



딥 러닝을 이용한 비디오 카메라 모델 판별 시스템

김동현*¹, 이수현*², 이해연**

Video Camera Model Identification System Using Deep Learning

Dong-Hyun Kim*¹, Soo-Hyeon Lee*², and Hae-Yeoun Lee**

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2017R1D1A1B03030432).

요 약

현대 사회에서 영상 정보 통신 기술이 발전함에 따라서 영상 획득 및 대량 생산 기술도 급속히 발전하였지만 이를 이용한 범죄도 증가하여 범죄 예방을 위한 법의학 연구가 진행되고 있다. 영상 획득 장치에 대한 판별 기술은 많이 연구되었지만, 그 분야가 영상으로 한정되어 있다. 본 논문에서는 영상이 아닌 동영상에 대한 카메라 모델의 판별 기법을 제안한다. 기존의 영상을 학습한 모델을 사용하여 동영상의 프레임을 분석하였고, 동영상의 프레임 특성을 활용한 학습과 분석을 통하여 P 프레임을 활용한 모델의 우수성을 보였다. 이를 이용하여 다수결 기반 판별 알고리즘을 적용한 동영상에 대한 카메라 모델 판별 시스템을 제안하였다. 실험에서는 5개 비디오 카메라 모델을 이용하여 분석을 하였고, 각각의 프레임 판별에 대해 최대 96.18% 정확도를 얻었으며, 비디오 카메라 모델 판별 시스템은 각 카메라 모델에 대하여 100% 판별률을 달성하였다.

Abstract

With the development of imaging information communication technology in modern society, imaging acquisition and mass production technology have developed rapidly. However, crime rates using these technology are increased and forensic studies are conducted to prevent it. Identification techniques for image acquisition devices are studied a lot, but the field is limited to images. In this paper, camera model identification technique for video, not image is proposed. We analyzed video frames using the trained model with images. Through training and analysis by considering the frame characteristics of video, we showed the superiority of the model using the P frame. Then, we presented a video camera model identification system by applying a majority-based decision algorithm. In the experiment using 5 video camera models, we obtained maximum 96.18% accuracy for each frame identification and the proposed video camera model identification system achieved 100% identification rate for each camera model.

Keywords

video camera model identification, forensic, deep learning, convolutional neural network

* 금오공과대학교 소프트웨어공학과

- ORCID¹: <https://orcid.org/0000-0002-0693-431X>

- ORCID²: <https://orcid.org/0000-0002-3372-5660>

** 금오공과대학교 컴퓨터소프트웨어공학과 교수
(교신저자)

- ORCID: <http://orcid.org/0000-0002-6081-1492>

• Received: May 27, 2019, Revised: Jun. 29, 2019, Accepted: Jul. 02, 2019

• Corresponding Author: Hae-Yeoun Lee

Dept. of Computer Software Engineering, Kumoh National Institute of
Technology, Korea

Tel.: +82-54-458-7548, Email: haeyeoun.lee@kumoh.ac.kr

1. 서 론

현대 사회의 영상 정보 통신 기술은 매우 급속도로 발전해오고 있다. 이러한 기술의 발달과 함께 멀티미디어 획득 장치의 성능과 영상의 대량 생산 기술 또한 빠르게 발전해왔다. 멀티미디어 획득 장치의 고도화는 기존보다 고품질의 영상을 얻을 수 있게 만들었고, 정지 영상을 넘어 동영상을 획득할 수 있게 되었다. 또한, 통신 기술의 비약적인 발전으로 SNS(Social Network Service)라 불리는 매체를 통해 영상 및 동영상 데이터는 복사, 확대, 변형되어 전 세계로 퍼져 나가게 되었다.

기술의 발전은 정보의 다양화, 확산, 그리고 공유를 통해 현대 사회의 삶의 질을 높여 왔다. 하지만 이러한 발전은 긍정적인 변화뿐만 아니라 부정적인 변화도 이끌어 왔다. 발전된 멀티미디어 획득 기술로 영상과 동영상을 누구나 손쉽게 얻을 수 있게 되었고, 이를 이용한 범죄가 많이 증가하게 되었다.

이러한 범죄를 예방 및 차단할 뿐만 아니라 실제 발생한 범죄에 대한 수사에서 획득한 범죄 증거들을 법정 증거로 활용하기 위해 멀티미디어 포렌식 기술의 필요성과 중요성이 증가 되고 있다. 멀티미디어 포렌식 기술은 미국, 유럽 등의 선진국에서는 이미 활발히 연구되어 오고 있으며, 많은 성과를 이루어 내고 있다. 이 중에서 영상을 이용한 카메라 장치 판별은 많은 연구가 진행되고 실제 적용 가능한 수준으로 유의미한 발전을 이루어 왔다. 또한, 최근에 많은 주목을 받는 딥 러닝 기술을 접목한 연구가 많이 진행되고 있으며 뛰어난 성과를 거두고 있다.

