

Treball final de grau

Estudi: Grau en Disseny i Desenvolupament de Videojocs

Títol: Tècniques de *Deep Learning* per la síntesi d'imatges. Aplicació a la restauració d'imatges.

Document: Resum

Alumne: Bryan Miguel Perez Ruchat

Tutor: Dr. LLADO BARDERA, XAVIER

Departament: ARQUITECTURA I TECNOLOGIA DE COMPUTADORS

Àrea: ARQUITECTURA I TECNOLOGIA DE COMPUTADORS

Convocatòria (mes/any) Juny 2022

Introducció i objectius

Aquest treball de fi de grau té com a principal objectiu entendre el funcionament de les xarxes neuronals i veure com s'apliquen en diferents formats, imatges, vídeos i jocs. Per entendre'n el funcionament, primer s'ha començat estudiant com funcionen amb imatges. El tema que s'ha decidit analitzar en aquest TFG és l'ús de les xarxes neuronals en la restauració d'imatges antigues així com la seva aplicació als videojocs, en concret en la capacitat de posar color a imatges/vídeos en blanc i negre. Els objectius d'aquest TFG han estat:

- Dissenyar i implementar una estratègia de *Deep Learning* per corregir soroll i per poder restaurar la qualitat de les imatges antigues.
- Dissenyar i implementar una estratègia *Deep Learning* per pintar imatges, cercar un mètode que permeti posar color a les imatges en blanc i negre.
- Realització del conjunt d'experiments a les dues xarxes, analitzant els paràmetres per tal d'optimitzar i obtenir el millors resultats possibles.
- Treballar el pintat d'imatges en vídeos, investigar com utilitzar la xarxa que s'ha emprat per introduir color a les imatges i utilitzar-ho en vídeos, comprovar si s'han de realitzar canvis.
- Realització de la *demo*, utilitzant *Unity*, intentat portar colors d'un altre joc, en aquest cas seria un similar al *Forza Horizon 5* per demostrar el seu ús en el camp dels videojocs.
- Anàlisi de resultats, amb avaluacions quantitatives i qualitatives, per tal de poder comparar els mètodes i observar quines solucions han aportat millors resultats.
- Redacció de la memòria, reunir tota la informació així com il·lustrar els resultats obtinguts en un document.

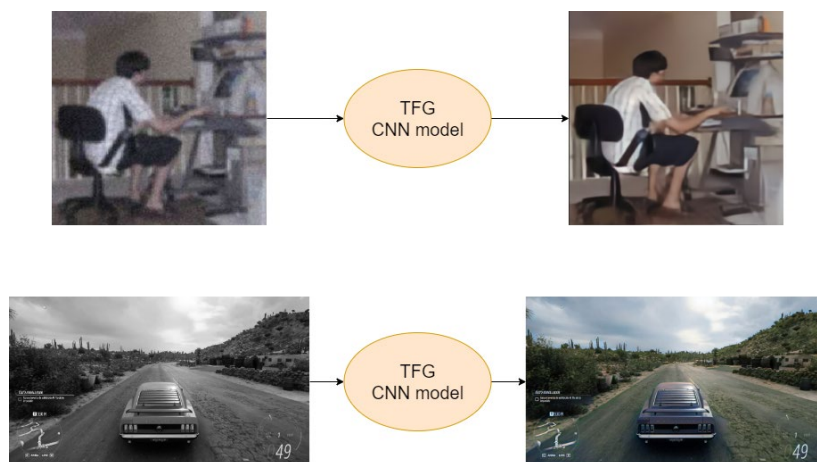


Figura 1 Representació de l'objectiu del TFG. A) Millora de la qualitat d'imatges B) Posar colors a les imatges

Desenvolupament

Les tècniques de *Deep Learning* usades per desenvolupar el projecte, són xarxes neuronals. Les xarxes neuronals són combinacions de neurones, simulant el funcionament del cervell humà per tal de solucionar un problema que no es té cap algoritme per solucionar-los. Per treballar en imatges, s'utilitzaran *CNNs* que són un tipus de xarxa neuronal en que es condensa la informació de la imatge usant l'operació de convolució. Finalment, per tal de poder generar les imatges i no tenir només la informació comprimida en poques dades s'utilitzaran *GANs* que permet codificar la imatge i descodificar-la d'una manera diferent per tal que al descodificar-la es té la imatge resultant i, aquesta es passa per un discriminador que l'intenta diferenciar d'imatges reals.

