



**Budapesti Műszaki és Gazdaságtudományi Egyetem**  
Villamosmérnöki és Informatikai Kar

Pongrácz Ádám (JGOB5J)

# **MODEL ENSEMBLE FOR MEDICAL IMAGE SEGMENTATION**

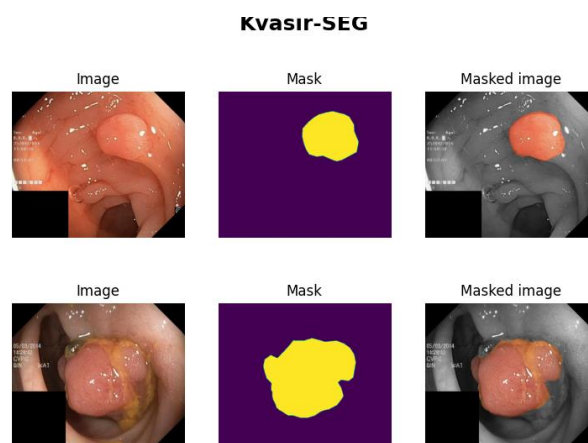
BUDAPEST, 2024

# 1 Feladat

Egy hatékony technika a modellek teljesítményének növelésére a model ensemble-k használata. Ez azt jelenti, hogy több modellt kombinálunk annak érdekében, hogy pontosabb és robusztusabb megoldást érjünk el, mint amit bármelyik modell önállóan nyújthat. A házi feladat során orvosi kép szegmentálással foglalkoztam, model ensemble alkalmazásával. Az orvosi kép szegmentálás (medical image segmentation) az a folyamat, amely során egy orvosi képből az egyes anatómiai struktúrákat vagy elváltozásokat külön-külön, jól körülhatárolt részekre bontják. Ennek célja, hogy a kép informatívabb és könnyebben feldolgozható legyen az orvosi diagnózisokhoz és elemzésekhez.

## 2 Adathalmaz: Kvasir-SEG

A Kvasir-SEG [1] [2] adathalmaz egy orvosi képadatbázis, amelyet főként a vastagbél endoszkópiás képeinek szegmentálására használnak. Az adathalmaz 1000 **endoszkópos képet**, melyek a vastagbél különböző szakaszait ábrázolják, valamint a hozzájuk tartozó **maszkot** tartalmaz, amik kijelölik a polipok és egyéb elváltozások helyét a képeken. A Kvasir-SEG célja, hogy az orvosok könnyebben felismerhessék a problémákat kolonoszkópiás képek elemzésekor. [3]



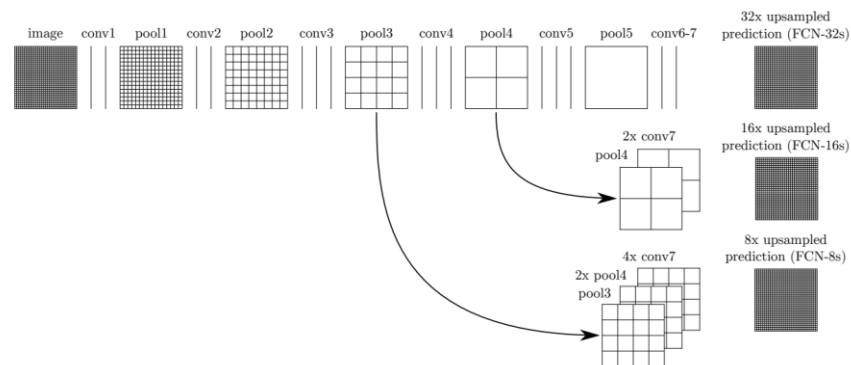
### 3 Modellek

### 3.1 Baseline

A baseline modellnek egy egyszerű 4 konvolúciós blokkból álló modellt használtam.

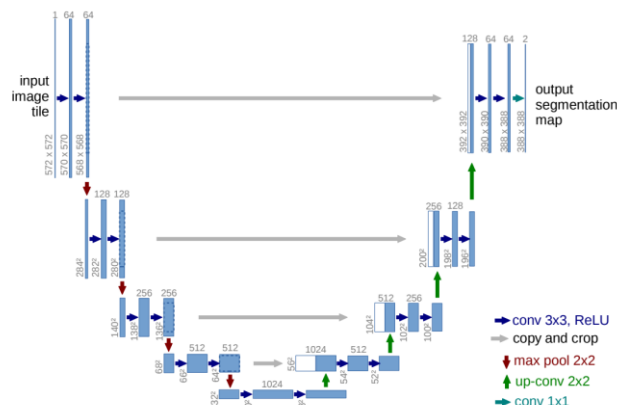
### 3.2 FCN

A **Fully Convolutional Networks (FCN)** [4], egy olyan architektúra, amelyet elsősorban szemantikai szegmentálásra használnak. Az architektúra konvolúció, pooling és upsampling rétegekből épül fel.



