
DATA 442: Neural Networks & Deep Learning

Dan Runfola – danr@wm.edu

icss.wm.edu/data442/



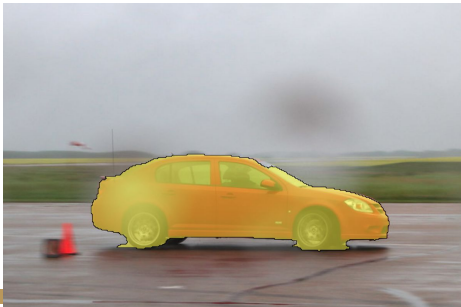
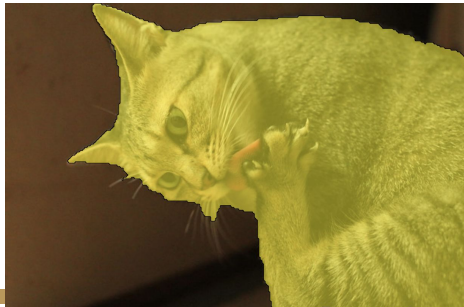
Summary

Total Loss=

$$\sum_{i=1}^{N=3} \{(x_i, y_i)\} \quad \frac{1}{N} \sum_i^N Loss_i(f(x_i, W), y_i) + \lambda R(W)$$

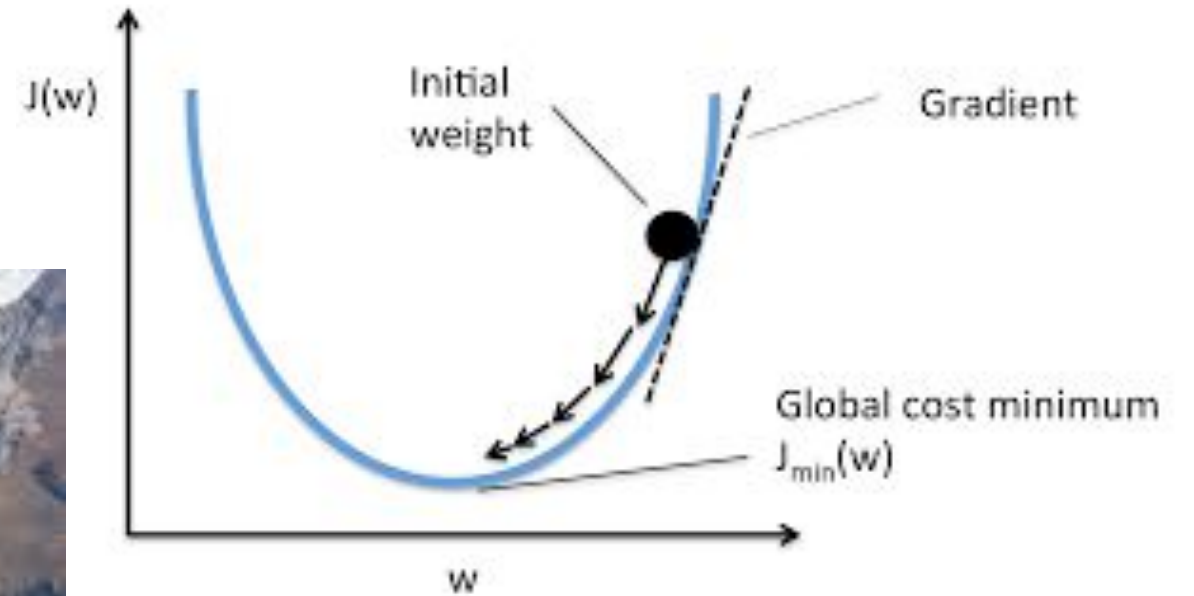
```
def predict(image, W):  
    return(W*image)
```

Cat	3.2	1.3	2.2
Car	5.1	4.9	2.5
Frog	-1.7	2.0	-3.1



$$L_i = -\log\left(\frac{e_k^s}{\sum_{j=1}^J e_j^s}\right)$$

Optimization



ITY CENTER OF EXCELLENCE

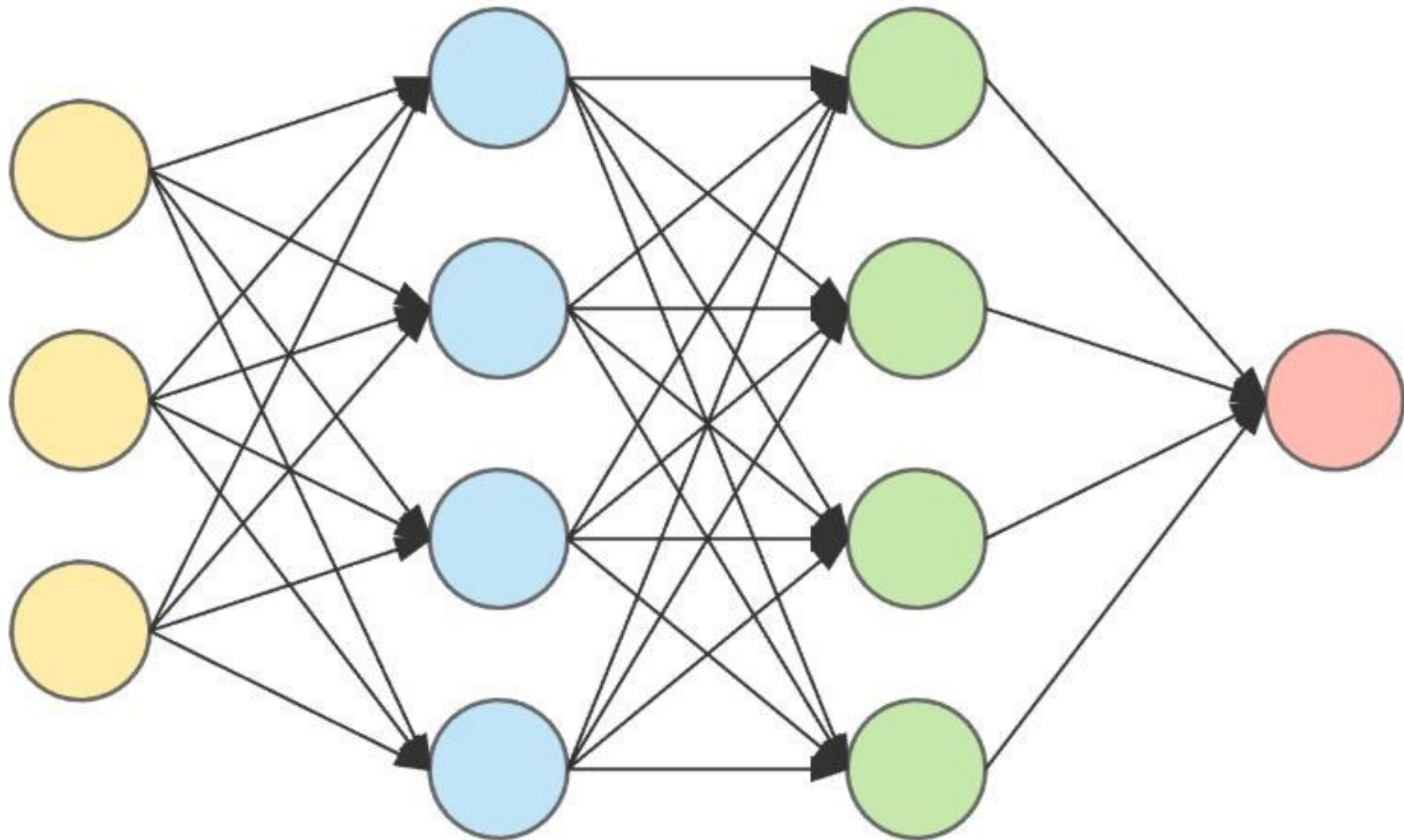
Analytic Gradient

$W = [0.34, -1.11, 0.78, 0.12 \dots 0.3, 0.77]$

$$dw = f(X, W)$$
$$\nabla f(X, W) = [\dots]$$

Gradient

$dW: [-2.5, 0.6, 4.3, 0.5 \dots 0, 0.3]$

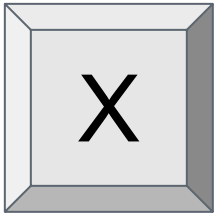


$$f(X, W) = \sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

$f(X, W)$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



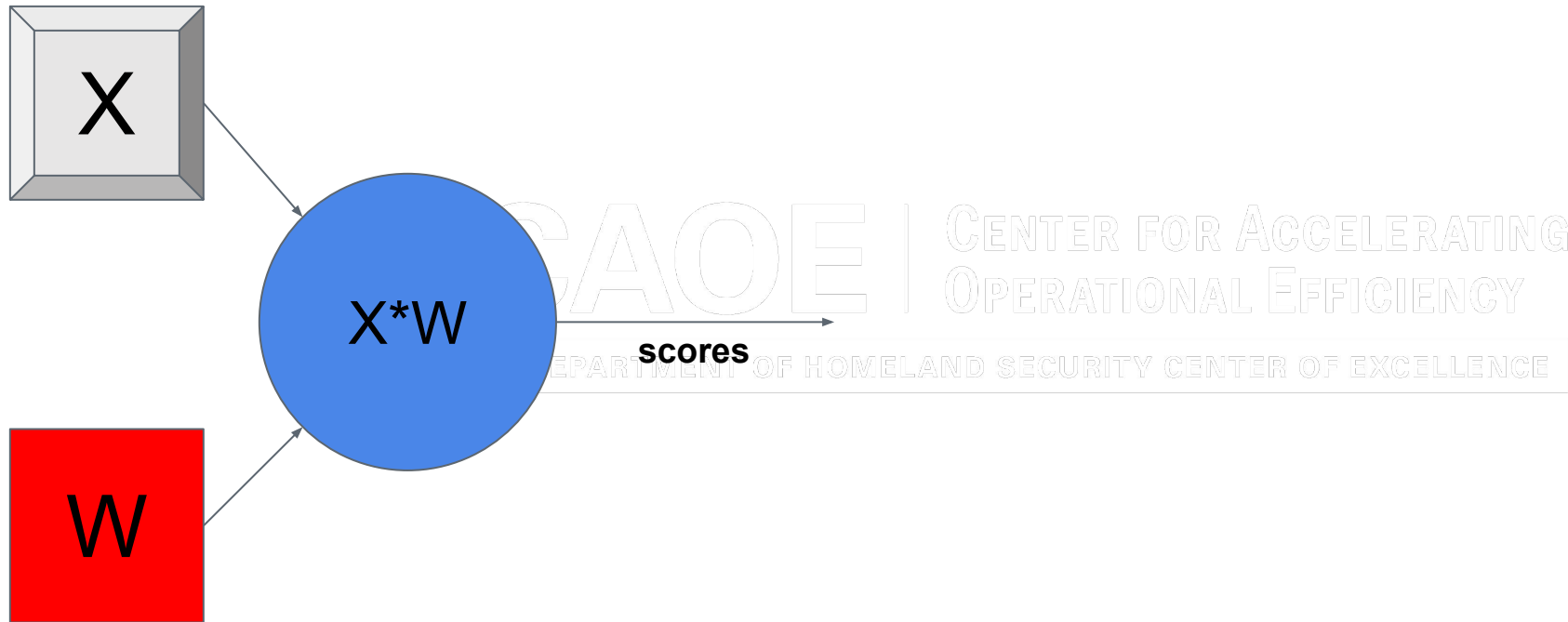
CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



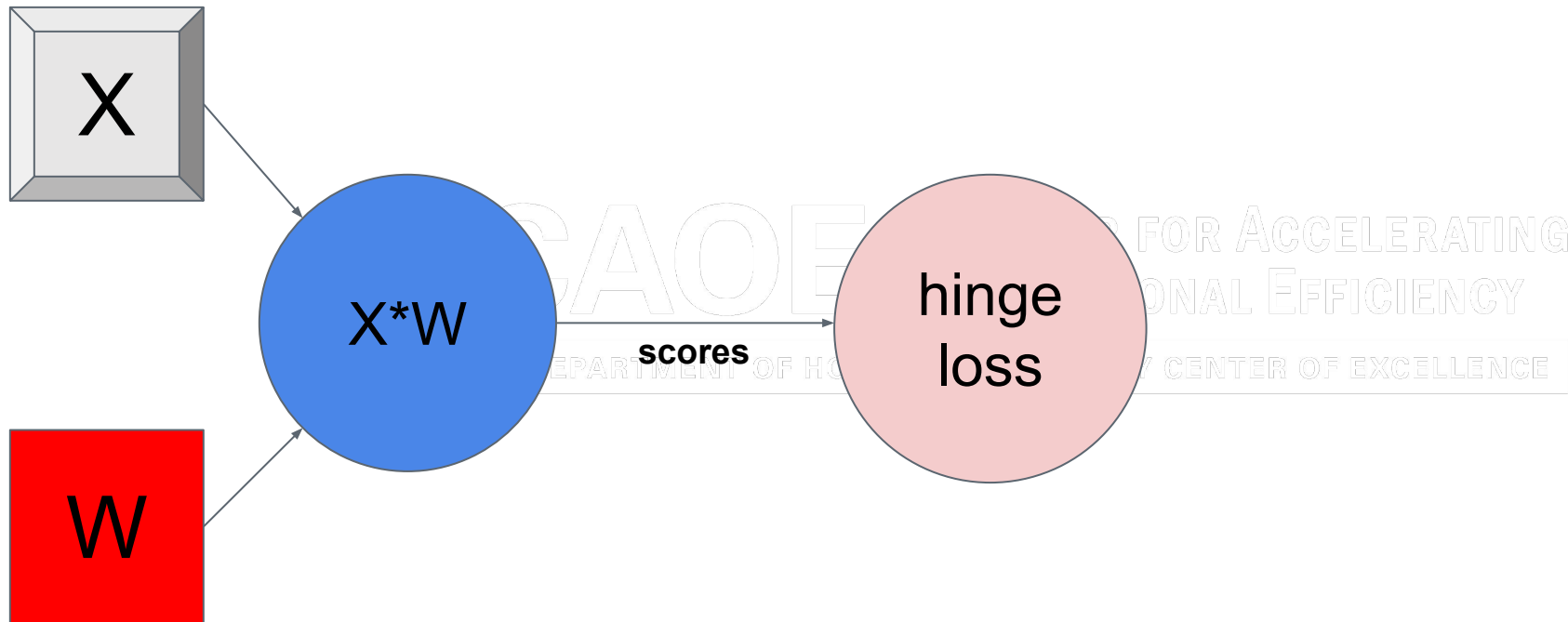
$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



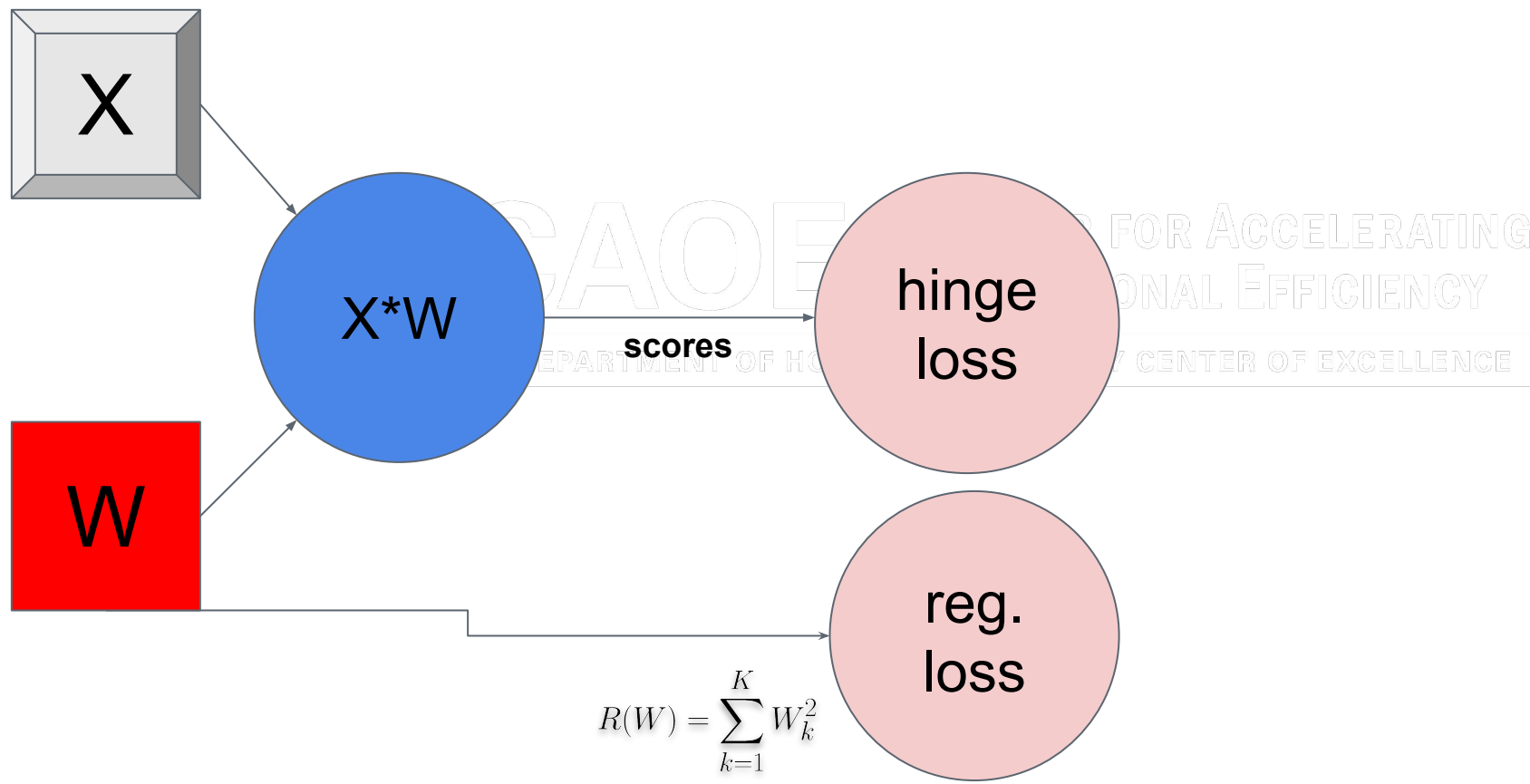
$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



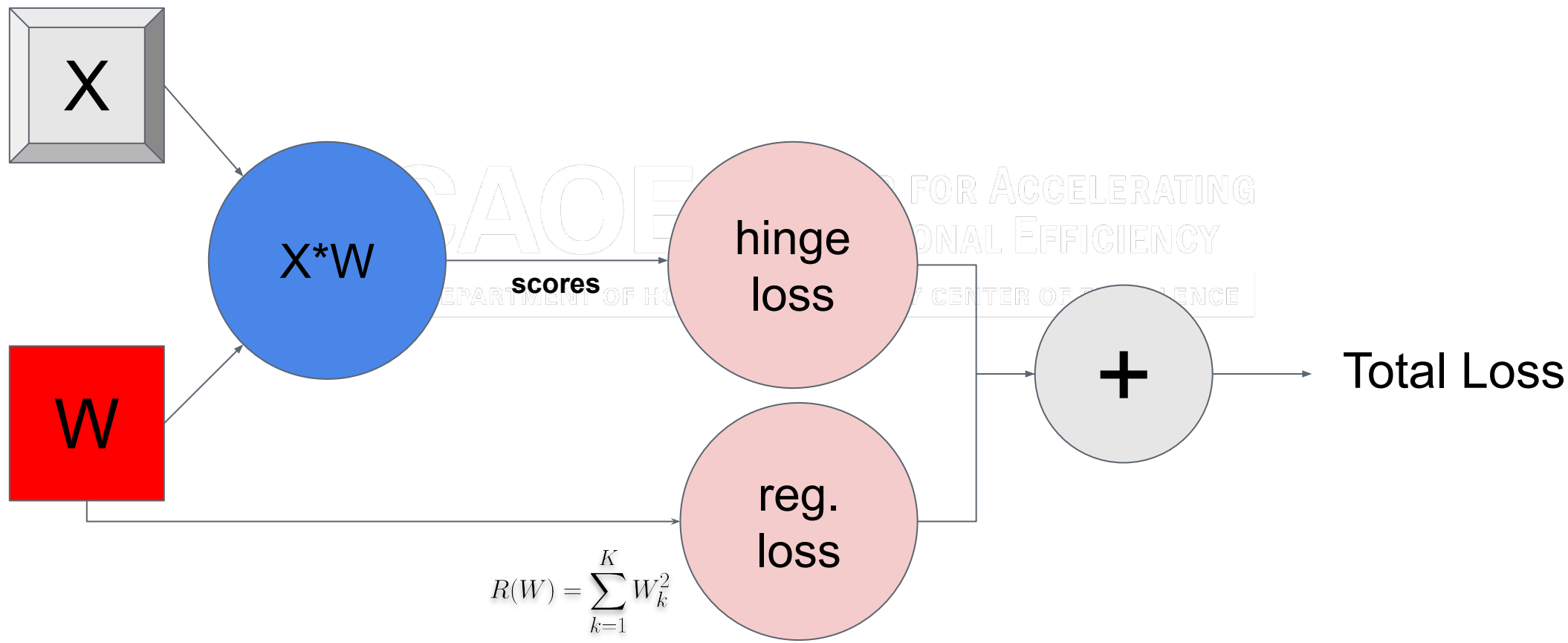
$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

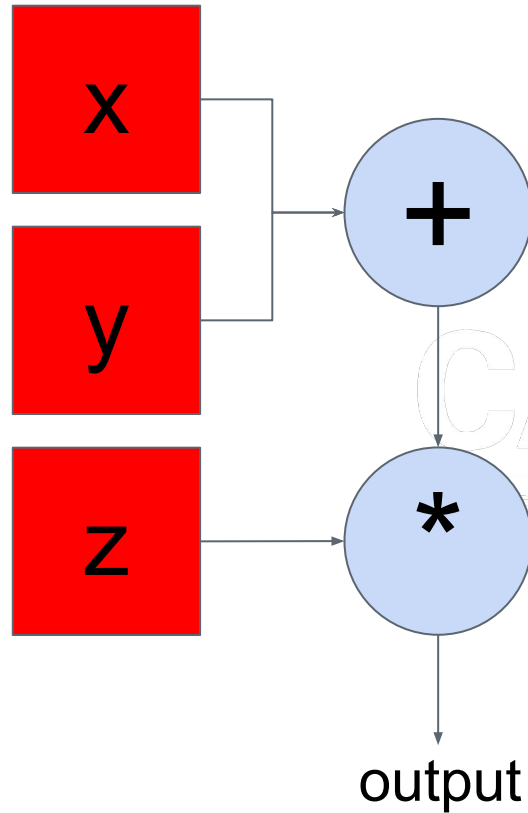


Backpropogation

$$f(x, y, z) = (x + y) * z$$

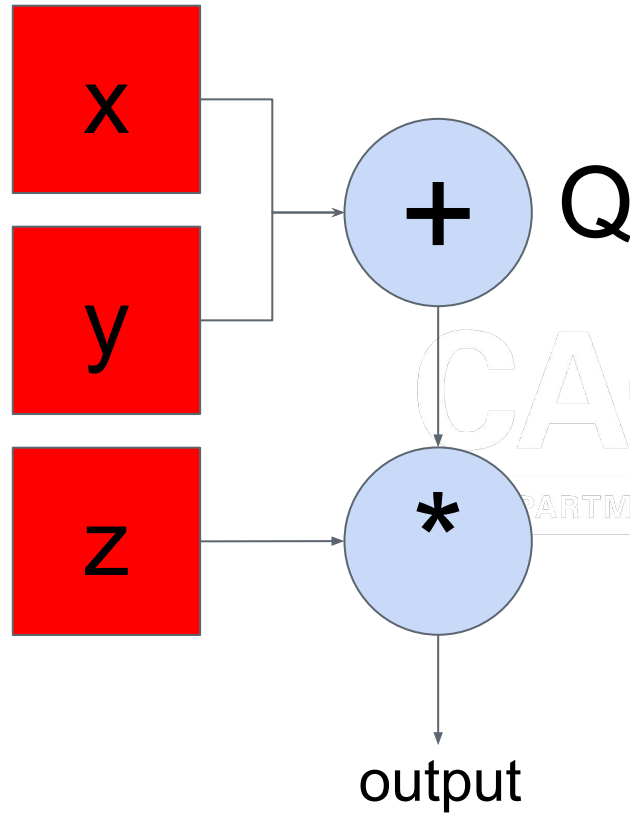
CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

$$f(x, y, z) = (x + y) * z$$



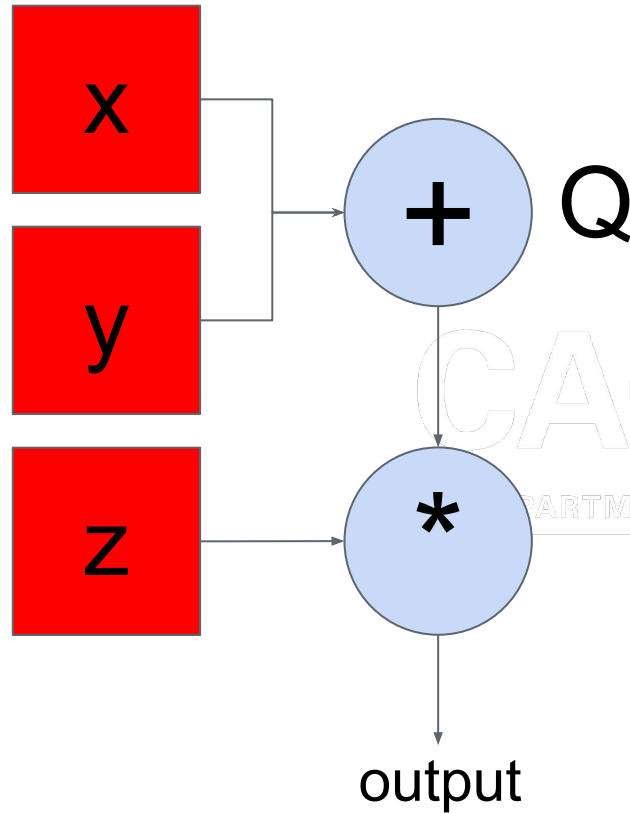
CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

