

ДЕТЕКЦИЈА НА ЛАЖНИ ПОТПИСИ СО ПОМОШ НА НАДГЛЕДУВАНО УЧЕЊЕ И НАОЃАЊЕ НА СЛИЧНОСТА НА ПОТПИСИТЕ

- семинарска работа по предметот Дигитално процесирање на слика -

Филип Самарџиски и Тамара Стојанова

Содржина

01. Вовед

02. Структура и тек на апликацијата

03. Надгледувано учење

- Препроцесирање
- Модели
- Точности на различните модели
- Заклучок

04. Споредба на сличност

01. Начини за наоѓање на сличност

- Наоѓање на сличност преку контури
- Наоѓање на сличност со SSIM
- Сличност меѓу два потписа
- Сличност меѓу еден и повеќе потписи

05. ДЕМО

06. Проблеми

01. Вовед

→ Дигитално процесирање
на слики

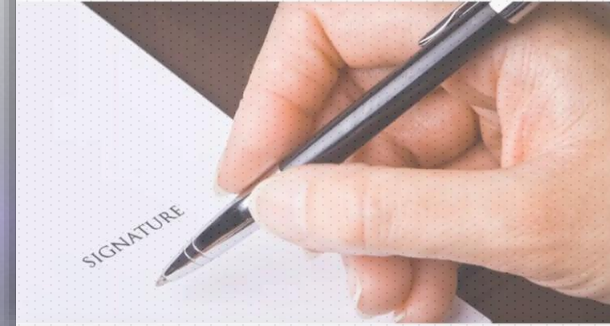
→ Злоупотреба на потписи

- Фалсификување
- Цели
- Детекција

This article is more than 2 years old

MPs pursue claims bank signatures were faked on UK court papers

Allegations over documents used for repossessions and to recover other debts



02. Структура и тек на апликацијата

структура



main.py



preprocessor.py



supervised_learning.py



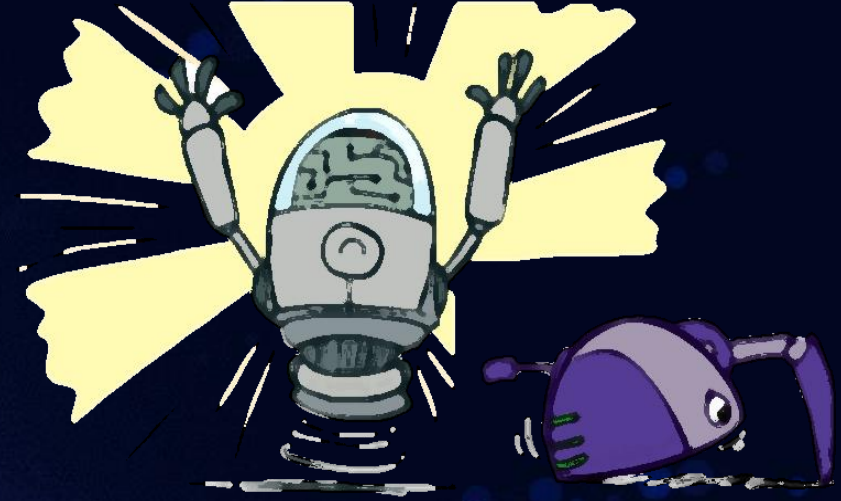
similarities.py

тек

- ☐ Дали потписот е веродостоен?
- ☐ Колку овие два потписи се слични?
- ☐ Колку е овој потпис сличен со останатите потписи?



Лажен?
Вистински?



