

A kamera még nálunk is többet lát?

Hol tart a mesterséges intelligencia alkalmazása a videotechnikában?

HTE előadás 2020 November 11.
Imre Gábor

AI (Artificial Intelligence) - azaz mesterséges intelligencia robbanásszerű növekedése

A mesterséges intelligencia területén zajló folyamatok az utóbbi években rendkívül felgyorsultak. Hónapról hónapra jelentős kutatási és gyakorlati eredmények látnak napvilágot. Öt év alatt ötszörös növekedés.

Ezek az eredmények a tudományos szférából egyre inkább a hétköznapi életünkbe csúsznak át, egyre több szoftver alkalmazásban jelennek meg

- Mindenki a mély tanulásról beszél.

Hol tartunk
most a
mesterséges
intelligencia
területén?

Mi a mélytanulás? (deep learning)

Gépi tanuló algoritmusok struktúrált
összessége, ahol több rétegen
keresztül próbáljuk az adatok
absztrakcióit kinyerni és modellezni.

A fejlődést gyorsító tényezők:

- ▶ Új algoritmusok + nagy adatmennyiség + párhuzamos feldolgozás GPU
- ▶ Egyre alacsonyabb belépési küszöb, sok tanulási lehetőség (internet)
Ingyenes szoftver könyvtárak
- ▶ Erősen „open source” orientált kutatói közösség, Open IA, Open Cog Foundation, Human Brain Project

Szoftver Architektúrák: TensorFlow, Torch7, Theano, Keras, Coffe

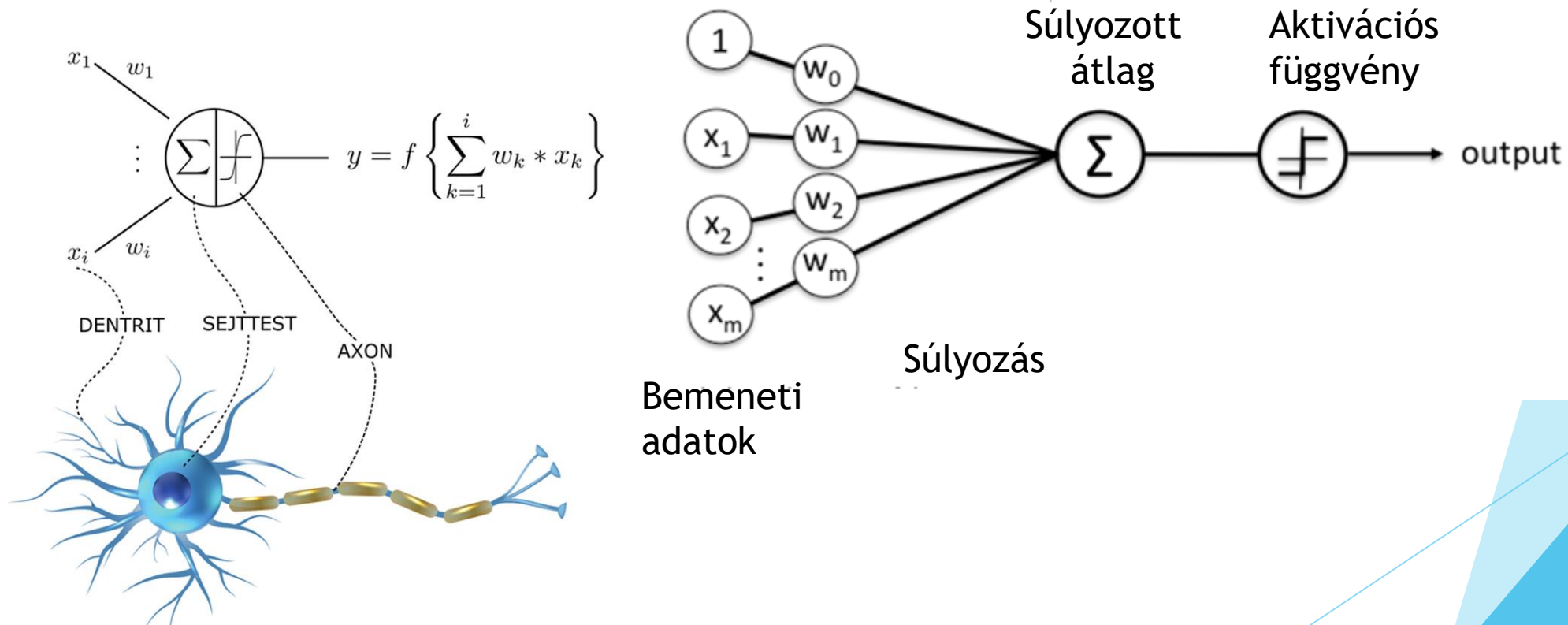
Tech óriások: Google, facebook, Microsoft, IBM, Apple, uber,
Amamazon, Badu ...

