



회/원/소/개

김선옥 교수

한국항공대학교 인공지능학과



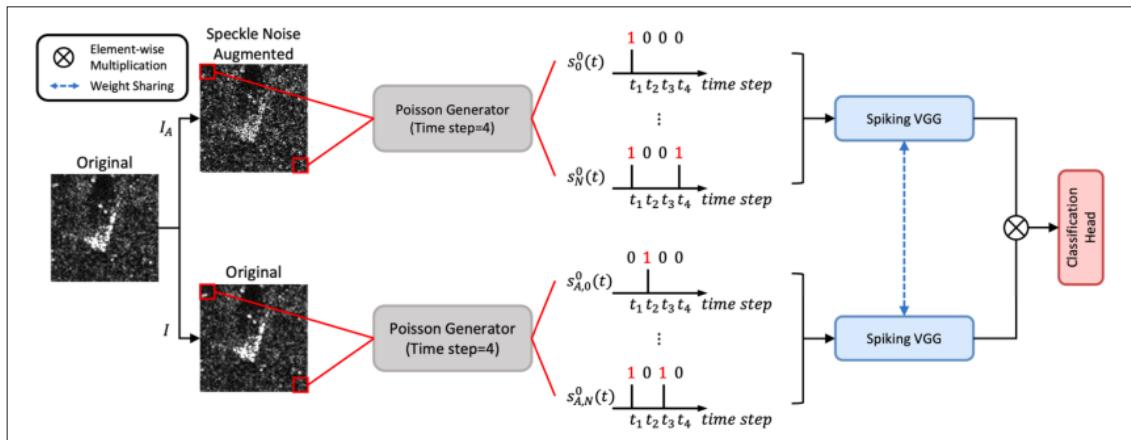
I. 저자 소개

김선옥 교수가 운영 중인 한국항공대학교 컴퓨터비전 연구실(CVLab.)은 컴퓨터비전 및 인공지능 분야 전반에 걸친 연구를 수행하고 있으며, 특히 위성 영상 분석과 인공지능 기반 우주 기술 개발을 중심으로 한 융합 연구에 주력하고 있다. 구체적으로 위성 영상의 이해 및 분석 기술을 핵심 연구 주제로 삼아, 최근에는 스파이킹 신경망(Spiking Neural Networks), 파운데이션 모델 등 인공지능의 핵심 기술을 위성 영상 분석에 접목하는 연구를 활발히 진행하고 있다. 이러한 연구 성과는 인공지능 및 원격탐사 분야의 최상위 국제 학술대회와 저명 SCIE 저널에 발표되고 있으며, 나아가 실제 산업 및 사회적 문제 해결을 목표로 국내 주요 연구소 및 산업체와의 협력 연구를 통해 긴밀한 산학 협력 관계를 구축하고 공동 연구를 수행하고 있다.

II. 연구 분야

1. 스파이킹 신경망을 활용한 SAR 영상 분류 연구

합성 개구 레이다(SAR) 영상 분류는 기상 및 조명 조건에 구애받지 않는 특성으로 인해 국방 및 감시 분야의 핵심 기술로 꼽히지만, 영상 고유의 스페클 잡음(Speckle Noise)과 라벨 데이터 부족 문제로 인해 고정밀 분류에 어려움이 있었다. 또한 기존의 인공 신경망(ANN)은 높은 연산 비용과 전력 소비로 인해 제한된 환경에서의 운용에 제약이 있고,



<그림 1> 제안하는 스파이킹 신경망 기반 SAR 영상 분류 모델

기존 스파이킹 신경망(SNN)은 얇은 구조로 인해 정확도 확보에 한계가 있었다. 이러한 한계는 실제 운용 환경에서 신뢰성 높은 SAR 영상 인식 기술의 적용을 어렵게 만드는 주요 요인으로 작용해 왔다. 본 연구실에서는 이러한 한계를 극복하기 위해 스파이킹 신경망을 활용하여 잡음에 강인한 SAR 영상 분류 프레임워크를 개발하였다. 제안 모델은 원본 영상과 스펙클 잡음이 증강된 영상을 쌍으로 입력받아, 두 입력 간의 공통된 특징을 융합하는 구조를 채택하였다(<그림1> 참조). 이를 통해 데이터가 부족한 환경에서도 잡음에 대한 강인성을 구조적으로 확보하였다. 특히 본 연구는 포아송 인코딩을 도입하여 스펙클 잡음을 스파이크 생성의 확률적 특성으로 흡수함으로써 정보 전달 효율을 높였다.

