

# Diminishing Unwanted Objects Based on Object Detection Using Deep Learning and Image Inpainting

Chanran Kim, Younkyoung Lee, and Jong-Il Park<sup>†</sup>

Computer and Software  
Hanyang University  
Seoul, Republic of Korea  
{chanrankim, dldusrud1995, jipark}@hanyang.ac.kr

Jaeha Lee

APS System Engineering Team  
Hyundai Mobis  
Seoul, Republic of Korea  
jaeha1223@mobis.co.kr

**Abstract**—Photography is not done in perfect circumstances. Especially, it is difficult to shoot the only desired object in many circumstances. This paper proposes a method to diminish unwanted objects in photographs. We propose a basic design method of diminished reality using deep learning. Proposed method composed of object detection using deep learning and exemplar-based image inpainting. Previous methods should specify an undesired range before or during operation. Our method has the advantage that it is possible to diminish with the object label to diminish without having to select the area to diminish. The experimental results show naturally diminished scene of the proposed method.

**Keywords**— *Diminished Reality, Object Detection, Image Inpainting, Deep Learning, Integrated System*

## I. INTRODUCTION

There are a lot of pictures that contain unwanted objects. It would be nice to have an environment where you can get rid of all the unwanted objects before shooting, but it is difficult in usual. This paper makes this process possible as a post-processing process.

Beginning with AlexNet [1], which has surpassed human capabilities in image classification, methods using Convolutional Neural Networks (CNNs) have been concentrated and expanded into various fields [2]. Object detection is to find objects in a scene and classify them. R-CNN has been proposed to enhance object detection performance using CNN, and fast R-CNN and YOLO have been proposed for the purpose of speed up [3, 4]. We use YOLO to find objects that exist in the scene.

Augmented reality is making a lot of progress. However, the process of understanding and preparing the scene for augmentation is very scarce. Augmenting an object in a crowded scene where there are many objects should be avoided for the completeness of the scene. Therefore, we need diminished reality to remove unnecessary objects from the scene. Jan Herling et al. proposed an object removal method for diminished reality that selects and erases the desired part and fills that part [5]. In contrast, in this paper, we propose a method to find and remove objects without selecting them.

## II. PROPOSED METHOD

We propose diminishing unwanted objects method for diminished reality. First, detect all possible objects in a scene using YOLO. Converts the region of objects of the specified label to black and make a mask for it. Finally, image inpainting using the image and the mask is performed to fill the empty(black) part of the image (Fig. 1).

### A. Object Detection

You only look once(YOLO) is a state-of-the-art and real-time object detection system. Ultimately, we perform object detection based on YOLO which shows good performance in terms of speed to perform diminished reality. YOLO computes the prediction information with the whole image at once in the test time [6]. Predictions are made very quickly by evaluating a single network. Nevertheless, it yields reasonable performance.

### B. Image Completion

We inpainted the area of unwanted objects which were found in object detection process. Image completion models using deep learning are introduced, but they work only in limited image sizes, or they are often performed only at a specified location. Also, there is no real-time oriented model like YOLO of object detection. Therefore, this paper uses exemplar-based method. Criminisi's method used a single image to propagate texture and structure information simultaneously [7]. It uses a block-based sampling process to speed up the process. Our image completion is based on Criminisi's method. Algorithmically, it is not much different from the Criminisi's method because we focus on interworking for the whole system.

### C. The Intermediate Process

This section covers the intermediate process for interlocking object detection with image completion. This intermediate process is divided into the process of removing the selected label area and the process of preparing for inpainting. The process of selected label area should be preprocessed. Determines the label on which unwanted objects are removed. When removing the determined label area, the proposed method forms a mask. This process is preparation for image completion.

