

논문 2022-17-37

웨이블릿 영역에서 회전 불변 에너지를 이용한 이중 브랜치 복사-이동 조작 검출 네트워크

(Dual Branched Copy-Move Forgery Detection Network Using Rotation Invariant Energy in Wavelet Domain)

박 준 영, 이 상 인, 엄 일 규*

(Jun Young Park, Sang In Lee, Il Kyu Eom)

Abstract : In this paper, we propose a machine learning-based copy-move forgery detection network with dual branches. Because the rotation or scaling operation is frequently involved in copy-move forger, the conventional convolutional neural network is not effectively applied in detecting copy-move tampering. Therefore, we divide the input into rotation-invariant and scaling-invariant features based on the wavelet coefficients. Each of the features is input to different branches having the same structure, and is fused in the combination module. Each branch comprises feature extraction, correlation, and mask decoder modules. In the proposed network, VGG16 is used for the feature extraction module. To check similarity of features generated by the feature extraction module, the conventional correlation module used. Finally, the mask decoder model is applied to develop a pixel-level localization map. We perform experiments on test dataset and compare the proposed method with state-of-the-art tampering localization methods. The results demonstrate that the proposed scheme outperforms the existing approaches.

Keywords : Copy-move forgery, Copy-move forgery localization, Dual branched network, Convolutional neural network, Rotation-invariant, Wavelet transform.

1. 서 론

영상은 때때로 어떤 사건을 명확히 하거나 검증하기 위한 중요한 정보로 사용된다. 그러나 다양한 영상 편집 툴의 발달로 영상의 조작이 쉬워졌다. 또한, 악의적 목적으로 조작된 영상을 사용하는 것은 사회에 부정적인 영향을 미칠 수 있다. 인간의 눈은 조작된 영상을 쉽게 감지하기가 어렵기 때문에, 신뢰할 수 있는 영상 조작 감지 방법이 필요하다. 이를 위해, 다양한 영상 위조 검출 방법에 관한 광범위한 연구가 수행되고 있다 [1, 2].

영상 조작에는 여러 가지 방법이 존재하지만, 하나의 영상 안에서 일정 영역을 복사하여 같은 영상의 다른 영역으로 붙이는 복사-이동 조작이 빈번히 사용된다. 배경을 복사하여 특정한 물체를 숨기거나, 물체를 여러 개로 보이기 위해 복사하는 방법이 흔히 사용된다. 그림 1은 복사-이동 조작 영상의 예를 나타내고 있다. 복사-이동 조작을 통하여, 영상이 가지는 정보나 의미가 숨겨지거나 바뀔 수 있으므로, 영상의 진위를 확인하고 복사-이동 조작영역을 찾아내는 것이 중요하다.

복사-이동 조작검출의 방법은 크게 블록, 키포인트, 기계



그림 1. 복사-이동 조작 이미지의 예
Fig. 1. Example of copy-move forgery image

학습 기반 방법으로 나눌 수 있다. 블록 기반 방법은 주로 다양한 블록분할과 영역분할 알고리즘을 통하여 블록 내부에서 크기와 회전 변환에 불변하는 특징들을 추출한다. 그리고 이러한 특징들을 블록 사이의 매칭을 통하여 조작된 부분을 찾아낸다. 키포인트 방법은 SIFT (scale invariant feature transform)와 같이 크기와 회전 변환에 강한 포인트들을 추출하여 매칭을 통해 조작영역을 검출한다. 최근 CNN (convolutional neural network)의 객체 감지 분야에서 보여주는 좋은 성능을 바탕으로 다양한 기계학습 기반의 복사-이동 조작검출의 방법들이 도입되어 유망한 탐지 결과를 보여주었다.

블록 기반의 복사-이동 조작검출 방법은 의심되는 이미지를 블록분할, 특징 추출, 추출된 특징을 매칭, 위조영역을 국부화하는 4단계로 이루어진다. 첫 번째 단계는 다양한 블록을 나누는 방법과 영역분할 방법을 이용하는 전처리 단계이다. 정사각형 블록을 중첩시키는 방법 [3], 정사각형 블록

*Corresponding Author (ikeom@pusan.ac.kr)
Received: Oct. 13, 2022, Revised: Nov. 22, 2022, Accepted: Nov. 26, 2022.
J. Y. Park: Pusan National University (Ph.D. Student)
S. I. Lee: Pusan National University (M.S. Student)
I. K. Eom: Pusan National University (Prof.)
※ 본 논문은 부산대학교 기본연구지원사업 (2년)에 의하여 연구되었음.

을 증첩시키지 않는 방법 [4] 또는 원형 블록 [5]의 방식을 사용할 수 있다. 특징 추출은 블록 기반의 방법에서 핵심적인 부분이다. 추출된 특징은 크기와 회전에 대하여 불변해야 하고, 잡음, 블러 등 여러 후처리에 대하여 견고해야 한다. 이 단계에서는 극좌표 코사인 변환 [6], 푸리에-멜린 변환 [7], 극복수 지수변환 [8], 푸리에, 이산코사인 [9], 웨이블릿 변환 [10] 등 여러 가지 변환들이 불필요한 정보 제거 및 특징 추출에 사용된다. 특징 매칭은 추출된 특징을 이용하여 원본 부분과 복사-이동 조작된 부분의 후보 쌍을 결정한다. 이 단계에서는 탐색과 유사성 측정 방법들을 이용한다. 탐색에는 다양한 정렬 방법 [10, 11] 및 해싱 프로세스 [12]가 사용된다. 복사-이동 조작의 마지막 단계는 조작영역의 국부화이다. 탐지 결과는 의심되는 영상에서 소스 부분과 조작된 영상 부분이 이진 맵으로 표현된다. 대부분의 블록 기반의 방법은 특징 수가 너무 많고 아핀 변환에 대해 약한 단점이 있다. 또한, 영상의 블록 부분의 값을 이용하기 때문에 블러와 뒤집는 변환에 대해 약한 단점이 있다.

