Ira A. Fulton Schools of Engineering
Arizona State University

CASCADE
ASU Center for Assured and SCAlable Data Engineering

SDMA: Saliency Driven Mutual Cross Attention for Multi-Variate Time Series

Yash Garg, and K. Selçuk Candan

International Conference on Pattern Recognition

ICPR 2020

Milan, Italy

January 2020

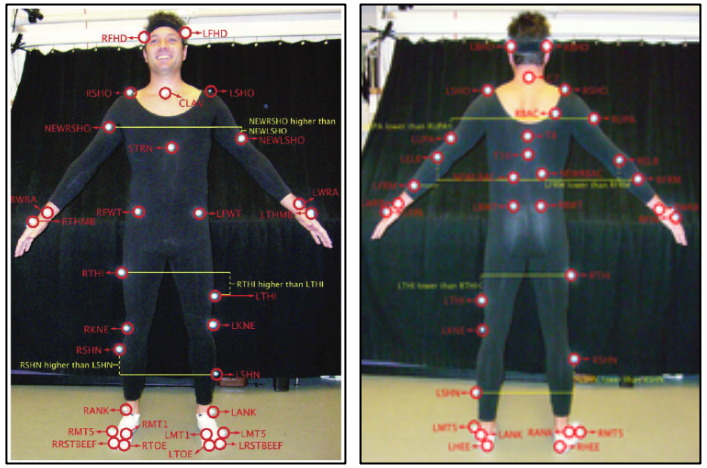
RESEARCH SUPPORTED BY:
NSF #1318788: "*Data Management for Real-Time Data Driven Epidemic Spread Simulation*"
NSF #1339835: "*E-SDMS: Energy Simulation Data Management System Software*"
NSF #1610282: "*DataStorm: A Data Enabled System for End-to-End Disaster Planning and Response*"
NSF #1629888: "*GEARS: An Infrastructure for Energy-Efficient Big Data Research on Heterogeneous and Dynamic Data*"
NSF #1633381: "*BIGDATA: Discovering Context- Sensitive Impact in Complex Systems*"
NSF #1827757: "PFI-RC: *Data-Driven Services for High Performance and Sustainable Buildings*"
NSF #1909555: "*pCAR: Discovering and Leveraging Plausibly Causal (p-causal) Relationships to Understand Complex Dynamic Systems*"
"*FourCModelling*": "*EUH2020 Grant Agreement No. 690817*"

emitlab

Motivation

SALIENCY-DRIVEN MUTUAL CROSS ATTENTION

- IoT systems record multiple attributes for time
- Time series may contain noisy data
  - Potentially deteriorating the model accuracy



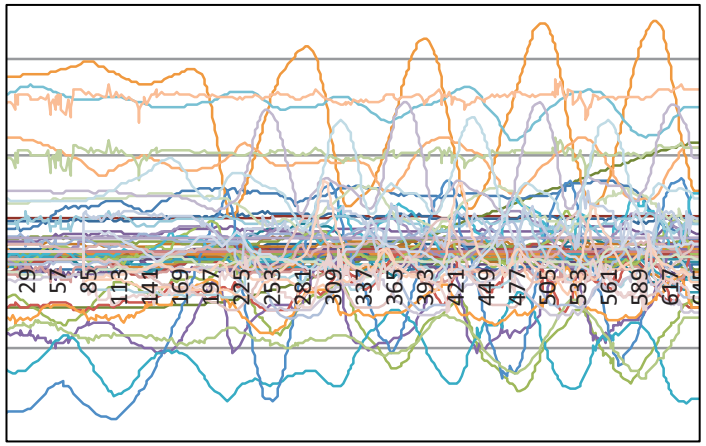
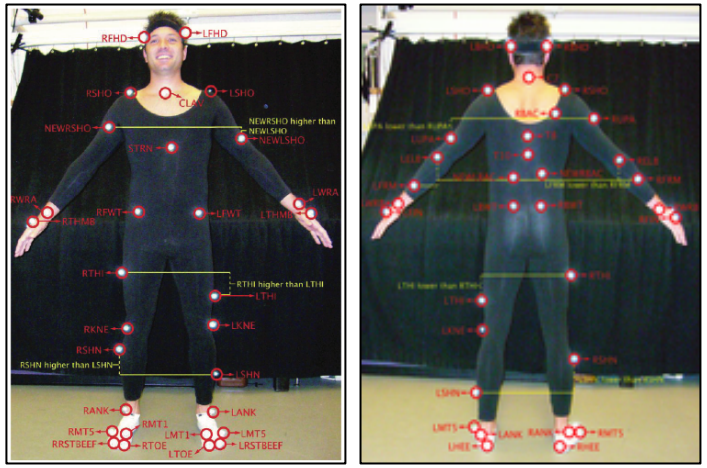


