

Konvolúciós hálózatok

Benyó Balázs

bbenyo@iit.bme.hu

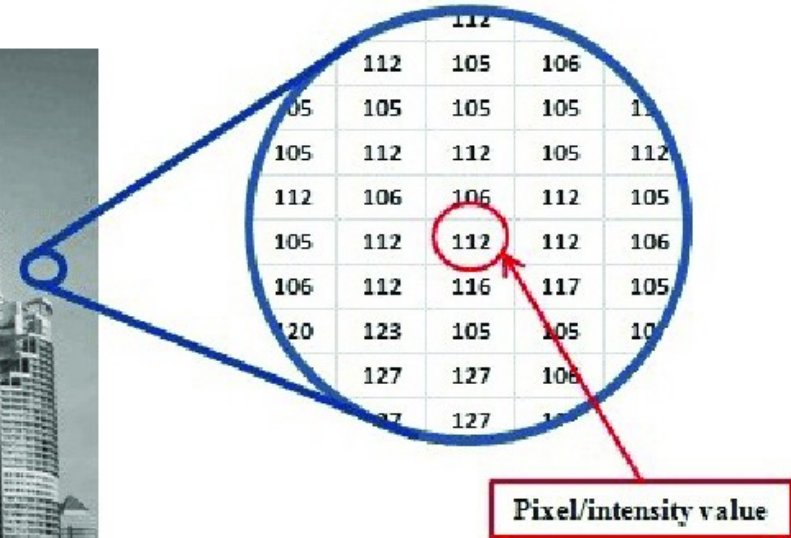
Motiváció

- Jelek feldolgozása neurális hálózatokkal
- Mit lát a számítógép?

- Képfeldolgozás / neurális hálózat alkalmazásának célja:
- Megtalálni, mely pixelek, mely **jellegzetességek**, ún. **feature-k** alapján tudom eldönteni a diagnosztikai kérdést, vagy kiszámolni a diagnosztikai paramétert!



Greyscale image



Probléma

- Minden pixel a feldolgozás szempontjából egy **független adat-dimenzió**
 - Egy 128x128-as kép esetén a bemeneti adatom 16384 dimenziós!
- **Kérdés:** Melyik releváns közülük?
- **Deep Learning:** Bízunk a hálózatra a relevancia elbírálását!

Példa a probléma megértéséhez:

- Tegyük fel, hogy a hálózat első rejtett rétegében 100 neuron van
- Ez összesen $100 * 16384$, azaz 1 638 400 súlyt jelent önmagában (bias nélkül)
- A **fully connected** hálózatok rosszul skálázódnak a felbontással:
 - A **paraméterter túl nagy** a tanítás ésszerű megvalósításához
 - Nagy lesz a veszélye a **túltanulásnak**:
hálózat megtanulja a mintákat, **kicsi lesz az általánosító képessége**

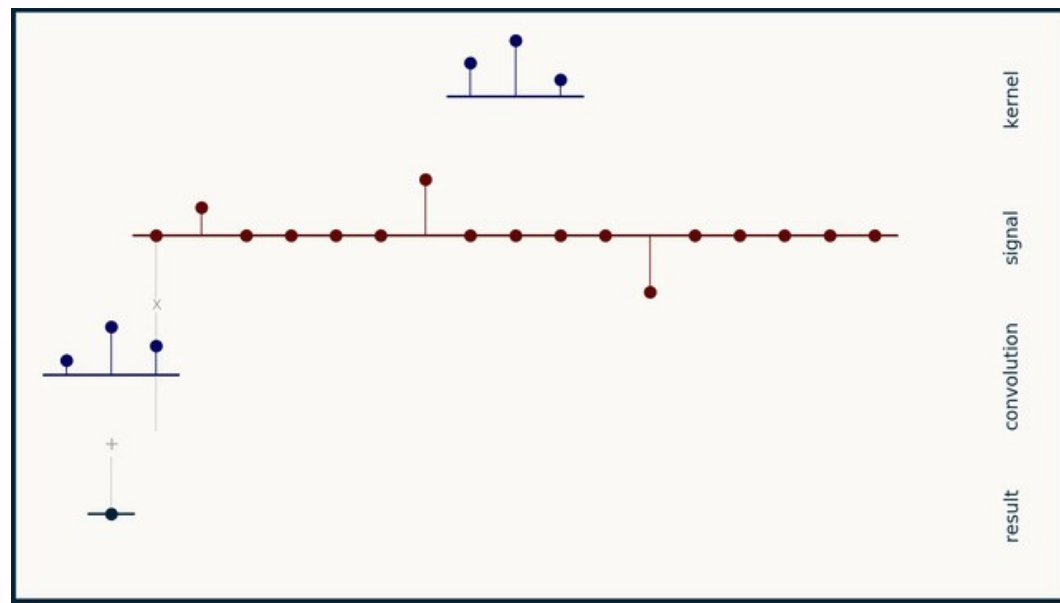
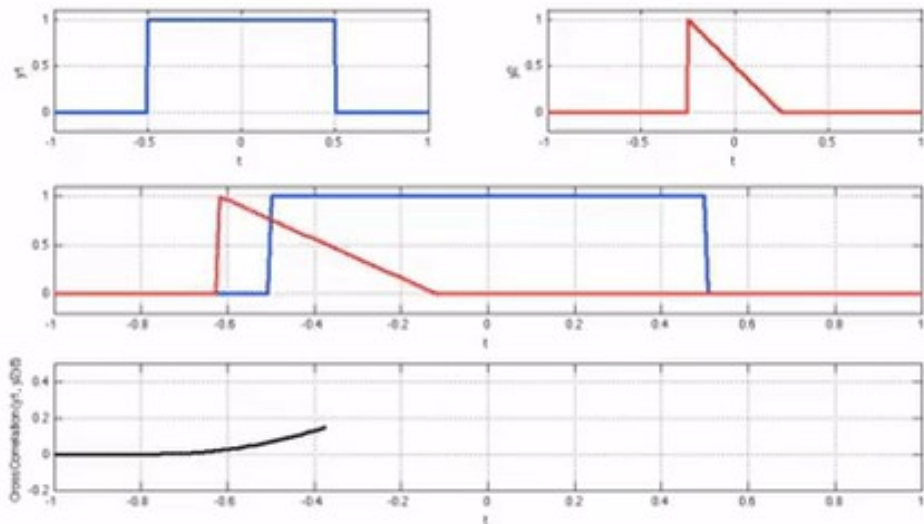


Konvolúció

The background features a dark green space filled with floating binary digits (0s and 1s) and glowing green particles. At the bottom, there are three undulating, wave-like structures composed of a grid of green dots and thin lines, creating a sense of depth and digital movement.

Konvolúció

- Matematikai művelet:
 - A feldolgozandó jeleket többször megszorozom egy másik jellel, úgy, hogy közben a szorzáshoz használt jelet eltolom
 - Az eredményt az **eltolás függvényében** ábrázolom



Konvolúció alkalmazása két dimenzióban

Strided Convolution



Input Image

1	9	8	4	4	5	7
4	8	6	7	9	1	7
4	0	5	9	3	8	4
7	3	5	9	0	5	4
7	4	1	1	8	1	2
7	6	6	9	8	7	6
3	6	3	5	4	2	7

Kernel

0	-1	0
-1	5	-1
0	-1	0

*

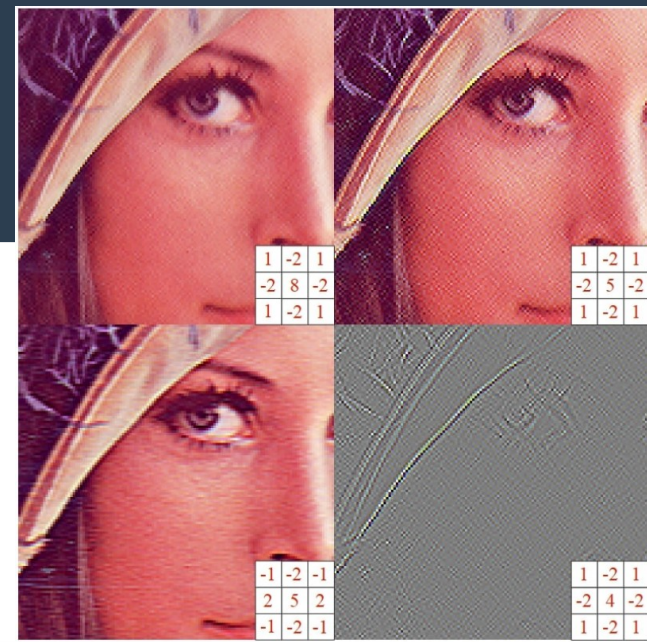
=

Feature Map



Konvolúció alkalmazása klasszikus

- Gyakori a klasszikus jelfeldolgozásban
- Segítségével a jel különböző tulajdonságai/jellemzői (pl. élek) kiemelhetők



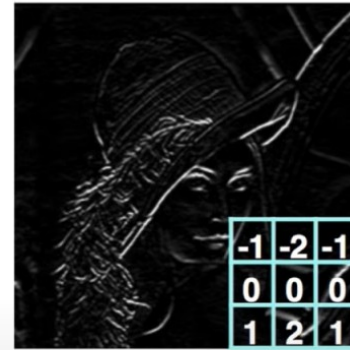
Original



Sharpen



Edge Detect



“Strong” Edge Detect



Konvolúció használata neurális hálózatokban

Dimenziócsökkentés:

- Konvolúció alkalmazásával a feldolgozott adathalmaz dimenziószáma csökken, pl. egy 3x3-as kernel 2 pixellel csökkenti a felbontást mindkét dimenzióban

Feldolgozási csatornák használata:

- Több kernel **párhuzamos** használatával **feldolgozási csatornák** hozhatók létre
- A csatornák különböző feature-re tanulhatnak rá

Pooling művelet:

- Konvolúció felbontás csökken, de ez még mindig nem elég hatékony
- Lineáris redukció helyett legyen **exponenciális** dimenziószám csökkentés:
- A **pooling** egy ilyen művelet



Pooling

The background features a dark, almost black, space filled with numerous small, glowing green and white particles. Some of these particles are arranged in a grid-like pattern, resembling binary code (0s and 1s). In the lower portion of the image, there are several undulating, wave-like structures composed of fine, green, grid-like lines. These waves appear to be made of many small, interconnected points, creating a sense of depth and movement. The overall aesthetic is futuristic and digital.

