Improving NRCan - VIA Rail Track Maintenance using Computer Vision

Kishan Patel

Graduate Student

Patel.ki@northeastern.edu

INTRODUCTION

PROBLEM DESCRIPTION

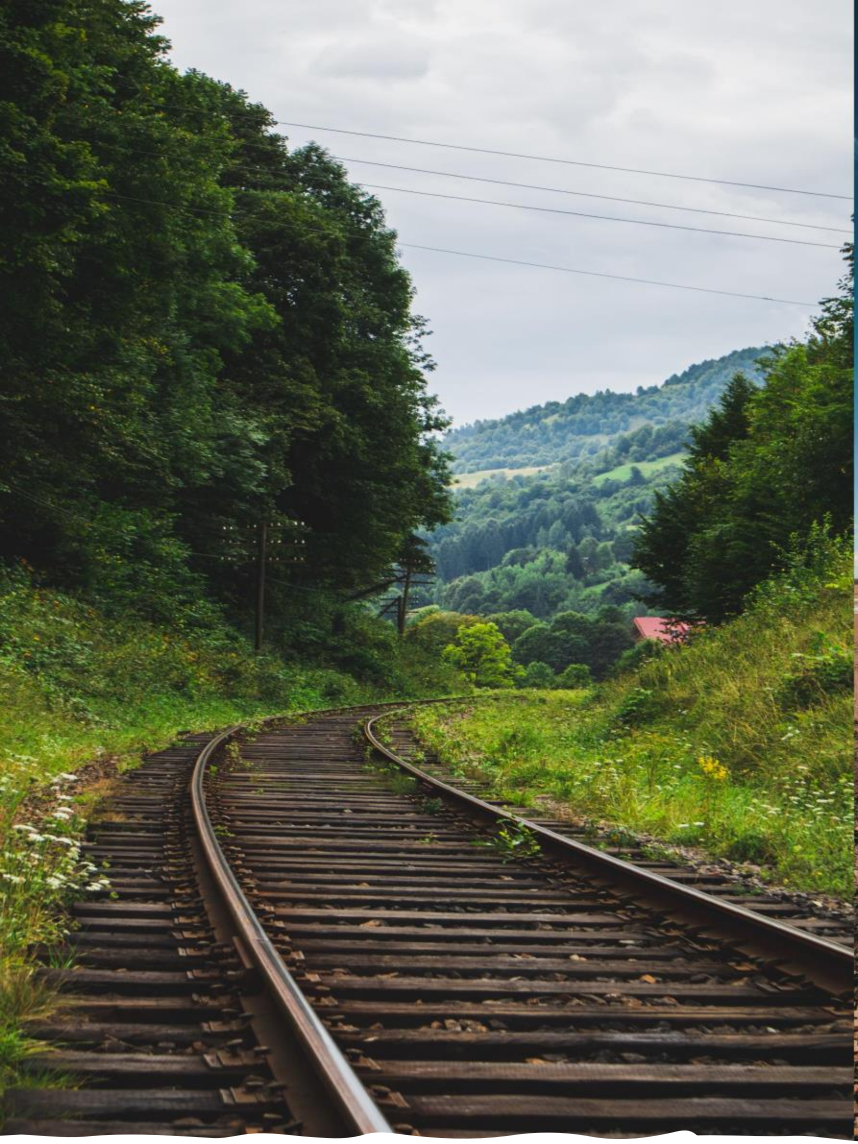
METHODOLOGY

RESULTS

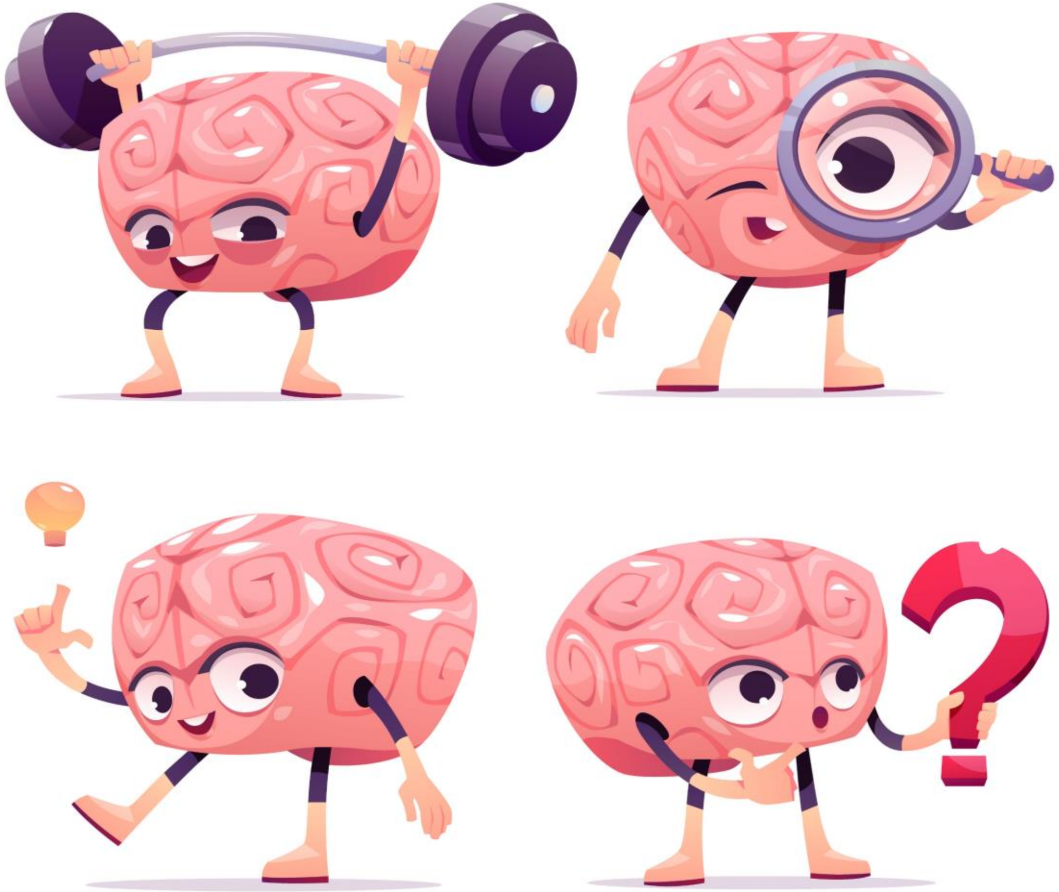
CONCLUSION

Agenda

INTRODUCTION



PROBLEM



RESEARCH

Key Points	Exploring Image Enhancement for Salient Object Detection in Low Light Images	A Crowdsourcing Repeated Annotations System for Visual Object Detection	Embedded YOLO	Real-time object detection method for embedded devices	Pruned-YOLO: Learning Efficient Object
Problem	It is impractical to assume that images are always captured in perfect light conditions. Low light images usually degrades with scene height and suffers from severe object information loss.	Large number of images with ground truth information in form of annotation is bottleneck for any object detection application.	Design a lightweight model that can balance the computational complexity and accuracy simultaneously	To improve the real-time of object detection, a fast object detection method on YOLOv4 based on YOLOv4-tiny to simplify the network structure and reduce parameters, which makes it be suitable for developing on the mobile and embedded devices.	Accurate real-time object detection plays a key role in various practical scenarios such as automatic driving and UAV surveillance. The memory limitation and poor computing power of edge devices hinder the deployment of high performance Convolutional Neural Networks (CNNs).
Proposed Solution	Physical based image enhancement enhances the image contrast by exploiting the relation among the atmosphere light and transmission map. Another subnet detects salient object from enhanced images.	Existing crowdsource techniques such as Amazon Mechanical Turk are studied and turn based annotation system consisting of three simple task (single object detection, a quality verification task, coverage verification task) is proposed.	A neural network model named Embedded YOLO which reduces number of parameters by using DSC_CSP module to replace the middle layers of YOLOv5	Replace two ResBlock-D modules with two CSPBlock modules in Yolov4 Tiny to reduce computation complexity. Implement auxiliary residual network block to extract more feature information of object to reduce detection error.	Iterative channel pruning is an effective method to obtain lightweight networks. In this paper, to measure the channel importance, they simultaneously consider the scale factor of batch normalization (BN) and the kernel weight of convolutional layers.
Evaluation Methodology	Used five SOD datasets such as DUT-OMRON, PASCAL-S, NTI-V1 are used to evaluate model for precision recall curves, F-measures and Mean absolute error.	Pascal VOC 2007 dataset is used for evaluation where 3 rounds of labelling are performed to deal with low annotation difficulty.	Used IVSLab and BDD100k dataset to evaluate performance of embedded Yolo architecture based on mAP, model size, GLOPs and speed	The mAP (mean value of average precision) , FPS (Frames per second) and GPU utilization are used to quantitatively evaluate the performance of different methods compared with YOLOv3, YOLOv4, YOLOv3-tiny, YOLOv4-tiny to test their performance in mAP and FPS.	Generous ablation studies are imported to investigate the properties of proposed method and compare it with similar state-of-the-art algorithms.
Results	Promising results were observed in NTI-V1 and public dataset with high volume of false postive due to noise interference.	Proposed technique shows improvement in accuracy by up to 8-10%. System can assist object detector to obtain higher accuracy in a cost effective manner.	Model achieved 0.59 mAP with 41 FPS on the mediaTek's Dimensity 1000 platform which confirmed suitability of proposed model on embedded systems	Simulation results show that the proposed method has faster object detection than YOLOv4-tiny and YOLOv3-tiny, and almost the same mean value of average precision as the YOLOv4-tiny.	Pruned-YOLOv3/v5, which is constructed via pruning YOLOv3/v5. The experimental results on the MS-COCO and VisDrone datasets show that the proposed model achieves a satisfactory balance between computational efficiency and detection accuracy.

Key Points	Exploring Image Enhancement for Salient Object Detection in Low Light Images	A Crowdsourcing Repeated Annotations System for Visual Object Detection	Embedded YOLO	Real-time object detection method for embedded devices	Pruned-YOLO: Learning Efficient Object
Problem	It is impractical to assume that images are always captured in perfect light conditions. Low light images usually degrades with scene height and suffers from severe object information loss.	Large number of images with ground truth information in form of annotation is bottleneck for any object detection application.	Design a lightweight model that can balance the computational complexity and accuracy simultaneously	To improve the real-time of object detection, a fast object detection method on YOLOv4 based on YOLOv4-tiny to simplify the network structure and reduce parameters, which makes it be suitable for developing on the mobile and embedded devices.	Accurate real-time object detection plays a key role in various practical scenarios such as automatic driving and UAV surveillance. The memory limitation and poor computing power of edge devices hinder the deployment of high performance Convolutional Neural Networks (CNNs).
Proposed Solution	Physical based image enhancement enhances the image constrast by exploiting the relation among the atmosphere light and transmission map. Another subnet detects salient object from enhanced images.	Existing crowdsource techniques such as Amazon Mechanical Turk are studied and turn based annotation system consisting of three simple task (single object detection, a quality verification task, coverage verification task) is proposed.	A neural network model named Embedded YOLO which reduces number of parameters by using DSC_CSP module to replace the middle layers of YOLOv5	Replace two ResBlock-D modules with two CSPBlock modules in Yolov4 Tiny to reduce computation complexity. Implement auxiliary residual network block to extract more feature information of object to reduce detection error.	Iterative channel pruning is an effective method to obtain lightweight networks. In this paper, to measure the channel importance, they simultaneously consider the scale factor of batch normalization (BN) and the kernel weight of convolutional layers.
