

일반논문 (Regular Paper)

방송공학회논문지 제23권 제6호, 2018년 11월 (JBE Vol. 23, No. 6, November 2018)

<https://doi.org/10.5909/JBE.2018.23.6.886>

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

CNN 기반 동영상의 프레임 삭제 검출 기법

홍진형^{a)}, 양윤모^{a)}, 오병태^{a)‡}

Detection of Frame Deletion Using Convolutional Neural Network

Jin Hyung Hong^{a)}, Yoonmo Yang^{a)}, and Byung Tae Oh^{a)‡}

요약

본 논문에서는 동영상의 압축 과정에서 발생하는 규칙성을 이용하여 동영상의 조작 여부를 검출하는 기법에 대해 소개한다. 제안 방식은 동영상의 이중 압축과 일부 영상의 조작에 의해 잃어버린 계층적 규칙성을 이용한다. 이러한 불규칙성을 추출하기 위해 HEVC의 기본 단위인 CU와 TU의 분할정보를 이용한다. 성능 향상을 위해 지역적인 정보를 활용하여 CU와 TU의 분할 지도를 제작한 뒤, GoP 단위로 묶어 입력 데이터를 제작한다. 효과적인 분류를 위하여 3차원 합성곱 신경망을 이용하여 동영상의 이중 압축 및 조작 여부를 판단한다. 실험 결과, 기존의 기계학습 알고리즘을 이용한 연구 결과에 비해 더욱 효과적으로 동영상의 조작 여부를 판단함을 확인하였다.

Abstract

In this paper, we introduce a technique to detect the video forgery by using the regularity that occurs in the video compression process. The proposed method uses the hierarchical regularity lost by the video double compression and the frame deletion. In order to extract such irregularities, the depth information of CU and TU, which are basic units of HEVC, is used. For improving performance, we make a depth map of CU and TU using local information, and then create input data by grouping them in GoP units. We made a decision whether or not the video is double-compressed and forged by using a general three-dimensional convolutional neural network. Experimental results show that it is more effective to detect whether or not the video is forged compared with the results using the existing machine learning algorithm.

Keyword : Video Forensics, Frame Deletion, HEVC, CNN, Coding Pattern

a) 한국항공대학교 항공전자정보공학부(Korea Aerospace University)

‡ Corresponding Author : 오병태(Byung Tae Oh)

E-mail: byungoh@kau.ac.kr

Tel: +82-2-300-0409

ORCID: <https://orcid.org/0000-0003-1437-2422>

※ 본 연구는 2016년도 정부(교육부)의 재원으로 한국연구재단 기초연구사업(NRF-2016R1D1A1B03930917)과 경기도 지역협력 연구센터 사업 (GRRC) (2017-B02, 이동형 미디어 디바이스를 위한 영상처리 및 UI 플랫폼 기술연구)의 지원을 받아 수행되었음.

※ This research was supported in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B03930917), and by the GRRC program of Gyeonggi Province [2017-B02, Study on image processing and UI platform for mobile media devices].

· Manuscript received September 11, 2018; Revised November 16, 2018; Accepted November 16, 2018.

Copyright © 2016 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. 서론

급격히 발전하는 과학 기술에 힘입어, 디지털 영상/동영상은 일상생활에서 빠질 수 없는 기술이 되었다. 일반적으로, 새로운 정보를 취득함에 있어서 시각을 이용하는 것은 다른 감각을 이용하는 것에 비해 효과적이다. 실제로 사람들은 수없이 다양한 경로를 통해 디지털 영상의 정보를 획득하고 있으며, 그 스스로가 정보 전달원이 되어 새로운 디지털 영상의 정보를 공유하고 있다. 이런 사회 현상 속에서 다양한 수단과 목적을 위해 디지털 정보가 수정, 조작되어 공유되는 경우가 있다. 따라서 이러한 과학 기술의 발전에 따른 사회적 변화 속에서 디지털 정보의 조작을 구별하는 디지털 포렌식 기술의 발전이 필요하다. 디지털 포렌식 기술 중에서도 비디오 포렌식 분야는 매년 다양하게 발전하는 동영상 포맷과 그 해상도에 따라 발전되어 왔다. 비디오 포렌식 기술에는 크게 동영상 내부의 시/공간적 특징 및 상관관계를 통한 조작 검출 방식과 물리적 불일치성을 통한 조작 검출 방식, 그리고 이중 및 다중 압축의 판별을 통한 원본 동영상 여부 검출 방식 등이 있다. 이 중에서 이중 및 다중 압축을 판별하여 동영상의 조작 여부를 검출하는 연구는 매우 다양한 분야에서 다양한 방식으로 발전되어 왔다^[1]. 과학 기술의 발전에 맞춰, 본 논문에서는 딥-러닝을 이용한 동영상의 이중 압축 여부를 검출하는 기법에 대해 소개하고자 한다.

실제로 법정에서는 영상이나 동영상과 같은 디지털 정보들의 위조 여부가 상당히 중요한 만큼 조작 여부를 판별하는 기술은 반드시 그에 맞춰 발전해나가야 한다. 영상 및 동영상은 매우 다양한 방식으로 조작되어 왔고, 조작 여부를 검출하는 기술 또한 다양하다. 예를 들어, 동영상의 조작 여부를 검출하기 위한 화질 평가 기술이 발전하였다^[2]. 또한 Photo Response Non-Uniformity를 기반으로 동영상의 출처를 식별하는 기술이 발전하였다^[3]. 동영상은 보다 높은 해상도와 다채로운 색감을 제공하기 위해 그 크기가 일반 영상에 비해 상당히 크다. 따라서 그에 따른 동영상의 압축 기술 또한 발전하였다. 이렇게 발생한 동영상 압축 파라미터를 분석하는 조작 검출 기법도 있다^[4].

