# Enhancing Syllabic Component Classification in Japanese Sign Language by Pre-training on Non-Japanese Sign Language Data

**Jundai Inoue, Makoto Miwa, Yutaka Sasaki, Daisuke Hara**
Toyota Technological Institute
2-12-1 Hisakata, Tempaku-ku, Nagoya, Aichi, 468-8511, Japan
{sd24410, makoto-miwa, yutaka.sasaki, daisuke}@toyota-ti.ac.jp

### Abstract

In sign languages, syllables are composed of syllabic components consisting of locations, movements, and handshapes; however, the rules of combinations of these syllabic components are still unclear. Decomposing existing syllables into syllabic components is necessary to clarify the rules. This study aims to construct an automatic syllabic component classification system for Japanese Sign Language (JSL) using deep learning. We propose a pre-training method using non-Japanese Sign Language data to achieve high performance in classifying syllabic components in a situation where the number of training JSL videos is limited. We also investigate multitask learning for syllabic component classification to share the information among the syllabic components. Experiments on the syllabic component classification for the dominant hand show that 1) pre-training with the American Sign Language (ASL) dataset improved classification performance for the movement and handshape components and 2) multitask learning did not contribute to the performance improvement of syllabic component classification. We also investigated the influence of pre-training on syllabic component classification by visualizing critical elements in videos to predict the components.

**Keywords:** Japanese Sign Language, Syllabic components, Pre-training, Multitask learning

## 1. Introduction

*Locations*, *movements*, and *handshapes* are the syllabic components in sign languages. Syllables of sign language are combinations of the syllabic components, and the composition rules for the syllables are still unclear (Hara, 2016). To analyze the rules of syllable composition in Japanese Sign Language (JSL), Hara (2019) proposed a syllable database with videos of syllables and their components that are decomposed by hand. However, manually decomposing a number of syllables that have not yet been registered in the database into syllabic components is costly. Therefore, it is needed to construct a system that can automatically recognize syllabic components from JSL videos. The syllabic component recognizer could be used not only to supplement the database but also to further analyze JSL using the system's prediction results.

Recently, deep learning approaches to sign language processing have been shown to be effective (Jiang et al., 2021; Chen et al., 2022; Zuo et al., 2023). Deep learning methods require a large amount of labeled training data to achieve high performance, but unfortunately, the number of JSL videos with labeled syllabic components is limited. On the other hand, there is a large amount of data of a non-Japanese Sign Language, such as American Sign Language (ASL), and the two sign languages share features in expressing signs with manual and non-manual signals. Although we can expect the improvement of classification performance for JSL by using the shared features, such an approach has yet to be investigated.

This study aims to construct an automatic syllabic component classification system from JSL videos. As the first step toward this goal, this study focuses on the location, movements, and handshape of the dominant hand. To address the problem of limited data in JSL, we propose pre-training using non-JSL datasets. We conduct training on JSL video data to classify syllabic components after initializing the parameters with those trained on a non-JSL dataset. We also introduce multitask learning in classifying location, movement, and handshape components by sharing the base classification model among the components.

The contributions of this study are summarized as follows:

- We constructed a system that automatically recognizes syllabic components of the dominant hand from JSL videos.
- We showed the effectiveness of using models pre-trained on a non-JSL dataset for the movement and handshape classification from JSL with limited data.
- We found that information sharing between tasks does not necessarily improve classification performance through multitask learning of syllabic components in JSL.

*Proceedings of the 11th Workshop on the Representation and Processing of Sign Languages*, pages 123–130

# 2. Related Work

## 2.1. Japanese Sign Language Dataset

Nagashima et al. (2018) constructed a versatile JSL database that can be used in the fields of linguistics and engineering. The database includes high-resolution video data capturing the actions of two native signers with a high-resolution camera from the front and diagonally forward from the left and right. Additionally, it incorporates 3D motion data obtained through optical motion capture and depth data from distance sensors. The dataset provides data on 4,873 glosses and ten dialogues.

