KNS Spring Meeting 2023. May 17-19, ICC Jeju

Reactor Physics & Particle Transport Computation Simulation Lab.

Application of a Vision Transformer for Prediction of Peaking Factors and Cycle Lengths in OPR1000

Seongjin Jeong
Professor : Hyun Chul Lee (hyunchul.lee@pusan.ac.kr)
Nuclear System Division, School of Mech. Eng.
Pusan National University

Research Purpose

- **Impossible to evaluate all possible Loading Patterns (LP)s in LP optimization**
  - Optimization algorithms such as simulated annealing have been developed so far.
  - However, it takes **tremendous computational time** for neutronics calculations.

- **Artificial Intelligence (AI) model could be applied to LP analyzation.**
  - **AI prediction is faster than code calculation**
    - Useful to avoid the code to calculate meaningless LPs
  - Outstanding AI model: **V**ision **T**ransformer(ViT)
    - Exhibit better accuracy for image recognition compared with every prior-developed AI model.

Research Background

➢ What is AI model?

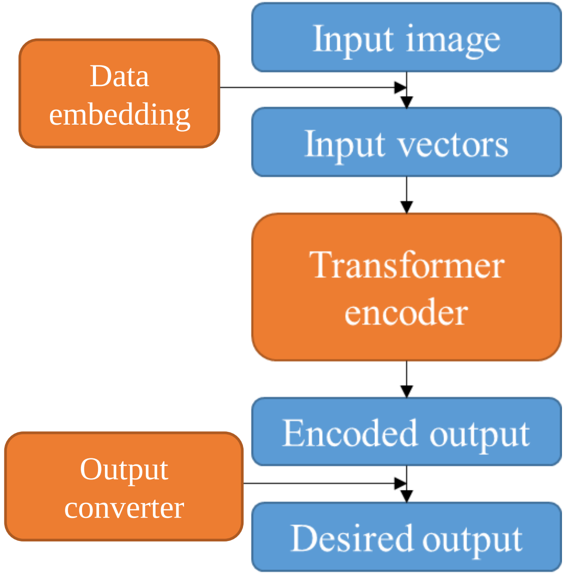
- **Prediction tool for a complicated phenomenon.**
  - Training AI model is to make higher prediction accuracy using data.
- Expressions in a training process.
  - **Parameter** is a value that consists AI model.
    - Training AI model is to optimize the parameters in AI model.
  - **Dataset** is a combination of **problems (inputs)** and the **answers (outputs).**
  - **Batch** is a sample of data that computing resources could handle.
    - The data is divided into batches and given to AI model sequentially.
  - **Epoch** is an iteration number that how many times AI model has seen the whole data.

➢ How can we make AI model?

- Prepare a dataset which contains many different situations.
- Tune the model to achieve better results with untrained data.

ViT Calculation

- Data embedding
  - Convert an image to vectors.
- **Transformer encoder**
  - Find the patterns in the vectors.
- Output converter
  - Arrange the encoded patterns into an output format.



Data-flow in the ViT

Data Embedding

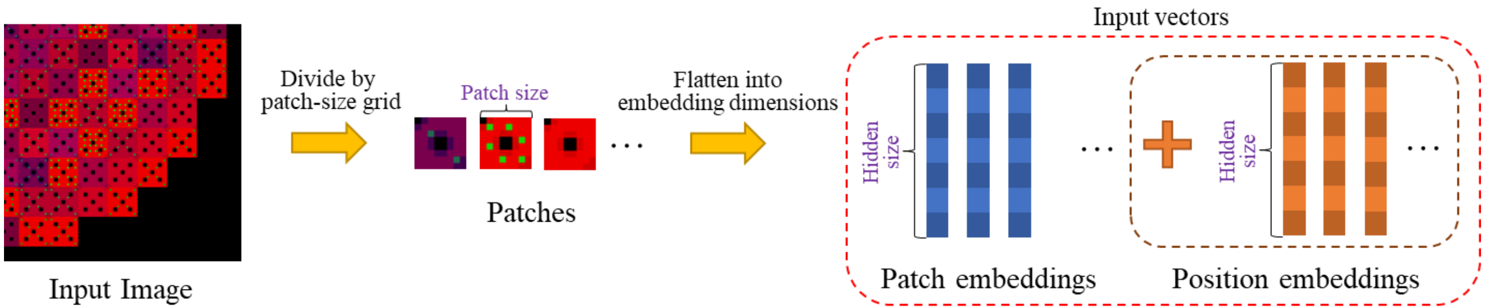
➢ Patch embedding

- The image is divided into patches.
  - Patch is a fixed-size square region extracted from an input image.
- Each patch corresponds to a patch embedding vector.

