

Aprendizaje Autosupervisado

Intro al Aprendizaje Profundo

LCD IIMAS UNAM

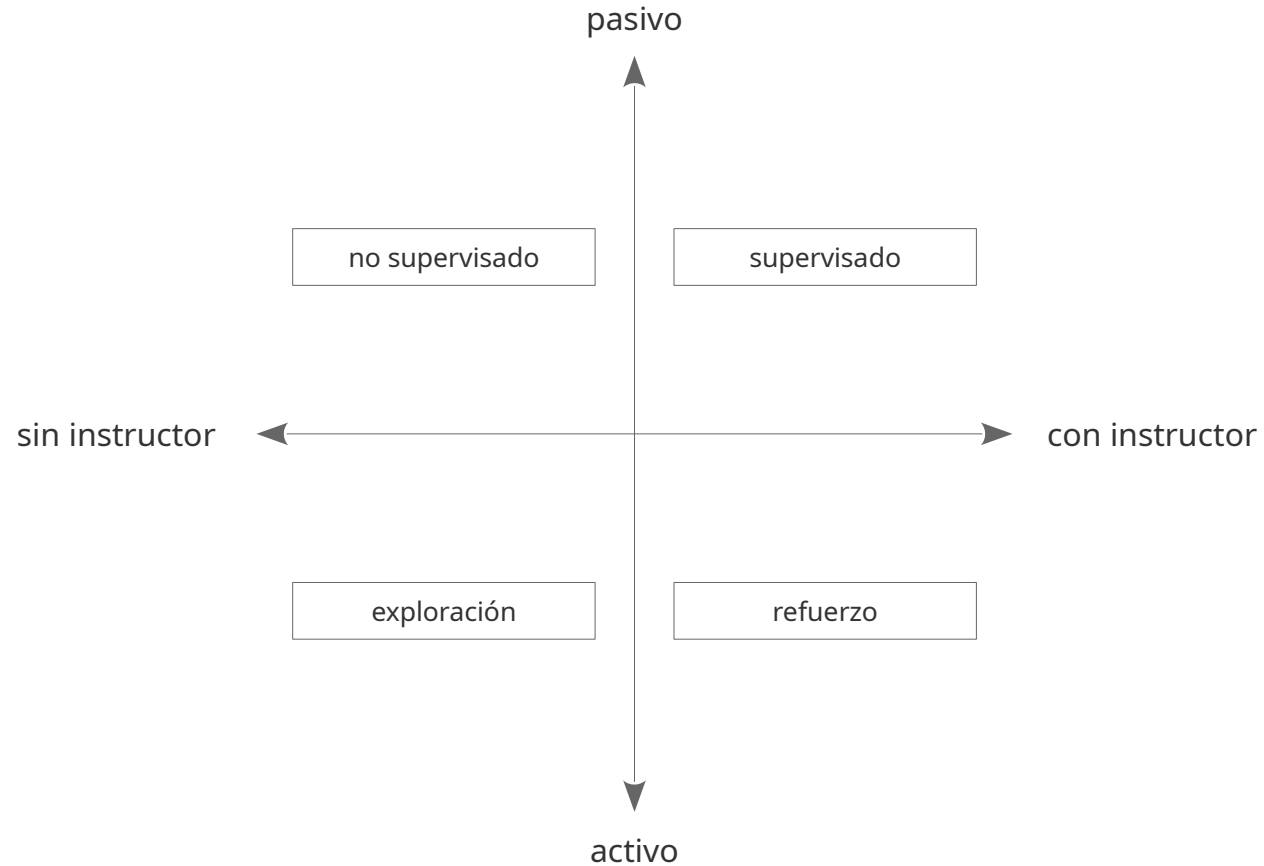
Profs. Bere & Ricardo Montalvo Lezama

Jun 2021

Contenido

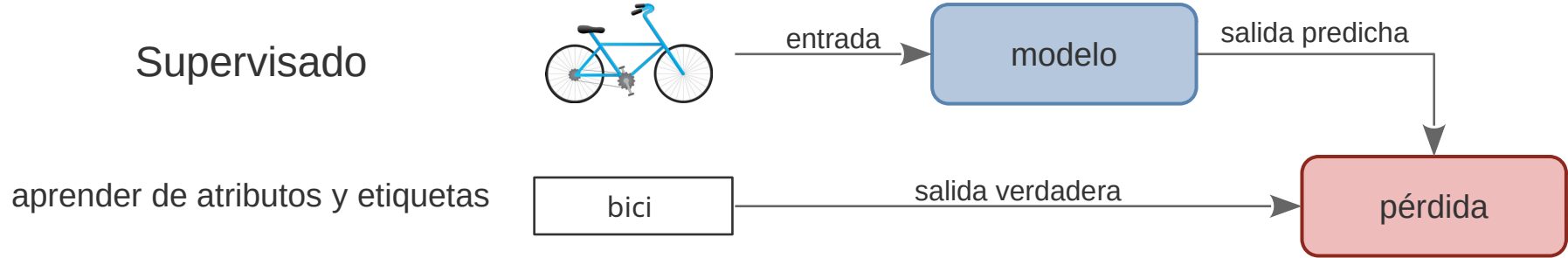
- Aprendizaje autosupervisado
 - Introducción
 - Ejemplo
 - Tareas pretexto
 - Estado del arte

