ENHANCING HEALTH DISORDER DISCRIMINATION SYSTEM USING CONVOLUTIONAL NEURAL NETWORKS OVER SUPPORT VECTOR MACHINE ALGORITHM

Ashutosh Tripathi[1], T. P. Anithaashri[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The Novel Health Disorder Discrimination system is developed for discrimination of diseases in human detected by blood cell using image processing of convolutional neural networks algorithm over support vector machine.

Materials & Methods: By using convolutional neural networks algorithm over support vector machine algorithm in image processing, diseases like malaria, pneumonia, breast cancer, skin cancer were identified on dataset having sample size of 80 each group and software tool python opencv and jupyter notebook is used for running application and accuracy scoring respectively. Accuracy values for identification of diseases are calculated to quantify performance of convolutional neural networks algorithm against support vector machine algorithm with t-test analysis.

Results and Discussion: The analysis on trained dataset and test dataset has been performed successfully using SPSS and acquired 95% accuracy for CNN algorithm compared to support vector machine method, which gave 90% accuracy with level of significance ($p<0.05$). The resultant data depicts reliability in independent sample tests.

Conclusion: On the whole process of prediction of accuracy by image processing, CNN model gives significantly better performance compared with support vector machine model.

Keywords: Enhancing Novel Health Disorder Discrimination, Convolutional Neural Network, Support Vector Machine, Image Processing, Disease Detection, Deep Learning.

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

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

1. Introduction

This project was created to provide safe, reliable, time saving efficient, comfortable and affordable application services for people to detect malaria, pneumonia, breast cancer, skin cancer using deep learning methods is seen as having accomplished this objective because of drawbacks that present in the existing system. The idea that would help people in a better way by author paper (Yadav and Jadhav 2019) to recognise disease if they have images. Furthermore customers need not pay for using this application and can take safety measures and diagnose quickly by identifying diseases through this image processing application. Deep learning based algorithms demonstrated remarkable outcomes in various fields such as computer-aided diagnosis, speech recognition etc (Shin et al. 2016). For the above stated idea behind disease detection, to construct a system using Convolutional Neural Network that detects diseases quickly (Li et al. 2014) and also predicts with mild error. The system is user-friendly with help of GUI, so that it can be used not only by medical professionals but also by population at large ((Tripathi and Kumar 2019) (Shin et al. 2016).

In existing methodology, the medical images are hard to collect, as collecting and labeling of medical data is confronted with both data privacy concerns and requirement for time-consuming expert explanations ((Li et al. 2014; Kermany et al. 2018). The most significant image processing method in disease detection was (Kermany et al. 2018) achieved 92% accuracy on a small pneumonia X-rays image dataset by transfer learning((Yadav and Jadhav 2019) used SVM application and CNN comparison to detect two types of specific pneumonia named bacteria and virus and achieved average accuracy of 74% for ORB & SVM and 88% with CNN. Novel health discrimination system was developed to overcome the classification of disease by images using CNN with higher accuracy, through users to upload their report image to detect their medical test results as either positive/normal or negative.

Our institution is keen on working on latest research trends and has extensive knowledge and research experience which resulted in quality publications (Rinesh et al. 2022; Sundararaman et al. 2022; Mohanavel et al. 2022; Ram et al. 2022; Dinesh Kumar et al. 2022; Vijayalakshmi et al. 2022; Sudhan et al. 2022; Kumar et al. 2022; Sathish et al. 2022; Mahesh et al. 2022; Yaashikaa et al. 2022). The research gap in existing systems paves way to an effective novel health disease detection system and provides a generic solution to problem of image classification (Kermany et al. 2018). The objective of proposed system (Beutel et al. 2000) is to provide an enhanced version of novel health disorder discrimination system to overcome difficulties of these research gaps using convolutional neural network algorithm. Aim is to implement a disease detection system by using convolutional neural network algorithms which can detect presence of diseases efficiently and provide a prediction of it with high accuracy in image processing. The overall process of proposed system is to improvise detection of human novel health diseases by data augmentation in image classification using convolutional neural networks over support vector machine

2. Materials and Methods

The Study setting of proposed work is done in Compiler Design Lab, Saveetha School of Engineering, Saveetha University. The number of groups identified are two. Group 1 is Convolutional neural network and group 2 is Support vector machine. Convolutional neural network method was iterated various number of times with a sample size of 80 for each group calculated from clinical website and SPSS analysis is carried out with level of significance $p<0.05$.

