Clearview AI

HISTORIES & ITERATIONS

Automating facial recognition has been challenging due to the complexity and meaningfulness of faces as visual stimuli, requiring research in computer vision, image processing, pattern recognition, and statistics.

Photo from RAND Corporation Archives

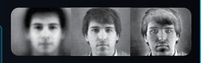
Kelly, Michael David. Visual identification of people by computer. Stanford University, 1971

viisage.com

VISIONICS CORPORATION

MIROS MOCEAN

On November 14, 2001, the Technology, Terrorism and Government Information Subcommittee of the Senate Judiciary Committee held a hearing on "Biometric Identifiers and the Modern Face of Terror: New Technologies in the Global War on Terrorism."



Google was processing upwards of 8 billion searches a year, and Facebook had connected over 150 million users worldwide. The combination of those twvo platforms and various others on the world wide web effectively overcame the dataset and preprocessing limitations of Woody's earlier work.

1969

Michale David Kelly produced a face recognition project at Stanford using three images to automatically extract head and body outlines and locate facial features.

The earliest research on facial recognition technology can be traced back to Panoramic Research, Inc. in California.

1970

The research was funded largely by the US Department of Defense and various intelligence agencies and was conducted by Woody Bledsoe, a co-founder of Panoramic Research.

Bledsoe developed the technique of "man-machine facial recognition," which involved manually entering the positions of facial feature points into a computer using a "rand tablet."

Panoramic's research datasets started with as little 122 photographs, growing slowly to 2,000 images over several years. Such data limitations were primarily due to the manual labor required to preprocess pictures as well as the lack of high quality publicly available images.

1991

Researchers in facial recognition became entrepreneurs and started companies to market facial recognition products.

The eigenface technique was developed at MIT Media Laboratory, digitizing a face and matching it against other images.

An image of Woody Bledsoe from a 1965 study.

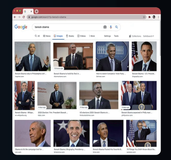
Photo from MIT Media Lab

MID 1990S

2001

The claim that facial recognition technology could have prevented the 9/11 terrorist attacks is a popular one, as it is believed that existing commercially available technology at the time could have quickly compared the surveillance footage of the attackers with photos of suspected terrorists.

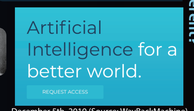
9/11 was a key moment for the conversation about Facial Recognition Technology.



2008

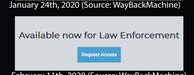
What's Different?

Distinctive features of Clearview's facial recognition technology include its ability to collect images of people's faces from various online sources, generate matches without head-on photos, and monitor searches within its software.



December 5th, 2019 (Source: WayBackMachine)

January 24th, 2020 (Source: WayBackMachine)



February 11th, 2020 (Source: WayBackMachine)

What's Groundbreaking?

Clearview AI collects images from the internet to create a facial recognition database for law enforcement use. Its technology is particularly useful in identifying suspects without prior criminal records. The Indiana State Police became Clearview AI's first paying customer after solving a case within 20 minutes of using the app.

INFRASTRUCTURE & STANDARDS

The National Institute of Standards and Technology (NIST), is a non-regulatory agency of the United States Department of Commerce. NIST's mission is to promote innovation and industrial competitiveness by advancing standards, technology, and metrology in various fields, including cybersecurity, biometrics, and engineering. NIST develops and publishes standards, guidelines, and best practices to improve the accuracy, reliability, and security of technology and measurement systems. Tests conducted by NIST have confirmed the superior accuracy and reliability of the Clearview AI facial recognition platform.

Security Standards--Encryption Procedures To Protect Data Subjects

1. Providing a Specific, Lawful Basis For Each Search Undertaken Using the System
2. Requiring the Preservation of Data & Metadata
3. Requiring Specialized Training to be Provided for All Users Authorized to Access the System
4. Prohibiting Purely Automated Matching, Requiring Investigative Process for Each Match

Two Different Ways to Understand Standards

Normative
- Ensuring Accuracy
- Preventing Discrimination
- Limiting Uses to What's Legal, Ethical, and in the Public Interest
  - Preventing Abuse
  - Protecting Human Rights

Technical
- Protecting Security
- Testing & Validation
- Protecting Against Misidentification
- Protecting the Privacy of Data Subjects
- Ensuring Accountability

Evaluation of Standards (Third Party Authentication)

Matching against a single record
Answers "Am I whom I claim to be?"

NIST FRVT 1:1

Given two face photos, the test evaluates the accuracy of a particular algorithm in correctly determining if they are the same person or not. Evaluation uses a database of mugshots, VISA photos, border photos, and WILD photos.

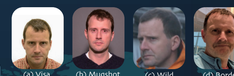
Matching against all records in the database
Answers "Who am I?"

