A Benchmark Methodology for Child Pornography Detection

João Macedo[1,2], Filipe Costa[3], Jefersson A. dos Santos[1]

[1]*Department of Computer Science, Universidade Federal de Minas Gerais, Brazil*
[2]*Brazilian Federal Police, Belo Horizonte, Minas Gerais, Brazil*
[3]*CPqD – Image and Speech Processing Management, Campinas, São Paulo, Brazil*
{joaomacedo,jefersson}@dcc.ufmg.br, fcosta@cpqd.com.br

***Abstract*—The acquisition and distribution of child sexual content are some of the most important concerns for legislative systems and law enforcement agencies around the world. There is a great demand for automatic detection of child pornography, mainly due to the large amount of existent data and the facility someone can share this content over the internet. Although there are some proposed methods to automatically detect child pornography content in the literature, there is no available dataset to assess and compare the performance of these methods due to legal restrictions, considering that in many countries the distribution or possession of this material is a crime by Law. To mitigate this problem, we work with the Brazilian Federal Police to structure and organize a benchmark methodology for child pornography to make it possible the comparison of distinct categories of child pornography detectors. Therefore, we present in this paper the used methodology for the creation of a new annotated dataset of images of child pornography. We also propose a child pornography detection step-wise methodology based on automatic age estimation combined with a pornography detector, which is evaluated using the described benchmark dataset. The proposed approach achieved results (79.84% accuracy) that overcome two tools currently used by the Brazilian Federal Police.**

I. Introduction

The ever-growing internet access around the world has been an enabling factor for the production and diffusion of all kind of data. Within this new reality, multimedia documents are published and shared effortlessly. This scenario easily leads to sharing of illegal or abusive contents, including child sex abuse material.

Although it is not a new problem, child pornography content has been broadly disseminated over the past decades, mainly because of the advent of the internet, mobile technologies, social media, and peer-to-peer networks. In an attempt to tackle this issue, law enforcement agencies have been approaching the problem in two ways [1]. The first is by internet monitoring, with focus on the distribution and consumption of already known illicit files or any suspect file based on automatic detection methods, aiming at restricting the access to such content and to identify consumers and distributors in peer-to-peer networks [2], file sharing applications, social media, and websites. The second occurs while performing forensic examination of suspected storage medium, with the goal of detecting the possession of child pornography material and any evidence of distribution or production of this content. This task has been traditionally made through visual analysis by a forensic expert. Because of this, it is very time-consuming, expensive and stressful for those who are constantly exposed to these content, with the risk of causing negative impacts on the involved professionals [3, 4] and generally under considerable time pressure.

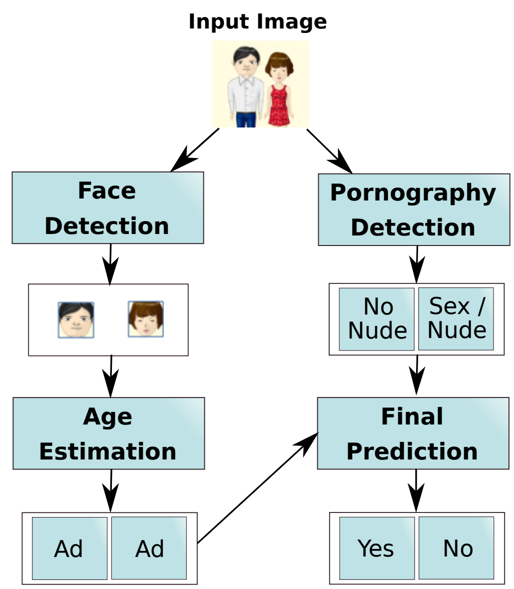

Fig. 1. Overview of the proposed methodology for child pornography detection. The method firstly detects if an image has pornographic content and then detects the faces in the image and estimates its corresponding ages to finally classify whether the image contains child pornography. The proposed method is based on the considerations that: a) many child pornography images and videos include the face of the victims; and b) pornography detection is an easier task than child pornography detection and assumes the hypothesis that a child pornography detector can be built through the combined analysis of a pornography detector and an age estimator.

In recent years, automatic tools [5] have been developed to support the analysis of child pornography content. These tools are based on distinct methods, which can be grouped as follows: (1) methods that search for files whose names are potentially correlated to child pornography; (2) methods that compare the hash of each file with a database of hashes of already known child pornography files; (3) methods that extract visual information from images and videos, such as color (skin color) and texture [1] or learned features [6]; and (4) methods that combine auxiliary classifiers [7], such as a face detector and an age estimator, to detect child pornography indirectly. The auxiliary classifiers are built with the same methods employed by the third group's methods.

