

Inspur E2E Smart Transportation CV Solution (using Insight Data-Intelligence Platform)

Jason Liu

Deputy General Manager of Inspur Cloud Computing and Big Data Product Line

Agenda

- **Motivations**

- Industry background

- **Insight's solution for Smart Transportation**

- Insight Data-Intelligence platform Introduction
- Insight End to End CV Solution architecture
- Data loading pipeline: building distributed reading of big data
- An introduction to the principle of smart transportation system
- Walkthrough: building distributed training tasks based on Analytics zoo

- **Insight's application for Intelligent Traffic**

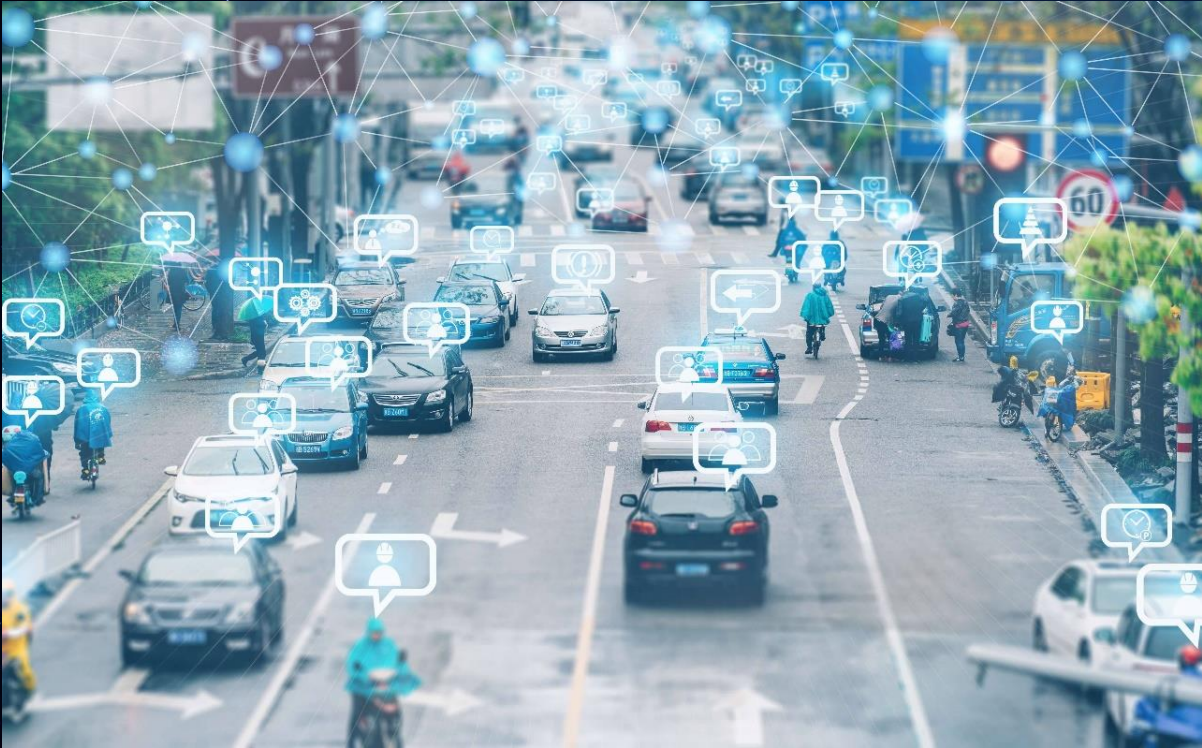
- **Summary**

Motivations

Industry Background

Industry Background

Analysis on trend of Smart Transportation in China



In recent years, China's transportation construction has developed rapidly and has made great achievements. But the ensuing problems are also more prominent, including three aspects: travel safety, traffic efficiency, and green environmental protection.

In order to solve these problems, the state has issued a series of related policies to support the construction and development of smart transportation.

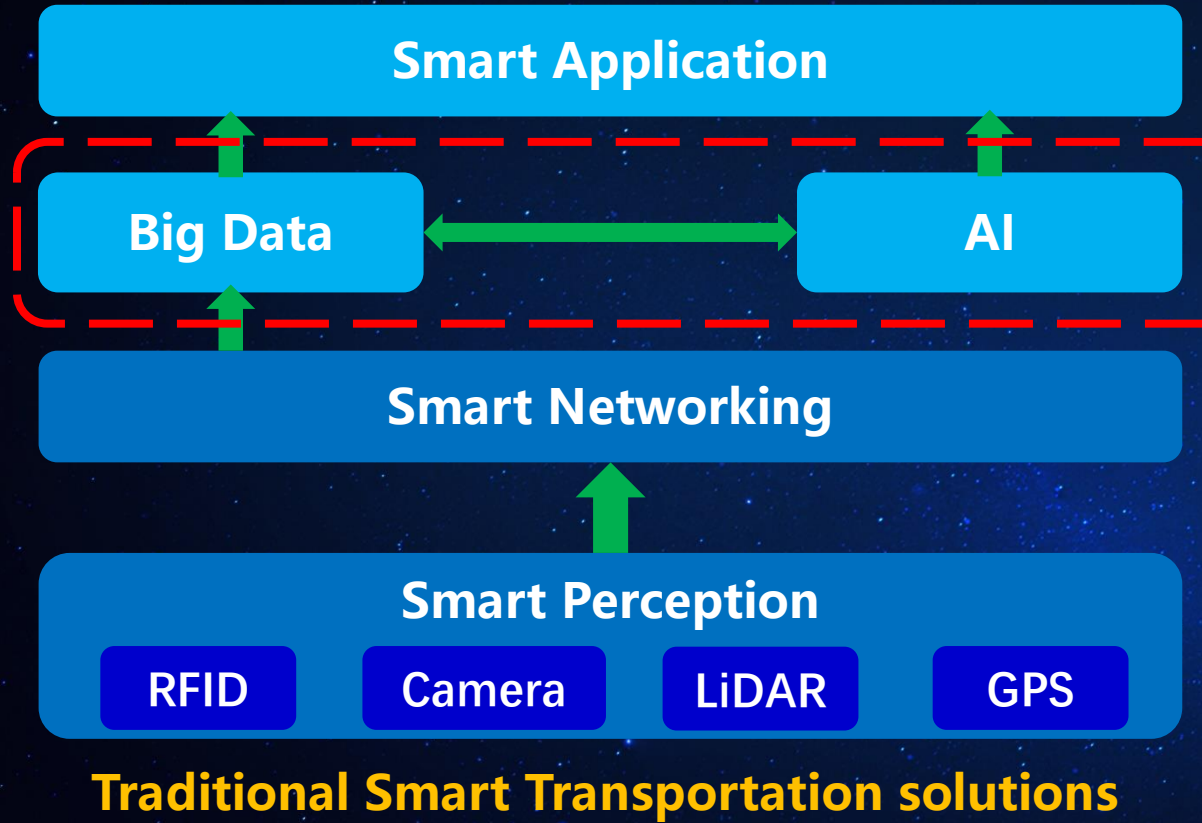
Industry Background

Smart Transportation includes:

- **Smart Perception:**
Use sensors to perceive environmental information.
- **Smart Networking:**
Achieve the interconnection of human, vehicle, road and environment.
- **Smart Decision-Making:**
It contains Big Data and Artificial Intelligence(AI) technology, which can process large amounts of data and decision-making in real time.
- **Smart Application:**
Upper-level AI applications based on Smart Decision-Making, such as Self-Driving cars, MOT(Multi-object tracking), and so on.

