Lecture 8

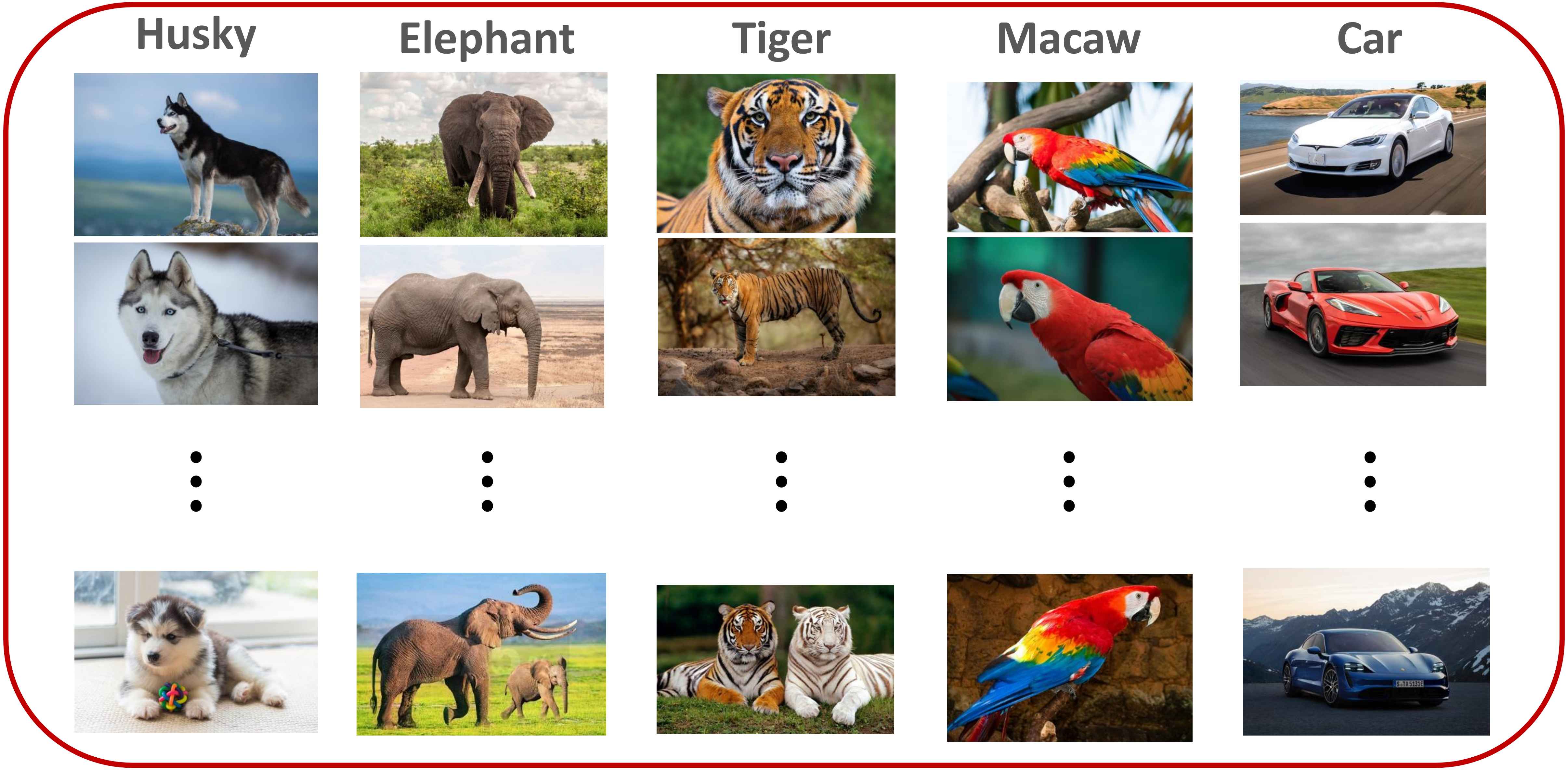
Metric Learning and Siamese Nets

Learning Pairwise Similarity Scores

Reference:

- Bromley et al. Signature verification using a Siamese time delay neural network. In *NIPS*. 1994.
- Koch, Zemel, & Salakhutdinov. Siamese neural networks for one-shot image recognition. In *ICML*, 2015.

Training Set



Training Data

Positive Samples	Negative Samples

Training Data

Positive Samples

Negative Samples

Training Data

Positive Samples	Negative Samples

Training Data

Positive Samples

Negative Samples

- (, , 1)

Training Data

Positive Samples

Negative Samples

- (, , 1)
- (, , 1)
- (, , 1)

Training Data

Positive Samples

- (, , 1)
- (, , 1)
- (, , 1)

Negative Samples

Training Data

Positive Samples

- (, , 1)
- (, , 1)
- (, , 1)

Negative Samples

Training Data

Positive Samples

- (, , 1)
- (, , 1)
- (, , 1)

Negative Samples

- (, , 0)

Training Data

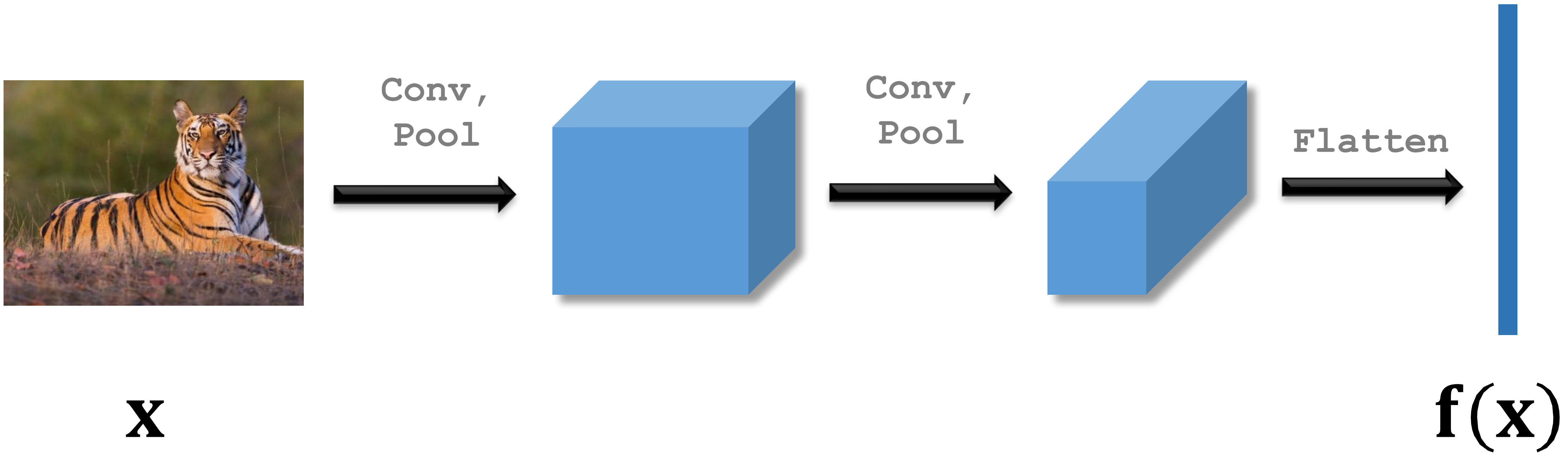
Positive Samples

- (, , 1)
- (, , 1)
- (, , 1)

Negative Samples

- (, , 0)
- (, , 0)
- (, , 0)

