

“Domain-Specific Data Expertise”

2025

AI Research and Development Company

**CODEVISION Inc.**



# With Core Expertise in Data & AI

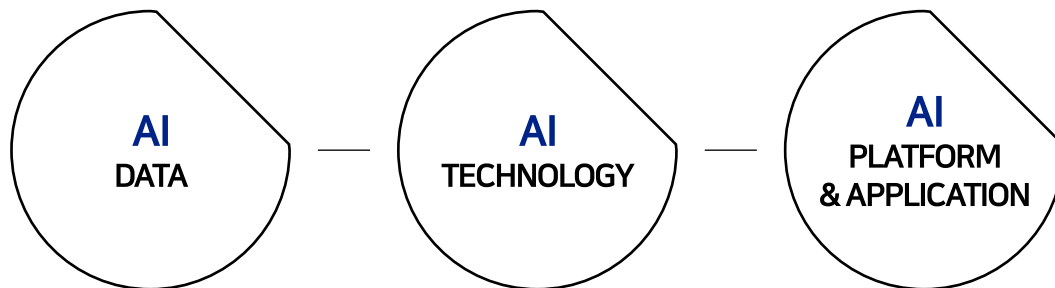
## **CODEVISION Inc.**

**"The innovation of AI technology enriches human society and contributes to a better future."**

CODEVISION : A specialized AI R&D company

- Leading technology in vision AI
- Developing optimal solutions driven by passion and innovation
  - Extensive R&D experience in collaborative projects with public and private enterprises

## **CODEVISION**



### **Company**

CODEVISION Inc.

### **C E O**

Eungyeol Song

### **Foundation Day**

November 10, 2018.

### **Headquarters**

Yonsei University S-Cube 304, 61 Yonsei-ro 2na-gil, Seodaemun-gu, Seoul, Republic of Korea

### **Business Area**

AI / Data / Cloud / SaaS

## CODEVISION

AI DATA  
AI TECHNOLOGY  
AI PLATFORM/APPLICATION

AI/Image processing research and development company



## PARTNERS



# Various R&D Experience History



## 2018 ~ 2020

- 2018.06. Order to technological innovation start-up business
- 2018.11. Established CODEVISION Inc.
- 2019.04. Selected the core technology development project in Robotics
- 2020.05 Selected for the AI Voucher Program
- 2020.05 Selected for the Data Voucher Program

## 2021

- 2021.04. Selected the 2021 AI Voucher Program
- 2021.05. Selected the 2021 Data Voucher Program  
Venture business approval
- 2021.06. Selected the 2021 ICT R&D Innovation Voucher Program
- 2021.07. Selected business with shipbuilding and marine industry  
technology project
- 2021.11. **Selected as a superior case in the 2021 AI Voucher  
Program (Top 8 in Korea)**

## 2022

- 2022.01. Established Affiliated Research Institute (Covi Lab)
- 2022.03. Selected the 2022 AI Voucher Program
- 2022.04. Selected the 2022 ICT R&D Voucher Program
- 2022.05. Selected the 2022 Data Voucher Program
- 2022.07. Approved as a professional R&D institution / engineering  
R&D institution
- 2022.11. Selected for the Startup growth technology development  
project in the field of a nuclear power plant

