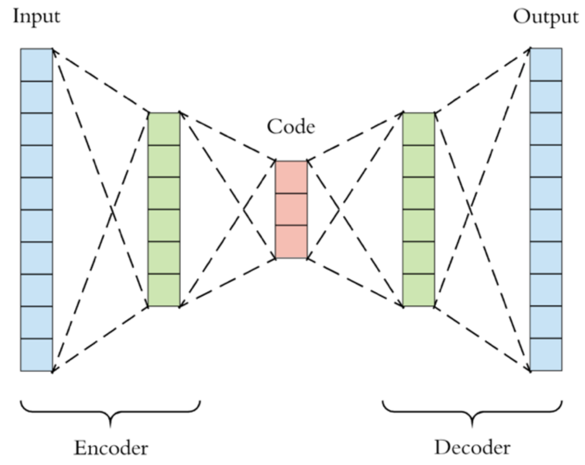

Artificial Lineup Facial Generation

Case Study: Northwestern MSAI

by Brandon Harris, Harper Pack, and Ikhlas Attarwala

An Overview

Eyewitnesses play an important role in the justice system. Testimony from an eyewitness can influence the direction of an investigation, shape the contours of a prosecutor's arguments, and - perhaps most importantly - sway the hearts and minds of jurors hearing a given judicial case. For better or for worse, eyewitness identification *acts* like proof.

"… there is almost nothing more convincing [to a jury] than a live human being who takes the stand, points a finger at the defendant, and says 'That's the one!'"
- U.S. Supreme Court Justice William J. Brennan

The Challenge

The United States has the world's largest prison population, both gross and per capita. The Innocence Project, a non-profit dedicated to exonerating the wrongly convicted, estimate approximately 20,000 inmates are wrongly incarcerated at this time. Each year, they estimate, the problem grows: between 1% and 5% of new convicts are innocent.

Further, the Project has found that eyewitness misidentification is the leading cause in over 70% of wrongful convictions, especially in the many cases where forensic evidence, such as DNA or fingerprints, was initially lacking or otherwise ambiguous.

Despite its shortcomings, eyewitness testimony is evidence, and thus will continue to occupy a prominent place within the criminal justice system.

Our Solution

Our team developed a solution with the potential to greatly reduce misidentification by presenting witnesses with a novel system that addresses many concerns regarding the proper procedural administration of pre-trial identification.

With recent advances in artificial intelligence and computer vision, data scientists have been able to generate high-fidelity, false portrait images that are indistinguishable from pictures of real humans. Generative neural networks, such as variational autoencoders, compress images to a smaller-dimensional space, allowing for randomization of representations that can be extrapolated and reconstructed into full images of brand-new faces.

Our method of pre-trial identification is similar to that of a photo array, where law enforcement personnel show sets of photographs to an eyewitness to ask if they recognize anyone from the photo set. Unlike a photo array that only uses real images, our solution contains a much larger majority of realistic, computer-generated portraits & mugshots.

First, the system takes a headshot of the suspected perpetrator. From that image, the system generates synthetic headshots that each bear a resemblance to the suspect. Finally, the system de-identifies the images, to guarantee that the lineup is blind for both the administrator and the participant.

When the eyewitness interacts with the system to review the lineup and possibly make an identification, they will read a stock set of instructions on-screen, which comply with the improvements we discuss in the following section.

This system can produce improved eyewitness lineups on-the-fly without an expensive or powerful system; common computer hardware is powerful enough to run the tool. This will yield faster lineups to speed investigations along, lead to fewer misidentifications giving prosecutors more confidence, and is best practice compliant to ensure investigative procedures are defensible and robust.

Improvements

The *National Center for State Courts* has recommendations for how to reduce bias and ensure fairness and validity in identification procedures, and from it the following standards of improvement were created for this system to abide by:

1. Lineups should be created for 1 suspect at a time.
2. Lineups should present generated images sequentially, one at a time.
3. Lineups should consistent of individuals bearing a close resemblance to the subject.
4. Administration of the lineup should be double-blind.
5. The witness should receive unbiased, standardized instructions, such as including that the suspect may or may not be present in the lineup.
6. Upon making an identification, the witness should document their confidence with their decision.

We believe each of these standards are critical to ensuring an optimal solution at reducing the false positive error rate known as misidentification.

Limitations

There are a few limitations that are inherent to this system. The most important is that procuring high-fidelity images is costly, both in terms of time and data. Additionally, dimensions within the latent space may not map to facial features as humans would categorize them.

Considerations

We spent time examining where this system could be improved further. Our system is predicated on a commitment to the integrity of the investigative process and the rights of individuals. As such, its role should be that of a tool, not an entity, such that it shouldn't influence or shape the investigation beyond its instrumentality. Our system must provide representative lineups regardless of the suspect's demographics or individual appearance so as to stay from demonstrating or affecting bias. Finally, the system must not compromise the privacy of the individuals whose facial data trained the system, which could invite legal repercussions or produce the appearance of impropriety.

Future Work

Given the duration of this project, we trained solely on white, male portrait images, and so of course we must acquire more data that includes a larger distribution of skin tones, facial structures, genders, and portraits with atypical facial features. A more flexible system that can take in a wider range of input images including variations in light, facial angle, expression and composition would also yield better results, as would a more flexible output to generate lineups with features such as beards, piercings, tattoos, etc.

We created a very simple architecture to reduce computational load, but with longer training and a deeper network, we could yield higher-fidelity results. Finally, further experimenting with the architecture of the VAE may prove useful; recent research has shown that progressively training from lower resolutions and adding higher resolutions with each layer solve problems commonly associated with generating high-quality, large images that appear more believable and realistic.

More on VAEs

Autoencoders (AE) have two main phases. In the first, they encode images into a smaller, low-dimensional vectorized "latent space", that learns latent features about the image. Afterwards, a decoder takes a point in the latent space and generates an image from it with the features encoded by the latent vector. In a variational autoencoder (VAE), a special loss function enforces meaningful structure in the latent space. This is so points in the latent space that do not correspond with training inputs can still be extrapolated to reasonable images. Proximity in the latent space corresponds approximately with having similar features in the decoded image.