딥페이크 및 사진 이미지 위변조 탐지 소프트웨어

KAICATCH 2.1

㈜디지탈이노텍

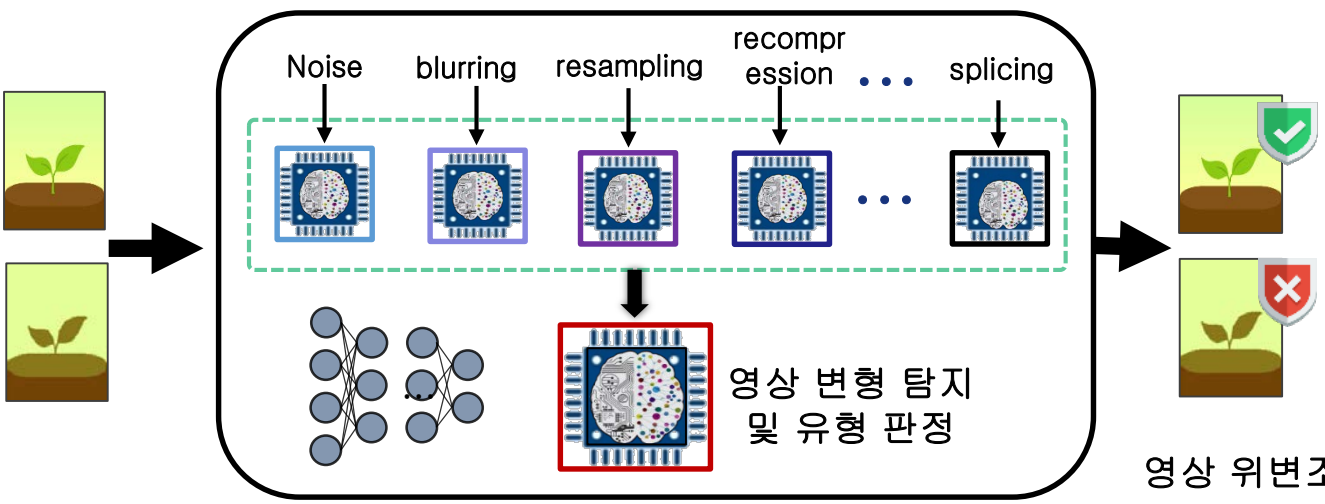
www.kaicatch.com

2022년

영상 위변조 탐지 소프트웨어 : 카이캐치 2.1

1. 영상 위변조 탐지란 ?

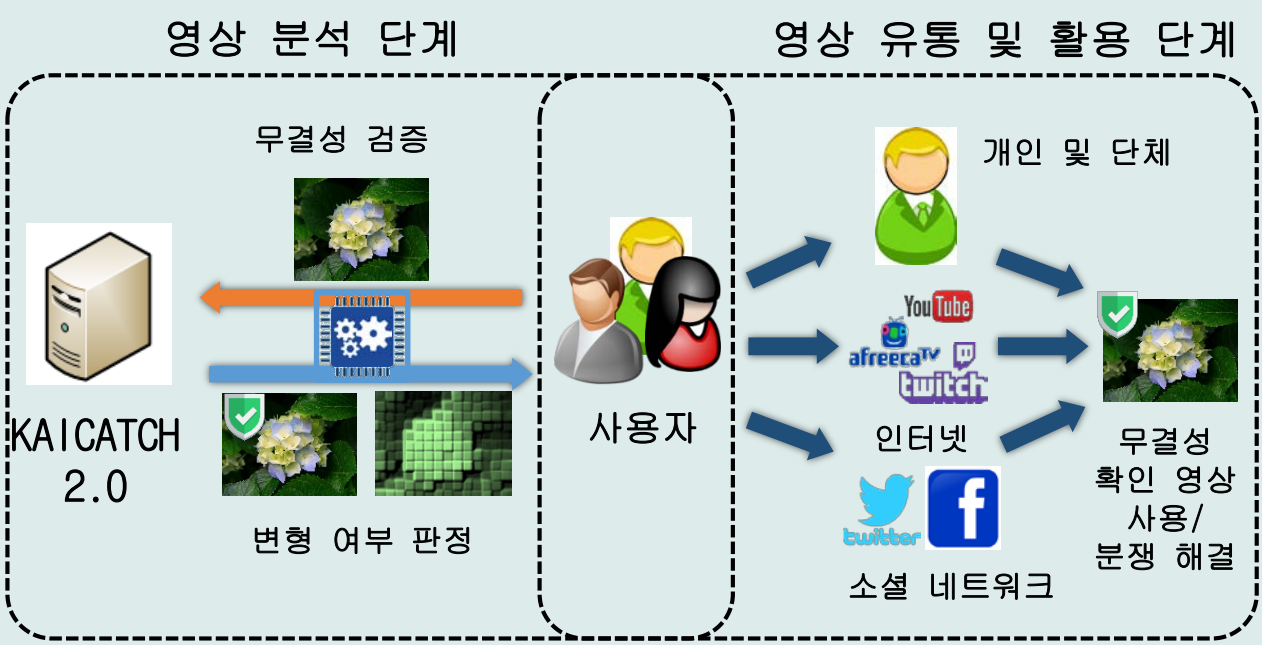
- 인공지능 기술을 이용하여 비디오내 등장하는 인물의 얼굴 교체, 얼굴 재현 및 속성 변환 등을 한 딥페이크 탐지와 디지털 사진에서 부분 및 전체 삭제, 절삭, 잘라 붙이기, 복사 붙이기, 객체 이동, 색조 변화, 구도 변화 등 이미지내 발생한 각종 변형을 탐지하는 기술을 통틀어 영상 위변조 탐지 기술이라 한다.
- 인공지능과 각종 딥페이크 생성 공개 소프트웨어, 그리고 포토샵 등 각종 영상 편집 도구들의 발달에 의해 누구나 편리하게 비디오나 사진을 편집할 수 있음으로 해서, 단순한 흥미 위주의 변형을 넘어 악의적인 변형 및 유통에 의해 사회적, 법적으로 큰 문제들을 야기시키고 개인의 행복권을 침해하는 사례들이 빈번하게 발생함에 따라 이러한 영상 위변조 탐지 기술의 필요성이 크게 증대하고 있다.



