Original Article

Accuracy Analysis for Image Classification and Identification of Nutritional Values Using Convolutional Neural Networks in Comparison with Logistic Regression Model

A.Satya Jnana Prakash[1], P. Sriramya[2]

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105.

[2]Project Guide, Corresponding Author, Department of Computer Science and Engineering, Saveetha School of Engineering, Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105.

Abstract

Aim: The goal is to raise public awareness about nutritional issues by using food images to predict nutritional analysis using a novel image classification technique. Methods and Materials: The proposed research will be conducted at our university, and a total of two groups have been formed. There are two types of neural networks: a convolutional neural network and a logistic regression network. The framework uses 10 samples per group to evaluate accuracy. Gpower of 80% was used to calculate the sample size. **Results:** Convolutional Neural Network algorithm has predicted the nutritional analysis with the accuracy of (83.84%) which is more compared with the Logistic algorithm (72.3%) in identifying the fruit, Calorie count, amount of protein content, total fat, and subsequently carbohydrates measurement and so on. There is no statistically significant difference with ($P = 0.092$, $>.05$) among the classification algorithms. Conclusion: The analysis shows that the Convolutional Neural Network is significantly better for the whole Nutrition Analysis process compared to the Logistic regression.

Keywords: Novel Image Classification, Convolutional Neural Network, Logistic Regression, Nutritional Analysis, Food, Calorie

DOI: 10.47750/pnr.2022.13.S04.067

INTRODUCTION

Calorie, protein, fat and carbohydrate counts are important in health care and diet management because they can be accurately predicted and analysed (Khan, Deshpande, and Tripathy 2019). Routine health information from the health management information system (HMIS) was analysed in this study to see whether it was being used effectively. Among other things, nutrition analysis was used to look at the aggregate prevalence of variables associated with chronic undernutrition (Tesema et al. 2021). Nutritional data may be predicted early and accurately, according to the study's findings (Kogias et al. 2018; Miranda 2020).

IEEE, Science Direct, and Google Scholar have published a total of 48, 44, and 8 articles related to the concept.(Ahuja et al. 2020) propose a new essential dietary assessment for maintaining a healthy diet. This model first penetrates the region proposal network and detects food images in dietary assessment by using the novel Image Classification model. The model is trained under different metrics, and gives the accuracy of 84.2% in dietary plans efficiently. This model proposed by(Kogias et al. 2018) helps the doctors and the patients to get high reasonable accuracy on the small food portions.(Attokaren et al. 2017) explains the image of food monitoring in day-to-day activities and approaches using CNN and retrieves the image through pixel itself, it has accuracy of 86.97% for the food recognition.(Kogias et al. 2018) proposed the food accurate estimation on a daily basis, and this proposed system achieves 84.18% and 85.94% accuracy at the first and second range of categorizing.(Khan, Deshpande, and Tripathy 2019) proposes human body nutrition development. Poor food reduces the immunity to build nutritional model recommendations using ML algorithms with accuracy of 78.62%.(Lei et al. 2018) applied data mining approaches in the diet of nutrition analysis.(Ahuja et al. 2020) proposed recommendation of a nutritional model in which diet is analyzed depending on the feedback of the user.Our team has extensive knowledge and research experience that has translate into high quality publications(Bhansali et al. 2021; Jayanth et al. 2021; Sudhakar, Ravel, and Perumal 2021; Sathiyamoorthi et al. 2021; Deepanraj et al. 2021; Raju et al.

2021; Arun Prakash et al. 2020; Kamath et al. 2020; Shanmugam et al. 2021; Rajasekaran et al. 2020; Adhinarayanan et al. 2020; Rajesh et al. 2020; Aurtherson et al. 2021)
Following the literature review, it was discovered that logistic regression is less accurate and efficient when it comes to predicting dietary outcomes. The project's goal is to improve accuracy by using new picture classification methods and doing nutritional analysis.

Materials And Methods

Image Processing Lab of the Computer Science Engineering Department of Saveetha School of Engineering, SIMATS, does this study. For the purposes of this investigation, CNN and Logistic Regression are the two methods being used. In all, 75 samples from each of the two groups are subjected to the suggested methodology's scrutiny. GPower statistical software was used to calculate the sample size, which had a power of 80%. An alpha and beta value of 0.05 and 0.2, respectively, are used in the FRUITS dataset used by the programme, which has 500 photos (Medus et al. 2021).

