

Multi-resolution Time-Series Transformer for Long-term Forecasting

<https://arxiv.org/abs/2311.04147>

0. Introduction

- 시계열 예측은 금융, 날씨, 전력망 등 다양한 분야에서 필수적임. 장기 예측에서는 긴 의존성과 다양한 주기성을 동시에 처리하기 어려움
- 기존 Transformer 기반 모델은 동일한 시간 해상도로 데이터 처리, 다양한 주기 패턴 동시 포착에 한계 있음
- 본 논문에서는 여러 시간 해상도에서 시계열 패치를 동시에 처리할 수 있는 모델 제안
- 멀티브랜치 구조와 상대적 위치 인코딩 활용, 장기 예측 성능과 주기 패턴 학습 능력 개선

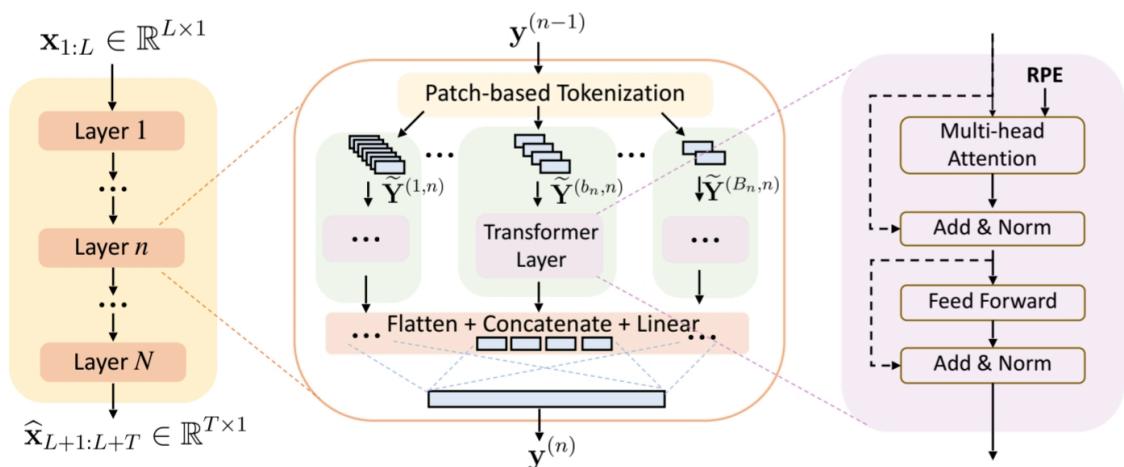
1. Overview

- 모델명 : Multi-resolution Time-Series Transformer (MTST)
- 핵심 아이디어 : 시계열을 다양한 크기의 패치로 나누어 각기 다른 주기와 해상도 패턴 학습, 병렬 브랜치 구조로 처리
- 구조 요약 :
 - 짧은 패치는 단기 변화, 긴 패치는 장기 추세와 계절성 패턴 추출
 - 각 패치 그룹마다 별도 브랜치에서 Transformer 블록 적용
 - 상대적 위치 인코딩으로 위치 정보 유연하게 반영
- 단변량과 다변량 시계열 모두 적용 가능

2. Challenges

- 동일 해상도나 패치 크기만 사용하는 기존 모델은 다양한 주기와 시간 해상도를 동시에 포착 어려움
- 단기 변화와 장기 추세, 계절성을 하나의 해상도에서 균형 있게 학습 어려움
- Transformer 구조에서 위치 인코딩이 주기적 특성을 충분히 반영하지 못할 수 있음
- 긴 예측 지평선을 커버해야 해서 계산 복잡도와 일반화 측면 부담 존재

3. Method



- 모델 구조 :
 - 입력 시계열을 여러 패치 크기로 나누어 다양한 주기 패턴 추출
 - 각 패치 그룹은 별도 브랜치에서 Transformer 블록 적용
- 위치 인코딩 :
 - 상대적 위치 인코딩 사용, 패치 간 위치 관계와 반복적 패턴 반영
- 병합 및 예측 :
 - 여러 브랜치에서 추출한 특징 통합, 최종 예측값 생성
 - 다양한 주기와 해상도 동시에 고려해 학습
- 입력 전처리와 학습 세부사항은 논문 본문에 자세히 설명됨, 핵심 설계 개념 중심으로 요약

