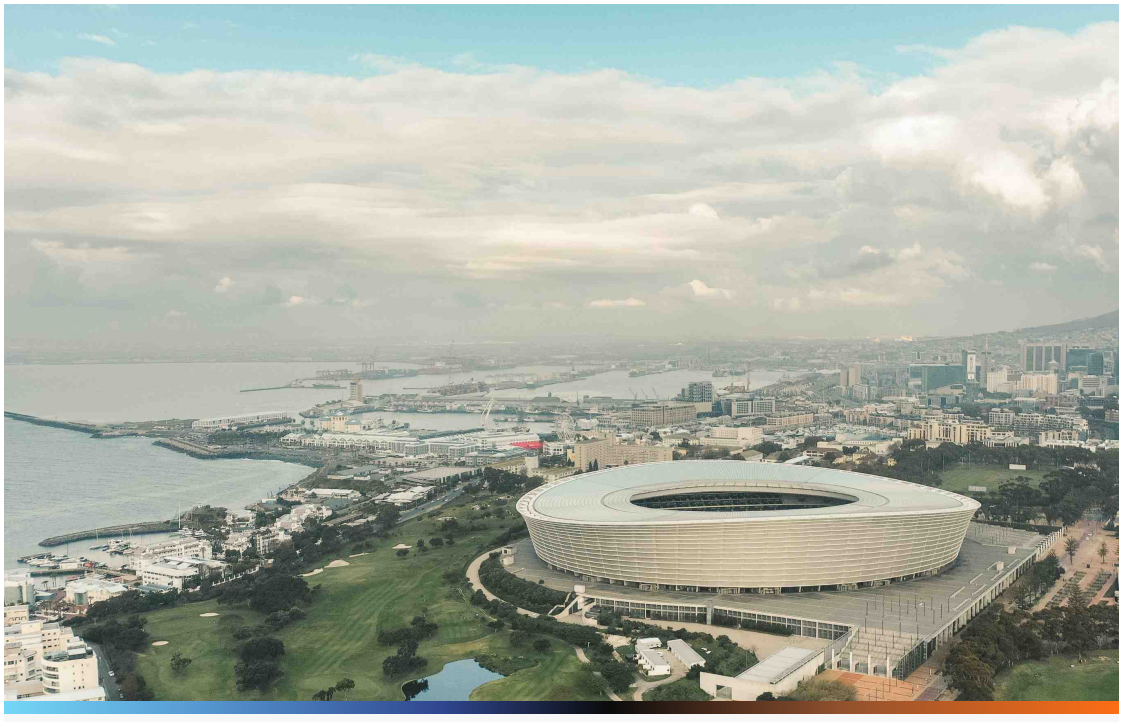
PARAVISION

ENTERPRISE-GRADE
1:N FACE RECOGNITION

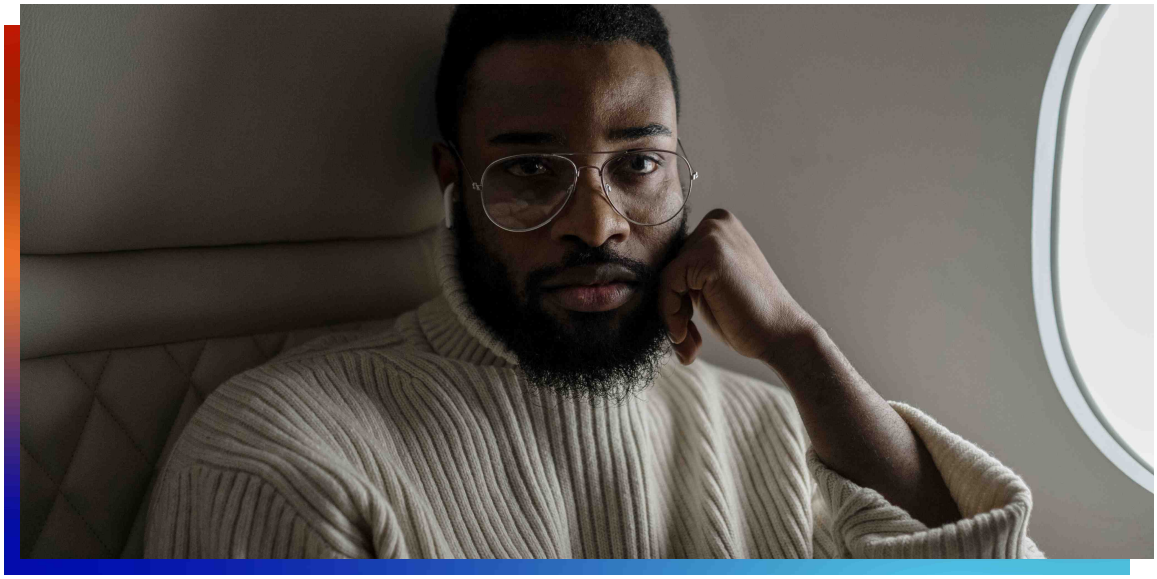
Face matching for large-scale use cases



→ paravision.ai

Trusted Vision AI

April 2023



Introduction

Biometric identification systems have become crucial for both governments and enterprises who need to confirm individual identities and uncover fraudulent attempts to gain identity privileges. Many of these biometric systems are small: for example, your smartphone only needs to manage one user identity, or at most the identities of a few users. Other biometric systems manage tens of millions of identities or even more.