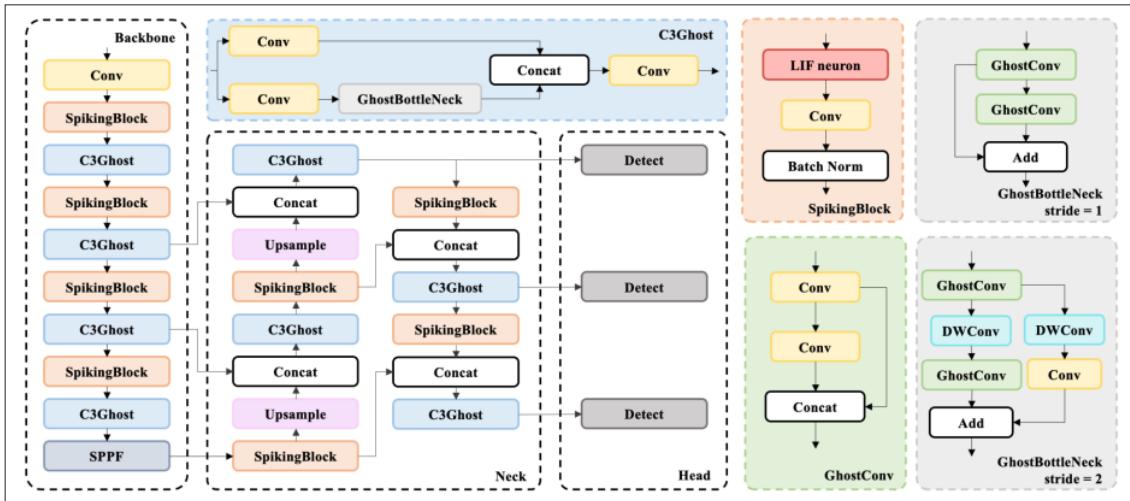
본 연구실에서는 위와 같이 고비용 연산이 요구되는 환경에서 필수적인 에너지 효율성과 분류 정확도를 동시에 만족하는 딥러닝 모델을 구현함으로써 향후 저전력 위성 시스템 및 실시간 원격탐사 응용을 위한 핵심 기반 기술을 실현하고자 목표하고 있다.

2. 위성 영상 객체 탐지 네트워크 개발 및 경량화

합성개구레이더(SAR) 영상 기반 선박 탐지는 해양 감시, 국방 및 재난 대응 등 다양한 위성 응용 분야에서 핵심적인 기술로 활용되고 있으나, 기존의 합성곱 신경망(CNN) 기반 탐지 기법들은 높은 전력 소모와 연산 복잡도로 인해 실시간 처리가 요구되는 위성 온보드 환경에 적용하는 데 근본적인 한계를 지니고 있다.

본 연구실에서는 이러한 문제를 해결하기 위하여, 스파이킹 신경망(Spiking Neural Network, SNN)의 연산 특성과 에너지 효율성을 이용하여 SNN 기반의 SAR 선박 탐지 네트워크를 제안하였다(<그림2> 참조). 제안한 방법은 YOLOv5 구조를 기반으로 기존 합성곱 연산을 스파이킹 기반 연산으로 대체하고, 경량화된 C3Ghost 모듈을 스파이킹 신경망 기반 객체 탐지 구조에 최초로 도입함으로써 연산량과 파라미터 수를 효과적으로 감소시키면서도 탐지 성능을 유지할 수 있도록 설계하였다. 또한, Leaky Integrate-and-Fire(LIF) 뉴런과 대체 기울기 학습 기법을 활용한 직접 학습 방식을 통해 기존 ANN-to-SNN 변환 방식의 한계를 극복하고, 스파이킹 신경망의 에너지 효율적 특성을 실제 객체 탐지 문제에 적용하였다.

