

일반논문 (Regular Paper)

방송공학회논문지 제30권 제3호, 2025년 5월 (JBE Vol.30, No.3, May 2025)

<https://doi.org/10.5909/JBE.2025.30.3.427>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

## 스파이크 신경망을 활용한 편광합성개구레이다 영상 기술 연구

김 선 옥<sup>a)\*</sup>, 최 환 효<sup>a)</sup>

### PolSAR Image Classification using Spike Neural Network

Sunok Kim<sup>a)\*</sup> and Hwanhyo Choi<sup>a)</sup>

#### 요 약

본 연구에서는 스파이킹 신경망(Spiking Neural Network, SNN)을 기반으로, 특징 융합(feature fusion) 기법을 활용하여 PolSAR(Polarimetric Synthetic Aperture Radar) 이미지 분류 성능을 비교·평가하였다. PolSAR 이미지는 다양한 편광 정보를 포함하고 있어 고차원적 특성 간 상호관계를 효과적으로 반영하는 기술이 요구된다. 기존의 분류 기법은 대부분 단일 채널 또는 전통적인 인공신경망(Artificial Neural Network, ANN)에 기반하여 설계되어, 계산 비용이 크거나 편광 정보 간의 상호작용을 충분히 활용하지 못하는 한계가 존재한다. 이를 해결하기 위해, 본 연구는 SNN의 에너지 효율성과 시간 정보 처리 능력을 바탕으로 초기 융합(Early Fusion) 및 후기 융합(Late Fusion) 기법을 적용하였다. 두 융합 방식은 다양한 편광 정보를 통합된 형태로 변환하며, 이를 통해 PolSAR 이미지 내 중요한 구조적 특징을 효과적으로 추출할 수 있도록 설계되었다. OpenSARShip 2.0 데이터셋을 활용한 실험 결과, 초기 융합 방식이 단일 채널 방식 대비 약 3% 이상의 정확도 향상을 보였으며, 전체적으로 특징 융합이 분류 성능 개선에 유의미한 효과를 미치는 것으로 나타났다. 본 연구는 PolSAR 분류 문제에 있어 SNN 구조의 실용 가능성과 고차원 편광 데이터 분석을 위한 특징 융합 기법의 적용 가능성을 실험적으로 입증하였다.

#### Abstract

In this study, based on the Spiking Neural Network (SNN), the PolSAR (Polarimetric Synthetic Aperture Radar) image classification performance was compared and evaluated using a feature fusion technique. Since the PolSAR image contains various polarization information, a technology that effectively reflects the interrelationship between high-dimensional characteristics is required. Most of the existing classification techniques are designed based on a single channel or traditional Artificial Neural Network (ANN), so there is a limitation in that the calculation cost is large or the interaction between polarization information cannot be fully utilized. To solve this problem, this study applied the Early Fusion and Late Fusion techniques based on the energy efficiency and time information processing ability of SNN. The two fusion methods convert various polarization information into an integrated form, and through this, they are designed to effectively extract important structural features in the PolSAR image. Experiments using the OpenSARShip 2.0 dataset showed that the initial fusion method improved accuracy by more than 3% compared to the single-channel method, and the overall feature fusion had a significant effect on improving the classification performance. This study experimentally demonstrated the practicality of SNN structure in PolSAR classification problems and the applicability of feature fusion techniques for high-dimensional polarization data analysis.

Keyword : Spiking Neural Networks, Feature Fusion, PolSAR Image, Classification

Copyright © 2025 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

## I. 서 론

최근 인공지능의 발전과 함께, 인공신경망(Artificial Neural Network, ANN)은 영상 처리, 자연어 처리, 음성 인식 등 다양한 분야에서 뛰어난 성능을 보여주고 있다. 그러나 ANN은 높은 연산량과 메모리 사용량, 그리고 지속적인 연산으로 인한 전력 소모가 커, 에너지 효율성이 중요한 임베디드 시스템이나 실시간 애플리케이션에는 적용에 제약이 있다<sup>[1]</sup>.

이를 보완하기 위한 대안으로 제안된 것이 스파이킹 신경망(Spiking Neural Network, SNN)이다. SNN은 생물학적 뉴런의 발화 메커니즘을 모방하여 이진 스파이크 신호를 이벤트 기반으로 전달하며, 이를 통해 계산량을 획기적으로 줄이고 에너지 소모를 최소화할 수 있는 장점이 있다<sup>[2]</sup>. 특히 시간 정보를 직접적으로 처리할 수 있다는 특성 덕분에, 동적 패턴 인식 및 시계열 분석에 적합한 구조를 가진다.

한편, PolSAR(Polarimetric Synthetic Aperture Radar) 이미지는 HH, HV, VV 등 다양한 편광 채널을 통해 지표의 전자기적 특성을 고차원적으로 표현하는 원격탐사 데이터이다. 이 데이터는 강우, 구름, 야간 등 기상 및 환경 조건에 관계없이 안정적으로 수집할 수 있어, 환경 모니터링, 해양 관측, 국방 등에서 널리 활용된다<sup>[3]</sup>. 그러나 PolSAR 데이터는 스펙클 노이즈(speckle noise), 높은 데이터 차원, 그리고 채널 간 복잡한 상호작용 등으로 인해 기존의 분류 기법만으로는 한계가 존재한다.

