co-one.co

DATA ANNOTATIONS FOR BETTER AI

We are driving the AI industry forward with our cutting-edge AI and crowdsource-based data platform, delivering unparalleled speed and accuracy in data annotations.

co-one.co

UPDATES ON AI

AI is growing

42%

every year.

(Grandview Research)

70% of companies in

2030

will increase their revenues by using AI.

(McKinsey)

AI Market will be valued at

$327B

in 2025.

(Grandview Research)

Generative AI

10M

users

BUT THERE IS A THING OVERLOOKED

co-one.co

NEEDS STILL MANUAL HUMAN WORK

co-one.co

VS

A human learns in x10

AI learns in x10.000

co-one.co

WHAT IS ANNOTATION?

Annotation is

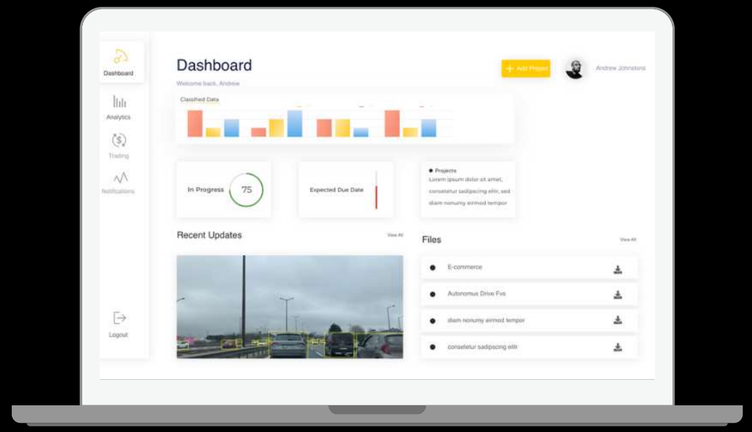
marking data

with human perception.

PRODUCTS

co-one.co

ANNOTATION MANAGEMENT DASHBOARD



FOR AI TEAMS

dashboard.co-one.co

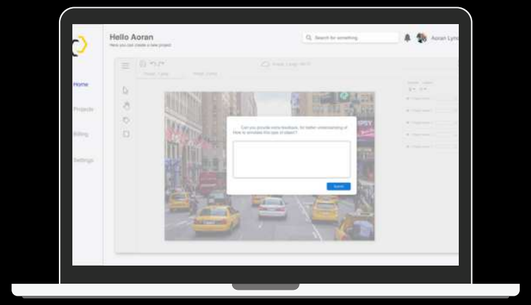
CROWDSOURCE ANNOTATION WORKFLOW

FOR ANNOTATORS

kovan.app

HOW CO-ONE WORKS?

co-one.co

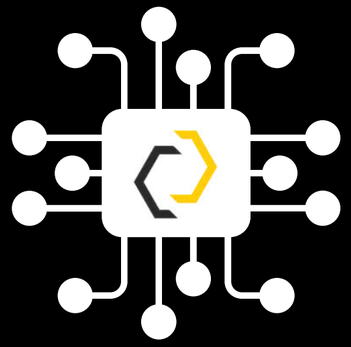


1

UPLOAD THE DATA

2

AI PRE-LABELS THE DATA

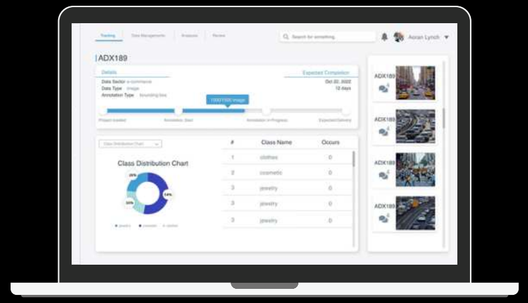


3

DISTRIBUTE ANNOTATORS

4

CROSS VALIDATION



5

TRACK THROUGH DASHBOARD

co-one.co

WHAT IS OUR SECRET SAUCE

AI & Human

Hybrid Annotation

Framework

Crowdsource

Control Mechanism

Super Easy Full

Mobile Annotation App

co-one.co

- The Importance of Data Annotation in AI
- Data Annotation Challenges
- Data Annotation Best Practices
- Case Study Results

(Gebru, T., et al. (2018). Datasheets for Datasets. arXiv preprint arXiv:1803.09010.)
(Lipton, Z. C. (2018). The mythos of model interpretability. arXiv preprint arXiv:1606.03490.)

co-one.co

THE IMPORTANCE OF DATA ANNOTATION IN AI

Raw data has limitations

Incomplete
Inconsistent
Unstructured

Annotation quality affects AI model accuracy

Annotated data provides context and semantics that help AI models understand and interpret data accurately.