Pooling

Feature Map

6	6	6	6
4	5	5	4
2	4	4	2
2	4	4	2

Max
Pooling

Average
Pooling

Sum
Pooling



Pooling

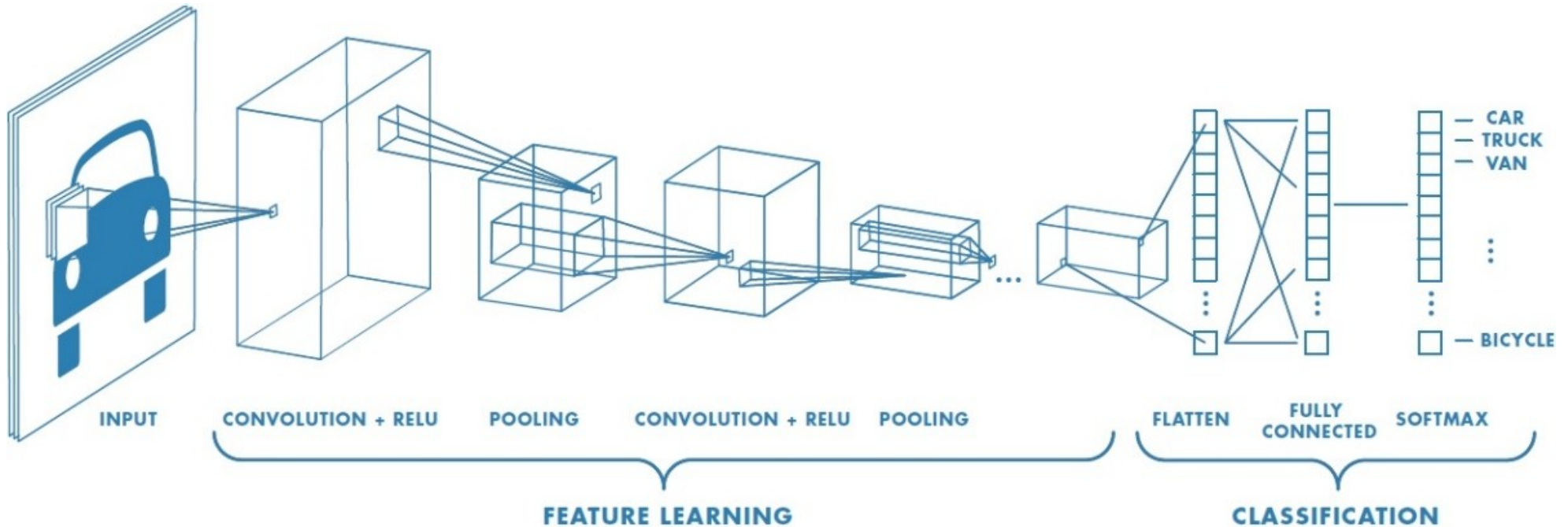
- Klasszikus pooling művelet - a felbontás felezése:
 - Az eredeti kép pixeleit 2x2-es csoportokban kiátlagolják
- Klasszikus képfeldolgozási analógia:
 - ún. rank szűrők: pixelértékek statisztikai tulajdonságai alapján szűr
- **Pooling tulajdonságai:**
 - Exponenciálisan redukálja a dimenziószámot
 - A kernel méretével arányos csökkentés
- **Pooling típusai:**
 - **Avg Pooling:** a pixelcsoport átlagát veszi
 - **Max Pooling:** a pixelcsoport maximumát veszi
 - **Min Pooling:** a pixelcsoport minimumát veszi



Hálózati architektúra

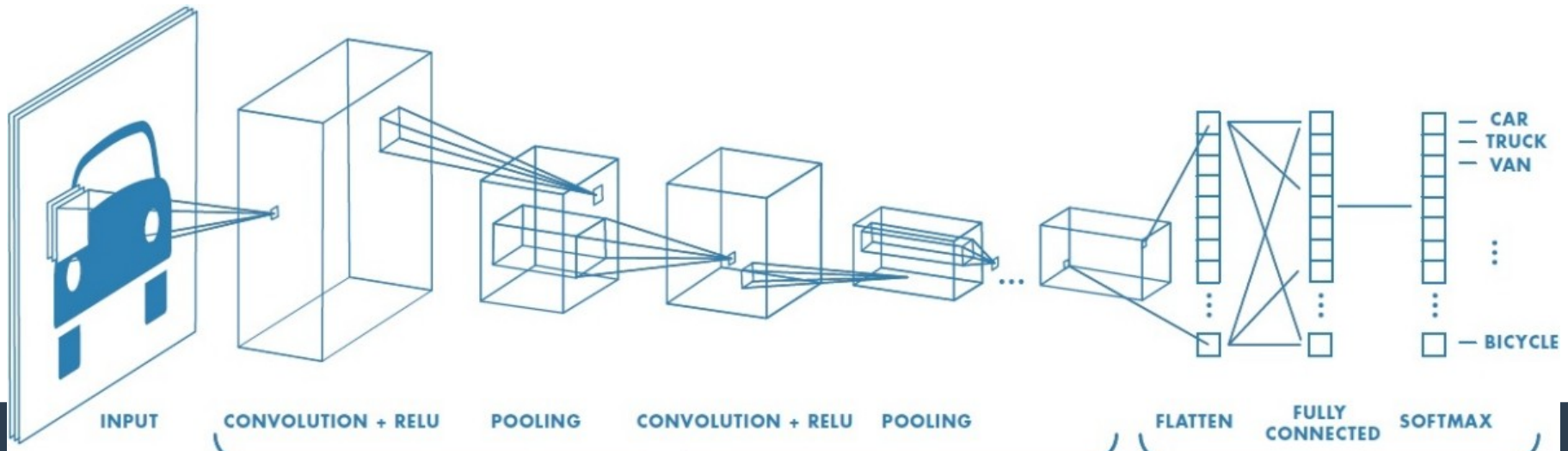
The background features a dark green to black gradient. In the foreground, there are several undulating, wave-like patterns composed of fine, glowing green lines and dots, resembling a digital data stream or network topology. Scattered throughout the scene are small, bright green and white particles, some of which are binary digits (0s and 1s), creating a sense of dynamic digital activity.

A konvolúciós neurális hálózat (CNN) jellemző felépítése



Egy alap architektúra

- 1. réteg: konvolúciós
- 2. réteg: pooling
- 3-4. réteg: konvolúciós + pooling rétegpáros (ha szükséges még redukció)
- Eredmény kinyerése: (fully connected architektúra)
- ún. flatten réteg: a 2D --> 1D konverzió;
- ún. softmax réteg: eredmény normalálása úgy (az összeg egy lesz)



A konvolúciós réteg

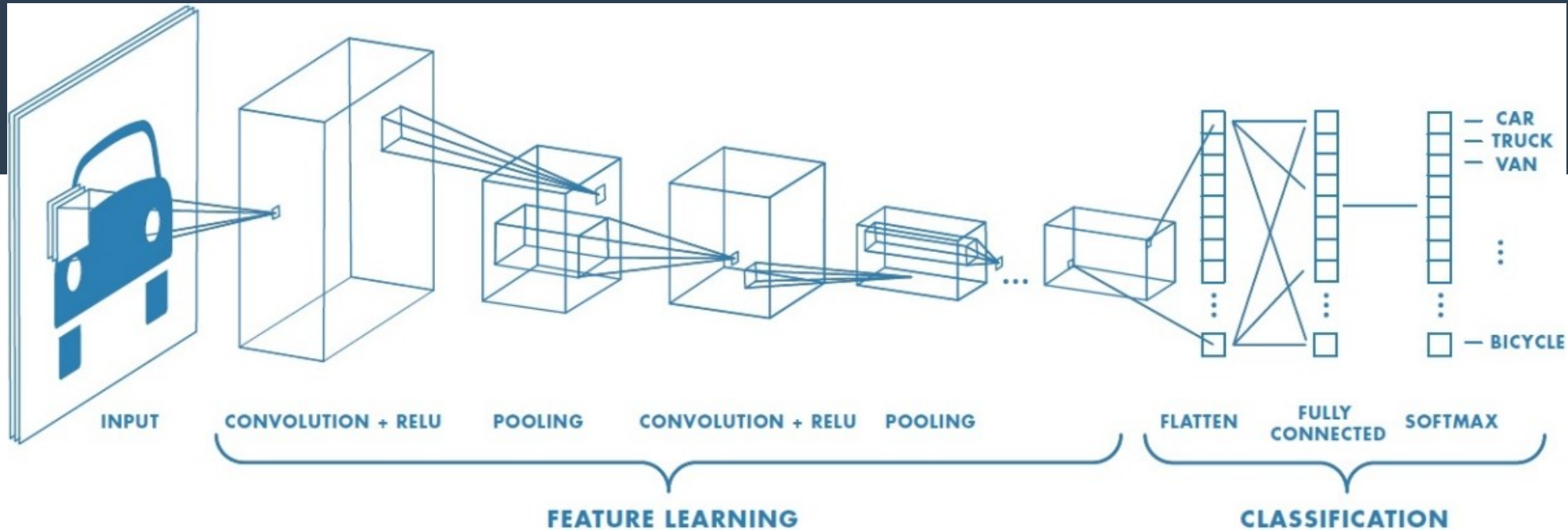
- A CNN hálózatokban több konvolúciós réteget használunk
- **Konvolúció réteg megadása:**
 - Nem a neuronok számát, hanem a **kernel méretét** kell megadni
- A hálózat a tanítás során a **kernelben szereplő súlyokat** fogjuk meghatározni
- A konvolúciós rétegnek is van aktivációs függvénye



