

# Identification of High-Resolution Images of Child and Adolescent Pornography at Crime Scenes

Pedro Monteiro da Silva Eleuterio<sup>(1)</sup> and Mateus de Castro Polastro<sup>(2)</sup>

(1) *Brazilian Federal Police. Email: pedro.pmse@dpf.gov.br*

(2) *Brazilian Federal Police. Email: polastro.mcp@dpf.gov.br*

**Abstract** - In several countries, the possession of files containing child and adolescent pornography is a crime. Law-enforcement officers must be able to detect this type of content at crime scenes. The NuDetective Forensic Tool was developed to assist forensic examiners in the timely identification of such files at crime scenes. NuDetective automatically detects nudity in images and employs other techniques to reduce the number of files potentially containing child and adolescent pornography that must be analyzed. High-resolution digital photographs can be produced with currently available digital cameras and camcorders; this technological capability may result in undesirable delays in file processing. This paper presents an approach of image resizing to optimize the runtime of the automatic nudity detection provided by the NuDetective Forensic Tool. The results of several experiments performed to evaluate the efficiency of this approach showed a runtime reduction of about 90%, with minimal changes in the accuracy of nudity detection.

**Keywords** - child pornography, nudity detection, image processing, computer forensics

## 1. Introduction

The production and propagation of illegal digital multimedia content, especially images and videos depicting child and adolescent pornography, have become increasingly common in crimes against children. The rapid evolution of computers and the popularization of the Internet have allowed more frequent production and subsequent exchange of illegal files among

individuals dispersed throughout the world, primarily through the World Wide Web (WWW), peer-to-peer networks [1], e-mail, and instant-messaging services.

These facts have encouraged the development of recent Brazilian legislation. Previously, Article 241 of the *Child and Adolescent Law* criminalized only the disclosure of pornographic material involving children or adolescents [2, 3]. *Law 11,829*, published on November 25<sup>th</sup>,

2008, criminalized the possession of such files in Article 241-B [4], which states:

*“Individuals may not acquire, possess or store, by any means, photo, video or other record that contains the scene of explicit sex or pornography involving children or adolescents”* [4].

Consequently, investigators have received more frequent requests for the forensic analysis of computer storage media, such as hard disk drives (HDDs), pen drives, optical media, and memory cards, to verify the presence and distribution of child pornography. In the authors' experience, the greatest difficulty encountered in such analysis is the identification of files containing child and adolescent pornography among the potentially hundreds of thousands of files stored on the devices. A recent study [5] showed that an analysis of more than 300,000 image files and 1,100 video files identified 148 files depicting the sexual abuse of a teenaged girl. This exemplifies the difficulty of finding such files, especially given the typically visual analytical methods applied to images and videos.

Amendments to Brazilian Law have expanded the need to detect files containing child and adolescent pornography to include not only seized devices, but also those present at crime scenes. To avoid unnecessary seizure of digital devices for subsequent laboratory analysis (often days or weeks later), forensic examiners must be able to identify these files in a timely manner so that investigators have the opportunity to arrest the suspect immediately. The NuDetective Forensic Tool [6] was developed to help forensic examiners conduct such rapid analysis at crime scenes. The Tool performs pixel and shape analysis to automatically detect files depicting human nudity (image analysis) and searches for file names that contain expressions typically used for child and adolescent pornography (file-name analysis).

However, the evolution of digital multimedia devices, such as cameras and camcorders, has allowed the production of high-resolution

multimedia content. Digital cameras currently marketed in Brazil have an average resolution of 1 megapixels (MP), and many exceed 20 MP. The increased computational effort required for the automated manipulation of such files may delay forensic examiners' analysis at the crime scene. The primary objective of the present study was the optimization of automatic nudity detection using the NuDetective Forensic Tool through the reduction of runtime required to process high-resolution images.

The comparison of file hash values is another widely used technique in forensic computing to identify illegal files. Hash values are calculated using a one-way mathematical function [7] and serve as a unique identifier of each file. In hash analysis, the hash values of all suspicious files are calculated and compared with a list of known hash values for previously identified files. Commercial tools, such the *Access Data Forensic Toolkit* [8] and *Guidance Encase* [9], have been used for this purpose, but their application at crime scenes has some limitations. Although hash analysis is an efficient method for the identification of known files, the calculation of hash values for thousands of files stored on an HDD, for example, may be a lengthy operation that cannot be performed at the crime scene. Moreover, the success of this method depends on the completeness of the list of hash values for known files. New and unknown files produced by pedophiles, for example, may not be successfully identified by hash analysis. However, depending on operational viability and the time required, hash analysis may be used in conjunction with the image and file-name analysis of the NuDetective Tool to improve the reliability and effectiveness of results. The present study thus sought to include hash analysis in the forensic analysis of high-resolution images potentially depicting child and adolescent pornography.