기존 PolSAR 이미지 분류 연구는 대부분 CNN, SVM, Random Forest 등 기존 머신러닝 기법에 기반하고 있으며, 이들은 고차원 편광 정보를 효과적으로 통합하거나 연산량을 제어하는 데 어려움을 겪고 있다<sup>[4]</sup>. 또한 이들 기법은

주로 단일 채널 또는 단순한 특징 기반 처리에 의존하는 경우가 많아, 복잡한 구조를 갖는 PolSAR 데이터의 정보 활용에 제약이 있다.

이러한 문제를 해결하고자 본 연구에서는 SNN 기반의 PolSAR 이미지 분류 프레임워크를 제안한다. 제안된 프레임워크는 다양한 편광 채널로부터 입력된 데이터를 스파이크 형태로 변환한 뒤, Spiking VGG-9 구조를 활용하여 시공간적 특징을 추출한다. 또한 입력 특징의 융합 방법으로 초기 융합(Early Fusion)과 후기 융합(Late Fusion)을 각각 적용하고, 동일한 학습 조건 하에서 성능을 비교 평가한다<sup>[5]</sup>.

본 연구의 목적은 다음과 같다. (1) SNN 구조가 PolSAR 이미지 분류에 적합한지를 검토하고, (2) 다양한 특징 융합 기법이 분류 성능에 미치는 영향을 분석하며, (3) 연산 효율성과 성능 향상 간의 트레이드오프를 실험적으로 규명하는 것이다. 본 연구는 PolSAR 분류 문제에 SNN을 적용한 연구 중 최초로, 다양한 융합 전략에 따른 성능 비교를 체계적으로 수행함으로써 향후 에너지 효율적인 원격탐사 이미지 분석 시스템 설계에 기여하고자 한다.

## II. 본 론

### 1. Spiking Neural Network

스파이킹 신경망(Spiking Neural Network, SNN)은 생물학적 신경망을 모방한 신경망 모델로, 신경 세포의 발화와 시간 정보를 활용하여 신경 신호를 처리한다. 기존의 인공 신경망(Artificial Neural Network, ANN)과 달리 이벤트 기반으로 작동하며, 계산 및 에너지 효율성이 높은 모델로 주목받고 있다. 스파이킹 뉴런 모델에는 Hodgkin-Huxley (H-H)<sup>[6]</sup> 모델, IF(Integrate-and-Fire)<sup>[7]</sup> 모델, LIF(Leaky Integrate-and-Fire)<sup>[8]</sup> 모델 등이 있다. 본 연구에서는 SNN의 동작 원리를 수학적으로 쉽게 모델링한 LIF를 사용한다.

LIF 모델은 뉴런의 전기적 특성을 단순화하여 모델링한 것으로, 입력 신호가 일정 임계값에 도달하면 뉴런이 발화하고 막전위를 초기화한다. 이 과정에서 뉴런은 입력 신호가 시간이 지남에 따라 점진적으로 감소하는 특성을 포함한다.

a) 한국항공대학교 인공지능학과(Dept. of Artificial Intelligence, Graduate School, Korea Aerospace University)

‡ Corresponding Author : 김선옥(Sunok Kim)

E-mail: sunok.kim@kau.ac.kr

Tel: +82-2-300-0262

ORCID: <https://orcid.org/0000-0002-9665-4214>

※ This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (NRF-2021R1C1C2005202). (Corresponding author: Sunok Kim)

· Manuscript March 14, 2025; Revised April 21, 2025; Accepted April 21, 2025.

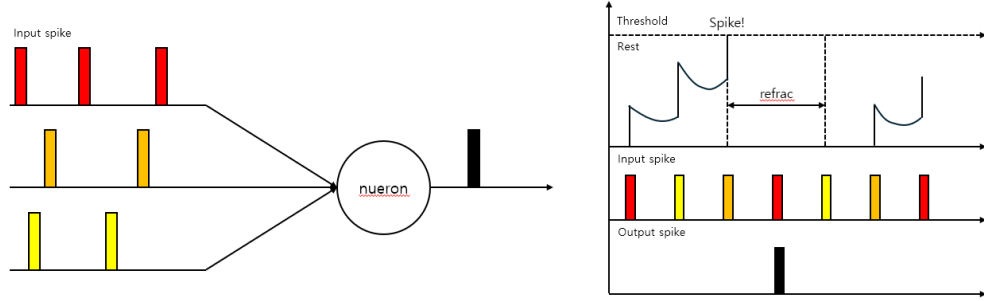


그림 1. LIF 모델 예시  
Fig. 1. Example of LIF Model

LIF 모델은 뉴런의 충전(Charging), 방전(Discharging), 휴식(Resting)을 가지고 있으며 다음과 같은 미분 방정식으로 막전위의 변화를 표현할 수 있다.

$$\tau_m \frac{dV(t)}{dt} = V_{rest} - V(t) + R_m I(t) \quad (1)$$

위의 식 (1)에서  $V(t)$ 는 시간  $t$ 에서의 뉴런 막전위(membrane potential),  $V_{rest}$ 는 안정 상태 전위(resting potential),  $R_m$ 은 뉴런의 입력 저항(input resistance),  $I(t)$ 는 시간  $t$ 에서의 입력 전류(input current),  $\tau_m = R_m C$ 은 막전위 시간 상수(membrane time constant)를 의미한다. 이 미분 방정식은 LIF 뉴런이 전류 자극에 따라 막전위가 변화하는 연속 동역학을 표현한다.

