International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Bibliographic Analysis on Image Classification using Transfer Learning

Dhatrika Bhagyalaxmi[1], B. Sekhar Babu[2]

Submitted: 13/11/2022 **Accepted:** 15/02/2023

Abstract*:* In this paper we conducted a systematic literature study on image classification using transfer learning techniques during last five years using bibliometric methods. Transfer learning is an important method to classify images using existing neural network architectures. These architectures were developed using Convolutional Neural Networks concept of Deep Learning. The analysis is carried on the standard SCOPUS dataset and analyzed using VOSviewer software. The study is limited to publications in English language on "Transfer Learning", "Deep Learning", "Convolutional Neural Network" and "Image Classification" covering Computer and Engineering subjects during the years 2017 to 2021. This paper's research will aid relevant researchers in understanding the current state of development and trends in this area. Only a few literature studies have tracked the growth of this field, and even fewer have used bibliometric approaches or scientific maps. As a result, this work provides an updated evaluation of this fast-growing subject, using a bibliometric technique to highlight new breakthroughs using scientific maps, and a unique visualization to depict the thematic network structure and progress.

Keywords: *Deep Learning, Convolutional Neural Networks, Transfer Learning, Image Classification, Bibliographic analysis*

1. Introduction

Bibliometric analysis is a statistical evaluation of published scientific papers, books, or book chapters, and it is a useful method of determining the impact of publishing in the scientific community [1-4]. The number of times a piece of research has been cited by other authors can be used to determine its academic impact [5-7]. Research in the field of image classification using transfer learning has grown in the last few years. So, in this research, Bibliometric analysis is employed in this field.

Bibliometric approaches are commonly used to evaluate scientific articles in order to discover research trends [22,23]. As a form of bibliometric approach, the citation index displays the number of times an article has been cited by other papers [24]. As a result, citation analysis aids researchers in gaining a preliminary understanding of publications and research that have an impact in a specific field of interest, and it is concerned with the evaluation of documents cited by scholarly works [25,26].

In this analysis, the important and essential research trends in the field of Image Classification, as well as the most relevant research fields on which Transfer Learning has a substantial impact, were determined using bibliometric approaches. Scopus database and VOSViewer software are used to conduct a systematic study and a thorough review of the literature with the purpose of determining the current state of the art in the application of ML, DL, and CNN in the diagnosis and classification of plant diseases, as well as identifying trends. As a result, the purpose of this study is to integrate and organize existing research by identifying fundamental and key works, obtain an understanding, and highlight future directions in this research area.

Koneru Lakshmaiah Education Foundation,Guntur,Andhra Pradesh,India
bhagyakmp519@gmail.com,
Research Scholar,, Department of CSE,Scholar,
Koneru Lakshmaiah Education Foundation,Guntur,Andhra Pradesh,India
sekharbabu@kluniversity.in

2. Literature Study

Deepak S. et.al, proposed a model for the classification of medical images which is a combination of Convolutional Neural Network (CNN) feature with and Support Vector Machine (SVM) [8]. This was applied on MRI images of Figshare dataset, which consists of three classes of brain tumors. This model attained an accuracy of 95.82% with less computation comparing with transfer learning.

To lower the computing cost while maintaining competitive classification, an evolutionary deep learning framework based on transfer learning is proposed by Wang

B. et.al, [9]. The proposal is to develop a CNN block from smaller datasets, then increase the evolved block's capacity to handle larger datasets. And the model produced good CNNs with less than 40 GPU hours.

In [10], Choudhary T. et.al presented a strategy on deep CNN-based transfer learning for histopathology image classification, along with structured filter pruning, to reduce the trained deep learning models' run-time resource requirements. The pruned model, which has fewer resource requirements, can easily be used in point-of-care devices for automated diagnosis.

The goal of the research in [11] is to classify malignant kinds of skin cancer using an image-processing-based Deep DCNN for human health. Using several pre-trained DCNN models and the transfer learning technique, this research effort automatically detects melanoma and non-melanoma skin cancer kinds.

According to the literature, in recent years, most of the researchers have studied Convolutional Neural Networks (CNN) in a variety of applications, including image classification, feature extraction, and picture segmentation. Plant disease detection is one of these applications, as plant disease is one of the most significant factors causing poor production in the agricultural industry [12-16].