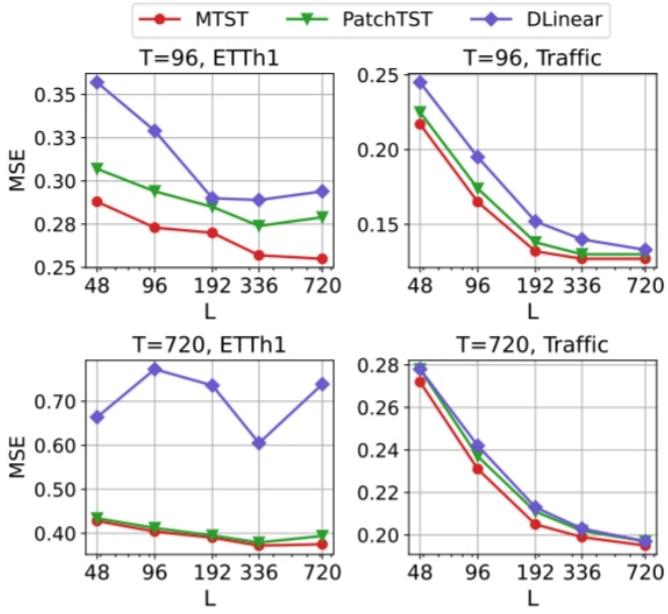
4. Experiments

- 데이터셋 : 여러 실제 장기 예측용 시계열 데이터셋 사용, 단변량과 다변량 모두 포함
- 비교 모델 : 기존 Transformer 기반 시계열 예측 모델과 기타 최신 모델
- 평가 지표 : MAE, MSE 등 장기 예측 적합 표준 지표 사용
- 실험 설계 : 패치 크기 다양성에 따른 성능 변화, 브랜치 수와 구조 변화에 따른 ablation study 수행

5. Results

Models		MTST		PatchTST		DLinear		MICN		TimesNet		Fedformer		Autoformer		Pyraformer	
Dataset	T	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Traffic	96	0.356*	0.244*	0.367	0.251	0.410	0.282	0.473	0.293	0.595	0.318	0.576	0.359	0.597	0.371	2.085	0.468
	192	0.375*	0.251*	0.385	0.259	0.423	0.287	0.483	0.298	0.615	0.326	0.610	0.380	0.607	0.382	0.867	0.467
	336	0.386*	0.256*	0.398	0.265	0.436	0.296	0.491	0.303	0.616	0.326	0.608	0.375	0.623	0.387	0.869	0.469
	720	0.425*	0.279*	0.434	0.287	0.466	0.315	0.559	0.327	0.655	0.353	0.621	0.375	0.639	0.395	0.881	0.473
Electricity	96	0.127*	0.222*	0.130	0.222	0.140	0.237	0.157	0.266	0.178	0.284	0.186	0.302	0.196	0.313	0.386	0.449
	192	0.144*	0.238*	0.148	0.240	0.153	0.249	0.175	0.287	0.187	0.289	0.197	0.311	0.211	0.324	0.386	0.443
	336	0.162*	0.256*	0.167	0.261	0.169	0.267	0.200	0.308	0.208	0.307	0.213	0.328	0.214	0.327	0.378	0.443
	720	0.199*	0.289*	0.202	0.291	0.203	0.301	0.228	0.338	0.245	0.321	0.233	0.344	0.236	0.342	0.376	0.445
Weather	96	0.150*	0.199*	0.152	0.199	0.176	0.237	0.178	0.249	0.163	0.219	0.238	0.314	0.249	0.329	0.896	0.556
	192	0.194*	0.240*	0.197	0.243	0.220	0.282	0.243	0.269	0.211	0.259	0.275	0.329	0.325	0.370	0.622	0.624
	336	0.246*	0.281*	0.249	0.283	0.265	0.319	0.278	0.338	0.286	0.311	0.339	0.377	0.351	0.391	0.739	0.753
	720	0.319*	0.333*	0.320	0.335	0.323	0.362	0.320	0.360	0.359	0.363	0.389	0.409	0.415	0.426	1.004	0.934
ETTh1	96	0.358*	0.390*	0.375	0.399	0.375	0.399	0.413	0.442	0.421	0.440	0.376	0.415	0.435	0.446	0.664	0.612
	192	0.396*	0.414*	0.414	0.421	0.405	0.416	0.451	0.462	0.511	0.498	0.423	0.446	0.456	0.457	0.790	0.681
	336	0.391*	0.420*	0.431	0.436	0.439	0.443	0.556	0.528	0.484	0.478	0.444	0.462	0.486	0.487	0.891	0.738
	720	0.430*	0.457*	0.449	0.466	0.472	0.490	0.658	0.607	0.554	0.527	0.469	0.492	0.515	0.517	0.963	0.782
ETTh2	96	0.257*	0.326*	0.274	0.336	0.289	0.353	0.303	0.364	0.366	0.417	0.332	0.374	0.332	0.368	0.645	0.597
	192	0.309*	0.361*	0.339	0.379	0.383	0.418	0.403	0.446	0.426	0.447	0.407	0.446	0.426	0.434	0.788	0.683
	336	0.302*	0.366*	0.331	0.380	0.448	0.465	0.603	0.550	0.406	0.435	0.400	0.447	0.477	0.479	0.907	0.747
	720	0.372*	0.416*	0.379	0.422	0.605	0.551	1.106	0.852	0.427	0.457	0.412	0.469	0.453	0.490	0.963	0.783
ETTm1	96	0.286*	0.338*	0.290	0.342	0.299	0.343	0.308	0.360	0.356	0.385	0.326	0.390	0.510	0.492	0.543	0.510
	192	0.327*	0.366	0.332	0.369	0.335	0.365	0.343	0.384	0.452	0.428	0.365	0.415	0.514	0.495	0.557	0.537
	336	0.362*	0.389	0.366	0.392	0.369	0.386	0.395	0.411	0.419	0.425	0.392	0.425	0.510	0.492	0.754	0.655
	720	0.414*	0.421	0.420	0.424	0.425	0.421	0.427	0.434	0.452	0.451	0.446	0.458	0.527	0.493	0.908	0.724
ETTm2	96	0.162*	0.251*	0.165	0.255	0.167	0.260	0.169	0.268	0.188	0.276	0.180	0.271	0.205	0.293	0.435	0.507
	192	0.220	0.291	0.220	0.292	0.224	0.303	0.247	0.333	0.242	0.310	0.252	0.318	0.278	0.336	0.730	0.673
	336	0.272*	0.326*	0.278	0.329	0.281	0.342	0.290	0.351	0.300	0.346	0.324	0.364	0.343	0.379	1.201	0.845
	720	0.358*	0.379*	0.367	0.385	0.397	0.421	0.417	0.434	0.391	0.403	0.410	0.420	0.414	0.419	3.625	1.451
#Rank-1st (total=28)		28	26	1	0	0	3	0	0	0	0	0	0	0	0	0	0

