Soft Computing (2023) 27:11295–11318
https://doi.org/10.1007/s00500-023-08605-y

DATA ANALYTICS AND MACHINE LEARNING

Deepfakes: evolution and trends

Rosa Gil[1] · Jordi Virgili-Gomà[1] · Juan-Miguel López-Gil[2] · Roberto García[1]

Accepted: 21 May 2023 / Published online: 15 June 2023

Abstract

This study conducts research on deepfakes technology evolution and trends based on a bibliometric analysis of the articles published on this topic along with six research questions: What are the main research areas of the articles in deepfakes? What are the main current topics in deepfakes research and how are they related? Which are the trends in deepfakes research? How do topics in deepfakes research change over time? Who is researching deepfakes? Who is funding deepfakes research? We have found a total of 331 research articles about deepfakes in an analysis carried out on the Web of Science and Scopus databases. This data serves to provide a complete overview of deepfakes. Main insights include: different areas in which deepfakes research is being performed; which areas are the emerging ones, those that are considered basic, and those that currently have the most potential for development; most studied topics on deepfakes research, including the different artificial intelligence methods applied; emerging and niche topics; relationships among the most prominent researchers; the countries where deepfakes research is performed; main funding institutions. This paper identifies the current trends and opportunities in deepfakes research for practitioners and researchers who want to get into this topic.

Keywords Deepfakes · Artificial intelligence · Deep learning · Bibliometrics

1 Introduction

Deepfake technology can be used to forge synthetic media that people cannot differentiate from true ones. It is a recent research area in which researchers in academia and industry have contributed deepfake databases, and synthesis and detection algorithms, which has made the deepfake popularity grow. Deepfakes are the product of artificial intelligence (AI) applications that merge, combine, replace, and superimpose images and video clips to create fake videos that appear authentic (Maras and Alexandrou 2019). Deepfakes use recent advances in deep neural networks to create hyper-realistic synthetic media. When deepfake technology is used on videos or images, the face of a person can be swapped with another face leaving little trace of manipulation (Chawla 2019). The emergence of deep learning has made previously existing fake face detection strategies vulnerable (Cho and Jeong 2017).

The availability of deepfake databases and synthesis and detection algorithms have made it possible for the community and even amateurish users to perform realistic deepfakes, which in turn has made the amount of popularity deepfake videos in the wild grow immensely (Pu et al. 2021a). Coupled with the reach and speed of social media, convincing deepfakes can quickly reach millions of people and have negative impacts on our society (Westerlund 2019).

The growth in deepfakes research has also been reflected in the amount of related scientific literature. Apart from technological aspects related to deepfake creation and detection, ethical, social, and legal aspects have also been carefully analyzed. There are already some reviews in specific fields, such as Creation and detection of deepfakes (Mirsky and Lee 2021), Law (da Silva 2021), Forensics (Amerini et al. 2021a), and Social impact (Hancock and Bailenson 2021a), to name a few. Still, none of them contemplates the full spectrum of research areas in deepfakes, which we believe can be very useful for researchers who wish to work on this

✉ Roberto García
roberto.garcia@udl.cat

Rosa Gil
rosamaria.gil@udl.cat

Jordi Virgili-Gomà
jordi.virgili@udl.cat

Juan-Miguel López-Gil
juanmiguel.lopez@ehu.eus

1 Universitat de Lleida, 25001 Lleida, Spain

2 University of the Basque Country, 20018 Donostia-San Sebastián, Spain

research topic. Despite its novelty, deepfakes research is a fast-growing research area, in which the research topics and their relationship is continuously changing over time and new trends appear. The different areas in which deepfakes research is performed indicate there are researchers with a wide variety of backgrounds. Apart from current trends, analyzing the funding opportunities is interesting to help focus the research effort.

The objective of this work is to get an overview of the current trends and evolution of deepfakes research, as well as to analyze the fields in which it is being applied. To this aim, all the empirical evidence that fits pre-specified eligibility criteria to answer the following six specific research questions was collated in Scopus and Web of Science databases: What are the main research areas of the articles in deepfakes? What are the main current topics in deepfakes research and how are they related? Which are the trends in deepfakes research? How do topics in deepfakes research change over time? Who is researching deepfakes? Who is funding deepfakes research? It has been decided which disciplines are developing, which are consolidating, and which are promising. The most studied areas of deep learning research, including the various artificial intelligence techniques used, have also been examined, along with emerging and niche topics. Relationships between the most well-known scientists, the nations where deepfakes research is conducted, and the major funding organizations have also been established. The prospects and trends in deepfakes research are identified in this article for practitioners and scholars who are interested in the subject.

The remainder of this paper is structured as follows. The next section presents the methods used to obtain the sample of articles to study that determine the focus, the specific research questions we seek to answer, and the software used to automate part of the process. In the results section, we expose the findings of specified research questions. After providing some reflections on the discussion, conclusions are drawn.

2 Methods

A systematic review attempts to collate all the empirical evidence that fits pre-specified eligibility criteria to answer a specific research question (Higgins et al. 2019). Therefore, the authors have ensured that the review addresses relevant questions to those who are expected to use and act upon its conclusions. More specifically, the research questions addressed by this review paper are:

- RQ1: What are the main research areas of the articles in deepfakes?
- RQ2: What are the main current topics in deepfakes research and how are they related?
- RQ3: Which are the trends in deepfakes research?
- RQ4: How do topics in deepfakes research change over time?
- RQ5: Who is researching deepfakes?
- RQ6: Who is funding deepfakes research?

Table 1 Records retrieved from Scopus and Web of Science in July and October 2021, between parentheses those in English

	Results (in English)		
	July 2021	October 2021	Growth
Scopus	242 (229)	331 (311)	89 (82)
Web of Science	8 (6)	12 (10)	4 (4)

Once the research questions were established, the starting point was a search carried out in Scopus in July 2021 and another in October of the same year. The specific query used in the case of Scopus was:

```
ALL (
  ( deepfake deep-fake "deep fake" ) AND
  ( ( action unit  OR  facial action unit
     coding system  OR facs ) OR ( video
     OR  clip  OR  image  OR  photogram )
  ) )
```

The same procedure was followed in Web of Science (WoS) also in July and October. The query in the case of WoS was:

```
TS=(
  ( deepfake deep-fake "deep fake" )
AND  ((Action Unit OR Facial Action
      Unit Coding System OR FACS) OR
(video OR clip OR image OR photogram))
```

As summarized in Table 1, the Scopus query retrieved a total of 242 records (229 in English) in July and 331 (311 in English) in October. The range of years for the retrieved records was from 2018 to 2021. There were no results before 2018 from any of the databases. In the case of Web of Science, the results were 8 in July (6 in English) and 12 in October (10 in English).

The first objective of these queries was to check if the same articles were being published in both databases and to estimate the rate of growth of the number of publications from the change between the July and October requests. Given the small number of results from Web of Science, and that just one of them is not present in the Scopus results, the detailed analysis focuses on the October results in English, i.e., 311 records from Scopus, from now on the SDO21 (Scopus Database October 2021) dataset. The dataset records are listed in Appendix A, divided into clusters based on their keywords, and available for download online.[1] It is also important to note the importance that conference publications have concerning deepfakes research as they are not included in the Web of Science. One hundred and seventy-nine of the records from Scopus are conference papers.

Given the size of the SDO21 dataset, the review has been automated using the Bibliometrix (Aria and Cuccurullo 2017) package for R, including the Biblioshiny application, as detailed in Sect. 3. Regarding transparent reporting of systematic review and meta-analysis, a PRISMA Flow Diagram [2] has not been considered necessary because the process has been simple. All the records retrieved have been considered, with the only exception of articles not in English, to facilitate the automated analysis using Bibliometrix. In any case, as observed in Table 1, the number of records that are not in English just represents between 5% and 6% of the results, in July and October, respectively.

3 Results

3.1 Main topics in deepfakes research

Regarding the first two research questions,RQ1: What are the main research areas of the articles in deepfakes? and RQ2: What are the main current topics in deepfakes research and how are they related?, our first exploration considers just the review papers, the focus of which is mainly placed on ethical and legal aspects as detailed next:

- Forensics (Amerini et al. 2021a; Castillo Camacho and Wang 2021a)
- Pornography (Karasavva and Noorbhai 2021a)
- Law (O'Donnell 2021; da Silva 2021; Aboueldahab and Freixo 2021a; Colon 2020a; Meskys et al. 2020a; Farish 2020a; Perot and Mostert 2020a)
- Theater (Fletcher 2018a)
- Social impact (Hancock and Bailenson 2021a)
- Social spam (Rao et al. 2021)
- Creation and detection of deepfakes (Mirsky and Lee 2021).

If we broaden to the whole set of 311 papers and just analyze the research areas they belong to, Computer Science is the most represented with 40.8% of the records related to this area. It is followed by Engineering (19,5%) and Social Sciences (9,4%), as shown in Fig. 1. It is important to note that papers might belong to more than one area, as defined by the corresponding literature database for each journal and year. We consider all areas when calculating these percentages as a way to recognize the interdisciplinary nature of deepfakes, with scientific journals aiming to promote interdisciplinary research and facilitate collaboration among researchers with diverse expertise.

To get deeper into the specific topics deepfakes research is dealing with, a knowledge discovery approach has been applied to identify the underlying conceptual structure. The keywords associated with each record in the SDO21 dataset have been analyzed with the Bibliometrix R package. The conceptual structure represents the relationship among the records' keywords. Keywords that appear together in a paper corresponding to a record are connected in the resulting co-keywords network. Keywords will be close in this network if a large proportion of papers have them together. Otherwise, they will be apart.

The process to create this co-keywords network that highlights the main research topics is first to create a co-occurrence symmetric matrix. As shown in Fig. 2, the elements in the diagonal k_{ii} correspond to the total amount of occurrences of each keyword in the whole SDO21 dataset. On the other hand, the element outside of the diagonal, k_{ij}, corresponds to how many times the keyword i and keyword j appear together in the same paper.

The co-keywords matrix is then used to generate the keywords network that highlights the research topics structure in deepfakes research. The network is an undirected graph where the nodes correspond to keywords and whose size depends on the keyword frequency, thus generated from the matrix's diagonal.

Then, two graph nodes are connected if the matrix cell for the corresponding keywords is greater than 0, and thus both keywords share at least one paper. The edges are weighted with the value of that cell, i.e., the number of papers where both keywords appear together as captured in the non-diagonal cells. Edges' weight is interpreted as a measure of the strength between two keywords, the higher they appear the closer they are on the graph. Based on this interpretation of the matrix, the graph can be rendered as shown in Fig. 3 and highlights the main research topics corresponding to the most frequent keywords. This technique processes keywords as text strings and thus does not include any kind of semantic similarity measure. It focuses on the keywords associated with each publication.

Co-occurrence networks use various measures to identify crucial nodes or vertices within the network. Among these measures, Betweenness (Table 2), Closeness (Table 3), and PageRank (Table 4) are used to provide notable insights. When considering the top 5 keywords for each metric, a sum of 8 unique keywords is obtained. This is consistent as each measure is capturing a different aspect of the network of keyword co-occurrences. Betweenness quantifies how often a node falls on the shortest paths between other nodes in

[1] Replication data, https://doi.org/10.34810/data750.

[2] http://prisma-statement.org.

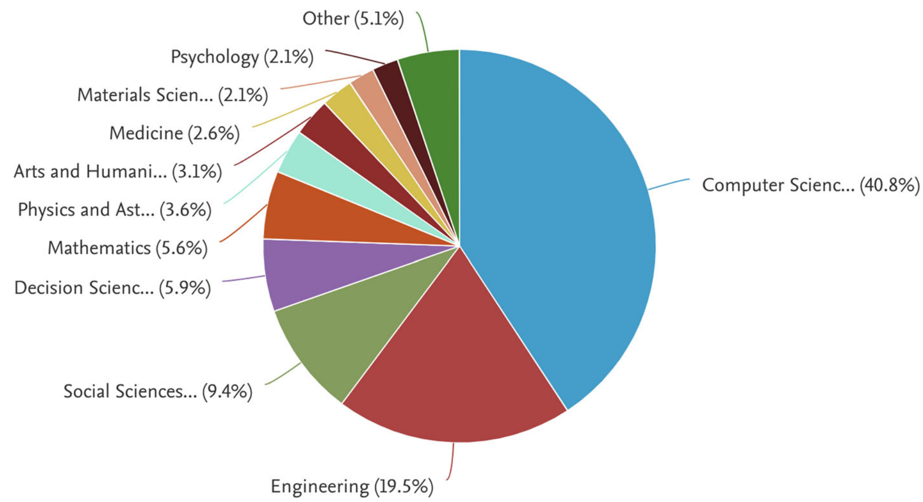


Fig. 1 Main research areas for the papers included in the SDO21 dataset (311 records retrieved from Scopus on October 2021)

$$\begin{bmatrix} k_{11} & k_{12} & k_{13} & \dots & k_{1n\text{-}1} & k_{1n} \\ k_{21} & k_{22} & k_{23} & \dots & k_{2n\text{-}1} & k_{2n} \\ k_{31} & k_{32} & k_{33} & \dots & k_{3n\text{-}1} & k_{3n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ k_{n\text{-}11} & k_{n\text{-}12} & k_{n\text{-}13} & \dots & k_{n\text{-}1n\text{-}1} & k_{n\text{-}1n} \\ k_{n1} & k_{n2} & k_{n3} & \dots & k_{nn\text{-}1} & k_{nn} \end{bmatrix}$$

Fig. 2 Co-keyword matrix used to generate the network highlighting the research topics in deepfakes research shown in Fig. 3

the network. Nodes with high Betweenness are critical since they connect different parts of the network, playing a vital role in the flow of information or resources between distinct groups of nodes. Closeness measures how closely connected a node is to all other nodes in the network. Nodes with high Closeness are significant since they have rapid access to a vast amount of information or resources and can disseminate them quickly throughout the network. PageRank assesses a node's importance based on the number and quality of incoming links it has. Nodes with high PageRank are crucial since they are highly connected to other important nodes in the network. In identifying key intermediaries or brokers in the network, Betweenness is the most critical measure. If the aim is to identify nodes that can quickly disseminate information throughout the network, Closeness is the most critical measure. Finally, to identify nodes that shape the network's overall behavior, PageRank is the most important measure. It is often useful to calculate all three measures to gain a comprehensive understanding of the network's structure and dynamics.

This representation makes it easier to visualize how the main research topics are organized in deepfakes research. Just the most representative topics, corresponding to the most used keywords, are shown. And they are more prominent the more present they are in the SDO21 dataset. Highly related topics, because they are covered jointly in many papers, are shown closer. This makes it also possible to apply a clustering algorithm that helps identify the main research topics and

Table 2 Top 5 keywords by co-occurrence (Betweenness)

Node	Cluster	Betweenness
Deep learning	5	334.138
Convolutional neural networks	2	172.340
Face recognition	3	123.595
Detection methods	5	85.786
Computer vision	3	68.919

Table 3 Top 5 keywords by co-occurrence (Closeness)

Node	Cluster	Closeness
Deep learning	5	0.0164
Convolutional neural networks	2	0.015
Detection methods	5	0.0144
Face recognition	3	0.014
Digital forensics	2	0.013

Table 4 Top 5 keywords by co-occurrence (PageRank)

Node	Cluster	PageRank
Deep learning	5	0.105
Convolutional neural networks	2	0.066
Detection methods	5	0.059
Face recognition	3	0.054
Adversarial networks	2	0.040

the central keywords giving the name to the corresponding topics:

- Red: deep learning, adversarial networks, learning systems, etc.
- Blue: face recognition, detection methods, forgery detections, etc.

