A SIMPLE APPROACH TO DIGITAL ID CARD IMAGE SECURITY

Hüseyin Bilal MACİT, Burdur Mehmet Akif Ersoy University, hbmacit@mehmetakif.edu.tr

INTRODUCTION

An identity card is used as a document proving who a person is. Turkish identity cards were used for the first time during the Ottoman Empire under the name of "Devleti Aliyye-i Osmaniyye Tezkiresi". These identity cards were distributed by the " Nüfus-u Umumiye Müdüriyeti" established in 1884 (Kubilay et al., 2010). The information on this ID included the person's name, mother's name, father's name, nationality, profession, eye color and height. Over time, ID cards took the form of notebooks. Until 1926, these notebooks were kept in the old alphabet. They have been written in the new alphabet since 1926. From 1928 onwards, photographs were also included in the ID notebooks to the extent of the possibilities of the period. As of January 01, 1976, it took the form of a single-page document and was called "TC Nüfus Cüzdanı" (https://www.nvi.gov.tr/tc-kimlik-karti). Figure 1 shows examples of old identity cards used in Turkey and the Ottoman Empire (https://tr.wikipedia.org/wiki/Türkiye_Cumhuriyeti_Kimlik_Kartı).

(a) (b) (c) (d)

Figure 1. (a) Devlet-i Aliyye-i Osmaniyye Tezkiresi (b) Old Alphabet "Nüfus Defteri" Including Photo (c) New Alphabet, Notebook Style "Nüfus Hüviyet Cüzdanı" (d) Gender Based "TC Nüfus Cüzdanı"

Recent technological developments have required the replacement of existing ID cards. The Republic of Turkey Identity Card (TCKK) Project has recently been prepared in order to use identity cards more actively in state affairs and to facilitate the work of citizens. With the TCKK project, a smart ID card was designed, which has a validity of 10 years, contains a contactless chip, complies with international standards and includes advanced security elements. Identity registration information, photo, signature and various biometric data belonging to the owner of the smart ID card are securely recorded on the chip on the smart card. Since 02 January 2017, only TCKK has been given to every citizen who wants to get a new identity card or renew the old one. Figure 2 shows the sample TCKK image shown by the Turkish Population and Citizenship Administration (https://www.nvi.gov.tr/tc-kimlik-karti).

(a) (b)

Figure 2. (a) Sample TCKK Front Side (b) Sample TCKK Back Side

TCKK is the same size as credit cards. It has many visual and electronic security elements on it to make it difficult to produce fake ID cards. It has multi-factor electronic verification technologies such as PIN, biometrics and certificate. Electronic signature can be loaded into it. In addition, it complies with many international standards such as ISO-7816, ISO-14443 and ICAO-9303 (https://tr.wikipedia.org/wiki/Türkiye_Cumhuriyet_Kimlik_Kartı). Similar to Turkey, smart ID cards are used in many countries of the world. Indonesia is the first country to use a smart ID card. Example smart ID cards used in different countries of the world are shown in Figure 3 (Castro, 2011)(https://www.bsi.bund.de) (https://en.wikipedia.org/wiki/Smart_card).

(a) **(b)** **(c)**

(d) **(e)** **(f)**

Figure 3. Smart ID Cards Used In (a) Estonia (b) Portugal (c) Belgium (d) Germany (e) Spain (f) Finland

The security elements on the TCKK are shown in Figure 4 (https://www.nvi.gov.tr/tc-kimlik-karti). Most of these safety elements can be distinguished by the Human Visual System (HVS). In this way, it may be easily possible to distinguish whether an ID card is real or fake. Optical Variable Ink (OVI), in particular, is a difficult technology to imitate. It is an iridescent foil with an image, such as a hologram. OVI can vary in size and shape, as well as the design of the interior image, they can also have various effects such as relief, mini texts, colored areas, matt areas, black or gold areas. (Taşçıoğlu et al., 2017).

