Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 7, S. 25-32, 5, 2023
© Telif hakkı IJANSER'e aittir
Araştırma Makalesi

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 7, pp. 25-32, 5, 2023
Copyright © 2023 IJANSER
Research Article

Examination of technologies that can be used for the development of an identity verification application

Dávid Fekete*, Pál Bárkányi [2]

[1]OmegaCode Hungary Software Development Ltd., Hungary
[2] Department of Informatics, Milton Friedman University, Hungary

(fekete.david@omegacode.hu Email of the corresponding author

(Received: 05 June 2023, Accepted: 20 June 2023)

(1st International Conference on Pioneer and Innovative Studies ICPIS 2023, June 5-7, 2023)

ATIF/REFERENCE: Fekete, D. & Bárkányi, P. (2023). Examination of technologies that can be used for the development of an identity verification application. *International Journal of Advanced Natural Sciences and Engineering Researches, 7(5), 25-32.*

Abstract – Artificial intelligence is becoming increasingly significant in online education, image processing, identity verification, and document integrity checking. We designed an approach for identifying the optimal method and tools for speeding and improving AI-assisted image processing. This paper compares appropriate, practical concepts to help you choose. Our article studies various AI models to recognise data by processing various photos, enhancing accuracy, minimising administrative time, and decreasing the likelihood of misreading and mistyping. In our paper, we presented character recognition and face identification based on deep learning through processing a personal identification document. We divided the process into parts, taught different artificial intelligence models, produced the data necessary for teaching, and then integrated it into our developed environment. The completed software, which can be adapted to any system in the form of an application, uses the image on the ID card and the device's camera to determine whether the application is being used by the person authorised to administer it and can optionally match the signature on the ID card with a digital sample requested by the application from its user.

Keywords – Artificial Intelligence; Image Processing; Comparison of Models; Identity Verification; Document Integrity

I. INTRODUCTION

Methods supporting online education are becoming more and more widespread in education. The spread of these learning support systems was written by Annuš et al. (2023) [1], among which solutions using artificial intelligence are playing an increasingly important role [2]. In addition to online education, it is increasingly important to support remote assessment, the limitation of which is the clear verification of exam eligibility and identity [3]. Therefore, we looked for a solution to simplify and make handling personal data and identity verification more efficient with the help of Artificial Intelligence.

We aimed to produce a program that can be used for general purposes, not only limited to identifying documents suitable for personal identification but also capable of handling other documents.

In this paper, we present what possible solutions we considered and how the practical solutions were selected.

II. MATERIALS AND METHOD

Online identity verification is not a new problem; we were not the first to think of a solution to the problem. "Know your customer" (KYC) are policies and regulations in financial services that require professionals to verify the identity, suitability and risks of maintaining a business relationship regarding customers. [4] The procedures against money laundering (Anti Money Laundering - AML) and the regulation against the financing of terrorism (Counter Terrorism Financing - CFT) fit into the broader scope.

KYC procedures usually involve collecting and verifying certain types of information, such as name, address, date of birth, and various identification documents. The information collected is used to assess the customer's risk level and ensure they do not engage in illegal activities. [5]

KYC requirements may vary depending on the country and the financial institution or business type involved. In some cases, KYC is required by law, while in others, it may be a voluntary best practice.

From the above, it is clear that this type of service is more for financial institutions and is used to check customers.

Illési and Honfi (2022) [6] also drew attention to the security risks of artificial intelligence solutions. Thus, when searching for identity verification methods, we must also consider the related dangers and weaknesses and try to avoid them. Our aim is to create a solution that is easy for users to use, pays attention to basic security procedures and can be operated sustainably [7].

When creating the system to be developed, we assume that the higher education students using it are aware of the dangers of cyberspace [8]. For identification to occur in the online examination system, we must be able to identify the facial image and the textual data on the presented ID card, read the textual data, and compare the image on the ID card and the user's face image displayed on the camera. For this, we need to select the appropriate image processing procedures, which can be used to process the scanned ID card data and separate the textual data and photographs. After that, the text data is processed, for which an optical character recognition solution must be selected and applied, and then face identification and a comparison of the ID image and the camera image must be performed. In the following, we will analyse these possible solutions, examining the machine learning solutions that can be used separately [9].

