TinyWM: 임베디드 환경에서 합성곱 신경망 기반 워터마킹 모델 최적화 및 성능 평가

허남정[1], 장우현[2], 김연재[2], 허재원[2], 박기웅[3,†]
[1]세종대학교 시스템보안연구실 (지능형드론 융합전공) 석사과정
[2]LIG 넥스원 연구원
[3,†] 세종대학교 정보보호학과 교수
uskawjdu@gmail.com, woohyun.jang2@lignex1.com, yeonjae.kim@lignex1.com,
jaewon.huh@lignex1.com, woongbak@sejong.ac.kr

TinyWM: Optimization and Performance Evaluation of CNN-Based Watermarking Model in Embedded Environments

Nam-Jung Heo[1], Woo-Hyun Jang[2], Yeon-Jae Kim[2],
Jae-Won Huh[2], Ki-Woong Park[3,†]
[1]SysCore Lab. (Convergence Engineering for Intelligent Drone),
Sejong University
[2]LIG NEX1 Researcher
[3,†] Dept. of Computer and Information Security, Sejong University

요 약

UAV는 다양한 분야에서 사용되면서, 점차 활용 범위가 증가하고 있다. UAV가 임무를 수행하면서 획득한 이미지에 대한 재산권을 보호할 수 있는 기술이 필요하다. 디지털 워터마킹 기술은 UAV가 촬영한 이미지에 재산권을 보호할 수 있는 워터마크를 삽입한다. 삽입된 워터마크에 대한 비가시성 및 강건성 평가를 통해 워터마크의 안정성을 평가할 수 있다. 또한 UAV는 임베디드 환경에서 동작할 수 있는 저전력 기술이 요구된다. 본 연구에서 OpenMV Cam RT1062에서 합성곱 신경망 기반의 워터마킹 기술을 설계하고 그에 따른 비가시성 및 강건성을 평가한다. 그 결과 충분한 비가시성 확보를 얻을 수 있었으며, 삽입된 워터마크의 강건성을 확인할 수 있다.

1. 서론

UAV(Unmanned Aerial Vehicle)는 해양 경비, 긴급 구조, 환경 모니터링 등 다양한 분야에서 사용될 수 있다. UAV가 임무를 수행하면서 획득한 이미지 데이터는 중요한 자산이며 이를 보호할 수 있는 기술로서 디지털 워터마킹 기술이 사용된다 [1].

UAV의 광범위한 활용은 기술적 진보를 수반하지만, 배터리 사용 시간의 제한이라는 문제점이 있다. 이러한 이유로 UAV는 하나의 임베디드 기기로 저전력 고효율의 데이터 처리가 필요로 하다.

본 연구에서는 UAV의 딥러닝을 활용한 디지털 워터마크 기술 도입 시, 저전력 시스템의 기술적 한계에 도전한다. 딥러닝 기반의 디지털 워터마크 기술은 강인한 비가시성과 강인성을 확보할 수 있다.

OpenMV Cam RT1062은 임베디드 환경에서 인공지능을 위한 저전력 카메라 모듈로 TinyML 기반의 시스템을 구현에 사용된다. 본 연구에서는 해당 카메라 모듈에 적합하게 신경망을 최적화하고 워터마크의 비가시성 및 강건성 평가에 초점을 맞춘다.

본 연구의 기여는 다음과 같다:

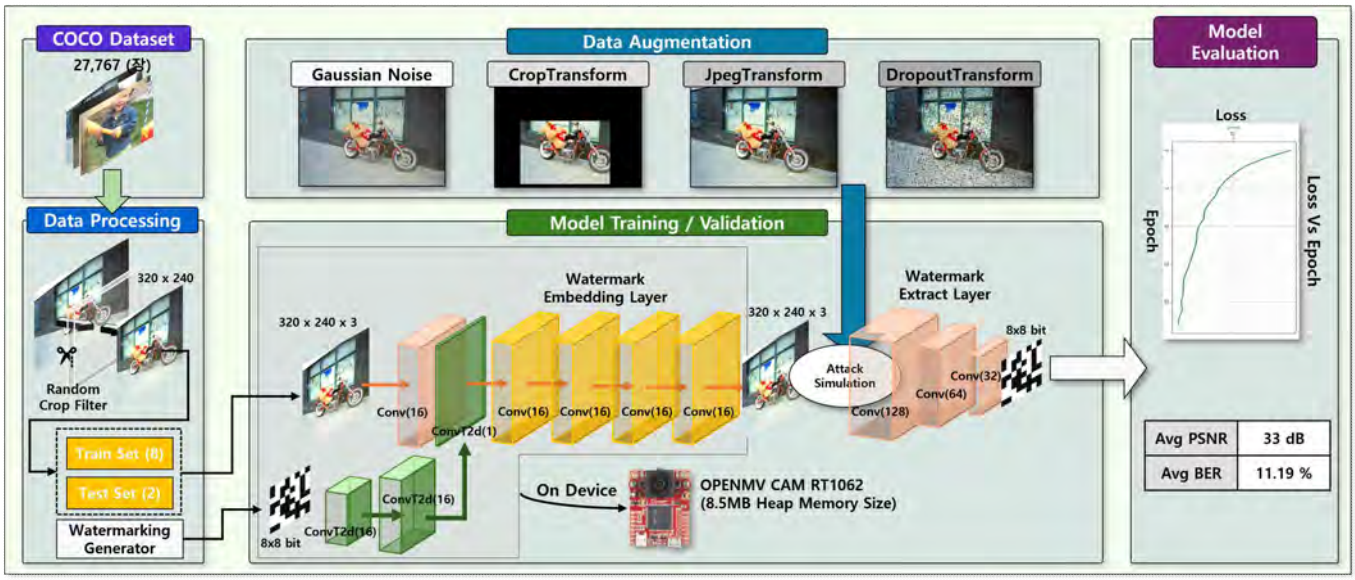
· OpenMV CAM 모듈을 사용하여 TinyML 기반 디지털 워터마킹 기술을 구현 방법을 제안한다.

· 구현된 워터마킹 기술의 비가시성 및 강건성에 평가를 수행한다.

본 논문의 구성은 다음과 같다. 제2장에서는 이미지 기반의 머신러닝 관련 연구를 정리하고 연구의 기여를 명확하게 한다. 제3장에서는 연구에서 제한한 모델 학습 및 아키텍처를 기술한다. 제4장에서는 구현된 워터마킹 기술의 강건성을 시각화한다.

2. 관련 연구

본 장에서는 <표1>과 같이 이미지 기반의 머신러닝 모델 연구 들을 인증, 분류, 탐지 3가지 분류로 구분한다. 이 중 TinyML 기반의 연구들은 식별한다. 본 연구는 이전 연구들에서 다루지 않은 TinyML 솔루션을 적용한 디지털 워터마킹 기술을



(그림 1) CNN 합성곱 신경만 워터마킹 모델 학습 아키텍처

제안한다. 이전 연구의 합성곱 기반의 신경망 모델은 크기가 커서 임베디드 장치에 직접 실행하기 어려운 한계점이 있다.

<표 1> Research on Image-Based Machine Learning

Paper	인증	분류	탐지	TinyML
Our	√	X	X	√
[2]	√	X	X	X
[3]	√	X	X	X
[4]	X	√	X	√
[5]	X	√	X	√
[6]	X	X	√	√

J.-E. Lee et al. [2]와 E. Rebahi et al. [3] 연구에서는 합성곱 신경망 기반의 워터마킹 모델을 제안한다. 인코더에서 워터마크와 원본 이미지 두 입력을 받아 디코더에서 워터마크가 삽입된 이미지를 출력한다. 추출기 모델을 사용하여 워터마크가 삽입된 이미지로부터 워터마크를 추출할 수 있다. 이러한 워터마킹 모델은 강인성, 비가시성의 특징을 지닌다.

