Quick Response Encoding of Human Facial Images for Identity Fraud Detection

Shweta Singh, Saheb Chhabra, Garima Gupta, Monika Gupta, Gaurav Gupta

Chapter 10

QUICK RESPONSE ENCODING OF HUMAN FACIAL IMAGES FOR IDENTITY FRAUD DETECTION

Shweta Singh, Saheb Chhabra, Garima Gupta, Monika Gupta and Gaurav Gupta

Abstract Advancements in printing and scanning technology enable fraudsters to tamper with identity documents such as identity cards, drivers' licenses, admit cards, examination hall tickets and academic transcripts. Several security features are incorporated in important identity documents to counter forgeries and verify genuineness, but these features are often lost in printed versions of the documents. At this time, a satisfactory method is not available for authenticating a person's facial image (photograph) in a printed version of a document. Typically, an official is required to check the person's image against an image stored in an online verification database, which renders the problem even more challenging.

This chapter presents an automated, low-cost and efficient method for addressing the problem. The method employs printed quick response codes corresponding to low-resolution facial images to authenticate the original and printed versions of identity documents.

Keywords: Facial images, documents, quick response codes, tamper detection

1. Introduction

Advancements in printing and scanning technology have made it easy for fraudsters to produce high-quality tampered documents. Indeed, changing or replacing the facial image of a person on an identity document is becoming very common.

Numerous frauds have been perpetrated by tampering with human facial images on documents. One example is the Vyapam scam in India [10], where a fraudster replaces the photograph of a student on an

examination admit card with that of an imposter who takes the exam on behalf of the student. Another example comes from China [15], where a broker finds a proxy to take an important exam such as the SAT, GRE or GMAT on behalf of a student. The broker then prepares a fake passport with the information of the student but with the image of the proxy.

Authenticating a tampered identity document requires an expert to manually analyze the document using sophisticated equipment such as a microscope or video spectral comparator. This process is time-consuming, inefficient and non-scalable. Also, this method for detecting tampered documents is not applicable to printed versions of documents because most security features are lost during the printing process.

Clearly, there is a need to develop an automated system that can authenticate a person's facial image on a document. This system should work for originals as well as printed versions of documents. Also, the system should be able to authenticate documents offline and without relying on a database of images. Additionally, if tampering is detected, the system should be able to reproduce the person's facial image that is similar to the original image in the document.

This chapter presents a method that employs printed quick response (QR) codes of low-resolution facial images to authenticate the original and printed versions of documents. The method supports image-to-image verification, which matches an image on an identity document against the image encoded as a quick response code on the same document. Also, it supports real-time person verification, which matches an image encoded as a quick response code on the identity document against a person's image captured in real time. Figure 1 shows an example of authenticating (and detecting the tampering of) a driver's license using the proposed method.

2. Related Work

Several researchers have proposed methods for detecting counterfeit documents. Gupta et al. [6] describe a method that considers the texture and unique color count in order to detect counterfeit documents. The method links a counterfeit document to its source scanner and printer. While the method can differentiate between the original and printed versions of documents, it is not effective at detecting tampering in printed documents.

Sarkar et al. [12] have proposed a method for detecting low-quality and high-quality counterfeit currency notes. They have also analyzed printed security features on currency notes.