A. *Image processing using artificial intelligence*

With AI technology, new dimensions of image processing have opened up, thanks to which many new solutions could be developed, for example, in medical imaging, autonomous driving or even facial recognition. Artificial intelligence algorithms such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN) or Generative Adversarial Networks (GAN) are effective image-processing tools that can be used for many tasks. Such tasks include image segmentation, object recognition, image correction, image restoration, or image generation [10].

Among the possible solutions, we must highlight the following:

Artificial Neural Networks (ANN) are computer processing systems largely inspired by the functioning of biological nervous systems (such as the human brain).

Convolutional Neural Network: CNN is a deep learning algorithm that can take an input image, assign “importance” (learnable weights and biases) to different aspects/objects of the image, and distinguish one from another. In the case of CNN, the required preprocessing is much smaller than in the case of other classification algorithms. The essence of the convolutional neural network is to reduce the images to an easier-to-process form without using the so-called information, and functions required for prediction would be lost.

Table 1. Comparing the different Neural Networks

Feature	**ANN**	**CNN**	**RNN**
Data type	Tabular and text data	Image data	Different series
Advantages	Can be operated with little knowledge, good fault tolerance	Accurate in image recognition	Memory management and self-learning
Disadvantages	Hardware requirement	The need for a huge learner data set	A slow and complex learning process
Area of use	Complex problem solving, such as predictive analytics	Machine vision including image recognition	Natural language processing including sentiment analysis and voice recognition

Recurrent Neural Networks (RNN) are unique because of their ability to process both past data and input data — and memorise things — and were developed to overcome the weaknesses of the feed-forward network [11].

The main properties of ANN, CNN and RNN solutions are compared in Table 1.

B. *Optical Character Recognition*

Optical Character Recognition (OCR) is a technology that converts printed texts into machine-readable digital text. The technology is widely used, from the digitisation of documents to the automation of data entry. The parameterisation and training of document recognition artificial intelligence presents many challenges. Each ID contains different fonts, there are differences in size and style, and different symbols and images make recognition difficult. Last but not least, the background is also designed to be difficult to fake. Thus, we face an armada of colours and shapes when we embark on such a project.

Based on the preliminary research, we found four approaches suitable for recognising personal documents. These approaches are the follows:

1) Template-based Matching: In this approach, we create a document template and compare the image of the scanned (photographed) document to this for character recognition. This approach can be effective for fixed-format documents such as a driver's license. To give good results, it is essential to create a well-designed template.
2) Feature Extraction: This approach involves the extraction of certain characteristics of the documents, such as the edges and corners of the card, to find and recognise characters. Feature extraction can be effective for documents of different formats.
3) Machine Learning-based Method: Deep learning methods like CNN are good for document identification and processing. These methods require a large amount of learning data, but they can accurately recognise characters on documents. Transfer learning, which involves fine-tuning pre-trained models on a given ID dataset, can effectively reduce the required training data.
4) Image preprocessing: This is not a separate approach since no character recognition takes place in this case. However, image preprocessing techniques (such as noise reduction or enhancement, contrast adjustment, and other image manipulation) can improve the image quality of documents to improve OCR accuracy.

If appropriate, more than one of the listed approaches must be used simultaneously.

Recurrent Neural Networks (RNN) are unique because of their ability to process both past data and input data — and memorise things — and were developed to overcome the weaknesses of the feed-forward network [11].

The main properties of ANN, CNN and RNN solutions are compared in Table 1.

C. Optical Character Recognition

Optical Character Recognition (OCR) is a technology that converts printed texts into machine-readable digital text. The technology is widely used, from the digitisation of documents to the automation of data entry. The parameterisation and training of document recognition artificial intelligence presents many challenges. Each ID contains different fonts, there are differences in size and style, and different symbols and images make recognition difficult. Last but not least, the background is also designed to be difficult to fake. Thus, we face an armada of colours and shapes when we embark on such a project.