H. Ai et al, [4] 연구에서는 LBP 특징 알고리즘 및 SVM 분류 인식을 사용하여 실시간 얼굴 인식을 구현한다. OpenMV Cam H7 카메라 모듈을 사용하여 대상 이미지를 획득하고 추론을 수행한다.

L. Santoro et al. [5] 연구에서는 OpenMV Cam H7 Plus을 활용하여 지정된 위치에 UAV를 착륙시키기 위한 비전 머신을 설계하고 성능을 평가한다. 연구 결과 17g에 소형 임베디드 기기인 OpenMV Cam을 활용하여 안정적으로 착륙할 수 있음을 보인다.

W. Li et al [6] 연구에서는 YOLO 알고리즘 기반의 탐지 시스템과 UAV 제어 시스템 통합을 제안한다. OpenMV Cam을 활용하여 대상을 식별하고 영역 정보를 획득하여 대상 위치를 파악할 수 있다.

이러한 관련 연구들은 OpenMV Cam 모듈이 UAV의 통합 가능성을 보인다.

3. 연구 방법

3.1. 모델 학습 환경

본 연구에서 학습에 사용된 환경은 다음과 같다:

- **CPU:** AMD Ryzen 5 PRO 4650G
- **GPU:** RTX 4060ti 16gb
- **RAM:** DDR4 32GB
- **Storatge:** m.2 ssd 2 TB

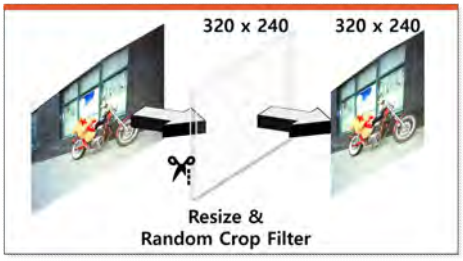
컴퓨터 시스템 자원의 한계로 학습 데이터 세트 중 일부만 학습에 사용한다. 평가 데이터를 제외한 학습에만 사용된 데이터는 22213장이며, 하나의 Epoch에는 대략 5분 정도의 시간이 소요된다.

3.2. 학습 데이터 구성 및 데이터 전처리

본 연구에서는 COCO(Common Objects in Context)에서 제공하는 “2017 Unlabeled images [123K/19GB]” 데이터를 학습에 사용한다 [7]. COCO 데이터 세트는 일상에서 볼 수 있는 객체들을 Object Detection 학습에 사용하기 위해서 공개된 데이터 세트이다. 일상적 객체를 촬영한 데이터는 UAV가 촬영하는 환경과 유사하기 때문에 학습 데이터로 선정한다. 해당 데이터 세트에는 다양한 사이즈와 비율의 컬러 및 흑백 사진이 섞여 있어 전처리 작업이 필요하다. [그림 1]은 전처리 및 전체

모델을 학습하고 평가하는 과정을 설명한다.

[그림 2]는 학습 데이터의 전처리 단계를 설명한다. 사진의 가로와 세로 중 더 가까운 크기에 맞추어 이미지를 Resize를 적용한 후에 320x240의 크기로 랜덤하게 Crop 한다.



(그림 2) Data preprocessing

모든 학습 데이터는 -1~+1로 Normalization을 수행한다. 워터마크 데이터는 랜덤하게 -1 또는 1의 8x8 데이터로 생성되며, 생성 알고리즘 및 학습 모델에 대해서는 깃허브에 공개한다 [8].

전체 데이터 세트를 8대 2로 나누어 학습 데이터 22213장과 테스트 데이터 5554장으로 설정한다. 모델이 학습할 때, 매 Epoch마다 테스트 데이터에 대해서 Loss를 계산하고 손실 평가를 수행한다.

3.3. 모델 훈련 및 데이터 증강

본 연구에서는 워터마크의 강건성 확보를 위하여 4가지 데이터 증강 기법을 랜덤으로 적용한다. 워터마크의 존재를 원치 않는 악의적인 행동자는 이미지를 일부 자르거나(Crop). 화질을 저하하는 등의 공격을 수행할 수 있다. 본 연구에서는 'Gaussian Noise', 'Crop Transform', 'Jpeg Transform', 'Dropout Pixels Transform'에 대해서 네 가지 데이터 증강 기법을 적용한다. 그리고 4장에서 실제 카메라로 획득한 데이터를 대상으로 악의적인 공격에 워터마크가 적절하게 추출이 되는지 강건성을 확인한다.

