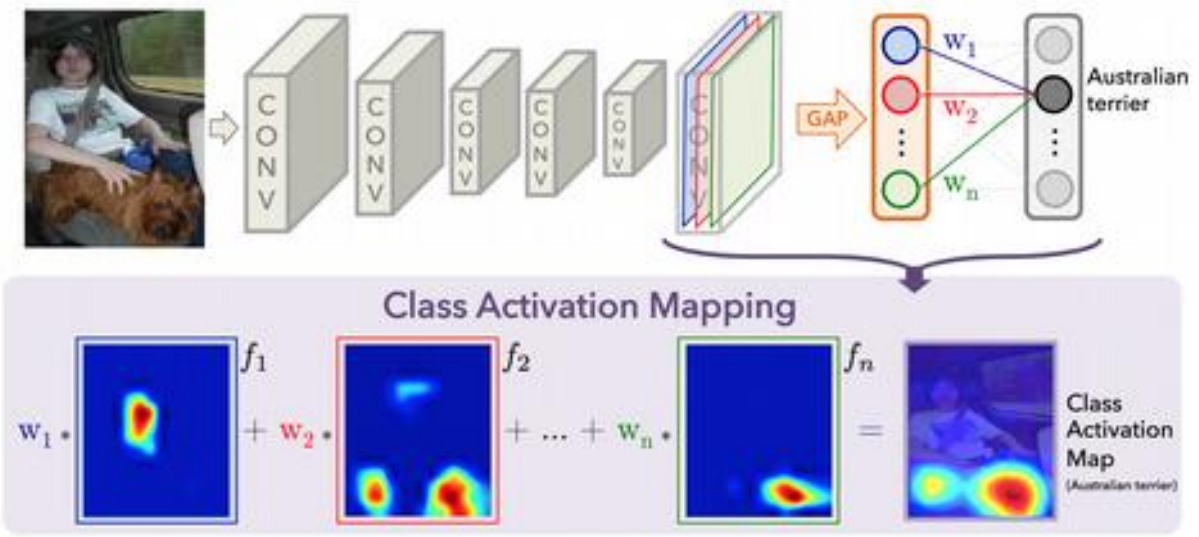
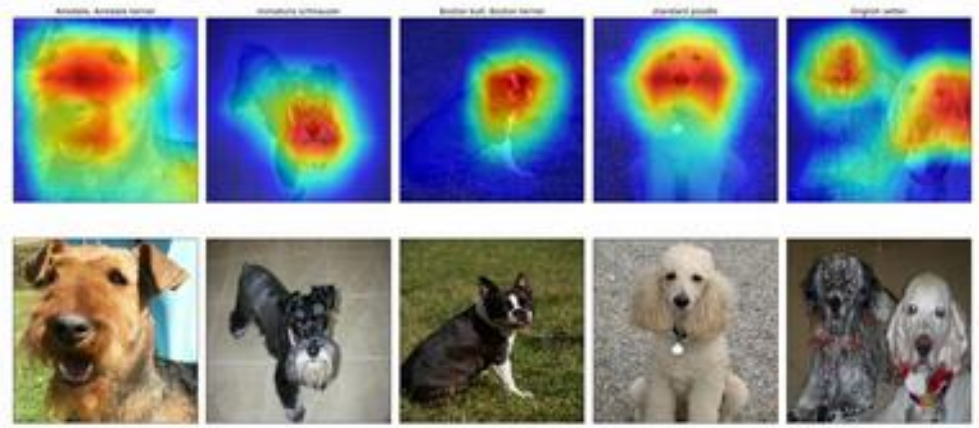
Siamese & Triplet Networks
Fully Convolutional Neural Nets
Class Activation Mapping

José M. Saavedra R.

