IMPLEMENTAÇÃO DO ALGORITMO DE DETECÇÃO DE CÓPIA/CLONAGEM NA ANÁLISE FORENSE DE IMAGENS DIGITAIS

Gustavo Aranha Araújo Costa dos Reis

Bacharel em Engenharia Elétrica pela Universidade de Brasília (UnB).
Especialista em Ciências Forenses IFAR/LS
E-mail: gustavoaranhareis@gmail.com

Palavras-chave: Análise forense de documentos digitais, Algoritmo de detecção de Cópia/colagem.

INTRODUÇÃO

O desenvolvimento tecnológico e seu uso no mundo contemporâneo crescem de forma bastante acelerada. Dentre os diversos produtos nascidos a partir do uso da tecnologia, pode-se afirmar que as imagens digitais compõem grande parcela. De acordo com Meeker (2013) estima-se que diariamente 500 milhões de imagens são compartilhadas nos mais diferentes sites e redes sociais por meio dos inúmeros canais de comunicação. Câmeras digitais e *softwares* de manipulação de imagens e vídeos são, na época atual, itens de baixo custo, simples uso e amplamente difundidos. Para Rocha e Goldenstein (2010), as mídias digitais podem ser manipuladas com dois fins básicos: o primeiro é a modificação inocente com intuito de melhorar a qualidade visual. O segundo tipo de adulteração visa iludir os observadores do documento digital a respeito do seu verdadeiro conteúdo. Os problemas surgem a partir do momento em que essas imagens e vídeos adulterados passam a implicar em questões legais e danosas, tais como: manipulação de imagens com viés político; falsificação de imagens nos meios de comunicação; apresentação de resultados forjados em trabalhos científicos como citado em Farid (2006); manipulação de imagens relacionadas à pornografia infantil; síntese ou adulteração de imagens e vídeos para incriminar ou inocentar sujeitos. Há ainda a hipótese de que a adulteração de imagens de eventos históricos afeta a memória das pessoas em relação a tais eventos (SACCHI; AGNOLI; LOFTUS,

2007). Para que a confiança na veracidade da produção e processamento dos conteúdos digitais possa ser reestabelecida, torna-se indispensável a utilização de técnicas e análises científicas que visam dirimir os conflitos existentes. Assim, surge, dentro da área da Computação Forense, a disciplina Análise Forense de Documentos Digitais – que abrange técnicas computacionais, desenvolvidas de acordo com cada problema específico – como uma alternativa científica ao restabelecimento da credibilidade das mídias digitais.

OBJETIVO

Tem-se como meta principal deste trabalho implementar um método válido e eficaz de análise forense de verificação de adulterações em imagens digitais denominada Cópia e colagem (*cloning*). O método de análise forense de imagens digitais proposto nesse trabalho baseia-se na técnica de detecção de cópia e colagem proposto por Popescu e Farid (2004). Esta técnica, por sua vez, pode ser materializada por meio da aplicação do algoritmo proposto pelos referidos autores.

METODOLOGIA

A implementação do algoritmo foi realizada em meio computacional com a utilização da ferramenta *MATLAB®* (*Mathworks*) versão R2016b, de teste. O *software* produzido pelos autores solicita ao usuário indicar qual imagem será analisada e, após o processamento, imprime ao mesmo uma imagem em preto e branco com as possíveis regiões de cópia e colagem. As imagens processadas pela implementação foram capturadas por meio de câmeras fotográficas de smartphones convencionais e a edição dessas imagens foi realizada – em formato *Portable Network Graphics* (PNG) – por meio do programa de manipulação de mídias *GIMP*, de uso livre e gratuito (*https://www.gimp.org*).

RESULTADOS E DISCUSSÃO

Dentre as diversas técnicas de adulteração de imagens e vídeos com viés fraudulento, o método de Cópia/Colagem, também denominado pela literatura como Clonagem (*cloning*) apresenta-se como uma das técnicas mais empregadas devido à sua simplicidade (ROCHA, GOLDENSTEIN; 2010). Essa manipulação consiste na cópia de determinadas porções de uma

