PS1 FINAL REPORT

ON

Dent and Scratch Annotation and License Disk Detection Model

BY

Raguram Venkatesan 2021A7PS0150H

Nischith P 2021A7PS3220H

Sandeep Y 2021A7PS3113H

Kamalesh Ram R 2021A7PS0286H

Daksh Shah 2021AAPS0962G

Raj Patel 2021AAPS2882H

Darsh Doshi 2021A7PS2836G

AT

Chaob Tech. , Carscan , Pune

A Practice School 1 Station of

Birla Institute of Technology and Science, Pilani

(July 2023)

A REPORT

ON

Dent and Scratch Annotation and License Disk Detection Model

BY

Name of Student	ID No.	Discipline
1.) Raguram Venkatesan	2021A7PS0150H	Computer Science
2.) Nischith P	2021A7PS3220H	Computer Science
3.) Sandeep Y	2021A7PS3113H	Computer Science
4.) Kamalesh Ram R	2021A7PS0286H	Computer Science
5.) Daksh Shah	2021AAPS0962G	Elec & Comm.
6.) Raj Patel	2021AAPS2882H	Elec & Comm.
7.) Darsh Doshi	2021A7PS2836G	Computer Science

Prepared in partial fulfillment of the

Practice School – I (COURSE NO: BITS F221)

AT

Chaob Tech. Carscan , Pune

A Practice School 1 Station of

Birla Institute of Technology and Science, Pilani

(July 2023)

ACKNOWLEDGEMENTS

We would like to express our gratitude to the following people for their help in the work leading to this report:

Name and Designation of Company Mentors:

Karthik Itharajula- Annotation Officer

Pushpak Bhoge- Machine Learning Engineer

Name and Designation of Faculty Mentors:

Dr. Ranjit S Patil- Assistant Professor, Department of Mechanical Engineering, BITS Pilani K.K. Birla Goa Campus

Farha Masroor- PhD Scholar, BITS Pilani K.K. Birla Goa Campus

- We would also like to thank the Practice School Division, BITS-PILANI, for conducting this Program.

1.) Raguram Venkatesan 2021A7PS0150H

2.) Nischith P 2021A7PS3220H

3.) Sandeep Y 2021A7PS3113H

4.) Kamalesh Ram R 2021A7PS0286H

5.) Daksh Shah 2021AAPS0962G

6.) Raj Patel 2021AAPS2882H

7.) Darsh Doshi 2021A7PS2836G

Station: Chaob Tech , Carscan , Pune (remote)

Duration: 53 Days (30th May 2023 to 21st July 2023)

Date of submission: 18th July 2023

Title of project: Dent and Scratch Annotation and License Disk Detection Model

ID NOs. / Names of students / Disciplines of the students:

Name of Student	ID No.	Discipline
1.) Raguram Venkatesan	2021A7PS0150H	Computer Science
2.) Nischith P	2021A7PS3220H	Computer Science
3.) Sandeep Y	2021A7PS3113H	Computer Science
4.) Kamalesh Ram R	2021A7PS0286H	Computer Science
5.) Daksh Shah	2021AAPS0962G	Elec & Comm.
6.) Raj Patel	2021AAPS2882H	Elec & Comm.
7.) Darsh Doshi	2021A7PS2836G	Computer Science

Name of PS Faculty: Prof. Ranjit Patil

Abstract

This report presents a comprehensive analysis of utilizing computer vision models to automatically detect and localize dents and scratches in car images. The study reviews existing literature on object detection and segmentation techniques for damage assessment in automobiles. An annotated dataset of diverse car images is created, encompassing various models, lighting conditions, and damage severity levels. Challenges including occlusions and complex backgrounds are addressed, along with strategies for improvement. The results demonstrate the effectiveness of the proposed models, with potential applications in automotive repair, insurance claims, and pre-owned car evaluations. This analysis contributes to advancing the field of computer vision and paves the way for intelligent systems automating the automotive damage assessment process.

Objectives Addressed

Submission of Assignment – 1

To submit an assignment based on the benefits of AI in our lives

Product Training Session

Led by Neha Verma, this meet summarised their products and

Introductory meet

Basic introduction and giving an idea of day-to-day operations of the company.

services, the steps a client (since it is a v2v company) must take when utilizing their service mainly through

a. Webapp
b. WhatsApp
c. Mobile App

We were eager to find out our roles and how we could be of service to the company. And a few days after our allotment we got a mail from Simran Thakur(designation) welcoming us to the company along with which, we were also asked to come up with a plan outlining how AI can be used to improve specific tech and product processes within the company. The mail also specifically mentioned that our plan should focus on practical and actionable ideas that can be implemented within the company's existing framework. We were encouraged to think about areas such as:

- Automating repetitive tasks to free up valuable time for our team members
- Enhancing data analysis and insights to drive informed decision-making
- Improving customer support and engagement through AI-powered solutions

- Optimizing product testing and quality assurance processes using AI algorithms

- Streamlining collaboration and communication among cross-functional teams

- Product documentation like technical, BRD, etc

Needless to say, creativity, innovation, and a well-reasoned approach were something that they valued as an institution and were consequently expected of us in our very first assignment. We were supposed to submit this assignment on the first day of our commencement before our orientation session began. So, we took it to our hearts and made it a goal to create an everlasting first impression with our presentation. We decided to have a meeting where we could brainstorm a few ideas and come up with a basic structure for how the presentation would look like. After this, we had 4 days' time to develop said ideas and come up with a unique proposal individually that we hoped would have a significant impact on our organization and provide valuable insights for future growth.

On 29th May we had to submit our assignments by 1 pm as we had our orientation call by 7 pm. This was done so that our work could be reviewed by the company before we get a chance to meet the head of departments in the company. Our orientation meet was headed by Avadhut Joshi (Head of Operations) and Shravan Raj (designation), who was working from South Africa at that time. We also had our faculty in charge guiding us through the initial steps of our learning experience by observing our virtual interactions. Both AJ and Shravan (They specifically asked us to use their first name to address them) were very friendly and enthusiastic about mentoring interns and we did our part by reciprocating this excitement from our side. We also got to know that Shravan Raj was a 2015 BITS MSc Chem+ Mech pass-out and that he always had a soft spot when it came to mentoring BITS interns. We were asked to introduce ourselves by giving our name, areas of interest, and hobbies. Overall, it was a light-hearted session where not a lot of technical stuff was discussed. We were yet to be put in a role and AJ gave us

assurance that this would be assigned to us in a day or two. There was a research project which required 2 people and a general team that would get involved in the inner workings and day-to-day of the company and could be involved in a variety of domains including back end, data annotations, AI modelling etc. Apart from this, we were also asked to look out for two training sessions on the, product and services &about the AI, that would be scheduled in the next couple of days. This meeting ended with a brief Q&A session.

As per the schedule that was given to us via mail, the next day, our training session took place on 31st May around 4 pm. Before the discourse started, Neha (designation), asked us about our understanding of the company and the services that we provided. We attempted to briefly summarise them. She told us that, in that meeting, we would look at the different journeys and modes a user can use to avail of the company's services. There are 3 different journeys:

- Web app
- WhatsApp
- Mobile app

And 2 modes in general. Namely:

- Portrait mode
- Landscape mode

The general use-case for the company, since it is a v2v company:

- Insurance company
- Car buy and sell company
- Fleet management company
- Rental car company

etc.