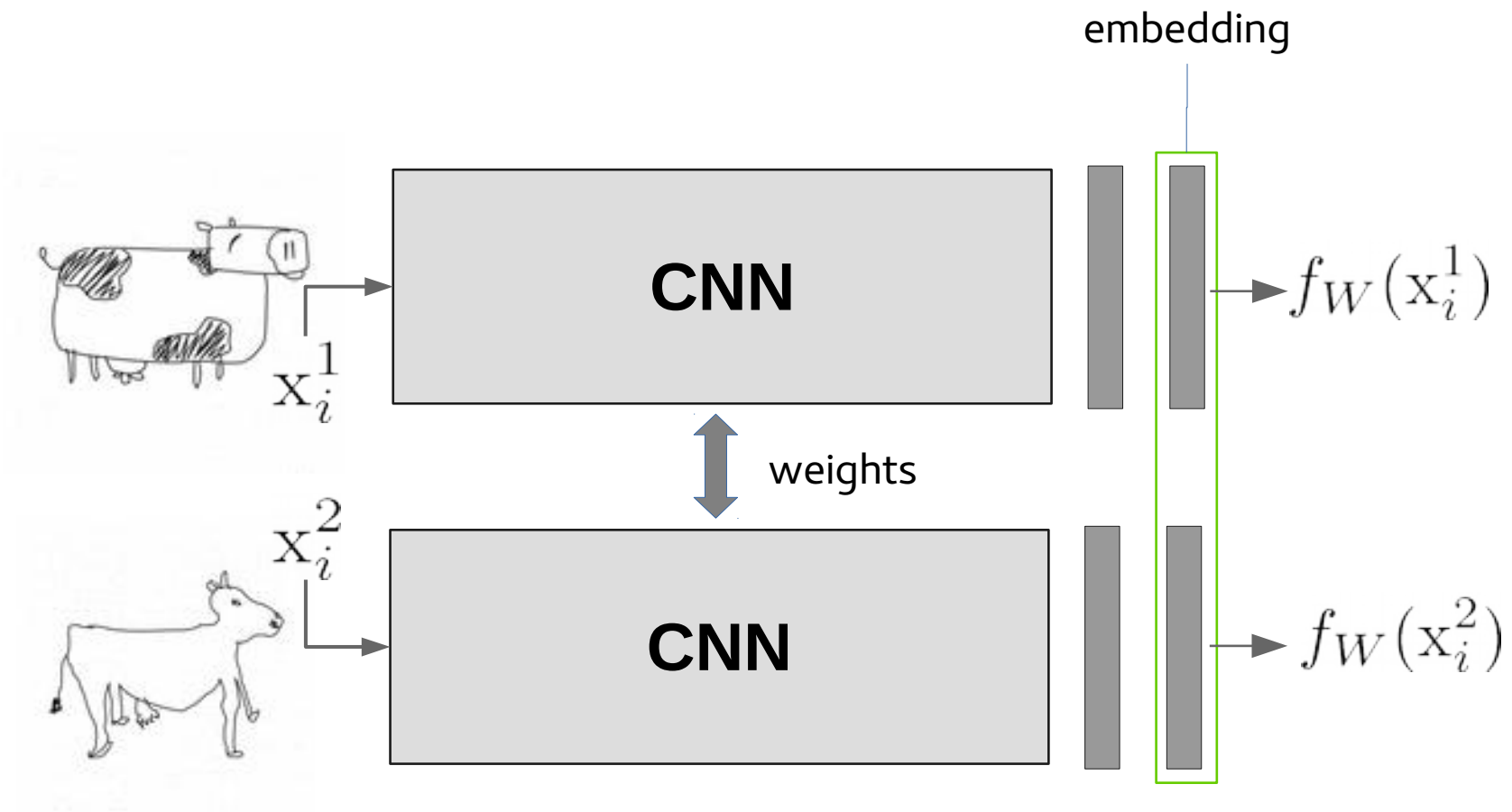
Class Activation Mapping



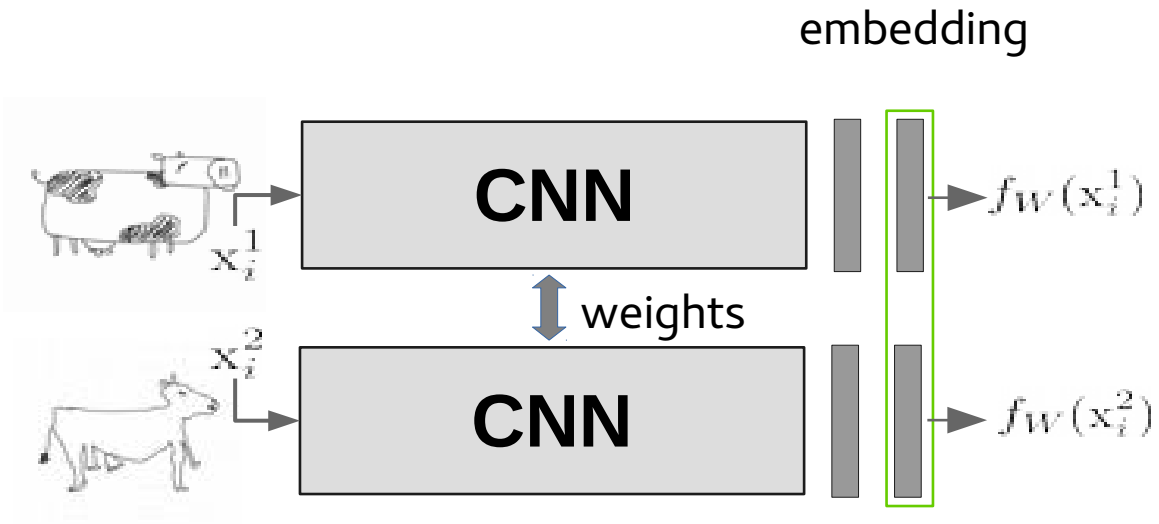
$$w_1 \cdot f_1 + w_2 \cdot f_2 + \ldots + w_{2048} \cdot f_{2048}$$



Siamese Network



Siamese Network

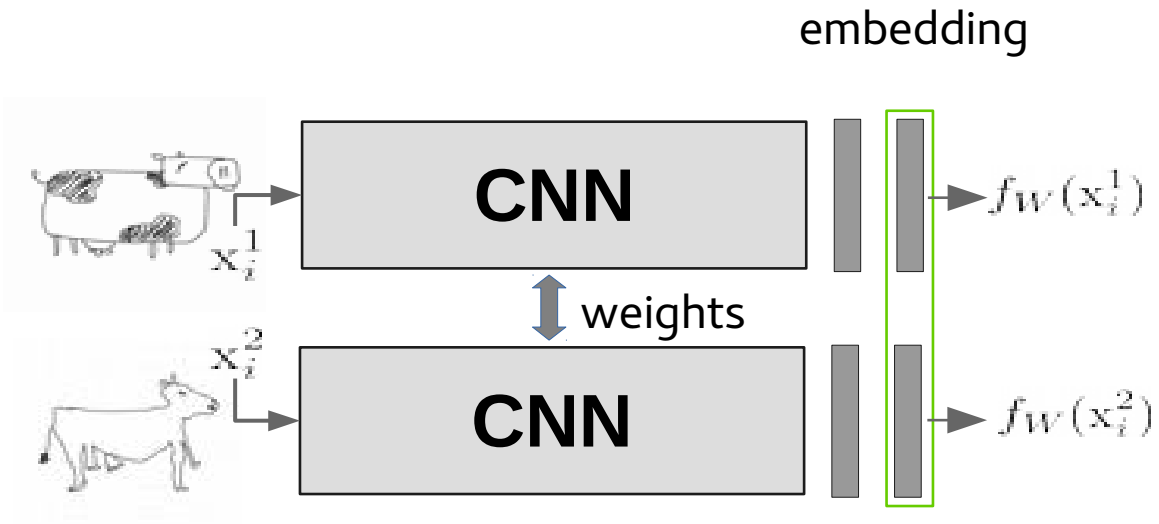


- La entrada se realiza de a pares (x_i^1, x_i^2)
- Cada para tiene asociado una etiqueta [1: similares, 0 : diferentes]

$$(x_i^1, x_i^2) \rightarrow y_i$$

- A diferencia de la clasificación, el modelo basado en *siameses* se enfoca en buscar *embeddings* que maximicen la semejanza entre objetos similares.