imagem para sobrescrever outras regiões do mesmo arquivo. Um dos principais objetivos desta técnica de adulteração é esconder objetos ou pessoas em uma imagem. Em outras palavras, de forma geral, visa-se omitir informações relevantes que estão sendo representadas por determinados elementos do documento digital. Nos casos em que a manipulação de cópia e colagem seja bem realizada (emenda imperceptível), a simples análise realizada pelo olho humano – mesmo com ajuda de ferramentas básicas, como o zoom – pode não ser suficiente para a detecção da fraude. Torna-se, assim, necessária a utilização de técnicas de análise computacional eficazes para a constatação da fraude. A escolha da técnica de detecção de Cópia/Colagem justifica-se por esta utilizar métodos estatísticos que reduzem a dimensionalidade da análise de dados (VASCONCELOS, SIMONE; 2011). O algoritmo, em resumo, propõe a análise de várias sub-regiões da imagem sob investigação. Cada sub-região é escrita a partir de sua composição de cores e tamanho em um grande número (índice). Ao final do processamento de cada sub-região, os números são armazenados em uma matriz que será ordenada coerentemente para posterior análise. Nesta análise, serão comparados os diversos índices entre si a fim de determinar as regiões com alto grau de similaridade: prováveis duplicações. As ferramentas empregadas juntamente com os procedimentos elaborados produziram resultados satisfatórios e coerentes com o que se esperava a partir da técnica proposta em Popescu e Farid (2004). Foram criados dois cenários de teste (ver apêndice A) e os resultados gerados a partir da análise do método de detecção de clonagem mostraram de forma precisa as regiões duplicadas em ambos os casos. A fim de complementar a discussão, é fundamental observar nos resultados uma outra consequência decorrente da análise computacional: a ocorrência de falsos positivos. Ou seja, em qualquer imagem ou fotografia podem existir regiões de cores, dimensões e formatos similares, mas que não são objetos de cópia/colagem. Posto isto, é possível concluir e observar que o algoritmo também imprime nas imagens dos resultados apresentados outras regiões de similaridades de conteúdo. Entretanto, o fato de existirem falsos positivos não invalida ou compromete a análise principal. Demonstra apenas que a análise computacional – que se mostrou válida e eficaz com taxas teóricas de acerto variando de 97,3 a 99,8% para os casos de estudo – deve ser complementada com uma posterior análise visual por parte daquele que investiga as imagens. No caso prático, as taxas

de acerto deverão ser calculadas empiricamente, caso a caso, pelo fato de não se saber de antemão qual foi a manipulação realizada. Uma outra observação importante a ser feita a respeito do método implementado diz respeito à suas limitações. A implementação não é robusta para clonagens em que a área copiada é rotacionada ou escalonada antes de ser colada. Essa limitação, apesar de ser passível de análise por outras técnicas (WANG; 2009), é previsível desde a fundamentação teórica da técnica em si. De forma objetiva, essa manipulação torna o fragmento rotacionado ou escalonado diferente do fragmento original ao ponto de ser interpretado pelo algoritmo como sendo uma região estranha acrescentada à imagem final, que passa a ter propriedades recém-adquiridas. Por outro lado, qualquer tipo de imagem, independente do formato digital ou da fonte a qual foi gerada, é perfeitamente apta a ser analisada pela implementação proposta neste trabalho.Um vídeo disponibilizado na *internet* pelo autor pode ser acessado pelo endereço: https://youtu.be/YXdR-TEMP94.

CONCLUSÃO

Pode-se concluir, por meio dos resultados observados, que implementação proposta gerou resultados coerentes e válidos para a detecção de fraudes em imagens digitais baseadas na técnica de Clonagem. Também foi possível constatar que apesar de eficiente a técnica deve ser complementada por uma análise visual por parte do investigador. Tendo-se que a soma das análises objetivas (algoritmo) e subjetivas (visual) aliada à aplicação do método dentro das condições corretas contribui para a determinação da veracidade ou falsidade de conteúdos digitais, além de fornecer uma técnica confiável e científica, passível de ser implementada e que pode ser de grande utilidade na esfera forense

REFERÊNCIAS

BAYRAM, Sevinc; SENCAR, Husrev Taha; MEMON, Nasir. A survey of copy-move forgery detection techniques. In: **IEEE Western New York Image Processing Workshop.** IEEE, 2008. p. 538-542.

COSTA, Marcelo Antonio Sampaio Lemos. **Computação forense.** 2011.

FARID, Hany. Exposing digital forgeries in scientific images. In: **Proceedings of the 8th workshop on Multimedia and security.** ACM, 2006. p. 29-36.

JOLLIFFE, Ian. **Principal component analysis.** John Wiley & Sons, Ltd, 2002.

