Semi-supervised Spectral Image Annotation

Dhruv Jain
Department of Computer Science and Engineering - IIT Bombay, India

***Abstract*—Semi-supervised image annotation involves labeling images using both labeled and unlabeled data. In this report, we discuss semi-supervised image annotation using multispectral imagery and unlabeled data. Due to the difficulty in obtaining large amounts of labeled data, we propose a solution using deep learning and active learning methods. Our approach involves using pre-trained models and image processing techniques to train an annotation model. We then select the most informative unlabeled samples for automatic labeling to improve the model's performance. Additionally, we explore the use of multispectral imagery in homes and public facilities, including creating datasets and annotating images using multispectral imagery-based segmentation. This involves generating synthetic infrared images and utilizing an infrared vegetation index for automatic annotation.**

I. Introduction

Multispectral imagery is a technology that captures light in different wavelength bands and allows visualization of information beyond the visible spectrum. Our report proposes a semi-supervised approach to image annotation that combines multispectral imagery and unlabeled data. Our methods are used to automate the annotation process and reduce the need for manual labeling.

The report discusses the importance and challenges of image annotation and proposes a solution using semi-supervised and unsupervised techniques. The approach involves selecting informative samples and training the annotation model with limited data using automatic labeling. Experimental results demonstrate the effectiveness of the proposed method in obtaining labeled data for multispectral imagery-based image processing. The approach can reduce the time and resources required for manual labeling and has various applications, such as detecting foreign matter in food factories and identifying early-stage fruit damage. (Fig. 1)

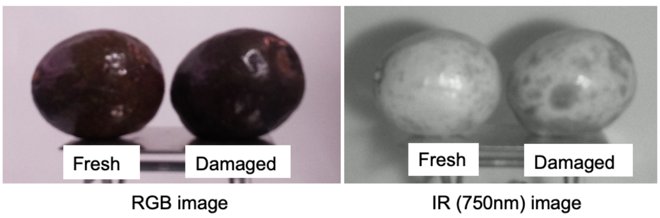

Fig. 1: Visibility of Rottenness

II. Background and Related Work

In machine learning, Inputs are the features of the data, and labels are the correct outputs the model should predict. Image annotation is the labeling of images with metadata, often for object recognition or segmentation. The manual annotation process is time-consuming, necessitating automation. Deep learning-based methods, like SegNet and Unet, have been used to segment crops from infrared images captured by drones. Object recognition tasks use far-infrared and near-infrared images with object detection algorithms like YOLO.

This paper has been written as part of a Paper Writing Skills Assignment.

III. Implementation

Discussed aspects of hardware and software requirements.

A. Camera configuration

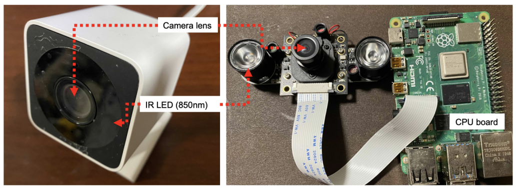

Fig. 2: Camera Setup

The decreasing cost of IP cameras and increasing processing power combined with faster mobile networks are making machine learning-based image processing on edge devices and via networks more technically easier. (Fig. 2)

B. Semi-Supervised annotation algorithm

To determine the region of interest in an image and annotate it automatically for vegetation, the algorithm uses RGB and Infrared images to calculate the NDVI index, which gets converted into a greyscale image using a thresholding filter. Contours of the vegetation portion of the image are found using masked regions of the generated binary image. Contours are saved in LABELME imaging format and converted into COCO format for instance segmentation tasks. The algorithm uses infrared images to automatically segment vegetation ROI using NDVI-image-process-filtering. (Fig. 3)

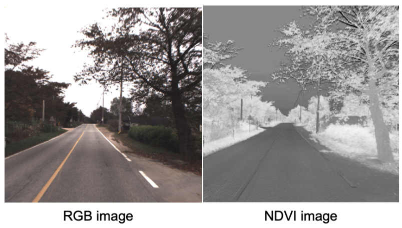

Fig. 3: Processing RGB into NDVI

C. Various formats and Explanations

LabelMe is an XML-based format that is easy to visualize and edit, and it stores the location of each object. COCO (Common Objects in Context) is a JSON-based format that is used for object detection and segmentation tasks. It also stores the location of each object and provides additional information such as masks, and key points. We trained an instance segmentation model using Mask-RCNN on a RANUS dataset which includes NIR images too. The same image subtraction technique can be used to annotate any object, and the annotated images can be stored in either LabelMe or COCO format depending on the application.

