장한얼 교수

국립한밭대학교 컴퓨터공학과

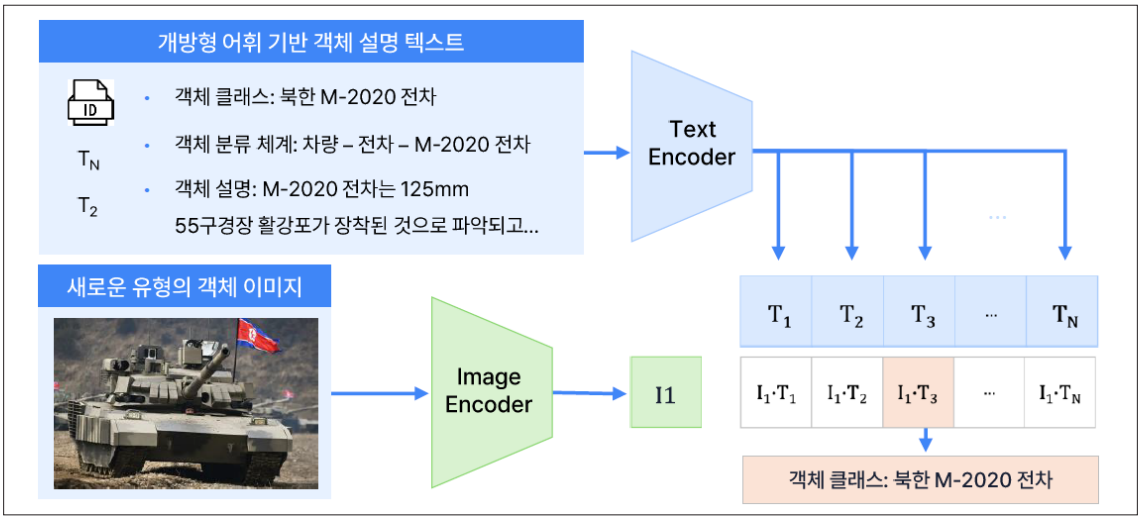
I. 인공지능미디어 연구실 소개

장한얼 교수가 운영 중인 인공지능미디어 연구실은 우주 및 국방 분야의 핵심 기술인 항공/위성 영상 처리와 디지털 콘텐츠의 신뢰성을 보장하는 멀티미디어 보안 분야의 연구를 활발하게 수행하고 있다. 주요 연구로는 제로샷 객체 탐지, 위성 영상 초해상화, 얼굴 위변조 탐지, 비가시적 워터마킹 등이 있다. 활발한 연구 활동을 통해 얻은 결과를 인공지능 분야의 최상위 국제 학술대회(AAAI, COLING 등)와 저명 SCIE 저널에 꾸준히 발표하고 있다. 또한, 핵심 원천 기술 확보를 위한 국내외 특허 출원 및 등록을 활발히 진행하고 있으며, 기술이전 성과도 다수 보유하고 있다. NAVER CLOUD, NAVER WEBTOON, 국가보안기술연구소 등 인공지능 분야 우수 기업과 정부 출연 연구소들과의 긴밀한 산학연 협력 연구를 통해 실용적인 기술을 개발하고 국가 기술 경쟁력 강화에 기여하는 것을 목표로 하고 있다.

II. 연구 분야

1. 항공/위성 영상 객체 탐지

항공 및 위성 영상은 국방, 재난 감시, 환경 모니터링, 자원 탐사 등 다양한 분야에서 그 중요성이 날로 증대되고 있다. 본 연구실에서는 방대한 양의 항공/위성 영상 빅데이터로부터 의미 있는 정보를 추출하고 정밀한 분석을 수행하기 위한 혁신적인 인공지능 알고리즘을 개발하고 있다. 특히, 드론이나 위성에서 촬영된 영상의 특정 객체나 지역을 식별하고 분류하는 객체 탐지 및 분할(Object Detection & Segmentation) 기술, 위성 영상의 해상도를 획기적으로 향



