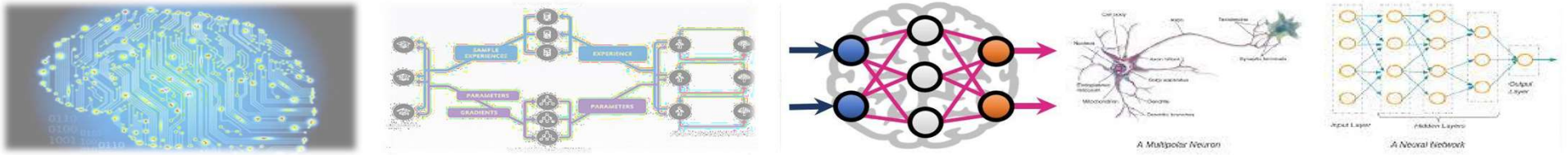


Comparison Long Short-Term Memory Network and Convolution Neural Network, ShuffleNet

LSTM 네트워크와 CNN기반 ShuffleNet 구현 및 비교

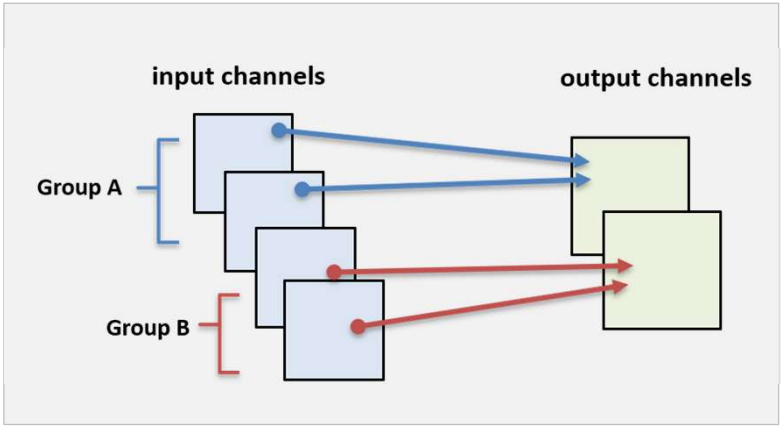
Geonil Yun,
rjsdlf45@gmail.com
2021. 06. 29.

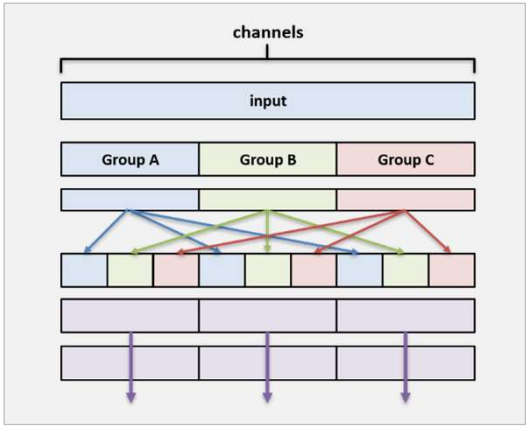
Introduction

- **Shuffle Net**

  - ✓ **Architecture for reduce computational cost**
  - ✓ **Use two key operations**

