

비식별화, 정량화 파이프라인 실습

정의석

AR3 팀
esjeong@neurophet.com

Overcome brain disorders
with advanced neuroscience.

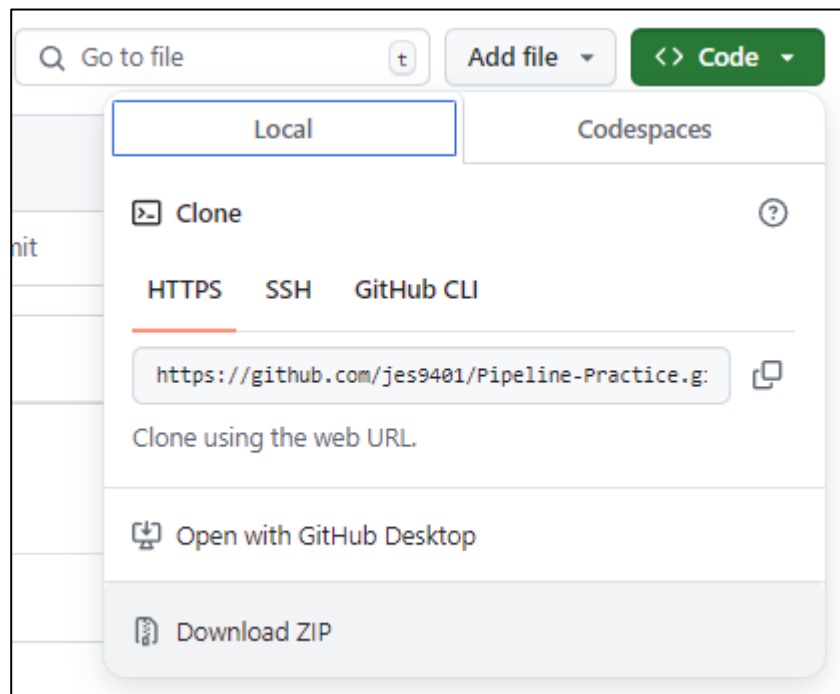
목차 CONTENTS

- 1 | 비식별화
- 2 | 정량화

준비

실습 파일 다운로드

- github.com/jes9401/Pipeline-Practice

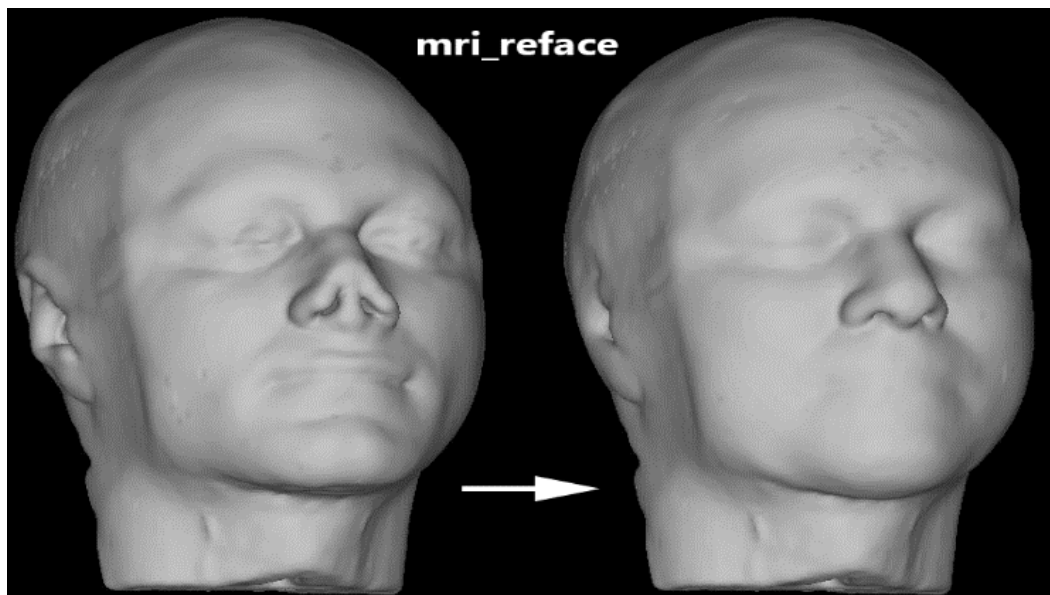


Chapter 1

비식별화

- 1 개요
- 2 구축 방법
- 3 실행 방법
- 4 결과 확인

개요

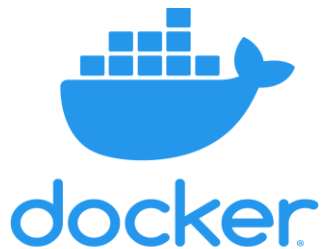


* 출처: https://www.nitrc.org/projects/mri_reface

Reface

- 얼굴 정보를 대체하여 식별을 방지하는 데 도움을 주는 도구
- NITRC(Neuroimaging Tools & Resources Collaboratory)의 오픈 소스 소프트웨어 사용
- 분석에 미치는 영향을 줄이기 위해 얼굴을 제거하지 않고 대체하는 익명화 방식
- 메타 데이터에서 환자를 식별할 수 있는 정보들도 함께 제거
- T1, T2, FLAIR, Amyloid PET, tau PET, FDG PET, CT 지원

