ENHANCING ACCURACY IN DETECTING HUMAN AND COUNTING USING CONVOLUTIONAL NEURAL NETWORK OVER SUPPORT VECTOR MACHINE

Prahalathan B[1], S.Christy[2*]

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

Abstract

Aim: To enhance the accuracy in human detection and human counting using a novel Convolutional Neural Network (CNN) over Support Vector Machine (SVM) algorithm.
Materials and Methods: This research study contains two groups, group one is novel convolutional neural network and group two is support vector machine algorithm. Each group consists of a sample size of 10 and the study parameters include alpha value 0.05, beta value 0.2, and the power value 0.8. Their accuracies are also compared with each other using different sample sizes.
Results: The novel convolutional neural network is 94.302 more accurate than the support vector machine algorithm of 784.302% in human detection and human counting.
Conclusion: The CNN model is significantly better than the SVM in detecting and counting humans. It can be also considered as a better option for human detection and counting the total number of humans in a frame.

Keywords: Novel Convolutional Neural Network, Support Vector Machine, Human Detection, Accuracy, Human Counting.

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

1. Introduction

A visual surveillance system's ability to reliably recognise human beings is critical for a variety of applications, including abnormal event detection, human gait characterisation, congestion analysis, person identification, gender categorization, and fall detection for the elderly. The initial stage in the detecting procedure is to identify a moving item (Fitzgibbon et al. 2012). Background subtraction, optical flow and spatio-temporal filtering methods might be used to identify objects (Hu 2005) when a moving item is discovered, it can be categorized as a person using shape-based (Jürgens 2018), texture-based, or motion-based characteristics (Yang and Ahuja 2012). Similar applications of Human detection and counting are Human Detection, Image Classification (Liu 2012).

There are about 82 articles in Google Scholar, Science direct and 21 in Scopus related to this study. In a study by Wen-Chang Wang, aims at various detection techniques used to detect and count humans using the convolutional neural network (Zhu 2007). (Wu 2012) paper tells about both supervised and unsupervised learning. These two learnings combined to find and count the objects (Fu, Ma, and Xiao 2012). This paper tells a simple and efficient bottom-up saliency detection model of a discriminative histogram feature metric by combining multiple color space and gradient magnitude channels to handle complex images (Zheng et al. 2016). Shape-based, motion-based, and texture-based methods (Rutishauser et al., n.d.; Koch and Walther 2010) are several types of object classification techniques (Wang 2017). The benchmark datasets' properties are discussed, as well as the most common uses of person detection.

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 human detection and counting is the availability of real time data sets and the accuracy to be improved. The selection of the algorithm also plays a vital role in Human Detection and Human Counting, So, this research focuses on improved accuracy in human detection using convolutional neural networks over support vector machines.

2. Materials and Methods

This work is carried out at the Data Analytics Lab, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS). The study consists of two sample groups namely convolutional neural network and support vector machine. Each group consists of 10 samples with pre-test power of 0.18. The sample size kept the threshold at 0.05, G power of 80%, confidence interval at 95% and enrolment ratio as 1.

Convolution Neural Network

Novel convolutional neural networks are a type of artificial neural network used to evaluate visual information. Based on the shared-weight architecture of the convolution kernels or filters that slide along input features and give translation equivariant responses known as feature maps. Surprisingly, most Novel Convolutional Neural Networks are only equivariant under translation, rather than invariant. Image and video recognition, recommender systems, image classification, image segmentation, medical image analysis, and natural language processing, among other things, were used in CNN shown in Fig. 1. The input to a CNN is a tensor with the above shape: (number of inputs) x (input height) x (input width) x (number of outputs) x (number of outputs) x (number of outputs) x (number of outputs) x (number of outputs) x (number of output (input channels). The image is abstracted to a feature map, also known as an activation map, after passing through a convolutional layer, with the following shape: (number of inputs) x (feature map height) x (feature map width) x (feature map height) x (feature map width) x (number of inputs) x (number of inputs) x (number of inputs) x (number of inputs) x (number of input (feature map channels). The input is convolved by convolutional layers, which then pass the result on to the next layer. The Pseudocode for CNN is given in Table 1. A cell in the visual brain responds to a given stimulus in a comparable way. Each convolutional neural only processes data for the receptive field in which it is located. Although fully linked feedforward neural networks can be used to learn features and categorize data, there are several limitations.

