ANALYTICAL AI

DATA TO INTELLIGENCE

www.analyticalai.com

AGENDA

BACKGROUND

SUMMARY

FEATURE SELECTION AND ENGINEERING

TRAINING METHODS

IMPORTANT FINDINGS

SIMPLE MODEL

DHS Challenge

Passenger Screening Algorithm Challenge

Winner Presentation
3rd Place Team

Participants

David Odaibo

Thomas Anthony

OVERVIEW

We utilized Convolutional and Recurrent Neural Networks and refined the annotations provided for the completion adding additional detail.

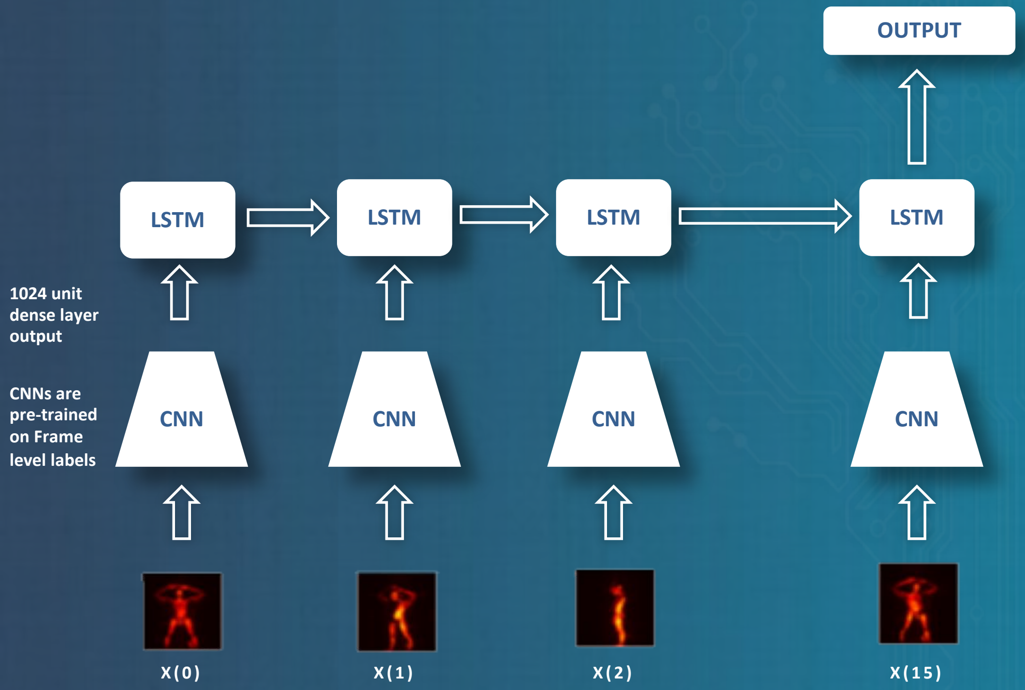
Our single best model was sufficient to maintain our 3rd place finish in the competition.

Our team has developed algorithms for managing 3D CT and MRI data

Our motivation for entering was to assess the generalizability of these algorithms in a new domain

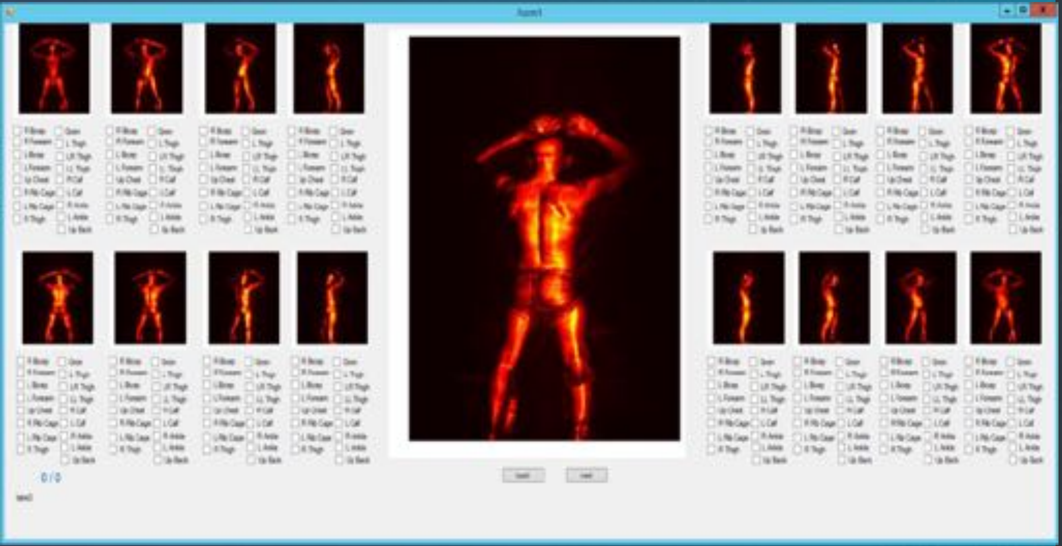
Participating gave us a benchmark of our algorithms compared to the field in one of our domains of expertise – 3D threat assessment

