

회화문화재 객체검출을 위한 학습용 이미지 데이터셋 구축 방안 연구

권도형*, 유정민**
 한국전통문화대학교 문화유산산업학과
 {20203022, jmyu}@nuch.ac.kr

A Study on the Construction of Image Datasets for Object Detection of Painting Cultural Heritage

Do-Hyung Kwon*, Jeong-Min Yu**
 Dept. of Cultural Heritage Industry, Korea National University of Cultural Heritage

요 약

본 연구는 회화문화재 속에 표현된 다양한 종류의 객체를 검출할 수 있는 딥러닝 모델생성을 위해 필요한 학습용 이미지 데이터셋 구축방안을 제안한다. 먼저 기존 동양화 기반의 회화문화재 이미지 데이터 및 객체 특징 분석을 진행하였고, 이를 바탕으로 Natural image에 Pose transfer 및 Style transfer를 적용한 새로운 방식의 회화문화재 이미지 데이터 생성 방법을 제안한다. 제안한 프레임워크를 통해 기존 문화재 분야에서 가지고 있던 제한된 데이터 구축문제를 극복하고, 검출모델 생성을 위한 대용량의 학습데이터 구축 가능성을 제시하였다.

1. 서론

딥러닝 기반 객체검출분야에서 대용량의 학습데이터 구축은 강건한 성능을 지닌 모델생성을 위해 필수적이다. 특히 회화작품 속에 묘사된 객체를 탐지하는 교차묘사문제에서는 동일한 객체 클래스일지라도 여러 작품에 걸쳐 상이한 화풍과 형태로 묘사되어 있으므로, 다양한 표현 범위를 아우르는 데이터 확보가 중요하다.

그러나 문화재 분야의 특성상 회화작품의 수가 제한되어 있고, 일반인들의 접근성 문제로 인해 데이터 구축과 공유의 한계점이 발생하여 이와 관련된 연구가 미비한 상황이다. 국외의 경우에는 회화작품 속에 표현된 객체를 탐지하기 위한 데이터셋[1][2]이 존재하지만 이는 동양화 기반의 회화문화재에 적용시키기에는 한계가 있다.

위와 같은 문제점을 해결하기 위해 본 연구에서는 한국 및 동양의 회화문화재 속에 표현된 객체를 검출하기에 적합한 데이터셋 구축 방안을 제안한다. 2장에서는 기존 회화문화재 이미지 데이터의 특성을 살펴보고, 3장에서는 객체의 Pose 및 Style 변환 기반 이미지 데이터 구축 기법에 대해 설명하며, 마지

막으로 4장에서는 제안하는 방법으로 데이터를 구축한 뒤의 예상결과를 제시한다.

