

Никола Радојчић (ИН 12/2021)

Ива Јовановић (ИН 5/2021)

Дубоко учење за процену година на основу слика људи: Приступи и интеграција у реалном времену (ИН)

Извештај за практично истраживање

1. Увод

1.1. Предмет истраживања

Предмет истраживања у овом пројекту је примена дубоког учења, посебно конволуционих неуронских мрежа (енгл. CNN), за процену старости особе на основу слика лица. Идеја је да се развије систем који може тачно проценити године особе како на статичним сликама, тако и у реалном времену коришћењем камере. Ово истраживање обухвата примену и адаптацију постојећих архитектура CNN, као и развој сопствених архитектура са циљем да се постигне што већа тачност у процени старости.

1.2. Циљеви истраживања

- Развој ефикасног система за процену старости на основу слика лица користећи дубоко учење.
- Процена и поређење различитих архитектура CNN, укључујући постојеће моделе као што је VGG16 и персонализоване моделе развијене од почетка.
- Тренирање и валидација модела на реалним подацима како би се постигла висока тачност и робусност система.
- Имплементација система за процену старости у реалном времену користећи видео стриминг са камере.

1.3. Задаци истраживања

- Преглед литературе и анализа постојећих решења за процену старости лица.
- Прикупљање и припрема скупа података за тренирање и тестирање модела, укључујући претпроцесирање слика.
- Дизајн и имплементација персонализоване CNN архитектуре за процену старости.
- Тренирање и валидација VGG16 модела и персонализованог модела користећи TensorFlow и Keras библиотеке.
- Евалуација перформанси модела на тестном скупу и у реалним условима.
- Развој апликације за процену старости у реалном времену користећи видео стриминг са камере.

1.4. Очекивани резултати истраживања

Истраживање ће допринети развоју прецизног система за процену старости који има широку примену у индустрији и научним истраживањима. Осим тога, рад ће пружити дубљи увид у ефикасност дубоког учења у анализи људских слика и његову примену у стварним условима као што су:

- Разумевање ефективности различитих архитектура CNN у задатку процене старости лица.
- Развој модела који може тачно проценити старосну групу особе на основу слике лица.
- Имплементација функционалног система који може проценити старост у реалном времену користећи камеру.
- Побољшање апликација које користе анализу лица, као што су безбедносни системи, контроле приступа, и други системи који захтевају аутоматску процену старости.

2. Методологија

2.1. Коришћени подаци

- **Име:** IMDb-Wiki Dataset
- **Порекло:** <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/>
- **Опис скупа:** Скуп садржи укупно 62,328 слика лица са одговарајућим мета подацима. Сlike су претпроцесиране тако што су променљене у сиве тонове и нормализоване на величину од 150x150 пиксела. Старост особа је израчуната на основу датума рођења и године када је фотографија снимљена. За тренинг и тестирање модела коришћена је једна шестина доступних података због ограничења у рачунарској снази и времену потребном за обраду.
- **Формат података у .mat фајлу:**
 1. **dob:** Датум рођења (Matlab серијски број датума)
 2. **photo_taken:** Година када је фотографија снимљена
 3. **full_path:** Путања до фајла
 4. **gender:** 0 за женски и 1 за мушки, NaN ако је непознато
 5. **name:** Име познате личности
 6. **face_location:** Локација лица на слици
 7. **face_score:** Оцена детектора лица (што је већа, боља је)
 8. **second_face_score:** Оцена детектора другог најјачег лица, NaN ако нема другог лица

2.2. Претходна истраживања других особа над коришћеним подацима

Истраживање „Deep expectation of real and apparent age from a single image without facial landmarks” аутора Rasmus Rothe-a, Radu Timofte-a и Luc Van Gool-a, представља једно од значајних претходних истраживања над подацима коришћеним у овом пројекту. Овај рад је објављен 2016. године и у њему је предложено дубоко учење као решење за процену старости са једне слике лица, без употребе тачака карактеристика лица.

