Dhayea et al. Iraqi Journal of Science, 2024, Vol. 65, No. 2, pp: 1102- 1116
DOI: 10.24996/ijs.2024.65.2.40

ISSN: 0067-2904

Human Skin Detection and Segmentation Based on Convolutional Neural Networks

Alyaa Mohsin Dhayea [1]*, Nidhal K. El Abbadi [2], Zainab Ghayyib Abdul Hasan [3]

[1] *Department of Computer Science, Faculty of Archaeology, University of Kufa, Iraq*
[2] *Department of Computer Techniques Engineering, Al-Mustaqbal University, Babylon, Iraq*
[3] *Department of Computer Science, Faculty of Education for Girls, University of Kufa, Iraq*

Received: 28/6/2022 Accepted: 25/12/2022 Published: 29/2/2024

Abstract

Human skin detection is the process of classifying the image pixels or regions as skin or non-skin. Skin detection has many applications, such as face tracking, skin diseases, nudity recognition, hand gestures, video surveillance, web content filtering, and pornographic content filtering. Skin detection is a challenging problem due to skin color variation, the human race, aging, gender, makeup, complex backgrounds, etc. This paper suggests detecting the skin region in the image and finding the location of the skin based on a convolutional neural network. In this proposal, the proposed CNN will be modified by adding two layers before the first layer of CNN and after the last layer of CNN. The main purpose of these layers is to prepare the input image by using a sliding window, which inputs an indexed small part of the image into the CNN. The network classifies each part as skin or non-skin and then sends the result into the second suggested layer. After processing all the image pixels, the non-skin blocks are mapped to the original image as black regions. The final image contains the skin regions with black in the background. The contribution of the proposed method is the ability to detect the skin from any part of the human body, unlike previous works, which focused on one part of the body. The input image is processed as blocks instead of the entire image as in the previous works, and then in the output, the original image is reconstructed. This method works well with most of the challenges that face the detection of skin, and finally, the designed network facilitates the localization and segmentation of the skin region almost accurately, while the previous networks focused on the classification of the image as skin or non-skin. The accuracy of the detection of skin when testing with images different from the training images was 95.4%.

Keywords: classification, convolution neural network, detection, image processing, human skin.

كشف جلد البشر وتحديده باستعمالCNN

علياء محسن ضايع[1*] , نضال العبادي[2], زينب غايب عبد الحسن[3]

[1*]قسم علوم الحاسوب، كلية الاثار ، جامعة الكوفة، محافظة النجف، العراق
[2]قسم هندسة تقنيات الحاسوب، جامعة المستقبل ، بابل ، العراق
[3]قسم علوم الحاسوب، كلية التربية للبنات، جامعة الكوفة، محافظة النجف، العراق

*Email: aalyam.zayi@uokufa.edu.iq

الخلاصة

الكشف عن جلد الإنسان هو عملية تصنيف نقاط الصورة أو المناطق على أنها جلد أو غير جلد. يحتوي اكتشاف الجلد على العديد من التطبيقات مثل تتبع الوجه والأمراض الجلدية والتعرف على العري وإيماءات اليد والمراقبة بالفيديو وتصفية محتوى الوب وتصفية المحتوى الإباحي. يعد اكتشاف الجلد مشكلة صعبة بسبب اختلاف لون الجلد، والجنس البشري، والشيخوخة، والجنس، والماكياج، والخلفية المعقدة، وما إلى ذلك. تقترح هذه الورقة اكتشاف منطقة الجلد في الصورة وإيجاد موقع الجلد بناءً على CNN. في هذا الاقتراح، سيتم تعديل CNN المقترحة بإضافة طبقتين قبل الطبقة الأولى من CNN وبعد الطبقة الأخيرة من CNN. الغرض الرئيسي من هذه الطبقة هو إعداد صورة الإدخال باستعمال نافذة تقوم بإدخال جزء صغير مفهرس من الصورة في شبكة CNN. تصنف الشبكة كل جزء على أنه جلد أو غير جلدي ثم ترسل النتيجة إلى الطبقة الثانية المقترحة. بعد معالجة جميع وحدات البكسل في الصورة، يتم تعيين الكتل غير الجلدية على الصورة الأصلية كمناطق سوداء. تحتوي الصورة النهائية على مناطق الجلد ذات اللون الأسود في الخلفية. تتمثل مساهمة الطريقة المقترحة في القدرة على اكتشاف الجلد من أي جزء من جسم الإنسان، على عكس الأعمال السابقة التي تركز على جزء واحد من الجسم، فإن عملية إدخال الصورة ككتل بدلاً من معالجة الصور بأكملها كما في السابق يعمل ثم في الإخراج يعيد بناء الصورة الأصلية، تعمل هذه الطريقة بشكل جيد مع معظم التحديات التي تواجه اكتشاف الجلد، وأخيرًا، تسهل الشبكة المصممة توطين وتجزئة منطقة الجلد بشكل شبه دقيق، بينما تركز الشبكات السابقة لتصنيف الصورة كجلد أو غير جلدي. كانت دقة الكشف عن الجلد عند الاختبار بصور مختلفة عن صور التدريب 95.4٪.

1. Introduction

We are awash with images. With the widespread use of cameras on smartphones and fast mobile networks, a huge number of images are uploaded every day to cloud storage, and most applications use images from social networks like Facebook, Twitter, Instagram, Google+, Flicker, etc. [1]. Human skin detection is used to differentiate between skin and non-skin pixels or regions, which is one of the important pre-processing steps in many image processing and computer vision tasks [2]. It has a wide range of applications such as face detection and human tracking [3], [4], [5], semantic filtering of web contents, nudity recognition, gesture analysis [6], image enhancement [7], surveillance systems [8], pornographic contents filtering [9], driver fatigue detection [10], dermatology diagnostics [11], etc.

