

2026 KNS 원자력 인공지능 강습회

2주짜리 CFD를 10초 만에:

Neural Operator와 원자력 Surrogate의 최전선

ICC 제주도

2026.05.06. 수

한양대학교 원자력공학과 송민섭 (THINKLAB)

hysms@hanyang.ac.kr

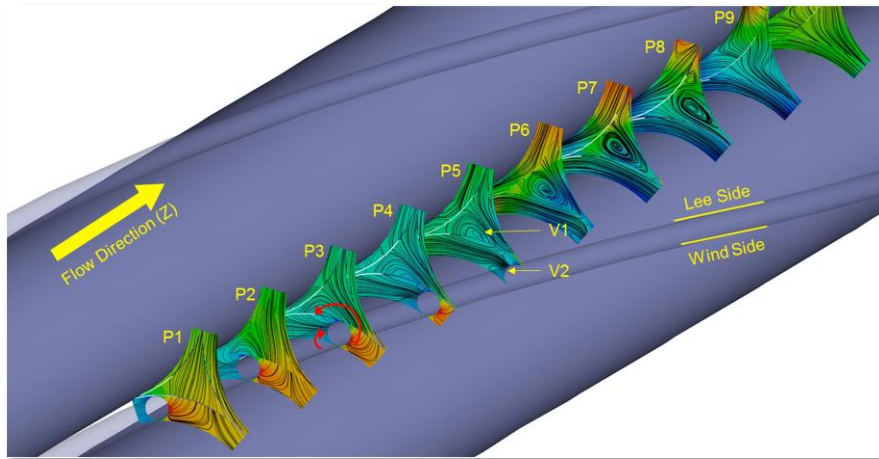
목 차

1. 왜 CFD Surrogate가 필요한가
2. Neural Operator의 개념과 FNO
3. 공학 분야 적용 사례
4. CFD 데이터셋 & THINK_BENCH
5. 검증, 난류 통계, 그리고 실험 데이터의 역할
6. 원자력 Surrogate의 향후 방향

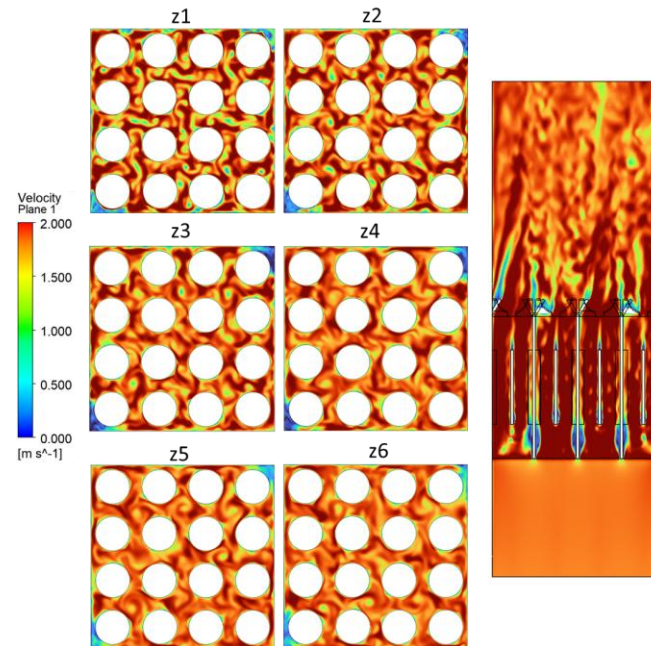
0. 원자력에서 CFD의 중요성은 왜 더 커지고 있는가?

■ 차세대 원자로로는 기존 경험식만으로 해석하기 어려운 3차원 복합 열유동 현상을 포함

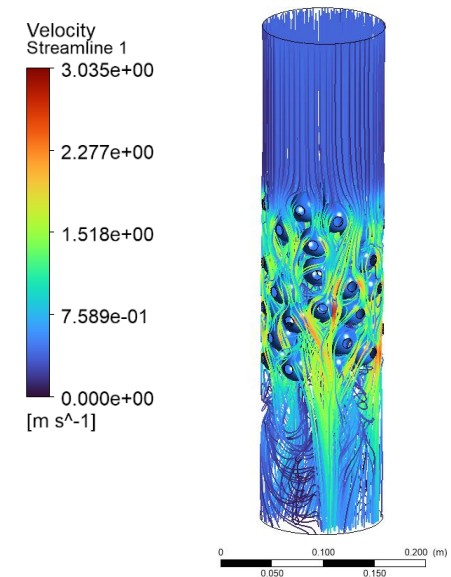
- SMR, 마이크로리액터, 비경수로형의 등장
 - 기존 대형 경수로 중심의 검증 경험을 그대로 적용하기 어려운 새로운 형상, 냉각재, 운전조건이 제시됨
 - Helical Coil Heat Exchanger, Wire-wrapped Bundle, Pebble Bed, PCHE 등 노심과 기기의 형상 복잡도가 증가
- 안전 여유도 평가의 정밀화가 요구됨
 - 평균 유량이나 1D 계통 해석 결과만으로는 국부적인 온도 상승, 혼합, 여유도 등을 충분히 평가하기 어려움
- CFD는 단순 열유동 해석이 아니라, Surrogate model, Uncertainty Quantification, Digital Twin을 학습시키는 고충실도 데이터 생성 수단이 됨



Sodium-cooled Fast Reactor
Cross-flow made by wire



Wake flow behind the Spacer Grid



Flow Mixing
Through Pebble Beds

1. 왜 CFD Surrogate가 필요한가?

■ CFD의 Surrogate 모델을 만들게 되면, 새로운 계산 비용이 0에 수렴하게 됨

- AI의 파급력은 생산 비용 (계산 비용)을 0에 가깝게 만든다는 것
 - 여전히 산업적으로 활용하는데 걸림돌이 되는 계산 비용 및 시간을 제로로 만들 수 있다면?
 - CFD는 격자, 난류 모델에 대한 민감도를 가짐
 - 그러나 CFD는 질량, 운동량, 에너지에 대한 보존을 맞추기 때문에, 물리적으로 타당한 속도 및 에너지 (온도) 분포를 도출할 수 있음. 이는 원자로 기기 설계, 안전 해석 중에 중요하게 작용



전통적 CFD

매 케이스마다 다시 계산

- 수억~수십억 격자
- 수일~수주 계산
- 파라미터가 바뀌면 처음부터 다시 진행해야 함

→ 설계 최적화·UQ·실시간 제어 어려움



Neural Operator Surrogate

훈련 후엔 새 케이스가 즉시 나옴

- 한 번 훈련 (며칠)
- 추론 0.01~1초 / 케이스
- $10^4 \sim 10^6$ 배 가속
- Mesh-independent (대부분)

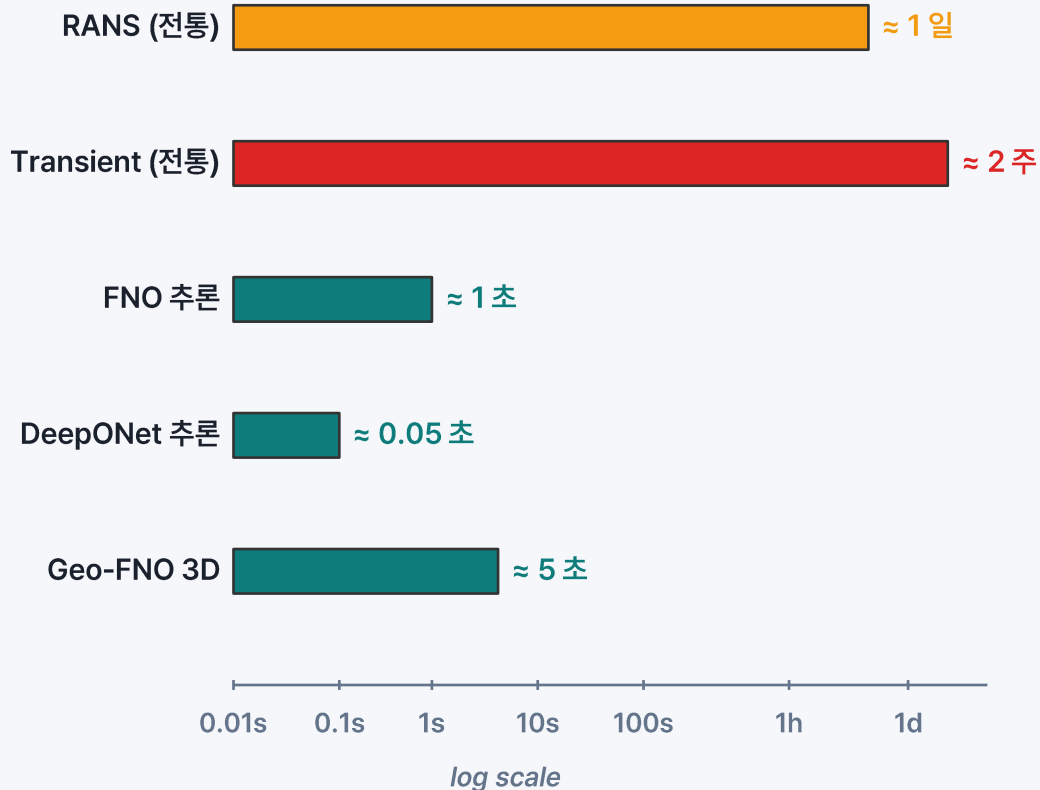
→ 설계 탐색·UQ·디지털트윈 가능

1. 왜 CFD Surrogate가 필요한가?

■ 통상 $10^4 \sim 10^6$ 배의 추론 가속화

- CFD Surrogate는 고비용 수치해석을 “매번 다시 푸는 문제”에서 “학습된 해를 즉시 예측하는 문제”로 전환

케이스당 계산시간 (log scale)



$10^4 \sim 10^6 \times$

전통 transient CFD 대비 추론 가속

\$\$ 비용 ↓

한 번 훈련해두면 추론은 거의 공짜

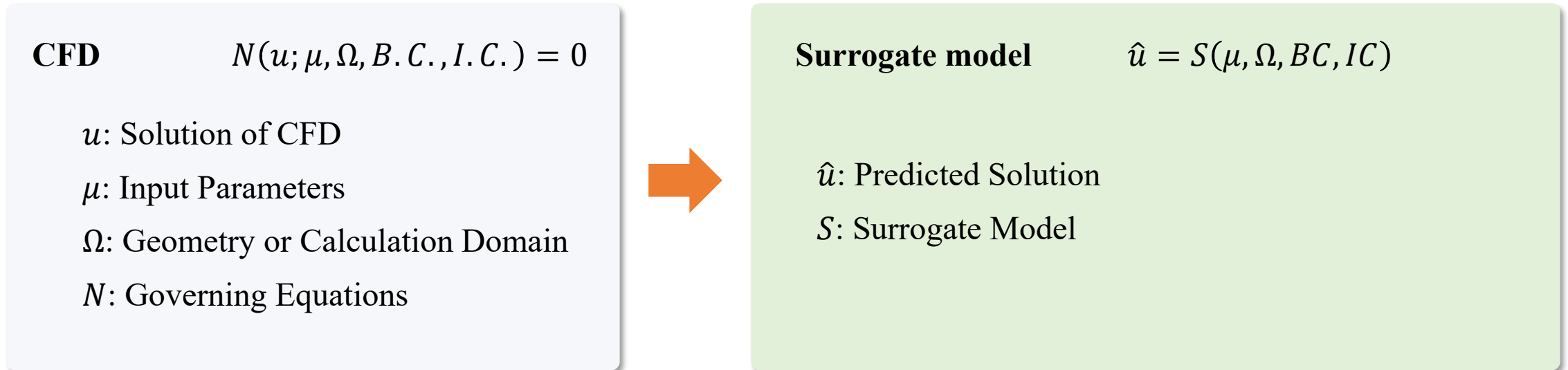
수백 케이스

보통 학습 데이터로 충분 (충분한 다양성 확보 시)

1. 왜 CFD Surrogate가 필요한가?

■ CFD의 Surrogate Model

- CFD Surrogate Model: "CFD 계산을 여러 번 진행하여 결과를 얻은 뒤, 이를 바탕으로 새로운 조건의 CFD 결과를 빠르게 예측하는 근사 모델"
 - Navier-Stokes 방정식과 에너지 방정식을 매번 직접 수치해석으로 풀지 않음.
 - 기존 CFD 계산 결과를 이용해 "입력 조건"과 "CFD 출력" 사이의 관계를 근사



*Surrogate model(대리 모델/대체 모델): 계산 비용이 높거나 복잡한 원본 시스템(블랙박스 모델)을 데이터 기반으로 근사(Approximation)하여 모방하는 보조 수학적 모델

1. 왜 CFD Surrogate가 필요한가?

