January 2020

Bibliometric Analysis of Passive Image Forgery Detection and Explainable AI

Kalyani Dhananjay Kadam
Symbiosis Institute of Technology ,Symbiosis International University, kalyanik@sitpune.edu.in

Dr. Swati Ahirrao Ahirrao
Symbiosis Institute of Technology , Symbiosis International University, swatia@sitpune.edu.in

Dr. Ketan Kotecha
Head, SCAAI ,Symbiosis International University, head@scaai.siu.edu.in

Bibliometric Analysis of Passive Image forgery Detection and Explainable AI

Abstract - Due to the arrival of social networking services such as Facebook and Instagram, there has been a vast increase in the volume of image data generated in the last decade. Use of image processing tools like GNU Gimp, Adobe Photoshop to create doctored images and videos is a major concern. These are the main sources of fake news and are often used in malevolent ways such as for mob incitement.Before move can be taken based on a fake image , we should confirm its realness.This paper shows systematic mappings of existing literature for image forgery detection using deep learning and explainable AI. This uses Scopus database for data analysis and various tools such as Sciencescape,Gephi,Tableau and VOS Viewer. The study discovered that largest number of reviews on image forgery detection using deep learning and explainable AI had explored very recently.It was observed that USA universities/institutions are foremost in the research studies focusing on this research topic .

Keywords Image forgery detection , deep Learning,explainable AI, bibliometric analysis.

1. Introduction -

Nowadays, there hardly exists any area where digital images are not used. We people will in general have faith in what we see as opposed to what we hear. Therefore visible information becomes more important for us rather than verbal data. Digital images are used in many fields such as the website of the newspapers, Facebook, Instagram, courts, medical systems, electronic media, military, and industries and so on over the internet[1].

Computers modernized the workmanship with early-articulated work by Hollywood, for example, Jurassic Park, and Terminator in mid 1990s.In 2001, after the episode of 9/11, a few videotapes of Osama container Laden were spread. After that [2] many forensic strategies are utilized for pulling out data from images and it additionally recognize how the pictures have been digitally altered.With the advancement of image processing software like Adobe Photoshop, Gimp, Coral Draw etc. it's become very easy to alter the image[3,4]. Iran did a trial of soaring of three missiles in 2008 though a picture of soaring four missiles was spread, which was later on proven to be manufactured [5,6].

These images often used as a source for spreading false information among the mob through social sites to mislead, cause emotional distress, or to purposefully influence decisions , approaches, and actions [7].As a result, the genuineness of digital images has

become a significant research area in the literature recently. XFake system[8], first explainable fake news detector is presented which helps end users to detect the news is fake or not.

Identifying forgery in the digital images is one of the difficulties of this invigorating digital area. Image forgery detection mainly classified into two approaches [9,10,11]: active approach and passive or blind approach. Active forgery detection approach uses a digital watermark or signature embedded in the original image to prove or reject the authenticity of the image. This approach having strong limitation that the watermark which is embedded must be performed either by a person authorized to process the image or by the acquisition device.

Passive approaches [9,10,11] gaining more attention of researcher as they don't require any information to detect tampered images.In passive approach as per literature copy move and image splicing techniques are gaining more attention. Copy move and image splicing are one of the most commonly performed manipulations on digital images. In copy-move technique a part of the image is copied and pasted somewhere else in the same image with an intention to cover an important image feature or object to create a duplicate of the object or to change the significance of image completely. There are a number of ways with the help of which copy move forgery can be performed. Based upon which it can categorized into plain copy move, multiple copy move and copy move with image inpainting .Image splicing is also known as image compositing is a process in which two or three images are merged to create a forged image. Splicing is alike to copy move forgery in some extent. In copy-move the region is copied into the same image but in the case of splicing it is pasted in different image.

In the light of previously mentioned situation the researchers were in this way spurred to complete bibliometric studies and understand in deep researches carried out on image forgery detection using deep learning and explainable AI. It is standard practice to compose list of sources at the end of book, article or report which are nothing but the number of sources utilized for creation of it. Bibliometric study in current time has grown probably the best practice which opens portal to new research subjects. This paper [12] characterized Bibliometric as the application of statistical and mathematical techniques to books and other media of communication. Bibliometric examinations can possibly produce an information driven vision of scientific research activities over various research areas and can display proof based depictions, correlations, and representations of research yields [15] . They have been perceived as an important tool for policy makers and researchers to see the status of research and science, and to help in future assignments dependent on priorities [16] and to shape the future research directions .

2. Bibliometric Analysis of Image tampering detection using Deep Learning :

While doing research in a particular area it is compulsory to have in depth knowledge of progressing research in that field and the number of authors who add to such research. Due to advancement in technologies vast amount of information is available in research, number of methods such as bibliometrics,webometrics,scientometric, and Hindex are used for identifying trends in research [14] . Bibliometric[13,21] study provide insights into the contribution of various countries, institutes, authors and journals in research area. This provides support to researchers in shaping up and enhancing further research actions and investments. To meet these goals, a bibliometric analysis based on data collected from Scopus database was carried out to identify a set of bibliometric enactment indicators such as quantitative indicators indicates productivity, and qualitative indicators specifies citations and Hirsch index (h -index).

2.1 Analysis of Keywords:

Table 1 List of Primary and Secondary keywords

Primary – Keyword	Image forgery detection	
Secondary - Keyword	(AND)	Deep Learning
	(OR)	" tampering" or "manipulation " OR "fabrication" OR "CNN" OR "Image Forensics" OR "Computer Vision"

Thus the query used to search the documents in Scopus is : "image" AND "forgery" OR "tampering" OR "manipulation" OR "fabrication" AND detection OR identification OR classification AND using AND deep AND learning**.**

2.2 Initial search results

Scopus database is the base of this research paper. By utilizing these keywords it gives all out 192 publications. Here we mainly concentrate on English language which gives 188 publications. (Table 2).

Table 2: Image forgery detection publishing languages trends

Publication Language	Publications
English	188
Chinese	3
Korean	1
Total publications	192

Source: http://www.scopus.com (assessed on 11th December 2019)

All types of published publications are considered comprising of journal papers, articles, book chapters, conference proceedings etc. for this survey. The researchers in image forgery detection using deep learning have publicized recent papers in conferences. 49.5% of conference papers and 31.8 % of journal articles were there. (Figure 1).

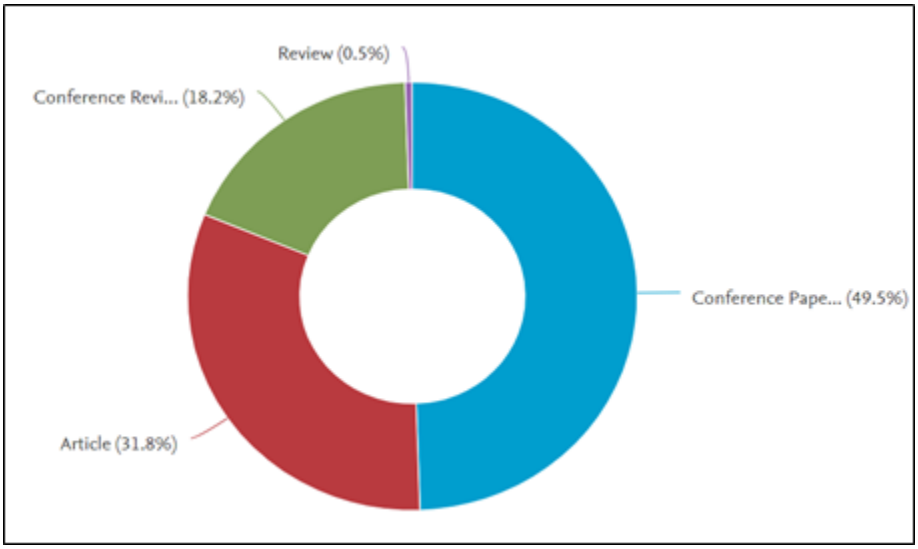

Figure 1 : Publication types in Image Forgery Detection Using Deep Learning

Source: http://www.scopus.com (assessed on 11th December 2019) Please take from Scopus database directly

2.3 Analysis by year:

The documents related to image forgery detection derived for the interval of ten years beginning from 2011 to 2019. Table shows trends in the number of publication count per year in this research field. By investigating this information, it very well may be effectively

broke down that the research in this field has been contributed more in the year from 2017 to 2019. However, very few researches were carried out in the span of year 2011 to 2016.

