

2026 02 05
발표 자료

광운대학교 로봇학과
FAIR Lab

김한서

이번 주 진행사항

- PMDformer
 - 논문 리뷰

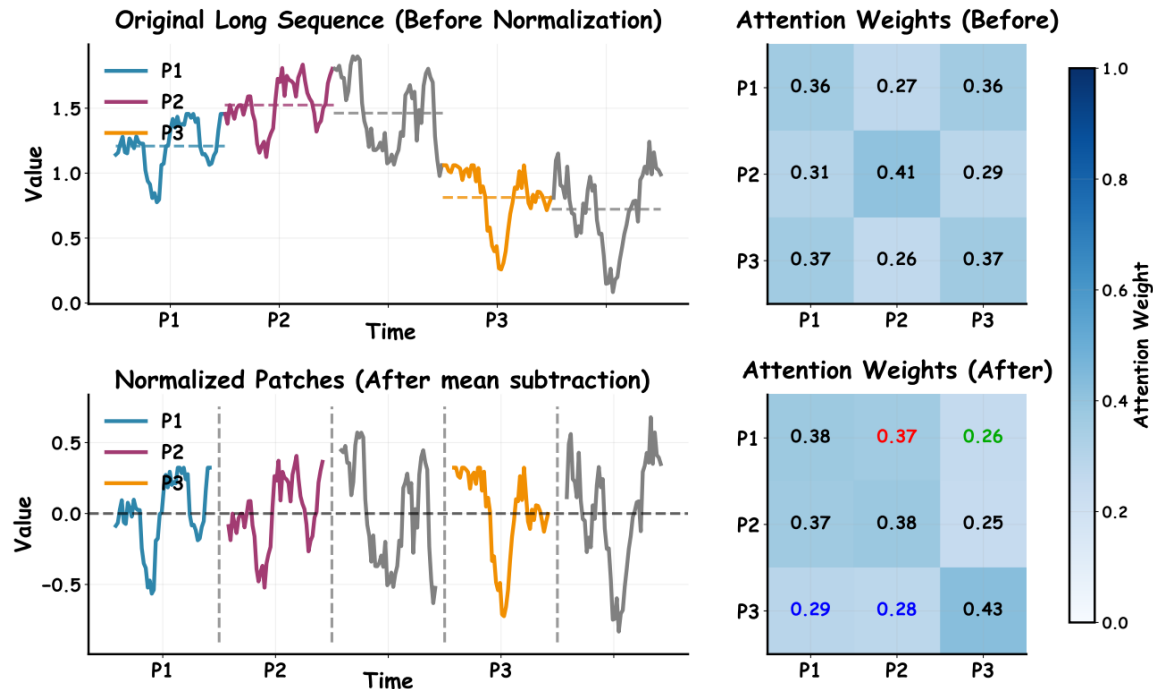
PMDFORMER: PATCH-MEAN DECOUPLING TRANSFORMER FOR LONG-TERM FORECASTING

Anonymous authors

Paper under double-blind review

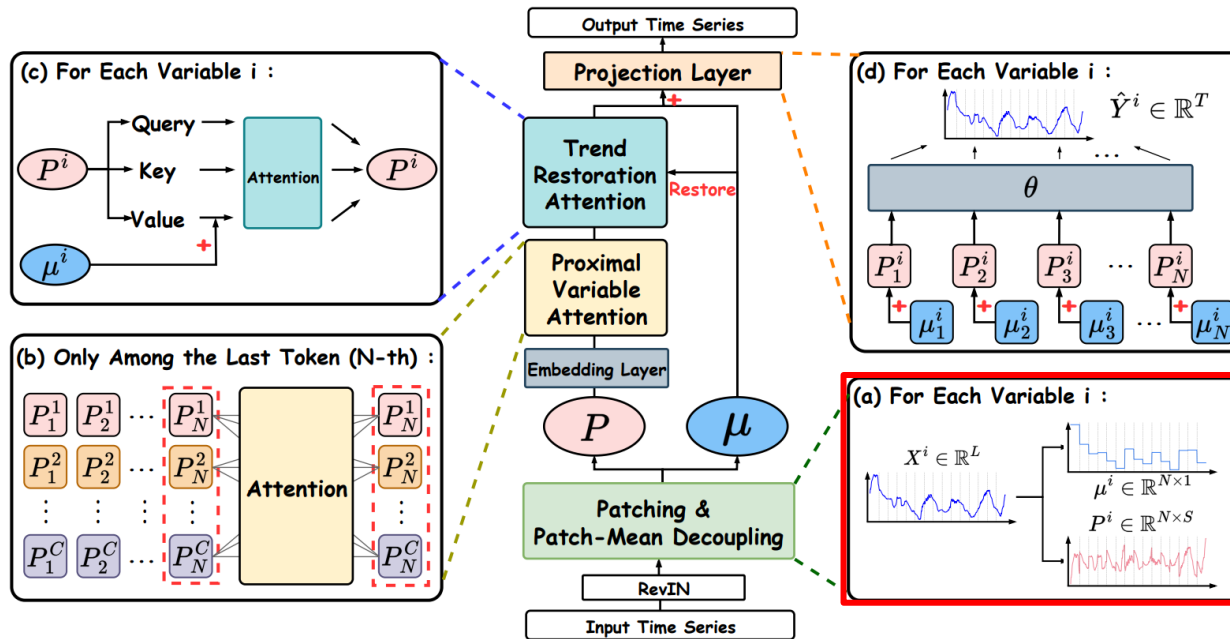
- 인용 수: 0회 (Google Scholar, 2026-02-03)