Inaccurate or biased annotations can lead to incorrect predictions and unreliable results. Therefore, it is essential to ensure high-quality annotations in the data annotation process.

Human annotation is still important

Human annotation is still necessary for tasks that require human judgement, such as image recognition and natural language processing.

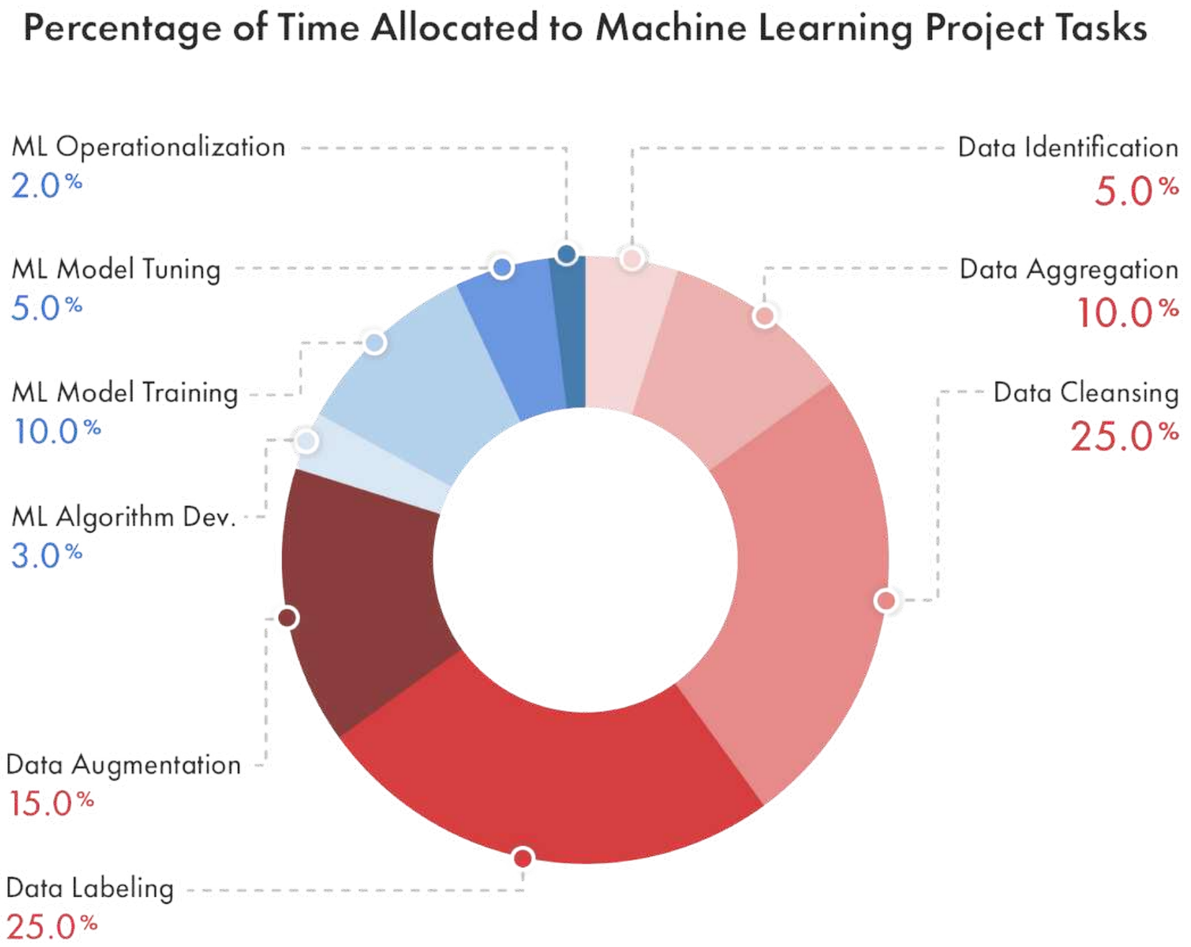
Accuracy and nuance that is difficult to achieve with AI-based tools alone

(Gebru, T., et al. (2018). Datasheets for Datasets. arXiv preprint arXiv:1803.09010.)
(Lipton, Z. C. (2018). The mythos of model interpretability. arXiv preprint arXiv:1606.03490.)