Logistic Regression Algorithm
This regression model relates the comparison of each object with other classes in a statistical manner, And each image is being fetched with an accurate probability in between 0 and 1. In logistic regression the beta values are determined by training data. It simplifies making predictions for classification problems. This regression technique has easily increased the classes. It mainly depends on the sigmoid function. The main advantage is it is very effective to train and does not overlap the overfitting problem. At last, it classifies very fast and accurately the unknown records.
Step 1: First step is to feature scaling of the dataset
Step 2: x=standardScaler() #scale around the same values
Step 3: xtrain=x.fit_transform(x train)
Step 4: xtest=x.fit_transform(x test).
Step 5: CNN relates to the data given.
Step 6: Classifier=LogisticRegression().
Step 7: classifier.fit(x train,x test).
Step 8: ypred=classifier.predict(xtest).
Step 9: result=confusion_matrix(test,ypred).

Convolutional Neural Network Algorithm
Convolutional Neural Network is a widely used detection model for image classification, Where CNN processes the images as grayscale images and outputs the multi-class labels. The layers are connected one with the other in a convolution way, and CNN trains the weight of its kernels(that runs on system hardware) on itself. The main advantage of CNN is not needing to format the images according to the cable of working on any kind of image.
For the classifier to do its nutrition analysis in the dataset, it needs training data. Training data may be created in a variety of ways. To perform a nutritional analysis, a collection of images is gathered and trained. Training and testing the model are accomplished via the use of the three functions X train, X test,and y train,y test The following is the suggested system's testing setup: Desktop with 32-bit, 30 GB disc space, 6 GB RAM, and software-Windows 10, Python 3.6 and above, Jupyter Notebooks. Now Split the acquired data into two separate datasets, one for training and the other for testing. ' When it comes to nutrition testing, data analysis is crucial since it informs both model construction and predictive modelling.
Step 1: Split the images differently based on the identifications.
Step 2: 70% of the dataset is used for training and building the model.
Step 3: 30% of the data is used for testing and removing the outliers of the model.
Step 4: Fit the model within the data,where the model should not overlap.
Step 5: Improved CNN is applied to the data entered.
Step 6: The nutrition analysis is done by model and calculates the calories effectively.
Now train the model and then evaluate the algorithm with the trained datasets and testing datasets. And train the novel image classification model based on the input dataset. Datasets are in image format.

Statistical Analysis
IBM SPSS version 21 was used to do the analysis. Ingredients, calories, fat, and fibre are the independent variables. Mean accuracy is the dependent variable. Iterations with a maximum of 10 samples were performed for both the proposed and current algorithms, and the projected accuracy for each iteration was documented for analysis accuracy. T-test was carried out using the values received from the iterations.

Results

In today's world nutrition diet plays an important role for healthy practices. As already proposed, the results indicated reasonable accuracy of the system in identifying the fruit, calorie count, amount of protein content, total fat and subsequently carbohydrates measurements. Table 1 gives the comparison of prediction accuracy between CNN and Logistic Regression for different iterations by varying test set size. The accuracy appears to vary for each iteration. Table 2 shows the Group Statistics of LR (mean accuracy of 85.4) and CNN (mean accuracy of 94.6) along with sample size, standard deviation and standard error.

Table 3 gives the statistical data analysis of the CNN and LR algorithms for the data set fixing confidence intervals as 95% and level of significance is 0.001. Figure 1 shows the bar graph of LR and CNN Algorithms. The comparison graph shows the Logistic Regression algorithm has an accuracy of 72.30% and CNN has an accuracy of 83.84% with +/-1 SD. The CNN algorithm technique achieved better performance than Logistic Regression.

Discussion

An innovative picture classification methodology was used in the system, and the categorised images were then analysed to determine their nutritional worth. CNN has an accuracy of 83.84 percent, whereas Logistic Regression has an accuracy of 72.30 percent, according to the results of this research.

The similar findings related to this work proposed by(Qiu 2019) have acquired an accuracy of 89% by using KNN algorithms but improved Conventional Neural Networks have obtained an accuracy of 92.8%. This model helps the doctors and the patients to get high reasonable accuracy on the small food portions by CNN model. In this paper the model can give the accurate accuracy of 84.2% in dietary plans efficiently(Kogias et al. 2018).(Attokaren et al. 2017) explains image food monitoring using CNN and it retrieves the image through pixels. Also it has an accuracy of 86.97% for food recognition. This study is a comparison between KNN and Conventional Neural Network algorithms. In nutrition value prediction and analysis it is referred to which can act as a monitoring system with the nutrition value using image classification and analysis(Srigurulekha and Ramachandran 2020). In the study of(Termritthikun and Kanprachar 2017) the data has shown that the conventional neural network has better accuracy than the algorithms like KNN, in efficient detection of the nutrition values using image classification and analysis. There are no opposing findings related to this work.

