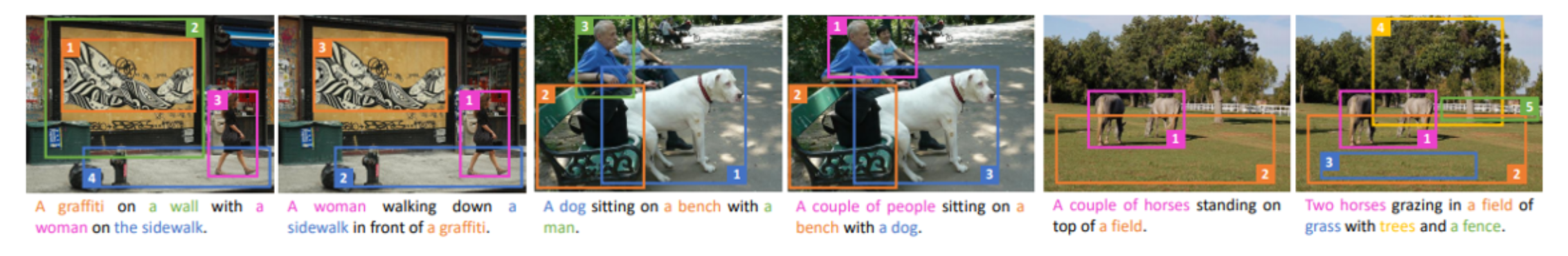
**시각장애인을 위한 그림 설명 도슨트 네트웍 개발 (이동명, 황혜경)**

**문헌 조사 :** [*Show, Control and Tell: A Framework for Generating Controllable and Grounded Captions*](https://arxiv.org/pdf/1811.10652.pdf)

요약

인공지능을 활용하여 사진(photo)이 아닌 그림(visual art of works)에 대한 객관적 사실(설명)을 제공하는 접근은 없다. **이에 그림에 대한** **설명을 다채롭게 제공**하는 알고리즘 제시 및 인공신경망을 구축하여, 전시 관람에서 시각장애인들에게 **더욱 깊이 있는 그림 감상**이 가능하도록 한다.

문제점



특정 사진에 대한 설명을 제시하는 인공신경망은 여럿 존재하나, 일반적인 사진에 대해 성능이 온전하지 못하여 현실에 적용하기 어려운 실정이다**(**explainability, **설명력 이슈)**. 더군다나, 사진이 아닌 그림에 대한 설명이 가능하도록 하는 접근법은 고안되고 있지 않은 실정이다. 또한, 정형화된 캡션 생성에서 벗어나 다양한 여러 표현을 조절하기 어렵다(controllability, **조절성 이슈)**. 이에 사진이 아닌 그림에 대해 설명이 가능하도록 확장하고, 현실적으로 시각장애인이 그림을 감상하는데 있어 더욱 더 다채로운 표현을 제시하여 도움이 될 수 있도록 한다.

접근

첫째로, **화풍 및 사조**에 따라 그림에 대한 설명이 가능하도록 데이터 셋을 정제한다. 또한 **control signal**은 표현력과 직접적으로 관계되므로, 그림에 대한 특징을 어떻게 신호에 실어 전달할 수 있을지 설계한다(그림에 대한 정성적인 정보를 정량적인 정보로 변환하여). 예를 들어, 그림에 대한 차가움이나 따뜻함의 정도를 정량화하여 감성을 전달할 수 있도록 signal을 설계를 한다. 신경망 설계 절차가 복잡한데, Statistical Learning Theory에 입각하여, auto-encoder, visual attention mechanism을 활용하여 **대상 단어의 그림과의 유관성**을 끌어 올릴 수 있도록 신경망을 설계 한다. **최종적으로, 그림에 대한 감성 인지 컴퓨팅이 가능하도록 만드는 것을 목표로 한다.**

