**Lower-Level Stimulus Features and FFA Activation**

2021311097

정 명 교

1. **Introduction**

Fusiform Face Area (FFA)는 얼굴을 인식하는데 주요한 역할을 한다고 알려져 있다. 하지만, [[1]](#endnote-1)FFA가 face의 lower-level stimulus, 가령 size, contrast, rotation, 그리고 viewpoint와 같은 특성을 처리하는 기능을 수행하는지는 명확히 알려진 바가 없다.

이번에 쓰인 연플리 movie data를 보면 여러 인물들의 얼굴이 상당히 자주 등장하는 것을 볼 수 있다. 이를 FFA와 관련지어 face의 [[2]](#endnote-2)lower-level stimulus을 영화 데이터에서 구분하고, 그것들 간의 차이가 FFA 내에서 유의미하게 나타나는지 분석하고 싶은 욕구가 생겼다.

따라서 lower-level stimulus 중에서 비교적 간단한 size of face 와 number (#) of face를 탐구 주제로 선정했다. 먼저, FFA에서 face의 size에 따른 선형적 관계성이 존재한다는 가설을 세웠고, multiple faces와 single face가 반응하는 강도의 세기가 다를 것이라는 가설도 세웠다. 이를 알아보기 위해 ROI analysis와 Whole brain analysis를 했다. 만약 FFA를 통한 ROI analysis에서 multiple faces와 single face 간의 유의미한 차이가 없다면, number or multiple (#)에 반응하는 brain region이 있지 않을까 하는 새로운 의구심이 들어 Whole brain analysis에서는 multiple faces에 대하여 통계적으로 유의미한 결과를 보이는 지점을 찾으려는 시도도 해보았다.

1. **Methods**

실험에 참가한 피실험자들은 남성 2명, 여성 2명인 총 4명으로 평균 나이는 26.75살(표준 편차 =1.09 년)이다. 각 피험자들은 스캐너에서 FOSS localizer 실험과 영화 실험에 대해 각각 이미지 (Face, Objects, Scenes, Scrambled)와 영상을 보았다.

FOSS는 “Face, Objects, Scenes, Scrambled”의 앞 글자를 딴 것으로, 얼굴, 장소, 물체 범주에 속하는 이미지를 봤을 때 선별적으로 반응하는 뇌 지역의 위치를 알아내기 위한 4개의 범주 block-design 실험 패러다임이다. 실험 디자인은 피험자가 예측할 수 없도록 배치된 각 14초의 자극들을 하나씩 보여줘 총 56초의 한 block이 있고 자극 block은 총 4개가 있고, 14초의 fixation block이 모든 자극 block을 둘러싼 형태이다. 첫번째 run의 순서는 [x, F, O, S, P, x, O, P, F, S, x, S, P, F, O, x, P, S, O, F, x], 두번째 run의 순서는 [x, P, S, O, F, x, S, P, F, O, x, O, P, F, S, x, F, O, S, P, x]이고 각 run당 총 TR은 147 TRs(14 x 5 + 56 x 4 = 4분 56초) 이다. 영화는 한국의 웹 드라마 “연애플레이리스트”의 3개의 에피소드를 자막에 검은 박스로 덮어 가린 채로 피험자들에게 보여주었다. 각 에피소드 앞에는 10초의 blank가 있은 후에 영상이 나오는데, 각 에피소드 마다 986초, 948초, 700초로, 총 2664(996 + 958+ 710) TRs이다.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Run | Experimental design (each is one block) | | | | | | | | | | | | | | | | | | | | |
| 1 | x | F | O | S | P | x | O | P | F | S | x | S | P | F | O | x | P | S | O | F | x |
| 2 | x | P | S | O | F | x | S | P | F | O | x | O | P | F | S | x | F | O | S | P | x |
| Notation | x-Fixation, F-Face images, O-Object images, S-Scrambled images, P-Place images- | | | | | | | | | | | | | | | | | | | | |