Skin detection is still considered a challenging problem due to a lot of variations in skin colors due to race, aging, gender, differences in illumination, makeup, a large area of tattoos, skin-like backgrounds, complex backgrounds, etc.

Most of the previous methods worked successfully to detect one part of human skin, mainly the face, but they failed to work with the entire human skin. Recently, a convolutional neural network (CNN) has been shown to provide good performance on many image classification and processing problems, but we could find only a few researchers that apply deep learning to skin detection problems and localize the skin region [12].

In this paper, we aim to design a CNN that can work efficiently with various skin detection challenges and generalize it for working on any part of the human skin, in addition to modifying the way of processing the image in a way that makes it suitable for localization and segmenting of the skin regions. Deep learning is a very good tool to solve complex problems in real-time that look difficult to solve with human brains. There have been various types of research to improve the algorithms of deep learning to solve such a set of problems in which image processing plays a main role. Deep learning extracts the essential features from the raw data [13].

The contribution of this paper is the detection and localization of skin regions in any part of the human body, while the previous works detected the skin in a specific part of the human body, almost the face. Also, this method detects skin regardless of age, race, skin color, variation of image illumination, skin with tattoos, skin-like backgrounds, or complex backgrounds. Finally, it's the ability to find almost accurately the location of skin in an image.

The rest of this paper is organized as follows: The related works are presented in Section 2. Section 3 focuses on the explanation of the proposed method. The results are introduced in Section 4, and finally, Section 5 concludes the suggested method and its results.

2. Related work

Many researchers have suggested methods for dealing with skin detection. Some of them are:

Kim et al. [12] proposed two convolutional neural networks (CNNs) and their training strategies for skin detection. The first CNN, consisting of 20 convolution layers with 3×3 filters, is a kind of VGG network. The second is composed of 20 networking network (NiN) layers, which can be considered a modification of the Inception structure. The first method focuses on local features such as skin color and texture, while the second focuses on human-related shape features as well as color and texture. The proposed sNiN architecture generally works better than the VGG network for skin detection; the proposed methods provide more plausible results regardless of illumination variation over the face; and the accuracy was 95.62%.

Zuo et al. [14] came up with a way to find skin that uses fully convolutional neural networks (FCNs) with layers of recurrent neural networks (RNNs). FCN layers capture generic local features, while RNN layers model the semantic contextual dependencies in images. The proposed approach has been validated for effectiveness based on the COMPAQ and ECU skin datasets. RNN layers effectively improve the stability of skin detection algorithms under complex backgrounds; the accuracy was 95.93%.

Paracchini et al. [15] suggested a method based on deep learning for facial skin detection. This proposal applies to low-resolution grayscale images. The author proposed to learn to solve the problem of detecting human skin by exploiting a colorization network to use and validate a transfer learning method. The performance of this method is reported based on different datasets. It works regardless of race, age, or ethnicity. Also, testing with various illuminations and occlusions gives good results; the disadvantage of this method is that it focuses on the face only and may not work well for other parts of the human body. The accuracy rate was 84%.

Yusuf et al. [16] proposed a study that is used in various fields of application in daily life. Each color model contains a specific representation space and components and can be converted from one color space to another using a standard formula. They concluded that the selection of a proper color model for a specific application depends on the characteristics of the model and the nature of the application. This study focuses on face detection using skin color segmentation. A discovery rate of up to 97.22% was acquired by using a standard database.

Mohammed et al. [25] Differentiate the pixels as skin and non-skin based on a metric that measures the distance of pixel colors to skin tones. In this work, the author proves that the YCbCr color space can give better performance than the RGB image due to the isolation of the illuminance channel. The segmentation of skin pixels is based on a histogram. This paper focuses on the detection of faces and hands only.

Salah et al. [2] introduced an approach based on convolutional neural networks for human skin detection. The fundamental objective of this study was to explore the potential of the CNN learning model for skin pixel classification. The performance of this work was good in general. It works well with a complex background, variation in illumination, and ethnicity. The method was efficient in working with the face and hand and in rejecting non-skin pixels in various situations. The accuracy was 93.57%.

3. The Methodology

This paper aims to design and implement a method that can detect human skin in any part of the human body, regardless of many of the challenges that face skin detection, and find the location of the skin region in the image based on CNN. We attempt to solve the problem of skin detection and segment the skin location based on a convolutional neural network. The general block diagram for the proposed system is shown in Figure 1.