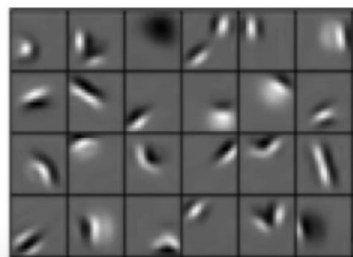
A pooling réteg

- A hálózati topológiában pooling réteg lesz
- A pooling rétegnek **nincs tanítható paramétere**
- A működési módját meg kell adni:
 - min,
 - max, vagy
 - avg.





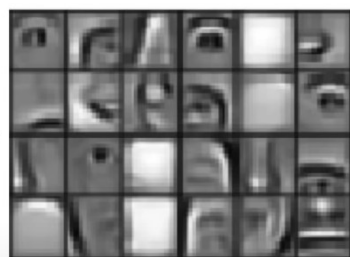
Low level features



Edges, dark spots

Conv Layer 1

Mid level features



Eyes, ears, nose

Conv Layer 2

High level features



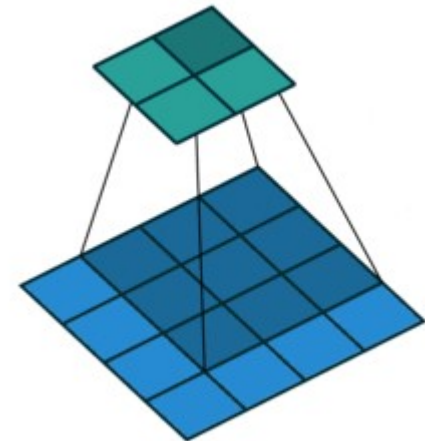
Facial structure

Conv Layer 3

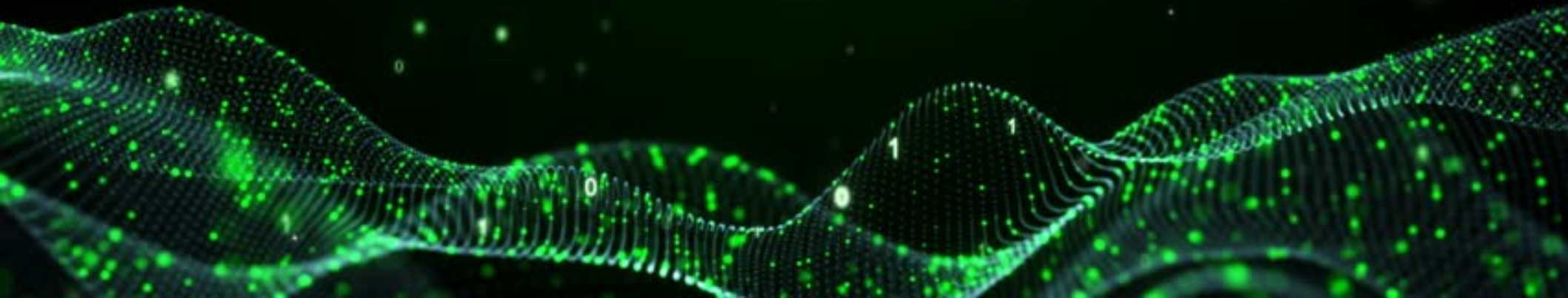


Számos CNN architektúra létezik

- Természetesen ettől az architektúrától eltérő megoldások is vannak
- Különböző **feladatokra specializált hálózati topológiák** léteznek
- Fontos CNN típus, amikor a feladat **kimenete is kép formátumú**:
 - ekkor a fully connected rész után szintén konvolúciót használunk, de „fordítva”.
 - Az ún. **dekonvolúció** segítségével állítjuk elő a feature-kból a kimeneti képet, néha az eredeti kép pixelértékeinek felhasználásával.

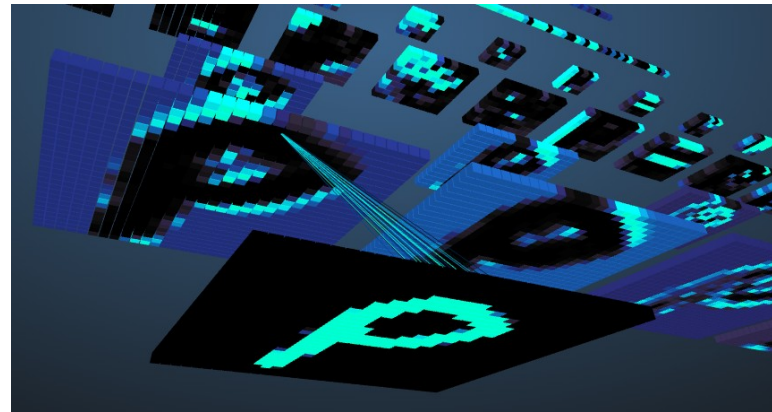
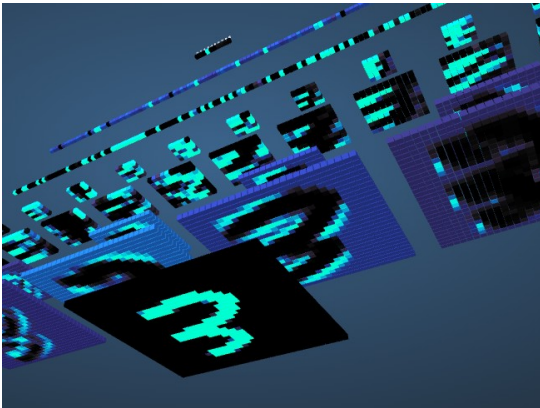


Interaktív online demonstrációs környezet



Interaktív online demonstrációs környezet

- Adam W. Harley:
An Interactive Node-Link Visualization of Convolutional Neural Networks*
- https://adamharley.com/nn_vis/
- Fully connected és konvolúciós hálózatok működésének vizualizációja és egyszerű összehasonlítása.



* A. W. Harley, "An Interactive Node-Link Visualization of Convolutional Neural Networks," in ISVC, pages 867-877, 2015

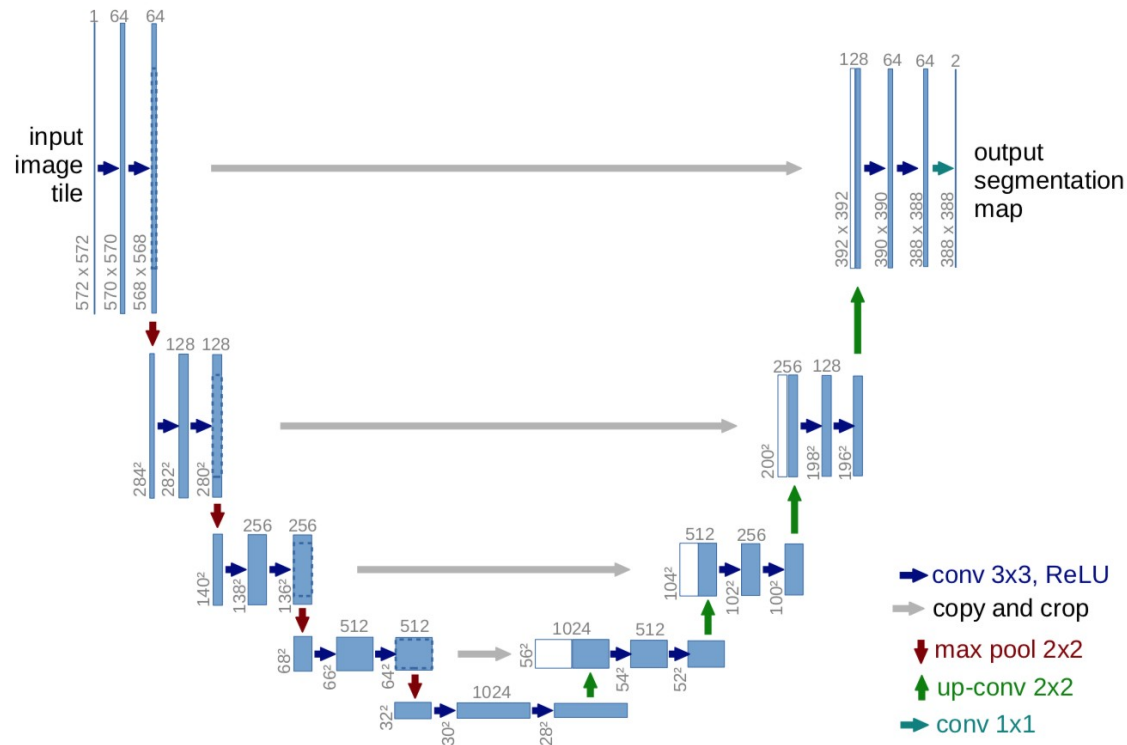


Konkrét architektúrák

The background features a dark green space filled with floating binary digits (0s and 1s) and glowing green particles. At the bottom, there are three undulating, wave-like structures composed of a grid of green dots and lines, creating a sense of depth and digital movement.

