

HMAX,HTM



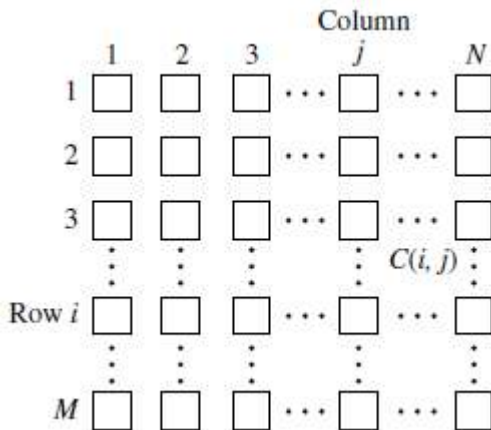
András Horváth

Mire jó egy CNN architektúra?

Turing teljes, de mit old meg hatékonyan?

2D $\rightarrow$ 2D reprezentáció. Térbeli és téridőbeli lenyomatok („feature”)-ök előállítása

A standard CNN architektúra cellák $M \times N$ -es reguláris hálózata



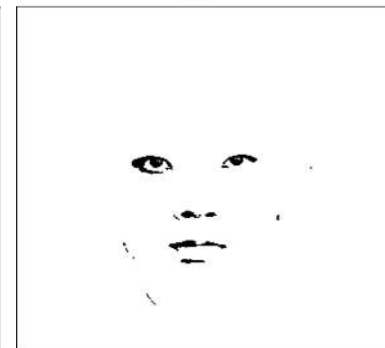
(a)



(a)



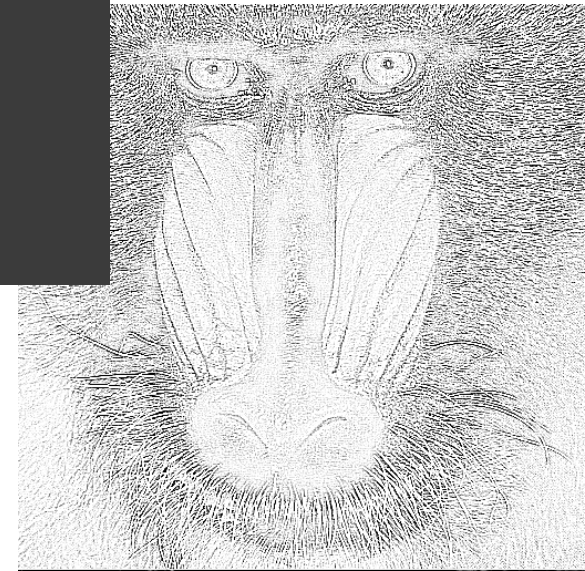
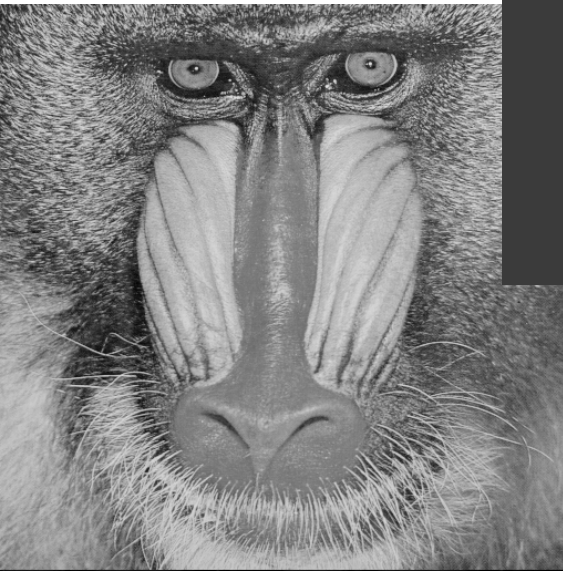
(b)



(b)

Térbeli lenyomat

Gradient



Retina felépítése

Gerinces retina: Inverz retina (ellentétben a fejlábúakkal, külön fejlődési ág) az érzékelő sejtek a leghátsó rétegben találhatóak

Csapok (cones):

