Object Detection from Multi Sensor Data for Autonomous Underground Mine Mapping

Jim Moraga – Colorado School of Mines
May 22, 2019

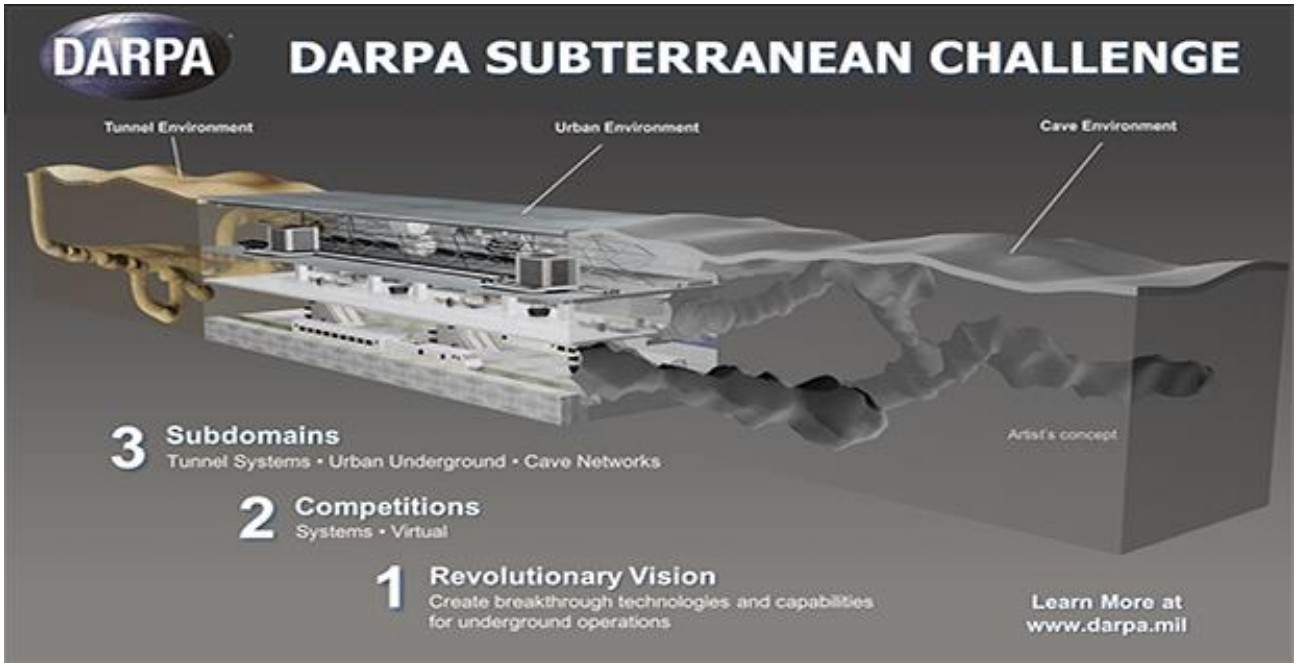
Why are we here?

Object Recognition in Underground environments

- Inspiration: SubT
- Sensors / Data
- Infrastructure
- Algorithms
  - Which ones
  - Data workflow
- Results
- Next Steps

Inspiration: SubT

SubT Challenge



- Goal of the DARPA Subterranean Challenge:
  - Discover innovative solutions that can rapidly and remotely map, navigate, and search complex environments, including human-made tunnel systems, urban and municipal underground infrastructure, and natural cave networks