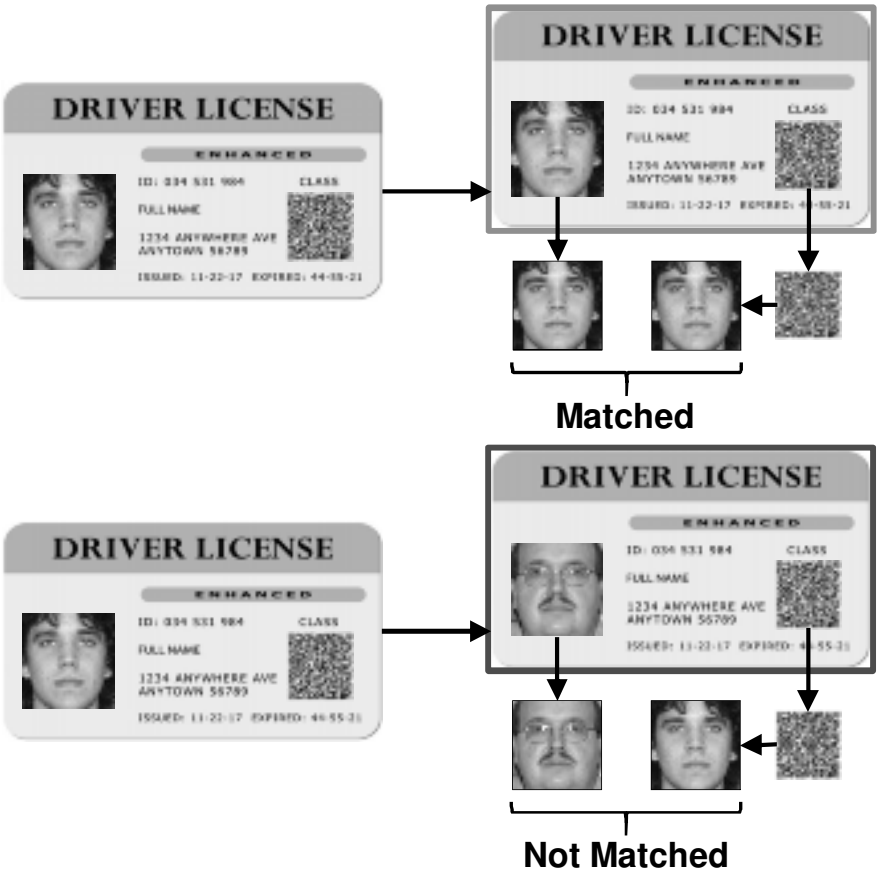

Figure 1. Authentication of a facial image on a driver's license.

Chhabra et al. [2] have developed a method for detecting fraudulent bank checks, including checks whose text has been altered using invisible ink. However, this method for alteration detection does not work well for printed documents.

The method proposed in this chapter employs quick response codes as information carriers. In the research literature, quick response codes have been used in applications ranging from physical document authentication to securing information and identifying leaked documents.

Warasart and Kuacharoen [16] have used quick response codes to authenticate text-based physical documents. Specifically, they created codes based on text in documents and verified them against codes created using text extracted from the documents by an optical character recognition system.

Espejel-Trujillo et al. [3] have employed a visual secret sharing scheme and quick response codes to authenticate documents. Their approach generates a binary encrypted message whose first part is encoded as a quick response code and printed on an identity document while the second part is encoded as a quick response code stored in a database. The authentication process extracts both the parts in order to verify the authenticity of a document.

Tkachenko et al. [14] have proposed a two-level quick response code for document authentication. The first-level code is the same as an ordinary quick response code that is readable by a typical quick response code decoder. The second private-level code contains textures and special patterns for document authentication.

Nayak et al. [8] have developed a font pixel manipulation approach for detecting the sources of leaks of hardcopy documents. A quick response code that encodes information in the font pixels of a unique document identifier is created and embedded in the document. During the detection process, this information is extracted from a leaked document to obtain its unique identifier.

Aygun and Akcay [1] have employed quick response codes to securely transmit biometric facial features used to authenticate e-government and e-passport applications. Seenivasagam and Velumani [13] have developed a method for authenticating medical images by embedding quick response codes containing patient identity details as watermarks in the images. More recently, Raval et al. [11] have proposed a method that protects the privacy of images using an adversarial perturbation mechanism and quick response codes.

Table 1 compares existing quick-response-code-based methods along with the proposed authentication method. The literature review indicates that no automated approach for authenticating a person's facial image in a document has been published previously. The proposed method addresses the deficiency by performing authentications in an offline manner. Low-resolution facial images are stored as quick response codes on documents when creating the documents. During authentication, the low-resolution image stored as a quick response code on a document is extracted and matched against a person's facial image on the document.

Two challenges are encountered when storing and verifying low-resolution images using quick response codes. First, because a quick response code has limited data storage capacity, an image is downsampled to a low resolution (e.g., 16×16 or 8×8), which leads to considerable information loss. Second, authentication requires the comparison of a low-resolution image against a high-resolution image. To address these challenges, the

Table 1. Comparison of tampering detection methods.

Authors	**Method**	**Documents**	**Facial Image Authentication**
Gupta et al., 2007 [6]	Texture and unique color count	Original	No
Chhabra et al., 2017 [2]	Texture	Original	No
Nayak et al., 2018 [8]	Font pixel manipulation	Printed	No
Espejel-Trujillo et al., 2016 [3]	Visual secret sharing	Printed	No
Singh et al., 2019 (proposed method)	Generative adversarial net	Original and printed	Yes

proposed method uses deep learning to enhance the low-resolution image that is used for authentication.

