

# Integración de Dinámicas Caóticas en Autocodificadores Convolucionales Generativos

M. Celeste Cebedio, Lucas A. Rabioglio y Luciana De Micco  
ICYTE, Depto. de Electrónica y Computación, Facultad de Ingeniería - UNMDP  
{celestecebedio,lucas.rabioglio,ldemicco}@fi.mdp.edu.ar

**Resumen**—En este trabajo se investiga el uso de mapas caóticos para el muestreo del espacio latente en autocodificadores convolucionales, con el objetivo de que estos puedan utilizarse como modelos generativos. Se analizan las distribuciones del espacio latente generado con los mapas y se evalúa la calidad de las imágenes sintéticas generadas, mediante las métricas Frechet Inception Distance e Inception Score.

**Palabras clave**—Caos, Autocodificador Generativo, Espacio Latente.

## I. INTRODUCCIÓN

Un autocodificador generativo convolucional (CAE) es una red neuronal que combina autocodificadores y capas convolucionales, aprendiendo una representación comprimida de los datos, llamada espacio latente. Esta representación permite reconstruir los datos originales con alta fidelidad y generar nuevos ejemplos similares. En aplicaciones de generación, el espacio latente desempeña un papel crucial al capturar las características esenciales de los datos de entrada, lo que facilita la síntesis de nuevos ejemplos con propiedades coherentes.

Tradicionalmente, se emplean autocodificadores variacionales (VAE) o redes generativas antagónicas (GAN) en tareas de generación, ya que los autocodificadores convolucionales presentan la desventaja de no permitir un control preciso sobre la distribución del espacio latente [1]. Esto dificulta la generación de nuevos puntos o el 'muestreo' eficiente del espacio latente [2]. Sin embargo, aunque los VAE y GAN superan ciertas limitaciones de los autocodificadores convolucionales, presentan algunas desventajas importantes. Los VAE, por ejemplo, requieren un entrenamiento más intensivo y es complejo, e incluso imposible en algunos casos, alcanzar errores de entrenamiento tan bajos como los obtenidos con autocodificadores convolucionales. Por otro lado, aunque las GAN cuentan con una convergencia teórica demostrada, su entrenamiento resulta complejo debido a la dificultad de equilibrar el rendimiento entre el generador y el discriminador. Además, tanto los VAE como las GAN demandan un elevado número de parámetros, lo que representa un desafío para su implementación en dispositivos con recursos limitados, donde la capacidad de procesamiento y memoria resulta crucial.

Este trabajo propone modelar la distribución del espacio latente de un autocodificador convolucional utilizando técnicas no lineales, con el objetivo de ajustar dicha distribución a la original y generar nuevas muestras en el espacio latente que permitan producir imágenes sintéticas.

## II. METODOLOGÍA PROPUESTA

La razón por la cual los autoencoders convolucionales no suelen emplearse como modelos generativos es que su espacio latente puede presentar cualquier distribución, lo que imposibilita realizar muestreos directos. En este trabajo, se propone utilizar el espacio latente del autoencoder como base para construir un generador caótico que produzca datos siguiendo la distribución aproximada del espacio latente, segmentada por tramos, suguiendo el método de Rogers et al. [3]. Este enfoque permite sintetizar un mapa caótico con una distribución invariante predefinida, basada en funciones lineales por tramos.

Por otro lado, se observa que las dimensiones del espacio latente presentan fuertes correlaciones, lo que indica una marcada dependencia entre ellas. Por este motivo, además de modelar la distribución individual de cada dimensión, se considera la generación de distribuciones correlacionadas.

## III. EVALUACIÓN DE LAS IMÁGENES GENERADAS

Para evaluar las imágenes ficticias generadas se emplean las métricas FID (Frechet Inception Distance) e Inception Score (IS). La FID compara las distribuciones de características de las imágenes generadas y las imágenes reales, proporcionando un valor que refleja la 'distancia' entre ambas distribuciones. Un FID más bajo indica mayor similitud entre las imágenes generadas y las reales, lo que sugiere que las imágenes sintéticas son más realistas. Por otro lado, el Inception Score mide la calidad visual y la diversidad de las imágenes, al evaluar la confianza con la que un modelo de clasificación reconoce objetos distintos en las imágenes generadas. Un Inception Score alto indica que las imágenes son variadas y claramente identificables en diferentes categorías, lo que implica una generación de alta calidad y diversidad.

