

Bírálat

Benda Krisztián "Feliratok illesztése képekre és videókra mesterséges intelligencia használatával"
című Diplomatervéről

1. A diplomaterv feladat tárgya, a célkitűzés időszerűsége.

A diplomaterv a képeken és videókon való feliratozás témakörét járja körbe és megoldásokat javasol képfeldolgozási algoritmusokkal és mesterséges intelligencia modellek segítségével.

A téma nagyon aktuális, napjainkban egyre nagyobb számban jönnek létre videó tartalmak, amelyeknek a felirat kézi pozicionálása nagyon sokba kerülne vagy nem is lenne rá idő, emberi kapacitás. A mesterséges intelligencia és képfeldolgozási eljárások már képesek komplex információ tartalmat kinyerni egy képből, így ezeknek a feliratok pozicionálásához való felhasználása egy igen jó kombinációnak tűnik.

2. A tartalom összefoglalása néhány mondatban.

A dolgozat a képeken és videókon található feliratozás témakörét és ennek nehézségeit mutatja be részletesen. A munka során egy saját rendszert ismertet és implementál, amely képeken és videókon is képes optimális helyet találni a feliratnak. Az optimális hely keresésére több típusú megoldást mutat be, amelyeket több fajta detektor segíti információkkal. A dolgozat végén a 3 különböző algoritmust kiértékeli futási idő, jósági tényező és felhasználói visszajelzések alapján.

3. A jelölt a kitűzött feladatot megoldotta-e?

A dolgozat a feladatkiírásnak megfelel. A dolgozat felépítése a tervezéstől a megvalósításig jól átgondolt és alapos, minőségi munkát mutat.

A diplomatervben külön szeretném kiemelni, hogy a videók publikus elérhetősége nagyban segítette a dolgozat eredményének szemléltetését. Illetve az eredmények kérdőíves tesztelése is jól mutatja, hogy a hallgató számításba veszi az emberi tényezőket is és nem csak a nyers, numerikus eredmények alapján értékeli ki a munkáját.

4. A kidolgozás értékelése.

4.1. Irodalomkutatás

Az irodalomkutatás során a hallgató a témakörbe illeszkedő, aktuális publikációkat és meglévő algoritmusokat mutatja be. Olyan embernek is könnyen érthető, aki még nem foglalkozott részletesebben ezzel a problémával.

4.2. A dolgozat szerkezete és stílusa

A dolgozat szerkezete és felépítése jó, kevés elgépelés és nyelvi pontatlanság található benne. Az ábrák jól értelmezhetőek és nagyban segítik a témakör megértését.

4.3. A dolgozat tartalma

A diplomaterv tartalma egy minőségi mérnöki munkát mutat. A rendszer megtervezésétől egészen a kiértékelésig a lépések jól követhetőek és következetes döntéseket találhatók benne. A dolgozatban kiemelendő, hogy nem csak 1 hanem 3 algoritmust is bemutat és meg is valósít a hallgató.

A detektor algoritmusok futási idejénél jó lett volna tudni, milyen hardveren lettek kiértékelve. Illetve az objektum detektáló Yolo és az arc detektor YoloV3-nál a két hálózat különbsége érdekes lehetett volna még. A többi algoritmus és a paraméter választások alaposan ki vannak fejtve.

A rendszer kiértékelése több tényező alapján is történik, ami jól mutatja, hogy a rendszer által létrehozott videókat nem elég csak jósági metrikák és futási idő alapján kiértékelni.

5. Összefoglaló értékelés

A diplomaterv a kiírásoknak megfelel, formátumra és tartalomra is egy minőségi mérnöki munkát mutat, a dolgozatot elfogadásra javaslom.

6. Kérdések

A Yolo hálózat futtatási sebessége milyen módon változna, ha egy erős GPU gyorsítással futna a kiértékelés? Így is kiemelkedne a többi algoritmus futási ideje közül?

Az arc detektálásra milyen alternatív megoldásokat lenne érdemes megpróbálni? (pl: Haar-cascade, SSD, stb..)

Dátum: 2021 március 30.

Javasolt jegy: 5



Gróf Attila
BME MSc, Citi