



The Report about CirCNN

authors: Tianma Shen

University of Shanghai for Science and Technology

December 8, 2017



Outline

- 1 Background and Motivation
- 2 Novelty of CirCNN
- 3 Related Knowledge
- 4 CirCNN Algorithms



Outline

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

1 Background and Motivation

2 Novelty of CirCNN

3 Related Knowledge

4 CirCNN Algorithms



Background

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

- Large-scale deep neural networks (DNNs)
- Limitations of computer performance
- Difficult tasks with big data



Motivation

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

- Reduce weight storage (model size)
- Accelerate the computation
- Maintain accuracy



Outline

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

1 Background and Motivation

2 Novelty of CirCNN

3 Related Knowledge

4 CirCNN Algorithms



Novelty of CirCNN

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

- Supporting both FC and CONV layers
- Block-circulant matrices



Outline

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

1 Background and Motivation

2 Novelty of CirCNN

3 Related Knowledge

4 CirCNN Algorithms



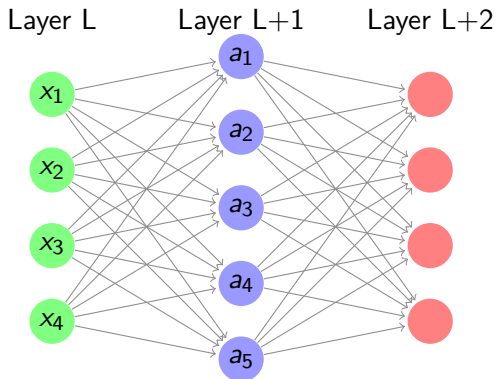
Full Connect Layers

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms



$$a_1 = f(W_{11} \cdot x_1 + W_{12} \cdot x_2 + W_{13} \cdot x_3 + W_{14} \cdot x_4) \quad (1)$$

$$a_5 = f(W_{51} \cdot x_1 + W_{52} \cdot x_2 + W_{53} \cdot x_3 + W_{54} \cdot x_4) \quad (2)$$



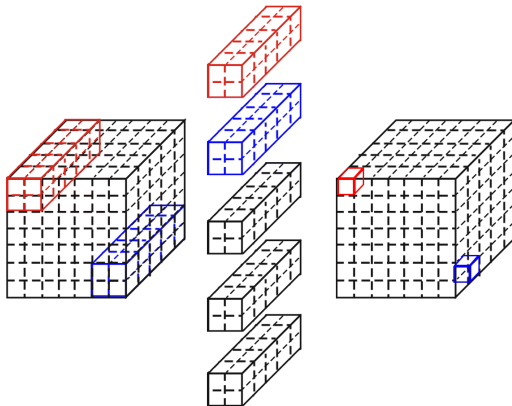
Convolution Layer

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms



$$Y(x, y, p) = \sum_{i=1}^r \sum_{j=1}^r \sum_{c=1}^c F(i, j, c, p) X(x + i - 1, y + j - 1, c) \quad (3)$$



Circulant matrices

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

$$\begin{bmatrix} W_{11} & W_{12} & W_{13} & \dots & W_{1n-2} & W_{1n-1} & W_{1n} \\ W_{1n} & W_{11} & W_{12} & W_{13} & \dots & W_{1n-2} & W_{1n-1} \\ W_{1n-1} & W_{1n} & W_{11} & W_{12} & W_{13} & \dots & W_{1n-2} \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ \cdot & & & & & & \cdot \\ W_{13} & \dots & W_{1n-2} & W_{1n-1} & W_{1n} & W_{11} & W_{12} \\ W_{12} & W_{13} & \dots & W_{1n-2} & W_{1n-1} & W_{1n} & W_{11} \end{bmatrix} \quad (4)$$



Discrete Fourier Transform

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

The discrete Fourier transform transforms a sequence of N numbers x_0, x_1, \dots, x_{N-1} into another sequence of complex numbers, X_0, X_1, \dots, X_{N-1} , which is defined by

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi k \frac{n}{N}} \quad (5)$$

$$= \sum_{n=0}^{N-1} x_n \left[\cos \left(2\pi k \frac{n}{N} \right) - i \cdot \sin \left(2\pi k \frac{n}{N} \right) \right] \quad (6)$$

where the last expression follows from the first one by Euler's formula.

