
Essays

Misconduct of images: guidance for biomedical authors and editors

Qing Ye

Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology), Hangzhou, China; jzus_yq@zju.edu.cn; ORCID 0000-0002-6681-4731

Hanfeng Lin

Journal of Zhejiang University-SCIENCE B (Biomedicine & Biotechnology), Hangzhou, China; jzus_lhf@zju.edu.cn; ORCID 0000-0002-7437-1037

DOI:10.20316/ESE.2019.45.19006

Abstract

Misconduct in terms of manipulation of images has become an increasingly serious issue for the scientific community, especially in biomedicine. Such misconduct takes many forms, the major categories being falsified or fabricated images, manipulated images, and plagiarized images. Different tools and techniques are briefly described to help authors and editors in detecting such misconduct, and guidance is offered on appropriate use of images under different situations. More specifically, Crossref Similarity Check, Motuin, and Droplets are proposed as the tools of choice for detecting similarity between images and their possible manipulation.

Keywords: image manipulation, image falsification, image plagiarism, detecting scientific misconduct

Introduction

Hundreds of research papers are being retracted each year now owing to misconduct related to the images used in these papers. In 2016, 375 papers were retracted on these grounds, accounting for 22.9% of the total of 1635 papers retracted that year (this figure was obtained by searching the Retraction Watch Database, <http://retractiondatabase.org>). In a global survey of images from 20,621 papers published from 1995 to 2014, 782 (3.8%) of the papers were found to have problematic figures, at least half of them showing some signs of manipulation.¹ Although duplicated text was considered a minor problem in a global survey of editors,² duplicated images are generally considered a fraudulent practice, except in cases of honest errors. Whether intentional or not, defective images may lead to wrong interpretations and, in turn, to waste of time and money. Perhaps even worse, incorrect results could lead to failure of clinical trials and a negative impact on the health of patients.^{3,4} Although plagiarized text in research papers has been discussed at length and many tools have been developed to detect similarity of text,⁵ detecting similarity of images continues to be tricky. The present paper is an attempt to highlight potential misconduct with respect to images and recommends some good practices related to appropriate use of images.

Falsified or fabricated images

Definition

Falsified or fabricated images are those obtained by falsifying experimental data or fabricating research results. In most cases, such images are graphs, charts, or similar images based on mathematical or statistical calculations.

Detection

Image forgery is a form of forged data, which is difficult for editors to detect. In most cases, the forgery is spotted by reviewers, members of editorial boards, or readers who have the necessary expertise and experience. To confirm such suspected forgeries, the most effective way is to ask the author to provide the original experimental data on which the images are based: reviewers, editors, or other experts can then check the data and the images by following the flowcharts and guidelines published by COPE (the Committee on Publication Ethics).⁶

More and more academic journals now encourage authors to provide source data or minimally processed data underlying the images at the submission stage. In 2017, the International Committee of Medical Journal Editors (ICMJE) published an editorial⁷ simultaneously in *Annals of Internal Medicine*, *BMJ (British Medical Journal)*, *Bulletin of the World Health Organization*, etc, titled "Data sharing statements for clinical trials", which stated that "As of 1 July 2018 manuscripts submitted to ICMJE journals that report the results of clinical trials must contain a data sharing statement." Because the traditional submission system does not have enough space to store the source data, many public repositories for storing and sharing data are being developed. Such depositories include GenBank (www.ncbi.nlm.nih.gov/genbank), BIGD (<http://bigd.big.ac.cn>), ArrayExpress (<https://www.ebi.ac.uk/arrayexpress>), and FAIRsharing (<https://fairsharing.org>). However, data accessibility, security, reliability, and traceability are the four most important considerations. All the operations or processes should be recorded, including storage, deletion, and modification of data and all with a time stamp. If necessary, any data element should be open to public scrutiny for checking its scientific integrity; this requirement for transparency is an effective means of preventing any misconduct related to images, meanwhile providing the author with an opportunity to prove his innocence.

Manipulated images

Definition

Image manipulation refers to excessive or improper modification of actual images that destroys the integrity and authenticity of source data. In most cases, such manipulated images consist of photographs of samples such as histological slides, Western blots, and DNA gels and photographs of patients showing disease symptoms, abnormalities, etc.