Hardver gyártók: Nvidia, Qualcomm, Intel/Nervana ...

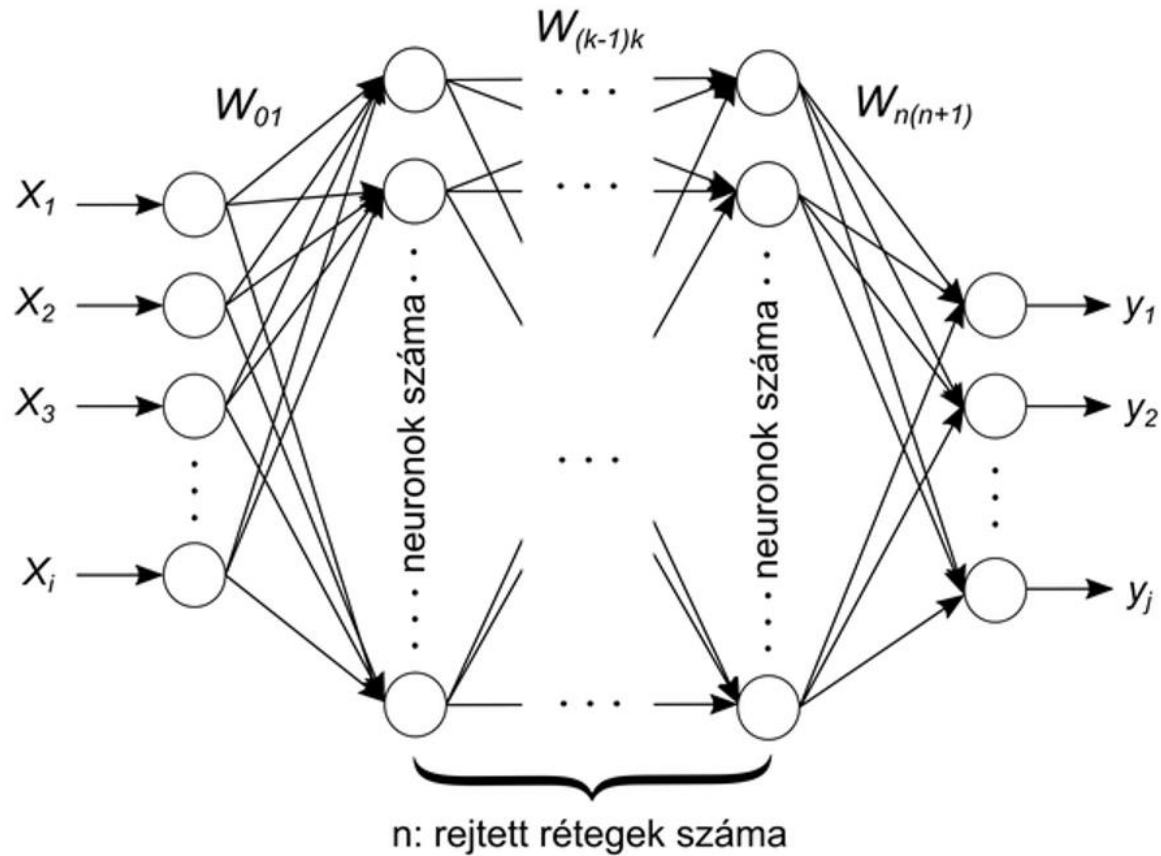
Kutatási központok, egyetemek: MIT, Oxford, NYU, Berkeley,
Stanford, Carnegie Mellon ...

