

Selective labeling using image super resolution for improving the efficiency of object detection in low-resolution oriental paintings

Hyeyoung Moon*, Namgyu Kim*

*Graduate Student, Graduate School of Business IT, Kookmin University, Seoul, Korea

*Professor, Graduate School of Business IT, Kookmin University, Seoul, Korea

[Abstract]

Image labeling must be preceded in order to perform object detection, and this task is considered a significant burden in building a deep learning model. Tens of thousands of images need to be trained for building a deep learning model, and human labelers have many limitations in labeling these images manually. In order to overcome these difficulties, this study proposes a method to perform object detection without significant performance degradation, even though labeling some images rather than the entire image. Specifically, in this study, low-resolution oriental painting images are converted into high-quality images using a super-resolution algorithm, and the effect of SSIM and PSNR derived in this process on the mAP of object detection is analyzed. We expect that the results of this study can contribute significantly to constructing deep learning models such as image classification, object detection, and image segmentation that require efficient image labeling.

▶ **Key words:** object detection, deep learning, image labeling, super resolution, SSIM

[요 약]

이미지에 레이블을 부착하는 레이블링은 객체 탐지를 수행하기 위해서는 반드시 선행되어야 하며 이러한 작업은 딥러닝 모델을 구축하는 데 있어서 큰 부담으로 여겨지고 있다. 딥러닝 모델을 훈련하기 위해서는 수 만장의 이미지가 필요하며 이러한 이미지에 인간 레이블러가 직접 레이블링을 진행하기에는 많은 한계가 있다. 이러한 어려움을 극복하기 위해 본 연구에서는 전체 이미지가 아닌 일부 이미지에 대한 레이블링을 통해서도 큰 성능의 저하 없이 객체 탐지를 수행하는 방안을 제안한다. 구체적으로 본 연구에서는 저품질 동양화 이미지의 객체 탐지를 위해 초고해상화 알고리즘을 이용하여 저해상도의 이미지를 고화질의 이미지로 변환하고, 이 과정에서 도출되는 SSIM과 PSNR이 객체 탐지의 mAP에 미치는 영향을 분석하여 객체 탐지 분석에 필요한 레이블링을 위한 최적의 샘플링을 수행하는 방안을 제안한다. 본 연구의 결과는 이미지 레이블링을 필요로 하는 이미지 분류, 객체 검출, 이미지 분할 등 딥러닝 모델 구축에 크게 기여할 수 있을 것으로 기대한다.

▶ **주제어:** 객체 탐지, 딥러닝, 이미지 레이블링, 초고해상화, SSIM

-
- First Author: Hyeyoung Moon, Corresponding Author: Namgyu Kim
 - *Hyeyoung Moon (data@kookmin.ac.kr), Graduate School of Business IT, Kookmin University
 - *Namgyu Kim (ngkim@kookmin.ac.kr), Graduate School of Business IT, Kookmin University
 - Received: 2022. 07. 27, Revised: 2022. 08. 29, Accepted: 2022. 09. 06.

I. Introduction

최근 딥러닝(Deep Learning)기술[1]의 발전으로 이미지 분류, 객체 탐지, 세그멘테이션, 강화학습, 그리고 GAN 등 다양한 방식으로 이미지를 분석하는 연구[2]가 활발하게 수행되고 있으며, 특히 딥러닝 기술을 이용하여 손실되었거나 저해상도로 저장된 오래된 이미지들을 선명한 고화질의 이미지로 복원하는 초고해상화에 대한 연구가 많은 관심을 받고 있다. 초고해상화 기술은 이미지 검색, CCTV복원, 위성 이미지, 의료 이미지, 영화 복원 등 광범위한 산업에서 널리 사용되고 있다. 특히 동양화를 포함한 많은 역사 사료들은 저해상도로 보관된 경우가 많은데, 해상도가 낮거나 손상된 부분을 사람이 직접 복원하기 위해서는 많은 시간과 비용과 노력이 소요된다. 따라서 손상이 복합적으로 존재하는 오래된 이미지를 보다 자연스럽고 선명한 이미지로 자동 복원하기 위해 최신 딥러닝 알고리즘을 접목한 초고해상화 기술이 사용된다.

저해상도 이미지에 대한 초고해상화 작업은 오래 전부터 꾸준히 연구가 이루어져 왔으며, 특히 최근 딥러닝 기반 초고해상화 알고리즘들은 기존의 전통적인 기법들에 비해 화질 개선 측면에서 우수한 성능을 보이고 있다. 딥러닝 기반 초고해상화 기법은 SRCNN[3]에서부터 시작하여 최근에는 생성적 적대 신경망(Generative Adversarial Network, GAN)을 초고해상화에 접목한 SRGAN[4], ESRGAN[5], 그리고 Real-ESRGAN[6] 등으로 계속된 발전을 거듭하고 있다. GAN[7]은 가짜 이미지를 만드는 생성자(Generator)와 주어진 이미지가 진짜인지 여부를 판별하는 판별자(Discriminator)를 사용하여, 진짜 이미지와 구분하기 힘든 가짜 이미지를 새롭게 생성하기 위해 생성자와 판별자를 동시에 학습하는 알고리즘이다. GAN 기반의 초고해상화 알고리즘들은 기존 알고리즘들에 비해 매우 우수한 성능을 나타내고 있다.

이미지 복원 성능에 대한 객관적인 평가는 복원 이미지와 원본 이미지의 비교를 통해 이루어지며, 구체적으로는 PSNR(Peak Signal to Noise Ratio)과 SSIM(Structural Similarity Index Measure)의 두 가지 평가 척도가 주로 사용된다. PSNR[8]은 MSE(Mean Squared Error)[9]를 통해 복원 이미지와 원본 이미지의 화소값이 얼마나 유사한지를 평가하고, SSIM[10]은 수치적인 차이가 아닌 인간이 시각적으로 느끼는 화질의 차이, 즉 이미지의 휘도(Luminance), 대비(Contrast), 그리고 구조(Structure)의 유사도를 평가한다. 이러한 척도를 통해 이미지 복원의 성능을 평가할 수 있다.

특히 오래된 영화와 같이 화질이 좋지 않은 이미지의 경

우 초고해상화에 의한 복원기술이 더욱 중요하게 사용될 수 있다. 초고해상화로 복원한 이미지를 분석에 사용하는 경우 원본 이미지에서는 인식이 거의 불가능한 객체도 인식할 수 있고, 전반적으로 대부분의 객체를 보다 선명하게 탐지하여 모델의 성능을 향상시키는 것으로 알려져 있다. 이미지에서 관심 객체를 배경과 구분해서 식별하는 기법인 객체 탐지는 딥러닝을 기반으로 한 AlexNet[11]을 시작으로 YOLO[12], Mobilenet-SSD[13], Faster RCNN[14] 등 다양한 알고리즘으로 구현되면서 발전을 거듭하고 있다. 특히 YOLOv5(You Only Look Once)는 원본 이미지를 동일한 크기의 그리드(Grid)로 나누고 위치와 클래스를 한번에 결정하는 단일-단계(One-Stage) 방식으로 객체 탐지를 수행하여, 속도와 정확도 측면에서 우수한 성능을 나타내고 있다.