Although the proposed algorithm yields better accuracy it has certain limitations. Convolutional Neural Network is slow sometimes due to max pooling technique. In future various applications can be made by working together with the combination of other algorithms, and the accuracy of the nutrition value can be more efficient by the image classification and analysis.

Conclusion

The results of the logistic regression classifier are compared to those of CNN in this study using the Nutrition analysis model. The logistic regression approach has a precision of 72.30 percent, whereas the CNN algorithm has a precision of 83.84 percent.

DECLARATIONS

Conflict of Interest

No conflict of Interest in this manuscript.

Authors Contribution

Author ASJP was involved in data collection, data analysis, manuscript writing. Author PS was involved in conceptualization, data validation, and critical review of manuscript.

Acknowledgements

The authors would like to express their gratitude towards Saveetha School of engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1. Biozone Pvt. Ltd.Chennai.
2. Saveetha University
3. Saveetha Institute of Medical and Technical Sciences
4. Saveetha School of Engineering

References

1. Adhinarayanan, Rajesh, Aravindh Ramakrishnan, Gopal Kaliyaperumal, Melvinvíctor De Poures, Rajesh Kumar Babu, and Damodharan Dillikannan. 2020. "Comparative Analysis on the Effect of 1-Decanol and Di-N-Butyl Ether as Additive with diesel/LDPE Blends in Compression Ignition Engine." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, June, 1–18.
2. Ahuja, Kashish, Mukul Goel, Sunil Sikka, and Priyanka Makkar. 2020. "WHAT-TO-TASTE: A FOOD RECOMMENDATION SYSTEM." *International Journal of Innovative Research in Computer Science & Technology*. https://doi.org/10.21276/ijircst.2020.8.3.6.
3. Arun Prakash, V. R., J. Francis Xavier, G. Ramesh, T. Maridurai, K. Siva Kumar, and R. Blessing Sam Raj. 2020. "Mechanical, Thermal and Fatigue Behaviour of Surface-Treated Novel Caryota Urens Fibre–reinforced Epoxy Composite." *Biomass Conversion and Biorefinery*, August. https://doi.org/10.1007/s13399-020-00938-0.
4. Attokaren, David J., Ian G. Fernandes, A. Sriram, Y. V. Srinivasa Murthy, and Shashidhar G. Koolagudi. 2017. "Food Classification from Images Using Convolutional Neural Networks." *TENCON 2017 - 2017 IEEE Region 10 Conference*. https://doi.org/10.1109/tencon.2017.8228338.
5. Aurtherson, P. Babu, Bhanu Teja Nalla, Karthikeyan Srinivasan, Kulmani Mehar, and Yuvarajan Devarajan. 2021. "Biofuel Production from Novel Prunus Domestica Kernel Oil: Process Optimization Technique." *Biomass Conversion and Biorefinery*, May. https://doi.org/10.1007/s13399-021-01551-5.
6. Bhansali, Karan J., Kamlesh R. Balinge, Subodh U. Raut, Shubham A. Deshmukh, M. Senthil Kumar, C. Ramesh Kumar, and Pundlik R. Bhagat. 2021. "Visible Light Assisted Sulfonic Acid-Functionalized Porphyrin Comprising Benzimidazolium Moiety for Photocatalytic Transesterification of Castor Oil." *Fuel* 304 (November): 121490.
7. Deepanraj, B., N. Senthilkumar, D. Mala, and A. Sathiamourthy. 2021. "Cashew Nut Shell Liquid as Alternate Fuel for CI Engine—optimization Approach for Performance Improvement." *Biomass Conversion and Biorefinery*, February. https://doi.org/10.1007/s13399-021-01312-4.
8. Jayanth, Bellappu Venkat, Melvin Victor Depoures, Gopal Kaliyaperumal, Damodharan Dillikannan, Dilipsingh Jawahar, Kumaran Palani, and Ganesha Prasad Meravanigee Shivappa. 2021. "A Comprehensive Study on the Effects of Multiple Injection Strategies and Exhaust Gas Recirculation on Diesel Engine Characteristics That Utilize Waste High Density Polyethylene Oil." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, June, 1–18.
