Journal of Soft Computing Paradigm (ISSN: 2582-2640)
www.irojournals.com/jscp/

An Integrated Approach for Crop Production Analysis from Geographic Information System Data using SqueezeNet

K. Geetha

Dean, Academics and Research, JCT College of Engineering and Technology, Coimbatore, India

E-mail: geetha.arulmani@gmail.com

Abstract

Predictions and estimations are very important for agriculture applications. The estimation results on crop production may have a huge impact in the economy of a country by changing their export and import data. The estimation of crop production was started by collecting information manually from the fields and analyzing it using a computer. However, the accuracy was not up to the mark due to the error caused by manual collection of data. The Geographic Information System (GIS) applications are developed to store the information observed from the satellite images on change detection in town planning, disaster management, business development and vegetation management. The proposed work estimates the crop production of Indian states from a GIS dataset with a SqueezeNet algorithm. The performance of the SqueezeNet algorithm is compared with the traditional Inception and ResNet algorithms.

Keywords: GIS data retrieval, GIS information processing, crop data prediction, vegetation planning, neural network analysis on agriculture

1. Introduction

The geographic information systems are developed to analyze the surfaces of the earth from a satellite image. The system acts like a computer software that has the ability to generate different kind of information in spatial and numerical form. Therefore, the system improves the decision-making capabilities of other supporting tools and mathematical models. The in-built software of the GIS model allows the user to create a customized map for their specific application [1]. Figure 1 represents the list of applications employed with GIS model. The GIS modules are loaded with some predetermined functions to create different layers of data. Each

DOI: https://doi.org/10.36548/jscp.2021.4.006
Received: 15.11.2021, received in revised form: 19.12.2021, accepted: 02.01.2022, published: 11.01.2022

layer may represent the street data, water resource data, building data or an integrated data. In most cases, the GIS modules analyze the map information by identifying the latitude and longitude information of an area and the data information is stored as vector or raster data [2, 3].

The vector data model operates with the collected information of x, y and z coordinates. However, the vector data information is not considered as 3D data as the z coordinate has either the height or the depth data of the captured scene [4, 5]. The vector data information is represented as point feature when the data is stored with a single point attribute and the same will be stated as polyline feature when the coordinates are indicated with two-point features. The point features are beneficial for the applications that require to mention the location of an object or a place. The polyline features are helpful in the place of boundary estimation and route mapping [6, 7].

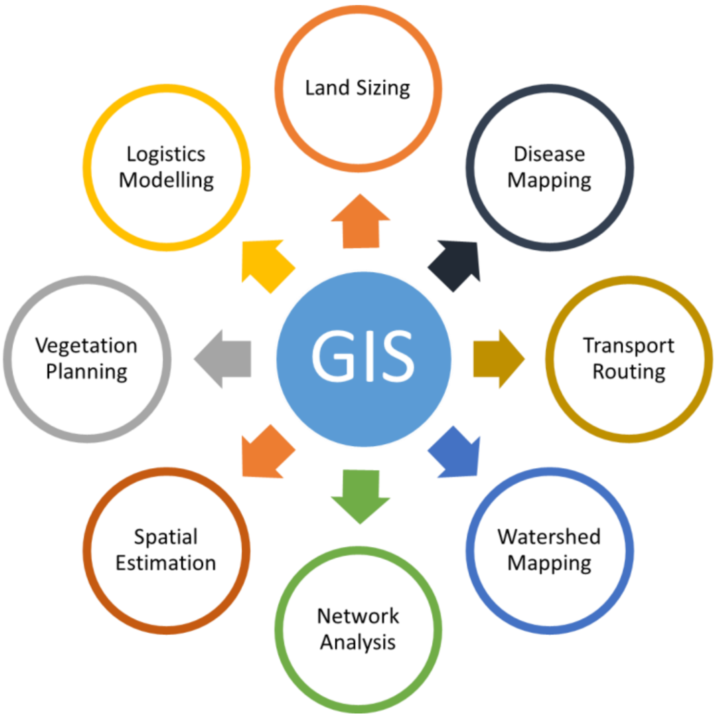

Figure 1. Applications of GIS

The raster data are the collective information of the data received from various satellites around the earth. The raster data information is always in the form of image model with a group of block values. The raster data can project the real-world information in an image format up to a certain extent [8, 9]. The values formed in the image blocks become meaningless when the image is over zoomed. In certain applications, the raster information is collected by aerial photography method where a digital camera is used for capturing images from any flying object like airplane and drone. The images taken by aerial photography are transmitted to the GIS station with a specific range of radio signals. The raster data which are collected by aerial photography is represented as remote sensing by the modern science [10].

