



Survey on Object Detection Using Synthetic Data

Janani J¹, Nanditha R², Yogalakshmi R³

^{1,2,3}Department of Information Technology, Sri Manakula Vinayagar Engineering College, Puducherry, India, 2022

ABSTRACT:

In the beyond decade, the wide variety of cars on the street has been growing rapidly. The great use of cars has multiplied the possibility of visitors guidelines violation, brought about surprising accidents, and caused visitors crimes. Object detection is a computer vision technique and prescient method for finding times of items in snap shots or videos. Object detection algorithms usually leverage gadget studying or deep studying to provide significant results. Object detection, a computer vision, solves the hassle of figuring out and localizing training in an image. Here we use Synthetic facts for object detection, As the name suggests, the data is artificially created instead of being generated via way of means of real events. It may be generated to satisfy unique desires or situations that aren't present in existing(real) facts. We use the equipment from Unity to generate and analyses artificial datasets with an illustrative instance of item detection.

I. INTRODUCTION

Artificial intelligence leverages computer systems and machines to imitate the problem-fixing and decision-making competencies of the human mind. At its most effective form, artificial intelligence is a field, which combines computer science and strong datasets, to allow problem solving. It additionally encompasses sub-fields of gadget mastering and deep mastering, which can be regularly referred to together with artificial intelligence.

These disciplines are AI algorithms which are seeking to create professional structures which make predictions or classifications primarily based totally on entered data.

The drastic boom in vehicular traffic at the roadways stimulates a large demand for traffic monitoring with the era for tracking and management. In this scenario, guide monitoring of motors going for walks speedy on the street is nearly now no longer feasible. There might be wastage of manpower and time. Even if it's operated manually, with a purpose to mirror big problems and big errors. There are already answers for monitoring the motors and wide variety of number plates regarding the use of gadget gaining knowledge of algorithms like machine learning. This set of rules extracts the unique information from a picture and offers it as an enter to the degree of automobile recognition.

Based on Artificial Intelligence, numerous experiments and reviews on datasets to assess our machine in hit upon, locate, and understand cars. The experimental consequences display that our proposed approach outperforms the latest industrial and educational structures in maximum difficult scenarios. Moreover, we progressed accuracy, so that you can understand one of a kind font faces and dynamic lengths of the cars.

The mission is able to detect numerous automobile positions in an actual international image, and is capable of hit upon all sorts of automobile in addition to with assorted foreground and heritage colours, from numerous angles and distances below on of a kind climate conditions. The experimental consequences display that our proposed approach outperforms the latest industrial and educational structures in maximum difficult scenarios. In this scenario, guide monitoring or manual monitoring of cars rapid on the street is almost now no longer feasible. This open supply package deal extends the Unity Editor and engine additives to generate flawlessly annotated examples for numerous computer vision tasks.

II. LITERATURE SURVEY

[1] In latest years, it offers an evaluation approximately programs and strategies wherein template matching is used the Unity Perception package deal which ambitions to simplify and boost up the method of producing artificial datasets for computer vision and prescient responsibilities with the aid of using supplying an easy to use and fairly customizable toolset. This open supply package deal extends the Unity Editor and engine additives to generate flawlessly annotated examples for numerous not unusual place computer vision and prescient responsibilities. A labeller makes use of this item statistics to generate floor reality statistics which includes 2D bounding containers or semantic segmentation masks. The produced floor reality is then captured at the side of related metrics in JSON files. Additionally, it gives an extensible Randomization framework that shall help the person to quickly assemble and configure randomised simulation parameters that allows you to introduce variant into the generated datasets.

[2] The implementation of a selected image/scene with one pass. It can simulate cameras, LIDAR and different information sources. Simulate the hardware your actual international use case utilises the primary engine temperature, gas quantity, and so on. It may even strive out if LIDAR information could enhance overall performance to your case earlier than making an investment in any equipment ,as a minimum to a few extent. So far it approximately matters which are simpler to perform with fabricated information however how approximately risky or not possible to accumulate information. A labeller

makes use of this item statistics to generate floor reality information along with 2D bounding packing containers or semantic segmentation masks. The produced floor reality is then captured along side related metrics in JSON files. Let's say you need to discover vehicle crashes on public roads. There is best a confined quantity of material you could gather, and it's far difficult to get more. it'd take per week or a month to generate sensible sufficient crashes in a 3D engine, however after that unlimited training data.