촬영한 MRI는 64 채널의 head coil이 장착된 3T Siemens Prisma MRI 스캐너이다. 다섯 개의 functional(T2\*) runs와 한 개의 anatomical run(T1)으로 구성된 한 scout을 collect했고 functional runs은 세 개의 영화 runs가 두 개의 FOSS localizer runs와 함께 구성돼 있다. Anatomical scan parameter는 1 x 1 x 1 mm voxel size, 256 sagittal slices, 2.44 ms TE, 2200 ms TR, 8° flip angle, 256 x 256 mm FOV이고, functional scan parameter는 FOSS와 영화 둘 다 같은 3 x 3 x 3 mm voxel size, 90° flip angle, axial slices, PROPixx 프로젝터(head coil에 고정된 거울에 투사되게 함)이다. 영화는 80 x 80 matrix size, 240 x 240 mm FOV, 30 ms TE, 48 slices이고, FOSS localizer는 84 x 72 matrix size, 252 x 216 mm FOV, 20 ms TE, 44 slices이다.

얼굴의 face size를 명확하게 인식하고 big face와 small face로 구분하기 위하여 OpenCV라는 오픈소스 라이브러리를 활용했다. 그 중에서도 Face detection을 하는데 뛰어난 성능을 보이는 Media pipe 라이브러리의 face mesh를 이용해서 사람의 얼굴을 자동으로 검출하고 얼굴의 크기에 따라 big face와 small face로 분류할 수 있는 프로그램을 구현했다.

Small face를 detection 할 때에는 yellow box를 만들어서 얼굴의 면적이 전체 프레임의 5% 이하의 비율을 차지할 때로 선정했다. Big face를 detection 할 때에는 blue box를 만들어서 얼굴의 면적이 전체 프레임의 15% 이상의 비율을 차지할 때로 선정했다. 5%와 15%로 구분한 이유는 다음과 같은 숫자를 기준점으로 잡았을 때, 우리가 직관적으로 느끼는 큰 얼굴과 작은 얼굴의 차이를 명확하게 구별 지을 수 있는 인지적 기능 구현이 가능했기 때문이다.

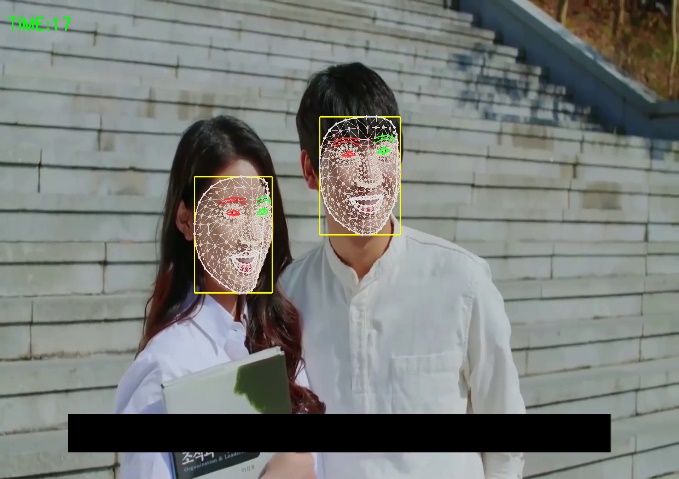
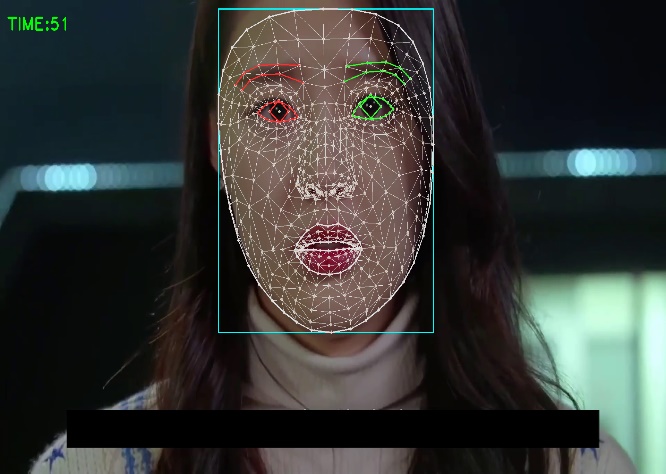
Mediapipe face mesh 구현을 통해 재구성된 영상을 보면, face를 detect는 하지만 박스가 형성이 안되는 경우를 볼 수 있는데 이 경우는 얼굴이 전체 frame의 5% 초과 15%미만인 경우로 얼굴이 등장하는 장면들이 너무 많기에 이들의 수를 줄이면서도 small face와 big face간의 차이를 명확하게 구분하기 위해서 design matrix에 포함시키지 않았다.

