

# Detecção de Nudez Utilizando Redes Neurais Artificiais

Pedro S. Franco, Daniel A. M. Sandoval, Alexandre Zaghetto, Célia Ghedini Ralha

**Resumo**—O presente artigo propõe um método baseado em redes neurais artificiais para detecção de nudez em imagens digitais a ser incorporado ao MADIK (*Multi-agent Digital Investigation ToolKit*) em um agente de processamento de imagens. O objetivo é utilizá-lo como ferramenta de auxílio a peritos em investigações que envolvem crimes como a pedofilia. O método proposto pode ser dividido em 5 etapas: (a) segmentação por cor de pele; (b) realização de operações morfológicas; (c) rotulação e contagem de elementos conectados; (d) medição da área do maior elemento conectado; (e) classificação por meio de uma rede neural artificial do tipo *feedforward*, treinada a partir do algoritmo Levenberg-Marquardt. Nosso conjunto de testes é formado por 1000 imagens, 500 contendo nudez e 500 não. Em nossos experimentos, foram treinadas 10 vezes cada uma das redes neurais com 1 a 15 neurônios na camada escondida, totalizando 150 redes. São calculadas as taxas de verdadeiros e falsos positivos e negativos e a F-measure é utilizada como métrica para seleção da rede que apresenta o melhor desempenho. A F-measure observada no melhor caso foi de 0.7967, para uma rede com 5 neurônios na camada escondida.

**Index Terms**—Pedofilia, Detecção de nudez, Redes Neurais Artificiais, *Multi-Agent Digital Investigation ToolKit*, Investigação forense.

## I. INTRODUÇÃO

Com o avanço tecnológico e a popularização de máquinas fotográficas digitais e dispositivos de armazenamento, vastas bibliotecas digitais de imagens digitais tornaram-se comuns. No combate a crimes como a pedofilia, a inspeção humana dessas vastas bibliotecas torna-se muito difícil, desgastante e por vezes inviável.

O processo investigativo de sistemas computacionais resulta em uma tarefa complexa, envolvendo a participação e cooperação de diversos especialistas forenses. No entanto, a excelência do processo pode ser limitada devido à falta de recursos, sejam eles técnicos, científicos ou computacionais. Visando uma maior agilidade e eficiência dos processos investigativos foi desenvolvido o MADIK (*Multi-Agent Digital Investigation ToolKit*) [1], [2]. Trata-se de uma ferramenta baseada na abordagem de sistema multiagente distribuído, o qual se revelou apropriado, principalmente devido à capacidade de cooperação de agentes autônomos especializados [3]. No MADIK foram implementados diferentes agentes, cada qual baseado em uma área específica de conhecimento técnico utilizando uma arquitetura composta de quatro camadas, tal como uma metáfora à hierarquia das organizações: estratégica, tática e operacional. Na camada operacional foram definidos cinco agentes especialistas:

*HashSetAgent*, *TimeLineAgent*, *FileSignatureAgent*, *FilePathAgent* e *KeywordAgent*. Os agentes especialistas interagem durante o processo investigatório, correlacionando achados sobre os arquivos em análise através do uso de um *blackboard* colaborativo, o qual permite uma avaliação mais detalhada de cada arquivo sob investigação.

Tendo em vista a importância de vestígios elucidativos em conteúdos digitais de imagens, este trabalho visa a extensão do MADIK com o desenvolvimento do agente especialista em processamento de imagens de alta resolução, denominado *ImageAgent*. De forma geral, o *ImageAgent* deve ser capaz de analisar o conteúdo de arquivos de tipos diversos de imagens, para permitir a identificação, extração e classificação de características de interesse de uma investigação criminal, sendo que o método proposto está mais especificamente ligado à detecção de imagens de nudez, com a finalidade de auxiliar no combate a pedofilia.

## II. MÉTODO PROPOSTO

O método proposto pode ser dividido em cinco etapas: (i) segmentação por cor de pele; (ii) realização de operações morfológicas; (iii) rotulação e contagem de elementos conectados; (iv) medição da área do maior elemento conectado; e (v) classificação por meio de uma rede neural artificial [4] do tipo *feedforward*, treinada a partir do algoritmo Levenberg-Marquardt [5]. Cada uma dessas etapas será detalhada a seguir.