3. Proposed Method

The proposed method has two steps: (i) document generation; and (ii) document authentication.

3.1 Document Generation

Given a document with a low-resolution facial image of a person, a quick response code corresponding to the image is created. This quick response code is then printed on the document. Figure 2(a) shows the steps involved in document generation.

3.2 Document Authentication

Two common situations are encountered when attempting to authenticate a person's facial image on a document: (i) image-to-image verification; and (ii) real-time person verification.

In image-to-image verification, the low-resolution image already encoded as a quick response code on a document is matched against the person's image on the document. This is required when the person's facial image on the document must be verified for possible tampering.

In real-time person verification, the facial image already encoded as a quick response code on a document is matched against the person's facial image captured in real time. This is required when it is difficult to verify

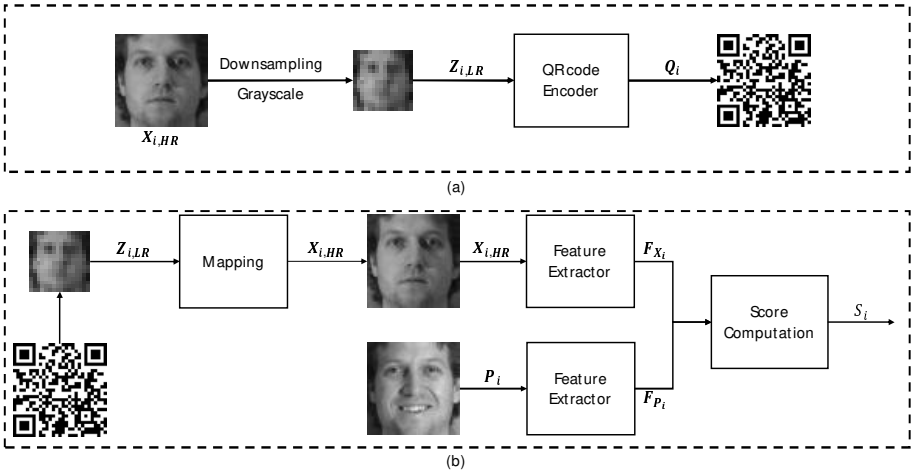

Figure 2. Proposed method.

the facial image on the document as a result of changes in the person's facial features due to age or injury or when the person's image on the document has been damaged. Figure 2(b) shows the steps involved in document authentication using the quick response code computed from a low-resolution image.

The embedded quick response code and the person's image on the document make it possible to perform the authentication process offline (i.e., without a database). The research literature demonstrates that quick response codes work well on printed documents.

The document authentication step has two components: (i) low-resolution to high-resolution image mapping; and (ii) image matching:

- **Low-Resolution to High-Resolution Image Mapping:** A generative adversarial network model [4] is employed to transform a low-resolution grayscale image encoded as a quick response code to a high-resolution color image. The approach involves training a generator model G to produce an output image x from a noise vector $\mathbf{z}$, and a discriminator D to distinguish between the real image y and a generated image x. The objective of the generator is to produce images such that the discriminator cannot distinguish between the real images and the generated images, where the discriminator has been trained to distinguish real images from generated images.

  The objective function of a generative adversarial network is given by:

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_y[logD(y)] + \mathbb{E}_z[log(1 - D(G(z)))] \quad (1)$$

where the generator model G tries to minimize the objective function and the discriminator D tries to maximize the objective function.

A conditional generative adversarial network is trained to transform the input low-resolution grayscale image to a high-resolution color image.

Let $\mathbf{X_{HR}}$ be a training set containing m high-resolution color images associated with a document. Let $\mathbf{Z_{LR}}$ be the corresponding low-resolution grayscale image set. Each image $\mathbf{Z_{i,LR}}$ in $\mathbf{Z_{LR}}$ is generated by downsampling the high-resolution color image $\mathbf{X_{i,HR}}$ using bicubic interpolation followed by grayscale conversion.

In order to learn the mapping between a low-resolution grayscale image and a high-resolution color image, it is necessary to first upsample the low-resolution grayscale image:

$$\mathbf{Z_{i,LR}} \xrightarrow[\text{Grayscale}]{\text{Upsampling}} \mathbf{Z_{i,HR}} \quad (2)$$

where, $\mathbf{Z_{i,HR}}$ is the upsampled high-resolution grayscale image.