Per aquesta treball, al restaurar les imatges s'ha centrat principalment en la eliminació del soroll de les imatges. Per tal de revisar si seria possible, per pintar també s'ha començat amb imatges per tal d'entendre el funcionament de la xarxa i trobar potencials limitacions. Amb aquestes tècniques, es podria arribar a portar imatges de fa anys enrere a una qualitat similar a la que es té ara mateix i, en el cas que es pugui, posar color a les que no en tinguin.

Un cop es tenen els coneixements necessaris al treballar en imatges s'ha analitzat quin procés es pot dur a terme per tractar els vídeos, arribant a posar color a un vídeo del joc *Forza Horizon 5* que prèviament s'havia passat a blanc i negre. Amb aquest sistema s'han pogut provar entrenant la mateixa xarxa amb diferents escenaris i vehicles per veure el seu comportament.

Amb el tractament de vídeos acabat s'ha passat a crear una *demo* en *Unity*. L'objectiu era tenir algunes zones que siguin similars a les del *Forza* per tal de provar de transferir-hi els colors. S'ha requerit investigar com es duria a terme la importació de la xarxa a *Unity* i amb quins paràmetres es pot treballar per tal d'obtenir el millor resultat.

Un cop s'ha vist que es possible treballar amb els diferents formats, s'han fet experiments en anàlisis quantitativus i qualitativus per tal de poder comparar els resultats. Amb aquesta comparació es pot definir quins són els millors paràmetres per tal de solucionar el problema en concret.

Resultats

Per solucionar el problema de l'eliminació del soroll s'han usat sets de 500, 1000 i 2000 i 400 imatges tant de imatges antigues a color, en blanc i negre i actuals. Per entrenar la part de pintar les imatges, s'ha usat el joc *Forza Horizon 5* en vídeos de 20 minuts que s'han passat a imatges.

Les següents figures 2, 3 i 4, mostren alguns dels resultats obtinguts utilitzant les *CNNs* entrenades dins del TFG.

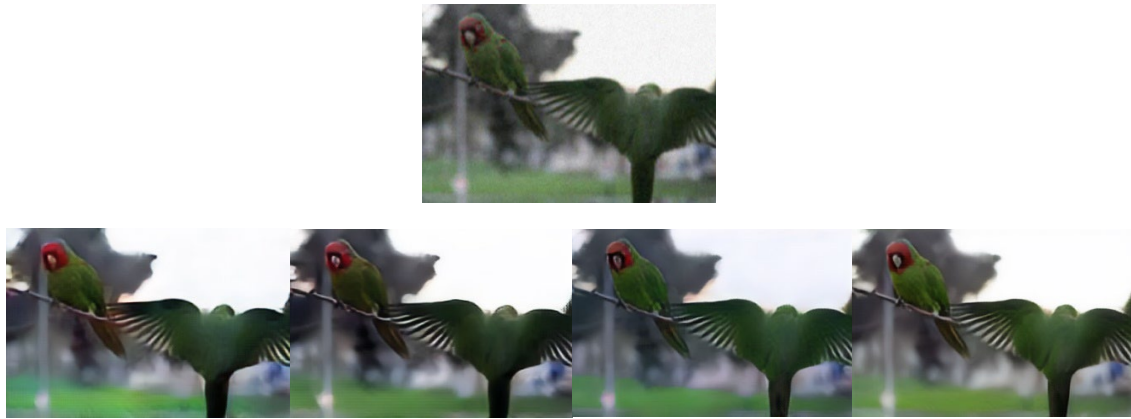


Figura 2 Resultats amb les comandes recomanades Bringing Old Photos Back To Life. a) Imatge d'entrada b) resultat en 500 imatges d'entrenament c) resultat en 1000 imatges d'entrenament d) resultat en 2000 imatges d'entrenament e) resultat en 4000 imatges d'entrenament

