

Machine Learning for Automated Floor Detection in Building Images for the Insurance Sector

1st Narendra Kumar
Thiruvananthapuram, India
nsparks.100@gmail.com

2nd Rahul Mahajan
Thiruvananthapuram, India
rsmahajan511@gmail.com

3rd Varsha Shaji
Thiruvananthapuram, India
varshas0611@gmail.com

4th Sangeeth G
Thiruvananthapuram, India
sangeethg944@gmail.com

Abstract—In the insurance sector, precise assessment of building properties is crucial for determining appropriate coverage and insurance premium amounts. Effective risk evaluation, formulating policies, and determining premiums depend on the accurate assessment of the relevant properties of the building to be insured. One such critical property is the number of floors in a building. Traditional methods of floor count estimation, such as surveys and sending an investigator, are often time-consuming and quite manual. Further, such methods are prone to inconsistencies and errors, which can be detrimental to the overall decision process. This paper presents a solution for the above challenge by utilising machine learning (ML) techniques for automating floor count detection using building images. In this paper, we propose an object detection model that can accurately estimate the number of floors in a building from front-view images of the building. The proposed solution is a multi-step methodology comprising data preprocessing, feature extraction, and ML model training. It involves the creation of a diverse dataset of annotated building images curated by encompassing various building heights and architectural styles curated explicitly for this task, preprocessing the images, and training and comparing multiple ML models to determine the floor count. The effectiveness of the proposed model is demonstrated in the experimental results, which show that the model can accurately determine the number of floors across diverse building types. Implementing this automated floor count detection system can help simplify property assessment within the insurance sector, reducing cost, processing time, and manual effort. Different object detection algorithms are compared to determine the efficient model for detecting the floors. In future, additional features will be considered in the property assessment. The solution mitigates human error and leads to more precise risk assessments.

Keywords—Building floor Detection, Object Detection, Deep Learning, Assessment

I. INTRODUCTION

In the insurance industry, precise assessment of property prior to insuring is crucial for effective risk evaluation, formulating policies and determining premiums. Prior to insuring a building, determining fundamental aspects of the building, such as the number of floors greatly influences the pricing and coverage decisions made by insurers. Conventional evaluation methods of manually counting floors from building images or cost and time-consuming methods of sending an investigator have been the norm. However, in recent times, with the advancement in technology, we are envisioning a future where automated floor count estimation along with other properties of the building are possible and an agent, from the comforts of his/her desk, would be able to accurately determine these

characteristics. This can highly improve accuracy, efficiency and risk assessment.

Estimating the count of floors from building images is a complex process as buildings can be found with a multitude of architectural styles, with building heights ranging from a few floors to high-storey skyscrapers, across a wide range of photographic perspectives. Addressing this complexity via manual methods has long since proven to be a time and resource-intensive process highly prone to errors. This research paper is meant to address these challenges by introducing a novel approach to automating floor count detection by applying machine learning models.

The objective of this paper is to compare and evaluate multiple machine learning models across different evaluation criteria to assess and determine the best-performing model for the task at hand. By exploiting the power of machine learning models, the proposed solution has the potential to significantly increase the accuracy of floor count detection while providing a way to improve the assessment process, thereby leading to better efficiency within the insurance industry.

In this paper, we dive deep into the complexity of the problem, emphasizing the limitations of traditional floor count estimation methods and putting a strong case forward for the need for a more accurate and efficient reliable solution. This paper proposes a multi-step methodology that includes data collection, pre-processing and machine learning model training.

The experimental results show the potential benefit of our proposed solution, as the comparisons reveal that the trained YOLO(You Only Look Once) model demonstrates robust performance across a wide variety of building types and architectural styles.

As data and machine learning become integral in reshaping the industrial landscape, this research paper can serve as a critical step to realizing the full potential and impact of implementing machine learning in the insurance sector while revolutionizing the way in which insurance firms conduct property assessments.

II. LITERATURE REVIEW

This section discusses the past research on machine learning in the insurance sector and the role of object detection in the implementation of automated floor detection and concludes

The deep learning-based detectors are classified as either One-stage detectors (regression-based) or two-stage detectors (classification-based). Two-stage detectors mark a set of regions of interest in the first stage of detection where the probability of objects is high and in the second stage of detection, detects objects in the marked region using Convolution Neural Networks. One-stage detectors like YOLO on the other hand, detect objects by predicting classes in one go making

in [21]. RCNN was a novel approach in object detection and worked by dividing an image into regions each of which is passed through a CNN to extract features which were then used to classify objects within the regions. Faster R-CNN was developed by Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun [22]. It combined Region proposal Network with a Fast R-CNN detection network leading to it being much more accurate and as the name suggests, faster than its predecessors.