Industry Background

Challenges



■ Fragmented

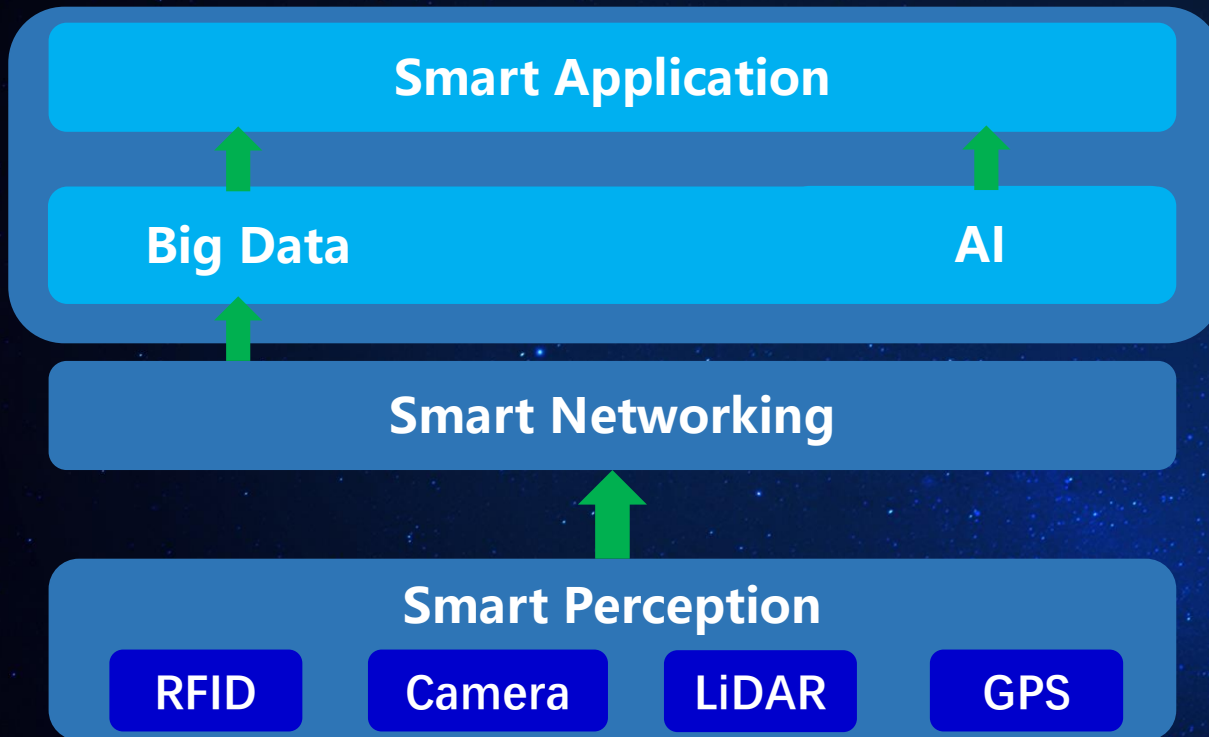
Traditional Smart Decision-Making layer doesn't integrate Big data and AI seamlessly, which makes Big Data Platforms cannot efficiently support AI Applications.

■ Inefficient

Data storage formats and structures are different for big data and AI, it difficult to provide an agile and efficient infrastructure platform for big data analysis and AI Applications. With the rapid growth of data volume and continuous iteration of AI Applications, we need a good solution for efficient data access and synchronization

Industry Background

Solution



Insight's solution for smart transportation

■ Unified

The solution provides a unified big data and AI analysis platform integrated with big data preprocessing, model training, inference, and existing big data processing workflows for an end-to-end big data analysis and deep learning application pipeline. efficiently support AI Applications.

■ Efficient

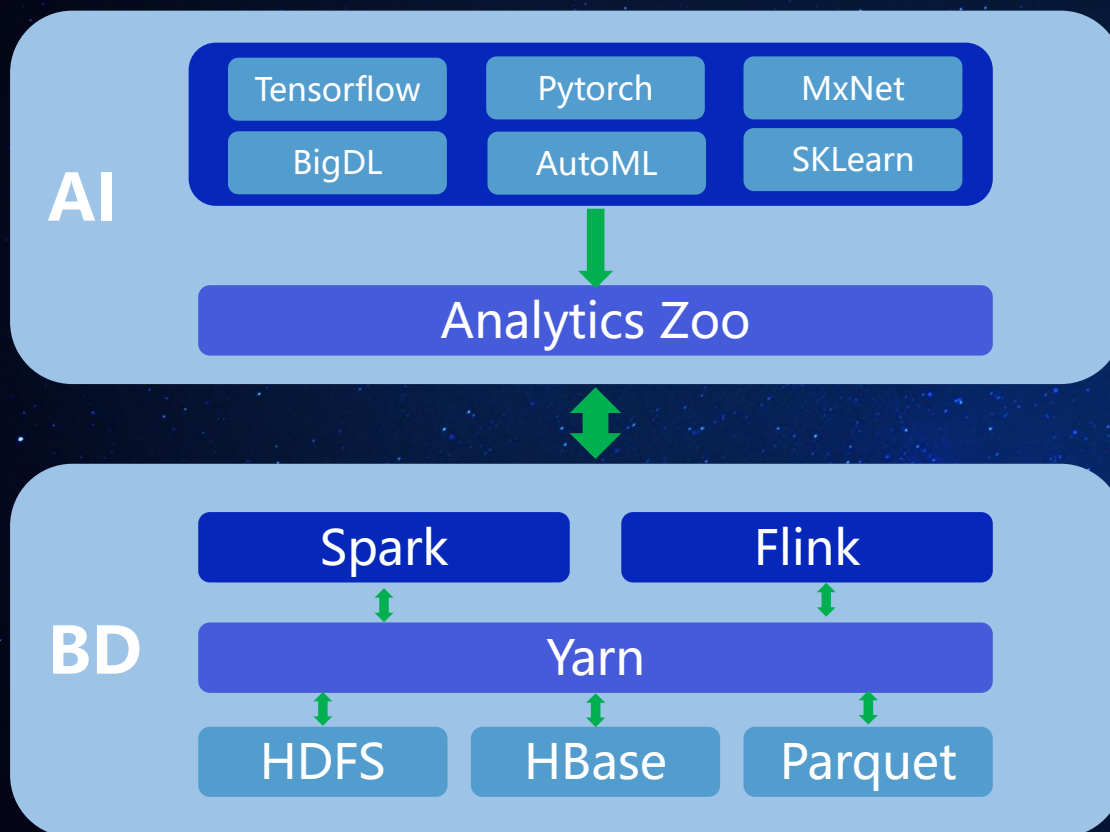
It significantly improves the deployment efficiency, resource utilization, and scalability of systems while reducing hardware management and CloudOps costs.

