

# Fast pornographic image recognition using compact holistic features and multi-layer neural network



I Gede Pasek Suta Wijaya <sup>a,1,\*</sup>, Ida Bagus Ketut Widiartha <sup>a,2</sup>, Keiichi Uchimura <sup>b,3</sup>,  
Muhamad Syamsu Iqbal <sup>c,4</sup>, Ario Yudo Husodo <sup>a,5</sup>

<sup>a</sup>Informatics Engineering Dept., Mataram University, Mataram, Lombok, Indonesia

<sup>b</sup>Electrical Engineering and Computer Science Dept., Kumamoto University, Kumamoto Shi, Japan

<sup>c</sup>Electrical Engineering Dept., Mataram University, Mataram, Lombok, Indonesia

<sup>1</sup>[gpsutawijaya@unram.ac.id](mailto:gpsutawijaya@unram.ac.id); <sup>2</sup>[widi@unram.ac.id](mailto:widi@unram.ac.id); <sup>3</sup>[uchimura@cs.kumamoto-u.ac.id](mailto:uchimura@cs.kumamoto-u.ac.id); <sup>4</sup>[msiqbal@unram.ac.id](mailto:msiqbal@unram.ac.id); <sup>5</sup>[ario@unram.ac.id](mailto:ario@unram.ac.id)

\* corresponding author

## ARTICLE INFO

### Article history

Received August 31, 2018

Revised May 20, 2019

Accepted May 20, 2019

Available online June 20, 2019

### Keywords

Pornographic image

Recognition system

Neural network

Holistic features

Frequency analysis

## ABSTRACT

The paper presents an alternative fast pornographic image recognition using compact holistic features and multi-layer neural network (MNN). The compact holistic features of pornographic images, which are invariant features against pose and scale, is extracted by shape and frequency analysis on pornographic images under skin region of interests (ROIs). The main objective of this work is to design pornographic recognition scheme which not only can improve performances of existing methods (i.e., methods based on skin probability, scale invariant feature transform, eigenporn, and Multilayer-Perceptron and Neuro-Fuzzy (MP-NF)) but also can works fast for recognition. The experimental outcome display that our proposed system can improve 0.3% of accuracy and reduce 6.60% the false negative rate (FNR) of the best existing method (skin probability and eigenporn on YCbCr, SEP), respectively. Additionally, our proposed method also provides almost similar robust performances to the MP-NF on large size dataset. However, our proposed method needs short recognition time by about 0.021 seconds per image for both tested datasets.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



## 1. Introduction

The spreading of pornographic contents (pictures, cartoon, animation, video, and text) is a big problem for a country like Indonesia due to their many adverse effects, particularly for kids and youths. A survey report in 2005 showed that 8% of emails, 12% of homepages browsing, 25% of the queries in search engines related to pornographic contents [1]. Concerning to the adverse effects of pornographic contents, the most one is an addiction on accessing pornographic contents. It means the addict always be flying or dependent on pornographic contents. Additionally, the early pregnancy or sex deviation behavior also can be triggered by pornographic contents addiction, especially on kids and teenagers who do not understand enough of the undesirable consequences of pornographic contents.

Consequently, a blocking scheme which can reject of browsing, accessing, and making the pornographic contents is required to minimize their adverse effects. In order to build a reliable blocking system of pornographic contents, a proper recognition or classification algorithm is needed. In this case, the function of the recognition or classification algorithm is to determine the correlation between the training set features and input features. If the recognition or classification algorithm returns a high similarity score (more than a given threshold score), the input features are concluded (recognized) as pornographic content then it will be blocked to be seen, saved, and shared. In order to contribute to the

rejection of pornographic contents, an alternative pornographic image recognition using compact holistic features and the neural network is presented. The compact holistic features which are invariant features against pose and scale is extracted by shape and frequency analysis on pornographic images under skin region of interests (ROIs). The skin ROI has been confirmed that it can handle the large pornographic images variability due to background variations [2], [3]. The classification based on a neural network is employed to overcome the recognition time. The main aim of the paper is to design new scenario of pornographic recognition using compact holistic features representing the shape and dominant skin information which can improve the performance of existing methods (i.e., methods based on skin probability, eigenporn, Multilayer-Perceptron and Neuro-Fuzzy, etc.). In this research, the performance indicators for examining our proposed methods are accuracy, false negative rate (FNR), false positive rate (FPR), and computational time.

Regarding to the previous works, there are three main groups of adult/pornographic image recognition scheme [4] namely, the approach based on color, shape, and local descriptor. While other researches [1] also have grouped approaches to pornographic filtering into three major classes: 1) based on text contents, 2) based on collection lists of adult website addresses that is blocked by internet firewall, and 3) based on image content analysis [1], [2], [5]–[7]. Regarding texts content-based method, it classifies the material to be pornographic using the probability/entropy of texts related to pornographic contents that are available in the material (i.e., websites). However, this method fails to block material having many pornographic images and videos. Next, the website URL based method rejects the accessing the adult image using internet firewall such as squid that has the rule to block the website addresses (URLs) list belonging to pornographic contents. However, adult website grows quickly, and the owner can easily rename the website address. Finally, the method based on image analysis performs the rejection of accessing website based on the images or videos that exist in the media. It can solve the weakness of the two above methods. As mentioned early, the last type of filtering method faces many obstacles because of the huge image variability due to the skin, pose, lighting, and background data.

Additionally, detection of pornographic images algorithms also can be categorized as based on region, contour [8], human skin probability [2], [6], [9], scalable color, edge histogram, and shape descriptors [10]. Both region and contour of the pornographic image were obtained by using skin color. Those schemes were recommended to solve the difficulties on recognition of the pornographic image. However, they still hold less achievement, especially high false negative and positive data due to the huge changeability of the input images. Commonly, the skin region scheme which was based on threshold model in YCbCr, RGB, and HSV color space [2], [9] and Gaussian mixture models [8] was employed for skin segmentation. Concerning to procedures of feature extraction, they can be classified into a local (eyes, nose, and mouth, genital), holistic, and shape feature extraction techniques. All of the procedures of feature extraction were commonly applied because they can work quickly and properly. The example approach of holistic feature extraction is content-based feature extraction using frequency analysis (FFT [11], DCT [12], and Wavelet [13], [14]), feature point descriptor using scale-invariant feature transforms (SIFT [15]–[17], eigenporn of HSV-ROI feature extraction [18], and descriptors of color, edge, and shape of pornographic images [10]. The freshest strategy for pornographic recognition was a modification of Convolutional Neural Network (CNN) called Deep Multicontext Network (DMN) [19]. In DMN frameworks, a deep CNN is applied to model blending features of sensible objects in images. It seems the DMN algorithm requires a complicated process which impacts to the computational cost.