More specifically, the manipulation involves (1) partially over-adjusting for colour, brightness, saturation, and contrast; (2) selective cropping, erasing, copying, rotating, scaling, or patching; and (3) improperly using image-processing software to modify the colour and contrast of the overall image, thereby misrepresenting the expressed information or suppressing some key information.⁸

Detection

With the development of computer technology, especially improvements in Adobe Photoshop® (PS), many modifications are hard to detect by the naked eye. Fortunately, tools are available to check for image modification. Such tools include Adobe Photoshop® Droplets (advised by the Office of Research Integrity, USA; <https://ori.hhs.gov/droplets>), Motuin (iPlagiarism, Beijing, China; <http://motuin.org>), and ImageJ (<http://imagej.nih.gov/ij>). However, to use these tools, some academic background and some knowledge of Photoshop is required. Some journals (for example, EMBO Journal, EMBO reports, EMBO Molecular Medicine, and Molecular Systems Biology) employ 'data detectives' to catch any manipulation of images or other related defects in images.⁹ These software programmes are described later in this article.

Plagiarized images

Definition

Image plagiarism is copying of an image or part of it and using it without any reference to its source.¹⁰

Typically, such plagiarism involves copying the target image or its part and making minor changes to its appearance such that its substance (idea) or the main message is unchanged: for example, redrawing a flowchart without any attribution of its source is image plagiarism.

Detection

Although plagiarism is easy to prove because it consists of outright copying, it requires a strong professional background to find the original source from which the image in question was copied. Most of the time, detection is based on communication from the plagiarized author (the true source of the image) or from a reviewer. Sometimes, Crossref Similarity Check (<https://www.crossref.org/services/similarity-check>) can be useful if the title, legend, or caption of the plagiarized image is identical to that of the original image. Motuin facilitates the comparison of two similar images; however, such comparison requires access to both

the images in question; in other words, the software package cannot retrieve potentially similar images from the internet or other databases. Although Google Images was launched in July 2001 and can search for similar images on the internet, most images used in research papers are too complex and detailed for a Google Image search to be of any practical use. Secondly, such potentially similar images are more likely to be behind a paywall or their sources accessible only with a subscription. Therefore, a different system is urgently needed to conduct similarity checks on images^{11,12} and will require a large repository of published images as well as a more efficient algorithm for comparing similar images.¹³

Guidance on appropriate use of images

Authors' obligations

Authors are responsible for preserving the minimally processed experimental data and images (source data) and making them available for evaluation when called for to settle a dispute or to prove a claim. If such source data are unavailable, the editorial, production, or publishing process will be stalled until the matter is resolved.¹⁴ If the image or images in question lead to suspicions of misconduct, the authors will also be requested to assist in the investigation and provide a reasonable explanation.⁶

General rules for image processing

Although such changes as cropping, resizing, rotating, and moderate adjustments of the overall resolution, colour, contrast, brightness, and saturation are permitted, these changes should not misrepresent the actual experimental findings and need to be made judiciously. For example, over-adjusted contrast that causes some details to either disappear or become unduly prominent should be avoided.⁸

Specifications for different types of images

Different types of images have different requirements and different community standards. Some journals, such as Nature, The New England Journal of Medicine, and EMBO Journal, offer detailed specifications for different types of images.^{15,16} The three most common image types in biology and medicine are those related to electrophoretic gels and blots and microscopy and those of clinical interest showing a patient (Table 1).

Table 1. Requirements for images in scientific journals

Electrophoretic gels and blots ¹⁵
<ul style="list-style-type: none"> • Positive and negative controls, as well as molecular size markers, should be included on each gel and blot. • Quantitative comparisons between samples on different gels or blots are discouraged. • Uncropped source data should be provided to complement cropped gels or blots. • Cropped blots in the body of the paper should retain at least six bandwidths above and below the band. • High-contrast gels and blots are discouraged, as overexposure may mask additional bands.

Microscopy¹⁵

- Authors should be prepared to supply the editors with source data on request, at the resolution collected, from which their images were generated.
- Cells from multiple fields should not be juxtaposed in a single field.

Clinical patient images¹⁶

- Authors should remove any information from photographs and manuscripts that might identify a patient.
- Informed consent should be obtained from all patients for whom identifying information is included in the article.

Dealing with academic misconduct related to images

Most academic misconduct related to images is discovered after publication. If serious misconduct is confirmed after investigation by the journal and admitted by the authors, a retraction note should be published immediately and the label 'Retracted' added to the original paper.¹⁷

If the disputed image was used unintentionally and does not affect the validity and the logic of the results, an erratum or correction should be published accordingly.