Insight's solution for smart transportation

Insight Data-Intelligence platform architecture

inspur

End-to-end smart computing solution based on Inspur's Insight big data platform and Intel® Analytics Zoo integrated.



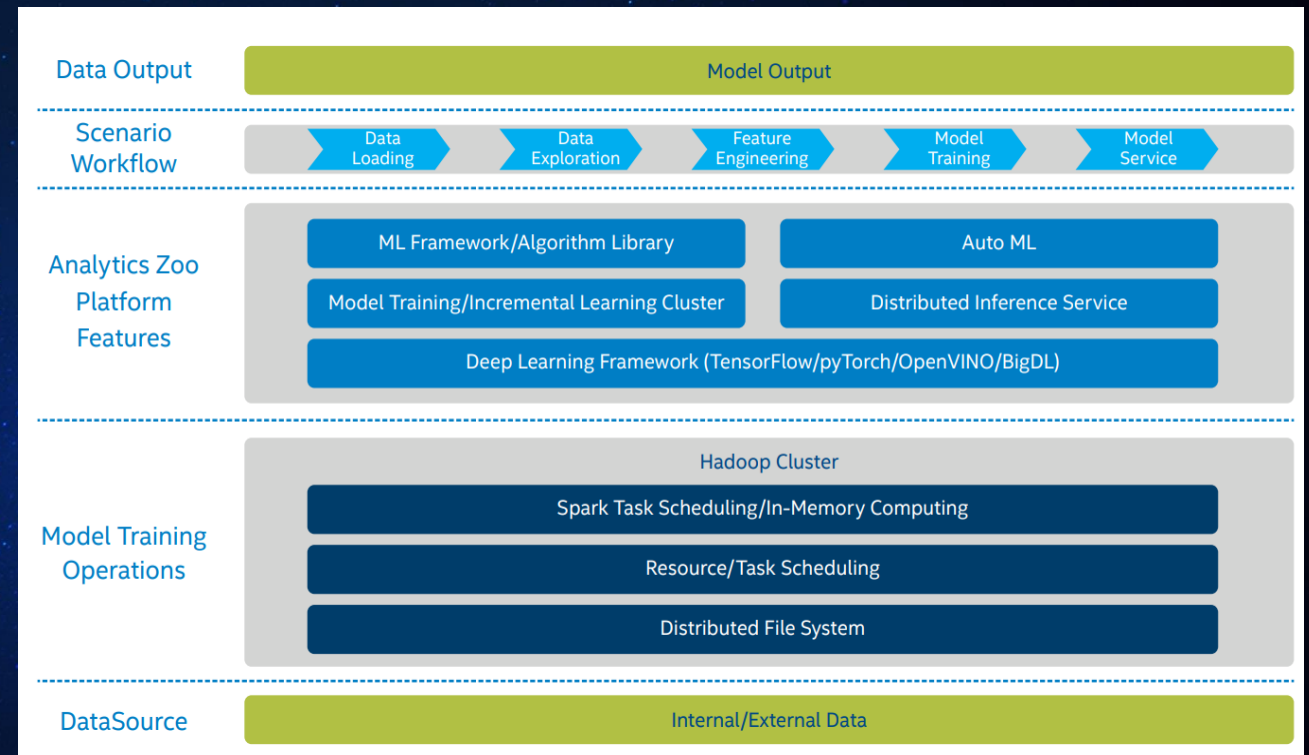
■ **Big data:** It integrates compute engines, storage services, data services, data workbenches, and management platforms on optimized infrastructure layers to build an end-to-end big data analysis and deep learning application pipeline.

■ **AI:** Analytics Zoo is the key component, which can seamlessly integrate Spark, Tensorflow, Keras and BigDL programs into a single integrated pipeline that can be transparently extended to largescale Apache Hadoop/Spark clusters for distributed training or inference.

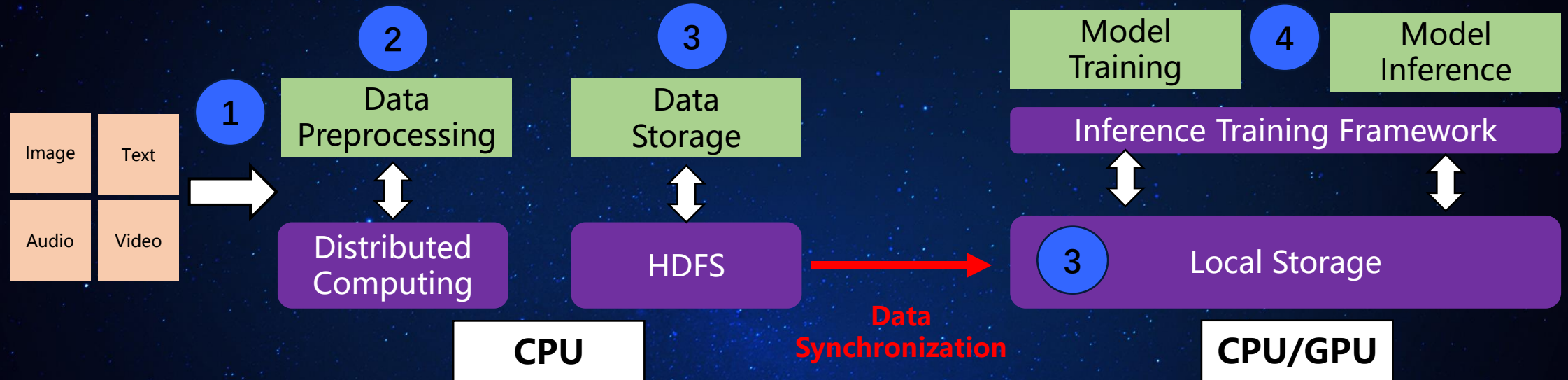
Insight Data-Intelligence platform features

Leveraging Inspur DIP to build one-stop and end-to-end AI applications

- A unified platform that integrates big data analysis and AI.
- Improve resource utilization and accelerate deployment of AI applications.
- Distributed training and inference improve system scalability.
- Better infrastructure efficiency and reduce TCO.

