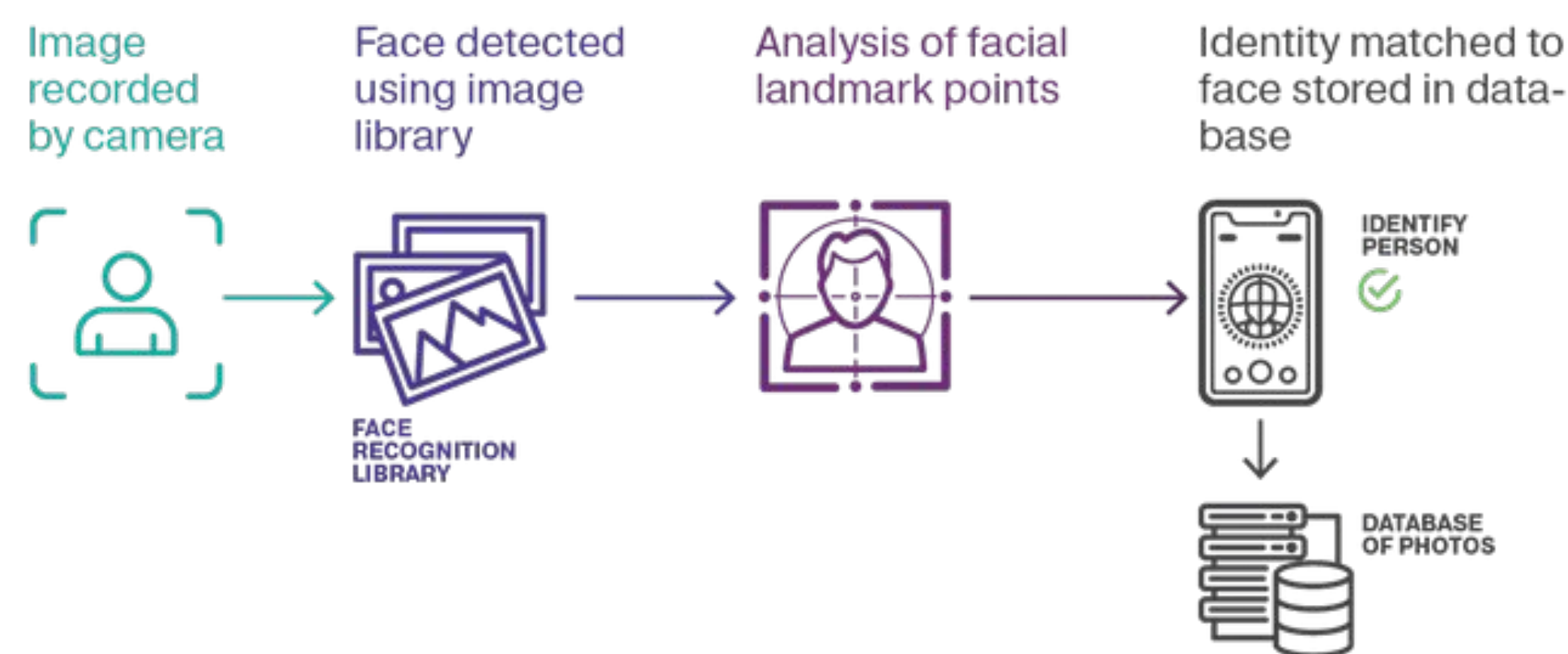# Facial Recognition Algorithms

Hee So Kim, Marisa Papagelis, Amy Fung

CS 231 | Fall 2021 | Final Poster

WELLESLEY

Background

- **Facial recognition** is a machine learning technique that learns and extracts the physical characteristics of humans and matches their features with tested images.
- Two main types: still images & live video


Motivation

- Common interest in facial recognition algorithms → compare 3 different ones
- Wanted to learn more about potential bias and ethical concerns of utilizing these algorithms → we see them used every day

Ethics

- Privacy concerns → (1) most facial recognition algorithms are trained on large datasets of images collected without (informed) consent; (2) data breach
- Crime → blurry line of when it is okay to use → terrorist under search warrent vs. suspected individual for a petty theft

