Choices:
1. PhotoNodes Project
2. Deep Learning-Driven 3D Reconstruction, Data Synthesis, and Size Estimation for Road Signs
3. Deep Learning for Estimating Lane Line Quality Using Retroreflectometer Ground Truth

Why I want my #1: I have always been interested in cybersecurity and OpSec - particularly the human element. Usually, the weakest link in any system is the one that isn't automated and is instead managed by a person with flaws. As has been seen time and time again, one thing people are notoriously bad about is posting pictures online that they really shouldn't. When these contain sensitive information, it can lead to data breaches. This project will help protect the US citizens from malicious agents seeking to commit identity fraud, or worse. I feel that this project can make a real difference in the world for the better.

Backgrounds/Qualifications: I have experience with using each of the two methods explicitly mentioned by the client (CNNs and Transformers) for neural classification purposes. I am sure that I could adapt the work I've previously done for 3D, time series MRI imaging to this 2D image task fairly well.

Strategy to go at the project: I feel that given this is intended to be used for high security operations, a heavyhanded approach should be taken to this problem with asymmetric cost - false negatives need to be more heavily penalized than false positives. After all, a false positive is some lost image, a false negative is a potential data breach. I would propose training two models - one specialized in detection of any form of text, and one specialized in detecting faces. Furthermore - I believe this is an effective approach as there is a multitude of massive free datasets online with labeled face and text, as many, many projects have been done for identifying faces and text, so we would have no shortage of data to work with, and would not need to manually create or label it.

As a baseline we would strive to eliminate any presence of those by randomizing the pixel color values within the bounding box of the identified item so that the data could not be unblurred or in any way uncorrupted. Then we would need a user interface that would let someone drag a photo in and allow the model to remove any presence of either detected phenomenon. If this gets done in time, we can add a user controlled slider that will allow them to dictate how aggressive they want the model to behave in eliminating what it believes to be sensitive data.