Private Body Part Detection using Deep Learning

André Tabone[1][a], Alexandra Bonnici[1][b], Stefania Cristina[1][c], Reuben Farrugia[2][d] and Kenneth Camilleri[1][e]

[1]*Department of Systems and Control, University of Malta, Malta*

[2]*Department of Computer Engineering, University of Malta, Malta*

Keywords: Deep Neural Networks, Pornographic Detection, Classifiers, Private Body Part Detector.

Abstract: Fast and accurate detection of sexually exploitative imagery is necessary for law enforcement agencies to allow for prosecution of suspect individuals. In literature, techniques which can be used to assist law enforcement agencies only determine whether the image content is pornographic or benign. In this paper, we provide a review on classical handcrafted-feature based and deep-learning based pornographic detection in images and describe a framework which goes beyond this, to identify the location of genitalia in the image. Despite this being a computationally complex task, we show that by learning multiple features, a MobileNet framework can achieve an accuracy of 76.29% in the correct labelling of female and male sexual organs.

1 INTRODUCTION

Every month, a new mass of data pertaining to pornography is made available on the Internet (Vitorino et al., 2018). Tracking such data to curtail the sharing of pornography and prosecute perpetrators is a critical aspect of cyber-crime activities. Software capable of fast detection of pornographic content is essential to law enforcement agencies (LEAs) and such needs have driven researchers to propose algorithms which aid LEAs in their fight against cybercrime. Thus, tools that detect pornographic content (Wehrmann et al., 2018), perform age estimation (Macedo et al., 2018b), and search for specific keywords in file names, amongst others, exist to aid LEAs. In the analysis of pornographic image content, algorithms which determine whether images contain exposed private body parts provide LEAs with essential descriptors of the image content. Such descriptors provide a better understanding of the reason why an images is pornographic without the need for manual inspection.

Deep learning algorithms are proving to be very effective in object detection and classification. To our knowledge, however, deep learning algorithms which analyse images specifically to determine whether these contain private body parts do not exist. The position taken by this paper is that by training a deep network on windows containing private body parts, we can create a system which detects and labels these parts. To this extent, we propose a two-step approach, in which the first step is a pornographic image detector, and the second step uses a windowing approach to detect private body parts within the image.

The rest of the paper is organised as follows. Section 2 gives an overview of the related work in the area, Section 3 describes our proposed approach, Section 4 presents the results obtained, while Section 5 concludes the paper.

[a] https://orcid.org/0000-0002-7144-5221
[b] https://orcid.org/0000-0002-6580-3424
[c] https://orcid.org/0000-0003-4617-7998
[d] https://orcid.org/0000-0001-8106-9891
[e] https://orcid.org/0000-0003-0436-6408

2 LITERATURE REVIEW

Human skin offers a practical feature for pornography detection since it is invariant to partial occlusion, as well as to changes in scaling and rotation (Naji et al., 2018). However, skin detection is an insufficient indicator of pornography since instances of benign images may also have large areas of exposed skin (Wang et al., 2009). Thus, skin detection often acts as a precursor to sexual organ detection, particularly, of the female breast (Fuangkhon and Tanprasert, 2005). In such cases, skin detection is followed by either model-based detection which uses ge-

Tabone, A., Bonnici, A., Cristina, S., Farrugia, R. and Camilleri, K.
Private Body Part Detection using Deep Learning.
DOI: 10.5220/0009101502050211
In *Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2020)*, pages 205-211
ISBN: 978-989-758-397-1; ISSN: 2184-4313

Table 1: Method and performance comparison of algorithms in previous literature.

Method	Aim	Approach	Performance
(Shen et al., 2010)	Sexual organ detection	Human pose model	True positive rate: 89%, False positive rate: 10%
(Choi et al., 2011)	Pornographic classification	Skin detection + Colour and texture features + MAP	True positive rate: 93.63%, False positive rate: 10.13% False negative rate: 6.9%
(Tian et al., 2018)	Sexual organ detection	Deformable Part Model + SVM	Overall precision: 80%, Recall: 82%, F-1 score: 18%
(Lv et al., 2011)	Pornographic classification	Semantic Tree model + X^2-kernel SVM	Accuracy: 87.6%
(Huang and Kong, 2016)	Pornography classification	Colour features + CNN	Accuracy: 78.29%
(Sigal et al., 2004)	Real-time skin segmentation	Bayes'	Detection rate: 86.84%
(Hajraoui and Sabri, 2014)	Face detection	Thresholding + Watershed	Detection rate 97.27%
(Taqa and Jalab, 2010)	Skin detection	ANN	Detection rate: 95.62%
(Kim et al., 2017)	Skin detection	CNN	Accuracy: 95.62%, Precision: 87.2%, Recall: 91.22% F-measure: 89.19%
(Zuo et al., 2017)	Skin detection	RNN + FCN	AUC (COMPAQ dataset): 95.93%, AUC (ECU dataset): 98.10% 1-EER (COMPAQ dataset): 90.18%, 1-EER (ECU dataset): 94.80%

