

Loss

Tensor

Loss
Layer

In

Data

Diff

Layer

Out

Data

Diff

In

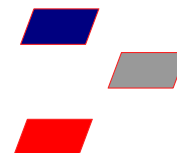
Data

Diff

Learnable
Parameters

Data
Diff

Cache



Layer

Out

Data

Diff

In

Data

Diff

Learnable
Parameters

Data
Diff

Cache