$$f(x, y, z) = (x + y) * z$$



$$Q = x + y$$

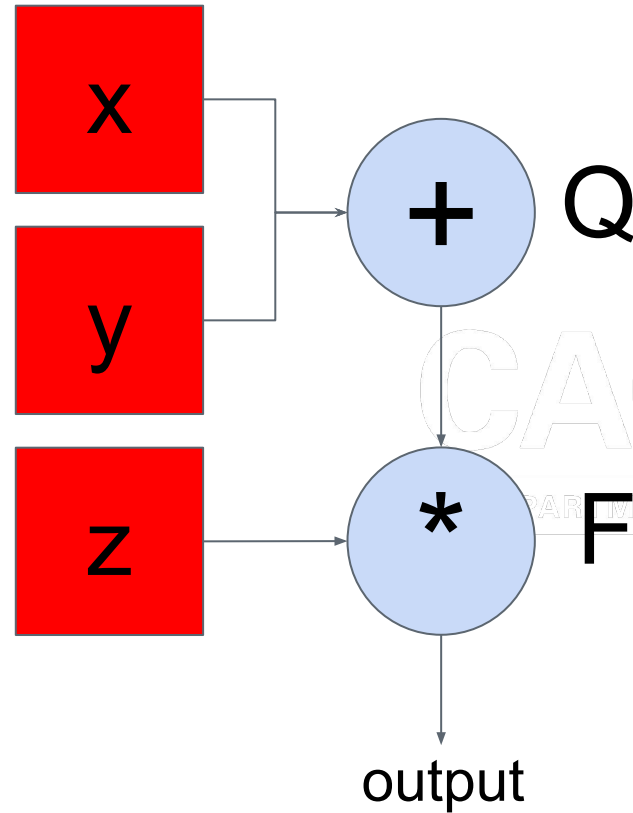
$$f(x, y, z) = (x + y) * z$$



$$Q = x + y$$

$$\frac{\partial q}{\partial x} = 1 \quad \frac{\partial q}{\partial y} = 1$$

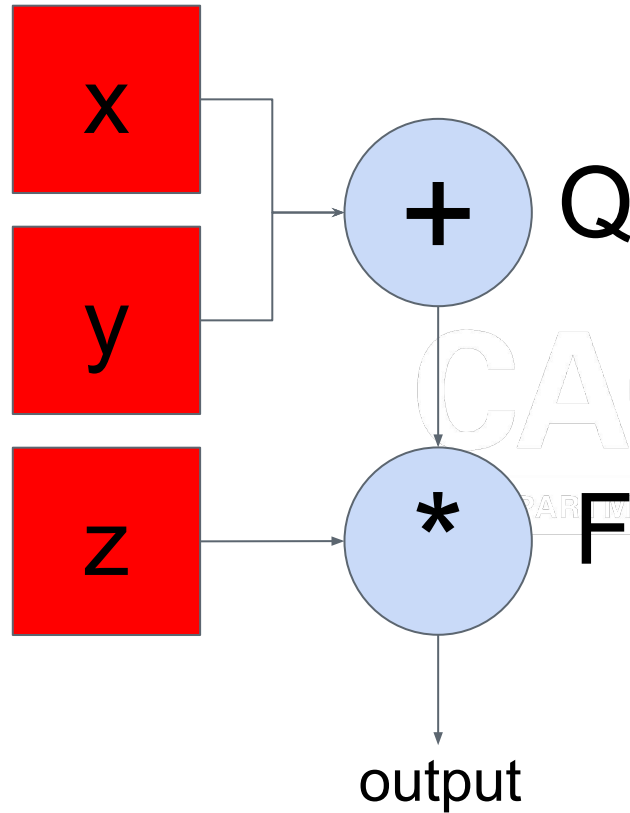
$$f(x, y, z) = (x + y) * z$$



$$F = qz$$

CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

$$f(x, y, z) = (x + y) * z$$

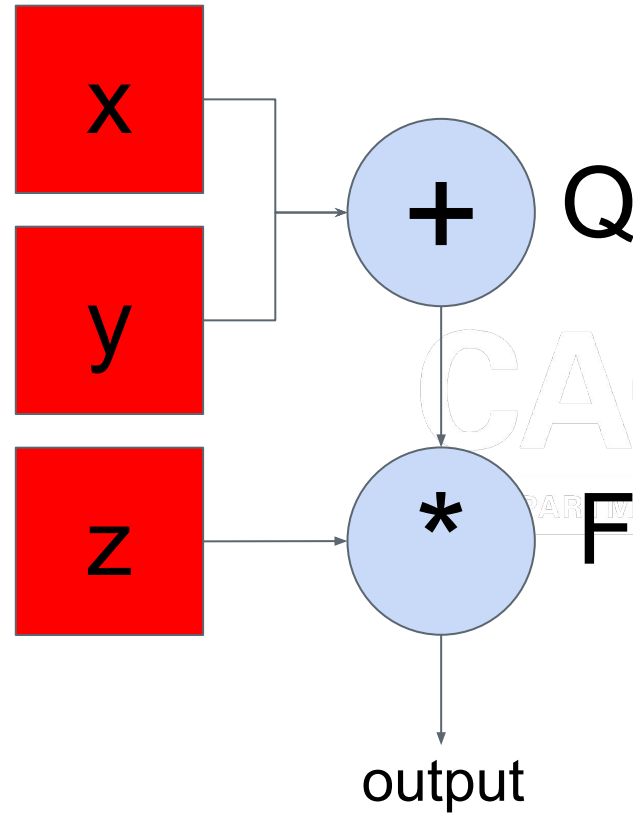


$$F = qz$$

CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

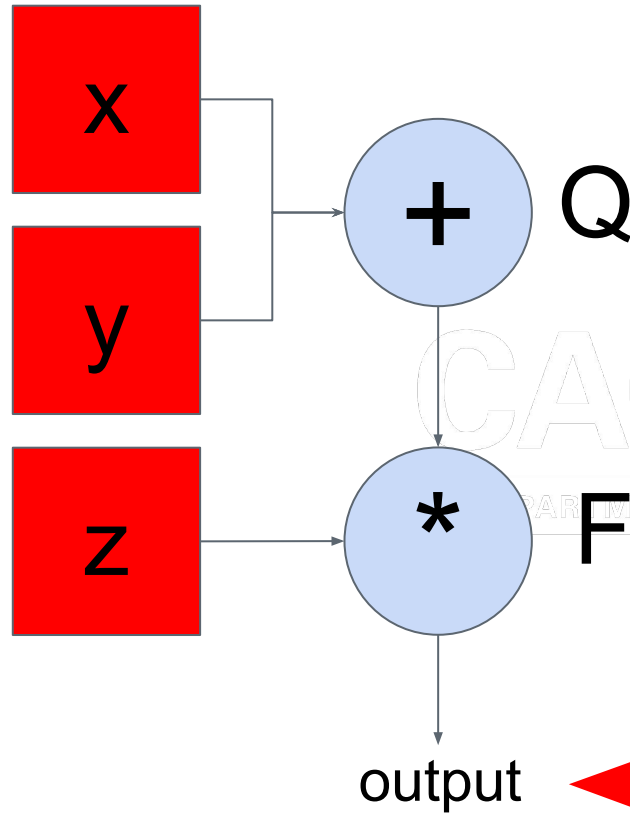
$$\frac{\partial f}{\partial Q} = z \quad \frac{\partial f}{\partial z} = Q$$

$$f(x, y, z) = (x + y) * z$$



$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

$$f(x, y, z) = (x + y) * z$$



CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

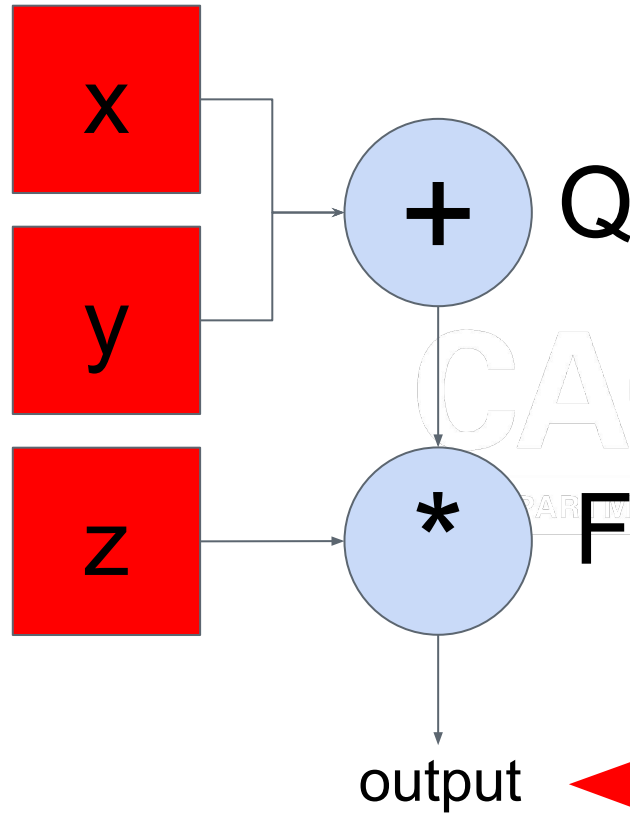
$$\frac{\partial F}{\partial F}$$



The Goal

$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

$$f(x, y, z) = (x + y) * z$$



CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

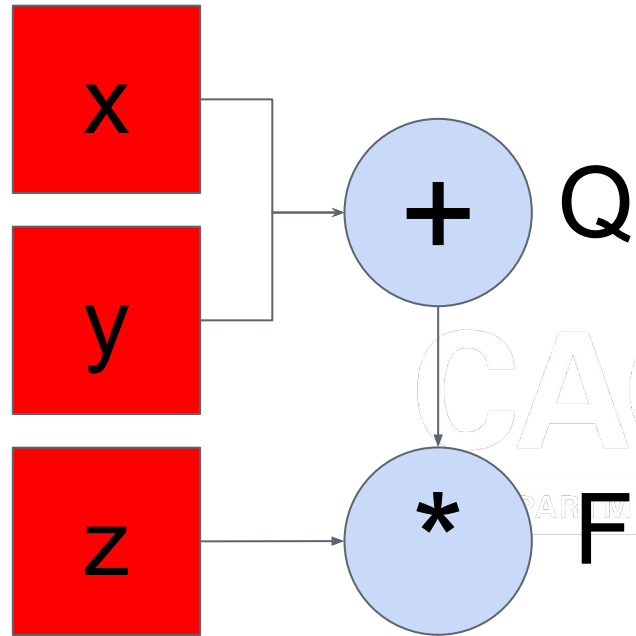
$$\frac{\partial F}{\partial F} = 1$$

←

The Goal

$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

$$f(x, y, z) = (x + y) * z$$



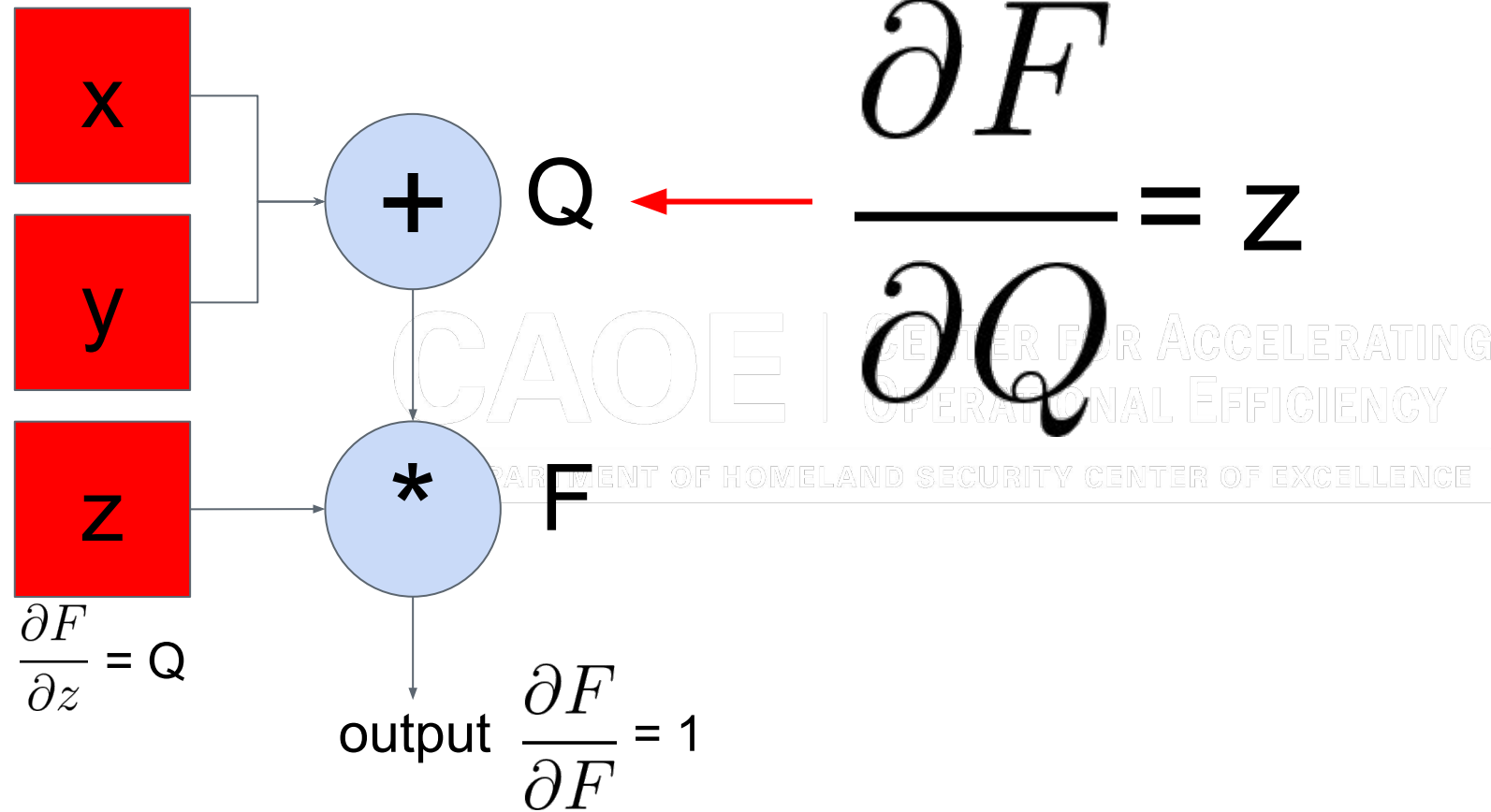
$$\frac{\partial F}{\partial z} = Q$$

output $\frac{\partial F}{\partial F} = 1$

The Goal

$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

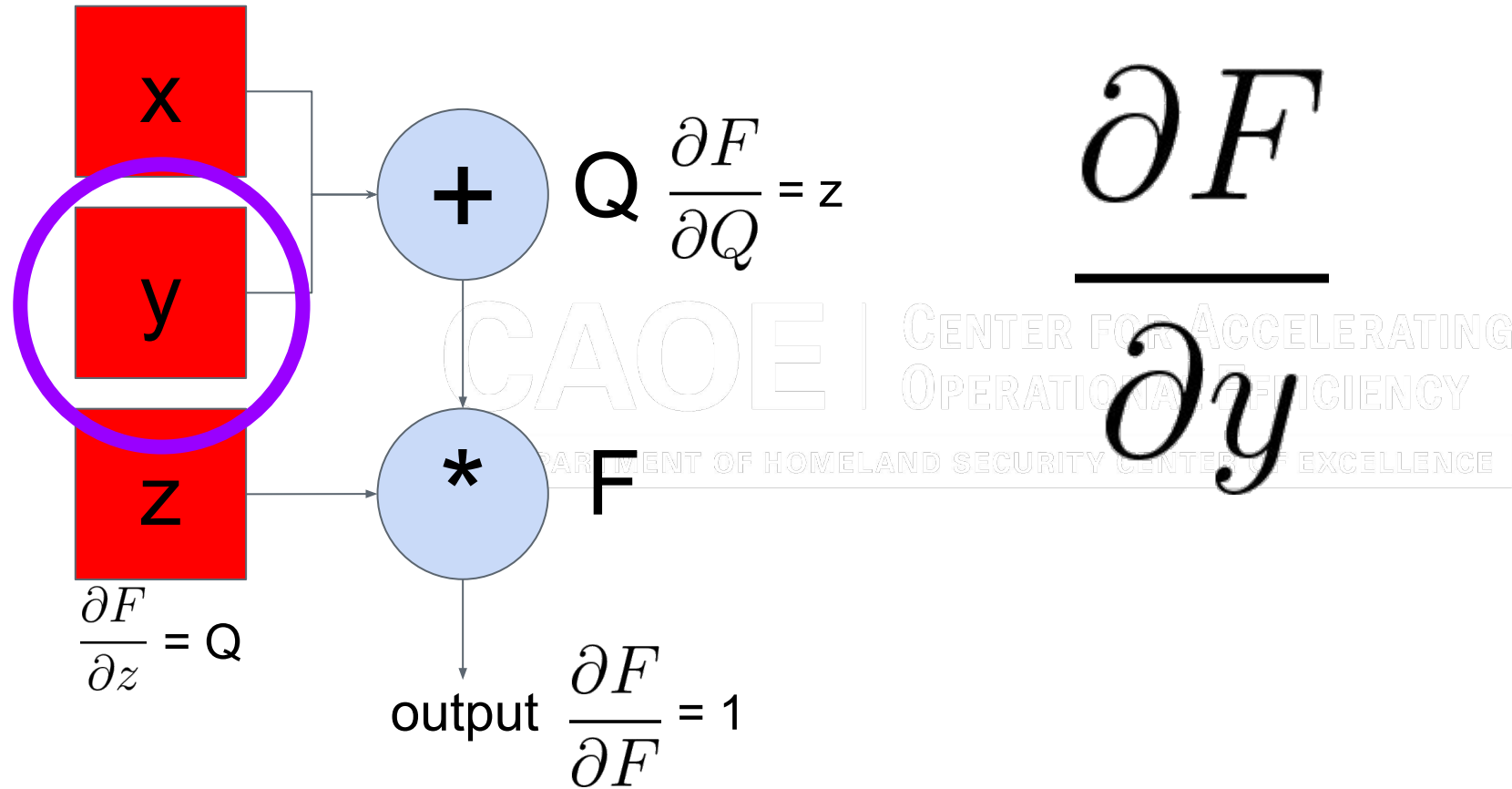
$$f(x, y, z) = (x + y) * z$$



The Goal

$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

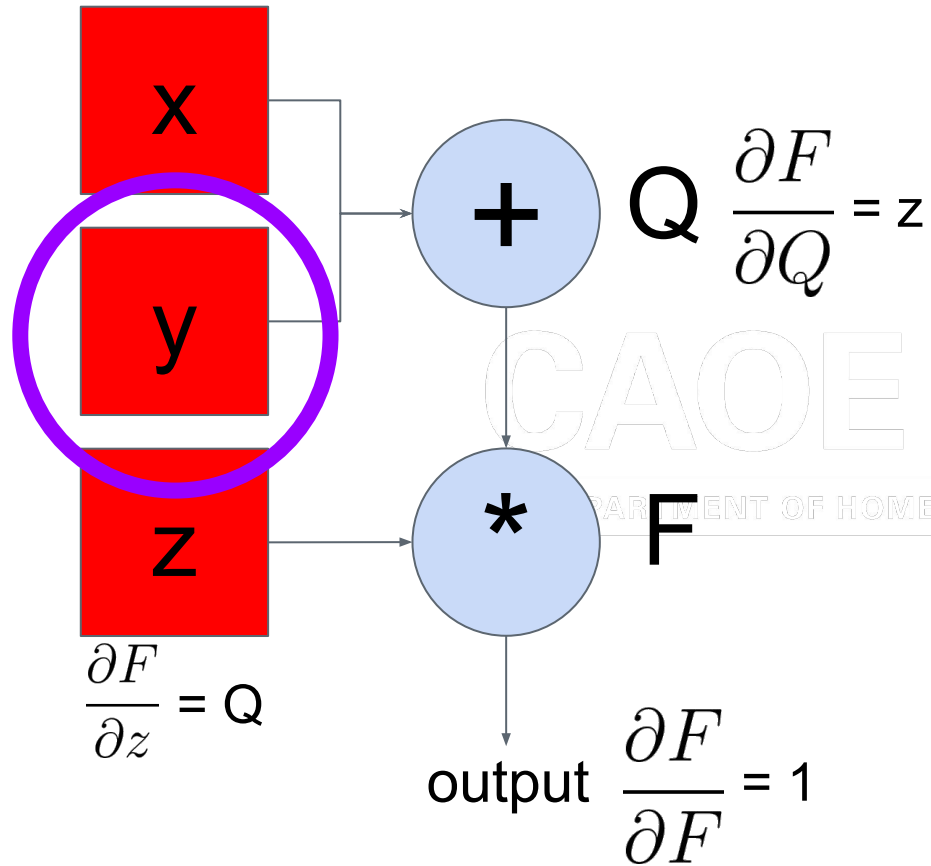
$$f(x, y, z) = (x + y) * z$$



The Goal

$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

$$f(x, y, z) = (x + y) * z$$

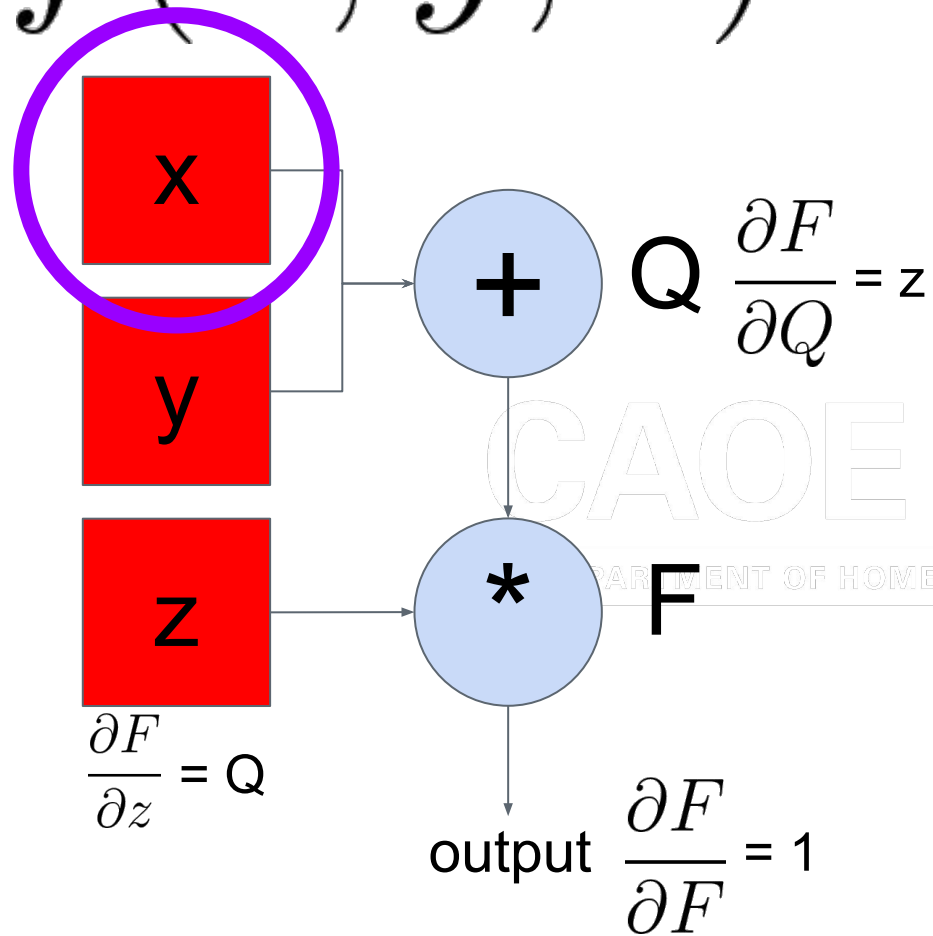


$$\frac{\partial F}{\partial y} = \frac{\partial F}{\partial Q} \frac{\partial Q}{\partial y}$$

The Goal

$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

$$f(x, y, z) = (x + y) * z$$

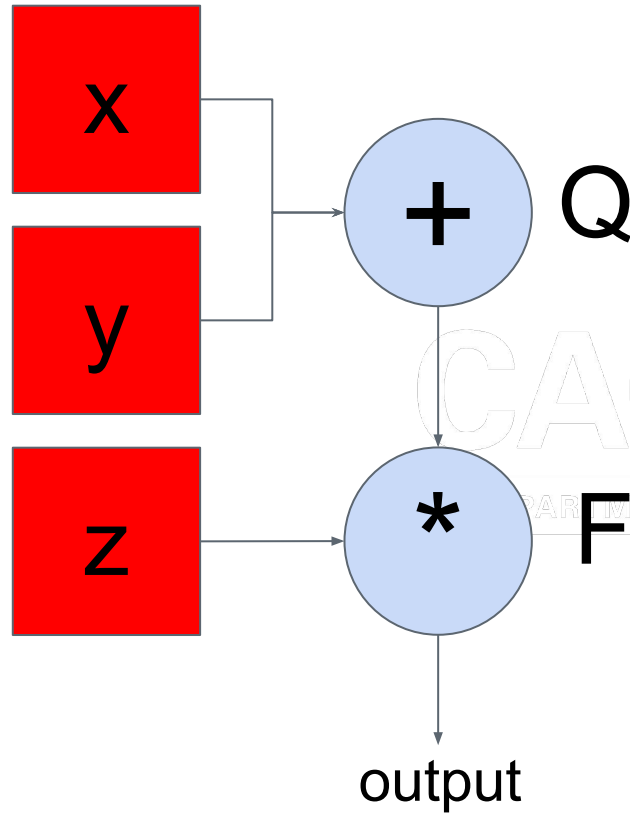


$$\frac{\partial F}{\partial x} = \frac{\partial F}{\partial Q} \frac{\partial Q}{\partial x}$$

