Generating Counterfactual Visual Explanations Using GAN

1. Generate images using GAN
   
   As shown in the following figure, using GAN can partially change the image content. So it's possible to change part of the bird pattern in the picture.

   ![](input-image)

   ![](disgusted-image)

   ![](fearful-image)

   ![](happy-image)

   ![](bird-image)

2. Use class definitions to generate images for rare classes

   Class definitions are helpful to get the accuracy of features.

   ![Class definition image]

   Description: This is a large bird with a white neck and a black back in the water.

   Class Definition: The Western Grebe is a waterbird with a yellow pointy beak, white neck and belly, and black back.

   Use both text and image information to generate counterfactual images. What if the number of images for the target class is small? -> it's difficult to generate counterfactual images.

   ![Counterfactual image generation diagram]