The diversity of pornographic recognition schemes show that the pornographic image recognition is an overwhelming task due to the huge variability of the images. It also means the pornographic image recognition still challenges the research topic. Furthermore, those approaches did not consider the computational time of the recognition process yet. Therefore, an alternative solution to pornographic recognition problem using compact holistic features representing shape and dominant skin information and multi-layer neural network (MNN) is proposed to improve the established pornographic recognition methods, which is potentially implemented for mobile devices. In order to confirm the robustness of the proposed method, the experimental data will be compared to those of the well-known schemes of pornographic image recognition methods (i.e skin probability, skin region [2], [3], [8], [9], eigenporn

of HSV ROI [18], fusion descriptor of YCbCr ROI(FD) [20], and skin probability and eigenporn on YCbCr (SEP) [17] and Multilayer-Perceptron and Neuro-Fuzzy (MP+NF) [1]).

## 2. Method

The diagram block of our proposed pornographic recognition is presented in Fig. 1. The main concern of this research is to design compact holistic features which consist of shape and dominant skin information of pornographic images. The compact holistic features are extracted by moment and Discrete Cosine Transforms on pornographic images under skin ROI. The classification process is performed using MNN. The main difference of this method to the eigenporn [17] and the MP+NF [1] is placed on the features and classification algorithms. The eigenporn based method implemented the eigenporn extracted by PCA [21], [22] and k-nearest neighbor [23], [24] for classification and MP+NF based method implemented the multi features and Multilayer-Perceptron and Neuro-Fuzzy for classification.

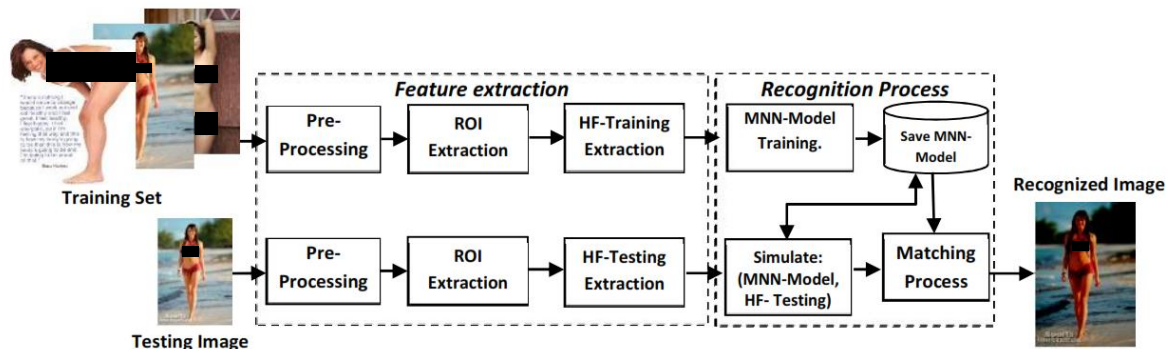


Fig. 1. Diagram block of the proposed pornographic image recognition

### 2.1. Features extraction

The compact holistic features consisting of shape and dominant skin information are presented by vectors representing the global information of pornographic or non-pornographic images. In this case, the compact holistic features, which are design to require limited memory space, are extracted from the chrominance components (Cb and Cr) of the input images. The intensity component (Y) is not included in the features because it is susceptible to lighting variations. Briefly, the features extraction starts from pre-processing, ROI extraction, and shape and frequency analysis.

### 2.2. Pre-Processing and ROI extraction

The pre-processing process relates to image resizing and lighting normalization. The input image is scaled into size 256 pixels by keeping the height and width ratio to decrease the computational time of the next process. In order to decrease the lighting effect on images, the histogram equalization is employed for normalization. The skin pixels of images that have much lighting effect fail to be classified as skin without lighting normalization. In other words, the histogram equalization is applied to handle large variability of lighting variations of the input images.

Next, the ROI extraction is started from pixels-based skin classification to obtain skin tone image. Some methods for pixel-based skin classification are presented by authors [1], [9], [25], [26]. Among the existing methods, the best performance was provided by pixels based skin classification on YCbCr color space [9], [17]. Therefore, pixels-based skin classification on YCbCr color space is employed for extracting skin tone of images as presented in Algorithm 1 (Fig. 2).

In this case, the Cb and Cr values are in the range of 16-240. These values are the output of the RGB to YCbCr transformation algorithm [27]. Additionally, the morphological also included in this process to remove false positive skin classification. This model has been proved to provide better performance for recognizing the pornographic images [17]. The output example of skin tone extraction algorithm is given in Fig. 3.

```

Algorithm 1 The process of skin tone extraction.
Input : image matrix (im)
Output : skinTone matrix
Process:
skinTone (im:rows; im:cols)
r ← 0; g←0; b← 0;
for i = 0; i < im:rows; i ++ do
  for j = 0; j < im:cols; j ++ do
    b ← im(l, j, 1); g ← im(l, j, 2); r ← im(l, j, 3);
    int cb ← int((-0.1481*r - 0.2908 *g + 0.4390 *b) +
128);
    int cr ← int((0.4391 * r - 0.3667* g - 0.0714*b) + 128);
    if (cb > 77&&cb < 127)&&(cr > 133&&cr < 173) then
      skinTone(l,j) ← 1;
    else
      skinTone(l, j) ← 0;
    end if
  end for
end for
return skinTone;

```

Fig. 2. The process of skin tone extraction



Fig. 3. The output example of skin tone extraction

From the skin tone image, the ROI is extracted by using vertical and horizontal projection [17]. After obtaining skin tone, the ROI extraction is started from performing the vertical (rows) and horizontal (column) integral projection to know the coordinates having large of skin and non-skin region using. Secondly, the vertical and horizontal projection probability having less than a defined threshold is removed. In this case, by trial and error, the best threshold can be set as 0.25 of maximum vertical and horizontal projection probability. Thirdly, the skin tone is cropped using the x and y coordinates where the vertical and horizontal projection probability is thresholded. Finally, the cropped skin tone is mapped to the original image to get the skin ROI image. In this case, the skin ROI itself is implemented to eliminate non-skin information of pornographic images and to minimize huge pornographic images variability because of variations of background.

