

Resum:

IDENTIFICACIÓ DE MELANOMA EN IMATGES DE LESIONS CUTÀNIES

Actualment, el càncer de pell és el tipus de càncer més freqüent en el món. La gran majoria de morts a causa d'aquest càncer es deuen al melanoma, que tot i representar només un 1% dels casos, és el responsable del 75% de les morts per càncer de pell.

Avui en dia, el diagnòstic del melanoma el duen a terme els dermatòlegs, que avaluen cadascuna de les pigues d'un pacient per identificar possibles lesions atípiques que tinguin més probabilitat de ser melanoma. Si es creu que podria tractar-se de càncer de pell, es faria una biòpsia de pell i s'enviaria la mostra per examinar-la sota el microscopi. Igual que amb altres càncers, una detecció precoç i precisa, amb l'ajuda potencial de la ciència de dades, implica una millor prognosi de la malaltia. Per tant, aquest tipus de projectes que impliquen intel·ligència artificial podrien ajudar a una diagnosi a temps i, en conseqüència, un millor pronòstic de la malaltia, ja que com més aviat es detecta la malaltia més aviat es pot començar el tractament.

A més a més, el dia d'avui els metges tenen una gran quantitat d'imatges per analitzar, tant de lesions cutànies com d'altres lesions com podrien ser fractures d'ossos (radiografies, ressonàncies magnètiques...), infeccions pulmonars, tomografies computaritzades del cervell (per tal de diagnosticar possibles malalties com ara l'alzheimer o el Parkinson), i, per tant, un algoritme al qual se li poden entrar una gran quantitat d'imatges i en fa una predicció acurada seria de gran ajuda.

En aquest projecte he utilitzat dues tècniques d'intel·ligència artificial per distingir entre una lesió benigna o maligna en imatges de lesions cutànies: aprenentatge automàtic supervisat i transfer learning.

APRENTATGE AUTOMÀTIC

Per fer la part d'aprenentatge automàtic supervisat, el primer que he hagut de fer ha estat la partició del dataset original d'imatges en dues carpetes: train (70% de les imatges) i test (30% restant). Posteriorment, he hagut de fer una altra subdivisió dins la carpeta de train per crear una altra carpeta de train amb el 70% de les imatges i una altra de validació amb el 30% de les imatges. El conjunt de train es fa servir per entrenar l'algoritme, mentre que el de validació serveix per validar-lo, és a dir, per anar testejant les característiques i anar-les modificant. El conjunt de test serveix per testejar el nostre model un cop validat l'algoritme (en aquest conjunt ja no s'hi poden fer modificacions).

Un cop he tingut les particions fetes, he fet un preprocessat a les imatges. Aquest preprocessat ha consistit a treure'ls-hi els pèls, ja que això podria haver afectat negativament els resultats, sobretot en realitzar la segmentació de la zona de la piga.

El següent pas ha estat el balancejament de les imatges. Això ha estat degut al fet que en les imatges inicials hi havia un 1% aproximat d'imatges malignes, quan hauria de ser un 50% benignes i 50% malignes per tal que l'algoritme aprengués correctament.

Per fer el balancejament el que he fet ha estat crear imatges sintètiques de cada imatge maligna fins a assolir el percentatge adequat. Per crear aquestes imatges semblants els hi he modificat les característiques de color (contrast, saturació i brillantor), els hi he afegit soroll i els hi he fet transformacions geomètriques. Tot i això, he vigilat de no fer massa modificacions alhora per tal de no sobrecarregar la imatge i que no quedés realista.

Després de fer el balancejament, ja hem pogut començar la validació del nostre model. Això consisteix en diversos passos: la correcció del fons d'algunes imatges, la segmentació, les combinacions de característiques i el classificador.

El primer pas ha estat una petita correcció del fons d'algunes imatges les quals al ser rodones, tenien un fons completament blanc que podien afectar en la segmentació. En la imatge en BW inicial, els píxels amb número 0 són negres i els 255 blancs, pel qual he realitzat proves amb 245, 250 i 255 per mirar quin era el píxel més adequat on fer el canvi de fons. En el lloc del fons blanc hi posava una mitjana del color de la imatge inicial, per tal que al fer la segmentació no ho considerés fons.

Un cop corregit això, hem segmentat les imatges mitjançant la funció `imbinarize`, que escull un llindar en el qual decideix si el píxel forma part de la regió o no.

Amb les imatges ja segmentades, hem dut a terme vàries proves de combinacions de característiques de forma mitjançant la funció `regionprops`, que n'hem fet un total de 7. D'entre totes, hem escollit el millor classificador, que ha estat l'Adaboost, ja que era el que donava millors resultats en la majoria dels casos. Dins els resultats proporcionats per aquest classificador hem escollit la millor combinació de característiques, que ha estat `Eccentricity + BoundingBox + EulerNumber`.