This paper is organized as follows: Section 2 presents the state of the art of pixel analysis and human nudity detection algorithms. A brief description of the basic features and nudity detection algorithms of the NuDetective Forensic

Tool is presented in Section 3. Section 4 presents the proposed approach to resize high-resolution digital images before submitting them to the automatic nudity detection provided by the NuDetective Forensic Tool, including several experiments that evaluated the efficiency of this approach. Section 5 briefly describes version 2.0 of the NuDetective Forensic Tool, which includes new features, such as hash analysis. Sections 6 and 7 present the conclusions of this study and discuss directions for future research, respectively.

## 2. Background

The detection of human nudity in digital images is based primarily on the identification of pixels that contain skin color and the analysis of relationships among identified skin regions. The identification of skin color requires the selection of an appropriate color space and the definition of relevant color parameters. The identification of nudity in an image relies primarily on comparison with human forms and analysis of the relationships among identified skin regions.

### 2.1. Color Space

A digital image can be described as a point matrix, wherein each point (pixel) represents a color. Like any mathematical representation of a physical phenomenon, colors may be represented in several ways, each associated with advantages and disadvantages [10]. The choice of color space thus depends on the application type.

The red-green-blue (RGB) color space was originally developed for use with cathode ray tubes (CRTs) and remains the most common method of storing digital images [11]. RGB defines color representation by combining the three primary colors: red, green, and blue [12]. The scale of each color component ranges from 0 to 255. Brightness and color are coupled in the RGB color space, making it unsuitable for color segmentation in images with unknown light conditions [13]. The normalization of

RGB, shown in equation 1, reduces the effect of luminance on color representation [12].

$$\mathbf{r} = \frac{R}{R+G+B}, \mathbf{g} = \frac{G}{R+G+B}, \mathbf{b} = \frac{B}{R+G+B} \quad (1)$$

Color spaces based on hue (H) and saturation (S) have been proposed to allow the intuitive description of colors [12]. Hue defines the dominant color and saturation defines the color based on brightness. Others color spaces like YCbCr (luminance, blue chrominance, and red chrominance), called orthogonal, seek to reduce redundancy in RGB color channels and to represent colors by using components as independently as possible [11].

### 2.2. Skin-Color Classification

The analysis of digital images to identify skin regions seeks to determine whether each pixel represents skin color. The classification of pixels representing skin color is difficult because several factors affect this color, including ambient light; the use of different digital cameras, which can produce different pixel colors of the same person under the same lighting conditions; and natural variation in human skin color [14]. Additionally, similarly colored objects, such as wood, leather, hair, and sand, may confound the classification of skin color in relation to the image background [15].

Human skin color, which depends mainly on hemoglobin and melanin concentrations and lighting conditions [16], is not located randomly within a color space. Skin colors are grouped within small areas of color space, which vary depending on the color space used [15].

Many studies have sought to develop methods for the classification of skin-color pixels using existing color spaces. One proposed method [17] for the detection of skin color in RGB color space uses simple rules to rapidly construct skin-color pixel classifiers [12]. This method uses a basic arithmetical formula (relationship 2) to describe the relationship between the three components of RGB color:

$$\mathbf{R} > 95 \text{ and } \mathbf{G} > 40 \text{ and } \mathbf{B} > 20 \text{ and } \max\{\mathbf{R}, \mathbf{G}, \mathbf{B}\} - \min\{\mathbf{R}, \mathbf{G}, \mathbf{B}\} > 15 \text{ and } |\mathbf{R} - \mathbf{G}| > 15 \text{ and } \mathbf{R} > \mathbf{G} \text{ and } \mathbf{R} > \mathbf{B}. \quad (2)$$

Other color spaces have also been used to define skin-color pixel classifiers. Areas of color space belonging to the skin-color set have been defined using the normalized RGB and HSV color spaces [18], as seen in relationships 3<sup>a</sup> and 4:

$$0.36 \leq r \leq 0.465, \quad 0.28 \leq g \leq 0.363, \quad (3)$$

$$0 \leq H \leq 50, \quad 0.20 \leq S \leq 0.68, \quad 0.35 \leq V \leq 1.0. \quad (4)$$

The YCbCr color space has been used to segment regions of skin color using chrominance information, resulting in high rates of detection in images with good lighting conditions [19].

