WeBank

FedVision: An Online Visual Object Detection Platform Powered by Federated Learning

Yang Liu, Anbu Huang, Yun Luo, He Huang,
Youzhi Liu, Yuanyuan Chen, Lican Feng, Tianjian
Chen, Han Yu, Qiang Yang

WeBank 微众银行

EXTREME VISION 极视角

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

香港科技大學 THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Two Secrets of AI’ s success : Computing Power and Big Data

Computing power

Intel i386

Intel i486

Intel Pentium

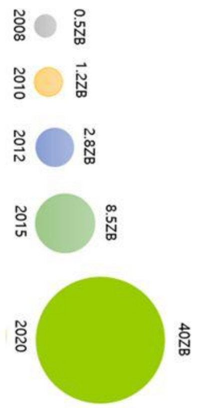
Intel Core

Nvidia GPU

Google TPU

Big data

1 ZB= 10^{21}Byte



——来自互联网数据中心（IDC）

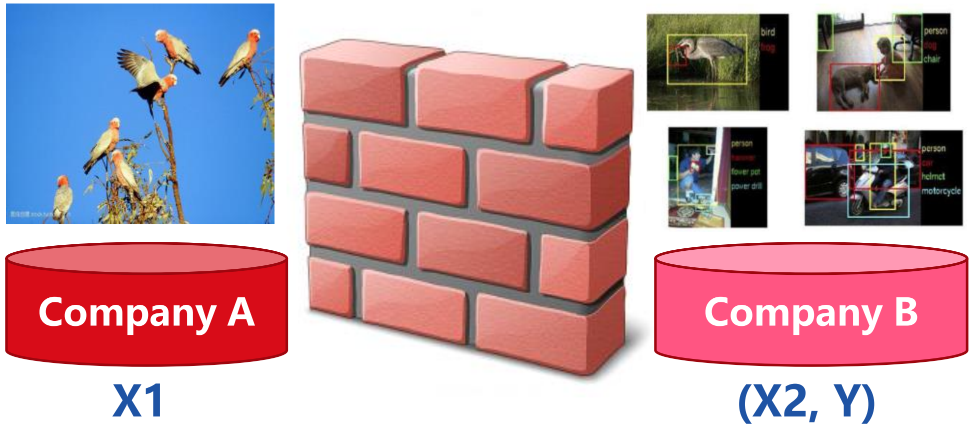
The New Rich

The world’s most valuable resource is no longer oil, but data.

The Economist - May 2017

WeBank

Motivation

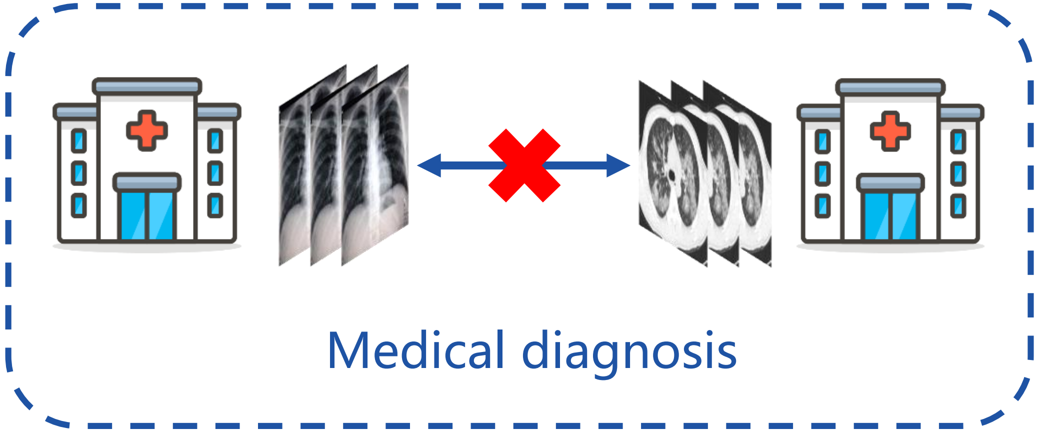


Company A

X1

Company B

(X2, Y)