[1] AUC(%): Area Under Curve
[2] ROC(%): Receiver Operating Characteristic [3] 1-EER(%):Equal Error Rate [4] MAP: Maximum a Posteriori [5] SVM: State Vector Machine [6] COMPAQ: (Bhoyar, 2010)

ometric models to describe the structure of the human body (Forsyth and Fleck, 1999), or region-based detection, which extracts local features for recognition (Hu et al., 2007).

The detection of sexual organs does not, however, require skin detection as a pre-processing step. For example, in (Wang et al., 2010), the shape and colour of the nipple are used to train the Viola-Jones algorithm with an AdaBoost classifier to detect potential nipple regions. The feature set used by Wang et al. in the AdaBoost classifier may be extended to include Haar-like features such as edge, line and centre-surround features (Lienhart and Maydt, 2002), and colour, texture and shape features obtained through colour moments, histogram of oriented gradients (HOG) and grey-level co-occurrence matrices (GLCM) (Kejun et al., 2012).

More elaborate human body models are required when taking into account other sexual organs. Here, in addition to the recognition difficulties introduced through pose, the localisation and classification problem should also consider that the genitalia may also be exposed in coitus (Lv et al., 2011). For example, (Shen et al., 2010) use a rudimentary human pose model based on the skin distribution using the location of the face and trunk to aid the classification of the nipple and pube regions.

Table 1 compares the performance of feature-based sexual organ detectors found in literature.

In light of the remarkable results achieved by deep learning architectures for various computer vision tasks, recent advances in pornography detection are also using deep learning approaches (Moustafa, 2015; Zuo et al., 2017), using convolutional neural networks (CNN) to learn pornographic features from image examples (Moustafa, 2015; Zuo et al., 2017). Such an approach has the advantage of learning the common global traits in pornographic images. However, these approaches do not specifically search for body parts in the image.

Moustafa describes one of the first deep-learning-based approaches to pornographic image content. Pre-trained AlexNet and GoogLeNet architectures are repurposed for pornographic image detection and denoted as ANet and GNet respectively. By modifying the third fully-connected layer as a two-way Softmax and training on the pornographic dataset (NPDI) these provide the probabilities that the input image is pornographic or benign. By considering the two architectures as providing complimentary class probability, Moustafa then adds two new architectures denoted as AGNet and AGbNet which fuse the results from ANet and GNet to improve the final classification. In AGNet, the fusion score is the average of the two probabilities while in AGbNet, the fusion score is the largest of the two. Using a receiver operating characteristic curve (ROC) curve Moustafa demonstrates that fusion improves the classification performance, with AGbNet performing slightly better than the AGNet (Moustafa, 2015). The use of two convolutional neural networks, however, comes at the cost of a large number of network parameters that need training, which in turn, requires a large image dataset to ensure proper training.

In (Huang and Ren, 2018), the controllability over the features learnt by the CNN is increased by introducing a colour feature histogram as an input to the CNN. This approach achieved an accuracy of 99.31% on the NPDI data set, outperforming a vanilla CNN by 2.67%. An additional contribution of this work is the integration of *bagging* into the CNN to solve the over-fitting problem and enhance generalisation (Huang and Ren, 2018). In (Vitorino et al., 2018), the over-fitting problem is addressed by adopting transfer learning approach, training the GoogLeNet CNN network first on a large dataset of non-pornographic content and then fine-tuning the network with pornographic content (Vitorino et al., 2018). Table 2 provides a comparison of deep learning approaches for pornographic image detection.

Many of the false-positives in pornography classification of these deep learning based approaches

Table 2: Comparison of Deep Learning approaches to pornographic image detection.