Probabilistic methods have also been used to estimate the distribution of skin color without the use of a color space, resulting in the production of skin probability maps (SPMs). These methods have been classified as parametric or nonparametric skin distribution modeling. Although non-parametric modeling SPMs can accurately detect skin color independently of the color space used, they require a large amount of storage space to be fitted and must use representative images to train the model [12]. Parametric modeling requires less storage space and has achieved high representation using the smallest number of training images [11]. However, in some methods, such as Gaussian mixture models, performance depends strongly on the shape of the skin-color distribution and training and classification can be extremely slow [12].

### 2.3. Nudity Detection

The detection of nudity in images begins with the identification of skin-color pixels. Once these points have been identified, it is necessary to determine whether the regions bounded by skin color correspond to shapes of the human body.

Previously employed methods have used image features to classify the presence of nudity or compared skin-color regions in segmented images with a database of pre-classified images. One study [20] has suggested several feature extraction techniques for nudity detection, like the comparison of the largest identified skin region with the image size, count the number of identified skin regions or the percentage of skin color in an image. Another study [21] used an algorithm to extract image features suggestive of nudity. This algorithm is based on the multiple and closely grouped skin tones and regions that are typical of nude images. The percentage of skin-color pixels among all image pixels is first calculated; through experimentation, the authors found that images with >15% skin-color pixels should be further evaluated. The continuity of the three major identified skin-color regions is then compared with various empirically obtained thresholds.

Other methods have used image databases to train the classifier. One study [22] trained a support vector machine (SVM) to detect images depicting human nudity. The authors first verified the correlation between regions containing skin color and those that did not, and then used this correlation to train the SVM. The classifier then used the information gained in training to determine whether a new image contained nudity. Another method [23] was similarly based on the comparison of suspicious images with a set of features obtained from the analysis of a training database.

## 3. The NuDetective Forensic Tool

This section reviews the main features of the NuDetective Forensic Tool [6], including pixel analysis and human nudity detection algorithms. The automatic detection algorithms of NuDetective were based on two main

<sup>a</sup> Relationship 3 does not present a value for b because the sum of the components r, g, and b is always equal to 1 in normalized RGB [12].

requirements: high processing speed and reduced number of false-negative results. The color space and skin-color identification parameters used in the tool met these essential requirements, as described below.

### **3.1. Color Space and Skin Color**

NuDetective uses the RGB color space to represent digital color, due to the ease and speed of color processing. Although the use of this color space does not allow the separation of luminance and color components, such separation does not necessarily lead to better results in skin-color classification [12]. Moreover, previous research [14] has found that the ability to detect skin color is independent of the color space used.

NuDetective uses the calculation described in relationship 2 (Section 2.2) to relate the three components of RGB color space [17]. The classification of each pixel as skin color or not applies only to color images, which does not represent a significant disadvantage because the vast majority of digital images depicting human nudity, including child and adolescent pornography, are colored.

### **3.2. Nudity Detection Algorithm**

The nudity detection algorithm used in the NuDetective Forensic Tool has produced excellent results [21], as described in Section 3.4. The computational rapidity of this algorithm met the proposed requirement for the NuDetective tool. Further, this method does not require the use of a database to classify images, instead using the percentage of skin color and computational geometry to identify nudity in digital images [21]. The choice of this algorithm allowed NuDetective to run without a large computing infrastructure, which is desirable at crime scenes.

### **3.3. Basic Features**

Developed using the *Java Standard Edition* (JSE) platform [24], the NuDetective Forensic

Tool automatically detects human nudity in images using the algorithms described above. Many other options can be configured by the forensic examiner to assist in the detection of child pornographic files.

The forensic examiner first chooses locals to search, such as folders and/or disk partitions, and configures search options. These options include the image types to be searched, the search method (by file extensions or signatures), filtering by image size, and the analysis of file names. NuDetective also allows the analysis of compressed files in ZIP format. NuDetective then scans all files of selected locals to search for images and submits identified images to image analysis with the aim of detecting nudity. The tool can compare the names of all files or only image files with a list of predefined keywords<sup>b</sup> commonly used to share child pornographic files on the Internet.