III. METHODOLOGY

This section discusses the methodology used including dataset curation and preparation and model training. The model training was done using Python in Google Colaboratory with GPU.

A. Data Curation and Preparation

The dataset was specially curated for the task at hand by combining building front-view images from multiple sources. We used the Paris Streetview dataset [23], the Zurich dataset which contains 14,900 training images and 100 test images. We also selected images from the Internet with different architectural styles and added them to the training and testing data. The curated dataset includes 244 training images. These images were then annotated using Labellmg and transferred to Google Drive to be prepared and trained in Google Colaboratory. The sample images of building architectures are in Fig.3



Fig. 3: Different building architectures

B. Evaluation Metrics

1) *Average Precision (AP)*: Average precision is used to measure how well the given model can correctly identify and locate objects within the image. The calculation of average precision involves taking the area under the interpolated precision-recall curve. The precision-recall curve is constructed for each class. Then the precision values are interpolated at various recall levels to create a smooth precision-recall curve. AP is particularly useful in the comparison of different models for object detection tasks.

2) *Mean Average Precision (mAP)*: If there are multiple classes detected, then the AP for each class is calculated and the mean of these values is computed for obtaining mAP. This gives a measure of the model's performance across all categories.

3) *Precision and Recall*: Precision is a measure of how accurately the model identifies objects. It is the ratio of true positives to the total number of predicted positives.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

Recall is a measure of how well the model is able to capture all the relevant objects in the image. It is the ratio of true positives to the total number of ground truth positives.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

IV. RESULTS AND DISCUSSION

The two different object detection techniques discussed in this paper are compared in this section. Table 1 compares the two object detection techniques discussed in this paper. Faster RCNN and YOLOv7 use different classification methods. YOLOv7 can run in real-time and is fast, but the technique struggles to find floors from low-resolution images. Resulting images after detection is shown in Fig.4

TABLE I: Comparison of models in terms of mAP@0.5 and mAP@0.5:0.95

Model	mAP@0.5	mAP@0.5:0.95
YOLOV7	87.7	60.2
Faster RCNN	78.7	47.9

TABLE II: Comparison of models in terms of Precision and Recall

Model	Precision	Recall
YOLOV7	80.7	85.3
Faster RCNN	69.3	66.3

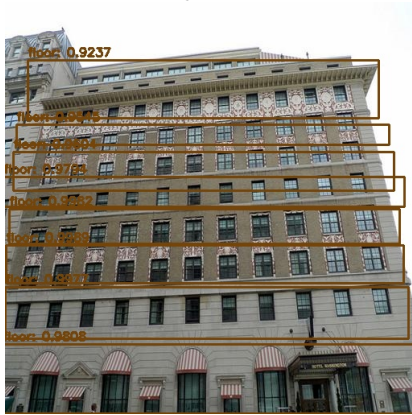
There is an issue when the structure of the building is different; detection omits a certain portion of the building structure. Even a single floor has multiple detections. This issue can be reduced by adding more diverse architectural buildings. Other factors in positioning, the angle of the image, are also vital to the detection. In some cases, trees or other objects will influence the training process. Lastly, both object detection methods can be trained further. Still, the issue is that training takes time, and results will vary depending on the image taken positions and the building structure method. This motivates further research to use a diversified set of building images.

V. CONCLUSION AND FUTURE SCOPE

This study has demonstrated the comparison of the YOLOv7 model and Faster RCNN using building image data for object detection. The results showed that YOLOv7 performs better as the mean Average Precision is higher than Faster RCNN. Future work that can be done is implementing this study with



((a)) Using Faster RCNN



((b)) Using YOLOv7

Fig. 4: Detected floors

more data enabling high-performance hardware. Also, as a product in automating the process of estimating the value of a property, area estimation and evaluating other features of the property will be crucial.