More broadly the sub-journeys were also classified into 3 again:

- Inspection
- Claim
- 360

And at the time of the meeting, we were also informed that the company was looking for ways to improvise and optimize the WhatsApp inspection journey. The web application's journey development was put on hold to achieve said goal. Neha shared her screen with us and started explaining the journey by giving us an example from the user and company side of the interface simultaneously. She showed us a demo video that was done by the company to explain the steps in detail. The first step was to trigger a link to the customer via WhatsApp to get a pop-up message which asked the customer if they would like to "claim now" or "later", setting a reminder for them to claim the service in 1,3 or 5 hours. After this, the customer must accept the terms and conditions of the company to proceed. The customer must choose the "Lets Capture Image" option to choose a damaged side of the car from the 4 given options which were:

- Front
- Back
- Passenger
- Driver

Moving along, the customer also must choose the specific damaged part on that side to get started. You must click pictures of all the parts on the chosen side and the necessary parts on the other side (for creating a better visual model for the ai to work on). The customer is also instructed to take a picture of their license disc and odometer to complete the process of taking pictures.

The final step is to give your preferred location to alert car companies in the vicinity. In the video, after this part, it immediately cut to the dashboard where this claim was present under columns such as

- SA Id (South African id -Unique identification)
- Policy No.
- Customer and AI report
- Vin no.
- GT Estimate
- Status

And now that this claim is submitted it is broadcast to the necessary companies in the vicinity. An AI report is instantly generated which finds out the damaged parts in the customer reports by using a trained model. This is called the first report and is marked as such. After this, the data annotations team marks it clearly in order to generate the final report.

After this, we moved on to the web app journey. And at the beginning of this, we were made aware of the 3 environments that were available in a hierarchical order so as to facilitate the inclusion of new updates systematically. We were also given a routine rundown of the usual things such as app permissions and information verification from users. After this, the app guides the user to click accurate pics of the car by providing a framework of the pic that is needed. And when all of this is done an assessment summary is created to submit all your info and images as a part of the claim.

This claim is then submitted to the dashboard as usual and we went on our way to look at the next sub-journey exclusive to the web app, 360. The manner in which the pictures are taken here is slightly different. You will have to position the camera as asked and the app will take photos for you. You will have to move around the car in circles during the process hence the name 360. After this process, you can see a 3d model of your car in 360 where you are allowed to annotate the hotspots so as to emphasize the

damages. And this claim when submitted will give the users a unique reference id to edit details later.

The session then ended with a Q&A session. It was very informative and helpful. A day after that we had our AI training session where we were informed that 7 of us are going to be a part of the Data annotations team headed by Karthik Itharajulu. They also said that we would be transferred to the model training team after we have had enough experience with our initial project. Other than this, there was nothing much of significance said in this meet.

TABLE OF CONTENTS

Chapter I

An Intro to Carscan.AI

1.1 Introduction to Carscan

We, students of BITS, are happy and excited to work in one of the data science-oriented companies, Chaob Technologies (Carscan).

Chaob Technologies is a carscan company, which uses AI/ML models to scan and detect the damages on the car such as scratches, dents, chipped etc. Carscan is an Augmented Reality app embedded with Artificial Intelligence. The platform assists Motor Insurance companies to expedite their claims, reconditioning, and trade-in processes by up to 90%, saving up to 30% of their cost respectively.

We helped M-sure reduce their claim process time by 80% and reduce their claims cost by 30%, solve governance challenges, and provide a 100% end-to-end digital inspection/claims adoption. The platform assists collision repairers and workshops to expedite their vehicle pre-inspections and upsell

processes respectively increasing their revenue by up to 30% and generating new business. The platform assists Car Rental Agencies and Car Retailers to expedite their vehicle pre-inspection and upsell processes respectively increasing their revenue by 30% and generating new business.

In order to make such an idea possible, we have to train the ML models to identify the defects.

1.1.1 THE WORKING

The Carscan app identifies such defects on the car and provides the customer with repair estimates by keeping track of the previous damages on the car and the latest damages .The user scans for the car with the damage part covered and submits, the ML models used in the application detect the damages and provide the cost and suggest solution to the damage along with nearby shops.

We, the students of BITS, are interested to work on such interesting projects involving ML models, which is our interest, which would help in learning in detail. This would offer us a good first time experience of doing a job. The company we interns will be working in, Chaob technologies located in South Africa is based on the above idea.

1.1.2 MORE ABOUT CHAOB

Chaob Technologies Private Limited is a Private incorporated on 15 June 2021. It is classified as Subsidiary of Foreign Company and is registered at Registrar of Companies, Pune. The company is highly active and has many young data science enthusiasts working throughout the day in a disciplined manner. They

are helpful and friendly with the interns. The company is open all the time for any doubts, any ideas to be discussed. Some of the employees in contact with us are Simran, Avadhut Joshi, Shravan Raj.etc. Neha Verma and Karthik are our head in AI team. Some of the other members of the AI team are Rushali Kodiwal, Tushar Kumar, Krishna, Rajeev Hegde,etc. The interns are divided into 3 teams namely,

- Research team
- AI team
- Product team

We interns are really interested in working on any of the teams since all the teams have something new to offer which is really useful to learn. Some interns work in the research team where they work on finding bots and ways to optimize testing of the web app and code reviews, AI team has to work on annotations and code.etc.

Chapter II

Literature Review

PART I - SuperAnnotate.com: Comprehensive Image Annotation and Dataset Management Platform

1. Introduction

SuperAnnotate.com is a powerful cloud-based platform that revolutionizes the process of image annotation and dataset management for computer vision tasks. This comprehensive solution offers a user-friendly interface, collaboration features, AI-assisted annotation capabilities, and efficient data organization, ensuring the creation of high-quality training datasets for machine learning models.

2. Image Annotation Tools

SuperAnnotate.com provides a wide range of annotation tools that empower users to annotate objects and regions of interest in images. These tools include bounding boxes, polygons, points, lines, and more. By utilizing these tools, users can label objects accurately and precisely, enabling tasks such as object detection, instance segmentation, and image classification.

The platform's annotation tools are intuitive and easy to use, allowing annotators to efficiently label datasets without extensive training or expertise. The interactive interface enhances the annotation experience, ensuring accuracy and reducing annotation time.

3. **Collaboration and Team Management**

SuperAnnotate.com offers robust collaboration and team management features, making it ideal for projects with multiple contributors. The platform allows teams to work together seamlessly on the same project simultaneously.

Users can assign annotation tasks to specific team members, track progress, and communicate through integrated collaboration tools. This streamlines the annotation workflow, enhances productivity, and ensures effective project management.

The platform supports task assignment and annotation review workflows, enabling efficient feedback loops between annotators and supervisors. This collaborative environment fosters accurate and consistent annotations, even for large-scale projects.