(1) Eigenfaces

- PCA to simplify face features using eigenvectors → called "eigenfaces"
- Each face has a unique set of weights
  - Mean face + the set of weights
- Matching by comparing the set of weights
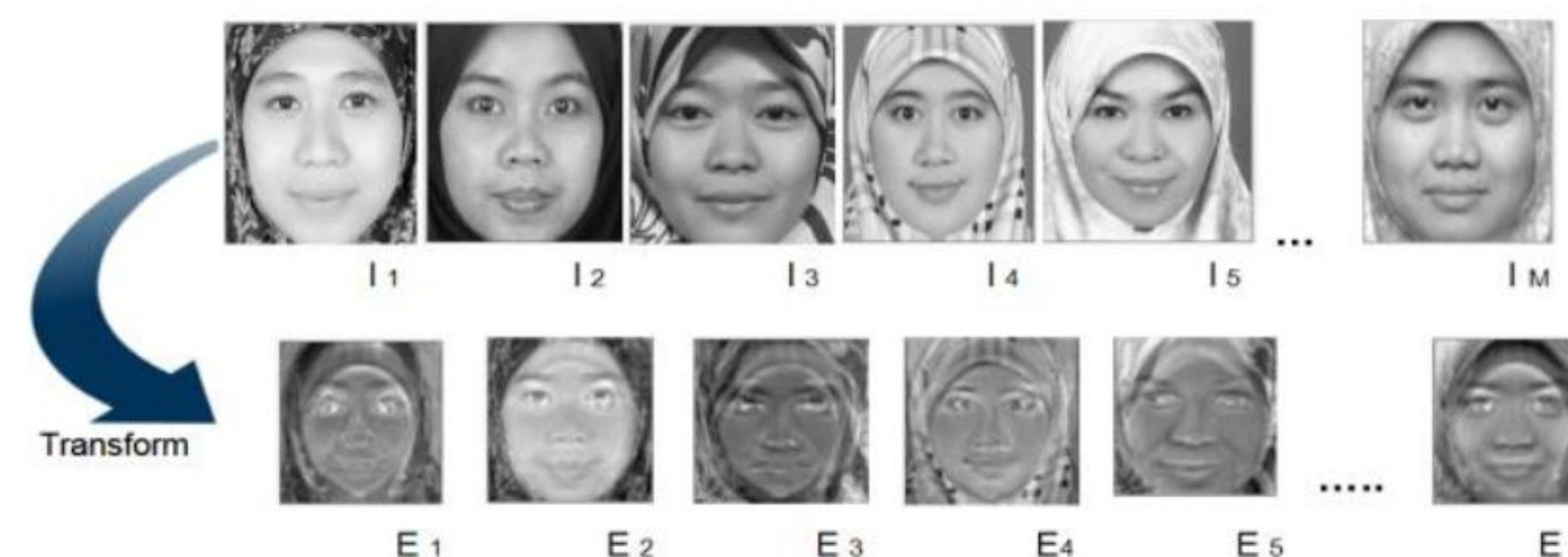
(2) Local Binary Patterns Histograms (LBPH)
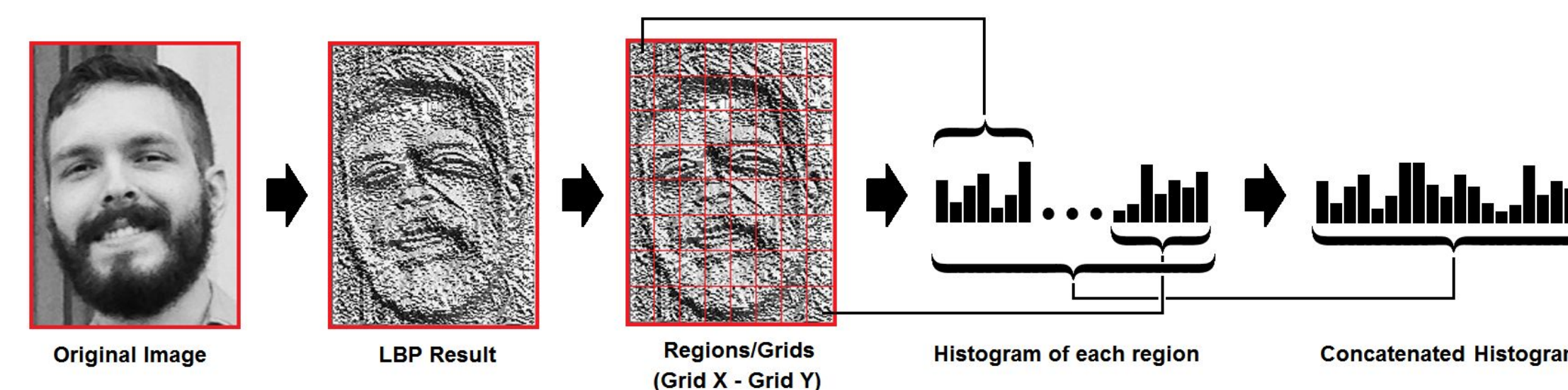
- 4 parameters: radius, neighbors, grid X,Y
- Simplifies image using binary values
- Divides simplified image into multiple grids
- Creates a histogram for each grid representing intensity patterns
- All histograms are concatenated → final histogram is matched to histograms of images in training dataset
- Confidence value to show correctness