$$V(t) = V_{rest} + RI_0(1 - \exp(-\frac{t}{\tau})) \quad (2)$$

위의 식 (2)는 일정한 전류  $I_0$ 가 입력될 때, 시간에 따라 막전위가 지수적으로 증가하는 충전 과정을 나타낸다. 초기 막전위가  $V_{rest}$  이고, 입력이 지속될수록 막전위는 점차 포화 상태에 도달한다. 이는 LIF 모델의 중요한 생물학적 특성을 반영한 식이다.

$$\begin{aligned} H[n] &= f(V[n-1], X[n]) \\ &= V[n-1] + \frac{1}{\tau} (V_{reset} - V[n-1] + X[n]) \\ S(n) &= \Theta(V[n] - V_{th}) \\ V[n] &= H[n] \cdot (1 - S[n]) + V_{reset} \cdot S[n] \end{aligned} \quad (3)$$

위의 식 (3)에서  $H[n]$ 은 충전 상태 예측값,  $X[n]$ 은 입력 자극,  $V[n]$ 은 시간  $n$ 에서의 뉴런 막전위,  $V_{reset}$ 은 스파이크 후 초기화 전위,  $S[n]$ 은 뉴런의 스파이크 발화 여부(0 또는 1),  $V_{th}$ 는 발화 임계값,  $\Theta(x)$ 는 계단 함수(Heaviside Step Function)이다. 이 수식은 뉴런이 입력을 받아 충전하고, 임계값 도달 시 발화 및 막전위 리셋 과정을 이산 시점에서 계산하는 과정을 나타낸다. 대부분의 딥러닝 프레임워크에서는 이 이산 모델을 기반으로 구현된다.

## 2. Spike-Timing-Dependent Plasticity

스파이크 시차 의존 가소성(Spike-Timing-Dependent Plasticity, STDP)<sup>[9]</sup>은 뉴런 간의 연결 강도(가중치)가 시냅스 전후의 발화 시간 차이에 따라 변화하는 학습 메커니즘이다. 이는 생물학적 신경망에서 관찰된 헤비안(Hebbian) 학습 원리를 SNN에 적용한 것으로 STDP는 스파이크 뉴런의 불연속적인 특성에 따른 역전파의 사용할 수 없는 점을 대신해서 SNN에서 학습과 적응을 가능하게 한다. STDP는 다음과 같은 규칙으로 작동한다.

STDP는 뉴런 간의 시공간적 관계를 학습하며, SNN의 자율적 학습 능력을 향상시킨다. 또한 입력 신호 패턴을 시간에 따라 동적으로 학습하고 적응할 수 있는 능력을 제공한다.

$$\Delta_w = \begin{cases} A_+(w) = \eta + \exp(w_{init} - w) \\ A_-(w) = \eta + \exp(w - w_{init}) \end{cases} \quad (4)$$

위의 식 (4)에서  $A_+(w)$ 는 LTP,  $A_-(w)$ 는 LTD를,  $\eta$

는 학습률(Learning Rate),  $w_{init}$ 는 초기 시냅스 가중치,  $w$ 는 현재 시냅스 가중치 값이다. 이 수식은 Spike-Timing- Dependent Plasticity(STDP) 학습 규칙을 간결하게 나타낸 것으로, 시냅스 전 뉴런의 발화가 시냅스 후 뉴런보다 먼저 발생할 경우 가중치가 증가하는 장기 시냅스 강화(Long-Term Potentiation, LTP)가 일어나며, 반대로 시냅스 후 뉴런이 먼저 발화할 경우 가중치가 감소하는 장기 시냅스 감소(Long-Term Depression, LTD)가 발생한다.

### 3. 인코딩 방법(Encoding Method)

SNN은 이벤트 기반 모델로 작동하기 때문에 입력 데이터를 스파이크 신호로 변환해야 한다. 이러한 변환 과정을 스파이크 인코딩(Spike Encoding)<sup>[10]</sup>이라 하며, SNN의 작동에 필수적이다. 정적인 데이터를 SNN이 처리할 수 있는 연속적인 형태로 변환해야 하며 이렇게 변환된 데이터는 입력 신호의 패턴을 시간 축으로 표현하여 SNN의 특성을 극대화할 수 있다. 이번 연구에서는 Rate Coding을 사용했으며, Rate Coding은 입력 신호의 강도를 발화 빈도로 변환하는 방법이다. 입력값의 크기에 비례하여 스파이크를 생성하는 방법으로 구현이 간단하며 대부분의 SNN에서 쉽게 적용할 수 있고 입력 신호의 강도를 직관적으로 표현할 수 있는 장점이 있지만, 스파이크 빈도를 기반으로 하여 시간 정보를 풍부하게 표현하지 못할 수 있다는 단점이 있다.

