Szegmentáló Neurális Hálózatok Robusztussága

Halmosi Levente Ferenc
II. évf. programtervező informatikus

Témavezető: Jelasity Márk

SZTE TTIK Számítógépes Algoritmusok és Mesterséges Intelligencia Tanszék

Napjainkban az ellenséges (adversarial) példák jelensége egy erősen kutatott terület a neuronháló alapú számítógépes látáson belül, amelynek lényege, hogy a bemenet szabad szemmel láthatatlan ellenséges perturbációival egy neuronháló kimenetét tetszőleges módon meg lehet változtatni. A kép klasszifikáció területén ez a téma egységes módszertannal rendelkezik, ebből adódóan, az ebben a témában megjelenő munkák összehasonlítása könnyedén lehetséges, az eredmények igazolhatóak. Ezzel szemben az ellenséges példák jelensége egy meglehetősen, nem egységes, alulkutatott terület, a neurális hálózat alapú képszegmentálás területén. A kutatás során létrehoztunk egy validáló algoritmus csomagot, amellyel sikerült igazolnunk azon hipotézisünket miszerint a jelenleg state of the art robusztus szegmentáló modell, valójában nem robusztus. Ezen túlmenően létrehoztunk egy olyan szegmentáló modellt, amely jelenleg a legnagyobb robusztussággal rendelkező szegmentáló háló. Mindezzel együtt megalkottunk egy módszertant, amely reményeink szerint alkalmas lesz a módszertani kohézió megteremtésére a témán belül.