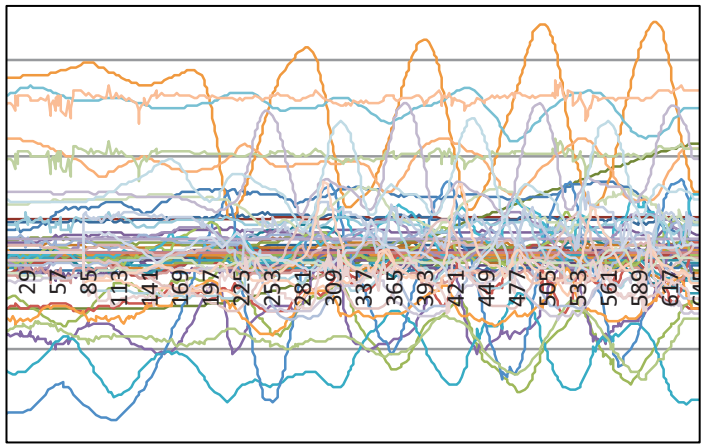
Image: CMU Motion Capture Dataset, 2015: http://mocap.cs.cmu.edu/marker_images/

Challenges

SALIENCY-DRIVEN MUTUAL CROSS ATTENTION

- How do we identify *salient* information; separating the noise?
- Can we generate a complementary input modality (saliency)?
- How to design network to leverage both input and saliency data?





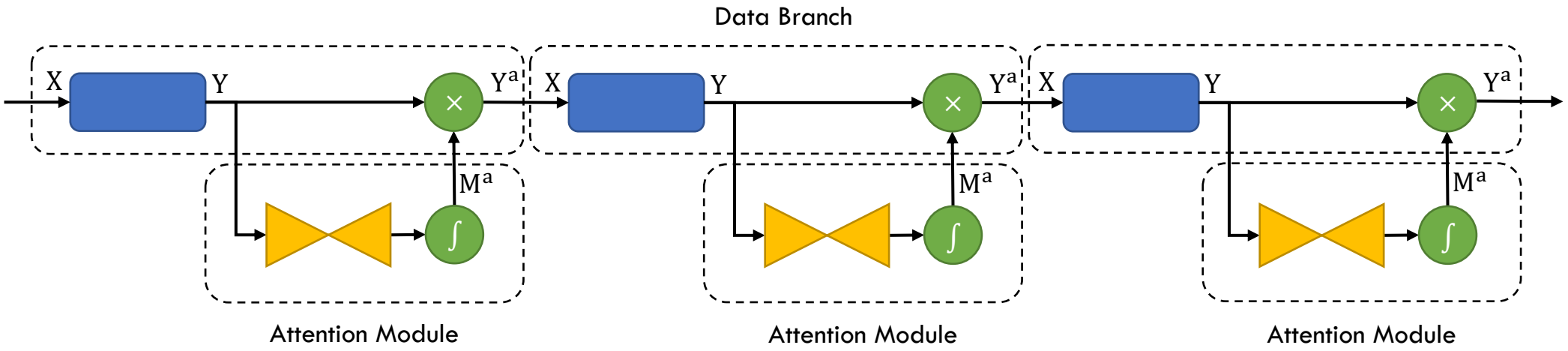
Problem Statement

PROBLEM SDMA AIMS TO SOLVE

Can we **HIGHLIGHT** the **RELEVANT** and **SUPPRESS** the **IRRELEVANT** information?

The Conventional Way

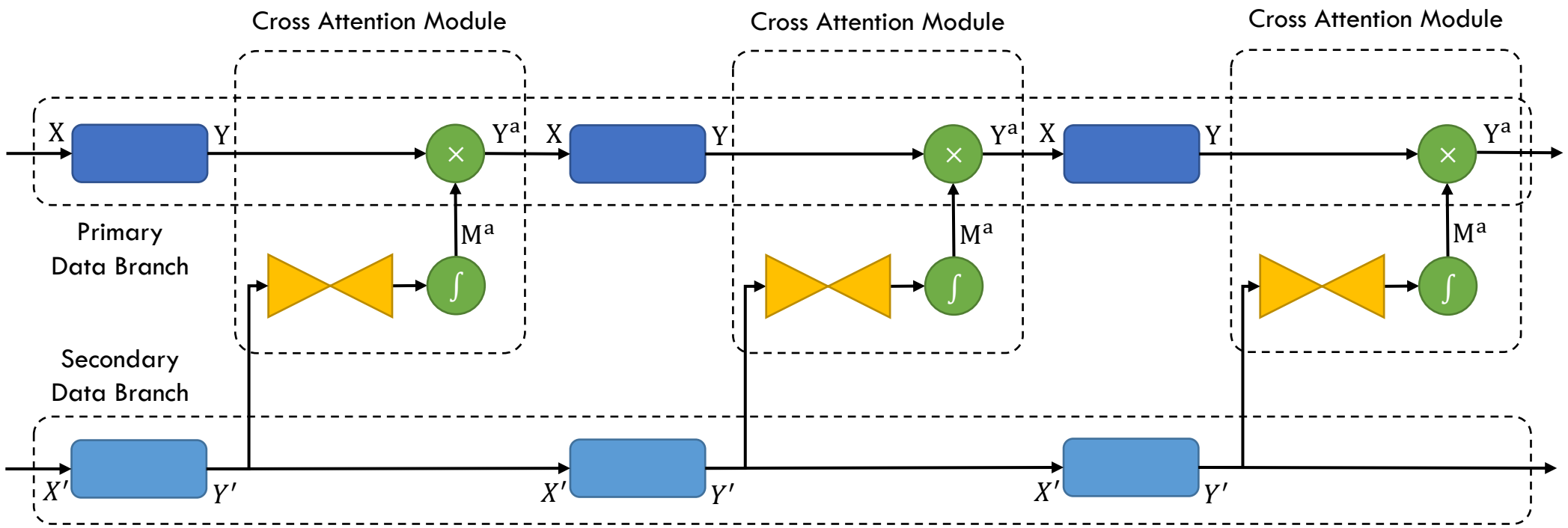
Finding important information through an **"attention"** block [1]