구축 방법

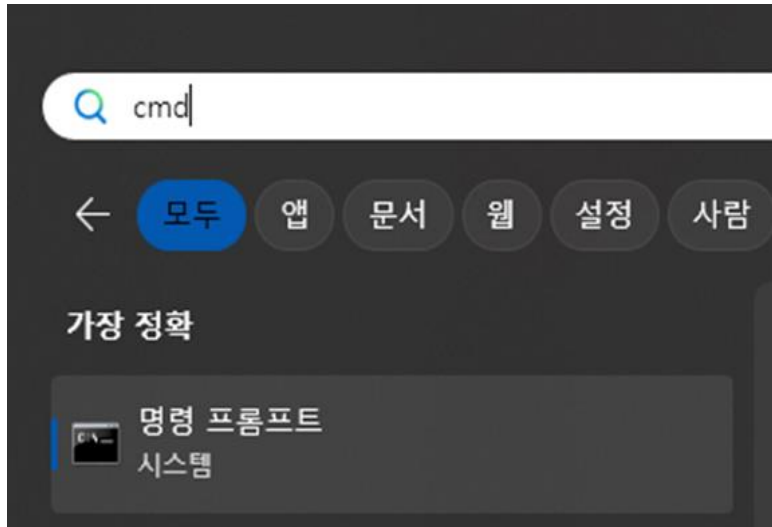


Windows 환경 설정

- 사용하는 소프트웨어들이 리눅스 기반이기 때문에 Windows 환경에서는 바로 사용할 수가 없음
- Windows의 WSL을 이용해 Ubuntu를 설치
 - WSL (Windows Subsystem for Linux)
- Docker Desktop 설치
 - 운영 체제와 독립적으로 작동할 수 있도록 사용
- Docker 이미지 설치

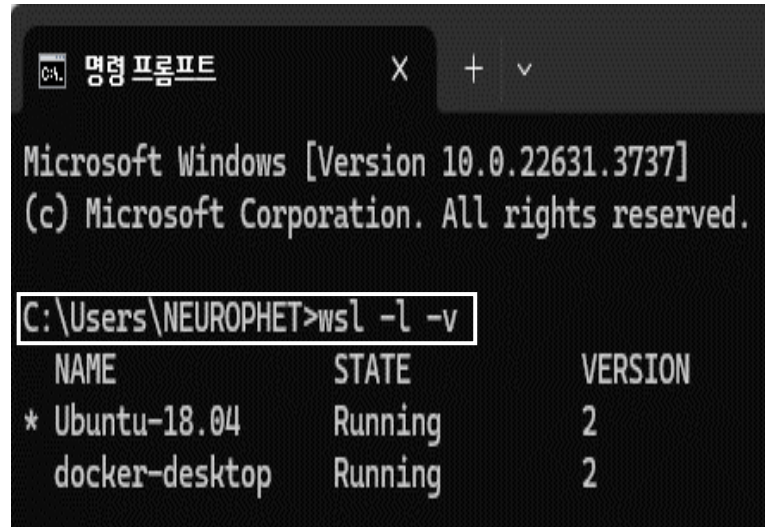
실행 방법

WSL 및 Docker 상태 확인



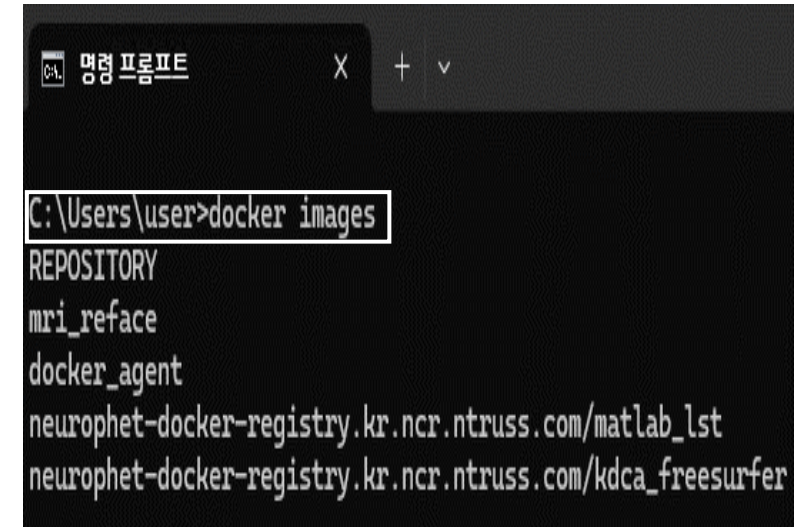
1. 명령 프롬프트 실행

- 윈도우 버튼을 누른 뒤 cmd를 검색한 후에 실행



2. WSL 상태 확인

- wsl -l -v 명령어를 입력
- 상태가 Running이라면 실행 중



3. 도커 이미지 확인

- docker images 명령어를 입력
- mri_reface와 freesurfer가 있으면 됨

실행 방법

CSV 파일 처리를 위한 패키지 설치

```
C:\Users\NEUROPHET>wsl  
user@DESKTOP-T8K4RUN:/mnt/c/Users/NEUROPHET$
```

1. 명령 프롬프트에서 WSL 접속

- wsl 명령어를 입력해 WSL에 접속

```
root@3bf91bc20461:/# pip3 list  
bash: pip3: command not found  
root@3bf91bc20461:/# apt update -y  
Get:1 http://security.ubuntu.com/ubuntu  
Hit:2 http://archive.ubuntu.com/ubuntu  
Get:3 http://archive.ubuntu.com/ubuntu  
Get:4 http://security.ubuntu.com/ubuntu  
Get:5 http://archive.ubuntu.com/ubuntu  
Get:6 http://archive.ubuntu.com/ubuntu  
Get:7 http://archive.ubuntu.com/ubuntu
```

2. pip3 설치 확인

- pip3 list 명령어를 입력
- bash: pip3: command not found 라고 뜬다면 아래 명령어 순서대로 실행
- sudo apt update -y
- sudo apt install python3-pip -y

```
root@3bf91bc20461:/# pip3 install pandas  
Collecting pandas  
  Downloading pandas-2.0.3-cp38-cp38-manylinux_2  
    | 12.4 MB  
Collecting python-dateutil>=2.8.2  
  Downloading python_dateutil-2.9.0.post0-py2.py  
    | 229 kB 9  
Collecting numpy>=1.20.3; python_version < "3.10  
  Downloading numpy-1.24.4-cp38-cp38-manylinux_2  
    | 17.3 MB
```

