

디지털 콘텐츠(이미지)의 신뢰성 향상을 위한 안내서





디지털 콘텐츠(이미지)의 신뢰성 향상을 위한 안내서

최근 디지털 기술의 발전에 따라 연구 논문에서 그래프, 사진 등 디지털 이미지 콘텐츠의 위·변조 또는 복제 등 연구부정행위가 발생하여 해당 논문이 철회되는 사례가 증가하고 있어 이를 방지할 수 있는 대안 마련이 필요한 상황입니다.

이에 본 안내서에서는 디지털 이미지 콘텐츠의 위·변조 및 복제에 대한 사전 검증과 그 결과를 활용하여 논문의 동료심사 및 연구과제 평가에 활용할 수 있는 가이드라인을 제시하였습니다.

또한, 최근 이슈가 되고 있는 페이퍼밀/논문공장(Paper Mill)의 피해를 예방하고자 「페이퍼밀 식별 체크리스트」를 안내하오니, 각 학술단체 및 연구기관의 연구자께서는 본 안내서를 참고하시어 연구부정행위를 예방하고, 디지털 콘텐츠 신뢰성 확보에 적극 활용하시기 바랍니다.

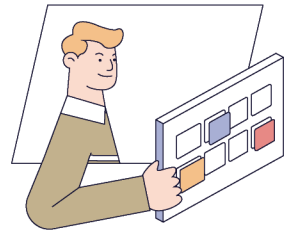
※ 동 안내서는 ‘연구 논문 디지털 이미지 콘텐츠의 신뢰성 검증 및 확보를 위한 가이드 개발(김수기, 연세대)’ 보고서에 근거하여 제시한 것으로, 한국연구재단의 공식적인 견해와는 다를 수 있습니다.

디지털 콘텐츠(이미지)의 신뢰성 향상을 위한 안내서

01

디지털 콘텐츠 유형 및 부정행위

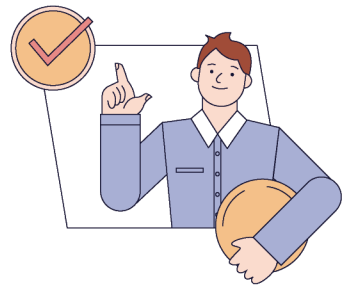
4p



02

디지털 콘텐츠 부정행위 탐지방법

10p

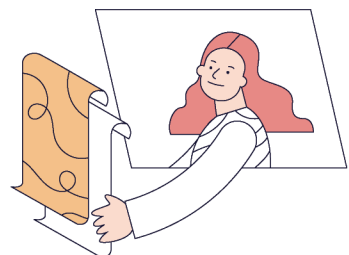


03

디지털 콘텐츠 부정행위 예방 체크리스트

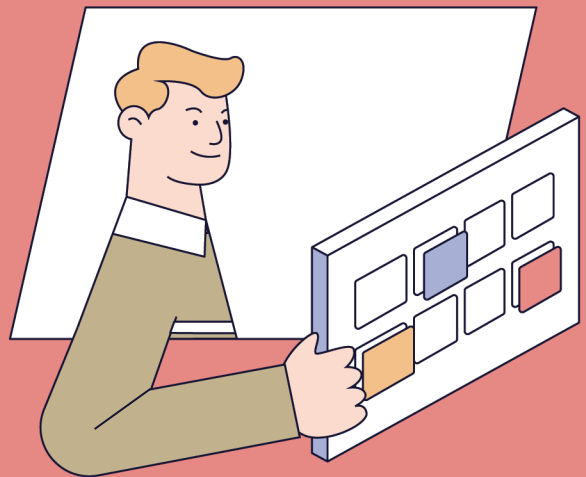
22p

- [1] 이미지 조작 식별 도구 및 판단 기준
- [2] 이미지 진실성 자가 체크리스트(연구자 용)
- [3] 페이퍼밀 식별 체크리스트(심사자 용)



01

디지털 콘텐츠 유형 및 부정행위



01

디지털 콘텐츠 유형 및 부정행위

1) 디지털 콘텐츠 유형

학술연구에서는 주로 사진과 그래프가 많으며, 최근에는 비디오 저널도 있어 디지털 콘텐츠 종류가 늘어날 것으로 전망된다 <표 1>.

<표 1> 이미지 유형의 구분

구분	내용
사진	가장 일반적인 이미지 유형이며, JPEG(.jpg), PNG(.png), TIFF(.tif), RAW(.raw, .cr2, .nef 등)와 같은 형식으로 저장됨
그래프	주로 학술 논문, 프레젠테이션, 비즈니스 보고서에서 사용되며, SVG(.svg)나 PDF(.pdf)와 같은 벡터 형식이나 JPEG(.jpg)나 PNG(.png) 같은 래스터 형식으로 저장됨
비디오	기술적으로 “이미지”는 아니지만, 비디오는 빠른 연속으로 재생되는 이미지(프레임)의 연속임(MP4(.mp4), AVI(.avi), MOV(.mov), MKV(.mkv) 등의 비디오 형식이 일반적임)
딥페이크	인공 신경망을 사용하여 기존 이미지나 비디오의 사람을 다른 사람의 모습으로 교체하는 합성 미디어임(딥페이크는 이미지와 비디오 모두의 형태가 가능하므로 JPEG(.jpg), PNG(.png), MP4(.mp4) 등의 형식으로 저장됨)
디지털 아트웍	디지털 도구를 사용하여 만든 이미지로, JPEG(.jpg)나 PNG(.png) 같은 비트맵 형식과 SVG(.svg)나 AI(.ai) 같은 벡터 형식으로 저장됨
인포그래픽	정보와 데이터의 시각적 표현임(그래프와 마찬가지로 SVG(.svg)나 PDF(.pdf)와 같은 벡터 형식이나 JPEG(.jpg)나 PNG(.png)와 같은 래스터 형식으로 저장됨)
3D 모델	전통적인 2D 이미지는 아니지만 그런 것으로 렌더링하여 볼 수 있음(3D 모델 파일은 보통 OBJ(.obj), STL(.stl), FBX(.fbx) 등의 형식으로 저장됨)

2) 디지털 콘텐츠와 관련한 부정행위

연구논문에서 디지털 이미지는 표, 그래프, 사진 등을 포함하며, 실험 과정에서 얻은 다양한 기법과 기기로 촬영하여, 최종적으로 소프트웨어로 다듬어져 디지털화된 문서를 의미한다.

