SSP 100 - 10
How Computers See

11 / 05 / 2026

Instructor: Michael Eckmann

Today's Topics

- Questions/Comments?
- Face recognition
  - History
  - Current state-of-the-art techniques
  - Some issues with certain datasets

Face Detection

- Face detection – determine if and where faces are in an image. Draw bounding box around them.

Face Detection example

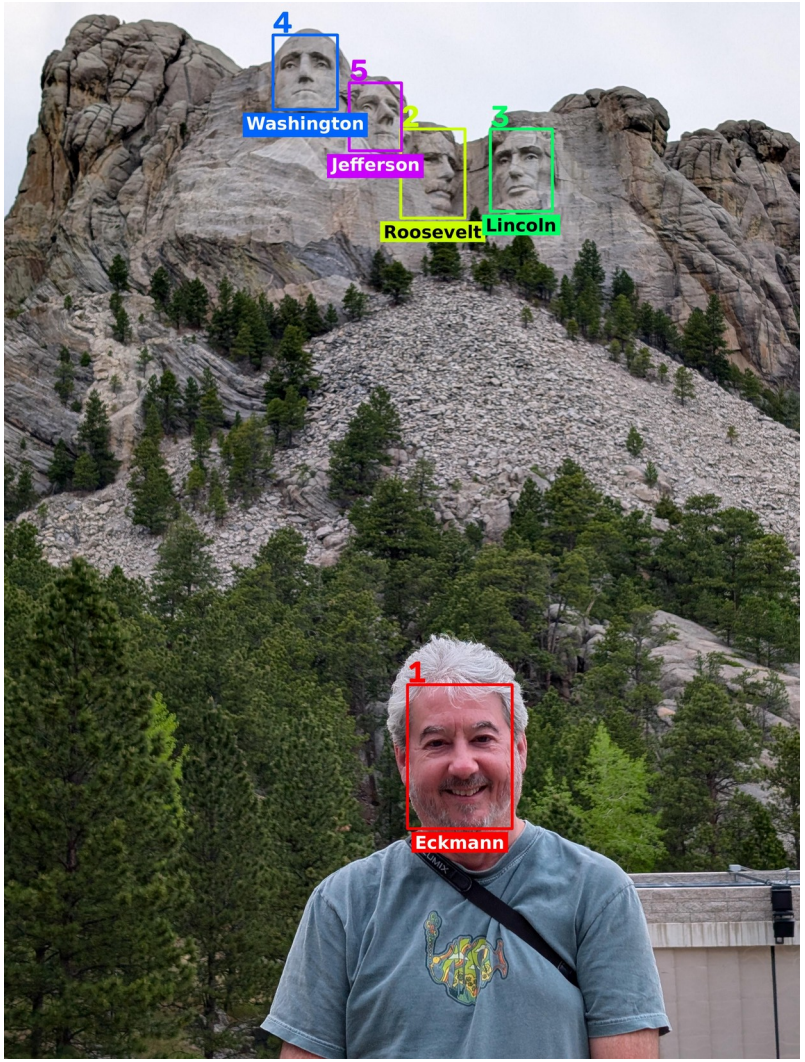


Michael Eckmann - Skidmore College - SSP 100 - Fall 2026

Face Recognition

- Face Recognition – determine who the face belongs to

Face Recognition example



Michael Eckmann - Skidmore College - SSP 100 - Fall 2026

Face detection uses

- Besides as a first step for face recognition, what other uses are there for face detection?

Face detection uses

- Besides as a first step for face recognition, what other uses are there for face detection?
  - Cameras typically have built in face detection and use it to make sure camera settings like focus, etc. are set to image the face well
  - Can be used for adding augmented reality affects to faces like cat ears, etc.

Thoughts on solutions

- Thoughts on how you would attempt a solution (or how you think researchers may have tackled these problems) to face detection and face recognition.

