Intuition Report 8

Ramtin Mojtahedi Saffari - 20307293

Assessed link 1: GAN Paint

Self-reflection

- **Short Summary**

The link provides a great interactive tool for exploring and experimenting with GAN Paint. GANpaint is a mobile application that works by directly activating and deactivating groupings of neurons in a deep neural network trained to produce pictures. Each of the buttons on the left (such as “door,” “brick,” and so on) corresponds to a group of 20 neurons. Using the interactive tool, it is understandable that the network learns about items such as trees, doorways, and roofs due to learning to draw.

- **Hypothesis and Expectation**

I hypothesized that the proposed GAN prevents drawing inappropriate objects in an unrelated area of the image. However, this is not the case for using the “remove” button, as it can be applied even to areas where the drawing hasn’t been done before, which may distort the image.

- **What I Achieved and Learned**

I tested all of the available images by using the seven provided features. The features of “Sky,” “Grass,” and “Cloud” are effective when the drawing is mostly done near the edge of the images. For example, applying “Grass” inside the sky can’t be done as the network has been trained to consider the downside of the network. In another experiment, I tried to add a door or brick to the sky. I understand that this can’t be done as when there is a pure blue sky. However, if there is a cloud or the drawing is applied near the edge, the drawing can be effective. However, even in a complex image, when the objects are close to each other or the color is complex, the output is vague and distorted.

In another experiment, I tried to use the “remove” button on the areas that had never applied any drawing. There was still distortion happening in the image. It is based on the majority of the color or object in that specific spot. Therefore, one of the provided tool’s shortcomings is that the remove button can’t work well on drawings done globally or cover most of the image.

The process of controlling the drawing is being done through sets of neurons. In this sense and according to the startling discovery, the same neurons govern a given object class in various settings, even though the ultimate appearance of the item differs significantly. The same neurons may activate the notion of a “door” regardless of whether a large stone wall necessitates a large, heavy door facing to the left or if a little hut necessitates the use of a small curtain door facing to the right. The network also recognizes when it is permitted to compose items and when it is not permitted. In the case of a door, for example, turning on neurons in the appropriate place of the building will result in the addition of a door. However,

doing the same thing in the sky or on a tree will almost always have no impact [1]. In this way, if the controlling neurons can't work well or unrelated neurons are activated, the results are distorted. This makes sense why it is challenging to apply GANs to complex images.

Suggestion and Filling the Gaps

I found the interactive link very interesting and helpful in learning how a GAN visual representation works. It is recommended to add in-detail short and publicly understandable explanations to the provided website, such as using information given in the reference [2].

Assessed link 2: Image-to-Image Demo

Self-reflection

- **Short Summary**

The provided link is an interactive tool for the pix2pix idea for image-to-image translation. The pix2pix model learns from pairs of pictures, such as labels on building facades and building facades, and then tries to produce the appropriate output image from whatever input image you provide. In this interactive tool, a user can reconstruct images for shoes, cats, buildings, and bags using the user's drawings.

- **Hypothesis**

I hypothesize that the reconstructed images are highly dependent on finding the similarity between the user's drawings and the actual images of the training dataset, which could be bag, building, cat, or shoes.

- **Testing the hypothesize and what I have learned**

I tried several experiments on different categories with different shapes and drawing. For example, when I draw a circle for the cat images, the reconstructed image searches to find similarities between the drawing and the actual shape of organs in a cat. This can result in a cat's nose, which is circular, or a small circular shape of a cat. For the categories of bag or shoes, I tried to draw different English alphabets, and the results were distorted and vague unless for those alphabets, which a part of their shapes can represent a part of the shoes or bag, such as "O," "U," "R," etc. Depending on the façade's type, the reconstructed image could be included in complex parts for the building category. For example, if the whole image is covered with a single object, such as window; then the reconstructed image can work like zoom and show a single big window. This confirms the dependency between the shape of the drawn image and the outcome of the reconstructed image.

The provided interactive tool uses conditional GANs. CGAN (Conditional GAN) is a GAN version in which the Generator and Discriminator are both conditioned on auxiliary input during training, such as a class label. Through this experiment, it is apparent that CGANs can be a useful tool in the translation of numerous images to and from other images, particularly when dealing with highly organized graphical outputs. These networks learn a loss tailored to the job and data at hand, which allows them to be used in a broad range of applications.

Suggestion and Filling the Gaps

Similar to what I mentioned before, I found the interactive link very interesting and helpful in learning-by-experiment with different given sentences. It is recommended to provide more explanation for how to get the most from interactive tools. This information can be used from the reference [3]

Self-evaluation:

In this intuition report, I have gone through an in-depth analysis of two of the provided links. In my assessment, I have completely considered the required expectation for deep exploration, including proposing a hypothesis and what I expected, reporting on what I achieved and explored, in-depth discussion of what I have learned, and providing gaps and recommendations to fill them. Considering the quality and assessment level, I deserve to get the full mark (4 points) for this intuition report.

In advance, thank you very much for your time and consideration of this report.

Best regards,

Ramtin

References

[1] Colyer, A. (2019). GAN Dissection and Datacenter RPCs. *Queue*, *17*(2), 22–23. https://doi.org/10.1145/3329781.3329783

[2] Lazarou, C. (2020, January 6). *This Will Change the Way You Look at GANs - Towards Data Science*. Medium. https://towardsdatascience.com/this-will-change-the-way-you-look-at-gans-9992af250454

[3] Abbasi, N. (2021, November 7). *What is a Conditional GAN (cGAN)?* Educative: Interactive Courses for Software Developers. https://www.educative.io/edpresso/what-is-a-conditional-gan-cgan