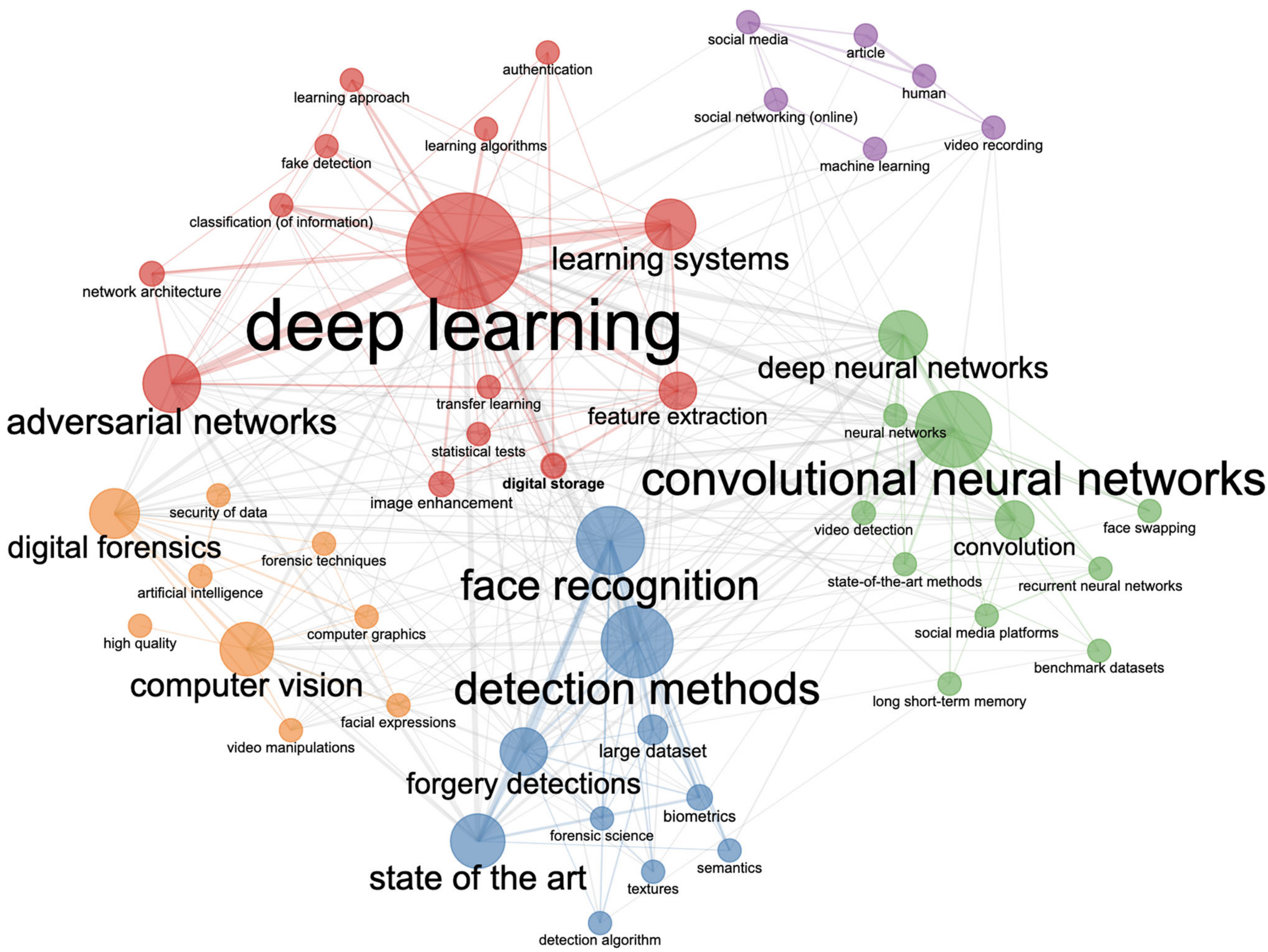


Fig. 3 Co-keywords graph representing the main research topics in deepfakes research

- Green: convolutional neural networks, deep neural networks, etc.
- Orange: computer vision, digital forensics, etc.
- Purple: social media, video recording, social networking, etc.

3.2 Trends and evolution of deepfakes research

In this section, we address the third and fourth research questions, RQ3: Which are the trends in deepfakes research? and RQ4: How do topics in deepfakes research change over time?. Despite the short time interval under study, the SDO21 dataset includes records from 2018 to 2021, it is possible to observe the evolution of the main research topics and identify their trends.

First of all, after applying a clustering algorithm to the keywords as detailed in the previous section, we can do more than just highlight the main topics of the deepfakes research domain. Each topic can be represented on a plot called Thematic Map (Cobo et al. 2011) as shown in Fig. 4.

This kind of plot classifies the cluster of keywords from the co-keyword network obtained in the previous section according to Callon's centrality and density measures (Callon et al. 1991):

- *Centrality*: measures the strength of the links to other topics, considering those from keywords included in a cluster to keywords in other clusters. Thus, it measures the importance of a topic in the context of the whole field of study.
- *Density*: is related to the strength of internal links among all keywords corresponding to the same topic cluster. It is interpreted as a measure of the topic's development degree.

Centrality and Density define the two axes of the Thematic Map and are used to divide it into four regions. The topics in these regions are associated with the following trends:

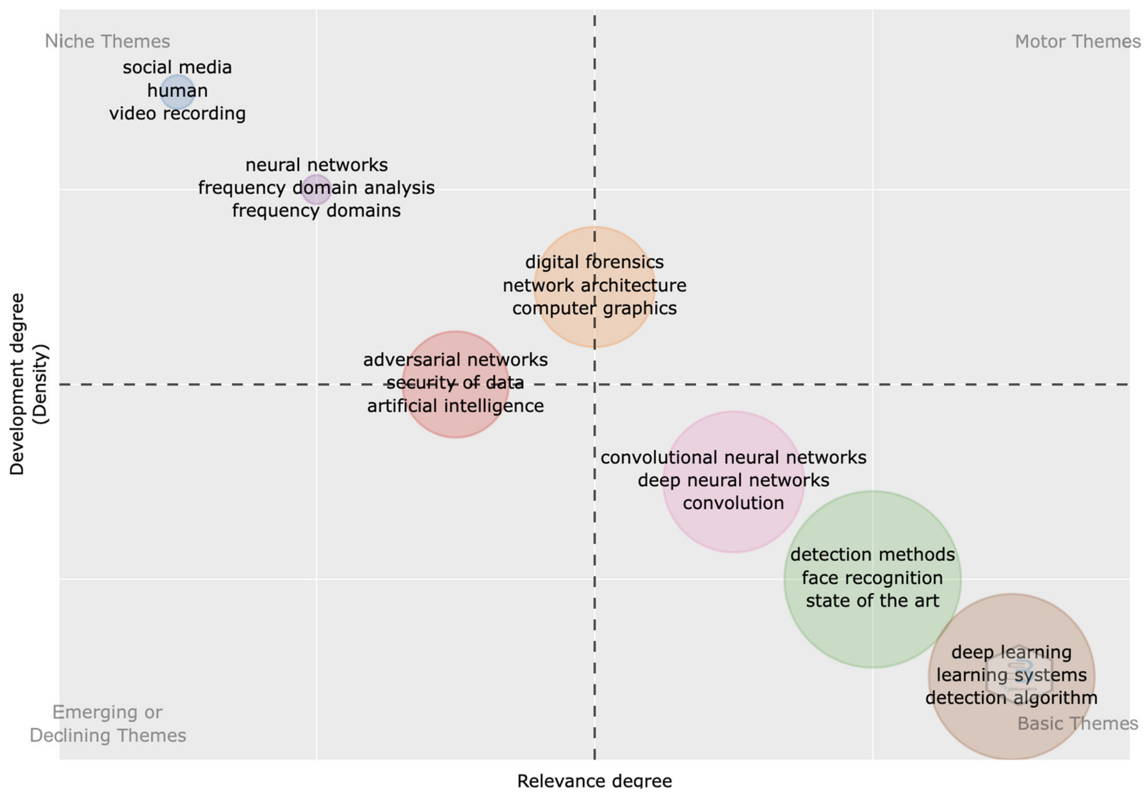


Fig. 4 Thematic Map the topic trends in deepfakes research. From top-left to bottom-right: Niche, Motor, Emerging, and Basic topics

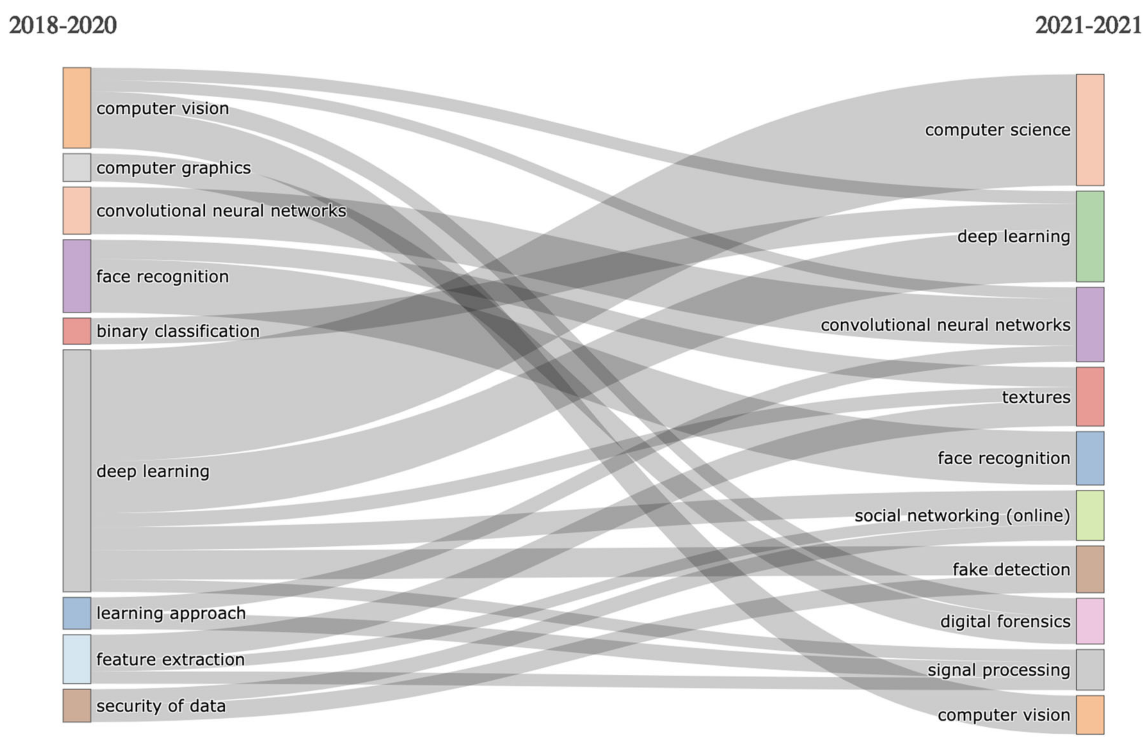


Fig. 5 This is the keywords plus graph. The colors represent the different clusters: deep learning (red), convolutional neural networks (green), computer vision (orange), and detection methods (blue)

- *Niche Topics (upper-left quadrant)*: well-developed internal ties but unimportant external ties and so, they have a marginal role in the development of the research field under study. Non-central but dense.
- *Motor Topics (upper-right quadrant)*: these topics are both well-developed and important in the context of the analyzed records. High centrality and density.
- *Emerging or Declining Topics (lower-left quadrant)*: they are both weakly developed and marginal, with low centrality and density.
- *Basic Topics (lower-right quadrant)*: they are important for a research field but are not developed, i.e., they show high centrality but low density.

The Motor and Basic topics are considered those that favor the development and consolidation of a research field due to their density and/or centrality. For the particular case of deepfakes research captured by the SDO21 dataset, there is a lack of clear Motor Topics. Most of them are Basic Topics related to the core of technologies used for deepfakes development, as is the case of convolutional neural networks or deep neural networks. This is also the case with detection methods such as facial recognition.

The only topics that are partially classified as Motor Topics, and thus are computer graphics, network architecture, and digital forensics. This seems related to the fact that, as noted at the beginning of Section 3.1, there are two reviews on the particular topic of forensics in the last four years.

On the other hand, the topics partially related to Emerging Topics (declining seems unfeasible given the youth of the discipline and the short time range) are those associated with artificial intelligence, data security, and adversarial networks. Finally, the more mature topics, though apart from the main efforts in this research domain, are those that have to do with the analysis in time and frequency to achieve better returns such as video recording or social networks.

It is important to note that what is being classified into these different trends are the keywords associated with the papers. Thus, quite related topics that might be even equivalent in some contexts, like "deep neural networks" and "neural networks," might be classified in different quadrants based on their use in the analyzed literature. The approach is thus completely agnostic regarding the interpretation of these keywords because they are highly contextual, like in the case of neural networks methods and applications (Samek et al. 2021).

In addition to the static view provided by the Thematic Map in Fig. 4, it is also possible to get an idea of the underlying dynamics using the Thematic Evolution diagram shown in Fig. 5. Thematic Maps for different periods are computed to identify topics' evolution over time. Topics at a particular period are then connected with those in the following one to create a stream of topics' evolution. Linking among topics is based on the percentage of keywords shared between the identified topics at each period. This way, it is possible to observe how initial topics might remain partially and split into other topics that then include the corresponding keywords.

For the SDO21 dataset, just two time periods have been defined given the short period, 2018–2020 and 2021. On the left of Fig. 5, there are the topics for the 2018–2020 period, including computer vision or computer graphics among others. On the right side, are those for 2021. The evolution of the topics is illustrated through the links connecting them, which are weighted based on the number of keywords shared by the topics in different periods.

For instance, the computer vision topic has split into many different ones in 2021, partially remaining as the same topic but less relevant because many of the associated keywords are now tied to other topics like deep learning, convolutional neural networks, or digital forensics. On the other hand, topics like computer graphics have disappeared and now the associated keywords are contributing to the digital forensics one, which has emerged from keywords from this topic combined with some from computer vision. Overall, Fig. 5 highlights the topics getting traction in deepfakes research and how they are consolidating from the topics that attracted the most attention just some years ago.

3.3 Deepfake technologies usage and funding

Regarding the last research questions, RQ5: Who is researching deepfakes? and RQ6: Who is funding deepfakes research?, they are addressed by analyzing the intellectual and social structures of the SDO21 dataset. First of all, and as can be observed in Table 5, the most relevant papers come from conferences, concretely from IEEE conferences and workshops. Forensics, signal processing, law, and blockchain are among the topics dealt with by the most cited articles about deepfakes research in Scopus between 2018 and 2021.

Going beyond this superficial analysis, the whole community that has generated the papers in SDO21 should be taken into account. It is for this reason that we have also carried out an analysis of the social structure to highlight how authors or institutions related to others in this particular research field. First of all through a co-authorship network, which is displayed in Fig. 6.

Many of the most referenced authors in Table 5 can be also identified in the co-authorship network, which also focuses on the most prominent authors. These authors appear in little clusters, like Amerini or Agarwal and their corresponding co-authors. This highlights that even highly cited authors' work collaborates in relatively closed circles and the overall community is quite fragmented from this perspective.

If we switch from individual researchers to their institutions and countries, we can also unveil the underlying

Table 5 The first 10 most cited articles in SDO21 between 2018 and 2021

Title	Publisher	Citations	References
FaceForensics++: Learning to Detect Manipulated Facial Images	IEEE	302	Rossler et al. (2019)
Exposing Deep Fakes Using Inconsistent Head Poses	IEEE	165	Yang et al. (2019a)
Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations	IEEE	136	Matern et al. (2019a)
Protecting World Leaders Against Deep Fakes	IEEE	105	Agarwal et al. (2019a)
Combating Deepfake Videos Using Blockchain and Smart Contracts	IEEE	80	Hasan and Salah (2019a)
Deep Fakes: A Looming Challenge for Privacy	California Law	72	Chesney and Citron (2018)
Celeb-DF: A Large-Scale Challenging Dataset for DeepFake Forensics	IEEE	64	Li et al. (2020d)
Detecting and Simulating Artifacts in GAN Fake Images	IEEE	49	Zhang et al. (2019b)
Media Forensics and DeepFakes: An Overview	IEEE	47	Verdoliva (2020)
Deepfake Video Detection through Optical Flow Based CNN	IEEE	44	Amerini et al. (2019a)

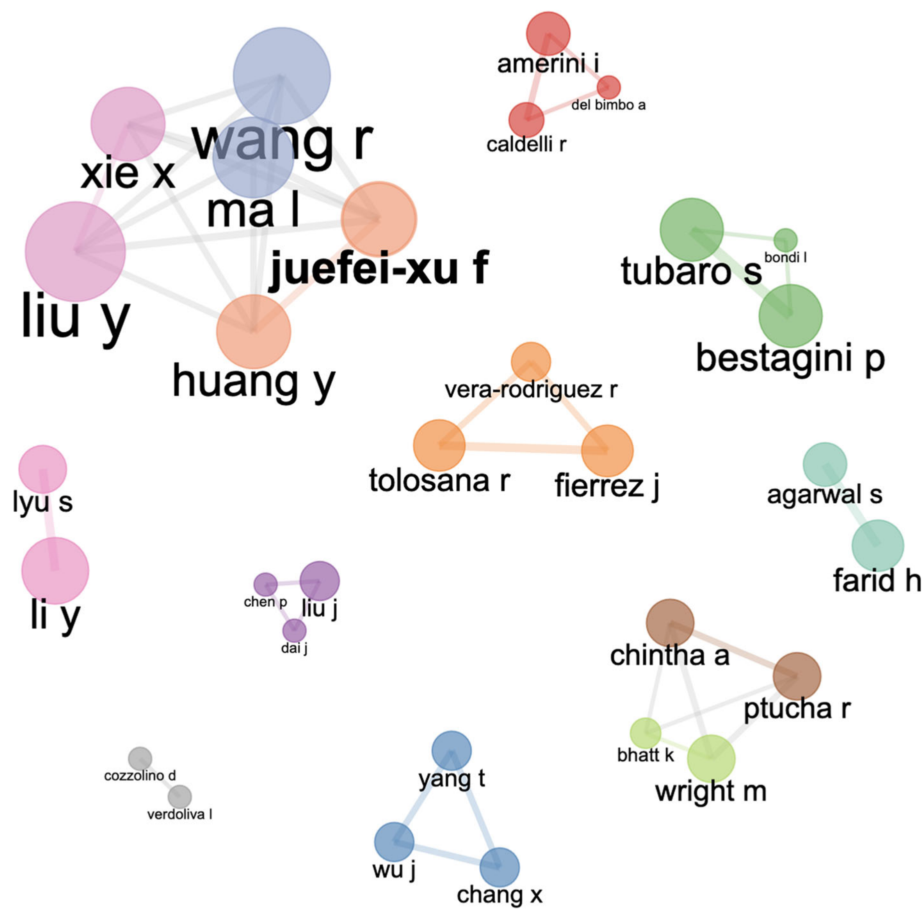


Fig. 6 Co-authorship network showing some of the more prominent authors in SDO21

social structures at these levels. Looking at the corresponding author countries, shown in Fig. 7, we can observe the great leadership that researchers from China have in this particular research area.