Evaluation Methodology	Used five SOD datasets such as DUT-OMRON, PASCAL-S, NTI-V1 are used to evaluate model for precision recall curves, F-measures and Mean absolute error.	Pascal VOC 2007 dataset is used for evaluation where 3 rounds of labelling are performed to deal with low annotation difficulty.	Used IVSLab and BDD100k dataset to evaluate performance of embedded Yolo architecture based mAP, model size, GLOPs and speed	The mAP (mean value of average precision) , FPS (Frames per second) and GPU utilization are used to quantitatively evaluate the performance of different methods compared with YOLOv3, YOLOv4, YOLOv3-tiny, YOLOv4-tiny to test their performance in mAP and FPS.	Generous ablation studies are imported to investigate the properties of proposed method and compare it with similar state-of-the-art algorithms.
Results	Promising results were observed in NTI-V1 and public dataset with high volume of false postive due to noise interference.	Proposed technique shows improvement in accuracy by up to 8-10%. System can assist object detector to obtain higher accuracy in a cost effective manner.	Model achieved 0.59 mAP with 41 FPS on the mediaTek's Dimensity 1000 platform which confirmed suitability of proposed model on embedded systems	Simulation results show that the proposed method has faster object detection than YOLOv4-tiny and YOLOv3-tiny, and almost the same mean value of average precision as the YOLOv4-tiny.	Pruned-YOLOv3/v5, which is constructed via pruning YOLOv3/v5. The experimental results on the MS-COCO and VisDrone datasets show that the proposed model achieves a satisfactory balance between computational efficiency and detection accuracy.

Key Points	Exploring Image Enhancement for Salient Object Detection in Low Light Images	A Crowdsourcing Repeated Annotations System for Visual Object Detection	Embedded YOLO	Real-time object detection method for embedded devices	Pruned-YOLO: Learning Efficient Object
Problem	It is impractical to assume that images are always captured in perfect light conditions. Low light images usually degrades with scene height and suffers from severe object information loss.	Large number of images with ground truth information in form of annotation is bottleneck for any object detection application.	Design a lightweight model that can balance the computational complexity and accuracy simultaneously	To improve the real-time of object detection, a fast object detection method on YOLOv4 based on YOLOv4-tiny to simplify the network structure and reduce parameters, which makes it be suitable for developing on the mobile and embedded devices.	Accurate real-time object detection plays a key role in various practical scenarios such as automatic driving and UAV surveillance. The memory limitation and poor computing power of edge devices hinder the deployment of high performance Convolutional Neural Networks (CNNs).
Proposed Solution	Physical based image enhancement enhances the image constrast by exploiting the relation among the atmosphere light and transmission map. Another subnet detects salient object from enhanced images.	Existing crowdsource techniques such as Amazon Mechanical Turk are studied and turn based annotation system consisting of three simple task (single object detection, a quality verification task, coverage verification task) is proposed.	A neural network model named Embedded YOLO which reduces number of parameters by using DSC_CSP module to replace the middle layers of YOLOv5	Replace two ResBlock-D modules with two CSPBlock modules in Yolov4 Tiny to reduce computation complexity. Implement auxiliary residual network block to extract more feature information of object to reduce detection error.	Iterative channel pruning is an effective method to obtain lightweight networks. In this paper, to measure the channel importance, they simultaneously consider the scale factor of batch normalization (BN) and the kernel weight of convolutional layers.
Evaluation Methodology	Used five SOD datasets such as DUT-OMRON, PASCAL-S, NTI-V1 are used to evaluate model for precision recall curves, F-measures and Mean absolute error.	Pascal VOC 2007 dataset is used for evaluation where 3 rounds of labelling are performed to deal with low annotation difficulty.	Used IVSLab and BDD100k dataset to evaluate performance of embedded Yolo architecture based on mAP, model size, GLOPs and speed	The mAP (mean value of average precision) , FPS (Frames per second) and GPU utilization are used to quantitatively evaluate the performance of different methods compared with YOLOv3, YOLOv4, YOLOv3-tiny, YOLOv4-tiny to test their performance in mAP and FPS.	Generous ablation studies are imported to investigate the properties of proposed method and compare it with similar state-of-the-art algorithms.
Results	Promising results were observed in NTI-V1 and public dataset with high volume of false postive due to noise interference.	Proposed technique shows improvement in accuracy by up to 8-10%. System can assist object detector to obtain higher accuracy in a cost effective manner.	Model achieved 0.59 mAP with 41 FPS on the mediaTek's Dimensity 1000 platform which confirmed suitability of proposed model on embedded systems	Simulation results show that the proposed method has faster object detection than YOLOv4-tiny and YOLOv3-tiny, and almost the same mean value of average precision as the YOLOv4-tiny.	Pruned-YOLOv3/v5, which is constructed via pruning YOLOv3/v5. The experimental results on the MS-COCO and VisDrone datasets show that the proposed model achieves a satisfactory balance between computational efficiency and detection accuracy.