압축 정보를 바탕으로 영상의 진위여부를 판단하는 비디오 포렌식 연구는 작은 크기의 해상도(cif, qcif등)의 동영상

에 대해서 진행되어 왔다. 하지만, 디지털 정보를 쉽게 얻을 수 있는 최근에는 작은 크기의 해상도를 갖는 동영상은 거의 없다. 그리고 HD(1920×1080), UHD(3840×2160)와 같은 크기의 해상도를 갖는 동영상을 획득 할 수 있게 되었다. 이렇게 큰 해상도의 동영상을 획득하기 위해 더욱 더 효율적인 동영상 압축 기술이 발전하였다. 그러므로, 대부분의 동영상이 H.264/AVC와 HEVC와 같은 고효율 비디오 코덱으로 압축되어 왔다. 그에 비해 HEVC로 압축된 동영상에 대한 비디오 포렌식 연구가 많지 않다. 기존의 MPEG-2 또는 H.264/AVC로 압축된 동영상에 대한 연구가 대부분이다^[4-11]. 따라서 본 논문에서는 보다 큰 HD 해상도의 HEVC로 압축된 동영상에 대하여 HEVC의 부호화 정보를 전처리하여 동영상 조작 여부를 검출하는 기술에 대해 소개한다.

전통적으로 기계학습을 이용하여 동영상의 조작을 검출해왔다. 더 나아가, 최근에는 딥-러닝 기술을 이용한 동영상 조작 검출 기술이 소개되기도 한다. 예를 들어, 이중 압축 시 재배열되는 I 프레임의 위치를 Convolutional Neural Network(CNN)을 이용하여 검출하여 조작 여부를 판단하기도 한다^[12]. 본 논문에서는 동영상의 압축 시 발생하는 부호화 패턴을 영상 형태로 전 처리하여 CNN에 입력함으로써 동영상의 이중 압축 여부를 판단하는 기술에 대해 소개한다.

본 논문의 구성은 다음과 같다. 2장에서는 동영상 위조 검출 알고리즘의 전통적인 기법에 대해 설명하고, 3장에서는 제안 알고리즘을 소개한다. 4장에서는 제안 알고리즘의 성능을 다양한 실험 결과에 따라 확인한 뒤, 5장에서 마지막으로 본 논문에 대한 결론을 맺는다.

II. 기존 연구

과거부터 현재까지, 동영상의 위조 여부를 판단하는 알고리즘은 다양한 방향으로 발전되어 왔다. 그 중에서, Wang과 Farid는 영상의 조작 여부를 판단하는 알고리즘을 응용하여 동영상의 화면 내 예측 프레임의 조작 여부를 판단하는 알고리즘을 제안하였다^[4]. 해당 연구는 MPEG-1과 MPEG-2로 이중 압축된 동영상에 대해 일부 영상의 삭제

또는 삽입 조작의 흔적을 찾는 방법을 보였다. 또한 MPEG-2 동영상 압축 과정에서 변환 과정의 DCT계수를 이용하여 이중 압축 여부를 판단하는 알고리즘이 있다¹⁵. 그 외에도 HEVC 이전의 코덱에 대한 동영상 압축의 기본 단위인 Macro Block(MB)단위로 동영상의 조작 여부를 판단하는 기법도 있다⁷. 이와 같이 MPEG-1또는 MPEG-2 환경에서 부호화 계수를 이용하여 조작된 위치에서의 특별한 분포를 찾아 동영상의 위조 여부를 검출하는 방식이 있다. 이처럼 동영상의 압축 단계에서 부호화 계수의 특별한 분포를 탐지하는 방식 이외에도 이중 압축과 일부 영상의 조작에 의해 발생하는 특별한 패턴을 찾아 새로운 변수를 제안하여 조작 여부를 판단하는 연구도 있다. 예를 들어 MPEG-2, MPEG-4 그리고 H.264 환경에서 Variation of Prediction Footprint(VPF)를 특징으로 추출하여 조작 여부를 판단하고 압축된 동영상의 첫 번째 압축 Group of Picture(GoP)의 크기를 추론하는 연구가 있다⁸. 또한 VPF와 화면 간 예측 프레임을 이용하여 동영상의 조작 여부를 검출하는 연구도 있다⁹. 마지막으로, Motion Compensated Edge Artifact (MCEA)를 추출하여 공간적인 동영상 특징과 함께 일부 영상의 삭제 또는 삽입 조작을 검출하는 연구가 있다^{10,11}.

위와 같이 다양한 방식으로 동영상의 조작 여부를 검출하는 비디오 포렌식 연구 분야 중에서도 특히 활발하게 진행되는 연구 분야가 바로 이중/다중 압축을 이용한 조작 동영상 검출 분야이다. 동영상은 최초로 기록될 때 보다 효율적으로 동영상을 기록하기 위해서 데이터의 첫 번째 압축이 수행된다. 그 뒤, 다른 목적을 가진 누군가가 해당 동영상을 조작하기 위해서는 첫 번째로 압축된 동영상을 복호화하고, 조작을 진행한 이후, 두 번째 압축을 수행된다. 따라서 첫 번째 압축 이외의 다중 압축 기록이 있을 때 해당 동영상은 더 이상 원본이라고 판단할 수 없게 된다. 이중/다중 압축을 이용한 조작 동영상 검출 분야는 이렇게 동영상의 이중 또는 다중 압축 여부를 판단하여 본 동영상의 원본 진위여부를 판단한다. 위에서 언급한 [4-11] 외에도, 마르코프 기반의 특징 추출 기법을 이용하여 MPEG-4 환경에서 압축된 동영상의 이중 압축 여부를 판단하는 알고리즘이 있다¹³.