Support Vector Machines

Support-vector machines (SVM), also known as support-vector networks, are supervised learning models that examine data for classification and regression analysis. SVMs, which are based on statistical learning frameworks and Chervonenkis, are one of the most reliable prediction systems. The Pseudocode for SVM is given in Table 2. An SVM

training algorithm creates a model that assigns new examples to one of two categories, making it a non-probabilistic binary linear classifier, given a series of training examples, each marked as belonging to one of two categories (although methods such as Platt scaling exist to use SVM in a probabilistic classification setting). Classifying data is a typical problem in machine learning as shown in Fig. 2. Assume that some data points are assigned to one of two classes, and the purpose is to determine which class a new data point will be assigned to. A data point is viewed as a displaystyle pp-dimensional vector (a list of displaystyle pp numbers) in support-vector machines, and we want to know if we can separate such points with a displaystyle (p-1)(p-1)-dimensional hyperplane.

The minimum requirement to run the softwares used here are intel core I3 dual core cpu, 4GB RAM,64 bit OS,1TB Hard disk space personal computer and Software specification includes Windows 8/10/11, Python 3.8 and MS-Office. The Human is Detected by the layering method. The moving object is found using different layers of image and the object is shown in a rectangular box and tracking them until they react out of the frame. To perform Human Detection and Counting the real time data sets used are images and videos. The input data sets for the proposed work is collected from kaggle.com. The data sets consist of 12 videos and 1461 images.

Statistical Package for the Social Sciences (SPSS) version 26 software tool was used for statistical analysis. An independent sample T-test was conducted for accuracy. Standard deviation, standard mean errors were also calculated using the SPSS Software tool. The independent attributes were frame, speed and images. The dependent attributes were accuracy and precision. An analysis is done with the collected data. The significance values of proposed and existing algorithms contain group statistical values of proposed and existing algorithms.

3. Results

The group statistical analysis on the two groups shows Novel Convolutional Neural Networks (Group 1) has more mean accuracy than Support-vector machines (Group 2) and the standard error mean is slightly less than Novel Convolutional Neural Networks. The Novel Convolutional Neural Networks algorithm scored an accuracy of 94.302% and Support-vector machines scored 83.422%. The accuracies are recorded by testing the algorithms with 10 different sample sizes and the average accuracy calculated for each algorithm is shown in Table 3.

In SPSS, the datasets are prepared using 10 as sample size for Convolutional Neural Networks and Support-vector machines. Group id is given as a grouping variable and Lot area is given as the testing variable. Group id is given as 1 for Convolutional Neural Networks and 2 for support vector machines. Group statistics is shown in Table 4, Two Independent Sample T-Tests are shown in Table 5. Comparison of novel convolutional neural networks and support vector machines in terms of mean accuracy. The mean accuracy of convolutional neural network better than support vector machines is calculated and the graph obtained is shown in Fig. 3.

4. Discussion

From the results of this study, novel convolutional neural networks are proved to be having better accuracy than the support-vector machines. CNN has an accuracy of 94.302% whereas SVM has an accuracy of 83.422% in detecting human and human counting. The group statistical analysis on the two groups shows that novel convolutional neural networks (Group 1) has more mean accuracy than support-vector machines (Group 2) and the standard error mean including standard deviation mean is slightly less than novel convolutional neural networks

For many years the researchers have focused on the Computer vision based detection system (Niu 2011).The planning and flow of image information is extract useful and precise, the human detection and counting using series of frames and and images in CNN the accuracy for the given image is 90%(Paul, Haque, and Chakraborty 2013). human detection and counting using a series of frames and images in SVM the accuracy for the given image is 90%(Rasmussen and Rouse 2013).