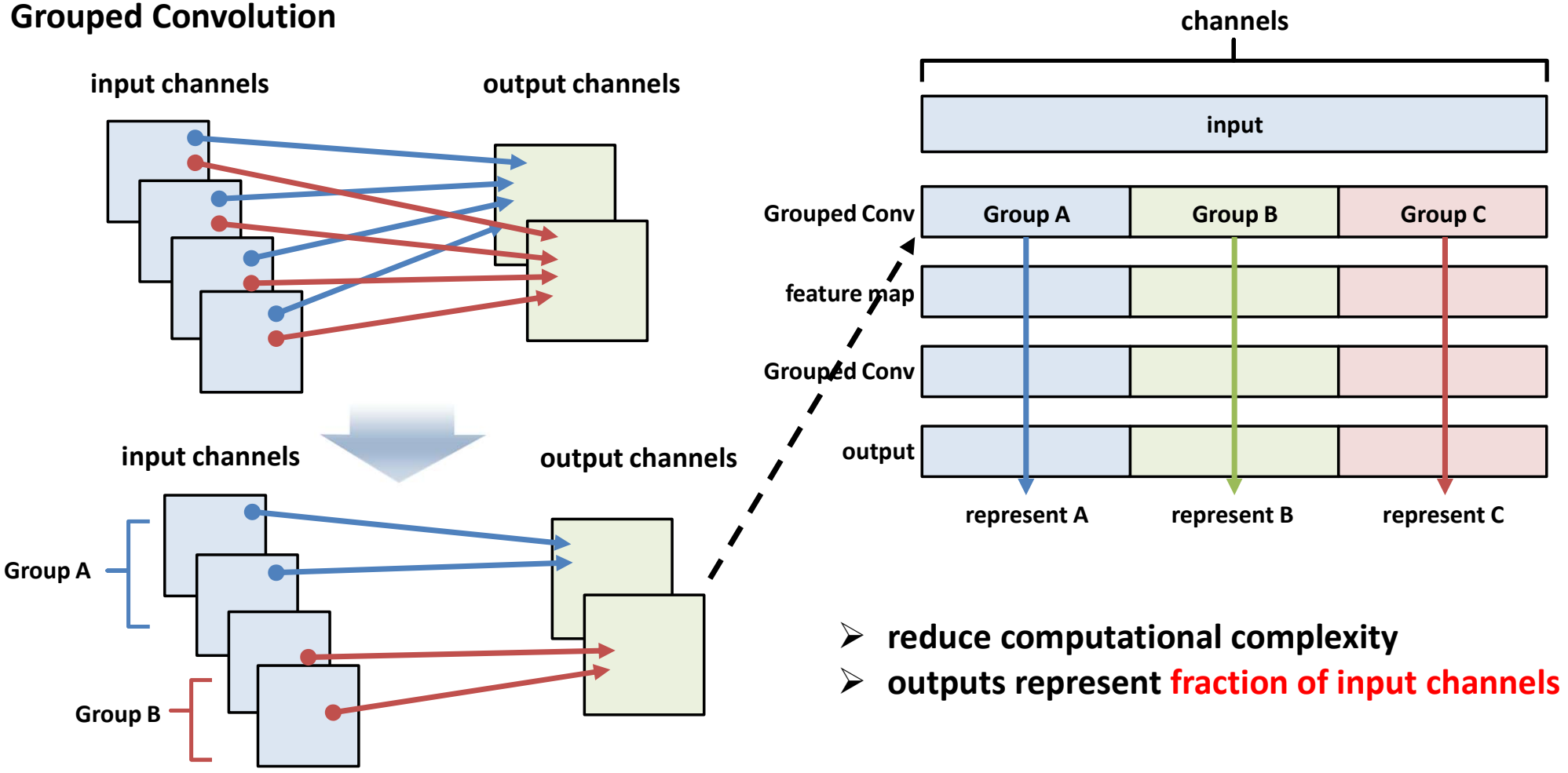
    ① **Pointwise group convolution**
    ② **Channel shuffle**





