information

MDPI

Article

Living in the Age of Deepfakes: A Bibliometric Exploration of Trends, Challenges, and Detection Approaches

Adrian Domenteanu [1], George-Cristian Tătaru [1], Liliana Crăciun [2], Anca-Gabriela Molănescu [2], Liviu-Adrian Cotfas [1] and Camelia Delcea [1,*]

[1] Department of Economic Informatics and Cybernetics, Bucharest University of Economic Studies, 0105552 Bucharest, Romania
[2] Department of Economics and Economic Policies, Bucharest University of Economic Studies, 0105552 Bucharest, Romania
* Correspondence: camelia.delcea@csie.ase.ro; Tel.: +40-769652813

Abstract: In an era where all information can be reached with one click and by using the internet, the risk has increased in a significant manner. Deepfakes are one of the main threats on the internet, and affect society by influencing and altering information, decisions, and actions. The rise of artificial intelligence (AI) has simplified the creation of deepfakes, allowing even novice users to generate false information in order to create propaganda. One of the most prevalent methods of falsification involves images, as they constitute the most impactful element with which a reader engages. The second most common method pertains to videos, which viewers often interact with. Two major events led to an increase in the number of deepfake images on the internet, namely the COVID-19 pandemic and the Russia–Ukraine conflict. Together with the ongoing "revolution" in AI, deepfake information has expanded at the fastest rate, impacting each of us. In order to reduce the risk of misinformation, users must be aware of the deepfake phenomenon they are exposed to. This also means encouraging users to more thoroughly consider the sources from which they obtain information, leading to a culture of caution regarding any new information they receive. The purpose of the analysis is to extract the most relevant articles related to the deepfake domain. Using specific keywords, a database was extracted from Clarivate Analytics' Web of Science Core Collection. Given the significant annual growth rate of 161.38% and the relatively brief period between 2018 and 2023, the research community demonstrated keen interest in the issue of deepfakes, positioning it as one of the most forward-looking subjects in technology. This analysis aims to identify key authors, examine collaborative efforts among them, explore the primary topics under scrutiny, and highlight major keywords, bigrams, or trigrams utilized. Additionally, this document outlines potential strategies to combat the proliferation of deepfakes in order to preserve information trust.

Keywords: deepfake; deepfake image detection; deepfake video detection; machine learning; bibliometric analysis; strategies

Citation: Domenteanu, A.; Tătaru, G.-C.; Crăciun, L.; Molănescu, A.-G.; Cotfas, L.-A.; Delcea, C. Living in the Age of Deepfakes: A Bibliometric Exploration of Trends, Challenges, and Detection Approaches. *Information* **2024**, *15*, 525. https://doi.org/10.3390/info15090525

Academic Editors: Arkaitz Zubiaga and Katsuhide Fujita

Received: 6 June 2024
Revised: 9 August 2024
Accepted: 20 August 2024
Published: 28 August 2024

1. Introduction

Starting in the 21st century, a trend emerged to transpose physical information found in books and newspapers to online sources, making the information more easily accessible for everyone. The creation of a public internet system was a great success, and nowadays, every human activity is dependent on the internet: chatting with friends and family using social media applications, talking using the internet all over the world, reading news, and many other actions. The benefits of internet implementation are incommensurable, but the threats that appeared in recent years affect the quality of the information. Deep fakes are an artificial intelligence (AI) technique used to manipulate videos and photos that look realistic [1]. Deepfakes evolved very fast, and it is getting more and more difficult to observe if a piece of information or a picture is fake or not, creating serious threats to society.

There are two main factors that contributed in a decisive manner to the spread of deepfake: the availability of data and the evolution of technology. One of the most common deepfake applications is face swapping, where facial features of a person are transposed onto another person's face in a photo or video, as Natsume et al. [2] found. Face reenactment is another application of deepfakes, which involves mapping the movements and facial expression of individuals onto other person in a synchronized manner, looking normal [3].

Deepfakes advanced in deep neural networks, generating hyper-realistic synthetic media, and thereby creating more data and interest from researchers, which is reflected in the amount of documents published [4]. In general, deepfake creation and detection, social, legal, and ethical elements have been discussed, but recently, researchers identified deepfake detection methods and quantified the social impact [5,6]. Deepfakes are closely related to the AI domain and differ depending on the communication technology. For instance, AI can behave like a human, chatbot, or virtual assistant [7].

Raghavendra et al. [8] presented the implementation of the electronic Machine-Readable Travel Document (eMRTD), which is available at this moment, especially at borders, where passports are used; a biometric reference image is stored for person verification, facilitating the border control process. The system is useful, but it can be subject to fraud, with manipulated face images generated for some individuals.

Deepfake technology is not limited only to visual content, it also offers the possibility of voice swapping, where audio can be adapted to another voice. A company was defrauded of USD 243,000 because voice software was used in order to convince the CFO that the CEO was requesting the money. Creating ideal messages in the marketing domain represents one of the major success factors, and most companies are using celebrities to promote their products and services. However, by using deepfakes and synthetic information, people can create fake advertising, as Lil Miquela created using computer-generated imagery (CGI). She has 2.7 million followers on Instagram and engages in endorsements with Samsung, Calvin Klein, Dior, and Prada. The tool has a predefined speech and focuses on the brand's objectives, and it is difficult for the companies to detect if the person behind the account is real or not [7].

Researchers have focused their efforts in recent years on defining methods for detecting deepfakes. Rossler et al. [9] created a tool called FaceForensic in March 2018 which detects deepfakes. The key for a good or bad model is mainly dependent on the type of dataset and the selected features. The data included in the analysis should be correlated, otherwise the accuracy could decrease significantly. In some cases, depending on the split of train and test datasets, the accuracy can rise from 50% to 98% [9]. The Face2Face approach has been used [3]; this is able to fully automatically render faces in a target video, obtaining a temporary face identity and calculating an identity fitting. Raghavendra et al. [10] proposed a detection model that uses pre-trained a deep convolution neural network (DCNN), VGG19, and AlexNet. The experiment proposed to identify digital and print-scanned versions of morphed face images. Initially, the Viola–Jones algorithm was used for face detection, and the detected area was then normalized to a specific size, similar to that of the DCNN's input data. The results differed depending on the dataset used. For digital images, the minimum detection equal error rate (D-EER) was 8.23%, for print-scanned (HP) it was 17.64%, and for print-scanned (RICOH) it was 12.47%.

Raghavendra et al. [8] developed a framework that is able to detect morphed face images, using binarized statistical image features (BSIFs) and support vector machines (SVMs). The minimum error calculated using the average classification error rate (ACER) was 1.73%.

Zhang et al. [11] analyzed generative adversarial network (GAN)-generated images. The approach was different, trying to detect artifacts using the frequency spectrum instead of pixel images, and was thus able to incorporate more image types. Using training and testing datasets, and an AutoGAN model, which simulates GAN pictures as real images, the algorithm was able to detect fake images based on their similarities.

Rana et al. [12] conducted a systematic literature review (SLR) by analyzing 112 articles between 2018 and 2020, in order to understand the deepfake detection solutions, and grouped documents into four different categories: deep learning-based methods, classical machine learning-based models, statistical techniques, and blockchain-based techniques. Using various performance metrics, several models were analyzed on different datasets, in order to find the best model for deepfake detection.

Due to the rapid development of AI, the deepfake technology has significant implications for the information landscape. As mentioned above, the development of deepfake research mirrors the trends one can observe in AI and information technology. Over the years, the evolution of the deepfake domain has been observed, from early explorations in face swapping and simple image manipulation, to more sophisticated situations materialized by voice cloning and video synthesis. As a result, when discussing the deepfake domain one can observe the dual nature of the technology advancements it incorporates, offering innovative solutions to various situations but providing, at the same time, challenges to information integrity and security.

The main purpose of this research is to present the evolution of deepfake domains, focusing on the main threats and the ways to encounter them, e.g., by using ML and AI tools to detect fake morphed images and videos. The secondary objectives will provide a complete perspective on the deepfake research domain by answering the following questions:

- How has the evolution of scientific production evolved during the analyzed period?
- Who are the most relevant authors, taking into account the number of papers published and the total number of citations?
- Which are the most notable journals in the domain of deepfakes?
- Which countries collaborated the most on developing scientific articles?
- What are the most used KeyWords Plus terms and author keywords?
- Which are the universities who published the most articles on deepfake domain?

Based on the above research questions, the present study aims to provide more insight into the deepfake research domain and how the field has evolved from a niche academic zone to an important aspect to be considered in various research areas, such as information systems, cybersecurity, and digital media. Thus, with the aim of mapping the evolution of this field, we aim to provide more understanding of the current state of the deepfake research by discussing the main research themes, the most cited papers, the most prominent authors, and the emerging trends that are very likely to emerge and influence the future developments of the field.

The article is structured in different sections, each one focusing on various key elements: Section 1 is the introduction to the domain, presenting the history of the domain. Section 2 presents the main steps of extracting the database from the Clarivate Analytics Web of Science Core Collection. Section 3 explores the data, and is divided into multiple subsections: data overview, sources, authors, countries and affiliations, words, and mixed analysis. Section 4 presents the limitations, while the last section focuses on discussion and conclusions.

2. Materials and Methods

In order to achieve the mentioned objectives, a bibliometric analysis was conducted through the use of a dataset extracted from the Clarivate Analytics Web of Science Core Collection [13] based on a series of keywords, as presented in the following, and analyzed through the Bibliometrix 4.0.0 package available in R Studio 4.3.2 [14]. The bibliometric analysis was chosen as the main technique as it offers a broad overview of a specific field in terms of authors, their collaborations, affiliations, country of origin, number of papers published, preferred sources, and citations [15–18]. According to Block and Fisch [19], it is important to distinguish between bibliometric analysis and review analysis. The latter focuses on summarizing the content and key findings within a field, while bibliometric analysis is primarily used to highlight the structure and development of a particular

field [18,19]. Additionally, in order to provide more insight into the deepfake field, a review of the most cited papers, along with a review of some of the papers with a large number of citations, was performed and is presented in the next section.

Although there are numerous article databases available, Bakir et al. [20] presented the reasons that justify the use of the Clarivate Analytics Web of Science Core Collection (also known as ISI Web of Science, or WoS) as the most suitable for a bibliometric analysis. Among these reasons, the authors state that the WoS database offers a great variety of indexed journals, most of which are highly appreciated and known by the scientific community, and are most frequently used in the scientific literature [21,22]. Using the bibliometric approach, numerous data can be extracted from the database, to find correlations between journals, authors, and countries, identify the most used keywords, and identify the key topics analyzed [23,24]. All available article databases (e.g., Scopus, Google Scholar, IEEE, Cochrane Library, and others) can be analyzed from a bibliometric point of view, but the subscription plans offered by the Clarivate Analytics Web of Science Core Collection are various, and Liu [25] and Liu [26] explained the importance of them. According to both authors, it is mandatory in a bibliometric analysis to present the indexes available using the subscription. Therefore, it shall be stated that, in our case, the database was searching through the following indexes:

- Emerging Sources Citation Index (ESCI)—2005–present;
- Current Chemical Reactions (CCR-Expanded)—2010–present;
- Book Citation Index—Science (BKCI-S)—2010–present;
- Arts and Humanities Citation Index (A&HCI)—1975–present;
- Book Citation Index—Social Sciences and Humanities (BKCI-SSH)—2010–present;
- Index Chemicus (IC)—2010–present;
- Conference Proceedings Citation Index—Science (CPCI-S)—1990–present;
- Science Citation Index Expanded (SCIE)—1900–present;
- Conference Proceedings Citation Index—Social Sciences and Humanities (CPCI-SSH)—1990–present;
- Social Sciences Citation Index (SSCI)—1975–present;

Table 1 summarize the main steps applied on the Clarivate Analytics Web of Science Core Collection in order to extract the database. In the first step, titles were filtered, looking for "deep_fake*", "deepfake*", or "deep-fake*", returning 918 documents. It shall be noted that we have used the asterisk at the end of the search keywords in order to retain in the analysis both the singular and the plural form of the keywords. In the second step, abstracts were analyzed, applying the same filters, extracting 1173 papers. In the third step, the filters were used for keywords, finding 870 articles. In step four, titles, abstracts, and keywords were filtered, with at least one required to have one specific word related to deepfakes, resulting in 1381 documents. In the fifth step, the articles published in any language other than English were removed, reducing the total database from 1381 to 1339 documents. The sixth filter is related to document type, keeping in the analysis only papers marked as "article", resulting in 707 articles. The last filter excludes the 2024 year, finally resulting in 584 papers. The decision to exclude the papers published in 2024 was based on the fact that the mentioned year was ongoing at the time of the analysis and the partial inclusion of some of the papers published in this year might have introduced bias, affecting the findings. Furthermore, it was observed that when aiming to reduce the inconsistency in the citations' indicators, the authors tend to use cut-off points related to the period under investigation for data collection [27,28]. The reader can consider the work of Liu [29] for an extensive discussion related to the online publication date versus the final publication dates for the papers included in the WoS database, which supports the use of the cut-off dates in such analyses [30].

Table 1. Data selection steps.

Exploration Steps	Questions on Web of Science	Description	Query	Query Number	Count
1	Title	Contains specific keywords related to deepfakes	(TI = (deep_fake*) OR TI = (deepfake*)) OR TI = (deep-fake*)	#1	918
2	Abstract	Contains specific keywords related to deepfakes	(AB = (deep_fake*) OR AB = (deepfake*)) OR AB = (deep-fake*)	#2	1173
3	Keywords	Contains specific keywords related to deepfakes	(AK = (deep_fake*) OR AK = (deepfake*)) OR AK = (deep-fake*)	#3	870
4	Title/Abstract/Keywords	Contains specific keywords related to deepfakes	#1 OR #2 OR #3	#4	1381
5	Language	Contains only documents written in English	(#4) AND LA = (English)	#5	1339
6	Document Type	Limit to Article	(#5) AND DT = (Article)	#6	707
7	Year published	Exclude 2024	(#6) NOT PY = (2024)	#7	584

3. Data Exploration and Analysis

In this section, the dataset is investigated from different perspectives: an initial overview, sources, authors, countries and affiliations, words, and mixed perspectives. The scope is to determine the most influential authors and affiliations, visualize the production evolution, and identify the most relevant words relating to the deepfake topic.

3.1. Dataset Overview

In order to have a complete overview of the data, a summarized analysis was performed on the dataset, extracting the main information, such as the number of documents, authors, references, and timespan.

Table 2 contains some descriptive statistics about the data that were used in the analysis. The period of time in which 584 documents from 284 sources were examined lies between 2018 and 2023; thus, we can notice a growing interest towards this area of literature in recent years. The spike in interest can also be observed from the remarkable number of average citations for each document (10.62). Moreover, the impact that this topic has on the international community is evident from the large number of references available in the papers (20,121).

Table 2. Data description.