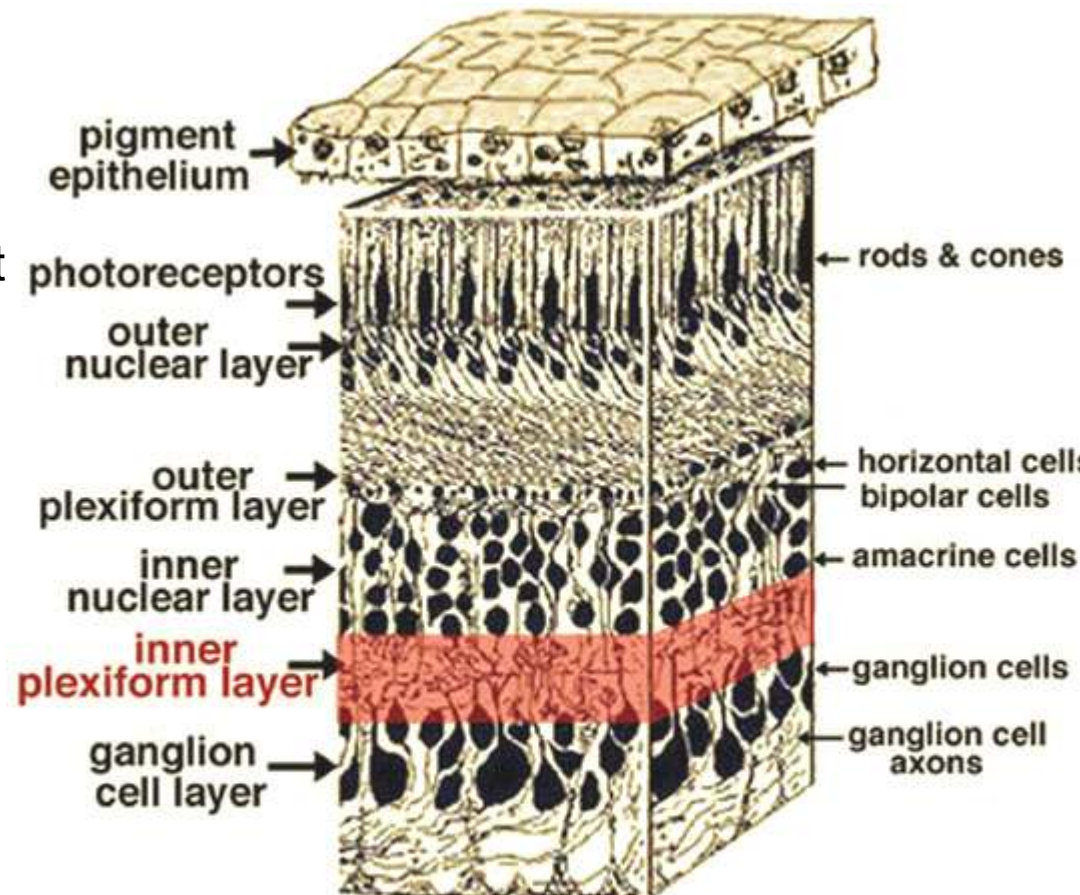
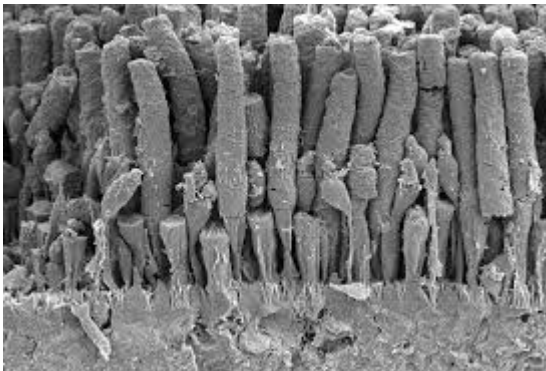
6 millió, Erős megvilágítás mellett

Fovea-ban, színlátás

Pálcikák (rods):

120 millió, gyenge megvilágítás mellett
környéki látás, egy szín

Ganglion, amacrin, bipoláris,
horizontális sejtek



Hierarchikus struktúra a retinában

100x több fotoreceptor, mint ganglion sejt:

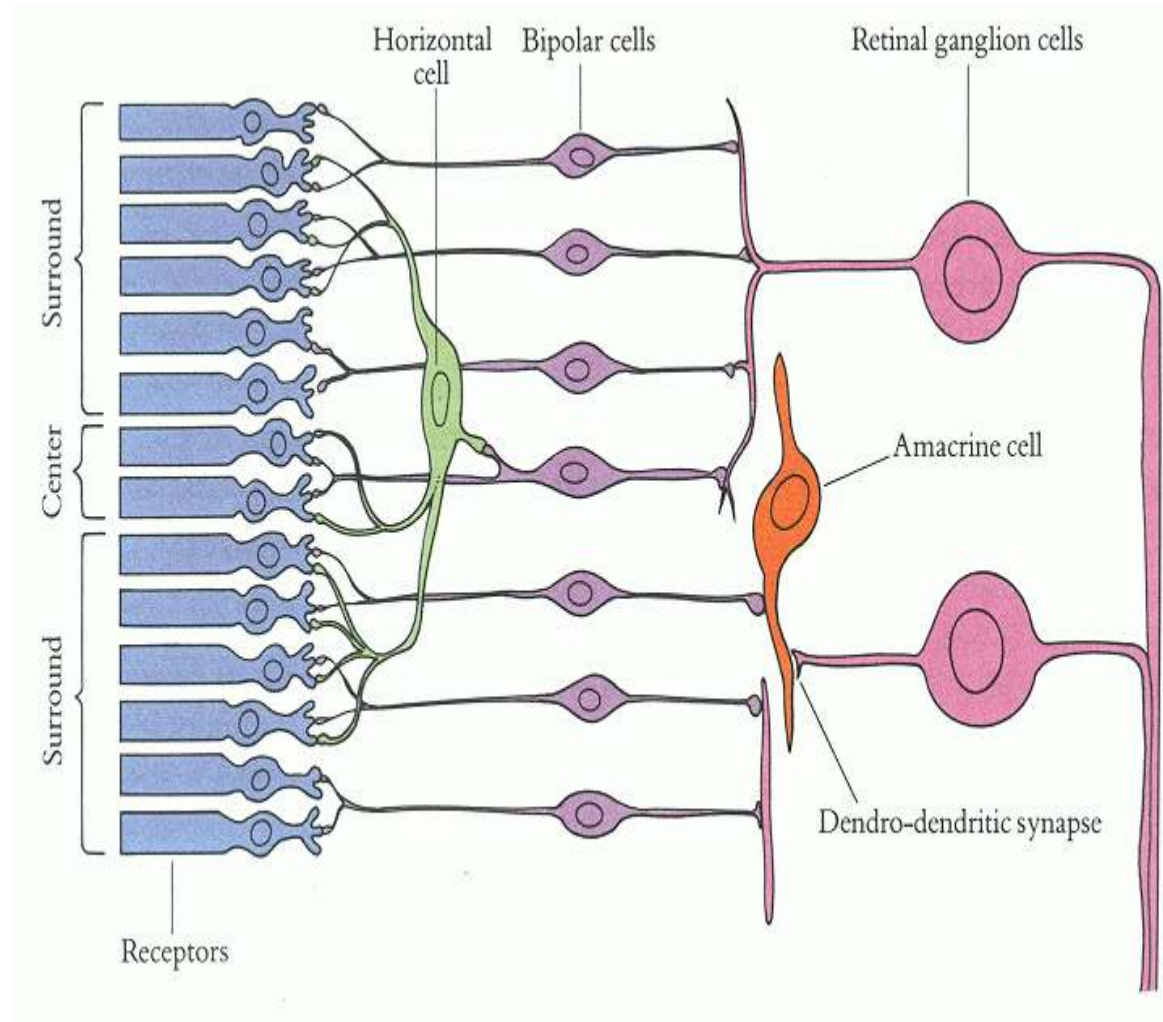
Nagymértékű tömörítésre van szükség

Nem pixel szinten látunk:
Területeket változásokat érzékelünk

Sárgafolt

On/Off

Serkentő, gátló régiók



Mit várunk egy felismerő rendszertől?

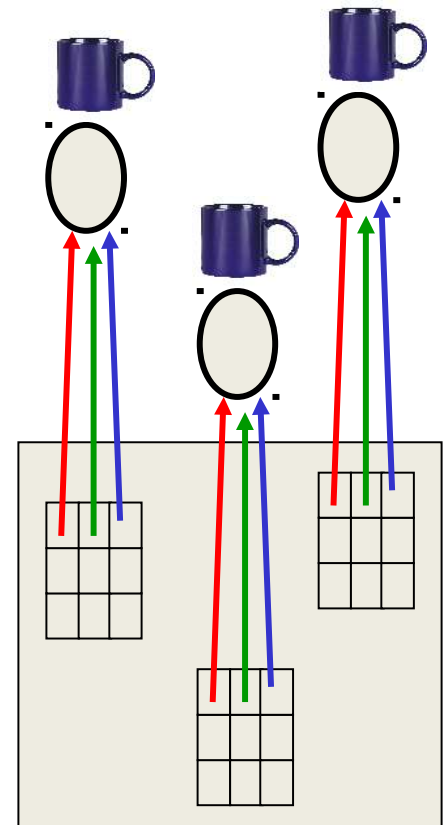
Ismételt lenyomatok (The replicated feature approach)

Nagyon népszerű neurálsi hálózatok esetében

És általában döntési folyamatokban

Ugyanazon lenyomat kinyerő struktúrát használjuk az adaton különböző 'helyzetekben'

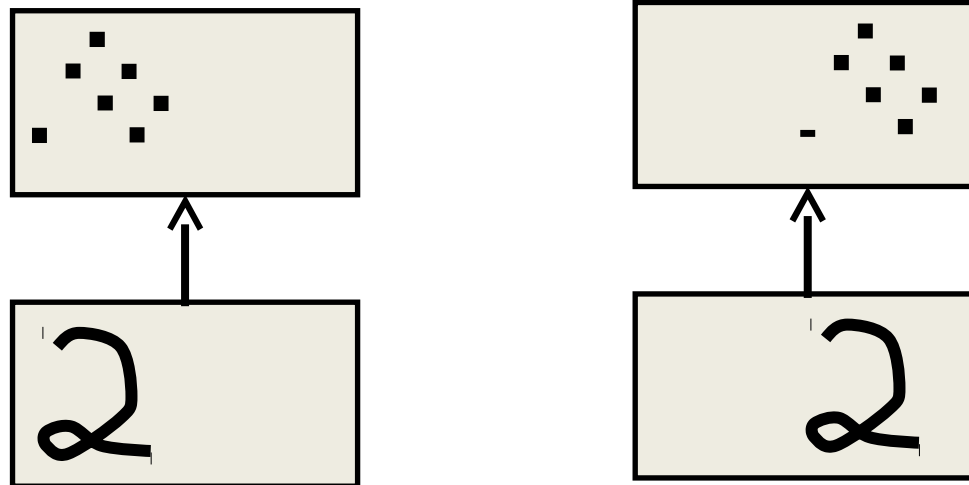
- Különböző pozíciókban azonos választ várnánk
- Azonos választ várnánk különböző méretű és orientációjú objektumokban (nehéz és számításigényes)
- Nagyon leegyszerűsíti a tanítást, sokkal kevesebb szabad paraméter lesz



Mit érünk el a lenyomat kiszámításának topografikus ismétlésével?

