

Una Introducción al Aprendizaje Autosupervisado

Ricardo & Bere Montalvo Lezama

github.com/richardtml/riiaa-20-aa

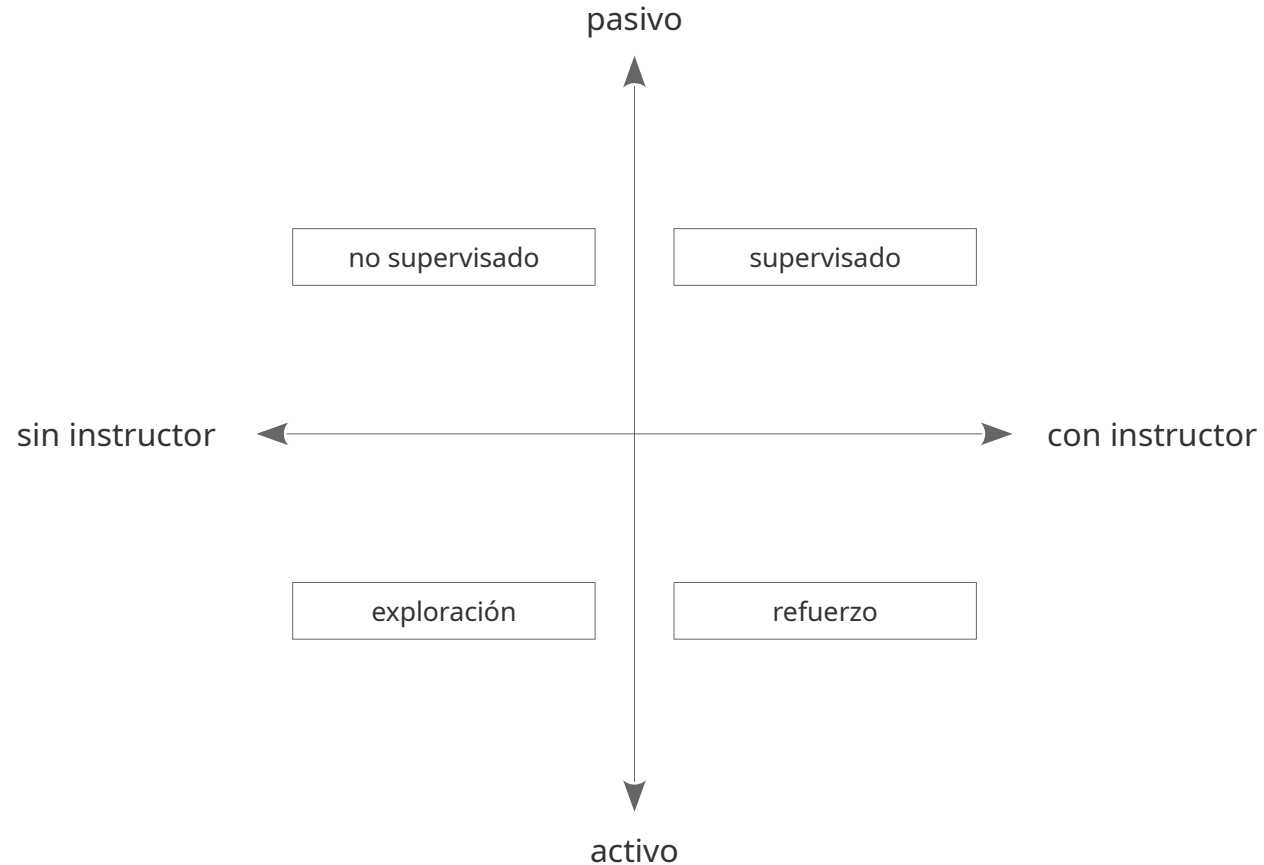


Agosto 2020

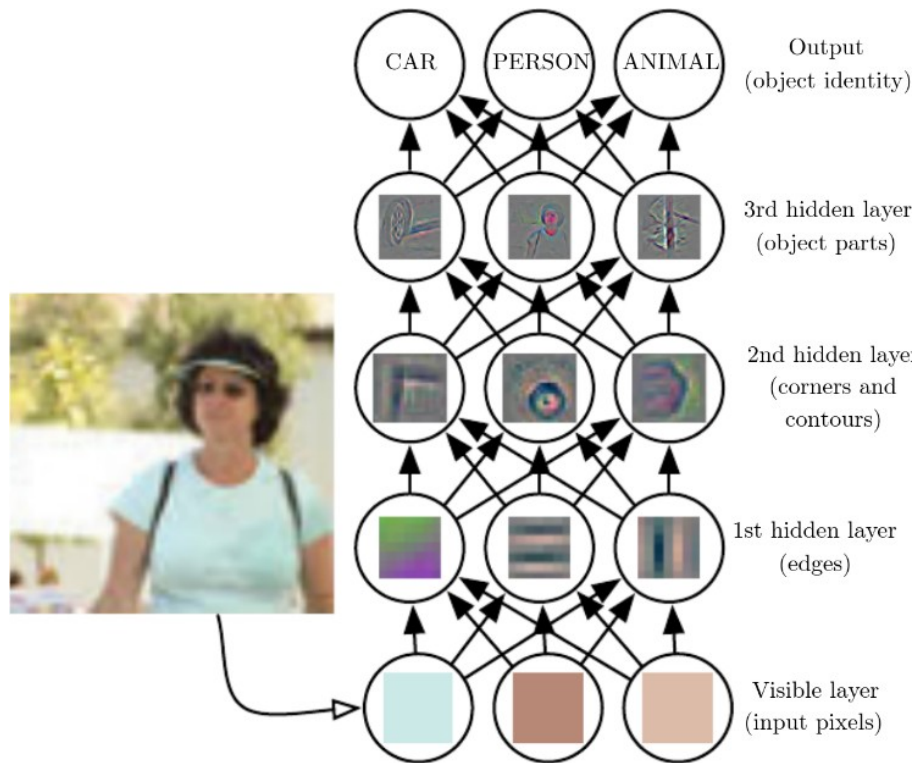
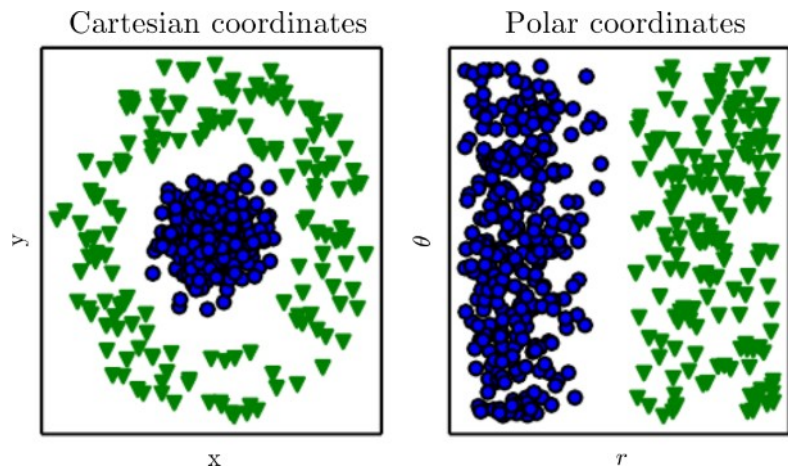
En este taller