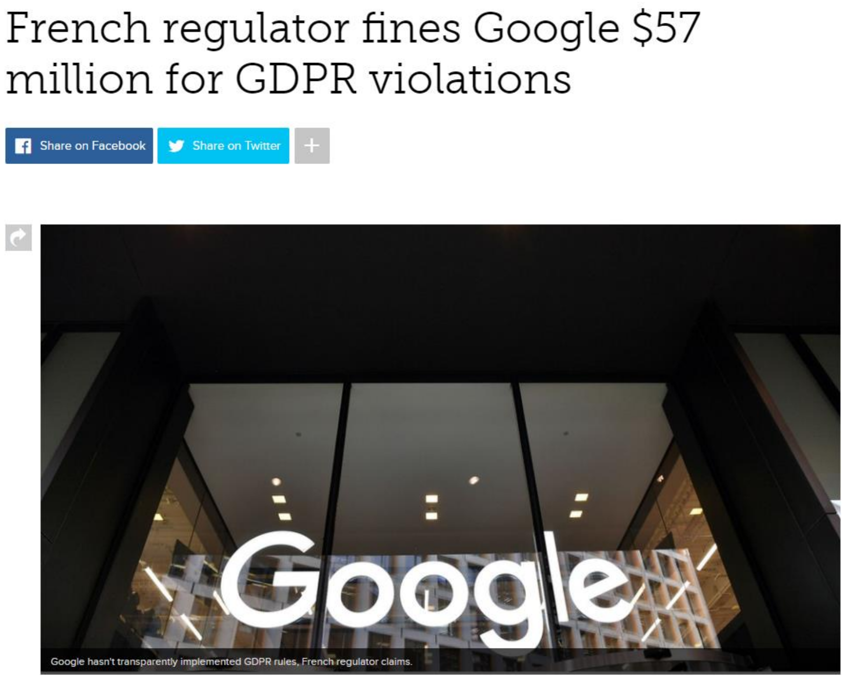
Medical diagnosis

- Data exists in the form of isolated islands
- Data integration between different departments of the same company faces heavy resistance
- Almost impossible to integrate the data scattered around the country and institutions
- Due to data privacy and data security, it is unfeasible to share sensitive data.

Motivation

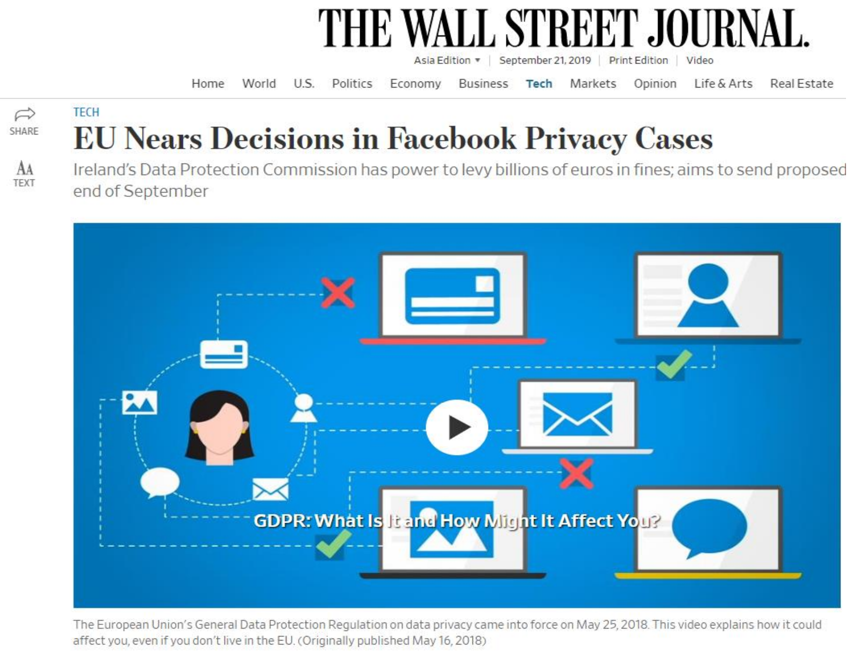
Legislation to protection of data security and privacy

- No Autonomous Modeling and Decision
- Interpretability of Model Decisions
- Users'Right for Data to be Forgotten
- Data Privacy By Design
- Explicit Consent for Data Usage



French regulator fines Google $57 million for GDPR violations

Google hasn't transparently implemented GDPR rules, French regulator claims.



THE WALL STREET JOURNAL.

Asia Edition ▾ | September 21, 2019 | Print Edition | Video
Home World U.S. Politics Economy Business Tech Markets Opinion Life & Arts Real Estate
SHARE
AA TEXT