## III. 제안된 네트워크 아키텍처

### 1. 백본 네트워크: Spiking VGG-9 with BNTT

제안된 프레임워크의 백본 네트워크는 스파이킹 버전의 VGG-9 구조로, 시간 정보 처리를 위한 학습 방식으로 BNTT(Backpropagation Through Time)를 채택하였다<sup>[11]</sup>. 본 구조는 스파이크 입력을 처리할 수 있는 합성곱 계층으로 구성되어 있으며, VGG 계열의 특징 추출 능력을 유지하면서도 SNN 특유의 에너지 효율성과 이벤트 기반 연산 구조를 통합하였다.

왼쪽부터 오른쪽까지 순서대로 Input Image, Spike Encoding, Spike Neural Network, Feature Fusion & Attention Mechanism, Multi-Layer Perceptron & Classifier 로 구성되어 있다.

### 2. 특징 융합과 어텐션 메커니즘(Feature Fusion & Attention Mechanism)

PolSAR 이미지의 고차원 편광 특성과 스펙클 노이즈 문제를 해결하기 위해, 본 연구에서는 특징 융합(feature fusion)과 CBAM(Convolutional Block Attention Module) 기반 어텐션 메커니즘을 적용하였다.

#### 2.1 특징 융합(Feature Fusion) 방법

특징 융합은 다양한 편광 채널의 정보를 하나의 통합된

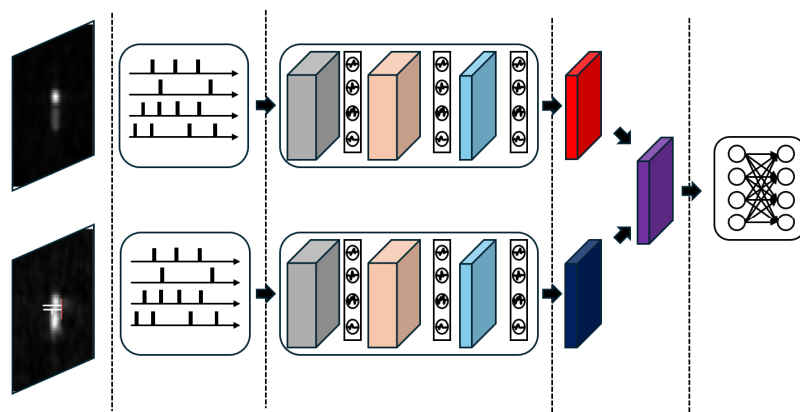


그림 2. 제안된 Spiking VGG-9 with BNTT 아키텍처

Fig. 2. Proposed spiking VGG-9 with BNTT architecture

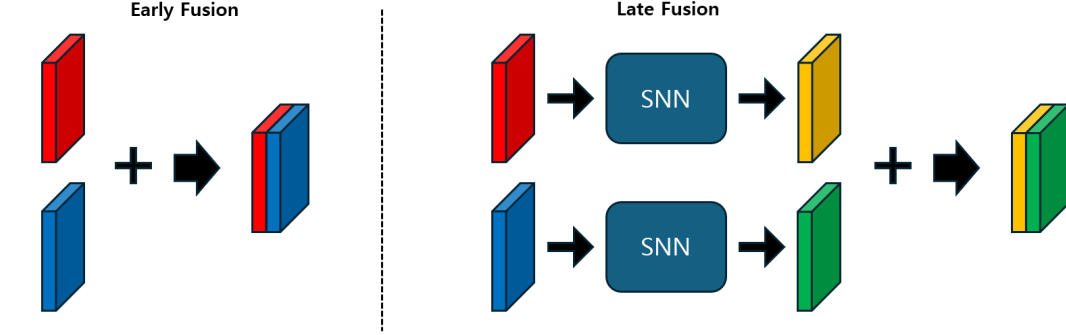


그림 3. 제안된 특징 융합 방법  
Fig. 3. Proposed feature fusion

표현으로 결합하는 방법으로, 본 연구에서는 Early Fusion과 Late Fusion 두 가지 전략을 비교하였다.

왼쪽 그림은 초기 융합, 오른쪽 그림은 후기 융합을 보여준다.

초기 융합(Early Fusion): 입력 단계에서 서로 다른 채널의 특징을 연결(concatenate)하여 하나의 통합 표현으로 구성한다. 이 방식은 연산 효율이 높고, 모델 전반에 걸쳐 융합된 정보를 학습할 수 있다.

대표적으로 그림 3과 같이 입력 특징  $F_A, F_B$ 를 연결하여 초기 융합 특징  $F$ 를 생성한다.

$$F_{early} = \text{concat}([F_A; F_B]) \quad (5)$$

위의 식 (5)에서  $F_{early}$ 는 연결된 통합 특징 맵(초기 융합 특징)이고,  $F_A, F_B$ 는 각각 입력 편광 채널 A와 B의 특징 맵이다. 위 과정을 거쳐서  $F_{early}$ 를 얻는다.

후기 융합(Late Fusion): 백본 네트워크를 통해 추출된 고차원 특징을 결합하는 방식으로, 연결(concatenate), 덧셈(add), 곱셈(multiply) 방식이 모두 고려되었다.

$$F_{late} = \text{concat}([SNN(F_A); SNN(F_B)]) \quad (6)$$

위의 식 (6)에서  $F_{late}$ 는 기 융합 결과로 생성된 통합 특징 맵이고,  $F_A, F_B$ 는 각각 입력 편광 채널 A와 B의 특징 맵,  $SNN$ 은 사용할 SNN 모델이다. 각 입력 특징  $F_A, F_B$

를 모델  $SNN$ 을 통과하여 얻은 특징들을 연결하여 그림 3의 오른쪽 그림과 같이 통합된 특징  $F_{late}$ 를 얻는다.