### 2.3. Shape and Frequency Analysis

Frequency analysis has been implemented to extract holistic features of an image such as a Fast Fourier Transform (FFT) [11], [28], Discrete Cosine Transforms (DCT) [1], [16], [17], and Discrete Wavelet Transforms (DWT) [13], [14]. In this research, we propose a different scheme from the mentioned methods in terms of the combination of shape information and frequency features of the skin ROI images. The shape information is extracted using invariant moment algorithm to figure out the huge pornographic images variability due to alterations of pose and the DCT is implemented to get the holistic skin information of ROI images which are robust to rotation and scale variations. In this case, the DCT algorithm that is employed to extract predominant frequency contents of ROI image ( $I$ ) having size  $N, M$  is presented by the Eq. (1).

$$y = Re \left[ e^{-jm\pi/2M} F \left( Re \left[ e^{-jn\pi/2N} F(I) \right]_n \right) \right]_m \quad (1)$$

where  $n=1, 2, \dots, N; m=1, 2, \dots, M$ , and  $F$  is Fast Fourier transforms.

From the two-dimensional DCT transformation coefficients, holistic skin information is selected from Cb-Cr color space by two processes: convert the DCT transformation coefficients to one-dimensional vector using zigzag rules as implemented in jpeg compression and select first 30 elements from each Cb and Cr vectors. The shape information is just extracted from the intensity component of the image because the most information of shape is available in this component. Finally, from both shape and skin information, the compact holistic features (HF) is composed by placing them as a vector as presented in Algorithm 2 (Fig. 4).

**Algorithm 2** The process of holistic features extraction.  
**Input** : image (im) and skinTone matrix  
**Output** : holistic features (hF) vector  
**Process** :  
 sProb ← sum(skin)=(im.rows \* im.cols);  
**if** sProb ≥ 0.01 **then**  
 [Y; Cb; Cr] ← rgb2ycbcr(im);  
 //calculate Invariant Moment  
 iMom ← getInvMoment(Y \* skin);  
 //calculate DCT Frequency  
 hfCb ← dct(Cb \* skin);  
 hfCr ← dct(Cr \* skin);  
 //Select the most significant values  
 hfCb ← zigzag(hfCb);  
 hfCr ← zigzag(hfCr);  
 hF ← [iMom hfCb(1 : 30) hfCr(1 : 30)];  
**end if**  
**return** hF

Fig. 4. The process of holistic features extraction

When the compact holistic features are evaluated from 1000 pornographic and 1000 non-pornographic images, it shows good enough discriminant information as distributed in two-dimensional space Fig. 5. The Fig. 5 indicates that the features of non-pornographic and pornographic are separated from one to each other. It means the compact holistic features are potential to be implemented for recognizing the pornographic image.

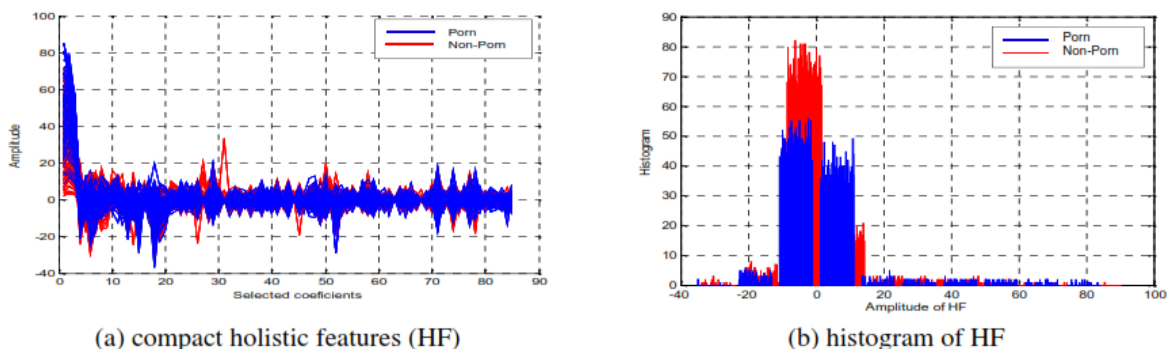


Fig. 5. The distribution of HF of 1000 pornographic and 1000 non-pornographic images

#### 2.4. Multi-Layers Neural Network (MNN) Model

There are many types of neural network that can be implemented for pattern recognition, which is distinguished by the architecture of neuron, kind of training, number of layers, etc. Generally, the MNN architecture is shown in Fig. 6 which consist of input vector ( $p$ ), bias ( $b$ ), weight matrices ( $W$ ), and transfer function ( $f$ ) [29], [30]. The weight matrices connecting to inputs vector is called as the input weight ( $IW$ ), while the weight connecting to outputs layer is called as the layer weights ( $LW$ ). In addition, superscripts for the various weights indicate the weight of the source/origin (second index) and the destination/target (first index). From Fig. 6, the output of the neural network is defined as (2).

$$a^3 = f^3(LW^{3,2} f^2(LW^{2,1} f^1(IW^{1,1} p + b^1) + b^2) + b^3) \quad (2)$$

