

Resumen

Se propone un modelo de red neuronal con el propósito de realizar correcciones de un pequeño conjunto de datos (dataset) de imágenes borrosas aplicando una técnica llamada transferencia de aprendizaje. Se implementaron 3 modelos diferentes bajo el mismo concepto del modelo propuesto con el propósito de realizar comparaciones. Los resultados muestran que la corrección visual generada por el modelo propuesto es superior a los modelos que se compararon y queda la motivación de realizar trabajos relacionados con el tema que logren brindar mejores resultados.

Introducción

Actualmente existe una gran aglomeración de información en internet, donde gran parte de esta información consiste en imágenes tomadas por miles de personas alrededor del mundo que terminan posteándose en redes sociales, sitios web, servidores de compañías entre otros. Sin embargo, muchas de estas imágenes fueron capturadas bajo malas condiciones climáticas o la escena se encontraba en movimiento generando distorsión, borrosidad y ruido.

En este proyecto se busca aplicar una de las técnicas para las redes neuronales denominada transferencia de aprendizaje (ver figura 1), que permite una mejor generalización de datos durante la reconstrucción resolviendo un problema similar al que se quiere resolver, al aplicar esta técnica se descubre que se obtienen mejores reconstrucciones visuales comparándola contra otras redes neuronales que siguen el mismo concepto.