텍스트이(가) 표시된 사진

자동 생성된 설명

Mediapipe face mesh 구현을 위한 코드

Single face에 대한 효과를 확인했다면, multiple faces와 single face 간의 차이가 나타나는 지도 분석하기 위해 Crowd를 single과 비교하는 실험도 수행했다. 이때, mutipiple faces에 대한 coefficient는 한 프레임 내에서 서로 다른 인물이 3명 이상 나올 때를 기준으로 하였다. Crowd에 대한 Baseline은 single face 중에서도 big face를 기준으로 하였다. Blue box에서는 인물이 1명만 detection 되지만, yellow box에서는 여러 명의 얼굴이 동시에 detection 되는 경우가 존재하기 때문이다.

Small face (5%) 와 Big face (15%)의 detection

모든 데이터는 fMRIprep을 통해 anatomical preprocessing과 functional preprocessing을 했고, 그 후 공통적으로 additional preprocessing을 했다. Anatomical preprocessing은 bias filed correction, skull-stripping, spatial normalization (MNI space)을 포함했고, functional preprocessing은 motion correction, co-registration, spatial normalization (MNI space), noise component extraction을 포함했고, additional preprocessing은 masking, detrending, spatial smoothing을 포함했다.

Localizer 데이터의 ROI는 FFA로 했고, Small Face – Empty (Baseline), Big Face – Empty, Big Face - Small Face, Crowd – Empty의 값들을 측정했다. 각 ROI를 만들 때 paired t-test를 했고, threshold는 공통으로 0.3보다 큰 좌표를 잡아서 만들었다.

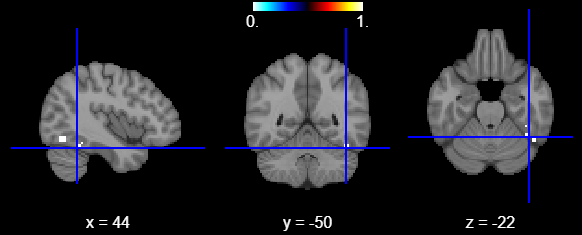
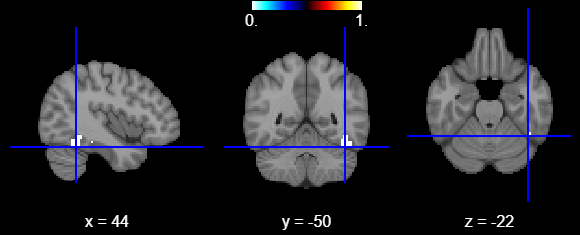
 

Figure 1. 영화 실험 ROI mask(왼쪽)와 FOSS localizer 실험 ROI mask(오른쪽)

또한 GLM에 가설을 입히기 위해 event regressor에 Small Face, Big Face, Crowd을 각각 11개의 trials이 들어가게 했다. Multi voxel의 pattern이 각 condition 내에서는 비슷하고 서로 다른 condition은 similarity가 낫을 것이라는 가설 아래에, RSA를 수행했다.

1. **Results**

functional images에 대한 motion correction을 수행하기 위해서, 모든 subjects의 모든 runs에 대한 motion parameters를 추정했다. 아래는 그들 중에서 subject 1번의 run 1에 대한 예시이다.