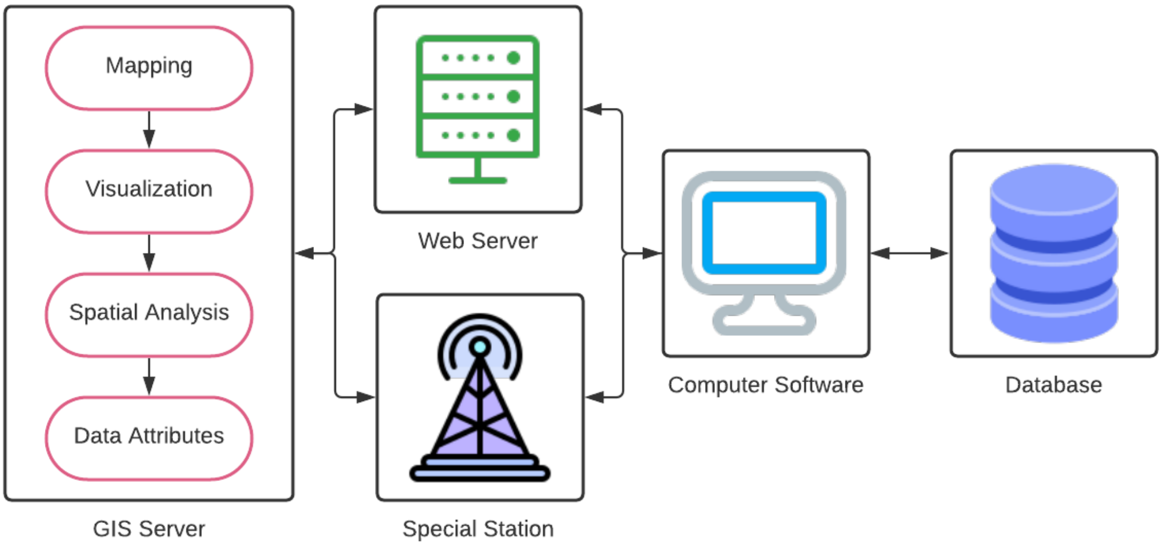

Figure 2. Architecture of GIS

The general architecture of a GIS module is shown in Figure 2. A computer software is employed in the architecture for a user interface. The area of mapping is given to the GIS server through a web server or radio signal station from the base station computer software. The GIS server decodes the received information and starts its visualization process by the help of powerful cameras and sensors. The collected data are converted into a spatial form for further feature analysis and that can be converted into a numerical matter as data attributes. A huge amount of data storage is required for GIS modules and that results in implementation of cloud storage in the latest GIS models [11-13].

2. Related Work

The GIS data was executed with a multi criteria decision making algorithm for estimating the choice of using the unused stagnant waters in the urban area for agriculture purposes. Land availability, soil type and water source availability distance were observed by the GIS module for the estimations. The Analytic Hierarchy Process (AHP) available in the

GIS software is utilized in the work indicating the weightages of the collected information [14]. A land suitability selection was made from the data collected through remote sensing and GIS setup. The analytical hierarchy process involved in the GIS software was merged with a pairwise comparison matrix for the selection process [15]. A land suitability prediction model was developed using GIS model for fruit crops in the central Portugal. The experimental work analyzed the soil type along with the climatic changes in the respective region. The GIS observed the annual rainfall, temperature, soil pH and soil organic matter for the analysis [16].

A fuzzy based AHP process was employed in a GIS framework for analyzing the suitability of land locations for wheat cultivation. The analytic work was merged in the process with geostatistics data which included terrain and soil information of the selected region. The experimental projection indicated that around 8.52% of area available in the selected region was not suitable for wheat farming [17]. The GIS and AHP setup was utilized to maintain soil nutrient availability in a selected field by giving suggestion on crop rotation at regular intervals. The experimental analysis was made by observing the eight major nutrient parameters available in normal soil. A semivariogram-based model was included in the work to analyze and give suggestion for planting new crop. The experimental analysis suggested to plant jute crop in the area that was planted with rice crop [18]. The literature work indicated that an inclusion of a data mining algorithm to the GIS framework results in a better prediction model. The following Table 1 indicates a list of data mining algorithm developed in the recent years for several applications.

