

Experimento 0

Introducción

En este experimento pretendemos comprobar si la red de clasificación binaria propuesta en “A Preliminary Study on Deep Transfer Learning Applied to Image Classification for Small Datasets” se ajusta a nuestro problema de clasificación múltiple y si es mejor entrenar una red mediante scratching o dividiendo el set de datos y aplicando transfer learning.

Metodología

Para saber si se cumplen las premisas anteriores:

1 - Para comprobar la primera hipótesis, replicamos la red (el código de la implementación se encuentra en el repositorio de github) y le hacemos una ligera modificación sustituyendo la capa de salida.

En el estudio, la capa de salida tiene solo una neurona con una función de activación sigmoide, lo cual es apropiado para una clasificación binaria, nosotros usaremos una capa de 7 neuronas (una por clase) y una función de activación softmax.

La red recibirá el nombre de CNN_SOCO y la entrenamos con el optimizador sugerido en el artículo, un detalle importante es que el lr está modificado tal y como se dice en el artículo usando como lr 0.0001 en lugar del por defecto.

2 - Para saber si se cumple la segunda hipótesis, dividiremos el set de datos en subconjuntos y compararemos el resultado con el obtenido en el punto anterior.

Para ello separaremos los datos en dos sets:

- `tf_data`: Consiste en la información que se usará para entrenar completamente el modelo. Consiste en un 60% del total del set de datos.
- `train_data`: Consiste en la información que se usará para entrenar las capas de salida del modelo.

Conclusión

En la siguiente tabla podemos ver los resultados de precisión y pérdida de la red.

| Red | Optimizador | Acc | Loss |
|-----------------|-----------------|--------|--------|
| CNN_SOCO | RMSPProp | 0.9283 | 0.3301 |

El resultado parece muy bueno y se puede considerar que por lo tanto la red se ajusta de forma correcta a nuestro problema.

En cuanto al transfer learning los resultados obtenidos son :

| Red | Modo de entrenamiento | Optimizador | Acc | Loss |
|-----------------|-------------------------------|-----------------|--------|--------|
| CNN_SOCO | 100 % del set de datos | RMSPProp | 0.9283 | 0.3301 |
| CNN_SOCO | 60% del set de datos | RMSPProp | 0.8036 | 0.6371 |
| CNN_SOCO | Transfer learning | RMSPProp | 0.9333 | 0.2542 |

Efectivamente, tal y como describe el artículo, la red ha mejorado su precisión con respecto a la misma entrenada por scratching.

Bibliografía:

[1] - "A Preliminary Study on Deep Transfer Learning Applied to Image Classification for Small Datasets"