■ CFD는 공간과 시간을 잘게 쪼개고, 매 점마다 반복적으로 풀어야 하기 때문에 느림

- 정상상태 (Steady-State) 계산은 하루 정도로 가능, 그러나 비정상상태 (Unsteady-State) 계산은 며칠 ~ 몇주의 시간을 요구

$$CFD\ Cost \sim N_{cell} \times N_{time} \times N_{iter} \times N_{var} \times C_{physics}$$



공간 해상도

벽면 근처는 더 작게

- 난류 경계층을 풀려면 $y^+ < 1$, 이는 1mm 이하 격자
- 속도 구배가 큰 영역에는 더 촘촘한 격자 필요
- LES인 경우 난류 와류 구조를 직접 해상하기 위해 더 작은 격자



시간 해상도

CFL · 빠른 현상

- 유속 10 m/s, 격자 1mm $\rightarrow \Delta t \approx 0.1ms$.
- 10초 시뮬레이션이면 10^5 스텝 필요.
- 제일 작은 현상에 맞춘 시간 스케일 필요



비선형 반복

수렴까지 여러 번

- Navier-Stokes는 대류항 때문에 비선형
- 압력-속도 결합 문제를 풀기 위해 SIMPLE 등의 알고리즘이 필요
- 각 반복마다 대규모 선형 방정식 시스템을 풀어야 함



물리모델 복잡도

다양한 물리적 현상

- 난류 모델 (k-e, SST k-w, RSM, LES 등)
- 열전달: 에너지 방정식, 공액열전달
- 다상유동: 계면 추적, 상간 운동량 교환, 표면장력
- 부력, 압축성, 화학반응, 상변화 등

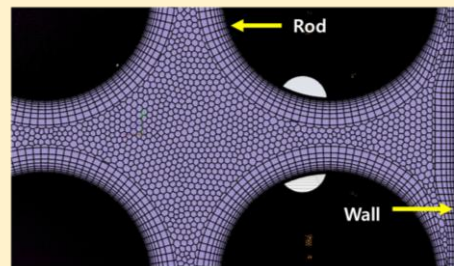
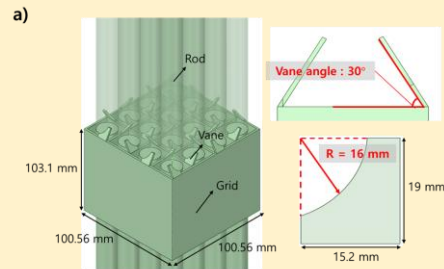
1.1 스케일 분석 - PWR 17x17 Fuel Assembly의 격자를 구축한다면?

■ PWR 17 x 17 연료 집합체는 '작은 유동 통로'가 수백 개 반복되는 다중 subchannel 구조

- 연료봉 사이의 좁은 gap, 벽면 경계층, spacer grid 주변 박리/재순환을 해상해야 함
 - 국부적으로 $y^+ < 1$ 수준의 벽면 근처 격자가 필요 (SST $k - \omega$)
- 단면에서 요구되는 격자 수가 10^6 수준이 되면, 축방향 적층만으로 전체 격자 수가 10^9 수준까지 증가
 - Full assembly 고해상도 CFD는 단일 워크 스테이션 계산 범위를 쉽게 초과

PWR 17x17 형상 정보 (예)

- 연료봉 외경 9.5 mm
- Pitch 12.6 mm
- 활성 연료 길이 3.66 m
- 연료봉 264개 + 안내관 25개
- Subchannel \approx 256개 (내부)
- Subchannel 한 변 \approx 3 mm



격자 수 견적 (RANS, $y^+ < 1$)

Subchannel 단면 셀 $\sim 10^3 \sim 10^4$
× 256 Subchannel = $2.6 \times 10^5 \sim 2.6 \times 10^6$ / slice
축방향: 3.66 m / 5 mm = 730 layer
전체 (full assembly) $\approx 1.8 \times 10^8 \sim 1.8 \times 10^9$ cell

full Assembly를 고충실도 계산하는 것은
현실적으로 불가능

1.2 스케일 분석 - 계산 방법 별 시간 비교

- CFD 계산 시간은 격자 수 뿐 아니라 물리시간, 반복 회수, 병렬 효율 등에 의해 결정됨
 - Steady 한 케이스에 며칠, Transient 한 케이스에 수주, 파라미터 Sweep은 수개월이 걸림

Subchannel 코드 (CTF/COBRA-TF) 1D 채널 기반 / 빠름	$10^4 \sim 10^5$ Control Volumes	수 초 ~ 수 분
RANS · 5×5 sub-bundle (1 m) 국부 상세 설계 검토 가능 (spacer grid, mixing vane, 압력 강하 평가 등)	5~50M cells	수 시간 ~ 1일
RANS · Full assembly (3.66 m) 한 case 분석 수준 (전체 집합체 분포 확인은 가능하나 반복 설계에는 부담)	100~500M cells	수 일
LES · Sub-bundle (transient) 고충실도 physics 연구용 (난류 구조, mixing, validation data 생성)	50~200M cells	1~3주
Full-core LES 현재 일반 설계 계산으로는 비현실적	$10^9 \sim 10^{10}$ cells	초고성능 HPC 필요 *100개 코어 기준

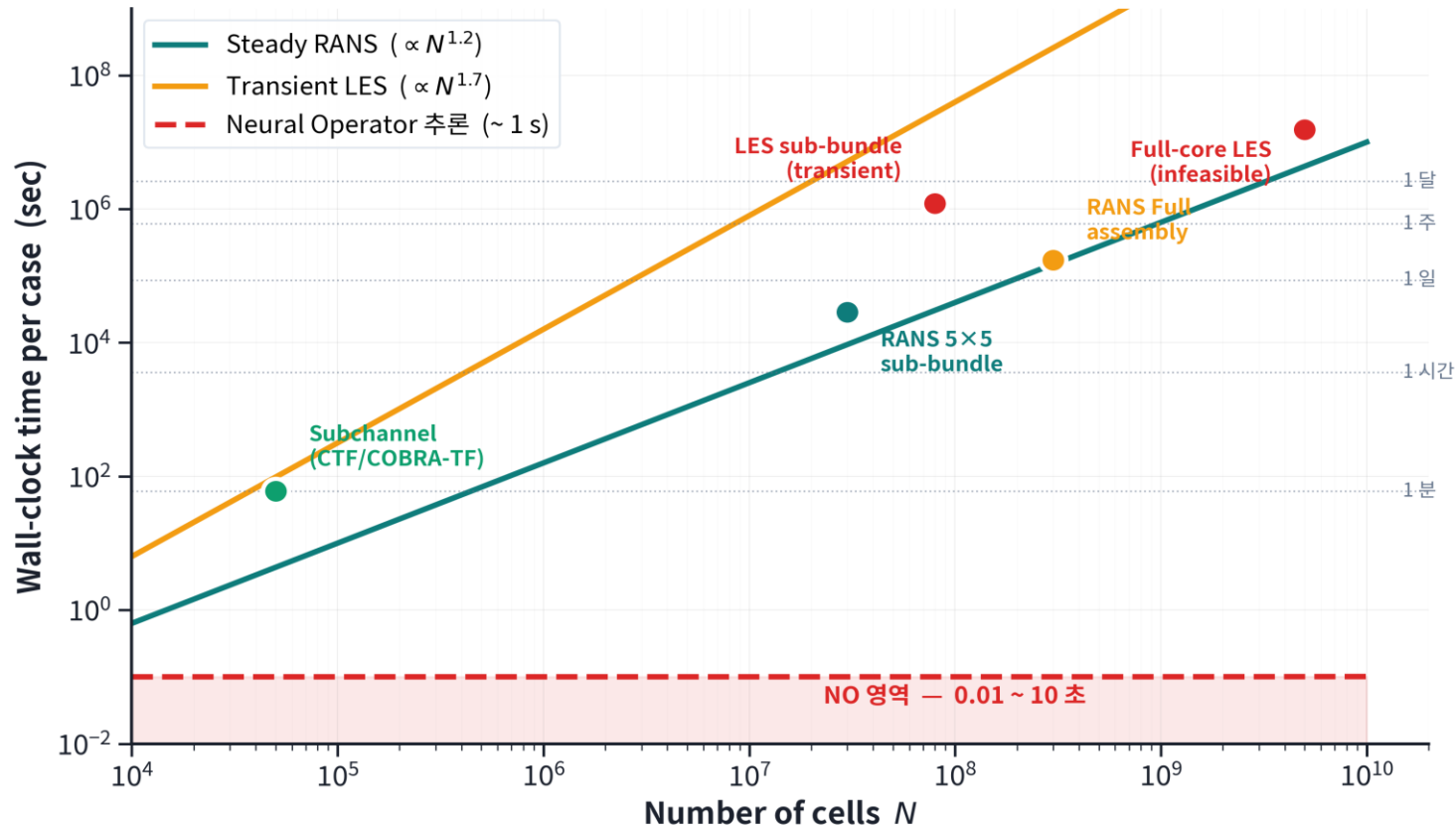
설계 최적화, UQ, 안전여유도 평가에는 수십~수백 케이스가 필요 → 모든 Case를 고충실도 CFD로 수행하는 것은 **사실상 불가능**

1.3 CFD 해석 비용과 Neural Operator 추론 비용의 스케일링 비교

■ CFD 해석 비용과 Neural Operator 추론 비용의 스케일링 비교

- CFD 해석 비용은 격자 수와 시간해상도 증가에 따라 지수적으로 증가하는 반면, 학습된 Neural Operator의 추론 비용은 입력 규모 증가에 대해 상대적으로 완만하게 증가
- 단일 정상상태 해석은 수시간-수일, 비정상 고충실도 해석 (LES)은 수주 이상이 소요될 수 있으며, Parameter Sweep에서는 계산 비용이 수개월 규모로 누적됨

CFD 계산 비용 스케일링 vs Neural Operator 추론



※ 약 1000 core 가정. Multigrid · HPC 최적화에 따라 ± 0.5 order 변동.

2. Neural Operator의 개념과 FNO

■ Neural Operator는 PDE의 해 연산자 (solution operator)를 학습하는 모델

- 일반 신경망은 유한 차원 벡터를 입력 받아 유한 차원 벡터를 출력
 - Neural Operator는 함수 자체를 입력 받아 함수 자체를 출력

Neural Network

벡터 → 벡터

- 입력: 고정 길이 벡터 / 픽셀
- 출력: 고정 길이 벡터 / 픽셀
- 예) Re=5000일 때 격자점 1024개의 속도
- → 격자가 1024 → 2048 되면 못 씀

Neural Operator

함수 → 함수

- 입력: 초기조건 / 경계조건 / 계수 "필드"
- 출력: 해(field) — 임의 위치에서 평가 가능
- 예) 입력열속분포 → 온도장($T(x,y,z)$)
- → 격자 해상도가 바뀌어도 그대로 평가
- → 학습 안 본 IC/BC도 일반화

2. Neural Operator의 개념과 FNO

■ CFD Surrogate 모델은 기존에도 존재하였음

- POD-ROM, 회귀분석, 일반 신경망 기반 Surrogate 모델이 활용되었으나, 적용 범위와 일반화 능력에는 명확한 한계가 존재



기존 ROM / Surrogate

- POD (Proper Orthogonal Decomposition)
[주성분 모드를 추출하여 저차원 공간에서 해를 근사]
- 회귀분석
[입력 파라미터와 출력 물리량 사이의 보간, 회귀 관계 구성]
- MLP / CNN
[고정된 벡터 또는 이미지 형태의 입력과 출력을 직접 매핑]



약점들

- 격자 해상도 변화에 취약
[모델은 대개 학습된 입,출력 차원에 고정됨]
- 기하 변화에 대한 일반화가 제한적
[새로운 형상에서 재학습 또는 별도 모델 구축 필요]
- 물리장 정체에 대한 외삽 성능이 약함
[훈련 분포를 벗어난 조건에서는 예측 신뢰도가 급격히 저하될 수 있음]

2. Neural Operator의 개념과 FNO – 범용적인 Neural Operator를 만들 수 있는가?

- Universal-NO는 사실상 모든 유동 문제의 해 공간을 학습해야 하기 때문에 현재로서는 어려움
 - 유동의 해는 Geometry에 의해 결정됨 -> 형상이 바뀔 때 해의 구조가 어떻게 바뀌는지 배워야 함
 - 경계 조건은 문제의 성격을 바꿈 (적용 물리식의 변경, 경계 조건의 수식 등)
 - 유동이 난류가 되면 여러 길이 스케일과 시간 스케일이 동시에 존재함 -> 다양한 스케일에 대해 모두 맞춰야 함
 - CFD는 격자의 노드에 해가 저장되는 반면, 실제 해는 연속 공간 상에 있음
 - 보존 법칙을 정확히 만족해야 함. -> Neural Operator에서는 이 지점이 약화될 수 있음
 - Universal-NO가 학습해야 할 공간은 사실상 무한
 - (원자력에서는) 결정론적 판단을 위해 정확한 결정값을 제시해야 함 (peak wall Temperature 등)

원자력분야에서는 Universal-NO가 아니라, 적용 범위가 명확하고 검증 가능한 특정 도메인의 surrogate AI 모델이 더 현실적이고 실용적

2. Neural Operator의 개념과 FNO

■ Fourier Neural Operator: 가장 유명한 Neural Operator (Zongyi Li et al, 2021, ICLR)