- Under review ICLR 2026
- 링크: <https://openreview.net/pdf?id=rfJ41gK9Ct>

Background



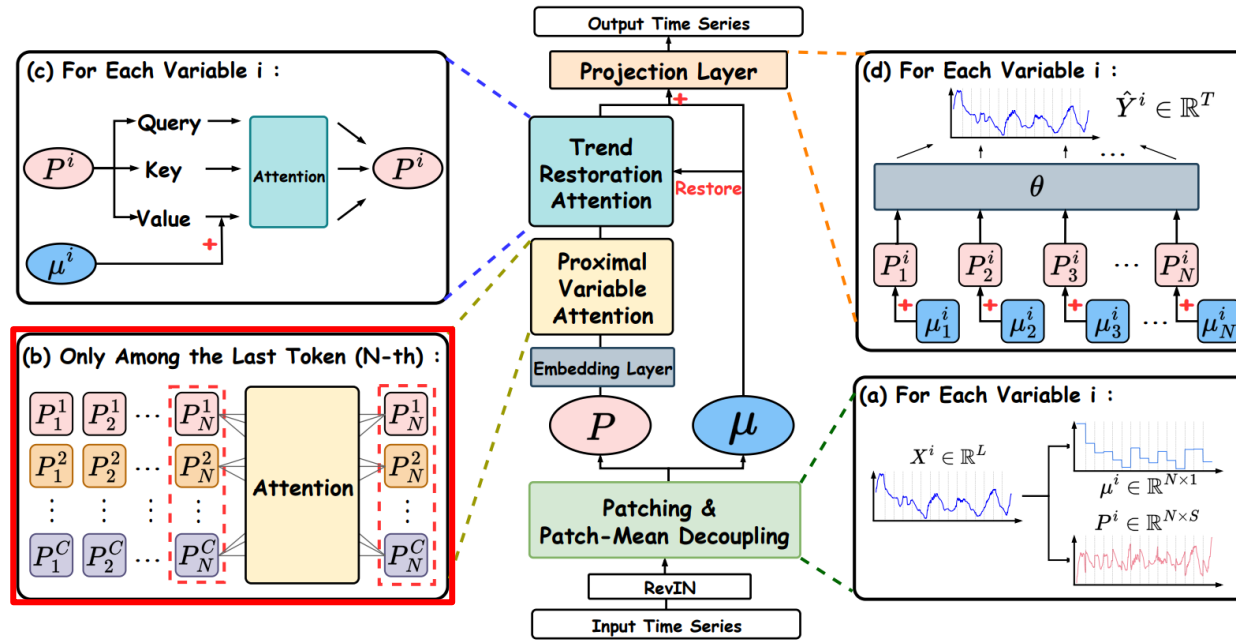
- 기존 방식에서 Shape가 비슷한 P1, P2보다, 단순히 수치적 범위가 가까운 P1, P3의 가중치가 더 높은 것을 발견
- 위의 Shape Similarity 파악을 방해하는 현상을 해결하기 위해 Patch-Mean Decoupling을 통해 각 패치에서 평균(Trend)를 분리하여 처리

