University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

2-3-2021

A Bibliometric Analysis of Plant Disease Classification with Artificial Intelligence based on Scopus and WOS

Shivali Amit Wagle
Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University) (SIU), Pune, India, kulkarni_shivali@yahoo.co.in

Harikrishnan R
Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University) (SIU), Pune, India, rhareish@gmail.com

A Bibliometric Analysis of Plant Disease Classification with Artificial Intelligence based on Scopus and WOS

Shivali Amit Wagle

Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University) (SIU), Pune, India
kulkarni_shivali@yahoo.co.in
and

Harikrishnan R

Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University) (SIU), Pune, India
rhareish@gmail.com

ABSTRACT

The maneuver of Artificial Intelligence (AI) techniques in the field of agriculture help in the classification of diseases. Early prediction of the disease benefits in taking relevant management steps. This is an important step towards controlling the disease growth that will yield good quality products to fulfill the global food demand. The main objective of this paper is to study the extent of research work done in this area of plant disease classification. The paper discusses the bibliometric analysis of plant disease classification with AI in Scopus and Web of Science core collection (WOS) database in analyzing the research by area, influential authors, institutions, countries, and funding agency. The 1125 research documents are extracted from 2010 to 9th Jan 2021 from both the database. Bibliometric analysis is the statistical analysis of the research published as articles, conference papers, and reviews, which helps in understanding the impact of publication in the research domain globally. The visualization analysis is done with open-source tools namely GPS Visualizer, Gephi, VOS viewer, ScienceScape, minivan, and wordcloud. The visualization aids in a quick and clear understanding of the different perspectives as mentioned above in a particular research domain search.

Keywords

bibliometric analysis, classification, deep learning, machine learning, plant disease

1. INTRODUCTION

Agriculture is the backbone of any country as it provides the food needs of the growing population across the globe. For the requirement of good quality and quantity of the agricultural produce, the agricultural products need to be protected from the disease at an earlier stage. Various viral and fungal diseases affect the plant. The weather conditions and seasonal changes cause variation in the level of temperature, humidity, wind, etc. These changes have an impact on the plants prone to certain diseases. The disease can cause a substantial amount of loss in the quality and quantity of the yield of the plants further resulting in a financial loss to the farmers. Certain steps should be taken by the farmers to protect the farm from the disease that can be prevented if the cause of the disease is known in advance. An appropriate care needs to be taken to avoid the major effects on plants as they, in turn, show their effect on the product quality, quantity, and productivity. The fast and accurate identification of disease severity will help to reduce yield losses. The traditional technique of inspection of the disease is cumbersome and affluent jobs (Arivazhagan et al. 2013; Bashish, Braik, and Bani-ahmad 2010). The diagnosis of diseases requires a large cultivating area to be inspected by the experts. If interpreted incorrectly, treating the plants may not be sufficient enough to save and reduce the diseases in them (Ferentinos 2018). Generally, the farmers used to spray pesticides or chemical fertilizers to get rid of the diseases. This harms the crop along with the person in contact with it. Sometimes using some basic things like plucking the diseased leaf and burning it or using organic fertilizers also help in solving the problem.

Artificial Intelligence plays a major role in overcoming the traditional techniques with fast, reliable, automated, cost-effective, and also most importantly accurate methods to detect disease in the plant. The computer-based image processing technology applied in agricultural engineering research is fruitful in terms of results. Automatic technique detecting the disease symptoms in the leaves of the plant is valuable as it reduces a lot of manual work and saves time (Singh et al., 2017). The powerful technique with the computational system of image processing and machine learning can detect and diagnose the disease caused to the plants (Mohanty et al., 2016; Yang & Guo, 2017).

The computation techniques that are used in the plant disease classification are Decision Trees, Support Vector Machines (SVMs), Artificial Neural Networks (ANNs), Machine Learning (ML), and Deep Learning (DL). The evolution of DL techniques has shown significantly better results compared to the shallow ML algorithm (H. Chen et al., 2018). The introduction to the deep learning model techniques in the field of classification and detection was introduced by Lecun et al. (Yann et al., 1998) with the basic deep learning tool of Convolutional Neural Networks (CNNs). CNNs is a dynamic model that helps in classification applications with a large amount of data.

There is a study on soil taxonomic class which helps in further understanding the soil for farmland purpose and further the condition of a pathogen that can affect the plant with the disease. Multiple machine learning models are compared for predicting the soil taxonomic classes in different areas of the USA (Brungard et al., 2015). These areas are at diverse locations and study helps in investigating digital soil mapping. Heung et al., 2016 compared CART, CART with bagging, Random Forest, k-nearest neighbor, nearest shrunken centroid, ANN, multinomial logistic regression, logistic model trees, and SVM for predicting soil taxonomic conditions. Topography, climate, and vegetation of the study area were some of the covariates in the analysis showing their impact on the soil-environmental relationships. The ML models of maximum entropy, SVM, and ANN were used to find the pattern in land activities that supports knowledge of predicting the land terrain. The statistical results showed that the spatial combination of the main drivers namely farmlands, 0.06–0.2 range in normalized difference vegetation index, rocks with interbedded limestone, and other susceptible classes therein can make at least a prone area of about 30% to land sliding in China (W. Chen et al., 2017). The evaluation of landslide ML models for the Himalayan region in India are built and validated with ensemble models and compared based on the Chi-square test (Pham et al., 2017).

Machine learning for precision agriculture utilizes supervised support vector machine (SVM), unsupervised networks like k-means, and Self-organizing map for the classification and clustering purpose respectively (Behmann et al., 2015). The CNN model developed by (Sladojevic et al., 2016) performs the distinguishing of plant leaf from the surrounding. This further is used to analyze if the plant leaf is healthy or diseased. The model achieved precision between 91% and 98%, for separate class tests, on average 96.3% accuracy. Different machine learning algorithms were developed by (Zhu et al., 2017) for the classification of tobacco mosaic virus disease and heathy class concerning effective wavelength, textural features, and data fusion. The performance was better with data fusion as compared to textural features and achieved an accuracy of 95% with BPNN. A machine learning framework with K-high resolution feature maps is capable to separate the visual symptoms comprising of stress severity, its intensity, and further quantification of the stress (Ghosal et al., 2018). The stress caused by the bacterial or fungal diseases and due to deficiency of the nutrient for 25,000 images of plant leaf was evaluated by the authors. Griffel et al., 2018 used SVM for the classification of potato plants with potato virus Y. The multiclass SVM was implemented with the color space segmentation of the images of the citrus plant in the automated detection of disease (Sharif et al., 2018).

The effectiveness of the LeNet CNN model in the classification of background concerning resolution, size, orientation was developed by (Amara et al., 2017). The VGG16 model with transfer learning is compared with the shallow networks with fine-tuning and from scratch models are compared by (Wang et al., 2017). The accuracy of 90.4% is achieved in classification with the VGG16 model. A CNN model with a high success rate is beneficial and can be used as an early warning tool. This approach helps to support real-time conditions in cultivation with an integrated supported system (Ferentinos, 2018). The fast and accurate results were obtained by (Rangarajan & Purushothaman, 2018) in the detection of the disease on the plant leaf. The botanists are now benefitted from the advances in science and technology with computer vision approaches in the plant identification task. Several perspectives have been proposed in the literature for the classification of plants. Too et al., 2019 compared VGG 16, Inception V4, ResNet with 50, 101, and 152 layers, and DenseNets with 121 layers for the classification of 38 different classes from the PlantVillage dataset with diseased and healthy classes. DenseNet performed extremely well with an accuracy of 99.75% amongst the other networks. Arnal Barbedo, 2019 in their work for the identification of multiple diseases affecting the same leaf of a plant using segmentation of data using deep learning techniques. The adapted model of 3D deep CNN for hyperspectral image data for soybean crops with economically important diseases affecting them (Nagasubramanian et al., 2019). A two-stage neural network proposed for plant disease classification focused on the largest dataset of 79265 leaf images comprising of inconsistent background and various weather conditions. The trained model achieved an accuracy of 93.67% (Arsenovic et al., 2019). Picon et al., 2019 in their work used adapted deep residual neural network-based algorithms to deal with the detection of multiple plant diseases in real-time acquisition conditions for wheat diseases. The hyperparameter of CNN was fine-tuned with particle swarm optimization for the correct classification of the model (Darwish et al., 2020). Residual learning for learning significant features and attention mechanism as the input to it was applied by (Karthik et al., 2020). For detecting the infection in tomato leaves. The deep learning process is strongly used in the detection of leaf disease in plants and further categorized the diseases (Vaishnnave et al., 2020). The main objective of the paper is to report a bibliometric analysis of the plant disease classification with deep learning and machine learning literature.

2. PRELIMINARY DATA COLLECTION

The publication database is usually categorized as open access and paid access ones (Ahumada & Villalobos, 2009) that can be viewed through the library portals of the university or by registering separately on individual websites (Chaudhari et al., 2019; Kadam et al., 2016). The widespread methods in retrieving the publication database are Scopus, Clarivate, SCImago, Mendeley, ScienceDirect, DBLP, Google Scholar, and Research Gate, etc. The citation database is with a wide variety of fields domains viz science, technology, engineering, medicine, social sciences, arts, and humanities. The data is fetched from the Scopus database and WOS on 9th January 2021. The tables, graphs, and clusters are based on the accessed date. The paper reflects the Scopus database and WOS for the bibliometric analysis with the significant keywords.

2.1 Significant keywords

The Significant keywords searched are “Plant disease” and “Classification” and “Deep Learning” and “Machine Learning”. The keywords in the search of Scopus and WOS are enumerated in Table 1. The correct incorporation of keywords benefits to target the significant research areas in terms of the number of publications (NoP). "Deep Learning" is the weightiest keyword that is used along with the significant keywords in this research area.

Table1: Keywords with the number of publications

Keywords (Scopus)	NoP	Keywords (WOS)	NoP
Deep Learning	114	Classification	84
Plant Disease	112	Prediction	63
Machine Learning	78	Model	44
Image Processing	66	Regression	43
Convolutional Neural Network	57	System	32
Neural Networks	57	Identification	30
Learning Systems	55	Selection	27
Plants (botany)	54	Spatial prediction	25
Classification (of Information)	53	Models	22

2.2 Preliminary Data Analysis

The literature documents were retrieved as articles, conference papers, conference reviews, review, book chapters, data paper, editorial material, meeting abstract, early access, and news item for the duration from 2010 to 9th January 2021. Trends for yearly publications are shown in Figure 1. The research in this area was not explored more till 2016 and there was a steep rise in research since the year 2017. There were 342 research papers published in WOS and 108 in Scopus in the year 2020.