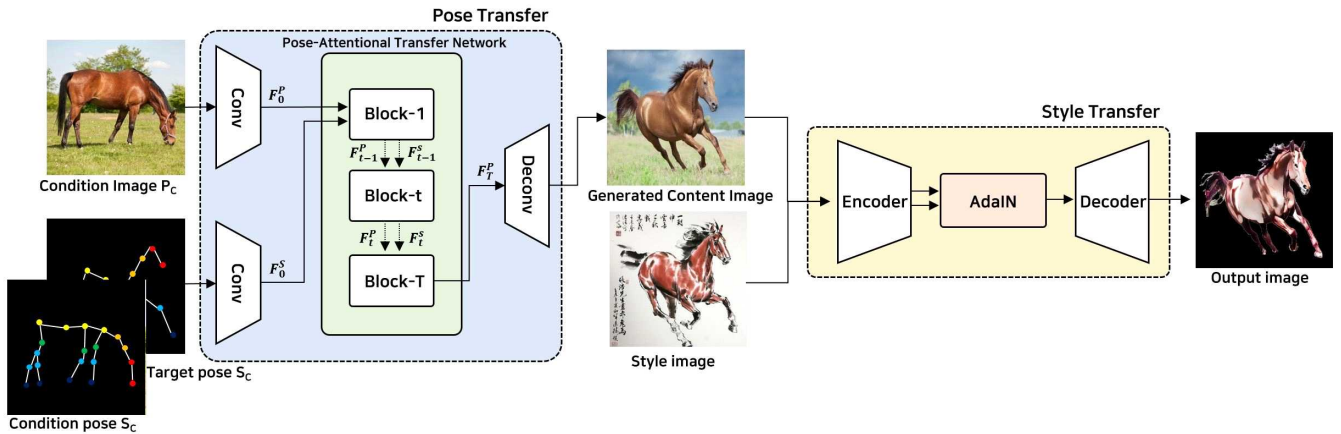
2. 회화문화재 이미지 데이터 및 객체표현 특성



일부분 훼손 식물과 상이한 형태 객체와 배경 간 경계 모호

(그림 1) 회화문화재 이미지 데이터 특성.

회화문화재에 묘사되는 객체의 범주는 크게 인물, 식물, 동물, 상상의 동물, 새/나비/곤충 등으로 구분할 수 있다. 이 중에서 특히 동물의 표현은 다른 객체 범주들과 달리 객체의 형태적 왜곡이 도드라지며, 색상과 무늬 등의 텍스처 표현도 다양하게 나타



(그림 2) 제안하는 회화문화재 이미지 데이터 구축 프레임워크.

난다. 또한 동양화는 서양화와 달리 객체가 표현된 전경과 배경의 경계가 모호한 경우가 많으며, 오랜 세월로 인해 일부분이 훼손되거나 전체적으로 부식이 일어나 형상이 흐릿하게 표현된 이미지가 다수 존재한다.

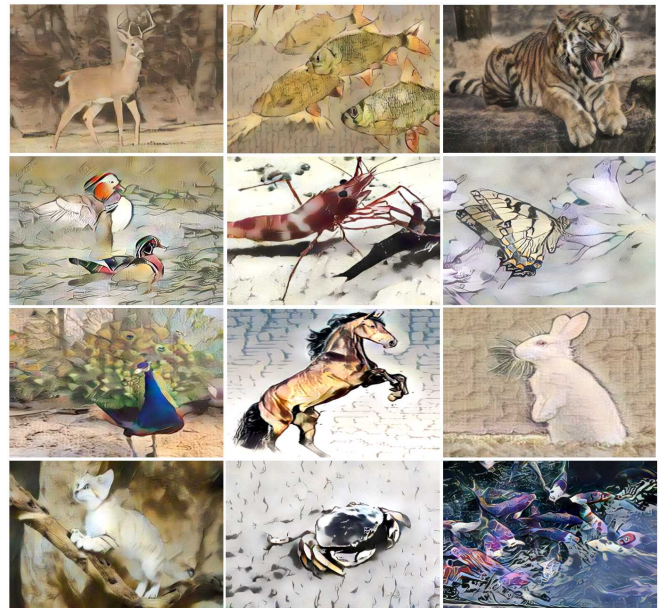
이러한 특성을 지닌 회화문화재 속 객체를 탐지하기 위해 Natural image 기반의 데이터셋을 사용하여 모델 구축을 진행할 수도 있으나, 객체 간의 Low-level 및 High-level feature variation으로 인해 강건한 검출 성능을 나타내기 어렵다. 따라서 회화문화재의 특성을 반영한 새로운 종류의 이미지 데이터셋 확보가 필수적이며, 풍부한 양의 이미지 데이터 구축을 위한 Image augmentation 및 generation 기법이 필요하다.

3. 제안하는 회화문화재 이미지 데이터 구축 기법

본 연구에서는 회화문화재 속 객체 검출 분야의 학습 데이터 불충분 문제를 해결하기 위해 Natural image에 Pose transfer[3]와 Style transfer[4]를 적용한 데이터 구축기법을 제안한다. 제안하는 프레임워크는 최근 GAN기반 Creative한 artwork 생성 연구의 접근방법들[5]과 달리 객체의 특정 자세와 화풍을 가진 이미지 생성을 목표로 한다.