Nagyon leegyszerűsíti a tanítást, sokkal kevesebb szabad paraméter lesz

Invariant knowledge: A tanítás esetén ha egy régióban hasznosnak bizonyult egy feature és ezt a rendszer megtanulta, akkor az összes pozícióban hasznos lesz.



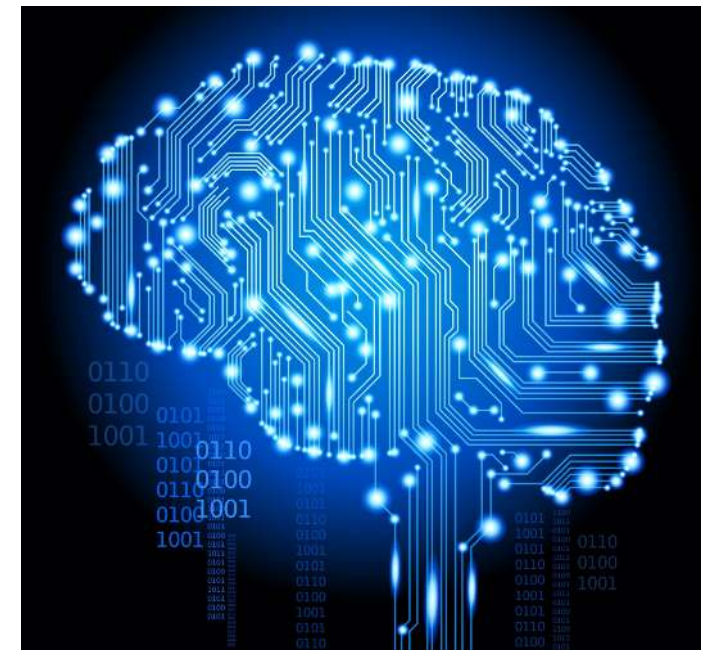
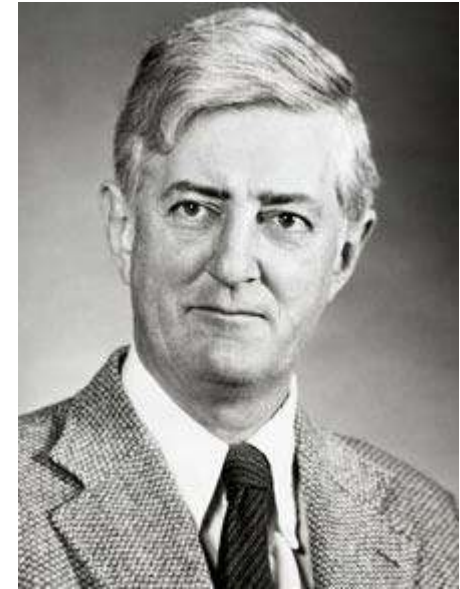
Ha a feature-ök detekcióját homogén CNN struktúrával hajtjuk végre, akkor a detekció biztosan invariáns lesz a translációra

