimentos

Classificação Binária

- Base reduzida
 - 4 transformações
 - Compara cada par de imagens parecidas e cada imagem da base com 4 novas imagens diferentes
- 30 execuções
- Métricas
- Matriz de confusão



Métrica	Binária	Multi-classe Reduzida
Precisão	99,234	75,571
Revocação	99,523	77,627

	Previsão	
	Parecidas	Diferentes
Parecidas	299	1
Diferentes	4	596

Conclusões

- Problema amplamente estudado, difícil de resolver e importante para várias áreas
- Mecanismo proposto é viável, pois apresentou bons resultados com as 2 abordagens de classificação
- Porém, o mecanismo proposto não escala e testamos numa base de apenas 1 categoria de imagens parecidas (transformações nas imagens) e com poucas instâncias

Trabalhos Futuros

- Usar outra base de dados, explorando outras categorias de imagens parecidas
- Usar mais características para representar a imagem, além de avaliar quais são mais relevantes para o mecanismo de detecção
- Desenvolver uma abordagem mais escalável
- Avaliar desempenho e memória

Perguntas

