AutoML Hyperparameter Tuning of Generative DNN Architecture for Nanophotonic Device Design

Toshiaki Koike-Akino, Keisuke Kojima, Ye Wang

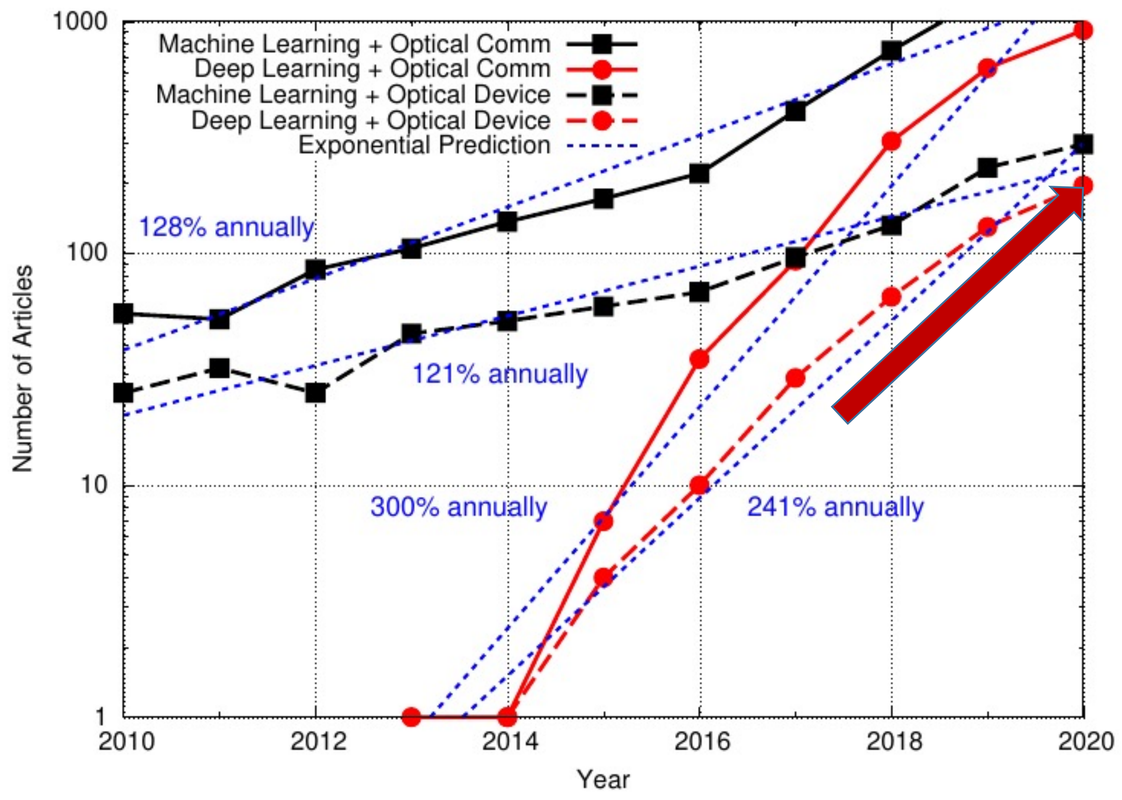
Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA.

Introduction

- **Deep neural networks (DNN)** have been widely used for photonic device design [1]– [10]
- The application has been exponentially growing by **240% every year**
- A **generative model** based on adversarial conditional variatinal autoencoder (**ACVAE**) can efficiently design nanophotonic device [4]
- However, hyperparameter selection requires great amount of **manual trial-and-error efforts**
- We use automated machine learning (**AutoML**) to optimize model hyperparameters [11]