TECH

EU Nears Decisions in Facebook Privacy Cases

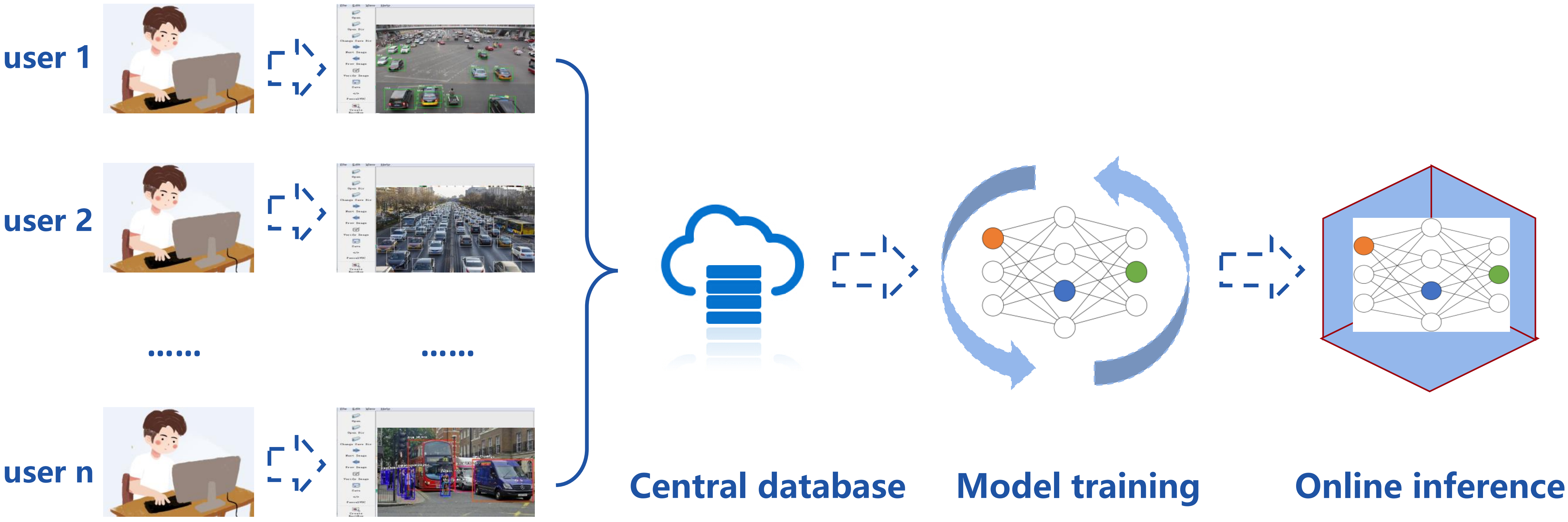
Ireland's Data Protection Commission has power to levy billions of euros in fines; aims to send proposed end of September

GDPR: What Is It and How Might It Affect You?

The European Union's General Data Protection Regulation on data privacy came into force on May 25, 2018. This video explains how it could affect you, even if you don't live in the EU. (Originally published May 16, 2018)

WeBank

Existing Approaches



user 1

user 2

......

......