NIST FRVT 1:N

Evaluation uses database of mugshots, webcam, profiles and WILD photos. Tests the accuracy of the algorithm to match a photo accurately out of over millions of other photos.
- Immigration-related images
- Law enforcement images

#1 IN U.S. & #2 IN WORLD

For the most difficult category"Wild Photos"



Clearview AI's submission to NIST's latest Face Recognition Vendor Test ("FRVT") ranked #1 in the U.S. for its performance in matching Visa Photos (99.81%), Mugshot Photos (99.76%), Visaborder photos (99.7%) and Border Photos (99.42%), and ranked in top five worldwide in all of these categories out of 650 algorithms.

Do Global Infrastructure and Legal Standards enable Clearview AI?

GDPR

Clearview AI violated Article 6 by collecting and using biometric data, and data access rights in Articles 12, 15 and 17. Also failed to respond to requests.

The General Data Protection Regulation (GDPR) calls itself the "toughest privacy and security law in the world" and protects EU residents from privacy and security violations.

Put into effect May 25, 2018.

BIPA

The Biometric Information Privacy Act (BIPA) protects Illinois residents from having their biometric data accessed without consent.

The 2020 ACLU settlement requires Clearview AI to restrict the sale of biometric data.

Meta has previously paid a $650 million fine relating to BIPA violations.

MOTIVATIONS & POWER PLAYERS

HOAN TON-THAT
CO-FOUNDER & CEO

"TALENTED PIONEER"

- Clearview AI raised over $7 million in funding from investors, deploying Republican officials to approach police forces.
- First paying customer: Indiana State Police.

- As a talented computer programmer and the founder of the company, Ton-That plays a critical role in Clearview AI.
- Ton-That managed software development and developed a vision for the company.

ABHINAV SOMANI
CHIEF OPERATING OFFICER

"THE QUIET PROFESSIONAL"

- Previously CEO and founder of Leverton AI.
- Experienced investment, financial, technology, business development, product development, and operations strategist.

Robinhood · Calm · coinbase · EPIC GAMES · DocuSign

RICHARD SCHWARTZ
CO-FOUNDER & PRESIDENT (?)

"SOCIAL CAPITAL BANK"

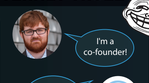
In 2016, Schwartz and Ton-That met at a Manhattan Institute event. Ton-That handled the development of the app while Schwartz leveraged his contacts to generate commercial interest.

- Schwartz played a crucial role in helping Clearview AI gain access to key decision-makers and stakeholders, ultimately contributing to the company's rapid growth and success in the industry.
- Schwartz was a Senior Advisor to Mayor Giuliani in the 1990s.

CHARLES "CHUCK" JOHNSON
DISPUTED CO-FOUNDER

"THE CONSERVATIVE INTERNET TROLL"

Ton-That met Johnson, who has been called an internet troll, in 2016 and attended the RNC convention together. That summer, Johnson introduced Ton-That to Peter Thiel, who provided the seed money. Two days later, Johnson introduced Ton-That to Schwartz.



From research project to commercial enterprise:
Began as a research project tool to help law enforcement, but pivoted to sell its technology to law enforcement and private companies.

Expansion of the client base and pivot to the international market
Expanded to countries such as Australia and Canada.

Legal and ethical challenges
Clearview AI faces growing criticism and legal challenges. Lawmakers, activists, and civil liberties groups raise concerns about ethical implications of using it for surveillance purposes.

PIVOTS

2020

EPIC complaint with FTC alleging that collection + use of facial recognition data without user consent is an unfair and deceptive trade practice in violation of Section 5 of the FTC Act.

Sued by ACLU violating BIPA

Fined by French data protection authority CNIL for $19.6 million for violating GDPR

Cease-and-desist letters from tech giants alleging the company violated their terms of service by scraping images.

Growing Competition
Faced growing competition from others offering facial recognition tech.

Partnerships and collaborations
Clearview AI has had several partnerships with law enforcement agencies which helped to expand its reach and influence.

USERS & NON-USERS

	WILLING	NON-WILLING	UNAUTHORIZED
USERS	Federal Law Enforcement/Entities, Public Defenders	N/A	Private Companies, Non-Law Enforcement
NON-USERS	People Who Can Affect How Their Data Is Used	People Who Cannot Affect How Their Data is Used, Law Enforcement Opt Out	N/A

	USERS	NON-USERS
Why they use the system?	Violent Crimes, Theft, Terrorism, Narcotics, Missing Persons, Human Trafficking, Fraud, Sexual Assault, Organized Crime	
How the group changed over time?	Department of Justice, Department of Homeland Security, Federal, Tribal, State & Local, Military	
What is needed to use the system?	Computer or mobile device	Permit to access to Clearview AI
What will make their use of the system unpleasant?	Wide opposition voice from the public: Privacy concerns; Accuracy and bias; Civil rights violations; Lack of regulation	

WHO CAN'T USE THIS SYSTEM?

In May 2022, under the terms of an ACLU settlement, Clearview AI agreed to a permanent ban from selling its facial recognition database to private companies.

There are no federal laws governing the use of facial-recognition technology, which has led states, cities, and counties to regulate it on their own in various ways, particularly when it comes to how law enforcement agencies can use it.