Key Points	Exploring Image Enhancement for Salient Object Detection in Low Light Images	A Crowdsourcing Repeated Annotations System for Visual Object Detection	Embedded YOLO	Real-time object detection method for embedded devices	Pruned-YOLO: Learning Efficient Object
Problem	It is impractical to assume that images are always captured in perfect light conditions. Low light images usually degrades with scene height and suffers from severe object information loss.	Large number of images with ground truth information in form of annotation is bottleneck for any object detection application.	Design a lightweight model that can balance the computational complexity and accuracy simultaneously	To improve the real-time of object detection, a fast object detection method on YOLOv4 based on YOLOv4-tiny to simplify the network structure and reduce parameters, which makes it be suitable for developing on the mobile and embedded devices.	Accurate real-time object detection plays a key role in various practical scenarios such as automatic driving and UAV surveillance. The memory limitation and poor computing power of edge devices hinder the deployment of high performance Convolutional Neural Networks (CNNs).
Proposed Solution	Physical based image enhancement enhances the image constrast by exploiting the relation among the atmosphere light and transmission map. Another subnet detects salient object from enhanced images.	Existing crowdsource techniques such as Amazon Mechanical Turk are studied and turn based annotation system consisting of three simple task (single object detection, a quality verification task, coverage verification task) is proposed.	A neural network model named Embedded YOLO which reduces number of parameters by using DSC_CSP module to replace the middle layers of YOLOv5	Replace two ResBlock-D modules with two CSPBlock modules in Yolov4 Tiny to reduce computation complexity. Implement auxiliary residual network block to extract more feature information of object to reduce detection error.	Iterative channel pruning is an effective method to obtain lightweight networks. In this paper, to measure the channel importance, they simultaneously consider the scale factor of batch normalization (BN) and the kernel weight of convolutional layers.
Evaluation Methodology	Used five SOD datasets such as DUT-OMRON, PASCAL-S, NTI-V1 are used to evaluate model for precision recall curves, F-measures and Mean absolute error.	Pascal VOC 2007 dataset is used for evaluation where 3 rounds of labelling are performed to deal with low annotation difficulty.	Used IVSLab and BDD100k dataset to evaluate performance of embedded Yolo architecture based mAP, model size, GLOPs and speed	The mAP (mean value of average precision) , FPS (Frames per second) and GPU utilization are used to quantitatively evaluate the performance of different methods compared with YOLOv3, YOLOv4, YOLOv3-tiny, YOLOv4-tiny to test their performance in mAP and FPS.	Generous ablation studies are imported to investigate the properties of proposed method and compare it with similar state-of-the-art algorithms.
Results	Promising results were observed in NTI-V1 and public dataset with high volume of false postive due to noise interference.	Proposed technique shows improvement in accuracy by up to 8-10%. System can assist object detector to obtain higher accuracy in a cost effective manner.	Model achieved 0.59 mAP with 41 FPS on the mediaTek's Dimensity 1000 platform which confirmed suitability of proposed model on embedded systems	Simulation results show that the proposed method has faster object detection than YOLOv4-tiny and YOLOv3-tiny, and almost the same mean value of average precision as the YOLOv4-tiny.	Pruned-YOLOv3/v5, which is constructed via pruning YOLOv3/v5. The experimental results on the MS-COCO and VisDrone datasets show that the proposed model achieves a satisfactory balance between computational efficiency and detection accuracy.

Key Points	Exploring Image Enhancement for Salient Object Detection in Low Light Images	A Crowdsourcing Repeated Annotations System for Visual Object Detection	Embedded YOLO	Real-time object detection method for embedded devices	Pruned-YOLO: Learning Efficient Object
Problem	It is impractical to assume that images are always captured in perfect light conditions. Low light images usually degrades with scene height and suffers from severe object information loss.	Large number of images with ground truth information in form of annotation is bottleneck for any object detection application.	Design a lightweight model that can balance the computational complexity and accuracy simultaneously	To improve the real-time of object detection, a fast object detection method on YOLOv4 based on YOLOv4-tiny to simplify the network structure and reduce parameters, which makes it be suitable for developing on the mobile and embedded devices.	Accurate real-time object detection plays a key role in various practical scenarios such as automatic driving and UAV surveillance. The memory limitation and poor computing power of edge devices hinder the deployment of high performance Convolutional Neural Networks (CNNs).
Proposed Solution	Physical based image enhancement enhances the image constrast by exploiting the relation among the atmosphere light and transmission map. Another subnet detects salient object from enhanced images.	Existing crowdsource techniques such as Amazon Mechanical Turk are studied and turn based annotation system consisting of three simple task (single object detection, a quality verification task, coverage verification task) is proposed.	A neural network model named Embedded YOLO which reduces number of parameters by using DSC_CSP module to replace the middle layers of YOLOv5	Replace two ResBlock-D modules with two CSPBlock modules in Yolov4 Tiny to reduce computation complexity. Implement auxiliary residual network block to extract more feature information of object to reduce detection error.	Iterative channel pruning is an effective method to obtain lightweight networks. In this paper, to measure the channel importance, they simultaneously consider the scale factor of batch normalization (BN) and the kernel weight of convolutional layers.
Evaluation Methodology	Used five SOD datasets such as DUT-OMRON, PASCAL-S, NTI-V1 are used to evaluate model for precision recall curves, F-measures and Mean absolute error.	Pascal VOC 2007 dataset is used for evaluation where 3 rounds of labelling are performed to deal with low annotation difficulty.	Used IVSLab and BDD100k dataset to evaluate performance of embedded Yolo architecture based mAP, model size, GLOPs and speed	The mAP (mean value of average precision) , FPS (Frames per second) and GPU utilization are used to quantitatively evaluate the performance of different methods compared with YOLOv3, YOLOv4, YOLOv3-tiny, YOLOv4-tiny to test their performance in mAP and FPS.	Generous ablation studies are imported to investigate the properties of proposed method and compare it with similar state-of-the-art algorithms.
Results	Promising results were observed in NTI-V1 and public dataset with high volume of false postive due to noise interference.	Proposed technique shows improvement in accuracy by up to 8-10%. System can assist object detector to obtain higher accuracy in a cost effective manner.	Model achieved 0.59 mAP with 41 FPS on the mediaTek's Dimensity 1000 platform which confirmed suitability of proposed model on embedded systems	Simulation results show that the proposed method has faster object detection than YOLOv4-tiny and YOLOv3-tiny, and almost the same mean value of average precision as the YOLOv4-tiny.	Pruned-YOLOv3/v5, which is constructed via pruning YOLOv3/v5. The experimental results on the MS-COCO and VisDrone datasets show that the proposed model achieves a satisfactory balance between computational efficiency and detection accuracy.