## 2023

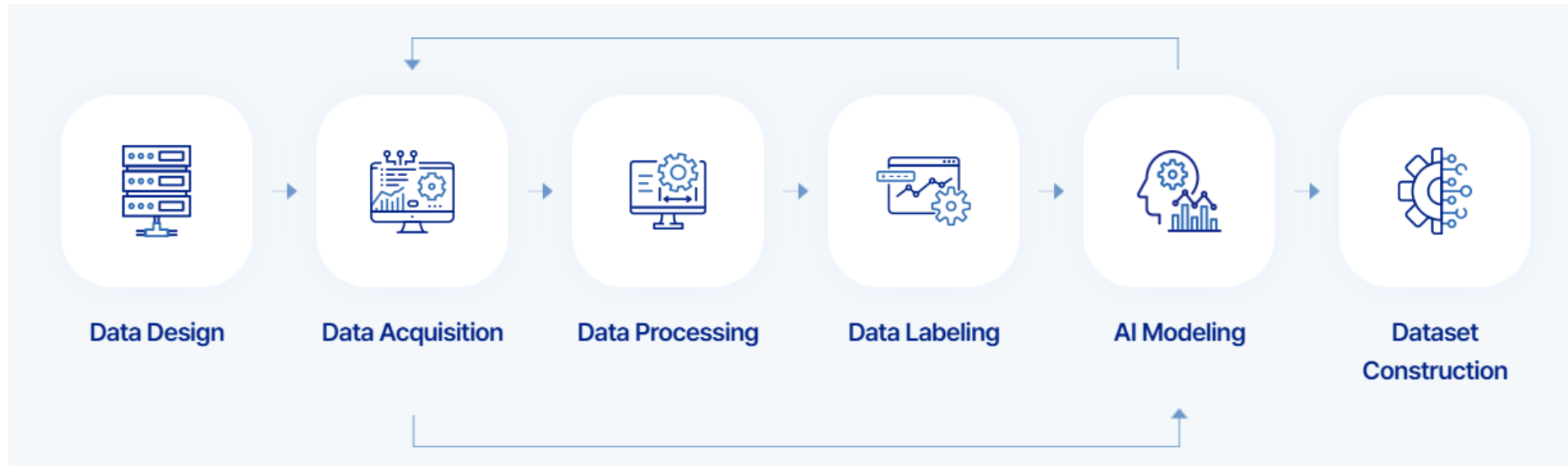
- 2023.03. Selected as a medical part for the 2023 AI Voucher  
Program
- 2023.04. Selected for AI industry convergence technology R&D  
project
- 2023.05. Selected the 2023 Data Voucher Program
- 2023.12. **Selected as a superior case in the 2023 Data Voucher  
Program**

## 2024

- 2024.05. Selected for the 2024 AI Voucher Project  
Selected for the 2024 Data Voucher Project
- 2024.06 Selected for the AI Convergence Technology Diffusion  
Project
- 2024.12. **Selected as a superior case in the 2024 AI Voucher  
Program**  
**Selected as a superior case in the 2024 Data Voucher  
Program**