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

Tipos de aprendizaje

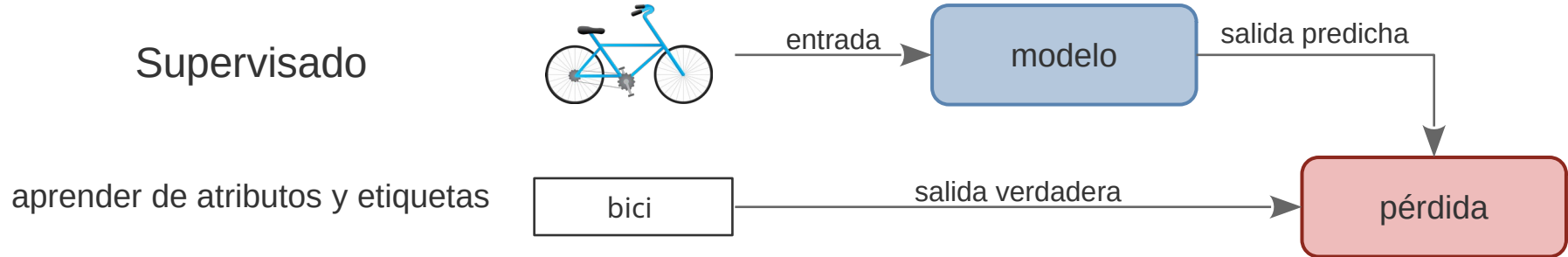


Aprendizaje de representaciones



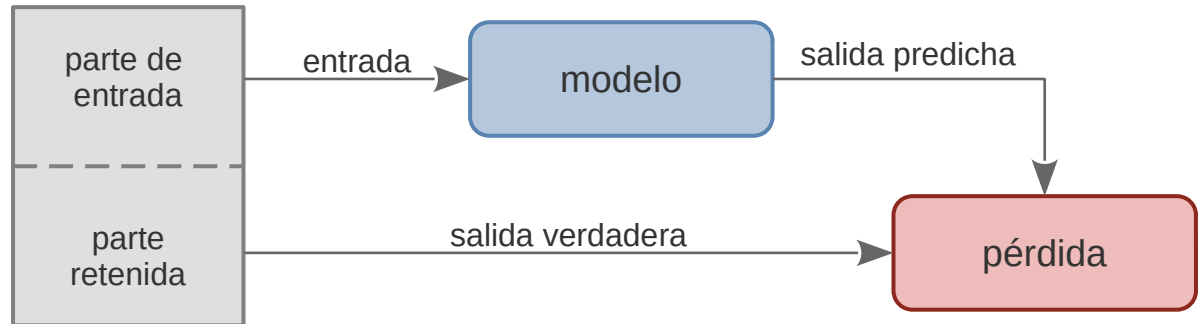
Autosupervisión

Supervisado



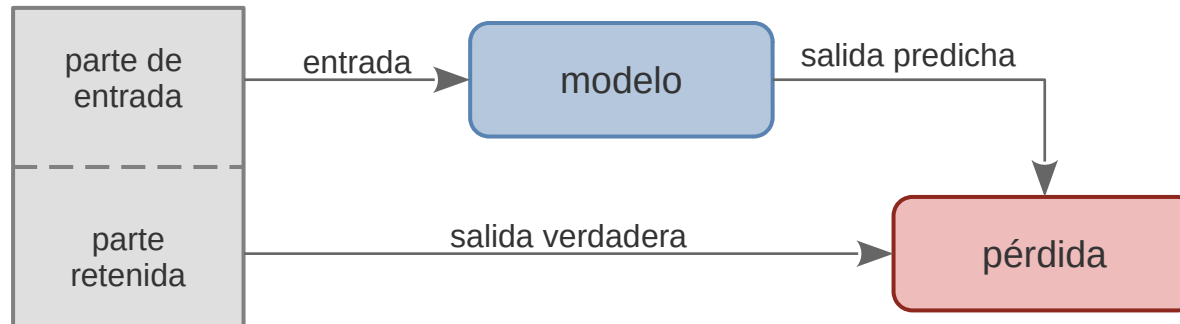
Autosupervisado

aprender solo de atributos



¿Qué es el AA?

- 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.
- Aprender representaciones iguales o mejores que las supervisadas.
- Explotar cantidades masivas de datos no etiquetados.



