Haladó adatelemzési módszerek labor

Készítette:

* Bálint Gergő
* Juhász Benedek László
* Tumay Ádám

[*Repository: https://github.com/Tadam16/TeethAndCariesDetection/*](https://github.com/Tadam16/TeethAndCariesDetection/)

**Projektterv**

Feladatunk felnőtt páciensek panoráma röntgen felvételein fogszuvasodás szegmentálása, detektálása. A megoldás motivációja segíteni a diagnózis felállítását, illetve a gyorsabb, megbízhatóbb döntéshozatalt az esetleges fogorvosi beavatkozások tekintetében.

A feladatot két részre bontjuk. Először a felvételeken szegmentáljuk a fogsort már meglévő módszerekkel képenként egy maszkot létrehozva, majd ezután az így előkészített adatokon készítünk modellt a fogszuvasodás(ok) lokalizációjára. Az első lépés így az adatelőkészítési folyamatunkban szerepel.

Az előszegmentációs modell tanításakor csak felnőtt páciensek mintáira tanítunk, gyerekek esetében komplexebb feladatról lenne szó a biológiai eltérések miatt (tejfogak és fejlődő csontfogak egyidejű megjelenése a felvételeken), mely a feladat mintaigényét túlságosan megnövelné.

**Irodalomkutatás**

A feladat irodalomkutatási fázisában a következő forrásokat vizsgáltuk:

* <https://www.kaggle.com/code/kaledhoshme/adult-tooth-segmentation-u-net-based-gan/notebook> (U-NET, GAN)
* <https://paperswithcode.com/paper/s-r2f2u-net-a-single-stage-model-for-teeth> (S-R2F2U-Net)
* <https://www.mdpi.com/2076-3417/13/23/12771/pdf?version=1701181876> (U-NET)
* <https://arxiv.org/pdf/2012.13666.pdf> (PaXNet)
* <https://www.di.ubi.pt/~hugomcp/doc/joao_oliveira_2008.pdf> (szakértői képfeldolgozási módszerek)

**Adatelőkészítés**

A modellek tanításához a következő adathalmazokat gyűjtöttük össze:

* <https://www.kaggle.com/datasets/humansintheloop/teeth-segmentation-on-dental-x-ray-images/data>
* <https://www.kaggle.com/datasets/thunderpede/panoramic-dental-dataset>

Az adatelőkészítés folyamán a következő lépéseket tettük:

* **Normalizációs lépés**: Képenként egyesével normalizáltunk, az intenzitás alsó és felső 3%-át clip-eljük, illetve 0-1 tartományba képzünk le.
* **Adattisztítás**: Mivel több különböző forrásból származó képekkel dolgozunk, melyekhez szegényes leírás áll rendelkezésre, ezért fontosnak tartottuk a fájlduplikátumok kiszűrését, melyeket fájlnév alapján nem lehet kiszűrni, így a fájlokat hash-eik alapján hasonlítottuk össze, és szűrtük ki az azonosakat.
* **Előtanítás**: Saját U-Net struktúrával (FancyUNet néven, egy UNet alapú DenseNet-szerű háló háló), illetve SegFormer architektúrával (egy transzformer alapú háló) az adathalmazon fog-szegmentációt végeztünk, a későbbi fogszuvasodás-keresés elősegítésére. A hálókat azért választottuk ilyenre, mert eltérő architektúrájúak, és mert korábbi tárgyban már vizsgáltuk a működésüket.

Az így elkészült modellek predikcióiból készült képek, melyeken bal oldalon a predikció a fogakra, jobb oldalon a valódi maszk látható, az 1. és 2. ábrán láthatók.

A képen röntgenfilm, művészet látható

Automatikusan generált leírás

*1 Figure: FancyUNet tanítás után*

A képen vázlat, művészet, röntgenfilm látható

Automatikusan generált leírás

*2 Figure: SegFormer tanítás után*

**Előszegmentáció bemutatás (Adatvizualizáció bemutatás helyett)**

Jól teljesítenek a fogakat előszegmentáló modelljeink, és numerikusan hasonló eredményt produkálnak (ezek az alábbi táblázatban láthatók). Ezen felül az eredmények kvalitatív vizsgálatakor azt találtuk, hogy az egyes modellek predikciói jól kiegészítik egymást (lásd az ehhez tartozó táblázat képein).