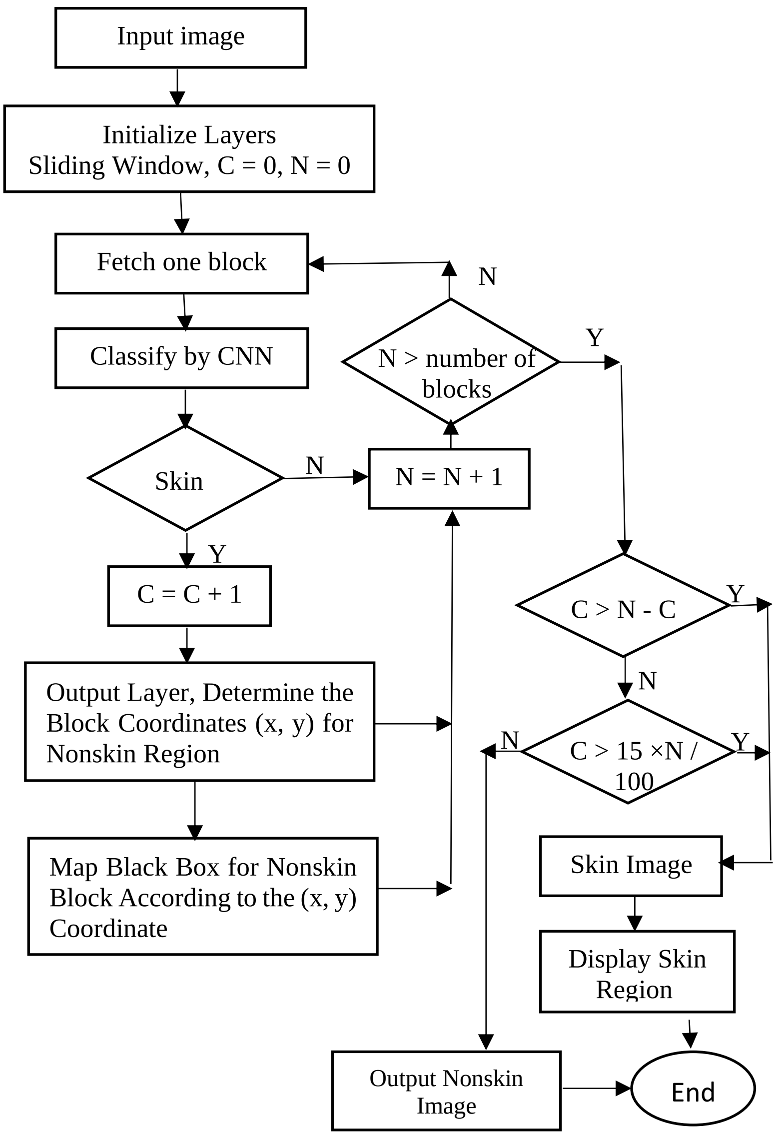

Figure 1: The block diagram of the proposed method

3.1. Dataset used

The dataset used in this proposal was collected by the authors and consists of 2200 color images (1100 with skin and 1100 without skin). The skin images are images with just the skin region used for the training. We focus on corrupting skin regions from general images that have human skin and resize this corrupted part into a 20 × 20 image size. The collected skin regions represent many different parts of the human body, such as the face, legs, hands, human abdominal, and human back regions. The skin images in this dataset included images of people from various ethnic groups and contained some images with various illuminations. A sample of the skin images in this dataset that were used only in the training stage is shown in Figure 2.

Figure 2: A sample skin image from the collected dataset used for training

While the non-skin images include images of nature, animals, and maybe objects similar to skin, resized into 20 × 20. In addition, there are some images with complex non-skin textures and images with multiple objects and complex backgrounds. Some of the non-skin images from the collected dataset are shown in Figure 3.

Figure 3: Images of non-skin from the collected dataset

3.2. Proposed Convolutional Neural Network (CNN)

A convolutional neural network (CNN) is a category of models dedicated to extracting features from 2D inputs. The proposed method comprises three parts: the initialize layer, CNN layers, and output layer. The initialized layer is responsible for preparing the input image, whereas a sliding, non-overlapping window scans the entire image, corrupting each block of the image and labeling it with an index before input to the CNN. CNN is a mathematical construct that is typically composed of three types of layers (or building blocks): convolution, pooling, and fully connected layers. The first two, the convolution and pooling layers, perform feature extraction, whereas the third is a fully connected layer that maps the extracted features into the final output, such as classification, as in the general purpose of CNN. The proposed CNN architecture consists of three convolution layers with kernels 8, 16, and 32, respectively. Two layers of max pooling, one after the first convolution layer and the other after the second convolution layer, and activation function (ReLU) with each of the three layers.

The filter size is 3×3. Finally, the fully connected layer classifies the input block as skin or non-skin and sends it to the output layer. In the output layer, the non-skin block classified in

the CNN is mapped into the original image as a black color block at the location according to the block index. Figure 4 shows the proposed architecture of CNN.
The proposed CNN is trained on skin and non-skin images. All the images in the training stage are of size 20×20 and each image is labeled as skin or non-skin. The training stage is summarized in Algorithm 1.

Algorithm 1- A training stage
Input: color image Output: trained network
Step 1: Load RGB images from the dataset. Step 2: Label each image as skin or non-skin. Step 3: Design the proposed CNN architecture. Step 4: Input images into CNN. Step 5:Train the CNN to classify the input image as skin or nonskin. Step 6: Check the performance.

When the training stage has been completed and the weight of CNN parameters has not changed more, the CNN network will be ready for testing. Testing steps are listed in Algorithm 2. For the testing stage, we used the same proposed CNN architecture that we used in the training stage with some modifications by adding an extra two layers before the first CNN layer (the initialization layer) and after the last layer (the output layer), as explained previously. The input image is loaded into the initialize layer, and a non-overlapping sliding window of size (20 x 20) will scan the entire image from left to right and top to bottom. The first block of pixels covered by a sliding window will be input into the first CNN layer as a corrupted segment of an image with an index (each block assigns an index starting from zero before input into the CNN). For each image, the CNN is implemented as many times as the number of blocks generated as a result of window scanning. At each iteration, one block of pixels covered by the sliding window will be input into the CNN and then classified as skin or non-skin. The results will be passed to the output layer, where the non-skin block will be mapped in the original image as a black block in the location determined by its index according to the Eqs. (1, 2, and 3). If the entire image has not been scanned yet, go back to the initialization layer to fetch another block. This process continues until all images are scanned by the sliding window.

NC = trunk (image width) / block size. (1)
x = trunk (block No. / NC) * block size (2)
y = (block No. mod NC) * block size (3)
where NC is the number of columns.
(x, y) is the top-left corner of the block in the original image. Algorithm 2 summarizes the testing stage.