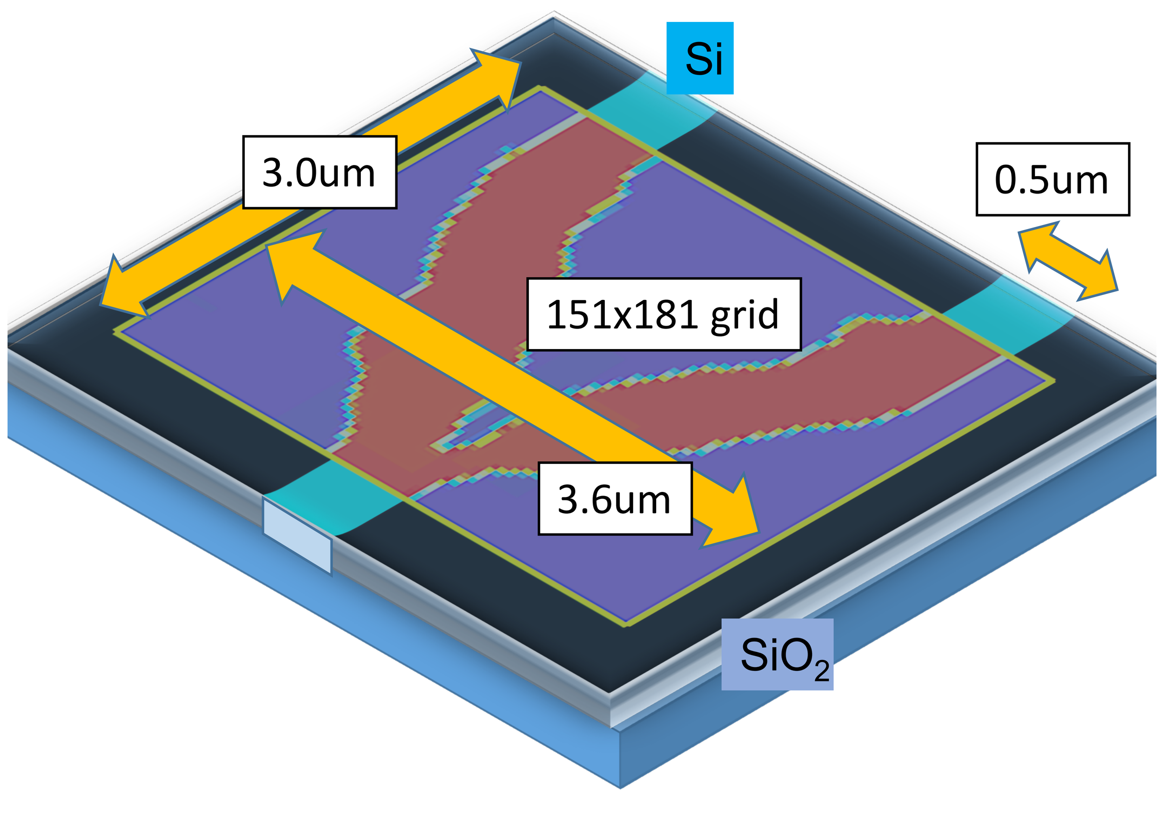
Learning for Photonic Device

- Application of DNN to optical devices has attracted much attention in the community
- The number of articles follows the **Moore's law**
- The annual growth rate is 240%