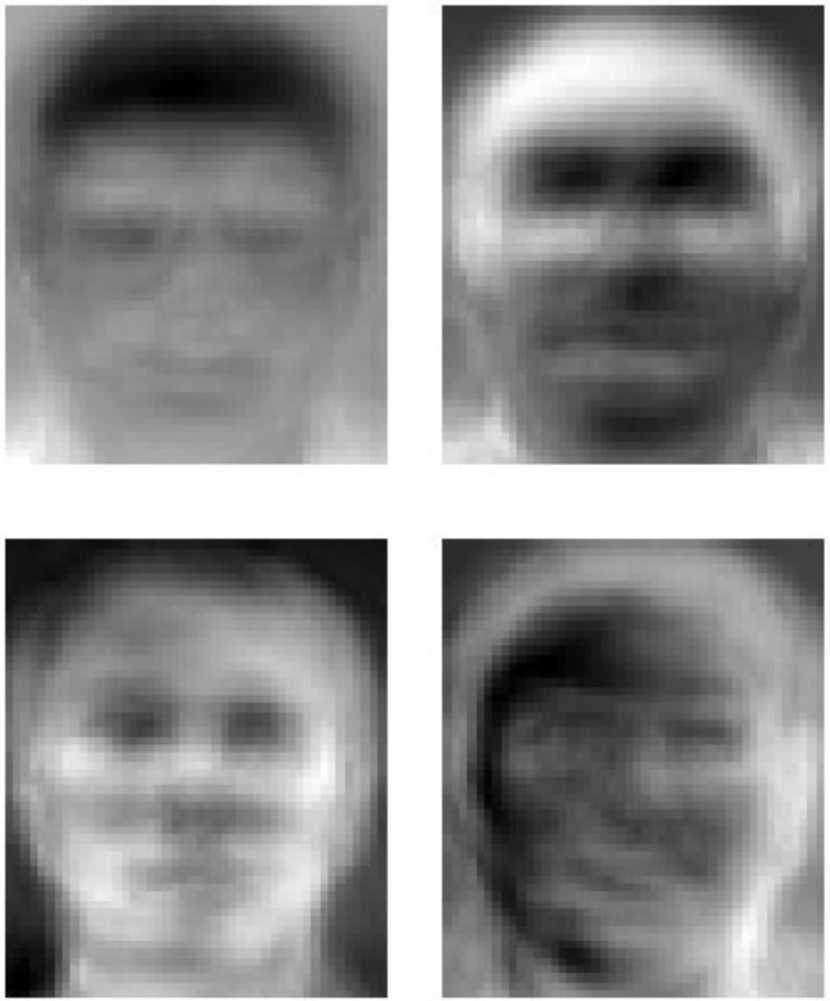
Early attempts

- 1960's Bledsoe, Wolf, and Bisson - human marked small number of landmarks/features (e.g. corners of face parts) on an early tablet of face images and compared distances between them to determine who the face was for a small set (tens of identities).
- Takeo Kanade in early 1970s, first to automate the detection of landmarks and tested system on hundreds of images and tens of identities with modest accuracy of recognition. Big improvement was automated feature detection rather than human required marking.

Later attempts

- The idea where functions can be represented as a linear combination of so-called basis functions (e.g. Fourier transforms).
- Late 1980s to late 1990s, researchers represented faces with a linear combination of the most significant basis faces.
- Early 1990s attempts had trouble with lighting changes, shadows, etc.
- Later solutions were more robust to lighting changes, etc.

Some eigenfaces from AT&T Laboratories Cambridge



Later attempts

- In early 2000's Viola & Jones made detection fast enough for real-time performance.
- Digital cameras were able to have Viola-Jones algorithm for face detection running on them to show a box around faces in the view finder before picture is taken.

Later attempts

- Early 2000s until about 2013 further improvements and different attempts were made for face recognition
- Such as hand crafted/engineered features (LBP, Gabor, SIFT, SURF) applied to faces that are then compared for similarity.
- Recall those SIFT features I talked about over a month ago when discussing the history of computer vision and the demo of matching points in images. Let me show you that demo again.