이러한 객체 탐지를 수행하기 위해서는 이미지에 레이블(Label)을 부착하는 레이블링(Labeling)[15] 작업이 반드시 선행되어야 하며, 이는 이미지 분석의 큰 부담으로 여겨지고 있다. 레이블링은 기능이나 목적에 부합하는 정보를 원본 이미지에 부착하는 작업을 의미하며, 딥러닝에서 모델을 훈련시키기 위해서는 수천, 수만장의 레이블링된 이미지가 필요하다. 레이블링 작업을 마친 방대한 이미지는 학습을 통해 딥러닝 모델을 구축하는 데 사용되며 추가적인 데이터가 필요한 경우에는 이러한 레이블링 작업을 다시 반복적으로 수행하게 된다.

이미지 레이블링 작업은 각 이미지에 대해 사람이 직접 수작업으로 진행하게 되어 많은 시간과 노력이 소요되며, 해결하고자 하는 문제가 복잡할수록 필요 데이터 셋의 복잡도와 레이블링 작업은 난이도도 증가하게 된다. 또한 이러한 레이블링 작업은 작업을 수행하는 사람의 경험, 역량, 그리고 주관적 판단의 영향을 받게 되므로, 여러 사람이 작업에 공동으로 참여하는 경우 결과의 일관성이 저하되어 결국 모델의 학습에도 부정적인 영향을 주게 된다. 연구자가 원활하게 데이터를 확보하기 어려운 경우 인공지능 기술개발은 더디게 진행될 수밖에 없으며, 따라서 선진 각국 정부는 주도적으로 데이터를 제작하고 공유하기 위한 노력을 기울이고 있다. 일반적인 AI 프로젝트에서 학습 데이터의 준비 과정이 전체 프로젝트의 시간의 약 80%를 차지하고 있으며, 데이터의 수요와 더불어 학습 데이터 제작을 위한 레이블링 시장은 지속적으로 성장하고 있다[16-18].

이처럼 딥러닝 모델을 개발하기 위해서는 대량의 학습 데이터 셋이 반드시 필요하며, 대량의 데이터 셋에 대한 레이블링하는 작업에는 전술한 바와 같이 많은 비용이 소요된다. 본 연구에서는 이러한 어려움을 극복하기 위한

효율적 레이블링 방안을 제시한다. 구체적으로 본 연구에서는 레이블이 구축되어 있지 않은 저해상도 이미지의 효율적 객체 탐지 방안을 제안한다. 제안 방법론은 초고해상화 알고리즘을 이용하여 저해상도의 이미지를 고화질의 이미지로 변환하고, 이 과정에서 도출되는 다양한 지표를 활용하여 객체 탐지 분석에 필요한 이미지 데이터의 레이블링을 위한 최적 샘플링을 수행한다.

본 논문의 이후 구성은 다음과 같다. 우선 다음 장인 2장에서는 딥러닝 기반 이미지 분석 관련 기존 연구를 소개한다. 다음으로 3장에서는 본 연구에서 제안하는 전체 방법론을 소개하며, 제안 방법론을 적용한 실험 결과는 4장에서 소개한다. 마지막 장인 5장에서는 본 연구의 기여와 한계를 요약한다.

II. Related Research

1. Super Resolution

초고해상화를 위한 GAN 모델 중 ESRGAN(Enhanced Super Resolution Generative Adversarial Networks)[4]은 이전의 SRGAN에서 사용한 배치 정규화(Batch Normalization)[19] 층을 제거하고 RRDB(Residual in Residual Dense Block)[20]를 깊게 쌓은 모델이다. 배치 정규화는 배치 단위로 평균과 분산을 구해서 정규화를 하는 방법으로 연산을 위해 많은 컴퓨터 자원을 사용해야 하는 단점이 있다.

Real-ESRGAN 모델은 ESRGAN의 Super Resolution 모델을 기반으로 다양한 전처리 열화기법 단계를 확장하였다. 또한 ESRGAN의 모델에 Pixel-Unshuffle를 기법을 적용하여, 입력되는 이미지의 사이즈를 줄이면서 채널의 수를 늘려서 GPU 소모량과 연산량을 줄이는 단계를 추가하였다 (Fig. 1)

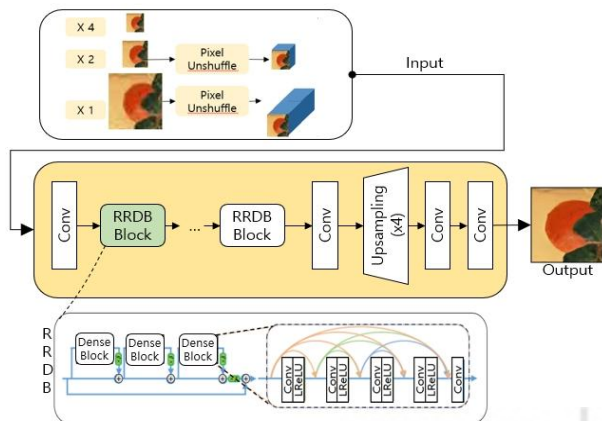


Fig. 1. Real-ESRGAN Generator Network

초고해상화로 생성된 이미지에 대한 객관적인 화질 평가 방법에는 대표적으로 PSNR[8]과 SSIM[10]이 있다. PSNR은 원본 이미지와 비교하여 수치적 차이로 이미지의 품질을 평가하는 방법으로 수식은 (식 1)과 같으며, 여기서 MSE(Mean Square Error)는 오차 제곱의 평균으로 두 이미지에서 같은 위치에 있는 데이터값의 분산을 나타낸다. MSE 값이 작다는 것은 원본과 변환본이 매우 유사하다는 것을 나타내며 MSE 값이 작을수록 PSNR은 높게 나타난다.

(식 1) PSNR

$$PSNR = 10 \times \log_2 \left(\frac{R^2}{MSE} \right)$$

$$\left(MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2, R : \text{maximum in pixel} \right)$$

SSIM은 압축 및 변환에 의해 발생하는 왜곡을 기준으로 유사도를 측정하는 방법으로, 휘도와 대비, 그리고 구조를 이용하여 계산한다. 빛의 밝기를 의미하는 휘도(l)는 평균을 이용하여 나타내며, 상대적인 밝기로서 이미지의 명암을 나타내는 대비(c)는 표준편차를 이용하여 계산한다. 구조(s)는 픽셀들의 상대적 위치로서 평균과 표준편차를 이용하여 계산한다. 위 3가지 항목을 묶어서 SSIM을 산출하며 수식은 (식 2)와 같다.

(식 2) SSIM

$$SSIM(x, y) = l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma$$

$$= \left(\frac{2u_x u_y + c_1}{u_x^2 + u_y^2 + c_1} \right)^\alpha \cdot \left(\frac{2\sigma_x \sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \right)^\beta \cdot \left(\frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \right)^\gamma$$

$$= \frac{(2u_x u_y + c_1)(2\sigma_x \sigma_y + c_2)}{(u_x^2 + u_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

$$\left(\begin{array}{l} l : \text{luminance}, c : \text{contrast}, s : \text{structure} \\ u_x : \text{average}, \sigma_x : \text{standard deviation}, \\ \sigma_{xy} : \text{covariance}, \sigma_x^2 : \text{variance} \\ c_1 = (k1 \times L)^2, c_2 = (k2 \times L)^2, c_3 = \frac{c_2}{2} \\ L : (\text{size of the range of pixel}, 2^{\text{bits per pixel}} - 1) \\ \text{default} : k1 = 0.01, k2 = 0.03, \alpha = \beta = \gamma = 1 \end{array} \right)$$