Neurális háló alapelemei

- ▶ Neuron vagy műveleti elem (processing element)
- ▶ Aktivációs függvény vagy transfer függvény (csomópont vagy node)



Többrétegű struktúra



$N > 0$ esetén mélytanuló hálózatról beszélünk

Mesterséges intelligencián alapuló rendszerek, melyeket ágenseknek nevezünk

Hasznosság alapú ágensek:

Az állapotokhoz rendelt hasznossági függvényt próbálja megtanulni

Q tanuló ágensek:

A cselekvés érték (Q) függvényt próbálja megtanulni a rendszer, nincs előre megadott modell

Reflex szerű ágens:

Megtanulja a rendszer az állapothalmazból, hogy az egyes állapotokra milyen reflexszerű cselekvéseket kell végeznie

Tanuló rendszerek, DNN

- ▶ Ellenőrzött vagy felügyelt tanulás (Tanító mintapárok, a folyamat végén visszajelzés)
- ▶ Nem ellenőrzött vagy felügyelet nélküli tanulás (nincs modell vagy kívánt válasz)
- ▶ Analitikus tanulás (Növeljük a rendszer „tudását”)
- ▶ Tanulási fázis (learning phase) > *hosszú folyamat*
- ▶ Előhívási fázis (recall phase) > *rövid reakcióidő*

Rétegtípusok

(mert többrétegű rendszerekről beszélünk)

Konvolúciós rétegek CNN:

Mozgóátlag képzéssel a jellemzők kiválasztása, feature extraction

Teljesen összekötött réteg Fully connected layer FC :

Feladata az összegzés, osztályozás, az egyes jellemzőkből a kívánt modell összerakása

Rekurens rétegek RNN:

Idősor elemzés végezhető vele, látja az aktuális bemeneti értékeket (T_n) és az előző pillanatban (T_{n-1}) felvett kimeneti értékeket ezek alapján dönt.
Long Short-Term Memory (LSTM)