Indicator	Value
Timespan	2018:2023
Sources	284
Documents	584
Average citations per documents	10.62
References	20,121
KeyWords Plus terms	445
Author's keywords	1607

Some other essential details about the dataset can be revealed by analyzing the author's keywords and KeyWords Plus terms. With an average of less than 1 keyword per document, KeyWords Plus terms highlight the usage of a fairly concentrated vocabulary. Such a small number of KeyWords Plus terms can be a sign of a very focused research paper. However, the average value of the author's keywords indicator is more than 4 times higher, suggesting that the author has approached subjects with a high level of complexity.

Table 3 summarizes information about the research body that published the papers that were analyzed in our study. As a first remark, we can see that the total number of authors who contributed to the deepfake literature is 1717 and only 93 of these were unique authors. Thus, we can observe that most of the papers that were published were the result of a collaborative effort. This tendency could be explained by the numerous research projects that entail knowledge from various areas, resulting in interdisciplinary teams. Two more arguments that sustain this hypothesis are the values of the documents per author (0.340) and co-authors per document (3.66) indicators.

Table 3. Author information.

Indicator	Value
Authors	1717
Authors of single-authored documents	93
Authors of multi-authored documents	1624
Documents per author	0.340
Co-authors per documents	3.66

Table 4 describes how many articles were published each year and how many citations these articles gathered. We can notice that the number of articles increased almost every year, rising from only 2 papers published in 2018 to 244 papers published in 2023. This indicates a growing interest in research papers dealing with the creation and analysis of deepfakes. However, we can observe a steep decrease in the number of citations, from a maximum value of 11.02 citations in 2019 to just 1.35 citations in 2023. Thus, we can conclude that the very rapid growth in the number of articles has negatively influenced the number of citations per article. Even though there is major interest in this area, there are still some challenges in maintaining relevance for these articles.

Table 4. Annual scientific production and average citations per year.

Year	Number of Documents Published	Average Citations per Year
2018	2	5.64
2019	10	11.02
2020	40	7.82
2021	120	3.71
2022	168	2.89
2023	244	1.35

Figure 1 displays the most prestigious journals from Zone 1 according to Bradford's law [31]. As we can notice, there are 18 journals in this zone, highlighting the impact of deepfakes in different aspects of our life. Firstly, the most popular area is engineering, represented by the "*IEEE Access*" journal with 31 published articles, and the multimedia area, represented by the "*Multimedia Tools and Applications*" journal with 21 published articles. The next journal in the list is "*IEEE Transactions on Information Forensics and Security*", a journal of interest in the security area. Another area of interest in deepfake research literature is artificial intelligence, being represented in the list by two journals, namely "*PeerJ Computer Science*" and "*Expert Systems with Applications*". Some other important journals from different areas are "*Applied Sciences-Basel*", "*IEEE Transactions on Circuits and Systems for video technology*", "*Synthese*", "*Sensors*", and "*Convergence-the International Journal of Research into new media Technologies*" with more than 10 published articles each. The remaining journals, with less than 10 published papers, are the following: "*Journal of imaging*", "*IEEE Signal Processing Letters*", "*Cyberpsychology Behavior and Social Networking*", "*Deep fakes, Fake news, and Misinformation in online Teaching and Learning Technologies*", "*Information Sciences*", and "*Scientific Reports*".

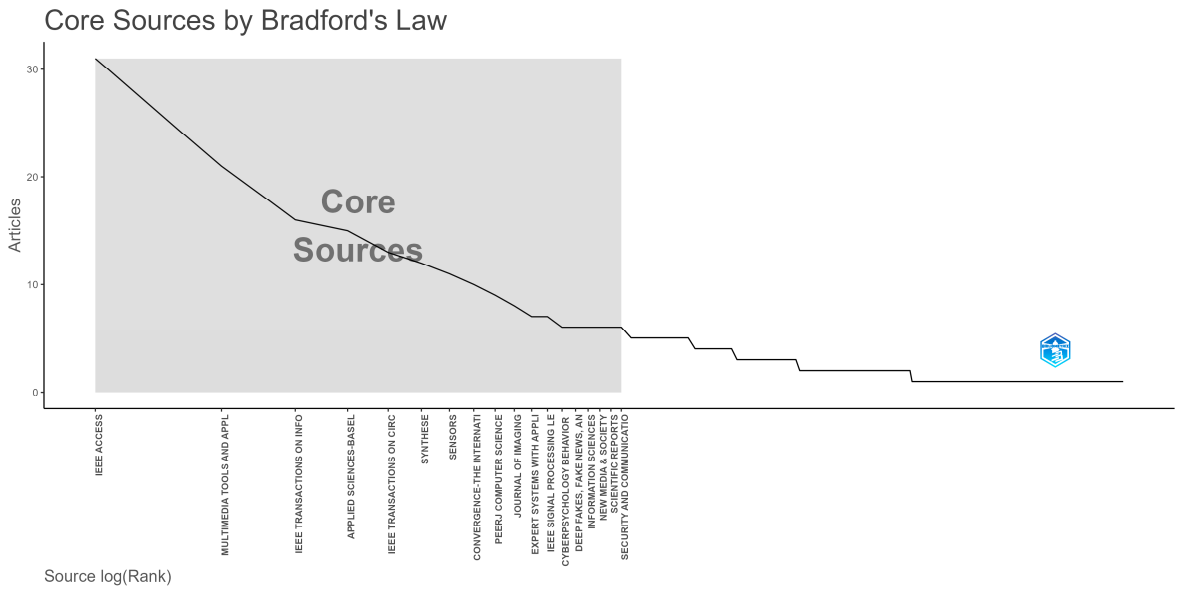

Figure 1. Bradford's law on source clustering.

3.2. Source Analysis

Journals are the main source of information related to any academic domain, and a key step in a bibliometric analysis is to determine the most relevant journals, how many documents they published, and the production during the analyzed timespan.

Figure 2 further highlights the most relevant journals in deepfake literature. As we previously stated, the fields in which these journals are published are diverse and generate great interest. Engineering and applied technologies are the most popular areas of study, followed by areas like multimedia, signal processing, and information security. The most popular journal, with 31 published articles is "*IEEE Access*", followed by "*Multimedia Tools and Applications*" with 21 published articles. Other journals of interest are "*IEEE Transactions on information forensics and security*" which is representative of the area of security, "*IEEE Transactions on Circuits and Systems for Video Technology*" which is representative of the area of computer vision, and "*IEEE Signal Processing Letters*" which is representative of the area of signal processing.

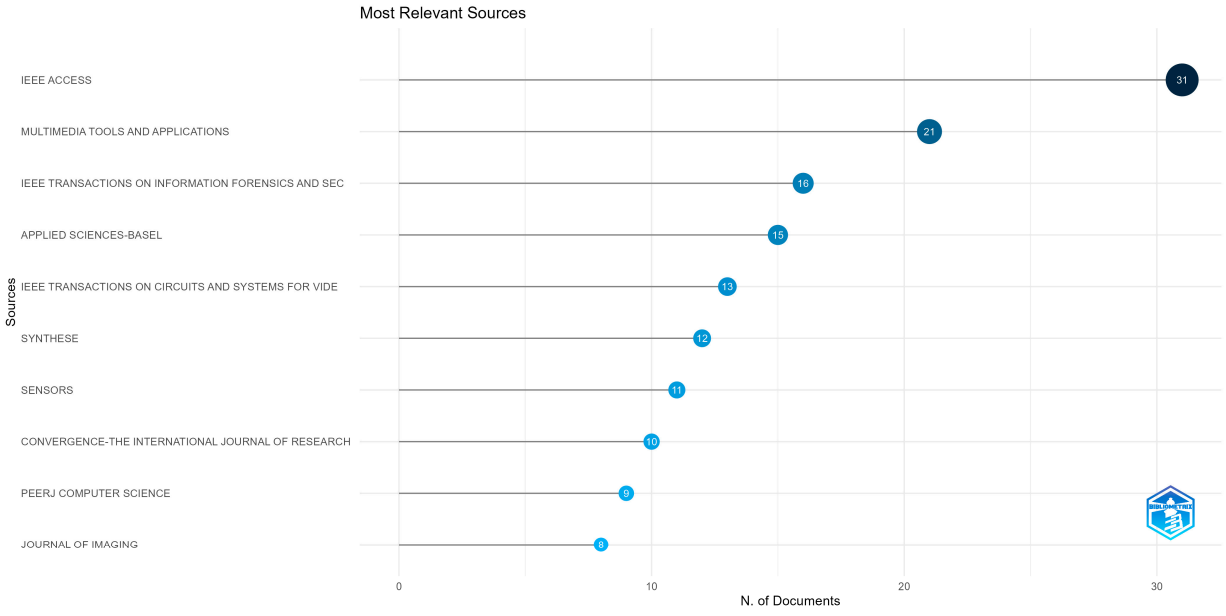

Figure 2. Top 10 most relevant sources.

Figure 3 is closely related to Figure 2, displaying a rank of journals according to the H-index. This value is representative not only of the number of published articles, but also of the impact that each paper generates.

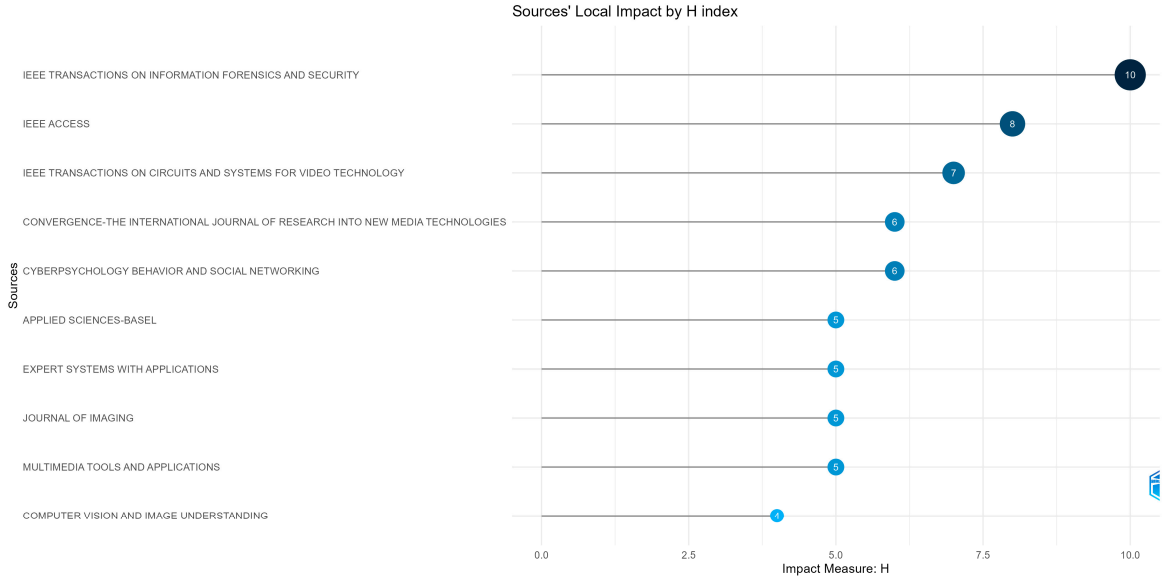

Figure 3. Top 10 sources based on the H-index.

For example, the first journal, "*Multimedia tools and applications*", has an H-index of 5. This result can be interpreted as this journal containing at least five papers with at least five citations each. Thus, we can notice that in opposition to Figure 2, we no longer find in our list journals like "*Synthese*" and "*Sensors*", for which, out of the 12 articles, only three have more than 12 citations. We also observe the new journal "*Cyberpsychology Behavior and Social Networking*", which is in fifth place with a H-index of 6.

Figure 4 presents some data about the number of articles published in different journals each year. As we could expect (taking into account the previous results from Figures 2 and 3), the journals with the highest increase in the number of published papers were "*IEEE Access*" with 31 published articles, and "*Multimedia Tools and Applications*" which has 21 articles.

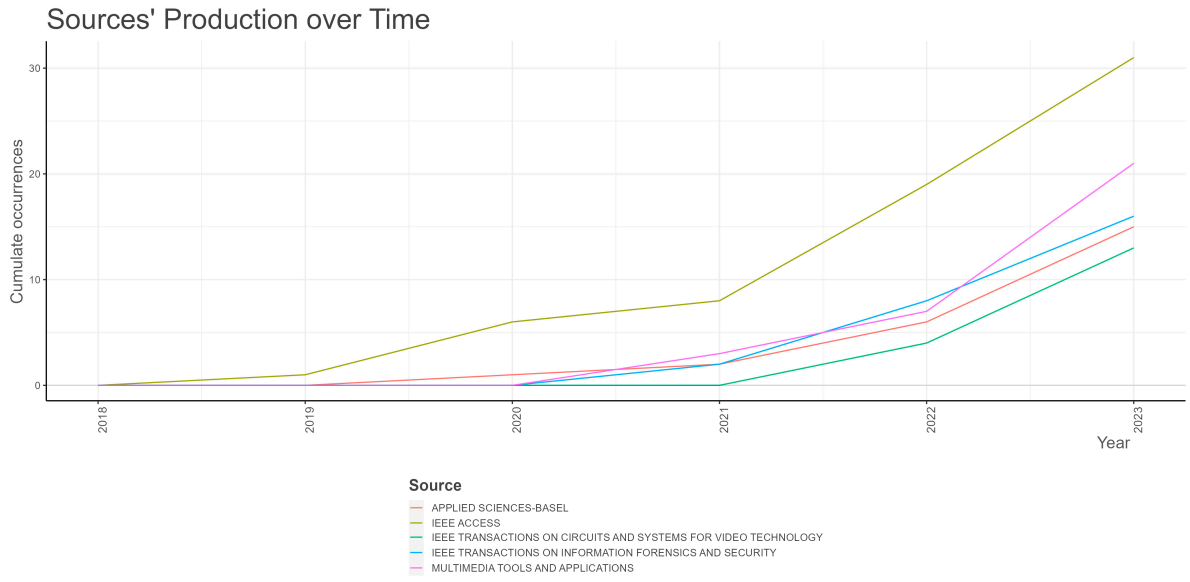

Figure 4. Source production over time.

3.3. Author Analysis

The most important authors, their annual production, and the impact of their papers are analyzed in this section.

Figure 5 presents the relevant authors for the deepfake domain. The most influential authors are Javed, A., Lu, W., Ahmed, S., Zhao, Y., Irtaza, A., Kietzmann, J., Xia, Z.H., and Yang, G.B., who share the first position in our ranking, with each of them contributing more than five research papers. These authors are followed by Cui, X.H. and Guo, Z.Q., with five articles each.