In the past, large-scale biometric identification systems were based upon fingerprint biometrics, or sometimes on iris biometrics. For example, the Aadhaar system in India, the world's largest biometric system, is primarily based upon fingerprint and iris biometrics. Historically, fingerprint and iris biometrics were used in these systems because those modalities were significantly more accurate and suited to scale than face recognition.

However, in the past few years, face recognition systems evolved significantly and are now on a par with or more accurate than systems based on other biometric modalities, and are becoming more popular in government and enterprise applications.

This white paper explores some of the reasons why face recognition is now suitable for large-scale government and enterprise-grade applications, and how newer face recognition systems can accurately search large databases with massive concurrency and low latency. The paper discusses the following:

- How newer face recognition systems can scale to process searches on hundreds of millions of records.
- How newer face recognition systems can maintain biometric accuracy even at scale.
- The new use cases that face recognition systems can support.
- The necessary components for an enterprise-grade face recognition system.

Key terms

Before delving into the capabilities of newer enterprise-grade face recognition systems, this paper will discuss critical terms used when discussing biometric recognition. The first topic that must be addressed is the idea of identity—who a person is—and the ability to confirm that identity. A person's identity may be confirmed by one or more of the following authentication factors:

Something you know, such as a password or a token

Something you have, such as a passport or driver's license

Something you are, such as the features of your face

Something you do, such as swiping a particular pattern on a screen

Somewhere you are, or your geolocation

Individuals can be uniquely identified by one of these factors, or by several factors in combination. Biometrics ("something you are") such as face recognition are one of the tools used to uniquely identify an individual. While biometrics may be captured in image form, they are not processed as images. Instead, biometric systems convert the original images to templates (referred to as "embeddings" in the world of AI). Templates provide two benefits over the original images.

1. Faster results due to smaller storage space and improved transmission time. In modern face recognition systems, a template is simply a vector that is on the order of 1 kB in size, vs. 100 kB for the original image.

2. Improved privacy due to retention of only the data needed to identify individuals. While a compromised face image can be used by anyone to identify an individual, a compromised proprietary template is much more difficult to reverse engineer, and even then cannot be used to re-create the actual appearance of a person.

Two other key terms are verification and identification, which require vastly different biometric methods and systems.

Verification (one-to-one, 1:1)	Answers the question "Are you the person that you claim to be?" If a person claims to be (for example) Lionel Messi, and Mr. Messi''s biometrics have been stored, then the person's face, fingers, iris, or other biometrics can be compared to the stored biometrics for Lionel Messi to prove or disprove the asserted identity. Verification only requires comparison against a single biometric template, and is also known as a one-to-one or 1:1 search.
Identification (one-to-many, 1:N)	Answers the question "Who are you?" If I am suffering from amnesia, then my biometrics can be compared to the stored biometrics for missing persons. Hopefully, my biometrics will match those of a known missing person, which would allow my identification. Identification requires comparison against multiple biometric templates, and is also known as a one-to-many or 1:N search.

When examining how artificial intelligence is applied to biometrics, it is worthwhile to consider how AI defines the terms training, inference, algorithm, and model. As described later in this paper, model training iteratively analyzes and detects patterns in data, resulting in a model which is often colloquially referred to as an algorithm even if that is not technically correct. Inference takes the algorithm and analyzes new data to produce insights.

Finally, when examining independent evaluations of face recognition accuracy, there are four terms that need explanation.

False Non-Match Rate (FNMR)	Used in one-to-one (verification) searches, the ***FNMR*** is the proportion of comparisons that incorrectly result in a non-match when they should have resulted in a match.
False Match Rate (FMR)	Used in one-to-one (verification) searches, the ***FMR*** is the proportion of comparisons that incorrectly result in a match when they should have resulted in a non-match.
False Negative Identification Rate (FNIR)	Used in one-to-many (identification) searches, the ***FNIR*** is the proportion of searches incorrectly resulting in a non-identification when they should have resulted in a positive identification.
False Positive Identification Rate (FPIR)	Used in one-to-many (identification) searches, the **FPIR** is the proportion of non-mated searches incorrectly producing one or more candidates when they should have resulted in none.

Typically, biometric systems are evaluated by establishing an acceptable false positive rate (FMR or FPIR for 1:1 or 1:N, respectively), and then mapping the threshold to a resulting false negative rate. FPIR is inherently more challenging than FMR because the possibility for false positives increases as the number of comparison points (i.e. database size) increases.