According to medical research, manual classification with human-assisted support, can lead to incorrect prediction and diagnosis. The variety and similarity of tumors and normal tissues are the key reasons behind this. Deep Learning (DL) algorithms have recently shown promise in improving the accuracy in identifying and classifying human diseases using Magnetic Resonance Imaging (MRI) [17-21] [27-29].

Zhang D. et.al introduced a transfer learning technique for remote sensing image categorization based on pre-trained EfficientNet models with a fine tuning strategy [30]. EfficientNet achieved current state-of-the-art performance with a much smaller number of parameters than other recent image classification models.

Unlike deep learning, which requires a huge number of training samples, transfer learning transfers the weights of a pre-trained deep neural network and only uses tiny sample data to achieve good results in UAV image identification. Furthermore, the transfer learning-based image detection model has produced good detection results in terms of accuracy, recall, and F1-score [31-33].

3. DATABASE

Popular sources like Scopus, Web of Science, Scimago, and Google Scholar are used to collect data about publications and their citations. In this paper SCOPUS database is used to gather the details of publications on Image Classification using Transfer Learning with additional keywords during the last five years. Following query used for the analysis and it resulted 526 documents.

(TITLE-ABS-KEY (“Transfer Learning”) AND TITLE-ABS-KEY (“Deep Learning”) AND TITLE-ABS-KEY (

“Convolutional Neural Network”) AND TITLE-ABS-KEY (“image classification”)) AND (LIMIT-TO (PUBSTAGE , “final”)) AND (LIMIT-TO (SUBJAREA

, “COMP”) OR LIMIT-TO (SUBJAREA , “ENGI”)) AND (LIMIT-TO (LANGUAGE , “English”)) AND (LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017))

4. Statistical Analysis

A.Documents by year of publications

In this analysis the publications are considered for the last five years 2017 to 2012. The publications include Articles, Book Chapters, and Conferences etc. The graph in Fig. 1 shows year wise statistics.

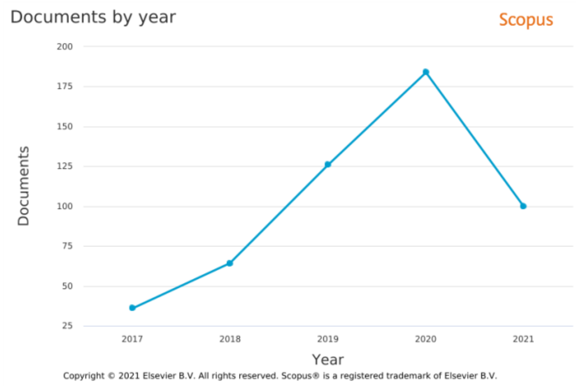

Fig. 1. Documents by year of publication

B. Documents by source

The documents collected from the Scopus database includes sources like IEEE Access, ACM etc., Among difference sources, Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics is with 29 documents followed by IEEE Access with 24 documents. The Fig. 2 shows details of other sources.

C. Documents by Author

Among the 526 selected publications, 159 authors have contributed their research articles with maximum five articles by Castelli, M., Chen, S.C., Kandel, I each.

D. Documents by Affiliation

In this analysis we found that there are 160 affiliations with a maximum of five documents from each including – Ministry of Education China, Florida International University, Chinese Academy of Sciences, Harbin Institute of Technology, etc.,

E. Documents by Country

In the Scopus database of 526 selected documents, researches from 69 countries have their publications with 102 authors from India and 99 from China.

F. Documents by Type

The publications under study are of five categories – Conference papers are of 295, Articles are 190 and the remaining from Review, Book Chapter, and Conference

Review. A pie chart showing this information is presented in the following Fig. 3.

