정확도 향상을 위한 머신러닝 기반 불법복제 영상 탐지 시스템

강 신 한*, 김 정 진, 오 정 훈
(Sin-Han Kang*, Jeong-Jin Kim, Jeong-Hoon Oh)

한국항공대학교 소프트웨어 전공

본 논문에서는 머신러닝 기법을 이용하여 편집이 심한 경우에도 불법복제 동영상 탐지확률이 높은 시스템을 제시하고 이를 구현하여 결과를 보인다. 제안하는 불법복제 시스템은 저작권(원본) 동영상에서 뽑은 프레임으로부터 편집을 가해 트레이닝데이터를 생성한다. 그리고 같은 프레임으로부터 나온 학습데이터간의 유사성(공통 특징)을 뽑아내기 위해 PCA을 적용 후 K-means를 통해 같은 프레임으로부터 나온 학습데이터들의 평균값을 계산해낸다. 이 평균값은 테스팅영상에서 나온 프레임들과 비교대상이 되어 거리가 가까우면 트레이닝 학습데이터와 비슷한 프레임이므로 불법복제 영상으로 탐지되는 시스템이다.

핵심용어: 불법복제 영상, 머신러닝, 이미지 프로세싱, 탐지 어플리케이션

I. 서 론

1-1. 연구 동기 및 필요성

최근에 미디어 영상의 이용률이 급증과 함께 영상 저작권 문제도 급증하고 있다. 2012년 기준 방송 및 영화 영상을 불법복제해 생긴 피해액만 약 9000억이 된다(한국저작권위원회 통계). 이후 점차적으로 미디어 영상의 사용은 나날이 증가되므로 현재 불법복제 영상 피해액은 9000억 이상이 될 것이다. 이런 피해를 기술적으로 막기 위해 불법복제 탐지 기술은 이전부터 존재해왔지만 현재 불법복제 영상 탐지를 피하기 위해 유포자들은 영상에 다양하고 심한 편집을 가하는 방법을 사용한다. 그래서 우리는 머신 러닝을 이용해서 영상에서 의미 있는 부분만을 뽑아 이를 이용하여 저작권 영상에 다양한 편집을 가 했더라도 탐지율 높은 시스템의 연구가 필요하다.

1-2. 연구 목적

이 연구의 목적은 머신 러닝을 기반으로 한 불법 복제 영상 탐지 시스템을 개발하는 것이다. 세부적으로 다음과 같은 3가지 목적을 위한 어플리케이션 개발을 주도하였다. 첫째, 저작권에 등록된 원본영상을 그대로 복제하여 올린 영상뿐만 아니라 간단한 편집, 다양하고 심한 편집이 가해 졌다고 하더라도 정확도 높은 시스템을 개발하는 것이 목적이다. 둘째, 머신 러닝 기법을 이용해서 이미지의 의미있는 특징을 분명하게 찾아내어 기존의 불법복제 영상 탐지 영상 기술과 다르게 다양한 편집이 들어간 영상에서도 탐지 가능한 시스템 구현이 목적이다.

책임저자: 강 신한 (kangsinhan@nate.com)
한국 항공대학교 소프트웨어 전공 4학년

II. 불법 복제 영상 탐지

2-1. 불법 복제 영상

브 및 다양한 미디어 사이트에서 현재 저작권이 있는 영상을 그대로 올리지 않고 저작권을 피하기 위해 다양한 편집기술을 이용하여 기존에 불법복제영상 탐지 기술을 우회하고 있다. 현재 불법복제 탐지 기술을 우회하기 위해 심하게 밝기를 높이기도 하고 줄이기도 하고 모자이크를 삽입 또는 의미없는 모형 삽입, 로고 삽입 등 여러가지 편집기술을 사용하여 우회하려고 하고 있다. 그래서 탐지 기술도 발전해 가지만 이를 우회하려고 하는 기술 또한 발전해가면서 좀 더 분명하게 탐지 해낼 수 있는 모델이 필요하다.