For instance, in the ongoing National Institute of Standards and Technology (NIST) Face Recognition Vendor Test (FRVT), the “Visa-Border” 1:1 test uses a 10-6 (1 in 1 million) FMR, which results in a 0.0027 FNMR for Paravison’s Gen 5 face recognition. For 1:N, NIST FRVT benchmarks FPIR at 0.003 (3 in 1,000) on a database of 1.6M records, for which Paravision Gen 5 delivers an FNIR of 0.0038.

Scaling face recognition systems

Systems that perform identification are necessarily more complex than systems performing verification, since a biometric template has to be compared against a database of biometric templates. The system needs to evaluate the results of all of these comparisons, and usually ranks the results of these comparisons in the order of matching probability.

The first face recognition systems in the 1990s were capable of identification, but were incapable of performing identification on databases of tens of millions of individual records. The computer hardware and architectures of the day were incapable of supporting these large workloads. Even later systems (as of 2010) included algorithms with low accuracy, insufficient to create useful results.

Since that time, systems have improved in the following ways:

- Accuracy

  Artificial intelligence has dramatically improved face recognition accuracy. These improvements over the last decade, and even over the last three years, will be discussed in more detail later in this paper.

- Architecture and Hardware

  Computing architectures and hardware components have improved. Previous computers and on-premise computing models have given way to accelerated parallel processors, support of elastic, horizontal scaling to increase system capabilities, and virtualized component-based architectures, including Docker and Kubernetes. These improvements support managed distribution of the computing workload among devices.

- Cloud Computing

  Computing can take place beyond the premises. Systems are no longer limited to a single computer room. Where needed, cloud computing enables deployment of systems of unlimited size that are not limited to a single building or even a single geographic location. As computing needs increase, system managers (or even better, the systems themselves) can “spin up” additional server instances in near-real time.

All of these improvements have dramatically increased the scalability of face recognition systems. As described below, the improvements not only contribute to improved processing speed, but also to improved accuracy.

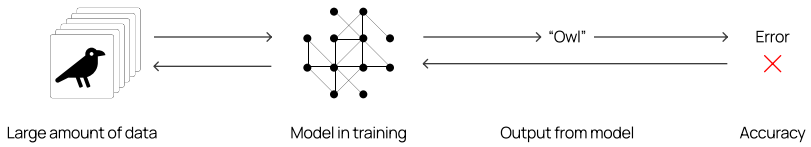
Providing accurate face recognition

One reason why early biometric systems relied on fingerprint and iris biometrics was because face recognition was not that accurate. Even as late as 2010, face recognition systems were not as accurate as other biometric systems. In that year, the National Institute of Standards and Technology (NIST) reported (NIST Interagency Report 7709) that for the best-performing face recognition algorithm "the chance of identifying the unknown subject (at rank 1) in a database of 1.6 million criminal records is about 92%."

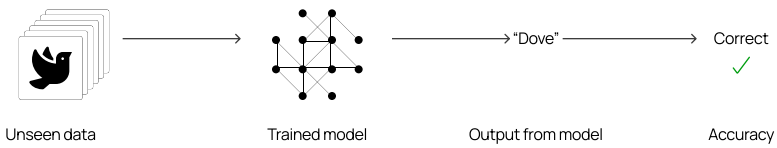
NIST testing in 2018 provided dramatically more accurate results. NIST Interagency Report 8238 demonstrated that "the most accurate algorithms (found) matching entries, when present, in galleries containing 12 million individuals, with error rates below 0.2%." These accuracy rates have continued to improve (in some cases with more than ten-fold reductions in errors since 2018) as NIST conducts its ongoing Facial Recognition Vendor Tests, and as organizations continue to submit newer algorithms.

Why did face recognition accuracy improve so dramatically over the past decade?
The introduction of artificial intelligence (AI) and machine learning (ML) to face recognition has resulted in dramatic accuracy improvements, as AI and ML experts have applied their knowledge to solving face recognition problems. As part of this, the experts have applied **model training** and **inference** to face recognition.