Hara (2019) defined a JSL coding manual and created a syllable database in which the syllables were broken down into location, movement, and handshape components. The database contains video clips representing the JSL syllables, recorded with a single signer. 1,086 syllable videos were included, each consisting of approximately 300 frames. The location components are classified into 22 categories to indicate the hand locations in space or on the body. The handshape components are divided into 69 categories. The location and handshape components are assigned to a single category label in the video. The location component signifies the starting position of the sign, and the handshape component indicates the shape of the hand. We should note that this database manually defines base handshapes so that each syllable can be represented by a single base handshape. We use this base handshape as the handshape component, and the changes in the handshape are represented by the movement component.

The movement components are distinguished into 55 ways of moving a hand, such as rightward movement and finger joint opening, with one to three categories assigned to each video. In addition to the components for dominant and non-dominant hands, more detailed decompositions of each syllabic component are attached, such as “contact,” “hand orientation,” and “metacarpal orientation.”

## 2.2. Sign language processing using machine learning and deep learning

Sign language processing using machine learning and deep learning, such as Sign Language Recognition (SLR) for predicting gloss (Jiang et al., 2021; Zuo et al., 2023) and sign language translation for translating signs into spoken language (Chen et al., 2022), has been actively conducted. Skeleton Aware Multi-modal SLR (SAM-SLR) (Jiang et al., 2021) is a framework that integrates body, motion, and depth information in addition to video and keypoint information. Video-Keypoint Network (VKNet) (Zuo et al., 2023) extracts features from 64 and 32 video frames and keypoints to account for different temporal information. VKNet consists of two sub-networks, VKNet-64 and VKNet-32. Each sub-network also contains video and keypoint encoders, and there are bidirectional lateral connections (Duan et al., 2022) to exchange information between each encoder. S3D (Xie et al., 2018), a 3D Convolutional Neural Network that can consider spatio-temporal information, is used as the encoder. After keypoints are estimated from the video using a learned pose estimation model, HRNet (Sun et al., 2019), 64 and 32 video frames and keypoints are input to VKNet-64 and VKNet-32, respectively. The combined representation vectors from each network are used to predict the gloss. VKNet performed well on several datasets for SLR.

Studies on sign languages considering syllabic components have also been conducted (Zhang and Duh, 2023; Tavella et al., 2022; Kezar et al., 2023; Hatano et al., 2016). To clarify the importance of the handshape component in SLR, Zhang and Duh (2023) constructed a dataset labeled with handshapes on an existing SLR dataset and proposed a model that predicts both glosses and handshapes simultaneously by extending the existing SLR model. The proposed model performs better than those that only use videos as input without considering handshapes. Tavella et al. (2022) and Kezar et al. (2023) have constructed datasets labeling multiple syllabic components in addition to gloss in sign language videos. Furthermore, Kezar et al. (2023) classified 16 different phonological features, which are close to fine-grained syllabic components, and demonstrated that learning the features through classification contributes to improving the performance of SLR. In JSL, Hatano et al. (2016) employed machine learning methods to recognize the location, movement, and handshape components and construct a SLR system based on the weighted sum of classification scores for each component. This method requires extracting the video’s features, such as coordinates, velocity, and acceleration.

# 3. Methods

This study proposes a method for classifying syllabic components in JSL videos using pre-training on a non-JSL dataset. This study focuses on the location, movement, and handshape components of the dominant hand, which are defined in the syllable database created by Hara (2019) and employs VKNet (Zuo et al., 2023) as the base deep learning model. We initialized the parameters of VKNet with those pre-trained on a non-JSL dataset to leverage information from non-JSL. The overall architecture of the proposed model is illustrated in Figure 1.