CNN for Feature Extraction



Training Siamese Network

$\mathbf{x}_1$

$\mathbf{x}_2$

Training Siamese Network

f

x1

f

x2

Training Siamese Network

$\mathbf{x1}$

$\xrightarrow{\mathbf{f}}$

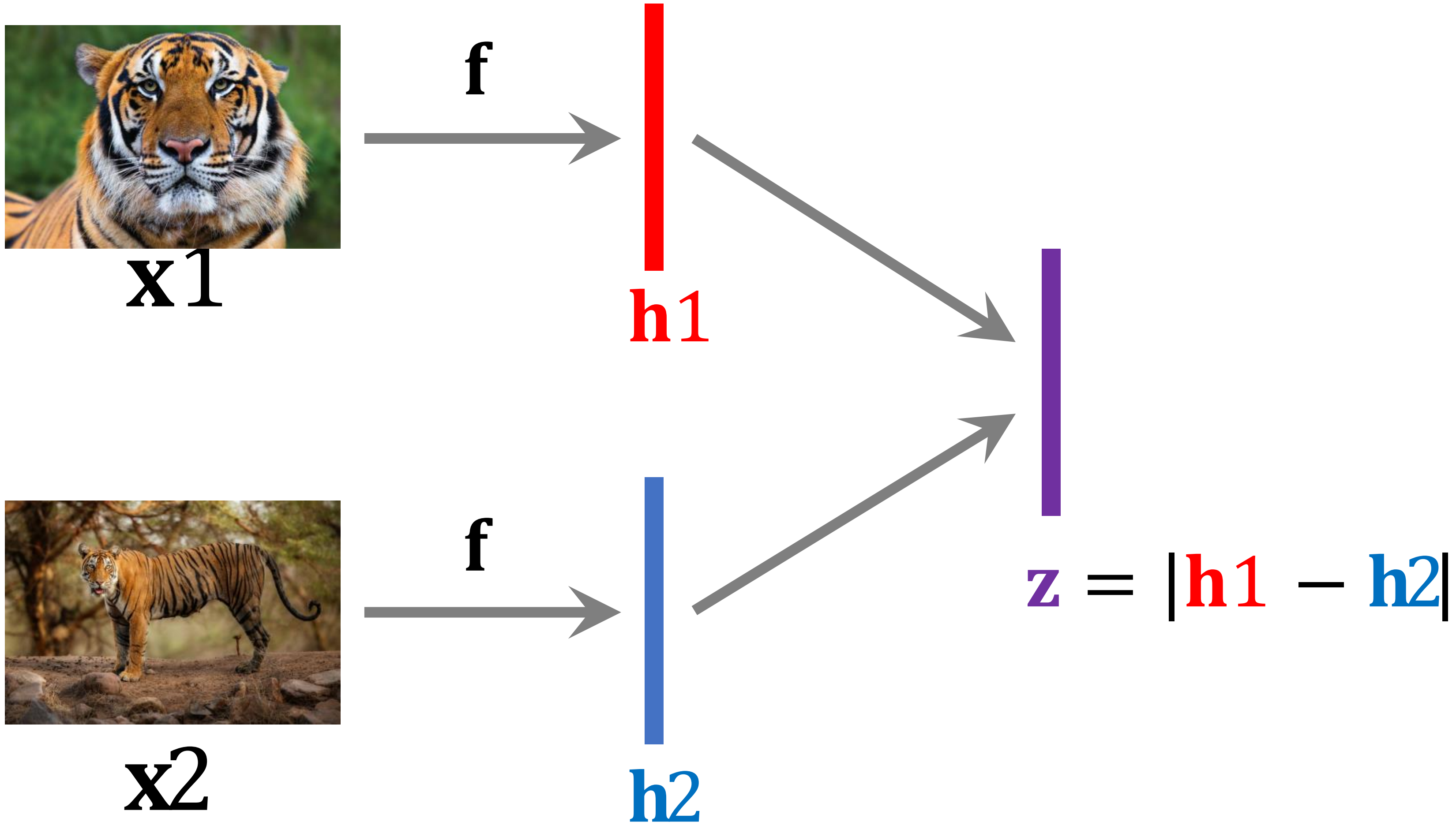
$\mathbf{h1} = \mathbf{f}(\mathbf{x1})$

$\mathbf{x2}$

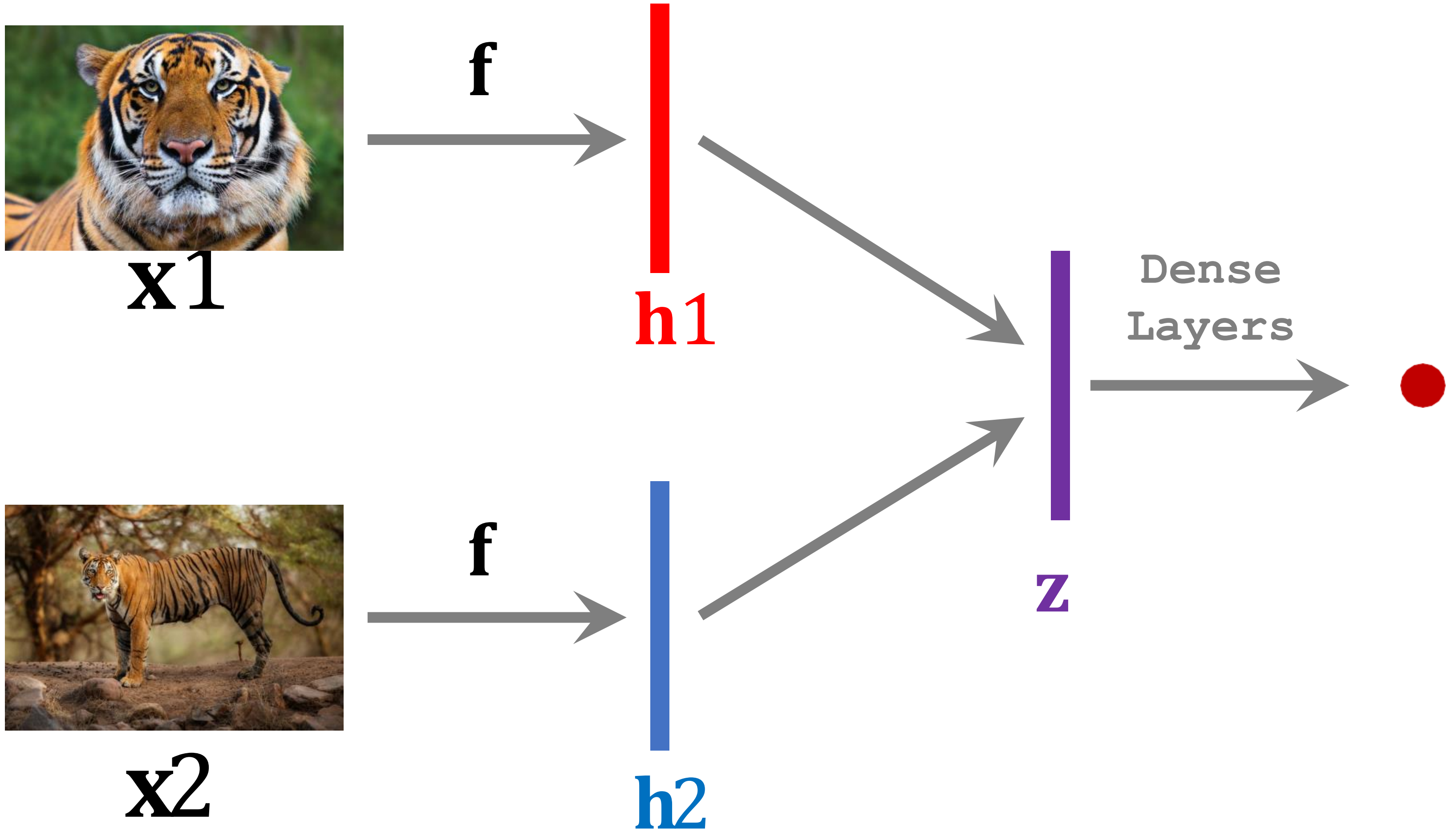
$\xrightarrow{\mathbf{f}}$

$\mathbf{h2} = \mathbf{f}(\mathbf{x2})$

