German University in Cairo
الجامعة الألمانية بالقاهرة

Assignment 2

DMET 1001: Image Processing

Submitted by:

Zeina Hezzah 52-2224

Submitted to: Dr. Mohamed Karam Gabr

May 5th, 2025

1) Project Overview

This project investigates the impact of different image preprocessing techniques on the object detection performance of the YOLOv8 model. The goal is to assess whether preprocessing can enhance detection accuracy by improving image clarity, emphasizing structure, or simplifying visual content prior to training and inference.

Three preprocessing techniques were explored: histogram equalization, Sobel edge detection, and K-means image segmentation. Each method was applied to the COCO128 dataset, and models were individually trained on the preprocessed images. The model performance was evaluated using standard object detection metrics including mAP50, mAP50-95, Precision, and Recall. The unprocessed dataset served as a baseline for comparison.

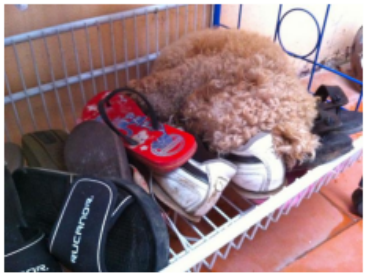
Each model was separately trained on the corresponding preprocessed datasets and their training metrics were observed and compared to observe quantitative differences. All 4 models (unprocessed and the 3 processed datasets) were tested on the same set of test images to observe how the models' perform under consistent conditions. The model outputs are also visualized to observe qualitative differences in prediction accuracy.

Colab Notebook Link: IP_Assignment2.ipynb

2) Raw Unprocessed Images

Sample Images



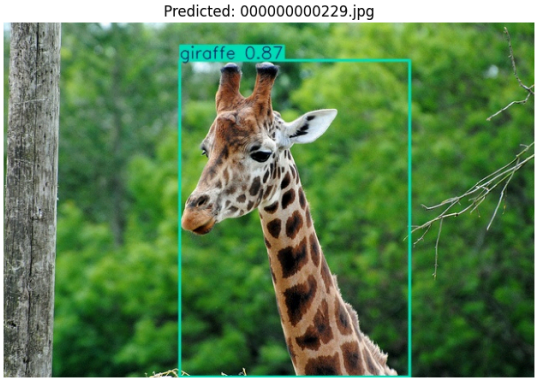


Model Training Metrics

mAP50	mAP50-95	Precision	Recall
0.64228	0.46769	0.63846	0.56557

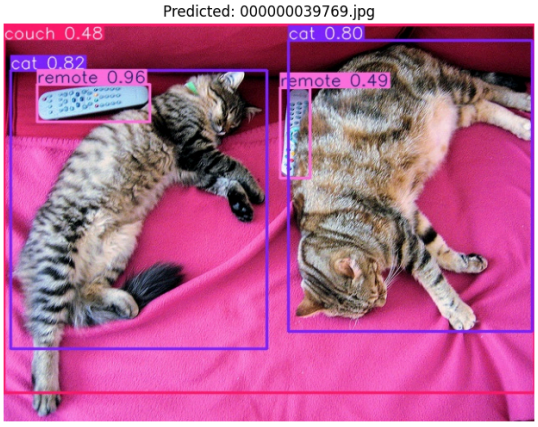
Model Output Samples

Predicted: 000000000229.jpg



Predicted: 000000252219.jpg

Predicted: 000000039769.jpg



3) Histogram Equalization

The first preprocessing technique we applied was histogram equalization, in which the images' intensity histograms are readjusted to equalize the distribution of color intensities across the image to span the full range of intensity values.

This technique is helpful to enhance images, specifically those with poor lighting, making objects in the image appear brighter and clearer, which will possibly help the model distinguish objects better.

Sample Images (Before and After)

It can be seen in the sample images, especially the 2nd image that histogram equalization enhanced the image colors and made some of the intensities brighter or more contrasting as can be seen in the vase as the ridges are more define or in how the shadows below it are more contrasting compared to the light spots.

