BEYOND MACHINE LEARNING

How Machine and Deep Learning Are Used in Identity Verification

acuant®

CONTENTS

I. INTRODUCTION: DEFINING ARTIFICIAL INTELLIGENCE, MACHINE LEARNING, AND DEEP LEARNING

In the world of computer science, Artificial Intelligence (AI) is one of the most significant advances that has changed the way people interact with technology today. The ability for computers to make human-like decisions creates efficiencies that were previously unachievable. Artificial Intelligence is a buzzword revolutionizing many businesses, however, the word has come to be used interchangeably with Machine Learning (ML) and Deep Learning (DL).

DEFINITIONS

ARTIFICIAL INTELLIGENCE: Computers perform tasks that normally require human levels of intelligence such as decision making, speech recognition, language translation, and visual perception. Artificial Intelligence can be broadly categorized into three groups: Narrow AI, artificial general intelligence (AGI), and superintelligent AI.

> MACHINE LEARNING IS A SUBSET OF ARTIFICIAL INTELLIGENCE: Computers that can complete tasks at human levels of intelligence make use of wide sets of data with many variables to find insights and make decisions. Machine Learning allows machines to take data and learn for themselves, recognizing patterns and making predictions.

>> DEEP LEARNING IS A SUBSET OF MACHINE LEARNING: Artificial Neural Networks that contain multiple layers learn and process information by mimicking human neurons. Deep Learning uses some Machine Learning techniques to solve real-world problems, it can take a huge amount of data—millions of images, for example—and recognize certain characteristics to aid in fraud detection, spam detection, and more.

AI can include anything from computers playing games, to voice recognition programs like Siri and Alexa, while Machine Learning and Deep Learning can power search engines, photo recognition technology, and self-driving cars among other tasks today. Machine Learning is generally defined as the ability for computers to learn without being explicitly programmed. Instead of using code, computers feed data into an algorithm which then builds logic based on that data. Deep Learning is what many consider the most exciting or radical of this set of buzzworthy and transformative technologies, as it can solve problems that require human or artificial thought.

II. MACHINE LEARNING IN IDENTITY AND ACCESS MANAGEMENT SOLUTIONS (IAM)

Many industries employ automated ID authentication for customer onboarding, access control, online transactions and card not present transactions. These include but are not limited to financial institutions, healthcare practices, eCommerce and sharing economy mobile apps, security systems and hospitality environments. Scanning customer IDs also allows businesses to get accurate information for customer relationship management systems, removing the inconvenience of manually entering in data, and improving customer experiences and conducting verification in one step.

The amount of fraudulent transactions, massive data breaches, and instances of identity theft continue to rise as hackers and fraudsters become more sophisticated. Couple this with today's burgeoning digital economy - where consumers are living their lives from mobile devices and conducting transactions remotely - and the need for strong, customer-friendly identity proofing solutions becomes critical.

Artificial Intelligence, and its subsets of Machine Learning and Deep Learning, make it possible to accurately process, verify, and authenticate identities at scale.

> "The Consumer Identity and Access Management Market (IAM) has a projected worth of **14.82 Billion** USD by 2021."
>
> *- MarketsandMarkets, Identity & Access Management Market by Component and Region (2017, February)*