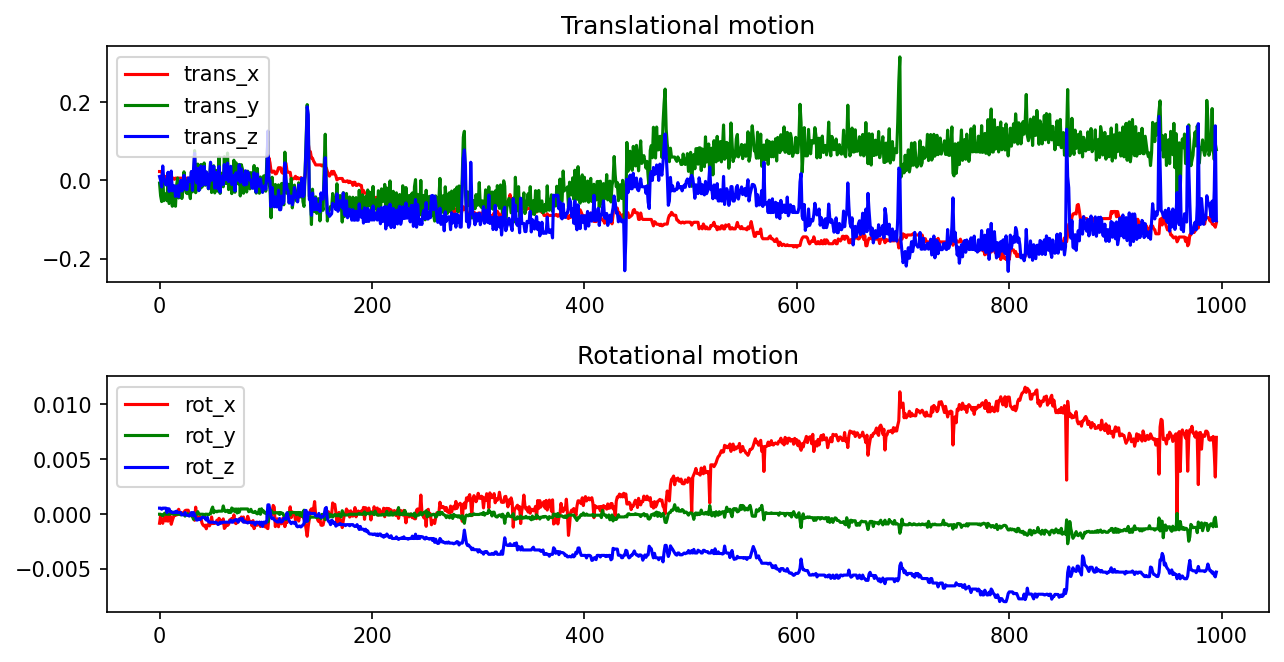


Figure 2. Subject 1번, run 1의 시간에 따른 motion parameters over time

Fig.3는 시간에 따른 motion parameter와 첫번째 피험자의 Control과 Experimental의 event regressor이다. Motion parameter에는 x, y, z축으로의 각각의 움직임, 회전, 미분 값, 제곱으로 이뤄져 있다. 그 아래 그림에서 Event regressor의 weight은 모두 같게 했다.

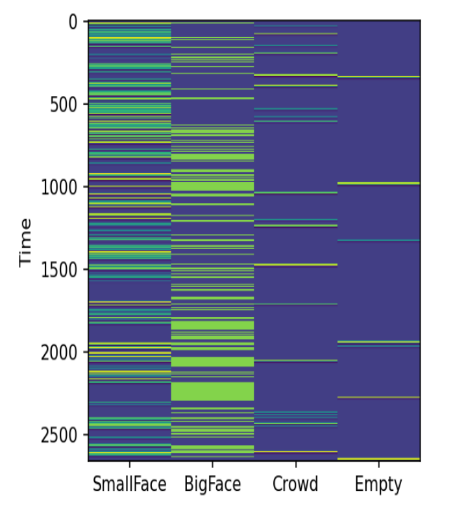
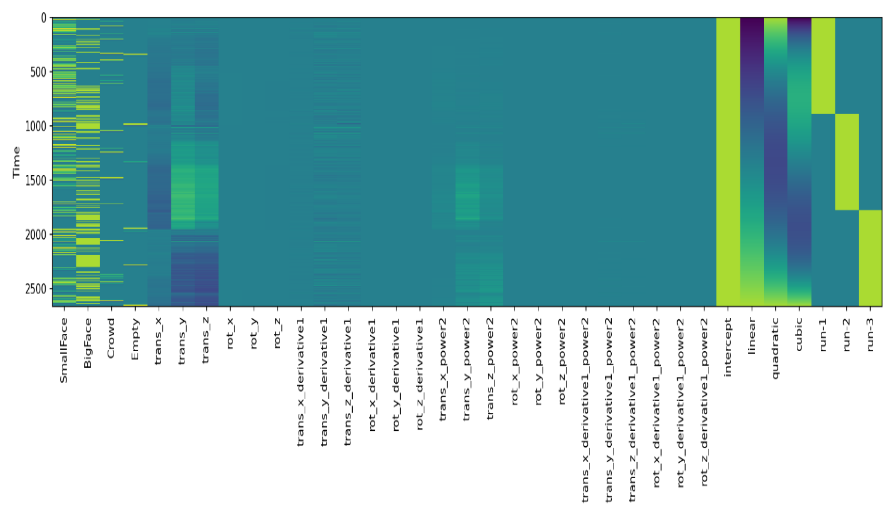


