A Prototype-Based Generalized Zero-Shot Learning Framework for Hand Gesture Recognition

Jinting Wu

Institute of Automation, Chinese Academy of Sciences

University of Chinese Academy of Sciences

Outline

- Introduction
- Methods
- Results
- Conclusion

Introduction

Motivation

- Most existing works can only recognize a limited number of categories that have been seen during training.
- Generalized Zero-Shot Learning (GZSL) provides a solution for tackling the above challenges. However, GZSL approaches for dynamic hand gesture recognition are less explored.

Introduction

Contributions

- We propose an end-to-end prototype-based GZSL framework for hand gesture recognition which consists of two branches.
- We establish a hand gesture dataset that specifically targets this GZSL task.

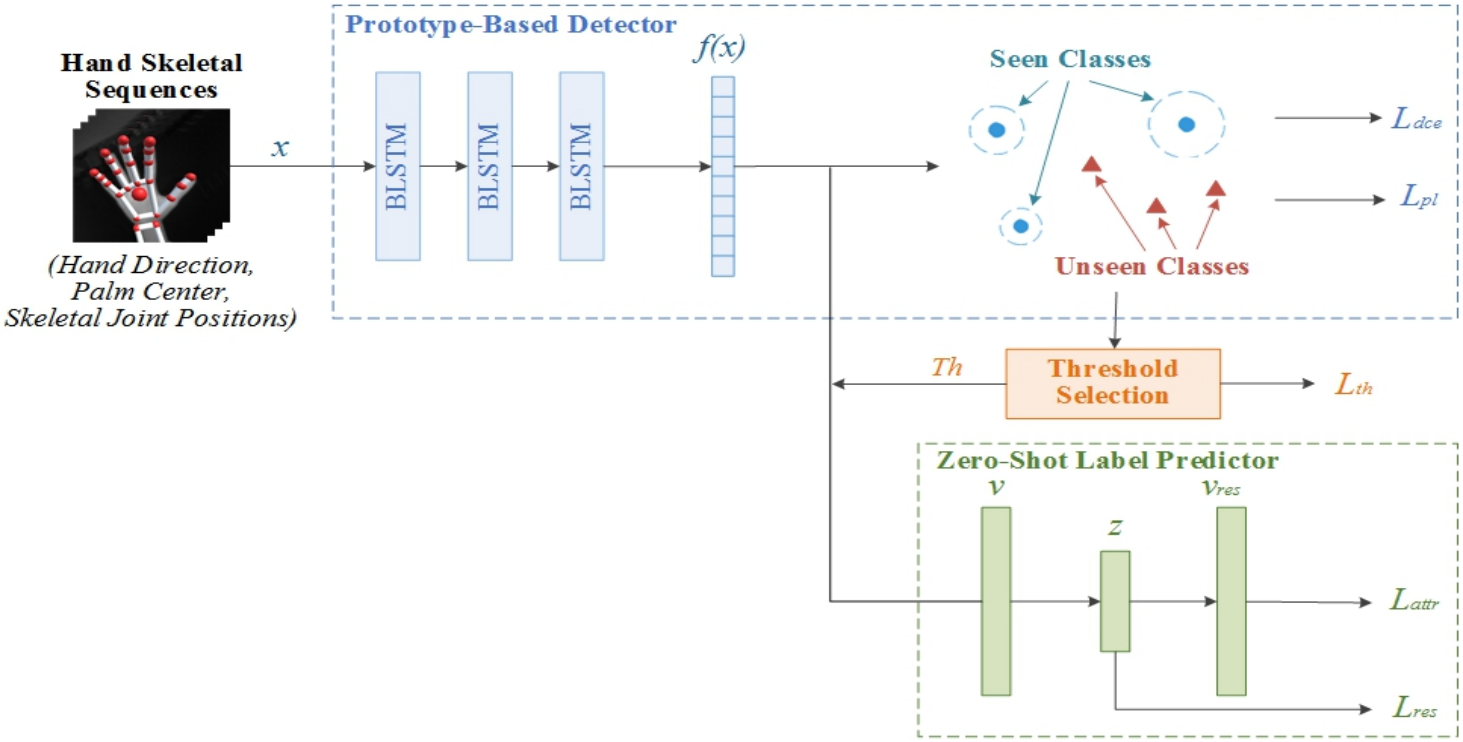
Outline

- Introduction
- Methods
- Results
- Conclusion

Methods

Overview of the Proposed Framework

- Two branches
- Jointly training



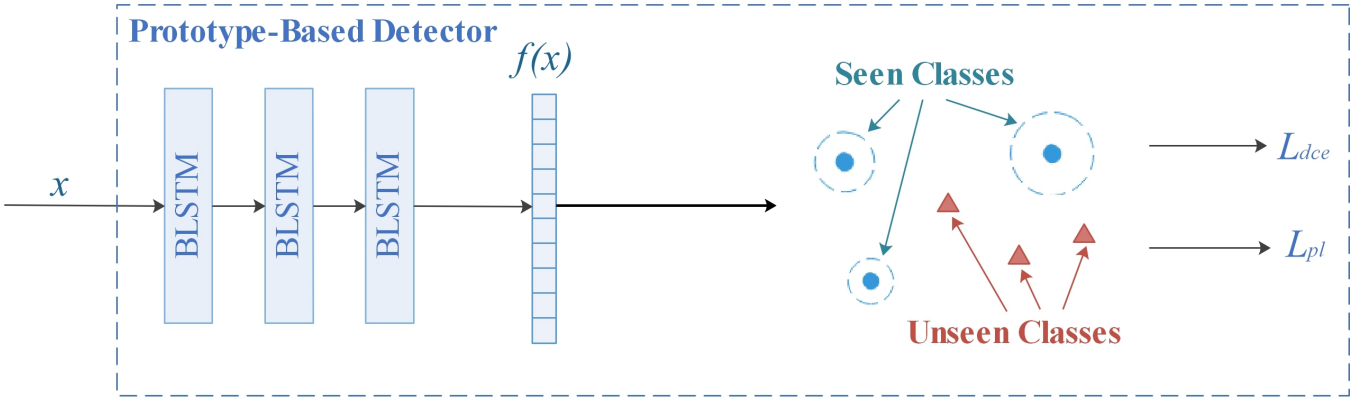
Methods

Prototype-Based Detector (PBD)

- Learning prototypes for each class
- Distance-based cross entropy loss and prototype loss

