Apache Spark 中国技术... 全员

该群属于“Apache Spark 中国技术交流社区”全员群，仅组织内部成员可以加入，如果组织外部人员收到此分享，需要先申请加入该组织。

Building AI to Play FIFA* Video Game Using Distributed TensorFlow* on Analytics Zoo

Shengsheng Huang, Shan Yu, Jason Dai

Collaborations with Shanghai Jiao Tong University

AGENDA

- Distributed TF on Apache Spark* using Analytics Zoo
- RL Platform for Playing FIFA18
- Playing FIFA18 using Imitation Learning & DRL
- Experimenting with GRF (Google Research Football*)

AGENDA

- **Distributed TF on Apache Spark using Analytics Zoo**
- RL Platform for Playing FIFA18
- Playing FIFA18 using Imitation Learning & DRL
- Experimenting with GRF (Google Research Football)

What is Analytics Zoo

BigDL

Distributed, High-Performance

Deep Learning Framework

for Apache Spark

https://github.com/intel-analytics/bigdl

ANALYTICS ZOO

Unified Analytics + AI Platform

Distributed TensorFlow, Keras, PyTorch and BigDL on Apache Spark

https://github.com/intel-analytics/analytics-zoo

Accelerating Data Analytics + AI Solutions At Scale

intel AI

INTEGRATED BIG DATA ANALYTICS AND AI

SEAMLESS SCALING FROM LAPTOP TO PRODUCTION

- Easily prototype the end-to-end pipeline
- "Zero" code change from laptop to distributed cluster
- Directly access production data without data copy
- Seamlessly deployed on production big data clusters

Analytics Zoo

Unified Big Data Analytics and AI Platform

ANALYTICS ZOO

Models & Algorithms	Recommendation	Time Series	Computer Vision	NLP
ML Workflow	AutoML for Time Series	Automatic Cluster Serving		
Integrated Analytics & AI Pipelines	Distributed TensorFlow & PyTorch on Spark	RayOnSpark		
	Spark Dataframes & ML Pipelines for DL	Model Serving		

Library & Framework

- Distributions (Cloudera/Databricks/....)
- Distributed Analytics (Spark/Flink/Ray/...)
- DL Frameworks (TF/PyTorch/...)
- Python Libraries (Numpy/Pandas/...)

https://github.com/intel-analytics/analytics-zoo

intel AI

DISTRIBUTED TENSORFLOW ON SPARK IN ANALYTICS ZOO

```
#pyspark code
train_rdd = spark.hadoopFile(…).map(…)
dataset = TFDataset.from_rdd(train_rdd,…)

#tensorflow code
import tensorflow as tf
slim = tf.contrib.slim
images, labels = dataset.tensors
with slim.arg_scope(lenet.lenet_arg_scope()):
   logits, end_points = lenet.lenet(images, …)
loss = tf.reduce_mean( \
   tf.losses.sparse_softmax_cross_entropy( \
   logits=logits, labels=labels))

#distributed training on Spark
optimizer = TFOptimizer.from_loss(loss, Adam(…))
optimizer.optimize(end_trigger=MaxEpoch(5))
```

Write TensorFlow code inline in PySpark program

MORE INFORMATION ON ANALYTICS ZOO

- Project website
  - https://github.com/intel-analytics/analytics-zoo
  - https://github.com/intel-analytics/bigdl
- Tutorials
  - CVPR 2018: https://jason-dai.github.io/cvpr2018/
  - AAAI 2019: https://jason-dai.github.io/aaai2019/
- "BigDL: A Distributed Deep Learning Framework for Big Data"
  - *In proceedings of ACM Symposium on Cloud Computing 2019 (SOCC'19)*
- Use cases
  - *Azure, CERN, MasterCard, Office Depot, Tencent, Midea, etc.*
  - https://analytics-zoo.github.io/master/#powered-by/

AGENDA

- Distributed TF on Apache Spark using Analytics Zoo
- **RL Platform for Playing FIFA18**
- Playing FIFA18 using Imitation Learning & DRL
- Experimenting with GRF (Google Research Football)

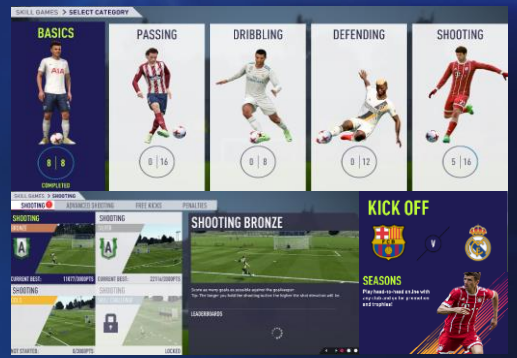
WHY FIFA18?

