



김선옥 한국항공대학교

## ‘인공지능 기반 컴퓨터 비전 및 영상처리 기술’ 특집호를 내며

이번 특집호에서는 인공지능 기반 컴퓨터 비전 및 영상처리 기술들을 소개하고자 합니다.

최근 몇 년간 인공지능 기술의 급격한 발전은 다양한 산업 분야에서 혁신을 이끌어 내고 있습니다. 그중에서도 컴퓨터 비전 및 영상처리 기술은 인공지능 기술의 발전을 이끈 핵심 응용 기술입니다. 컴퓨터 비전 기술은 컴퓨터가 이미지나 영상을 통해 정보를 이해하고 해석하는 기술을 의미하며, 이를 위해 디지털 이미지를 변환 및 개선하는 영상처리 기술과 함께 다양한 응용 분야에서 자율 주행, 보안 및 감시, 엔터테인먼트, 미디어, 의료, 스포츠 등 많은 분야에서 관심을 받고 있습니다. 전통적인 컴퓨터 비전 기술은 수학적 이론이나 알고리즘에 의하여 특징자(feature)를 추출하고 이를 수학적으로 모델링합니다. 이러한 방법은 사전에 데이터 학습이 필요없다는 장점이 있지만 특정 작업에 최적화되어 있고, 보편적으로 잘 작동하는 특징자를 추출하는 데는 한계를 가지고 있었습니다. 최근 인공지능 기술의 발전으로 컨볼루션 신경망(Convolutional Neural Network), 트랜스포머(Transformer) 등 다양한 딥러닝 기반의 모델을 사용하여 자동적으로 최적화된 특징자를 추출하고 대규모의 데이터셋을 활용함으로써 모델이 복잡한 패턴과 변화에 대해서도 잘 학습할 수 있도록 합니다. 나아가 영상 뿐만 아니라 음성 및 텍스트 도메인의 정보를 인공지능 기술을 이용하여 융합하는 기술도 활발히 연구되고 있습니다. 본 특집호에서는 컴퓨터 비전 기술에 활용되는 데이터 처리, XAI 방법에 따른 성능 분석을 포함하여 발화 영상 생성 기술, 3D 얼굴 영상 생성 기술, 영상 기반 사람 검색 기술 등을 소개한 컴퓨터 비전 및 영상처리 기술 전문가들의 원고 5편을 준비하였습니다.

첫 번째 기고문(“다중 센서를 활용한 회귀 모델 진행 시 효율적인 데이터 전처리 기법”)에서는 다중 센서를 통해 습득한 데이터를 회귀 모델에 사용하기 위한 다양한 전처리 기법을 실험하고 그 결과를 비교합니다. 인공지능 기술을 활용하기 위해서는 고품질의 데이터셋의 확보와 데이터 전처리가 중요한데, 특히 다양한 센서를 활용하여 획득한 데이터는 더욱 다이나믹한 전처리 기법의 활용이 요구됩니다. 본 기고문에서는 여러 종류의 센서 데이터가 존재할 때 이를 학습하기 위하여 데이터 확인, 데이터 가공 및 특이값 제거, STL decomposition, noise reduction, feature selection 단계로 진행되는 데이터 전처리 기법을 설명합니다.

두 번째 기고문(“Distance Method를 바탕으로 한 eXplainable Artificial Intelligence(XAI) 기술의 정량 비교 동향”)에서는 기존의 XAI 방법들에 대한 기술에 대해 소개하고, 이러한 방법들을 정량적으로 평가하는 방법을 소개합니다. XAI란 기존 블랙박스로 여겨져 해석이 어려웠던 인공지능 결과를 설명 가능하게 해석하여 설명하는 방법입니다. 인공지능 기술들을 설명하기 위하여 많은 XAI 기술들이 제시되고 있는데 각각의 알고리즘은 다른 근거로 학습된 모델의 예상값에 대한 해석을 시도하는 문제점을 가지고 있습니다. 본 기고문에서는 XAI 알고리즘을 평가하기 위해 정확성, 일관성,

민감성, 안정성 등의 기준을 사용하여 평가하는 다양한 방법들을 소개하고 그중에서도 도장(stamp) 기반의 distance method를 사용하여 MNIST 데이터셋에서 도장의 존재 여부를 예측하는 딥러닝 모델을 XAI 방법을 이용하여 평가하고 성능을 비교합니다.

세 번째 기고문(“개인화된 3차원 생성 모델을 활용한 고품질 발화 영상 생성”)에서는 정적인 장면에서 3차원 장면을 렌더링하는 신경방사망(Neural Radiance Fields, NeRF) 기술을 오디오 기반의 발화 영상 생성에 활용하는 방법에 대해 설명합니다. 실제 환경에서 발생할 수 있는 다양한 객체와 장면에서 자연스러운 발화 영상을 생성하기 위하여 오디오 정보와 3차원 기하 정보를 합성할 수 있는 방법에 대하여 소개합니다. 인공지능 기술을 접목시킨 3D 생성형 모델 및 고품질의 3D 발화 영상 생성을 위한 프레임워크인 Talk3D 기술을 이용하여 우수한 성능의 개인화된 오디오 기반 발화 영상 생성이 가능함을 보여줍니다.

네 번째 기고문(“음성 기반 말하는 3D 얼굴 애니메이션 생성 연구 동향”)에서는 음성 신호로부터 3D 얼굴 영상을 생성하는 기술인 Speech-driven 3D Facial Animation 기술의 최근 동향과 향후 연구 방향에 대하여 소개합니다. 3D 디지털 인간은 증강현실이나 가상현실을 이용한 화상 회의, 영화, 게임에서의 애니메이션 생성과 같이 몰입형 애플리케이션에서 높은 활용도로 큰 관심을 받고 있습니다. 최근 인공지능 기술에 기반한 신경 렌더링(Neural Rendering)의 발전과 함께 실제 인간처럼 합성하는 기술이 많이 발전하였습니다. 본 기고문에서는 인공지능의 발전과 더불어 3D 얼굴 영상 생성 기술이 어떻게 발전하였는지 소개하고 각 기술들의 성능을 분석하였습니다.

다섯 번째 기고문(“영상 기반 사람 검색 주요 기술과 현황”)에서는 최근 컴퓨터 비전 분야에서 활발하게 연구되고 있는 분야 중 하나인 영상 기반 사람 검색 분야의 주요 기술과 현재 동향을 소개합니다. 영상 기반 사람 검색 기술은 사람 재식별(re-identification)과 객체 탐지(detection) 기술에 기반을 두고 있으며, 전통적인 컴퓨터 비전 접근 방식에서는 주로 특징자 추출과 매칭 기술을 사용해 왔습니다. 최근 인공지능 기술을 활용하여 트랜스포머 및 확산 모델을 통해 우수한 성능을 보이고 있는데, 이에 대한 다양한 접근 방식을 비교하고 분석하였습니다.

이번 “인공지능 기반 컴퓨터 비전 및 영상처리 기술” 특집호 출간을 위해서, 바쁘신 와중에도 원고를 준비해 주신 저자분들께 깊이 감사드립니다. 아무쪼록 본 특집호가 인공지능 기반 컴퓨터 비전 및 영상처리 기술에 대한 동향 파악과 관련 연구에 조금이나마 도움이 되기를 기대합니다.