The Goal

$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

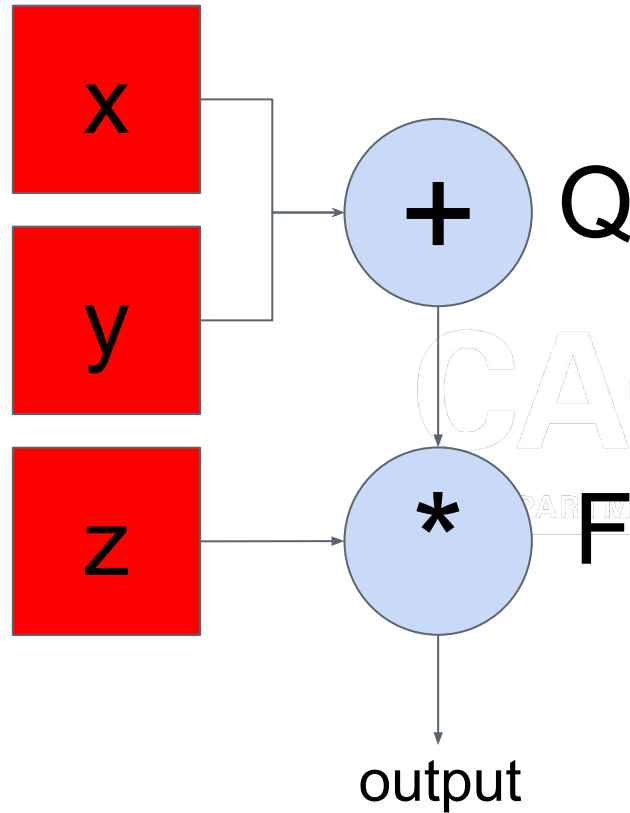
$$f(x, y, z) = (x + y) * z$$



$$\frac{\partial F}{\partial x} = \frac{\partial F}{\partial q} \frac{\partial Q}{\partial x}$$

$$\frac{\partial F}{\partial Q} = z$$

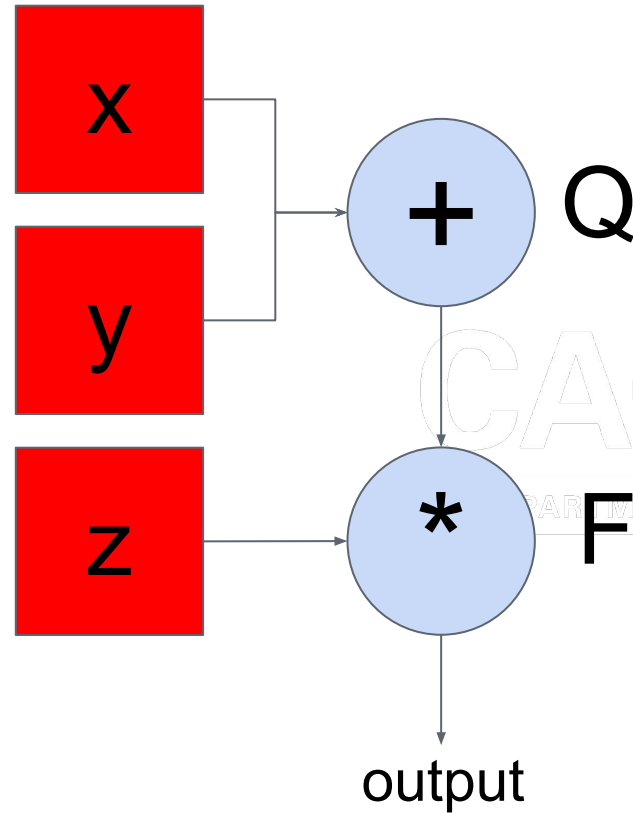
$$f(x, y, z) = (x + y) * z$$



$$\frac{\partial F}{\partial x} = z \frac{\partial Q}{\partial x}$$

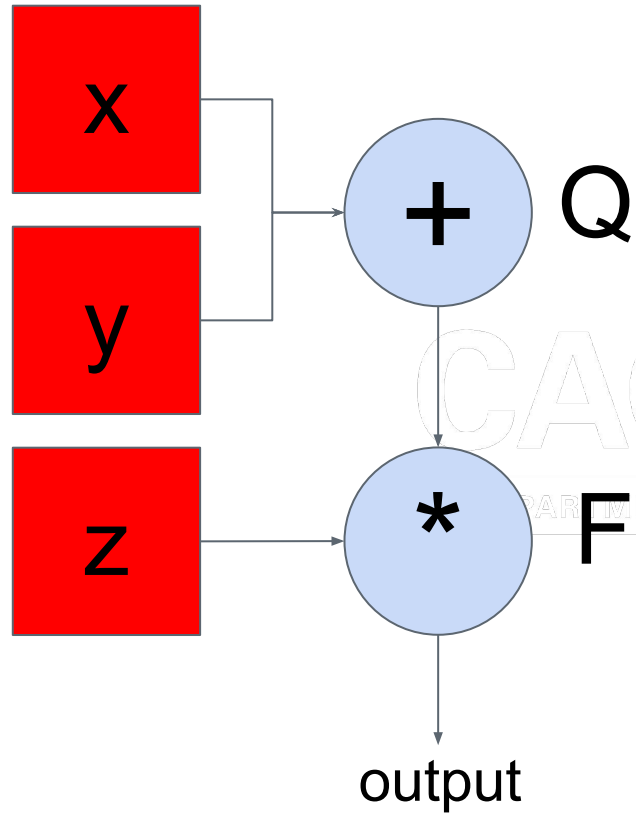
CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

$$f(x, y, z) = (x + y) * z$$



$$\frac{\partial F}{\partial x} = z * 1$$

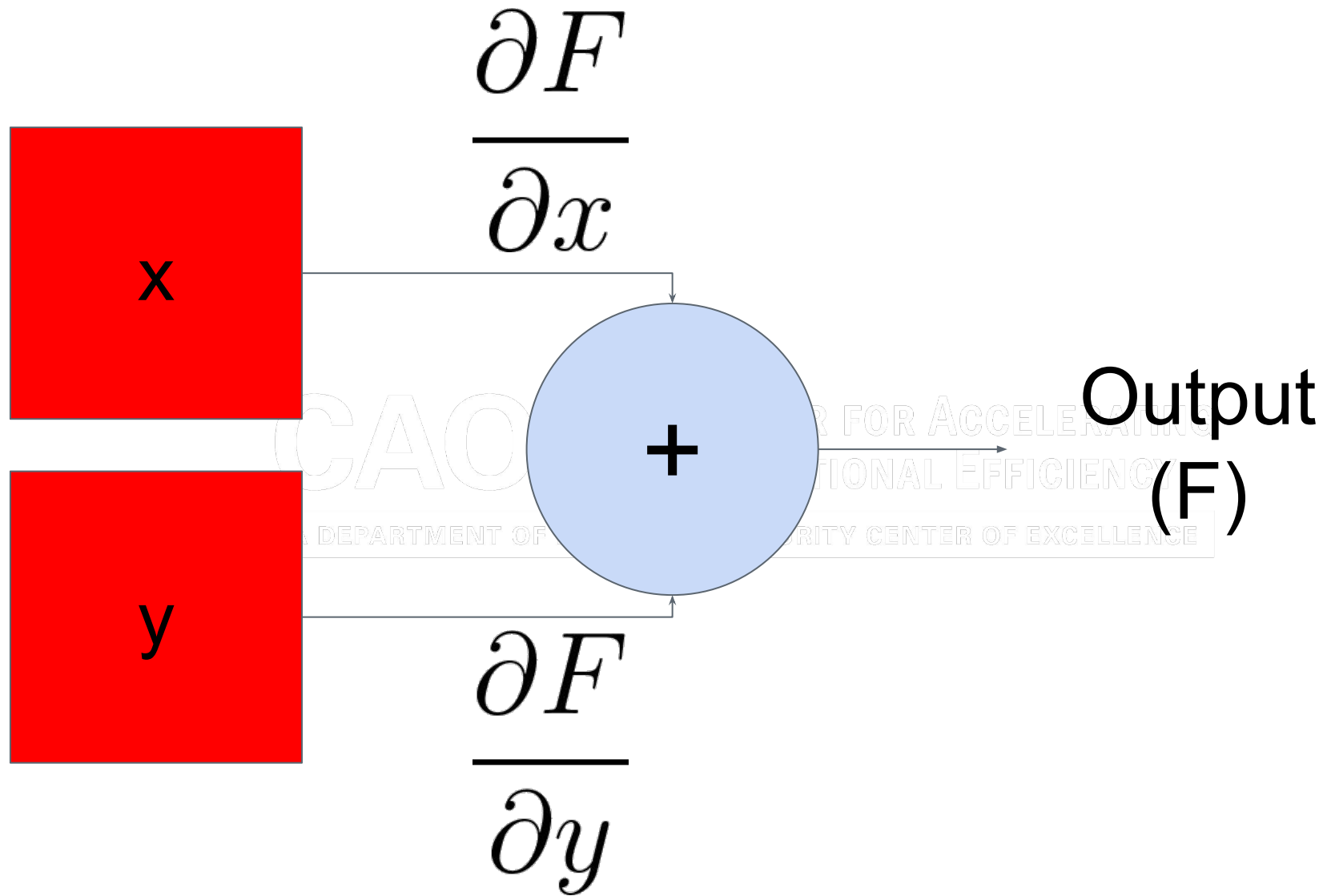
$$f(x, y, z) = (x + y) * z$$



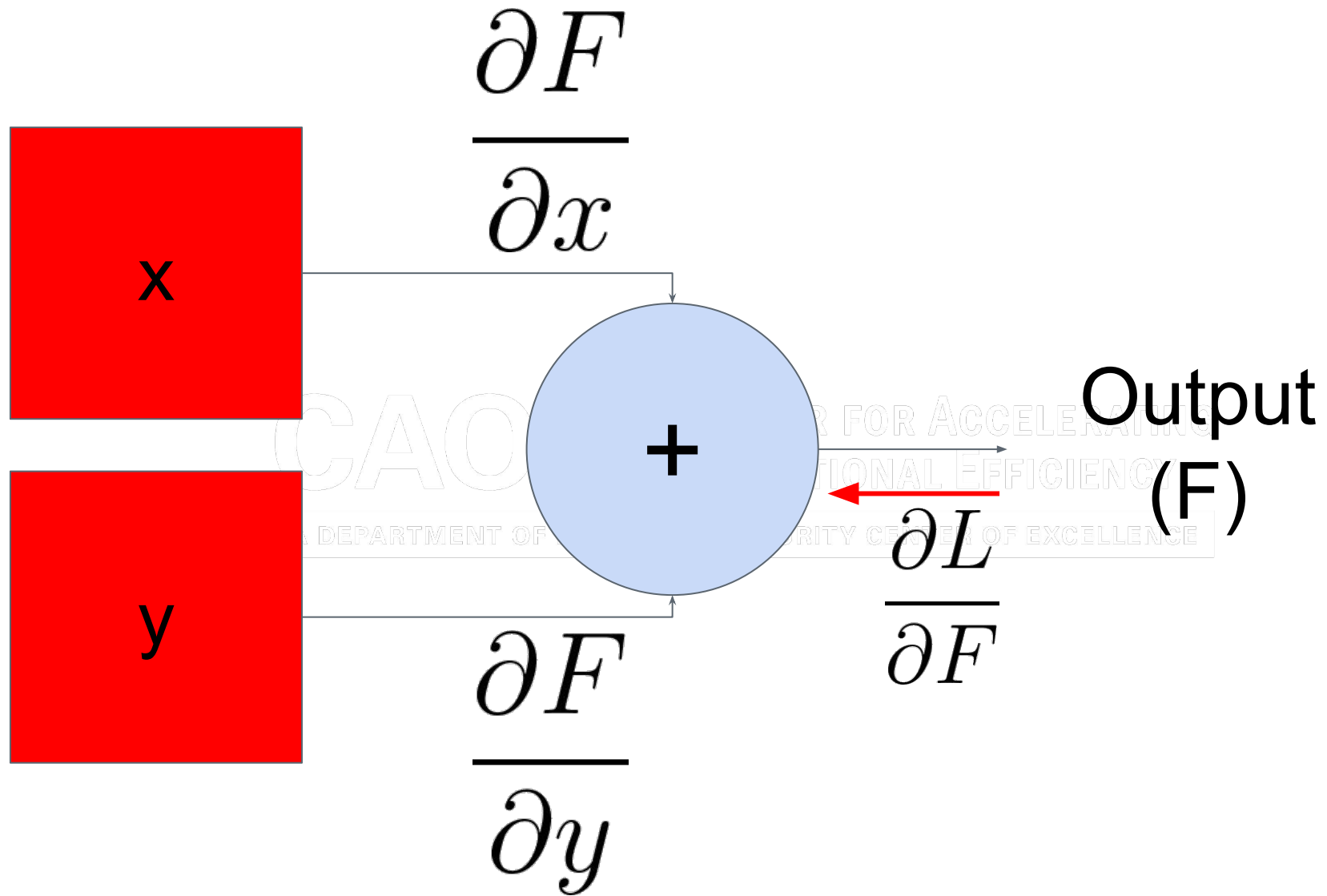
The Goal

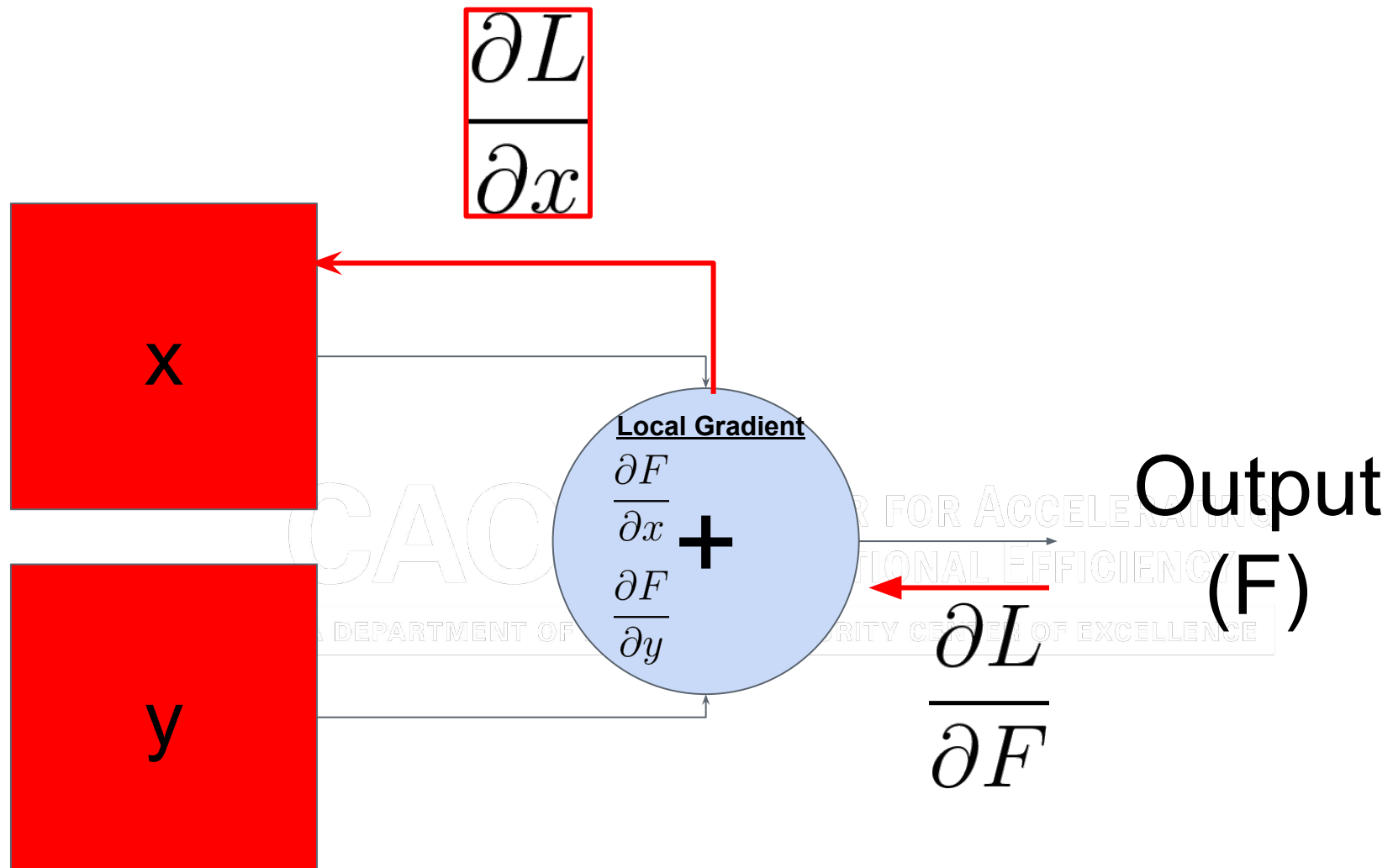
$$\frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \quad \frac{\partial F}{\partial z}$$

Local Gradient

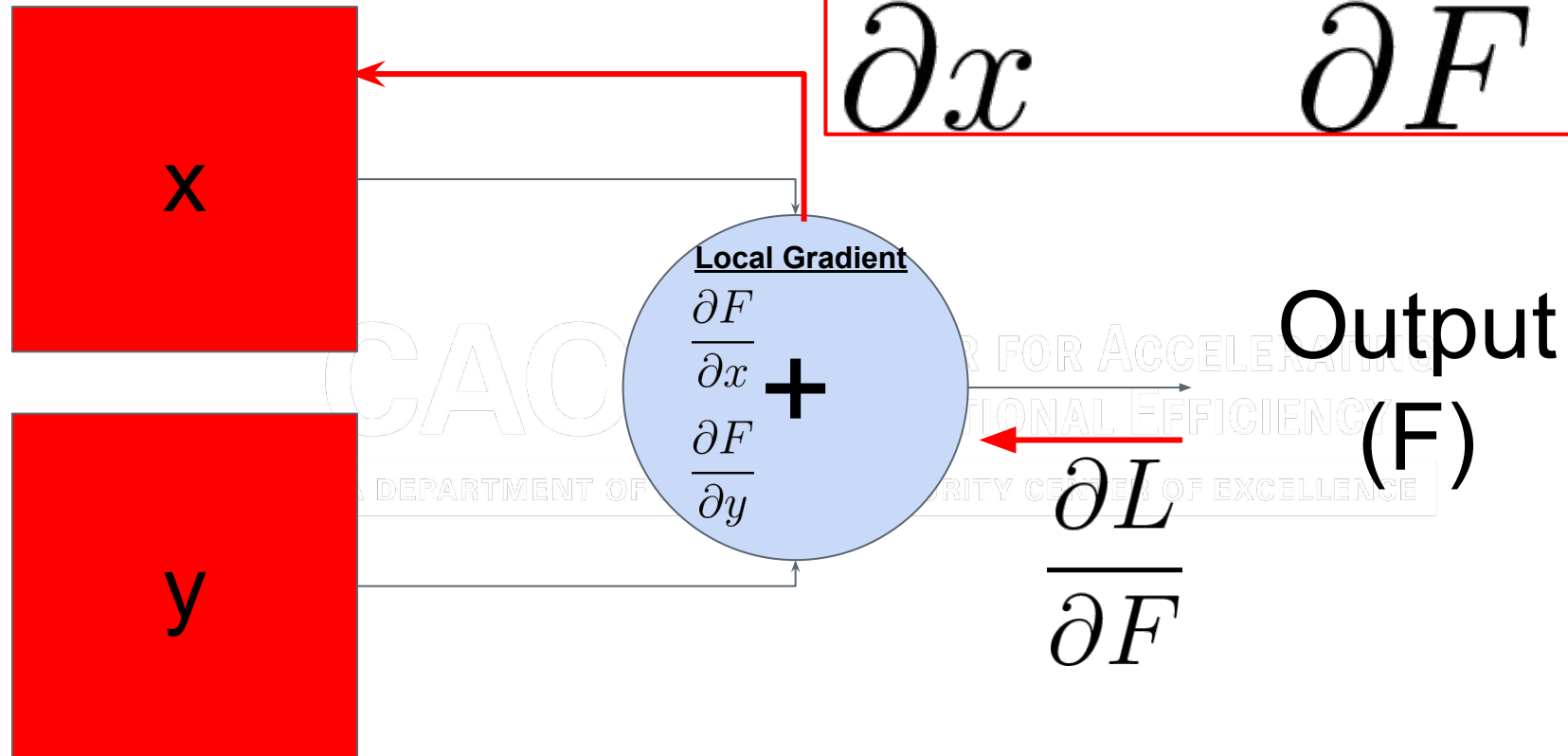


Local Gradient

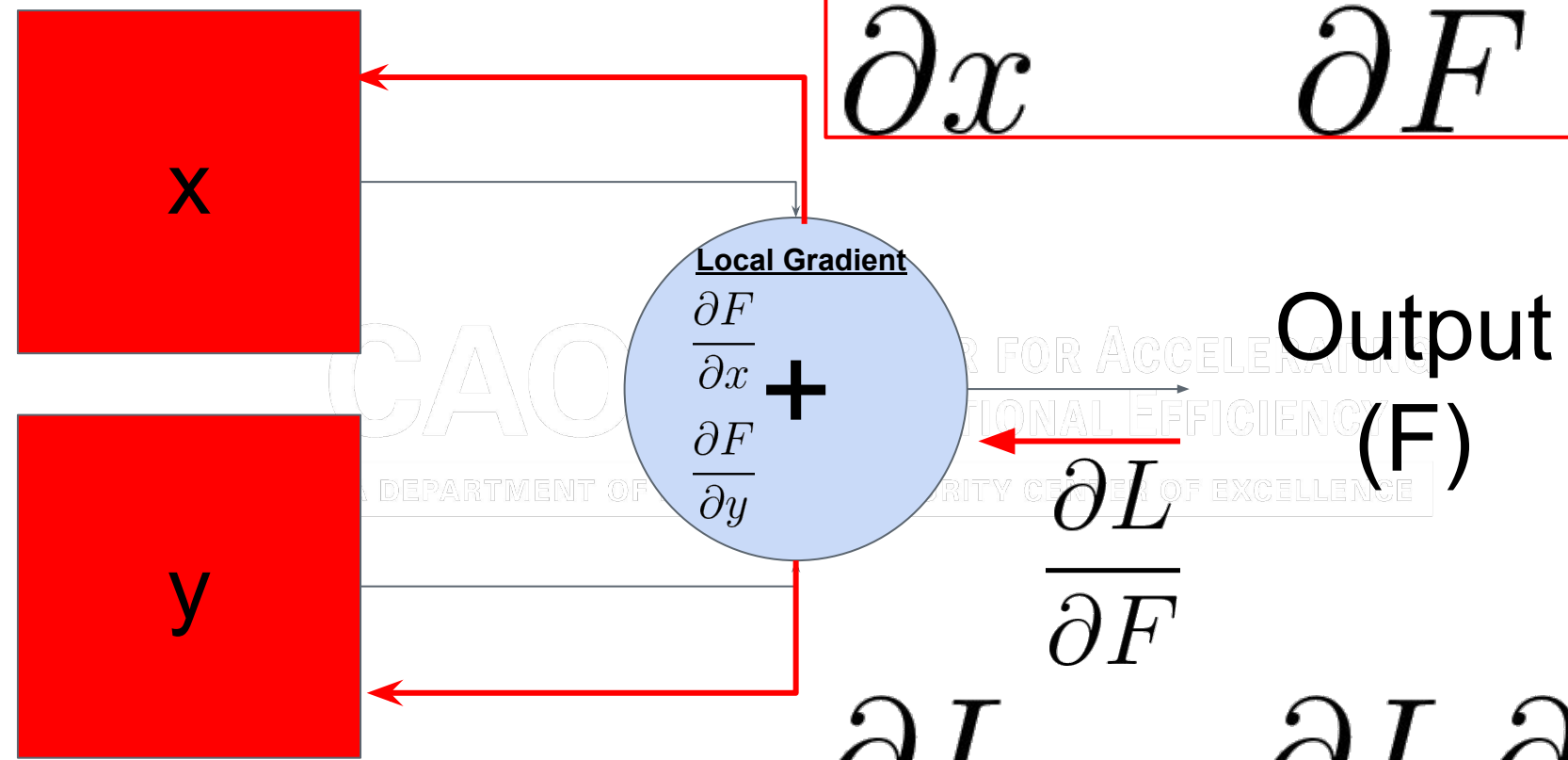




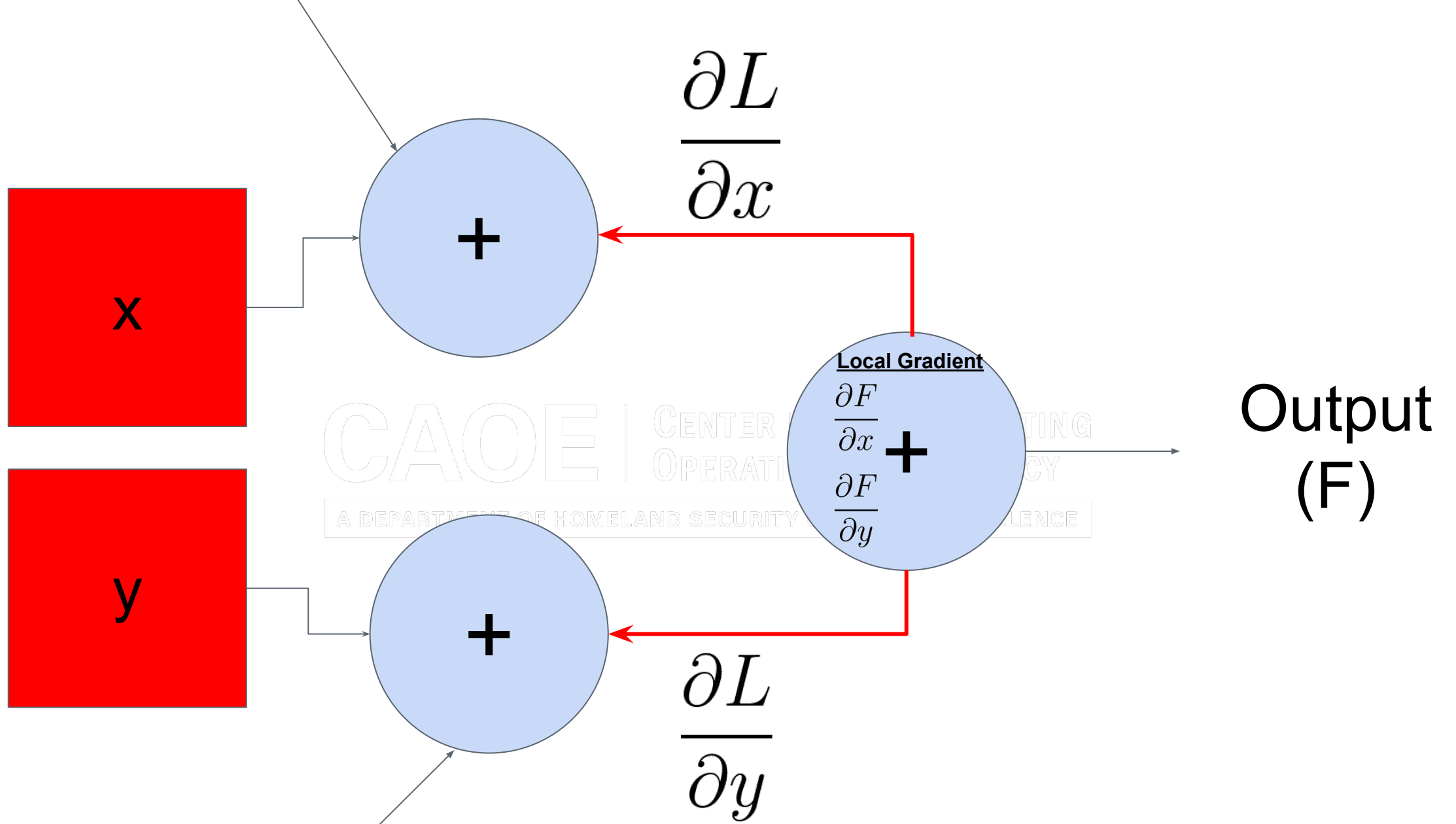
$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial F} \frac{\partial F}{\partial x}$$



$$\frac{\partial L}{\partial x} = \frac{\partial L}{\partial F} \frac{\partial F}{\partial x}$$