Pierre Sermanet
@psermanet

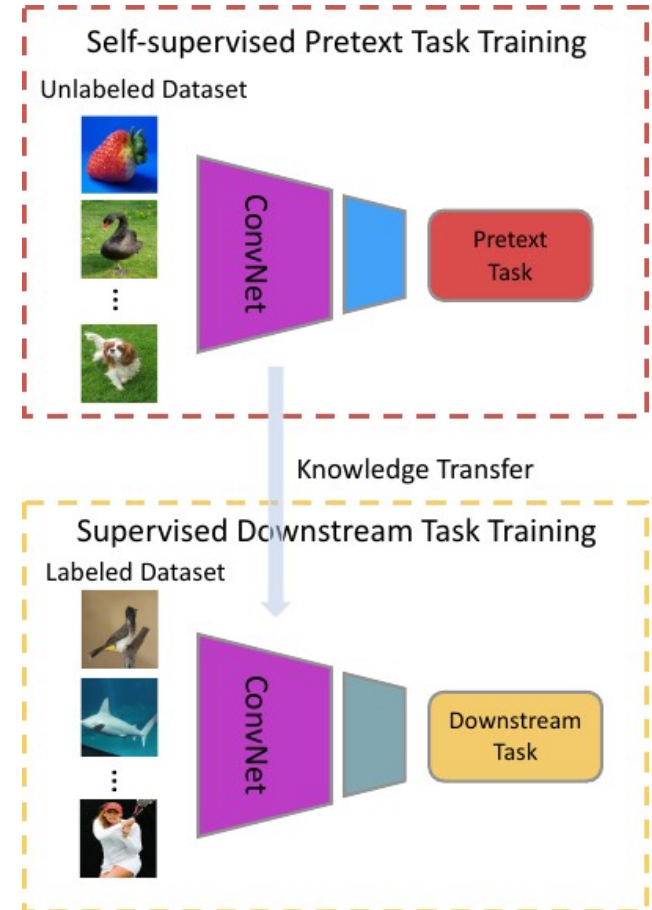


Give a robot a label and you feed it for a second;
teach a robot to label and you feed it for a lifetime.

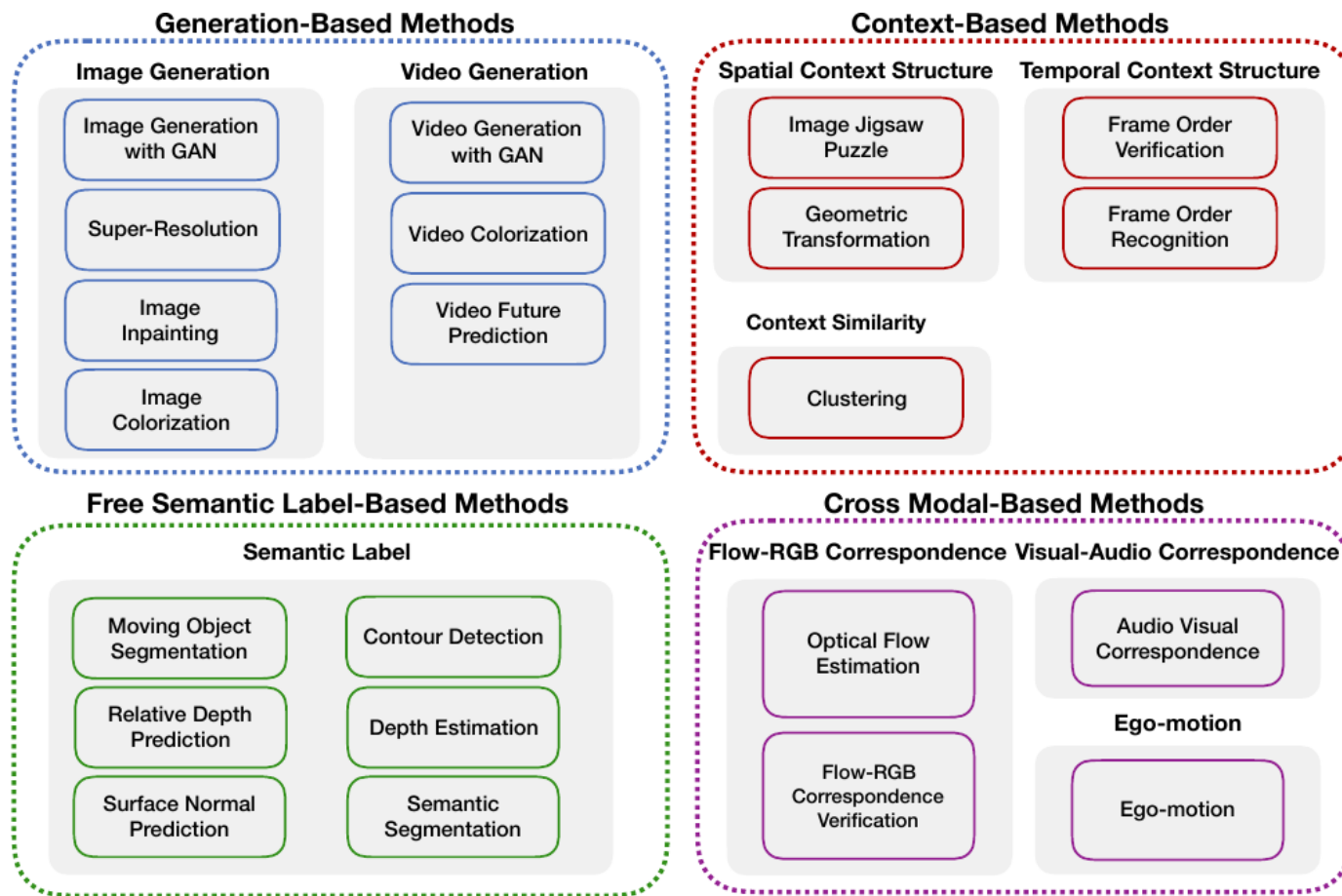
3:37 PM · 3 déc. 2018 · [Twitter for Android](#)

