CLASSIFICACIÓ D'IMATGES CNN AMB KERAS

Esther Sanromà Segura IT ACADEMY

Abstract

En aquest projecte, s'ha utilitzat una xarxa neuronal convolucional (CNN) per classificar imatges del dataset "Places" mitjancant la llibreria Keras. El dataset es va obtenir de Kaggle i conté imatges de diferents llocs al voltant del món. En aquesta ocasió, s'ha utilitzat un subconjunt de 6380 imatges per entrenar el model, 1821 per validar-lo i 924 per provar-lo. S'han creat diversos models de xarxes neuronals i s'ha ajustat el millor d'ells per obtenir la millor precisió. Malgrat això, els resultats no han estat favorables i s'ha conclòs que s'han de revisar els paràmetres i augmentar la quantitat de dades per millorar l'entrenament.

Introducció

La classificació d'imatges és una tasca clau en molts camps, com la vigilància, la robòtica o la visió per computador. En aquest projecte, es pretén classificar imatges del dataset "Places" utilitzant una xarxa neuronal convolucional (CNN) mitjançant la llibreria Keras. Això permetrà classificar correctament una imatge segons el lloc que representa.

State of Art

En els últims anys, les xarxes neuronals convolucionals s'han convertit en la principal eina per a la classificació d'imatges. Això és degut a la seva eficàcia en la detecció de patrons i característiques en les imatges.

Els models de CNN són capaços de classificar amb precisió imatges de diferents categories, gràcies a la seva capacitat per aprendre a través de les capes de la xarxa.

Metodologia

En aquest projecte, s'ha utilitzat una xarxa neuronal convolucional (CNN) per classificar les imatges del dataset "Places". S'ha fet servir la llibreria Keras per a la implementació de la CNN i la manipulació de les imatges.

En primer lloc, s'ha descarregat el dataset "Places" de Kaggle, que conté més de 180.000 imatges de diferents llocs al voltant del món. Per reduir la complexitat del problema i disminuir el temps d'entrenament de la xarxa neuronal, s'ha utilitzat un subset del dataset original, anomenat "places_reduced". Aquest subset conté només 23 carpetes, de la "a" a la "w", que contenen imatges etiquetades. Després, s'ha seleccionat la carpeta "a" del dataset "places reduced" per a la classificació d'imatges. S'ha dividit aquesta carpeta en tres conjunts: train, test i validació, amb una proporció del 70%, 20% i 10%, respectivament. En total, hi ha 6380 imatges a la carpeta de train, 1821 a la de test i 924 a la de validació.

Per preprocessar les imatges, s'han redimensionat totes les imatges a 118x118 píxels amb 3 canals (RGB) utilitzant una funció de Keras. Això ajuda a disminuir la quantitat de dades a processar i millora l'eficiència del model.

Per a la implementació de la xarxa neuronal, s'ha utilitzat la llibreria Keras. S'han creat diferents models de CNN, variant el nombre de capes, el nombre de neurones per capa i els hiperparàmetres de a xarxa neuronal, per aconseguir la millor precisió en la classificació d'imatges. S'ha utilitzat l'ImageGenerator de Keras per dividir les imatges en els tres conjunts (train, test i validació).

Després d'haver entrenat el model de CNN, s'ha fet un ajust final del model per optimitzar els paràmetres del model.

Després d'aquest ajust, s'ha calculat el percentatge de precisió del model utilitzant les dades de test per avaluar la precisió del model en imatges que no ha vist abans.

Per a la visualització de l'exactitud (precisió) i la pèrdua (loss) en el temps, s'ha creat una gràfica del model. Això ajuda a identificar si el model està sobreajustant (overfitting) o no.

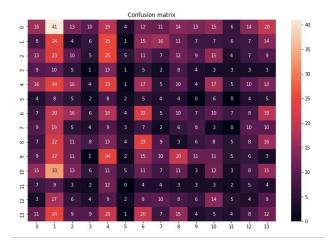
Finalment, s'ha creat una matriu de confusió per avaluar el rendiment del model en la classificació de cada classe. Això permet identificar quines classes es classifiquen amb major o menor precisió i quines són les que presenten més dificultats per a la xarxa neuronal.

Resultats

Els resultats obtinguts en aquest projecte no han estat favorables per a la classificació de les imatges. El model de xarxa neuronal convolucional (CNN) que s'ha creat no ha aconseguit una precisió alta en la classificació de les imatges. Després d'entrenar diferents models de xarxes neuronals convolucionals (CNN) amb diferents paràmetres, s'ha observat que el model amb 4 capes convolucionals, regulador de Kernel L2, BatchNormalization i Dropout del 20% ha obtingut una precisió del 44,65% en la classificació de les imatges de la carpeta "a".

	model	num_conv_layers	batchnorm	dropout	L2	epochs	test_loss	test_accuracy
0	model26.h5	4	False	-	L2	100	3.606509	0.332784
1	modell27.h5	4	False	0.2	L2	100	2.033662	0.423943
2	model30.h5	4	True	0.2	L2	200	3.779453	0.446458
3	model01.h5	8	True	0.5	-	70	1.907737	0.425041

A més, s'ha observat que el model tendeix a sobreajustar-se (overfitting), ja que l'exactitud en el conjunt d'entrenament és molt alta, mentre que l'exactitud en el conjunt de validació és relativament baixa. Tot i que els resultats no han estat satisfactoris, es pot concloure que aquest projecte ha servit per aprendre i experimentar amb la classificació d'imatges utilitzant una xarxa neuronal convolucional (CNN) i la llibreria Keras. S'ha pogut comprendre l'estructura bàsica d'una xarxa neuronal convolucional (CNN) i l'efecte que tenen els diferents paràmetres en la precisió de la classificació.



Conclusió

En conclusió, aquest projecte ha estat una oportunitat per experimentar amb la classificació d'imatges utilitzant una xarxa neuronal convolucional (CNN) i la llibreria Keras. S'ha pogut aprendre sobre l'estructura bàsica d'una xarxa neuronal convolucional (CNN), la importància del pre-processament de les dades, la selecció dels paràmetres adequats i com avaluar la precisió del mode Encara que els resultats obtinguts no han estat satisfactoris, aquest projecte ha estat un aprenentatge valuós per entendre el funcionament d'una xarxa neuronal convolucional (CNN) i com aplicar-la en la classificació d'imatges.

Referències

Keras: https://keras.io/
 TensorFlow: https://www.tensorflow.org/

Kaggle: https://www.kaggle.com/
 Places: http://places.csail.mit.edu/
 user/