

Medical Text Extractor er et verktøy utviklet for å hjelpe helsepersonell med å raskt hente ut administrasjonsinstruksjoner for medisiner fra bilder. Applikasjonen overvåker utklippstavlen for bilder, bruker OCR (Optical Character Recognition) for å trekke ut tekst, identifiserer medisinnavn og viser administrasjonsinstruksjoner automatisk.

Før du installerer Medical Text Extractor, må du ha følgende programvare installert:

Merk: Du vil bli bedt om å oppgi passord for å klonere repositoret.

Opprett Conda-miljø:

```
powershell
```

```
powershell -ExecutionPolicy Bypass -File setup_conda_env.ps1
```

Dette vil opprette et Conda-miljø med navnet med_text_env i d:\Conda.env\med_text_env.

Aktiver Conda-miljøet:

```
powershell
```

```
.\activate_env.bat
```

Verifiser installasjonen:

```
powershell
```

```
python -c "import pytesseract; print('Tesseract er tilgjengelig')"
```

Hvis du ikke ser noen feilmeldinger, er installasjonen vellykket.

For å starte Medical Text Extractor med systemstatusikon:

Aktiver Conda-miljøet:

```
powershell
```

```
.\activate_env.bat
```

Start den forbedrede utklippstavleovervåkingen:

```
powershell
```

python start_enhanced_clipboard_service.py

Et systemstatusikon vil vises i systemstatusfeltet. Høyreklikk på ikonet for å se alternativer:

Trykk tasten "Prt Scrn" (Print Screen) for å ta bilde av skjermen og aktivere programmet. Dersom programmet oppdager legemiddel i bildet av Alfa-skjermen, vil det kopiere bruksanvisningen (administreringen). Den kan kopieres ut ved å trykke Ctrl+V eller høyre knapp på musen og velge "Lim inn" i Alfa.

Alternativt kan du kopiere et bilde som inneholder medisinsk informasjon til utklippstavlen (f.eks. ved å ta et skjermbilde eller kopiere et bilde fra en annen applikasjon)

Applikasjonen vil automatisk oppdage bildet, trekke ut tekst og analysere den

Hvis medisinformasjon blir funnet, vil en dialogboks vise administrasjonsinstruksjonene

Administrasjonsinstruksjonene blir også kopiert til utklippstavlen for enkel liming inn i andre applikasjoner

Applikasjonen overvåker kontinuerlig utklippstavlen for nye bilder. Når et bilde oppdages, blir det behandlet for å trekke ut medisinsk informasjon.

Optical Character Recognition (OCR) brukes for å trekke ut tekst fra bilder. Applikasjonen bruker Tesseract OCR, som er en kraftig OCR-motor som støtter flere språk, inkludert norsk.

Den ekstraherte teksten analyseres for å identifisere viktig medisinsk informasjon, inkludert:

- HPR-nummer (Helsepersonellregisternummer)
- Rekvirent (lege eller annet helsepersonell)
- Legemiddelnavn

Applikasjonen lagrer og henter administrasjonsinstruksjoner for medisiner fra en lokal database. Dette gjør det mulig å raskt hente instruksjoner for kjente medisiner.

Et systemstatusikon gir enkel tilgang til applikasjonens funksjoner, inkludert:

- Starte/stoppe overvåking av utklippstavlen
- Avslutte applikasjonen

Applikasjonen viser dialogbokser med administrasjonsinstruksjoner når medisiner blir identifisert.

Dette gir umiddelbar tilgang til viktig informasjon.

Følgende Python-pakker er nødvendige:

- opencv-python - for bildebehandling
- pillow - for bildemanipulering
- pyautogui - for skjermbilledeteksjon og automatisering
- pynput - for tastatur-/museovervåking
- pyperclip - for utklippstavleoperasjoner
- pytesseract - Python-wrapper for Tesseract OCR
- win10toast - for Windows-varsler
- pystray - for systemstatusikon
- pywin32 - for Windows API-tilgang
- psutil - for prosessovervåking

Velg "Start overvåking"

Stopp overvåking:

Hvis du ønsker å behandle et bilde manuelt:

For å legge til en ny medisin i databasen:

Aktiver Conda-miljøet:

```
powershell
```

```
.\activate_env.bat
```

Kjør skriptet for å legge til medisiner:

```
powershell
```

```
python add_medication_improved.py
```

Følg instruksjonene for å legge til medisinformasjon:

For å vise alle medisiner i databasen:

Aktiver Conda-miljøet:

```
powershell
```

```
.\activate_env.bat
```

Kjør skriptet for å vise medisiner:

```
powershell
```

```
python view_medications.py
```

Problem: Applikasjonen starter ikke når du kjører start_enhanced_clipboard_service.py.

Løsning:

1. Sjekk at Conda-miljøet er aktivert:

```
powershell
```

```
.\activate_env.bat
```

2. Sjekk at alle nødvendige pakker er installert:

```
powershell
```

```
pip list
```

3. Sjekk loggfilen service_starter.log for feilmeldinger:

```
powershell
```

```
Get-Content -Path service_starter.log -Tail 20
```

Problem: Applikasjonen oppdager bilder, men trekker ikke ut tekst korrekt.

Løsning:

1. Sjekk at Tesseract OCR er installert og konfigurert korrekt
2. Sjekk at bildet har god kvalitet og er lesbart
3. Prøv å behandle bildet manuelt med debug-skriptet:

```
powershell
```

```
python debug_clipboard_monitor.py
```

4. Sjekk loggfilen clipboard_monitor.log for feilmeldinger:

```
powershell
```

```
Get-Content -Path clipboard_monitor.log -Tail 20
```

Problem: Applikasjonen trekker ut tekst, men finner ikke medisininformasjon.

Løsning:

1. Sjekk at teksten inneholder medisinnavn i et format som applikasjonen kan gjenkjenne
2. Sjekk at medisinen finnes i databasen:

```
powershell
```

```
python lookup_medication.py [medisinnavn]
```

3. Prøv å legge til medisinen manuelt:

```
powershell
```

```
python add_medication_improved.py
```

Applikasjonen genererer flere loggfiler som kan være nyttige for feilsøking:

For spørsmål, tilbakemeldinger eller støtte, vennligst kontakt:

- Automatisk overvåking av utklippstavlen: Oppdager automatisk når bilder kopieres til utklippstavlen
- OCR-behandling: Trekker ut tekst fra bilder ved hjelp av Tesseract OCR
- Tekstanalyse: Identifiserer norske medisinske termer som "rekvirent" og "Legemiddel"

- Databaseintegrasjon: Lagrer og henter administrasjonsinstruksjoner for medisiner
- Systemstatusikon: Gir enkel tilgang til å starte/stoppe overvåkingen og avslutte applikasjonen
- Brukernotifikasjoner: Viser dialogbokser med administrasjonsinstruksjoner når medisiner blir identifisert
- Prosessor: Minimum 2 GHz dual-core
- Minne: Minimum 4 GB RAM
- Lagring: Minimum 500 MB ledig diskplass
- Operativsystem: Windows 10 eller nyere
- Python: Version 3.9
- Conda: For håndtering av virtuelle miljøer
- Tesseract OCR: For OCR-funksjonalitet
- clipboard_monitor.log: Loggfil for utklippstavleovervåkingen
- service_starter.log: Loggfil for oppstartsprosessen
- extracted_text.txt: Inneholder tekst som er trukket ut fra det siste behandlede bildet
- clipboard_results.txt: Inneholder resultater fra den siste behandlingen
- GitHub: github.com/tskaret/MedicalTextExtractor