However, the upsampled image becomes blurred due to significant information loss. Inspired by the work of Isola et al. [7], the problem is overcome using a Pix2Pix generative adversarial network to learn the mapping between the upsampled grayscale image and the high-resolution color image.

Specifically, a conditional generative adversarial network may be used to learn the mapping from input images to the corresponding output images (i.e., image-to-image translation). In this work, it is required to learn the mapping from the input upsampled grayscale image set $\mathbf{Z_{HR}}$ to the output high-resolution color image set $\mathbf{X_{HR}}$. Therefore, the objective function is written as:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{\mathbf{Z_{HR}},\mathbf{X_{HR}}}[logD(\mathbf{Z_{HR}}, \mathbf{X_{HR}})] + \mathbb{E}_{\mathbf{Z_{HR}},z}[log(1 - D(G(\mathbf{Z_{HR}}, z)))] \quad (3)$$

As mentioned above, a low-resolution image becomes blurred after upsampling. Therefore, the $L1$ norm (Manhattan distance) is employed to encode sharpness in the output images. The new objective function is given by:

$$\mathcal{L}_{cGAN}(G,D) = \mathbb{E}_{\mathbf{Z_{HR}},\mathbf{X_{HR}}}[logD(\mathbf{Z_{HR}},\mathbf{X_{HR}})]+ \mathbb{E}_{\mathbf{Z_{HR}},z}[log(1-D(G(\mathbf{Z_{HR}},z)))] + \lambda\mathcal{L}_{L1}(G) \quad (4)$$

- **Image Matching:** The verification task is to determine if an image encoded as a quick response code matches the input image associated with the document or matches the image captured in real time.

  Let $\mathbf{P_i}$ be the input image to be matched against the image encoded in the quick response code. For this purpose, the low-resolution grayscale image in the quick response code is extracted and mapped to the high-resolution color image $\mathbf{X_{i,HR}}$. Next, the two images are input to a pre-trained facial model to obtain the output facial representation vectors for the two images. The Euclidean distance between the two output facial representation vectors is computed to obtain the image matching score S_i, which is given by:

$$S_i = ||R(\mathbf{P_i}) - R(\mathbf{X_{i,HR}})||_F \quad (5)$$

  where $R(.)$ is the function that extracts facial features from the input images and $||.||_F$ is the Frobenius norm.

4. Experiments and Results

The proposed method was evaluated using the Multi-PIE 51 dataset [5]. The evaluation employed 16×16 and 8×8 low-resolution images.

The Multi-PIE 51 dataset contains 50,248 images of 337 subjects. The images of each subject have differing illuminations, poses and expressions. The dataset was partitioned into a training set and testing set with 202 (60%) subjects and 135 (40%) subjects, respectively.

Two experiments were conducted: (i) image-to-image verification; and (ii) real-time person verification. Table 2 shows the details of the the experiments, which were conducted with 16×16 and 8×8 resolution images. The Light CNN-29 [17] and VGGFace [9] facial representation models were employed to extract features from images.

In order to encode images as quick response codes, the low-resolution images were converted to unicode and then stored as quick response codes. The steps were reversed to decode the images. The Pix2Pix generative adversarial network model was trained for 150 epochs using the 16×16 and 8×8 resolution images.

Table 2. Experiments conducted with 16×16 and 8×8 resolution images.

Experiment	**Resolution**	**Training Samples**	**Testing Samples**
Image-to-Image Verification	16×16	33,613	16,635
	8×8	33,613	16,635
Real-Time Person Verification	16×16	33,613	16,635
	8×8	33,613	16,635

4.1 Performance Evaluation

The proposed method was evaluated for image-to-image verification and real-time person verification. For both the evaluations, 16,635 genuine and imposter pairs were generated using test dataset. In the case of image-to-image verification, two same images (regardless of their resolutions) were considered to be a genuine pair. In the case of real-time person verification, two images of the same subject were considered to be a genuine pair.

4.2 Image-to-Image Verification

Image-to-image verification compared images in the documents against the images encoded as quick response codes in the documents. Figures 3 and 4 show the receiver operating characteristic (ROC) curves obtained for the Light CNN-29 and VGGFace models, respectively. Note that the curves on the left-hand sides of the figures are for 16×16 resolution images whereas the curves on the right-hand sides of the figures are for 8×8 resolution images. The ROC curves in the two figures indicate that image-to-image verification yields better results with the 16×16 resolution images for both the models. Also, the Light CNN-29 model performs better with 16×16 resolution images whereas the VGGFace performs better with 8×8 resolution images.