Siamese Network

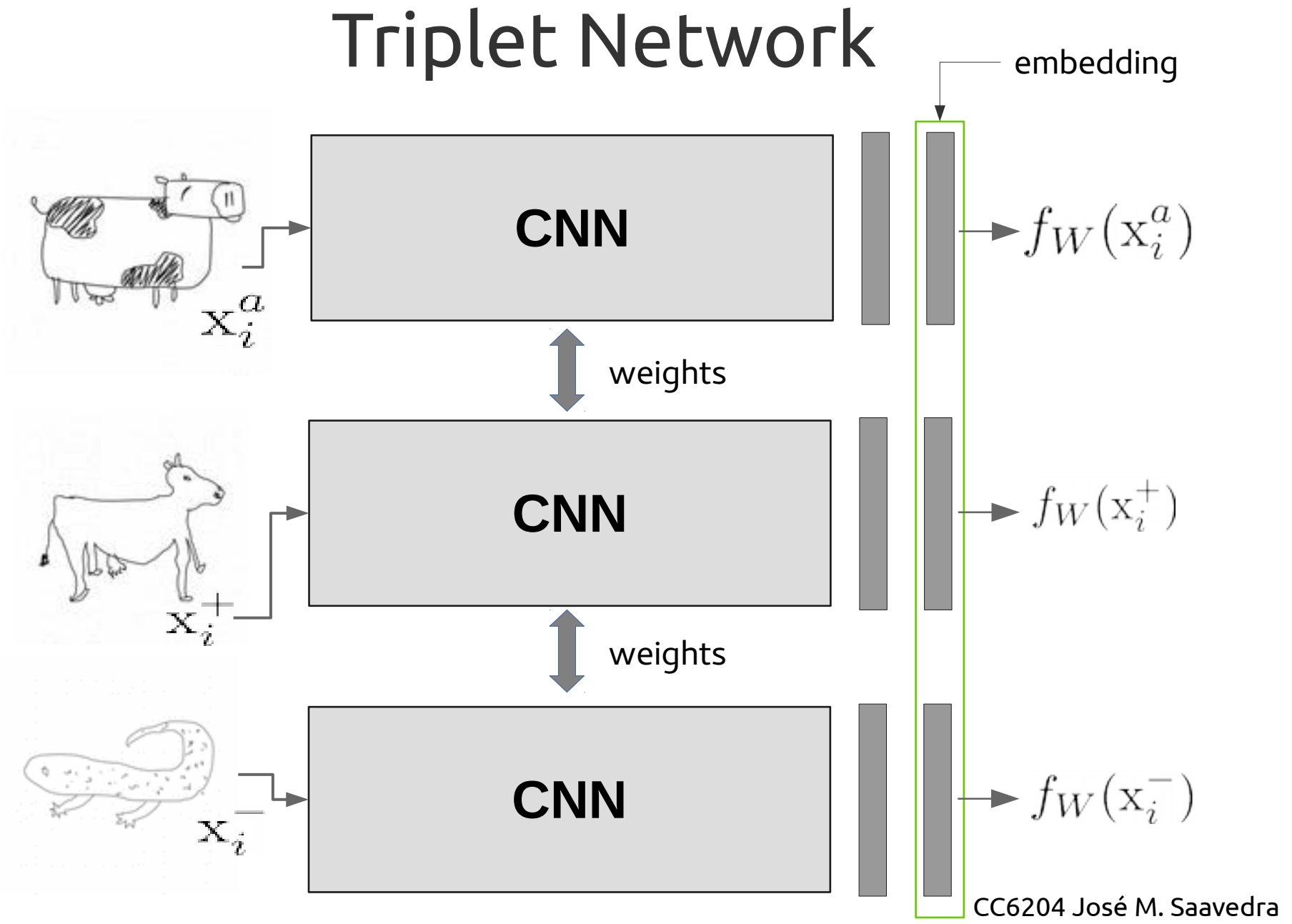


$$D_w^i = L_2(f_w(x_i^1), f_w(x_i^2))$$

$$L_i = y_i {D_w^i}^2 + (1 - y_i)\{max(0, \lambda - D_w^i)\}^2$$

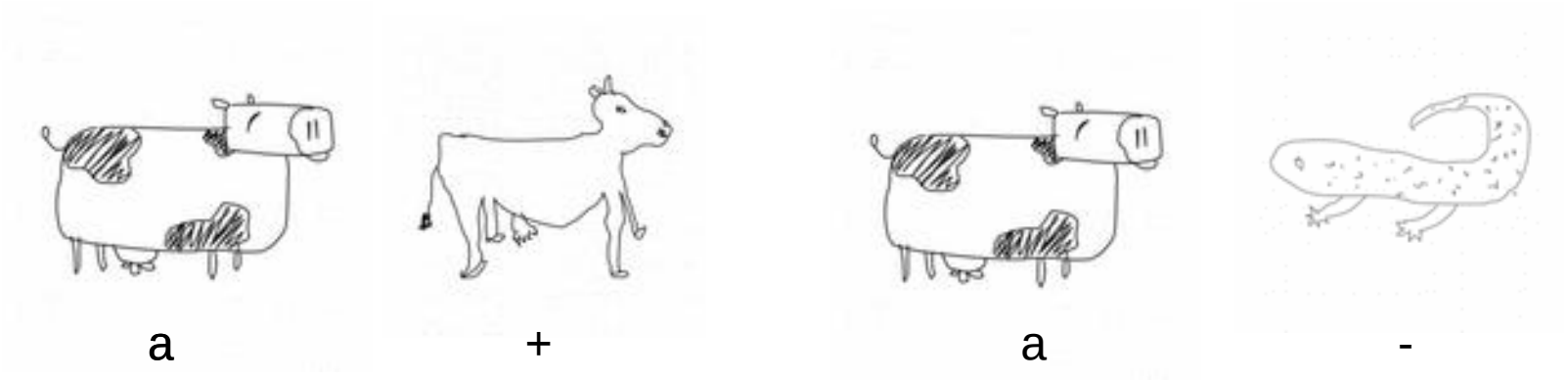
λ : margen

Triplet Network