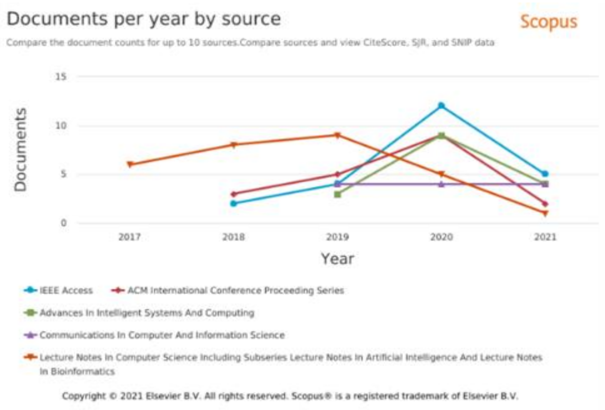

Fig. 2. Documents by Source

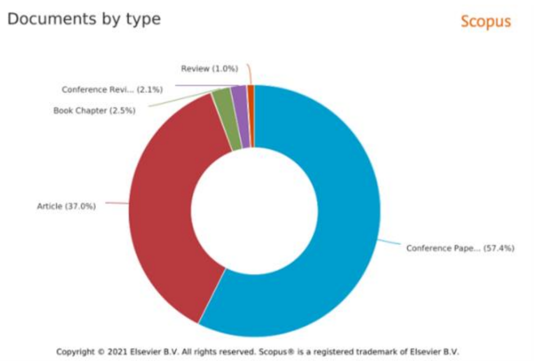

Fig. 3. Documents by Type

G.Documents by Subject Area

Scopus database has resulted 1267 for the query and this analysis is limited to Computer Science and Engineering. These subjects filtered the publications to 710 with Computer Science 444 and Engineering 266. The following pie chart in Fig. 4 shows particulars of the remaining subjects.

This part of the analysis is based on – Co-authorship, Co-occurrences, Citations, Bibliographic Coupling, and Co-citation.

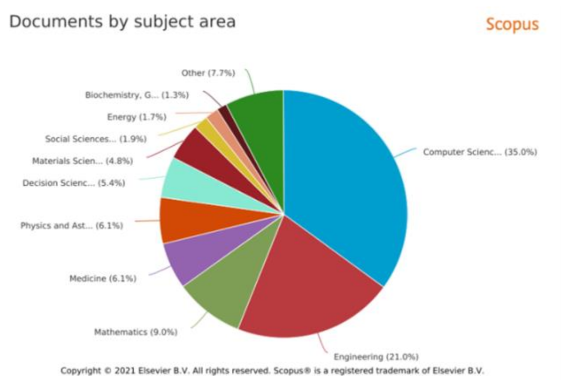

Fig. 4. Documents by Subject Area

5. Network Analysis

This part of the analysis is based on – Co-authorship, Co-occurrences, Citations, Bibliographic Coupling, and Co-citation.

A. Co-authorship analysis

The Co-authorship is analyzed in combination with

Authors, Organizations, and Countries.

1) Co-authorship in combination with Authors: In analyzing the Co-authorship with Authors, the threshold set for documents of an Author are 2. For this threshold, 178 authors meet out of 1715 authros. The largest set of connected items consists of 89 out of 178 and the network diagram is shown in Fig. 5.

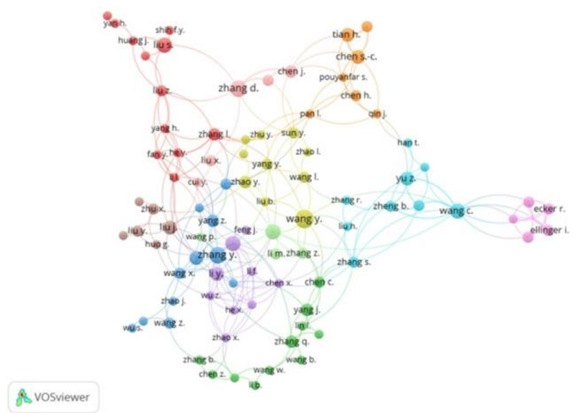

Fig. 5. Network Visualization of Co-authorship with Authors

2) Co-authorship in combination with Organizations: The publications under study are from 1009 organizations. There 31 organizations that meet the threshold of documents of an organzation as 2. But some of these 31 itmes are not connected to each other. The largest set of connected items consists of 4 items and that is shown in Fig. 6.

3) Co-authorship in combination with Countries: In this study the publications are from 71 countries and 42 countries meet if the threshold of documents of a country is three. In these 42 countries, the largest set of connected items consists of 38 items and the network is shown in the Fig. 7.