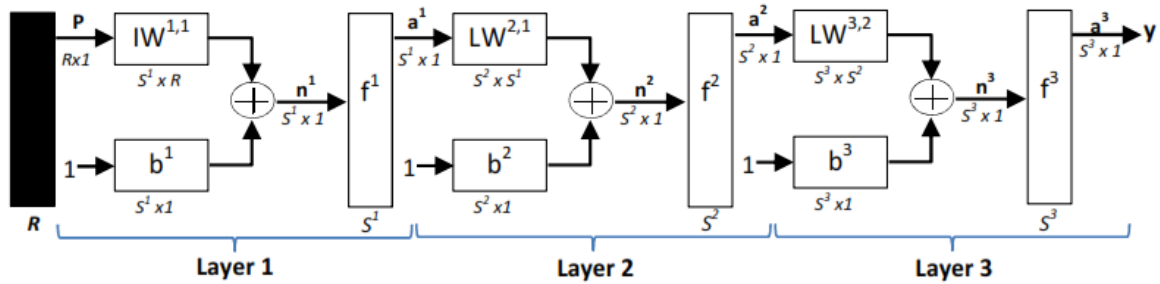


Fig. 6. The General multi-layer Neural Network Model

In this paper, an MNN model is employed for classification because it can work powerfully and quickly. For example, two layers neural network, where the first and the second layer are sigmoid and linear respectively, can be trained to estimate any function/task (with a limited number of discontinuities/disruption) adequately. However, the best variation of the layer and transfer function (how many layers and what the transfer functions are) has to be investigated by performing several experiments.

### 2.5. Recognition Process

Recognition process has two main stages: training and matching. The training needs a dataset consisting of both non-pornographic and pornographic images. The compact HF's of the input data is extracted by using Algorithm 2 (Fig. 4). Next, features selection are carried out to remove similar features of extracted compact HF's by intersection operation (3).

$$HF_{P,N} = (HF_P \cup HF_N) - (HF_P \cap HF_N) \quad (3)$$

Where  $HF_{P,N}$  is final trained HF,  $HF_N$  and  $HF_P$  are compact HF's of non-pornographic and pornographic images, respectively. From these sets, the global mean and standard deviation vectors of each  $HF_P$  and  $HF_N$  called as  $\mu_P$ ,  $\sigma_P$  and  $\mu_N$ ,  $\sigma_N$  are determined, respectively. Next, in order to obtain the minimum difference of the mean and standard deviation of both training set, the distance  $\mu_P$ ,  $\sigma_P$  and  $\mu_N$ ,  $\sigma_N$  are calculated. The HF's having smallest score are concluded as shared information which is removed for getting most discriminant information.

Next, the MNN model is trained using the  $HF_{P,N}$ , which is supervised by two targets vector. The first target for pornographic  $HF_P$  is [1 1 1 1 1 1 1] and for non-pornographic  $HF_N$  is [-1 -1 -1 -1 -1 -1 -1]. For instance, a two-layers neural network with linear, log-sigmoid, and tan-sigmoid transfer functions could achieve the goal of setting error when it was trained using HF (size 64 elements) and defined targets vector.

Finally, the classification is performed by simulating the query HF using the obtained MNN. From the simulation output, if the output vector is close to the first target vector, the query HF is concluded as a pornographic image and otherwise as non-pornographic image. This classification process is supposed to work very fast, which is the most benefit of the classification method because the query HF does not need to be compared to all trained HF vectors.

### 3. Results and Discussion

In order to know the achievement of the proposed method, some investigations were taken by using two datasets called UNRAM [17] and Kia datasets [1]. The UNRAM dataset consists of 687 pornographic and 712 non-pornographic images. While Kia dataset has 18354 images which 9295 and 9059 images are non-pornographic and pornographic, respectively. The images of both datasets were downloaded from the Internet using some downloader tools. The pornographic images of both datasets have large variability in terms of people, pose, skin. While non-pornographic images contain objects

which are similar to human skin, such as flower, wood, tiger, dessert, and etc. The data treatment for the experiments was performed as follows:

- For each dataset, 50% of non-pornographic and pornographic images were randomly taken as the training set, and their remaining were applied as testing.
- The accuracy, false negative rate (FNR), and false positive rate (FPR) parameters were used for performance indicators (using Eq. (4, 5, and 6)), and
- The evaluation was performed on a pc Intel Core i3-2370M, 2.4 GHz with 8 GB RAM.

The performance indicators were determined by the formulas that were derived from the confusion table [19], [31] (Table 1).

$$\text{Accuracy}(\%) = \frac{T_P + T_N}{N_P + N_N} \times 100 \quad (4)$$

$$\text{FNR}(\%) = \frac{F_N}{N_P} \times 100 \quad (5)$$

$$\text{FPR}(\%) = \frac{F_P}{N_N} \times 100 \quad (6)$$

where  $N_P$  is total of pornographic testing images, and  $N_N$  is total of non-pornographic testing images.

Table 1. Confusion value

		Actual Value	
		<i>Porn</i>	<i>Non-Porn</i>
Prediction Outcome	<i>Porn</i>	True Positive (TP): porn images that were correctly classified as porn	False Positive (FP): non-porn that were incorrectly labeled as porn
	<i>Non-Porn</i>	False Negative (FN): porn images that were incorrectly marked as non-porn	True Negative (TN): non-porn images that were correctly classified as non-porn

The first test was executed on UNRAM dataset to demonstrate that the proposed compact HF can be used to discriminate between non-pornographic and pornographic images. Furthermore, this experiment also investigated what size of HF was sufficient for pornographic image recognition. The matching process in the first experiment was performed by Euclidean distance, and the smallest distance deduces the best likeness. The test results show that the proposed compact HF, which consists of shape and dominant skin information, gives high enough accuracy, as shown in Fig. 7.

