Identificação de modelos de câmera usando Local Binary Patterns e Sensor Pattern Noise

Mateus Coelho *

Abstract

Com a popularização de câmeras digitais quase sempre incluídas em smartphones, há uma preocupação crescente com relação à segurança e à origem de imagens criadas com esses aparelhos. De fato, ao longo da última década a comunidade forense tem desenvolvido vários métodos para identificar modelos de câmeras fotográficas, dos quais dois foram analisados neste projeto: local binary patterns (LBP) e sensor pattern noise (SPN). Aquele é um método conhecido e amplamente utilizado para descrever texturas de imagens e este foi mais recentemente desenvolvido e leva em conta o ruído produzido pelo sensor do aparelho fotográfico para criar features a serem utilizadas no classificador. Então, com essas duas técnicas foram extraídas features que aplicadas em uma regressão logística fornecem um modelo de classificação de câmeras. Para treinar esse modelo foi utilizado um data set com 2750 imagens de 10 modelos de câmeras. O melhor resultado obtido foi empregando o LBP com um score na plataforma Kaggle de 0.531666.

1. Introdução

Os avanços nas tecnologias de imagens digitais têm provocado uma revolução no mercado de câmeras fotográficas, já que estas estão a cada dia mais baratas e modernas. Inclusive, a popularização de smartphones contribuiu signifcantemente nessa questão. Cresceu também o número de aplicações que envolvem o uso de imagens digitais, desde uso médico até uso bancário. Porém, com esses avanços tem aumentado também a preocupação com a origem das imagens geradas por esquipamentos digitais, como máquinas fotográficas e scanners. A popularização de softwares de manipulação de imagens é um fator ainda mais crítico nesse aspecto pois permite que pessoas possivelmente desonestas possam utilizá-los para falsificar dados.

Assim, em meio a este cenário, este projeto analisou dois métodos de criar features para alimentar um classificador. O resultado são modelos que tem como entrada uma imagem e como saída o tipo da câmera usada para tirar tal foto.

Esse paper está organizado da seguinte forma: segunda seção mostra soluções já estudadas e presentes na literatura; terceira seção expõe em mais detalhes os métodos utilizados; quarta seção discute os experimentos e resultados obtidos; quinta seção apresenta a conclusão.

2. Atividades

A principal dificuldade em relação ao problema de identificar modelos de câmeras digitais é encontrar features que descrevam com qualidade as imagens. Três métodos de extração de features bastante conhecidos são local binary patterns, wavelet features e sensor pattern recognition.

Xu e Shi [**?**] conseguiu uma acurácia de 98.08% empregando como features a distribuição resultante do uniform local binary patterns para as três bandas de cor. Além disso usou vetores de erro de predição correspondentes e subbandas da parte diagonal de um wavelet de primeiro nível. Já Wang et al. [**?**] optou por extrair features de wavelet de alta ordem e features de coocorrência de coeficientes de wavelets. Seu resultado após aplicar Sequential Forward Feature Selection e multi-class SVM foi acurácia de 98%. Por fim, Filler et al. [**?**] estimou e classificou fingerprints (Sensor photo-response non-uniformity) de várias câmeras e modelos diferentes. Para isso extraiu features a partir dos fingerprints que refletem diferenças, por exemplo, no CFA e no sensor de transferência de sinal. A acurácia obtida foi de 90.8%.

Existem ainda outros jeitos de abordar o problema de identificação de câmeras que se baseiam em etapas específicas do processo de aquisição de uma imagem, mas vários esbarram em problemas como acurácia reduzida para câmeras da mesma marca [**?**], transformações geométricas e ataques de ruído.

3. Soluções Propostas

Este paper visa analisar duas maneiras de extrair features de imagens que são local binary patterns e sensor pattern noise. Mais informações sobre eles serão mostrados nas seções abaixo.

*__Contact__: mateus_coelho@live.com

3.1. Local Binary Patterns

Local binary patterns é tipo de descritor visual amplamente utilizado na área de visão computacional para classificar texturas de imagens digitais. Este método pode ser descrito pelas equações 1 e 2, onde R é o raio de um círculo que representa a vizinhança e P é o número de amostras em volta do círculo. g_p e g_c são os valores das intensidades no pixels no centro do círculo e em um ponto da amostra do círculo. A diferença desses valores é aplicada na função s e o resultado é computado criando um número binário que representa a comparação do pixel central com seus vizinhos. Para parâmetros $P = 8$ e $R = 1$, por exemplo, temos a comparação do pixel central com seus 8 vizinhos mais próximos.

$$\text{LBP}_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \tag{1}$$

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0. \end{cases} \tag{2}$$

Computando esse valor para cada pixel de uma imagem tem-se uma distribuição de 2^P possíveis valores. O histograma dessa distribuição é, então, o vetor de features que será empregado no classificador. Se a imagem tem três bandas de cor, esse algoritmo deve ser aplicada para a matriz de cada banda.