This is even more evident when we realize that, despite it might seem that part of this leadership comes from collaborations with other countries because it is the country with the highest amount of inter-country collaborations, these collaborations are really with Chinese researchers based in other countries. This is illustrated in Fig. 8, which shows the connection between researchers and countries, and then from countries to research topics. Therefore, although inter-country collaboration is indeed very high in China, it is because these researchers work in other countries, in most cases in the USA as shown in Fig. 8.

Finally, focusing on RQ6: Who is funding deepfakes research?, the main research funding organization of the reviewed publications is the National Natural Science (Foun-

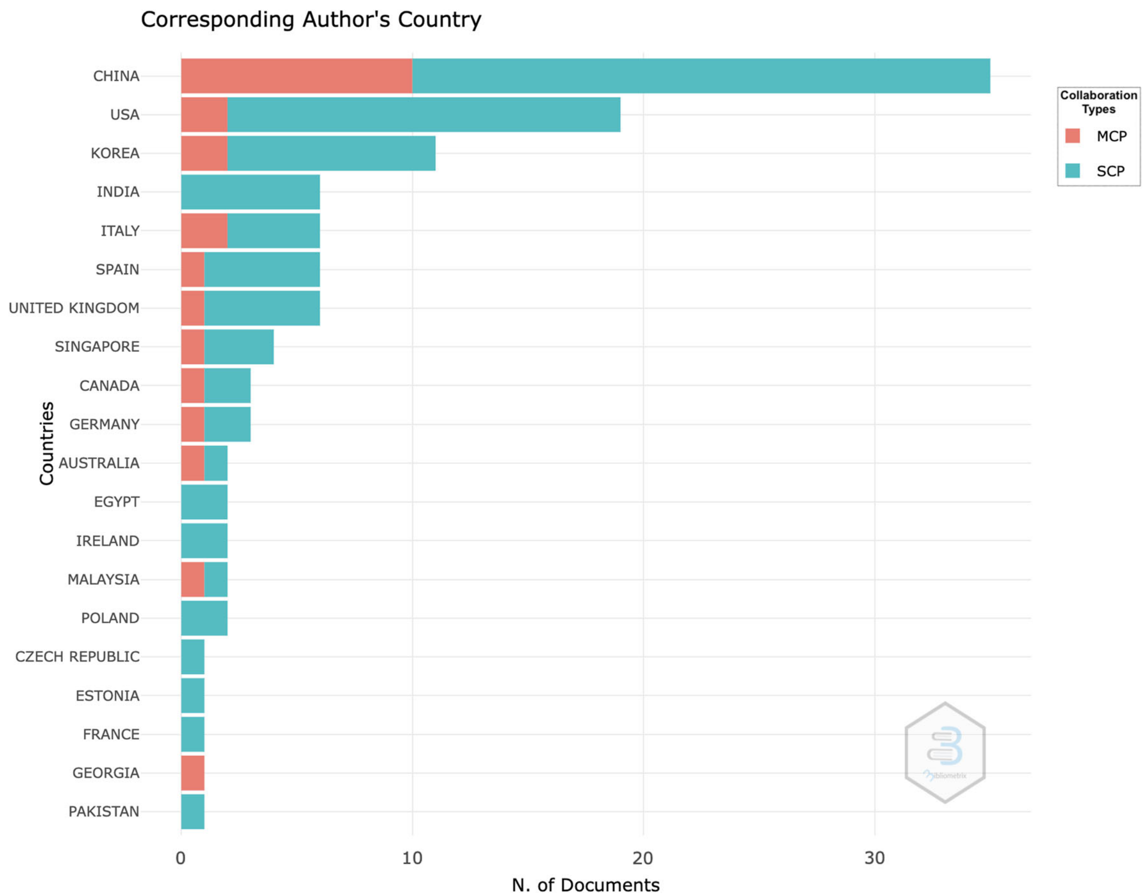


Fig. 7 Corresponding author's countries, including intra-country (SCP) and inter-country (MCP) collaborations in SDO21

dation of China) with 30 publications, followed by the Defense Advanced Research Projects Agency (US DARPA) with 22, and the National Key Research and Development Program of China with 13. Then, with 12 publications, we find the US Air Force Research Laboratory and the US National Science Foundation. A complete table with the top 10 funding organizations is shown in Table 6. Therefore, China is leading the investigation as a country, mostly from institutions related to the military and defense sectors. And as shown in Fig. 9, which displays the collaborations among institutions, these collaborations are kept at the national level.

4 Discussion

This paper employs metadata analysis to investigate the trends and tendencies related to deepfake research. It is important to note that our objective was not to conduct a literature review, but to analyze its metadata. However, it may be valuable to include this section in the paper that provides further insights into the representative results of the included publications.

Deepfakes is a field of research that has gained significant attention in recent years due to its potential implications in manipulating digital media. Following the content found in the lower-right quadrant of Fig. 4, which contains "topics that are important for the research field but are not yet fully developed" learning systems, detection methods, and algorithms are the key and future directions in the topic. One of the most common approaches used in Deepfakes is generative adversarial networks (GANs) (Hu et al. 2021). These techniques consist of two neural networks, one that generates fake data and another that evaluates the generated data authenticity. The results obtained using GANs have shown remarkable progress in generating highly realistic images and videos. Another popular method is the use of autoencoders (Singh et al. 2021), neural networks that are trained to reconstruct the input data. The encoded representation of the input is then used to generate new data. The results obtained using autoencoders have shown promise in generating high-quality Deepfakes.

In addition to GANs and autoencoders, there are other methods that have been used in Deepfakes, such as variational autoencoders (Zendran and Rusiecki 2021), deep belief networks (Iacobucci et al. 2021), and convolutional neural

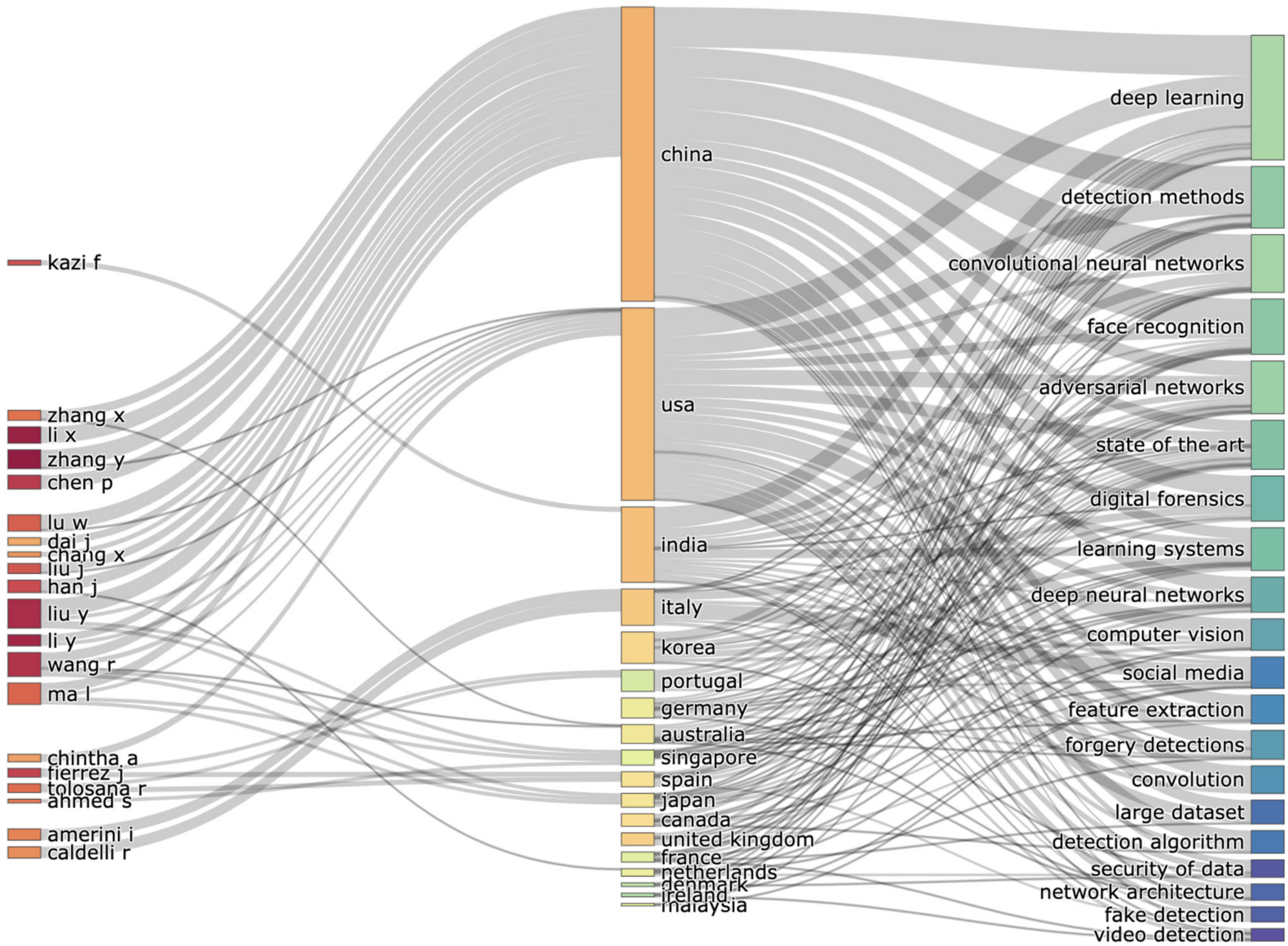
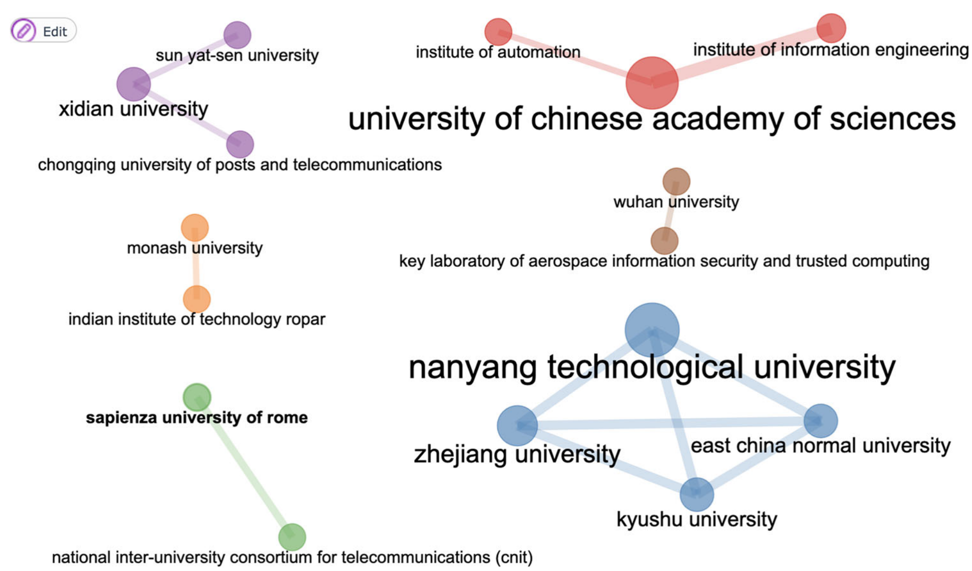


Fig. 8 Relationships among the most prominent researchers, the countries where they conduct research, and the main research topics per country

Fig. 9 Collaboration among institutions as derived from the records in the SDO21 dataset



networks (Agrawal and Sharma 2021). Each of these methods has shown varying degrees of success in generating Deepfakes. Of course, these methods are improving by applying not only new approaches but combining known techniques in a new way, as Zheng et al. (2018) proposes a novel two-stage training process for deep convolutional neural networks (CNNs) that improves their generalization ability by implicit regularization, particularly when the training data is limited.

Practical cross-area applications can be found in works like (Yao et al. 2021), where a method is proposed to automatically separate compound figures in biomedical research articles. It uses a deep learning model that is trained to separate the subfigures based on their visual features and is augmented with a "side loss" to ensure that the model also considers the context and layout of the subfigures. This article is a good example of how a single publication can show insights into distant topics from upper-left Fig. 4 (frequency domain analysis) and lower-right (detection methods) at the same time.

Despite the progress made in deepfakes, there are still limitations to the current state of the art. The primary challenges are the ability to generate realistic and high-quality deepfakes without significant artifacts (Matern et al. 2019b) and paradoxically, the ability to detect and prevent the spread of deepfakes in the public domain (Rossler et al. 2019).

Finally, regarding funding, the top five funding institutions are either government agencies (NSFC, DARPA, AFRL, and NSF) or state-sponsored programs (NKRDPC and USNCF) that prioritize funding for research projects that are strategically important to their respective countries (see Table 6). As these projects may include those with military applications or those that promote the development of key industries, it is reasonable to infer that these strategic priorities may account for the low inter-country collaboration ratio (MCP) presented in Fig. 7. This could be because research with strategic importance often challenges collaboration due to national security concerns, funding restrictions (in some cases, funds may be restricted for international collaborations), and intellectual property issues.

5 Conclusions

It has been found that growth since 2018 has skyrocketed regarding research publications in the area of deepfakes. The queries for Web of Science and Scopus did not retrieve any results before 2018 but accumulated 311 results, after less than four years, in 2021. The specific findings for each of the research questions are discussed in the next paragraphs.

RQ1: What are the main research areas of the articles in deepfakes? Deepfakes research includes many different research areas. Our analysis identified 10 different areas with at least 2% of the articles about the topic. All 10 combined represent roughly 95% of the papers. However, there is a big imbalance as just 3 of them accumulate almost 70% of the results. Computer Science is the most represented with 40.8%, followed by Engineering (19,5%). Thus, these technological research areas are those with the biggest percentage of articles. The third area is Social Sciences (9,4%), so deepfakes research is also noticeable in social sciences-related topics.

RQ2: What are the main current topics in deepfakes research and how are they related? Regarding the most studied topics, a knowledge discovery approach has been applied to identify the underlying conceptual structure starting from the keywords associated with the analyzed articles. Using a clustering algorithm, five main sets of topics have been identified, being the most representative topics in each cluster: deep learning, face recognition, convolutional neural networks, computer vision, and social media. Other relevant topics in each cluster are presented in Fig. 3. As can be observed, overall, deep learning stands out. And more specifically, adversarial and convolutional neural networks. It is also relevant to the research on forgery detection and the literature related to face recognition.

RQ3: Which are the trends in deepfakes research? The main topics identified using clustering have been analyzed using a Thematic Map, shown in Fig. 4. This kind of plot classifies the clusters of keywords obtained in the previous section according to Callon's centrality and density measures (Callon et al. 1991). Based on these measures, we can identify:

- *Niche Topics*: well-developed but with a marginal role in the development of the research field, like Social Media related to Video Recording or Neural Networks in the context of Frequency Domain Analysis.
- *Emerging or Declining Topics*: these are weakly developed and still marginal topics. Given the youth of the deepfakes discipline, they should be mainly emerging topics. Though the analysis does not identify clear emerging topics, research related to adversarial networks in the context of security might be considered an emerging area with potential relevance in the future.
- *Motor Topics*: these are both well-developed and important in the context of deepfakes. As previously stated, the youth of the discipline causes a lack of clear candidates. Just topics related to computer graphics, network architecture, and digital forensics might be classified as Motor.
- *Basic Topics*: these are the topics on which research should be focused. They are important for deepfake research but have not been developed yet. Here, we can find the bulk of the research. The most promising topics are convolutional deep neural networks and detection methods based on face recognition or deep learning.