user n

Central database

Model training

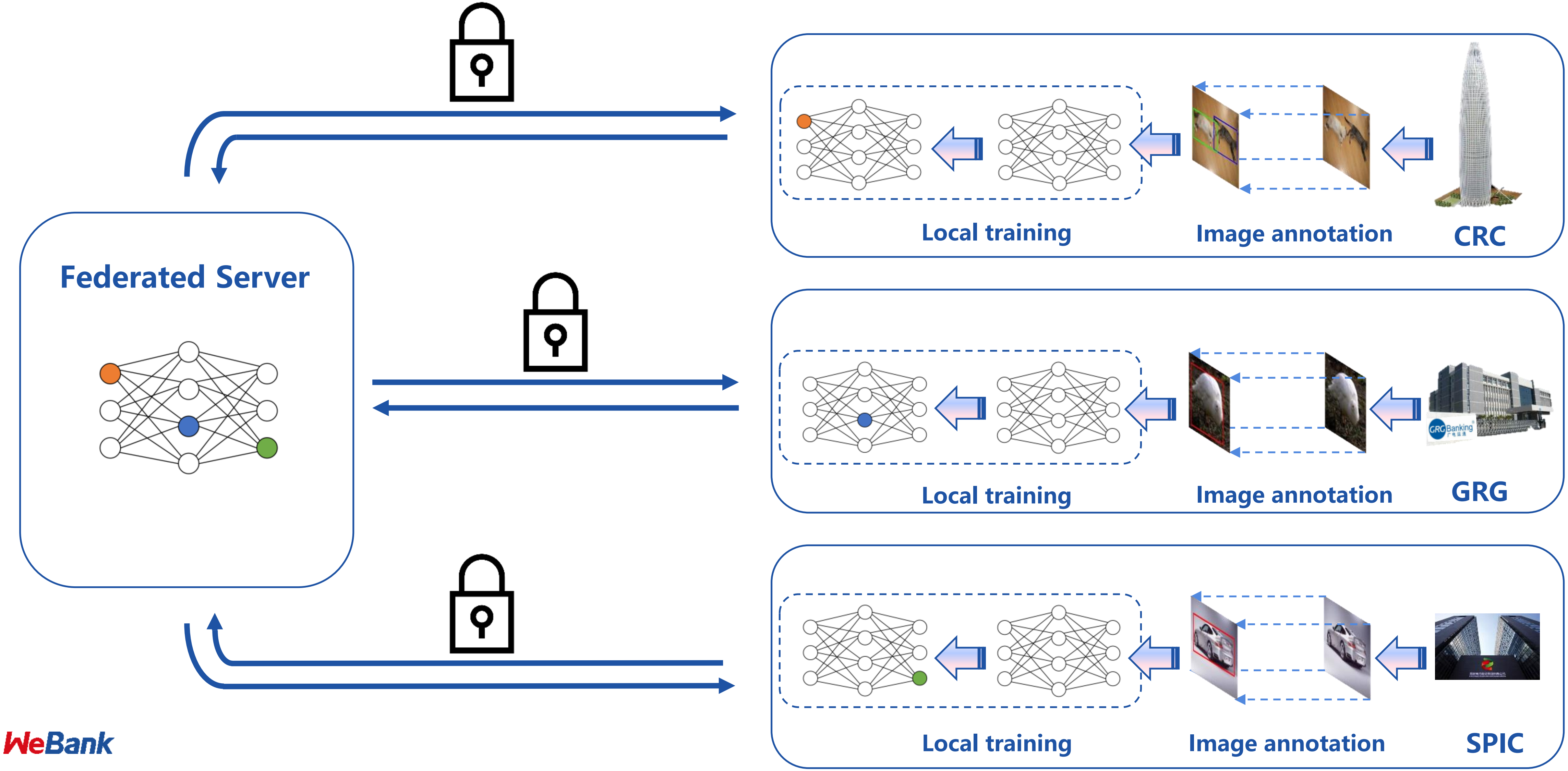
Online inference

Image Annotation

Fedvision – new machine learning framework

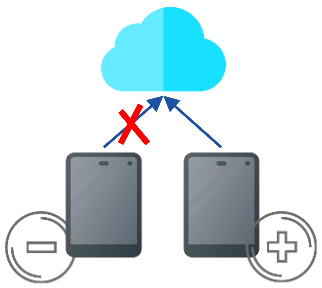
Decouple the need for model training with the need to store the data in the cloud or central database