RESEARCH QUESTIONS

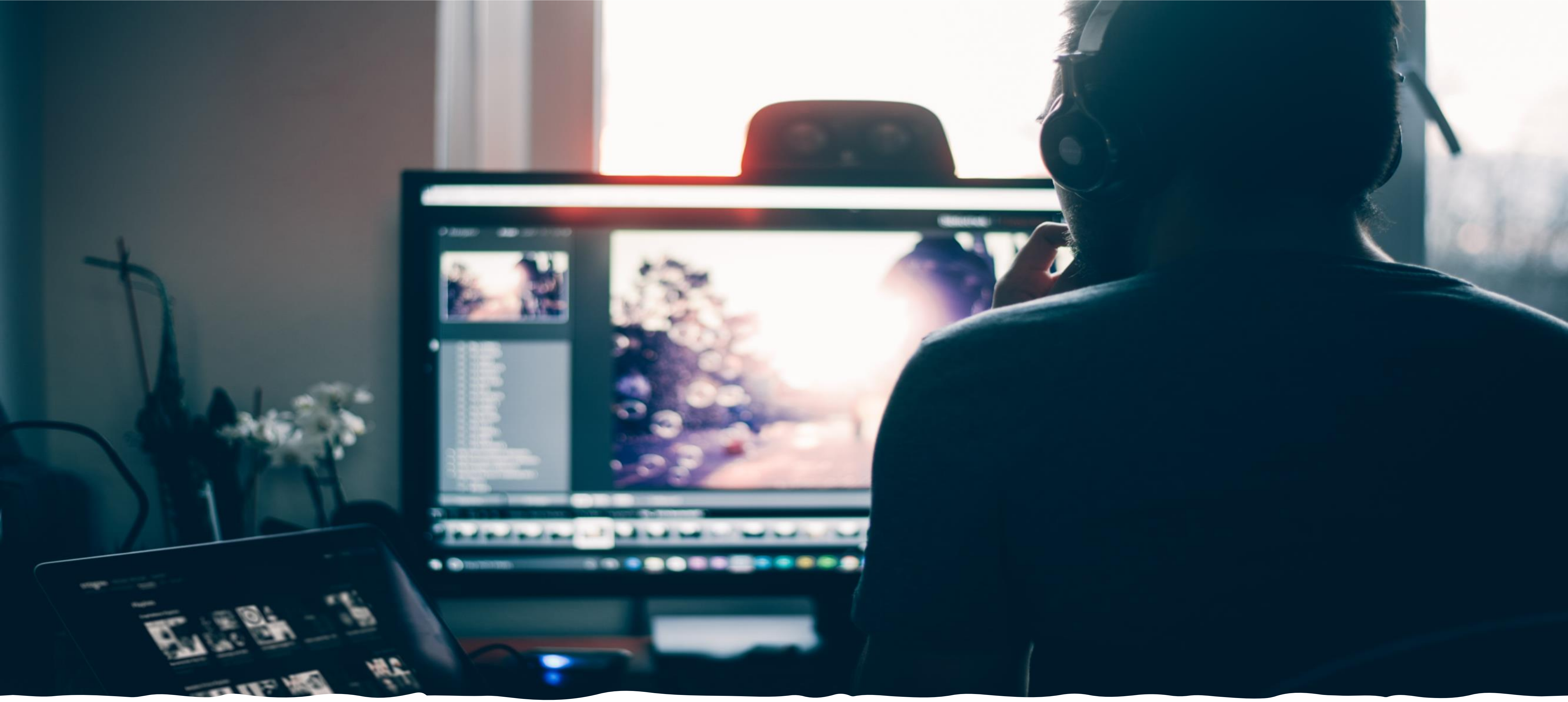


IMAGE MANIPULATION

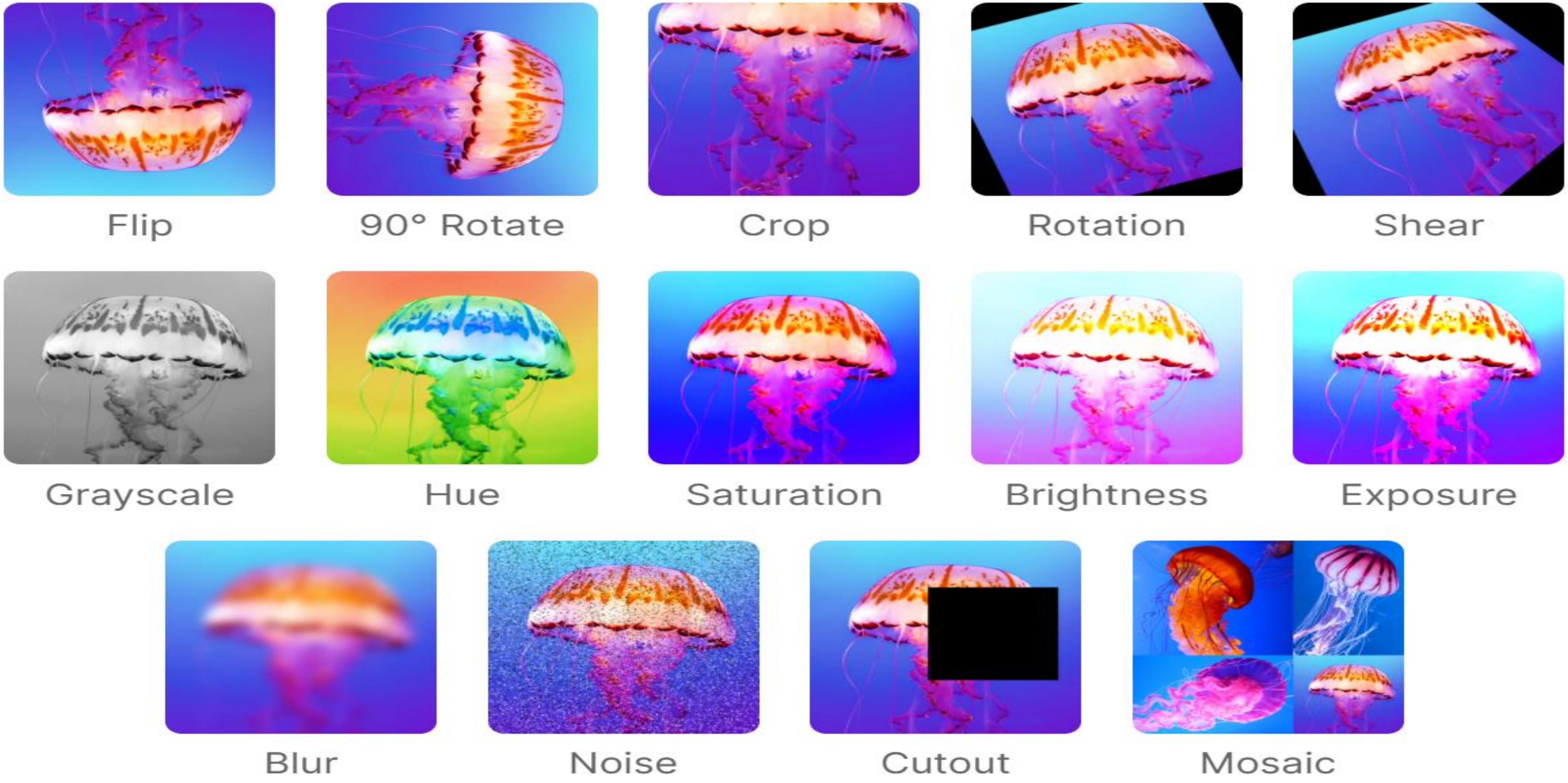
