SUMMARY

The main objective of this project was to segment a given creative of a company and recover the background from the creative. There is a segmentation model and an impainting model associated with the end to end approach.

The segmentation task has two components associated with it, one is the object segmentation where a binary mask is generated for humans/models and another component is the text segmentation where the text is detected from the creative and the binary mask is created.

After both masks from the components are generated, both of them are combined together to get a single mask, which is fed to the impainting model along with the original image.

At last the impainting model recovers the background from the image.
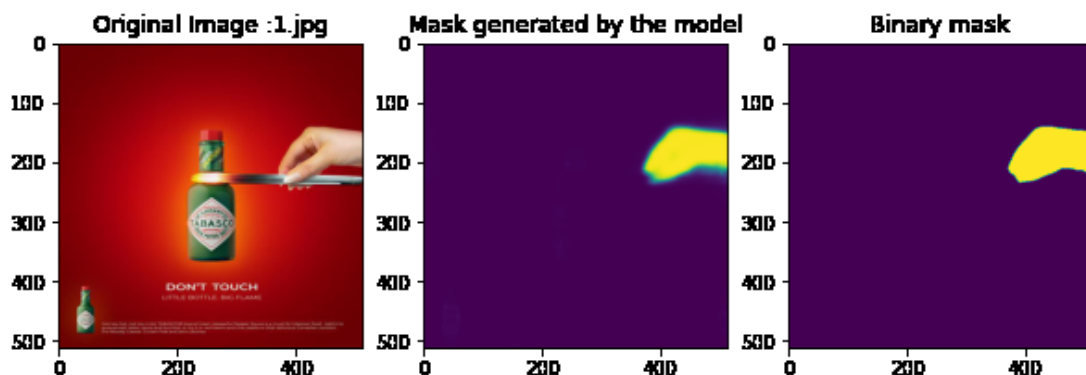
Models Used -
Object segmentation - pretrained fcn_resnet50
Text segmentation - pretrained psenet_text_detector
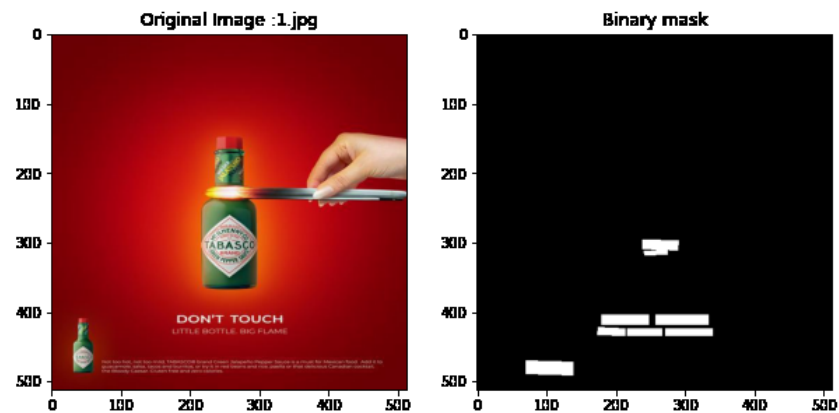Impainting - pretrained lama
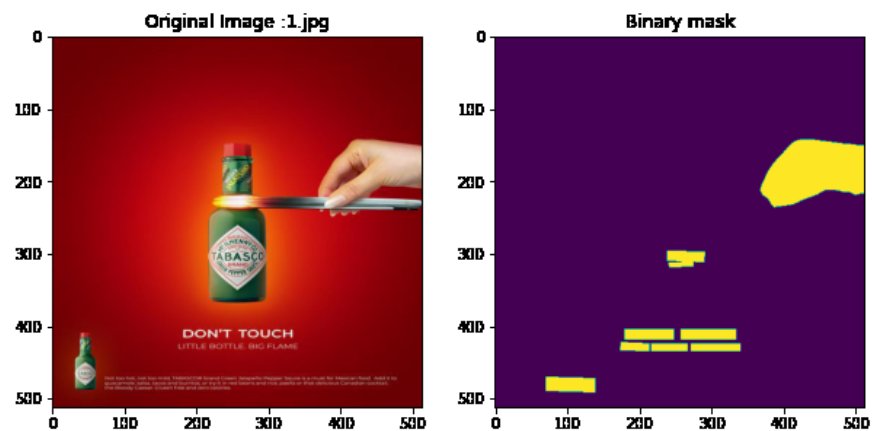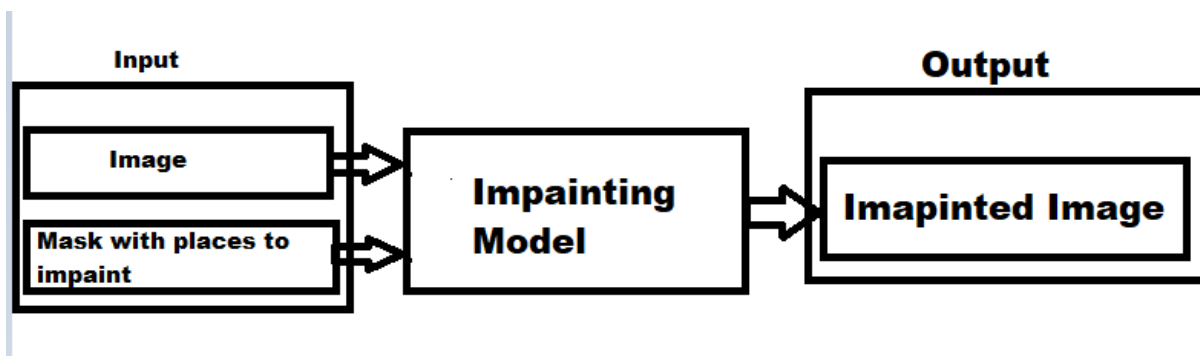
RESULTS
Object detection -
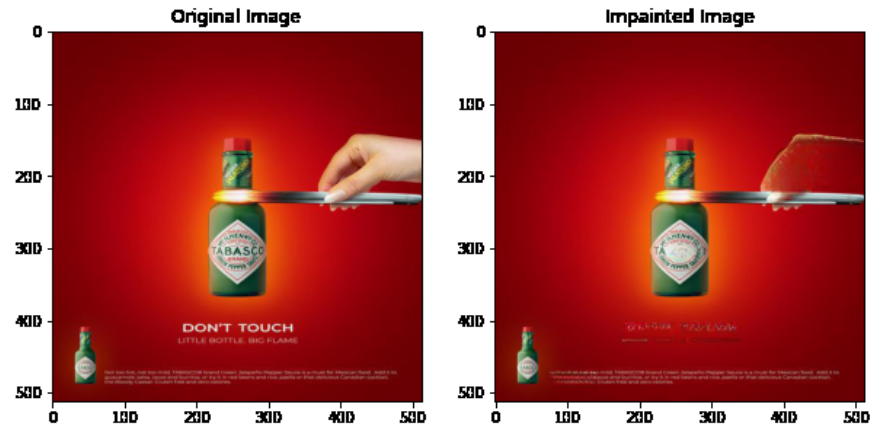
FCN