Tipos de aprendizaje



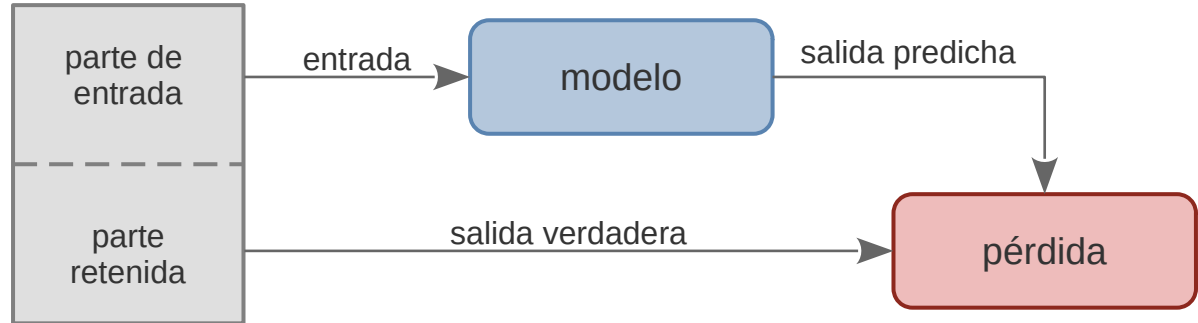
Autosupervisión

Supervisado

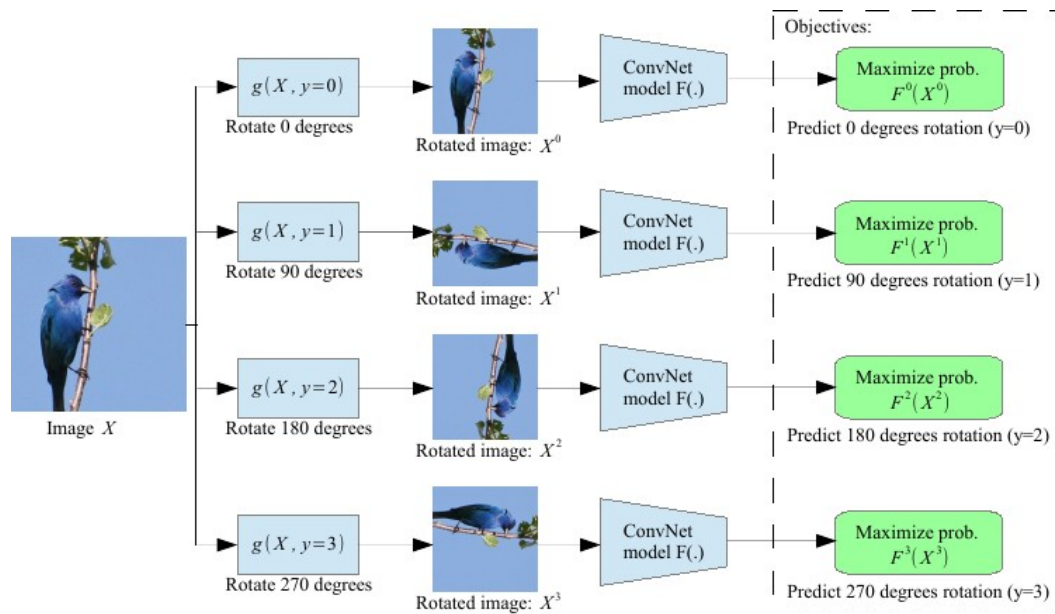
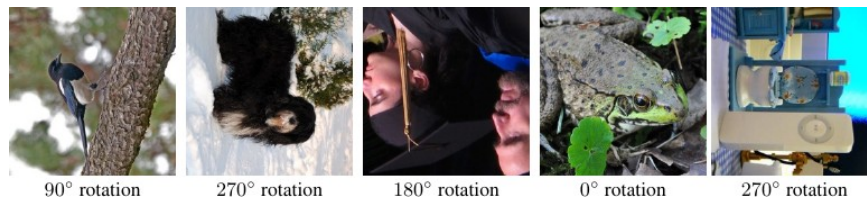


Autosupervisado

aprender solo de atributos

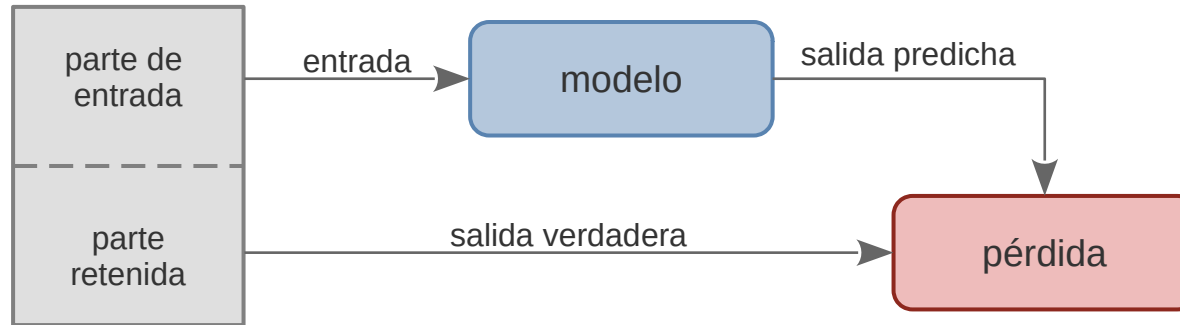


¿Cuál es la rotación de la imagen?



¿Qué es el autosupervisión?

- Aprender de forma supervisada en datos sin etiquetas.
- Tareas pretexto: retener una parte de la entrada y aprender a predecirla.
- La tarea pretexto determina el tipo de representaciones que se aprenden.



¿Por qué es importante?

- Etiquetar conjuntos de datos es costoso.
 - Contratar humanos, etiquetado manual, implementar tuberías y GUIs.
- Una buena supervision puede ser infactible: medicina.
- Explotar cantidades masivas de datos no etiquetados.
- Aprender representaciones iguales o mejores que las supervisadas.



Pierre Sermanet
@psermanet

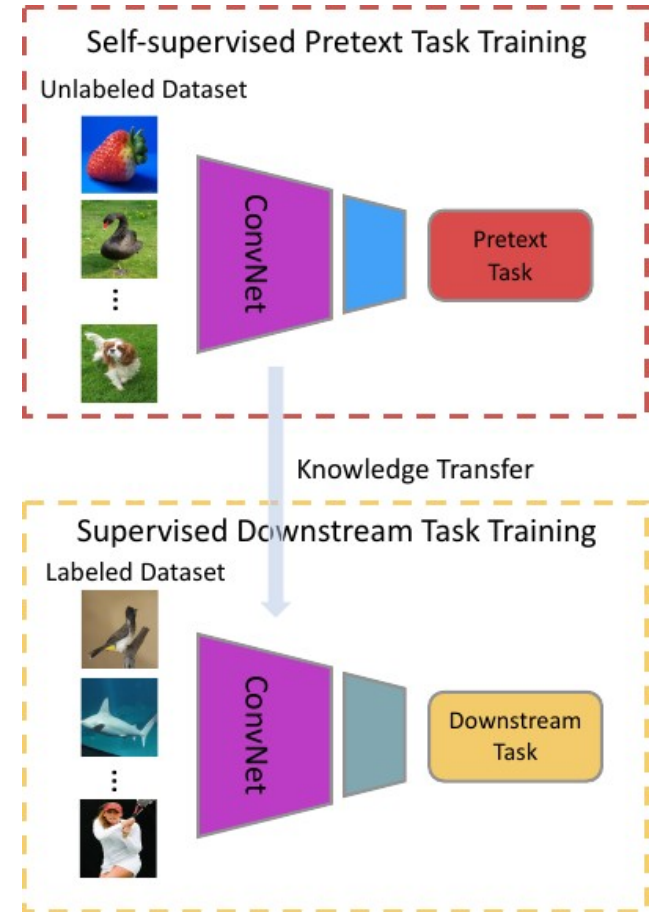


Give a robot a label and you feed it for a second;
teach a robot to label and you feed it for a lifetime.