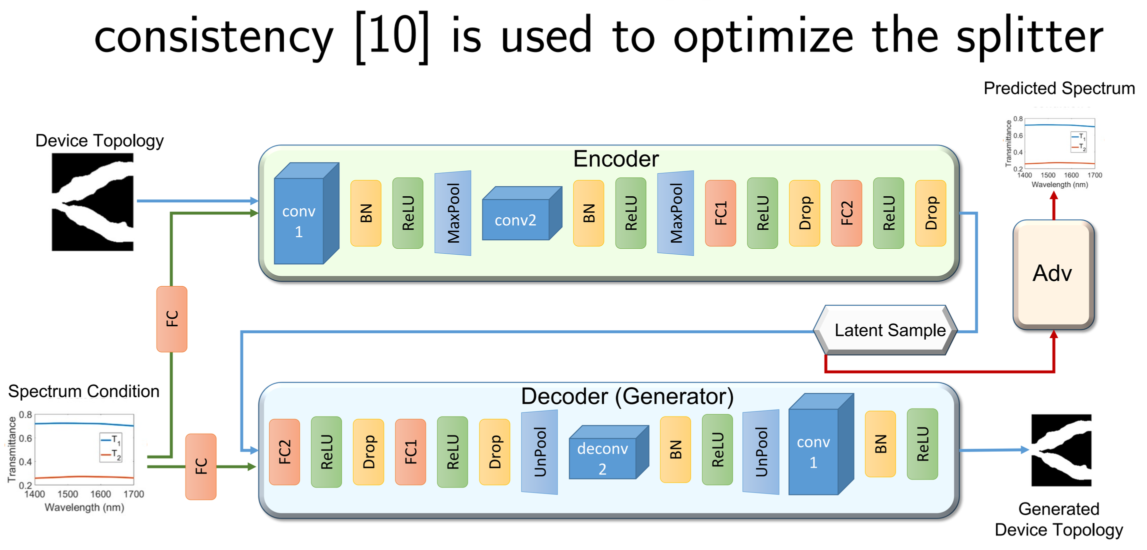
Nanophotonic Power Splitter

- We consider to design compact nanophotonic power splitter
- Silicon-on-insulator (SOI) platform
- Wideband wavelengths: 1450nm–1650nm



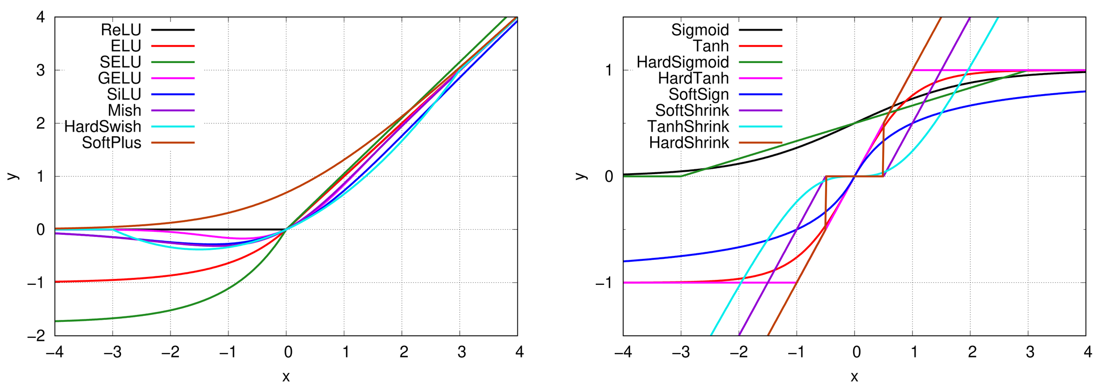
Generative Adversarial CVAE

- Generative ACVAE model [4] with cycle consistency [10] is used to optimize the splitter



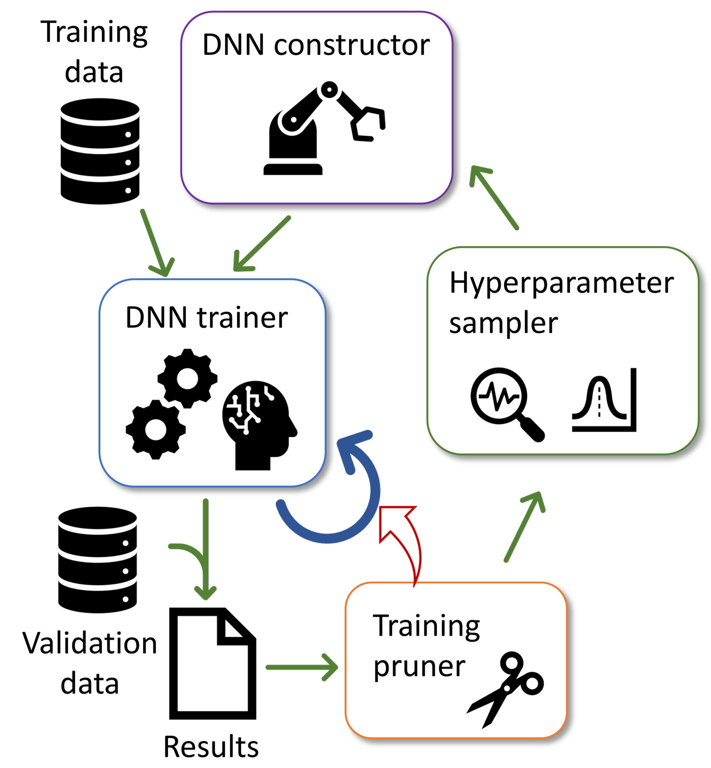
DNN Hyperparameters

- Selection of hyperparameters for DNN model is cumbersome
- **Huge search area**: number of hidden layers/nodes, activation functions, convolutional kernel sizes/strides/dilations/channels, latent size, adversarial coefficient, learning rates, etc.
- A large amount of manual efforts in trial-and-error exploration is required to find best hyperparameters