하지만 기존의 단순한 영상뿐 아니라 동영상까지 멀티미디어 콘텐츠로 활용되고 있으며, 영상에 대한 카메라 모델 판별 연구는 많이 진행되었으나 동영상에 대한 카메라 장치 판별 기술에 관한 연구는 많이 부족한 실정이다.

제안하는 알고리즘은 멀티미디어 포렌식 기술 중 하나인 카메라 모델 판별을 영상이 아닌 동영상을 대상으로 하였다. 기본적으로 카메라 장치 판별은 카메라가 대상물로부터 영상을 획득할 때, 렌즈의 광학 센서를 이용하고 이러한 센서는 카메라 장치마다 서로 달라 장치별 고유한 노이즈 특성을 가지

게 하는 점을 이용한다. 이러한 센서의 노이즈 성분을 추출하고 카메라를 판별하기 위해 딥 러닝 기술을 적용하는 방향으로 실험을 설계, 수행한다.

본 논문의 구성은 다음과 같다. 2장에서는 카메라 장치 판별에 관한 국내외의 연구 동향과 딥 러닝 기술에 대한 설명을 기술하고, 3장에서는 기존의 영상을 이용한 모델의 한계를 분석하고, 제안하는 알고리즘을 기술한다. 4장에서 실험 결과에 대해 제시하고, 5장에서 결론을 맺는다.

II. 관련 연구

2.1 카메라 장치 판별 연구

카메라 장치 판별에 관한 연구는 범행을 입증하기 위한 포렌식의 관점에서 매우 중요하기에 많은 연구가 국내외에서 진행되고 있다.

Lukas et. al은 카메라에서 영상을 획득할 때 사용하게 되는 광학 센서의 필연적인 비균일성에 의해 생기는 노이즈 성분인 PRNU(Photo Response Non Uniformity)를 추출하여 이를 이용, 통계학적인 유사도를 이용하여 카메라 장치를 판별하였다[1]. Choi et. al은 카메라 렌즈의 수차에 의한 왜곡을 이용하여 카메라를 판별하는 연구를 수행하였다[2]. Bayram et. al은 CFA 보간을 이용하여 카메라 장치를 판별하였다[3]. 또한, Lee는 다양한 SPN(Sensor Pattern Noise)를 이용하여 카메라 장치 판별 연구를 수행했다[4]. Lee et. al은 모폴로지 필터링 기반의 SPN을 통한 동영상 획득 장치 판별 연구를 수행하였다[5].

2.2 딥 러닝 기술

딥 러닝은 현재 가장 주목받고 있는 기술로 다양한 연구를 통해 빠르게 발전하고 있으며, 광범위한 분야에 활용되고 있다. 딥 러닝은 인공신경망을 기반으로 하여 발전된 기술이다. 기존에 연구되었던 인공신경망이 충분히 많은 계층을 가진다면 단순한 문제뿐 아니라 복잡한 문제도 해결할 수 있음이 밝혀짐으로써 다양한 분야와의 접목을 위한 연구가 진행되고 있다. 또한, 단순히 쌓은 형태가 아니라

이용하는 데이터와 목적에 따라 다양하게 변형된 구조들이 존재한다. 단순하게 완전 연결 계층들을 깊게 쌓은 DNN(Fully Connected Deep Neural Network), 영상과 음성 등에 높은 성능을 보이는 컨볼루션을 이용한 CNN(Convolution Neural Network), 순차적인 시계열 데이터에 적합한 것으로 알려진 RNN(Recurrent Neural Network) 뿐만 아니라 LSTM, RBM, DBN, GAN, ResNet, DenseNet 등 이 존재한다[6]-[9].

본 논문에서는 영상 데이터에 대해 높은 성능을 보이는 것으로 검증된 CNN을 활용하여 동영상 카메라 장치 판별에 이용하였다. CNN은 컨볼루션 연산을 초기 계층에 추가한 인공 신경망 모델로 입력 계층, 컨볼루션 계층, 완전 연결 계층, 출력 계층 등으로 이루어졌다.

2.3 딥 러닝을 이용한 카메라 장치 판별 연구

현재 많은 연구팀은 딥 러닝 기술을 이용하여 이를 다양하게 변형하고 적용하여 카메라 모델 판별에 이용 중이다. Tuama et. al은 기존의 PRNU를 이용하는 방법과 딥 러닝을 이용한 방법을 비교하였다. CNN을 이용하였고, High-Pass 필터와 Wiener 필터를 이용하여 둘을 비교하였다. 다양한 카메라 모델과 장치를 이용해 피사체를 촬영한 Dresden Image

Database와 개인용 카메라들로 찍은 영상들을 이용하여 성능을 평가하고 비교하였다[10]. Wang et. al은 LBP(Local Binary Pattern)라는 밝기변화에 무관한 특징을 이용하여 카메라 장치 판별 연구를 수행하였다[11]. 또한, Kamal et. al은 State-of-the-art에 해당하는 딥 러닝 모델인 DenseNet을 이용하여 연구를 수행하고 있다[12].