3:37 PM · 3 déc. 2018 · [Twitter for Android](#)

Escenarios de uso

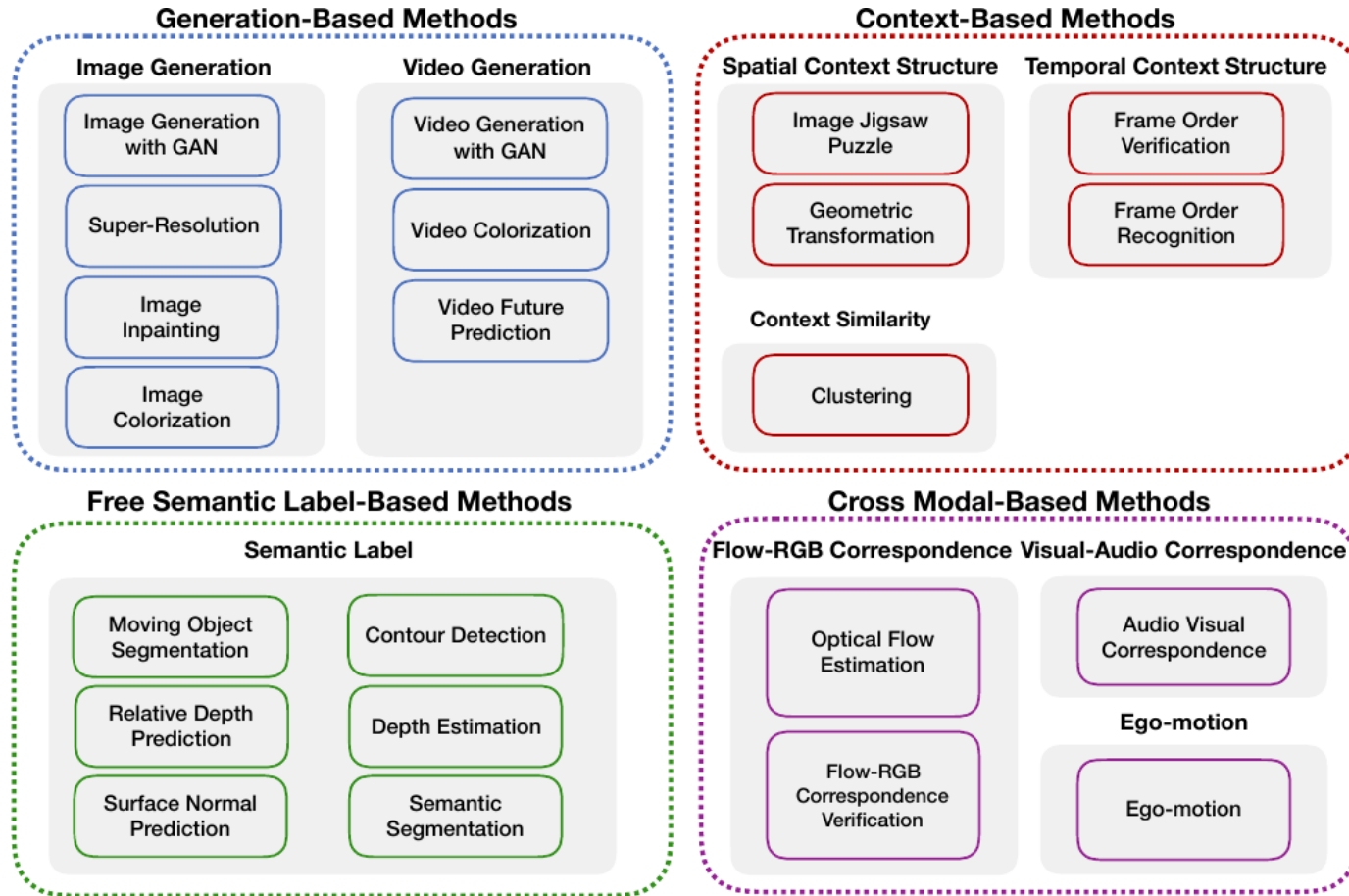
- Tareas de naturaleza autosupervisada
- Preentrenamiento y transferencia de conocimiento





¡tiempo de programar!
5a_auto_rot.ipynb

Tareas pretexto en VC



Generación de Imágenes

Restaurar colores ¹

GT



ChromaGAN



w/o Class



Chroma Net.



Input (a)



Shift-net (b)



Contextual Attention (c)



Our Result (d)



Ground Truth (e)



Restaurar parches ²

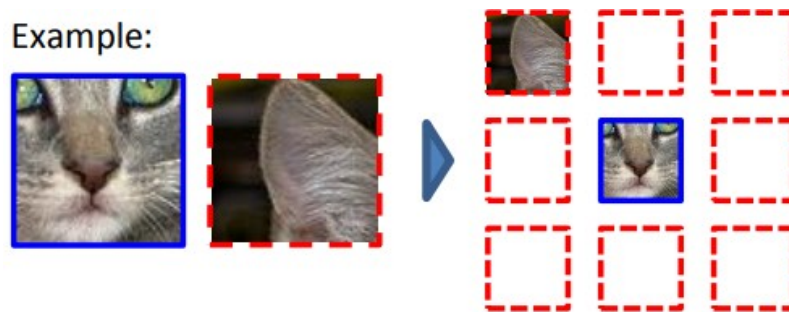
1. ChromaGAN: Adversarial Picture Colorization with Semantic Class Distribution. 2020.

2. Coherent Semantic Attention for Image Inpainting. 2019.

Estructura de contexto espacial

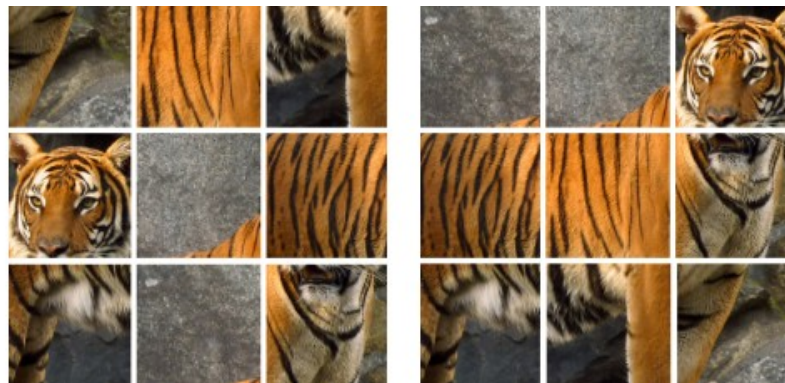
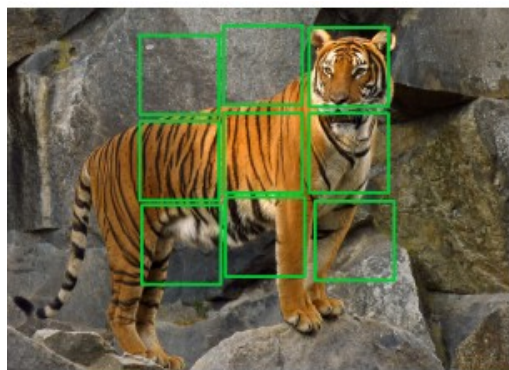
¿Cuál es la posición
relativa de los parches? ¹