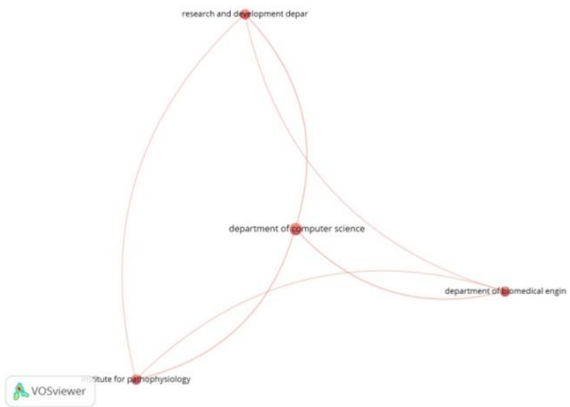

Fig. 6. Network Visualization of Co-authorship with Organizations

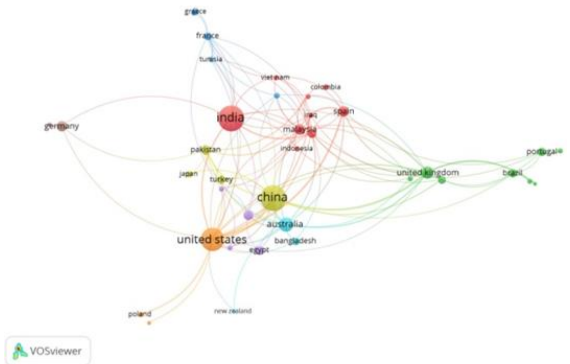

Fig. 7. Network Visualization of Co-authorship with Countries

B. Co-occurrence analysis

The Co-occurrence analysis consists of analysis on All keywords, Author keywords, and Index keywords.

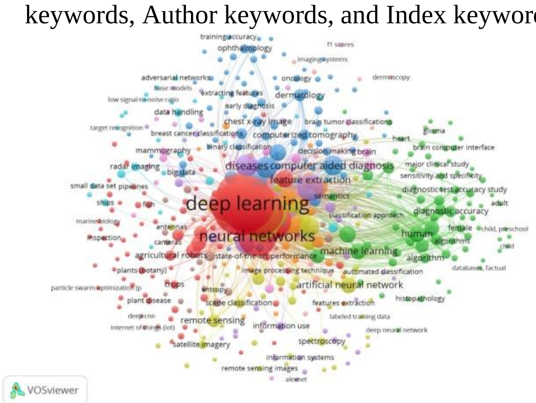

Fig. 10. Network Visualization of Co-occurrence with Index Keywords

1) Co-occurrence in combination with All keywords: There are 3489 keywords in the selected publications and the no. of occurrences of keyword is set to 3. The network is shown in Fig. 8.

2) Co-occurrence in combination with Author keywords: There are 1050 keywords in the selected publications and the no. of occurrences of keyword is set to 3. There are 91 entries that meet the threshold and the network is shown in Fig. 9.

3) Co-occurrence analysis in terms of index keywords:

For index keywords, out of 2806 keywords in the selected publications, 449 meet the threshold of min. no. of occurrences of a keyword 3. The network shown in Fig. 10, the keyword Deep Learning is having max. links 448 and total link strength 4975 with occurrences 404. Another keyword Transfer Learning with links 448 and total link strength 4807.