WHO IS REJECTING THIS SYSTEM?

Where companies stand on sale of face recognition:

Google IBM	Committed to not provide face surveillance producvts to local and federal law enforcement
Microsoft Amazon	Committed to not provide face surveillance products to local law enforcement
Clearview Dataworks Idemia, NEC	Has not committed to stop selling face surveillance products

Localities that have stepped law enforcement use of face recognition:

ADDITIONAL TYPES OF NON-USERS

RESISTERS
People who want to be moved from Clearview AI database but can't legally request this.

REJECTORS
People who can ask to be removed from database (California, Virginia, Illinois). Organizations who reject the sale of facial recognition.

EXCLUDED
People who have avoided getting their face captured excluded from database. The general public cannot be granted access.

EXPELLED
People who have been removed from database and/or private companies.

WHAT'S IT MADE OF?

CLOUD STORAGE

Servers: Computers that store the data in cloud storage.
Storage arrays: Devices that connect multiple servers together.
Network switches: Devices that connect the servers and storage together.
Routers: Devices that route traffic between different networks.

Firewalls: Protect the servers and storage from attacks.
Load balancers: Distribute traffic evenly across the servers.
Domain Name System: Translates domain names into IP addresses
Network Operator: Communications service that controls network infrastructure.

Billions of Images from Social Media Platforms

The Public Internet

Online Mugshots and Other Criminal Databases
Publicly Available Blogs
Open-Source Social Media Posts
National & Local News Websites

HOW'S IT THINK?

A PATENTED PROCESS

MAIN PROCESS

Face image capture and transmission
- Facial image capture using network camera or onboard camera using mobile devices
- (Optional) Image processing on mobile devices (face detection, cropping, etc.)
- Transmit images to a remote server (e.g., cloud) for further processing and facial recognition

Facial recognition
- Facial recognition based on neural network (e.g., deep convolutional neural network)
- Match facial images with those in databases (e.g., crime database) based on nearest neighbor classifier

Personal information output and delivery
- Retrieve personal information or online profile of the person with matching facial images
- Transmit and display the information on mobile devices or send alert message (when crime databases are used)

SEARCH ALGORITHM

Stage 1 - User interface
1. Take an input photo
2. Crop out the face
3. If there are multiple faces, user selects one

Stage 2 - Search Process
1. Face image is turned into a vector with neural network
2. Vector is searched across the NNDB cluster
3. The matching face identifiers are looked up in SQL database for metadata

Stage 3 - Processing
1. Extract canonical links (profiles, social media)
2. Extract the name associated with that face
3. Extract age, biography, education, location, etc

CRAWLING ALGORITHM

Stage 1 - Web Crawler
1. Web crawler is seeded with thousands of starting URLs
2. Each image is extracted and put on the download queue
3. Each link on the page is followed. Duplicate links are tracked so that we don't end in an infinite loop

Stage 2 - Face Extraction
1. Each image is downloaded and checked for faces
2. Each face rectangle is stored with the image id in SQL
3. If the image has a face, the metadata like title, description are stored in SQL

Stage 3 - Processing to Databases
1. Face image is sent to the GPU cluster and turned into a vector
2. Face vector is added to the NNDB
3. Metadata: title, description and name added to SQL

TRAINING DATA SETS AND BIAS:

The Clearview AI algorithm was trained on a large, diverse, custom-built dataset, as opposed to pre-packaged training galleries, which may favor a certain gender or ethnicity and impact the accuracy of the results. As a result, the racial bias and accuracy disparities that have affected other applications are substantially eliminated.

CONVOLUTIONAL NEURAL NETWORKS (CNN)

Convolutional Neural Networks, are a deep learning algorithm that has a specialization in picking out or detecting patterns in images. This process detect patterns actual faces and biometric identification.

First, it uses multiple filters across multiple layers to detect simple objects like specific edges, corners, and shapes.

In later layers, the filters detect sophisticated objects like eyes, ears, noses, etc.

A filter converts all the pixels in its receptive field into a single value, it does this for the whole image before creating the output.

The output is used by another layer and filter, repeatedly across pixels in the entire image.

WHAT'S IT KNOW?

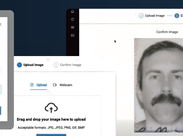
Clearview collects publicly available photos and information to:
-Provide their Products and Services
-Improve their Products and Services
-To train their algorithms

Photo Scraping the World Wide Web → Cloud Based Data Storage → Clearview AI API

Category of Personal Information
Face vectors and photos, and such metadata as image files may contain (sensitive personal information).

Sources
From the Internet

Disclosure and Recipients of Personal Information
Clearview AI may have sold this category of personal information to law enforcement, governmental agencies, security and national security professionals. Clearview does not provide any third party with access to face vectors Clearview produces.



Required In-Take Form is completed Crime Type Case Number

Image or Face is uploaded into the API

Search is performed

Facial image results returned with source image URL

METAL & BITS