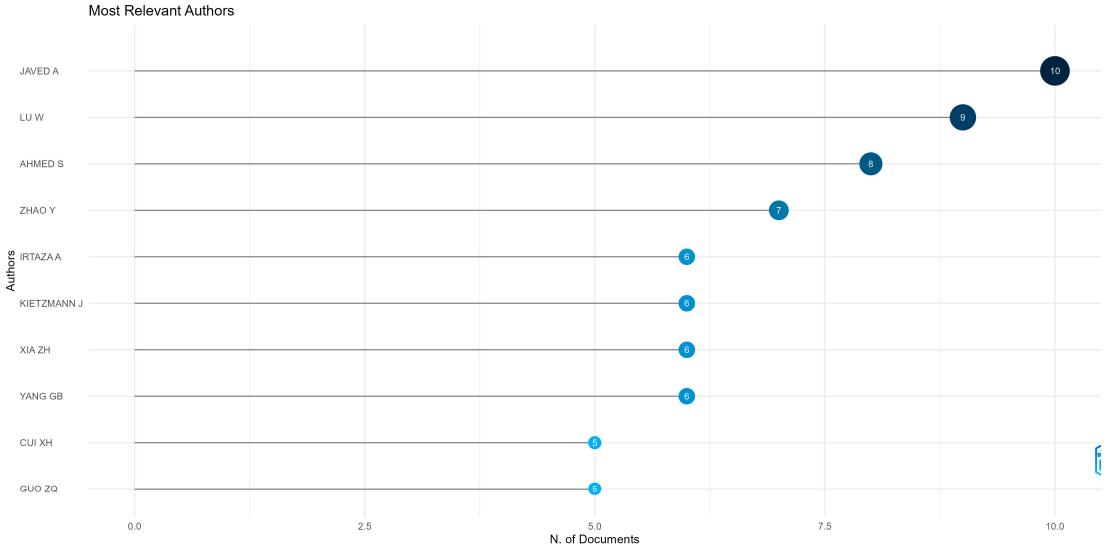

Figure 5. Top 10 most relevant authors.

Figure 6 represents Lotka's law applied to author's productivity [32].

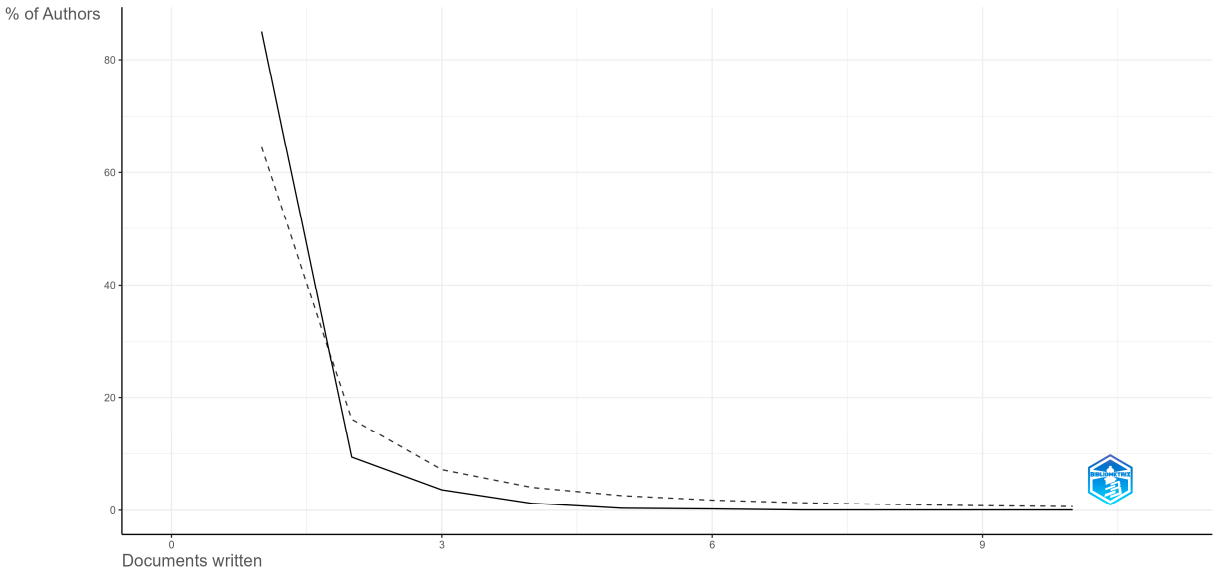

Figure 6. Author productivity through Lotka's law.

This law states that the number of authors that publish a number of articles, n, is inversely related to n^2. Figure 6 reports that 85% of the sample, meaning 1459 authors, have written a single article, while 9.4% of them (161 authors) have been involved in writing two articles, and 3.6% of the authors have written three articles. These results indicate that there are more authors with a low level of productivity than those with a high level of

productivity. Therefore, this law helps us understand how to better allocate the resources in order to support the most active researchers.

Figure 7 displays the evolution of authors' productivity from 2020 to 2023. For a better understanding of the figure, we first need to define a few elements. The circles represent data about the number of articles published and the average number of citations for each author. A higher intensity of the color stands for a higher number of articles published by the author, and a larger diameter indicates a higher number of citations. Thus, according to these two parameters, we can conclude that Javed A has the highest average number of citations (38.5) and also the highest average number of articles published (7). On the other hand, there are also some authors (e.g., Zhao, Y.) with a large number of published articles but with a quite reduced average number of citations (11.5). The red line displays the period of activity in the field of each author. This shows that most of the authors published in the field for at least two years, except for Zhao Y who published articles only in 2023.

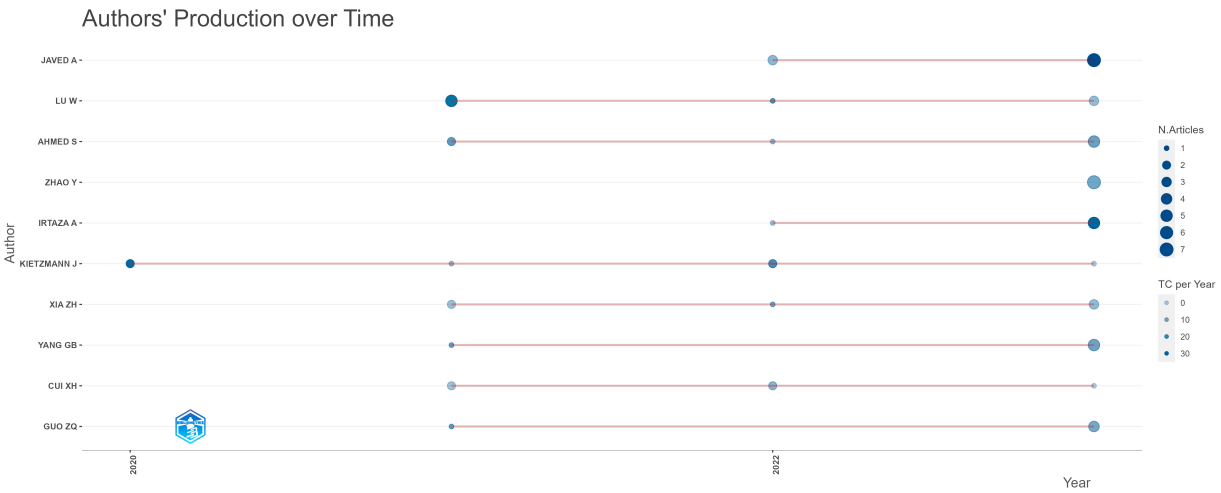

Figure 7. Top 10 authors' production over time.

Figure 8 highlights the most impactful authors in the deepfake literature according to the H-Index. If we compare these results with the ones from Figure 5, we notice that although some authors like Lu W or Xia ZH have many published articles, they are not so well cited as other authors with a lower number of published articles.

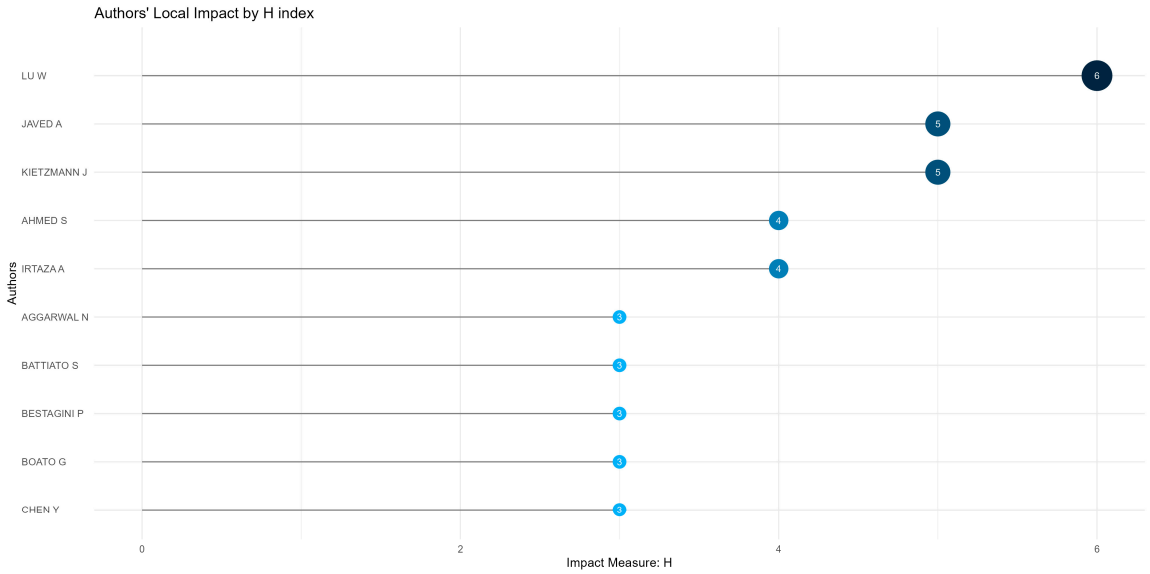

Figure 8. Top 10 authors' impact based on the H-index.

3.4. Country and Affiliation Analysis

Countries and journals are a major part of a bibliometric analysis, providing relevant information about the importance of the topic in specific regions, specifically in terms of the most interested universities, and how the process of publication evolved for each journal.

The most relevant affiliations are presented in Figure 9. We can observe that most of the universities from this list are from China and Egypt, with a total of 8 out of 10 affiliations. The top is predominantly led by Chinese affiliates, the first being "Chinese Academy of Sciences" with 24 articles, "Institute of information engineering, CAS" with 10 articles, and "Nanyang Technological University" with 10 articles. Egypt is represented among the top affiliations by "Egyptian knowledge bank (EKB) "with 22 articles and "Al Azhar University" with 11 articles.

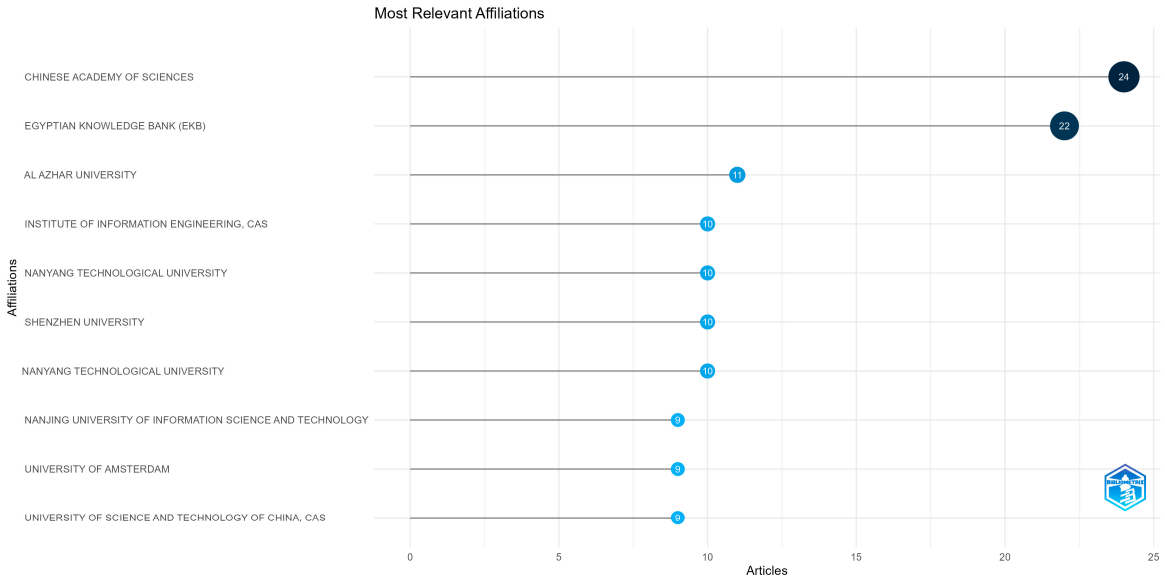

Figure 9. Top 10 most relevant affiliations.

Only one European university is included in our list, the "University of Amsterdam" from the Netherlands. Despite the fact that there are only three countries present in the list, they all come from separate development regions, thus suggesting that the deepfake issue is a global one. A collaboration of these universities would lead to a sharp increase in the development of the research field of deepfake detection.

In line with the previous graph, the countries with the most cited papers are the United States of America (USA) and China (shown on Figure 10). Although the number of citations of articles from China is almost 120% higher than that from Italy, the average number of citations per article in Italy is 19 compared to only 7.5 in China, resulting in a greater influence on the research literature. Some other countries from Europe, Asia, and Africa have a large influence in the domain, such as the United Arab Emirates with 66 average citations, Canada with 29.1 citations, and Spain with 23.2 citations.

Figure 11 highlights the geographical distribution of the scientific contribution. It shall be noted that in Figure 11 we have used various shades of blue to depict each country's contribution in terms of number of published papers, with darker blue colors symbolizing higher contribution and lighter blue colors symbolizing lower contribution. The remainder of the countries—colored in grey—have provided no contribution to the field in terms of number of published papers. We can notice that this subject generates interest all over the world, with the first five countries by the number of publishers being from four different continents. Most of the publishers come from China (390), closely followed by USA (204). Some other countries that are actively contributing to the literature are India (120), U.K. (82), and Australia (66).