Table 1. Review on data mining algorithms on various applications

Citation	Methodology	Dataset	Application	Attainments
Hammer et al. (2020) [19]	RF, Boosting, SVM	Real data set from sugar mills	Sugarcane Yield Prediction	RMSE RF = 19.702 Boosting = 20.025 SVM = 19.725

Anand (2021) [20]	CNN, LSTM, RNN, MLP, SVM	NSE	Stock Price Prediction	Better accuracy observed in CNN with ARIMA model
Werner et al. (2020) [21]	Decision Tree	NDVI and EVI	Cotton Field Mapping	Accuracy 96.08%
Sungheetha (2021) [22]	Data-Driven Epidemic Intelligence Strategies	SEIR-DPTT	COVID 19	Reduced Computational Speed
Rajagopal et al. (2021) [23]	Discrete DBN-VGGNet	Self-made dataset on soil, rainfall, and crop production	Prediction of Crop Production	Accuracy 97%
Karthigaikumar (2021) [24]	PCA and Decision Stump	SECOM	Industrial Quality Prediction	Better accuracy observed with rough set model
Pant et al. (2021) [25]	GBR, RF, DT	FAOSTAT	Crop Yield Prediction	Accuracy GBR = 89.65% RF = 6842 % DT = 96%
Shakya (2021) [26]	Decision Tree	Self-made hybrid dataset	Customer Segregation in Banking Model	Accuracy 77%

Agarwal et al. (2021) [27]	LSTM+RNN+ SVM	huge generic crop dataset from Kaggle	Crop Yield Prediction	Accuracy = 97%
Vivekanandam (2021) [28]	Wolf Optimization	Self-made Android dataset	Malware Detection	Accuracy = 92.3%

3. Proposed Work

The proposed work is employed with a hybrid dataset that consists of historical data of each Indian states from the year 1997 to 2014. Similarly, the nature data are downloaded from a GIS framework for the same years and the preprocess work is done on both datasets for filling-up the missing and mismatch data available in the dataset. In some places the data available in the dataset are converted into the same metric for a comfortable manipulation. Further the dataset information between the years 1997 to 2013 is considered for the training process and 2014 data is taken for the testing process. Figure 3 indicates the architectural block diagram of the proposed model.