Unet

- Fully convolutional neural network: Teljesen **konvolúciós**, nem tartalmaz fully connected réteget
- Encoder-decoder részekre tagolódik
 - **Encoder** rész:
 - dimenzió csökkentés
 - **Decoder** rész:
 - Dimenzió növelés
- **Szemantikus** szegmentációs feladatok megoldására
- Kifejezetten **orvosi területre** szántál
- Alkalmazási példák:
 - Foggyökér szegmentáció
 - Máj szegmentáció
 - Vese szegmentáció



Szemantikus szegmentáció

- Összetett képfeldolgozási feladat
- **Szegmentáció:**
 - Egy kép bizonyos **tulajdonságú** pixelinek megjelölése
 - A hálózat szempontjából felfogható **pixelenkénti osztályozásként**
 - Következmény:
 - a hibafüggvény is pixelenként számított érték
- **Szemantikus** szegmentáció:
 - Komplex szabályrendszerrel definiálható, az adott pixelek mély elemzése, ill. azokon „mért jel” értelmezése alapján lehet döntést hozni

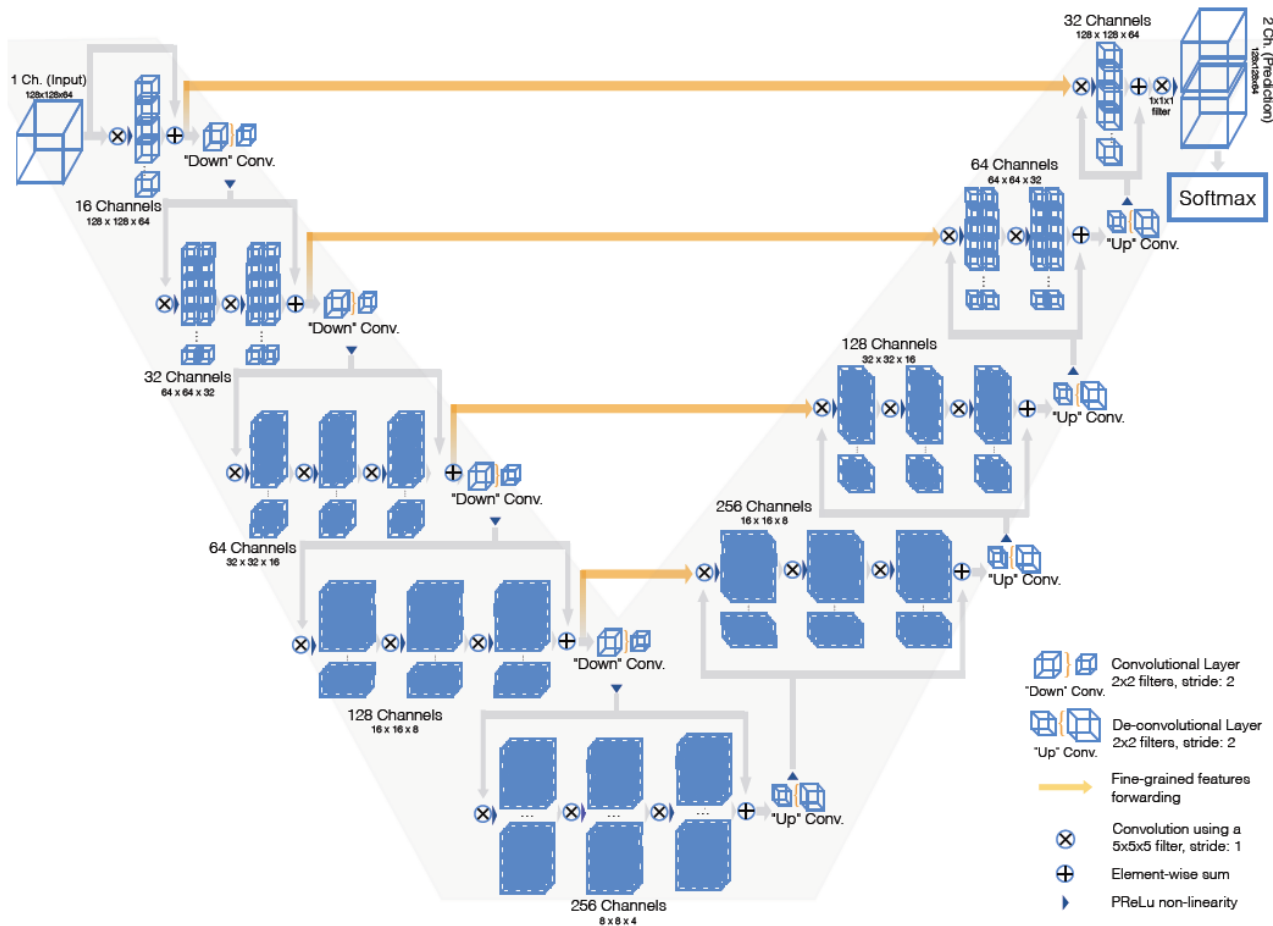


Unet – szegmentálási példa, foggyökércsatorna felismerése

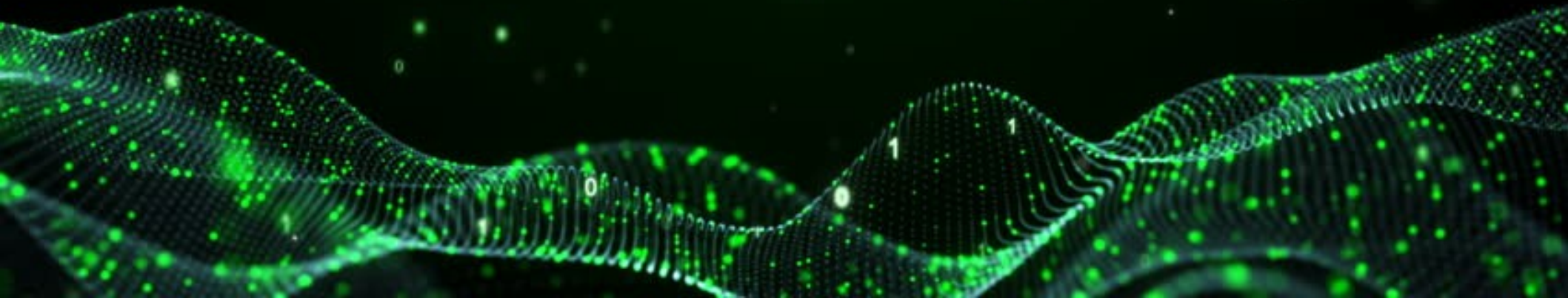
Bemenet	Tanító maszk	Hálózat kimenete	Hiba
			
			

Vnet

- Ugyan az az elv, mint az U-net-nél
- 3D adatok feldolgozására optimális
 - Bemenet és kimenet 3D adathalmaz



Visszacsatolt neurális hálózatok



Visszacsatolt hálózatok: adatszekvenciák elemzésére

- Adat szekvenciák feldolgozására kidolgozott hálózattípus
 - **Adat szekvencia:**
 - Futó paraméterhez (pl. idő) tartozó adatelemek
 - Az adatelemek az adatsorban előttük elhelyezkedő adatelemektől **függenek**
 - Egyszerű példa: COVID fertőzöttek száma Magyarországon
- Hasonló feladat a klasszikus **feed forward network**-nek **nehézséget** okoz:
 - Pl: *Elutaztam a barátaimmal Kanadába.* vs *Kanadába utaztam el a barátaimmal.*
 - Tartalmában nagyon hasonló, az FFN háló számára teljesen különböző
- A **FFN** erősen **pozíció érzékeny!**
- A **visszacsatolt hálózatok** hatékonyan tudják elemezni az **adatszekvenciákat**
 - *Másik megoldás lehet konvolúció használata, ami eltolás invariáns, de erről máskor...*