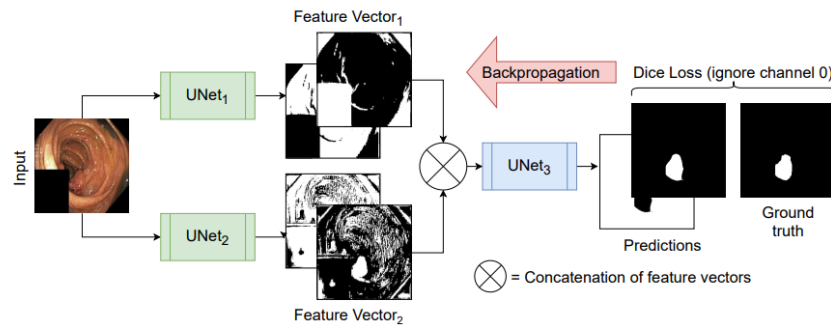
### 3.3 UNet

A **UNet** [5] egy mély tanulási architektúra, amelyet kifejezetten képszegmentálásra fejlesztettek ki, és különösen jól alkalmazható orvosi képek szegmentálására. A modell nevét a U-alakú szerkezetéről kapta. A hálózat 2 részből áll: lemenő részből, ami a tipikus konvolúciós blokkokból épül fel, valamint a felmenő részből, ami Upsampling blokkokkal egészül ki.



### 3.4 TriUNet

A **TriUNet** [6] architektúra a UNet továbbfejlesztése. A modell 3 UNet-ből épül fel. Az első lépésben a képet keresztülküldjük két párhuzamos UNet modellen. A kapott eredményeket ezután a modell összefüzi és átküldi egy harmadik UNet modellen, ami kimenetként előállítja a prediktált maszkot.



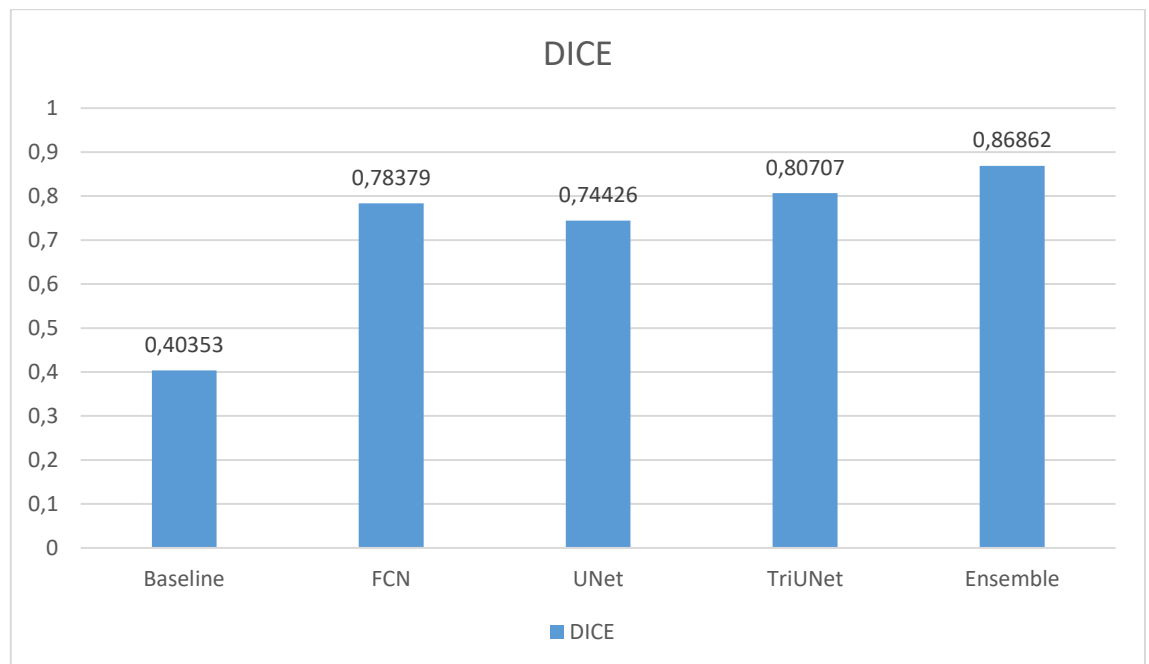
### 3.5 Ensemble

Az Ensemble modell egy olyan gépi tanulási megközelítés, amely több modellt egyesít a jobb teljesítmény és a robusztusság elérése érdekében. Az általam előállított modell 3 részből áll: egy FCN, egy UNet, valamint egy TriUNet modellekből. Az előbbiek eredményét átlagolva állítja elő a prediktált maszkot.

## 4 Eredmények

A modellek kiértékeléséhez DICE metrikát használtam. A DICE metrika (más néven DICE koefficiens vagy F1-score) egy mérőszám, amelyet elsősorban a képszegmentálásban használnak annak meghatározására, hogy a modell milyen jól végzi el a predikciót egy adott objektum vagy régió szegmentálásakor.

Az alábbi ábrán látható a betanított modellek tesztadatokon elért eredménye. Mint látható a legjobb eredményt a Ensemble modell érte el 0,86-os DICE értékkel.



## 5 Irodalomjegyzék

- [1] „Kvasir-SEG: A Segmented Polyp Dataset,” [Online]. Available: <https://arxiv.org/pdf/1911.07069v1>.
- [2] „Kvasir SEG,” [Online]. Available: <https://datasets.simula.no/kvasir-seg/>.
- [3] „A multi-centre polyp detection and segmentation,” [Online]. Available: <https://arxiv.org/pdf/2106.04463>.
- [4] „Fully Convolutional Networks,” [Online]. Available: <https://arxiv.org/pdf/1605.06211>.
- [5] „U-Net: Convolutional Networks for Biomedical,” [Online]. Available: <https://arxiv.org/pdf/1505.04597v1>.
- [6] „DivergentNets: Medical Image Segmentation by,” [Online]. Available: <https://arxiv.org/pdf/2107.00283>.