- FNO: 공간 영역에서의 복잡한 컨볼루션 연산을 푸리에 변환을 통해 주파수 영역으로 변환한 뒤, 저주파 성분에 가중치를 곱하고 역변환
 - 직접 PDE를 푸는 것보다 $10^3 \sim 10^6$ 배 빨라짐
 - 글로벌한 패턴 (Eddy, 진동)을 빠르게 작음 / 고주파, 충격파 등의 해석에 약함. 복잡 형상에 추가 기법이 필요
 - Resolution Invariance: 학습 시 해상도와 무관하게 추론이 가능하며, 낮은 해상도 데이터로 학습 후 높은 해상도에서 기능을 수행할 수 있음

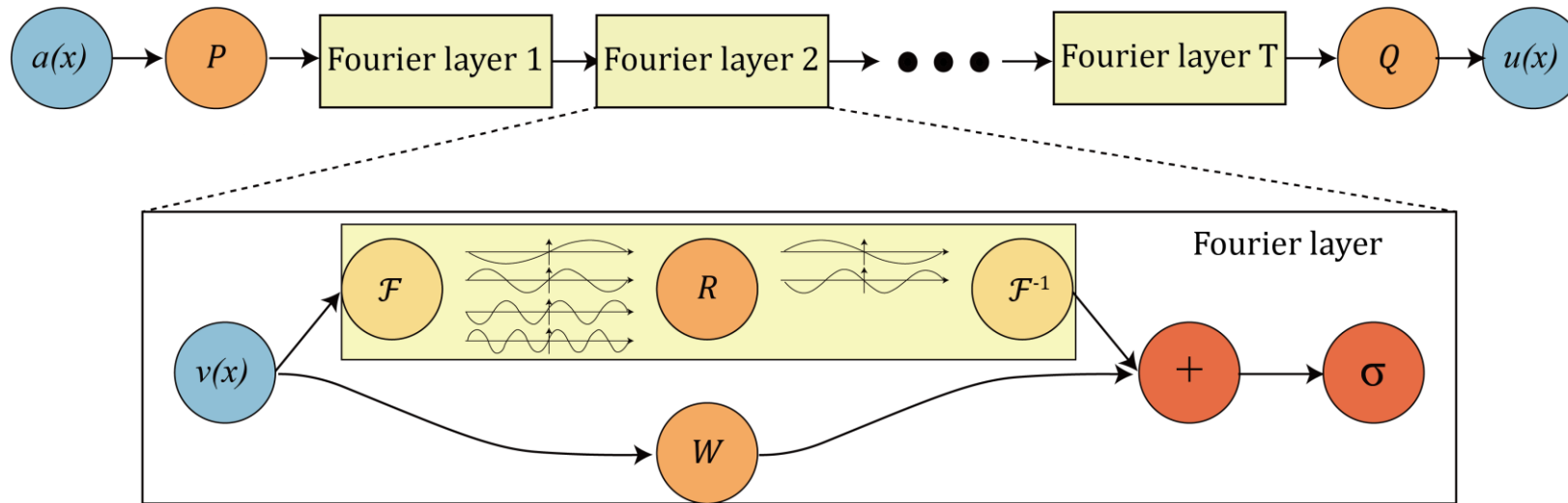
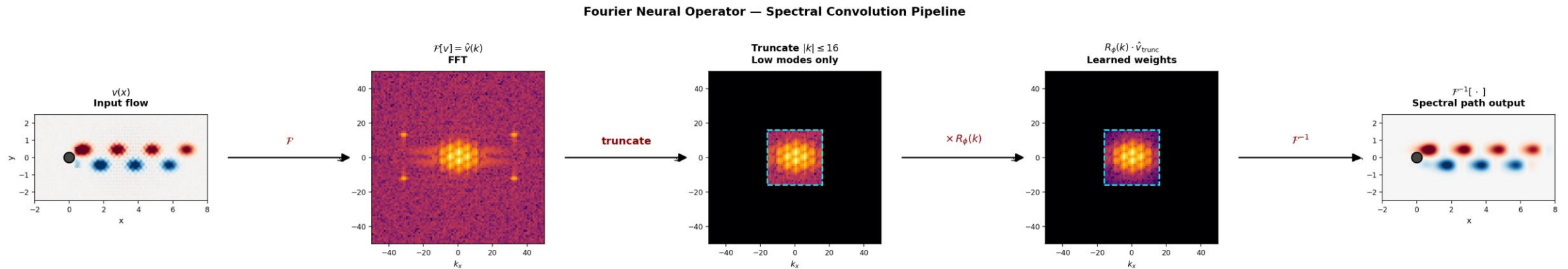


Figure 2: **top:** The architecture of the neural operators; **bottom:** Fourier layer.

2. Neural Operator의 개념과 FNO

■ Fourier Neural Operator: 가장 유명한 Neural Operator (Zongyi Li et al, 2021, ICLR)

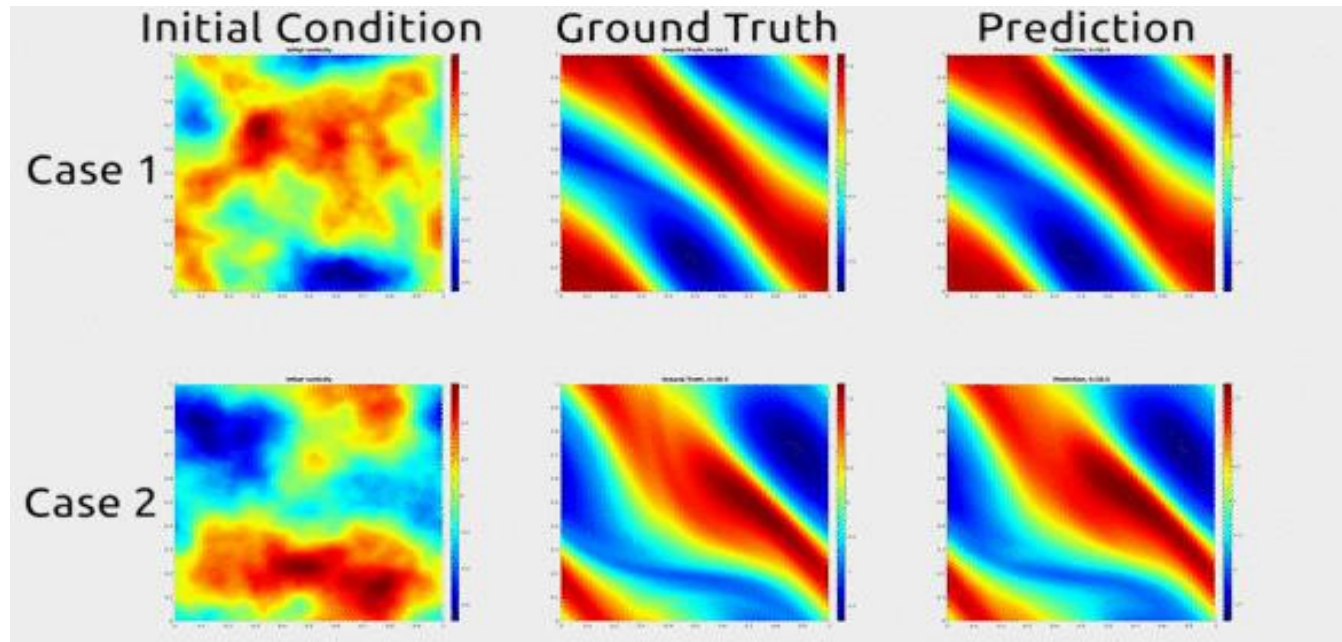
- FNO: 공간 영역에서의 복잡한 컨볼루션 연산을 푸리에 변환을 통해 주파수 영역으로 변환한 뒤, 저주파 성분에 가중치를 곱하고 역변환
 - 직접 PDE를 푸는 것보다 $10^3 \sim 10^6$ 배 빨라짐
 - 글로벌한 패턴 (Eddy, 진동)을 빠르게 작음 / 고주파, 충격파 등의 해석에 약함. 복잡 형상에 추가 기법이 필요
 - Resolution Invariance: 학습 시 해상도와 무관하게 추론이 가능하며, 낮은 해상도 데이터로 학습 후 높은 해상도에서 기능을 수행할 수 있음



2. Neural Operator의 개념과 FNO

■ Fourier Neural Operator: 가장 유명한 Neural Operator (Zongyi Li et al, 2021, ICLR)

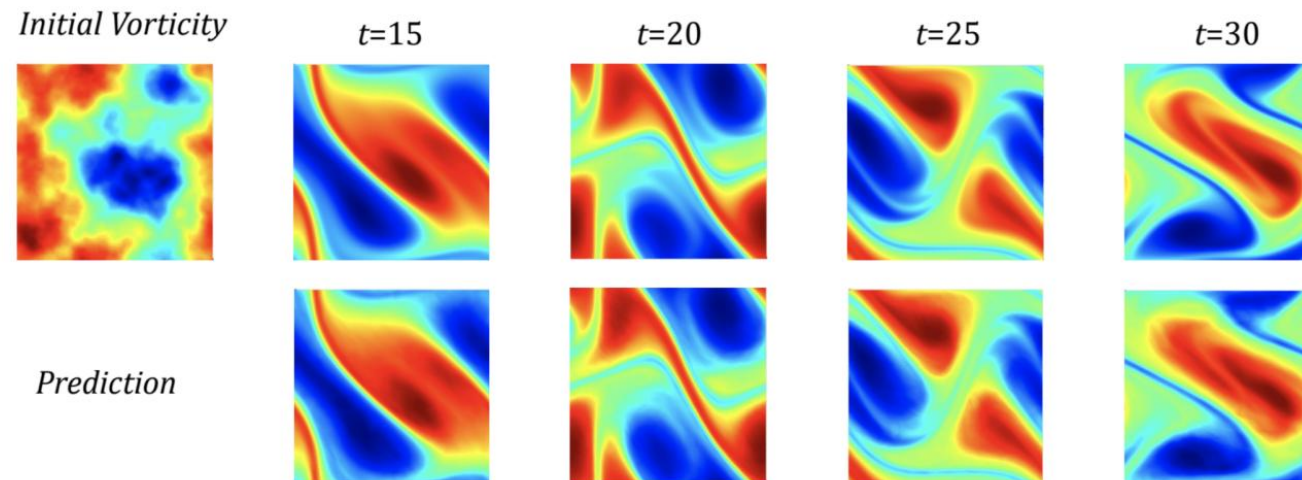
- FNO: 공간 영역에서의 복잡한 컨볼루션 연산을 푸리에 변환을 통해 주파수 영역으로 변환한 뒤, 저주파 성분에 가중치를 곱하고 역변환
 - 직접 PDE를 푸는 것보다 $10^3 \sim 10^6$ 배 빨라짐
 - 글로벌한 패턴 (Eddy, 진동)을 빠르게 작음 / 고주파, 충격파 등의 해석에 약함. 복잡 형상에 추가 기법이 필요
 - Resolution Invariance: 학습 시 해상도와 무관하게 추론이 가능하며, 낮은 해상도 데이터로 학습 후 높은 해상도에서 기능을 수행할 수 있음



2. Neural Operator의 개념과 FNO

■ Fourier Neural Operator: 가장 유명한 Neural Operator (Zongyi Li et al, 2021, ICLR)

- FNO: 공간 영역에서의 복잡한 컨볼루션 연산을 푸리에 변환을 통해 주파수 영역으로 변환한 뒤, 저주파 성분에 가중치를 곱하고 역변환
 - 직접 PDE를 푸는 것보다 $10^3 \sim 10^6$ 배 빨라짐
 - 글로벌한 패턴 (Eddy, 진동)을 빠르게 작음 / 고주파, 충격파 등의 해석에 약함. 복잡 형상에 추가 기법이 필요
 - Resolution Invariance: 학습 시 해상도와 무관하게 추론이 가능하며, 낮은 해상도 데이터로 학습 후 높은 해상도에서 기능을 수행할 수 있음



Zero-shot super-resolution: Navier-Stokes Equation with viscosity $\nu = 1e-4$; Ground truth on top and prediction on bottom; trained on $64 \times 64 \times 20$ dataset; evaluated on $256 \times 256 \times 80$ (see Section 5.4).

Figure 1: **top:** The architecture of the Fourier layer; **bottom:** Example flow from Navier-Stokes.