What is FIFA18*?

- **A real-time 3D soccer simulation video game by Electronic Arts***

Why FIFA18?

- **It's fun ☺**
- **It's challenging**
  - **Complex (esp. full-court game) and non-deterministic**
  - **Large action space (16 basic keys w/ combinations)**
- **Many modes available**
  - **Full-court, mini-games, skill games, etc.**



SHOOTING BRONZE: OUR EXPERIMENT ENVIRONMENT

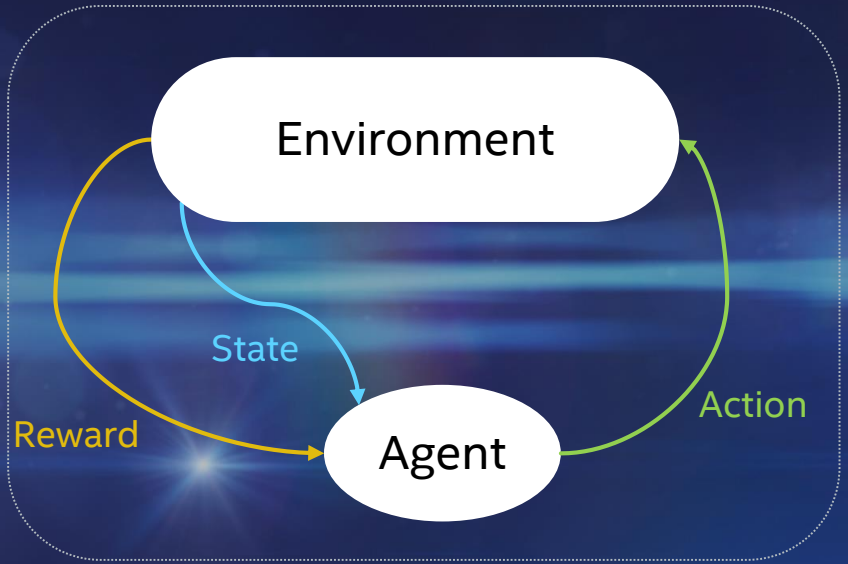
Shooting is one of the mini-games in FIFA18,
Bronze is the easiest level

Game mode
- Player & goalkeeper 1v1
- Goal: get higher score in 44s

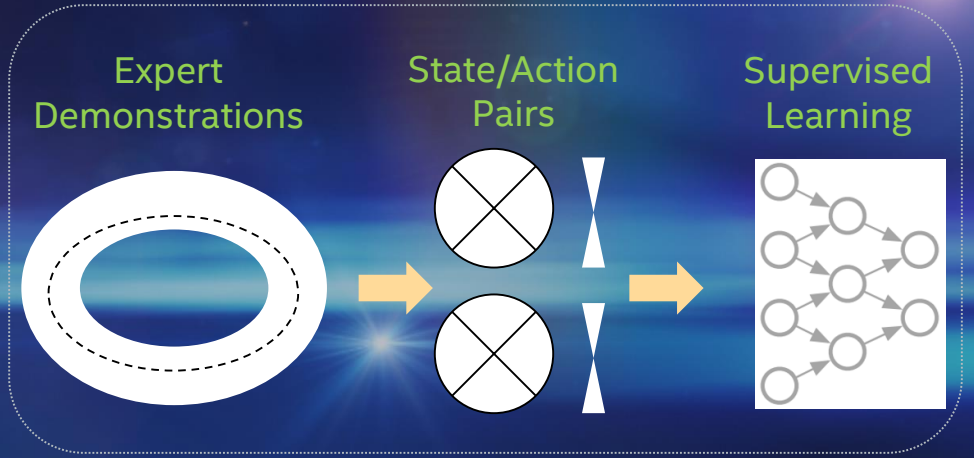
Evaluation
- Single shoot: score ≤ 200 for miss; 200<score<1200 for goal
- Accumulated scores after the game