이중 압축을 이용하여 동영상의 조작을 탐지하는 주목할 만한 논문이 바로 Shanableh가 소개한 영상 삭제 검출 기술이다¹⁴. 해당 논문에서는 MPEG-2 환경에서 압축된 동영

상을 복호화 하는 과정에서 일부 영상 삭제 여부를 판단하기에 적합한 8개의 특징 벡터를 추출한다. 그렇게 추출된 8개의 특징 벡터 중에서 삭제된 영상의 수에 따라 적합한 특징 벡터를 단계적 회귀 분석을 통해 선택한다. 선택된 특징 벡터의 차원을 Linear Discriminant Analysis(LDA)를 이용하여 효과적으로 감소시켰다. 그 뒤, 다양한 기계학습 알고리즘을 이용하여 동영상의 이중 압축 여부를 분류하였다. 그러나 해당 논문에서는 기계학습에 사용된 동영상 모두 매우 작은 크기의 해상도(cif, qcif)를 가지고 있으며, 일부 특징 벡터가 H.264 다음 세대의 코덱인 HEVC에는 알맞지 않은 단점이 있다. 다른 주목할 만한 연구는 Yu가 제안한 동영상의 일부 영상 삭제 여부를 검출하는 기술이다¹⁵. 해당 논문에서는 예측 잔차 신호와 Intra MB의 개수를 특징 벡터로 추출하여 전 처리를 통해 하나의 매개변수로 합성하여 동영상에서 일부 영상의 삭제 여부를 검출한다. 이렇게 추출되어 전 처리된 매개변수는 실험적으로 정해진 문턱값을 이용하여 조작 영상 여부를 판단하게 된다. 그러나 해당 논문에서도 이전에 언급한 논문과 마찬가지로 실험에 사용된 동영상의 해상도가 매우 작고, Intra MB는 HEVC 코덱에서 본 알고리즘을 적용하는데 어려움이 있다. 또한, 경험적으로 선택된 문턱값을 이용하는 방식은, 경험하지 못한 새로운 동영상의 조작 여부 판단에 있어서 상당히 불안정하다. 마지막으로 HEVC환경의 부호화 패턴을 분석하여 동영상의 조작 여부를 검출하는 연구가 있다¹⁶. 해당 알고리즘은 검증 동영상의 부호화 패턴 중 분류하기 용이한 패턴들을 추출하여 기계학습 알고리즘을 통해 분류 기준을 모델링한다. 이 과정에서 사용되는 k-Nearest Neighbor(kNN) 분류기는 경험적으로 k값을 직접 정해야하며, GoP단위로 특징을 추출할 때 지역적인 정보는 고려하지 않았다. 또한 임의의 접근 모드에서의 성능을 알 수 없다. 이러한 논문의 단점들을 해결하기 위해서 본 논문에서는 HD 이상의 고해상도 영상에도 적용할 수 있도록 설계하였다. 또한, 경험적으로 선택하는 문턱값 방식 대신 보다 일반적인 특징 벡터를 추출하고 지역적인 정보를 활용하여 분류하기 위해 합성곱 신경망을 이용한 알고리즘을 소개한다. 본 논문에서는 현재 많은 분야에서 널리 사용되고 있는 HEVC 압축 코덱의 환경에서 동작할 수 있도록 HEVC의 부호화 계수를 이용하여 동영상의 조작 여부를 판단한다.

III. 제안 알고리즘

제안하는 알고리즘에서는 인위적으로 조작된 동영상의 특징을 추출하기 위해서 이중 압축 검출에 중점을 두었다. 조작된 동영상의 경우 앞서 언급했던 대로 이중 압축이 필수 불가결하다. 예를 들어 최초에 동영상이 압축된 후 동영상의 일부 영상이 삭제된다면, 두 번째 동영상의 압축 과정이 반드시 필요하다. 이 때, 동영상 압축 과정에서 부여되는 영상의 종류가 변화하는 영상이 발생한다. 아래 그림 1은 해당 과정을 자세하게 나타낸 것이다. 영상의 삭제로 인해 첫 번째에서 I 프레임이었던 영상은 두 번째 압축 시 그 종류가 변화하여 B 프레임이 된다. 이러한 변화는 이어지는

영상에서 지속적으로 발생하여 압축 시 생기는 패턴의 변화가 생기게 된다.

제안하는 알고리즘에서는 부호화 패턴의 변화를 분석하기 위해 HEVC의 부호화 정보를 이용하여 특징으로 추출한다. 그 뒤, 적절한 전 처리 과정을 거쳐 미리 구성된 합성곱 신경망의 입력으로 제공한다. 아래 그림 2는 본 논문에서 제안하고자하는 전반적인 알고리즘의 순서도이다.