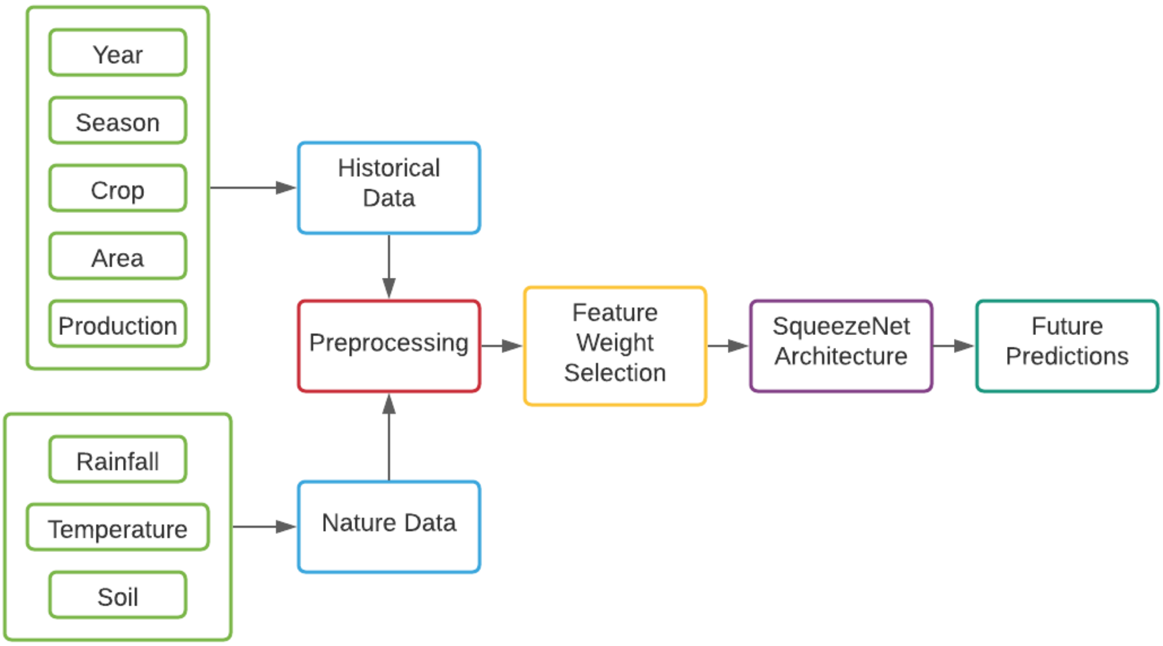

Figure 3. Block diagram of the proposed model

The different metrics of data available in the datasets are included in the work with different weightages. The year and seasonal information are considered in the work with a separate and equal weightage when compared to the other data. However, the weightages of features belong to the crop, and production is taken for the training process with an individual data attribute value. The information of soil is combined in the work for the analysis, and the values that belong to rainfall and temperature were averaged for all the training years. The features extracted with different weightages are forwarded to a SqueezeNet model for the training process. The SqueezeNet estimates the future growth, by understanding the crop growth rate observed among the previous rates, with respect to the climatic changes observed on the same year. The work is involved with a supervised training model, where the year wise information is forwarded to the neural architectures of the SqueezeNet one by one. This improves the training capacity of the network algorithm.

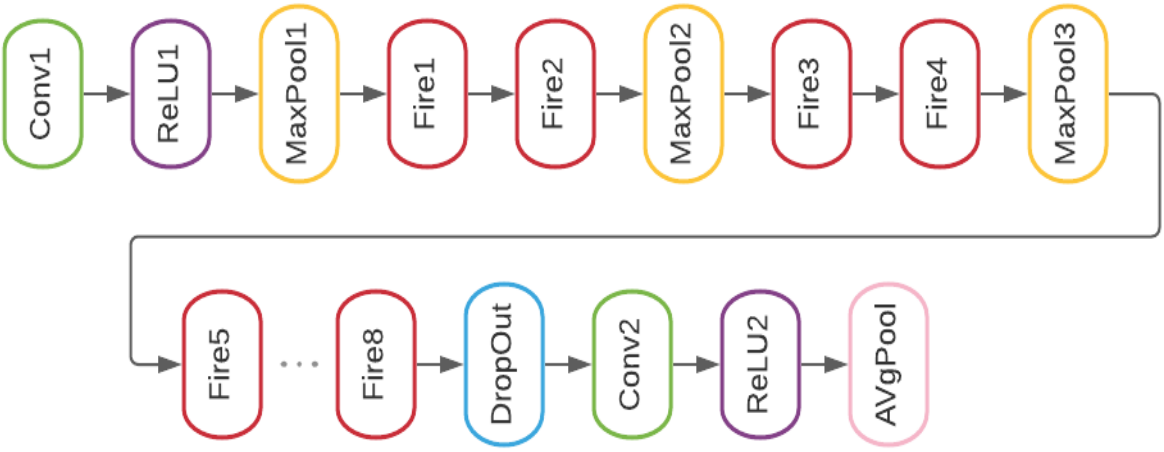

Figure 4. Architecture of the SqueezeNet algorithm

The SqueezeNet was developed to minimize the computational complexity of a regular neural network. Therefore, the model was developed with less number of training and tuning parameters and it fits the system even with a usual memory space. The parametric size of the SqueezeNet is 5MB whereas in the regular neural network like AlexNet, the parametric data size ranges around 240MB. The size reduction is achieved in the SqueezeNet by minimizing the filter sizes from 3x3 to 1x1 along with a reduction in input channels. Figure 4 explores the architectural view of the developed SqueezeNet model, and the fire model present in the architecture explores the squeezed blocks of convolution filters that varies from 1x1 to 3x3.

The convolution filters included in the fire blocks are tunable metrics same as that available in an Inception neural network. The accuracy of the SqueeezeNet is achieved by increasing the filter count in each fire block.

4. Experimental Analysis

The performance of the proposed work is compared with the Inception and ResNet algorithms in terms of crop, area and production accuracies. The downloaded dataset [29] contains the information of more than 21 Indian states. However, the work utilizes the information of only seven states. The SqueezeNet architecture is developed to overcome the computational complexity of the Inception and ResNet algorithms. Therefore the work compares the performances of the SqueezeNet with Inception and RestNet models.

