6th International Conference on Nanotechnologies and Biomedical Engineering
Proceedings of ICNBME-2023, September 20–23, 2023, Chisinau, Moldova
Volume 2: Biomedical Engineering and New Technologies for Diagnosis, Treatment, and Rehabilitation

Multimodal Machine Learning for Sign Language Prediction

Yassèr Khalafaoui, Nistor Grozavu, Basarab Matei, Nicoleta Rogovschi

https://doi.org/10.1007/978-3-031-42782-4_26

Abstract

Numerous applications, including translation tools, interpreting services, video remote interpreting, human-computer interaction, online hand tracking of human communication in desktop settings, real-time multi-person recognition systems, games, virtual reality settings, robot controls, and natural language communications, benefit from sign language recognition advantages. Multimodal data contains information from different sources such as video, sensors, electrocardiograms (ECGs), while emotions refer to the non-verbal cues that accompany language use, such as facial expressions and body posture. Integrating these additional sources of information helps to better understand the user's intent, which improves the performance of the sign language recognition model. To build such a model, a set of multimodal data and emotions must be collected. This data set should be differentiated and cover different individual/isolated signs, emotions and body gestures. The model is designed to integrate multimodal data and emotions, which would involve combining different machine and deep learning algorithms adapted to different types of data. In addition, the model will need to be trained to recognize the different emotions that accompany sign language. Once the model is trained, it can be tested on the test dataset to assess its performance and also plan for a test on real data (with signing people). In this paper we propose a study to use the multi-modal machine learning for sign recognition language.

Keywords: sign language, multimodal data, multimodal machine learning (MML)

6th International Conference on Nanotechnologies and Biomedical Engineering
Proceedings of ICNBME-2023, September 20–23, 2023, Chisinau, Moldova
Volume 2: Biomedical Engineering and New Technologies for Diagnosis, Treatment, and Rehabilitation

References

1. Bengio,Y., Courville,A.,Vincent, P.: Representation learning: a reviewand newperspectives. IEEE Trans. Pattern Anal. Mach. Intell. **35**(8), 1798–1828. IEEE Computer Society (2013)
2. Guo,W.,Wang, J.,Wang, S.: Deep multimodal representation learning: a survey. IEEE Access **7**, 63373–63394 (2019). https://doi.org/10.1109/ACCESS.2019.2916887
3. Yan, A.,Wang,W., Ren, Y., Geng, H.: A clustering algorithm for multi-modal heterogeneous big data with abnormal data. Front. Neurorobot. **15**, 64 (2021)
4. Pedrycz, W., Hirota, K.: A consensus-driven fuzzy clustering. Pattern Recogn. Lett. **29**(9), 1333–1343 (2008)
5. Zhao, B., Kwok, J.T., Zhang, C.: Multiple Kernel clustering. In: Proceedings of the 2009 SIAM International Conference on Data Mining, pp. 638–649. Society for Industrial and Applied Mathematics (2009)
6. Bickel, S., Scheffer, T.: Multi-view clustering. Proc. ICDM **4**(2004), 19–26 (2004)
7. Simon, T., Joo, H., Matthews, I., Sheikh, Y.: Hand keypoint detection in single images using multiview bootstrapping. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1145–1153 (2017)
8. Yan, S., Xia, Y., Smith, J.S., Lu,W., Zhang, B.:Multiscale convolutional neural networks for hand detection. Appl. Comput. Intell. Soft Comput. **2017**, 1–13 (2017)
9. Rao, G., Syamala, K., Kishore, P.V.V., Sastry, A.S.C.S.: Deep convolutional neural networks for sign language recognition. In: Conference on Signal Processing and Communication Engineering Systems (SPACES), India (2018)
10. Koller, O., Ney, H., Bowden, R.: Deep learning of mouth shapes for sign language. In: Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 85–91 (2015)
11. Newell, A., Yang, K., Deng, J.: Stacked hourglass networks for human pose estimation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016, Proceedings, Part VIII 14, vol. 9912, pp. 483–499. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46484-8_29
12. Pu, J., Zhou, W., Li, H.: Dilated convolutional network with iterative optimization for continuous sign language recognition. IJCAI **3**, 7 (2018)

13. Cui, R., Liu,H., Zhang, C.:Adeep neural framework for continuous sign language recognition by iterative training. IEEE Trans. Multimedia **21**(7), 1880–1891 (2019)

14. Chen, H., Tong, R., Chen, M., Fang, Y., Liu, H.: A hybrid CNN-SVM classifier for hand gesture recognition with surface EMG signals. In: 2018 International Conference on Machine Learning and Cybernetics (ICMLC), vol. 2, pp. 619–624. IEEE, July 2018

15. Tzeng, E., Hoffman, J., Saenko, K., Darrell, T.: Adversarial discriminative domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7167–7176 (2017)

16. Nguyen, X.S., Brun, L., Lézoray, O., Bougleux, S.: A neural network based on SPD manifold learning for skeleton-based hand gesture recognition. In: CVPR, pp. 12036–12045 (2019)

17. Nguyen, X.S., Brun, L., Lézoray, O., Bougleux, S.: Learning recurrent high-order statistics for skeleton-based hand gesture recognition. In: ICPR, pp. 975–982 (2020)

18. Nguyen, X.S.: GeomNet: a neural network based on Riemannian geometries of SPD matrix space and Cholesky space for 3D skeleton-based interaction recognition. In: ICCV, pp. 13379–13389 (2021)

19. Nguyen, X.S.: A Gyrovector space approach for symmetric positive semi-definite matrix learning. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) ECCV 2022. LNCS, vol. 13687, pp. 52–68. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19812-0_4

20. Nguyen, X.S.: The Gyro-structure of some matrix manifolds. In: NeurIPS (2022)