SUMMARY



- Combined Convolutional and Recurrent Neural Networks
- APS (projected image angle sequence data) gave us the best results.
- It was important for us to create image level annotations and pre-train CNN before training CNN and RNN jointly
- We utilized heavy data augmentation (salt/pepper noise, various distortions) this helped with over fitting
- We did not observe any over fitting, and training loss was closely correlated with our validation loss, so we train for as long a we could. (many days)

FEATURES SELECTION/ENGINEERING



The competition data annotations were provided on a subject level

Using image level annotation enabled us work with neural network architectures that utilized 2d convolutions rather that 3d convolutions, making a solution feasible at the original image resolution

Our prior experience in medical image analysis led us to believe we needed to create image level annotations

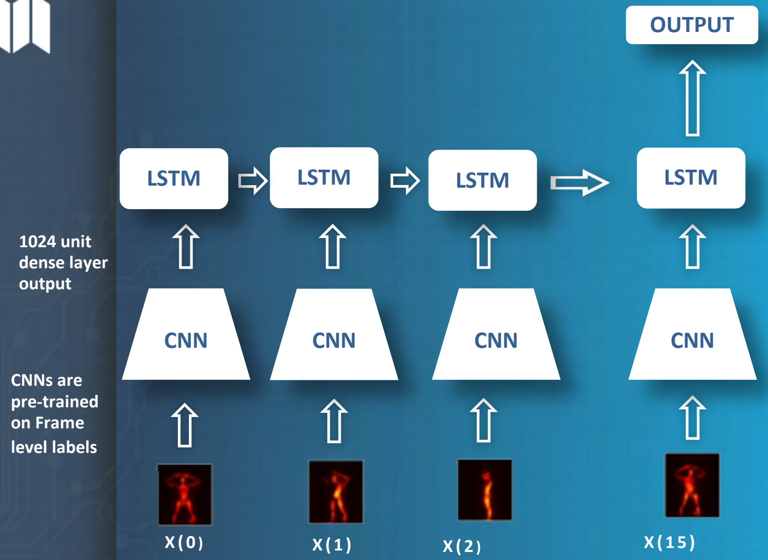
We created a utility to help create image level annotations of the training data