1. HEVC 부호화기의 구조

앞서 언급한 것처럼, HEVC에서의 동영상의 일부 영상 삭제를 검출하기 위해 HEVC의 기본 구조를 이용할 필요

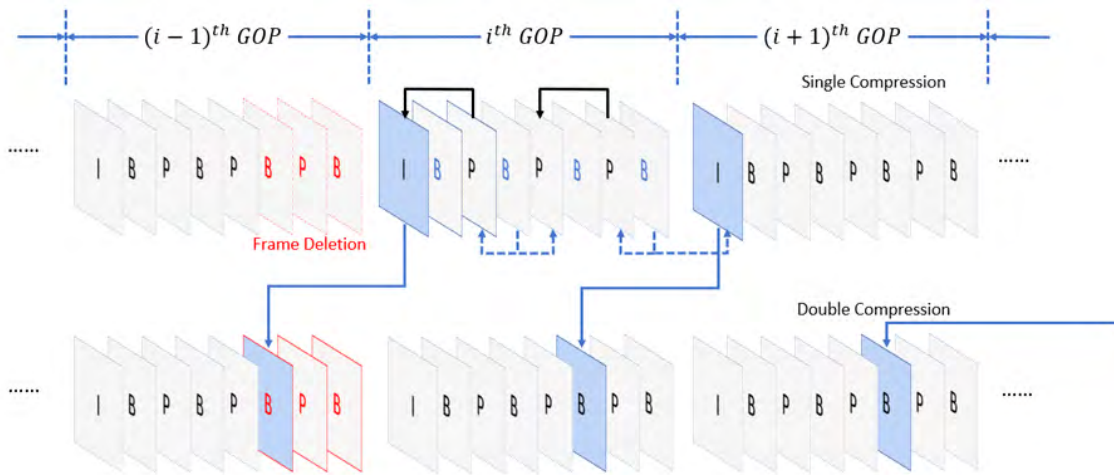


그림 1. 영상 삭제 시 발생하는 프레임 타입의 변화
 Fig. 1. Change of the frame type by frame deletion

