Analyzing How the General Public Rates Images Based on a Computational Model

Ramsey Boyce[1], Darlene Feldstein[2]

[1]Henry M. Gunn High School, [2]independent

INTRODUCTION

Appealing images are important for everything from blogging to marketing. To make things more efficient, we could try to rate images with a computer. However, we can't use regular computer programs to do this; we need to use a neural network, a program that mimics a part of the human brain.

In this project, I used a neural network to model how the general public rates images. The network was trained using a set of images called the AVA dataset (see Figure 1). This dataset was used because it contains a large variety of diverse images--a quarter of a million in total. After training, the network was analyzed to determine what the "best" and "worst" features are in images.

RESEARCH METHODOLOGIES

First, the images were converted to codes by using a different neural network. The codes store the images' features by converting them into a string of 4096 numbers, making the network more accurate because the network didn't have to detect the features. The network is trained to rate an image based on the codes and the ratings that they correspond to.

When the network is trained, numerical weight values in it are calculated to optimize the network's ability to rate images—each input feature has its own weight that it is multiplied with. The final rating is the sum of the products of the features and the weights. Therefore, the input features with the highest positive weight values can be said to be good because they affect the final rating in a positive way, and the features with the lowest negative weights bad because they lower the final rating.

Images with high values for the best and worst features will be visually analyzed to determine what the features correspond to in an image.

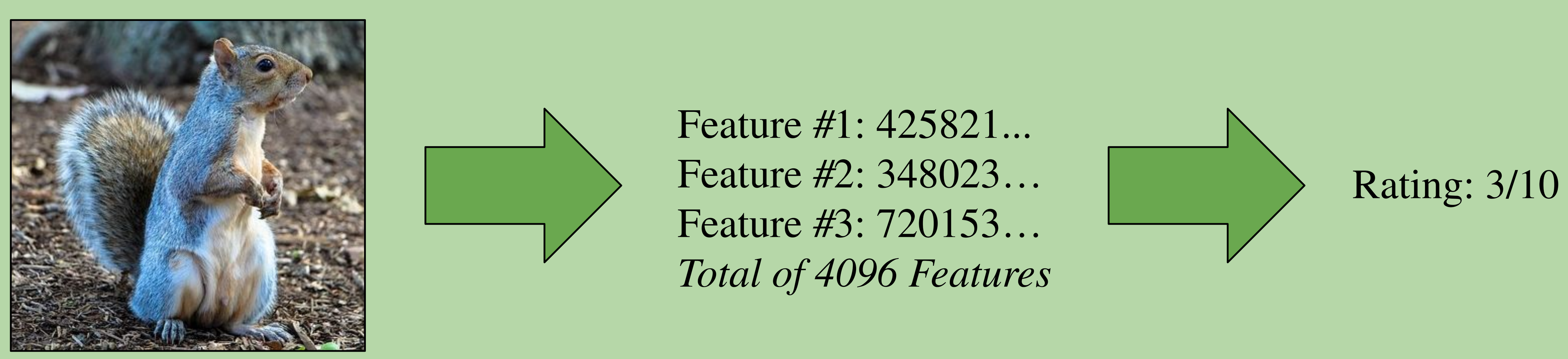

Figure 1: How an Image Becomes a Rating

CONCLUSIONS AND ANALYSIS

There might be inherent randomness in image ratings because the network was more accurate when the distribution of ratings it generated was "crushed". This suggests that guessing close to the median rating is a good, safe strategy for rating images—maybe people have different internal scales for image ratings?

When looking at the good/bad features, unfiltered nature seemed to be mostly good, except for squirrels.

Some features appeared on both the positive and negative sides: cups and images with the blue-green filter are two examples—these features are hard to place since some people really like them and others dislike them.

Other features were already known to be good or bad, like images with too many colors. However, the features are still hard to understand as humans did not pick them.

The identified good/bad features should be useful for photographers or anyone who needs a good image.