Figure 3. Motion parameters와 event regressor

Big Face, Small Face, Crowd의 condition에 각각 hrf를 convolve했다. hrf는 binary regressors로 조건을 만족하는 지의 여부에 따라 1 또는 0의 값을 output으로 내보낸다.

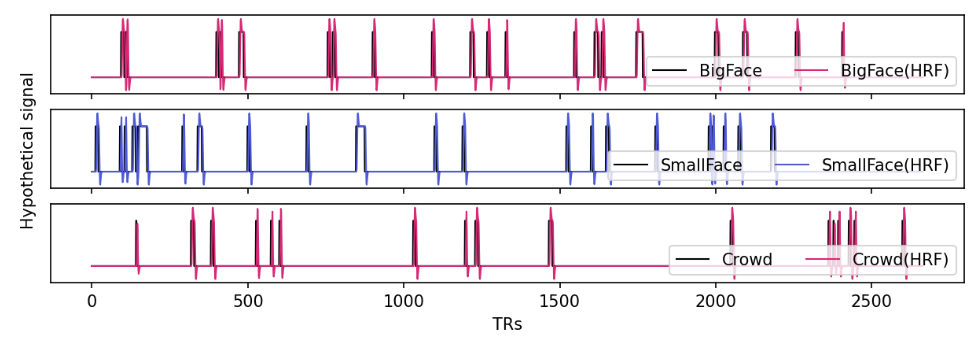


Figure 4. event regressor에 hrf를 convolve한 것

Whole brain analysis는 Experimental beta에서 Control beta에 대해 paired t-test를 p<0.05를 기준으로 T map에 나타냈다. Big Face – Empty, Small Face – Empty는 FFA에서 그 차이를 확인했지만 Big Face – Small Face는 FFA에서 유의미하 차이를 발견하지 못했다. 대신, [[3]](#endnote-3)Amygdala와 Hippocampus에서 차이를 발견할 수 있었는데, 얼굴이 크게 나오는 장면이 인물이 화를 내거나, 짝사랑하는 상대를 미련 가득한 눈빛으로 쳐다보거나, 눈물을 흘리는 신인데 이런 부분에서 표정이나 감정이 명확하게 나타나기 때문에 이러한 결과가 나온 것 같다. 또한 아무래도 Big Face가 한 인물에 focus를 하다 보니 Small Face에 비해서 인물이 재등장할 때 기억이나 회상이 더 활성화되기에 활성화가 나타나는 것으로 보인다.

Crowd Face – Single Face은 FFA에서 유의미한 차이를 발견하지 못했다. 대신, 차이를 보이는 지점의 좌표를 neurosynth을 통해 확인해보니 공통적으로 place, spatial과 관련된 keyword가 나왔는데 여러 face가 나오면 face 그 자체에만 집중하기 보다는 spatial한, 그리고 place에도 attention이 생기기에 다음과 같은 결과가 나온 것으로 보인다.

다음은 Whole brain analysis의 결과이다.