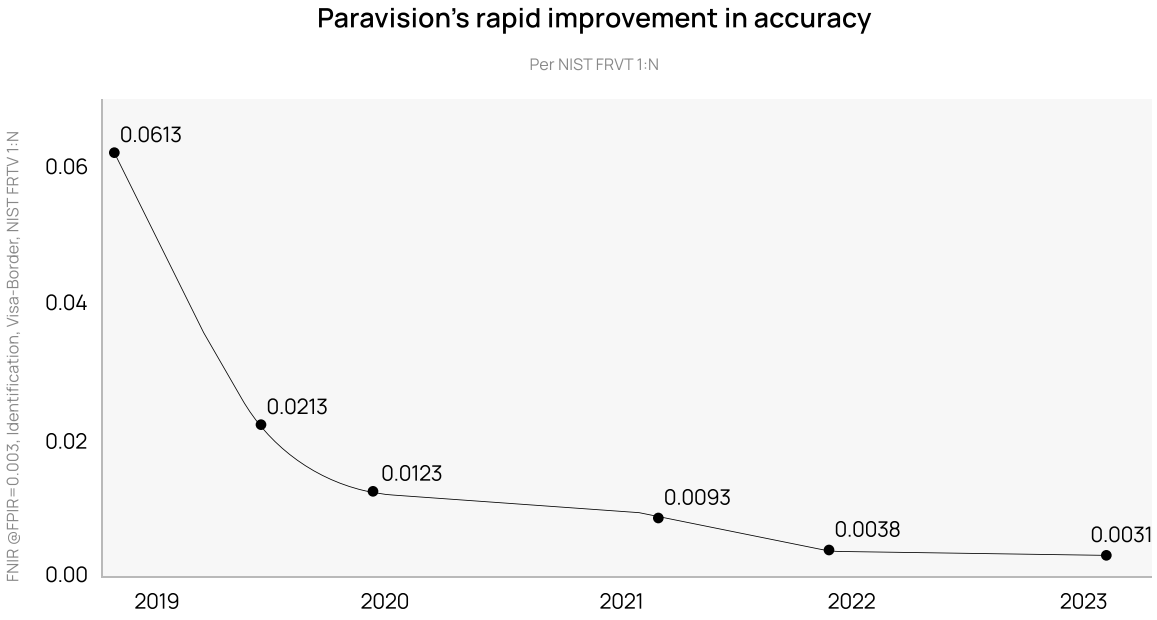
The first component of AI, model **training**, requires an iterative process of analyzing a large amount of data, detecting patterns in that data, and generating a model which reflects what was learned about detecting meaningful patterns. The model is checked for accuracy, and the process is repeated.



The second component of AI, **inference**, takes the model (often referred to colloquially as an "algorithm") generated by training and analyzes new data to produce insights. In some cases, these insights can form subsequent inputs to the iterative process described above.



As AI and ML experts continuously apply model training and inference to face recognition, accuracy continues to improve. The graphic below documents how AI and ML have improved the accuracy of Paravision face recognition. It displays a dramatic improvement in the false negative identification rate (FNIR) over the course of just three years. These results are taken from the National Institute of Standards and Technology (NIST) Ongoing FRVT 1:N Identification test, results of which are continuously updated on the NIST website.



Applying face recognition to use cases

These improvements in face recognition accuracy are coupled with other benefits of face recognition:

- **Ease of acquisition:** Faces are by far the easiest biometrics to acquire. They can be acquired at a distance, when a person is moving, in either visible or infrared light, and (with proper camera placement) without requiring any special effort or interaction from the person.

- **Ease of deployment:** Most face recognition systems do not require any special capture equipment, and can work with standard still or video cameras supporting a variety of formats. Many entities already have the necessary cameras, and well-designed face recognition systems can integrate with these existing camera systems.

Face recognition supports use cases not easily supported by other biometrics:

Identity Verification (IDV) and Know Your Customer (KYC)	Digital identity verification has exploded in popularity as a way of enabling self-service and real-time onboarding for financial services, gig economy, and government programs that demand strong identity. Face recognition is uniquely appropriate here due to availability of high quality smartphone selfie cameras to capture and match with faces on identity documents.
Travel and Borders	On every passport, state ID, and drivers license, the face is the ideal identity credential for travel and border. Meanwhile, in the airport or other border crossing, the ability to deliver fully touchless, intuitive, and fast biometric authentication makes face recognition the right tool for the job.
Enterprise physical access control	Fast, frictionless, touchless and secure facility access, from compact biometric readers that work with existing access control infrastructure. Face recognition enables highly secure biometric authentication from users who are in motion and at-a-distance.

Ethical Deployment

It should be specifically noted that just because face recognition can be applied to a very wide range of identification use cases does not mean that it should be applied to them. Achieving the level of accuracy noted here is not sufficient justification for deployment. There are many ways to look at limiting use cases. Paravision has taken the stand of creating a set of AI Principles to which all employees adhere. These excerpts from the company's AI Principles set the stage for Paravision's approach to ethical deployment.

> We believe in AI that is both Ethically Trained and Conscientiously Sold … We fully vet every potential partner or customer with whom we do business. Our leadership team understands and approves of the use of the technology and the level of thoughtfulness around issues of privacy, security, bias and human rights. To help ensure our technology is always conscientiously sold, we … **Limit distribution by use case.** We will only sell to law enforcement, defense, and intelligence agencies when we believe sufficient legislation or process is in place to govern its use, and we will never sell the technology for use cases intended to discriminate against any individual or group based on protected characteristics. We believe computer vision, and specifically face recognition, can be a powerful tool for law enforcement to solve and prevent crime and for defense and intelligence agencies to protect our nations, provided the face recognition solution used: 1) complies with applicable laws and is used for properly-defined purposes; 2) requires rigorous training for users on its legitimate and lawful use; and 3) operationally includes thorough human review and analysis before any consequential decision is made or action is taken. In no event will we permit the use of our technology for lethal autonomous weapons systems (LAWS).