THE IMPORTANCE OF DATA ANNOTATION IN AI

co-one.co

Raw Data

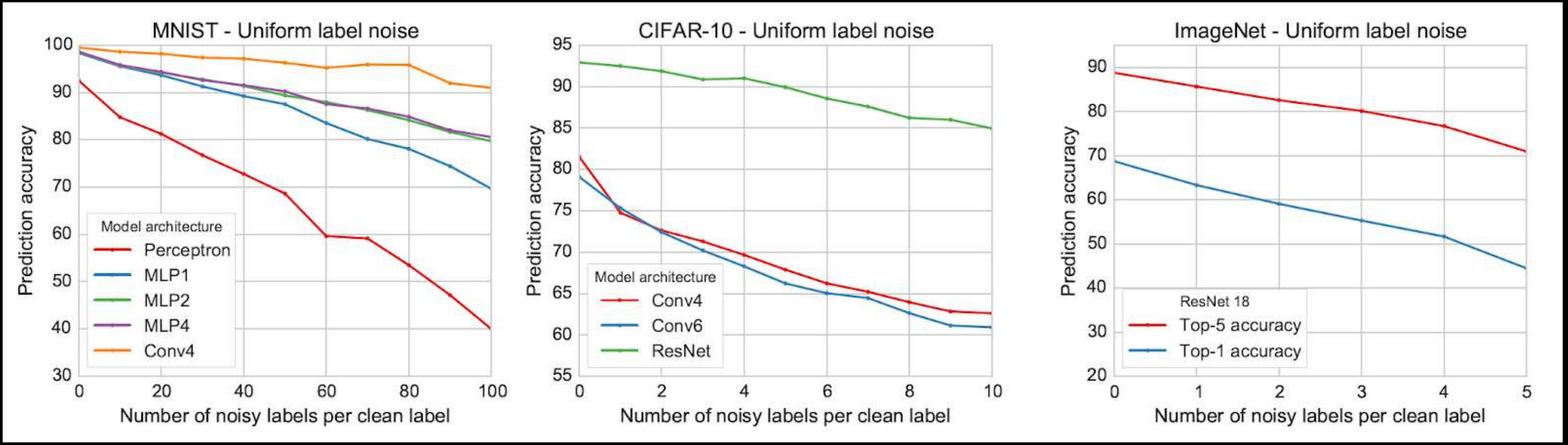


(Source: Cognilytica)

co-one.co

THE IMPORTANCE OF DATA ANNOTATION IN AI

Annotation quality vs. model accuracy



(https://www.superannotate.com/blog/how-to-detect-mislabeled-annotations)

co-one.co

DATA ANNOTATION CHALLENGES

Ambiguity in the data

Sometimes the data being annotated can be ambiguous or difficult to understand, leading to inconsistency in the annotations

Quality control

Ensuring the quality and consistency of annotations can be challenging.

Scalability

As datasets grow larger, it can become difficult to manage the annotation process in an efficient and effective manner.

AMBIGUITY

co-one.co

Uncertain Images

Figure 4.4a : Uncertain Data Point.

Figure 4.4b : Uncertain Data Point.

Kayhan A., (2021), ACTIVE LEARNING BASED HUMAN IN THE LOOP DEEP OBJECT DETECTION FOR SCALABLE DATA ANNOTATION

AMBIGUITY

co-one.co

Uncertain Images - Earthquake Collapsed Buildings

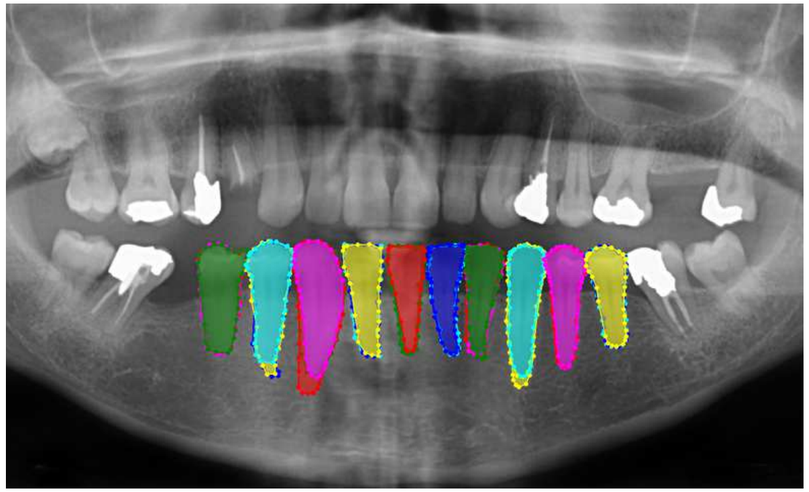
Kayhan A., (2021), ACTIVE LEARNING BASED HUMAN IN THE LOOP DEEP OBJECT DETECTION FOR SCALABLE DATA ANNOTATION

QUALITY CONTROL

co-one.co

Annotation Quality

False Annotations



Annotator Confidence

- Annotator Training
  - Active Confidence Control
  - Long Term Confidence

Spam Annotation

- Is annotator trustable?
  - Malicious User
  - Intentional False Annotations

SCALABILITY

co-one.co

Requirement for larger datasets

Complex Annotation Requirements

- Human annotation can be time-consuming and expensive.
- Models trained on smaller datasets may not perform as well on larger datasets due to issues such as overfitting or bias
- Inadequate resources, such as a shortage of trained annotators, can also slow down the annotation process and make it difficult to handle large-scale projects.