FASTEST GROWING INDUSTRIES IN IDENTITY TECHNOLOGY

BANKING

HEALTHCARE

GOVERNMENT

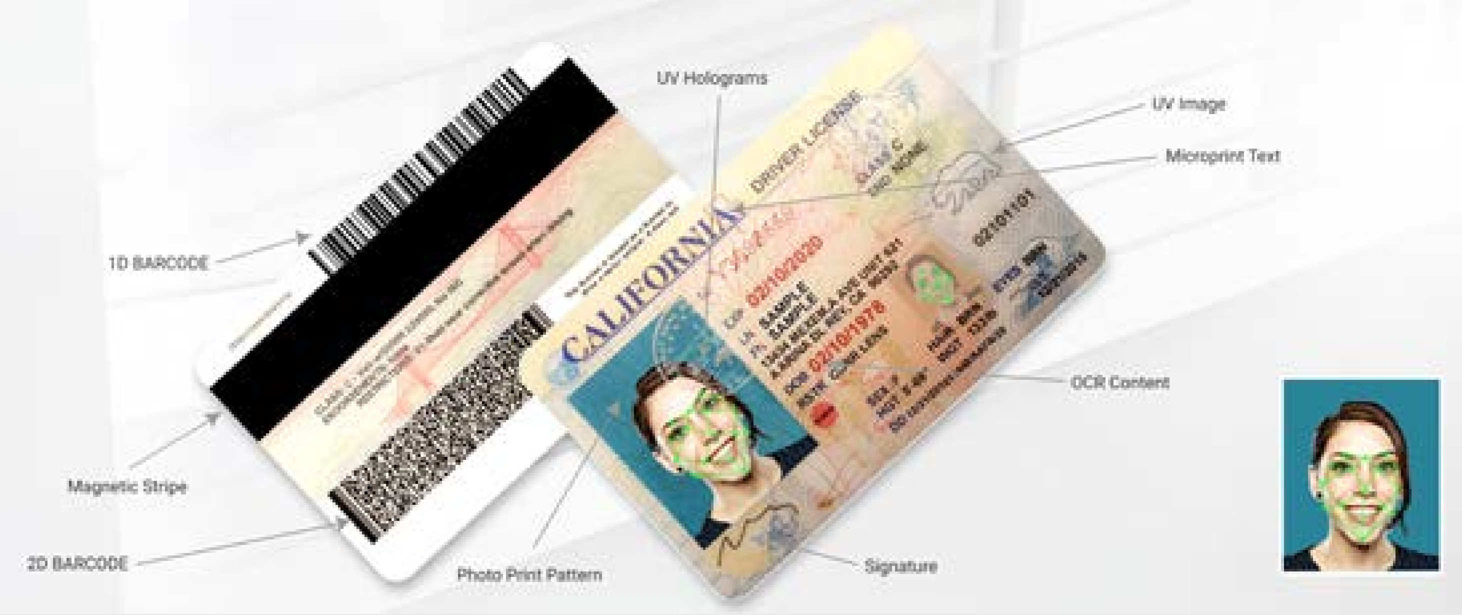
A. IDENTITY DOCUMENT AUTHENTICATION

There are various strengths of ID scanning solutions; some simply scan the ID's barcode while more robust software performs forensic and biometric tests to ensure that an ID is not forged. Identity documents, such as driver's licenses and passports, are scanned either on premise or remotely with mobile devices to test various elements of an ID.

For example, shining a UV light on a hologram may prove that the ID passes this test, while a simple scan of a barcode may show that the card appears to have valid data. Businesses seeking to authenticate identity documents should look for multiple tests depending on the use case and level of risk associated with the transaction. The stronger the tests, the easier it is to approve (or deny) transactions.

SAMPLING OF AUTHENTICATION TESTS:

- ✓ CONFIRMATION OF GENUINE MICROPRINT TEXT & SECURITY THREADS
- ✓ VALIDATION OF SPECIAL PAPER & INK, AND INK PATTERNS
- ✓ COMPARISON BETWEEN MRZ AND VISUAL INSPECTION ZONES
- ✓ COMPARISON BETWEEN OCR AND BARCODES, AND MAGNETIC STRIPES
- ✓ OPTICALLY VARIABLE DEVICES TEST FOR STATE SEALS & HOLOGRAMS
- ✓ ELECTRONIC CHIP CONTENT CROSSCHECK WITH OCR CONTENT
- ✓ BIOMETRICS FOR FACIAL RECOGNITION
- ✓ DATE VALIDITY TEST
- ✓ MANUAL REVIEW



SAMPLE OF DOCUMENT ELEMENTS TO CHECK

These tests and automated capture technology are much more accurate and efficient than an untrained human looking at a document, or even a trained human authenticating all the various elements on an ID. Today's identity authentication solutions are advanced and catch fakes that trained humans could not. In seeking solution providers, consider the number of tests performed and speed of the testing. Preferred solutions will be automated and provide multiple tests in a matter of seconds to keep customer experience frictionless and results optimal. Tests performed by machines, as opposed to humans who will need training and could miss something, will also handle a higher volume and be more consistently accurate.

i. SCALING ID AUTHENTICATION WITH MACHINE LEARNING

> "Recent advances in data science, machine learning and other artificial intelligence techniques provide significant user experience improvements in identity proofing processes."
>
> *- Gartner, Inc., Identity Proofing is the Cornerstone of Trust in a Digital Relationship Rabinovich, P., Robins, A. E. (2016, October 13)*

Preferred solutions must apply Machine Learning in identity document authentication to continually improve. Solutions should contain an anonymous internal data collection mechanism capable of storing information about the operation and performance of the software. This data should be transmitted to the provider automatically on a regular basis. This process, if automated, saves time and improves the quality of the results without user intervention.