2. Object Detection

객체 탐지의 대표적 모델인 YOLOv5[21]는 마치 사람처럼 어디에 무엇이 있는지를 한 번에 판단하는 ‘You

Only Look Once'를 구현한 것으로, 객체의 분류(Classification)와 위치 파악을 동시에 진행한다. 구체적으로 해당 모델은 ConvNet과 소프트맥스(Softmax)함수를 이용하여 객체 분류를 수행한다. 또한 객체의 위치(Object Localization)는 객체의 주변에 바운딩 박스를 표시하고 바운딩 박스의 좌표를 신경망 출력에 포함시켜서 나타낸다. 구체적으로 바운딩 박스는 [pc, bx, by, bw, bh, c1, c2, c3, c4, ..., cn]과 같은 형식으로 표현되는데, pc는 주어진 바운딩 박스에 객체가 존재할 확률, (bx, by, bw, bh)는 바운딩 박스의 위치 정보, 그리고 (c1, c2, c3, ...cn)은 확인된 객체가 각 클래스(Class)에 속할 확률을 나타낸다. YOLOv5 모델은 이미지를 $s \times s$ 의 그리드 셀(Grid Cell)로 나누는데, 하나의 객체가 다수의 그리드 셀에 동시에 걸쳐서 존재할 수 있기 때문에 각 객체당 여러 바운딩 박스가 대응될 수 있다. 하나의 객체에 대해 여러 개의 바운딩 박스가 대응되는 경우, 일반적으로 NMS(Non Maximum Suppression)[22] 알고리즘의 IOU(Intersection Over Union)에 기반을 두어 가장 높은 신뢰도를 갖는 바운딩 박스 하나만 선택한다. IOU는 바운딩 박스 a_1 과 a_2 의 교집합 면적을 합집합 면적으로 나눈 값으로 정의되며, 이 값을 a_n 까지 반복하면서 산출한다[23]. NMS 알고리즘은 이러한 IOU 값들 중 가장 큰 값을 갖는 바운딩 박스를 선택하여 최종적으로 하나의 객체에 대해 하나의 바운딩 박스 정보만 남긴다(Fig. 2).

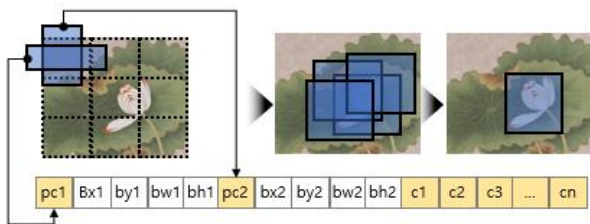


Fig. 2. An Example of Bounding Box Selection in YOLOv5

3. Image Labeling

이미지 레이블링은 학습 데이터를 기반으로 모델을 훈련시키는 딥러닝 모델 구축을 위해 필수로 선행되어야 하는 작업이다. 사람이 수작업으로 진행하는 레이블링은 많은 시간과 노력을 요구할 뿐 아니라, 담당자의 주관과 경험에 따라 일관적이지 않은 레이블링 기준이 적용되는 경우가 비일비재하다는 점에서 큰 한계를 나타내고 있다. 따라서 객관적인 기준에 따라 정확하게 이미지에 레이블을 부착하기 위한 연구가 다수 이루어지고 있으며, 가장 대표적인 개념으로 바운딩 박스(Bounding Box), 키 포인트

(Key Point), 폴리곤(Polygon), 그리고 폴리라인(Polyline) 등을 들 수 있다[18].

바운딩 박스는 이미지에 존재하는 객체의 위치 정보를 표시하기 위해 공백을 최소화하면서 특정 객체를 감싸는 형태로 드래그하는 기법으로, 이미지 레이블링 작업에서 가장 많이 사용되고 있다. <Fig. 3>의 예와 같이 바운딩 박스 레이블링을 진행하면 해당 이미지에 대한 클래스 번호, 바운딩 박스의 중심 좌표, 가로 길이, 세로 길이 등의 정보가 기록된다.

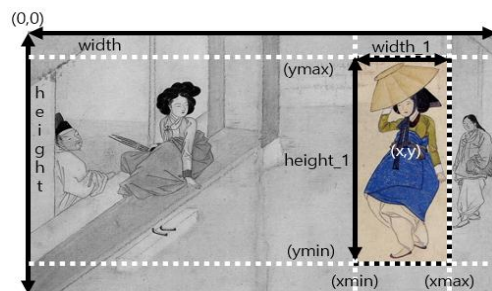


Fig. 3. Image Labeling(Bounding Box)

키 포인트는 객체의 특징점을 레이블링하는 방식으로 이미지 매칭, 골격추출 등에 널리 사용된다. 또한 폴리곤은 다각형 모양의 외곽선을 따라 점을 찍으면서 객체의 경계선을 정교하게 추출하는 방법이며, 폴리라인은 객체를 따라 여러 개의 점으로 연속선을 그리는 레이블링 기법이다(Fig. 4).

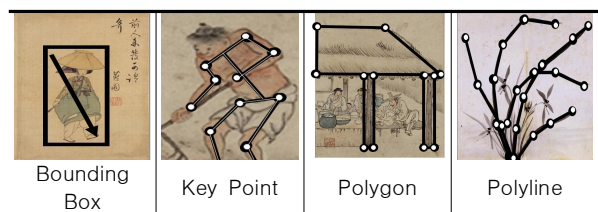


Fig. 4. Representative Image Labeling Techniques

이러한 다양한 이미지 레이블링 기법들은 구축하고자 하는 모델의 특성에 따라 적용되어야 한다. Lee et al.(2021)의 연구[18]에서는 객체 개수를 2개에서 10개까지 포함하는 10장의 이미지를 대상으로 실험을 진행하여 바운딩 박스와 폴리곤의 레이블링 시간을 비교하였다. 실험 결과 바운딩 박스와 폴리곤 레이블링에 소요된 시간은 각각 4.52분, 19.01분으로 나타나, 바운딩 박스보다 복잡한 폴리곤 레이블링에 약 4.1배의 시간이 소요되었음을 확인하였다.

특히 배경 정보가 복잡하거나 외곽 정보가 두드러지게

나타나지 않은 경우 이미지 레이블링에 더욱 많은 시간이 소요된다. Jeon and Rhee(2021)의 연구[24]에서는 212개의 채소 열매와 잎에 레이블을 지정하기 위하여 2주일간 하루 5시간씩 5명이 투입되었음을 확인하였으며, 또 다른 연구[25]에서는 약 100,000개의 의료 이미지로 구성된 데이터 셋에 인간 레이블러가 수작업으로 레이블을 지정하기 위해서는 약 24년이라는 기간이 필요하다고 예상하였다. 특히 구축하고자 하는 모델이 복잡한 경우에는 더욱 많은 시간과 노력이 필요하게 된다[26].

레이블링 이미지 데이터의 확보가 어려운 경우 딥러닝 기술 개발은 더욱 더디게 진행되기 때문에, 최근 다양한 관점에서 이미지 레이블링을 효율적으로 하는 방안이 지속적으로 연구되고 있다[27, 28, 29]. 이러한 기존 연구들은 대부분 최초 학습에는 레이블이 부착된 이미지를 사용하므로, 레이블이 부착된 이미지 셋이 전혀 없는 경우에는 이들 레이블링 기법을 적용할 수 없다. 따라서 기존의 이미지 레이블링 기법을 사용하더라도 초기 학습을 위해서는 수작업 레이블링이 반드시 필요하며, 방대한 이미지에 대해 레이블링을 수행할수록 모델의 정확도는 상승하지만, 레이블링 비용은 증가하게 된다.