Last 15 or so years

- Convolutional Neural Networks trained in a supervised manner.
- Images of faces with known identities used for training --- that is supervised learning because the data (images of faces) and labels (known identities) are provided.
- Number of output nodes in the network would be the number of identities the network can detect.

Some datasets (table generated by Claude)

Current mainstream training sets (real, web-scraped)

Dataset	Identities	Images	Imgs/ID	Derived from	Role today
MS1MV2	85K	5.8M	~68	MS-Celeb-1M, cleaned	Long-time standard baseline; most papers still report results on it
MS1MV3	93K	5.2M	~56	MS-Celeb-1M, re-cleaned	Cleaner drop-in replacement for MS1MV2
Glint360K	360K	17M	~47	MS1M + Celeb-500K, merged and cleaned	InsightFace's large set (2021), popular for big-backbone training
WebFace4M	200K	4.2M	~21	10% subset of WebFace42M	Very widely used now; same scale as MS1M but trains better models
WebFace12M	600K	12M	~20	30% subset of WebFace42M	Mid-scale option; AdaFace releases checkpoints on it
WebFace42M	2M	42M	~21	WebFace260M, auto-cleaned (CAST)	Largest public cleaned training set
WebFace260M	4M	260M	~65	Raw web scrape	Noisy source pool; not trained on directly

VGGFace2

- VGGFace2 9.1K 3.3M ~362 2018
- By researchers in Visual Geometric Group (VGG) at University of Oxford
- Dataset itself was retracted in 2019 and is no longer publicly available, but models trained on this data still are in use
- Dataset was created from scraping Google Image Search mostly for celebrities, politicians, etc.
- Let’s look at some figures in the paper.
- Why do you think the dataset was retracted?

VGGFace2

- Retracted due to ethical concerns of lack of consent and privacy concerns mainly.
- As with any pretrained model, there are also concerns about bias and under or over representation of certain groups which could lead to more errors with recognizing people of those underrepresented groups.

MS-Celeb-1M

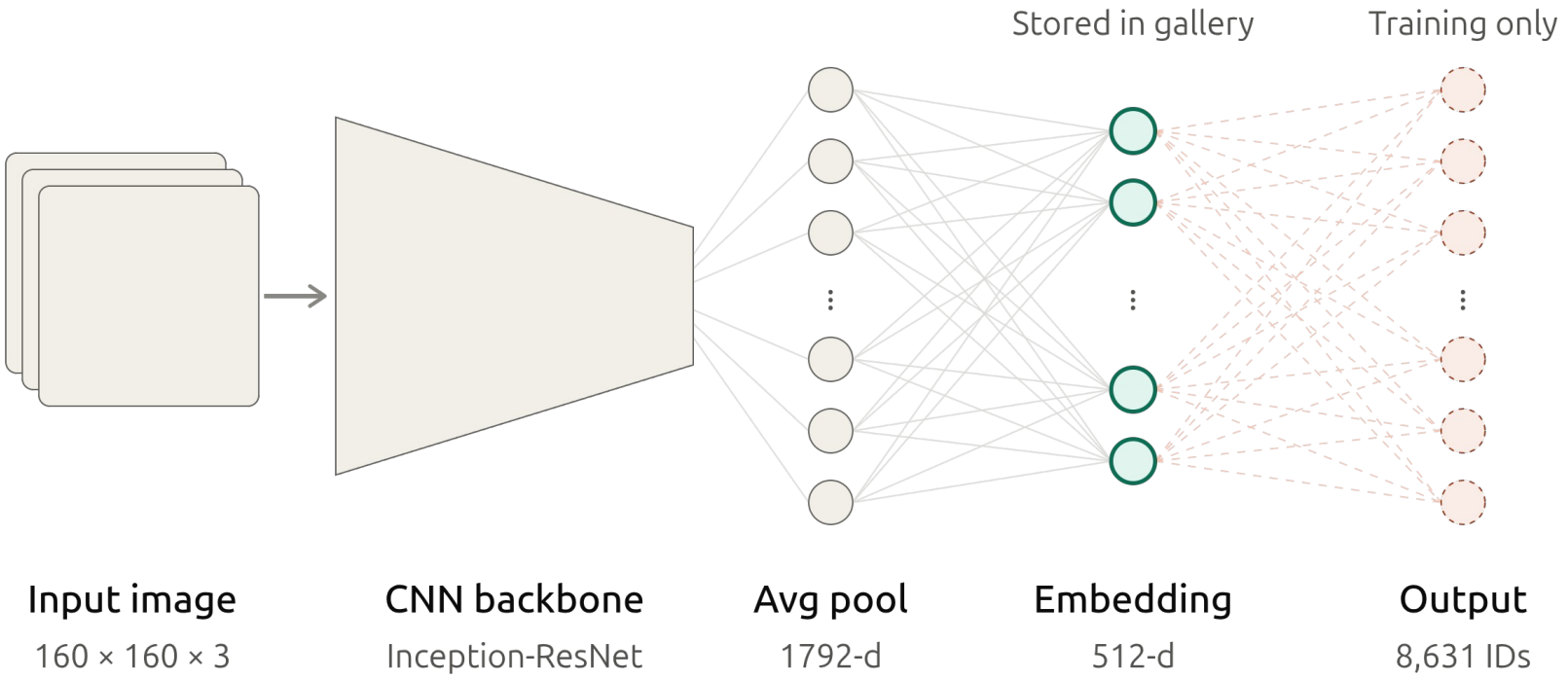
- MS-Celeb-1M by Microsoft also was taken down after complaints that it included other public figures that are not celebrities.