4. **AI-Assisted Annotation**

SuperAnnotate.com integrates artificial intelligence (AI) to assist in the annotation process. The platform leverages machine learning algorithms to analyze previously labeled data and provide suggestions for annotations. This AI-assisted annotation significantly speeds up the process, reduces manual effort, and improves productivity.

The AI-assisted annotation feature learns from user interactions, adapting to specific project requirements and annotation patterns. This intelligent system improves over time, continuously enhancing the accuracy and efficiency of annotation tasks.

Users have the flexibility to accept or modify AI-generated suggestions, maintaining full control over the annotation process. This combination of human expertise and AI assistance ensures high-quality annotations while reducing the time required to complete projects.

5. **Data Management and Organization**

SuperAnnotate.com offers robust data management and organization features to streamline the annotation process. Users can easily upload, organize, and manage their image datasets within the platform. The intuitive interface allows for efficient navigation and search, simplifying dataset organization.

The platform supports various file formats, making it compatible with diverse image datasets. It also enables users to track the status of annotation tasks, monitor progress, and ensure timely completion of projects.

SuperAnnotate.com facilitates the export of annotated data in formats compatible with popular machine learning frameworks. This seamless integration allows users to seamlessly integrate annotated data into their preferred training pipelines.

6. **Security and Privacy**

SuperAnnotate.com prioritizes data security and privacy. The platform implements robust security measures, including encryption and secure data storage, to protect user data from unauthorized access.

User authentication and access controls ensure that only authorized individuals have access to sensitive project data. SuperAnnotate.com adheres to industry best practices to maintain the highest standards of data security and privacy.

7. **Conclusion**

SuperAnnotate.com is a game-changing platform for image annotation and dataset management in the field of computer vision. Its comprehensive set of annotation tools, collaboration features, AI-assisted annotation capabilities, and efficient data organization make it an ideal solution for both individual annotators and teams.

By leveraging SuperAnnotate.com, users can streamline the annotation process, improve annotation accuracy, and create high-quality training datasets for machine learning models

. The platform's user-friendly interface, coupled with its collaboration and team management features, ensures efficient project execution and effective communication among team members.

Furthermore, the integration of AI-assisted annotation enhances productivity by reducing manual effort and providing intelligent suggestions. SuperAnnotate.com's robust data management capabilities enable seamless organization and export of annotated data, facilitating its integration into machine learning workflows.

Overall, SuperAnnotate.com empowers users to tackle complex image annotation tasks with ease, accelerating the development of accurate and robust computer vision models.

Chapter III

Data Annotation

3.1 Introduction:

Data annotation is a fundamental step in the field of computer vision, playing a vital role in training and improving machine learning models. In computer

vision tasks, such as object detection, image classification, semantic segmentation, and instance segmentation, the process of data annotation involves labeling or tagging images or videos with relevant information to provide the ground truth for training algorithms.

The significance of data annotation in computer vision cannot be overstated. High-quality annotated data serves as the foundation for supervised learning algorithms, enabling them to recognize patterns, make accurate predictions, and generalize to diverse examples. Without accurate and comprehensive annotations, computer vision models would struggle to understand and interpret visual data effectively.

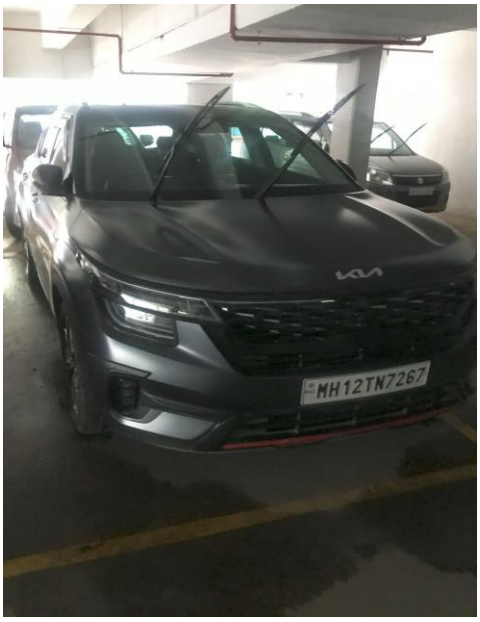
3.2 The Process of Data Annotation

The data annotation process typically involves human annotators who carefully examine each image or frame and apply appropriate labels or annotations. These annotations can include bounding boxes, object categories, keypoints, semantic segmentation masks, or any other relevant information depending on the specific computer vision task.

The data annotation process can be divided into several steps:

1. Dataset Collection: The first step is to gather a diverse and representative dataset that covers various scenarios and conditions relevant to the target application. This dataset can be collected from existing databases, public repositories, or generated through data acquisition methods such as image or video capture. As a part of the internship projects at Carscan, we were required to perform data collection in two phases of the project.

a) Car Images in Low Lighting Conditions:



b) Car Images with Finger Obstacle.

Both of these dataset collection activities were aimed at improving the accuracy of the model, exploring diverse lighting and challenging photographic conditions.

2. Annotation Guidelines: Before starting the annotation process, clear and detailed annotation guidelines need to be established. These guidelines define the labelling criteria, annotation types, and any specific instructions for annotators to follow. Consistency and standardisation are crucial to ensure reliable and uniform annotations across the dataset.

These annotations were shared with us by our mentor Karthik Itharajula, and we were clearly briefed about the difference in types of damage on the car body.
The types of damages include:

a) Dents
b) Scratches
c) Minor Damages
d) Chipped Parts.

Even though we were briefed about minor damages and chipped parts, our project only included annotating the Scracthes and Dents classes. Initially, we faced a few challenges in distinguishing between the various classes of damages, but with practice we managed to learn how to differentiate between these classes.
We were also instructed by our mentor, Karthik Itharajula, as to how the damages are to be demarcated on the images. There was an emphasis on the quality of annotations in terms of shape, number of vertices, and smoothness. This was done to ensure higher accuracy during model training.

3. Annotation Tools: Annotators utilize specialized annotation tools and software to perform the labeling task efficiently. These tools provide a user-friendly interface with features like drawing bounding boxes, creating

polygons, selecting object categories, or placing keypoints. Some popular annotation tools include Superannotate Labelbox, VGG Image Annotator (VIA), and RectLabel.

For the majority of our project, we made use of the SuperAnnotate tool for annotating images. SupperAnnotate is widely used in the industry due to its easy-to-use interface and diverse functionality. The tool allowed us to mark and classify the damages on car images. It also was paired with the system in such a way that images could be sent directly to Quality Check for further processing.

4. Annotation Types: The choice of annotation types depends on the computer vision task at hand. Common annotation types include:

- Bounding Box: Annotators draw rectangular boxes around objects of interest, defining their location and size.

- Object Category: Each annotated object is assigned a specific category label, indicating its class or type.

- Keypoints: Annotators mark specific points or landmarks on objects, enabling the model to understand their structure or pose.

- Semantic Segmentation: Annotators label each pixel in an image, assigning it to a specific object or background class, creating detailed masks.

- Instance Segmentation: Similar to semantic segmentation, but each object instance is labeled separately, enabling the model to distinguish between individual objects.