NuDetective uses Java threads, which allows the parallel execution of some tasks. One thread is responsible for searching the selected locals, a second conducts image analysis to detect nudity, and a third is responsible for displaying the results in real time using a graphical user interface (GUI). The forensic examiner can thus begin to analyze the results before the tool has completed the search of all files in the selected locals.

### **3.4. Previous Results**

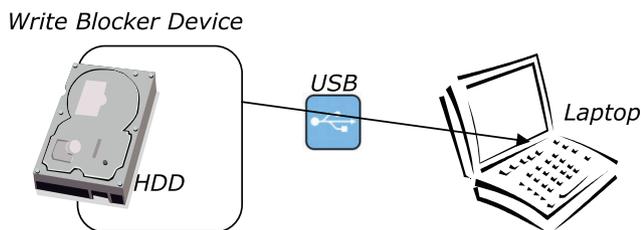
As previously described [6], two experiments were conducted to measure and evaluate the nudity detection rate of the NuDetective Forensic Tool. In the first experiment, a controlled database was created with several images obtained from *imageafter.com* [25] and with nude images obtained from randomly accessed Internet sites, including those depicting Asian, Black, and Caucasian individuals. The second experiment

---

<sup>b</sup> The authors defined the list of keywords according to their professional experience.

used an HDD seized during an operation to combat pedophilia in Brazil. The HDD was placed in a Forensic Talon that was connected to the USB port of a laptop computer (Fig. 1). This configuration, which is widely used by forensic examiners at crime scenes, ensured the preservation of all data in the HDD because the Talon acted as a write blocker of the storage device.

The results of the first experiment proved the ability of the NuDetective Forensic Tool to automatically recognize files that may contain nudity. The results of the second experiment proved that the use of NuDetective was feasible, highly reliable, and significantly reduced the time spent by forensic examiners to detect illegal content. NuDetective searched and analyzed all files (>300,000) on the HDD in less than 13 minutes. Nudity detection rates in these experiments were approximately 95%, with low rates of false positives [6].



**Fig. 1.** Configuration widely used by forensic examiners in the analysis of digital storage devices: a write-protected hard disk drive (HDD) connected to the USB port of a laptop computer through a write blocker device.

## 4. Optimization of Automatic Nudity Detection

The authors observed lengthy runtimes for automatic nudity detection with the application of the NuDetective Forensic Tool to material including high-resolution digital photographs, as described in Section 4.2. The use of NuDetective to search for child pornographic material on an HDD containing numerous high-resolution images could thus take more time than desirable at a crime

scene. This section describes our approach to runtime optimization, including experiments conducted to evaluate the success of this approach.

### 4.1. Materials

All experiments in this study were performed using a ThinkPad T60 laptop computer with a 2.33-GHz Intel Centrino processor with 2 GB of RAM and a Mobility Radeon X1400 display adapter running Windows XP (SP3). The results of runtime experiments are directly dependent on the hardware configuration used.

### 4.2. Runtime of Automatic Nudity Detection in High-Resolution Images

To calculate the runtime of automatic nudity detection in high-resolution images using NuDetective, the authors performed experiments with digital photos of different resolutions. Two hundred random photos with 12-MP resolution, some containing human nudity, were initially selected for the experiments. The resolutions of these photos were then reduced to approximately 8.0, 4.0, 2.0, 0.7, and 0.075 MP using Multiple Image Resizer.NET software. We thus performed experiments using six sets of images containing 200 photos each.

To calculate runtimes, these image sets were copied to six different folders. Each folder was submitted to the image analysis provided by NuDetective. The file-name analysis function was disabled. Table 1 shows the results of these trials, including the average runtime for an individual photo in each set. All times presented in Table 1 include the time spent to read the images stored on the HDD. This experiment showed that a single high-resolution image required more than 5 seconds of processing time, which may be undesirable at crime scenes. However, NuDetective rapidly analyzed medium- and low-resolution images, which are most frequently shared through the Internet.