Method	Aim	Approach	Performance
(Zuo et al., 2017)	Skin detection	RNN	98.1% AUC 94.8% 1-EER
(Moustafa, 2015)	Pornographic vs Non pornographic classification	AlexNet & GoogLeNet	$94.2 \pm 2\%$ accuracy
(Huang and Ren, 2018)	Pornographic vs Non pornographic classification	colour feature histogram & GoogLeNet CNN	97.51% accuracy on NPDI dataset
(Vitorino et al., 2018)	Detect sexually exploitative imagery of children & adults from innocuous images	colour feature histogram & GoogLeNet CNN	91.5 % accuracy on Pornograpy-2K dataset 86.5% accuracy on a SEIC dataset
(Wang et al., 2016)	Detect exposed body parts	CNN with MIL	98.4% accuracy on NPDI dataset
(Jin et al., 2018)	region-based recognition of sexual organs	GoogLeNet	97.5% accuracy on NPDI[1] dataset
(Macedo et al., 2018b)	age estimation & pornography detection	ResNet-50 architecture & VGG-16 for age estimation	79.84% accuracy on NSFW[2] & RCPD datasets
(Perez et al., 2017)	video pornography detection	CNN GoogLeNet architecture fusing MPEG and still image features	96.3% accuracy on Porn-2k dataset

[1] MPEG: Moving Picture Experts Group [2] RCPD: Region-bsed annotated child pornography dataset (Macedo et al., 2018a)

stems from the lack of training on private body parts (Wehrmann et al., 2018). This issue can be resolved by either designing separate classifiers for different body parts (Wehrmann et al., 2018) or by using a generic pornographic content detector trained to detect the private body parts (Wang et al., 2016; Jin et al., 2018). In (Wang et al., 2016), a CNN is trained using multiple instance learning (MIL), thus, using a generic detector which nevertheless recognises an image as pornographic if it contains at least one exposed private body part. The results are compared with traditional methods for pornography detection, based on image retrieval and bag-of-features techniques, as well as with variants of the proposed method, by either training the CNN on entire images rather than multiple instances, or training the CNN on images of the breast and genitalia separately. Wang et al. report that the generic detector trained on multiple image instances of breasts and genitalia achieved better pornographic/benign image classification.

From this review, we note that although the approaches described in (Wang et al., 2016) and (Jin et al., 2018) use image instances with private body parts to recognise images as pornographic/benign, there is a lack of deep learning approaches which locate and assign a label to the different body parts. However, methods that label instances of body parts within the image would be of benefit to help law enforcement agencies obtain a better understanding of the content of the image and the reason why the image is considered pornographic. We address this research gap in our proposed approach.

3 METHODOLOGY

The aim of this work is the identification and labelling of private body parts, namely, female breasts, and female and male sexual organs. To achieve this goal, we propose a two-step process in which the first step classifies the image as pornographic/benign while the second step performs a more in-depth analysis on only those images identified as pornographic, locating and labelling instances of private body parts. Such a two-step process allows us to employ the multi-class classifier to only those images that require further analysis. This approach is necessary considering the speed requirement for law enforcement agencies. Our pornographic image classifier consists of a CNN, which takes as its input the entire image. To detect and label private body parts within the pornographic image we adopt a windowing approach, whereby each image is divided into smaller windows, checking each of these windows with a multi-class CNN.

3.1 Pornography Classification

The first step of our approach entails the classification of an image as pornographic/benign, thus, a binary classifier is used. Keeping in mind the speed and efficiency requirements, the lightweight MobileNet architecture was adopted as it uses depth-wise separable convolutions to a build fast, light-weight deep neural network (Howard et al., 2017). Our evaluations demonstrate that the accuracy reached by this model is comparable to that achieved by other models while having a faster computation time.

We adopt a transfer learning approach on a pretrained MobileNet by replacing its classification layer by a custom classifier consisting of a Dense ReLU layer followed by a Softmax output layer. Fine tuning was used to find the right balance between using the pretrained feature extractors and the learning of new features.

To train this model, we use a subset of the large-scale pornographic image corpus created by the University of Leon[3] (UL). This database was generated by crawling around two million images from five popular pornography websites. Since the data set has 1656 benign and 16033 pornographic images, training using this data introduces a bias. Thus, additional benign images were introduced from the VOC2012[4] dataset.

3.2 Private Body Part Classification

Once an image is considered to be of a pornographic nature, it is tessellated into non-overlapping 128×128 windows, passing each of these windows to our classifier to identify the location, if any, of four different body parts, namely *buttocks* (Bt), *female breasts* (FBr), *female genitalia* (FG), *male genitalia* (MG). We note that the appearance of female genitalia can differ considerably between instances of posing and instances of sexual activity. Thus, we redefine these body parts as *female genitalia posing* (FGP) and *female genitalia active* (FGA). This helps the model learn each class better by reducing the variation in observations of each object. Moreover, we note that some pornographic images contain sex toys. While these may appear similar to genitals, they should not be labelled as such. Thus, we introduce a sixth class which corresponds to *sex toys* (ST). A seventh, *benign* (Bn) class is required to describe image content which is not a private body part. Thus, body-part classification requires a multi-class classifier.