2-2. 불법 복제 영상 기술 현황

현재 기존의 불법 복제 영상 기술 탐지 기술로는 텍스트 필터링, 워터마크 기법, 블록 히스토그램 및 동적 매칭을 이용한 기술이 있다. '텍스트 필터링' 기술은 제목 또는 확장자의 제한을 두어 복제 컨텐츠를 공유할 시 제한을 두는 방식이다. 하지만 이 방식은 제목의 경우 띄어쓰기, 특수문자 삽입, 다른 단어로의 치환을 통해서 우회할 수 있고, 확장자 제한의 경우 다른 파일 형식으로 변환하거나 완전히 다른 확장자로 올리더라도 다운로드 후 다시 원래의 확장자로 변경하면 파일이 원래 그대로 인식을 할 수 있는 방법으로 피할 수 있다. '워터 마크 기법'는 단순히 화면에 이미지를 추가하는 수준에서 벗어나 각 프레임 색상 비트 단위에서 워터 마킹 데이터를 추가 후 암호화하는 식으로 복잡하게 이루어진 방식도 있다. 이 기술은 마킹 데이터의 손실이 없는 선에서 데이터가 변경 될 경우 탐지 불가능한 단점을 가진다. 마지막으로 '블록 히스토그램 및 동적 매칭'을 이용한 기술은 블록 히스토그램 방식으로서 동영상의 정보를 생성하고 이를 비교하는 방식으로 복제 동영상을 탐지한다. 프레임 추출을 위해 부분 동영상을 생성한다. 부분 동영상이란 화면의 내용이 바뀌는 지점, 장면 전환점을 기준으로 나눈 동영상을 말하며, 이 부분 동영상에서 일정한 간격으로 프레임을 추출하였다. 그 후 한 프레임을 나눈 각 블록의 색상 히스토그램을 생성한다. 색상 변환이나 삽입이 이루어지면 모든 히스토그램에서 일정한 변화가 생기기 때문에 이러한 변화를 없애도록 히스토그램의 교집합을 제거한다. 이 논문에서는 테스트 영상의 편집기술이 간단하다는 면에서 현재 다양하고 심한 편집을 가한 테스트 영상에서는 탐지 못한다는 단점을 가진다. 그림 1은 이 논문에서 사용한 편집 기술이다.

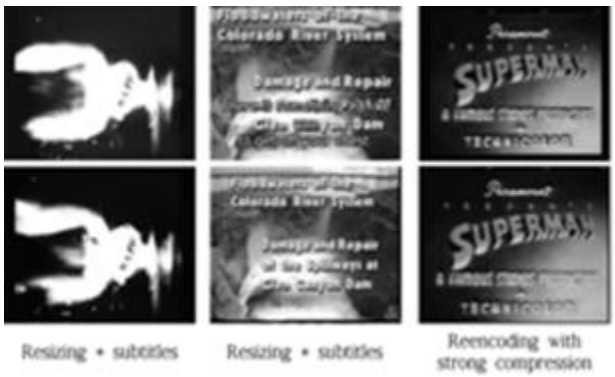

그림 1. 블록 히스토그램 기술을 이용한 논문에서 사용한 편집 기술

2-3. 머신 러닝 기반 불법복제 영상 탐지

이에 비해 본 연구는 머신 러닝이라는 기법을 이용하여 영상에 의미있는 부분을 분명하게 찾아 낼 수 있다는 장점을 근거하여 연구를 진행하였다. 컴퓨터가 트레이닝 데이터를 통해 학습을 하여 인간이 찾아내야 했던 불법복제 영상을 대신해서 찾아 주는 시스템을 만들도록 한다.

III. 제안하는 머신 러닝 기반 불법복제 영상 탐지

연구에서는 불법 복제 영상 탐지를 위해 다음과 같은 트레이닝 , 테스팅 모듈을 나누어 설명한다.

1) 트레이닝 모듈
   : 원본 영상으로부터 특정 프레임들을 추출하여 이를 대상으로 트레이닝 데이터를 만들어 학습시킨다.
2) 테스팅 모듈
   : 트레이닝 모듈에서 나온 결과물과 질의 영상의 프레임들과 비교하여 불법복제 영상인지를 판단한다.