<그림 1> 개방형 어휘 기반 제로샷 및 퓨샷 탐지 연구

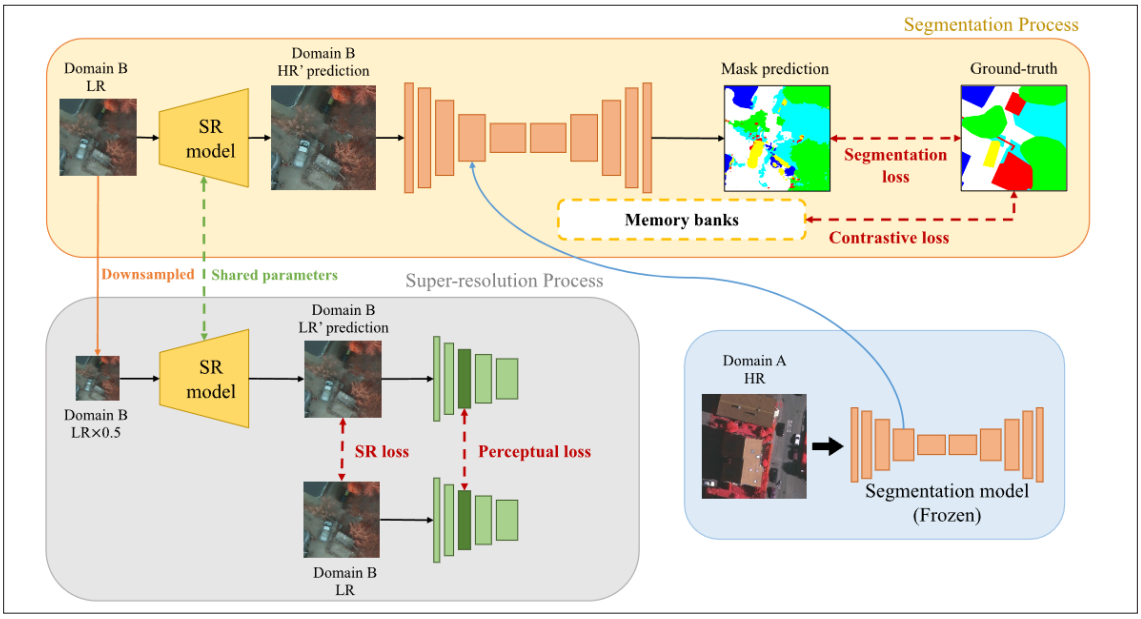
상시키는 초해상화(Super-Resolution) 기술에 대한 깊이 있는 연구를 수행하고 있다.

본 연구실에서는 군사적 환경의 특수성으로 인해 표적 영상 데이터 확보가 제한적인 상황을 극복하기 위한 개방형 어휘 기반 제로샷 및 퓨샷 탐지 연구를 수행하였다(<그림 1> 참조). 자연어 특징과 시각적 특징 간의 군용 표적 대상 임베딩 정렬을 최적화하고 이를 통해 Cross-Modality Fusion 능력을 강화하였다. 객체의 클래스 정보, 분류 체계, 객체의 시각적 특징 정보를 결합하여 텍스트 임베딩함으로써 군용 표적 탐지 성능을 개선시켰다.

2. 항공/위성 영상 초해상화

위성이나 항공기에서 촬영한 원격 감지(Remote Sensing) 영상은 토지 피복 분석, 도시 계획 등 다양한 분야에 필수적이다. 딥러닝 기반의 의미론적 분할(Semantic Segmentation) 기술은 각 픽셀이 어떤 종류(건물, 나무, 도로 등)에 속하는지 분류하여 정밀한 분석을 가능하게 한다. 하지만, 특정 지역의 고해상도 영상으로 학습한 AI 모델을 해상도가 낮은 다른 지역의 영상에 적용하면, 데이터의 분포 차이(도메인 격차)와 해상도 차이로 인해 성능이 크게 저하되는 문제점이 있다.

본 연구실에서는 해상도와 도메인 격차 문제를 해결하기 위해 단순히 저해상도 이미지의 화질을 개선하는 것을 넘어, '분할'이라는 최종 과제의 성능을 극대화하는 방향으로 초해상화 모델을 학습시키는 과업 지향적인 초해상화 프레임워크를 제안하였다(<그림 2> 참조). 즉, 기존의 고해상도로 학습한 분할 모델은 그대로 둔 채, 해당 모델이 가장 잘 인식할 수 있는 형태로 이미지를 변환해 주는 '맞춤형 안경'과 같은 초해상화 모델을 개발하는 것이다. 초해상화 모델 학습 시, 픽셀 단위의 정확도를 높이는 분할 손실(Segmentation Loss), 영역별 특징을 명확히 구분하는 대조 손실(Contrastive Loss), 사람의 눈으로 보기에 자연스러운 이미지를 만드는 지각 손실(Perceptual Loss), 원본 고해상도 이미지의 정보를 유지하는 픽셀 손실(Pixel Loss)을 모두 함께 사용하는 복합 손실 함수(Compound Loss)를 제안하여 최적의 결과를 도출하였다.