3. pandas 패키지 설치

- pip3 install pandas 명령어를 통해 설치
- pip3 show pandas 명령어를 입력했을 때 패키지 관련 정보가 뜬다면 설치 성공

실행 방법

```
root@DESKTOP-T8K4RUN: /mnt/d/neurophet/reface
root@DESKTOP-T8K4RUN: /mnt/d/neurophet/reface# docker ps
CONTAINER ID   IMAGE          COMMAND                  CREATED        STATUS        PORTS   NAMES
58050445728b   mri_reface     "sleep infinity"        11 seconds ago Up 8 seconds   mri_reface
```

도커 컨테이너 확인

- 컨테이너 이름을 mri_reface 로 고정 했기 때문에 이미 같은 이름의 컨테이너가 실행 중이라면 삭제하고 진행해야 함
- "docker ps" 명령어로 컨테이너가 존재하는지 확인
- 존재한다면 "docker rm -f mri_reface" 명령어로 삭제

실행 방법

	A	B
1	input_path	type
2	/mnt/c/pipeline_test/test_data/chris_t1_1.nii	T1
3	/mnt/c/pipeline_test/test_data/chris_t1_2.nii	T1
4	/mnt/c/pipeline_test/test_data/chris_t1_3.nii	T1
5	/mnt/c/pipeline_test/test_data/chris_t1_4.nii	T1
6	/mnt/c/pipeline_test/test_data/chris_t1_5.nii	T1

CSV 파일 생성

- 여러 개의 파일을 한 번에 처리하기 위해 파일을 리스트업하고, 이를 CSV 파일로 저장하여 데이터 Refacing을 진행
- 입력 파일의 경로와 modality를 지정
 - DICOM의 경우 폴더까지 지정, NIFTI 파일의 경우 nii 확장자만 가능
 - 이미지 타입은 T1, T2, FLAIR, FDG, PIB, FBP, TAU, CT, AUTO 중 하나로 지정
- WSL 기반의 파일 경로로 작성
 - D드라이브: /mnt/d
 - C드라이브: /mnt/c
 - 외장하드: /mnt/e /mnt/f 등으로 자동 연결됨
 - 연결되지 않았다면 "wsl --shutdown" 명령어로 WSL을 재시작 하면 연결됨

실행 방법

Argument	Description
--csv_path	- 입력 파일과 타입이 지정되어 있는 csv 파일 경로
--output_path	- 결과 파일들이 저장 될 경로 - output_path 하위에 입력 파일 이름으로 폴더가 생성
--multi	- 프로세스의 개수 - 기본 값은 1이며 지정하지 않아도 됨
--threads	- 프로세스 당 사용할 스레드의 개수 - 기본 값은 1이며 지정하지 않아도 됨

python 스크립트 실행

- scripts 폴더에 저장되어 있는 run_mri_reface.py 파일 사용
- WSL에 접속 후 옵션을 지정하여 실행하면 순차적으로 처리됨
- ex) `python3 /mnt/c/pipeline_test/scripts/run_mri_reface.py --csv_path=/mnt/c/pipeline_test/csv/practice/reface_input.csv --output_path=/mnt/c/pipeline_test/result/practice/reface --multi=1 --threads=4`

실행 방법

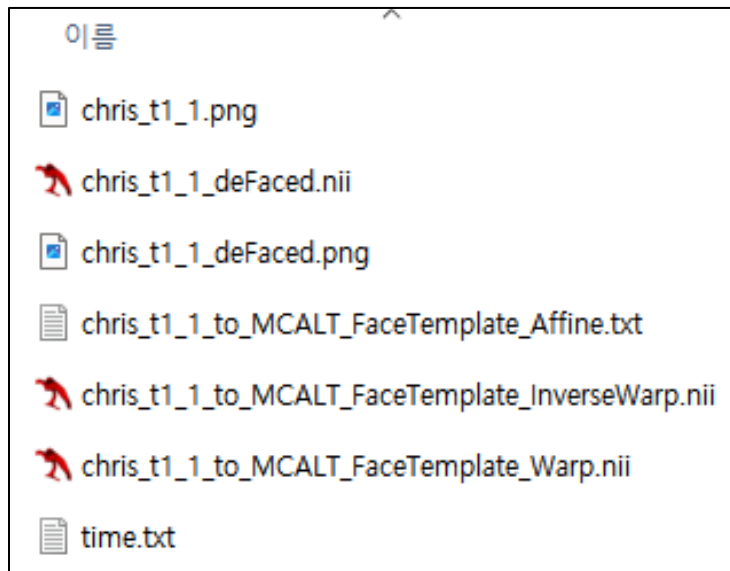
```
user@DESKTOP-T8K4RUN:/mnt/c/Users/NEUROPHET$ python3 /mnt/c/pipeline_test/scripts/run_mri_reface.py --csv_path=/mnt/c/pipeline_test/csv/practice/reface_input.csv --output_path=/mnt/c/pipeline_test/result/practice/reface --multi=1 --threads=4
docker_run_command = docker run --name=mri_reface --mount type=bind,src\=/tmp/tmp.tZjqbjJSxF,target=/tmp/tmp.tZjqbjJSxF neurophet-docker-registry.kr.ncr.ntruss.com/mri_reface sleep infinity
-----
Running mri_reface, version 0.3.2
By: Christopher G. Schwarz schwarz.christopher@mayo.edu
Temp directory: /tmp/tp46f66d40_c8ad_46c5_822f_ee7c577b711d
Using 4 threads.
Target file: /tmp/tmp.tZjqbjJSxF/inputs/chris_t1_1.nii
Output directory: /tmp/tmp.tZjqbjJSxF/outputs/chris_t1_1
Target image type: T1
Template: /root/.mcrCache9.12/mri_re0/mri_reface/images/MCALT_FaceTemplate_T1.nii
Face atlas: /root/.mcrCache9.12/mri_re0/mri_reface/images/MCALT_FaceMask.nii

Calculating affine parameters to template. This will take some time.
Using face-optimized affine.
Estimating warp to template using ANTS. This will take a while.
```