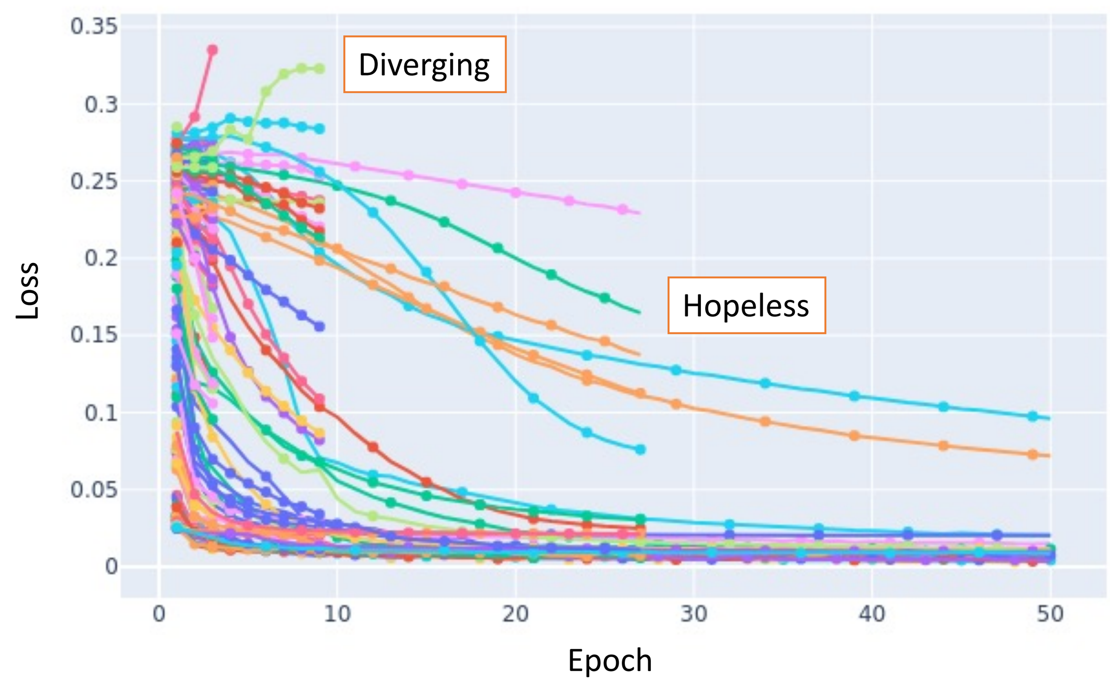
AutoML Tuning

- We use **AutoML** framework called Optuna [11]
- AutoML automatically searches for good hyperparameters
- **Bayesian optimization (BO)**-based efficient search