As explained in §2.1, there are 22, 55, and 69 categories for location, movement, and handshape

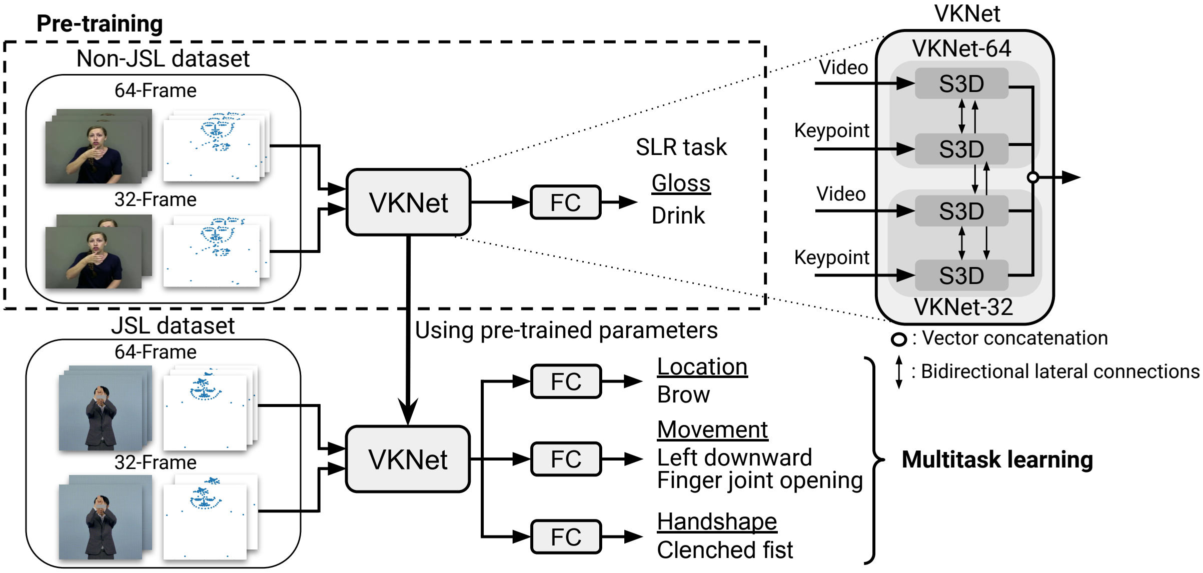


Figure 1: The overview of syllabic component classification through pre-training using non-JSL dataset

components, respectively. We added three fully connected (FC) layers corresponding to individual components to the VKNet pre-trained on the non-JSL to classify each syllabic component.

A softmax function is applied to the output vector of the FC layers for the location and handshape components, where a single label is assigned from multiple categories. This function enables multi-class classification, where the class with the highest predicted probability is considered the prediction. By contrast, a sigmoid function is applied to the output vector of the FC layers for the movement component, which involves multiple labeled movements. This function allows for binary classification for each movement type; movements with predicted probabilities higher than a threshold are considered the prediction in the multi-label classification.

The loss function includes cross-entropy and asymmetric losses (Ridnik et al., 2021). The cross-entropy loss is used for location and handshape classification, while the Asymmetric Loss (AsLoss) is applied to the movement classification. Since there are only up to three movements for each syllable in the database, the classification problem is highly imbalanced, with few positive and many negative examples. The AsLoss addresses this imbalance by calculating a weighted sum in which the weight of the loss in positive examples is larger than that in negative examples. It is defined as:

$$\text{AsLoss} = \begin{cases} -(1-p)^{\gamma^+}\log(p) & \text{if} \quad y=1 \\ -p_m^{\gamma^-}\log(1-p_m) & \text{otherwise} \end{cases} \tag{1}$$

where $p_m$ is defined in Equation (2) to ignore negative examples that can be classified easily.

$$p_m \quad = \quad \max(p-m, 0) \tag{2}$$

Note that $p$ is the network's output probability and hyperparameters $\gamma^-$ and $\gamma^+$ are sets such that $\gamma^- > \gamma^+$ to emphasize the contribution of positive examples. $m$ represents the threshold value.

During training, multitask learning is performed to share the information among syllabic components. Specifically, VKNet is shared, and the loss function is the sum of classification losses for each syllabic component.

## 4. Experimental settings

We evaluated the proposed method using the syllable database created by Hara (2019). We randomly split the 1,072 instances annotated with the location, movement, and handshape components into 750, 161, and 161 instances for training, development, and testing, respectively. The statistics for the top-10 instances of each component are presented in Table 1. The table shows that syllable instances are highly imbalanced among the categories. To avoid highly challenging classification problems, we excluded instances with the categories with fewer than five instances in the training data, treating them as false-negative predictions. We adopted the micro F-score as the evaluation metric.

