

Sensorformer: Cross-patch attention with global-patch compression is effective for high-dimensional multivariate time series forecasting

<https://arxiv.org/abs/2501.03284>

0. Introduction

- 고차원 다변량 시계열 예측에서 기존 Transformer는 변수 간 동적 의존성을 충분히 포착하지 못함
- 변수 수가 많고 데이터 길이가 길면 연산 복잡도와 메모리 문제 발생
- Sensorformer는 교차 패치 어텐션(cross-patch attention)과 글로벌 패치 압축(global-patch compression)을 도입
- 이를 통해 변수 간 및 시간 간 상관관계를 동시에 모델링하면서 계산 효율성 확보
- 실험 결과 Sensorformer는 여러 고차원 시계열 데이터셋에서 기존 Transformer 대비 예측 정확도 향상
- 모델 설계는 계산 효율성과 예측 성능을 모두 고려, 다양한 산업 및 센서 데이터 적용 가능

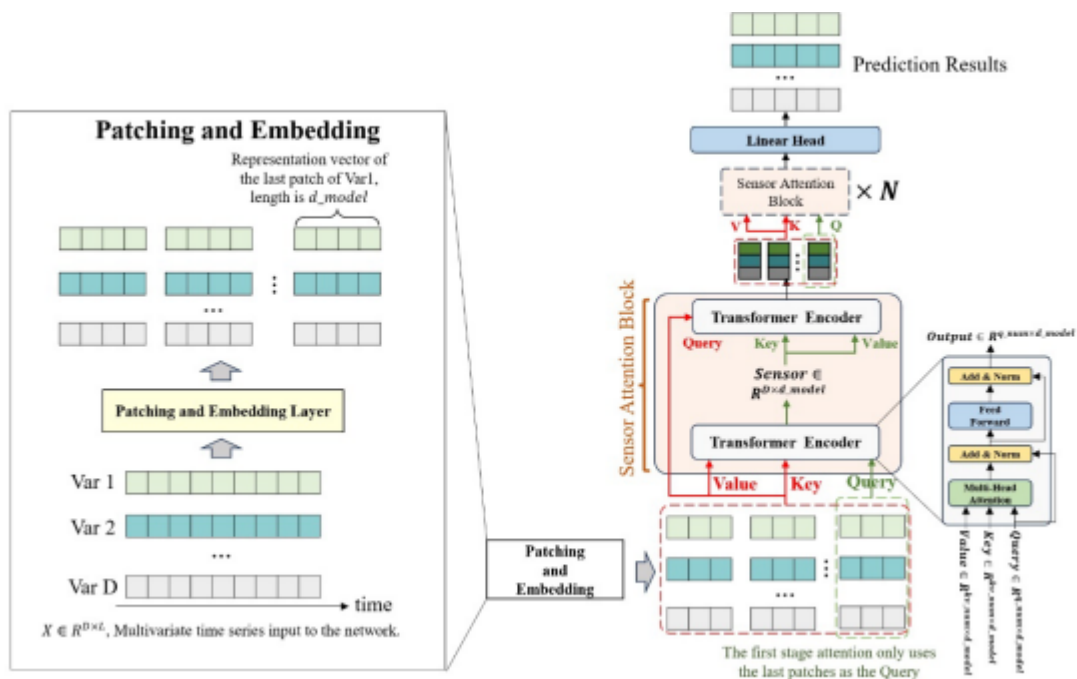
1. Overview

- 고차원 다변량 시계열에서 변수 간 및 시간 간 의존성을 효과적으로 포착하는 것이 목표
- 다양한 도메인(산업 센서, 금융, 환경 등)에 적용 가능하도록 설계
- 계산 효율성과 예측 성능 동시 확보를 지향

2. Challenges

- 고차원 다변량 시계열에서는 변수 수가 많고 데이터 길이가 길어 기존 Transformer 계산 복잡도와 메모리 부담이 큼
- 변수 간 동적 의존성과 시간 간 관계를 동시에 학습하기 어려움
- 패치 단위로 입력을 나누더라도 정보 손실 가능성 존재
- 다양한 도메인과 센서 환경에서 일반화 가능한 모델 설계 필요
- 기존 Transformer 기반 모델은 일부 변수나 특정 시간 구간만 주목하는 경향이 있어 장기 예측 성능 제한
- 계산 효율성과 정확도를 동시에 만족시키는 설계가 어려움