KAICATCH 2.1

2. 영상 위변조 탐지 기술 응용 분야

- 영상 위변조 탐지 기술이 필요한 기관, 단체 및 개인
  - 영상을 법 집행시 이용하는 공공기관
  - 각종 영상을 수집하고 활용하는 언론기관
  - 개인 및 기관간 분쟁을 다루는 법률회사
  - 영상을 서비스하는 콘텐츠 서비스 분야
  - 콘텐츠 솔루션 산업계
  - 각종 분쟁 해결을 위한 개인, 단체 및 기관들



카이캐치 2.1
소프트웨어 활용 예

3. 영상 위변조 탐지 기술 현황

- 영상 위변조 탐지 자체는 오랜 역사를 지니고 있다. **60~70** 년 이전부터 사진을 여러 형태로 변조, 변형하여 왔고, 전문가들은 사진내의 색감 차이, 배경 정보, 등장 인물들의 적합성 등을 활용하여 위변조 여부를 판정하여 왔다. 디지털카메라의 등장과 함께 영상들이 디지털 형태로 촬영, 저장됨에 따라 **2000**년 초부터는 본격적인 영상 포렌식 기술에 대한 연구를 수행하기 시작하여, 현재까지 수많은 영상 포렌식 관련 연구 논문들이 발표되었다.
- 포토샵 등 영상 편집도구들의 성능이 나날이 발전하여 최근에는 누구나 손쉽고, 간편하게 영상들을 편집, 수정할 수 있는 단계에 이르고 있다. 더욱이 뉴럴 네트워크에 기반한 인공지능의 등장으로 영상 변형의 수단이 고도화하여 감에 따라 딥페이크가 등장하는 등 급속하게 발전하는 **IT** 기술의 부작용 역시 점차 커지고 있다.
- 영상 포렌식은 요소 기술로 크게 분류 할 때는, 픽셀, 포맷, 광학 및 기하학 방식으로 나눌 수 있다. 픽셀 기반 기술로는 엣지, 텍스쳐, 색감 분석 등이 가능하고, 광학 기반 기술로는 명암, 그림자, 조명 분석 등이 가능하다. 기하학 기반 기술로는 형태, 크기, 위치 분석이 가능하며, 포맷 기반 기술로는 압축 횟수 등의 조사가 가능하다.
- 그러나 변형의 유형 분류가 무한할 정도로 많은데다가 인공지능에 의해 이를 모두 학습한다는 것 자체도 어렵고, 또한 인공지능에 의해 이를 회피하는 학습도 하므로 **100%** 확답을 줄 수 있는 포렌식 기술 확보를 위해서는 많은 연구가 필요하다. 영상 포렌식 기술의 현 상태는 다음과 같다:
  - 이론적으로 완벽한 기술은 아직 등장하지 않고 있다.
  - 인공지능과 신호처리 기술을 사용하여 일정 수준의 확률적인 성능을 지닌 포렌식 도구에 의한 위변조 판정 결과 제시는 비전문가들도 일정 수준 이상의 신뢰도를 갖는 분석 결과를 일차적으로 손쉽게 확보 할 수 있다는 점에서 분쟁 사전 예방이나 해결시 큰 도움을 준다.