|  |  |  |  |
| --- | --- | --- | --- |
|  | train | validation | test |
| UNET | **0.916** | **0.905** | **0.913** |
| Segformer | **0.916** | **0.915** | **0.917** |

Megfigyeltük továbbá, hogy az egyes képeken a detekció szempontjából kritikus a megfelelő kontraszt-viszony. Emellett a nagyfrekvenciájú komponensek kiszűrése nem okoz nehézséget az egyes képletek felismerésében.

|  |  |  |
| --- | --- | --- |
|  | Segformer | FancyUNet |
| Test |  |  |
| Validation |  |  |
| Carries detection dataset |  |  |
| Other |  |  |

Az előszegmentáló (fogszegmentáció, presegment) részen végeztünk további méréseket. A fogszegmentációink a szuvasodásos területeknek mekkora részét fedik le, ha 50% threshold mellett küszöböljük. Ezt külön megnéztük a UNet modellre, a Segformer modellre, és a kombinált modellre, ahol a kimenet a két modell kimenetének elemenkénti maximuma.

Az eredményeink:

|  |  |
| --- | --- |
| FancyUNet caries coverage | 0.92 |
| Segfomer coverage | 0.95 |
| Combined coverage | 0.98 |
| Bad predictions (under 30% dice score) | 0 |

**A szuvasodás szegmentáló modellek tanítása**

A következő lépésünk a szuvasodás szegmentáló eljárás elkészítése. Ennek bemenete három elemből áll, ezek a kép és a két preprocesszáló modell kimeneti maszkjai. Ezzel a módszerrel biztosítjuk annak lehetőségét, hogy a modell figyelembe vehesse mindkét pixel szintű osztályozó kimenetét. A tervezett modellek architektúrái a bemenetet leszámítva megegyeznek az előfeldolgozást végző architektúráival.

Betanítottuk a FancyUNet és a Segformer modelljeinket a Panoramic Dental Dataset adathalmazon.

Az eredményeket dice-score metrika tekintetében nem találtuk kielégítőnek, ezért további adathalmazokat kerestünk. A Roboflow adatmegosztó oldalon találunk két adathalmazt, amelyek “Vzrad2”, valamint a “New final dataset” névre hallgatnak, de nem találtunk róluk több leírást.

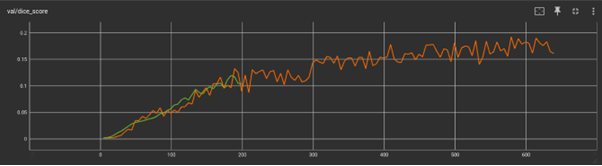
A duplikátumok hash collision alapú kiszűrése után újból elvégeztük a tanítást a kibővített adathalmazra (~7000 minta), azonban nem sikerült javulást elérni. Modelljeink gyengébb teljesítményt értek el. Ennek egyik okaként az adatok túlzott mértékű augmentációját azonosítottuk, melynek eredménye, hogy a már az eredeti háttéreloszlástól eltérő adathalmazt kaptunk.

Egy következő kísérletben csak augmentáció nélküli mintákkal tanítottunk, melyet úgy értünk el, hogy az adathalmazok teszt- és validációs mintáit használtuk fel (~1000 minta). Így is lassabb konvergenciát figyeltünk meg, ám képesek voltunk utolérni, sőt, meghaladni az eredeti adathalmazon tanított modelleket.

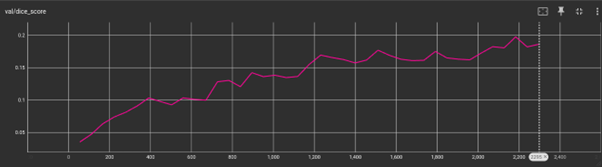
Végül az egyik új halmaz tanító mintáiból is sikeresen kiszűrtük az augmentált eseteket, szakértői módszerek segítségével, amely a fekete keret detektálásán alapult.