Model Performance Metrics

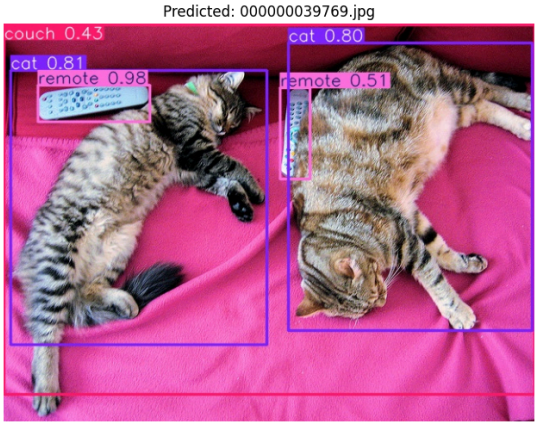
mAP50	mAP50-95	Precision	Recall
0.60002	0.43779	0.57998	0.55656

Model Output Samples

Predicted: 000000000229.jpg

Predicted: 000000252219.jpg

Predicted: 000000039769.jpg

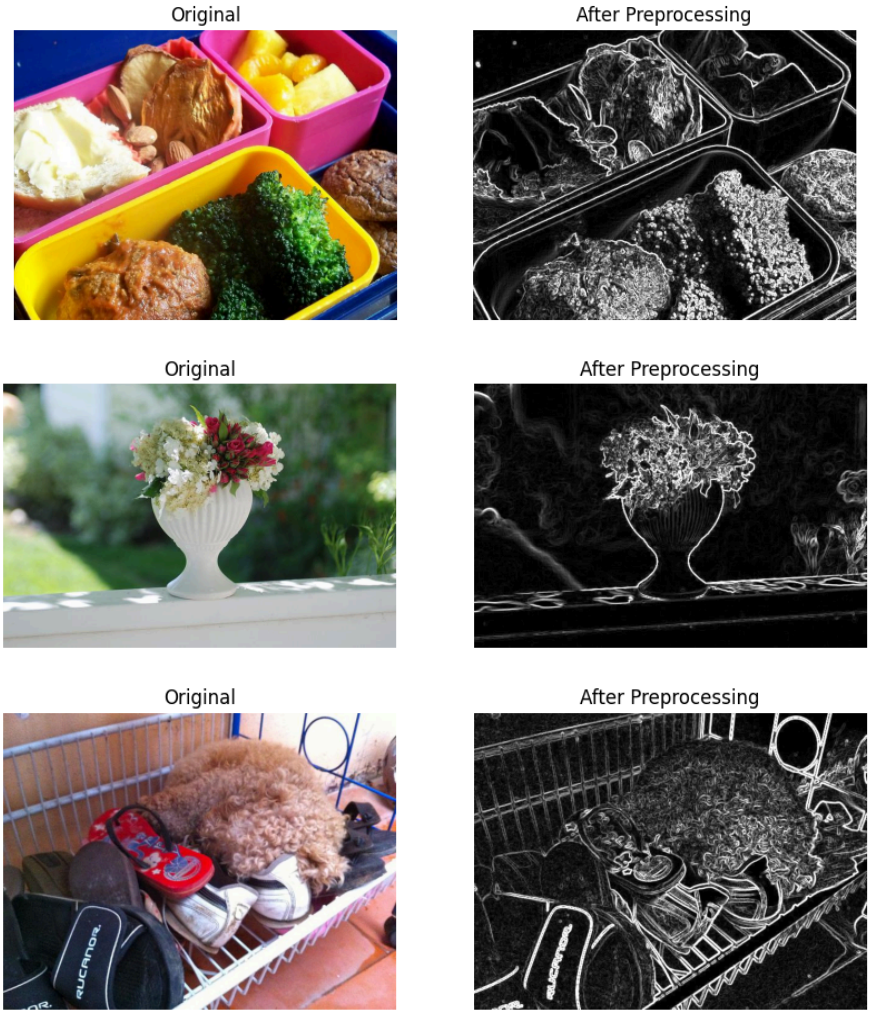


4) Sobel Edge Detection

The second preprocessing technique we applied was edge detection using the Sobel filter, which is a technique used to highlight object boundaries by detecting locations where the intensity of pixels changes abruptly.

Since object boundaries often correspond to edges, applying the Sobel operator can enhance the structural information in images, potentially helping the detection model separate objects from the background, especially in cases with low contrast, therefore allowing it to identify objects better.