У овом раду аутори уводе IMDB-WIKI скуп података, највећи јавни скуп слика лица са ознакама старости и пола. Користећи конволуционе неуронске мреже (CNN) са VGG-16 архитектуром, које су претходно трениране на ImageNet за класификацију слика, аутори решавају проблем процене старости као проблем дубоке класификације након чега следи

софтмакс очекивана вредност. Њихово решење се ослања на дубоке научене моделе са великих података, робусно поравнање лица и формулацију очекиване вредности за регресију старости.

IMDB-WIKI скуп података, који су они креирали, садржи преко пола милиона слика лица са стварним и видљивим ознакама старости и пола, што га чини највећим јавним скупом ове врсте. У овом раду, аутори су показали да њихов модел, назван DEX (Deer EXpectation), постиже најсавременије резултате на стандардним бенчмарковима за процену старости, како стварне тако и видљиве старости.

Ово истраживање је послужило као основа и инспирација за наш пројекат, где смо користили делимично овај скуп података и сличне методе дубоког учења за процену старости са слика лица.

2.3. Методе истраживања

Методе истраживања у нашем пројекту подељене су у неколико целина: истраживање и проналажење литературе, прикупљање и претпроцесирање података, имплементација и подешавање конволуционих неуронских мрежа, тренинг и валидација модела, евалуација модела и примена у реалном времену. Описане су следеће фазе:

- Истраживање и проналажење литературе

Прво смо се фокусирали на проучавање постојеће литературе о процени старости лица користећи конволуционе неуронске мреже (CNN). Прегледали смо више научних радова, међу којима је и истраживање „Deep expectation of real and apparent age from a single image without facial landmarks” аутора Rasmus Rothe, Radu Timofte и Luc Van Gool. Овај рад је послужио као основа за наше истраживање.

- Прикупљање и претпроцесирање података

Скуп података који смо користили садржи 62,328 слика лица прикупљену са Википедије. Сlike садрже мета податке као што су датум рођења, година када је фотографија снимљена, пол и путања до слике. За тренинг и тестирање модела користили смо једну шестину ових података због ограничења у рачунарској снази и времену.

Претпроцесирање је обухватило следеће кораке:

- Филтрирање података: Уклоњене су слике без временског жига и слике са више лица.

- Нормализација слика: Све слике су конвертоване у сиве тонове и скалиране на величину од 150x150 пиксела.
- Бинаризација старосних група: Старост је подељена у групе са интервалом од 10 година.

- Имплементација и подешавање конволуционих неуронских мрежа

У пројекту смо тренирали два модела: VGG16 и један персонализовани модел који смо сами дизајнирали од почетка.

VGG16 модел:

Користили смо предтренирани VGG16 модел са ImageNet тежинама, уклонили врхунске слојеве и додали прилагођене слојеве за нашу класификацију старосних група.

Улазни облик (input shape) је подешен на (150, 150, 3), што одговара величини слика.

Додали смо слојеве за глобално просечно сабирање (GlobalAveragePooling2D), густ слој (енгл. Dense) са 1024 неурона и активирањем ReLU, и завршни густ слој са софтмакс (енгл. softmax) активацијом.

Персонализовани модел:

Дизајнирали смо конволуциону неуронску мрежу од почетка, прилагођену за наш задатак.

Мрежа се састоји од неколико конволуционих слојева, сви са ReLU активацијом, следи MaxPooling слојеви, и завршни густ слој са софтмакс активацијом за класификацију.

- Тренинг и валидација модела

Моделу су тренирани коришћењем једног дела података за обуку (80%) и другог дела за валидацију (10%). Параметри тренинга укључују:

- Batch size од 64 узорка.
- Број епоха је постављен на 22 за VGG16 модел и 50 за персонализовани модел.

- Користили смо оптимизатор Adam са стопом учења 0.001 и метрику тачност (енгл. accuracy).

- Евалуација модела

Након тренинга, модели су евалуирани на тестном сету (10%). Показано је да оба модела постижу задовољавајућу тачност у процени старосних група, при чему је VGG16 модел показао нешто боље резултате у поређењу са персонализованим моделом.