Kayhan A., (2021), ACTIVE LEARNING BASED HUMAN IN THE LOOP DEEP OBJECT DETECTION FOR SCALABLE DATA ANNOTATION

co-one.co

DATA ANNOTATION TECHNIQUES & BEST PRACTICES

Data Type and Annotation Tools

Different data types require different annotation types.

- Images
- Videos
- Text
- LIDAR
- Audio

Creating clear annotation guidelines

Clear and concise guidelines help annotators understand what needs to be annotated and how. Guidelines should include detailed instructions, examples, and definitions of terms.

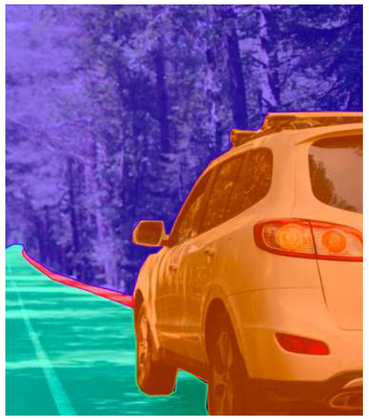
Quality Control Measures

Quality control measures, such as inter-annotator agreement and spot-checking, help ensure that annotations are accurate and consistent across annotators.

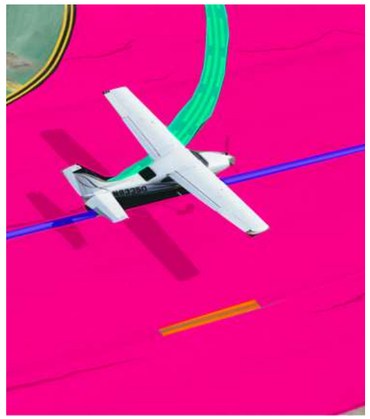
ANNOTATION TYPES

co-one.co

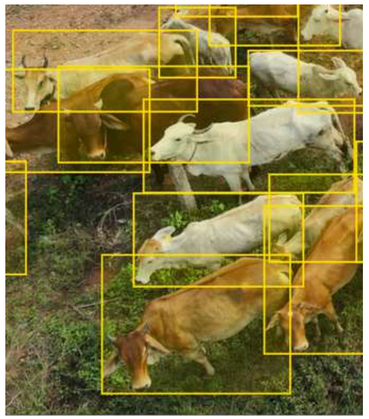
ADAS - Semantic



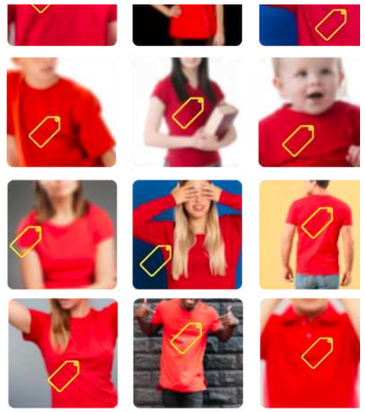
Aerospace - Semantic



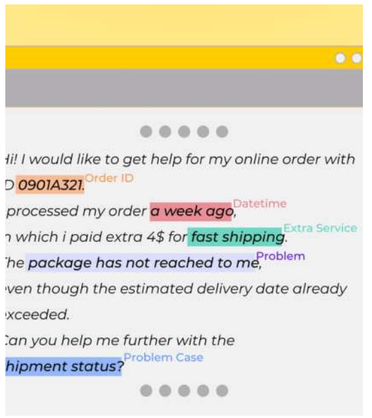
Agriculture- Bounding Boxes