텍스트, 실내이(가) 표시된 사진

자동 생성된 설명 텍스트, 실내이(가) 표시된 사진

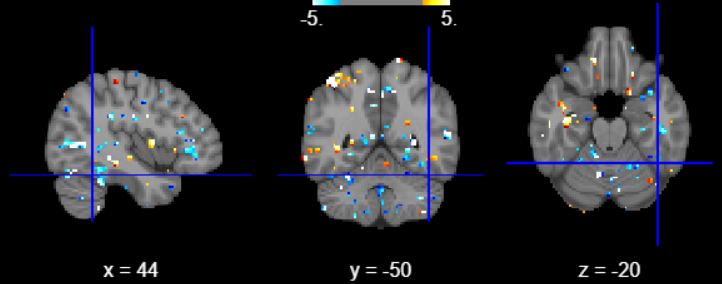
자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트이(가) 표시된 사진

자동 생성된 설명

Figure 5. 왼쪽부터 차례로 Big Face – Empty, Small Face – Empty

 텍스트, 다른, 여러개이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트이(가) 표시된 사진

자동 생성된 설명

Figure 6. 왼쪽부터 차례로 Big Face – Small Face의 FFA, Amygdala / Hippocampus

텍스트, 다른이(가) 표시된 사진

자동 생성된 설명 지도이(가) 표시된 사진

자동 생성된 설명

텍스트이(가) 표시된 사진

자동 생성된 설명 텍스트이(가) 표시된 사진

자동 생성된 설명

Figure 7. Crowd Face – Single Face (Big Face)

ROI analysis는 Big Face 와 Small Face, 그리고 Crowd와 Big Face에 대해서 paired t-test를 해보았다. Big Face가 Small Face에 비해, face의 size에 따라 linear한 관계가 있을 것이라는 가설과는 정반대의 결과가 나와 design matrix를 다시 reconstruction하여 결과를 도출할 필요성이 있어 보인다. 그에 반해, crowd는 single face에 비해 더 큰 효과를 발견할 수 있었다.

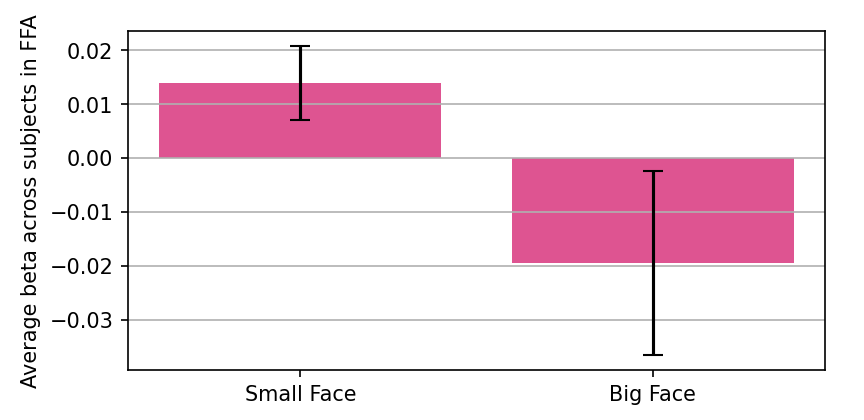
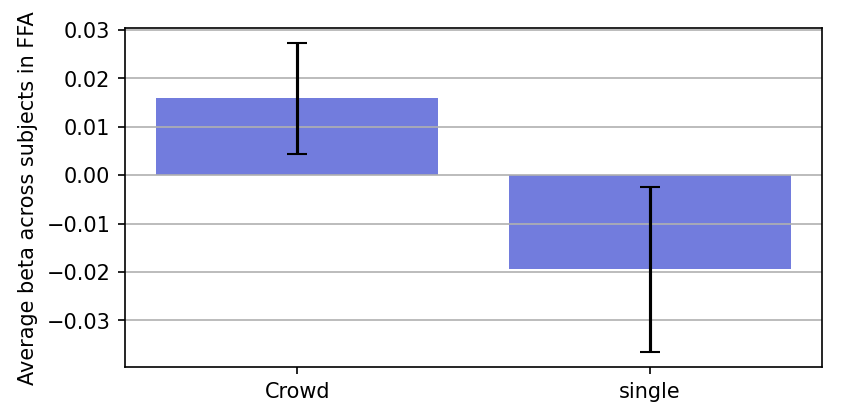
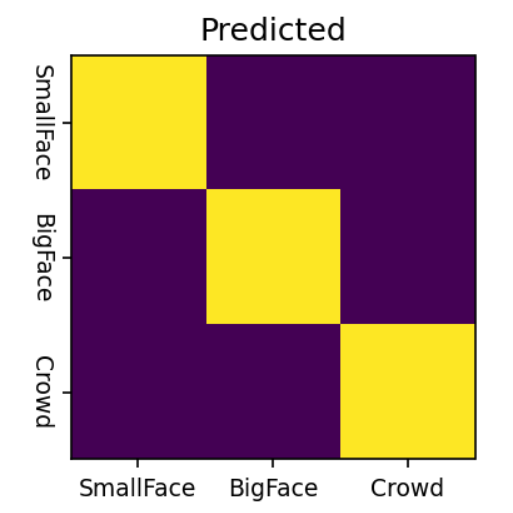
 