Based on the preliminary research, we found four approaches suitable for recognising personal documents. These approaches are the follows:

5) Template-based Matching: In this approach, we create a document template and compare the image of the scanned (photographed) document to this for character recognition. This approach can be effective for fixed-format documents such as a driver's license. To give good results, it is essential to create a well-designed template.
6) Feature Extraction: This approach involves the extraction of certain characteristics of the documents, such as the edges and corners of the card, to find and recognise characters. Feature extraction can be effective for documents of different formats.
7) Machine Learning-based Method: Deep learning methods like CNN are good for document identification and processing. These methods require a large amount of learning data, but they can accurately recognise characters on documents. Transfer learning, which involves fine-tuning pre-trained models on a given ID dataset, can effectively reduce the required training data.
8) Image preprocessing: This is not a separate approach since no character recognition takes place in this case. However, image preprocessing techniques (such as noise reduction or enhancement, contrast adjustment, and other image manipulation) can improve the image quality of documents to improve OCR accuracy.

If appropriate, more than one of the listed approaches must be used simultaneously.

D. Face Identification and Comparison

During the processing of personal identification documents, in addition to character recognition, another important step is to compare the photo on the ID with the face of the person who identifies. Many factors can affect the comparison between the image on the document and the face captured by the camera, such as differences in lighting, pose, and facial expression. To ensure accurate results, high-quality images and well-chosen algorithms are essential aspects [12].

The face comparison process can be broken down into the following steps:

1) Recognition and extraction of the ID photo on the document.
2) Recognition and extraction of the face in the selfie image taken with the camera.
3) Comparison of the two portraits.

We have several models available for this, the most popular of which are available:

a. FaceNet: a deep neural network developed by Google that maps facial features into a multidimensional space, enabling accurate face comparison and recognition. [13]
b. OpenFace: a neural network developed by Carnegie Mellon University researchers, capable of performing high-accuracy face recognition and comparison tasks. [14]
c. DeepFace: a convolutional neural network developed by Facebook AI researchers, which can perform facial recognition and comparison with high accuracy. [15]
d. VGGFace: a deep neural network developed by the Visual Geometry Group of the University of Oxford, which can also identify and compare faces with high accuracy. [16]
e. ArcFace: the deep neural network developed by the Chinese company Megvii, which is capable of performing face recognition and comparison tasks with high accuracy despite changing light conditions and pose changes. [17]
f. Facenet-pytorch: open-source solution that is part of PyTorch. [18]
g. Dlib Face Recognition: a popular open-source facial recognition library that uses deep learning algorithms. [19]

III. RESULTS

During the execution, the software processes must be planned, which are the following:

1. The user uploads or uses the camera to take a picture of their identity card and address card according to the program's instructions.
2. The program starts processing the uploaded images according to the defined image pre-processing processes.
3. An artificial intelligence model (object detection model) finds the exact locations of the texts found in the image and the location of the identity card image.
4. Another artificial intelligence model, Natural Language Processing (NLP), identifies the texts in marked places.
5. If the text recognition and ID card image extraction are unsuccessful, the process returns to step 1 and asks the user to try to change the lighting conditions, possibly the position of the document, and take a new image.
6. In case of successful identification, the program asks the user to look into the device's camera while it takes a picture for face identification.
7. The software compares the created image with the ID image extracted from the document.
8. If the two images do not match, it returns to step 6 and asks the user to try to take a picture from a different angle under more optimal lighting conditions.
9. In case of an acceptable match, an information screen with the user's data is displayed.