Triplet Network

Triplet Loss



$$D_W^i(a,+)$$

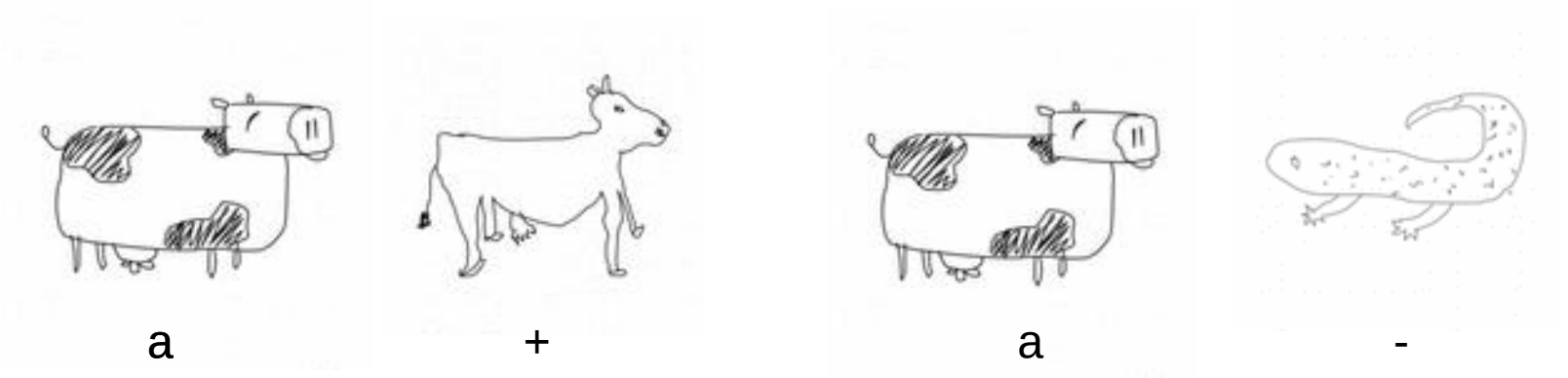
$$D_W^i(a,-)$$

$$D_W^i(a,+) + \lambda < D_W^i(a,-)$$

[1] O. M. Parkhi, A. Vedaldi, A. ZissermanDeep Face Recognition British Machine Vision Conference, 2015