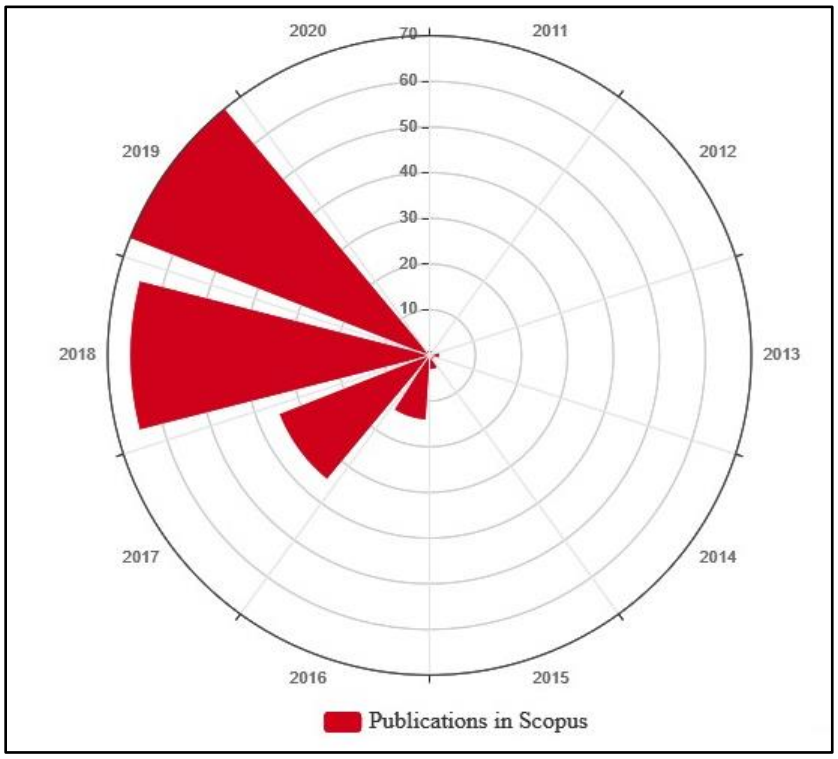

Figure 2 : Year wise Publications

Source: http://www.scopus.com (assessed on 11th December 2019).

2.4 Analysis by geographic location:

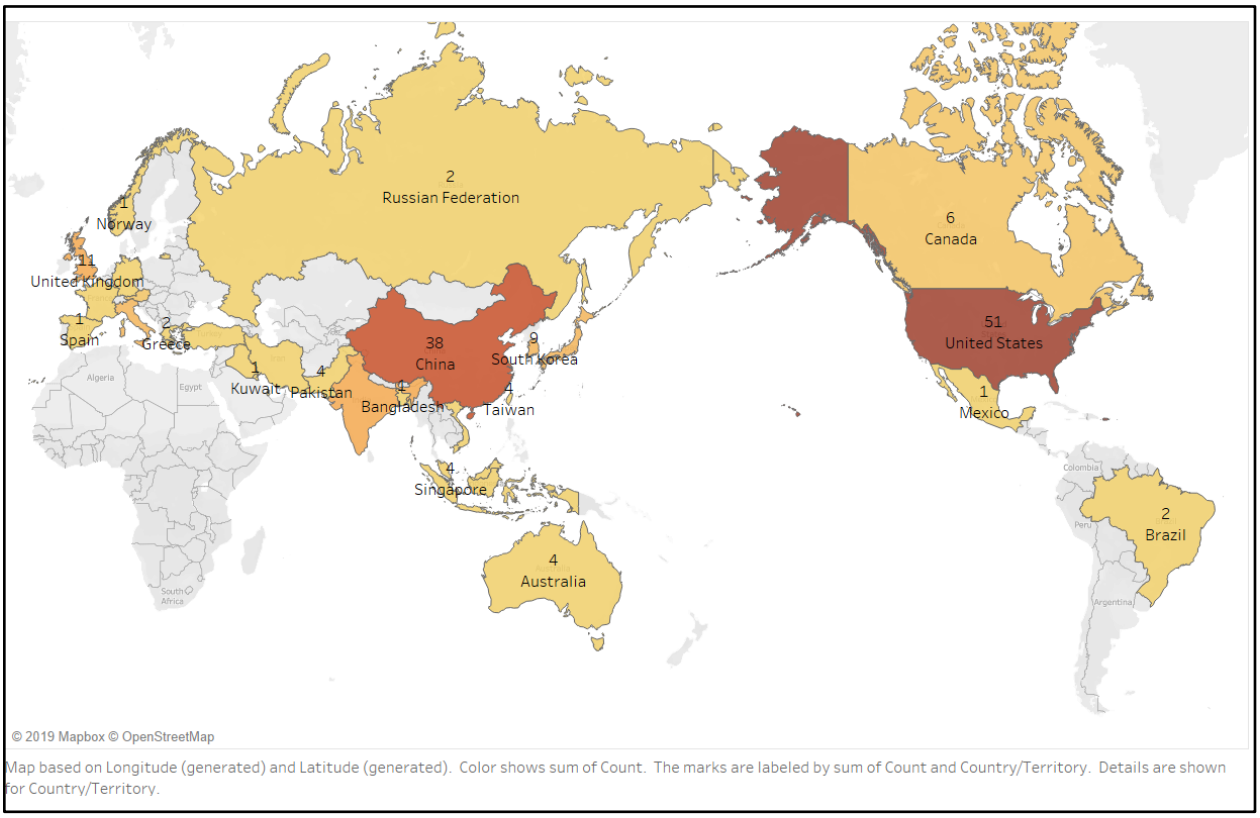

Figure 3 : Geographic locations of research related to Image forgery detection using Deep Learning (Please change colour of this figure)

Figure 3 is drawn using Tableau showing geographical regional location clusters of published papers. This map shows countries with their research counts. The maximum research has taken place in the USA.The other countries include China, India, and United Kingdom etc.

2.5 Analysis by Subject area:

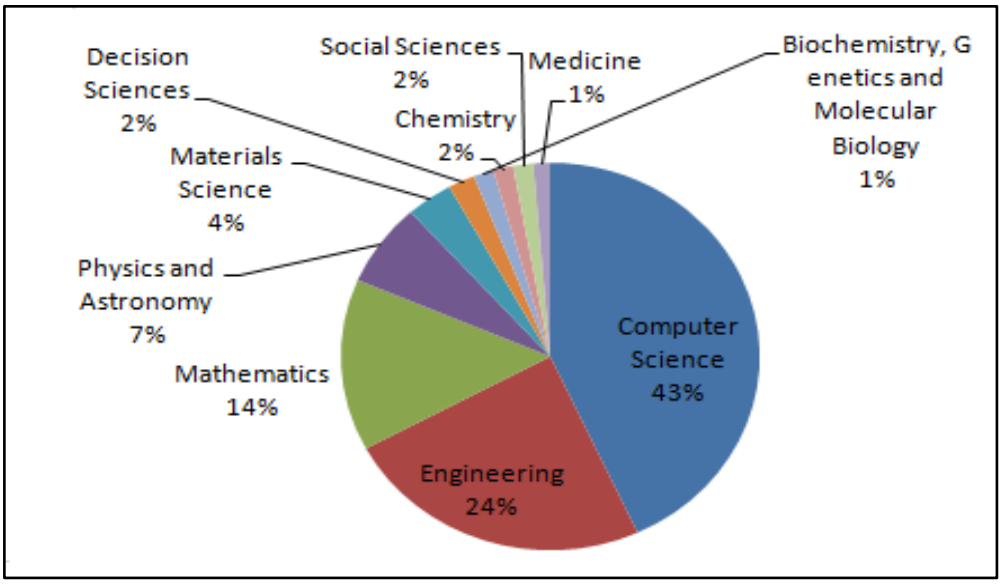

Fig. 4 Top ten most popular subject areas. Source: https://www.scopus.com/ (accessed on 11th December 2019)

Figure 4 shows the subject area-wise compartmentalization for research area . It is clear from the figure that maximum research has been carried out in computer science followed by engineering and mathematics.

2.6 Network Analysis

Network analysis is a graphical representation technique used for understanding the association between various statistical factors . Tools used for this are ScienceScape,Table2Net ,Gephi and VOSViewer. Figures 5–8 show different types of networks that are drawn with different parametric combinations image forgery detection using deep learning from the data extracted from Scopus assessed on 11th December 2019.In figure 5 clusters are built based on number of publications that author is having in that research area. Figure 6 indicates affiliation statistics.