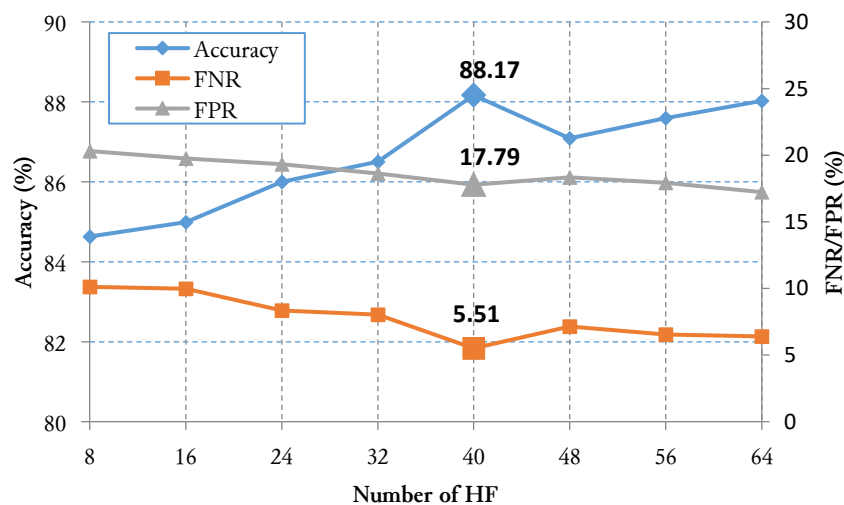


Fig. 7. The effect of HF size on accuracy of the proposed pornographic image recognition methods

These experimental results prove that the compact HF can be used to discriminate between the non-pornographic and pornographic images. It can be achieved because the compact HF has good enough discriminant information, as shown in Fig. 5. Additionally, the best HF size for performing pornographic recognition is 40 elements which are shown by the highest accuracy and small enough FPR and FNR (see Fig. 7). In detail, the best accuracy is by about 88.17% and the FNR and FPR by about 5.51% and 17.79% respectively. This experimental result also proves that the HF requires small memory space for representing the pornographic image, which implies to the computational cost of the recognition process. For further evaluation, the next experiments will be performed by the best size of HF (40 elements).

The second and third tests were conducted to find the best MNN model parameters (hidden layers and transfer functions) for compact HF classification. In these experiments, the variation of hidden layers and transfer functions were investigated to obtain their best combination of MNN model for pornographic image classification. The results of the second test denote that the best achievement is given by the MNN model having two hidden layers (2 HLs), which is indicated by the highest accuracy and the smallest FNR and FPR, as presented in Fig. 8. These achievements agree to the theory of MNN that it can estimate any function/task/problem (with a limited number of discontinuities/disruption) arbitrarily well. It means the MNN can provide crisp classification hyperplane for pornographic image recognition.

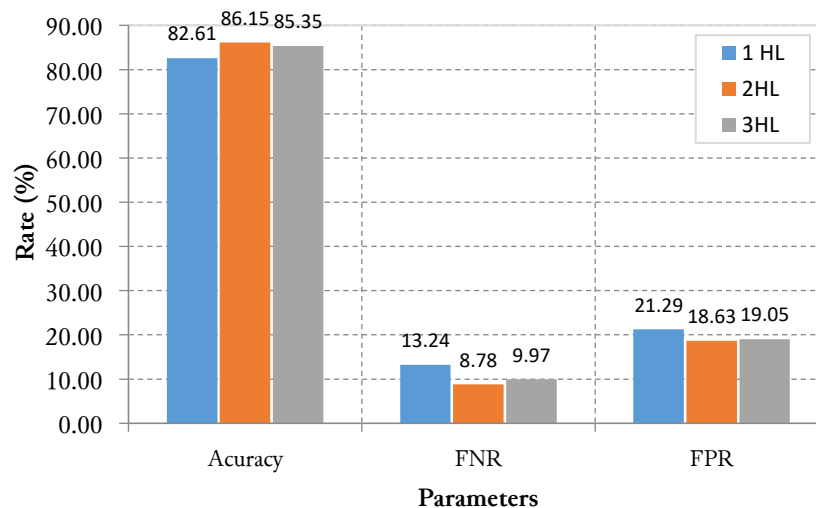


Fig. 8. The performance of pornographic recognition in some MNN models

Next, in order to discover the best variation of transfer functions for the MNN model having 2 HLs, the third experiment was performed using the same dataset as carried out in the second experiment. The transfers functions that were evaluated in this experiment were linear (L), log-sigmoid (S), and tan-sigmoid (T). The experimental results show that the best variation transfer functions are T, S, and L, as shown in Fig. 9. It means that the best parameters of MNN model for classifying the compact HF are two HLs and TSL transfer functions. Furthermore, the third experimental results also support the second experimental achievement in term of the powerfulness of MNN for classifying the compact HF of non-pornographic and pornographic images.

In order to analyze the achievement of the combination between compact HF and MNN (HF+MNN) for pornographic image recognition to the previous methods (skin probability (SP), skin region (SR), fusion descriptor (FD on YCbCr), eigenporn on HSV ROI (Ep on HSV), and SEP methods ([2], [3], [8], [9], [17], [18], [20])), the fourth experiment was done on UNRAM dataset using the best size HF and the best MNN model from the prior examinations. The test achievements present that HF+MNN gives greater accuracy and smaller FPR than those of established systems, as shown in Fig. 10. In detail, the HF+MNN method increases 0.30% of the accuracy and decreases 6.60% of the FNR. However, its FPR increases by about 5.90% of that of the best existing method (SEP [17]). Even though the FPR of



HF+MNN is higher than that of the SEP method, but the HF+MNN takes very short computational time (0.21 seconds) among the existing methods. Therefore, it can be concluded that HF+MNN outperforms among existing methods. Overall, these experimental results are in-line with all previous achievements which the compact HF of skin ROIs images can be implemented to discriminate the non-pornographic and pornographic images. This performance can be achieved because the compact HF consist most significant shape and dominant skin information of skin ROI images.