III. 비디오 카메라 모델 판별 시스템

본 장에서는 제안하는 비디오 카메라 모델 판별 시스템에 대하여 설명한다. 기존의 딥 러닝 모델은 하나의 영상을 입력으로 하여 영상의 클래스를 분류한다. 본 시스템에서는 동영상이 영상의 집합임을 이용하여 판별 시스템을 구성하였다. 먼저 3.1절에서 카메라 모델 판별 알고리즘을 소개하고, 3.2절에서는 딥 러닝 모델을 소개한다. 3.3절에서는 동영상 프레임 특성에 대한 분석 결과를 설명한다.

3.1 비디오 카메라 모델 판별 알고리즘

본 절에서는 제안하는 비디오 카메라 모델 판별 알고리즘에 대해 설명한다. 그림 1은 제안하는 비디오 카메라 모델 판별 알고리즘을 나타낸 것으로 크게 학습 과정과 판별 과정으로 나누어 처리된다.

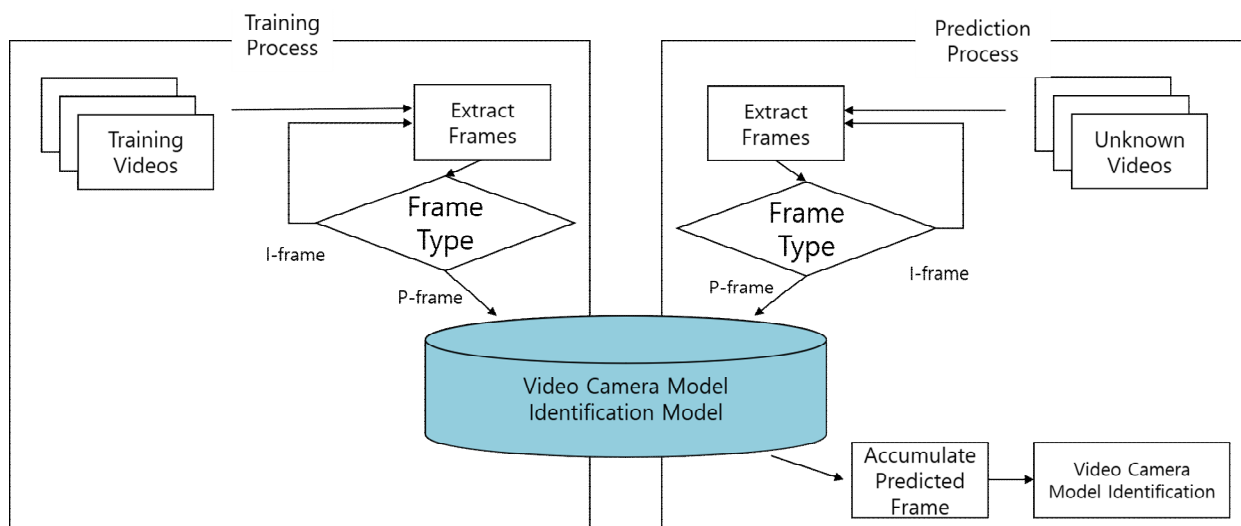


그림 1. 비디오 카메라 모델 판별 시스템
Fig. 1. Video camera model identification system

4 딥 러닝을 이용한 비디오 카메라 모델 판별 시스템

두 과정 모두 동영상에서 프레임을 추출하고 이를 실제 딥 러닝 모델에 입력으로 활용하게 된다. 본 모델에서는 프레임을 추출한 후 P 프레임만을 선택하여 학습과 판별에 이용한다. 판별 과정에서는 추가로 각각의 동영상에서 얻은 프레임들의 결과를 모두 합산하여, 가장 높은 빈도로 추측된 클래스로 판별한다. 상기의 과정에서 비디오 카메라 판별을 위한 딥 러닝 모델은 3.2절에서 설명하며, I 프레임을 사용하지 않고 P 프레임만 사용하여 설계한 이유는 3.3절에서 기술한다.

3.2 비디오 카메라 판별을 위한 딥 러닝 모델

동영상은 GoP(Group of Pictures)라 부르는 프레임들의 집합으로 구성되어 있다. 이들 프레임은 종류에 따라 크게 I(Intra), P(Predicted), B(Bidirectional predicted) 프레임으로 나눌 수 있다. 무손실 압축된 영상은 다양한 코덱을 통해 프레임들로 인코딩되고 실제 재생 시에 각각의 프레임은 디코딩 과정을 거쳐 영상으로 복원된다. 따라서 동영상을 구성하는 프레임은 디코딩 과정을 거쳐 영상으로 변환할 수 있으므로 각각의 프레임을 영상으로 보아 기존의 카메라 모델 판별 기술을 이용할 수 있다. 그림 2는 기존 카메라 장치 판별 연구에서 우수한 성능을 나타낸 CNN 모델의 계층 구성을 나타낸다.