[1] Bahdanau, D et al, "Neural Machine Translation by Jointly Learning to Align and Translate." ICLR 2015

The Cross-Attention Way

Finding important information through secondary modality: **"cross attention"** block [1]

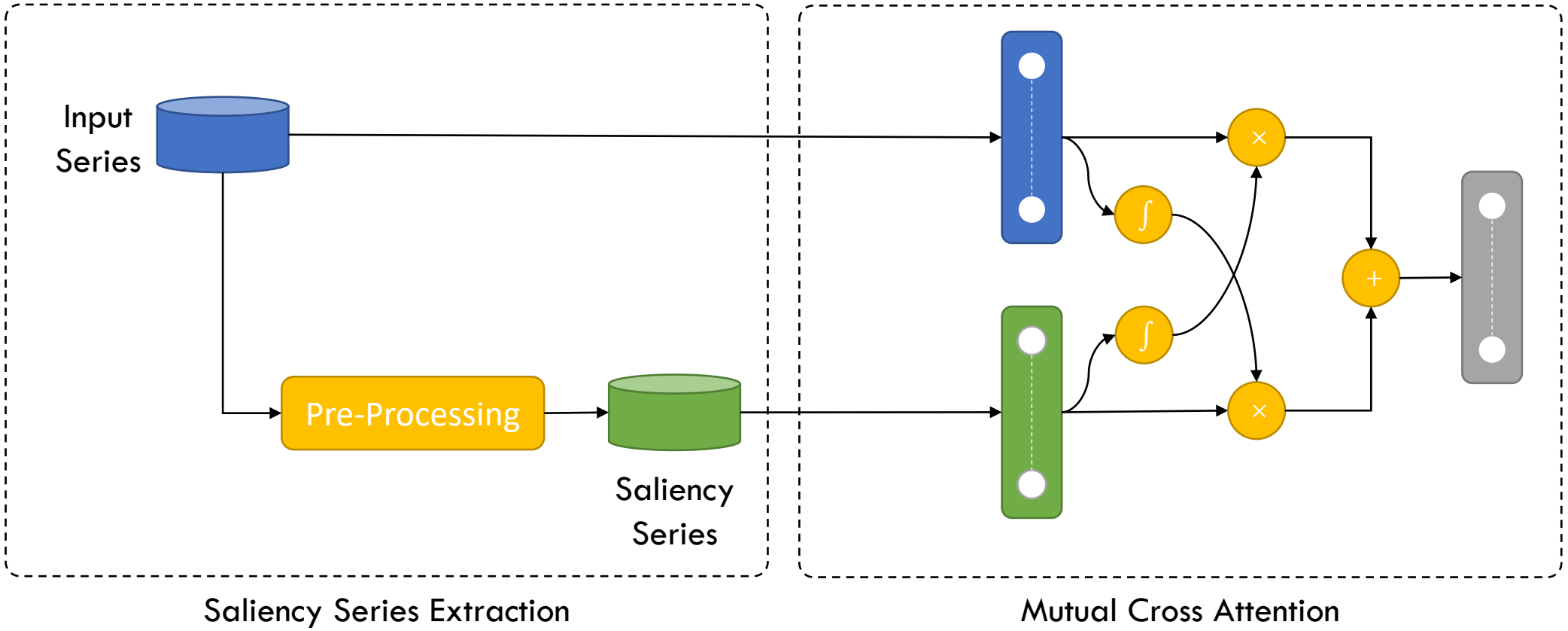


Secondary modality may itself have noisy information

[1] Mohla et al., "FusAtNet: Dual Attention based Multi-Model Fusion Network" CVPR Workshops 2020.