## 2.2 어텐션 메커니즘(Attention Mechanism) 방법

어텐션 메커니즘(Attention Mechanism)<sup>[12]</sup>: 본 연구에서는 CBAM(Convolutional Block Attention Module)을 네트워크에 통합하여, 중요한 특징에 가중치를 부여하는 방식으로 모델의 표현력을 강화하였다. CBAM은 채널 주의(channel attention)와 공간 주의(spatial attention) 모듈로 구성되며, 이를 순차적으로 적용한다.

채널 어텐션은 평균 풀링과 최대 풀링을 통해 채널 간 중요도를 계산한 후, 다층 퍼셉트론(MLP)을 통해 각 채널별 가중치를 산출한다.

공간 어텐션은 2D 특징 맵에 대해 평균/최대 풀링 후, 7×7 컨볼루션을 적용하여 중요한 위치 정보를 강조한다.

$$\begin{aligned} F' &= M_c(F) * F \\ F'' &= M_s(F') * F' \\ M_c(F) &= \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \\ M_s(F) &= \sigma(f^{7 \times 7}([AvgPool(rF); MaxPool(F)])) \end{aligned} \quad (7)$$

여기서  $F'$ 은 채널 주의 어텐션을 사용하여 통합된 특징이고,  $F''$ 은 공간 주의 어텐션을 사용하여 통합된 특징,  $M_c$ 는 채널 주의 어텐션(Channel Attention)을 적용한 특징 맵,  $M_s$ 는 공간 주의 어텐션(Spatial Attention)을 적용한 특징 맵이다.  $M_c$ 는 평균 풀링(avg pooling), 최대 풀링(max