2. Neural Operator의 개념과 FNO

■ 문제 특성에 따라 선택되는 Neural Operator 계열

- Neural Operator는 격자 구조, 형상 복잡도, 데이터 규모, 예측 목적에 따라 선택해야 함

DeepONet

Branch + Trunk Network 기반 Operator Learning

강점

- 입력 함수와 출력 위치를 분리하여 학습
-> 임의 위치에서 해 예측 가능
- sensor/query 기반 데이터에 적합
-> 비정형 샘플링 데이터 활용 가능
- 비교적 단순한 PDE 및 저차원 물리에서 안정적

약점

- 3D 고해상도 CFD에서는 branch/trunk 조합이 커져 계산량 증가
- 복잡 형상, 난류, 강한 비선형 유동에서는 일반화 어려움

Graph Neural Operator

Geometry-aware / Graph-based Operator Learning

강점

- 비정형 mesh, 복잡 형상, 실제 CFD grid 입력 가능
- geometry embedding 또는 graph 구조를 활용하여 형상 변화에 대응 가능
- CAD 기반 복잡 형상 문제에 적합

약점

- graph 구성, neighborhood 정의, geometry encoding에 따라 성능이 민감
- FNO보다 학습 구조가 복잡하고, 계산 비용이 증가할 수 있음
- 다양한 형상 일반화를 위해 대량의 고품질 CFD 데이터가 필요함

Transformer-NO

Attention-based Multi-physics / Multi-condition Operator Learning

강점

- Attention mechanism으로 장거리 상관관계와 전역 구조 학습 가능
- 여러 물리량, 경계조건, 입력 형태를 동시에 처리하기 유리
- 대규모 데이터 기반 다물리, 다조건 surrogate로 확장 가능

약점

- 고해상도 3D field에서는 memory와 계산 비용이 매우 큼
- 데이터 부족 시 과적합과 학습 불안정성이 발생하기 쉬움
- 모델 구조가 복잡하여 물리적 해석 가능성이 낮아질 수 있음

3. 공학 분야 적용 사례

■ 열유체, 노물리, 핵융합 등으로 적용 범위가 확장되고 있음

- Neural Operator는 아직 범용 CFD 해석을 대체하지는 못하지만, 특정 형상, 조건, 물리 영역에서 고속 surrogate model 및 digital twin 구성 요소로 검증되고 있음

유체역학

Fourier Neural Operator 기반 Surrogate Model을 이용하여 3차원 난류 채널 유동의 장시간 거동을 예측

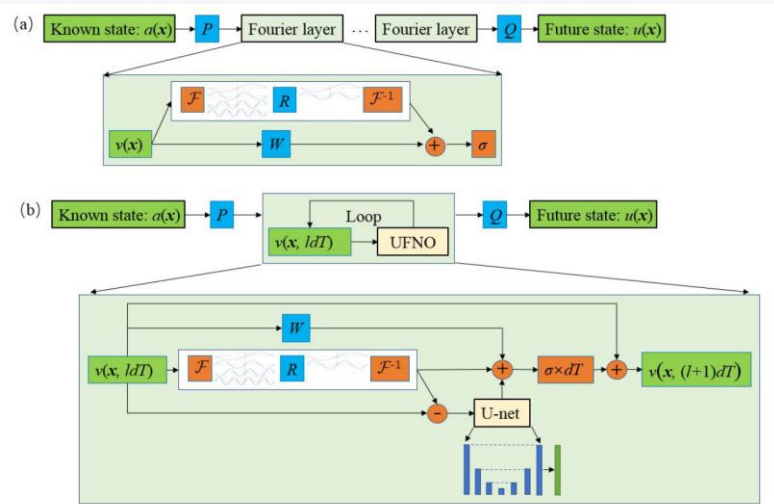


FIG. 1. The configurations of (a) Fourier neural operator (FNO), and (b) implicit U-Net enhanced Fourier neural operator (IUFNO).

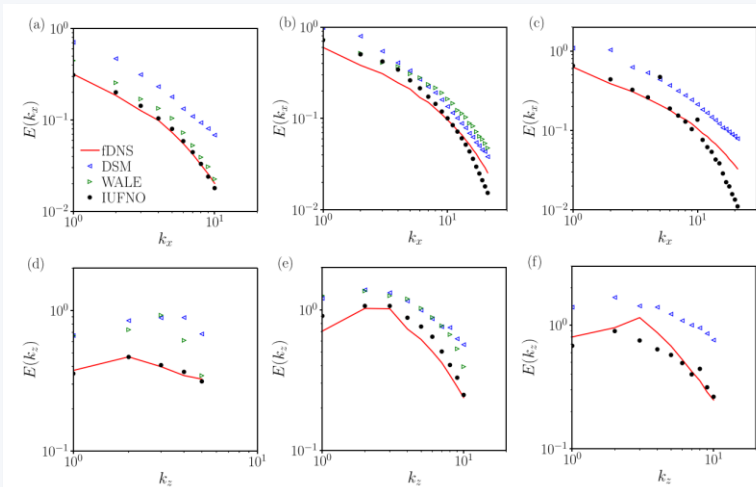


FIG. 13. The streamwise energy spectrum at various Reynolds numbers: (a) streamwise spectrum, $Re_\tau \approx 180$; (b) streamwise spectrum, $Re_\tau \approx 395$; (c) streamwise spectrum, $Re_\tau \approx 590$; (d) spanwise spectrum, $Re_\tau \approx 180$; (e) spanwise spectrum, $Re_\tau \approx 395$; (f) spanwise spectrum, $Re_\tau \approx 590$.

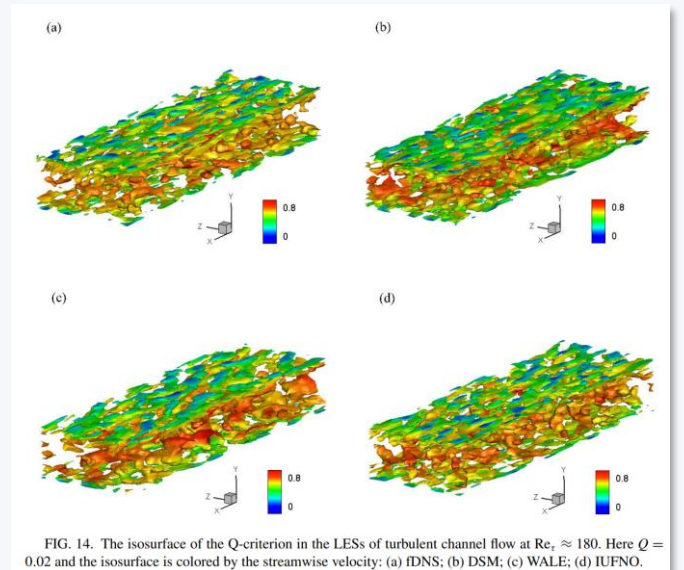


FIG. 14. The isosurface of the Q-criterion in the LESs of turbulent channel flow at $Re_\tau \approx 180$. Here $Q = 0.02$ and the isosurface is colored by the streamwise velocity: (a) DNS; (b) DSM; (c) WALE; (d) IUFNO.

Ref: Prediction of turbulent channel flow using Fourier neural operator-based machine-learning strategy, Y. Wang et al., Phys. Rev. Fluids 9, 084604 (2024)

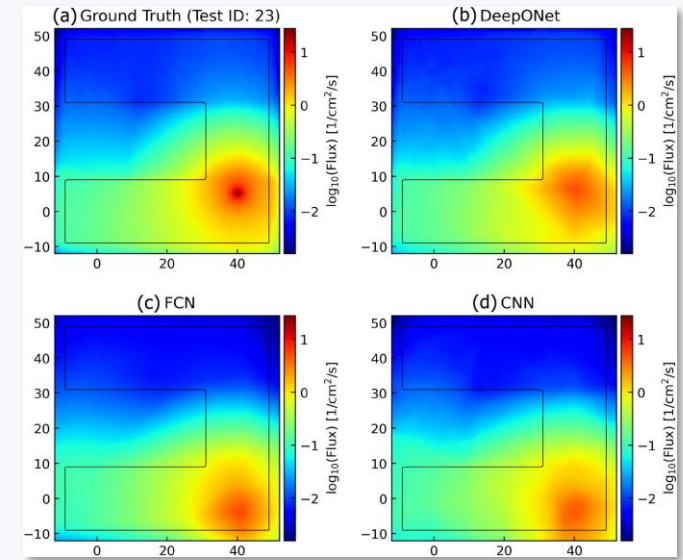
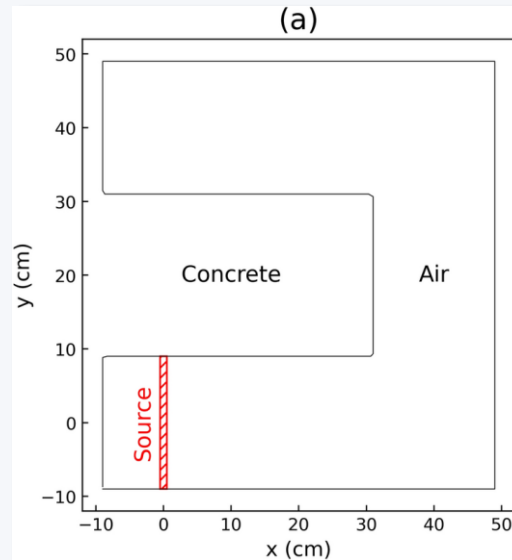
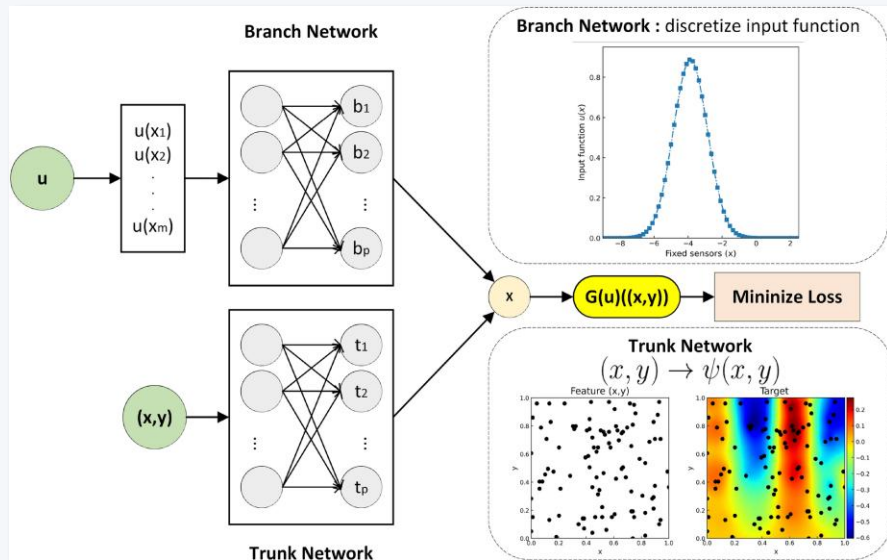
3. 공학 분야 적용 사례

■ 열유체, 노물리, 핵융합 등으로 적용 범위가 확장되고 있음

- Neural Operator는 아직 범용 CFD 해석을 대체하지는 못하지만, 특정 형상, 조건, 물리 영역에서 고속 surrogate model 및 digital twin 구성 요소로 검증되고 있음

노물리

DeepONet을 이용하여 중성자 소스 분포 함수 -> 공간 플럭스 분포 함수의 연산자를 학습함으로써, 고비용 수송해석을 surrogate model로 대체



Ref: Deep neural operator-driven real-time inference to enable digital twin solutions for nuclear energy systems, Kobayashi et al, 2024, scientific reports, <https://www.nature.com/articles/s41598-024-51984-x>

3. 공학 분야 적용 사례

■ 열유체, 노물리, 핵융합 등으로 적용 범위가 확장되고 있음

- Neural Operator는 아직 범용 CFD 해석을 대체하지는 못하지만, 특정 형상, 조건, 물리 영역에서 고속 surrogate model 및 digital twin 구성 요소로 검증되고 있음

열수력 (V-SMR)

DeepONet을 이용하여 Karman Vortex, Plenum 유동에 대한 Surrogate Model제작 (포항공대, 한양대)

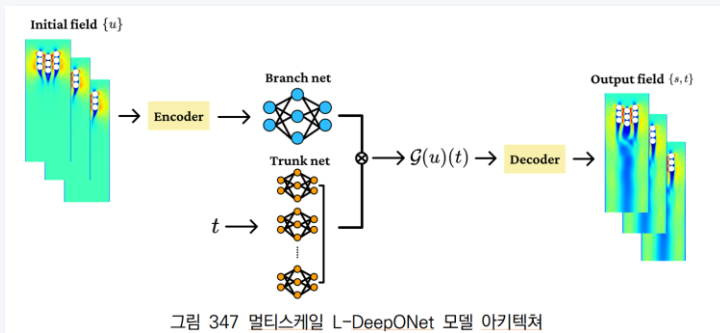
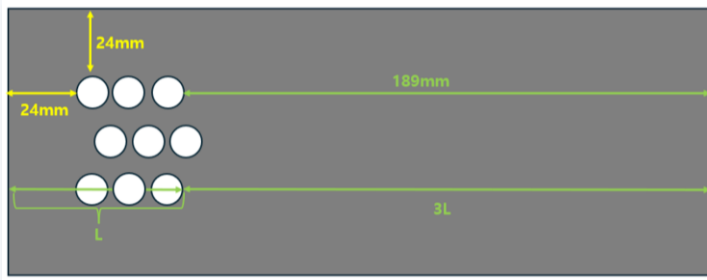


