

# Predlog diplomskog rada

## Definicija problema/ cilj projekta

Cilj projekta je detekcija melanoma na koži sa akcentom na jednakost izmedju boja kože. Potrebno je da model podjednako dobro klasificuje maligne/benigne lezije na svim bojama i tipovima kože. Ulaz u sistem je dermatoskopska slika lezije zajedno sa meta-podacima (podaci o slici i pacijentu). Izlaz iz sistema je binarna klasifikacija - da li je lezija maligna ili benigna.

## Motivacija problema rešavanog u projektu

Rana detekcija melanoma značajno povećava šanse za uspešno lečenje i smanjenje smrtnosti. Međutim, aktuelni modeli često ispoljavaju pristrasnost prema tonovima kože, naročito kod najtamnijih i najsvetlijih nijansi, usled nedovoljne zastupljenosti tih slučajeva u dostupnim skupovima podataka. Ovakva nejednakost može dovesti do nepouzdanih rezultata za deo pacijenata. Rešavanje ovog problema omogućilo bi razvoj robusnijih i pravednijih sistema za detekciju melanoma, sa većom kliničkom relevantnošću.

## Relevantna literatura

### [Literatura 1: Towards Fairness in AI for Melanoma Detection: Systemic Review and Recommendations](#)

#### 1. Zadatak

Rad je sistematski pregled literature (2013-2024) koji se fokusira na AI / deep learning metode za detekciju melanoma, sa posebnim naglaskom na zastupljenost i fer tretman različitih tonova kože. Cilj je identifikovanje pristrasnosti u trenutnim modelima, analiza kakvi su skupovi podataka (dataset-i), koje skin tone skale se koriste, i davanje preporuka kako da se unapredi pravičnost i inkluzivnost u budućim modelima.

#### 2. Metodologija

Primarna metodologija je pregled sistema prateći PRISMA smernice. Istražuju postojeće studije/eksperimente, analiziraju dataset-e, prikazuju koje metode AI su primenjene u tim radovima

#### 3. Skup podataka

Rad ne koristi jedan skup, vec analizira vise skupova podataka. Primeri dataset-a: PH2 (200 slika), Edinburgh Dermofit Library (~1300), HAM10000 (~10015), ASAN (~120,780 slika), Diverse Dermatology Images (DDI) dataset (ima slike sa FST I-II, III-IV, V-VI tonovima kože, sa benignim i malignim lezijama). Skupovi podataka se razlikuju po: broju slika, da li su uključeni demografski podaci (starost, pol, etnička pripadnost, boja kože), tip & lokacija

lezije, tipu slike, rezoluciji. Neki dataset-i imaju vrlo malo informacija o boji kože / tonovima kože; neki dataset-i nemaju uopšte

#### 4. Evaluacija rešenja

Pošto je ovo pregledni rad, evaluacija modela je indirektna — kroz upoređivanje onoga što su autori drugih radova izvestili: koje mere performansi su korišćene (accuracy, sensitivity, specificity, ROC-AUC, itd.), da li su ti rezultati raspoređeni po tonovima kože ili drugim demografskim varijablama, koliko su dataset-i uravnoteženi po tonovima kože

#### 5. Rezultati

Pronađeno je da AI metode za detekciju melanoma imaju značajnu **pristrasnost prema svetlijim tonovima kože**, jer su modeli trenirani i testirani većinom na slikama svetlih ili neodređenih tonova kože, dok tamniji tonovi ostaju nedovoljno zastupljeni.

#### 6. Šta po meni nedostaje

Bilo bi korisno dopuniti sa empirijskim studijama kojima se upoređuju performanse modela kad se uključe tamni i svetli tonovi kože.

### Literatura 2: Detecting Melanoma Fairly: Skin Tone Detection and Debiasing for Skin Lesion Classification

#### 1. Zadatak

Cilj je da se otkrije i smanji pristrasnost u klasifikaciji melanoma u zavisnosti od tona kože. Autori predlažu algoritam za automatsko određivanje tona kože slika lezija, označavaju ISIC skup podataka tim tonovima, i primenjuju tehnike „unlearning” da bi ublažili razlike u performansama između svetlih i tamnih tonova kože.

#### 2. Metodologija

Koristi se CNN (ResNetXt-101, EfficientNet), tehnike debiasing-a “Learn Not To Learn” i “Turning a Blind Eye” kao i “Turning a Blind Eye” + gradient reversal. Za procenu tona kože koristi se ITA.

#### 3. Skup podataka

- *Fitzpatrick17k* dataset sa ljudskim anotacijama Fitzpatrick tipova kože ( $\approx 16.577$  slika, od čega  $\sim 4.316$  slika neoplastičnih lezija) koristi se za trening i verifikaciju algoritma za procenu tona kože.
- *ISIC (2017 + 2020) skup* ( $\sim 35.574$  slika) koristi se za treniranje modela sa automatski označenim tonovima kože.
- Za testiranje/generalizaciju koriste se *MClass*, *Interactive Atlas of Dermoscopy* i *ASAN datasets*, koji dolaze iz različitih populacija i sadrže i dermatoskopske i kliničke slike.

