Original Article

A Real Time Object Detection in Integral Part of Computer Vision using Novel Image Classification of Faster R-CNN Algorithm over Fast R-CNN Algorithm

M.Srikar[1], Malathi K[2]

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Tamil Nadu, India, PinCode: 602105
[2]Project Guide, Corresponding Author, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Tamil Nadu, India, PinCode: 602105

Abstract

Aim: The objective of the work is to increase the precision of object detection using novel image classification using machine learning algorithms. **Materials and Methods:** The categorising is performed by adopting a sample size of n = 10 in Faster R-CNN (RPN) and sample size n = 10 in Fast R-CNN (Selective Search) algorithms with a sample size = 10 and the G-Power analysis was carried out with 80% and confidence interval 95%. **Results and Discussion:** The observation of the outcomes shows that the Faster R-CNN using region proposal networks has a high accuracy of 81.72% in comparison with the Selective Search based Fast R-CNN 79.61%. A statistically significant difference exists between the research groups with p=0.028 (2 tailed) ($p<0.05$). **Conclusion:** Detection of objects with high accuracy using machine learning algorithms shows that the regional proposal network based Faster R-CNN generates higher accuracy than the Selective search (Fast RCNN) algorithm.

Keywords: Object Detection, Region Proposal Networks, Convolutional Neural Network, Novel Image Classification, Softmax Layer, Translation-Invariant Anchors.

https://doi.org/10.47750/pnr.2022.13.S04.202

INTRODUCTION

In today's fast-paced environment, computer vision has a critical role in identifying things with higher speed and precision. The purpose of this work is to increase the accuracy of recognizing items in computer vision through image classification using machine learning algorithms by means of region proposal networks (Lv et al. 2018). Computer vision's major process is object detection, it is quite a complex problem for a system (Arulprakash and Aruldoss 2021). Since robotic vision is a major thing in the trending world with driverless vehicles, traffic detections, tracking items and many security purposes (Jiang et al. 2018). It was focused to develop a low cost object recognition system and enhance the accuracy and speed in order to identify items with accuracy and identifying correctly (Novotny and Matas 2015).

Predicting items using item identification algorithms for over past years and several surveys have been published in the last years over 17,600 articles from Google Scholar, 7761 journals IEEE Xplore digital library, 975 research articles from ScienceDirect. Among all the articles and journals, the most cited paper is (Ren et al. 2017). The model produced by the (Ren et al. 2017) is very accurate and much more efficient compared to the Selective search model. The classifier and regressor are part of region proposal networks. The probability of the desired object in the input image is recognized by the classifiers. Through Regressor the coordinates of the model output will be regressed. (Zheng, Chen, and Hu 2019). The concepts of Anchors are introduced in this model. The main function of the sliding window is Anchor (Laban et al. 2019). As a result, this algorithm is resistant to translations, and translational invariance is one of its important qualities (Soni 2019).

Our institution is passionate about high quality evidence based research and has excelled in various fields (Parakh et al. 2020; Pham et al. 2021; Perumal, Antony, and Muthuramalingam 2021; Sathiyamoorthi et al. 2021; Devarajan et al. 2021; Dhanraj and Rajeshkumar 2021; Uganya, Radhika, and Vijayaraj 2021; Tesfaye Jule et al. 2021; Nandhini, Ezhilarasan, and Rajeshkumar 2020; Kamath et al. 2020). This method which was used before has less accuracy on detecting objects. It is very important to predict and determine the objects in minimum time to prevent issues. For example, driverless vehicles need to identify the object within a fraction of seconds and analyse the situation to move forward, otherwise there will be many consequences. The research gap found to be the low accuracy and constant speed to identify the object is lacking in the previous model. Different attributes proposed in the previous system have become less efficient as the environment changes. In order to sequence the methods and techniques in this research generally fairs better than Selective search (Fast R-CNN) which is based on the softmax layer (Girshick 2015). It also takes a lot more time to render all the images to train the model compared to Faster R-CNN (SNMS) (Cai and Vasconcelos 2018). The objective of the work is to increase the precision of detecting items using region proposal networks over Selective search machine learning algorithms such as Faster region convolutional neural network over Fast region convolutional neural network to improve accuracy.