통상 국내외 법규나 출판윤리 가이드에선 위조(Forgery or Fabrication), 변조(Tampering or Falsification), 표절(Plagiarism)로 구분하며, 최근 개정된 「연구윤리 확보를 위한 지침(시행 2023.7.17.)」에서는 이들을 다음과 같이 정의하고 있다.

- “**위조**”는 존재하지 않는 연구 원자료 또는 연구자료, 연구결과 등을 허위로 만들거나 기록 또는 보고하는 행위
- “**변조**”는 연구 재료·장비·과정 등을 인위적으로 조작하거나 연구 원자료 또는 연구자료를 임의로 변형·삭제함으로써 연구 내용 또는 결과를 왜곡하는 행위
- “**표절**”은 일반적 지식이 아닌 타인의 독창적인 아이디어 또는 창작물을 적절한 출처표시 없이 활용함으로써, 제3자에게 자신의 창작물인 것처럼 인식하게 하는 행위

한편, 학문 분야나 법적인 문서, 작품 등에서 부정행위를 가리키는 관련 용어로 각 다음과 같이 쓰이기도 한다.

- **위조(Forgery/Fabrication)** : 주로 문서나 서류를 불법으로 복제하거나 조작하여 진본으로 가장하는 행위로 보통 법적 문서나 서명, 통화, 예금 증서, 연구논문 등과 관련이 있다.
- **변조/조작(Tampering/Falsification)** : 이미 존재하는 문서나 정보를 조작하거나 수정하여 원래의 의도와는 다르게 나타나게 하는 행위로, 주로 문서, 기록, 전자 파일 등에 적용되며, 원본의 내용을 왜곡하거나 변경함으로써 정보를 속이는 것을 목적으로 한다. 또는 연구재료, 장비, 과정 등을 인위적으로 조작하거나 데이터를 임의로 변형 삭제하는 것을 말한다.
- **표절(Plagiarism)** : 다른 사람의 아이디어, 문장, 작품, 연구 결과 등을 부당하게 도용하여 자신의 것처럼 사용하는 행위로, 학문적으로는 다른 사람의 연구 논문, 책, 노트 등에서 아이디어를 가져와 자신의 연구나 글에서 출처를 명시하지 않고 사용하는 것을 말한다.
- **복제(Replication)** : 표절과 유사하나 표절은 주로 표적물의 글을 일부 또는 전체를 도용하나, 복제는 주로 연구논문에서 이미지를 도용하는 데 사용된다.

이런 용어들을 이미지 조작 등에 적용할 수 있으나 국내외로 학술연구에서 이미지에만 특화하여 정해진 규정이 별로 없어 논의가 필요하다.

요약하면, 학술연구에서 이미지 조작은 **논문 투고자의 이기적 욕심으로 최종 디지털화된 원본의 이미지를 다양한 방식으로 바꾸는 것을 말한다.** 국제적으로 흔히 쓰는 세부적인 용어는 다음과 같다 <표 2>.

<표 2> Retraction Watch database에서 분류하는 이미지 관련 문제점

구분	설명
이미지에 대한 우려/문제	이미지의 유효성에 대한 질문, 논란 또는 논쟁
이미지 복제	'자기 표절'이라고도 함. 원본 기사의 한 명 또는 모든 저자가 작성한 항목의 이미지가 적절한 인용 없이 원본 기사에서 반복되는 경우 사용
이미지 오류	이미지 준비 또는 인쇄 시 발생한 실수
이미지의 위조/변조	실제 데이터를 나타내지 않도록 이미지를 의도적으로 변경하는 행위
이미지 표절	원저작물의 저작자 1인 또는 모든 저작자가 작성하지 않은 항목의 이미지가 적절한 인용 없이 원저작물에서 반복되는 경우
신뢰할 수 없는 이미지	이미지의 정확성이나 유효성이 의심스러움

※ 출처 : <https://retractionwatch.com/retraction-watch-database-user-guide/retraction-watch-database-user-guide-appendix-b-reasons>
에서 이미지 문제로 논문이나 보고서가 철회되는 이유를 6개의 이미지 핵심어와 설명으로 나타냄

3) 학술연구에서 디지털 콘텐츠의 중요성 및 피해내용

학술연구에서 이미지는 연구 결과를 시각화하여 가독성과 이해력을 높이며, 이를 통해 저자의 주장을 강화하며, 독자와 심사자를 설득하는 데 강력한 수단이다. 하지만 기술 발전으로 이미지를 조작하고 수정하는 것이 쉬워졌지만, 가짜 이미지 조작과 표절도 횡행하게 되었다.

고의로 제작된 이미지 조작 사진 및 영상으로 국가 간, 국가 내, 조직과 개인에게 엄청난 피해를 줄 수 있는 대표적인 피해 내용은 다음과 같다.

이미지 조작으로 인한 대표적인 피해 내용

- 조작된 이미지를 통해 가짜 뉴스를 만들어 경제적 이득을 취하거나 군사적 목적으로 허위 선전을 할 수 있음
- 조작된 이미지를 통해 거짓 증거를 만들어 경찰 수사나 재판에 사용할 수 있음
- 조작된 이미지를 성 관련 범죄 등 악의적으로 사용해 모욕이나 명예 훼손을 할 수 있음
- 조작된 이미지를 사용하여 연구 결과를 조작하여 편향된 결론을 내거나, 기존 이미지를 복제 후 수정하여 사용하여 의미 없는 논문을 출판하는 등 학계에 피해를 줄 수 있음

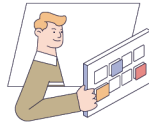
상기 예시의 마지막 항목은 최근 학술 분야에 심각한 문제를 일으키고 있으며, 학술논문 외에 연구보고서에 이미지 조작의 사례가 국내외 유명 학술지에 계속 보고되고 있다.

