Neural Network Based Objectionable Image Detection

Handaru Jati[1], Dhanapal Durai Dominic[2]
Computer and Information Sciences, Universiti Teknologi PETRONAS
[1]handaruj@yahoo.com, [2]pdddominic@yahoo.com

Abstract- **In this paper we propose objectionable image detection based on neural network. The proposed system is a new method in that shape information is used to classify the objectionable images. First, the proposed system finds face regions using image feature vector. The image feature vector is used as an input to a neural network, and using the same procedure is used to find objectionable image. Experimental results show that the proposed system can achieve a good classification performance.**

I. INTRODUCTION

The growth of the Internet, the reducing price of storage devices and people's interest on images has been contributing to create the Internet an enormous image library. However, among these images, some are offensive and objectionable, for example, pornographic images. It is now an urgently important duty to prevent people, especially children from accessing this kind of objectionable material. The research reported here proposed on discriminating objectionable images from non-objectionable images based on the shape features in the images. Software image content filtering is very complex and uses many approaches and methods. In principally, in this method, the filtering software has a function to examine the content of the websites when user tries to open it (1). One of the methods is automatic system to detect adult images based on computer vision and pattern recognition. Firstly, images are filtered by skin color model and then the output images are classified by support vector machine (2, 3). The other method marks skin like pixels using combined color and texture properties. These skin regions are then fed to a specialized grouper, which attempts to group a human figure using geometric constraints on human structure. If the grouper finds a sufficiently complex structure, the system decides a human is present. The approach is effective for a wide range of shades and colors of skin and human configurations (4). Adaptive skin detection methods is methods in software content filtering which allows modeling and detection of the true skin color pixels with significantly higher accuracy and flexibility than previous methods (5). This process performs a rough skin classification using a generic skin model, which defines the skin similar space. The skin similar faces often contains many non skin pixels due to the inevitable overlap in the color space between skin pixels and some non-skin pixels under the generic skin model (5). A new method was developed using web robot to collect objectionable web contents. This method shortens the update period of the database, increases the number of uniform resource locator (URL) in the database, and enhances the accuracy of the information in the database (6). The rest of the paper is organized as follows: In section 2, we briefly review methodology for detecting objectionable image. Details of neural network system are described in section 3. Experimental results are shown in section 4 and section 5 concludes the paper.

II. METHODOLOGY

The proposed system contains two major modules as in figure 1. The face detection and specific object (woman breast) image are used to categorize objectionable images. The image feature vector is generated from the image and then used as an input to a neural network image processing which determines whether images are "objectionable" or not.

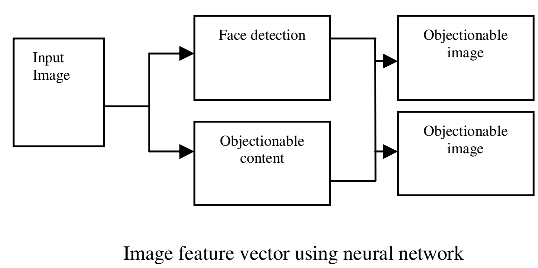

Figure1. Proposed objectionable image detection system

A. *Face detection and objectionable content*

The images, containing faces and objectionable content as illustrated in figure 2, can be categorized as objectionable images. To resolve this problem, the proposed system uses the face detection and objectionable content using neural network algorithm.

Figure2. Face and objectionable content
Detection results

After performing face detection and objectionable content, the image is classified as objectionable image. Face detection and objectionable content results are shown in figure 2.

B. Neural Network

As discussed in the introduction, there are many different approaches to objectionable image detection, each with their own techniques. One such approach is that of Neural Networks. This section gives a brief explanation to the theory of neural networks and presents neural network-based object recognition. Neural Nets are essentially networks of simple neural processors, arranged and interconnected in parallel. Neural Networks are based on our current level of knowledge of the human brain. Since the early stages of development in the 1970's, interest in neural networks has spread through many fields, due to the speed of processing and ability to solve complex problems. As with all techniques though, there are limitations. They can be slow for complex problems, are often susceptible to noise, and can be too dependent on the training set used, but these effects can be minimized through careful design. Neural Nets can be used to construct systems that are able to classify data into a given set or class, in the case of face detection and objectionable content, a set of images containing face and objectionable content, and a set of images that contains non face and objectionable content. Neural Networks consist of parallel interconnections of simple neural processors. Figure 3 shows an example of a single neural processor, or neuron. Neurons have many weighted inputs, that is to say each input (p1, p2, p3… pm) has a related weighting (w1, w2, w3… wm) according to its importance. Each of these inputs is a scalar, Representing the data.