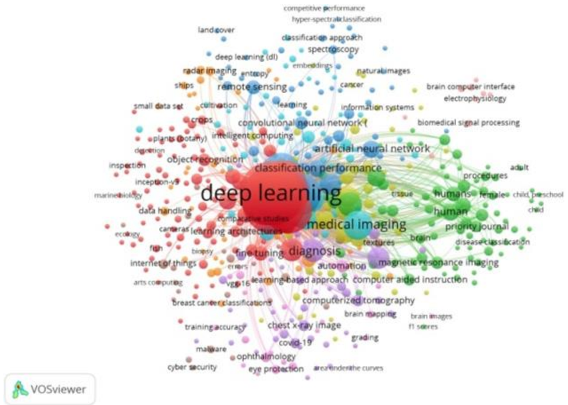

Fig. 8. Network Visualization of Co-occurrence with All keywords

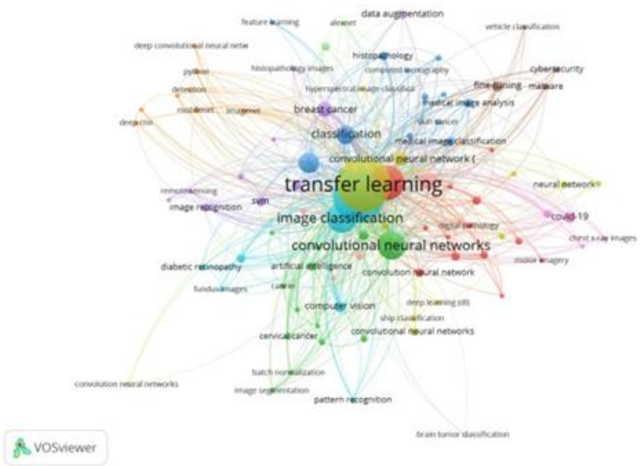

Fig. 9. Network Visualization of Co-occurrence with Author Keywords

C. Citation analysis

This citation analysis is carried with unit of analysis –

Documents, Sources, Authors, Organizations, and Countries.

1) Citation analysis of Documents: Among 526 publications, 169 documents meet the threshold of minimum 4 citations per document. In these, Apostolopoulos i.d. is in first position with 385 citations followed by Scott g.j. with 173 citations. But, some of the

items in the network are not connected to each other. The network in Fig. 11 shows all items.

2) Citation analysis of Source: The documents are collected from 289 sources. Lecture Notes in Computer Science is with 29 and IEEE Access is with 24 documents are the top two in the list.

3) Citation analysis of Author: In 526 documents 1715

authors have contributed their research work. Considering min. no. of documents of an author as two, we have 178 authors. Want Y. is in the top of the list with 9 documents with 30 citations. An interesting observation is with author Zhang H. with 6 documents with maximum 232 citations.

4) Citation analysis on Organization: There are 1009 organizations from where the Authors have their constributions. School of Computing and Information Sciences, Florida International University is with 5 documents and 34 citations.

5) Citation analysis on Country: The documents selected for this analysis are from 71 countries. In these, India, China and United States are the top 3 with 102, 100 and 84 documents respectively. But interestingly citations of these countries are 446, 1037, and 1133 respectively.

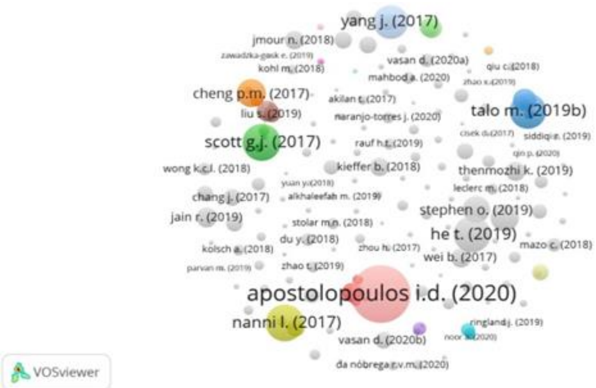

Fig. 11. Citation analysis of documents

D. Bibliographic coupling analysis

Kessier defined bibliographic coupling as a "measure to establish a similarity of link between documents, institutions, and authors." The bibliographic coupling network may be built and visualized using VOSviewer. This analysis is also having unit of analysis similar to Citation analysis – Documents, Sources, Authors, Organizations, and Countries.

1) Documents: Of the 526 documents, there are 246 documents which meet the threshold of min. no. of citations of documents as 2. The author Alzubaidi I. is in the top of the list with 14 citations and 490 link strength followed by Dhillon A. with 52 citations and 429 link strength. The coupling network is shown in Fig. 12.

2) Sources: In the 516 documents, top 5 journals, no. of articles, citations, and their link strength are shown in the Table 1. From the table it is evident that, IEEE Access and Lecture notes in Computer Science are the important sources for publication of research articles on Image Classification and Transfer Learning. There are 55 documents which meet the threshold of min. no. of documents as 2. The coupling network is shown in Fig. 13.

