Convolutional Neural Networks are a novel method for enhanced mammogram image segmentation and classification when compared to Support Vector Machines with improved accuracy.

Golla Vamsi Krishna[1], John Justin Thangaraj[2*]

[1]Researcher Scholar, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Science, Saveetha University, Chennai, Tamilnadu, Pincode:602105

[2*]Project Guide, Corresponding Author, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Science, Saveetha University, Chennai, Tamilnadu, Pincode: 602105

ABSTRACT

Aim: The objective of this study is to frame a mammogram breast image detection model using the Novel Convolutional Neural Network in comparison with the Support Vector Machine. In Image pre-processing, tumour segmentation, feature extraction, and detection are the functional phases of the proposed breast cancer detection. **Materials and Methods:** Novel Convolutional Neural Network with Support Vector Machine is a widely used neural network algorithm for mammogram image segmentation and classification as it remembers some important data which is received by the input and helps them to image formation of the next output accurately. It is mostly used in sequential data. The Novel Convolutional Neural Network algorithm with sample size 950, was computed with number of iterations N = 10 and Support Vector Machine algorithm with sample size 950, was computed with number of iterations N = 10, were evaluated many times to predict the accuracy percentage. **Results:** Convolutional Neural Network Algorithm has better accuracy (90.94%) when compared to Support Vector Machine (88.54%). The statistical significance difference (two-tailed) is 1.0471 (α = 0.001) value ($p<0.05$) with a confidence level of 95%. **Conclusion:** The study proves that Novel Convolutional Neural Networks exhibit better accuracy than Support Vector Machinein enhanced mammogram image segmentation and classification .

Keywords : Breast Cancer, Novel Convolutional Neural Network Algorithm, Image Segmentation, Mammography, Support Vector Machine.

INTRODUCTION

Breast cancer has the potential to spread to neighbouring tissues in the breast. It can also spread when cancer cells enter the bloodstream and migrate through the lymphatic system. (Campilho, Karray, and ter Haar Romeny 2018). The radiologist can analyse the pattern in a mammogram and detect the different types of breast cancer by analytical research on image texture (Song et al. 2020). Beginning in the 1980s, screening mammography led to significant advancements in breast cancer detection (Zhang et al. 2019). Mammography is a type of medical imaging that scans the breasts with low-dose x-ray equipment (Deepak., John Justin Thangaraj, and Rajesh Khanna 2020). A mammogram is a type of mammography test that aids in the early

detection and diagnosis of breast disorders in women. Application Breast cancer screening mammography is used to detect breast cancer in women who have no symptoms. Diagnostic mammograms are performed when a screening mammography yields abnormal results or when some indicators of breast cancer prompt the clinician to examine the tissue (Mata, Beena Ullala Mata, and Meenakshi Dr.M 2018). The applications of Mammograms are used as a screening tool to detect early breast cancer in women experiencing no symptoms. They can also be used to detect and diagnose breast disease in women experiencing symptoms such as a lump, pain, skin dimpling or nipple discharge.

This study is referenced in 75 IEEE xplore publications and 52 ScienceDirect articles. Image segmentation is crucial in lowering the cost of screening by improving the digital mammography outcome, especially in cases with tiny breast lesions (Al-Najdawi, Biltawi, and Tedmori 2015). The first major aspects of breast cancer are a mass with a defined location, texture, boundary, and form, all of which are straightforward to investigate using image processing techniques and the MATLAB application (Carmi et al. 2019). The concept of image processing entails transforming the result picture into a matrix as a binary image, followed by a number of operations until the right result is obtained (Mishra 2019), which aids in the diagnosis of breast lesions (Aghdam and Heravi 2017).(Venu and Appavu 2021; Gudipaneni et al. 2020; Sivasamy, Venugopal, and Espinoza-González 2020; Sathish et al. 2020; Reddy et al. 2020; Sathish and Karthick 2020; Benin et al. 2020; Nalini, Selvaraj, and Kumar 2020)

Previously our team has a rich experience in working on various research projects across multiple disciplines(Venu and Appavu 2021; Gudipaneni et al. 2020; Sivasamy, Venugopal, and Espinoza-González 2020; Sathish et al. 2020; Reddy et al. 2020; Sathish and Karthick 2020; Benin et al. 2020; Nalini, Selvaraj, and Kumar 2020).The disadavantages previously existing models show the less accurate results in detection of breast cancer image segmentation. Research gap analysis has reviewed and critically assessed enduring issues and new challenges emerging from recent research, and proposes strategies for translating solutions into practice. So the current paper aims is to find the image of which is breast effected using Convolutional Neural Network Algorithm and Standard Vector Machine Algorithm with the comparatively higher improved accurate results by modifying the models and choosing the largest dataset with the more number of parameters and more diverse result these help in determining patterns much better compared to previous models. The aim is to improve the accuracy rate using an enhanced Convolutional Neural Network in comparison with the Support Vector Machine detection.