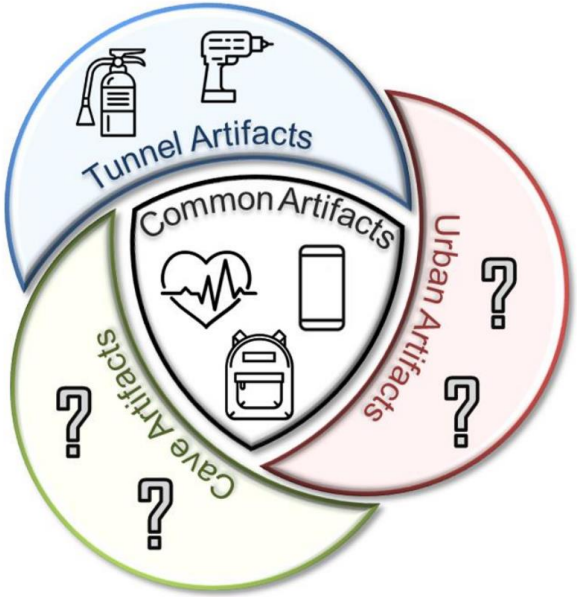
SubT Challenge

- Environment
  - The primary scenario of interest for the competition is providing rapid situational awareness to a small team of operators preparing to enter unknown and dynamic subterranean environments. The layout of the environment is unknown, could degrade or change over time (i.e., dynamic terrain), and is too high-risk to send in personnel.
- Scoring

Technology Area	Relevant Metrics	Stretch Goals
Autonomy	Number of manual interventions during mission	Zero interventions over four-hour mission
Perception	Resolution of multimodal 3D mapping	<10cm map detail
	Geo-localization of physical agents over mission duration	<1m error over 1km traversed
Networking	Latency in situational awareness updates per traversal distance	<1s per 500m path length
Mobility	Effective, team-aggregated mission endurance	Four-hour effective endurance

Object Detection

- Objects associated with test environment: survivors, extinguisher, backpack, etc
- Expected degraded sensing and communications: darkness, smoke, dust…
- Requires:
  - multiple sensors
  - multiple recognition modes
  - data fusion