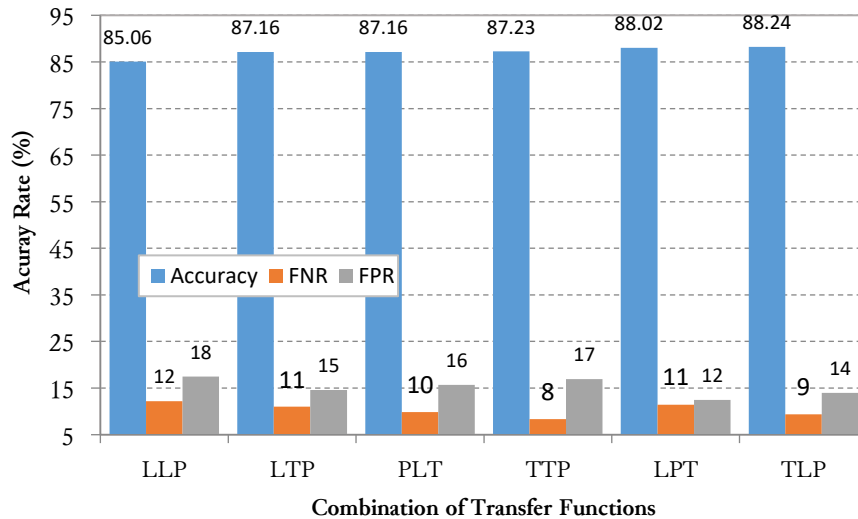


Fig. 9. The transfer functions variations versus recognition achievements

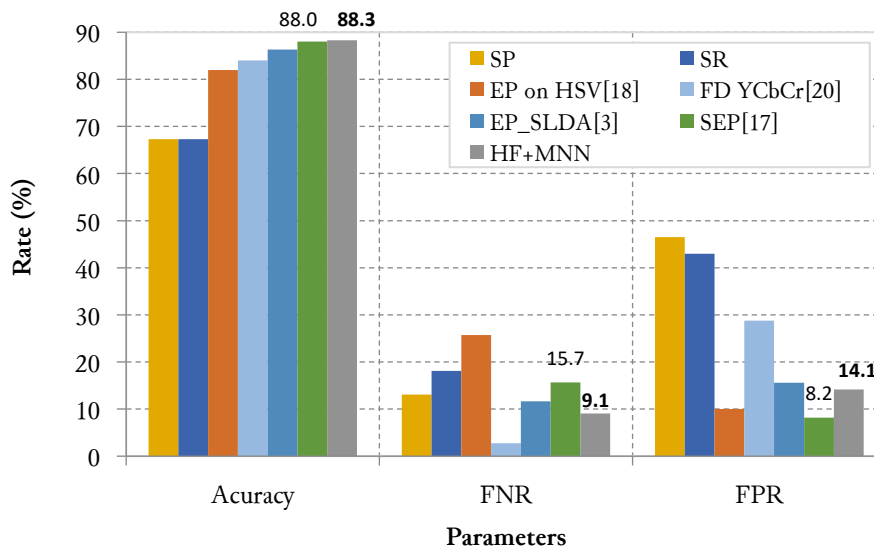


Fig. 10. The performance comparison to the some existing methods for UNRAM dataset

The next experiment was performed on the large size dataset (Kia dataset [1]) to know the robust performance of the HF+MNN method over large variability pornographic images. In this case, the HF+MNN method is compared to the latest existing method (MP+NF [1]). The experimental result shows that the HF+MNN provides similar performance to the latest existing method (almost 90% of accuracy, 10% and 7% of FNR and FPR, respectively), as presented in Fig. 11. It re-proves that the combination of compact HF and MNN gives good enough achievement for recognizing pornographic images. The variability of the large image due to skin-like and clothes variations such as images with skin-like background, dressing in transparent and mini clothes, partly porn, and dominant background, causes false classification. Additionally, false recognition also happened due to dressing in skin-like clothes. The Fig. 12, is the images examples which were miss classification by the proposed recognition algorithm.

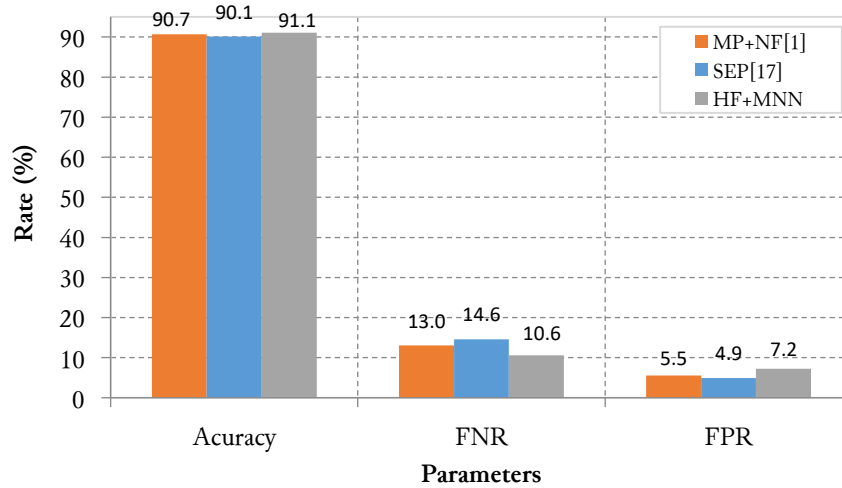


Fig. 11. The performance comparison to the recent methods for Kia dataset

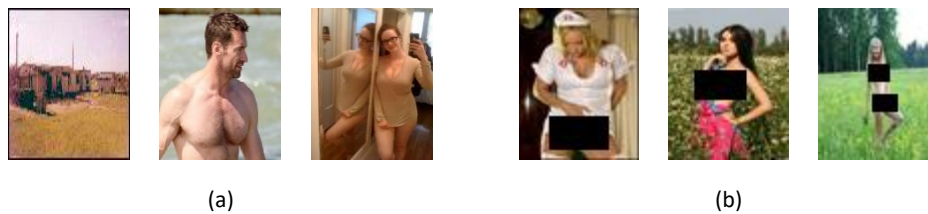


Fig. 12. Miss classification images examples: (a) FP and (b) FN

In order to know, whether the proposed recognition system can work fast, the last experiment was carried out. The experimental result shows that the HF+MNN method requires much shorter recognition time by about 0.021 seconds per image than that of existing methods, which is the fastest recognition process (Fig. 13). It can be achieved because the compact HF has a small size (40 elements) for each image. However, the weakness of HF+MNN is the training time by about 194.52 seconds for UNRAM dataset. Long training time is contributed by MNN, which is well known as the main weakness of the neural network. In practice, this problem can be handled by separating the training and recognition processes. From these achievements, HF+MNN method is potential to be implemented for pornographic rejection system for a mobile platform.

