Latent Independent Excitation for Generalizable Sensor-based Cross-Person Activity Recognition (AAAI-21)

Hangwei Qian, Sinno Jialin Pan, Chunyan Miao

Nanyang Technological University, Singapore

Dec. 2020

Outline

Activity Recognition with Edge Devices

Ubiquitous edge devices

- Mobile phones
- Smart watches
- Smart glasses
- Sports bracelets

Activity Recognition with Edge Devices

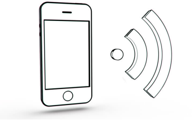
Numerous applications: elderly assistance, healthcare, fitness tracking, smart building, gaming, etc

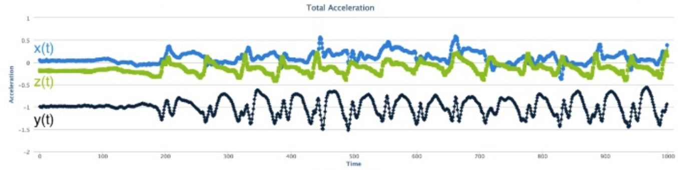
Photo Credit: https://www.betterhealth.vic.gov.au/health/HealthyLiving/Walking-the-benefits-for-older-people; https://www.digitalauthority.me/resources/healthcare-marketing/; https://lepfitness.co.uk/the-benefits-of-fitness-tracking/; https://blog.bosch-si.com/projects/bosch-singapore-campus-smart-building-concept-turned-reality/

Human Activity Recognition

A multi-class classification problem

- Input: data $X^{M \times N}$ collected from accelerometers
- Output: activity labels $y^{N \times 1} \in \{1, ..., n_c\}$





Outline

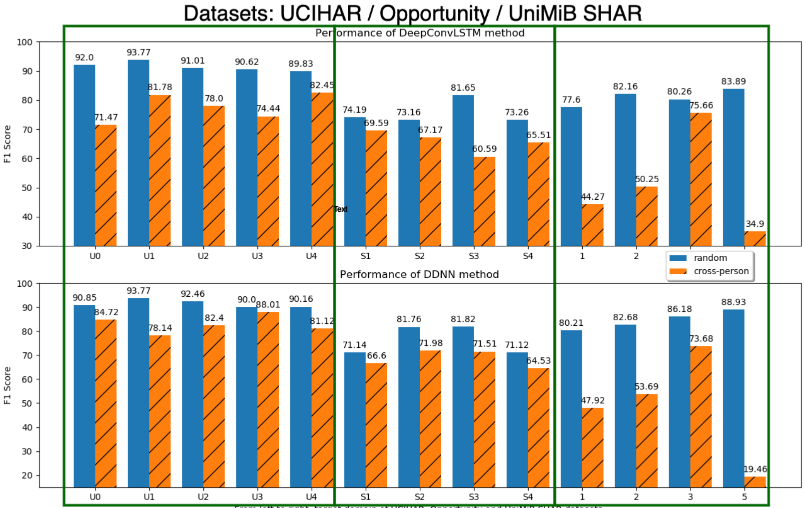
Performance Degradation Problem

Classification performance drops when a model is trained on young people while tested on elder people

Photo Credit: https://www.ramblers.org.uk/advice/facts-and-stats-about-walking.aspx
https://www.betterhealth.vic.gov.au/health/HealthyLiving/Walking-the-benefits-for-older-people

Performance Degradation Problem

- train/test split: overlapping VS non-overlapping
- Larger diversity → more severe performance drop



From left to right: target domain of UCIHAR, Opportunity and UniMiB SHAR datasets.

Performance Degradation Problem

Photo Credit: https://www.ramblers.org.uk/advice/facts-and-stats-about-walking.aspx
https://www.betterhealth.vic.gov.au/health/HealthyLiving/Walking-the-benefits-for-older-people

1. Domain gaps
   - Different ages, health status, movement patterns
   - Different environments, constraints, emergency situations
2. Dataset bias
   - Cover a subset of population

Performance Degradation Problem

Photo Credit: https://www.ramblers.org.uk/advice/facts-and-stats-about-walking.aspx
https://www.betterhealth.vic.gov.au/health/HealthyLiving/Walking-the-benefits-for-older-people