3. Method



- 시계열 데이터를 패치 단위로 나누어 각 패치를 임베딩
- 글로벌 패치 압축(Global-patch compression)으로 전체 시계열 정보를 압축, 연산 효율성 확보
- 교차 패치 어텐션(Cross-patch attention)을 통해 변수 간과 시간 간 의존성을 동시에 학습
- 기존 Transformer 구조를 크게 변경하지 않고, 임베딩과 어텐션 모듈 중심으로 설계

- 모델은 패치 압축과 교차 어텐션을 반복하여 장기 의존성과 변수 간 상관관계를 효과적으로 포착
- 다양한 고차원 시계열 데이터셋에서 일반화 가능하도록 설계

4. Experiments

- 여러 고차원 다변량 시계열 데이터셋에서 모델 성능 평가
- 데이터셋 도메인 산업 센서, 환경, 금융 등
- 기존 Transformer 기반 모델과 비교하여 교차 패치 어텐션과 글로벌 패치 압축 효과 분석
- Ablation study 진행 패치 압축 제거, 교차 어텐션 제거, 임베딩 방식 변경 시 성능 비교
- 장기 예측 성능, 계산 효율성, 일반화 능력 평가
- Sensorformer는 대부분 데이터셋에서 기존 모델 대비 예측 정확도 향상과 계산 효율성 개선 확인

