

ENGN4528 Project Proposal

Topic: Face detection and recognition.

Group members: Kuangyi Xing (u5817313), Sam Toyer (u5568237); both undergraduate

Proposed approach: We intend to tackle project option 3: write a program which can detect which students are present or absent from a lecture, given separate photos of each student and a single photo of the lecture theatre. Our proposed pipeline for this task consists of several stages:

1. Apply Viola-Jones face detector to extract faces of students in the group into separate images.
 - a. **Extension:** replace Viola-Jones detector with a CNN-based approach, most likely YOLO (Redmon et al., 2016).
2. Align faces so that prominent features appear at a fixed location in each cropped image. Initially, we will simply use eye locations found with another Viola-Jones detector.
 - a. **Extension:** as with the face detector, replace the eye detector with a CNN.
3. Decompose each aligned face into constituent eigenfaces to get a set of small feature vectors.
 - a. **Extension:** instead of obtaining a feature vector using eigenfaces, use the latent representation produced by a variational autoencoder (Kingma & Welling, 2013) as a feature vector. Novel strategies will be required to ensure that this latent vector primarily contains information relevant to the person's identity, rather than encoding irrelevant features of the background, etc. One group member has a fair amount experience with variational autoencoders, and we have some strategies in mind for approaching this problem, but will have to experiment to see which works best.
4. Compare features extracted from the image with features corresponding to faces in the student profile dataset to associate a student with each detected face.

Note the items marked "extension": initially, we intend to implement the simplest system we can. As we come to understand the weaknesses and strengths of the system, we will choose extension tasks to implement, replacing the original components. We hope that this will allow us to turn in a reasonably ambitious assignment at the end of the term, without running the risk of being unable to turn in a working system after spending too much time on novel components.

References:

Kingma, D., and Welling, M. "Auto-encoding variational Bayes." *arXiv:1312.6114* (2013).

Redmon, J, et al. "You only look once: unified, real-time object detection." *CVPR*. 2016.