5. Annotation Quality Control: Maintaining annotation quality is critical for producing reliable and accurate training data. Quality control measures include regular feedback and communication with annotators, inter-annotator agreement checks, and periodic reviews of annotations. In some cases, a subset of the dataset may be annotated by multiple annotators to assess consistency and resolve any discrepancies.

6. Iterative Process: Data annotation is an iterative process that involves continuously improving and refining the annotations. As models are trained on annotated data and evaluated, new insights and challenges may arise, leading to iterative feedback loops with annotators to enhance the annotations and address any limitations.

3.3 Significance of Data Annotation in Computer Vision

The significance of data annotation in computer vision extends beyond model training. Annotated datasets serve as benchmarks for evaluating model performance, comparing different algorithms, and driving advancements in the field. They facilitate research and development in computer vision, enabling researchers to push the boundaries of what machines can perceive and understand.

Moreover, annotated data is crucial for real-world applications of computer vision, such as autonomous driving, surveillance systems, medical imaging, and industrial automation. Accurate and comprehensive annotations allow these systems to make informed decisions, detect objects, recognise patterns, and ensure safety and efficiency in various domains.

However, data annotation can be a time-consuming and resource-intensive process, especially for large-scale datasets. Addressing this challenge, researchers are exploring automated or semi-automated annotation techniques. These methods leverage approaches like active learning, weak supervision, or data synthesis to reduce the human annotation effort while maintaining annotation quality. Active learning focuses annotators' attention on the most informative examples, while weak supervision leverages imperfect or noisy labels to generate annotations at a larger scale.

3.4 Conclusion

In conclusion, data annotation is an essential step in computer vision, enabling machines to understand and interpret visual information accurately. Through the careful labeling and annotation of images or videos, the process provides ground truth data that drives the training and evaluation of machine learning models. The significance of data annotation extends to research, development, and real-world applications, empowering computer vision systems to perceive, analyze, and make informed decisions in various domains.

3.5 SuperAnnotate Tool:

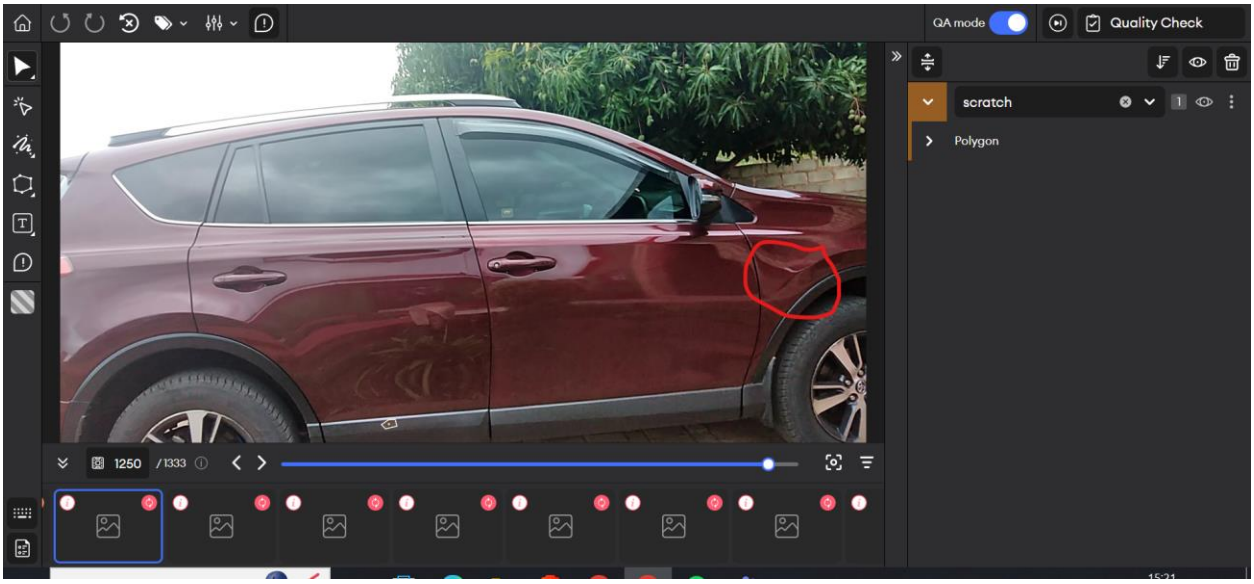
3.5.1 Introduction

SuperAnnotate is a cutting-edge annotation tool that has gained significant popularity in the industry for its exceptional capabilities and user-friendly interface. It is extensively used by professionals in the field of machine learning and computer vision to annotate images and videos for various applications, such as object detection, image segmentation, and keypoint recognition.

At its core, SuperAnnotate provides a comprehensive set of annotation features that cater to diverse annotation needs. It supports a wide range of annotation types, including bounding boxes, polygons, keypoints, and semantic segmentation. This versatility allows users to annotate different types of objects and scenes accurately, providing essential training data for machine learning algorithms.

One of the standout features of SuperAnnotate is its intuitive and user-friendly interface. The tool is designed to optimize annotator productivity, with a clean and organized layout that allows users to efficiently navigate and annotate images and videos. This ease of use makes SuperAnnotate accessible to both experienced annotators and beginners, reducing the learning curve associated with annotation tools.

Collaboration is a crucial aspect of many annotation projects, and SuperAnnotate excels in this area. It offers robust collaborative features that enable multiple annotators to work on the same project simultaneously. Annotations are synchronized in real-time, allowing teams to collaborate seamlessly and expedite the annotation process. Moreover, SuperAnnotate provides tools for annotator communication, such as commenting and feedback, fostering efficient teamwork and ensuring annotation quality.



A snapshot of the SuperAnnotate Tool Interface.

- The left pane has various annotation and marking types.
- The right pane allows the user to select between various classes of damages.
- The top pane has options of undo, redo and skip image and sent to Quality Check.
- Comments can be added to the image as well.
- The lower image pane allows the user to scroll between images, and they have specific markings indicating whether the image has in-progress, skipped, or sent to Quality Check.

As this was our first instance of annotating images and datasets, we were given a few instructions and a guideline document. Initially, we were allotted hourly targets of 20 images per hour. Gradually, the target was increased to annotating and rectifying 125 images per day. Towards the later stages of the project, our efficiency increased, and we were instructed to complete a target of annotating 225 images daily. The gradual increase in targets ensure that the quality of annotation and labelling stayed up to the mark as per spicified standards.

To enhance productivity further, SuperAnnotate incorporates machine learning algorithms into its workflow. The tool includes an auto-annotation feature that leverages pre-trained models to automatically generate annotations, reducing the manual effort required. While auto-annotation is not suitable for all scenarios, it can significantly speed up the annotation process for certain tasks and datasets.