Convolutional Neural Network (CNN)

The architecture of a CNN is designed to take advantage of 2D structure of an input in image processing. This is achieved with local connections and tied weights followed by some form of pooling which results in translation invariant features. CNN's have fewer parameters than fully connected networks with same number of hidden units. CNN does image classification based on features such as pixel intensity and image edges. Following are steps of convolutional neural network algorithm :

Algorithm

Step 1: Start the process
Step 2: Load datasets path
Step 3: Prepare dataset for training
Step 4: shuffle the data
Step 5: assigning labels and features
Step 6: Normalizing X and converting labels to categorical data
Step 7: Split X and Y for use in CNN
Step 7: Define, compile and train CNN Model
Step 8: Accuracy and Score of model
Step 9: Stop the process

Support vector machine: **(SVM)**

SVM is most popular supervised learning algorithm in field of image processing, which is used for Classification as well as Regression problems. SVM chooses extreme points/vectors that help in creating hyperplanes. These extreme

cases are called support vectors, and hence algorithm is termed as Support Vector Machine. Following are steps of support vector machine algorithm:

Algorithm

Step 1: Start the process
Step 2: import dataset.
Step 3: Explore data to figure out what they look like.
Step 4: Pre-process data.
Step 5: Split data into attributes and labels.
Step 6: Divide data into training and testing sets.
Step 7:Train SVM algorithm.
Step 8: Check predictions.
Step 9: End the process

Anaconda navigator is used for execution of project code. It helps to manage and access notebook files and any kind of python files. By giving python environment a command prompt can provide easy access to code and execution. Main tools that need to be installed in python environment are keras and tensorflow. Minimum of 4GB RAM is required to compile and execute project code. Preferred operating systems are windows and ubuntu. Above mentioned method is for users using windows OS. Using anaconda navigator software and anaconda prompt that install necessary modules.

Statistical analysis

To check data accuracy & reliability an SPSS a statistical software is used with a default alpha value of 0.05 for sample size of 80 of each group. The independent variables for dataset were blur, varying lighting condition,shadowing effects,image size of images. Many potential variables are dependent in image classification like spectral signatures, vegetation indices, transformed images, textural or contextual information, multitemporal images, multisensor images, and ancillary data. Image is segmented and binarized to build a function that contains an interest area for detection. Bar graph and error graph were generated for comparison of differences between CNN and SVM algorithms.

3. Results

Enhancing Novel health disorder discrimination system is used to detct diseases such as skin cancer, malaria, breast cancer, pneumonia using image processing in efficient way. For this purpose a convolutional neural network with Support Vector Machine By applying Method 1 gives significantly better accuracy of 95% than Method 2. The results are collected by performing multiple iterations of experiment for identifying different scales of accuracy rate. Further, performed statistical calculations using SPSS tool and obtained accuracy from experimented data and independent samples t-test is performed.

Table 1 represents comparison of accuracy in disease detection systems by using CNN method and SVM method, by iterating in disease detection systems for various numbers of times.

Table 2 represents sample size (N=80), Mean, Standard deviation and Standard error mean are classified based on accuracy and loss of data. Accuracy 95% of CNN is significantly higher compared to Support vector machine

Table 3 represents significance of data and standard error difference, where significance of convolutional neural network (CNN) and SVM with confidence interval as 95% and level of significance of <0.05.

Figure 1, represents comparison of mean accuracy of convolutional neural networks (CNN) The comparison of accuracy gained. Accuracy of group 1 is 95% and group 2 is 90%. The CNN has significantly performed better when compared to Support vector machine Group 1 appears to produce most consistent results with its standard deviation ranging from lower 93's to higher 95's. Group 2 appears to produce most variable results with its standard deviation ranging from 85's to 90's. There is a significant difference between convolutional neural networks (CNN) and Support vector machines.

4. Discussion

In this study of disease detection systems, convolutional neural networks have higher accuracy of 95% in comparison to support vector machine algorithms having accuracy of 90%. convolutional neural networks have better significance ($p<0.05$) than support vector machines and while using independent sample t-tests.