# An Expert Approach to AI Data

## Systematic Data Processing

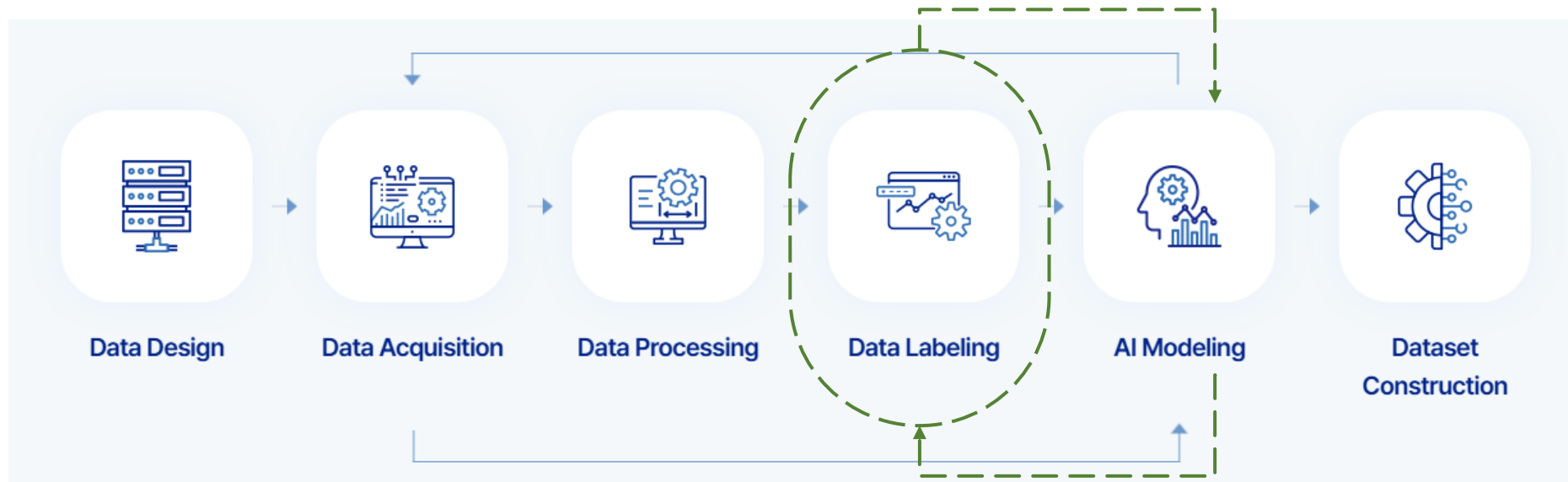


### Ensuring Data Reliability through Structured Management

- Managing data quality through iterative inspections
- Enhancing performance by reviewing edge cases frequently misclassified by the model