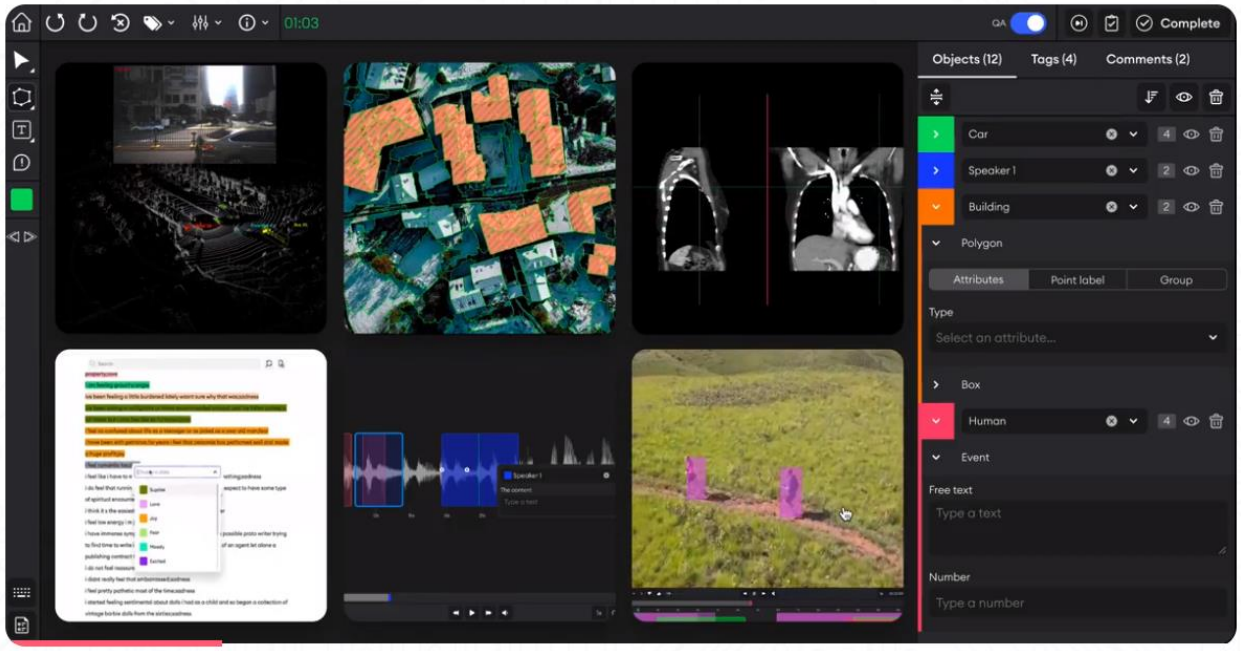
Quality control is paramount in annotation projects, and SuperAnnotate offers a range of quality assurance mechanisms. Annotators can review and correct annotations, verify each other's work, and leave comments and feedback for improved collaboration and error detection. The tool also allows users to measure annotation accuracy metrics, providing insights into the quality of the annotations.

SuperAnnotate goes beyond being just an annotation tool. It offers comprehensive project management capabilities, enabling users to organise and track annotations effectively. Users can assign tasks to specific annotators, monitor progress, and set deadlines. This project management functionality streamlines the annotation workflow, ensuring efficient allocation of resources and timely completion of projects.

The popularity of SuperAnnotate in the industry is a testament to its effectiveness and utility. It is widely used by professionals in various domains, including autonomous vehicles, medical imaging, agriculture, and robotics, among others. The tool's versatility and extensive annotation features make it suitable for a broad range of applications and industries.

SuperAnnotate's integrations with popular machine learning frameworks, such as TensorFlow and PyTorch, facilitate a seamless transition from annotation to training. Annotations can be easily exported in the required formats, enabling users to directly use them for model training and

evaluation. This integration eliminates the need for complex data preprocessing and conversion, saving time and effort for researchers and developers.



3.5.2 Key Features of the SuperAnnotate Tool:

1. Versatile Annotation Types: SuperAnnotate supports a wide range of annotation types, including bounding boxes, polygons, keypoints, and semantic segmentation, allowing for accurate and detailed annotations.

2. Intuitive User Interface: The tool offers an intuitive and user-friendly interface, making it easy for both experienced annotators and beginners to navigate and work efficiently.

3. Collaborative Environment: SuperAnnotate provides robust collaboration features, enabling multiple annotators to work on the same project simultaneously. Real-time synchronization and communication tools facilitate efficient teamwork.

4. Auto-Annotation: SuperAnnotate includes an auto-annotation feature that utilizes machine learning algorithms to automatically generate annotations, reducing manual effort and speeding up the annotation process.

5. Quality Assurance Mechanisms: The tool offers comprehensive quality control features, allowing annotators to review, correct, verify, and leave comments and feedback on annotations. This ensures high annotation accuracy.

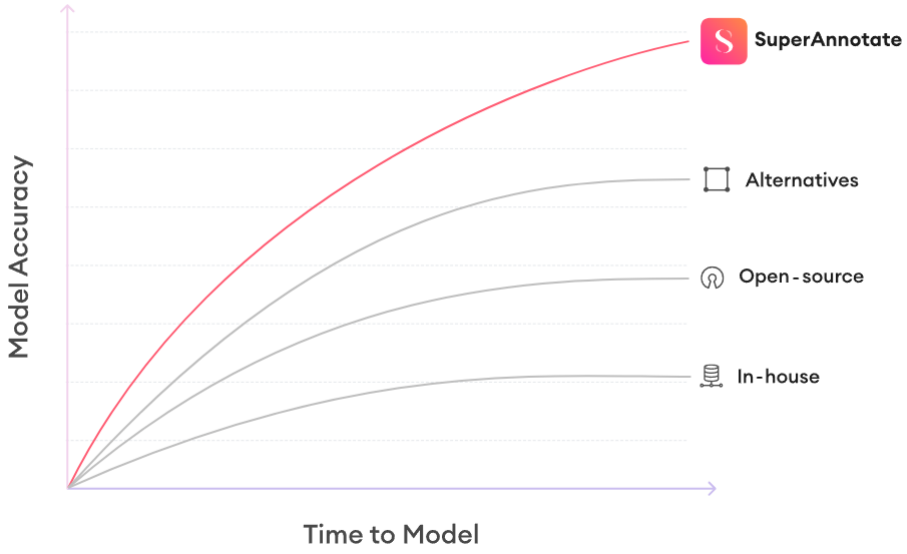
6. Project Management Capabilities: SuperAnnotate provides project management functionalities to organize and track annotations effectively. Users can assign tasks, monitor progress, and set deadlines, streamlining the annotation workflow.

7. Integration with Machine Learning Frameworks: SuperAnnotate seamlessly integrates with popular machine learning frameworks, such as TensorFlow and PyTorch. Annotations can be exported in the required formats, facilitating a smooth transition from annotation to model training.

8. Data Augmentation: SuperAnnotate includes data augmentation capabilities, allowing users to generate synthetic data variations from annotated images. This feature enhances model performance and generalization.

9. Image and Video Annotation: SuperAnnotate supports both image and video annotation, enabling annotators to annotate frames within videos for tasks like action recognition or tracking.

10. Versatile Industry Applications: SuperAnnotate is widely used across various industries, including autonomous vehicles, medical imaging, agriculture, and robotics, among others. Its extensive annotation capabilities make it suitable for diverse application domains.



