

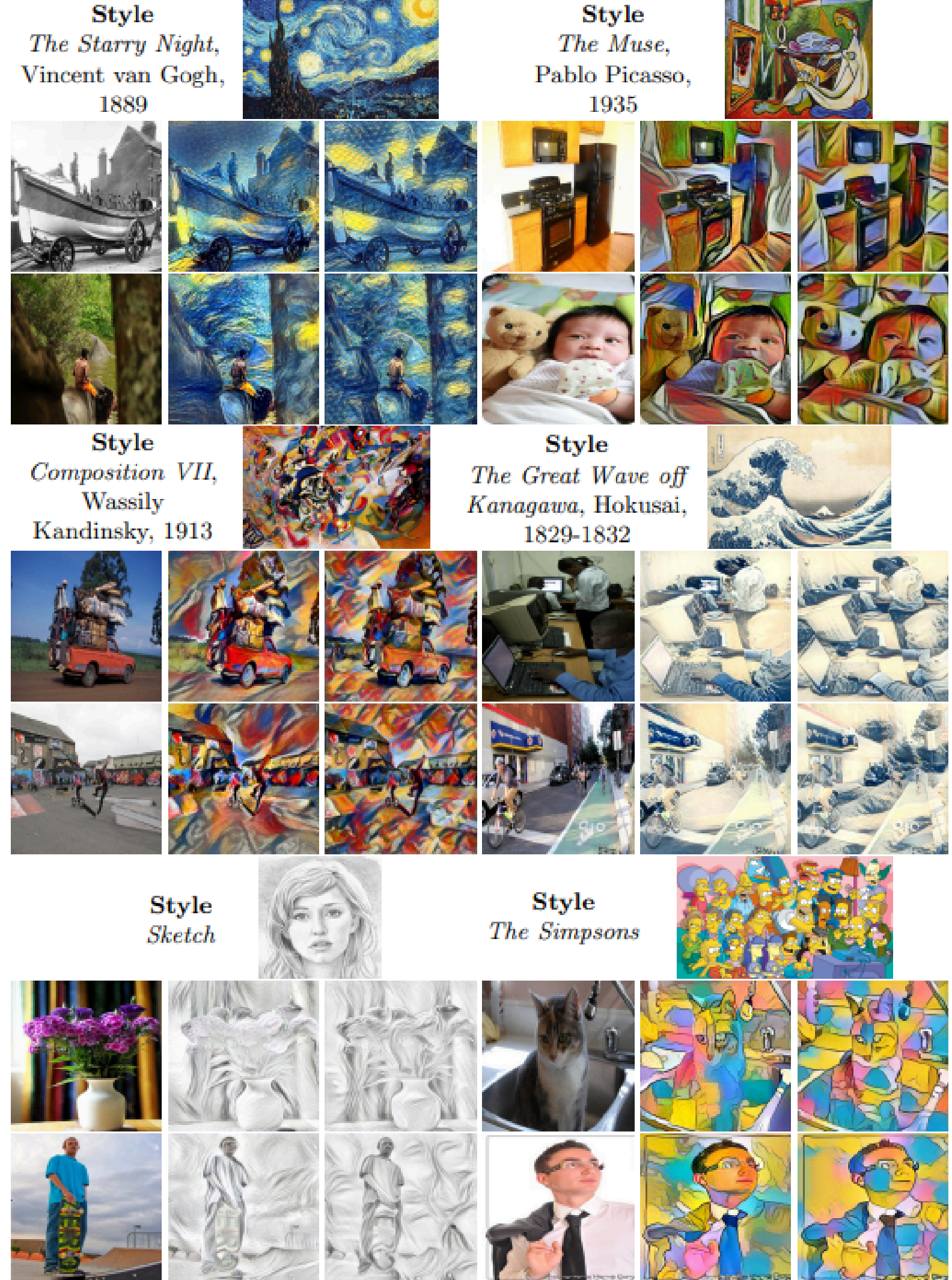
Ramiro Sanes y Joaquín Guerra

STYLE TRANSFER CON PERCEPTUAL LOSS

**Obligatorio Inteligencia Artificial Generativa
Universidad ORT**

Índice

- ¿Qué es Style Transfer?
- Contexto
- Introducción Teórica
- Motivación
- Nuestra Implementación



¿Qué es Style Transfer?



