Vol 4 (2023) No 3

Jitsi Jurnal Ilmiah Teknologi Sistem Informasi

URL: http://jurnal-itsi.org ISSN 2722-4619 (Print) ISSN 2722-4600 (Online)

A Bibliometrics Analysis of Multimedia Forensics and Deep Learning Research Based on Scopus Index

Erika Ramadhani[#],Dedy Hariyadi[*]

[#] *Department of Informatics, Universitas Islam Indonesia, Indonesia*
[*] *Universitas Jenderal Achmad Yani Yogyakarta, Yogyakarta, Indonesia*
E-mail: erika@uii.ac.id

A B S T R A C T S

Forensics of digital data are concerned with identifying, acquiring, processing, analysing, and reporting on electronic data. Multimedia forensics focuses on investigating computer crimes using forensic methods. The analysis of multimedia evidence is the role of multimedia forensics, on the other hand. In the analysis, digital evidence is evaluated scientifically to maintain its integrity, to find its source, and to authenticate it. There are several methods in multimedia forensics, such as the implementation of deep learning. The purpose of this research is to conduct bibliometric analysis in multimedia forensics and deep learning by using the data gathered from Scopus index on the keyword "multimedia forensics and deep learning". The result is 68 relevant papers were found in the range of 2017-2022. The results of this research can be used by researchers as a reference when conducting research and determining the research themes to be pursued.

Manuscript received September 16, 2023; revised August 25, 2023. accepted September 17, 2023 Date of publication September 30, 2023 International Journal, JITSI : Jurnal Ilmiah Teknologi Sistem Informasi licensed under a Creative Commons Attribution-Share Alike 4.0 International License

Keywords / Kata Kunci — ***Digital forensics; multimedia forensics; deep learning and cnn; bibliometric analysis; data analysis***

1. INTRODUCTION

As a standard tool in science policy and research management, bibliometrics is a set of methods for analysing or measuring texts and information [1]. Especially in academia, citation analysis is increasingly being used to determine hiring, promotion, tenure, and funding decisions [2]. In scholarly publishing, bibliometric studies typically analyse metadata elements like author, title, subject, citations, etc. relating to publications within a discipline. The results of this type of analysis provide useful indicators of scientific productivity, trends, research emphasis, and researchers' publication preferences [3]. Bibliometrics is the organization, classification, and quantitative evaluation of macro-communications and their authorships through mathematical and statistical calculations, according to [4].

A digital forensics expert is someone who identifies, acquires, processes, analyses, and reports on data stored electronically[5]. Law enforcement investigations require digital forensic support because electronic evidence is a component of almost all criminal activities. There are a number of sources of electronic evidence, including computers, smartphones, remote storage, unmanned aerial systems, shipborne equipment, and more. Digital forensics consists of recovering data from electronic evidence, processing it into actionable intelligence, and presenting the findings to a prosecutor. To ensure the findings are admissible in court, all processes use sound forensic techniques[6].

Multimedia forensics is the science in which only a particular digital asset is analysed in order to attempt to provide an assessment on the content and to extract information that will address and support a specific investigation related to the scene represented in that specific digital document. The study of multimedia forensics involves the analysis of digital multimedia content, such as photos, videos, and audio, which are used in forensic investigations. Forensic image analysis involves investigating still images through authentication and integrity

checks (Forgery Detection) and reconstruction of the history of the image since its acquisition (Image Ballistics) [7].

In machine learning and artificial intelligence (AI), deep learning mimics how humans gain knowledge. Data science includes statistics and predictive modelling, which includes deep learning. The use of deep learning can be extremely beneficial to data scientists who are responsible for collecting, analysing, and interpreting large amounts of data. Essentially, deep learning is a way to automate predictive analytics. The complexity and abstraction of deep learning algorithms are stacked in a hierarchy, as opposed to traditional machine learning algorithms which are linear in nature.

Since 2004, Elsevier has been publishing abstracts and citations in Scopus. A total of 36,377 titles are covered by Scopus (22,794 active titles and 13,583 inactive titles) from approximately 11,678 publishers, including 34,346 peer-reviewed journals in top-level subject fields: life sciences, social sciences, physical sciences, and health sciences. The database covers three types of sources: book series, journals, and trade journals. In the Scopus database, all journals indexed are reviewed annually to ensure sufficiently high quality. This is done using four types of numerical measures, namely h-Index, CiteScore, SJR, and SNIP. In Scopus, patents are also searched[8]. Previous studies regarding bibliometric analysis in various fields have been conducted. Many papers discussed about bibliometric analysis, such as: (1) digital forensics[9], (2) social media marketing[10], (3) Covid-19 pandemic[11], etc. The aim of this study is to conduct bibliometric analysis in multimedia forensics and deep learning. This study used a qualitative descriptive method. The results of this study were described using data. The journal data we collected is based on research published in Scopus-indexed journals. Data from the journal concerned multimedia forensics and deep learning. As a reference, we used the Publish or Perish application system to gather the data

2. RESEARCH METHODOLOGY

2.1. Data Sources

An analysis of bibliometric data was performed by using the Scopus database. In this article, the keywords are digital forensics and deep learning. Scopus uses the terms "multimedia forensics" and ("deep learning" or CNN). The year of the search was not determined during the search process. Moreover, CNN was used as a keyword in the search.