A. *Definition of Artificial Intelligence Models*

The steps in the process that require artificial intelligence models:

1. Accurate determination of the document's location (object detection). The first step is to identify the document itself in the photograph and determine its exact location. Several object recognition models are available for this, the most popular of which are available:

a) YOLO family (YOLOv3 to YOLOv8): high-accuracy real-time object detection model. [20]

b) SSD: popular for its speed, accuracy and reliability. [21]

c) Faster R-CNN: robust and accurate object detection model, more computationally demanding than the ones listed, but promising a more accurate result in return. [22]

The preliminary comparison is difficult since there are no reports on how the individual models perform, so we trained the models on our generated data set until we found the optimal one. The dataset consisted of 1,000 images in the first round and 5,000 in the second round. In the case of YOLO, we used version 5, including Yolov5s. We divided the teaching and validation datasets in the ratio of 80%-20%, which means that for 1000 images, we used 800 teachings and 200 validation images. For 5000 images, we used 4000 teaching and 1000 validation images. The model performs well, as indicated by very low box_loss and obj_loss values, perfect P and R values, and high mAP50 and mAP50-95 values. The box_loss and obj loss values for the entire training process are shown in Figures 1 and 2.

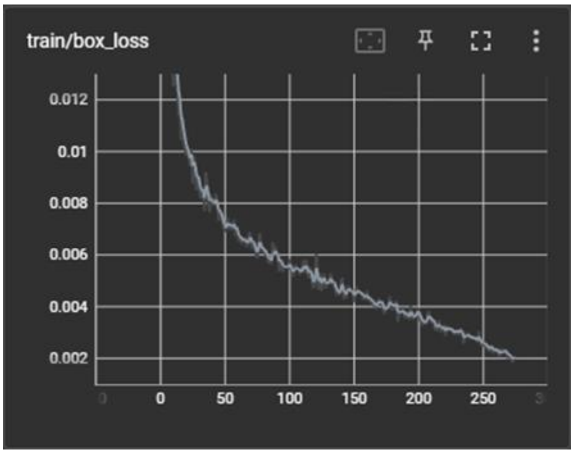

Fig. 1 Evolution of box_loss

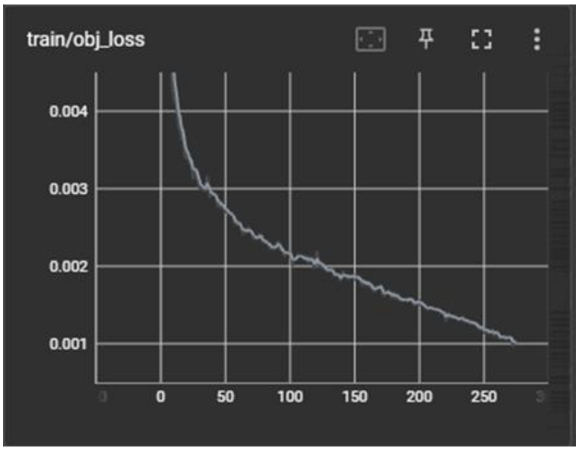

Fig. 2 Evolution of obj_loss

2. Image processing: removing the background, rotating the ID horizontally, and cropping the document (image preprocessing). The library uses built-in AI models for various image

transformations; these are trained, well-functioning models, and we just have to use them.

a) Crop the image based on the coordinates of the bounding box defined by the YOLO model.
b) We look for the contours of the image and try to level the image as best as possible based on them.

3. Finding and cropping a passport photo (face detection). There are many ways to identify the ID photo. We can choose from several pre-taught models that need not be finely tuned. We will use the "dlib face recognition" library for the solution because it is the easiest to implement based on preliminary surveys and fully meets expectations. It is important to note that the library is not necessarily suitable for processing large real-time images or videos, as it is not optimised for speed or memory management. It is recommended to use models with special hardware acceleration for these tasks. Both its speed and resource requirements are suitable for the task. It has not one but two built-in procedures for face identification:

a) HOG (Histogram of Oriented Gradients) based face recognition focuses on capturing colour gradients within an image to represent the object's shape. It achieves this by dividing the image into small cells, computing the gradient histogram for each cell, normalising these into larger blocks and concatenating them to form a final feature vector.
b) CNN-based face recognition uses a hierarchical approach to learn complex features from the input image using convolutional, pooling, and fully connected layers.