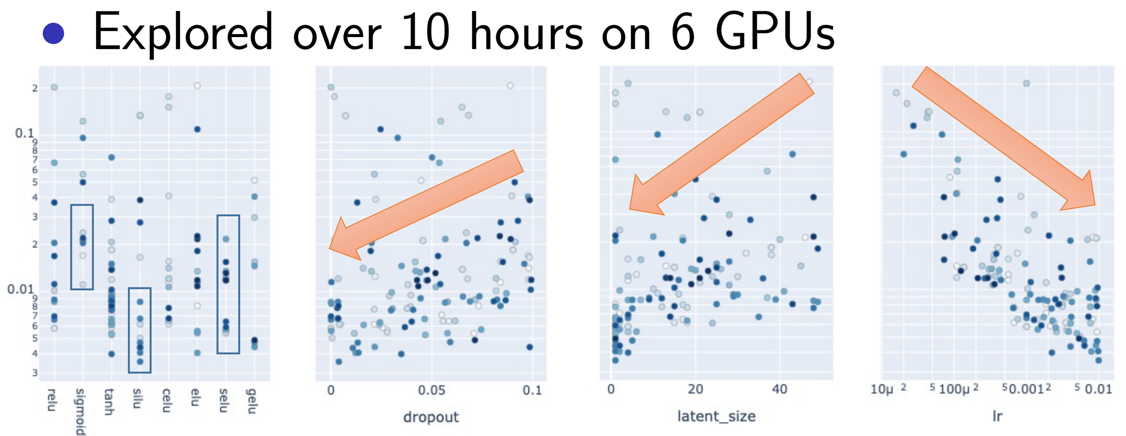
Early Pruning

- We use **hyperband pruning**
- Efficient search by discarding hopeless hyperparameter candidates
- 80% cases are pruned to save optimization time



