Magic of Multiple Modalities in Clinical Risk Prediction

Department of Anesthesiology

Sandhya Tripathi, Christopher R King

Washington University in St. Louis
School of Medicine

Why should we care about postoperative complications?

- One in six patients undergoing elective surgery develop at least one postoperative complication [1].
- Reduce life expectancy and quality of life, increase treatment costs, need for critical care, inhospital mortality and readmission.

Electronic health records (EHRs) can be used for designing clinical risk prediction algorithms.

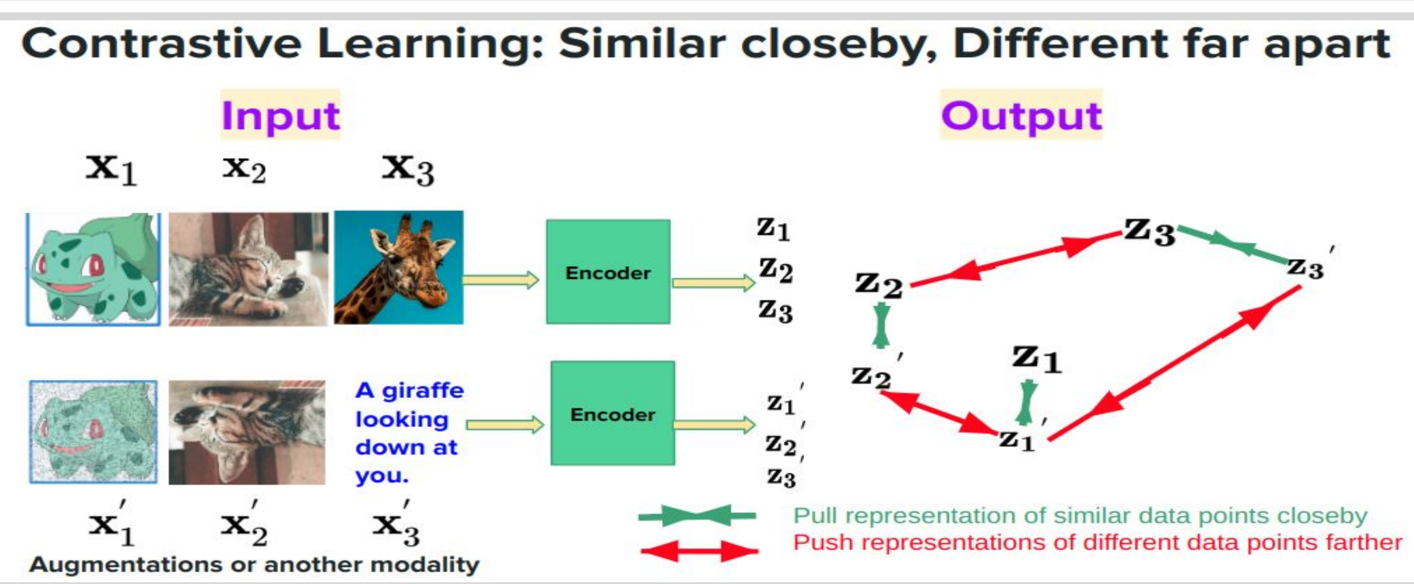
What are the challenges with current ML methods?

- Simultaneously learning from highly complex multi modal data including time series, unstructured notes, etc.
- Reduced effective sample size due to highly skewed and inconsistent quality labels.

What could be a possible solution?

- Self supervised multimodal learning particularly contrastive learning (CL) that bypasses the need for explicit supervision [2].

Objective: Use CL techniques to exploit similarities or complementary relationships between different modalities, such as home medications and problem lists for learning time series representations which can be used for diverse prediction tasks.



Existing work

- Concatenating tabular data with statistical features from time series data or complex time series representation.
- CL on imaging and tabular data [3].
- LLMs with only time series (UniTS, TinyTimeMixers, Lag-Llama) or only tabular (MediTab) data.