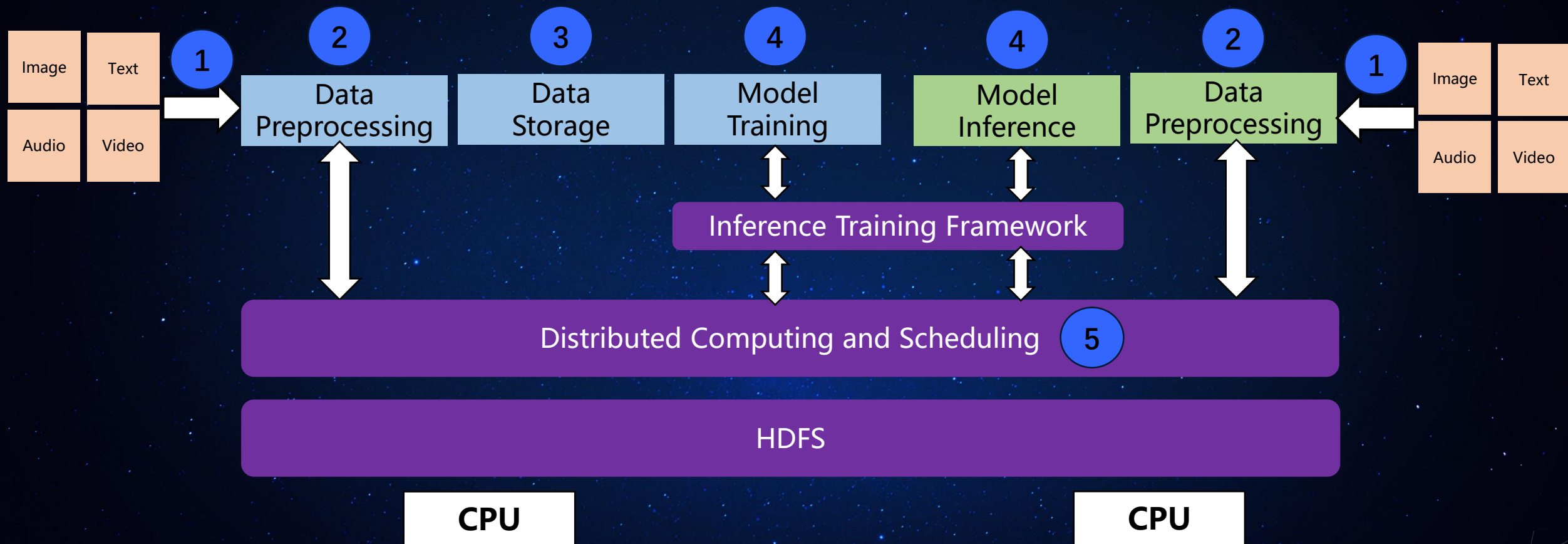


Insight Data-Intelligence platform features



Traditional Compute Mode

Insight Data-Intelligence platform features



One-Stop End-to-End Smart Computing Mode

Insight End to End CV Solution architecture



- 1 Data Loading
- 2 Data Preprocessing
- 3 Developing ML/DL models
- 4 Distributed Training/Inference

Data loading pipeline

Building distributed reading of big data

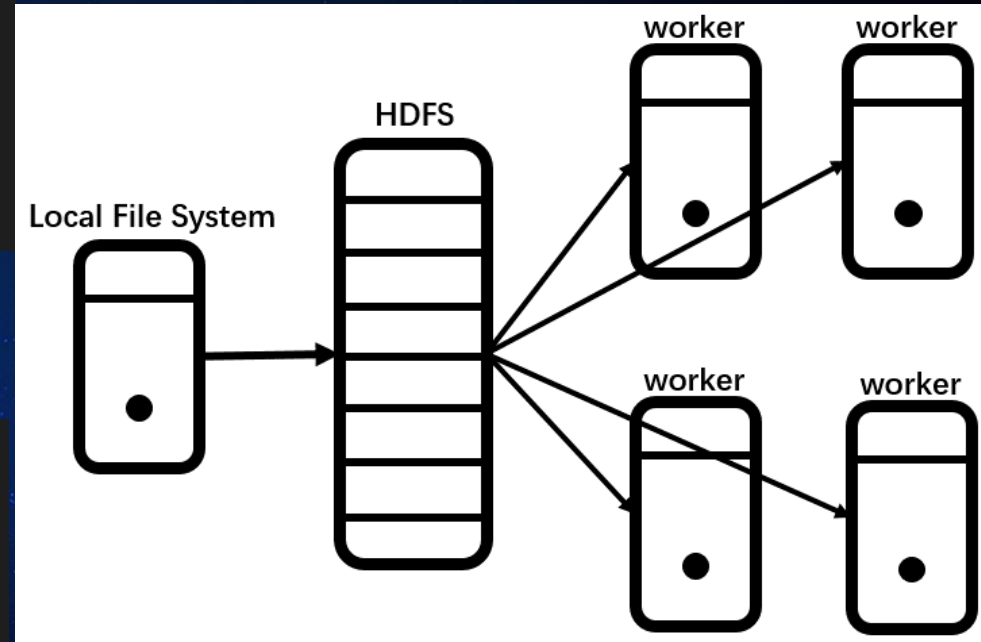
1. Use *write_parquet* to convert local data to parquet and save it to HDFS.

```
from zoo.orca.data.image.parquet_dataset import write_parquet

write_parquet('kitti_tracking',
              output_path="hdfs://output_dir",
              image_dir="~/kitti_tracking/training/image_02",
              anno_file="./annotations/tracking_train.json")
```

2. Use *read_parquet* to get *tf.data.Dataset* from HDFS load parquet data.

```
from zoo.orca.data.image import read_parquet
import tensorflow as tf
train_datasets = read_parquet(format="tf_dataset",
                              path="hdfs://output_dir",
                              output_types=(tf.float32, tf.float64),
                              output_shapes=((516,516,3),(129,129,9)))
# train_datasets = read_parquet(format="dataloader",
#                               path="hdfs://output_dir")
```