The methods from the first group are very faulty. For instance, since they do not perform any analysis of the file's content, they usually do not detect renamed files or files with generic names. The methods from the second group can be very efficient, but they fail to detect unknown files or modified versions of known files. The methods from the third and fourth groups are more general and hence are better suited to detect child pornography. It must be said, however, that the simpler methods that rely on skin color analysis usually lead to high false positive rates. More sophisticated methods can be more accurate and can bring along other advantages, like context or scenario cues, and additional information about individuals in the image, such as age ranges, gender, and ethnic group.

The greatest barrier to the development of new techniques [7] in this area is the unavailability of suitable datasets, due to the illegal nature of the content, since the possession of this material is forbidden by law in many countries. Thereby, researchers have no common base to evaluate or compare their methods with the research community results.

To fill the gap of child pornography test dataset unavailability, this paper introduces a region-based annotated dataset of images related to child pornography. Furthermore, a child pornography detection solution is proposed (Figure 1). The benchmark dataset introduced in this paper was used for evaluating the proposed approach, that achieved an accuracy of 79.84% in child pornography detection.

The main contributions of this work are:

- The introduction of a region-based annotated dataset belonging to the Brazilian Federal Police, that may be used for testing new methods and can stimulate further research in this area.
- The evaluation of an approach to child pornography detection based on the combination of auxiliary classifiers.
- The development of an age estimation technique that is used as one of the auxiliary classifiers in the proposed child pornography detection method.

The remainder of the paper is organized as follows. Section II presents the relevant literature. In Section III, we describe the region-based annotated child pornography dataset. Section IV presents the proposed approach for child pornography detection. Section V describes the experimental evaluation and Section VI discusses the results. The conclusion is presented in Section VII.

II. Related Work

There are few works specifically developed for child pornography detection. The detection methods are usually built as a single-tier classifier, which directly detects child pornographic content, or as a multi-tier classifier, that combines auxiliary classifiers and detectors, such as a pornography detector and an age estimator, to predict child pornography content.

In this section, we present the main works related to child pornography detection. Since the multi-tier methods, such as the one proposed in this paper, rely on pornography detection and age estimation methods, we also review the main contributions in these areas.

A. Child Pornography Detection

The first works on child pornography detection through content analysis were based on statistical classification of texture and color information. Sae-Bae et al. [1] proposed an approach which considers skin color filters to identify human skin tones and explicit content. Then the faces and their corresponding landmarks are detected and their ages are estimated using a SVM classifier. This method was evaluated with a private dataset and achieved an accuracy of 82.86% in detecting pornographic content and a true positive rate of 89.53% in the classification of the detected child faces.

Yiallourou et al. [7] built a synthetic dataset of images with possibly suspicious content and devised a method that performs face detection using Haar feature-based cascade classifiers. Then the age and gender of the faces are estimated, and the lighting intensity of the image is calculated. The combination of these steps results in five image features that denote child presence, number of people, age diversity, gender distribution and lightning. These features were used to train a regression model to classify the images as appropriate, neutral or inappropriate, achieving an accuracy of 48% in testing.

Vitorino et al. [6] developed a method that uses *convolutional neural networks* (CNN) to predict child pornography from images directly, i.e., without using a combination of correlated predictors. The work distinguishes normal images from adult pornographic content and from child pornographic content, indicating levels of pornography. As far as we are concerned, this is the unique method that uses CNN to predict child pornography. One of the reasons behind this is the fact that datasets with child pornographic content are not available, since the possession of this material is illegal. They only exist in police agencies, such as the Child Exploitation Obscene Reference File [8], maintained by the FBI.

B. Age Estimation from Face Images

Since the pioneering works in aging estimation from face images, many different methods and techniques have been proposed to target the problem. The approaches adopted in those methods somehow reflect the advances of the visual pattern recognition techniques. The first attempts to the age estimation problem were mainly based on image processing techniques aimed at extracting visual information such as facial landmarks [9] and its ratios and proportions, which were directly used to predict the age.

Instead of searching for local features, such as geometric proportions and wrinkles, part of the research community started to apply efforts to develop matching techniques based on models. These techniques usually adopt a statistical approach to represent the face through dimensionality reduction and in a second step apply regression techniques to perform age estimation. The active appearance model (AAM) is one of these techniques, which tries to understand new images based on similar synthetic images, using a parametrized model of appearance [10]. The method proposed in [11] generates a statistical model of facial appearance based on the face images and performs Principal Component Analysis (PCA) for dimensionality reduction.

In the AGES (AGing pattErn Subspace) [12], the method aims at building aging patterns of different people in order to form a representative subspace of aging patterns. In the Bio-Inspired Features methods [13], the age estimation method is based on an object recognition technique that models the visual processing in the cortex proposed by [14] and later extended by [15]. Some other methods make use of general purpose descriptors, such as Local Binary Patterns (LBP) and Histograms of Oriented Gradients (HOG), and then choose one machine learning algorithm, such as Support Vector Machines (SVM) for training a classifier with the extracted features.