Sample Images (Before and After)



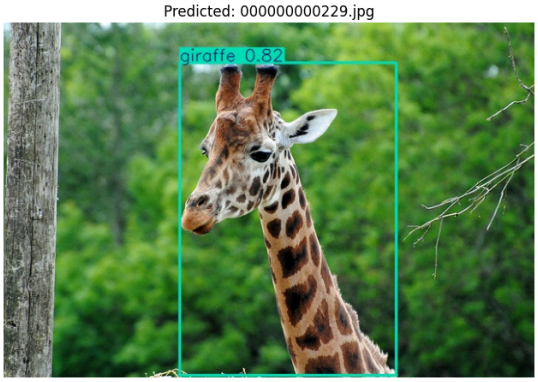
The processed images clearly show defined edges of the objects in the image retaining only structural information disregarding any intensity variations in the images only recognizing edge pixels (white) and non-edge pixels (black)

Model Performance Metrics

mAP50	mAP50-95	Precision	Recall
0.2951	0.19888	0.64009	0.2408

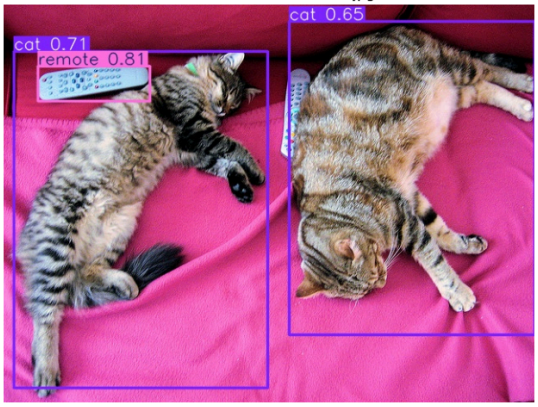
Model Output Samples

Predicted: 000000000229.jpg



Predicted: 000000252219.jpg

Predicted: 000000039769.jpg

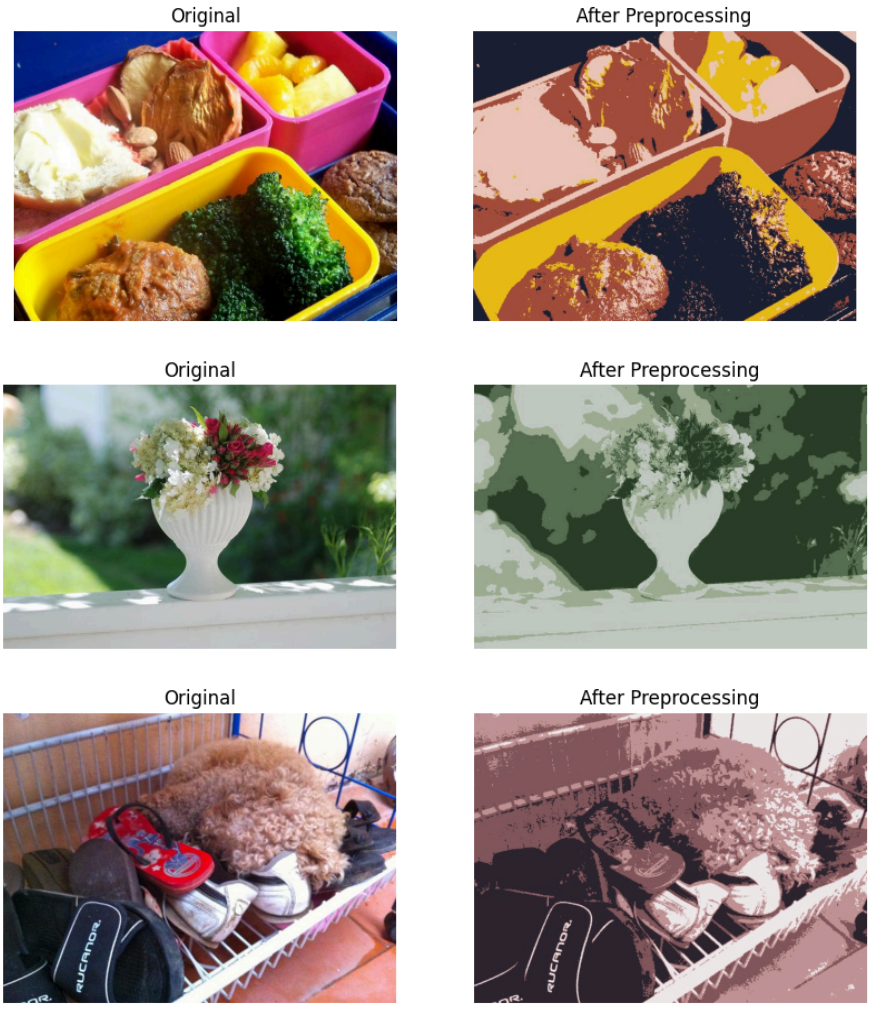


5) K-Means Image Segmentation

The third preprocessing technique we applied is K-Means image segmentation, in which the image pixels are clustered into groups based on color similarity. This reduces color variation and simplifies the image by flattening similar regions into the same color group and reducing noise while preserving object regions.

By simplifying textures and grouping similar regions, K-Means segmentation can make object boundaries more distinct and reduce background distractions. This may help the detector focus on the actual objects rather than getting caught up in fine background details.

Sample Images (Before and After)



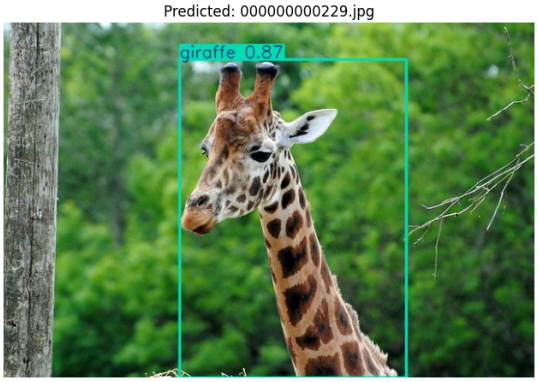
The processed images show the images segmented into much fewer color intensities representing the major intensity clusters in the image. This takes away from some of the finer details in the image providing a general view of the objects in the image which gives a balance between structural composition and intensities. However, it is noticeable in the 2nd and 3rd image that objects with similar intensities to the background somewhat blend into it.