➢ Position embedding

- The position information of each patch corresponds to a position embedding vector.

➢ The input vector consists of the patch and position embedding vectors.



The data embedding process

Transformer Encoder

➢ Multi-head attention

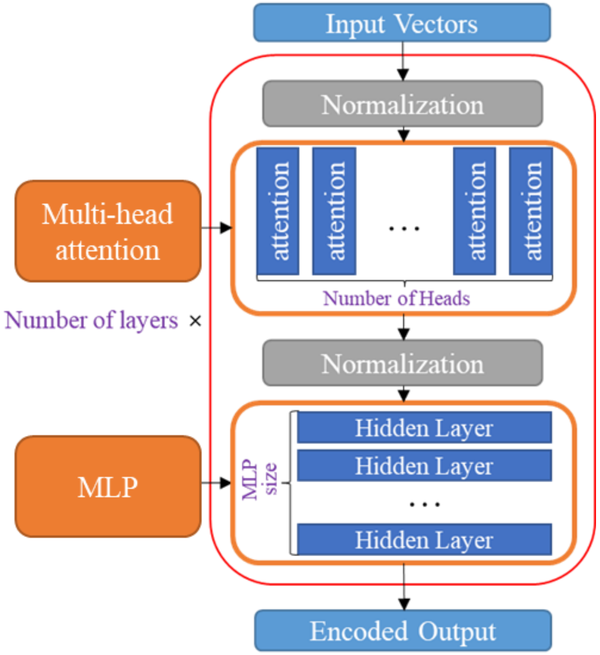
- Attention is to calculate the dependencies among input vectors.
- Multi-head attention is to process multiple sight of view in parallel.
- **Crucial factor of ViT performance.**

➢ **M**ulti-**L**ayer **P**erceptron (MLP)

- The hidden layers consist of perceptron.
- Perceptron used to change the dependencies into encoded values.

➢ A layer in transformer encoder

- a set of Multi-head attention and MLP
- repeated for more detailed interpretation



One layer in transformer encoder

Generated LPs by Applying Random Deviations

- Used the LPs described in the reference.
  - "Nuclear Design Report for Yonggwang Unit 3 Cycle 1/2"
  - Includes 9 types of fuel assemblies in cycle 1, and 12 types in cycle 2.
- Changed the parameters of the fuel pins in the one-eighth fuel assembly.
  - Modified parameters
    - **Fuel enrichment (wt.%)**
    - **Burnable Poison (BP) fraction (wt.%)**
- Arranged the resulting fuel assemblies in the core in a random pattern.
  - Changed Parameter is **fuel assembly burnups (MWd/kgU).**
    - based on the values of fuel assembly types in BOC
  - All pins in an assembly have the same burnup.
- Each change of parameters is a random value within ±20%.
  - The value of 20% was determined from sensitivity tests.

