NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

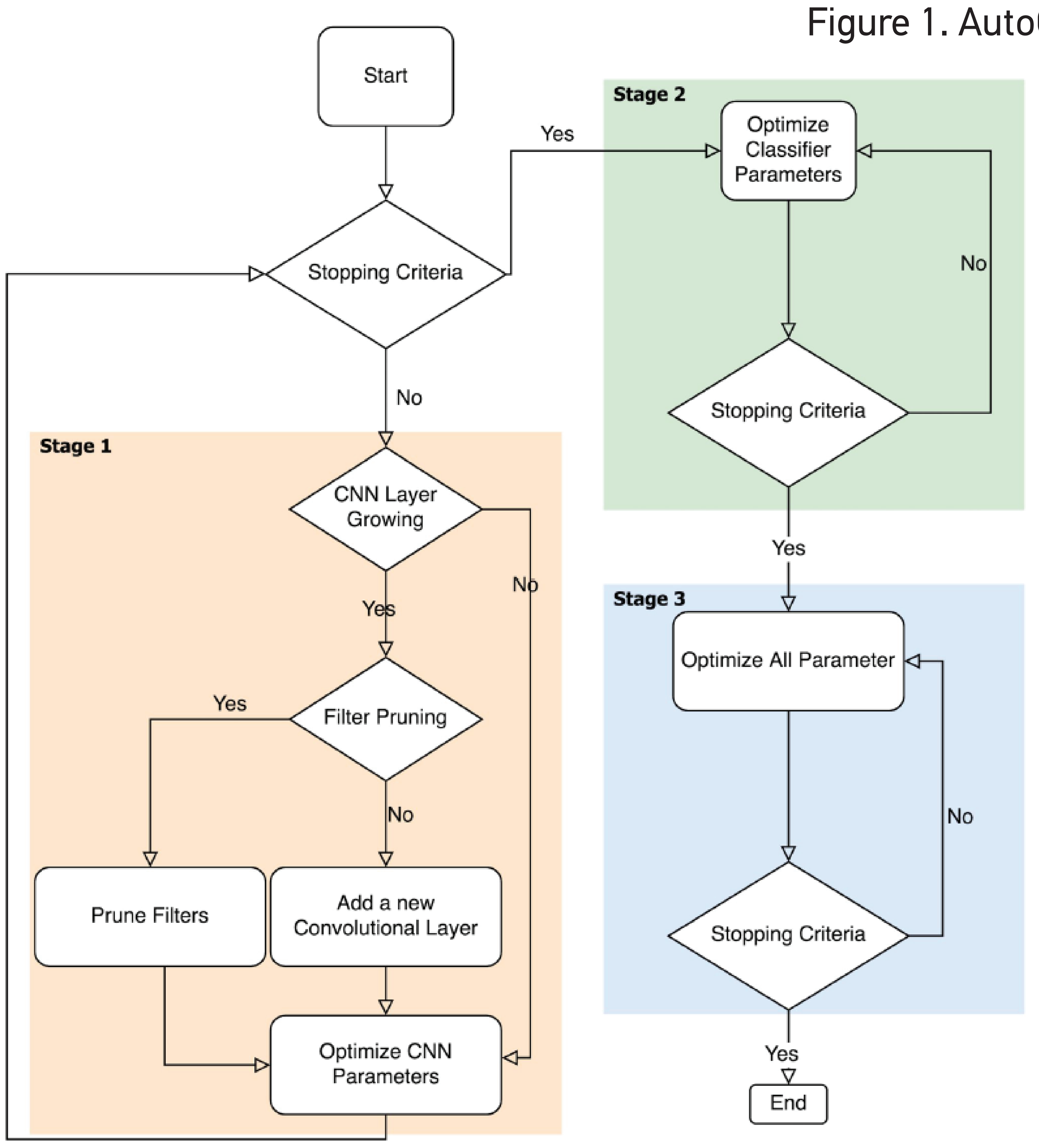
School of Computer Science and Engineering
College of Engineering