Uma imagem quando é adquirida pela câmera fotográfica passa por um pipeline de algoritmos, como filtragem e compressão JPEG, que alteram a imagem original com característica localizadas. Segundo Xu e Shi [**?**] este método pode capturar efetivamente essas características geradas pelo pipeline e até certo ponto suprimir a influência do conteúdo da imagem propriamente.

3.2. Sensor Pattern Noise

Segundo Lukáš et al. [**?**], na captura de uma imagem por uma câmera fotorgráfica, ela passa por um pipeline de etapas até que o resultado final seja obtido. Algumas etapas são conversão de photons para voltagens, passagem pelas lentes, aplicação de um filtro antialiasing, vetor de filtro de cor etc. Ao longo desse processo ruído introduzido. Este pode ser no final das contas dividido em duas parcelas: FPN e PRNU. O PRNU (photo-response nonuniformity noise) é parte dominante e quando obtido pode ser utilizado para especificar uma câmera específica. Esse ruído é considerado como a principal parte do fingerprint (”digital”) de uma câmera.

Uma aproximação do PRNU pode ser obtido tirando a média do ruído de várias imagens. O ruído de uma imagem é calculado simplesmente subtraindo a imagem original do resultado de um filtro de redução de ruído aplicado na imagem original. Para que a aproximação seja boa, devem ser utilizadas no mínimo 50 ruídos de imagens. No caso desse trabalho, foram utilizadas 50 imagens com um filtro de redução de ruído a base de wavelets com as características presentes na tabela 1. As imagens 1 e 2 mostram, respectivamente o ruído de uma imagem e o ruído médio de 50 imagens retiradas do data set do trabalho.

Com os fingerprints em mãos, existem várias maneiras de retirar features para relacionar o ruído médio com o ruído de cada imagem.

Table 1.

Filtro de redução de ruído wavelet	
Sigma	0.019
Wavelet	8-tap Daubechies QMF
Níveis de decomposição	4

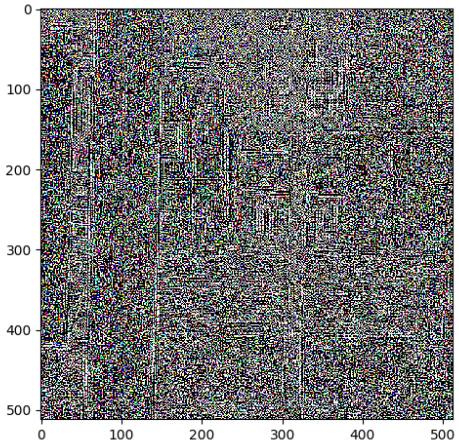

Figure 1. Ruído de uma imagem, obtida pela subtração da original pela original filtrada

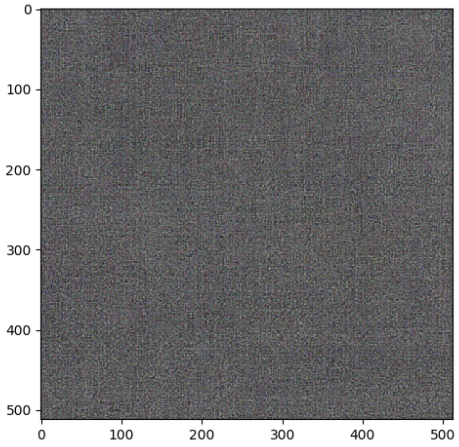

Figure 2. Ruído médio de 50 imagens.

4. Experimentos e Discussão

5. Conclusões e Trabalho Futuro

References

[1] Kai Ni, Anitha Kannan, Antonio Criminisi, and John Winn. Epitomic location recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, AK, USA, 2008. 3

[2] Kai Ni, Anitha Kannan, Antonio Criminisi, and John Winn. Epitomic location recognition. *Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 31(12):2158–2167, 2009. 3

[3] Fulano Silva and Beltrano Souza. Hey! this is my paper. In *European Conference on Nothing (ECN)*, pages 000–007, Graz, Austria, 2010. 3

[4] Fulano Silva. A paper on everything useless. In *European Conference on Nothing (ECN)*, pages 008–014, Graz, Austria, 2010. 3

[5] Fulano Silva, Beltrano Souza, and Sicrano Rocha. Revisiting the classical publishing problem. In *European Conference on Nothing (ECN)*, pages 015–021, Graz, Austria, 2010. 3