MEEKER, Mary; WU, Liang. KPCB Internet trends 2013. In: **Internet Trends D11 Conference**. 2013.

NETO, JM Moita; MOITA, Graziella Ciaramella. Uma introdução à análise exploratória de dados multivariados. **Química Nova**, v. 21, n. 4, p. 467-469, 1998.

PEREIRA, Evandro et al. Forense Computacional: fundamentos, tecnologias e desafios atuais. **VII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais**, p. 3-53, 2007.

POPESCU, A. C.; FARID, H. Exposing digital forgeries by detecting duplicated image region [Technical Report]. 2004-515. **Hanover, Department of Computer Science, Dartmouth** College**. USA**, p. 32, 2004.

ROCHA, Anderson; GOLDENSTEIN, S. CSI: análise forense de 16. documentos digitais. **Belo Horizonte: Sociedade Brasileira de Computação (SBC),** p. 263-317, 2010.

ROCHA, Anderson et al. Vision of the unseen: Current trends and challenges in digital image and video forensics. **ACM Computing Surveys (CSUR),** v. 43, n. 4, p. 26, 2011.

SACCHI, Dario LM; AGNOLI, Franca; LOFTUS, Elizabeth F. Changing history: Doctored photographs affect memory for past public events. **Applied Cognitive Psychology,** v. 21, n. 8, p. 1005-1022, 2007.

SILVA, Ewerton Almeida; ROCHA, Anderson. Análise forense de documentos digitais: além da visão humana. **Saúde, Ética & Justiça**, v. 16, n. 1, 2011.

VASCONCELOS, SIMONE. Análise de componentes principais. **Availabe at: http://www. ic. uff. br/aconci/PCA-ACP. pdf,** 2011.

WANG, Junwen et al. Detection of image region duplication forgery using model with circle block. In: **2009 International Conference on Multimedia Information Networking and Security**. IEEE, 2009. p. 25-29.

AGRADECIMENTOS

Ao professor Msc. Ian Alves Ulian pelas contribuições ao conteúdo deste trabalho.

Não foram declarados conflitos de interesse associados à publicação deste artigo.

APÊNDICE A

Figura 1 - Imagem original de objeto a ser duplicado.
Fonte: Próprio autor (capturada pelo *smartphone* Moto G - 2ª geração).

Figura 2 - Imagem editada de objeto duplicado.
Fonte: Próprio autor (editada pelo GIMP).

Figura 3 - Resultado da análise de Detecção de Cópia/Colagem da figura 2.
Fonte: Próprio autor (gerada pelo MATLAB®).

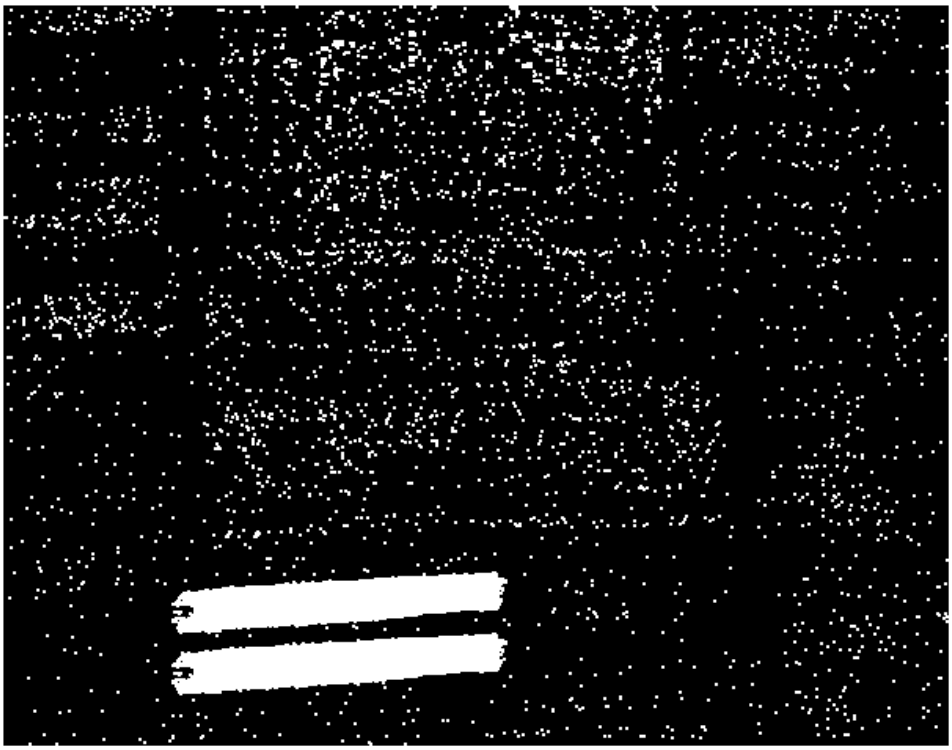

Figura 4 - Imagem original de objeto a ser ocultado
Fonte: Próprio autor (capturada pelo *smartphone* Moto G - 2ª geração).