Models		MTST (Base)		w/o Low-RES.		w/o High-RES.	
Dataset	T	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.358	0.390	0.373	0.402	0.372	0.400
	192	0.396	0.414	0.397	0.418	0.399	0.424
	336	0.391	0.420	0.397	0.428	0.399	0.424
	720	0.430	0.457	0.435	0.460	0.430	0.457
ETTh2	96	0.257	0.326	0.260	0.329	0.266	0.335
	192	0.309	0.361	0.311	0.364	0.317	0.370
	336	0.302	0.366	0.304	0.369	0.311	0.376
	720	0.372	0.416	0.373	0.417	0.380	0.422
ETTm1	96	0.286	0.338	0.290	0.341	0.285	0.338
	192	0.327	0.366	0.334	0.372	0.322	0.364
	336	0.362	0.389	0.367	0.389	0.365	0.386
	720	0.414	0.421	0.422	0.421	0.419	0.423
Weather	96	0.150	0.199	0.151	0.199	0.151	0.199
	192	0.194	0.240	0.196	0.240	0.196	0.243
	336	0.246	0.281	0.246	0.280	0.247	0.280
	720	0.319	0.333	0.322	0.335	0.324	0.335
Traffic	96	0.356	0.244	0.357	0.244	0.362	0.249
	192	0.375	0.251	0.379	0.253	0.378	0.253
	336	0.386	0.258	0.388	0.257	0.391	0.261
	720	0.425	0.281	0.425	0.279	0.427	0.283
#Rank-1 (total=20)		18	15	2	7	3	6
MTST w/		RPE (ours)		SinAPE		LenaredAPE	
Dataset	T	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	0.358	0.390	0.362	0.392	0.363	0.394
	192	0.396	0.414	0.400	0.417	0.400	0.420
	336	0.391	0.420	0.389	0.419	0.390	0.419
	720	0.430	0.457	0.440	0.464	0.435	0.462
ETTh2	96	0.257	0.326	0.258	0.328	0.258	0.328
	192	0.309	0.361	0.309	0.363	0.311	0.365
	336	0.302	0.366	0.305	0.371	0.306	0.373
	720	0.372	0.416	0.377	0.422	0.375	0.420
ETTm1	96	0.286	0.338	0.297	0.346	0.290	0.342
	192	0.327	0.366	0.332	0.370	0.330	0.370
	336	0.362	0.389	0.369	0.390	0.366	0.388
	720	0.414	0.421	0.419	0.423	0.420	0.425
Weather	96	0.150	0.199	0.150	0.197	0.150	0.198
	192	0.194	0.240	0.194	0.240	0.194	0.240
	336	0.246	0.281	0.247	0.281	0.247	0.282
	720	0.319	0.333	0.326	0.339	0.320	0.333
#Rank-1 (total=20)		19	17	4	4	2	4



