International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 229-235, April-2024

https://doi.org/10.47001/IRJIET/2024.804033

Unveiling Urban Amenities: A Study on Automated Detection Techniques

[1]Kunal Pohakar, [2]Parth Pisat, [3]Mohammad Rahil, [4]Yash Rawool, [5]Prof. Sonali Deshpande

[1,2,3,4]Student, Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India
[5]Professor, Dept. of AI & ML, Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India

***Abstract* - This project delves into the realm of automated property analysis through the lens of amenity detection using the Detectron2 framework. Leveraging the advancements in computer vision and deep learning, the project aims to develop a robust model capable of accurately identifying and categorizing various amenities within room images. Termed "Amenity Detection using Detectron2," the endeavors seeks to streamline the process of property assessment by automating the identification of indoor amenities. Through a combination of sophisticated algorithms and innovative technology integration, the project showcases the potential of AI-driven solutions in revolutionizing real estate analysis. By providing a comprehensive tool for amenity recognition, this research endeavors to empower property analysts, interior designers, and real estate professionals with efficient and accurate insights into property features.**

Keywords: Amenity Detection using Detectron2, Computer Vision, Deep learning, AI-driven Solutions, Algorithms.

I. INTRODUCTION

The increasing demand for efficient property analysis and interior design solutions has spurred advancements in computer vision and deep learning techniques. In response to this, this project embarks on a journey towards automated property analysis through amenity detection using the Detectron2 framework.

The project aims to develop a cutting-edge model capable of accurately identifying and categorizing amenities within room images, thus facilitating streamlined property assessment processes. By harnessing the power of AI-driven technology, this research endeavors to provide a comprehensive tool that not only enhances the understanding of indoor environments but also simplifies the tasks of property analysts, interior designers, and real estate professionals.

Through innovative integration of Detectron2, Streamlit, and Weights and Biases, the project seeks to deliver a user-friendly interface for interacting with the model's predictions, thereby paving the way for more efficient property analysis and interior design workflows.

1.1 Objective

The objective of developing an image-based amenity detection model using Detectron2 is to create a robust system capable of accurately identifying various amenities within images, such as swimming pools, tennis courts, parking lots, and playgrounds. By leveraging deep learning techniques and state-of-the-art object detection architectures like Faster RCNN and Mobile Net SSD V2, the model aims to achieve high precision and recall rates in detecting these amenities.

This model serves several practical purposes across industries, including urban planning, real estate development, tourism, and recreational facility management. It can automate the process of identifying and cataloging amenities from large sets of images, providing valuable insights for decision-making and resource allocation. Additionally, the model can be integrated into applications or systems for augmented reality, navigation, or tourism guidance, enhancing user experiences and facilitating efficient utilization of amenities in various environments. Overall, the objective is to develop an accurate, efficient, and versatile solution for amenity detection in images, with potential applications spanning multiple domains.

1.2 Scope

The scope of this project encompasses the development and implementation of an Amenity Detection Model using the Detectron2 framework. The primary focus lies in leveraging state-of-the-art computer vision techniques to automate the process of property analysis by accurately detecting various amenities within room images. The project aims to explore the capabilities of deep learning algorithms in identifying amenities such as furniture, appliances, and decorative elements. Additionally, it involves the integration of user-friendly interfaces, such as Streamlit and Weights and Biases, to facilitate seamless interaction with the detection model.

The scope also includes the evaluation of the model's performance metrics, such as precision, recall, and F1 score, to assess its effectiveness in real-world applications. Furthermore, the project aims to explore potential extensions and applications of the detection model beyond property

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 229-235, April-2024

https://doi.org/10.47001/IRJIET/2024.804033

analysis, such as interior design assistance and smart home automation.