복사-이동 조작에서 키포인트에 관련된 연구는 활발하게 진행되었다. SIFT는 크기 변환과 회전 변환에 강한 키포인트를 뽑아내기 때문에 복사-이동 조작영역 검출에 적합하다 [13-17]. SIFT 기반 복사-이동 조작영역 검출의 첫 번째 단계는 영상 피라미드에 의한 스케일 공간을 통하여 다양한 스케일에서 추출되는 스케일 불변의 키포인트를 생성한다. 두 번째 단계는 키포인트의 회전 불변성을 위하여 키포인트에 방향을 지정하고 주변 픽셀을 128차원의 설명자로 표현한다. 매칭 과정에서 키포인트의 설명자는 다른 모든 키포인트의 설명자와 비교되어 매칭 여부를 결정한다. 하지만 이러한 매칭에서 잘못된 매칭은 항상 발생하기 때문에 이러한 매칭을 제거해주기 위해 J-linkage [18], 거리 기반 [19] 및 계층적 클러스터링 [20]과 같은 다양한 클러스터링 알고리즘이 사용된다. 복사-이동 조작 부분이 작거나 배경 영역을 포함하는 경우 SIFT 기반의 알고리즘은 성능이 저하될 수 있다 [15, 17]. 또한, SIFT 기반의 방식은 키포인트 주변의 광역 정보를 얻는 것은 어렵다. 따라서 복사-이동 조작된 부분과 원본의 경계면에 키포인트가 존재하거나, 영상을 압축하는 경우 성능이 저하될 수 있다.

최근 기계학습은 영상 처리 및 컴퓨터 비전 작업에서 획기적인 성능을 보였다. 따라서 기계학습 기반의 복사-이동 조작 연구도 활발히 이루어졌다. 대표적으로 Wu 등은 Simi-Det 및 Mani-Det을 포함하는 이중 병렬 구조의 CNN 방식 (BusterNet) [21]을 제안했다. Simi-Det은 유사성 감지를 위해 설계되었으며 Mani-Det은 원본과 조작영역 구별을 위해 설계되었다. 그러나 각 BusterNet의 구조가 조작영역을 정확하게 찾지 못하면 BusterNet은 본연의 목적인 원본 지역과 대상 지역을 구분할 수 없다. Zhong 등은 다차원 밀집 특징 연결 네트워크인 Dense-InceptionNet을 사용하는 복사-이동 조작 검출 알고리즘 [22]을 제안했다. Dense-InceptionNet은 피라미드 특징 추출기, 특징 상관 매칭 및 계층적 후처리 모듈로 구성된다. 네트워크가 유사한 역할을 하는 여러 모듈을 사용하여 탐지 정확도를 향상했지

만, 훈련 매개변수의 수가 급격히 증가하여 계산적 비용이 크다는 단점이 있다. Chen 등은 유사성 및 조작 감지를 위해 직렬 구조의 복사-이동 조작 검출 방법 [23]을 제안했다. 이 방법은 VGG16 [24]의 최종 컨볼루션 계층에서 atrous convolution [25]을 사용하였고 계층적 특징 비교를 위해 상관 모듈에서 이중 상관계수를 사용했다. 또한, 조작영역의 위치를 정확하게 검출하기 위한 attention mechanism을 제안하였다. 마지막으로 영상 레벨 네트워크를 사용하여 유사성 탐지 구조에서 얻은 영역이 변조되었는지를 결정했다. 실험 결과는 제안한 알고리즘이 BusterNet보다 우수한 성능을 달성했음을 보여주었다. 그러나 이 알고리즘은 여전히 성능 면에서 개선될 필요가 있다. 최근에 Lee 등은 회전 불변 특성을 가지는 웨이블릿 고주파 계수의 에너지를 이용한 조작 검출 CNN 구조 [26]를 제안하였다. 이 방법은 특히 회전 변환이 추가된 복사-이동 검출 조작에 대해 우수한 성능을 보였다.

본 논문에서는 이중 브랜치 구조를 가지는 복사-이동 조작 검출 네트워크를 제안한다. 제안된 네트워크는 SWT (stationary wavelet transform) 기반의 특징 분리 모듈, CNN을 사용한 특징 추출 모듈, 유사성을 확인하는 상관분석 모듈, 이진 검출 맵의 크기를 증가시키는 디코딩 모듈, 이중 가지 구조의 결과 결합을 위한 결합 모듈 총 5가지 모듈로 구성된다. 제안된 네트워크의 구조는 입력으로 RGB 채널의 색 정보를 활용하는 것보다 웨이블릿 계수를 활용하였기 때문에 크기 조정 및 회전에 더 적합한 입력이라 할 수 있다. 실험 결과를 통하여 제안된 네트워크의 입력 및 구조가 기존 방법들에 비하여 우수한 복사-이동 조작 위치 검출 결과를 나타냄을 보이고자 한다.

논문의 나머지 부분은 다음과 같이 구성된다. 2장에서는 제안 방법의 구조에 대해 설명한다. 3장에서는 제안 방법의 결과를 다른 방법들과 비교하고 4장에서는 결론을 서술한다.

II. 제안 방법

본 논문에서 제안하는 복사-이동 검출 네트워크는 영상의 스케일 정보를 가진 웨이블릿 저주파 밴드의 계수와 회전-불변 정보를 가진 고주파 계수의 에너지 정보를 각각 학습시켜 결합하는 이중 가지 구조를 기반으로 한다. 네트워크 구조는 특징 분리 모듈, 특징 추출 모듈, 상관계수 모듈, 디코딩 모듈로 구성되고 최종적으로 각 구조에서 나온 결과를 합치는 결합 모듈로 구성된다.