Fig.4 – ROI analysis 결과

추가적인 분석으로 RSA를 사용했는데, Small Face, Big Face, Crowd끼리는 유사도가 높고, 서로 간에는 유사도가 낮은 model matrix를 기대했지만 완전히 같지는 않았다. 그나마 Small Face끼리는 높은 유사성을 보였고 Small Face 와 Big Face 간에도 어느정도 유사성이 있다는 것을 확인할 수 있다. 이를 통해 single face와 multiple face를 구분하는 그 기저가 서로 다른 독립된 경로라는 것을 유추할 수 있다.

 텍스트, 점수판, 디스플레이이(가) 표시된 사진

자동 생성된 설명

Fig.5 – RSA를 통해 보인 Small Face, Big Face, Crowd의 유사도

1. **Conclusions**

ROI, Whole brain, RSA 분석 결과에 따르면 ace의 size에 따른 선형적 관계성의 존재는 잘 맞아 떨어지지는 않는다. 이는 추후 코드 개선과 face의 size의 재설정을 통해 실험을 재구성해야 할 필요가 있음을 의미한다. 또다른 가설인 multiple faces와 single face의 강도 세기의 차이는 발견하지 못했지만 이 둘을 처리하는 brain pathway가 서로 독립적인 system으로서 작동할 수 있다는 점은 어느정도 유추할 수 있었다. 따라서 FFA는 어느정도는 lower-level stimulus를 처리하는 기능을 수행한다고 볼 수 있다.

하지만 분석을 하는 과정에서 multiple comparison problem이 있을 뿐만 아니라, 가설 혹은 가설의 trial을 정의하는 데에 문제가 많다. Trial을 비교할 때 covariate의 존재 가능성이 분명히 있고, face의 size와 crowd가 되는 기준을 자의적으로 설정했기 때문에 일반성이 떨어질 수 있다. 따라서 이 실험으로부터 도출한 결론은 타당성이 높다고는 말하기 어렵다.

1. **References**

1. Rossion, B(2015). Face Perception. *Brain Mapping: An Encyclopedic Reference*, 515-522. https://doi.org/10.1016/B978-0-12-397025-1.00037-3 [↑](#endnote-ref-1)
2. Yue X, Cassidy BS, Devaney KJ, Holt DJ, Tootell RB. Lower-level stimulus features strongly influence responses in the fusiform face area. Cereb Cortex. 2011 Jan;21(1):35-47. doi: 10.1093/cercor/bhq050. Epub 2010 Apr 7. PMID: 20375074; PMCID: PMC3000561. [↑](#endnote-ref-2)
3. Baxter MG, Croxson PL. Facing the role of the amygdala in emotional information processing. Proc Natl Acad Sci U S A. 2012 Dec 26;109(52):21180-1. doi: 10.1073/pnas.1219167110. Epub 2012 Dec 14. PMID: 23243143; PMCID: PMC3535668. [↑](#endnote-ref-3)