Example:



Question 1:

Question 2:

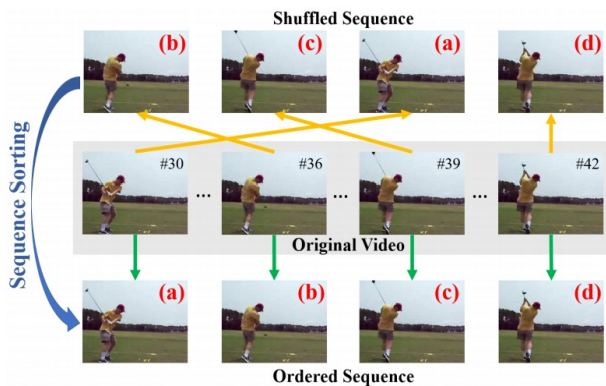
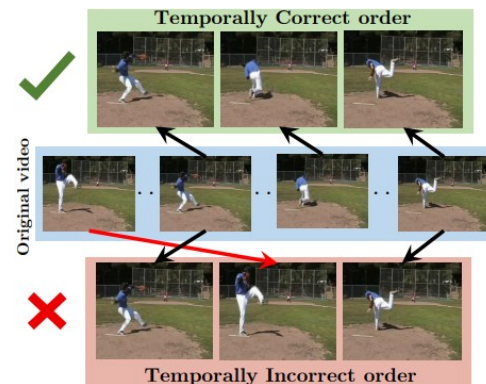


¿Cuál es el orden
de los parches? ²

1. Doersch et al. Unsupervised Visual Representation Learning by Context Prediction. 2015.
2. Kim et al. Unsupervised Learning of Visual Representations by Solving Jigsaw Puzzles. 2016.

Estructura de contexto temporal

¿El orden de los cuadros es correcto? ¹



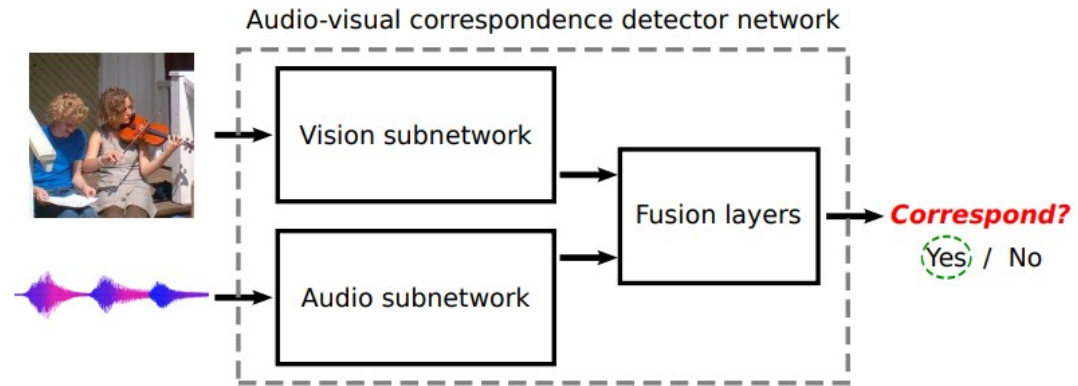
¿Cuál es el orden de los cuadros? ²

1. Shuffle and Learn: Unsupervised Learning using Temporal Order Verification. 2016.

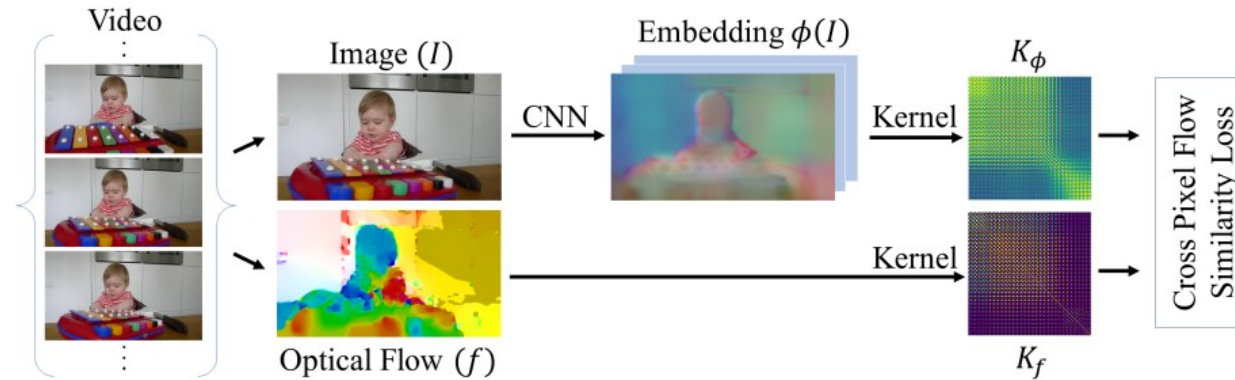
2. Unsupervised Representation Learning by Sorting Sequences. 2017.