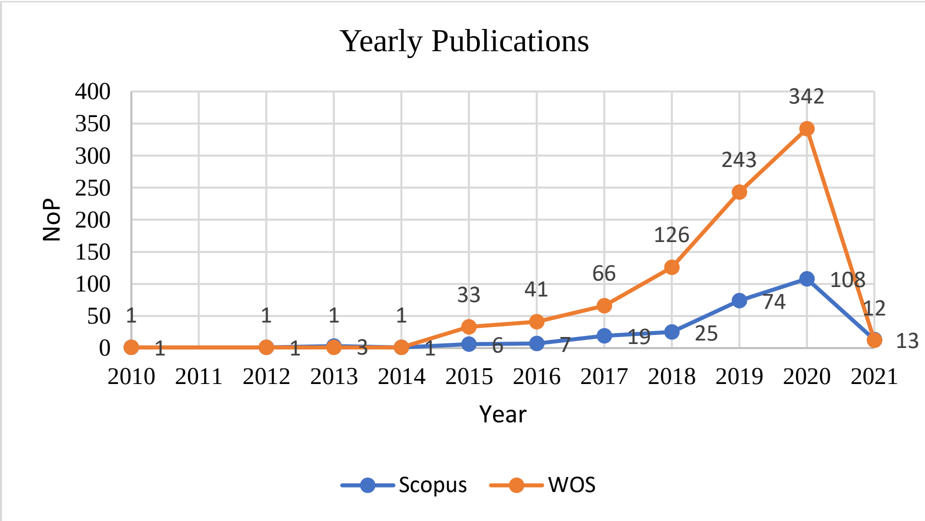


Figure 1: Annual publishing movement in plant disease classification

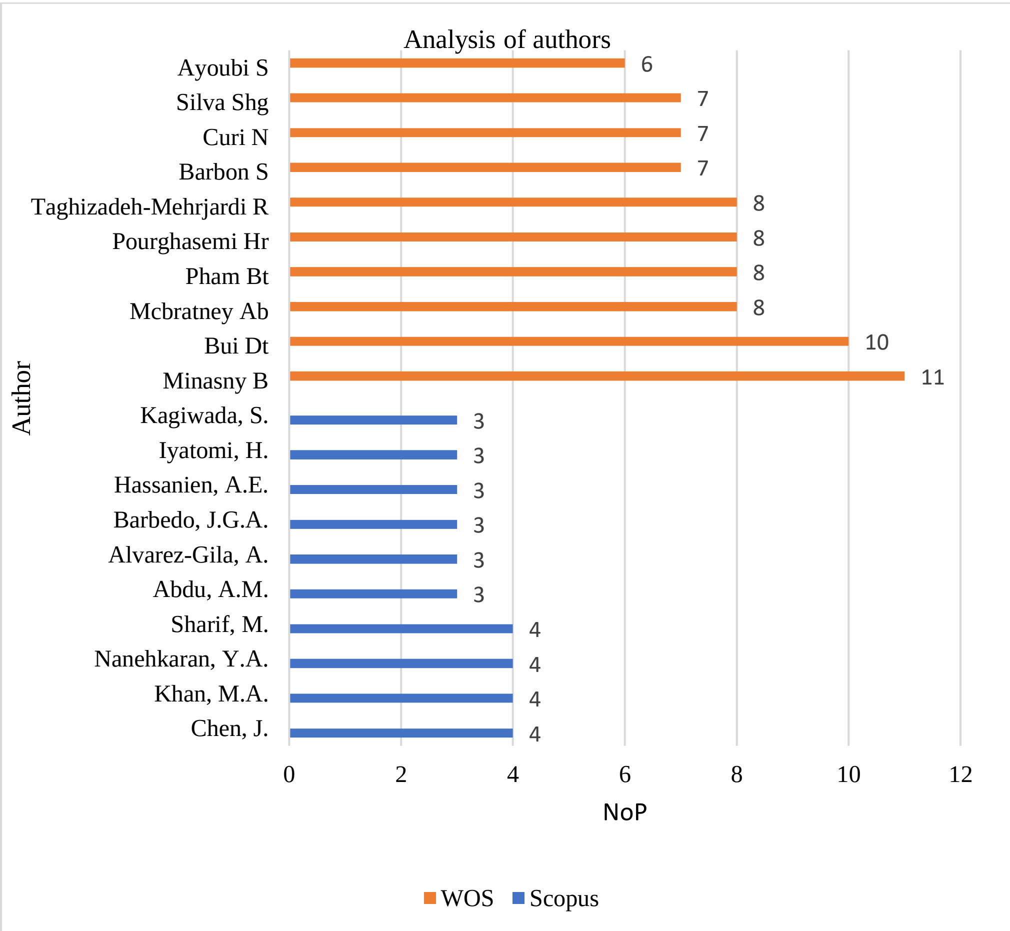


Figure 2: Key contributing authors in Scopus and WOS in plant disease classification.

The top ten influential authors contributing to the area of plant disease classification in Scopus and WOS are portrayed in Figure 2. Minasny B is an influential author with 11 publications followed by Bui Dt with ten publications in WOS from 2010 to 2020; and Chen, J., Khan, M.A., Nanehkaran, Y.A., and Sharif published four papers each in Scopus in this research domain.

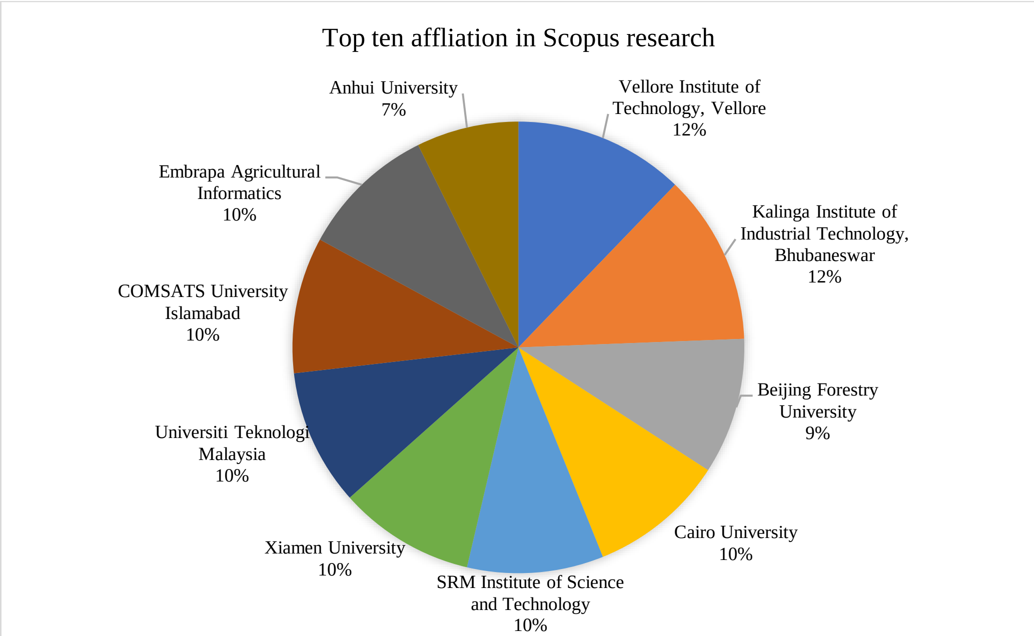


Figure 3: Top ten affiliations in plant disease classification in Scopus

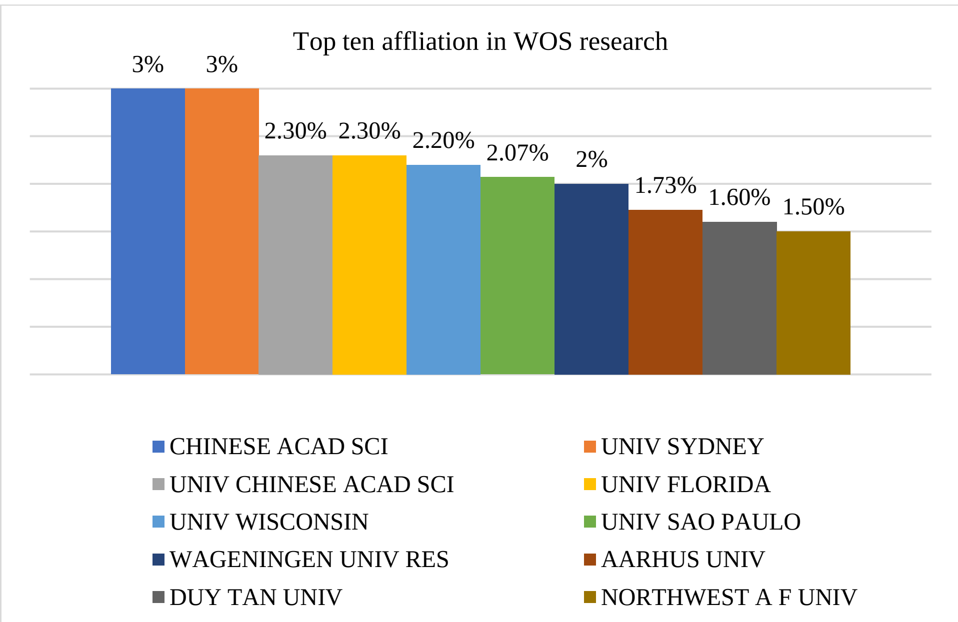


Figure 4: Top ten affiliations in plant disease classification in WOS

Figure 3 and Figure 4 show the top ten affiliations in Scopus and WOS respectively. Vellore Institute of Technology, Vellore, and Kalinga Institute of Industrial Technology, Bhubaneswar is the top affiliated institutes publishing in the Scopus database on plant disease classification. Chinese ACAD SCI and University Sydney are top affiliated institutes publishing in WOS. The types of documents for publication are shown in Figure 5 and Figure 6 for Scopus and WOS respectively. The publications are highest in article type followed by documents in the conference paper, conference review, review, book chapter, data paper in Scopus. The publications are maximum in an article in WOS with 811 papers. The other category consists of papers in a book chapter, editorial material, meeting abstract, and news item.