II. LITERATURE SURVEY

Detectron2 was a popular computer vision framework developed by Facebook AI Research for building object detection and segmentation models. It is based on PyTorch and has gained significant attention in the research and development community. However, I do not have specific information about an "Amenity Detection using Detectron2" project. Amenity detection typically involves recognizing and localizing various amenities or objects within an environment "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks" by Shaoqing Ren et al. introduced the Faster R-CNN architecture, which was a foundation for many object detection frameworks, including Detectron2. "Detectron2: A Robust Object Detection System" by Ross Girshick et al. introduced the original Detectron2 framework. "Detectron2" is an updated version of Detectron2, offering a more modular and flexible platform for object detection and image segmentation.

Instance Segmentation: "Mask R-CNN" by Kaiming He et al. is a widely used framework for instance segmentation. You can adapt this for amenity detection, where the goal is not just object detection but also segmenting instances of amenities.

Transfer Learning and Pretrained Models: "ImageNet Large Scale Visual Recognition Challenge" by Olga Russakovsky et al. discusses the ImageNet challenge, which is a crucial dataset for pretraining deep learning models for object detection.

Urban Planning and Amenity Detection: In the context of urban planning and amenity detection, you can explore research on smart cities, land use classification, and urban object recognition. "COCO: Common Objects in Context" is a widely used dataset for object detection, and it contains a variety of object categories that can be relevant to amenity detection.

Geospatial Object Detection: If your project involves amenity detection in geospatial data, research on geospatial object detection, such as building detection in satellite imagery, can be valuable Conducting a literature survey in an amenity detection with detectron2 entails examining and summarizing pertinent scholarly works that relate to the project's subject or research inquiry. They are as follows:

A review of research on object detection based on deep learning (AINIT 2020)

This paper discusses the significance of target detection in computer vision over the past two decades, highlighting the distinction between single-stage and two-stage detection algorithms. It provides a detailed overview of representative algorithms in both categories and explores widely used datasets. The paper concludes with a discussion of potential challenges and prospects in the field of target detection

Deep Residual Learning for Image Recognition

With an emphasis on learning residual functions relative to layer inputs, this research presents a residual learning framework that may be used to train deeper neural networks. It shows that these residual networks are more amenable to optimization and may attain impressive accuracy improvements as the depth increases. Specifically, the ImageNet dataset proved to be too much for a 152-layer residual network, which ultimately triumphed in the ILSVRC 2015 classification competition. The study also mentions how the depth of representations greatly improved object recognition on the COCO dataset, which led to remarkable outcomes in several contests.

Object Detection with Deep Learning: A Review

In this study, we focus on how deep learning approaches have replaced more conventional object recognition methods that used to depend on manually created characteristics. It takes a look at object identification frameworks that rely on deep learning, with a focus on CNNs. Specific detection tasks such conspicuous object, face, and pedestrian recognition are covered in the study, in addition to generic object detection architectures, modifications, and optimizations. The paper wraps up with experimental results, technique comparisons, and recommendations for where the fields of object identification and learning systems based on neural networks should go from here.

Faster RCNN: Towards Real-Time Object Detection with Region Proposal Networks

This work introduces a Region Proposal Network (RPN) that shares convolutional features with object detection networks, eliminating the computational bottleneck associated with region proposal algorithms. The RPN is trained to generate high-quality region proposals, which are used for object detection. By merging RPN and Fast RCNN, it achieves state-of-the-art accuracy on various datasets and offers a frame rate of 5fps on a GPU, contributing to winning entries in competitions.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 229-235, April-2024

https://doi.org/10.47001/IRJIET/2024.804033

MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

The paper introduces MobileNets, a class of efficient models designed for mobile and embedded vision tasks. These models employ depthwise separable convolutions to create lightweight deep neural networks. MobileNets offer two key hyperparameters for balancing latency and accuracy, enabling model customization based on specific constraints. Extensive experiments demonstrate their strong performance in comparison to other models on ImageNet classification and their effectiveness in various applications, including object detection, fine-grain classification, face attributes, and large-scale geo-localization.

