

El *Global Average Pooling* (GAP) es una técnica de reducción espacial utilizada principalmente en redes neuronales convolucionales (CNN) para reducir la dimensionalidad de las características extraídas de las capas convolucionales sin agregar demasiados parámetros adicionales, lo que ayuda a evitar el sobreajuste.

¿Cómo funciona?

Después de que una imagen pasa a través de varias capas convolucionales y de pooling, se generan mapas de características de alto nivel que capturan patrones relevantes. En lugar de aplanar estos mapas (como se haría para conectarlos a una capa completamente conectada), el GAP aplica una operación de promediado a cada uno de los mapas de características.

Supongamos que tienes un tensor de salida de forma (h, w, d) donde:

- h es la altura del mapa de características,
- w es el ancho del mapa de características,
- d es el número de filtros o profundidad.

El GAP calcula el promedio de todos los valores en cada uno de los mapas de características a lo largo de las dimensiones espaciales (h y w), y genera un vector de tamaño d . Por lo tanto, reduce cada mapa de características a un solo valor promedio.

Beneficios de GAP:

1. **Reducción de parámetros:** A diferencia de las capas densas o fully connected, que añaden muchos parámetros, el GAP no introduce nuevos pesos, haciendo la red más simple y menos propensa al sobreajuste.
2. **Mejor interpretable:** Al reducir cada mapa de características a un solo valor, se mantiene una conexión más clara entre los filtros de la última capa convolucional y las clases predichas.
3. **Eficiencia computacional:** GAP simplifica las operaciones y puede acelerar el entrenamiento, especialmente en redes profundas.

Ejemplo:

Si tienes un tensor de salida de tamaño $(7, 7, 1024)$ de una CNN, el GAP convierte esto en un vector de tamaño $(1024,)$, tomando el promedio de los valores en cada uno de los 1024 mapas de características de tamaño 7×7 .

En resumen, *Global Average Pooling* reduce el espacio dimensional de las características conservando la estructura de los canales, facilitando la transición hacia la capa de clasificación final, normalmente una capa softmax para la clasificación.