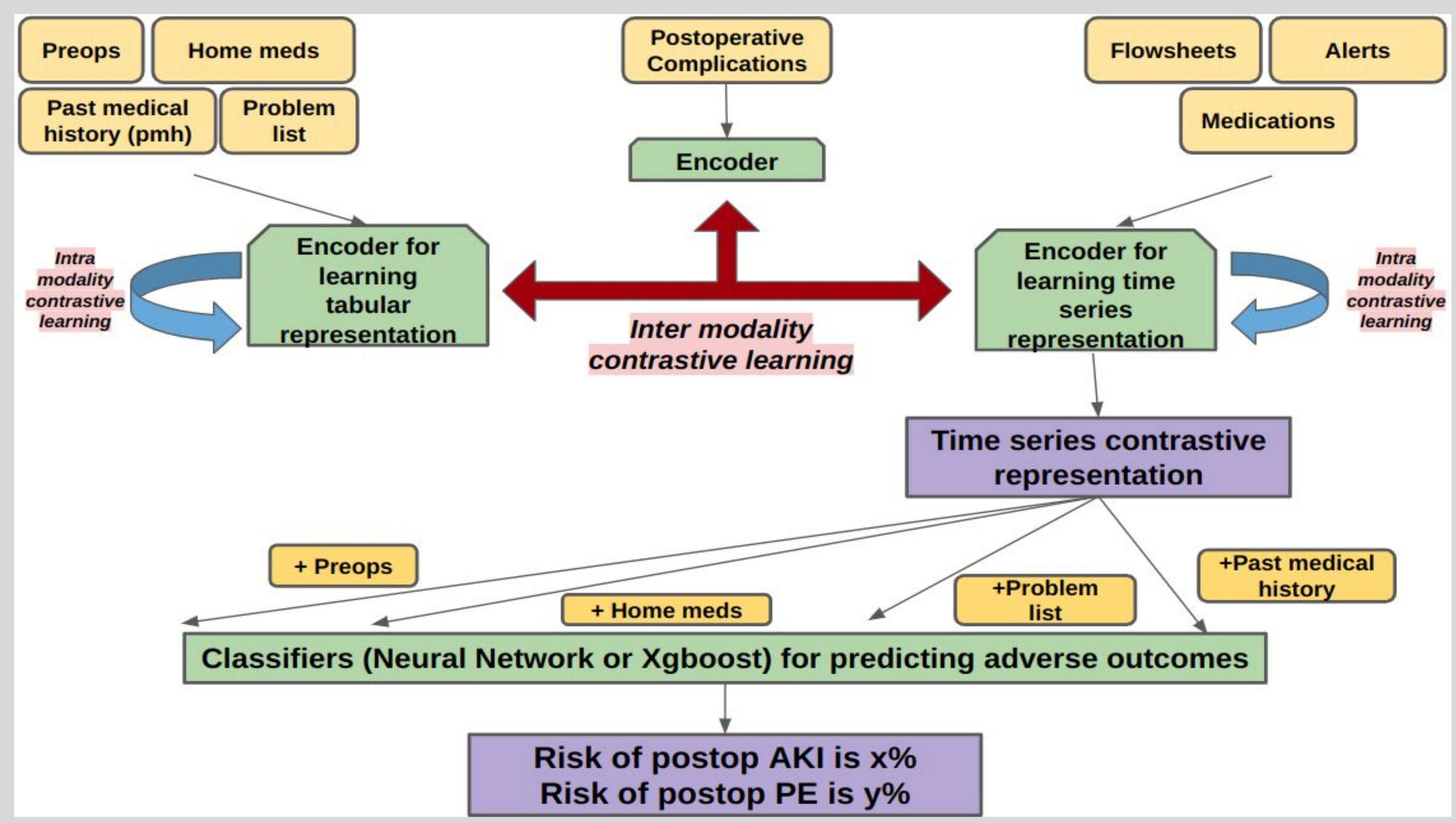
Experimental setup

- **Modalities of interest:** preoperative data, home medications, problem lists, past medical history, intraoperative data (flowsheets, medications and alerts), and postoperative complications.
- **Datasets**: ACTFAST-Epic with 69K train & 15K test data points[4].
- **Models**: TS2Vec for time series CL [5], SCARF for tabular CL [6], XGBoost for classifier training (with isotonic calibration post facto).
- Results reported as averaged over 5 random seed initializations.