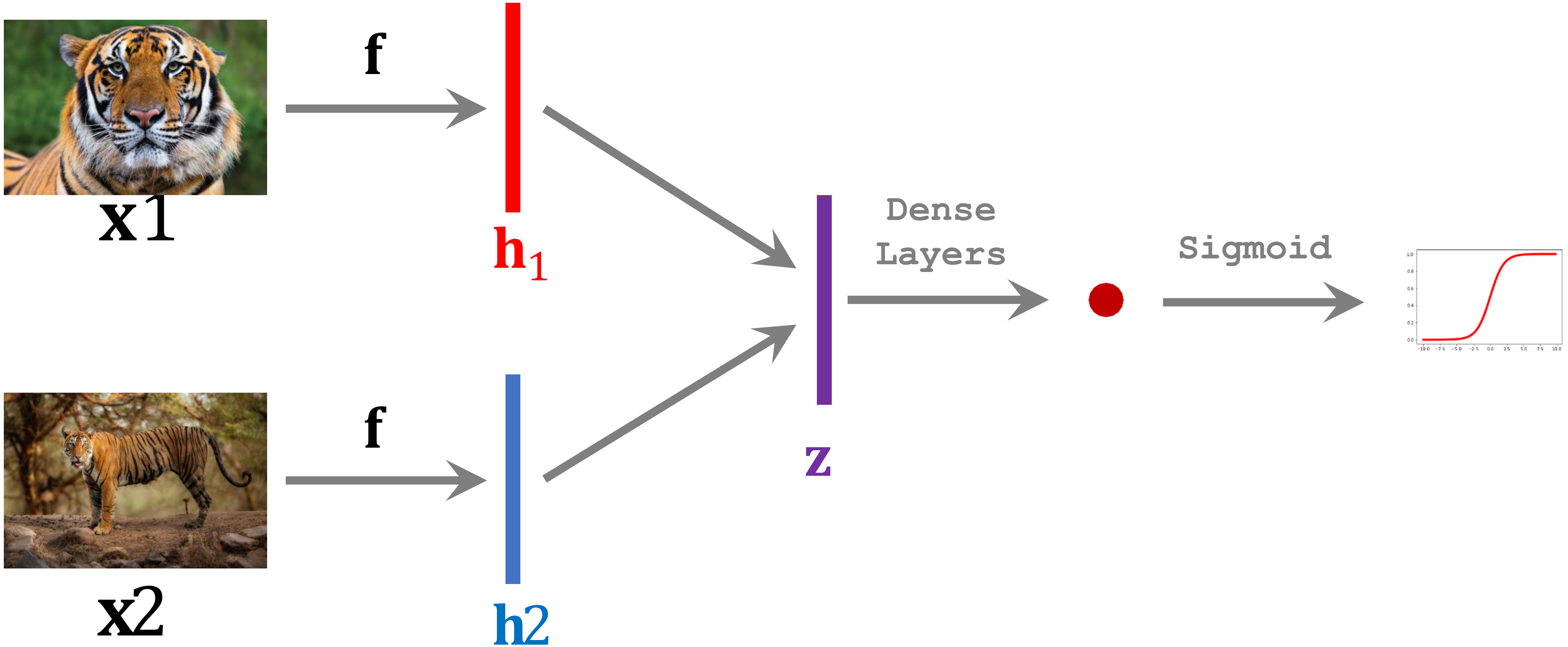
Training Siamese Network



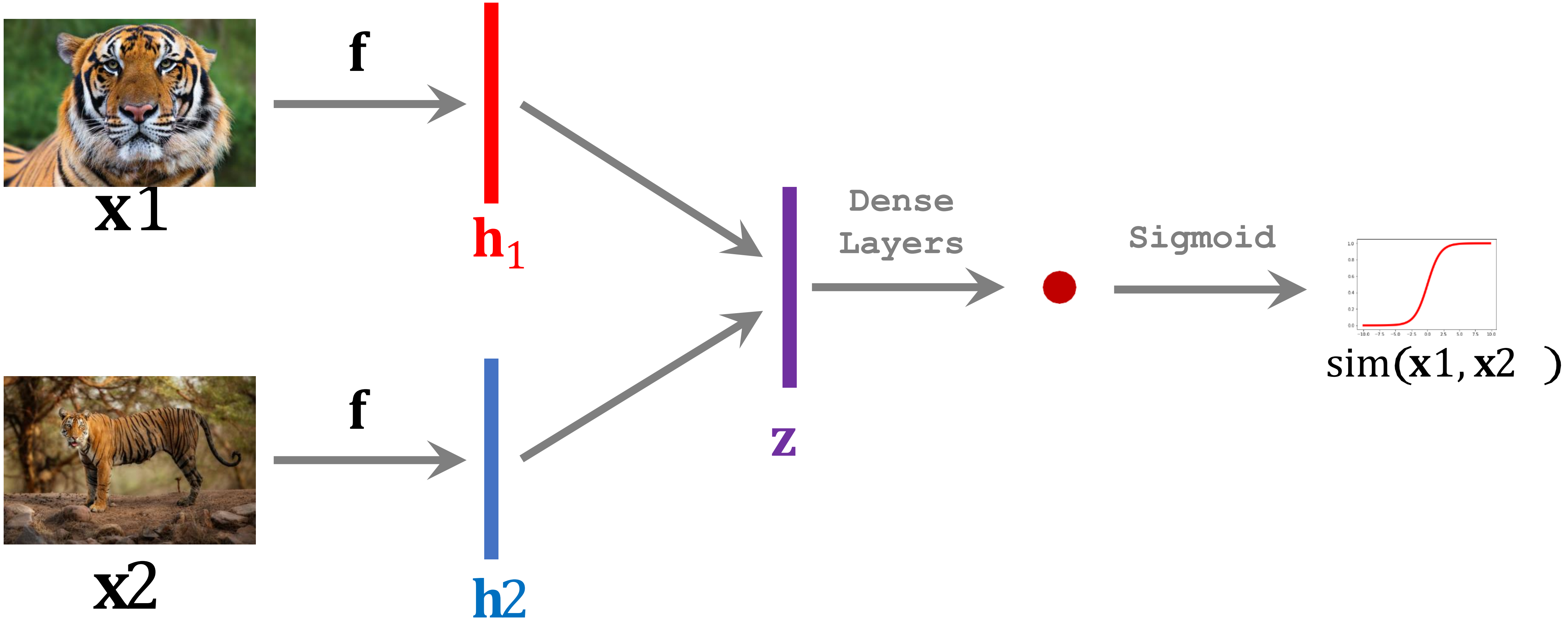
Training Siamese Network



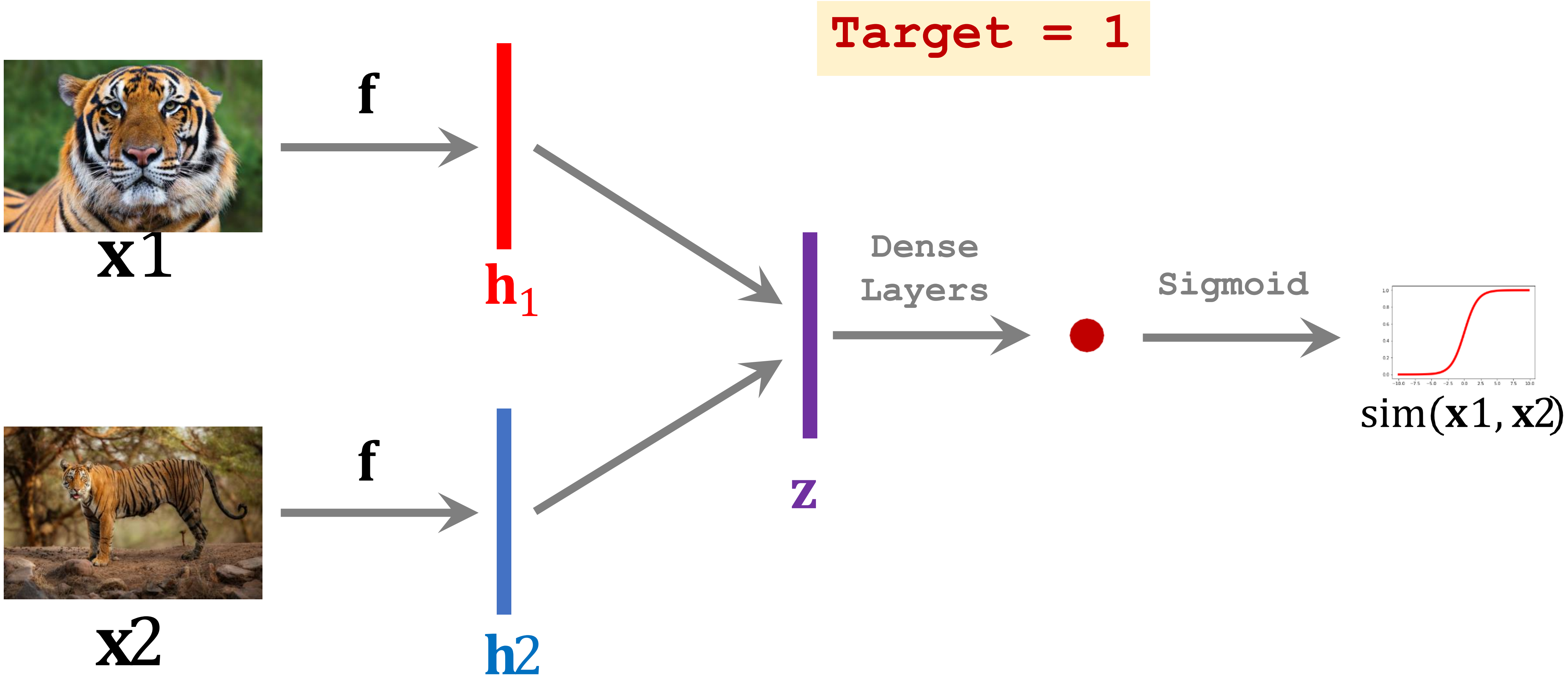
Training Siamese Network



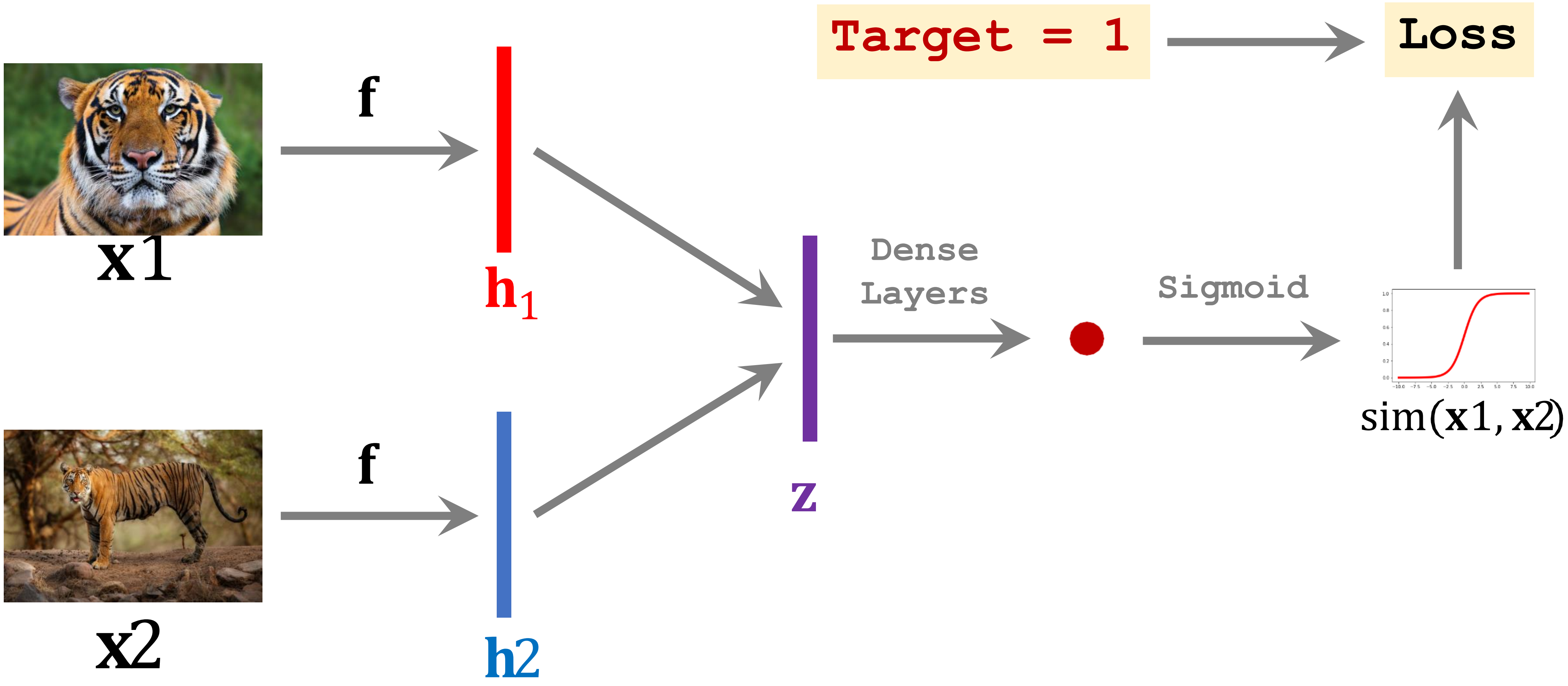
Training Siamese Network



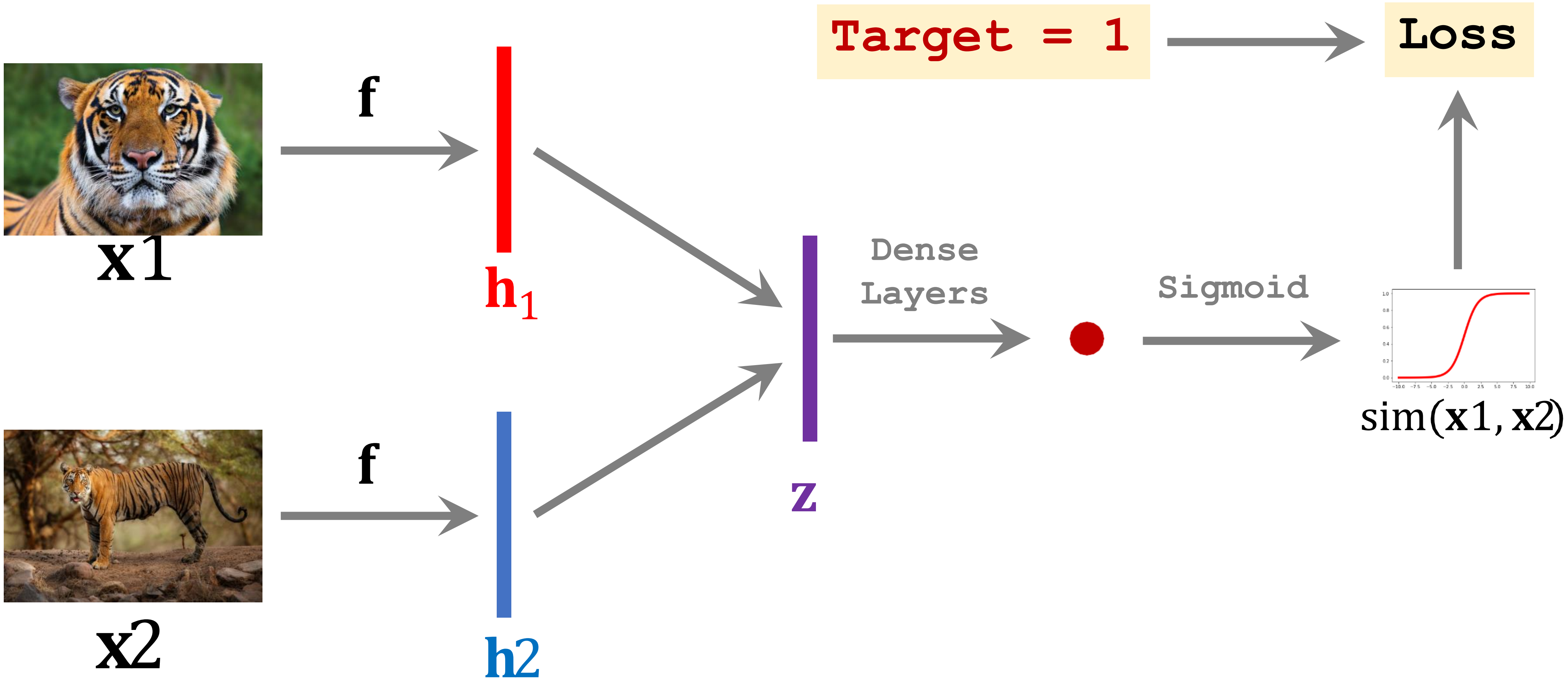
Training Siamese Network



Training Siamese Network



Training Siamese Network