Keyboard control
- A/S/W/D: left/right/up/down
- Space: shoot

REINFORCEMENT LEARNING



Sequential Decision Making



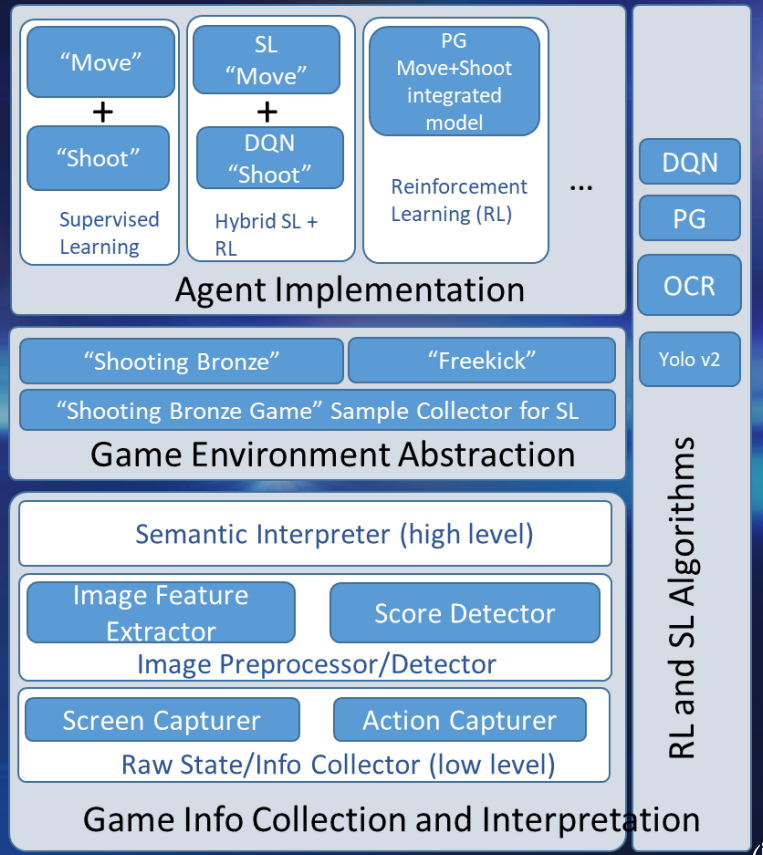
Imitation Learning

RL PLATFORM FOR PLAYING FIFA18

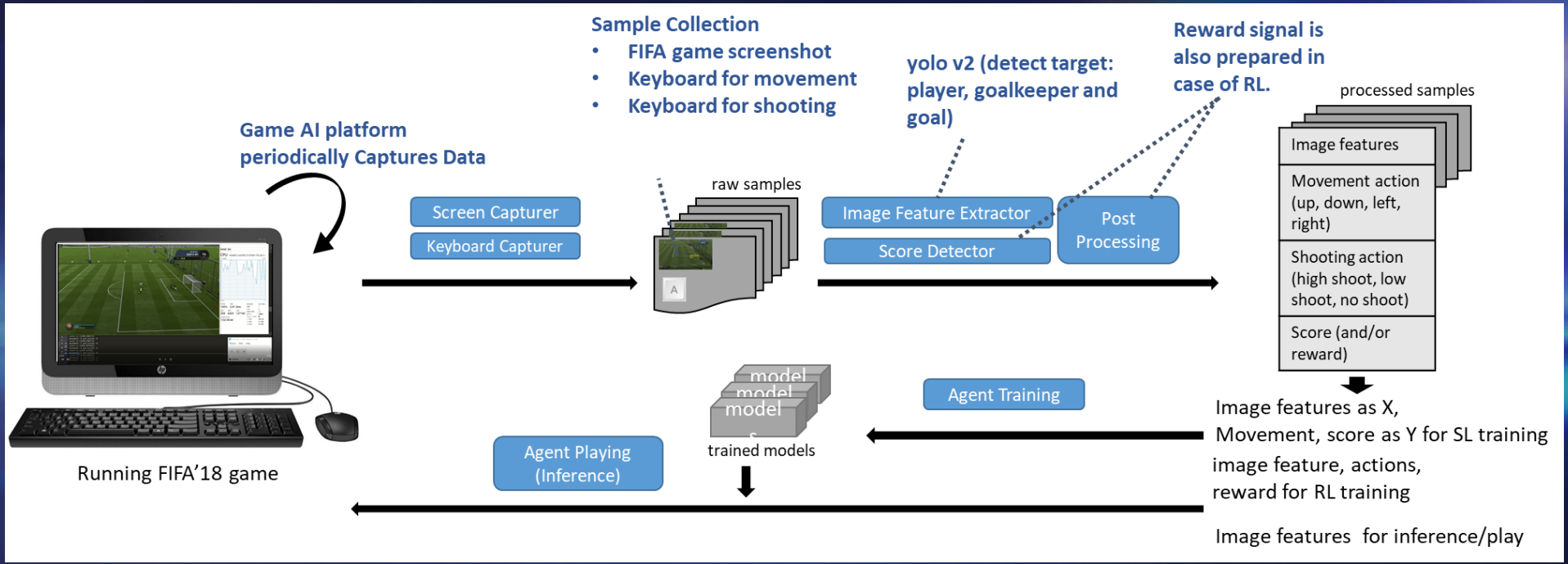
Experiment platform for RL agents and algorithms for FIFA18

Major components

- Game info collection & Interpretation
- Game Environment Abstraction
- Agent Implementation
  - Imitation learning / supervised learning (SL)
  - Reinforcement learning (RL)
  - Hybrid (SL+RL)



