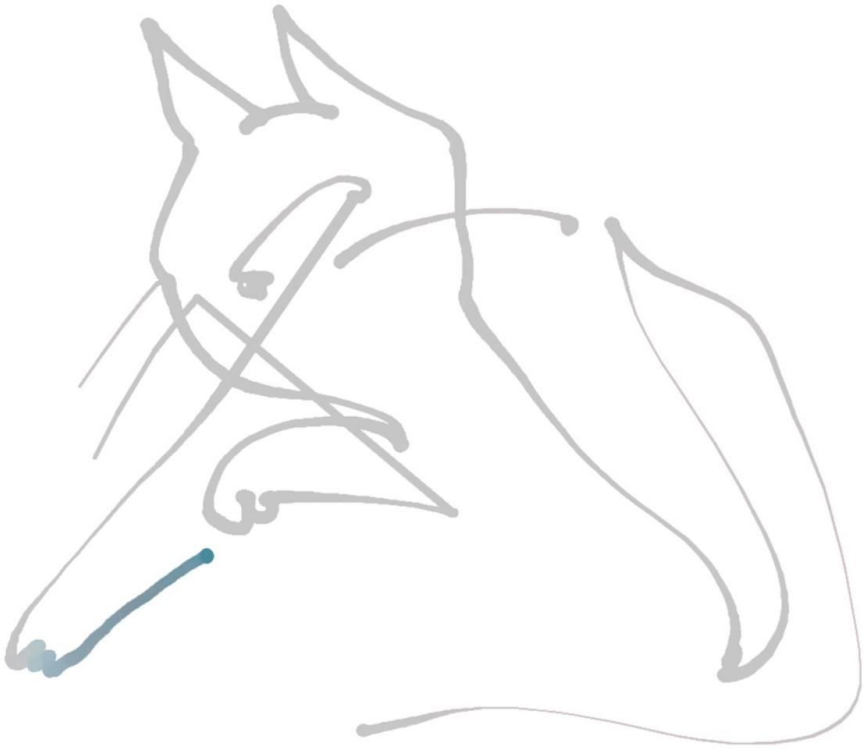
ART AND MACHINE LEARNING
CMU 2019 SPRING
PROJECT 3

Doodle Match



Marisa Lu
Mira Murali
Loki Ravi
Izzy Stephen
Yuhan Xiao

DESCRIPTION

1. **Concept:**

   Doodle Match is an interactive web app designed to foster fun, lighthearted artistic collaboration between humans and AI. Artists who use machine learning regard algorithms as a tool or medium, like oil paint. Meanwhile critics, who often lack understanding of machine learning, misconstrue these works as art "made by robots" that think and create autonomously. This tool shows that the truth lies somewhere in the middle - humans can work with artificial intelligence to make artwork that's greater than the sum of its parts.

   Through latent space arithmetic, users can manipulate the latent vector outputted by the encoder network of Sketch-RNN. This sheds light on the features the model uses to separate its representations of different objects and concepts it's been exposed to. Latent vectors of similar/dissimilar sketches from a given category of sketches can be added, subtracted, multiplied or divided to enhance their features or identify the differences between them. Across different categories (heterogenous subsets), sketches can be interpolated to produce surprising and amusing results reminiscent of GANbreeding. Experimenting with latent space arithmetic allows the user to delve into the model's understanding of the world, seeing familiar objects through unfamiliar eyes.

2. **Technique:**

   a. ***Problem Formulation***

      In this project, we use Sketch-RNN to explore two ideas:

      i. Creating an interactive experience of collaborating with an AI to create a doodle where each adds a single stroke in a back and forth game of "what-next?"

      ii. Experimenting with latent space arithmetic in the seq2seq model to add two dissimilar sketches or subtract similar sketches to observe whether the final result contains the essence of both sketches (in the first case) or retains the common features between the sketches (latter case)

   b. ***Training and Datasets***

      We use the pre-trained models presented in the official Sketch-RNN repository to implement the ideas. We look to past experiments using the model to collect datasets on top of the demo datasets given in the repository. Some of the sketches in these datasets include those of bees, flowers, ants, yoga poses. The majority of our data comes from the Google QuickDraw Dataset

3. **Process:**

   Doodle Match:

   Sketch-RNN is a seq2seq model with a bidirectional RNN for an encoder and an autoregressive RNN for the decoder. During prediction, given the first stroke, the decoder generates the subsequent strokes to create the final result.