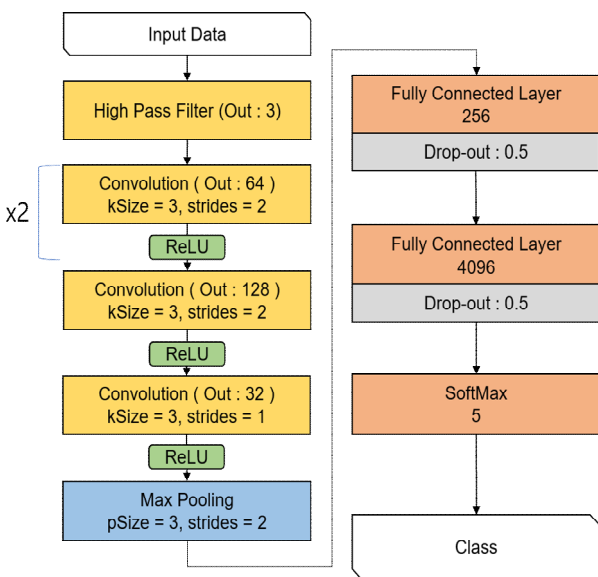


그림 2. 컨볼루션 뉴럴 네트워크 계층 구조
Fig. 2. CNN layers structure

먼저 입력된 영상을 HPF(High-Pass-Filter)를 통과시킨 후에 컨볼루션 계층에 입력한다. 3개의 컨볼루션 계층을 거친 후에는 Max-Pooling을 이용하여 Feature Map의 크기를 줄인다. 이후 0.5의 드롭아웃을 가진 단순 완전 연결 계층 2개에서 산출된 값을 soft-max 계층에 입력하였고 출력 값을 통해서 판별하였다. 모두 ReLU 활성화 함수를 이용하였다.

표 1은 영상에 대한 카메라 판별 연구에 사용된 장치의 모델 이름과 해상도, 사진의 수를 나타낸다.

표 1. 카메라 장치 목록

Table 1. List of camera device

Model name	Resolution	Count
Canon 650D	5184x3456	114
Canon EOS_500D	4770x3178	114
Canon EOS_M	5208x3476	87
Canon IXUS_160	5152x3864	162
Nikon COOLPIX_S33	4160x3120	101

그림 2의 모델을 이용하여 영상을 이용한 카메라 모델 판별 실험을 수행하였다. 실험 영상은 원본 영상을 256x256 크기 패치로 잘라서 이용하였다. 기존 PRNU를 이용한 방법은 균일한 영상에서 센서 패턴 노이즈의 통계학적 계수를 구했으나, 딥 러닝의 다양한 데이터에 대한 학습의 특성을 생각하여 균일한 영상과 비균일한 영상을 섞어 사용하였다. 이 중 학습에 80%, 테스트에 20%로 나누어 사용하였으며, 실험 결과로 92.5%의 성능을 확인하였다.

다음으로 영상에 대하여 학습된 카메라 모델 판별 딥 러닝 모델을 동영상의 각 프레임에 적용하는 실험을 수행하였다. 표 1에서 사용된 모델과 동일한 카메라 장치로 촬영한 동영상을 이용하였으며 I 프레임과 P 프레임으로 분리하여 입력에 사용하여 실험을 수행하였다. 프레임에 따른 판별 결과는 표 2와 같이 예상과는 다르게 낮은 결과로 나타났다.

표 2. 동영상 프레임 판별 결과

Table 2. Video frame identification accuracy

Frame type	Accuracy
I frame	12.9%
P frame	21.9%

기존의 영상 판별 결과에 반해 현저히 낮은 결과로 보아 기존의 영상을 이용한 카메라 모델 판별이 동영상에 적용할 수 없음을 확인할 수 있었다.

이러한 결과가 나타나는 이유는 영상과 동영상이 저장될 때 사용되는 압축방법의 차이에 따른 결과로 보인다. 실험에서 사용된 영상 데이터는 압축되지 않은 원본 영상이거나 JPEG 압축을 이용하여 압축된 영상이고, 동영상 데이터는 AVC(Advanced Video Coding)을 이용하여 압축된 영상이다. 또한, 동영상은 일반적으로 영상에 비해 높은 압축률이 요구되고 그에 따라 더 많은 정보가 수정, 삭제되므로 영상과는 차이가 있는 것으로 분석된다. 또한, 영상의 해상도가 동영상의 해상도보다 일반적으로 크기 때문에 영상과 동영상 간의 센서 노이즈의 크기와 위치가 다르다. 그 때문에 위치 동기화가 깨지게 되어 좋은 결과가 나오지 않은 것으로 분석된다.

위에서 언급한 바와 같이 동영상은 기존의 학습 모델을 이용할 경우 전혀 판별할 수 없는 정확도가 나오는 것으로 보아 기존의 학습 모델을 바로 이용할 수 없을 것으로 보인다.