MATERIALS AND METHODS

This research work was performed at the Cyber Forensic Laboratory, Saveetha School of Engineering, SIMATS (Saveetha Institute of Medical and Technical Sciences). The proposed work contains two groups such as group 1 is taken as Convolutional Neural Network

and group 2 as Support Vector Machine algorithm. The Convolutional Neural Network Algorithm and Support Vector Machine algorithm were evaluated a different number of times with a sample size of 950. The ethical approval does not arise, since the doctor has to detect the disease, the patient himself is giving his concern wholeheartedly. The calculation is performed utilising G-power 0.80 with an alpha value of 0.05 and beta value of 0.95 with a certainty stretch at 95%.

After dataset collection from the kaggle, the unused and not important content which is wrong in the datasets were removed by pre-processing and data cleaning steps. After cleaning and preprocessing the data, it opens the data sets and the accuracy of both Convolutional Neural Network algorithm and Support Vector Machine algorithm is calculated.

The clustering process of Convolutional Neural Network and Support Vector Machine algorithms are given below. MATLAB is used as an implementation for this work. The code was done in that notebook itself. Hardware configuration of the system consists of 8GB RAM and ROM of 1TB HDD SSD with a processor of 11^{th} gen intel(R) core(i5).

Convolutional Neural Networks Algorithm

In sample preparation group 1, The Convolutional Neural Network Algorithm is used to identify breast cancer image formation, The Novel Convolutional neural networks are feedforward neural networks that are made up of neurons. Local connection, weight sharing, and pooling operation are among its features. The network's complexity is effectively lowered, as well as the number of training parameters. Translation, scaling, and torsion are all features of convolutional neural networks.

Pseudocode

Input- Enhanced Mammogram Image Segmentation
Output- Accuracy of the model
Step 1. Begin
Step 2. Import the data from the required library
Step 3. Outliers decrease the effectiveness of the model
Step 4. Outliers are detected using quartile functions
Step 5. Remove the outliers from the dataset.
Step 6. From sklearn import train samples, then test the samples.
Step 7. Divide the dataset into 2 parts as training samples and testing samples
Step 8. Training constitutes 30% of data and is required to build the model.
Step 9. Remaining 70% is used to test the model
Step 10. Import Convolution Neural Network Classifier and fit the training data into it.
Step 11. It creates enhanced Mammogram Image Segmentation for each of the testing
Step 12. From which the accuracy is calculated.
Step 13. N_estimators parameter which denotes number of segmentations.
Step 14. Max_features number of features the model considers to split a node
Step 15. Min_sample_images number of images required to split an internal node.
Step 16. The **c**urve is considered from which accuracy score is predicted.

Step 17. Accuracy is achieved through means of all Recurrent Neural Networks.
Step 18. End.

Support Vector Machine

In sample preparation group 2, the Support Vector Machine algorithm is preferred when all the features are continuous in order to work with this import Support Vector Machine algorithm for solving classification problems. It is a tree structured classifier. Although the name of the algorithm suggests that a Support Vector Machine algorithm is a collection of trees, it is more efficient and produces more accurate results compared to Logistic Regressions. The accuracy obtained for the Support Vector Machine Algorithm is 89.67% via this method.

Pseudocode

Input- Enhanced Mammogram Image Segmentation
Output- Accuracy of the model
Step 1. Begin
Step 2. Give directory of the dataset in the csv extension file.
Step 3. This can also be the average value or mid value of the total values
Step 4. for each image processing let us consider there are a total of 950 images
Step 5. Use SVM for the importing and Sequential models.
Step 6. predicted the value for all the hundred images.
Step 7. fit the image segmentation based on the residual and predict the residuals
Step 8. The prediction has been updated for each of the values of the images.
Step 9. Now build a second image, compute the prediction using the second image.
Step 10.The best prediction is by minimising the sum of squared residuals.
Step 11.End.