Introduction

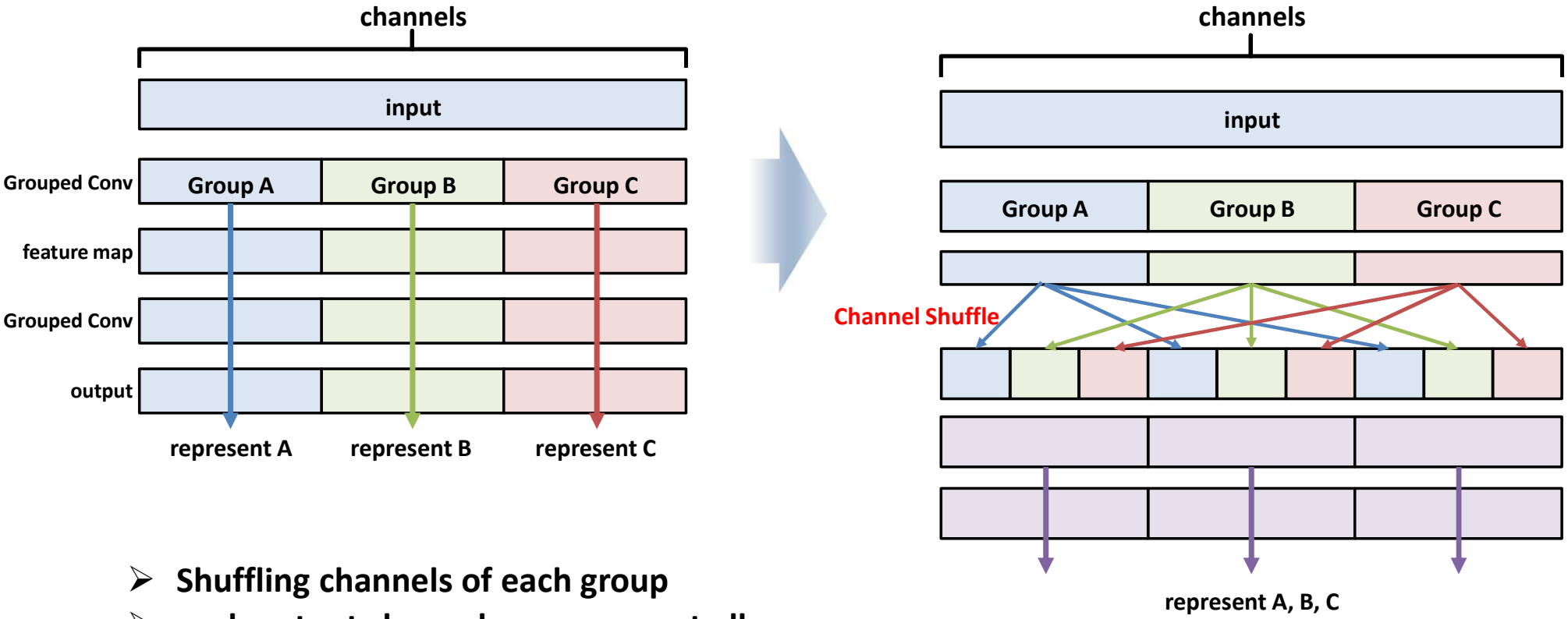
- Grouped Convolution



- reduce computational complexity
- outputs represent fraction of input channels

Introduction

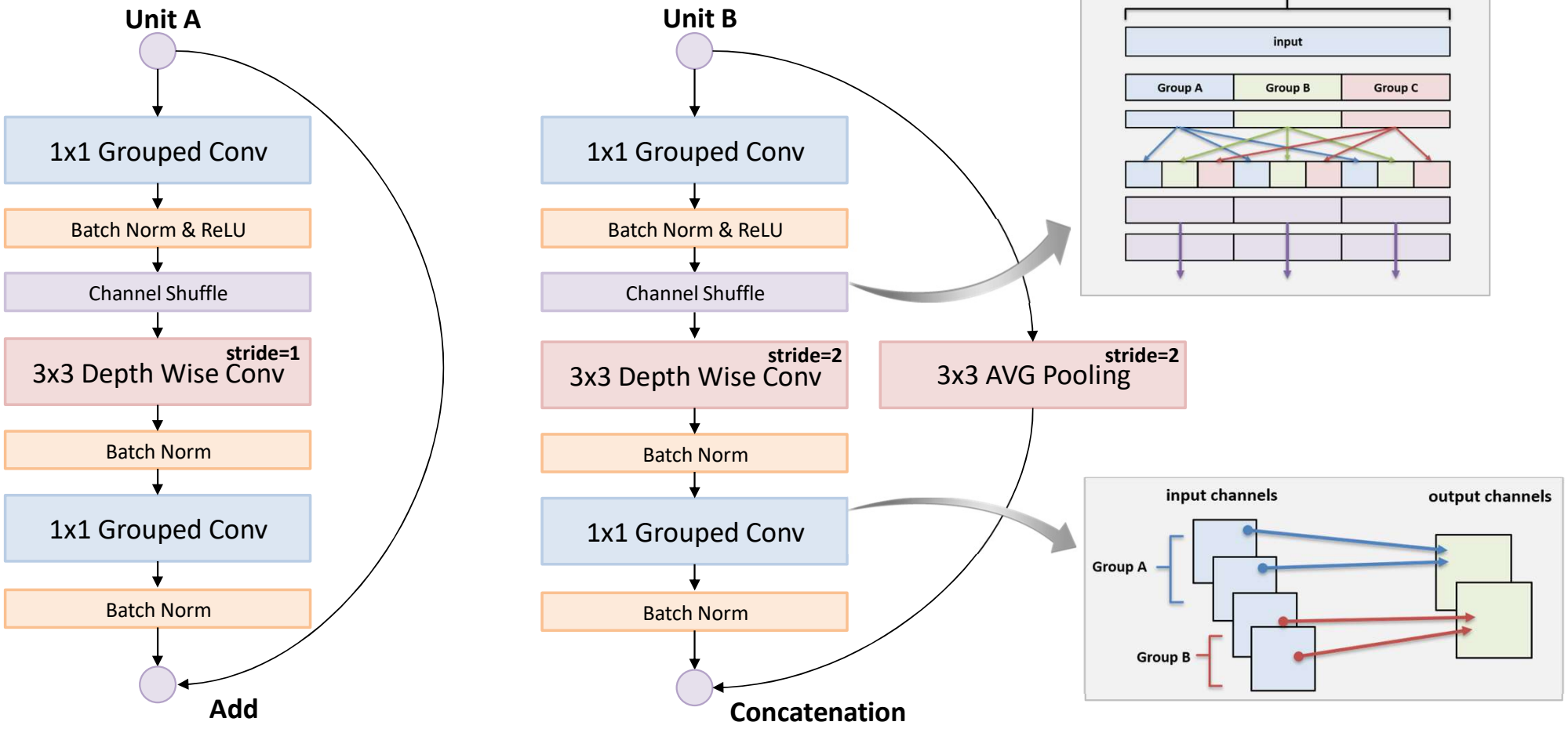
- Channel Shuffle



- Shuffling channels of each group
- each output channel can represent all groups