3.3 동영상 프레임 특성에 따른 성능 분석

본 논문에서는 위에서 제시한 영상 기반 카메라 판별 기술의 딥 러닝 구조와 하이퍼 파라미터 등을 유지하고 학습 데이터로 영상이 아닌 동영상의 프레임을 사용하여 문제를 해결하였다. 특히, 본 연구에서는 I 프레임, P 프레임과 I와 P 프레임을 함께 사용한 데이터 등을 이용하여 서로 교차로 학습 및 테스트를 수행하여 데이터 간의 관계를 분석하였다.

실험 결과에 제시한 것과 같이 다양한 프레임별 분류를 통하여 실험을 수행하였다. 평균적으로 P 프레임으로 학습과 판별을 하는 것이 가장 높은 성능을 보였고 모델에 따른 판별률의 차이 또한 적은 것을 확인할 수 있었다. 반면에 I 프레임은 카메라 판별에 있어서 P 프레임보다 상대적으로 중요도가 낮은 것으로 확인하였다. 영상에서 일반적으로 I 프레임의 수보다 P 프레임의 수가 월등히 많다. 이는 많은 데이터에 의존하는 딥 러닝 프로세스에서 불리하게 작용 되기 때문에 I 프레임 사용에 대한 이점이 존재하지 않고 오히려 악영향을 끼치는 것을

확인할 수 있었다.

따라서 P 프레임만을 활용하는 것이 최적의 성능을 나타내는 것으로 분석되어 그림 2에서 제시한 것과 같이 P 프레임을 활용한 비디오 카메라 모델 판별 알고리즘을 설계하였다.

IV. 실험 결과

4.1 실험 환경 및 데이터

비디오 카메라 모델 판별을 위한 딥 러닝 학습 환경은 다음과 같다. CPU는 Intel i7-7700을 사용하였고, 실제적인 딥 러닝 학습을 담당하는 그래픽카드는 Nvidia Titan XP(VRAM 12GB)를 사용하였다. 메모리는 16GB로 구성하였으며, 윈도우 10 환경에서 가장 범용적이고 빠른 구현이 가능한 딥 러닝 프레임워크인 Tensorflow-GPU(1.10.0)을 사용하였다.

본 논문에서는 표 1의 기존의 카메라 장치에서 사용하였던 모델과 같은 장비에서 동영상을 획득하였다. 각각의 동영상들은 평균적으로 20초 전후의 길이를 가진 동영상이며, 표 3은 각각 장비에서 획득한 동영상의 해상도와 각각의 프레임의 수를 보여 준다.

그림 3은 표 3의 각각의 프레임의 영상 일부를 보여 준다. 각각의 프레임들은 크기가 매우 크고, 딥러닝 환경에 적합하지 않아서 학습과 테스트 시에 프레임을 256×256의 크기의 패치로 나누어 사용한다. 또한, 마지막으로 비디오 카메라 모델 판별 알고리즘 실험을 위해서 총 103개의 동영상 영상의 P 프레임만을 이용하여 데이터셋을 구축하였다.

Model name	I frame	P frame
Canon 650D		
Canon EOS_500D		
Canon EOS_M		
Canon IXUS_160		
Nikon COOLPIX_S33		

그림 3. 동영상 프레임 예시
Fig. 3. Video frame examples

6 딥 러닝을 이용한 비디오 카메라 모델 판별 시스템

표 3. 동영상 프레임 판별 결과

Table 3. Video frame identification accuracy

Model name	Resolution	I frame	P frame
Canon 650D	1920x1088	759	8257
Canon EOS_500D	1920x1088	820	7274
Canon EOS_M	1280x720	1517	16585
Canon IXUS_160	1280x720	6666	7215
Nikon COOLPIX_S33	1920x1080	309	8961

4.2 학습 진행에 따른 결과 및 분석

프레임별 영향과 데이터 간의 관계를 분석하기 위해 I 프레임과 P 프레임, 그리고 그 둘을 합친 데이터 셋을 이용하여 각각에 대해 학습을 시키고 학습된 3개의 모델에 대해 같은 데이터 셋에서 추출한 테스트 영상을 이용하여 정확도를 구하는 방식으로 실험을 수행하였다. 학습은 최대 200 epoch로 수행하였으며, 20 epoch마다 중단점을 두어 학습된 모델을 검증하였으며, 에포크와 각각의 프레임별에 대한 분석을 한 결과는 표 4, 표 5 및 표 6에 각각 나타나 있다.

표 4는 I 프레임으로 학습하고 각각의 프레임에 대한 정확도를 나타낸 표로 단일 I 프레임에 대해서 높은 결과를 보이며 표 5는 P 프레임으로 학습하고 각각의 프레임에 대한 정확도를 나타낸 표로 단일 P 프레임에 대해서 높은 결과를 보인다. 표 6은 I와 P 프레임을 모두 이용하여 학습한 결과의 정확도로 예상과는 달리 단일 P 프레임에서 높은 정확도를 보이고 있다.