참고로 아래 웹사이트에서 다양한 분야에 이미지 조작 사례를 확인할 수 있다 <표 3>.

<표 3> 이미지 조작 확인 웹사이트

Institution	URL
Forensic Tools of The Office of Research Integrity	https://ori.hhs.gov/forensic-tools
Image Data Integrity, Inc.	http://www.imagedataintegrity.com
VerifEye	https://verifeye.online
Image Analysis Collaboratory at Harvard Medical School	https://iac.hms.harvard.edu
법영상분석연구소	https://www.imageid.co.kr/default

※ 출처 : 의학논문 출판윤리 가이드라인 제3판 (2019, 50쪽); https://www.kamje.or.kr/board/view?b_name=bo_publication&bo_id=13



**디지털 콘텐츠(이미지)의
신뢰성 향상을 위한 안내서**

02

디지털 콘텐츠 부정행위 탐지방법



02

디지털 콘텐츠 부정행위 탐지방법

1) 학술연구논문 관련 이미지 조작/변조에 대한 정보

일차적으로 RetractionWatch(<https://retractionwatch.com>)와 PubPeer(<https://pubpeer.com>) 사이트를 조회하면 이미지 조작 논문과 저자에 관한 상세한 정보를 얻을 수 있다.

특히 RetractionWatch의 하위 메뉴에서 여러 Database 중 이미지 관련 핵심어 (+Falsification/Fabrication of Image OR +Error in Image OR +Manipulation of Images OR +Plagiarism of Image OR +Unreliable Image OR +Duplication of Image)를 넣고, 각 분야 별 예로 인문 예술 사회학 분야 [(HUM) Arts - Biography OR (HUM) Arts - Film Studies OR (HUM) Arts - General OR (HUM) Arts - Literature/Poetry OR (SOC) Communications OR (SOC) Education OR (SOC) Ethics/Bioethics OR (SOC) Law/Legal Issues OR (SOC) Philosophy OR (SOC) Political Science OR (SOC) Sociology OR (SOC) Sports and Recreation OR (SOC) Psychology)] 관련 핵심어를 넣으면 게재 취소된 논문을 수집할 수 있다.

아래는 학술연구논문 관련 이미지 조작/변조의 탐지에 관련된 기관 (공립, 사립, 비영리 국제기구, 대학/연구소), 기술 및 플랫폼의 웹사이트 주소들이다.

[참고 1] 학술연구논문 관련 이미지 조작/변조의 탐지 관련 참고 사이트

사이트명	사이트 개요	주소
Retraction Watch	과학 논문의 철회와 관련 주제에 대해 보도하는 사이트	<ul style="list-style-type: none"> • https://retractionwatch.com/retraction-watch-database-user-guide/retraction-watch-database-user-guide-appendix-b-reasons
STM Association	신뢰할 수 있는 연구를 발전시키기 위해 회원들과 협력하는 학술 출판 산업 관련 사이트	<ul style="list-style-type: none"> • https://www.stm-assoc.org
COPE(Committee on Publication Ethics)	저널 출판의 윤리적 문제를 다루기 위한 모범 사례에 대한 조언과 지침을 제공	<ul style="list-style-type: none"> • https://publicationethics.org
GitHub	보안 소프트웨어를 구축, 확장 및 제공하기 위한 완전한 개발자 플랫폼	<ul style="list-style-type: none"> • https://github.com/shobhittya/ForensicNet
The Office of Research Integrity	미국 국립보건원(NIH)의 자금 지원을 받는 연구자들의 과학적 부정행위 등에 대한 조사를 담당하는 기관	<ul style="list-style-type: none"> • https://ori.hhs.gov • https://ori.hhs.gov/droplets
For Better Science	연구 윤리, 생명의학 윤리 및 학술 출판에 대한 사이트	<ul style="list-style-type: none"> • https://forbetterscience.com/?s=paper+mill
ProofigAI	과학 논문의 그림에서 윤리 문제를 감지하기 위한 AI 기반 소프트웨어 제공	<ul style="list-style-type: none"> • https://www.proofig.com
Imagetwin		<ul style="list-style-type: none"> • https://imagetwin.ai
ImaChek		<ul style="list-style-type: none"> • https://www.imachek.com
Forensically	디지털 이미지 포렌식을 위한 무료 사이트로, 클론 탐지, 오류 수준 분석, 메타 데이터 추출 등이 포함	<ul style="list-style-type: none"> • https://29a.ch/photo-forensics/#forensic-magnifier
FotoForensics	오류 수준 분석, 메타데이터 및 튜토리얼을 포함한 디지털 그림 분석을 위한 도구 및 교육을 제공	<ul style="list-style-type: none"> • https://fotoforensics.com
Image-integrity	과학 논문의 불법 이미지 조작 문제를 감지하고 이에 대한 설명을 제공	<ul style="list-style-type: none"> • https://image-integrity.com/about.html
Image Data Integrity	생명의공학 연구의 이미지 조작에 대한 컨설팅 서비스 제공	<ul style="list-style-type: none"> • https://www.imagedataintegrity.com
(주)디지털이노텍	딥페이크 및 사진 이미지 위변조 탐지 소프트웨어 제공 사이트	<ul style="list-style-type: none"> • https://kaicatch.com

2) 페이퍼밀의 식별 지표

이미지 조작은 대부분 페이퍼밀과 관련이 있으며, 페이퍼밀 논문 식별과 관련된 대표적 특징 20가지는 <표 4>와 같다. 이러한 특징은 PM 식별에 매우 중요하며, 실제 서로 다른 국제 학술지의 심사자나 편집위원들이 필드에서 경험하는 양상과 같거나 유사하다. 그러나 최근 이를 벗어나는 경우도 많아 세심한 주의가 요구되고 있다.