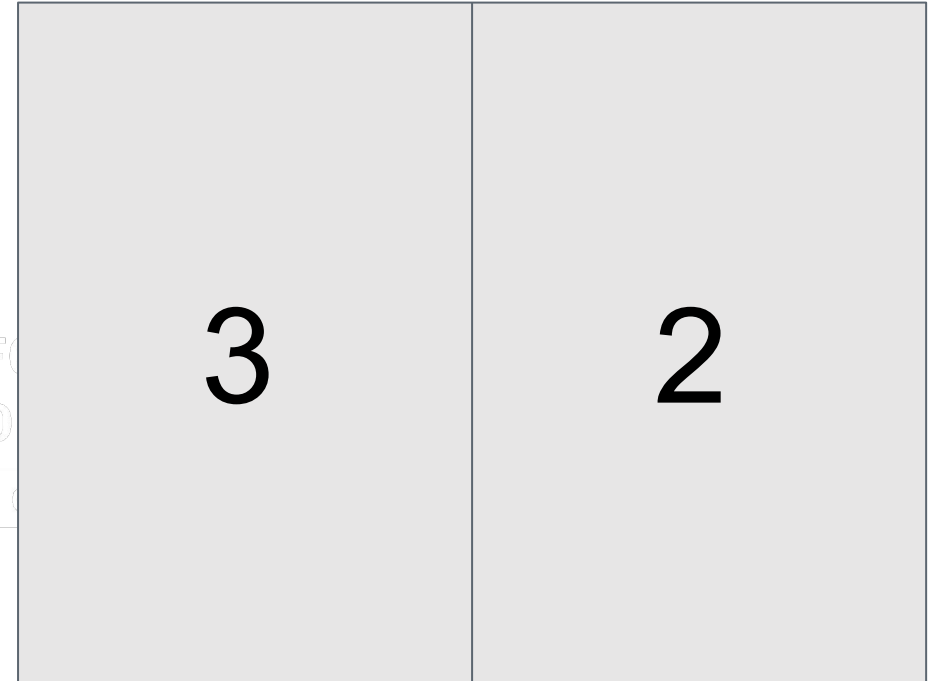


$$\frac{\partial L}{\partial y} = \frac{\partial L}{\partial F} \frac{\partial F}{\partial y}$$





CENTER FOR
OPERATIONS
AND SECURITY





Center for Accelerating Operational Efficiency

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

W1_1 (Bird)
-2

W1_2 (Bird)
-1

W2_1 (Car)
1

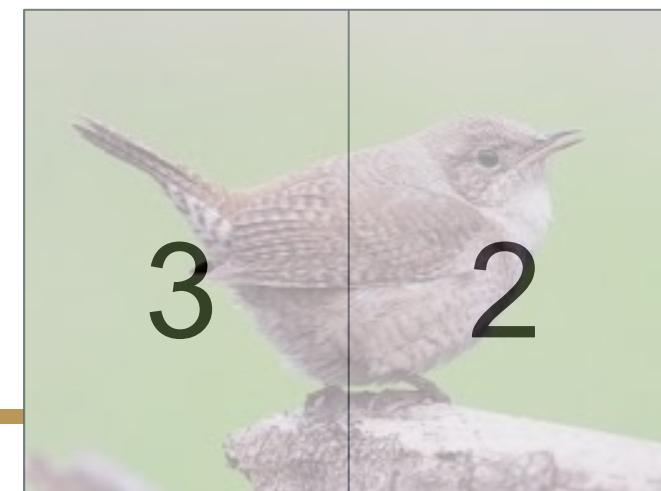
W2_2 (Car)
-5

$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



W1_1 (Bird)
-2

W1_2 (Bird)
-1

Pixel_1
3

Pixel_2
2

W2_1 (Car)
1

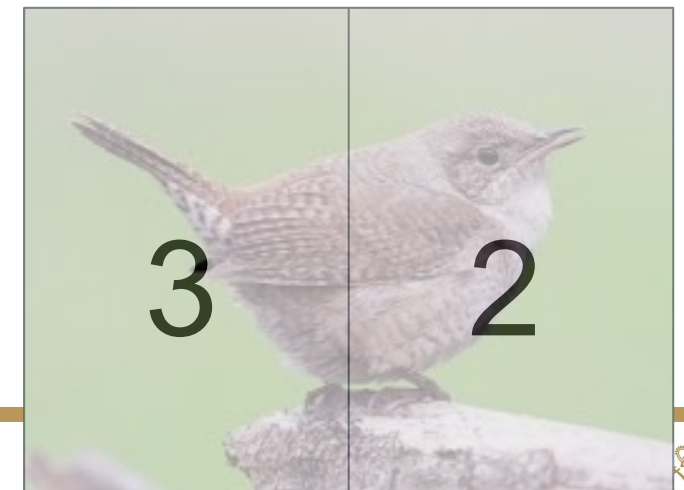
W2_2 (Car)
-5

$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



W1_1 (Bird)
-2

Pixel_1
3

W2_1 (Car)
1

W2_2 (Car)
-5

Pixel_2
2

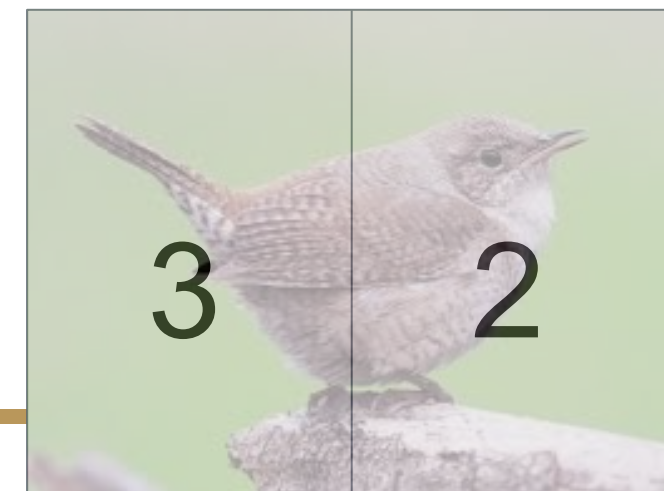
W1_2 (Bird)
-1

$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

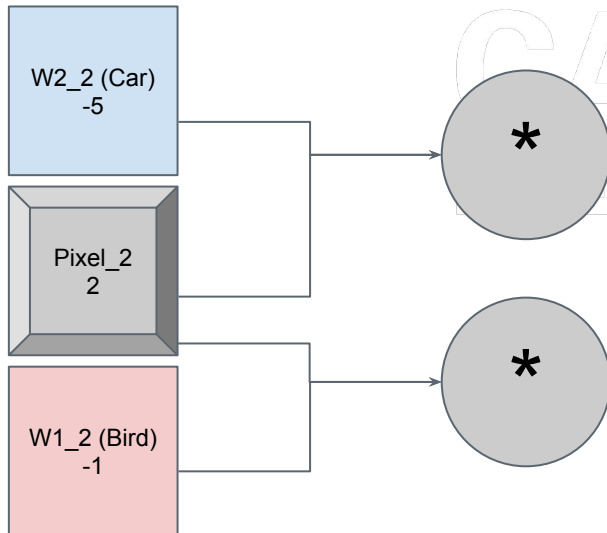
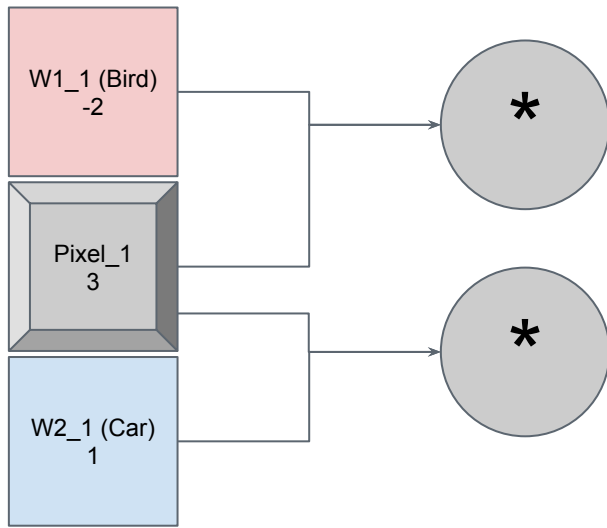
CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

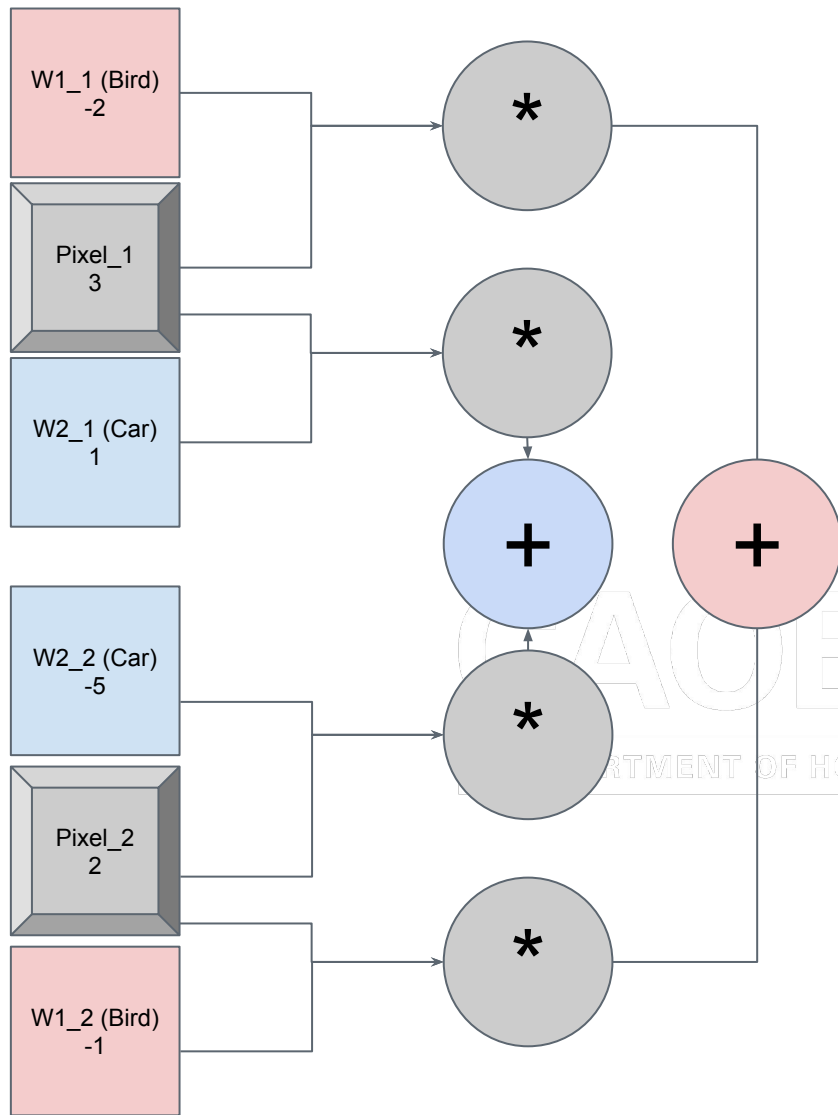


CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



$$f(X, W)$$

$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

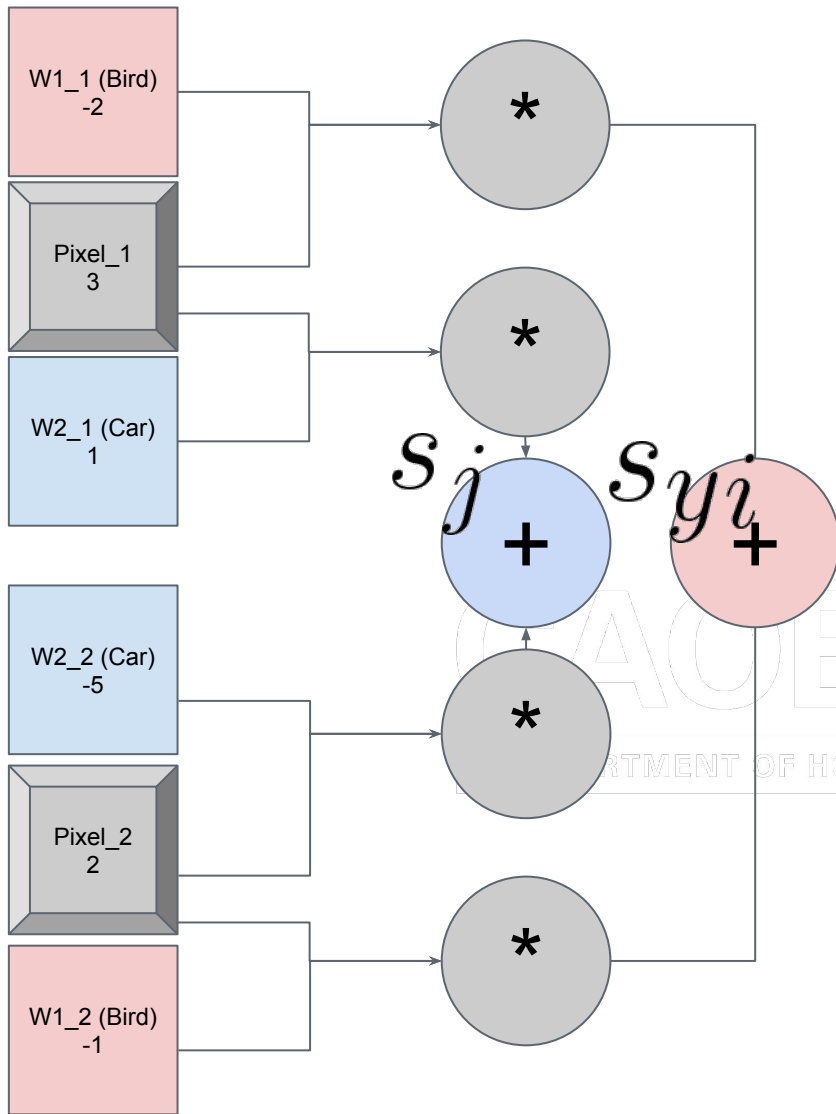


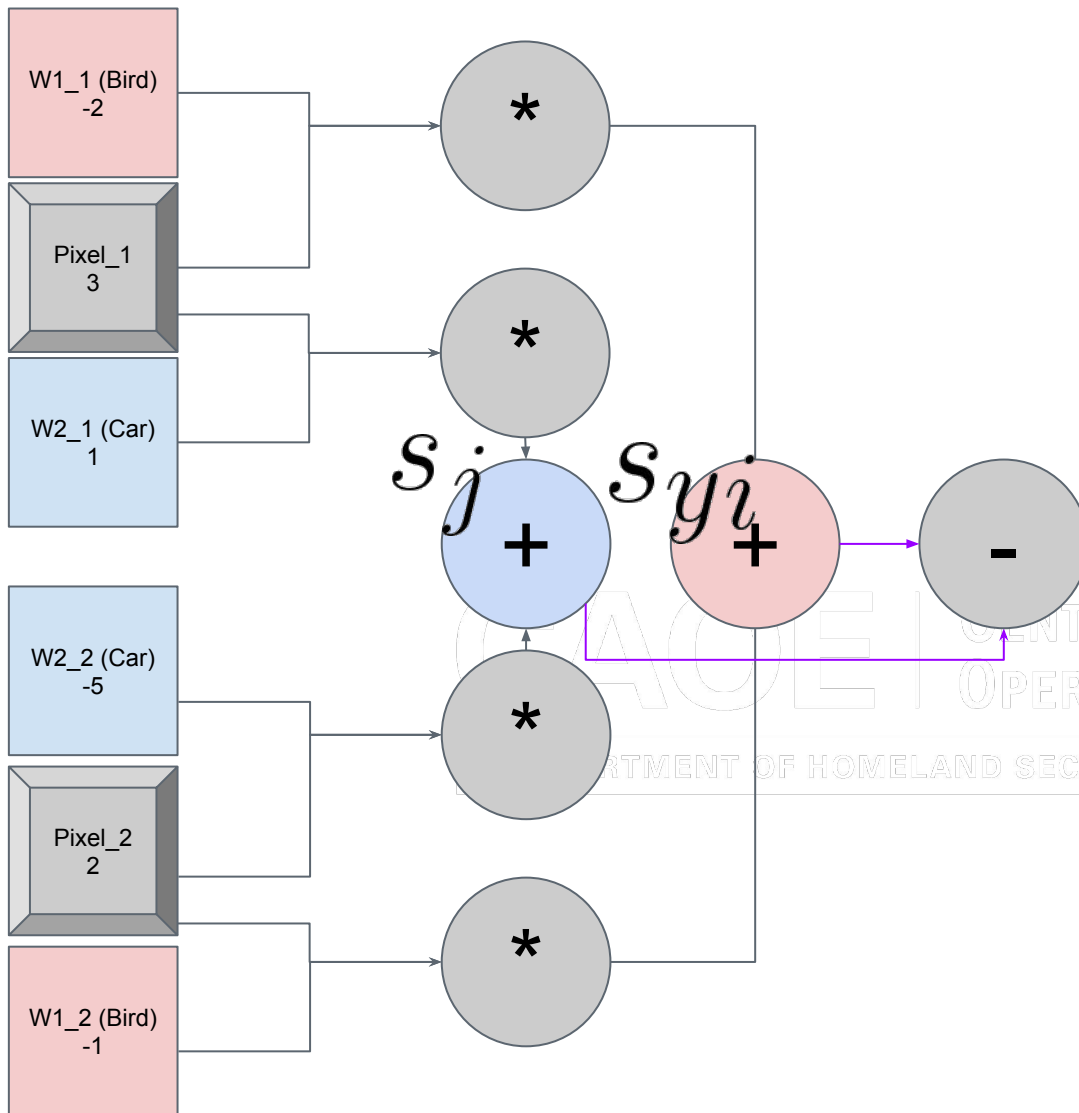
CACOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



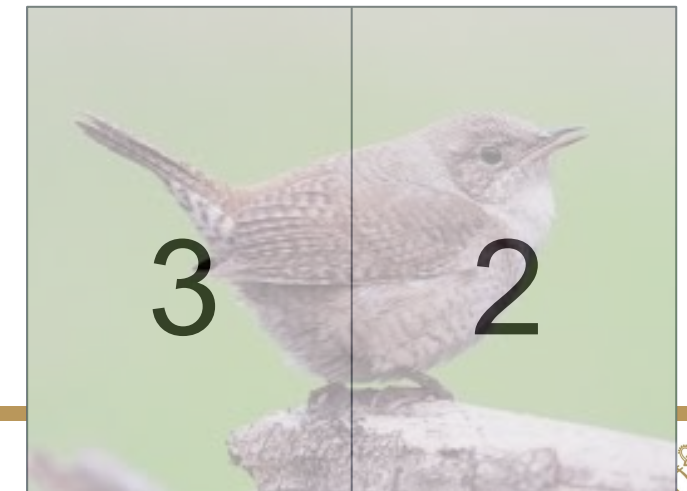
$$f(X, W)$$

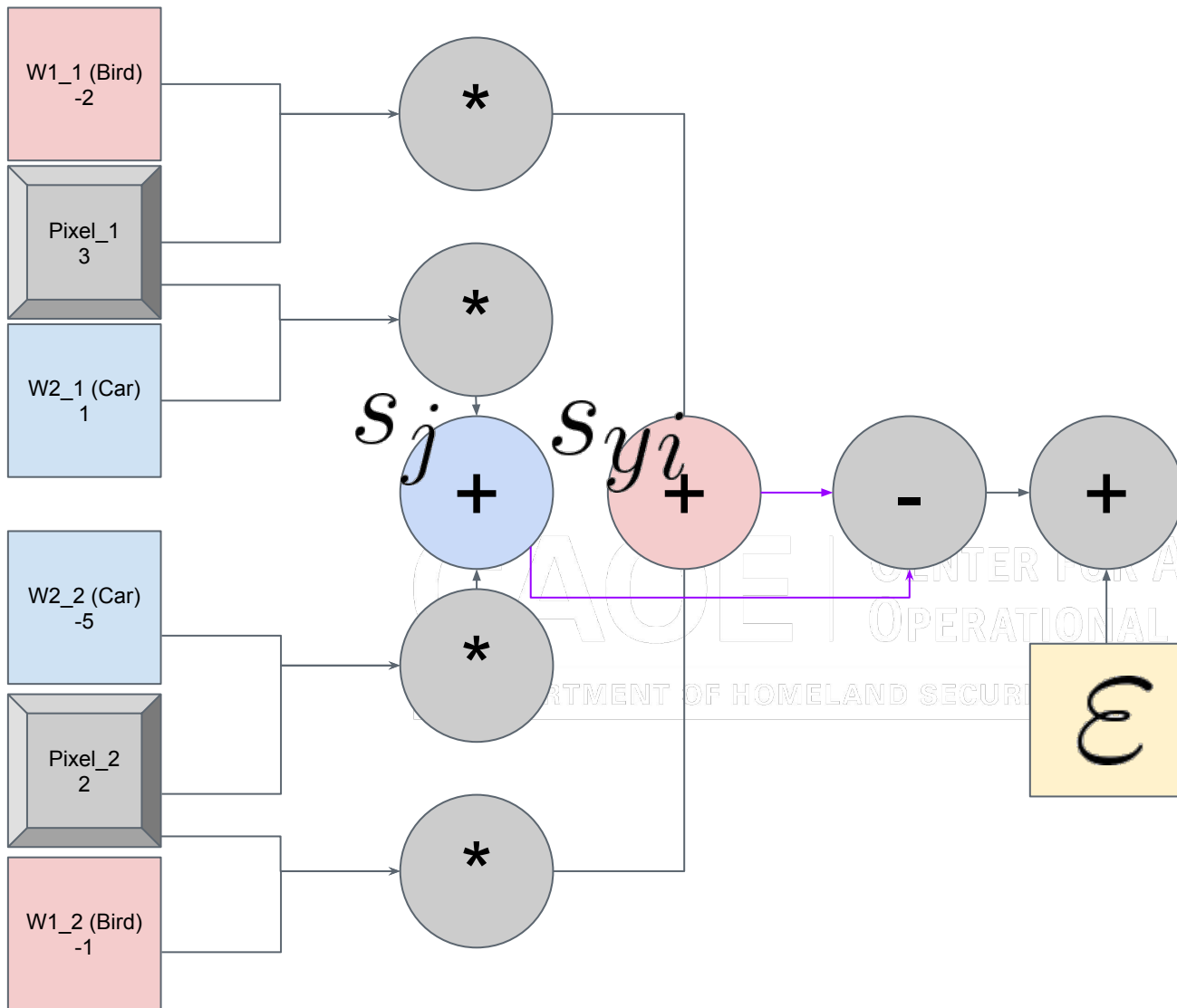
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



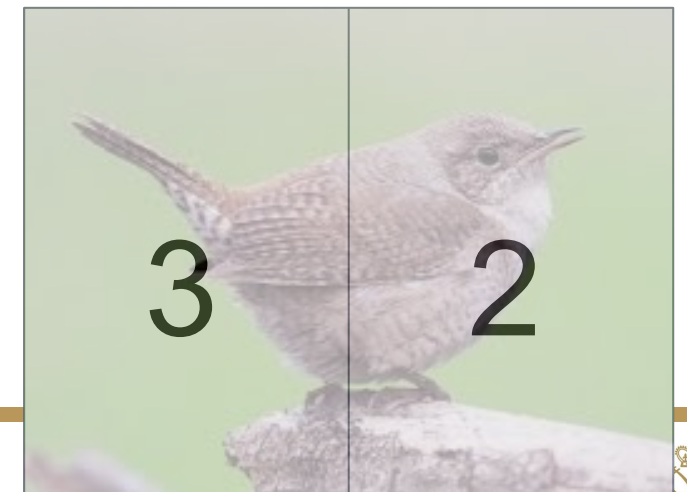


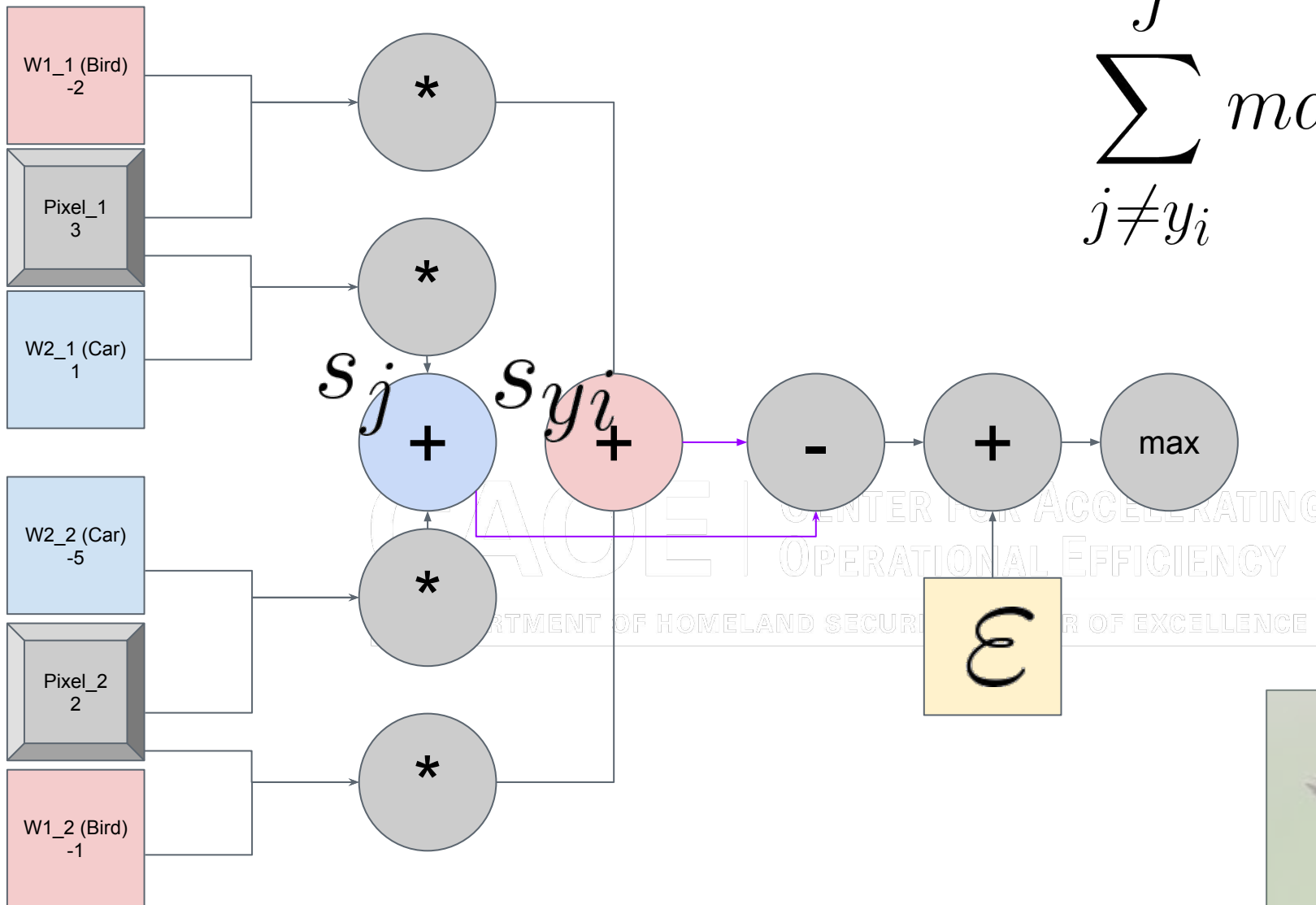
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