첫 번째로 Pose transfer에서는 Pose-Attentional Transfer Blocks(PATBs)을 기반으로 Public dataset 및 web crawling으로 구축한 동물 이미지를 회화 문화재 속 동물들이 취하고 있는 몇 가지의 특정 target pose로 점진적으로 변형시킨다. 이미지와 함께 Input으로 입력되는 Condition 및 Target pose로는 동물의 Pose estimation에 주로 사용되는 관절 위치 Key point heatmap이며, 이를 위해서는

Animal pose estimator 구축이 우선적으로 필요하다. 다음으로 Style transfer에서는 Adaptive Instance Normalization(AdaIN)을 기반으로 특정한 Target pose로 변형되어 출력된 동물 이미지(Generated content image)를 회화문화재의 표현재료, 기법, 소재 등에 적절하게 화풍 변환을 실시하여 최종적인 이미지 데이터셋을 구축한다.



(그림 3) 구축된 회화문화재 이미지데이터 예상결과

제안한 프레임워크에 따라 생성된 데이터셋이 실제 활용에 적합한지 평가하기 위한 방법으로는 Fréchet inception distance(FID)[6]가 있다. FID를 구하는 공식은 아래와 같다.

$$d^2((m, C), (m_w, C_w)) = \|m - m_w\|_2^2 + \text{Tr}(C + C_w - 2(CC_w)^{\frac{1}{2}})$$

FID는 GAN을 통해 생성된 이미지의 품질을 평

가하기 위한 방법으로 생성된 이미지와 실제 이미지 간의 특징 벡터 간 거리 계산을 통해 유사성을 측정하며, FID score가 낮을수록 생성된 이미지의 품질이 좋은 것을 의미한다.

위의 방법을 통해 동물의 Natural image와 회화 문화재 속 객체 간 Feature 차이를 감소시킬 수 있는 데이터셋을 구축하여 Cross domain 문제를 원활하게 해결 할 수 있다.

4. 결론

그동안 회화문화재를 대상으로 진행된 객체 검출 연구는 일반인의 제한된 접근성과 한정된 작품 수 등의 문제로 학습 데이터셋 구축에 한계점이 존재하였다. 본 연구에서는 이를 극복하기 위해 기존 회화 문화재에 표현된 객체의 특성을 반영하여 이미지를 생성하는 프레임워크를 제안하여, 새로운 방법의 이미지 데이터 구축 가능성을 제시하였다. 향후 연구로는 본문에서 제안한 프레임워크를 기반으로 대규모의 이미지 데이터셋을 구축한 뒤 강건한 성능의 검출 모델생성 연구를 진행할 계획이다.

Acknowledgements

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2021 년도 문화기술연구개발 지원 사업으로 수행되었음. [R2020040045, 실감형 문화유산 체험을 위한 애셋 기반 지능형 큐레이션 및 서비스 운영 기술 개발]

참고문헌

- [1] Westlake, N., Cai, H., Hall, P. "Detecting people in artwork with cnns", European Conference on Computer Vision(ECCV), pp 825-841, 2016
- [2] Wilber, MJ., Fang, C., Jimm H., Hertzmann, A., Collomosse, J., Belongie, S. "BAM! The Behance Artistic Media Dataset for Recognition Beyond Photography", IEEE International Conference on Computer Vision (ICCV), pp 1211-1220, 2017
- [3] Zhu, Z., Huang, T., Shi, B., Yu, M., Wang, B., Bai, X. "Progressive Pose Attention Transfer for Person Image Generation", IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp 2342-2351, 2019
- [4] Huang, X., Belongie, S. "Arbitrary Style

Transfer in Real-time with Adaptive Instance Normalization", IEEE International Conference on Computer Vision (ICCV), pp 1510-1519, 2017

[5] Xue, A. "End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks", IEEE Winter Conference on Applications of Computer Vision (WACV), pp 3862-3870, 2021

[6] Huesel, M., Ramsauer, H., Unterthiner, T., Nessler, B. "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium", Proceedings of the 31st International Conference on Neural Information Processing Systems(NIPS), pp 6629-6640, 2017