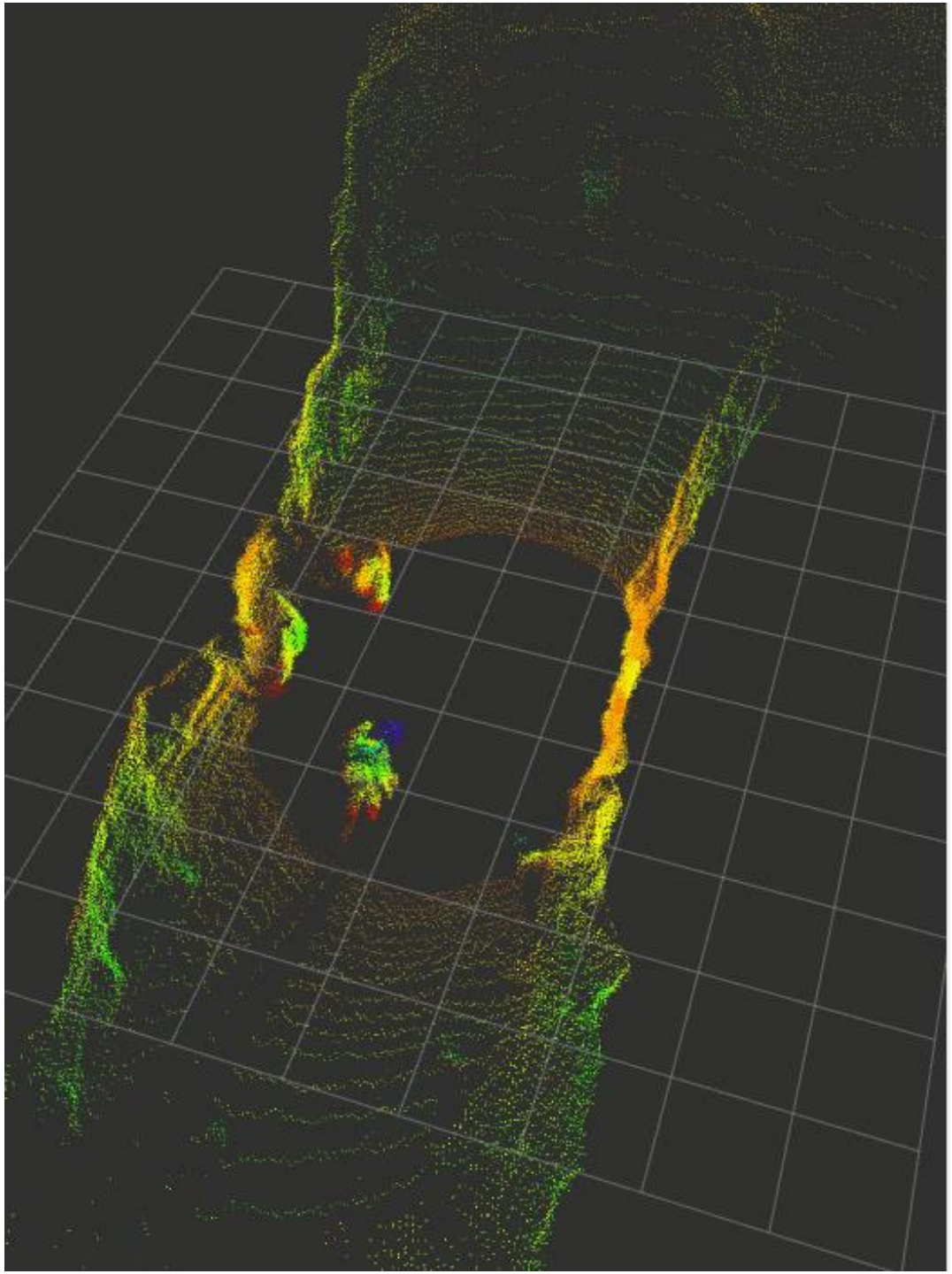


Sensors / Data

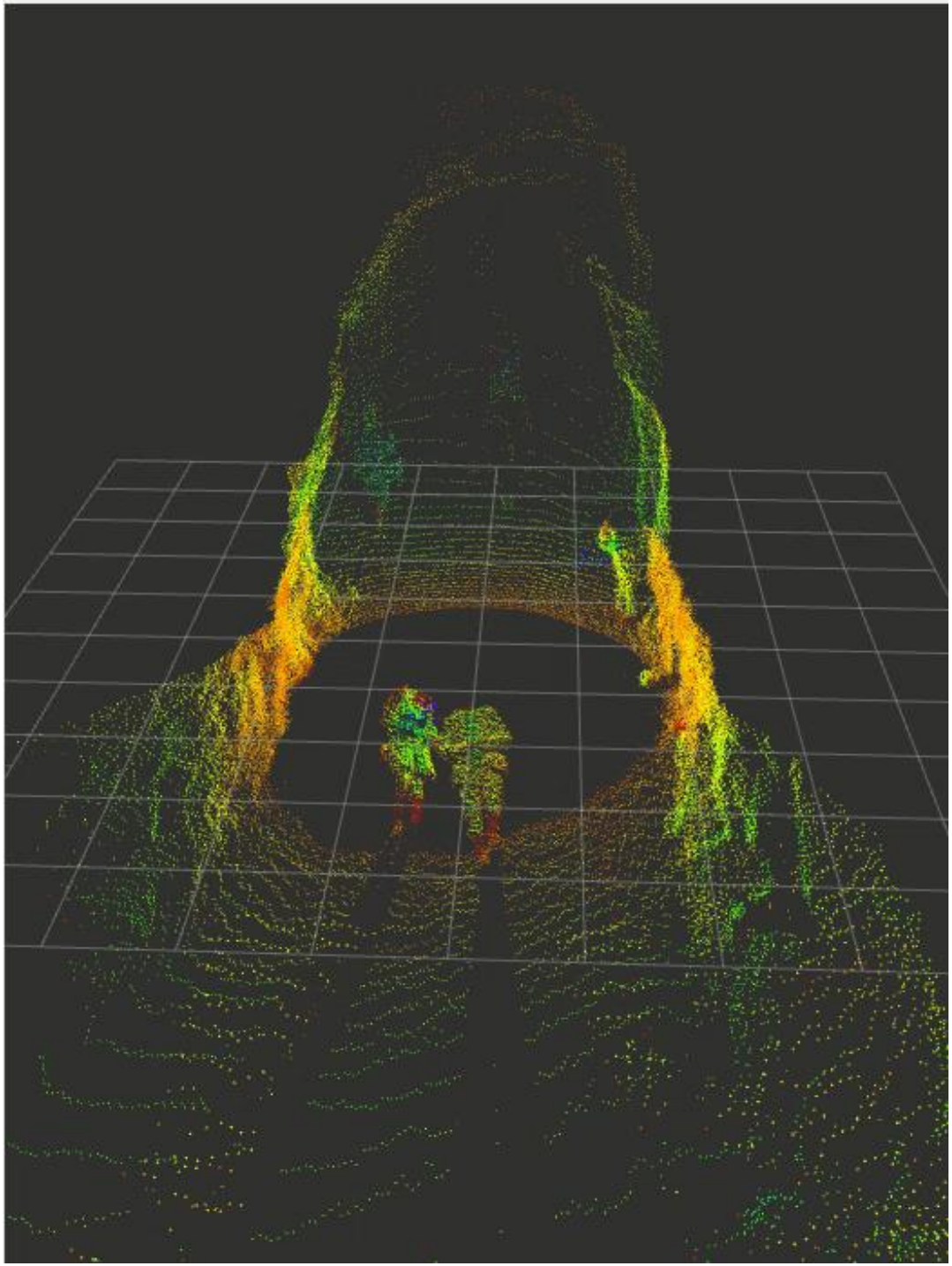
3D + 2D data capture in-situ

2D Objects – in degraded environment

3D: LIDAR