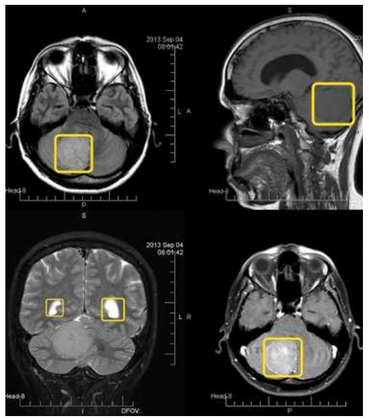
E-commerce - Data Enrichment



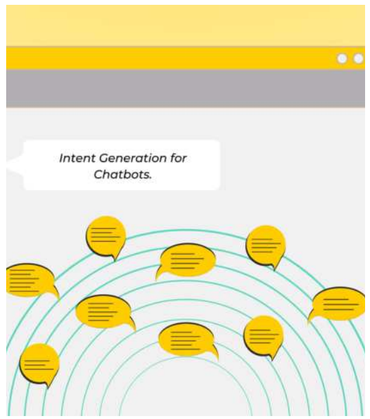
NLP - Named Entities



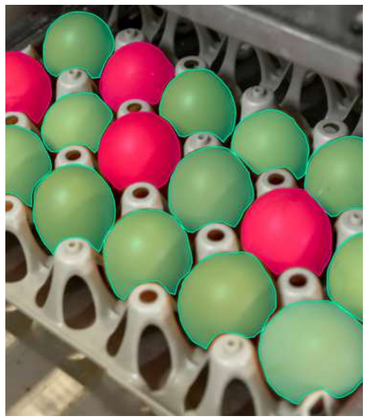
Health - Bounding Boxes



Chatbots - Intent Generation



Manufacturing- Polygons



Retail - Bounding Boxes

Security & Surveillanve

CLEAR GUIDELINES

co-one.co

Annotation File Format

.JSON

Annotation Format

.COCO

Data Format

.PNG

Annotation Type

.POLYGONS

3.1.1. DETAILS OF OBJECTS

a.wrecked_buildings

Brief: They are wrecked buildings.

Example Annotations:

CLEAR GUIDELINES

co-one.co

c.wrecked_facility
Brief: Wrecked facilities. It covers wrecked structures outside the building.
For example, facility, factory, mosque etc.

d.undamaged_facility
Brief: They are undamaged facilities. It covers undamaged structures
outside the building. For example, facility, factory, mosque etc.

b.undamaged_buildings

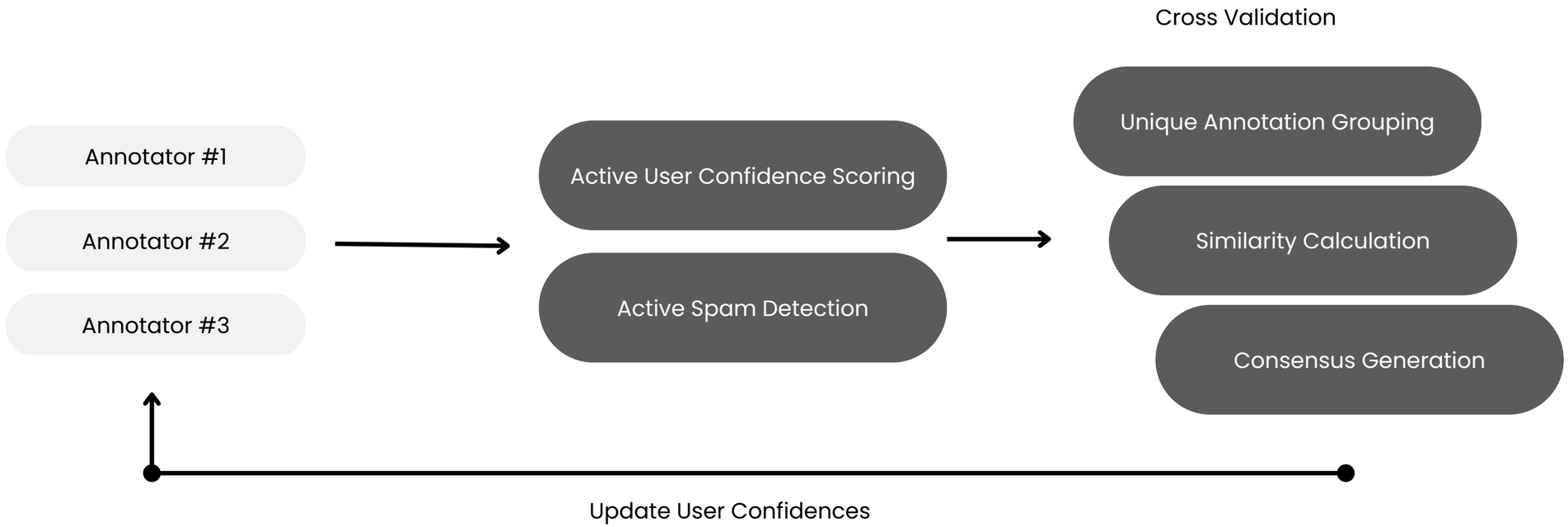
Brief: These are buildings that have not been damaged by the earthquake.

