

# Day 9: Vision Capstone

## Image Classification Goal and Notes

- Face or not face
- Ada, Charles, or someone else?
  - Extraction, Standardize size
- Cross-entropy loss
- Using face descriptors
  - A vector that represents a feature of the face
  - Steps to obtain:
    1. Initial convolution
    2. Convolution/Max Pool
    3. Flatten
    4. Dense (feedforward/full-connected/"standard" NN)
    5. Return resulting array
  - Triplet Loss
    - $\vec{d}_0 \cdot \vec{d}_1 > \vec{d}_0 \cdot \vec{d}_2$
    - $\vec{d}_0 \cdot \vec{d}_1 > \vec{d}_1 \cdot \vec{d}_2$
    - Geometric:  $\cos(\vec{d}_0, \vec{d}_1) = \cos(\vec{d}_0 \cdot \vec{d}_2)$
- **facenet**
  1. Face detection: multi-task cascaded convnet (MTCCN)
  2. Face description: ResNet
    - These two tools will return a list of bounding boxes (shape: (N, 4)) as well as face probabilities (shape: (N)), landmarks, and descriptor vectors (returned by the ResNets)

# Capstone

- Database
  - name → profiles
    - name
    - descriptors for person (need to be trained to extract)
  - Face clustering algorithm
    - Chinese Whispers