In summary, SuperAnnotate is a powerful and widely adopted annotation tool that significantly simplifies and streamlines the process of annotating images and videos for machine learning and computer vision tasks. Its user-friendly interface, collaborative features, and emphasis on quality assurance make it a top choice for professionals in the industry. With its comprehensive annotation capabilities, project management functionalities, and integration with machine learning frameworks, SuperAnnotate empowers researchers and developers to accelerate their projects and achieve accurate and reliable results.

Chapter IV

Vehicle License Disk Object Detection

4.1: Optimizing Object Detection with YOLOv5

YOLOv5 is a computer vision model for object detection and localization. It is the latest version of the YOLO (You Only Look Once) series of real-time object detection algorithms. YOLO models are known for their speed and accuracy in detecting objects in images and videos.

YOLOv5 was developed by Ultralytics and released in June 2020. It builds upon the previous versions (YOLOv1, YOLOv2, YOLOv3) by introducing several improvements and optimizations. The key features of YOLOv5 include:

- Improved architecture: YOLOv5 introduces a new architecture that is more streamlined and efficient. It replaces the Darknet architecture used in previous versions with a more modular and flexible design.
- Various model sizes: YOLOv5 offers different sizes, such as YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. These models vary in terms of size and accuracy, allowing users to choose the trade-off that best suits their needs.
- Efficient inference: YOLOv5 leverages advanced techniques such as model pruning, quantization, and automatic mixed precision to optimize inference speed and reduce memory consumption. This enables real-time or near-real-time object detection on both CPUs and GPUs.
- Training pipeline: YOLOv5 provides an end-to-end training pipeline that simplifies the process of training custom object detection models. It supports various data augmentations, transfer learning, and

hyperparameter tuning, making it easier for researchers and developers to adapt the model to their specific use cases.

- Versatile usage: YOLOv5 is designed to work well in various applications, including autonomous driving, surveillance systems, robotics, and more. Its ability to detect and localize objects in real time makes it suitable for both static images and video streams.

It's worth noting that while YOLOv5 has gained popularity for its performance and ease of use, it is an independent project developed by Ultralytics and is not an official continuation of the original YOLO series by Joseph Redmon.

YOLOv5 has gained recognition for its impressive speed and precision, finding applications in various fields such as healthcare, security surveillance, and self-driving cars. The Ultralytics team has dedicated their efforts to enhancing this model since 2015, resulting in multiple subsequent versions. In this article, we will delve into an overview of YOLOv5, focusing on its fifth iteration.

4.1.1. High-level architecture for single-stage object detectors

YOLO is a single-stage object detector architecture that is comprised of three main components: A backbone, a Neck, and a Head to make dense predictions

Model Backbone: The backbone refers to a pre-trained network employed to extract high-quality feature representations from images. This process involves reducing the spatial resolution while enhancing the feature (channel) resolution.

Model Neck: The model neck is responsible for extracting feature pyramids, enabling the model to effectively handle objects of varying sizes and scales.

Model Head: The model head executes the final operations, including applying anchor boxes to feature maps and generating the ultimate output, which includes classes, objectness scores, and bounding boxes.

4.1.2. Cross-Stage Partial Network

In YOLOv5, the CSP-Darknet P53 serves as the backbone architecture. CSP-Darknet53 is essentially the convolutional network Darknet53, which was originally used as the backbone for YOLOv3. However, the authors of YOLOv5 have implemented the Cross Stage Partial (CSP) network strategy on top of Darknet53 to enhance its performance and effectiveness.

YOLO is a deep neural network that incorporates residual and dense blocks to facilitate the smooth flow of information to deeper layers and address the issue of vanishing gradients. However, the utilization of dense and residual blocks can introduce redundant gradients. To mitigate this problem, the Cross Stage Partial (CSPNet) technique is employed, which effectively truncates the gradient flow and helps overcome the issue of redundant gradients. By adopting the CSPNet strategy, YOLOv5 leverages the benefits of DenseNet's feature reuse characteristics while effectively reducing redundant gradient information. The application of the CSPNet strategy involves partitioning the feature map of the base layer into two sections and subsequently merging them using a cross-stage hierarchy, as depicted in the accompanying figure. This approach offers significant advantages to YOLOv5, including a reduction in the number of parameters and computational requirements (FLOPS). As a result, the inference speed of the model, a crucial factor in real-time object detection, is increased.

4.1.3. Neck of YOLOv5

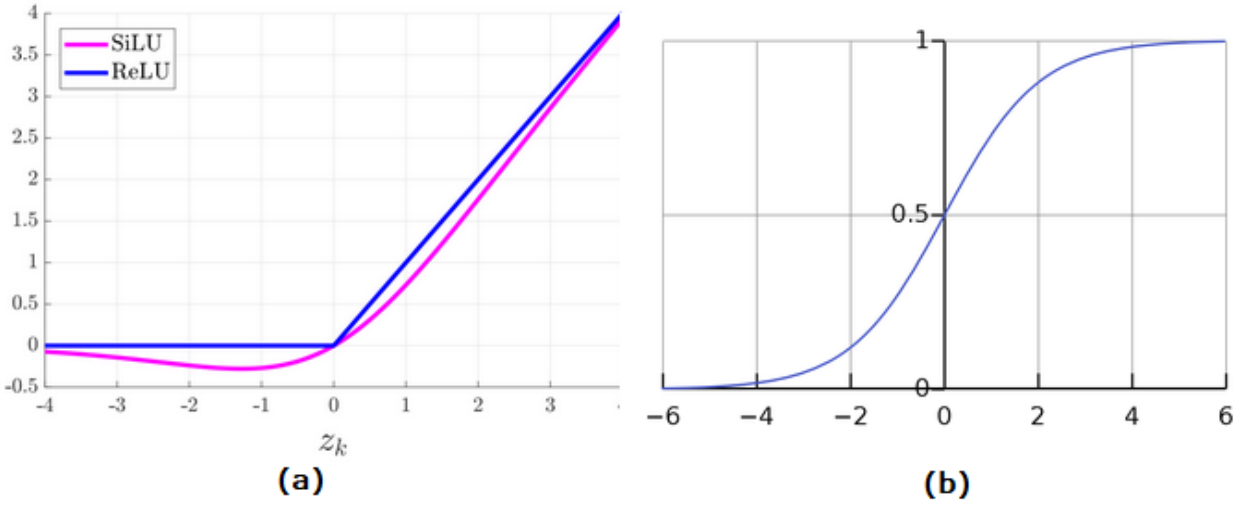
YOLOv5 introduces two significant changes to the model neck. Firstly, a variant of Spatial Pyramid Pooling (SPP) has been incorporated, and secondly, the Path Aggregation Network (PANet) has been modified by integrating the BottleNeckCSP into its architecture. PANet, a feature pyramid network, was

previously used in YOLOv4 to enhance information flow and assist in accurate pixel localization for mask prediction tasks.
The SPP block performs information aggregation from the input and produces an output of fixed length. This feature provides the advantage of significantly increasing the receptive field and isolating the most relevant contextual features without compromising network speed.

4.1.4. Head of the network

YOLOv5 utilizes the same head structure as YOLOv3 and YOLOv4. This head consists of three convolutional layers responsible for predicting the bounding box location parameters (x, y, height, width), object scores, and object classes.