[3] In this approach further, using Unity Perception, PeopleSansPeople, and a third device known as MakeHuman to generate labelled training statistics of synthetic humans for machine learning. The educational is geared in the direction of the ML practitioner and assumes no revel in with 3D artwork or game development. Note that it assuming a Windows based system for this tutorial, however all of the equipment need to be like compatible with Linux.

[4] The Unity Perception bundle allows a brand new workflow in Unity for producing artificial datasets and helps each Universal and High Definition Render Pipelines. In this primary release, it gives equipment for dataset seize and includes four number one features: object labelling, labellers, photograph seize, and custom metrics. The bundle gives a easy interface to enter object-label associations, which can be picked up robotically and fed to the labellers. A labeller makes use of this item records to generate floor fact statistics inclusive of 2D bounding containers or semantic segmentation masks. The produced floor fact is then captured in conjunction with related metrics in JSON files.

[5] The factor in general used were speaking approximately matters which can be less difficult to perform with fabricated statistics however how approximately risky or not possible to gather statistics. A labeller makes use of this item records to generate floor fact statistics inclusive of 2D bounding containers or semantic segmentation masks. The produced floor fact is then captured in conjunction with related metrics in JSON files. Additionally, it gives an extensible Randomization framework that shall we the person speedy assemble and configure randomised simulation parameters with a view to introduce version into the generated datasets. It may assist to lessen decision or best stages to healthy the best of the cameras and so on, relying to your use-case. Real blockers stand up whilst predictions consciousness on small information inclusive of facial recognition. approximately risky or not possible to gather statistics. A labeller makes use of this item records to generate floor fact statistics inclusive of 2D bounding containers or semantic segmentation masks. Simulate the hardware your actual global use-case utilises the principle engine temperature, gas amount, and so on.

[6] This method is used to summarise the unrestricted-area spontaneous presentation, and it's far analysed the usage of each goal and subjective measurements. It has been verified that the proposed techniques are a success in taking pictures the essence of spontaneous speech. It is made of wheels which could rotate of their very own axis and reply in another way to exclusive manipulate instructions. The clever shared manipulate machine is made of three essential manipulate modules in addition to a selection maker that chooses which action must be carried out. The second essential one is a discrete occasion controller, which makes use of the records accumulated with the aid of using the sensor to guide the robot away far from potential dangers as it walks.

[7] The algorithms from supervised, unsupervised and regression getting to know have been selected to decorate accuracy and decrease the quantity of functions on the time of training. Ensemble getting to know calls for excessive computation power, as more than one fashions want to be built. Moreover, deciding on a big quantity of functions influences the detection performance of an set of rules in IoT environments. Therefore, in the proposed technique, we present an efficient feature choice mechanism. Moreover, to lessen the computational overhead, we applied a one-elegance classifier to acquire a best detection charge at the side of low false-alarm rates.

[8] Anomaly primarily based totally detection processes attempt to stumble on botnets primarily based totally on some of site visitors anomalies consisting of excessive community latency, excessive volumes of site visitors, site visitors on uncommon ports and uncommon machine conduct that might display life of bots withinside the community. Binkley et al. proposed a powerful set of rules that mixes TCP-primarily based totally anomaly detection with IRC tokenization and IRC message facts to create a machine that may definitely stumble on patron vehicles. However, this method may be without difficulty defeated with the aid of using absolutely the usage of a trivial cipher to encode the IRC commands. Ensemble getting to know calls for excessive computation power, as more than one models want to be built. Moreover, deciding on a big quantity of functions

[9] Detection the usage of community visitors and their consciousness changed into on car detection using the device's usage and temperature, reminiscence consumption, and numerous running tasks. The highest F-score for their experiment using a 1-s time interval was .We preprocessed the dataset with a 1-s time interval to test botnet network traffic and regular traffic for our experiment. The maximum hard a part of designing a behavior primarily based totally detection version is the function selection. It isn't always sincere to recognise which capabilities ought to be used and a way to extract the pattern. Meanwhile, the motive of the second one module changed into to locate the car to server to take down the botnet through blockading the supply IP from coming into the community. In this module, the network traffic changed into extracted and aggregated