실행 확인

- 정상적으로 동작하는 경우 다음과 같은 로그가 출력됨
- 사용하는 옵션에 따라 시간 차이가 발생하며 1건 당 약 10분 정도 소요됨

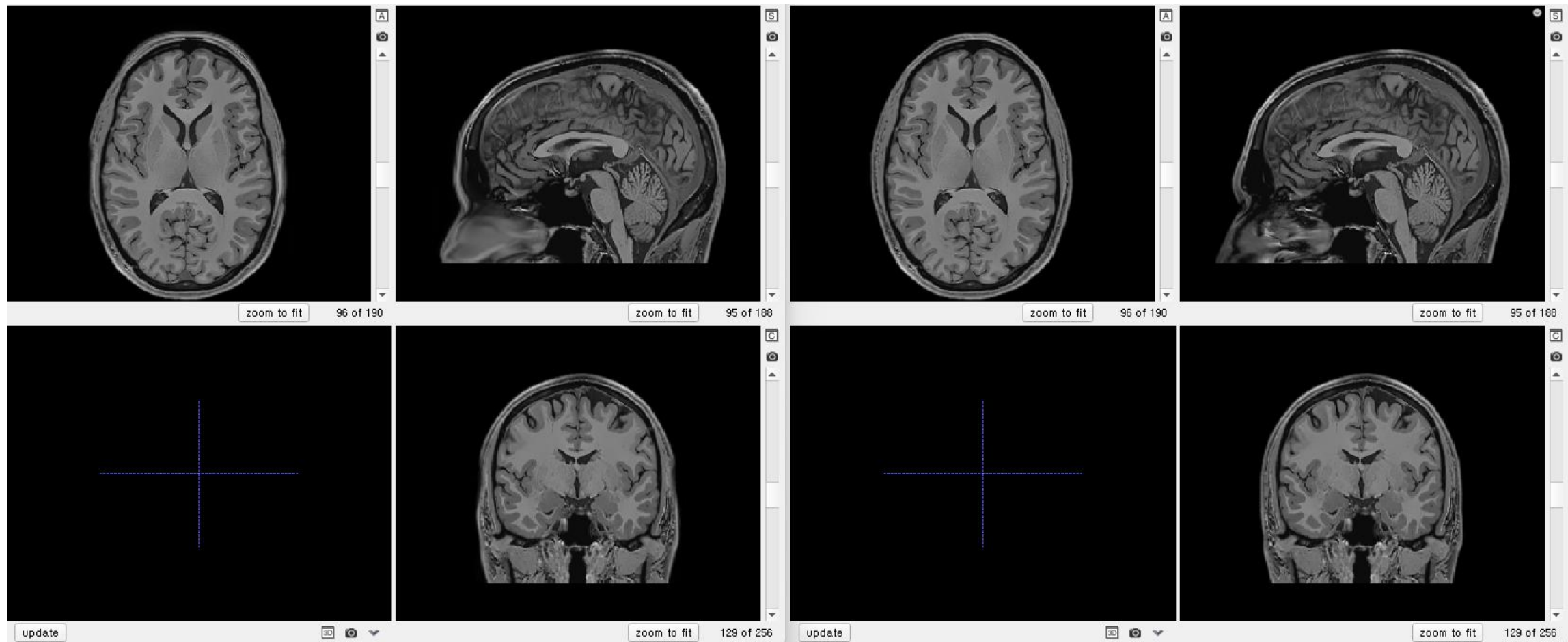
결과 확인



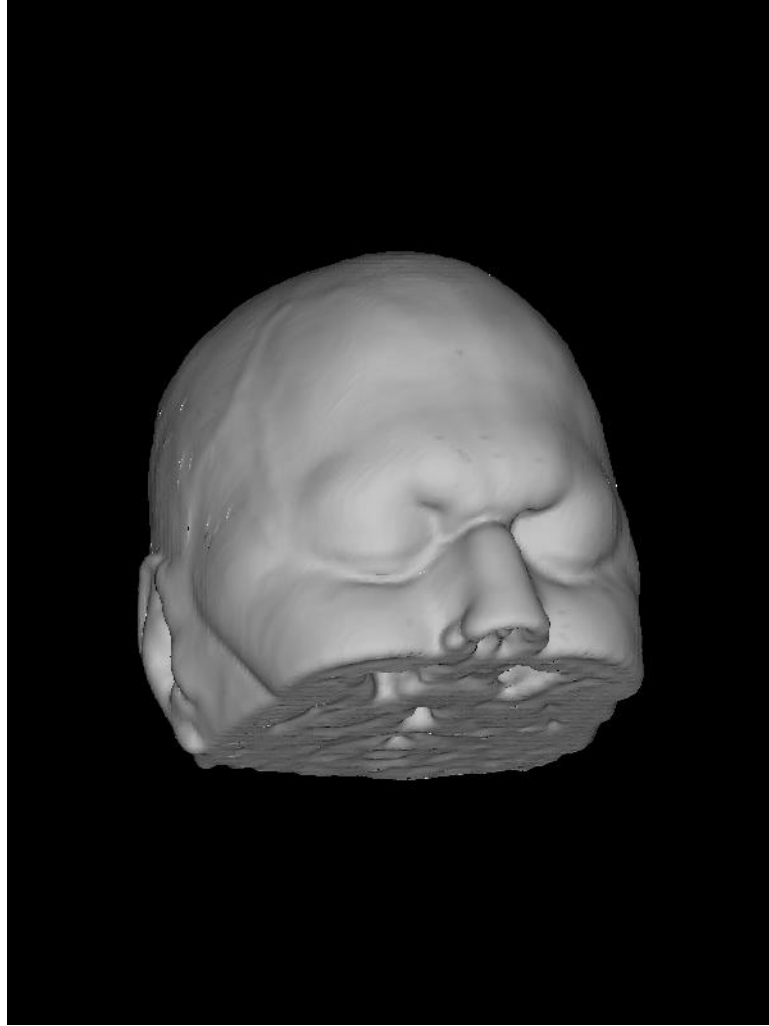
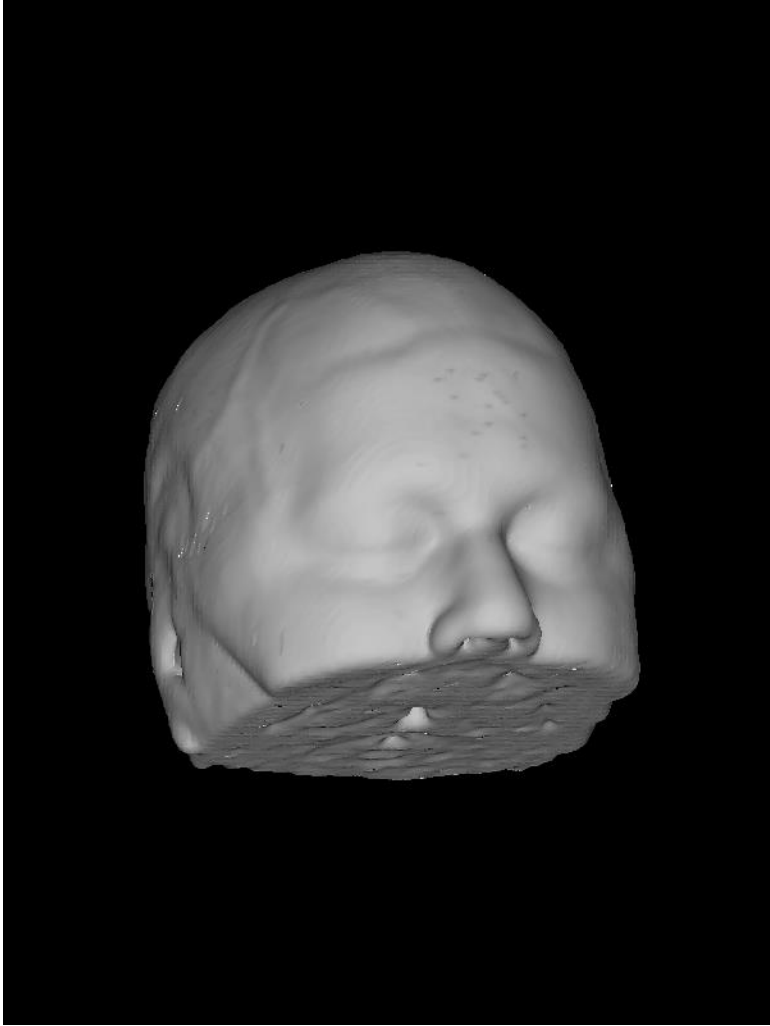