The SDMA Framework

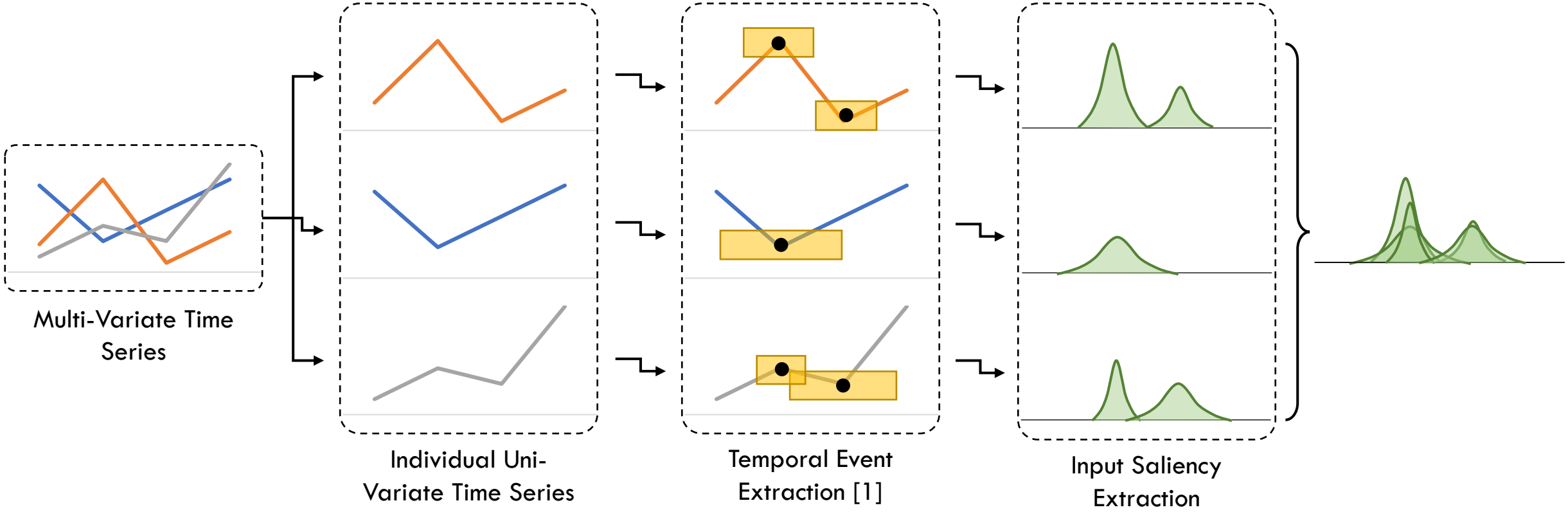
SDMA is a two-step framework: Saliency Extraction and Mutual Cross Attention



Secondary modality is not always available

Saliency Series Extraction

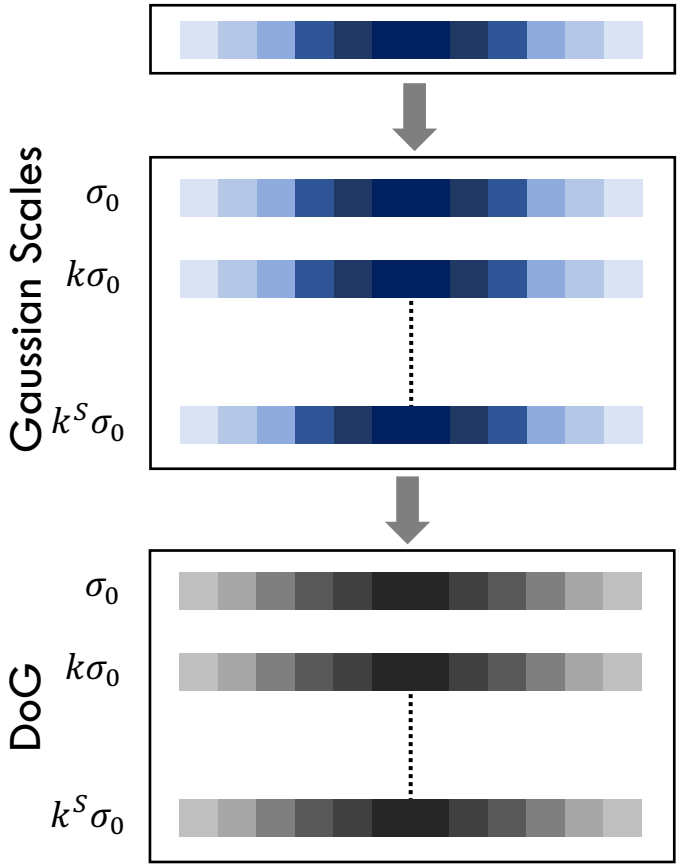
Highlighting the relevant information and suppressing the irrelevant



[1] Candan, K. Selçuk, et al. "sDTW: Computing DTW Distances using Locally Relevant Constraints based on Salient Feature Alignments." VLDB 2012

