Kontextus alapu érzelemdetekci6

Keszitette: Szaniszlo Csongor Adam

Konzulens: Dr. Hullam Gabor, Revy Gabor

Kibóvitett Absztrakt

Az rzelemdetektalasnak szamos felhasznalasi területe van, ilyen peldaul az egeszsegügy, ahol mentalis betegsegeket tudnak detektalni ezaltal,a vevöszolgalat, ahol ennek segitsegeve szemelyre szólobb ajanlasokat tudnak adni a cégek vagy akar az oktatas is, ahol peldaul egy tana kaphat bovebb visszajelzest az orajarol. Igy tehat az erzelemdetektalasnak, mint feladatnak szamo motivacioja van.

Ugyanakkor mig ez emberek szamara egy egyszerübb feladat, hiszen emberkent ezt születesünk ta ,tanuljuk, addig a szamitogepek szamara ez mar közel sem egy egyertelm megoldassal rendelkezo problema. Ennek nagyon sok oka lehet, peldaul, hogy a bemeneti kepek vagy videofelvetelek minösege nem feltetlen megfelel, amely megneheziti a gepek feladatat. A neuralis halók pontosabban a Convolution Neural Network (CNN) megjelenese jelentösen elosegitette a jobb eredmenyek elereset ezen a területen.

Az utobbi években szamos kutatas foglalkozott melyebben ezzel a temaval, melyek soran a fo cel annak megallapitasa volt, hogy különboz jellemzok/bemeneti informaciók, mint peldaul egy kep az arcrol vagy egy hangfelvetel hogyan segitik el az erzelmek helyes detekciojat. A temalaboratoriumi munkam soran is peldaul egy ilyen kutatasnak olvastam utana, lletve probaltam ki a cikk altal javasolt megoldast. Ennek soran a FER2013 (Facial Emotion Recognition)] adathalmazzal_foglalkoztam,amely az erzelemdetektalasra hasznalt_ adathalmazok_ kozül kiemelkedoen fontos, gyakorlatilag egy mercejeve valt az arc alapu erzelemdetektalasnak.

Ugyanakkor pszichológiailag is alatamaszthato, hogy az erzelemdetektalas soran tobb aspektust is figyelni kell ahhoz, hogy helyesen tudjuk behatarolni a vizsgalando szemely erzelmi allapotat. Igy peldaul az elozöekben emlitett FER2013-ra visszautalva nem eleg az, ha csupan egy szemely arckifejezeset vizsgaljuk

1. es 2. abra: Kivagott arckép az EMOTIC(2] adathalmazbol, illetve a hozza tartozo kép, kontextussal

Peldaul a fenti kep eseten is lathato, hogy ha csupan az arcat vizsgaljuk a szemelynek, akkor valamilyen negativ érzelemmel asszocialnank az erzelmi allapotat, mig ha a kontextust is figyelembe vesszük, akkor mar evidens, hogy a szemely valoszinüleg a sikerenek orvend.

Ezt figyelembe veve tehat az &rzelemdetektalas soran erdemes az arcon kivul mast is elemezni. Ilyen peldaul a kontextus, amelyben az alany reszt vesz, hiszen ebbol megallapithato az, hogy a szemely boldog, ha peldaul egy koncerten van. Egy masik aspektus, amit erdemes fontolora venni, az az alanyrol kivagott kep, amit a kontextust tartalmazo kepbl nyerünk ki. Ezen kep alapjan elemezheto az alany testtartasa, ami ugyanügy jelentosen hozzajarulhat az erzelemdetektalashoz.

Az onallo laboratoriumi munkam soran tehat a kontextus alapu erzelemdetektalassal foglalkoztam A motivacio az volt, hogy megvizsgaljam, hogy a kontextus tenylegesen mennyire befolyasolja az erzelemdetektalast.

A munkam soran eloszor egy publikacioval, illetve az irok altal letrehozott adathalmazzal, az EMOTIC-kal foglalkoztam. Az EMOTIC egy olyan adathalmaz, amely minden egyes vizsgaland szemelyhez ket kepet tarsit. Az els egy olyan kep, amely a szemelyt egy adott kontextusban jeleniti meg. A masodik pedig az els kepbol egy kivagott kep, ez magat a szemelyt tartalmazza.