이에 본 연구에서는 수작업 레이블링을 효율적으로 수행하기 위한 샘플링 방안, 즉 대량의 데이터로부터 소량의 샘플만 추출하여 이에 대해서만 레이블링을 수행하면서도 분석 모델의 성능을 크게 저하시키지 않는 방안을 제시하고자 한다.

III. Proposed Method

1. Overall Research Process

본 장에서는 초고해상화 알고리즘을 활용하여, 효율적인 객체 탐지를 위해 선택적 레이블링을 수행하는 방안을 제시한다. 본 방법론은 레이블링이 전혀 이루어지지 않은 이미지의 초기 레이블링을 효율적으로 수행하기 위해 제안된 것이므로, 사진 데이터에 비해 레이블링이 훨씬 덜 이루어진 명화 데이터를 분석 대상으로 선정하였다. 특히 오래된 동양화의 경우 해상도가 낮은 경우가 많다는 점에 주목하여, 동양화를 분석 대상으로 선정하였다.

해상도가 낮은 이미지에 대해 초고해상화를 수행하는 과정에서 SSIM과 PSNR 지표를 산출할 수 있다. 기존 연구에서는 SSIM과 PSNR을 주로 이미지 복원 알고리즘의 품질 평가 기준으로 사용하였다. 즉 원본 이미지가 완벽하다고 전제하고, 불완전한 이미지를 복원하여 얼마나 원

본 이미지에 가깝게 만들었는지 평가하기 위한 기준으로 SSIM과 PSNR을 사용하였다. 이와 달리 본 연구에서는 SSIM과 PSNR을 레이블링을 위한 최적 이미지 선정 기준으로 사용한다. 즉 지나치게 높은 SSIM과 PSNR 값은 원본 이미지의 해상도 개선이 거의 이루어지지 않았음을 암시하고, 지나치게 낮은 값은 초고해상화 과정에서 원본 이미지에 대한 왜곡이 발생했을 가능성이 존재함을 암시한다. 따라서 본 연구에서는 이러한 두 가지의 현상을 동시에 고려하여 가장 적절하게 초고해상화가 이루어진 지점을 찾고자 하며, 이를 위해 SSIM과 PSNR을 활용한다.

제안 방법론의 아이디어는 <Fig. 5>의 실제 예를 통해 설명할 수 있다. <Fig. 5>에서 좌측의 이미지 쌍은 SSIM이 매우 높은 경우를 나타낸다. 이 경우 원본과 초고해상화 이미지는 매우 유사함을 확인할 수 있으며, 이는 초고해상화를 통해 원본의 해상도가 크게 개선되지 않았음을 의미한다. 한편 가장 우측의 이미지 쌍은 SSIM 값이 매우 낮게 나타났지만, 초고해상화를 통해 원본 이미지의 해상도가 오히려 낮아진 상황을 보인다. 반면 가운데의 이미지 쌍은 중간 정도의 SSIM 값을 나타내는 변환으로, 초고해상화를 통해 원본 이미지의 해상도가 크게 개선되었음을 알 수 있다. 이렇듯 본 연구에서는 초고해상화를 통한 이미지 품질 개선의 정도를 SSIM 값의 변화를 통해 측정할 수 있을 것으로 예상된다. 구체적으로 본 연구에서는 품질 개선이 가장 크게 이루어진 SSIM 구간을 찾기 위해, SSIM의 구간별로 모델을 각각 구축하고 각 모델의 객체 인식 정확도를 비교하는 방식으로 고품질의 이미지 셋을 찾아내고자 한다.

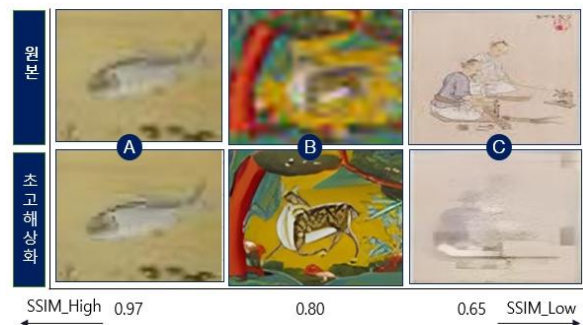


Fig. 5. SSIM of Some Image Pairs

제안 방법론에 따라 진행되는 이미지 셋 구성과 명칭은 <Fig. 6>과 같다. 방대한 이미지(All_img)가 가지고 있는 대표적인 특성을 파악하기 위하여 원본 이미지 일부(Ori_img)를 추출하여 초고해상화를 진행한다. 초고해상화된 이미지(Sr_img)를 사용하여 제안 방법론에 따라 모

델을 구축하고, 최적의 이미지를 선정하는 기준안을 마련한다. 파악된 기준안에 따라 Phase 2를 진행하게 된다.

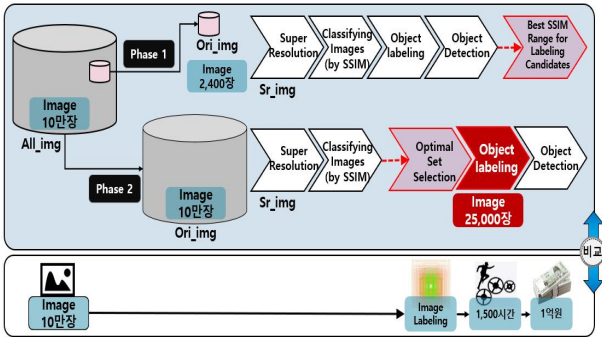


Fig. 6. Image Set Configuration

제안 방법론의 전체 개요는 <Fig. 7>과 같다. <Fig. 7>의 Phase 1은 선택적 레이블링을 위한 고품질 이미지 셋을 선정하는 기준을 설정하는 과정이고, Phase 2는 앞에

서 설정한 기준을 적용하여 새로 주어진 이미지 셋 (All_img)에 대해 선택적 레이블링을 수행하는 과정이다.

우선 Phase 1의 Process 1은 실험에 사용되는 원본 이미지(Ori_img)에 대해 Real-ESRGAN 알고리즘을 이용하여 초고해상화를 수행하고, 원본 이미지(Ori_img)와 새롭게 생성된 초고해상화 이미지(Sr_img)를 비교하여 SSIM과 PSNR을 산출하는 과정이다. Process 2에서는 초고해상화 이미지(Sr_img)를 SSIM 값에 따라 N개의 구간으로 구분하고, 각 구간에 속한 이미지들로 구간별 이미지 셋을 구성한다. <Fig. 7>에서 음영으로 구분된 Process 3 ~ 5의 과정은 이미지 셋의 수만큼 반복되는 과정이다. 우선 Process 3은 N개의 이미지 셋에서 하나의 셋을 반복적으로 선택하는 과정이며, 선택된 셋에 대해 객체의 레이블을 부여하는 작업은 Process 4에서 수행된다. Process 5는 레이블이 부여된 이미지 셋에 대해 YOLOv5 모델을 통해 객체 탐지를 수행하는 과정이다. 이러한 과정은 N번 반복