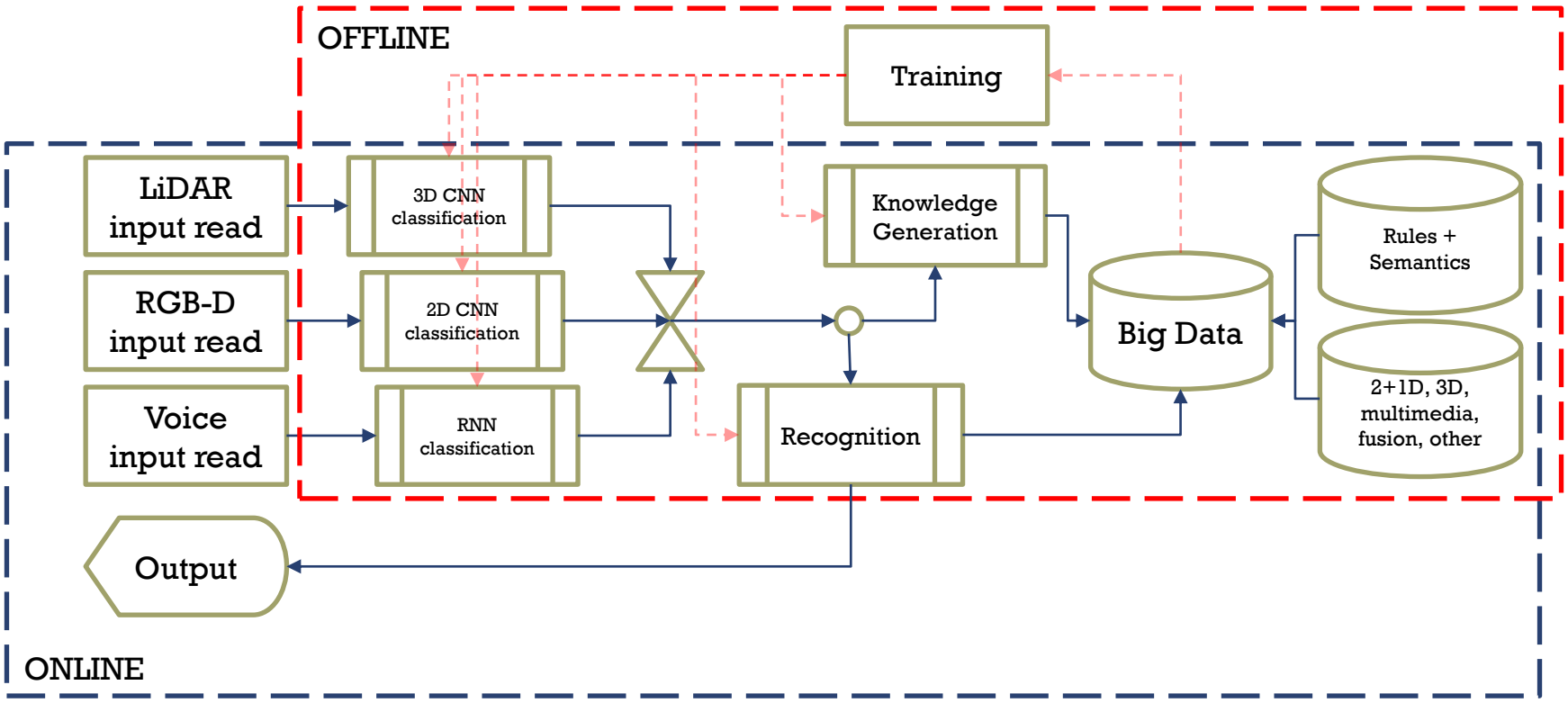


3D: LIDAR

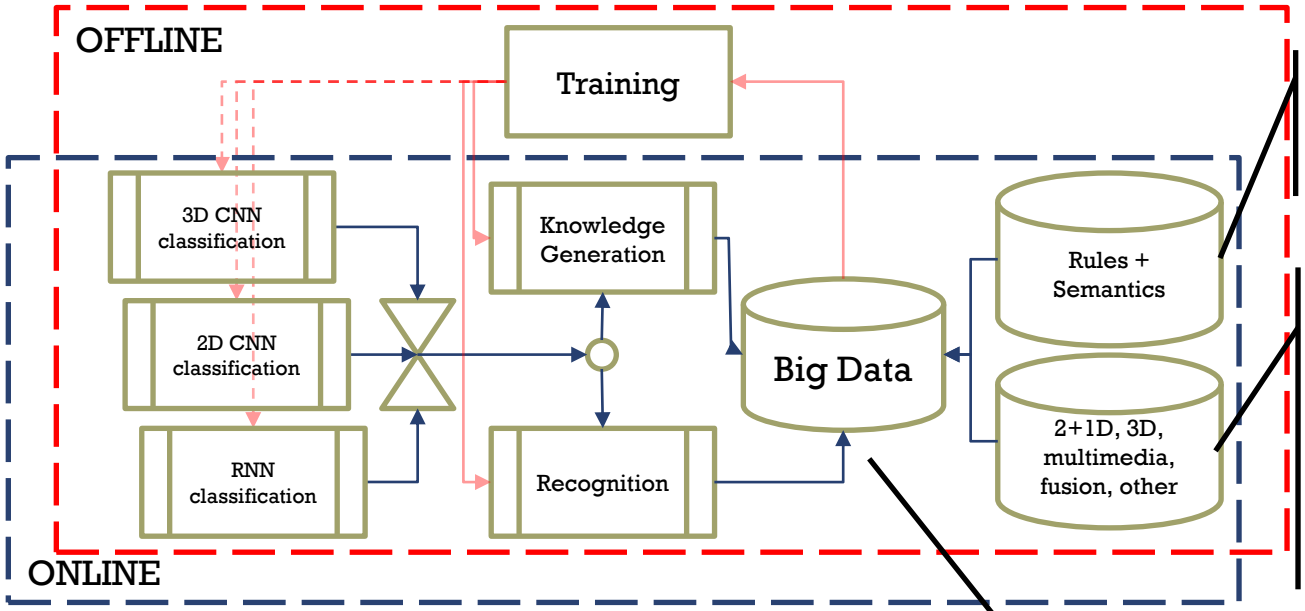


Infrastructure

Desired high level architecture



Detailed architecture – Data management



- Structured
- Graph/tree
- Small size

- Unstructured
- Audio & visual data
- Time stamp and data-fusion requirements
- Potential ML capabilities required
- NoSQL
- Large/very large size