A tanításaik eredményei:

|  |  |  |
| --- | --- | --- |
| (validációs dice score) | FancyUNet | Segformer |
| Panoramic Dental Dataset | 0.189 | 0.11 |
| Bővített, augmentáció-szűrt | 0.197 | - |

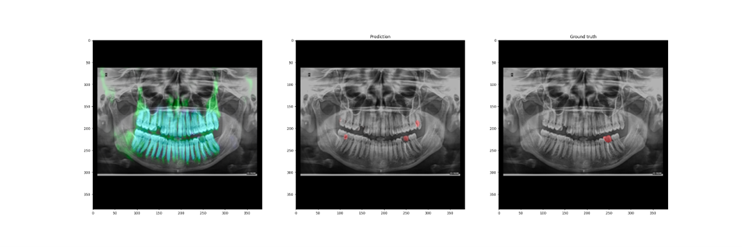


Tanulási görbék a Panoramic Dental Dataset adathalmazon (narancs: FancyUNet, zöld: Segformer)

Tanulási görbe a bővített, augmentáció-szűrt adathalmazon.

|  |  |
| --- | --- |
|  |  |

Kép és augmentált változata a szűrhető tanító adathalmazból

FancyUNet [bemenet a két előszegmentációval, kimenet, ground truth] a tanítás után

**Értékelés**

A tanító adat felhasználásával tanítottunk újra, de az eredmények elmaradtak az általunk várttól, talán a nagyobb variancia miatt, ezért úgy döntöttünk, hogy ezeket a képeket csak validációs halmazban használjuk.

Megmértük akövetkező módosítások hatásait:

1. Fekete rész levágás: Levágtuk a képek széléről a fekete keretet. Nem volt hasznos.
2. Tükrözés: Megtükröztük a felvételeket a függőleges tengely mentén, ennek motivációja, hogy az ember arca szimmetrikus, így a tükrözés és átlapolódás elvileg segíthet a modellnek. A gyakorlatban azonban nem tapasztaltunk javulást.
3. Felbontás csökkentése és batch-méret növelése: A hálózat gyorsabban tanult, de nem ért el jobb eredményt. Érdekes azonban mefigyelni, hogy az eredménye még igen alacsony képfelbontás mellett is közel azonos volt, ez is arra hívja fel a figyelmünket, hogy a háló csak a nagyobb szuvasodásokat találja meg.
4. Focal-loss használata: Dinamikus hibafüggvény, mely az osztályok kiegyenlítetlenségét hivatott egyensúlyozni. Nem sikerült megfelelően úgy a paramétereit, hogy javuljanak az eredmények.

Kvantitatív eredmények:

|  |  |
| --- | --- |
| Dice-score | 0.21 |
| Átlagos Hamming-távolság | 157 pixel (384x384 képen). A képek 10%-án nem értelmezhető, mert 50% küszöb mellett nincs is fogszuvasodás |

Azon kívül, hogy a kvantitatív eredményeink gyengeségét levontuk, kvalitatívan is elemeztük a predikciókat. A detektált régiók a fogszuvasodás potenciális zónái (pl.: fognyak, korona érintkezési rész), vagyis a modell az alakzatfelismerésen kívül az esetek többségében nem tud döntést hozni a szuvasodásról. Nem szuvasodást, hanem fogat ismer fel.

Néhány mintánál azonban egészen pontosan lokalizálja a szuvasodásokat, és epoch-ok között összehasonlítva a modelljeinket azt tapasztaljuk, hogy bizonyos szuvasodásokat alacsony varianciával több epoch-ban is detektál. Ez arra utal, hogy egy megfelelő ensemble használata, vagy akár Stochastic Weighted Averaging is alkalmazható lehet.

A modellünk összességében közvetlen használatra alkalmatlannak találjuk, mert pontatlanul javasol régiókat.