Visszacsatolt hálózatok használatának háttere

- Az ötlet ugyanaz, mint a konvolúciós hálózatoknál a **kernel** használata
- Tegyük fel, hogy létezik egy f függvény (generális függvény, paramétere megosztott), ami segítségével a sorozat minden eleme ($s^{(t)}$) előállítható az előző elemből

$$s^{(t)} = f(s^{(t-1)}; \theta),$$

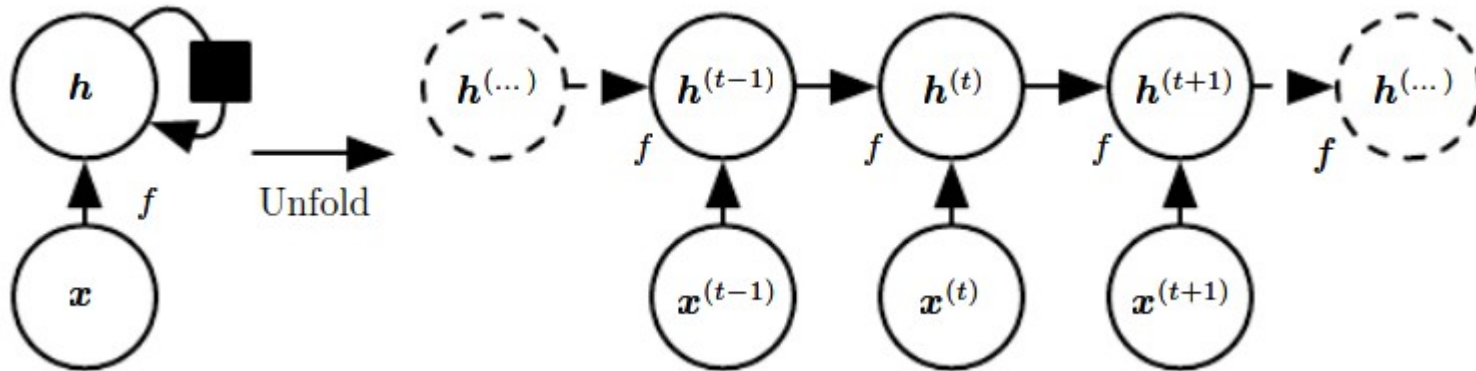
- Ilyenkor a sorozat generálható az f függvény **rekurzív** alkalmazásával, ami a hálózat **visszacsatolásának** felel meg
- Pl.: $s^{(3)}$ előállítása:

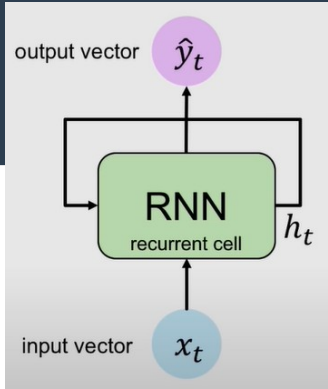
$$\begin{aligned} s^{(3)} &= f(s^{(2)}; \theta) \\ &= f(f(s^{(1)}; \theta); \theta). \end{aligned}$$



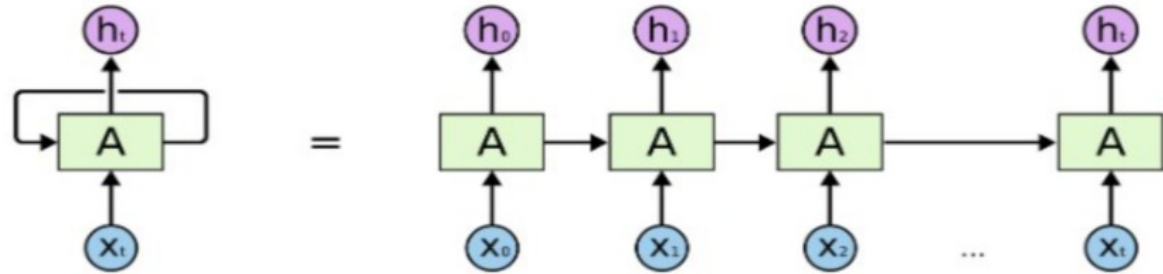
Visszacsatolt hálózatok működése: kifejtés (unfold)

- Hogyan működik a visszacsatolt hálózat?
- Egy adott véges példányban a hálózatot **kifejtjük** a visszacsatolás szerint
 - Ez olyan mintha a hálózat **annyi példányban** futtatnánk, amilyen hosszan kifejtettük
 - A kifejtéskor előállított példányok mindegyike **bemenetként** megkapja az **előző** csomópont állapot **kimenetét**





Recurrent Neural Networks



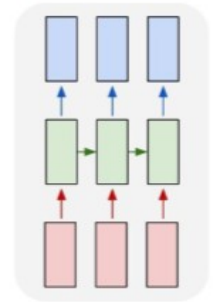
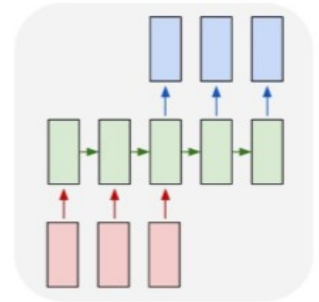
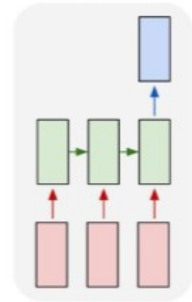
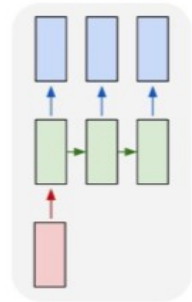
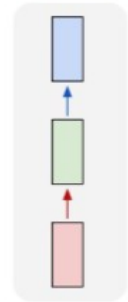
one to one

