

SAVANNAH THAIS - PRINCETON UNIVERSITY

IRIS-HEP RETREAT - 09/12/19

INTRODUCTION AND PLANS

INTRODUCTION

- ▶ New post-doc with IRIS-HEP and Princeton CMS Group
- ▶ Interests:
 - ▶ Machine Learning
 - ▶ Algorithm development, interpretability, integration
 - ▶ Trigger
 - ▶ HL-LHC high occupancy, parallelization
 - ▶ Supporting students in exploring ML
 - ▶ Computing best practices within HEP
 - ▶ Integrating development with physics education
 - ▶ Intersection of policy and emerging technology

BACKGROUND

- ▶ PhD from Yale with Sarah Demers and Keith Baker

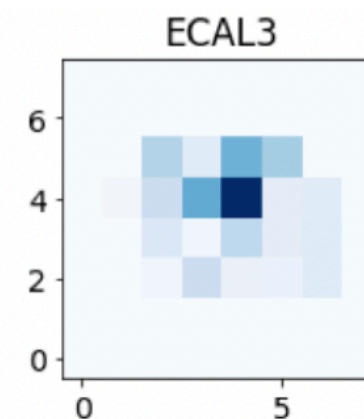
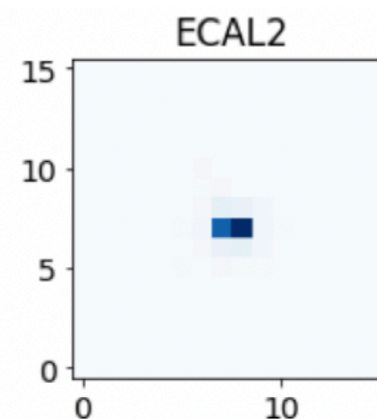
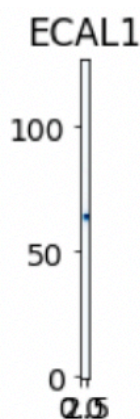
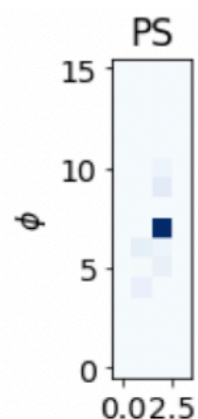
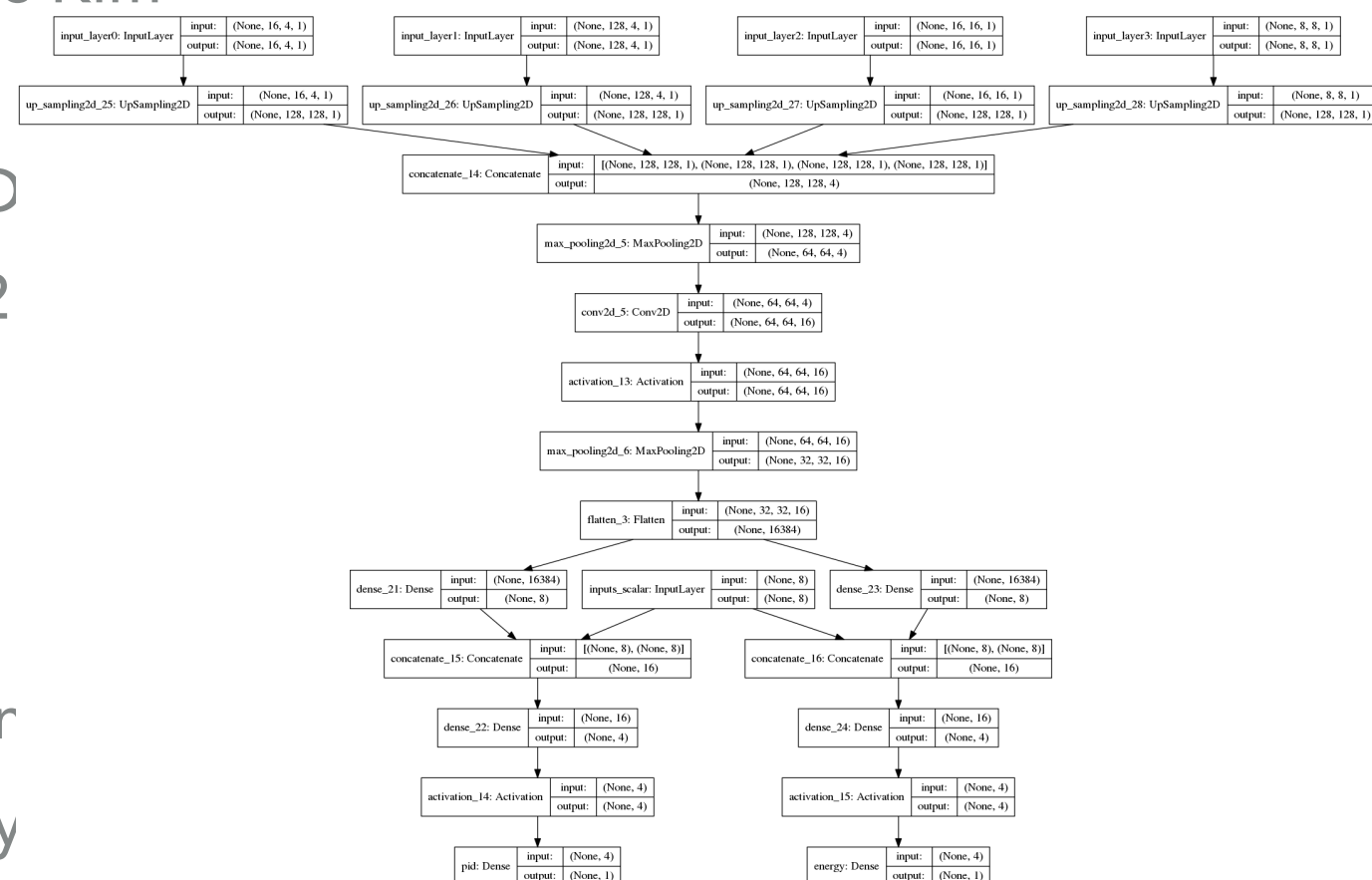
- ▶ Undergrad at Uchicago with Young-Kee Kim