표 4. I 프레임 학습 시 판별 정확도

Table 4. Identification accuracy with I frame training

Epoch	I frame	P frame	IP frame
0	0.200882	0.179334	0.180949
20	0.426958	0.260076	0.272577
40	0.511427	0.375774	0.385935
60	0.590217	0.411245	0.424651
80	0.660920	0.493820	0.506337
100	0.701951	0.557533	0.568351
120	0.727412	0.583745	0.594507
140	0.745122	0.595568	0.606771
160	0.763700	0.603696	0.615682
180	0.777065	0.606380	0.619166
200	0.792970	0.603609	0.617794

표 5. P 프레임 학습 시 판별 정확도

Table 5. Identification accuracy with P frame training

Epoch	I frame	P frame	IP frame
0	0.195937	0.185460	0.186245
20	0.590684	0.753469	0.741275
40	0.710572	0.847321	0.837078
60	0.750735	0.887062	0.876850
80	0.773456	0.908366	0.898260
100	0.793972	0.920704	0.911210
120	0.804197	0.931077	0.921572
140	0.817629	0.938956	0.929867
160	0.827987	0.944973	0.936210
180	0.833868	0.949551	0.940885
200	0.840417	0.953496	0.945025

표 6. I, P 프레임 학습 시 판별 정확도

Table 6. Identification accuracy with I & P frame training

Epoch	I frame	P frame	IP frame
0	0.094493	0.080996	0.082007
20	0.592756	0.791602	0.776707
40	0.706830	0.868994	0.856846
60	0.749465	0.900152	0.888864
80	0.785819	0.917900	0.908006
100	0.808541	0.930087	0.920982
120	0.829190	0.938182	0.930017
140	0.846832	0.944188	0.936896
160	0.864809	0.949746	0.943383
180	0.876504	0.953442	0.947678
200	0.881382	0.957430	0.951733

일반적으로 key 프레임이라고 부르는 I 프레임의 경우에 많은 특성이나 정보가 남아 있을 것으로 생각했으나 본 실험에서 나타난 것과 같이 I 프레임만 학습시킨 경우에 전반적으로 낮은 정확도를 보이는 것을 알 수 있다. 또한, I와 P 프레임을 모두 학습시킨 경우에 I 프레임만 학습시켰을 때에 비해서 I 프레임에 대한 정확도가 높은 것으로 보아, 동영상 프레임의 특성에 따른 정확도뿐만 아니라 많은 데이터의 수 또한 정확도에 영향을 끼치는 것으로 나타났다.

그림 4는 모든 실험 결과를 하나의 그래프로 도시하였으며, ‘·’의 좌변은 학습에 사용된 프레임, 우변은 검증에 사용된 프레임을 나타낸다.

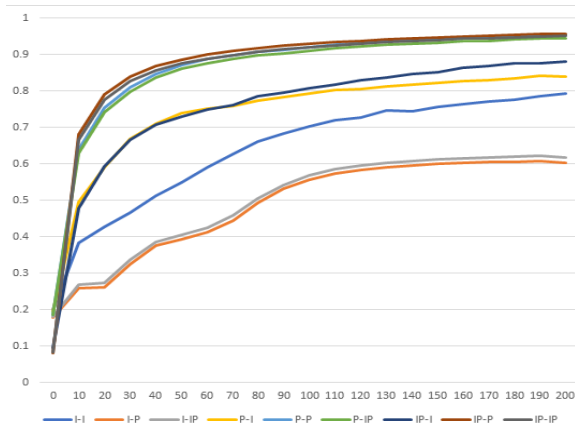


그림 4. 학습-예측 정확도
Fig. 4. Training-prediction accuracy

4.3 프레임 종류에 따른 결과 및 분석

본 논문의 최종 목표인 비디오 카메라 모델 판별을 위해서 I와 P 프레임의 다양한 학습/예측 중에서 최적의 모델을 찾기 위해 모델 각각에 대한 정확도를 구하였다.

I 프레임을 학습시킨 후 I, P, I/P 프레임을 판별 했을때의 정확도, P 프레임을 학습시켰을 때 I, P, I/P 프레임을 판별 했을때의 정확도, I와 P 프레임 모두를 학습시켰을 때 I, P, I/P 프레임을 판별 했을 때의 정확도를 분석하였고 표 7에 각각의 학습/예측 모델의 정확도를 정리하였다.

I 프레임 학습의 경우, 정확도가 다른 경우에 비해 낮은 것을 확인할 수 있다. 또한, I와 P 프레임 학습의 경우에도 I 프레임의 영향으로 인해 정확도가 P 프레임 학습에 비해 낮은 것을 알 수 있다. 따라서 P 프레임 학습과 P 프레임 예측의 경우가

장 높은 정확도를 보이는 것을 알 수 있다. 뿐만 아니라 비슷한 평균 정확도를 보이는 I와 P 프레임 학습과 I와 P 프레임 예측은 정확도의 편차가 나타나는 것을 볼 수 있다. 결론적으로 비디오 카메라 모델 판별을 위해서 P 프레임 학습 및 P 프레임 예측을 이용하는 것이 우수한 것으로 나타났다.