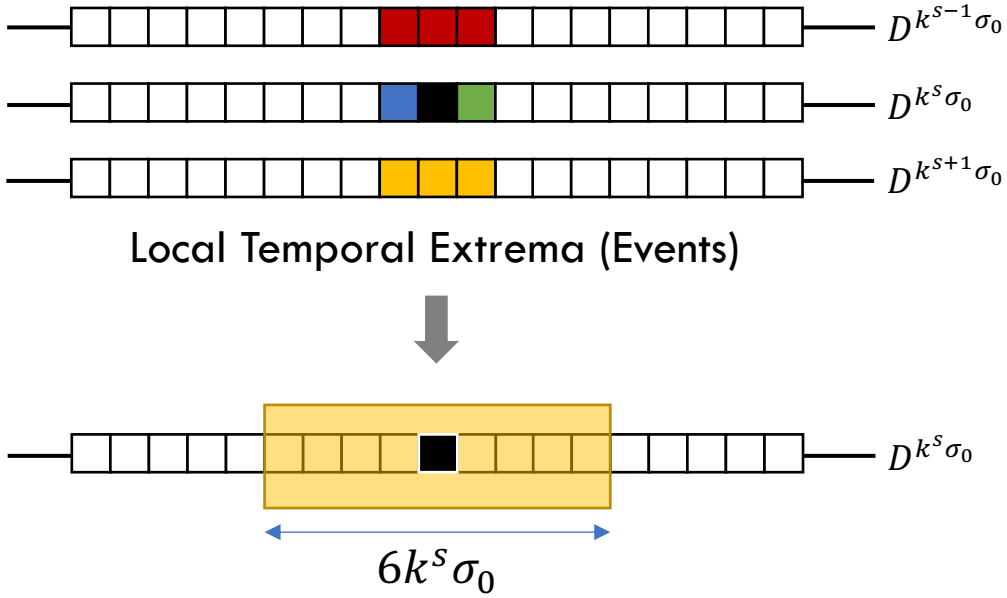
Temporal Event Extraction

Extracting robust multi-scale saliency temporal subsequences [1,2]



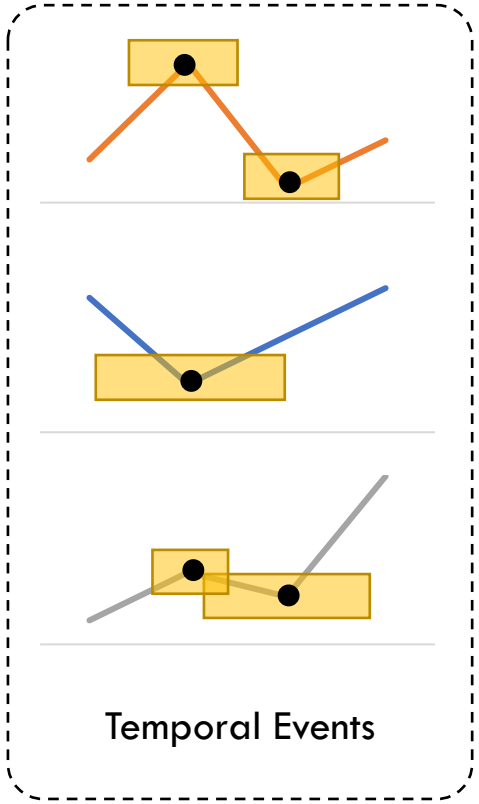
Eliminating noisy information

For a given Resolution and Scale



Local Temporal Extrema (Events)