Model Performance Metrics

mAP50	mAP50-95	Precision	Recall
0.35586	0.24039	0.60011	0.33486

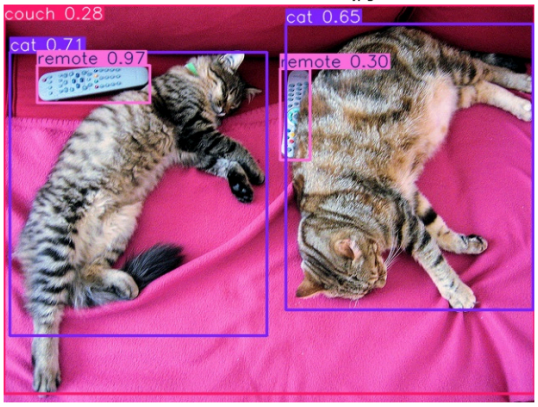
Model Output Samples

Predicted: 000000000229.jpg



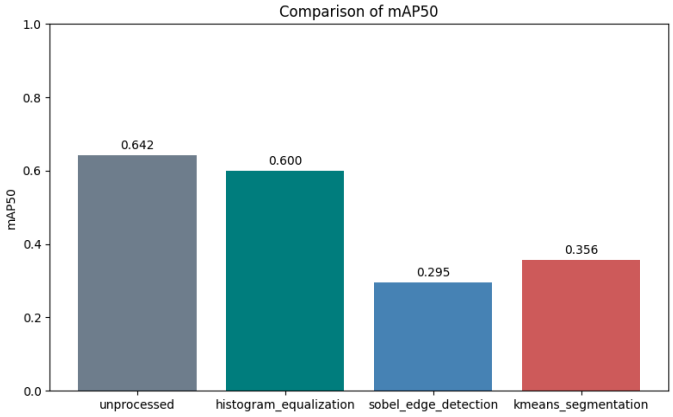
Predicted: 000000252219.jpg

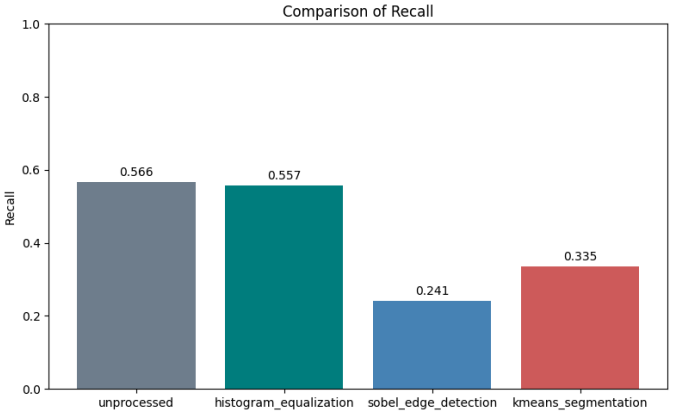
Predicted: 000000039769.jpg

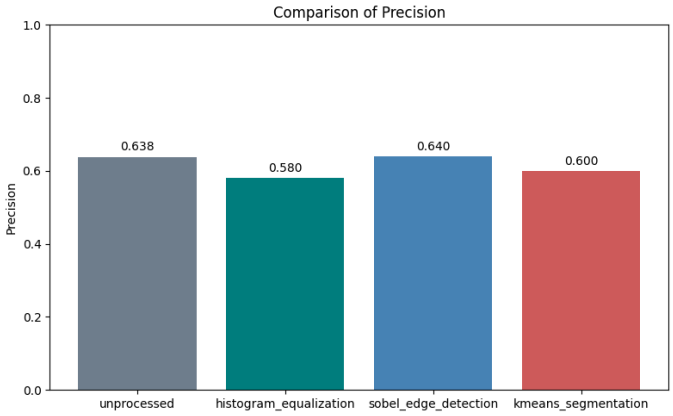


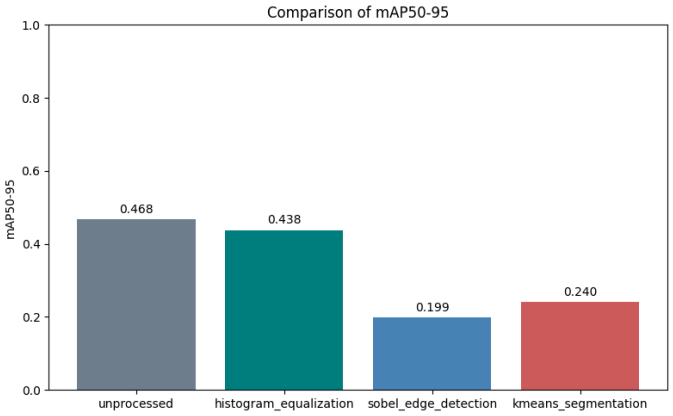
6) Metrics Comparison

	mAP50	mAP50-95	Precision	Recall
Unprocessed	0.64228	0.46769	0.63846	0.56557
Histogram Equalization	0.60002	0.43779	0.57998	0.55656
Sobel Edge Detection	0.29510	0.19888	0.64009	0.24080
K-Means Segmentation	0.36865	0.24216	0.54966	0.35670