**Table 1.** Runtime of image analysis using the NuDetective Forensic Tool to process six sets of 200 photos each.

Resolution (Megapixels)	Total Time	Average Time per Image
12.0	17m 21s	5.20s
8.0	12m 46s	3.83s
4.0	5m 48s	1.74s
2.0	3m 02s	1.11. 0.91s
0.7	1m 09s	0.34s
0.075	0m 09s	0.04s

### 4.3. Proposed Optimization Approach

Given these results, the authors then sought to reduce the runtime of image analysis using the NuDetective Forensic Tool by reducing image resolution before submitting images to the nudity detection algorithms. The success of this approach depended on the resolution of three major challenges:

- (i) Develop an algorithm that reduces image resolution without contributing to total runtime;
- (ii) Ensure the similarity of nudity-detection results in images with original and reduced resolutions; and

- (iii) Establish an optimal minimum image resolution for use with the NuDetective Forensic Tool.

### 4.4. Implementation of Proposed Optimization

To achieve the main objective of this study, the three major challenges listed in Section 4.3 must be resolved. The approach to each of these challenges is described below, including the development of the algorithm and the experiments performed to evaluate the optimization.

#### 4.4.1. Reducing Image Resolution

The authors developed an algorithm to reduce image resolution based on the use of Java classes **BufferedImage** and **Graphics2D** (Fig. 2). Images with original resolution are represented by **originalImage** and those with reduced resolution are represented by **reducedImage**.

```
BufferedImage reducedImage = new BufferedImage (newImageWidth,
    newImageHeight, originalImage.getType());
Graphics2D graphics = img.createGraphics();
graphics.setRenderingHint(RenderingHints.KEY_INTERPOLATION,
    RenderingHints.VALUE_INTERPOLATION_BILINEAR);
graphics.drawImage(originalImage, 0, 0, newImageWidth, newImageHeight, 0,
    0, originalImageWidth, originalImageHeight, null);
graphics.dispose();
```

**Fig. 2.** Main lines of Java code for the algorithm developed to reduce image resolution.

Experiments using this algorithm to reduce image resolution were performed to calculate the runtime of this procedure. Most of the images used in the trials had 12-MP resolutions. The results showed that the runtime ( $t$ ) to reduce the resolution of an image loaded into the Java program's memory was less than a millisecond ( $t < 0.001s$ ). These results were considered satisfactory because the runtime was very low and did not add significant delays to the image analysis provided by the NuDetective Forensic Tool.

#### 4.4.2. Nudity-Detection Results for Different Image Resolutions

The authors performed experiments using the six sets of 200 photographs to evaluate the similarity of nudity-detection results in images with original and reduced resolutions. The results of NuDetective image analysis are expressed in a binary manner as positive (nudity present) or negative (nudity absent; Table 2). Results for original-resolution (12-MP) images were compared with the results of reduced-resolution sets.

As shown in Table 2, the results of nudity detection by NuDetective were very similar for images of different resolutions, varying by a maximum of 1%. These results were considered satisfactory because they indicated that the reduction of image resolution before NuDetective image analysis did not significantly change detection rates, as shown in Table 2.

**Table 2.** Results of image analysis using the NuDetective Forensic Tool for each set of 200 images.

Resolution (Megapixels)	Positive (Nudity Present)	Negative (Nudity Absent)	Variation (%)
12.0	33	167	(base)
8.0	33	167	0%
4.0	34	166	0.5%
2.0	34	166	0.5%
0.7	33	167	0%
0.075	35	165	1%

#### 4.4.3. Optimal Image Resolution

The authors conducted several other experiments to determine the minimum image resolution successfully analyzed by the nudity detection algorithms of NuDetective without significantly changing the results (less than 1%). After exhaustive tests using various resolutions smaller than 1 MP, the authors concluded that the resolution of 0.01 MP (10 Kilopixels) was most appropriate in terms of runtime and nudity detection results. Thus, the image analysis algorithms reduce image resolutions to a maximum height or width of 100 pixels while maintaining the aspect ratio. With this configuration, we obtained the best runtime and low variation in results (Table 3). The results presented in Table 3 show a 1% variation in nudity detection results obtained with the optimization function enabled, in comparison with those obtained without optimization.

We thus incorporated a new option into the NuDetective Forensic Tool that allows the user to optimize image analysis. All images used in this experiment had original resolutions of 12 MP and were reduced in real time by the new reduction algorithm provided by the Tool. Total

runtime included the time to read the original image files from the HDD, which is a relatively lengthy operation.

**Table 3.** Runtimes and nudity detection results using the optimization option in version 2.0 of the NuDetective Forensic Tool.