FedVision approach



WeBank

System Challenges



Client change dynamically

1

Resource constraints

2

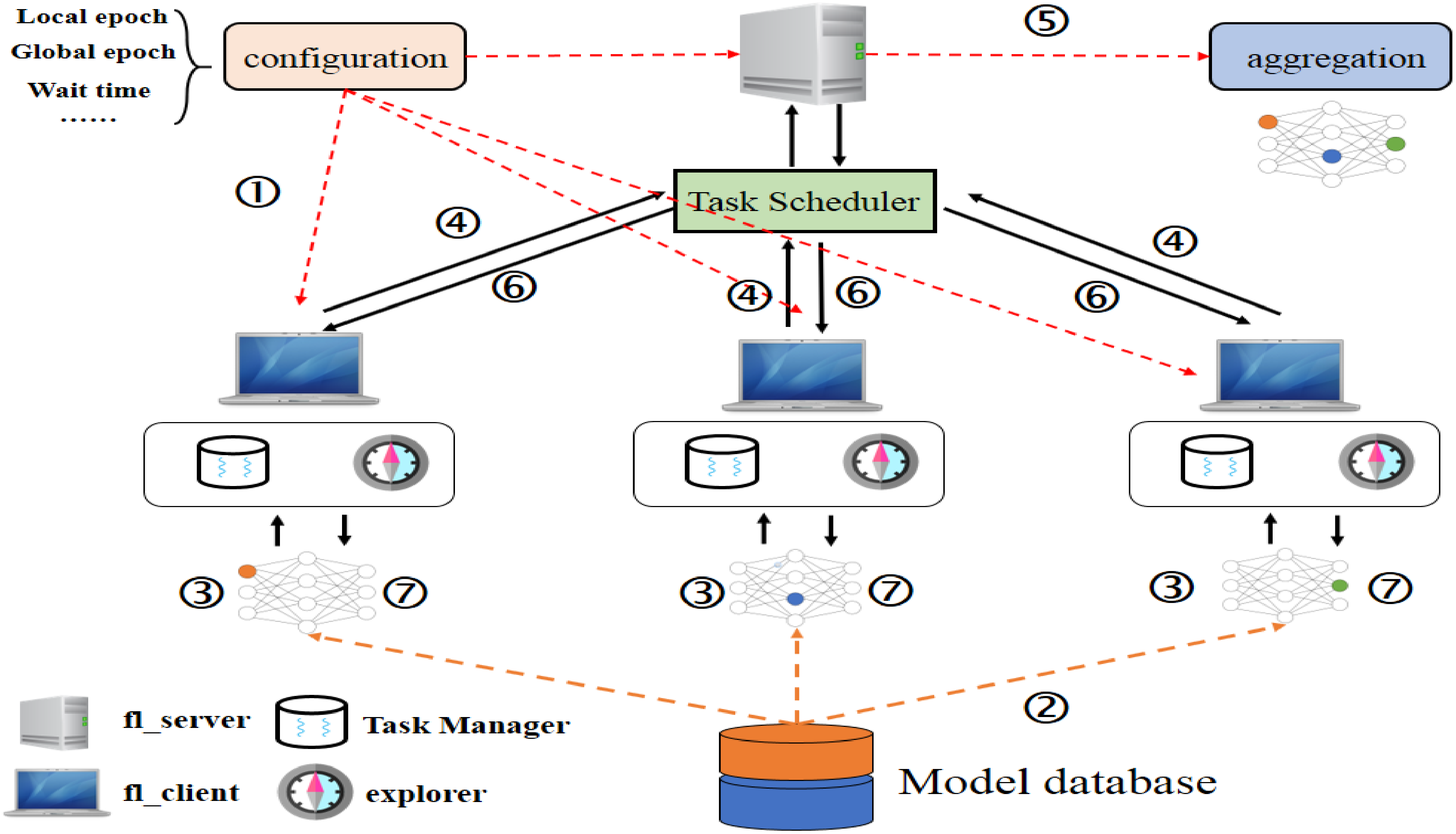
Device diversity

3

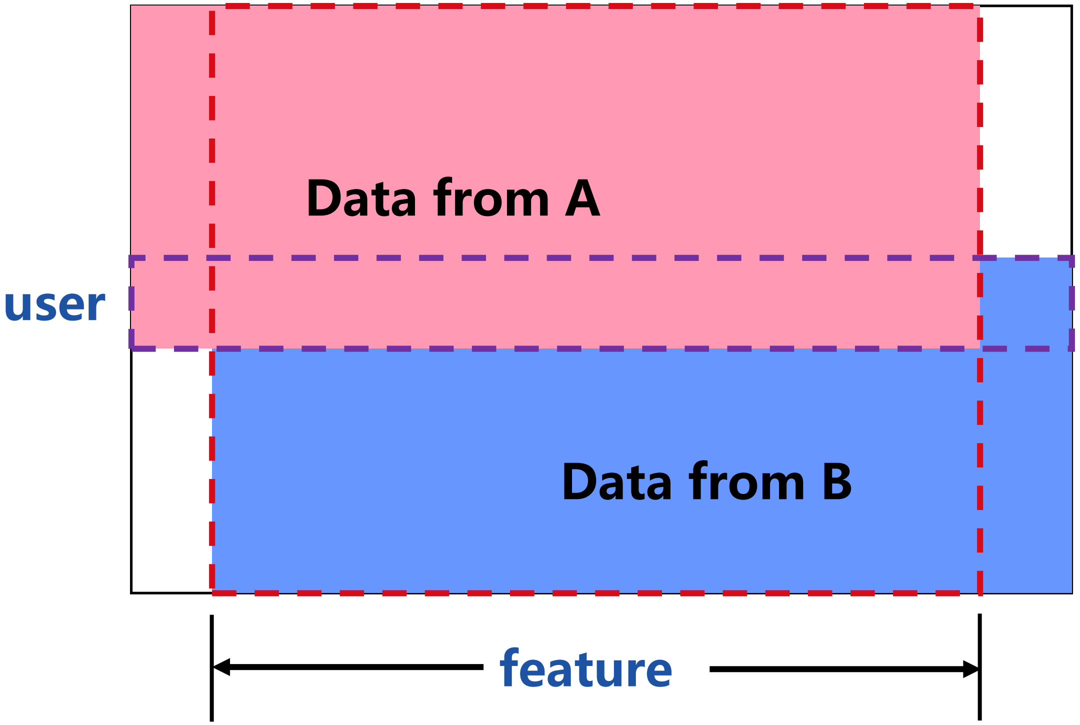
Network diversity

4

System Architecture



Use of AI Technology



- Enables different participants to collaboratively train machine learning model
- FL distributes the machine learning process over to the edge
- Keep dataset on device locally
- Each client has the same input features and model structure