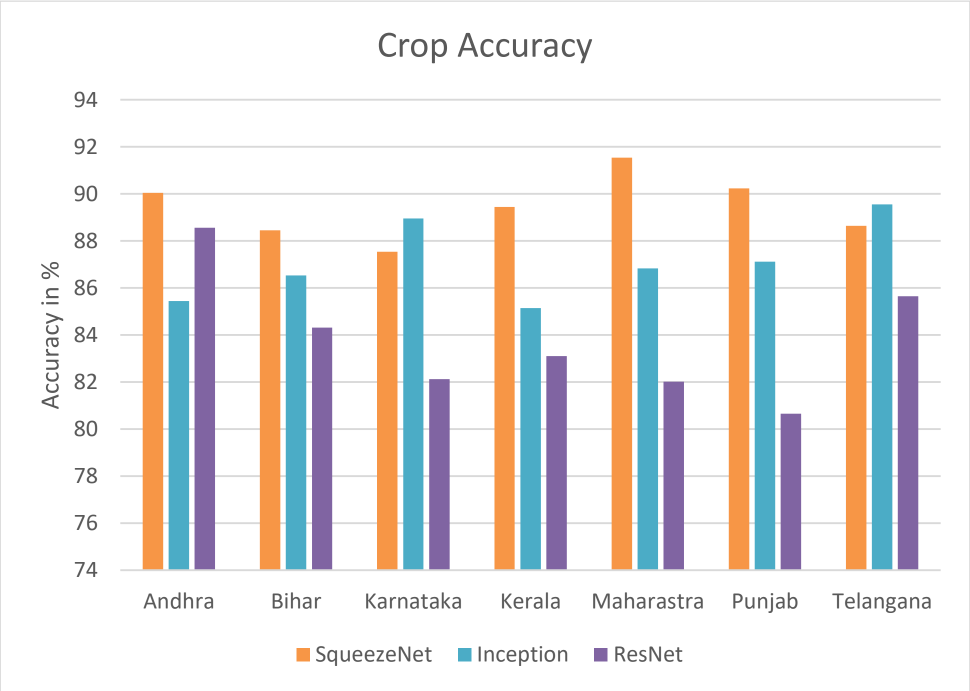

Figure 5. Crop accuracy comparison among the verified algorithms

Figure 5 explores the accuracy of the crop prediction data of the year 2014 available in the dataset. The experimental projection indicates that the accuracy of the SqueezeNet comes to the top for majority of the states except the Karnataka and Telangana states. The Inception network shows a betterment in those two state's data and the results of ResNet shows a betterment over the Inception algorithm in the state of Andhra.

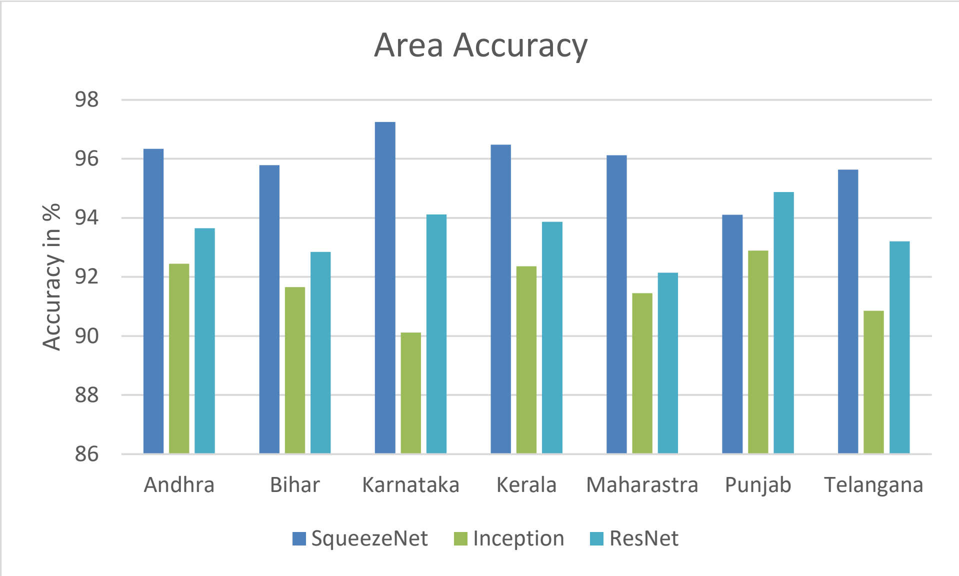

Figure 6. Area accuracy comparison among the verified algorithms

Figure 6 represents the area of cultivation accuracy of the verified algorithms, and it shows a collective betterment of ResNet in many states over the Inception and in few states than the proposed SqueezeNet. However, the SqueezeNet shows an overall betterment among the other two algorithms in average.