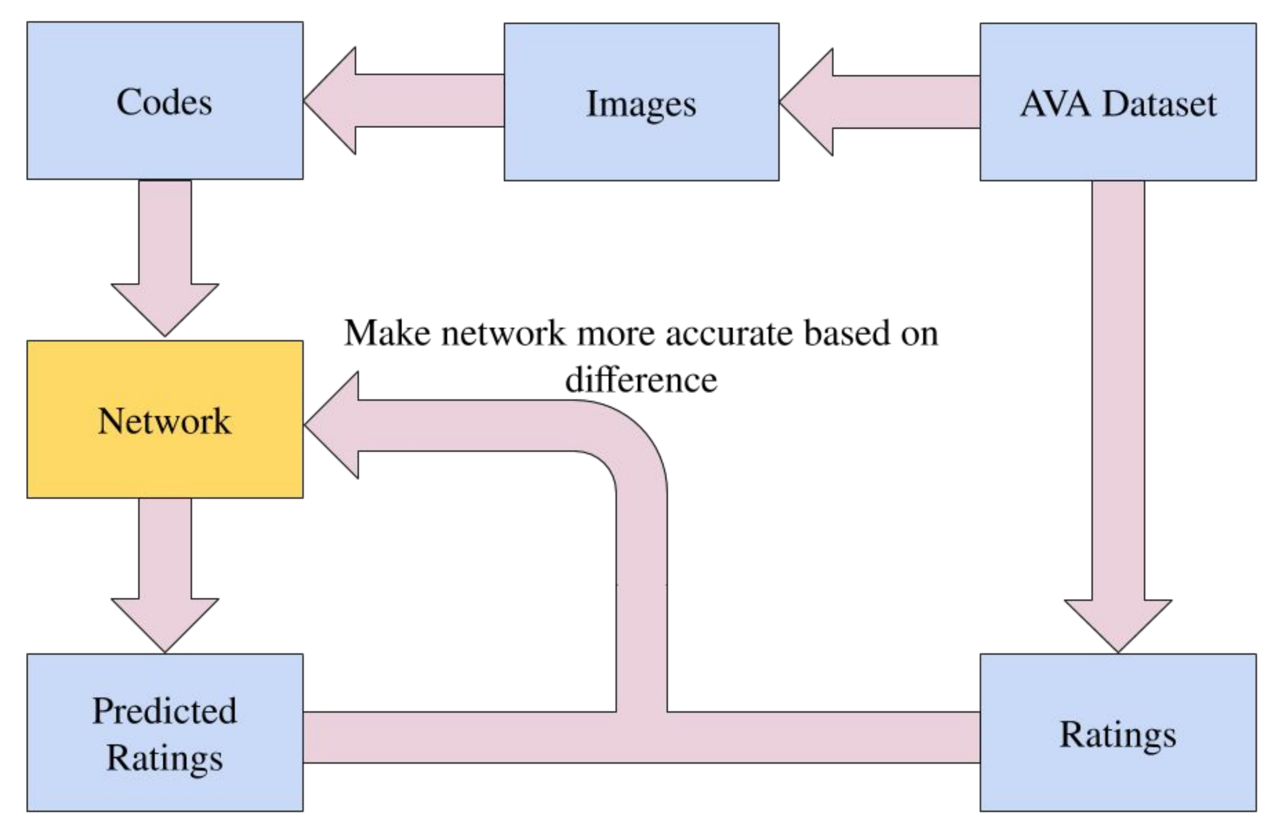

Figure 2: How a Network is Trained

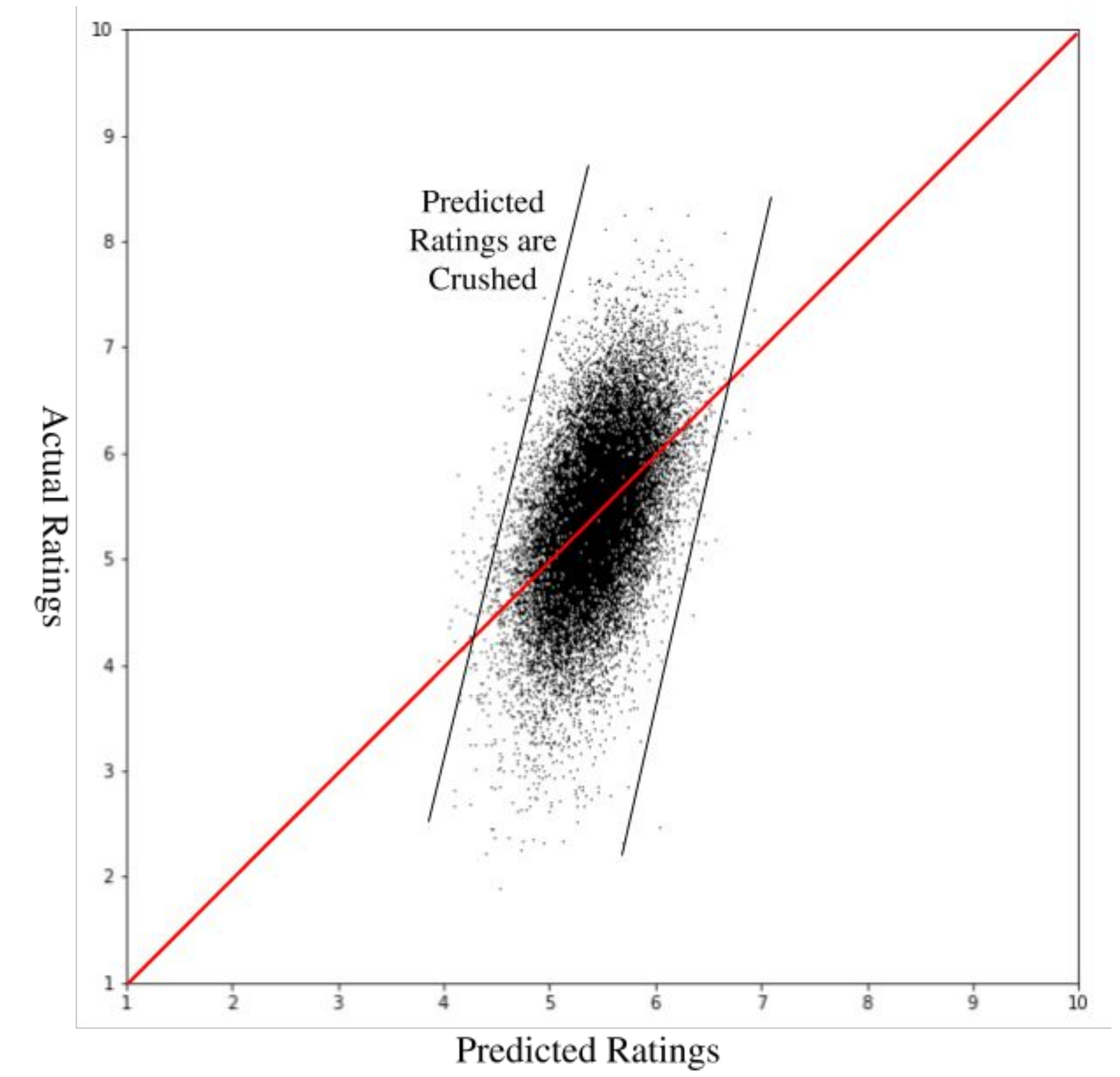

Figure 7: Optimal Network's Predicted Ratings vs Actual Ratings

IMPLICATIONS AND NEXT STEPS

In the future, more analysis could be done on the features in the codes. If there is a better way to determine what the common feature is from a group of images, we could get a better understanding of the things that people like/dislike.

More work could also be done to determine whether the subject or the effects in an image are more important. Squirrels were a negative feature, which seems to show that the subject matter of an image is more important, but the blue-green filter seems to show that the effects visible in an image are more important. The importance of each is unknown.

DATA AND FINDINGS

	Place of Feature	First Possibility for Feature	Second Possibility for Feature
Top 5 Features	1 (Best)	Frogs, snakes, fish (see figure 6)	Speckled metal
	2	Drinking cups	Sheep
	3	Horses with blue-green filter (see figure 5)	Birds
	4	Dogs	Frogs
	5	Flowers	Wispy objects
Lowest 5 Features	4092	Too many colors	Wispy objects
	4093	Cups/silverware	Shiny objects
	4094	Squirrels	Collection of small, round objects
	4095	Dogs with blue-green filter (see figure 5)	Cats with blue-green filter (see figure 5)
	4096 (Worst)	Fruits/nuts	Shiny animals