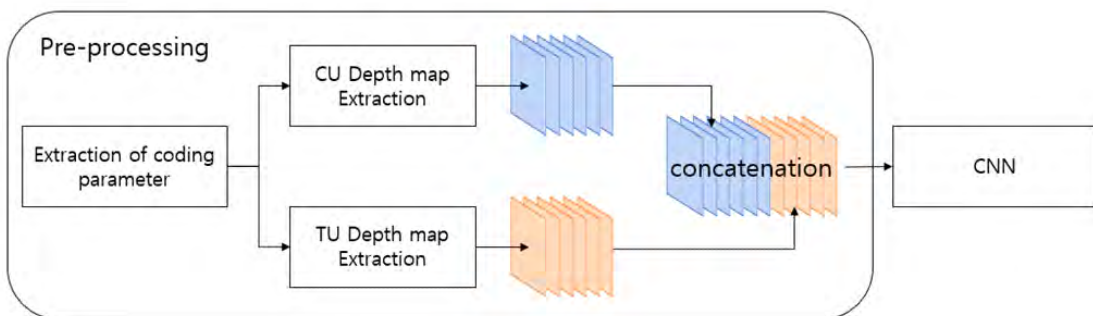


그림 2. 제안 알고리즘의 순서도
 Fig. 2. Block diagram of proposed algorithm

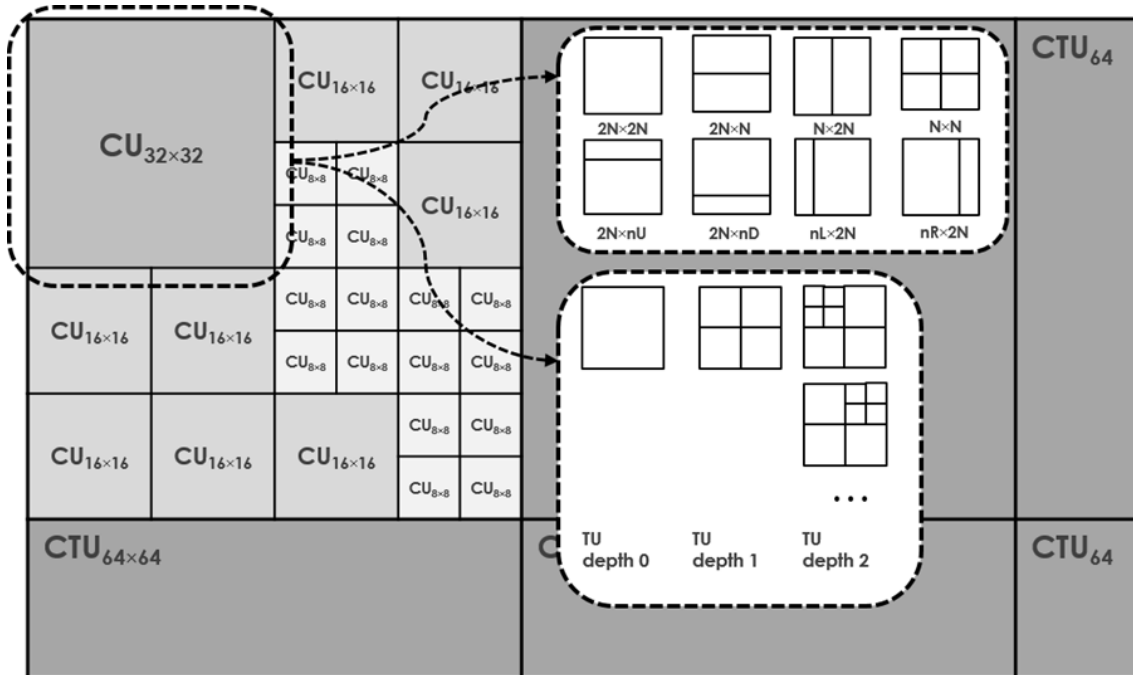


그림 3. HEVC의 주요 부호화 단위
Fig. 3. Main encoding unit in HEVC

가 있다. 아래 그림 3은 HEVC에서 정의한 압축 기본 단위를 그림으로 나타낸 것이다. HEVC 부호화기에서는 MPEG-2, H.264/AVC와 달리 매크로블록(Macro Block)을 이용하지 않고, 더 큰 단위인 Coding Tree Unit (CTU)와 그 세부 단위인 Coding Unit (CU), Prediction Unit (PU), Transform Unit (TU)을 이용하여 동영상을 부호화 한다. CTU는 모든 압축과정에서 CU로 분할된다. CU는 압축 효율에 따라 크게 64×64 에서 작게는 8×8 까지 다양한 크기로 분할된다. 각각의 CU는 화면 내 예측 또는 화면 간 예측 모드 중 하나로 압축된다. 화면 내 예측의 경우 영상 내에서만 참조하고 화면 간 예측의 경우 이웃의 영상을 참조한다. 각각의 CU는 필요한 경우에 한해 변환 및 양자화 과정을 거치는데, 이 때 TU를 기준으로 해당 과정이 수행된다. 이 두 가지 CU와 TU는 뒤에서 제안하는 알고리즘의 특징 추출에서 사용하게 된다.

2. 특징 추출

동영상은 압축과정에서 정해진 수의 일련의 프레임들을

GoP로 묶어 압축된다. 압축 과정에서는 GoP내에 속해있는 영상들의 종류에 따라 각각의 영상을 계층적인 분포로 구성한다. 이러한 구조는 다양한 부호화 파라미터가 갖게 되는 규칙성을 야기한다. 제안 방식에서는 동영상 조작 여부 검출에 이용하기 위해 GoP 단위의 이러한 규칙성을 특징으로 사용한다. 삭제 이전에 존재하는 각각의 GoP는 모두 영상의 계층적인 구조에 의해 비슷한 규칙성을 가지고 있다. 반면, 삭제 이후에는 첫 번째 압축 때와는 다른 종류의 영상과 계층적인 구조의 변화에 따라 원래 가지고 있던 규칙성을 잃게 된다. 이러한 사실은 동영상의 조작 여부를 판단하는데 큰 역할을 하게 된다.

제안하는 알고리즘은 이중 압축과 일부 영상 조작에 의해 잃어버린 계층적 규칙성을 이용하여 분류 문제를 수행한다. 따라서 우리는 부호화 계수 중에서 압축의 기본 단위인 CU와 변환과 양자화의 기본 단위인 TU에 주목하였다. 위에서 언급한 대로, CU와 TU는 부호화 과정에서 가장 기본이 되는 단위이다. 규칙성을 잃는다는 것은 비디오 코덱 관점에서 이전 부호화 파라미터와는 전혀 다른 양상으로 동영상을 압축할 수 있다는 가능성을 내포한다. 그러므로

논문 제안 방식에서는 GoP의 계층적 부호화에 의한 규칙성의 정보를 이용하기 위해 부호화 파라미터 중에서 CU와 TU를 선택하였다.

앞서 설명한 대로, 동영상의 조작 여부를 판단하기 위해서 CU와 TU를 GoP 단위로 추출한다. 동영상의 일부 영상이 삭제되었을 때, 실제로 조작 여부 판단의 큰 영향을 주는 것은 시간의 흐름에 따라 이동하는 물체라고 할 수 있다. 예를 들어 일반적인 CCTV의 동영상의 경우 기본적으로 배

경은 움직임이 거의 없다. 이러한 정적인 동영상의 경우 일부 영상이 삭제되더라도 알아차리기가 어려울 수 있다. 하지만, 움직임이 없는 배경 속에서 이동하는 물체가 있다면 그 시점에서 일부 영상이 삭제되었는지를 판단하는 것은 자명하다. 따라서 CU와 TU는 분할정보를 이용하여 동영상의 조작 여부를 판단하는데 있어서 해당 CU와 TU의 지역적인 정보를 이용하는 것은 본 알고리즘의 목적 달성의 큰 도움이 될 수 있다. 그러므로 본 논문에서는 CU와 TU의



그림 4. 영상 삭제 조작에 의해 시간적으로 떨어진 동영상에서의 지역 정보
 Fig. 4. Local information that has fallen in time due to the frame deletion