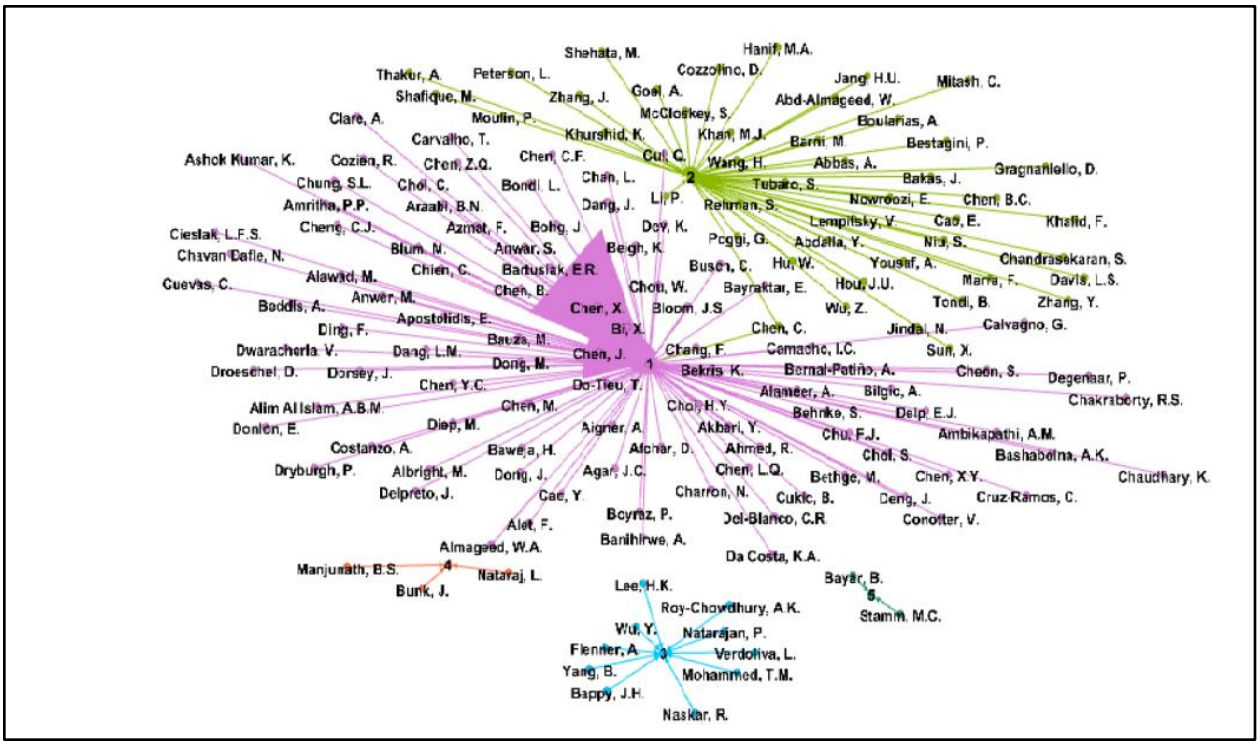

Figure 5 :Clusters as per Author Publications Count in Scopus

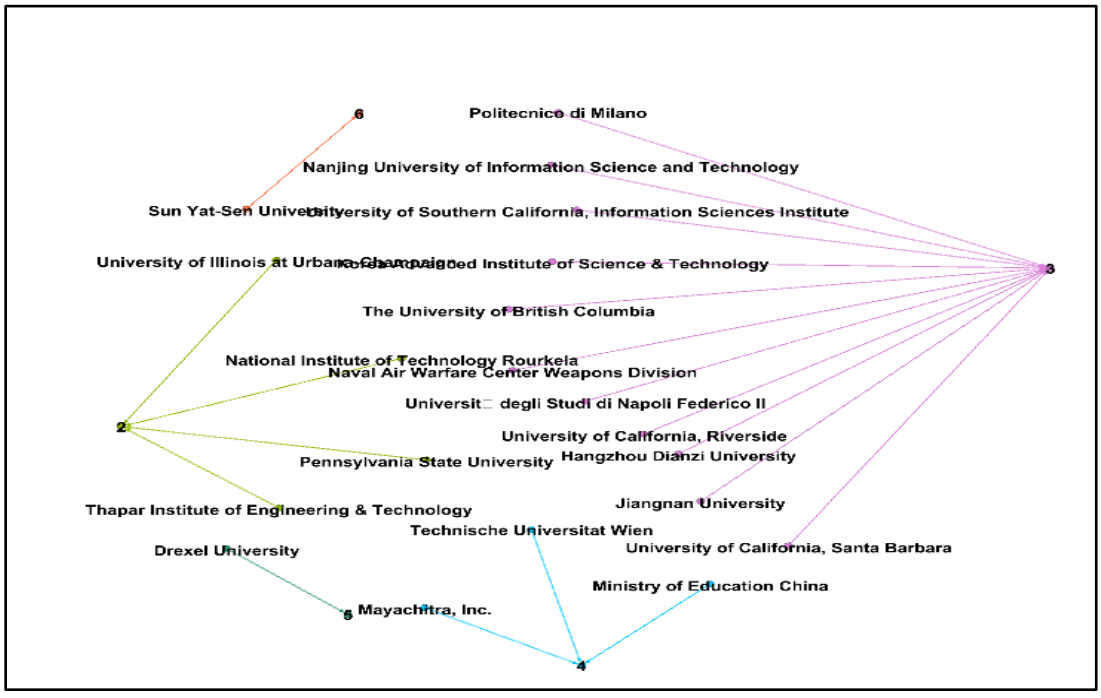

Figure 6:A cluster of affiliation for top twenty contributing universities/organizational affiliations.

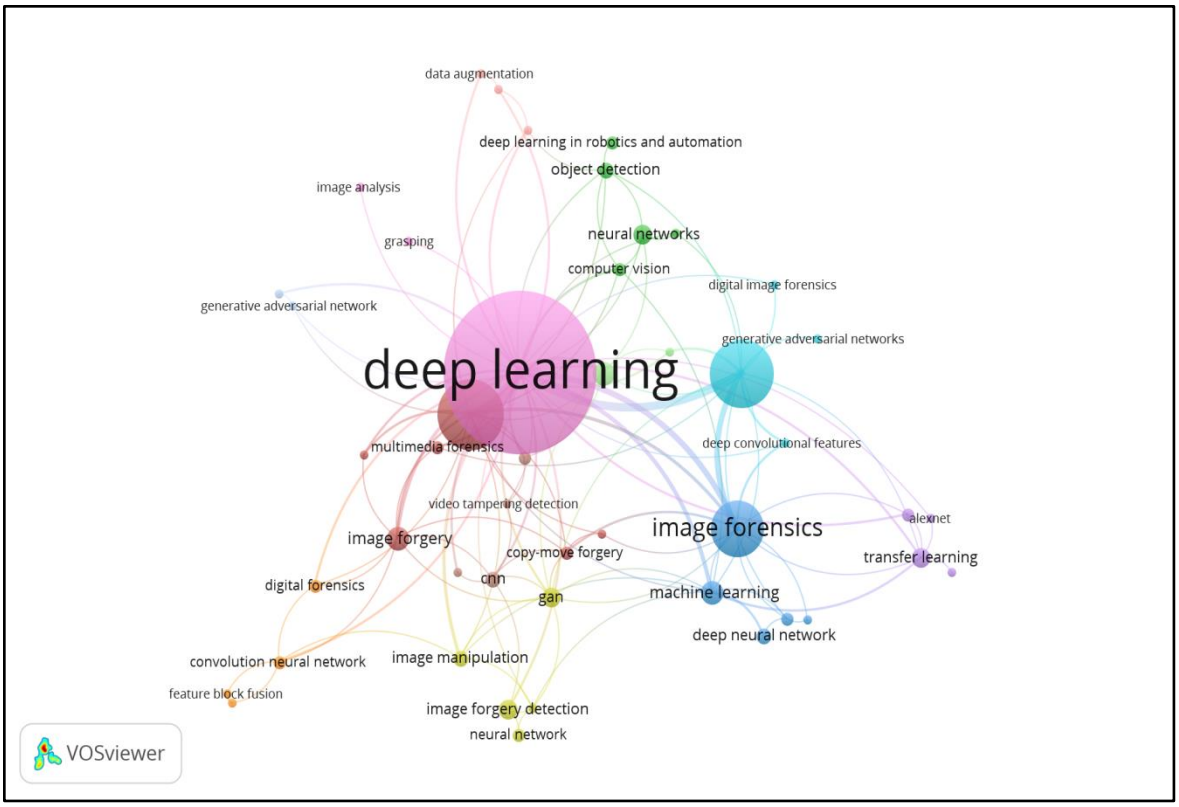

Figure.7.Network visualization map in association with published research that utilized Deep Learning for image forgery detection. Analysis of co-occurrence of author keywords, minimum number of occurrences of a keyword was set to 2, of the 379 keywords: 47 keywords meet the threshold. For each of the 47 keywords, the number of co-occurrence links was calculated. The keywords with the largest number of links are selected. Keywords which are having the most intra-cluster co-occurrence relations are arranged in the same cluster (in this case, there were 12 clusters) Source: https://www.scopus.com/ (accessed on 11th December 2019)