Escenarios de uso

- Tareas de naturaleza autosupervisada
- Preentrenamiento y transferencia de conocimiento



Tareas pretexto en VC



Generación de Imágenes

Restaurar los colores ¹

GT



ChromaGAN



w/o Class



Chroma Net.



Input (a)



Shift-net (b)



Contextual Attention (c)



Our Result (d)



Ground Truth (e)



Restaurar los parches ²

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

2. Coherent Semantic Attention for Image Inpainting. 2019.

Estructura de contexto espacial

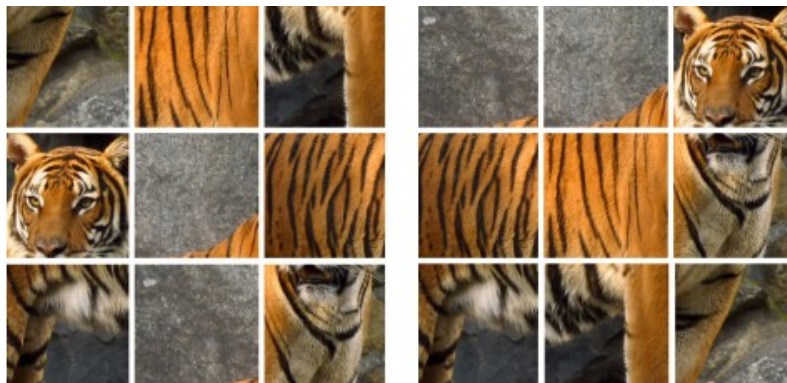
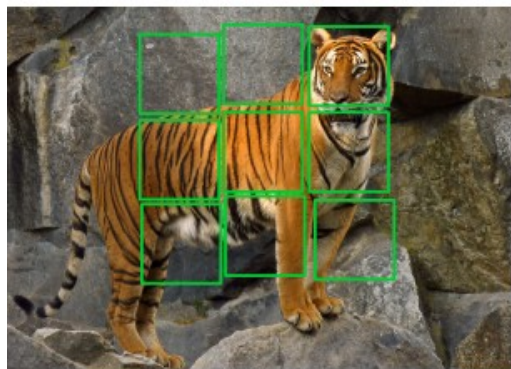
¿Cuál es la posición
relativa de los parches? ¹

Example:



Question 1:

Question 2:

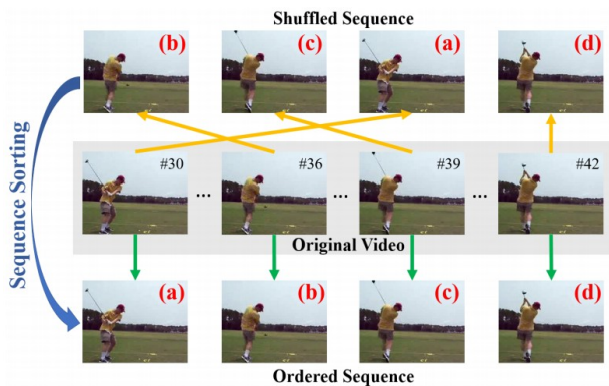
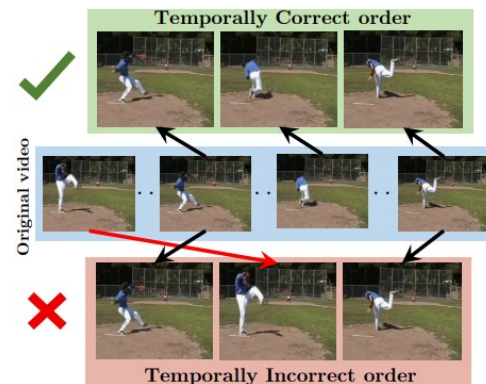