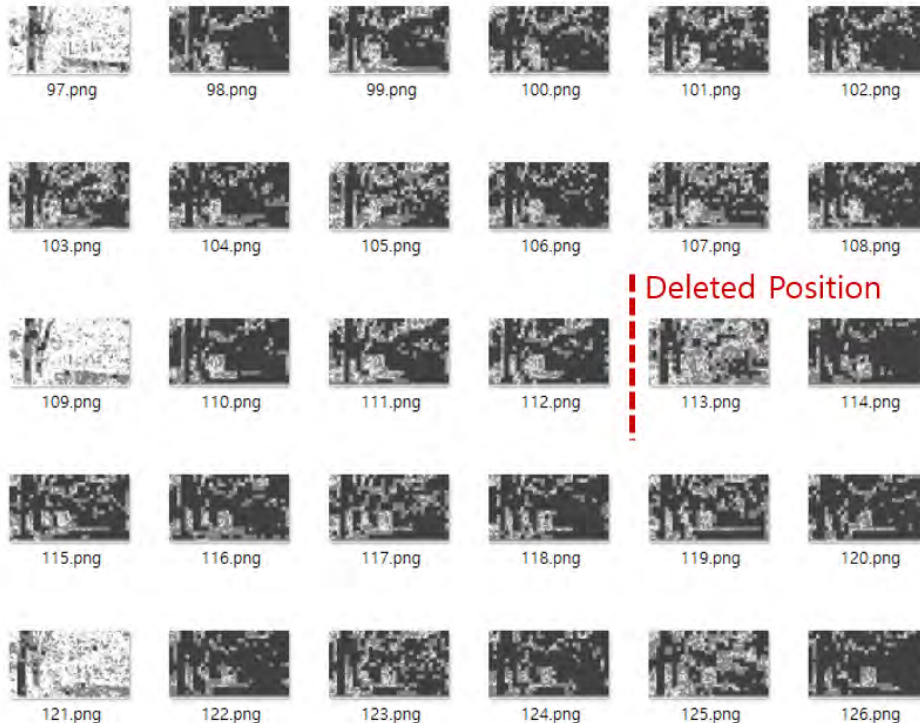


그림 5. 추출된 CU로부터 제작된 2차원 CU 분할 지도
 Fig. 5. Two-dimensional CU depth map produced from extracted CU

지역적인 정보를 이용할 수 있도록 해당 특징을 전 처리한다. 이렇게 CU와 TU를 2차원 영상 형태의 정보로 재가공하여 2차원 정보를 분류할 수 있는 합성곱 신경망을 이용해 동영상의 조작 여부를 판단한다. 아래 그림 4는 영상의 삭제 조작에 대한 지역적인 정보의 상태를 나타낸 것이다.

CU와 TU는 지역적인 정보를 이용하기 위해 2차원의 형태로 전 처리하는 과정이 필요하다. 그림 2에서와 같이 압축된 비트스트림 동영상에서 CU와 TU의 위치 정보와 크기 정보를 추출하여 동영상의 해상도와 동일한 2차원 분할 지도를 제작한다. 분할 지도는 분할 정도에 따라 4단계로 구분하며, 지도로 제작하기 위해 4단계에 해당하는 분할 정보를 0~255 사이의 화소 값에 매핑한다. 각각의 분할 값에 대한 화소 값은 아래 식 (1)과 같다.

$$Value = \frac{255}{4} \times depth, \quad depth = 1, 2, 3, 4 \quad (1)$$

식 (1)의 *depth*는, 각각의 CU와 TU가 분할하는 단계를 나타낸다. 예를 들어, 분할된 CU의 크기가 32일 경우에 *depth*는 2이며, 크기가 16일 경우에 3, 마지막으로 크기가 8일 때 4의 값을 갖는다. 이렇게 제작된 각각의 2차원 분할 지도는 GoP 단위로 묶어 하나의 3차원 특징 벡터로 구성한다. 이 두 특징 벡터를 각각의 채널로 이어 붙여 4차원의 구조를 갖는 입력 데이터로 제작한다. 아래 그림 5는 추출된 CU의 정보로부터 제작된 2차원 분할 지도를 나타낸 것이다. 이 때, 각각의 분할 지도는 입력영상과 동일한 크기를 갖지만, 시스템의 메모리 및 계산 복잡도를 줄이기 위해 제안 방식에서는 입력영상을 줄여 사용하였다. 예를 들어 CU와 TU의 경우 그 분할 범위의 최소 크기가 $N \times N$ 이라고 한다면 실제 $H \times W$ 의 입력영상을 $(H/N) \times (W/N)$ 으로 줄

여 사용해도 그 정보를 온전히 유지하는 것이 가능하다. 따라서 이렇게 제작된 입력 데이터는 아래 그림 6처럼 구성된 3차원 합성곱 신경망을 이용하여 특징 추출 단계를 거친다. 제안하는 네트워크는 마지막에 연결된 완전연결 층에 의해 조작 여부를 출력하게 된다. 그림 6의 3차원 합성곱 신경망에서 처음 두 층을 제외하고는 더 이상 정보를 줄이지 않기 위해 풀링을 수행하지 않았다. 맥스 풀링과 3×3 크기의 커널을 사용하였으며, 계산과정의 간소화를 위해 ReLU (Rectifier Linear Unit) 함수를 활성화 함수로 적용하였다.

IV. 실험 결과

1. 실험 환경

시간적 특성을 이용하기 위해 GOP단위로 데이터 셋을 구성하였다. 또한 합성곱 신경망을 사용하여 실험을 수행하기 위해 많은 수의 동영상을 사용하였다. 본 논문에서는 MPEG 표준에서 사용하고 있는 동영상을 포함하여 총 58개의 동영상을 사용하였다¹⁷⁻¹⁹. 훈련에 사용된 동영상은 무작위로 50개이며, 평가에 사용된 동영상은 나머지 8개로 진행되었다.

일부 영상이 삭제되면 두 번째 압축에서 해당 위치의 영상의 예측 모드 타입이 변한다. 따라서 GOP 내 삭제 위치를 바꿔가며 해당 GOP 끝까지 삭제하는 방식으로 삭제 길이를 조절하였다. 예를 들어 GOP크기가 8인 동영상의 GOP내 5번째 위치에서 영상의 삭제가 발생한다면, 총 4장의 영상이 삭제되는 방식이다. 이와 같이 영상 삭제 경우를 다양하게 준비하여 여러 가지 상황에서도 강인하게 영상 조작 여부를 검출할 수 있도록 실험 환경을 조성하였다.