If misconduct cannot be confirmed, or the author cannot provide an acceptable explanation, the editorial office may consider contacting the author's university or a relevant professional association to assist in the investigation and publish an expression of concern in the journal.^{14,18}

Checking for misconduct in images

Crossref Similarity Check

Crossref Similarity Check (formerly Crosscheck) is widely used in looking for similarities in blocks of text in manuscripts and published literature. Although inapplicable to images, some textual elements of an image, such as titles, captions, and legends, can be matched and may help in locating the relevant source image. To judge whether any misconduct is involved, the best way is to compare the two images in question is to do so manually. An example is shown in Fig. 1(a), drawn from an instance in which one submitted paper was subjected to a Crossref Similarity Check. The comparison highlighted the similarity between the figure caption as submitted and that accompanying an earlier published figure – Figure 1(a) and (b), respectively – which enabled us, after reviewing and comparing with the highlighted source, to trace the figure to a book that had been published earlier.¹⁹

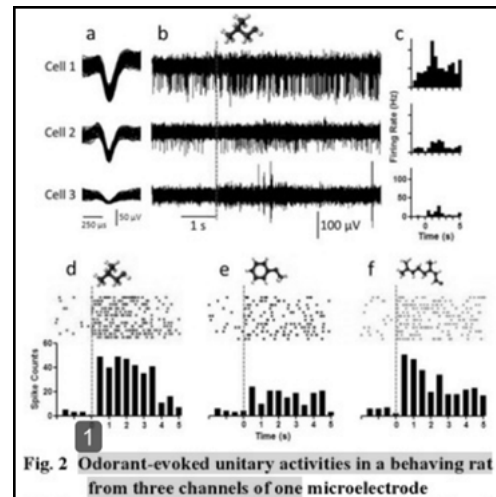


Fig. 2 Odorant-evoked unitary activities in a behaving rat from three channels of one microelectrode

(a)

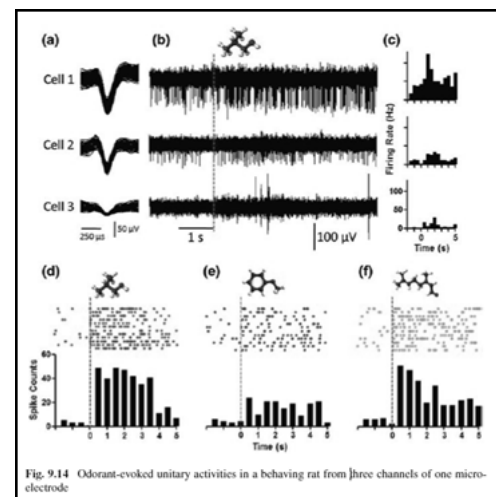


Fig. 1.14 Odorant-evoked unitary activities in a behaving rat from three channels of one microelectrode

(b)

Figure 1. Crossref Similarity Check report (a) and the similar source (b)

(a) The submitted figure, when screened for text similarity, with the figure caption highlighted and (b) a similar figure traced to a published book. Reprinted with permission from Springer Nature: *In Vivo Bioelectronic Nose* by Zhuang *et al.* (2015)

Motuin

Motuin is a software package developed by a Chinese company (iPlagiarism, Beijing), which makes it possible to examine the authenticity of an image for tampering, forgery, and similarity. A reviewer reported improper duplication in a submitted image – Figure 2(a) – and the allegation of misconduct was confirmed by Motuin. The similarity report showed a series of parallel diagonal green lines, which can be seen in Figure 2(b). Bright green lines that straddle the two images being compared suggest similar contents, pointing to possible misconduct involving copy-paste or copy-move forgery.

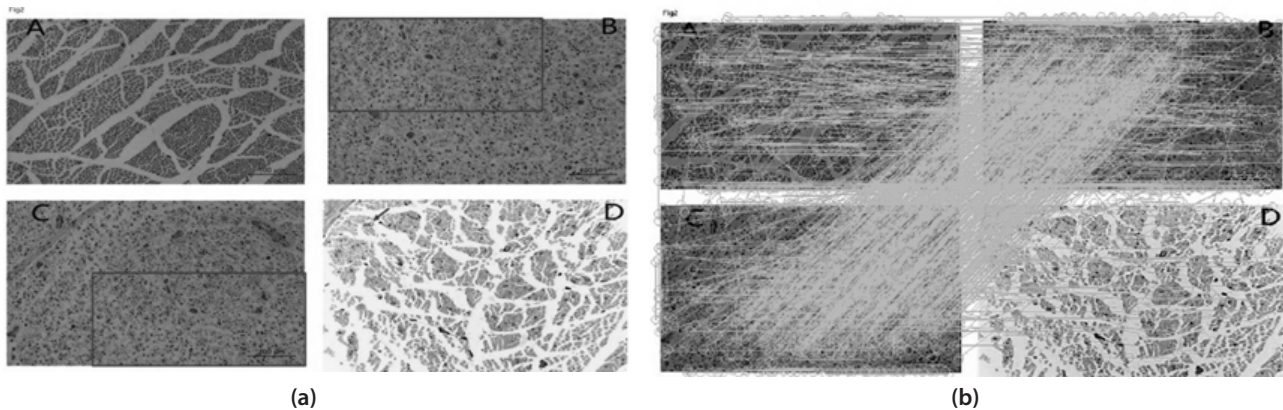


Figure 2. Improper duplication confirmed by reviewer (a) and Motuin (b)

(a) Duplicated regions pointed out by reviewer, marked by (red) rectangle. (b) Detection of similarity by Motuin. Bright (green) lines that extend across two images indicate similarity of contents and possible forgery involving copy-paste or copy-move.