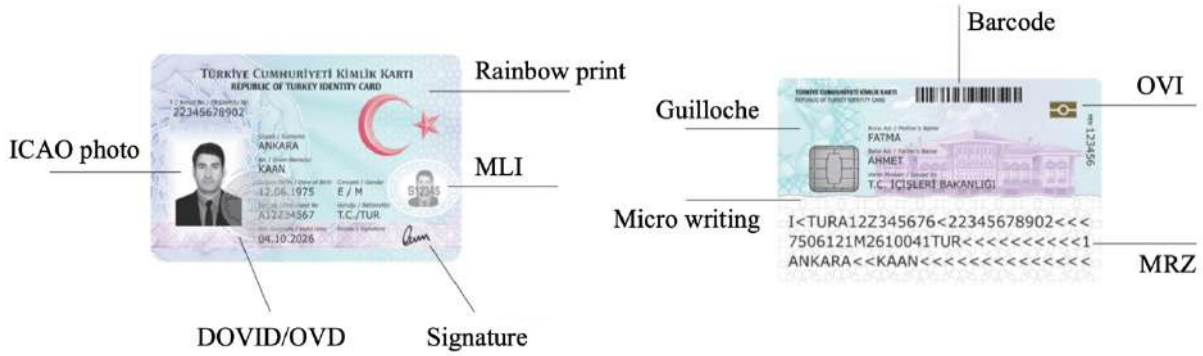

Figure 4. TCKK Visual Security Elements

OVI provides good security on physical media, but it is mostly invisible when the smart ID card is scanned to digital media. Smart ID cards are transferred to digital media and stored in many areas such as banking, education and health. Different technologies are used to ensure the security of stored ID card images. If the

storage medium is accessible to the internet, hackers can access these identities and modify the identity data. Fake ID images can be used in illegal activities that may cause property damage and forensic cases. In this study, a self-security method is proposed for the digitized TCKK.

METHOD

The proposed method consists of two stages. The first stage is the "save stage", which includes the algorithm of saving the TCKK image for the first time. The second step is the "load stage", which includes the steps of reading the image from the database and confirming its authenticity (Figure 5).

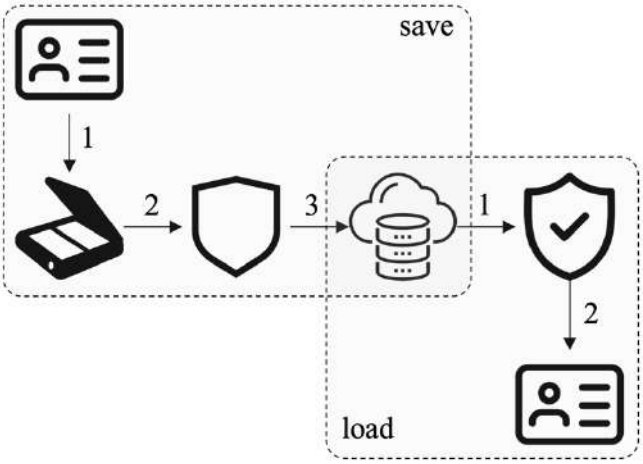

Figure 5. Save And Load Stages Of Proposed Method

In the continuation of the article, the TCKK images used for the demonstration of the proposed method are the sample images given on the official website of the TC Population and Citizenship Directorate and Wikipedia. They are not identity card images of any real people.

Save Stage

Hackers can manipulate important information such as photo, name, TC Identification number, date of birth of the ID card in the digital environment. The security method proposed in this study is based on matching the text information in the TCKK image with the photograph on it. In the TCKK front view, cardholder information is placed in standard places. According to the measurements made in this study, TC ID Number, Name, Surname and Date of Birth information are in the rectangles obtained with the coordinate factors shown in Table 1 on the 2-dimensional plane.