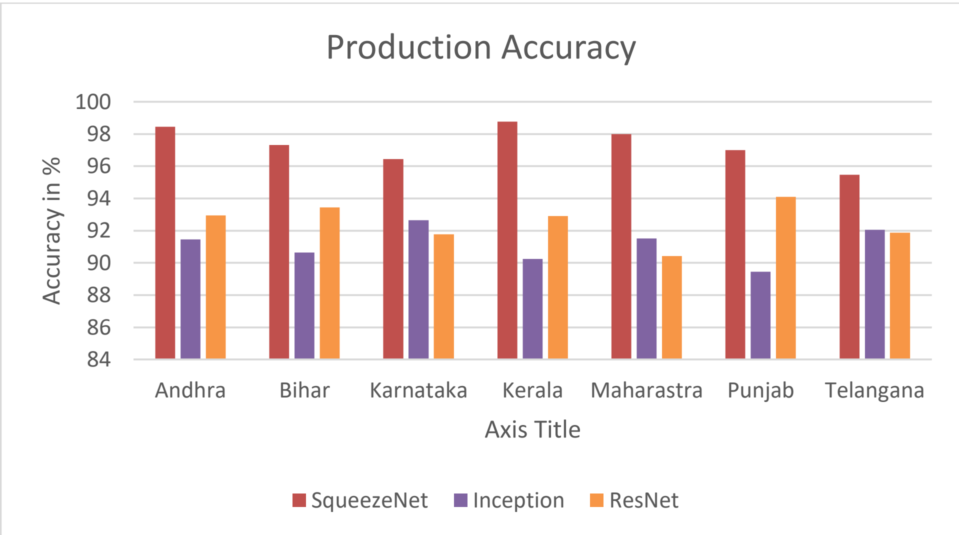

Figure 7. Production accuracy comparison among the verified algorithms

Figure 7 indicates the production accuracy of the SqueezeNet, Inception and ResNet models. It represents a good accuracy rate for SqueezeNet on all the state data. This is comparatively higher in rate when compared to the area and crop accuracies. The competition between the accuracies of ResNet and Inception algorithms continues to the production accuracy as well though the accuracy of ResNet explores a major betterment than the Inception algorithm.

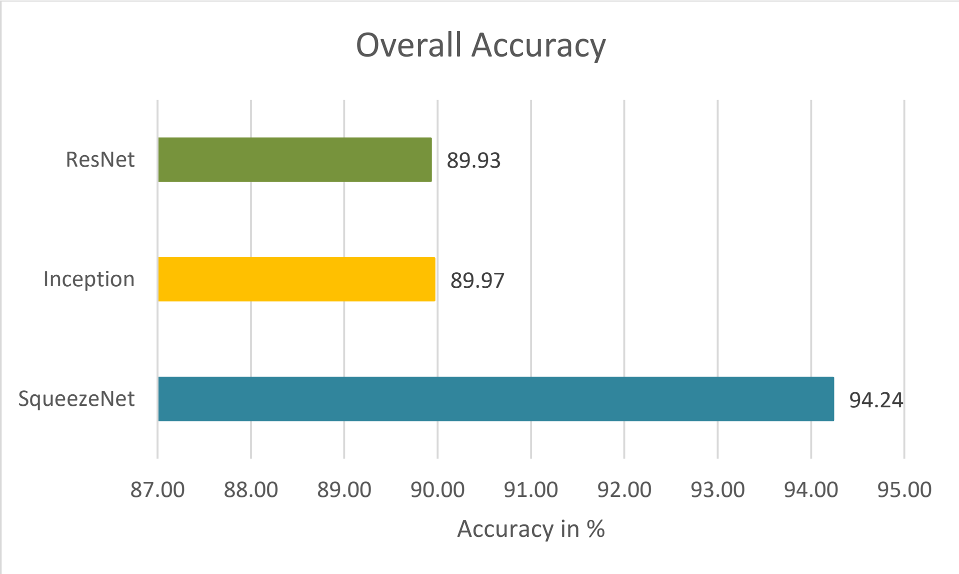

Figure 8. Overall accuracy comparison among the verified algorithms

The overall accuracy differences among the verified algorithms are shown in Figure 8 and it indicates a huge accuracy variation on SqueezeNet in average. The accuracies of Inception and ResNet are closer to each other with just 0.04% difference. The prediction work also found a reduced computational analysis time on SqueezeNet over the traditional Inception and ResNet.