Similar findings showed that human detection and counting was performed using convolutional neural network based categorization which gave an accuracy of 94 (El Moataz et al. 2020)Attention based Human Detection and Counting paper provided an accuracy of 74.6% (Institute of Electrical and Electronics Engineers 2007).

The limitation in this model is that the accuracy of SVM requires full labeling of input data. Most of the data is simulated from nature which is far from reality effective data preprocessing techniques, of CNN to make the frame processing even accurate and precise detection in the future

5. Conclusion

Based on the experimental results, the Novel Convolutional Neural Networks has been proved to detect humans more significantly better than Support-vector machines.

Declarations

Conflicts of Interest

No conflicts of interest in this manuscript.

Author Contributions

Author PB was involved in data collection, data analysis, data extraction, manuscript writing. Author CS was involved in conceptualization, data validation, and critical review of the 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 organizations for providing financial support that enabled us to complete the study.

1. Best Enlist Private LTD.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

6. References

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.

El Moataz, Abderrahim, Driss Mammass, Alamin Mansouri, and Fathallah Nouboud. 2020. Image and Signal Processing: 9th International Conference, ICISP 2020, Marrakesh, Morocco, June 4–6, 2020, Proceedings. Springer Nature.

Fitzgibbon, Andrew, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid. 2012. Computer Vision – ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012. Proceedings, Part VII. Springer.

Fu, Huiyuan, Huadong Ma, and Hongtian Xiao. 2012. "Real-Time Accurate Crowd Counting Based on RGB-D Information." In 2012 19th IEEE International Conference on Image Processing. IEEE. https://doi.org/10.1109/icip.2012.6467452.

Hu, Anqi. 2005. Human Detection in Security Video.

Institute of Electrical and Electronics Engineers. 2007. Special Issue on Human Detection and Recognition.

Jürgens, Arne Gerd. 2018. Human Counting in Public Transport.

Koch, C., and D. Walther. 2010. "Bottom-up Visual Attention to Salient Proto-Object Regions." Journal of Vision. https://doi.org/10.1167/6.6.826.

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.

Liu, Dongyou. 2012. Molecular Detection of Human Parasitic Pathogens. CRC Press.

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.

Niu, Xiaoxiao. 2011. Fusions of CNN and SVM Classifiers for Recognizing Handwritten Characters.

Paul, Manoranjan, Shah M. E. Haque, and Subrata Chakraborty. 2013. "Human Detection in Surveillance Videos and Its Applications - a Review." EURASIP Journal on Advances in Signal Processing 2013 (1): 1–16.

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.

Rasmussen, Jens, and William B. Rouse. 2013. Human Detection and Diagnosis of System Failures. Springer Science & Business Media.

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.

Rutishauser, U., D. Walther, C. Koch, and P. Perona. n.d. "Is Bottom-up Attention Useful for Object Recognition?" Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. https://doi.org/10.1109/cvpr.2004.1315142.

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.

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.

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.

Wang, Lu. 2017. Three-Dimensional Model Based Human Detection and Tracking in Crowded Scenes.

Wu, Z. 2012. Human Detection and Pose Estimation for Motion Picture Logging and Visualisation.

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.

Yang, Ming-Hsuan, and Narendra Ahuja. 2012. Face Detection and Gesture Recognition for Human-Computer Interaction. Springer Science & Business Media.

Zheng, Ling, Shuhan Chen, Xuelong Hu, and Lifeng Zhang. 2016. "A Specified Saliency Detection for Human Detection." Proceedings of the 4th IIAE International Conference on Intelligent Systems and Image Processing 2016. https://doi.org/10.12792/icisip2016.064.

Zhu, Qiang. 2007. Human Detection in Images and Videos. ProQuest.

TABLES AND FIGURES

Table 1. Pseudocode for Novel Convolutional Neural Networks.