4.1.5. Activation Function



The activation functions used in YOLOv5 can be visualized as (a) SiLU Function Graph: SiLU, short for Sigmoid Linear Unit, is also known as the swish activation function. It is utilized in conjunction with the convolution

operations within the hidden layers of YOLOv5.(b) Sigmoid Function Graph: The Sigmoid activation function is employed with the convolution operations in the output layer of YOLOv5.

Please note that the exact shape of the activation function graphs cannot be displayed in this text-based format, but the descriptions provided give an understanding of their usage and characteristics in YOLOv5.

4.2 Use of Makesense.ai

Makesense.ai is an online tool that allows users to label photos without any complicated installation process. By simply visiting the website, users can start utilizing the tool right away. It is designed to be cross-platform, meaning it works seamlessly on any operating system. The tool is particularly useful for small computer vision deep learning projects, as it simplifies and accelerates the process of dataset preparation. After labeling the photos, users can download the prepared labels in multiple supported formats. The application is built using TypeScript and relies on the React/Redux framework.

makesense.ai aims to significantly reduce the time spent on photo labeling. To achieve this, we employ various AI models that provide recommendations and automate repetitive and tedious tasks.
It utilizes the SSD model pre-trained on the COCO dataset, which assists in drawing bounding boxes on photos and, in some cases, suggesting labels. Additionally, the PoseNet model, a vision model, estimates the pose of a person in an image or video by identifying key body joints.

4.3 Roboflow

Roboflow offers a comprehensive suite of functionalities to support computer vision workflows. It provides tools for efficient dataset management, allowing users to upload, organize, and version datasets. Roboflow also offers powerful data preprocessing capabilities, enabling users to apply transformations and augmentations to improve data quality. With its annotation tools, users can easily label and annotate images with bounding boxes, polygons, key points, and segmentation masks. The platform supports model training using popular deep learning frameworks like TensorFlow and PyTorch, along with model deployment through APIs and integrations with deployment platforms. Roboflow provides evaluation metrics for assessing model performance, facilitates collaboration and sharing among team members, and seamlessly integrates with existing workflows and platforms. Overall, Roboflow streamlines the entire computer vision pipeline, from dataset management and preprocessing to annotation, training, deployment, and collaboration.

4.3.1 Dataset annotation

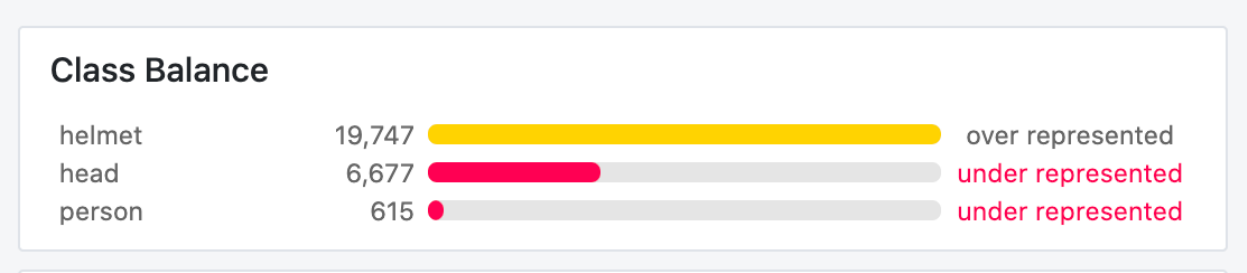
We first upload the dataset that requires annotation. Roboflow supports various formats for both images and videos, including JPG, PNG, BMP, MOV, and MP4. For annotation, you can choose from multiple formats specifically designed for object detection tasks, such as JSON, XML, CSV, and TXT. Once the dataset is annotated, you have the flexibility to export it in the format of your choice.

4.3.2 Dataset Preprocessing

Before model fitting, dataset preprocessing plays a crucial role in achieving effective training, validation, and testing results while mitigating the risk of overfitting. This tool provides various preprocessing capabilities, including resizing, converting to grayscale, and augmentation, to enhance the dataset quality and suitability for model training.

4.3.3 Dataset Health check:

With this tool, you can visualize the health of your dataset, which can have a significant impact on the accuracy of your model. We can obtain inferences on various dataset parameters such as the number of images and annotations and class Balancing.



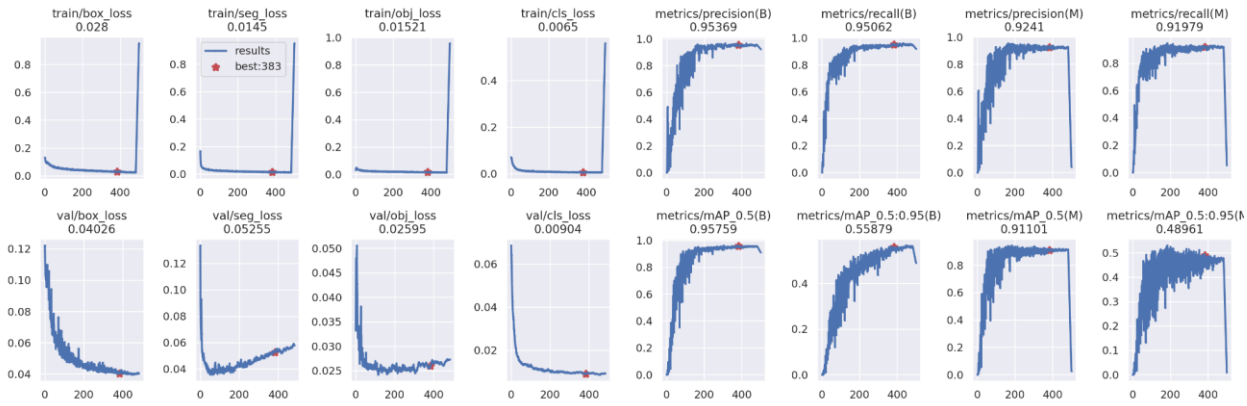
In our project focused on detecting license plate details on cars in Africa, we divided our team into two groups: Team - 1 (comprised of Kamalesh Ram, Darsh Doshi, Raj Patel, and Raguram Venkatesan) and Team - 2 (comprised of Nischith P, Sandeep YSP, and Daksh Shah). Team - 1 was assigned the task of developing a model with high accuracy to predict the details on a license plate image when the QR code cannot be captured properly.

Initially, we worked with a dataset of 300 images. Out of these, 150 were used for prediction, 99 for training, and 53 for validation. Since annotation was not required for the prediction images, we focused on annotating the 150 images using makesense.ai. We classified the different objects into 11 classes and annotated them accordingly. To facilitate the annotation process, we divided the 152 images into two sets of 38 images each. Subsequently, we converted the annotations from the COCO JSON format to a YOLOV5 format using Roboflow. Roboflow also generated a YAML file based on the COCO JSON file. Using the open-source code from https://github.com/ultralytics/yolov5 on Google Collab, we trained the model.