Figure 3: The Five Best and the Five Worst Features

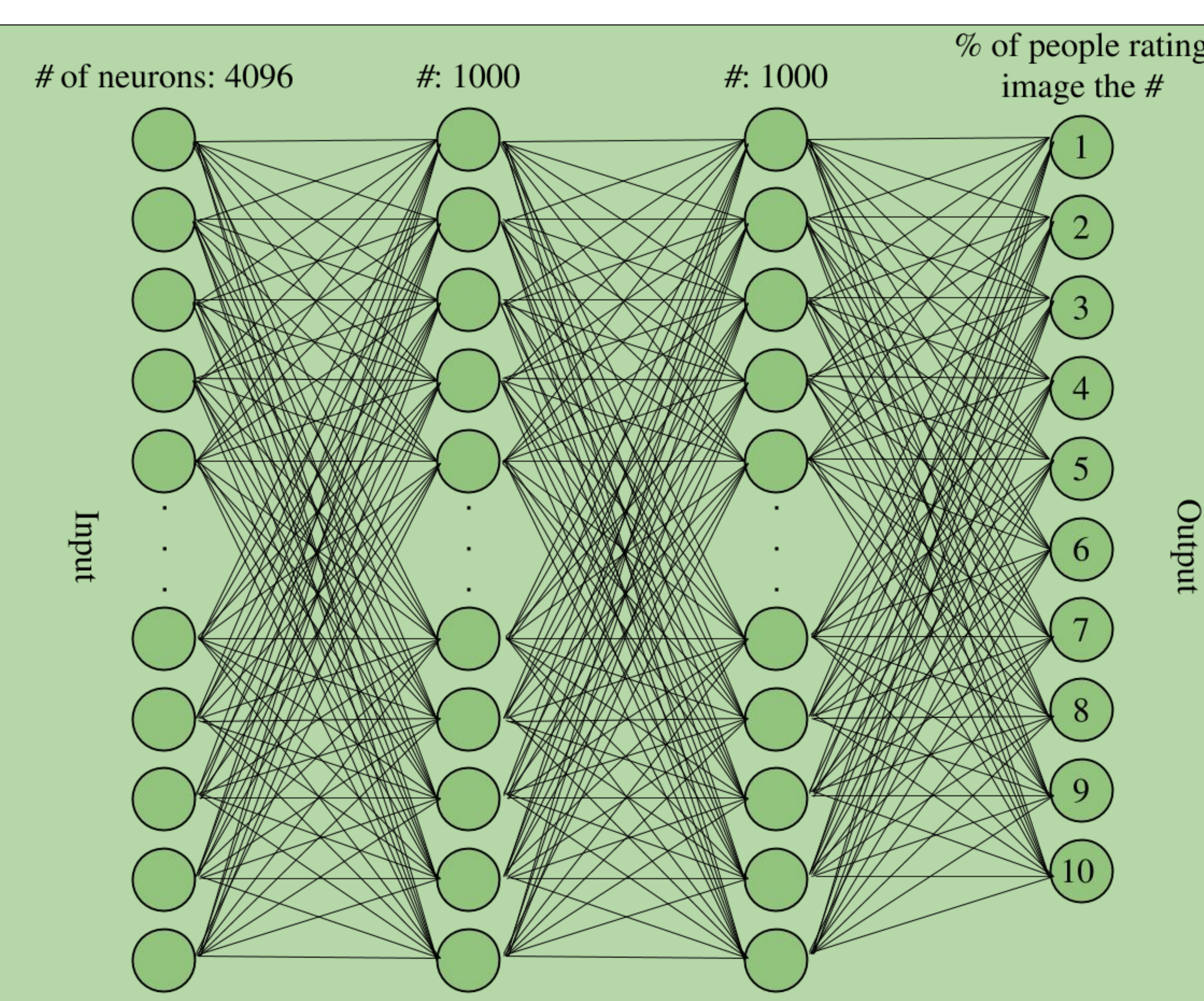

Figure 4: Final network structure - each circle represents a neuron and the lines represent connections

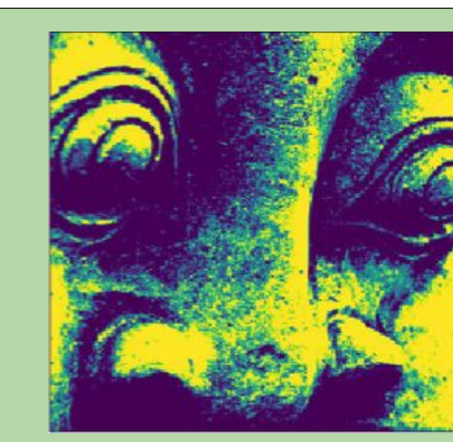

Figure 5: Blue-Green Filter from Figure 3

Figure 6: Frog from Figure 3

ACKNOWLEDGEMENTS / REFERENCES

Joshi, D., Datta, R., Fedorovskaya, E., Luong, Q. T., Wang, J. Z., Li, J., & Luo, J. (2011). Aesthetics and emotions in images. IEEE Signal Processing Magazine, 28(5), 94-115.

Kong, S., Shen, X., Lin, Z., Mech, R., & Fowlkes, C. (2016, October). Photo aesthetics ranking network with attributes and content adaptation. In European Conference on Computer Vision (pp. 662-679). Springer, Cham.

Lu, X., Lin, Z., Jin, H., Yang, J., & Wang, J. Z. (2014, November). Rapid: Rating pictorial aesthetics using deep learning. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 457-466). ACM.

Nguyen, A. T., & Kelle, J. Personalized Image Aesthetic Prediction.

Park, K., Hong, S., Baek, M., & Han, B. (2017, March). Personalized Image Aesthetic Quality Assessment by Joint Regression and Ranking. In Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on (pp. 1206-1214). IEEE.

Ren, J., Shen, X., Lin, Z. L., Mech, R., & Foran, D. J. (2017, October). Personalized Image Aesthetics. In ICCV (pp. 638-647).

Wang, G., Yan, J., & Qin, Z. (2018). Collaborative and Attentive Learning for Personalized Image Aesthetic Assessment. In IJCAI (pp. 957-963).

Yeh, C. H., Ho, Y. C., Barsky, B. A., & Ouhyoung, M. (2010, October). Personalized photograph ranking and selection system. In Proceedings of the 18th ACM international conference on Multimedia (pp. 211-220). ACM.