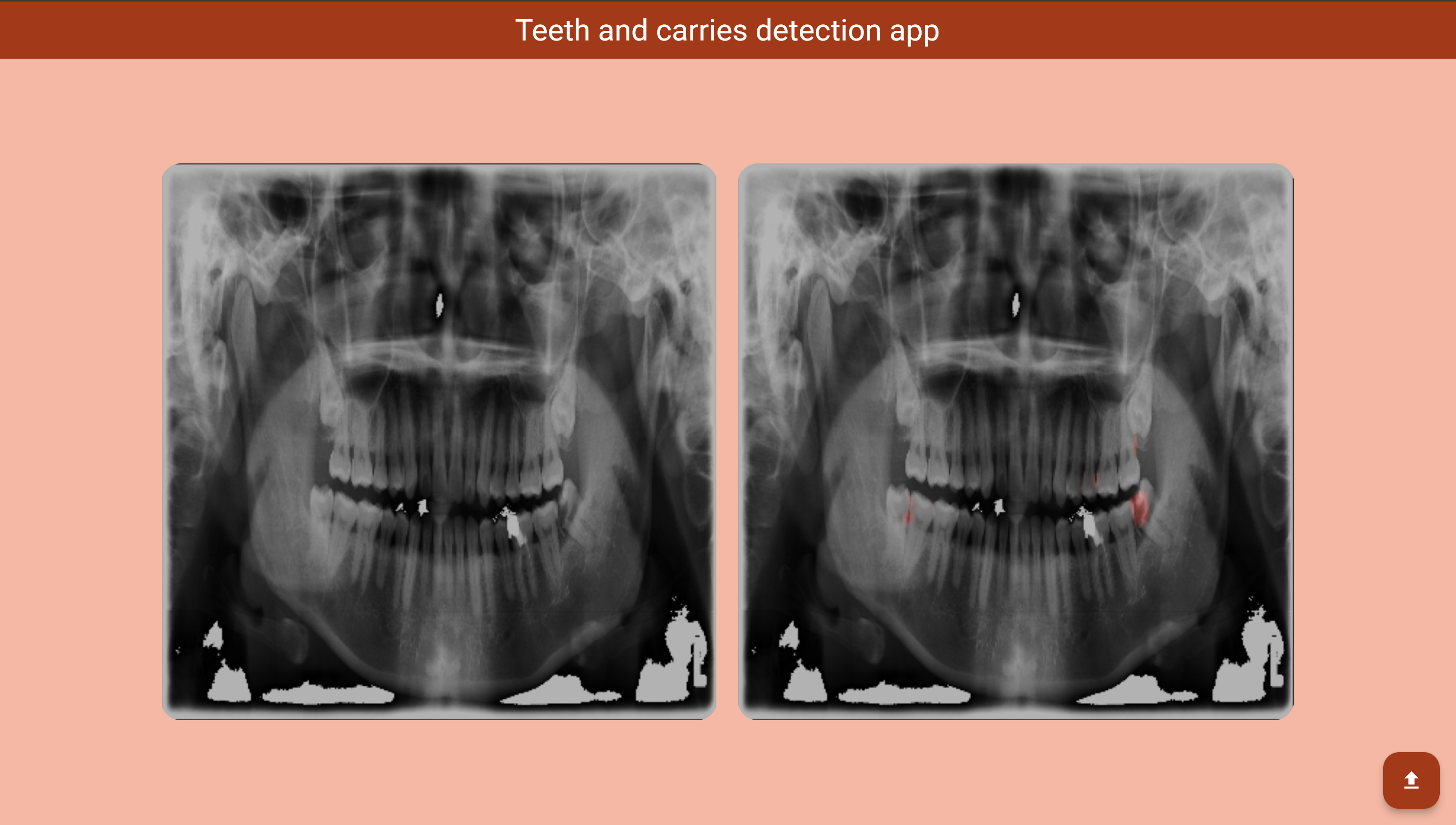
Kitekintésül feltételezzük, hogy amennyiben olyan hálót alkalmaznánk, amelyek ROI-k predikciójára alkalmasabbak lennének (pl- detection felírás), akkor használhatóbb modelleket kapnánk.

**Alkalmazás**

Az algoritmus/modell alkalmazásba ágyazásának célja a fogorvos munkájának segítése. Ez a gyakorlatban azt jelenti, hogy az orvos feltölti a páciens panoráma röntgen-felvételét és a modell predikcióját kapja vissza, mely a panoráma röntgen-felvételen színessel jelölve jelenik meg a szuvasnak minősített területeken.

Mivel konténerizált tanításunk van, kézenfekvő egy REST API-n kiszolgálni a felhasználót. Mivel a képek feltöltésére tipikusan asztali környezetben van lehetőség, egy webalkalmazással szolgáltuk ki a kéréseket, amelyeket az általunk ismert Flutter technológiával készítettünk el.

Az alkalmazásban feltölthetünk egy röntgen képet, és megtekinthetjük a fogszuvasodás szegmentációs maszkját. A háttérben a backend elkészíti a két legjobb előszegmentáló modellünkkel a fogak maszkját, és a legjobb szegmentáló modellünkkel előállítja a predikciót.

Az alkalmazás használat közben