Tasks	Optimization Disabled	Optimization Enabled
Total runtime of 200 12-Megapixel images	17 m 21 s	1 m 52 s
Average runtime per image	5.20 s	0.56 s
Number of positive (nudity present)/negative (nudity absent) results	67/133	65/135

#### 4.5. Discussion

Our approach to optimization, which included the reduction of image resolution to 0.01 MP by adjusting the height/width to a maximum of 100 pixels, produced very good results. The optimization function added to the NuDetective Forensic Tool reduced total runtime of image analysis by approximately 90% (17 m 21 s to only 1 m 52 s) without causing significant variation in results (only 2 different results in 200 images). This significant reduction in runtime increases the feasibility of using the NuDetective Forensic Tool to detect child and adolescent pornography at crime scenes.

We also found that images of different resolutions could be successfully processed using the algorithms proposed by Kovac et al. and Ap-apid [17, 21] without significantly changing nudity detection rates (only 1% variation). This variation can be explained by the joining or segmentation of identified skin regions caused by the reduction of image resolution.

### 5. NuDetective 2.0: New Features

Using the findings of this study, the authors developed version 2.0 of the NuDetective Forensic Tool. This section describes the new features of this version and the main improvements made to the tool.

### 5.1. Hash Analysis Implementation and Evaluation

As described in Section 1, the calculation of file hash values is widely used method of searching for previously known content in computer forensics. We thus implemented a new hash analysis feature in version 2.0 of NuDetective. This feature allows the forensic examiner to import a text file containing a list of known hash values for comparison with the hash values of files in selected locals (folders and/or disk partitions). NuDetective 2.0 also allows the user to calculate the hash values of only image/video files, which are most likely relevant to the detection of child and adolescent pornography. The forensic examiner can also choose the hash type to be used: MD5, SHA-1, SHA-256, or SHA-512.

To measure the performance of hash analysis, we submitted the original set of 200 high-resolution (12-MP) images to optimized image analysis with each type of hash analysis provided by the tool. The main goal of this experiment was to assess whether hash analysis introduced runtime delays. The experiment used a text file with a list of 1,000 hash values. Table 4 shows the results of this hash analysis experiment. The MD5 hash type had the fastest runtime, introducing only a 4-second delay (vs. optimized image analysis, Table 3) in the analysis of 200 high-resolution images. MD5 is the recommended hash type for use at crime scenes. Because NuDetective 2.0 was implemented with Java threads, the delay introduced by hash analysis was minimal.

Table 4. Results of the experiment to verify the performance of hash analysis provided by NuDetective 2.0.

Hash Type	Runtime of Optimized Image Analysis with Hash Analysis	Delay Introduced by Hash Analysis
MD5	1 m 56 s	4 s
SHA-1	2 m 03 s	11 s
SHA-256	2 m 05 s	13 s
SHA-512	2 m 09 s	17 s

### 5.2. Other Features

Version 2.0 of NuDetective allows the forensic examiner to combine image, hash, and file-name analyses according to the forensic needs of each case. The settings for each analysis type can also be changed by the user.

Figure 3a illustrates the GUI of version 2.0, which includes new options for the optimization of image and hash analyses. This GUI will support the analysis of video files under development by the authors, as discussed in Section 7. The GUI table of results (Figure 3b) features a new column indicating the results of hash analysis (Hs), which is similar to the presentation of image (Im) and file-name analysis (FN) results. These columns indicate the analysis type by which each file was classified as suspicious.

Version 2.0 of the NuDetective Forensic Tool incorporates many other fixes and internal improvements, including optimization of the source code to search files, updated statistical function and status messages, and new control and error logs.