Recently, some researchers have proposed deep-learning based approaches [16–21]. Huerta et al. [16], proposed a fusion of well-known descriptors that were compared to a Deep ConvNet approach. The authors used the LeNet CNN [22] as the base architecture for the proposed network, and modified its parameters, the number of convolution layers and the fully connected layers to create a new architecture, better suited for the age estimation task. The best architecture achieved a mean absolute error (MAE) of 3.88 for the MORPH dataset [23] and of 3.31 for the FRGC dataset.

Levi and Hassner [17] proposed a small CNN with three convolutional layers followed by three fully connected layers to predict age and gender using the Adience dataset [24], achieving an accuracy of 50.7 ± 5.1 using a 5-fold cross-validation protocol. Duan et al. [18] combined a CNN and an extreme learning machine to perform age and gender classification using the Morph-II and Adience benchmarks, achieving an accuracy of 52.3 ± 5.7 in the latter one.

In [19], the authors exploited the ordinal relation between ages, proposing an architecture that consists of a series of CNNs, one for each age or age group, where each CNN yields a binary output that tells if the face's age is higher or lower than a certain value. All the binary outputs are aggregated to make the final age prediction. Age order information is also exploited in [21]. Liu et al. proposed a label-sensitive deep metric learning method which uses a deep residual network to learn distances between ages. Ranjan et al. [20] extended a previous work [25] and designed a multi-purpose CNN model that performs face detection, landmarks localization, pose estimation, gender recognition, smile detection and age estimation, exploring the idea of joint learning correlated tasks, sharing parameters in the lower layers.

C. Pornography detection

In the same way as the child pornography detection methods, the works on pornography detection were primarily based on statistical classification of texture and color information [26, 27]. The main disadvantage of these methods is the high false positive rate, because many things in an image can have similar color to the human skin. Some more sophisticated works were developed, such as the combination of low-level descriptors with mid-level representations [28] to detect pornographic content in videos. More recently, some works based on deep learning approaches were developed, such as Nian's et al. [29] proposed architecture, with five convolutional layers followed by three fully connected layers; Yahoo!'s open-sourced architecture [30], based on the ResNet-50 architecture with half number of filters in each layer, which was pre-trained with ImageNet and fine-tuned with a proprietary not suitable/safe for work (NSFW) dataset, and Perez et al. [31] work, that exploits motion information in video files to detect pornography with a convolutional neural network.

III. Region-based Annotated Child Pornography Dataset

The region-based annotated child pornography dataset (RCPD) described in this paper was created in collaboration with the Brazilian Federal Police. The child pornographic images within it were gathered and labeled by computer forensic experts under our guidance and its contents cannot be illustrated in this paper or made public due to its illicit nature. It must be emphasized that the authors, as researchers, never accessed, visualized or possessed the images.

The aim of the dataset is to assess and compare the performance of child pornography detection methods. We hope that this initiative will boost the development of new approaches to this important application of the forensic field. Researchers may submit their algorithms following the instructions provided at http://www.patreo.dcc.ufmg.br/rcpd.

The dataset consists of 2138 images, including 508 images with no people and 1630 images containing individuals showing or not the face. It is important to have a reasonable number of images not containing any person to allow the evaluation of false positive rates. Face occlusions are also important to make the dataset more robust and realistic.

The individuals in the 1630 images may have multiple labels, such as age, gender and nudity exposure. The face and relevant parts of the individuals are also labeled, and all these information are used to determine the label of the images. Thereby, for each image it is possible to know the age of the youngest person, the age of the youngest shown face and the most severe nudity exposure, that can vary from no one to seminude, nude and sex. From the image labels, it is possible to answer many queries, such as if there is any child in the image, or even more complex queries, such as if the image has nude or sexual content and a child showing her face.

The dataset's total number of images and the quantity of images in the two major categories are shown in Table I. The numbers returned by some queries of interest to the dataset's objectives are shown in the nine last lines of the table.

TABLE I
IMAGES IN DATASET

Category	Number of images
All	2138
No person	508
Person	1630
Person showing face	1455
Child	1238
Child face	1065
Seminude, nude or sex	1407
Face + seminude, nude or sex	1233
Child + seminude, nude or sex	1051
Child face + seminude, nude or sex	879
Child + nude or sex	836
Child face + nude or sex	664

The age ranges in the dataset correspond to the age ranges used in the Adience dataset [17]: $0-2$, $4-6$, $8-13$, $15-20$, $25-32$, $38-43$, $48-53$ and $60-$. Tables III and IV illustrate how the nudity categories in the dataset's images are distributed by the youngest person in the images and by the youngest shown face in the images, respectively.