However, in order to make the process interactive, we use teacher forcing at time of inference to feed user input at every alternate time step. The user creates the first stroke, based on which the decoder predicts the next stroke. Instead of letting the decoder continue to generate subsequent strokes based on its previous strokes, we let the user intervene and provide the input every other stroke.

Latent Space Arithmetic:

The output of the encoder is projected onto a fully connected layer, which output two vectors. These vectors are combined in a random fashion to generate the 128-dimensional latent vector.

The latent vector essentially captures the key features of each sketch. We are interested in the two basic arithmetic operations: addition and subtraction. We hope that by adding two vectors, the features of both will be amplified and that by subtracting them, only the dissimilarities between them would be retained.

However, we realized that the 128-dimensional vector is not de-correlated, i.e., each dimension in the vector does not represent an independent feature.

4. **Result:**

We present the results of our latent space manipulation below. We experiment with homogeneous datasets (which contains images from a single class of sketches) and heterogeneous datasets (which contains images from multiple classes of sketches).

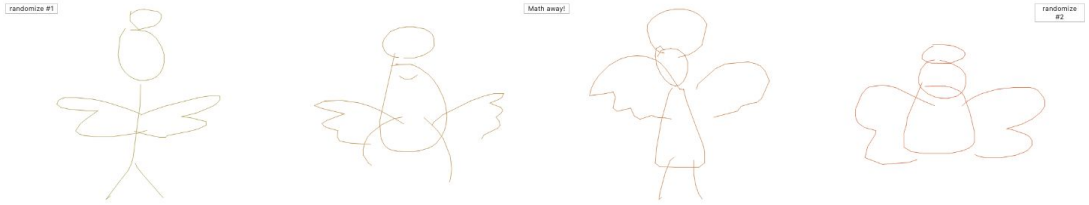


Fig 1 Body Shaming Angels: Result of adding and subtracting the latent vectors of two sketches from homogeneous datasets(angels). The sketch on the far left is the first selection(thin angel) and the sketch on the far right is the second selection(fat angel). The second sketch from the left is the sum of the two latent vectors as decoded by the model. We observe that this angel has bent legs and has a rounder shape. The third sketch from the left is the difference of the two latent vectors as decoded by the model. We observe that this angel goes to the gym.

Fig 2 Cool Dude: Result of adding and subtracting the latent vectors of two sketches from heterogeneous datasets(man-bicycle). The sketch on the far left is the first selection(man) and the sketch on the far right is the second selection(bicycle). The second sketch from the left is the sum of the two latent vectors as decoded by the model. We observe that this is a man doing hip hop adopting the pose of a bicycle. The third sketch from the left is the difference of the two latent vectors as decoded by the model. We observe that this is a man riding a segway.

5. **Reflection:**

The results of the latent space arithmetic manipulation were unexpected. For instance, the result below was impressive. Although neither of the source images contained numbers, the addition result added numbers to the clock sketch, making it a more realistic doodle. This may be attributed to the fact that the model was exposed to many clock sketches containing the numbers and understood to amplify those particular features.

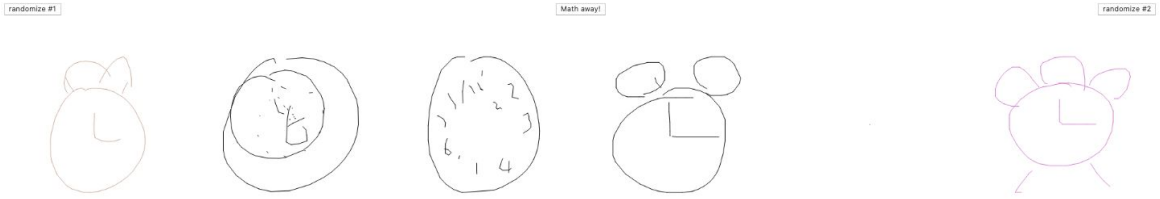


Fig 3: Example of addition, subtraction and multiplication results from a homogeneous dataset(clocks)

CODE: www.github.com/mira-murali/rom-comp

CONTRIBUTIONS:

- Artistic meaning and concept: Izzy Stephen
- Interactive Web Interface: Marisa Lu, Yuhan Xiao
- Latent Space Arithmetic: Loki Ravi, Mira Murali, Yuhan Xiao
- Project report, presentation: All