OpenMV Cam RT1062 카메라는 8.5MB의 Heap 메모리를 지원한다. 이 크기는 작은 수치이기 때문에, 모델을 설계할 때는 모델 크기의 제약을 고려하여 설계해야 한다. 본 연구에서는 최소한의 파라미터만으로 '워터마크 임베딩 레이어'와 '워터마크 추출 레이어'를 구현하기 위해서 Convolution Layer만을 사용한다. 특히 '워터마크 임베딩 레이어'는 카메라 모듈 내부에서 동작하기 때문에 최소한 필터 수(filters=16)만으로 구성한다.

본 연구에서는 이전 연구에서 사용된 파라미터 값을 그대로 사용하여 [2], 학습을 수행했을 때 25 Epoch 이하로는 테스트 데이터의 Loss 값이 변화가 거의 없는 것을 확인하고 학습을 중단한다.

3.4. 모델 경량화

훈련된 모델은 Tensorflow Lite 변환 과정을 거쳐서 추론 모델이 생성된다. 원래 생성된 모델의 크기는 62KB 크기이다. OpenMV Cam에서는 신경망 연산을 Heap에 수행하기 때문에, 하드웨어의 제약을 고려해야 한다 [5]. 따라서 전체 정수의 양자화를 수행하여 네트워크 가중치와 활성화를 부동 소수점 대신 정수형 8비트로 표현한다. 메모리 요구사항을 줄이고 신경망의 실행 속도를 향상시킨다. 대신 성능의 저하가 발생할 수 있다. 모델의 양자화 결과 추론 모델의 크기는 28KB으로 줄어든다.

4. 성능 평가

본 연구에서는 훈련된 모델의 비가시성 및 강건성을 평가하기 위하여 테스트 데이터에 대해서 PSNR (Peak Signal-to-Noise Ratio) 및 BER(Bit Error Rate) 계산한다. <표 2>에서 전체 평균 PSNR 및 BER을 확인할 수 있다.

<표 2> Evaluation Table

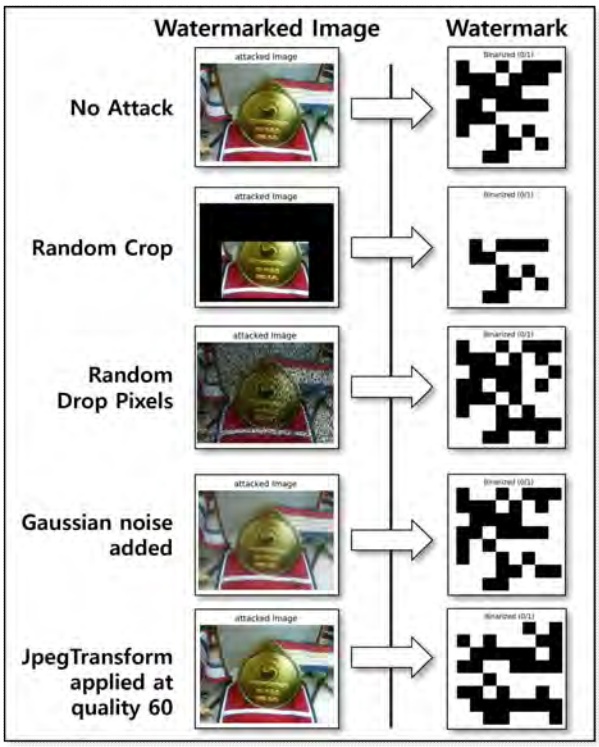
Avg PSNR	33 dB
Avg BER	11.194%
Execution Time	3 seconds

PSNR 이란 워터마크가 삽입된 이미지와 원본 이미지 간의 유사성 즉, 비가시성을 평가한 지표로 그 값이 30dB 초과하면 높은 비가시성을 보장한다 [3]. 본 연구에서는 테스트에 사용된 전체 데이터에 대해서 워터마크와 원본 이미지 간의 평균 PSNR 지표를 계산한다. 그 결과 33dB으로 비가시성을 확보한다.