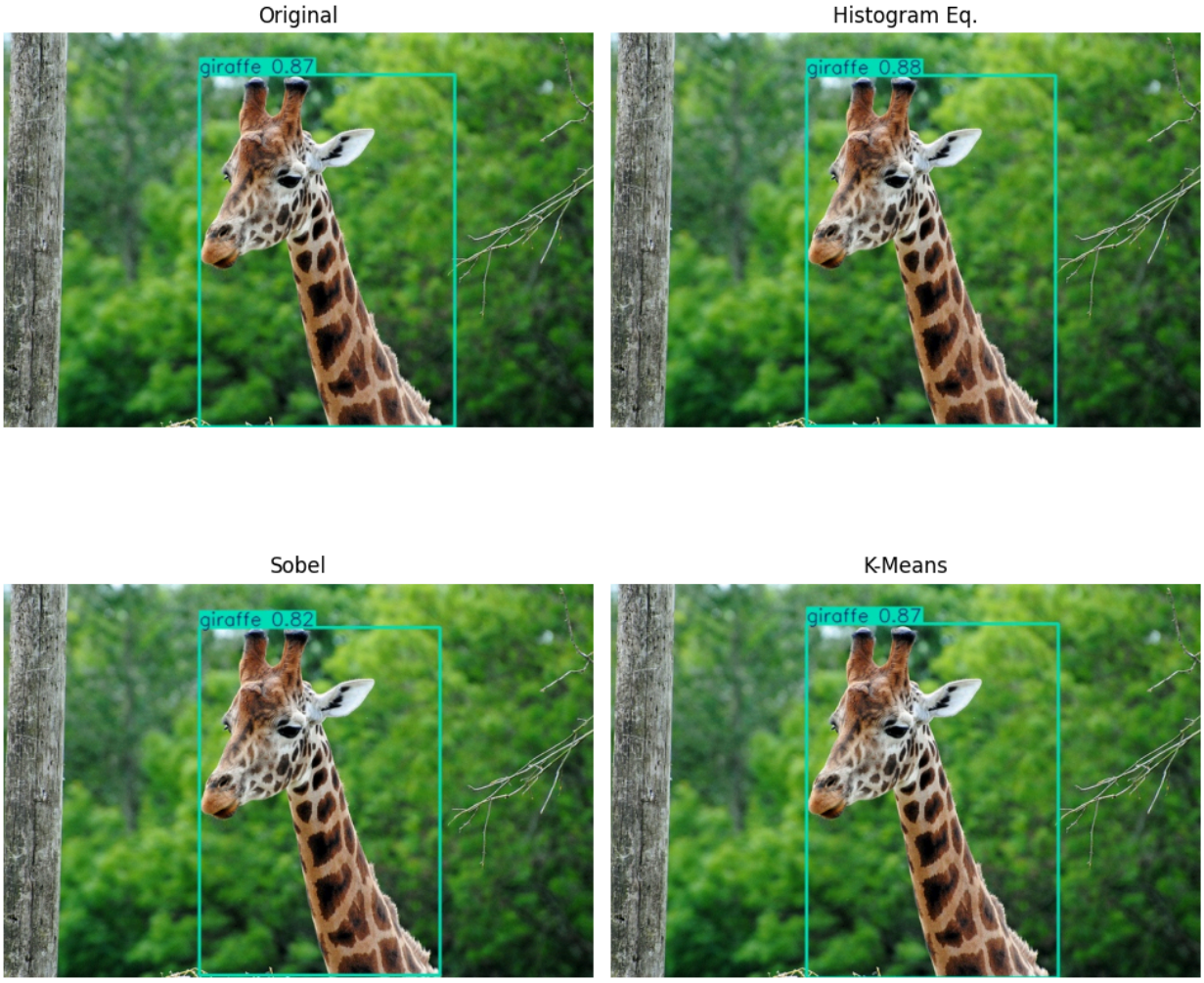




It is clear that although all preprocessing techniques did not improve the model's performance compared to the unprocessed images, histogram equalization resulted in the best training metrics, followed by k-means segmentation, followed by Sobel edge detection. This implies that the model relies on intensity variations during training and that relying only on structural information of the objects in the image is not enough for the model to learn to identify and recognize objects accurately.