The bounding boxes of the individuals and their parts are annotated, totaling 5111 objects in the whole dataset. The annotated regions are the body, face, breast and private parts.

Figure 2 presents a scheme of the annotation process of two people and their faces. The annotation indicates the bounding boxes or regions of a 25-year-old man and of a 20-year-old woman, and two smaller bounding boxes of their faces. With this information, that image has no nude person, the younger person depicted has 20 years old and the younger face depicted belongs to a 20-year-old.

In Figure 3, the annotation indicates a bounding box of a 25-year-old man with the *nude* tag, represented in the figure by the greater rectangle, and the bounding boxes of his face, breast and privates.

TABLE II
PARTS ANNOTATIONS IN THE DATASET

Age	Individuals	Faces	Breasts	Privates	All
0-2	91	71	7	82	251
4-6	210	195	42	101	548
8-13	963	923	309	383	2578
15-20	232	226	161	78	697
25-32	186	180	130	111	607
38-43	67	55	42	58	222
48-53	25	23	11	14	73
60-	46	46	22	21	135
Total	1820	1719	724	848	5111

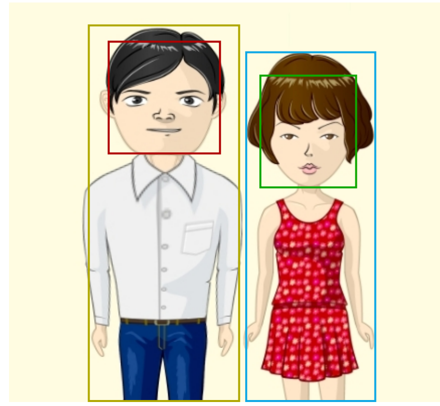

Fig. 2. Illustration of region-based annotations of the bodies and faces of a man and a woman.

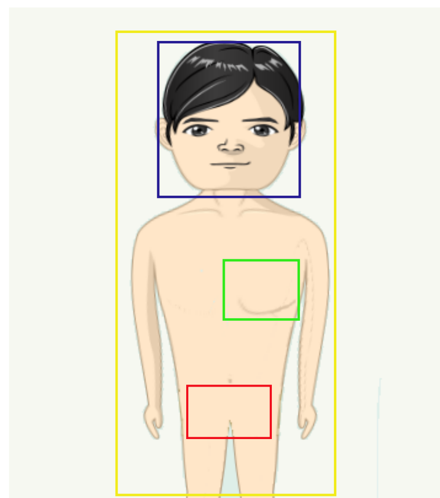

Fig. 3. Annotation representation for nude parts of an image. The dataset has bounding boxes for the face, breast and genitals.

The choice over the design of the labels allows to assess child pornography methods that perform whole image classification and methods that combine facial age classification with pornography detection. It is also possible to assess methods designed to detect sensitive body parts.

IV. PROPOSED METHODOLOGY FOR CHILD PORNOGRAPHY DETECTION

A. Overview

The proposed method combines a pornography detector and a child face detector to determine whether an image has child pornography. If the pornography detector indicates that an input image is pornographic and the face detector finds a child's face, it concludes that the image contains child pornography.

The child face detector has two auxiliary methods: a face detector and an age estimator, which tells whether a detected

TABLE III
NUDITY EXPOSURE IN AGE GROUPS

Category	0-2	4-6	8-13	15-20	25-32	38-43	48-53	60-	All
Nonude	9	21	157	22	5	1	1	7	223
Seminude	5	25	185	46	16	4	5	2	288
Nude	80	128	454	100	108	25	8	9	912
Sex	44	51	78	8	14	4	0	8	207
Total	138	225	874	176	143	34	14	26	1630

TABLE IV
NUDITY EXPOSURE IN AGE GROUPS BY FACE

Category	0-2	4-6	8-13	15-20	25-32	38-43	48-53	60-	All
Nonude	9	20	157	22	5	1	1	7	222
Seminude	5	25	185	46	16	4	5	2	288
Nude	42	100	404	98	107	26	8	10	795
Sex	10	35	73	8	12	4	0	8	150
Total	66	180	819	174	140	35	14	27	1455

face belongs to a child. It is important to notice that all the thresholds and parameters mentioned were selected after preliminary experiments that were not reported due to space restrictions.

Figure 1 depicts the pipeline of the proposed approach, presenting an example of how an image with two adults showing their faces would be processed. In the first step, the image is submitted to a pornography detector, that yields a number between 0 and 1 corresponding to the pornography probability of the image. The proposed method uses a predefined threshold τ (in our experiments, $\tau = 0.3$) to identify whether the predicted probability indicates pornography in the image: the image is considered not pornographic if the probability is lower than τ, and pornographic otherwise. The threshold τ was empirically chosen after evaluating results on a dataset of pornography images.