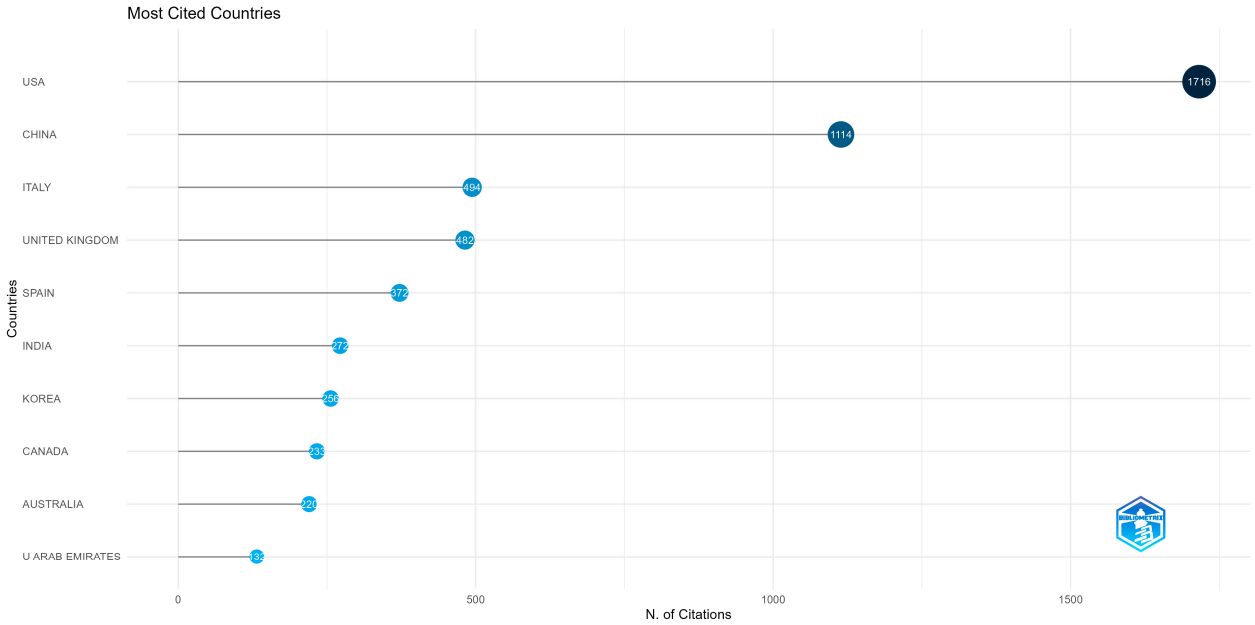

Figure 10. Top 10 most cited countries.

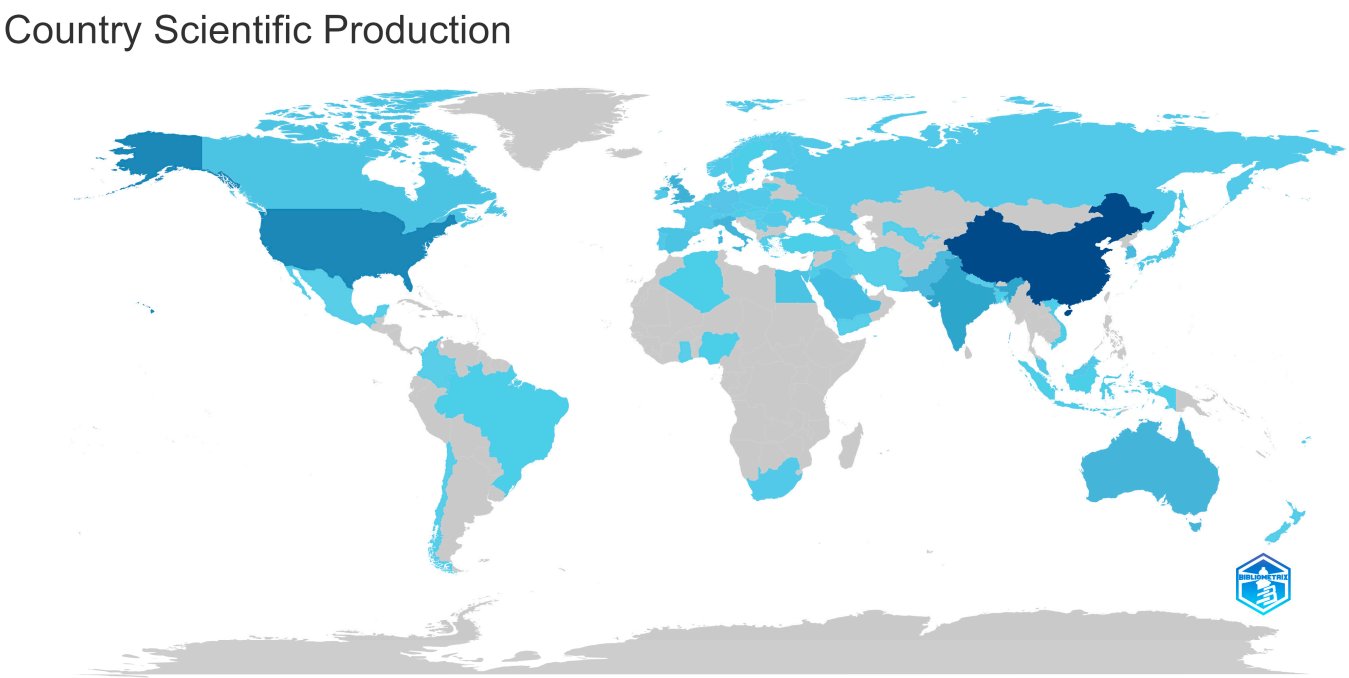

Figure 11. Country scientific production.

Figure 12 displays two important indicators for measuring the level of scientific international collaboration. SCP (Single Country Publication) counts how many articles were published by authors affiliated to universities from one single country, in opposition to MCP (Multiple Country Publication), which counts how many articles were published by authors affiliated to universities from more than one country.

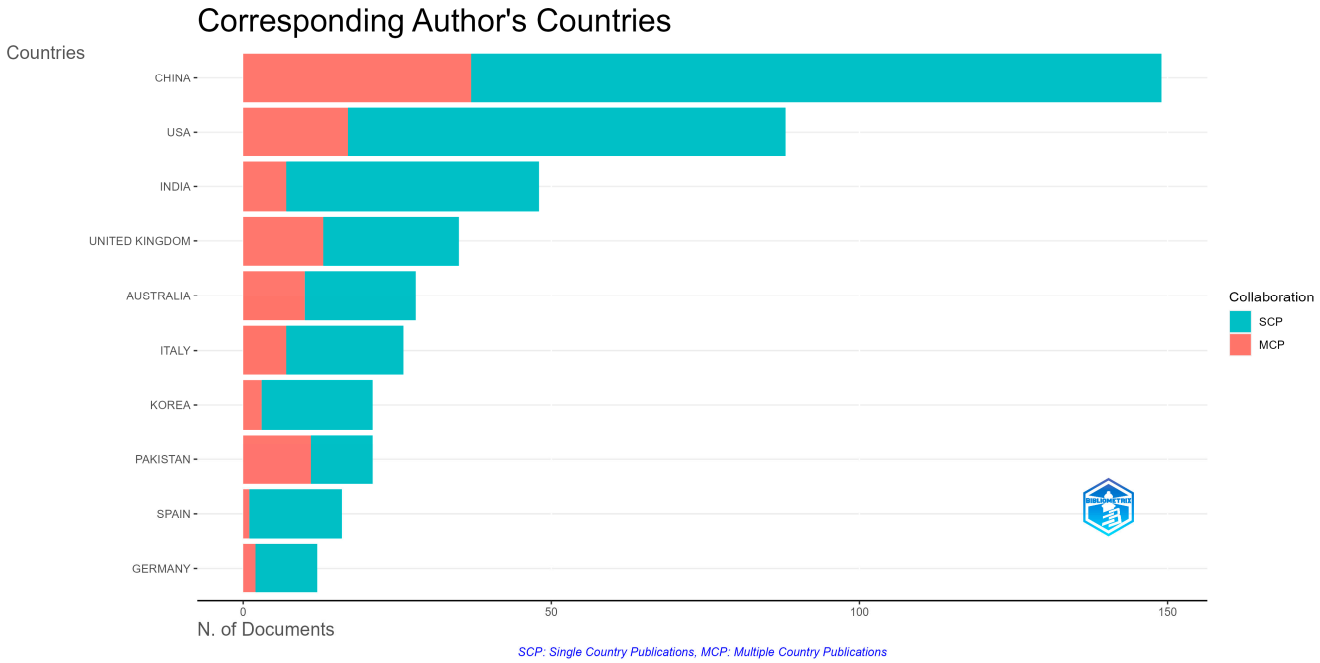

Figure 12. Top 10 most important corresponding author's countries.

We can notice that most of the countries have a higher SCP indicator, suggesting that there is still a limited propensity for inter-country collaboration. Interpreting the SCP and MCP indicators in absolute values, we observe that the most prolific countries without collaborations are China with 112 articles, the USA with 71, and India with 41, and the countries with the highest MCP are China with 37, the USA with 17, and U.K with 13. Moreover, analyzing the relative values we notice that the countries with the highest average MCP are countries like Pakistan, Egypt, Canada, and Malaysia, which have over 50% of their publications written in collaboration. However, there are also some countries with all articles published without any collaboration, such as Russia and Estonia.

Figure 13 displays all the international collaborations established by authors from each country. The aim of international collaboration is to diversify research perspectives, facilitating the development of solutions. As in Figure 11, in Figure 13 the colors of the countries represent the contribution of each country (expressed in number of published papers) to the analyzed field—ranging from dark blue (higher contribution) to light blue (lower contribution), or grey color (no contribution), while the lines in red symbolize the intensity of the collaboration among the authors from the depicted countries. As we could expect by taking into consideration Figure 12, the most international collaborations are established between authors from USA and China. Some other countries with a high MCP indicator are Australia, UK, and Canada. The countries that collaborated the most with each other are USA and UK with 12 collaborations, China and USA with 11 collaborations, China and Singapore with 10 collaborations, and Pakistan and Saudi Arabia, also with 10 collaborations.

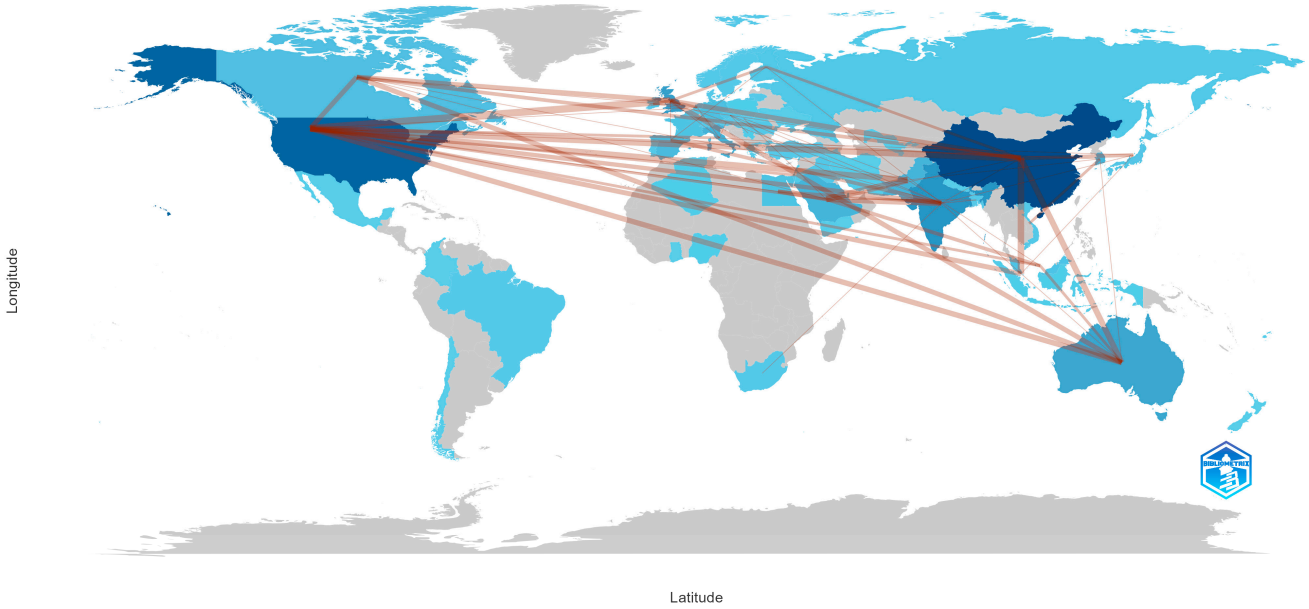

Figure 13. Country collaboration map.

3.5. Word Analysis

To be able to understand the researchers' purposes, the most used words and groups of words were analyzed, in order to extract the main information from titles, abstracts, and keywords.

Table 5 shows the most frequently used KeyWords Plus terms. "networks" is the most frequent word, with 42 appearances, followed by "images" with a frequency of 35. The third most common keyword plus is "fake news" with 20 occurrences, and "face manipulation" has 12 appearances. The fifth and sixth each have 11 appearances and are similar: "disinformation" and "information", two of the main topics in the deepfake domain. The rest of the words have a frequency of 10: "artificial-intelligence", "location", "media", and "recognition". All KeyWords Plus terms described in Table 5 describe the areas of deepfakes, which are the effects such as "disinformation", "face manipulation", and "fake news", and the methods of sharing the information: "networks, location, media". A list of synonyms was used, merging into "images" the singular "image" and plural "images" forms.

Table 5. Top 10 most used KeyWords Plus terms.

Keyword Plus Terms	Frequency
Networks	42
Images	35
Fake news	20
Face manipulation	12
Disinformation	11
Information	11
Artificial intelligence	10
Location	10
Media	10
Recognition	10

Table 6 contains the most used author keywords in documents related to the deepfake domain. A list of synonyms was used in order to group words from similar domains or with a similar spelling. "Deepfakes detection" is the most frequently used keyword, with a frequency of 367, and it contains the following words: "deepfake detection", "deepfakes", and "deepfake". The second most used keyword is "deep learning" with 91 appearances,

followed by "Artificial intelligence" with 54 occurrences, both referring to the technologies that facilitate the apparition and evolution of deepfakes. The rest of the words—"face", "fake news", "feature extraction", "disinformation", "Machine learning", "Generative adversarial networks", and "misinformation"—have lower frequency, and all are related to the creation and detection of fake images and videos.

Table 6. Top 10 most used author keywords.

Author Keywords	Value
Deepfakes detection	367
Deep learning	91
Artificial intelligence	54
Faces	35
Fake news	34
Feature extraction	32
Disinformation	30
Machine learning	30
Generative adversarial networks	27
Misinformation	24

Table 7 groups the top 10 most used title bigrams. In first place is "deepfakes detection" with 185 appearances, which contains six bigrams: "deepfakes detection", "deepfake videos", "deep fake", "deepfake detection", "deepfakes videos", and "video detection", all of which refer to the methods of detecting image and video changes. The second most used bigram is "Deep learning" with 26 occurrences, representing the technology that facilitates the creation of deepfakes. The third bigram is "forgery detection", which is another method of deepfake recognition, and has 25 occurrences. The rest of the bigrams have a frequency of 14 or less and their impact is less significant.

Table 7. Top 10 most used title bigrams.

Title Bigrams	Frequency
Deepfakes detection	185
Deep learning	26
Forgery detection	25
Artificial intelligence	14
Neural network	14
Generative adversarial	12
Social media	12
Convolutional neural	11
Fake news	11
Deep fakes	10

Table 8 presents the most relevant abstract bigrams. A list of synonyms is used, similar to the one presented for Table 7. The most used bigram is "deepfakes detection" with a frequency of 501 appearances, followed by "deep learning" with 183 occurrences, representing one of the main technologies that facilitated the creation of deepfakes, together with the third bigram, "Artificial Intelligence" [33], having a frequency of 111. "Social Media", including the platforms where almost all fake news is distributed, has a frequency of 111, and the rest of the bigrams have a lower impact, with a number of occurrences of less than 100.