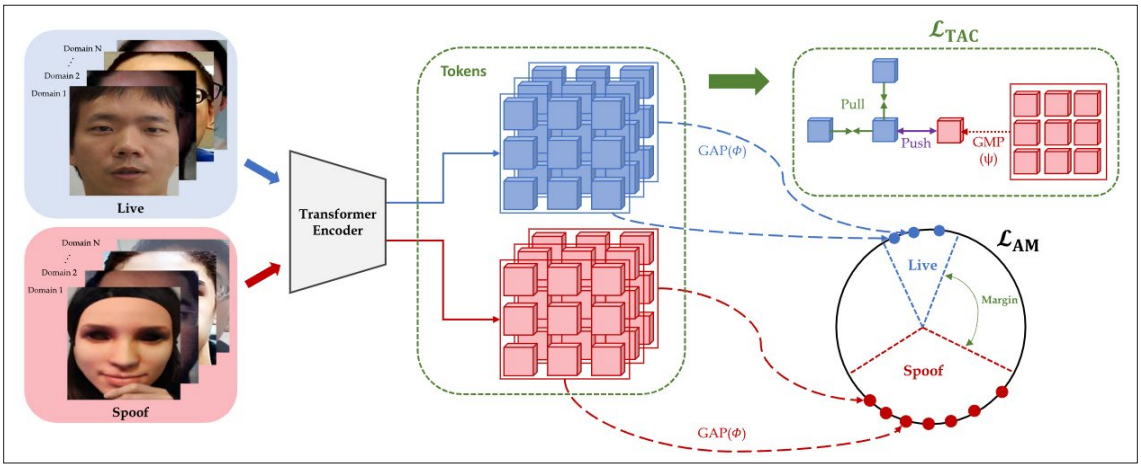
<그림 2> 도메인 적응을 위한 과업 지향적 초해상화 연구

3. 얼굴 위조 방지

얼굴 인식이 스마트폰 잠금 해제부터 보안 시스템까지 널리 사용되면서, 사진, 동영상, 3D 마스크 등을 이용한 위조 공격(Spoofing Attack)이 심각한 위협이 되고 있다. 이를 방지하기 위한 얼굴 위조 방지(Face Anti-Spoofing, FAS) 기술의 중요성이 높아지고 있다. 기존 FAS 기술들은 학습 과정에서 경험한 '알려진' 공격 유형은 잘 탐지하지만, 새롭게 등장하는 정교하고 '알려지지 않은(Unknown)' 공격 유형에는 매우 취약한 한계를 보인다. 또한, 기존 연구들은 실제 얼굴과 위조 얼굴의 특징을 대칭적으로 학습하려 했으나, 위조 얼굴은 공격 방식에 따라 특징이 매우 다양하여 이러한 접근 방식은 일반화 성능에 한계가 있었다.

본 연구실에서는 비대칭 대조 학습, 토큰 단위 학습, 각도 마진 손실을 효과적으로 결합하여 알려지지 않는 공격 유형에 대한 탐지율을 크게 향상시켰다(<그림 3> 참조). 비대칭 대조 학습(Asymmetric Contrastive Learning)의 핵심은 '실제 얼굴(Live) 특징은 최대한 조밀하게 모으고, 위조 얼굴(Spoof) 특징은 넓게 흩어지도록' 학습하는 것이다. 이를 통해 실제 얼굴만의 고유하고 일관된 특징 공간을 형성하여, 어떤 유형의 새로운 위조 공격이 들어오더라도 이 공간 밖에 위치하도록 만들어 탐지율을 높인다. 토큰 단위 학습(Token-wise Learning)은 얼굴 전체를 한 번에 분석하는 대신, 이미지를 여러 개의 작은 조각(토큰)으로 나누어 각 조각에 담긴 미세한 '생동감 단서(Liveness Cue)'에 집중하는 것이다. 이를 통해 얼굴의 형태나 특정 인물 정보가 아닌, 위조 여부를 판단할 수 있는 본질적인 패턴을 학습한다. 각도 마진 손실(Angular Margin Loss)은 학습된 실제 얼굴 특징과 위조 얼굴 특징 사이에 더 넓은 '안전 마진'을 두어 경계선에 애매하게 걸치는 경우를 줄이고 판별 성능을 극대화한다.

토큰 단위 비대칭 대조 학습 기반 얼굴 위조 방지 기술은 기존 기술 대비 알려지지 않은 새로운 공격 유형에 대한 방어 능력을 획기적으로 향상시켰다. 또한, 더 현실적이고 어려운 환경을 가정한 새로운 평가 프로토콜을 제시하고 최



<그림 3> 토큰 단위 비대칭 대조 학습 기반 얼굴 위조 방지 연구

고 성능을 달성하였다.