<표 4> 페이퍼밀 논문의 20가지 특징

구분	내용
상업용 이메일 주소	모든 논문에서 상업용 이메일 주소를 쓰고, 소속 기관 이메일 주소를 쓰지 않음(진짜 학술 기관은 학술용 이메일 주소를 제공. 편의를 위해 상업용 이메일 주소를 사용은 보안 공백으로 페이퍼 밀에 의해 약용된다.
순조로운 동료 심사 과정	투고 논문은 문제없이 iThenticate 소프트웨어를 사용한 표절 검사와 조작된 연구자료에 대한 의심 없이 1차 심사를 통과한다.
가짜 성과의 “과학적” 근거를 과대 홍보	투고 논문은 화학적으로 정의된 천연 화합물을 다룸. 대부분 식물에서 추출한 “매우 중요”, “매우 관련성” 또는 “의학적 서비스가 부족한” 질병과 관련이 있다. “아직 치료법 없음” 및 “높은 의학적 필요” 등 약리학 분야의 큰 성과와 기대를 과장한다.
서비스 실험실에 대한 보고 없음	몇몇 논문에는 소위 서비스 실험실의 데이터가 포함되어 있지만 논문은 재료 및 방법 섹션에서 이 사실을 언급하지 않는다.
불특정 “서비스 연구소/자료생산랩”	저자는 “문제 데이터”를 제공한 “서비스 연구소/자료생산랩”을 언급하고, 비난하지만 저널 측은 “문제 데이터”의 정확한 특성에 대한 정보를 얻을 수 없고, 원 자료 생산 연구원과 소통이 안된다.
데이터가 전혀 없음	모든 경우에 저자가 원본 데이터(원시 데이터)를 제공할 수 없다.
아름다운 이미지 잘라내기 및 붙여넣기	특히 유세포 분석 실험, 형광 세포 이미지, 웨스턴 블롯 및 조직학과 관련된 가짜 데이터를 논문에 사용한다.
이메일 서명 없음	어떠한 경우에도 교신저자가 이메일에 전문(기관) 이메일 서명을 사용하지 않는다.

구분	내용
이메일에 사용된 매우 열악한 영어	이메일에 사용된 영어는 매우 열악하고 놀라울만큼 논문에 사용된 영어보다도 훨씬 열악한 경향을 보인다.
철회에 대한 신속한 비공식 동의	놀랍게도, 철회 경고시 교신저자가 매우 신속하게(매우 몇 시간 내에 거의 동시에!) 실제로 페이퍼밀을 공개적으로 인정하지 않고 철회에 동의한다. 저자들은 문제가 무엇인지 구체적으로 밝히지 않고 "데이터에 문제가 있다"는 일반적인 표현을 사용한다.
다른 사람 탓하기	어떤 경우에는 교신저자가 "문제가 있는" 데이터를 생성한 대학원생 중 한 명(이를 밝히지 않음)을 비난했으며, 해당 학생은 이미 교육 기관에서 결과("처벌")에 직면했다고 한다.
고통스러운 공식 철회 과정	공식 철회 과정은 저자 측에서 매우 더디며, 전문적으로 처신하지 못한다.
내용 없는 커뮤니케이션	교신저자들과의 이메일 커뮤니케이션은 회피와 허위정보 수법이 만연하다. 편집장과의 소통이 거의 안되며, 저자들은 자신의 영어 실력이 "불행히도" 편집장이나 출판사의 질문을 이해하기에는 너무 부족함을 전달하고 싶은 태도를 보인다.
수수께끼 같은 이메일 클러스터	놀랍게도, 여러 페이퍼밀 의심 투고에서 편집장은 마치 동일한 이메일을 짧은 시간(분에서 몇 시간)에 분명히 다른 그룹의 저자로부터 이메일을 받았으며, 텍스트가 다소 비슷하다는 점도 발견된다. 이는 한 명의 페이퍼밀 직원이 여러 개의 가짜 투고 원고를 처리하고 있다고 추정할 수 있다.
느린 이메일 통신(없는 경우도 있음)	이메일 주소(특히 공동 저자)가 작동하지 않거나 저자가 응답하지 않아 통신이 전혀 되지 않는 경우가 있다.
학술 기관 없음	논문을 작성한 연구자가 소속된 기관과의 연락이 불가능하다.
ORCID ID 없음	ORCID ID를 사용한 저자는 거의 없다.
가짜 전문심사자	저자가 제안한 전문심사자를 사용할 때, 경우에 따라 리뷰가 비정상적으로 짧은 시간에 전달된다.
같은 논문, 다른 저자	같은 논문인데 저자가 다른 경우도 있다.
Paper Mill의 지리적 기원	주로 투고처는 중국의 대형 병원의 저자들이다.

※ 출처 : Seifert, R. How Naunyn-Schmiedeberg's Archives of Pharmacology deals with fraudulent papers from paper mills. Naunyn-Schmiedeberg's Arch Pharmacol 394, 431-436 (2021). (<https://doi.org/10.1007/s00210-021-02056-8>)

한편 COPE & STM (The International Association of Scientific, Technical and Medical Publishers: 국제 과학, 기술 및 의학 출판사 협회) 은 최근 발간한 보고서를 통해 페이퍼밀을 탐색하는 주요 지표를 다음과 같이 제시하고 있다.

COPE가 제시한 페이퍼밀 탐색 지표

- **과학적 주제**
논문은 주로 세포 및 분자 생물학 분야인 경우가 많지만, 다른 분야도 있다.
- **실험**
일반적으로 웨스턴 블롯 실험, 세포 분석, 조직학/세포 염색 사진을 사용한다.
- **실험 데이터**
웨스턴 블롯은 특히 배경이 너무 깨끗한 경우가 많으며, 세포 분석도 배경이 너무 깨끗하다; 분자량 마커는 일반적으로 웨스턴 블롯 실험에 표시되지 않는다.
- **레이아웃**
논문의 레이아웃이 매우 유사(그래프, 통계적 오차 막대, 그림의 글꼴 등)
- **소속**
저자의 소속이 특정 대학을 표시하지 않는다. 때때로 언급된 학과가 논문 주제와 일치하지 않는다.
- **저자**
특정 저널에 출판한 기록이 없는 저자가 제출 또는 다른 곳에서 제출한다. 기관이 아닌 이메일 주소가 많이 사용되며, 매번 투고 시 새로운 ORCID가 생성된다.
- **실험 설계**
면밀한 평가 결과, 실험 설계에 결함이 있다. 예를 들어 실험 데이터가 실험 방법의 설명과 일치하지 않거나 시약이 설명된 목적과 다르다.
- **윤리적 승인**
동물 실험에 대한 윤리적 승인이 없다.
- **저자 변경**
수정 또는 교정 과정에서 저자 목록이 크게 변경된다.
- **복제 이미지 발견**
동일 이미지가 다른 논문에서 발견된다. 예로 동일하거나 약간 조작된 이미지, 웨스턴 블롯 이미지, 세포분석법 분석 또는 현미경 세포 이미지 등이 있다.

