



Broadcast自动扩展

主讲人：龙良曲

Broadcasting

- Expand
- without copying data

前面和后面部分相同
剩下的没有或为一

Key idea

- Insert 1 dim ahead
 - Expand dims with size 1 to same size
 - Feature maps: [4, 32, 14, 14]
 - Bias: [32, 1, 1] => [1, 32, 1, 1] => [4, 32, 14, 14]
-

$$\begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 10 & 10 & 10 \\ \hline 20 & 20 & 20 \\ \hline 30 & 30 & 30 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 10 & 10 & 10 \\ \hline 20 & 20 & 20 \\ \hline 30 & 30 & 30 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline \end{array} =$$



$$\begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 10 & 10 & 10 \\ \hline 20 & 20 & 20 \\ \hline 30 & 30 & 30 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 10 & 10 & 10 \\ \hline 20 & 20 & 20 \\ \hline 30 & 30 & 30 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline 10 & 11 & 12 \\ \hline 20 & 21 & 22 \\ \hline 30 & 31 & 32 \\ \hline \end{array}$$



$$\begin{array}{|c|} \hline 0 \\ \hline 10 \\ \hline 20 \\ \hline 30 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline 0 & 0 & 0 \\ \hline 10 & 10 & 10 \\ \hline 20 & 20 & 20 \\ \hline 30 & 30 & 30 \\ \hline \end{array} + \begin{array}{|c|c|c|} \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline 0 & 1 & 2 \\ \hline \end{array} =$$

Why broadcasting

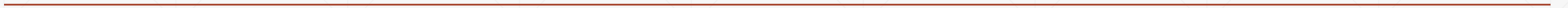
- 1. for actual demanding
 - [class, students, scores]
 - Add bias for every students: +5 score
 - [4, 32, 8] + [4, 32, 8]
 - [4, 32, 8] + [5.0]
 - 2. memory consumption
 - [4, 32, 8] => 1024
 - [5.0] => 1
-

Is it broadcasting-able?

- Match from **Last** dim!
 - If current dim=1, expand to same
 - If either has no dim, insert one dim and expand to same
 - otherwise, NOT broadcasting-able
-

Situation 1:

- $[4, 32, 14, 14]$
- $[1, 32, 1, 1] \Rightarrow [4, 32, 14, 14]$



Situation 2

- $[4, 32, 14, 14]$
 - $[14, 14] \Rightarrow [1, 1, 14, 14] \Rightarrow [4, 32, 14, 14]$
-

Situation 3

- [4, 32, 14, 14]
 - [2, 32, 14, 14]
 - Dim 0 has dim, can NOT insert and expand to same
 - Dim 0 has distinct dim, NOT size 1
 - NOT broadcasting-able
-

How to understand this behavior?

- When it has no dim
 - treat it as all own the same
 - $[class, student, scores] + [scores]$
- When it has dim of size 1
 - Treat it shared by all
 - $[class, student, scores] + [student, 1]$



It's effective and critically, intuitive

- $[4, 3, 32, 32]$

- $+ [32, 32]$

- $+ [3, 1, 1]$

- $+ [1, 1, 1, 1]$

Thank You.