Droplets

The US Office of Research Integrity recommends Droplets for a quick examination of scientific images. Droplets are created by recording and saving sequences of actions in Photoshop: the areas of concern are displayed in a colour different from the background colour with clearly marked boundaries. Common forms of manipulation, such as cropping, erasing, copying, rotating, and patching, can be detected (Figure 3).

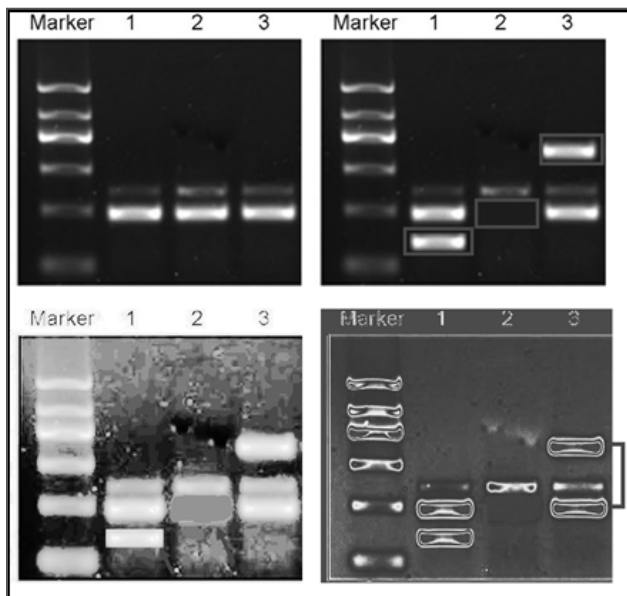


Figure 3. Detecting image manipulation by using 'Droplets' toolkit.

(Top left) Original image. (Top right) Manipulation of original image (red rectangles): 1, copying; 2, erasing; 3, patching by stamp tool in Photoshop. (Bottom left) Manipulated zones (red arrows) detected by using the 'Features in Dark or Light Areas' of Droplets (<https://ori.hhs.gov/droplets>). (Bottom right) Copied bands (lines) with similar graphic structures are detected by using the 'Forensic Gradient Map' of Droplets.

Conclusion

Images should accurately reflect the source data and research results. This broad topic needs more in-depth discussion on such aspects as the role of open data in open science, newly developed algorithms for image comparison, and the proper course of action to deal with possible misconduct. Although current methods of detecting image misconduct continue to depend on professional domain expertise, are not fully automated, and take up a great deal of time, detecting such misconduct is essential to maintaining scientific integrity. We urge authors, editors, and even publishers to pay more attention to academic misconduct involving fabricated, plagiarized, or manipulated images and believe that greater demand for tools to detect such misconduct will spur their development.

References

- 1 Bik EM, Casadevall A, Fang FC. The prevalence of inappropriate image duplication in biomedical research publications. *MBio*. 2016, 7(3):e00809-16. DOI: 10.1128/mBio.00809-16.
- 2 Smart P, Gaston T. How prevalent are plagiarized submissions? Global survey of editors. *Learned Publishing*. 2019, 32(1):47–56. DOI: 10.1002/leap.1218.
- 3 The Lancet Editors. Retraction—engineered whole organs and complex tissues. *The Lancet*. 2018, 392(10141):11. DOI: 10.1016/S0140-6736(18)31560-5.
- 4 PubPeer. Retracted: engineered whole organs and complex tissues. Available: <https://pubpeer.com/publications/910C2D0D4837D572131F744F0BD575#> (accessed 11 June 2019)
- 5 Zhang YH. Against plagiarism: a guide for editors and authors. Springer, 2016. DOI: 10.1007/978-3-319-24160-9.
- 6 COPE. English: all flowcharts. Available: <https://publicationethics.org/resources/flowcharts-new/translations> (accessed 11 June 2019)
- 7 Taichman DB, Sahni P, Pinborg A, et al. Data sharing statements for clinical trials. *BMJ*, 2017, 357:j2372. DOI: 10.1136/bmj.j2372.
- 8 Rossner M, Yamada KM. What's in a picture? The temptation of image manipulation. *Journal of Cell Biology*, 2004, 166(1):11–15. DOI: 10.1083/jcb.200406019.
- 9 Van Noorden R. The image detective who roots out manuscript flaws. *Nature News*. 2015. Available: DOI: 10.1038/nature.2015.17749 (accessed 11 June 2019).

References continued on page 71