데이터 구성

- *_deFaced.nii: 입력 파일에 refacing이 적용 된 파일
- png: 원본의 대표 이미지와 refacing된 파일의 대표 이미지

결과 확인



결과 확인

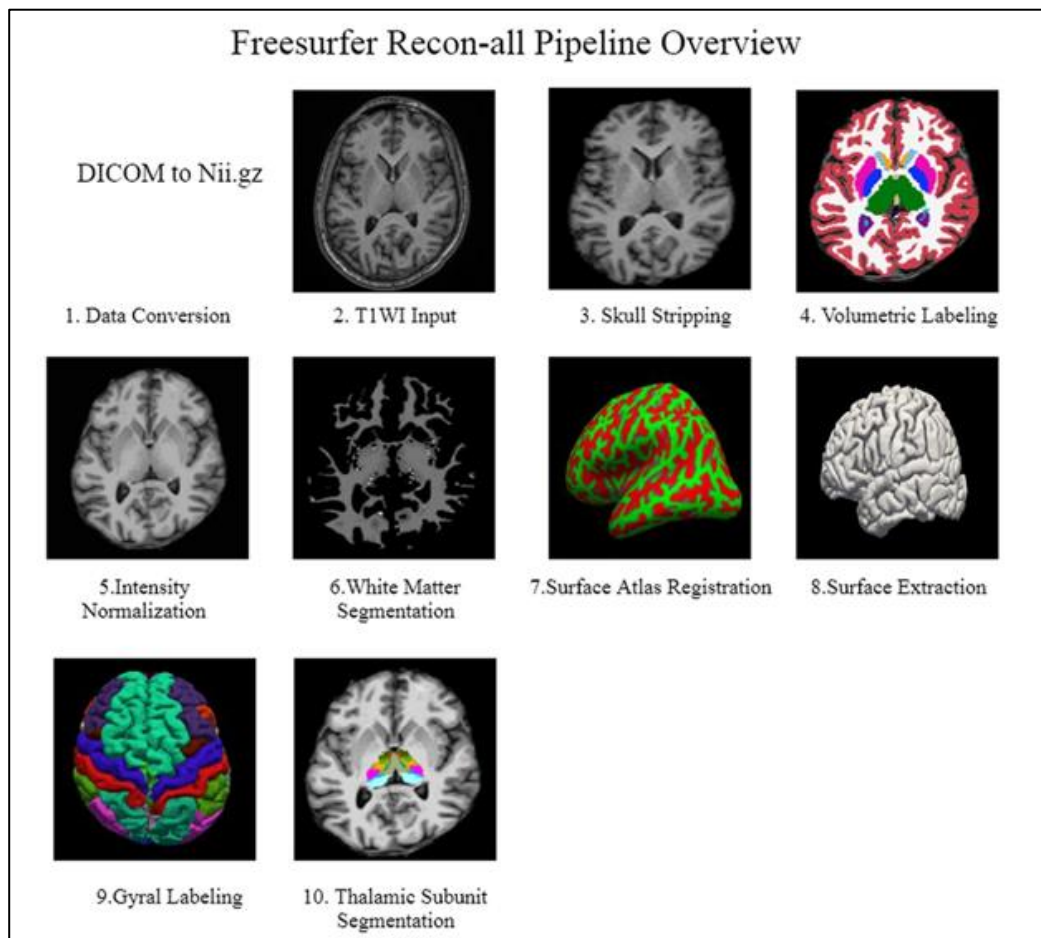


Chapter 2

FreeSurfer (+PETSurfer)