Results

Modalities								Outcomes to predict																	
Case #	Preops	Home meds	PMH &PL	Intraop Flowsheets	Intraop Medications	Intraop Alerts	Postoperative complications	30-day mortality (incidence rate 2.1%)			Acute Kidney Injury KDIGO grade 2+ (incidence rate 5.6%)			Cardiac (new CHF or MI incidence rate 1.1%)			Pulmonary Embolism (PE, incidence rate 0.3%)			Pneumonia (incidence rate 0.3%)			Unplanned ICU (incidence rate 8%)		
								AUROC	AUPRC	Imp	AUROC	AUPRC	Imp	AUROC	AUPRC	Imp	AUROC	AUPRC	Imp	AUROC	AUPRC	Imp	AUROC	AUPRC	Imp
1	✅							0.918	0.336		0.939	0.705		0.866	0.082		0.713	0.018		0.724	0.011		0.889	0.545	
2	✅	✅						0.924	0.367	++	0.941	0.702	+	0.880	0.102	++	0.668	0.023		0.871	0.050	++	0.905	0.601	++
3	✅	✅	✅					0.920	0.435	+	0.940	0.696		0.887	0.104		0.859	0.102	++	0.891	0.067	++	0.904	0.591	
4				✅				0.839	0.158		0.799	0.196		0.750	0.031		0.602	0.005		0.788	0.015		0.921	0.634	
5				✅	✅			0.881	0.227	++	0.880	0.379	++	0.776	0.035	+	0.688	0.007	+	0.821	0.018	+	0.929	0.680	+
6				✅	✅	✅		0.886	0.234	+	0.885	0.381	++	0.779	0.039		0.682	0.007		0.828	0.019		0.928	0.672	
7	✅	✅	✅	✅				0.837	0.164		0.798	0.194		0.756	0.032		0.641	0.006		0.801	0.015		0.920	0.636	
8	✅	✅	✅	✅	✅			0.936	0.484	++	0.954	0.735	++	0.869	0.098	++	0.828	0.039	++	0.872	0.027	++	0.957	0.782	++
9	✅	✅	✅	✅	✅	✅		0.940	0.471	+	0.952	0.729		0.874	0.106	++	0.843	0.045	++	0.874	0.026		0.956	0.781	
10	✅	✅	✅	✅	✅	✅	✅	0.938	0.478		0.952	0.730		0.880	0.108	+	0.849	0.050	+	0.888	0.031	++	0.956	0.780	

Table reading notes: + and ++ in **Imp** column denotes at least second decimal improvement without last data modality (previous row) in only AUROC/AUPRC and both AUROC and AUPRC respectively. Intraop alerts and postoperative complications are used only during representation learning and not in classifier training. PMH: Past medical history, PL: Problem list



Conclusion

- Adding time series modality improves performance for 30-day mortality, AKI and Unplanned ICU (Comparing Case# 3 and 10).
- Both home meds and intraop medications modality lead to better representations and further improved classifier performance (Case#2, 5, 8).

Limitations

- No external dataset validation.
- Postoperative complications data modality only uses an MLP encoder so might not be very expressive (no CL).

Future directions

- Patching techniques for tokenizing & foundation models for time series in CL.
- Evaluation on more datasets (MOVER from UC Irvine and latest data extract from BJH).
- Using tabular data representation (instead of raw tabular data) for classifier training.

References

1) "Global patient outcomes after elective surgery: prospective cohort study in 27 low-, middle-and high-income countries." BJA: British Journal of Anaesthesia 117, no. 5 (2016): 601-609.
2) Muse, E.D. and Topol, E.J., 2024. Transforming the cardiometabolic disease landscape: Multimodal AI-powered approaches in prevention and management. Cell Metabolism.
3) Hager, P., Menten, M.J. and Rueckert, D., 2023. Best of both worlds: Multimodal contrastive learning with tabular and imaging data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 23924-23935).
4) Xue, B., Li, D., Lu, C., King, C.R., Wildes, T., Avidan, M.S., Kannampallil, T. and Abraham, J., 2021. Use of machine learning to develop and evaluate models using preoperative and intraoperative data to identify risks of postoperative complications. JAMA network open, 4(3), pp.e212240-e212240.
5) Yue, Z., Wang, Y., Duan, J., Yang, T., Huang, C., Tong, Y. and Xu, B., 2022, June. Ts2vec: Towards universal representation of time series. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 8, pp. 8980-8987).
6) Bahri, D., Jiang, H., Tay, Y. and Metzler, D., 2021. Scarf: Self-supervised contrastive learning using random feature corruption. arXiv preprint arXiv:2106.15147.