To determine the architecture that best suits our application, multiple pretrained models were trained for this classifier, training all classifiers with image windows of 128×128 pixels. We note that during the fine tuning stage, the best performance was achieved when none of layers were excluded from training. This may be due to the fact that we are working with windows rather than complete images.

To train the multi-class classifier, points of interest on the pornographic images from the UL dataset were manually labelled. Five 128×128 windows were taken around each labelled point, one centred on the labelled point and four perpendicularly offset. The *benign* class training set was generated separately by extracting 128×128 windows from random points on benign images. The seven classes and the corresponding number of samples gathered are shown in Table 3. From this, we note that some classes have very few examples. Thus, data augmentation in the form of image scaling with a scaling factor of up to 0.2, image rotations with rotation angles in the range of $[0, \frac{\pi}{2}]$ and horizontal flips were used to boost the size of this dataset.

Table 3: Window dataset classes and their size.

Class Name	Number of windows
Benign (Bn)	9895
Buttocks (Bt)	6490
Female breasts (FB)	19258
Female genitalia posing (FGP)	4865
Female genitalia active (FGA)	7460
Male genitalia (MG)	2395
Sex toys (ST)	1635

4 RESULTS AND EVALUATION

The results were generated on a Windows computer running a 64-bit operating system, and featuring an Intel®Core$^{\text{TM}}$i7-8750H CPU at 2.20GHz, 16Gb of DDR4 memory and a GeForce GTX 1070 graphics card. Python was used as the development language, and Tensorflow and Keras[5] as open source packages to build the networks. Similar to (Vitorino et al., 2018) and others, we evaluate our classifiers by using the sensitivity, specificity, F1-score and accuracy metrics.

The evaluation of the first-stage pornographic image classifier, we use a test set of 50 pornographic and 50 benign images. The first row of Table 4 shows the parameters used to train the MobileNet classifier which takes 5.178s to load and 0.017s to classify an image. For the test set used, we obtained a sensitivity of 0.95 and a specificity of 0.95 which are satisfactory for the purpose of our application.

The models used to test the performance of the private body part classifier are shown in Table 4 along the parameters of the best version obtained for each. To compare the performance of these models, we manually selected windows centred around private body parts from 50 pornographic images. The F1 score of each was calculated. Since windows might contain more than one sexual organ, both top-1 and top-3 accuracy were considered along the computation and loading time. These values are shown in Table 5 and were used to select the most suitable classifier for our intended application.

[3]http://gvis.unileon.es/dataset/apd-2m/

[4]https://bit.ly/33Jd2rp

[5]https://keras.io/

Table 4: Training Parameters for all the Classifiers used.

Training Parameters					Classifier Used						
Model Architecture	Batch Size	Optimiser	Epochs	Validation Accuracy	Drop	Dense 'ReLU'	Drop	Dense 'ReLU'	Drop	Dense 'ReLU'	Dense 'SoftMax'
MobileNet	50	RmsProp lr=4e-7	25	0.9417	/	/	/	/	0.5	128	2
MobileNet	200	RmsProp lr=2e-5	50	0.7350	/	/	0.4	128	0.4	128	7
VGG-16	300	RmsProp lr=1e-5	30	0.6860	/	/	0.5	128	0.5	128	7
ResNet	150	RmsProp lr=5e-5	40	0.7663	0.5	128	0.7	128	0.7	128	7
Inception V3	300	RmsProp lr=4e-5	25	0.7529	/	128	0.4	128	0.4	64	7

Table 5: Model Accuracy and Computation time.

Model	Accuracy Top-1	Accuracy Top-3	Computation time (ms)	F1 score
ResNet	0.8011	0.9569	9.824	0.7258
Inception-v3	0.8010	0.9731	13.08	0.8011
MobileNet	0.7527	0.9731	7.07	0.7258
VGG-16	0.7097	0.9570	13.02	0.7097

Proving a suspect guilty may entail the processing of large amounts of data in a limited time frame, hence the speed at which each image is checked is of utmost importance. Thus, the MobileNet model was chosen since it performs faster compared to other models and suffers only from a slight tradeoff in accuracy. To better visualise this model's performance, a confusion matrix is generated and shown in Figure 2. The manually selected windows set is made up of labelled sexual organs and hence has no *benign* class. The majority of misclassifications are caused by the *male genitalia* class which may be attributed to the limited number of examples used during training in combination with the resulting relatively large amplification of its loss when balancing the weights of each class.