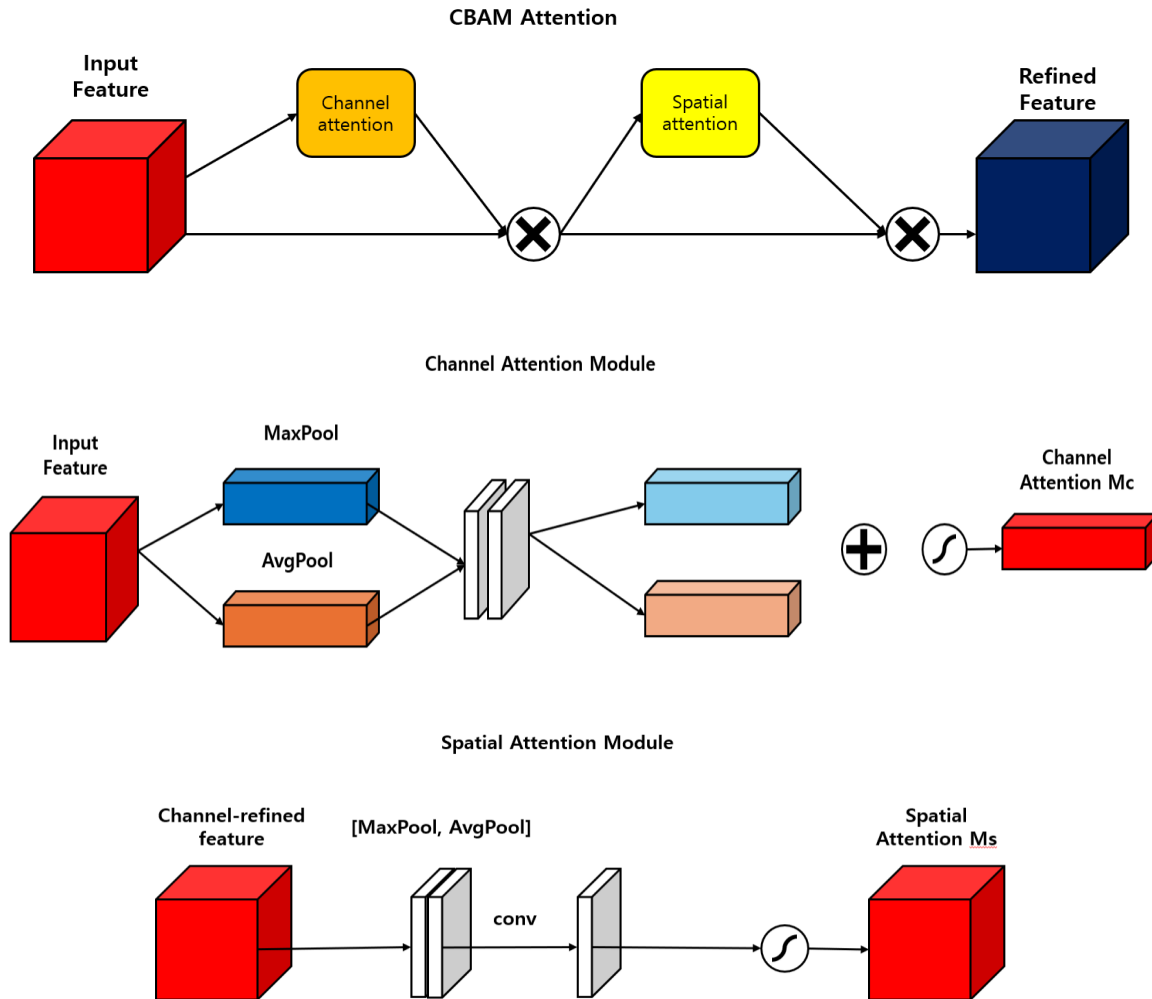


그림 4. 제안된 CBAM 어텐션 모듈  
Fig. 4. Proposed CBAM attention module

pooling), 다층 퍼셉트론(MLP)을 통해 채널별 중요도를 계산하고,  $M_s$ 는  $7 \times 7$  컨볼루션과 시그모이드 함수를 통해 공간적 중요도를 추출한다. MLP는 다층 퍼셉트론, AvgPool은 평균 풀링, MaxPool은 최대 풀링,  $\sigma$ 는 시그모이드 함수,  $f^{7 \times 7}$ 은  $7 \times 7$  풀링을 의미한다.

CBAM은 비교적 경량화된 구조로 다양한 네트워크에 쉽게 통합 가능하며, 본 연구에서는 각 시간 단계에서 합성곱 연산 전 위치에 삽입되어 적용되었다. 참고로 본 연구에서는 다른 어텐션 구조들(spatial-only, triplet 등)은 실험에 포함하지 않았으며, CBAM의 적용 효과만을 중심으로 분석

하였다.

### 3. 분류기 및 손실함수(Classifier and Loss Function)

최종적으로 융합된 특징은 의사결정을 위해 다층 퍼셉트론(MLP) 분류기에 전달된다. 네트워크 학습을 위해 다중 클래스 분류 작업에 적합한 크로스 엔트로피 손실함수(Cross-Entropy Loss)를 사용하였다. 이 손실함수는 올바른 클래스 확률을 강조하고 잘못된 예측에 페널티를 부여하여 네트워크를 효과적으로 최적화한다.

## IV. 실험 및 결과

### 1. 실험 설정

제안한 SNN 기반 PolSAR 이미지 분류 프레임워크의 성능을 평가하기 위해 OpenSARShip 2.0 데이터셋을 사용하여 다양한 융합 전략에 따른 성능 비교 실험을 수행하였다. 사용된 데이터는 HH(수직 편광), VV(수평 편광) 채널로 구성되어 있으며, 모델은 동일한 네트워크 구조(Spiking VGG-9 with BNNT)를 기반으로 설계되었다.

실험은 Python 3.8, PyTorch 1.12.1 환경에서 진행되었으며, 하드웨어는 NVIDIA GeForce RTX 3070 GPU를 사용하였다. 입력 이미지는 64×64 해상도(Image Size)이며, 시간 단계(Time Step)는 10, 배치 크기(Batch Size)는 16으로 설정하였다. 최적화 기법(Optimizer)으로는 SGD(Stochastic Gradient Descent)를 사용하였고, 학습률(Learning Rate)은

표 1. 실험 환경  
Table 1. Experiment Setting

	Detail
Model	Spiking VGG9 with BNNT
Number of Training Samples	7296
Number of Test Samples	1824
Image size	64
Batch Size	16
Num of Epochs	200
Learning Rate	0.01
Optimizer	SGD
Time Steps	10

0.01로 고정하였다. 전체 학습은 200 epoch 동안 수행하였으며, 훈련/검증 데이터(Train/Test data)는 8:2의 비율로 분할하였다.

본 실험에서는 정적 이미지를 입력으로 사용하기 때문에, Spike Encoder로는 Rate Coding을 적용하였다. 모든 실험은 같은 조건 하에서 진행되어 비교의 공정성을 확보하였다.

### 2. 실험 결과 및 분석

본 연구는 OpenSARShip 2.0 데이터셋의 두 편광 채널(수직 편광 HH, 수평 편광 VV)을 활용하여, 특징 융합과 어텐션 메커니즘이 PolSAR 이미지 분류 성능에 미치는 영향을 정량적으로 평가하였다. 실험은 다음과 같은 네 가지 전략으로 수행되었다. 1) 단일 채널 기반 모델(HH, VV) 2) 초기 융합(Early Fusion) 3) 후기 융합(Late Fusion - Add, Multiply, Concatenate) 4) 초기 융합에 어텐션 메커니즘(CBAM)을 적용한 방식

각 실험은 동일한 Spiking VGG-9 with BNNT 구조를 기반으로 진행되었으며, 공정한 비교를 위해 동일한 하이퍼파라미터와 학습 조건을 적용하였다. 주요 실험 결과는 다음 표 2와 같다.