Correspondencia de modalidad

Audio y video¹



Flujo óptico y video²

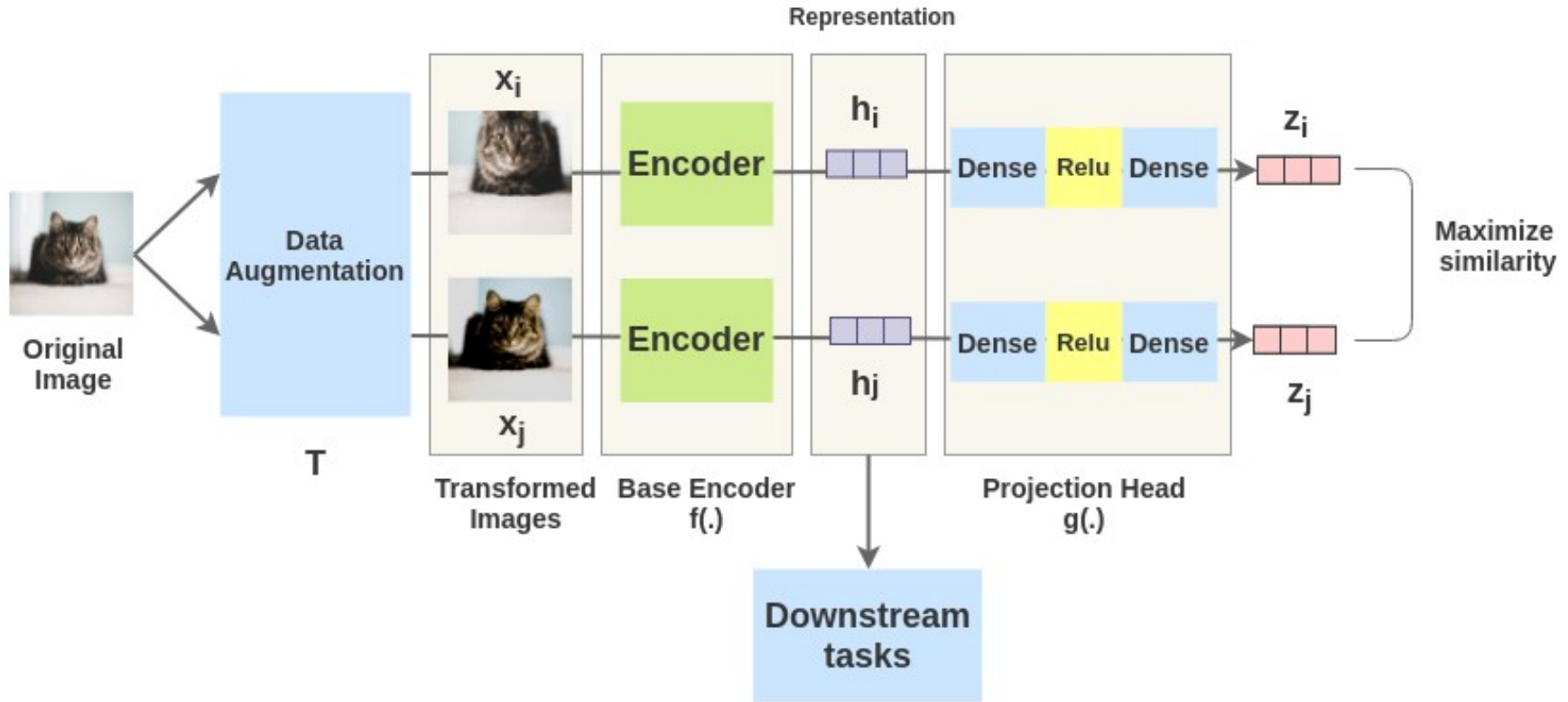


1. Look, Listen and Learn. 2017.

2. Cross Pixel Optical Flow Similarity for Self-Supervised Learnings. 2018.

SimCLR

SimCLR Framework



Aprendizaje contrastivo

Match the correct animal

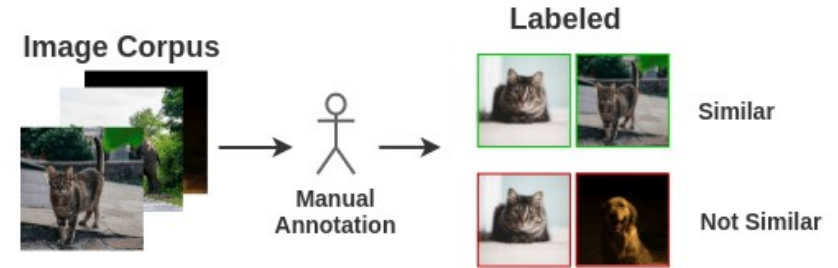


¿Cómo podemos generar la supervisión?

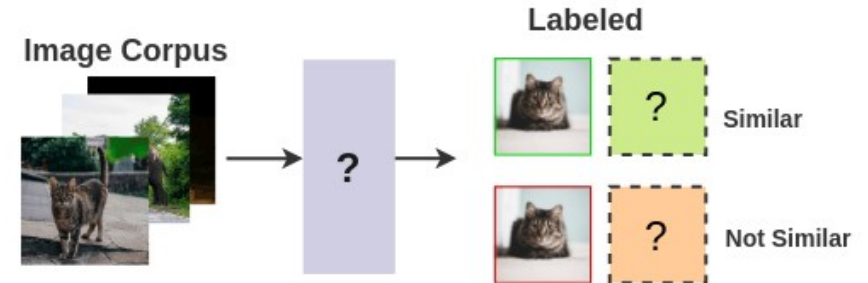
Need similar and different examples



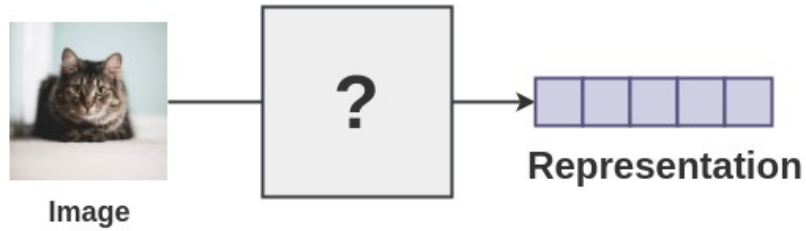
Supervised Approach



How can we automatically generate pairs?