그림 347 멀티스케일 L-DeepONet 모델 아키텍처

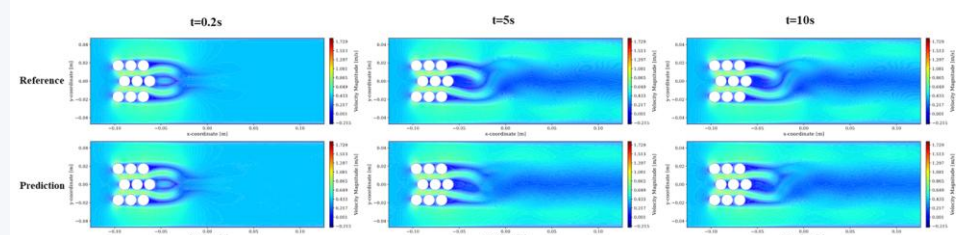


그림 354 내삽 조건에서 멀티스케일 L-DeepONet (MLP-AE) 모델의 속력장 예측 결과

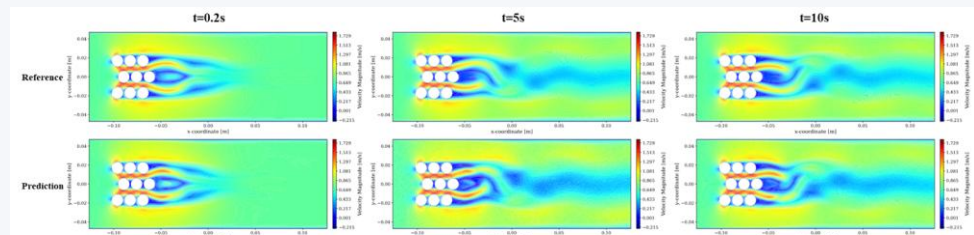


그림 356 외삽 조건에서 멀티스케일 L-DeepONet (MLP-AE) 모델의 속력장 예측 결과

Ref: SMR 가상원자로 플랫폼 개발 전략연구단

4. CFD 데이터셋 & THINK_BENCH

■ CFD 데이터가 확보되면 Neural Operator 기반 Surrogate Model을 직접 구축할 수 있음

- 핵심은 범용 모델을 바로 만드는 것이 아니라, 공개 데이터와 자체 CFD 데이터를 결합해 특정 설계 도메인에 특화된 모델을 단계적으로 개발하는 것
 - 공개 데이터셋은 베이스라인 학습과 코드 검증에 활용, 자체 CFD 데이터는 실제 관심 형상과 운전조건을 반영하는데 사용

Step 1

공개 데이터로 베이스라인 구축

- 공개된 데이터셋으로 FNO 학습
- 전처리 학습- 평가 파이프라인을 확보

Step 2

자체 CFD 데이터 생성

- 관심 형상 100~500개 이상 케이스 생성
- LHS 기반 파라미터 샘플링

Step 3

도메인 특화 surrogate 구축

- FNO/GINO/PINO 등 적용
- 속도, 압력, 온도장 예측 및 물리 제약 추가

Step 4

검증 및 활용

- 범위 밖 데이터에 대해 테스트하고, 실험 데이터와 비교
- 설계 최적화, UQ, 디지털 트윈에 활용

4. CFD 데이터셋 & THINK_BENCH

■ Neural Operator 기반 Surrogate Model을 만들 시 필수로 고려해야 하는 사항

- CFD 기반 Surrogate 모델은 데이터의 양보다 다양성, 일관성, 물리성, 검증성 등이 중요

1

데이터 다양성 & 충분성

- 수백~수천 케이스 확보
- 단순히 케이스 수만 늘리는 것은 일반화 성능을 보장하지 않음
- 형상, 운전조건, 물성, 경계 조건의 파라미터 공간을 LHS 기반으로 균형있게 샘플링해야 함 (Latin Hyperbolic Sampling)

2

표현 일관성

- 모든 케이스는 동일한 도메인 정의, 정규화 방식, 변수 정의, 좌표계, 단위계를 가져야 함.
- 격자가 서로 다른 경우에는 공통 격자 보간 또는 mesh-free 표현 적용이 필요

3

물리 제약 / Physics-informed

- Surrogate 모델이 질량, 운동량, 에너지 보존을 심각하게 위반하면 공학적으로 사용하기 어려움
- 손실 함수에 PDE residual, 경계조건, 보존 오차를 포함하는 방식이 필요

4

검증 데이터셋 분리

- 훈련 분포 내의 테스트뿐 아니라, 훈련 범위 밖 조건에 대한 검증 필요
- 훈련 데이터에만 잘 맞는 모델은 실제 설계용 surrogate로 사용하기 어려움

5

불확실도 정량화

- 예측값과 신뢰도를 함께 제시해야 함.
- 원자력에서는 예측값도 중요하지만, 오차 범위와 보수성이 더 중요



원자력 분야 적용 시 추가 고려 사항

규제 검증 가능성 / 보수성 마진 / 안전 등급 / 설명가능성 / 불확실성 정량화

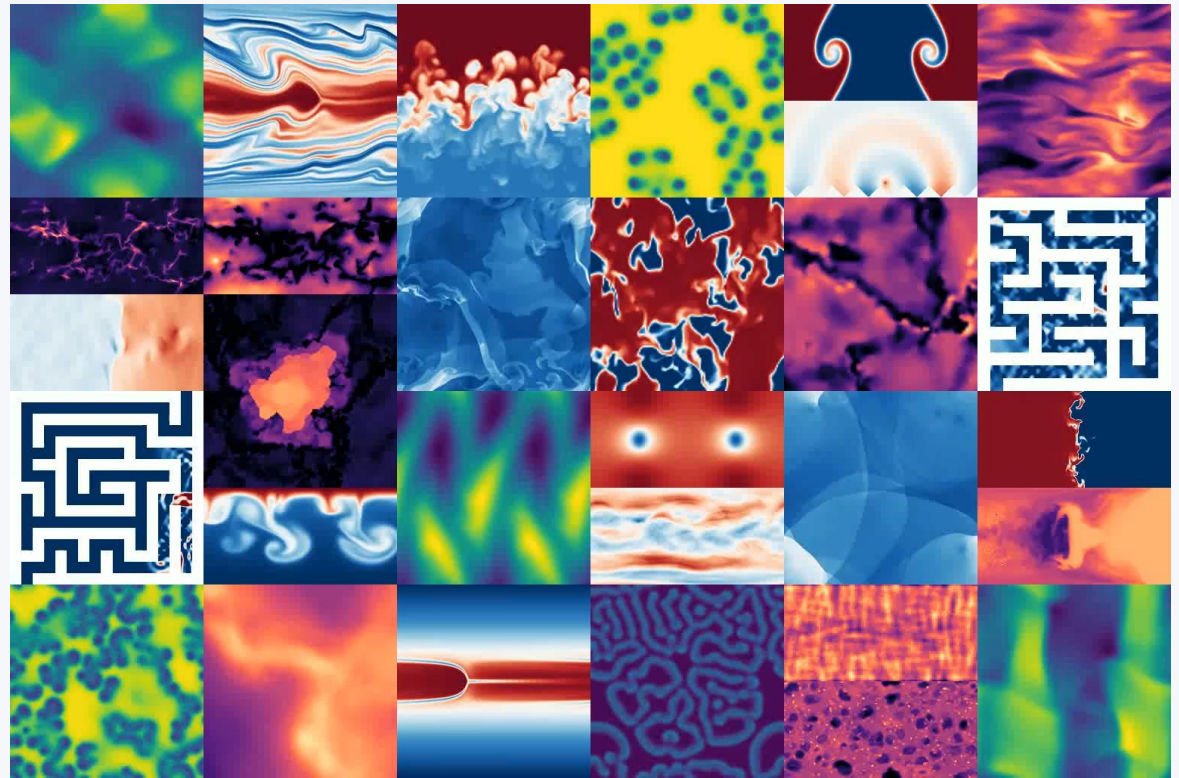
4. CFD 데이터셋 & THINK_BENCH

- 현재 다양한 연구 그룹에서 CFD/시뮬레이션 데이터셋을 공개하고 있음
 - 데이터셋을 선택 시 단순한 데이터 크기보다 "학습 가능한 물리 정보"가 충분한지를 확인해야 함.
 - 문제 유형, 물리 현상의 다양성, 해상도, 메타데이터, 검증 가능성 확인 필요

The Well *Copyright: Polymathic AI (NeurIPS 2024)*

- 다양한 물리 현상에 대한 수치 시뮬레이션을 포함하는 대규모 머신러닝 데이터셋
- 유체 흐름, 자기유체역학, 초신성 폭발 등
- 16개 물리 시뮬레이션 도메인 통합. 약 15TB. Foundation model 학습용.

polymathic-ai.org/the_well



4. CFD 데이터셋 & THINK_BENCH

■ 현재 다양한 연구 그룹에서 CFD/시뮬레이션 데이터셋을 공개하고 있음

- 데이터셋을 선택 시 단순한 데이터 크기보다 "학습 가능한 물리 정보"가 충분한지를 확인해야 함.

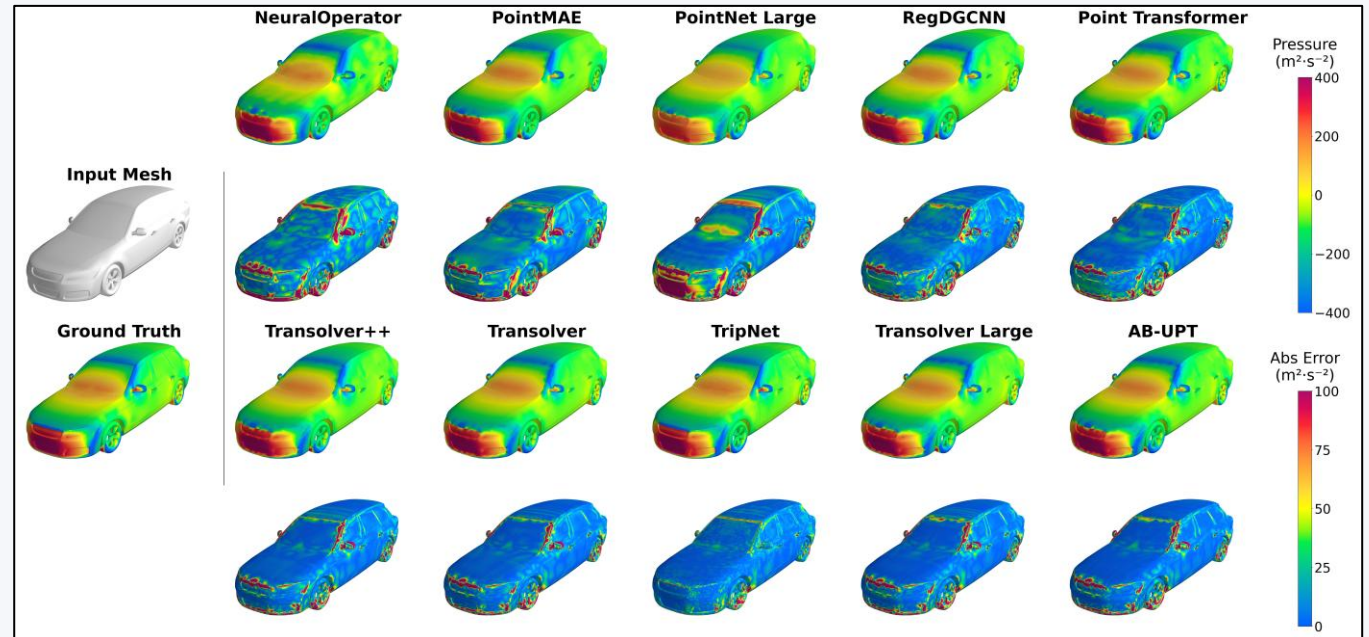
➢ 문제 유형, 물리 현상의 다양성, 해상도, 메타데이터, 검증 가능성 확인 필요

Car Bench *Copyright: MIT, Toyota Research Institute*

- 대규모 3D 자동차 공기역학에 특화된 종합 벤치마크 도구
- 8,000개 이상의 고정밀 CFD 시뮬레이션 데이터 제공 (DrivAerNet++)
- 압력, 속도, 벽면 전단 응력에 대한 체적 데이터와 표면 압력 및 전단 응력 데이터 포함

<https://mohamedelrefaie.github.io/CarBench/>

"Car Bench의 경우 Transformer 기반 예측이 더 뛰어남"



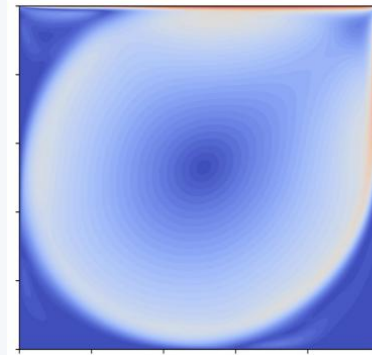
4. CFD 데이터셋 & THINK_BENCH

- 현재 다양한 연구 그룹에서 CFD/시뮬레이션 데이터셋을 공개하고 있음
 - 데이터셋을 선택 시 단순한 데이터 크기보다 "학습 가능한 물리 정보"가 충분한지를 확인해야 함.
 - 문제 유형, 물리 현상의 다양성, 해상도, 메타데이터, 검증 가능성 확인 필요