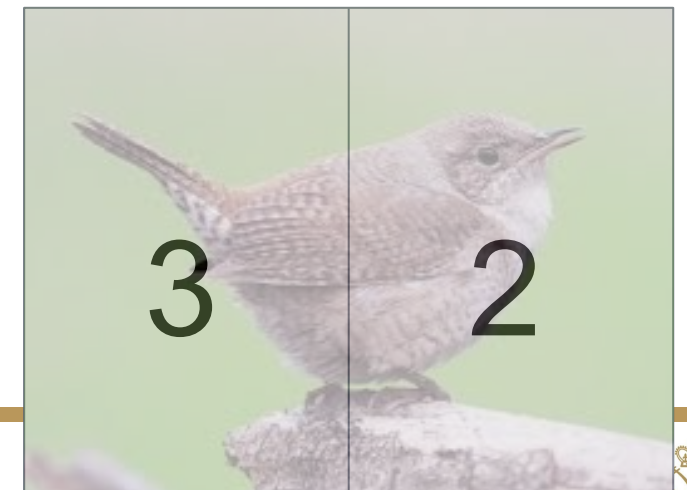


$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \epsilon)$$

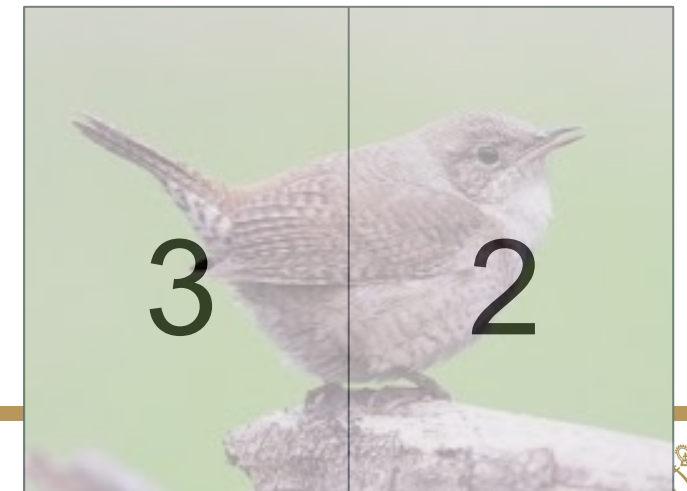
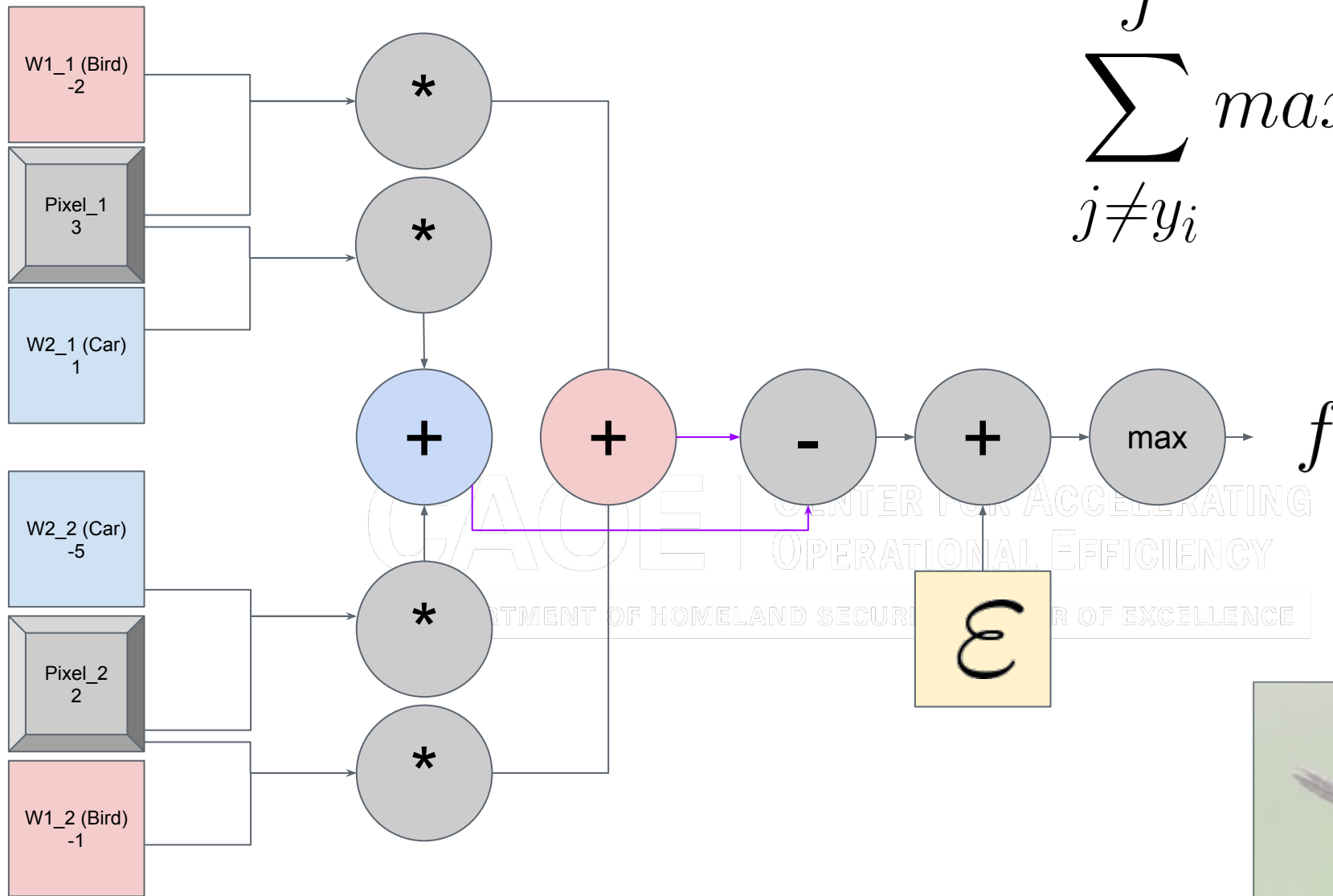


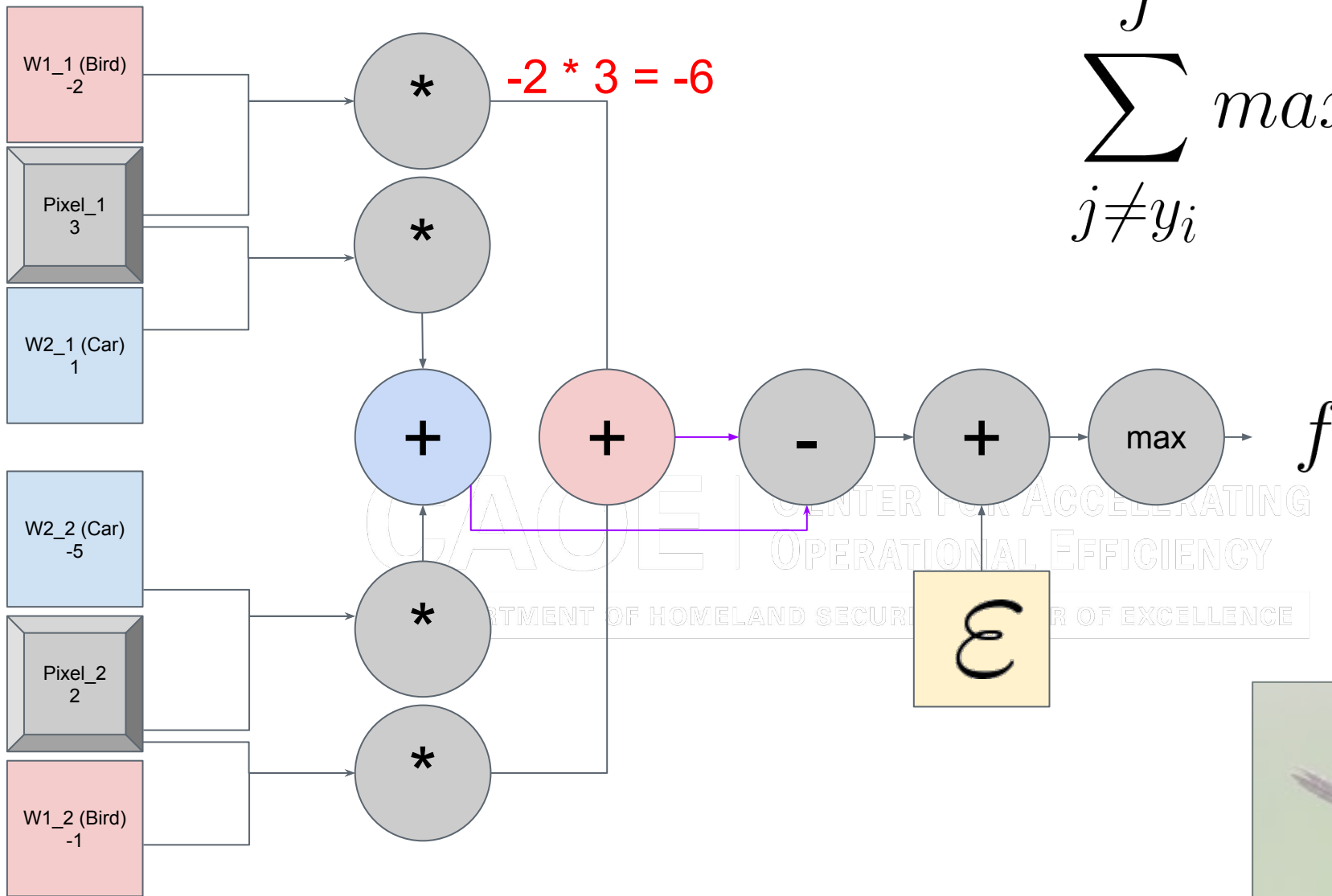


$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \epsilon)$$

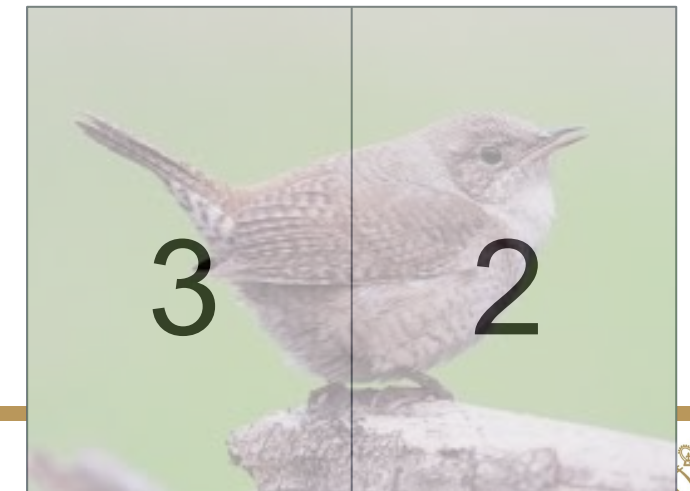


$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

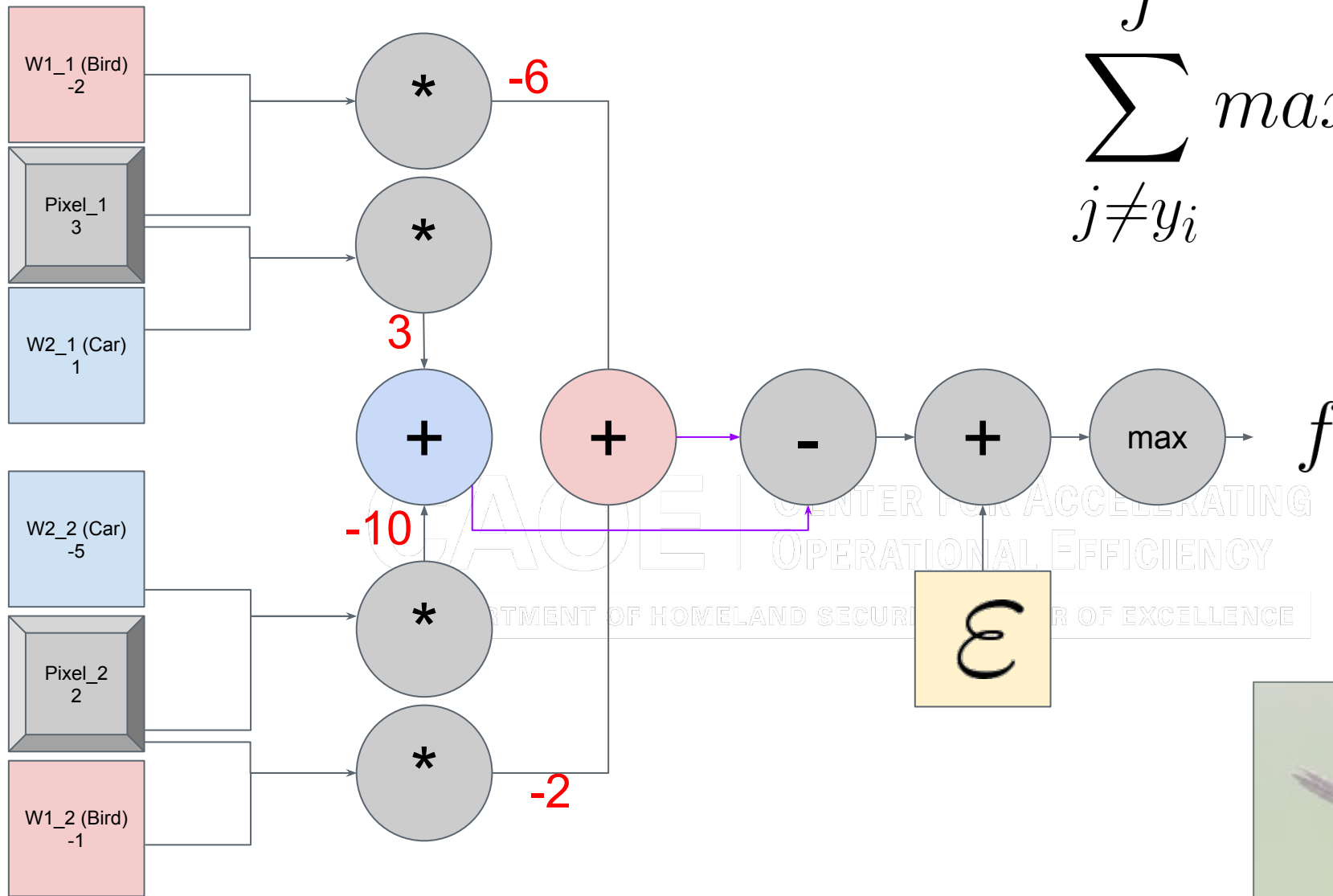




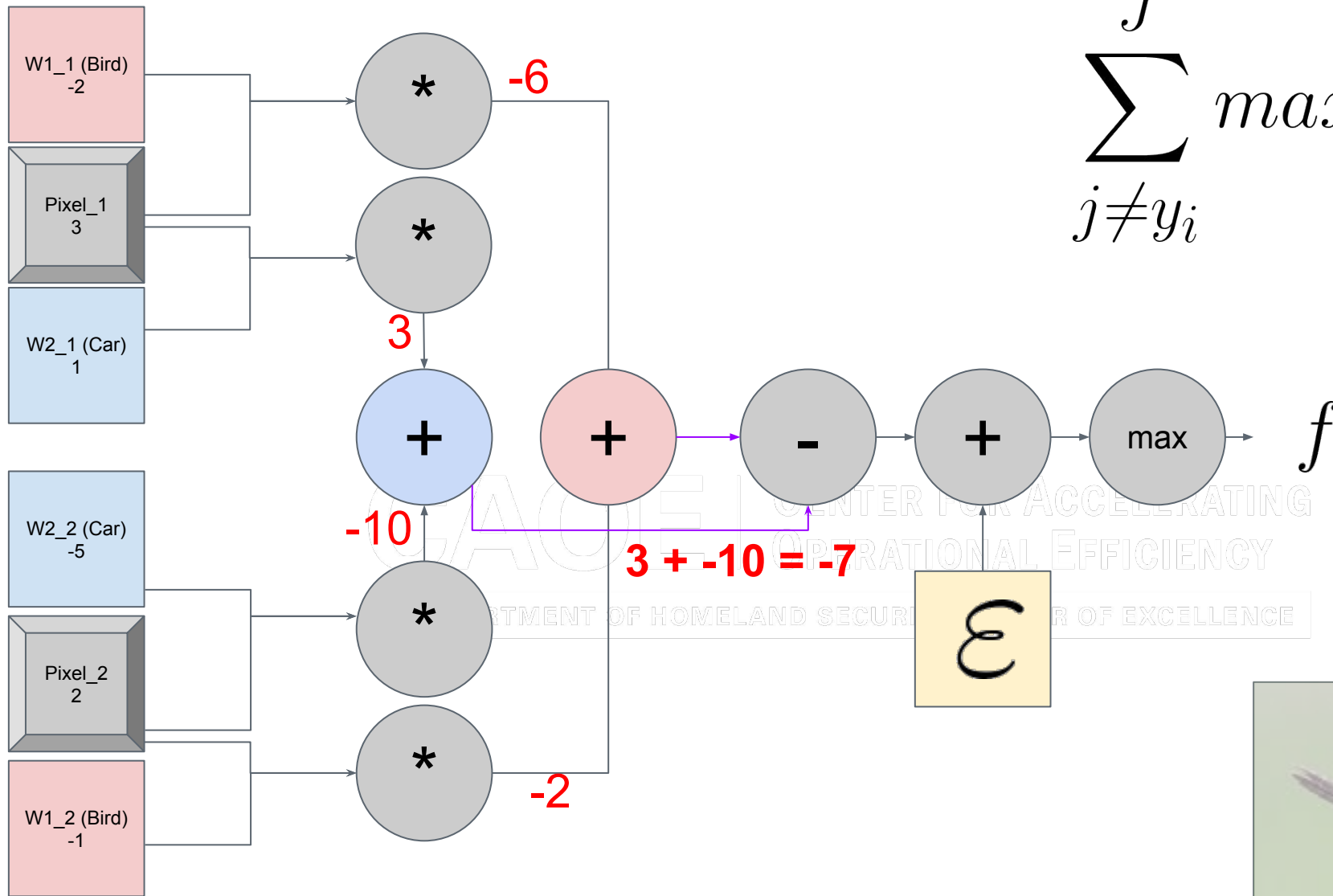
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \epsilon)$$



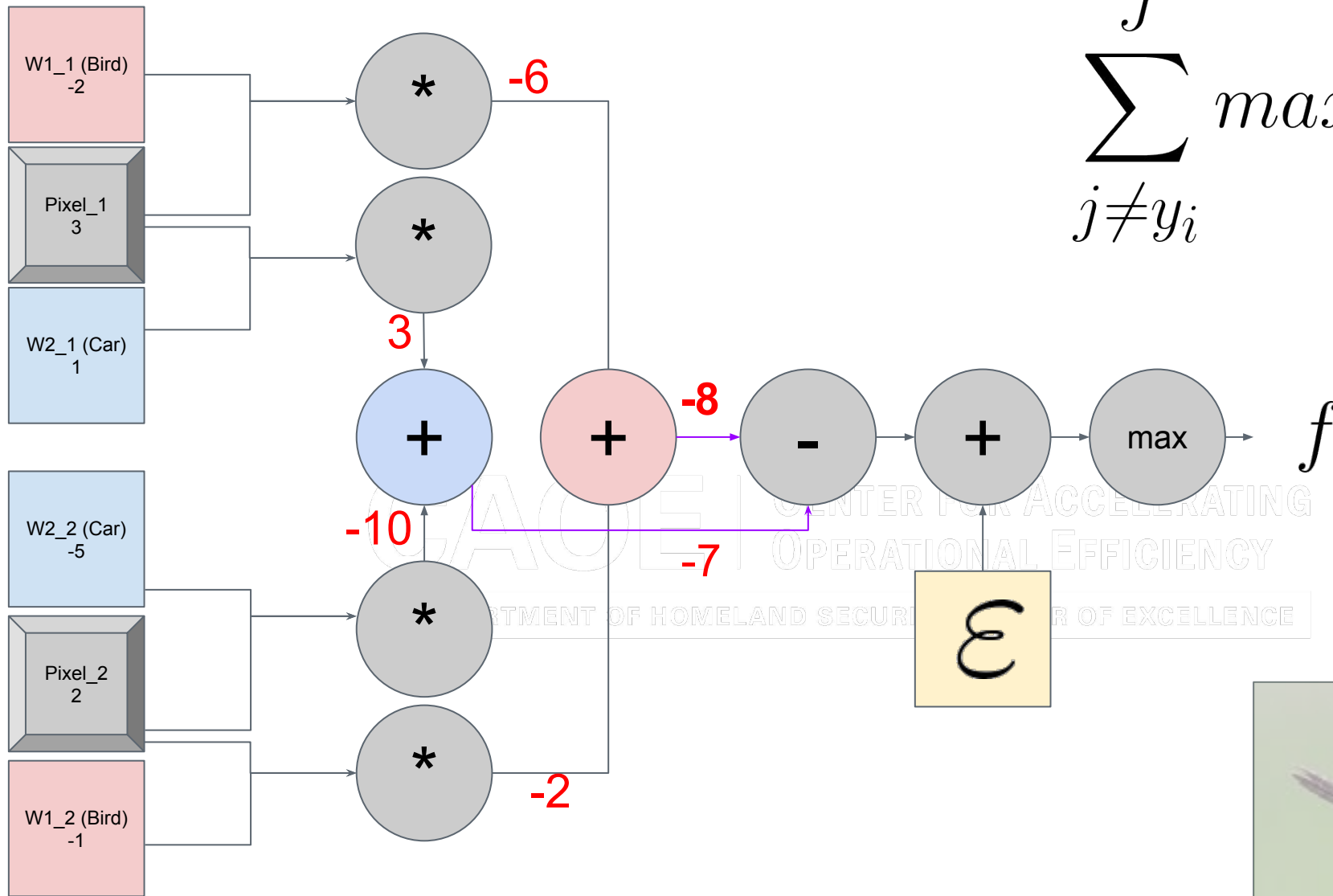
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



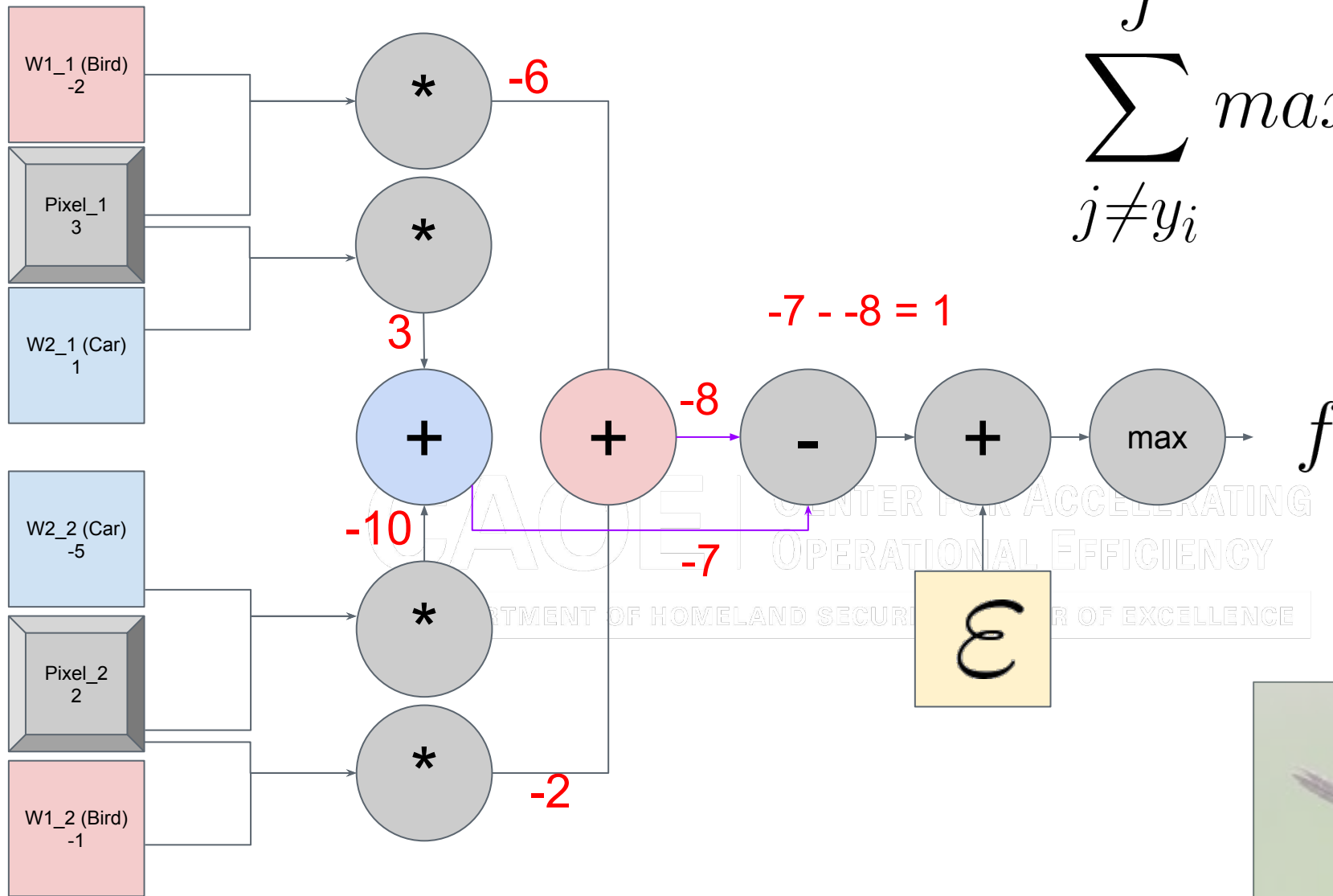
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