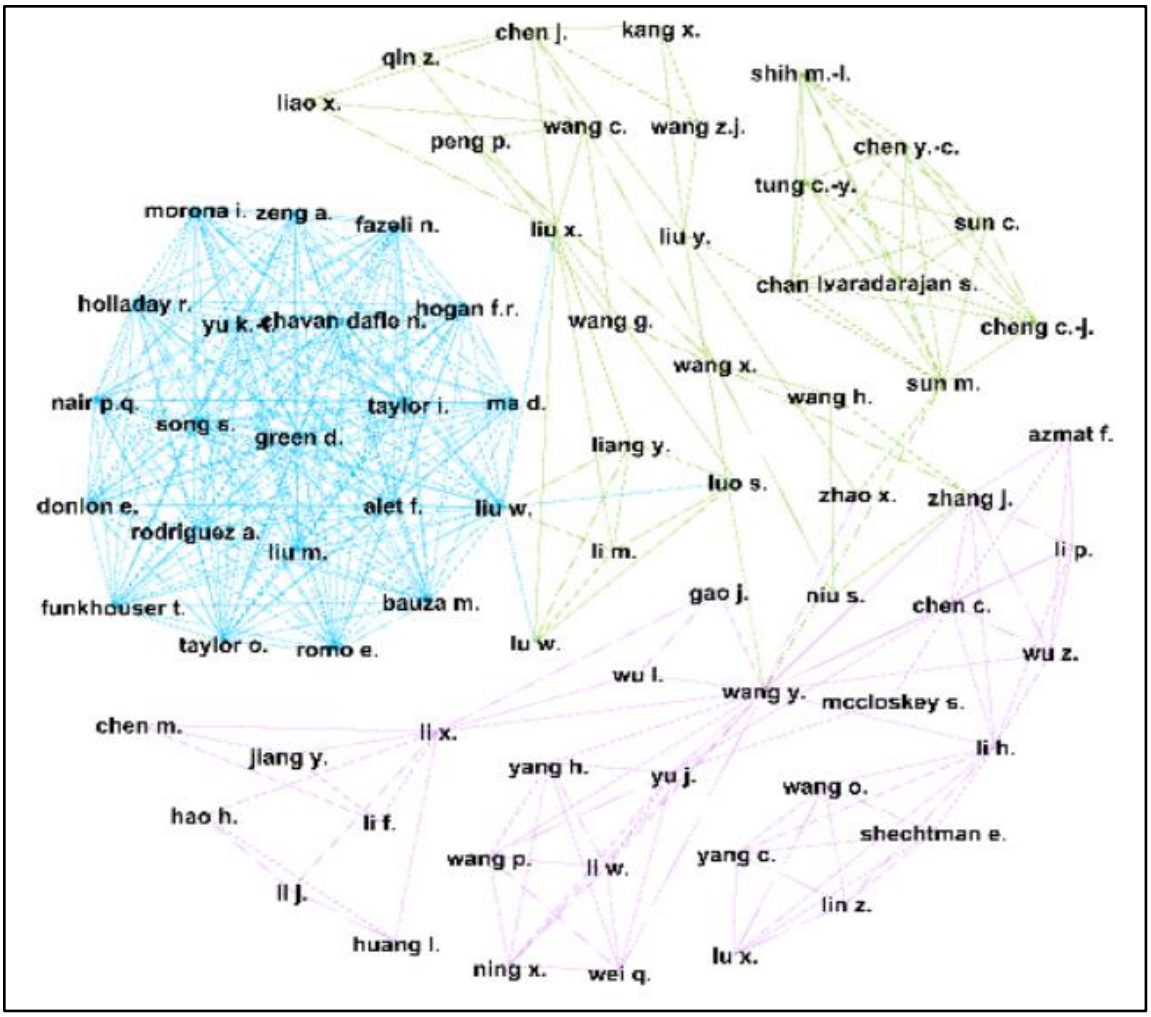

Figure.8 .shows network of authors linked by co-publication which is 414 nodes and 1168 edges reduced network with filter Remove nodes < 3 links, then disconnected nodes. The size of node and label indicates maximum co-publication link with different authors.

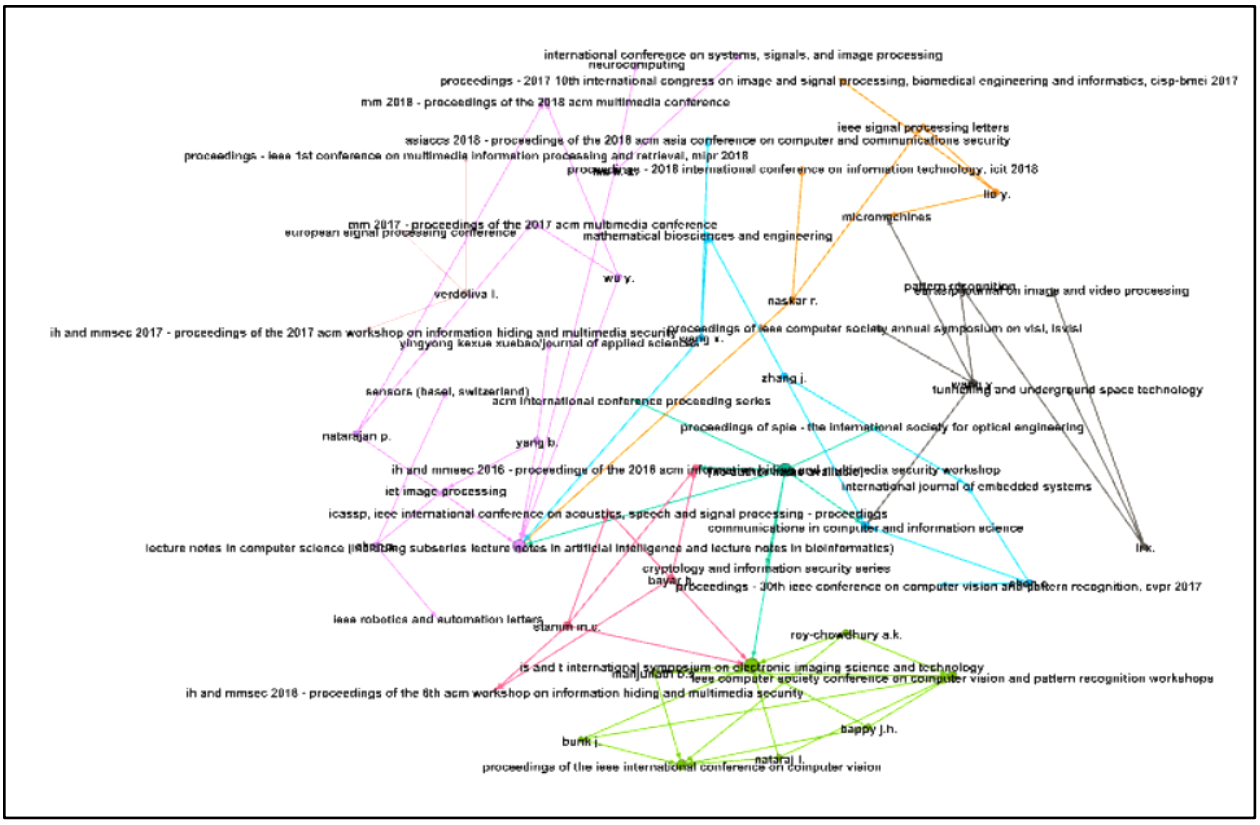

Figure.9 is drawn for network of authors and source titles, co-appearing in the same papers with Nodes 55 and Edges: 70 with filtering condition as Remove nodes < 3 links, then disconnected nodes

2.7 Citation analysis

The significance of the publications is being comprehended through the idea of citation analysis which shows the reputation of the publication. Citation analysis is done by counting the number of times the individual publication is cited or referred by other publications. Table 3 shows the citation in the area of research. The total citation count of 192 publications is 1260, to date. The list of the first ten papers along citations received to them till date for this research is shown in Table 4.

Table 3 Analysis of citations for publications related to Image forgery detection using Deep Learning . Source: https://www.scopus.com/ (accessed on 11 Dec 2019)

Year	<2015	2015	2016	2017	2018	2019	Subtotal	>2019	Total
No. of Citations	0	0	10	110	427	670	1217	43	1260

Table 4 A citation analysis of the top ten publications related to XAI . Source: https://www.scopus.com/ (accessed on 11 Dec 2019)

Publication Title	<2015	2015	2016	2017	2018	2019	Subtotal	>2019	Total Citations
Robust Physical-World Attacks on Deep Learning Visual Classification	0	0	0	0	6	43	49	1	50
Constrained Convolutional Neural Networks: A New Approach Towards General Purpose Image Manipulation Detection	0	0	0	0	7	21	28	1	29
What-and-where to match: Deep spatially multiplicative integration networks for person re-identification	0	0	0	0	10	39	49	4	53
Modern Trends in Hyperspectral Image Analysis: A Review	0	0	0	0	9	30	39	1	40
High-resolution image inpainting using multi-scale neural patch synthesis	0	0	0	6	47	69	122	2	124
A deep learning approach to detection of splicing and	0	0	0	8	32	28	68	4	72

copy-move forgeries in images									
Design principles of convolutional neural networks for multimedia forensics	0	0	0	3	18	11	32	0	32
Safety verification of deep neural networks	0	0	0	3	41	51	95	0	95
A deep learning approach to universal image manipulation detection using a new convolutional layer	0	0	1	29	53	44	127	3	130
Median Filtering Forensics Based on Convolutional Neural Networks	0	0	7	31	53	49	140	1	141

Table 5 : Citation Analysis of Top Ten Journal in Image forgery detection in Deep Learning

Journal Title	<2015	2015	2016	2017	2018	2019	Subtotal	>2019	Total Citations
Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition	0	0	0	0	6	43	49	1	50

Pattern Recognition	0	0	0	0	10	39	49	3	53
IEEE Access	0	0	0	0	9	30	39	1	40
Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017	0	0	0	6	47	69	122	2	124
Journal of Digital Imaging	0	0	0	9	15	16	40	1	41
8th IEEE International Workshop on Information Forensics and Security, WIFS 2016	0	0	0	8	32	28	68	4	72
IS and T International Symposium on Electronic Imaging Science and Technology	0	0	0	3	18	11	32	0	32
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	0	0	0	3	41	51	95	0	95
IH and MMSec 2016 - Proceedings of the 2016 ACM Information Hiding and Multimedia	0	0	1	29	53	44	127	3	130

Security Workshop									
IEEE Signal Processing Letters	0	0	7	31	53	49	140	1	141