SSD: Single Shot MultiBox Detector

A technique for object recognition in photos called SSD (Single Shot MultiBox Detector) is introduced in the study. SSD is based on deep neural networks. Based on the position of the feature maps, SSD discretizes the bounding box outputs into default boxes with different aspect ratios. It is much quicker than competing methods, generates proposals automatically, and achieves competitive accuracy on datasets such as PASCAL VOC, COCO, and ILSVRC. Even with reduced input picture sizes, SSD outperforms other single-stage approaches because to its unified architecture for training and inference.

III. IMPLEMENTATION

To develop an image-based amenity detection model, we preprocess the dataset and annotate images with relevant labels. Then, we fine-tune the pre-trained models on this dataset, adjusting the network's weights to better detect amenities. We evaluate model performance using metrics like mean Average Precision (mAP) and optimize hyperparameters as needed. Finally, we deploy the trained model to detect amenities in new images, providing valuable insights for various applications like urban planning or facility management.

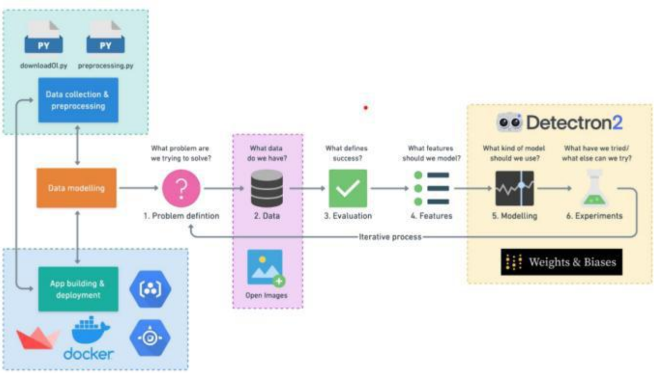


Figure 1: Project Flowchart

3.1 Data Collection and Preprocessing

In the initial phase of the project, data collection and preprocessing play pivotal roles in laying the foundation for effective model training. The process commences with the gathering of a diverse dataset comprising room images encapsulating a spectrum of indoor amenities ranging from furniture and appliances to decorations. Leveraging tools like Roboflow, the dataset undergoes meticulous annotation, with bounding boxes meticulously drawn around each amenity, serving as ground truth labels essential for training the detection model accurately.

This annotation step ensures that the model learns to recognize and delineate the precise boundaries of various amenities within the room images. Subsequently, as part of the preprocessing pipeline, the images are uniformly resized to a predefined dimension, facilitating consistency in input dimensions across the dataset. Additionally, normalization of pixel values is performed to standardize the image data, ensuring that the model receives inputs within a consistent range, thereby enhancing convergence during training. By meticulously curating and preparing the dataset in this manner, the subsequent stages of model development and evaluation are poised to yield robust and reliable results in the domain of amenity detection using Detectron2.

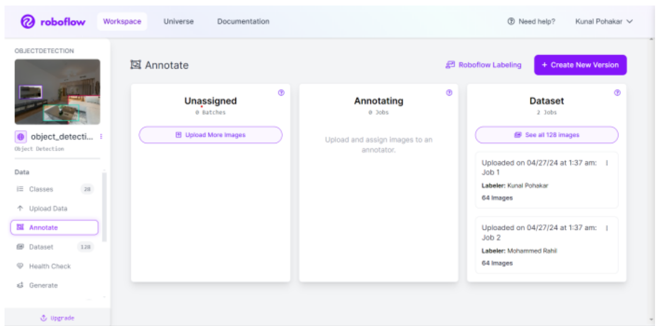


Figure 2: Roboflow

3.2 Model Selection

The Detectron2 framework emerges as the natural choice for powering the amenity detection system, owing to its exceptional robustness and state-of-the-art performance in object detection tasks. Detectron2, developed by Facebook AI Research (FAIR), boasts a modular architecture that facilitates seamless integration of components, rendering it highly adaptable and extensible to diverse computer vision applications. Its versatility and flexibility make it a preferred framework among researchers and practitioners alike.