4.4 동영상 판별 결과 및 분석

본 논문에서 제안하는 비디오 카메라 모델 판별 알고리즘의 경우 프레임별 분석이 아닌 입력으로 사용된 패치의 출력값을 누적 계산하여 가장 많은 클래스의 모델을 해당 모델이라 판별한다. 표 8에는 이러한 결과를 간략히 나타내었으며, True는 정확하게 찾은 패치의 수를, All은 전체 입력으로 사용된 패치의 수를 나타낸다. 실험 결과에 따르면 모든 동영상에 대해 전체 패치 수 중에서 정확하게 찾은 패치의 수가 잘못 판별된 경우보다 많음을 알 수 있다. 본 알고리즘을 이용하여 동영상을 촬영한 비디오카메라 모델 판별에 대해서는 100%의 정확도를 달성하였다.

V. 결론 및 향후 과제

현대 기술이 발전함에 따라 멀티미디어 획득 장치 기술 또한 빠르게 발전해왔다. 이들 장치는 이제 단순히 영상뿐만 아니라 동영상을 획득할 수 있게 되었다. 이러한 기술의 발전으로 멀티미디어 콘텐츠가 불법적으로 이용되는 경우가 늘어나고, 포렌식 기술에 대한 수요가 증가하고 있다.

표 7. 학습/예측 모델 정확도 평균
Table 7. Training/prediction model accuracy average

Camera	Training/prediction								
	I/I	I/P	I/IP	P/I	P/P	P/IP	IP/I	IP/P	IP/IP
Canon 650D	69.43%	54.46%	55.72%	85.01%	93.03%	92.35%	89.94%	95.75%	95.26%
Canon EOS_500D	88.70%	23.38%	30.00%	93.84%	93.66%	93.68%	92.53%	90.83%	91.01%
Canon EOS_M	84.31%	65.84%	67.39%	77.17%	97.82%	96.09%	80.82%	97.53%	96.13%
Canon IXUS_160	79.03%	51.39%	53.74%	93.73%	96.69%	96.44%	97.39%	96.35%	96.44%
Nikon COOLPIX_S33	70.05%	94.77%	93.94%	72.18%	99.72%	98.80%	86.18%	99.88%	99.43%
Average	78.30%	57.97%	60.16%	84.39%	96.18%	95.47%	89.37%	96.07%	95.65%

표 8. 모델 별 비디오 카메라 패치 판별 결과

Table 8. Result of video camera patch identification for each model

No.	Canon 650D		Canon EOS_500D		Canon EOS_M		Canon IXUS_160		Nikon COOLPIX_S33	
	True	All	True	All	True	All	True	All	True	All
0	10891	11256	9578	10080	9382	9850	205	220	19484	19488
1	14133	15484	11571	11844	7072	7080	629	790	12163	12180
2	8537	10192	9978	10276	7578	7580	3831	4030	12073	12180
3	7902	10640	9576	9912	7913	8000	3533	3840	12112	12180
4	11232	12320	7592	9156	8090	8290	3530	3580	12065	12180
5	9194	10752	8558	9772	7135	7220	3431	3490	12117	12180
6	9315	11032	9630	10332	7338	7470	3436	3440	12114	12180
7	10141	12124	8538	10836	8888	8900	3433	3440	12067	12180
8	12257	13356	9597	10696	8052	8370	3559	3560	12113	12180
9	9490	10444	6505	9548	10396	11110	3484	3490	12119	12180
10	9698	9912	8426	8456	7041	7090	3495	3500	12128	12180
11	9857	10052	8552	8596	8711	8750	3546	3550	11364	11368
12	10644	10976	10050	10080	8993	9380	3719	3720	12173	12180
13	9861	10052	11037	11144	8270	8320	3408	3490	12173	12180
14	9927	10388	8578	8652	8526	8580	3456	3460	12170	12180
15	9992	10108	8948	9072	7392	7490	3424	3430	11361	11368
16	9798	9940	12469	12600	7352	7380	3384	3490	12175	12180
17	10839	11368	12631	13020	7355	7440	3297	3600	12170	12180
18	10090	10164	8697	8736	8732	8970	3300	3500	13786	13804
19	9270	9576	10719	10864	8192	8580	3182	3550	12174	12180
20	11021	11060					3203	3480		
21							3117	3500		

본 논문에서는 동영상을 이용한 비디오 카메라 모델 판별을 위한 방법을 제안하였다. 먼저 기존의 영상에 대한 카메라 모델 판별 기술을 바로 이용할 수 없음을 확인하였다. 하지만 영상에 대한 카메라 모델 판별에 사용한 모델의 구조를 그대로 이용하여 동영상의 각각의 프레임을 학습하고 예측한 결과를 토대로 이를 이용할 수 있음을 확인하였고, 이 중 평균적으로 최고의 성능을 나타내는 P 프레임 학습, P 프레임 예측을 이용하여, 각각의 동영상에 대해 전체 패치 중에서 가장 많이 분류된 패치를 모델로 예측하여 5개 비디오카메라로 촬영한 동영상에 대해 100%의 판별 정확도를 얻음을 확인할 수 있었다.