### A. Segmentação por Cor de Pele

Dentre os espaços de cores com os quais se pode operar, o HSV (*Hue, Saturation, Value*) é indicado pela literatura como um dos mais adequados para efeito de segmentação por cor de pele [6], [7], sendo o plano determinado por *H* e *S* utilizado para tal finalidade. Uma vez que a componente *V* não é levada em consideração, a interferência da luminância no processo de identificação das regiões de interesse é minimizada [8].

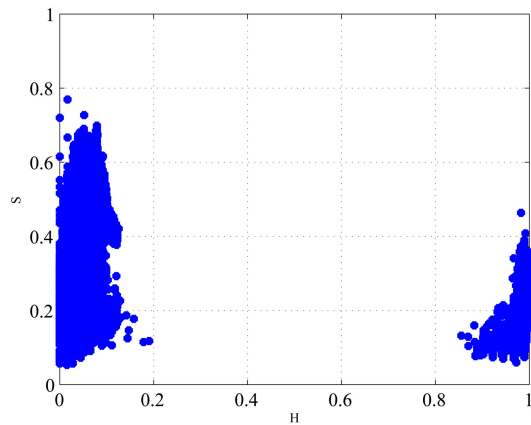
As cores típicas para pele humana foram primeiro estudadas por Felix Ritter Von Luschan [9], que desenvolveu uma escala de classificação denominada Escala de Von Luschan. A Figura 1 mostra um instrumento antropométrico que utiliza a classificação proposta por Von Luschan [10]. Em um cenário mais realista uma imagem composta por segmentos de fotografias digitais de pele humana por ser utilizada como referência na avaliação da escala proposta por Von Luschan. Um exemplo de imagem, bem como o plano  $H \times S$  correspondente, são mostrados na Figura 2. A partir da análise



Fig. 1. Instrumento antropométrico construído a partir da Escala de Von Luschan<sup>1</sup>.



(a)



(b)

Fig. 2. (a) Recortes de fotografias de pele; e (b) localização das cores no plano  $H \times S$ .

de gráficos dessa natureza, pode-se estabelecer os intervalos de interesse  $H_i = [0 \ 0.11]$  e  $S_i = [0.2 \ 0.7]$  para as componentes  $H$  e  $S$ , respectivamente [11]. *Pixels* cujos valores de  $H$  e  $S$  estiverem contidos nesses intervalos são marcados em uma máscara binária com o valor 1. Os demais são marcados com 0. Trata-se de um método de classificação trivial que pode em trabalhos futuros ser substituído por métodos mais robustos.

Por fim, dado que a complexidade computacional é um problema a ser considerado, a máscara binária resultante é subamostrada por um fator de 0.5 antes de ser submetida às etapas seguintes de processamento.

### B. Realização de Operações Morfológicas

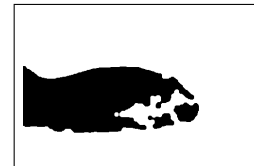
Após a binarização e a subamostragem descritas na etapa anterior, a imagem passa pelos processamentos morfológicos de abertura e fechamento [12]. O objetivo é diminuir a quantidade de regiões que erroneamente foram classificadas como regiões de interesse e conectar outras que por



(a)



(b)



(c)

Fig. 3. Binarização e operações morfológicas: exemplo de imagem original (a), e suas versões binarizadas antes (b) e depois (c) do processamento morfológico sugerido.

ventura tenham sido desconectadas pela binarização. A Figura 3 mostra um exemplo de imagem original (a), e suas versões binarizadas antes (b) e depois (c) do processamento morfológico sugerido.

### C. Rotulação e Contagem de Elementos Conectados

Considerando que as imagens binárias são representadas pelos bits 1 (branco) e 0 (preto), a metodologia adotada para se rotular os elementos de interesse está resumida no algoritmo *ROTULA*, onde  $M$  é a imagem binária e  $R$  é a matriz de rótulos.