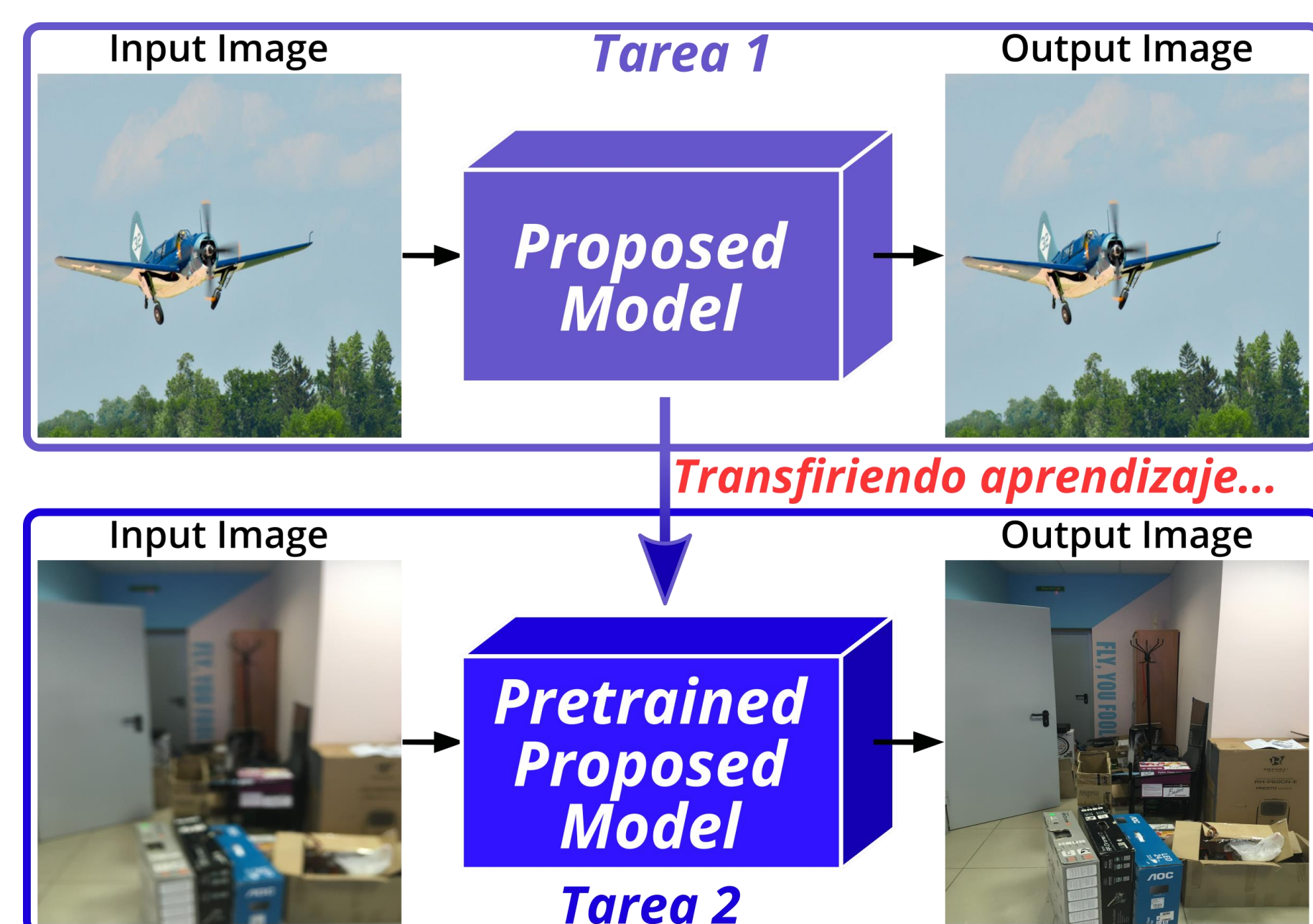


Figura 1. Concepto de transferencia de aprendizaje usando como ejemplo el modelo propuesto.

Proceso y método

- **Modelo propuesto:** una arquitectura de autoencoder basada en U-Net (ver figura 2). La arquitectura se compone de los bloques: Double Conv2D, contribuye al mapeo de características más relevantes del dataset de entrenamiento. Bloque Downsample, permite reducir espacialmente el tamaño del mapa de características. Bloque Upsample, permite aumentar la muestra del mapa de características de entrada a un mapa de características de salida deseado.
- **Hiperparámetros:** los hiperparámetros usados para todos los modelos implementados son el optimizador Adam con tasa de aprendizaje de 1×10^{-3} y 100 épocas para reforzar el uso de la transferencia de aprendizaje.
- **Métricas:** se usó la proporción Máxima de Señal a Ruido $PSNR = 10\log_{10}(\frac{\|I_n\|_{\infty} NM}{\|I_e - I_n\|_F^2})$ y el error cuadrático medio $MSE = \frac{1}{n} \sum_{i=1}^n (I_n^i - I_e^i)^2$ donde se comparan las imágenes pixel a pixel para medir la calidad de reconstrucción.
- **Datasets:** Se usaron dos datasets para resolver el modelo propuesto con una configuración establecida como se observa en la figura 5, el dataset de la derecha contiene pocas imágenes y estas no fueron tomadas pixel-a-pixel.
- **Procedimiento:** consta de dos tareas (ver figura 1), la primera tarea es aprender los pesos óptimos del modelo propuesto usando de dataset de imágenes naturales (ver figura 5), esto se hace con propósito de mitigar el problema del sobre ajuste de este tipo de modelos cuando se entrenan con pocas imágenes. Una vez el modelo propuesto es entrenado, se procede a resolver la segunda tarea que consiste en congelar la primera mitad de sus pesos (ver figura 2) ya que estos son los que recuerdan las características más generales de las imágenes como líneas y curvas evitando el sobre ajuste mencionado y así solo se entrenarían los pesos no congelados del modelo usando el dataset de imágenes borrosas.
- Los modelos entrenados para comparar contra el modelo propuesto fueron: (1) Modelo preentrenado sin congelamiento de capas, (2) modelo sin transferencia de aprendizaje y (3) modelo Xception modificado para realizar la tarea de autoencoder.