BER(Bit Error Rate)이란 원본 워터마크 비트와 추출된 워터마크의 비트열을 차이를 계산한다. 본 연구에서는 악의적인 공격(데이터 증강 기법)이 적용된 이미지로부터 워터마크 정보를 추출할 때, 원본 워터마크와 차이에 대해서 평가한다.

이전에 다른 연구에 따르면 BER의 수치가 10% 이하일 때, 워터마크의 보호를 받을 수 있다 [2]. 전체 테스트 데이터에 대하여 랜덤으로 데이터 증강 기법을 적용하고 BER을 계산한 결과 전체 평균 11.194% 수치를 보인다. OpenMV Cam으로 촬영한

이미지에 워터마크가 삽입된 상황에서 악의적인 공격(데이터 증강 기법)이 적용될 때, 워터마크 추출의 패턴 변화를 확인한다. [그림 4]는 워터마크를 적용된 원본 이미지와 워터마크 추출 정보에 대해서 각각의 공격 상황에 추출된 워터마크의 패턴 변화를 시각화한다.



(그림 3) 카메라 모듈에서 획득한 사진의 강건성 비교

OpenMV Cam 모듈이 추론을 수행할 때 전력 소비를 측정한다. 그 결과 1.131W(4.71V/0.24A)를 소비하며, 이는 라즈베리 파이의 공식 사양인 15W(5V/3A) 전력 소비와 비교했을 때 작은 수치이다. 또한 모델의 추론 시간은 3초 소요된다.

5. 결론

본 논문의 실험 결과 임베디드 환경에서 비가시성 및 강건성을 갖춘 워터마킹 기술을 확인할 수 있다. 워터마킹 기술은 UAV를 사용하는 사용자와 기업의 중요한 지적 재산권 보호를 수행할 수 있을 것이다. 이러한 UAV는 저전력 시스템인 점을 고려하여 적합한 모델 설계 및 학습이 중요하다. 또한 실제 UAV와 통합하여 제안된 구조의 효과성을 입증할 수 있을 것이라 기대한다.

Acknowledge

이 논문은 2022년 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임 (KRIT-CT-22-051)

참고문헌

[1] 허남정 and 박기웅, “군사 드론작전운영시데이터 보호를위한 멀티모달 기반 디지털워터마킹분류체계”, 한국디지털포렌식학회 하계학술대회, 2025.

[2] J.-E. Lee, Y.-H. Seo and D.-W. Kim, “Convolutional neural network-based digital image watermarking adaptive to the resolution of image and watermark”, Applied Sciences, vol.10, no.19, 2020.

[3] E. Rebahi, M. Hemis and B. Boudraa, “Image Watermarking Technique Using Convolutional Autoencoder”, 2023 International Conference on Advances in Electronics, Control and Communication Systems (ICAECCS), BLIDA, Algeria, pp.1-6, 2023.

[4] H. Ai, H. Xia, W. Chen and B. Yang, “Face Tracking Sign-in System Based on LBP Feature Algorithm″, 2020 Chinese Automation Congress (CAC), Shanghai, China, pp.2257-2262, 2020.

[5] L. Santoro et al, “A Plug-and-Play TinyML-based Vision System for Drone Automatic Landing”, 2023 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT), pp.293-298, 2023.

[6] W. Li, W. Huang, Y. Li and L. Wang, “Design of Joint Control System for Target Detection UAV and Ground Vehicle Based on YOLO Algorithm”, In Proceedings of the 2023 11th International Conference on Computer and Communications Management (ICCCM '23), pp.70-75, 2023.

[7] COCO Dataset, Available: https://cocodataset.org/#download.

[8] WaterMarking Model github, Available: https://github.com/CherryPichu/WaterMark_Model.