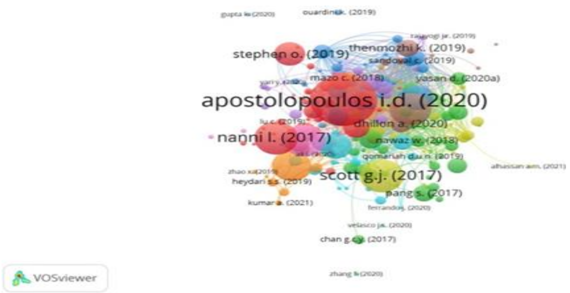

Fig. 12. Coupling network of documents

3) Authors: The observations made during the coupling analysis with authors are, there are 1715 authors have participated in 516 publications. The authors Al-Shamma O., Alzubaidi I., Duan Y., Fadhel M.A., Santamria J., and Zhang

J. are with 2 documents, 29 citations, and 8679 link strength. The author Wang Y. is having maximum documents 9, and Zhang H. is with maximum citations 232. The network is shown in Fig. 14 is with min. no. of documents threshold set to 2.

4) Countries: In this analysis researchers from 71 countries have participated and if the threshold min. no. of documents per country is set to two, then 50 countries meet the threshold. The countries China, United States and India are in the front line with 25810, 23298, and 19723 link strength. The network analysis is shown in Fig. 15.

4. Conclusion

The bibliometric analyis on Image Classification and Transfer Learning exposed the research trends and potential of this domain. The keywords used in this analysis to extract papers from the SCOPUS database are Deep Learning, Transfer Learning, Convolutional Neural Network, Image Classification. There are 526 papers found from 2017 to 2021 by limiting the language to English. Major contributions are from Conference proceedings and Journals with subject areas Computer Science and Engineering. The results revealed that the China, India and United States were the most productive countries in this study, with the highest number of published articles in the globe. In the analysis of authors, there were 1715 authors in all papers with a maximum of 7 documents of an author. The maximum number of citations of an author is 232 with link strenght of 14 and United States and China authors from are top in the citations. The analysis highlighted the most up-to-date analytical methodologies, such as Neural Networks and Transfer Learning, as well as the most popular and emerging issue areas, such as Deep Convolutional Networks and its accuracy in Image Classification applications like identifying cancer cells, blockages, tumors, and crop diseases.

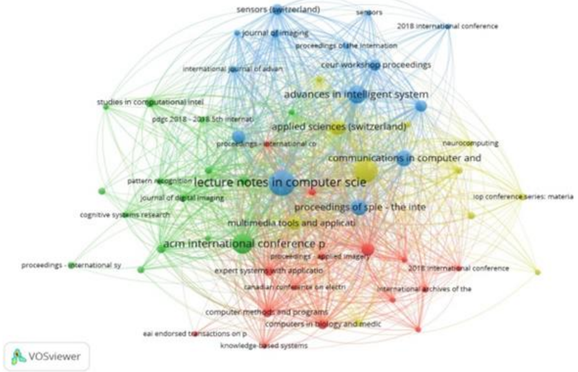

Fig. 13. Coupling network of documents

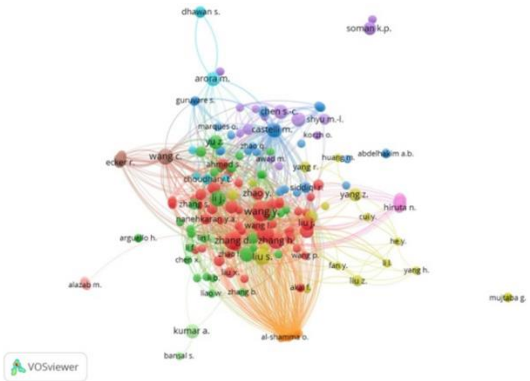

Fig. 14. Coupling network of Authors

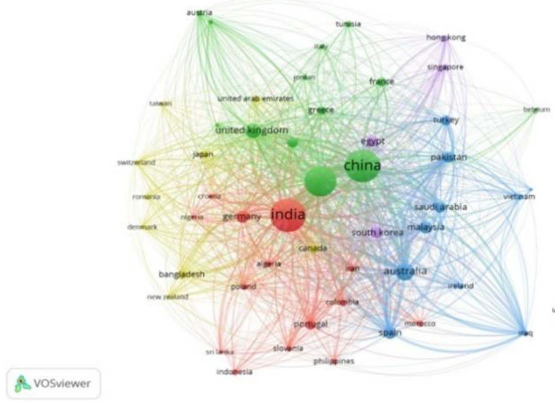

Fig. 15. Coupling network of Countries

Table 1. Top 5 Sources Based On The Link Strength

Document Source	Documents	% of documents	Citations	Total link strength
IEEE Access	24	4.65	194	4050
Lecture notes in Computer Science	29	5.62	165	2867
Applied Sciences	11	2.13	71	2270
ACM International Conference Proceeding Series	19	3.68	37	1858
Proceedings of SPIE – the International Society for Optical Engineering	12	2.32	16	1283