다만 본 논문에서는 동영상의 용량이 방대하고 딥러닝 하드웨어의 제약으로 인하여 사용한 카메라 모델의 수가 적고, 다양한 경우에 대해서는 실험을 수행하지 못한 문제가 있으며, State-of-the-art 기술이 많이 발표되고 있는데 이를 적용하지 못하여 아쉬운 점이 있다. 하지만 본 연구를 통하여 기존의

영상 획득 장치 판별을 위한 딥러닝 모델의 구조를 이용하여 동영상 프레임 특성을 활용하여 학습하면 동영상에 대해서도 판별을 할 수 있음을 확인할 수 있었고, 나아가 동영상 파일 각각의 패치의 정확도를 누적하여 높은 성능의 판별률을 달성할 수 있었다. 차후 연구는 State-of-the-art한 기술을 프레임 예측에 도입하는 것과 보다 많은 비디오 카메라 장치를 학습하여 보다 많은 장치에서도 본 알고리즘이 효율적으로 수행되는 것을 목표로 한다.

References

- [1] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise", *IEEE Trans. on Information Forensics Security*, Vol. 1, No. 2, pp. 205-214, Jun. 2006.
- [2] K. S. Choi, E. Y. Lam, and K. Y. Wong, "Source camera identification using footprints from lens

aberration", Proc. of SPIE, Digital Photography II, Vol. 6069, No. 1, pp. 172-179, Feb. 2006.

[3] S. Bayram, H. T. Sencar, and N. D. Memon, "Improvements on source camera model identification based on cfa interpolation", Proc. of IFIP Int. Conf. on Digital Forensics, pp. 289-299, 2006.

[4] H.-Y. Lee, "Imaging Device Identification using Sensor Pattern Noise Based on Wiener Filtering", The Trans. of the Korean Institute of Electrical Engineers, Vol. 65, No. 12, pp. 2153-2158, Sep. 2016.

[5] S.-H. Lee, D.-H. Kim, T.-W. Oh, K.-B. Kim, and H.-Y. Lee, "Digital Video Source Identification Using Sensor Pattern Noise with Morphology Filtering", KIPS Trans. on Software and Data Engineering, Vol. 6, No. 1, pp. 15-22, Jan. 2017.

[6] J. Schmidhuber, "Deep learning in neural networks: An overview", Neural networks, Vol. 61, pp. 85-117, Jan. 2015.

[7] A. Fischer and C. Igel, "An introduction to restricted Boltzmann machines", Proc. of Iberoamerican Congress on Pattern Recognition, pp. 14-36, Sep. 2012.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", Prof. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 770-778, Jun. 2016.

[9] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks", Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp. 4700-4708, Jul. 2017.

[10] A. Tuama, F. Comby, and M. Chaumont, "Camera model identification with the use of deep convolutional neural networks", Proc. of IEEE Int. Workshop on Information Forensics and Security, Abu Dhabi, United Arab Emirates, pp. 1-6, Dec. 2016.

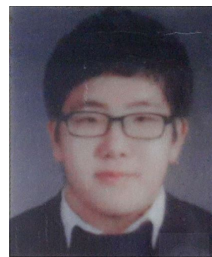
[11] B. Wang, J. Yin, S. Tan, Y. Li, and M. Li,

"Source camera model identification based on convolutional neural networks with local binary patterns coding", Signal Processing-Image Communication, Vol. 68, pp. 162-168, Oct. 2018.

[12] U. Kamal, A. M. Rafi, R. Hoque, S. Das, M. A. Abrar, and M. K. Hasan, "Application of DenseNet in Camera Model Identification and Post-processing Detection", arXiv preprint arXiv:1809.00576, Sep. 2018.

저자소개

김 동 현 (Dong-Hyun Kim)



2016년 2월 : 금오공과대학교
컴퓨터소프트웨어공학과 (학사)
2016년 2월 ~ 현재 : 금오공과
대학교 소프트웨어공학과
석사과정
관심분야 : 이미지 처리, 포렌식,
딥 러닝

이 수 현 (Soo-Hyeon Lee)



2018년 2월 : 금오공과대학교
컴퓨터소프트웨어공학과 (학사)
2018년 2월 ~ 현재 : 금오공과
대학교 소프트웨어공학과
석사과정
관심분야 : Image Processing,
Deep Learning

이 해 연 (Hae-Yeoun Lee)



1997년 : 성균관대학교 정보공학과
(학사)
1999년 : 한국과학기술원 전산학과
(공학석사)
2006년 : 한국과학기술원
전자전산학과 (공학박사)
2008년 ~ 현재 : 금오공과대학교
컴퓨터소프트웨어공학과 교수
관심분야 : Digital Forensics, Image Processing, IoT