Table 6 Top 10 Research Funding Organizations

Institution	Funded Projects
National Natural Science (Foundation of China)	30
Defense Advanced Research Projects Agency, DARPA	22
National Key Research and Development Program of China	13
Air Force Research Laboratory, AFRL	12
US National Science Foundation	12
Google	5
Nvidia	4
Ministry of Science and ICT, South Korea, MSIT	4
Ministero dell'Istruzione, dell'Università e della Ricerca	4
National Research Foundation of Korea, NRF	4

RQ4: How do topics in deepfakes research change over time? In addition to the dynamics of deepfakes research captured by the previous trends analysis, it is also possible to visualize the underlying dynamics using a Thematic Evolution chart, as shown in Fig. 5. We use Thematic Maps for different periods, which are then connected with those in the following one to create a stream of topics' evolution based on the percentage of keywords shared between the identified topics at each period. An insight that can be derived from this diagram is the diversification of the research around deep learning, which remains one of the main topics but with clear applications to texture analysis, fake detection, or online social networking. The same can be said about computer vision, which gets out of the main focus even more than deep learning. On the contrary, from a technical perspective, convolutional neural networks are getting more attention from recent research compared to the beginning of the analyzed period.

RQ5: Who is researching deepfakes? and RQ6: Who is funding deepfakes research? It is China as a country the one that directs the investigations, being the one that contributes the most in all regards, including funding through the Natural Science Foundation of China and NKRDPC. Researchers are mainly from this country, though many of them perform their research in the USA. On the other hand, the collaboration communities in this research area are still small and fragmented as observed when studying the co-authorship network. Usually, they are formed by just 2 or 3 authors, except for the most prolific Chinese researchers that are organized in a community of 6 authors. The same happens at the country level, most collaborations are among institutions of the same country. Additionally, though authors might be based on centers in different countries, we do not observe inter-country collaborations.

In addition to the conclusions regarding the different research questions, we have identified some missing research topics that we think should already be in the literature, such as research on the repercussions of deepfakes on marketing or online negotiation processes. These kinds of risks have been tangentially addressed in the context of studies about identity usurpation, which have been the topic of some law journals. In any case, we believe that considering the emerging risks of deepfakes in connection with tasks like online meetings is crucial.

As a limitation of this work, the number of articles found on deepfakes research made it impossible to perform a systematic literature review or meta-analysis on the whole area of deepfakes research. On the other hand, this type of study can be carried out by focusing on more specific aspects of the area identified by this work, such as the different artificial intelligence techniques used to synthesize or analyze deepfakes.

To conclude, the research articles retrieved about deepfakes serve to provide a complete overview of deepfakes. The main insights of this work include the various areas in which deepfakes research is being conducted, focusing on which areas are emerging, those that are considered basic, and those that currently have the greatest potential for development. The most studied topics in deepfakes research, including the various artificial intelligence methods employed, are analyzed together with emerging and niche topics, to provide insight into the current trends.

The relationships among the most prominent researchers, together with the countries in which deepfakes research is conducted and the main funding sources, complete the outlook regarding the people who carry out research in that area and the options for collaboration and obtaining existing funds.

Overall, this article discusses current trends and opportunities in deepfakes research for practitioners and researchers interested in this field. Future research directions emerging from the review point in the direction of the identified "Basic Topics": convolutional deep neural networks and detection methods based on face recognition or deep learning.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00500-023-08605-y.

Author Contributions Rosa Gil and Juan-Miguel López-Gil were involved in conceptualization and methodology; Jordi Virgili-Gomà and Roberto García helped in data curation; Roberto García contributed to funding acquisition; Jordi Virgili-Gomà was involved in validation; Rosa Gil helped in visualization; Rosa Gil, Juan-Miguel López-Gil, and Roberto García contributed to writing—original draft; Jordi Virgili-Gomà, Rosa Gil, Juan-Miguel López-Gil, and Roberto García helped in writing—review & editing.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work has been partially supported by the project "ANGRU: Applying kNowledge Graphs to research data ReUsability" with reference PID2020-117912RB-C22 and funded by MCIN/AEI/10.13039/501100011033. Additionally, this research benefits from funding from the Research Group program of the University of the Basque Country under contract GIU21/037.

Data Availability The datasets generated and analyzed during the current study are available online at https://drive.google.com/file/d/1Attj4yMnsYJB1rx9kYIdVVoQeMhqhW7k/view and are in the process of being published in the CORA-RDR repository, https://dataverse.csuc.cat.

Declarations

Competing Interests The authors have no relevant financial or non-financial interests to disclose.

Appendix A References in SDO21 Dataset

The following table lists all references in the SDO21 dataset of records retrieved from Scopus as detailed in Sect. 1. They are divided into 4 clusters centered on the keywords associated with each of them.

Cluster	References
deep learning 27.2% detection methods 18.5% face recognition 18.5%	Li et al. (2020c), Dang et al. (2020), Ciftci et al. (2020), Lyu (2020), Nguyen et al. (2021), Carlini and Farid (2020), Tursman et al. (2020), Kaur et al. (2020), Masi et al. (2020), Feng et al. (2020), Maksutov et al. (2020), Amerini et al. (2019b), Javed et al. (2021), Kuang et al. (2021), Patil et al. (2021), Megías et al. (2021), Patil and Chouragade (2021), Xu et al. (2021b), Jiang et al. (2021), Bonomi et al. (2021), Shang et al. (2021), Ling et al. (2021), Fung et al. (2021), England et al. (2021), Fazheng et al. (2021), Zhao et al. (2021b), Valenzuela et al. (2021), Xiang et al. (2021), Hosler et al. (2021), Caldelli et al. (2021), Pan et al. (2021), Khalil and Maged (2021), Demir and Ciftci (2021), Li and Lyu (2021), Whler and Zembaty (2021), Kohli and Gupta (2021), Lv (2021), Dondero (2021), Guo et al. (2021), Carter et al. (2021), Tu et al. (2021), Shelke and Kasana (2021), Pokroy and Egorov (2021), Tjon et al. (2021), Sun et al. (2021), Sankaranarayanan et al. (2021), Li et al. (2021a), Yang et al. (2021b), Kawa and Syga (2021), Gong et al. (2021), Hussain et al. (2021), Godulla et al. (2021), Tolosana et al. (2021), Jeong et al. (2021), Hernandez-Ortega et al. (2021), Zhang et al. (2021a), Deshmukh and Wankhade (2021), Caporusso (2021), Amerini and Caldelli (2020), Zhu et al. (2020a), Bonettini et al. (2020), Kukanov et al. (2020), Bondi et al. (2020), Nasar et al. (2020), Mittal et al. (2020a), Du et al. (2020a), El Rai et al. (2020), Ramadhani and Munir (2020), Gupta et al. (2020), Du et al. (2020b), Lewis et al. (2020), Chugh et al. (2020), Shah et al. (2020), Tarasiou and Zafeiriou (2020), Ross et al. (2020), Nguyen and Derakhshani (2020), Khodabakhsh and Loiselle (2020), Li et al. (2020e), Zotov et al. (2020), Wu et al. (2020a), Hongmeng et al. (2020), Chowdhury and Lubna (2020), Suratkar et al. (2020a), Hewage and Ekmekcioglu (2020), Zhao et al. (2020b), Younus and Hasan (2020a), Alattar et al. (2020), Li et al. (2020a), Peng et al. (2020), Ivanov et al. (2020), Zhao et al. (2020a), Cozzolino et al. (2019)

Cluster	References
deep learning 11.3% adversarial networks 7.5% artificial intelligence 6.6%	Kietzmann et al. (2020), Ahmed et al. (2021), Jung et al. (2020), Wang et al. (2020a), Mirsky and Lee (2021), Meskys et al. (2020a), Chesney and Citron (2019), Zhang et al. (2020), Fallis (2020), Maddocks (2020), Farish (2020a), Fletcher (2018a), Rao et al. (2021), Khormali and Yuan (2021), Lai and Patrick Rau (2021), Yu et al. (2021), Pavis (2021), Lees et al. (2021), de Seta (2021), Bode et al. (2021), Holliday (2021), Mihailova (2021), Bode (2021), Ayers (2021), Hayward and Maas (2021), Ahmed (2021c), Kim et al. (2021a), Sybrandt and Safro (2021), Diakopoulos and Johnson (2021), José and García-Ull (2021), Huber et al. (2021), Tahir and Batool (2021), Nygren et al. (2021), Fagni et al. (2021), Medoff and B.K. (2021), Pu et al. (2021b), Castillo Camacho and Wang (2021b), Mcglynn and Johnson (2021), D'Alessandra and Sutherland (2021), Freeman (2021), Karasavva and Noorbhai (2021a), Brooks (2021), Iacobucci et al. (2021), Hancock and Bailenson (2021a), Johnson (2021), Choraś et al. (2021), Tesfagergish et al. (2021), O'Donnell (2021), da Silva (2021), Thaw et al. (2021), Frick et al. (2021), Aboueldahab and Freixo (2021a), Hänska (2021), Zhang et al. (2021b), de Ruiter (2021), Ahmed (2021b), Murphy and Flynn (2021), Zhao et al. (2021a), Wahl-Jorgensen and Carlson (2021), Pavlíková et al. (2021), Dasilva et al. (2021), Vizoso et al. (2021), Johnson and Diakopoulos (2021), Chi et al. (2021), Kietzmann et al. (2021), Dobber et al. (2021), Kwok and Koh (2021), Kikerpill (2020), Perot and Mostert (2020a), Hasan and Salah (2019a), Aliman and Kester (2020), Xie et al. (2020), Pan et al. (2020), Kozyreva et al. (2020), Partadiredja et al. (2020), Tulk Jesso et al. (2020), Colon (2020a), Gosse and Burkell (2020), Lomnitz et al. (2020), Wang et al. (2020b), Katarya and Lal (2020), Kaye and Johnson (2020), Gandhi and Jain (2020), Chang et al. (2020), Šepec and Lango (2020), Hosier and Stamm (2020), Hartmann and Giles (2020), Gong et al. (2020), Jongman (2020), Pertsch et al. (2020), Shahar and Hel-Or (2020), Houde et al. (2020), Zhu et al. (2020b), Jeong (2020), Amelin and Channov (2020), Kang and Park (2020), Pashentsev (2020), Davis and Fors (2020), Hazan (2020), Bore (2020)

Cluster	References
deep learning 27.7% convolutional neural networks 23.1% detection methods 16.9%	Yang et al. (2019b), Jiang et al. (2020), Agarwal et al. (2019b), Agarwal et al. (2020a), Mittal et al. (2020b), Zi et al. (2020), Agarwal et al. (2020b), Chen et al. (2020), Montserrat et al. (2020), Ahmed (2021a), Marcon et al. (2021), Ajoy et al. (2021), Liang and Deng (2021), Ru et al. (2021), Tran et al. (2021), Ismail et al. (2021), Yavuzkilic et al. (2021), Fei et al. (2021), Yang et al. (2021a), Siegel et al. (2021), Agarwal and Farid (2021), Masood et al. (2021), Singh et al. (2021), Sanghvi et al. (2021), Xu et al. (2021a), Agrawal and Sharma (2021), Zendran and Rusiecki (2021), Trinh et al. (2021), Li et al. (2021b), Lu et al. (2021), Su et al. (2021), Luo et al. (2021), Biswas et al. (2021), Jiang et al. (2021), Korshunov and Marcel (2021), Chen and Tan (2021), Jin et al. (2021), Khalil et al. (2021), Gu et al. (2021), Yang et al. (2021c), Chintha et al. (2020a), Younus and Hasan (2020b), Korshunov and Marcel (2019), Suratkar et al. (2020b), Mitra et al. (2020), Liang et al. (2020), Burroughs et al. (2020), Huang et al. (2020a), Chintha et al. (2020b), Ki Chan et al. (2020), Zhu et al. (2020c), Yang et al. (2020), Wu et al. (2020b), Jafar et al. (2020), Şener (2020), Malolan et al. (2020), Fernandes and Jha (2020), Pantserev (2020a), Wang et al. (2020c), Zeng et al. (2020), Pantserev (2020b), Karandikar et al. (2020), Megahed and Han (2020), Albahar and Almalki (2019), Sohrawardi et al. (2019)
deep learning 36.4% adversarial networks 22. 7% computer vision 20.5%	Li et al. (2020b), Tolosana et al. (2020), Verdoliva (2020), Rossler et al. (2019), Matern et al. (2019b), Prajwal et al. (2020), Khalid and Woo (2020), Neves et al. (2020), Fernandes et al. (2020), Yang et al. (2021d), Swathi and Saritha (2021), Amerini et al. (2021a), Agarwal et al. (2021), Dal Cortivo et al. (2021), Schwarcz and Chellappa (2021), Kim et al. (2021b), Bailer et al. (2021), Tariq et al. (2021), Hu et al. (2021), Ahmed and Sonuç (2021), Laishram et al. (2021), Goebel et al. (2021), Han and Gevers (2021), Echizen et al. (2021), Fernando et al. (2021), Kubanek et al. (2021), Huang et al. (2020b), Baek et al. (2020), Zhang et al. (2019c), Pu et al. (2020), Pham et al. (2020), Wang and Dantcheva (2020), Mi et al. (2020), Ranjan et al. (2020), Frank et al. (2020), Yang and Lim (2020), Hashmi et al. (2020), Ranjith Kumar et al. (2020), Xuan et al. (2019), Guan et al. (2019), Kharbat et al. (2019), Bose and Aarabi (2019), Zhang et al. (2019a), Ward (2019)

References

Aboueldahab S, Freixo I (2021) App-generated evidence: a promising tool for international criminal justice? Int Crim Law Rev 21(3):505–533. https://doi.org/10.1163/15718123-bja10061

Agarwal S, Farid H (2021) Detecting deep-fake videos from aural and oral dynamics. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 981–989, https://doi.org/10.1109/CVPRW53098.2021.00109

Agarwal S, Farid H, El-Gaaly T, et al. (2020a) Detecting deep-fake videos from appearance and behavior. In: 2020 IEEE International Workshop on Information Forensics and Security, WIFS 2020, https://doi.org/10.1109/WIFS49906.2020.9360904

Agarwal S, Farid H, Fried O, et al. (2020b) Detecting deep-fake videos from phoneme-viseme mismatches. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 2814–2822, https://doi.org/10.1109/CVPRW50498.2020.00338

Agarwal S, Farid H, Gu Y, et al. (2019a) Protecting world leaders against deep fakes. pp 38–45, conference of 32nd IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2019 ; Conference Date: 16 June 2019 Through 20 June 2019; Conference Code:159074

Agarwal S, Farid H, Gu Y, et al. (2019b) Protecting world leaders against deep fakes. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 38–45

Agarwal H, Singh A, Rajeswari D (2021) Deepfake detection using svm. In: Proceedings of the 2nd International Conference on Electronics and Sustainable Communication Systems, ICESC 2021, pp 1245–1249, https://doi.org/10.1109/ICESC51422.2021.9532627

Agrawal R, Sharma D (2021) A survey on video-based fake news detection techniques. In: Proceedings of the 2021 8th International Conference on Computing for Sustainable Global Development, INDIACom 2021, pp 663–669, https://doi.org/10.1109/INDIACom51348.2021.00117

Ahmed S (2021) Fooled by the fakes: cognitive differences in perceived claim accuracy and sharing intention of non-political deepfakes. Personal Individ Differ. https://doi.org/10.1016/j.paid.2021.111074

Ahmed S (2021) Navigating the maze: deepfakes, cognitive ability, and social media news skepticism. New Media Soc. https://doi.org/10.1177/14614448211019198

Ahmed S (2021) Who inadvertently shares deepfakes? analyzing the role of political interest, cognitive ability, and social network size. Telemat Inform. https://doi.org/10.1016/j.tele.2020.101508

Ahmed S, Sonuç E (2021) Deepfake detection using rationale-augmented convolutional neural network. Appl Nanosci (Switzerland). https://doi.org/10.1007/s13204-021-02072-3

Ahmed M, Miah M, Bhowmik A, et al. (2021) Awareness to deepfake: A resistance mechanism to deepfake. In: 2021 International Congress of Advanced Technology and Engineering, ICOTEN 2021, https://doi.org/10.1109/ICOTEN52080.2021.9493549

Ajoy A, Mahindrakar C, Gowrish D, et al. (2021) Deepfake detection using a frame based approach involving cnn. In: Proceedings of the 3rd International Conference on Inventive Research in Computing Applications, ICIRCA 2021, pp 1329–1333, https://doi.org/10.1109/ICIRCA51532.2021.9544734