A tanuló rendszerek ezekből épülnek fel

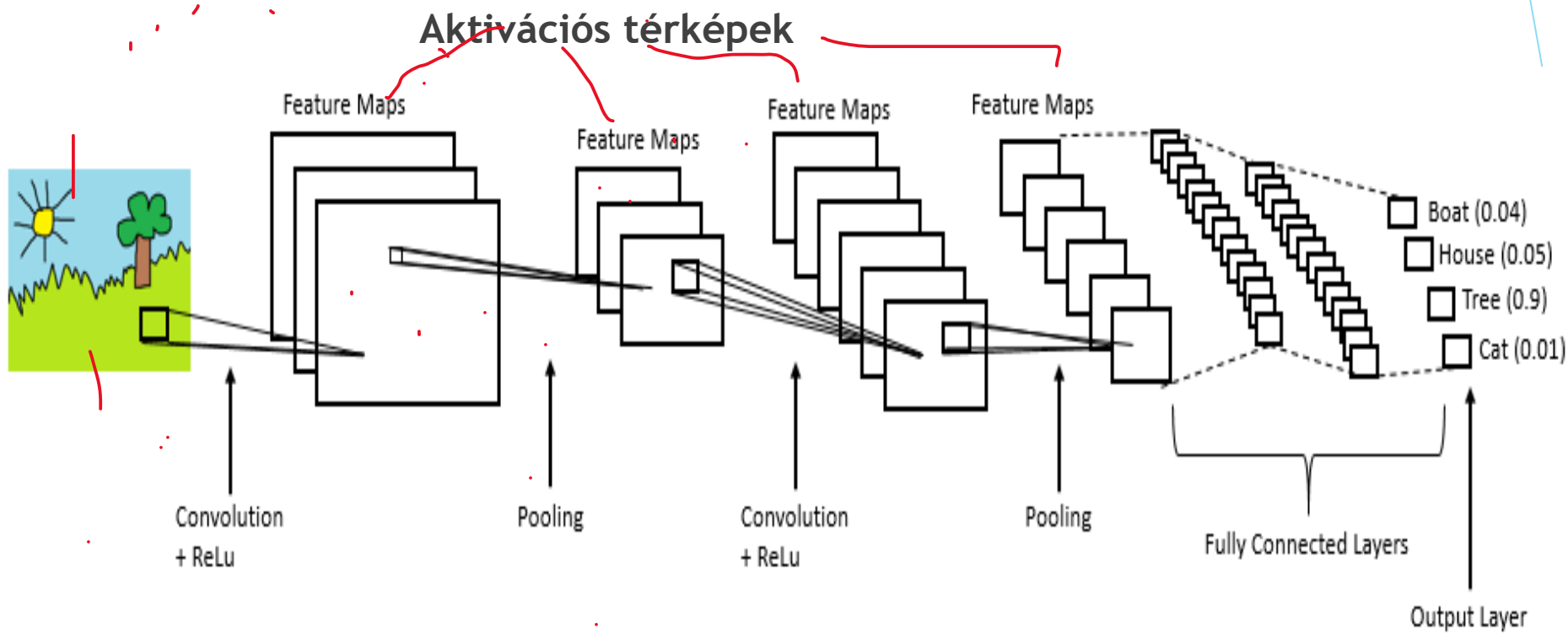
Például: CNN + FC

LSTM + FC

CNN + LSTM

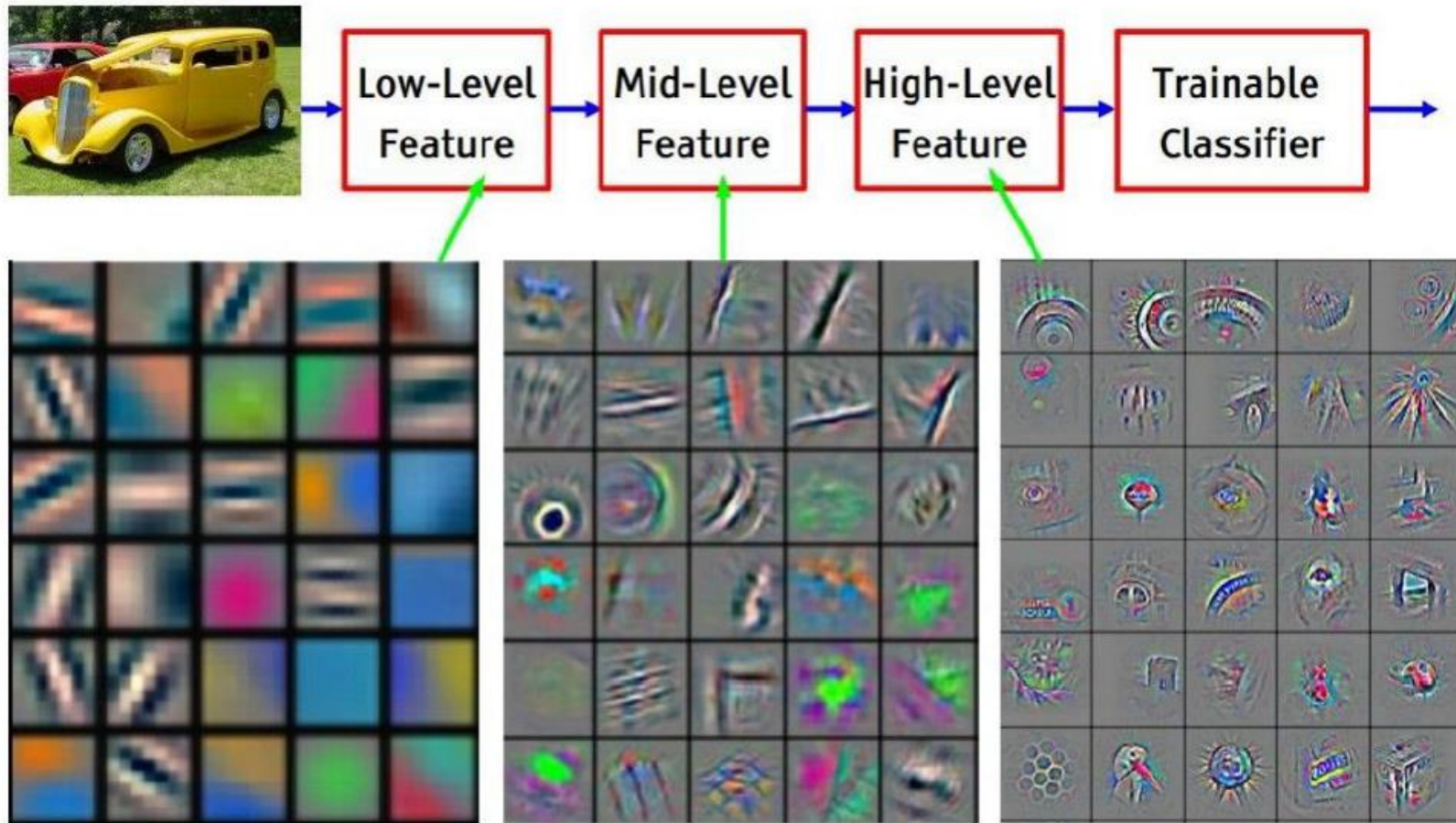
Konvolúciós Neurális Hálózatok CNN

A képfeldolgozási rendszerek alapja
a konvolúció itt súlyozott, mozgó átlagot jelent