4. KAICATCH 2.1 솔루션의 특징

- 일반을 상대로 하는 앱 기반의 딥페이크 탐지 및 사진 위변조 탐지 서비스는 국내 최초로 시도하는 서비스 기술
- 2015년에 시작한 영상 조작 탐지 웹서비스를 통해 약 30 여 만장의 실 환경 사진들을 보유하고 있으며 이를 기반으로 1,000 개가 넘는 비표준화된 양자테이블에 의한 압축 이미지들을 포함한 각종 유형의 디지털 사진 분석이 가능한 인공지능 기반 사진 위변조 탐지 실용 소프트웨어를 국내 최초 출시
- 국내 : 체계적인 영상 위변조 탐지 기술 솔루션 출시는 이루어지지 못하고 있다.
- 국외 : 미국의 Amped software, Fourandix, Hacker factor 등이나 영국의 Forensic pathway 등이 있다. 해당 기술들은 영상의 메타데이터와 JPEG 정보를 이용하여 단순한 조작 가능성을 제시하거나, 제한된 환경에서만 조작을 탐지하는 기술들이 주류를 이루고 있다.

- 본 영상 위변조 탐지 S/W 카이캐치 2.0 의 특징은 다음과 같다 :
  - ✓ 미세 변형 신호 흔적과 미세 이상 신호 흔적 탐지 기술을 적용한 신호처리 및 인공지능 기술을 이용한 딥페이크 탐지 기능
  - ✓ 특정 딥페이크 기법에 특화된 기술이 아니라 딥페이크 영상 특성 탐지를 통해 딥페이크 탐지 정확도를 크게 높임
  - ✓ 실환경에서 유통되는 다양한 유형의 사진 이미지 분석 기능
  - ✓ 무압축, 비손실 압축, JPEG 압축 사진 이미지 분석 가능
  - ✓ 표준 및 비표준 JPEG 압축 방식의 이미지 분석 가능
  - ✓ 다양한 변형에 개개 대응하는 기술이 아니라 일반 변형시 발생하는 필수 변이 탐지를 통한 변형 탐지 기능
  - ✓ 영상 편집 영역 추정 기능
  - ✓ 신호처리 변형 유형 추정 기능
  - ✓ 영역, 유무 판정이 매우 용이

5. KAICATCH 2.1 세부 기술

- 핵심 기술
  - ✓Avi나 MP4 포맷 비디오내 얼굴 딥페이크 여부 판별 가능
  - ✓딥페이크 탐지율 95% 내외
  - ✓30 여 만장의 보유 실환경 영상을 활용한 기술 (세계 최초)
    - 일반 인터넷에서 유통되는 임의의 디지털 사진 처리 가능
  - ✓실환경 고신뢰 사진 위변조 탐지/무결성 검증 기능 (세계 최초)
    - 90 ~ 98% 내외의 높은 정확도
  - ✓비디오 편집 변형 탐지 기능

- 학습 이미지 DB
  - ✓딥페이크 : DFDC, MegaFace, CelebA-HQ 등
  - ✓사진 이미지 : 실유통 이미지 DB, 스풀라이싱 이미지 DB 7종, ALASKA 스테그 챌린지 DB, 20개 이상의 필수변이 유형들에 의해 생성된 이미지 DB 등
  - ✓변형 유형에 따라 QF50-100 또는 QF75-95 사이의 이미지 분석

- 분석 가능 영상 포맷
  - ✓무압축(BMP 등), 무손실 압축(TIF, PNG 등), JPEG 압축 이미지
  - ✓50 여개의 표준 양자테이블 이외 천 여개의 비표준 양자테이블에 의한 JPEG 압축 이미지
  - ✓변형 유형에 따라 QF50-100 또는 QF75-95 이미지
  - ✓Full HD (1920×1080) 기준