**ROTULA( $M$ )**

```

1   $R(i, j) \leftarrow 0$ 
2  for each  $M(i, j) = 0$  do
3      if  $M(i, j - 1) = 1$  and  $M(i - 1, j) = 1$ 
4          then  $Indice = Indice + 1$ ;
5           $R(i, j) \leftarrow Indice$ ;
6      elseif  $M(i, j - 1) = 0$  and  $M(i - 1, j) = 1$ 
7          then  $R(i, j) \leftarrow R(i, j - 1)$ ;
8      else  $R(i, j) \leftarrow R(i - 1, j)$ ;
9  return  $R$ 
```

Em uma segunda passagem, estabelece-se a equivalência entre rótulos, quando regiões vizinhas possuem rótulos diferentes. Após essa etapa, tem-se a informação de quantos elementos conectados há na imagem. A Figura 4 ilustra o resultado do processo de rotulação. Nesse exemplo foram rotuladas 4 regiões: o fundo (preto) mais três regiões de cor de pele (demais tons de cinza).

### D. Medição da Área do Maior Elemento Conectado

Nessa etapa, calcula-se a área em *pixels* para cada elemento conectado identificado na etapa anterior e, em seguida,

<sup>1</sup>Museo di Storia Naturale, Università degli Studi di Firenze - <http://www.msn.unifi.it/CMpro-v-p-69.html>.



Fig. 4. Exemplo de rotulação: foram rotuladas 4 regiões: o fundo mais três regiões de cor de pele.

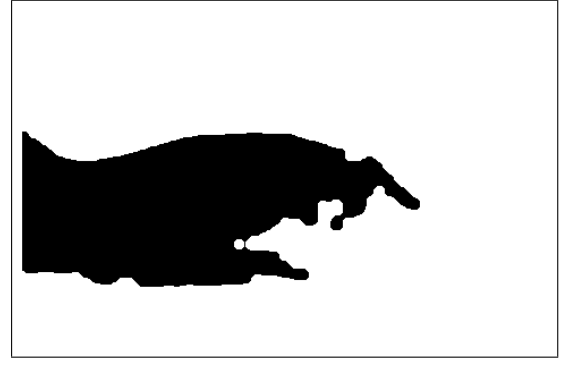


Fig. 5. Destaque do elemento conectado de maior área: aproximadamente 23% da área total da imagem.

procura-se pelo maior valor observado. É importante ressaltar que são considerados como candidatos nessa etapa apenas os elementos conectados que encontram-se em sua maioria dentro da área classificada como cor de pele na etapa de segmentação. Dessa forma, evita-se contar o fundo como um possível elemento conectado. Por fim, calcula-se a razão entre a área do maior elemento conectado e a área total da imagem. A Figura 5 destaca a região de maior área para o exemplo em questão. Esse elemento conectado corresponde a aproximadamente 23% da área total da imagem.

#### E. Classificação Utilizando Redes Neurais Artificiais

O objetivo do classificador proposto é identificar imagens candidatas a conterem nudez. Optou-se por utilizar uma rede neural artificial [4] do tipo *feedforward*, treinada a partir do algoritmo Levenberg-Marquardt [5]. Foi proposta uma arquitetura de 3 camadas: entrada, saída e oculta. Na camada de entrada os neurônios não realizam qualquer computação. Já nas camadas de saída e oculta, a função de ativação tangente hiperbólica é empregada. Um único neurônio compõe a camada de saída, enquanto o número de neurônios na camada oculta variou de 1 a 15 durante o treinamento.

O vetor de entrada  $P$  para o classificador é composto por dois parâmetros: (i) o número de elementos conectados encontrados na imagem; e (ii) a área relativa do maior elementado conectado. Esses dois valores são extraídos nas etapas descritas nas Subseções II-C e II-D, respectivamente.

A rede é treinada para responder com o valor  $T = 1$ , se o vetor de entrada  $P$  corresponde a uma imagem candidata a conter nudez, e com  $T = -1$ , caso contrário. É importante ressaltar que na operação os neurônios de saída não respondem exatamente com 1's ou -1's, mas apenas com valores aproximados. Por isso, é necessário incluir uma etapa de limiarização. Caso a resposta do neurônio de saída seja maior que 0, a imagem é marcada como contendo nudez. Por outro lado, se a resposta for menor ou igual 0, a imagem não é tida como suspeita.