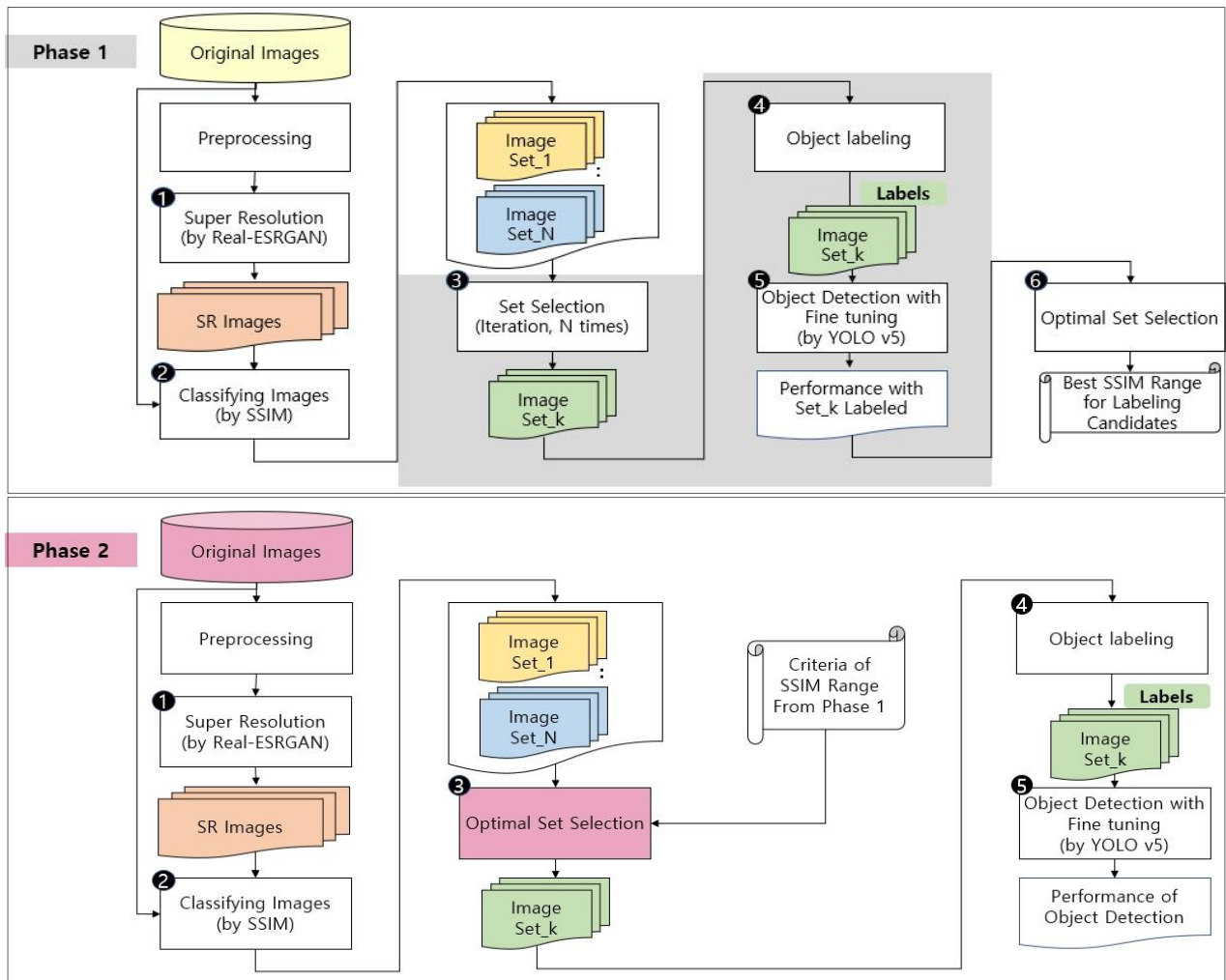


Fig. 7. Overall Research Process

되어 그 결과로 총 N개의 이미지 셋 각각에 대한 mAP가 측정되며, Process 6은 이러한 결과를 토대로 객체 인식의 성능이 가장 우수하게 나타나는 이미지 셋의 SSIM 구간, 즉 최적 SSIM 범위를 설정한다.

Phase 2는 앞 단계에서 설정된 최적 SSIM 범위를 사용하여 방대한 이미지(All_img)에 대해 선택적 레이블링을 수행하는 과정을 소개한다. Process 1과 2는 Phase 1과 동일한 방식으로 수행된다. Process 3에서는 N개의 이미지 셋 중 하나의 셋만을 레이블링 대상으로 선정하며, 이때 Phase 1의 결과인 최적 SSIM 범위를 참고한다. 그 결과로 단 하나의 셋에 대해서만 레이블링을 수행하게 되며, 해당 셋에 대해서만 이후 과정인 Process 4와 5를 수행한다.

본 장 이후 절에서는 제안 방법론의 구체적인 내용을 상세히 소개하며, 제안 방법론을 실제 이미지에 적용하여 실험을 수행한 결과는 다음 장인 4장에서 소개한다.

2. Super Resolution(by Real-ESRGAN)

본 절에서는 <Fig. 7>의 단계 중 Phase 1의 Process 1, 2에 해당하는 과정을 소개한다. 먼저 Process 1은 원본 이미지(Ori_img)에 대해 Real-ESRGAN 알고리즘을 이용하여 초고해상화를 수행한다. 다음으로 새롭게 생성된 초고해상화 이미지(Sr_img)와 원본 이미지(Ori_img)를 비교하여 SSIM과 PSNR을 산출할 수 있다. 이미지 특성별로 SSIM과 PSNR의 전체적인 분포는 다르게 나타날 수 있다. 본 연구에서는 파일럿 실험 결과 SSIM에 의한 구간별 특징에서는 SSIM값의 변화에 따라 모델의 정확도가 규칙적으로 변화하는 현상을 발견하였지만, PSNR에 의해 분류된 구간에서는 PSNR의 변화에 따른 모델 정확도의 변화가 일관적으로 나타나지 않았다. 따라서 구간을 나누는 평가지표로 SSIM을 선택하여 이후 실험을 진행하였다. 원본 이미지와 변환 이미지의 차이가 큰 경우 SSIM은 작게 나타나며, 초고해상화가 잘 이루어지지 않아서 원본 이미지와 변환 이미지의 차이가 거의 없는 경우 SSIM은 매우 크게 나타나게 된다.

Process 2는 산출된 SSIM을 기준으로 N개의 구간으로 이미지 셋을 구성하는 과정이다. 즉 초고해상화 이미지(Sr_img)를 SSIM에 따라 정렬하여 2개의 구간, 4개의 구간 또는 10개의 구간 등으로 나눌 수 있다. 예를 들어 SSIM을 기준으로 전체 이미지를 4개의 구간으로 구분하여 4개의 이미지 셋을 구성할 수 있다. 이후 절에서 소개할 객체 탐지 모델의 학습에는 레이블이 부착된 데이터만 사용되는 반면, Phase 1의 선택적 레이블링 단계에서는

레이블이 부착되지 않은 데이터(Unlabeled Data)도 함께 사용된다.

3. Object Detection (by YOLOv5)

본 절에서는 <Fig. 7>의 단계 중 Phase 1의 Process 3~5에 해당하는 과정을 소개한다. 먼저 Process 3은 N개의 이미지 셋에서 하나의 셋을 선택한다. 다음으로 Process 4는 선택된 하나의 셋에 대해 수동적인 이미지 레이블링을 진행하며, Process 5는 생성된 레이블과 이미지 쌍을 이용하여 객체 탐지 모델을 구축한다. 객체 탐지에는 수 만장의 이미지와 80개 클래스에 대해 이미 충분히 사전학습이 이루어진 모델인 YOLOv5를 사용한다. 이러한 사전학습 모델을 기반으로 미세 조정을 통해 동양화에 적합하게 하이퍼 파라미터를 수정하는 방식으로 새로운 객체 탐지 모델을 구축하였다. <Fig. 8>은 동양화에 대해 객체 탐지를 진행한 결과의 일부 예시이다.