[10] After deciding on capabilities, the information had been aggregated to be the enter withinside the following process, which for the primary module changed into clustering, and for the second one changed into class. For this study, we used a k-method set of rules. The clustering changed into performed via Weka, a system studying device and library, and the consequences proceeded to the assessment process. The second module is the class module to locate the Command and Control server via the supply . Ensemble studying calls for excessive computation power, as more than one modules need to be built. Moreover, deciding on a massive wide variety of capabilities influences the detection performance of a set of rules in IoT environment. Volumes of visitors, traffic on uncommon ports and uncommon machine behaviour that might display lifestyles of bots withinside the community. Binkley et al. proposed a powerful set of rules that mixes TCP-primarily based totally anomaly detection with IRC tokenization and IRC message facts to create a machine which can definitely detected. The second module is the class module to locate the Command and Control server via the supply . Ensemble studying calls for excessive computation power, as more than one modules need to be built. The dataset changed into break up alongside a 70-30 ratio,

wherein it changed into the education set and it changed into the checking out set. The assessment and prediction had been run at the checking out dataset simplest stays the cost to the reference. it gives an extensible Randomization framework that shall we the person quick assemble and configure randomised simulation parameters which will introduce version into the generated datasets.

III. SYNTHETIC DATA

The important undertaking of fabricated datasets is getting it to shut sufficient similarity with the real international use-case; especially video. It would possibly assist to lessen decision or first-class ranges to in shape the first-class of the cameras and so on, relying to your use-case. Real blockers get up whilst predictions awareness on small info along with facial recognition.

Also, like stated above, rendering lengthy photo-realistic videos may be extraordinarily useful resource but heavy. You can get round that with the aid of using handiest drawing the segmentation maps, in particular in case you are changing the raw photos to segmentation maps in any case.

And finally, you may simply be lacking the sources to investigate and construct a records production pipeline withinside the first place, even though it'd store time within side the lengthy run.

First matters first, we want a Unity utility that renders and saves photographs. For example, I'll render a rotating vehicle in the front of a brick wall. In a real use-case, you'll have maximum probable need to store vehicle display screen coordinates and such from the scene, however I'll hold this case simple.

The Unity utility in action, taking screenshots of a vehicle in numerous angles and lights conditions. Nothing too fancy taking place here, the entire utility is much less than one hundred fifty traces of code. You have a scene you need to apply for and contact `ScreenCapture.CaptureScreenshot()` . with Unity, you may use the executable on Valohai.

It hosted the binary on S3, however you may additionally construct the entire Unity utility at the platform, however we will pass that for brevity. Next, we will use the standalone Unity utility to store a few photographs and redirect the ones photographs to some other execution which masses and makes use of the photographs.

The very last piece is to run the mock Python education. In our example, we offer the generated dataset, it simply prints the wide variety of photographs acquired and outputs semi-random education values, however the script ought to do any Python operations along with running Keras, TensorFlow or PyTorch.

That is a high-degree evaluation of a way to run Unity packages and tie them to a records technological know-how pipeline. You ought to effectively run more than one photograph turbines or education in parallel with a click on of some buttons.

IV. THE UNITY PERCEPTION PACKAGE

The Unity Perception package deal extends the Unity Editor with equipment for producing artificial datasets that encompass floor fact annotations. In addition, the package deal helps area randomization for introducing range into the generated datasets. Out of the box, the package deal helps computer vision and prescient responsibilities inclusive of 2D/three-D item detection, semantic segmentation, example segmentation, and key points (nodes and edges connected to 3D objects, beneficial for responsibilities along with human-pose estimation).

It depicts pattern outputs for those tasks. In addition, customers can increase the package deal the usage of C# to guide new floor reality labelling techniques and randomization techniques. Additionally, at the side of the functionality to generate datasets locally, the Perception package deal has easy-to-use integrated guide for running dataset technology jobs withinside the cloud the usage of Unity Simulation, making it feasible to generate hundreds of thousands of annotated photographs particularly quickly, and without the want for effective nearby computing resources.

V. RELATED WORKS

a) Rule based Approach

The computer vision community has invested remarkable sources to create datasets along with PASCAL VOC [, NYU-Depth V2, MS COCO, and SUN in studies on complicated obligations along with semantic segmentation of indoor scenes, they can't cover all of the situations researchers are inquisitive about or offer a course for others to create new datasets. Some researchers have diagnosed the capability for artificial facts and feature executed remarkable consequences in unique obligations and domain names. Examples of this encompass item detection of groceries, controlling robot palms to transport blocks , and fine-grained manipulation of a Rubik's cube; however, those simulation environments are regularly now no longer publicly available, which makes the studies tough to breed and not possible to extend. Others have diagnosed this undertaking and created simulation environments along with SYNTHIA, CARLA, AI2-THOR, Habitat, and Gibson.

These simulators growth the range of examples researchers can use to educate their computer vision techniques and facilitate reproducibility; however, they do now no longer offer a general- cause platform as they couple the simulator with unique domain names and obligations. NVIDIA Isaac Sim takes those efforts and steps up to add and offers a platform constructed to allow a huge variety of robotics simulations, in place of a particular area or task.