In addition to mining information about the document template, image, and personal data, there are often overlooked attributes of the document that will factor into the machine's decision making.

MACHINE LEARNING MODELS FOR ID AUTHENTICATION MUST INCLUDE:

- ✓ PERSONALIZATION OF THE DOCUMENT
- ✓ VARIATIONS IN MANUFACTURING
- ✓ VARIATIONS DUE TO WEAR AND AGING OF THE DOCUMENT
- ✓ TAMPERING AND COUNTERFEITING

ii. THE IMPORTANCE OF AN IDENTITY DOCUMENT LIBRARY

A typical program relies on the collection of metadata on the document recognition and authentication process. Instead of containing information about the document being processed, this metadata contains information about the processes that are run and the details and outcome of those processes. By analyzing this information, the software trains itself to detect complex patterns and output a prediction/result. It optimizes the performance of the software and library to improve the reliability of the document read and authentication processes.

Using Machine Learning to identify between good and bad IDs is extremely efficient. However, without supervision, the logic developed by the algorithm may exclude IDs that are not fraudulent. There are many reasons why an ID may not pass even though it is valid. For example, IDs are often not printed in the same location by the same machine. There may be several printers creating IDs which can develop anomalies based on printing quality, misprints, or unclear images. In addition, a computer must be taught that IDs which are worn or damaged are still valid. Accessing a program that allows the collection of operational and performance metrics for the software is useful for improving the recognition and authentication of documents supported by the document library of the provider.

A robust document library to compare captured IDs against is vital. A comprehensive and regularly updated library cuts down the time that machines must process data on their own, and maximizes data extraction and authentication capabilities. Semi-supervised Machine Learning enables adjustment of the direction of the logic without interfering with the insights that authenticate documents or slowing down data processing.

Automatically passing worn or damaged IDs that otherwise would be marked as bad is a key technology that minimizes friction for both the business and customer. Ultimately more good transactions are approved. If a good customer is unable to make a high value purchase, is denied credit or a loan, or is unable to gain access (physical or digital), everyone loses and there is little benefit to automation. It is instances such as this where semi-supervised machine learning benefits are evident. Another feature to look for in providers besides a substantial, automated document library is a Customer Experience Program to ensure that document libraries are used in tandem with input from the real world.

