

Daniel Kim

HW 5

Github: <https://github.com/DanielKim512/Intro2DL.git>

1. L0 norm counts all non-zero values. This is a helpful property in determining only the countable values of the reconstruction error that have a difference. That way, the zero values (which means 0 error) would not be considered and only the non-zero values would be primarily focused. L1 norm takes the sum of all absolute values. This property from first glance in terms of reconstruction error has equal weighting for all errors whether they are outliers or not. This could be helpful if we don't want to emphasize outliers as highly in our reconstruction error calculations. L1 also can produce multiple solutions. Finally, L1 norm could be useful due to its sparsity. Unlike L2 norm using regularization methods such as ridge regression that does not create zero coefficients, L1 norm methods like Lasso do set coefficients to 0 which could reduce loss. This built-in feature selection is another helpful property when used in reconstruction error. L2 norm takes the square root of the sum of squared values. Due to squaring, L2 norms penalize outliers compared to L1. This is beneficial if we are highly sensitive to outliers and would like to see them more in the reconstruction error. L2 norm is a Euclidean distance that has only one solution. L2's property attempts to minimize loss, which is beneficial. While L2 is more computationally difficult, modern technology assuages these concerns and in return, provide better predictive properties compared to L1. For example, when performing feature selection, L2 norm would combine significant features together to create the best predictions, while L1 would select the highest feature.
2. High contrast geometric images are quite complex and require penalties for large enough distances in the pixel to pixel comparison from the reconstruction error. L0 norm would be the most ideal due to being the most penalizing. Because of its property of simply identifying non-zero elements, using L0 norm helps to quickly identify whether a pixel is incorrectly reconstructed or not.
3. Wildlife images or any other type of images where geometrical patterns do not exist can use L2 norm for example to calculate reconstruction error. This is because now using the MSE as our error type, can compute similarities between pixels without necessarily needing to binary identify them as right or wrong. For example, when looking pixel by pixel of grass, the distance between each grass is short since they are all neighboring. One can also claim that L1 is also a way of measuring the distance pixel to pixel with equal weight compared to L2 which penalizes outliers heavier due to squaring the differences.
4. Assuming p distribution contains the complete data, doing an E-M algorithm gives a density estimation for q which would help make q close to p , where first starting with the expectation step and finding the expected value of the log likelihood function of θ . Then, perform the maximization step where we want to maximize θ . Through an iterative process, eventually converging the value to a fixed point. This iterative step is done by first choosing values for an unknown parameter and then estimating one group of unknown parameters then repeatedly estimating one parameter with the other.
5. Translation of images using Cycle-Consistent Adversarial Networks: The research paper goes over what to do in situations where paired training data is unavailable. To translate an image, they instead implement adversarial loss along with inverse mapping and use cycle consistency

loss to enforce. The focus of the paper is to do an image-to-image translation by taking special characteristics of one image and applying it to another image. One example they use is with Monet's intricate painting style and the ability to translate paintings to Monet's style. To do this, they introduce a concept of cycle consistency loss having two mappings of X to Y and Y to X are inversely related to one another and create a cycle when mapped simultaneously. In terms of the usage of adversarial loss, the researchers implemented GAN's allowing for generated images to be indistinguishable when compared to the original photos. Unlike many methods mentioned that either use paired image to image translation or specific assumptions made in unpaired image to image translation, the researchers developed a method with more general purpose. To formulate, the researchers put all what was said into effect. By creating two mappings G and F in addition to the adversarial discriminators, they can achieve adversarial losses as well and cycle consistency losses which are mentioned above. In terms of implementing, the network used were 3 convolutions, residual blocks, 2 fractionally strided convolutions as well as one convolution mapping to RGB. For 128×128 images, they used 6 blocks while using 9 blocks for 256×256 images. For the discriminator network, they used 70×70 PatchGANs. To finalize, they replaced the negative log likelihood with a least square loss create stable and higher quality results as well as updating discriminators with historically generated images. Using their methods, the researchers were able to achieve great results in terms of being able to produce translations that were of similar quality to the if the images were under full supervision using pix2pix. The methods used were outperforming baselines. Some failures are shown in Figure 17 due to CycleGan being unable to make major changes to the input image. Overall though, the paper provides greater insight in the possibilities of going further beyond in unsupervised learning.