Text detection -

Combined mask-

Impainting -

Alternate Approach-
If we have more data customized to the task, an end to end approach might be to train the object detection model on the masks of the creatives as the ground truth then using an impainting model. That way we will not need 2 models for segmentation for text and object.

Similarly one more approach might be an end to end impainting model which is trained on the background as the ground truth and the input might be the image.

How to tackle edge cases-
To tackle this issue we can integrate some changes into the segmentation model. To address this issue I reviewed some papers which deal with this problem.

Papers Reviewed -
1. "SSUL: Semantic Segmentation with Unknown Label for Exemplar-based Class-Incremental Learning", Link - https://proceedings.neurips.cc/paper/2021/file/5a9542c773018268 8fc6271f7afeea969-Paper.pdf
2. "Class-independent sequential full image segmentation, using a convolutional net that finds a segment within an attention region,

given a pointer pixel within this segment", Link - https://arxiv.org/ftp/arxiv/papers/1902/1902.07810.pdf

Review on Findings -

The paper "SSUL: Semantic Segmentation with Unknown Label for Exemplar-based Class-Incremental Learning" discusses about the class incremental problem which is learning incrementally as new classes arrive and the plasticity problem associated with it, which is defining unknown classes within the background class to help to learn future classes. The process they followed which is relevant to the question goes like this: They used saliency map to assign an additional "Unknown" class label to the objects in the background. Then they use a base feature extractor to distinguish the representations of the potential future class objects and the actual background region by augmenting the background label with this additional class. They classified background into three factors: future object classes that the model does not yet observe, past object classes that are already learned, and the true background. After getting this unknown class map using this approach, they concatenate with the ground truth mask and train the model.
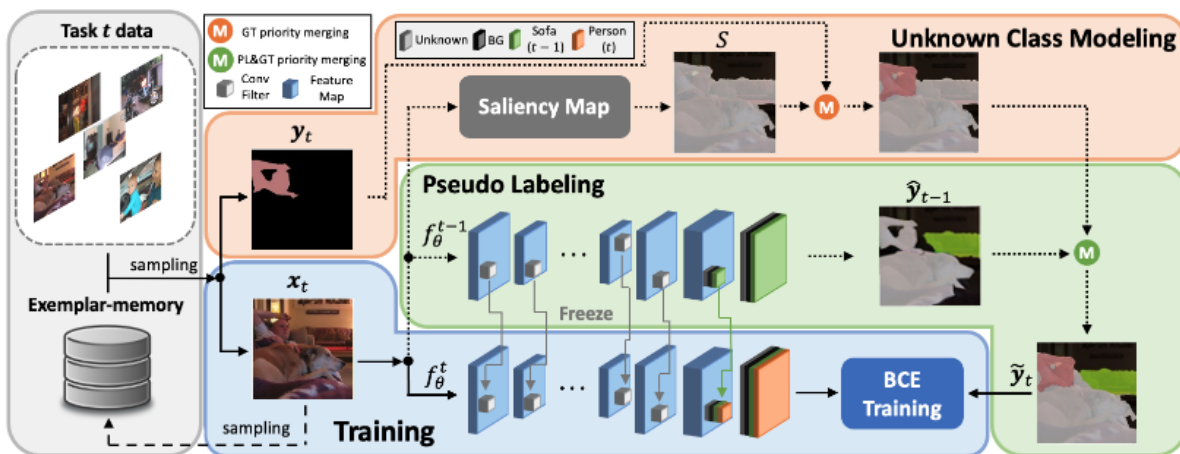
Credits: Image given in the paper.

The paper "Class-independent sequential full image segmentation, using a convolutional net that finds a segment within an attention region, given a pointer pixel within this segment" uses a pointer base classification approach, where they train their segmentation model with the image along with a point inside the object they want to segment and a region of interest mask. The method is trained on a large number of segments corresponding to different categories, and returns the region of the segment as a binary mask, regardless of the segment class. As a result, they claim that the method learns a general class-agnostic segmentation pattern and can segment even unknown categories that were not encountered during the training stage. Since the method is class independent.
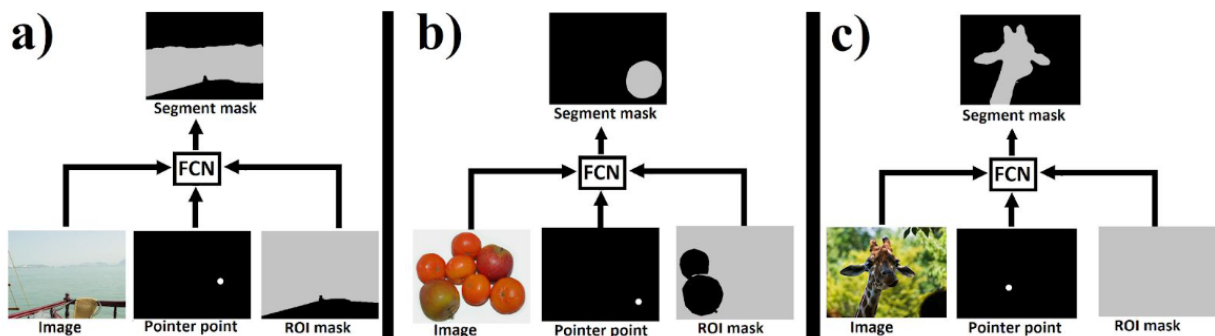
In purpose we can integrate this research in the model to classify unknown objects in the creatives.