모델 구조



- Patching & Patch-Mean Decoupling
 - 입력 시계열 데이터를 일정 길이 s 인 겹치지 않는 N 개의 패치(P_j^i)로 나눔
 - Decoupling을 통해 나뉜 각 패치에서 평균(Trend) μ^i 을 빼 Shape P^i 만 남김
- Embedding Layer
 - Shape에 Linear Projection과 Positional Embedding을 더해 토큰을 만듦

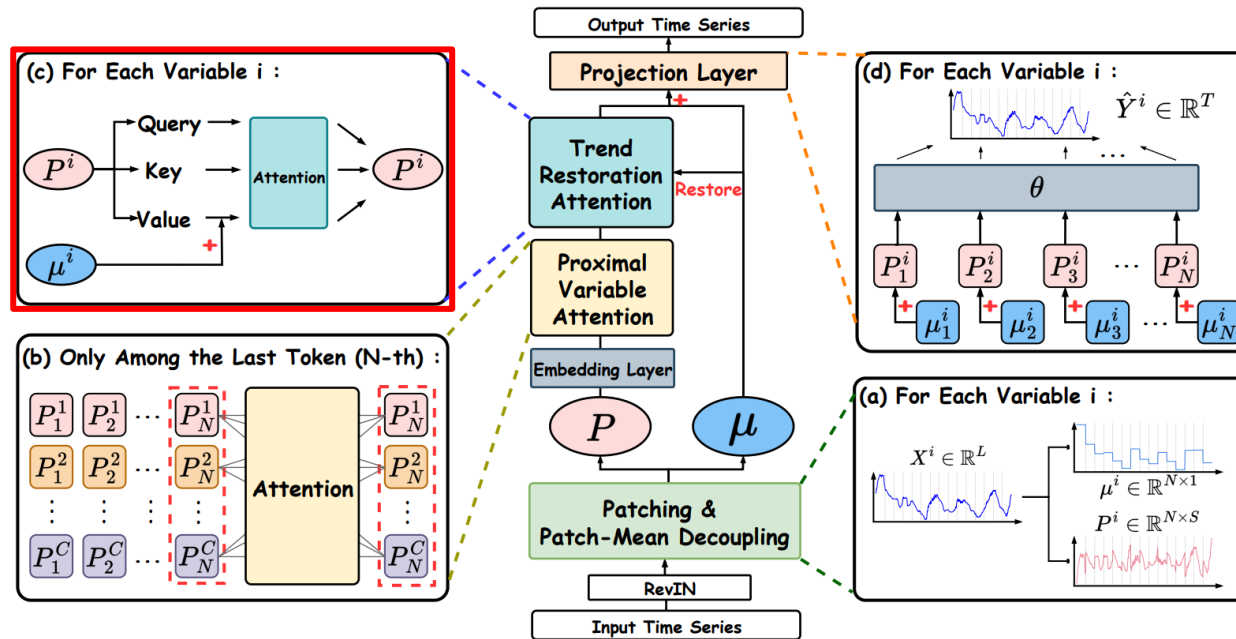
모델 구조



- Proximal Variable Attention

- PMD 과정을 거친 모든 변수 C 의 마지막 토큰들을 추출해 Attention 수행
- 이후 FFN을 거친 뒤, Attention을 수행하지 않은 기존 패치들과 Concat하여 결합