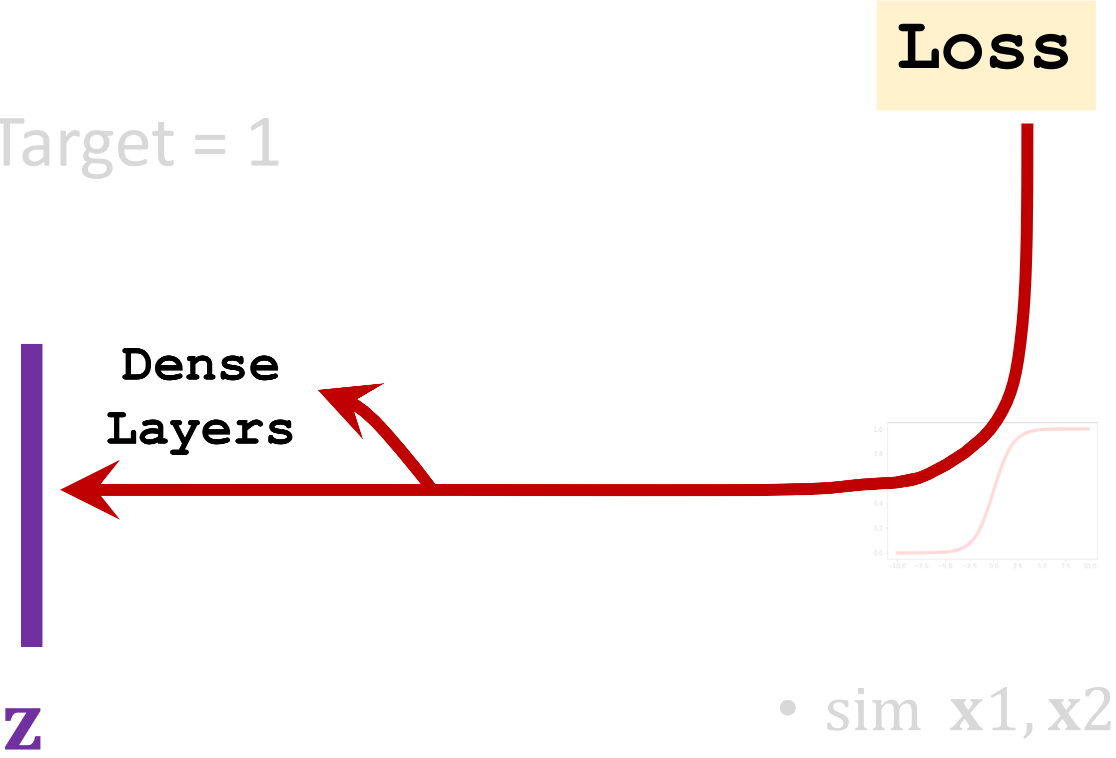


Training Siamese Network

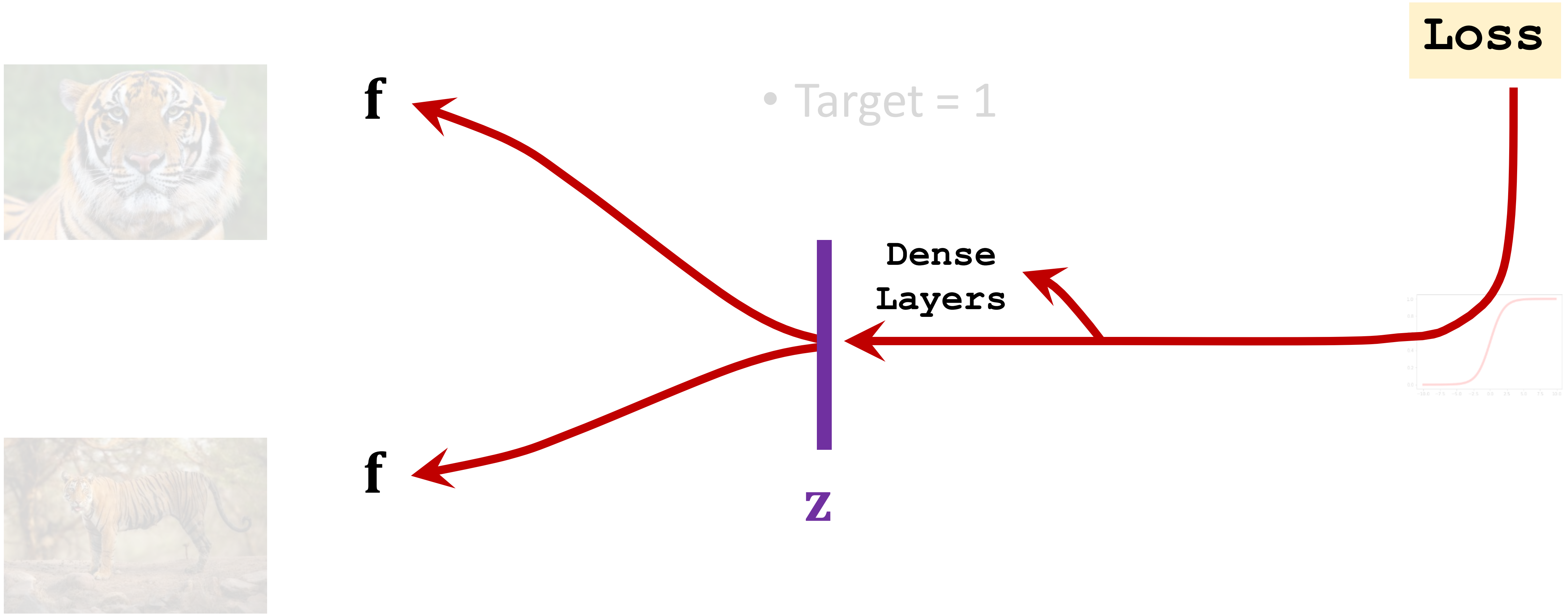
f

f

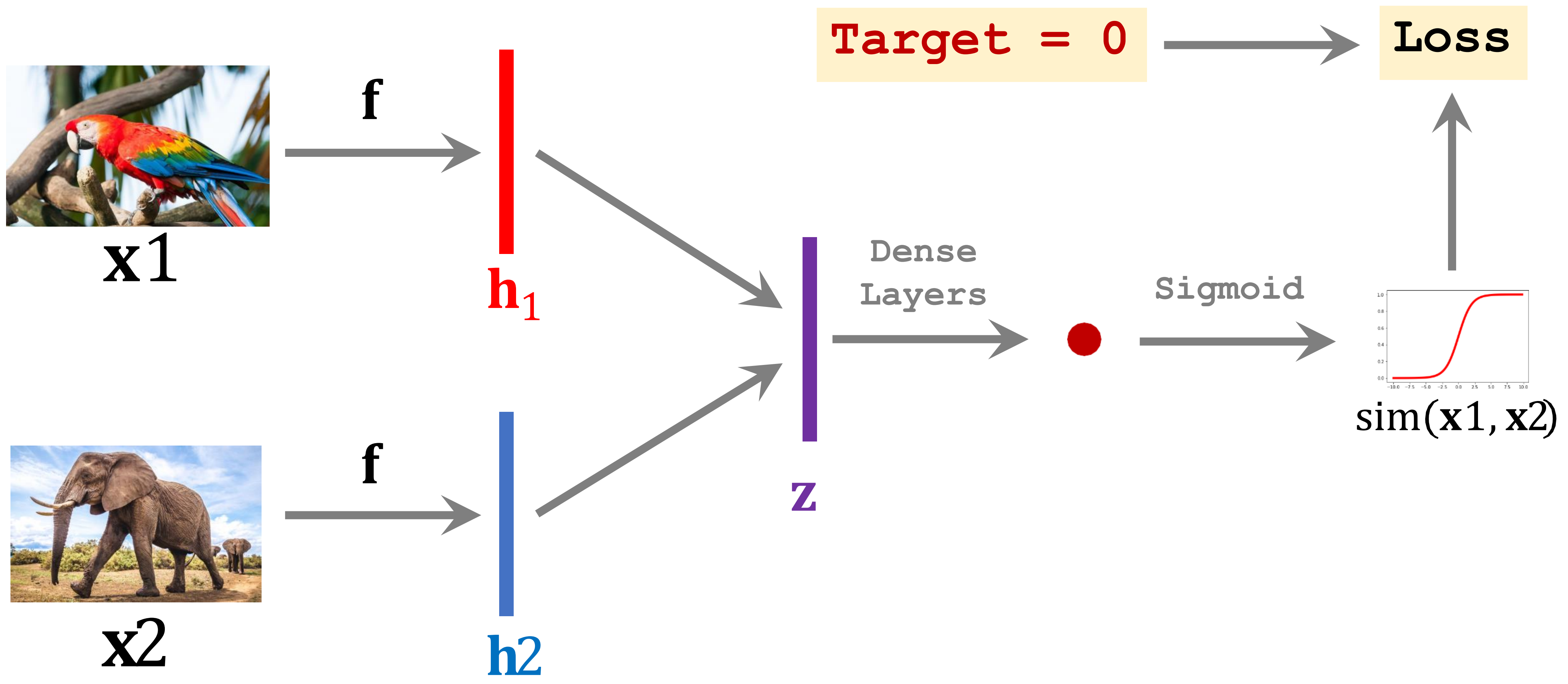
Sigmoid



Training Siamese Network



Training Siamese Network



One-Shot Prediction

- 6-way 1-shot prediction: support set has 6 test classes; each class has 1 sample.
- The training data (for the Siamese network) does not contain the 6 classes.