- 1 개요
- 2 실행 방법
- 3 결과 확인
- 4 결과 추출

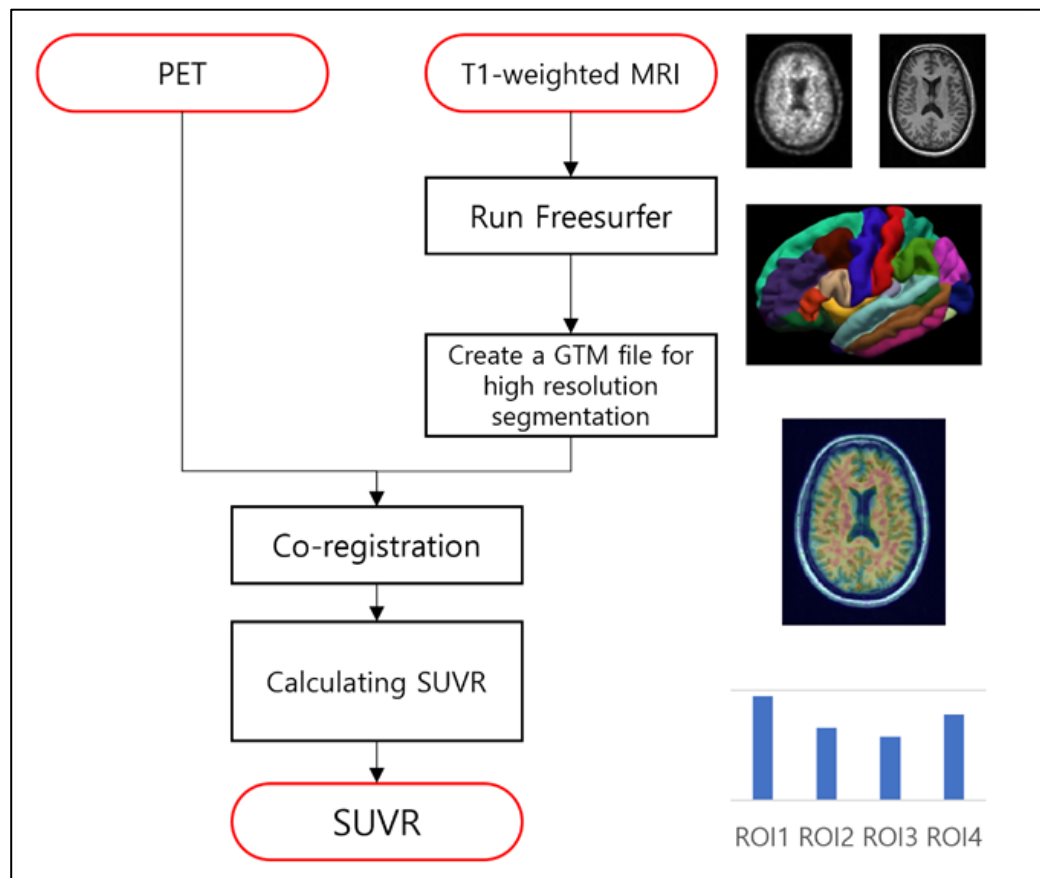
개요



FreeSurfer

- MRI 스캔 데이터를 분석하여 뇌의 해부학적 및 기능적 특성을 추출하는 소프트웨어 패키지
- 뇌의 부피, 표면 모양, 회색질과 백색질의 해부학적 특성을 추정하고 해부학적 레이블을 지정하는 등의 기능을 제공하여 뇌 영상 연구 및 임상 평가에 활용됨

개요



PETSurfer

- FreeSurfer 내에 포함되어 있는 MRI-PET 분석을 위한 도구
- motion correction, PET-MRI registration 및 기준 영역 kinetic modeling, 볼륨 보정 등의 과정을 통해 SUVR을 추출
 - Standardized Uptake Value Ratio (표준섭취계수율)

실행 방법

도커 컨테이너 생성

- FreeSurfer의 경우 임시 폴더가 아닌 실제로 지정한 경로에서 작업이 실행되기 때문에, 컨테이너를 미리 생성하고 진행
- `docker run --name={컨테이너 이름} -d -v {로컬 경로}:{컨테이너 내부의 경로} {이미지 ID} sleep infinity`
- ex) `docker run --name=fs_710 -d -v /mnt/c/pipeline_test:/mnt/c/pipeline_test 0799 sleep infinity`
 - -d: Detached 모드, 백그라운드로 실행하는 옵션
 - -v: 호스트와 컨테이너의 폴더를 연결하기 위한 옵션
 - --name: 컨테이너 이름 설정 (fs_710으로 고정하였음)

```
user@DESKTOP-T8K4RUN:/mnt/c/Users/NEUROPHET$ docker run --name=fs_710 -d -v /mnt/c/pipeline_test:/mnt/c/pipeline_test 0799 sleep infinity
9f6866c44ffc00ab71afe9f848a81eb8ea58ab83d77e390eae3ec1aff55e73ee
```