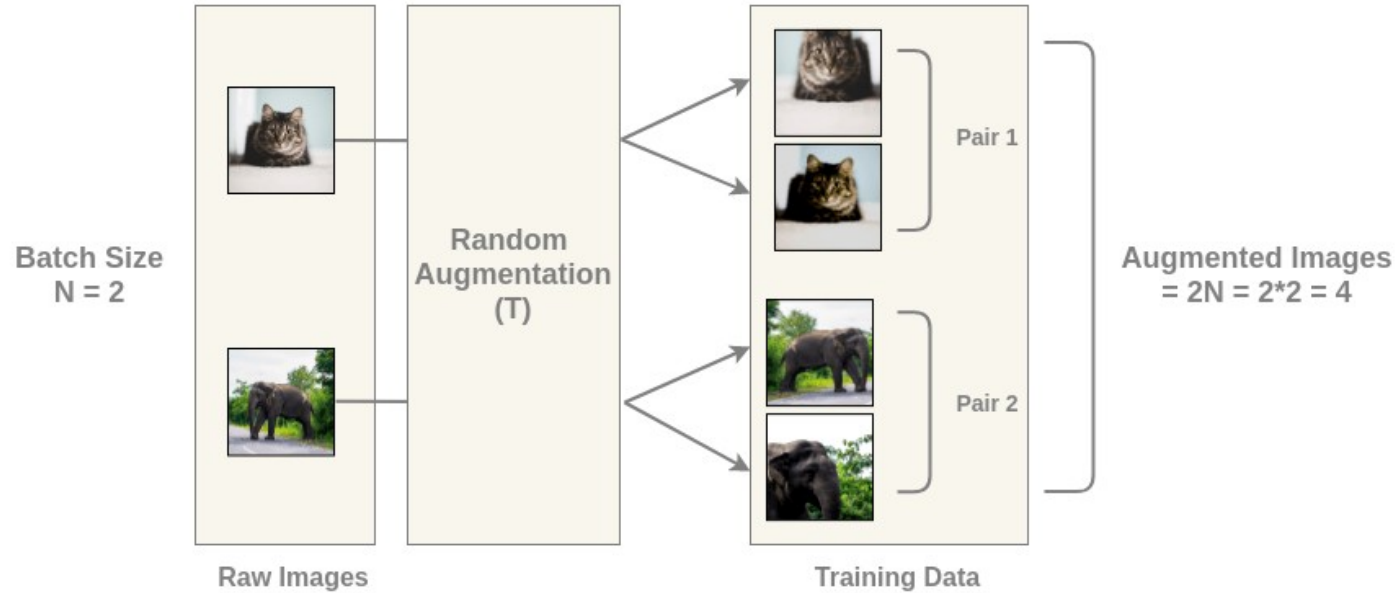
Representaciones y similitud



similarity(, )

Similitud por aumentado

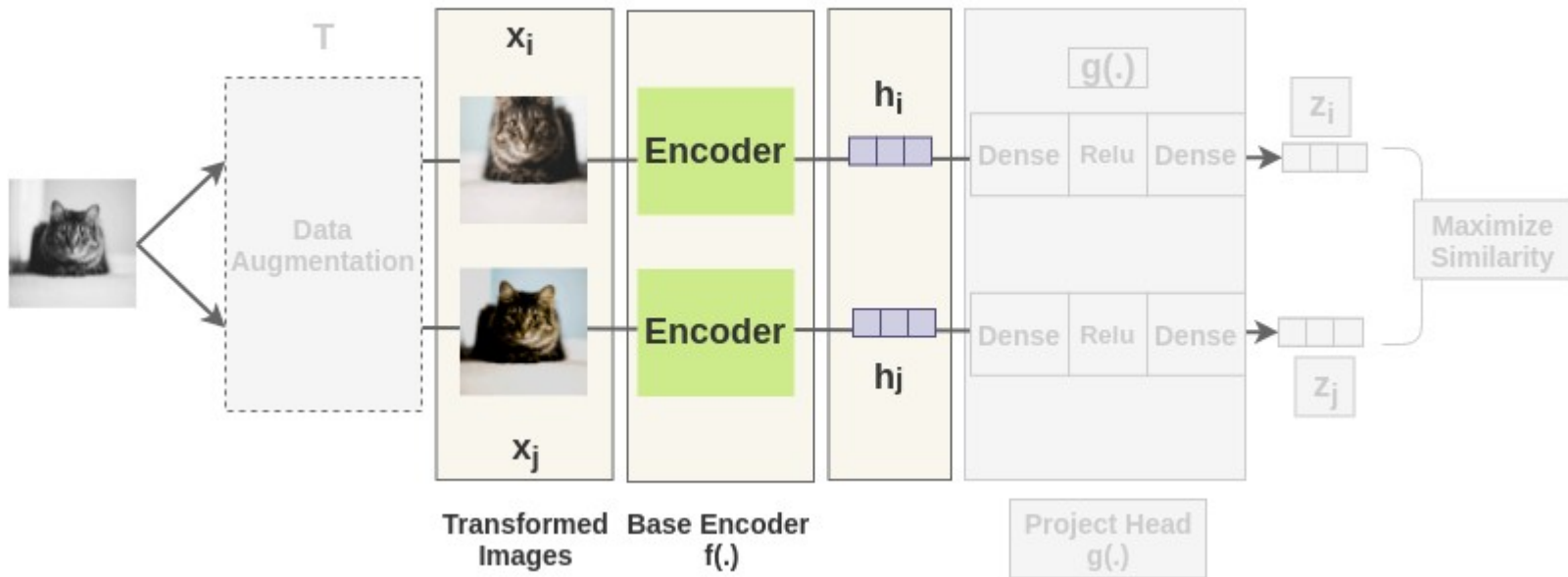
Preparing similar pairs in a batch



transformación = random (crop + flip + color jitter + grayscale)