The whole pipeline is then tested by tessellating the images identified as pornographic by the pornographic classifier into windows. The confusion matrix in Figure 2, gives the overall performance of the private-body part classification. A seven-class, pure-chance classifier would have a detection accuracy of 1/7. The accuracy achieved by our classifier stands at 0.62 which is four times better than a random selector. Nevertheless, we note that there is a decline in the overall detection accuracy when using the full pipeline. This cutback may be attributed to a misclassification of benign windows with large skin patches as private body parts and the misclassification of windows containing body parts not centred in the window. Moreover, the imbalance between classes during training was attempted to be catered for by weighting the error of each class, though this does not compare to the improvements that can be made by using more examples of body parts which differ in pose and position within the window.

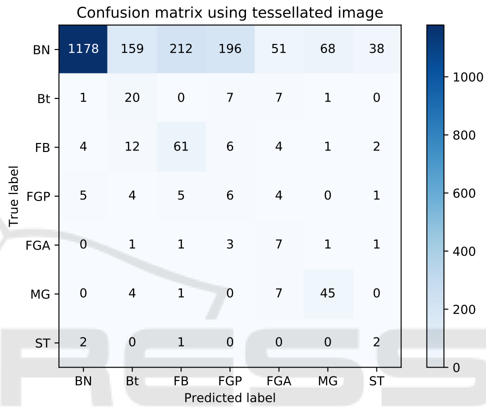

Figure 1: Confusion matrix using whole pipeline.

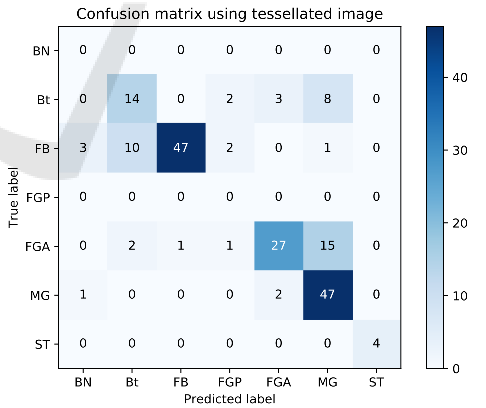

Figure 2: Confusion matrix with manual windowing.

5 CONCLUSIONS

In this paper we review various classifiers that use both hand-crafted features and deep learning techniques for classifying images as pornographic/benign. However, no known deep learning approaches were found to label sexual organs inside these images. Hence, we propose a two-step approach, which uses a first classifier to classify an image as pornographic or benign, then tessellates pornographic images into windows. A second classifier checks these windows for sexual organs and augments the initial classification of the pornographic image by describing the explicit content in that image.

Our results show that the MobileNet architecture has similar classification performance as the other networks tested but reduces the computation time by at least 20%. Thus, it was considered suitable for our system and achieved an accuracy of 95% in the first step of our pipeline. This result implies that the pornographic image classifier correctly distinguishes between benign and pornographic images.

The second step of our pipeline reached an accuracy of 72.58% for windows centred around the private body parts. This demonstrates that deep learning approaches can be used to detect and label private body parts in pornographic images. When evaluating the full pipeline, an accuracy of 62% was achieved. While this is larger than a pure chance classifier, we note that this classifier can be further improved by having more training data and introducing smarter window selection schemes. We propose that window selection can be achieved by extracting a feature map from the first pornography classifier and using this map to localise regions of interest presumably containing private body parts.

ACKNOWLEDGEMENTS

This research has been funded with support from the European Commission under the 4NSEEK project with Grant Agreement 821966. This publication reflects the views only of the authors, and the European Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

Bhoyar (2010). Skin color detection model using neural networks and its performance evaluation. *Journal of Computer Science*, 6(9):963–968.

Choi, B., Han, S., Chung, B., and Ryou, J. (2011). Human body parts candidate segmentation using Laws texture energy measures with skin color. In *13th International Conference on Advanced Communication Technology (ICACT2011)*, pages 556–560.

Forsyth, D. A. and Fleck, M. M. (1999). Automatic detection of human nudes. *International Journal of Computer Vision*, 32(1):63–77.

