

R&D Report

Enhanced Al

Augmenting Identity Verification with Artificial Intelligence





Table of Contents

| 02 | Scaling ID Verificat |
|----|----------------------|
| 03 | Machine Learning |
| 05 | Natural Language Pro |
| 06 | Deep Learning |
| 07 | Verification Experts |
| 08 | Enhancements in S |

tion with Artificial Intelligence

ocessing

s to Supplement Technology

Shufti Pro's Technology

Overview

It's no secret that identity-related fraud is a growing problem, especially in the digital world, where certain limitations exist in the physical verification of endusers. A record of \$16.9 million was lost in digital fraud during 2019-2020, which is 13% more than the previous year [1].

Fraudulent transactions, data breaches, and identity theft instances continue to rise as fraudsters become more sophisticated. Online ID verification solutions are considered as one of the ways to cope with identity-related frauds. From scanning simple ID documents and MRZ codes to performing sophisticated biometric checks, verification solutions have multiple benefits.

Subdomains of artificial intelligence like machine learning and deep learning have paved way for effective means of ID verification. The technology is used to successfully authenticate, verify, and process identities at scale.



1473 data breaches occurred in 2019 exposing nearly 165 million sensitive records (Review42)

Scaling ID Verification with Artificial Intelligence

To check various elements against an ID, either remotely or on-premises, information from ID documents such as national identity cards, driver's licenses and passports is extracted. Some of these identity checks include identifying the country of document origin, confirming the authenticity of microprint text and security threads, validation of special ink, and comparison with data extracted through OCR and MRZs. Moreover, it also includes validating data and verifying the facial features to link individuals to ID documents provided for verification.

With the power of machine learning, automating operations makes the process accurate and efficient rather than solely depending on human verification experts. These ID solutions collect internal data through robust mechanisms for recordkeeping of software performance and operation. The acquired data is regularly transmitted to the provider using automated systems. The process, if automated, significantly improves the accuracy and turnaround time.

Identity document verification using AI checks the input image against all authentic classes of documents to verify its originality. Parameters such as consistency in content, visual correspondence, and resemblance in layout are assessed to perform verification.



Shufti Pro adopted artificial intelligence to enhance verification process accuracy. But to understand how Shufti Pro uses artificial intelligence to scale ID verification, you should first understand these concepts: **Machine Learning (ML), Deep Neural Networks, and Natural Language Processing (NLP).**

Machine Learning

'Machine learning algorithms find and apply patterns in data. And they pretty much run the world.' [2]

In simple terms, machine learning is twofold: (i) Make the computer learn by showing it a set of examples (ii) Perform predictions based on the learned information. Some good examples could be detecting an apple or banana in a given photo, assessing movements of pedestrians for a self-driving car, accurately recognising speech to generate captions for a YouTube video, or identifying documents, which in our case, is the main purpose of using machine learning algorithms. These algorithms are only capable of performing tasks they are built for and nothing else.



Shufti Pro



In our case, we are particularly interested in the classical applications of machine learning and computer vision. Shufti Pro incorporates computer vision to generate ID templates for specific types of IDs (e.g. a UK passport or Canadian driving license). Our team has created classifiers for document edges and coordinates of document images. All these classifiers are used to develop algorithms that help in determining elements of an ID document including the type and country of origin of the document.

Identifying type and origin of the document is important, but is not the only essential thing to verify identity. One of the other important things for identity verification includes data extraction from the document. Even though OCR technology is used to extract data from paper documents, traditional OCR engines aren't that effective when it comes to ID documents. This is because most of the ID documents contain text in local languages from the country of origin. For recognising text in multiple languages, traditional OCR doesn't provide the accuracy required for such critical tasks. This is why natural language processing, another subset of AI, is of utmost importance.





Natural Language Processing

Definitionally, Natural Language Processing (NLP) is a branch of AI that helps computers to understand, comprehend and manipulate human languages [3].

As a human, you may speak and write in English, Spanish or Chinese. But a computer's native language – known as machine code or machine language – is largely incomprehensible to most people. At your device's lowest levels, communication occurs not with words but through millions of zeros and ones that produce logical actions [3].