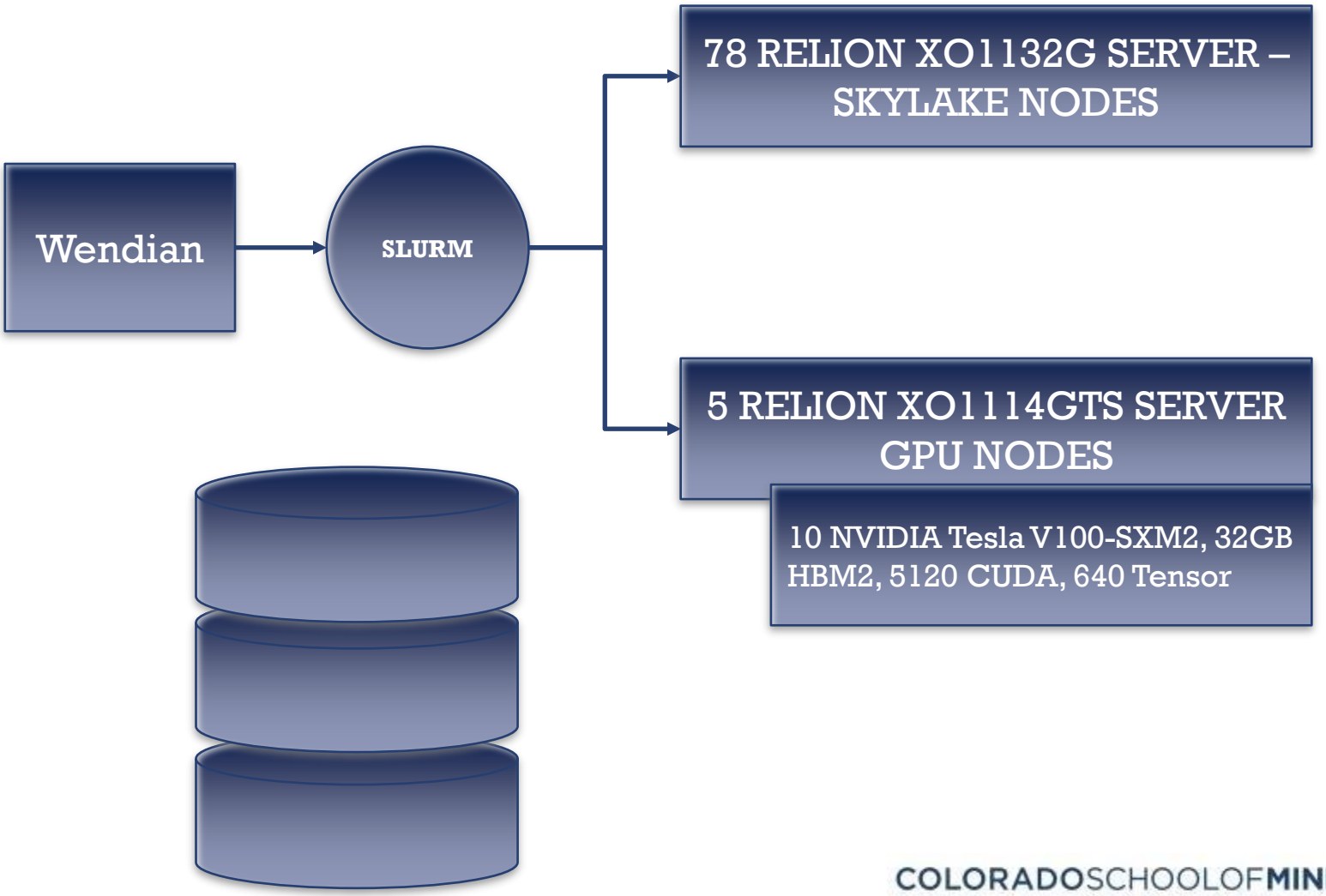
- Scalable
- High Throughput
- NoSQL capable
- Big Data