Alattar A, Sharma R, Scriven J (2020) A system for mitigating the problem of deepfake news videos using watermarking. In: Adnan M. A.M. GGNasir D. N.D. (ed) IS and T International Symposium on Electronic Imaging Science and Technology, https://doi.org/10.2352/ISSN.2470-1173.2020.4.MWSF-117

Albahar M, Almalki J (2019) Deepfakes: threats and countermeasures systematic review. J Theor Appl Inf Technol 97(22):3242–3250

Aliman NM, Kester L (2020) Malicious design in aivr, falsehood and cybersecurity-oriented immersive defenses. In: Proceedings - 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality, AIVR 2020, pp 130–137, https://doi.org/10.1109/AIVR50618.2020.00031

Amelin R, Channov S (2020) On the legal issues of face processing technologies. Commun Comput Inf Sci 1242:223–236. https://doi.org/10.1007/978-3-030-65218-0_17

Amerini I, Anagnostopoulos A, Maiano L et al (2021) Deep learning for multimedia forensics. Found Trends Comput Gr Vis 12(4):309–457. https://doi.org/10.1561/0600000096

Amerini I, Caldelli R (2020) Exploiting prediction error inconsistencies through lstm-based classifiers to detect deepfake videos. In: IH and MMSec 2020 - Proceedings of the 2020 ACM Workshop on Information Hiding and Multimedia Security, pp 97–102, https://doi.org/10.1145/3369412.3395070

Amerini I, Galteri L, Caldelli R, et al. (2019a) Deepfake video detection through optical flow based cnn. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE Computer Society, Los Alamitos, CA, USA, pp 1205–1207, https://doi.org/10.1109/ICCVW.2019.00152, https://doi.ieeecomputersociety.org/10.1109/ICCVW.2019.00152

Amerini I, Galteri L, Caldelli R, et al. (2019b) Deepfake video detection through optical flow based cnn. In: Proceedings - 2019 International Conference on Computer Vision Workshop, ICCVW 2019, pp 1205–1207, https://doi.org/10.1109/ICCVW.2019.00152

Aria M, Cuccurullo C (2017) bibliometrix: an r-tool for comprehensive science mapping analysis. J Informetr 11(4):959–975. https://doi.org/10.1016/j.joi.2017.08.007

Ayers D (2021) The limits of transactional identity: whiteness and embodiment in digital facial replacement. Convergence 27(4):1018–1037. https://doi.org/10.1177/13548565211027810

Baek JY, Yoo YS, Bae SH (2020) Generative adversarial ensemble learning for face forensics. IEEE Access 8:45,421-45,431. https://doi.org/10.1109/ACCESS.2020.2968612

Bailer W, Thallinger G, Backfried G, et al. (2021) Challenges for automatic detection of fake news related to migration : Invited paper. In: Proceedings - 2021 IEEE International Conference on Cognitive and Computational Aspects of Situation Management, CogSIMA 2021, pp 133–138, https://doi.org/10.1109/CogSIMA51574.2021.9475929

Biswas A, Bhattacharya D, Kumar K (2021) Deepfake detection using 3d-xception net with discrete fourier transformation. J Inf Syst Telecommun 9(35):161–168

Bode L (2021) Deepfaking keanu: youtube deepfakes, platform visual effects, and the complexity of reception. Convergence 27(4):919–934. https://doi.org/10.1177/13548565211030454

Bode L, Lees D, Golding D (2021) The digital face and deepfakes on screen. Convergence 27(4):849–854. https://doi.org/10.1177/13548565211034044

Bondi L, Daniele Cannas E, Bestagini P, et al. (2020) Training strategies and data augmentations in cnn-based deepfake video detection. In: 2020 IEEE International Workshop on Information Forensics and Security, WIFS 2020, https://doi.org/10.1109/WIFS49906.2020.9360901

Bonettini N, Bondi L, Cannas E, et al. (2020) Video face manipulation detection through ensemble of cnns. In: Proceedings - International Conference on Pattern Recognition, pp 5012–5019, https://doi.org/10.1109/ICPR48806.2021.9412711

Bonomi M, Pasquini C, Boato G (2021) Dynamic texture analysis for detecting fake faces in video sequences. J Vis Commun Image Represent. https://doi.org/10.1016/j.jvcir.2021.103239

Bore J (2020) Insider threat. Adv Sci Technol Secur Appl. https://doi.org/10.1007/978-3-030-35746-7_19

Bose A, Aarabi P (2019) Virtual fakes: Deepfakes for virtual reality. In: IEEE 21st International Workshop on Multimedia Signal Processing, MMSP 2019, https://doi.org/10.1109/MMSP.2019.8901744
Brooks C (2021) Popular discourse around deepfakes and the interdisciplinary challenge of fake video distribution. Cyberpsychol Behav Soc Netw 24(3):159–163. https://doi.org/10.1089/cyber.2020.0183
Burroughs S, Gokaraju B, Roy K, et al. (2020) Deepfakes detection in videos using feature engineering techniques in deep learning convolution neural network frameworks. In: Proceedings - Applied Imagery Pattern Recognition Workshop, https://doi.org/10.1109/AIPR50011.2020.9425347
Caldelli R, Galteri L, Amerini I et al (2021) Optical flow based cnn for detection of unlearnt deepfake manipulations. Pattern Recognit Lett 146:31–37. https://doi.org/10.1016/j.patrec.2021.03.005
Callon M, Courtial JP, Laville F (1991) Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. Scientometrics 22(1):155–205. https://doi.org/10.1007/BF02019280
Caporusso N (2021) Deepfakes for the good: a beneficial application of contentious artificial intelligence technology. Adv Intell Syst Comput. https://doi.org/10.1007/978-3-030-51328-3_33
Carlini N, Farid H (2020) Evading deepfake-image detectors with white- and black-box attacks. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 2804–2813, https://doi.org/10.1109/CVPRW50498.2020.00337
Carter M, Tsikerdekis M, Zeadally S (2021) Approaches for fake content detection: strengths and weaknesses to adversarial attacks. IEEE Internet Comput 25(2):73–83. https://doi.org/10.1109/MIC.2020.3032323
Castillo Camacho I, Wang K (2021) A comprehensive review of deep-learning-based methods for image forensics. J Imaging. https://doi.org/10.3390/jimaging7040069
Castillo Camacho I, Wang K (2021) A comprehensive review of deep-learning-based methods for image forensics. J Imaging. https://doi.org/10.3390/jimaging7040069
Chang X, Wu J, Yang T, et al. (2020) Deepfake face image detection based on improved vgg convolutional neural network. In: Fu J. SJ (eds). Chinese Control Conference, CCC, pp 7252–7256, https://doi.org/10.23919/CCC50068.2020.9189596
Chawla R (2019) Deepfakes: how a pervert shook the world. Int J Adv Res Dev 4(6):4–8
Chen B, Tan S (2021) Featuretransfer: Unsupervised domain adaptation for cross-domain deepfake detection. Security and Communication Networks. https://doi.org/10.1155/2021/9942754
Chen P, Liu J, Liang T, et al. (2020) Fsspotter: Spotting face-swapped video by spatial and temporal clues. In: Proceedings - IEEE International Conference on Multimedia and Expo, https://doi.org/10.1109/ICME46284.2020.9102914
Chesney RM, Citron DK (2018) Deep fakes: a looming challenge for privacy, democracy, and national security. Calif Law Rev 107:1753
Chesney B, Citron D (2019) Deep fakes: A looming challenge for privacy, democracy, and national security. Calif Law Rev 107(6):1753–1820. https://doi.org/10.15779/Z38RV0D15J
Chi H, Maduakor U, Alo R et al (2021) Integrating deepfake detection into cybersecurity curriculum. Adv Intell Syst Comput 1288:588–598. https://doi.org/10.1007/978-3-030-63128-4_45
Chintha A, Thai B, Sohrawardi S et al (2020) Recurrent convolutional structures for audio spoof and video deepfake detection. IEEE J Sel Top Signal Process 14(5):1024–1037. https://doi.org/10.1109/JSTSP.2020.2999185
Chintha A, Rao A, Sohrawardi S, et al. (2020a) Leveraging edges and optical flow on faces for deepfake detection. In: IJCB 2020 - IEEE/IAPR International Joint Conference on Biometrics, https://doi.org/10.1109/IJCB48548.2020.9304936
Cho M, Jeong Y (2017) Face recognition performance comparison between fake faces and live faces. Soft Comput 21(12):3429–3437. https://doi.org/10.1007/s00500-015-2019-4
Choraś M, Demestichas K, Giełczyk A et al (2021) Advanced machine learning techniques for fake news (online disinformation) detection: a systematic mapping study. Appl Soft Comput. https://doi.org/10.1016/j.asoc.2020.107050
Chowdhury S, Lubna J (2020) Review on deep fake: A looming technological threat. In: 2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020, https://doi.org/10.1109/ICCCNT49239.2020.9225630
Chugh K, Gupta P, Dhall A, et al. (2020) Not made for each other- audio-visual dissonance-based deepfake detection and localization. In: MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia, pp 439–447, https://doi.org/10.1145/3394171.3413700
Ciftci U, Demir I, Yin L (2020) How do the hearts of deep fakes beat? deep fake source detection via interpreting residuals with biological signals. In: IJCB 2020 - IEEE/IAPR International Joint Conference on Biometrics, https://doi.org/10.1109/IJCB48548.2020.9304909
Cobo M, López-Herrera A, Herrera-Viedma E et al (2011) Science mapping software tools: review, analysis, and cooperative study among tools. J Am Soc Inf Sci Technol 62(7):1382–1402. https://doi.org/10.1002/asi.21525
Colon M (2020) How can iowans effectively prevent the commercial misappropriation of their identities? why iowa needs a right of publicity statute. Iowa Law Rev 106(1):411–454
Cozzolino D, Poggi G, Verdoliva L (2019) Extracting camera-based fingerprints for video forensics. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 130–137
da Silva R (2021) Updating the authentication of digital evidence in the international criminal court. Int Crim Law Rev. https://doi.org/10.1163/15718123-bja10083
Dal Cortivo D, Mandelli S, Bestagini P et al (2021) Cnn-based multi-modal camera model identification on video sequences. J Imaging. https://doi.org/10.3390/jimaging7080135
D'Alessandra F, Sutherland K (2021) The promise and challenges of new actors and new technologies in international justice. J Int Crim Justice 19(1):9–34. https://doi.org/10.1093/jicj/mqab034
Dang H, Liu F, Stehouwer J, et al. (2020) On the detection of digital face manipulation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 5780–5789, https://doi.org/10.1109/CVPR42600.2020.00582
Dasilva J, Ayerdi K, Galdospin T (2021) Deepfakes on twitter: which actors control their spread? Media Commun 9(1):301–312. https://doi.org/10.17645/MAC.V9I1.3433
Davis M, Fors P (2020) Towards a typology of intentionally inaccurate representations of reality in media content. IFIP Adv Inf Commun Technol 590:291–304. https://doi.org/10.1007/978-3-030-62803-1_23
de Ruiter A (2021) The distinct wrong of deepfakes. Philos Technol. https://doi.org/10.1007/s13347-021-00459-2
de Seta G (2021) Huanlian, or changing faces: Deepfakes on chinese digital media platforms. Convergence 27(4):935–953. https://doi.org/10.1177/13548565211030185
Demir I, Ciftci U (2021) Where do deep fakes look? synthetic face detection via gaze tracking. In: S.N. S (eds) Eye Tracking Research and Applications Symposium (ETRA), https://doi.org/10.1145/3448017.3457387
Deshmukh A, Wankhade S (2021) Deepfake detection approaches using deep learning: a systematic review. Lect Notes Netw Syst 146:293–302. https://doi.org/10.1007/978-981-15-7421-4_27
Diakopoulos N, Johnson D (2021) Anticipating and addressing the ethical implications of deepfakes in the context of elec-

tions. New Media Soc 23(7):2072–2098. https://doi.org/10.1177/1461444820925811
Dobber T, Metoui N, Trilling D et al (2021) Do (microtargeted) deepfakes have real effects on political attitudes? Int J Press/Polit 26(1):69–91. https://doi.org/10.1177/1940161220944364
Dondero M (2021) Composition and decomposition in artistic portraits, scientific photography, and deep fake videos1. Lexia 2021(37–38):439–454. https://doi.org/10.4399/978882553853321
Du C, Duong L, Trung H, et al. (2020a) Efficient-frequency: A hybrid visual forensic framework for facial forgery detection. In: 2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020, pp 707–712, https://doi.org/10.1109/SSCI47803.2020.9308305
Du M, Pentyala S, Li Y, et al. (2020b) Towards generalizable deepfake detection with locality-aware autoencoder. In: International Conference on Information and Knowledge Management, Proceedings, pp 325–334, https://doi.org/10.1145/3340531.3411892
Echizen I, Babaguchi N, Yamagishi J et al (2021) Generation and detection of media clones. IEICE Trans Inf Syst E104D(1):12–23. https://doi.org/10.1587/transinf.2020MUI0002
El Rai M, Al Ahmad H, Gouda O, et al. (2020) Fighting deepfake by residual noise using convolutional neural networks. In: 2020 3rd International Conference on Signal Processing and Information Security, ICSPIS 2020, https://doi.org/10.1109/ICSPIS51252.2020.9340138
England P, Malvar H, Horvitz E, et al. (2021) Amp: Authentication of media via provenance. In: MMSys 2021 - Proceedings of the 2021 Multimedia Systems Conference, pp 109–121, https://doi.org/10.1145/3458305.3459599
Fagni T, Falchi F, Gambini M et al (2021) Tweepfake: about detecting deepfake tweets. PLoS ONE. https://doi.org/10.1371/journal.pone.0251415
Fallis D (2020) The epistemic threat of deepfakes. Philos Technol. https://doi.org/10.1007/s13347-020-00419-2
Farish K (2020) Do deepfakes pose a golden opportunity? considering whether english law should adopt california's publicity right in the age of the deepfake. J Intell Prop Law Pract 15(1):40–48. https://doi.org/10.1093/jiplp/jpz139
Fazheng W, Yanwei Y, Shuiyuan D, et al. (2021) Research on location of chinese handwritten signature based on efficientdet. In: 2021 IEEE 4th International Conference on Big Data and Artificial Intelligence, BDAI 2021, pp 192–198, https://doi.org/10.1109/BDAI52447.2021.9515222
Fei J, Xia Z, Yu P et al (2021) Exposing ai-generated videos with motion magnification. Multimed Tools Appl 80(20):30,789-30,802. https://doi.org/10.1007/s11042-020-09147-3
Feng D, Lu X, Lin X (2020) Deep detection for face manipulation. Commun Comput Inf Sci 1333:316–323. https://doi.org/10.1007/978-3-030-63823-8_37
Fernandes S, Jha S (2020) Adversarial attack on deepfake detection using rl based texture patches. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12535 LNCS:220–235. https://doi.org/10.1007/978-3-030-66415-2_14
Fernandes S, Raj S, Ewetz R, et al. (2020) Detecting deepfake videos using attribution-based confidence metric. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 1250–1259, https://doi.org/10.1109/CVPRW50498.2020.00162
Fernando T, Fookes C, Denman S et al (2021) Detection of fake and fraudulent faces via neural memory networks. IEEE Trans Inf Forensics Secur 16:1973–1988. https://doi.org/10.1109/TIFS.2020.3047768
Fletcher J (2018) Deepfakes, artificial intelligence, and some kind of dystopia: the new faces of online post-fact performance. Theatre J 70(4):455–471. https://doi.org/10.1353/tj.2018.0097
Frank J, Eisenhofer T, Schönherr L, et al. (2020) Leveraging frequency analysis for deep fake image recognition. In: Daume H. SA (eds) 37th International Conference on Machine Learning, ICML 2020, pp 3205–3216
Freeman L (2021) Weapons of war, tools of justice: using artificial intelligence to investigate international crimes. J Int Crim Justice 19(1):35–53. https://doi.org/10.1093/jicj/mqab013
Frick R, Zmudzinski S, Steinebach M (2021) Detecting deepfakes with haralick's texture properties. In: Adnan M. A.M. GGNasir D. N.D. (eds) IS and T International Symposium on Electronic Imaging Science and Technology, https://doi.org/10.2352/ISSN.2470-1173.2021.4.MWSF-271
Fung S, Lu X, Zhang C, et al. (2021) Deepfakeucl: Deepfake detection via unsupervised contrastive learning. In: Proceedings of the International Joint Conference on Neural Networks, https://doi.org/10.1109/IJCNN52387.2021.9534089
Gandhi A, Jain S (2020) Adversarial perturbations fool deepfake detectors. In: Proceedings of the International Joint Conference on Neural Networks, https://doi.org/10.1109/IJCNN48605.2020.9207034
Godulla A, Hoffmann C, Seibert D (2021) Dealing with deepfakes - an interdisciplinary examination of the state of research and implications for communication studies [der umgang mit deepfakes - eine interdisziplinäre untersuchung zum forschungsstand und implikationen für die kommunikationswissenschaft]. Stud Commun Media 10(1):73–96. https://doi.org/10.5771/2192-4007-2021-1-72
Goebel M, Nataraj L, Nanjundaswamy T, et al. (2021) Detection, attribution and localization of gan generated images. In: Adnan M. A.M. GGNasir D. N.D. (eds) IS and T International Symposium on Electronic Imaging Science and Technology, https://doi.org/10.2352/ISSN.2470-1173.2021.4.MWSF-276
Gong D, Goh O, Kumar Y et al (2020) Deepfake forensics, an ai-synthesized detection with deep convolutional generative adversarial networks. Int J Adv Trends Comput Sci Eng 9(3):2861–2870. https://doi.org/10.30534/ijatcse/2020/58932020
Gong D, Kumar Y, Ye O et al (2021) Deepfakenet, an efficient deepfake detection method. Int J Adv Comput Sci Appl 12(6):201–207. https://doi.org/10.14569/IJACSA.2021.0120622
Gosse C, Burkell J (2020) Politics and porn: how news media characterizes problems presented by deepfakes. Crit Stud Media Commun 37(5):497–511. https://doi.org/10.1080/15295036.2020.1832697
Guan H, Kozak M, Robertson E, et al. (2019) Mfc datasets: Large-scale benchmark datasets for media forensic challenge evaluation. In: Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2019, pp 63–72, https://doi.org/10.1109/WACVW.2019.00018
Guo Z, Yang G, Chen J et al (2021) Fake face detection via adaptive manipulation traces extraction network. Comput Vis Image Underst. https://doi.org/10.1016/j.cviu.2021.103170
Gupta P, Chugh K, Dhall A, et al. (2020) The eyes know it: Fakeet- an eye-tracking database to understand deepfake perception. In: ICMI 2020 - Proceedings of the 2020 International Conference on Multimodal Interaction, pp 519–527, https://doi.org/10.1145/3382507.3418857
Gu Y, Zhao X, Gong C, et al. (2021) Deepfake video detection using audio-visual consistency. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12617 LNCS:168–180. https://doi.org/10.1007/978-3-030-69449-4_13
Hancock J, Bailenson J (2021) The social impact of deepfakes. Cyberpsychol Behav Soc Netw 24(3):149–152. https://doi.org/10.1089/cyber.2021.29208.jth
Han J, Gevers T (2021) Mmd based discriminative learning for face forgery detection. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture