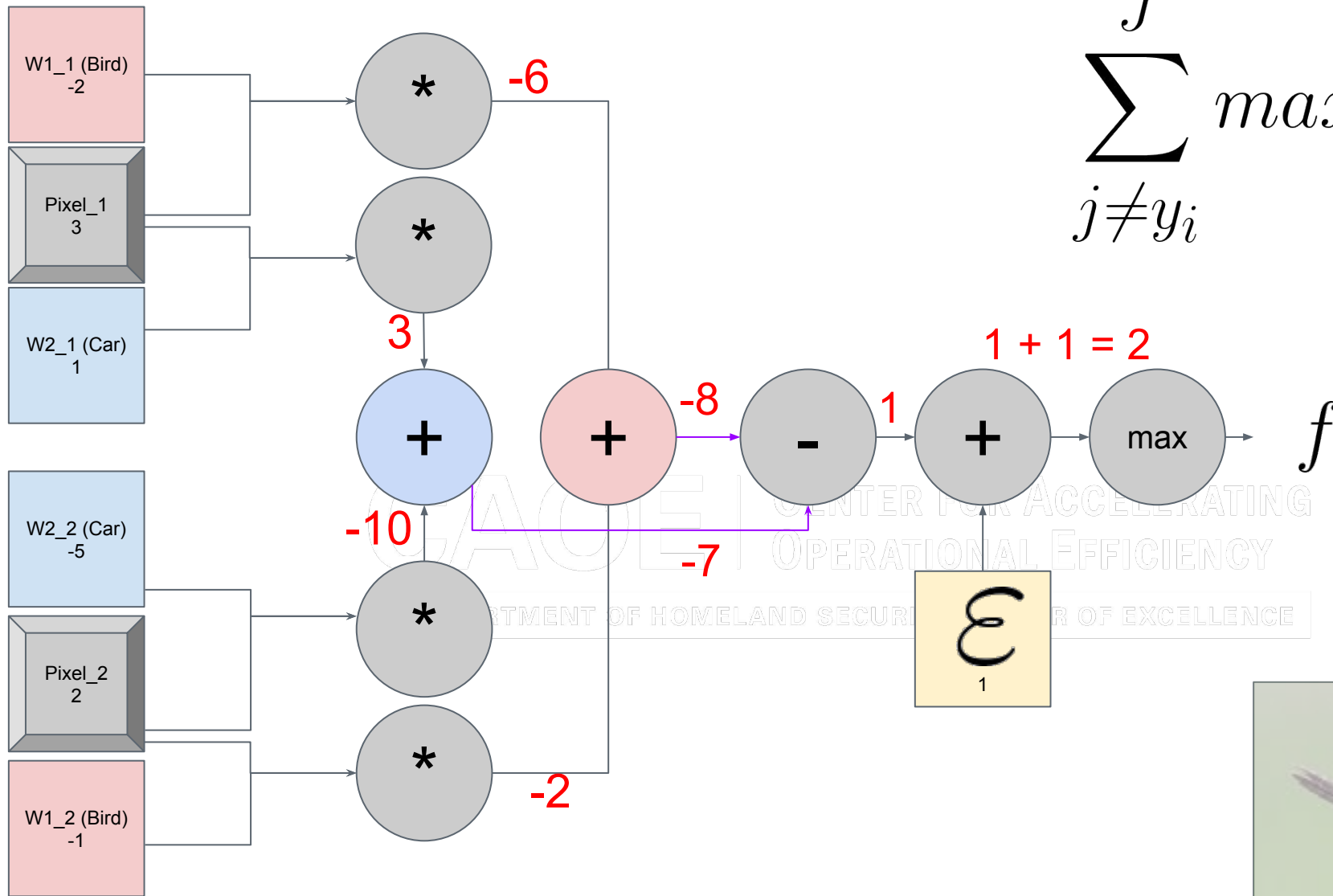
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



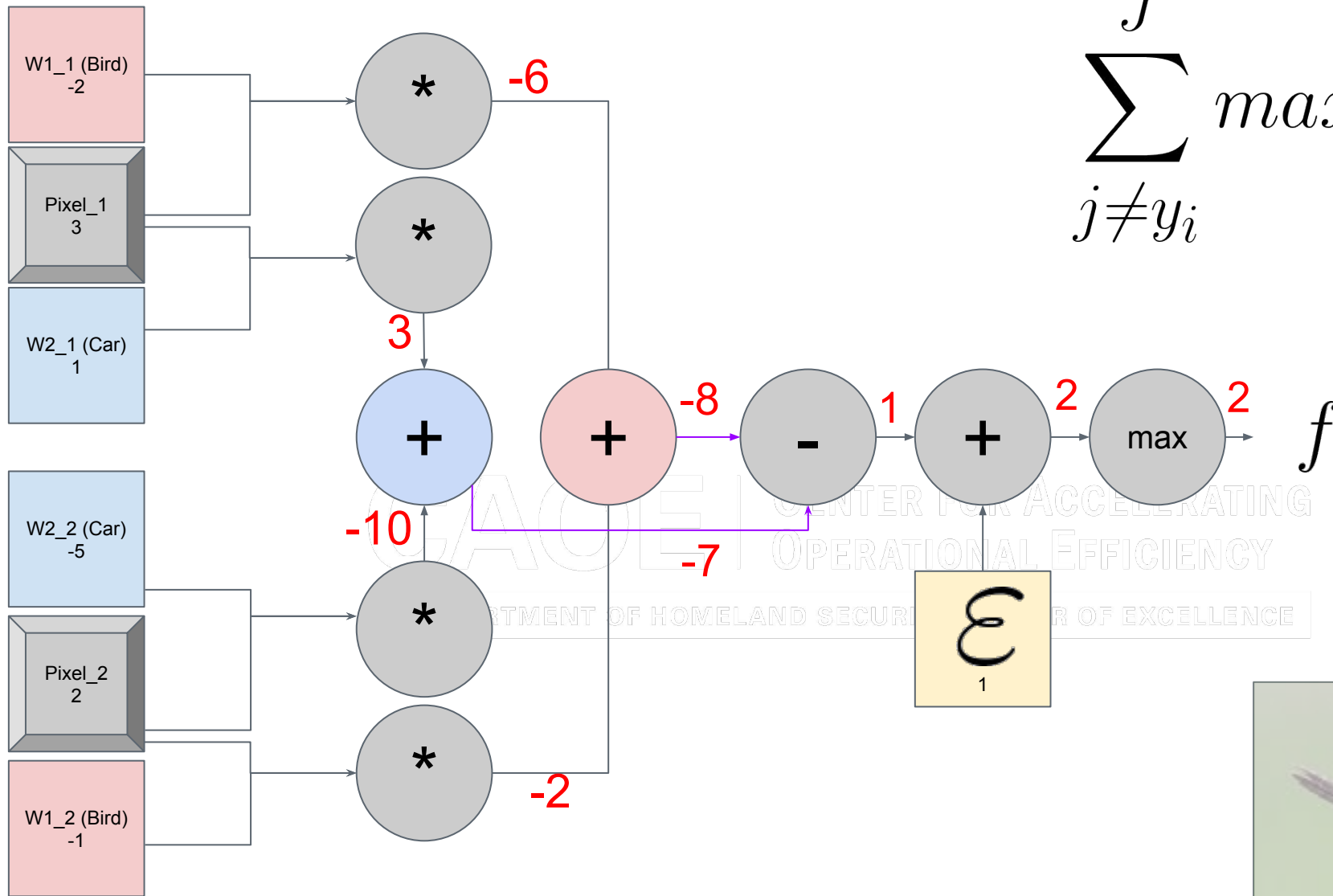
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

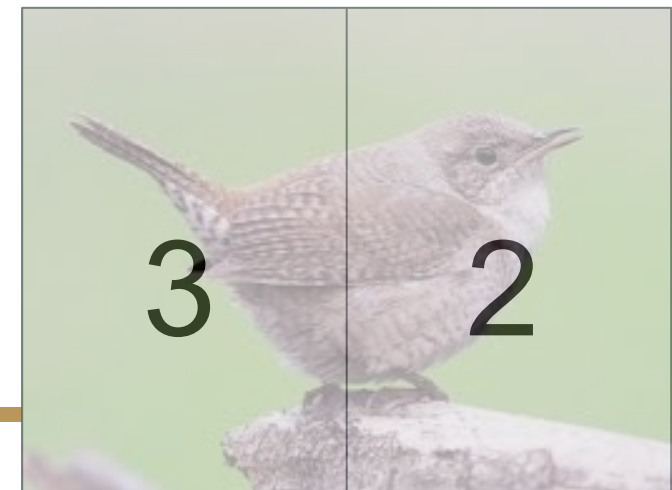


$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

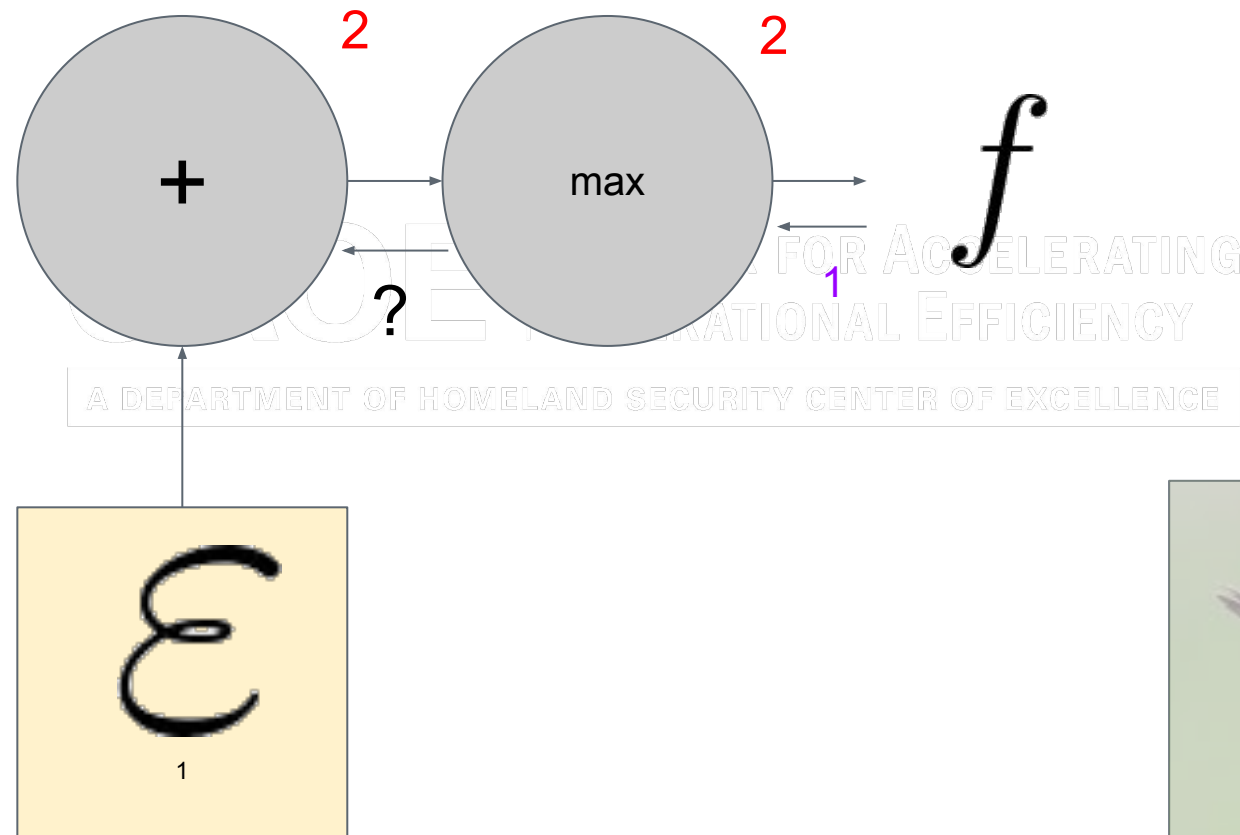


$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

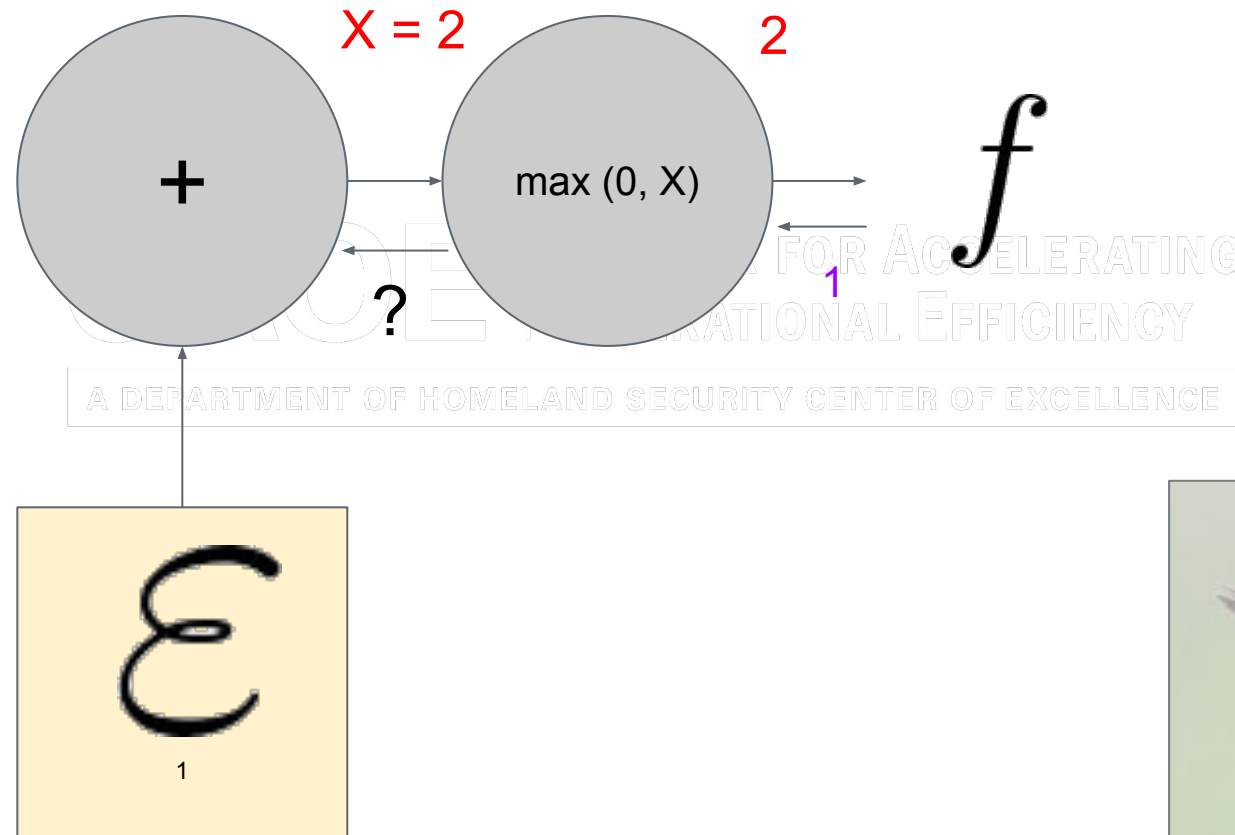




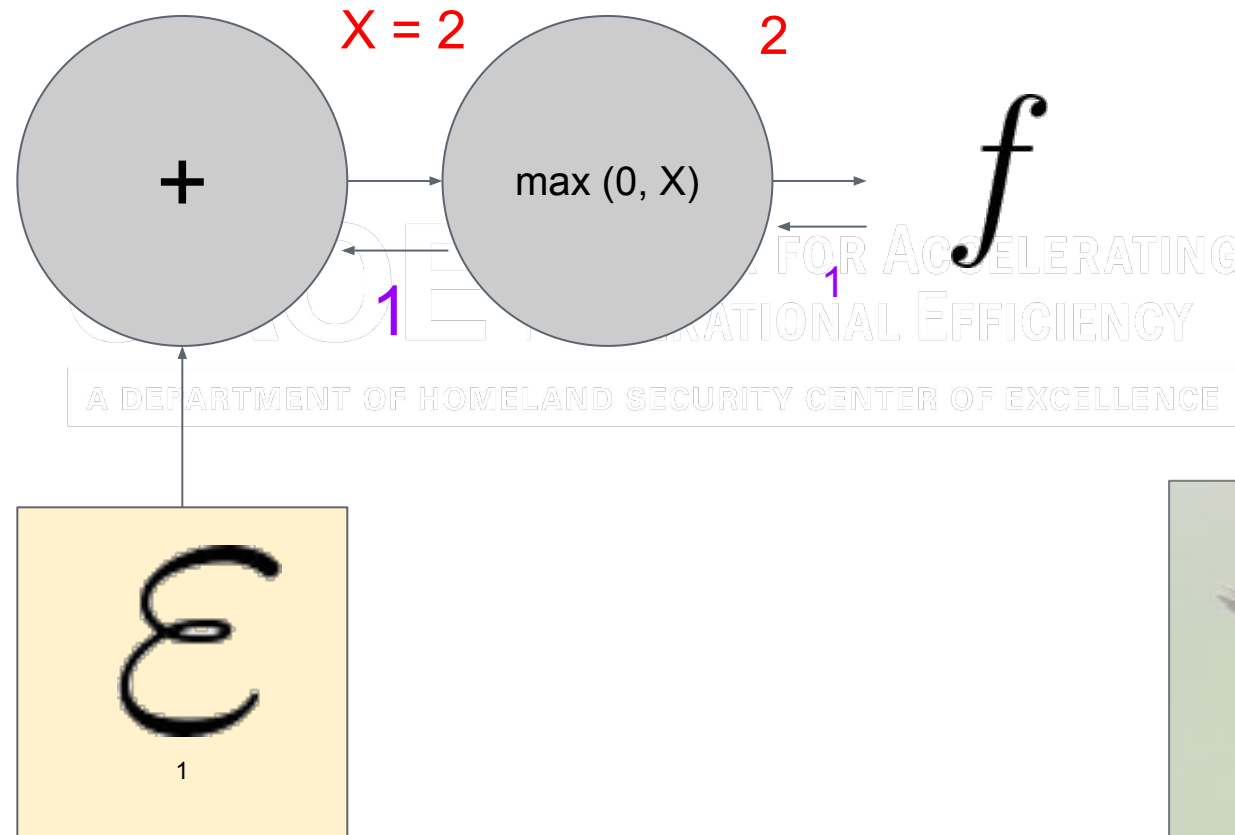
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



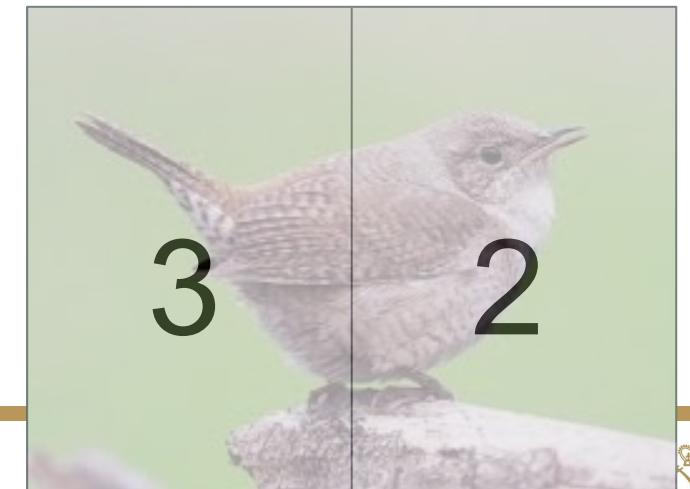
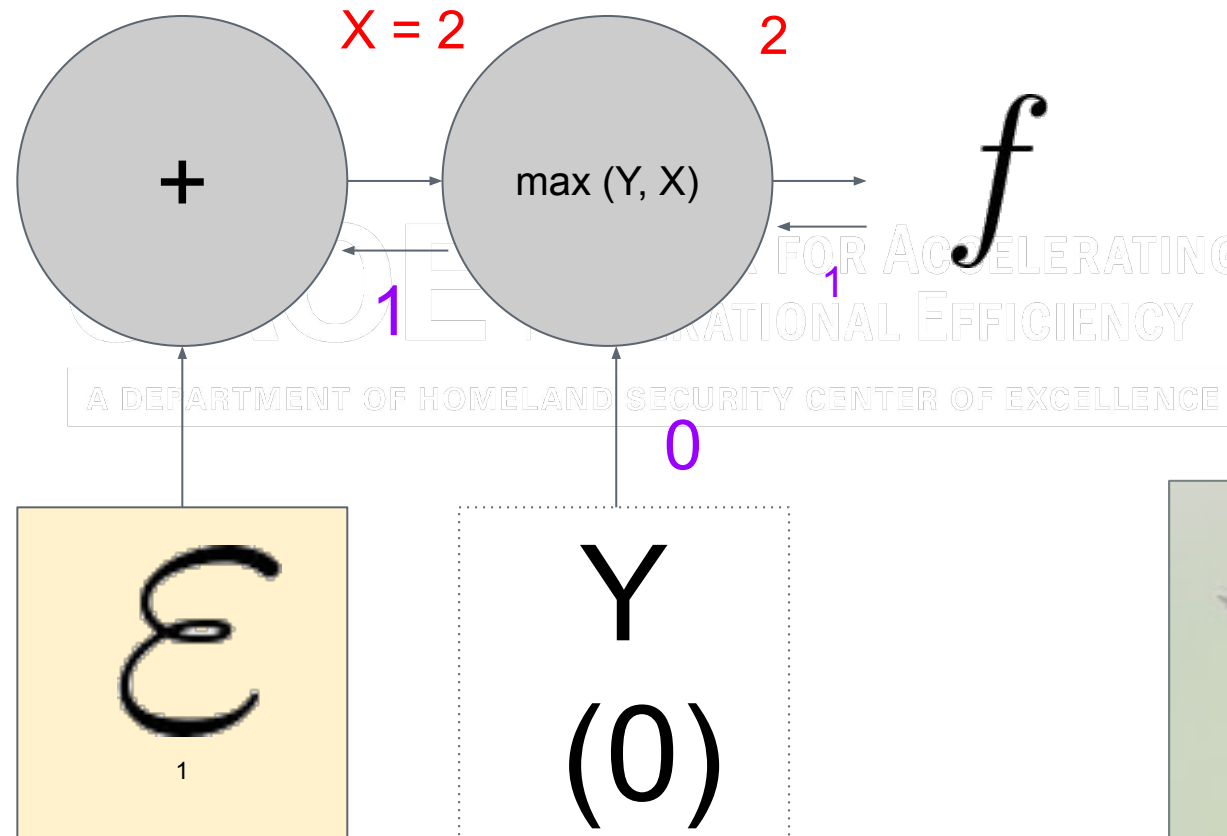
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



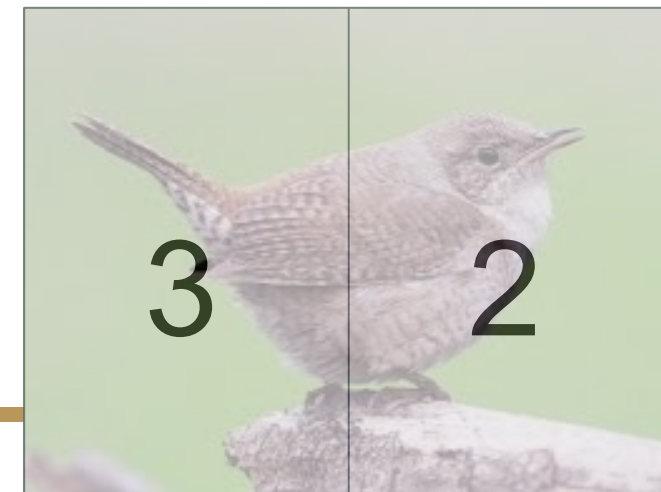
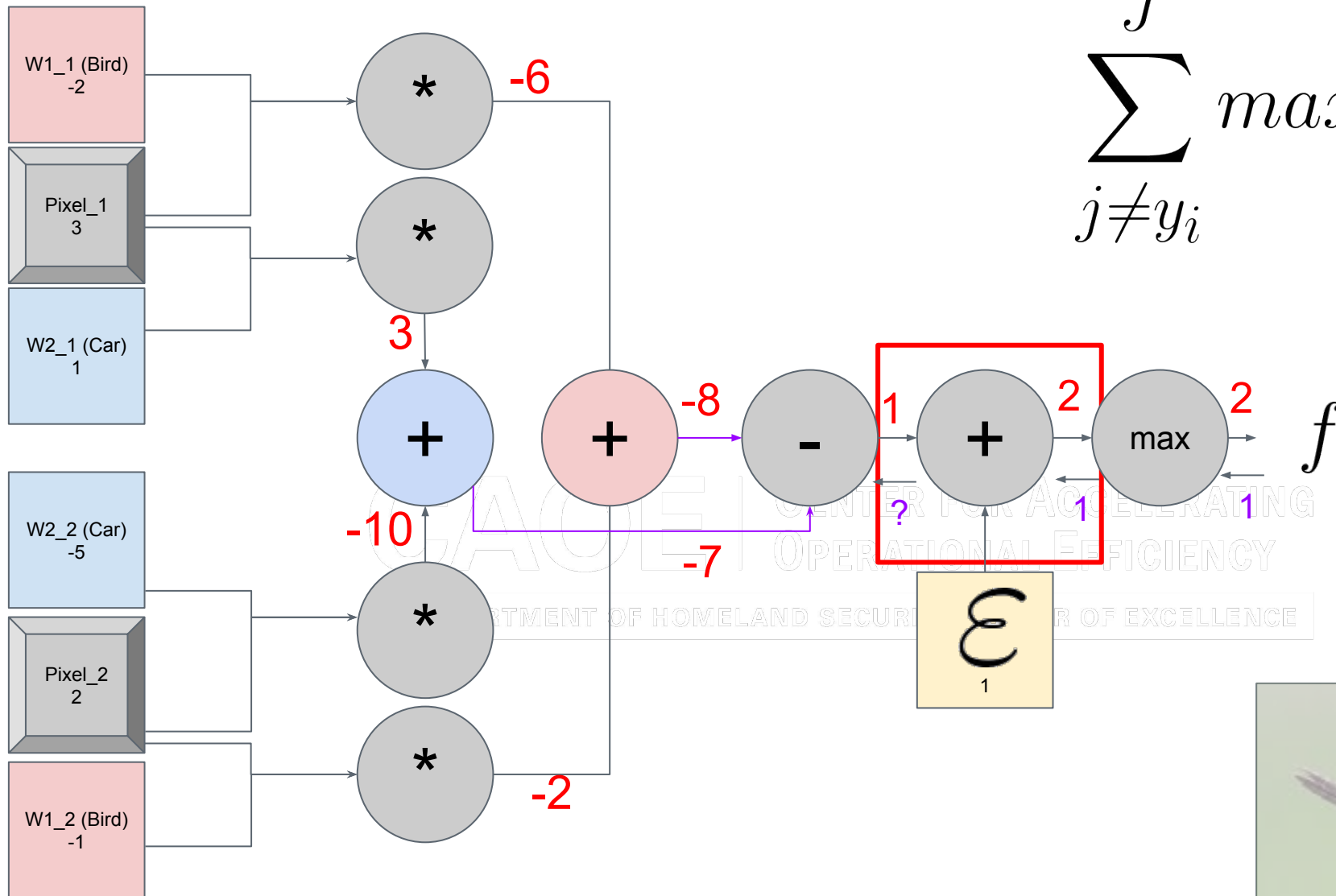
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$