Oliver Selfridge

1926- 2008

A gép észlelés (Machine Perception) és a mesterséges intelligencia úttörője

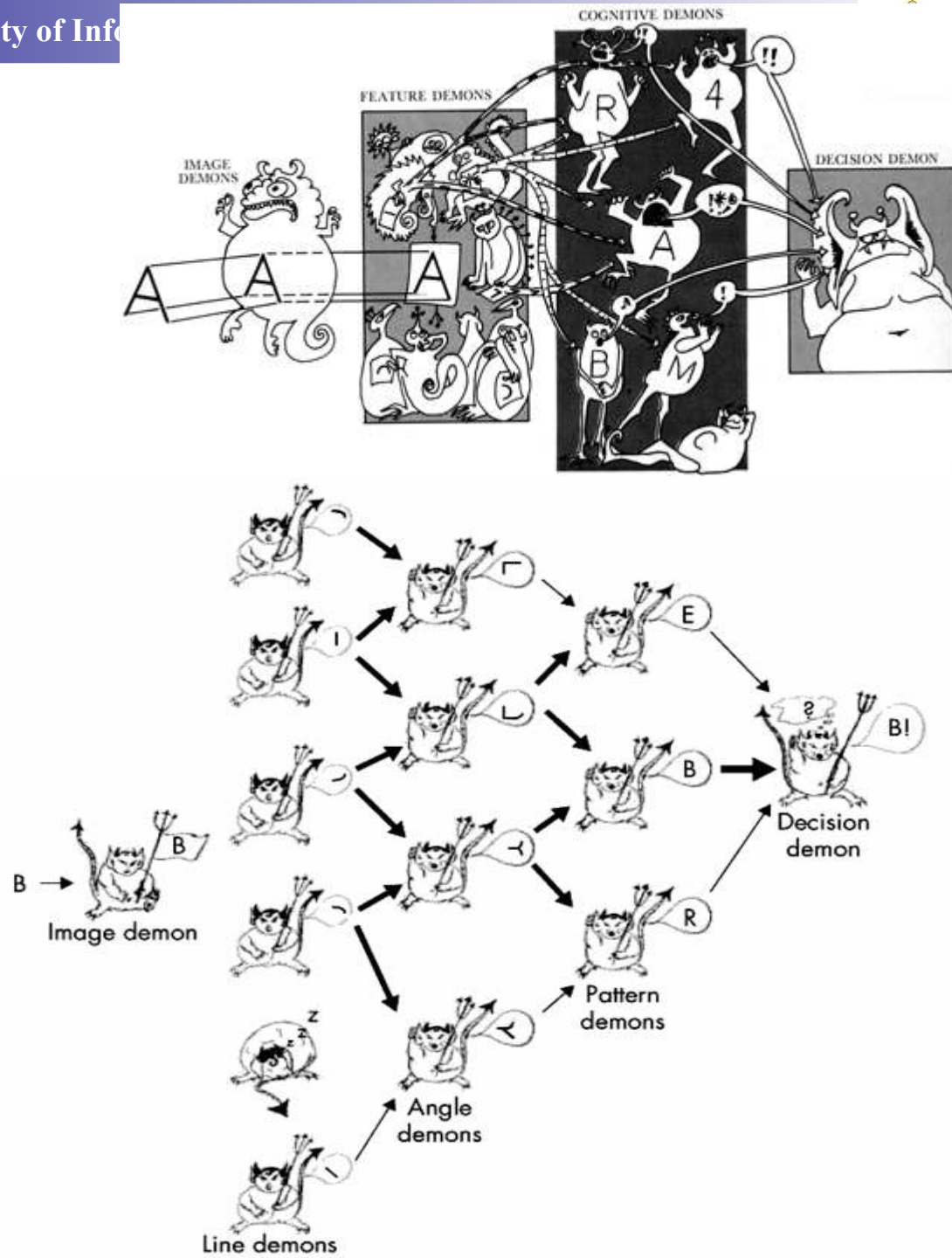
Magas szintű elírásokat adott a gépi észlelés módszereiről és az azzal szemben támasztott elvárásokról



Pandemonium paper

Egy Top-down leírást adott a gépi látás objektum felismerés folyamatáról, anélkül, hogy konkrét módszereket implementált volna. Eszerint a következőkre van szükség:

- képi szint
- feature szint (élek, irányok, minták)
- kognitív szint
- döntési szint



HMAX Algoritmus

Motiváció: Biológiai (Tomasso Poggio)

„Egy általánosítása az agyban található egyszerű sejtektől a komplex rendszerek” jelenségnek.

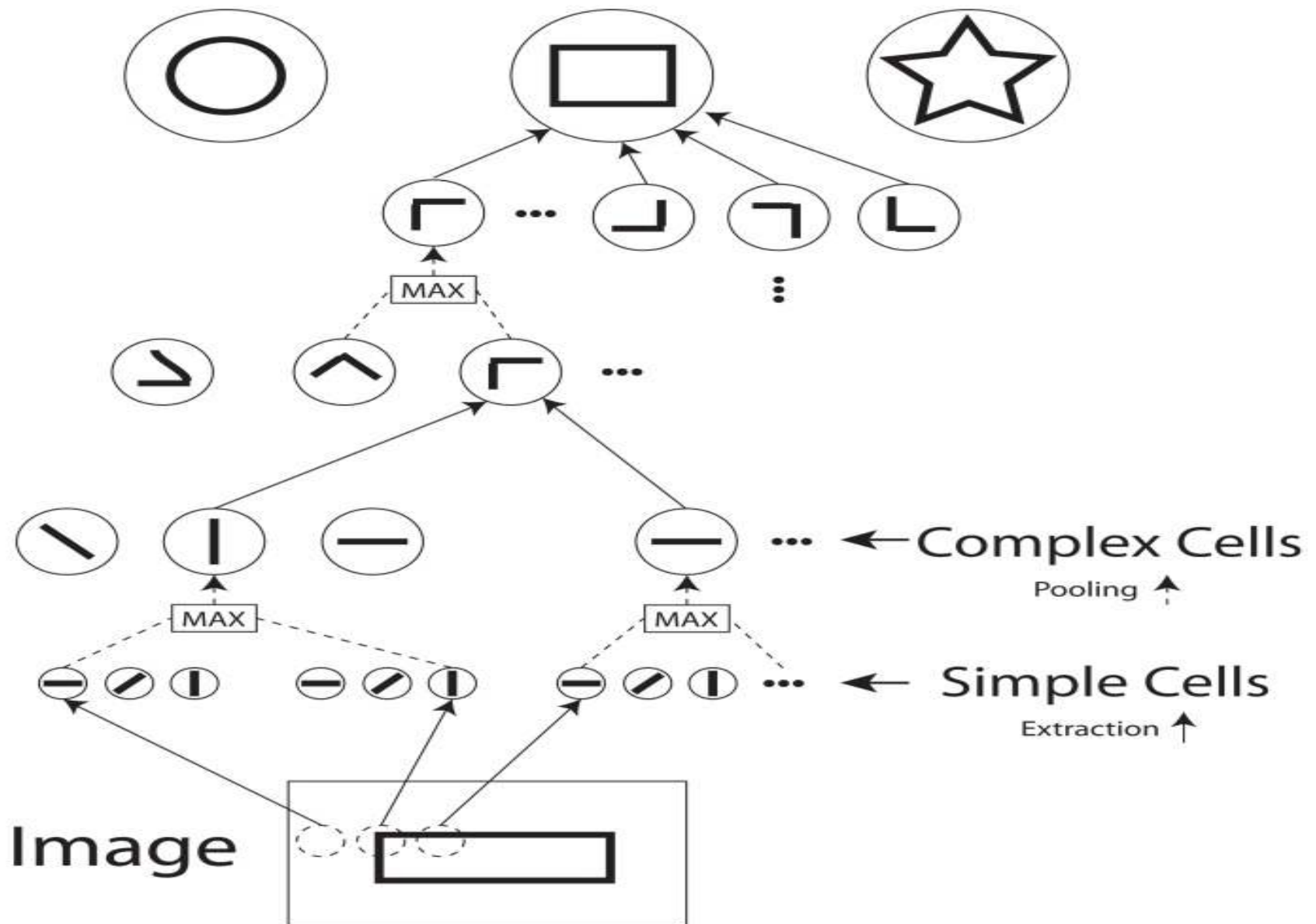
A modell változó rétegekre épül (S -simple és C complex)