Style Transfer es una técnica que permite combinar:

- **Contenido** de una imagen (su forma y estructura)
- **Estilo** de otra (colores, patrones, texturas)

Resultando en una nueva imagen que mantiene los objetos de la primera, pero capturando la estética de la segunda

Algunos ejemplos clásicos son: super-resolución, denoising, colorización



Contexto

Métodos Tradicionales

Entrenaban redes feed-forward usando funciones de pérdida pixel a pixel que no logran capturar diferencias perceptuales, fundamentales para este tipo de problemas

Métodos de Optimización (Gatys)

Generan imágenes de alta calidad usando pérdidas perceptuales, pero son **lentos computacionalmente** ya que requieren resolver un problema de optimización para cada imagen en inferencia.

Que se busca lograr?

Combinar la velocidad de los métodos tradicionales con la calidad de los métodos perceptuales.

Contenido + Estilo → Imagen Resultante

Content :



Style :



RESULT

Solución Propuesta

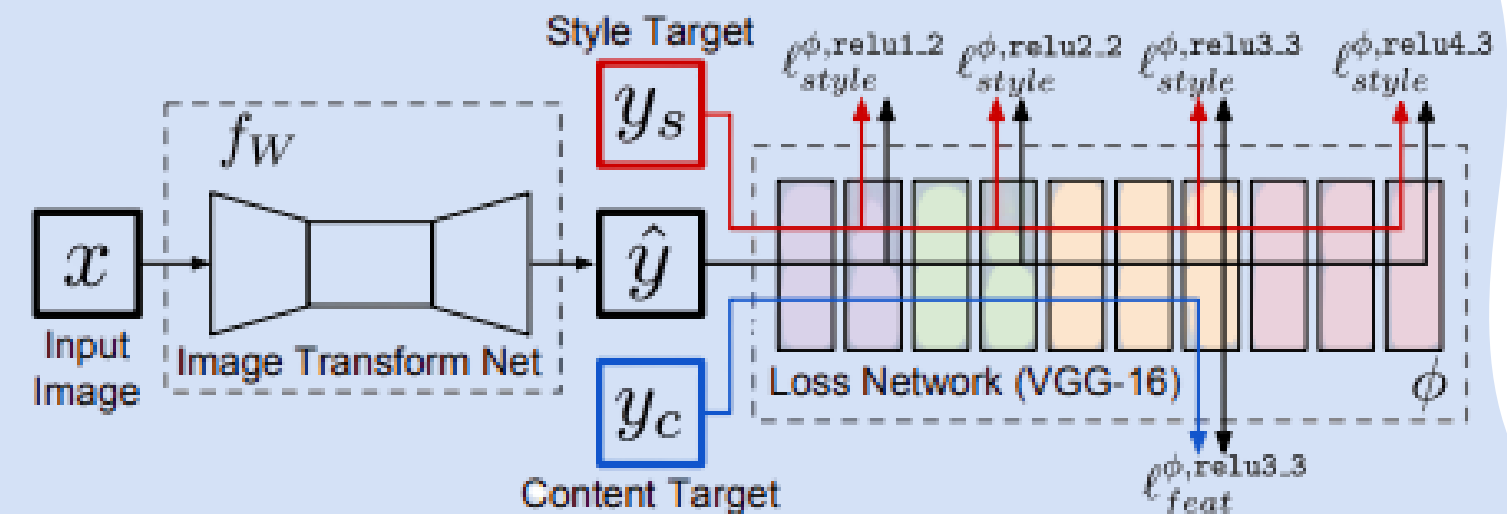
La idea propuesta por el paper de referencia (Johnson et al., 2016) consiste en entrenar una red feed-forward que en lugar de comparar imágenes pixel a pixel use Perceptual Losses:

- Pérdida **basada en características internas de una red preentrenada** (VGG)
- Permite medir **contenido y estilo**