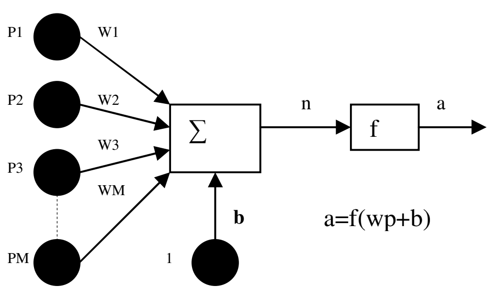

Figure3. A single neuron example neural network

In the case of face detection and objectionable content, the shade of GRAY of each pixel could be presented to the neuron in parallel (thus for a 18x27 pixel image, there would be 486 input lines p1 to p486, with respective weightings w1 to w486, corresponding to the 486 pixels in the input image). The weighted inputs are combined together, and if present, a bias (b) is added. This is then passed as the argument to a transfer function (typically a pure linear, hardlimit, or log-sigmoid function), which outputs a value (a) representing the chosen classification. Problems that are more complex can be realized by adding more neurons, forming multiple layers of several neurons, interconnected via a weighted matrix. Additional layers of neurons not connected directly to the inputs or the outputs are called hidden layers. Once the architecture is established, the network must be trained. A labeled representative set of examples from each class is presented to the network, which attempts to classify each example. The weights and biases are initialized with small random values and updated incrementally, such that the performance of the detector improves producing a more accurate decision boundary for the problem. Once trained, the network can be used to classify previously unseen images, indicating whether they contain faces and objectionable content based-on the 'location' of the input relative to the decision boundary formed during training.

C. Sample Data

The first modification regards the sample training data. This data set contains 111 training examples, 69 faces and 55 non faces, and also 60 objectionable images and 55 non objectionable images. By resizing the images to 27x18 pixels, converting each image to an 8 bit Grayscale PNG, and renaming them. The operation of the objectionable detection system can be broken down into three main areas:

- Initialization (design and creation of a neural network)
- Training (choice of training data, parameters, and training)
- Classification (scanning images to locate faces)

A feed forward neural network is created which is trained using back propagation. The training set used contains examples of both face and non-face images and objectionable and non objectionable, and the classifier is trained to output a value between 0.9 and -0.9 (0.9 firmly indicating the presence of a face, -0.9 firmly indicating the absence of a face). When a new image is presented to the network, the image is rescaled and divided into windows which are individually presented to the network for classification. All the experimental work is to be carried out in Matlab.

The growth of the Internet, the reducing price of storage devices and people's interest on images has been contributing to create the Internet an enormous image library. However, among these images, some are offensive and objectionable, for example, pornographic images. It is now an urgently important duty to prevent people, especially children from accessing this kind of objectionable material. The research reported here proposed on discriminating objectionable images from non-objectionable images based on the shape features in the images. Software image content filtering is very complex and uses many approaches and methods. In principally, in this method, the filtering software has a function to examine the content of the websites when user tries to open it (1). One of the methods is automatic system to detect adult images based on computer vision and pattern recognition. Firstly, images are filtered by skin color model and then the output images are classified by support vector machine (2, 3). The other method marks skin like pixels using combined color and texture properties. These skin regions are then fed to a specialized grouper, which attempts to group a human figure using geometric constraints on human structure. If the grouper finds a sufficiently complex structure, the system decides a human is present. The approach is effective for a wide range of shades and colors of skin and human configurations (4). Adaptive skin detection methods is methods in software content filtering which allows modeling and detection of the true skin color pixels with significantly higher accuracy and flexibility than previous methods (5). This process performs a rough skin classification using a generic skin model, which defines the skin similar space. The skin similar faces often contains many non skin pixels due to the inevitable overlap in the color space between skin pixels and some non-skin pixels under the generic skin model (5). A new method was developed using web robot to collect objectionable web contents. This method shortens the update period of the database, increases the number of uniform resource locator (URL) in the database, and enhances the accuracy of the information in the database (6). The rest of the paper is organized as follows: In section 2, we briefly review methodology for detecting objectionable image. Details of neural network system are described in section 3. Experimental results are shown in section 4 and section 5 concludes the paper.