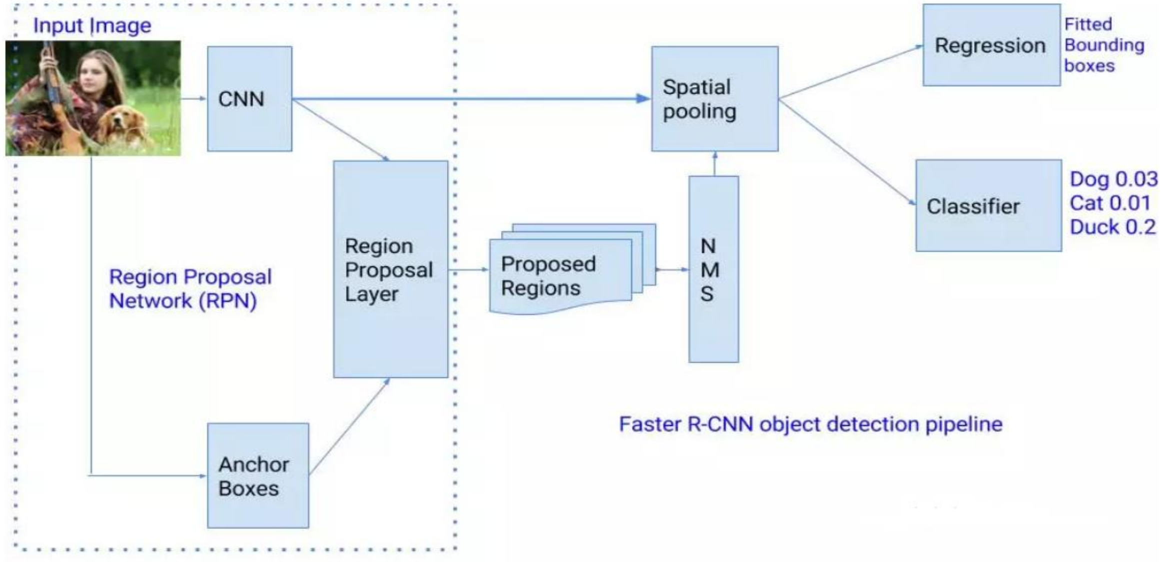
Horizontal federated learning

WeBank

Use of AI Technology

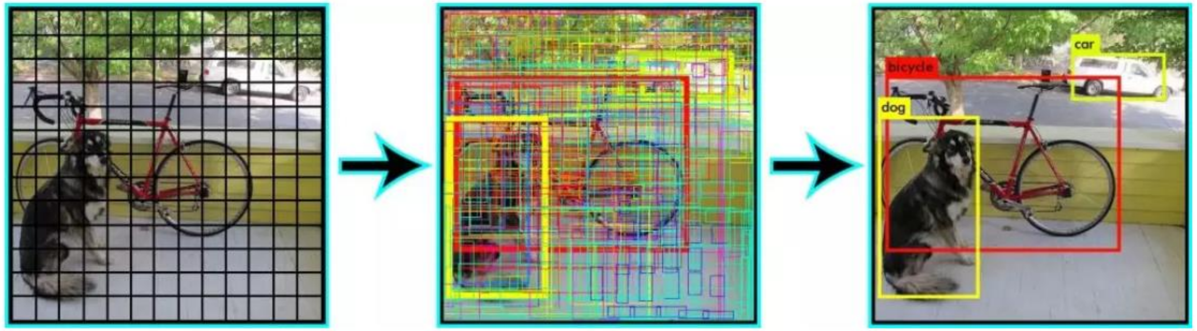
Object Detection

- Two-stage algorithm
  - Stage 1: select region proposal
  - Stage 2: execute bbox regression and classification
  - R-CNN, SPP-Net, Faster R-CNN ……



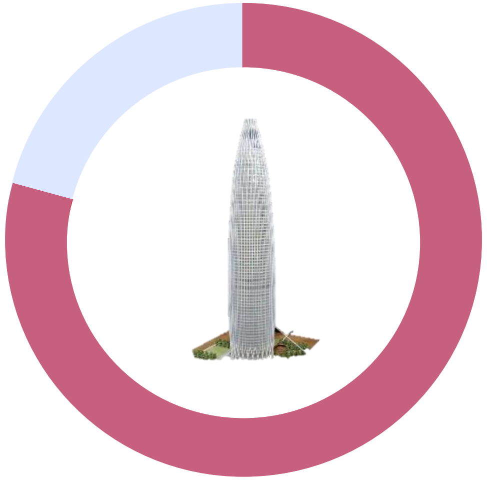
Faster R-CNN object detection pipeline

- One-stage algorithm
  - Take an input image and learns the class probabilities and bbox coordinates simultaneously
  - SSD, YOLO ……



Deployment and Payoff

70%



Efficiency

- CentVision: At least half a month to process and deploy.
- FedVision： real-time; he system administrator can finish labeling the images by himself.