AutoCNN: A Data-Driven Architecture Learning Approach

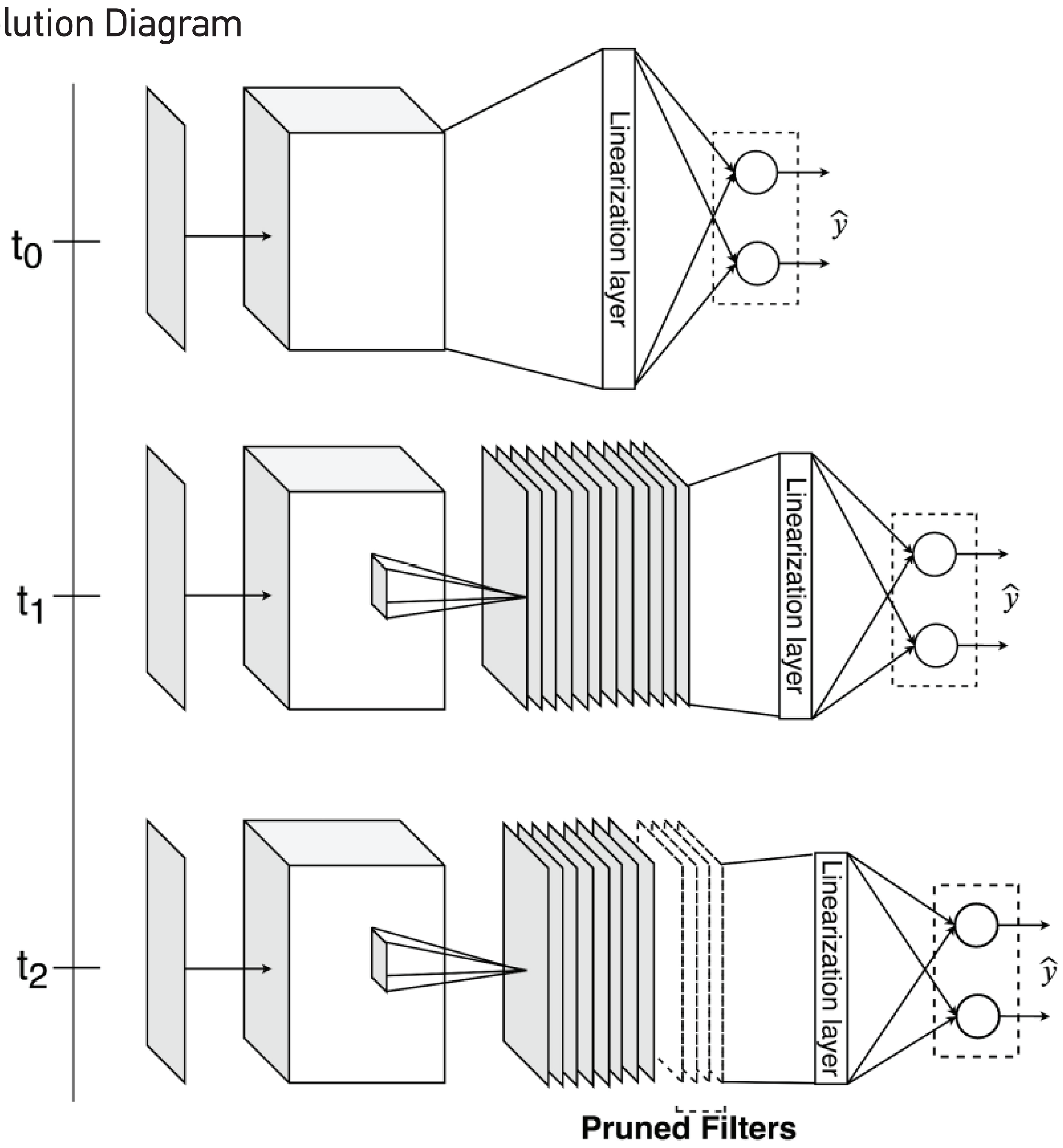
Student: Fienny Angelina **Supervisor:** Asst Prof Mahardhika Pratama

AutoCNN is a novel data-driven CNN architecture learning algorithm. Figure 1.a. shows the AutoCNN learning policy which is divided into 3 stages: CNN Adaptation, Classifier Optimization and All Parameter Optimization. Stage transformation is governed by the stopping criteria, based on the relative decrease of the exponentially weighted average of the training error. CNN Adaptation process is further visualized in Figure 1.b which shows the CNN Growing and Filter Pruning strategy. CNN Growing occurs when the network performance has saturated. Filter Pruning is used as it is impossible to determine the right number of filter at the start, thus some redundancy is introduced and highly similar filters are pruned. During the Classifier Optimization Stage, all neuron weights except those of the Fully Connected Layer is frozen. The last stage, the All Parameter Optimization trains all parameter until stopping criteria is reached.

Figure 1. AutoCNN Evolution Diagram



a. Learning Policy of AutoCNN



b. **Time t_0:** Initialization, **Time t_1:** CNN Growing, **Time t_2:** Filter Pruning

Experiment & Results

This novel algorithm is tested on 6 datasets: MNIST, MNIST Fashion, MNIST-rot-back-image, K-MNIST, CIFAR-10 and ADHD200. Each experiment is run on a single server with 2 GPU with model number GTXP2. All experiments are completed in 8 hours, except for the CIFAR-10 which takes 3 days to complete. It achieved the state-of-the-art accuracy on 2 datasets (MNIST rot-back-image and ADHD) and the second highest accuracy on another 2 datasets. Auxiliary studies are also conducted to measure the reproducibility and the capability of AutoCNN to improve on existing architecture, all of which show promising results.

Conclusion

AutoCNN is a novel data-driven CNN architecture learning method which is computationally inexpensive and able to generate network with superior performance and high immunity towards noise. Moreover, AutoCNN is not only capable to generate network, but also to improve upon existing network.

Dataset	Model	Accuracy
MNIST	AutoCNN	99.61%
	IPPSO	98.79%
	EvoCNN	98.82%
	FCCNN	97.57%
	DCNN+GFE	99.5%
	SOPCNN	99.83%
MNIST rot-back-image	AutoCNN	87.40%
	IPPSO	65.50%
	EvoCNN	62.62%
	FCCNN	66.40%
MNIST Fashion	AutoCNN	94.33%
	EvoCNN	94.53%
	EDEN	90.6%
	CNN2	92.54%
	C-CoNN	90.84%

Dataset	Model	Accuracy
Kuzushiji MNIST	AutoCNN	98.09%
	VGG8B+LL+CO	99.01%
	PARN-MM	98.83%
	HED NN+	85.12%
CIFAR-10	AutoCNN	90.07%
	PENAS	97.70%
	ResNet	94.77%
	Gpipe	99.00%
ADHD200	AutoCNN	77.60%
	3D CNN	69.15%
	DTM	70.36%

Table 1. AutoCNN Performance in Comparison with other algorithms