As the pre-training parameters, we utilized the pre-trained VKNet parameters,[1] which was trained on the 14,289 training instances with 2,000 glosses of Word-Level American Sign Language (WLASL) dataset for SLR in ASL (Li et al., 2020).

We conducted two comparisons in the experiments. The first comparison is to investigate the ef-

[1] https://github.com/FangyunWei/SLRT/tree/main/NLA-SLR

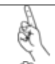
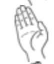
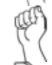
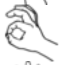
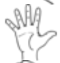
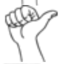
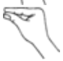
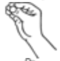
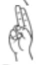
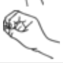
| Movement | # | Handshape | # | Location | # |
|---|---|---|---|---|---|
| Rightward movement of a hand | 142 |  | 138 | * | 835 |
| Forward movement of a hand | 135 |  | 125 | Temples | 40 |
| Wrist rotation: outward rotation of a wrist with the little finger as the axis | 120 |  | 57 | Mouth | 32 |
| Downward movement of a hand | 117 |  | 55 | Chest | 23 |
| Flexion of finger joints with handshape changes | 80 |  | 53 | Brow | 22 |
| Extension of finger joints with handshape changes | 77 |  | 48 | Eyes | 17 |
| Circular or semicircular movement on a horizontal plane | 69 |  | 42 | Face | 16 |
| Upward movement of a hand | 64 |  | 40 | Elbow | 13 |
| Leftward movement of a hand | 61 |  | 40 | ** | 13 |
| Non-linear movement (trajectory) of a hand | 51 |  | 34 | Abdomen | 12 |

Table 1: Numbers (#) of top-10 instances for the location, movement, handshape components, icons from McKee et al. (2011). * and ** in the location component represent the neutral space in which the sign is made in front of the body or face, respectively.

| Method | Syllabic component | | |
|---|---|---|---|
| | Location | Movement | Handshape |
| VKNet | 80.75 (± 1.02) | 38.29 (± 2.54) | 39.54 (± 1.05) |
| + Pre-training | 81.16 (± 2.05) | 52.41* (± 0.86) | 44.72* (± 3.55) |
| + Multitask learning | 81.99 (± 0.00) | 45.76* (± 0.82) | 42.23† (± 1.34) |

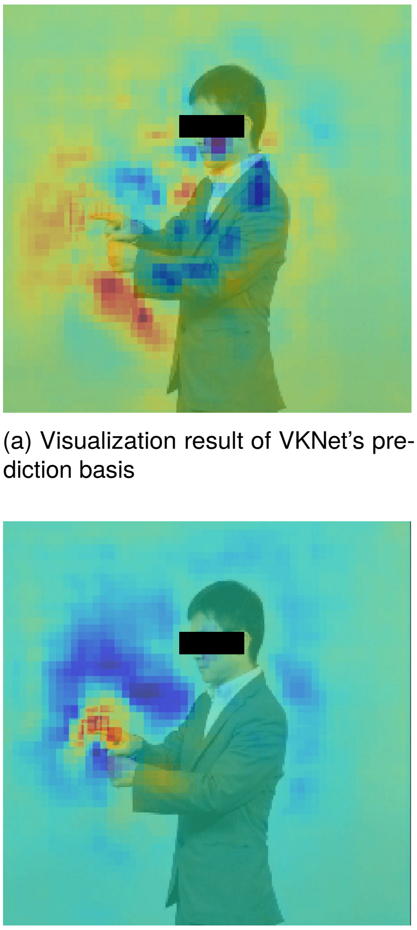
Table 2: Results of syllabic component classification. The means of three runs are shown as the final micro F-scores (%). The numbers in parentheses are standard deviations. * and † denote significance levels of 0.05 and 0.1 compared with the results directly above.