2.2 Data Extraction and Analysis

The document will be downloaded to the computer in the .RIS and .CSV formats using a predetermined search query in Scopus. Bibliometric data can be more easily analysed when documents are formatted in .RIS and .CSV. A bibliometric analysis provides a variety of parameters, such as publication trends, analyses of contributing publishers and authors, and keyword co-occurrence networks.

Based on our search, 69 papers have been published between 2017 and 2022. Our analysis of the data was further refined to the specific topic and 68 documents were found. In the process of cleaning, one paper was found to be irrelevant.

2.3 Network Map

The VOSViewer application is used to conduct bibliometric analysis based on title, abstract, and relevant keywords. Clustering results are visualized using a network map. Several points on these maps show the frequency of scientific publications as well as the colour of the network showing the relevance of research publications.

3. RESULTS AND DISCUSSION

3.1 Publications trends

In Figure 1, a search query without a range of years returns a collection of articles from 2017-2022. A total of five proceedings, lecture notes, and journals will be published in 2017. Detection of image manipulation, audio steganalysis, design principles for multimedia forensics, and content recapture detection are the topics covered. Various methods are used, including deep neural networks and convolutional neural networks.

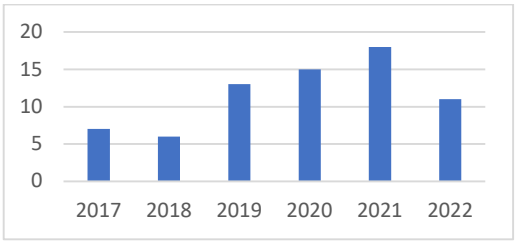

FIG 1. Digital forensics and deep learning publication trends

Most of the papers in 2018 were published in proceedings. There is also a focus on image manipulation, the localization of tempered regions in images, the detection of source similarity, and the detection of contrast adjustments. Deep learning and CNN are the methods used.

There has been an increase in research into multimedia forensics in 2019. There are nine proceedings published by IEEE publisher, three lecture notes, and a handbook. Topics covered include splicing detection in digital images, steganographic attacks, camera model identification, adversarial attacks, steganalysis on GPUs, social network identification, image deblocking detection, and tempered video detection.

In 2020 - 2022, the topics vary based on the method for detecting forgeries in digital images. In the majority of cases, digital images are forgeries and manipulations that are detected, along with the camera model as well. There are primarily CNNs and deep learning methods used.

3.2 Keyword Co-occurrence network

Based on the keyword co-occurrence network in Figure 2, the article can be used as a research interest. A total of four clusters (red, green, blue, and yellow) were associated with different topics. A density visualization of the keyword co-occurrence network can be seen in Figure 3. Figure 3 shows an increase in research numbers as the colours become more concentrated.

According to Figure 1-2, the keywords often appearing were multimedia forensics, digital forensics, deep learning, convolutional neural networks, and neural networks. Data from these sources can be used to investigate the existence of multi-media forensics and deep learning research. By looking at keyword density, it appears that there has been little work done by other researchers on multimedia forensics in terms of texture. Perhaps this is due to the fact that the topic area does not need to be researched

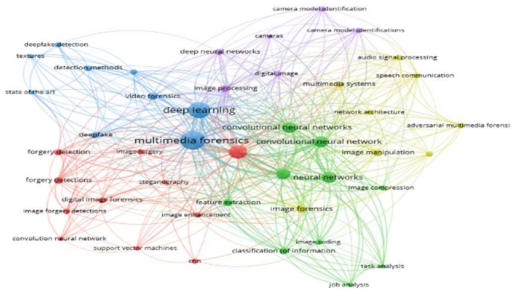

FIG 2. Keyword cooccurrence network

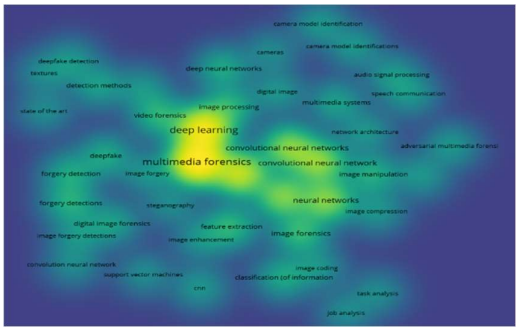

FIG 3. Keyword cooccurrence density

There was a limitation on the number of keywords that could be linked to a research topic. An analysis of the relationship between the terms was conducted using a data set of research articles. In Table 1, six keyword clusters are identified as a result of the analysis of the keyword co-occurrence network in articles. The clusters describe potential subfields of research as well as the relationship between each subfield and the other fields. It is helpful for researchers to know possible keywords so that they can find references to support their research