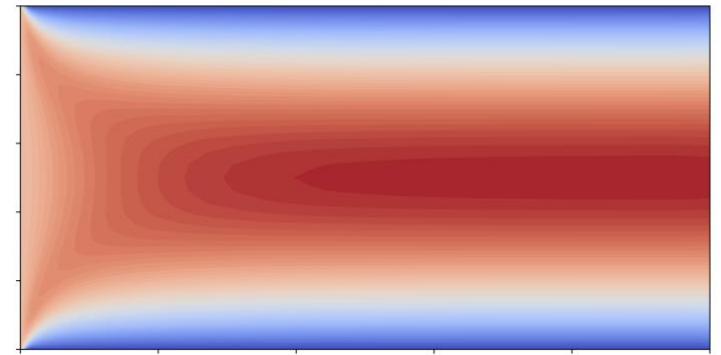
CFDBench *Copyright: Tsinghua University*

- 다양한 경계 조건, 물리적 특성 및 영역 형상을 가진 유체 역학 분야 머신러닝 방법 평가를 위한 최초 대규모 벤치마크 데이터
- 네가지 고전적 CFD 문제 포함
- 보간 데이터 (~13.4GB), 원본 데이터 (~460GB)

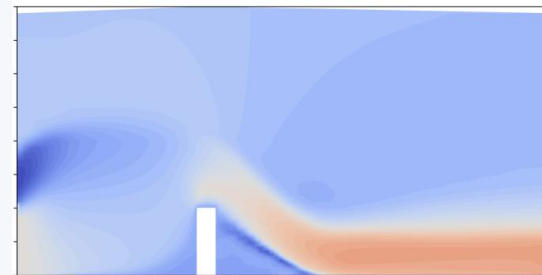
github.com/luo-yining/CFDBench



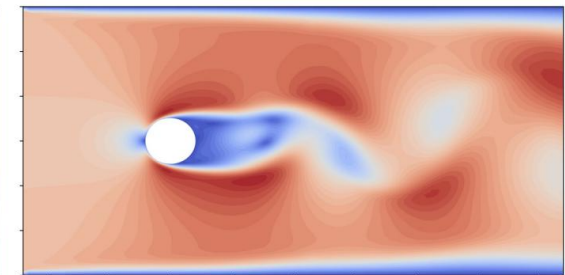
Cavity Flow



Tube Flow



Dam Flow



Cylinder Flow

4. CFD 데이터셋 & THINK_BENCH

■ 현재 다양한 연구 그룹에서 CFD/시뮬레이션 데이터셋을 공개하고 있음

- 데이터셋을 선택 시 단순한 데이터 크기보다 "학습 가능한 물리 정보"가 충분한지를 확인해야 함.

➢ 문제 유형, 물리 현상의 다양성, 해상도, 메타데이터, 검증 가능성 확인 필요

CFDBench *Copyright: Tsinghua University*

- 다양한 경계 조건, 물리적 특성 및 영역 형상을 가진 유체 역학 분야 머신러닝 방법 평가를 위한 최초 대규모 벤치마크 데이터
- 네가지 고전적 CFD 문제 포함
- 보관 데이터 (~13.4GB), 원본 데이터 (~460GB)

github.com/luo-yining/CFDBench

"유동의 종류에 따라 최적의 예측 모델이 달라짐"

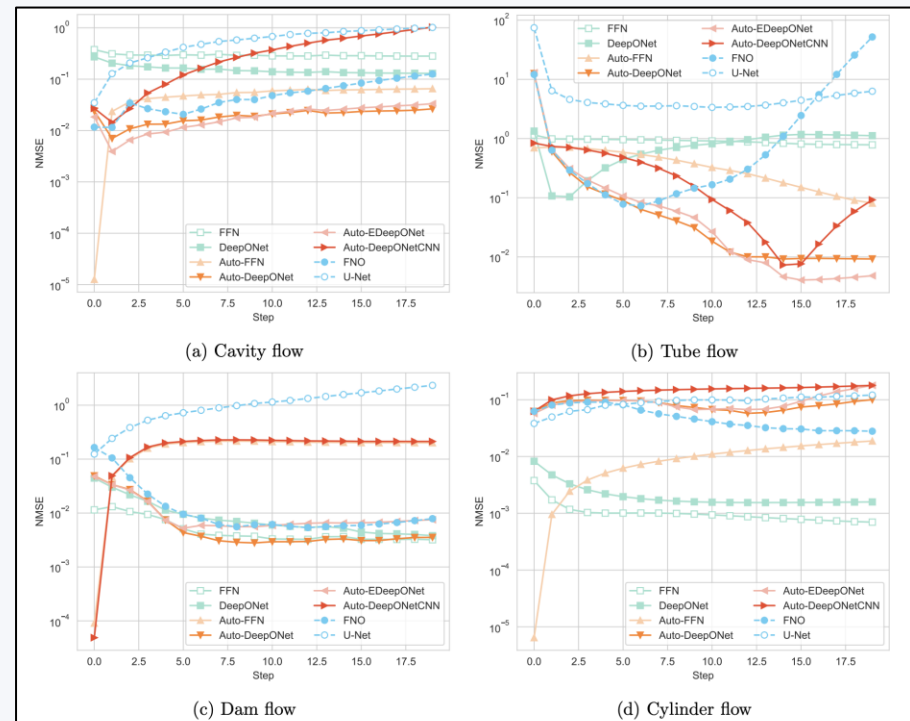


Figure 8: The error of different baselines as a function of time steps, given only the operating parameters Ω and the initial conditions. The y -axis is on a logarithmic scale. ResNet is not included because its error is too high and including its line would make the figure less intelligible.

4. CFD 데이터셋 & THINK_BENCH

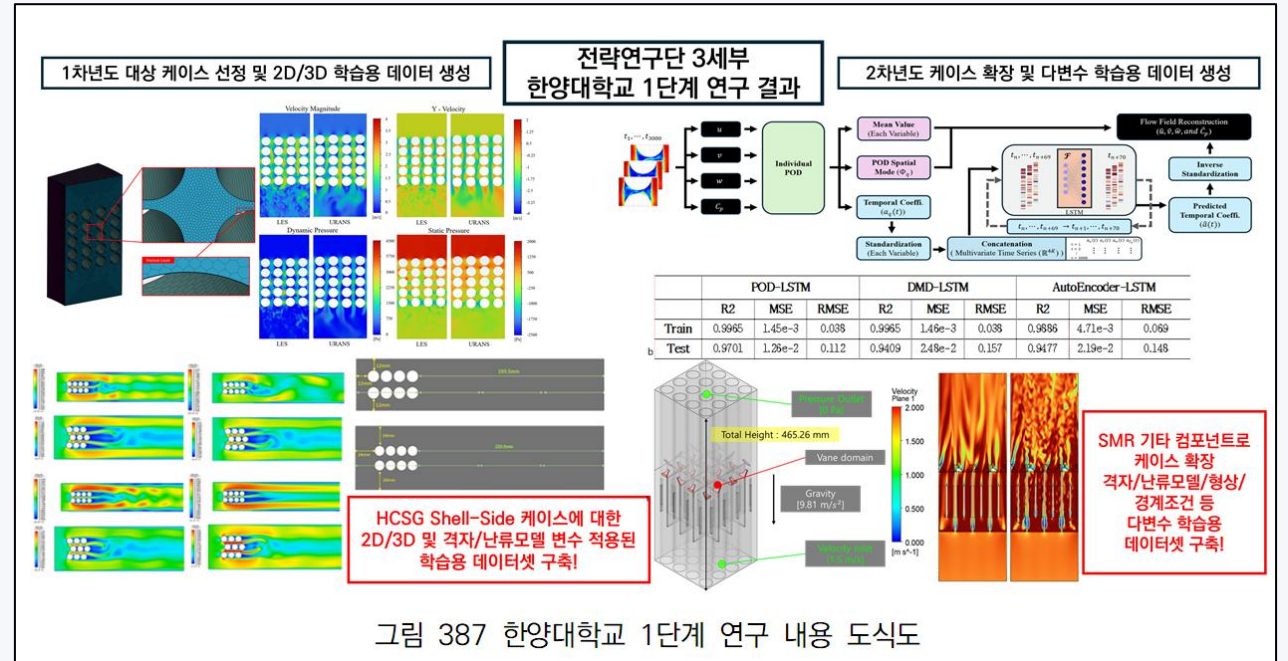
■ SMR 도메인 데이터 생성: 가상원자로 플랫폼 개발 전략연구단 3세부

- 한양대학교에서 데이터 개발 -> 포항공대, 카이스트
 - SMART 원자로 계통, Spacer Grid, Lower Plenum, Rod Bundle

THINK_Bench Copyright: V-SMR (2026)

- SMR 내부 대표 유동 현상에 대한 케이스 선정 및 데이터셋 구축
- 헬릭스 (Helical Coil Heat Exchanger), 유동 혼합 (Spacer Grid, Lower Plenum) 등 선정
- 차원 (2D/3D), 격자 및 난류 모델 변경 등 여러 조건을 고려하여 250개 데이터셋 생성
- 해석 자동화 프로세스 기반 마련

가상원자로 플랫폼 개발 전략연구단 과제 참조



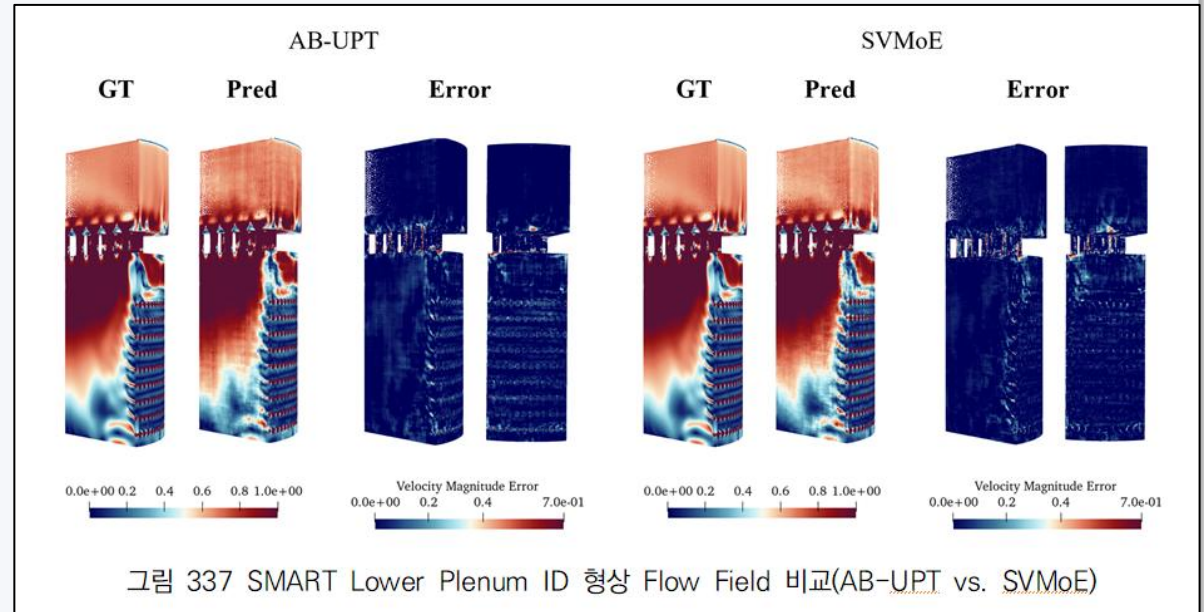
4. CFD 데이터셋 & THINK_BENCH

- SMR 도메인 데이터 생성: 가상원자로 플랫폼 개발 전략연구단 3세부
 - 한양대학교에서 데이터 개발 -> 포항공대, 카이스트
 - SMART 원자로 계통, Spacer Grid, Lower Plenum, Rod Bundle

THINK_Bench *Copyright: V-SMR (2026)*

- SMR 내부 대표 유동 현상에 대한 케이스 선정 및 데이터셋 구축
- 횡류 (Helical Coil Heat Exchanger), 유동 혼합 (Spacer Grid, Lower Plenum) 등 선정
- 차원 (2D/3D), 격자 및 난류 모델 변경 등 여러 조건을 고려하여 250개 데이터셋 생성
- 해석 자동화 프로세스 기반 마련

가상원자로 플랫폼 개발 전략연구단 과제 참조



5. 검증, 난류 통계, 그리고 실험 데이터의 역할

■ 아직 원자력 안전 계통에 바로 쓰기 어려운 이유

● 작은 난류 구조 예측의 한계

- Surrogate 모델은 큰 유동 패턴을 잘 예측
- 작은 와류, 국부 혼합 등의 고주파 변동 성분은 놓칠 수 있음
- 원자력 열수력에서는 이러한 변동 성분이 압력 강하, 열전달, 국부 온도 상승과 연결됨
- 따라서 평균 오차뿐 아니라 난류 통계, 스펙트럼 까지 검증해야 함