본 연구실에서는 기존의 스파이킹 신경망 기반 객체 탐지 연구들이 주로 이벤트 기반 영상이나 자연 영상에 국한되

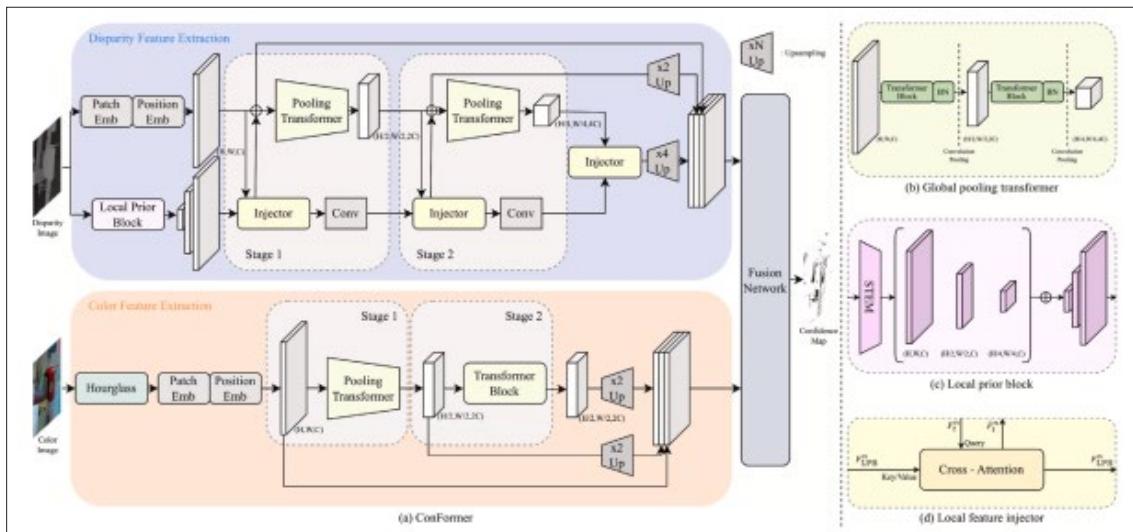


<그림 2> 제안하는 스파이킹 신경망 기반의 선박 탐지 네트워크

어 있던 한계를 넘어, 복잡한 산란 특성과 잡음을 포함하는 SAR 영상 환경에서도 스파이킹 신경망이 효과적으로 활용될 수 있음을 확인하고자 다양한 연구를 진행하고 있다. 이를 통해 위성 온보드 실시간 영상 처리 환경을 고려한 구체적인 모델 설계 및 구현 가능성을 제시함으로써, 스파이킹 신경망이 생물학적 모델링을 넘어 고성능과 에너지 효율이 동시에 요구되는 위성 시스템의 실질적인 대안 기술로 확장하고자 한다.

3. 스테레오 매칭의 신뢰도 추정 기술 연구

스테레오 매칭을 통해 획득한 깊이 영상은 반사 표면, 무텍스처 영역, 반복 패턴, 가림 현상 등 다양한 요인으로 인해 정확한 추정이 어려우며, 이를 추정하기 위한 신뢰도(confidence) 추정 기술은 깊이 영상 품질 향상의 핵심 요인으로 작용한다. 기존의 딥러닝 기반 신뢰도 추정 방법들은 주로 합성곱 신경망(CNN)에 기반하여 국소적인 수용 영역 내 특징을 학습함으로써, 장거리 상관관계나 전역적 문맥 정보를 충분히 반영하는데 한계가 있었다. 본 연구실에서는 이러한 한계를 극복하기 위해 트랜스포머(Transformer) 구조를 최초로 적용하였다(<그림 3> 참조). 본 연구실에서 제안된 방법은 자기 주의(self-attention) 메커니즘을 활용하여 영상 전반의 전역적 관계를 모델링함으로써, 기존 CNN 기반 접근법에서 포착하기 어려웠던 의미적 상관관계를 효과적으로 학습하였다. 특히, 깊이 영상에 대해서는 전역 풀링 트랜스포머(global pooling transformer)를 통해 장거리 의존성을 학습하는 동시에, 지역 사전 정보를 주입하는 구조를 도입하여 미세한 경계와 세부 구조 정보까지 함께 반영할 수 있도록 설계하였다. 또한, 컬러 영상의 경우 깊이 정보와 상이한 특성을 고려하여 국소 특징 주입을 배제하고, 전역 문맥 이해에 초점을 둔 Transformer 기반 특징 추출 방식을 채택함으로써 불필요한 시각적 잡음의 영향을 최소화하였다. 이후 깊이 영상 기반 신뢰도 특징과 컬러 영상 기반 신뢰도 특징을 가중치 기반 주의(attention) 융합 방식으로 통합함으로써, 각 모달리티의 상대적 중요도를 반영한 고성능의 신뢰도 맵을 획득하였다. 이는 단순 결합 방식이나 기존 융합 전략 대비, 모달리티 간 상호보완적 정보를 효과적으로 활용할 수 있다는 점에서 우수성을 지닌다. 본 연구실에서는 스테레오 신뢰도 추정에서 전역 문맥 정보



<그림 3> 제안하는 트랜스포머 기반의 신뢰도 추정 네트워크

의 중요성을 실험적으로 규명하고, 트랜스포머 기반 구조가 해당 문제에 실질적인 이점을 제공함을 확인함으로써 전역-국소 정보 통합 전략을 이용하여 다양한 분야에서 입력 정보를 효율적으로 융합할 수 있는 연구로 확장하고자 한다.

III. 연구 실적

- 인공지능 분야 BK21 우수학술대회 6편 발표(CVPR, ICCV 등) 및 저명 SCIE 저널 15편 게재
- 국방분야 연구과제 4건 수행(한화시스템, LIG넥스원, 대한항공 등)

저자 소개



김선옥

- 2019년 : 연세대학교 전기전자공학과 박사
- 2019년 : 로잔연방과대학(EPFL), 스위스, 방문연구원
- 2019년 ~ 2021년 : 연세대학교 박사후연구원
- 2021년 ~ 현재 : 한국항공대학교 인공지능학과 부교수
- 주관심분야 : 위성영상분석, 컴퓨터비전, 인공지능, 스파이크 신경망