- 주요 탐지 기능
  - ✓다종 신호처리 변형(블러링, 노이즈, 리샘플링, 콘트라스트 변화, 모핑 등 변형에 수반되는 필수 변이 유형 20종)의 유형 추정

5. KAICATCH 2.1 세부 기술

▪ 주요 탐지 기능

✓다종 신호처리 변형(블러링, 노이즈, 리샘플링, 콘트라스트 변화, 모핑 등 변형에 수반되는 필수 변이 유형 20종)의 유형 추정

✓정상 및 변형 여부 판별 기능

✓잘라 붙이기(스플라이싱), 복사 붙이기, 물체 이동 등 특정 영역에 가해진 영상 편집 영역 추정

▪ 주요 탐지 성능

✓신호처리 변형 유형 추정

- 평가지표 : 정확도
- 평가방법 : ALASK dataset으로 생성한 800,000개의 정상/변형 이미지로 측정
- 정상 및 변형 탐지 정확도 : 90% 내외
- 공격 유형 분류 정확도 : 94 내외

✓영상 편집 영역 추정

- 평가지표 : mIoU (mean intersection of union)
- 평가방법 : 공개 데이터셋에 따른 측정
  1) NIST16(미국 국립표준기술연구소) Dataset: 67.18
  2) Columbia(Columbia University) Dataset: 83.05
  3) Carvalho Dataset: 67.18
  4) CASIAv2(중국과학원) Dataset: 87.63

6. 보유 기술 및 분석 의뢰 절차

- 보유중인 영상 변형 탐지 인공지능 엔진
  - 딥페이크 이미지 및 사진 위변조 탐지 기본 엔진, KAICATCH 2.1
  - 사진 전체영역 및 부분영역 변형 심층 분석 엔진 (별도 엔진)
  - 스테고 미세 신호 변이 탐지 엔진 (별도 엔진)
  - 딥페이크 비디오 탐지 엔진 (별도 엔진)
  - 비디오 편집 변형 탐지 엔진 (별도 엔진) 등

- 위변조 변형 여부 (추가) 분석 의뢰 절차
  - 디지털 이미지, 사진 및 비디오 등의 (추가) 분석을 통한 변형 여부 판정이 필요한 경우,
  - 이메일, 우편 등으로 분석이 필요한 디지털 형태의 영상전달 및 분석 요청

    대표 이메일 : kaicatch@naver.com heunglee@kaist.ac.kr

    대표주소 : 대전광역시 유성구 대학로 291 한국과학기술원 전산학부 E3-1동, 1428호실, 34141 이흥규 교수

    (보내실때 디지털 영상 이외, 고객님의 이메일, 휴대폰 등 연락처 기재 요망합니다)

  - 보내주신 영상물 (추가) 분석 타당성 검토를 위한 검토후,
  - 고객님에게 분석 가능/불가능 여부와 비용을 알려 드립니다.
  - 분석 가능일 경우, 비용 선지급후 분석 결과 이메일, 카톡 등으로 전달
  - 분석 결과 및 그 사용과 책임은 ㈜디지탈이노텍 홈페이지에 있는 사용약관을 따르는 것으로 합니다.