● 학습 범위 밖 조건에서 정확도 저하

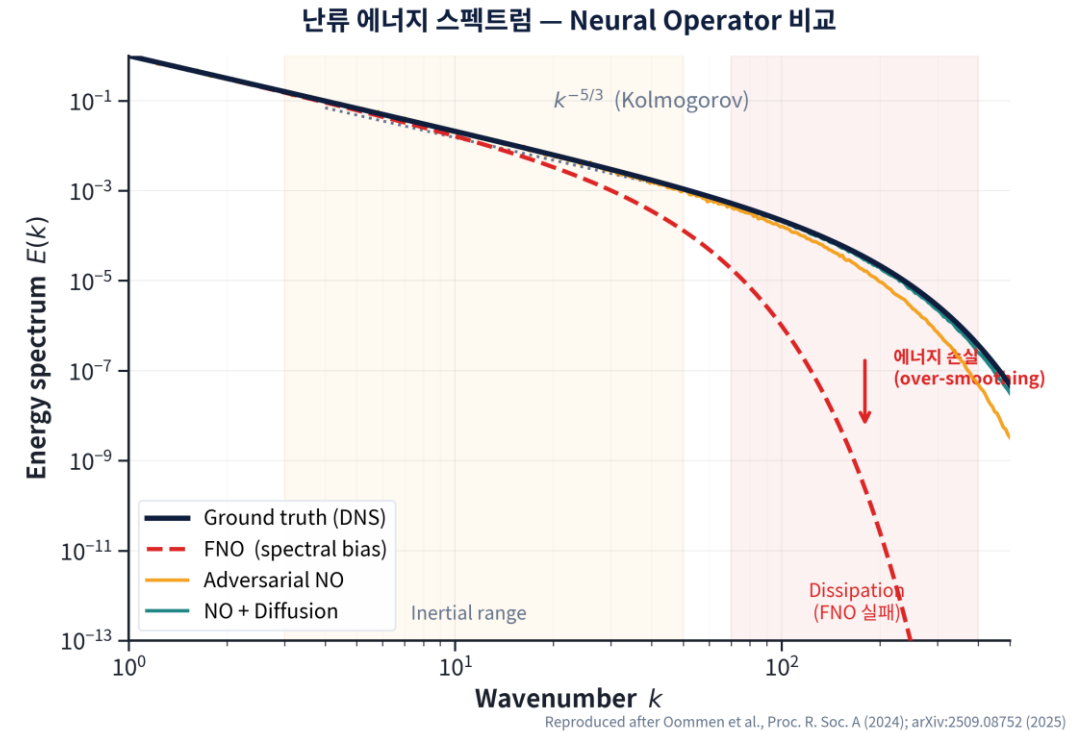
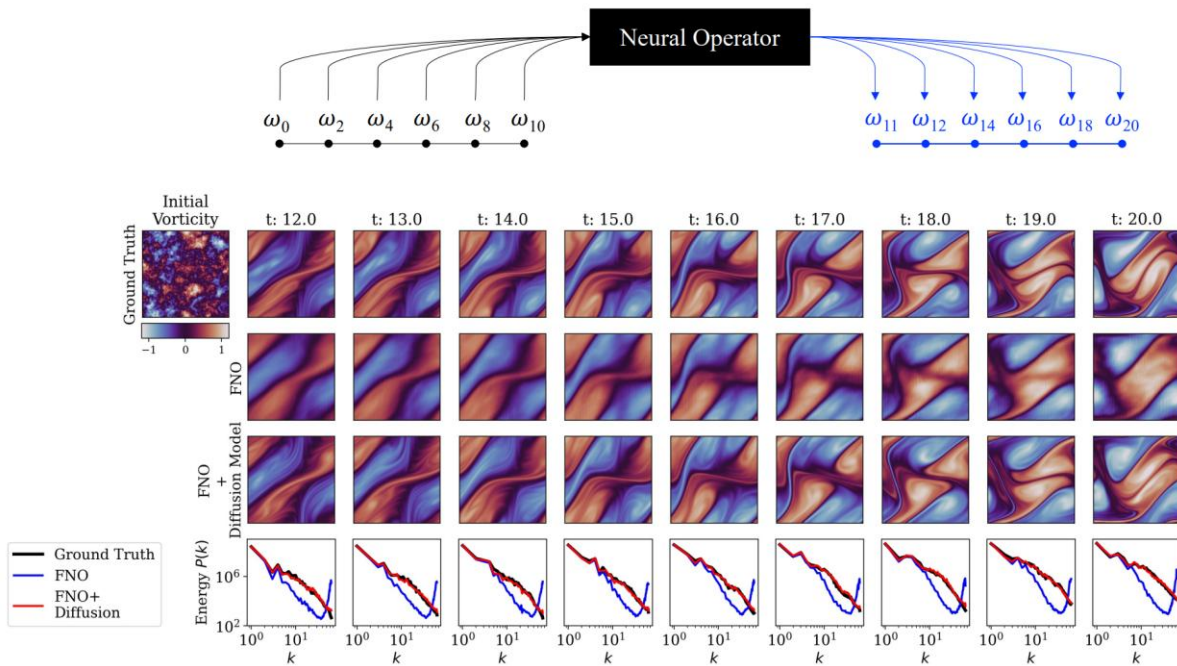
- Surrogate Model은 기본적으로 학습 데이터안의 패턴을 바탕으로 예측
- 학습되지 않은 형상, 조건에서는 정확도가 떨어질 수 있음
- 원자력 안전 해석은 정상상태가 아닌 과도 상태도 포함
- 따라서 모델이 어느 범위까지 신뢰가 가능한지 명확히 제시해야 함

5. 검증, 난류 통계, 그리고 실험 데이터의 역할

■ 아직 원자력 안전 계통에 바로 쓰기 어려운 이유

● 작은 난류 구조 예측의 한계의 예

➢ 순수 FNO를 이용한 Surrogate Model은 난류의 고주파 성분에 대한 예측이 떨어짐



스펙트럼: FNO는 dissipation range에서 수 order 작은 에너지 — 작은 와류 정보 손실.

5. 검증, 난류 통계, 그리고 실험 데이터의 역할

■ 아직 원자력 안전 계통에 바로 쓰기 어려운 이유

● 보존 법칙 위반 가능성

- 순수 데이터 기반 모델은 질량, 운동량, 에너지 보존을 만족하지 않음
- 원자력 안전해석에서 이런 물리적 불일치가 안전 판단에 영향을 줄 수 있음
- Physics-informed model이 필요

● Surrogate Model이 복잡해질수록 Black-Box에 가까워짐

- 이는 규제 검토에 바로 쓰기 어렵게 만드는 요소가 됨
- 모델 정확도, 적용 범위, 한계, 불확실성을 함께 제시해야 함

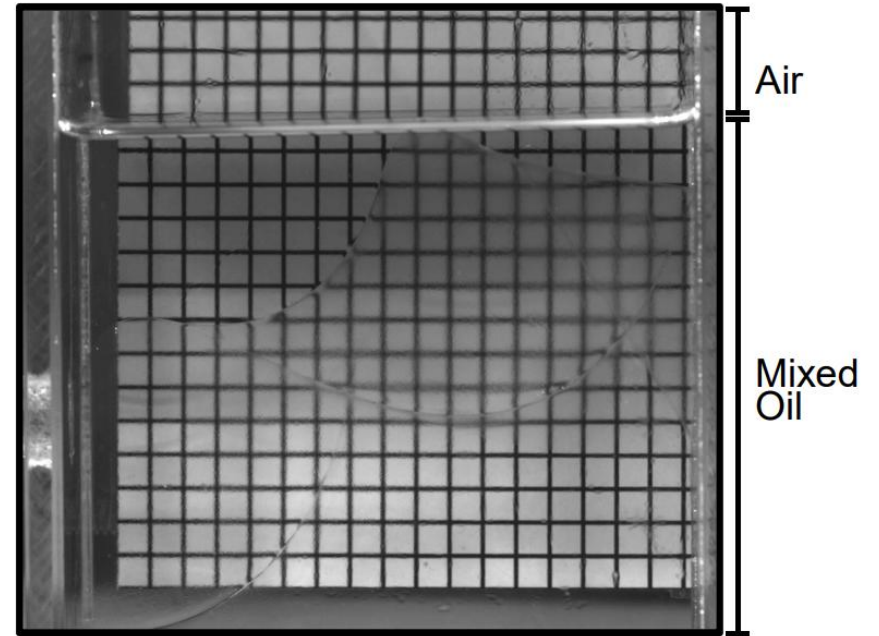
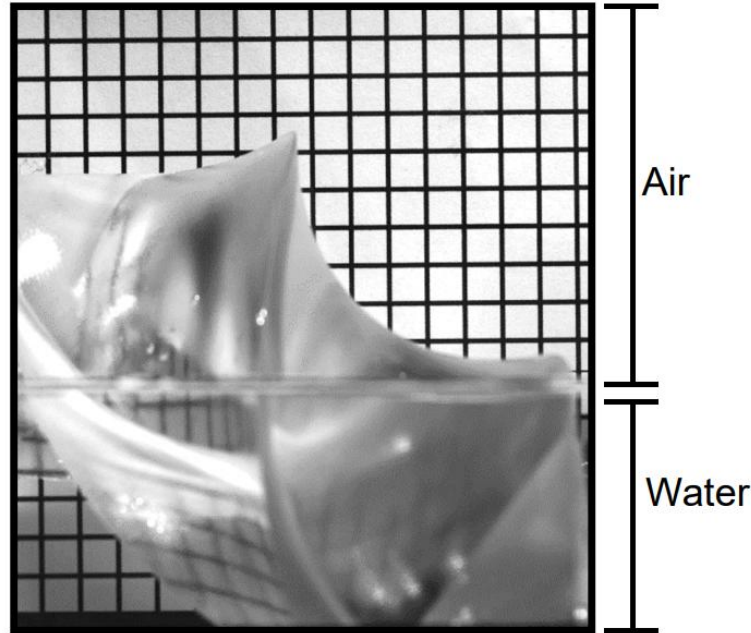
● 고품질 데이터 생산이 필요

- 신뢰성 높은 CFD 데이터 및 실험 데이터가 필요
- Surrogate model을 학습시키기 위한 CFD 데이터 생성의 가속화가 필요함 (GPU 병렬화 등)

5. 검증, 난류 통계, 그리고 실험 데이터의 역할

■ CFD 검증을 위한 유동 가시화 데이터의 필요성

- Particle Image Velocimetry, Magnetic Resonance Velocimetry 등은 CFD의 난류 모델 민감도, 격자 민감성 등에 대한 검증 데이터를 제공할 수 있음



3D Printed Twisted Prism in Working Fluid

[Ref. Song et al, NED (2015)]

그림. 3D 프린팅 모델과 고굴절률 용액 간 굴절률 일치 개념

5. 검증, 난류 통계, 그리고 실험 데이터의 역할

■ CFD 검증을 위한 유동 가시화 데이터의 필요성

- Particle Image Velocimetry, Magnetic Resonance Velocimetry 등은 CFD의 난류 모델 민감도, 격자 민감성 등에 대한 검증 데이터를 제공할 수 있음

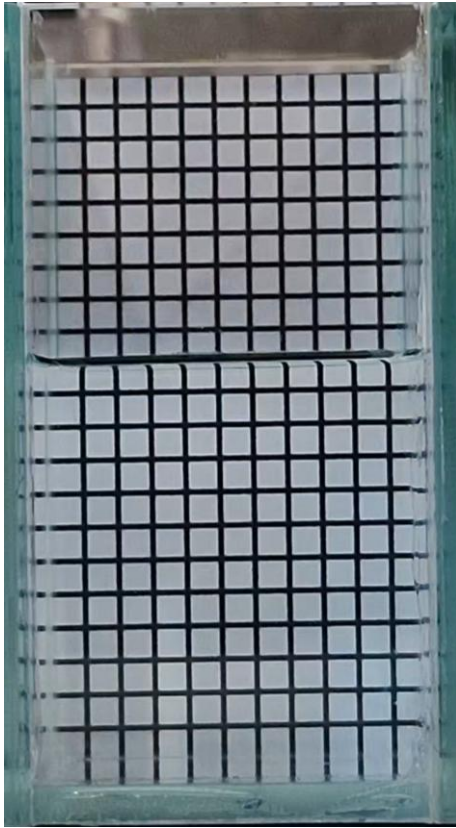


그림 1. Pebble bed 정면 view



그림 2. Twisted prism 정면 view



그림 3. Wire wrapped rod 정면 view



그림 4. Tube 정면 view

5. 검증, 난류 통계, 그리고 실험 데이터의 역할

■ CFD 검증을 위한 유동 가시화 데이터의 필요성

- Particle Image Velocimetry, Magnetic Resonance Velocimetry 등은 CFD의 난류 모델 민감도, 격자 민감성 등에 대한 검증 데이터를 제공할 수 있음