5. Conclusion

Prediction algorithms can achieve a better accuracy rate when the collected data are integrated with some pre-available information. The proposed work is incorporated with a historical vegetation data and nature data for improving the prediction accuracy. The work utilizes a SqueezeNet algorithm because of its light weighted nature on the running process. A manual preprocessing work is also incorporated in the work with a change in feature selection rate. An experimental analysis was performed in the work to prove the efficacy of the proposed model on three categories and the results show a lesser accuracy rate on crop prediction model and higher accuracy rate on production estimation model. The reduction of accuracy in crop estimation is observed due to the wide range availability of crops. The same reduction of accuracy in crop estimation is observed on the traditional Inception and ResNet algorithms as well. The accuracy attainments of SqueezeNet algorithm can be improved to a certain extent when the feature attribute's weightages are fine tuned in the fire block.

References

[1] Adimalla, Narsimha, and Ajay Kumar Taloor. "Hydrogeochemical investigation of groundwater quality in the hard rock terrain of South India using Geographic Information System (GIS) and groundwater quality index (GWQI) techniques." Groundwater for Sustainable Development 10 (2020): 100288.

[2] Bansal, Malti. "Low-Noise Amplifier for Wireless Local Area Network Applications." In International Conference on Communication, Computing and Electronics Systems, pp. 667-680. Springer, Singapore, 2020.

[3] Subramanian, D. Venkata, R. Sugumar, N. Dhipikha, R. Vinothini, S. Kavitha, and A. Harsha Anchaliya. "Comprehensive and Comparative Study of Efficient Location Tracking Based on Apriori and Dijkstra Algorithms." In International conference on Computer Networks, Big data and IoT, pp. 48-59. Springer, Cham, 2019.

[4] Chen, Joy Iong Zong. "Modified Backscatter Communication Model for Wireless Communication Network Applications." IRO Journal on Sustainable Wireless Systems 3, no. 2: 107-117.

[5] Mainalli, Sanjay, and Kalpana Sharma. "An Efficient Technique for Lossless Address Data Compression Using Novel Adaptive SPIHT Algorithm in WSN." In International Conference on Mobile Computing and Sustainable Informatics, pp. 81-91. Springer, Cham, 2020.

[6] Suma, V. "Community Based Network Reconstruction for an Evolutionary Algorithm Framework." Journal of Artificial Intelligence 3, no. 01 (2021): 53-61.

[7] Ogato, Gemechu Shale, Amare Bantider, Ketema Abebe, and Davide Geneletti. "Geographic information system (GIS)-Based multicriteria analysis of flooding hazard and risk in Ambo Town and its watershed, West shoa zone, oromia regional State, Ethiopia." Journal of Hydrology: Regional Studies 27 (2020): 100659.

[8] Raj, Jennifer S., and Mr C. Vijesh Joe. "Wi-Fi Network Profiling and QoS Assessment for Real Time Video Streaming." IRO Journal on Sustainable Wireless Systems 3, no. 1 (2021): 21-30.

[9] Naik, Deepa, Pothumudi Sireesha, and Tanmay De. "Machine Learning-Based Green and Energy Efficient Traffic Grooming Architecture for Next Generation Cellular Networks." In Evolutionary Computing and Mobile Sustainable Networks, pp. 261-269. Springer, Singapore, 2021.

[10] Manjunath, H. R., and C. D. Guruprakash. "Energy Efficient Heterogeneous Wireless Sensor Networks-Recent Trends & Research Challenges." In International Conference on Computer Networks and Inventive Communication Technologies, pp. 146-151. Springer, Cham, 2019.

[11] Rahimunnisa, Dr K. "LoRa-IoT Focused System of Defense for Equipped Troops [LIFE]." Journal of Ubiquitous Computing and Communication Technologies 2, no. 3 (2020): 153-177.

[12] Andi, Hari Krishnan. "Analysis of Serverless Computing Techniques in Cloud Software Framework." Journal of IoT in Social, Mobile, Analytics, and Cloud 3, no. 3 (2021): 221-234.

[13] Hossain, M. S., C. K. Gadagamma, Y. Bhattacharya, M. Numada, N. Morimura, and K. Meguro. "Integration of smart watch and Geographic Information System (GIS) to identify post-earthquake critical rescue area part. I. Development of the system." Progress in Disaster Science 7 (2020): 100116.