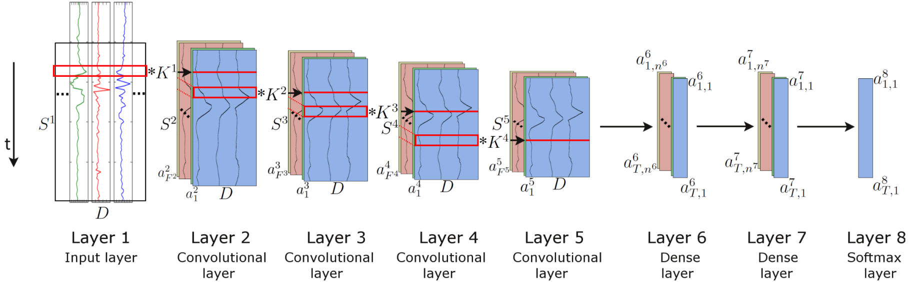
1. Domain gaps
   - Different ages, health status, movement patterns
   - Different environments, constraints, emergency situations
2. Dataset bias
   - Cover a subset of population

Q: How to train a model that is readily generalizable to unseen target domains?

Outline

Improving Model Capacity

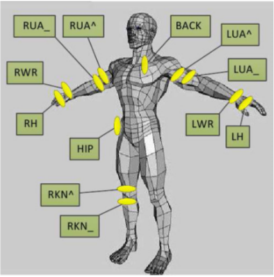
CNN_Yang model and DeepConvLSTM model [2]

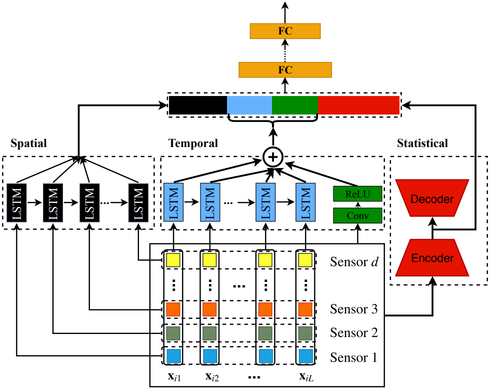


Improving Model Capacity

Distribution-embedded Deep Neural Network [3]

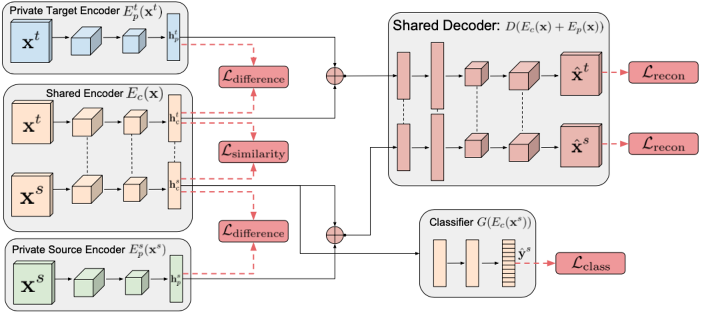
- Temporal module
- Statistical module: moments features in deep model
- Spatial module with a different perspective: spatial correlations and constraints between sensor placements





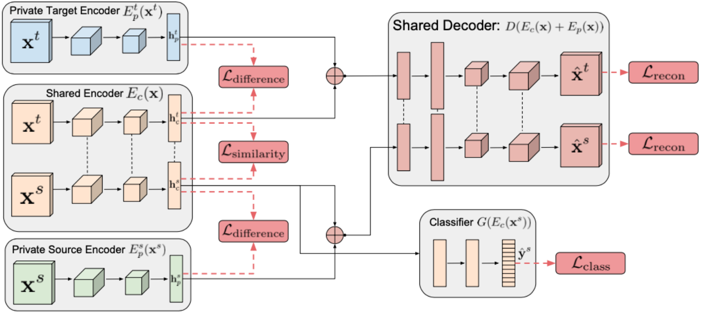
Reducing Domain Gaps

Domain Separation Networks [1]: shared-private network structure



Reducing Domain Gaps

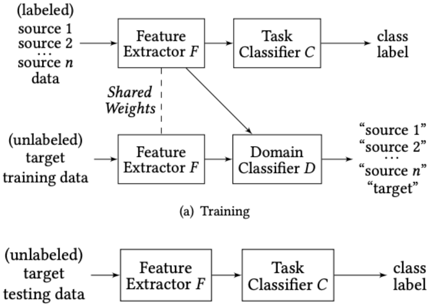
Domain Separation Networks [1]: shared-private network structure



- ✗ single source and target domain
- ✗ target domain data needs to be accessible at training phase
- ✗ target domain is fixed, otherwise the model should be re-trained

Reducing Domain Gaps

Multi-source domain adaptation with weak supervision [4]
Assume that the target domain's label proportion is available