Notes in Bioinformatics) 12626 LNCS:121–136. https://doi.org/10.1007/978-3-030-69541-5_8
Hänska M (2021). Communication against domination: Ideas of justice from the printing press to algorithmic media. https://doi.org/10.4324/9780429280795
Hartmann K, Giles K (2020) The next generation of cyber-enabled information warfare. In: International Conference on Cyber Conflict, CYCON, pp 233–250, https://doi.org/10.23919/CyCon49761.2020.9131716
Hasan H, Salah K (2019) Combating deepfake videos using blockchain and smart contracts. IEEE Access 7:41,596-41,606. https://doi.org/10.1109/ACCESS.2019.2905689
Hashmi M, Ashish B, Keskar A et al (2020) An exploratory analysis on visual counterfeits using conv-lstm hybrid architecture. IEEE Access 8:101,293-101,308. https://doi.org/10.1109/ACCESS.2020.2998330
Hayward K, Maas M (2021) Artificial intelligence and crime: a primer for criminologists. Crime Media Cult 17(2):209–233. https://doi.org/10.1177/1741659020917434
Hazan S (2020) Deep fake and cultural truth - custodians of cultural heritage in the age of a digital reproduction. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12215 LNCS:65–80. https://doi.org/10.1007/978-3-030-50267-6_6
Hernandez-Ortega J, Tolosana R, Fierrez J, et al. (2021) Deepfakesonphys: Deepfakes detection based on heart rate estimation. In: CEUR Workshop Proceedings
Hewage C, Ekmekcioglu E (2020) Multimedia quality of experience (qoe): current status and future direction. Future Internet. https://doi.org/10.3390/FI12070121
Higgins JP, Thomas J, Chandler J et al (2019) Cochrane handbook for systematic reviews of interventions. John Wiley & Sons
Holliday C (2021) Rewriting the stars: surface tensions and gender troubles in the online media production of digital deepfakes. Convergence 27(4):899–918. https://doi.org/10.1177/13548565211029412
Hongmeng Z, Zhiqiang Z, Lei S, et al. (2020) A detection method for deepfake hard compressed videos based on super-resolution reconstruction using cnn. In: ACM International Conference Proceeding Series, pp 98–103, https://doi.org/10.1145/3409501.3409542
Hosier B, Stamm M (2020) Detecting video speed manipulation. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 2860–2869, https://doi.org/10.1109/CVPRW50498.2020.00343
Hosler B, Salvi D, Murray A, et al. (2021) Do deepfakes feel emotions? a semantic approach to detecting deepfakes via emotional inconsistencies. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 1013–1022, https://doi.org/10.1109/CVPRW53098.2021.00112
Houde S, Liao V, Martino J, et al. (2020) Business (mis)use cases of generative ai. In: Geyer W. SSMKhazaeni Y. (ed) CEUR Workshop Proceedings
Huang R, Fang F, Nguyen H, et al. (2020a) Security of facial forensics models against adversarial attacks. In: Proceedings - International Conference on Image Processing, ICIP, pp 2236–2240, https://doi.org/10.1109/ICIP40778.2020.9190678
Huang Y, Juefei-Xu F, Wang R, et al. (2020b) Fakepolisher: Making deepfakes more detection-evasive by shallow reconstruction. In: MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia, pp 1217–1226, https://doi.org/10.1145/3394171.3413732
Huber E, Pospisil B, Haidegger W (2021) Modus operandi in fake news : Invited paper. In: Proceedings - 2021 IEEE International Conference on Cognitive and Computational Aspects of Situation Management, CogSIMA 2021, pp 127–132, https://doi.org/10.1109/CogSIMA51574.2021.9475926
Hu S, Li Y, Lyu S (2021) Exposing gan-generated faces using inconsistent corneal specular highlights. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp 2500–2504, https://doi.org/10.1109/ICASSP39728.2021.9414582
Hussain S, Neekhara P, Jere M, et al. (2021) Adversarial deepfakes: Evaluating vulnerability of deepfake detectors to adversarial examples. In: Proceedings - 2021 IEEE Winter Conference on Applications of Computer Vision, WACV 2021, pp 3347–3356, https://doi.org/10.1109/WACV48630.2021.00339
Iacobucci S, De Cicco R, Michetti F et al (2021) Deepfakes unmasked: the effects of information priming and bullshit receptivity on deepfake recognition and sharing intention. Cyberpsychol Behav Soc Netw 24(3):194–202. https://doi.org/10.1089/cyber.2020.0149
Ismail A, Elpeltagy M, Zaki M et al (2021) A new deep learning-based methodology for video deepfake detection using xgboost. Sensors. https://doi.org/10.3390/s21165413
Ivanov N, Arzhskov A, Ivanenko V (2020) Combining deep learning and super-resolution algorithms for deep fake detection. In: S. S (ed) Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, EIConRus 2020, pp 326–328, https://doi.org/10.1109/EIConRus49466.2020.9039498
Jafar M, Ababneh M, Al-Zoube M, et al. (2020) Digital forensics and analysis of deepfake videos. In: 2020 11th International Conference on Information and Communication Systems, ICICS 2020, pp 53–58, https://doi.org/10.1109/ICICS49469.2020.239493
Javed A, Jalil Z, Zehra W et al (2021) A comprehensive survey on digital video forensics: Taxonomy, challenges, and future directions. Eng Appl Artif Intell. https://doi.org/10.1016/j.engappai.2021.104456
Jeong D (2020) Artificial intelligence security threat, crime, and forensics: taxonomy and open issues. IEEE Access 8:184,560-184,574. https://doi.org/10.1109/ACCESS.2020.3029280
Jeong Y, Choi J, Kim D, et al. (2021) Dofnet: Depth of field difference learning for detecting image forgery. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12627 LNCS:83–100. https://doi.org/10.1007/978-3-030-69544-6_6
Jiang L, Li R, Wu W, et al. (2020) Deeperforensics-1.0: A large-scale dataset for real-world face forgery detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 2886–2895, https://doi.org/10.1109/CVPR42600.2020.00296
Jiang J, Wang B, Li B, et al. (2021) Practical face swapping detection based on identity spatial constraints. In: 2021 IEEE International Joint Conference on Biometrics, IJCB 2021, https://doi.org/10.1109/IJCB52358.2021.9484396
Jin X, Ye D, Chen C (2021) Countering spoof: towards detecting deepfake with multidimensional biological signals. Secur Commun Netw. https://doi.org/10.1155/2021/6626974
Johnson J (2021) ‘catalytic nuclear war’ in the age of artificial intelligence & autonomy: emerging military technology and escalation risk between nuclear-armed states. J Strateg Stud. https://doi.org/10.1080/01402390.2020.1867541
Johnson D, Diakopoulos N (2021) What to do about deepfakes. Commun ACM 64(3):33–35. https://doi.org/10.1145/3447255
Jongman B (2020) Recent online resources for the analysis of terrorism and related subjects. Perspect Terror 14(1):155–190
José F, García-Ull GU (2021) Deepfakes: the next challenge in fake news detection. Analisi 64:103–120. https://doi.org/10.5565/REV/ANALISI.3378
Jung T, Kim S, Kim K (2020) Deepvision: deepfakes detection using human eye blinking pattern. IEEE Access 8:83,144-83,154. https://doi.org/10.1109/ACCESS.2020.2988660

Kang M, Park J (2020) Contragan: Contrastive learning for conditional image generation. In: Advances in Neural Information Processing Systems

Karandikar A, Deshpande V, Singh S et al (2020) Deepfake video detection using convolutional neural network. Int J Adv Trends Comput Sci Eng 9(2):1311–1315. https://doi.org/10.30534/ijatcse/2020/62922020

Karasavva V, Noorbhai A (2021) The real threat of deepfake pornography: a review of canadian policy. Cyberpsychol Behav Soc Netw 24(3):203–209. https://doi.org/10.1089/cyber.2020.0272

Katarya R, Lal A (2020) A study on combating emerging threat of deepfake weaponization. In: Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020, pp 485–490, https://doi.org/10.1109/I-SMAC49090.2020.9243588

Kaur S, Kumar P, Kumaraguru P (2020) Deepfakes: temporal sequential analysis to detect face-swapped video clips using convolutional long short-term memory. J Electron Imaging. https://doi.org/10.1117/1.JEI.29.3.033013

Kawa P, Syga P (2021) Verify it yourself: A note on activation functions' influence on fast deepfake detection. In: di Vimercati S.De.C. SP (ed) Proceedings of the 18th International Conference on Security and Cryptography, SECRYPT 2021, pp 779–784, https://doi.org/10.5220/0010581707790784

Kaye B, Johnson T (2020) Appsolutely trustworthy? perceptions of trust and bias in mobile apps. Atl J Commun 28(4):257–271. https://doi.org/10.1080/15456870.2020.1720023

Khalid H, Woo S (2020) Oc-fakedect: Classifying deepfakes using one-class variational autoencoder. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 2794–2803, https://doi.org/10.1109/CVPRW50498.2020.00336

Khalil S, Youssef S, Saleh S (2021) Article icaps-dfake: an integrated capsule-based model for deepfake image and video detection. Future Internet. https://doi.org/10.3390/fi13040093

Khalil H, Maged S (2021) Deepfakes creation and detection using deep learning. In: 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference, MIUCC 2021, pp 24–27, https://doi.org/10.1109/MIUCC52538.2021.9447642

Kharbat F, Elamsy T, Mahmoud A, et al. (2019) Image feature detectors for deepfake video detection. In: Proceedings of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA, https://doi.org/10.1109/AICCSA47632.2019.9035360

Khodabakhsh A, Loiselle H (2020) Action-independent generalized behavioral identity descriptors for look-alike recognition in videos. In: BIOSIG 2020 - Proceedings of the 19th International Conference of the Biometrics Special Interest Group

Khormali A, Yuan JS (2021) Add: Attention-based deepfake detection approach. Big Data Cognitive Comput. https://doi.org/10.3390/bdcc5040049

Ki Chan C, Kumar V, Delaney S, et al. (2020) Combating deepfakes: Multi-lstm and blockchain as proof of authenticity for digital media. In: 2020 IEEE / ITU International Conference on Artificial Intelligence for Good, AI4G 2020, pp 55–62, https://doi.org/10.1109/AI4G50087.2020.9311067

Kietzmann J, Lee L, McCarthy I et al (2020) Deepfakes: trick or treat? Bus Horiz 63(2):135–146. https://doi.org/10.1016/j.bushor.2019.11.006

Kietzmann J, Mills A, Plangger K (2021) Deepfakes: perspectives on the future reality of advertising and branding. Int J Advert 40(3):473–485. https://doi.org/10.1080/02650487.2020.1834211

Kikerpill K (2020) Choose your stars and studs: the rise of deepfake designer porn. Porn Studies 7(4):352–356. https://doi.org/10.1080/23268743.2020.1765851

Kim KS, Sin SC, Yoo-Lee E (2021) Use and evaluation of information from social media: a longitudinal cohort study. Libr Inf Sci Res. https://doi.org/10.1016/j.lisr.2021.101104

Kim M, Tariq S, Woo S (2021b) Fretal: Generalizing deepfake detection using knowledge distillation and representation learning. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 1001–1012, https://doi.org/10.1109/CVPRW53098.2021.00111

Kohli A, Gupta A (2021) Detecting deepfake, faceswap and face2face facial forgeries using frequency cnn. Multimed Tools Appl 80(12):18,461-18,478. https://doi.org/10.1007/s11042-020-10420-8

Korshunov P, Marcel S (2019) Vulnerability assessment and detection of deepfake videos. In: 2019 International Conference on Biometrics, ICB 2019, https://doi.org/10.1109/ICB45273.2019.8987375

Korshunov P, Marcel S (2021) Subjective and objective evaluation of deepfake videos. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp 2510–2514, https://doi.org/10.1109/ICASSP39728.2021.9414258

Kozyreva A, Lewandowsky S, Hertwig R (2020) Citizens versus the internet: confronting digital challenges with cognitive tools. Psychol Sci Public Interest 21(3):103–156. https://doi.org/10.1177/1529100620946707

Kuang Z, Guo Z, Fang J et al (2021) Unnoticeable synthetic face replacement for image privacy protection. Neurocomputing 457:322–333. https://doi.org/10.1016/j.neucom.2021.06.061

Kubanek M, Bartłomiejczyk K, Bobulski J (2021) Detection of artificial images and changes in real images using convolutional neural networks. Advances in Intelligent Systems and Computing 1267 AISC:197–207. https://doi.org/10.1007/978-3-030-57805-3_19

Kukanov I, Karttunen J, Sillanpaa H, et al. (2020) Cost sensitive optimization of deepfake detector. In: 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2020 - Proceedings, pp 1300–1303

Kwok A, Koh S (2021) Deepfake: a social construction of technology perspective. Curr Issue Tour 24(13):1798–1802. https://doi.org/10.1080/13683500.2020.1738357

Lai X, Patrick Rau PL (2021) Has facial recognition technology been misused? a user perception model of facial recognition scenarios. Comput Hum Behav. https://doi.org/10.1016/j.chb.2021.106894

Laishram L, Rahman M, Jung S (2021) Challenges and applications of face deepfake. Commun Comput Inf Sci 1405:131–156. https://doi.org/10.1007/978-3-030-81638-4_11

Lees D, Bashford-Rogers T, Keppel-Palmer M (2021) The digital resurrection of margaret thatcher: creative, technological and legal dilemmas in the use of deepfakes in screen drama. Convergence 27(4):954–973. https://doi.org/10.1177/13548565211030452

Lewis J, Toubal I, Chen H, et al. (2020) Deepfake video detection based on spatial, spectral, and temporal inconsistencies using multimodal deep learning. In: Proceedings - Applied Imagery Pattern Recognition Workshop, https://doi.org/10.1109/AIPR50011.2020.9425167

Li H, Li B, Tan S et al (2020) Identification of deep network generated images using disparities in color components. Signal Process. https://doi.org/10.1016/j.sigpro.2020.107616