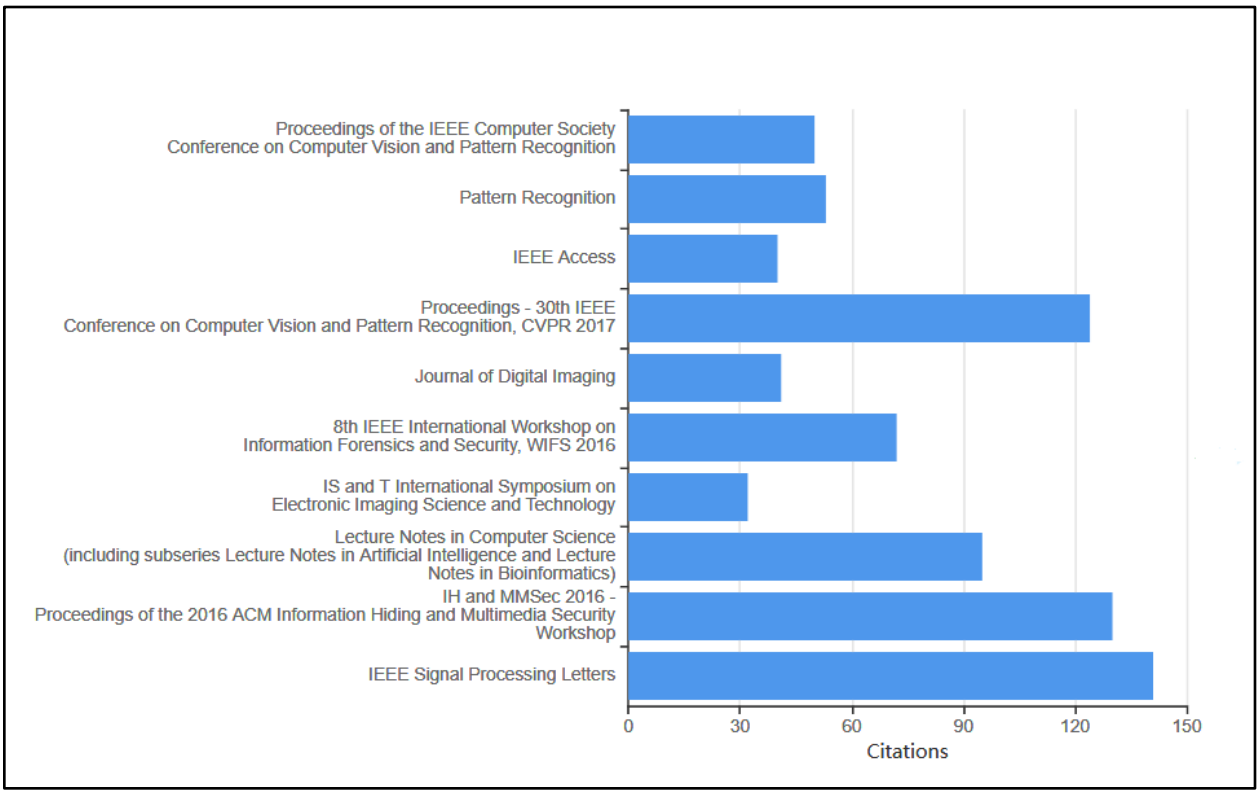

Fig. 10 Journal statistics for publications on image forgery detection using deep learning in the top ten most popular journals. Source: https://www. scopus.com/ (accessed on 11 Dec 2019)

3. Bibliometric Analysis of Explainable AI

The query used to search the documents in Scopus is : explainable AND ai.

3.1 Keywords statistics

First ten keywords are given in Table 6 from publications related to explainable AI. Keyword specifies what the researchers want to search. The right blend of keywords targets foremost research areas.

Table 6 First ten keywords related to Explainable AI . Source: https://www.scopus.com/ (accessed on 11 Dec 2019).

Keywords	Number of publications	Keywords	Number of publications
Explainable AI	124	Decision Making	40
Artificial Intelligence	83	Interpretability	29
Machine Learning	67	Deep Neural Networks	28
Deep Learning	49	User Interfaces	26
Learning Systems	42	Explainable Artificial Intelligence	24

3.2 Initial search results

By using these keywords it gives all out 281 publications. Here we considering on English language which gives 280 publications. All types of published publications are considered comprising of journal papers, articles, book chapters, conference proceedings etc. for this survey. The researchers in explainable AI have publicized recent papers in conferences. 73.7% of conference papers and 16.7 % of journal articles were there. (Table 7).

Table 6 : Publication types in Explainable AI (Source: http://www.scopus.com (assessed on 11th December 2019)

Source type	Number of publications	Percentage of publication (%)
Conference Paper	207	73.7 %
Article	47	16.7 %
Conference Review	11	3.9 %
Review	10	3.6 %
Book Chapter	3	1.1 %
Editorial	2	0.7 %
Undefined	1	0.4

3.3 Analysis by year:

The documents related to explainable AI derived for the duration beginning from 1990 to 2020 . Figure shows trends in the number of publication count per year in this research field. By considering this information, it shows that the research in this field has been contributed more in the year from 2017 to 2019. However for duration ranging from 1990 to 2016 very less research were carried out.

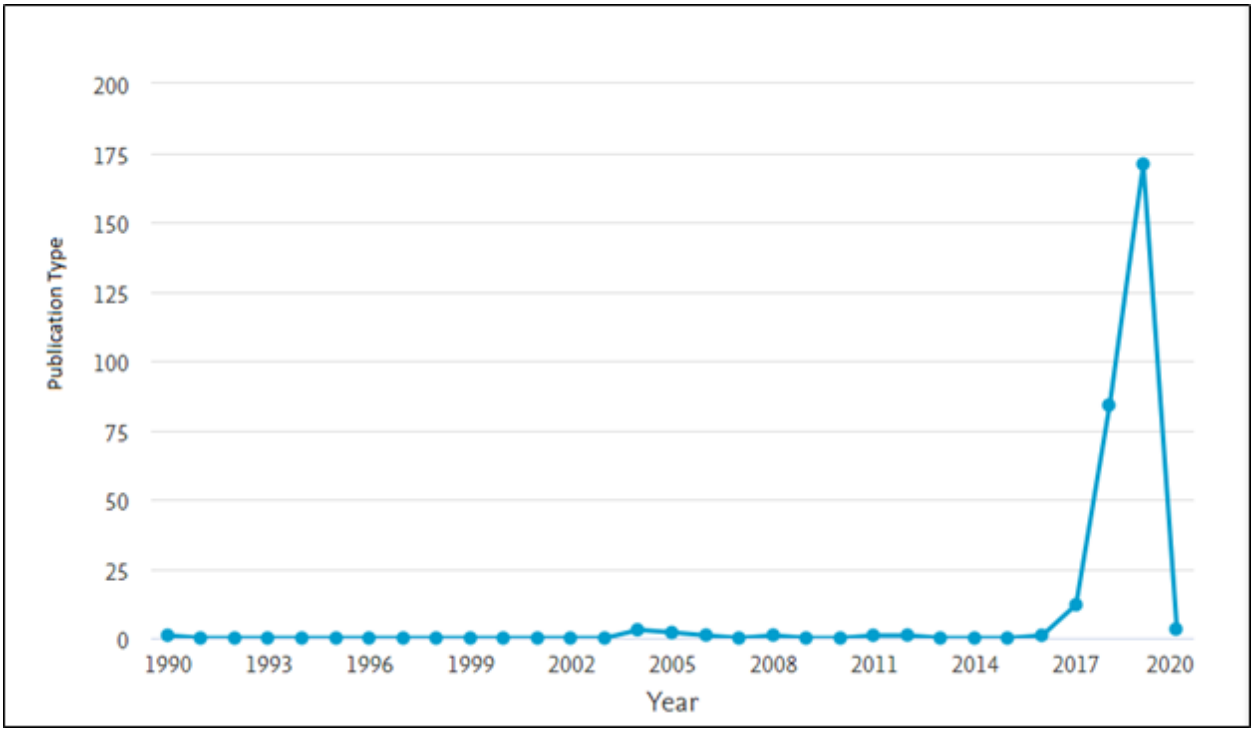

Figure 11 : Year wise Publications (Source: http://www.scopus.com (assessed on 11th December 2019).

3.4 Analysis by Subject area:

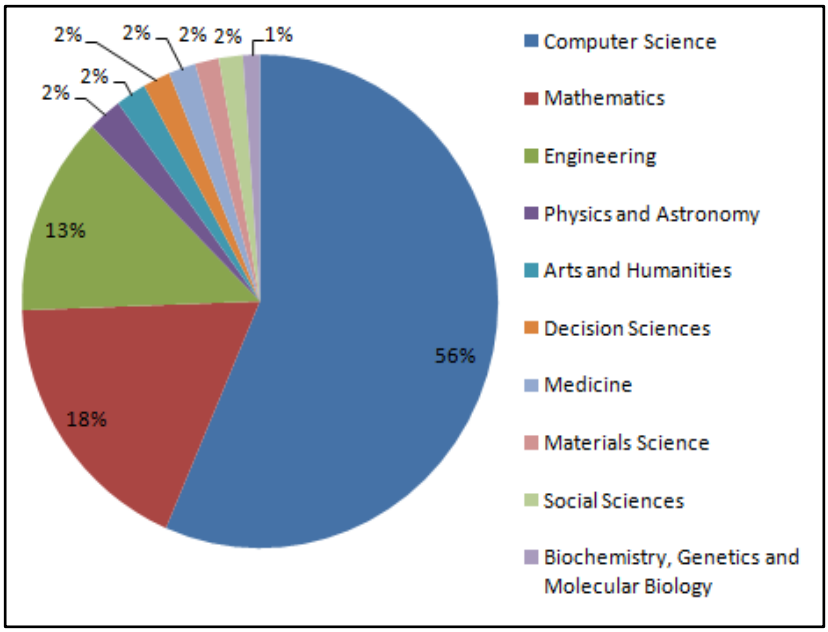

Fig. 11 Top ten most popular subject areas. Source: https://www.scopus.com/ (accessed on 11th December 2019)

Figure 11 shows the subject area-wise distribution for research area . From the figure it is clear that utmost research has been carried out in computer science followed by mathematics and engineering.