Algorithm 2- A testing stage
Input: color image Output: classification results (skin, or non-skin)
Step 1: Input RGB image from the dataset. Set the counter (C) to zero. Where (C) represents the number of skin blocks. Step 2: Scan the image with a nonoverlapped sliding window, with a size of (20×20). Step 3: Label each block under the sliding window with a sequence index number (from left to right and top to down) starting from zero. Step 4: Process each block with CNN layers. Step 5: The CNN classified the processed image (block) as skin or non-skin. Step 6: Change the nonskin region of the original Image with coordinate (x .. x+20, y .. y+20) to black color. (x, y) determined by equations (1, 2, and 3) Step 7: Increase the counter (C). Step 8: If the sliding window does not reach the end of the image size, go to step 2. Step 10: Check the count with the threshold. Step 11: If the count number is larger than the threshold then decide whether the image contains skin. Step 12: Determine the location (x, y) of the skin region in the original image. Step 14: Display the image.

The architecture of the proposed CNN is shown in Figure 4. The best window size is 20×20. This size was achieved experimentally; we tested many window sizes and found that the 20×20 size gives the best accuracy and minimum time for each image. There is a counter (C) to count the number of blocks that are classified as skin.

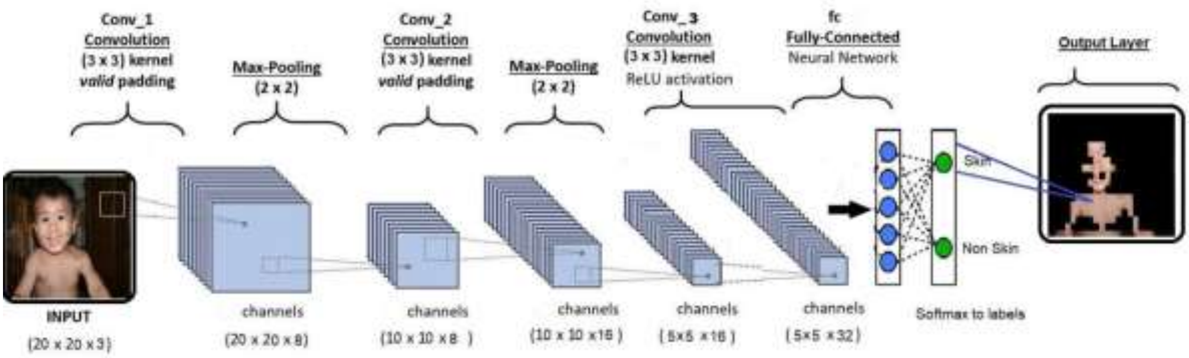

Figure 4: The proposed CNN architecture

When all the blocks are classified, then we have to decide if the image contains skin when one or both of the following conditions in Eqs. (4) and (5) are met:

$$C > (N - C) \quad (4)$$
$$C > 15 * N / 100 \quad (5)$$

Where (N) is the number of blocks for the image.
C is the number of skin blocks.

4. The Results

In the current proposal, we aim to detect the skin in images and localize the location of the skin. We proposed a CNN network with some modifications to achieve this aim. The architecture of the proposed CNN is designed precisely, and each important parameter in the network has been carefully determined to choose the optimum value. Images used in the test stage are from the collected dataset; the images contain human skin with various backgrounds, as shown in Figure 5.

The number of images used for testing was 300, of which 50 were related to the challenges of this field of study.

Figure 5: Sample images used in the test stage

The first part of testing is focused on calculating the optimum values for the CNN parameters and determining the performance of the proposed method.

The number of layers for the CNN network that gives the best accuracy was determined. Each layer of CNN extracts features and returns a feature map that is used as input into the next layer; this process is repeated for all network layers. Finally, the vector of features created contains as many elements as there are classes. So, the number of layers is a crucial part of designing the CNN network. In this test, we design many networks with a different number of layers and then determine the accuracy for each one, as shown in Figure 6. We found that a network with three layers is the best design.

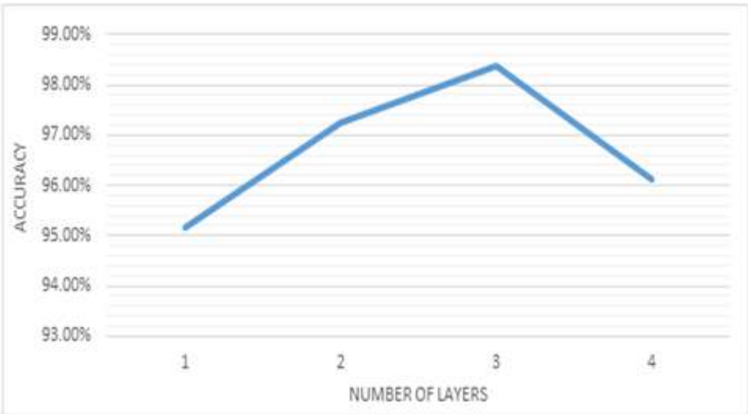

Figure 6: The accuracy of various CNN networks (each with a different number of layers)

Although the filter size of 3×3 is a good choice for most CNN designers, that does not mean it is the optimal choice for all CNN applications. For that reason, we selected the filter size experimentally. The accuracy is determined for the network when using different filter sizes, as shown in Figure 7.