The transform is sometimes denoted by the symbol \mathcal{F} , as in $\mathbf{X} = \mathcal{F}\{\mathbf{x}\}$ or $\mathcal{F}(\mathbf{x})$ or $\mathcal{F}\mathbf{x}$.



Convolution theorem

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

Let \mathcal{F} denote the Fourier transform operator, so $\mathcal{F}\{f\}$ and $\mathcal{F}\{g\}$ are the Fourier transforms of f and g , respectively. Then

$$\mathcal{F}\{f * g\} = \mathcal{F}\{f\} \cdot \mathcal{F}\{g\} \quad (7)$$

$$f * g = \mathcal{F}^{-1}\{\mathcal{F}\{f\} \cdot \mathcal{F}\{g\}\} \quad (8)$$

where \cdot denotes point-wise multiplication.

$*$ denotes convolution.

\mathcal{F}^{-1} denotes inverse Fourier transform.



Outline

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

1 Background and Motivation

2 Novelty of CirCNN

3 Related Knowledge

4 CirCNN Algorithms



Forward propagation process

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

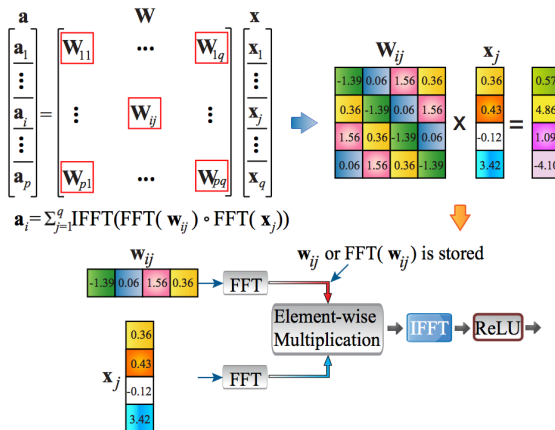


Figure: Illustration of the calculation of W_x in the inference process



Block-circulant matrices

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

0.36	-1.39	0.06	0.43	-0.24	3.42	-0.12	1.56
1.56	0.36	-1.39	0.06	0.43	-0.24	3.42	-0.12
-0.12	1.56	0.36	-1.39	0.06	0.43	-0.24	3.42
3.42	-0.12	1.56	0.36	-1.39	0.06	0.43	-0.24
...
...
...
...

0.36	-1.39	0.06	0.43	-0.24	3.42	-0.12	1.56
-1.39	0.36	0.43	0.06	3.42	-0.24	1.56	-0.12
0.06	1.22	1.72	0.08	1.45	-1.42	0.57	1.47
1.22	0.06	0.08	1.72	-1.42	1.45	1.47	0.57

Figure: when the numbers of inputs and outputs are not equal



Result

Background
and
Motivation

Novelty of
CirCNN

Related
Knowledge

CirCNN
Algorithms

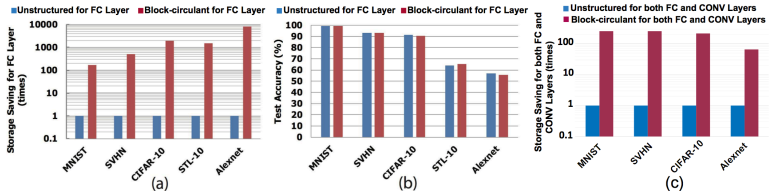


Figure: (a) Storage saving and (b) test accuracy after using block-circulant FC layer for DCNN models on different datasets. (c) Storage saving after using both block-circulant FC layer and block-circulant CONV layer for DCNNs on MNIST, SVHN, CIFAR-10, and ImageNet datasets.