Running on several samples, it can generally be said about both methods that, although according to the documentation, the CNN-based method is more resource-demanding, the HOG-based method produced the same quality in 0.24 thousandths of a second, and the CNN-based method in 0.07 thousandths of a second.

4. Comparison of passport photo and selfie photo (face compare). We also use the "face recognition" library to compare the ID and selfie pictures. The two methods we can use are "compare_faces" and "face_distance". The former returns a simple true/false value (whether the two images match or not), while the latter calculates the Euclidean distance between the two images. The developer can determine what threshold value is used to check the match. This distance can be a value between 0 and 1, which means an increasing similarity as it approaches 0 while an increasing difference as it approaches 1.

5. Detection of document texts and their exact location (text detection). There are also pre-taught models to accurately determine the location of the text on the ID card, but these did not yield acceptable results during the preliminary tests. One of the following pre-trained OCR engines can be a good basis for creating your own fine-tuned model:

a) Tesseract OCR: the OCR engine developed by Google is open-source and supports over a hundred languages, including Hungarian.
b) OCRopus: also developed by Google, it was developed more for the analysis of historical and scientific documents, uses the Tesseract engine, but also has its own layout analysis and text recognition components.
c) CRAFT: a text recognition model based on deep learning that works well on complex and crowded images.
d) EAST: real-time text recognition model that can handle different text orientations, sizes and languages. Another advantage is that it is fast and accurate.
e) TextBoxes++: this model is also based on deep learning, capable of detecting text of various shapes and orientations.
f) PaddleOCR: it can also be used for text recognition, character recognition, and layout analysis and is designed to be easy to use and customisable. It supports several languages, including Hungarian too.

Preliminary comparison and evaluation of the listed models is difficult, as their performance varies depending on the data set, image quality and applied image processing techniques.

6. Recognition of detected texts (text recognition). It is important to note that Tesseract and OCRopus, PaddleOCR are so-called end-to-end OCR models that include both text detection and text recognition models. In contrast, the other three can be used "only" for text recognition. Using such an end-to-end model is appropriate in our case, so we chose PaddleOCR.

Table 2. Comparison of text recognition and OCR models

Model	Speed	Accuracy	Teaching is required	Language support	Sensitivity to preprocessing
Tesseract	Fast	Medium	No	High	High
OCRopus	Medium	Medium	No	High	High
CRAFT	Medium	High	Yes	Medium	Medium
EAST	Fast	High	Yes	Medium	Medium
TextBoxes++	Medium	High	Yes	Medium	Medium
PaddleOCR	Fast	High	No	Medium	High

IV. DISCUSSION

In our paper, we presented character recognition and face identification based on deep learning through processing a personal identification document. We divided the process into parts, taught different artificial intelligence models, produced the data necessary for teaching, and then integrated it into our developed environment.

The completed program uses different Artificial Intelligence models to recognise the data on the identity documents, thanks to which the possibility of misreading and mistyping is reduced to almost 0%, increasing accuracy and reducing administrative time.

V. CONCLUSION

The most suitable technology for the development of the program is the so-called web application, which can be used via the Internet using a browser, and regardless of device type, it can work on anyone's smartphone, tablet, laptop or even desktop computer.

The completed software, which can be adapted to any system in the form of an application, uses the image on the ID card and the device's camera to determine whether the application is being used by the person authorised to administer it and can optionally match the signature on the ID card with a digital sample requested by the application from its user.

REFERENCES

[1] Annuš, N., Csóka, M., Paksi, D.: Learning Management Systems and Their Possibilities in Education - Case of Slovakia. 17th International Technology, Education and Development Conference, 2023. 6981-6986

[2] J. Udvaros and N. Forman. (2023) Artificial Intelligence and Education 4.0, INTED2023 Proceedings, pp. 6309-6317. https://doi.org/10.21125/inted.2023.1670.

