**Generative Model**

I observed that 2/3 of the generated images illustrated a black refrigerator, in a supermarket background, as expected. However, the refrigerator was placed in an unrealistic location, such as the middle of the store. I chose the image that made the most sense (the refrigerator was placed on the side) because I had already spent 4-5 hours generating images (due to errors in the code that I couldn't spot until trying to generate images, which took about an hour). As an improvement, I could have specified that the store products should not be black, but I found the selected image to be good enough and appropriate for a real-world scenario.

**Voice Synthesis Module**

I couldn't figure out how to make it pause while speaking, so I tried to "fit" the words so they would sync with the video. I also considered creating multiple audio files for each picture with pauses in between, and then combining them. Additionally, I noticed that it reads "." as "dot."

**CIE XYZ Color Space**

To convert to this color space, I transformed the images from BGR format to RGB and then multiplied each pixel by a matrix, as seen here or here. This color space is used for device calibration and is especially useful in industry. For instance, the RGB color space can look different from screen to screen.

This color space encompasses all the sensations perceived by the human eye. The eye has three types of cone cells, and the way they are stimulated represents what we perceive as color. The correspondence between colors and their effect on these cells has been established empirically.

**Applying the Mask for Color Extraction**

I took an example from the internet for creating and applying the mask. Through trial and error, I selected the HSV range to extract all pixels that were not black (or almost black). Very few pixels in the image were truly black, so I tried multiple versions until I found the one that best removed the black refrigerator, even if that meant removing other darker pixels in the image.

**XVID Codec**

XVID is an open-source, configurable video codec. The part of creating a video with a specified codec gave me the most trouble. I could create a video, as in Lab 7, with the specified codec, but it took me a while to figure out how to add the audio and save the complete video with that codec. I used this resource to help.

**Results:**

O imagine care conține text, raft, carte, captură de ecran

Descriere generată automat

<https://www.geeksforgeeks.org/python-opencv-cv2-rectangle-method/>

<https://stackoverflow.com/questions/44947505/how-to-make-a-movie-out-of-images-in-python>

<https://www.geeksforgeeks.org/python-opencv-cv2-puttext-method/>