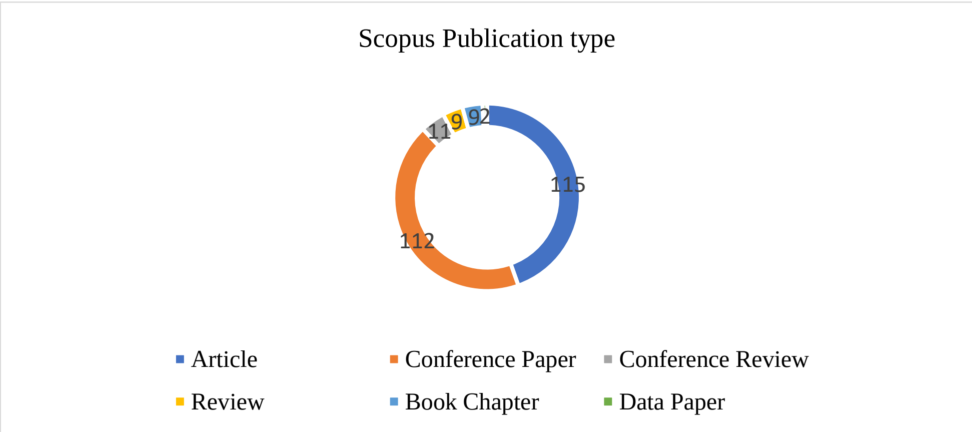


Figure 5: Types of documents for publications in Scopus in plant disease classification.

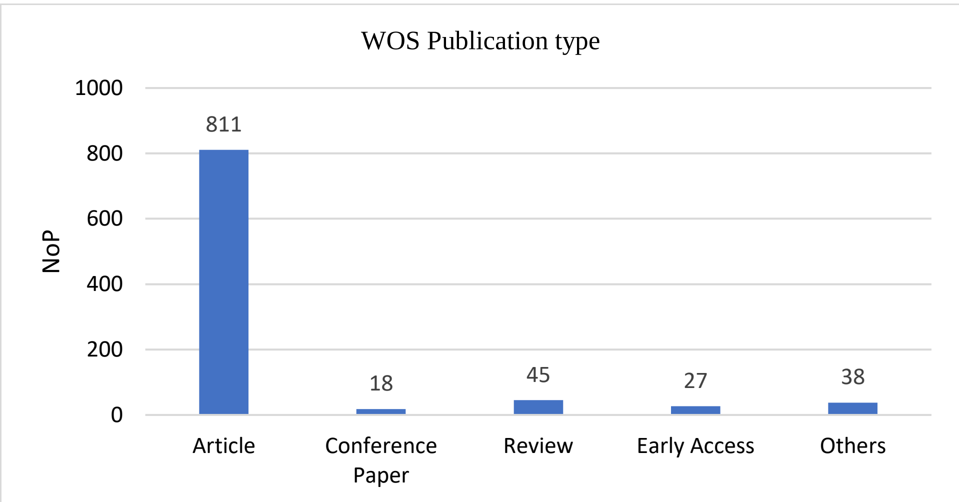


Figure 6: Types of documents for publications in WOS in plant disease classification.

The top ten countries partaking the publications in the area of plant disease classification in Scopus and WOS respectively are shown in Figure 7. India leads in the publication with 42.24% trailed by China with 16.28% and the US with 8.14% in terms of Scopus publications and US lead with 22.37% followed by China with 17.18% and Australia with 12.23% for WOS publications.

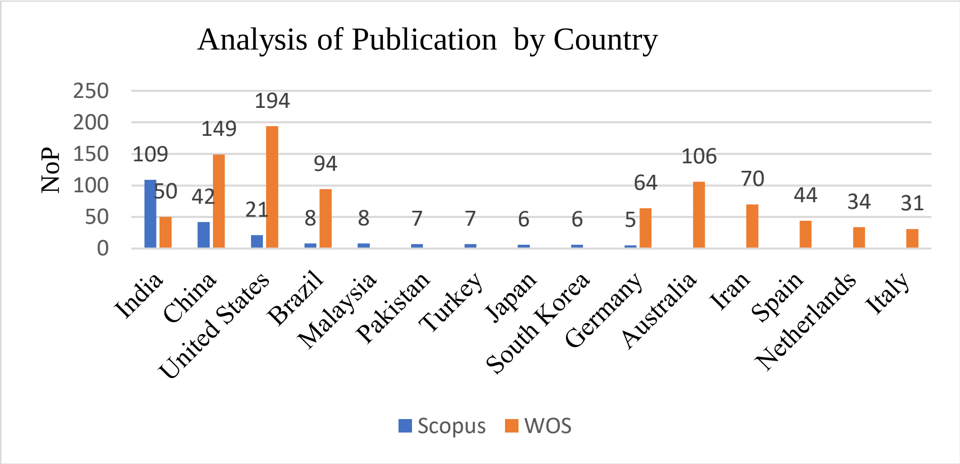


Figure 7: Top ten countries partaking in plant disease classification from Scopus and WOS publications.

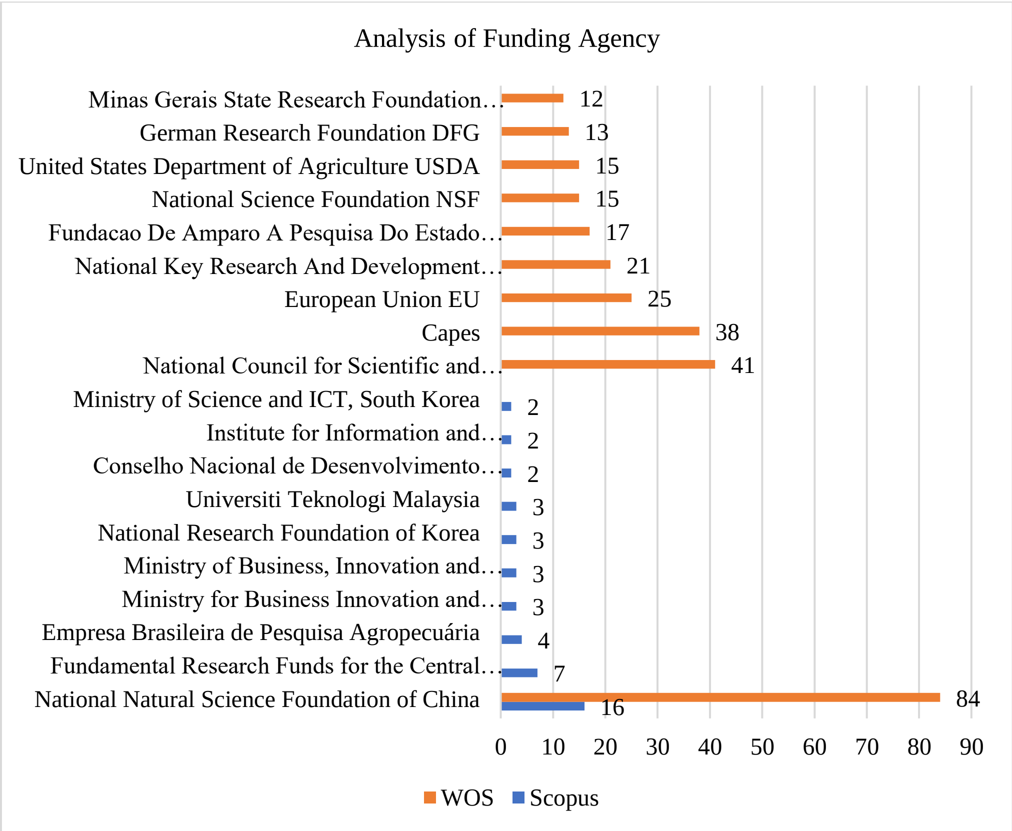


Figure 8: Top ten Funding agency plant disease classification in Scopus and WOS

The analysis of the top ten funding agency for plant disease classification is shown in Figure 8. The papers under maximum funding are by the National Natural Science Foundation of China with 84 and 16 publications in WOS and Scopus respectively.

Table 4: Top ten subject areas of extracted literature.

Subject Area (Scopus)	NoP	Subject Area (WOS)	NoP
Computer Science	185	Computer Science	234
Engineering	109	Water Resources	68
Agricultural and Biological Sciences	61	Food Science Technology	62
Mathematics	45	Plant Sciences	51
Physics and Astronomy	24	Geology	43
Decision Sciences	23	Meteorology Atmospheric Sciences	34
Biochemistry, Genetics and Molecular Biology	16	Forestry	31
Medicine	12	Environmental Sciences Ecology	30
Environmental Science	11	Veterinary Sciences	23
Energy	10	Chemistry	17

The subject area in plant disease classification publications is shown in Table 4. The maximum research is in “Computer Science” followed by “Engineering” and “Agriculture and Biological Sciences”. “Computer Science” followed by “Water Resources” and “Food Science Technology” in WOS.

The document by source analysis for the top five sources is shown in Figure 9. The CiteScore for 2010-2019 is shown in Figure 9(a). “Computers and Electronics in Agriculture” is having the highest CiteScore of 6.7 in the year 2019. The CiteScore of this journal is high throughout. SCImago journal rank per year (SJR) is shown in Figure 9(b). “Computers and Electronics in Agriculture” is having an SJR value of 1.058 in the year 2019. Source normalized impact per paper by year (SNIP) is shown in Figure 9(c). “Computers and Electronics in Agriculture” is having the highest SNIP of 2.356 in the year 2013. The journal

“Computers and Electronics in Agriculture” is the top journal for publishing high-quality research.

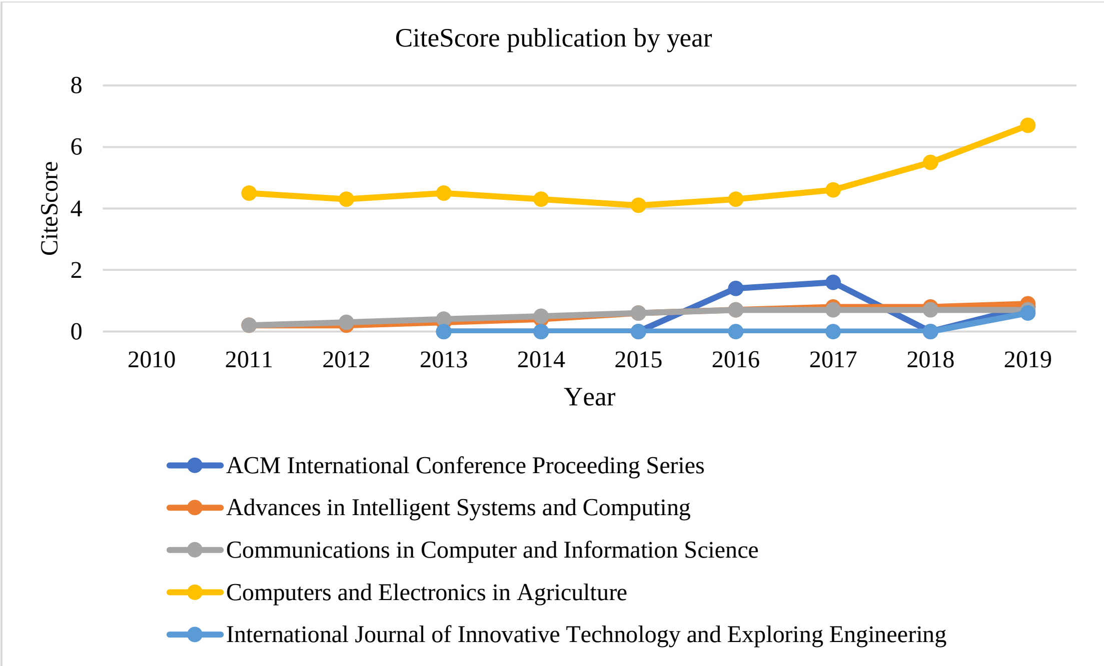


Figure 9 (a): CiteScore publication by year in plant disease classification for Scopus.