MATERIALS AND METHODS

The work was carried out in the Image Processing Lab, Department of Computer Science and Engineering, Saveetha School of Engineering, SIMATS. Basically, it is considered with two groups of classifiers namely Faster R-CNN (Regional Proposal Network) and Fast R-CNN (Selective search), it is used to identify objects in images with various image datasets and labels. Group 1 is the Faster R-CNN with the sample size of 10 and Group 2 is the Fast R-CNN with sample size of 10 and it was used to compare for more accuracy score and loss values for choosing the best algorithm to detect objects correctly. A Sample size of 20 different values has been calculated with the G power having 80%, 95% confidence interval, alpha as 0.05 and beta as 0.2 values and standard deviation for Faster R-CNN = 1.68131 and Fast R-CNN = 2.81191.

The dataset used in the model for prediction process has been collected from the Kaggle website. The dataset can be found as the Fruit Recognition dataset created by the owner (Chris Gorgolewski). The dataset contains different label images around 660 files of apple class labels and equally other class labels named Class A and Class B respectively.The images were collected in various kinds of conditions to obtain accurate results even under different low light conditions. The data is further processed for training and testing datasets and has been split up in the ratio of 70% and 30% respectively. Adapting Google Collaboration as the software which will provide a wide Range of GPUs. The Gpu used is the Nvidia Tesla K80. The configurations of the system used are Intel i5, 5th Gen CPU @ 2.8GHZ, 8 GB RAM, and 64-bit OS.

Faster R-CNN

Faster R-CNN is one of the primary techniques. It is a deep convolutional network which serves as a core approach by importing regions and employs Translation-Invariant Anchors with Novel Image Classification technique (Faster R-CNN). The full system acts as a sole, cohesive model with heed processes. This model informs the region-convolutional neural layer to process the image. The Region Proposal Network includes a picture as input later converts into a collection of rectangles, and object suggestions, each with a defined value. This is accomplished through the use of a fully connected layer including certain additional layers with a Fast R-CNN system, an item identification model that employs Translation-Invariant Anchors. The areas are created by a tiny network on a complex feature map. The final layer shares the feature map as the outcome of the input image.

Pseudocode

Step 1: Import tensorflow
Step 2: from weights import weightsPath
Step 3: Import region based classifiers
Step 4: Blob → cv2.dnn.blobFromImage(frame)
Step 5: net.setInput(blob) #passing the blob as an input to the ConvNets
Step 6: box → boxes[0,0,i], left → int(frameW*box[3]) #Acquiring bounding boxes
Step 7: cv2.rectangle(frame, color,2) #drawing bounding boxes
Step 8: Resize and reshape the image and form a cluster pixel
Step 9: Accuracy of the Faster R-CNN.

Fast R-CNN

The second approach is developed with selective search (Fast R-CNN) and the primary approach is Faster R-Convolutional Neural Network which developed using region proposal networks. The Fast R-CNN is given a picture with a set of item suggestions as input. The model takes the input and analyses the entire picture with a pipeline of neural networks with softmax layers on top to obtain a feature map. After that it derives some item propositions from the feature map based on the region of interest (ROI) layer. Individual feature propositions will undergo a multiple sequence of fully connected (FC) layers, which eventually part into two related output layers. This yields a softmax layer chance to predict for K object categories as well as to catch required details of the image, and another layer that yields four integer values for each K object category. Each four-value set encodes revised bounding-box coordinates for one of the K categories (Wang, Shrivastava, and Gupta 2017).

Pseudocode
Step 1: Import selective search layer
Step 2: Import tensorflow
Step 3: Conf Threshold → 0.5, maskThreshold → 0.3
Step 4: Import classifiers
Step 5: From config import configPath,configPath → .../rcnn_inception_v2_coco_Data.pbtxt
Step 6: net.setInput(blob) #passing the blob as an input to the ConvNets
Step 7: for k in range(numDetections):
Step 8: box → boxes[0,0,k]
Step 9: cv2.rectangle(frame,(startX,startY),(endX,endY),color,2) #drawing bounding boxes
Step 10: Accuracy of the Fast R-CNN regional proposal network.

