

Konvolúciós hálózatok általánosító képességének javítása 3D augmentációs technikákkal logó osztályozási feladatokon

*Kardos Péter, Váradi Endre Tamás
II. évf. Programtervező informatikus, Msc*

*Témavezető: Megyeri István
SZTE TTIK Számítógépes Algoritmusok és Mesterséges Intelligencia Tanszék*

Kijelenthetjük, hogy az utóbbi időkben a képi információ automatikus értelmezése egyre nagyobb figyelmet kap. Ezen belül is igen felkapott a logó felismerés témája, ahol jól meghatározott alakzatok osztályozása a cél, mint például márka jelzések besorolása. Ilyen rendszereket több helyen tudunk alkalmazni, többek között termék megjelenések ellenőrzésére, szerzői jogok megsértésének észlelésére, hirdetések személyre szabására. A problémát igazán nehezzé teszi, hogy egyetlen logónak több variációja is létezhet, továbbá a valós képeken a körülmények (fények, háttér, kamera beállítás) nagyban befolyásolják a felismerés hatékonyságát. Népszerű megközelítés hasonlóan komplex feladatok megoldására neurális hálók alkalmazása. Azonban betanításuk nagy mennyiségű adatot igényel és a tanítás során nem látott körülményekre nehezen általánosít. Az említett problémákra megoldást jelenthet különböző adat augmentációs módszerek használata a tanítás során.

A dolgozatunkkal a célunk egy 3D-s térben augmentáló program bemutatása, mellyel versenyképes eredményeket tudunk felmutatni a logó felismerés problémában. A 3D-s térben való augmentálás lehetőséget ad a valós képek készítésekor fellépő torzítások közelítésére, ezzel is jóval több lehetőséget adva a képek átalakításához az elterjedt transzformátorokhoz képest. A transzformátorunk kimenetén tanítunk egy neurális hálót, amelynek feladata a transzformált képek besorolása a megfelelő osztályokba. Használt háló architektúrának a manapság legelterjedtebb előtanított modelleket választottuk, melyek közül többön is méréseket végeztünk. A megoldásunkat több logó osztályozási feladaton is kiértékeljük és összehasonlítottuk a State of the Arttal, mint például Logo2k+, melyen jobb eredményt értünk el a ma publikált legjobb rendszerrel szemben. Ezen felül megmutatjuk, hogy az általunk definiált augmentációkat alkalmazva a kapott modell a nem látott transzformációkon is jelentős javulást hoz.

Keywords: Logó felismerés, Képosztályozás, kép augmentálás