Options:
- Hadoop, Hive, Bigtable
- Hadoop tasks, Spark, Dataproc
- Must be Local

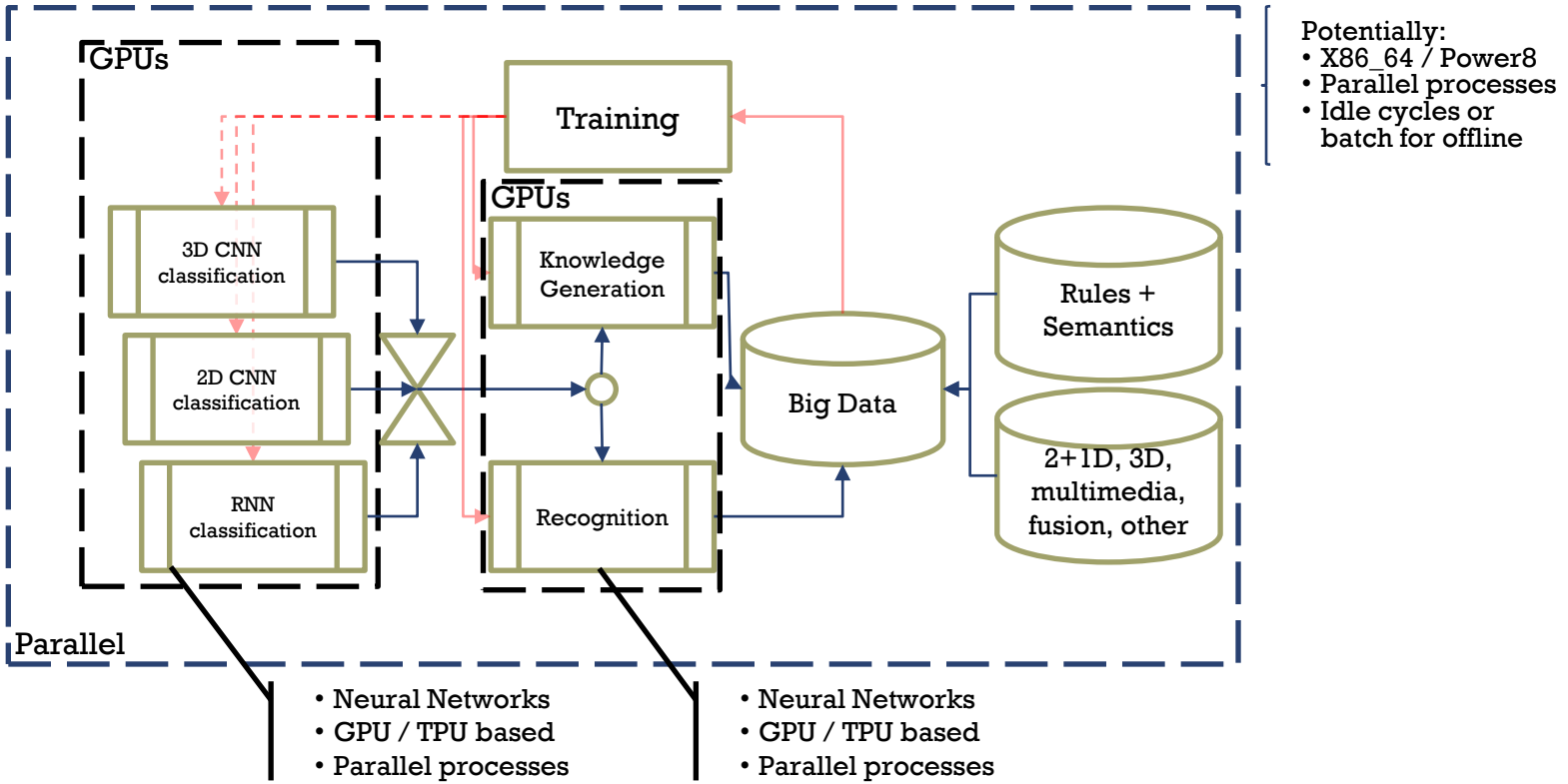
CSM – HPC Infrastructure

- MIO (150+ Tflops)
  - 200+ NODES – 8-28 CORES PER NODE
  - 2 POWER8 with 10 NVIDIA Tesla GPUS
- WENDIAN (350+ Tflops)
  - 78 RELION XO1132G SERVER – SKYLAKE NODES
  - 5 RELION XO1114GTS SERVER GPU NODES
    - 10 NVIDIA Tesla V100-SXM2, 32GB HBM2, 5120 CUDA, 640 Tensor