END-TO-END WORKFLOW



ANALYTICS ZOO

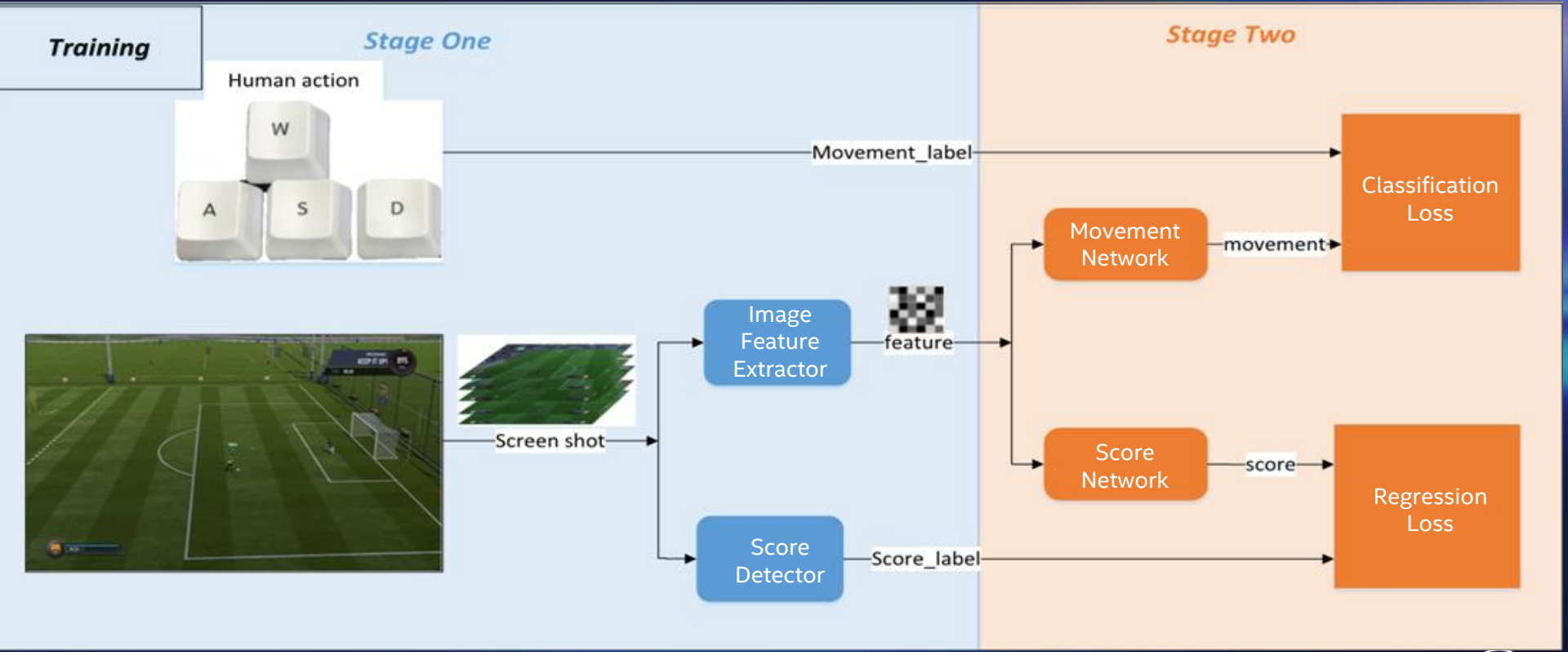
tfpark: Distributed TensorFlow on Spark

intel AI

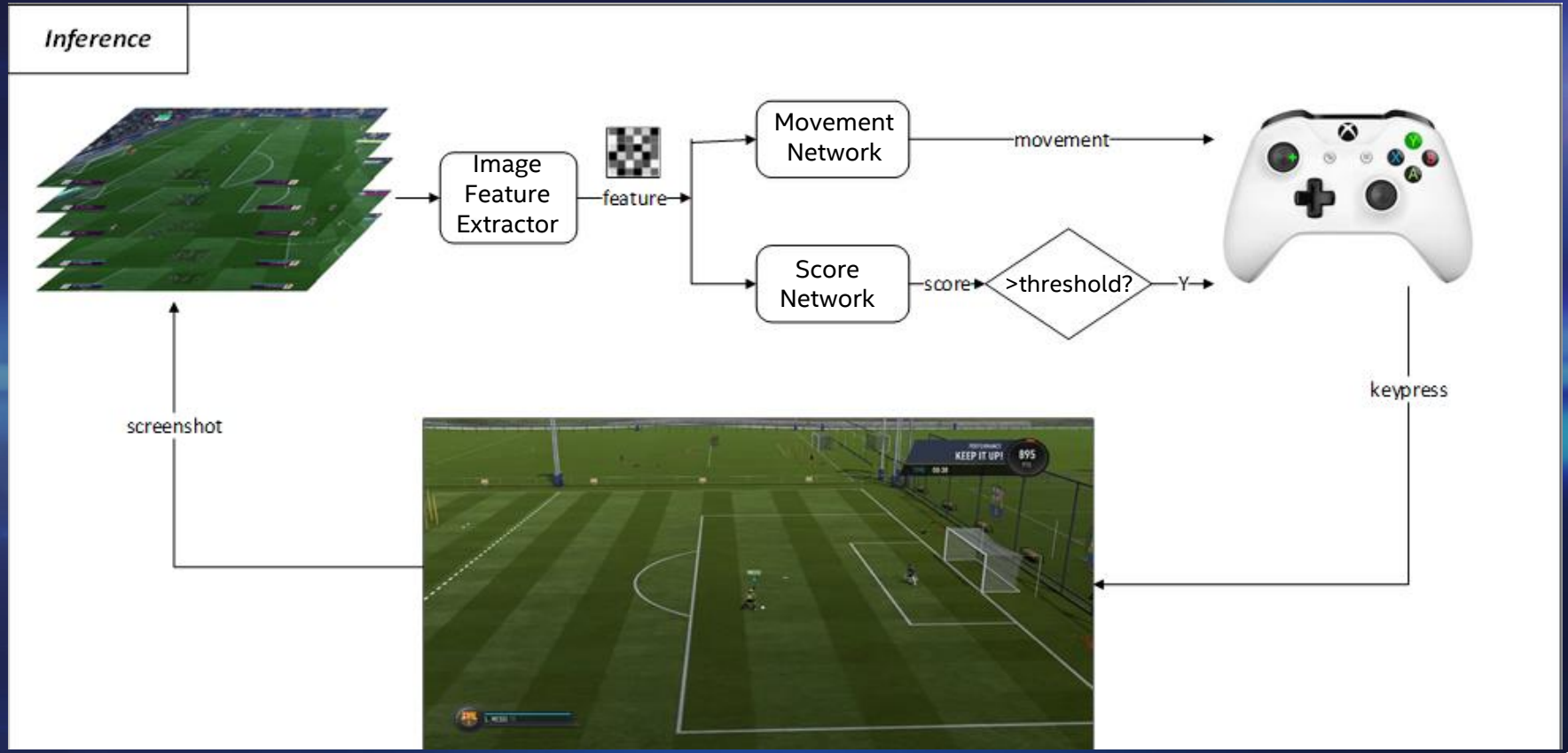
AGENDA

- Distributed TF on Apache Spark using Analytics Zoo
- RL Platform for Playing FIFA18
- **Playing FIFA18 using Imitation Learning & DRL**
- Experimenting with GRF (Google Research Football)