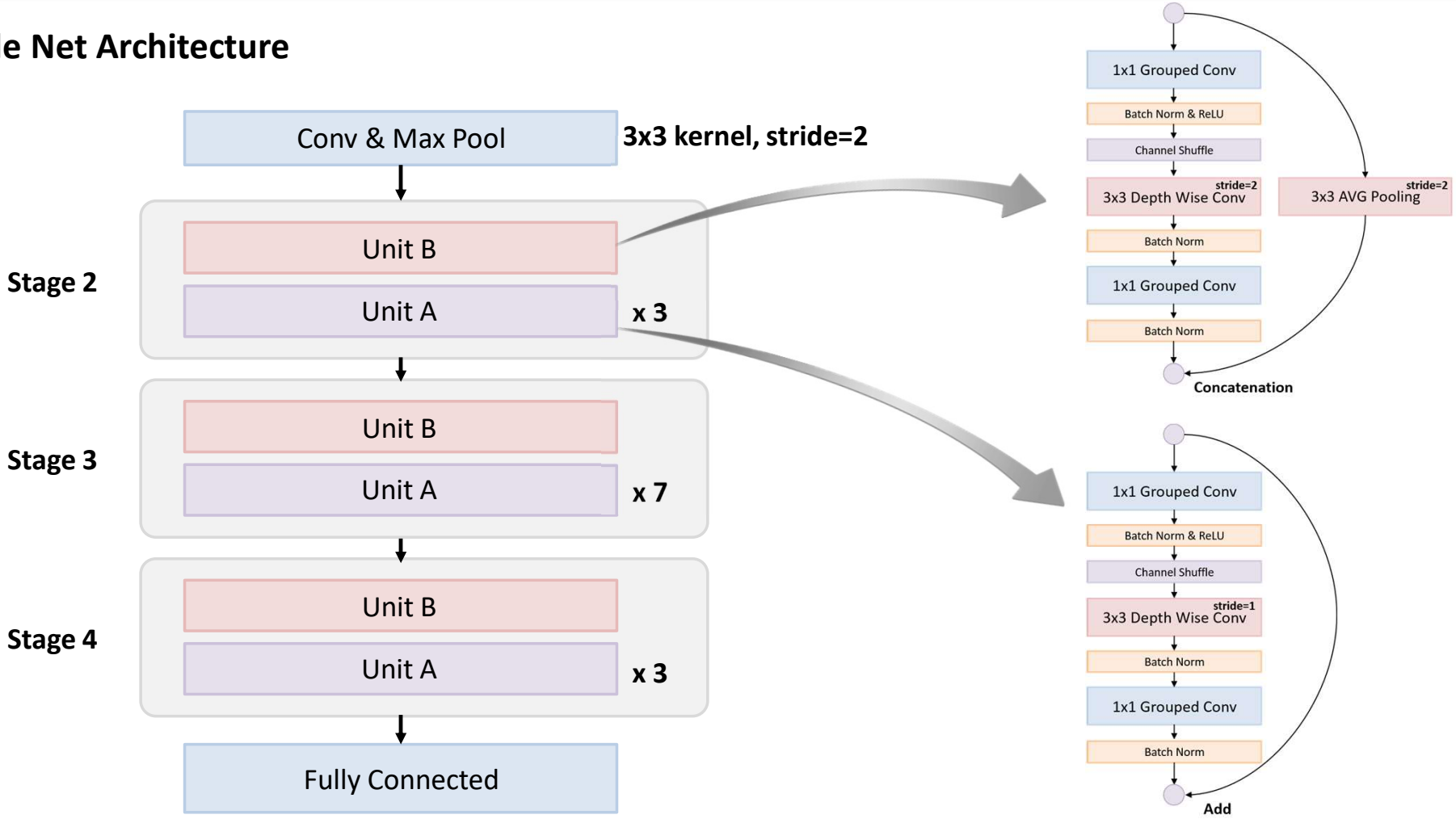
Proposed Method

- **Shuffle Net Units**



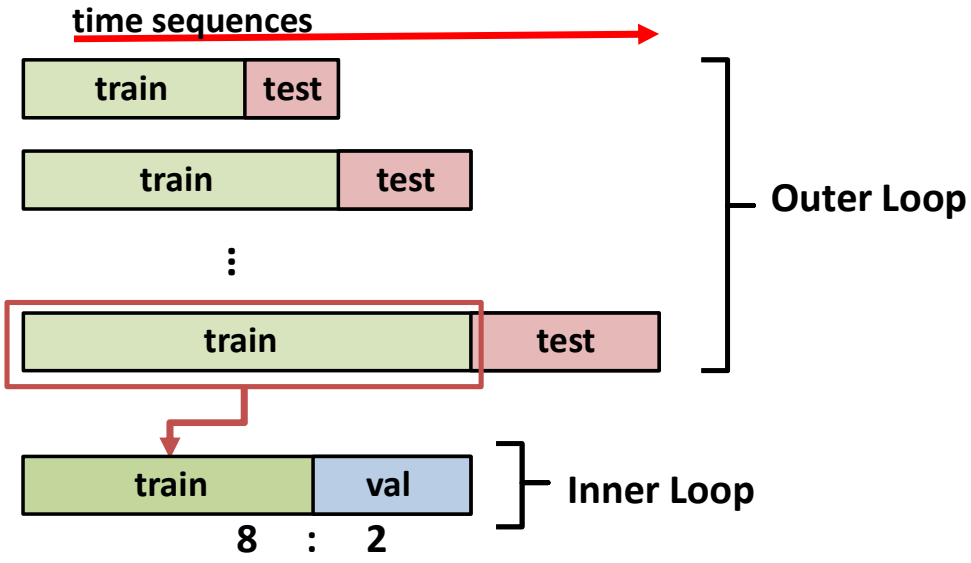
Proposed Method

- **Shuffle Net Architecture**



Experimental setting

- **Nested Cross Validation for time series**



- ➢ **train past data & predict future data**
- ➢ **handling time series data in Nested CV**

- hyperparameter

Model	CNN (ShuffleNet)	LSTM
fold num	10	
hidden size	64	
input length	24	
output length	1	
epoch	30	
batch size	32 or 64	
learning rate	**0.01, 0.001**	

- ➢ batch size 32 for **ISE, Chickenpox** dataset
- ➢ batch size 64 for others

Experimental setting

- **Statistics**

Dataset		Mean	Standard Deviation	Skewness	Kurtosis	Jarque Bera	ADF
Air Quality		1099.833	217.080	0.755	0.334	897.811	-9.823
Appliances Energy		97.694	102.524	3.386	13.664	191240.781	-21.616
Beijing Air Pollution		98.613	92.050	1.802	4.768	62162.743	-20.606
Hungarian Chickenpox		101.245	76.354	0.948	1.400	121.010	-6.933
Electric Consumption		1.483	1.194	1.141	1.732	3421.560	-5.360
Eye State		0.448	0.497	0.205	-1.957	2497.788	-4.803
Istanbul Stock Exchange	ISE	0.0016	0.016	-0.094	1.391	44.022	-22.746
	SP	0.0006	0.014	-0.090	3.068	211.013	-11.031
	DAX	0.0007	0.014	-0.106	2.104	99.894	-23.056
Metro Volume		3259.818	1986.860	-0.089	-1.309	3506.110	-28.016
Occupancy Detection		0.212	0.408	1.406	-0.020	2686.285	-3.430
SML 2010		20.493	3.312	-0.169	-0.105	21.792	-4.314

*ISE : Istanbul Stock Exchange, *SP : Standard & Poor's, *DAX : Germany

Comparison results

- RMSE (Root Mean Squared Error)