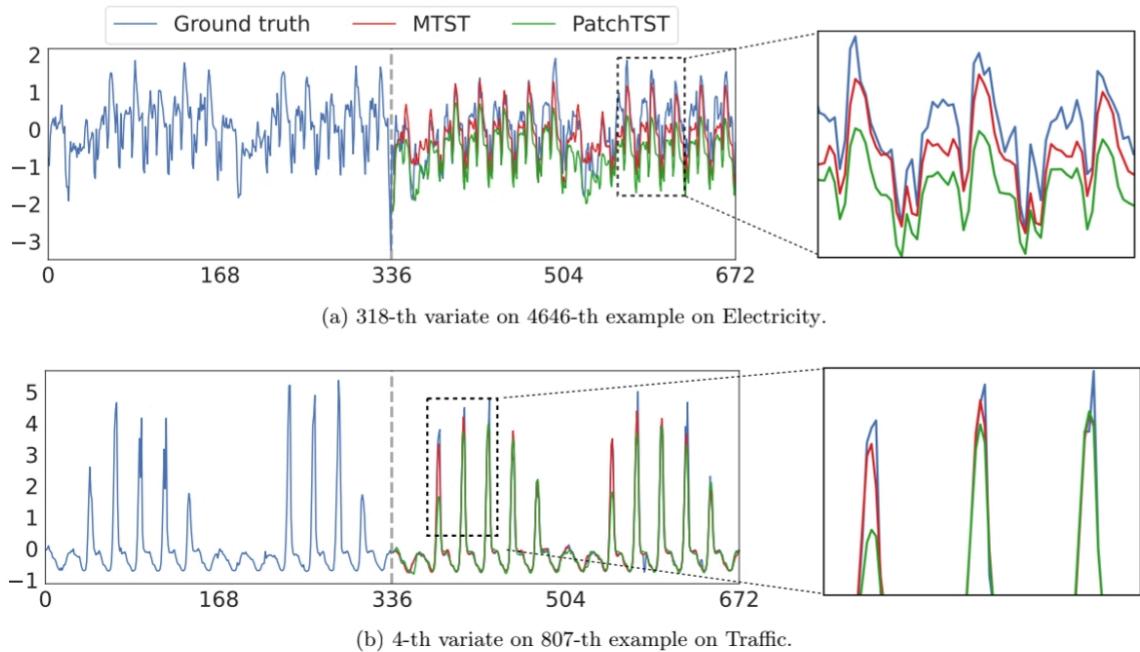


Figure 6: Visualization of forecasts for PatchTST and MTST ($L=336$ and $T=336$).

- MTST 모델, 여러 벤치마크 데이터셋에서 기존 모델 대비 우수한 성능
- 다양한 패치 크기와 멀티브랜치 구조가 함께 작동할 때 예측 정확도 향상
- 상대적 위치 인코딩, 위치 민감성과 주기성 학습에 긍정적 기여
- ablation study 결과, 브랜치 수나 패치 크기 다양성 줄이면 성능 저하, 설계 요소 중요성 확인
- 모델 복잡도와 계산량 측면에서는 기존 단일 해상도 모델 대비 부담 가능성 있음

6. Insight

- 다양한 시간 해상도 동시에 고려하는 구조, 장기 시계열 예측에서 유망
- 위치 인코딩 설계, 절대 위치 방식보다 반복 패턴 학습에 효과적
- 실무 적용 시, 패치 크기 선택, 브랜치 수, 계산 리소스 신중히 설계 필요
- 향후 연구 방향 :
 - 연산 효율화 위한 경량화 브랜치, 파라미터 공유 구조
 - 불규칙 샘플링, 결측치 처리, 다해상도 전환 고려
 - 다양한 도메인에서 적용 및 일반화 평가