모델 구조

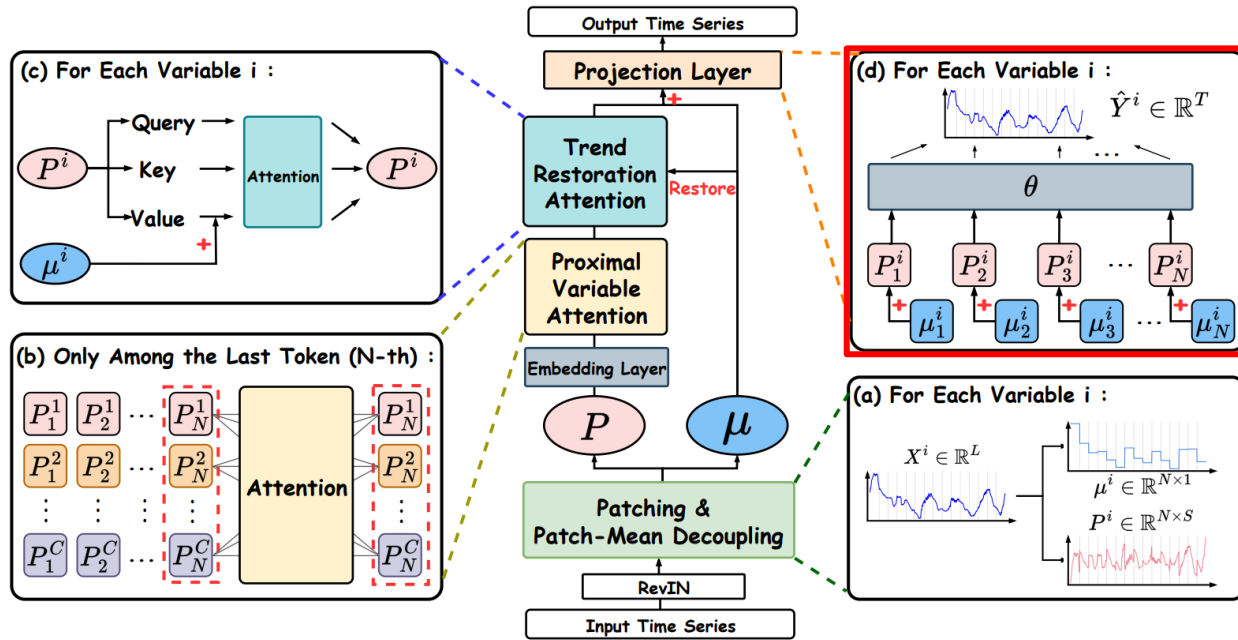


• Trend Restoration Attention

- Query와 Key는 Shape Embedding을 사용, Value는 Shape Embedding에 Trend를 더한 뒤 사용
- 이후 각 Query, Key, Value를 통해 Attention 수행