7. 포렌식 기술 관련 보유 논문

- 국내외 저명 저널 및 학술 회의에 영상/동영상 포렌식 관련 논문을 다수 제출 하여 국내외적으로 높은 기술 수준 인정

가. 국제저널

- C. H. Choi, H. Y. Lee, H. K. Lee, “Estimation of Color Modification in Digital Images by CFA Pattern Change,” *Forensic Science International*, Vol. 226(1-3), 10 March 2013, pp. 94-105. (SCI = 4.882)
- S. J. Ryu, M. Kirchner, M. J. Lee, H. K. Lee, “Rotation Invariant Localization of Duplicated Image Regions Based on Zernike Moments,” *IEEE Trans. on Information Forensics and Security*, Vol. 8, No. 8(August 2013), pp. 1355-1370. (SCI = 7.178)
- D. K. Hyun, S. J. Ryu, H. Y. Lee, H. K. Lee, “Detection of Upscale-Crop and Partial Manipulation in Surveillance Video Based on Sensor Pattern Noise,” *Sensors*, Vol. 13, No. 9(September 18 2013), 12605-12631. (SCI = 3.576)
- J. H. Choi, H. Y. Lee, H. K. Lee, "Color Laser Printer Forensics based on Noisy Feature and Support Vector Machine Classifier," *Multimedia Tools and Applications,* Vol. 67, No. 2(November, 2013), pp 363-382. (SCI = 2.757)
- S. J. Ryu, H. K. Lee, “Estimation of Linear Transformation by Analyzing the Periodicity of Interpolation,” *Pattern Recognition Letters*, Vol. 36, 15 January 2014, pp. 89-99. (SCI = 3.756)
- D. G. Kim, H. K. Lee, “Color Laser Printer Identification Using Halftone Texture Fingerprint,” *Electronics Letters*, Vol. 51, No. 13(25 June 2015), pp. 981-983. (SCI = 1.314)
- D. J. Jung, D. K. Hyun, H. K. Lee, “Recaptured Video Detection based on Sensor Pattern Noise,” *EURASIP Journal on Image and Video Processing*, 3 December 2015, 2015 : 40, 1:14. (SCI = 1.789)

가. 국제저널(계속)

- J. S. Park, D. K. Hyun, J. U. Hou, D. G. Kim, H. K. Lee, "Detecting Digital Image Forgery in Near-Infrared Image of CCTV," *Multimedia Tools and Applications*, Vol. 76, No. 14(July 2017), DOI: 10.1007/s11042-016-3871-7, pp. 15817-15838. (SCI = 2.757)
- J. U. Hou, H. K. Lee, "Detection of Hue Modification Using Photo Response on Nonuniformity," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 27, No. 8(August 2017), pp. 1826-1832. (SCI= 4.685)
- H. Y. Choi, D. K. Hyun, S. Choi, H. K. Lee "Enhanced Resampling Detection based on Image Correlation of 3D Stereoscopic Images," *EURASIP Journal on Image and Video Processing* , December 2017, 2017 : 22, (SCI = 1.789)
- D. K. Kim, H. U. Jang, S. M. Mun, S. Choi, H. K. Lee, "Median Filtered Image Restoration and Anti-Forensics Using Adversarial Networks," *IEEE Signal Processing Letters*, Vol. 25, No. 2(February 2018), pp. 278-282. (SCI = 3.109)
- H. U. Jang, H. Y. Choi, J. Son, D. K. Kim, J. U. Hou, S. Choi, H. K. Lee, "Cropping-Resilient 3D Mesh Watermarking based on Consistent Segmentation and Mesh Steganalysis," *Multimedia Tools and Applications*, Vol. 77, No. 5(March 2018), pp 5685-5712. (SCI = 2.757)
- D. J. Jung, H. K. Lee, "Frame-rate Conversion Detection based on Periodicity of Motion Artifact," *Multimedia Tools and Applications*, Vol. 77, No. 5(March 2018), pp. 6095-6116. (SCI = 2.757)
- H. K. Kim, J. S. Park, D. K. Kim, H. K. Lee, "Two-stream Neural Networks to Detect Manipulation of JPEG Compressed Images," *Electronics Letters*, Vol. 54, No. 6(22 March 2018), pp. 354-355. (SCI = 1.314)

KaiCatch

가. 국제 저널 (계속)

- J. S. Park, H. K. Kim, D. K. Kim, I. J. Yu, H. K. Lee, "Paired Mini-batch Training: A New Deep Network Training for Image Forensics and Steganalysis," *Signal Processing - Image Communication*, Vol. 67, September 2018, pp. 132-139. (SCI = 3.256)
- D. K. Kim, J. U. Hou, H. K. Lee, "Learning Deep Features for Source Color Laser Printer Identification based on Cascaded Learning," *Neurocomputing*, Vol. 365, No. 1 (6 November 2019), pp. 219-228. (SCI = 5.719)
- W. Ahn, S. H. Nam, M. Son, H. K. Lee, "End-to-End Double JPEG Detection with a 3D Convolution Network in the DCT Domain," *Electronics Letters*, Vol. 56, No. 2, pp. 82-85. (SCI = 1.314)
- W. Ahn, H. U. Jang, S. H. Nam, I. J. Yu, H. K. Lee, "Local-Source Enhanced Residual Network for Steganalysis of Digital Images," *IEEE ACCESS*, Vol. 8, August 2020, pp. 127477-127490. (SCI = 3.367)
- I. J. Yu, W. Ahn, S. H. Nam, H. K. Lee, "BitMix : Data Augmentation for Image Steganalysis," *Electronics Letters*, Vol. 56, No. 24(26 Nov. 2020), pp. 1311-1314. (SCI = 1.314).
- I. J. Yu et. al., "Manipulation Classification for JPEG Images Using Multi-Domain Features," *IEEE ACCESS*, Vol. 8(Decem. 2020), pp. 210837-210857. (SCI = 3.367)
- D. H. Kim, et. al., "End-to-End Anti-Forensics Network of Single and Double JPEG Detection," *IEEE ACCESS*, Vol. 9(Jan. 2021), pp. 13390-13402. (SCI = 3.367)
- S. H. Nam., et. al., "Deep Convolutional Neural Network for Identifying Seam-Carving Forgery," *IEEE Trans. On Circuits and Systems for Video Technology*, Vol. 31, No. 8(August 2021), pp. 3308-3326. (SCI=4.686)