References

[1] Diagnostic criteria and classification of hyperglycemia first detected in pregnancy: a World Health Organization guideline. http://repositorio.uchile.cl/bitstream/handle/2250/129629/Diagnostic- criteria-and-classification-of-hyperglycaemia-first-detected-in- pregnancy-A-World-Health-Organization-Guideline.pdf?sequence=1. Diabetes Res Clin Pract. 2014:341–363.

[2] Proceedings of the fourth international work-shop-conference on gestational diabetes mellitus. Metzger BE, Coustan DR. https://www.ncbi.nlm.nih.gov/pubmed/9704245 Diabetes Care. 1998; Suppl 2:0–167.

[3] Scientometric data files. A comprehensive set of indicators on 2649 journals and 96 countries in all major science fields and subfields 1981-1985. Schubert A, Glänzel W, Braun T. Scientometrics. 1989; 16:3–478.

[4] Coverage analysis of Scopus: a journal metric approach. de Moya- Anegón F, Chinchilla-Rodríguez Z, Vargas-Quesada B, Corera- Álvarez E, Muñoz-Fernández F, González-Molina A. Scientometrics. 2007; 73:53–78.

[5] The role of energy nutrients, foods, and dietary patterns in the development of gestational diabetes mellitus: a systematic review of observational studies. Schoenaker D, Mishra G, Callaway L, Soedamah-Muthu S. Diabetes Care. 2015; 39:16–23.

[6] 100 top-cited scientific papers in limb prosthetics. Eshraghi A, Osman N, Gholizadeh H, Ali S, Shadgan B. Bio Med Eng OnLine. 2013; 12:119.

[7] Trends in the quality of highly cited surgical research over the past 20 years. Brooke B, Nathan H, Pawlik T. Ann Surg. 2009; 249:162–167.

[8] Deepak, S., Ameer, P.M. Automated Categorization of Brain Tumor from MRI Using CNN features and SVM. J Ambient Intell Human Comput 12, 8357–8369 (2021). https://doi.org/10.1007/s12652-020- 02568-w

[9] Bin Wang, Bing Xue, and Mengjie Zhang. 2021. A transfer learning based evolutionary deep learning framework to evolve convolutional neural networks. In Proceedings of the Genetic and Evolutionary Computation Conference Companion (GECCO '21). Association for Computing Machinery, New York, NY, USA, 287–288. DOI:https://doi.org/10.1145/3449726.3459455

[10] Choudhary, Tejalal & Mishra, Vipul & Goswami, Anurag & Sarangapani, Jag. (2021). A Transfer Learning with Structured Filter Pruning Approach for Improved Breast Cancer Classification on Point-of-Care Devices. Computers in Biology and Medicine. 134. 10.1016/j.compbiomed.2021.104432.

[11] V. Mishra, A. K. V and M. Arora, "Deep convolution neural network based automatic multi-class classification of skin cancer from dermoscopic images," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 800- 805, doi: 10.1109/ICICCS51141.2021.9432160.

[12] E. Hirani, V. Magotra, J. Jain and P. Bide, "Plant Disease Detection Using Deep Learning," 2021 6th International Conference for Convergence in Technology (I2CT), 2021, pp. 1-4, doi: 10.1109/I2CT51068.2021.9417910.

[13] Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk and Darko Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification", Computational Intelligence and Neuroscience, 2016.

[14] Sue Han Lee, Hervé Goëau, Pierre Bonnet and Alexis Joly, "New perspectives on plant disease characterization based on deep learning", Computers and Electronics in Agriculture, vol. 170, 2020, ISSN 0168-1699.

[15] Jayme G.A. Barbedo, Factors influencing the use of deep learning for plant disease recognition Biosystems Engineering, vol. 172, no. 2018, pp. 84-91, ISBN 1537–5110.