Liang T, Chen P, Zhou G, et al. (2020) Sdhf: Spotting deepfakes with hierarchical features. In: Alamaniotis M. PS (ed) Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI, pp 675–680, https://doi.org/10.1109/ICTAI50040.2020.00108

Liang J, Deng W (2021) Identifying rhythmic patterns for face forgery detection and categorization. In: 2021 IEEE International Joint Conference on Biometrics, IJCB 2021, https://doi.org/10.1109/IJCB52358.2021.9484400

Li L, Bao J, Zhang T, et al. (2020b) Face x-ray for more general face forgery detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 5000–5009, https://doi.org/10.1109/CVPR42600.2020.00505

Li M, Liu B, Hu Y, et al. (2020c) Exposing deepfake videos by tracking eye movements. In: Proceedings - International Conference on Pattern Recognition, pp 5184–5189, https://doi.org/10.1109/ICPR48806.2021.9413139

Li M, Liu B, Hu Y, et al. (2021a) Deepfake detection using robust spatial and temporal features from facial landmarks. In: Proceedings - 9th International Workshop on Biometrics and Forensics, IWBF 2021, https://doi.org/10.1109/IWBF50991.2021.9465076

Li Y, Lyu S (2021) Obstructing deepfakes by disrupting face detection and facial landmarks extraction. Advances in Computer Vision and Pattern Recognition pp 247–267. https://doi.org/10.1007/978-3-030-74697-1_12

Ling H, Huang J, Zhao C, et al. (2021) Learning diverse local patterns for deepfake detection with image-level supervision. In: Proceedings of the International Joint Conference on Neural Networks, https://doi.org/10.1109/IJCNN52387.2021.9533912

Li W, Wang Q, Wang R, et al. (2021b) Exposing deepfakes via localizing the manipulated artifacts. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12919 LNCS:3–20. https://doi.org/10.1007/978-3-030-88052-1_1

Li Y, Yang X, Sun P, et al. (2020d) Celeb-df: A large-scale challenging dataset for deepfake forensics. pp 3204–3213, https://doi.org/10.1109/CVPR42600.2020.00327, conference of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020 ; Conference Date: 14 June 2020 Through 19 June 2020; Conference Code:162261

Li Y, Yang X, Sun P, et al. (2020e) Celeb-df: A large-scale challenging dataset for deepfake forensics. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 3204–3213, https://doi.org/10.1109/CVPR42600.2020.00327

Lomnitz M, Hampel-Arias Z, Sandesara V, et al. (2020) Multimodal approach for deepfake detection. In: Proceedings - Applied Imagery Pattern Recognition Workshop, https://doi.org/10.1109/AIPR50011.2020.9425192

Lu Y, Liu Y, Fei J et al (2021) Channel-wise spatiotemporal aggregation technology for face video forensics. Secur Commun Netw. https://doi.org/10.1155/2021/5524930

Luo Y, Ye F, Weng B et al (2021) A novel defensive strategy for facial manipulation detection combining bilateral filtering and joint adversarial training. Secur Commun Netw. https://doi.org/10.1155/2021/4280328

Lv L (2021) Smart watermark to defend against deepfake image manipulation. In: 2021 IEEE 6th International Conference on Computer and Communication Systems, ICCCS 2021, pp 380–384, https://doi.org/10.1109/ICCCS52626.2021.9449287

Lyu S (2020) Deepfake detection: Current challenges and next steps. In: 2020 IEEE International Conference on Multimedia and Expo Workshops, ICMEW 2020, https://doi.org/10.1109/ICMEW46912.2020.9105991

Maddocks S (2020) 'a deepfake porn plot intended to silence me': exploring continuities between pornographic and 'political' deep fakes. Porn Stud 7(4):415–423. https://doi.org/10.1080/23268743.2020.1757499

Maksutov A, Morozov V, Lavrenov A, et al. (2020) Methods of deepfake detection based on machine learning. In: S. S (ed) Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, EIConRus 2020, pp 408–411, https://doi.org/10.1109/EIConRus49466.2020.9039057

Malolan B, Parekh A, Kazi F (2020) Explainable deep-fake detection using visual interpretability methods. In: Proceedings - 3rd International Conference on Information and Computer Technologies, ICICT 2020, pp 289–293, https://doi.org/10.1109/ICICT50521.2020.00051

Maras MH, Alexandrou A (2019) Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. Int J Evid Proof 23(3):255–262

Marcon F, Pasquini C, Boato G (2021) Detection of manipulated face videos over social networks: a large-scale study. J Imaging. https://doi.org/10.3390/jimaging7100193

Masi I, Killekar A, Mascarenhas R, et al. (2020) Two-branch recurrent network for isolating deepfakes in videos. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12352 LNCS:667–684. https://doi.org/10.1007/978-3-030-58571-6_39

Masood M, Nawaz M, Javed A, et al. (2021) Classification of deepfake videos using pre-trained convolutional neural networks. In: 2021 International Conference on Digital Futures and Transformative Technologies, ICoDT2 2021, https://doi.org/10.1109/ICoDT252288.2021.9441519

Matern F, Riess C, Stamminger M (2019a) Exploiting visual artifacts to expose deepfakes and face manipulations. pp 83–92, https://doi.org/10.1109/WACVW.2019.00020, conference of 19th IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2019 ; Conference Date: 7 January 2019 Through 11 January 2019; Conference Code:145024

Matern F, Riess C, Stamminger M (2019b) Exploiting visual artifacts to expose deepfakes and face manipulations. In: Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2019, pp 83–92, https://doi.org/10.1109/WACVW.2019.00020

Mcglynn C, Johnson K (2021) Cyberflashing: Recognising Harms, Reforming Laws

Medoff N, B.K. K (2021) Interconnected by the internet. https://doi.org/10.4324/9781003020721-5

Megahed A, Han Q (2020) Face2face manipulation detection based on histogram of oriented gradients. In: Proceedings - 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2020, pp 1260–1267, https://doi.org/10.1109/TrustCom50675.2020.00169

Megías D, Kuribayashi M, Rosales A, et al. (2021) Dissimilar: Towards fake news detection using information hiding, signal processing and machine learning. In: ACM International Conference Proceeding Series, https://doi.org/10.1145/3465481.3470088

Meskys E, Liaudanskas A, Kalpokiene J et al (2020) Regulating deep fakes: legal and ethical considerations. J Intell Proper Law Pract 15(1):24–31. https://doi.org/10.1093/jiplp/jpz167

Mi Z, Jiang X, Sun T et al (2020) Gan-generated image detection with self-attention mechanism against gan generator defect. IEEE J Sel Top Sign Proces 14(5):969–981. https://doi.org/10.1109/JSTSP.2020.2994523

Mihailova M (2021) To dally with dalí: deepfake (inter)faces in the art museum. Convergence 27(4):882–898. https://doi.org/10.1177/13548565211029401

Mirsky Y, Lee W (2021) The creation and detection of deepfakes. ACM Comput Surv 54(1):7

Mitra A, Mohanty S, Corcoran P, et al. (2020) A novel machine learning based method for deepfake video detection in social media. In: Proceedings - 2020 6th IEEE International Symposium on Smart Electronic Systems, iSES 2020, pp 91–96, https://doi.org/10.1109/iSES50453.2020.00031

Mittal T, Bhattacharya U, Chandra R, et al. (2020b) Emotions don't lie: An audio-visual deepfake detection method using affective cues. In: MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia, pp 2823–2832, https://doi.org/10.1145/3394171.3413570

Mittal H, Saraswat M, Bansal J, et al. (2020a) Fake-face image classification using improved quantum-inspired evolutionary-based feature selection method. In: 2020 IEEE Symposium Series on

Computational Intelligence, SSCI 2020, pp 989–995, https://doi.org/10.1109/SSCI47803.2020.9308337

Montserrat D, Hao H, Yarlagadda S, et al. (2020) Deepfakes detection with automatic face weighting. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 2851–2859, https://doi.org/10.1109/CVPRW50498.2020.00342

Murphy G, Flynn E (2021) Deepfake false memories. Memory. https://doi.org/10.1080/09658211.2021.1919715

Nasar B, Sajini T, Lason E (2020) Deepfake detection in media files - audios, images and videos. In: 2020 IEEE Recent Advances in Intelligent Computational Systems, RAICS 2020, pp 74–79, https://doi.org/10.1109/RAICS51191.2020.9332516

Neves J, Tolosana R, Vera-Rodriguez R et al (2020) Ganprintr: improved fakes and evaluation of the state of the art in face manipulation detection. IEEE J Sel Top Sign Proces 14(5):1038–1048. https://doi.org/10.1109/JSTSP.2020.3007250

Nguyen X, Tran T, Le V et al (2021) Learning spatio-temporal features to detect manipulated facial videos created by the deepfake techniques. Forensic Sci Int Dig Investig. https://doi.org/10.1016/j.fsidi.2021.301108

Nguyen H, Derakhshani R (2020) Eyebrow recognition for identifying deepfake videos. In: BIOSIG 2020 - Proceedings of the 19th International Conference of the Biometrics Special Interest Group

Nygren T, Guath M, Axelsson CA et al (2021) Combatting visual fake news with a professional fact-checking tool in education in france, romania, spain and sweden. Information (Switzerland). https://doi.org/10.3390/info12050201

O'Donnell N (2021) Have we no decency? section 230 and the liability of social media companies for deepfake videos. Univ Ill Law Rev 2021(3):701–740

Pan Z, Ren Y, Zhang X (2021) Low-complexity fake face detection based on forensic similarity. Multimed Syst 27(3):353–361. https://doi.org/10.1007/s00530-021-00756-y

Pan D, Sun L, Wang R, et al. (2020) Deepfake detection through deep learning. In: Proceedings - 2020 IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, BDCAT 2020, pp 134–143, https://doi.org/10.1109/BDCAT50828.2020.00001

Pantserev K (2020a) Deepfakes as the new challenge of national and international psychological security. In: F. M (ed) Proceedings of the European Conference on the Impact of Artificial Intelligence and Robotics, ECIAIR 2020, pp 93–99, https://doi.org/10.34190/EAIR.20.003

Pantserev K (2020) The malicious use of ai-based deepfake technology as the new threat to psychological security and political stability. Adv Sci Technol Secur Appl. https://doi.org/10.1007/978-3-030-35746-7_3

Partadiredja R, Serrano C, Ljubenkov D (2020) Ai or human: The socio-ethical implications of ai-generated media content. In: I. W (ed) 13th CMI Conference on Cybersecurity and Privacy - Digital Transformation - Potentials and Challenges, CMI 2020, https://doi.org/10.1109/CMI51275.2020.9322673

Pashentsev E (2020) Malicious use of deepfakes and political stability. In: F. M (ed) Proceedings of the European Conference on the Impact of Artificial Intelligence and Robotics, ECIAIR 2020, pp 100–107, https://doi.org/10.34190/EAIR.20.025

Patil U, Chouragade P (2021) Deepfake video authentication based on blockchain. In: Proceedings of the 2nd International Conference on Electronics and Sustainable Communication Systems, ICESC 2021, pp 1110–1113, https://doi.org/10.1109/ICESC51422.2021.9532725

Patil U, Chouragade P, Ambhore P (2021) An effective blockchain technique to resist against deepfake videos. In: Proceedings of the 3rd International Conference on Inventive Research in Computing Applications, ICIRCA 2021, pp 1646–1652, https://doi.org/10.1109/ICIRCA51532.2021.9544854

Pavis M (2021) Rebalancing our regulatory response to deepfakes with performers' rights. Convergence 27(4):974–998. https://doi.org/10.1177/13548565211033418

Pavlíková M, Šenkýřová B, Drmola J (2021) Propaganda and disinformation go online. Political Campaigning and Communication pp 43–74. https://doi.org/10.1007/978-3-030-58624-9_2

Peng C, Zhang W, Liu D, et al. (2020) Temporal consistency based deep face forgery detection network. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12488 LNCS:55–63. https://doi.org/10.1007/978-3-030-62463-7_6

Perot E, Mostert F (2020) Fake it till you make it: an examination of the us and english approaches to persona protection as applied to deepfakes on social media. J Intell Proper Law Pract 15(1):32–39. https://doi.org/10.1093/jiplp/jpz164

Pertsch K, Rybkin O, Ebert F, et al. (2020) Long-horizon visual planning with goal-conditioned hierarchical predictors. In: Advances in Neural Information Processing Systems

Pham KL, Dang KM, Tang LP, et al. (2020) Gan generated portraits detection using modified vgg-16 and efficientnet. In: Bao V.N.Q. TTVan Vu N. (ed) Proceedings - 2020 7th NAFOSTED Conference on Information and Computer Science, NICS 2020, pp 344–349, https://doi.org/10.1109/NICS51282.2020.9335837

Pokroy A, Egorov A (2021) Efficientnets for deepfake detection: Comparison of pretrained models. In: S. S (ed) Proceedings of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering, ElConRus 2021, pp 598–600, https://doi.org/10.1109/ElConRus51938.2021.9396092

Prajwal K, Mukhopadhyay R, Namboodiri V, et al. (2020) A lip sync expert is all you need for speech to lip generation in the wild. In: MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia, pp 484–492, https://doi.org/10.1145/3394171.3413532

Pu J, Mangaokar N, Kelly L et al (2021) Deepfake videos in the wild: analysis and detection. Proceedings of the Web Conference 2021:981–992

Pu J, Mangaokar N, Kelly L, et al. (2021b) Deepfake videos in the wild: Analysis and detection. In: The Web Conference 2021 - Proceedings of the World Wide Web Conference, WWW 2021, pp 981–992, https://doi.org/10.1145/3442381.3449978

Pu J, Mangaokar N, Wang B, et al. (2020) Noisescope: Detecting deepfake images in a blind setting. In: ACM International Conference Proceeding Series, pp 913–927, https://doi.org/10.1145/3427228.3427285

Ramadhani K, Munir R (2020) A comparative study of deepfake video detection method. In: 2020 3rd International Conference on Information and Communications Technology, ICOIACT 2020, pp 394–399, https://doi.org/10.1109/ICOIACT50329.2020.9331963

Ranjan P, Patil S, Kazi F (2020) Improved generalizability of deepfakes detection using transfer learning based cnn framework. In: Proceedings - 3rd International Conference on Information and Computer Technologies, ICICT 2020, pp 86–90, https://doi.org/10.1109/ICICT50521.2020.00021

Ranjith Kumar M, Prabhu A, Asthana S et al (2020) Denet: a deepfake visual media detection network. J Adv Res Dyn Control Syst 12(2):792–799. https://doi.org/10.5373/JARDCS/V12I2/S20201098

Rao S, Verma A, Bhatia T (2021) A review on social spam detection: challenges, open issues, and future directions. Expert Syst Appl. https://doi.org/10.1016/j.eswa.2021.115742

Ross A, Banerjee S, Chowdhury A (2020) Security in smart cities: a brief review of digital forensic schemes for biometric data. Pattern Recogn Lett 138:346–354. https://doi.org/10.1016/j.patrec.2020.07.009

Rossler A, Cozzolino D, Verdoliva L, et al. (2019) Faceforensics++: Learning to detect manipulated facial images. In: Proceedings of the IEEE International Conference on Computer Vision, pp 1–11, https://doi.org/10.1109/ICCV.2019.00009

Ru Y, Zhou W, Liu Y, et al. (2021) Bita-net: Bi-temporal attention network for facial video forgery detection. In: 2021 IEEE International Joint Conference on Biometrics, IJCB 2021, https://doi.org/10.1109/IJCB52358.2021.9484408

Samek W, Montavon G, Lapuschkin S et al (2021) Explaining deep neural networks and beyond: a review of methods and applications. Proc IEEE 109(3):247–278. https://doi.org/10.1109/JPROC.2021.3060483

Sanghvi B, Shelar H, Pandey M, et al. (2021) Detection of machine generated multimedia elements using deep learning. In: Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021, pp 1238–1243, https://doi.org/10.1109/ICCMC51019.2021.9418008

Sankaranarayanan A, Groh M, Picard R, et al. (2021) The presidential deepfakes dataset. In: Aimeur E. HHDiaz Ferreyra N.E. (ed) CEUR Workshop Proceedings, pp 57–72

Schwarcz S, Chellappa R (2021) Finding facial forgery artifacts with parts-based detectors. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 933–942, https://doi.org/10.1109/CVPRW53098.2021.00104

Şener O (2020) New literacies for disinformation and manipulation through digital sound and video

Šepec M, Lango M (2020) Virtual revenge pornography as a new online threat to sexual integrity. Balk Soc Sci Rev 15(15):117–134

Shahar H, Hel-Or H (2020) Fake video detection using facial color. In: Final Program and Proceedings - IS and T/SID Color Imaging Conference, pp 175–180, https://doi.org/10.2352/issn.2169-2629.2020.28.27