Temporal Scope of Events

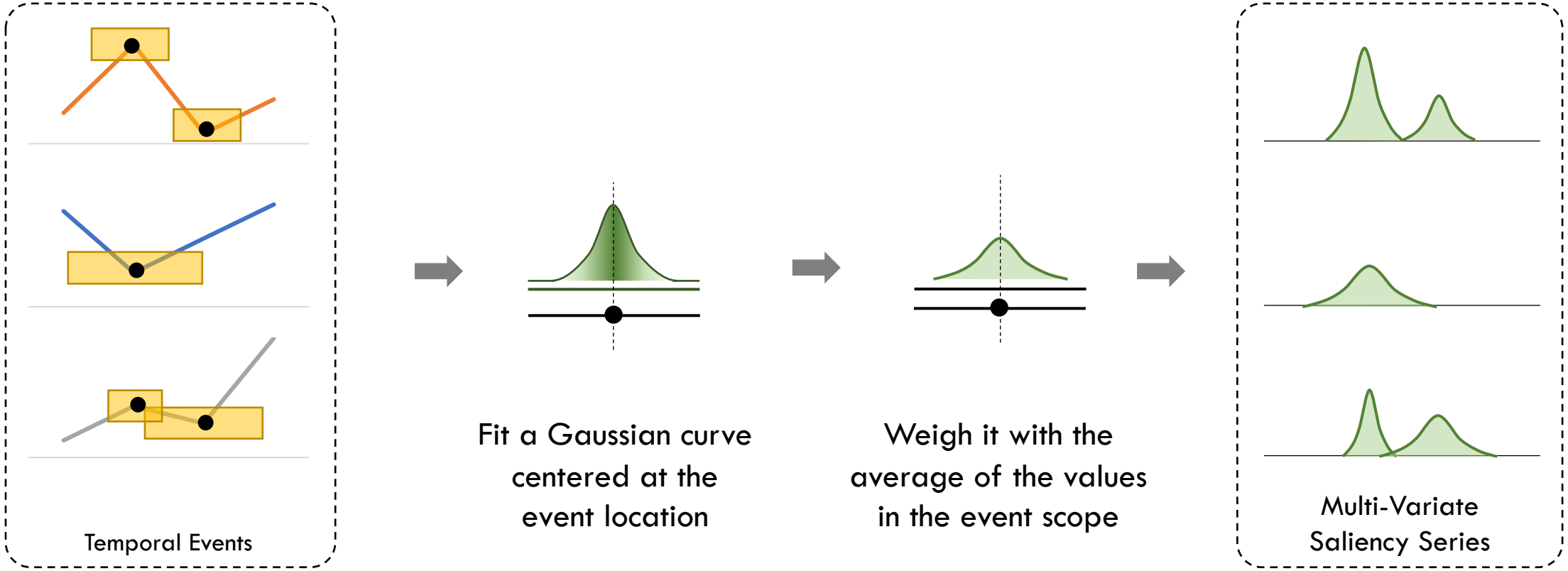


[1] Lowe, David "Distinctive Image Features from Scale-Invariant Keypoints" IJCV 2004.

[2] Candan, K. Selçuk, et al. "sDTW: Computing DTW Distances using Locally Relevant Constraints based on Salient Feature Alignments." VLDB 2012