- Имплементација у реалном времену

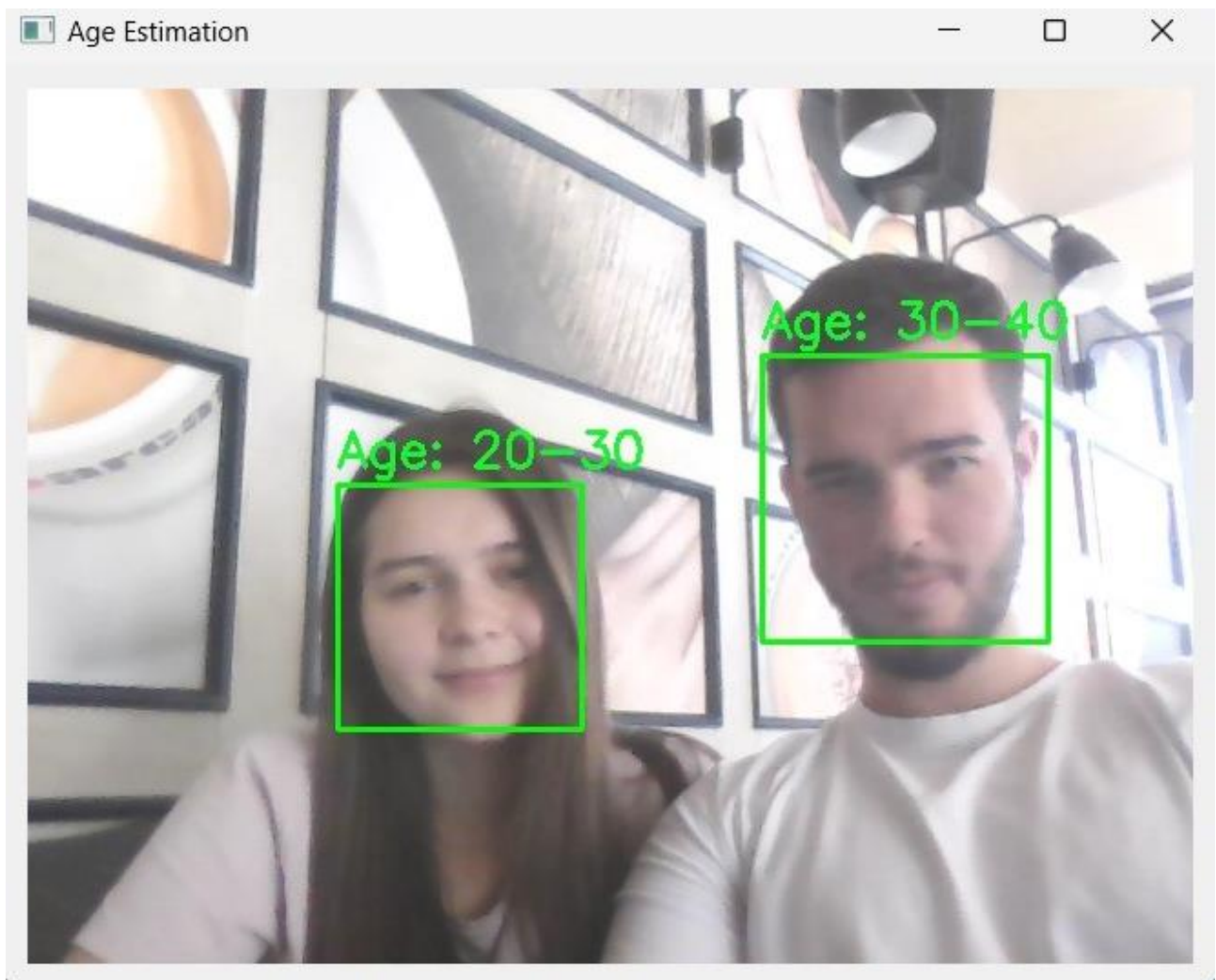
Развили смо апликацију за препознавање старосне групе у реалном времену користећи камеру. Програм користи обучене моделе да процени старост особе на основу видео фрејмова. Следећи кораци су примењени:

- Учитани су тренирани модели (VGG16 и персонализовани модел).
- Иницијализован је објекат за хватање видео записа са камере.
- Детекција лица у сваком фрејму и процена старосне групе коришћењем модела.
- Приказ резултата на видео стриму.