Table 8. Top 10 most used abstract bigrams.

Abstract Bigrams	Frequency
Deepfakes detection	501
Deep learning	183
Artificial Intelligence	111
Social media	111
Detection methods	89
Convolutional neural	77
Fake news	77
Neural networks	77
Experimental results	74
Machine Learning	73

Table 9 contains the most used title trigrams. A list of synonyms is used in order to group similar trigrams into a generic form. For instance, "convolutional neural networks" contains the following trigrams: "convolutional neural networks", "convolutional neural network", "convolution neural networks", and "convolution neural network"; and "deepfake detection methods" contains "deepfake detection methods"", deepfake video detection", "deepfake detection method", "deepfake image detection", "deepfake detection based", "deepfake detection algorithm", "deepfake videos detection", "deep fake detection", and "detection algorithm based". "Deep learning methods" is formed by the following trigrams: "deep learning models", "deep learning algorithms", "deep learning techniques", "deep learning approach", "deep learning methods", "deep learning model", and "deep learning networks"; "forgery detection methods" contains "image forgery detection, "video forgery detection", "forgery video detection", "video deepfake detection", and "video detection based"; and "generative adversarial networks" is formed by "generative adversarial networks" and "generative adversarial network". The most frequently used trigram is "deepfake detections methods" appearing 44 times, followed by "forgery detection methods" with only 19 appearances. Third is "convolutional neural networks" with 13 occurrences, while "deep learning methods" and "generative adversarial networks" have 11 and 10 appearances. The rest of the trigrams have a frequency of 3 or 2, and their impact is significantly reduced. Trigrams present the various methods of detection that are currently used, or other representative terms for the deepfake domain.

Table 9. Top 10 most used title trigrams.

Title Trigrams	Frequency
Deepfake detection methods	44
Forgery detection methods	19
Convolutional neural networks	13
Deep learning methods	11
Generative adversarial networks	10
detecting compressed deepfake	3
Capsule dual graph	2
Compressed deepfake videos	2
Conditional generative adversarial	2
Deep fake attacks	2

Table 10 presents the most used trigrams in abstracts extracted from the analyzed dataset. A list of synonyms is included in the analysis, the same as that described in Table 9, in order to group the similar trigrams into a common form. "Convolutional neural networks" is the most frequently used trigram, appearing 84 times, followed by "deepfake detection methods" with 80 occurrences, and "deep learning methods" with 63 appearances. Fourth is "generative adversarial networks" with 57 occurrences, and the rest of the trigrams have a frequency less than or equal to 30, also having a smaller impact.

All trigrams present the deepfake domain from different perspectives, i.e., describing the possible solutions of detecting fake images or videos, such as "deepfake detection methods", "deep learning methods", "generative adversarial networks", "neural networks cnn", or "deep neural networks", or the elements that facilitate the evolution and distribution of fakes, such as "artificial intelligence ai" and "social media platforms".

Table 10. Top 10 most used abstract trigrams.

Abstract Trigrams	Frequency
Convolutional neural networks	84
Deepfake detection methods	80
Deep learning methods	63
Generative adversarial networks	57
Artificial intelligence ai	30
Adversarial networks gans	23
Forgery detection methods	22
Neural networks cnn	17
Social media platforms	16
Deep neural networks	12

3.6. Mixed Analysis

Mixed analysis focuses on extracting the main information from the five most cited documents in the deepfake domain (these documents each have more than 200 citations) and 10 more papers having a number of citations between 30 and 200, which were randomly extracted. Additionally, an analysis of the connections between countries, authors, keywords, and affiliations is provided.

3.6.1. Five Most Cited Papers' Review and Overview

The best way to understand a domain is to analyze the most globally cited documents, citations being one of the best performance metrics, together with the number of publications for each author [34].

Table 11 contains the five most cited documents, together with some performance metrics, such as total citations, total citations per year, and normalized total citations.

In 2019 in "*California Law Review Journal*", Chesney and Citron [35] published the most cited document globally at the time of this research. With 258 citations, an average of 43 citations per year, and 3.9 normalized total citations, the document discusses the deepfake area, where video and audio of people can be created, making them talk and do things they did not say or do, thanks to the machine learning evolution. Deepfake technology has the advantage of rapidly spreading. The scope of the analysis is to explain the main factors and risks of the technology evolution, and how the deepfakes can be detected, by creating a questionnaire and collecting information from various persons, asking them about the criminal penalties, civil liability, economic sanctions, importance of technological solutions, and many other questions. Researchers at the University of Washington presented a deepfake using a neural network tool, altering videos and making the speakers say things that were different from what they initially said, and presenting a video of Barack Obama where he appeared discussing things that he never talked about. The evolution of machine learning into neural network methods increased, in a significant manner, the appearance of images and audio. Generative adversarial networks, also known as GANs, combine two neural networks that work simultaneously. The first neural network draws on the dataset by producing a mimicked sample, and the second neural network evaluates the success of the previous network. The audio and video deepfakes will have a much greater impact in the future, and social media platforms will facilitate their distribution and increase their effects. At this moment, the legal and policy laws are not optimal, and because of that, the risks are high. Deep learning tools are the technology that created the

deepfake and fake media. Forensic methods are one of the main applications that is trying to detect fake images and videos.

In 2020, Verdoliva published a paper in "*IEEE Journal of Selected Topics in Signal Processing*", which has 245 citations, an average of 49 citations per year, and normalized total citations of 6.27. The paper explains the evolution of techniques that manipulate content and are able to edit information in a very realistic way. Technology has a lot of benefits, offering new features in advertising, arts, film production, or video games, but it is also extremely dangerous for society. The software applications are free and very simple to use, offering the possibility to almost everyone to create fake images and videos. The risks are a serious threat, with deepfakes being able to manipulate public opinion during elections, and discredit or blackmail individuals, or facilitate the fraud. It is necessary to find, as quickly as possible, tools that automatically detect false multimedia, reducing the spread of false information. The research focuses on presenting various methods of verification of visual media integrity and detecting manipulated videos and images. A lot of money is invested in major research initiatives in order to find the best methods of deepfake detection, but it is difficult to forecast if the efforts will ensure strong information security in the future. At this moment there are numerous methods that combat deepfakes, but these methods are not strong enough to identify all of them.

In 2020 in "*Information Fusion*", Tolosana et al. [36] published one of the most cited documents, having 206 citations, an average of 41.20 citations per year, and 2.57 normalized total citations. The availability of free databases, together with the evolution of deep learning algorithms, especially generative adversarial networks, created realistic fake images and videos, having a serious impact on society. The focus of the research is to review the techniques of face image manipulation, including also deepfake methods, analyzing four different methods: entire face synthesis, expression swapping, attribute manipulation, and identity swapping. For each one, the available public databases, the manipulation algorithms, and the metrics to evaluate the detection rate are presented. Entire face synthesis creates a non-existent face image. Using generative adversarial networks, great results are achieved, and high-quality facial images, very similar to the real pictures, are provided. Identity swapping is the technique of replacing the face of a person in a video with the face of someone else. This type of deepfake is mostly used in the film industry, and in numerous cases, the usage of this approach is for bad purposes. Attribute manipulation represents face editing by changing some attributes of the face, such as the gender, age, or color of the hair. One of the most well-known applications that is used for attribute manipulation is FaceApp, which is available on mobile, and allows users to add various cosmetics, makeup, glasses, or hairstyles to the pictures. Expression swapping edits the facial expression of a person, and the most well-known techniques are Face2Face and NeuralTextures. One of the most common expression-swapping examples is when Mark Zuckerberg appeared in a video and discussed things that he never talked about in reality.

In 2020 in "*Social Media + Society*", Vaccari and Chadwick [37] published a paper having 204 citations, with an average of 40.80 citations per year, and 5.22 normalized total citations. Artificial intelligence offers the possibility of mass creation, using synthetic information, of videos and images that are very close to reality. Political deepfakes are very popular on the internet, leading to disinformation, and having a serious impact on journalism and affecting the quality of democracy. Deepfakes are one of the main methods of disinformation, increasing uncertainty, reducing the trust in news on social media, increasing the cost of obtaining real information, and increasing the level of uncertainty. The research performed an experimental analysis on political deepfakes, and the results explained the trend of people's increasingly reduced trust in news on social media, with people having less interest in cooperating in contexts where the trust level is low. This behavior can lead to less collaborative and responsible social media users, and this will constantly reduce the trust in news on social media. Citizens can also refuse to read news in order to reduce their stress level. It will become more and more difficult to perform

public debates, as the society must be vigilant and able to observe every possibility of manipulation. Another problem is the trend of illiberal policies, which promises to clear the internet of deepfakes.

In 2021 in "*ACM Computing Surveys*", Mirsky and Wenke [5] published a paper having 203 citations, an average of 50.75 citations, and normalized total citations of 13.67. In 2018, the technology of generative deep learning was used for malicious applications and spreading misinformation, and since then deepfakes have evolved significantly. The deepfake is a combination of "deep learning" and "fake", representing fake content created by artificial neural networks. The most common use of deepfakes is the creation of videos and images. Deep learning methods also have some useful and productive applications, such as reanimating historical figures and dubbing foreign films in a realistic manner. At the end of 2017, a deepfake video appeared on Reddit for the first time in which a celebrity appeared in an adult movie; since then, the number of deepfakes has increased exponentially. In 2018, a video with Barack Obama was presented by BuzzFeed, where the former president talked about a subject that in reality was never discussed, raising serious concerns over identity theft. The deepfake is also used to clone voices. In just five seconds, a CEO of a company was scammed out of USD 250,000. Deep learning can also generate realistic human fingerprints, offering the possibility of unlocking devices. It is important to understand that not all technology is dangerous, and the purpose is not always bad, but the key is to identify the best methods to detect fake news.

Table 11. Five most cited papers.

No.	Paper (First Author, Year, Journal, Reference)	Number of Authors	Region	Total Citations (TC)	Total Citations per Year (TCY)	Normalized TC (NTC)
1	Chesney, Bobby, 2019, *California Law Review* [35]	2	USA	258	43	3.90
2	Verdoliva, Luisa, 2020, *IEEE Journal of Selected Topics in Signal Processing* [38]	1	Italy	245	49	6.27
3	Tolosana, Ruben, 2020, *Information Fusion* [36]	5	Spain	206	41.20	2.57
4	Vaccari, Cristian, 2020, *Social Media + Society* [39]	2	UK	204	40.80	5.22
5	Mirsky, Yisroel, 2021, *ACM Computing Surveys* [5]	2	USA	203	50.75	13.67

3.6.2. Review and Overview of 10 Randomly Selected Papers

In order to reflect the research areas of the other papers included in the database, we randomly selected 10 papers that have a number of citations between 30 and 200. We decided to select papers with at least 30 citations as these papers represent the top 8% of the papers included in the database with respect to the number of citations, thus attracting relatively high interest from the research community. Furthermore, as there are 47 papers with more than 30 citations in the database, it would have been difficult to provide a review of all these papers.

Table 12 contains the information regarding the 10 selected papers. Only two countries are presented, China and USA, showing their interest in deepfakes, which are a great risk for their economics and politics. The most cited document has 69 citations, which is expected for a domain that is still new to the academic community; this document also has small values for total citations per year and normalized total citations. Similar to the analysis of Rana et al. [12], the most used database was FaceForesnics++, and CNN models are the most commonly used

Table 12. Overview of the 10 selected documents.

No.	Paper (First Author, Year, Journal, Reference)	Number of Authors	Region	Total Citations (TC)	Total Citations per Year (TCY)	Normalized TC (NTC)
1	Chesney, Robert, 2019, *Foreign Affairs* [39]	2	USA	69	11.5	1
2	Bimber, Bruce, 2020, *New Media & Society* [40]	2	USA	66	13.2	3.51
3	Yang, Jiachen, 2021, *IEEE Transactions on Information Forensics and Security* [41]	5	China	60	15	3.51
4	Fletcher, John, 2018, *Theater* [42]	1	USA	54	7.71	1
5	Guo, Zhiging, 2021, *Computer Vision and Image Understanding* [43]	4	China	52	13	3.12
6	Yang, Jianchen, 2022, *IEEE Transactions on Circuits and Systems for Video Technology* [44]	6	China	48	16	6.41
7	Rini, Regina, 2020, *Philosophers' Imprint* [45]	1	USA	44	8.8	2.34
8	Yang, Jianchen, 2021, *Future Generation Computer Systems* [46]	5	China	34	8.5	2.04
9	Yu, Peipeng, 2022, *IEEE Transactions on Information Forensics and Security* [47]	5	China	33	11	4.41
10	Johnson, James, 2022, *Journal of Strategic Studies* [48]	1	USA	32	10.67	4.28

Chesney and Citron [39] presented, in a philosophical manner, what a photo can express, but nowadays, video and audio recordings are more relevant. Audio and video recordings offer the possibility for people to be a witness to an event, even if they were not physically at that event. Thanks to social media platforms, now it is much easier to publish, share, or capture a video or a photo. However, people have to decide if they trust every single person who has access to their posts; otherwise, they could face serious problems, which could affect their social life. Chesney and Citron provided some examples of what deepfakes can produce: a private conversation between an Israeli prime minister and a colleague, in which they are discussing an imminent attack on Tehran, or an audio recording where an Iranian official describes a plan to attack a region in Iraq. All these could be faked using various deep learning tools, making them almost impossible to distinguish from the real ones. One of the most used algorithms is the generative adversarial network, or GAN, which has a pair of algorithms. The first creates content based on the source data, and the second tries to find the artificial content, picking out the fake content. Deepfakes also have numerous benefits, changing the audio or video of historical figures, or even restoring speech to people who have lost their voice. Unfortunately, deepfakes are commonly used for darker purposes, such as putting people's faces into dangerous situations, without their consent, or trying to blackmail, intimidate, or sabotage them.

Bimber and de Zuniga [40] believe democracy could be affected by fake information, with social media being one of the major factors in the sharing of deepfakes. The Xinhua news agency created, using AI, synthetic-video news, which looked realistic, using Barack Obama's speech. In 2018, the Wall Street Journal was the first news company to announce the risks of a fake video, creating a dedicated team that investigates if the photos or videos that could be presented by the company are fake or not. The solutions to counter deepfakes could be mainly applied by the social media companies, by introducing an authentication of users when an account is created and requiring the individuals to publish their identity. Social media creates public opinion about what people are thinking and discussing, what they like, and what they want to do, and numerous methods of exploiting the vulnerabilities of individuals exist. Anonymity and pseudonymity in social media are the main reasons for creating deepfakes, by facilitating deception. Democracies are weaker because of deepfakes since messages are manipulated. The British Conservative Party, in 2018, used social media armies to promote their messages and to manipulate.