9. Kamath, Manjunath, Subha Krishna Rao, Jaison, Sridhar, Kasthuri, Gopinath, Sivaperumal, and Shantanu Patil. 2020. "Melatonin Delivery from PCL Scaffold Enhances Glycosaminoglycans Deposition in Human Chondrocytes – Bioactive Scaffold Model for Cartilage Regeneration." *Process Biochemistry* 99 (December): 36–47.
10. Khan, Asmabee, Sachin Deshpande, and Amiya K. Tripathy. 2019. "Optimizing Nutrition Using Machine Learning Algorithms-a Comparative Analysis." *2019 International Conference on Nascent Technologies in Engineering (ICNTE)*. https://doi.org/10.1109/icnte44896.2019.8946091.
11. Kogias, K., I. Andreadis, K. Dalakleidi, and K. S. Nikita. 2018. "A Two-Level Food Classification System For People With Diabetes Mellitus Using Convolutional Neural Networks." *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2018 (July): 2603–6.
12. Lei, Zhenfeng, Shuangyuan Yang, Han Liu, Saba Aslam, Jinyu Liu, Halefom Tekle, Emmanuel Bugingo, and Defu Zhang. 2018. "Mining of Nutritional Ingredients in Food for Disease Analysis." *IEEE Access*. https://doi.org/10.1109/access.2018.2866389.
13. Qiu, Bojia. 2019. *Food Recognition and Nutrition Analysis Using Deep CNNs*.
14. Rajasekaran, S., D. Damodharan, K. Gopal, B. Rajesh Kumar, and Melvin Victor De Poures. 2020. "Collective Influence of 1-Decanol Addition, Injection Pressure and EGR on Diesel Engine Characteristics Fueled with diesel/LDPE Oil Blends." *Fuel* 277 (October): 118166.
15. Rajesh, A., K. Gopal, De Poures Melvin Victor, B. Rajesh Kumar, A. P. Sathiyagnanam, and D. Damodharan. 2020. "Effect of Anisole Addition to Waste Cooking Oil Methyl Ester on Combustion, Emission and Performance Characteristics of a DI Diesel Engine without Any Modifications." *Fuel* 278 (October): 118315.
16. Raju, P., K. Raja, K. Lingadurai, T. Maridurai, and S. C. Prasanna. 2021. "Glass/Caryota Urens Hybridized Fibre-Reinforced nanoclay/SiC Toughened Epoxy Hybrid Composite: Mechanical, Drop Load Impact, Hydrophobicity and Fatigue Behaviour." *Biomass Conversion and Biorefinery*, March. https://doi.org/10.1007/s13399-021-01427-8.
17. Sathiyamoorthi, Ramalingam, Gomathinayakam Sankaranarayanan, Dinesh Babu Munuswamy, and Yuvarajan Devarajan. 2021. "Experimental Study of Spray Analysis for Palmarosa Biodiesel-diesel Blends in a Constant Volume Chamber." *Environmental Progress & Sustainable Energy* 40 (6). https://doi.org/10.1002/ep.13696.
18. Shanmugam, Rajasekaran, Damodharan Dillikannan, Gopal Kaliyaperumal, Melvin Victor De Poures, and Rajesh Kumar Babu. 2021. "A Comprehensive Study on the Effects of 1-Decanol, Compression Ratio and Exhaust Gas Recirculation on Diesel Engine Characteristics Powered with Low Density Polyethylene Oil." *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects* 43 (23): 3064–81.
19. Srigurulekha, K., and V. Ramachandran. 2020. "Food Image Recognition Using CNN." *2020 International Conference on Computer Communication and Informatics (ICCCI)*. https://doi.org/10.1109/iccci48352.2020.9104078.
20. Sudhakar, M. P., Merlyn Ravel, and K. Perumal. 2021. "Pretreatment and Process Optimization of Bioethanol Production from Spent Biomass of Ganoderma Lucidum Using Saccharomyces Cerevisiae." *Fuel* 306 (December): 121680.
21. Termritthikun, Chakkrit, and Surachet Kanprachar. 2017. "Accuracy Improvement of Thai Food Image Recognition Using Deep Convolutional Neural Networks." *2017 International Electrical Engineering Congress (iEECON)*. https://doi.org/10.1109/ieecon.2017.8075874.