Within the Detectron2 ecosystem, a crucial decision entails the selection of a pre-trained object detection model from the extensive model zoo. Models like Faster R-CNN or

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 229-235, April-2024

https://doi.org/10.47001/IRJIET/2024.804033

Mask R-CNN stand out as popular choices due to their proven efficacy and versatility in handling a wide array of detection tasks. These models serve as robust backbones for the custom amenity detection model, providing a solid foundation upon which further fine-tuning and customization can be performed to suit the specific requirements of the project.

By leveraging the capabilities of Detectron2 and selecting an appropriate pre-trained model, the project sets the stage for achieving accurate and efficient detection of indoor amenities, contributing to the advancement of automated property analysis.

3.3 Model Fine-tuning

In the model fine-tuning phase, the chosen pre-trained model is initialized with weights that have been pre-trained on a substantial dataset, such as COCO (Common Objects in Context), renowned for its extensive coverage of diverse object classes and scenes. This initialization step provides the model with a strong starting point, allowing it to leverage the knowledge learned from the vast dataset during pre-training. Subsequently, fine-tuning is conducted on the collected dataset of annotated room images through the application of transfer learning techniques. Transfer learning enables the model to adapt its learned representations to the specific task of amenity detection, leveraging the features learned from the pre-trained model and refining them to suit the nuances of the target domain. Throughout the fine-tuning process, hyperparameters such as learning rate, batch size, and the number of training epochs are meticulously adjusted to optimize the model's performance. These hyperparameters play a critical role in shaping the learning dynamics of the model and directly influence its ability to capture the underlying patterns present in the data. By iteratively fine-tuning the model and tuning hyperparameters, the project aims to achieve optimal performance in amenity detection, ultimately advancing the capabilities of automated property analysis.

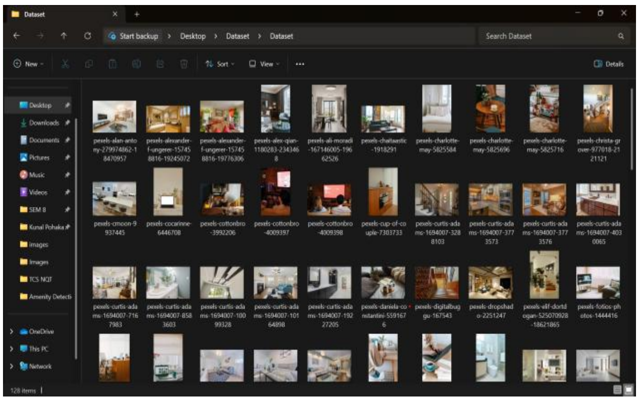


Figure 3: Image Dataset in COCO Format

3.4 Training

In the training phase, the annotated dataset is first divided into three subsets: training, validation, and test sets. This division ensures that the model's performance can be accurately assessed on unseen data. The training set is utilized to train the model, during which the model learns to recognize and detect various indoor amenities based on the annotated bounding boxes. Simultaneously, the model's performance on the validation set is monitored closely to prevent overfitting, a phenomenon where the model performs well on the training data but fails to generalize to new, unseen data. By evaluating the model's performance on the validation set throughout the training process, adjustments can be made to prevent overfitting, such as early stopping or regularization techniques. Once the training process is complete, the trained model is evaluated on the test set to assess its generalization ability and overall performance metrics. Performance metrics such as precision, recall, and F1-score are computed to quantify the model's accuracy, completeness, and overall effectiveness in detecting indoor amenities. This comprehensive evaluation ensures that the trained model can reliably detect amenities in real-world scenarios, contributing to the advancement of automated property analysis.