Example Annotations:

QUALITY CONTROL

co-one.co

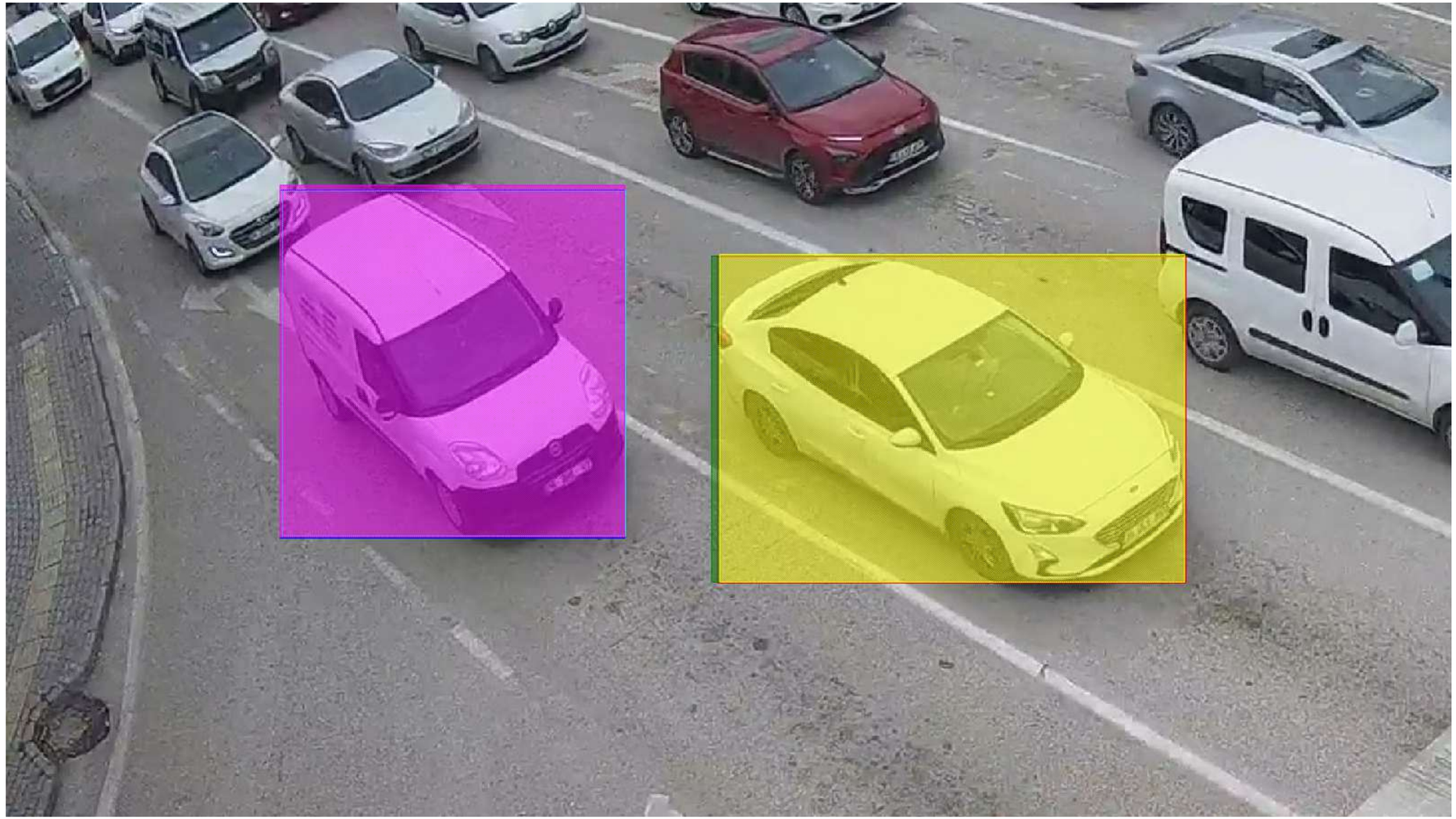
Annotation Validation
Inner Agreement Scoring



QUALITY CONTROL

co-one.co

Annotation Validation Inner Agreement Scoring



co-one.co

DATA ANNOTATION TECHNIQUES & BEST PRACTICES

Diversity and representativeness of data

The annotated data should be diverse and representative of the real-world scenarios that the AI model is expected to encounter.

Privacy and Security of Data

The annotated data should be kept secure and confidential, and the privacy of individuals in the data should be protected.