※ 출처 : <https://publicationethics.org/sites/default/files/paper-mills-cope-stm-research-report.pdf>

3) 이미지 조작의 유형

학술연구에서 이미지 조작은 두 가지로 나눌 수 있다.

Papermill-related image forgery(페이퍼밀 관련 이미지 조작)와 Non-papermill-related image forgery(페이퍼밀 무관한 이미지 조작)로 나눌 수 있고 이 차이는 다음과 같다.

<표 5> 연구논문에서 이미지 조작의 유형별 차이

항목 \ 유형	Papermill 관련 이미지 조작	Non-papermill 관련 이미지 조작
주체 (조작)	논문을 의뢰한 PI(연구책임자) (Non-lab base)	PI(연구책임자), 주로 연구원이 관여 (Lab base)
동기	이기적 의도 (사업 이익, 승진, 임용, 인센티브)	이기적 의도 또는 실수(비의도적)
건수/상습성	대규모/상습성, 고의성	소규모/산발적, 우발적, 상습성, 고의성은 적음
연구의 창의성, 독창성	거의 없음	Papermill에 비하여 창의성과 독창성이 현저함
피해 규모/ 파급 정도	대규모/파급력	소규모/파급력
제어 주체	국가, 출판사, 학술지, 학회	출판사, 학술지, 학회, 저자가 속한 Lab(연구실)
관련 국가	중국, 러시아, 중동, 유럽	특정 국가 없음

4) 이미지 조작 탐지를 위한 가이드

<표 2>를 통해 Retraction Watch database에서 분류하는 이미지 관련 문제점과 논문 철회 이유를 설명하였고, 추가로 STM과 COPE에서 제시한 이미지 조작의 수준은 아래와 같다 <표 6>.

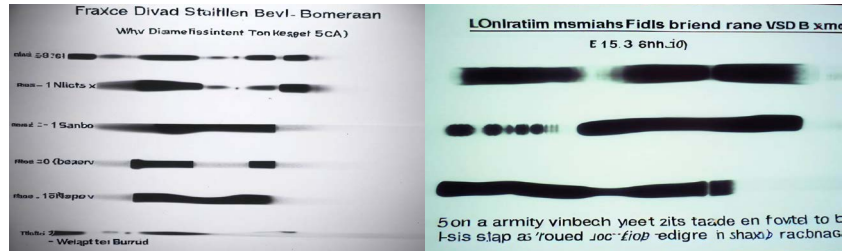
<표 6> (STM이 제시한) 이미지 조작의 수준 가이드

구분	설명	예시
레벨 1	이미지의 이상은 실제 또는 이미지 패널, 원본 데이터의 일부에만 한정되며, 이는 데이터 처리 오류나 확장 이미지 처리로 인해 발생하고, 이해에 영향을 줄 수 있다. 원본 데이터는 쉽게 접근할 수 있고, 이미지 이상을 설명할 수 있으며, 이러한 이상이 결과에 큰 영향을 미치지 않을 경우, 적절히 수정된 버전이나 대체 데이터를 제공 시 수정이 가능하다.	<ul style="list-style-type: none"> - 이미지의 전체 또는 일부의 중복(동일한 내부, 다른 그림, 다른 논문에서) - 겹/블롯 레인의 수직적인 분할 (동일한 블롯/겹에서 비교 가능한 데이터를 합치거나, 다른 블롯/겹에서 비교 불가능한 데이터를 결합) - 같은 블롯/겹에서 실행되지 않은 대조군 겹/블롯의 로딩
레벨 2	중요한 이미지를 '미화' 또는 실제 실험이 안된 이미지/데이터 조작은 객관적인 데이터 제시에 위반되며, 학문적인 관행과도 상충되며, 연구 논문의 핵심 데이터에 대한 과학적 결론을 바꾼다. 심사자의 오해를 살 의도성이 있으며, 공식적인 추가 조사를 배제할 수 없다.	<ul style="list-style-type: none"> - 중요한 특징의 양적 정보에 영향을 미치거나 이미지를 불분명하게 극단적으로 대비되게 조정한다. 이 양적 정보가 논문의 핵심 결론에 결정적이라면 이는 '위조'라 할 수 있다. - 대조군 레인이나 대조군 패널의 재사용(잠재적으로 '위조/조작')
레벨 3	심각한 이미지 조작, 일반적으로 한 개 이상의 이미지 패널에서 오해를 낳는 조작하는 의도에 대한 실질적인 증거, 그리고 강력하고 인증된 원본 데이터가 결여되어 있다. 3단계 논문은 일반적으로 2단계 또는 3단계 이미지에 이상 있는 여러 개의 개별 그림 패널을 보인다.	<ul style="list-style-type: none"> - 별도의 실험, 클로닝, 이미지 부분 (신호 또는 배경) 또는 그래프의 삽입, 삭제, 재배치, 반전, 패널 내 요소 또는 그래프 뒤집기; 이미지 합성이 있다. - 잘못된 라벨링 및 이미지의 다른 부분의 중복이 있다. - 이미지 또는 이미지 세트 내에서 이미지 처리가 불균등하여 데이터가 잘못 해석될 수 있다. - 데이터를 대표하지 않는 이미지의 선택적 보고 또는 자르기가 있다.