// I : Input dataset records
1. Import the required packages.
2. Convert the image into machine readable after the extraction feature.
3. Assign the image to the output variables.
4. Using the model function, assign it to the variables.
5. Compiling the model using metrics as accuracy.
6. Evaluate the output
7. Get the accuracy of the model.
OUTPUT :///Accuracy

Table 2. Pseudocode for Support Vector Machines

// I : Input dataset image
INPUT: Capture Image
Step 1. Pre-process the image of the particular humans
Step 2. Segment and normalize the images.
Step 3. Extract the feature vector of each normalized candidate
Step 4. Train SVMs based on a saved sample database.
Step 5. Recognize the human by the set of SVMs trained in advance.
Step 6. If there are no more unclassified samples, then STOP.
Step 7. Add these test samples into their corresponding database for further training. OUTPUT: Numbers of humans detected/counted
OUTPUT : //Accuracy

Table 3. Accuracy of human detection using Novel Convolutional Neural Networks and Support Vector Machines

Iteration	CNN Accuracy	SVM Accuracy
Test 1	91.77	83.86
Test 2	94.72	85.96
Test 3	94.48	85.75
Test 4	91.36	86.55
Test 5	97.72	84.86
Test 6	92.89	86.55
Test 7	91.34	86.98
Test 8	94.98	86.99
Test 9	93.21	92.98
Test 10	94.08	95.98

Table 4. Group Statistical analysis for convolutional neural network and support vector machines Mean, Standard Deviation and standard error mean are determined

Group Statistics					
	Group	N	Mean	Std.Deviation	Std.Error Mean
Accuracy	CNN	10	94.302	2.56845	0.81221
	SVM	10	83.422	1.81097	0.57268
Loss	CNN	10	5.698	2.56845	0.81221
	SVM	10	16.578	1.81097	0.57268

Table 5. Independent sample T-test t is performed on two groups for significance and standard error determination. p value is less than 0.05 (0.004) and it is considered to be statistically significant with 95% confidence interval

		Levene's test for Equality Of variance		T-Test for equality of mean						
				t	df	Sig(2 tailed)	Mean difference	Std.Error Difference	95% confidence of Difference	
		F	Sig						Lower	Upper
Accuracy	**Equal variances assumed**	.238	.031	-3.300	18	.004	-5.83600	1.76860	-9.55169	-2.12031
	Equal Variances not assumed			-3.300	17.88	.004	-5.83600	1.76860	-9.55341	-2.11859
Loss	**Equal variances assumed**	.238	.031	-3.300	18	.004	5.83600	1.76860	2.12031	9.55169
	Equal Variances not assumed			-3.300	17.88	.004	5.83600	1.76860	2.11859	9.55341

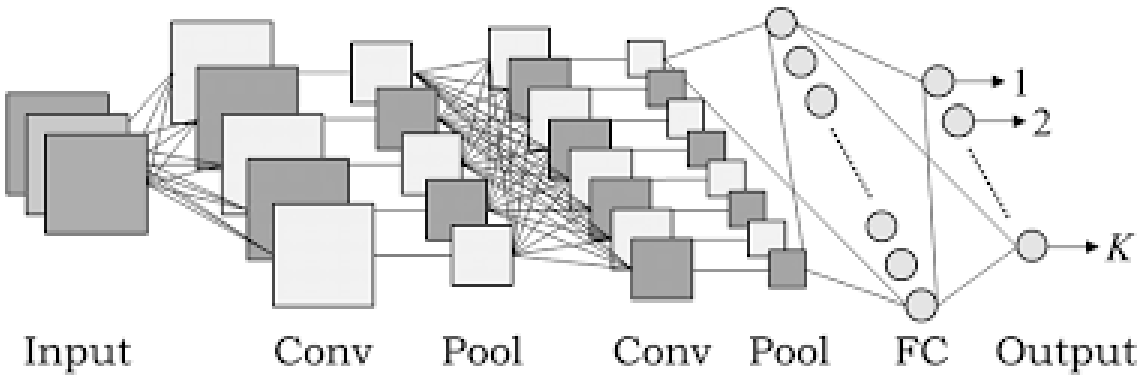

Fig. 1. The architecture of a CNN is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