Table 1. 2D Coordinate Factors Of The Data To Be Used On The TCKK Image

Data	X axis		Y axis	
	Start point	End point	Start point	End point
T.C. ID No	0,039	0,246	0,325	0,375
Name	0,344	0,984	0,453	0,515
Surname	0,344	0,984	0,325	0,375
Date of Birth	0,344	0,541	0,593	0,656
Serial	0,344	0,541	0,734	0,797

When the coordinate factors in Table 1 are multiplied by corresponding dimensions of scanned TCKK image, the yellow-colored areas in Figure 6 are selected as information areas. All information in yellow-colored areas are recognized as text using a standard Optical Character Recognition (OCR) function. Each read text is added to the string named "data".

Figure 6. Display Of The Location Of The Data Recognized By OCR

TCKK owner information string created with the sample TCKK image is: “12345678902HAKANACAR220519757505297F”. The resulting string consists of "variable width encoding" characters. Since the proposed method works at the bit level, in the next step the information is converted to a binary string. For this, the 9-bit equivalent of each character in “data” string is read in sequence and added to the “ArrBin” array. The size of the ArrBin array is 9 times the “data” string. In the next step, the photo in the TCKK image is extracted. Photo size is based on many different factors, such as hair length, physical size, shooting distance of the photo. For this reason, determining the location of the photo by coordinates may cause extra unnecessary areas to be selected. There are numerous algorithms for distinguishing the face image on an image. Viola Jones algorithm is used in this study. Viola-Jones face detection method is the first framework based on object detection that provides good detection rates in real time by Paul Viola & Michael Jones in 2001 (Viola & Jones, 2001). Viola - Jones algorithm includes 3 techniques for the detection of facial areas; the Haar like features for the feature extraction, Ada boost machine-learning method for detecting and Cascade classifier used to combine many of the features efficiently (Vikram & Padmavathi, 2017). In this study, the algorithm is implemented in Matlab software using the vision.CascadeObjectDetector() method. The face area detected in the sample TCKK image with the Viola Jones algorithm is shown as yellow-colored in Figure 7.

Figure 7. Detected Face Area By Viola Jones Algorithm

Thus, face area and a binary array containing the personal character data of the ID card is obtained. We used image steganography to combine these two different types of data. Image Steganography is the process of hiding information which can be text, image or video inside a cover image (Subramanian et al., 2020). Steganography can be applied to an image in either the spatial domain or the frequency domain. In this study, we applied steganography in spatial domain of the extracted area. A color digital image of size mxn

consists of 3 color layers in the form of Red-Green-Blue (RGB), that is, it can be mathematically expressed as 3 matrices of size mxn. Each cell of the matrices represents a decimal number in the range 0-255, which expresses the color intensity level of that pixel. That is, each cell is represented by 8 bits. The rightmost bit of these 8 bits is the Least Significant Bit (LSB) for color intensity. Changing the LSB does not cause any significant changes to the image. This method of hiding a bit of confidential data is called LSB steganography. At this stage, we replace the ArrBin array we created earlier with the LSB bits of the 3 matrices sequentially.

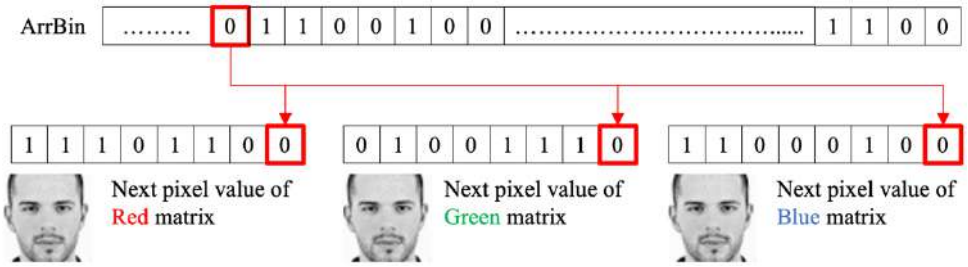

Figure 8. Embedding Of ArrBin To Each Color Layer

The length of the ArrBin binary array will always be less than mxn. By the end of the ArrBin array, not all pixels of the face image have been processed. All pixels must be processed for good security. Therefore, the LSB steganography process starts again from the first element of the ArrBin array until it reaches the cell showing the last pixel (P_(m,n)) and continues until the last element. The save stage is finished when the elements of the ArrBin array are written to the LSBs of all pixels of the face area. Now the TCKK digital image can be safely stored in the database.