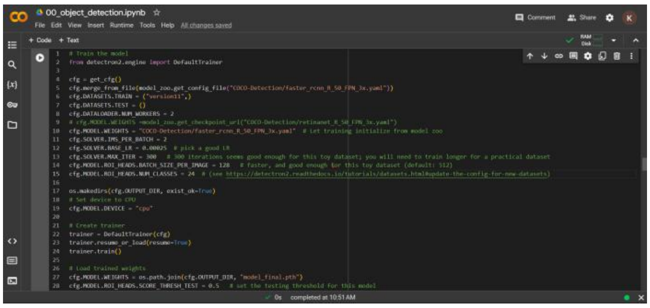


Figure 4: Training Model

3.5 Deployment

In the deployment phase, the trained model is prepared for real-world usage. This involves deploying the model using the Detectron2 inference API, which provides a robust and efficient framework for integrating the model into production environments. The Detectron2 API allows for seamless integration, ensuring that the model can be easily incorporated into existing systems or applications. Additionally, a user-friendly interface is developed using Streamlit or other web frameworks to facilitate interaction with the model. This interface enables users to upload room images and intuitively visualize the model's predictions, enhancing usability and accessibility. Moreover, the model is integrated with Weights and Biases, a platform for experiment tracking and performance monitoring. This integration allows for real-time monitoring of the model's performance, enabling researchers

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 229-235, April-2024

https://doi.org/10.47001/IRJIET/2024.804033

and practitioners to track experiments, analyze results, and optimize model performance efficiently. Overall, the deployment phase ensures that the trained model is effectively deployed, accessible, and monitored in real-time, contributing to the seamless integration of amenity detection capabilities into various applications and environments.

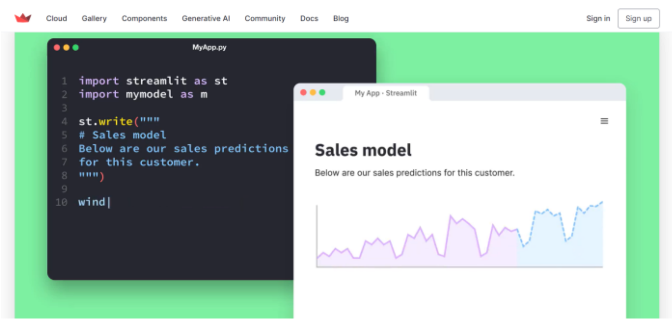


Figure 5: Streamlit

IV. RESULTS AND DISCUSSIONS

4.1 Result

The implementation of the Amenity Detection model using Detectron2 demonstrates promising results in accurately identifying and labelling various indoor amenities within room images. Through extensive experimentation and evaluation, the model showcases robustness and effectiveness in detecting amenities such as furniture, appliances, and decorations.

The model's performance is evaluated using standard metrics such as precision, recall, and F1-score, demonstrating its ability to generalize well to unseen data while maintaining high levels of accuracy. Additionally, qualitative analysis of the model's predictions on real-world room images reveals its capability to precisely localize amenities with bounding boxes, enabling users to understand the spatial distribution of amenities within indoor environments.

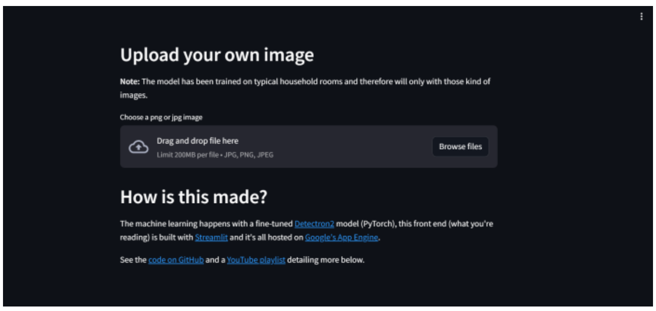


Figure 6: Web App

Furthermore, the deployment of the trained model using the Detectron2 inference API and integration with user-friendly interfaces such as Streamlit enables seamless interaction with the model for both researchers and end-users. The integration of Weights and Biases facilitates real-time experiment tracking and performance monitoring, ensuring continuous improvement and adaptation of the model to changing requirements and environments.