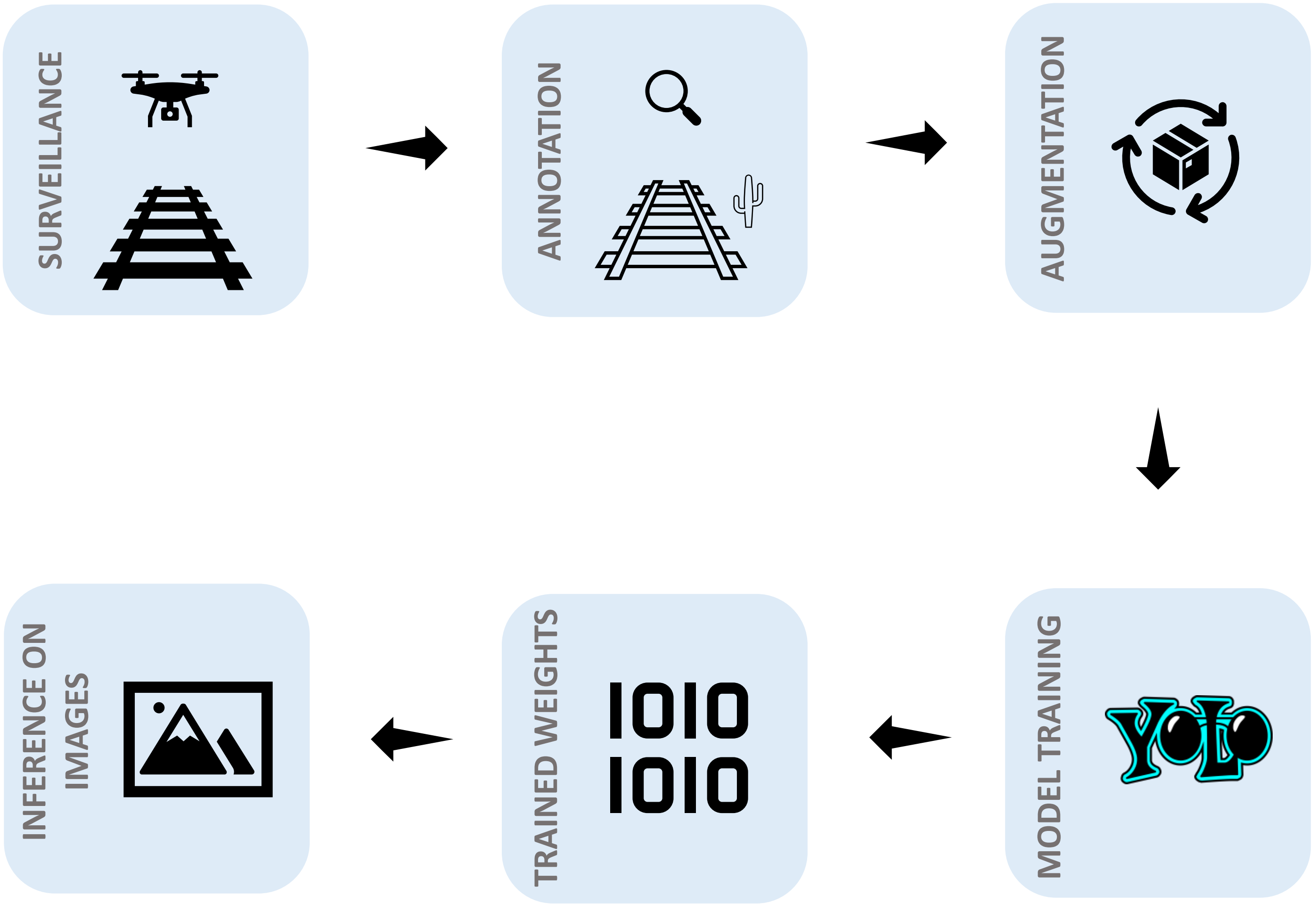
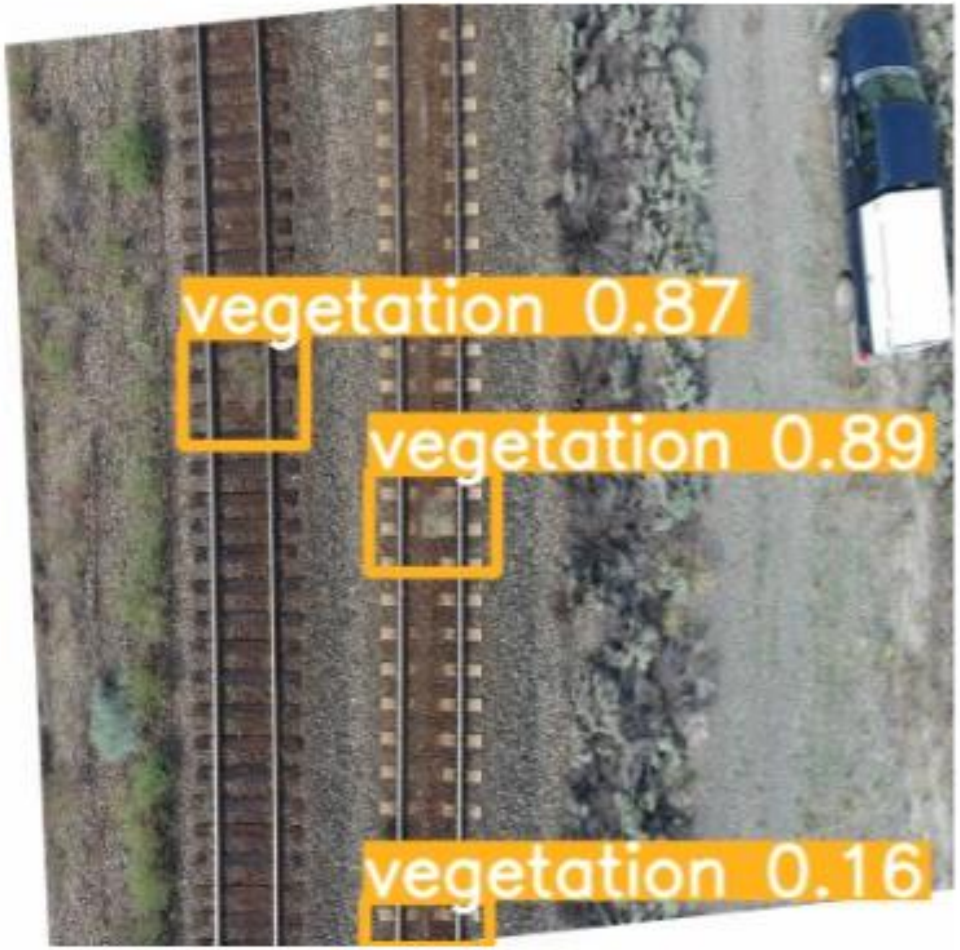