$$L_{dce}\left((x,y)\middle|\theta,M\right)=-\log\sum_{j=1}^{K}\frac{e^{-\gamma dis\left(p_{pbd}(x),m_{yj}\right)}}{\sum_{k=1}^{C}\sum_{l=1}^{K}e^{-\gamma dis\left(p_{pbd}(x),m_{kl}\right)}}$$

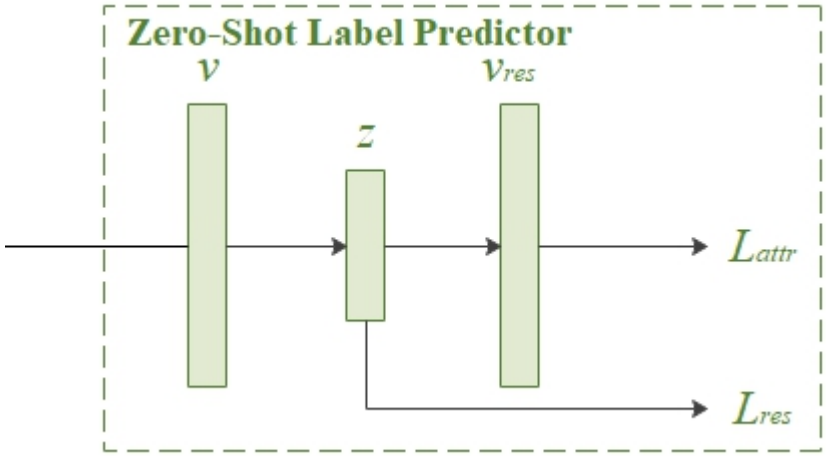
$$L_{pl}\left((x,y)\middle|\theta,M\right)=\left\|p_{pbd}(x)-m_{yj}\right\|_2^2$$



Methods

Zero-Shot Label Predictor

- Using a multi-layer Semantic Auto-Encoder (SAE) to predict the unseen gestures
- Attribute loss and reconstruction loss



$$L_{attr}\left(\left(x,z_s\right)\middle|\theta,\phi\right)=\left\|z-z_s\right\|_2^2$$

$$L_{res}\left(\left(x,z_s\right)\middle|\theta,\phi\right)=\left\|v-v_{res}\right\|_2^2$$

Methods:

End-to-End Learning Objective

- $L\left(\left(x, \mathrm{y}, z_s\right) \middle| \theta, M, \phi\right) = L_{dce} + \lambda_1 L_{pl} + \lambda_2 L_{attr} + \lambda_3 L_{attr}$

Label Prediction

- **Comparing the minimum distance in the prototype space $d_m(x)$ with the thresholds $Th(x)$.**
- **Seen categories: PBD result $\varepsilon(x)$**
- **Unseen categories: SAE result $\varepsilon_u(x)$**

$$label\left(x\right) = \begin{cases} \varepsilon\left(x\right), d_m\left(x\right) \le Th\left(x\right) \\ \varepsilon_u\left(x\right), d_m\left(x\right) > Th\left(x\right) \end{cases}$$