TABLE 1. Networking map of keyword cooccurrence

Cluster	Keywords	Description
Cluster 1	cnn, convolutional neural network, digital forensics, digital image forensics, forgery detection, forgery detections, image enhancement, image forgery, image forgery detection, steganography, support vector machine	Multimedia forensics activity such as digital image forgery detection using cnn, svm, image enhancement, and steganography.
Cluster 2	classification of information, convolutional neural network, convolutional neural networks, feature extraction, image coding, image compression, job analysis, neural network, task analysis	Analysis of digital image such as image compression, image coding, job analysis, task analysis, classification of information using cnn, neural network, and feature extraction
Cluster 3	deep learning, deepfake, deepfake detection, detection methods, multimedia forensics, semantics, state of the art, textures, video forensics	Any activity in multimedia forensics such as semantics, state of the art, textures, video forensics, deepfake, deepfake detection using some detection method and deep learning
Cluster 4	adversarial digital forensics, audio signal processing, image forensics, image manipulation, image manipulation detection, multimedia systems, network architecture, speech communication	Multimedia forensics in image manipulation, multimedia system by using audio signal processing, speech communication
Cluster 5	camera model identification, camera model identifications, cameras, deep neural networks, digital image, image processing	Identification of source model such as camera by using image processing and deep neural network

3.3 Visualization country of study

Besides bibliometric analysis, we also examine author information, journal information, and the country of study as well as language used in the writing. Figure 4 shows bibliometric analysis for the country of study.

Multimedia Forensics and Deep Learning in the Scopus database provides 69 results for the term country. A closer analysis reveals, however, that only seven countries have conducted such a study. Chinese researchers have conducted the most studies, followed by Americans and Italians.

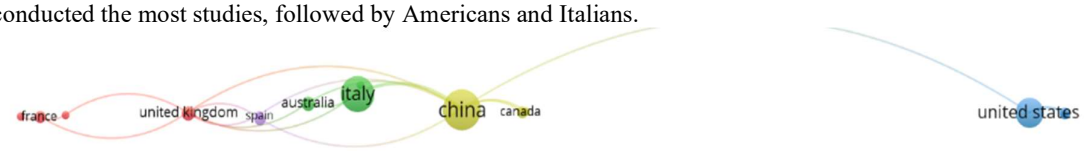

FIG 4. Country of study

3.4 Most cited articles

Digital forensics and deep learning articles cited most frequently are represented in Table 2. In the article 'Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection', 193 citations were found. A novel approach to image manipulation detection is presented in this article published in IEEE Transactions on Information Forensics and Security.

TABLE 2. Most cited paper reinforcement

Cites	Title	Source	Year
193	Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection [12]	IEEE Transactions on Information Forensics and Security	2018
126	Noiseprint: A CNN-Based Camera Model Fingerprint[13]	IEEE Transactions on Information Forensics and Security	2020
78	Design principles of convolutional neural networks for multimedia forensics[14]	IS and T International Symposium on Electronic Imaging Science and Technology	2017
37	Forensic Similarity for Digital Images[15]	IEEE Transactions on Information Forensics and Security	2020
36	Emotions Don't Lie: An Audio-Visual Deepfake Detection Method using Affective Cues[16]	MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia	2020

4. CONCLUSIONS

A bibliometric analysis of the literature on digital forensics and deep learning was conducted in this study. There are two categories of digital forensics in this study: computer forensics and multimedia forensics. As part of the search process, the keyword "multimedia forensics and deep learning" was chosen, which was based on a topic area with titles, keywords, and abstracts. Because CNN is a component of deep learning, it is also used in the search process. We found 69 relevant articles after conducting the search. Once the mapping procedure had been completed, VOSviewer was used to complete the process. According to the results of analysis and mapping with VOSviewer, multimedia forensics and deep learning research were the most studied topics in 2019-2021.

In the field of multimedia forensics, CNN was the most extensively researched. According to the VOSviewer analysis, multimedia forensics and deep learning are related to five keyword clusters. There are main terms in each cluster, which are linked with other terms. Cluster 1 included terms such as image forgery detection and CNN. In Cluster 2-6, terms included were classification of information, adversarial digital forensics, detection of image manipulation, deepfake detection, and camera identification.