Table 3 shows the true positive rates for three false positive rates (0.01, 0.1 and 0.2) for the Light CNN-29 and VGGFace models with the two image resolutions. The first row of the table shows the image-to-image verification results. Note that the Light CNN-29 model yields a better true positive rate with 16×16 resolution images for a false positive rate of 0.01 whereas the VGGFace model yields better true positive rates for false positive rates of 0.1 and 0.2. In the case of 8×8 resolution images, the VGGFace model yields better true positive rates for all three false positive rates.

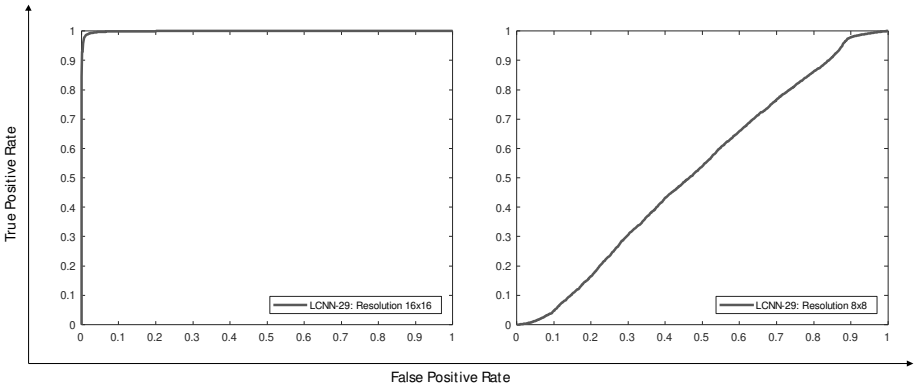

Figure 3. ROC curves for image-to-image verification (Light CNN-29 model).

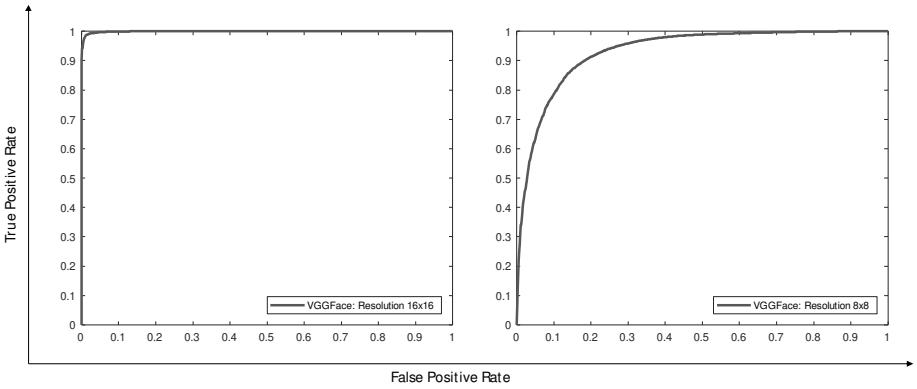

Figure 4. ROC curves for image-to-image verification (VGGFace model).

Table 3. True positive rates for false positive rates of 0.01, 0.1 and 0.2.

Experiment	Resolution	Light CNN-29			VGGFace		
		0.01	**0.1**	**0.2**	**0.01**	**0.1**	**0.2**
Image-to-Image Verification	16×16	0.9823	0.9988	0.9996	0.9818	0.9992	0.9999
	8×8	0.0461	0.0461	0.1642	0.3068	0.7857	0.9132
Real-Time Person Verification	16×16	0.0598	0.6439	0.9396	0.0538	0.4582	0.6933
	8×8	0.0087	0.0977	0.1476	0.0429	0.3070	0.4864

Figure 5 shows image samples obtained using the Pix2Pix generative adversarial network model (i.e., mapped from low-resolution grayscale images to high-resolution color images). The first row shows the original images, the second row shows the 16×16 resolution images encoded as quick response codes and the third row shows the images obtained

Figure 5. Images generated using Pix2Pix for 16×16 resolution images.

using Pix2Pix. Note that the visual appearances of subjects are almost completely preserved in the 16×16 resolution images. This demonstrates that the proposed method is able to reproduce images that are similar to the original images.