※ 출처 : STM Recommendations for handling image integrity issues, <https://doi.org/10.17605/OSF.IO/XP58V>; <https://osf.io/kgyc6>

이 표를 요약하면 (레벨) 1, 2는 의도적이 아닐 수 있고, 논문 게재취소의 사유로는 미흡하나, (레벨) 3인 경우 페이퍼밀 관련 또는 무관련 이미지 조작에서 현저하며, 의도적이며 때로는 상습적이며 논문 거절 및 이미 출판된 논문이라도 게재취소의 사유로 충분하다.

그러나 현재 COPE에서 제시된 레벨 3까지 이미지 조작으로만 심사자가 기존의 기준으로 잡아 낼 수 없다. 즉 페이퍼밀에서 기존의 이미지를 조작하거나 복제하는 수준을 넘어서 가짜 데이터를 생성해서 만든 이미지나 시를 이용하여 만든 이미지등이 그 예다 <그림 1>.



<그림 1> 인공지능으로 생성한 웨스턴 블롯 밴드*

* 프롬프트 : "Draw only western blot band results taken in camera! draw only western blot band results taken in camera! Band should be five different sample with discriminative density with two fold decrease such as very faint, faint, intermediate, and, more intermediate, and strong within same horizontal gel. Band shape such as short tadpole tail or rectangular should be flexible, and short diagonal in black with white background. Band size should be similar."

이러한 점을 고려하여 본 가이드에서는 이미지 조작의 수준을 탐지하기 위한 추가적인 가이드(레벨 4, 5)를 아래와 같이 제시한다 <표 7>.

<표 7> 이미지 조작을 탐지하는 추가 가이드

구분	설명	예시
레벨 4	의도적, 상습적, 고의적이다 (저자의 한 논문에서만 국한), 가짜 데이터를 생성해서 만든 이미지나 시를 이용하여 만든 이미지로 구성된다.	조작된 이미지 자료의 숫자가 종류와 무관하게 적어도 2개 이내다. 결론에 기여하는 이미지 결과의 정도가 상당하다.
레벨 5	매우 고의적이고, 실수의 가능성 전혀 없다, 저자의 적어도 한 논문 이상에서 발견될 수 있다. 가짜 데이터를 생성해서 만든 이미지나 시를 이용하여 만든 이미지로 구성된다.	조작된 이미지 자료의 숫자가 종류와 무관하게 적어도 3개 이상이다. 결론에 기여하는 이미지 결과의 정도가 결정적이다.

5) 이미지 조작의 참지용어

이미지 조작과 관련하여, 여러 문헌에서 발췌해 정리한 것으로 종류별 용어 및 설명은 다음과 같다 <표 8>.

< 표 8 > 이미지 조작의 세부 전문용어 예시 및 설명

용어	영문	설명
베끼기/복제 (부분, 전체/배경, 표적 마커)	Cloning/Replication/ Duplication (C/R1/D)	표적 그림인 경우 부분 또는 전체가 가능하며, 배경이나 표적 마커 등을 베낌
꾸미기/조정/미화 (크기/양적, 질적 음영)	Adjustment/Contrast /Beutification (A/C1/B)	주로 표적 마커를 늘리거나 줄이고, 음영을 조절하여 대조군에 비하여 의미 있는 자료 생산
돌리기/회전	Rotation (R)	주로 표적 마커를 여러 각도로 돌려 실험 없이 부족한 자료를 보충
뒤집기/반전	Flip/Inversion (F/I)	주로 표적 마커를 뒤집어 다르게 보여 실험 없이 부족한 자료를 보충
지우기/제거	Erasion/Deletion (E/D1)	주로 원하지 않는 표적 마커를 없앴
쪼개기/분열	Splicing/Splitting (S/S1)	주로 원하지 않는 표적 마커를 쪼개서 없애거나 다른 마커로 대체하기 위한 과정
더하기/삽입	Insertion/Addition (I/A1)	주로 원하는 표적 마커를 더함
기타 (섞기/혼합)	Mixing/Hybrid/ Distortion (M/H/D2) C+A or R/F/E/S/I R+C or A/F/E/S/I F+C or A/R/E/S/I E+C or A/R/F/S/I S+C or A/R/F/E/I I+C or A/R/F/E/S/I	위의 여러 세부 조작이 한 개 이상 섞인 경우

※ 출처 : <https://image-integrity.com/Examples.html>; <https://www.proofing.com/>

6) 이미지 조작의 대처: 가이드 사례

이미지 조작에 대하여 「이공계 연구윤리 및 출판윤리 매뉴얼(2014)」에 따르면 “디지털 이미지 파일, 그림, 표 등에 대한 처리 및 심사 기준을 두고 규정에 명확하게 제시해야 한다.”고 되어 있으며, 「의학논문 출판윤리 가이드라인」 제3판에 나온 이미지 조작 관련 체크리스트는 다음과 같다.

의학논문 출판윤리 가이드라인(제3판) (2019)

- 학술지 투고 규정에 이미지 조작과 관련한 설명이 포함되어 있는가?
- 투고한 그림이나 사진에 어떤 형태로든 수정을 한 경우 이를 별도로 기술하였는가?
- 이미지에 문제가 제기되었을 경우 수정하지 않은 원본 이미지를 제출할 수 있는가?
- 이미지 조작을 밝히지 않으면 연구윤리 위반으로 제소될 수 있음을 알고 있는가?

※ 출처 : 의학논문 출판윤리 가이드라인 제3판 (2019, 50쪽)
 발행처 : 대한의학학술지편집인협회(KAMJE), 발행일: 2019년 3월 28일, ISBN : 978-89-93453-53-9 93510
https://www.kamje.or.kr/board/view?b_name=bo_publication&bo_id=13

국의 JCB (Journal of Cell Biology) 한 저널에서는 “디지털 이미지가 중요한 역할을 하는 경우가 많은 학술지에서는 투고된 논문의 디지털 이미지가 적절하게 처리되었는지를 확인하는 절차와 인력을 마련해야 한다.”고 되어 있다.