(3) Convolutional Neural Network (CNN)

- Learns independently the images' features by layer with feature maps that convolve around the images
- Convolutional, pooling, and fully connected layer types - each layer composed of learnable filters
- Much lower pre-processing
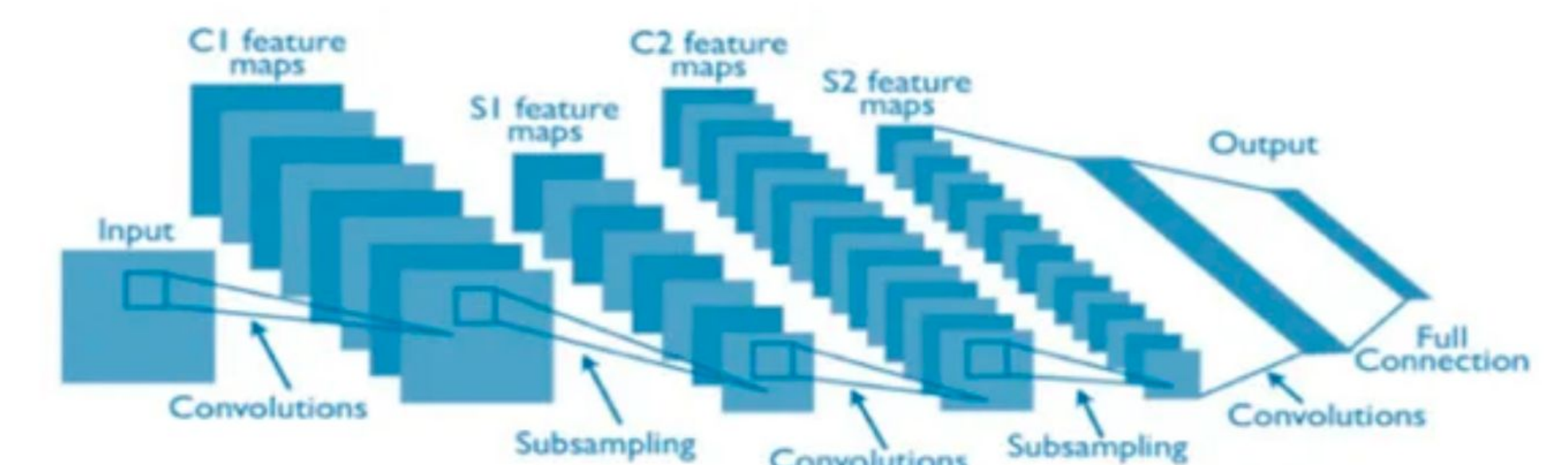
Similarities & Differences

- Eigenfaces vs LBPH: differences in feature extraction methods, similarities in matching methods
- CNN downsamples the image whereas Eigenfaces and LBPH uses all pixels
- Features from LBPH can be used in CNN (ex. LBP → CNN)

Bias

- Bias against people of color and women, especially women of color
- If input is biased, algorithm will also be biased (lack of input data)