Ivan Miller - Problem Set 3 for Adversarial AI. Tuesday, Oct. 11, 2022

1. Consider x = (1,0,0,1) and y = (0,1,1,0) as real-valued, 4-dimensional vectors. Given an arbitrary real-valued vector z = (zz1, zz2, zz3, zz4), how can you tell if z can be written as a linear combination of x and y or not? Describe an algorithm that takes in an input vector z and outputs "yes" if it can be written as a linear combination of x and y, and outputs "no" otherwise.

First, we could write the linear combination we'd like to check in the following form:

$$z = a \cdot x + b \cdot y$$

Then we could solve it for each element of z:

$$zz_1 = a \cdot 1 + 0 = a$$
$$zz_2 = 0 + b \cdot 1 = b$$
$$zz_3 = 0 + b \cdot 1 = b$$
$$zz_4 = a \cdot 1 + 0 = a$$

So in order to check whether vector z could be written as a linear combination of x and y we need to confirm that $(zz_1 = zz_4) != (zz_2 = zz_3)$

2. Now consider a process that takes linear combinations of x and y and then adds small noise values εε1, εε2, εε3, εε4 in each coordinate. (specifically, suppose each noise value has absolute value < 0.01). Describe an algorithm that takes in an arbitrary vector z in 4 dimensions and outputs "yes" if it is a possible output of this process and "no" otherwise.

After adding a noise vector to the above equation we would need to check the following equation:

$$z = a \cdot x + b \cdot y + c \cdot \varepsilon$$

See below for each element in z that we'd need to check:

$$zz_1 = a + c \cdot \varepsilon\varepsilon_1$$
$$zz_2 = b + c \cdot \varepsilon\varepsilon_2$$
$$zz_3 = b + c \cdot \varepsilon\varepsilon_3$$
$$zz_4 = a + c \cdot \varepsilon\varepsilon_4$$

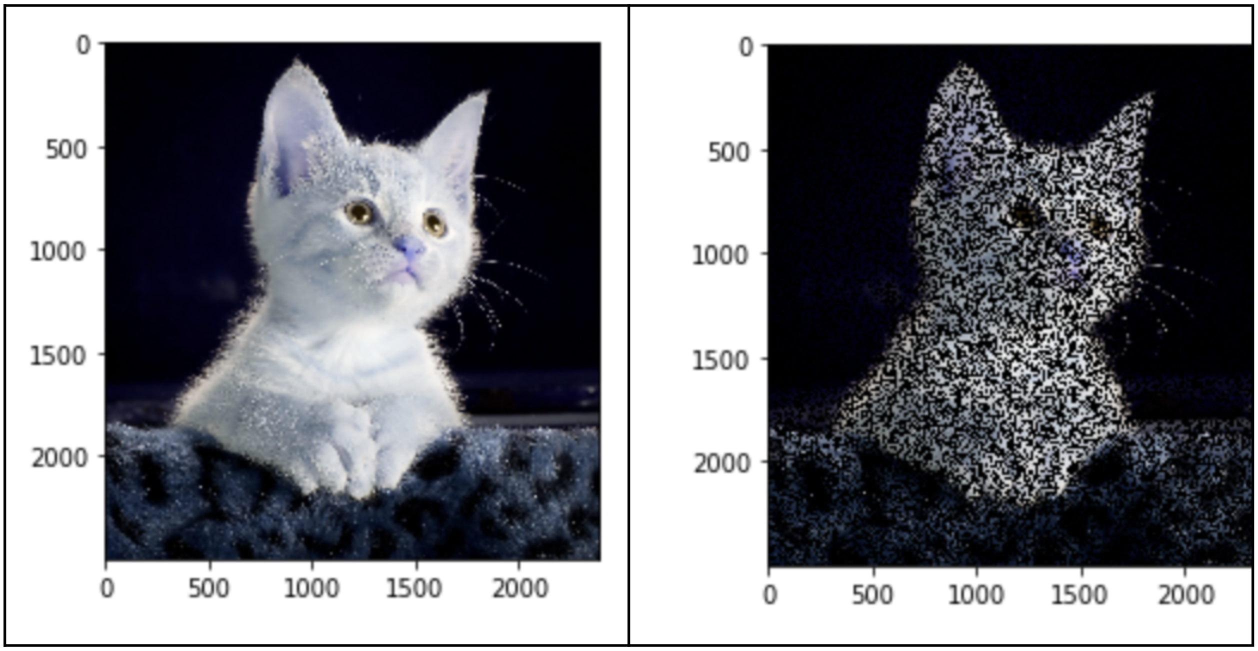
Which brings us to two possible scenarios:

1) Noise value added to each coordinate is constant
   In that case our algorithm basically stays unchanged and we only need to check if
   $(zz_1 = zz_4) != (zz_2 = zz_3)$ to output "yes"

2) Noise value is unique for each element of ε
   In that scenario each element of z will be unique which will make it impossible for us to determine whether vector z is in fact a linear combination of x and y.

3. Consider a simpler proposal to "hide" the sensitive content of images by selecting a random half of the pixels and turning them to black. Do you think this would prevent the images from being decently reconstructed? Why or why not?

It is not possible to completely prevent reconstruction of the original images but the success may be dependent on how much of the original image was black to begin with. For example, if the original image had a lot of black pixels even the removal of 50% at random would likely preserve a recognizable outline of the picture and will make the reconstruction easier:



One strategy to attempt reconstruction could be to fill all black pixels with average value for non-black pixels in the modified image.