- 디지털 이미지 자료의 중요성을 인지하고, 전문 편집자를 통해 투고된 논문의 이미지 자료를 검토한다.
- 특히, 흑백 사진에서 밝기와 대비의 조정만으로도 이미지 조작을 확인할 수 있다.
- JCB는 저자로부터 디지털 이미지의 원본 파일 제출을 요구하며, 모두 검토한다.
- 이와 같은 절차는 이미지 자료의 적절한 처리를 보장하기 위함이다.
- JCB는 이미지 처리 원칙에 따라, 동일한 이미지 내 일부의 조작은 허용되지 않으며, 여러 이미지의 결합은 명확하게 표시되어야 한다.
- 또한, 밝기, 대비, 색채균형 등의 조정은 전체 이미지에 적용되며, 원래의 정보를 손상시키지 않는 범위에서만 허용된다.

※ 출처 : https://www.owlnet.rice.edu/~bios311/bios311/jcb_imagemanip.pdf
<https://rupress.org/jcb/article/166/1/11/34064/What-s-in-a-picture-The-temptation-of-image>
<https://rupress.org/jcb/pages/fig-vid-guidelines>

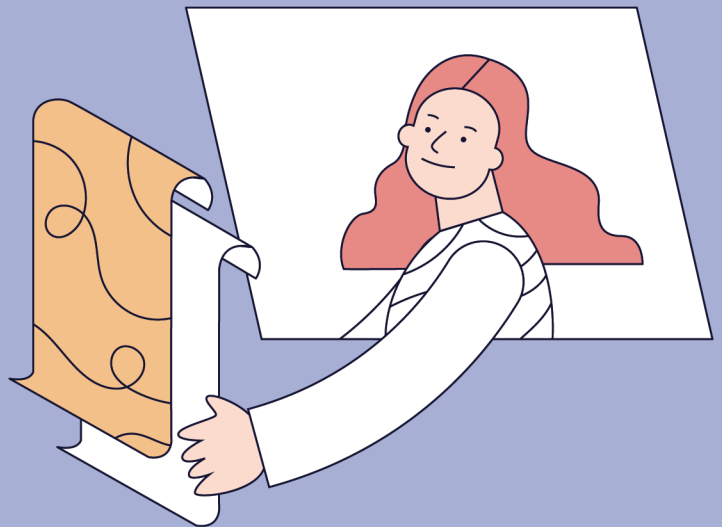
또한, 세계적인 과학저널인 사이언스가 올해부터 인공지능(AI)을 이용한 이미지 분석 도구인 ‘프루픽’(Proofig)을 활용하여 논문의 이미지 조작을 걸러내기 시작했다. 프루픽은 이미지 복제, 회전, 접합, 배율 왜곡 등 비정상적인 점을 탐지하고, 문제가 있는 이미지를 발견하면 논문 저자에게 해명을 요구하거나 논문 게재를 거부하거나 철회하는 등의 조치를 취한다.

그리고 미국과학편집인협회(Council of Science Editors, CSE)는 이미지 조작에 대한 지침을 제시하고, 연구자들에게 이미지의 특정 기능을 향상시키거나, 가리거나, 이동하거나, 제거하거나 도입하지 않도록 권고하고 있다. 또한, 이미지를 수정한 경우에는 그림 설명에 명시적으로 언급하도록 권장하고 있다.

03

디지털 콘텐츠 부정행위 예방 체크리스트

- [1] 이미지 조작 식별 도구 및 판단 기준
- [2] 이미지 진실성 자가 체크리스트(연구자 용)
- [3] 페이퍼밀 식별 체크리스트(심사자 용)



03

디지털 콘텐츠 부정행위 예방 체크리스트

이미지 조작을 예방하기 위해서는 연구자들의 윤리적 자각과 책임감이 가장 중요하고, 연구 기관, 출판사, 편집자, 심사자, 독자 등 모든 이해 당사자들의 협력과 감시가 필요하다.

이미지 조작은 과학의 발전을 저해하고, 공동체의 신뢰를 손상시키는 행위다. 연구자를 포함하여 이해 당사자들은 이미지 조작을 예방하고 규제하기 위해 서로 협력해야 한다.

연구 논문의 이미지 조작 문제에 대응하기 위해서는

저널의 지침과 요구사항을 숙지하고,
이미지 조작의 유형을 이해하며,
이에 대응하는 도구를 익히고,
전문가들의 지식을 공유하고,
연구 윤리에 관한 최신 정보를 얻고,
부적절한 연구를 발견하고 토론하며,
연구 맥락과 데이터 사용 방법에 대해 숙지하고,
이미지를 주의 깊게 살피며,
이미지 검증 도구를 사용하고,
발견된 문제를 저널 편집자에게 보고하는 것이 필요하며,
이를 지속하는 것이 효과적임을 인지하는 것이다.

아무쪼록 디지털 콘텐츠 부정사용 방지를 위하여 아래 세 가지 체크리스트를 제공하오니 적극 활용하시기 바란다.

[1]
이미지 조작
식별 도구 및
판정 기준

이용 방법

항목 순으로 점검을 권하며, AI (유/무료) 사용시 페이퍼밀, 비페이퍼밀 상관없이 이미지 복제나 조작을 식별 시 사용 가능하며, 탐지의 확률을 높이기 위해 다른 AI 플랫폼과 페이퍼밀 식별 알고리즘 혼용을 권장함.