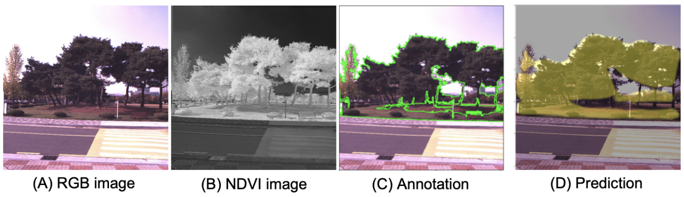

Fig. 4: Auto-annotation using NDVI

D. Comparision and Error Correction

Image Type	Average Accuracy
RGB (Not using NDVI)	65%
NIR (Using NDVI)	81.8%
Mask-RCNN-GAN Model	86%

TABLE I: Comparison Results

Table 1. shows the comparison results for different image types. The highest average accuracy was achieved on RGB+NIR images, with a score of 86%. The average RMSE for all the image types was 2.97. These results suggest that by using NIR, we can not only automate the process of annotation but also improve the accuracy of detection models.

Formula:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{1}$$

The report examines different methods to validate and fix errors in image annotations. Pixel thresholding involves setting a threshold value to check the accuracy of the annotation. Contour selection involves selecting the contours of the object, while binary translation converts the image into a binary format for easier analysis. Manual control involves human experts manually inspecting annotated images for errors. The report also examines methods to handle cases where no labels are found in the image. These include manual intervention, relaxing threshold constraints, selective operations, or discarding the sample. It is recommended to use these methods to ensure minimized errors while annotating images. (Fig. 4 check the steps)

IV. Imagery applications

Multispectral imagery has a range of applications in household and commercial settings. In households, multispectral imagery can be combined with machine learning to identify and predict the freshness of food on a dining table, which may be useful for elderly parents (Fig. 5). In commercial applications, multispectral imagery can be used to continuously monitor the damage to structures such as bridges and roads, by identifying the progress of corrosion in concrete and rust on steel. This has the potential to improve maintenance and prolong the lifespan of these structures.

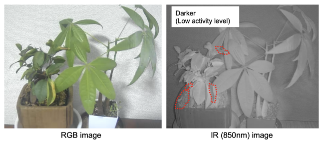

Fig. 5: Household Plants Monitory

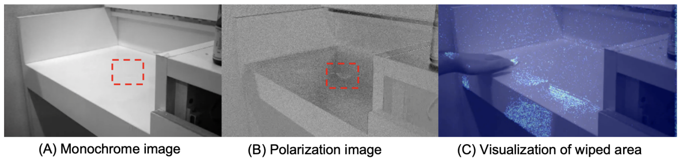

Fig. 6: Commercial Restaurants Monitory

V. Conclusions and FUTURE WORK

To conclude, the research findings are promising and suggest various avenues for future exploration. These include the development of instance segmentation models, exploration of multispectral imagery in different applications, investigation of hyperspectral cameras, innovative techniques for image annotation, evaluating system performance under various conditions, and exploring additional machine learning techniques.

The study also proposed methods for making multispectral cameras more affordable for use in homes and public facilities, allowing for the detection and identification of a more extensive range of objects in various environments. The potential for expanding the diversity of objects analyzed by multispectral imagery was demonstrated through plant annotation with multispectral imagery-based segmentation. By extracting images with moderately less noisy patterns, the range of analyzable objects can be expanded at a minimal cost. (Fig. 4)

With further advancements in technology, it is expected that multispectral imagery and machine learning-based object detection will continue to revolutionize the way we observe and interact with our surroundings. (Fig. 6 Technology in-use)

References

[1] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. 2017. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Transactions on pattern analysis and machine intelligence 39, 12 (2017), 2481–2495.

[2] T Barros, P Conde, G Gonçalves, C Premebida, M Monteiro, CSS Ferreira, and UJ Nunes. 2021. Multispectral Vineyard Segmentation: A Deep Learning Approach. arXiv preprint arXiv:2108.01200 (2021)

[3] Barbara Boldrini, Waltraud Kessler, Karsten Rebner, and Rudolf W Kessler. 2012. Hyperspectral imaging: a review of best practice, performance and pitfalls for in-line and on-line applications. Journal of near-infrared spectroscopy 20, 5 (2012), 483–508.