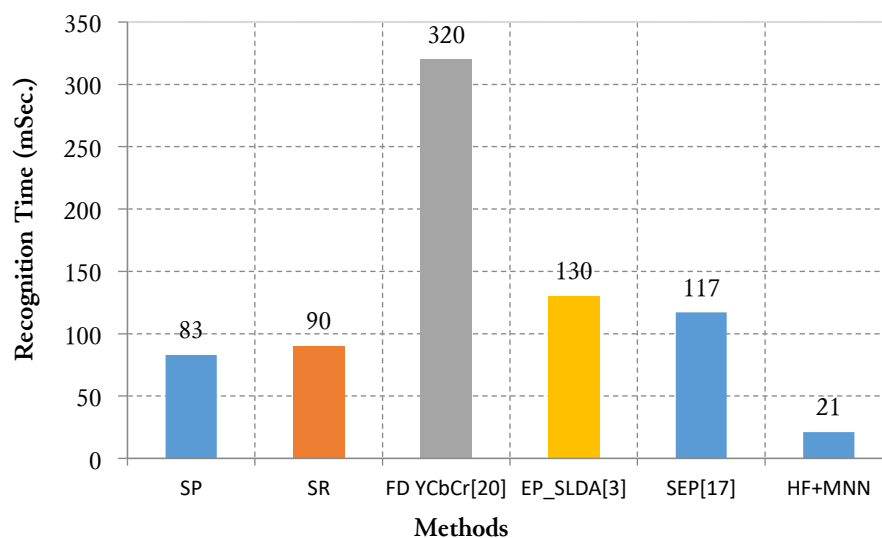


Fig. 13. The proposed method computational time compared to that of existing methods

#### 4. Conclusion

The proposed compact HF, which consists of shape and dominant skin information is powerful features for discriminating against pornographic images. The combination of compact HF and MNN (HF+MNN) has been demonstrated to provide excellent performances for pornographic image recognition. It indicates the proposed method is potential to be developed for real-time denial of pornographic images. The HF+MNN method also can improve the best existing method (SEP) that is shown by increasing its accuracy by about 0.30% and decreasing its FNR by about 6.6%, respectively. Furthermore, the HF+MNN method gives similar robust performance over large size dataset to the latest existing method (MP+NF) by about almost 90% of accuracy, 10% and 7% of FNR and FPR, respectively. However, our proposed method needs short recognition time by about 0.021 seconds per images. It means our proposed method can work very fast, which is potential to be applied for blocking the pornographic image in smart mobile media.

The results need to be improved by including the genital and edge information to improve the performances, especially for decreasing the FNR and FPR. Additionally, the proposed method must be tested by considering the context information of the image for the robustness evaluation. A simple deep learning based recognition system for pornographic image and video will be designed, which will be able to be implemented for mobile devices because it has reported providing the best performance.

#### Acknowledgment

This research is sponsored by The Ministry of Research, Technology, and Higher Education under the scheme applied research and contract No.: 801/UN18.L1/PP/2018. Moreover, we would like to express my great appreciation to the Image Media Laboratory of Kumamoto University, which provide hardware and software to realize this paper.

#### References

- [1] S. M. Kia, H. Rahmani, R. Mortezaei, M. E. Moghaddam, and A. Namazi, "A Novel Scheme for Intelligent Recognition of Pornographic Images," in *Computer Vision and Pattern Recognition, Cornell University Library*, 2015, available at: <https://arxiv.org/abs/1402.5792v3>.
- [2] J. A. Marcial-Basilio, G. Aguilar-Torres, G. Sánchez-Pérez, L. K. Toscano-Medina, and H. M. Pérez-Meana, "Detection of Pornographic Digital Images," *Int. J. Comput.*, vol. 5, no. 2, pp. 298–305, 2011, available at: [Google Scholar](#).
- [3] I. G. P. S. Wijaya, I. B. K. Widiartha, K. Uchimura, and G. Koutaki, "Pornographic image rejection using eigenporn of simplified LDA of skin ROIs images," in *2015 International Conference on Quality in Research (QiR)*, 2015, pp. 77–80, doi: [10.1109/QiR.2015.7374899](https://doi.org/10.1109/QiR.2015.7374899).
- [4] C. Ries and R. Lienhart, "A survey on visual adult image recognition," *Multimed. Tools Appl.*, vol. 69, no. 3, pp. 661–688, 2014, doi: [10.1007/s11042-012-1132-y](https://doi.org/10.1007/s11042-012-1132-y).
- [5] L.-H. Lee and C.-J. Luh, "Generation of pornographic blacklist and its incremental update using an inverse chi-square based method," *Inf. Process. Manag.*, vol. 44, no. 5, pp. 1698–1706, 2008, doi: [10.1016/j.ipm.2008.05.001](https://doi.org/10.1016/j.ipm.2008.05.001).
- [6] R. Mustafa and D. Zhu, "Objectionable Image Detection in Cloud Computing Paradigm-A Review," *J. Comput. Sci.*, vol. 9, no. 12, pp. 1715–1721, 2013, doi: [10.3844/jcssp.2013.1715.1721](https://doi.org/10.3844/jcssp.2013.1715.1721).
- [7] A. A. Zaidan, N. N. Ahmad, H. A. Karim, M. Larbani, B. B. Zaidan, and A. Sali, "On the multi-agent learning neural and Bayesian methods in skin detector and pornography classifier: An automated anti-pornography system," *Neurocomputing*, vol. 131, pp. 397–418, 2014, doi: [10.1016/j.neucom.2013.10.003](https://doi.org/10.1016/j.neucom.2013.10.003).
- [8] W. Hu, O. Wu, Z. Chen, Z. Fu, and S. Maybank, "Recognition of Pornographic Web Pages by Classifying Texts and Images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 1019–1034, 2007, doi: [10.1109/TPAMI.2007.1133](https://doi.org/10.1109/TPAMI.2007.1133).
- [9] J. A. M. Basilio, G. A. Torres, G. S. Pérez, L. K. T. Medina, and H. M. P. Meana, "Explicit image detection using YCbCr space color model as skin detection," *Appl. Math. Comput. Eng.*, pp. 123–128., 2011, available at: [Google Scholar](#).
- [10] J.-L. Shih, C.-H. Lee, and C.-S. Yang, "An adult image identification system employing image retrieval technique," *Pattern Recognit. Lett.*, vol. 28, no. 16, pp. 2367–2374, 2007, doi: [10.1016/j.patrec.2007.08.002](https://doi.org/10.1016/j.patrec.2007.08.002).