3.5 Analysis by geographic location:

Figure 12 : Geographic locations of Explainable AI

Figure 12 is drawn using Tableau showing geographical locations of XAI. This map shows countries with their research counts. The maximum research has taken place in the USA.The other countries include the United Kingdom, Germany , and Netherland etc.

3.6 Network Analysis

Figures 13–17 show various kinds of networks that are drawn with various parametric blends in XAI from the data extracted from Scopus considered on 11th December 2019.

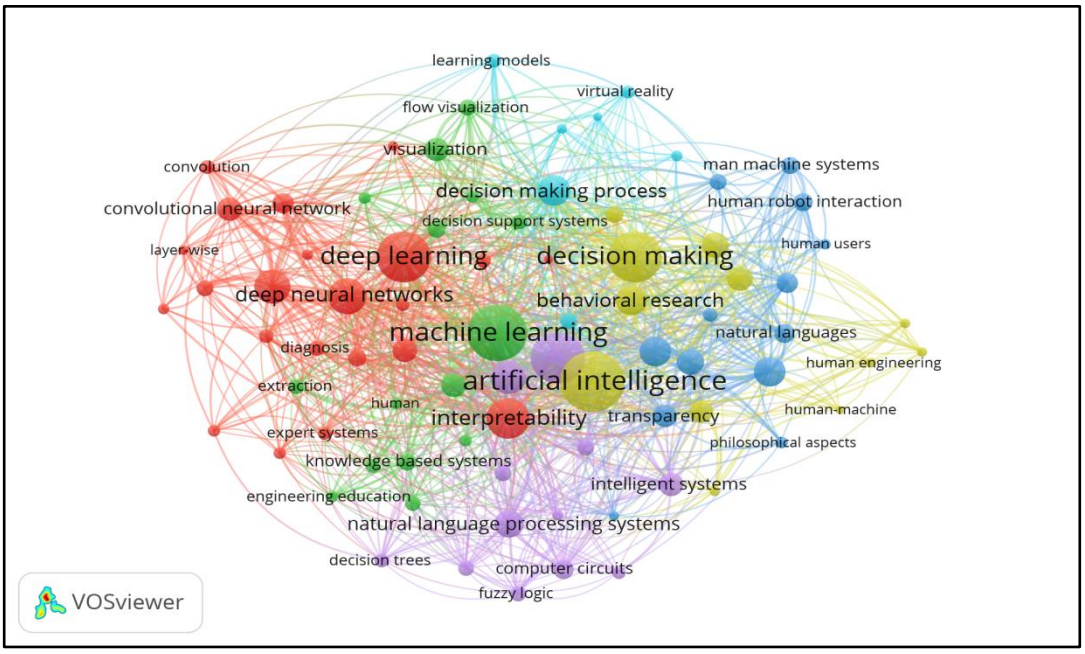

Figure.13 .Network visualization map for XAI . Analysis of co-occurrence of index keywords, minimum number of occurrences of a keyword was set to 5, of the 1795 keywords: 81 keywords meet the threshold. For each of the 81 keywords, the number of co-occurrence links was calculated. The keywords with the largest number of links are selected. Keywords which are having the most intra-cluster co-occurrence relations are arranged in the same cluster (in this case, there were 6 clusters) Source: https://www.scopus.com/ (accessed on 11th December 2019)