그림 2는 제안된 복사-이동 검출 네트워크의 구조를 나타낸다. 특징 분리 모듈은 복사-이동 조작에서 발생할 수 있는 스케일과 회전 변환을 검출하기 위한 것이고, 두 개의 특징 추출 모듈에서 각각의 특징들을 추출한다. 상관분석 모듈은 추출된 특징들의 유사성을 판단하여 복사-이동된 영역의 후보자는 선택한다. 디코딩 모듈은 입력 영상과 같은 크기의 복사-이동 조작 영상의 상세한 정보를 복원하기 위해 사용되며, 마지막으로 결합 모듈에서 회전 변환과 스케

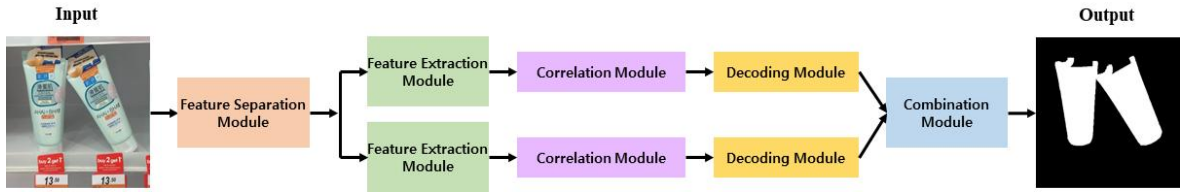


그림 2. 제안된 네트워크 구조
Fig. 2. Proposed network architecture

일 변환이 가해진 조작 영상에 대한 복사-이동 영역을 결합하여 최종적인 검출 결과를 얻는다.

1. 특징 분리 모듈

기존의 CNN은 다양한 크기의 필터를 사용하여 컨볼루션이 수행되기 때문에 다양한 스케일 특징을 추출할 수 있다. 그러나 회전-불변성은 기본 CNN 구조에서 과생될 수 없기 때문에, CNN 회전 불변성을 만들기 위한 다양한 시도가 있었다 [27]. 논문에서는 CNN에 회전 불변성을 제공하기 위해 웨이블릿 변환을 사용하여 스케일 특징을 추출하고, 웨이블릿 계수의 에너지를 이용하여 회전-불변 특징을 추출하는 방법을 제안한다.

어떤 스케일 $i(i=0,1,2,\dots)$ 에서 주어진 저주파 부밴드 영상 $W_{i,LL}$ 입력 영상을 웨이블릿 변환하면 다음과 같다.

$$W_{i+1,o} = SWT(W_{i,LL}). \quad (1)$$

여기서 $SWT(z)$ 는 부밴드 영상 z 의 웨이블릿 변환이고, $o(o=LL, LH, HL, HH)$ 는 부밴드의 방향을 나타낸다. LL 는 저주파 부대역을 나타내고, LH, HL 및 HH 는 각각 수평, 수직 및 대각 방향을 가지는 고주파 부밴드를 나타낸다. $i=0$ 는 입력 영상을 표현한다.

제안된 네트워크에서 사용되는 특징 분리 모듈은 그림 3과 같이 구성된다. 특징을 분리하기 전 먼저, 입력 영상을 256×256 크기로 변환한다. 입력 영상에 대하여 Haar 필터에 기반한 2단계 변환을 통해 두 개의 저주파 부밴드 $W_{1,LL}$ 와 $W_{2,LL}$ 을 생성한다. 저주파 부밴드와 함께 형성된 고주파 부밴드 계수를 이용하여 본 논문에서는 회전-불변 특성을 생성하기 위하여 에너지 특징을 사용한다. 제안된 에너지 특징은 고주파 부밴드 에너지의 평균 제공근으로 다음과 같이 정의된다. 즉,

$$E_i(x,y) = \sqrt{(W_{i,LH}^2(x,y) + W_{i,HL}^2(x,y) + W_{i,HH}^2(x,y))/3}. \quad (2)$$

여기서 $E_i(x,y)$ 는 i -단계 웨이블릿 부밴드의 (x,y) 위치에서의 평균 에너지의 제공근이다. 그리고 $W_{i,LL}^2(x,y)$, $W_{i,HL}^2(x,y)$, 및 $W_{i,HH}^2(x,y)$ 는 각각 (x,y) 위치에서 수평, 수직 및 대각 성분 에너지값을 나타낸다. 회전 변환이 생기더라도 웨이블릿 부밴드의 전체 에너지는 변하지 않으므로 식 (2)의 에너지 특징은 회전 변환에 강인하게 작용할 수 있다.

그림 3은 특징 분리 모듈을 나타낸다. 웨이블릿 통해 얻어

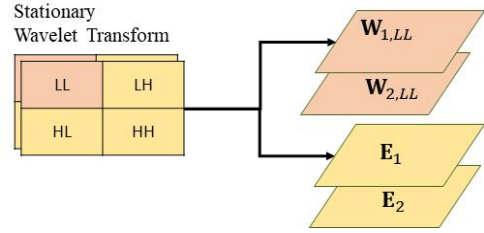


그림 3. 특징 분리 모듈
Fig. 3. Feature separation module

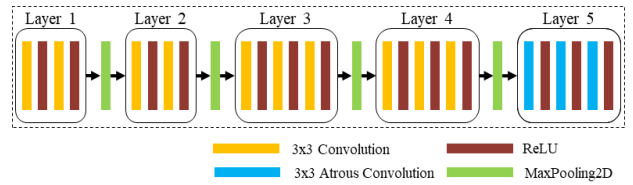


그림 4. 특징 추출 모듈
Fig. 4. Feature extraction module

낸 두 개의 저주파 부밴드 $W_{1,LL}$ 및 $W_{2,LL}$ 은 스케일이 다른 영상을 얻을 수 있어서, 복사-이동 조작에서 복사 영역의 스케일 변화의 특징을 추출할 수 있다. 또한, 고주파 부밴드의 에너지 특징은 회전 변환에 강인하다. 이러한 특징을 활용하여 본 논문에서는 이중 가치를 가지는 네트워크를 구성하여 더욱 강한 복사-이동 조작 검출 결과를 얻고자 한다.