An introduction to the principle of smart transportation

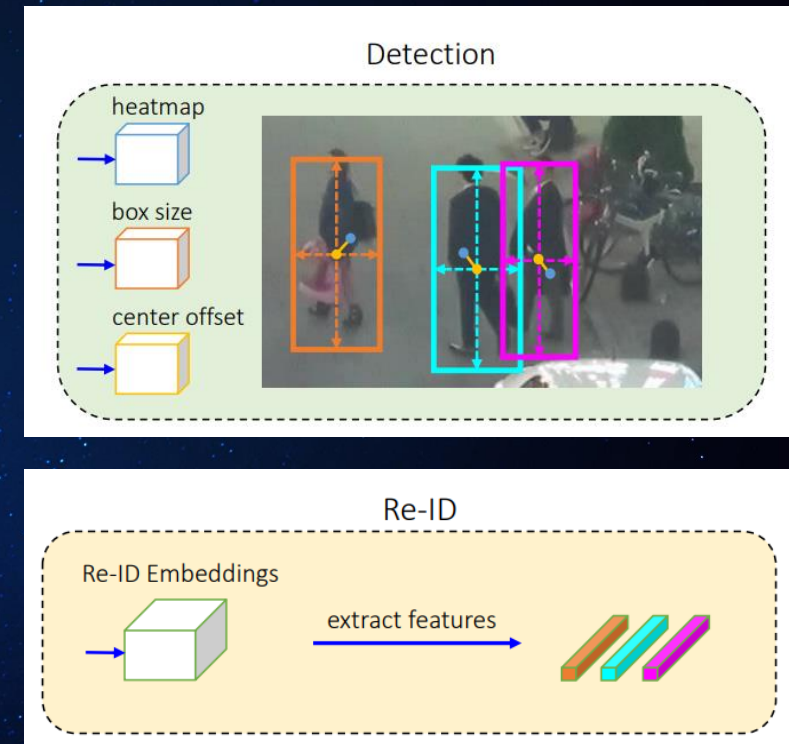
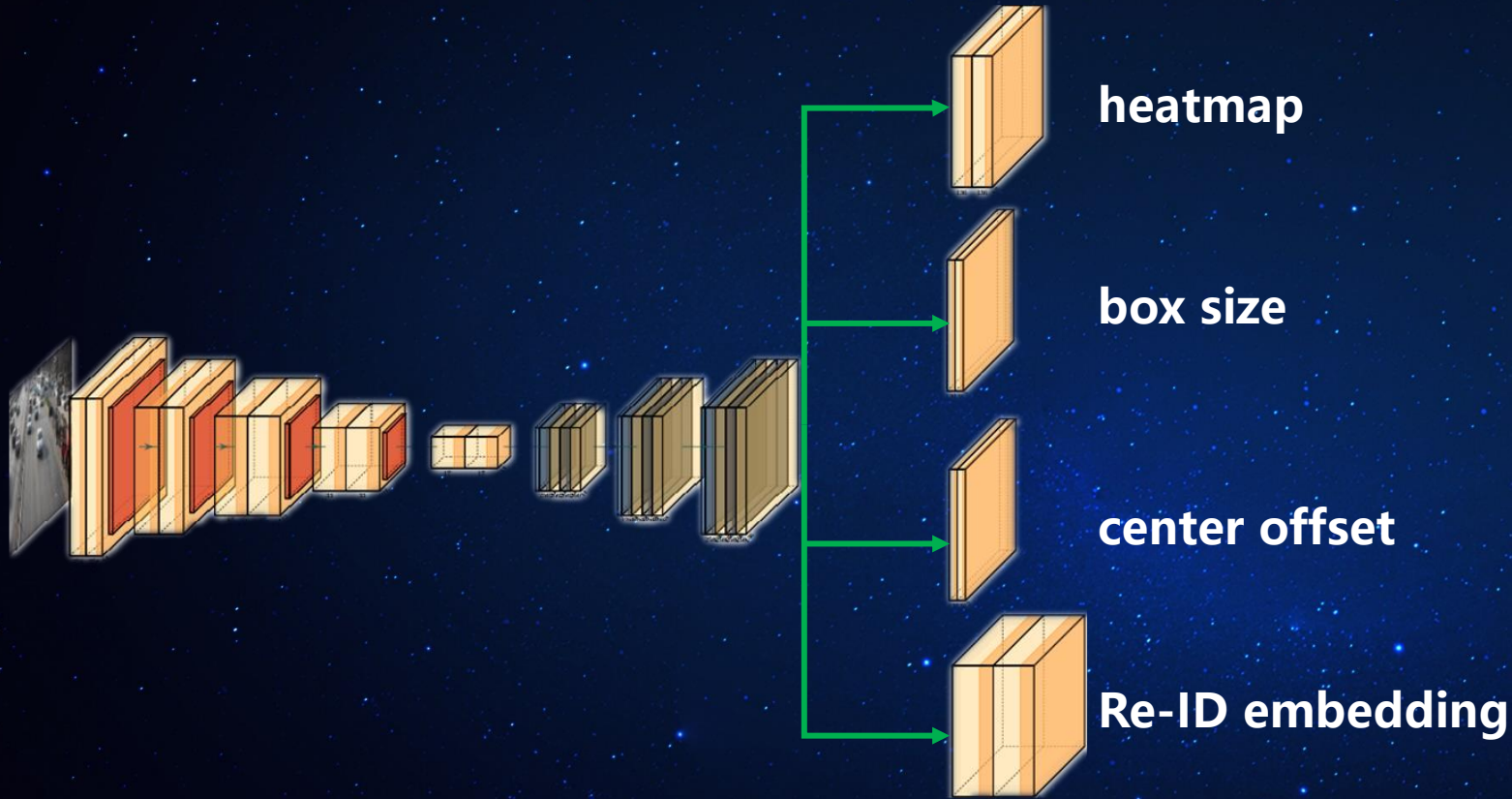
Multi-object tracking, OCR recognition



- The video stream data of the big data platform inputs to Multi-object tracking model for inference, and the location and ID information of the target in the picture are predicted.
- The predicted vehicle location information is utilized to intercept vehicle data and the vehicle data is pass into the OCR model to identify license plate information.
- The vehicle trajectory binding the vehicle ID information and license plate information is send to the Insight big data platform.

Multi-object tracking algorithm introduction

Object Detection, Re-ID



Multi-object tracking algorithm introduction



T-1 image



T-1 result of detection



T-1 feature vector

Time T-1 is the image of the previous frame of the video. Through algorithm prediction, the target category, location and feature vector of the object are obtained.