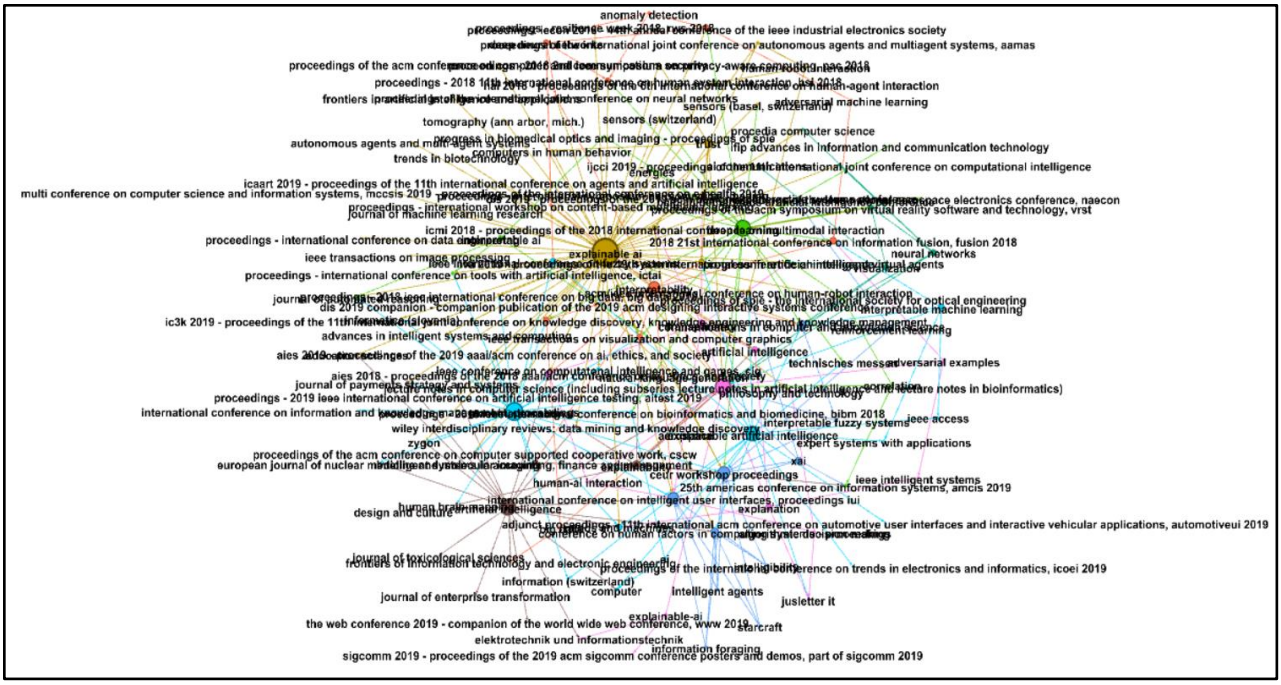

Figure 14 Clustering networks with author keywords and source title

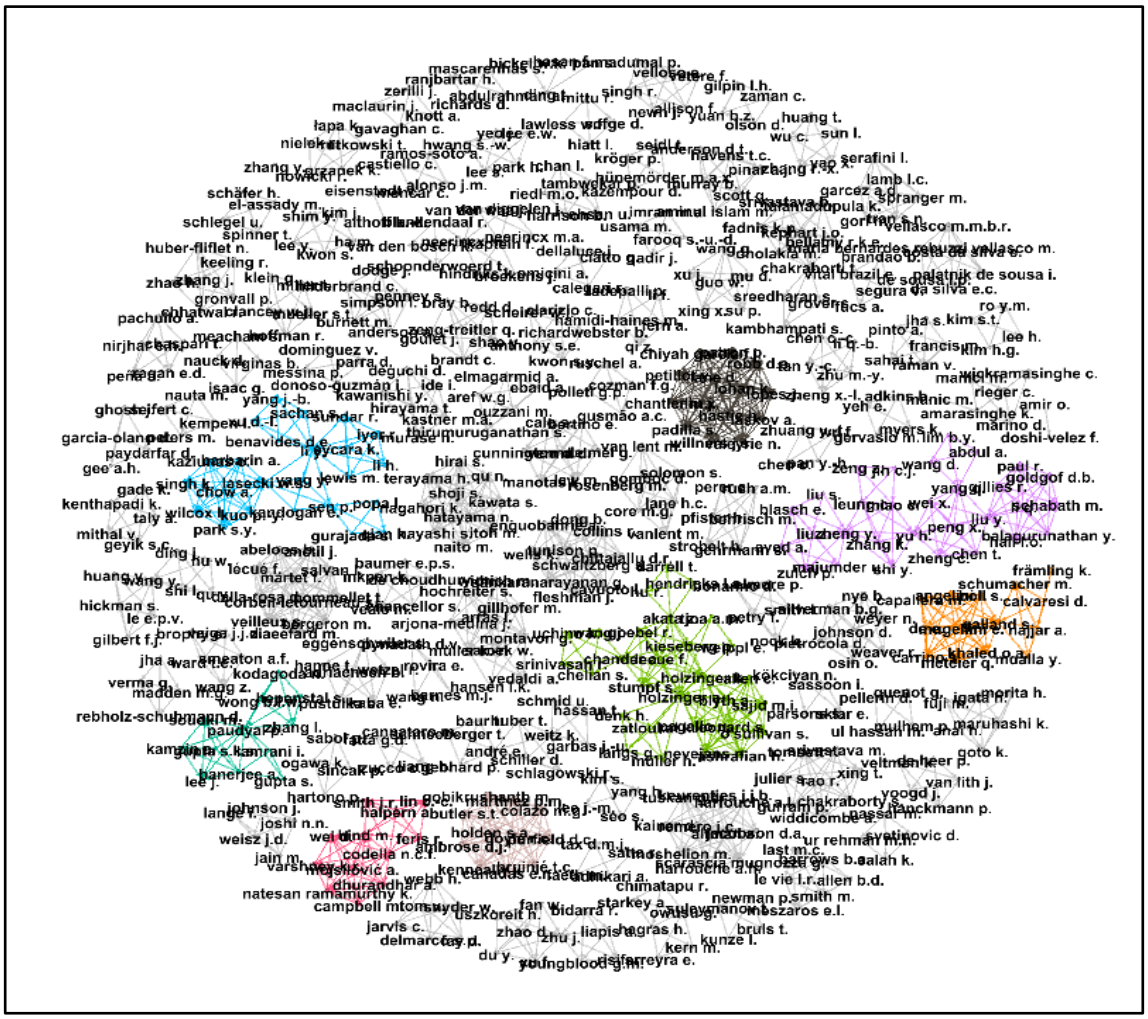

Figure 15 Network of authors linked by co-publication

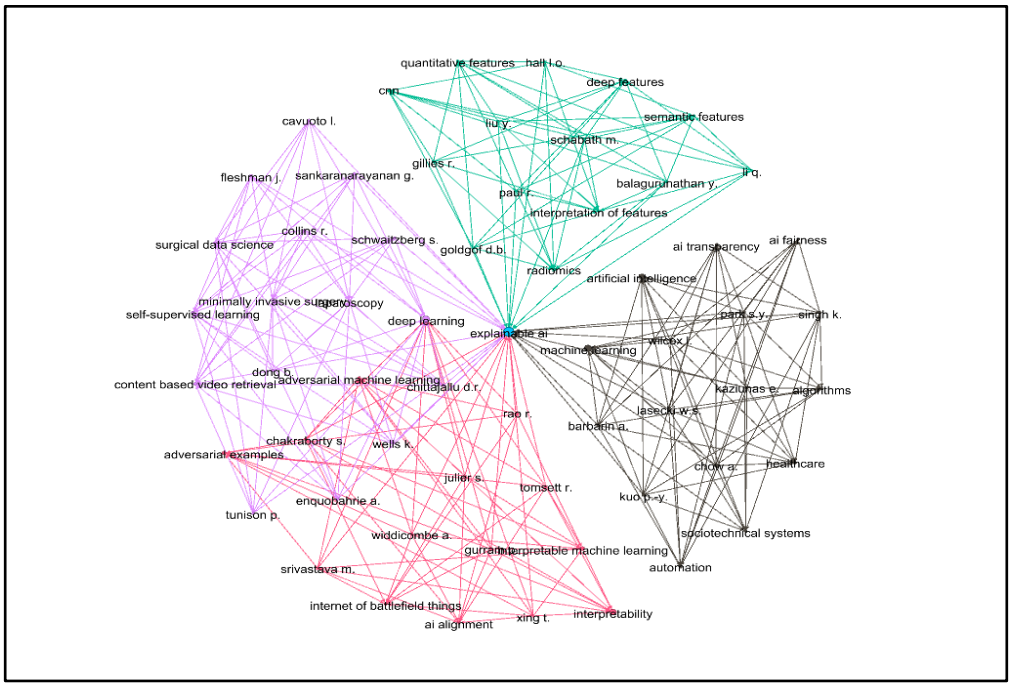

Figure 16 Network diagram for Authors and author Keywords, co-appearing in the same papers with Nodes 1083 and Edges 3260 with filtering condition as Remove nodes < 3 links, then disconnected nodes with this filter 37 nodes again removed.

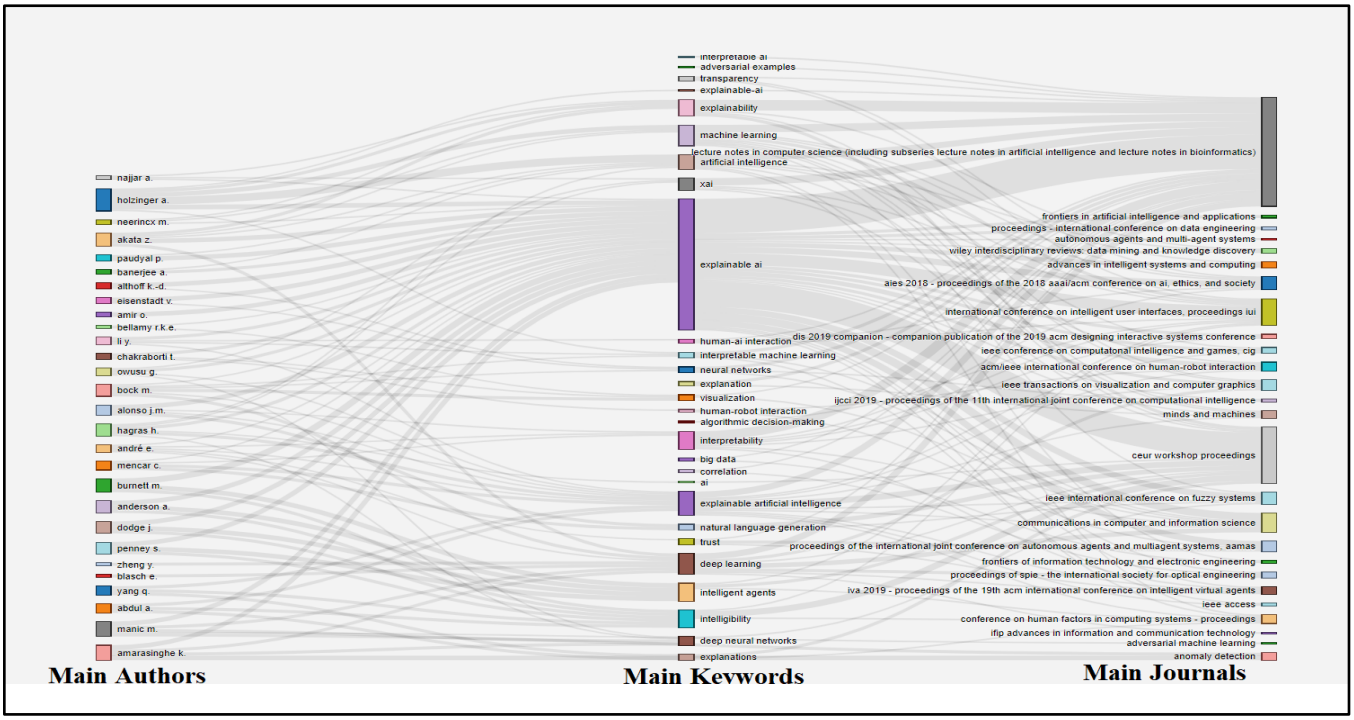

Figure 17 Network of main author -main keywords-main journals

3.7 Citation analysis

The significance of the publications is being comprehended through the idea of citation analysis which shows the reputation of the publication. Citation analysis is done by counting the number of times the individual publication is cited or referred by other publications. Table 3 shows the citation in the area of research. The total citation count of 281 publications is 605, to date. The list of the first ten papers along with citations are shown in table 4.

Table 7 Analysis of citations for publications related to XAI . Source: https://www.scopus.com/ (accessed on 11 Dec 2019)

Year	<2015	2015	2016	2017	2018	2019	Subtotal	>2019	Total
No. of Citations	66	6	6	20	84	401	583	22	605

Table 8 A citation analysis of the top ten publications related to XAI. Source: https://www.scopus.com/ (accessed on 18 Dec 2019)

Publication Title	<2015	2015	2016	2017	2018	2019	Subtotal	>2019	Total Citations
Explanation in artificial intelligence: Insights from the social sciences	0	0	0	0	0	47	47	1	48
Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)	0	0	0	0	0	41	41	3	44
Building explainable artificial intelligence systems	14	2	1	3	12	5	23	0	37

An explainable artificial intelligence system for small-unit tactical behavior	16	0	0	0	11	7	18	0	34
Transparent, explainable, and accountable AI for robotics	0	0	0	5	9	19	33	0	33
Challenges and opportunities: from big data to knowledge in AI 2.0	0	0	0	5	3	14	22	0	22
Design recommendations to support automated explanation and tutoring	19	0	1	0	1	0	2	0	21
Explainable AI: The new 42?	0	0	0	0	3	14	17	2	19
Rich socio-cognitive agents for immersive training environments: Case of NonKin Village	8	3	4	1	1	2	11	0	19
The impact of POMDP-generated explanations on trust and performance in human-robot teams	0	0	0	4	9	5	18	0	18