A métrica de desempenho utilizada foi a *F-measure* [13], definida pela Equ. 1,

$$F\text{-measure} = 2 \cdot \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \quad (1)$$

onde *precision* e *recall* são definidas por,

$$\text{precision} = \frac{tp}{(tp + fp)} \quad \text{e} \quad \text{recall} = \frac{tp}{(tp + fn)},$$

e  $tp$ ,  $fp$  e  $fn$  são as taxas de verdadeiros positivos (*true positive*), falsos positivos (*false positive*) e falsos negativos (*false negative*), respectivamente. De uma maneira geral, a medida *precision* de um sistema de reconhecimento de padrões representa a relação entre a quantidade de classificações verdadeiramente positivas ( $tp$ ) retornadas pelo sistema e a quantidade total de classificações retornadas como positivas ( $tp + fp$ ). Em outras palavras, representa a probabilidade de uma classificação retornada pelo sistema como positiva ser uma ocorrência verdadeiramente positiva. Já a medida *recall* representa a relação entre a quantidade de classificações verdadeiramente positivas retornadas pelo sistema ( $tp$ ) e a quantidade de classificações que deveriam ter sido retornadas como positivas ( $tp + fn$ ). Ou seja, representa a probabilidade de uma ocorrência verdadeiramente positiva ser classificada como tal pelo sistema. A *F-measure* é, dessa forma, a média harmônica entre *precision* e *recall*.

### III. RESULTADOS

Nossa base de testes é formada por 1000 imagens baixadas da Internet, 500 contendo nudez e 500 não. A Figura 6 mostra a distribuição dos vetores de entrada  $P$ , definidos a partir dessas imagens. O eixo vertical representa o número de elementos conectados e o eixo horizontal representa a área relativa do maior elemento conectado. Para o treinamento da rede, as duas classes de imagens, nudez e não-nudez, são divididas em três subconjuntos: treinamento, validação e teste, cada qual contendo aproximadamente 1/3 das imagens. O conjunto de treinamento é utilizado para se computar o gradiente e atualizar os pesos e *bias* da rede. O conjunto de validação é utilizado para interromper o processo de aprendizado assim que a rede começa a apresentar super-aprendizado (*overfitting*). O conjunto de teste é utilizando na avaliação do desempenho final da rede.

Durante o treinamento, variou-se a quantidade de neurônios na camada escondida de 1 a 15 e cada rede foi treinada 10 vezes, com seus pesos iniciados aleatoriamente a cada treinamento. O melhor desempenho observado (*F-measure* = 0.7967) foi obtido com 5 neurônios da camada escondida. As

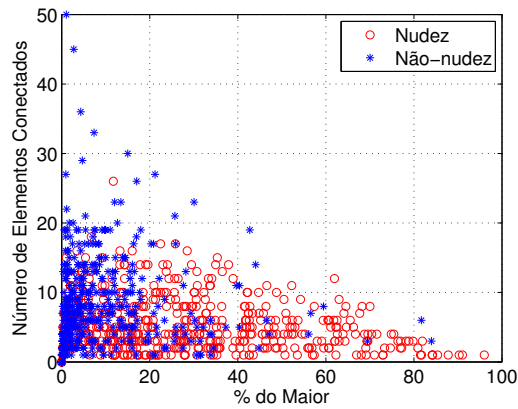
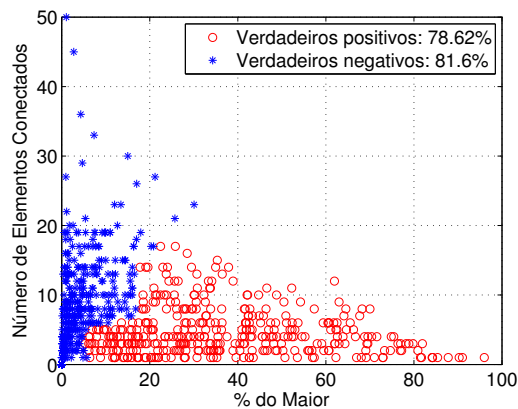
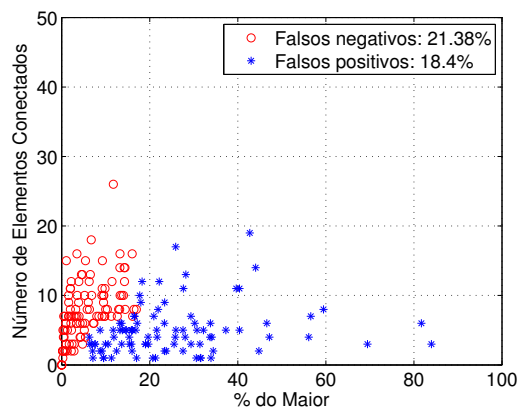


Fig. 6. Distribuição dos vetores de entrada  $P$ , definidos a partir das imagens da base de teste. O eixo vertical representa o número de elementos conectados e o eixo horizontal representa a área relativa do maior elemento conectado.



