ECS709P Introduction to Computer Vision Coursework 1 Discussion

Aditya Gupta (Student Number: 240754763)

November 2024

1 Understanding

During this coursework, I gained a deeper understanding of image transformations, convolution, and object detection techniques in computer vision. In the section on transformations, I applied rotation and skewing, which highlighted the importance of matrix formulation and how interpolation affects image quality. Through experimenting with different filters in convolution tasks, I discovered how the choice of kernel impacts image characteristics like smoothness and edge sharpness, reinforcing the connection between kernel structure and image processing outcomes. The use of histogram intersection in video segmentation provided valuable insight into frame-to-frame similarity, demonstrating a simple yet effective method to detect scene changes. Additionally, texture classification using LBP (Local Binary Patterns) descriptors emphasized how local texture patterns can be used to differentiate between image classes. Object detection through frame differencing and flood filling illustrated fundamental techniques for analyzing motion. I also learned how to use powerful image analysis tools like numpy and opencv. Overall, these methods have equipped me with a robust set of skills that I am excited to apply in the future.

2 Analysis

The results from each task revealed several important insights. In the transformations task, I observed that the order in which rotation and skew were applied significantly impacted the final output. This was due to the non-commutative nature of matrix multiplication, which illustrated how the sequence of transformations influences the results. Although I had learned this concept in my linear algebra courses, seeing it visualized and realizing its practical applications was very exciting. The convolution task demonstrated the effects of applying filters in different sequences. For instance, applying Gaussian blur before edge detection resulted in smoother but less distinct edges, whereas applying edge detection first emphasized details even after blurring. In video segmentation, the normalized histogram intersection values proved effective for detecting major scene changes, although this method may be limited in situations where frames have similar color distributions. The texture classification task indicated that using smaller window sizes for Local Binary Patterns (LBP) descriptors increased classification accuracy by capturing finer details. In contrast, larger window sizes tended to produce poorer representations that missed subtle features. Lastly, in the object detection task, denoising methods like Gaussian blur were successful in reducing false positives, proving effective for noise reduction. Overall, these observations underscore the importance of method selection and parameter tuning in achieving accurate results in computer vision.

3 Challenges

One significant challenge I faced during the course was the misinterpretation of tasks. The instructions were worded ambiguously and could be understood in multiple ways. In hindsight, I realize that some of the misinterpretations were my own fault, but this issue wasted a lot of time and forced me to experiment more than usual to achieve the same results as others.

During the convolution task, I faced an overflow and underflow issue with the colour channels during convolution. In one attempt to create a convolution function that avoided boundary issues, I inadvertently caused the colour channels to overflow and underflow, resulting in unwanted artefacts in the image. After several attempts, I resolved this by correctly assigning pixel values based on the proper dimensions and using the appropriate padding. This experience highlighted the importance of carefully managing colour channels in image processing.

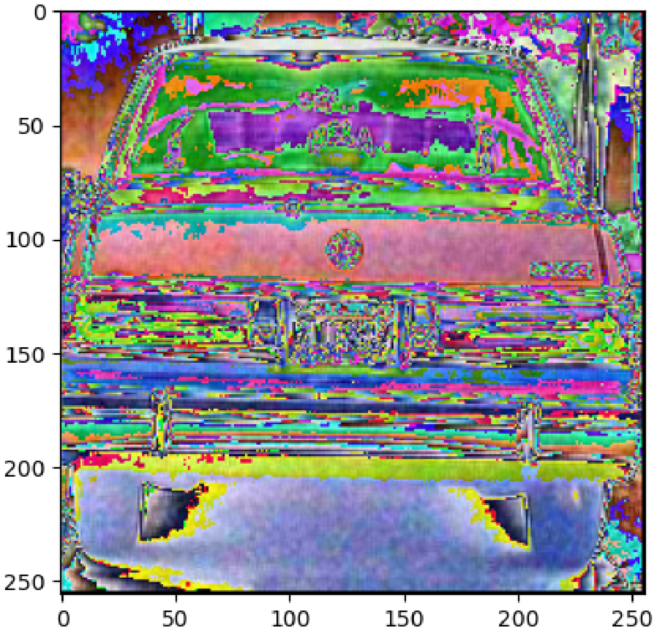

Figure 1: Result of Failed convolution filter which underflows the color channels

4 Mistakes

During this coursework, several mistakes helped refine my approach. In transformations, early attempts with image scaling led to parts of the image falling outside the frame, highlighting the need for better scaling and coordinate adjustments, even then the methods I used are not robust to multiple transformations being applied in succession. In convolution, I initially overlooked normalization after applying certain kernels, such as the Laplacian filter, which led to invalid pixel values. Correcting this by applying thresholding was necessary to produce usable images. In object detection, relying solely on frame differencing without noise reduction produced inflated object counts. Adding denoising with Gaussian blur helped reduce false positives. These mistakes provided valuable learning experiences in ensuring accuracy and robustness in image processing tasks.

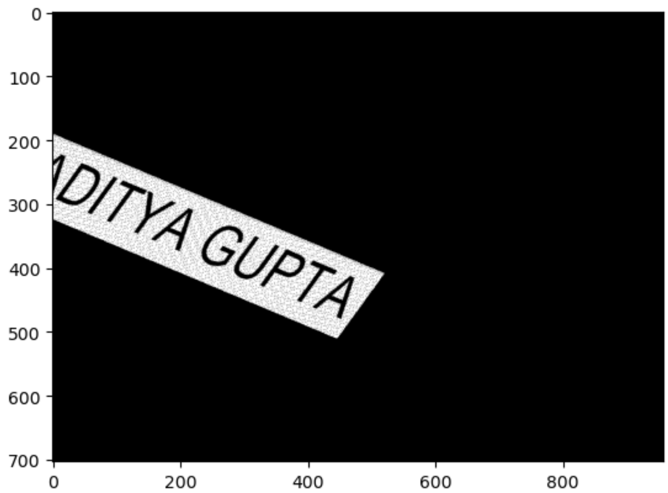

Figure 2: Image clipping due to invalid frame size

5 Discoveries

This coursework led to several discoveries about practical computer vision techniques. For video segmentation, the histogram intersection technique proved surprisingly effective for scene change detection without complex models, showing the potential of simple methods for temporal analysis. In texture classification, the effectiveness of LBP descriptors at capturing local patterns indicated that even basic descriptors can achieve high accuracy when configured correctly. The success of denoising methods in object detection revealed the importance of preprocessing in minimizing errors, especially in dynamic scenes with background noise. These discoveries highlight the practical applicability of basic computer vision methods and how fine-tuning can significantly enhance performance.