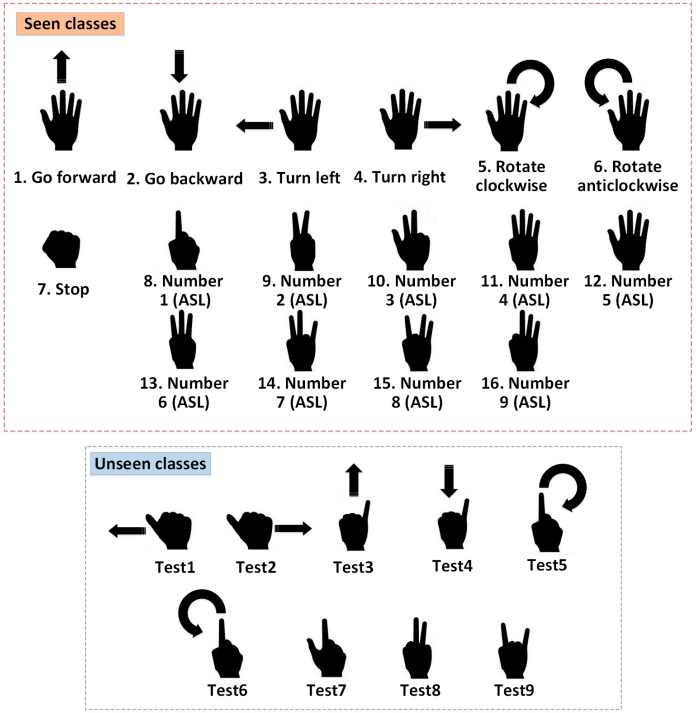
Outline

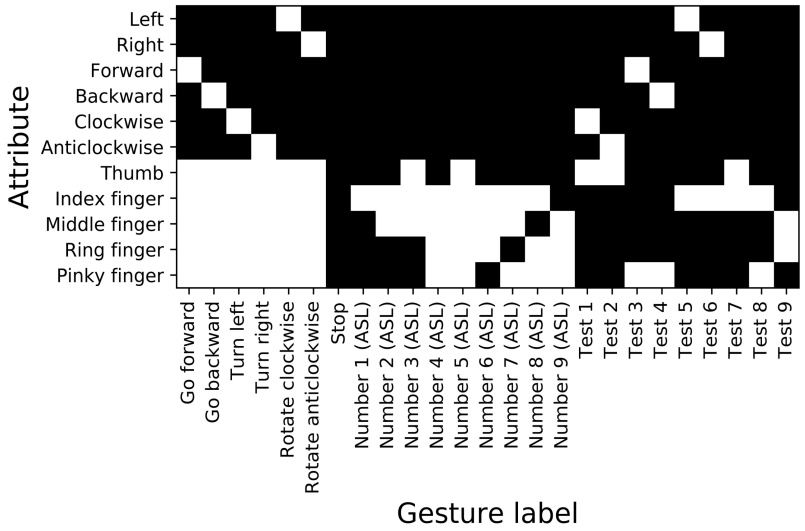
- Introduction
- Methods
- Results
- Conclusion

Results

Dataset

- 16 seen gestures and 9 unseen gestures
- 11 attributes including hand movement and finger bending states





Results

Experimental Results

State-of-the-art Comparisons

- Zero-shot gesture recognition method: ESZSL[1]
- Generalized zero-shot object recognition method: CADA-VAE[2] and f-CLSWGAN[3]

Methods	Acc_s	Acc_u	H
ESZSL [15]	77.81%	13.89%	23.57%
CADA-VAE [11]	80.00%	53.89%	64.40%
f-CLSWGAN [12]	79.79%	55.00%	65.08%
End-to-End Framework (Ours)	**89.06%**	**58.33%**	**70.49%**

1. Madapana, Naveen, and Juan Wachs. Zsgl: zero shot gestural learning.
2. Schonfeld, Edgar, et al. Generalized zero-and few-shot learning via aligned variational autoencoders.
3. Xian, Yongqin, et al. Feature generating networks for zero-shot learning.

Results

Experimental Results

Ablation Analysis

- The traditional SAE[1] without the prototype-based detector
- The framework with a fixed threshold
- The framework where two branches are trained separately

Methods	Acc_s	Acc_u	H	Test Time
BLSTM+SAE [6]	**91.88%**	15.00%	25.79%	0.023s
End-to-End Framework (Fixed Threshold)	84.69%	50.56%	63.31%	0.022s
PBD+SAE	90.63%	57.22%	70.15%	0.026s
End-to-End Framework	89.06%	**58.33%**	**70.49%**	**0.022s**

1. Kodirov, Elyor, Tao Xiang, and Shaogang Gong. Semantic autoencoder for zero-shot learning.

Outline

- Introduction
- Methods
- Results
- Conclusion

Conclusion

- A prototype-based GZSL framework for hand gesture recognition
  - An end-to-end framework with two branches
  - A novel hand gesture dataset
  - Comprehensive experiments demonstrate the effectiveness of our proposed approach

Thanks for your attention!

- Jinting Wu
- Institute of Automation, Chinese Academy of Sciences; University of Chinese Academy of Sciences
- E-mail: wujinting2016@ia.ac.cn