one to many

many to one

many to many

many to many



Output Vector

$$\hat{y}_t = W_{hy}^T h_t$$

Update Hidden State

$$h_t = \tanh(W_{hh}^T h_{t-1} + W_{xh}^T x_t)$$

Input Vector

x_t

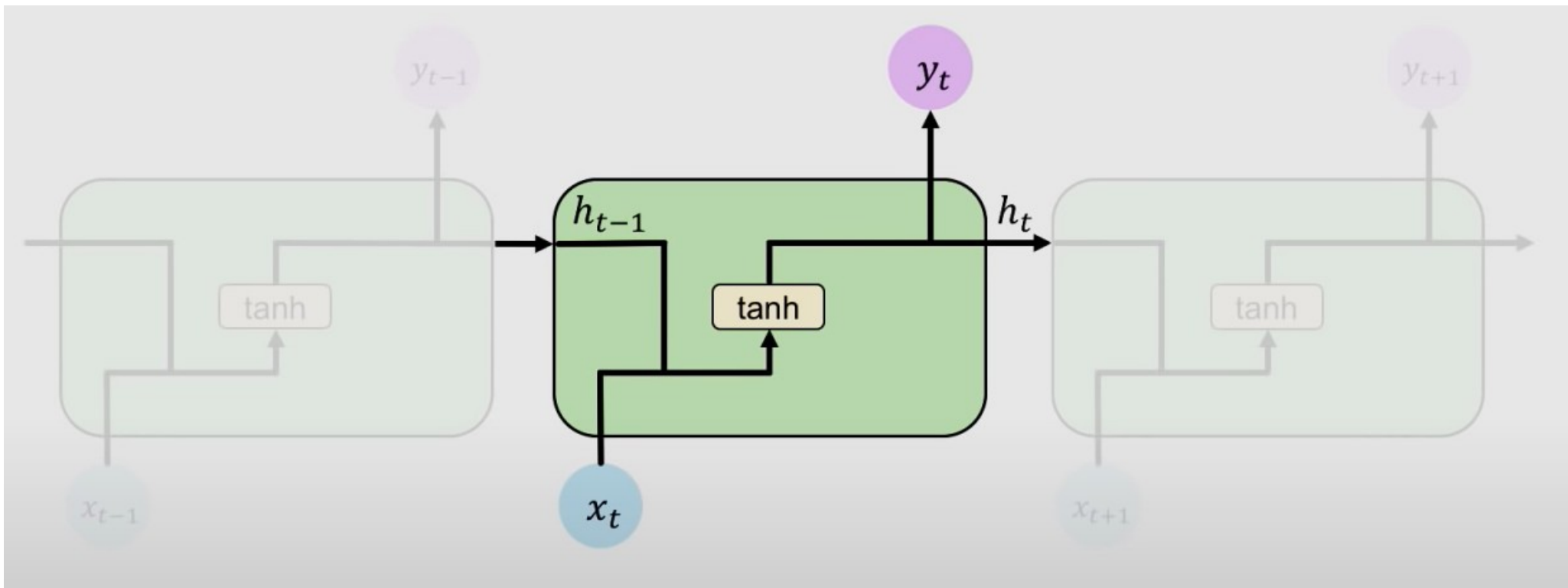


Visszacsatolt hálózati típusok

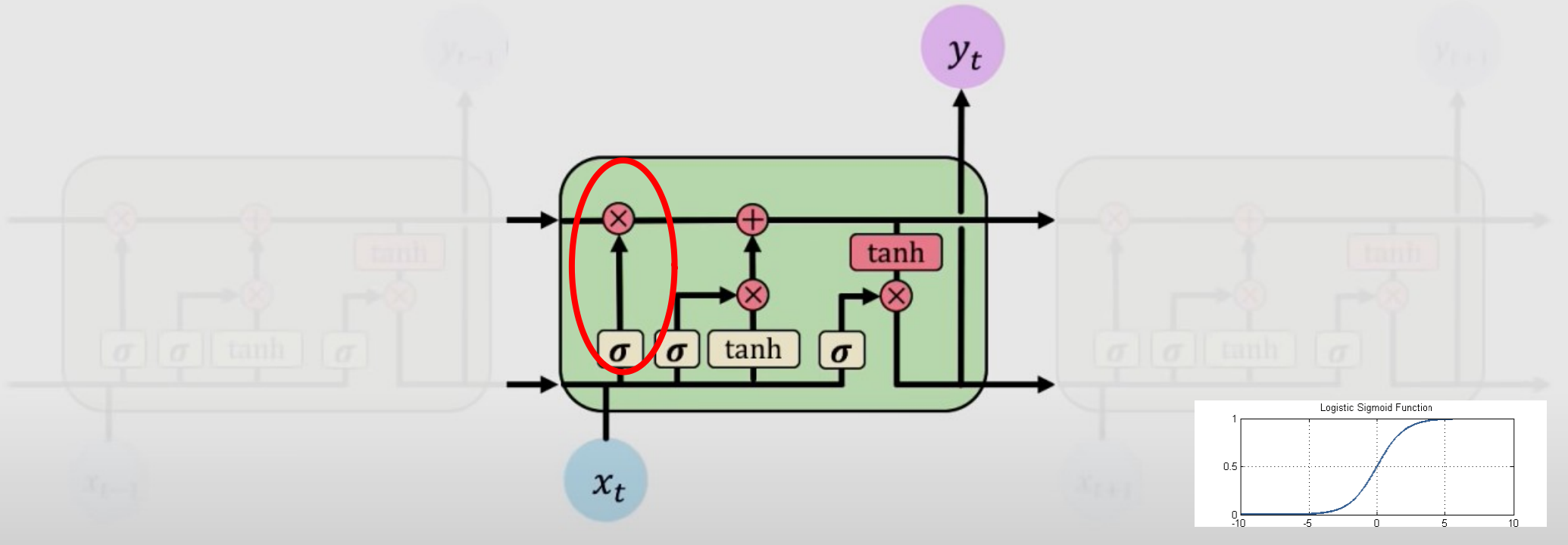
- **RNN: Recurrent Neural Network**
 - $\tanh()$ alkalmazó visszacsatolt hálózat
 - Probléma, hogy **milyen** hosszú a **hálózat memóriája**, hány elemig képes az RNN figyelembe venni a korábbi bemeneteket?
- **LSTM: Long Short Term Memory**
 - A **visszacsatolt RNN** egységben különböző **aktivációs** függvények felhasználásával számoljuk ki az állapotot tartalmazó kimenetet
 - Ezek kapcsolatait is **súlyozzuk**
 - Az aktivációs függvény lehetővé teszi a hálózat **különböző időtávon** történő felejtést
 - Pl. ha egy sorozat sok-sok **alsorozatból** áll, elég csak az alsorozatban akkumulálni az információt
 - A csomópontok működését **kapuk** szabályozzák, amiket a bemeneti adat működtet



Standard RNN



Long Short Term Memory (LSTM)

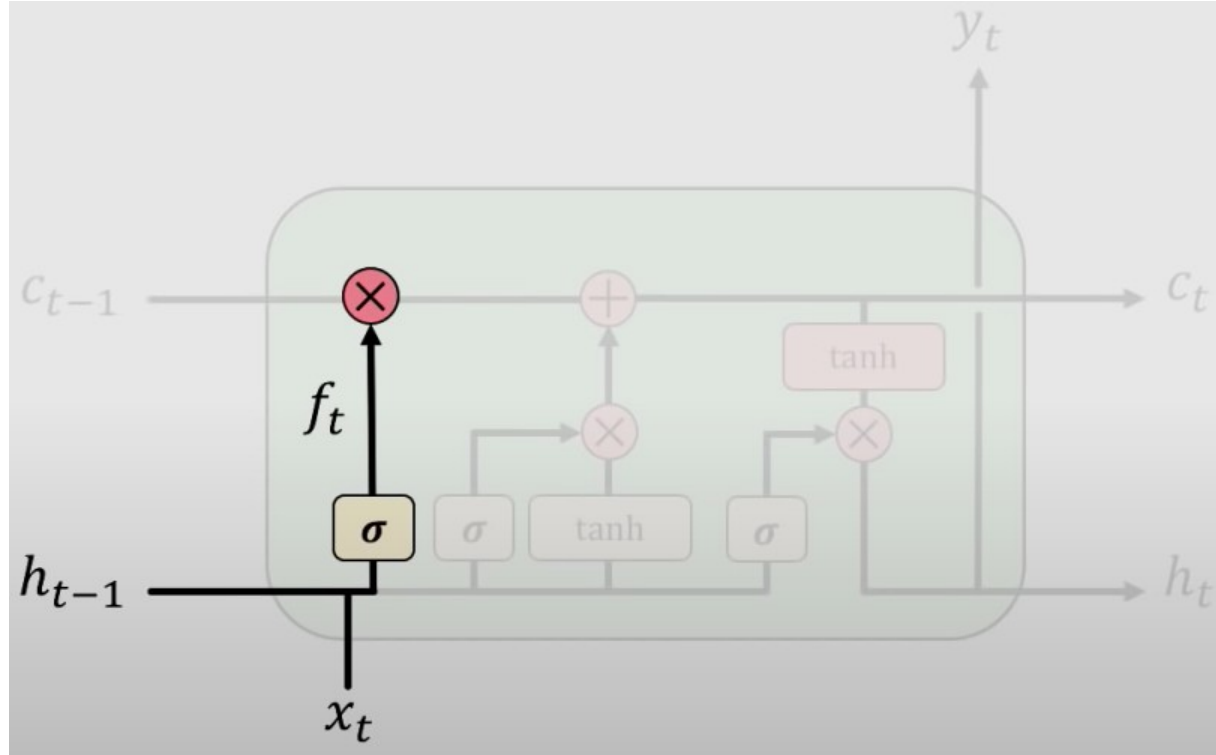


Szerepel egy ún. **forget gate** (σ) ami feltételesen engedi az információt át.
(sigmoid layer + pointwise multiplication)



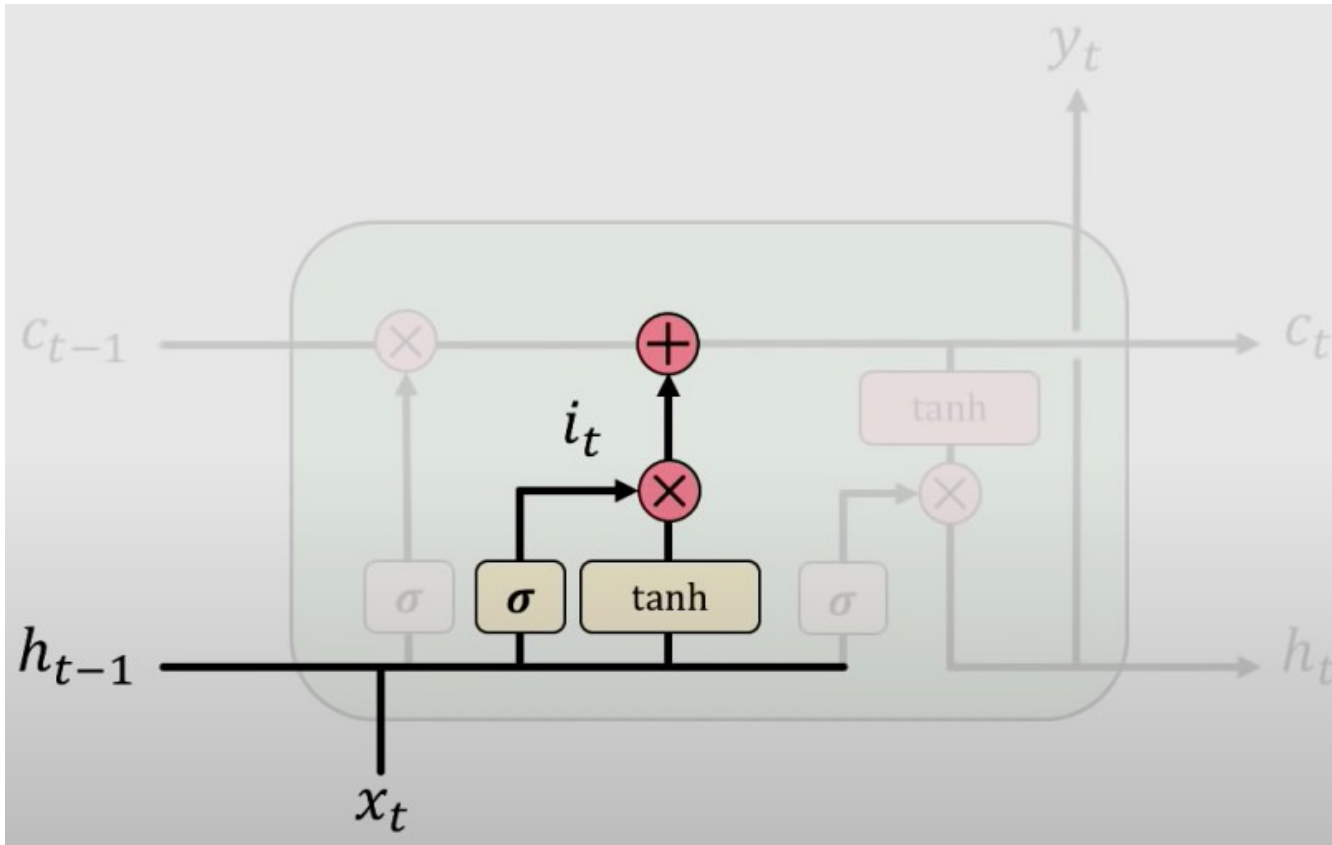
Long Short Term Memory (LSTM)

1. **Forget gate** "letiltja" a nem fontos információ további felhasználását az állapot kimeneten