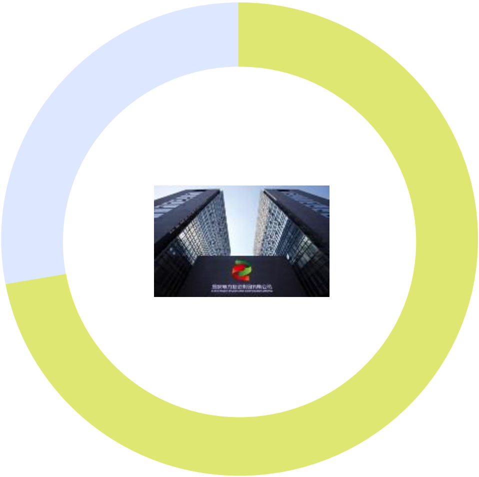
80%

GRG Banking

Privacy

- CentVision：send raw data to database, which had been proven unsafe and vulnerable to data leakage.
- FedVision： keep dataset on device locally, which can significantly mitigate many of the systemic privacy risks.

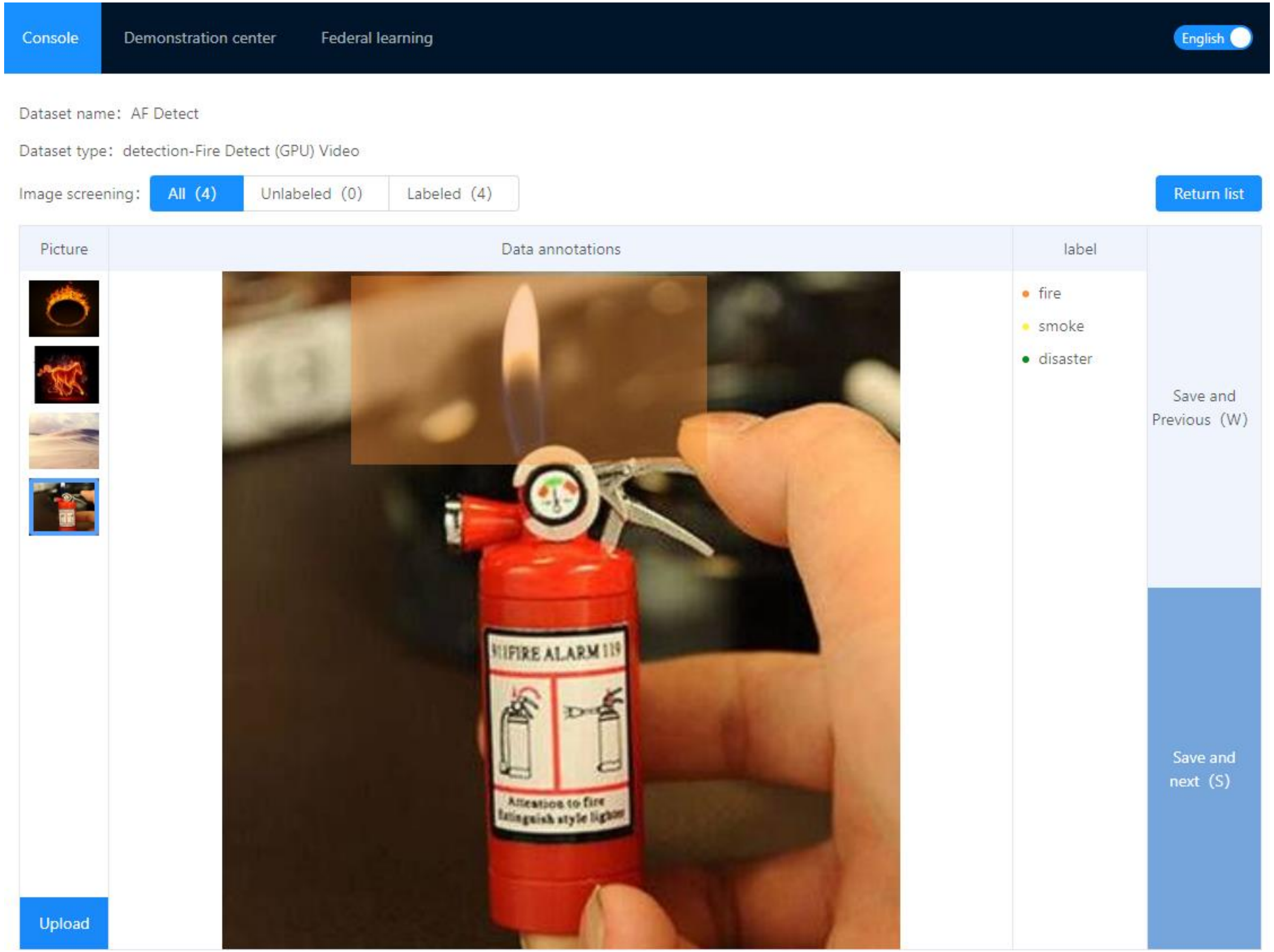
60%



Cost

- CentVision：a total of 100 channels required, these 100 channels require at least 50 MB/sec of network .
- FedVision： the network bandwidth required for model updateis significantly reduced to less than 1 MB/sec.