2. 특징 추출 모듈

VGGNet [24]은 적은 학습 파라미터 수와 3×3 컨볼루션 필터만으로도 강력한 성능을 보여준다. 또한, 기존의 CNN 기반 복사-이동 조작 검출에서 VGG16을 이용한 방법들이 사용되었다 [21, 23, 26]. 본 논문 또한 기존 방법과의 성능 비교를 위해 다양한 백본 네트워크 중 VGG16을 특징 추출 모듈 사용하였다. 본 논문에서는 저주파 부밴드 및 웨이블릿 에너지를 입력으로 하는 두 개의 특징 추출 모듈을 그림 2와 같이 병렬로 사용한다. 즉, 하나의 VGG16 특징 추출 모듈에는 웨이블릿 저주파 부밴드 입력 $W_{1,LL}$ 와 $W_{2,LL}$ 가 사용되고, 또 다른 VGG16 특징 추출 모듈에는 E_1 및 E_2 가 입력으로 사용된다.

그림 4는 특징 추출 모듈의 구성을 나타낸다. 각 모듈은 다섯 개의 층으로 이루어져 있으며 3×3 크기의 필터, ReLU 활성화 함수를 교차로 사용하면서 영상의 특징을 넓

은 영역에서 찾아낸다. 또한, 특징 맵의 크기를 줄이고 파라미터의 수를 줄이면서 특징을 추출하기 위해 2×2 크기의 최대 풀링을 이용한다. 모든 층의 구성은 기존의 VGG16 구조와 비슷하나 마지막 층은 수용 영역의 증가를 위해 atrous convolution [25]을 이용해 특징 맵을 분석한다.

3. 상관분석 모듈

두 가지의 특징 추출 모듈에서 얻어낸 스케일 및 회전 변환에 강인한 특징 맵은 복사-이동된 영역의 유사도의 점수를 얻기 위해 상관분석 모듈에 입력된다. 특징 추출 모듈에서 추출된 두 개의 특징 맵을 각각 f_i 및 f_j 벡터로 표현하면, 두 개의 특징 벡터에 대한 상관계수 $\rho(i, j)$ 는 다음과 같다.

$$\rho(i, j) = \frac{\tilde{f}_i^T \tilde{f}_j}{512}. \quad (3)$$

여기서 \tilde{f}_i 는 정규화된 특징 벡터로서 다음과 같이 계산된다. 즉,

$$\tilde{f}_i = \frac{f_i - \mu(f_i)}{\sigma(f_i)}. \quad (4)$$

여기서 $\mu(f_i)$ 및 $\sigma(f_i)$ 는 f_i 는 각각 f_i 의 평균 및 표준편차이다.

그림 5는 제안된 네트워크의 상관분석 모듈을 나타낸다. 상관분석 모듈은 BusterNet [21]의 구성과 유사하지만 각 채널의 집중하는 CAM (channel attention module) [28]을 사용하는 대신 조작된 화소 위치에 집중하기 위하여 특징 맵 내의 특정 화소에 중요한 정보가 있는지 집중하는 모듈인 SAM (spatial attention module) [28]을 이용하였다. 복사-이동 조작된 영역의 화소 간의 유사 점수를 파악하기 위해 BusterNet [21]에서 이용한 피어슨 상관관계수에 의한 자기 상관도 방법을 이용한다. 유사 점수를 내림차순으로 정렬하고 필요한 정보를 추출하는 percentile pooling을 이용하여 유사 점수가 높은 상위 100개의 특징 맵을 추출한다. 마지막으로 층이 깊어지면서 발생할 문제를 대비하기 위해 배치 정규화를 진행한다.

4. 디코딩 모듈

입력 영상이 특징 추출 모듈과 상관분석 모듈을 거치며 16×16 크기의 낮은 해상도를 가진 복사-이동 조작된 영역에 대한 대략적인 정보를 가진 특징 맵이 생성된다. 입력 영상과 같은 크기의 복사-이동 조작 영상의 상세한 정보를 가진 특징 맵으로 복원하기 위해 디코딩 모듈을 적용한다. 그림 6은 디코딩 모듈을 나타낸다.

영상은 다양한 보간법에 따라 확대될 수 있다. 본 논문에서는 쌍선형 보간법을 이용해 16배의 확대 과정을 거친다. 기존의 방법 [21]과 같이 BN-Inception을 모든 확대 과정에 사용하는 작은 적은 영역에 있는 오류를 확대시키는 경향이 있다. 따라서 본 논문에서는 전체의 확대 과정이 아니라 128×128 과 256×256 크기의 확대 과정에만 BN-Inception을



그림 5. 상관분석 모듈
Fig. 5. Correlation module

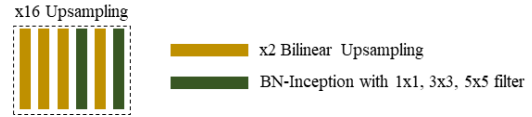


그림 6. 디코딩 모듈
Fig. 6. Decoding module

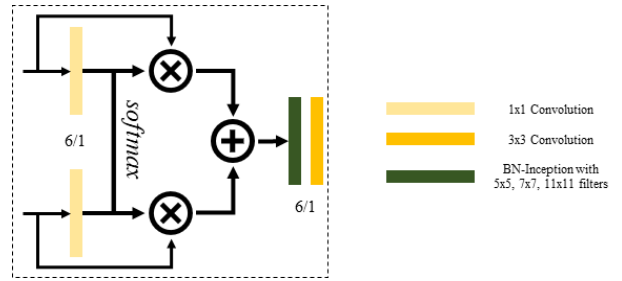


그림 7. 결합 모듈
Fig. 7. Combination module

적용하였다. 이러한 방법은 제대로 검출된 영역이 확대 과정에서 다소 부정확할 수 있지만, 마지막의 결합 모듈에서 개선될 수 있으므로 제안된 디코딩 모듈에서 쌍선형 보간법을 더 많이 사용하였다.