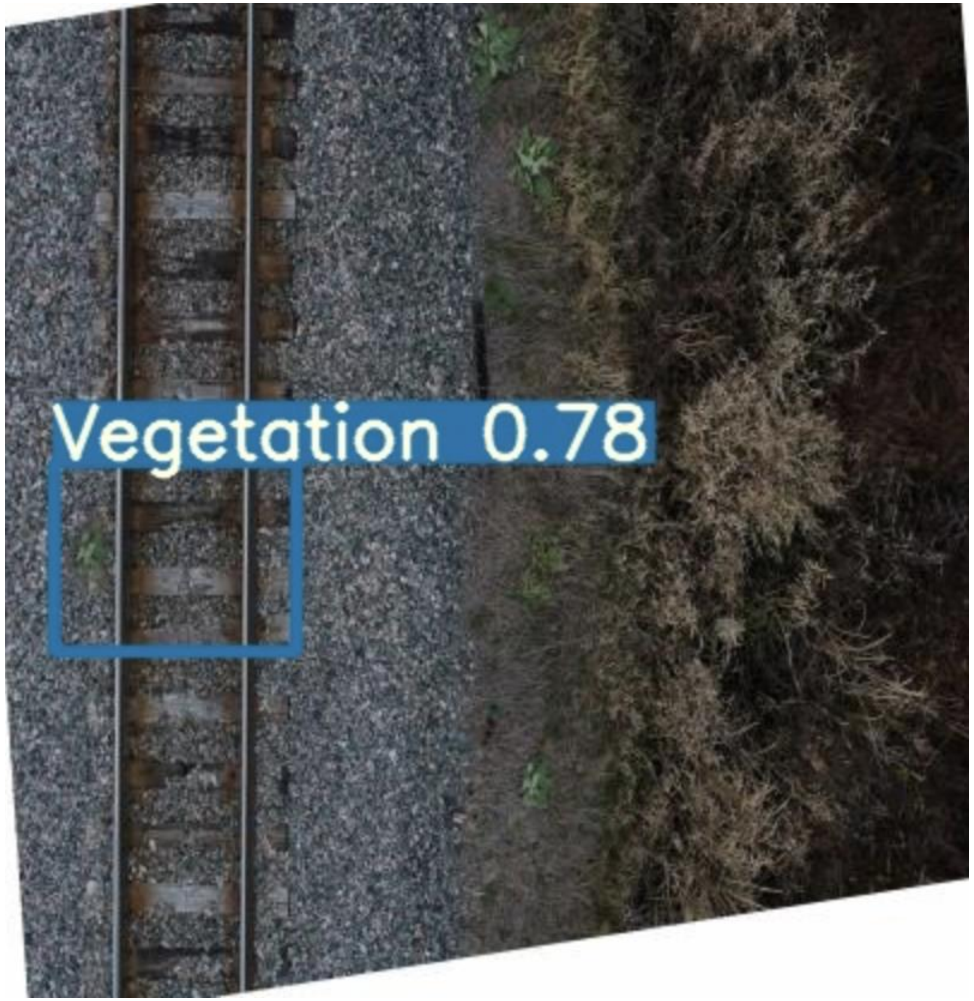
IMAGE LEVEL AUGMENTATION

YOLO VERSIONS

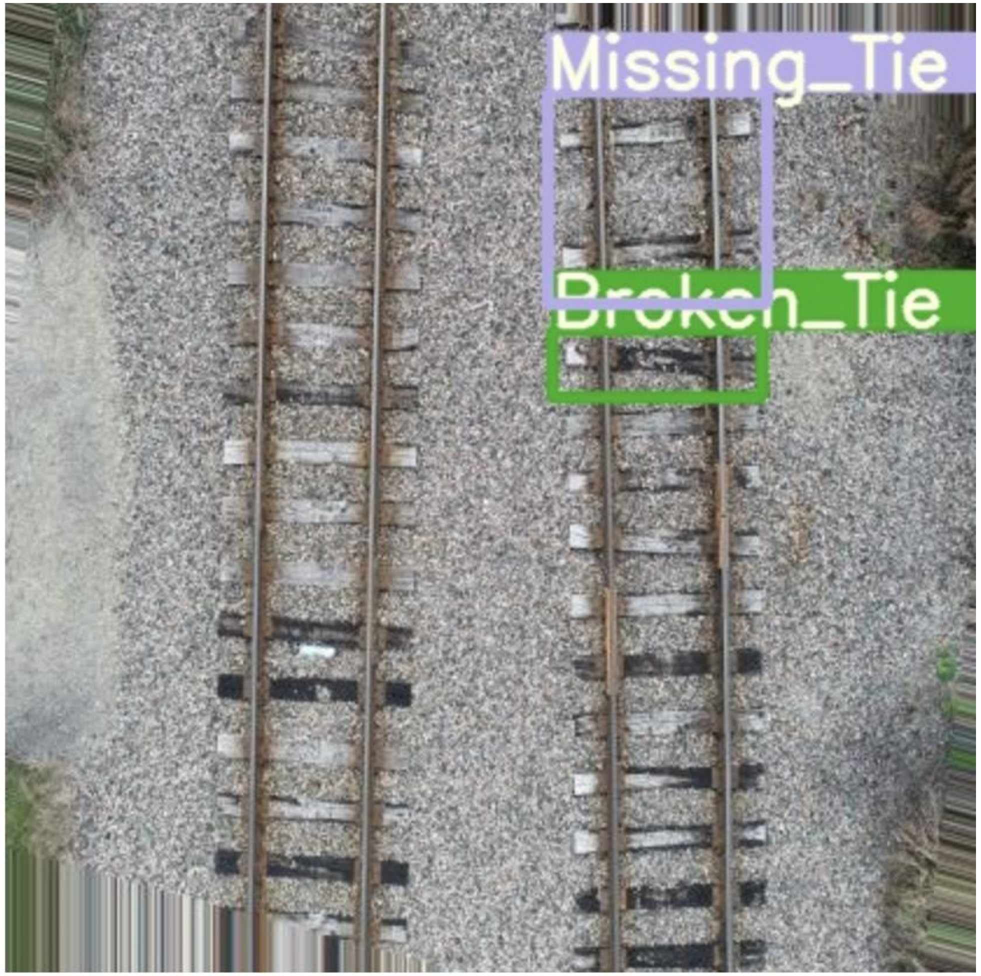


HOW DOES IT WORK?