실행 방법

	A	B	C	D	E	F
1	subject_id	input_path	output_path	openmp	pet_path	psf
2	chris_t1_1	/mnt/c/pipeline_test/result/sample/reface/chris_t1_1/chris_t1_1_deFaced.nii	/mnt/c/pipeline_test/result/sample/freesurfer	4		
3	chris_t1_2	/mnt/c/pipeline_test/result/sample/reface/chris_t1_2/chris_t1_2_deFaced.nii	/mnt/c/pipeline_test/result/sample/freesurfer	4		
4	chris_t1_3	/mnt/c/pipeline_test/result/sample/reface/chris_t1_3/chris_t1_3_deFaced.nii	/mnt/c/pipeline_test/result/sample/freesurfer	4		
5	chris_t1_4	/mnt/c/pipeline_test/result/sample/reface/chris_t1_4/chris_t1_4_deFaced.nii	/mnt/c/pipeline_test/result/sample/freesurfer	4		
6	chris_t1_5	/mnt/c/pipeline_test/result/sample/reface/chris_t1_5/chris_t1_5_deFaced.nii	/mnt/c/pipeline_test/result/sample/freesurfer	4		

CSV 파일 생성

- 여러 개의 파일을 한 번에 처리하기 위해 CSV 파일로 저장하여 데이터 분석을 진행
- subject_id: 분석 대상을 구분하기 위한 ID, 결과 폴더 이름으로 사용됨
- input_path: MRI 파일의 경로를 지정, DICOM일 경우 시리즈 1개의 파일을 지정
- output_path: 결과를 저장할 경로, 해당 경로에 subject_id로 지정한 이름으로 폴더가 생성됨
- openmp: 병렬 처리를 위해 지정하는 옵션, 사용할 CPU 코어의 수를 지정하면 됨
- pet_path(선택): PETSurfer를 진행하는 경우에만 사용, MRI와 pair 인 PET 파일 경로를 지정
- psf(선택): PETSurfer를 진행하는 경우에만 사용하며 pvc를 진행하지 않을 거라면 0으로 지정하면 됨
 - pvc: partial volume correction
 - psf: point-spread function

실행 방법

Argument↵	Description↵
--csv_path↵	- 분석 옵션이 지정되어 있는 csv 파일 경로↵
--multi↵	- 한 번에 몇 개의 분석을 진행할지 설정하는 옵션↵ - 기본값은 1↵ ↵

python 스크립트 실행

- scripts 폴더에 저장되어 있는 run_freesurfer_for_test.py 파일 사용
 - 전체 프로세스가 아닌 일부분만 실행하도록 수정된 파일
 - <https://surfer.nmr.mgh.harvard.edu/fswiki/recon-all>
- WSL에 접속 후 옵션을 지정하여 실행
- ex) python3 /mnt/c/pipeline_test/scripts/run_freesurfer_for_test.py --csv_path=/mnt/c/pipeline_test/csv/practice/freesurfer_input.csv --multi=1

실행 방법

```
user@DESKTOP-T8K4RUM: /m  x + v
First Sorting
Computing Slice Direction
Vs: -0.024981 0.037616 0.873834
Vs: -0.0285497 0.0429897 0.998668
Second Sorting
IsDWI = 0, IsPhilipsDWI = 0
Counting frames
nframes = 1
nslices = 256
ndcmfiles = 256
INFO: rescale not needed
PE Dir = UNKNOWN (dicom read)
Loading pixel data
TR=0.00, TE=0.00, TI=0.00, flip angle=0.00
i_ras = (-0.999561, -0.00914398, 0.0281819)
j_ras = (0.00792003, -0.999033, -0.0432401)
k_ras = (0.0285497, -0.0429897, 0.998668)
writing to /mnt/d/neurophet/kdca/test/sub_0001/mri/orig/001.mgz...
#
### MotionCor Thu May 2 07:26:08 UTC 2024
Found 1 runs
/mnt/d/neurophet/kdca/test/sub_0001/sub_0001/mri/orig/001.mgz
Checking for (invalid) multi-frame inputs...
Only one run found so motion
correction will not be performed. I'll
copy the run to rawavg and continue.

cp /mnt/d/neurophet/kdca/test/sub_0001/sub_0001/mri/orig/001.mgz /mnt/d/neurophet/kdca/test/sub_0001/sub_0001/mri/rawavg.mgz

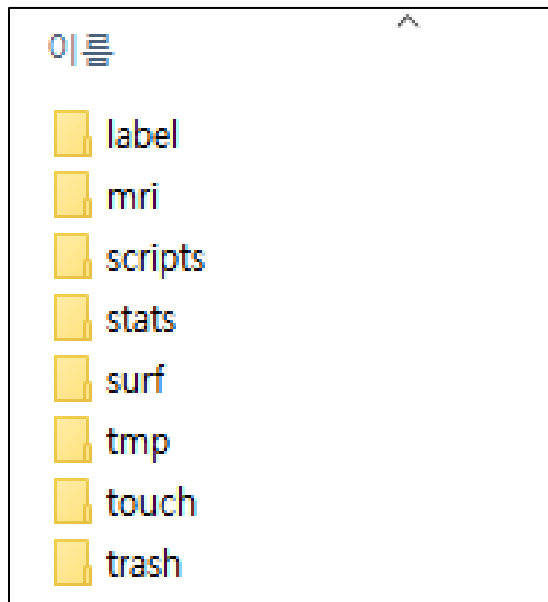
/mnt/d/neurophet/kdca/test/sub_0001/sub_0001

mri_convert /mnt/d/neurophet/kdca/test/sub_0001/sub_0001/mri/rawavg.mgz /mnt/d/neurophet/kdca/test/sub_0001/sub_0001/mri/orig.mgz --confor
m
```