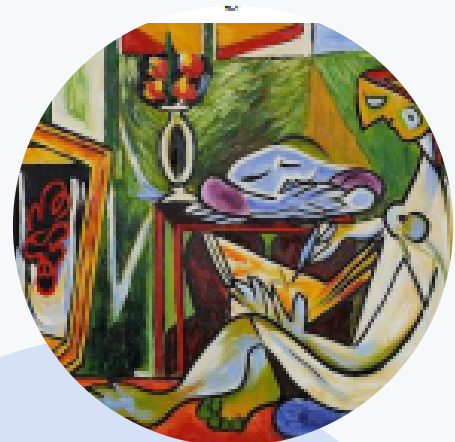
Arquitectura

El sistema se compone de dos elementos clave:

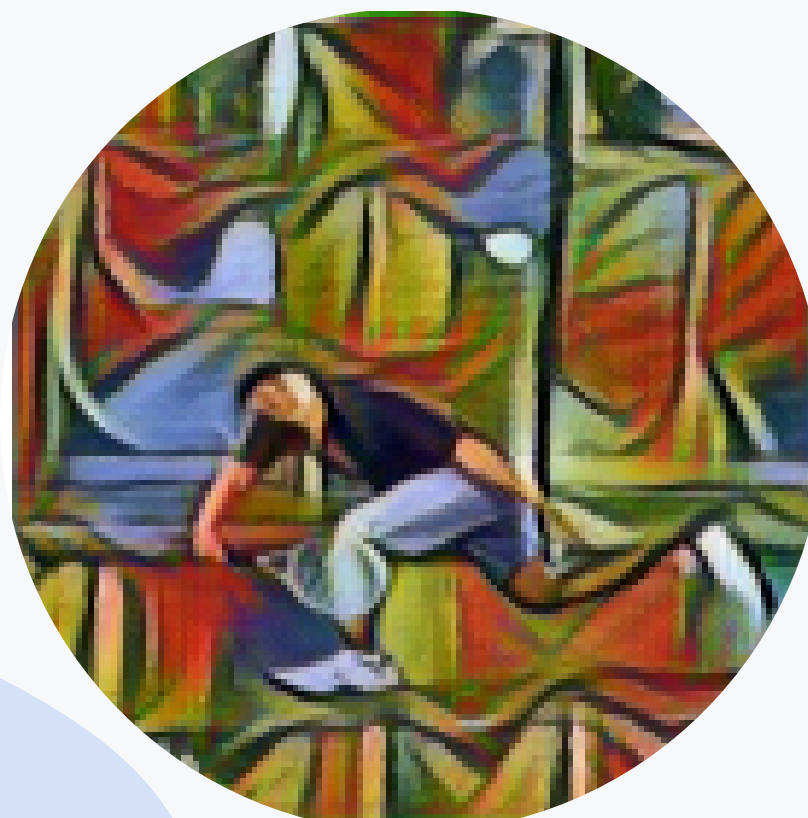
- **Red de transformación de imágenes** (f_W) residual convolucional
- **Red de pérdida** (ϕ) convolucional profunda, **fija** (pre-entrenada para clasificación). Esta red define las funciones de pérdida transfiriendo conocimiento semántico a la Red de Transformación.



Funciones de Pérdida



+



Pérdida de Reconstrucción de Contenido (L_{feat})

Mide las diferencias de contenido entre la imagen de salida y el objetivo de contenido