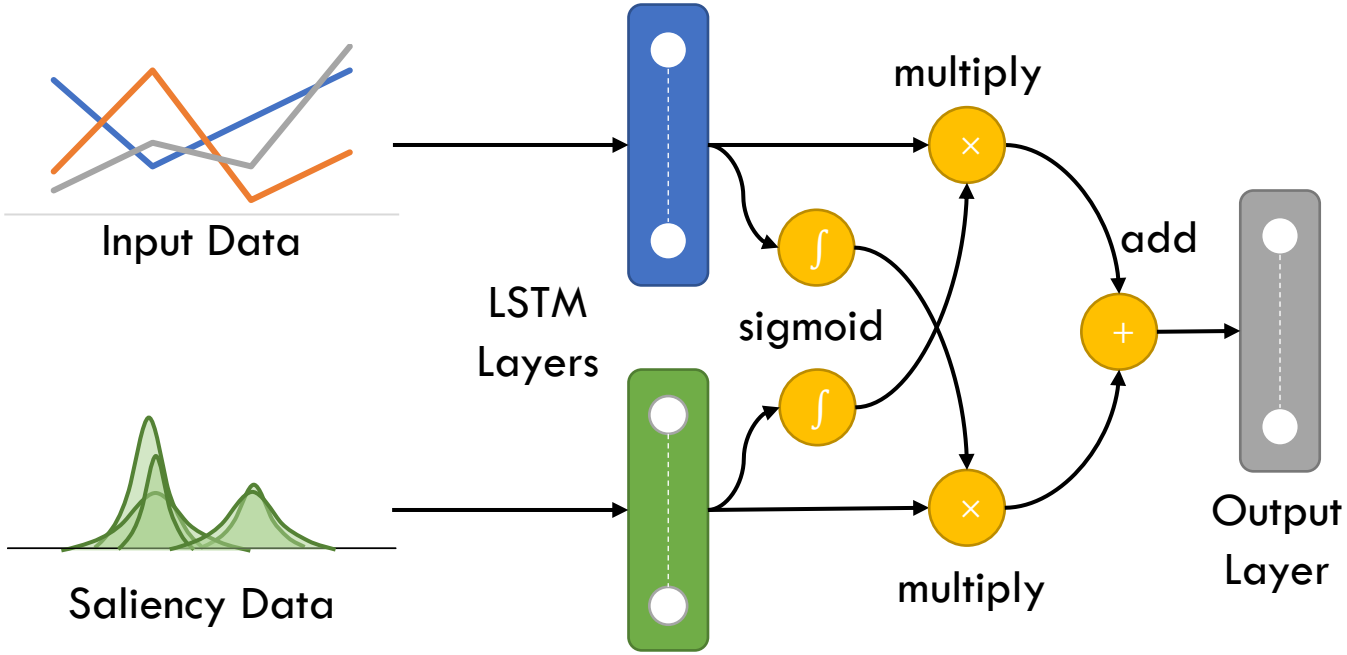
Input Saliency Extraction

Transforming temporal event in saliency series



Single Layer Configuration

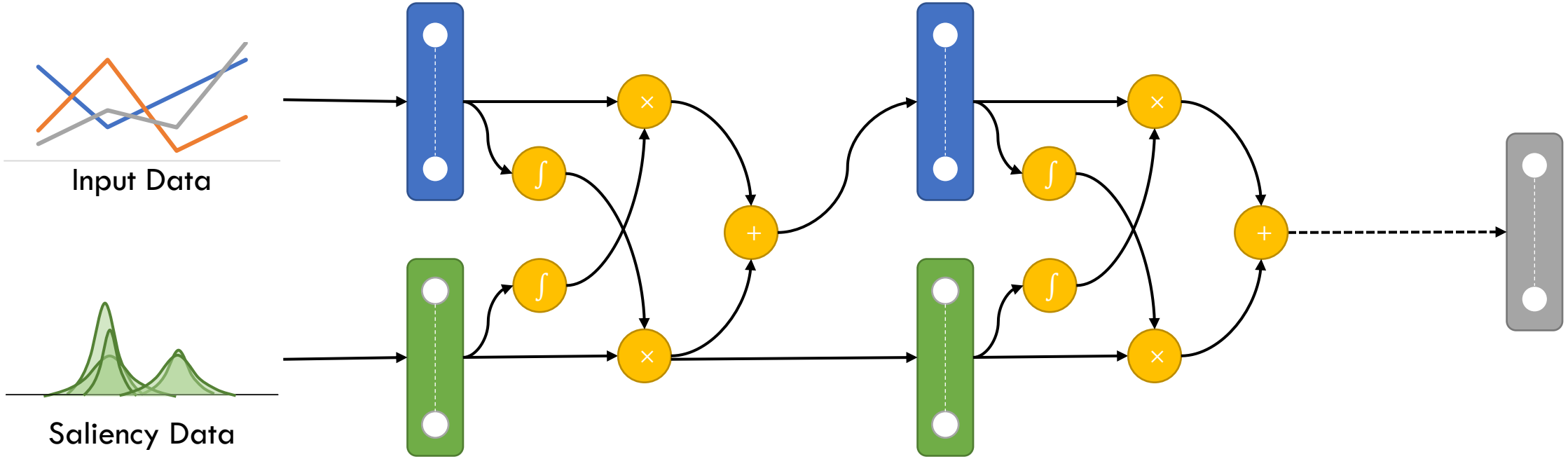
MUTUAL CROSS ATTENTION



Saliency series provides a rough approximation of key information of interest

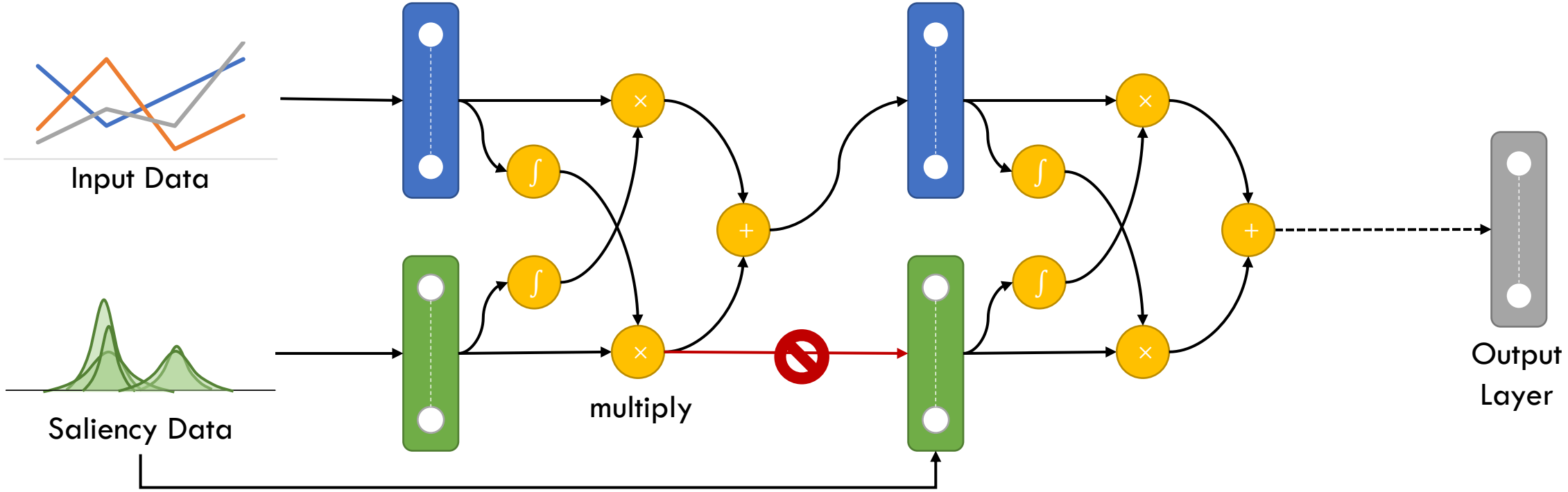
Multi Layer Configuration (Type1)

MUTUAL CROSS ATTENTION



Multi Layer Configuration (Type2)