TRAINING METHODS

The Models were trained in 2 phases

Phase I
Train CNN with image level labels

Phase I - trained the CNN with our image level annotations on 2d images

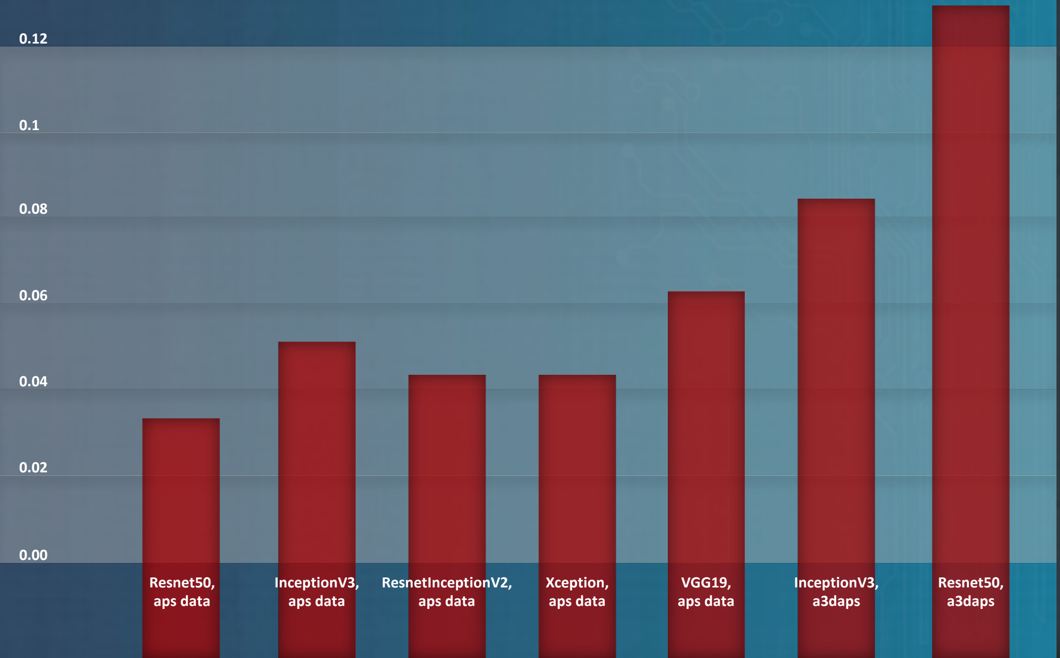


Phase II
Initialize CNN with weights from phase 1 and add LSTM

Phase II - added a RNN to capture temporal features in the 3rd dimension and initialized the CNN with the weights from Phase 1

FEATURES SELECTION/ENGINEERING

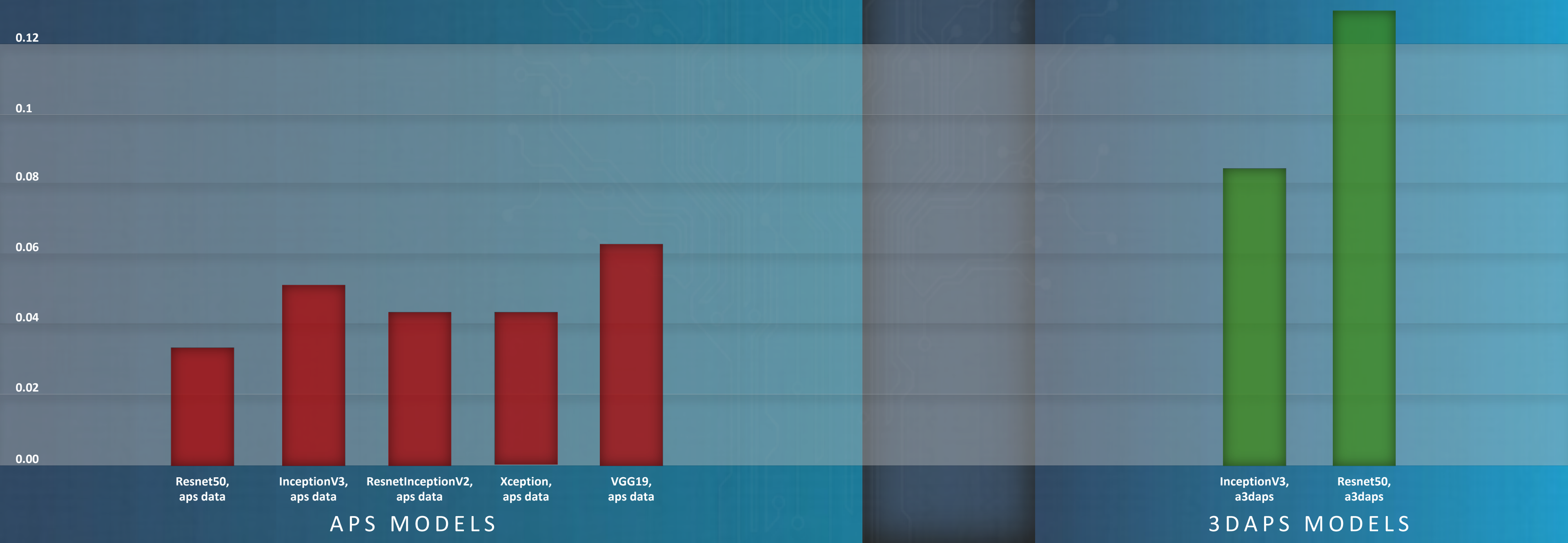
Log-loss of 0.03042 is approximate to an estimate of 1 false positive in 1135 passengers with exactly one threat in one of 17 possible body zones.