Software specifications are concerned with the resources that must be installed in the target system in order to get an application to work. The minimal software specifications for this model to work are Windows operating system version 7/8/10, matlab programming

STATISTICAL ANALYSIS

In the current Study it is used a Statistical tool called IBM SPSS. Using this software's descriptive and group statistics for the accuracy values are calculated. Independent sample tests are taken and significance values are calculated. According to the analysis done between Convolutional Neural Network Algorithm and Support Vector Machine, Convolutional Neural Network appears to perform better than Convolutional Neural Network in all the platforms.Independent variables are distinct attributes that are helpful in prediction and dependent variables are improved accuracy values.The dependent variables are age, education, income, age when gave birth to first child, menopause, etc., and the independent variables are low parity, lactation, obesity, genetics, and the family history of breast cancer.

RESULTS

Table 1, shows descriptive statistics for accuracy for both the algorithms Novel Convolutional Neural Network Algorithm and Support Vector Machine. Table 1, also shows group statistics which gives the accuracy mean of 90.94 for Convolutional Neural Network Algorithm appears to be more when compared with Support Vector Machinewhich has only 88.54, Standard

deviation and mean errors are calculated (Standard error mean for novel Convolutional Neural Network Algorithm is 0.5100 and Support Vector Machine is 0.08207). Table 2, shows Independent test analysis, it gives significance 0.003. The bar chart Fig,1. shows the mean accuracy between Convolutional Neural Network and Support Vector Machine. From the results it is clearly evident that the Novel Convolutional Neural Network Algorithm is performing better when compared to the Support Vector Machine Algorithm.

DISCUSSION

In the current study it is observed that the supervised learning Novel Convolutional Neural Network Algorithm appears to have a higher success rate than the Standard Vector Machine (p=0.01, Independent sample Test). The improved accuracy of the novel Convolutional Neural Network algorithm than the Support Vector Machine.

The Comparison of novel Convolutional Neural Networks with the Support Vector Machinethat Convolutional Neural Networks are faring better than the Support Vector Machine (Mendel et al. 2019). Convolutional Neural Networks have accuracy of 90.94% and Support Vector Machines have 88.54%. Advantages compared Standard Vector Machine with the Convolutional Neural Networks and the results show that Convolutional Neural Networks are performing better than the Support Vector Machine. It has also shown similar results finally (Ma and Guo 2014). Also performed mammogram image segmentation on supervised machine learning algorithms and the results shows that Convolutional Neural Networks perform better and have the highest accuracy among all other algorithms (Wang and Ferguson 2017). Based on the research work done, 4 works showed similar findings and 1 work showed opposite findings (Yao et al. 2019). Moreover, based on the above discussions and findings (Aghdam and Heravi 2017). One can conclude that the Convolutional Neural Networks algorithm appears to have better accuracy and performance in all conditions than the Support Vector Machine algorithm (Kose and Alzubi 2020).

There are some limitations with the Convolutional Neural Networks algorithm that consists of clusters of large numbers which takes more time to get executed compared to other machine learning algorithms for breast cancer image processing (Rout, n.d.). In the future scope the model will be improved with better features and least running time possible and getting more precise results. This might have a better future as the number of victims has been increasing every day (Tsihrintzis, Sotiropoulos, and Jain 2018).

CONCLUSION

In this current paper it is predicted breast cancer with two different algorithms, Convolutional Neural Networks Algorithm and Support Vector Machine algorithm. Convolutional Neural Networks algorithm shows higher accuracy rate and performed better at a more significant rate than that of the Support Vector Machine.

DECLARATIONS

Conflicts of Interests

No conflicts of interests in the manuscript.

Authors Contribution

Author GVK was involved in data collection, data analysis, and manuscript writing. Author CMV 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 the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organisations for providing financial support that enabled us to complete the study.