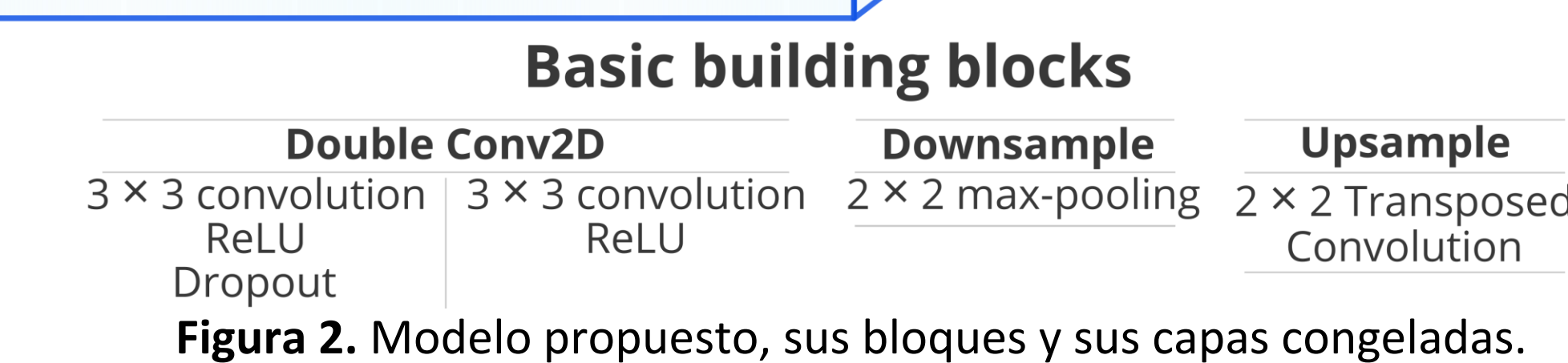
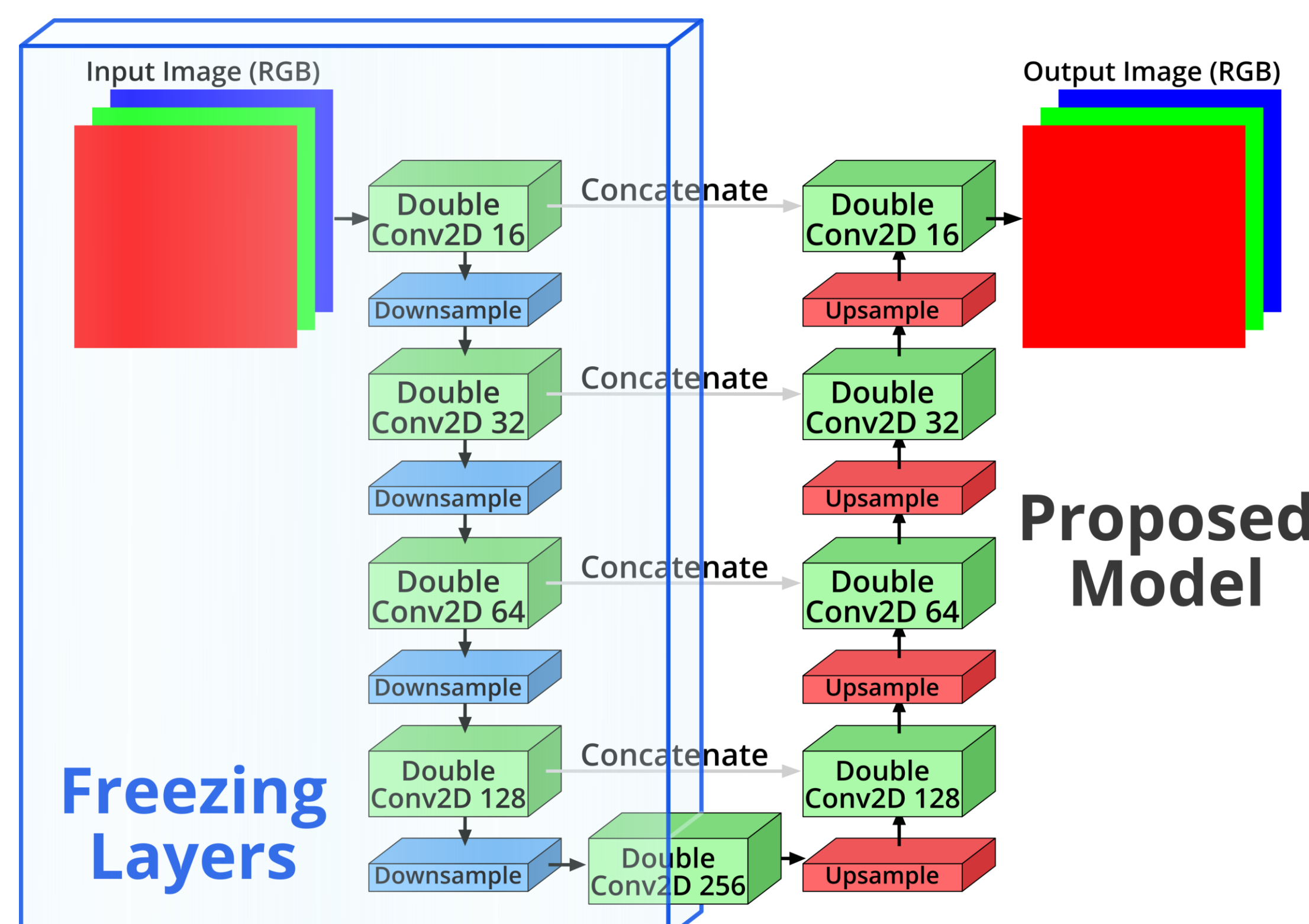


Figura 2. Modelo propuesto, sus bloques y sus capas congeladas.