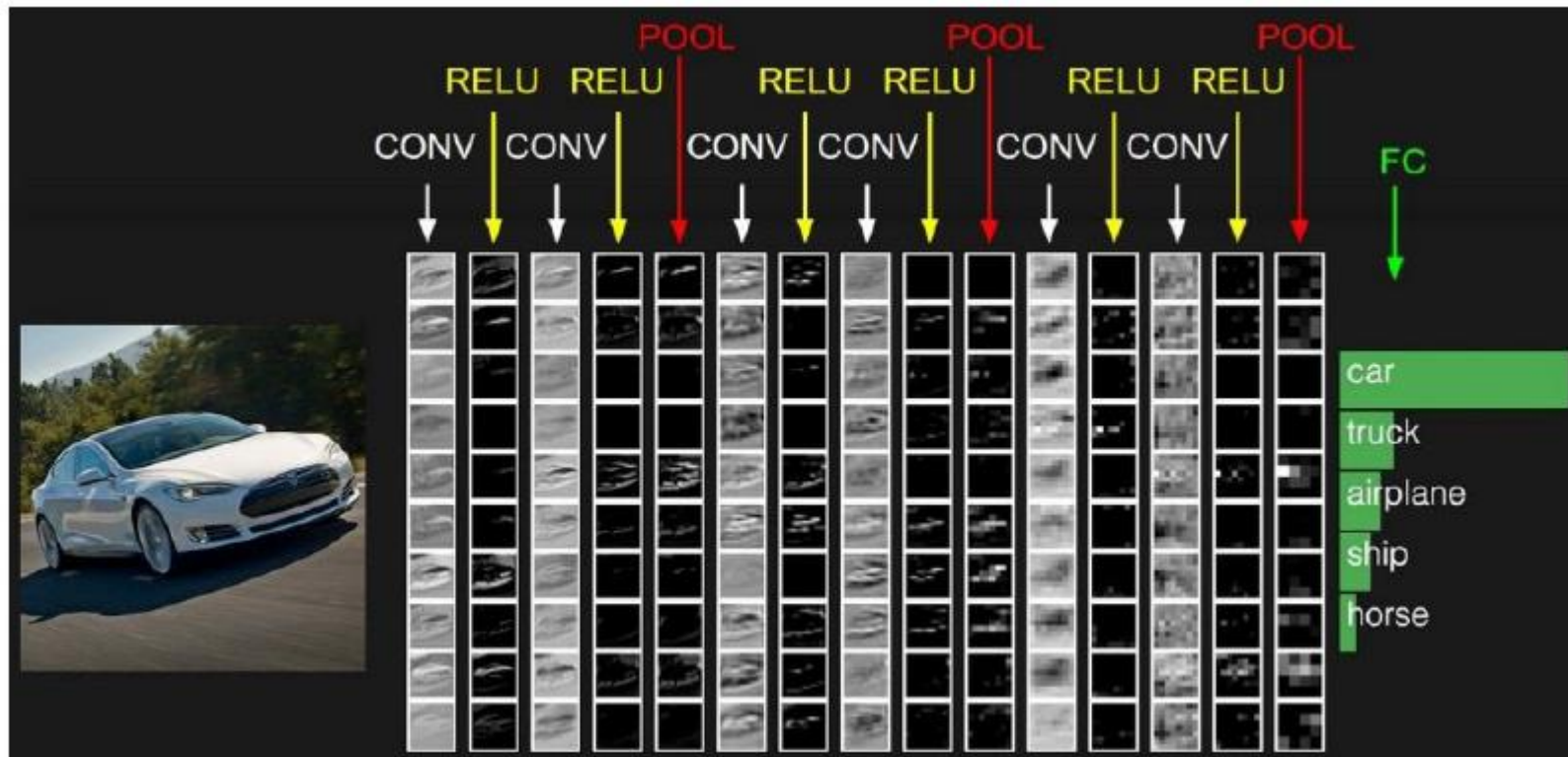


CNN + FC

A szűrők mint tulajdosság érzékelők



Teljesen összekötött réteg (Fully Connected Layer)



- ▶ Az összes konvolúciós réteg minden eleme összevan kötve az FC réteg összes „neuronjával”
- ▶ A konvolúciós (CNN) rétegek a jellemzők kiemelését végzik az FC rétegek pedig az osztályozást

Mélytanulás egy gyakorlati alkalmazása például az arcfelismerés

- ▶ A tanulási folyamat során a magasabb konvolúciós rétegek egyre több arcszerű jellemzőt tartalmaznak
- ▶ A mélytanuló rendszerek képesek hatalmas, többdimenziós adatállományokat kezelni