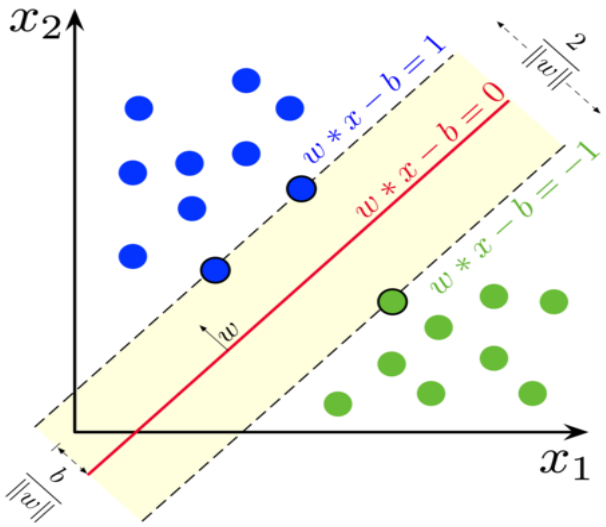

Fig. 2. Hyperplane for Support Vector Machine, The SVM algorithm's objective is to find the optimal line or decision boundary for categorizing n-dimensional space into classes so that additional data points may be readily placed in the proper category in the future. A hyperplane is the name for the optimum choice boundary

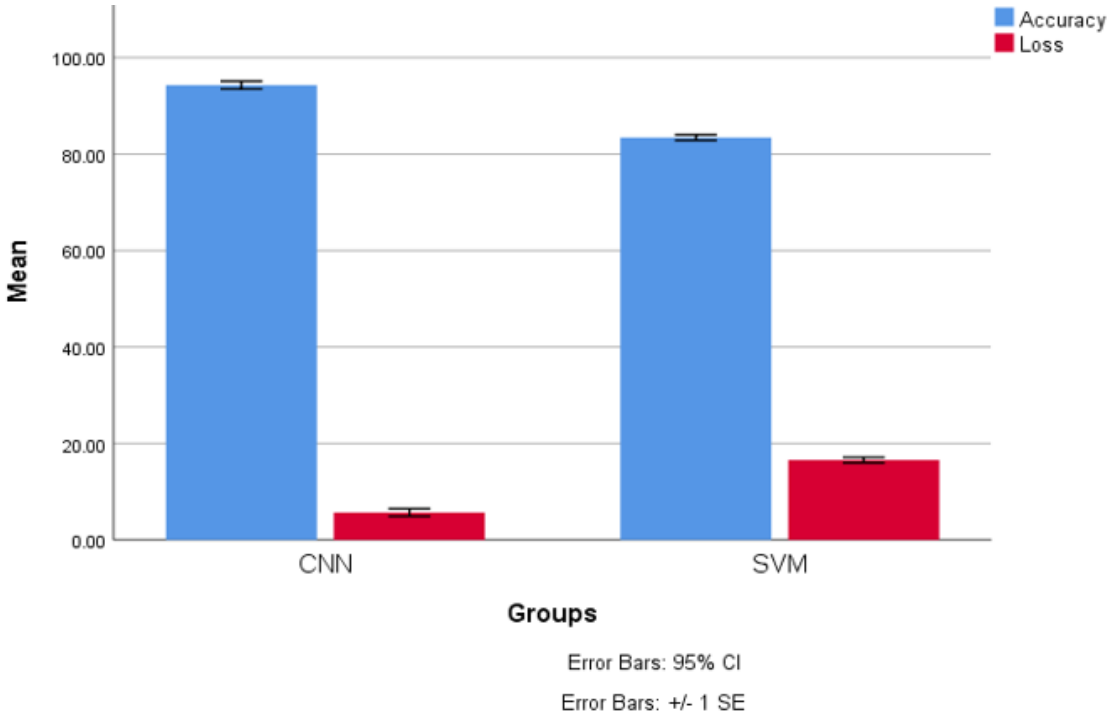

Fig. 3. Comparison of novel convolutional neural networks and support vector machines in terms of mean accuracy. The mean accuracy of is convolutional neural network better than support vector machines. The standard deviation of novel convolutional neural networks is slightly better than X Axis: convolutional neural network vs support vector machines Algorithm. Y Axis: Mean accuracy of detection ± 1 SE.