III. Result and Discussion

The most important part of the system lies the classifier, the object that actually makes the decision as to whether an image is receives contains an objectionable object and face or not. Initially tests were carried out to investigate how well the classifier could classify the data set on which it was trained. Although this is not indicative of true performance, it serves as a guide to how well the network is learning from the training data and classifies 27x18 pixel images as either face or non-face images, producing a numerical value between -0.9 and 0.9 (0.9 strongly indicating the presence of a objectionable image and face, -0.9 indicating the absence of a face). The addition of a threshold value allows the classifier to be tuned somewhat, such that an image is marked as when the numerical output value exceeds the threshold. Upon completion the classification rate and number of incorrect classifications for both the face and non-face examples are displayed. Figure 4 shows the detection result for this system.

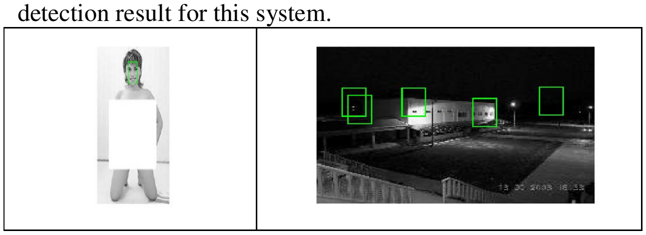

Figure4. Result of the detection system

TABLE I
Comparison of Performance statistics

	Number of objectable	Number of non object detected	Correct rate
Test 1	5	0	100%
Test 2	13	0	86.6%
Test 3	8	0	75%
Test 4	5	1	50%
Test5(non object)	0	1	0%
Test6(non object)	0	1	0%
Test7(non object)	0	1	0%
Test8(non object)	0	10	0%

The threshold value based from this experiment between 0.4 and 0.6 is optimal. The results (figure5) were encouraging with a high result in detection rates, although an accompanying with the high number in the false detections.

IV. Conclusion

Object detection has a many applications including internet content filtering applications, until recently much of the work in the field of computer vision has focused on object recognition. Human face detection is often the first-step in the recognition process as detecting the objectionable image content, prior to attempting recognition can focus computational resources on the face area of the image. Although a trivial task that human performs effortlessly, the task of face detection is a complex problem in computer vision, due the great multitude of variation present across faces. Feature-Based Approaches, using neural network often applied to real-time systems, are often reliant on a priori knowledge of the face, which is used explicitly to detect features. Based on the test sample, on average 77.7% of the face examples were classified correctly.

References

[1]. Kherfi ML, Ziou D, Bernardi A. Image Retrieval from the World Wide Web: Issues, Techniques, and Systems. ACM Computing Surveys (CSUR) 2004;36(1):35-67.

[2]. Duan L, Cui G, Gao W, Zhang H. Adult image detection method base-on skin color model and support vector machine. The Fifth Asian Conference on Computer Vision (ACCV) 2002:780–97.

[3]. Rowley HA, Jing Y, Baluja S. Large Scale Image -Based Adult-Content Filtering. International Conference on Computer Vision Theory and Applications, February 2006.

[4]. Forsyth DA, Fleck MM. Automatic Detection of Human Nudes. International Journal of Computer Vision 1999;32(1):63-77.

[5]. Zhu Q, Wu CT, Cheng KT, Wu YL. An adaptive skin model and its application to objectionable image filtering. Proceedings of the 12th annual ACM international conference on Multimedia 2004:56-63.

[6]. Choi SG, Han SW, Jeong CY, Nam TY. Specialized Web Robot for Objectionable Web Content Classification. V International Enformatika Conference(IEC 2005) 2005:18-21.