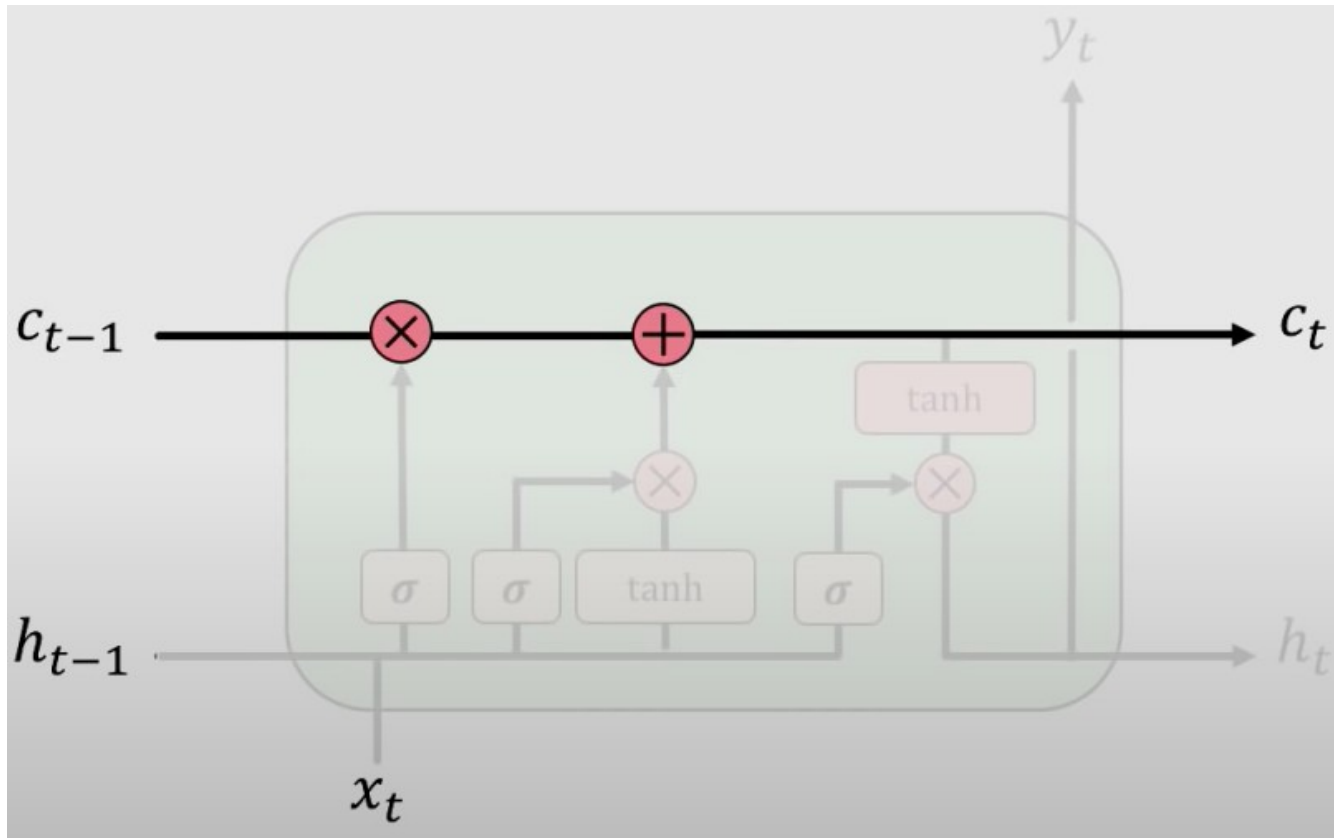
Long Short Term Memory (LSTM)

2. **Store** a fontos információt az előző kimenetről és a bemenetről hozzáadja a csomópont állapotához



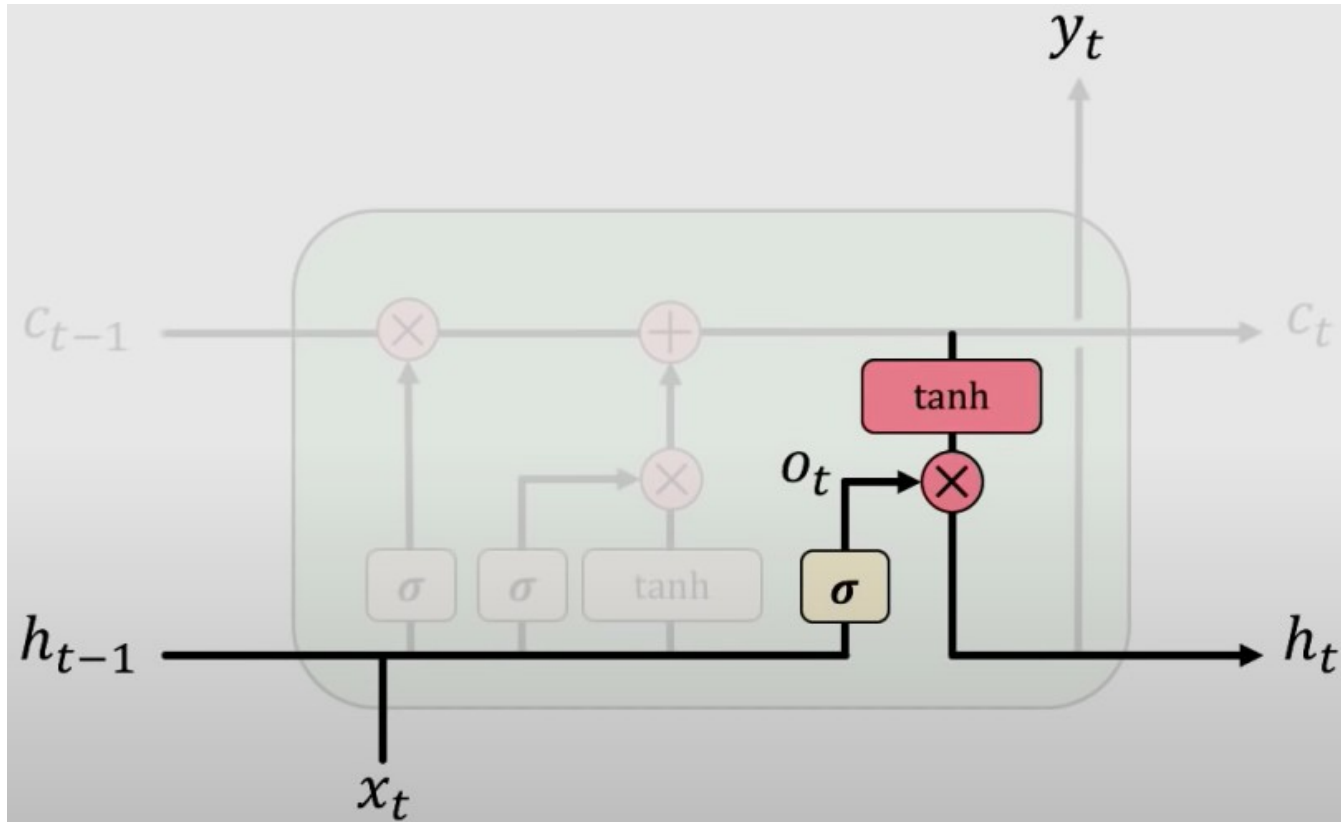
Long Short Term Memory (LSTM)

3. Szelektíven Frissíti csomópont állapotát



Long Short Term Memory (LSTM)

4. **Output** a csomópont állapotának megszürt értékét jeleníti meg a kimeneten, (ill. használja fel később)