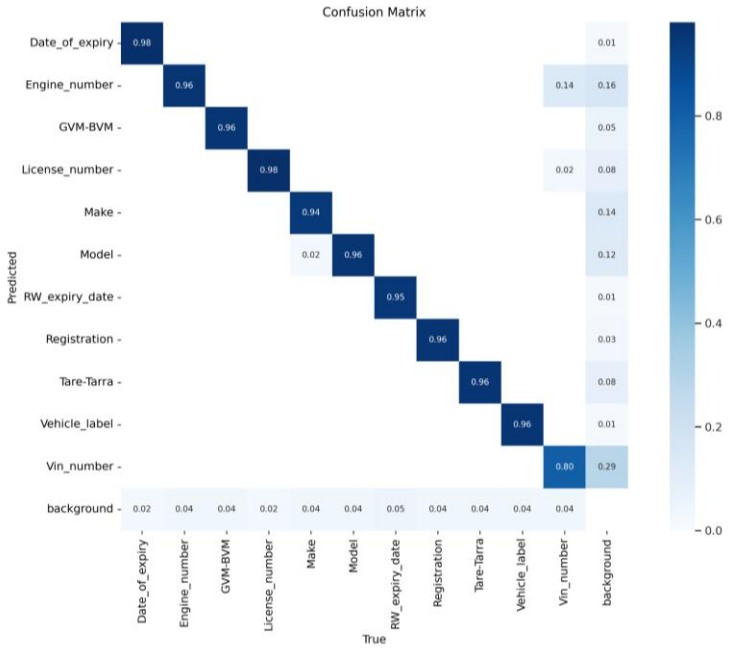
To improve accuracy, we adjusted the hyperparameters, eventually settling on 500 epochs as the optimal value. The resulting best.pt file, containing the best-weighted parameters, was used for predicting the remaining 150 images. Validation of the model indicated an accuracy of 92%.

In the second iteration of the License Disk model, our team aimed to enhance the accuracy of the previously developed model. For this purpose, we trained the model again, this time utilizing a larger dataset comprising 3000 images. The predictions for these images were generated in the yolov5 annotations format using a pre-trained model. To upload the images to makesense.ai, we converted the TXT file annotations to the COCO JSON format. Furthermore, we meticulously reviewed the annotations to identify any incorrect predictions, reannotating them for missing fields and making necessary corrections. Finally, we trained the model using the newly acquired dataset.

4.3.4 Results:



These confidence curves represent the accuracy progression of the model over 500 epochs, reaching 92% accuracy.



Conclusion

Initially, as interns, we were uncertain about the work, the project, and what the next two months would hold for us. However, our enthusiasm for the company and the exciting project ahead was unwavering. Despite the initial uncertainty, we were eager to dive in and make the most of this valuable opportunity at Chaob Technologies. Right from the beginning, the supportive and welcoming environment made us feel at home. The mentorship and guidance provided by the experienced professionals at Chaob Technologies were instrumental in helping us find our footing in the work domain.

Total number of students alloted to this company is 10. We were divided into 3 teams namely, AI team, Research team and Product team. Only Divyaroop was working in the research team. Raguram, Kamalesh Ram, Darsh Doshi, Raj Patel, Nischith P, YSP Sandeep, Daksh shah were a part of the AI team. Shubham Birla and Shree Ganesh worked in the product team.

AI team

We were introduced to amazing mentors like Neha, Kartik, Siddant and Pushpak. They guided us throughout our internship. They were always available to solve any doubts we had or discuss any issues or ideas.

We worked on 2 projects mainly-

1. Dent and Scratch
2. Licence disk field identification

At first all the interns worked on the same project dent and scratch under Kartik. We were asked to annotate
all the images of the car alloted by Karthik. Annotation involves identification of dents and scratches on the car and marking and labelling them. To execute this task, we did this on the website called superannotate.com .

Karthik showed us how to work and how to identify dents and scratches with regular meetings, which was quite informative and helpful. The working hours were 11am to 7pm, we were supposed to complete our work within the working hours. We had a daily online meeting for 30 min in the morning at 11am. The daily target(number of images to be annotated) gradually increased from 25 images to 225 images in two months.

In July the interns were divided into two teams-

Team 1 - Raguram, Kamalesh Ram, Darsh Doshi, Raj Patel

Team 2 - Nischith P, YSP Sandeep, Daksh Shah

Team 1 had a change in the project for a week. They were asked to work under Siddant on licence disk field management. Team 2 continued to work in dent and scratch project under Karthik. In the licence disk field management project, our team of interns embarked on a fascinating endeavour to accurately label the various components present on a licence disk using the advanced makesense.ai platform. Through meticulous labelling, we aimed to create a reliable dataset that would serve as the foundation for training powerful Machine Learning (ML) models. By

employing the state-of-the-art yolov5 framework, we harnessed the potential of deep learning to predict and validate the labelled data effectively. This enabled us to build robust ML models capable of recognizing and marking the labelled components on fresh, previously unseen data. Throughout the project, collaboration and teamwork were key. Our interns worked cohesively, exchanging ideas and insights, which facilitated a comprehensive understanding of the challenges and opportunities in the licence disk field management domain. The collective efforts of our team ensured that the dataset was well-curated, and the models were fine-tuned to achieve impressive levels of accuracy and precision. After a week team 2 were called to work on this project and team 1 were allotted the dent n scratch project. Even team 2 worked on the same thing for a week.

During our internship, we were encouraged to maintain a daily diary provided by our college, where we diligently recorded our experiences and the tasks we accomplished. This initiative proved to be immensely beneficial as it helped us keep track of our work, reflect on our progress, and ensure we were on the right path to achieving our goals. Throughout the internship, we had the privilege of being mentored by two exceptional individuals, Ranjith Sir and Farhah Masroor from BITS Goa. We are sincerely grateful for their unwavering support and guidance throughout the entire journey. Their expertise and experience were invaluable as they constantly provided us with valuable insights, tips, and tricks on how to excel in our workplace and make the most of our internship opportunity. One of the standout aspects of their mentorship was their dedication to conducting timely and well-structured evaluations. They encouraged us to reach out to them whenever we encountered challenges or needed assistance.

We do believe that this internship was a valuable experience, since we learnt many things out of it. This internship taught us teamwork, time management, distribution of work, leadership skills, etc. During our internship, many of us were stepping into the professional world for the first time, and as expected,

we encountered numerous challenges and obstacles along the way. However, this invaluable experience also presented us with opportunities to learn and grow, equipping us with essential skills to navigate the complexities of a professional environment. The initial phase of the internship was undoubtedly challenging, as we grappled with new responsibilities and expectations. Over time, we learned to proactively tackle problems, seeking innovative solutions and leveraging the expertise of our mentors and colleagues. The internship also instilled a strong sense of discipline within us. The requirement to maintain a daily diary not only facilitated our work progress tracking but also encouraged a structured approach to our tasks. Meeting deadlines and adhering to timelines became second nature, enhancing our time management skills.

References

Yolov5 and Roboflow

- https://github.com/ultralytics/yolov5
- https://iq.opengenus.org/yolov5/
- https://medium.com/red-buffer/roboflow-d4e8c4b52515
- https://colab.research.google.com/github/ultralytics/yolov5/blob/master/segment/tutorial.ipynb#scrollTo=zR9ZbuQCH7FX

Annotations

- https://app.superannotate.com/
- https://www.makesense.ai/