Fig. 8. Object Detection

Process 3~5는 이미지 셋의 수만큼 N번 반복되는 과정으로, 총 N개의 이미지 셋 각각에 대해 객체 탐지 모델을 구축한다. 이후 과정에서 구축된 각각의 객체 탐지 모델에 대해 mAP를 이용하여 정확도를 비교한다.

4. Optimal Set Selection(by SSIM)

본 절에서는 <Fig. 7>의 단계 중 Phase 1의 Process 6과 Phase 2에 해당하는 과정을 소개한다. Phase 1의 Process 6은 N개의 이미지 셋 각각에 대한 mAP를 비교하여 객체 탐지의 성능이 가장 우수하게 나타나는 SSIM 구간의 최적 이미지 셋을 선정하는 과정이다.

Phase 2는 방대한 이미지(All_img)를 대상으로 진행하는 과정이다. 먼저 Process 1과 2는 Phase 1과 동일한 방식으로 초고해상화를 진행하고, 그 결과 산출된 SSIM을 기준으로 정렬하여 구간별 이미지 셋을 구분한다. Process 3은 Phase 1에서 선정된 최적의 SSIM 범위를 참고하여 하나의 이미지 셋을 선별하는 과정이다. Process 4는 선별된 최적의 이미지 셋에 대해서만 레이블링을 수행하는 과정이며 Process 5는 해당 이미지 셋을 학습 데이터로 사용하여 객체 탐지 모델을 구축한다.

IV. Experiment

1. Experiment Overview

본 절에서는 3장에서 소개한 제안 방법론을 실제 데이터에 적용한 실험 결과를 소개한다. 실험에는 동양화 이미지 2,400장을 수집하고, “carp”, “cranebird”, “deer”, “lotusflower”, “turtle”, 그리고 “person”의 6가지 객체에 대한 레이블링을 직접 수행하였다. 각 클래스별 이미지의 수는 400장으로 통일하였다. 실험은 Python 프로그램을 사용하여 GPU:Tesla V100, CPU:Xeon (R) 5120 2.20GHz, Memory:180 GB의 환경에서 진행하였다. 초고해상화는 Real-ESRGAN 알고리즘을 이용하여 3.5배 Upsampling으로 진행하였으며 Residual Block 16개, RRDB블록 23개를 지정하였다. 초고해상화 이미지와 원본 이미지(Ori_img)를 비교하여 SSIM과 PSNR을 산출하였다.

2. Results of Super Resolution

본 실험에 앞서 초고해상화가 객체 탐지의 정확도에 미치는 영향을 확인하기 위해 원본(Ori_img) 이미지와 초고해상화(Sr_img) 이미지에 대해 각각 객체 탐지 모델을 구축하였다. 객체 탐지 모델은 YOLOv5 모델을 사용 구축하였고, Batch Size는 32, Epoch은 100, Img_size는 416으로 지정하여 동양화 이미지로 미세 조정을 진행하였다. 사용된 이미지는 전체 2,400장의 이미지 중 Train 1,680장, Validation 360장이다. Test는 360장의 이미지를 사용하여 실험을 진행하였으며, 이는 이후 과정에서도 동일하다.

객체 탐지 모델의 객관적 성능은 예측 결과가 얼마나 정확한지로 평가하며, 이를 위해 정밀도(Precision)와 재현율(Recall)을 이용한 mAP(mean Average Precision) 척도가 주로 사용된다. 정밀도는 예측된 객체 중 실제 정답의 비율이고 재현율은 검출되어야 할 객체 중 제대로 검출된 것의 비율을 나타낸다(Fig. 9). 또한 AP(Average Precision)는 정밀도와 재현율을 나타낸 그래프의 면적을 구한 값이고, 각 클래스별 AP의 평균을 구한 값이 mAP이다.

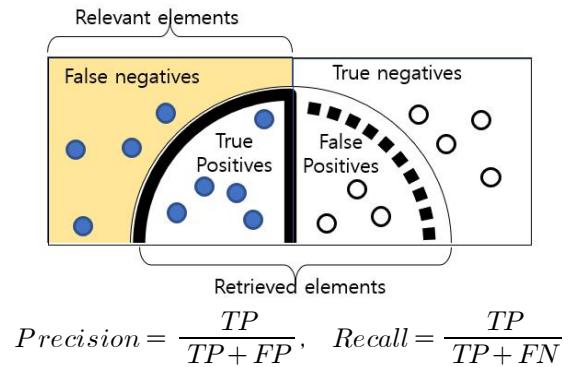


Fig. 9. Precision and Recall

원본 이미지(Ori_img)와 초고해상화(Sr_img) 이미지로 각각 객체 탐지모델을 별도로 구축하여 mAP를 비교한 실험 결과는 <Table 1>과 같으며, 실험 결과 초고해상화된 이미지를 사용하여 구축한 객체 탐지 모델의 정확도가 원본 이미지를 사용한 모델에 비해 높게 나타남을 확인하였다.

Table 1. mAP(Ori_img vs. Sr_img)

Image Set	mAP @.5:.95
Ori_img(a0)	0.6310
Sr_img(a1)	★ 0.6470

3. Results of Optimal Image Set Selection

초고해상화된 동양화에서 최적의 이미지 셋을 추출하기 위해 테스트 이미지 360장을 제외한 2,040장에 대한 SSIM 값을 산출하였으며, SSIM에 따라 이미지를 정렬하고 이를 2개의 구간, 4개의 구간, 10개의 구간으로 구분하여 각각의 이미지 셋을 별도로 구성하였다. 이미지 셋은 a_j^i 로 나타내며 j는 구간의 수, i는 이미지 셋의 순서를 나타낸다. 즉, a_2^1 는 SSIM에 따라 2개의 구간으로 나눈 이미지 셋 중 첫 번째 셋을 의미한다. 이미지 셋을 2개의 구간으로 구분한 경우, 각 구간에 속한 이미지 1,020장씩에 대한 SSIM 값의 최대, 최소, 평균값은 <Table 2>와 같다.

Table 2. SSIM of Each Set (Divided into 2 Sets)

Image Set	Max	Min	Average
a_2^2	0.9890	0.9248	0.9475
a_2^1	0.9246	0.2729	0.8745

이미지 셋을 4개의 구간으로 구분한 경우와 10개의 구간으로 구분한 경우에 대한 SSIM 값의 최대, 최소, 평균 값은 각각 <Table 3>, <Table 4>와 같다.

Table 3. SSIM of Each Set (Divided into 4 Sets)

Image Set	Max	Min	Average
a_4^4	0.9890	0.9459	0.9592
a_4^3	0.9459	0.9248	0.9357
a_4^2	0.9246	0.8946	0.9105
a_4^1	0.8945	0.2729	0.8383

Table 4. SSIM of Each Set (Divided into 10 Sets)

Image Set	Max	Min	Average
a_{10}^{10}	0.9890	0.9602	0.9689
a_{10}^9	0.9600	0.9505	0.9550
a_{10}^8	0.9504	0.9419	0.9461
a_{10}^7	0.9418	0.9342	0.9381
a_{10}^6	0.9341	0.9250	0.9294
a_{10}^5	0.9249	0.9141	0.9196
a_{10}^4	0.9141	0.9012	0.9080
a_{10}^3	0.9012	0.8855	0.8941
a_{10}^2	0.8855	0.8550	0.8723
a_{10}^1	0.8541	0.2729	0.7786

이상의 구간별 이미지 셋 각각에 대해 별도의 객체 탐지 모델을 구축하고, 이들 모델의 성능을 mAP를 사용하여 비교하였다. 객체 탐지에는 저해상도 원본 데이터가 아닌 초고해상화를 거친 이미지를 사용하였다. 이후의 모든 표에서 제시되는 mAP는 Test 데이터로부터 산출된 값이다.