Table 5: Citation Analysis of Top Ten Journal in Image forgery detection in Deep Learning

Journal Title	<2015	2015	2016	2017	2018	2019	Subtotal	>2019	Total Citations
Artificial Intelligence	0	0	0	0	0	47	47	1	48
IEEE Access	0	0	0	0	0	41	41	3	44
Proceedings of the National Conference on Artificial Intelligence	14	2	1	3	12	5	23	0	37
Proceedings of the National Conference on Artificial Intelligence	16	0	0	0	11	7	18	0	34
Science Robotics	0	0	0	5	9	19	33	0	33
Frontiers of Information Technology and Electronic Engineering	0	0	0	5	3	14	22	0	22
Simulation Interoperability Standards Organization - 14th Conference on Behavior Representation in Modeling and Simulation 2005	19	0	1	0	1	0	2	0	21
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial	0	0	0	0	3	14	17	2	19

Intelligence and Lecture Notes in Bioinformatics)									
Autonomous Agents and Multi-Agent Systems	8	3	4	1	1	2	11	0	19
Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS	0	0	0	4	9	5	18	0	18

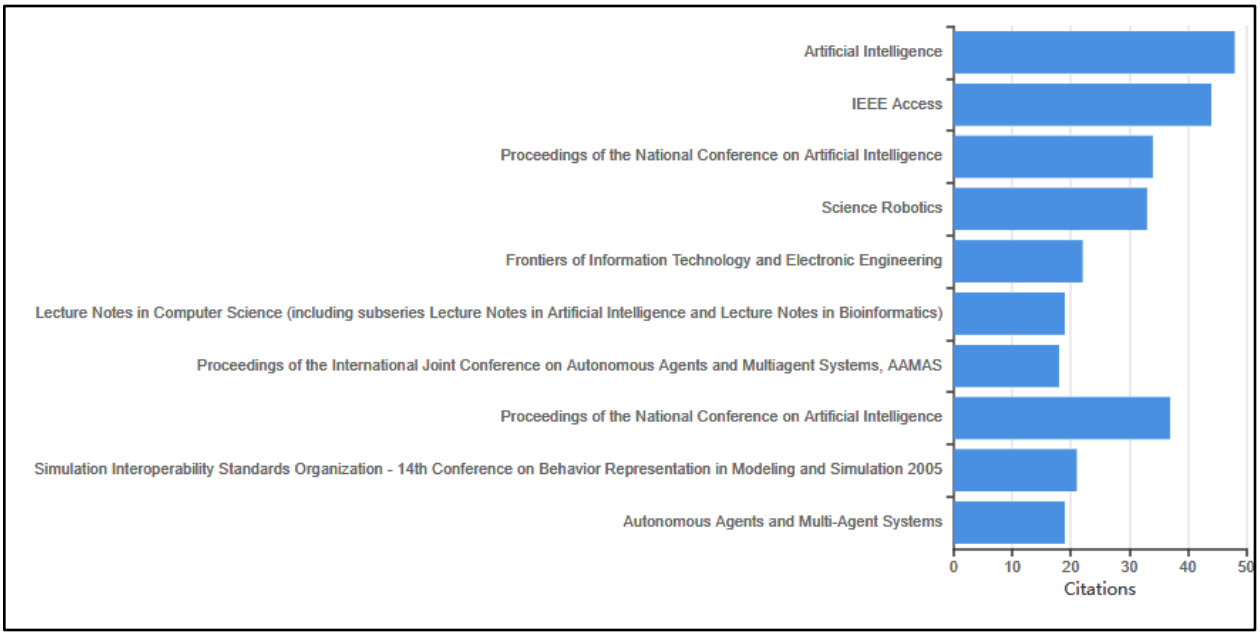

Fig. 18 Journal statistics for publications on topic XAI in the top ten most popular journals. Source: https://www. scopus.com/ (accessed on 18 Dec 2019)

4. Conclusion and future work

From the bibliometric study it is observed that many researchers are working in the domain of image vision forgery detection using deep learning and explainable AI. The earlier information presented with the help of graphs, it is clear that many journals are publishing their work in this research domain. Now a days visual information is widely spreading across the globe so it is important to check the truthfulness of such information due to which image forgery detection is the most valued field of research today. The review process shows that research gaps are there for image region tampering or identification detection and also the percentage by which the region is tamped or forged and also in the field of explainable AI for computer vision. But still, a wide range of things can be assimilated in this field and work should stay in development forever.

5. Limitations of the present study

This bibliometric study consider just Scopus based publications resulting with the blending of keywords used by researcher. There may be other publications and journals from databases such as Web of Science, Google Scholar and PubMed which are not considered during data analysis of this study, so they could not be incorporated in this study. Next limitation is with the counting of citations which were elicited from Scopus database only. Different research databases mostly shows different statistics of citations. In spite of that, Scopus database continues as one of the main search databases which are available in the field of exploring, comparing and tracking citations [17]. And lastly this research is limited to just English language.

References

1. Krawetz, N. (2007) 'A pictures worth digital image analysis and forensics', Black Hat Briefings, pp.1–31.
2. Piva, A. (2013) 'An overview on image forensics', ISRN Signal Processing.
3. Bayram, S., Sencar, H.T. and Memon, N. (2008) 'A survey of copy-move forgery detection techniques', IEEE Western New York Image Processing Workshop, Citeseer.
4. Nizza, M. and Lyons, P.J. (2008) 'In an Iranian image, a missile too many', The Lede, The New York Times News Blog.
5. Lian, S. and Zhang, Y. (2010) 'Multimedia forensics for detecting forgeries', Handbook of Information and Communication Security, Springer, pp.809–828.
6. https://timesofindia.indiatimes.com/times-fact-check/news/fake-viral-photo-of-pm-modi-touching-sonia-gandhis-feet/articleshow/68850648.cms
7. Fan Yang, Shiva K. Pentyala , Sina Mohseni , Mengnan Du , Hao Yuan , Rhema Linder , Eric D. Ragan , Shuiwang Ji , Xia (Ben) Hu ,"XFake : Explainable Fake News Detector with Visualizations",Proceeding WWW '19 The World Wide Web Conference Pages 3600-3604San Francisco, CA, USA — May 13 - 17, 2019.
8. Savita Walia & Krishan Kumar, "Digital image forgery detection: a systematic scrutiny", Australian Journal of Forensic Sciences, 2018 Pages 488-526.
9. Tarun Kumar , Gourav Khurana "Towards recent developments in the field of digital image forgery detection", Int. J. Computer Applications in Technology, Vol. 58, No. 1, 2018 Pages 1-16.
10. Gajanan K.Birajdar,Vijay H.Mankar,"Digital image forgery detection using passive techniques: A survey",Digital Investigation Volume 10, Issue 3, October 2013, Pages 226-245.
11. Pritchard, A. (1969) Statistical Bibliography, an Interim Bibliography, Clearing House for Federal Scientific and Technical Information. Springfield, 112-124.
12. De Bakker, F. G., Groenewegen, P., & Den Hond, F. (2005). A bibliometric analysis of 30 years of research and theory on corporate social responsibility and corporate social performance. Business & Society, 44(3), 283-317.tronic Business, 13(1), 38-69.
13. Kadam, S., Bandyopadhyay, P.K., Patil, Y.: Mapping the field through bibliometric analysis of passenger centric railway transportation. Int. J. Autom. Logist. 2, 349–368 (2016).
14. Rosas, S. R. , Kagan, J. M. , Schouten, J. T. , Slack, P. A. , & Trochim, W. M. K. (2011). Evaluating research and impact: A bibliometric analysis of research by the NIH/NIAID HIV/AIDS clinical trials networks. PLoS ONE, 6 (3) Article number (e17428) .

15. Smith, K. , & Marinova, D. (2005). Use of bibliometric modelling for policy making. Mathematics and Computers in Simulation, 69 , 177–187 (1-2 SPEC. ISS.) .
16. De Granda-Orive, J. I. , Alonso-Arroyo, A. , & Roig-Vázquez, F. (2011). Which database should we use for our literature analysis? Web of Science versus SCOPUS. Archivos de Bronconeumologia, 47 (4), 213 .
17. Yisroel Mirsky, Tom Mahler, Ilan Shelef,Yuval Elovici ,"CT-GAN: Malicious Tampering of 3D Medical Imagery using Deep Learning",Published in the 28th USENIX Security Symposium (USENIX Security 2019).
18. https://www.theweek.in/news/health/2019/04/04/Researchers-hack-CT-scans-to-create-fake-cancers-in-imaging.html.
19. Shimon Kogan,Tobias J. MoskowitzMarina Niessner,"Fake News: Evidence from Financial Markets",SSRN's 84 Pages Posted: 31 Aug 2018 Last revised: 22 Apr 2019.
20. Zhen Zhang, Yuan Ren, Xi-Jian Ping, Zhi-Yong He, Shan-Zhong Zhang, 'A survey on passive-blind image forgery by doctor method detection', International Conference on Machine Learning and Cybernetics, IEEE , 05 September 2008.
21. Chaudhari, A., Joshi, R. R., Mulay, P., Kotecha, K., & Kulkarni, P. (2019). Bibliometric Survey on Incremental Clustering Algorithms. *Library Philosophy and Practice*, 1-23.