A Deep Neural Network for Unsupervised Anomaly Detection and Diagnosis in Multivariate Time Series Data

Haakon Jonsson

Motivation

- Monitoring complex systems is important for many industries.
- This generates a lot of multivariate time series data.
- Anomaly detection allows operators to take action early which prevent expensive downtime.
- Systems will often recover from Short term anomalies by themselves so we want to provide operators with a severity score.
- There is little to no anomaly labeled data.

Challenges

- Anomaly detection
- Severity score
- Root cause detection
- Robustness to noise.

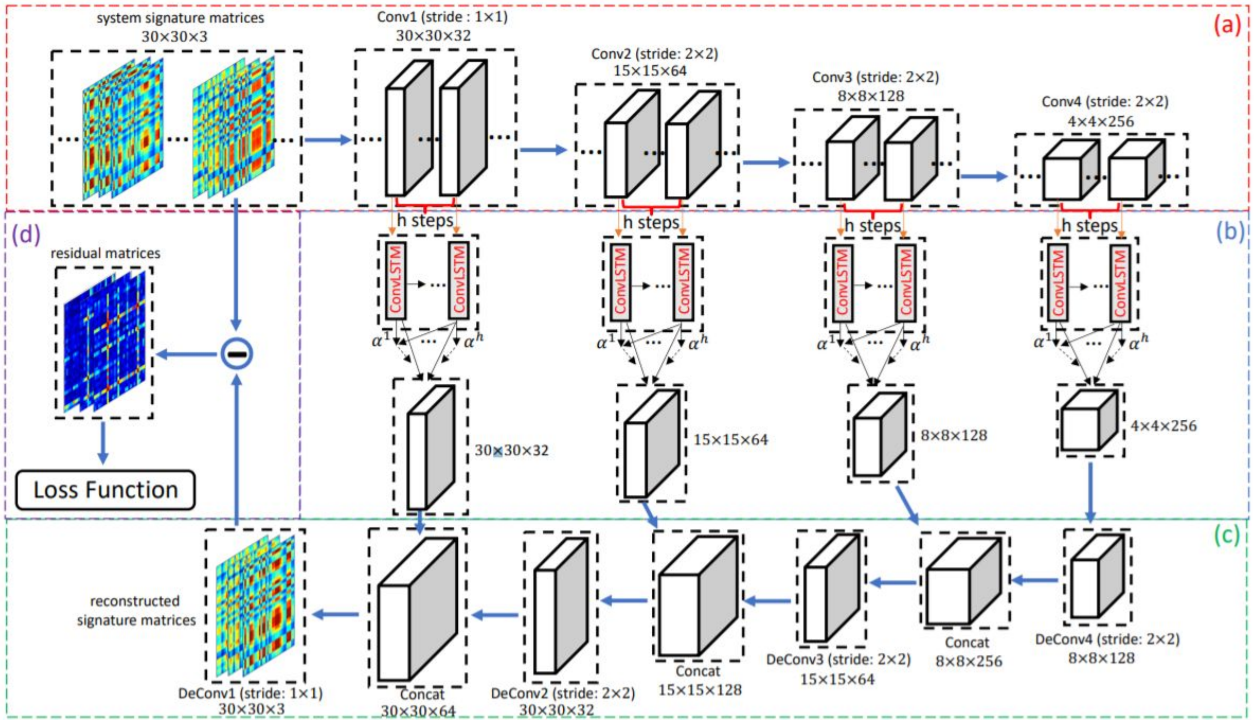
Related work

- k-Nearest Neighbor (kNN)
- Clustering Models
- One-Class SVM (OC-SVM)
- Autoregressive Moving Average
- Correlation Methods
- Ensemble Methods
- Deep Autoencoding Gaussian Mixture Model (DAGMM)
- LSTM encoder-decoder (LSTM-ED)

They cannot jointly consider the temporal dependency, noise resistance, and the interpretation of severity of anomalies.

