Vision Transformers for End-to-End Particle Reconstruction for the CMS Experiment

Personal info

Name: Ka Wa Ho
Email: kaho@cern.ch
University: University of Notre Dame
Major: High Energy Physics (Ph.D.)
GitHub: https://github.com/kawaho

Projection description

The Compact Muon Solenoid (CMS) is a particle detector for the world's largest and most energetic particle collider, the Large Hadron Collider. CMS collects hundreds of physics events per second with tens of particles overlay per event. Previously, convolutional neural networks (CNN) were applied to identify physics events [1] and particles [2] [3] from the CMS open data. The performance of these networks were promising but they could potentially be improved upon with the novel vision transformer (ViT) network [4]. ViT has outperformed state-of-the-art (SOTA) CNNs both in terms of accuracy and computational cost for image classification. Recently, ViT has also shown improved results on jet-tagging compared to the SOTA graph neural network based-models [5].

The goal of the project is to apply and develop "end-to-end" vision transformer (ViT)-based networks for jet-flavor identification (b quark vs c quark vs light quark vs gluon jets) in simulated di-jet QCD events from the CMS open data and to benchmark their accuracy performance. This is expected to expand the previous work [2] focused only on light quark vs gluon jets tagging with CNNs. The "end-to-end" here refers to a direct use of images recorded by the particle sub-detectors at the CMS for classification, unlike the standard approach to carry out particle reconstructions beforehand, where useful information could be lost.

The project is valuable for the open source development of an end-to-end particle classifier for the CMS open data. It strips away the barrier to understand the complex reconstruction algorithms and object definitions used by the CMS [6] and invites ideas outside of the particle physics community. This aligns perfectly with Google's commitment with the growth of an open source ecosystem.

Proposed timeline

The proposed timeline is tentative and flexible according to the availability of the mentors. The planned duration is 175 hours over the course of 5 weeks, with an average of 35 hours per week:

- **MAY 4 - 28 (COMMUNITY BONDING PERIOD)**
  - Review understanding of previous publications on similar topics [1] [2] [3].
  - Design data selection/processing drawn upon previous studies.

- **JUNE 5 - 11 (WEEK 1 OF CODING)**
  - Carry out data processing on the CMS open data to generate images for the network trainings; target simulated data sets include:
    1. Di-jet QCD production for jet-flavor identification (b quark vs c quark vs light quark vs gluon jets)
    2. (Optional) Multi-jet QCD production and decay of the Higgs boson to a bottom quark pair from gluon-gluon fusion for boosted Higgs boson vs QCD jets tagging.
  - Implement and train a ResNet-based architecture [7] and a baseline ViT model [4] as the first benchmark.

- **JUNE 12 - 18 (WEEK 2 OF CODING)**
  - Fine-tune hyper-parameters of the baseline models to optimize their accuracy performance, meeting at least the previous published results [2].
  - Adjust the data selection/processing and apply data augmentation if needed.
  - Implement SOTA variants of the ViT model that are representative of the different approaches to improve the baseline ViT, including but not limited to
    1. DeiT: adding an extra token to learn upon a strong "teacher" model [8]
    2. CvT: adapt CNN-like shrinking (expanding) spatial (feature) dimension structure and use convolutional layers for the tokenization of image patches and projection in the self-attention layers [9]

3. Swin: adapt CNN-like shrinking (expanding) spatial (feature) dimension structure and use sliding attention windows to facilitate cross-patch connection [10].

- **JUNE 19 - 25 (WEEK 3 OF CODING)**
  - Continue implementation of the ViT-based models and start training.
  - Fine-tune hyper-parameters to optimize the models' performance and benchmarking.
  - Adjust the data selection/processing and apply data augmentation if needed.

- **JUNE 26 - JULY 2 (WEEK 4 OF CODING)**
  - Continue the fine-tuning and possible adjustment of the data selection/processing/augmentation.
  - Study and identify the strength and weakness of each trained models.
  - (Optional) Investigate new architecture or techniques to reach even better performance based on the studies. This could come from combining the existing models or testing novel solutions to resolve the bottleneck of the existing models.

- **JULY 3 - 9 (WEEK 5 OF CODING)**
  - (Optional) Continue the study, train, and optimize the new architectures/techniques developed.
  - Organize the studied results and clean code.

- **JULY 10 - END OF CODING PERIOD**
  - Buffer time for debugging or in the face of unexpected events.

About me

I am a PhD student in high energy physics at the University of Notre Dame. I conduct research at the CMS experiment at CERN and is passionate about research related to the use of machine learning in particle collider data analysis and the development of software/firmware for the readout electronics at particle calorimeters.

My day-to-day work involves the extensive use of python and C++ for big data (hundreds of TB) analysis and visualization. I was involved in the publication of two data analyses with 3 years of collected data at the CMS to search for new physics signal [11] [12]. Both analyses involved machine learning techniques of using pre-trained neural networks for particle identification and training of boosted decision trees to identify potential signal from backgrounds in data.

In my free time, I work on small projects in machine learning, mostly to understand and implement myself interesting paper I found. Some examples are projects on generative adversarial networks [13], CNNs for image segmentation, and uncertainties assessments in CNNs [14].

I believe my publication history and personal projects are good indication that I am qualified and passionate about the project I am applying. My familiarity with high energy physics, the CMS detector, and deep learning together makes me a well-suited candidate on this task with the minimal time needed to familiarize with the CMS data format. And as a CMS collaborator myself, I believe this task would help flourish the openness of the CMS data and attract interests from experts outside of the collaboration.

References

[1]: Andrews, Michael, et al. "End-to-end physics event classification with CMS open data: Applying image-based deep learning to detector data for the direct classification of collision events at the LHC." *Computing and Software for Big Science* 4 (2020): 1-14.

[2]: Andrews, Michael, et al. "End-to-end jet classification of quarks and gluons with the CMS Open Data." Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 977 (2020): 164304.

[3]: Andrews, Michael, et al. "End-to-end jet classification of boosted top quarks with the CMS open data." *Physical Review D* 105.5 (2022): 052008.

[4]: Dosovitskiy, Alexey, et al. "An image is worth 16x16 words: Transformers for image recognition at scale." *arXiv preprint arXiv:2010.11929* (2020).

[5]: Qu, Huilin, Congqiao Li, and Sitian Qian. "Particle transformer for jet tagging." *International Conference on Machine Learning*. PMLR, 2022.

[6]: Sirunyan, Albert M., and CMS collaboration. "Particle-flow reconstruction and global event description with the CMS detector." *JINST* 12.10 (2017): P10003.

[7]: He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[8]: Touvron, Hugo, et al. "Training data-efficient image transformers & distillation through attention." *International conference on machine learning*. PMLR, 2021.

[9]: Wu, Haiping, et al. "Cvt: Introducing convolutions to vision transformers." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.

[10]: Liu, Ze, et al. "Swin transformer: Hierarchical vision transformer using shifted windows." *Proceedings of the IEEE/CVF international conference on computer vision*. 2021.

[11]: Sirunyan, Albert M., et al. "Search for lepton-flavor violating decays of the Higgs boson in the μ τ and e τ final states in proton-proton collisions at s=13 TeV." *Physical Review D* 104.3 (2021): 032013.

[12]: CMS collaboration. “Search for the lepton flavor violating decay of a Higgs boson in the eμ final state in proton-proton collisions at s=13 TeV." Technical report, CERN, 2023

[13]: https://github.com/kawaho/W-GP-DCGAN

[14]: https://github.com/kawaho/MLproj