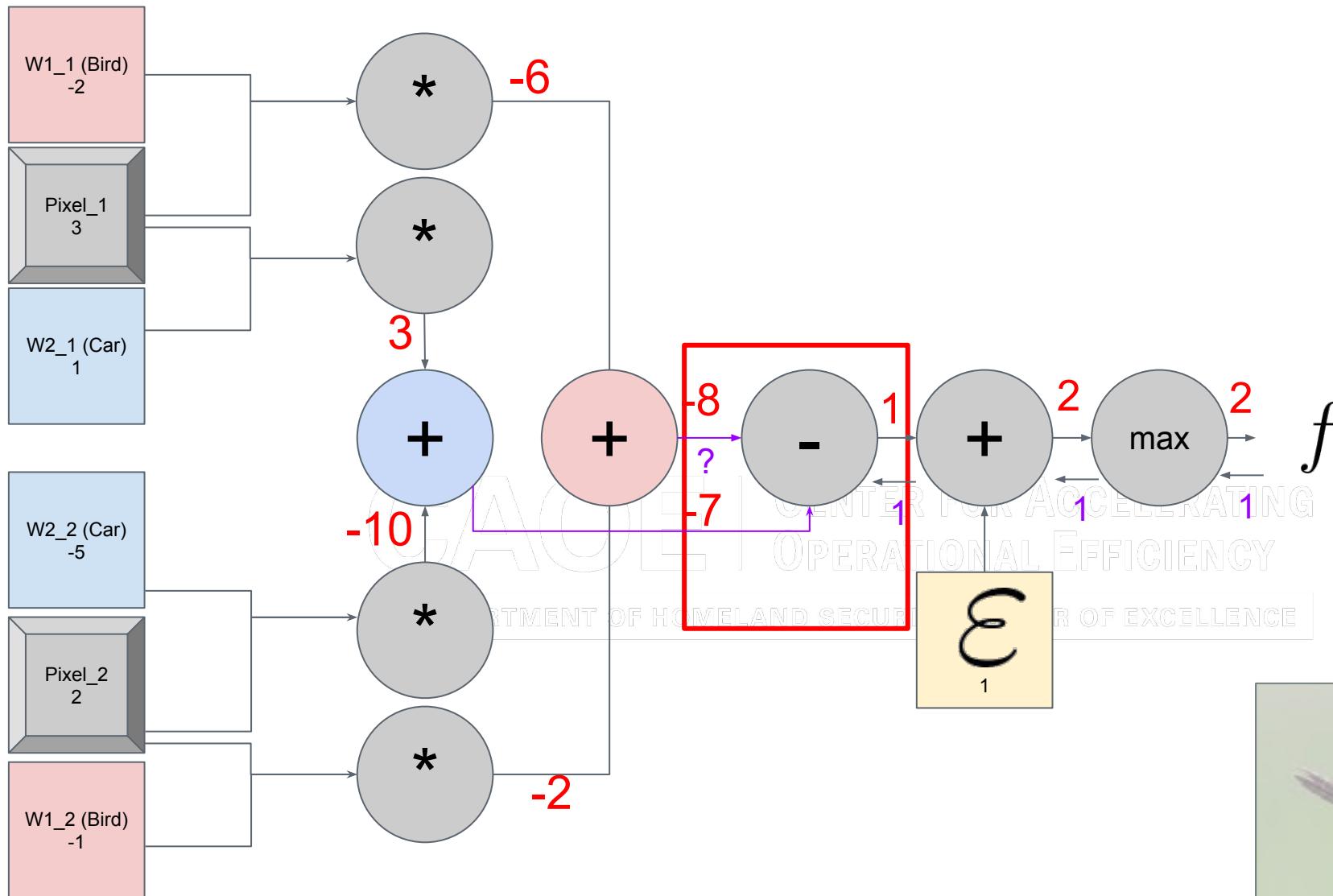
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

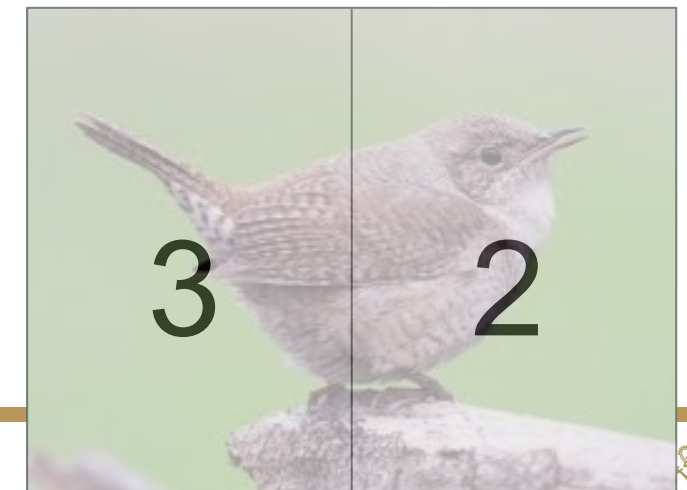
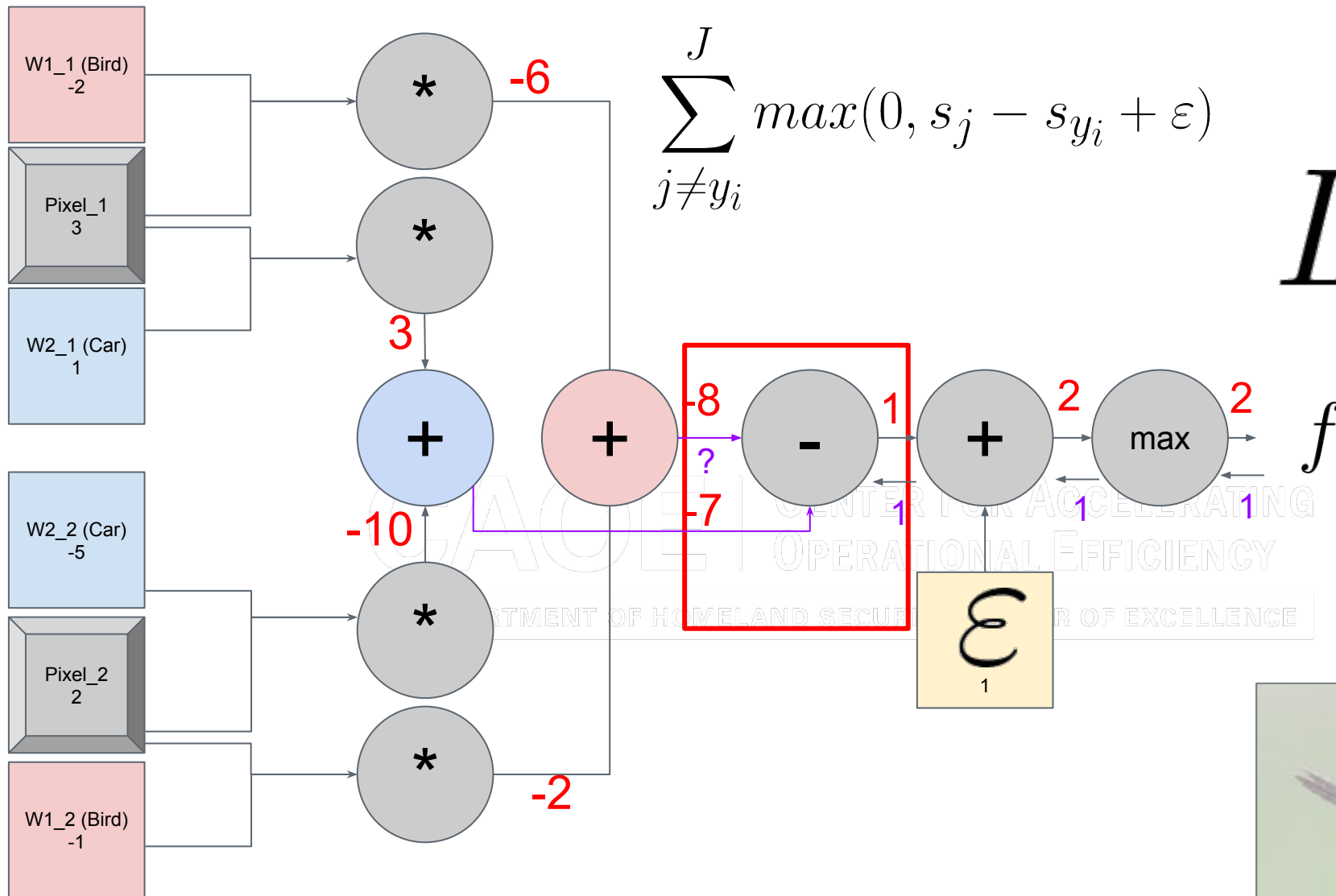


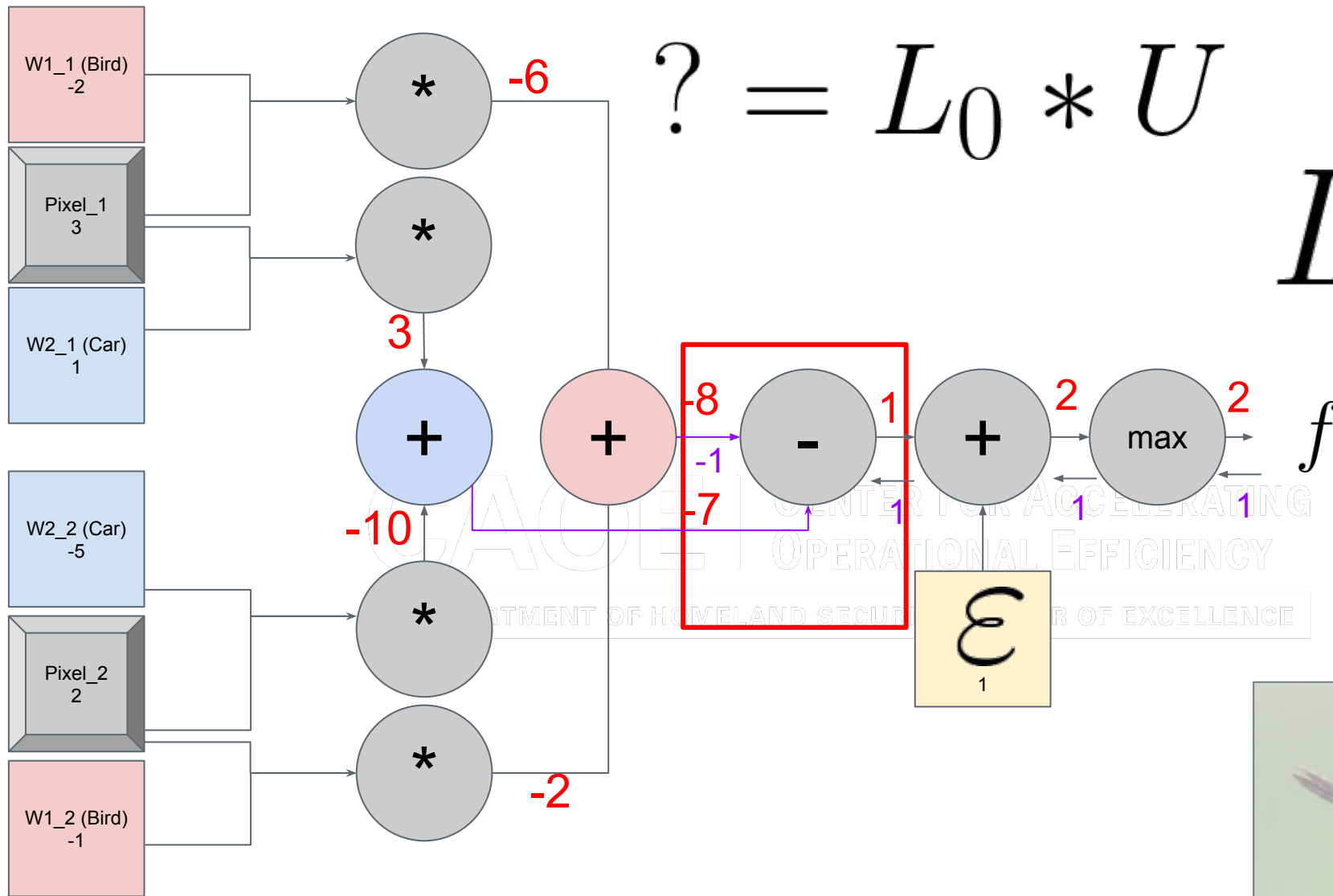
$$\sum_{j \neq y_i}^J \max(0, s_j - s_{y_i} + \varepsilon)$$