¿Cuál es el orden
de los parches? ²

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



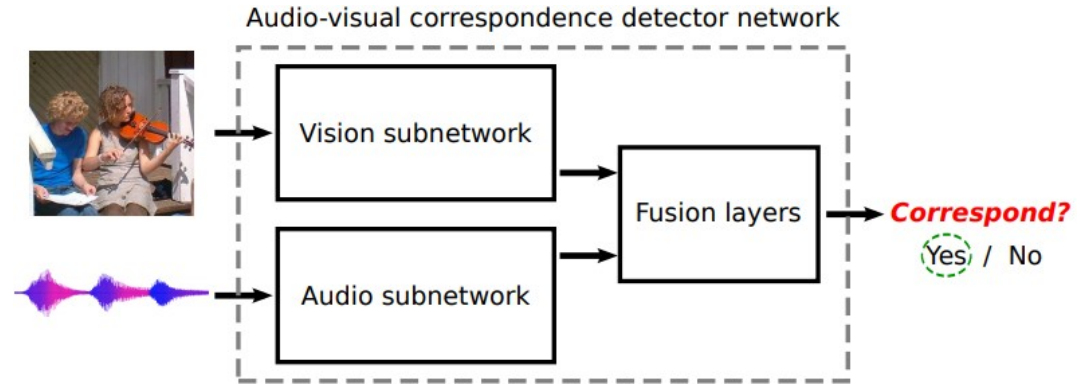
¿Cuál es el orden de los cuadros? ²

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

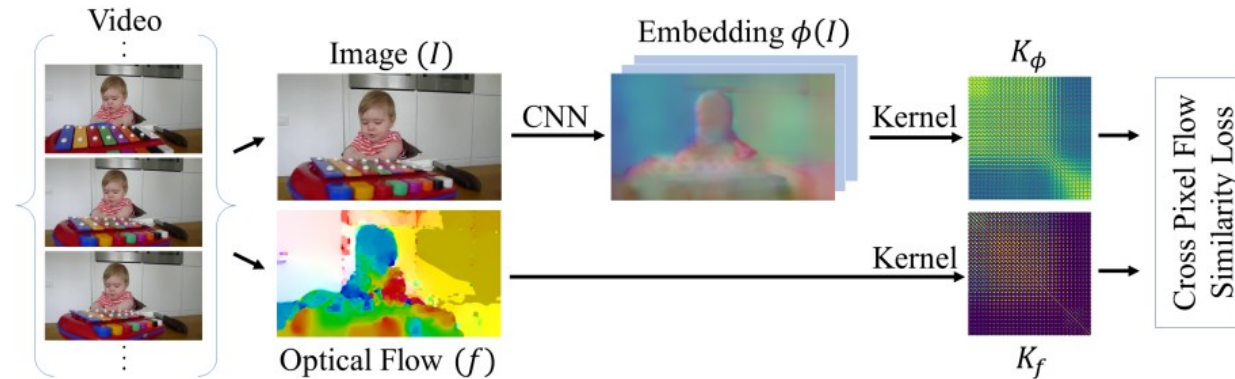
2. Unsupervised Representation Learning by Sorting Sequences. 2017.