5. Results

Models	Sensorformer (Ours)		iTransformer		PatchTST		Crossformer		TDE		TimesNet		DLinear		SCINet		FEDformer		Stationary		Autoformer		
	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	
ETTm1	96	0.329	0.369	0.334	0.368	0.334	0.374	0.404	0.426	0.364	0.387	0.338	0.375	0.345	0.372	0.418	0.418	0.379	0.419	0.386	0.398	0.395	0.475
	192	0.371	0.395	0.374	0.399	0.376	0.392	0.45	0.451	0.398	0.404	0.374	0.387	0.38	0.395	0.439	0.45	0.426	0.441	0.459	0.444	0.555	0.496
	336	0.6	0.41	0.426	0.42	0.404	0.413	0.532	0.515	0.428	0.425	0.41	0.411	0.413	0.413	0.49	0.485	0.445	0.459	0.469	0.464	0.621	0.537
	720	0.499	0.447	0.491	0.459	0.468	0.447	0.666	0.589	0.487	0.461	0.478	0.45	0.474	0.457	0.595	0.55	0.543	0.48	0.585	0.518	0.871	0.561
	Avg	0.39	0.403	0.406	0.409	0.398	0.407	0.513	0.496	0.419	0.419	0.4	0.406	0.403	0.407	0.485	0.481	0.448	0.452	0.481	0.456	0.588	0.517
ETTm2	96	0.176	0.263	0.18	0.263	0.176	0.262	0.287	0.266	0.207	0.205	0.187	0.287	0.193	0.292	0.286	0.377	0.283	0.287	0.192	0.274	0.255	0.339
	192	0.346	0.366	0.25	0.369	0.243	0.366	0.414	0.492	0.28	0.364	0.248	0.369	0.284	0.362	0.399	0.445	0.349	0.378	0.28	0.339	0.241	0.34
	336	0.311	0.351	0.311	0.348	0.306	0.346	0.597	0.542	0.377	0.422	0.321	0.351	0.369	0.427	0.637	0.591	0.325	0.366	0.334	0.361	0.399	0.372
	720	0.413	0.408	0.412	0.406	0.407	0.405	1.73	1.042	0.558	0.524	0.468	0.495	0.554	0.522	0.96	0.735	0.421	0.415	0.417	0.413	0.435	0.432
	Avg	0.266	0.317	0.288	0.317	0.293	0.31	0.757	0.61	0.358	0.404	0.301	0.313	0.35	0.405	0.571	0.517	0.385	0.349	0.306	0.347	0.327	0.371
ETTh1	96	0.381	0.4	0.386	0.405	0.414	0.419	0.423	0.448	0.459	0.464	0.384	0.402	0.396	0.4	0.454	0.599	0.376	0.419	0.511	0.441	0.489	0.459
	192	0.43	0.478	0.441	0.456	0.46	0.445	0.471	0.474	0.525	0.492	0.436	0.479	0.437	0.432	0.519	0.631	0.42	0.448	0.534	0.584	0.5	0.682
	336	0.465	0.446	0.487	0.458	0.501	0.466	0.57	0.546	0.565	0.515	0.481	0.469	0.481	0.459	0.778	0.659	0.429	0.465	0.588	0.535	0.521	0.496
	720	0.504	0.498	0.503	0.491	0.5	0.488	0.653	0.621	0.594	0.558	0.523	0.5	0.519	0.516	0.836	0.699	0.586	0.507	0.643	0.606	0.514	0.512
	Avg	0.435	0.441	0.454	0.447	0.469	0.454	0.579	0.532	0.541	0.507	0.458	0.45	0.456	0.437	0.547	0.647	0.44	0.446	0.57	0.517	0.496	0.487
ETTh2	96	0.291	0.343	0.297	0.348	0.302	0.345	0.345	0.384	0.4	0.44	0.34	0.374	0.332	0.387	0.397	0.621	0.358	0.397	0.476	0.458	0.346	0.388
	192	0.373	0.394	0.38	0.4	0.388	0.4	0.577	0.658	0.528	0.509	0.462	0.414	0.477	0.478	0.86	0.689	0.429	0.459	0.512	0.483	0.456	0.452
	336	0.426	0.43	0.428	0.432	0.426	0.433	1.043	0.731	0.643	0.571	0.452	0.452	0.594	0.541	1	0.744	0.466	0.487	0.552	0.551	0.482	0.488
	720	0.432	0.448	0.431	0.448	0.437	0.446	1.104	0.763	0.876	0.679	0.462	0.468	0.831	0.657	1.349	0.818	0.463	0.474	0.562	0.56	0.515	0.51
	Avg	0.381	0.404	0.384	0.407	0.387	0.407	0.645	0.684	0.611	0.55	0.414	0.427	0.559	0.515	0.854	0.773	0.437	0.449	0.526	0.518	0.43	0.459
ECL	96	0.19	0.278	0.197	0.268	0.193	0.261	0.219	0.214	0.227	0.229	0.184	0.217	0.197	0.282	0.247	0.345	0.193	0.308	0.169	0.273	0.201	0.517
	192	0.195	0.285	0.194	0.278	0.196	0.285	0.251	0.222	0.236	0.23	0.184	0.289	0.196	0.285	0.257	0.355	0.281	0.315	0.182	0.286	0.232	0.534
	336	0.212	0.302	0.212	0.296	0.212	0.301	0.246	0.237	0.249	0.244	0.188	0.3	0.209	0.301	0.269	0.369	0.234	0.329	0.2	0.374	0.231	0.538
	720	0.244	0.334	0.246	0.332	0.244	0.334	0.28	0.243	0.248	0.271	0.222	0.32	0.245	0.313	0.299	0.39	0.246	0.355	0.227	0.371	0.254	0.561
	Avg	0.212	0.3	0.212	0.294	0.214	0.3	0.244	0.234	0.251	0.244	0.182	0.295	0.212	0.3	0.268	0.365	0.214	0.327	0.192	0.296	0.227	0.518