INFERRED IMAGES



INFERRED IMAGES

TRAIN / TEST SPLIT

Training Set 69%	Validation Set 20%	Testing Set 10%
1.2k images	350 images	175 images

PREPROCESSING

Auto-Orient: Applied

Resize: Stretch to 416×416

AUGMENTATIONS

Outputs per training example: 3

Saturation: Between -50% and +50%

DETAILS

Version Name: AUG_SATURATION 2:29am

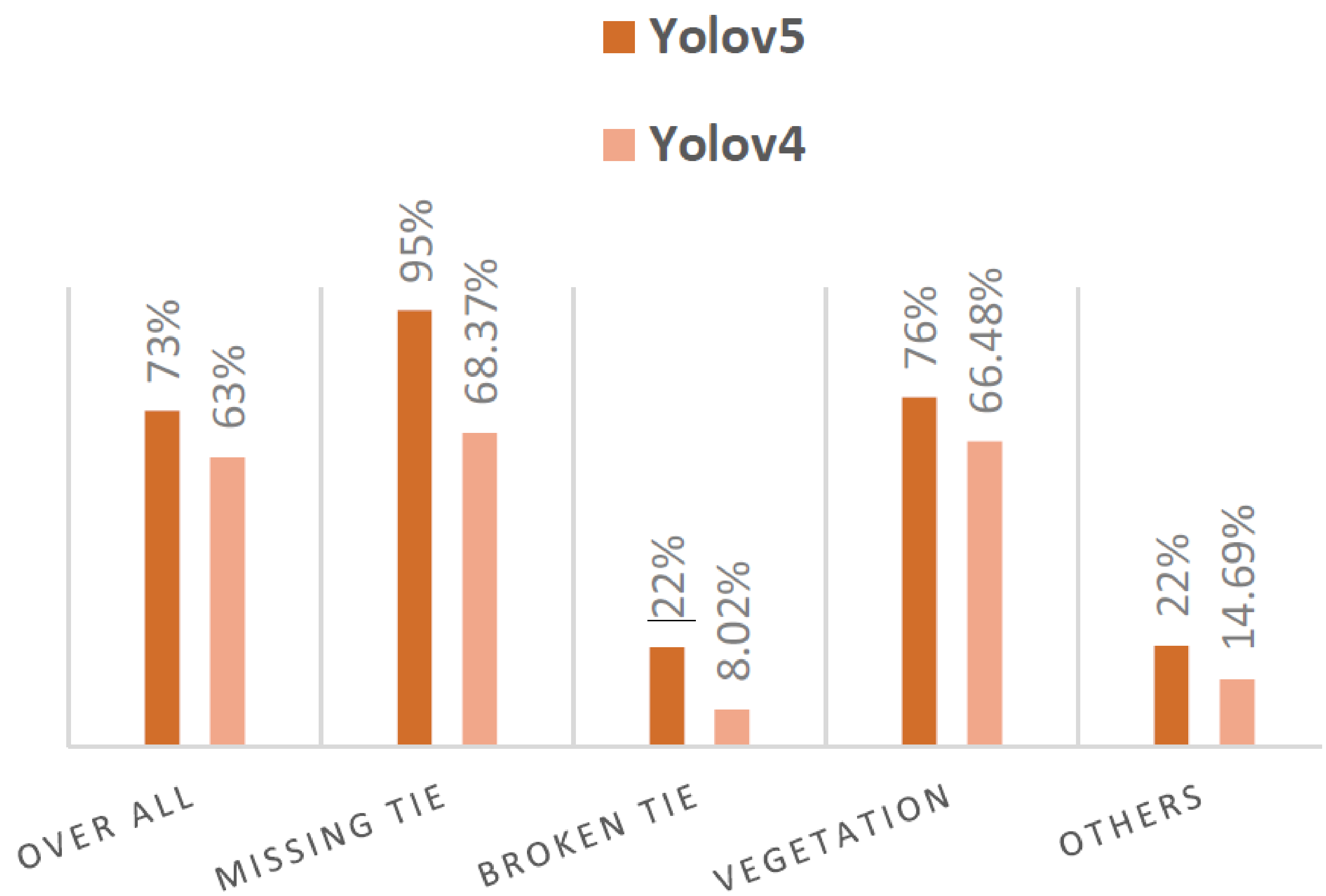
Version ID: 4