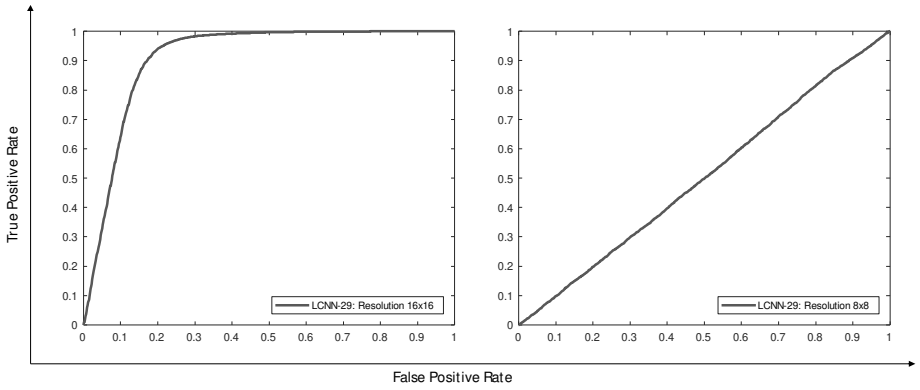

Figure 6. ROC curves for real-time person verification (Light CNN-29 model).

4.3 Real-Time Person Verification

Real-time person verification compared images encoded as quick response codes in the documents against persons' images captured in real time. Figures 6 and 7 show the receiver operating characteristic curves obtained for the Light CNN-29 and VGGFace models, respectively. Once again, the curves on the left-hand sides of the figures are for 16×16 resolution images whereas the curves on the right-hand sides of the figures

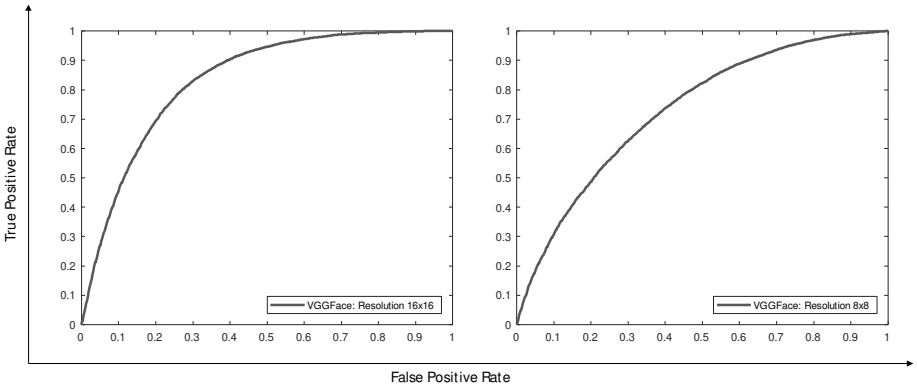

Figure 7. ROC curves for real-time person verification (VGGFace model).

are for 8×8 resolution images. The ROC curves reveal significant drops in real-time person verification compared with image-to-image verification – these are due to variations in illumination, poses and expressions in the images. The ROC curves for both models indicate that real-time person verification is better for 16×16 resolution images. Furthermore, the Light CNN-29 model performs better with 16×16 resolution images whereas the VGGFace performs better with 8×8 resolution images.

Table 3 also shows the true positive rates for three false positive rates (0.01, 0.1 and 0.2) for real-time person verification with the Light CNN-29 and VGGFace models for the two image resolutions. The second row of the table shows the real-time person verification results.

Figure 8 shows image samples obtained using the Pix2Pix generative adversarial network model (i.e., mapped from low-resolution grayscale images to high-resolution color images). The first row shows the original images, the second row shows the 8×8 resolution images encoded as quick response codes and the third row shows the images obtained using Pix2Pix. Note that the visual appearances of the subjects are almost completely distorted in all the images. This demonstrates that the 8×8 resolution images are not suitable for real-time person verification.

5. Conclusions

The availability of sophisticated, yet inexpensive, printing and scanning equipment makes it easy for fraudsters to tamper with facial images on important documents such as identity cards, drivers' licenses, admit cards, examination hall tickets and academic transcripts. Several solutions have been developed to verify the authenticity of identity documents, but they are usually limited to verifying document original-

Figure 8. Images generated using Pix2Pix for 8×8 resolution images.

ity. Moreover, their underlying techniques often fail when applied to printed versions of identity documents. Some researchers have developed techniques for detecting the tampering of text in photocopied and printed versions of identity documents, but they do not verify the facial identities of persons from the documents.