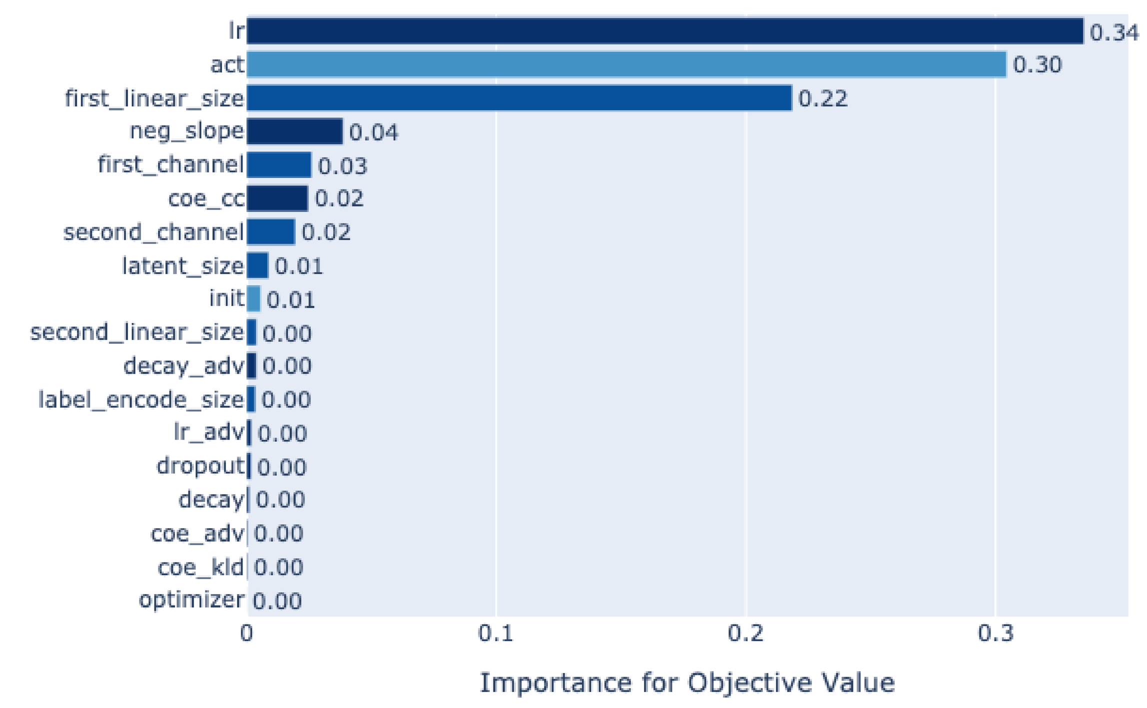
Hyperparameter Exploration

- Explored over 10 hours on 6 GPUs



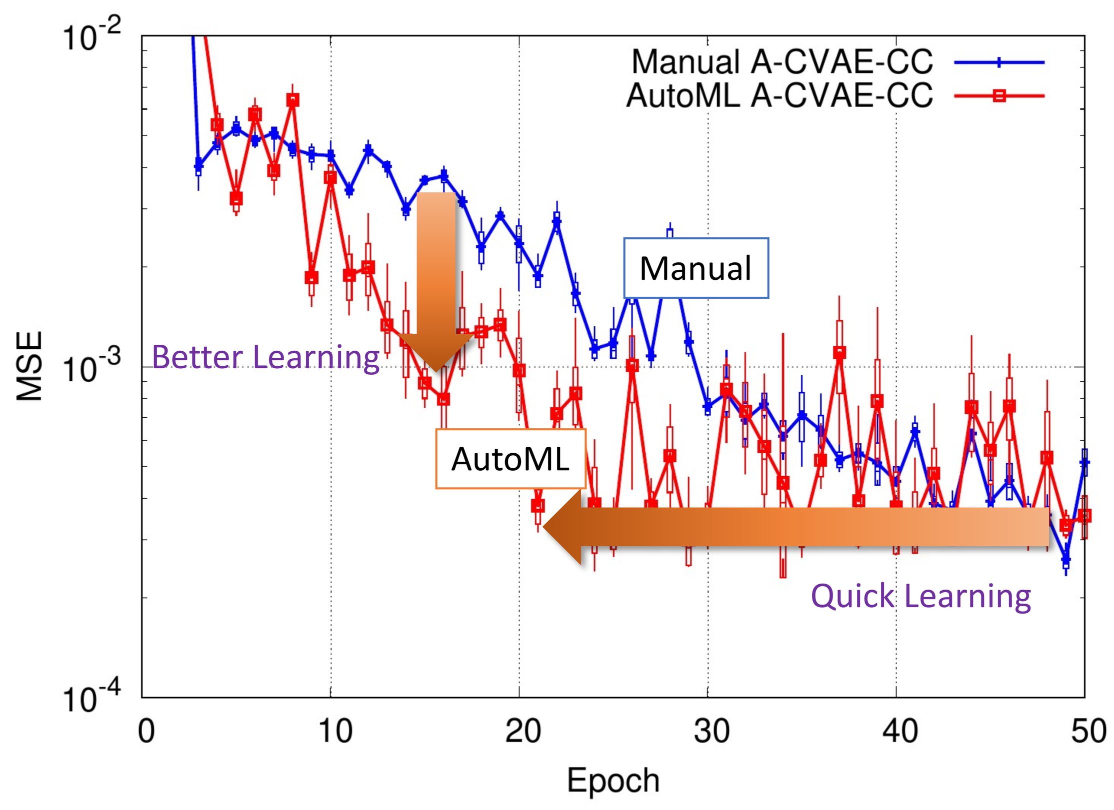
fANOVA Importance Score

- Hyperparameter **importance score**
- Major factors: Learning rate and activation



Performance

- AutoML found good model hyperparams
- Optimized model outperforms manual tuning
- It achieved quick and better learning
- Insertion loss of **0.1 dB** is achieved
- ACVAE model takes just **20 seconds** to generate 250 good devices



Conclusions

- We demonstrated that hyperparameter tuning of a generative DNN model can be efficiently automated and accelerated via AutoML

References

[1] K. Kojima *et al.*, "Acceleration of FDTD-based inverse design using a neural network approach", *IPR*, July 2017.

[2] M. TaherSima *et al.*, "Deep neural network inverse design of integrated photonic power splitters", *Scientific Reports*, Feb. 2019.

[3] M. TaherSima *et al.*, "Deep neural network inverse modeling for integrated photonics", *OFC*, Mar. 2019.

[4] Y. Tang *et al.*, "Generative deep learning model for a multi-level nano-optic broadband power splitter", *OFC*, Mar. 2020.

[5] K. Kojima *et al.*, "Deep neural networks for designing integrated photonics", *OFC*, Mar. 2020.

[6] Y. Tang *et al.*, "Generative deep learning model for inverse design of integrated nanophotonic devices", *Laser Photon. Rev.*, Oct. 2020.

[7] K. Kojima *et al.*, "Inverse design of nanophotonic devices using deep neural networks", *ACP*, Sep. 2020.

[8] K. Kojima *et al.*, "Deep neural networks for inverse design of nanophotonic devices", *JLT*, Jan. 2021.

[9] K. Kojima *et al.*, "Application of deep learning for nanophotonic device design", *Photon. West*, Mar. 2021.

[10] Y. Tang *et al.*, "Nano-optic broadband power splitter design via cycle-consistent adversarial deep learning", *CLEO*, May 2021.

[11] T. Akiba *et al.*, "Optuna: A next-generation hyperparameter optimization framework," *SIGKDD*, 2019.