Support Set:

Fox

Squirrel

Rabbit

Hamster

Otter

Beaver

One-Shot Prediction

Query:

Support Set:

Fox

Squirrel

Rabbit

Hamster

Otter

Beaver

One-Shot Prediction

Query:

Fox	Squirrel	Rabbit	Hamster	Otter	Beaver
					

One-Shot Prediction

Query:

$\text{sim} = 0.2$

Fox

Squirrel

Rabbit

Hamster

Otter

Beaver

One-Shot Predictic

Query:

sim = 0.2

sim = 0.9

Fox

Squirrel

Rabbit

Hamster

Otter

Beaver

One-Shot Predictio

Query:

sim = 0.2	sim = 0.9	sim = 0.7	sim = 0.5	sim = 0.3	sim = 0.4
Fox	**Squirrel**	**Rabbit**	**Hamster**	**Otter**	**Beaver**

One-Shot Predictio

Query:

sim = 0.2

Fox

sim = 0.9

Squirrel

sim = 0.7

Rabbit

sim = 0.5

Hamster

sim = 0.3

Otter

sim = 0.4

Beaver

Triplet Loss

Reference:

- Schroff, Kalenichenko, & Philbin. Facenet: A unified embedding for face recognition and clustering. In *CVPR*, 2015.

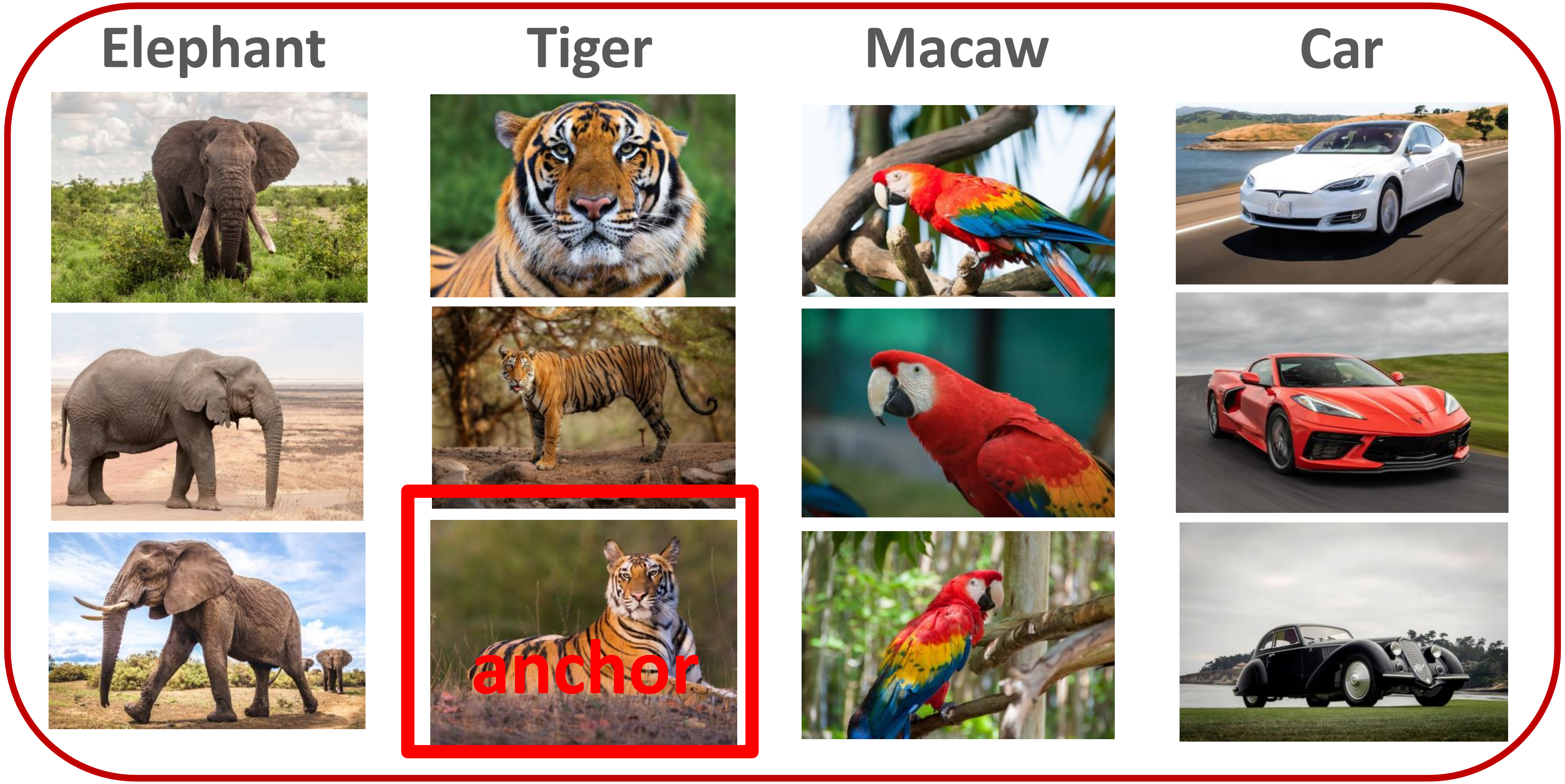
Data for Training Siamese Network

Training Set

Elephant	Tiger	Macaw	Car

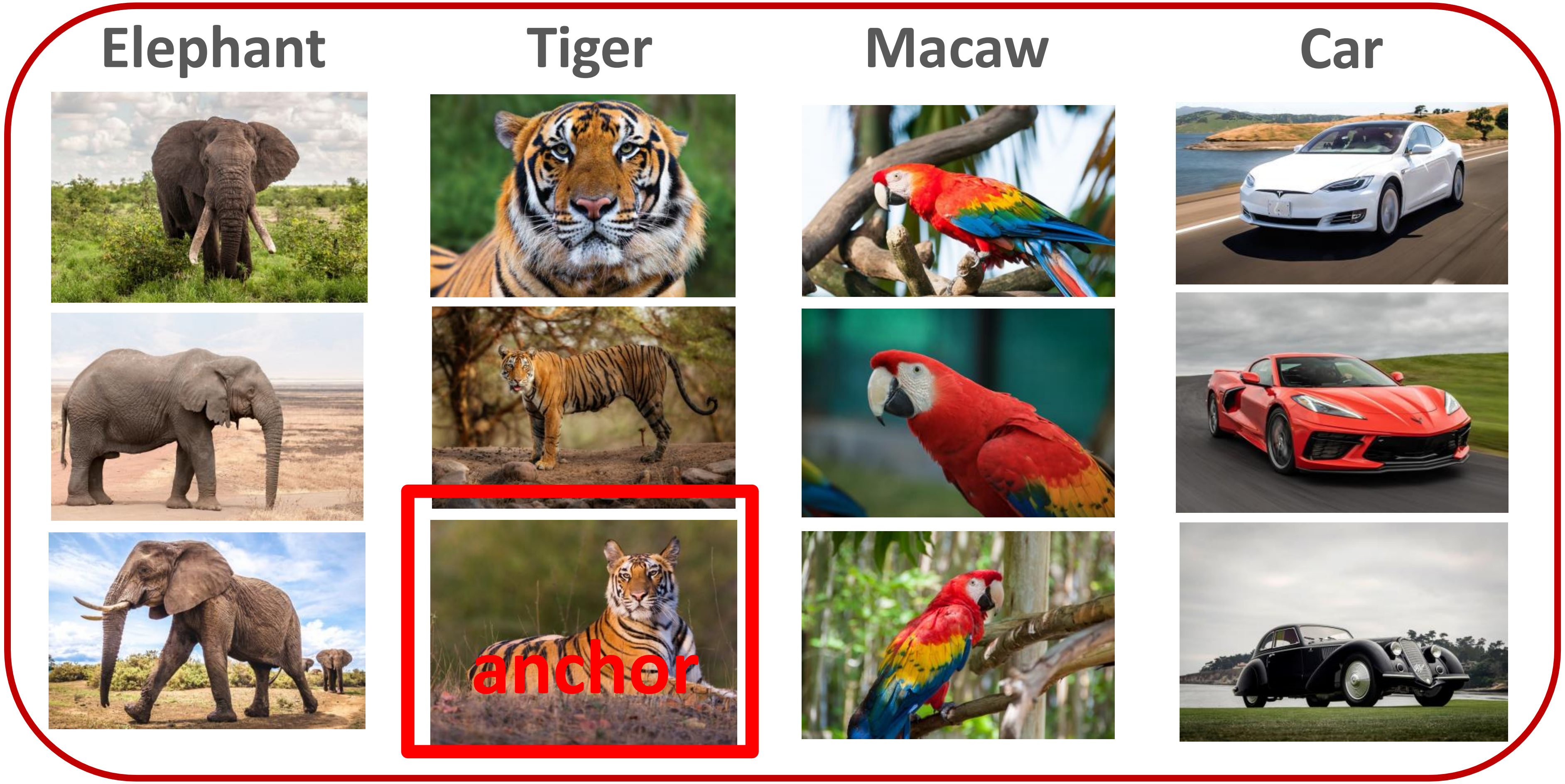
Data for Training Siamese Network

Training Set



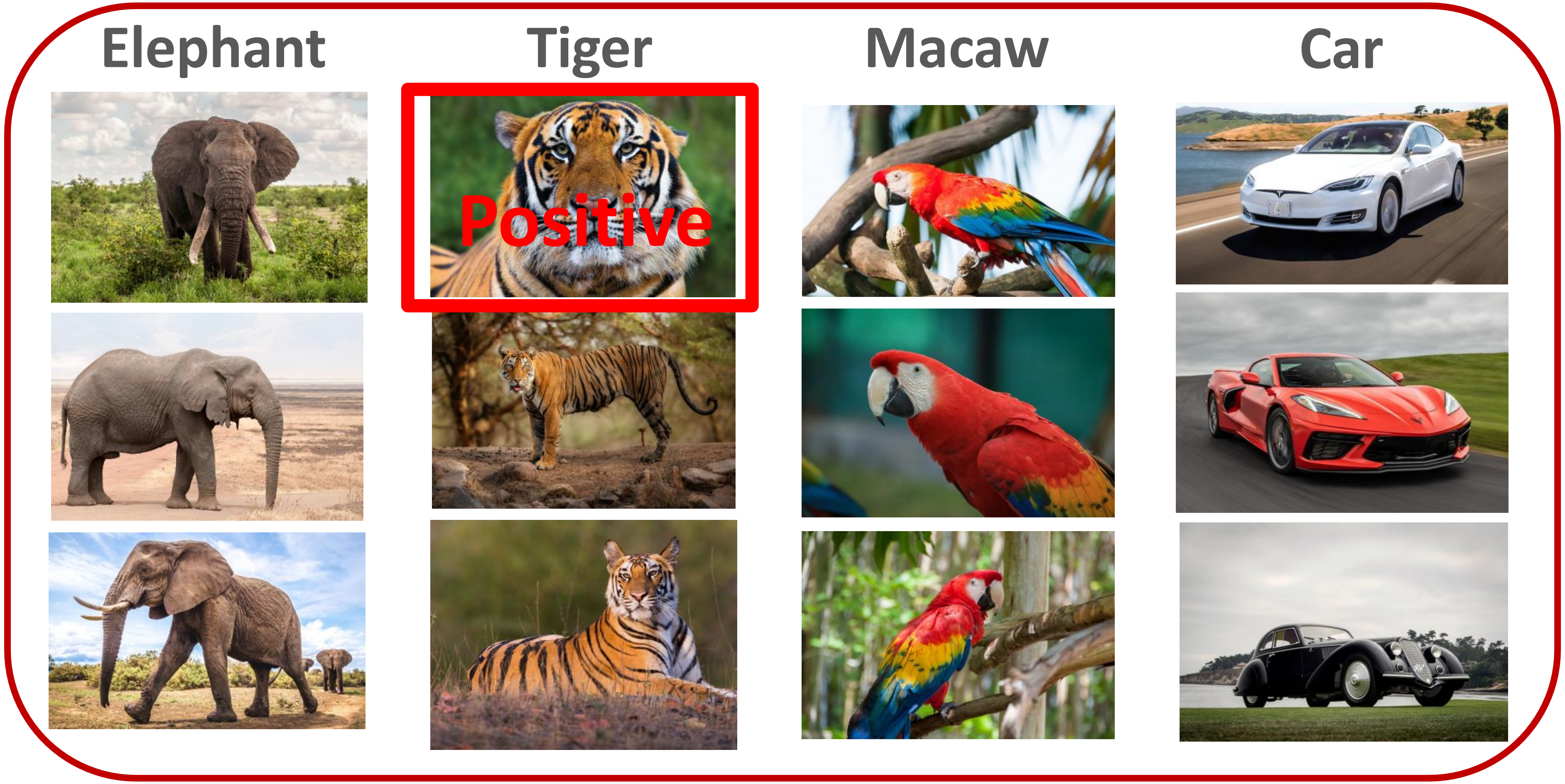
Data for Training Siamese Network

Training Set



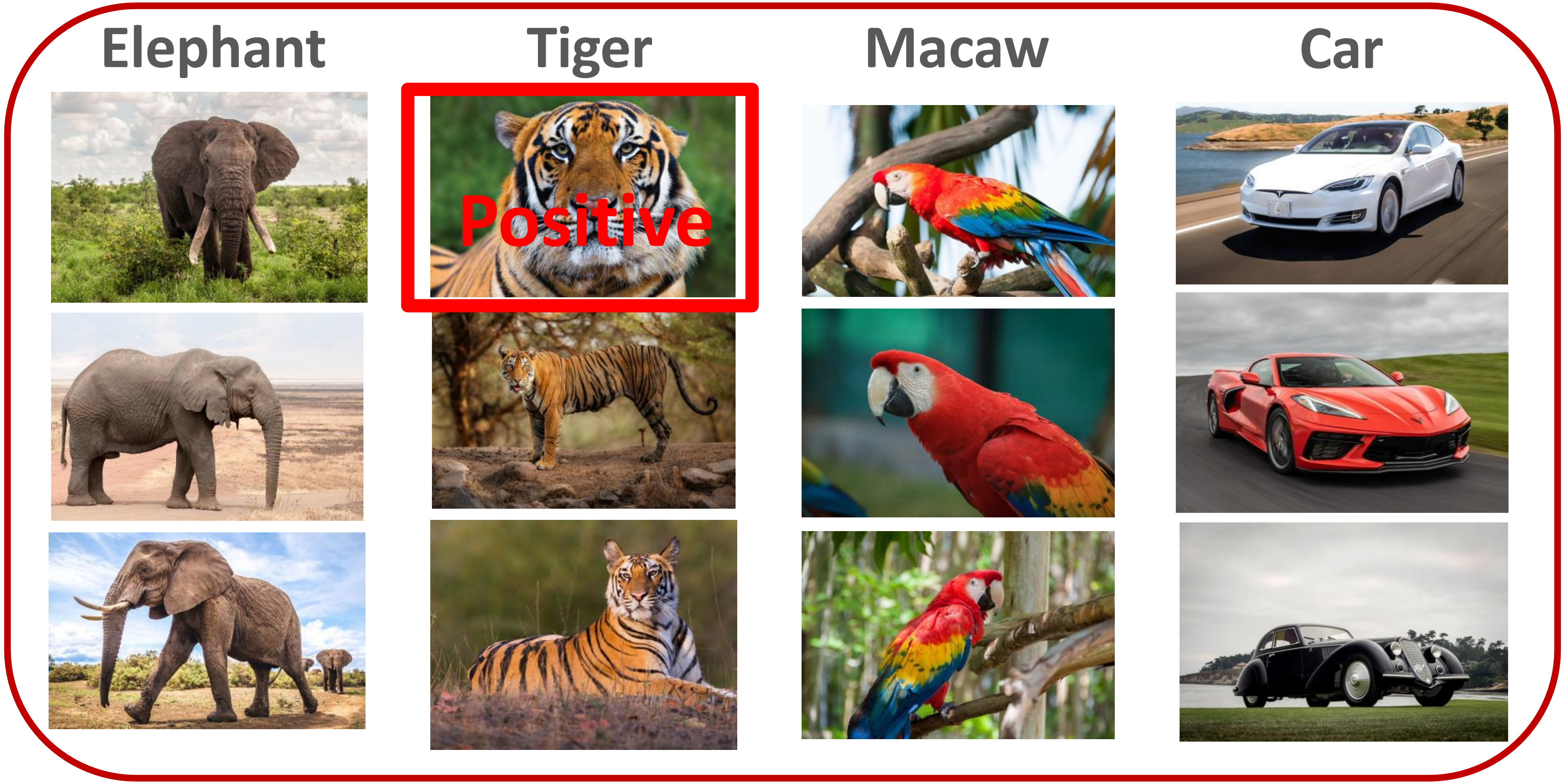
Data for Training Siamese Network

Training Set



Data for Training Siamese Network

Training Set



Data for Training Siamese Network

Training Set

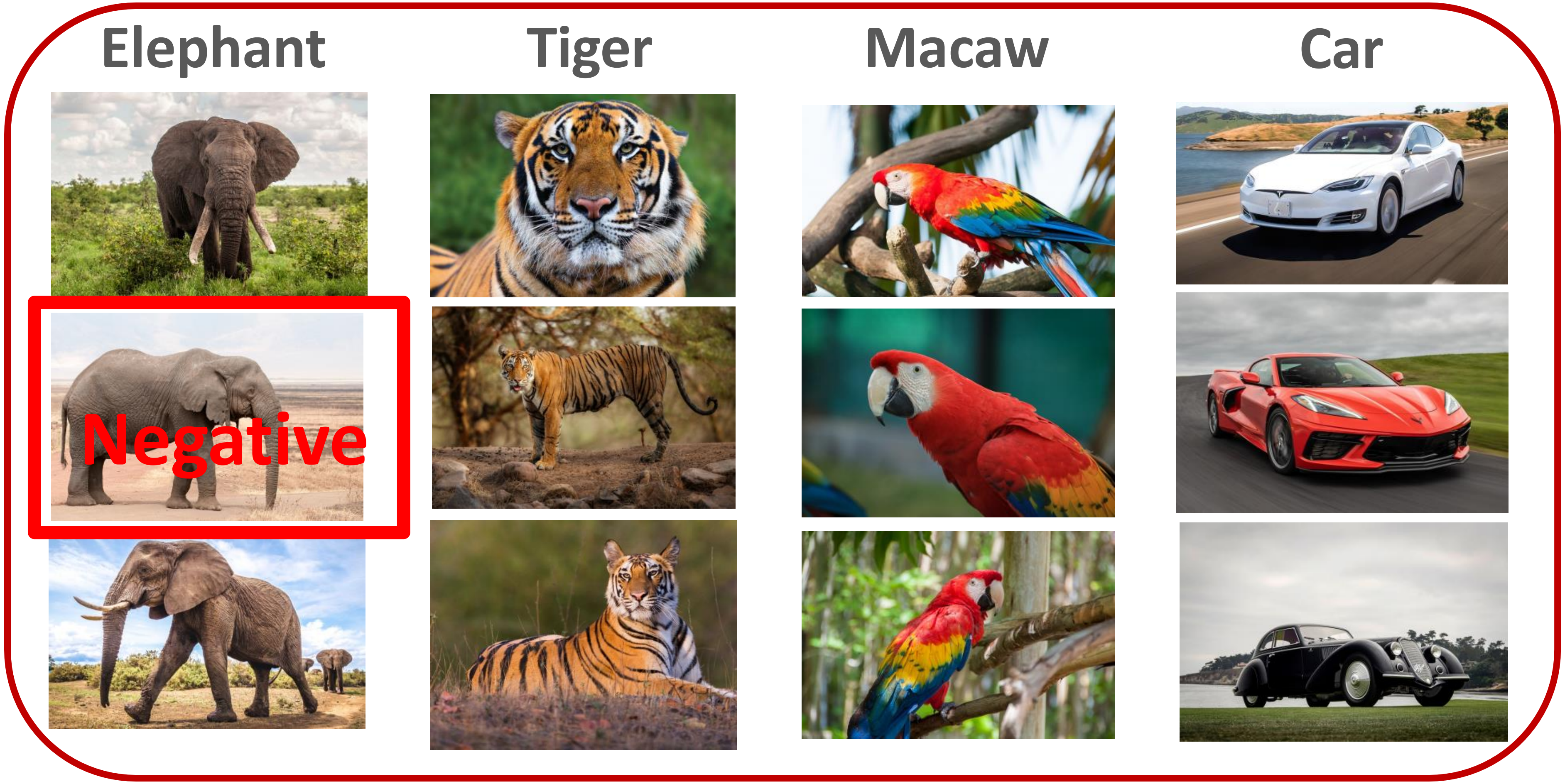
Elephant

Macaw

Car

Data for Training Siamese Network

Training Set



Triplet Loss

x+

(positive)

xa

(anchor)

(negative)

x-

Triplet Loss

x+

(positive)

$\xrightarrow{f}$

xa

(anchor)

$\xrightarrow{f}$

x-

(negative)