<sup>†</sup>Corresponding author

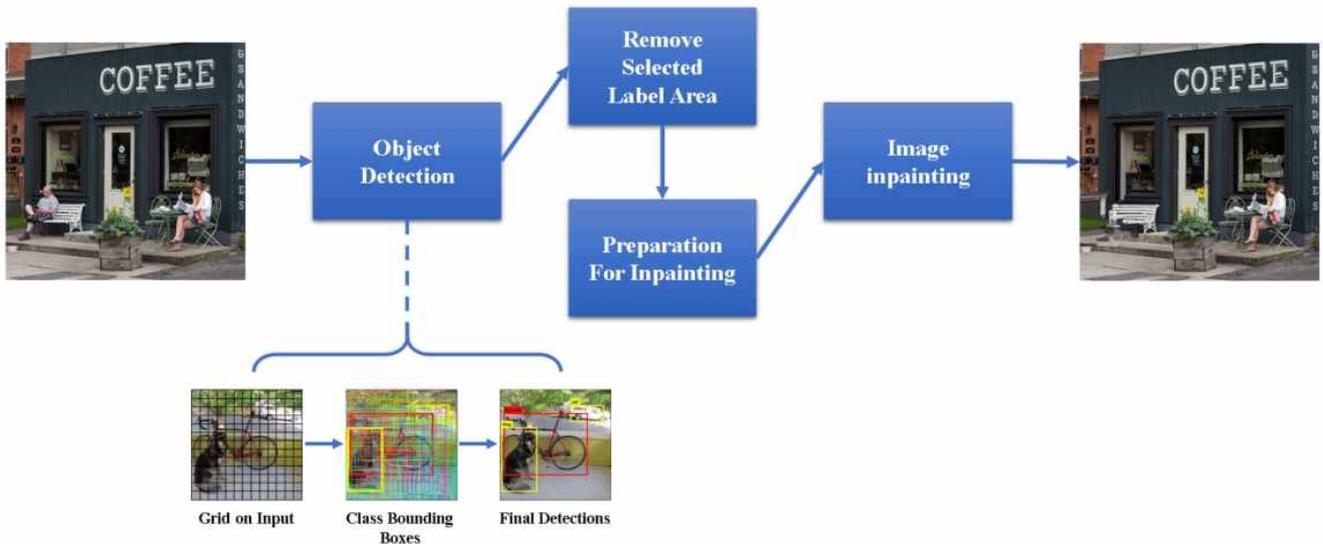


Fig 1. A Diagram of Proposed Method

### III. EXPERIMENTS

To test the proposed method, it is necessary to prepare an appropriate image. You must enter an image that contains an object that can be removed because it cannot be removed except for the learned or learned label. The input image is labeled with the detected object through the object detection method, and the area corresponding to the unwanted label is rearranged through the intermediate process. The proposed method is terminated by image completion. Figure 2 shows the method step by step. It cannot work in real time, but you can diminish unwanted objects using this model every frame of the image.

Figure 3 shows a failure case. When the unwanted object is



Fig 2. A Process of Proposed Method

overlapped by a lot of weight, it is hard to separate. Object detection is incomplete and image inpainting fails noticeably.

### IV. CONCLUSION

In this paper, we proposed a method for diminishing unwanted objects using object detection and image inpainting. Used object detection method is based on the YOLO and used image completion method is exemplar-based. Experiments show that the proposed method works well. We represent a basic design method of diminished reality by integrating object detection and image completion. our future work is to create a network in which object detection and image completion have an organic relationship.

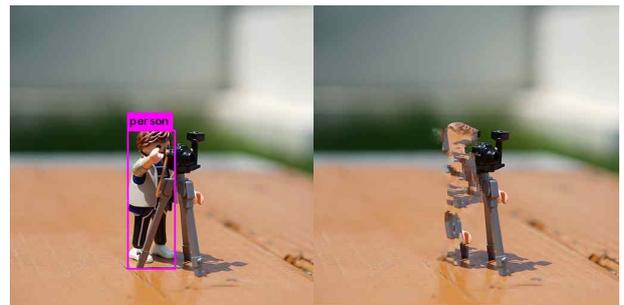


Fig 3. A Failure Case

### ACKNOWLEDGEMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT). (No. 2017-0-01669, Development of original technology for practical application of AR in various real environment)

This research was supported by BK21 Plus project of the National Research Foundation of Korea Grant.

## REFERENCES

- [1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," In Proc. of Advances in Neural Information Processing Systems (NIPS), pp. 1097-1105, 2012.
- [2] Jürgen Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85-117, 2015.
- [3] Ross Girshick, "Fast r-cnn," In Proc. of the IEEE International Conference on Computer Vision (ICCV), pp. 1440-1448, 2015.
- [4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You only look once: Unified, real-time object detection," In Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.
- [5] Jan Herling, and Wolfgang Broll, "Advanced self-contained object removal for realizing real-time diminished reality in unconstrained environments," In Proc. of the 9<sup>th</sup> IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 207-212, 2010.
- [6] Joseph Redmon and Ali Farhadi, "YOLO9000: better, faster, stronger," arXiv preprint arXiv:1612.08242, 2016.
- [7] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting", IEEE Transactions on image processing, 13(9), pp. 1200-1212, 2004.