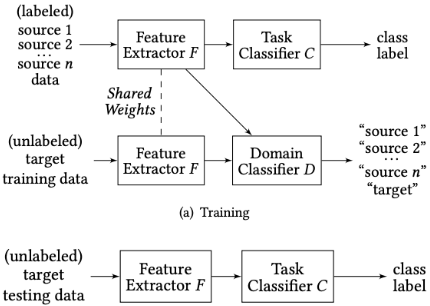


(a) Training

(b) Testing

Reducing Domain Gaps

Multi-source domain adaptation with weak supervision [4]
Assume that the target domain's label proportion is available



(a) Training

(b) Testing

✗ target domain data and label proportion need to be accessible at training phase

✗ target domain is fixed, otherwise the model should be re-trained

Outline

The Proposed Method

Input: labeled data from D source domains $\{(X^d, y^d) \sim \mathbb{P}^d(x, y)\}_{d=1}^{D}$
Goal: to train a deep model f to generalize well to unseen target domain data $X^{\tilde{d}}$

Generalizable Independent Latent Excitation (GILE) model

- Backbone: VAE
- Feature disentanglement
- Independent excitation mechanism

✔ Does not require access to any information from target domain
✔ Readily generalizable to any unseen target domain

Feature Disentanglement

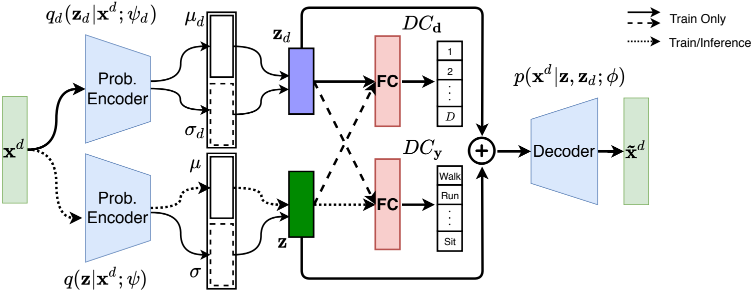
Disentangle features into two types

- domain-agnostic features z
  - to capture the commonality of conducting the same activity among different people
  - to generalize well to unseen target domains
- domain-specific features z_d of domain d

Feature Disentanglement

Disentangle features into two types

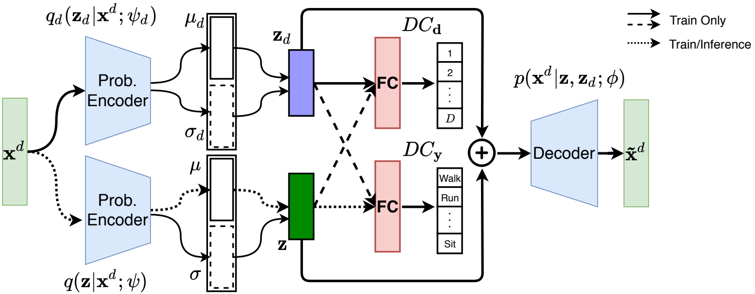
- domain-agnostic features z
  - to capture the commonality of conducting the same activity among different people
  - to generalize well to unseen target domains
- domain-specific features z_d of domain d



Feature Disentanglement

VAE backbone loss:

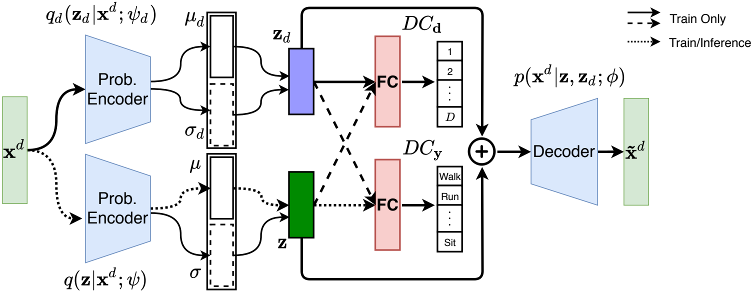
$$\begin{aligned}\mathcal{L}_{\mathrm{elbo}} &= \mathbb{E}_{\mathrm{d},\mathrm{q}_\mathrm{d}(\mathrm{z}_\mathrm{d}|\mathrm{x}^\mathrm{d};\psi_\mathrm{d}),\mathrm{q}(\mathrm{z}|\mathrm{x}^\mathrm{d};\psi)}[\log \mathrm{p}(\mathrm{x}^\mathrm{d}|\mathrm{z}_\mathrm{d},\mathrm{z};\phi)] \\ &\quad -\mathrm{KL}(\mathrm{q}_\mathrm{d}(\mathrm{z}_\mathrm{d}|\mathrm{x}^\mathrm{d};\psi_\mathrm{d})||\mathrm{p}(\mathrm{z}_\mathrm{d})) \\ &\quad -\mathrm{KL}(\mathrm{q}(\mathrm{z}|\mathrm{x}^\mathrm{d};\psi)||\mathrm{p}(\mathrm{z})),\end{aligned}$$