A publikacio iroi egy modellt is bemutattak, amin megvizsgaltam, hogy az eszkoz milyen eredmenyeket kepes elerni az adathalmazon. A modell az elobb emlitett ket bemeneti kep alapjan hozott ket dontest. Az egyik dontes 26 &rzelmi kategoriabol tetszoleges mennyisegübe sorolta be a kepet, mig a masodik dontes 3 folytonos ertek prediktalasara vonatkozott, amely ertekek a VAD/PAD[3] modell szerint értekeltek a szemely érzelmi allapotat. A kiertekelés soran a kapott

eredmenyeket_osszehasonlitottam a publikacioban leirt eredmenyekkel. Bar ez nem sokban különbozott (az alkalmazott metrika szerint 2%-os volt az elteres), megis arra a megfontolasra jutottam, hogy a modell altal elert eredmenyek viszonylag alacsonyak, igy megprobaltam javitani az eszkoz teljesitmenyen.

Ennek soran szamos megkozelitessel probalkoztam. Eloszor megprobaltam image padding-et hasznalni. Ennek motivacioja az volt, hogy az adat elofeldolgozas soran ugy iteltem meg, hogy a csupan csak a szemelyt tartalmazo kep eseteben a kep atmeretezesenek hatasara tul sok informacio veszik el, amely megneheziti a modell feladatat a dontesben. Ezenkivül megvizsgaltam az adathalmazt es azt figyeltem meg, hogy a kepek nagy reszénel a vizsgaland szemely arca reszben, vagy teljes egeszeben nem latszodik. Ugy iteltem meg, hogy bar nem elegseges feltetele az erzelemdetektalasnak az arc elemzese, de ugyanakkor szükseges, igy az adathalmazbol kiszürtem az arcot ,nem tartalmazo képeket. A harmadik megkozelités pedig egy j modell integralasa volt, amely a FER adathalmazon lett elötanitva. Ezzel a celom az volt, hogy a kontextuson, illetve a testtartason kivil a szemelyek arcat is elemezze a modell, ezzel segitve azt a dontesben.

adathalmazt is annak remenyeben, hogy a modell azon jobban teljesit majd. Ez az adathalmaz a HECO[4) volt. A HECO nagyon hasonlit az EMOTIC-hoz, az egyetlen különbseg az, hogy 26 erzelem helyett 8 erzelmi cimke van, illetve minden kephez egy darab erzelmet tarsitottak az alkotok.

A HECO-n val6 kiértekeles soran sajnos ugyanugy teljesitmennyel kapcsolatos problemakba utkztem, igy az elozoekben bemutatott megkozelitesekkel probaltam a modell teljesitmeny javitani. Ebben az esetben sem tapasztaltam semmilyen javulast.

Az eredmenyek hatasara ügy dontottem, hogy celszerü az alkalmazott modell, illetve a hasznalt adathalmazok melyebb vizsgalata. A modellek vizsgalatanal arra jutottam, hogy a vizsgalando szemelyt elemz modell, illetve az arcot elemz modell egyik adathalmaz eseteben sem volt kepes relevans jellemzoket kiemelni. Ennek az elobbi esetben az volt az oka, hogy a modell egy olyan adathalmazon lett eltanitva, amely egyaltalan nem kapcsolodik az erzelemdetektalashoz, mig utobbi eseteben a problema a modell komplexitasaval (EMOTIC) volt vagy az adathalmaz inkonzisztens cimkézes&vel (HECO). A kontextust elemz modell vizsgalataval igy arra jutottam, hogy a modell feladata erzelem-kontextus parok osszekapcsolasa volt, viszont megvizsgaltam azt

is, hogy a modell milyen kepre milyen kontextus kimenetet adja, es arra jutottam, hogy semmilyen korrelacio nincs az erzelmek es a kontextusok kozott

Az adathalmazok vizsgalata soran felfedeztem, hogy mindket adathalmazzal tobb problema volt. Peldaul hibas bounding box-ok (egy dobozon belül tobb szemely is van), inkonzisztens cimkezes, illetve az annotalt szemelyek arcanak nagy resze nem latszodik bar utobbi javithato, ugyanakkor a kepek kiszürese nem cel, ha azt szeretnenk tesztelni, hogy egy modell hogyan teljesit egy adott adathalmazon

Forrasok

[1] FER2013 - https/www.kaggle.com/datasets/msambare/fer2013

[2] EMOTIC website - https://s3.sunai.uoc.edu/emotic/index.html

[31 VAD/PAD model - https://en.wikipedia.org/wiki/PAD_emotional_state_model

[4] HECO website - https://heco2022.github.io/