1) The names of the student(s) in your group.

Caroline Cutter and Julia Fairbank

2) Title of your project.

reCAPTCHA: Using Human Cognition to Supplement Optical Character Recognition with Small World Neural Networks to supplement the CNN-ECOC approach.

3) The general research area (e.g. Machine Learning).

Machine Learning, Neural Networks, Optical Character Recognition (OCR)

4) A summary of the related literature you reviewed and any background necessary to understand your research problem. Remember to include citations.

Handwritten Character Decoding and the CNN-ECOC Method

Bora, Mayur Bhargab, et al. "Handwritten Character Recognition from Images Using CNN-ECOC." *Procedia Computer Science*, vol. 167, 2020, pp. 2403–2409., https://doi.org/10.1016/j.procs.2020.03.293.

OCR is a process that uses feature extraction and classification to classify optical patterns that are present in an image. There are many uses in OCR which includes decoding scriptures, optical scanning of cards, bank checks, etc., plate recognition in car, scanning academic records for databases, and many more. There has been recent study showing that neural networks can be used to implement OCR algorithms to aid in deep learning. One of the challenges that researchers have faced is handwritten character recognition and with recent advancements in the use of neural networks in object recognition, this paper proposes that OCR for handwritten characters can be achieved by using Convolutional Neural Networks. Convolutional Neural Networks (CNN) are large networks (similar to graphs) of interconnected neurons that have weights and biases that learn as it sees/uses more information. The networks of neurons are organized in many layers – an input layer, hidden layers and an output layer – in 3D space. The hidden layers are to receptive fields of the input layer from the previous layer, which reduces the number of connections to save time and space. This network works in object recognition by holding the image in the input layer, extracting features in the convolution layer, reducing in the ReLU and pooling layer and then making connections in the fully connected layer. The CNN uses training by forward and back passing and then once the data is trained, it can be used to solve a specific problem. This paper describes how a combination of CNN (for feature

extraction) and ECOC (used for classification) architectures can help with OCR, specifically handwritten character recognition.

Captchas:

Gugliotta, Guy. "Deciphering Old Texts, One Woozy, Curvy Word at a Time." *The New York Times*, The New York Times, 28 Mar. 2011, https://www.nytimes.com/2011/03/29/science/29recaptcha.html.

Richard, Diane. "ReCAPTCHA: Deciphering Old Texts I - Mosaicrpm." *Mosiac RPM*, http://www.mosaicrpm.com/files/reCAPTCHA_final.pdf.

Captchas, which stands for "completely automated public Turing test to tell computers and humans apart" is a form of security that many online systems have used to ensure a human is using the computer instead of a machine. These consists of distorted words and letters that a human must recognize and type to "beat" the security. The reason why these have been so effective in telling humans and machines apart is because machines have a lot of trouble in recognizing handwritten or distorted letters. However, recently people have been using machine/deep learning techniques to help computers recognize this. The backstory behind Captchas is that the Dr. von Ahn from Carnegie Melon received the images from the NY times of words that their digitalization project could not detect. Dr. von Ahn then turned these undetectable characters/words into a security system (Captchas) and used them on the internet to have people all over the world solving them. It is estimated that humans decode at least 200 million Captchas per day, which allows machines to be learning and getting better at detecting handwriting and distorted letters at a very fast pace.

Small World Neural Networks

Javaheripi, Mojan, et al. "SWNet: Small-World Neural Networks and Rapid Convergence." *UC San Diego*, 2019. https://arxiv.org/pdf/1904.04862.pdf.

Deep learning models can be computationally costly because of over-learning and excess number of trained parameters. From the first paper, we saw the use of CNN in deep learning in order to decipher/ digitalize handwritten letters from photos. This paper proposes a Small-World Network in addition to CNN to change the topology of the CNN architecture to optimize the connectivity of the graph in an efficient manner. Small World Network architecture allow for a connective structure among nodes that makes it more efficient to communicate across short and long distances. Small-World networks require fewer training parameters which can increase the

efficiency of the CNN. This experiment showed a 2.1 x improvement of convergence speed in the CNN

Example Study: Deep Learning CAPTCHA Solver

Team, Towards AI Editorial. "Deep-Learning-Based Automatic CAPTCHA SOLVER." *Towards AI*, 28 Sept. 2020, https://towardsai.net/p/deep-learning/deep-learning-based-automatic-captcha-solver.

Although this is not a scientific article, this project uses some similar data and structure that we would like to follow for our final project. This includes methods on how to scrape the web for CAPTCHAS, how the data was analyzed and an overall good structure. We don't as much want to use the information that this study found, but the ways in which they found this data using our proposed neural network in CAPTCHA solving.

** **Note** ** Although we did not include it in our proposal, we would also like to include some real neural network (and small-world network usage in the brain) information to this study. classification) architectures can help with OCR, specifically handwritten character recognition.

5) The specific research problem you will be trying to solve and how it is novel from work that has already been done in the past.

Proposing two different research problems, would love your recommendation/direction on which to move forward with.

1. Ways CAPTCHA can be used to solve other problems.
    - Deciphering and digitizing old texts
    - Decode Google Maps street view addresses
    - Facial recognition

Most other work around CAPTCHA examines the payoff between security and user accessibility/ease, we are interested in examining the ways in which the data from reCAPTCHA can be used in the digitization of other fields.

2. Improving the CAPTCHA algorithm using small world networks and Convolutional Neural Networks (CNN).

We believe that we could apply small world networks and CNN to increase the efficiency of encoding CAPTCHAs. Currently, there is little/no research on using small world network graph in the CAPTCHA algorithm.

6) A description of the set of methods you plan to use to try to solve your problem.

We are not exactly sure how we are going to implement the combined Small-World network with CNN yet, but here are some methods we found that could be useful in our project in terms of collecting and analyzing CAPTCHAs:

Team, Towards AI Editorial. "Deep-Learning-Based Automatic CAPTCHA SOLVER." *Towards AI*, 28 Sept. 2020, https://towardsai.net/p/deep-learning/deep-learning-based-automatic-captcha-solver.

Data Collection: From the paper above, there are data sets of thousands of CAPTCHA images that could be used in our study. https://www.kaggle.com/datasets/aadhavvignesh/captcha-images is a website that provides them.

Data Processing: This is where we would try to combine the systems of Small World Network and CNN code in order to more efficiently find CATPCHAs

Another research question that we were interested in is seeing the different ways in which Captcha solving algorithms could be use in different ways – so we could propose this more efficient algorithm and then see what other areas of Machine Learning they could be used in because it can be a fairly general algorithm.

7) A description of how you will go about evaluating the results of the methods described in the previous section.

Ideally, we would like to be able to compare run times of the algorithm that we are proposing and see that the one we are proposing is more efficient. However, we are not sure if we will be coding this or just using information from previous journals that we studied to estimate the runtime of said algorithm. We also want to focus on extending this study and how it can be used in different areas of machine learning. One way that we could analyze this is by seeing how the algorithm could be generalized to different problems rather than just handwriting – such as facial recognition.

If we find it too difficult to combine these two methods, we also may compare different methods in neural network object recognition instead.