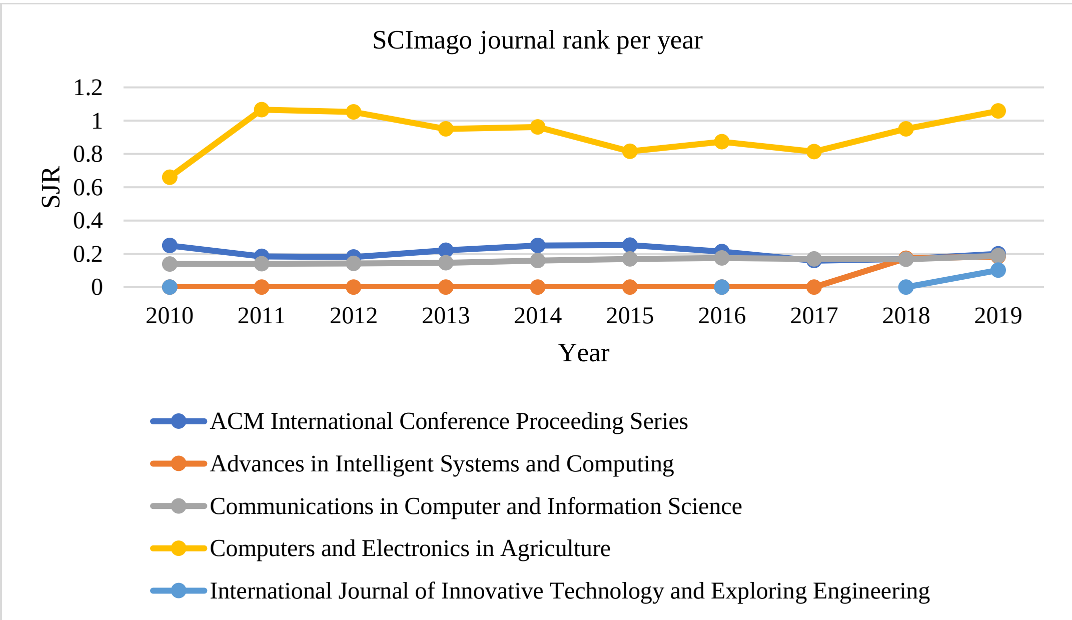


Figure 9 (b): SCImago journal rank per year in plant disease classification for Scopus.

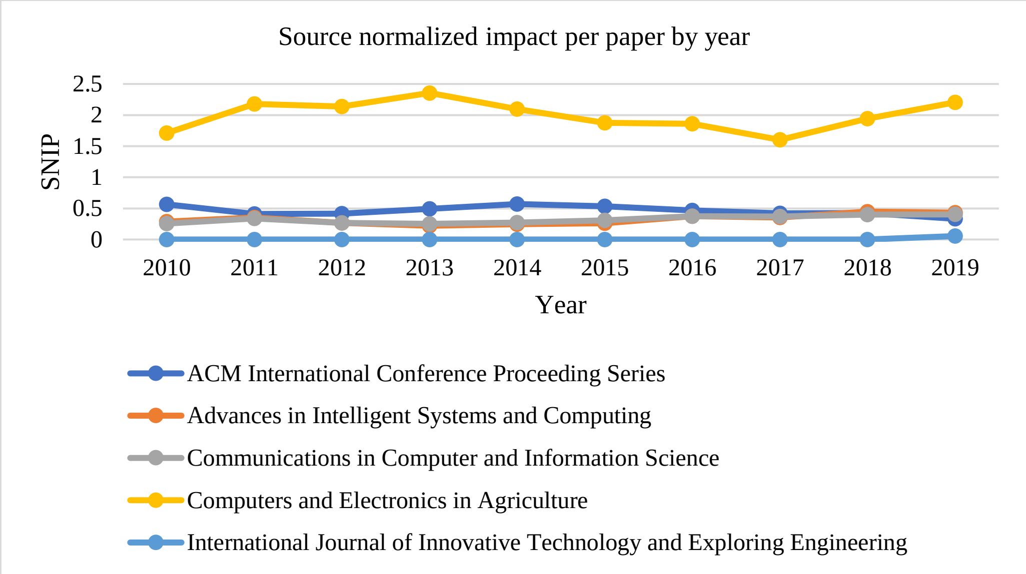


Figure 9 (c): Source normalized impact per paper by year in plant disease classification for Scopus.

3. BIBLIOMETRIC ANALYSIS

The plant disease classification, with its uniqueness of literature and prominent researches with involved metrics therein, the bibliometric analysis is listed below:

3.1 Geographical Region Analysis

Figure 10. Geographic locations of the study plant disease classification

The maximum number of publications are from India, China, and the United States (US) in Scopus and the US, China, and Australia in WOS. The geographical locations worldwide where the research on plant disease classification is carried out is shown in Figure 10 using the GPS Visualizer tool of gpsvisualiser.com. This gives a global outlook towards the research.

3.2 Network Visualization

The divergent statistical parameters relationship is visualized graphically with “Gephi” an open-source software. Gephi enables the clustering of the network data with the use of certain filtering, navigation, and management. The various parameters that are used are the year of publication, publication title, authors, keywords used by them, journal, citations of the publication are revealed with nodes and edges. The "Fruchterman Reingold" and "Yifan Hu" layout is used with different manual adjustments. The cluster with more nodes and edges are quite dense. ScienceScape and minivan are used to cluster with parameters of author and author keyword; source title and author keyword coappearing in the same paper. The VOS viewer tool is also used for visualization of co-occurrences of the keywords in the publication. The words that there used in the publication in the plant disease classification are visualized with wordcloud.com This gives information about the significant terms in the research domain. Figure 11 – Figure 22 shows the different parametric amalgamations for a plant disease classification for the data extracted from Scopus and WOS.

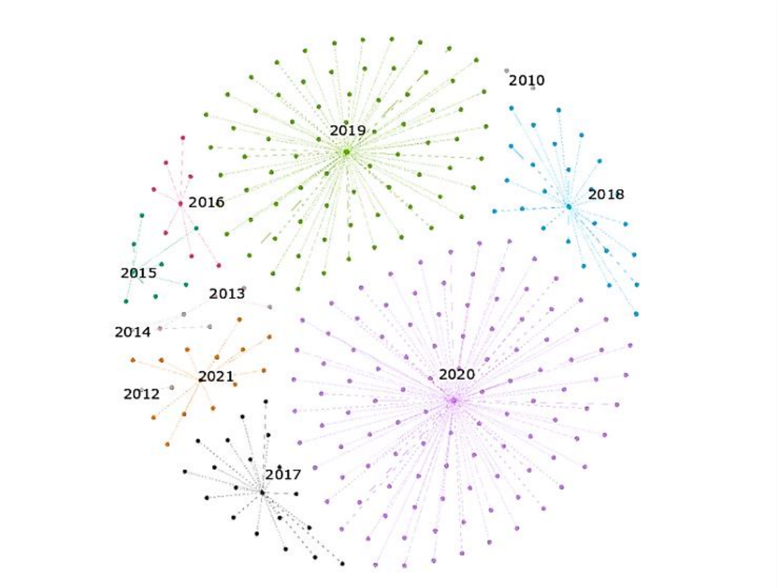


Figure 11: The Cluster of year and publication title Scopus

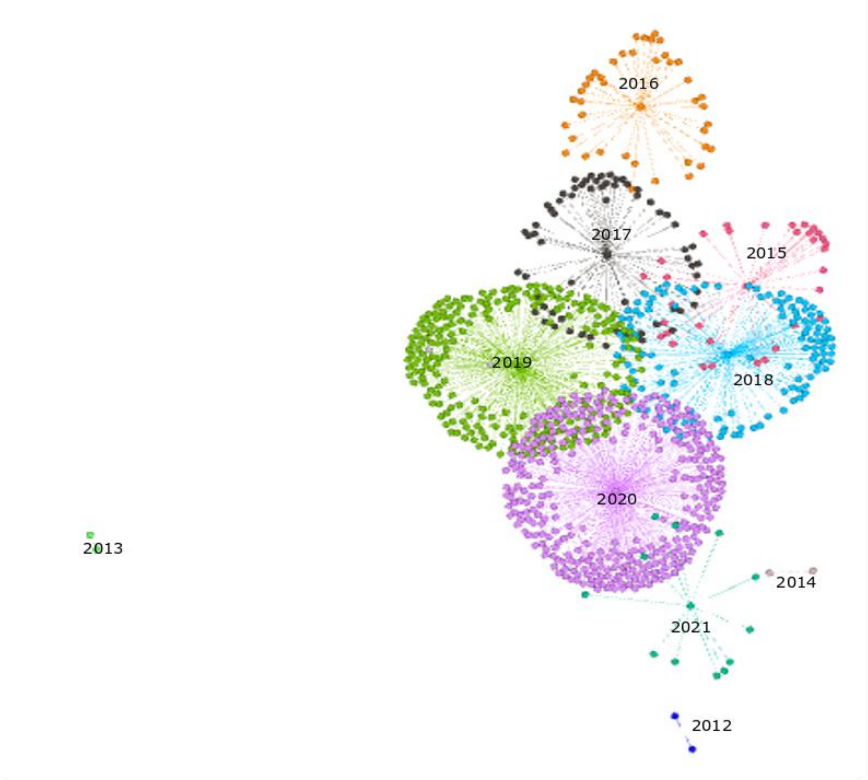


Figure 12: The Cluster of year and publication title WOS

Figure 11 and Figure 12 shows the cluster of yearly publication in Scopus and WOS respectively. There were few publications till 2016 in Scopus and WOS. The was an increase in the research since 2016. In both cases, maximum publications were seen in the year 2020. Figure 13 shows the Sankey Graph – of main authors, main keywords, and main journals in the publications. The first column shows the influential authors, the second column shows the significant keywords they use and the third column shows the journal where the research is published. The interconnection of the terms is seen through the link. Figure 14 shows the tabular information of the Sankey Graph of Figure. 13. The details under each author, keyword, and journal with the NoP can be evaluated. Each of the elements is having a hyperlink, which when clicked gives access to the content. Author Chen J has published 7 research papers in this area. The link gets access to the 7 papers he authored. Similarly, papers can be accessed for other authors, keywords, and journals.