실행 확인

- 정상적으로 동작하는 경우 다음과 같은 로그가 출력됨
- 단일로 분석하는 경우 1건 당 2.5 ~ 4시간 정도 소요됨

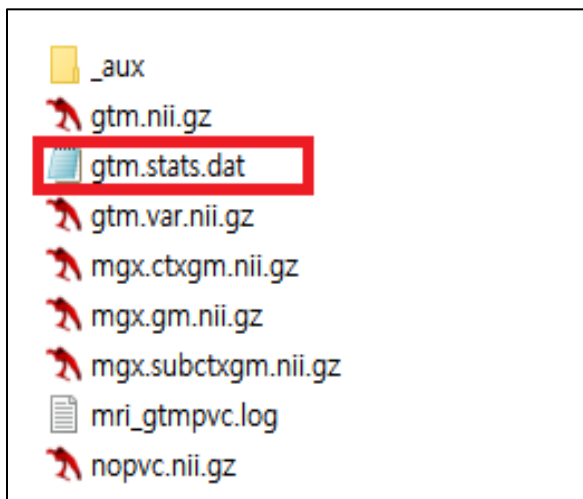
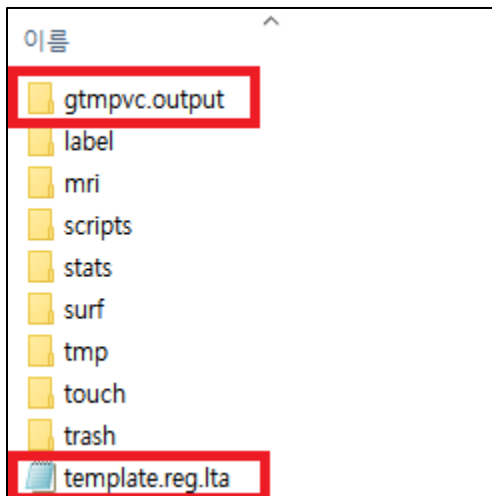
결과 확인



FreeSurfer 결과

- label : 영상 분석 결과에 대한 레이블링 정보
- mri : 비식별화된 MRI 영상과 그 정보를 포함
- scripts : 분석 과정을 기록한 스크립트
- stats : 영상 분석 결과에 대한 통계 정보 (뇌 영역별 부피, 두께, 면적 등의 평균, 분산, 최소-최대값 정보)
- surf : 뇌의 surface에 해당하는 분석 결과 및 정보
- tmp : 임시 파일들
- touch : 실행한 명령어 목록
- trash : 결과에 필요하지 않은 파일들

결과 확인



PETSurfer 결과

- csv에 pet 경로를 지정했다면 PETSurfer의 결과까지 같이 생성됨
- gtmpvc.output, template.reg.lta를 제외한 8개의 폴더는 FreeSurfer의 결과
 - template.reg.lta: registration 정보가 저장된 파일
 - gtmpvc.output: GTM pvc 결과 폴더
- suvr 값은 gtmpvc.output/gtm.stats.dat 파일에 저장되어 있음

결과 추출 - FreeSurfer

segmentation 파일 사용

- mri 폴더에 저장된 aparc+aseg.mgz 파일을 읽고 각 라벨 별 volume을 계산하고 저장
- 사용 스크립트: get_freesurfer_result.py
- 필요 패키지: nibabel, numpy, pandas
 - pip3 install nibabel

Argument	Description
--result_path	- FreeSurfer 결과들이 저장되어 있는 경로
--csv_save_path	- volume 값이 정리된 csv 파일을 저장할 경로
--package_path	- python 패키지가 설치되어 있는 경로 - 이미 패키지가 설치되어 있다면 지정하지 않아도 됨

	A	B	C	D	E	F
1	id	0	2	4	5	7
2	sub_0001	15564497.86	254832.0304	16303.00194	609.0000726	13508.00161
3	sub_0002	15890935	153913	28157	1167	9008
4	sub_0003	15807049	210035	16630	1003	12016
5	sub_0004	15689122	207510	15569	865	14257
6	sub_0005	15799445.12	194308.9768	18401.99781	531.9999366	10272.99878

python 스크립트 실행

- scripts 폴더에 저장되어 있는 get_freesurfer_result.py 파일 사용
- WSL에 접속 후 옵션을 지정하여 실행
- ex) python3 /mnt/c/pipeline_test/scripts/get_freesurfer_result.py --result_path=/mnt/c/pipeline_test/result/sample/freesurfer --csv_save_path=/mnt/c/pipeline_test/freesurfer_result.csv
- <https://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial/AnatomicalROI/FreeSurferColorLUT>

결과 추출 - PETSurfer

gtm.stats.dat 파일 사용


- FreeSurfer 결과 중 gtmpvc.output 폴더에 있는 gtm.stats.dat 파일을 읽어 결과 추출
- 사용 스크립트: get_petsurfer_result.py
- 필요 패키지: pandas

Argument	Description
--result_path	- FreeSurfer 결과들이 저장되어 있는 경로
--csv_save_path	- volume 값이 정리된 csv 파일을 저장할 경로
--package_path	- python 패키지가 설치되어 있는 경로 - 이미 패키지가 설치되어 있다면 지정하지 않아도 됨

python 스크립트 실행

- scripts 폴더에 저장되어 있는 get_petsurfer_result.py 파일 사용
- WSL에 접속 후 옵션을 지정하여 실행
- ex) python3 /mnt/c/pipeline_test/scripts/get_petsurfer_result.py --result_path=/mnt/c/pipeline_test/result/sample/freesurfer --csv_save_path=/mnt/c/pipeline_test/petsurfer_result.csv

	A	B	C	D	E	F
1	id	left_cerebral_white_matter_wm	left_cerebellum_white_matter_wm	left_cerebellum_cortex_subcort_gm	left_thalamus_subcort_gm	left_caudate_subcort_gm
2	sub_0001	1.204	1.119	1.047	1.208	1.297
3	sub_0002	1.397	1.151	0.927	1.309	1.574
4	sub_0003	1.093	1.057	0.897	1.186	1.497
5	sub_0004	1.053	1.08	0.977	1.211	1.216

neurophet
Overcome brain disorders
with advanced neuroscience

www.neurophet.com

Tel +82 2.6954. 7971

Fax +82 2.6954. 7972

© 2024. NEUROPHET Inc. All Rights Reserved.