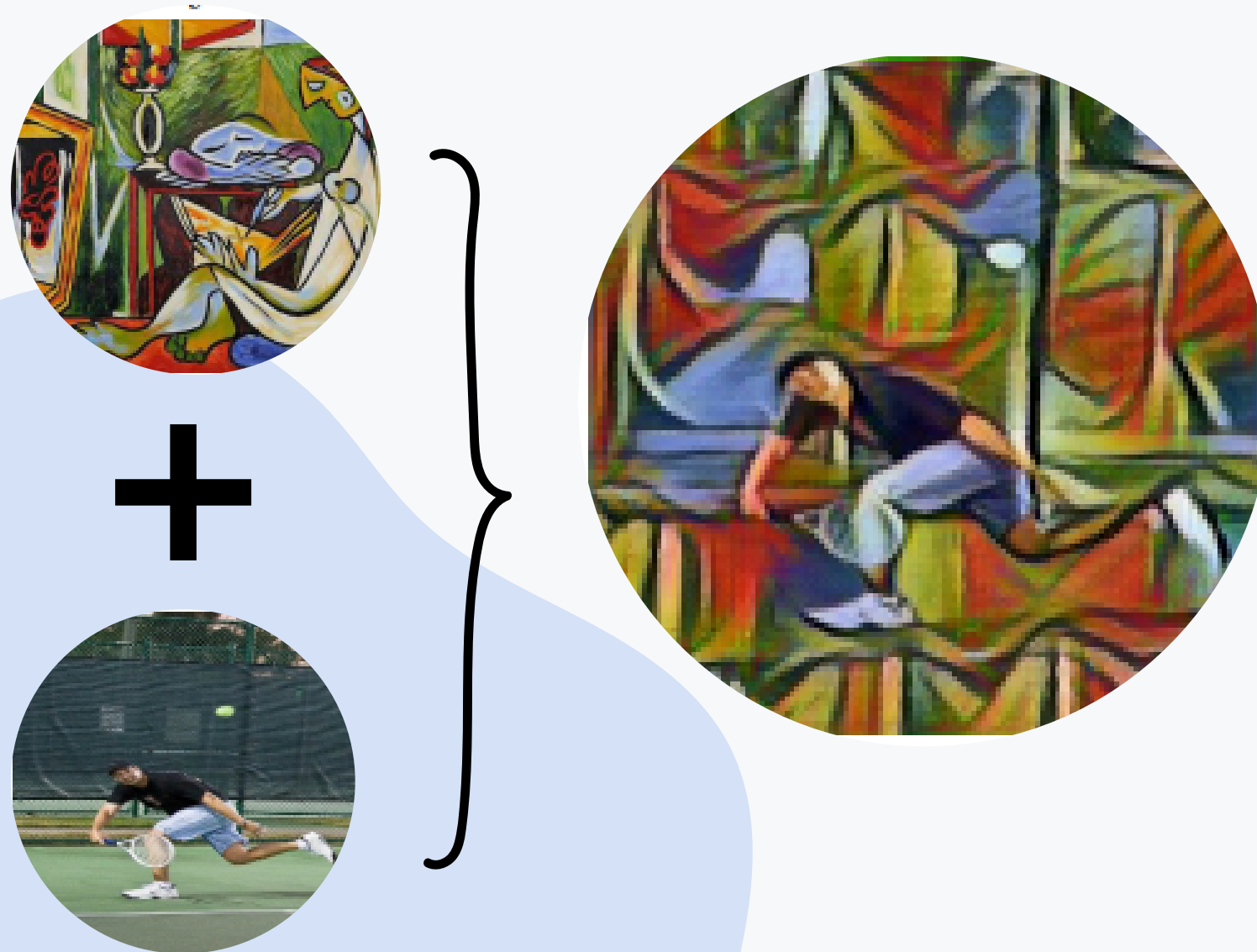
Pérdida de Reconstrucción de Estilo (L_{style})

Penaliza las diferencias en estilo (colores, texturas, patrones).

Se basa en la **Matriz de Gram** que indica que características tienden a activarse conjuntamente.

Se preservan las características estilísticas pero no la estructura espacial.

Funciones de Pérdida



¿Por qué no pérdida por píxeles?

La pérdida por píxeles **penaliza diferencias exactas en color**, mientras que la Perceptual Loss penaliza diferencias en características semánticas.

Al desplazar 1 píxel una imagen, la perceptual loss entiende que sigue siendo la misma escena mientras que **la pérdida por píxeles identifica un error grande**.

Como **no existe una imagen objetivo** correcta para comparar, necesitamos medir similitud perceptual en lugar de similitud pixel a pixel. La Perceptual Loss captura forma, textura y estilo usando representaciones internas de VGG, que están entrenadas para reconocer contenido semántico.

Motivación

No existe una solución correcta única. No hay un 'ground truth'.

Resultados visuales: La calidad del resultado es algo que se percibe, no solo algo que se mide con una métrica. Trabajar con funciones de pérdida que buscan asemejarse al ojo humano.

Tiene aplicaciones **prácticas claras**, como filtros de Instagram, TikTok o herramientas de edición que vemos en el día a día.

Nuestra implementación

- Implementar la red transformadora (fW).
- Usar VGG19 para extraer contenido y estilo.
- Entrenar un modelo para cada estilo elegido.
- Comparar resultados en tiempo y calidad contra los mencionados en el paper.

¡Gracias!