Save stage algorithm

```
I ←TCKK Front
face ← vision.CascadeObjectDetector()
location ← step(face,I)
face ← I(location(x,y))
TCNo ← ocrResult.coordinate(0.039m, 0.984m, 0.325n, 0.375n) ←append(data)
name ← ocrResult.coordinate(0.344m, 0.246m, 0.453n, 0.515n) ←append(data)
surname ← ocrResult.coordinate(0.344m, 0.246m, 0.325n, 0.375n) ←append(data)
dob ← ocrResult.coordinate(0.344m, 0.541m, 0.593n, 0.656n) ←append(data)
serial ← ocrResult.coordinate(0.344m, 0.541m, 0.734n, 0.797n) ←append(data)
ArrBin ← charToBinary(data)
R, G, B ← face(:,:,:)
pointer = 1
for i←location(x)
   for j←location(y)
      LSB(R(i,j)) ←ArrBin(pointer)
      LSB(G(i,j)) ←ArrBin(pointer)
      LSB(B(i,j)) ←ArrBin(pointer)
      pointer++
      if endOf(ArrBin)
         pointer = 1
Merge(R,G,B)
```

Load Stage

The load process of the proposed method is applied to check whether there is an attack on the TCKK image stored in the database or cloud. Part of the load stage algorithm is the same as the save stage. The TC ID Number, Name, Surname and Date of Birth information on the image read by OCR methods in the save stage is written to the “data" string. The 9-bit equivalent of each element of the “data” string is written into the “ArrBin” array. Again, the vision.CascadeObjectDetector() method is used to obtain the face area of the TCKK image. The CheckArrBin binary array is created to transfer the hidden data in the LSBs of the face area. The face area is split into R, G, and B layer matrices. These matrices are processed simultaneously to calculate average of overlapping cells. These average results are floored and added to next CheckBinArray array cell as shown in Figure 9.

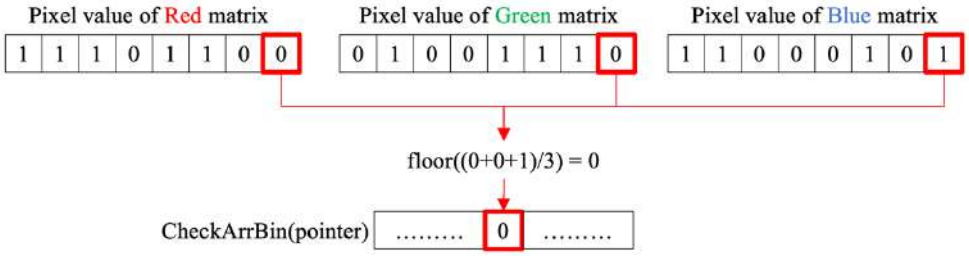

Figure 9. Construction Of CheckArrBin

If the personal data hidden in the LSBs of the face area and the data read by OCR function are the same, it means that the originality of the image has been verified. If the image has been subjected to processes such as lossy compression, there may have been changes in pixel values. In this case, an original image can be perceived as fake. In the load process, we interpret whether the image is original or not by calculating the percentage according to the value of the variable named ProbOrg. For this, ArrBin array and CheckArrBin array are compared. But the CheckArrBin array is much larger in size. Therefore, the ArrBin array is hovered over the CheckArrBin array and each element of the array is compared in order. For bits that are equal to each other, the value of the ProbOrg variable is incremented by 1. When the compare process is finished, the probability that the TCKK digital image is original is calculated as ProbOrg/sizeOf(CheckArrBin).