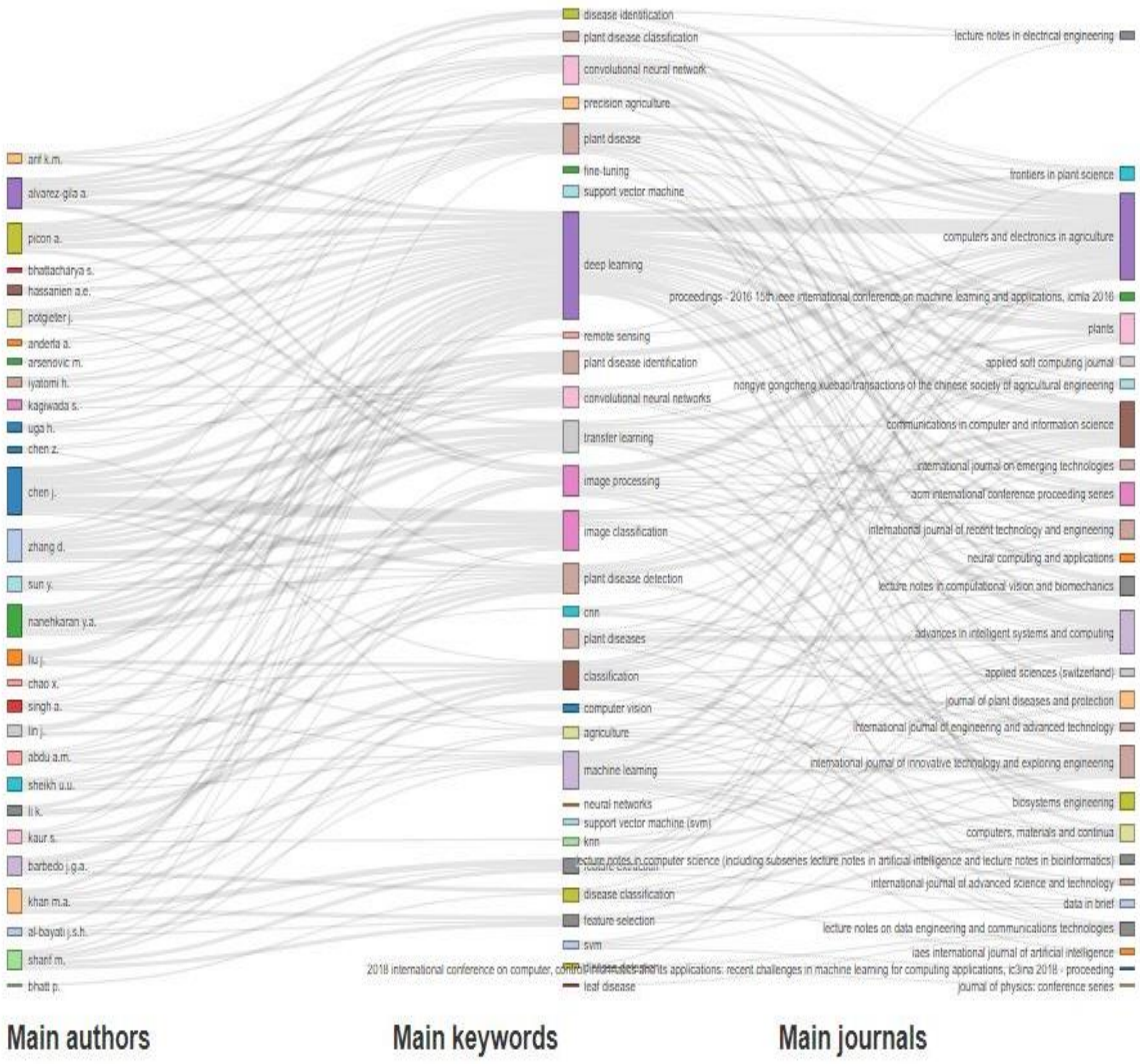


Figure 13: Sankey Graph -Main Authors, main keywords, and main journals.

Main authors

- [no author name available] (11 papers)
- chen j. (7 papers)
- zhang d. (6 papers)
- liu j. (5 papers)
- khan m.a. (4 papers)
- nanehkaran y.a. (4 papers)
- sharif m. (4 papers)
- singh a. (4 papers)
- abdu a.m. (3 papers)
- alvarez-gila a. (3 papers)
- barbedo j.g.a. (3 papers)
- chen z. (3 papers)
- hassanien a.e. (3 papers)
- iyatomi h. (3 papers)
- kagiwada s. (3 papers)
- kaur s. (3 papers)
- li k. (3 papers)
- lin j. (3 papers)
- picon a. (3 papers)
- potgieter j. (3 papers)
- sheikh u.u. (3 papers)
- sun y. (3 papers)
- uga h. (3 papers)
- al-bayati j.s.h. (2 papers)
- anderla a. (2 papers)
- arif k.m. (2 papers)
- arsenovic m. (2 papers)
- bhatt p. (2 papers)
- bhattacharya s. (2 papers)
- chao x. (2 papers)

Main keywords

- deep learning (78 papers)
- machine learning (43 papers)
- classification (35 papers)
- plant disease (33 papers)
- image processing (27 papers)
- convolutional neural network (26 papers)
- feature extraction (18 papers)
- plant disease detection (18 papers)
- transfer learning (18 papers)
- convolutional neural networks (15 papers)
- image classification (15 papers)
- disease classification (13 papers)
- agriculture (12 papers)
- cnn (11 papers)
- plant diseases (11 papers)
- support vector machine (10 papers)
- disease detection (9 papers)
- computer vision (8 papers)
- svm (8 papers)
- support vector machine (svm) (7 papers)
- neural networks (6 papers)
- plant disease identification (6 papers)
- precision agriculture (6 papers)
- disease identification (5 papers)
- feature selection (5 papers)
- fine-tuning (5 papers)
- knn (5 papers)
- leaf disease (5 papers)
- plant disease classification (5 papers)
- remote sensing (5 papers)

Main journals

- computers and electronics in agriculture (18 papers)
- advances in intelligent systems and computing (16 papers)
- communications in computer and information science (9 papers)
- international journal of innovative technology and exploring engineering (6 papers)
- acm international conference proceeding series (5 papers)
- international journal of recent technology and engineering (5 papers)
- plants (5 papers)
- lecture notes in computational vision and biomechanics (4 papers)
- applied sciences (switzerland) (3 papers)
- biosystems engineering (3 papers)
- computational intelligence and neuroscience (3 papers)
- computers, materials and continua (3 papers)
- frontiers in plant science (3 papers)
- lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) (3 papers)
- lecture notes in electrical engineering (3 papers)
- lecture notes on data engineering and communications technologies (3 papers)
- nongye gongcheng xuebao/transactions of the chinese society of agricultural engineering (3 papers)
- 2018 international conference on computer, control, informatics and its applications: recent challenges in machine learning for computing applications, ic3ina 2018 - proceeding (2 papers)
- applied soft computing journal (2 papers)
- data in brief (2 papers)
- iaes international journal of artificial intelligence (2 papers)

Figure 14: Tabular Information of Sankey Graph in Figure. 13.

2010

- plant disease diagnosis 1 paper
- cnns 1 paper
- residual structure 1 paper
- severity estimation 1 paper
- shuffle units 1 paper
- deep learning 0 paper
- machine learning 0 paper
- classification 0 paper
- plant disease 0 paper
- image processing 0 paper

2011

- deep learning 0 paper
- machine learning 0 paper
- classification 0 paper
- plant disease 0 paper
- image processing 0 paper
- convolutional neural network 0 paper
- feature extraction 0 paper
- plant disease detection 0 paper
- transfer learning 0 paper
- convolutional neural networks 0 paper

2012

- machine learning 1 paper
- hyperspectral 1 paper
- cultivar discrimination 1 paper
- phytophthorainfestans 1 paper
- reflectance spectroscopy 1 paper
- deep learning 0 paper
- classification 0 paper
- plant disease 0 paper
- image processing 0 paper
- convolutional neural network 0 paper

2013

- classification 2 papers
- deep learning 1 paper
- machine learning 1 paper
- feature extraction 1 paper
- transfer learning 1 paper
- image classification 1 paper
- disease classification 1 paper
- cnn 1 paper
- disease detection 1 paper
- feature selection 1 paper

2014

- deep learning 1 paper
- plant disease 1 paper
- image processing 1 paper
- convolutional neural network 1 paper
- precision agriculture 1 paper
- disease identification 1 paper
- early pest 1 paper
- multi-label classification 1 paper
- contextual meta-data 1 paper
- contextual meta-data conditional neural network 1 paper

2015

- deep learning 3 papers
- plant disease 2 papers
- image processing 2 papers
- convolutional neural network 2 papers
- precision agriculture 2 papers
- machine learning 1 paper
- convolutional neural networks 1 paper
- plant disease identification 1 paper
- disease identification 1 paper
- data augmentation 1 paper

2016

- machine learning 2 papers
- image processing 2 papers
- transfer learning 2 papers
- image classification 2 papers
- convolutional neural networks 1 paper
- disease classification 1 paper
- disease detection 1 paper
- support vector machine (svm) 1 paper
- remote sensing 1 paper
- deep neural nets 1 paper

2017

- deep learning 5 papers
- classification 4 papers
- machine learning 3 papers
- image processing 3 papers
- disease classification 3 papers
- plant disease 2 papers
- transfer learning 2 papers
- agriculture 2 papers
- texture analysis 2 papers
- convolutional neural network 1 paper

2018

- deep learning 6 papers
- plant disease 4 papers
- machine learning 3 papers
- convolutional neural networks 3 papers
- feature extraction 2 papers
- classification 1 paper
- image processing 1 paper
- convolutional neural network 1 paper
- plant disease detection 1 paper
- transfer learning 1 paper

2019

- deep learning 17 papers
- machine learning 13 papers
- classification 8 papers
- plant disease 8 papers
- image processing 8 papers
- plant disease detection 7 papers
- support vector machine 6 papers
- convolutional neural network 5 papers
- feature extraction 5 papers
- transfer learning 5 papers