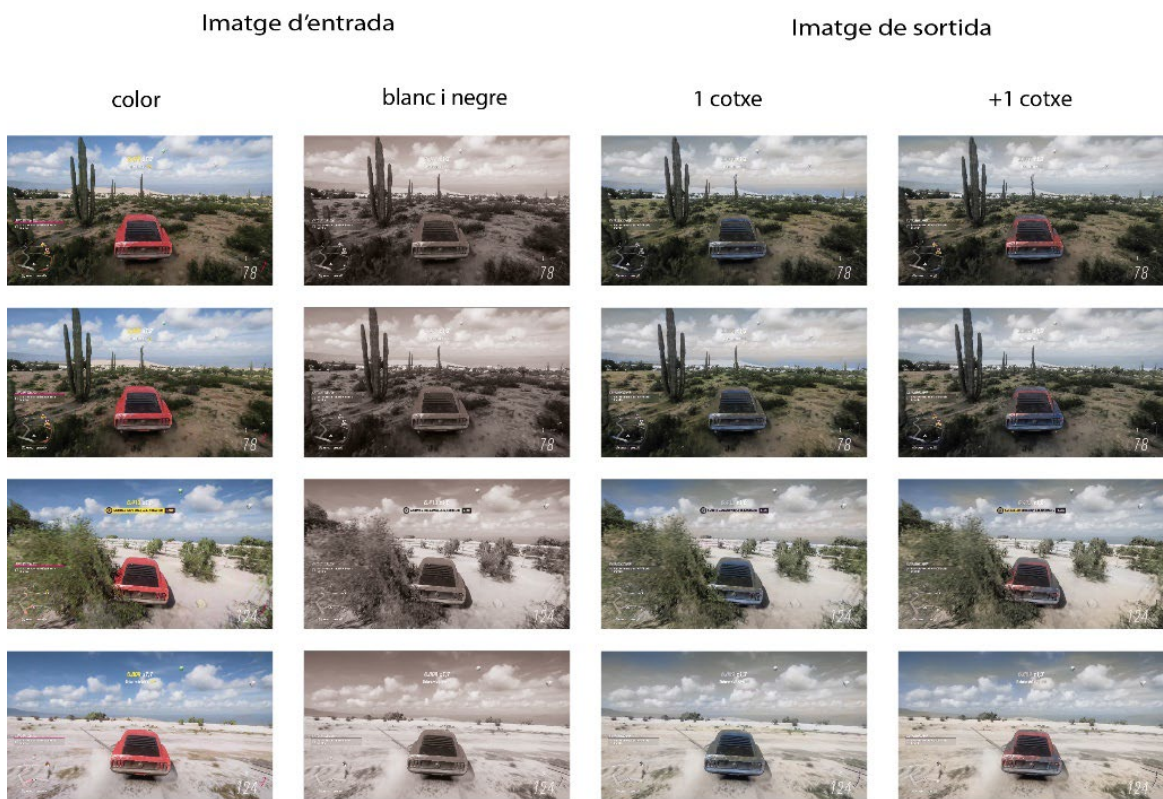


Figura 3 Frames del joc Forza Horizon 5 en color, blanc i negre, resultat de l'entrenament amb un cotxe i un escenari i resultat de l'entrenament de 3 cotxes i tres escenaris diferents



Figura 4 Processat en temps real dins del joc, passant la xarxa neuronal amb dimensió de 512x512 píxels. A) Xarxa entrenada amb 1 cotxe i 1 escenari b) Xarxa entrenada amb diferents cotxes i diferents escenaris.

Amb aquests resultats qualitius que també van acompanyats en resultats quantitius, que comparen les imatges resultants en les imatges d'entrada. Aquests resultats mostren que hi ha la possibilitat de fer una síntesi de colors per als jocs podent-se emprar, per exemple, en sistemes de *cloud computing* per tal de reduir la quantitat de dades que s'envien cap als jugadors. A part de reduir la quantitat de dades, també podria ser útil per mostrar diferents escenaris a diferents jugadors només necessitant canviar la xarxa, de manera que no es necessiti redissenyar tot l'escenari un segon cop.

Conclusions i treball futur

Amb el que s'ha desenvolupat, s'han pogut complir tots els objectius que s'havien fixat en un inici. Aquest treball ha sigut clau per entendre el funcionament de les xarxes neuronals, descobrir un conjunt d'arquitectures i com desenvolupen la feina cadascuna d'elles. Una part molt interessant ha sigut veure que es poden exportar aquestes xarxes i importar-les a *Unity* per poder fer tot tipus de síntesis d'imatge, en aquest cas transferir els colors de un joc, el *Forza*, a una petita *demo*.

De treball futur, han quedat un conjunt de tasques que podrien ser un nou TFG per algun futur estudiant:

- Revisar el treball recent *Bringing Old Films Back to Life* desenvolupat pels que han realitzat el *Bringing Old Photos back To Life*. Aquest mètode funciona directament amb vídeos i seria interessant revisar el seu funcionament en les proves fetes dins del TFG.
- Una altra qüestió que es podria millorar com a treball futur és el funcionament de la xarxa integrada dins del joc de *Unity*. Trobant un punt d'equilibri entre velocitat i qualitat de les imatges sintetitzades.
- Un altre punt interessant, seria l'estudi de com integrar aquestes eines dins d'una plataforma *cloud computing*. Per saber quines dades es passarien al client i quina part, si es que en tindria alguna el client, analitzant les diferents opcions i veient quins avantatges i inconvenients té cadascuna.