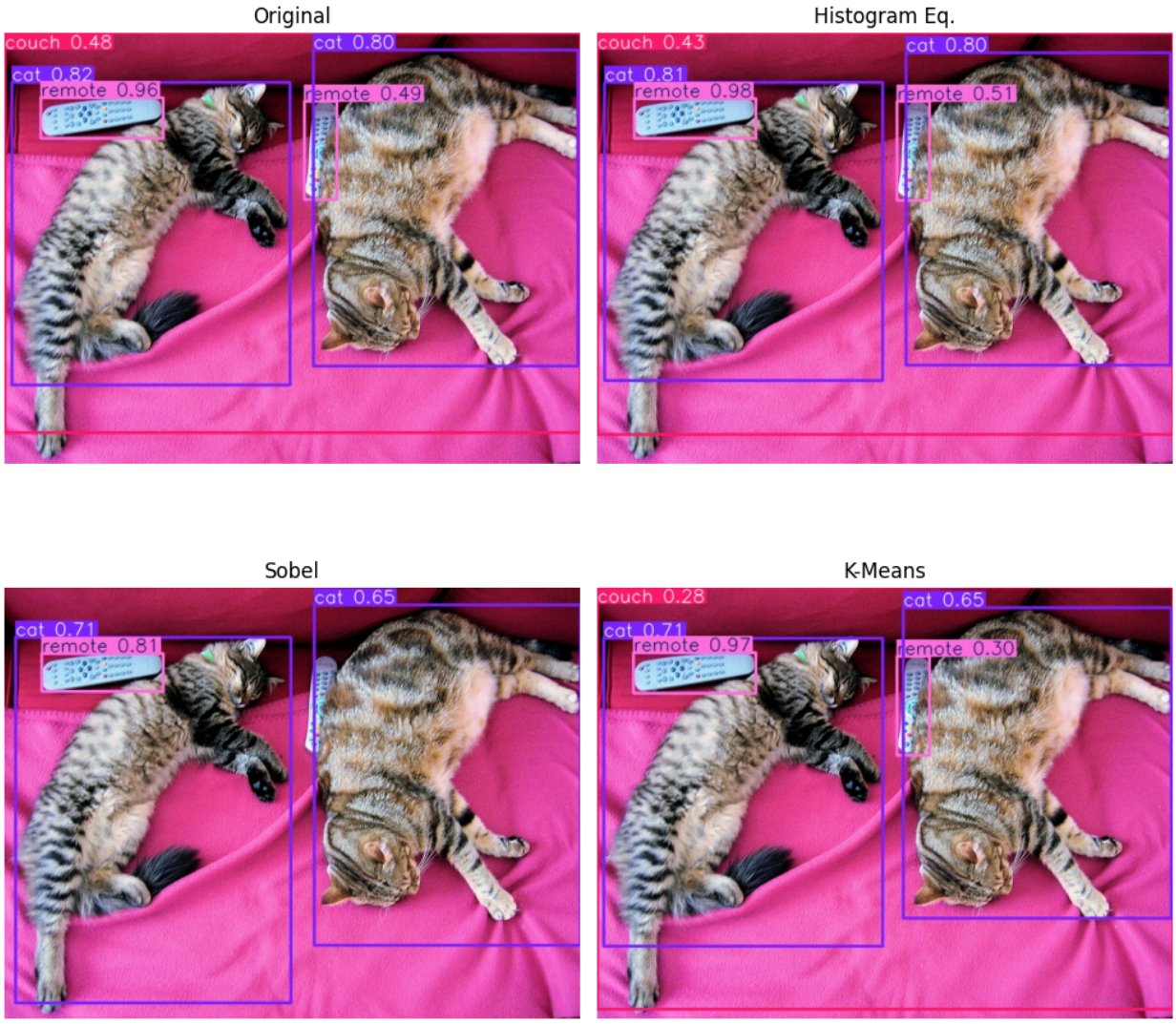
7) Output Comparisons

Test Image 1



All 4 models accurately recognize the giraffe but with different certainties with histogram equalization performing the best and the Sobel model performing the worst. However the differences are insignificant since this is a very clear picture with no distracting background objects, so all 3 preprocessing techniques most likely produced clear images with distinct object identification.

Test Image 2



All 4 models accurately recognize the cats, however the remotes are not recognized by all of them. For all 4 models, the left remote is more certainly identified, while the right remote isn't as accurately detected since it is slightly covered by the cat so its composition is not fully recognizable. The Sobel model failed to even recognize the 2nd remote's existence at all, possibly due to the fact that the intensity changes aren't a factor so the edges weren't properly detected as they are hidden by the cat. It also failed to recognize the couch since it is likely that the couch edges weren't highlighted and it was simply considered the background of the image.

Test Image 3

Original

Histogram Eq.

Sobel

K-Means

All 4 models seem to perform well at recognizing people as all identify the 3 people in the image with high certainty. The 4 models also recognize the traffic light and umbrella with similar certainties. Interestingly, the K-means model identified the shopping cart as a chair although it was unlabeled by all other models. This is likely due to the fact that after segmenting the image, the fine details of the cart were flattened and the general cluster shape resembled that of a chair causing it to inaccurately recognize it as such. The other models however clearly recognize the edges and details of the cart and therefore cannot classify it as a chair as it contains features that aren't visible in chairs.