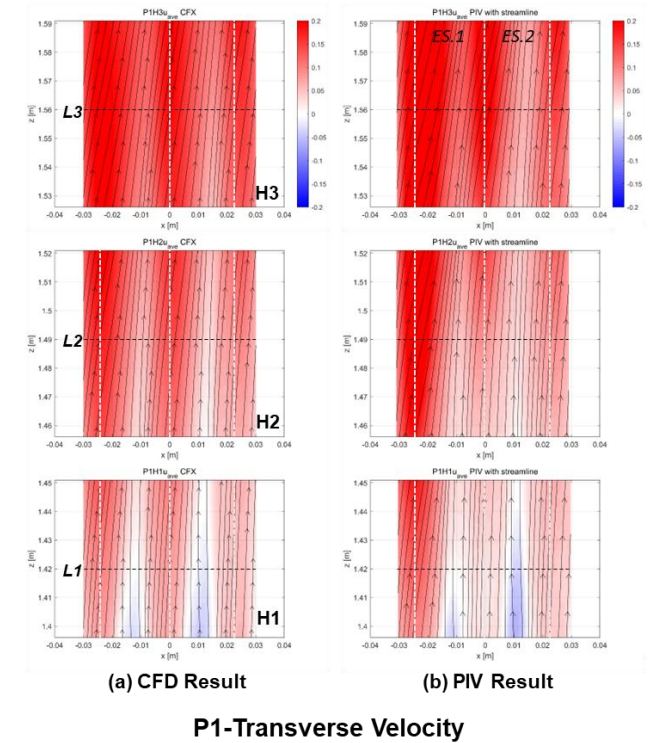
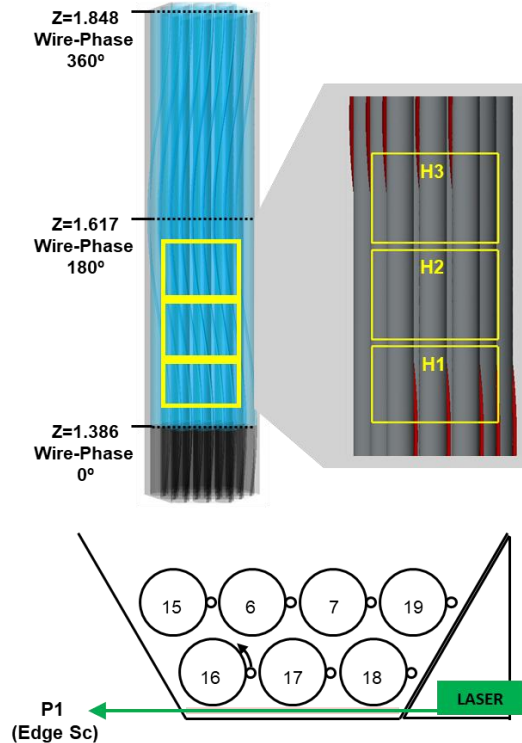
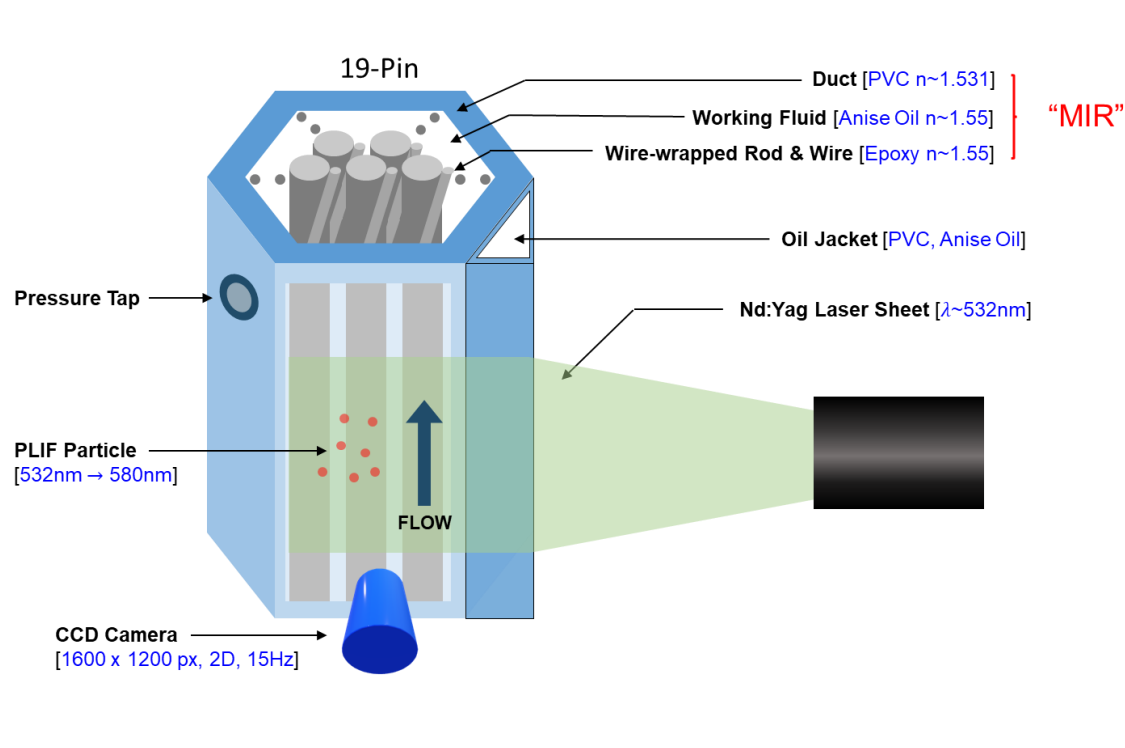


그림. SFR Wire-wrapped fuel bundle에서의 2차원 속도장 검증 예

5. 검증, 난류 통계, 그리고 실험 데이터의 역할

■ CFD 검증을 위한 유동 가시화 데이터의 필요성

- Particle Image Velocimetry, Magnetic Resonance Velocimetry 등은 CFD의 난류 모델 민감도, 격자 민감성 등에 대한 검증 데이터를 제공할 수 있음

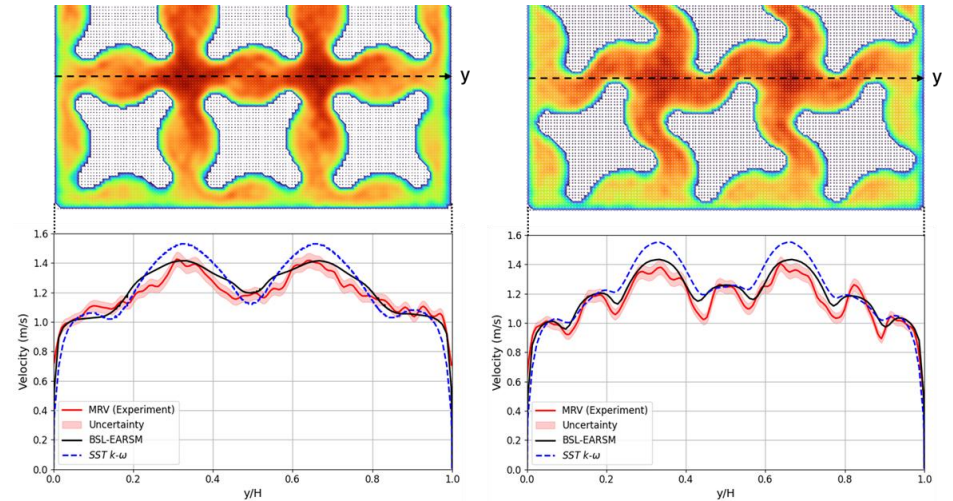
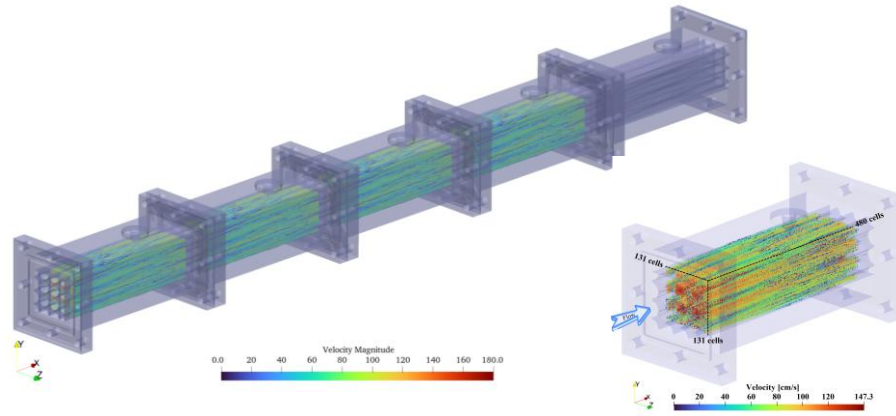
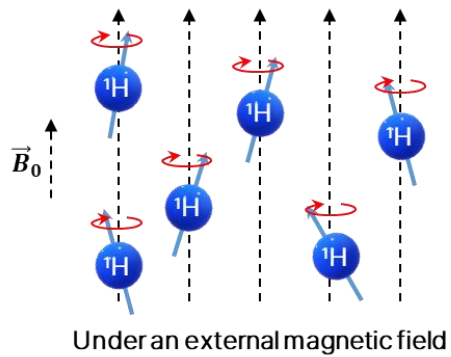
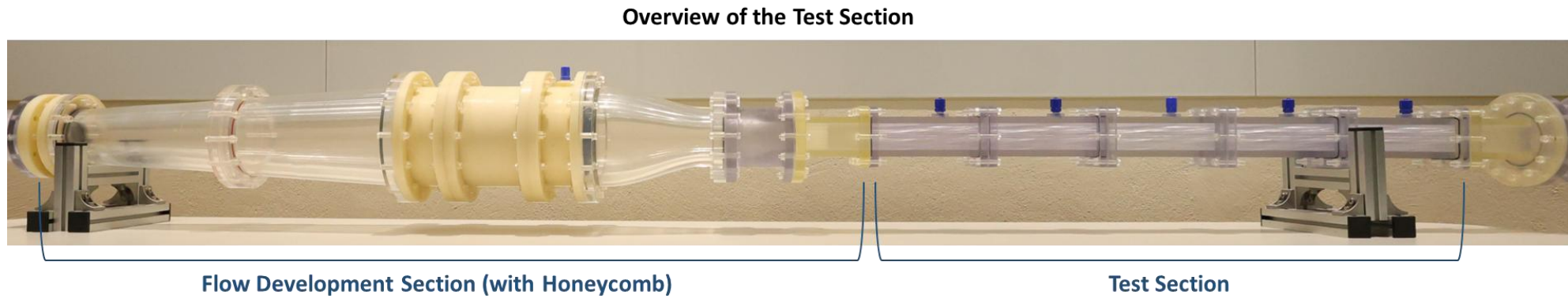


그림. MRV를 이용한 유동 데이터 검증 예

6. 원자력 Surrogate의 향후 방향

■ 원자력 Surrogate 모델은 어디로 가야 하는가

- 범용적인 CFD 모델이 아니라, 현상, 형상 도메인 특화 모델로 개발해야 함
 - 모든 형상과 운전 조건을 대체하는 universal surrogate는 아직 현실적으로 어려움.
 - Ex: 핵연료 집합체, spacer grid, low plenum, helical coil heat exchanger, pebble bed 등
- 대규모 CFD 데이터셋 기반의 표준 프레임워크 구축
 - 원자력분야 CFD Surrogate 모델은 신뢰성 있는 CFD 데이터셋이 없이는 성립하기 어려움. 검증 기반 데이터셋 구축이 필요
 - 다양한 형상, 경계조건, 출력 등을 포함하는 체계적 데이터 베이스 개발이 필요
 - 공개 가능한 CFD / 실험 데이터셋을 구축하는 것이 향후 경쟁력을 가질 수 있음
- Neural Operator 기반의 학습 사례 추가
 - FNO, DeepONet, Transformer 등 데이터의 형상 및 조건에 따른 장단점을 판단하고, 최적의 모델 구축 사례를 추가 해야 함
- Physics-informed 모델과 결합
 - Surrogate Model의 불완전한 보존성을 보장하기 위해 지배방정식, 경계조건, 난류 특성 등을 모델 구조에 반영해야 함

6. 원자력 Surrogate의 향후 방향

■ Surrogate의 실용적 사용을 위하여

● Multi-fidelity / Multi-scale 모델을 구축

- 원자력 열수력 해석은 subchannel, RANS, LES, DNS, 실험 데이터 (유동 가시화 등)의 다양한 Fidelity를 가짐
- 하나의 surrogate가 모든 층위의 Fidelity를 대체하는 것이 아닌, Fidelity level에 따른 surrogate model 제작이 필요

● 검증, 불확실성 정량화, 표준화 체계 확보

● 디지털 트윈의 엔진으로 사용

- Surrogate model은 설계 최적화 뿐 아니라, 감시, 이상 진단, 사고 시나리오의 빠른 평가 등에도 사용 가능
- 실제 운전 데이터와 CFD surrogate의 결합을 통해 SMR 등의 현 상태 추정과 제어 지원이 가능

감사합니다