Load stage algorithm

```
I ←TCKK Front
face ← vision.CascadeObjectDetector()
location ← step(face,I)
face ← I(location(x,y))
TCNo ← ocrResult.coordinate(0.039m, 0.984m, 0.325n, 0.375n) ←append(data)
name ← ocrResult.coordinate(0.344m, 0.246m, 0.453n, 0.515n) ←append(data)
surname ← ocrResult.coordinate(0.344m, 0.246m, 0.325n, 0.375n) ←append(data)
dob ← ocrResult.coordinate(0.344m, 0.541m, 0.593n, 0.656n) ←append(data)
serial ← ocrResult.coordinate(0.344m, 0.541m, 0.734n, 0.797n) ←append(data)
ArrBin ← charToBinary(data)
R, G, B ← face(:,:,:)
pointer = 1
```

```
for i←location(x)
   for j←location(y)
      CheckArrBin(pointer) ←floor((LSB(R(i,j))+ LSB(G(i,j))+ LSB(B(i,j)))/3)
      pointer++
while (!endOf(CheckArrBin))
   for pointer←sizeOfArrBin
      if ArrBin(pointer)=CheckArrBin(pointer) ProbOrg++
print(ProbOrg/sizeOf(CheckArrBin))
```

RESULTS

We got the sample TCKK images we used for the experimental study from the official website of the Turkish Republic Population and Citizenship Directorate and Wikipedia. After processing the sample image with the save stage, we stored it in Portable Network Graphics (PNG) format with lossless compression. In the absence of any manipulation, the load algorithm showed 100% probability that the image was original. We used a graph editing software to calculate the accuracy of the method. After the save stage, we applied lossy jpeg compression to the image. In this case, the load stage algorithm calculated the ProbOrg value of 94.1%. After that, we cut off the TC ID number field in the image and moved it from another TCKK image. In this case, we calculated the ProbOrg value as 30.56%. When we changed the photo on the sample image, the ProbOrg value is calculated as 7.3% (Table 2).

Table 2. Forgery Experiments And Results

Forgery type	Not forged .png	Jpeg compression	Text manipulation	Photo manipulation
TCKK digital image				
Result	100%	94.1%	30.56%	7.3%

The results shown in Table 2 clearly show that the proposed algorithm ensures that the TCKK digital image is safely stored in the database. In addition, low processing overhead and high accuracy are achieved.

REFERENCES

Castro, D. (2011). Explaining International Leadership: *Electronic Identification Systems*

Deutschland Federal Office for Information Security – Retrieved from https://www.bsi.bund.de

Kubilay, M.A., Adalıer, O. & Karademir, A. (2010). Türkiye'nin e-kimlik yolculuğu, *Tübitak UEKAE*, Vol.2(4), pp.6-25.

TC General Directorate of Population and Citizenship Affairs, Retrieved from https://www.nvi.gov.tr/tc-kimlik-karti

Retrieved from https://tr.wikipedia.org/wiki/Türkiye_Cumhuriyeti_Kimlik_Kartı

Retrieved from https://en.wikipedia.org/wiki/Smart_card

Subramanian, N., Al-Maadeed, S. & Bouridane, A. (2020). Image Steganography: A Review of the Recent Advances, *Digital Object Identifier*, Vol 9.

Taşçıoğlu, M., Yıldız, C. & Erdoğan Aydın, D. (2017), Designing University Diplomas, *Anadolu University Journal of Art and Design*, Vol.7(1), ISSN:2146-7692, pp.91-103.

Vikram, K & Padmavathi S. (2017). Facial Parts Detection Using Viola Jones Algorithm, *International Conference on Advanced Computing and Communication Systems*, India.

Viola, P. & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features, *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, DOI: 10.1109/CVPR.2001.990517.