Shah Y, Shah P, Patel M, et al. (2020) Deep learning model-based multimedia forgery detection. In: Proceedings of the 4th International Conference on IoT in Social, Mobile, Analytics and Cloud, ISMAC 2020, pp 564–572, https://doi.org/10.1109/I-SMAC49090.2020.9243530

Shang Z, Xie H, Zha Z et al (2021) Prrnet: pixel-region relation network for face forgery detection. Pattern Recogn. https://doi.org/10.1016/j.patcog.2021.107950

Shelke N, Kasana S (2021) A comprehensive survey on passive techniques for digital video forgery detection. Multimed Tools Appl 80(4):6247–6310. https://doi.org/10.1007/s11042-020-09974-4

Siegel D, Kraetzer C, Seidlitz S et al (2021) Media forensics considerations on deepfake detection with hand-crafted features. J Imaging. https://doi.org/10.3390/jimaging7070108

Singh R, Sarda P, Aggarwal S, et al. (2021) Demystifying deepfakes using deep learning. In: Proceedings - 5th International Conference on Computing Methodologies and Communication, ICCMC 2021, pp 1290–1298, https://doi.org/10.1109/ICCMC51019.2021.9418477

Sohrawardi S, Seng S, Chintha A, et al. (2019) Poster: Towards robust open-world detection of deepfakes. In: Proceedings of the ACM Conference on Computer and Communications Security, pp 2613–2615, https://doi.org/10.1145/3319535.3363269

Su Y, Xia H, Liang Q et al (2021) Exposing deepfake videos using attention based convolutional lstm network. Neural Process Lett. https://doi.org/10.1007/s11063-021-10588-6

Sun P, Yan Z, Shen Z, et al. (2021) Deepfakes detection based on multi scale fusion. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12878 LNCS:346–353. https://doi.org/10.1007/978-3-030-86608-2_38

Suratkar S, Johnson E, Variyambat K, et al. (2020a) Employing transfer-learning based cnn architectures to enhance the generalizability of deepfake detection. In: 2020 11th International Conference on Computing, Communication and Networking Technologies, ICCCNT 2020, https://doi.org/10.1109/ICCCNT49239.2020.9225400

Suratkar S, Kazi F, Sakhalkar M, et al. (2020b) Exposing deepfakes using convolutional neural networks and transfer learning approaches. In: 2020 IEEE 17th India Council International Conference, INDICON 2020, https://doi.org/10.1109/INDICON49873.2020.9342252

Swathi P, Saritha S (2021) Deepfake creation and detection:a survey. In: Proceedings of the 3rd International Conference on Inventive Research in Computing Applications, ICIRCA 2021, pp 584–588, https://doi.org/10.1109/ICIRCA51532.2021.9544522

Sybrandt J, Safro I (2021) Cbag: conditional biomedical abstract generation. PLoS ONE. https://doi.org/10.1371/journal.pone.0253905

Tahir R, Batool B (2021) Seeing is believing: Exploring perceptual diferences in deepfake videos. In: Conference on Human Factors in Computing Systems - Proceedings, https://doi.org/10.1145/3411764.3445699

Tarasiou M, Zafeiriou S (2020) Extracting deep local features to detect manipulated images of human faces. In: Proceedings - International Conference on Image Processing, ICIP, pp 1821–1825, https://doi.org/10.1109/ICIP40778.2020.9190714

Tariq S, Lee S, Woo S (2021) One detector to rule them all: Towards a general deepfake attack detection framework. In: The Web Conference 2021 - Proceedings of the World Wide Web Conference, WWW 2021, pp 3625–3637, https://doi.org/10.1145/3442381.3449809

Tesfagergish S, Damaševičius R, Kapočiūtė-Dzikienė J (2021) Deep fake recognition in tweets using text augmentation, word embeddings and deep learning. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12954 LNCS:523–538. https://doi.org/10.1007/978-3-030-86979-3_37

Thaw N, July T, Wai A et al (2021) How are deepfake videos detected? an initial user study. Commun Comput Inf Sci 1419:631–636. https://doi.org/10.1007/978-3-030-78635-9_80

Tjon E, Moh M, Moh TS (2021) Eff-ynet: A dual task network for deepfake detection and segmentation. In: Lee S. IRChoo H. (ed) Proceedings of the 2021 15th International Conference on Ubiquitous Information Management and Communication, IMCOM 2021, https://doi.org/10.1109/IMCOM51814.2021.9377373

Tolosana R, Vera-Rodriguez R, Fierrez J et al (2020) Deepfakes and beyond: a survey of face manipulation and fake detection. Inf Fusion 64:131–148. https://doi.org/10.1016/j.inffus.2020.06.014

Tolosana R, Romero-Tapiador S, Fierrez J, et al. (2021) Deepfakes evolution: Analysis of facial regions and fake detection performance. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12665 LNCS:442–456. https://doi.org/10.1007/978-3-030-68821-9_38

Tran VN, Lee SH, Le HS et al (2021) High performance deepfake video detection on cnn-based with attention target-specific regions and manual distillation extraction. Appl Sci (Switzerland). https://doi.org/10.3390/app11167678

Trinh L, Tsang M, Rambhatla S, et al. (2021) Interpretable and trustworthy deepfake detection via dynamic prototypes. In: Proceedings - 2021 IEEE Winter Conference on Applications of Computer Vision, WACV 2021, pp 1972–1982, https://doi.org/10.1109/WACV48630.2021.00202

Tu Y, Liu Y, Li X (2021) Deepfake video detection by using convolutional gated recurrent unit. In: ACM International Conference Proceeding Series, pp 356–360, https://doi.org/10.1145/3457682.3457736

Tulk Jesso S, Kennedy W, Wiese E (2020) Behavioral cues of humanness in complex environments: How people engage with human and artificially intelligent agents in a multiplayer videogame. Fron-

tiers in Robotics and AI 7. https://doi.org/10.3389/frobt.2020.531805
Tursman E, George M, Kamara S, et al. (2020) Towards untrusted social video verification to combat deepfakes via face geometry consistency. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 2784–2793, https://doi.org/10.1109/CVPRW50498.2020.00335
Valenzuela A, Segura C, Diego F, et al. (2021) Expression transfer using flow-based generative models. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 1023–1031, https://doi.org/10.1109/CVPRW53098.2021.00113
Verdoliva L (2020) Media forensics and deepfakes: an overview. IEEE J Sel Top Sign Proces 14(5):910–932. https://doi.org/10.1109/JSTSP.2020.3002101
Vizoso A, Vaz-Álvarez M, López-García X (2021) Fighting deepfakes: media and internet giants' converging and diverging strategies against hi-tech misinformation. Media Commun 9(1):291–300. https://doi.org/10.17645/MAC.V9I1.3494
Wahl-Jorgensen K, Carlson M (2021) Conjecturing fearful futures: journalistic discourses on deepfakes. J Pract 15(6):803–820. https://doi.org/10.1080/17512786.2021.1908838
Wang Y, Dantcheva A (2020) A video is worth more than 1000 lies. comparing 3dcnn approaches for detecting deepfakes. In: Struc V. GFF (ed) Proceedings - 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2020, pp 515–519, https://doi.org/10.1109/FG47880.2020.00089
Wang R, Juefei-Xu F, Huang Y, et al. (2020a) Deepsonar: Towards effective and robust detection of ai-synthesized fake voices. In: MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia, pp 1207–1216, https://doi.org/10.1145/3394171.3413716
Wang R, Juefei-Xu F, Ma L, et al. (2020b) Fakespotter: A simple yet robust baseline for spotting ai-synthesized fake faces. In: C. B (ed) IJCAI International Joint Conference on Artificial Intelligence, pp 3444–3451
Wang X, Yao T, Ding S, et al. (2020c) Face manipulation detection via auxiliary supervision. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12532 LNCS:313–324. https://doi.org/10.1007/978-3-030-63830-6_27
Ward J (2019) 10 things judges should know about ai. Judicature 103(1):12–18
Westerlund M (2019) The emergence of deepfake technology: A review. Technol Innov Manag Rev 9(11)
Whler L, Zembaty M (2021) Towards understanding perceptual differences between genuine and face-swapped videos. In: Conference on Human Factors in Computing Systems - Proceedings, https://doi.org/10.1145/3411764.3445627
Wu J, Feng K, Chang X, et al. (2020a) A forensic method for deepfake image based on face recognition. In: ACM International Conference Proceeding Series, pp 104–108, https://doi.org/10.1145/3409501.3409544
Wu X, Xie Z, Gao Y, et al. (2020b) Sstnet: Detecting manipulated faces through spatial, steganalysis and temporal features. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp 2952–2956, https://doi.org/10.1109/ICASSP40776.2020.9053969
Xiang Z, Horvath J, Baireddy S, et al. (2021) Forensic analysis of video files using metadata. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp 1042–1051, https://doi.org/10.1109/CVPRW53098.2021.00115
Xie D, Chatterjee P, Liu Z, et al. (2020) Deepfake detection on publicly available datasets using modified alexnet. In: 2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020, pp 1866–1871, https://doi.org/10.1109/SSCI47803.2020.9308428
Xu B, Liu J, Liang J et al (2021) Deepfake videos detection based on texture features. Comput Mater Continua 68(1):1375–1388. https://doi.org/10.32604/cmc.2021.016760
Xuan X, Peng B, Wang W, et al. (2019) On the generalization of gan image forensics. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11818 LNCS:134–141. https://doi.org/10.1007/978-3-030-31456-9_15
Xu Y, Jia G, Huang H, et al. (2021b) Visual-semantic transformer for face forgery detection. In: 2021 IEEE International Joint Conference on Biometrics, IJCB 2021, https://doi.org/10.1109/IJCB52358.2021.9484407
Yang CZ, Ma J, Wang S et al (2021) Preventing deepfake attacks on speaker authentication by dynamic lip movement analysis. IEEE Trans Inf Forensics Secur 16:1841–1854. https://doi.org/10.1109/TIFS.2020.3045937
Yang J, Li A, Xiao S et al (2021) Mtd-net: Learning to detect deepfakes images by multi-scale texture difference. IEEE Trans Inf Forensics Secur 16:4234–4245. https://doi.org/10.1109/TIFS.2021.3102487
Yang J, Xiao S, Li A et al (2021) Detecting fake images by identifying potential texture difference. Futur Gener Comput Syst 125:127–135. https://doi.org/10.1016/j.future.2021.06.043
Yang C, Ding L, Chen Y, et al. (2021a) Defending against gan-based deepfake attacks via transformation-aware adversarial faces. In: Proceedings of the International Joint Conference on Neural Networks, https://doi.org/10.1109/IJCNN52387.2021.9533868
Yang X, Li Y, Lyu S (2019a) Exposing deep fakes using inconsistent head poses. pp 8261–8265, https://doi.org/10.1109/ICASSP.2019.8683164, conference of 44th IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2019 ; Conference Date: 12 May 2019 Through 17 May 2019; Conference Code:149034
Yang X, Li Y, Lyu S (2019b) Exposing deep fakes using inconsistent head poses. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, pp 8261–8265, https://doi.org/10.1109/ICASSP.2019.8683164
Yang C, Lim SN (2020) One-shot domain adaptation for face generation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp 5920–5929, https://doi.org/10.1109/CVPR42600.2020.00596
Yang T, Wu J, Liu L, et al. (2020) Vtd-net: Depth face forgery oriented video tampering detection based on convolutional neural network. In: Fu J. SJ (ed) Chinese Control Conference, CCC, pp 7247–7251, https://doi.org/10.23919/CCC50068.2020.9188580
Yao T, Qu C, Liu Q, et al. (2021) Compound figure separation of biomedical images with side loss. arXiv:2107.08650
Yavuzkilic S, Sengur A, Akhtar Z et al (2021) Spotting deepfakes and face manipulations by fusing features from multi-stream cnns models. Symmetry. https://doi.org/10.3390/sym13081352
Younus M, Hasan T (2020a) Abbreviated view of deepfake videos detection techniques. In: Proceedings of the 6th International Engineering Conference "Sustainable Technology and Development", IEC 2020, pp 115–120, https://doi.org/10.1109/IEC49899.2020.9122916
Younus M, Hasan T (2020b) Effective and fast deepfake detection method based on haar wavelet transform. In: Proceedings of the 2020 International Conference on Computer Science and Software Engineering, CSASE 2020, pp 186–190, https://doi.org/10.1109/CSASE48920.2020.9142077
Yu M, Zhang J, Li S et al (2021) Deep forgery discriminator via image degradation analysis. IET Image Proc 15(11):2478–2493. https://doi.org/10.1049/ipr2.12234
Zendran M, Rusiecki A (2021) Swapping face images with generative neural networks for deepfake technology - experimental study. In:

Procedia Computer Science, pp 834–843, https://doi.org/10.1016/j.procs.2021.08.086
Zeng Y, Guo X, Yang Y, et al. (2020) Dfdm - a deepfakes detection model based on steganography forensic network. Communications in Computer and Information Science 1253 CCIS:536–545. https://doi.org/10.1007/978-981-15-8086-4_51
Zhang K, Liang Y, Zhang J et al (2019) No one can escape: a general approach to detect tampered and generated image. IEEE Access 7:129494–129503. https://doi.org/10.1109/ACCESS.2019.2939812
Zhang W, Zhao C, Li Y (2020) A novel counterfeit feature extraction technique for exposing face-swap images based on deep learning and error level analysis. Entropy. https://doi.org/10.3390/e22020249
Zhang H, Lu ZM, Luo H et al (2021) Restore deepfakes video frames via identifying individual motion styles. Electron Lett. https://doi.org/10.1049/ell2.12015
Zhang Y, Gao F, Zhou Z, et al. (2021b) A survey on face forgery detection of deepfake. In: Jiang X. FH (ed) Proceedings of SPIE - The International Society for Optical Engineering, https://doi.org/10.1117/12.2600889
Zhang X, Karaman S, Chang SF (2019b) Detecting and simulating artifacts in gan fake images. https://doi.org/10.1109/WIFS47025.2019.9035107, conference of 2019 IEEE International Workshop on Information Forensics and Security, WIFS 2019 ; Conference Date: 9 December 2019 Through 12 December 2019; Conference Code:158617
Zhang X, Karaman S, Chang SF (2019c) Detecting and simulating artifacts in gan fake images. In: 2019 IEEE International Workshop on Information Forensics and Security, WIFS 2019, https://doi.org/10.1109/WIFS47025.2019.9035107
Zhao B, Zhang S, Xu C et al (2021) Deep fake geography? when geospatial data encounter artificial intelligence. Cartogr Geogr Inf Sci 48(4):338–352. https://doi.org/10.1080/15230406.2021.1910075
Zhao Z, Wang P, Lu W (2021) Multi-layer fusion neural network for deepfake detection. Int J Digit Crim Forensics 13(4):26–39. https://doi.org/10.4018/IJDCF.20210701.oa3
Zhao Y, Ge W, Li W, et al. (2020a) Capturing the persistence of facial expression features for deepfake video detection. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11999 LNCS:630–645. https://doi.org/10.1007/978-3-030-41579-2_37
Zhao Z, Wang P, Lu W (2020b) Detecting deepfake video by learning two-level features with two-stream convolutional neural network. In: ACM International Conference Proceeding Series, pp 291–297, https://doi.org/10.1145/3404555.3404564
Zheng Q, Yang M, Yang J et al (2018) Improvement of generalization ability of deep cnn via implicit regularization in two-stage training process. IEEE Access 6:15,844-15,869. https://doi.org/10.1109/ACCESS.2018.2810849
Zhu B, Fang H, Sui Y, et al. (2020a) Deepfakes for medical video de-identification: Privacy protection and diagnostic information preservation. In: AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, pp 414–420, https://doi.org/10.1145/3375627.3375849
Zhu H, Fu C, Wu Q, et al. (2020b) Aot: Appearance optimal transport based identity swapping for forgery detection. In: Advances in Neural Information Processing Systems
Zhu K, Wu B, Wang B (2020c) Deepfake detection with clustering-based embedding regularization. In: Proceedings - 2020 IEEE 5th International Conference on Data Science in Cyberspace, DSC 2020, pp 257–264, https://doi.org/10.1109/DSC50466.2020.00046
Zi B, Chang M, Chen J, et al. (2020) Wilddeepfake: A challenging real-world dataset for deepfake detection. In: MM 2020 - Proceedings of the 28th ACM International Conference on Multimedia, pp 2382–2390, https://doi.org/10.1145/3394171.3413769
Zotov S, Dremliuga R, Borshevnikov A, et al. (2020) Deepfake detection algorithms: A meta-analysis. In: ACM International Conference Proceeding Series, pp 43–48, https://doi.org/10.1145/3421515.3421532