(a)



(b)

Fig. 7. (a) Verdadeiros positivos e negativos; e (b) Falsos positivos e negativos.  $F\text{-measure} = 0.7967$ .

Figuras 7 (a) e (b) mostram a distribuição de verdadeiros positivos e negativos; e falsos positivos e negativos, respectivamente.

#### IV. CONCLUSÕES

Foi apresentado um método baseado em redes neurais artificiais para detecção de nudez em imagens digitais. Pretende-se, com isso, auxiliar na investigação de crimes de pedofilia. Os resultados mostraram a viabilidade do método proposto, no entanto sugerem a necessidade de se incluir descritores de imagem mais adequados, em substituição ou além daqueles que foram utilizados no presente trabalho, que são: o número de elementos conectados e a área do maior elemento conectado. Em trabalhos futuros pretende-se, ainda, melhorar a segmentação por cor de pele, além de incluir detectores de face, de formas corporais e órgãos genitais. A versão final do método será incorporada à ferramenta de investigação forense MADIK em um agente de processamento de imagens.

#### REFERÊNCIAS

- [1] Bruno W. P. Hoelz, Célia Ghedini Ralha, Rajiv Geeverghese, and Hugo C. Junior, "A cooperative multi-agent approach to computer forensics," in *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Sydney, NSW, Australia, December 9-12, 2008*, 2008, pp. 477–483, IEEE Computer Society.
- [2] Bruno W. P. Hoelz, Célia Ghedini Ralha, Rajiv Geeverghese, and Hugo C. Junior, "Madik: A collaborative multi-agent toolkit to computer forensics," in *On the Move to Meaningful Internet Systems: OTM 2008 Workshops, OTM Confederated International Workshops and Posters, ADI, AWeSoMe, COMBEK, EI2N, IWSSA, MONET, OnToContent + QSI, ORM, PerSys, RDDS, SEMELS, and SWWS 2008, Monterrey, Mexico, November 9-14, 2008. Proceedings*, 2008, vol. 5333 of *Lecture Notes in Computer Science*, pp. 20–21, Springer-Verlag.
- [3] Bruno W. P. Hoelz, Célia Ghedini Ralha, and Rajiv Geeverghese, "Artificial intelligence applied to computer forensics," in *Proceedings of the 2009 ACM Symposium on Applied Computing (SAC), Honolulu, Hawaii, USA, March 9-12, 2009*, 2009, pp. 883–888, ACM.
- [4] Simon Haykin, *Neural Networks: A Comprehensive Foundation (2nd Edition)*, Prentice Hall, 2 edition, July 1998.
- [5] M. T. Hagan and M. Menhaj, "Training feedforward networks with the marquardt algorithm," *IEEE Transactions on Neural Networks*, vol. 5, pp. 989–993, 1994.
- [6] Benjamin D. Zarit, Boaz J. Super, and Francis K. H. Quek, "Comparison of five color models in skin pixel classification," in *In ICCV 99 Int. Workshop on*, 1999, pp. 58–63.
- [7] Juwei Lu, Qian Gu, K. N. Plataniotis, and Jie Wang, "A comparative study of skin-color models," in *Proceedings of the Second international conference on Image Analysis and Recognition*, 2005, ICIAR'05, pp. 729–736.
- [8] Vladimir Vezhnevets, Vassili Sazonov, and Alla Andreeva, "A survey on pixel-based skin color detection techniques," in *IN PROC. GRAPHICON-2003*, 2003, pp. 85–92.
- [9] F. von Luschan, *Völker, Rassen, Sprachen : Anthropologische Betrachtungen*, Deutsche Buchgemeinschaft, Berlin, 1927.
- [10] Università degli Studi di Firenze, "Strumenti antropometrici," 2011.
- [11] Anjali A. Shejul and Umesh L. Kulkarni, "A Secure Skin Tone based Steganography Using Wavelet Transform," *International Journal of Computer Theory and Engineering*, vol. 3, no. 1, pp. 16–22, February 2011.
- [12] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, EUA, 2nd edition, 2002.
- [13] C. J. van Rijsbergen, *Information Retrieval*, Butterworth-Heinemann, 2 edition, Mar. 1979.