Statistical Analysis
IBM SPSS (version 26) statistical tool is used for analysis. 10 iterations were done with a sample size of 10 for each of the algorithms and predicted accuracy was noted for analysis. And the sample size was calculated using G power of 80% for each group and confidence interval of 95% for the two groups. The value obtained from the iterations of a total 20 samples 10 iterations from two individual algorithms and conducted an Independent Sample T-test using SPSS statistical tool. The dependent data sets are ImageNet. The independent values are AlexNet, VGGNet (Ren et al. 2017). The fragmented analysis has been done with independent variables labelled images, bounding box coordinates and dependent variables are accuracy, duration, and feature graded object graph to find the objects with more accuracy and speed.

RESULTS

The model has been trained through different label images. Group statistics of Faster R-CNN by Fast R-CNN by grouping with iterations sample size of 10 is shown in Table 1. It was observed that Faster R-CNN is having greater precision value over Fast R-CNN model.

In Table 2, The statistical analysis of two independent groups shows that the Faster R-CNN have higher accuracy mean 81.72% compared to selective search based Fast R-CNN with accuracy 79.61%. Standard Deviation = 1.68131, Standard Error Mean = .53168. Descriptive Independent Sample Test of Accuracy is applied for the dataset in SPSS.

In Table 3, The Significant value= 0.600, Mean Difference= 2.11000 and confidence interval = (-.06662 - 4.28662) of Faster R-CNN and Fast R-CNN respectively. The significance value is $p=0.028$ ($p<0.05$) with an independent sample T-Test. Images, labels and tested datasets independent variables.

Figure 2, represents the simple bar graph comparison of mean accuracy on Faster R-CNN and Fast R-CNN.

DISCUSSION

This work on identifying items in pictures, subsequently termed object detection, is very essential in many industries in order to comprehend different scenarios by computer systems (Pathak, Pandey, and Rautaray 2018). Identifying items using Faster R-CNN based on the region proposal network (Ren et al. 2017) is pragmatically proven to be highly effective than Fast R-CNN which is based on selective search with softmax layer, the same can be observed using the Independent sample T-test. The mean accuracy is 81.72% for Faster R-CNN and 79.61% for Fast R-CNN. And the statistical 2-tailed significant difference in accuracy for two models is 0.028 ($p<0.05$). The core argument is to prove that detection of objects in various conditions, Faster R-CNN through Novel image classification is a better method. And in many of the recent findings, it has been observed that the Region proposal network with Translation-Invariant Anchors is the most focused and better method of detecting objects with more accuracy than Fast R-CNN (Girshick 2015).

Object identification uses a bounding box approach and also through a softmax layer to acknowledge and localise each object instance . It is commonly employed in driverless cars as a classic challenge in vision through computers (Akhtar and Mian 2018) and assistive robots (Subudhi 2009). Scale invariant feature transform is one of the commonly supported algorithms for identifying entities (“Real-Time Object Detection and Localization with SIFT-Based Clustering” 2012). and histogram of oriented gradient (HOG) (Patel et al. 2020). These approaches extract the thin features and scan the picture for locations with the highest class-specific response. However, these strategies are susceptible to noise and work effectively only on limited object categories. These issues restrict the applicability of standard object identification models.
The success of cascade for identifying items (Cai and Vasconcelos 2018) and instance classification (Lin et al. 2014) had other difficulties and limitations that are demonstrated, which are having slow process speed due to additional layers in the pipeline model that creates a huge process time to compile the picture where likely multilayered object detection might not be a future paradigm for quick and precision trade off. The most common application of bounding boxes is to evaluate general item recognition methods using innovative image classification and Translation-invariant anchors instead of softmax layers (Bauckhage and Tsotsos 2005), therefore this approach is adopted in this work. However, as the future enhancement, the research community progresses from picture level classification to single object localisation, generic object recognition and pixel wise object segmentation, future problems are predicted to be at the pixel level.

CONCLUSION

The objective of the research is to increase the precision of object detection using novel image classification using computer vision algorithms. The accuracy and speed is improved significantly by use of the region proposal network method. The outcome of the Faster R-CNN based on region proposal networks showed higher accuracy 81.72% than the Selective search based Fast R-CNN 79.61%.

DECLARATION

Conflicts of Interest

No conflicts of interest in this manuscript.

Authors Contributions

Author MS was involved in data collection, data analysis, algorithm framing, implementation and manuscript writing. Author KM was involved in the designing the workflow, guidance and Critical review of manuscript.

Acknowledgements

The authors would like to extend their appreciation to Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure and constant help to carry out this work successfully.