T image



T result of detection



T feature vector

Time T is the image of the current frame of the video. After the model predicts, it will be fused with the prediction result (category, position, feature vector) of the previous frame to match the same target.

Walkthrough

Building distributed training tasks based on Analytics zoo

1. Build a data generator: load the training dataset and the validation dataset.

```
def train_data_creator(config):  
    #load data  
    train_datasets = read_parquet(format="tf_dataset",  
                                  path="hdfs://output_dir",  
                                  output_types=(tf.float32, tf.float64),  
                                  output_shapes=((516,516,3),(129,129,9)))  
  
    #augment data  
    datasets = datasets.map(transform).shuffle(100).batch(config['batch_size'])  
    return datasets  
  
def valid_data_creator(config): ...
```

2. Build ML/DL models

```
def model_creator(config):  
    #using keras api build training model  
    net = FairMot(config)  
    return net
```

Walkthrough

Building distributed training tasks based on Analytics zoo

3. Initialization of distributed environment

```
init_orca_context(cluster_mode="yarn-client",  
                  num_nodes=8, cores=20,  
                  init_ray_on_spark=True,  
                  memory="20g", driver_memory="16g",  
                  hadoop_user_name='hdfs',  
                  hadoop_conf="/etc/hadoop/3.0.1.0-187/0/")
```

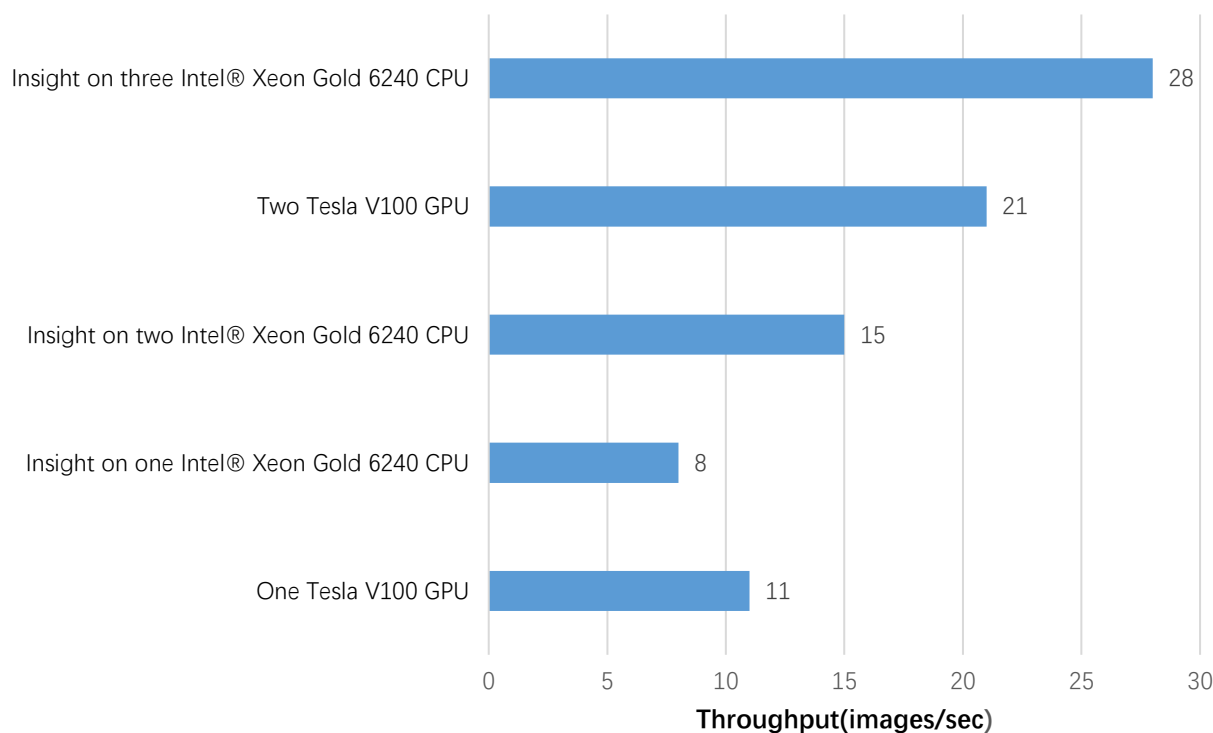
4. Distributed training ML/DL models

```
#Estimator load keras Model  
est = Estimator.from_keras(model_creator, config=config)  
  
#distributed training keras Model  
est.fit(train_data_creator,  
        epochs=10,  
        steps_per_epoch=train_num // batch_size,  
        validation_data_creator=valid_data_creator,  
        validation_steps=valid_num // batch_size)
```