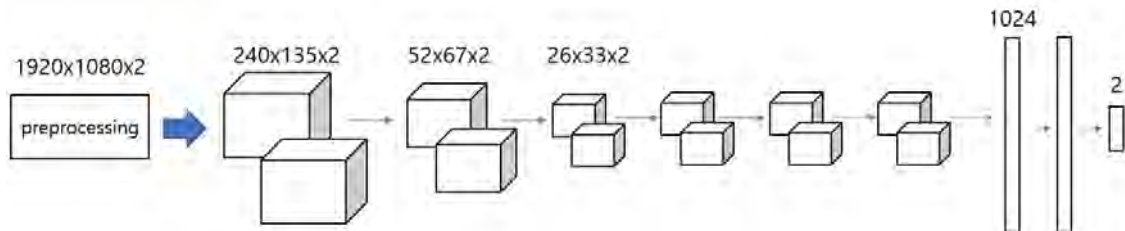


그림 6. 3차원 합성곱 신경망과 완전 연결 층의 구조
Fig. 6. 3-D Convolutional neural network and fully-connected layer

2. 성능 평가

앞서 언급한 대로, 제안방식은 GoP 내부의 계층적인 구조의 변화를 검출하는 기술이다. 그러므로, 계층적 구조가 서로 다른 저-지연 P 모드 (Low Delay P mode)와 임의의 접근 모드 (Random Access mode)로 압축된 두 경우를 각각 실험하여 결과를 도출하였다. 보다 객관적인 성능 평가를 위해 아래의 표 1과 같이 분류 분야에서 일반적으로 사용하는 성능 평가 지수인 혼동행렬을 사용하였다. 표 1의 결과, 제안하는 알고리즘은 두 가지 부호화 모드 모두 참-긍정 (True-positive) 과 참-부정(True-Negative) 수치가 매우 높은 것을 확인할 수 있다.

표 1. 동영상 삭제 여부 판단 결과
Table 1. Results for the detection of frame deletion

Actual	Low Delay P mode		Random Access mode	
	Positive	Negative	Positive	Negative
Positive	312	4	442	6
Negative	5	311	5	443

또한 혼동행렬을 이용하여 결과 성능 수치를 비교하기 쉬운 지표인 F1-score를 계산하였다. 자세한 F1-score 계산 과정은 아래 (2)-(4)과 같다.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1\text{-score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

이 때, TP, FP, FN은 각각 참-긍정, 거짓-긍정, 거짓-부정으로, 참-긍정의 경우 예측 결과가 긍정일 때 실제 결과와 동일한 것을 나타낸다. 거짓-긍정의 경우 예측 결과가 긍정일 때 실제 결과와 다른 것을 나타낸다. 거짓-부정은 예측 결과가 거짓일 때, 실제 결과와 다른 것을 나타내는 수치이다. 아래 표 2의 결과, 저-지연 P 모드의 재현율 (Recall)은 0.9873, 정확률 (Precision)은 0.9842, F1-score는 0.9857이

다. 그리고, 임의의 접근 모드의 재현율은 0.9866, 정확률은 0.9888, F1-score는 0.9877이다. 기존의 [16] 방식은 저-지연 모드에서 정확률 0.9741, 재현율 0.9328, F1-score는 0.9530으로 제안하는 알고리즘의 성능이 더 높은 것을 확인할 수 있다. [16]의 결과가 본 논문에서 제안하는 알고리즘의 성능에 비해 낮은 것은, 합성곱 신경망을 이용했기 때문이다. [16]의 네트워크 구조에서는 일반적인 다중 퍼셉트론 (Multi-layer Perceptron: MLP) 구조를 구성하였다. 하지만, 제안하는 알고리즘의 합성곱 신경망은, 분할 정보만으로 제작된 분할 지도 데이터로부터 특징을 추출하는 역할을 한다. 그 뒤, 완전 연결 층(Fully Connected Layer)을 통해 이진분류가 수행된다. 그러나 [16]의 성능이 제안하는 알고리즘의 성능과 큰 차이를 보이지 않는다. 이는 인공적으로 설계한 특징의 분류와, 합성곱 신경망으로 제작된 특징의 분류가 크게 차이 나지 않는 것을 의미한다. 결국, 인공적으로 설계된 특징의 분류 결과를 통해 해당하는 특징의 적합성이 높으나 합성곱 신경망을 통해 학습된 특징보다 조금 낮은 성능을 보이는 것을 알 수 있다. 따라서 제안하는 알고리즘은 HEVC 부호화 정보의 지역적 정보를 잘 이용하여 고해상도 동영상의 조작 검출을 모드에 강인하게 효과적으로 검출하고 있음을 알 수 있다.

표 2. 동영상 삭제 검출성능 비교
Table 2. Comparisons of the methods for the detection of frame deletion

Method	proposed		[16]
	Low Delay P mode	Random Access mode	Low Delay P mode
Precision	0.9842	0.9888	0.9741
Recall	0.9873	0.9866	0.9328
F1-score	0.9857	0.9877	0.9530

V. 결론

본 논문에서는 HEVC로 압축된 고해상도 동영상의 삭제 조작 여부와 이중 압축 여부를 판단하기 위해 HEVC의 부호화 계수인 CU와 TU를 추출하여, 동영상 압축의 계층적인 구조와 일부 영상 삭제 시 발생하는 지역적인 특징을

이용하기 위해 2차원 분할 지도를 제작하였다. 이렇게 제작된 2차원 분할지도를 시간적 특성을 이용하기 위해 GoP 단위로 분류를 진행하였다. 제안된 3차원 합성곱 신경망을 통해 동영상의 삭제 여부와 이중 압축 여부를 판단하는 알고리즘을 제안하였다. 실험 결과 계층적인 특징이 다른 두 가지 모드에 대해 강인하면서도 고해상도의 동영상의 조작 여부를 효과적으로 판단함을 보였다. 추후 연구에서는 보다 일반적인 상황에서 동영상의 조작 여부를 판단할 수 있도록 제한된 정보만을 이용하여 동영상의 조작 여부를 판단하는 알고리즘에 대한 연구를 진행할 예정이다.