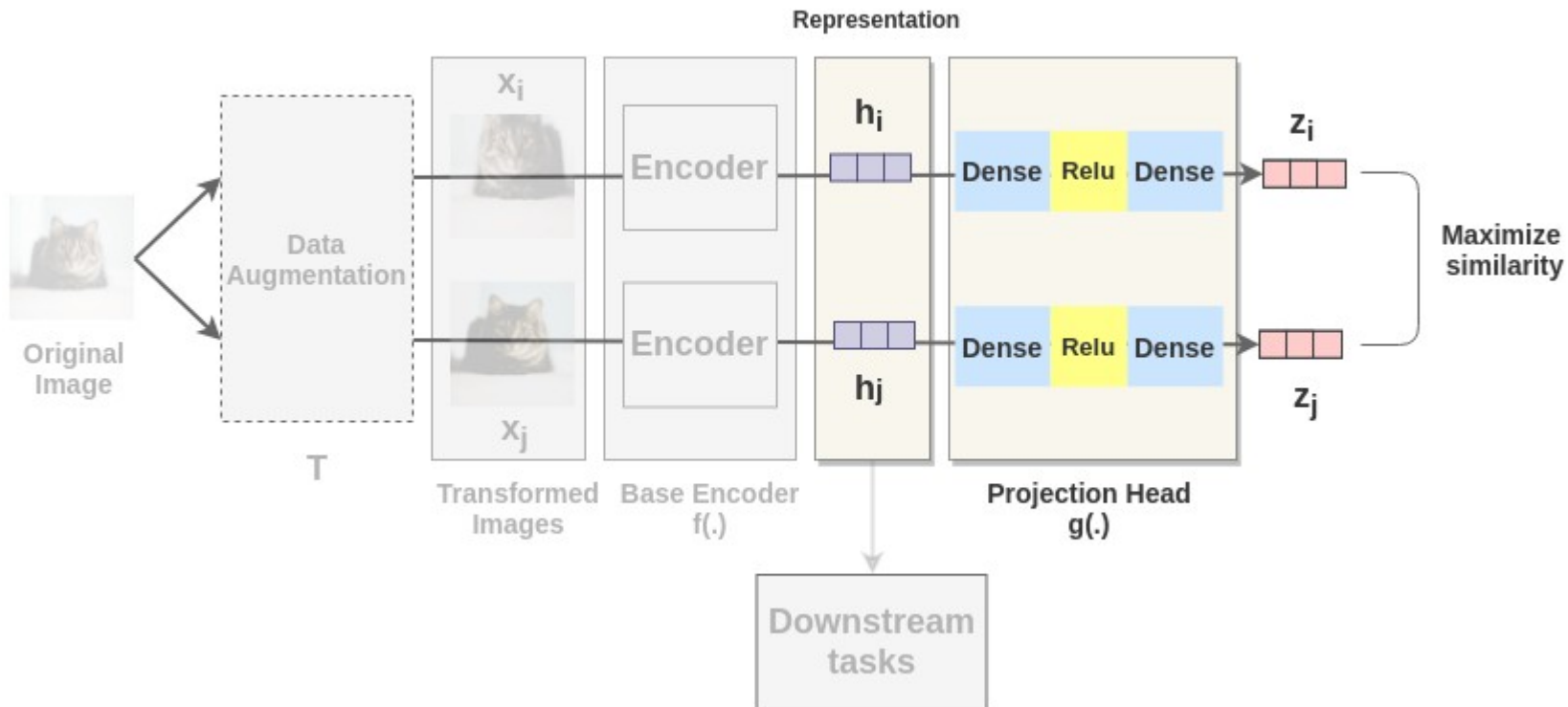
Cómputo de representaciones

Encoder Component of Framework



Proyección a vectores de comparación

Projection Head Component



Cálculo de similitud

Similarity Calculation of Augmented Images

$$\text{similarity}(x_i, x_j) = \text{cosine similarity}(z_i, z_j)$$

$$s_{i,j} = \frac{z_i^T z_j}{(\tau ||z_i|| ||z_j||)}$$

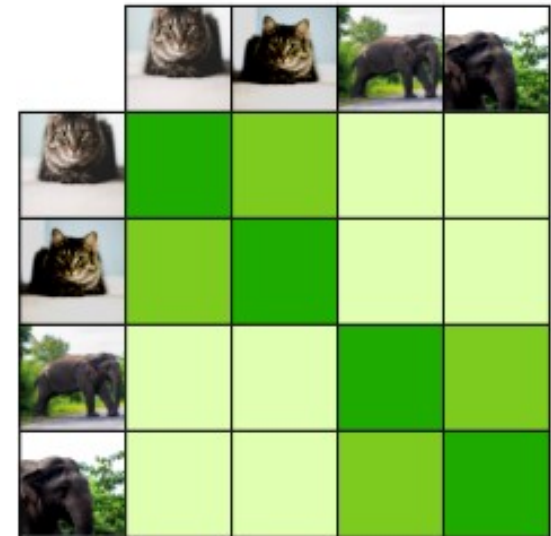
-1: opuesto idéntico

1: idéntico

0: distinto (ortogonal)

otros: valores intermedios de similitud

Pairwise cosine similarity



Clasificador de similaridad

Softmax =

Diagram illustrating the similarity classifier process:

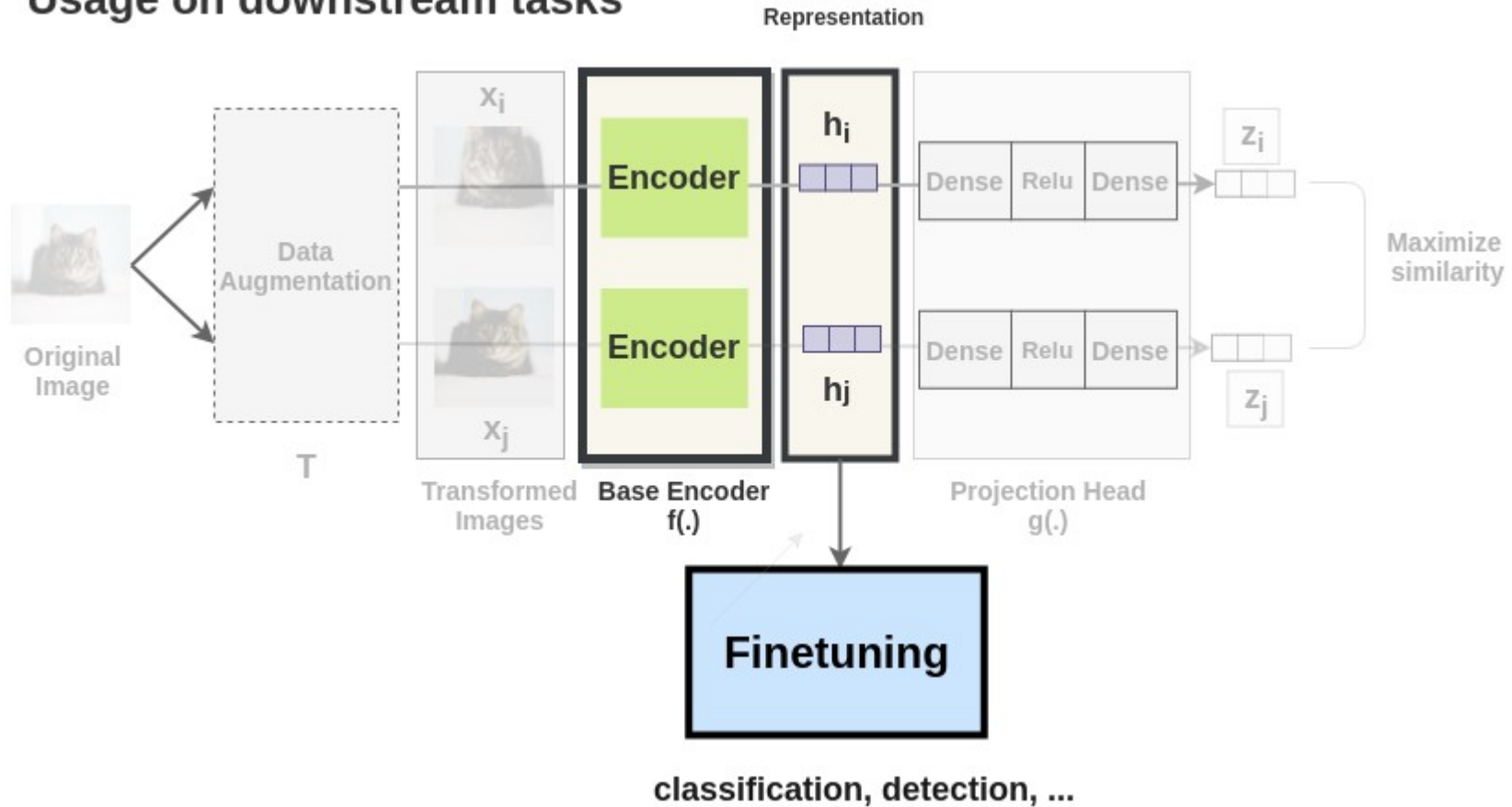
Pair 1 (Two images of cats) is input to the classifier.