KaiCatch

가. 국제 저널(계속)

- W. G. Bae, et. al., “Dual-Path Convolutional Neural Network for Classifying Fine-Grained Manipulations in H.264 Videos,” *Multimedia Tools and Applications*, Vol. 80, pp. 30879-30906, 2021. (SCI = 2.757)
- M. J. Kwon, et. al., “Learning JPEG Compression Artifacts for Image Manipulation Detection and Localization,” Springer, *Int. J. of Computer Vision*, will appear in 2022. (SCI = 7.410)

나. 국제 학술 회의

- “Detecting Composite Image Manipulation based on Deep Neural Networks,” IEEE & EURASIP, 24th Int. Conf. on Systems, Signals and Image Processing(IWSSIP’2017), 22-24 May 2017, Poland.
- “Identifying Photorealistic Computer Graphics using Convolutional Neural Networks,” IEEE Int. Conf. on Image Processing(ICIP’2017), 17-20 September 2017, China, pp. 4093-4097.
- “Double JPEG Detection in Mixed JPEG Quality Factors using Deep Convolutional Neural Network,” 2018 European Conf. on Computer Vision(ECCV’2018), 8-14 September 2018, Germany, pp. 656-672.
- “Content-Aware Image Resizing Detection Using Deep Neural Network,” IEEE Int. Conf. on Image Processing(ICIP’2019), 22-26 September 2019, Taiwan, pp. 106-110.
- “Two-Stream Network for Detecting Double Compression of H.264 Videos,” IEEE Int. Conf. on Image Processing(ICIP’2019), 22-26 September 2019, Taiwan, pp. 111-115. 외 다수

회사 개요

- 일반사항
  - 회사명 : (주)디지탈이노텍
  - 법인등록번호: 160111-0099162
  - 사업자등록번호: 314-81-33864
  - 대표 이메일 : heunglee@kaist.ac.kr
  - 주 소 : 대전광역시 유성구 대학로 291 E3-1동 1428호실/
    대전광역시 유성구 문화원로 119, 7층 #714
  - 품목 : 딥페이크 및 사진 위변조 탐지 S/W 개발 및 서비스

- 연 혁
  - 2000. 5. ㈜디지탈이노텍 설립, KAIST 창업 승인 및 KAIST 전산학동 1440호 입주
  - 2000. 6. KAIMark 1.0 출시
  - 2000. 8. KAIST와 창업 지원 계약 체결 및 무상기술실시권 협약
  - 2001. 12. 정지영상, 동영상, 오디오 통합 솔루션 KAIMark 2.0 출시
  - 2002. 4. 소프트웨어 국가 시범사업에 참여
  - 2003. 문화부 국가 시범사업 참여
  - 2007. KAIMark 5.0 출시
  - 2010. KAIST와 영상 변조 및 변형탐지 기술 이전 계약 체결
  - 2010. 이미지, 비디오 저작권 진위판별 솔루션 출시
  - 2012. 웹기반 진위 판별 솔루션 출시 및 국가기관, 산업체 등에 다수 납품
  - 2015. 영상 변형 탐지 기술 개발(KAICATCH 1.0)
  - 2018. 동영상 포렌식 연구 수행
  - 2020. 11. 디지털 사진 위변조 탐지 소프트웨어 KAICATCH 2.0 출시
  - 2021. 3. 안드로이드 기반 카이캐치 앱 출시
  - 2022. KAICATCH 2.1 출시, 비디오 변형 탐지 포함