Az egyszerű S rétegek egy-egy feature-t, lenyomatot számlolnak ki, míg a C rétegek ezen S tulajonságokat összegzik

A legfelső réteg nem neurális hálózat, hanem egy döntési mechanizmus

Általában SVM, de lehet decision tree is

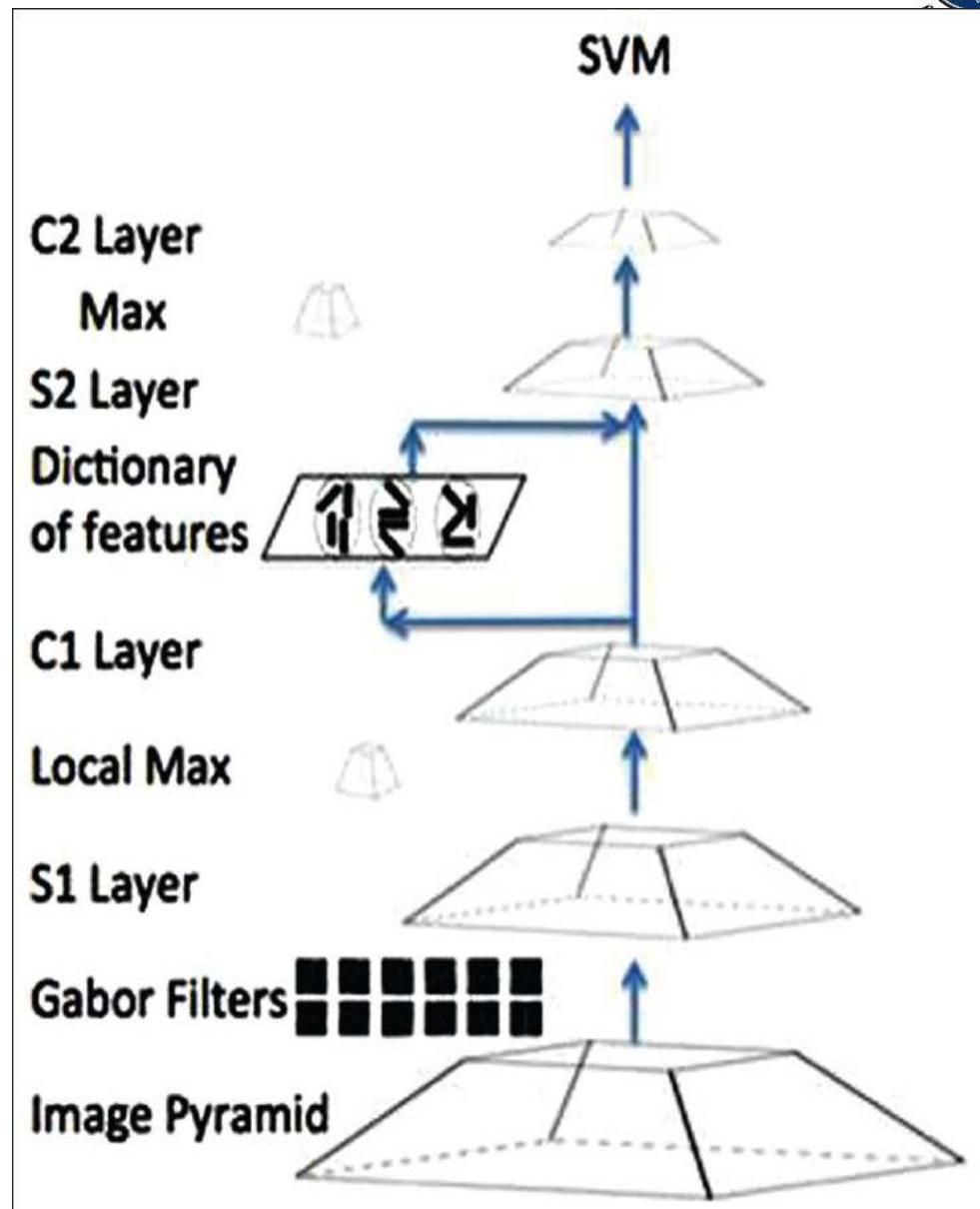
HMAX Algorithmus



HMAX Algoritmus

Deep Learning esetében hosszú tanítás, gyors döntés. A feature-ök száma és az architektúra is kötött.