MUTUAL CROSS ATTENTION

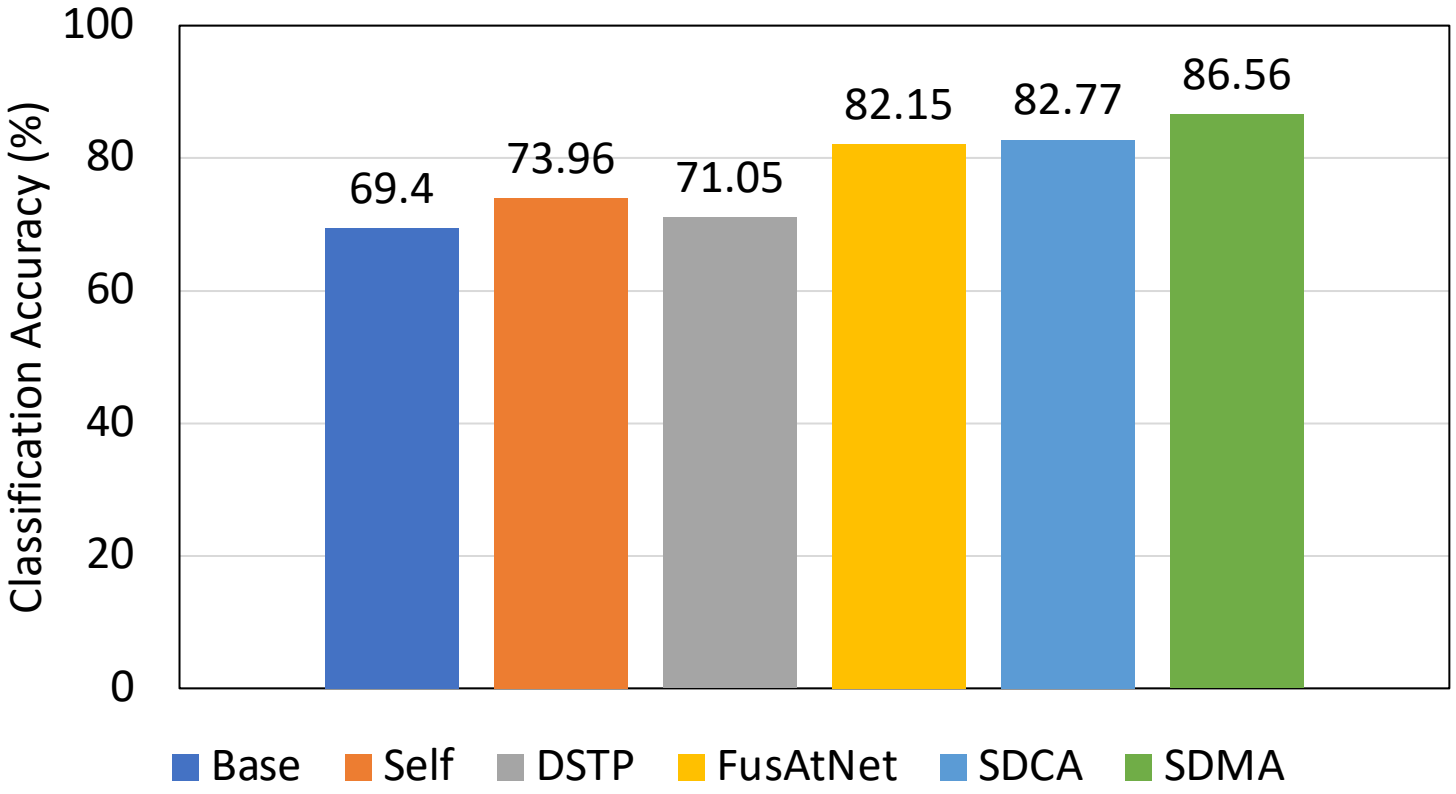


Experimental Evaluation

- Architectures considered:
  - Single Layer, Multi Layer (Type 1, Type 2)
- Datasets considered:
  - MOCAP, AUSLAN, SML
- State-of-the-art considered:
  - Self, DSTP, FusAtNet, SDCA (Naïve Cross Attn Version)

[Self] Luong et al., "Effectives Approaches to Attention-based Neural Machine Translation", EMNLP 2015
[DSTP] Liu et al., "A Dual-Stage Two-phase Attention-based Recurrent Neural Network for Long-Term and Multi-Variate Time Series Prediction", Expert Systems and Applications2020
[FusAtNet] Mohla et al., "FusAtNet: Dual Attention based Multi-Model Fusion Network" CVPR Workshops 2020.
[SDCA] Self Implementation of Naïve Cross Attention

Classification Accuracy (AUSLAN, Type 2)



[Self] Luong et al., "Effectives Approaches to Attention-based Neural Machine Translation", EMNLP 2015
[DSTP] Liu et al., "A Dual-Stage Two-phase Attention-based Recurrent Neural Network for Long-Term and Multi-Variate Time Series Prediction", Expert Systems and Applications 2020
[FusAtNet] Mohla et al., "FusAtNet: Dual Attention based Multi-Model Fusion Network" CVPR Workshops 2020.
[SDCA] Self Implementation of Naïve Cross Attention

Conclusion

- Intelligently extracts saliency series to highlight relevant information
- Novel mutual cross attention combines input and saliency series

Thank You