## IV. RESULTADOS PRELIMINARES

Para evaluar el método propuesto, se entrenó un autocodificador convolucional con un espacio latente de dimensión 4 utilizando el conjunto de datos NSynth [4].

El entrenamiento generó una matriz de espacios latentes de dimensión  $4 \times 55.000$ . La Fig. 1.a muestra los histogramas condicionales entre combinaciones de dimensiones:  $D1$  vs  $D2$ ,  $D1$  vs  $D3$ ,  $D1$  vs  $D4$ ,  $D2$  vs  $D3$ ,  $D2$  vs  $D4$ , y  $D3$  vs  $D4$ , mientras que la Fig. 2.a ilustra las correlaciones entre pares de dimensiones. Se observa que el espacio latente no

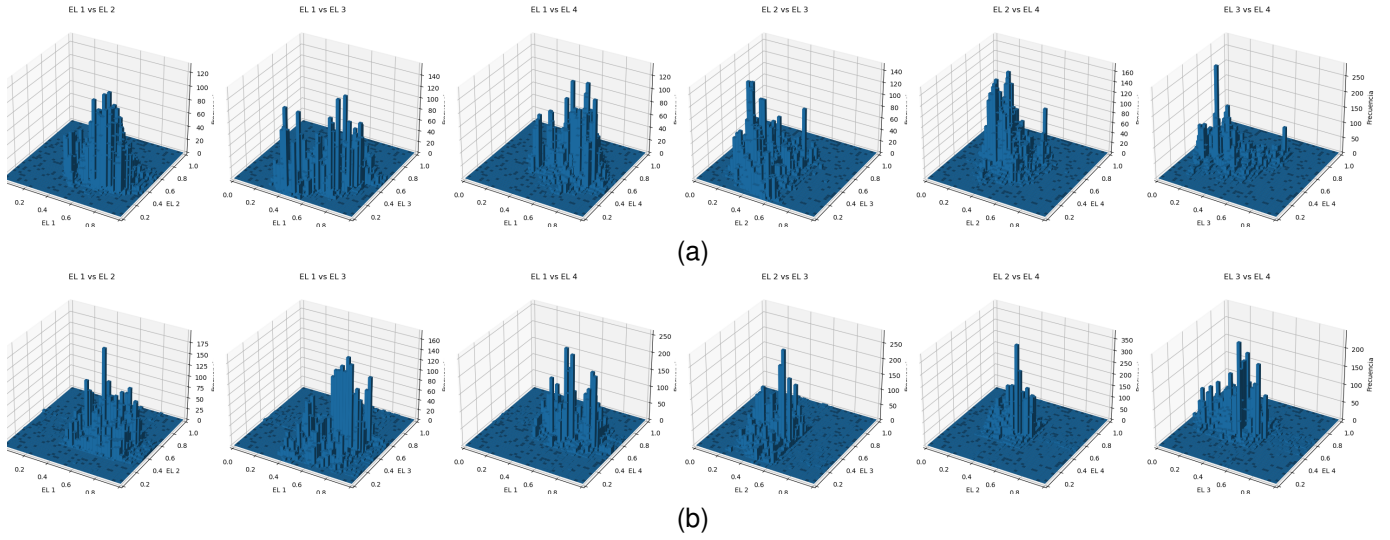


Figura 1. Distribuciones condicionales de espacio latente D=4. a) Espacio latente original. b) Espacio latente generado por mapas caóticos.

