

Transformers

- [ROFORMER: ENHANCED TRANSFORMER WITH ROTARY POSITION EMBEDDING](#)
- [📄 Depth Anything](#)
- [📄 CLIP](#)
- [📄 Open-Vocabulary RGB-Thermal Semantic Segmentation](#)
- [📄 Efficient Vision Transformers with Partial Attention](#)

ROFORMER: ENHANCED TRANSFORMER WITH ROTARY POSITION EMBEDDING

📋 Integrantes:

- 📋 Miguel Ángel Molina G.
- 📋 Sneider Sánchez

📋 Material de apoyo:

- 📋 **Diapositivas:** [Diapositivas](#)
 - 📋 **Paper:** [Artículo](#)
-

📋 Objetivos

Se plantean 3 objetivos para esta exposición:

- 📋 Analizar las limitaciones de los métodos tradicionales de encoding posicional (absoluto y relativo).
 - 📋 Justificar la necesidad de una mejor representación posicional en modelos de lenguaje.
 - 📋 Explicar cómo la matriz de rotación preserva la información relativa entre tokens.
-

📋 Resultados Esperados

📋 Se espera obtener al final de la sesión:

- Diferenciación entre RoPE y otros métodos de encoding posicional.
 - Explicar cómo la multiplicación por una matriz de rotación mantiene la magnitud de los embeddings inalterada, asegurando que la información posicional no distorsione las representaciones originales del modelo.
-

□□ Referencias

□□ Enlaces que pueden servir de ayuda para el lector:

□□ [Video complementario #1](#)

□□ [Video complementario #2](#)

□□ Depth Anything

□□ Integrantes:

- □□ Guillermo
- □□ Jorge

□□ Material de apoyo:

- □□ **Diapositivas:** [Depth Anything](#)
 - □□ **Paper:** [Depth Anything \(arXiv\)](#)
 - □□ **Código externo:** [Repositorio oficial en GitHub](#)
-

□□ Objetivos

□□ Como objetivos de la presente sesión, se plantean:

- Comprender el funcionamiento general de **Depth Anything** como modelo de estimación monocular de profundidad.
 - Reconocer sus ventajas frente a modelos supervisados y métodos tradicionales.
-

□□ Resultados Esperados

□□ Se espera que al final de esta sesión se:

- Comprenda la arquitectura y funcionamiento del encoder preentrenado y del decoder monocular.
 - Observar como se pueden utilizar datos no etiquetados para mejorar el funcionamiento de la red.
-

□□ Referencias

□□ Esta sección recopila enlaces a recursos relevantes sobre visión por computadora y estimación de profundidad:

[Space Depth Anything](#)

[Página del proyecto con demos](#)

CLIP

Integrantes:

- Juan Calderón
- César Vanegas

Material de apoyo:

- **Diapositivas:** [Ver presentaciones](#)
 - **Paper:** [Learning Transferable Visual Models From Natural Language Supervision](#)
 - **Código externo:** [Repositorio de código](#)
-

Objetivos

Resultados Esperados

Referencias

Esta sección recopila enlaces a recursos relevantes para la sesión:

Open-Vocabulary RGB-Thermal Semantic Segmentation

Integrantes:

- Julián León
- Miguel Pimiento

Material de apoyo:

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

Objetivos

En esta sección se definen los objetivos de la sesión:

- ¿Por qué es importante este tema?
 - ¿Qué se espera lograr durante la sesión?
-

Resultados Esperados

Esta sección describe de manera general lo que se espera obtener al final de la sesión:

- Mayor comprensión del tema tratado.
 - Identificación de conceptos clave.
 - Recopilación de información relevante para futuras implementaciones.
-

Metodología

▢ Aquí se explicarán todos los temas tratados en la sesión con mayor detalle. Esta sección se completará después de la sesión e incluirá:

- ▢ Explicaciones detalladas del proceso.
- ▢ Análisis de los conceptos presentados.
- ▢ Ejemplos prácticos y fragmentos de código.

▢ **Ejemplo de código en Python:**

```
import cv2
import matplotlib.pyplot as plt

imagen = cv2.imread(".images/ejemplo.png")
plt.imshow(cv2.cvtColor(imagen, cv2.COLOR_BGR2RGB))
plt.show()
```

▢ **Uso de imágenes**

⚠ Solo utilizar imágenes disponibles en internet debido a las limitaciones de almacenamiento.

▢ **Ejemplo de imagen adjunta:**

Ejemplo de imagen

▢ También puedes ajustar el tamaño y alineación de las imágenes:

drawing

drawing

▢ **Ejemplo de tabla:**

▢ A	▢ B	▢ C
✓ Uno	Texto de prueba	▢

▢ Referencias

▢ Esta sección recopila enlaces a recursos relevantes sobre procesamiento de imágenes:

[▢▢ Documentación de OpenCV](#)

[▢▢ Guía de NumPy](#)

[▢▢ Artículo sobre procesamiento de imágenes](#)

Efficient Vision Transformers with Partial Attention

Integrantes:

- Brayan Quintero
- Valentina Pérez

Material de apoyo:

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

Objetivos

En esta sección se definen los objetivos de la sesión:

- ¿Por qué es importante este tema?
 - ¿Qué se espera lograr durante la sesión?
-

Resultados Esperados

Esta sección describe de manera general lo que se espera obtener al final de la sesión:

- Mayor comprensión del tema tratado.
 - Identificación de conceptos clave.
 - Recopilación de información relevante para futuras implementaciones.
-

Metodología

▢ Aquí se explicarán todos los temas tratados en la sesión con mayor detalle. Esta sección se completará después de la sesión e incluirá:

- ▢ Explicaciones detalladas del proceso.
- ▢ Análisis de los conceptos presentados.
- ▢ Ejemplos prácticos y fragmentos de código.

▢ **Ejemplo de código en Python:**

```
import cv2
import matplotlib.pyplot as plt

imagen = cv2.imread(".images/ejemplo.png")
plt.imshow(cv2.cvtColor(imagen, cv2.COLOR_BGR2RGB))
plt.show()
```

▢ **Uso de imágenes**

⚠ Solo utilizar imágenes disponibles en internet debido a las limitaciones de almacenamiento.

▢ **Ejemplo de imagen adjunta:**

Ejemplo de imagen

▢ También puedes ajustar el tamaño y alineación de las imágenes:
drawing

drawing

▢ **Ejemplo de tabla:**

▢ A	▢ B	▢ C
✓ Uno	Texto de prueba	▢

Referencias

▢ Esta sección recopila enlaces a recursos relevantes sobre procesamiento de imágenes:

[▢▢ Documentación de OpenCV](#)

[▢▢ Guía de NumPy](#)

[▢▢ Artículo sobre procesamiento de imágenes](#)