# Efficient Data Labeling with AI

## Systematic Data Processing



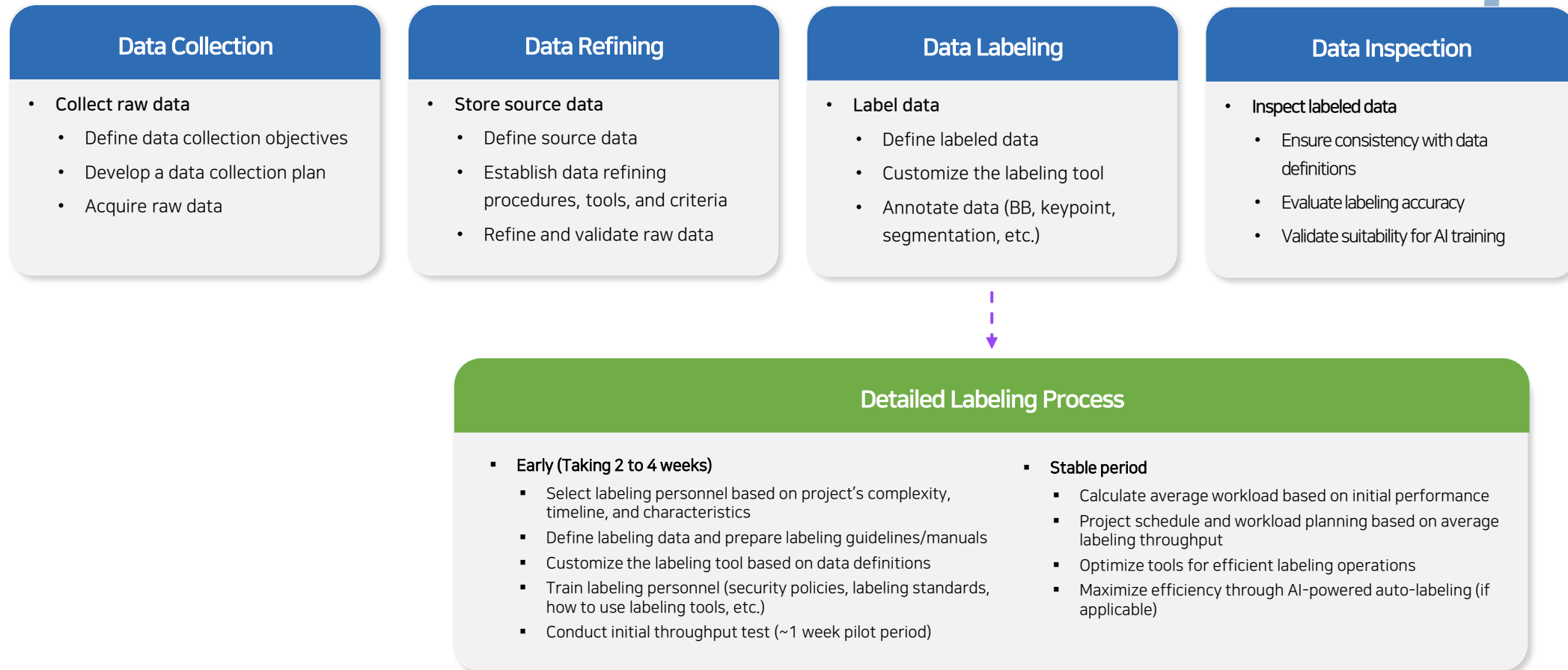
### Smart Data Labeling via AI

- Automating data labeling using high-accuracy models
- Ensuring labels accurately reflect data characteristics through expert validation

Train initial models on human-labeled data, then scale labeling using advanced AI

**Boost efficiency in processing large-scale data**

# End-to-End Data Processing Workflow



# Optimized Data Solutions for Industrial Environments

## Tailored Industrial Data Services

### CODEVISION Data Technology

Providing data optimized to the needs of each industry

Expert consulting by data specialists

Customized labeling platform

Build data suitable for the industry

#### 01\_ Customized Consulting

- Professional data consulting
- In-depth analysis of customer needs and site-specific requirements
- Tailored data design and construction plans derived from consultation

#### 02\_ Customized Labeling Platform

- Labeling tool tailored to your data
  - Secure handling of sensitive information (e.g., PII)
- Flexible configuration to meet customer requirements (labeling type, platform, etc.)

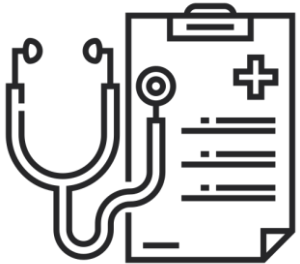
#### 03\_ On-Site Optimized Data

- Develop data tailored to each field environment
- Ensure seamless integration with existing workflows
- Proven expertise in building data for diverse industries

# Trusted Through Experience

## Successful Data Cases

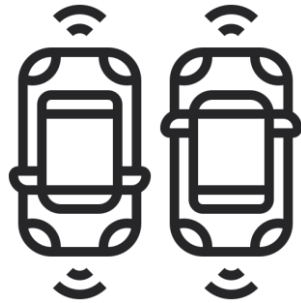
"Extensive Experience  
with 200,000+ Data Samples"



17,500+ Data Samples

### Medical/Healthcare

- Sleep data
- Biometric data (tongue)



25,000 + Data Samples

### Mobility

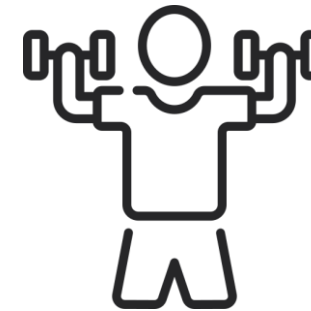
- Autonomous vehicle: Human/pet data
- Automotive crash tests: dummy data
- Vehicle collision detection: sensor data



54,000 + Data Samples

### Industry (Manufacturing)

- Defective car oil filter data
  - Ultrasound-based nondestructive testing data
- Power systems inspection ultrasonic data
  - Defective solar panel data
- 14 types of underwater env. data
- Vibration data from MOV actuators