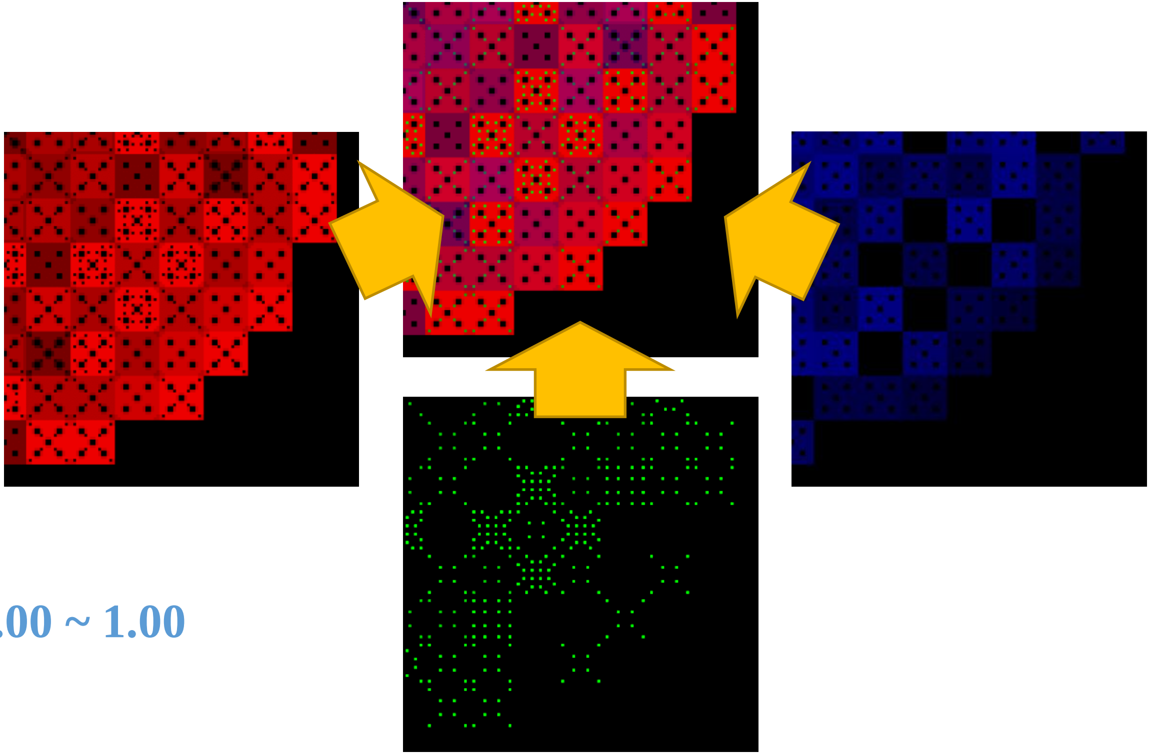
Generated Datasets for LPs.

- Analyzed LPs by using a two-step code system ST/R2.
  - STREAM 2D
    - Generated group constants from 2D transport calculation
    - XS/Depletion library: E71JD40
    - Branch option: Hot-state case matrix created with coolant temperature of ±20 K change
    - Final burnup: 80MWd/kgHM
  - RAST-K 2.0
    - Solved 3D LWR core problems by employing the diffusion equation.
    - Calculate the LP properties
      - Cycle length: EFPD when CBC becomes 10 ppm.
      - Peaking factor: maximum value of F-xy
- The Dataset was generated from a total of 110k LPs, with 55k LPs in each cycle.
  - Input: LP parameters for each pin in the LP
  - Output: LP properties analyzed by ST/R2

Interpreted Each LP as a Color Input of ViT

- Red channel
  - Fuel enrichment
  - Range: 0.00 ~ 6.00 wt.%
    - 0.00 for non-fuel region
    - 0.711 wt.% for BP pin.
- Green channel
  - BP fraction
  - Range: 0.00 ~ 10.0 wt.%
    - 0.00 for non-BP
- Blue channel
  - Assembly burnup
  - Range: 0.00 ~ 30.0 MWd/kgU
- **Every values normalized into 0.00 ~ 1.00**

Fuel enrichments (red), BP fractions(green), and assembly burnups(blue) combines into a color image to insert in a ViT calculation



The method to optimize ViT Model

➢ Estimated dataset prediction errors.

- A difference between the outputs from ST/R2 and ViT calculation.
- It could have some bounce while training, there should be no signs of error increasing trend.