User Interaction Design



Dataset name：AF Detect

Dataset type：detection-Fire Detect (GPU) Video

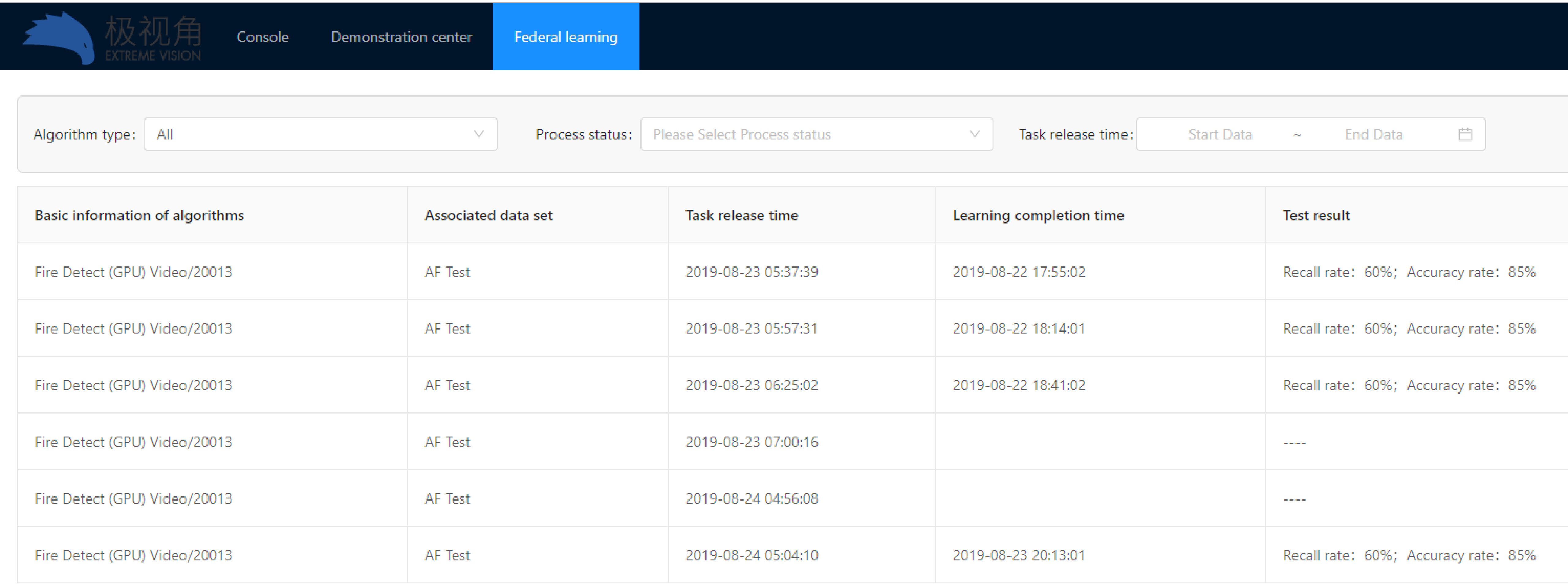
Image screening：All（4） Unlabeled（0） Labeled（4）

● fire

● smoke

● disaster

User Interaction Design



Algorithm type：All

Process status：Please Select Process status

Task release time：Start Data ~ End Data

Basic information of algorithms	Associated data set	Task release time	Learning completion time	Test result
Fire Detect (GPU) Video/20013	AF Test	2019-08-23 05:37:39	2019-08-22 17:55:02	Recall rate：60%；Accuracy rate：85%
Fire Detect (GPU) Video/20013	AF Test	2019-08-23 05:57:31	2019-08-22 18:14:01	Recall rate：60%；Accuracy rate：85%
Fire Detect (GPU) Video/20013	AF Test	2019-08-23 06:25:02	2019-08-22 18:41:02	Recall rate：60%；Accuracy rate：85%
Fire Detect (GPU) Video/20013	AF Test	2019-08-23 07:00:16		----
Fire Detect (GPU) Video/20013	AF Test	2019-08-24 04:56:08		----
Fire Detect (GPU) Video/20013	AF Test	2019-08-24 05:04:10	2019-08-23 20:13:01	Recall rate：60%；Accuracy rate：85%