Blender and NVISII proportion comparable dreams of seasoned viding an API to generate training examples for some computer vision technique obligations, however don't focus on physics or simulation.

The Unity Perception package deal builds on pinnacle of the Unity Editor to be a constructing block of such robotics computer vision projects. The first assignment to apply this package deal became Synthes, that is and end-to-end solution for detecting grocery items the use of artificial datasets and studying version overall performance beneath diverse mixtures of actual and artificial data.

b) Model based approach

Model primarily based totally structures are extra effective due to the fact they depend upon a deep knowledge of the device. Model-primarily based totally reasoning structures are inference techniques in professional structures that are primarily based totally on a version of the bodily world.

The knowledge or skilled individual who understands the bodily, mechanical, facts flow, or different information of the complicated device develops the version-primarily based totally device. This method takes extra time to construct the version. This version is quantitative or qualitative and it's far primarily based totally on mathematical equations or cause/impact models. In this method, an appropriate version is advanced after which it's far delicate till it produces the de- sired results.

In a version- primarily based totally device, human information is encoded via way of means of a hand coded representation. In the hand coded version, best qualitative variables are used to outline bodily behaviour of the device. This version is physics-primarily based totally device or AI primarily based totally device including fuzzy common sense, neural network, etc. T. M Beerse carried out fuzzy common sense to compute a fault speculation within side the version primarily based totally diagnostic device.

c) Data Driven approach

Data-driven approach refers to studying structures which might be famous in cognitive modelling i.e. modelling hassle fixing technique. The device of statistics pushed version is constructed at once from automatically found device running statistics. Learning structures are primarily based totally on the subsequent assumption: statistical traits of the statistics are stable, however for any fault occasion happens withinside the device. It does now no longer constitute physics of the modelled technique however builds the connection among enter and output variables. This version is beneficial for fixing a hassle whilst there's a large amount of statistics is to be had approximately the hassle.

d) Ground truth generation

As a child node of P. The node P is labelled with check standards T and walls $P_1, P_2 \dots P_m$ are in addition partitioned recursively in a comparable way. A partition wherein all of the information be- longs to equal elegance isn't partitioned in addition, and the leaf node similar to it's miles labelled with the elegance.

However, an imperfect, smaller choice tree frequently achieves extra accuracy than a super tree. The cause is this best choice tree is incredibly sensitive to statistical irregularities of the schooling dataset. Hence pruning (delete nodes iteratively), the second one phase, is per- shaped after the development of a tree to attain an correct tree than a super tree. A choice tree T1 is stated to over suit schooling records if any other tree T2 exists that has a bigger blunders than T1 whilst confirmed at the schooling records, and smaller blunders than T1 whilst confirmed at the whole dataset.

The set of rules that is primarily based totally at the KNN class technique is pro- posed. The interleaving of SVM and KNN, a brand new classifier is proposed to enhance the accuracy of the classifier. This new classifier plays properly on balanced records. It is touchy to unbalanced education samples and the overall performance of the classifier declines significantly.

The Unity Perception bundle, which gives a number of handy and customizable equipment that assist accelerate and simplify the manner of producing labelled artificial datasets for laptop imaginative and prescient problems. If there may be no authentic records in a dataset, then it's far a very artificial records set. If there may be any authentic records in a newly generated records set, it's far in part a artificial records set. Only sensitive data is regenerated the use of artificial records era strategies in a in part artificial records collection.

Majorly used 3D assets and Unity belief bundle to acquire the records and create a scene that is then made to stimulate. After simulation, artificial dataset is created wherein we can create a dataset for insights and dataset data the use of the amassed records. An Autonomous System is gift wherein we are able to locate the actual and simulated records. In this machine the manner of coordination, control, making plans and belief takes place.

The outside records like HD maps and V2X are amassed from eventualities and despatched to the Autonomous machine. As we recognize Unity computer vision and prescient datasets are primarily based totally on artificial records wherein data is generated artificially primarily based totally on actual time eventualities. After the data is amassed from eventualities they go through diverse techniques within the Simulation environment.

e) Randomization Tools

The Unity Perception bundle affords a randomization framework that simplifies introducing variant into artificial environments, main to numerous data. An entity referred to as the Scenario controls and coordinates all randomizations withinside the scene. This includes triggering a hard and fast of Randomizers in a predetermined order. Each Scenario's execution is referred to as an Iteration and every Iteration can run for a consumer described wide variety of frames. Users can configure Randomizers to behave at numerous timestamps in the course of every Iteration, consisting of on the begin and give up or consistent with every frame.