Correspondencia de modalidad

Audio y video¹



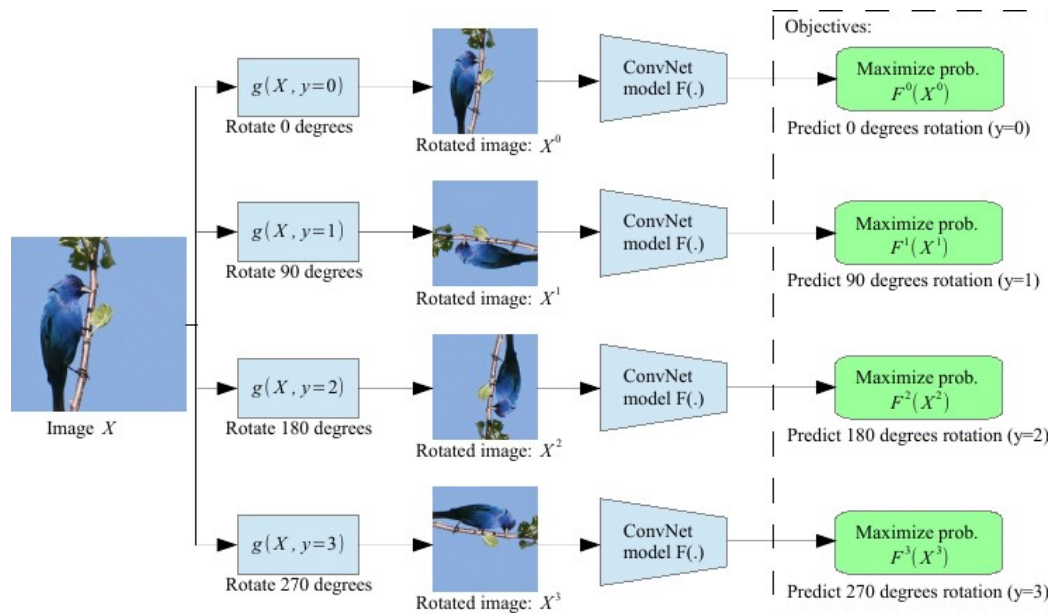
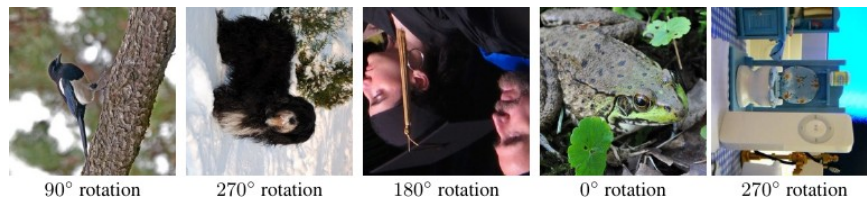
Flujo óptico y video²



1. Look, Listen and Learn. 2017.

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

¿Cuál es la rotación de la imagen?