정확도 분석 결과, 초기 융합(Early Fusion) 방식이 65.0767%로 가장 높은 성능을 보였으며, 단일 채널 VV 방식은 62.2807%로 가장 낮은 성능을 기록하였다. 융합 방식에 따른 정확도 순위는 초기 융합, 초기 융합 + 어텐션 메커니즘(Attention), 후기 융합(Late Fusion), 단일 채널 방식 순

표 2. 각 특징 융합 및 어텐션 메커니즘에 대한 실험 결과  
Table 2. Experimental results for each feature fusion and attention mechanism

	Accuracy	Precision	Recall	F1-Score
Early fusion	65.0767	0.6343	0.6343	0.6343
Single vh	62.6644	0.6312	0.6312	0.6312
Single vv	62.2807	0.5906	0.5906	0.5906
Late fusion add	63.0482	0.6968	0.6968	0.6968
Late fusion mul	63.5416	0.7156	0.7156	0.7156
Late fusion concatenate	64.0350	0.6375	0.6375	0.6375
Early CBAM Attention before convolution	64.2543	0.6468	0.6468	0.6468
Early CBAM Attention after convolution	63.5416	0.6343	0.6343	0.6343

으로 나타났다. 한편, 정밀도(Precision), 재현율(Recall), F1-Score는 데이터의 클래스 불균형 문제로 인해 micro 방식으로 계산되었으며, 이들 지표에서는 후기 융합 방식이 가장 우수한 값을 기록하였다. 전체적으로 볼 때, 단일 채널 방식 대비 초기 융합 방식은 정확도가 약 3% 향상되었으며, 후기 융합 방식은 정밀도 및 재현율에서 약 10% 이상의 성능 향상을 보였다. 이러한 결과는 특징 융합과 어텐션 메커니즘이 PolSAR 이미지 분석 성능 향상에 효과적으로 기여함을 실험적으로 입증한 것이다.

3. 계산 복잡도 및 모델 비교

본 연구에서는 제안된 Spiking Neural Network 기반 분류 모델의 연산 복잡도(FLOPs), 파라미터 수(Parameters), 평균 추론 시간(Latency)을 측정하여 기존 ANN 구조와 비교 분석하였다. 실험에는 VGG9(ANN 버전), Spiking VGG9 with BNNT, 다양한 특징 융합 방식 및 CBAM 어텐션 모듈이 적용된 모델이 포함되었다.

표 3에서 확인할 수 있듯이, SNN 구조는 ANN 구조에 비해 FLOPs 측면에서는 약 20배 이상 높은 연산량을 요구하지만, 이벤트 기반 연산 방식의 특성상 평균 추론 시간은 실시간 처리가 가능한 수준(8~14ms)으로 유지된다. 이는 시간 단계 단위의 계산을 분산 처리하는 SNN의 구조적 이점에 기인한다.

특히 CBAM 어텐션 모듈이나 Late Fusion 기법은 FLOPs와 파라미터 수를 다소 증가시키지만, 성능 향상과 비교할 때 연산 비용 대비 효율이 높은 것으로 분석된다. 예를 들어, Late Fusion(Multiply)의 경우 상대적으로 높은 정밀도 및 F1-Score를 기록하면서도 추론 지연은 13ms 수

준에 머무는 등 실용 가능성이 충분함을 보여준다.

이러한 결과는 SNN이 단순히 에너지 효율적인 구조를 넘어서, 고차원 PolSAR 이미지와 같이 복잡한 데이터를 처리하는 데 있어서도 실시간성과 성능을 양립할 수 있음을 의미한다. 따라서 SNN은 향후 원격탐사 분야에서 저전력 기반 고성능 분석 플랫폼으로 활용 가능성이 크다.

V. 결 론

본 연구에서는 PolSAR(Polarimetric Synthetic Aperture Radar) 이미지 분류를 위한 스파이킹 신경망(Spiking Neural Network, SNN) 기반 프레임워크를 제안하고, 다양한 특징 융합 기법 및 어텐션 메커니즘의 적용 효과를 정량적으로 비교·분석하였다. 다양한 편광 채널을 통합하는 특징 융합(feature fusion) 기법으로 초기 융합(Early Fusion)과 후기 융합(Late Fusion)을 각각 설계하여 성능을 평가하였으며, 어텐션 메커니즘으로는 CBAM(Convolutional Block Attention Module)을 적용하였다. 모든 실험은 동일한 Spiking VGG-9 with BNNT 구조와 조건 하에서 수행되어 비교의 공정성을 확보하였다.

실험 결과, 정확도 측면에서는 초기 융합 방식이 가장 우수한 성능(65.08%)을 보였으며, 정밀도 및 재현율에서는 후기 융합(Late Fusion)이 가장 높은 값을 기록하였다. 어텐션 메커니즘(CBAM)의 경우 일부 실험에서 성능 향상에 기여하였으나, 초기 융합 단독 방식보다 성능이 다소 낮게 나타나는 한계도 존재하였다.