Exchange	96	0.096	0.194	0.095	0.207	0.097	0.206	0.256	0.267	0.094	0.218	0.107	0.234	0.088	0.218	0.267	0.396	0.148	0.278	0.111	0.235	0.197	0.523
	192	0.184	0.302	0.176	0.302	0.178	0.301	0.47	0.508	0.184	0.307	0.226	0.344	0.176	0.315	0.351	0.459	0.271	0.315	0.219	0.335	0.3	0.369
	336	0.254	0.418	0.236	0.422	0.239	0.427	1.208	0.883	0.349	0.451	0.367	0.448	0.313	0.427	1.324	0.853	0.46	0.427	0.421	0.476	0.309	0.524
	720	0.451	0.697	0.447	0.693	0.455	0.697	1.707	1.068	0.852	0.998	0.64	0.746	0.639	0.695	1.488	0.797	1.092	0.789	1.447	0.941	0.81	0.941
	Avg	0.264	0.406	0.251	0.406	0.245	0.406	0.94	0.707	0.37	0.413	0.416	0.443	0.359	0.434	0.75	0.636	0.579	0.429	0.465	0.454	0.313	0.516
Traffic	96	0.507	0.334	0.511	0.345	0.548	0.353	0.811	0.503	0.805	0.491	0.593	0.321	0.85	0.386	0.768	0.499	0.583	0.366	0.612	0.339	0.613	0.588
	192	0.506	0.338	0.519	0.348	0.534	0.345	0.817	0.508	0.786	0.474	0.617	0.336	0.598	0.37	0.789	0.505	0.684	0.373	0.613	0.34	0.616	0.582
	336	0.521	0.343	0.539	0.359	0.546	0.349	0.83	0.517	0.782	0.477	0.628	0.336	0.605	0.373	0.797	0.508	0.621	0.383	0.618	0.324	0.622	0.537
	720	0.555	0.349	0.565	0.378	0.582	0.366	0.859	0.524	0.719	0.449	0.64	0.35	0.649	0.344	0.841	0.523	0.626	0.382	0.653	0.355	0.66	0.468
	Avg	0.523	0.345	0.534	0.358	0.553	0.353	0.824	0.513	0.76	0.473	0.62	0.336	0.623	0.383	0.804	0.506	0.61	0.376	0.624	0.34	0.626	0.579
Weather	96	0.186	0.223	0.189	0.229	0.184	0.224	0.19	0.265	0.202	0.261	0.172	0.22	0.196	0.255	0.221	0.306	0.217	0.296	0.173	0.223	0.266	0.336
	192	0.232	0.263	0.235	0.266	0.232	0.263	0.236	0.268	0.242	0.298	0.219	0.241	0.237	0.298	0.261	0.34	0.238	0.336	0.245	0.285	0.307	0.567
	336	0.285	0.302	0.288	0.306	0.285	0.303	0.292	0.323	0.287	0.335	0.258	0.306	0.263	0.335	0.309	0.378	0.239	0.338	0.323	0.338	0.359	0.593
	720	0.358	0.35	0.369	0.353	0.361	0.351	0.39	0.412	0.351	0.386	0.345	0.359	0.345	0.381	0.377	0.477	0.489	0.438	0.414	0.41	0.419	0.478
	Avg	0.265	0.285	0.27	0.289	0.266	0.283	0.277	0.317	0.271	0.32	0.258	0.317	0.267	0.317	0.292	0.363	0.289	0.34	0.328	0.314	0.338	0.582
Solar Energy	96	0.233	0.276	0.235	0.276	0.239	0.281	0.31	0.331	0.312	0.399	0.25	0.292	0.29	0.378	0.257	0.344	0.242	0.342	0.213	0.249	0.884	0.711
	192	0.266	0.297	0.279	0.304	0.278	0.305	0.354	0.325	0.239	0.416	0.296	0.318	0.32	0.398	0.28	0.38	0.285	0.334	0.254	0.272	0.434	0.602
	336	0.291	0.312	0.311	0.325	0.304	0.321	0.375	0.335	0.308	0.43	0.319	0.331	0.353	0.415	0.394	0.389	0.282	0.376	0.29	0.294	0.841	0.723
	720	0.292	0.313	0.311	0.325	0.304	0.319	0.369	0.365	0.37	0.425	0.318	0.337	0.356	0.413	0.388	0.388	0.357	0.427	0.285	0.295	0.882	0.711
	Avg	0.271	0.3	0.284	0.308	0.281	0.307	0.341	0.339	0.347	0.417	0.301	0.319	0.333	0.401	0.282	0.375	0.291	0.381	0.261	0.281	0.885	0.711
1st Count	16	19	3	7	3	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2nd Count	15	14	16	13	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Sum of last and 2nd	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

- Sensorformer는 다양한 고차원 시계열 데이터셋에서 기존 Transformer 대비 예측 정확도 향상
- 교차 패치 어텐션과 글로벌 패치 압축이 장기 의존성과 변수 간 상관 관계 학습에 기여
- Ablation study에서 패치 압축이나 교차 어텐션 제거 시 성능 하락, 설계 중요성 확인
- 계산 효율성 측면에서도 기존 Transformer보다 메모리 사용량과 연산량 감소
- 도메인별 분석에서 산업 센서와 금융 데이터에서 가장 큰 성능 향상, 일부 환경 데이터에서는 개선폭 작음

6. Insight

- Sensorformer는 교차 패치 어텐션과 글로벌 패치 압축으로 변수 간과 시간 간 의존성을 동시에 학습 가능
- 장기 예측에서 안정적 성능과 계산 효율성을 동시에 확보
- 패치 압축과 교차 어텐션 설계가 예측 성능에 큰 영향
- 도메인 특성에 따라 성능 차이 존재 산업 센서와 금융 데이터에서 효과가 특히 큼
- 향후 발전 가능성으로 더 다양한 데이터 유형과 센서 환경 확장, 실시간 예측, 다중 시계열 적용 연구 가능