PMDformer

모델 구조



- **Projection Layer**

- Shape Embedding에 패치별 Trend를 더해준 뒤, 모든 패치 토큰을 Flatten하여 θ 를 통과시켜 최종 예측값 출력

$$Y^i = (P^i + \mu^i)W_o + b_o$$

주요 모델 성능 비교

- 대부분의 데이터셋에서 기존 모델들보다 더 뛰어난 성능을 보임

Models		PMDformer		TQNet		TimeBase		SOFTS		SparseTSF		ModernTCN		iTransformer		TimeMixer		PatchTST	
Year		ours		2025		2025		2024		2024		2024		2024		2024		2023	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ECL	96	0.122	0.214	0.143	0.244	0.139	0.231	0.133	0.229	0.139	0.239	0.131	0.227	0.135	0.233	0.142	0.247	0.141	0.240
	192	0.140	0.231	0.151	0.247	0.153	0.255	0.160	0.255	0.155	0.250	0.145	0.241	0.155	0.253	0.159	0.256	0.156	0.256
	336	0.152	0.245	0.166	0.261	0.169	0.262	0.182	0.277	0.171	0.265	0.162	0.261	0.169	0.267	0.169	0.270	0.172	0.267
	720	0.177	0.272	0.194	0.291	0.207	0.294	0.224	0.310	0.208	0.300	0.193	0.289	0.204	0.301	0.209	0.313	0.208	0.299
	Avg	0.148	0.241	0.164	0.261	0.167	0.258	0.175	0.268	0.180	0.264	0.158	0.255	0.166	0.264	0.170	0.272	0.169	0.266
Traffic	96	0.338	0.212	0.398	0.297	0.394	0.267	0.355	0.255	0.389	0.268	0.382	0.267	0.374	0.273	0.396	0.294	0.363	0.250
	192	0.367	0.227	0.397	0.277	0.403	0.271	0.365	0.258	0.399	0.272	0.393	0.271	0.393	0.283	0.404	0.295	0.382	0.258
	336	0.379	0.235	0.403	0.279	0.417	0.278	0.390	0.278	0.417	0.279	0.406	0.277	0.409	0.292	0.419	0.302	0.399	0.268
	720	0.426	0.262	0.448	0.304	0.456	0.298	0.429	0.294	0.449	0.299	0.452	0.305	0.450	0.314	0.458	0.309	0.432	0.289
	Avg	0.378	0.234	0.412	0.289	0.418	0.279	0.385	0.271	0.414	0.280	0.408	0.280	0.407	0.291	0.419	0.300	0.394	0.266
Weather	96	0.141	0.181	0.160	0.213	0.146	0.198	0.165	0.219	0.174	0.231	0.155	0.210	0.159	0.212	0.163	0.223	0.149	0.199
	192	0.185	0.226	0.212	0.261	0.185	0.241	0.213	0.258	0.216	0.267	0.205	0.254	0.203	0.252	0.201	0.254	0.193	0.243
	336	0.236	0.274	0.260	0.299	0.236	0.281	0.272	0.305	0.260	0.299	0.255	0.290	0.253	0.291	0.258	0.300	0.240	0.281
	720	0.305	0.323	0.328	0.343	0.309	0.331	0.380	0.371	0.325	0.345	0.317	0.336	0.317	0.337	0.329	0.348	0.312	0.334
	Avg	0.217	0.251	0.240	0.279	0.219	0.263	0.258	0.288	0.244	0.286	0.233	0.273	0.233	0.273	0.238	0.281	0.224	0.264
Solar	96	0.160	0.193	0.181	0.242	0.179	0.248	0.192	0.239	0.205	0.241	0.196	0.258	0.217	0.255	0.232	0.271	0.205	0.239
	192	0.178	0.211	0.203	0.261	0.213	0.252	0.197	0.259	0.215	0.265	0.224	0.280	0.208	0.257	0.238	0.293	0.227	0.280
	336	0.190	0.218	0.219	0.272	0.222	0.261	0.212	0.273	0.213	0.276	0.240	0.288	0.238	0.309	0.234	0.301	0.225	0.290