The results of proposed algorithm are better in both experimental and statistical analysis, there are certain limitations in work. Main limitation in this experiment is that attributes in dataset contain very few to predict accuracy (%) of disease as images considered are clear pixels. Further research should consider other deep learning algorithms to ascertain more efficient ways to perform better for large datasets. It is recommended that further research should be carried out on other parameters that can improve accuracy of detection

5. Conclusion

It is inferred that the accuracy in detection of disease through extracted images of blood cells using the Convolutional Neural Network algorithm gives significantly better results than Support Vector Machine algorithm. CNN shows the accuracy of 95% using and SVM shows the accuracy of 90% in detecting the diseases.

Declarations

Conflict of Interests

No conflict of interest in this manuscript.

Author contribution

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

Acknowledgement

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 necessary infrastructure to carry out this work successfully.

Funding

We thank following organizations for providing financial support that enabled us to complete this study:

1. MANAC Infotech pvt.Ltd., Hyderabad
2. Saveetha University
3. Saveetha School Of Engineering
4. Saveetha Institute Of Medical Technical Sciences.

6. References

Beutel, Jacob, Harold L. Kundel, Yongmin Kim, Richard L. Van Metter, and Steven C. Horii. 2000. *Handbook of Medical Imaging*. SPIE Press.

Caicedo, Juan C., Angel Cruz, and Fabio A. Gonzalez. 2009. "Histopathology Image Classification Using Bag of Features and Kernel Functions." *Artificial Intelligence in Medicine*. https://doi.org/10.1007/978-3-642-02976-9_17.

Dinesh Kumar, M., V. Godvin Sharmila, Gopalakrishnan Kumar, Jeong-Hoon Park, Siham Yousuf Al-Qaradawi, and J. Rajesh Banu. 2022. "Surfactant Induced Microwave Disintegration for Enhanced Biohydrogen Production from Macroalgae Biomass: Thermodynamics and Energetics." *Bioresource Technology* 350 (April): 126904.

Kermany, Daniel S., Michael Goldbaum, Wenjia Cai, Carolina C. S. Valentim, Huiying Liang, Sally L. Baxter, Alex McKeown, et al. 2018. "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning." *Cell* 172 (5): 1122–31.e9.

Kim, Yoonsik, Insung Hwang, and Nam Ik Cho. 2017. "Convolutional Neural Networks and Training Strategies for Skin Detection." *2017 IEEE International Conference on Image Processing (ICIP)*. https://doi.org/10.1109/icip.2017.8297017.

Kumar, J. Aravind, J. Aravind Kumar, S. Sathish, T. Krithiga, T. R. Praveenkumar, S. Lokesh, D. Prabu, A. Annam Renita, P. Prakash, and M. Rajasimman. 2022. "A Comprehensive Review on Bio-Hydrogen Production from Brewery Industrial Wastewater and Its Treatment Methodologies." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123594.

Li, Qing, Weidong Cai, Xiaogang Wang, Yun Zhou, David Dagan Feng, and Mei Chen. 2014. "Medical Image Classification with Convolutional Neural Network." In *2014 13th International Conference on Control Automation Robotics & Vision (ICARCV)*. IEEE. https://doi.org/10.1109/icarcv.2014.7064414.

Mahesh, Narayanan, Srinivasan Balakumar, Uthaman Danya, Shanmugasundaram Shyamalagowri, Palanisamy Suresh Babu, Jeyaseelan Aravind, Murugesan Kamaraj, and Muthusamy Govarthanan. 2022. "A Review on Mitigation of Emerging Contaminants in an Aqueous Environment Using Microbial Bio-Machines as Sustainable Tools: Progress and Limitations." *Journal of Water Process Engineering*. https://doi.org/10.1016/j.jwpe.2022.102712.

Mohanavel, Vinayagam, K. Ravi Kumar, T. Sathish, Palanivel Velmurugan, Alagar Karthick, M. Ravichandran, Saleh Alfarraj, Hesham S. Almoallim, Shanmugam Sureshkumar, and J. Isaac JoshuaRamesh Lalvani. 2022. "Investigation on Inorganic Salts K2TiF6 and KBF4 to Develop Nanoparticles Based TiB2 Reinforcement Aluminium Composites." *Bioinorganic Chemistry and Applications* 2022 (January): 8559402.

Oppenheim, Dor, Guy Shani, Orly Erlich, and Leah Tsror. 2019. "Using Deep Learning for Image-Based Potato Tuber Disease Detection." *Phytopathology* 109 (6): 1083–87.