TABLES AND FIGURES

Table 1. Comparison of prediction accuracy between CNN and Logistic Regression for different iterations.

Execution	CNN	LR
1	92.8	76.4
2	89.00	75
3	88.34	74.34
4	85.33	73.46
5	83.00	73.11
6	82.00	73.08

7	80.98	71.79
8	79.95	70.57
9	79.00	69.22
10	78.00	66.00

Table 2. Group Statistics of LR (mean accuracy of 85.4) and CNN (mean accuracy of 94.6) with sample size, mean, standard deviation and standard error

	Algorithm	**N**	**Mean**	**Std.Deviation**	**Std.error mean**
Accuracy	LR	10	72.30	3.042	0.962
	CNN	10	83.84	4.880	1.543

Table 3. Statistical Data Analysis of the CNN and LR algorithms for the data set fixing confidence interval as 95% and level of significance less than 0.05

	Levene's Test for Equality of Variances		T-test for Equality of Means						
			t	df	Sig	Mean Difference	Std.Error Difference	95% Confidence interval of the Difference	
	F	Sig.						Lower	Upper
Equal variances assumed	3.175	.092	6.348	18	.001	11.543	1.818	7.723	15.363
Equal not variances assumed			6.348	18	.001	11.543	1.818	7.669	15.417

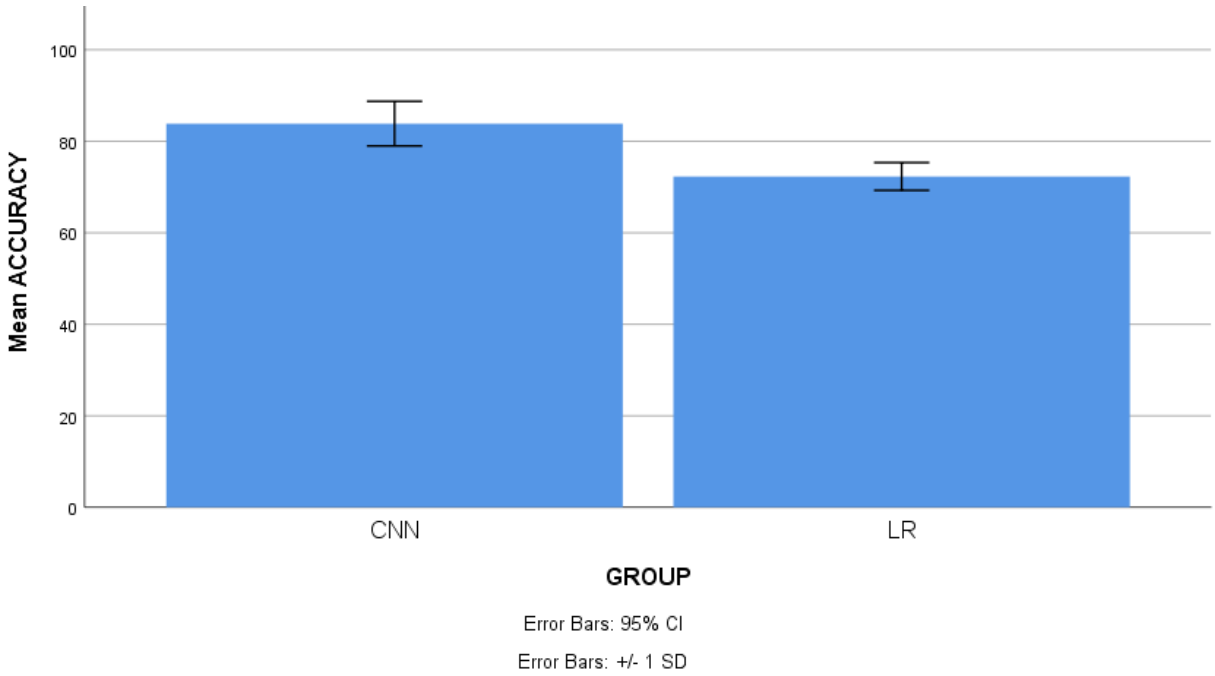

Fig. 1. Bar Graph of LR and CNN Algorithms: The comparison graph of the Logistic Regression algorithms which has an accuracy of 72.30% and CNN has an accuracy of 83.84%. X-axis:CNN vs LR. Y-axis: Mean Accuracy of detection +/-1 SD.