Feature Disentanglement

Two disentangling classifiers $\mathrm{DC_y}$ and $\mathrm{DC_d}$:

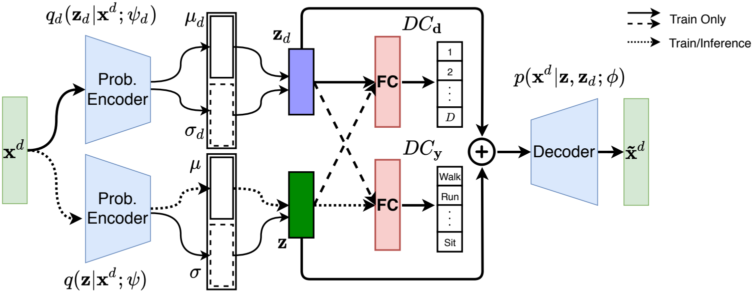
$$\mathcal{L}_{\mathrm{DC}} = \frac{1}{\mathrm{N_S}} \sum_{\mathrm{d}=1}^{\mathrm{D}} \sum_{\mathrm{i}=1}^{\mathrm{N_d}} [\ell(\mathrm{y_i^d}, \mathrm{DC_y}(\mathrm{z}; \mathrm{w_y})) + \ell(\mathrm{d}, \mathrm{DC_d}(\mathrm{z_d}; \mathrm{w_d}))]. \tag{1}$$



Independent Excitation Mechanism

Minimize the correlations between domain-agnostic and domain-specific features

$$\mathcal{L}_{\mathrm{IE}} = -\frac{1}{\mathrm{N_S}} \sum_{\mathrm{d}=1}^{\mathrm{D}} \sum_{\mathrm{i}=1}^{\mathrm{N_d}} [\ell(\mathrm{y_i^d}, \mathrm{DC_y}(\mathrm{z_d}; \mathrm{w_y})) + \ell(\mathrm{d}, \mathrm{DC_d}(\mathrm{z}; \mathrm{w_d}))]. \quad (2)$$



Experimental Settings

Datasets

- UCIHAR
  - 5 domains, 9 features
  - 6 daily activities: walking, sitting, laying, standing, walking upstairs, walking downstairs
- Opportunity
  - 4 domains, 77 features
  - 18 gestures: open / close dishwasher / fridge /drawer1 /door1 / drawer2 / door2 / drawer3, move cup, clean table, null
- UniMiB SHAR
  - 4 domains, 3 features
  - 9 types of daily living and 8 types of falls

Evaluation: Leave-One-Subject-Out

Measure: F1 score

Overall Results

Evaluations on UCIHAR and Opportunity dataset

Source	Target	VAE	β-VAE	DIVA	DDNN	DeepConvLSTM	CoDATS	GILE
1,2,3,4	0	51.87	53.31	75.00	<u>84.72</u>	71.47	81.27	**85.15**
0,2,3,4	1	44.70	44.37	77.18	78.14	**81.78**	55.63	<u>81.56</u>
0,1,3,4	2	64.22	62.17	71.61	<u>82.40</u>	78.00	77.42	**86.97**
0,1,2,4	3	36.91	49.21	81.87	<u>88.01</u>	74.44	60.57	**94.37**
0,1,2,3	4	39.07	58.28	79.68	81.12	<u>82.45</u>	66.23	**92.81**
Ave.		47.35	53.47	77.07	<u>82.88</u>	77.63	68.22	**88.17**

Table 1: The overall performance on the UCIHAR dataset (unit: %). The best performance is highlighted in bold, and the second best performance is underlined.