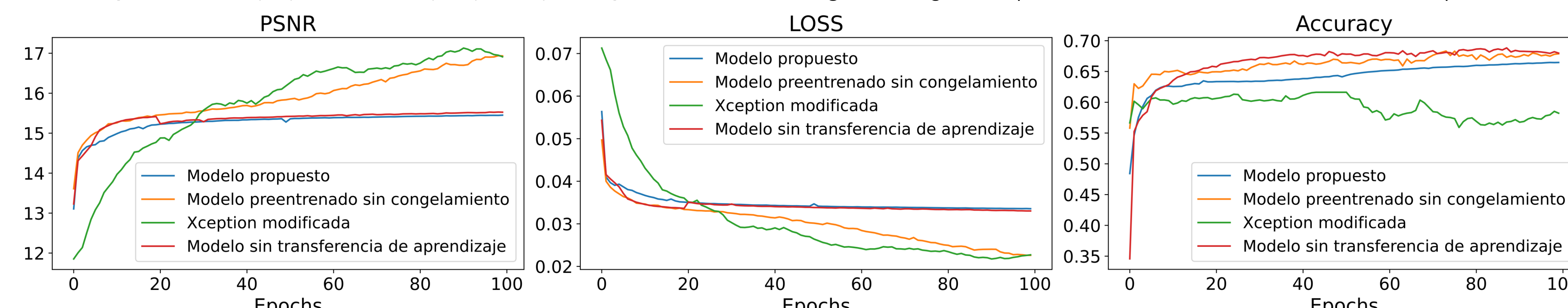


Figura 4. Historial de reconstrucción del modelo propuesto vs los modelos para comparación mediante las métricas PSNR, MSE (LOSS) y Accuracy (0 – 100%).