5. 결합 모듈

지금까지의 과정으로 저주파 부밴드를 사용한 스케일 특징과 고주파 에너지를 이용한 회전-불변 특징을 이용하여 복사-이동 조작된 영상의 특징 맵 두 개를 생성하였다. 각각의 가지들은 서로 다른 특징으로 학습되었기 때문에 두 맵을 결합함으로써 두 가지 특징의 장점을 가진 맵을 만들 수 있다.

그림 7은 이를 위한 모듈의 구성을 나타낸다. 그림 7에서 보는 바와 같이, 각 특징 맵의 화소 정보를 유지하고 채널의 수를 줄이기 위해 1×1 필터 한 개로 연산을 진행한다. 생성된 각 특징은 화소 단위의 softmax를 통해 각 가중치를 정한다. 디코딩 모듈에서 얻어낸 특징 맵에 가중치를 화소별로 곱해준 후 BN-Inception을 적용하여 최종적인 복사-이동 영역 이진 맵을 생성한다.

6. 손실 함수

본 논문에서는 학습을 위한 손실 함수로 이진 교차 엔트로피를 다음과 같이 사용한다.

$$L = \frac{1}{m} \sum_{k=1}^m [y_k^t \log(y_k) + (1 - y_k^t) \log(1 - y_k^t)]. \quad (5)$$

여기서 L 은 손실 값, m 은 학습 영상의 개수, y_k^t 는 k 번째 실제 영상 그리고 y_k 는 k 번째 결과 영상이다.

III. 실험 및 결과

1. 실험 데이터

본 실험에서는 기존의 기계학습 기반 복사-이동 조작 검출 방법 [21, 23, 26]에서 사용되었던 것과 마찬가지로 100,000장의 테스트 영상 (조작 및 원본 각 100,000장씩)을 사용하였다. 이 영상들은 8: 1: 1의 비율로 나누어져 각각 훈련, 검증, 및 테스트에 사용되었다.

2. 실험 환경 및 세부 사항

본 실험은 학습 및 훈련을 TensorFlow v2.4.0 환경에서 진행하였다. 최적화 기법으로는 0.001의 학습률 (learning rate)을 가진 Adam (adaptive moment estimation)을 사용하였다. 배치 크기는 16, 에포크는 100회로 설정하였다. 또한, 제안 방법의 정확한 성능 평가를 위하여 교차 검증을 10회 실시하였다. 학습 환경은 Intel i7-11700K @ 3.60 GHz CPU, NVIDIA GeForce RTX 3080 Ti 12GB GPU, 64GB RAM, Windows 10 OS 기기를 사용하였다.

3. 성능 평가

검출 성능 평가 지표로써 화소 단위에서 Precision, Recall 및 F1-score를 사용하였다. 평가 지표를 계산하기 위한 TP, FP 및 FN은 각각 실제 정답을 정답이라 예측한 횟수, 실제 오답을 정답이라 예측한 횟수 그리고 실제 정답을 오답이라 예측한 횟수를 말한다. Precision은 예측 결과 정답 중 실제 정답의 비율, Recall은 실제 정답 중 예측 결과가 정답인 비율이며, F1-Score는 Precision과 Recall의 조화 평균이다. 위의 세 가지 성능 평가 지표는 다음과 같다.

$$\text{Precision} = \frac{T_p}{T_p + F_p}, \quad (6)$$

$$\text{Recall} = \frac{T_p}{T_p + F_N}, \quad (7)$$

$$\text{F1} = 2 \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (8)$$

4. 실험 결과

제안된 복사-이동 검출 네트워크의 성능을 평가하기 위하여 회전 변환이 없는 테스트 영상 집합, 회전 변환이 있는 영상 집합을 구분하여 결과를 도출하였다. 성능 비교를 위해 세 가지의 SIFT 기반 방법들 즉, SIFT+FM (SIFT and feature matching) [13], SIFT+Seg (SIFT and adaptive segmentation) [16] 및 SIFT+RLBP (SIFT and reduced local binary pattern) [17]을 비교하였다. 또한, BusterNet

표 1. 회전 변환이 없는 테스트 영상 집합에 대한 복사-이동 조작 검출 성능 비교

Table 1. Performance comparison of copy-move forgery detection for test images without rotation attack

| Method | Precision | Recall | F1-Score |
|----------------|-----------|--------|----------|
| SIFT+FM [13] | 0.105 | 0.040 | 0.054 |
| SIFT+Seg [16] | 0.249 | 0.028 | 0.045 |
| SIFT+RLBP [17] | 0.426 | 0.358 | 0.370 |
| BusterNet [21] | 0.608 | 0.508 | 0.511 |
| SeNet [23] | 0.725 | 0.655 | 0.657 |
| RIWCNN [26] | 0.494 | 0.609 | 0.507 |
| Proposed | 0.534 | 0.576 | 0.531 |

표 2. 회전 변환이 있는 테스트 영상 집합에 대한 복사-이동 조작 검출 성능 비교

Table 2. Performance comparison of copy-move forgery detection for test images with rotation attack

| Method | Precision | Recall | F1-Score |
|----------------|-----------|--------|----------|
| SIFT+FM [13] | 0.119 | 0.049 | 0.067 |
| SIFT+Seg [16] | 0.346 | 0.073 | 0.108 |
| SIFT+RLBP [17] | 0.359 | 0.321 | 0.326 |
| BusterNet [21] | 0.467 | 0.481 | 0.420 |
| SeNet [23] | 0.495 | 0.421 | 0.412 |
| RIWCNN [26] | 0.791 | 0.770 | 0.765 |
| Proposed | 0.817 | 0.802 | 0.801 |