Funding

We would like to express our gratitude to the following organisations for giving financial support that allowed us to finish the work.

1. Qbec Infosol, Chennai.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

References

1. Akhtar, Naveed, and Ajmal Mian. 2018. “Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey.” *IEEE Access*. https://doi.org/10.1109/access.2018.2807385.
2. Arulprakash, Enoch, and Martin Aruldoss. 2021. “A Study on Generic Object Detection with Emphasis on Future Research Directions.” *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2021.08.001.
3. Bauckhage, C., and J. K. Tsotsos. 2005. “Bounding Box Splitting for Robust Shape Classification.” *IEEE International Conference on Image Processing 2005*. https://doi.org/10.1109/icip.2005.1530096.
4. Cai, Zhaowei, and Nuno Vasconcelos. 2018. “Cascade R-CNN: Delving Into High Quality Object Detection.” *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. https://doi.org/10.1109/cvpr.2018.00644.
5. Devarajan, Yuvarajan, Beemkumar Nagappan, Gautam Choubey, Suresh Vellaiyan, and Kulmani Mehar. 2021. “Renewable Pathway and Twin Fueling Approach on Ignition Analysis of a Dual-Fuelled Compression Ignition Engine.” *Energy & Fuels: An American Chemical Society Journal* 35 (12): 9930–36.
6. Dhanraj, Ganapathy, and Shanmugam Rajeshkumar. 2021. “Anticariogenic Effect of Selenium Nanoparticles Synthesized Using Brassica Oleracea.” *Journal of Nanomaterials* 2021 (July). https://doi.org/10.1155/2021/8115585.
7. Girshick, Ross. 2015. “Fast R-CNN.” *2015 IEEE International Conference on Computer Vision (ICCV)*. https://doi.org/10.1109/iccv.2015.169.
8. Jiang, Xiaoyue, Abdenour Hadid, Yanwei Pang, Eric Granger, and Xiaoyi Feng. 2018. *Deep Learning in Object Detection and Recognition*. Springer.
9. Kamath, S. Manjunath, K. Sridhar, D. Jaison, V. Gopinath, B. K. Mohamed Ibrahim, Nilkantha Gupta, A. Sundaram, P. Sivaperumal,