In the example, if the predicted pornography probability of the image is lower than τ, the execution will finish at this point with a negative answer to child pornography. If the predicted pornography probability is greater than τ, the child detection module illustrated in the left branch of the pipeline will be executed. In this module, the faces would be detected, extracted, aligned and submitted to the age estimator, that would report two adults in the image. Next, the final prediction would combine the results from the pornography detector and from the age estimator, concluding that there is no child pornography in the image, because even having detected pornography in the image, no child face was detected by the age estimator. In the following subsections, the auxiliary classifiers and detectors will be described in more detail.

B. Face Detection

For the face detection task of the pipeline, the proposed method uses the MTCNN face detector [32], that employs a cascaded architecture consisting of three CNNs, each one responsible for executing one of the three consecutive steps of the face detection method. In the first step, one CNN predicts potential faces, in the second step, the initial detection is refined, and in the last step, the third CNN further refines the results and adds faces landmarks.

C. Age Estimation

The age estimation task that integrates the child pornography method is performed by a convolutional neural network adapted for this work, fine-tuned and evaluated on the Adience dataset [17] using the five-fold cross-validation protocol described by Eidinger et al. [24].

The Adience dataset contains real-world unconstrained images with faces classified by age groups ($0-2$, $4-6$, $8-13$, $15-20$, $25-32$, $38-43$, $48-53$, $60-$) and by gender. There is no label for child and adult, and this information was inferred by the age groups. For the purposes of this work, we considered the three first age groups as child ($0-2$, $4-6$, $8-13$) and the other five age groups as adult.

Due to the small number of samples in the target dataset, we started with a VGG-16 [33] architecture pre-trained for the ImageNet dataset [34] to avoid over-fitting and to accelerate the learning process. The three last fully connected layers were replaced by two fully connected layers with 4096 channels followed by three soft-max layers to simultaneously learn the age groups, child detection and gender classification. The overview of the architecture, showing the three final prediction layers, is depicted in Figure 4.

D. Pornography Detection

For the pornography detection task of the pipeline, we used a recently published adult content classification network open sourced by Yahoo [30]. The network detects adult content and outputs a probability from 0 to 1 for each input image. It is based on the ResNet-50 architecture [35] with half the number of filters in each layer. The network was pre-trained

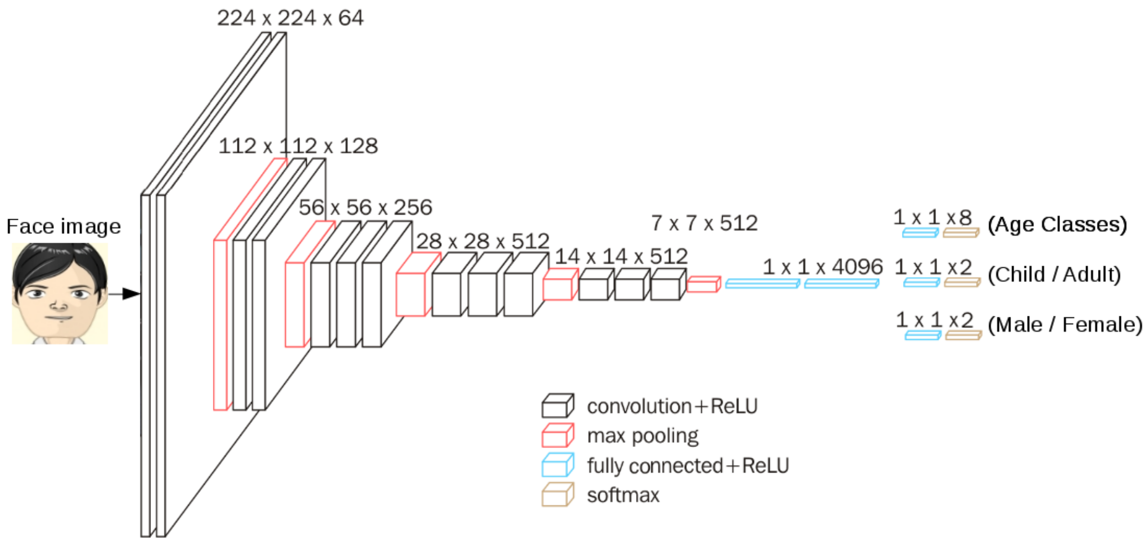

Fig. 4. VGG-16 adapted architecture.

with ImageNet dataset and fine-tuned with a proprietary not suitable/safe for work (NSFW) dataset.