#### 4. Evaluacija rešenja

Primarna metrika je AUC - Area Under ROC Curve

#### 5. Rezultati

Algoritam za automatsko etiketiranje tona kože imao je ~60.61 % tačnosti (unutar  $\pm 1$  nivoa na Fitzpatrick skali)

6. Šta po meni nedostaje

Treba vise istraziti efikasnost metoda u uslovima promenljivog osvetljenja, raznih uređaja za snimanje, i realnim slikama iz kliničke prakse (ne samo javnih datasetova).

### [Literatura 3: Data-Driven Color Augmentation Techniques for Deep Skin Image Analysis](#)

1. Zadatak

Cilj je da se poveća robusnost dubokih neuronskih mreža za analizu dermatoskopskih slika kroz tehnike augmentacije boje, tako da modeli bolje podnose varijacije u osvetljenju prilikom snimanja. Rad primenjuje ove tehnike na zadatke segmentacije i klasifikacije lezija u okviru ISIC 2017 izazova.

2. Metodologija

Autori koriste *Shades of Gray* tehniku za procenu ilumininanta slike, zatim vrše bijelo balansiranje, i tokom treninga nasumično primenjuju boju osvetljenja (sample-ovane iluminante) kako bi simulirali različite uslove osvjetljenja; uz to koriste i korekciju gama i geometrijske augmentacije.

3. Skup podataka

Koristi se ISIC 2017 dataset za obu-ku i evaluaciju (za zadatke segmentacije i klasifikacije), bez eksternih dermatoloških baza podataka.

4. Evaluacija rešenja

Modeli su trenirani i evaluirani sa standardnom podelom ISIC 2017: trening set + validacioni set + test set, mere performansi su npr. dice koeficijent, Jaccard, accuracy, AUC, specificity, sensitivity.

5. Rezultati

Za zadatak segmentacije, model je ostvario DICE koeficijent ~0.846 i ACC ~0.948 na validacionom skupu; za klasifikaciju melanoma vs ostale lezije AUC oko 0.791, i prilično dobru specifičnost/senzitivnost na seborrheic keratosis klasifikaciji.

6. Šta po meni nedostaje

Metoda ne adresira eksplisitno pravičnost između tonova kože.

Dodatne literature:

- [Literatura 4: FairDisCo: Fairer AI in Dermatology via Disentanglement Contrastive Learning](#)
- [Literatura 5: A patient-centric dataset of images and metadata for identifying melanomas using clinical context](#)
- [Literatura 6: Fairness Metrics in AI: Your Step-by-Step Guide to Equitable Systems](#)

## Skup podataka

Skup podaka za izradu diplomskog rada bio bi [ISIC 2020](#), ali zbog njegove nebalansiranosti dodali bi se skupovi ISIC (2017-2020).

- Referenca na rad u kome je skup konstruisan: [A patient-centric dataset of images and metadata for identifying melanomas using clinical context](#)
- [Link](#) sa kog se skup može preuzeti
- Ciljna oznaka: binarna klasifikacija – *malignan (melanoma) vs benign*.
- Raspodela klase: dataset je veoma neuravnotežen: od 33,126 slika za trening, samo oko 1.8 % su *maligne* slike, dok je ~98.2 % benignih.

Za svaku instancu/sliku dostupni su sledeći atributi:

- Slika (dermatoskopska / dermoscopic image) – glavni input.
- Metadata / metapodaci:
  1. **patient\_id** – identifikator pacijenta.
  2. **lesion\_id** – identifikator lezije (u nekim verzijama metapodataka).
  3. **sex** – pol pacijenta.
  4. **age** – starost pacijenta.
  5. **general\_anatomic\_site** – opšte područje tela gde se lezija nalazi (npr. noga, leđa, ruka itd.).
- U nekim slučajevima postoji „detailed diagnosis“ metapodatak (više klasa dijagnoze) u treningu, ali glavni target je binarna malignost vs benignitet.

## Metodologija

Za izradu diplomske zadatke bi se podelio u 2 faze.

Faza pripreme podataka - Monk Skin Tone Scale bi se koristila kao referenca za labeliranje slika po tonovima kože. Koristila bi se Augmentacija podataka da se postigne raznovrsnost osvetljenja, oblika i tonova.

Faza klasifikacije - Koristila bi se CNN mreža, konkretno EfficientNet koja prima sliku i meta podatke, koristi funkciju gubitka da penalizuje klasifikaciju po tonovima kože u zavisnosti od zastupljenosti određenog tona kao i tačnosti u klasifikovanju

## Metod evaluacije

**Accuracy** – koliko je tačno klasifikovanih uzoraka (ali kod neuravnoteženih klasa nije pouzdana).

**Precision, Recall, F1-score** – korisno kada jedna klasa (npr. melanoma) ima mnogo manje uzoraka.

**AUC-ROC (Area Under ROC Curve)** – standard u medicinskoj dijagnostici jer meri sposobnost modela da razlikuje pozitivne i negativne slučajeve nezavisno od praga.

**Balanced accuracy / sensitivity / specificity** – za neuravnotežene datasete.

**Fairness metričke** (ako je fokus na pravičnosti) – npr. razlike u AUC između grupa (po tipu kože, polu, starosti).

**Napomena:**

Projekat je već urađen, rađen je u timu od 2 člana: Katarina Krstin SV57/2021 i Jovan Vučković SV64/2021. Otvoreni smo za proširenja i poboljšanja projekta ukoliko je potrebno. Na predmetu PIGKUT zajedno pišemo na temu ovog projekta.