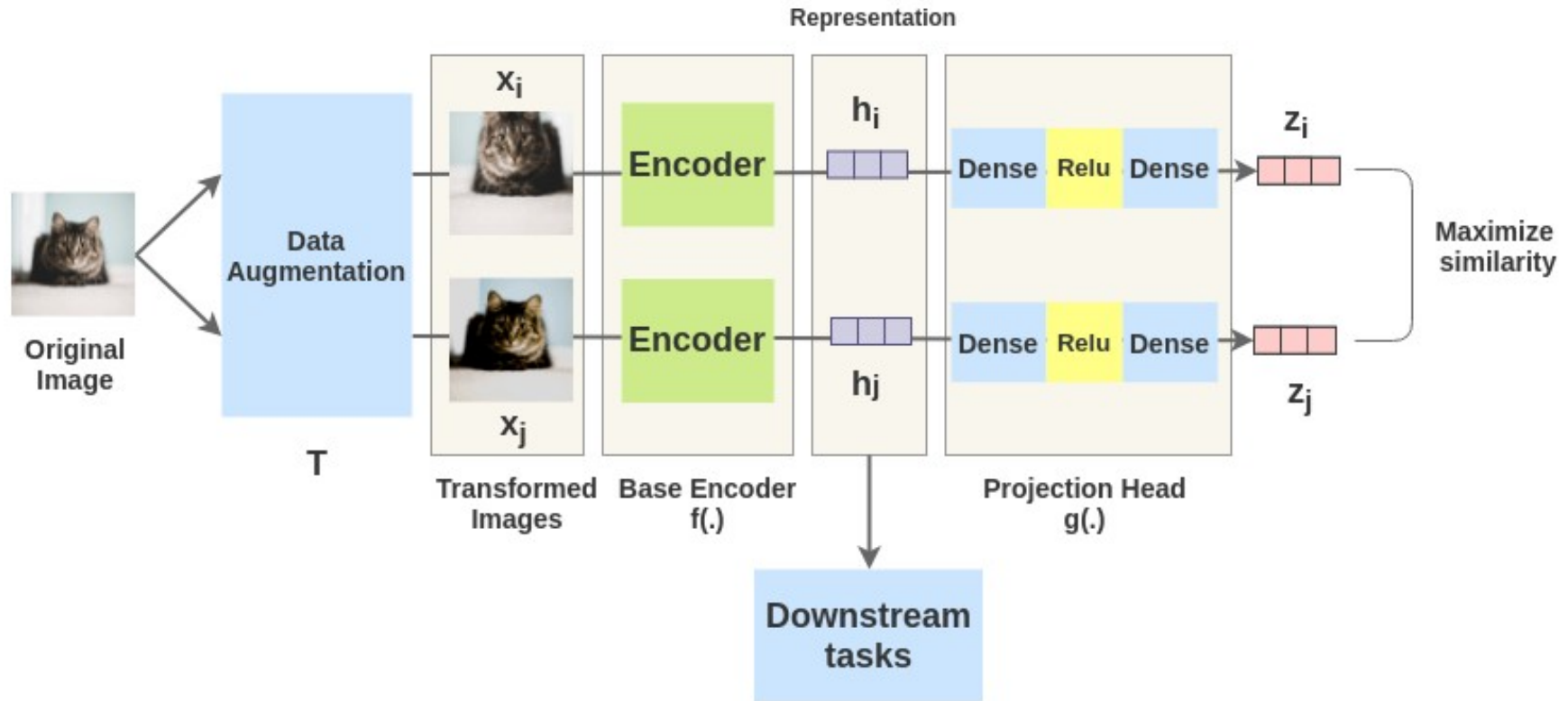




¡tiempo de programar!
3a_auto.ipynb

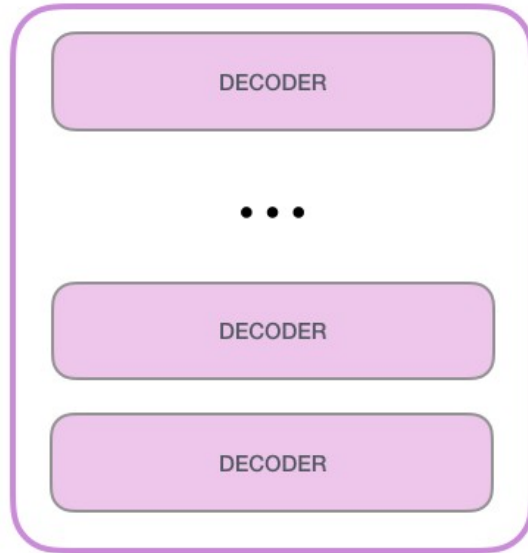
SimCLR

SimCLR Framework

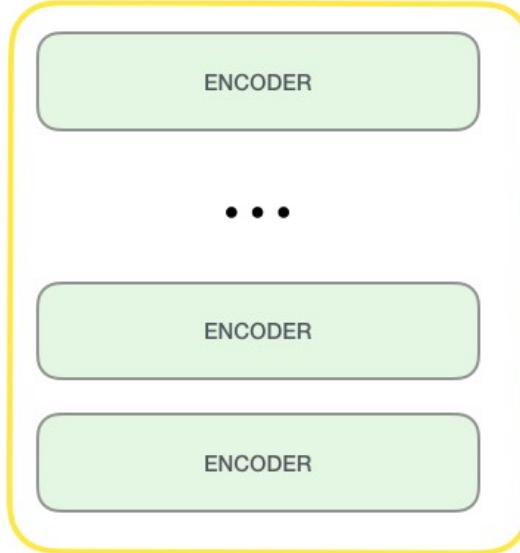


Modelos PLN

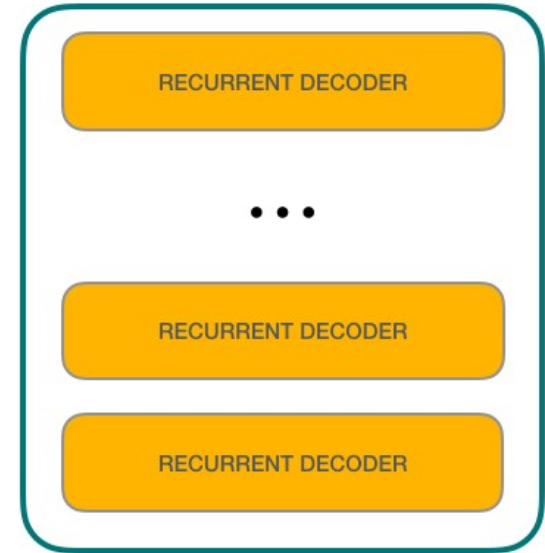
 GPT-2



 BERT



 TRANSFORMER XL





¡Gracias!

Ricardo Montalvo Lezama

<http://turing.iimas.unam.mx/~ricardoml/>
ricardoml@turing.iimas.unam.mx

Bere Montalvo Lezama

<http://turing.iimas.unam.mx/~bereml/>
bereml@turing.iimas.unam.mx