[3] J. Udvaros and O. Takáč. (2022) Technical IT solutions in teaching, INTED2022 Proceedings, pp. 4047-4052. https://doi.org/10.21125/inted.2022.1107.

[4] PYMNTS (2018) "Businesses Can't Just KYC, They Must Also KYCC". PYMNTS.com. [Online] Available: PYMNTS.com

[5] Directive (EU) 2015/849 of the European Parliament and of the Council of 20 May 2015 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, amending Regulation (EU) No 648/2012 of the European Parliament and the Council, and repealing Directive 2005/60/EC of the European Parliament and of the Council and Commission Directive 2006/70/EC

[6] Illési, Z., & Honfi, V. (2022). A Security Assessment of AI, Related to the Financial Institutions. In Security-Related Advanced Technologies in Critical Infrastructure Protection: Theoretical and Practical Approach (pp. 85–94).

[7] Nyikes, Z., Kovács, T. A., Honfi, V., & Illési, Z. (2022). Digital Competence and Security Awareness from the Perspective of Sustainability. In Security-Related Advanced Technologies in Critical Infrastructure Protection: Theoretical and Practical Approach (pp. 139–150).

[8] Honfi, V., Neuhoffer, H., & Barna, R. (2008). Szükséges-e a számítógépes bűnözés oktatása pénzügy szakon? In Informatika a felsőoktatásban 2008.

[9] Fülöp, M.T.; Gubán, M.; Gubán, Á.; Avornicului, M. Application Research of Soft Computing Based on Machine Learning Production Scheduling. Processes 2022, 10, 520. https://doi.org/10.3390/pr10030520

[10] M. Uřičář, P. Křížek, D. Hurych, I. Sobh, S. Yogamani, P. Denny, "Yes, we GAN: Applying adversarial techniques for autonomous driving" in Proc. IS&T Int'l. Symp. on Electronic Imaging: Autonomous Vehicles and Machines Conference, 2019, pp 48-1 - 48-17.

[11] M. Memon: “ANN vs CNN vs RNN: Neural Networks Guide” 2022 Levity website [Online]. Available https://levity.ai/blog/neural-networks-cnn-ann-rnn

[12] O.Takáč, L. Végh (2021) CREATION OF 3D MODELS OF REAL OBJECTS IN THE TEACHING OF COMPUTER SCIENCE, ICERI2021 Proceedings, pp. 5723-5727.

[13] Sandberg, David. Github. https://github.com. (2023) [Online] Available: https://github.com/davidsandberg/facenet

[14] Baltrusaitis, Tadas. Github (2023) available Ahttps://github.com. [Online] https://github.com/TadasBaltrusaitis/OpenFace

[15] Serengil. Github. https://github.com (2023) [Online] available https://github.com/serengil/deepface

[16] Rcmalli. Github. https://github.com [Online] available: https://github.com/rcmalli/keras-vggface.

[17] Cochard, David. Medium. https://medium.com (2021) [Online] Available: https://medium.com/axinc-ai/arcface-a-machine-learning-model-for-face-recognition-5f743cdac6fa

[18] Timesler. Github. https://github.com (2023) [Online] Available: https://github.com/timesler/facenet-pytorch

[19] Geitgey, Adam. Github (2023) https://github.com. [Online] Available: https://github.com/ageitgey/face_recognition

[20] Chablani, Manish. Medium. https://medium.com.(2017) [Online] Available: https://medium.com/towards-data-science/yolo-you-only-look-once-real-time-object-detection-explained-492dc9230006

[21] Hui, Jonathan. Medium. https://medium.com. (2018) [Online] Available: https://medium.com/@jonathan-hui/ssd-object-detection-single-shot-multibox-detector-for-real-time-processing-9bd8deac0e06

[22] Gao, Hao. Medium. https://medium.com. (2017) [Online] Available: https://medium.com/@smallfishbigsea/faster-r-cnn-explained-864d4fb7e3f8