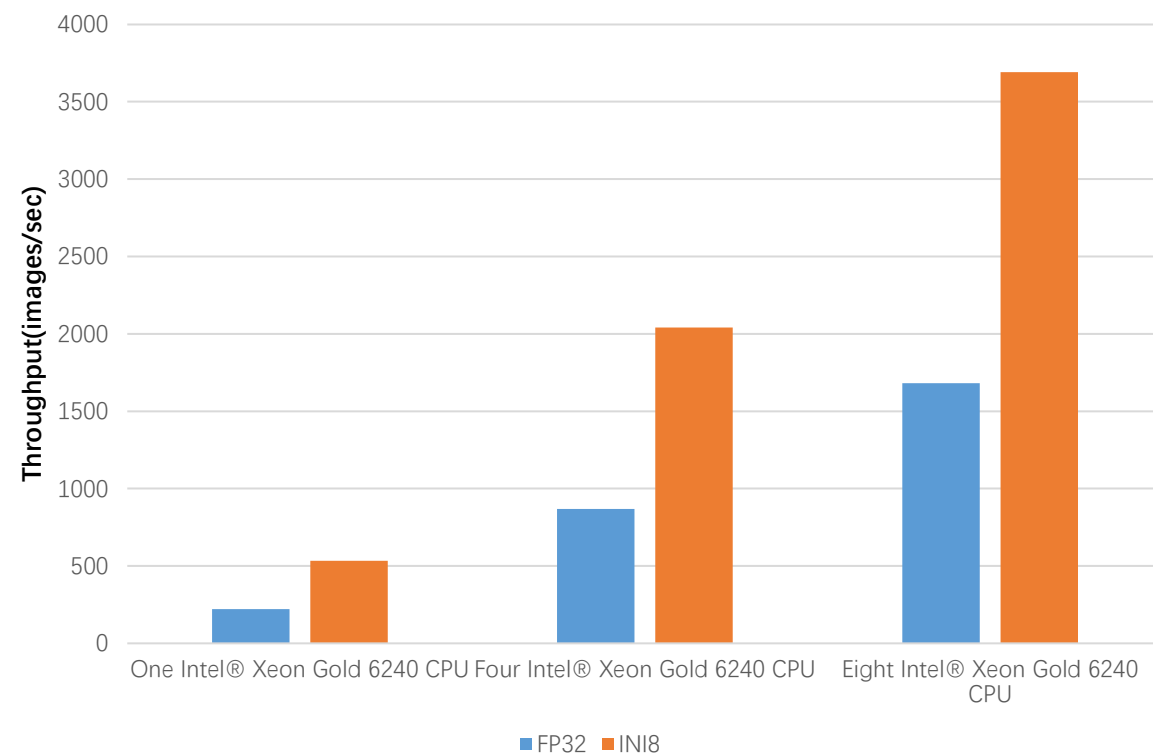

Performance and Benefits

High Efficiency, Low Cost

Training Throughput



Insight's distributed inference performance

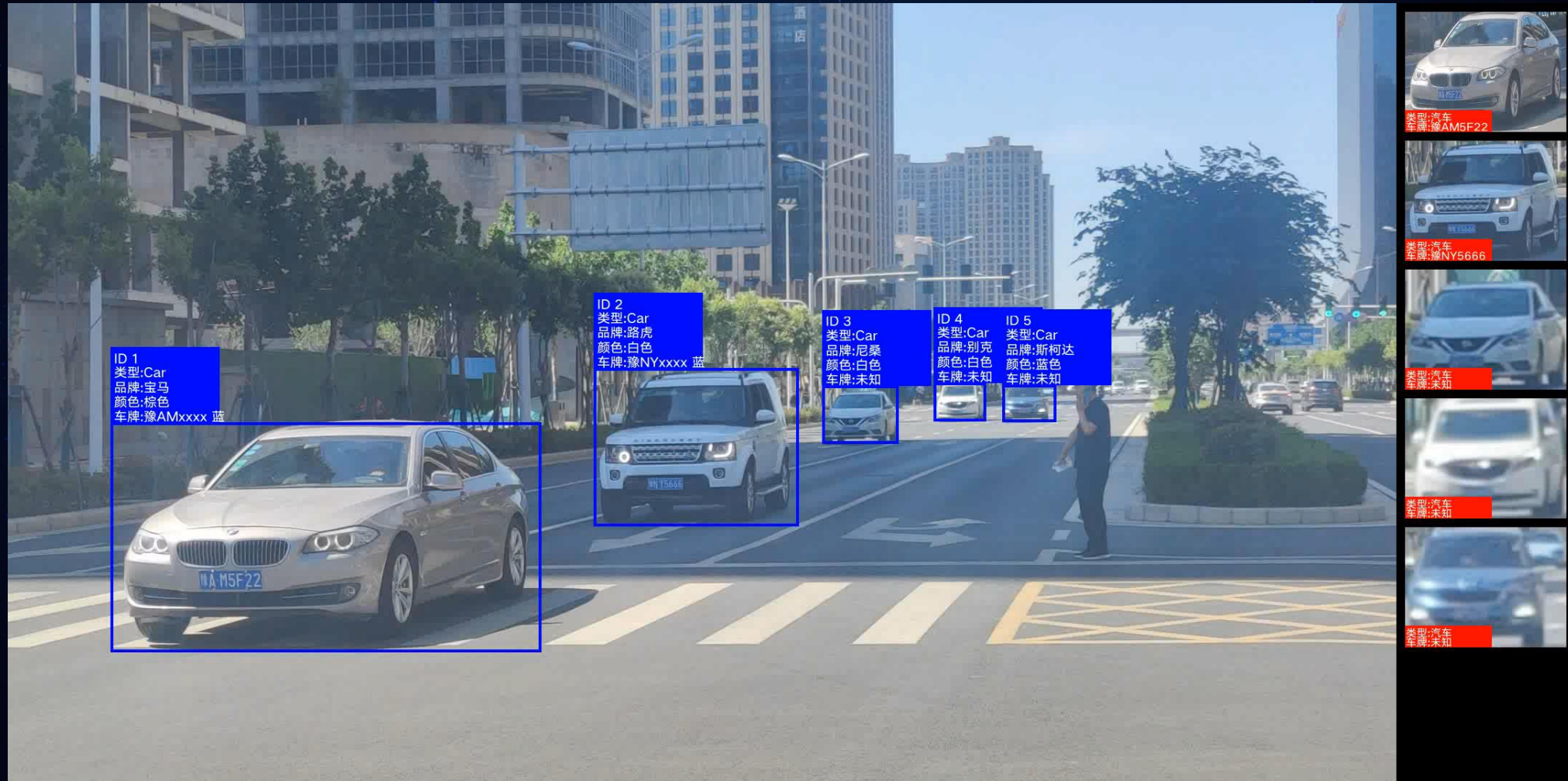


get 30% higher performance and 40% lower cost

Insight's application for Intelligent Traffic

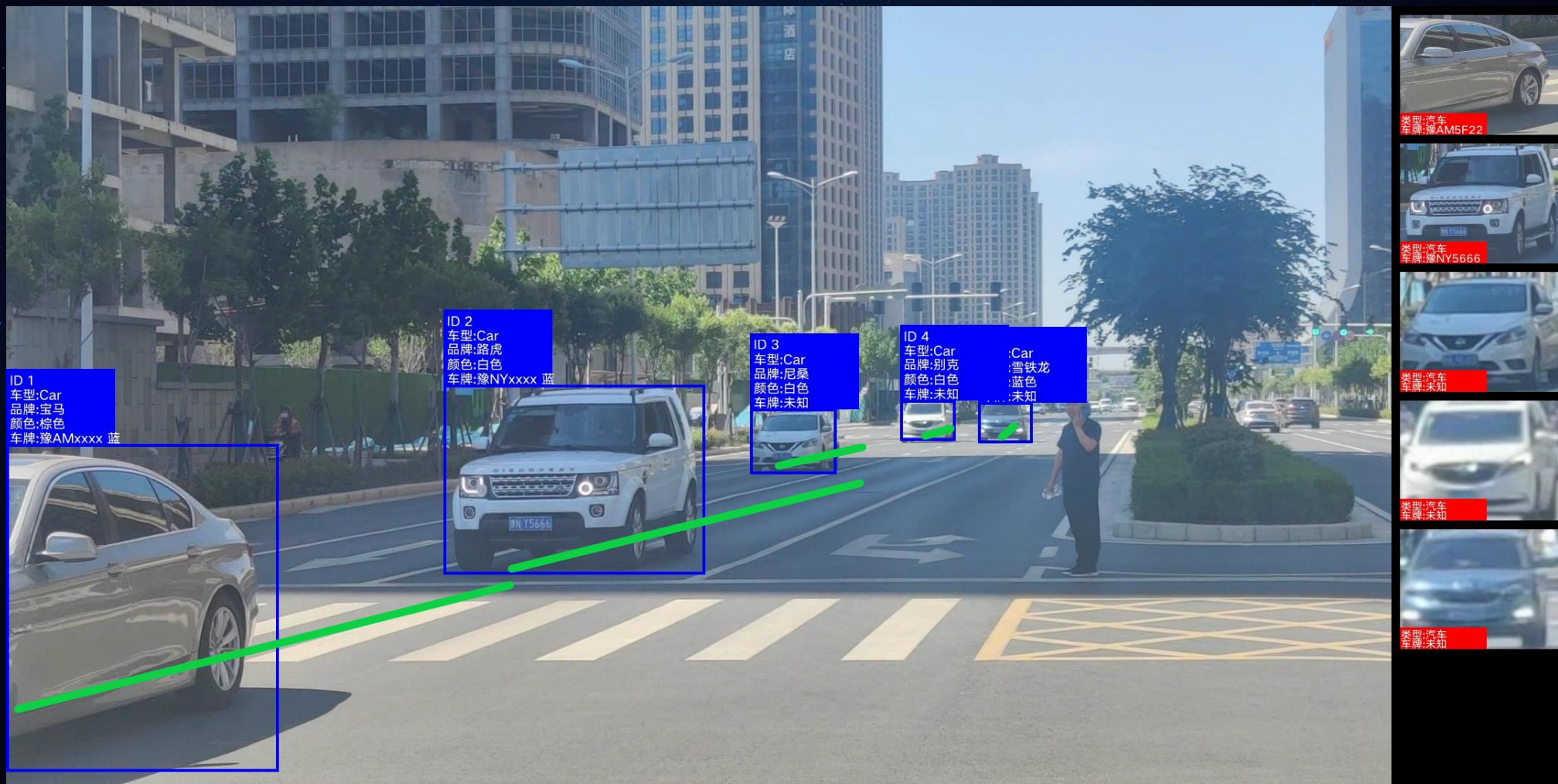
Insight's application for Intelligent Traffic

