Recognizing a face based on its facial sketch along with the natural language description is an important, yet challenging problem in the face recognition community. We propose to translate faces into large scale photo database by generating synthetic facial images with descriptive attributes and minimal facial sketches.

**Contributions:**
1. To the best of our knowledge, this is the first attempt to perform face image generation and recognition jointly utilizing the information provided from both sketch and descriptive attributes.
2. We propose a generative model (MMC-GAN) to generate photorealistic facial images from sketch and high level descriptive information. Achieving facial attribute manipulation given a specific attribute value.
3. Face recognition can be implemented effectively based on the synthetic faces since we preserve the facial structures and local details during the generation.

**Methods**

**Multi-Modal Generator**

The generator integrates the minimal sketch and descriptive attributes into a hidden incorporation, which complements each other for face synthesis. Then the global and detailed information on a facial sketch is transcribed into synthetic faces while preserving the facial structures and local details during the generation.

The local detail generator recovers landmark details with skip connections, and the global generator integrates the minimal sketch and descriptive attributes into a hidden incorporation, which complements each other for face synthesis. Then the global and detailed information on a facial sketch is transcribed into synthetic faces while preserving the facial structures and local details during the generation.

**Conditional Discriminator**

The conditional discriminator produces an output vector representing the probability of the input image being real or fake, considering the facial attribute labels.

**Results**

Figure 4: Synthesis results based on the proposed MMC-GAN conditioned on different attributes. Given sketch and different attributes, we only flip one attribute value for each synthetic face image. The labels are: Black hair, Blond hair, Bald, Bangs, Beard, Eyeglasses, Hat, Old, Gender, respectively.

Verification accuracy(%) of the proposed MMC-GAN and compared methods in Sketch-LFWA.

**Conclusions**

An effective Multi-Modal Conditional GAN (MMC-GAN) is proposed to generate photorealistic facial images from sketch and high level descriptive attributes. Simultaneously, we make the global structures and local details of a face respectively join in different channels of the generator, and then fuse them together with the guidance of the corresponding descriptive attributes, which achieve better transferability and controllability. Experimental results demonstrate the effectiveness of the proposed MMC-GAN in tasks of sketch recognition and generation.