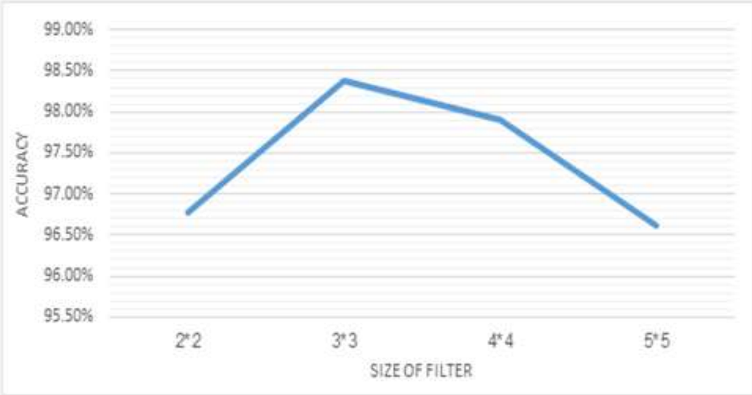

Figure 7: The performance of the CNN network when using various filter sizes

The learning rate is the most important hyperparameter in the model because it controls the speed at which it learns. Each time the weights are updated, the learning rate controls the amount of error that occurs when the weights of the model are updated. Unfortunately, for any model with a specific dataset, we cannot analytically determine the optimal learning. Instead, trial and error is the only way to discover a good learning rate. In this test, many learning rate values are used to measure the corresponding model accuracy. Figure 8 shows a sample of these values. We argue that the value of 0.001 gives the best accuracy.

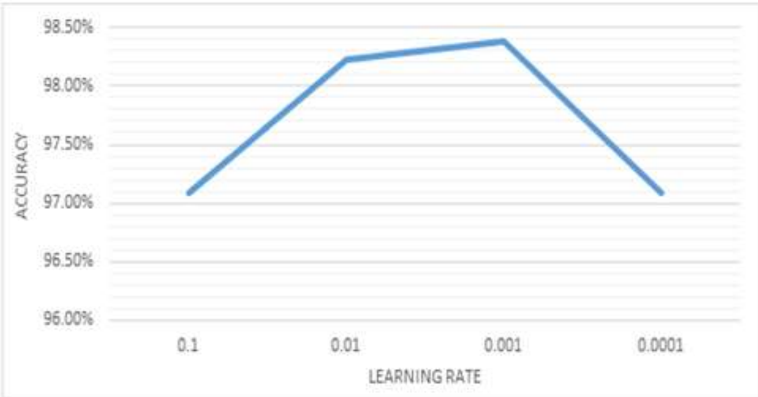

Figure 8: The relationship between learning rate and network performance

Another factor that has a high effect on network efficiency is the number of epochs, which determines how many times the weights of the network must be changed. Increasing the number of epochs means increasing the number of times the weights in the neural network change, and the boundary goes from underfitting to optimal to overfitting to determine the optimal number of epochs. Figure 9 shows the results of this test. From this test, we found that twenty epochs give very good detection for images with skin, while the twenty-five epochs give the best detection for non-skin images and have less accuracy for detecting the images with skin. For our proposal, we prefer detecting skin over non-skin, so we selected twenty epochs for our model.

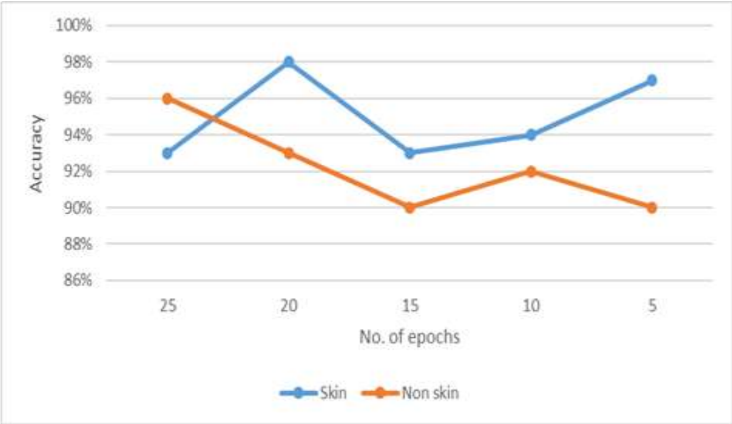

Figure 9: The performance of the proposed model for detection images with skin and non-skin when using a various number of epochs

The best model results are obtained when the number of epochs equals twenty. The training accuracy = 100%, and the validation accuracy = 98.38%, as shown in Figure 10. While training loss = 0%, and validation loss = 0.1%, as shown in Figure 11.

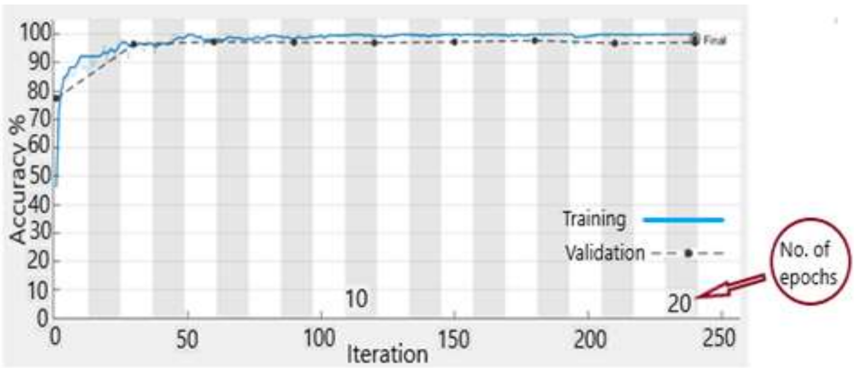

Figure 10: The relationship between No. of epochs and network efficiency.

The graph is drawn at the training stage by the suggested model.