Model	Final Test Data log loss
CNN Model 1, aps data	0.03576 (best single model)
CNN Model 2, aps data	0.04341
CNN Model 3, aps data	0.04403
CNN Model 4, aps data	0.04896
CNN Model 5, aps data	0.06704
CNN Model 6, a3daps data	0.08744
CNN Model 7, a3daps data	0.13288
Ensemble of 7 models	0.03042 (ensemble of 7 models)

IMPORTANT AND INTERESTING FINDINGS

The smaller APS dataset performed much better than the larger 3DAPS and A3D data



Strengths, Weakness, and Possible Improvements

A major strength of our approach is that we do not exploit any weakness in the competition design and data collection process.

Some top finishing teams exploited the fact that the same volunteers were used in the training and test datasets multiple times, and were able to engineer features based on this fact.

We believed algorithms that exploited embedded design weaknesses will experience a significant drop in performance in production environments where the exploited assumptions don't hold.

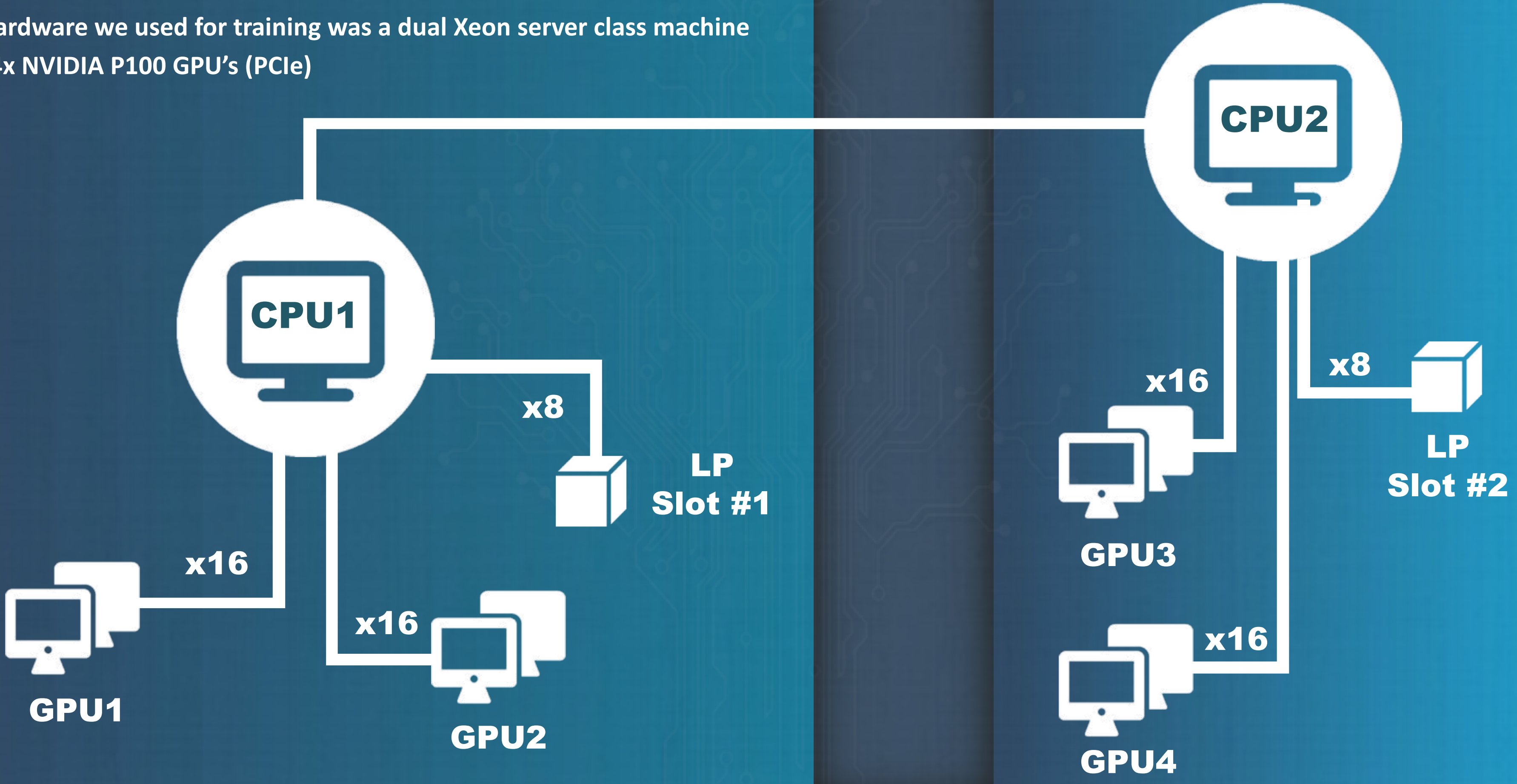
A major strength of our algorithm is that our single best model is sufficient to maintain a 3^{rd} place finish. This makes it more amenable to production deployment.