Triplet Network

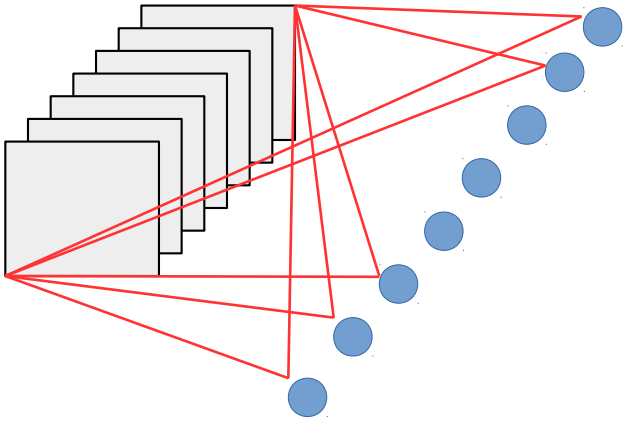
Triplet Loss



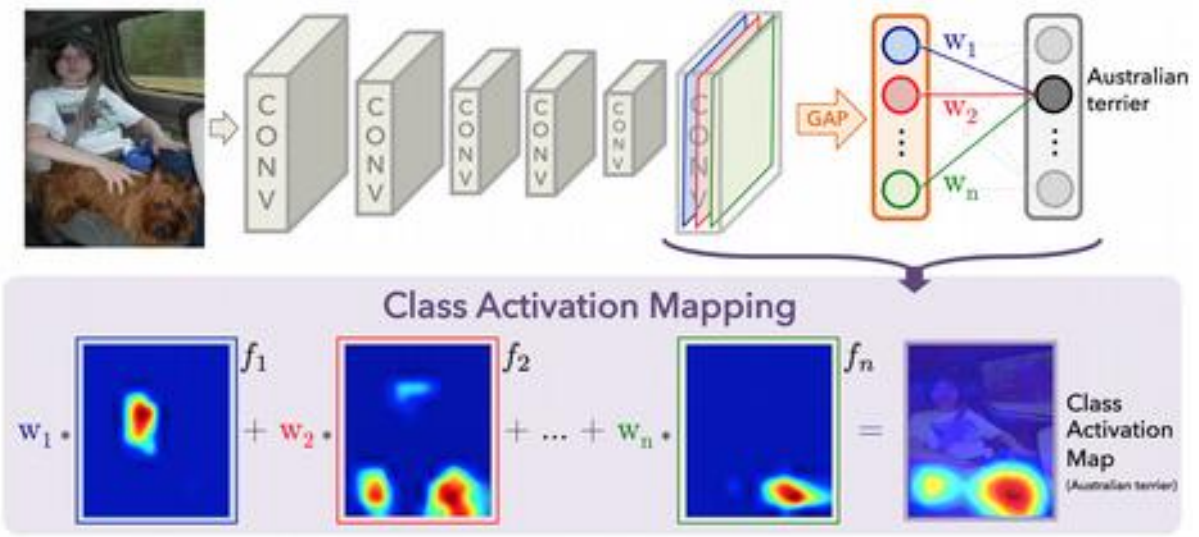
$$L_i = max(D_W^i(a,+) - D_W^i(a,-) + \lambda, 0)$$

Fully Convolutional Neural Networks

- Todas las capas son convolucionales
- Capas fully-connected se convierten en convolucionales ¿Cómo?



Class Activation Mapping



$$w_1 \cdot f_1 + w_2 \cdot f_2 + \ldots + w_{2048} \cdot f_{2048}$$