HMAX esetében a rétegek száma fix, de a feature-ök száma a tanítás során változik, így a tanítási folyamat lerövidül, azonban minél több mintát tanítottunk, általában annál több feature kell egy felismerés lefuttatásához



HTM Algoritmus

Jeff Hawkins and Dileep George (Numetna)

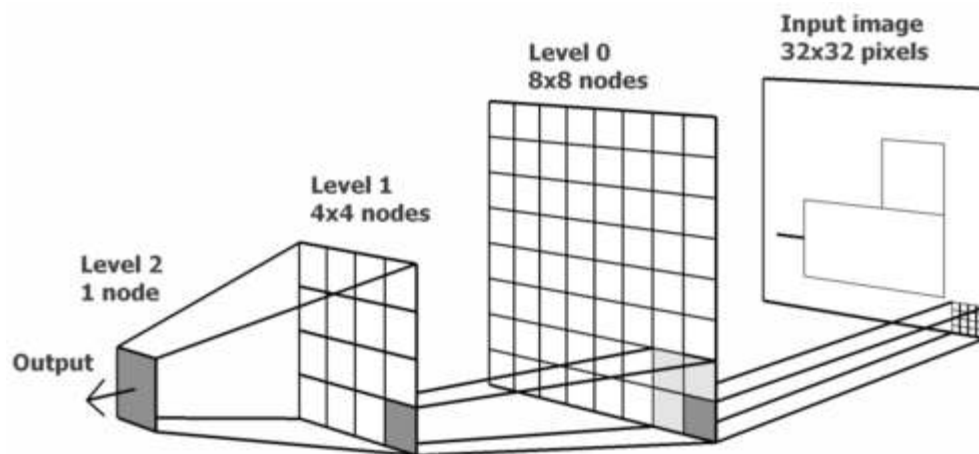
biomimetic model based on the memory-prediction

Minden szint azonos funkcióval bír

A legfelső szint, általában egyetlen elem, ami a lehetséges mintázatokat kódolja

Motiváció: a neocortex III-ik rétegében található sejtek

HTM: szigorúan bináris mintázatok és szinapszisok



HTM Algoritmus

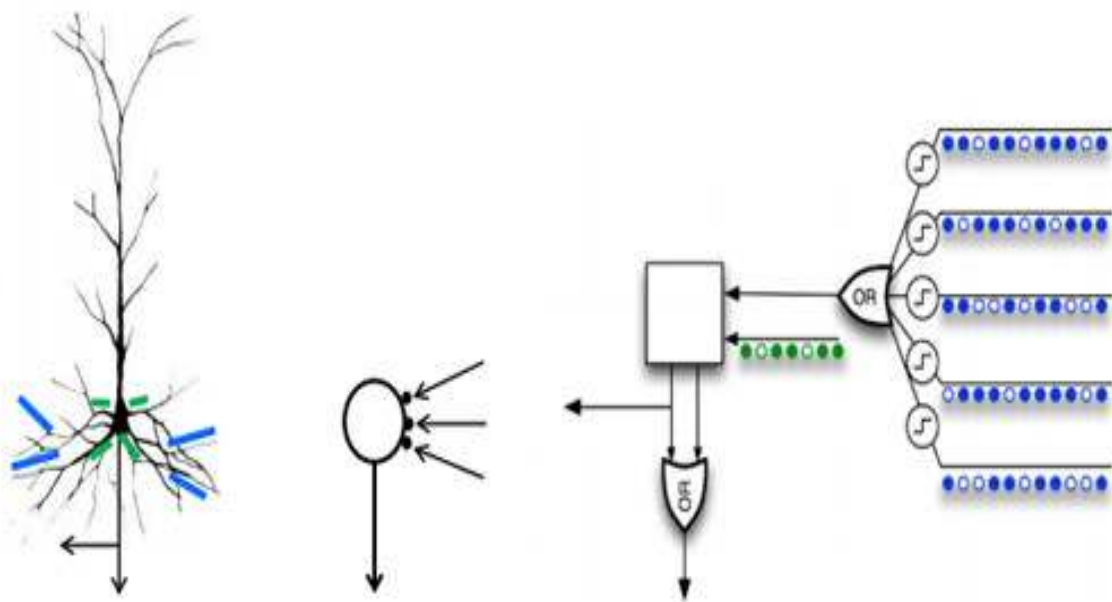
HTM sequence learning

Ha térben és időben közeli mitnázatok tűnnek fel együtt, akkor a szinapszisok megerősödnek