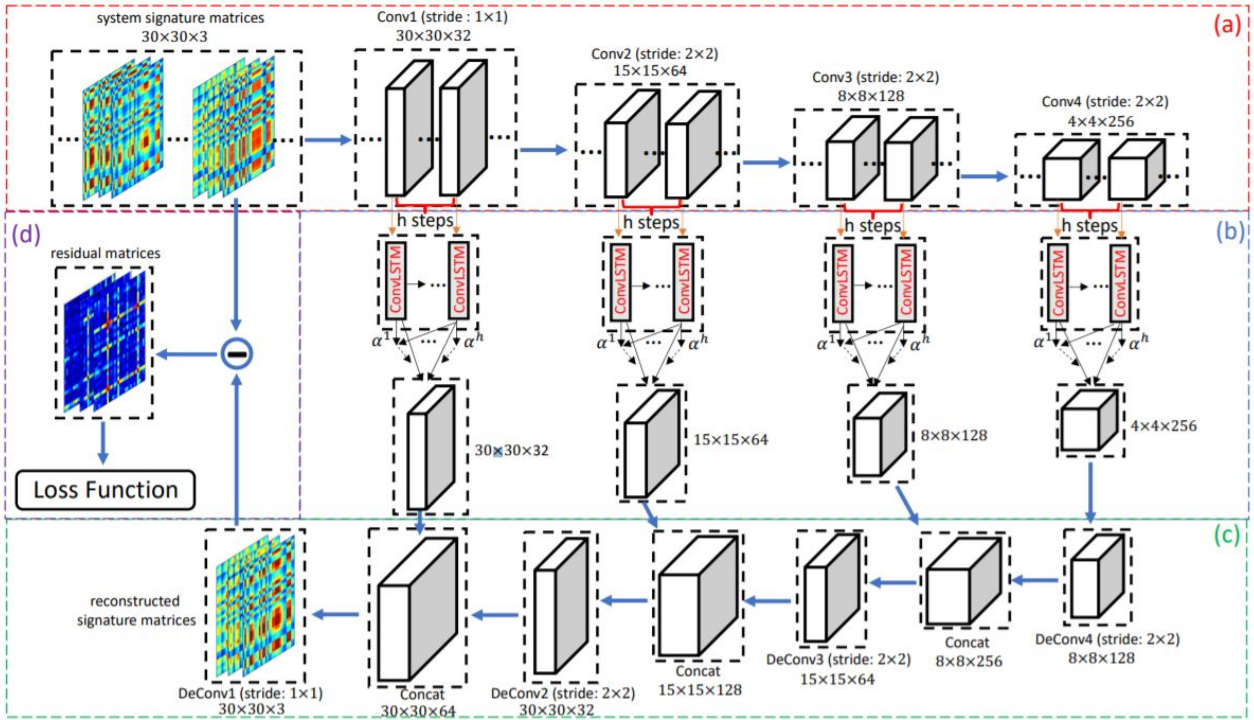
Methods

- Compute signature (correlation) matrices for various timescales.
- Convolutional encoder
- ConvLSTM
- Attention mechanism
- Convolutional decoder
- Residual matrices



Methods

- Anomalies causes more residual.
- Residual matrices from different timescales can be used to estimate severity.
- Top-k series identified as root causes.



Experiment

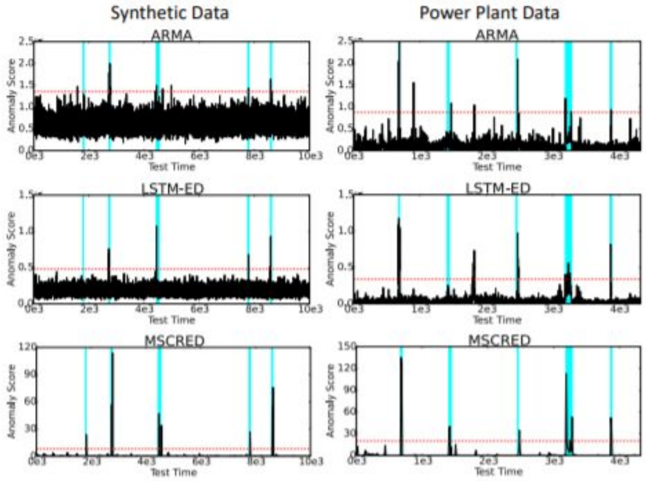
- A synthetic dataset and a power plant dataset.
- Baseline
  - OC-SVM
  - DAGMM
  - ARMA
  - LSTM-ED
- MSCRED variants
  - Attention module missing.
  - Attention and 2 ConvLSTM layers missing.
  - Attention and 3 ConvLSTM layers missing.

Results

MSCRED improves on the state of the art for anomaly detection.

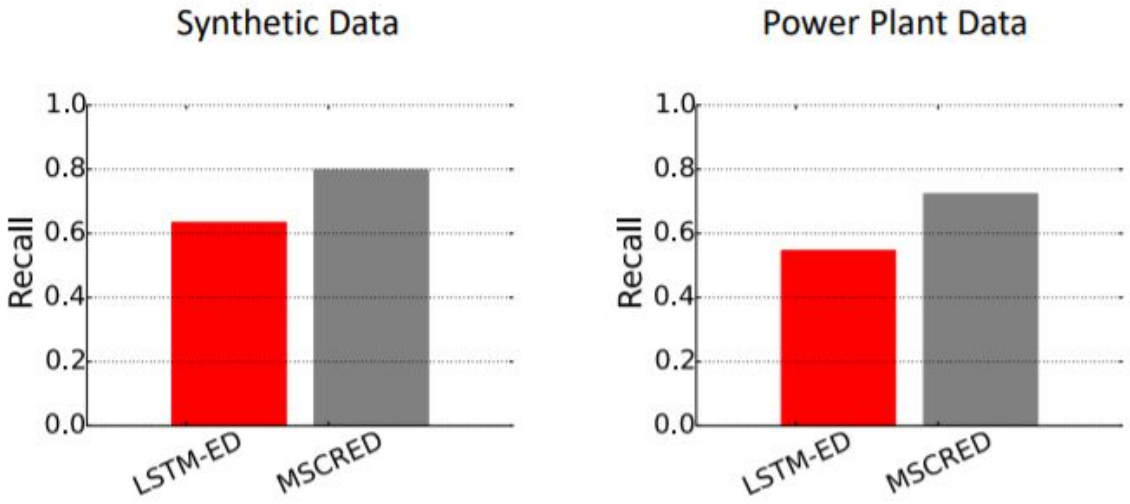
Table 2: Anomaly detection results on two datasets.