Példák



Osztályozás konvolúciós neurális hálózattal

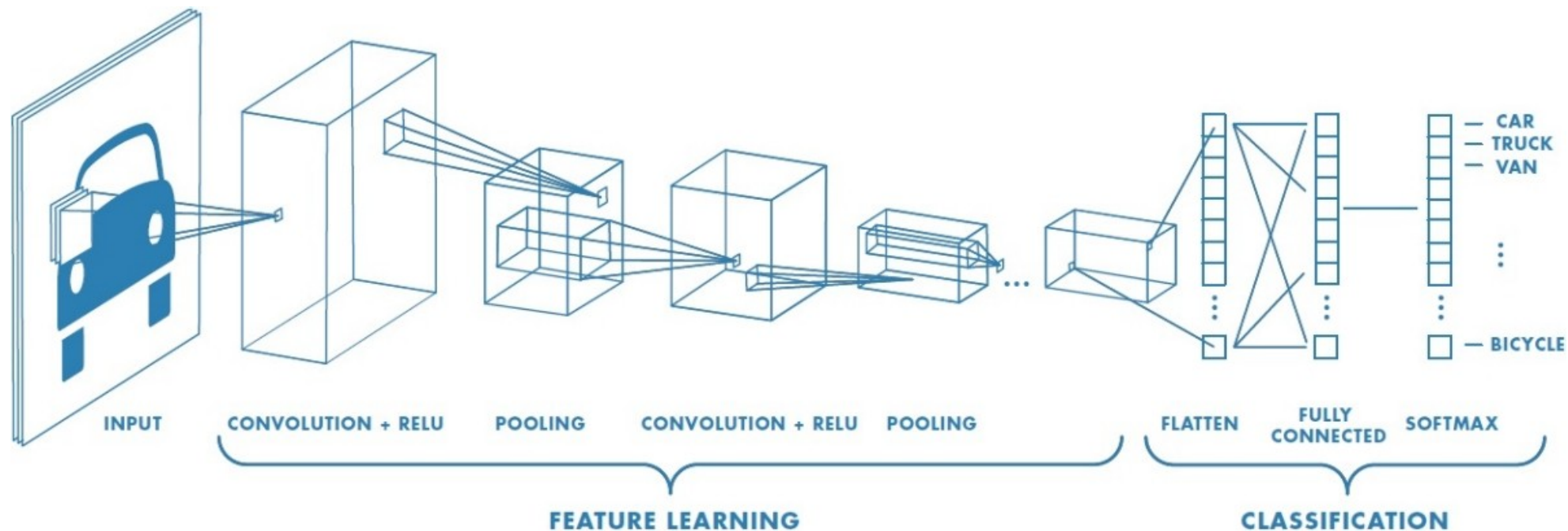


Osztályozás konvolúciós neurális hálózattal: Példák bemeneti adatokra

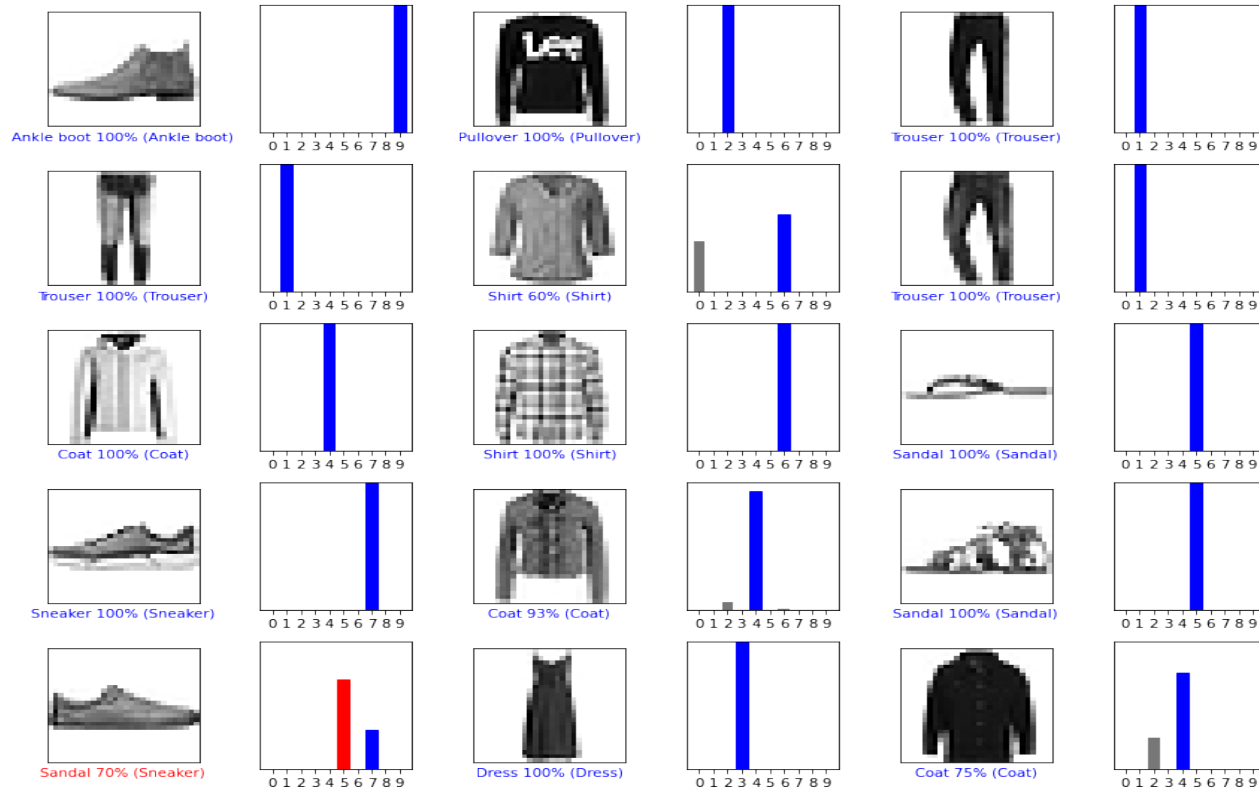
- **Tanító adathalmaz:**
60000 kép
- **Teszt adathalmaz:**
10000 kép
- **Kimenet:**
10 kimeneti osztály:
 - **Ruházati cikk típus**
 - **pl.: ing, pulóver, kabát, cipő stb.**



Hálózat



Osztályozási példák - hálózat osztályozásának eredménye



Konfúziós mátrix:

Hány bemeneti adatot osztályozott az egyes osztályokba a háló a vizsgált minták közül?



Tanulás visszacsatolt neurális hálózattal - példa



RNN példa

- **Tanítsuk meg egy visszacsatolt hálózatnak a drámaírást!**
- **Legyen a tanító halmaz a Rómeó és Júlia**
- **Bemenet egy ún. seed sorozat (pl. egy szó), a hálózat pedig folytassa a történetet**
- **RNN hálózat:**
 - Rómeó és Júlia regénnyel tanítjuk
 - Hálózat kimenete: minden betűhöz hozzárendel egy valószínűséget, ami megmondja, hogy az adott karakter milyen valószínűséggel lesz a következő karakter



Tanító halmaz: Shakespeare: Rómeó és Júlia

First Citizen:

Before we proceed any further, hear me speak.

All:

Speak, speak.

First Citizen:

You are all resolved rather to die than to famish?

All:

Resolved. resolved.

First Citizen:

First, you know Caius Marcius is chief enemy to the people.



Tanítás nélküli kimenet

Input:

```
b'hou diest.\n\nShepherd:\nI cannot speak, nor think\nNor dare to know that which I know. O sir!\nYou have '
```

Next Char Predictions:

```
b'ie3qWPT!NYFMF?eRVx, aYF!TKCcLfz-BPNFpQ3IWqgK\nvbvNvviP&.gEkTgWq?Iau,$DPPrHLWwZTFznJn?LBQP dwhLF,Z0tPte'
```



Kimenet rövid tanítás után

ROMEO:

We le destoradow?

No hen lead, and on me leiss ant;

Thill kinghate allimening greath has as hove to de reats,

Tho his as has

I in saycus he be ar wefhires thous feracfir, for thie the tord ame come a einss like charke:

I and wrees, denthand thimbous Quhered of

I pead rigtote Henese in the marsel-

Juk Hercaine: gru gas; ard queak,

'tisperm, leas a tith bept, af the wite Louberar earde,

Our, shreus no beat of from'st you tis wo dechise

And that is in it muepteen?

SRORIUCEME SICHIONGLERCEN:

A,

And of than out thie preirmend pethy centlesem be me bening:

Teen forlr wilk nos thes true all frebate

Thingededsenchour,

Mecpore, and commantifines, in your mepalo.

Whou, is andtar,

All rourdosed you anceing,

Ay ard your and bo vend ganth they of the tree-

Ond the ant he sornal.

LURI GLARD: ETay: Gie pried, Chave tiamss winge--o

my you, fay.

Come age ge

Bithe you suth ray, and thear Ixe?

KINGRCARTEN:

I ke makt, I we ba rein my giser be knisted:

If bous calve tit them oorins upl?

ASINGO: Shecfu



Kimenet hosszabb tanítás után

ROMEO:

My lawder mind power, that says made, for this news;
But what is to be called upon a tay.

BENVOLIO:

And in good under? if it passable to speak boy,
The lose chid resoluty to these payen shalt be sound.
Ay, and you are do bescuee; for it sourself.
You has a liaral, properate.

BRUTUS:

Heat now me, say, or it breath;
And therefole, weblet a man all happy treasonsome know
Thes that do you callariolage; you are excused!

AUTOLYCUS:

Yet ban, holder, fair hate? thus trumple of his entian';
If I know the beardiss of his salotion
That Kind and I will breek upon:
Let Bitin, your grace I have a king: I say we show
these doth may both causin it an hung to the Lord
Merry to my touldst men!
Our poor cry Warwick;
If these verd dishonour betber Bain and all what is pleasanis.
Thou art tells but the way this not at the quiate,
And will, but alike abroads, by capory
As well and I can there sell it but out in put of all this action
Do comerop trues: the coving Kath; am displace
To spide upon thy deea



Kimenet további tanítás után

ROMEO:

She will be an ask the old prover band away.

KING EDWARD IV:

Let them have see the passible to her made
As he that sought to be ingrateful woe.

FRIAR LAURENCE:

Not a solemn heir, nay, not, nor banishment,
That Richmond is the prompte of my mouth, meal,
In prosperity sex'd-clep-rate much happily are on:
Condemn'd out an entrails of them; but he use his eyes
Are flaxt out for many vouchsafery. Call hither Barnard,
Lace not to be the footing of his sort,
And live on can sweet eye dear malice.

KING RICHARD II:

What know'st thou not have touch'd his fruit-till
To you that came to lie will yet aride.

HASTINGS:

I do demandation; pierced with a tal.

WARWICK:

Grandam, off without dun, stir for the Duke of day or would,
As if that joy, untappinutary
And the entertainment in his broken.
I think for this, boy, to must be prepared.

DUKE VINCENTIO:

I do beseech you, mark-how me to kill a witnior.

PRINCE EDWARD:

I thank the matter with my traitor to the crown;
And I'll send for policting afte



Demó

- **Google Colab link a demóra:**
- **<https://colab.research.google.com/drive/1IZbCIAjD3YYv3lQQQAPkokrHzC3C8mk0?usp=sharing>**
- **A notebook utasításait a böngészőben végre lehet hajtani!**



RNN demó



Szegmentálás Unet segítségével

- **Képek szegmentálása, macskák és kutyák azonosítása**
- **Google CoLab link a demóra:**
- **<https://colab.research.google.com/drive/1L20NSqBj1vc--X4CIhE07Y5apaEh1LUQ?usp=sharing>**
- **A notebook utasításait a böngészőben végre lehet hajtani**



Köszönöm a figyelmet!