Yang et al. [41] describes convolutional neural network discriminators and discusses the multi-scale texture difference. This is one of the key elements in face forgery detection, and significantly facilitates the process of identifying a fake photo and a real one with high accuracy. Because of technological advances, for humans it is impossible to detect fake and real photos and videos. A new multi-scale texture difference model has been created, known as MTD-Net, which was used for face forgery detection. The approach of the model is to leverage central difference convolution (CDC) and atrous spatial pyramid pooling (ASPP). CDC merges pixel intensity information and gradient information, offering

a stationary description of texture difference information. The analysis was performed on multiple databases, i.e., Faceforensics++, Deeper Forensics-1.0, Celeb-DF, and DFDC. The results of the experiment showed great potential, having a higher accuracy compared with existing methods, showing that MTD-Net is more robust for image distortion.

Fletcher [42] presented the historical evolution of deepfakes. The ML tool that creates deepfakes appeared in late 2017, and it took a few months for governments and social media companies to start understanding the possible impact of deepfakes. Face swapping effects can be easily achieved using AI, thanks to various applications that appeared starting in January 2018. Deep learning is a specialized form of ML, and their algorithms operate as neural nets, having a similar approach to biological neurons. Deep neural nets could be trained to indicate a correct medical diagnosis, create more efficient medicines, or completely change urban development. FakeApp is a desktop application that makes the face-swapping process extremely easy for videos, such that even users without coding knowledge can use it. Users just have to upload two videos, and in a few hours the face change process is completed. The program called Lyrebird uses deep learning techniques to create fake speeches using famous voices, such as those of Donald Trump or Barack Obama. Numerous research institutions make their AI technology available and publish open-source software libraries. One of the best-known examples is Google's TensorFlow, which offers the possibility of innovation to any programmer. The evolution of ML algorithms is visible in the daily activities of individuals, mainly on social media platforms, where the users receive suggested ads or videos to watch.

Guo et al. [43] described face image manipulation (FIM) techniques such as Face2Face and Deepfake, which helped spread fake images on the internet, creating serious problems and concerns for society. Researchers have progressed significantly with fake face detector algorithms, but there are numerous elements to be improved, since FIM is more and more complex. CNNs learn content representation of images but have some limitations, learning only parts of manipulation traces. An adaptive manipulation trace extraction network (AMTEN) represents the pre-processing of suppressed image content, showing the manipulation traces, focusing on convolutional layers, and predicting manipulation traces of pictures. AMTENnet, a fake face detector, was built to facilitate the integration of AMTEN with a CNN. The CNN models are divided into three different categories: stacking standard CNN modules for a certain fake image, using hand-crafted residual features by different models, and improving the form of the convolution layer, forcing the CNN to learn features from tampering traces. The results of the analysis showed good results for AMTEN, while AMTENnet had an average accuracy of 98.52%, outperforming state-of-the-art works. When the dataset contains face images with unknown post-processing operations, the algorithm had an average accuracy of 95.17%. Even if the forensics cases were simulated using post-processing methods, there are significant differences from real cases, with AI-generated images sent all over social media platforms.

Yang et al. [44] pointed out the risks of fake videos presented on the internet, which affect individual's activities and relationships, and also pollute the web environment and trigger public opinion. In some cases, they can become a national security threat. The majority of the existing algorithms are based on convolutional neural networks, which learn the feature differences between real and fake frames. The purpose of the analysis is to create a multi-scale self-texture attention generative network (MSTA-Net) that is able to track the potential texture trace in images and to eliminate the interference of deepfake post-processing elements. Initially, a generator was created, which performs encoding–decoding and disassembling, in order to visualize the traces, and finally merging generated trace images with the original ones as input into a classifier with Resnet. The second part of the tool is the self-texture attention mechanism (STA), which skips the existing connection between the encoder and decoder. The final step is to propose a loss function known as Prob-tuple loss restricted, which finds the probability of amending the generation of forgery traces. To check the performance of the model, several experiments were performed,

showing that the algorithm performs well on FaceForensics++, Celeb-DF, Deeperforensics, and DFDC databases, having a high level of feasibility.

Rini [45] provided an explanation of deepfakes and their effects, and introduced the idea of an epistemic backstop. Video and image recordings are an epistemic backstop, having great availability and regulating testimonial practices. Deepfakes could affect the democracy of information sharing and debates when key people could be deepfaked. Unfortunately, deepfakes spread outside of the journalistic domain, entering computer science, legal, and even academic domains. The recordings can facilitate the process of correcting errors in past testimony and regulating ongoing practices. In the summer of 2019, Deeptrace, a digital security company, tracked over 15,000 deepfakes on the web, almost double compared with an earlier period of that year, of which 96% were videos. The most dangerous applications of deepfakes are in politics. A few journalistic companies published fake political news, where, for instance, Barack Obama called Donald Trump in an offensive manner. In May 2018, the Flemish Socialist Party published a deepfake video where Donald Trump insisted on the withdrawal of Belgium from environmental treaties. Later, the Party tried to explain the purpose of the video, which was only a method of increasing interest in the subject, and was not to fool somebody. In January 2018, John Winseman described what could have been the first deepfake attempt with a political purpose, discussing gay rights in Russia.

Yang et al. [46] believes fake detection is an acute problem, and discovered the texture variations between real and fake images. Due to technology, our life has improved significantly, but there are also threats, one of which is cybercrime. Deepfakes fabricate fake events which are shared on the internet, causing problems with a big impact and chaos. Using forensic methods, an algorithm for deepfake detection has been discovered, which compares real and fake images in image saliency, extracting the face texture differences. Resnet18, a classification network, was trained to identify the differences between images and tested to find real and fake face images, and the accuracy was also evaluated. The process is divided into two parts: the first focuses on full image training, while the second has only face images, which are also added into the training dataset, and, after that, are added to the Resnest18 algorithm. The evaluation was performed on 14,000 images and 140 videos, taking into consideration 2800 real images and 11,200 fake images. The Xception trained model has an accuracy of approximately 0.52, while that of Mesonet is 0.72 and that of Cozzolino is only 0.34. The Guided model, which was created by the researchers, has a performance of 0.8.

Yu et al. [47] tried to improve face video forgery detection, in order to improve the generalization. There are already numerous face forgery algorithms that provide similarities in forgery trace videos. The purpose of the research is to understand a completed generalization in the detection of unknown forgery methods. Initially, a model called Specific Forgery Feature Extractors (SFFExtractors) was trained separately for each of the given forgery algorithms. Using the U-net structure, with various possible losses, the SFFExtractors were tested to detect corresponding forgery methods. In the next step, another algorithm, Common Forgery Feature Extractor (CFFExtractor), was trained, taking into consideration the results of SFFExtractors, and the similarities between forgery methods were explored. The results obtained by models on FaceForensic++ showed a great success of SFFExtractors in face forgery detection. CFFExtractor was also run on multiple databases, and the results proved that commonality learning is a good approach for improving generalization and developing an effective strategy.

Johnson [48] explores the impact of AI in strategic decision-making processes and its stability, presenting the risks and the adaptability of military forces to the latest technology. An advantage of AI is that it replaces humans, who could be affected by empathy, creativity, intuition, or other external events, in making decisions. An AI generative adversarial network (GAN) could create deepfakes, which could create a crisis between two or more nuclear countries by just creating an image or a video of a military leader with fake orders, creating tensions and confusion between states. Deepfakes are already a tool used for

disinformation and deception. It is very difficult, during a crisis, to understand the purpose of the attacker. China's fear of a US attack has made the Chinese to prioritize false-negative scenarios instead of false-positive scenarios. False negative refers to misidentifying a nuclear weapon as non-nuclear, and false positive is misidentifying a non-nuclear weapon as a nuclear one. Technology not just provides deepfakes, but also bots and fake news, which are used to exploit human psychology by creating false narratives, intensifying false alarms, and trying to destabilize.

Table 13 presents the 10 selected documents with a number of citations greater than 30, together with the title of the papers, data that were utilized in the research, and the scope of the documents.

Table 13. Brief summary of the content of 10 selected documents.

No.	Paper (First Author, Year, Journal, Reference)	Title	Data	Purpose
1	Chesney, Robert, 2019, *Foreign Affairs* [39]	Deepfakes and the New Disinformation War	Datasets found on the internet, for instance cat pictures, or any other large datasets	To present the benefits and threats of deepfake technology, by providing numerous examples, explaining the GAN algorithm
2	Bimber, Bruce, 2020, *New Media & Society*, [40]	The unedited public sphere	No datasets have been used	To make people aware of the potential of deepfakes, which could affect the health of democracy, presenting examples of deepfakes and how social media is helping to spread the information
3	Yang, Jiachen, 2021, *IEEE Transactions on Information Forensics and Security* [41]	MTD-Net: Learning to Detect Deepfakes Images by Multi-Scale Texture Difference	Faceforensics++, DeeperForensics-1.0, Celeb-DF and DFDC databases have been included in the analysis	To facilitate the process of detecting deepfakes, by creating a new model called Multi-scale Texture Difference (MTD-Net) used for robust face forgery detection
4	Fletcher, John, 2018, *Theater* [42]	Deepfakes, Artificial Intelligence, and Some Kind of Dystopia: The New Faces of Online Post-Fact Performance	No data were used	To present a historical evolution of the domain and when it was used for the first time; it is one of the most used applications for face swapping; and how the technological evolution helped to create increasingly complex deepfakes
5	Guo, Zhiging, 2021, *Computer Vision and Image Understanding* [43]	Fake face detection via adaptive manipulation traces extraction network	CelebA, CelebA-HQ datasets have been included	To create a detection model for deepfakes called adaptive manipulation trace extraction network (AMTEN), which has been tested and showed great results
6	Yang, Jianchen, 2022, *IEEE Transactions on Circuits and Systems for Video Technology* [44]	MSTA-Net: Forgery Detection by Generating Manipulation Trace Based on Multi-Scale Self-Texture Attention	FaceForensic++, DFDC, and CelebDF and Deeperforesnics datasets have been utilized	To design a deepfake model detector called multi-scale self-texture attention generative network (MSTA-Net) since the actual algorithms are not strong enough to detect the latest deepfakes, which, thanks to the actual technology, became more and more complex
7	Rini, Regina, 2020, *Philosophers' Imprint*	Deepfakes and the Epistemic Backstop	No dataset has been used	To explain the risks of deepfakes, and the epistemic backstops, providing various real and unreal examples, with various impacts
8	Yang, Jianchen, 2021, *Future Generation Computer Systems* [46]	Detecting fake images by identifying potential texture difference	FaceForensic++, Face2Face and FaceSwap datasets have been included in the analysis;	To create a new algorithm for deepfake images, based on a new approach, based on the Resnet18 algorithm. The results of the model are compared with traditional models
9	Yu, Peipeng, 2022, *IEEE Transactions on Information Forensics and Security* [47]	Improving Generalization by Commonality Learning in Face Forgery Detection	FaceForensic++, DFDC, and CelebDF have been used to train and test methods	To generalize face video forgery detection, using Specific Forgery Feature Extractors and Common Forgery Feature Extractor, which showed great performances
10	Johnson, James, 2022, *Journal of Strategic Studies* [48]	Delegating strategic decision-making to machines: Dr. Strangelove Redux?	No data were used	To express the trend of decision-making processes, where humans have been replaced by machines, and the impact of deepfakes, fake news, and bots in society

Considering the entire database, a thematic map was generated based on the titles of the papers using bigram word extraction. The results of this approach are visualized in Figure 14.

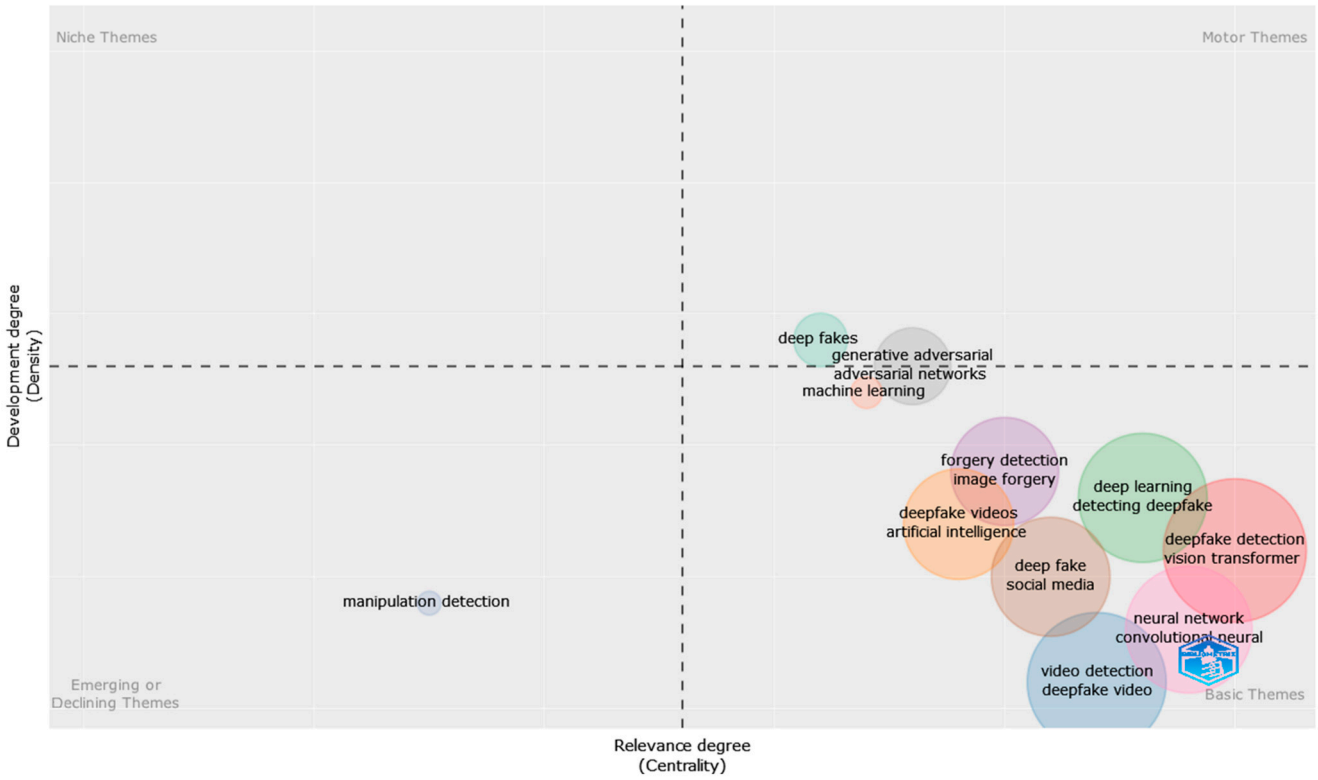

Figure 14. Thematic map.