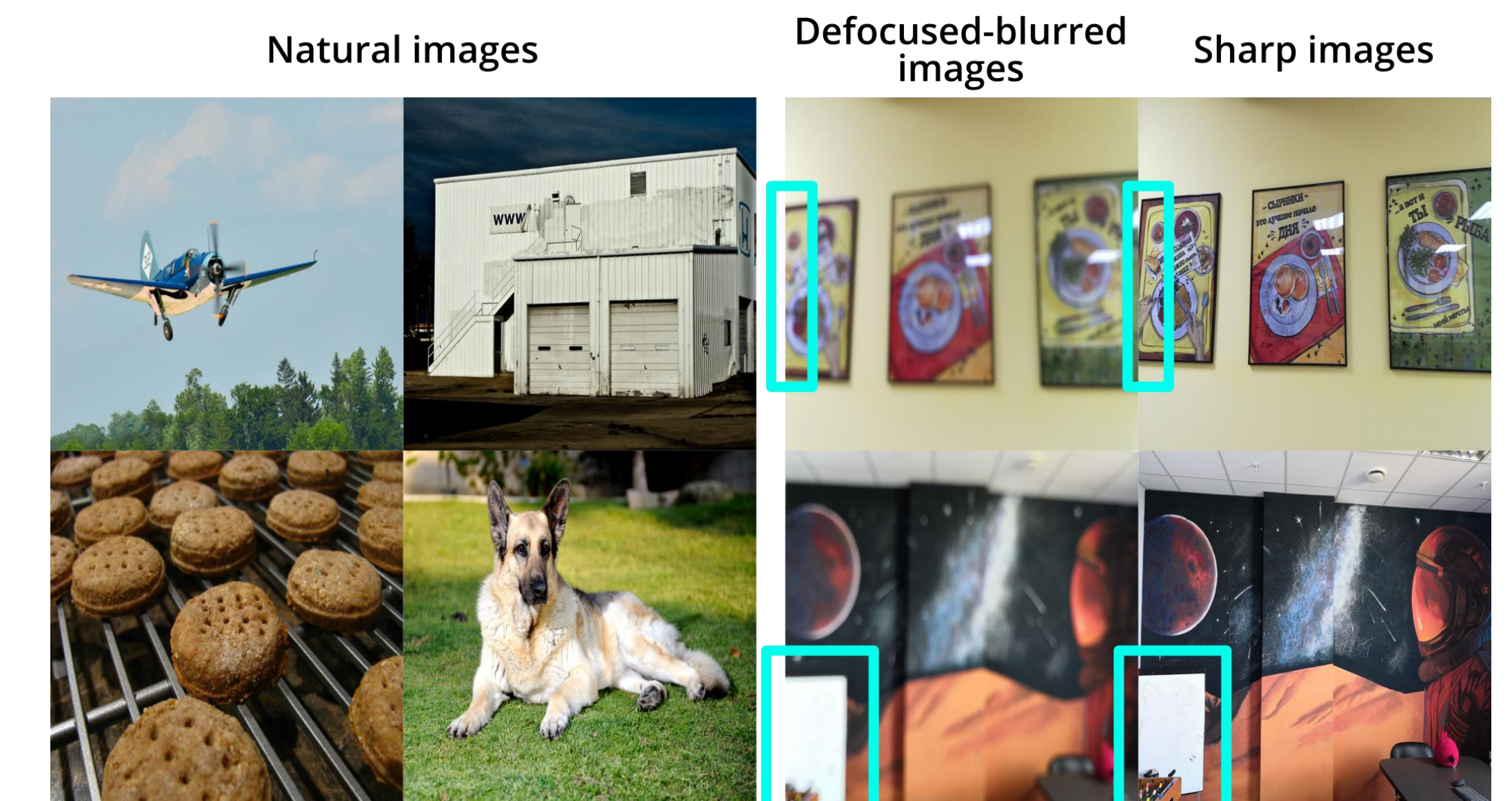


Figura 5. Datasets usados con sus configuraciones establecidas.

Dataset	Images	Preprocessing
Open Images V6 (RGB)	2000	$512 \times 512 \times 3$ [0, 1]
Blur images	350	$512 \times 512 \times 3$ [0, 1]

Resultados

Se resumen de los resultados cuantizados por PSNR en la tabla 1, las restauración de las imágenes borrosas se observa en la figura 3 y la calidad de reconstrucción del modelo durante el entrenamiento se observa en la figura 4.

	Propuesto	Sin congelar	Sin TF	Xception modificado
count	70.000000	70.000000	70.000000	70.000000
mean	14.107340	14.127115	14.085491	11.793798
std	2.581096	2.524316	2.581738	1.440310
min	7.462991	7.589792	7.464261	8.418376
25%	12.492813	12.508280	12.400735	10.804779
50%	13.966845	14.026080	13.895311	11.867144
75%	15.418158	15.418882	15.522024	12.901189
max	19.787668	21.049421	20.203077	14.756302

Tabla 1. Resumen de los resultados obtenidos para cada uno de los modelos cuantizados mediante la métrica de PSNR.

Conclusiones

A pesar de que el método propuesto obtuvo un segundo lugar respecto a la media en la tabla 1, podemos observar que en la figura 3 que fue la que mejores resultados visuales generó, esto quiere decir que la técnica de transferencia de aprendizaje cumplió su objetivo de generalizar los datos entrenados en la tarea 1 para evitar el sobre ajuste, lo cual no ocurrió en las demás reconstrucciones realizadas.

Trabajo Futuro

Este trabajo es un marco de referencia que busca la exploración de técnicas para redes neuronales que aún no han sido exploradas en profundidad, por lo que este campo de investigación aún tiene mucho que aportar a la sociedad y al campo de la fotografía.

Información de contacto

Emmanuel David Martínez Estrada, Email: enmartz@hotmail.com
María Paula Riveros Gómez, Email: paulariveros0321@gmail.com

Profesor
Fabio Martínez Carrillo, Email: famarcarr@saber.uis.edu.co

Referencias Bibliográficas

1. Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., Pont-Tuset, J., ... & Ferrari, V. (2018). The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale. arXiv preprint arXiv:1811.00982.
2. Kwentar/blur_dataset. (2020). Retrieved 6 September 2020, from https://github.com/Kwentar/blur_dataset.
3. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
4. Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258).