fectiveness of pre-training using the ASL dataset in syllabic component classification for JSL; we compared the classification performance of VKNet with parameters initialized from the pre-trained model and VKNet with randomly initialized parameters. The second comparison is to evaluate multitask learning. We compared the classification performance when simultaneously or independently addressing each task to understand the impact of information sharing between tasks. We used the Adam optimization method (Kingma and Ba, 2015), setting the learning rate to $5{\times}10^{-5}$ and applied cosine annealing as a scheduler to change the learning rate per epoch. We set the hyperparameters $\gamma^{-}$, $\gamma^{+}$, and $m$ of the AsLoss to 4, 1, and 0.05, respectively. To suppress overfitting, we employed dropout (Srivastava et al., 2014) and regularization, setting their values to 0.2 and $10^{-3}$, respectively.

## 5. Results

The results of syllabic component classification from JSL videos in test data are shown in Table 2. The results of syllabic component classification using VKNet with parameters pre-trained on the WLASL dataset as initial values showed that the micro F-scores for the location, movement, and handshape components were improved compared to those using VKNet with random parameters as initial values. The results evaluated on the development and test data are summarized in appendix B. We conducted a significance difference test with the bootstrap method to verify the improvement in classification performance of the pre-trained VKNet. As a result, we confirm that the pre-training method effectively improved the classification of the movement and handshape components of JSL.

Multitask learning improved the micro F-score of the location component but decreased those of the movement and handshape components. The significance test showed a significant decrease in the classification of the movement component, while there was no significant difference for the location and handshape components. This result indicates that multitask learning is ineffective or harmful in classifying syllabic components of JSL.



(a) Visualization result of VKNet's prediction basis

(b) Visualization result of pre-trained VKNet's prediction basis

Figure 2: Visualization results (classification of the movement component)

## 6. Discussion

To verify the influence of pre-training on the syllabic components of VKNet, we visualized the parts of the video VKNet focused on while predicting syllabic components using Adaptive Occlusion Sensitivity Analysis (AOSA) (Uchiyama et al., 2023), one of the methods of explainable AI techniques. The AOSA results were visualized with colors from red to blue to indicate their importance; the areas with high importance are shown in red. The example of the movement component that could not be classified by VKNet but could be classified by the pre-trained VKNet is visualized in Figure 2. From these results, we can see that the right hand, which is the dominant hand, is more focused after pre-training. This change in the focus suggests that the pre-trained VKNet can make more accurate predictions than the VKNet by focusing on the dominant hand and classifying syllabic components.

## 7. Conclusions

This study proposed the classification of the syllabic component for the dominant hand using parameters of a model pre-trained on a non-JSL dataset as a first step to construct a method for syllabic component classification based on JSL videos. We also introduced multitask learning for sharing information among syllabic component classification. We evaluated the proposed method based on the VKNet model using the JSL database in the experiments. Experimental results show that pre-training with the ASL dataset significantly improves the classification performance of the movement and handshape components from a limited number of the JSL videos. On the other hand, the classification performance with multitask learning did not improve the performance of syllabic component classification in JSL. We also investigated the effect of pre-training on syllabic component prediction by visualizing the predictive basis of VKNet using AOSA. The visualization results suggest that the proposed pre-training enabled the focus on the target hand. Future work includes investigating the models and training methods to improve the classification and classification performance of syllabic components for both the dominant and non-dominant hands.

## 8. Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 23H00626.

## 9. References

Yutong Chen, Fangyun Wei, Xiao Sun, Zhirong Wu, and Stephen Lin. 2022. A simple multi-modality transfer learning baseline for sign language translation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5120–5130.

Haodong Duan, Yue Zhao, Kai Chen, Dahua Lin, and Bo Dai. 2022. Revisiting skeleton-based action recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2969–2978.

Daisuke Hara. 2016. *18. An information-based approach to the syllable formation of Japanese Sign Language*, pages 457–482. De Gruyter Mouton, Berlin, Boston.

Daisuke Hara. 2019. New Japanese Sign Language Coding Manual. (In Japanese).

Mika Hatano, Shinji Sako, and Tadashi Kitamura. 2016. Real-time sign language recognition by kinect v2 based on three elements of sign language. *IEICE technical report*, 115(491):59–64.