1. Vehicle detection and tracking system at traffic intersections



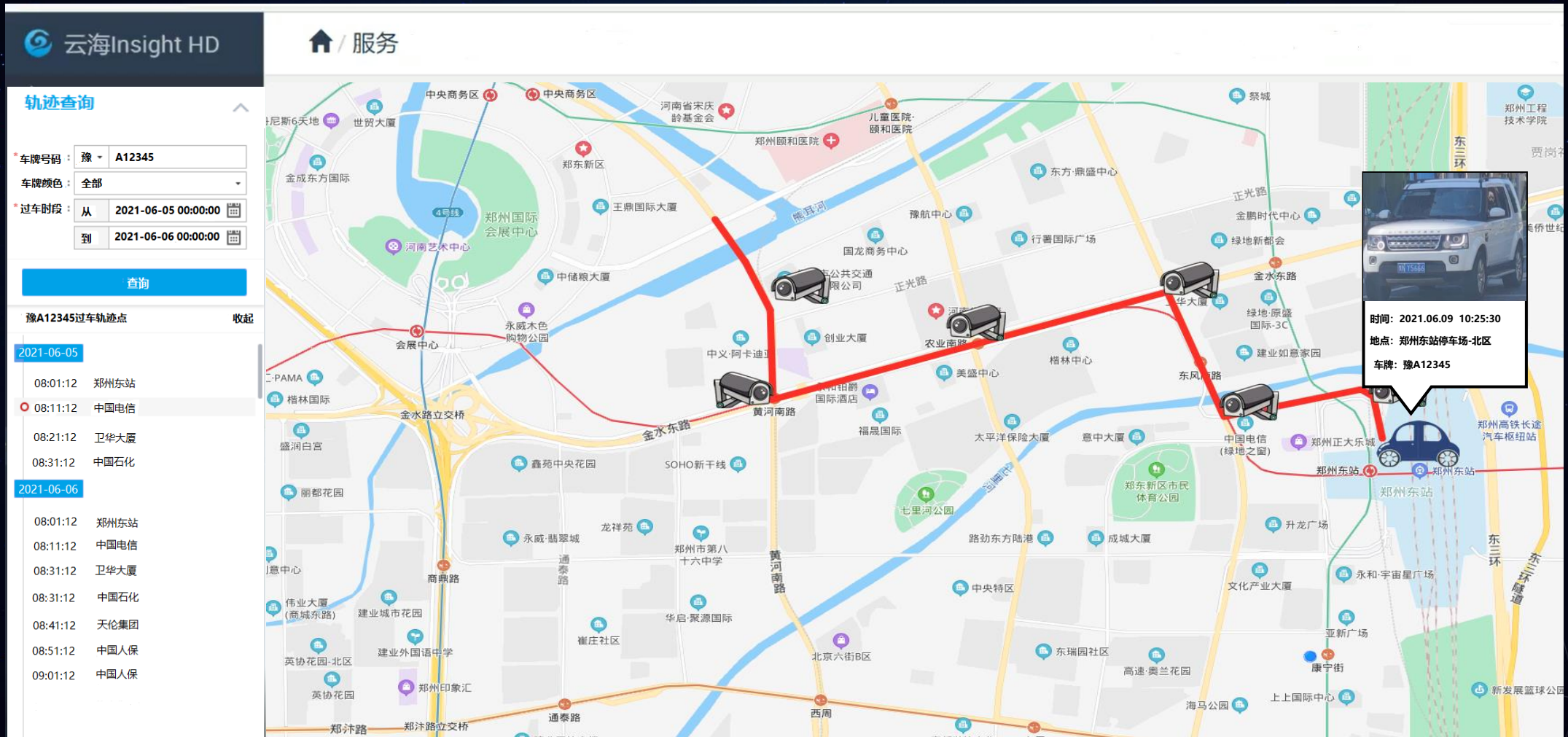
Insight's application for Intelligent Traffic

2. Vehicle trajectory recording system at traffic intersection



Insight's application for Intelligent Traffic

3. Vehicle trajectory tracking system based on big data platform



Contact me:

- Email: liujianbj@inspur.com
- Wechat: 18610069959 (same with mobile phone)

Thanks