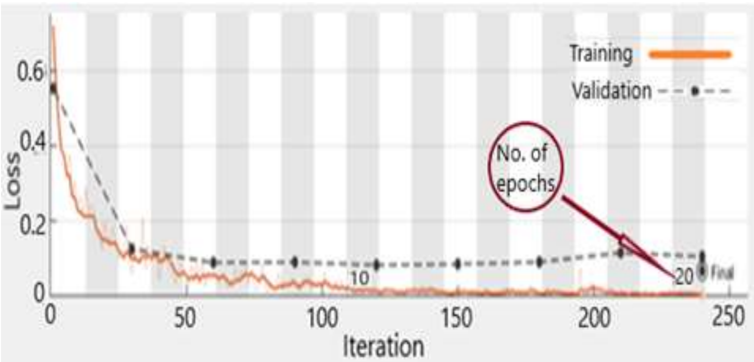

Figure 11: The loss function of the proposed model

Another factor that has a high effect on the proposed network efficiency is the size of the sliding window. As we mention in the methodology, the input image is scanned with a non-overlapping sliding window. The sliding window is changed into different sizes and corrupts a block of pixels from the input image to be imaged into the CNN. So we have to decide what the best window size is that can give high accuracy when the number of epochs is twenty. As shown in Figure 12, we conclude that the window size of 20×20 gives the best accuracy in general (this window size is for training and testing). While Figure 13 shows the accuracy corresponding to the window size for the classification of skin or non-skin, Figures 14 and 15 show the time for training and testing.

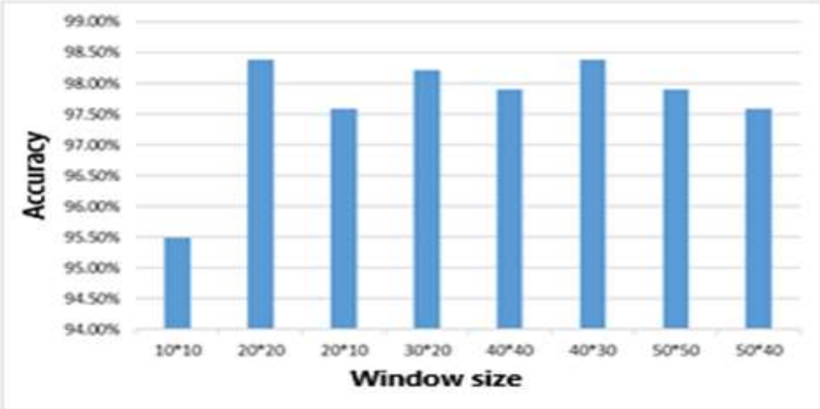

Figure 12: The relationship between the size of the image and network efficiency

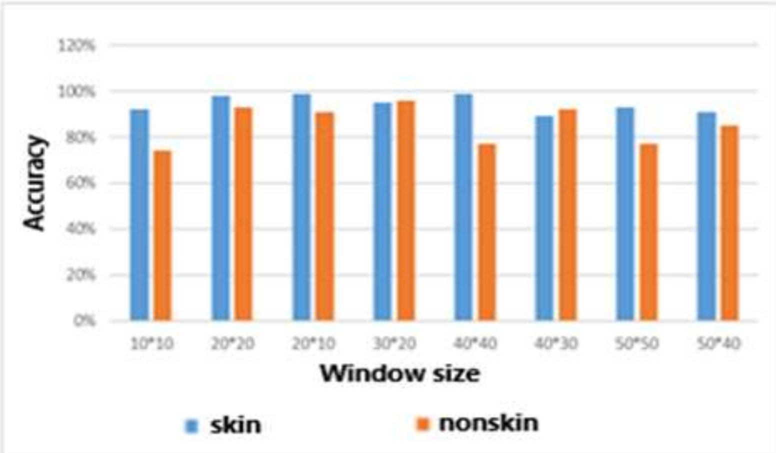

Figure 13: The model's performance in detecting skin and non-skin when using various sizes of blocks of the image

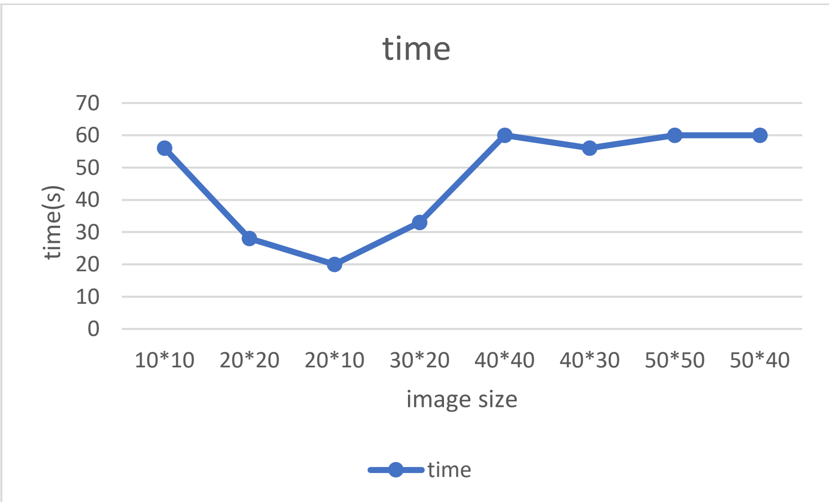

Figure 14: The relationship between the size of the image and the time for training