Ram, G. Dinesh, G. Dinesh Ram, S. Praveen Kumar, T. Yuvaraj, Thanikanti Sudhakar Babu, and Karthik Balasubramanian. 2022.

"Simulation and Investigation of MEMS Bilayer Solar Energy Harvester for Smart Wireless Sensor Applications." *Sustainable Energy Technologies and Assessments*. https://doi.org/10.1016/j.seta.2022.102102.

Rezaeilouyeh, Hadi, Ali Mollahosseini, and Mohammad H. Mahoor. 2016. "Microscopic Medical Image Classification Framework via Deep Learning and Shearlet Transform." *Journal of Medical Imaging (Bellingham, Wash.)* 3 (4): 044501.

Rinesh, S., K. Maheswari, B. Arthi, P. Sherubha, A. Vijay, S. Sridhar, T. Rajendran, and Yosef Asrat Waji. 2022. "Investigations on Brain Tumor Classification Using Hybrid Machine Learning Algorithms." *Journal of Healthcare Engineering* 2022 (February): 2761847.

Sathish, T., V. Mohanavel, M. Arunkumar, K. Rajan, Manzoore Elahi M. Soudagar, M. A. Mujtaba, Saleh H. Salmen, Sami Al Obaid, H. Fayaz, and S. Sivakumar. 2022. "Utilization of Azadirachta Indica Biodiesel, Ethanol and Diesel Blends for Diesel Engine Applications with Engine Emission Profile." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123798.

Shin, Hoo-Chang, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers. 2016. "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning." *IEEE Transactions on Medical Imaging* 35 (5): 1285–98.

Srivastava, Shrey, Amit Vishvas Divekar, Chandu Anilkumar, Ishika Naik, Ved Kulkarni, and V. Pattabiraman. 2021. "Comparative Analysis of Deep Learning Image Detection Algorithms." *Journal of Big Data* 8 (1). https://doi.org/10.1186/s40537-021-00434-w.

Sudhan, M. B., M. Sinthuja, S. Pravinth Raja, J. Amutharaj, G. Charlyn Pushpa Latha, S. Sheeba Rachel, T. Anitha, T. Rajendran, and Yosef Asrat Waji. 2022. "Segmentation and Classification of Glaucoma Using U-Net with Deep Learning Model." *Journal of Healthcare Engineering* 2022 (February): 1601354.

Sundararaman, Sathish, J. Aravind Kumar, Prabu Deivasigamani, and Yuvarajan Devarajan. 2022. "Emerging Pharma Residue Contaminants: Occurrence, Monitoring, Risk and Fate Assessment – A Challenge to Water Resource Management." *Science of The Total Environment*. https://doi.org/10.1016/j.scitotenv.2022.153897.

Tripathi, Shyava, and Rishi Kumar. 2019. "Image Classification Using Small Convolutional Neural Network." *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. https://doi.org/10.1109/confluence.2019.8776982.

Varshni, Dimpy, Kartik Thakral, Lucky Agarwal, Rahul Nijhawan, and Ankush Mittal. 2019. "Pneumonia Detection Using CNN Based Feature Extraction." *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*. https://doi.org/10.1109/icecct.2019.8869364.

Vijayalakshmi, V. J., Prakash Arumugam, A. Ananthi Christy, and R. Brindha. 2022. "Simultaneous Allocation of EV Charging Stations and Renewable Energy Sources: An Elite RERNN-m2MPA Approach." *International Journal of Energy Research*. https://doi.org/10.1002/er.7780.

Yaashikaa, P. R., P. Senthil Kumar, S. Jeevanantham, and R. Saravanan. 2022. "A Review on Bioremediation Approach for Heavy Metal Detoxification and Accumulation in Plants." *Environmental Pollution* 301 (May): 119035.

Yadav, Samir S., and Shivajirao M. Jadhav. 2019. "Deep Convolutional Neural Network Based Medical Image Classification for Disease Diagnosis." *Journal of Big Data*. https://doi.org/10.1186/s40537-019-0276-2.

Tables and Figures

Table 1. Accuracy table for Convolutional Neural Networks and Support vector machine, the accuracy of Method 1 is 95% and Method 2 is 90%.

