

راهنما

استخراج و دیجیتالی کردن سیگنال‌های ECG از عکس‌های اسکن شده

پارسا حسینی

۱۴۰۰/۹/۲

نصب و راه اندازی:

برنامه از [این لینک](#) قابل دانلود است. توضیحات نصب، راه اندازی و اجرا در همان صفحه گیتهاب موجود است.

راهنما:

منبع اصلی کد [این پروژه](#) است. در زیر فایل های مهمی که به پروژه اضافه کردم توضیح دادم.

config.ini

paperECG\src\main\python\Main.py

paperECG\src\main\python\controllers\MainController.py

paperECG\src\main\python\metadata.py

paperECG\src\main\python\headerOCR.py

• فایل config.ini:

قبل از هر کاری باید ۴ متغیر این فایل را ست کنید.

RECORDS_PATH: فولدري که عکس های اسکن ECG قرار دارند. نباید فایل دیگری در آن نباشد.

حتما به فرمت آدرس قبلی توجه کنید.

OUT_PATH: فولدر خروجی. برای هر عکس یک فولدر به اسم آن درست خواهد شد و خروجی مربوط

درون آن ذخیره می شود. باز هم حتما به فرمت آدرس نمونه توجه کنید.

METADATA_PATH: یک فایل جیسون نمونه که برای همه عکس ها استفاده شود. این فایل نشان

می دهد در هر عکس هر lead کجا قرار گرفته است. برای ساختن این فایل ها راحت ترین راه استفاده از

[این نرم افزار](#) است. (نسخه باینری برای ویندوز را دانلود کنید.) یک نمونه از این فایل ها در پوشه

metadata قرار داده شده است.

tesseract_path: بعد از اینکه tesseract و pytesseract را نصب کردید، باید در

system variables هم تغییرات مناسب ایجاد کنید. برای اینکه نیازی به این تغییرات نباشد،

کافی است جایی که tesseract.exe را نصب کردید را در اینجا قرار دهید.

- فایل `Main.py`:

با اجرای این فایل برنامه شروع می‌شود. برای اجرا از دستور زیر استفاده کنید:

```
py -3.6 main.py
```

خروجی نهایی برنامه برای هر عکس، یک فایل `CSV` بوده که هر ستون آن خروجی دیجیتالیز شده یک `lead` است.

- فایل `MainController.py`:

تغییرات اصلی برای استفاده از کد آماده در این قسمت داده شده است. 3 تابع `setImageFile` و `digitization` و `closeImageFile` کاملاً جدید هستند.

- فایل `metadata.py`:

این فایل برای تولید متادیتا مربوط به هر عکس است. به فرمت جیسون آنها اگر دقت کنید می‌بینید که در آن هم اسم عکس و هم دایرکتوری آن وجود دارد. این کد این موارد را تصحیح می‌کند.

- فایل `headerOCR.py`:

قسمت اصلی برای خواندن اطلاعات بالای عکس‌ها است. اطلاعات مهم با توجه به ریجکس استخراج شده‌اند. در کل برای هر عکس 2 فایل تکست خروجی داده می‌شود. اولی `raw.txt` که تمام متنی که خوانده شده به صورت به هم ریخته در آن است. دومی `info.txt` که در آن اطلاعات مهم تمیز ذخیره شده‌اند.