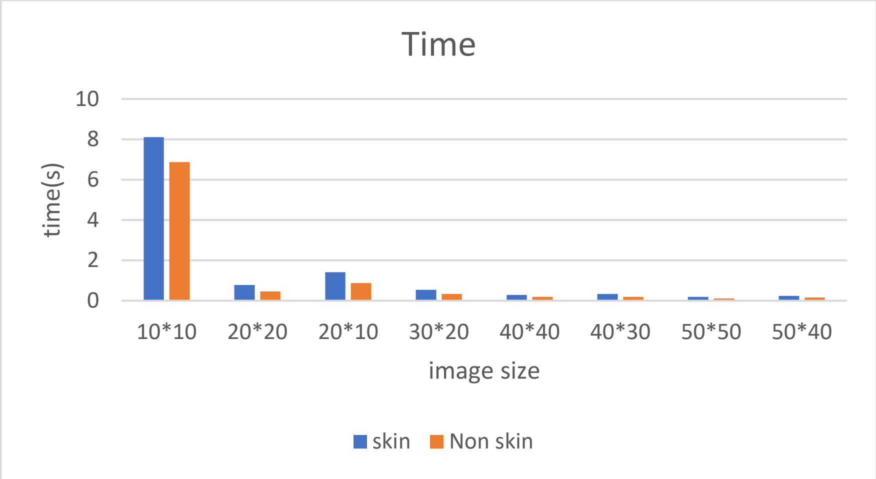

Figure 15: The relationship between the size of the image and the time for testing

The proposed method was tested with images different from the training images—images not seen by the model previously. The performance measurements are used to measure the model's performance. The value of these metrics was recall = 93%, precision = 98%, F1-score = 95.4, and accuracy = 95%. A sample of skin images detected by the proposed model with the corresponding skin localization image is shown in Figure 16.

Figure 16: Sample of skin images segmented by the proposed model. The first row is the original image,
while the second row is the segmented skin for the corresponding images in the first row.

One of the important tests is to determine the performance of the suggested model in the detection and localization of skin regions in images containing challenging human skin. The results were promised, as shown in Figure 16, where the image in Figure 17 (A) is skin with tattoos. The model detects the skin and isolates the tattoos successfully. Human age is detected successfully, as in the images in Figure 17 (B and C). Also, an image in Figure 17 (D) shows the ability of the proposed model to detect a very difficult case (the hands and face of the cameraman, which are blurred). Faces with makeup are easily detected, as in Figure 17 (E). The last image is a hand with colored tattoos discovered successfully.

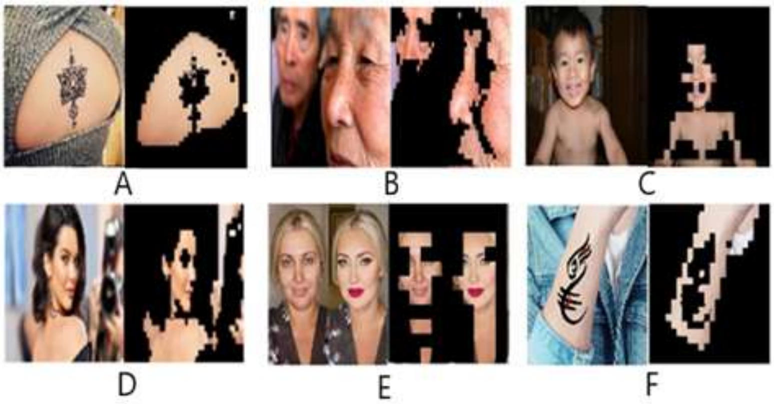

Figure 17: A sample of challenging skin images detected by the proposed model

Finally, the proposed model is compared with other similar methods, as shown in Table 1. We argue that the proposed model provides better performance than the other methods.

Table 1: Comparison of the proposed method with other methods

No	Author	Method used	Recall %	Precision %	F1 %	Accuracy %	Dataset used
1	Khawla Ben Salah et al. [2021](2)	convolutional neural network	97.30	98.01	97.65	93.57	Pratheepan and SFA
2	Amna Shifa et al. [2019](17)	color-space	83.30	72.69	77.63	91.78	TDSD
3	Iptehaj Alhakam et al. [2014](18)	An Improved Probability Density Function (PDF)				96.05	Caltech (California Technology Institute)
4	S. Kolkur et al. [2017](19)	Color Models		89.33		94.43	the dataset is downloaded randomly from Google
5	Mohammadreza Hajiarbabi et al. [2015](20)	Skin Deep Learning	75.38	46.05	57.17	88.81	VT-AAST
6	Marco Paracchini et al. [2020](15)	deep learning model			84		Helen and MUCT
7	Hani K. Al-Mohair et al. [2015](21)	Neural Network and K-Means Clustering Technique	88.00	87.65	87.82		ECU
8	Yong Luo et al. [2017](22)	hybrid color space strategy	87.42	83.23	85.28	93.74	TDSD
9	Yoonsik Kim et al. [2017](12)	convolutional neural networks (CNN)	91.22	87.20	89.17	95.62	ECU and Pratheepan
10	Ma, C et al. [2018](23)	FCNN	89.81	84.80	86.78	94.99	Pratheepan
11	Hwang, I et al. [2017](24)	SPSD	93.28	76.59	84.12	87.82	Pratheepan
	Proposed method	CNN	93	98	95.4	95	collected dataset

5. Conclusion

In this work, a proposed approach based on modified convolutional neural networks for human skin detection and localization in an image is presented. The fundamental objective of this study was to explore the potential of the CNN learning model for skin pixels or region classification and determine their location in the image by modifying the way the image is processed in CNN. The first contribution of this paper is adding an initialized layer before the first CNN layer with a sliding window to input the image as blocks instead of the entire image, and another layer after the last layer of the CNN. This gives the network the ability to detect and localize the skin region (segment the skin region). In this work, the CNN was modified to work as a segmentation network in addition to classification. The proposed CNN design involves selecting most of its parameters experimentally. This method has a very good ability to segment the skin region in an image regardless of the many challenges faced by skin detection methods such as skin color variation, aging, race, gender, makeup, large regions of tattoos, complex backgrounds, etc. The results obtained from this method were compared with other works, and they were better. The accuracy of the detection of skin and non-skin images was 95.4%. For future work, we suggested working on other challenges that occur in the images, such as the presence of thick hair, rain, and snow, covering the skin with mud, the skin being wet with water, etc. Also, we recommend using this proposal for the segmentation of other objects.