➢ Two modes to find the optimum ViT model

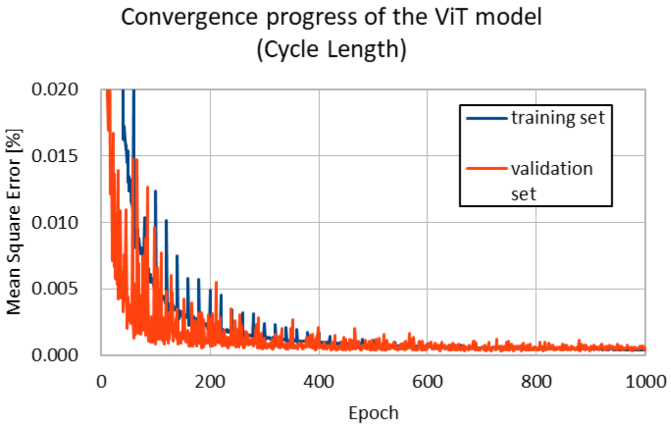
- Training mode
  - Allow to modify the parameters in ViT model.
  - Use output of dataset to optimize it.
  - **training set** is used for training in an epoch.
- Evaluation mode
  - Limit to change the parameters in ViT model.
  - Check the ViT model whether prediction is accurate on untrained data.
  - **Validation set** is used for evaluating ViT model in every epoch.
  - **Test set** is used after the training process is done.

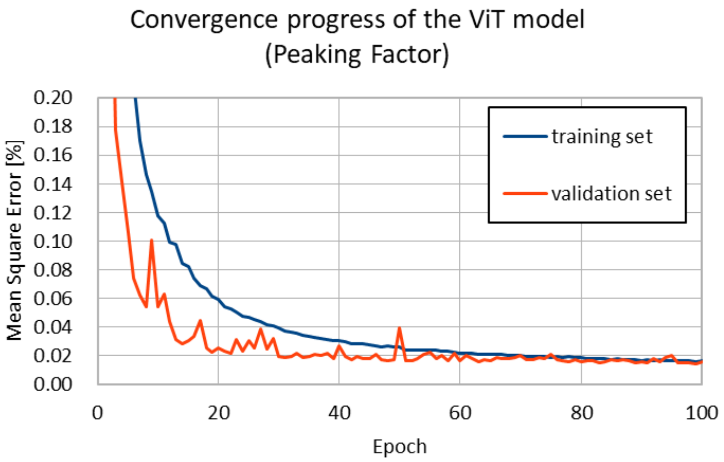
Calculation descriptions of training ViT Model

- Computing resource used for the training
  - CPU: Intel Core i7-9700F @ 3.00GHz
  - GPU: NVIDIA GeForce RTX 2080 Ti 12GB
  - Memory: 62GB RAM
- Model properties
  - Total number of parameters in ViT model: 1.1M
  - Dataset memory: 20 GB
- Subset size
  - Training set: 90k
  - Validation set: 10k
  - Test set: 10k
- It took 50 sec for one epoch of ViT model training.

Convergence of Dataset Prediction Error

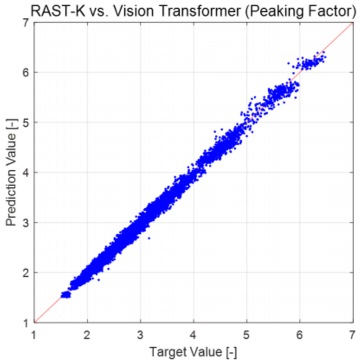
- Set a convergence criteria of dataset prediction error.
  - The average of the validation errors in the last 10 steps is higher than that in the 10 steps.
- The dataset prediction errors for validation were lower than those for training.
  - Due to the low batch size restricted by GPU memory size





Prediction Accuracy of ViT model

- The prediction values were compared with the ST/R2 code calculation results.
- **Standard Deviation** of relative error
  - The deviation of relative prediction error
  - Under 3% for the cycle length and peaking factor.
- Maximal value of relative error
  - Largest prediction error that ViT model could be.
  - About 3.75% for the cycle length prediction model
  - However, it is over 15% for the peaking factor prediction model



Error types	Std. Dev. [%]	Max Rel. Err. [%]
Cycle length	0.24	3.75
Peaking factor	2.51	16.29

Conclusions

- Suggests a guideline to apply ViT model in LP evaluation.
  - Make a specific method to generate dataset with LPs of OPR1000.
  - Achieve the RMS of relative error below 5%.

- Future work
  - Optimize ViT model for evaluating peaking factor
    - Find the reason of relatively large maximum of relative error.
  - Set a new large dataset for ViT model
    - A larger dataset makes it possible to generate more generalized evaluation model.
  - Compare with other AI models
    - Train other AI models with the dataset we produced.

Thank you ☺