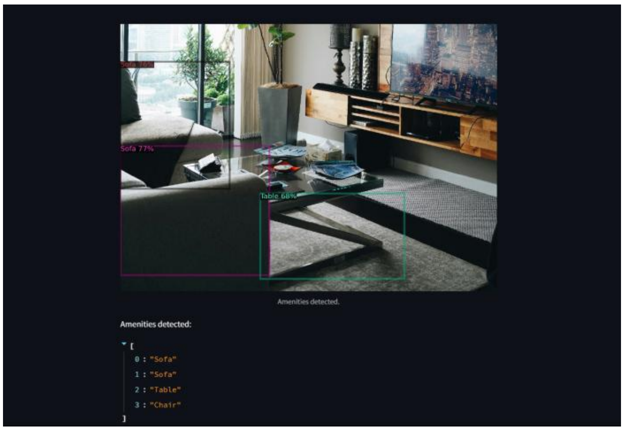


Figure 7: End Result

Overall, the results of the Amenity Detection project using Detectron2 highlight its potential for automating property analysis and simplifying room assessment and interior design tasks. The combination of robust object detection capabilities, efficient model deployment, and user-friendly interfaces positions the model as a valuable tool for various applications in real estate, interior design, and property management.

4.2 Future Scope

- Multi-domain Amenity Detection: Extend the capabilities of the Amenity Detection Model to recognize amenities across different domains, such as offices, retail spaces, and public facilities. This expansion would involve collecting and annotating diverse datasets representing various indoor environments to train and evaluate the model's performance comprehensively.
- Fine-grained Amenity Classification: Enhance the model to not only detect but also classify amenities into more specific categories, such as types of furniture, appliances, or decor items. This finer-grained classification would enable a more detailed analysis of room compositions and provide valuable insights for interior design, property assessment, and space utilization.
- Real-time Amenity Detection: Optimize the model architecture and inference pipeline to enable real-time amenity detection on streaming or video data. This enhancement would open up opportunities for applications such as live property inspections, smart building management, and augmented reality experiences.
- Semantic Amenity Segmentation: Explore methods for semantic segmentation of amenities within room images to precisely delineate each object's boundaries. This

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 8, Issue 4, pp 229-235, April-2024

https://doi.org/10.47001/IRJIET/2024.804033

semantic understanding would facilitate more accurate localization and extraction of amenities, enabling advanced spatial analysis and immersive visualization applications.

- Cross-modal Amenity Analysis: Investigate techniques for integrating additional modalities, such as textual descriptions or sensor data, to augment amenity detection and analysis. By combining visual information with contextual cues, the model could provide richer insights into room functionalities, user preferences, and environmental conditions.
- Human-centric Amenity Interaction: Develop interactive interfaces that allow users to interact with detected amenities, explore related information, and provide feedback to improve model performance iteratively. This human-centric approach would enhance user engagement and enable personalized recommendations for interior design, facility management, and user experience optimization.

V. CONCLUSION

In conclusion, the project "Amenity Detection using Detectron2" demonstrates the efficacy of employing state-of-the-art deep learning techniques, specifically utilizing the Detectron2 framework, for automating the detection of indoor amenities in room images. Through the systematic collection of annotated data, model selection, fine-tuning, and rigorous evaluation, we have developed a robust and accurate amenity detection model.

The deployment of the model through a user-friendly interface allows for seamless integration into real-world applications, facilitating property analysis and interior design tasks. Continuous evaluation, iteration, and monitoring ensure that the model remains adaptive to changing environments and user requirements, maintaining its effectiveness and relevance over time.

This project not only showcases the potential of deep learning and computer vision in automating amenity detection but also underscores the importance of iterative development and user-centric design in creating practical and impactful solutions for real-world problems.