[14] Zolfaghary, Parvin, Mahdi Zakerinia, and Hossein Kazemi. "A model for the use of urban treated wastewater in agriculture using multiple criteria decision making (MCDM) and geographic information system (GIS)." Agricultural Water Management 243 (2021): 106490.

[15] Iliquín Trigoso, Daniel, Rolando Salas López, Nilton B. Rojas Briceño, Jhonsy O. Silva López, Darwin Gómez Fernández, Manuel Oliva, Lenin Quiñones Huatangari, Renzo E. Terrones Murga, Elgar Barboza Castillo, and Miguel Ángel Barrena Gurbillón. "Land Suitability Analysis for Potato Crop in the Jucusbamba and Tincas Microwatersheds (Amazonas, NW Peru): AHP and RS–GIS Approach." Agronomy 10, no. 12 (2020): 1898.

[16] Quinta-Nova, L. C., and Dora Ferreira. "Land suitability analysis for emerging fruit crops in central Portugal using GIS." Agriculture and Forestry/Poljoprivreda i Sumarstvo 66, no. 1 (2020): 41-48.

[17] Tashayo, Behnam, Afshin Honarbakhsh, Aliasghar Azma, and Mohammad Akbari. "Combined fuzzy AHP–GIS for agricultural land suitability modeling for a watershed in southern Iran." Environmental Management 66, no. 3 (2020): 364-376.

[18] Singha, Chiranjit, Kishore Chandra Swain, and Sanjay Kumar Swain. "Best crop rotation selection with GIS-AHP technique using soil nutrient variability." Agriculture 10, no. 6 (2020): 213.

[19] Hammer, Ralph G., Paulo C. Sentelhas, and Jean CQ Mariano. "Sugarcane yield prediction through data mining and crop simulation models." Sugar Tech 22, no. 2 (2020): 216-225.

[20] Anand, C. "Comparison of Stock Price Prediction Models using Pre-trained Neural Networks." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 3, no. 02 (2021): 122-134.

[21] Werner, J. P. S., SR De M. Oliveira, and J. C. D. M. Esquerdo. "Mapping cotton fields using data mining and MODIS time-series." International Journal of Remote Sensing 41, no. 7 (2020): 2457-2476.

[22] Sungheetha, Akey. "COVID-19 Risk Minimization Decision Making Strategy Using Data-Driven Model." Journal of Information Technology 3, no. 01 (2021): 57-66.

[23] Rajagopal, Aghila, Sudan Jha, Manju Khari, Sultan Ahmad, Bader Alouffi, and Abdullah Alharbi. "A Novel Approach in Prediction of Crop Production Using Recurrent Cuckoo Search Optimization Neural Networks." Applied Sciences 11, no. 21 (2021): 9816.

[24] Karthigaikumar, P. "Industrial Quality Prediction System through Data Mining Algorithm." Journal of Electronics and Informatics 3, no. 2 (2021): 126-137.

[25] Pant, Janmejay, R. P. Pant, Manoj Kumar Singh, Devesh Pratap Singh, and Himanshu Pant. "Analysis of agricultural crop yield prediction using statistical techniques of machine learning." Materials Today: Proceedings (2021).

[26] Shakya, Subarna, and S. Smys. "Big Data Analytics for Improved Risk Management and Customer Segregation in Banking Applications." Journal of ISMAC 3, no. 03 (2021): 235-249.

[27] Agarwal, Sonal, and Sandhya Tarar. "a Hybrid Approach for Crop Yield Prediction Using Machine Learning and Deep Learning Algorithms." In Journal of Physics: Conference Series, vol. 1714, no. 1, p. 012012. IOP Publishing, 2021.

[28] Vivekanandam, B. "Design an Adaptive Hybrid Approach for Genetic Algorithm to Detect Effective Malware Detection in Android Division." Journal of Ubiquitous Computing and Communication Technologies 3, no. 2 (2021): 135-149.

[29] "Crop Production Prediction." n.d. Kaggle.com. https://www.kaggle.com/anandsubbu007/crop-production-prediction/data.

Author's biography

K. Geetha is presently working as a Dean-Academics and Research at JCT College of Engineering and Technology, Coimbatore, India. Her area of research includes image processing, computer vision, artificial intelligence, optimization techniques, power systems, process control and automations.