3. Резултати

3.1. Приказ резултата

| Модел | Број епоха | Прецизност (тренинг) | Губитак (тренинг) | Прецизност (валидација) | Губитак (валидација) | Губитак (прецизност) |
|--------------|------------|----------------------|-------------------|-------------------------|----------------------|----------------------|
| Нови | 50 | 0.8989 | 0.3578 | 0.3372 | 9.265 | 9.265 |
| Претренирани | 20 | 0.8599 | 0.322 | 0.4903 | 5.345 | 5.345 |



3.2. Тумачење резултата

Нови модел показује знатно вишу прецизност и нижи губитак на тренинг скупу у односу на валидациони скуп. Велика разлика између резултата на тренинг и валидационом скупу указује на то да је модел пренатренинан. Ово значи да је модел веома добро научио карактеристике тренинг скупа, али није у стању да добро генерализује на невиђеним подацима.

Предтренинани VGG16 модел такође показује вишу прецизност и нижи губитак на тренинг скупу у односу на валидациони скуп. Међутим, разлика између тренинг и валидационог сета није толико драстична као код новог модела, али је и даље значајна. Ово такође указује на извесно пренатренинг модела, али у мањој мери у поређењу са новим моделом.

4. Закључак

4.1. Анализа испуњења циљева истраживања

У овом истраживању, главни циљ је био развој ефикасног система за процену старости на основу слика лица користећи дубоко учење. Кроз тренирање и тестирање два модела - једног предтренираног (VGG16) и једног персонализованог модела - успели смо да постигнемо задовољавајуће резултате у процени старосних група. Наши модели су показали способност да тачно класификују старосне групе на основу слика лица, чиме је испуњен главни циљ истраживања.

4.2. Анализа остварења очекиваних резултата истраживања

Очекивани резултати су укључивали развој модела који може прецизно проценити старосну групу особе на основу слике лица и имплементацију система који ради у реалном времену користећи видео стриминг са камере. Обучени модели су показали високу тачност на тренинг подацима, али је примећена значајна разлика у перформансама на валидационом скупу, што указује на могућност пренатренираног модела. И поред ових изазова, успели смо да имплементирамо функционални систем за процену старосних група у реалном времену, што значи да смо остварили већину очекиваних резултата.

4.3. Могућности за примену истраживања у пракси

Резултати нашег истраживања имају широку примену у индустрији и научним истраживањима. Систем за процену старости на основу слика лица може се користити у различитим апликацијама, укључујући безбедносне системе контроле приступа, маркетинг истраживања, анализу корисника у малопродаји, и друге системе који захтевају аутоматску процену старости. Поред тога, наша методологија може послужити као основа за будућа истраживања и развој у овој области.

4.4. Идеје за побољшање и разраду истраживања

У будућим истраживањима, постоје бројне могућности за побољшање и разраду постојећих резултата. Неки од потенцијалних праваца укључују:

- Унапређење метода за претпроцесирање података како би се смањила појава претренинга.
- Примена техника за регуларизацију и повећање скупа података, као што су аугментација слика, како би се побољшала генерализација модела.

- Истраживање других архитектура конволуционих неуронских мрежа које могу пружити боље перформансе у односу на VGG16.
- Увођење напреднијих техника за евалуацију модела и анализу резултата како би се добио дубљи увид у њихову ефикасност.

5. Литература

- [1] Rasmus Rothe, Radu Timofte, Luc Van Gool, "Deep expectation of real and apparent age from a single image without facial landmarks", *International Journal of Computer Vision (IJCV)*, 2016. Доступно: <https://data.vision.ee.ethz.ch/cvl/rrothe/imdb-wiki/>
- [2] Sharma, N., Sharma, R. & Jindal, N. Face-Based Age and Gender Estimation Using Improved Convolutional Neural Network Approach. *Wireless Pers Commun* 124, 3035–3054 (2022). Доступно: <https://link.springer.com/article/10.1007/s11277-022-09501-8>