이는 PolSAR 데이터셋의 클래스 불균형 문제와 더불어, 기존 CNN 기반 어텐션 구조가 스파이킹 신경망(SNN)의

표 3. 모델별 계산 복잡도 및 추론 성능 비교

Table 3. Comparison of computational complexity and inference performance by model

Model	SNN	FLOPs	Number of Parameters	Average Inference Time (ms)
VGG9 (ANN)	no	625M	17.93M	0.54
BNNT VGG9	yes	12.49B	18.52M	8.01
BNNT + CBAM (pre-conv)	yes	12.49B	18.53M	8.28
BNNT Late Fusion (Add)	yes	12.21B	18.53M	14.00
BNNT Late Fusion (Mul)	yes	12.21B	18.53M	13.17
BNNT Late Fusion (Concatenate)	yes	12.28B	35.31M	12.43



시간 기반 처리 방식과 완전히 호환되지 않았기 때문으로 해석된다. 이러한 결과는 PolSAR 이미지 분류 문제에 있어 SNN 구조의 적용 가능성과 실용성을 실험적으로 검증한 사례로, 에너지 효율성과 시공간 정보 처리 능력을 동시에 만족시킬 수 있는 대안이 될 수 있음을 시사한다. 또한, 고차원 편광 정보를 효과적으로 통합하는 특징 융합 기법이 성능 향상에 중요한 역할을 함을 입증하였다. 향후 연구에서는 다양한 어텐션 모듈 간의 비교 실험을 포함하고, PolSAR 데이터에 특화된 SNN 구조 설계를 통해 어텐션의 효율성과 구조적 적합성을 개선하는 방향으로 확장할 예정이다.

## 참 고 문 헌 (References)

- [1] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, "Efficient Processing of Deep Neural Networks: A Tutorial and Survey," *Proceedings of the IEEE*, vol. 105, no. 12, pp. 2295 - 2329, Dec. 2017.  
doi: <https://doi.org/10.1109/JPROC.2017.2761740>
- [2] G. Datta, S. Kundu, and P. A. Beerel, "Training Energy-Efficient Deep Spiking Neural Networks with Single-Spike Hybrid Input Encoding," in *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, Jul. 2021, pp. 1 - 8.  
doi: <https://doi.org/10.1109/IJCNN52387.2021.9534306>
- [3] T. Zhang, T. Zeng, and X. Zhang, Eds., *Synthetic Aperture Radar (SAR) Meets Deep Learning*. MDPI, 2023.  
doi: <https://doi.org/10.3390/books978-3-0365-6383-1>
- [4] N. Wang, W. Jin, H. Bi, C. Xu, and J. Gao, "A Survey on Deep Learning for Few-Shot PolSAR Image Classification," *Remote Sensing*, vol. 16, no. 24, p. 4632, Dec. 2024.  
doi: <https://doi.org/10.3390/rs16244632>
- [5] X. Jiang, H. Xie, Z. Lu, and J. Hu, "Energy-Efficient and High-Performance Ship Classification Strategy Based on Siamese Spiking Neural Network in Dual-Polarized SAR Images," *Remote Sensing*, vol. 15, no. 20, p. 4966, Oct. 2023.  
doi: <https://doi.org/10.3390/rs15204966>
- [6] A. Hodgkin and A. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *Bulletin of Mathematical Biology*, vol. 52, no. 1 - 2, pp. 25 - 71, 1990.  
doi: [https://doi.org/10.1016/S0092-8240\(05\)80004-7](https://doi.org/10.1016/S0092-8240(05)80004-7)
- [7] A. P. Georgopoulos, "Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems," *The Quarterly Review of Biology*, vol. 79, no. 1, pp. 113 - 113, Mar. 2004.  
doi: <https://doi.org/10.1086/421681>
- [8] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, and A. Maida, "Deep learning in spiking neural networks," *Neural Networks*, vol. 111, pp. 47 - 63, Mar. 2019.  
doi: <https://doi.org/10.1016/j.neunet.2018.12.002>
- [9] K. Yamazaki, V.-K. Vo-Ho, D. Bulsara, and N. Le, "Spiking Neural Networks and Their Applications: A Review," *Brain Sciences*, vol. 12, no. 7, p. 863, Jun. 2022.  
doi: <https://doi.org/10.3390/brainsci12070863>
- [10] Y. Kim, H. Park, A. Moitra, A. Bhattacharjee, Y. Venkatesha, and P. Panda, "Rate Coding or Direct Coding: Which One Is Better for Accurate, Robust, and Energy-Efficient Spiking Neural Networks?," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, May 2022, pp. 71 - 75.  
doi: <https://doi.org/10.1109/ICASSP43922.2022.9747906>
- [11] H. Kim and P. Panda, "Revisiting Batch Normalization for Training Low-Latency Deep Spiking Neural Networks From Scratch," *Frontiers in Neuroscience*, vol. 15, Dec. 2021.  
doi: <https://doi.org/10.3389/fnins.2021.773954>
- [12] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," *Computer Vision - ECCV 2018*, pp. 3 - 19, 2018.  
doi: [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)

## 저 자 소 개



### 김 선 옥

- 2009년 ~ 2014년 : 연세대학교 전기전자공학과 학사
- 2014년 ~ 2019년 : 연세대학교 전기전자공학과 박사
- 2019년 ~ 2021년 : 연세대학교 박사후연구원
- 2021년 ~ 현재 : 한국항공대학교 소프트웨어학과 조교수
- ORCID : <https://orcid.org/0000-0002-9665-4214>
- 주관심분야 : 컴퓨터비전, 인공지능, 3차원 영상처리

---

저 자 소 개



**최 환 효**

- 2024년 2월 : 한국항공대학교 소프트웨어학과 학사
- 2024년 3월 ~ 현재 : 한국항공대학교 인공지능학과 석사과정
- ORCID : <https://orcid.org/0009-0003-3174-386X>
- 주관심분야 : 컴퓨터비전, 인공지능