6. References

[1] S. S. Farfade, M. J. Saberian, and L. J. Li, "Multi-view face detection using deep convolutional neural networks," in *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, pp. 643-650, 2015.

[2] K. B. Salah, M. Othmani, and M. Kherallah, "A novel approach for human skin detection using convolutional neural network," *The Visual Computer*, vol. 38, no. 5, pp. 1833-1843, 2022.

[3] R. L. Hsu, M. Abdel-Mottaleb and A. K. Jain, "Face detection in color images," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 696-706, 2002.

[4] S. J. A. Al-Atroshi and A. M. Ali, "Facial Expression Recognition Based on Deep Learning: An Overview," *Iraqi Journal of Science*, vol. 64, no. 3, pp. 1401-1425, 2023.

[5] H. K. Dhahir and N. H. Salman, "A Review on Face Detection Based on Convolution Neural Network Techniques," *Iraqi Journal of Science*, vol. 63, no. 4, pp. 1823-1835, 2022.

[6] H. Francke, J. Ruiz-del-Solar and R. Verschae, "Real-time hand gesture detection and recognition using boosted classifiers and active learning, " *In Pacific-rim symposium on image and video technology*, pp. 533-547, Springer, Berlin, Heidelberg, 2007.

[7] B. Zafarifar, E. B. Bellers and P. H. de With, "Application and evaluation of texture-adaptive skin detection in TV image enhancement, " *in Proceedings of the 2013 IEEE International Conference on Consumer Electronics (ICCE)*, IEEE, pp. 88-91, 2013.

[8] Z. Zhang, H. Gunes and M. Piccardi, "Head detection for video surveillance based on categorical hair and skin color models, "*in Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP)*, IEEE, pp. 1137-1140, 2009.

[9] M.J. Jones and J.M. Rehg, "Statistical color models with application to skin detection, " *International Journal of computer vision*, vol. 46, no. 1, pp. 81-96, 2002.

[10] M. S. Devi and P.R. Bajaj, "Driver fatigue detection based on eye tracking, " *in Proceedings of the 2008 First International Conference on Emerging Trends in Engineering and Technology*, IEEE, pp. 649-652, 2008.

[11] G. Schaefer, R. Tait, and S. Y. Zhu, "Overlay of thermal and visual medical images using skin detection and image registration, "*in Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, pp. 965-967, 2006.

[12] Y. Kim, I. Hwang, and N. I. Cho, "Convolutional neural networks and training strategies for skin detection," *in Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, IEEE, pp. 3919-3923, 2017.

[13] E. M. T. A. ALSAADI and N. K. El Abbadi, " An automated classification of mammals and reptiles animal classes using deep learning," *Iraqi Journal of Science*, vol. 61, no. 9, pp. 2361-2370, 2020.

[14] H. Zuo, H. Fan, E. Blasch and H. Ling, "Combining convolutional and recurrent neural networks for human skin detection, " *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 289-293, 2017.

[15] M. Paracchini, M. Marcon, F. Villa, and S. Tubaro, "Deep skin detection on low-resolution grayscale images, " *Pattern Recognition Letters*, vol. 131, pp. 322-328, 2020.

[16] A. A. Yusuf, F. S. Mohamad, and Z. Sufyanu, "Human face detection using skin color segmentation and watershed algorithm, "*American Journal of Artificial Intelligence*, vol. 1, no. 1, pp. 29-35, 2017.

[17] A. Shifa, M. B. Imtiaz, M. N, Asghar and M. Fleury, "Skin detection and lightweight encryption for privacy protection in real-time surveillance applications, " *Image and Vision Computing*, vol. 94, 2020.

[18] I. Alhakam and N. H. Salman, "An Improved Probability Density Function (PDF) for Face Skin Detection," *Iraqi Journal of Science*, vol. 63, no. 10, pp. 4460-4473, 2022.

[19] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat and J. Jatakia, "Human skin detection using RGB, HSV, and YCbCr color models, " *arXiv preprint arXiv:1708.02694*, 2017.

[20] M. Hajiarbabi and A. Agah, "Human skin detection in color images using deep learning," *International Journal of Computer Vision and Image processing (IJCVIP)*, vol. 5, no. 2, pp.1-13, 2015.

[21] H. K. Al-Mohair, J.M. Saleh and S.A. Suandi, "Hybrid human skin detection using neural network and K-means clustering technique, " *Applied Soft Computing*, vol. 33, pp. 337-347, 2015.

[22] Y. Luo and Y. P. Guan, "Adaptive skin detection using face location and facial structure estimation," *IET Computer Vision*, vol. 11, no. 7, pp. 550-559, 2017.

[23] C. H. Ma and H. C. Shih, "Human skin segmentation using fully convolutional neural networks," *in Proceedings of the 2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)*, IEEE, pp. 168-170, 2018.

[24] I. Hwang, Y. Kim and N.I. Cho, "Skin detection is based on multi-seed propagation in a multi-layer graph for regional and color consistency," *in Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 1273-1277, 2017.

[25] S.G. Mohammed, A.H. Majeed, E. K. Ali Aldujaili and S.S. Hassan, "Image Segmentation for Skin Detection," *Journal of Southwest Jiaotong University*, vol. 55, no. 1, 2020.