- GDPR
- Data anonymization
- Blurring

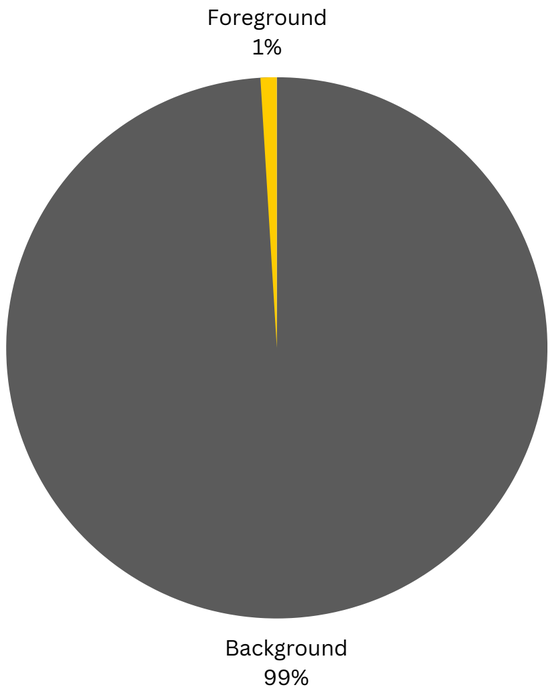
Hybrid, Human in the Loop Systems

AI and Human annotation can be employed together within active learning framework.

DIVERSITY AND BALANCED DATASET

co-one.co

Imbalanced Dataset Problem



Sampling can be applied

https://neptune.ai/blog/imbalanced-data-in-object-detection-computer-vision
https://arxiv.org/pdf/1909.00169.pdf

DIVERSITY AND BALANCED DATASET

co-one.co

Dataset Diversity

Different than model diversification techniques.

Can be achieved during **data collection** and **data annotation** stages.

IDEALLY -> Diverse, Balanced and Edge Case Covering.