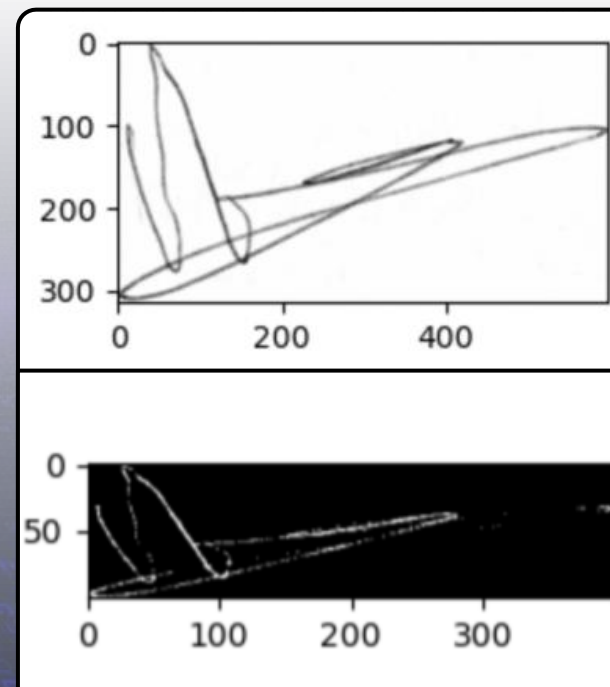
03. Надгледувано учење





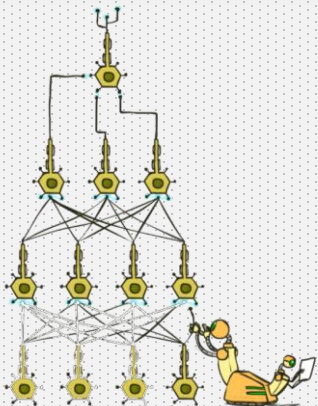
Препроцесиране на податоците

- Вчитување на сликите – `imread()`
- Намалување на шум – `fastNIMeansDenoising()`
- Thresholding со инверзна бинаризација
- Пронаоѓање на ненулеви пиксели и преградување околу нив
- Промена на големина и претворање во еднодимензионална низа – `flatten()`
- Збир на колони и редици и сооднос на сликите

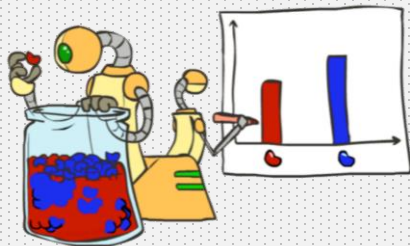


Слика - потпис пред и по фазата на препроцесирање

Модели при надгледувано учење



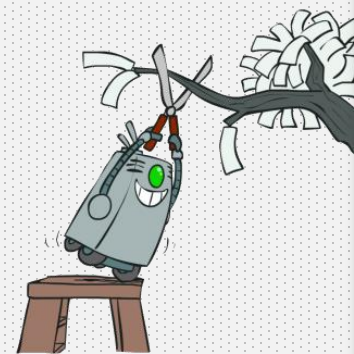
Невронски
мрежи



Наивен
Баесов



Дрва
на одлука



Шума од
дрва

Точности на различните модели

* се споредува точноста од тестирачкото множество, затоа што врз тренирачкото множество сите модели имаат точност од 100%.

Модел	Точност над тестирачко множество
Невронска мрежа	0,7777777778
Наивен Баесов класификатор	0,80555555555513
Дрво на одлучување	0,90277777778
Шума од дрва на одлучување	0,972222222222467

Точност над тестирачкото множество според модели и автори			
Модел	Avtor1	Avtor2	Avtor3
Невронска мрежа	0.833333333333334	0.875	0.625
Наивен Баесов класификатор	0.9166666666666666	0.583333333333334	0.9166666666666666
Дрво на одлучување	0.875	1.0	0.833333333333334
Шума од дрва на одлучување	0.958333333333334	1.0	0.958333333333334



Заклучок

2. Невронската мрежа дава најниска точност во најголем дел од случаите.

3. Шумата од дрва на одлука дава најстабилни резултати и доста висока точност.

1. Разновидни и неочекувани резултати

ЗОШТО?

✓ недоволно препроцесирање на сликите

✓ неразновидност на податоците во тренирачкото множество

✓ *overfitting* на тренирачкото множество

✓ неоптималност на поставените параметри

✓ случајната распределба на податоци и нивни атрибути кај шумата од дрва

✓ принцип на независност на карактеристиките на податоците кај наивен баесов

04. Споредба на сличност

Начини за наоѓање на сличноста на потписите

КОНТУРИ

Пронаоѓање на контурите на соодветните слики и нивна споредба со помош на готовата функција `cv2.matchShapes`.

SSIM

Подобри резултати

Функција за споредба на слики која го оценува структурниот и текстурниот контекст на сликата, вклучувајќи ги вредностите за осветленост и контраст.

Резултати за сличноста помеѓу:

ДВА ПОТПИСИ

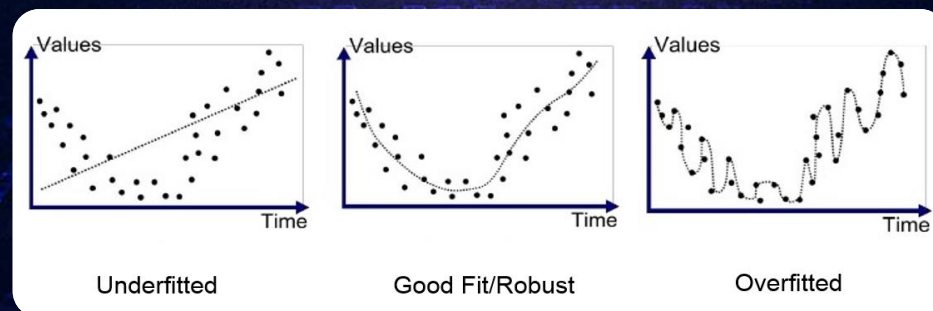
Прилично неконзистентни резултати за оваа функционалност.

ЕДЕН И ПОВЕЌЕ ПОТПИСИ

Не е целосно точен и стабилен, меѓутоа сепак дава одредена одделеност на вистинските од лажните потписи со одредени исклучоци.

05. ДЕМО

06. Проблеми



**Ви благодариме на
вниманието.**

Дали имате некои прашања?