$\xrightarrow{f}$

Triplet Loss

x+

(positive)

$x+ \xrightarrow{f} f(x+)$

xa

(anchor)

$xa \xrightarrow{f} f(xa)$

x-

(negative)

$x- \xrightarrow{f} f(x-)$

Triplet Loss

$\mathbf{x+}$ (positive) $\xrightarrow{\mathbf{f}}$ $\mathbf{f}(\mathbf{x+})$

$\mathbf{xa}$ (anchor) $\xrightarrow{\mathbf{f}}$ $\mathbf{f}(\mathbf{xa})$

$$d^{+} = ||\mathbf{f}(\mathbf{x+}) - \mathbf{f}(\mathbf{xa})||_2^2$$

$\mathbf{x-}$ (negative) $\xrightarrow{\mathbf{f}}$ $\mathbf{f}(\mathbf{x-})$

Triplet Loss

$\mathbf{x+}$ (positive) $\xrightarrow{\mathbf{f}}$ $\mathbf{f}(\mathbf{x+})$

$\mathbf{xa}$ (anchor) $\xrightarrow{\mathbf{f}}$ $\mathbf{f}(\mathbf{xa})$

$\mathbf{x-}$ (negative) $\xrightarrow{\mathbf{f}}$ $\mathbf{f}(\mathbf{x-})$

$$d^{+} = \left|\left|\mathbf{f}(\mathbf{x+}) - \mathbf{f}(\mathbf{xa})\right|\right|_2^2$$

$$d^{-} = \left|\left|\mathbf{f}(\mathbf{xa}) - \mathbf{f}(\mathbf{x-})\right|\right|_2^2$$

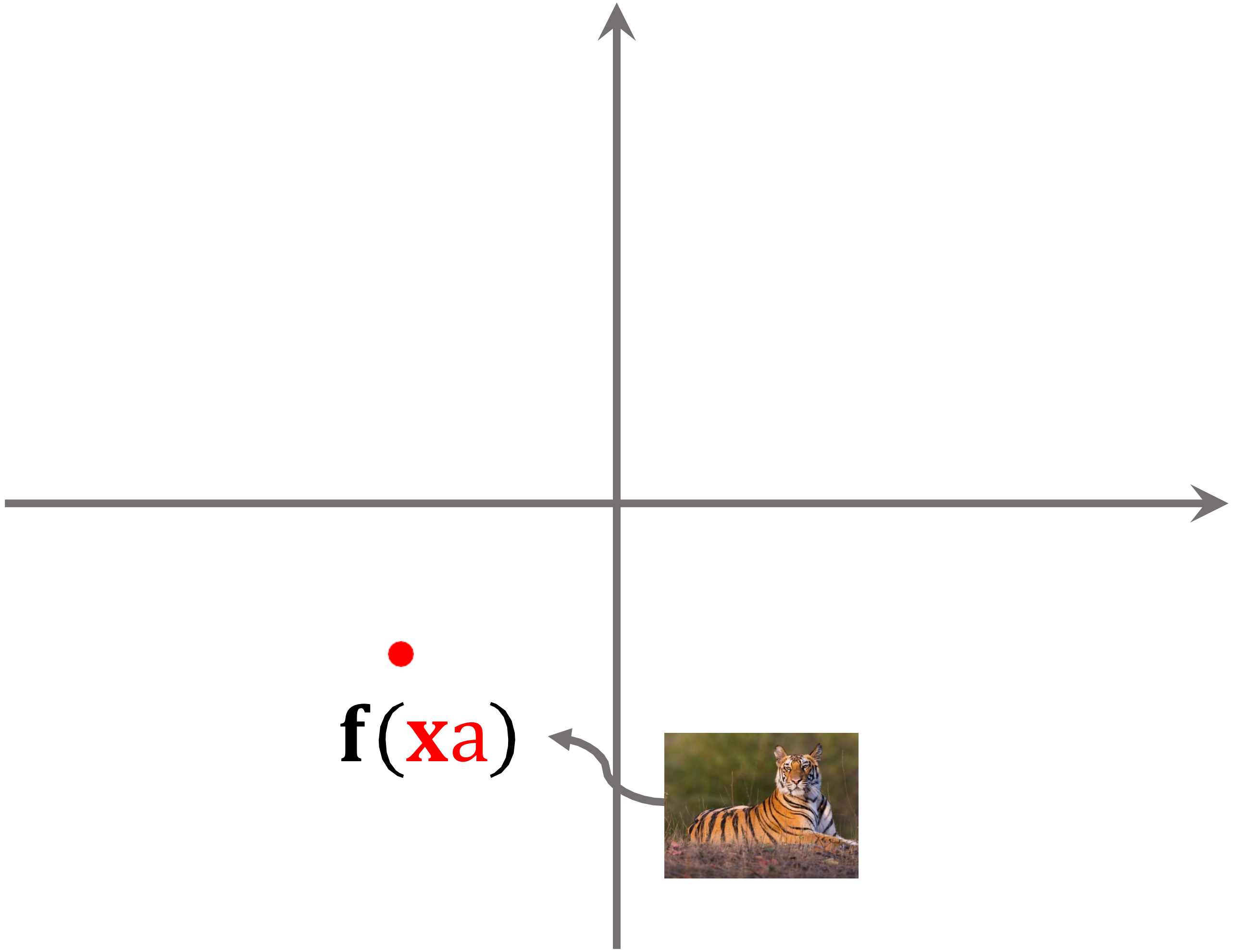
Triplet Loss

x+
(positive)

xa
(anchor)

x-
(negative)

Feature Space

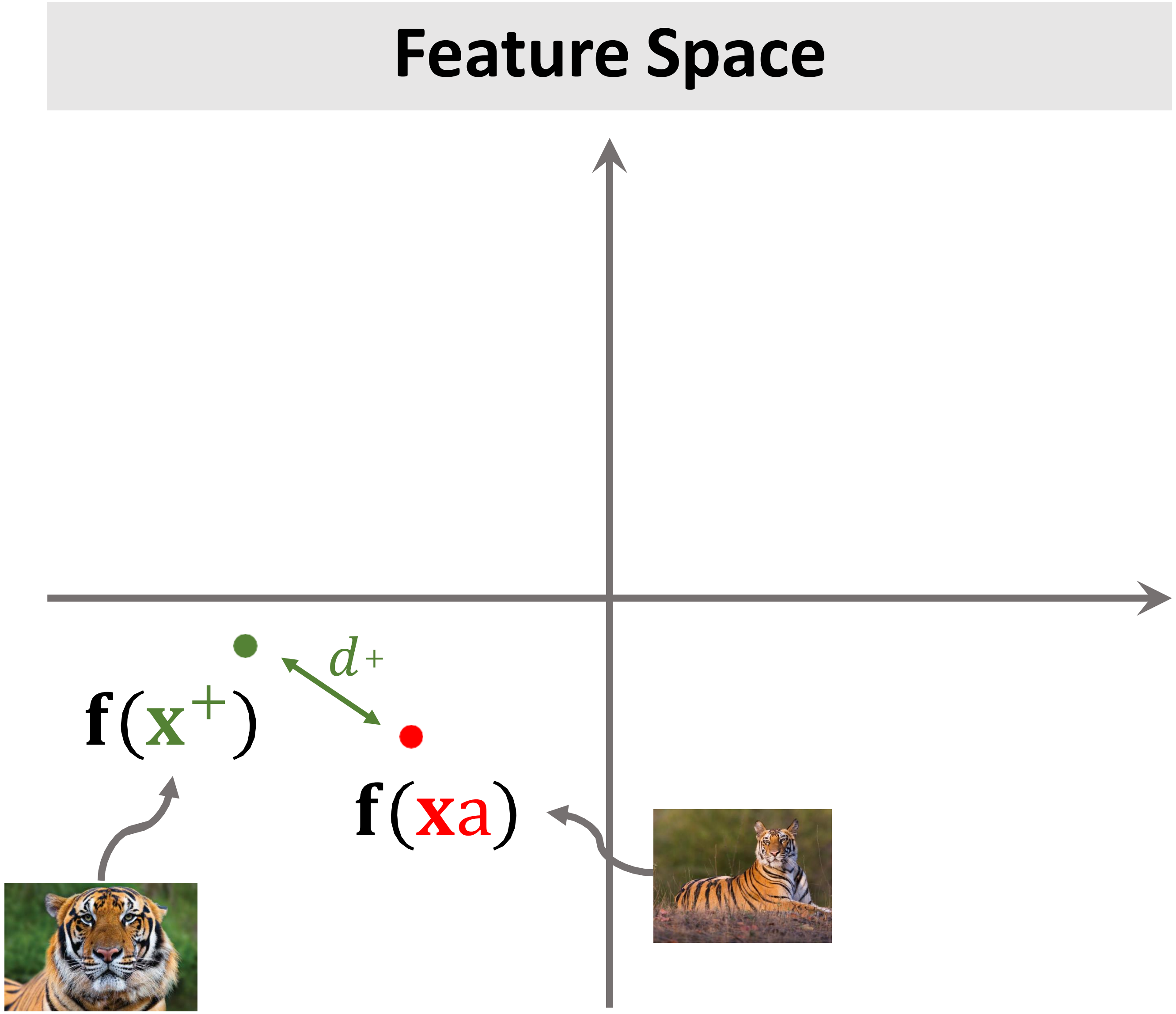


Triplet Loss

$\mathbf{x}+$
(positive)

$\mathbf{xa}$
(anchor)

$\mathbf{x}$-
(negative)