The classifier outputs the similarity score for the pair, represented as:

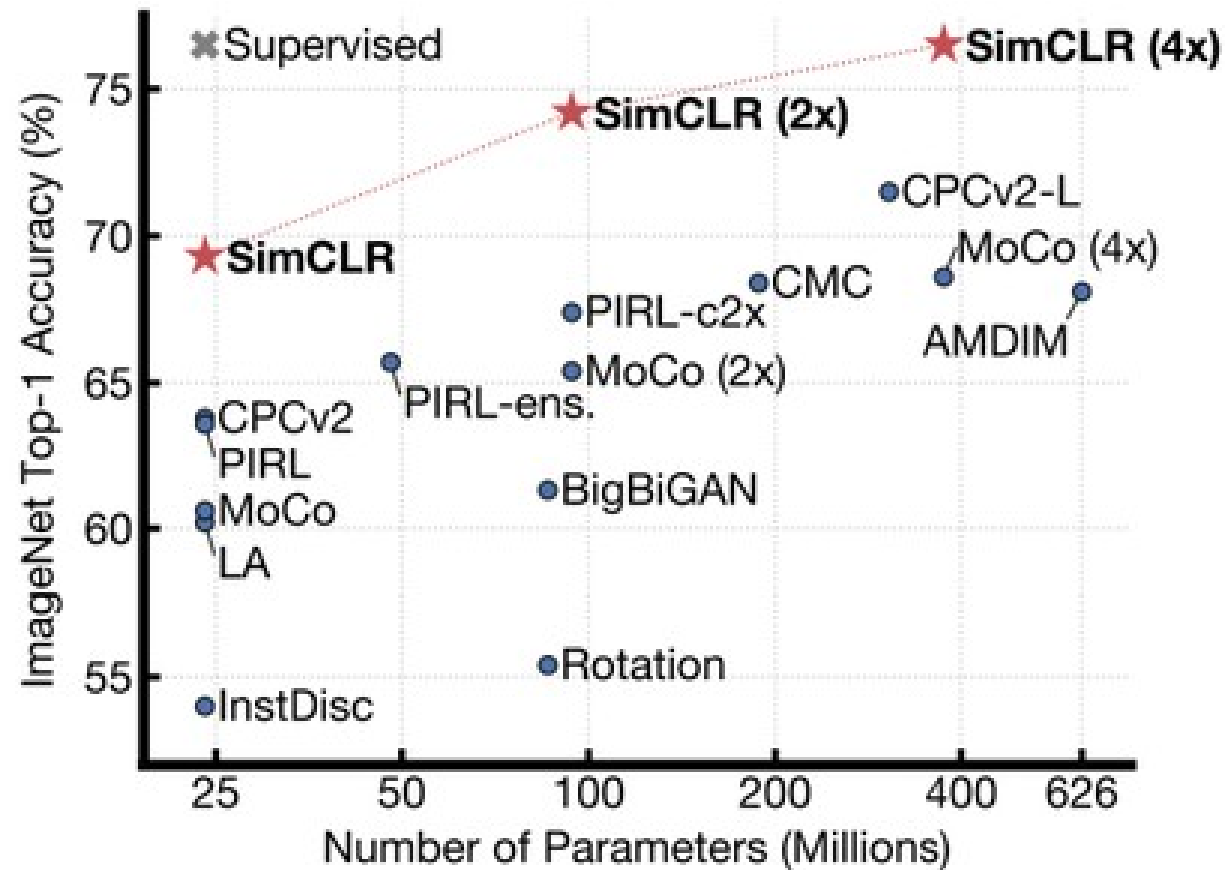
$$\frac{e^{\text{similarity}(\text{cat}_1, \text{cat}_2)}}{e^{\text{similarity}(\text{cat}_1, \text{cat}_2)} + e^{\text{similarity}(\text{cat}_1, \text{elephant}_1)} + e^{\text{similarity}(\text{cat}_1, \text{elephant}_2)}}$$

SimCLR

Usage on downstream tasks



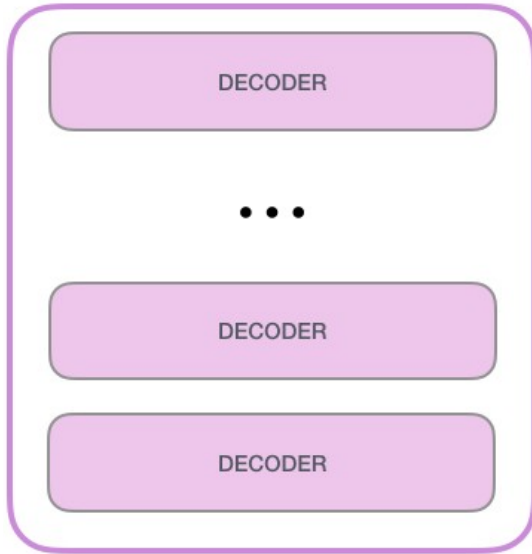
SimCLR



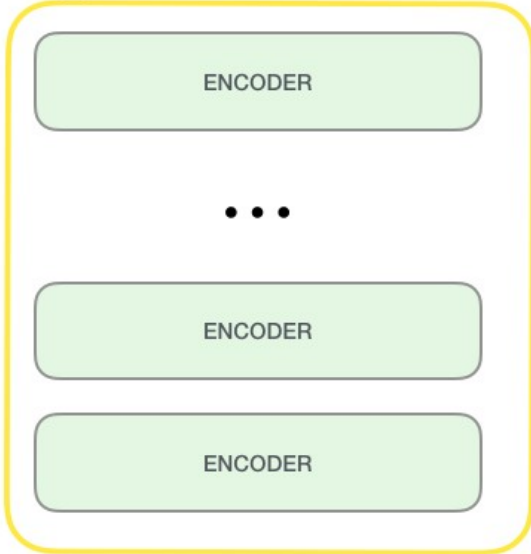
Modelos PLN



GPT-2



BERT



TRANSFORMER XL