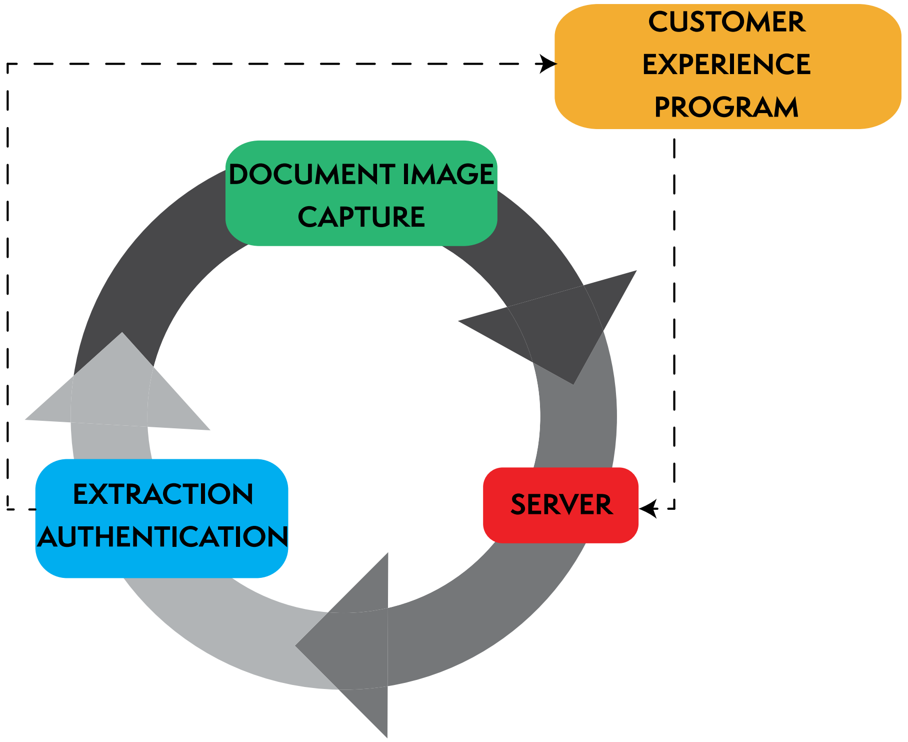
iii. MODELING AND FEEDBACK LOOP

Preferred ID Authentication Solutions providers should be using the following models:

DATA MINING: Examining large databases to turn raw data into useful information. Make sure the provider also extracts clean data to save time and increase efficiency of the solution

SEMI-SUPERVISED LEARNING: Completely automated Machine Learning might fail documents in real-life situations such as wear and tear on a document or manufacturing errors

REGRESSION ANALYSIS: A method to continually test and analyze outcomes to improve the algorithm



Training and re-training is a key part of Machine Learning and referred to here as the FEEDBACK LOOP. In this loop, new data is fed into the algorithm to test that outcomes are consistent and improving; the outcomes are then fed back into the algorithm so that the computer continues to learn. The more data fed into the algorithm, the more types of situations can be addressed and the more precisely decisions can be made.

B. BIOMETRICS: FACIAL RECOGNITION AND MATCH

Biometric identity verification methods implement a biometric measure, such as facial or voice recognition, to strengthen the identity proofing process. Biometric verification is more passive for consumers, reducing the need to manually enter ID information/passwords or answer time-consuming, knowledge based authentication (KBA) that any fraudster might know, guess, or hack. Facial recognition match technology makes it easier and more user-friendly to verify identities and integrate with existing apps or systems. Customers can simply snap selfies to verify that they match the photo on their government issued ID from a webcam or mobile device. This extra step prevents fraudulent transactions by further binding a person to their identity. Biometric applications can be used in any industry for initial or recurring transactions to match a person to their ID on file, or to confirm with liveness detection and image spoofing tests that a real person that matches the ID on file is trying to access information, service or location.

USE OF FACIAL RECOGNITION FOR IDENTITY ACCESS MANAGEMENT SHOULD BE:

✓ Fast and Convenient

✓ Contain real time checking and processing

✓ Match face biometrics to the picture of the ID document

✓ Test for face liveness

Biometric security solutions utilize Deep Learning to mimic the way human neurons process extremely difficult information like faces and language. With Deep Learning technology, the software provider can model large amounts of complex data, such as many images and faces. Facial Recognition technology utilizes Deep Learning to learn to match the image on the ID to a person's face. The basic process to do so is by scanning the document and taking a selfie. Then the algorithm looks for certain patterns such as basic shapes (eyes, mouth, nose) and complex shapes (complete faces and distinctive shapes), and finally returns an output that indicates whether the image matches the IDs face or not.

III. CONCLUSION: AI TECHNOLOGY POWERS EVERY LAYER OF IDENTITY AND ACCESS MANAGEMENT SECURITY

Artificial Intelligence increases customer security and prevents fraud. Identity Proofing providers are a trusted source to collect evidence of fraud, and through machine learning, that evidence can be analyzed and applied continuously to ensure higher accuracy. This automation makes it possible for ID verification and authentication to provide a seamless layer of security across every type of transaction for any industry and on any platform.

Because the pass/fail thresholds for documents can be extremely black and white, the combination of both Machine Learning and human insight is needed to produce the best results. Human insight can be applied automatically by the computer and continue adapting as more data gets processed. Leveraging both methods enables solutions that provide the scale needed to verify and authenticate millions of transactions at once. A manual review option by humans to catch flagged anomalies is also a good idea as a last check, and could allow for approval of more valid IDs. Lastly, Deep Learning for photo and facial recognition match technology adds another layer of security to businesses and consumers to add further protection.