2020

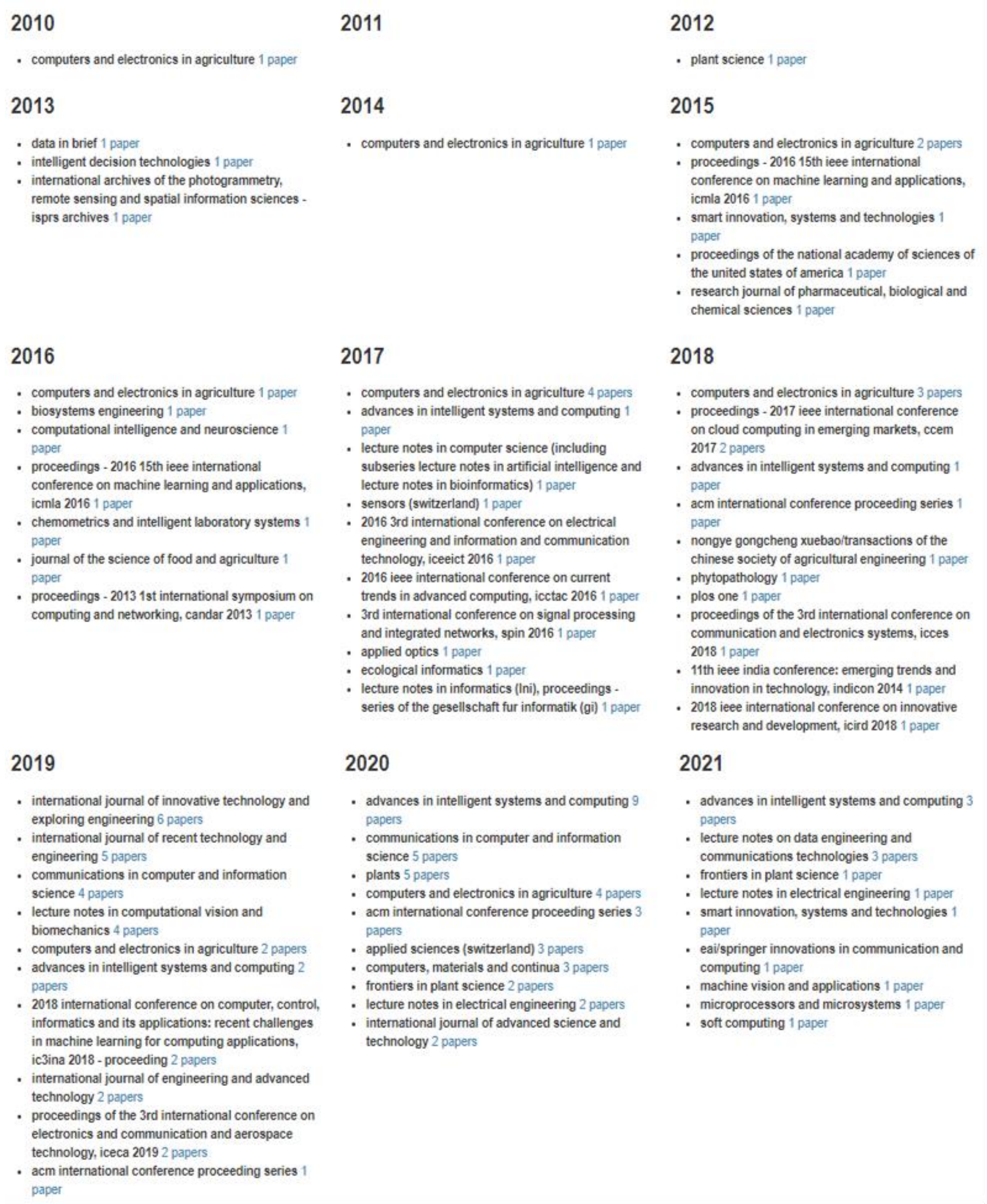
- deep learning 40 papers
- machine learning 19 papers
- classification 19 papers
- plant disease 15 papers
- convolutional neural network 14 papers
- feature extraction 10 papers
- image processing 9 papers
- plant disease detection 9 papers
- image classification 8 papers
- transfer learning 7 papers

2021

- deep learning 5 papers
- convolutional neural network 2 papers
- agriculture 2 papers
- plant diseases 2 papers
- svm 2 papers
- glcm 2 papers
- classification 1 paper
- plant disease 1 paper
- image processing 1 paper
- plant disease detection 1 paper

Figure 15: Top keywords by year.

Figure 15 shows the top keywords by year. The significant keywords used in the research in each year and the NoP with that keywords are seen under them. Since 2015, the deep learning keyword is been used extensively in this research. Figure 16 shows the top journals in each year where publications are done in Scopus. There are 1,2 and 9 papers in Advances in Intelligent Systems and Computing in the year 2018, 2019, and 2020 respectively.



2010
- computers and electronics in agriculture 1 paper

2011

2012
- plant science 1 paper

2013
- data in brief 1 paper
- intelligent decision technologies 1 paper
- international archives of the photogrammetry, remote sensing and spatial information sciences - isprs archives 1 paper

2014
- computers and electronics in agriculture 1 paper

2015
- computers and electronics in agriculture 2 papers
- proceedings - 2016 15th ieee international conference on machine learning and applications, icmla 2016 1 paper
- smart innovation, systems and technologies 1 paper
- proceedings of the national academy of sciences of the united states of america 1 paper
- research journal of pharmaceutical, biological and chemical sciences 1 paper

2016
- computers and electronics in agriculture 1 paper
- biosystems engineering 1 paper
- computational intelligence and neuroscience 1 paper
- proceedings - 2016 15th ieee international conference on machine learning and applications, icmla 2016 1 paper
- chemometrics and intelligent laboratory systems 1 paper
- journal of the science of food and agriculture 1 paper
- proceedings - 2013 1st international symposium on computing and networking, candar 2013 1 paper

2017
- computers and electronics in agriculture 4 papers
- advances in intelligent systems and computing 1 paper
- lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) 1 paper
- sensors (switzerland) 1 paper
- 2016 3rd international conference on electrical engineering and information and communication technology, iceeict 2016 1 paper
- 2016 ieee international conference on current trends in advanced computing, icctac 2016 1 paper
- 3rd international conference on signal processing and integrated networks, spin 2016 1 paper
- applied optics 1 paper
- ecological informatics 1 paper
- lecture notes in informatics (lni), proceedings - series of the gesellschaft fur informatik (gi) 1 paper

2018
- computers and electronics in agriculture 3 papers
- proceedings - 2017 ieee international conference on cloud computing in emerging markets, ccem 2017 2 papers
- advances in intelligent systems and computing 1 paper
- acm international conference proceeding series 1 paper
- nongye gongcheng xuebao/transactions of the chinese society of agricultural engineering 1 paper
- phytopathology 1 paper
- plos one 1 paper
- proceedings of the 3rd international conference on communication and electronics systems, icces 2018 1 paper
- 11th ieee india conference: emerging trends and innovation in technology, indicon 2014 1 paper
- 2018 ieee international conference on innovative research and development, icird 2018 1 paper

2019
- international journal of innovative technology and exploring engineering 6 papers
- international journal of recent technology and engineering 5 papers
- communications in computer and information science 4 papers
- lecture notes in computational vision and biomechanics 4 papers
- computers and electronics in agriculture 2 papers
- advances in intelligent systems and computing 2 papers
- 2018 international conference on computer, control, informatics and its applications: recent challenges in machine learning for computing applications, ic3ina 2018 - proceeding 2 papers
- international journal of engineering and advanced technology 2 papers
- proceedings of the 3rd international conference on electronics and communication and aerospace technology, iceca 2019 2 papers
- acm international conference proceeding series 1 paper

2020
- advances in intelligent systems and computing 9 papers
- communications in computer and information science 5 papers
- plants 5 papers
- computers and electronics in agriculture 4 papers
- acm international conference proceeding series 3 papers
- applied sciences (switzerland) 3 papers
- computers, materials and continua 3 papers
- frontiers in plant science 2 papers
- lecture notes in electrical engineering 2 papers
- international journal of advanced science and technology 2 papers

2021
- advances in intelligent systems and computing 3 papers
- lecture notes on data engineering and communications technologies 3 papers
- frontiers in plant science 1 paper
- lecture notes in electrical engineering 1 paper
- smart innovation, systems and technologies 1 paper
- eai/springer innovations in communication and computing 1 paper
- machine vision and applications 1 paper
- microprocessors and microsystems 1 paper
- soft computing 1 paper

Figure 16: Top journals per year.

The Cluster of co-occurrences of the keywords in the Scopus publication is shown in Figure 17. Plant disease, image processing, and machine learning are the significant and most co-occurring keywords that are linked together in this research.

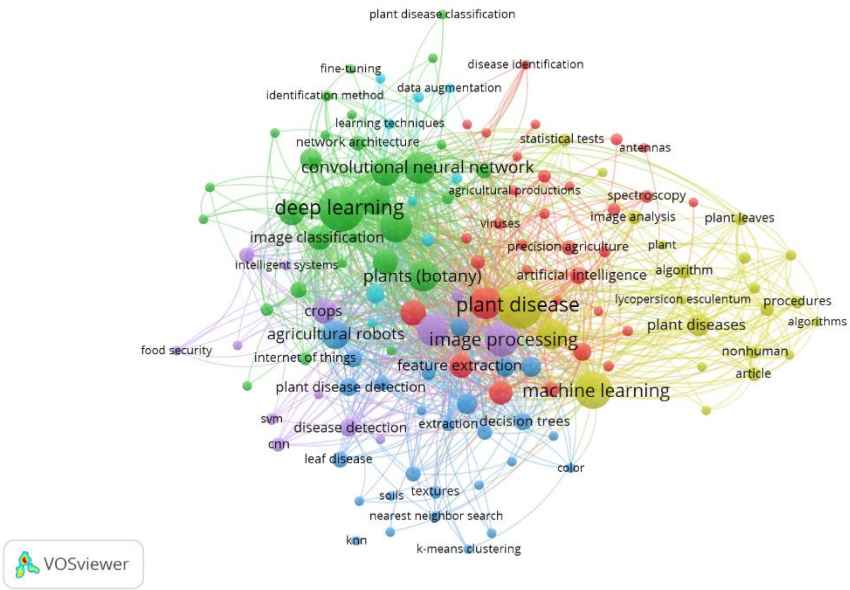


Figure 17: The Cluster of co-occurrences of the keywords in the Scopus publication.