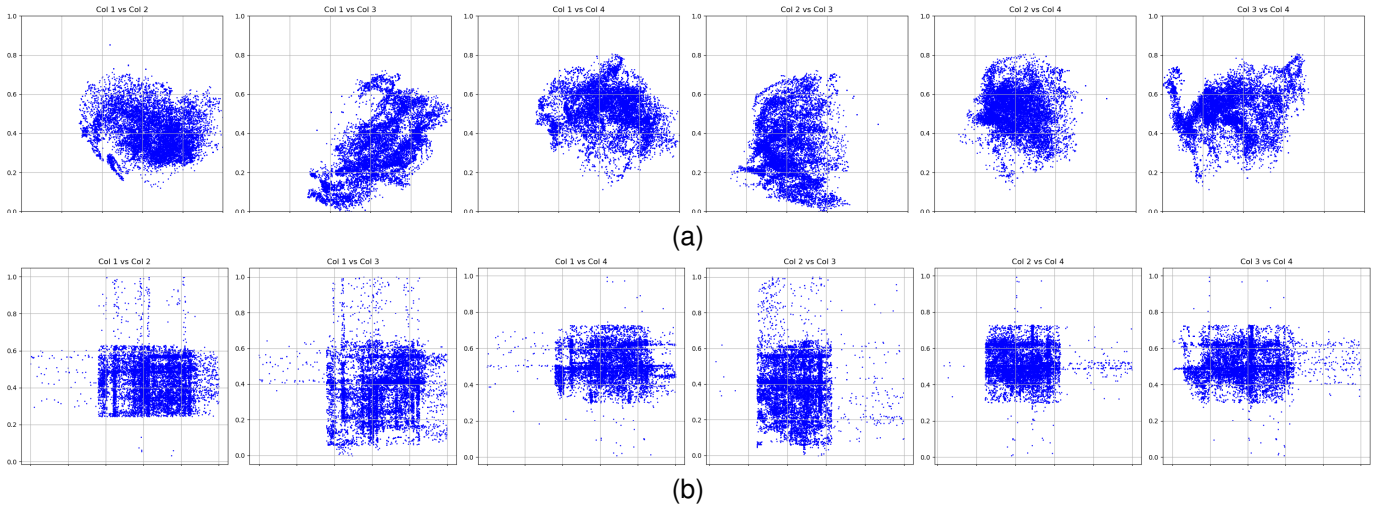


Figura 2. Puntos de espacio latente D=4. a) Espacio latente original. b) Espacio latente generado por mapas caóticos.

sigue una distribución gaussiana y sus dimensiones están correlacionadas. Usando el método propuesto, se generaron muestras a partir de mapas caóticos correlacionados. Las distribuciones condicionales resultantes se presentan en la Fig. 1.b y las correlaciones en la Fig. 2.b. Los resultados muestran que el método propuesto preserva de manera consistente la estructura y las correlaciones del espacio latente original.

La Tabla I resume las métricas obtenidas para distintos muestreos en CAE y VAE bajo las mismas condiciones. Se

Autocodificador	FID	IS
VAE	1.0043	1.9885 ± 0.0236
CAE mapa no corr B=50	0.5460	2.5759 ± 0.0278
CAE mapa no corr B=500	0.5432	2.8107 ± 0.0601
CAE KDE 50	0.5463	2.8278 ± 0.0382
CAE mapa corr B=5	0.4639	2.8764 ± 0.0517
CAE mapa corr B=10	0.4220	3.3594 ± 0.0651

Cuadro I

COMPARACIÓN DE FID E IS ENTRE DIFERENTES MODELOS.

observa que CAE supera a VAE. Además, aumentar los tramos para aproximar la distribución deseada (B) no mejora los resultados, indicando que mayor complejidad en los mapas no aporta beneficios, una cualidad deseable para reducir los requisitos de hardware en sistemas de bajos recursos.

## REFERENCIAS

- [1] Q. Xu, Z. Wu, Y. Yang, and L. Zhang, "The difference learning of hidden layer between autoencoder and variational autoencoder," in *29th Chinese Control And Decision Conference*, 2017.
- [2] T. White, "Sampling generative networks," *arXiv preprint arXiv:1609.04468*, 2016.
- [3] A. Rogers, R. Shorten, and D. M. Heffernan, "Synthesizing chaotic maps with prescribed invariant densities," *Physics Letters A*, vol. 330, no. 6, pp. 435–441, 2004.
- [4] J. Engel, C. Resnick, A. Roberts, S. Dieleman, D. Eck, K. Simonyan, and M. Norouzi, "Neural audio synthesis of musical notes with wavenet autoencoders," 2017.