Ha egy új mintázat tűnik fel, akkor a szinapszis gyengül

N darab tér és időbeli mintát tud tárolni

Appendix A: A Comparison between Biological Neurons and HTM Cells



HTM Algoritmus

Tanítás során az elemek időbeli sorozatokat kapnak

Spatial pooling: térrbeli mintázatokat jegyez meg. Azon mintázatok melyek együtt tűnnek fel, eltárolásra kerülnek. Ha egy ehhez hasonló mintázat érkezik újra, a mitna megerősödik

Temporal pooling: Azok a mintázatok, amelyek időben egymás után következnek ugyanígy eltárolsára kerülnek

A felsőbb szinteken tárolt minták sokkal lassabban változnak, mint az alsók

Felismerés esetében egy elem egy valószínűséget (súlyt) ad vissza,hogy mennyire valószínű, hogy az éppen megfigyelt mintázat egy osztályba tartozik.

HTM Algoritmus

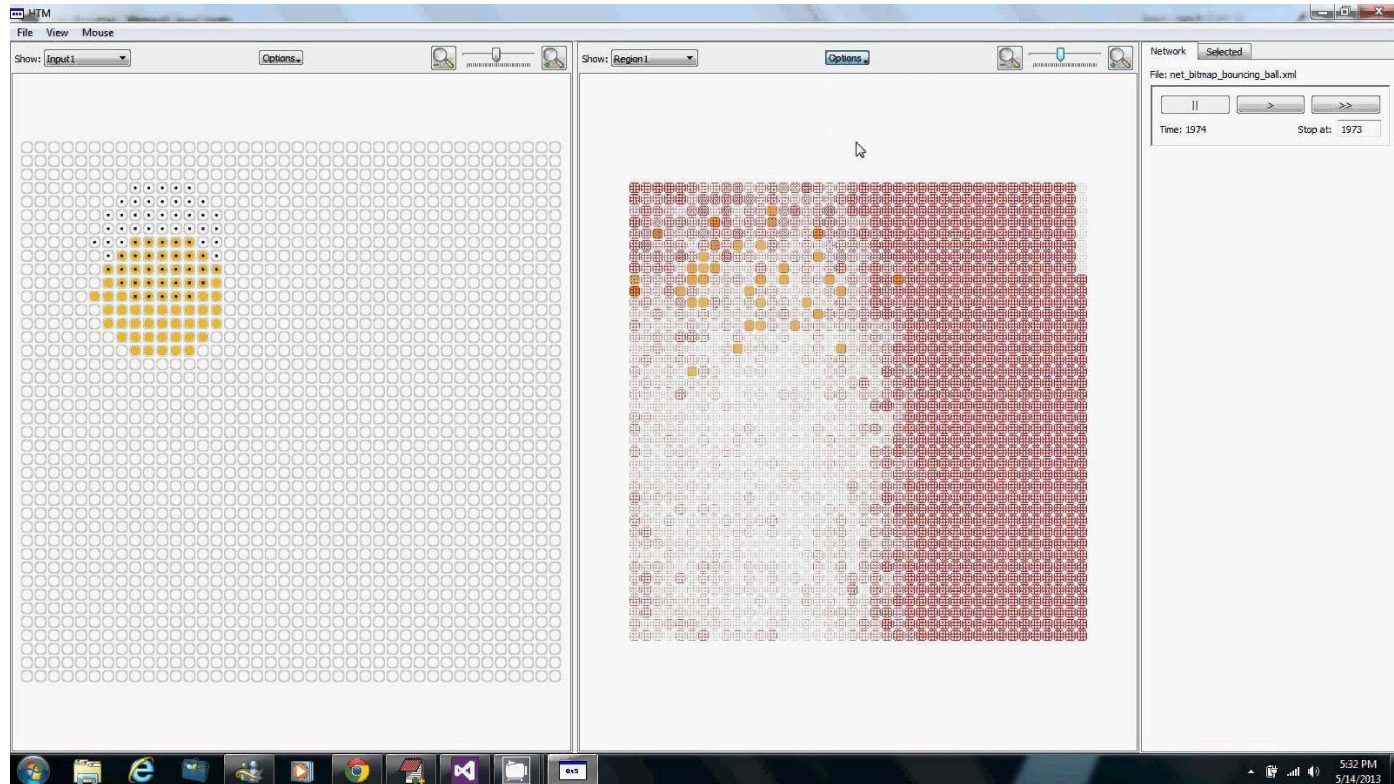
A tanítás során különböző súlyok tárolódnak a hálózatban, de az eltárolt minták nem egyszerű feature-ök, hanem tér és időbeli sorozatok.

Ezáltal képes predikcióra is, vagyis, ha már felismert egy mintát – egy adott időpont után a legnagyobb súlyú feature-öket visszük tovább a következő körökben- Akkor az eltárolt mintázatot kapjuk vissza

HTM Algoritmus

Egyszerű modell tanítása HTM- mel (bouncing ball)

<https://www.youtube.com/watch?v=YeBC9eew3Lg>



HTM Algoritmus

Numenta: Vitamin D → Sighthound