Source	Target	VAE	β-VAE	DIVA	DDNN	DeepConvLSTM	CoDATS	GILE
S2,S3,S4	S1	77.21	11.48	75.86	66.6	69.59	<u>83.58</u>	**83.86**
S1,S3,S4	S2	73.94	61.02	73.54	71.98	67.17	<u>81.04</u>	**81.65**
S1,S2,S4	S3	15.65	31.72	65.81	71.51	60.59	<u>78.11</u>	**78.66**
S1,S2,S3	S4	75.86	13.65	73.43	64.53	65.51	<u>80.60</u>	**81.41**
Ave.		60.67	29.47	72.16	68.66	65.72	<u>80.83</u>	**81.40**

Table 2: The overall performance on the Opportunity dataset (unit: %). The best performance is highlighted in bold, and the second best performance is underlined.

Overall Results

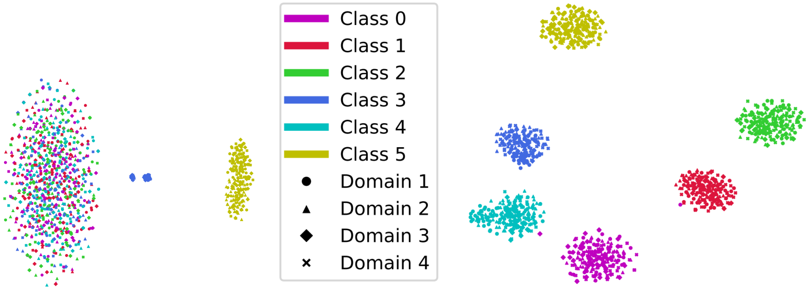
Evaluations on UniMiB SHAR dataset

Source	Target	VAE	β-VAE	DIVA	DDNN	DeepConvLSTM	CoDATS	GILE
2,3,5	1	11.72	15.63	<u>48.17</u>	47.92	44.27	42.71	**55.72**
1,3,5	2	32.76	32.76	39.06	<u>53.69</u>	50.26	46.66	**54.06**
1,2,5	3	22.37	26.97	61.87	<u>73.68</u>	**75.66**	61.51	70.31
1,2,3	5	29.19	30.20	<u>38.43</u>	19.46	34.90	31.88	**42.81**
Ave.		24.01	26.39	46.88	48.69	<u>51.27</u>	45.69	**55.61**

Table 3: The overall performance on the UniMiB SHAR dataset (unit: %). The best performance is highlighted in bold, and the second best performance is underlined.

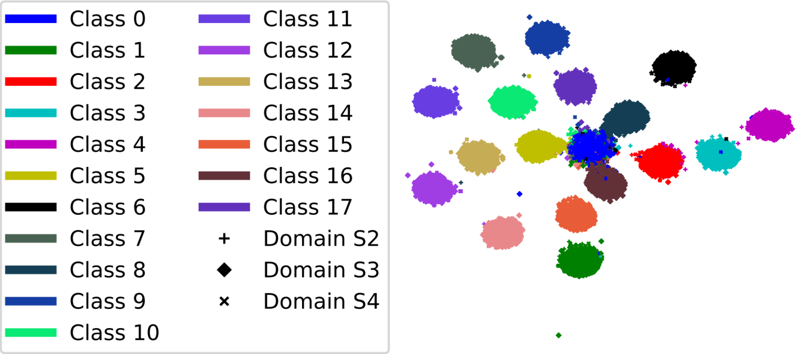
Latent Feature Space Visualization

The t-SNE embeddings of the learned domain-agnostic features by 1) VAE and 2) GILE method on UCIHAR dataset



Latent Feature Space Visualization

The t-SNE embedding of the learned domain-agnostic features by GILE method on Opportunity dataset



Outline

Conclusion

The GILE model for the cross-person activity recognition task

- Feature disentanglement
- Independent excitation mechanism

The model's advantages:

✔ Does not require access to any information from target domain
✔ Readily generalizable to any unseen target domain
✔ Consistently outperforms state-of-the-art methods empirically

Questions?

More info: http://hangwei12358.github.io/
Discussion and cooperation are welcome!

References

[1] Konstantinos Bousmalis et al. "Domain Separation Networks". In: NIPS. 2016, pp. 343–351.

[2] Francisco Javier Ordóñez Morales and Daniel Roggen. "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition". In: Sensors 16.1 (2016), p. 115.

[3] Hangwei Qian et al. "A Novel Distribution-Embedded Neural Network for Sensor-Based Activity Recognition". In: IJCAI 2019, Macao, China, August 10-16, 2019. ijcai.org, 2019, pp. 5614–5620.

[4] Garrett Wilson, Janardhan Rao Doppa, and Diane J. Cook. "Multi-Source Deep Domain Adaptation with Weak Supervision for Time-Series Sensor Data". In: KDD. ACM, 2020, pp. 1768–1778.