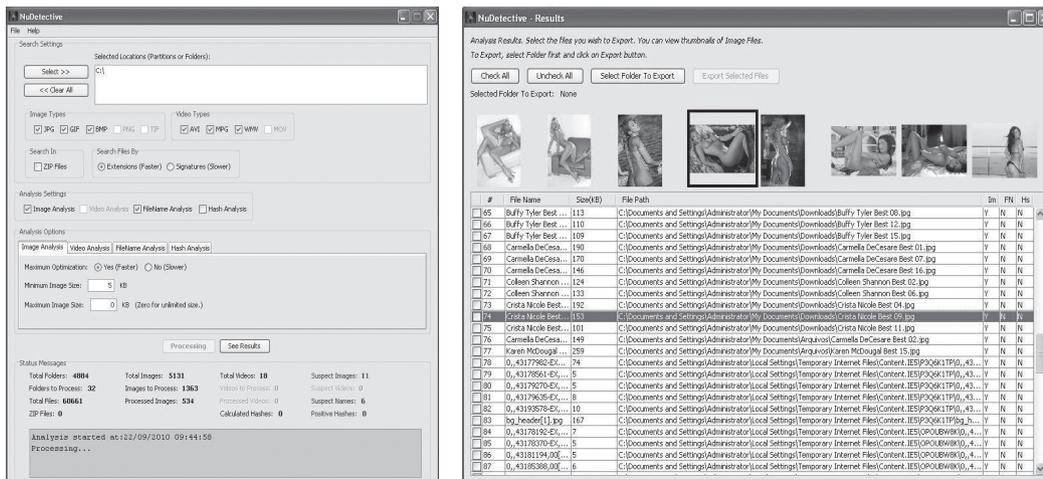


Fig. 3. (a) Main graphical user interface and (b) graphical user interface results table in version 2.0 of the NuDetective Forensic Tool.

## 6. Conclusions

A previous study [6] proved the feasibility of the use of the NuDetective Forensic Tool to detect child and adolescent pornography. The tool's internal algorithms to detect human skin [17] and nudity [21] achieved successful detection rates of 95%, with rapid processing speeds. However, the previous version of NuDetective was weakened by lengthy runtimes (>5 s) in the analysis of high-resolution images, such as 12-MP digital photographs.

With the evolution of digital cameras and camcorders, anyone with such equipment can produce high-quality and high-resolution multimedia content. High-resolution photographs appear with increasing frequency on seized computers and storage devices examined by forensic experts. We thus developed an optimization technique to reduce the runtime of automatic nudity detection in high-resolution digital images with the NuDetective Forensic Tool.

The proposed approach reduces the resolution of images analyzed by the tool's detection algorithms [17, 21]. Through the development of version 2.0 of NuDetective, this study resolved the three main challenges described in Section 4:

- (i) An algorithm to reduce image resolution was implemented in Java using `BufferedImage` and `Graphics2D` classes; this algorithm required less than 0.001 s to reduce a high-resolution image, which was considered to be a very satisfactory result;
- (ii) The results of automatic nudity detection using NuDetective 2.0 showed a maximum variation of 1% between images with original and reduced resolutions, indicating that the proposed approach did not significantly change results; and
- (iii) Image resolution of 0.01 MP reduced runtime and provided good detection results, and was thus selected as the optimal reduced resolution of images submitted to image analysis using NuDetective.

## 7. Future Developments: Video Analysis

Although the NuDetective Forensic Tool has been proven effective for use in forensic cases, it cannot currently be used to analyze video content, a typical medium of storage for material containing nudity. The authors have thus begun a project to implement nudity detection in videos through frame extraction, thereby further improving the NuDetective software. As shown in Fig. 3a, the main GUI of NuDetective 2.0 is ready to support this new feature.

We will attempt to detect nudity in video files by applying the image-analysis algorithms to extracted video frames. In this manner, we will create a multivariate equation that considers variables such as the number of frames containing nudity, the distribution of these frames within the video (timeline), and the content of video titles and subtitles [identified using optical character recognition (OCR)]. Other future improvements have been previously described [6].

### Endnote

The NuDetective Forensic Tool is free and available only to law enforcement and public entities. For instructions on acquiring this tool, please contact the authors or send an e-mail to [nudetective@gmail.com](mailto:nudetective@gmail.com).

### Acknowledgements

The authors thank the Brazilian Federal Police, especially in Mato Grosso do Sul state, for financial and logistic support.

### References

- [1] Saroiu, S. et. al. "A Measurement Study of Peer-to-Peer File Sharing Systems". Proceedings of the Multimedia Computing and Networking (MMCN). San Jose, Jan, 2002.