Since identity documents are in multiple languages and the specified format, extracting the information accurately requires a lot more than simple OCR. So natural language processing algorithms are useful to enhance the OCR engine. For example, in the identity documents, it is important to verify that the expiry date is accurately identified. With NLP, we could verify that only numbers are extracted from this field. Shufti Pro enhanced its verification process by developing a purpose-built OCR engine.



Computer vision and NLP are good, but not great enough - especially when the identity image in the document is not captured in an ideal condition. Due to their error-prone nature, classical computer vision and image processing couldn't compete with human intelligence, until recent advances brought innovation to the table. Due to this, deep learning algorithms play a key role in these operations.

Deep Learning

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called Artificial Neural Networks (ANN) [4]. The design of an ANN is inspired by neural networks in the human brain. Andrew Ng, a renowned deep learning supporter and mind behind Google Brain, said that deep neural networks are a step towards real AI.

"Using brain simulations, hope to:

- Make learning algorithms much better and easier to use.
- Make revolutionary advances in machine learning and AI.
- I believe this is our best shot at progress towards real AI"





^[4] What is Deep Learning?, machinelearningmastery.com

Shufti Pro uses deep learning models to identify fraudulent activities related to ID documents. Deep learning helps us identify fake IDs such as understanding the unique font of an ID document, security features of IDs, forgery, etc. Apart from this, liveness detection and anti-spoof deep learning algorithms for 3D detcetion help in accurately verifying that the person taking a picture is the same on the document and is a real person.



03

Purpose-built OCR engine having accuracy over 90%



Verification Experts to Supplement the Technology

However, these algorithms are trained using data and self-taught from previous inputs, there are still instances where machine learning, NLP or deep learning models may not be able to correctly identify issues with the ID. This is where human verification experts come in handy. The expert reviews verification operations and identify potential errors so that they could be avoided in future, and ID verification could be fully automated. This method creates a suitable environment for learning and helps Shufti Pro enhance its identity verification procedure.

Unlike fully automated systems, Shufti Pro takes into account the accuracy of online identity verification. Incorrect verifications could result in a major loss in capital and market reputation for the enterprise since it could provide unauthorized access to sensitive data. This is why Shufti Pro relies on a hybrid approach where human verification experts verify suspicious IDs and analyse the reasons for acceptance and failure. The causes of a failed authentication by verification experts are deeply analyzed to better identify unique features missed by identification systems, which helps us in refining our AI models.

Having realised the power of Artificial Intelligence and a vision to make online marketplaces secure, Shufti Pro is constantly improving its technology and Enhanced AI is our bet to make identity verification secure and reliable.



Incorporate Enhanced AI to optimise Identity Verification





www.shuftipro.com





Get Information

sales@shuftipro.com

Shufti Pro True Identity Builds Trust

Expanding services to 230+ countries and territories in a short period of time, Shufti Pro envisioned playing a pivotal role in creating cyberspace where every transaction is verifiable and secure. With enough experience in technologies like Machine Learning (ML), OCR, artificial intelligence, and Natural Language Processing (NLP), Shufti Pro strives to provide the best identity verification services to verify customers and businesses online.

Shufti Pro's cost-effective solutions help businesses to prevent fraud and illicit crimes that can ruin the integrity and brand reputation of your business. Our perfect solution suite consisting of KYC verification, AML screening, ID verification, Facial Recognition, Biometric Authentication, Video KYC, OCR, and KYB helps to improve your company's fraud prevention, Know your Customer (KYC) and Anti Money Laundering (AML) regulatory efforts by automating the workflow. With single API integration, Shufti Pro empowers you to verify customers with document checks from 3000+ ID templates and business entities from 200 million companies data.

Disclaimer: No warranty or claim is herein provided that information contained in this document is accurate, up-to-date, and/or complete. All information provided in this document is limited for general informational purposes only. In no circumstance(s), does such information constitute as legal or any other advice. Any individual or company who intends to use, rely, pass-on, or re-publish the information contained herein in any way is solely responsible for the same and any likely outcomes. Any individual or company may verify the information and/or obtain expert advice independently if required.