Triplet Loss

$x+$

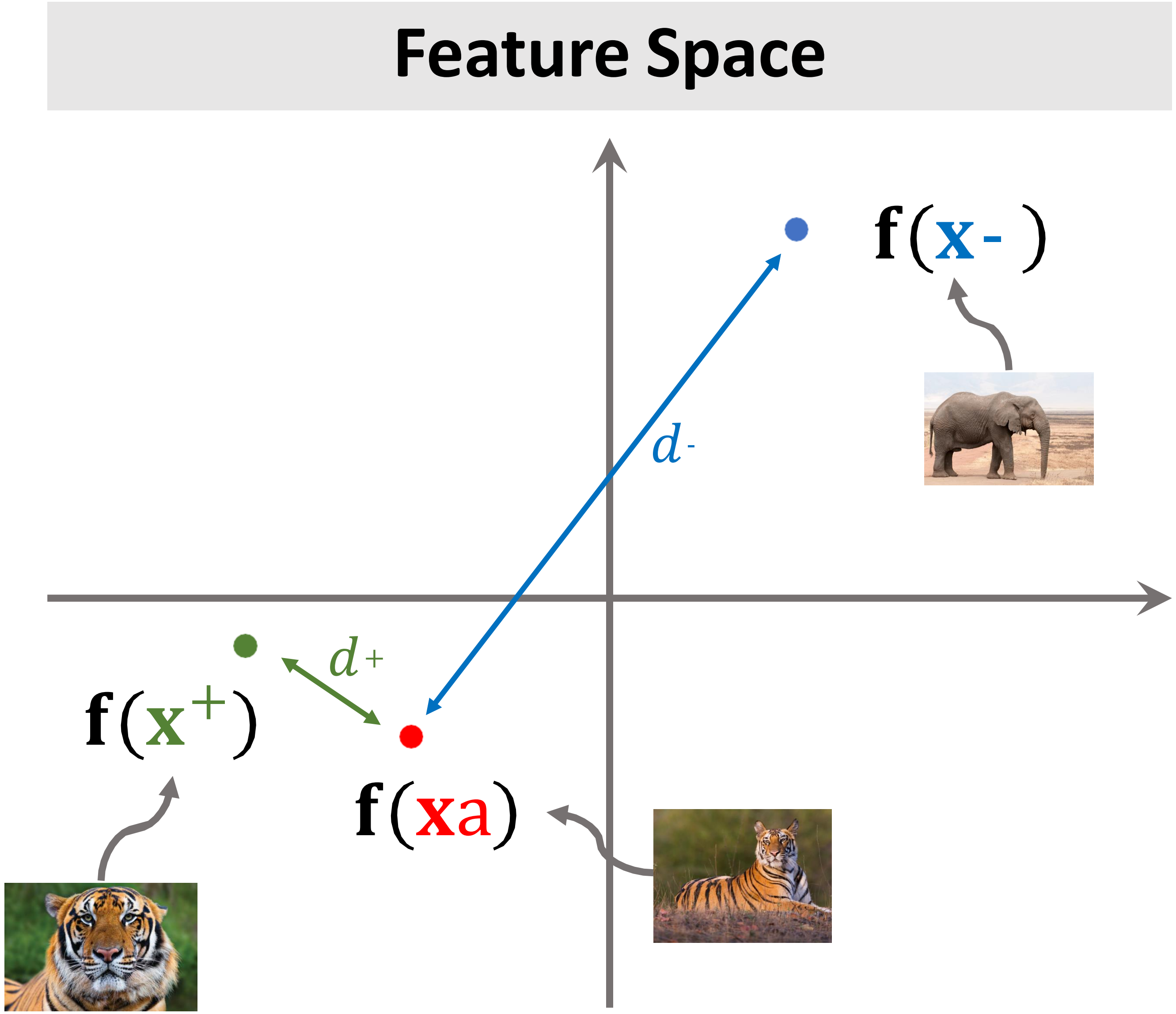
(positive)

xa

(anchor)

$x-$

(negative)



Triplet Loss

$\mathbf{x}+$

(positive)

$\mathbf{x}\mathrm{a}$

(anchor)

$\mathbf{x}-$

(negative)

- Encourage $d^+ = ||\mathbf{f}(\mathbf{x}+) - \mathbf{f}(\mathbf{x}\mathrm{a})||_2^2$ to be small.
- Encourage $d^- = ||\mathbf{f}(\mathbf{x}^\mathrm{a}) - \mathbf{f}(\mathbf{x}^-)||_2^2$ to be big.

Triplet Loss

$\mathbf{x}^{+}$

(positive)

$\mathbf{x}^{a}$

(anchor)

$\mathbf{x}^{-}$

(negative)

- Encourage $d^{+} = \|\mathbf{f}(\mathbf{x}^{+}) - \mathbf{f}(\mathbf{x}^{a})\|_2^2$ to be small.
- Encourage $d^{-} = \|\mathbf{f}(\mathbf{x}^{a}) - \mathbf{f}(\mathbf{x}^{-})\|_2^2$ to be big.
- If $d^{+} \geq d^{-} + \alpha$, then no loss. ($\alpha > 0$ is margin.)
- Otherwise, the loss is $d^{+} + \alpha - d^{-}$.

Triplet Loss

$\mathbf{x}+$

(positive)

$\mathbf{x}^{\mathrm{a}}$

(anchor)

$\mathbf{x}-$

(negative)

- Encourage $d^{+} = \left\|\mathbf{f}(\mathbf{x}+) - \mathbf{f}(\mathbf{x}\mathrm{a})\right\|_2^2$ to be small.
- Encourage $d^{-} = \left\|\mathbf{f}(\mathbf{x}\mathrm{a}) - \mathbf{f}(\mathbf{x}-)\right\|_2^2$ to be big.
- If $d^{+} \geq d^{-} + \alpha$, then no loss. ($\alpha > 0$ is margin.)
- Otherwise, the loss is $d^{+} + \alpha - d^{-}$.
- $\mathrm{Loss}(\mathbf{x}\mathrm{a}, \mathbf{x}+, \mathbf{x}-) = \max\{0,\ d^{+} + \alpha - d^{-}\}$.
- Update the CNN (function $\mathbf{f}$) to decrease the loss.

One-Shot Predictio

Query:

Support Set:

Fox	Squirrel	Rabbit	Hamster	Otter	Beaver
					

One-Shot Predicti

Query:

dist = 231

Fox	Squirrel	Rabbit	Hamster	Otter	Beaver
					

One-Shot Predictio

Query:

dist = 231

dist = 19

Fox

Squirrel

Rabbit

Hamster

Otter

Beaver

One-Shot Prediction

Query:

dist = 231	dist = 19	dist = 138	dist = 76	dist = 122	dist = 94
Fox	**Squirrel**	**Rabbit**	**Hamster**	**Otter**	**Beaver**






One-Shot Predictio

Query:

$\text{dist} = 231$

Fox

$\text{dist} = 19$

Squirrel

$\text{dist} = 138$

Rabbit

$\text{dist} = 76$

Hamster

$\text{dist} = 122$

Otter

$\text{dist} = 94$

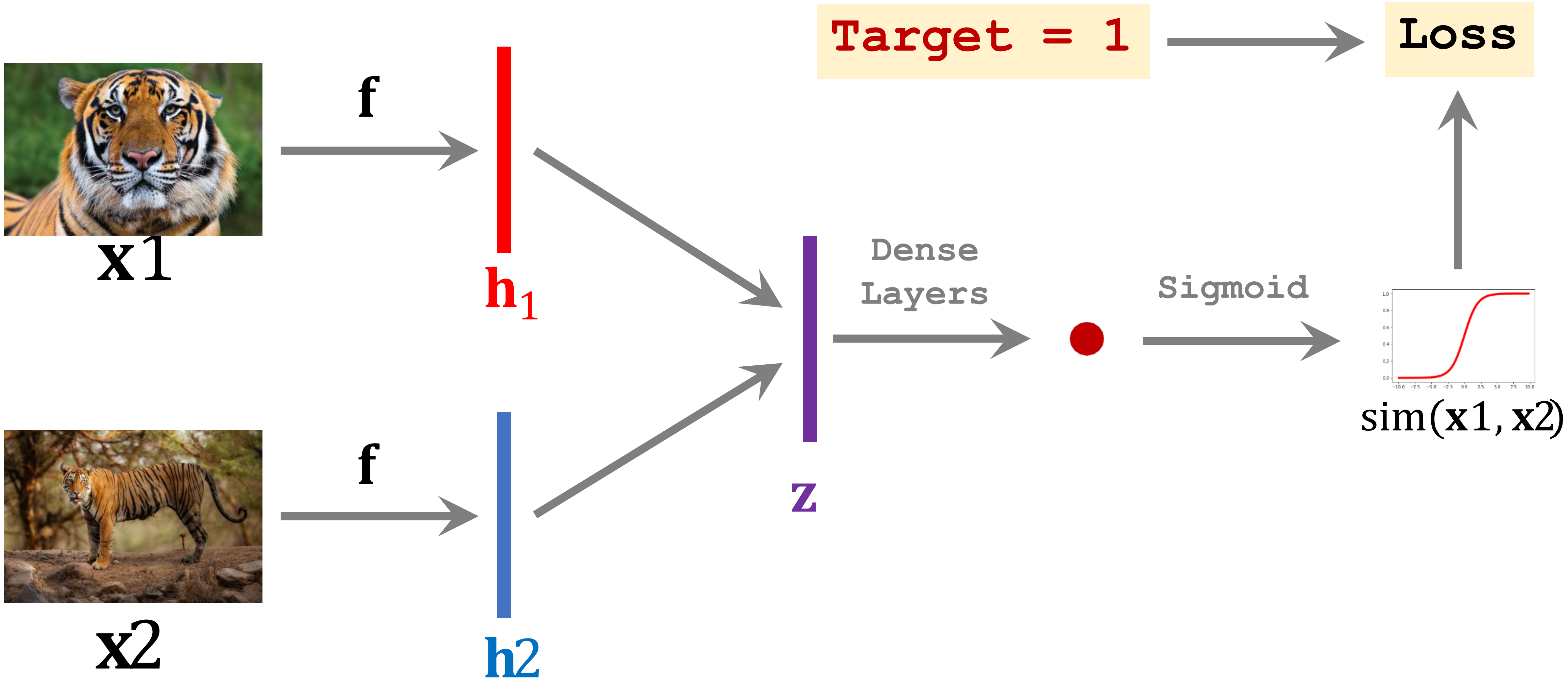
Beaver

Summary

Basic Idea of Few-Shot Learning

- Train a Siamese network on large-scale training set.
- Given a support set of k-way n-shot.
  - k-way means k classes.
  - n-shot means every class has n samples.
  - The training set does not contain the k classes.
- Given a query, predict its class.
  - Use the Siamese network to compute similarity or distance.

Siamese Network for Pairwise Similarity



Siamese Network with Triplet Loss

$\mathbf{x}+$
(positive)

$\mathbf{x}+ \xrightarrow{\mathbf{f}} \mathbf{f}(\mathbf{x}+)$

$\mathbf{xa}$
(anchor)

$\mathbf{xa} \xrightarrow{\mathbf{f}} \mathbf{f}(\mathbf{xa})$

$\mathbf{x}^-$
(negative)

$\mathbf{x}^- \xrightarrow{\mathbf{f}} \mathbf{f}(\mathbf{x}-)$

$$d^+ = ||\mathbf{f}(\mathbf{x}+) - \mathbf{f}(\mathbf{xa})||_2^2$$

$$d^- = ||\mathbf{f}(\mathbf{xa}) - \mathbf{f}(\mathbf{x}-)||_2^2$$