The Randomizers disclose the surroundings parameters for randomization and make use of samplers to select out random values for those parameters. The mixture of the Scenario and its Randomizers permit the consumer to outline problematic and deterministic schedules for dominations to arise in the course of a simulation.

f) Package generation

To aid each use cases, we built dataset IO modules that permit customers to parse, load, and rework datasets in memory. The protected information cover factors inclusive of general and consistent with body item count, seen pixels consistent with item, and body through body visualization of the captured floor truth. These, in conjunction with aid for extending the toolset to consist of new information, make it viable to recognize and affirm the generated datasets earlier than the usage of them for version training. For instance, information can assist customers to determine whether or not a bigger dataset or one with distinctive area randomizations implemented is wanted for the precise hassle they're looking to solve. They additionally function a debugging device to identify troubles withinside the dataset. Statistics from the Synthesis mission.

VI. CONCLUSION

Training a version with artificial information that plays withinside the actual international provides numerous demanding situations associated with the way you generate a beneficial dataset. For item detection, the principle demised artificial dataset containing 400,000 photographs and 2D bounding container annotations. We additionally amassed and annotated an actual-international dataset, named UnityGroceries Real4 , which incorporates 1267 photographs of the sixty three goal grocery items. Difficulties are the way to create a virtual dual of the goal objects, diversify your artificial dataset, and teach your version the usage of this dataset. Using asset scanning answers lets in you to create an artificial model of your goal property in Unity. Then the usage of Unity Simulation, you could parametrically range components of the property and surroundings and generate artificial information at scale.

Training a version with artificial information that plays withinside the actual international provides numerous demanding situations associated with the way you generate a beneficial dataset. It used 3-D property and Unity notion package deal to accumulate the information and create a scene that's then made to stimulate. After simulation, artificial dataset is created in which we are able to create a dataset for insights and dataset information the usage of the amassed information. An Autonomous System is gift in which we will locate the actual and simulated information. In this machine the manner of coordination, control, making plans and notion takes place.

For object detection, the principle problems are the way to create a virtual dual of the goal objects, diversify your artificial dataset, and teach your version the usage of this dataset. Using asset scanning answers lets in you to create an artificial model of your goal property in Unity. Then the usage of

VII. REFERENCES

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, RealTime Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
- [2] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single Shot MultiBox Detector," in Computer Vision – ECCV 2016, pp. 21–37, 2016.
- [3] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal Loss for Dense Object Detection," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), Oct. 2017.
- [4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.
- [5] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), Oct. 2017.
- [6] M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 10778– 10787, 2020.
- [7] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.
- [8] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 4, pp. 834– 848, 2017.
- [9] O. Ronneberger, P. Fischer, and T. Brox, "UNet: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 (N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds.), (Cham), pp. 234– 241, Springer International Publishing, 2015.
- [10] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

-
- [11] W. Kuo, A. Angelova, J. Malik, and T.-Y. Lin, "ShapeMask: Learning to Segment Novel Objects by Refining Shape Priors," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Oct. 2019.
- [12] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-time Instance Segmentation," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019.
- [13] P. Voigt and A. v. d. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*. Springer Publishing Company, Incorporated, 1st ed., 2017.
- [14] S. Ren, *California Consumer Privacy Act (CCPA): An implementation guide*. IT Governance Publishing, 2019.
- [15] Unity, "Unity Technologies," 2021.
- [16] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge," *International Journal of Computer Vision*, vol. 88, pp. 303–338, June 2010.
- [17] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor Segmentation and Support Inference from RGBD Images," in *Computer Vision – ECCV 2012*, (Berlin, Heidelberg), pp. 746–760, Springer Berlin Heidelberg, 2012.
- [18] T.-Y. Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollar, "Microsoft COCO: Common Objects in Context," in *Computer Vision – ECCV 2014*, (Cham), pp. 740–755, Springer International Publishing, 2014.
- [19] S. Song, S. P. Lichtenberg, and J. Xiao, "SUN RGB-D: A RGB-D scene understanding benchmark suite," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.
- [21] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 2481–2495, Dec. 2017. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.
- [22] S. Hinterstoisser, O. Pauly, H. Heibel, M. Marek, and M. Bokeloh, "An Annotation Saved is an Annotation Earned: Using Fully Synthetic Training for Object Instance Detection," in Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, Oct. 2019.