Generated: Dec 12, 2021

Annotation Group: Points-of-interests

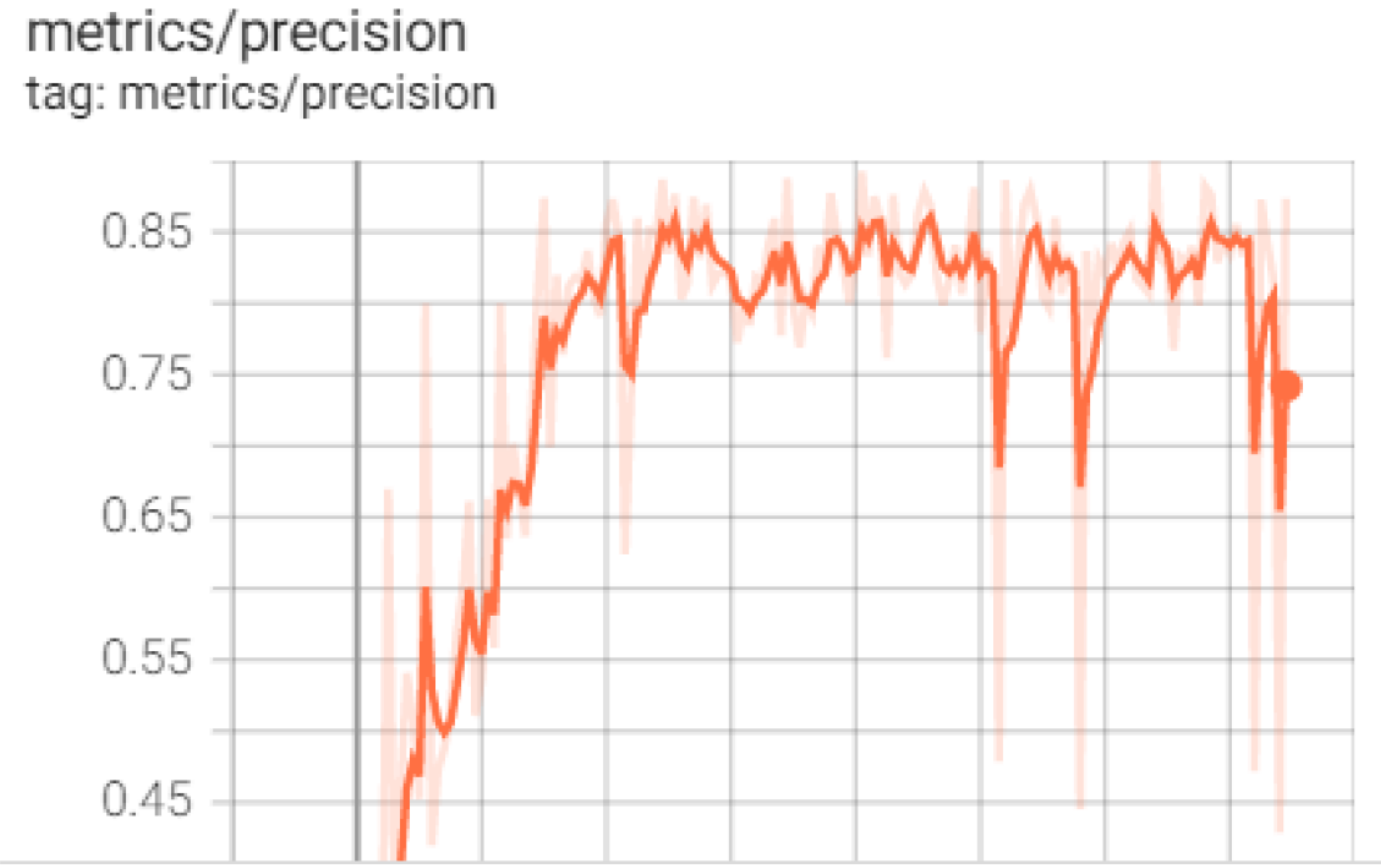
SIMULATION SETUP

Accuracy Comparison for Points of Interest



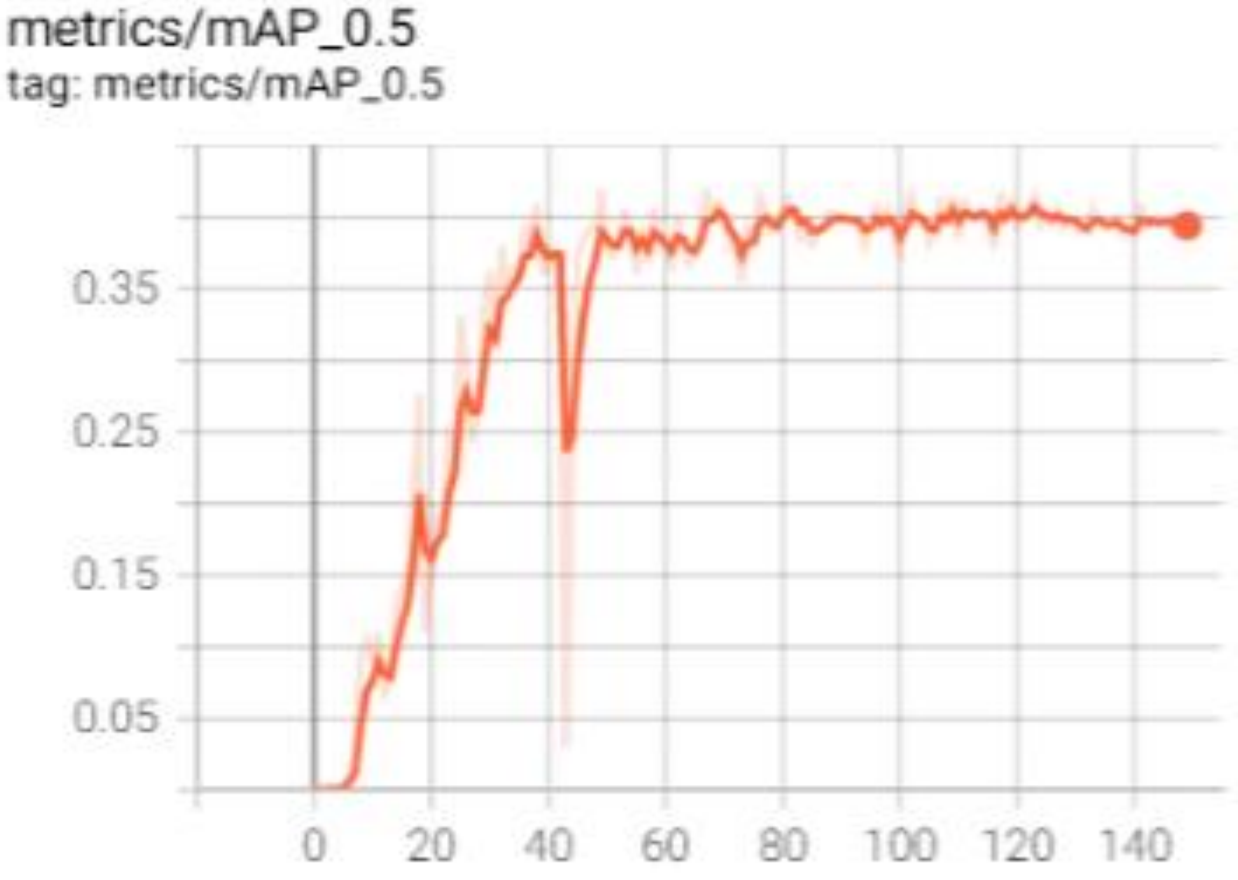
- Accuracy

Precision

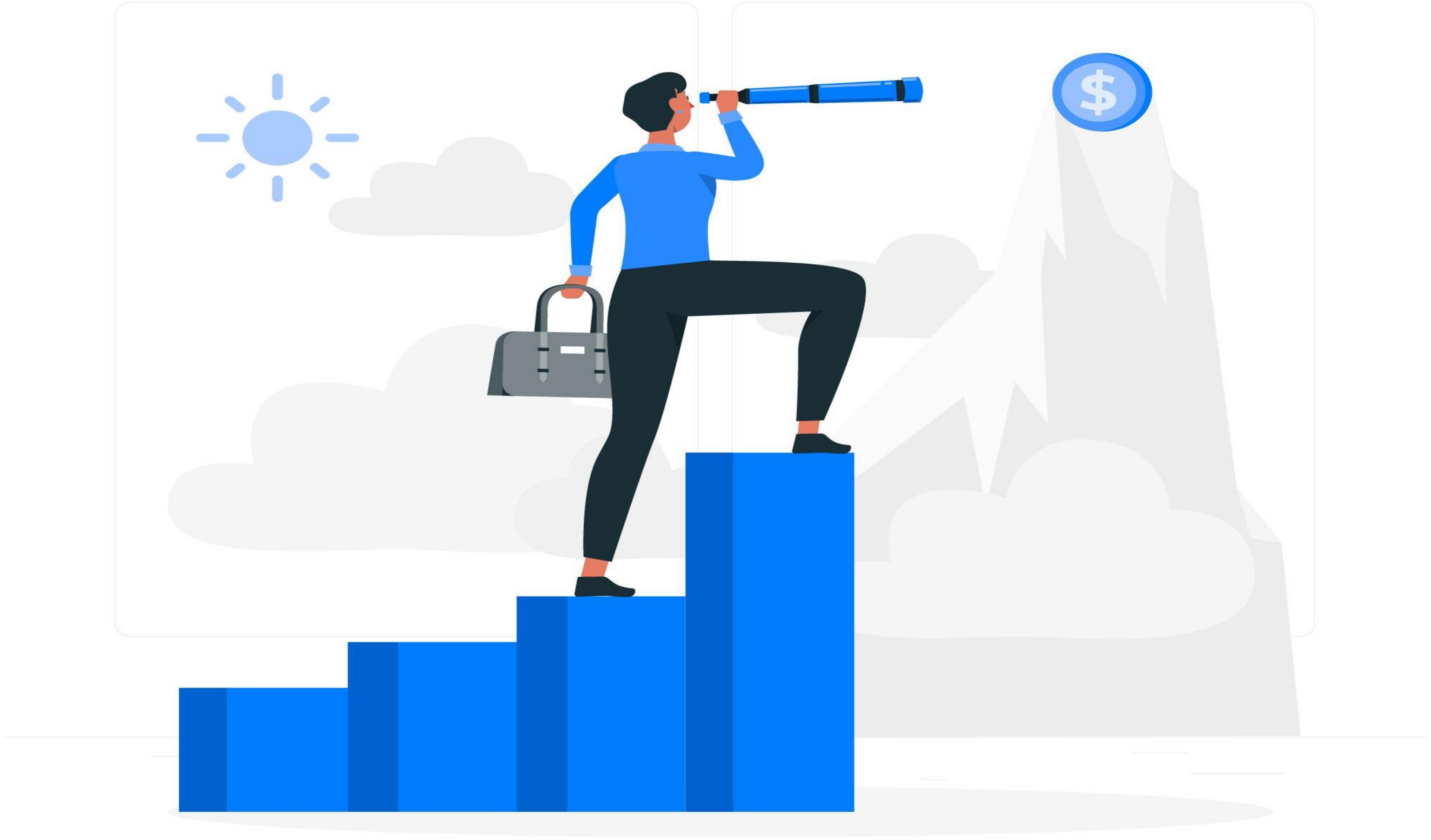

metrics/precision
tag: metrics/precision
0.85
0.75
0.65
0.55
0.45

Recall

Mean Average Precision



CONCLUSION



FUTURE WORK


THANK
YOU