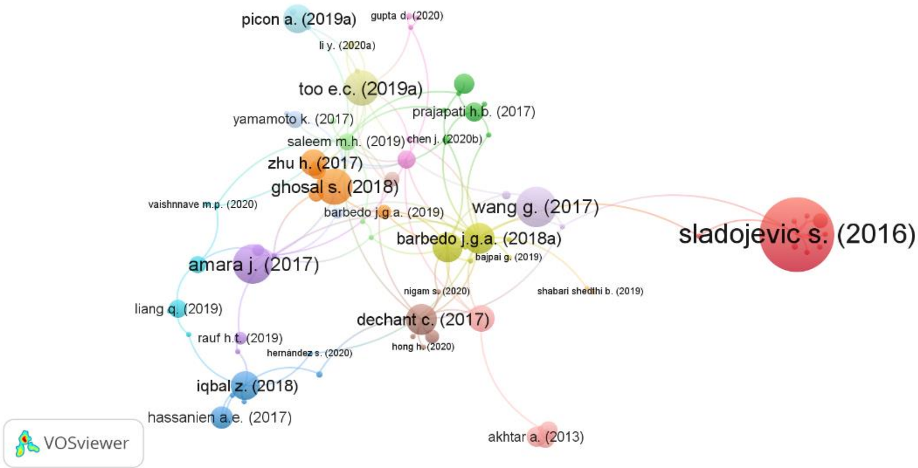


Figure 18: The Cluster of citations of documents in the Scopus publication.

The Cluster of citations of documents in the Scopus publication is shown in Figure 18. The interconnection between documents cited by authors can be seen here. The publication titled "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification" by author Sladojevic S is cited 380 times to date. Figure 19 shows the cluster of author and author keywords coappearing in the same paper. Deep learning, plant disease classification, and machine learning are seen to be the most significant keywords among the author keywords that are extensively used in the literature.

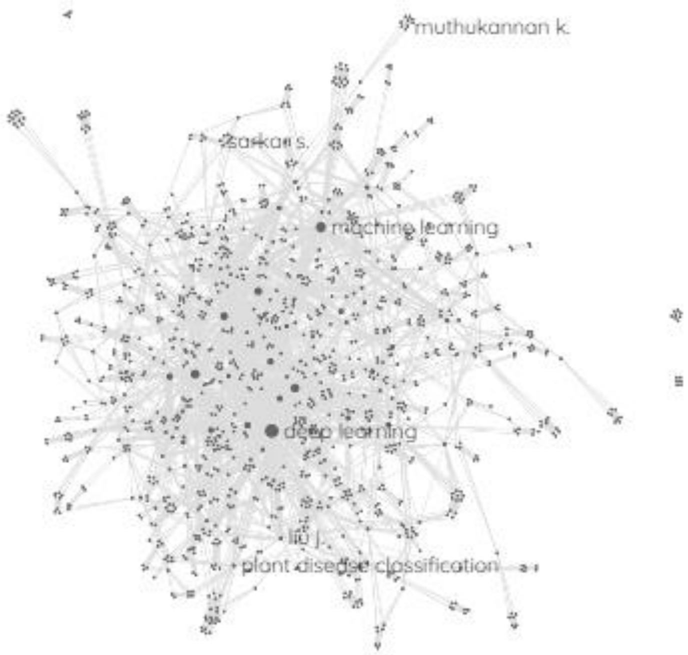


Figure 19: Author and author keyword coappearing in the same paper.

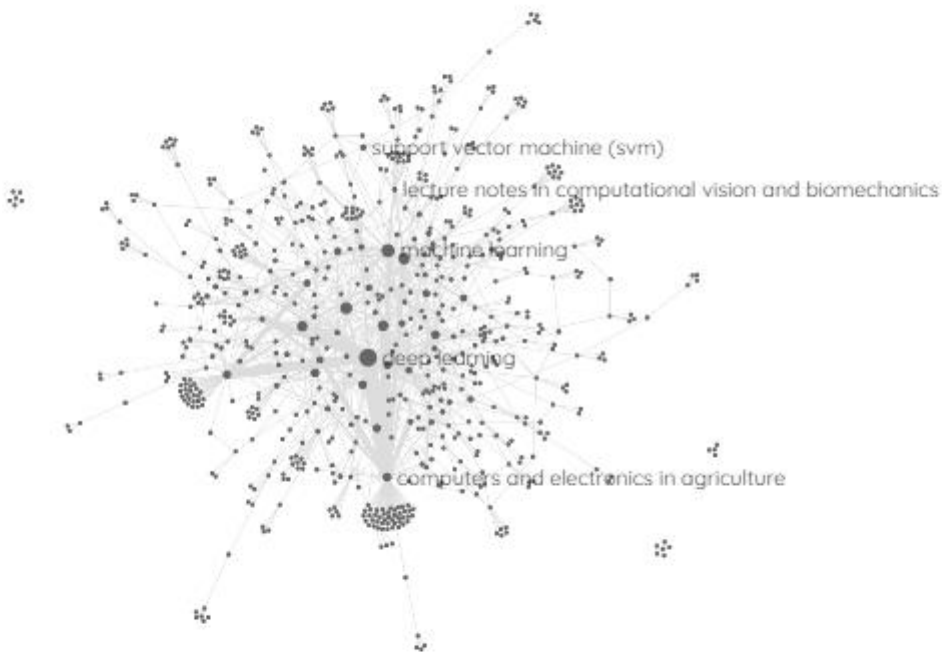


Figure 20: Source title and author keyword coappearing in the same paper.

Figure 20 shows the cluster of source tile and author keyword coappearing in the same paper. Computers and Electronics in Agriculture is seen to be amongst the top five significant source titles. These top five source titles can be seen in Figure 9 showing the CiteScore, SNIP, and SJR details.

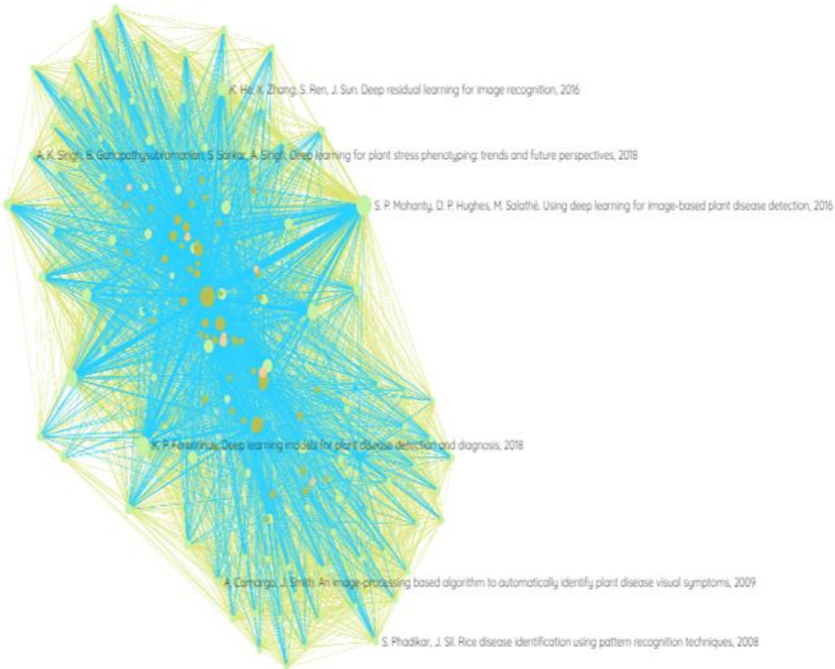


Figure 21: Reference-scape for Scopus database

Reference-scape shown in Figure 21 highlights some of the crucial references. This figure is useful for the researchers who want to contribute to the area of plant disease classification as they can be the most valuable references for formulating their research articles. The most used words that are frequently seen in this research publication are grouped in a cluster showcasing them in the publications for the plant disease classification. The visualization of words is through wordcloud as shown in Figure 22. The significant keywords that are used more in the research are shown with higher font size. The most used words are "plant", "images", "diseases", "neural network", "leaves", "CNN", "model" etc to name a few.

Figure 22: The words in the publication are visualized with wordcloud

The visualization tools discussed here shows the deep insights of information for the budding researcher to follow the different parametric incorporations in the field of plant disease classification for the data extracted from Scopus and WOS. The growth of research is seen in both Scopus and WOS. The link between the authors, the keywords they used in their research, and the source journal where they publish their work is showing a brief direction for choosing a more niche area and then follow and implement it through this visualization. The interlinkage between the keywords helps in understanding the broad scope in this area. The influential authors to whom the other researchers follow that can give them a better direction towards the research is revealed through the visualization.

3.3 Statistical analysis of publication citation

The yearly citation of the publications is shown in Table 5 with the publications in the area of plant disease classification. The total citation for the 258 publications is 2618 for Scopus and 867 publications with 8251 citations in WOS data. Table 6 shows the list of the topmost five papers with citations received by them in both the database. The first five (Sr no. one to five) are research papers from Scopus and the next five (Sr no. six to ten) are research papers from WOS that have received the maximum citations. “Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification” is the most cited paper.

Table 5: Analysis of citations for publications in plant disease classification.

Year	<2017	2017	2018	2019	2020	2021	Total
No of citations (Scopus)	64	83	249	720	1380	122	2618
No of citations (WOS)	174	351	861	2300	4403	162	8251

Table 6: A citation analysis of the top ten publications in plant disease classification in Scopus and WOS

Sr No	Publication title	Yearly citations received by the publication						
		<2017	2017	2018	2019	2020	2021	Total
1	"Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification"	1	23	51	130	160	15	380
2	"Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning"	0	2	12	32	64	5	115
3	"A deep learning-based approach for banana leaf diseases classification"	0	2	16	47	41	5	111
4	"A review of advanced machine learning methods for the detection of biotic stress in precision crop protection"	8	10	25	23	29	2	97
5	"An explainable deep machine vision framework for plant stress phenotyping"	0	0	6	33	50	4	93
6	"Deep learning models for plant disease detection and diagnosis"	0	0	19	67	109	3	198
7	"Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS"	0	13	25	67	80	1	186
8	"An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping"	3	18	18	39	47	6	131
9	"Machine learning for predicting soil classes in three semi-arid landscapes"	9	13	28	31	36	3	120
10	"Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques"	0	7	21	38	49	0	115

3.4 Statistical analysis of document source

The document source statistics of the top ten publication sources in plant disease classification is shown in Figure 23. The highest numbers of publications are from "Computers and Electronics in Agriculture" with 27% in WOS and 7% in Scopus.