Data		LSTM	Standard Deviation	CNN (ShuffleNet)	Standard Deviation
Air Quality		104.545	24.832	145.833	25.764
Appliances Energy		73.348	22.123	75.966	9.782
Beijing Air Pollution		26.776	8.101	49.456	10.976
Hungary Chickenpox		57.977	18.244	66.446	17.458
Electric Consumption		0.294	0.238		
Eye State		3.010	4.771	0.616	0.090
Istanbul Stock Exchange	ISE	6.654	8.861		
	SP	6.497	8.738		
	DAX	6.016	7.917		
Metro Interstate Traffic Volume		522.602	129.710		
Occupancy Detection		1.981	2.641		
SML 2010		0.285	0.192		

Comparison results

- MAE (Mean Absolute Error)

Data		LSTM	Standard Deviation	CNN (ShuffleNet)	Standard Deviation
Air Quality		80.097	24.090	111.932	22.299
Appliances Energy		42.697	20.164	37.614	7.429
Beijing Air Pollution		12.167	0.899	32.908	6.590
Hungary Chickenpox		41.321	15.165	48.717	11.758
Electric Consumption		0.121	0.086		
Eye State		2.960	4.789	0.491	0.083
Istanbul Stock Exchange	ISE	6.652	8.862		
	SP	6.495	8.739		
	DAX	6.014	7.919		
Metro Interstate Traffic Volume		340.367	72.239		
Occupancy Detection		1.713	2.118		
SML 2010		0.233	0.148		

Thank you

Appendix : Shuffle Net Architecture

Layer	Output size	KSize	Stride	Repeat	Output channels (g groups)				
					$g = 1$	$g = 2$	$g = 3$	$g = 4$	$g = 8$
Image	224×224				3	3	3	3	3
Conv1	112×112	3×3	2	1	24	24	24	24	24
MaxPool	56×56	3×3	2						
Stage2	28×28		2	1	144	200	240	272	384
	28×28		1	3	144	200	240	272	384
Stage3	14×14		2	1	288	400	480	544	768
	14×14		1	7	288	400	480	544	768
Stage4	7×7		2	1	576	800	960	1088	1536
	7×7		1	3	576	800	960	1088	1536
GlobalPool	1×1	7×7							
FC					1000	1000	1000	1000	1000
Complexity					143M	140M	137M	133M	137M

Appendix : Code

Model	view with code
LSTM	https://colab.research.google.com/drive/1oNW4HzBzmrjyWcWZbjFMtksP1vTV9IO2?usp=sharing
CNN	https://colab.research.google.com/drive/1Ccd7l65bJwiD2HW8XmOKTHO1PEUMuQI3?usp=sharing
LSTM-Beijing air pollution	https://colab.research.google.com/drive/1De0Tn_iD5NQBKAu_eKr3PWZwvcnNTu97?usp=sharing
CNN-Beijing air pollution	https://colab.research.google.com/drive/1HmjdyXbwUASZxzS40v4opc9myAuuhcrF?usp=sharing
LSTM-Hungarian chickenpox	https://colab.research.google.com/drive/1eeLLtCJGzQ2u_V09pXF0dHg2Oz4YHfxn?usp=sharing
CNN-Hungarian chickenpox	https://colab.research.google.com/drive/1E4PywHa-xACYhHTiMiHRwEkCHT9eBz4A?usp=sharing
LSTM-Power consumption	https://colab.research.google.com/drive/1HVLLJMPGJeixJ4H9O3IvzfUhUWp8DAXN?usp=sharing
CNN-Power consumption	https://colab.research.google.com/drive/1d7wu9u7zo_URAXYd2E01uROGcsdoSc87?usp=sharing

Appendix : Dataset

Dataset	references
Air Quality	https://archive.ics.uci.edu/ml/machine-learning-databases/00360/
Appliances Energy	https://archive.ics.uci.edu/ml/machine-learning-databases/00374/
Beijing Air Pollution	https://archive.ics.uci.edu/ml/machine-learning-databases/00381/
Hungarian Chickenpox	https://archive.ics.uci.edu/ml/machine-learning-databases/00580/
Electric Consumption	https://archive.ics.uci.edu/ml/machine-learning-databases/00235/
Eye State	https://archive.ics.uci.edu/ml/machine-learning-databases/00264/
Istanbul Stock Exchange	https://archive.ics.uci.edu/ml/machine-learning-databases/00247/
Metro Volume	https://archive.ics.uci.edu/ml/machine-learning-databases/00492/
Occupancy Detection	https://archive.ics.uci.edu/ml/machine-learning-databases/00357/
SML 2010	https://archive.ics.uci.edu/ml/machine-learning-databases/00274/

Appendix : References

[1] : Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 6848-6856).