첫 번째 실험은 이미지 셋을 SSIM 값에 따라 2개의 구간으로 나누어 수행하였으며, 실험 결과 mAP는 $a_2^2(0.550)$, $a_2^1(0.585)$ 으로 a_2^1 셋이 더 높게 나타났다(Table 5).

Table 5. mAP of Each Set (Divided into 2 Sets)

Image Set	mAP @.5:.95
a_2^2	0.550
a_2^1	★ 0.585

4개의 구간에 대한 실험에서는 mAP가 $a_4^4(0.486)$, $a_4^3(0.517)$, $a_4^2(0.533)$, $a_4^1(0.510)$ 으로 a_4^2 가 가장 높게 나타

났으며(Table 6), 10개의 구간에 대한 실험에서는 a_{10}^4 의 mAP가 가장 높게 나타났다(Table 7). 이러한 방식으로 2개, 4개, 그리고 10개로 구분한 이미지 셋에 대해 가장 높은 객체 탐지 정확도를 갖는 SSIM 값의 구간을 식별할 수 있다.

Table 6. mAP of Each Set (Divided into 4 Sets)

Image Set	mAP @.5:.95	Image Set	mAP @.5:.95
a_4^4	0.486	a_4^2	★ 0.533
a_4^3	0.517	a_4^1	0.510

Table 7. mAP of Each Set (Divided into 10 Sets)

Image Set	mAP @.5:.95	Image Set	mAP @.5:.95
a_{10}^{10}	0.398	a_{10}^5	0.434
a_{10}^9	0.387	a_{10}^4	★ 0.477
a_{10}^8	0.403	a_{10}^3	0.449
a_{10}^7	0.385	a_{10}^2	0.391
a_{10}^6	0.427	a_{10}^1	0.365

4. Results of Selective Labeling

본 절에서는 방대한 이미지로부터 무작위 추출을 통해 레이블링을 위한 샘플을 구성한 경우와, 제안 방법론에 따라 레이블링을 위한 최적 샘플을 구성한 경우의 객체 탐지 모델 정확도를 비교한 결과를 소개한다. 제안 방법론에 따라 선택된 최적 이미지 셋과 Train, Validation, Test 데이터의 개수는 <Table 8>과 같으며, 동일한 개수의 무작위 샘플링 이미지 셋을 구성하여 비교 실험을 수행하였다. <Table 8>에서 Data Size는 전체 데이터에 대한 샘플 데이터의 비율을 나타낸다. 즉 100%는 전체 데이터에 대한 레이블링 수행한 경우이며, 10%는 전체 데이터 중 10%에 대해서만 레이블링을 수행한 경우이다.

Table 8. Optimal Image Set Configuration

Data Size	Optimal Image Set	Number of Data		
		Train	Val.	Test
100%	(a_1^1)	1,680	360	360
50%	(a_2^1)	816	204	
25%	(a_4^2)	408	102	
10%	(a_{10}^4)	164	40	

<Table 8>의 데이터에 대해 객체 탐지 실험을 수행한 결과는 <Table 9>와 같다.

Table 9. Performance Comparison (mAP)

Data	mAP @ 5:.95		
	Random sampling(a)	Optimal Image Set(b)	(b)-(a)
100%	0.647		0
50%	0.564	0.585	0.021
25%	0.507	0.533	0.026
10%	0.396	0.477	0.081

<Table 9>에서 제안 방법론에 따라 구축한 객체 탐지 모델의 mAP가 무작위 추출 방법론에 의한 객체 탐지 모델에 비해 모든 Data Size에 대해 높게 나타났다. 정확도의 차이는 소량의 샘플을 사용한 경우 더욱 크게 나타나며, 이러한 추세는 <Table 9>를 그래프로 나타낸 <Fig. 10>에서 더욱 명확하게 확인할 수 있다.

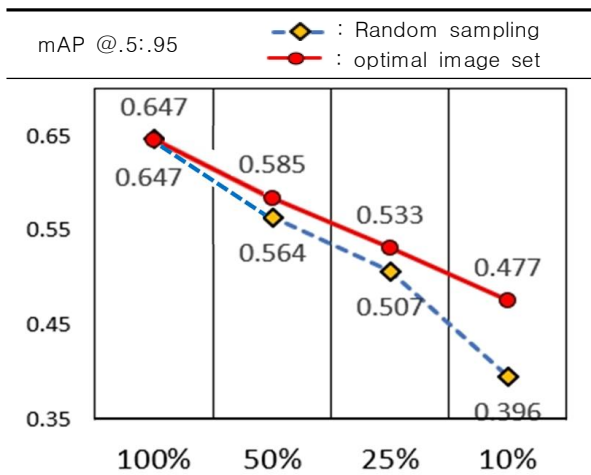


Fig. 10. Performance Comparison (mAP)

V. Conclusion

방대한 양의 이미지에 사람이 직접 레이블링을 진행하는 것은 매우 많은 시간과 노력을 필요로 한다. 이러한 어려움을 극복하기 위해 본 연구에서는 방대한 이미지가 아닌 일부 이미지에 대한 레이블링을 통해서도 큰 성능의 저하 없이 객체 탐지를 수행하는 방안을 제안하였다. 또한 직접 구축한 동양화 이미지에 대해 실험을 수행한 결과, 제안 방법론에 따른 이미지 샘플링 방법이 무작위 샘플링에 비해 객체 탐지 모델 구축에 적합한 레이블링 샘플을 추출함을 확인하였다.

본 연구의 기여는 다음과 같다. 우선 본 연구에서는 초고해상도 과정에서 산출되는 SSIM 값과 객체 인식의 정확도를 측정하는 척도인 mAP와의 관계를 분석하고, 이를 활용하여 객체 인식 모델 구축에 가장 큰 도움을 줄 수 있

는 이미지 셋을 식별하였다. 향후 다양한 분석을 통해 SSIM을 활용하여 mAP를 향상시키기 위한 연구가 다수 수행될 수 있을 것으로 기대하며, 이는 본 연구의 학술적 기여로 인정받을 수 있다. 또한 이미지의 양이 너무 방대해서 레이블링이 불가능한 여러 분야의 데이터에 대해 제안 방법론을 즉시 적용하여 효율적인 레이블링을 수행할 수 있으며, 이는 본 연구의 실무적 기여가 될 수 있다.

본 연구의 한계는 다음과 같다. 우선 본 연구에서는 동양화 이미지 2,400장에 대한 실험을 진행하였는데, 이는 방법론의 성능을 안정적으로 검증하기에 충분한 데이터라고 하기 어렵다. 향후 데이터 증강 등의 기법을 활용하고 더욱 다양하고 방대한 데이터를 수집한 후 추가 실험을 수행함으로써, 제안 방법론의 견고함을 확인할 필요가 있다. 특히 동양화와는 상이한 특징을 갖는 다양한 이미지에 대해서도 동일한 실험을 반복 수행하여, 제안 방법론의 성능을 더욱 엄밀하게 평가할 필요가 있다. 또한 향후 연구에서는 SSIM 값이 지나치게 높거나 낮게 나타나는 이미지의 특성을 보다 면밀히 고찰하여, 이러한 특징이 객체 탐지 모델 학습에 영향을 미치는 메커니즘을 규명할 필요가 있다.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5A2A01061459).