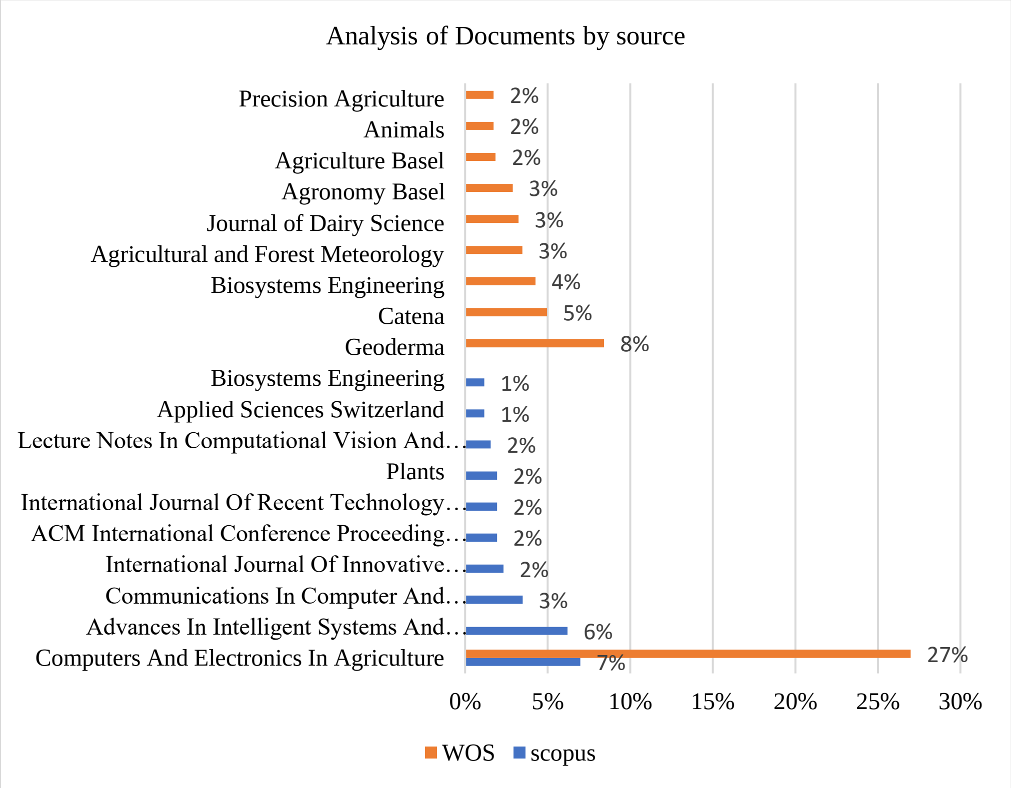


Figure 23: Source statistics for publications in plant disease classification in Scopus and WOS

4. Inferences Drawn from the Research Study

The research in the domain of plant disease classification is carried out globally. This research is important as the yield from the plants helps in fulfilling the increasing demand for food. The accurate and fast classification of disease helps in understanding and taking the preventive management of the plant to prevent the further growth of the disease. This results in good quality and quantity of the yield output. The study enlightens the contribution by resourceful authors in the plant disease classification through 115 research articles, 123 conference papers, 9 reviews in Scopus and 811 research articles, 45 reviews, and 18 conference papers in WOS. There is a steep growth in the research work since 2017. The majority of the literature is published in the English language. The maximum targeted keywords in the research community are deep learning, plant disease, machine learning, image processing, classification, and prediction. The statistical analysis shows that the top contribution in research is carried out by India, China, the US, and Australia. The maximum funding for this research domain is from the National Natural Science Foundation of China. The maximum citations are received in the year 2020 and very few in the year 2016. With the arising need to explore the domain of plant disease classification, researchers can explore this area which was unheeded in the past.

4.1 The Constraint in the Study

Different combinations of keywords were used to explore the Scopus and WOS database for bibliometric review. A few important journals and articles which were not available in these databases couldn't be fused in this study. Also, the study confines the research papers to the English language only.

CONCLUSIONS

The bibliometric study of plant disease classification with artificial intelligence is based on the research publication data fetched from the Scopus and WOS database. The bibliometric study highlights the maximum contribution of publication of 342 and 108 publications in WOS and Scopus in the year 2020. The articles were published in "Computer Science" followed by "Engineering", and "Agricultural and Biological Sciences" in Scopus. The contribution from “Computer Science”, “Water Resources” and “Food Science Technology” in WOS is maximum. The key research contributors in the articles and conference papers are from India, China, the United States, and Australia amongst the other countries. National Natural Science Foundation of China is having the maximum funded research work on plant disease classification in both Scopus and WOS. Minasny B is an influential author with 11 publications in WOS from 2010 till date in plant disease classification. The journal “Computers and Electronics in Agriculture” is the top journal for publishing high-quality research in Scopus and WOS, showing a good CiteScore, SJR and SNIP in the time frame chosen. The network visualization tools help in highlighting the significant keywords with influential authors publishing in the source title, year of publication in the research area. Plant disease classification is an important aspect as it succors to the disease management of the crop and leads to the good quality and quantity of the yield. This benefits society with the increasing demands of food globally. This bibliometric survey is very useful for new upcoming researchers to find the research gap, ongoing research, funding agencies, eminent researchers, and their affiliations in the field of plant disease classification.

REFERENCES

Ahumada, O., & Villalobos, J. R. (2009). Application of planning models in the agri-food supply chain: A review. *European Journal of Operational Research*, *196*(1), pp 1–20.

Amara, J., Bouaziz, B., & Algergawy, A. (2017). A deep learning-based approach for banana leaf diseases classification. *Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft Fur Informatik (GI), 266*, pp 79–88.

Arnal Barbedo, J. G. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, *180*(2016), pp 96–107.

Arsenovic, M., Karanovic, M., Sladojevic, S., Anderla, A., & Stefanovic, D. (2019). Solving current limitations of deep learning based approaches for plant disease detection. *Symmetry*, *11*(7). pp 1 -21.

Behmann, J., Mahlein, A. K., Rumpf, T., Römer, C., & Plümer, L. (2015). A review of advanced machine learning methods for the detection of biotic stress in precision crop protection. *Precision Agriculture*, *16*(3), pp 239–260.

Brungard, C. W., Boettinger, J. L., Duniway, M. C., Wills, S. A., & Edwards, T. C. (2015). Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*, *239*,pp 68–83.

Chaudhari, A., Joshi, R. R., Mulay, P., Kotecha, K., & Kulkarni, P. (2019). Bibliometric survey on incremental clustering algorithms. *Library Philosophy and Practice*, *2019*.pp 1-24.

Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., & Blaschke, T. (2018). The rise of deep learning in drug discovery. *Drug Discovery Today*, *23*(6),pp 1241–1250.

Chen, W., Pourghasemi, H. R., Kornejady, A., & Zhang, N. (2017). Landslide spatial modeling: Introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. *Geoderma*, *305*(June), pp 314–327.

Darwish, A., Ezzat, D., & Hassanien, A. E. (2020). An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm and Evolutionary Computation*, *52*(July 2019), pp 1-12.

Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and Electronics in Agriculture*, *145*(September 2017), pp 311–318.

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(18), pp 4613–4618.

Griffel, L. M., Delparte, D., & Edwards, J. (2018). Using Support Vector Machines classification to differentiate spectral signatures of potato plants infected with Potato Virus Y. *Computers and Electronics in Agriculture*, *153*(May),pp 318–324.

Kadam, S., Bandyopadhyay, P. K., & Patil, Y. (2016). Mapping the field through bibliometric analysis of passenger centric railway transportation. *International Journal of Automation and Logistics*, *2*(4), pp 349-368.

Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., & Menaka, R. (2020). Attention embedded residual CNN for disease detection in tomato leaves. *Applied Soft Computing Journal*, *86*, pp 1-24.

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease Detection. *Frontiers in Plant Science*, *7*(September), pp 1–10.

Nagasubramanian, K., Jones, S., Singh, A. K., Sarkar, S., Singh, A., & Ganapathysubramanian, B. (2019). Plant disease identification using explainable 3D deep learning on hyperspectral images. *Plant Methods*, *15*(1), pp 1–10.

Pham, B. T., Tien Bui, D., Prakash, I., & Dholakia, M. B. (2017). Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. *Catena*, *149*, pp 52–63.

Picon, A., Alvarez-Gila, A., Seitz, M., Ortiz-Barredo, A., Echazarra, J., & Johannes, A. (2019). Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild. *Computers and Electronics in Agriculture*, *161*(March 2018), pp 280–290.

Rangarajan, A. K., & Purushothaman, R. (2018). Tomato crop disease classification using pre-trained deep learning algorithm. *Procedia Computer Science*, *133*, pp 1040–1047.

Sharif, M., Khan, M. A., Iqbal, Z., Azam, M. F., Lali, M. I. U., & Javed, M. Y. (2018). Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection. *Computers and Electronics in Agriculture*, *150*(April),pp 220–234.

Singh, U. P., Chouhan, S. S., Member, S., & Jain, S. (2017). Multilayer Convolution Neural Network for the Classification of Mango Leaves Infected by Anthracnose Disease. *IEEE Access*, pp 1–9.

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Computational Intelligence and Neuroscience*, *2016*. pp 1-11

Too, E. C., Yujian, L., Njuki, S., & Yingchun, L. (2019). A comparative study of fine-tuning deep learning models for plant disease identification. *Computers and Electronics in Agriculture*, *161*(March 2018), pp 272–279.

Vaishnnave, M. P., Suganya Devi, K., & Ganeshkumar, P. (2020). Automatic method for classification of groundnut diseases using deep convolutional neural network. *Soft Computing*, *24*(21), pp 16347–16360.

Wang, G., Sun, Y., & Wang, J. (2017). Automatic Image-Based Plant Disease Severity Estimation Using Deep Learning. *Computational Intelligence Neuroscience*, *2017*.pp1- 8

Yang, X., & Guo, T. (2017). Machine learning in plant disease research. *European Journal of BioMedical Research*, *3*(1), 6. https://doi.org/10.18088/ejbmr.3.1.2017.pp6-9

Yann, L., Leon, B., Yoshua, B., & Patrick, and H. (1998). Gradient-Based Learning Applied to Document Recognition. *PROCEEDINGS OF THE IEEE*, *86*(11), pp2278-2324

Zhu, H., Chu, B., Zhang, C., Liu, F., Jiang, L., & He, Y. (2017). Hyperspectral Imaging for Presymptomatic Detection of Tobacco Disease with Successive Projections Algorithm and Machine-learning Classifiers. *Scientific Reports*, *7*(1), pp1–12.