From the search result, it can be concluded that multimedia forensics research does not grow significantly each year. The topic of research has also always been forgery detection in image forensics. There is a lack of research in audio forensics, video analysis, and task analysis. A gap in research can be considered a recommendation for further research or as confirmation that the topic is not expanding every year due to a lack of need for further research. A gap in research can be considered a recommendation for further research or as confirmation that the topic is not expanding every year due to a lack of need for further research.

ACKNOWLEDGMENT

The authors acknowledge the support received from Universitas Islam Indonesia and Universitas Jendral Achmad Yani Yogyakarta for publishing this paper. The authors also acknowledge to the team of reviewer**.**

REFERENSI

[1] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, How to conduct a bibliometric analysis: An overview and guidelines. J. Bus. Res., vol. 133, pp. 285–296, 2021, doi: https://doi.org/10.1016/j.jbusres.2021.04.070.

[2] P. Weingart, Impact of bibliometrics upon the science system: Inadvertent consequences?, Scientometrics, vol. 62, no. 1, pp. 117–131, 2005, doi: 10.1007/s11192-005-0007-7.

[3] D. Jacobs and P. Ingwersen, A Bibliometric Study of the Publication Patterns in the Sciences of South African Scholars 1981–96, Scientometrics, vol. 47, no. 1, pp. 75–93, 2000, doi: 10.1023/A:1005617825947.

[4] D. F. Thompson and C. K. Walker, A descriptive and historical review of bibliometrics with applications to medical sciences, Pharmacother. J. Hum. Pharmacol. Drug Ther., vol. 35, no. 6, pp. 551–559, 2015.

[5] T. Authors, T. Url, P. Date, and A. This, Digital forensics trends and future, 2014.

[6] Y. Prayudi and A. Ashari, A Study on Secure Communication for Digital Forensics Environment, Int. J. Sci. Eng. Res., vol. 6, no. 1, pp. 1036–1043, 2015, doi: 10.14299/ijser.2015.01.010.

[7] S. Battiato, O. Giudice, and A. Paratore, Multimedia Forensics : discovering the history of multimedia contents, no. June, pp. 23–24, 2016.

[8] A. V Kulkarni, B. Aziz, I. Shams, and J. W. Busse, Comparisons of Citations in Web of Science, Scopus, and Google Scholar for Articles Published in General Medical Journals, JAMA, vol. 302, no. 10, pp. 1092–1096, Sep. 2009, doi: 10.1001/jama.2009.1307.

[9] M. Kusuma, D. Hariyadi, Fazlurrahman, and M. A. Nugroho, The Bibliometric Analysis the Digital Forensics Researcher in Indonesia Based on Garba Rujukan Digital: 2008–2020. in 2021 IEEE Mysore Sub Section International Conference (MysuruCon), Oct. 2021, pp. 13–17. doi: 10.1109/MysuruCon52639.2021.9641641.

[10] D. Salimi, K. Tavasoli, E. Gilani, M. Jouyandeh, and S. Sadjadi, The impact of social media on marketing using bibliometrics analysis, Int. J. Data Netw. Sci., vol. 3, no. 3, pp. 165–184, 2019.

[11] A. B. D. Nandiyanto, D. N. Al Husaeni, and D. F. Al Husaeni, A bibliometric analysis of chemical engineering research using vosviewer and its correlation with Covid-19 pandemic condition, J. Eng. Sci. Technol., vol. 16, no. 6, pp. 4414–4422, 2021.

[12] B. Bayar and M. C. Stamm, Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection, IEEE Trans. Inf. Forensics Secur., vol. 13, no. 11, pp. 2691–2706, 2018, doi: 10.1109/TIFS.2018.2825953.

[13] D. Cozzolino and L. Verdoliva, Noiseprint: A CNN-Based Camera Model Fingerprint, IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 144–159, 2020, doi: 10.1109/TIFS.2019.2916364.

[14] B. Bayar and M. C. Stamm, Design principles of convolutional neural networks for multimedia forensics, IS T Int. Symp. Electron. Imaging Sci. Technol., pp. 77–86, 2017, doi: 10.2352/ISSN.2470-1173.2017.7.MWSF-328.

[15] O. Mayer and M. C. Stamm, Forensic Similarity for Digital Images, IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 1331–1346, 2020, doi: 10.1109/TIFS.2019.2924552.

[16] T. Mittal, U. Bhattacharya, R. Chandra, A. Bera, and D. Manocha, Emotions Don’t Lie: A Deepfake Detection Method using Audio-Visual Affective Cues, ArXiv, vol. abs/2003.06711, 2020