[16] Maeda-Gutiérrez, Valeria & Galván Tejada, Carlos & Zanella Calzada, Laura & Celaya Padilla, Jose & Galván Tejada, Jorge & Gamboa-Rosales, Hamurabi & Luna-Garcia, Huizilopoztli & Magallanes-Quintanar, Rafael & Carlos, Guerrero-Mendez & Olvera- Olvera, Carlos. (2020). Comparison of Convolutional Neural Network Architectures for Classification of Tomato Plant Diseases. Applied Sciences. 10. 1245. 10.3390/app10041245.

[17] Logeswari, T., Karnan, M. An improved implementation of brain tumor detection using segmentation based on hierarchical self organizing map (2010) International Journal of Computer Theory and Engineering, 2 (4), p. 591.

[18] Tandel, Gopal S., Mainak Biswas, Omprakash G. Kakde, Ashish Tiwari, Harman S. Suri, Monica Turk, John R. Laird, Christopher K. Asare, Annabel A. Ankrah, N. N. Khanna, B. K. Madhusudhan, Luca Saba, and Jasjit S. Suri 2019. "A Review on a Deep Learning Perspective in Brain Cancer Classification" Cancers 11, no. 1: 111. https://doi.org/10.3390/cancers11010111

[19] Zeineldin, R.A., Karar, M.E., Coburger, J. et al. DeepSeg: deep neural network framework for automatic brain tumor segmentation using magnetic resonance FLAIR images. Int J CARS 15, 909–920 (2020). https://doi.org/10.1007/s11548-020-02186-z

[20] Swati, Zar & Zhao, Qinghua & Kabir, Muhammad & Ali, Farman & Ali, Zakir & Ahmad, Saeed & Lu, Jianfeng. (2019). Brain tumor classification for MR images using transfer learning and fine-tuning. Computerized Medical Imaging and Graphics. 75. 34-46. 10.1016/j.compmedimag.2019.05.001.

[21] M. Gurbină, M. Lascu and D. Lascu, "Tumor Detection and Classification of MRI Brain Image using

Different Wavelet Transforms and Support Vector Machines," 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), 2019, pp. 505-508, doi: 10.1109/TSP.2019.8769040.

[22] Zitt M, Ramanana-Rahary S, Bassecoulard E. Relativity of citation performance and excellence measures: from cross-field to cross-scale effects of field-normalization. Scientometrics. 2005;63(2):373–401.

[23] Li LL, et al. Global stem cell research trend: bibliometric analysis as a tool for mapping of trends from 1991 to 2006. Scientometrics. 2009;80(1):39–58.

[24] Fooladi M, et al. Do criticisms overcome the praises of journal impact factor? Asian Soc Sci. 2013; 9(5):176–82.

[25] Ale Ebrahim N, et al. Equality of google scholar with web of science citations: case of Malaysian engineering highly cited papers. Mod Appl Sci. 2014; 8(5):63–9.

[26] Gomez-Jauregui V, et al. Information management and improvement of citation indices. Int J Inf Manage.2014; 34(2):257–71.

[27] Andrea Caroppo, Alessandro Leone, Pietro Siciliano, Deep transfer learning approaches for bleeding detection in endoscopy images, Computerized Medical Imaging and Graphics, Volume 88, 2021, 101852, ISSN 0895-6111, https://doi.org/10.1016/j.compmedimag.2020.101852.

[28] Ghosh, T., Chakareski, J. Deep Transfer Learning for Automated Intestinal Bleeding Detection in Capsule Endoscopy Imaging. J Digit Imaging 34, 404–417 (2021). https://doi.org/10.1007/s10278-021- 00428-3.

[29] H.C. Shin, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R.M. Summers, Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. IEEE Transactions on Medical Imaging 35(5), 1285 (2016).

[30] D. Zhang, Z. Liu and X. Shi, "Transfer Learning on EfficientNet for Remote Sensing image Classification," 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), 2020, pp. 2255-2258, doi: 10.1109/ICMCCE51767.2020.00489.

[31] W. Meng and M. Tia, "Unmanned Aerial Vehicle Classification and Detection Based on Deep Transfer Learning," 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), 2020, pp. 280-285, doi: 10.1109/ICHCI51889.2020.00067.

[32] C Aker and S. Kalkan, "Using deep networks for drone detection", In: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), pp. 1-6, 2017

[33] C Tan, F Sun, T Kong, W Zhang, C Yang and C. Liu, A survey on deep transfer learning. International Conference on Artificial Neural Networks. ICANN, pp. 270-279, 2018.