V. Experiments

This section describes the experiments of the proposed method on the RCPD dataset introduced in this work. We organize the experiments into two subsets. The first one is focused on the evaluation of the child face detector. The last one groups the pornography detector evaluation and the combined approach for child pornography detection.

A. Child Face Detection

The child face detection is composed by a face detection method and an age estimation method. For the face detection task, the MTCNN face detector [32] was chosen. This detector was compared to other methods commonly used for this task, such as Haar Feature-based Cascade face detection and Dlib [36], and it was chosen for its better detection rate and good performance. For the age estimation task, a CNN was adapted for this work and fine-tuned on the Adience dataset [17] to simultaneously learn the age groups, child detection and gender classification. To train the adapted network, the MTCNN face detector was used to extract the faces from the Adience dataset and a data augmentation strategy based on flipping and slightly rotating the faces was employed, aiming at avoiding over-fitting and improving the classification performance.

The results of the adapted network for the age estimation task are presented in Table V. Although the adapted network accuracy is below the state-of-art results, it should be noticed that the above 60.0% accuracy methods perform an extra fine-tuning step using the IMDB-WIKI-101 dataset [37], which is the largest in the wild face images dataset up to date, with more than 500.000 images. It is important to note, however, that for child/adult classification, which is the main interest for

TABLE V
Age estimation results on the Adience benchmark

Method	Accuracy $\pm$ Std. Dev. (%)
Eidinger et al. [24]	45.1 ± 2.6
Levi and Hassner [17]	50.7 ± 5.1
Rothe et al. [38] w/o IMDB-WIKI pretrain	55.6 ± 6.1
Rothe et al. [38] w/ IMDB-WIKI pretrain	64.0 ± 4.2
Hou et al. [39] w/ IMDB-WIKI pretrain	67.3
Best from Zhang [32]	67.3 ± 3.6
Proposed method	56.8 ± 6.0

this work, the adapted model achieved an accuracy of 94.1 ± 2.3.

B. Child Pornography Detection

The proposed approach was evaluated on the RCPD dataset described in this work. Since the pornography detector outputs a probability to each image, we set a threshold $\tau = 0.3$ to filter out images containing nude or sexual content and a threshold $\tau = 0.1$ to filter out images containing seminude, nude or sexual content.

The accuracy in the detection of child pornography on the RCPD dataset with this configuration ($\tau = 0.3$) was 79.84%. The auxiliary classification models used in the pipeline had an accuracy of 82.55% for the child detection task and 85.78% for the pornography (nude or sex) classification task step, as shown in Table VI.

We evaluated the proposed method to perform a looser classification to also identify seminude content, besides nude and sexual content. This configuration uses a lower threshold ($\tau = 0.1$) for the NSFW classifier and targets child related images containing seminude, nude or sexual content (seminude+). This kind of broader classification can be of interest in specific forensic scenarios where it may be necessary to care about near pornography pictures involving children. The results are shown in Table VII.

TABLE VI
CHILD PORNOGRAPHY DETECTION (NUDE AND SEX)

Metric	Child Porn	Child Detection	Porn Detection
Accuracy (%)	79.84	82.55	85.78
Precision (%)	68.64	90.80	80.99
Recall (%)	64.61	72.30	95.17
F1-score (%)	66.56	80.50	87.51

Two forensic tools from the Brazilian Federal Police were evaluated against the RCPD dataset. The first of them was a tool named NuDetective [27], which is based on skin color analysis, and the other was a tool named LED, which aims to find already known digital evidence. NuDetective achieved an accuracy of 57.43%, while LED achieved a better accuracy of 76.47%, with lower precision and higher recall rates. The results are presented in Table VIII, where we added the results of our proposed method for comparison purposes.

TABLE VII
CHILD PORNOGRAPHY DETECTION (SEMINUDE, NUDE AND SEX)

Metric	Child Seminude+	Child Detection	Seminude+
Accuracy (%)	81.15	82.55	94.11
Precision (%)	83.90	90.80	94.45
Recall (%)	67.01	72.30	96.73
F1-score (%)	74.50	80.50	95.57

The experiments were performed on an Intel Xeon E5-2630@2.6Hz with an NVIDIA GTX 1080 TI graphics card. The proposed method takes an average time of 0.38 s per image: 0.05 s on the pornography classification task, 0.32 s on the face detection and alignment tasks, and 0.01 s on the age estimation task.

VI. DISCUSSION

Experimental evaluation with the proposed method on the RCPD dataset yielded an accuracy of 79.84%. It demonstrates the suitability of the proposed face-based child detection combined with a pornography detector. At the same time, it shows that there is room for improvement in this challenging application.

Although the method succeeded in achieving better results than the analyzed forensic tools, it requires more resources to process the images, which we do not consider a limitation factor, and it has a weakness when it comes to child pornography images without child's faces.

