# Data Augmentation by Using Object Shape Reconstruction
# Popular Scientific Summary

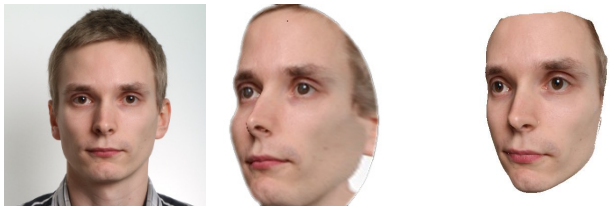
Filip Johannesson
Mats Granholm

***Introduction*—In many machine learning tasks it is important to have large datasets in order to create good models, especially for neural networks. Google's FaceNet is one of the more extreme examples, where over 100 million images were used to train the classification model. Manual data-gathering for datasets of these magnitudes can be very time and resource consuming, so a lot research is conducted in different augmentation techniques, where a small dataset can be expanded with synthetic images. In this thesis, several advanced augmentation techniques are investigated and evaluated on the task of face recognition.**

Existing data augmentation methods revolve around manipulating face images in 2D, such as rotating or flipping the image, in order to create more images. These methods are easy to use and robust. However, they tend to produce images that look too similar which is not ideal for training face recognition models. Our proposed solution is to utilize more advanced image synthesis schemes that can provide more variation.

Three synthesizing techniques are evaluated. Two of them are based on 3D models, and the third utilizes a neural network to generate images. The first model employed is called the *Surrey Face Model* (SFM). This is a so-called morphable face model, which is a 3D-model constructed in such a way so that major characteristics of the face shape and expression can adjusted by a small set of parameters. The second model is dubbed the *Million Faces Model* (MFM) and utilizes very dense and pre-posed 3D-models. Examples can be seen in Fig. 1. The third model is called a *Generative Adversarial Network* (GAN) which is capable of generating realistic looking face images from a short set of parameters. Images generated using the procedure can be seen in Fig. 2.

The augmentation techniques utilizing the SFM and the MFM are capable of improving the performance of face recognition in a predictable way. In particular these methods could be useful in order to improve performance of face recognition in surveillance camera footage, as captured images can be compared against similarly posed synthetic images. However, the augmentation techniques used are quite complex and run the risk of failing at times, which can reduce recognition performance. Creating more robust augmentation procedures is a key component for future work in this area.



(a) Original (b) MFM (c) SFM

Fig. 1: 3D-based image synthesis.

Fig. 2: Images generated using the GAN.

The GAN is capable of producing natural looking face images. However, it is difficult to control the appearance of the generated image which makes data augmentation problematic. A simple method to solve the problem is proposed, but in order to achieve state-of-the-art performance another method is required. Discovering ways to control the output of the GAN is another possible path for future work.