Figura 5 - Imagem da edição de objeto a ser ocultado
Fonte: Próprio autor (captura de tela).

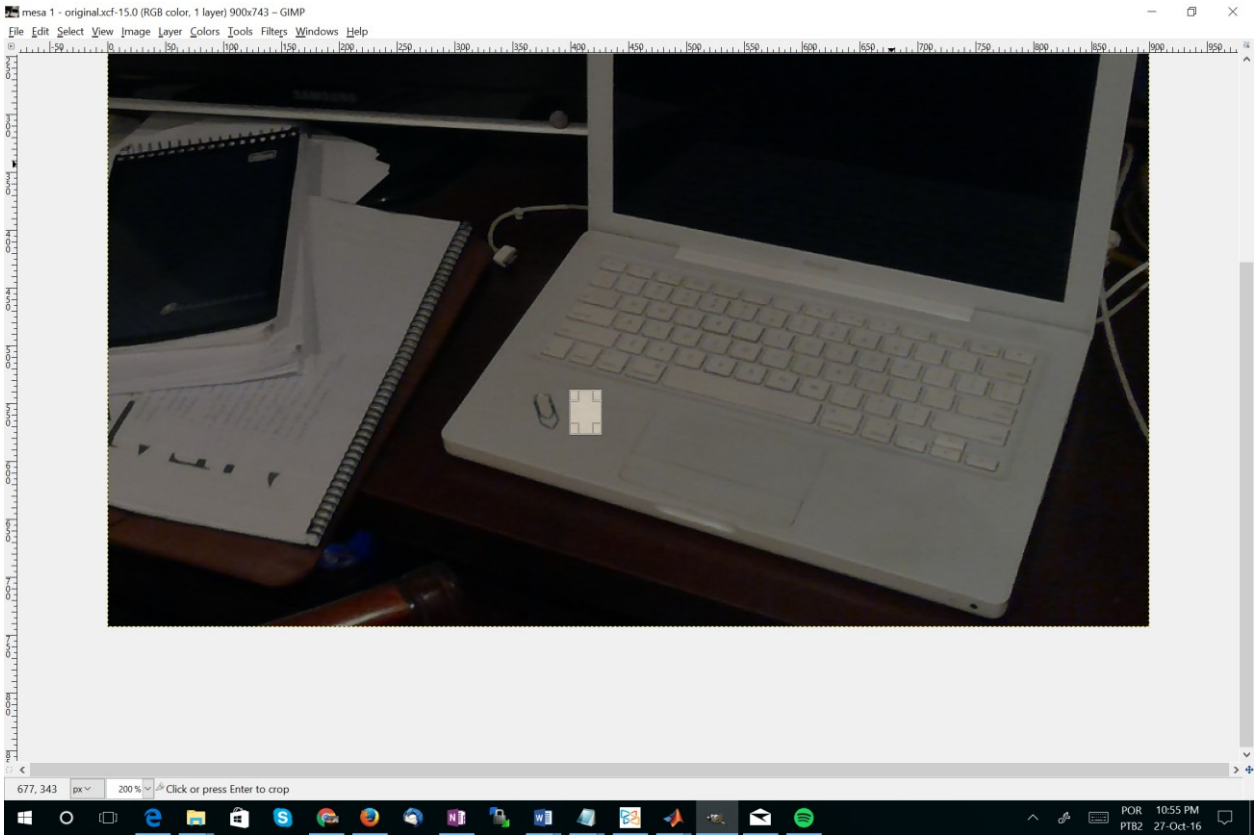

Figura 6 – Imagem editada de objeto ocultado.
Fonte: Próprio autor (editada pelo GIMP).

Figura 7 - Resultado da análise de Detecção de Cópia/Colagem da imagem 6.
Fonte: Próprio autor (gerada pelo MATLAB®).