This chapter has presented an effective and low-cost solution for verifying facial images from original and printed versions of identity documents. A person's facial image is converted to a quick response code that is embedded in an identity document during its creation. In image-to-image verification, the image on an identity document is compared against the image encoded as a quick response code on the document. In real-time person verification, the image encoded as a quick response code on an identity document is compared against the person's facial image captured in real time.

Future work will implement the proposed method in real-world environments. Also, research will focus on detecting fraudulent identity documents where the facial features of two persons are amalgamated to create facial images.

References

[1] S. Aygun and M. Akcay, Securing biometric face images via steganography for QR code, *Proceedings of the Eighth International Conference on Information Security and Cryptology*, pp. 128–133, 2015.

[2] S. Chhabra, G. Gupta, M. Gupta and G. Gupta, Detecting fraudulent bank checks, in *Advances in Digital Forensics XIII*, G. Peterson and S. Shenoi (Eds.), Springer, Cham, Switzerland, pp. 245–266, 2017.

[3] A. Espejel-Trujillo, I. Castillo-Camacho, M. Nakano-Miyatake and H. Perez-Meana, Identity document authentication based on VSS and QR codes, *Procedia Technology*, vol. 3, pp. 241–250, 2012.

[4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville and Y. Bengio, Generative adversarial nets, *Proceedings of the Twenty-Seventh Annual Conference on Neural Information Processing Systems*, pp. 2672–2680, 2014.

[5] R. Gross, I. Matthews, J. Cohn, T. Kanade and S. Baker, Multi-PIE, *Image and Vision Computing*, vol. 28(5), pp. 807–813, 2010.

[6] G. Gupta, S. Saha, S. Chakraborty and C. Mazumdar, Document frauds: Identification and linking fake documents to scanners and printers, *Proceedings of the International Conference on Computing: Theory and Applications*, pp. 497–501, 2007.

[7] P. Isola, J. Zhu, T. Zhou and A. Efros, Image-to-Image Translation with Conditional Adversarial Networks, arXiv:1611.07004 (`arxiv.org/abs/1611.07004`), 2018.

[8] J. Nayak, S. Singh, S. Chhabra, G. Gupta, M. Gupta and G. Gupta, Detecting data leakage from hard copy documents, in *Advances in Digital Forensics XIV*, G. Peterson and S. Shenoi (Eds.), Springer, Cham, Switzerland, pp. 111–124, 2018.

[9] O. Parkhi, A. Vedaldi and A. Zisserman, Deep face recognition, *Proceedings of the British Machine Vision Conference*, pp. 41.1–41.12, 2015.

[10] A. Rai, I have been asked to shut my mouth, but work will go on – An interview with the whistleblower who exposed Madhya Pradesh Vyapam scam, *The News Minute*, February 25, 2015.

[11] N. Raval, A. Machanavajjhala and L. Cox, Protecting visual secrets using adversarial nets, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1329–1332, 2017.

[12] S. Sarkar, R. Verma and G. Gupta, Detecting counterfeit currency and identifying its source, in *Advances in Digital Forensics IX*, G. Peterson and S. Shenoi (Eds.), Springer, Berlin Heidelberg, Germany, pp. 367–384, 2013.

[13] V. Seenivasagam and R. Velumani, A QR code based zero-watermarking scheme for authentication of medical images in teleradiology cloud, *Computational and Mathematical Methods in Medicine*, article no. 516465, 2013.

[14] I. Tkachenko, W. Puech, C. Destruel, O. Strauss, J. Gaudin and C. Guichard, Two-level QR code for private message sharing and document authentication, *IEEE Transactions on Information Forensics and Security*, vol. 11(3), pp. 571–583, 2016.

[15] P. Tyre, How sophisticated test scams from China are making their way into the U.S., *The Atlantic*, March 21, 2016.

[16] M. Warasart and P. Kuacharoen, Paper-based document authentication using digital signature and QR code, *Proceedings of the Fourth International Conference on Computer Engineering and Technology*, pp. 94–98, 2012.

[17] X. Wu, R. He, Z. Sun and T. Tan, A light CNN for deep face representation with noisy labels, *IEEE Transactions on Information Forensics and Security*, vol. 13(11), pp. 2884–2896, 2018.