In order to ensure compliance with its AI Principles, Paravision holds a regular Use Case Review, where potentially sensitive applications are discussed, measured against best practices, and in many cases rejected as they do not meet the wording or intent of the Paravision AI Principles.

Characteristics of a superior enterprise-grade face recognition system

Regardless of the particular applications, there are six key features that a superior enterprise-grade face recognition system requires.

01 The face recognition system must deliver **outstanding accuracy** in challenging, real world environments and across demographics including age, gender, and race. The system vendor must provide transparency in performance with statistically significant, real-world benchmarks. NIST FRVT is singular in this regard, and the system should use the same technology benchmarked in NIST FRVT. Other benchmarks, such as LFW, do not offer sufficient transparency and / diversity and challenge of data.

02 The face recognition system vendor must be committed to ethical development and deployment of the technology, ensuring that the face recognition system has been developed and tested in the right way, and that it will not be deployed in such a manner that is contrary to privacy or other usage laws or standards.

03 The face recognition system must support a **variety of deployment models**, including on-premises, cloud, mobile, edge, and hybrid solutions. Again, enterprises have taken advantage of the wide variety of deployment solutions available today. A face recognition system must operate in on-premises, cloud, mobile, edge and hybrid environments, and must support the unique security requirements of each environment for both enterprise and government users.

04 The face recognition system must be **based upon open systems standards**. Almost all face recognition systems are integrated into other systems, such as identity databases and security systems. Therefore, the system must easily integrate with these other systems.

05 The face recognition system must work with a **variety of common image formats**. Enterprises have sunk significant investments into existing systems, and do not want to replace these systems just to satisfy the needs of a new face recognition solution.

06 The vendor of the face recognition system must be **committed to continuous improvement** of its solution. This includes a commitment to improving accuracy and performance, and a willingness to incorporate new hardware and software technologies into its solution as they are developed. Ideally, the vendor's customers will realize continuous accuracy and speed benefits as the vendor improves its artificial intelligence, architecture, and other capabilities.

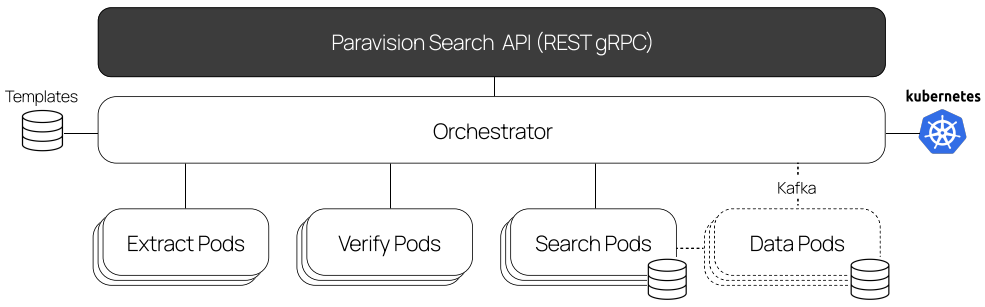
Deploying 1:N Face Recognition with Paravision

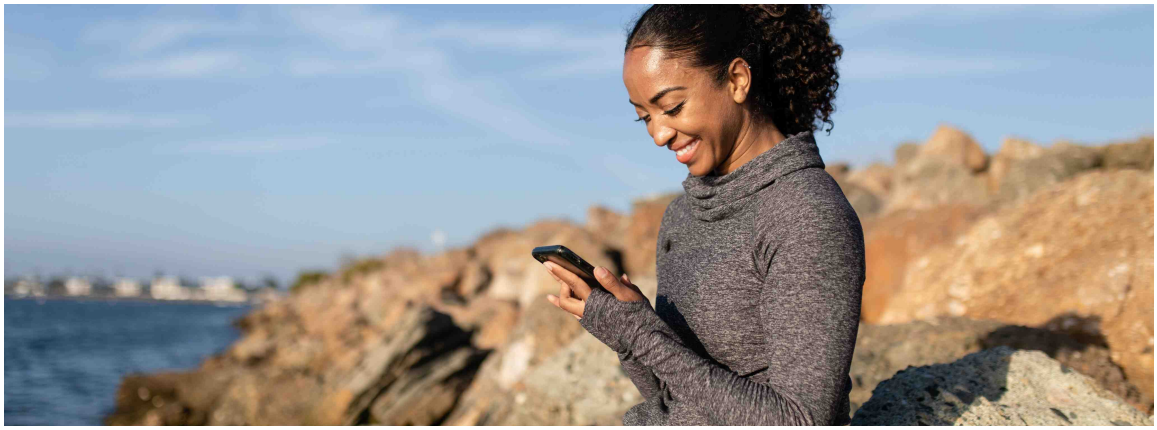
From Cloud to Edge, Paravision's comprehensive face recognition product suite offers all of the tools necessary to develop and deploy mission-critical biometric identification and authentication solutions across a wide range of applications, including identity verification, physical access control, air travel and government services. Paravision's toolset is modular and interoperable, enabling partners to develop the specific solution that meets their customers' needs. In addition, Paravision Search makes Paravision's technology easier to deploy and scale than ever, in particular supporting 1:N face recognition with high speed, superior accuracy, and fully elastic scalability.