Method	**Synthetic Data**			**Power Plant Data**		
	Pre	Rec	F_1	Pre	Rec	F_1
OC-SVM	0.14	0.44	0.22	0.11	0.28	0.16
DAGMM	0.33	0.20	0.25	0.26	0.20	0.23
HA	0.71	0.52	0.60	0.48	0.52	0.50
ARMA	0.91	0.52	0.66	0.58	0.60	0.59
LSTM-ED	<u>1.00</u>	<u>0.56</u>	<u>0.72</u>	<u>0.75</u>	<u>0.68</u>	<u>0.71</u>
$\text{CNN}_{ConvLSTM}^{ED(4)}$	0.37	0.24	0.29	0.67	0.56	0.61
$\text{CNN}_{ConvLSTM}^{ED(3,4)}$	0.63	0.56	0.59	0.80	0.72	0.76
$\text{CNN}_{ConvLSTM}^{ED}$	0.80	0.76	0.78	0.85	0.72	0.78
MSCRED	**1.00**	**0.80**	**0.89**	**0.85**	**0.80**	**0.82**
Gain (%)	–	30.0	23.8	13.3	19.4	15.5



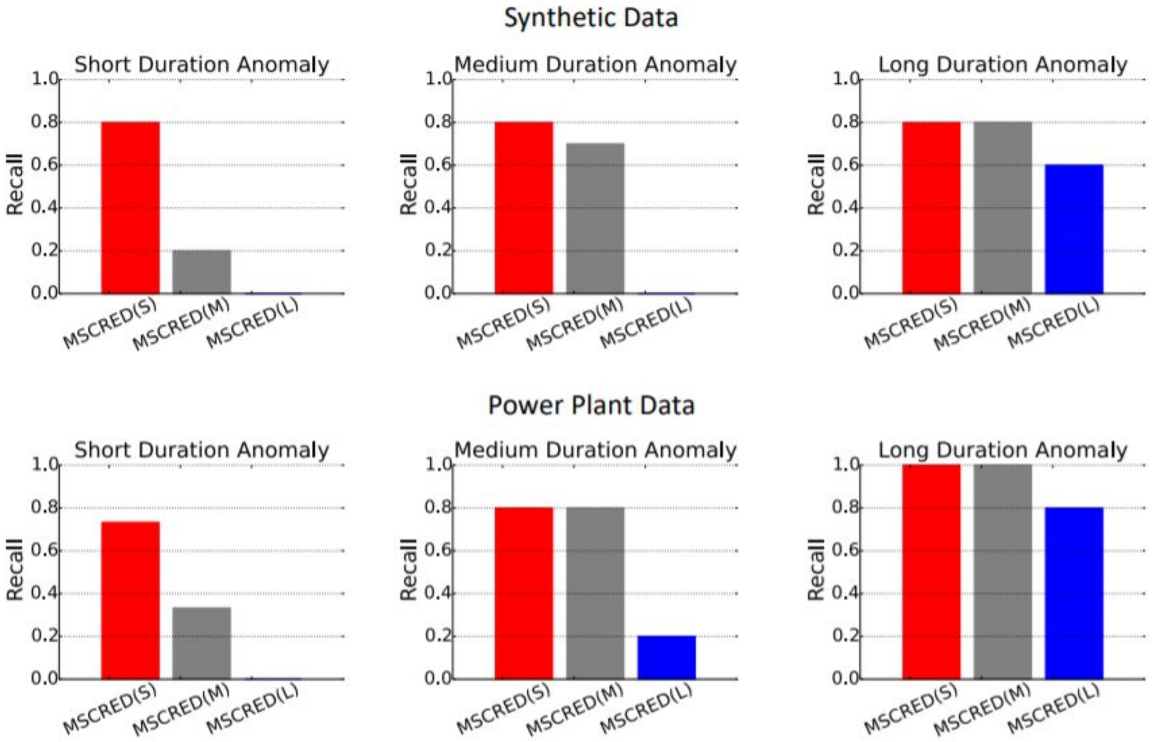
Results

MSCRED beat LSTM-ED on root cause identification.



Results

MSCRED can be used to score anomaly severity.



Conclusions

- MSCRED improves on anomaly detection.

- MSCRED imporves on root cause analysis.

- MSCRED can be used to score anomaly severity

- MSCRED is robust to noise.

- The authors seems to have achieved what it set out to do.
- Could have explained why they used CNN on the signature matrices.