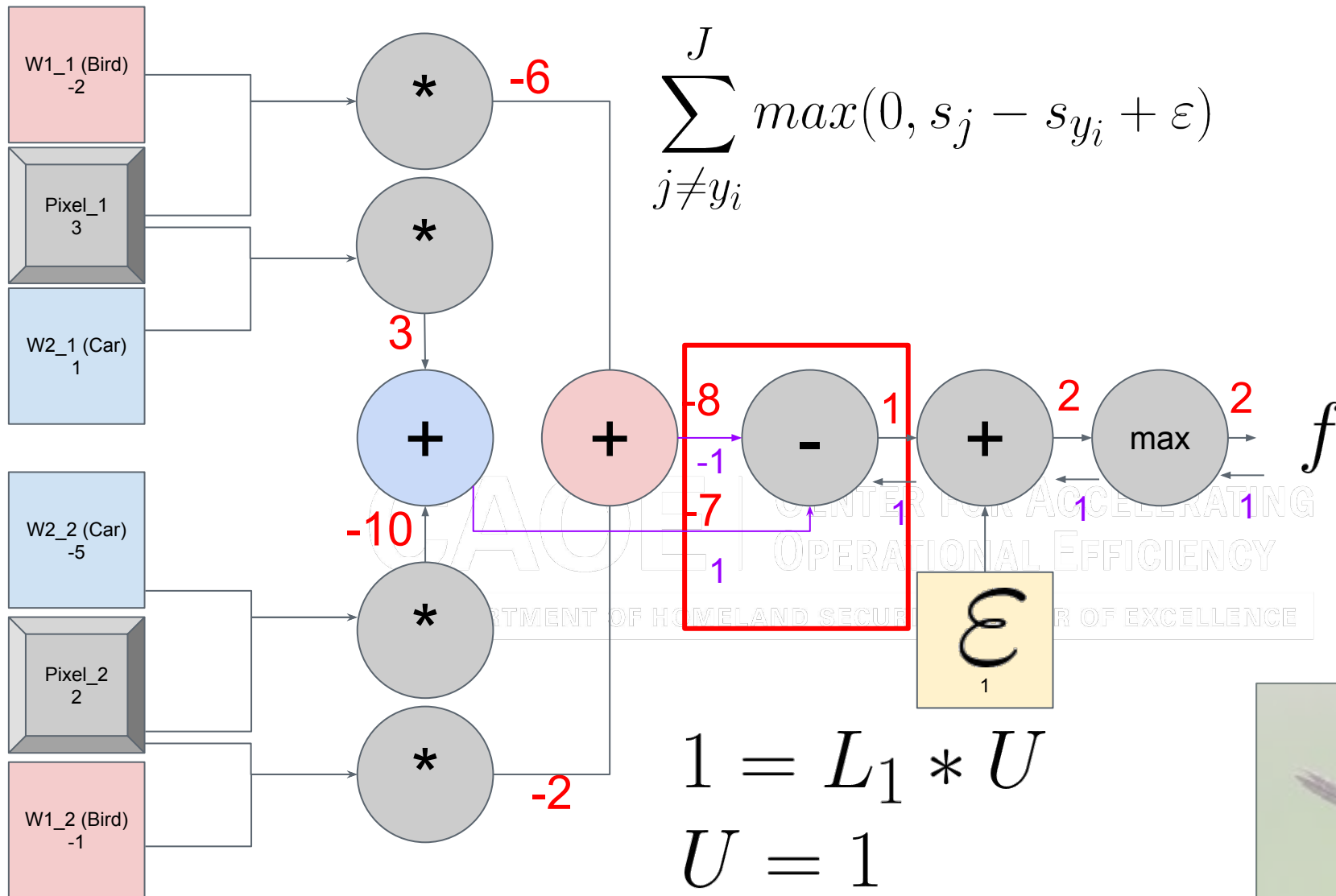


$$U = 1$$





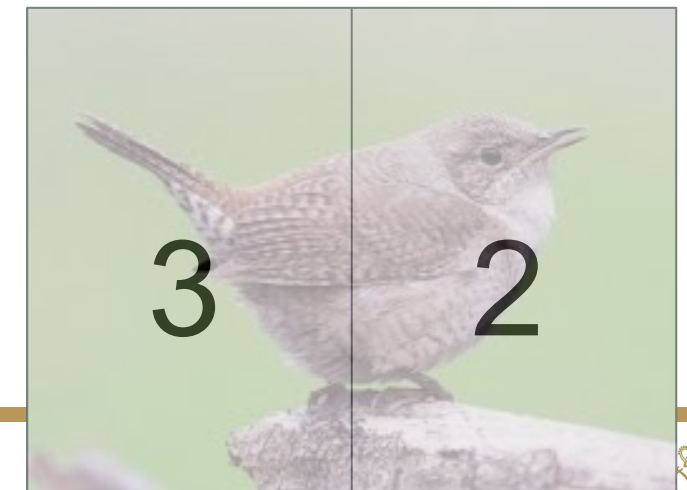


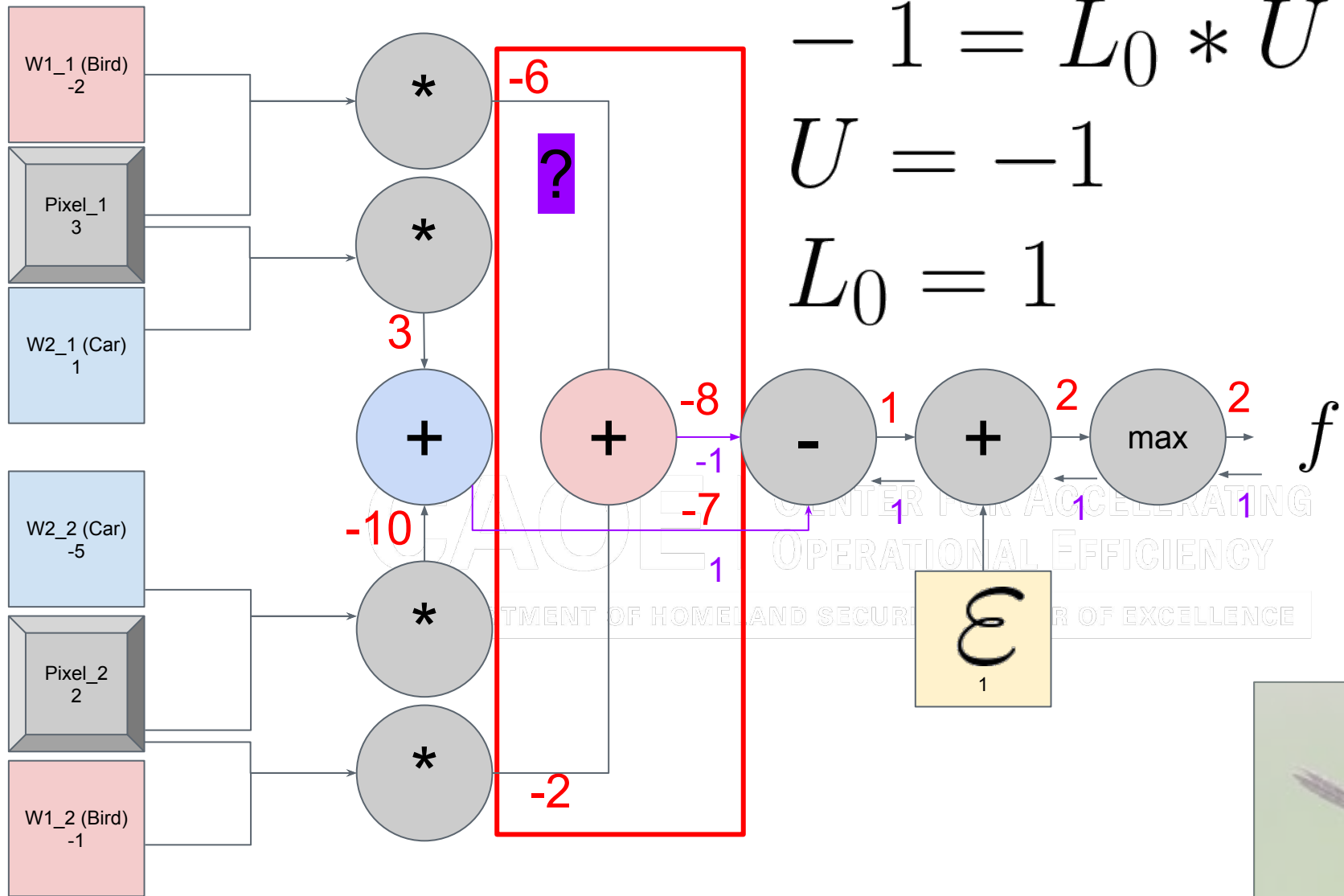


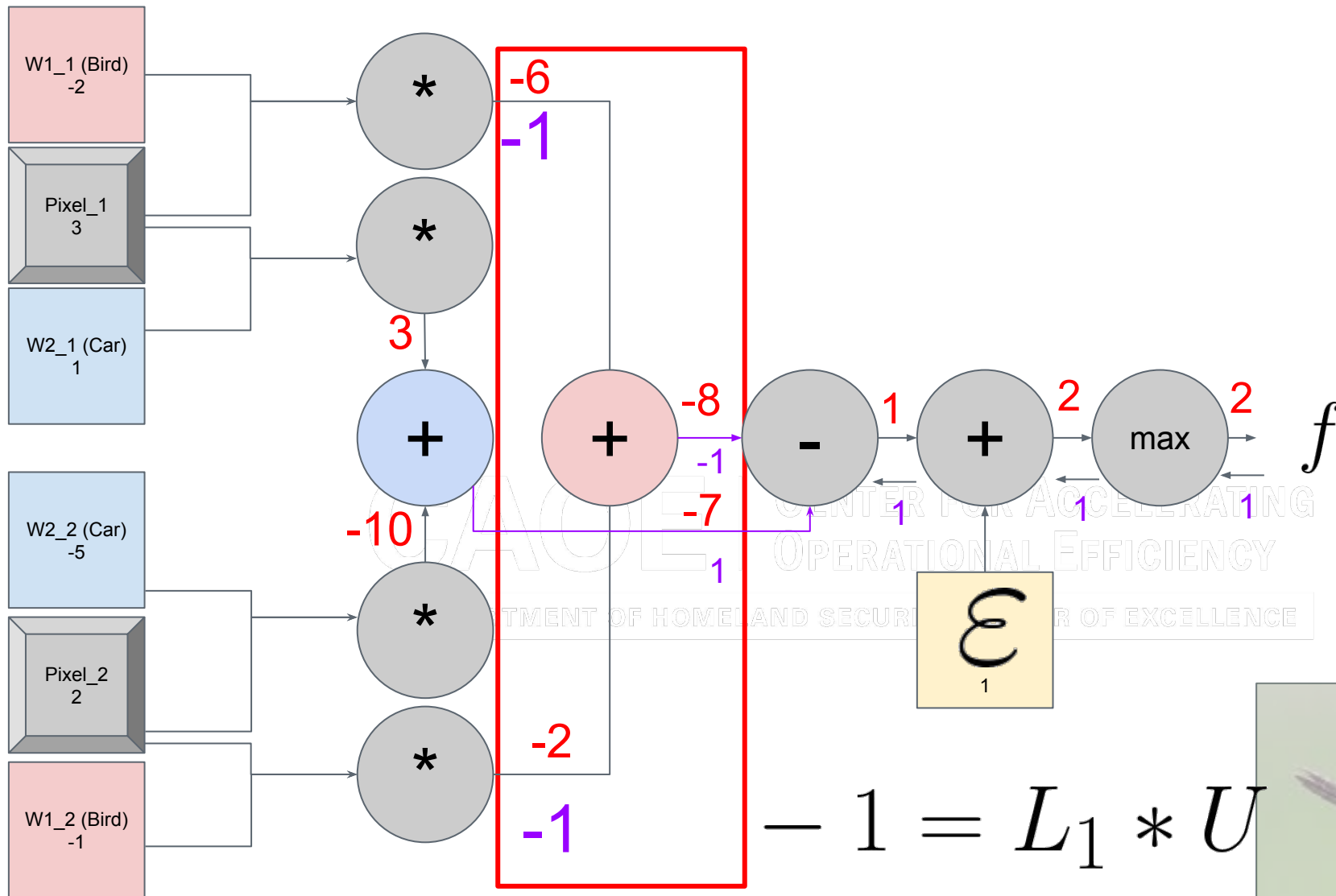
$$1 = L_1 * U$$

$$U = 1$$

$$L_1 = 1$$



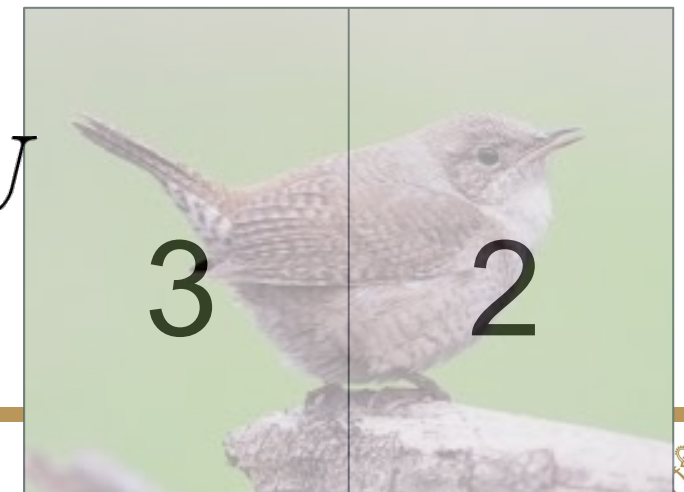


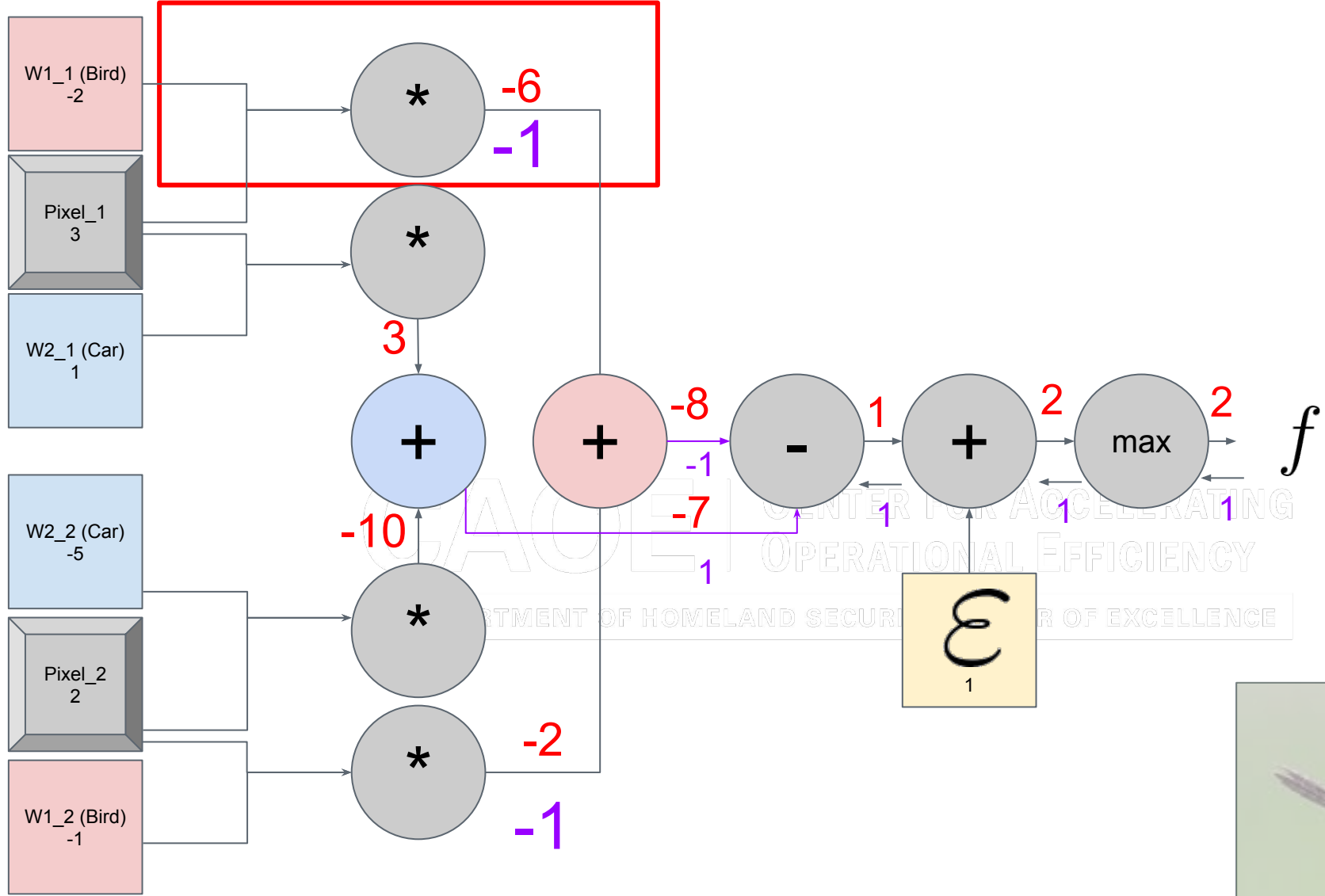


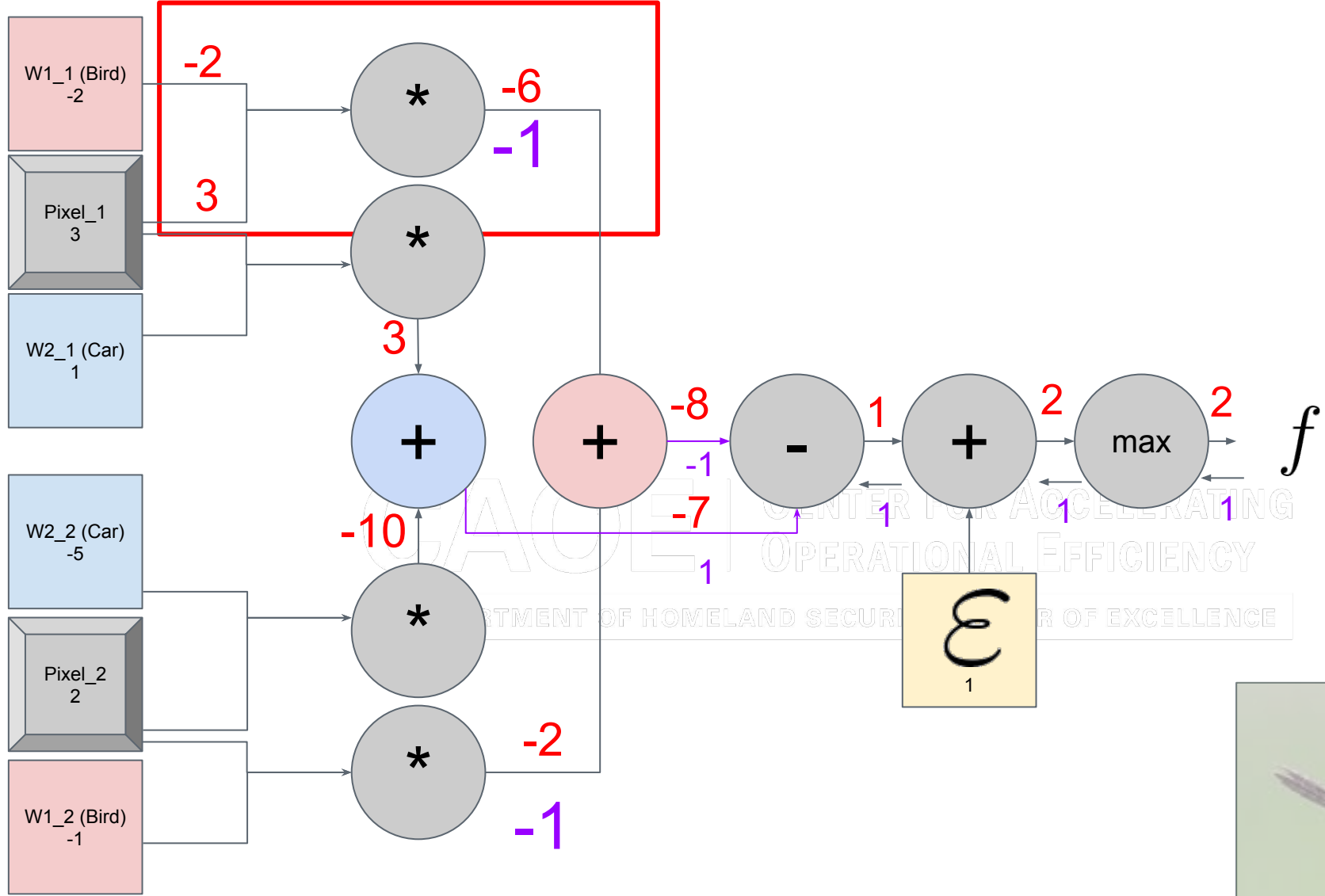
$$-1 = L_1 * U$$

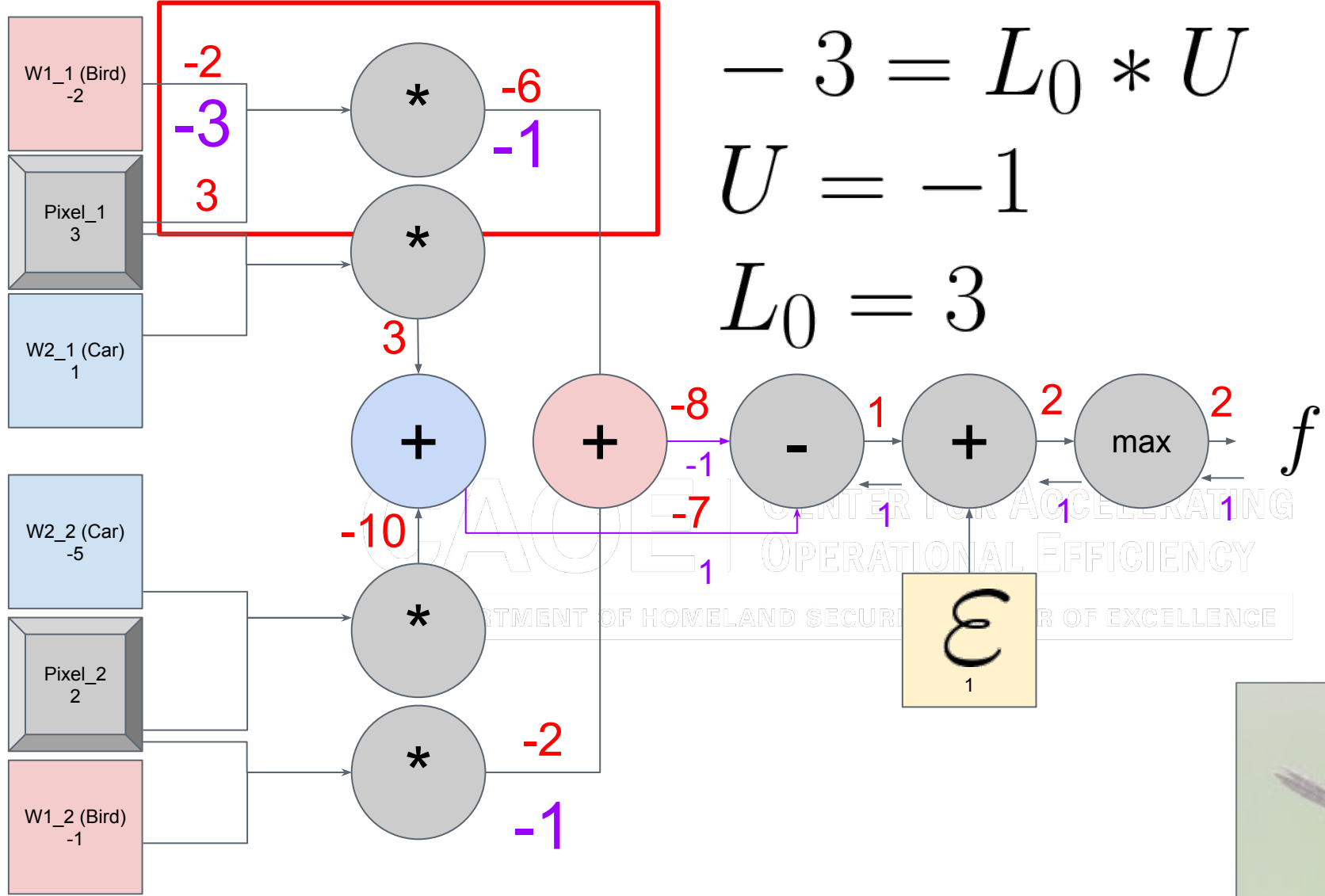
$$U = -1$$

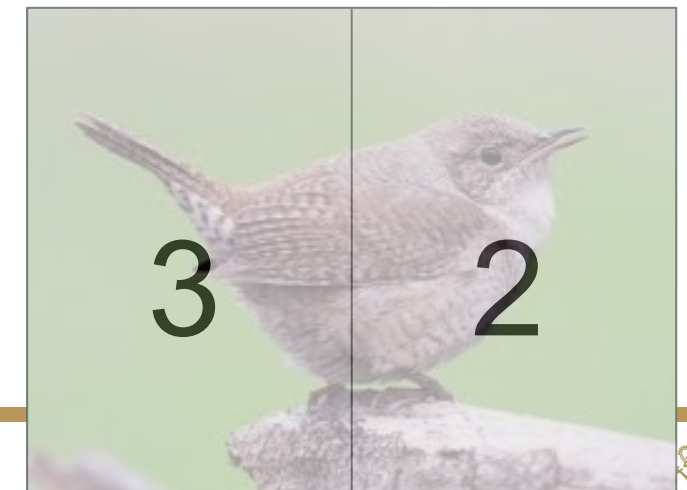
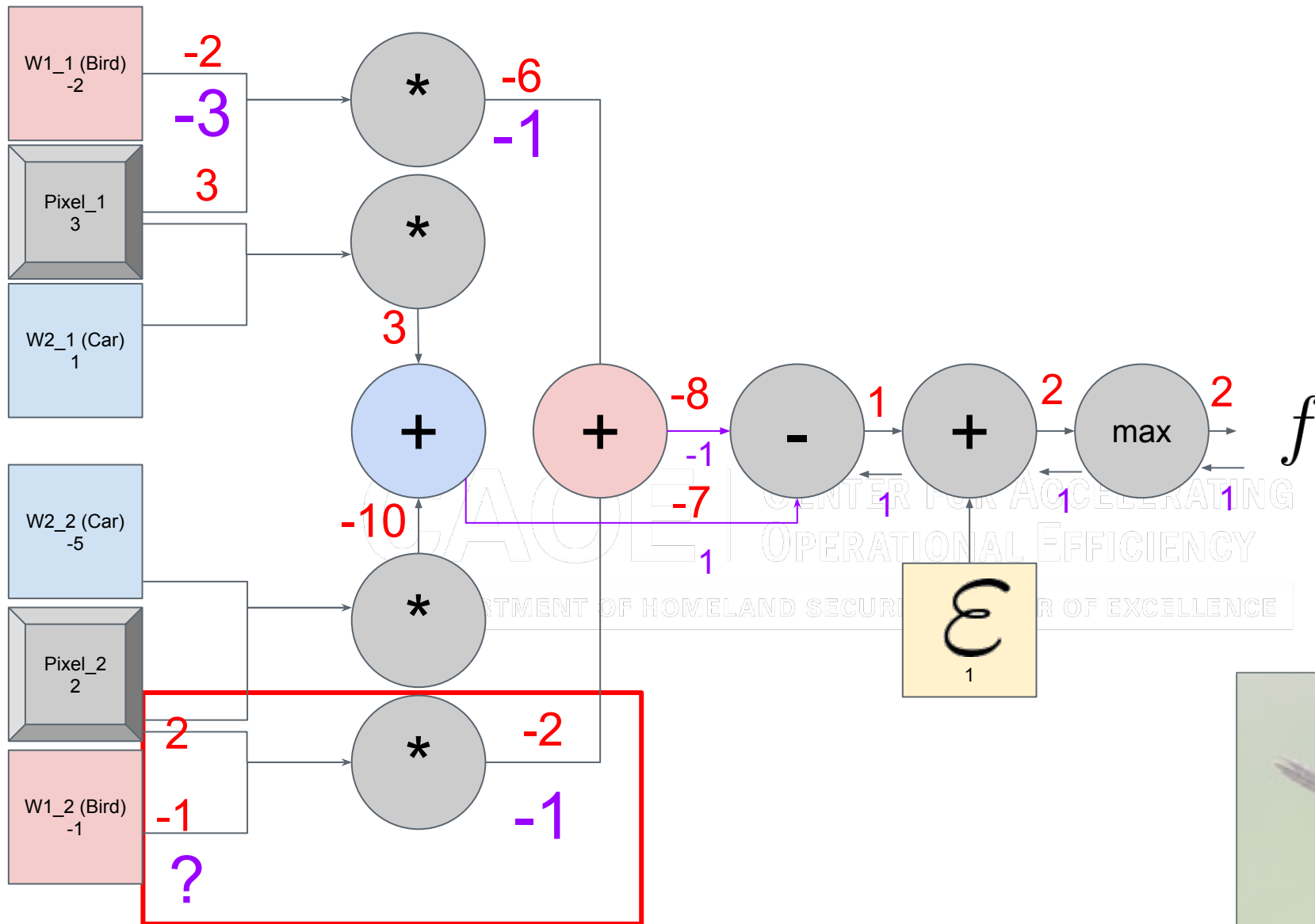
$$L_1 = 1$$

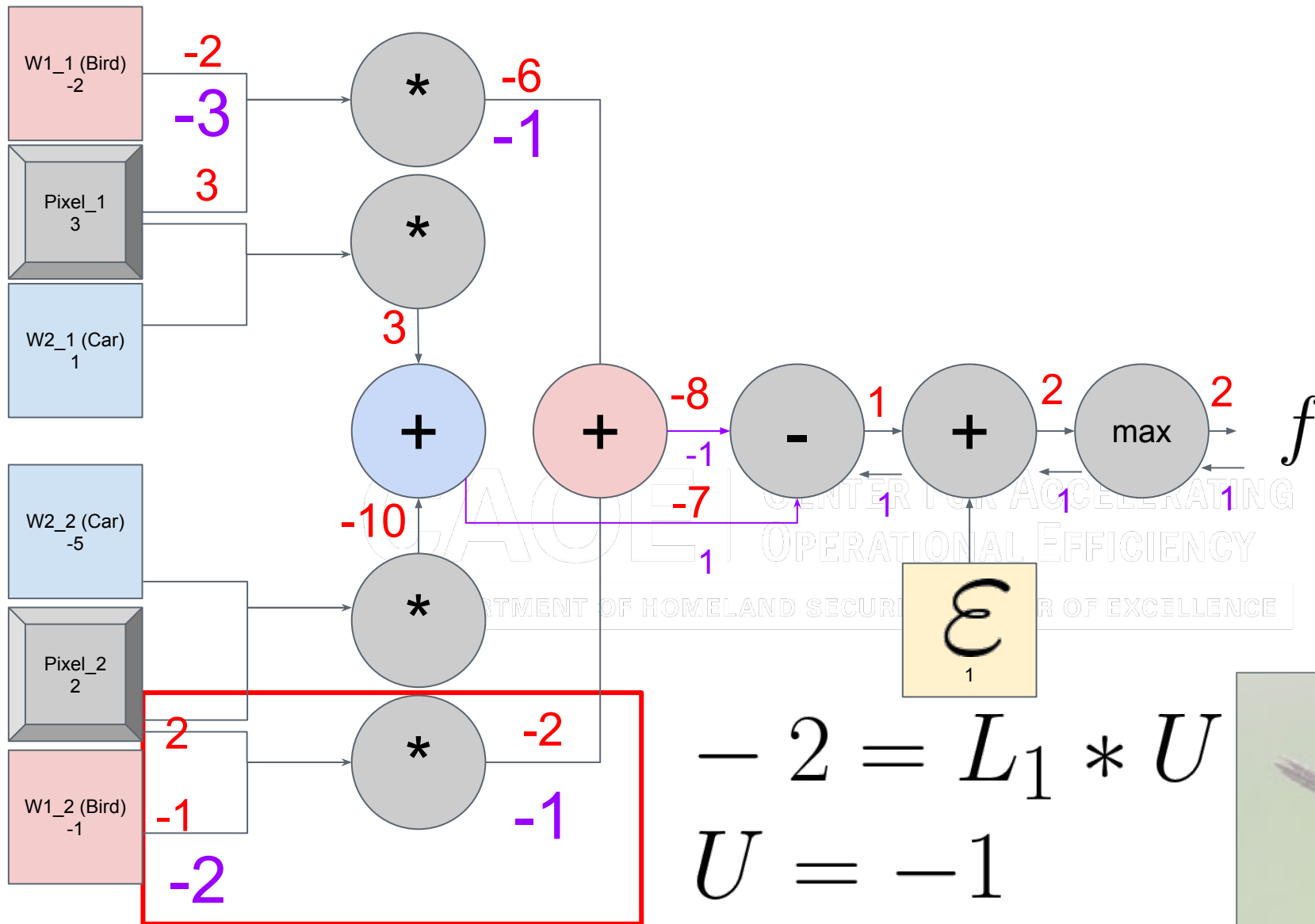










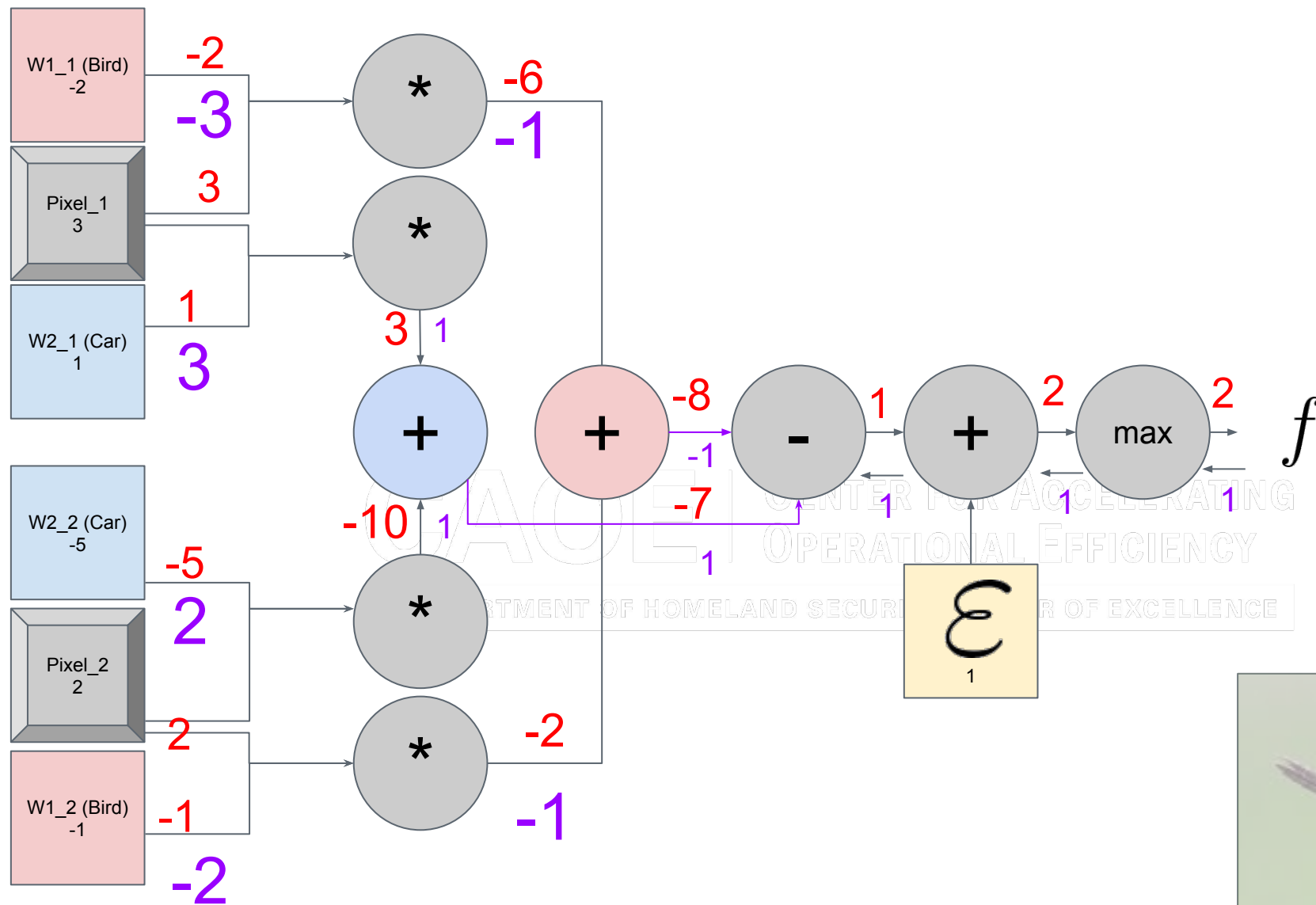


$$-2 = L_1 * U$$

$$U = -1$$

$$L_1 = 2$$





Weights

Gradient

W1_1
(Bird)
-2

-3

W2_1
(Car)
1

3

W2_2
(Car)
-5

2

W1_2
(Bird)
-1

-2

For these weights, in our forward pass,
the **Car** score was **-7**.

The **bird** score was **-8**.

**So in our
classification, we
guess Car. That isn't
what we want (i.e.,
it's Bad).**



Weights Gradient

W1_1
(Bird)
-2 -3

W2_1
(Car)
1 3

W2_2
(Car)
-5 2

W1_2
(Bird)
-1 -2

For these weights, in our forward pass,
the **Car** score was **-7**.

The **bird** score was **-8**.

Our **loss** was **2**.

CACOEI CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



Weights Gradient

W1_1
(Bird)
-2 -3

W2_1
(Car)
1 3

W2_2
(Car)
-5 2

W1_2
(Bird)
-1 -2

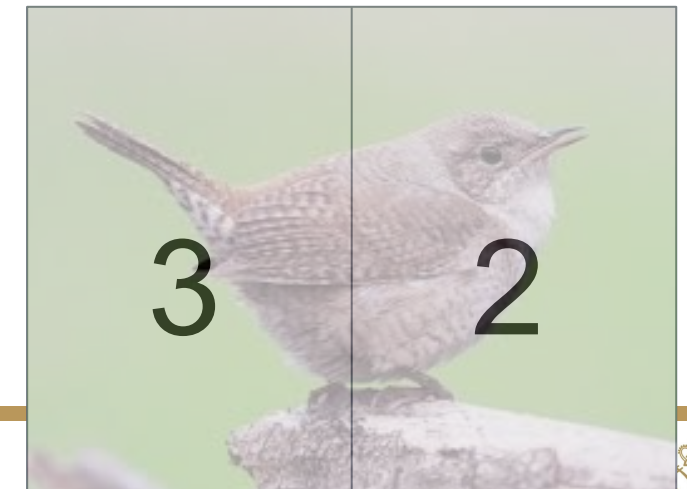
For these weights, in our forward pass,
the **Car** score was **-7**.

The **bird** score was **-8**.

Our **loss** was **2**.

CACOEI CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE



Weights	Gradient	Weights + (-1 * Gradient)
W1_1 (Bird) -2	-3	$-2 + (-1 * -3) =$ $-2 + 3 =$ 1
W2_1 (Car) 1	3	
W2_2 (Car) -5	2	
W1_2 (Bird) -1	-2	

CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY

A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

For these weights, in our forward pass,
the **Car** score was **-7**.

The **bird** score was **-8**.

Our **loss** was **2**.



Weights	Gradient	Weights + (-1 * Gradient)
---------	----------	---------------------------

W1_1 (Bird) -2	-3	1
----------------------	----	---

W2_1 (Car) 1	3	-2
--------------------	---	----

W2_2 (Car) -5	2	-7
---------------------	---	----

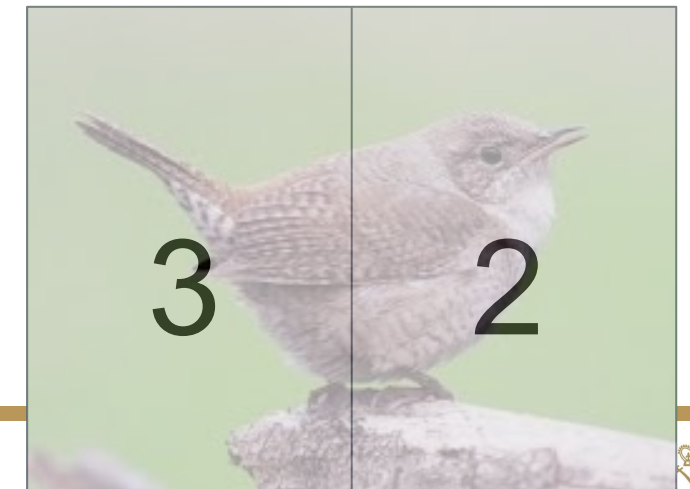
W1_2 (Bird) -1	-2	1
----------------------	----	---

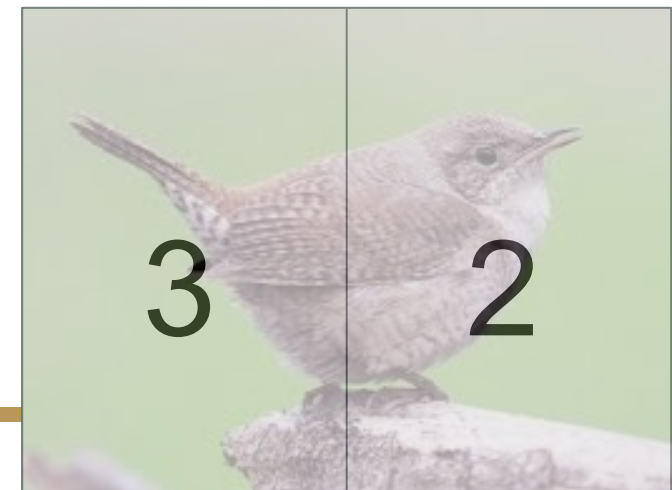
CAOE | CENTER FOR ACCELERATING
OPERATIONAL EFFICIENCY
A DEPARTMENT OF HOMELAND SECURITY CENTER OF EXCELLENCE

For these weights, in our forward pass,
the **Car** score was **-7**.

The **bird** score was **-8**.

Our **loss** was **2**.





Summary

- Computational Graphs
- Gradients & Partial Derivatives
- Backpropagation with a small example
- Next time: Matrix and vectorized backpropagation