S. Padmapriya, and S. Shantanu Patil. 2020. "Fabrication of Tri-Layered Electrospun Polycaprolactone Mats with Improved Sustained Drug Release Profile." *Scientific Reports* 10 (1): 18179.
10. Laban, Noureldin, Bassam Abdellatif, Hala M. Ebeid, Howida A. Shedeed, and Mohamed F. Tolba. 2019. "Convolutional Neural Network with Dilated Anchors for Object Detection in Very High Resolution Satellite Images." *2019 14th International Conference on Computer Engineering and Systems (ICCES)*. https://doi.org/10.1109/icces48960.2019.9068145.
11. Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. 2014. "Microsoft COCO: Common Objects in Context." *Computer Vision – ECCV 2014*. https://doi.org/10.1007/978-3-319-10602-1_48.
12. Lv, Xiaogang, Xiaotao Zhang, Yinghua Jiang, and Jianxin Zhang. 2018. "Pedestrian Detection Using Regional Proposal Network with Feature Fusion." *2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA)*. https://doi.org/10.1109/ipta.2018.8608159.
13. Nandhini, Joseph T., Devaraj Ezhilarasan, and Shanmugam Rajeshkumar. 2020. "An Ecofriendly Synthesized Gold Nanoparticles Induces Cytotoxicity via Apoptosis in HepG2 Cells." *Environmental Toxicology*, August. https://doi.org/10.1002/tox.23007.
14. Novotny, David, and Jiri Matas. 2015. "Cascaded Sparse Spatial Bins for Efficient and Effective Generic Object Detection." *2015 IEEE International Conference on Computer Vision (ICCV)*. https://doi.org/10.1109/iccv.2015.137.
15. Parakh, Mayank K., Shriraam Ulaganambi, Nisha Ashifa, Reshma Premkumar, and Amit L. Jain. 2020. "Oral Potentially Malignant Disorders: Clinical Diagnosis and Current Screening Aids: A Narrative Review." *European Journal of Cancer Prevention: The Official Journal of the European Cancer Prevention Organisation* 29 (1): 65–72.
16. Patel, Chirag I., Dileep Labana, Sharnil Pandya, Kirit Modi, Hemant Ghayvat, and Muhammad Awais. 2020. "Histogram of Oriented Gradient-Based Fusion of Features for Human Action Recognition in Action Video Sequences." *Sensors* 20 (24). https://doi.org/10.3390/s20247299.
17. Pathak, Ajeet Ram, Manjusha Pandey, and Siddharth Rautaray. 2018. "Application of Deep Learning for Object Detection." *Procedia Computer Science*. https://doi.org/10.1016/j.procs.2018.05.144.
18. Perumal, Karthikeyan, Joseph Antony, and Subagunasekar Muthuramalingam. 2021. "Heavy Metal Pollutants and Their Spatial Distribution in Surface Sediments from Thondi Coast, Palk Bay, South India." *Environmental Sciences Europe* 33 (1). https://doi.org/10.1186/s12302-021-00501-2.
19. Pham, Quoc Hoa, Supat Chupradit, Gunawan Widjaja, Muataz S. Alhassan, Rustem Magizov, Yasser Fakri Mustafa, Aravindhan Surendar, Amirzhan Kassenov, Zeinab Arzehgar, and Wanich Suksatan. 2021. "The Effects of Ni or Nb Additions on the Relaxation Behavior of Zr55Cu35Al10 Metallic Glass." *Materials Today Communications* 29 (December): 102909.
20. "Real-Time Object Detection and Localization with SIFT-Based Clustering." 2012. *Image and Vision Computing* 30 (8): 573–87.
21. Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. 2017. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (6): 1137–49.
22. 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.
23. Soni, Divyanshu. 2019. "Translation Invariance in Convolutional Neural Networks." Medium. November 13, 2019. https://divsoni2012.medium.com/translation-invariance-in-convolutional-neural-networks-61d9b6fa03df.
24. Subudhi, Bidyadhar. 2009. *Computational Intelligence, Control and Computer Vision in Robotics and Automation*.
25. Tesfaye Jule, Leta, Krishnaraj Ramaswamy, Nagaraj Nagaprasad, Vigneshwaran Shanmugam, and Venkataraman Vignesh. 2021. "Design and Analysis of Serial Drilled Hole in Composite Material." *Materials Today: Proceedings* 45 (January): 5759–63.
26. Uganya, G., Radhika, and N. Vijayaraj. 2021. "A Survey on Internet of Things: Applications, Recent Issues, Attacks, and Security Mechanisms." *Journal of Circuits Systems and Computers* 30 (05): 2130006.
27. Wang, Xiaolong, Abhinav Shrivastava, and Abhinav Gupta. 2017. "A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection." *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2017.324.
28. Zheng, Jingye, Dihu Chen, and Haifeng Hu. 2019. "Multi-Scale Proposal Regression Network for Temporal Action Proposal Generation." *IEEE Access*. https://doi.org/10.1109/access.2019.2933360.

Tables And Figures

Table 1. Comparison of accuracy values of Faster R-CNN and Fast R-CNN models. The Faster R-CNN obtained accuracy of 81.72% compared to Fast R-CNN having 79.61%. The significant 2-tailed value for the two groups is P = 0.028 (p<0.05).

Sample (N)	Dataset size	Faster R-CNN	Fast R-CNN
1	662	80.35	78.48
2	554	80.25	78.56
3	510	79.37	79.57
4	494	82.96	78.88
5	384	82.46	79.86
6	287	83.56	80.35
7	250	81.57	78.57

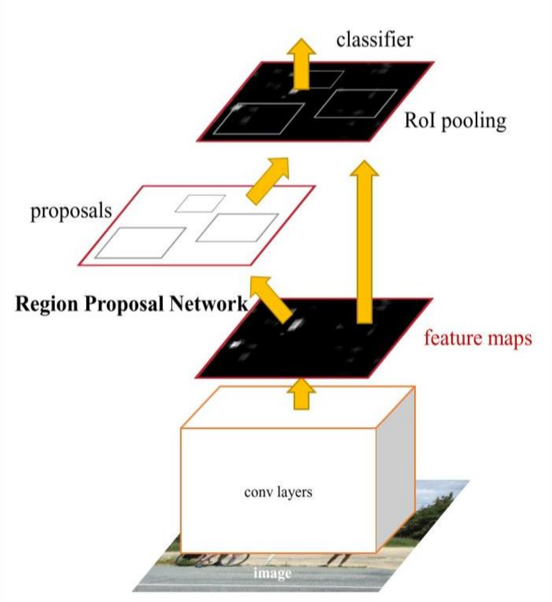
8	150	82.36	80.61
9	134	79.35	79.95
10	87	83.94	79.44

