

FACULTAD DE INFORMÁTICA

GENERACIÓN DE IMÁGENES CON TÉCNICAS DE APRENDIZAJE AUTOMÁTICO. APLICACIONES EN RECONOCIMIENTO DE LENGUA DE SEÑAS

Rios, Gastón Gustavo

Ronchetti, Franco (Dir.), Hasperué, Waldo (Codir.)

Instituto de Investigación en Informática (III-LIDI).

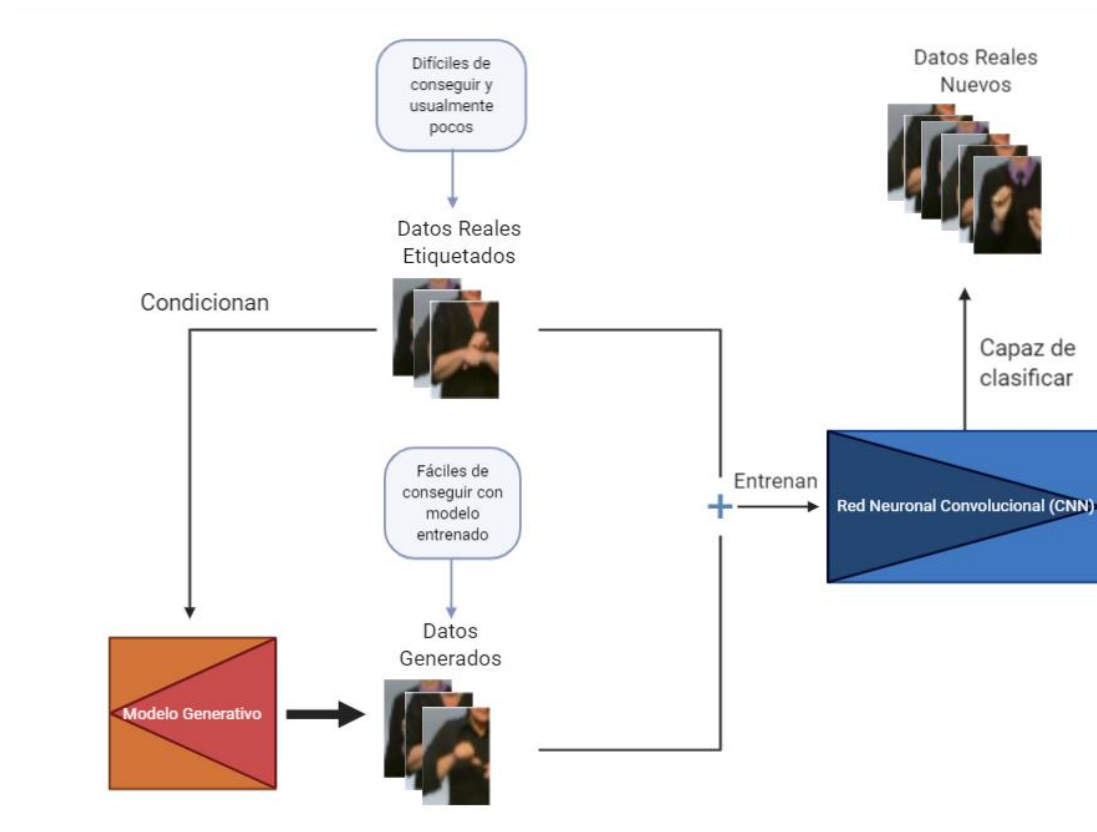
grios@lidi.info.unlp.edu.ar

PALABRAS CLAVE: aprendizaje automático, redes neuronales, sistemas inteligentes e IA, visión por computadora, generación de imágenes.

IMAGE GENERATION USING MACHINE LEARNING. APPLIED TO SIGN LANGUAGE RECOGNITION

KEYWORDS: artificial intelligence, artificial neural networks, computer vision, machine learning, image generation.

Resumen gráfico



Resumen

El objetivo general de este plan de investigación es el de crear modelos y técnicas de entrenamiento capaces de generar imágenes artificiales realistas y variadas en entornos con pocos datos etiquetados. Con estos modelos y técnicas de entrenamiento se facilitará el acceso a múltiples nuevas áreas de aplicación para modelos generativos. El énfasis estará puesto en la generación de imágenes de gestos de lengua de señas, permitiendo de esta forma entrenar modelos discriminadores precisos que utilicen aprendizaje profundo a partir de pocos datos etiquetados. Los objetivos son los siguientes: 1) Estudiar y analizar las bases de datos públicas sobre reconocimiento de gestos, incluyendo bases de datos específicas para lengua de señas, y de grandes bases de datos de diversos dominios para ser utilizadas en las técnicas de entrenamiento. Particularmente LSA16 y LSA64, bases de datos para la Lengua de Señas Argentina desarrollada en el III-LIDI. 2) Crear nuevos modelos de generación de imágenes que posean coherencia semántica. Particularmente se estudiarán los autoencoders y las Generative Adversarial Networks (GAN) ya que son los modelos más utilizados en el estado del arte. 3) Crear nuevas técnicas de entrenamiento que permitan el entrenamiento de modelos generadores utilizando una cantidad de datos limitada. En particular se estudiarán transfer learning, data

augmentation y aprendizaje semi supervisado. 4) Analizar y comparar los resultados obtenidos utilizando los modelos generativos y técnicas de entrenamiento desarrollados. Se usarán métricas comúnmente utilizadas en este tipo de problemas como Frechet Inception Distance (FID) e Inception Score (IS). Adicionalmente se realizarán evaluaciones semánticas para comprobar la variabilidad y coherencia de las imágenes generadas. 5) Desarrollar modelos clasificadores de imágenes aplicado al reconocimiento de gestos de la lengua de señas. Particularmente, se estudiarán Redes Neuronales Convolucionales (CNN) específicas para problemas de clasificación de objetos en imágenes. Aplicación, análisis y comparación de los resultados obtenidos utilizando los nuevos datos generados. Con esta investigación se espera realizar un aporte significativo al estado del arte en cuanto a modelos generativos de imágenes artificiales y técnicas de entrenamiento. Particularmente se espera desarrollar modelos que permitan la generación de nuevas imágenes para gestos de la lengua de señas. Esto permitirá diversificar las bases de datos existentes, que suelen poseer pocos datos etiquetados. Esto podría ser aplicado en un futuro a otros dominios con datos etiquetados limitados.