Paravision face recognition has been optimized for leading chipsets, from Intel CPU to NVIDIA GPUs to Ambarella SoCs, supports leading operating systems including Windows, Linux, and iOS, Android, and can be deployed at any level, from SDK to API-driven Docker containers.

Here, we will look at 4 different use cases where 1:N face recognition is useful, and how Paravision's product suite can be integrated to meet the use case need. While the specific architecture for Paravision Search is included in each use case, please note that Paravision's SDKs or Dockers can be optionally used in the same way in the case that a system does not require the integrated benefits of Paravision Search.

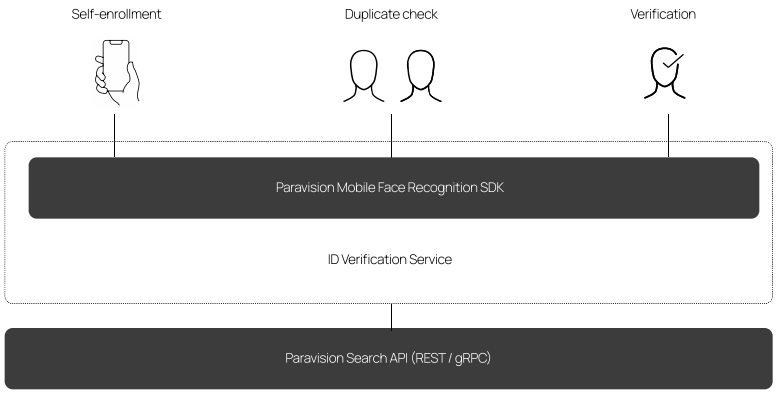
Paravision Search Architecture:





1. Digital Identity Verification

Identity verification ("IDV") has become a critical part of accessing trusted online services, enabling the gig economy, ensuring safety around age-restricted goods, and more. In many cases, IDV is enabled by a combination of 1:1 verification and 1:N identification. 1:1 verification is typically used for initial account generation, where a face may be matched against a government-issued identity document to ensure that the person registering for the account is the actual holder of that document. 1:N identification may be deployed alongside 1:1 verification to check fraud datasets. For example, fraudsters may be using the same face repeatedly with different names to create multiple accounts. Checking a database of known fraudulent images can offer an additional layer of identity security.