- Diverse age
- Gender
- Race
- Ethnicity
- Backgrounds

HYBRID ANNOTATION

co-one.co

Active Learning - Human In The Loop

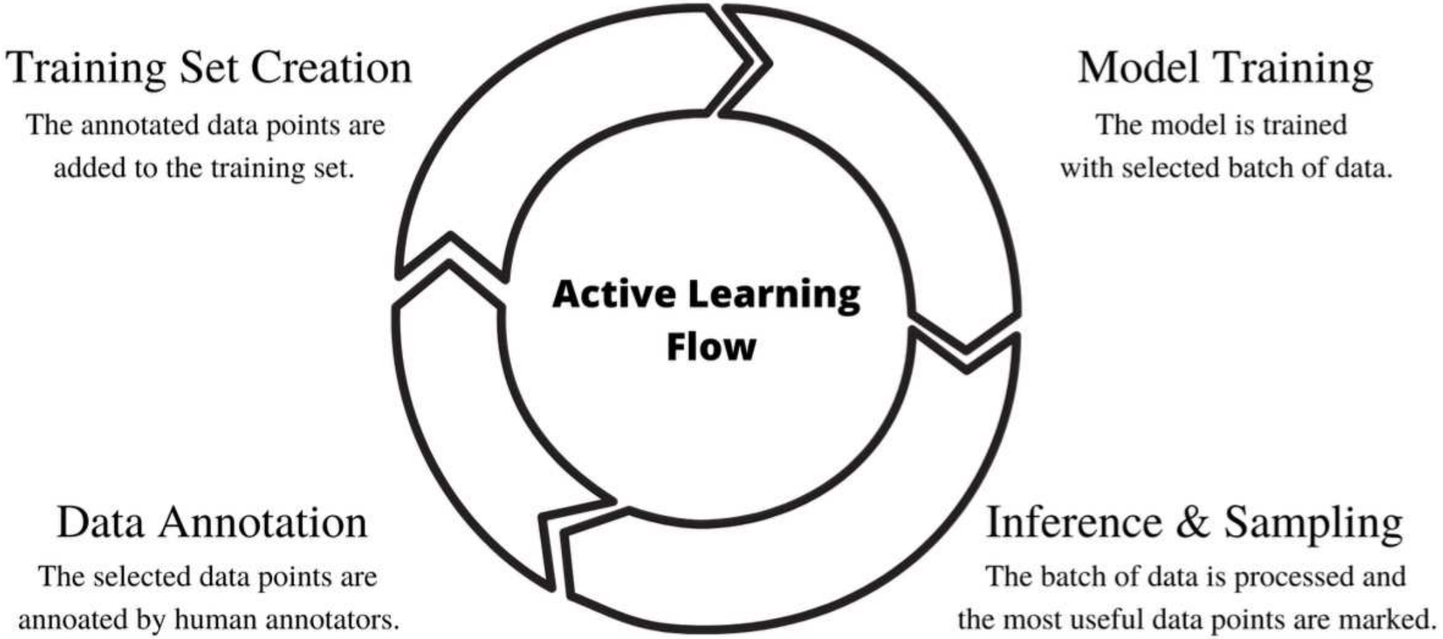


Figure 3.1 : Model Architecture

Kayhan A., (2021), ACTIVE LEARNING BASED HUMAN IN THE LOOP DEEP OBJECT DETECTION FOR SCALABLE DATA ANNOTATION

co-one.co

CASE STUDIES & RESULTS

CYBELLE MEDIKAL

AI based hair detection auto laser epilation

310K
annotation is provided

15%
increase in dark hair detection performance

18%
increase in thin hair detection performance

co-one.co

CASE STUDIES & RESULTS

ALTERNA CX

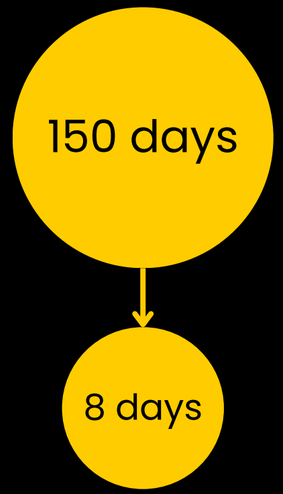
AI based experience management software

220K

text classification is provided

X18

Faster data annotation process



LESS COST

Dedicated team is dismissed

co-one.co

CASE STUDIES & RESULTS

ETIYA

AI based digital business services, chatbots

~10K

text generation is provided

13%

increase in model performance

co-one.co

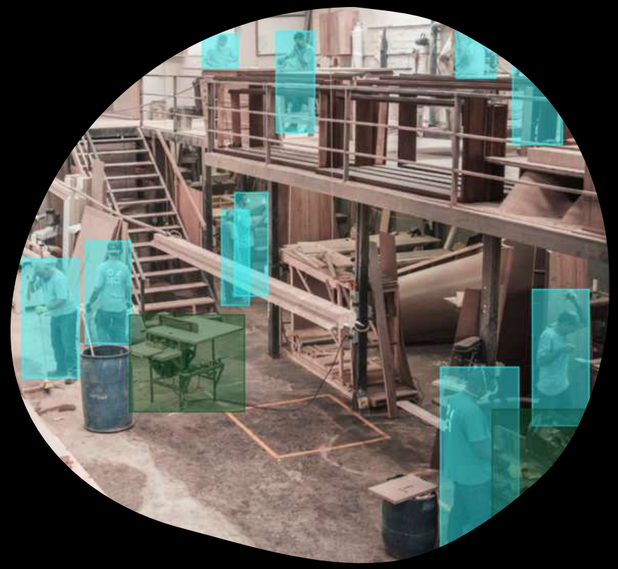
BETTER DATA LEADS BETTER AI

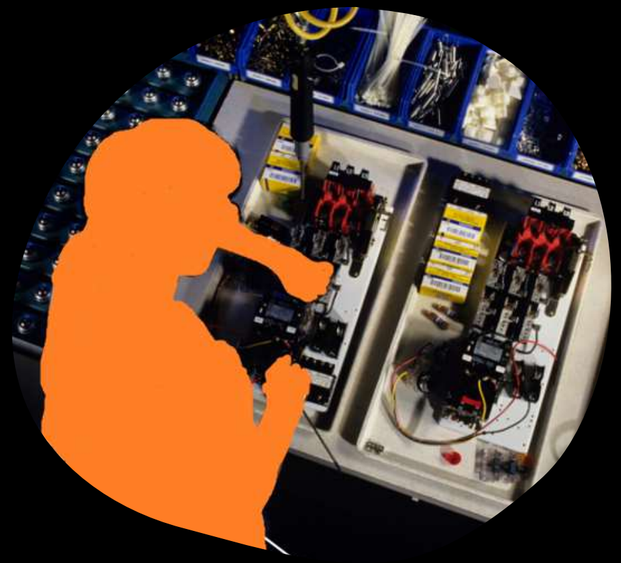
BETTER
self-drivig cars

SAFER
workplaces

FASTER
RPA

BETTER
shopping experinece





co-one.co

OUR CUSTOMERS

OUR SUPPORTERS

co-one.co

WHO ARE WE?

A. Arman KAYHAN

AI & Tech. Management (Master)
Aerospace Engineering

4+ Years Tech Team Lead
2+ Years AI Developer
4+ Years Product Management

Mert MENEKSE

Engineer Management (Master)
Aerospace Engineering

2+ Years Project Team Leadership
2+ Years Sales Engineer
4+ Years Business Development
2+ Innovation Management

INVESTORS

leap. into the future

MAXIS

TechOne

LIMA VENTURES

DOMiNO VENTURES

TRANGELS

ADVISORS

Erdem Yörük, Phd
CTO
vispera

Ömer Orkun Düztaş
CEO
büyütech

Nazım Kemal Üre, Phd
Vice President
İTÜ

Thank you

A. Arman KAYHAN
arman.kayhan@co-one.co

Mert MENEKSE
mert.menekse@co-one.co

info@co-one.co
co-one.co
Co-one
@cooneAI
@coone.co

Tallin, Estonia
Istanbul, Turkey