With the Machine Learning capabilities available today, businesses looking to provide a more seamless, secure experience for customers no longer need to worry about managing high volumes of transactions while protecting customers.

WHEN EVALUATING IDENTITY VERIFICATION SOLUTIONS, ENSURE PROVIDERS:

- ✓ USE A LAYERED APPROACH AND OFFERINGS ON AUTHENTICATION SOFTWARE THAT ADDRESSES ALL CHANNELS IN MULTIPLE INDUSTRIES - ESPECIALLY THOSE WITH HIGH VOLUME AND GROWTH
- ✓ ARE CUSTOMER CENTRIC, BUILDING SOLUTIONS WITH THEIR PARTNERS' NEEDS AND THEIR END USERS' EXPERIENCES AS THE FOCUS
- ✓ CONTINUOUSLY ADAPTING TO EVOLVING PHYSICAL AND DIGITAL THREATS

ADDITIONAL RESOURCES AND FURTHER READING

Read more about the technology behind identity-based transactions, where the Identity and Access Management Industry is trending, and how to better protect businesses with these additional resources:

IDENTITY FRAUD PREVENTION IN TODAY'S DIGITAL ECONOMY

https://www.acuantcorp.com/identity-fraud-prevention-in-todays-digital-economy/

WHAT WE LEARNED AT THE K(NO)W IDENTITY CONFERENCE: PART ONE

https://www.acuantcorp.com/what-we-learned-at-the-know-identity-conference-part-one/

WHAT WE LEARNED AT THE K(NO)W IDENTITY CONFERENCE: PART TWO

https://www.acuantcorp.com/what-we-learned-at-the-know-identity-conference-part-two/

JUNIPER PREDICTS 600M+ MOBILE DEVICES WILL USE VOICE AND FACIAL RECOGNITION BY 2021

https://www.acuantcorp.com/biometrics-boom-juniper-predicts-600m-mobile-devices-will-use-voice-and-facial-recognition-by-2021/

ACUANT MAINTAINS UPTIME DURING DDOS ATTACK ON DYN

https://www.acuantcorp.com/acuant-maintains-100-uptime-during-dyn-ddos-attack/

IDENTITY AND ACCESS MANAGEMENT MARKET WORTH 14.82 BILLION USD BY 2021

http://www.marketsandmarkets.com/Market-Reports/identity-access-management-iam-market-1168.html

GARTNER REPORT: MARKET GUIDE FOR USER AUTHENTICATION

https://www.gartner.com/doc/3210517/market-guide-user-authentication

ABOUT ACUANT

Acuant Inc. is a global technology company that provides comprehensive Identity Verification solutions responding to various levels of risk and assurance requirements. Utilizing patented technology Acuant transforms data intake into a business-enhancing proposition while instantly reducing the risk of fraudulent transactions.

Acuant employs a Sample Management System (SMS) that contains millions of document image samples to train and test against for thousands of unique document types. In addition, the Acuant Customer Experience Program yields over two million new records a month allowing our products to be fine-tuned by actual field level experience. These are key ingredients to our unique testing methodology.

Acuant's AssureID™ Customer Experience Program collects operational and performance metrics for the AssureID software. This data contains information that is useful for diagnosing operational performance and improving the recognition and authentication of documents supported by the AssureID document library. The data collected is 100% anonymous. Absolutely no personally identifiable information (PII) about the customer, the operator of the software, or those whose documents are authenticated by the software is stored or transmitted to Acuant under any circumstances. The collected data does not include any images or partial images of the documents that are processed. The only information collected that relates to the document being processed is the type of the document (e.g. "USA Passport, Series 2006e"). This information is required in order to correlate performance issues with the document type so that improvements can be made to the AssureID document library and specifically, to improve the performance, recognition, and authentication reliability of documents encountered by Acuant's customers.

Acuant's intelligent engineering is made to work in any industry in any environment with compatibility for Windows, iOS, Android, Hybrid and HTML 5, and built to allow meeting the highest-level security requirements and regulations such as KYC, PII, HIPAA and AML. Partners include start-ups, Fortune 500 and FTSE 350 organizations.

Contact us

info@acuantcorp.com | acuantcorp.com

Tel: 213.867.2625
6080 Center Drive Suite 850
Los Angeles, CA 90045

acuant®