참 고 문 헌 (References)

- [1] S. Milani, M. Fontani, P. Bestagini, M. Barni, A. Piva, M. Tagliascchi, and S. Tubaro, "An overview on video forensics," *APSIPA Transactions on Signal and Information Processing*, 1, 2012.
- [2] T. Shanableh, "No-reference PSNR identification of MPEG video using spectral regression and reduced model polynomial networks," *IEEE Signal Processing Letters*, 17.8, 2010.
- [3] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, 1.2: 205-214, 2006.
- [4] W. Wang, and H. Farid, "Exposing digital forgeries in video by detecting double MPEG compression," In: *Proceedings of the 8th workshop on Multimedia and security*. ACM, 37-47, 2006.
- [5] W. Wang, and H. Farid, "Exposing digital forgeries in video by detecting double quantization," In: *Proceedings of the 11th ACM workshop on Multimedia and security*. ACM, 39-48, 2009.
- [6] Y. Su and J. Xu, "Detection of double-compression in MPEG-2 videos," *Intelligent Systems and Applications (ISA)*, 2010 2nd International Workshop on. IEEE, 1-4, 2010.
- [7] T. Shanableh, "Prediction of structural similarity index of compressed video at a macroblock level," *IEEE Signal Processing Letters*, May, 18.5, 2011
- [8] D. Vazquez-Padin, M. Fontani, T. Bianchi, P. Comesaña, A. Piva, and M. Barni, "Detection of video double encoding with GOP size estimation," In: *Information Forensics and Security (WIFS)*, 2012 IEEE International Workshop on. IEEE, 151-156, 2012.
- [9] A. Gironi, M. Fontani, T. Bianchi, A. Piva, and M. Barni, "A video forensic technique for detecting frame deletion and insertion," In: *Acoustics, Speech and Signal Processing (ICASSP)*, 2014 IEEE International Conference on. IEEE, 6226-6230, 2014.
- [10] Q. Dong, G. Yang, and N. Zhu, "A MCEA based passive forensics scheme for detecting frame-based video tampering," *Digital Investigation*, 9.2: 151-159, 2012.
- [11] Y. Su, J. Zhang, J. Liu, "Exposing digital video forgery by detecting motion-compensated edge artifact," In: *Computational Intelligence and Software Engineering*, 2009. CiSE 2009. International Conference on. IEEE, 1-4, 2009.
- [12] P. He, X. Jiang, T. Sun, S. Wang, B. Li, and Y. Dong, "Frame-wise detection of relocated I-frames in double compressed H.264 videos based on convolutional neural network," *Journal of Visual Communication and Image Representation* 48: 149-158, 2017.
- [13] X. Jiang, W. Wang, T. Sun, YQ. Shi, and S. Wang, "Detection of double compression in MPEG-4 videos based on Markov statistics," *IEEE Signal Processing Letters*, 20.5: 447-450, 2013.
- [14] T. Shanableh, "Detection of frame deletion for digital video forensics," *Digital Investigation*, 10.4: 350-360, 2013.
- [15] L. Yu, H. Wang, Q. Han, X. Niu, SM. Yiu, J. Fang, and Z. Wang, "Exposing frame deletion by detecting abrupt changes in video streams," *Neurocomputing*, 205: 84-91, 2016.
- [16] J. H. Hong, Y. Yang, B. T. Oh, "Detection of frame deletion using coding pattern analysis," *Journal of Broadcast Engineering*, 22.6, 734-743, 2017.
- [17] <https://www.cdvl.org/find-videos/index.php>
- [18] <https://media.xiph.org/video/derf/>
- [19] http://ivc.univ-nantes.fr/en/databases/1080i_Videos/

저 자 소 개



홍진형

- 2017년 2월 : 한국항공대학교 전자및항공전자공학 학사
- 2017년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 석사과정
- ORCID : <http://orcid.org/0000-0003-1887-1704>
- 주관심분야 : 영상처리, 영상포렌식, 기계학습

저 자 소 개



양 윤 모

- 2015년 2월 : 한국항공대학교 전자및항공전자공학 학사
- 2015년 3월 ~ 2017년 2월 : 한국항공대학교 항공전자정보공학과 석사
- 2017년 3월 ~ 현재 : 한국항공대학교 항공전자정보공학과 박사과정
- ORCID : <http://orcid.org/0000-0003-2816-1685>
- 주관심분야 : 영상처리, 영상포렌식, 기계학습



오 병 태

- 2003년 8월 : 연세대학교 전기전자공학부 학사
- 2009년 8월 : Univ. of Southern California (USC), Dept. of Electrical Eng. 석사 및 박사
- 2009년 ~ 2013년 : 삼성종합기술원 전문연구원
- 2013년 ~ 현재 : 한국항공대학교 항공전자정보공학부 부교수
- ORCID : <http://orcid.org/0000-0003-1437-2422>
- 주관심분야 : 영상처리, 비디오압축, 영상포렌식, 3차원영상