+ Wendian


Wendian
SLURM
78 RELION XO1132G SERVER – SKYLAKE NODES
5 RELION XO1114GTS SERVER GPU NODES
10 NVIDIA Tesla V100-SXM2, 32GB HBM2, 5120 CUDA, 640 Tensor

Detailed architecture – Processing



But… there are limitations in processing power for detection

+

Algorithms

Algorithms

- 2D
  - Yolo v3
    - Pros
    - Cons
- 3D
  - PointPillars
    - Pros
    - Cons
  - Frustum Pointnets

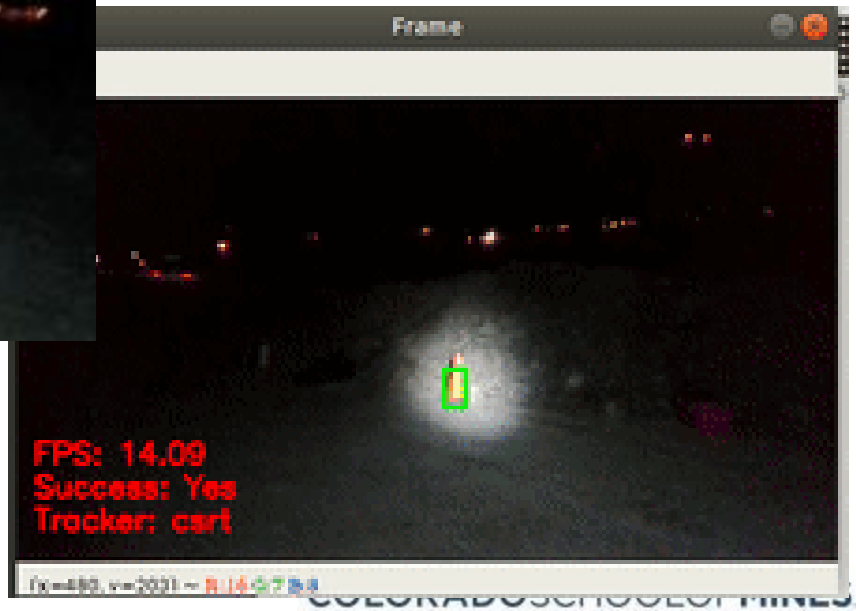
Data annotation workflow

- Data Capture:
  - RGB Camera, Velodyne Lidar
- Data Extraction (from ROSBAG)
  - JPG / PNG
  - Point clouds
- Data Selection
  - Manual process
- Manual Annotation
  - CVAT
- Conversion to Yolo format

Results

2D Objects – Recognition

- 2D object recognition can find objects in low light environments
- Finds objects that are hard to detect even for us



2D Objects – Latest Results

Next Steps

Objects in UG Mine Challenge

Common throughout all	Tunnel Circuit
Survivor	Hand Drill
Backpack	**Fire Extinguisher**
Cell Phone	

Gaps in data & processing

- Annotating Edgar Mine point clouds
  - 6,000+ images with multiple objects
- Annotating Edgar Mine pictures
  - Should also include 3D rotation / translation matrices
- Data augmentation
  - Virtual versions of backpack, hand drill, survivor, cell phone
  - Generation of multiple poses for each object
  - Simulation to generate 2D and 3D data and annotations
- Big data infrastructure
- Algorithms:
  - 3D object detection (with or without RGB images)
  - 2D object detection (accelerate recognition from ~2Hz)

+

Questions?

Jim Moraga
jmoraga@mines.edu