- ▶ CP Work: Electron Identification

- ▶ Implemented and optimized low pT ID
 - ▶ Moved to fully data-driven ID for Run 2
 - ▶ Explored CNN image-based ID

- ▶ Analysis: $VH, H \rightarrow \tau\tau$

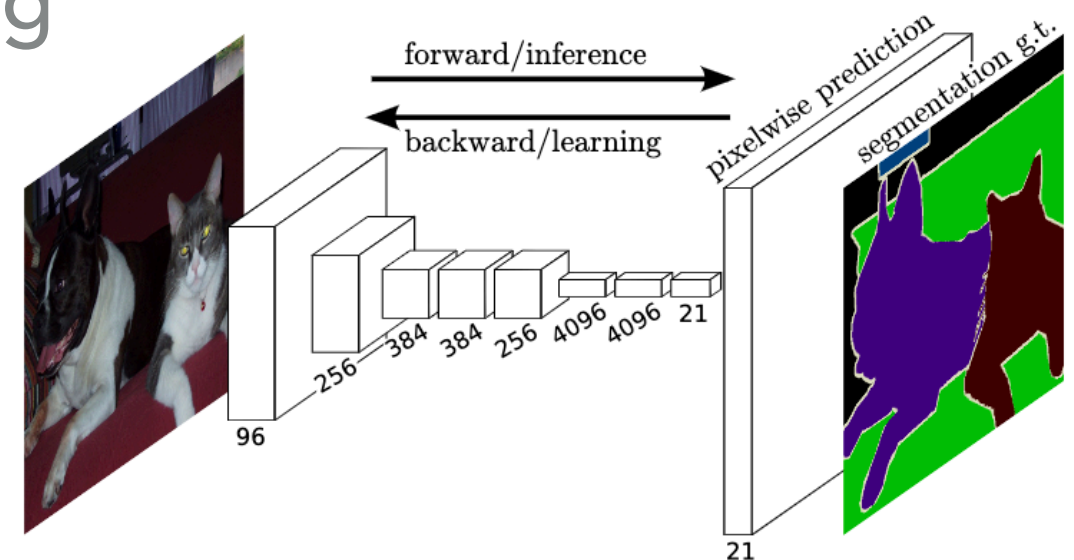
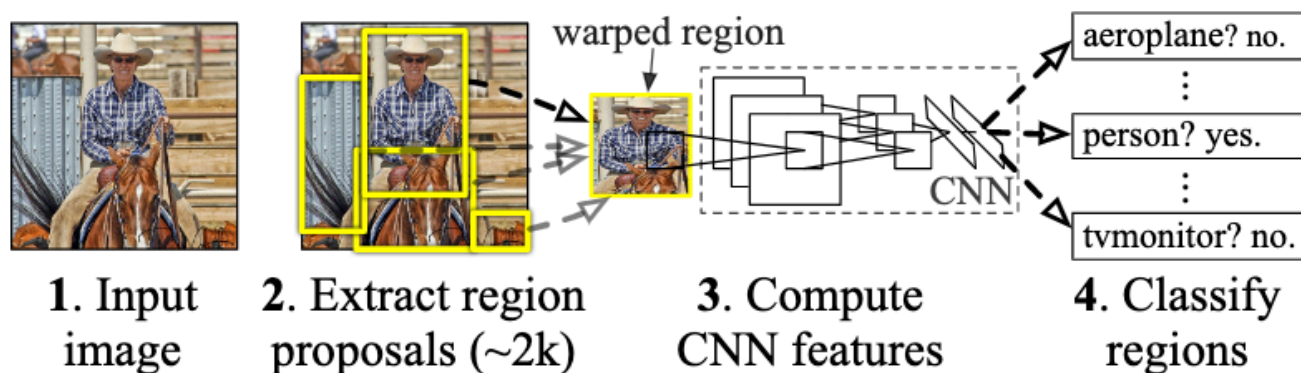
- ▶ Improved event selection efficiency
 - ▶ Developed data-driven fake background
 - ▶ Began migrating to fully pythonic analysis



PLANS: PIXEL TRACKING

- ▶ Clustering in the pixel detector with semantic segmentation
 - ▶ R-CNN to integrate pixel labeling and clustering **and/or**
 - ▶ FCN to label pixels then CNN with templates
 - ▶ Working with FastML group to implement on FPGAs for HLT
 - ▶ Fits in G4.4/G4.9
- ▶ Pixel track-let finding with GNNs
 - ▶ Assess GNN feasibility on a limited-scope problem
 - ▶ Interest in full detector track finding

R-CNN: *Regions with CNN features*



PLANS: STRIP TRACKING

- ▶ [MkFit](#) is a CMS project to parallelize full tracking
 - ▶ Takes all reconstructed hits as input
- ▶ Currently all hits are constructed serially
 - ▶ Working with Bei, David, and others to parallelize strip clustering on GPUs
 - ▶ Focused on data unpacking
 - ▶ Exploring existing methods for GPU-based spatial sorting

INTERESTS/PLANS: EDUCATION AND TRAINING

- ▶ A lot of experience developing trainings and outreach programs
- ▶ Plan to integrate with the Blueprint meeting planning and best practices development
- ▶ Potential joint CMS/ATLAS training in November
 - ▶ Agnostic mornings, separate breakouts
- ▶ Interested in how we can integrate with formal education
- ▶ AS: common plotting tools and documentation (stats)

PLANS: POLICY AND OUTREACH

- ▶ [Modeling region-specific opioid abuse indicators in Appalachia](#)
 - ▶ Integrating legal, lobbying, and infrastructure information into traditional models
 - ▶ Long-term goal: partner with local organizations to inform intervention
- ▶ MIT Media Lab [City Science Opportunity Zone](#) Project
 - ▶ Building API to geographically request investment and employment information
- ▶ Serving on [Forum on Physics and Society](#) for APS
 - ▶ Organizing session on AI and policy for March Meeting
 - ▶ Developing policy statements (currently gun violence research funding)
- ▶ NeurIPS: co-organizing [ML and Physical Sciences Workshop](#), serving on [WiML](#) Policy and Research Committee