- [2] BRAZIL. 1990. Estatuto da Criança e do Adolescente. Lei número 8.069 de 13 de Julho de 1990. Brasil : s.n., 1990; [in portuguese].
- [3] BRAZIL. 2003. Lei número 10.764, de 12 de Novembro de 2003. Altera partes do Estatuto da Criança e do Adolescente. Brasil : s.n., 2003; [in portuguese].
- [4] BRAZIL. 2008. Lei número 11829 de 25 de novembro de 2008. Brasil : s.n., 2008; [in portuguese].
- [5] Eleuterio, P. M. S.; Machado, M. P. Identificação de autoria e materialidade em crimes de abuso sexual de criança/adolescente a partir da análise de arquivos multimídia. IV International Conference On Forensic Computer Science (ICoFCS 2009), Natal/RN, Brazil, 2009; [In portuguese].
- [6] Polastro, M. C.; Eleuterio, P. M. S. NuDetective: a Forensic Tool to Help Combat Child Pornography through Automatic Nudity Detection. Proceedings of the Twenty-First International Workshop on Database and Expert Systems Applications (DEXA'10 – VISM'10) - pgs. 349-353. IEEE Computer Society. Bilbao, Spain. Aug, 2010.
- [7] Menezes, A. et. al. "Handbook of Applied Cryptography". CRC Press, 1996.
- [8] Access Data Forensic ToolKit (FTK). Available from <http://www.accessdata.com>. [Visited Dez, 2010].
- [9] Guidance EnCase Forensic. Available from <http://www.guidancesoftware.com/computer-forensics-ediscovery-software-digital-evidence.htm>. [Visited Dez, 2010].
- [10] Pascale, D. A Review of RGB Color Spaces ... from xyY to R'G'B'. The BabelColor Company. Montreal, Canada, 2003.
- [11] Kakumanu, P.; Makrogiannis, P.; Bourbakis, N. A survey of skin-color modeling and detection methods. Pattern Recognition 40, 2007.
- [12] Vezhnevets, V.; Sazonov, V.; Andreeva, A. A Survey on Pixel-Based Skin Color Detection Techniques. Proceedings Graphicon, 2003.
- [13] Wong, K.; Lam, K.; Siu, W.. A robust scheme for live detection of human faces in color images. Signal Processing: Image Communication 18, 2003.
- [14] Albiol, A.; Torres, L.; Delp, E. J. Optimum color spaces for skin detection. International Conference on Image Processing, volume 1, pages 122–124, 2001.
- [15] Elgammal, A.; Muang, C.; Hu, D. Skin Detection - a Short Tutorial. Encyclopedia of Biometrics, Verlag, 2009.
- [16] Jones, M. J., Rehg, J. M. Statistical color models with application to skin detection. International Journal of Computer Vision (IJCV), pages 81–96, 2002.
- [17] Kovac, J.; Peer, P.; Solina, F. Human skin colour clustering for face detection. International Conference on Computer as a Tool EUROCON 2003, Eds. B. Zajc, Ljubljana, Slovenia, September 2003.
- [18] Wang, Y.; Yuan, B. A novel approach for human face detection from color images under complex background. Pattern Recognition 34, pages 1983–1992, 2001.
- [19] Wang, H.; Chang, S. A highly efficient system for automatic face region detection in MPEG video, IEEE Trans. Circuits Systems Video Technol., pages 615-628, 1997.
- [20] Carlsson, A.; Eriksson, A.; Isik, M. Automatic detection of images containing nudity. Master thesis in intelligent systems Design, University of Gothenburg, Sweden, 2008.
- [21] Ap-apid, R. An Algorithm for Nudity Detection. Proceedings of the 5th Philippine Computing Science Congress, Cebu City, 2005.
- [22] Lin, Y.; Tseng, H.; Fuh, C. Pornography Detection Using Support Vector Machine. 16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP). Kinmen, ROC, 2003.
- [23] Wang, J.; Li, J.; Wiederhold, G.; Firschein, O. System for Screening Objectionable Images Using Daubechies' Wavelets and Color Histograms. Proc. of the International Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services, pages 10–30, 1997.
- [24] Java SE Overview. Available from <http://java.sun.com/javase/>; [Visited Dez, 2010].
- [25] Image After. Available from <http://www.imageafter.com>; [Visited Dez, 2010].



**Pedro Monteiro da Silva Eleuterio** is a Computer Engineer at the Federal University of Sao Carlos (UFSCar), Brazil. He obtained his Master's degree in Computer Science with a specialization in hypermedia at the University of Sao Paulo (USP), Brazil. Since 2006, he has worked as a criminal forensic expert for the Brazilian Federal Police (DPF).



**Mateus de Castro Polastro** obtained his bachelor's degree in Computer Science at the University of Campinas (Unicamp), Brazil. He is currently enrolled in the M.Sc. program in Computer Forensics at the University of Brasilia (UnB), Brazil. Since 2007, he has worked as a criminal forensic expert for the Brazilian Federal Police (DPF).