	720	0.196	0.221	0.231	0.281	0.235	0.264	0.217	0.274	0.232	0.272	0.246	0.299	0.270	0.319	0.273	0.319	0.249	0.291
	Avg	0.181	0.211	0.209	0.264	0.216	0.254	0.201	0.264	0.216	0.264	0.227	0.281	0.233	0.285	0.244	0.296	0.227	0.275
ETTh1	96	0.356	0.388	0.379	0.404	0.349	0.384	0.389	0.417	0.362	0.389	0.380	0.405	0.389	0.421	0.410	0.441	0.377	0.408
	192	0.397	0.416	0.429	0.441	0.387	0.410	0.427	0.443	0.404	0.412	0.418	0.428	0.424	0.446	0.448	0.465	0.413	0.431
	336	0.420	0.432	0.454	0.455	0.408	0.418	0.446	0.458	0.435	0.428	0.453	0.450	0.456	0.469	0.482	0.490	0.436	0.446
	720	0.432	0.456	0.499	0.506	0.439	0.446	0.468	0.491	0.426	0.448	0.480	0.484	0.545	0.532	0.475	0.500	0.455	0.475
	Avg	0.401	0.423	0.440	0.452	0.396	0.415	0.433	0.452	0.407	0.419	0.433	0.442	0.454	0.467	0.454	0.474	0.420	0.440
ETTh2	96	0.269	0.329	0.288	0.354	0.292	0.345	0.309	0.365	0.294	0.346	0.273	0.341	0.305	0.361	0.315	0.380	0.276	0.339
	192	0.333	0.373	0.377	0.403	0.339	0.387	0.378	0.405	0.340	0.377	0.337	0.385	0.405	0.421	0.383	0.415	0.342	0.385
	336	0.357	0.396	0.377	0.415	0.358	0.410	0.460	0.461	0.360	0.398	0.369	0.414	0.411	0.436	0.385	0.438	0.364	0.405
	720	0.390	0.429	0.424	0.452	0.400	0.448	0.441	0.467	0.383	0.425	0.408	0.448	0.448	0.470	0.432	0.471	0.395	0.434
	Avg	0.337	0.382	0.367	0.406	0.347	0.398	0.397	0.425	0.344	0.387	0.347	0.397	0.392	0.422	0.379	0.426	0.344	0.391
ETTm1	96	0.279	0.328	0.296	0.349	0.311	0.351	0.303	0.361	0.314	0.359	0.313	0.357	0.315	0.369	0.332	0.384	0.298	0.352
	192	0.323	0.358	0.337	0.374	0.338	0.371	0.336	0.377	0.348	0.376	0.343	0.377	0.349	0.388	0.355	0.398	0.335	0.373
	336	0.361	0.383	0.369	0.393	0.364	0.386	0.384	0.407	0.368	0.386	0.372	0.393	0.381	0.409	0.386	0.416	0.366	0.394
	720	0.421	0.416	0.447	0.434	0.413	0.414	0.438	0.438	0.419	0.413	0.420	0.420	0.437	0.439	0.452	0.457	0.420	0.421
	Avg	0.346	0.371	0.362	0.388	0.357	0.381	0.365	0.396	0.362	0.384	0.362	0.387	0.371	0.401	0.381	0.414	0.355	0.385
ETTm2	96	0.155	0.240	0.169	0.257	0.167	0.259	0.188	0.274	0.167	0.259	0.179	0.269	0.179	0.274	0.192	0.285	0.165	0.260
	192	0.213	0.282	0.231	0.299	0.219	0.297	0.256	0.317	0.219	0.297	0.243	0.312	0.239	0.314	0.253	0.329	0.219	0.298
	336	0.267	0.319	0.282	0.337	0.271	0.330	0.334	0.366	0.271	0.330	0.270	0.330	0.309	0.356	0.307	0.362	0.268	0.333
	720	0.347	0.373	0.371	0.398	0.353	0.380	0.392	0.406	0.353	0.380	0.362	0.393	0.387	0.407	0.380	0.412	0.352	0.386
	Avg	0.246	0.304	0.263	0.323	0.253	0.317	0.293	0.341	0.253	0.317	0.264	0.326	0.279	0.338	0.283	0.347	0.251	0.319
1st Count		32	33	0	0	7	4	1	0	2	3	0	0	0	0	0	0	0	0