표 3. 모든 테스트 영상 집합에 대한 복사-이동 조작 검출 성능 비교

Table 3. Performance comparison of copy-move forgery detection for all test images

| Method | Precision | Recall | F1-Score |
|----------------|-----------|--------|----------|
| SIFT+FM [13] | 0.112 | 0.045 | 0.061 |
| SIFT+Seg [16] | 0.302 | 0.053 | 0.079 |
| SIFT+RLBP [17] | 0.389 | 0.338 | 0.346 |
| BusterNet [21] | 0.531 | 0.493 | 0.461 |
| SeNet [23] | 0.599 | 0.527 | 0.522 |
| RIWCNN [26] | 0.657 | 0.697 | 0.648 |
| Proposed | 0.689 | 0.700 | 0.679 |

[21], SeNet (serialized network) [23]과 RIWCNN (rotation invariant wavelet CNN) [26] 등의 세 가지 기계학습 기반 방법의 결과를 비교하였다.

표 1은 회전 변환이 없는 조작 영상에 대한 성능 평가 지표를 나타낸다. 표 1에서 볼 수 있는 바와 같이, SIFT 기반 방법들은 F1-Score가 모두 0.4 이하로 대체로 낮은 성능을 보인다. SeNet이 가장 높은 F1-Score 0.657을 도출하고 있고, 제안 방법은 두 번째로 높은 F1-Score 0.531을 가진다. 결론적으로 SeNet은 회전 변환이 없는 조작 영상의 검출에 대하여 제안 방법보다 우수한 결과를 보임을 알 수 있다. 그림 8은 회전 변환이 없는 조작 영상의 검출 결과의 예를 나타낸다.

표 2는 회전과 스케일 변환이 동시에 있는 조작 영상에 대한 검출 결과를 나타낸다. 표 2에서 보는 바와 같이, 회전



그림 8. 회전 변환이 없는 테스트 영상에 대한 복사-이동 조작 영역 검출 결과
 Fig. 8. Copy-move forgery localization results for test images without rotation attack

변환이 없는 경우 0.657의 F1-Score를 가졌던 SeNet은 회전 변환이 있는 경우 F1-Score가 0.412로 매우 감소하였다. 반면에, 회전 불변성을 고려한 RIWCNN 및 제안 방법은 F1-Score가 각각 0.507 및 0.531에서 0.765 및 0.801로 상당히 증가하였다. SIFT 기반의 세 가지 검출 방법은 회전 변환이 있는 경우 모두 F1-Score가 감소하였음을 알 수 있다. 따라서 복사-이동 조작 검출에 대하여 회전 변환에 강인한 특징을 사용한 방법들은 회전 변환에 상당히 강인한 결과를 보인다는 것을 알 수 있다. 회전 불변에 대한 고려의 반대급부로 스케일 변환만 있을 때, 성능이 떨어지는 단점이 있지만, 일반적으로 복사-이동 조작을 할 경우 회전 변환을 많이 사용하기 때문에, 회전 불변성을 고려하는 것이 전체적인 검출 성능을 높이는데 효과적이라 할 수 있다. 그림 9는 회전 변환과 스케일 변환 조작이 동시에 가해진 경우의 조작 영상에 대한 검출 결과의 예를 나타낸다.

표 3은 전체 테스트 영상에 대한 검출 성능을 종합한 것이다. 표 3에서 볼 수 있는 바와 같이, 제안 방법의 F1-Score가 0.679로 가장 높으며, RIWCNN과 SeNet이 각

각 0.648 및 0.522로 두 번째 및 세 번째로 높은 F1-Score를 가진다. 또한, 제안 방법의 성능은 예측 결과 정답 중 실제 정답의 비율을 표현하는 Precision이 0.689이고, 실제 정답 중 예측 결과가 정답인 비율을 나타내는 Recall 값이 0.700이기 때문에, 다른 모든 방법들에 비하여 좋은 Precision 및 Recall의 균형을 이루고 있다.

IV. 결론

본 논문에서는 이중 브랜치를 가지는 CNN을 이용한 복사-이동 조작 검출 방법을 제안하였다. 복사-이동 조작을 할 때 회전 또는 스케일 변환이 일반적으로 동시에 이루어지기 때문에, 기존의 CNN은 조작을 검출하기에 효과적이지 않다. 따라서 제안 방법은 웨이블릿 계수를 회전에 불변한 부밴드와 스케일 변환에 불변한 특징으로 나누었다. 각각의 특징들을 동일한 구조를 갖는 서로 다른 브랜치에 입력하여, 결합 모듈에서 융합하였다. 각 브랜치는 특징 추출, 상관분석, 디코딩 모듈로 구성되며, 각 브랜치의 최종 디코딩 모듈