The thematic map is divided into four quadrants representing niche, motor, emerging or declining, and basic themes. Within the basic themes, several key areas can be identified. Some focus on various types of deepfake detection, such as video deepfake detection (e.g., "video detection", "deepfake video/videos" bigrams) and image forgery detection (e.g., "forgery detection", "image forgery" bigrams). Others discuss the methods used to address deepfakes, including machine learning, deep learning, artificial intelligence, neural networks, and convolutional neural networks (e.g., "machine learning", "deep learning", "artificial intelligence", "neural network", "convolutional neural" bigrams).

Motor themes prominently feature generative adversarial networks, identified through bigrams such as "generative adversarial" and "adversarial networks", positioned at the borderline between motor and basic themes. Additionally, manipulation detection is highlighted as an emerging theme.

3.6.3. Three-Field Plots

The multifaceted relationships are described in Figures 15 and 16, showing the connections between countries, authors, and keywords, and respectively between affiliations, authors, and keywords.

As shown in Figure 15, it can be observed that the contribution of China to the research area dedicated to deepfakes represents a large number of the overall contributions made in the scientific literature.

Additionally, a series of keywords is listed, most of which are from the deepfake field (e.g., "deepfake", "deepfakes", "deepfake detection"), and the domains in which the deepfake can be encountered (e.g., "video"), along with the means to detect them (e.g., "machine learning", "feature extraction", "artificial intelligence").

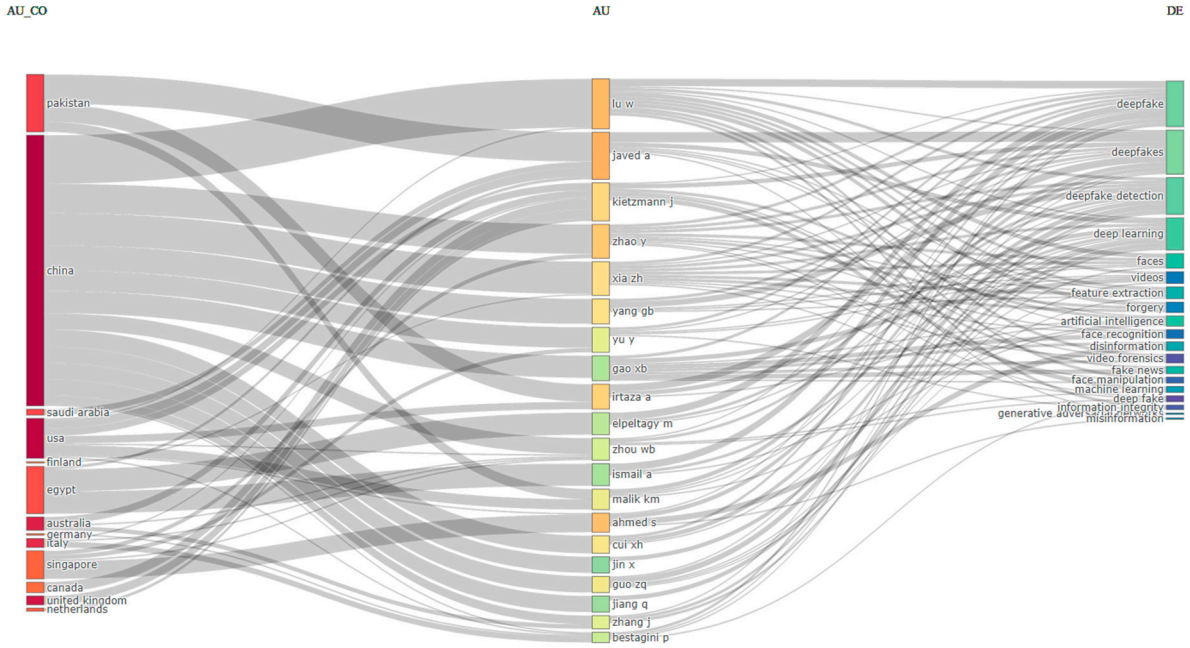

Figure 15. Three-field plot of countries (**left**), authors (**middle**), and journals (**right**).

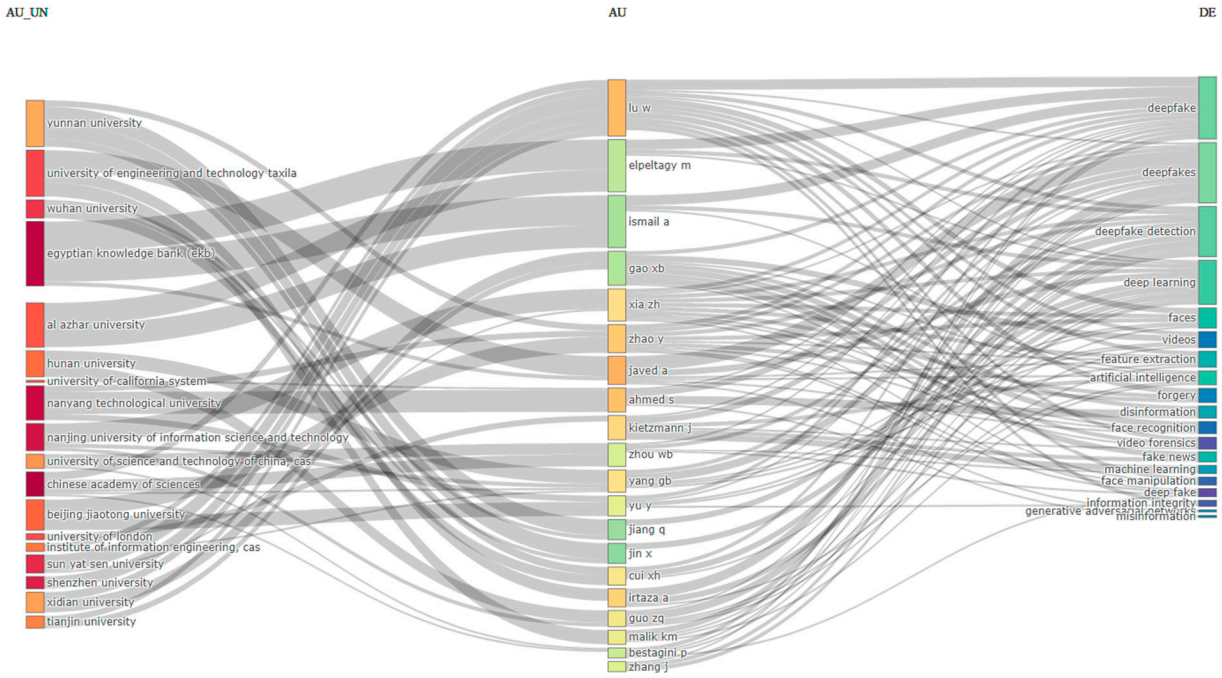

Figure 16. Three-field plot of affiliations (**left**), authors (**middle**), and keywords (**right**).

4. Discussion and Limitations

The bibliometric analysis provided an overview of the deepfake domain. Despite the recent appearance of the domain, it captured the interest of the academic community, who have publishing their ideas in dedicated journals such as "*IEEE Access*", "*Applied Sciences-Basel*", and "*IEEE Transactions on Information Forensics and Security*". There are also journals that recently started to accept documents related to deepfakes, such as "*Multimedia tools and applications*" or "*IEEE Transactions on Circuits and Systems for Video Technology*".

Taking into consideration the number of publications, a series of universities from China are holding the first positions in terms of contributors, namely, Chinese Academy of Sciences, Institute of Information Engineering, CAS, and Nanyang Technological Univer-

sity. Egypt also has publications on the deepfake domain, mainly published by Egyptian Knowledge Bank (EKB) and Al Azhar University.

The USA is the country with the highest number of citations on the deepfake domain, with 1716 citations, almost 50% more than China, who has 1114 citations. Even though the rest of the countries have fewer publications, the number of average citations is significantly larger than the number of citations that China has. Thus, in Table 11, which presents the five most cited papers in the world, we can observe that countries like the United Arab Emirates ranks in sixth place with 131 citations. Moreover, in Table 11, we can notice countries from two main regions, Europe and North America. Europe is represented by three countries: Spain, Italy, and UK, while North America is represented by the USA.

In terms of content of the papers, it has been observed that most of the globally cited documents focused on presenting the actual risks of deepfakes and provided various solutions for detecting changes in images, voices, or videos. The power of deepfakes is mainly thanks to artificial intelligence and deep learning, which facilitated the appearance of software tools that, in a few hours, can change faces and voices or create fake conversations, generating instability and fear in society, which could lead to political crisis. Chesney and Citron [34] present the case of Emma Gonzalez, an activist and advocate for gun control. She stepped into the spotlight when a fake image and video of her spread across the internet. The fake image portrayed Emma ripping up the Constitution of the USA, instead of ripping up a bullseye target. Due to this fake image, her speeches can no longer have a positive impact, and her image was affected. Fortunately, as the deepfake technology was still in its infancy, the people rapidly realized that they were manipulated. Nevertheless, the incident of Emma Gonzales illustrates the negative impact that deepfakes can have on our lives.

Furthermore, based on the information provided in this paper, it can be observed that different approaches for studying deep forgeries can be considered, mostly taken from computer science and machine learning, but also from other related research areas, such as social sciences, or legal and ethical approaches to the matter. Given the complexity of the issue, in order to address the prevalence of deepfakes, it might be necessary to leverage strengths from multiple disciplines and to foster the collaborations between researchers from various research areas and with various backgrounds. Developing a better comprehension of deep forgeries could have an impact on better dealing with the challenges imposed by the occurrence of this situation.

Nevertheless, it shall be stated that there are also some limitations to the research conducted in the presented papers which might affect the results of this study. The most important constraint is the database from which the dataset was extracted. Even though the Clarivate Analytics Web of Science Core Collection (WoS) is recognized internationally by the academic community, the exclusion of other databases could impact the variety of subjects, reducing the dimension of the dataset. As Liu [49] explained, the keywords used for filtering the WoS database also have some limitations regarding the availability of the information, as WoS gathers data from journals that might not be fully available all the time. This situation might occur due to two main factors: first, for the papers published years ago, some information was not requested by the journals from the authors, and therefore, is not available in some cases; and second, for all the papers in the database, some information may not have been fully extracted from the journals when included in the database. As the papers in our dataset are recently published papers—due to the novelty of the field—the first situation does not apply in our case. As for the second situation, it was observed that in the dataset we had 51.71% of the KeyWords Plus terms missing (302 KeyWords Plus terms), 12.33% of the keywords missing (72 keywords), 5.82% of DOIs missing (34), 2.4% of the abstracts missing (14), three cited references missing (1.03%), and 0.68% of the corresponding authors and affiliations (four each) missing. As a result, it should be stated that these missing elements might have an impact on the obtained results.

Other limitations that could have altered the results of the analysis were related to the language of the papers, because only documents published in English were kept. Dis-

cussing the language of the documents included in the popular databases, Vera-Baceta [50] presented a comparison between WoS and Scopus databases, and demonstrated that there is no discrepancy between WoS and Scopus databases for English papers, and that a significant difference exists for non-English documents, with Scopus having a greater coverage than WoS for non-English articles.

Furthermore, we retained in the analysis only the documents that were marked as "articles" by the Clarivate Analytics Web of Science Core Collection. In connection with the type of document, it shall be mentioned that the database includes in the "article" type of documents all the relevant research; thus, it does not necessarily exclude the conference papers [51].

5. Conclusions

In this paper, we tried to assess the level of interest and attention exhibited towards deepfake technologies in recent years. Even though the first step in researching the artificial intelligence field was made in the 1950s, only after the recent developments in hardware infrastructure could a steep increase in popularity for this subject be noticed. Currently, the models using artificial intelligence are so advanced that it has become difficult for us to distinguish between what is real and what is artificially generated. Unfortunately, artificial intelligence was also used for improper or even harmful purposes, especially by falsifying information with deepfakes. These techniques are designed to manipulate the masses of people and influence their decisions.

Based on the research, the following can be highlighted:

- The evolution of the papers included in the dataset showed an upward trend in the analyzed period, with a peak in 2023, when 244 papers were included in the WoS database, an increase of 45.23% compared to the previous year.
- The most prominent authors based on the number of papers published in the area of deepfakes are Javed A, Lu W, Ahmed S, Zhao Y, Irtaza A, Kietzmann J, Xia ZH, Yang GB, Cui XH, and Guo ZQ, with papers numbering between 5 and 10. Furthermore, considering the yearly number of publications of the most prominent authors, it can be noticed that most of them contributed in 2023.
- The most notable journals are "*IEEE Access*", "*Multimedia Tools and Applications*", "*IEEE Transactions on information forensics and security*", "*IEEE Transactions on Circuits and Systems for Video Technology*", and "*IEEE Signal Processing Letters*". By analyzing the content of the papers published in these sources and the scope of the journals, it can be observed that most of them are in the areas related to multimedia and information security. Moreover, the top-placed journal based on the number of published papers, namely "*IEEE Access*", provides a fast peer review and publication process that might contribute to its popularity among the researchers in this area, which is characterized by a fast evolution in terms of usage and creation of new detection approaches.
- The countries that have collaborated the most in developing scientific articles are USA and UK, followed by China and USA, China and Singapore, and Pakistan and Saudi Arabia.
- The most used KeyWords Plus terms are related to social media and networks ("networks", "media"), the area of deepfake usage ("images"), and the methods used to address this phenomenon ("artificial intelligence"). In terms of author keywords, it can be observed that the most frequent words are related to deepfake detection, followed by the method and/or techniques to address them ("deep learning", "artificial intelligence", feature extraction", "machine learning"), as well as the type of information spread ("fake news", "disinformation", "misinformation").
- The universities who have published the most articles in the deepfake domain in a series of affiliations have been identified from various parts of the world. For example, in China "Chinese Academy of Sciences", "Institute of information engineering, CAS", and "Nanyang Technological University" were identified; Egypt is represented by

"Egyptian knowledge bank (EKB) " and "Al Azhar University"; and Netherlands by "University of Amsterdam".

As demonstrated in this paper, there is significant global interest and extensive research on the topic of deepfakes generated by artificial intelligence. Our analysis underscores that deepfakes are a worldwide concern, drawing attention from various research communities. Furthermore, through the use of the thematic map analysis, key areas, such as deepfake detection methods, including video and image forgery detection, generative adversarial networks, manipulation detection, and techniques involving machine learning, deep learning, and artificial intelligence, are identified as either basic, motor, or emerging themes.

The results obtained through this research could be of interest for the readers in the field of information as it underscores the importance of multidisciplinary research and collaboration in addressing the highly changing challenges posed by the deepfake field nowadays and in the future. Given the expected continuous evolution of the deepfake field, it is expected that the implications for connected fields, such as information systems, digital communications, and trust in social environments, will continually grow. This will require the interested parties to make continuous efforts to stay informed and engaged with the latest developments in the field.