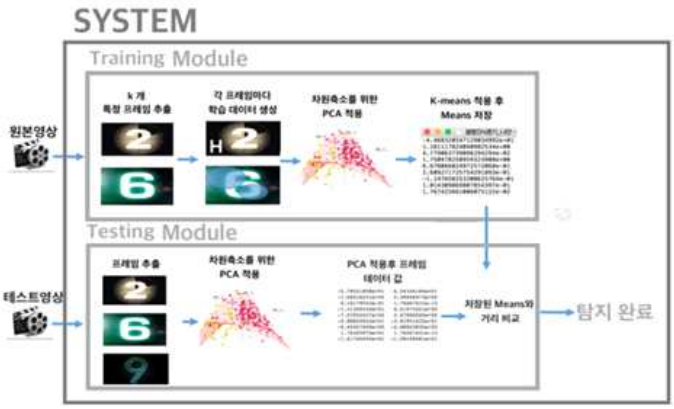

그림 2. 제안하는 시스템 개념도

먼저 트레이닝 모듈에서 k개(실험에서는 5개로 잡음)의 특정 프레임을 추출하게 된다. 이 프레임들은 학습데이터를 만들기위한 데이터가 되어 프레임 각각에 편집기술을 이용하여 학습데이터를 만들게 된다. 그리고 그 데이터들은 차원축소 및 특징 추출을 위해 PCA를 거치게 되고 이후 데이터들은 K-means clustering 알고리즘을 이용해서 clustering되고 그 clustering의 중심점인 평균값이 트레이닝의 산출물이 된다.

그 다음 테스팅 모듈에서 원본영상의 프레임을 시간순으로 뽑아 내어 PCA를 적용하게 되고 이 데이터값들은 트레이닝 모듈에서 나온 산출물인 Means와 거리비교를 하게되어 탐지하게 된다.

3-1 트레이닝 모듈

저작권이 등록된 원본 영상으로부터 랜덤으로 프레임을 뽑는다. 3분에서 10분사이의 영상을 기준으로 영상안의 5개의 프레임을 랜덤으로 추출하게 된다. 5개의 프레임은 트레이닝 데이터를 만드는데 사용되는 기준 데이터가 되어 이 5개의 프레임을 기반으로 트레이닝 데이터를 만들게 된다. 5개의 프레임 각각 여러 편집기술을 통

해 변형을 가하여 3000개 이상의 트레이닝 데이터를 만든다. JAVA의 기본 API와 JCodec API를 이용하여 MP4 영상의 프레임을 추출하였으며 추출 후 편집 기술로는 그림 3과 같이 8개의 종류로 나누었다.

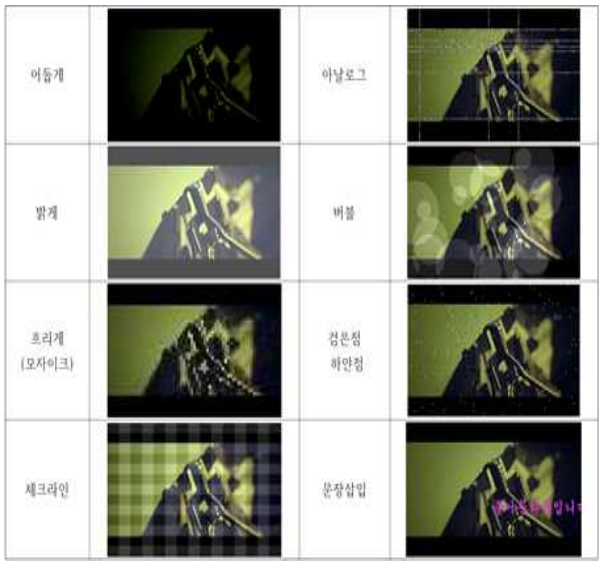

그림 3. 학습데이터 생성을 위한 편집 기술

원본 프레임으로 그림 3과 같은 편집 효과를 랜덤 위치, 크기로 추출하였다. 원본 파일의 RGB값의 정보를 읽어 변경 후 이미지를 저장하여 추출된 각 프레임당 3000개의 트레이닝 데이터를 만들었다.

총 15000개의 이미지 데이터에 대해서 학습(트레이닝)시키는 것은 계산량이 많아지게 되면서 트레이닝 시간이 길다는 점에서 머신 러닝 기법인 PCA를 사용한다. PCA를 사용함으로써 이미지의 주요 특징 추출을 하게 되고 차원축소를 통해 이미지 데이터크기가 줄어든다. PCA 기법은 component 갯수에 따라 이미지의 특징 추출이 얼마나 정교하게 될지 안될지 정해지는 파라미터가 된다. 즉, component갯수는 이미지의 특징추출 갯수와 비례하게된다. 다음 그림 4 는 PCA component 갯수를 5,25,125개로 정해 놓고 이미지에 PCA적용한 후 다시 그 특징으로 이미지를 재생성해놓았을 때의 데이터이다.

