

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](#)
-

Revision #4

Created 1 October 2025 05:53:28 by Valentina Perez

Updated 1 October 2025 06:06:11 by Valentina Perez