

# Transformers

- [TTTTT OUTRAGEOUSLY LARGE NN: THE SPARSELY-GATED MoE LAYER](#)
- [TT Deep Image Prior](#)
- [Making Convolutional Networks Shift-Invariant Again](#)
- [TT ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders](#)
- [TT Swin Transformer: Hierarchical Vision Transformer using Shifted Windows](#)
- [TT Hiera: A Hierarchical Vision Transformer without the Bells-and-Whistles](#)
- [TT Masked-attention Mask Transformer for Universal Image Segmentation](#)

# OUTRAGEOUSLY LARGE NN: THE SPARSELY- GATED MoE LAYER

## Integrantes:

- Valentina Pérez
- Sneider Sánchez

## Material de apoyo:

- **Diapositivas:** [Ver presentaciones](#)
- **Paper:** [Ver artículos académicos](#)

# ☐☐ Deep Image Prior

Deep Image Prior Teaser

## ☐☐ Integrantes:

- ☐☐ Jorge Andrey Garcia Vanegas
- ☐☐ Mateo Delgado

## ☐☐ Material de apoyo:

- ☐☐ **Diapositivas:** [Canva](#)
- ☐☐ **Paper:** [arxiv](#)

## ☐☐ Referencias

- [Artículo medium explicación](#)
- [Repositorio de código de Deep Image Prior](#)

# Making Convolutional Networks Shift-Invariant Again

Deep Image Prior Teaser

## ☐☐ **Integrantes:**

- ☐☐ Sebastian Solano
- ☐☐ Cristian Tristancho

## ☐☐ **Material de apoyo:**

- ☐☐ **Diapositivas:** [Ver presentaciones](#)
- ☐☐ **Paper:** [arxiv](#)

# 👤👤 ConvNeXt V2: Co-designing and Scaling ConvNets with Masked Autoencoders

## 👤👤 Integrantes:

- 👤👤 Valentina Pérez Díaz
  - 👤👤 Juan Camilo Arias Sarabia
- 

## Introducción

ConvNeXt V2 es una nueva familia de modelos de Redes Convolucionales (ConvNets) diseñada para mejorar su rendimiento significativamente a través del aprendizaje auto-supervisado. Aunque los modelos ConvNets modernos, representados por ConvNeXt, han demostrado un rendimiento sólido, los autores descubrieron que su simple combinación con técnicas de aprendizaje auto-supervisado como los autoencoders enmascarados (MAE) resulta en un rendimiento deficiente. Esto ocurre porque el diseño de codificador-decodificador de MAE está optimizado para el procesamiento de secuencias de transformers y no es compatible con las ConvNets estándar. Para superar este desafío, el trabajo propone un Fully Convolutional Masked Autoencoder (FCMAE), que utiliza convoluciones sparse para procesar eficientemente solo las partes visibles de la entrada enmascarada y reducir el costo del pre-entrenamiento, y la adición de una nueva capa llamada Global Response Normalization (GRN) a la arquitectura ConvNeXt para mejorar la competencia de características inter-canal. Esto dando como resultado ConvNeXt V2, que logra un rendimiento significativamente mejorado para ConvNets puros en diversas tareas de reconocimiento, incluyendo la clasificación ImageNet, la detección COCO y la segmentación ADE20K

---

# ☐☐ **Material de apoyo:**

- ☐☐ **Diapositivas:** [Ver presentaciones](#)
- ☐☐ **Paper:** [Ver artículos académicos](#)

# □□ Swin Transformer: Hierarchical Vision Transformer using Shifted Windows

## □□ Integrantes:

- □□ Samuel Penilla
- □□ Sneider Sánchez

## □□ Material de apoyo:

- □□ **Diapositivas:** [Ver presentaciones](#)
- □□ **Paper:** [Ver artículos académicos](#)

## □□ Referencias

- [GitHub del paper](#)
- [Video Explicativo](#)

# □□ Hiera: A Hierarchical Vision Transformer without the Bells-and-Whistles

## □□ Integrantes:

- □□ Brayan Yesid Quintero Santander
- □□ Juan Jose Ardila Aragón

## □□ Material de apoyo:

- □□ **Diapositivas:** [Ver presentaciones](#)
- □□ **Paper:** [Ver artículos académicos](#)

## □□ Referencias

- [Youtube Video](#)

# ☐☐ Masked-attention Mask Transformer for Universal Image Segmentation

## ☐☐ Integrantes:

- ☐☐ Nicolás Rivera
- ☐☐ Guillermo Pinto

## ☐☐ Material de apoyo:

- ☐☐ **Diapositivas:** [Ver presentaciones](#)
- ☐☐ **Paper:** [Ver artículos académicos](#)

## ☐☐ Referencias

- [Project's repository](#)
- [MaskFormer paper](#)