CS 639 Project Proposal

Team members:

Emma Liu, liu763@wisc.edu
Yufei Wang, ywang2395@wisc.edu

Problem Statement

Object detection and image processing for visual content creation.

Manual image processing of a recurring object of choice from a massive amount of digital media can be very tedious for the visual content creation process. Automated object detection and image processing can free visual content creators from doing repetitive tasks, such as blurring out the same object that occurs in more than 100 pictures, and allow them to focus more on the creative process. Therefore, we are planning to implement our own program which could detect user-specified objects in large amounts of images, and apply various image processing algorithms on them.

Why is the problem Important

With the expansion of social media and graphics software, visual content creation becomes accessible to non-professional content creators. Visual content creators usually use image processing softwares to make their content more visually pleasant to their target audiences or keep them safe from legal issues. However, when it comes to massive amounts of digital images, it would be scary for creators to manually process each image over the same task, and human error may occur. For example, a photographer who took a lot of photos of cars in the city may need to blur out license plates for all the cars. Manually applying mosaic on these pictures one by one is very time consuming, so a software that can automatically detect license plates and blur them would save this photographer a lot of time. We hope that by creating an automated object detection and image processing pipeline, we can make the visual content creation more accessible and less scary for people who want to be content creators, and help them focus on their creative process or content branding.

Automated object detection and image processing pipeline is a popular demand for non-content creators as well. In the medical field, a patient can understand better about their medical imaging result if the area of concern has been automatically detected and marked out by a colored rectangle border with text explanation. Traffic officers would easily spot a vehicle violating traffic laws if it's automatically detected with its license plate highlighted. Object detection and image processing pipeline can help people reduce extra labor and confusion in their daily routines. We may explore more possible applications of object detection and image processing in non-visual content creation fields if we have extra time.

Possible Solution

The problem could actually be solved in two stages - detect the object in the images and then do the image processing on the specified parts.
The problem definition of object detection is to determine where the object is located in a given image and which category each object belongs to.[1] Or in other words, it also consists of two parts: object localization and object classification. The location of an object is usually represented by its bounding box, which is the rectangle boundary of the object in the given image. Nowadays, a lot of deep learning frameworks have been developed for dealing with object detection problems. For instance, YOLO (You Only Look Once) is a real time object detection algorithm which could find the location of all objects in an image and classify them in just one pass.[2] Other framework like Mask RCNN [3] which could not only detect the object, but also find its semantic segmentation by giving its pixel-level mask. For solving our problem, we firstly need to detect the recurring objects in large amounts of images, **therefore we are planning to investigate the existing algorithms and re-use them, or even implement our version of the model and make some improvements if we have enough time.**
After we find the object in the images, we could do some image processing on the certain area by referring to the bounding box of objects. We have some rough ideas on the image processing part, like mosaic the detected parts - this is a common need when processing videos, for instance in some situations we want to blur all human faces in the video. We could also implement more algorithms for this part, like adding stickers or replacing the texture of the object. This seems easy but it could also be challenging, as we only want to process part of the image and this part should still be consistent with the whole image after processing.

Rough Time Table

10/1 - 10/10 paper reviews and find the algorithm we want to use or implement
10/10 - 10/30 implement the object detection algorithm and test it on our own dataset
11/1 - 11/15 implement the image processing algorithms and write programs to handle video processing
11/16 - 11/30 Wrap all things and write the final report

References

[1] Zhao, Zhong-Qiu, et al. "Object detection with deep learning: A review." IEEE transactions on neural networks and learning systems 30.11 (2019): 3212-3232.
[2] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[3] He, Kaiming, et al. "Mask r-cnn." Proceedings of the IEEE international conference on computer vision. 2017.