- **https://exposing.ai/msceleb/**

You too can do face recognition!

- How?
- Publicly available pre-trained models exist
- Take the model's trained weights up to but not including the last layer. The output before the output layer is called an embedding.

You too can do face recognition!



You too can do face recognition!

- Take a pretrained model that converts images of faces into an embedding (exclude the weights after the penultimate layer and the output layer).
- Run your dataset through it to compute an embedding per identity.
- Then send a test face image through to compute the embedding of that face and compare the embedding to your database of embeddings.

Enrollment

- Enroll your dataset
  - Have a small number of reference images per person
  - For each person:
    - Compute the embedding of each reference image of a person
    - Store a normalized average embedding per person

Probe & Match attempt

- An image of a person is captured
  - Detect the face within the image
  - Do some processing to the face portion of the image (e.g. straighten it) and make it the correct dimensions and number of pixels required (e.g. 160x160)
  - Compute the embedding
  - Compare this embedding to the known embeddings for recognition

Matching

- Think about potential answers a probe could give.

Matching

- Think about potential answers a probe could give.
- Face will be recognized as someone in your database, but what if the person detected isn’t in your database?

Matching

- Think about potential answers a probe could give.
- Face will be recognized as someone in your database, but what if the person detected isn’t in your database?
- An unknown person should not be recognized incorrectly as a person in your database.
- This is the issue with closed set recognition.
- Open set recognition should allow an additional output label of “Unknown”.

Matching

- For open set recognition to allow an additional output label of “Unknown” what is typically done is to first get the highest matching score for a person within the database.
- If this highest matching score is above some preset threshold, then answer that person identity.
- If the highest matching score is below the threshold, output “Unknown.”
- There are other techniques done when that simple way doesn’t work.

Matching

- For a face recognition system you'd want to test it and record a variety of different errors:
- FPIR – False positive identification rate
  - when non-enrolled people are identified as an enrolled person
- FNIR – False negative identification rate
  - What do you think this could mean?

Matching

- For a face recognition system you'd want to test it and record a variety of different errors:
- FPIR – False positive identification rate
  - when non-enrolled people are identified as an enrolled person
- FNIR – False negative identification rate
  - when enrolled people are not identified correctly

Discussion

- Group discussions