In respect to the forensic tools, one of the main drawbacks of pornography or child pornography detection methods based on skin color analysis is the high rate of false positives, and this behavior was observed in the results of NuDetective: the tool selected 1456 images, while there are just 836 files related to child pornography. It classified as positive 155 files that did not contain any person, and had a relatively low recall rate, which is not desirable for this kind of application.

On the other hand, the analysis of LED results indicates that the tool had a good performance on the dataset, achieving an accuracy of 76.47% with a relatively good precision and a recall rate better than the one achieved by NuDetective, while presented just a small number of false positives. These results demonstrate the power of this category of tools when the search is conducted against already known files, as it happens with RCPD dataset, that had its files gathered from real cases. In these situations, hash-based tools can achieve good results with the additional advantage of being faster than any other approach, which makes it especially suited for use in the field.

Despite the good performance of LED, it is important to emphasize that hash-based approaches usually fail when there are few known files. In these cases, a method like the proposed in this work usually has better results. In the specific case of the RCPD, the proposed method achieved a better accuracy, with higher recall rate, even considering a dataset that had its images selected from real cases.

TABLE VIII
EVALUATION OF FORENSIC TOOLS

Metric	NuDetective [27]	LED	Proposed Approach
Accuracy (%)	57.43	76.47	**79.84**
Precision (%)	**78.74**	75.34	68.64
Recall (%)	41.24	57.21	**64.61**
F1-score (%)	54.13	66.30	**66.56**

VII. CONCLUSION

In this work, we introduced a child pornography region-based annotated dataset (RCPD) belonging to the Brazilian Federal Police. The dataset contains images with child pornography content gathered and labeled by computer forensic experts. These images were mixed up with groups of images not related to child pornography, including images with no person, images with adults, images with children and pornographic images with adults.

We also proposed a combined method to detect child pornography using a child face detection module and a pornography detector. The proposed method classifies images as related to child pornography only if the images contain pornography and if a child face is detected. The child face detection module used in the proposed method was built through the integration of a face detector and an age estimation method that was specifically developed for this work. The proposed method was evaluated against the RCPD dataset and achieved an accuracy of 79.84%, which was better than the results achieved by two forensic tools evaluated in the same dataset, LED and NuDetective. The experiments also demonstrated the usability of the described dataset to compare child pornography detection methods, and although the potential of the dataset was not fully exploited, what could include metrics related to the objects' annotations, the simple detection analysis already showed that the existence of a benchmark dataset in this area can form a basis for further research, comparison and optimization of methods. These results also showed the viability of the strategy of combining different classifiers and detectors to perform the child pornography detection task, but there is still much room for improvement, either in the age estimation task as in the pornography classification task. The proposed method can be extended to videos using a sampling approach to select a subset of frames.

Acknowledgment

The authors would like to thank the Brazilian Federal Police and CPqD foundation for the collaboration. We are grateful to NVIDIA for the donation of the GPUs that allowed the execution of our experiments. We also thank CAPES, CNPq, and FAPEMIG (APQ-00449-17) for the financial support provided for this research project.

References

[1] N. Sae-Bae, X. Sun, H. T. Sencar, and N. D. Memon, "Towards automatic detection of child pornography," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5332–5336.

[2] R. Hurley, S. Prusty, H. Soroush, R. J. Walls, J. Albrecht, E. Cecchet, B. N. Levine, M. Liberatore, and B. Lynn, "Measurement and analysis of child pornography trafficking on gnutella and emule," *Amherst, US: Department of Computer Science, University of Massachusetts. Retrieved from http://scholar. googleusercontent. com/scholar*, 2012.

[3] M. Krause, "Identifying and managing stress in child pornography and child exploitation investigators," *Journal of Police and Criminal Psychology*, vol. 24, no. 1, pp. 22–29, 2009.

[4] R. J. Edelmann, "Exposure to child abuse images as part of one's work: Possible psychological implications," *The Journal of Forensic Psychiatry & Psychology*, vol. 21, no. 4, pp. 481–489, 2010.

[5] M. Thakor, "How to look: Apprehension, forensic craft, and the classification of child exploitation images," *IEEE Annals of the History of Computing*, vol. 39, no. 2, pp. 6–8, 2017.

[6] P. Vitorino, S. Avila, M. Perez, and A. Rocha, "Leveraging deep neural networks to fight child pornography in the age of social media," *Journal of Visual Communication and Image Representation*, vol. 50, pp. 303–313, 2018.

[7] E. Yiallourou, R. Demetriou, and A. Lanitis, "On the detection of images containing child-pornographic material," in *Proc. 24th Int. Conf. Telecommunications (ICT)*, May 2017, pp. 1–5.