More specifically, considering the elements highlighted through the thematic map analysis, the present study illustrates the current state in the deepfake research, but also provides a future perspective on the themes that are likely to shape future research and technological developments, which can be of interest for the readers and researchers in information science. As a result of the thematic map analysis, it can be observed that the basic themes identified—namely deepfake detection (both video and image forgery) and the methods employed to counter these manipulations (e.g., machine learning, deep learning, and convolutional neural networks)—serve as a basis for the research papers in the field, highlighting the efforts made by researchers worldwide to preserve the information integrity when dealing with the advancements in deepfake technology. Furthermore, the presence of generative adversarial networks (GANs) at the verge of the motor and basic themes suggests their dual role in the research associated with the deepfake domain—the core component of deepfake creation [52]—but also as a possible central point for developing countermeasures in the face of the deepfake avalanche [53].

Considering all the above-mentioned elements, we contend that addressing the adverse effects of deepfakes necessitates vital collaboration among global research centers. Thus, increasing interdisciplinary collaborations could represent a starting point in achieving this objective. In terms of developing collaborations, for instance, technology experts could develop advanced detection systems, legislators could establish comprehensive regulatory frameworks, psychologists could examine the societal and individual impacts, and educators could enhance awareness and promote digital literacy to mitigate the recognition of deepfakes. This collaborative approach aims to leverage diverse expertise to effectively tackle the challenges posed by deepfake technology.

Future analyses may include a wider range of applications, due to the continuous growth in deepfake technologies. To conclude, we believe that understanding how these technologies will evolve and devising methods to protect us against their harmful effects has become a major requirement, which can only be achieved by sustained research in the artificial intelligence field.

Furthermore, as future research directions, more in-depth analyses can be conducted on various subareas of deepfakes—such as deepfakes in images or videos—in order to better capture the specific interests of the researchers in each area and to gain more knowledge on the particular methods used for deepfake detection and on the means to counteract them.

Author Contributions: Conceptualization, A.D., G.-C.T., L.-A.C. and C.D.; Data curation, A.D., G.-C.T., L.C. and C.D.; Formal analysis, A.D., G.-C.T., L.C. and A.-G.M.; Investigation, A.D., G.-C.T., L.C., A.-G.M. and L.-A.C.; Methodology, A.D., G.-C.T., L.-A.C. and C.D.; Resources, A.D., G.-C.T., L.C. and A.-G.M.; Software, A.D., G.-C.T. and L.-A.C.; Supervision, L.-A.C. and C.D.; Validation, A.D., G.-C.T., L.C., A.-G.M. and C.D.; Visualization, A.D., G.-C.T. and A.-G.M.; Writing—original draft, A.D. and G.-C.T.; Writing—review and editing, L.C., A.-G.M., L.-A.C. and C.D. All authors have read and agreed to the published version of the manuscript.

Funding: The work is supported by a grant of the Romanian Ministry of Research and Innovation, project CNFIS-FDI-2024-F-0302—"Development and adaptation of collaborative processes in excellence research conducted at BUES, in the context of modern challenges brought by Open Science and Artificial Intelligence (eXROS)". The work is also supported by the project "Analysis of the Economic Recovery and Resilience Process in Romania in the Context of Sustainable Development" coordinated by the Bucharest University of Economic Studies.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kaushal, A.; Kumar, S.; Kumar, R. A Review on Deepfake Generation and Detection: Bibliometric Analysis. *Multimed. Tools Appl.* **2024**. [CrossRef]
2. Natsume, R.; Yatagawa, T.; Morishima, S. RSGAN: Face Swapping and Editing Using Face and Hair Representation in Latent Spaces. In Proceedings of the ACM SIGGRAPH 2018 Posters, New York, NY, USA, 12 August 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 1–2.
3. Thies, J.; Zollhöfer, M.; Stamminger, M.; Theobalt, C.; Nießner, M. Face2Face: Real-Time Face Capture and Reenactment of RGB Videos 2020. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
4. Gil, R.; Virgili-Gomà, J.; López-Gil, J.-M.; García, R. Deepfakes: Evolution and Trends. *Soft Comput.* **2023**, *27*, 11295–11318. [CrossRef]
5. Mirsky, Y.; Lee, W. The Creation and Detection of Deepfakes: A Survey. *ACM Comput. Surv.* **2021**, *54*, 1–41. [CrossRef]
6. Hancock, J.T.; Bailenson, J.N. The Social Impact of Deepfakes. *Cyberpsychology Behav. Social. Netw.* **2021**, *24*, 149–152. [CrossRef] [PubMed]
7. Whittaker, L.; Letheren, K.; Mulcahy, R. The Rise of Deepfakes: A Conceptual Framework and Research Agenda for Marketing. *Australas. Mark. J.* **2021**, *29*, 204–214. [CrossRef]
8. Raghavendra, R.; Raja, K.B.; Busch, C. Detecting Morphed Face Images. In Proceedings of the 2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS), Niagara Falls, NY, USA, 1–6 September 2016; IEEE: Niagara Falls, NY, USA, 2016; pp. 1–7.
9. Rossler, A.; Cozzolino, D.; Verdoliva, L.; Riess, C.; Thies, J.; Nießner, M. FaceForensics: A Large-Scale Video Dataset for Forgery Detection in Human Faces. *arXiv* **2018**, arXiv:1803.09179.
10. Raghavendra, R.; Raja, K.B.; Venkatesh, S.; Busch, C. Transferable Deep-CNN Features for Detecting Digital and Print-Scanned Morphed Face Images. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; IEEE: Honolulu, HI, USA, 2017; pp. 1822–1830.
11. Zhang, X.; Karaman, S.; Chang, S.-F. Detecting and Simulating Artifacts in GAN Fake Images 2019. In Proceedings of the 2019 IEEE International Workshop on Information Forensics and Security (WIFS), Delft, The Netherlands, 9–12 December 2019.
12. Rana, M.S.; Nobi, M.N.; Murali, B.; Sung, A.H. Deepfake Detection: A Systematic Literature Review. *IEEE Access* **2022**, *10*, 25494–25513. [CrossRef]
13. WoS Web of Science. Available online: https://www.webofscience.com/wos/woscc/basic-search (accessed on 9 September 2023).
14. Aria, M.; Cuccurullo, C. Bibliometrix: An R-Tool for Comprehensive Science Mapping Analysis. *J. Informetr.* **2017**, *11*, 959–975. [CrossRef]
15. Firdaniza, F.; Ruchjana, B.; Chaerani, D.; Radianti, J. Information Diffusion Model in Twitter: A Systematic Literature Review. *Information* **2021**, *13*, 13. [CrossRef]
16. Dewamuni, Z.; Shanmugam, B.; Azam, S.; Thennadil, S. Bibliometric Analysis of IoT Lightweight Cryptography. *Information* **2023**, *14*, 635. [CrossRef]
17. Rejeb, A.; Rejeb, K.; Treiblmaier, H. Mapping Metaverse Research: Identifying Future Research Areas Based on Bibliometric and Topic Modeling Techniques. *Information* **2023**, *14*, 356. [CrossRef]

18. Sandu, A.; Ioanăș, I.; Delcea, C.; Geantă, L.-M.; Cotfas, L.-A. Mapping the Landscape of Misinformation Detection: A Bibliometric Approach. *Information* **2024**, *15*, 60. [CrossRef]
19. Block, J.H.; Fisch, C. Eight Tips and Questions for Your Bibliographic Study in Business and Management Research. *Manag. Rev. Q.* **2020**, *70*, 307–312. [CrossRef]
20. Bakır, M.; Özdemir, E.; Akan, Ş.; Atalık, Ö. A Bibliometric Analysis of Airport Service Quality. *J. Air Transp. Manag.* **2022**, *104*, 102273. [CrossRef]
21. Cobo, M.J.; Martínez, M.A.; Gutiérrez-Salcedo, M.; Fujita, H.; Herrera-Viedma, E. 25years at Knowledge-Based Systems: A Bibliometric Analysis. *Knowl. Based Syst.* **2015**, *80*, 3–13. [CrossRef]
22. Mulet-Forteza, C.; Martorell-Cunill, O.; Merigó, J.M.; Genovart-Balaguer, J.; Mauleon-Mendez, E. Twenty Five Years of the Journal of Travel & Tourism Marketing: A Bibliometric Ranking. *J. Travel Tour. Mark.* **2018**, *35*, 1201–1221. [CrossRef]
23. Domenteanu, A.; Delcea, C.; Chiriță, N.; Ioanăș, C. From Data to Insights: A Bibliometric Assessment of Agent-Based Modeling Applications in Transportation. *Appl. Sci.* **2023**, *13*, 12693. [CrossRef]
24. Delcea, C.; Chirita, N. Exploring the Applications of Agent-Based Modeling in Transportation. *Appl. Sci.* **2023**, *13*, 9815. [CrossRef]
25. Liu, W. The Data Source of This Study Is Web of Science Core Collection? Not Enough. *Scientometrics* **2019**, *121*, 1815–1824. [CrossRef]
26. Liu, F. Retrieval Strategy and Possible Explanations for the Abnormal Growth of Research Publications: Re-Evaluating a Bibliometric Analysis of Climate Change. *Scientometrics* **2023**, *128*, 853–859. [CrossRef]
27. Jigani, A.-I.; Delcea, C.; Florescu, M.-S.; Cotfas, L.-A. Tracking Happiness in Times of COVID-19: A Bibliometric Exploration. *Sustainability* **2024**, *16*, 4918. [CrossRef]
28. Sandu, A.; Cotfas, L.-A.; Stănescu, A.; Delcea, C. Guiding Urban Decision-Making: A Study on Recommender Systems in Smart Cities. *Electronics* **2024**, *13*, 2151. [CrossRef]
29. Liu, W. A Matter of Time: Publication Dates in Web of Science Core Collection. *Scientometrics* **2021**, *126*, 849–857. [CrossRef]
30. Sandu, A.; Cotfas, L.-A.; Delcea, C.; Crăciun, L.; Molănescu, A.G. Sentiment Analysis in the Age of COVID-19: A Bibliometric Perspective. *Information* **2023**, *14*, 659. [CrossRef]
31. Yang, J.-M.; Tseng, S.-F.; Won, Y.-L. A Bibliometric Analysis on Data Mining Using Bradford's Law. In Proceedings of the 3rd International Conference on Intelligent Technologies and Engineering Systems (ICITES2014), Kaohsiung, Taiwan, 1 December 2014; Juang, J., Ed.; Springer International Publishing: Cham, Switzerland, 2016; pp. 613–620.
32. Kushairi, N.; Ahmi, A. Flipped Classroom in the Second Decade of the Millenia: A Bibliometrics Analysis with Lotka's Law. *Educ. Inf. Technol.* **2021**, *26*, 4401–4431. [CrossRef] [PubMed]
33. Leibowicz, C.R.; McGregor, S.; Ovadya, A. The Deepfake Detection Dilemma: A Multistakeholder Exploration of Adversarial Dynamics in Synthetic Media. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, Virtual, 30 July 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 736–744.
34. Narin, F.; Hamilton, K.S. Bibliometric Performance Measures. *Scientometrics* **1996**, *36*, 293–310. [CrossRef]
35. Chesney, B.; Citron, D. Deep Fakes: A Looming Challenge for Privacy. *Calif. L. Rev.* **2019**, *107*, 1753. [CrossRef]
36. Tolosana, R.; Vera-Rodriguez, R.; Fierrez, J.; Morales, A.; Ortega-Garcia, J. Deepfakes and beyond: A Survey of Face Manipulation and Fake Detection. *Inf. Fusion* **2020**, *64*, 131–148. [CrossRef]
37. Vaccari, C.; Chadwick, A. Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Soc. Media + Soc.* **2020**, *6*, 2056305120903408. [CrossRef]
38. Verdoliva, L. Media Forensics and DeepFakes: An Overview. *IEEE J. Sel. Top. Signal Process.* **2020**, *14*, 910–932. [CrossRef]
39. Chesney, R.; Citron, D. Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics. *Foreign Aff.* **2019**, *98*, 147.
40. Bimber, B.; Gil de Zúñiga, H. The Unedited Public Sphere. *New Media Soc.* **2020**, *22*, 700–715. [CrossRef]
41. Yang, J.; Li, A.; Xiao, S.; Lu, W.; Gao, X. MTD-Net: Learning to Detect Deepfakes Images by Multi-Scale Texture Difference. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 4234–4245. [CrossRef]
42. Fletcher, J. Deepfakes, Artificial Intelligence, and Some Kind of Dystopia: The New Faces of Online Post-Fact Performance. Available online: https://muse.jhu.edu/article/715916 (accessed on 14 May 2024).
43. Guo, Z.; Yang, G.; Chen, J.; Sun, X. Fake Face Detection via Adaptive Manipulation Traces Extraction Network. *Comput. Vis. Image Underst.* **2021**, *204*, 103170. [CrossRef]
44. Yang, J.; Xiao, S.; Li, A.; Lu, W.; Gao, X.; Li, Y. MSTA-Net: Forgery Detection by Generating Manipulation Trace Based on Multi-Scale Self-Texture Attention. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 4854–4866. [CrossRef]
45. Rini, R. Deepfakes and the Epistemic Backstop. *Philos. Impr.* **2020**, *20*, 1–16.
46. Yang, J.; Xiao, S.; Li, A.; Lan, G.; Wang, H. Detecting Fake Images by Identifying Potential Texture Difference. *Future Gener. Comput. Syst.* **2021**, *125*, 127–135. [CrossRef]
47. Yu, P.; Fei, J.; Xia, Z.; Zhou, Z.; Weng, J. Improving Generalization by Commonality Learning in Face Forgery Detection. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 547–558. [CrossRef]
48. Johnson, J. Delegating Strategic Decision-Making to Machines: Dr. Strangelove Redux? *J. Strateg. Stud.* **2022**, *45*, 439–477. [CrossRef]
49. Liu, W. Caveats for the Use of Web of Science Core Collection in Old Literature Retrieval and Historical Bibliometric Analysis. *Technol. Forecast. Soc. Chang.* **2021**, *172*, 121023. [CrossRef]

50. Vera-Baceta, M.-A.; Thelwall, M.; Kousha, K. Web of Science and Scopus Language Coverage. *Scientometrics* **2019**, *121*, 1803–1813. [CrossRef]
51. WoS Document Types. Available online: https://webofscience.help.clarivate.com/en-us/Content/document-types.html (accessed on 3 December 2023).
52. Nguyen, T.T.; Nguyen, Q.V.H.; Nguyen, D.T.; Nguyen, D.T.; Huynh-The, T.; Nahavandi, S.; Nguyen, T.T.; Pham, Q.-V.; Nguyen, C.M. Deep Learning for Deepfakes Creation and Detection: A Survey. *Comput. Vis. Image Underst.* **2022**, *223*, 103525. [CrossRef]
53. Sadhana, P.; Ravishankar, N.; Ashok, A.; Ravichandran, R.; Paul, R.; Murali, K. Enhancing Fake Image Detection: A Novel Two-Step Approach Combining GANs and CNNs. *Procedia Comput. Sci.* **2024**, *235*, 810–819. [CrossRef]