An addressable weakness of our model is that we did not fully utilize the higher dimensional datasets.

Single best model score (0.03576)
Ensemble score(0.03042)

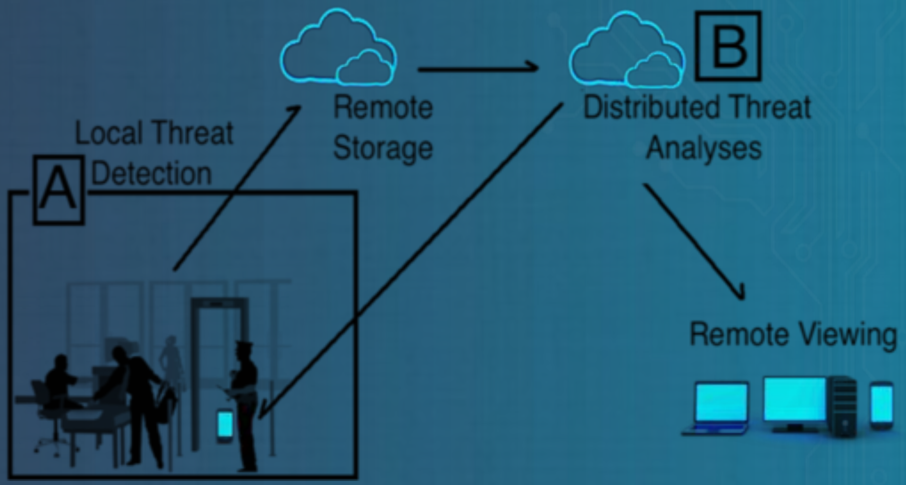
HARDWARE ARCHITECTURE

The hardware we used for training was a dual Xeon server class machine with 4x NVIDIA P100 GPU's (PCIe)



THE FUTURE OF AIRPORT SECURITY

We envision a future in which airport security devices, facial recognition, and other threat detection technologies (include luggage screening) will be linkable to provide aggregate threat assessments locally and across distributed settings



Current security algorithms are proprietary, expensive, and often released in long cycle times

These approaches are not amenable to change in the face of new threats, nor are they amenable to improvement in a field in which technology and speed of computation continually advances

AI-based threat detection can continuously learn and develop along with technology, and is nimble in the face of developing threats

Embedded systems will provide threat assessment in real time (A) and with the context of retained memory of prior events and detections (B)

Future Work

AnalyticalAI was founded prior to this competition, to develop threat detection technologies in the medical domain and other domains

We are open to commercial partnerships with vendors, and have an established relationship with one commercial vendor in the threat detection domain

AnalyticalAI specializes in two principle domains current:

- Threat or resource detection in 2D, 3D, and higher dimensional data
- Financial technology

COMPANY BACKGROUND

DAVID ODAIBO

Co-Founder, Chief Data Officer & Chief Data Scientist

5+ years experience with AI design and implementation. Extensive experience with software design and implementation.

THOMAS ANTHONY

Co-Founder, Chief Technology Officer

Director , UAB Big Data Lab. Chief Big Data consultant – Fortune 500 Company. 10+ years experience in building classical HPC systems and GPU HPC architecture

MARK FROEHLICH

Co-Founder, CEO

Extensive business experience, including executive positions at Thompson Caterpillar, Secureworks, and CapitalSouth Banks, as well as private equity experience.

DR. FRANK SKIDMORE

Co-Founder, Chief Science Officer

NIH funded investigator with 7+ years experience in data analytics and medical imaging.

DR. RAJBEER SINGH SANGHA

Investor, Board Member

Investor www.analyticalai.com

ANALYTICAL AI

www.analyticalai.com