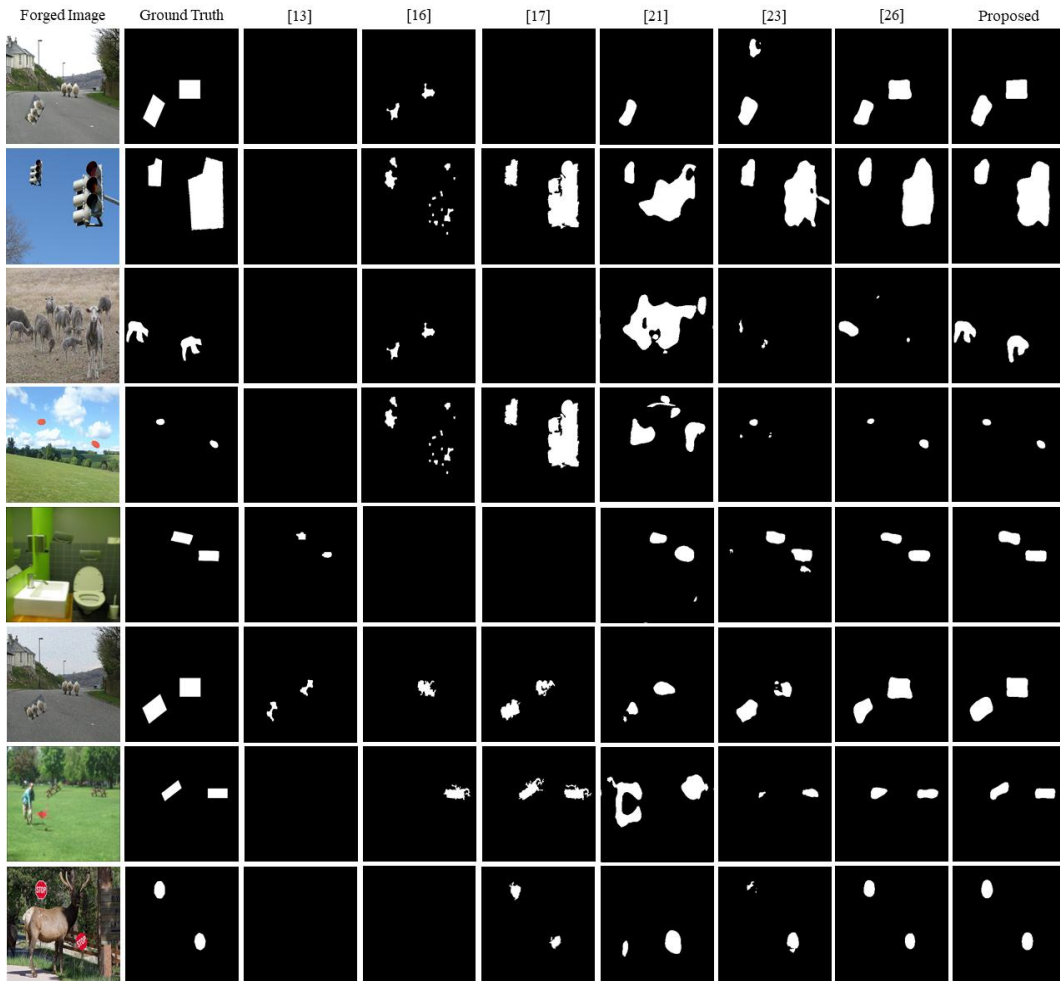


그림 9. 회전 변환이 있는 테스트 영상에 대한 복사-이동 조작 영역 검출 결과
 Fig. 9. Copy-move forgery localization results for test images with rotation attack

의 결과는 결합 모듈에서 융합된다. 융합된 결과로부터 복사-이동 조작된 영역을 추출하였다. 제안 방법의 성능을 평가하기 위하여 기존의 복사-이동 검출 방법들과 비교를 수행하였다. 실험 결과 제안된 네트워크는 회전 변환이 없는 조작 영상에 대하여 두 번째로 높은 F1-Score를 얻었고, 회전 변환이 가해진 조작 영상에 대해서는 가장 높은 성능을 보였다. 종합적으로 전체 테스트 영상에 대하여 가장 높은 Precision, Recall 및 F1-Score 값을 얻었다.

References

[1] S. Teerakanok, T. Uehara, "Copy-move Forgery Detection: A State-of-the-art Technical Review and Analysis," *IEEE Access.*, Vol. 7, pp. 40500-40568, 2019.
 [2] R. Thakur, R. Rohilla, "Recent Advances in Digital Image Manipulation Detection Techniques: A Brief Review," *Forensic Sci. Int.*, Vol. 312, 110311, 2020.
 [3] G. Lynch, F. Y. Shih, H. Y. Liao, "An Efficient Expanding Block Algorithm for Image Copy-move

Forgery Detection," *Inf. Sci.*, Vol. 239, pp. 2253-265, 2013.
 [4] Y. Sun, R. Ni, Y. Zhao, "Nonoverlapping Blocks Based Copy-move Forgery Detection," *Secur. Commun. Netw.*, Vol. 2018, Article ID 1301289, 2018.
 [5] Y. Gao, T. Gao, L. Fan, Q. Yang, "A Robust Detection Algorithm for Copy-move Forgery in Digital Images," *Forensic Sci. Int.*, Vol. 214, No. 1, pp. 33-43, 2011.
 [6] Y. Li, "Image Copy-move Forgery Detection Based on Polar Cosine Transform and Approximate Nearest Neighbor Searching," *Forensic Sci. Int.*, Vol. 224, No. 1-3, pp. 59-67, 2013.
 [7] R. Dixit, R. Nakar, "Copy -move Forgery Detection Utilizing Fourier - Mellin Transform Log-polar Features," *J. Electron. Imaging.*, Vol. 27, No. 2, 023007, 2018.
 [8] K. M. Hosny, H. M. Hamza, N. A. Lashin, "Copy-move Forgery Detection of Duplicated Objects Using Accurate PCET Moments and Morphological Operators," *Imag. Sci. J.*, Vol. 66, No. 6, pp. 330-345, 2018.
 [9] M. H. Alkawas, G. Sulong, T. Sabas, A. Rehman, "Detection of Copy-move Image Forgery Based on