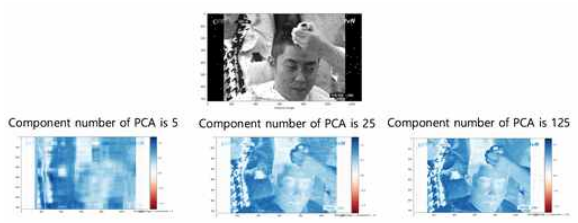

그림 4. PCA 개수에 따른 이미지 특징 추출 정도

그림 4에서 볼 수 있듯이 component갯수에 따라 원본 이미지가 PCA 적용 후 특징을 얼마나 잘 보존하는 지 알 수 있다. 그래서 component갯수가 크게 할 경우 원본 영상의 이미지를 특징을 정확하게 잡아내어 원본 이미지와 거의 다름이 없는 것을 확인가능하고 component갯수를 너무 작게하면 원본 이미지의 특징을 거의 추출해내지 못하여 원본이미지와 확연히 다르다는 것을 확인가능하다. 우리는 PCA를 사용함으로써 편집 기술을 가했더라도 PCA component 갯수를 적절히 잡아 이미지를 뭉개어 어떤 편집을 가했더라도 다 비슷한 이미지 데이터로 만들어 주어 이후 K-means clustering을 할 때 데이터 군집이 잘 되도록 한다. 그래서 군집이 잘되도록 하기위해서는 적절한 component갯수를 목적이다. 우리는 실험적으로 최적의 component갯수를 찾아 내었다.

이후 적절한 component갯수를 찾아 PCA를 적용 후에 K-means를 적용하게 된다. K-means에서 K는 프레임 추출 갯수와 동일하다. 그래서 각각 5개의 프레임에서 추출된 3000개이상의 트레이닝 데이터들이 같은 군집내에 속하도록 한다. K-means 알고리즘을 통해 각 군집내의 평균 값을 뽑아낼 수 있다는 장점을 살려 이 평균값들은 트레이닝 모듈의 마지막 산출물이 되고 테스트 모듈에서 테스트(질의) 영상에서 나온 프레임들과 비교 대상이 된다.

3-2 테스팅 모듈

테스트 모듈에는 테스트 할 영상이 시스템에 주어지게 되고 테스트 영상에서 JCodec을 이용하여 영상의 모든 프레임을 시간 순차적으로 추출하게 된다. 먼저 시간순차적으로 프레임하나가 추출되었다면 그 프레임은 트레이닝 모듈에서와 같은 component갯수를 가지고 PCA를 적용하고 그 데이터와 트레이닝 모듈에서 나온 산출물인 평균값 들과 유클리디안 거리 비교를 하게 된다.테스트 영상에서 나온 프레임의 데이터값이 평균값들과 거리 비교를 했을 때 차이가 크게 없다면 비슷한 데이터로 판단되어 불법복제 영상으로 탐지 되는 것이고 프레임의 데이터값이 평균값들과 차이가 크다면 다른 이미지 데이터로 판단되어 다음 프레임을 추출하고 반복적으로 위의 과정을 반복한다. 그래서 모든 프레임이 뽑힐 때까지 비교했을 때 평균값과 거리비교를 했을 때 비교값이 작은 프레임이 없다면 결국 불법복제 영상이 아님으로 탐지가 된다.

IV. 결과고찰

4. 성능측정

테스트 영상에 사용된 편집기술은 그림 1 과 같고 테스트 영상에는 일부러 트레이닝 데이터 즉, 트레이닝 프레임이 들어간 영상을 부분 추출하였다. 실험적으로 PCA 갯수에 따라 불법복제 영상 탐지 정확도를 계산한 결과는 다음 그림 5 와 같다.