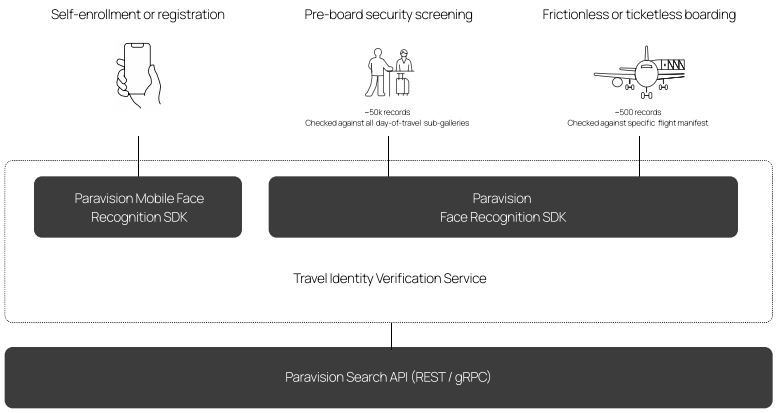


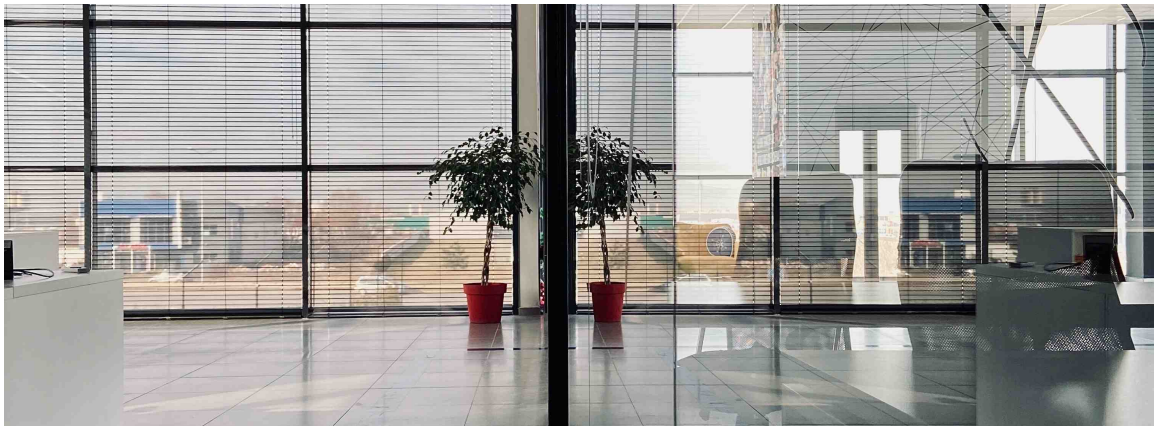
2. Air Travel

There are many applications for biometric identification and authentication within air travel, which Paravision has covered extensively in its white paper "Face Recognition and the Future Air Travel Experience." Here, we will look specifically at curb-to-gate "seamless" travel, where a traveler can drop their bags, go through security, and board an aircraft with their face as their boarding pass and ID.

The process starts with a registration and biometric opt-in process, typically at home, with a mobile device. The flight systems determine which travelers are coming to the airport on a given day, and load a series of day-of-travel sub-galleries associated with each flight. The system can deliver 1:N identification at bag drop or security (i.e. pre-board screening) but looking across all flight sub-galleries. Then, at boarding, 1:N can be performed even more precisely for the sub-gallery associated with the given flight. The sub-galleries can be wiped on a daily basis or as frequently as policy dictates.

This cascading approach to galleries helps to improve accuracy and efficiency. 1:N face recognition is performed only on the relevant data. The impacts of this approach can be seen in NIST FRVT Paperless Travel, among other places.

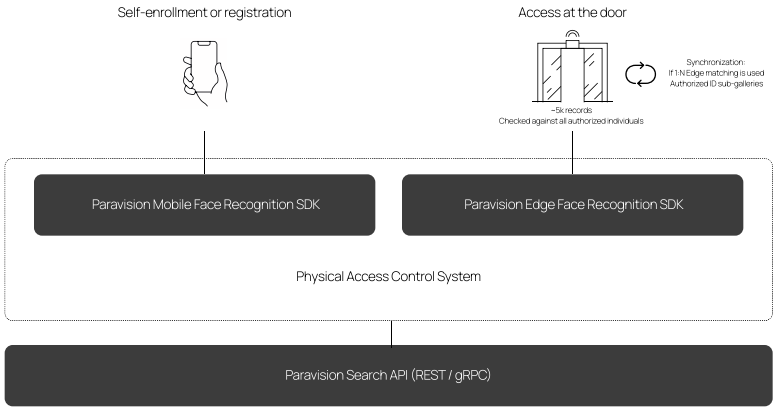


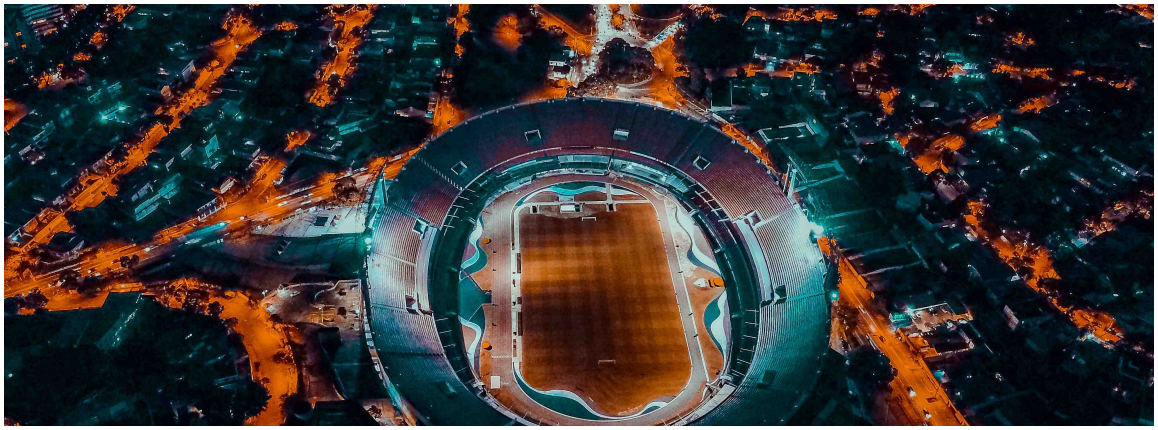


3. Enterprise access control

Just as face recognition has largely replaced the use of PINs as the primary means of smartphone security, it is poised to replace conventional access cards as the primary means of security in enterprise access control.

