

컴퓨터학콜로키움 (COSE-405) :: 5주차 김승룡 교수님

:: On the Dense Matching in a Deep Learning Era

강의 : 2022-10-05 / 작성 날짜 : 2022-10-05

고려대학교 컴퓨터학과 2017320108

고재영

현대에 살아가는 우리는 여러 과학 기술의 비약적인 발전에 따라서 크고 작은 여러 변화의 물결을 맞아 왔다. 가까운 몇 년 동안 가장 중대한 물결을 하나 꼽자면 역시나 인공지능과 기계학습 분야라고 볼 수 있다. 이러한 비약적인 발전은 해당 개념이 도래하기 이전과 이후를 비교하여 우리 삶을 송두리째 바꿔 놓았고 더 빠른 속도로 진행중이라고 말하더라도 과언이 아닐 것이다. 이러한 인공지능 분야 안에서도, 전통적인 모델로부터 멀티 레이어 구조의 인공 신경망에 기반한 학습 기반인 **Deep Neural Network**까지 뿌리가 뻗어났다. 오늘 이 시간에는 김승룡 교수님께서 딥러닝과 특히 컴퓨터 비전에 관련하여 강연을 하셨다.

교수님께서 이번 시간 크게 강조하신 내용은 **Neural Twins**란 개념이었다. 실제 세상을 표현하고자 할 때, **neural network**로 만들어지는 필드와 같은 개념으로, 이전에 제시되었던 **3D Reconstruction**이란 개념과 유사성을 띄고 있으나 좀 더 큰 개념이다. 특히 컴퓨터 비전이란 학문의 궁극적 목표이자 교수님의 연구실의 목표 방향성이 가상 현실에서 실제 세상을 모방할 수 있는 필드를 만들어 그 안에서 뉴럴 시뮬레이터, 자율 주행, **AR/VR**, 메타버스 등 많은 것을 할 수 있게 **Neural Twins**를 구현하는 것이라고 한다.

실제 세상을 이미지로 나타낸다고 할 때, 세 가지 관점으로 **Geometric modeling**, **Neural Field modeling**, 그리고 **Generative modeling**으로 나누어 볼 수 있다. 오늘 이 강의에선 그 중에서도 **Geometric modeling**의 **Image Matching**에 대해 초점을 맞추었다. 이미지 매칭은 간단히 생각한다면 두 이미지 사이에 어떠한 대응 관계가 있느냐에 주목하여 해당 이미지 간의 연결을 찾아가는 과정이다. 단순히 본다면 간단하지만, 같은 장소나 대상에 대해 낮과 밤의 영향으로 **appearance** 차이가 큰 경우,

시간의 흐름에 따른 차이가 있는 경우, 구조물과 같이 일정한 패턴이 무수히 반복되는 경우, 고양이나 강아지와 같이 동일한 종에 대해서 다양성으로 인해 상이한 차이를 보이는 **intra-class variation** 등의 경우에 해결하는 것은 그리 간단한 작업은 아니다. 이러한 **Dense Image Matching**에 있어서 **feature vector**의 추출에 대해, **Regularizer**에 대해, **Search space**에 대해, 그리고 **Optimization**으로 4가지 화두를 꼽을 수 있다. **Handcrafted** 방법으로 직접 **feature**를 지정하는 전통적인 방법과 다르게 **deep learning**을 적용하여 **feature**를 추출하는 식의 첫 번째 패러다임 변화가 존재한다. 두 번째 패러다임은 **matching**에 관련해서도 **single global transformation**의 제한적인 방법 대신, 픽셀 단위의 매칭으로 세부 디테일이 더 잘 매칭되도록 학습하는 것이다. 매칭에 있어서 크게 생각해봐야 할 매칭 포인트의 지역적인 윈도우 사이즈에 대한 결정과 주변부와 얼마나 유사한지에 관한 인터랙션, 이렇게 두 가지의 관점에 대해 토큰화의 방법으로 **vision transformer**를 이용해 **feature**와 **cost**에 대한 **aggregation**을 적용한다. 세 번째 패러다임은 학습의 데이터 효율성에 관련해서, 기존의 지도학습은 **supervision**을 무한히 생성해낼 수 있으나 **synthesized**된 이미지에 대한 한계점을 타파하기 위해, 정답을 직접 알려주기 보다는 **augmentation**의 강도를 달리하여 **strong branch**가 **weak branch**의 결과를 따라 가도록 **pseudo label**을 주면서 학습시키는 형태의 발전이다. 마지막 패러다임은 **Search space**에 관련하는데, 2D 좌표에 관해 매칭되기 때문에 결국에 **quadratic**의 복잡도를 가지는 매칭 필드에 의한 메모리 규모에 대한 문제가 있는데, 이러한 매칭 필드에 관련하여서도 **neural learning**을 적용하여 포인트 지점을 **neural network**가 직접 디코딩하도록 학습시키는 것으로 개선시켰다.

필자는 컴퓨터 학과의 공학도로서 현재까지 약 8학기에 달하는 학부 강의를 듣게 되면서 개인적으로는 비주얼적으로 드러나고 결과를 보다 눈에 가시적으로 볼 수 있는 분야를 좋아한다는 것을 알게되어 컴퓨터 비전 및 이미지 프로세싱과 그래픽스와 관련하여 선호도가 높은 편이다. 이번 기회를 통해 컴퓨터 비전에 있어서 이미지 매칭과 관련하여 4가지 핵심적인 부분에 표적으로 딥러닝을 접목시킨 현황을 보게 되어 매우 흥미로웠다. 앞으로도 우리 삶에서 어떤 원리나 매커니즘을 보게 되더라도 혹시나 **neural network**를 접목시켜 학습으로 증강을 시킬 수 있는가를 바라보는 추가적인 시각을 견지하는 것도 좋을 것 같다는 생각이 든다.