PCA component갯수가 30일 때 가장 좋은 성능을 내었으며 이는 기존논문 '블록 히스토그램을 이용한 불법복제 영상탐지 논문' 의 성능측정에서 True positive가 92%

라는 점에서 우리가 제안하는 모델과 정확도 성능은 비슷한 우리의 연구에서는 다양하고 심한 편집에서도 정확도가 높다는 점에서 가치를 가진다.

PCA 개수	10	20	30	40	50
True Positive	94%	90%	92%	84%	78%
False Positive	6%	10%	8%	16%	22%
True Negative	66%	78%	80%	83%	86%
False Negative	34%	22%	20%	17%	14%

True Positive = 불법복제영상을 불법복제영상으로 탐지
False Positive = 불법복제영상을 불법복제 영상이 아닌것으로 탐지
True Negative = 불법복제영상이 아닌것을 불법복제 영상이 아닌것으로 탐지
False Negative = 불법복제영상이 아닌것을 불법복제 영상으로 탐지

그림 5. PCA component 개수에 따른 정확도 비교

V. 결 론

5-1. 결론 및 향후 연구 계획

제안하는 머신 러닝 기반의 불법복제 영상 탐지 시스템은 불법복제영상을 탐지해내는 데 있어서 최적의 PCA component 갯수를 찾아내었고 K-means알고리즘은 주로 clustering을 위해 사용되지만 우리는 K-means에서 means를 이용해 시스템을 구현해내었다는 독창성을 가진다. 향후에는 정확도를 높히기 위해 트레이닝 데이터의 수를 늘리거나 시계열 분석을 통해 불법 복제로 탐지된 프레임 순서가 원본 영상에서 뽑힌 프레임 순서와 동일한 지를 확인하는 방법을 계획중이다. 그래서 불법 복제로 탐지된 프레임 순서가 트레이닝 모듈에서 뽑힌 프레임 순서와 같다는 것을 확인하여 True Negative의 정확도 개선을 할 예정이다. 또한 현재 우리 모델은 원본영상의 크기를 줄이고 뒷 배경을 다른 이미지로 채우는 편집기술에 대해 탐지가 불가능 하다는 점에서 테스트 영상의 프레임을 블록으로 나누어 테스팅하도록 한다.

5-2. 기대 효과

제안하는 불법 복제 영상 탐지 기술은 영상 이미지에 우리가 트레이닝 시킨 이미지가 포함되어 있는지 없는지 확인하는 방법이다. 이 방법과 GAN과 같은 머신러닝 기법을 같이 이용한다면 범죄자의 얼굴이나 대포 차량과 같은 이미지를 가지고 트레이닝 시킨후 테스팅으로 cctv에 그 이미지가 발견되면 탐지되는 기술로 발전 될 가능성이 있다. 그래서 현재 경찰에서 범죄자나 대포차량을 사람이 직접찾아 확인하지만 이 기술이 구현된다면 cctv를 통해 실시간으로 범죄자나 대포 차량이 어디에 나타나는 지 확인가능하다. 이는 범죄검거율이 더 높아질것이며 인간이 수작업으로 해야할 일을 대신해서 해준다는 면에서 효율성을 가진다.

참 고 문 헌

[1] 김주섭, 남제호 "불법 복제 콘텐츠 필터링 기술 동향 분석" 방송공학회지 제12권 제4호, 2007.12
[2] 임여선, 배건태, 임용, 어정, 변혜란 "블록 히스토그램 및 동적 매칭을 이용한 중복 동영상의 빠른 검출" 정보과학회논문지 : 소프트웨어 및 응용 제 40 권 제 2 호, 2013.2

강 신 한 (Sin-Han Kang) 1995년 6월 23일생

2014년 2월 : 한국 항공대학교 소프트웨어 전공 (공학사)
관심 분야 : 머신러닝, 컴퓨터 보안
특 기 : 파이썬, C, 머신러닝

김 정 진 (Jeong-Jin Kim) 1992년 7월 9일생

2011년 2월 : 한국 항공대학교 소프트웨어 전공 (공학사)
관심 분야 : 안드로이드, 웹
특 기 : 자바, SQL

오 정 훈 (Jeong-Hoon Oh) 1993년 6월 3일생

2012년 2월 : 한국 항공대학교 소프트웨어 전공 (공학사)
관심 분야 : 파이썬, 이미지 처리
특 기 : 자바