Fuangkhon, P. and Tanprasert, T. (2005). Nipple detection for obscene pictures. In *Proceedings of the 5th WSEAS International Conference on Signal, Speech and Image Processing*, SSIP'05, pages 242–247.

Hajraoui, A. and Sabri, M. (2014). Face detection algorithm based on skin detection, watershed method and gabor filters. *International Journal of Computer Applications*, 94:33–39.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., and Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications.

Hu, W., Wu, O., Chen, Z., Fu, Z., and Maybank, S. (2007). Recognition of pornographic web pages by classifying texts and images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:1019–1034.

Huang, L. and Ren, X. (2018). Erotic image recognition method of bagging integrated convolutional neural network. In *Proceedings of the 2nd International Conference on Computer Science and Application Engineering*, CSAE '18, pages 107:1–107:7, New York, NY, USA. ACM.

Huang, Y. and Kong, A. W. K. (2016). Using a CNN ensemble for detecting pornographic and upskirt images. In *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–7.

Jin, X., Wang, Y., and Tan, X. (2018). Pornographic image recognition via weighted multiple instance learning. *IEEE Transactions on Cybernetics*, pages 1–9.

Kejun, X., Jian, W., Pengyu, N., and Jie, H. (2012). Automatic nipple detection using cascaded AdaBoost classifier. In *2012 Fifth International Symposium on Computational Intelligence and Design*, volume 2, pages 427–432.

Kim, Y., Hwang, I., and Cho, N. I. (2017). Convolutional neural networks and training strategies for skin detection. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3919–3923.

Lienhart, R. and Maydt, J. (2002). An extended set of haar-like features for rapid object detection. In *Proceedings. International Conference on Image Processing*, volume 1, pages I–I. IEEE.

Lv, L., Zhao, C., Lv, H., Shang, J., Yang, Y., and Wang, J. (2011). Pornographic images detection using high-level semantic features. In *2011 Seventh International Conference on Natural Computation*, volume 2, pages 1015–1018.

Macedo, J., Costa, F., and dos Santos, J. A. (2018a). A benchmark methodology for child pornography detection. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. IEEE.

Macedo, J., de Oliveira Costa, F., and dos Santos, J. A. (2018b). A benchmark methodology for child pornography detection. *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 455–462.

Moustafa, M. N. (2015). Applying deep learning to classify pornographic images and videos. In *7th Pacific-Rim Symposium on Image and Video Technology (PSIVT)*.

Naji, S., Jalab, H. A., and Kareem, S. A. (2018). A survey on skin detection in colored images. *Artificial Intelligence Review*, pages 1–47.

Perez, M., Avila, S., Moreira, D., Moraes, D., Testoni, V., Valle, E., Goldenstein, S., and Rocha, A. (2017). Video pornography detection through deep learning techniques and motion information. *Neurocomputing*, 230:279 – 293.

Shen, X., Wei, W., and Qian, Q. (2010). A pornographic image filtering model based on erotic part. In *2010 3rd International Congress on Image and Signal Processing*, volume 5, pages 2473–2477.

Sigal, L., Sclaroff, S., and Athitsos, V. (2004). Skin color-based video segmentation under time-varying illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(7):862–877.

Taqa, A. Y. and Jalab, H. A. (2010). Increasing the reliability of skin detectors. *Scientific Research and Essays*, 5(17):2480–2490.

Tian, C., Zhang, X., Wei, W., and Gao, X. (2018). Color pornographic image detection based on color-saliency preserved mixture deformable part model. *Multimedia Tools and Applications*, 77(6):6629–6645.

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

Wang, X., Hu, C., and Yao, S. (2009). An adult image recognizing algorithm based on naked body detection. In *2009 ISECS International Colloquium on Computing, Communication, Control, and Management*, volume 4, pages 197–200.

Wang, Y., Jin, X., and Tan, X. (2016). Pornographic image recognition by strongly-supervised deep multiple instance learning. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 4418–4422.

Wang, Y., Li, J., Wang, H. L., and Hou, Z. (2010). Automatic nipple detection using shape and statistical skin color information. In *MMM*.

Wehrmann, J., Simões, G. S., Barros, R. C., and Cavalcante, V. F. (2018). Adult content detection in videos with convolutional and recurrent neural networks. *Neurocomputing*, 272:432–438.

Zuo, H., Fan, H., Blasch, E., and Ling, H. (2017). Combining convolutional and recurrent neural networks for human skin detection. *IEEE Signal Processing Letters*, 24(3):289–293.