- Discrete Cosine Transform,” *Neural Comput. Appl.*, Vol. 30, pp. 183–192, 2016.
- [10] T. Mahmood, A. Irtaza, Z. Mehmood, M. T. Mahmood, “Copy - move Forgery Detection Through Stationary Wavelets and Local Binary Pattern Variance for Forensic Analysis in Digital Images,” *Forensic Sci. Int.*, Vol. 279, pp. 8–21, 2017.
- [11] M. Zandi, A. Mahmoudi-Aznavah, A. Talebpour, “Iterative Copy-move Forgery Detection Based on a New Interest Point Detector,” *IEEE Trans. Inf. Forensics Secur.*, Vol. 11, No. 11, pp. 2499–2512, 2016.
- [12] C. M. Pun, J. L. Chung, “A Two-stage Localization for Copy-move Forgery Detection,” *Inf. Sci.*, Vol. 463–464, pp. 33–55, 2018.
- [13] X. Pan and S. Lyu, “Region Duplication Detection Using Image Feature Matching,” *IEEE Trans. Inf. Forensics Secur.*, Vol. 5, No. 4, pp. 857–867, Dec. 2010.
- [14] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, G. Serra, “A SIFT-based Forensic Method for Copy - move Attack Detection and Transformation Recovery,” *IEEE Trans. Inf. Forensics Secur.*, Vol. 6, No. 3, pp. 1099–1110, 2011.
- [15] V. Christlein, C. Riess, J. Jordan, C. Riess, E. Angelopoulou, “An Evaluation of Popular Copy-move Forgery Detection Approaches,” *IEEE Trans. Inf. Forensics Secur.*, Vol. 7, No. 6, pp. 1841–1854, 2012.
- [16] C. M. Pun, X. C. Yuan, X. L. Bi, “Image Forgery Detection Using Adaptive Oversegmentation and Feature Point Matching,” *IEEE Trans. Inf. Foren. Secur.*, Vol. 10, No. 8, pp. 1705–1716, 2015.
- [17] J. Y. Park, T. A. Kang, I. K. Eom, “Copy-move Forgery Detection Using Scale Invariant Feature and Reduced Local Binary Pattern Histogram,” *Symmetry*, Vol. 12, No. 4, 492, 2020.
- [18] G. Jin, X. Wan, “An Improved Method for SIFT-based Copy-move Forgery Detection Using Non-maximum Value Suppression and Optimized J-Linkage,” *Signal Process. Image Commun.*, Vol. 57, pp. 113–125, 2017.
- [19] D. M. Uliyan, H. A. Jalab, A. W. A. Wahab, S. Sadeghi, “Image Region Duplication Forgery Detection based on Angular Radial Partitioning and Harris Key-points,” *Symmetry*, Vol. 8, No. 7, pp. 62, 2016.
- [20] N. B. A. Warif, A. W. A. Wahab, M. Y. I. Idris, R. Salleh, F. Othman, “SIFT-symmetry: A Robust Detection Method for Copy-move Forgery with Refection Attack,” *J. Vis. Commun. Image Represent.*, Vol. 46, pp. 219–232, 2017.
- [21] Y. Wu, W. Abd-Almageed, P. Natarajan, “BusterNet: Detecting Copy-move Image Forgery with Source/target Localization,” *Proceedings of the European Conference on Computer Vision*, Munich, Germany, pp. 168 - 184, 2018.
- [22] J. L. Zhong, C. M. Pun, “An End-to-end Dense-InceptionNet for Image Copy-move Forgery Detection,” *IEEE Trans. Inf. Forensics Secur.*, Vol. 15, pp. 2134–2146, 2020.
- [23] B. Chen, W. Tan, G. Coatrieux, Y. Zheng, Y. Q. Shi, “A Serial Image Copy-move Forgery Localization Scheme with Source/target Distinguishment,” *IEEE Trans. Multimed.*, Vol. 23, pp. 3506–3517, 2021.
- [24] K. Simonyan, A. Zisserman, “Very Deep Convolutional Networks for Large-scale Image Recognition,” *Proceedings of 3rd International Conference on Learning Representations*, pp. 1 - 14, 2015.
- [25] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, “DeepLab: Semantic Image SSegmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 40, No. 4, pp. 834 - 848, 2018.
- [26] S. I. Lee, J. Y. Park, I. K. Eom, “CNN-based Copy-move Forgery Detection Using Rotation-invariant Wavelet Feature,” *IEEE Access*, Vol. 10, pp. 106217–106229, 2022.
- [27] G. Cheng, J. Han, P. Zhou, D. Xu, “Learning Rotation-invariant and Fisher Discriminative Convolutional Neural Networks for Object Detection,” *IEEE Trans. Image Process.*, Vol. 28, No. 1, pp. 265–278, 2019.
- [28] S. Woo, J. Park, J.-Y. Lee, I. S. Kweon, “CBAM: Convolutional Block Attention Module,” *Proceedings of European Conference on Computer Vision*, pp. 3–19, 2018.

Jun Young Park (박 중 영)



2017 Department of Electronics Engineering from Pusan National University (BS.)

2017~Department of Electronics Engineering from Pusan National University (Ph.D.)

Field of Interests: Image enhancement, Computer Vision, Copy-Move Forgery Localization & Machine Learning
Email: jyp9917140@pusan.ac.kr

Sang In Lee (이 상 인)



2021 Department of Electronics Engineering from Pusan National University (BS.)

2021~Department of Electronics Engineering from Pusan National University (MS.)

Field of Interests: Image Processing, Copy-Move Forgery Localization & Machine Learning
Email: silee96@pusan.ac.kr

Il Kyu Eom (엄 일 규)



1990 Department of Electronics Engineering from
Pusan National University (B.S.)

1992 Department of Electronics Engineering from
Pusan National University (M.S.)

1998 Department of Electronics Engineering from
Pusan National University (Ph.D.)

2006~Department of Electronics Engineering from Pusan National
University (Prof.)

Field of Interests: Image Processing, Computer Vision, Digital
Image Forensic & Machine Learning

Email: ikeom@pusan.ac.kr