- [11] M. Mironovova and J. Bila, "Fast fourier transform for feature extraction and neural network for classification of electrocardiogram signals," in *2015 Fourth International Conference on Future Generation Communication Technology (FGCT)*, 2015, pp. 1–6, doi: [10.1109/FGCT.2015.7300244](https://doi.org/10.1109/FGCT.2015.7300244).
- [12] W. Yongmao and Z. Shan, "Improved method based on DCT and locality preserving projection for face recognition.," *J. Huazhong Norm. Univ.*, vol. 48, no. 2, 2014, doi: [10.3724/SP.J.1087.2012.00528](https://doi.org/10.3724/SP.J.1087.2012.00528).
- [13] Z. Abidin and A. Alamsyah, "Wavelet based approach for facial expression recognition," *Int. J. Adv. Intell. Informatics*, vol. 1, no. 1, pp. 7–14, 2015, doi: [10.26555/ijain.v1i1.7](https://doi.org/10.26555/ijain.v1i1.7).
- [14] M. Wang, H. Jiang, and Y. Li, "Face Recognition based on DWT/DCT and SVM," in *Computer Application and System Modeling (ICCSM), 2010 International Conference on*, 2010, vol. 3, pp. V3--507, doi: [10.1109/ICCSM.2010.5620666](https://doi.org/10.1109/ICCSM.2010.5620666).
- [15] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94).
- [16] W. Chen, M. J. Er, and S. Wu, "PCA and LDA in DCT domain," *Pattern Recognit. Lett.*, vol. 26, no. 15, pp. 2474–2482, 2005, doi: [10.1016/j.patrec.2005.05.004](https://doi.org/10.1016/j.patrec.2005.05.004).
- [17] I. G. P. S. Wijaya, I. B. K. Widiartha, and S. E. Arjarwani, "Pornographic Image Recognition Based on Skin Probability and Eigenporn of Skin ROIs Images," *TELKOMNIKA (Telecommunication Comput. Electron. Control.)*, vol. 13, no. 3, pp. 985–995, 2015, doi: [10.12928/telkomnika.v13i3.1476](https://doi.org/10.12928/telkomnika.v13i3.1476).
- [18] I. G. P. S. Wijaya, K. Uchimura, and G. Koutaki, "Pornographic Image Recognition using Eigenporn of HSV Skin Segmented Image," in *Proceeding of the 2015 IEICE General Conference*, 2015, doi: [10.1109/QiR.2015.7374899](https://doi.org/10.1109/QiR.2015.7374899).
- [19] X. Ou, H. Ling, H. Yu, P. Li, F. Zou, and S. Liu, "Adult image and video recognition by a deep multicontext network and fine-to-coarse strategy," *ACM Trans. Intell. Syst. Technol.*, vol. 8, no. 5, p. 68, 2017, doi: [10.1145/3057733](https://doi.org/10.1145/3057733).
- [20] I. G. P. S. Wijaya, I. B. K. Widiartha, K. Uchimura, and G. Koutaki, "Pornographic Image Recognition Using Fusion of Scale Invariant Descriptors," in *Proc. of the 21st Korea-Japan joint Workshop on Frontiers of Computer Vision (FCV 2015)*, doi: [10.1109/FCV.2015.7103754](https://doi.org/10.1109/FCV.2015.7103754).
- [21] W. Chen, M. J. Er, S. Wu, and , DCT, "PCA and LDA in Pattern Recognition Letter, Volume pp.," vol. 26, pp. 2474–2482, 2005, doi: [10.1016/j.patrec.2005.05.004](https://doi.org/10.1016/j.patrec.2005.05.004).
- [22] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991, doi: [10.1162/jocn.1991.3.1.71](https://doi.org/10.1162/jocn.1991.3.1.71).
- [23] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012, available at: [Google Scholar](https://scholar.google.com/).
- [24] T. K. Hazra, D. P. Singh, and N. Daga, "Optical character recognition using KNN on custom image dataset," in *2017 8th Annual Industrial Automation and Electromechanical Engineering Conference (IEMECON)*, 2017, pp. 110–114, doi: [10.1109/IEMECON.2017.8079572](https://doi.org/10.1109/IEMECON.2017.8079572).
- [25] Z. Ma and A. Leijon, "Human Skin Color Detection in RGB Space with Bayesian Estimation of Beta Mixture Models," in *Proc. of 18th European Signal Processing Conference (EUSIPCO-2010)*, 2010, pp. 1204–1208, available at: <https://ieeexplore.ieee.org/abstract/document/7096733>.
- [26] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques," in *Proc. Graphicon-2003*, 2003, pp. 85–92, available at: [Google Scholar](https://scholar.google.com/).
- [27] T. M. Mahmoud, "A New Skin Color Detection Technique," *World Acad. Sci. Eng. Technol.*, vol. 2, no. 7, pp. 434–438, 2008, available at: [Google Scholar](https://scholar.google.com/).
- [28] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006, doi: available at: <https://dl.acm.org/citation.cfm?id=1076432>.
- [29] H. Demuth, M. Beale, and M. Hagan, "Neural network toolbox," 1992, available at: [Google Scholar](https://scholar.google.com/).
- [30] B. Sharma and K. Venugopalan, "Comparison of neural network training functions for hematoma classification in brain CT images," *IOSR J. Comput. Eng.*, vol. 16, no. 1, pp. 31–35, 2014, doi: [10.9790/0661-16123135](https://doi.org/10.9790/0661-16123135).
- [31] M. Latah, "Human action recognition using support vector machines and 3D convolutional neural networks," *Int. J. Adv. Intell. Informatics*, vol. 3, no. 1, pp. 47–55, 2017, doi: [10.26555/ijain.v3i1.89](https://doi.org/10.26555/ijain.v3i1.89).