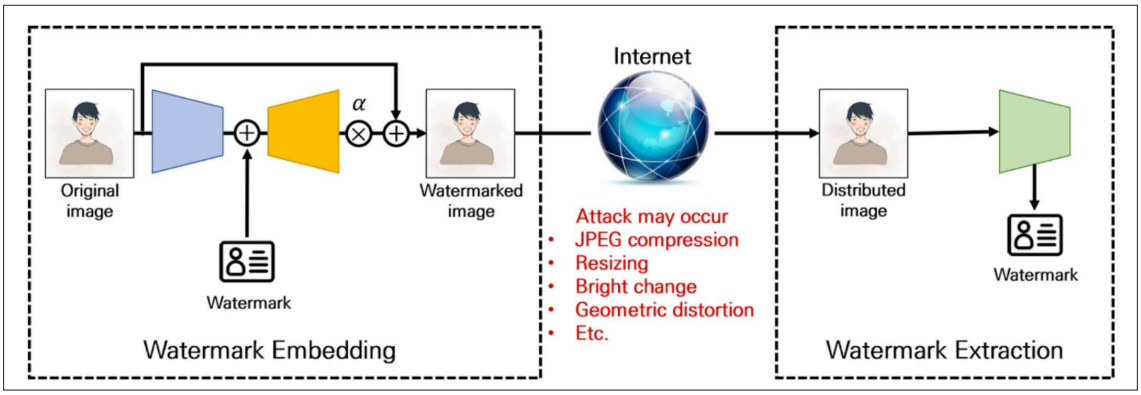
4. 비가시적 이미지 워터마킹

웹툰, 영화 등 K-콘텐츠의 세계적인 인기와 함께 불법 복제 및 유통으로 인한 저작권 침해 문제도 심각해지고 있다. 비가시적 이미지 워터마킹은 콘텐츠에 육안으로 식별 불가능한 저작권 정보를 삽입하여 불법 유포 시 출처를 추적하는 핵심적인 저작권 보호 기술이다.

불법 유포자들은 워터마크를 제거하기 위해 JPEG 압축, 이미지 잘라내기, 회전, 밝기 조절, 노이즈 추가 등 여러 공격을 복합적으로 사용한다. 또한, 학습 과정에서 고려되지 않은 새로운 방식의 디노이징 필터 등을 적용하기도 한다. 기존 기술들은 이처럼 정교하고 복합적인 공격에 매우 취약한 문제점이 있다.

본 연구실에서는 공격 시뮬레이션 네트워크 설계, 복합 및 미지의 공격 대응, 커리큘럼 학습 방법을 제안하여 공격자의 정교하고 복합적인 공격에 대한 강인성을 크게 높였다(<그림 4> 참조). 공격 시뮬레이션 네트워크는 딥러닝 학습 과정에 '가상의 공격자' 역할을 하는 네트워크이다. 이 공격 네트워크는 JPEG 압축, 기하학적 왜곡, 디노이징 등 다양한 공격을 시뮬레이션하며 워터마킹 모델이 이러한 공격을 예상하고 이에 대한 방어 능력을 갖추도록 훈련시킨다. 복합 및 미지의 공격 대응하기 위해 다양한 종류의 공격을 조합한 '복합 공격'과 학습에 사용되지 않은 '알려지지 않은 공격' 시나리오를 모두 학습에 반영하여, 어떤 종류의 훼손 시도에도 워터마크가 강인하게 살아남도록 하였다. 초기에는 약한 공격부터 시작하여 점차 강하고 복잡한 공격으로 난이도를 높여가는 '커리큘럼 학습' 방식을 적용하여, 모델이 안정적으로 높은 강인성을 확보하도록 하였다.

이 비가시적 워터마킹 기술은 불법 유포자를 효과적으로 추적하고, 콘텐츠 불법 복제로 인한 경제적 피해를 크게 감소시키는데 기여할 수 있다. 또한, 웹툰 플랫폼을 비롯한 다양한 콘텐츠 산업의 핵심 기술로 활용되어 K-콘텐츠 산업의 지속 가능한 성장을 뒷받침할 것이다.



<그림 4> 복합 공격에 강인한 딥러닝 기반 이미지 워터마킹 연구

Ⅲ. 주요 실적

- 인공지능 분야 BK21 우수학술대회 2편 발표(AAAI, COLING) 및 저명 SCIE 저널 12편 게재
- 인공지능 경진대회 수상 9건(Kaggle, AI CONNECT, 연구개발특구 등)
- 국방분야 연구과제 4건 수행(합동참모본부, 국방기술진흥연구소, 에스아이에이 등)
- 보안분야 연구과제 15건 수행(국가보안기술연구소, 네이버웹툰, 한국원자력통제기술원 등)
- 제로샷 객체 탐지 및 심층신경망 암호화 관련 기술이전 3건

저 자 소 개

장 한 얼

- 2018년 : 한국과학기술원 전산학부 박사
- 2018년 : 네이버 클로바 인공지능 연구원
- 2018년 ~ 2020년 : 국가보안기술연구소 사이버보안본부 선임연구원
- 2020년 ~ 현재 : 국립한밭대학교 부교수
- 주관심분야 : 위성/항공영상처리, 멀티미디어 보안, 컴퓨터비전, 멀티모달