42,500 + Data Samples

### Sports

- Exercise data from 10 types of gym equipment
- Golf action and pose data



32,500 + Data Samples

### Digital Life Style

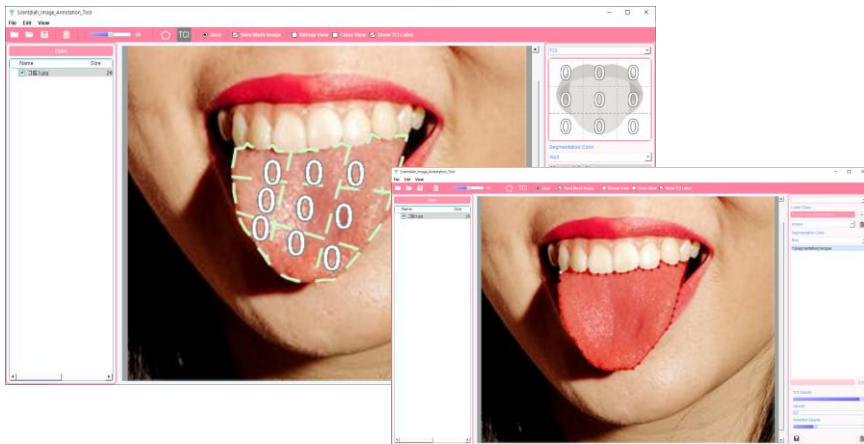
- Background removal (human segmentation) : person foreground data
- Video data of pet cats using litter boxes

# Medical/Healthcare Data Case

## Tongue Biometric Data for Health Diagnosis

### Customer Requirements

- Building image datasets to train AI for oral health diagnosis based on tongue surface features
  - Dividing the tongue into nine regions and labeling bacterial membrane accumulation levels
- Collecting tongue data from diverse individuals (age, gender, nationality) across various environments



CODEVISION tongue data labeling platform

### CODEVISION Solution

#### 01\_Privacy

- Only the tongue region was captured, limiting facial exposure
- Obtained legal consent for the collection, use, and sharing of personal data
- Built a local PC-based labeling system to process sensitive information without external transmission

#### 02\_Data Diversity

- Collected tongue data from 459 individuals aged 20 to 80
- Captured images under two lighting conditions (on/off)
- Captured multiple angles through guided subject movement

#### 03\_Accurate Consulting & Labeling

- Established labeling standards based on oral professionals' criteria for bacterial accumulation
- Segmentation required to identify and label all nine tongue regions
- Labeled a total of **15,000 tongue images** through segmentation and TCI scoring

#### 04\_Customized Labeling Platform

- Developed a custom labeling tool tailored to TCI (Tongue Coating Index) standards
- Algorithm visualizes 9 adaptive tongue segments based on tongue shape
- Each region was annotated with TCI values: 0(none), 1(moderate), 2(severe)

# Manufacturing Defect Data Case (1)

## Defect Inspection Data for Oil Filter Plate Assembly

### Customer Requirements

- Building AI training data for automatic inspection of oil filter plate assembly components
  - Handling various defect cases caused by lower plate welding in the element assembly process



(Left) Original + 5 augmented images  
 (Top right) Camera installation scene  
 (Bottom right) Labeling interface and sample result

### CODEVISION Solution

#### 01\_Field-Friendly Setup

- Collected data while the welding machine was in actual operation
- Collected data by small cameras to avoid interfering with field operations
- Installed the camera at an angle where the target area for checking for faulty parts is clearly visible

#### 02\_Systematic Data Processing

- Built a dataset of 5 cases: 1 normal and 4 defective
- Collected extra defect data due to low occurrence
- Built **100,000 refined video data** and **12,336 labeled samples**

#### 03\_Optimal Camera Selection

- Selected rugged, focus-adjustable cameras with high resolution and low distortion
- Created a candidate for the camera and pre-check data quality
- Identified the optimal camera and angle for real-world data collection

#### 04\_Data Diversity

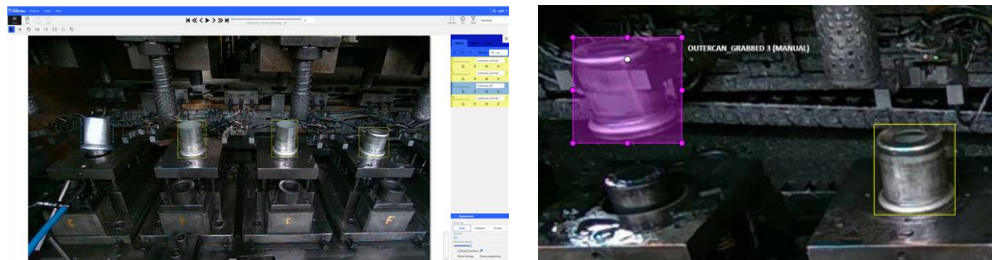
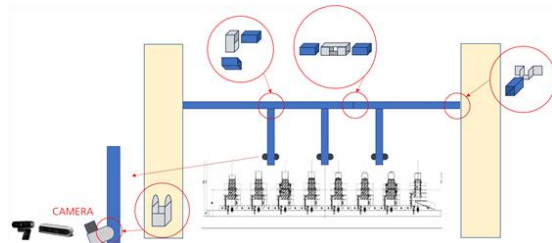
- Used dual webcams to capture from multiple angles
- Collected static frames at different times of day to capture changes in lighting and workflow
- Applied 5 augmentation techniques to overcome physical limitations: Brightness up/down, horizontal flip, clockwise/counterclockwise rotation

