MD5: Progressive Growing of GANs

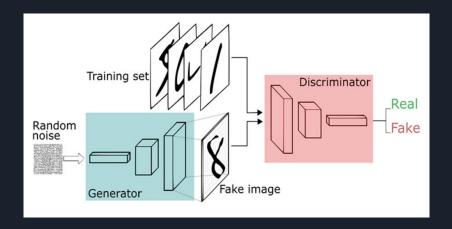
Progressive Growing of GANs for Improved Quality, Stability, and Variation

> Github Link Paper Link

Juan Carlos Soriano Valle - 1493037 Jordi Gimenez Perez - 1493035

Introducció

- Les arquitectures GAN (Generative Adversarial Networks) consisteixen en dues xarxes neuronals, la Generativa i la Discriminativa. La primera genera mostres artificials d'allò que vulguem i la segona determina si aquelles mostres generades són suficientment bones com per a passar per reals.
- El paper presenta un nou mètode d'entrenament per a GANs que es bassa en anar fent a cada pas més gran tant el generador com el discriminador. A cada capa que s'afegeix, el model aconsegueix reproduir cada vegada millor els detalls.
- Obtenció de resultats rècord.
- Els resultats s'obtenen en temps molt inferiors en comparació a anteriors tècniques que aconsegueixen resultats comparables (encara que no tan bons).



Explicació del Dataset



- Xarxes Neuronals importades amb el mètode Pickle.
- El projecte funciona amb dades obtingudes dels datasets CELEBA, LSUN i CIFAR10. Amb aquests datasets, els models de NN testejen el resultat obtingut.
- La principal font de dades és el dataset CELEBA, format per 202.599 retrats de 10.177 persones diferents amb 40 atributs binaris per cada instància.
- CELEBA té resolució de 43 x 55 fins a 6732 x 8984 píxels: Es necessita una normalització.
- CELEBA final 30.000 imatges (1024 x 1024 píxels)



airplane	🚟 🔉 📈 🤛 😁 💆 💥 🚤 🐱
automobile	and the second s
bird	
cat	S S S S S S S S S S S S S S S S S S S
deer	
dog	🔅 🧥 🦝 🚮 🉈 🧸 📢 🗥 🥸

Codi i mòduls

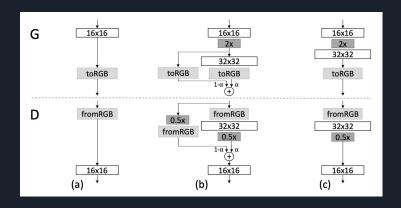
- config.py: Que permet canviar diferents hiperparàmetres dels models així com canviar diferents opcions que afecten al entrenament i a com s'emmagatzemen els resultats. Hi ha una configuració predeterminada, però hi ha moltes línies comentades que es poden utilitzar per realitzar canvis ràpidament i que venen donades pels programadors.
- dataset.py: Trobem classes que serveixen per representar tant el dataset d'entrenament com aquells que podem generar com a resultat. Està adaptat per treballar amb els TFRecords de Tensorflow. Aquestes classes permeten generar els minibatches i labels aleatoris.
- loss.py: Tot el necessari referent a la funció de cost dels diferents algorismes. Trobem: Generador de funció de cost i el discriminador. S'utilitzen els tipus WGAN, WGAN-GP i AC-GAN
- misc.py: Diferents funcions per treballar les imatges, els outputs a consola, carregar models entrenats, etc...
- **networks.py:** Funcions que implementen les capes de les xarxes amb les diferents operacions que realitzen. També hi ha funcions que retornen xarxes exactament iguals a les presentades al paper.

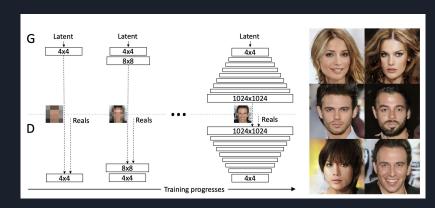
Codi i mòduls

• train.py: Dintre aquest mòdul es troben les funcions necessàries per a un correcte entrenament de les dades. Hi ha funcions per configurar la mida i els paràmetres de les miniatures de les imatges, el processament final de les imatges abans de passar-les per les xarxes neuronals, la funció que s'encarrega d'avaluar i guardar tots els registres i paràmetres dels entrenaments per la seva posterior representació. Finalment també ens trobem la funció principal d'entrenament, la més important de tot el mòdul.

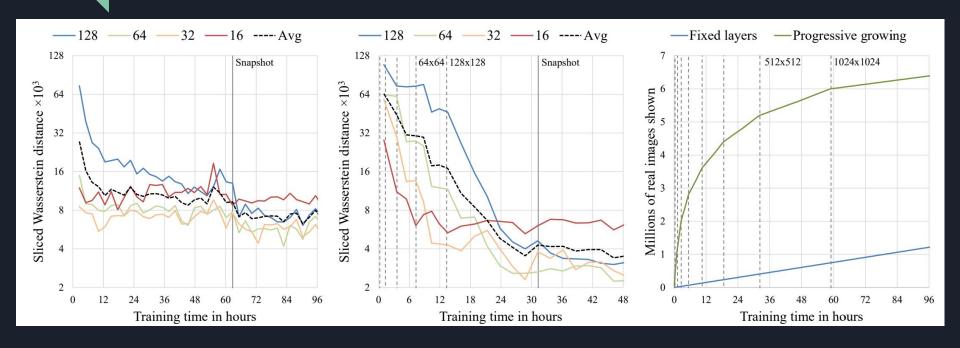
Mètode utilitzat i teoría

- Augmenta el nombre de capes de les dues xarxes augmentat a la vegada la resolució de les imatges generades. (NxN a la imatge és la mida de les capes convolucionals).
- La majoría d'iteracions de l'entrenament es fan a resolucions baixes, augmentant això la velocitat de l'entrenament. (De 2 a 6 vegades més ràpid).
- Es realitza un up-sample del resultat de la capa anterior utilitzant KNN i s'envía com a input per la nova capa. Això permet una integració més suau de cada una de les noves capes.
- El paràmetre alpha permet determinar quin pes té aquest input.
- S'utilitza la desviació estàndard dels minibatch (de 16 imatges) i s'utilitza com a input per millorar els resultats.

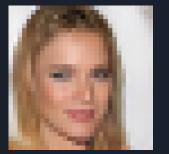




Resultats



Resultats (CELEBA)



1 dia



2 dies



5 dies



10 dies



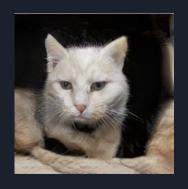
18 dies

Resultats (LSUN)









Conclusions

- Es poden obtenir resultats realment impressionants depenent del dataset on s'aplica.
- Un gran avenç degut a les millores en l'entrenament que s'aconsegueixen, fent que l'experimentació amb GANs sigui una mica més accessible sense comprometre (inclús millorar) els resultats.
- Es poden trobar múltiples aplicacions per a aquest model.
- En general, els millors resultats obtinguts són amb l'aplicació sobre el dataset CELEBA-HQ.

- Com ells mateixos diuen, encara hi ha un gran treball per davant fins a aconseguir resultats fotorealístics acceptables independentment dels datasets.
- Seria interessant afegir algun component de comprensió semàntica i de restriccions relacionades amb els datasets, com ells també indiquen, per aconseguir resultats més realistes en datasets diferents al de les celebrities.
- Pot ser una eina considerablement perillosa com ja s'ha demostrat últimament amb aplicacions semblants.