TRAINING THE AGENT USING IMITATION LEARNING



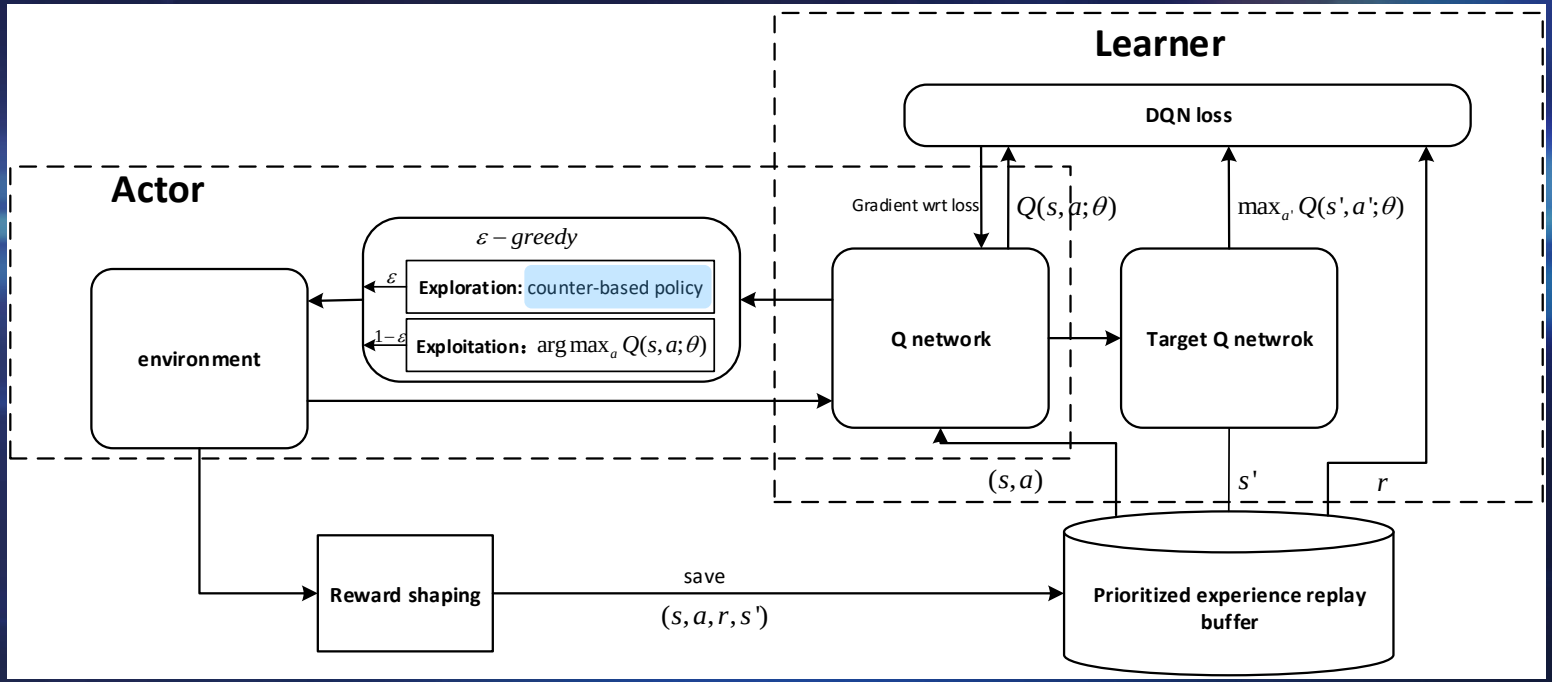
GAME PLAYING (INFERENCE) FOR IMITATION LEARNING



HYBRID APPROACH FOR TRAINING AGENT

Movement network: trained with Imitation Learning

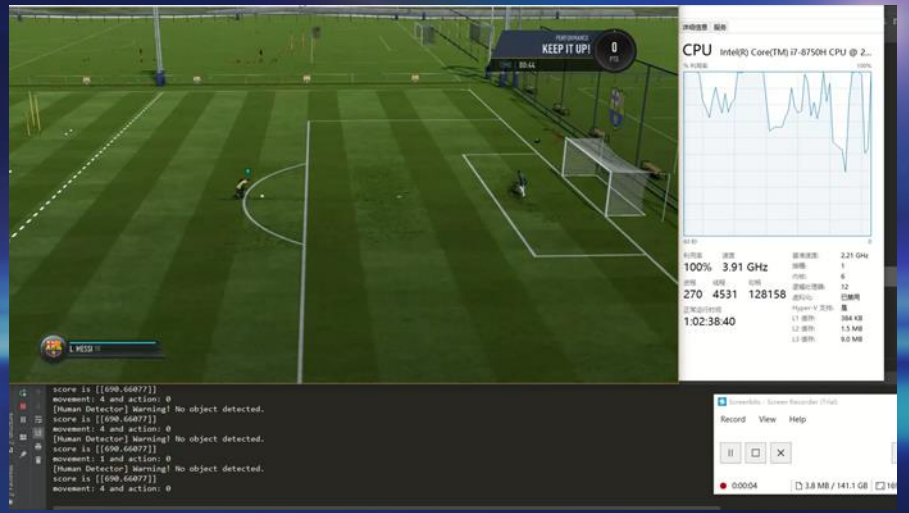
Shoot network: Double DQN



DEMO

https://drive.google.com/file/d/13dBsGOiGbCYOS5TgVAI95Qd-YszAHTW6/view

Human (demonstrator)



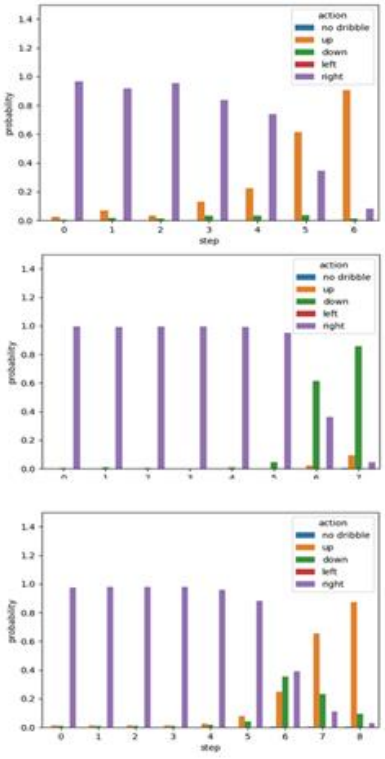
https://drive.google.com/file/d/1JVZjlDSyX8YtUy6qOuGRD_VN4RSZw0U8/view