REFERENCES

- [1] W.D. Larson, T.M. Sinclair, K. Maehashi, M. Shintani, Nowcasting unemployment insurance claims in the time of COVID-19, *International Journal of Forecasting*, 10.1016/j.ijforecast.2021.01.001, 2021, 2021
- [2] Kohei Maehashi, Mototsugu Shintani, Macroeconomic forecasting using factor models and machine learning: an application to Japan, *Journal of the Japanese and International Economies*, Volume 58, 2020, 101104, ISSN 0889-1583, <https://doi.org/10.1016/j.jjie.2020.101104>.
- [3] Seema Rawat, Aakankshu Rawat, Deepak Kumar, A. Sai Sabitha, Application of machine learning and data visualization techniques for decision support in the insurance sector, *International Journal of Information Management Data Insights*, Volume 1, Issue 2, 2021, 100012, ISSN 2667-0968, <https://doi.org/10.1016/j.jjime.2021.100012>.
- [4] C. Blier-Wong, H. Cossette, L. Lamontagne, and E. Marceau, "Machine Learning in P& C Insurance: A Review for Pricing and Reserving," *Risks*, vol. 9, no. 1, p. 4, Dec. 2020, doi: 10.3390/risks9010004.
- [5] Faheem Aslam, Ahmed Imran Hunjra, Zied Fütü, Wael Louhichi, Tahira Shams, Insurance fraud detection: Evidence from artificial intelligence and machine learning, *Research in International Business and Finance*, Volume 62, 2022, 101744, ISSN 0275-5319, <https://doi.org/10.1016/j.ribaf.2022.101744>.
- [6] Kaushik K, Bhardwaj A, Dwivedi AD, Singh R. Machine Learning-Based Regression Framework to Predict Health Insurance Premiums. *Int J Environ Res Public Health*. 2022 Jun 28;19(13):7898. doi: 10.3390/ijerph19137898. PMID: 35805557; PMCID: PMC9265373.

- [7] L. N. Tournas and D. M. Bowman, "AI Insurance: Risk Management 2.0," in *IEEE Technology and Society Magazine*, vol. 40, no. 4, pp. 52-56, Dec. 2021, doi: 10.1109/MTS.2021.3123750.
- [8] Xiao, Y., Tian, Z., Yu, J. et al. A review of object detection based on deep learning. *Multimed Tools Appl* 79, 23729–23791 (2020). <https://doi.org/10.1007/s11042-020-08976-6>
- [9] Chen X, Xiang S, Liu C-L, Pan C-H (2013) Vehicle detection in satellite images by parallel deep convolutional neural networks. In: *Asian conference on pattern recognition (ACPR)*
- [10] Q. Fan, L. Brown and J. Smith, "A closer look at Faster R-CNN for vehicle detection," *2016 IEEE Intelligent Vehicles Symposium (IV)*, Gothenburg, Sweden, 2016, pp. 124-129, doi: 10.1109/IVS.2016.7535375.
- [11] Taigman Y, Yang M, Ranzato MA, Wolf L (2014) Deepface: closing the gap to human-level performance in face verification. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp 1701–1708
- [12] Huang W, Qiao Y, Tang X (2014) Robust scene text detection with convolution neural network induced msr trees. In: *European conference on computer vision (ECCV)*, pp 497–511, 282
- [13] Zhang X, Zhang C, Shen W, Yao C, Liu W, Bai X (2016) Multi-oriented text detection with fully convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp 4159–4167
- [14] Ouyang W, Wang X (2013) Joint deep learning for pedestrian detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp 2056–2063
- [15] Zhang L, Lin L, Liang X, He K (2016) Is faster r-cnn doing well for pedestrian detection? In: *European conference on computer vision (ECCV)*, pp 443–457
- [16] Kaur, J., Singh, W. Tools, techniques, datasets and application areas for object detection in an image: a review. *Multimed Tools Appl* 81, 38297–38351 (2022). <https://doi.org/10.1007/s11042-022-13153-y>
- [17] Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. *Commun ACM* 60(6):84–90
- [18] Girshick R, Donahue J, Darrell T, Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp 580–587
- [19] Redmon J, Divvala S, Girshick R, Farhadi A (2016) You only look once: unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, pp 779–788
- [20] <https://www.v7labs.com/blog/yolo-object-detection-two-stage-detectors>
- [21] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, pp. 580-587, doi: 10.1109/CVPR.2014.81.
- [22] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks"
- [23] Yang, C., Lu, X., Lin, Z., Shechtman, E., Wang, O. and Li, H. High-resolution image inpainting using multi-scale neural patch synthesis. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 6721-6729), 2017.