Amb totes les propietats escollides, i, per tant, amb l'algoritme validat, hem testejat l'algoritme i ens ha donat una `accuracy` de 0.8054.

TRANSFER LEARNING

El transfer learning consisteix bàsicament a fer servir una xarxa preentrenada i substituir-ne les capes finals (en les quals l'algoritme aprèn les característiques més específiques) per noves capes per aprendre del nostre dataset d'imatges.

Per fer això, primer hem hagut d'ajuntar totes les imatges (amb el preprocessat i el balancejament) en una sola carpeta, ja que després l'algoritme fa automàticament la partició entre `train` i `test` en el percentatge que li indiquis (70 i 30 en el nostre cas).

El segon pas és carregar la xarxa preentrenada (nosaltres n'hem provat 3: `GoogleNet`, `ResNet-18` i `SqueezeNet`) i mirar la mida d'imatges d'entrada que requereix. Un cop hem sabut la mida que requereix, hem fet un `resize` de totes les imatges.

Ara ja tenim les imatges llestes. El següent pas ha estat substituir les capes finals (l'última capa d'aprenentatge i la capa de classificació final) per noves capes sense etiquetes (es posaran les etiquetes automàticament quan entrenem la xarxa).

En quart lloc, hem congelat les capes inicials i hem tornat a connectar totes les capes. Un cop preparada la xarxa, ja podem entrenar-la amb les opcions i els conjunts de `train` i `test` que li passem.

Per acabar, ja amb la xarxa entrenada, només ens falta fer la classificació i visualitzar els resultats, que han estat els següents: `GoogleNet` = 0.9767; `ResNet-18` = 0.8960, `SqueezeNet` = 0.9720.

Com podem veure, tot i tenir molts bons resultats d'accuracy, no sabem d'on provenen, és a dir, en quines característiques s'ha fixat la xarxa per fer la classificació. Aquí entra el GradCAM.

GRADCAM

El mètode GradCAM consisteix principalment en un mapa d'activació de classe ponderada pel gradient (d'aquí el nom). Produeix un mapa de calor que destaca les regions importants d'una imatge emprant els gradients de l'objectiu de la capa convolucional final.

Aquest mètode és una tècnica popular de visualització útil per comprendre com la xarxa neuronal ha arribat a prendre la seva decisió de classificació. Sobretot en cas d'error, és un mètode útil per entendre on es troba la dificultat de la xarxa convolucional, a més a més de fer l'algoritme més transparent.

Per tant, nosaltres hem creat un mapa de calor per algunes de les imatges, tant les ben classificades com les mal classificades per saber en què es fixava la xarxa en els dos casos.

Fent un resum, en general, en les imatges ben classificades podem observar una molt bona segmentació de la zona de la lesió, amb una lleugera diferència entre les benignes i les malignes: en les benignes està molt ben segmentada la regió de la lesió, el qual podria indicar que es fixa en les propietats de l'interior de la pigada (color, textura), i en el contorn. En canvi, en les imatges malignes veiem com la zona on es focalitza són zones dels contorns on es perd la circularitat o on hi ha canvis de textura.

Per altra banda, en les imatges mal classificades, en general segueix un patró: no identifica correctament la zona de la lesió, especialment en les imatges amb el fons blanc, el qual identifica com a zona principal.

CONCLUSIONS

Tenint en compte els objectius inicials, que eren la diagnosi de melanoma mitjançant imatges de lesions cutànies i estudiar tècniques d'intel·ligència artificial explicables, s'han assolit satisfactòriament.

Hem pogut veure com la intel·ligència artificial pot ajudar a diagnosticar el melanoma de manera precoç mitjançant diferents tècniques com l'aprenentatge supervisat o el transfer learning.

Tot i els resultats obtinguts, es podrien fer moltes millores, començant per partir amb un conjunt de dades més ben balancejat, fet que ens hagués proporcionat millors resultats en tenir més exemples de cada classe. A més a més, per la part d'aprenentatge supervisat es podrien introduir millores mitjançant més comprovacions de combinacions de característiques i optimitzant els classificadors. Si haguéssim disposat de les imatges segmentades, podríem haver comprovat si nosaltres la fèiem correctament i, per tant, hauríem millorat l'algoritme.

De cara a la part de transfer learning, podríem haver provat més xarxes, tot i que trobo que les millores principals s'haurien de fer de cara a la seva explicabilitat, ja que tot i que amb el GradCAM podem fer-nos una idea d'en què s'ha basat l'algoritme, es podria trobar una manera de saber en quines característiques específiques s'ha basat. A més a més, podríem haver canviat els fons blanc de les imatges rodones tal com s'havia fet en la part d'aprenentatge automàtic, per evitar els errors que ha provocat.

Respecte als objectius i treballs futurs, es podrien encarar d'una manera més ambiciosa provant més tècniques d'intel·ligència artificial o, a part d'identificar si es tracta d'una lesió benigna o maligna, identificar-ne també el tipus.