

Métodos para súper-resolución de imágenes

Madrigal-Custodio Jesús A., Tevera-Ruiz Alejandro, Torres-Martínez Luis Á.

Departamento: Robótica y Manufactura Avanzada

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional

Resumen

En el presente documento se explican los fundamentos, metodología y proceso de implementación para el desarrollo de algoritmos de súper resolución bajo diversos enfoques con el objetivo...

Palabras Clave

Súper Resolución, Redes Convolucionales, Inteligencia Artificial

I. INTRODUCCIÓN

I-A. SRGAN

Las GAN's (Generative Adversarial Networks) son un tipo de redes cuyo funcionamiento está basado en la estimación de modelos generadores, esto es posible gracias al entrenamiento de un modelo *generador* (G) que obtiene la distribución de los datos y un modelo *discriminador* (D) el cual se encarga de estimar la probabilidad de que la muestra provenga del dataset de entrenamiento en lugar del modelo G . [1]

II. ANTECEDENTES

Bajo un enfoque *clásico*, existen tres formas de mejorar la resolución de una imagen:

- Amplificación de detalles existentes
- Suma de múltiples frames
- Único frame

Para el primero de ellos, se realiza una amplificación de las frecuencias altas (donde se encuentran los detalles existentes de la imagen) dada la variación local entre los píxeles vecinos. La amplificación de detalles existentes resulta bastante sencillo de aplicar. Sin embargo, ante imágenes con una cantidad considerable de ruido puede no ser la mejor opción a tomar. Además, al potencializar las frecuencias ya existentes de la imagen, el resultado estará definido por el detalle previo en la imagen de entrada.

El segundo de los métodos considera que el frame de alta resolución es el resultado de una secuencia de frames de baja resolución que permiten obtener las frecuencias altas de la imagen resultante para mejorar su resolución. Esto es conveniente cuando ya se cuenta con el conjunto de imágenes requeridas y se planea realizar una reconstrucción de la imagen en una mejor calidad.

Por otro lado, el tercer método basado en un único frame o imagen busca aproximar las frecuencias altas (detalles) que no se encuentra en la entrada del algoritmo y que evidentemente no puede obtenerse sólo amplificando las frecuencias altas como lo que ocurre con el primero de los métodos.

II-1. Interpolación

Para mejorar la calidad se busca aumentar la densidad de píxeles de la imagen con el objetivo de hacer la imagen más grande y mejorar sus detalles a partir de la predicción de píxeles que no se encuentran en la imagen visiblemente, pero que podrían aproximarse al buscar que se mantenga una consistencia en la imagen modificada de acuerdo a la vecindad de los píxeles.

Esto permite proponer el uso de algoritmos de interpolación que buscan predecir los píxeles vecinos y con ello aumentar la densidad de píxeles de la imagen de entrada. Con base en [2], dichos algoritmos pueden agruparse en dos categorías: adaptativos y no adaptativos. Los primeros cambian dependiendo de lo que se está interpolando (bordes o texturas suaves) pixel por pixel con el objetivo de minimizar los errores antiestéticos de los algoritmos de interpolación como el desenfoque o pérdida de detalles en regiones evidentes. Ejemplos de ellos pueden ser los softwares de licencia como *Qimage*, *PhotoZoom Pro*, *Genuine Fractals*, etc.

Mientras que los métodos no adaptativos tratan todos los píxeles por igual dada la predicción de un pixel central de acuerdo a sus píxeles adyacentes. Esto involucra que entre más vecinos se consideren en la interpolación, una mejor aproximación se tendrá del pixel a predecir, pero de manera proporcional aumentarán los recursos computacionales necesarios. Dentro de los algoritmos se incluyen: *vecino más cercano*, *bilineal*, *bicúbica*, *spline*, entre otros.

A continuación se describirán algunos de los algoritmos no adaptativos para interpolación que serán utilizados en los diferentes métodos para *Súper Resolución*:

- **Vecino más cercano** - Dado un pixel considera sólo un pixel adyacente para la interpolación, lo que resulta en un menor tiempo de procesamiento pero resultados poco consistentes al observar al conjunto de pixeles interpolados.
- **Bilineal** - Considera una vecindad 2x2 correspondiente al pixel a predecir con su correspondiente promedio ponderado de acuerdo a la distancia del pixel desconocido. Esto da como resultado un aspecto más suave que el vecino más cercano.
- **Bicúbica** - Valora una vecindad 4x4 de pixeles conocidos para la predicción del pixel central considerando el mismo procedimiento de la interpolación bilineal. Como resultado, se alcanzan imágenes más nítidas que los métodos anteriores. Logrando así un equilibrio entre la calidad de salida y el tiempo de procesamiento. Lo anterior promueve que sea un estándar en muchos programas de edición de imágenes, controladores de impresoras e interpolación en cámaras.

NOTA: Agregar comparativo de interpolaciones con un parche y escalado x2

II-2. Example Based Super Resolution

En particular, [3] propone un parchado de la imagen rescalada a partir de un conjunto de entrenamiento o diccionario de parches en pares de alta y baja resolución. Dichos parches permiten construir una imagen con frecuencias altas que no están en la imagen de entrada con el objetivo de sumar la imagen original interpolada con las frecuencias altas que buscan mejorar su resolución al realzar sus detalles. En la Figura II.1 puede observarse de manera específica el algoritmo propuesto por [3].

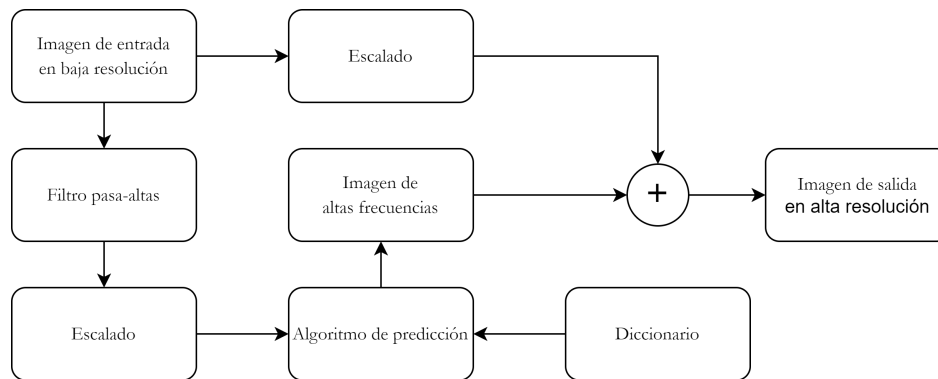


Figura II.1. Algoritmo de súper resolución

Observe que la imagen de entrada debe pre-procesarse mediante la aplicación de un filtro pasa-altas y el escalado mediante alguna técnica de interpolación.

II-A. Redes Convolucionales

Las Redes neuronales convolucionales son un tipo de redes neuronales artificiales donde las *neuronas* corresponden a campos receptivos de una manera muy similar a las neuronas en la corteza visual primaria (V1) de un cerebro biológico. Este tipo de red es una variación de un perceptrón multicapa, sin embargo, debido a que su aplicación es realizada en matrices bidimensionales, son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones.

II-B. SRCNN

Como se menciona en [3],

III. IMPLEMENTACIÓN

IV. RESULTADOS

V. DISCUSIÓN

VI. CONCLUSIONES

buenas buenas

REFERENCIAS

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *ArXiv*, 06 2014.
- [2] S. McHugh, "Digital image interpolation," 2005. [Online]. Available: <https://www.cambridgeincolour.com/tutorials/image-interpolation.htm>
- [3] W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *Image-Based Modeling, Rendering, and Lighting*, 2002.