[8] Forensic database tech digital evidence table. [Online]. Available: https://www.nist.gov/forensics/forensic-database-tech-digital-evidence-table

[9] Y. Kwon and N. Lobo, "Locating facial features for age classification," in *Proceedings of SPIE-the International Society for Optical Engineering Conference*, 1993, pp. 62–72.

[10] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer vision and image understanding*, vol. 61, no. 1, pp. 38–59, 1995.

[11] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation," *Part B (Cybernetics) IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, no. 1, pp. 621–628, Feb. 2004.

[12] X. Geng, Z.-H. Zhou, Y. Zhang, G. Li, and H. Dai, "Learning from facial aging patterns for automatic age estimation," 2006.

[13] G. Guo, G. Mu, Y. Fu, and T. S. Huang, "Human age estimation using bio-inspired features," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Jun. 2009, pp. 112–119.

[14] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," pp. 1019–1025, 1999.

[15] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 411–426, Mar. 2007.

[16] I. Huerta, C. Fernández, C. Segura, J. Hernando, and A. Prati, "A deep analysis on age estimation," *Pattern Recognition Letters*, vol. 68, pp. 239–249, 2015.

[17] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops (CVPRW)*, Jun. 2015, pp. 34–42.

[18] M. Duan, K. Li, C. Yang, and K. Li, "A hybrid deep learning cnn–elm for age and gender classification," *Neurocomputing*, vol. 275, pp. 448–461, 2018.

[19] S. Chen, C. Zhang, M. Dong, J. Le, and M. Rao, "Using ranking-CNN for age estimation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 742–751.

[20] R. Ranjan, S. Sankaranarayanan, C. D. Castillo, and R. Chellappa, "An all-in-one convolutional neural network for face analysis," in *Proc. 12th IEEE Int. Conf. Automatic Face Gesture Recognition (FG 2017)*, May 2017, pp. 17–24.

[21] Z. Niu, M. Zhou, L. Wang, X. Gao, and G. Hua, "Ordinal regression with multiple output CNN for age estimation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 4920–4928.

[22] Y. Lecun, B. L., Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[23] K. Ricanek and T. Tesafaye, "Morph: a longitudinal image database of normal adult age-progression," in *Proc. 7th Int. Conf. Automatic Face and Gesture Recognition (FGR06)*, Apr. 2006, pp. 341–345.

[24] E. Eidinger, R. Enbar, and T. Hassner, "Age and gender estimation of unfiltered faces," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2170–2179, Dec. 2014.

[25] R. Ranjan, V. M. Patel, and R. Chellappa, "Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1, 2017.

[26] S. J. Sathish and S. H. Sengamedu, "Texture-based pornography detection," Jul. 3 2008, uS Patent App. 11/715,051.

[27] M. d. C. Polastro and P. M. d. S. Eleuterio, "Nudetective: A forensic tool to help combat child pornography through automatic nudity detection," in *Database and Expert Systems Applications (DEXA), 2010 Workshop on*. IEEE, 2010, pp. 349–353.

[28] C. Caetano, S. Avila, S. Guimares, and A. d. A. Arajo, "Pornography detection using bossanova video descriptor," in *Proc. 22nd European Signal Processing Conf. (EUSIPCO)*, Sep. 2014, pp. 1681–1685.

[29] F. Nian, T. Li, Y. Wang, M. Xu, and J. Wu, "Pornographic image detection utilizing deep convolutional neural networks," *Neurocomputing*, vol. 120, pp. 283–293, 2016.

[30] J. Mahadeokar and G. Pesavento. (2016) Open sourcing a deep learning solution for detecting nsfw images. [Online]. Available: https://yahooeng.tumblr.com/post/151148689421/open-sourcing-a-deep-learning-solution-for

[31] M. Perez, S. Avila, D. Moreira, D. Moraes, V. Testoni, E. Valle, S. Goldenstein, and A. Rocha, "Video pornography detection through deep learning techniques and motion information," *Neurocomputing*, vol. 230, pp. 279–293, 2017.

[32] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[33] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[34] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778.

[36] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

[37] R. Rothe, R. Timofte, and L. V. Gool, "Dex: Deep expectation of apparent age from a single image," in *Proc. IEEE Int. Conf. Computer Vision Workshop (ICCVW)*, Dec. 2015, pp. 252–257.

[38] R. Rothe, R. Timofte, and L. Van Gool, "Deep expectation of real and apparent age from a single image without facial landmarks," *International Journal of Computer Vision*, vol. 126, no. 2-4, pp. 144–157, 2018.

[39] L. Hou, D. Samaras, T. Kurc, Y. Gao, and J. Saltz, "Convnets with smooth adaptive activation functions for regression," in *Artificial Intelligence and Statistics*, 2017, pp. 430–439.