항목	검사명/분석명	사용 도구	판정 기준	해당 유/무
1	이미지 확대 검사	(AI 무사용)	PDF 파일을 그림 부분을 120% 이상 확대 시 심한 Blurring (흐릿함) 발생 시 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
2	Error Level Analysis (ELA; 에러 수준 분석)	[무료 AI 플랫폼] https://29a.ch/photo-forensics/#forensic-magnifier	ELA 분석 시 조작 의심되는 비교질성과 객체가 주변과 다른 진한 음영이 발견 시 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
3	Clone, detection, Noise analysis, Level sweep analysis, Luminance gradient, Principal component analysis 등 여러 필터를 사용	[무료 AI 플랫폼] https://29a.ch/photo-forensics/#forensic-magnifier	5가지 필터 중 1개라도 가이드라인 해당 예시에 합치되면 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
4	이미지 복제 및 조작 탐지	[유료 AI 플랫폼] Imagetwin (https://imagetwin.ai ; https://app.imagetwin.ai)	가이드라인 해당 예시에 합치되면 “예” 서로 다른 저널의 이미지 복제도 탐지하고, 같은 원고의 같은 쪽, 다른 쪽의 그림 복제를 탐지하고, 가장 속도가 빨라 시를 이용 우선 선별 시 추천, 단 위음성도 있어 다른 AI 플랫폼과 PM 식별 알고리즘 혼용을 권장	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
5		[유료 AI 플랫폼] ImaCheck (https://www.imachek.com ; (https://alpha.imachek.com))	가이드라인 해당 예시에 합치되면 “예” 같은 원고의 같은 쪽, 다른 쪽의 그림 복제를 탐지하고, 가장 정확도가 높아 시를 이용 이미지 조작 확진 시 추천, 단 속도가 느려 필요시 다른 AI 플랫폼과 PM 식별 알고리즘 혼용을 권장	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
6		[유료 AI 플랫폼] Proofig (https://www.proofig.com ; https://app.proofig.com)	가이드라인 해당 예시에 합치되면 “예” 수동모드로 이미지를 선택하여 그림 복제 및 조작을 탐지하여 정확도가 높으나 작동법이 어렵고, 선별 시간이 오래 걸림. 특히 웨스턴 밴드 분석 시 유용함. 필요시 다른 AI 플랫폼과 PM 식별 알고리즘 혼용을 권장	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
판정	무료 AI 플랫폼에서 1개는 이미지 조작 확률/관련성을 중, 2개 이상은 상, 유료 AI 플랫폼에서 1개 이상은 상으로 판정			

[2]
 이미지진실성
 자가체크리스트
 (연구자용)

항목	내용	기입 요령	해당 유/무
1	이미지를 만든 저자를 사인과 함께 모두 적으셨습니까?	각 그림 순서대로 저자 기입, 대필 금지, 필요시 별지 사용	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
2	이미지의 Raw data (원본 자료)가 있습니까?	해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
3	2와 관련 자료 요청 시 복사본이나 증빙을 연구노트와 같이 보내줄 수 있습니까?		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
4	이미지의 원본 자료는 실제 실험한 자료입니까?		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
5	이미지의 해상도는 원고와 PDF로 만들 시 해상도가 적정합니까?		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
6	이미지 제작 시 사용한 기기와 소프트웨어가 있습니까?		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
7	그림의 의미를 강조하기 위하여 이미지 조작/변조가 없음을 맹세합니까?		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
8	그림의 자료 중 AI를 이용 생성한 것은 없습니까?		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
9	무료나 유료 AI 플랫폼을 써서 이미지 조작이 없는 증빙서를 요청 시 보내주실 수 있습니까? ※ 요청이 없더라도 AI(유/무료) 플랫폼을 사용하여 이미지 복제나 조작이 없는 증빙서 제출		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
10	교신저자가 논문의 그림의 조작이 없음을 확인하고 보증합니까? 교신저자 서명과 사인과 함께 적으시오		“예”, 대필 금지, 서명과 사인을 기입. 교신저자가 다수 시 별지 사용
판정	10개 항목 모두 “예”로 나와야 적합함, 기입 내용도 모두 완수해야 함.		

[3]
 페이퍼밀
 식별체크리스트
 (심사자용)

항목	구분	판정 기준	해당 유/무
1	이메일 주소	기관 이메일을 사용하지 않으면, “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
2	ORCID (Open Researcher and Contributor Identifier) 등록 여부와 등록 시점	ORCID 등록이 없거나, 투고전 등록 시 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
3	윤리적 선언 증빙의 부재나 조작	IRB 또는 IACUC 등 증빙이 없거나, 요청 시 조작이 의심 시 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
4	패턴화된 제목	가이드라인 해당 예시에 합치되면 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
5	신규성이나 독창성/가치성 (Originality/ novelty)의 부재		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
6	선행 출판 기록의 부재	유사한 연구의 저자의 선행 출판 기록이 없을 시 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
7	텍스트 표절 및 데이터 구조의 동일/유사 패턴	원 보고서나 기 페이퍼밀의 해당 예시에 합치되면 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
8	완벽한 통계		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
9	데이터 배열 패턴		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
10	이미지 조작/조작	원 보고서나 기 페이퍼밀의 해당 예시 (AI 사용 없이)에 합치되면 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
11	영문 교정 증명서	원 보고서나 기 페이퍼밀의 해당 예시에 합치되면 해당 유/무에 “예”로 표시	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
12	작성자 서명		<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
판정	“예”가 8개 이상 해당되면 페이퍼밀 확률이 고(높고), 5개 이상 8개 이하이면 (중), 5개 이하이면 (저)로 평가		

디지털 콘텐츠(이미지)의 신뢰성 향상을 위한 안내서

| 저자

김수기 연세대학교 교수

| 감수

윤철희 서울대학교 교수

한동수 한양대학교 교수

김옥주 서울대학교 교수

| 기획

한국연구재단 윤리정책팀

| 발행/인쇄일자

2024년 8월 27일

| 디자인/인쇄

디자인심원 042) 486-5777

| 발행 및 문의처

한국연구재단 윤리정책팀 042) 869-6648



이 안내서는 교육부 연구윤리활동지원사업 지원을 통해 도출된 결과물로, 저작권은 한국연구재단에 있습니다. 한국연구재단의 허락 없이 영리 목적으로 판매하는 등의 행위는 저작권법 등 관련 법에 따라 처벌될 수 있습니다.

※ 자세한 사항은 한국연구재단 연구윤리정보포털(<https://cre.nrf.re.kr/>)*에서 다운로드 가능합니다.

* 학술이카이브- 최종보고서(연구 논문 디지털 이미지 콘텐츠의 신뢰성 검증 및 확보를 위한 가이드 개발(김수기, 연세대), 2024)

디지털 콘텐츠(이미지)의
신뢰성 향상을 위한 안내서