No. of iterations	Convolutional neural network (CNN)	Support vector machine
1	93. 0	88. 0
2	93. 2	88. 2

3	93. 4	88. 4
4	93. 6	88. 6
5	93. 8	88. 8
6	94. 0	89. 0
7	94. 2	90. 0
8	94. 6	90. 2
9	94. 8	89. 5
10	95. 0	89. 0

Table 2. Statistical Analysis of Mean, Standard Deviation, and Standard Error of accuracy of convolutional neural network (CNN) and Support vector machine. There is a statistically significant difference in accuracy between the methods. The convolutional neural network (CNN) method has the highest accuracy (95%) and the Support vector machine has the least accuracy (90%).

Group	N	Mean	Std. Deviation	Std.Error Mean
Algorithms				
Convolutional Neural network (CNN)	80	95. 9600	0 .68508	0.21664
Support Vector Machine	80	90. 0600	0. 69570	0.22000

Table 3. Comparison of Significance Level with value $p<0.05$. Both CNN and Support vector machines have a confidence interval of 95% with the level of significance <0.05.

		F	sig.	t	df	sig. (2-tailed)	Mean difference	Std. Error Difference	95% Confidence interval of the difference Lower	95% Confidence interval of the difference Upper

Accuracy	Equal variance assumed	3.27	0.032	15.98	18	0.003	4. 90	0.3087	4. 29132	5. 58868
Accuracy	Equal variances not assumed			15.98	1799	0.003	4. 94	0. 3087	4. 29131	5. 58869

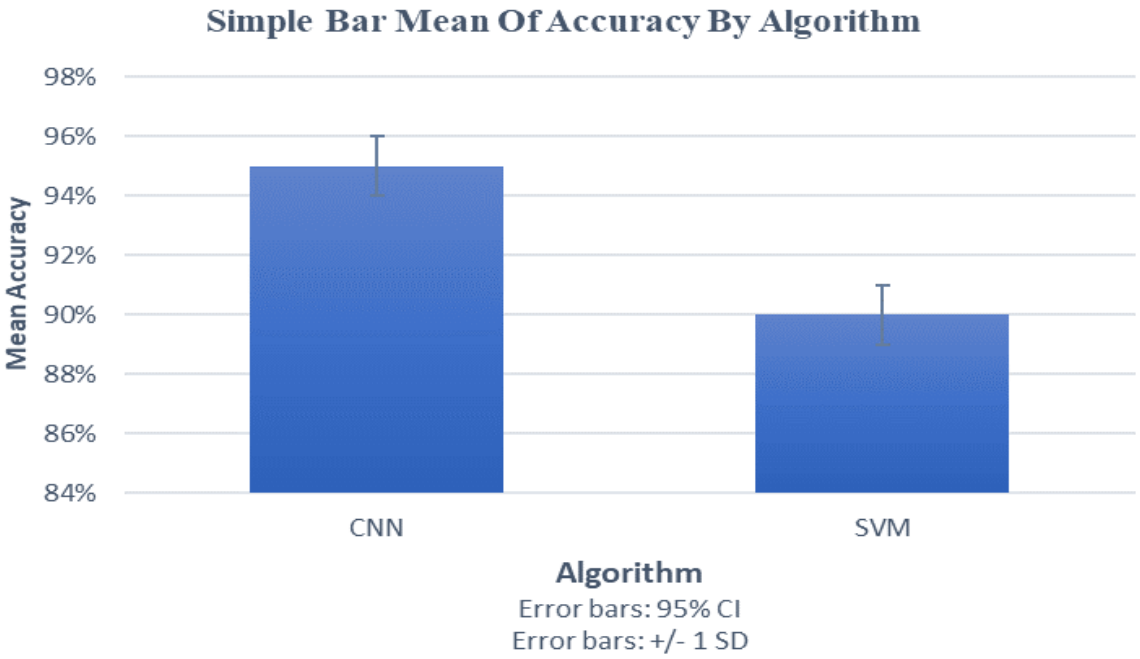

Fig. 1. Comparisons of Convolutional Neural Networks (CNN) that appear to produce the most consistent results with minimal standard deviations. There is a significant difference between Convolutional Neural Network (CNN) and Support vector machine. X Axis: Convolutional Neural Network (CNN) vs Support vector machine (SVM) Y Axis: Mean accuracy of detection ± 1 SD