# Manufacturing Defect Data Case (2)

## Defect Detection Data for Outer Can Metal Molding Process

### Customer Requirements

- Data for automated detection of ejection defects during the press molding of outer cans (automotive components)
- Handling various outer can shape cases related to ejection defects during deep-drawing press operation



(Upper) Data collection system hardware diagram  
 (Lower left) Example of data labeling process  
 (Lower right) Example of data labeling results

### CODEVISION Solution

#### 01\_ Proper Equipment Installation

- Collected data during machine operation
- Installed cameras in consideration of locations that do not interfere with press machine and operator
  - Positioned cameras to clearly capture defect-prone areas

#### 02\_ Stable Data Collection

- Used two types of imaging devices optimized for the capture environment
- Recorded fixed-length videos with each camera and automatically transmitted and stored them
- Captured high-resolution videos suitable for AI training

#### 03\_ On-Site Data Design

- "Lighting conditions varied (e.g., natural light, LED lighting)
- Press machine vibration caused image instability
- Side movement may obscure some cans from the camera
- Based on the above analysis, systematically identifies and filters potentially low-quality data

#### 04\_ Real Defect-Based Data

- Built datasets for 5 representative press cases, including 2 normal and 3 defect scenarios
- Collected additional defect data to address class imbalance
- Constructed **150 GB of refined video** and **19,486 labeled samples**

# Mobility Data Case

## Vehicle Collision Detection Sensor Data

### Customer Requirements

- Building data to develop high-accuracy deep-learning models for enhanced dashcam collision detection
- Designing diverse impact scenarios (e.g., door dings, hammer strikes, jigs) and collecting sensor data under real parking lot conditions



그림. 전방 MTC센서

그림. 전방 G센서

그림. 센터 G센서



그림. 후방 G센서 (상단, 하단)



그림. 후방 G센서 (트렁크 내부)



(Left) Sensor placement inside the vehicle  
(Right) Example of data collection setup

### CODEVISION Solution

#### 01\_ Systematic Collection Scenario

- Defined scenarios considering driving/parking status, collision type, and location variety
- Increased data volume for critical scenarios like parking crashes
- Specified sensor installation positions within vehicles

#### 02\_ Environmental Factors

- Tested with designated vehicles (e.g., Genesis G90) and impact equipment
- Logged sensor data with environmental info (temp, humidity, location, weather)
  - For door ding collisions, ensured collection of vehicle-to-vehicle distance and related impact details

#### 03\_ AI-Ready Data

- Included impact type, intensity, direction, impact phases' timestamp, vehicle info and number of occupants
- Collected and processed sensor data in real time through an integrated interface connected to the vehicle
- Built a final dataset with **30,000+ collision events**

#### 04\_ Inspection & Additional Collection

- Performed visual data inspection using dedicated QC software
- Removed problematic segments (e.g., sensor dropout, excessive noise) to ensure dataset integrity
- Used inference verification to detect edge cases and expand the dataset with new samples

# Why CODEVISION Data Solutions?

## More Than Just Meeting Customer Needs

### Tailored to Customer's Data Maturity Level

- For those without existing data who need to begin collection
- For those with data but lacking knowledge in utilization
- Provide optimal solutions tailored to each customer's data maturity level

