

# ▼ SwiFT v2 논문 전략서 (v.3) ( $\Rightarrow$ v2.3에 N x T 법칙이 추가된 버전)

## Title: Breaking the Scaling Wall: A Blueprint for Compute-Optimal fMRI Foundation Models via Factorized Spatio-Temporal Attention

(스케일링의 벽을 넘어서: 팩터라이즈드 시공간 어텐션을 통한 fMRI 파운데이션 모델의 청사진)

### 1. Introduction: The Crisis & The Four Pillars

"We pushed the limits, hit the wall, and found a new path."

- **The Status Quo (현황):** LLM은 스케일링 법칙을 통해 비약적으로 발전했다. fMRI 분야도 BrainLM 등 대규모 모델이 등장하고 있다.
- **The Crisis (위기 - 핵심 추가):** "하지만 fMRI 모델도 무작정 키우면 성능이 오를까?" 우리는 이 질문에 답하기 위해 기존 SOTA인 4D Swin Transformer(SwiFT v1)를 **3B 파라미터까지 확장하는 실험을 감행했다.**
  - **결과:** 성능 향상은커녕 학습이 발산하고 붕괴(Collapse)하는 **'Scaling Wall'**을 세계 최초로 실증했다.
- **The Solution (해결):** 이 실패를 통해 우리는 4D 연산의 비효율성과 불안정성을 확인했고, 이를 해결하기 위해 **'Factorized Attention (SwiFT v2)'**을 고안했다.
- **Our Contributions (기여):**
  1. **Identification of Scaling Wall:** 4D 모델 확장의 한계를 실증적으로 규명.
  2. **The Engine (SwiFT v2):** 벽을 넘어 매끄럽게 확장되는 아키텍처 제안.
  3. **The Discovery (N x T Law):** 데이터 학습의 새로운 물리 법칙 발견.
  4. **The Proof (SOTA):** 임상 및 외부 벤치마크 압도적 1위 달성.

### 2. Results Part I: The Engine (Engineering Verification)

"Turning Failure into Success."

## [Figure 1: The Scaling Laws]

- **X축:** Compute (FLOPs) & Parameters / **Y축:** **Validation Loss (Reconstruction Error)**
- **(A) The Failure vs. Success:**
  - **Red Line (4D Baseline):** 3B에서 Loss가 줄지 않고 튀어 오르는 (Exploding/Oscillating) 모습 시각화.  $\rightarrow$  "이것이 우리가 발견한 벽이다."
  - **Blue Line (Ours - v2):** 9M  $\rightarrow$  66M  $\rightarrow$  264M로 갈수록 Loss가 예측 가능한 로그 스케일로 감소.  $\rightarrow$  "우리는 그 벽을 넘었다."
- **(B) Data Scalability:** 데이터 양(10%  $\rightarrow$  100%)에 따른 선형적 Loss 감소.
- **메시지:** "SwiFT v2는 단순한 모델 개선이 아니라, 확장 가능성(Scalability) 그 자체를 확보한 엔진이다."

## 3. Results Part II: The Discovery (Scientific Investigation)

┆ "Unveiling the  $N \times T$  Law." (Strategy 1.5 적용)

## [Figure 2 & 3: The Iso-Token Experiment]

- **실험 설계 (Real Subject Control):** 총 토큰 수 고정 (Iso-Token).
  - **Cond A (High N):** 1,000명  $\times$  300 Frames (UKB Style - 얇고 넓게)
  - **Cond B (Balanced):** 500명  $\times$  600 Frames
  - **Cond C (High T):** 250명  $\times$  1,200 Frames (HCP Style - 좁고 깊게)
- **지표 (Y축): Internal Linear Probing Accuracy** (HCP Held-out Set에서의 Subject ID / Sex Prediction).
  - *Loss 대신 Accuracy를 써서 '표상 능력'을 강조.*
- **결과:** Cond C (High T) > Cond A (High N).
- **메시지:**
  - "fMRI 파운데이션 모델의 핵심은 '개인차(N)'를 외우는 게 아니라, '뇌의 동역학(Dynamics, T)'을 이해하는 것이다."

- "250명만 깊게 봐도 1,000명을 훑어본 것보다 뇌를 더 잘 이해한다."

## 4. Results Part III: The Proof (Clinical Utility)

| "From Lab to Clinic."

### [Figure 4: The Grand Benchmark]

- **지표: External Clinical Task Accuracy / AUROC** (6개 데이터셋: AD, Depression, etc.).
- **비교:**
  - **Baseline:** BrainLM, Brain-JEPA, NeuroSTORM (기존 SOTA).
  - **Ours (Full Data):** 모든 데이터를 다 쓴 모델.
  - **Ours (N x T Optimized):** (옵션) 법칙에 따라 효율적으로 학습한 모델.
- **결과:** SwiFT v2가 모든 Task, 특히 데이터가 적은 **Low-shot Clinical Regime**에서 압도적 SOTA.
- **메시지:**
  - "Internal Representation Quality(Fig 2,3)가 좋으면, External Generalization(Fig 4)은 필연적으로 따라온다."
  - "우리의 법칙은 실험실용 이론이 아니라, **실전 승리 공식**이다."

## 5. Results Part IV: The Reason (Neuroscience Interpretability)

| "Physics-Informed Deep Learning."

### [Figure 5: Why did 4D fail and v2 succeed?]

- **(A) Spectral Analysis (PSD):**
  - 4D (3B) 모델: **\*\*White Noise ( $\beta \approx 0$ )\*\***에 과적합됨. (이래서 망했음)
  - SwiFT v2: **\*\*Scale-free Dynamics ( $\beta \approx 1$ )\*\***를 정확히 복원. (이래서 성공함)
- **(B) Hurst Exponent:** Long-range Memory 보존 능력 시각화.
- **메시지:** "SwiFT v2의 성공 비결은 **\*\*뇌 신호의 물리학(Physics)\*\***을 존중했기 때문이다. 4D 모델은 연산량은 높았지만 물리적 특성을 파괴했다."

## 6. Discussion: Guidelines & Conclusion

| "A New Standard for fMRI AI."

- **Scaling Insight:** 무작정 모델만 키우지 마라(Don't just scale up). 구조를 바꿔야 한다(Scale smart).
- **Data Policy:** 바이오뱅크는 이제 \$N\$ 경쟁을 멈추고 \*\*Deep Phenotyping (\$T\$)\*\*에 투자해야 한다. 우리의 \$N \times T\$ 법칙이 그 증거다.
- **Impact:** 우리는 실패(Scaling Wall)를 딛고, fMRI AI가 나아가야 할 \*\*명확한 청사진(Blueprint)\*\*을 제시했다.

### [작성 팁: 이 설계도의 장점]

#### 1. 기승전결의 완벽함:

- **Intro (위기):** "3B까지 키워봤는데 망했다." (독자 충격)
- **Result 1 (해결):** "구조 바꾸니까 해결되더라." (안도)
- **Result 2 (발견):** "알고 보니 데이터 법칙도 있더라." (흥미)
- **Result 3 (증명):** "그래서 다 이겼다." (확신)

#### 2. 약점의 무기화:

- 4D 모델의 실패를 감추지 않고 \*\*'Scaling Wall 규명'\*\*이라는 업적으로 포장했습니다.
- \$N\$이 적은 HCP 데이터의 한계를 \*\*'Deep T의 중요성 규명'\*\*이라는 강점으로 전환했습니다.

이제 이 설계도만 있으면, 에디터는 \*\*"이 논문은 fMRI AI 분야의 교과서가 될 것"\*\*이라고 평가할 것입니다. 건투를 빕니다!