Ablation study

Design	Norm	ETTh2		ETTm1		Weather		Traffic		Solar	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
PMDformer	PMD	0.337	0.382	0.346	0.371	0.217	0.251	0.378	0.234	0.181	0.211
Replace	w/ stdev	0.354	0.392	0.347	0.370	0.218	0.252	0.396	0.259	0.205	0.221
	SAN	0.360	0.403	0.353	0.380	0.225	0.275	0.392	0.273	0.182	0.235
	✗	0.359	0.394	0.347	0.370	0.223	0.260	0.397	0.258	0.199	0.212

(a) PMD module

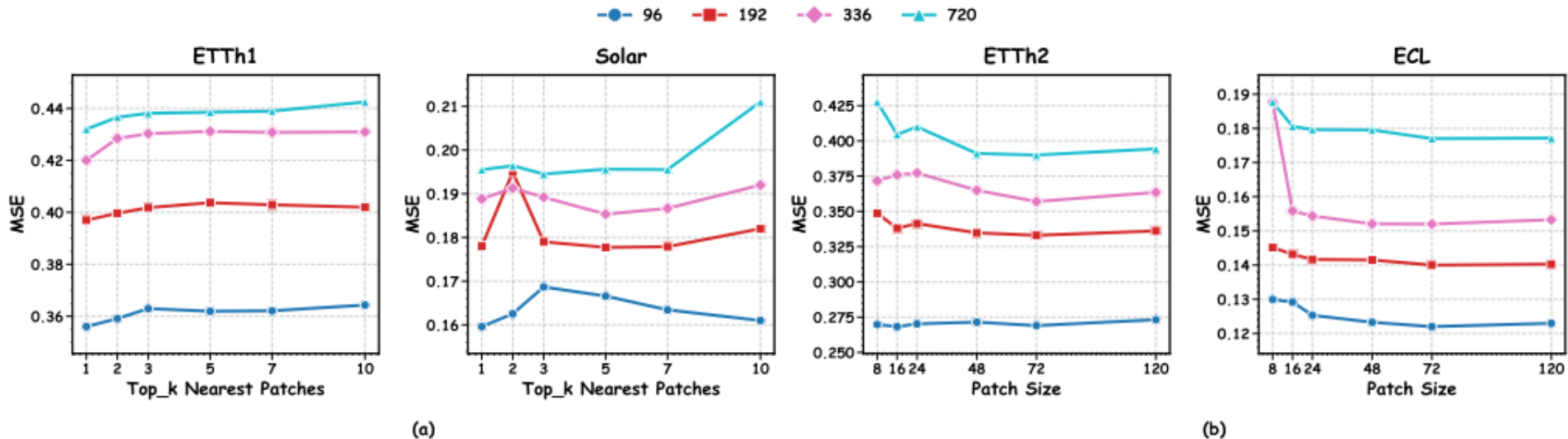
Design	TRA	PVA	ETTh2		ETTm1		Traffic		Solar	
			MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
PMDformer	✓	Last Token	0.337	0.382	0.346	0.371	0.378	0.234	0.181	0.211
Replace	✓	All Token	0.340	0.384	0.354	0.375	0.380	0.239	0.186	0.214
	Self-attention	Last Token	0.345	0.386	0.352	0.372	0.388	0.251	0.196	0.217
	Swap Order ⇌		0.342	0.385	0.350	0.372	0.379	0.235	0.188	0.216
w/o	✗	Last Token	0.344	0.381	0.347	0.372	0.410	0.270	0.215	0.226
	✓	✗	0.340	0.383	0.347	0.371	0.386	0.240	0.194	0.214
	✗	✗	0.346	0.384	0.351	0.372	0.426	0.288	0.222	0.230

(b) TRA, PVA module

- 핵심 모듈 PMD, TRA, PVA를 대체 또는 제거했을 때의 성능보다 모두 적용했을 때의 성능 향상 폭이 큼

PMDformer

파라미터별 성능 비교



- (a) Top_k 개수별 MSE 성능 비교, PVA에서 Attention할 가장 마지막 패치의 개수 실험
- (b) Patch Size별 MSE 성능 비교, 24~72 사이의 패치 크기가 가장 성능이 좋았음