### CODEVISION DATA

### AI Data Expert-Driven Solutions

- Building data based on consultations with experienced AI experts
  - Systematic process and reliable execution
    - Expert handling of edge cases where standard models show lower accuracy

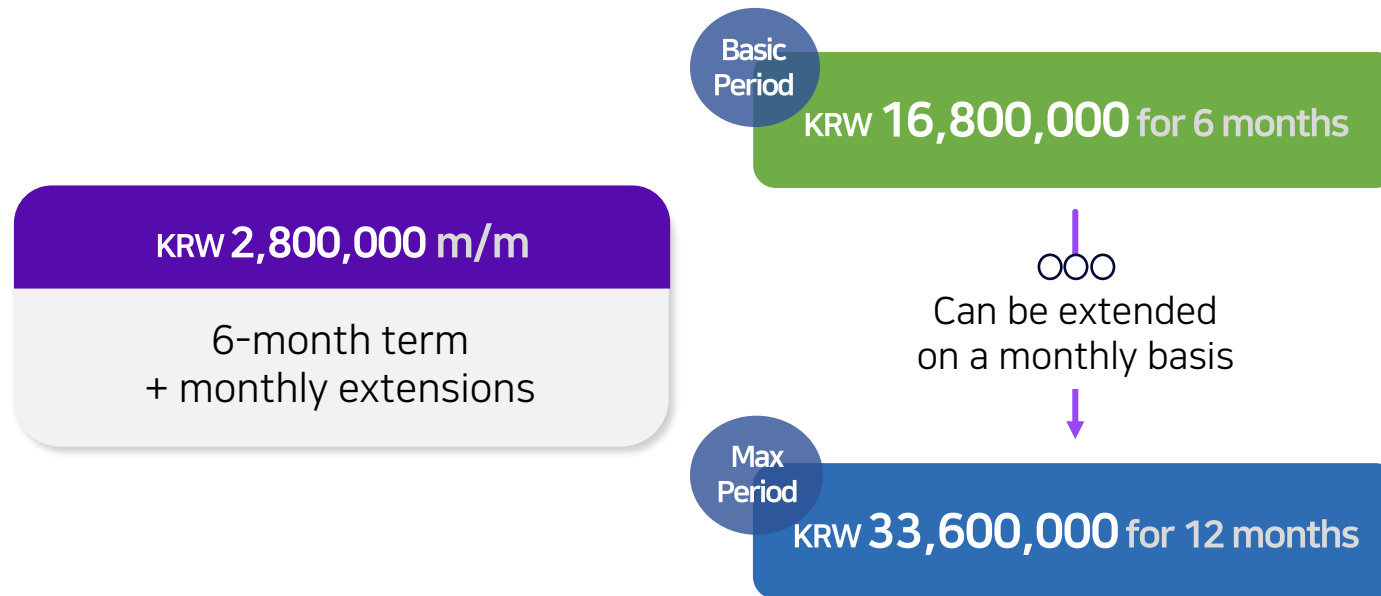
### Sustainable Data Support

- For customers needing long-term data management, not one-off projects
- For those handling diverse types of data over time
- Data partnership for scalable, long-term strategies

# Concentrated Data Solution

Suggestions for Those Who Need Long-Term Data Solutions

Key to handling large and varied data:  
**Long-term support from AI data experts**



- N-month contract (minimum 6 months)
- Tasks can be adjusted during the contract period (e.g., BB, Segmentation, Point, etc.)
- Change of assigned personnel is possible
- Customer requirements can be reflected (e.g., additional labeling after inspection)

[Example]

- 2 people for 6 months : KRW 33,600,000
- 3 people for 12 months : KRW 100,800,000

Depending on the amount of data, the difficulty of the task, etc., the number of people and the length of time required may vary.

Unlike typical solutions based on processing volume, our pricing is based on monthly labor costs and the contract duration.  
 (Please contact us for more information.)

# Thank you



[www.codevision.kr](http://www.codevision.kr) / (02) 2088-6624

Copyright © All Rights Reserved by CODEVISION Inc.