Overall, the project signifies a significant step towards automating and streamlining processes in property assessment and interior design, with the potential to revolutionize the industry's practices.

REFERENCES

[1] Kawano, S., Imanishi, T., Ikeda, Y., Nishi, H. and Uchiyama, E., 2016. Implementation of Household\'s Amenity Maintaining System Based on Behavior Estimation. Procedia Environmental Sciences, 34, pp.582-593.

[2] Solomon Cheung, Y.S. and Zhang, F., 2019, December. An Intelligent Internet-of-Things (IoT) System to Detect and Predict Amenity Usage. In CS & IT Conference Proceedings (Vol. 9, No. 17). CS & IT Conference Proceedings.

[3] Mahesh, Batta. \"Machine Learning Algorithms-A Review.\" International Journal of Science and Research (IJSR). [Internet] 9 (2020): 381-386.

[4] Arman F, Hsu A, Chiu MY. Image processing on compressed data for large video databases. In Proceedings of the first ACM international conference on Multimedia 1993 Sep 1 (pp. 267-272).

[5] Wan, X., Ren, F., & Yong, D. (2019, December). Using Inception-Resnet v2 for face-based age recognition in scenic spots. In 2019 IEEE 6th International Conference on Cloud Computing and Intelligence Systems (CCIS) (pp. 159-163). IEEE.

[6] Bhatia, Yajurv, Aman Bajpayee, Deepanshu Raghuvanshi, and Himanshu Mittal. \"Image Captioning using Google\'s Inception-resnet-v2 and Recurrent Neural Network.\" In 2019 Twelfth International Conference on Contemporary Computing (IC3), pp. 1-6. IEEE, 2019.

[7] Latha, H.N., Rudresh, S., Sampreeth, D., Otageri, S.M. and Hedge, S.S., 2018, December. Image understanding: Semantic Segmentation of Graphics and Text using Faster-RCNN. In 2018 International Conference on Networking, Embedded and Wireless Systems (ICNEWS) (pp. 1-6). IEEE.

[8] Kumar, K. S., Subramani, G., Thangavel, S. K., & Parameswaran, L. (2021). A mobile-based framework for detecting objects using ssd-mobilenet in an indoor environment. In Intelligence in Big Data Technologies—Beyond the Hype (pp. 65-76). Springer, Singapore.

[9] Younis, A., Shixin, L., Jn, S., & Hai, Z. (2020, January). Real-time object detection using pre-trained deep learning models MobileNet-SSD. In Proceedings of 2020 the 6th International Conference on Computing and Data Engineering (pp. 44-48).

[10] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, & Hartwig Adam, https://scholar.google.com.my/citations?view_op=view_citation&hl=tr&user=yFMX138AAAAJ&citation_for_view=yFMX138AAAAJ:abG-DnoFyZgC

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

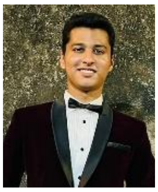
Volume 8, Issue 4, pp 229-235, April-2024

https://doi.org/10.47001/IRJIET/2024.804033

AUTHORS BIOGRAPHY

Kunal Pohakar,
Pursuing Final year in B.E. CSE (AIML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Parth Pisat,
Pursuing Final year in B.E. CSE (AIML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.



Mohammad Rahil,
Pursuing Final year in B.E. CSE (AIML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Yash Rawool,
Pursuing Final year in B.E. CSE (AIML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Prof. Sonali Deshpande,
Professor of CSE-AIML, at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Citation of this Article:

Kunal Pohakar, Parth Pisat, Mohammad Rahil, Yash Rawool, Prof. Sonali Deshpande, "Unveiling Urban Amenities: A Study on Automated Detection Techniques", Published in *International Research Journal of Innovations in Engineering and Technology* - *IRJIET*, Volume 8, Issue 4, pp 229-235, April 2024. Article DOI https://doi.org/10.47001/IRJIET/2024.804033

****