Typical enterprise physical access control systems (PACS) involve a fairly straightforward form of 1:N face recognition. As with the cases above, a person can opt-in and register their face through a mobile app via a selfie. This could optionally include some of the IDV workflow as described above. Faces are typically stored in an enterprise identity database. The enterprise PACS system then determines which identities might have access to which door, and can either (i) synchronize the appropriate sub-galleries with Edge PACS readers or (ii) perform 1:N matching in the backend from faces captured on edge devices. Here, the benefits of synchronizing the 1:N database or sub-gallery to the edge include reduced latency, improved user experience, and full uptime even in the case of network disruption.

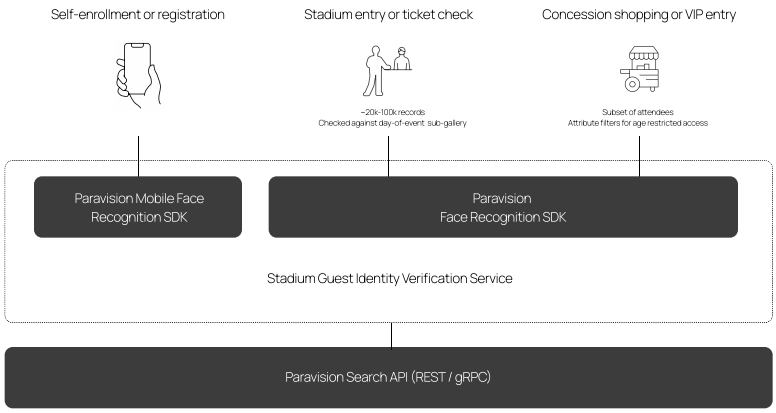




4. Sporting Events

Biometric authentication and identification are becoming increasingly common for sporting events as a way of offering superior fan experiences and limiting queue lengths and wait times. For fans who have opted in, access can be fast, touchless, and effortless. From a biometric perspective, the system architecture can be very similar to Air Travel as noted above: Fans register their identity, consent to use the system, and then their faces are uploaded to a day-of-event sub-gallery that optimizes speed and accuracy relative to a full database search.

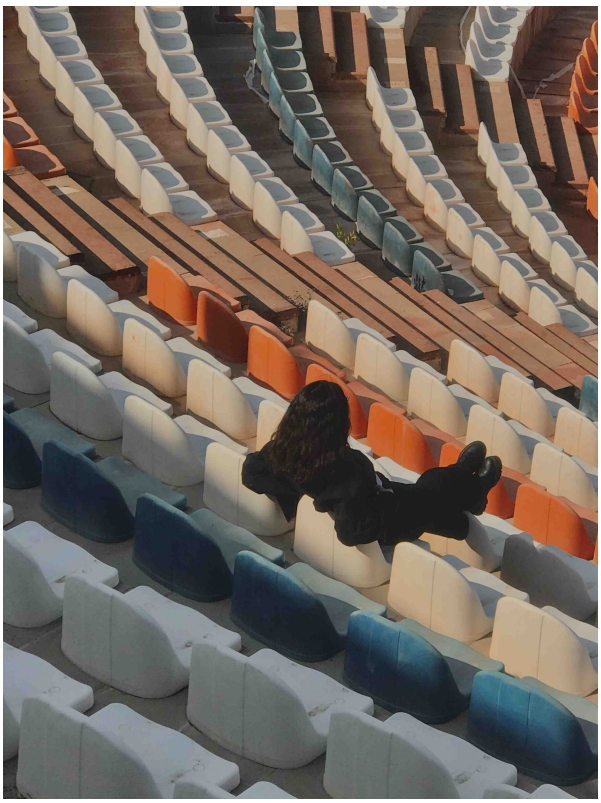
Use cases involve stadium access (with 1:N identification typically on the order of 5-50,000 records), as well as VIP lounge access and purchase of goods (especially those that are age restricted, such as alcoholic beverages). In addition to the optimization of a day-of-event sub-gallery, the system can use attribute filters (such as date of birth or VIP access) to enable further accuracy and speed benefits with the goal of delivering the best user experience for whatever the use case might be.



In Conclusion

The application of AI and ML to face recognition, as well as other technological advances, have expanded the use of face recognition to a variety of 1:N use cases that cannot be accomplished with other biometric modalities, enabling customer, employee, traveler, or fan experiences that were not possible a few short years ago.
Enterprises can implement these new capabilities to improve their business results, increase customer satisfaction, and ensure that the enterprises are operating properly.

With modular, open architecture systems and software that works on a wide range of leading platforms from Cloud to Mobile to Edge, Paravision is proud to help our partners realize these goals, regardless of the size of the enterprise database or the number of transactions the enterprise performs.



PARAVISION

Trusted Vision AI

For more information or to schedule a demo, please contact us at:

info@paravision.ai