This work was supported by the 'High Performance Computing Support' project of the Ministry of Science and ICT of the Republic of Korea and the National IT Industry Promotion Agency.

REFERENCES

- [1] J. Yu, Y. Han, J. Kim, and H. Hahn, "Ensemble Deep Network for Dense Vehicle Detection in Large Image," The Korean Society Of Computer And Information, Vol. 26, No. 1, pp. 45-55, January 2021.
- [2] J. Song, S. Lee, and S. Park, "A Study on the Industrial Application of Image Recognition Technology," Journal of the korea contents association, Vol. 20, No. 7, pp. 86-96, July 2020.
- [3] Y Li, C Ma, T. Zhang, J. Li, Z. Ge, Y. Li, and S Serikawa, "Underwater Image High Definition Display Using the Multilayer

- Perceptron and Color Feature-Based SRCNN,” IEEE Access, pp. 83721-83728, June 2019. DOI: 10.1109/ACCESS.2019.2925209
- [4] Y. Lee and H. Park, “A Study of Lightening SRGAN Using Knowledge Distillation,” Journal of Korea Multimedia Society, Vol. 24, No. 12, pp. 1598-1605, December 2021.
- [5] X. Wang, K. Yu1, and S. Wu, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks,” arXiv:1809.00219 v2, September 2018.
- [6] X. Wang and L. Xie, C. Dong, and Y. Shan, “Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data,” International Conference on Computer Vision Workshops IEEE/CVF, July 2021.
- [7] C. Han, H. Hayashi, L. Rundo, R. Araki, “GAN-based synthetic brain MR image generation,” 2018 IEEE 15th International Symposium on Biomedical Imaging, May 2018. DOI: 10.1109/ISBI.2018.8363678
- [8] J. Lee, “PSNR Analysis of Ultrasound Images for Follow-up of Hepatocellular Carcinoma,” Journal of the Korean Society of Radiology, Vol. 9, No. 5, pp. 263-267, August 2015.
- [9] U. Sara, M. Akter, and MS. Uddin, “Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study,” Journal of Computer and Communications, Vol.7 No.3, March 2019.
- [10] Q. Junlong, J. Shin, and J. Ko, “A Study on Energy Consumption Prediction from Building Energy Management System Data with Missing Values Using SSIM and VLSW Algorithms,” The transactions of The Korean Institute of Electrical Engineers, Vol. 70, No. 10, pp. 1540-1547, October 2021.
- [11] M. Laavanya and V. Vijayaraghavan, “Residual Learning of Transfer-learned AlexNet for Image Denoising,” IEIE Transactions on Smart Processing and Computing, Vol. 9, No. 2, April 2020. DOI : 10.5573/IEIESPC.2020.9.2.135
- [12] S. Jeon, D. Kim, and H. Jung, “YOLO-based lane detection system,” Journal of the Korea Institute of Information and Communication Engineering, Vol. 25, No. 3, pp. 464-470, March 2021.
- [13] W. Kim, F. Dehghan, and S. Ch, “Vehicle License Plate Recognition System using SSD-Mobilenet and ResNet for Mobile Device,” Smart media journal, Vol. 9, No. 2, pp. 92-98, September 2020.
- [14] K. Shin, J. Lee, and J. Kim, “Thermal Image Processing and Synthesis Technique Using Faster-RCNN,” Journal of Convergence for Information Technology, Vol. 11, No. 12 , pp. 30-38, December 2021.
- [15] S. Hong, D. Kim, and B. Kim, “Image Labeling Technology Analysis and Training Set Generation Model for Detecting Damage and Cracks in Road Pavement,” Journal of Korean Society for Geospatial Information Science, Vol. 28, No. 4, pp. 119-125, December 2020.
- [16] H. Gu, J. Seo, and S. Choo, “A Development of Façde Dataset Construction Technology Using Deep Learning-based Automatic Image Labeling,” Architectural Institute of Korea, Vol. 35, No. 12, pp. 43-53, December 2019.
- [17] K. Choi, “Conceptual Design of Data Labeling System based on Metaverse for the Military Artificial Intelligence using the Modified Functional Analysis Method,” Korean Journal of Military Art and Science, Vol. 78, No. 1, pp. 375-390, February 2022.
- [18] R. Lee, R. Jang, and M. Park, “An Auto-Labeling based Smart Image Annotation System,” Journal of the Korea contents association, Vol. 21, No. 6, pp. 701-715, June, 2021.
- [19] J. Yoon, T. Kim, and Y. Choe, “GAN based Single Image Super-Resolution via Spatially Adaptive De-normalization,” The Transactions of the Korean, Vol. 70, No. 2, pp. 402-407, February 2021.
- [20] C. Lee, S. Youn, and C. Cho, “License Plate Image Enhancement Based on Enhanced Super-resolution Generative Adversarial Networks,” Journal of Next-generation Convergence Technology Association, Vol. 6, No. 1, pp. 5-11, January 2022.
- [21] J. Kang and J. Gwak, “Adaptive Face Mask Detection System based on Scene Complexity Analysis,” Journal of The Korea Society of Computer and Information, Vol. 26, No. 5, pp. 1-8, May 2021.
- [22] H. Lim, “Overview of Image-based Object Recognition AI technology for Autonomous Vehicles,” Journal of the Korea Institute of Information and Communication Engineering, Vol. 25, No. 8, pp. 1117-1123, August 2021.
- [23] H. Kim and D. Choi, “A Perimeter-Based IoU Loss for Efficient Bounding Box Regression in Object Detection,” Journal of KIISE, Vol. 48, No. 8, pp. 913-919, August 2021.
- [24] W. Jeon and S. Rhee, “An Annotation Method of Vegetable Fruits and Leaves using a Depth Map,” Journal of Korean Institute of Intelligent Systems, Vol. 31, No. 6, pp. 465-474, December 2021.
- [25] Aitimes, “Don't worry about running out of data,” <https://www.aitimes.com/news/articleView.html?idxno=143907>, April 2022.
- [26] Y. Cho and S. Kim, “Labeling of Facility Crops Using Instance Segmentation of Deep Learning,” Journal of Korean Institute of Intelligent Systems, Vol. 31, No. 4, pp. 305-310, August 2021.
- [27] T. Mensink, J. Verbeek, and G. Csurka, “Learning structured prediction models for interactive image labeling,” CVPR, August 2011.
- [28] LC. Chen, S. Fidler, and AL. Yuille, “Beat the MTurkers: Automatic Image Labeling from Weak 3D Supervision,” CVPR, pp. 3198-3205, 2014.
- [29] A. Diaz-Pinto, S. Alle, A. Ihsani, and M. Asad, “Monai label: A framework for ai-assisted interactive labeling of 3d medical images,” arXiv:2203.12362, March 2022.

Authors



Hyeyoung Moon received the M.S. degree in Information Computer Education from Graduate School of Education, Konkuk University in 2010 and Ph.D. degree in Engineering from Graduate School of

Business IT, Kookmin University in 2022. She is interested in text mining, deep learning, and data modeling.



Namgyu Kim received the B.S. in Computer Engineering from Seoul National University in 1998, M.S. and Ph.D. degrees in Management Engineering from KAIST, Korea, in 2000 and 2007, respectively.

Dr. Kim joined the faculty of the School of Management Information Systems at Kookmin University, Seoul, Korea, in 2007. He is currently a professor in the Graduate School of Business IT at Kookmin University. He is interested in text mining, deep learning, and data modeling.