Imitation Learning (better score than demonstrator)

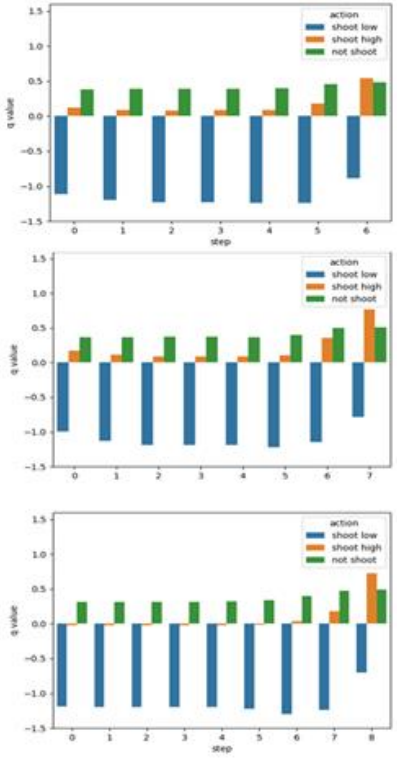
TYPICAL TRAJECTORY ANALYSIS (HYBRID)

Typical Trajectories

Movement Policy (SL)



Shoot Q-Value (RL)



RESULTS

		Score	Goal Ratio	Convergence speed
Human	beginner	5846.69	50%	-
	master	10112.78	92%	-
	demonstrator	7284.98	84.96%	-
Agent	Imitation Learning	10345.18	92.54%	-
	RL (Policy Gradient)	5606.31	40.25%	1069.5 epochs
	Hybrid	10514.43	95.59%	749.6 epochs

For more complete information about performance and benchmark results, visit www.intel.com/benchmarks.

AGENDA

- Distributed TF on Apache Spark using Analytics Zoo
- RL Platform for Playing FIFA 2018
- Playing FIFA using Imitation Learning & DRL
- Experimenting with GRF (Google Research Football)

GOOGLE RESEARCH FOOTBALL (GRF)

An open source RL environment for playing soccer from Google Brain

- https://github.com/google-research/football

A great RL environment for playing soccer

- More state and reward info & controls
- Customizable scenarios, players, rewards and observations, etc.
- More useful features such as accelerated speed, self-play, multi-agent, etc.
- Easy to dump traces and replay

Transfer between FIFA18 and GRF?

Google Research Football: A Novel Reinforcement Learning Environment (https://arxiv.org/abs/1907.11180)

EARLY EXPERIMENTS ON GRF

https://drive.google.com/file/d/1bNO5rpUhCeCZY9zPGgVCzgUlqH9QF39n/view

Trained using PPO in OpenAI* baseline

FUTURE WORK

Ray* support in Analytics Zoo
- E.g., *RayOnSpark*

Support for *Google Research Football*
- *E.g., transfer between GRF and FIFA?*

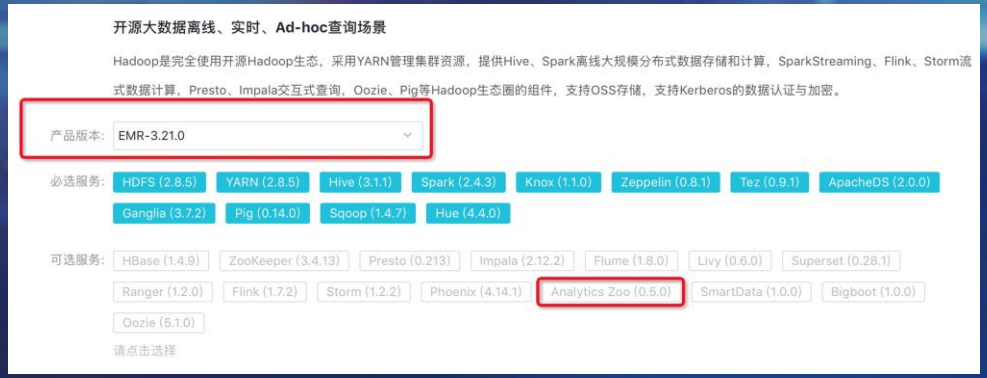
Additional algorithms/models and scenarios
- E.g., *full-court* game

https://medium.com/riselab/rayonspark-running-emerging-ai-applications-on-big-data-clusters-with-ray-and-analytics-zoo-923e0136ed6a

Analytics Zoo on Ali E-MR

ANALYTICS ZOO + Alibaba Cloud aliyun.com

Analytics Zoo is already out-of-box on Ali EMR:

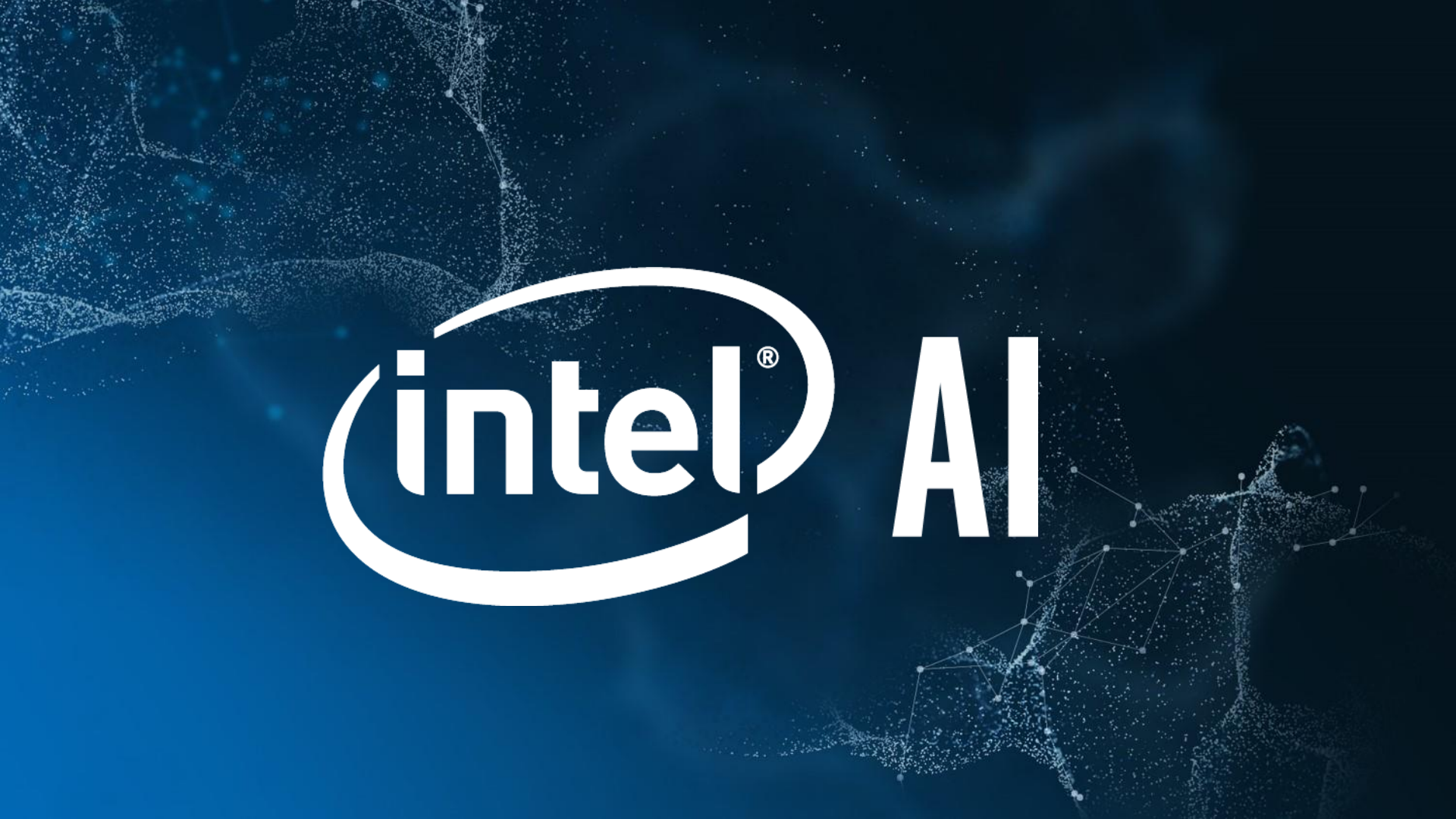


* Version upgrade for Analytics Zoo is on-going.

For more information and support, contact Wesley:

Email: wesley.du@intel.com
DingTalk:

intel AI


intel®
AI

LEGAL NOTICES AND DISCLAIMERS

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit intel.com/performance.

Intel does not control or audit the design or implementation of third-party benchmark data or websites referenced in this document. Intel encourages all of its customers to visit the referenced websites or others where similar performance benchmark data are reported and confirm whether the referenced benchmark data are accurate and reflect performance of systems available for purchase.

Optimization notice: Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software, or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com/benchmarks.

intel AI

Apache Spark 中国技术... 全员

该群属于“Apache Spark 中国技术交流社区”全员群，仅组织内部成员可以加入，如果组织外部人员收到此分享，需要先申请加入该组织。