Table 2. Group Statistics of regional proposal network based Faster R-CNN and Fast R-CNN by grouping the iterations with sample size 10, Mean = 81.7280 and 79.6180 respectively, Standard Deviation = 1.68131. Descriptive Independent Sample T-Test of Accuracy is applied for the dataset in SPSS.

	Group	N	Mean	Std. Deviation	Std.Error Mean
Accuracy	Faster R-CNN	10	81.7280	1.68131	.53168
	Fast R-CNN	10	79.6180	2.81191	.88920

Table 3. Independent Sample T-Test is applied for the dataset fixing confidence as 95% and level of significance as P = 0.028 ($p<0.05$).

	Levene's test for equality of variances		**t-test for Equality of Means**						
Accuracy	**F**	**Sig.**	**t**	**df**	**Sig.(2-tailed) p**	**Mean Difference**	**Std. Error Difference**	**95% Confidence Interval of the Difference**	
								Lower	**Upper**
Equal variances assumed	.285	.600	2.037	18	0.028	2.1100	1.03603	-.06662	4.28662
Equal variances not assumed			2.037	14.706	0.028	2.11000	1.03603	-.10210	4.32210



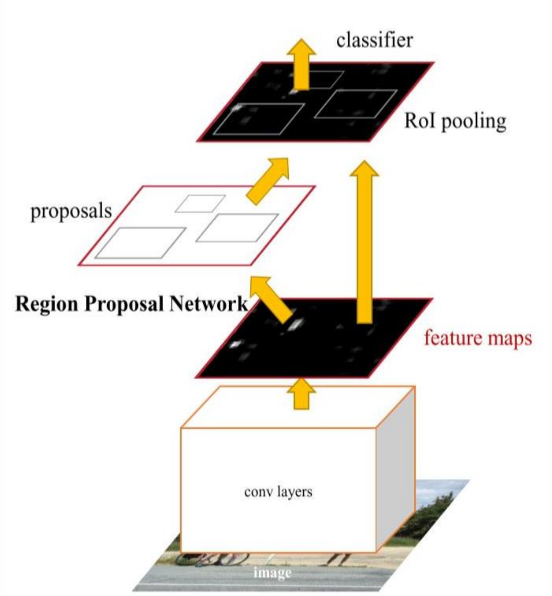


Fig. 1. The Architecture of Faster Region based - Convolutional Neural Network.

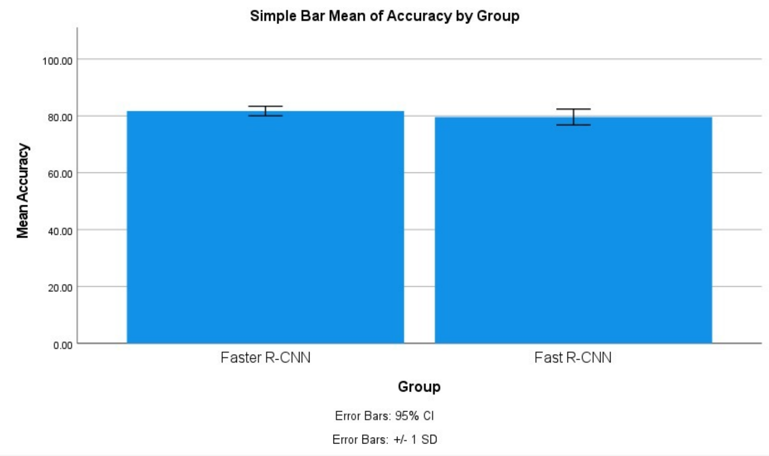


Fig. 2. Comparison of regional proposal network based Faster R-CNN in terms of mean accuracy. As the graph shows the mean accuracy of the Faster R-CNN is greater than the Fast R-CNN. X-Axis: Faster R-CNN and Fast R-CNN models, Y-Axis: Mean Accuracy of detection ± 1 SD.