



Xarxes Neuronals Convolucionals

SAR — Enginyeria de Sistemes TIC

Andreu Oulqaid Garcia

19 de desembre de 2025

Índex

- 1 Expliqueu breument com funciona una xarxa neuronal convulacional. Relacioneu-ho amb el codi que hem fet servir a classe. Quins avantatges i inconvenients hi ha comparat amb el mètode de l'exercici de vores i cantonades fet amb matlab? 2

1 Expliqueu breument com funciona una xarxa neuronal convulucional. Relacioneu-ho amb el codi que hem fet servir a classe. Quins avantatges i inconvenients hi ha comparat amb el mètode de l'exercici de vores i cantonades fet amb matlab?

Una xarxa neuronal convulucional (CNN) és un tipus de xarxa neuronal dissenyada per processar dades amb estructura espacial, com ara imatges. El seu funcionament es basa en l'aplicació de filtres de convolució que recorren la imatge d'entrada i extreuen característiques locals com vores, textures o formes. Aquestes característiques passen per funcions d'activació, habitualment ReLU, que introdueixen no-linealitat simulant el comportament aproximat de les neurones biològiques: les respostes negatives s'anul·len i les positives es mantenen de forma lineal. Posteriorment, les capes de *pooling* redueixen la dimensionalitat dels mapes de característiques, fent el model més robust a petites variacions espacials. Finalment, les característiques extretes s'aplanen i s'utilitzen en capes completament connectades per realitzar la classificació o predicció.

Aquest esquema teòric es reflecteix directament en el codi treballat a classe amb MNIST: s'introdueix la imatge d'entrada de mida 28×28 . Les capes *conv1* i *conv2* apliquen respectivament 10 i 20 filtres per detectar patrons locals; ReLU introdueix la no-linealitat (aproxima el comportament d'una neurona); el *max pooling* redueix la resolució i aporta invariància a petites translacions; i la capa completament connectada combina les característiques finals per classificar els dígit del 0 al 9. Mitjançant l'entrenament amb *CrossEntropyLoss* i descens de gradient, els filtres passen de ser aleatoris a capturar patrons significatius.

Conceptualment, aquesta aproximació contrasta amb els mètodes clàssics de visió artificial basats en la detecció de vores i cantonades (com Canny o Harris), on les característiques són definides explícitament pel programador i la decisió final es pren mitjançant regles geomètriques i llindars. Aquests mètodes són interpretables i eficaços en entorns controlats, però poc flexibles davant canvis de rotació, escala o soroll. En canvi, les CNN aprenen automàticament quines característiques són rellevants a partir de les dades, sent més robustes i adaptatives, tot i perdre interpretabilitat directa. En síntesi, l'enfocament clàssic és determinista i dissenyat a mà, mentre que la CNN és un model aprenent i generalitzable.