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

REFERENCES

1. Aghdam, Hamed Habibi, and Elnaz Jahani Heravi. 2017. *Guide to Convolutional Neural Networks: A Practical Application to Traffic-Sign Detection and Classification*. Springer.
2. Al-Najdawi, Nijad, Mariam Biltawi, and Sara Tedmori. 2015. "Mammogram Image Visual Enhancement, Mass Segmentation and Classification." *Applied Soft Computing*. https://doi.org/10.1016/j.asoc.2015.06.029.
3. Benin, S. R., S. Kannan, Renjin J. Bright, and A. Jacob Moses. 2020. "A Review on Mechanical Characterization of Polymer Matrix Composites & Its Effects Reinforced with Various Natural Fibres." *Materials Today: Proceedings* 33 (January): 798–805.
4. Campilho, Aurélio, Fakhri Karray, and Bart ter Haar Romeny. 2018. *Image Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings*. Springer.
5. Carmi, Nehory, Azaria Cohen, Mireille Avigal, and Anat Lerner. 2019. "A Storyteller's Tale: Literature Audiobooks Genre Classification Using CNN and RNN Architectures." *Interspeech 2019*. https://doi.org/10.21437/interspeech.2019-1154.
6. Deepak., S. John Justin Thangaraj, and M. Rajesh Khanna. 2020. "An Improved Early Detection Method of Autism Spectrum Anarchy Using Euclidean Method." In *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. IEEE. https://doi.org/10.1109/i-smac49090.2020.9243361.
7. Gudipaneni, Ravi Kumar, Mohammad Khursheed Alam, Santosh R. Patil, and Mohmed Isaqali Karobari. 2020. "Measurement of the Maximum Occlusal Bite Force and Its Relation to the Caries Spectrum of First Permanent Molars in Early Permanent Dentition." *The Journal of Clinical Pediatric Dentistry* 44 (6): 423–28.
8. Kose, Utku, and Jafar Alzubi. 2020. *Deep Learning for Cancer Diagnosis*. Springer Nature.
9. Mata, B. N. Beena Ullala, B. Beena Ullala Mata, and Meenakshi Dr.M. 2018. "Mammogram Image Segmentation by Watershed Algorithm and Classification through K-NN

Classifier." *Bonfring International Journal of Industrial Engineering and Management Science*. https://doi.org/10.9756/bijiems.8352.

10. Ma, Yunqian, and Guodong Guo. 2014. *Support Vector Machines Applications*. Springer Science & Business Media.
11. Mendel, Kayla, Hui Li, Deepa Sheth, and Maryellen Giger. 2019. "Transfer Learning From Convolutional Neural Networks for Computer-Aided Diagnosis: A Comparison of Digital Breast Tomosynthesis and Full-Field Digital Mammography." *Academic Radiology* 26 (6): 735–43.
12. Mishra, Pradeepta. 2019. "CNN and RNN Using PyTorch." *PyTorch Recipes*. https://doi.org/10.1007/978-1-4842-4258-2_3.
13. Nalini, Devarajan, Jayaraman Selvaraj, and Ganesan Senthil Kumar. 2020. "Herbal Nutraceuticals: Safe and Potent Therapeutics to Battle Tumor Hypoxia." *Journal of Cancer Research and Clinical Oncology* 146 (1): 1–18.
14. Reddy, Poornima, Jogikalmat Krithikadatta, Valarmathi Srinivasan, Sandhya Raghu, and Natanasabapathy Velumurugan. 2020. "Dental Caries Profile and Associated Risk Factors Among Adolescent School Children in an Urban South-Indian City." *Oral Health & Preventive Dentistry* 18 (1): 379–86.
15. Rout, Jitendra Kumar. n.d. *Machine Learning for Intelligent Decision Science*. Springer Nature.
16. Sathish, T., and S. Karthick. 2020. "Gravity Die Casting Based Analysis of Aluminum Alloy with AC4B Nano-Composite." *Materials Today: Proceedings* 33 (January): 2555–58.
17. Sathish, T., D. Bala Subramanian, R. Saravanan, and V. Dhinakaran. 2020. "Experimental Investigation of Temperature Variation on Flat Plate Collector by Using Silicon Carbide as a Nanofluid." In *PROCEEDINGS OF INTERNATIONAL CONFERENCE ON RECENT TRENDS IN MECHANICAL AND MATERIALS ENGINEERING: ICRTMME 2019*. AIP Publishing. https://doi.org/10.1063/5.0024965.
18. Sivasamy, Ramesh, Potu Venugopal, and Rodrigo Espinoza-González. 2020. "Structure, Electronic Structure, Optical and Magnetic Studies of Double Perovskite Gd2MnFeO6 Nanoparticles: First Principle and Experimental Studies." *Materials Today Communications* 25 (December): 101603.
19. Song, Dong, Xiaxia Man, Meng Jin, Qian Li, Han Wang, and Ye Du. 2020. "A Decision-Making Supporting Prediction Method for Breast Cancer Neoadjuvant Chemotherapy." *Frontiers in Oncology* 10: 592556.
20. Tsihrintzis, George A., Dionisios N. Sotiropoulos, and Lakhmi C. Jain. 2018. *Machine Learning Paradigms: Advances in Data Analytics*. Springer.
21. Venu, Harish, and Prabhu Appavu. 2021. "Experimental Studies on the Influence of Zirconium Nanoparticle on Biodiesel–diesel Fuel Blend in CI Engine." *International Journal of Ambient Energy* 42 (14): 1588–94.
22. Wang, Jun, and Mark K. Ferguson. 2017. *Atlas of Minimally Invasive Surgery for Lung and Esophageal Cancer*. Springer.
23. Yao, Hongdou, Xuejie Zhang, Xiaobing Zhou, and Shengyan Liu. 2019. "Parallel Structure Deep Neural Network Using CNN and RNN with an

Attention Mechanism for Breast Cancer Histology Image Classification." *Cancers*. https://doi.org/10.3390/cancers11121901.

24. Zhang, Xiaohui, Yaoyun Zhang, Qin Zhang, Yuankai Ren, Tinglin Qiu, Jianhui Ma, and Qiang Sun. 2019. "Extracting Comprehensive Clinical Information for Breast Cancer Using Deep Learning Methods." *International Journal of Medical Informatics* 132 (December): 103985.

Tables and Graphs

Table 1: Accuracy of CNN and SVM. The CNN algorithm is 2.4% more accurate than the SVM algorithm.

ITERATION NO.	CNN (%)	SVM (%)
1	93.14	91.52
2	92.59	90.20
3	92.19	90.28
4	91.72	89.78
5	91.18	89.19
6	90.88	88.76
7	90.20	88.05
8	89.67	87.68
9	89.09	87.17
10	88.80	86.59

Table 2.Group Statistics results (Mean of Convolutional Neural Networks is 90.94 is more compared to Convolutional Neural Network is 88.54 and Standard error mean for Convolutional Neural Networks is 0.51003 and Support Vector Machine is 0.08207).

Group Statistics					
	GROUPS	**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
Accur	**Convolutional**	10	90.9470	2.6128	0.5100

acy	**Neural Network**				
	Support Vector Machine	10	88.5400	1.2595	0.0820

Table 3. Independent Sample Test for significance and standard error determination. P value is 0.003(less than 0.005) considered to be statistically significant and 95% confidence interval was considered.

Independent Samples Test										
Levene's Test for Equality of Variances				**t-test for Equality of Means**						
		F	**Sig.**	**t**	**df**	**Sig. (2-tailed)**	**Mean Difference**	**Std. Error Difference**	**95% Confidence Interval of the Difference**	
									Lower	**Upper**
Accuracy	**Equal variances assumed**	0.045	0.001	4.272	18	<0.001	2.2070	0.5165	1.1216	3.2923
	Equal variances not assumed			4.272	18	<0.001	20.2070	0.5165	1.0471	3.3669

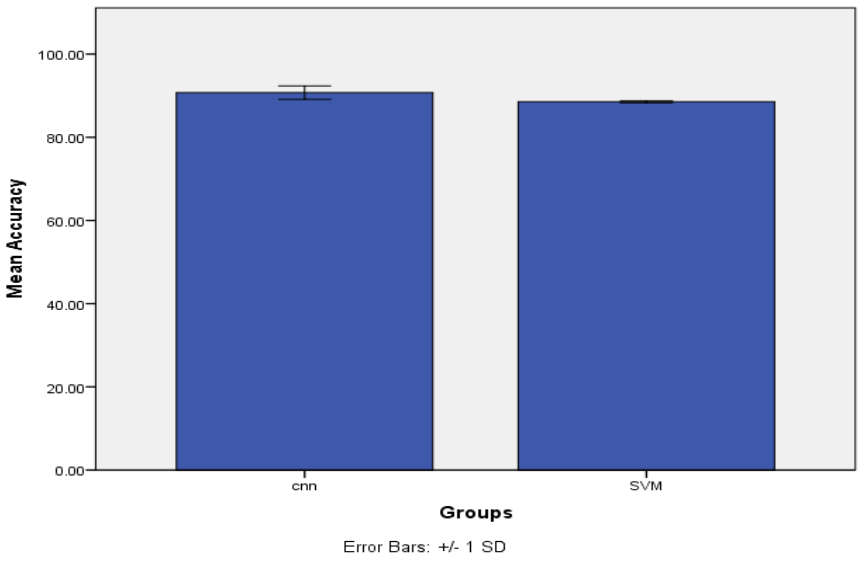

Fig. 1: Bar Chart representing the comparison of Mean Accuracy of Convolutional Neural Networks and Support Vector Machine. Mean accuracy of Convolutional Neural Network is 90.94% which appears to be better than the Support Vector Machine which is 88.54%. The X-axis represents Convolutional Neural Network and Support Vector Machine and Y-axis represents the mean accuracy ± 1 SD