Songyao Jiang, Bin Sun, Lichen Wang, Yue Bai, Kunpeng Li, and Yun Fu. 2021. Skeleton aware multi-modal sign language recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Lee Kezar, Jesse Thomason, Naomi Caselli, Zed Sehyr, and Elana Pontecorvo. 2023. The sem-lex benchmark: Modeling asl signs and their phonemes. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '23, New York, NY, USA. Association for Computing Machinery.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*.

Dongxu Li, Cristian Rodriguez Opazo, Xin Yu, and Hongdong Li. 2020. Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*.

D. McKee, R. McKee, S. Pivac Alexander, L. Pivac, and M Vale. 2011. Online dictionary of new zealand sign language. `https://www.nzsl.nz/`.

Yuji Nagashima, Daisuke Hara, Yasuo Horiuchi, Shinji Sako, Rituko Kikusawa, Akira Ichikawa, Keiko Watanabe, and Naoto Kato. 2018. Development of the super high-definition and high-precision japanese sign language database available for various research fields. In *Proceedings of Language Resources Workshop*, volume 3, pages 148–155. National Institute for Japanese Language and Linguistics. (In Japanese).

Tal Ridnik, Emanuel Ben-Baruch, Nadav Zamir, Asaf Noy, Itamar Friedman, Matan Protter, and Lihi Zelnik-Manor. 2021. Asymmetric loss for multi-label classification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 82–91.

Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958.

Ke Sun, Bin Xiao, Dong Liu, and Jingdong Wang. 2019. Deep high-resolution representation learning for human pose estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Federico Tavella, Viktor Schlegel, Marta Romeo, Aphrodite Galata, and Angelo Cangelosi. 2022. WLASL-LEX: a dataset for recognising phonological properties in American Sign Language. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 453–463, Dublin, Ireland. Association for Computational Linguistics.

Tomoki Uchiyama, Naoya Sogi, Koichiro Niinuma, and Kazuhiro Fukui. 2023. Visually explaining 3d-cnn predictions for video classification with an adaptive occlusion sensitivity analysis. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 1513–1522.

Saining Xie, Chen Sun, Jonathan Huang, Zhuowen Tu, and Kevin Murphy. 2018. Rethinking spatiotemporal feature learning: Speed-accuracy trade-offs in video classification. In *Proceedings of the European conference on computer vision (ECCV)*, pages 305–321.

Xuan Zhang and Kevin Duh. 2023. Handshape-aware sign language recognition: Extended datasets and exploration of handshape-inclusive methods. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 2993–3002, Singapore. Association for Computational Linguistics.

Ronglai Zuo, Fangyun Wei, and Brian Mak. 2023. Natural language-assisted sign language recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14890–14900.

## A. Impact of data imbalance on location component classification

For the location component classification, neutral space instances, the first row in Table 1, cover most of the dataset. To examine its impact on the classification results, we used a pre-trained VKNet and evaluated it by excluding the instances. The evaluation results on the development data are shown in Table 3. When excluding neutral space instances from the dataset, the performance significantly dropped. This result suggests that the model was affected by the bias in the dataset and fitted to the neutral space class. This performance degradation indicates that, to improve the generality of the model, the bias in the dataset needs to be addressed by sampling data or changing the loss function.

## B. Overall result

In this study, we set four learning conditions to compare the effects of pertaining VKNet and multitask learning: (1) no pertaining VKNet, no multitask learning, (2) pertaining VKNet, no multitask learning, (3) no pertaining VKNet, multitask learning, (4) pertaining VKNet and multitask learning. We performed syllabic component classification for each condition using the development and test data. The results are shown in Table Table 4

## C. Hyperparameter tuning in multitask learning

we conducted additional experiments to optimize the coefficients of the loss functions for each task in multitask learning. Previously, we summed the losses for each syllabic component. Still, this time, we introduced weighting coefficients for the loss of each syllabic component and attempted to optimize these coefficient values using a Bayesian optimization. Specifically, the value of each coefficient was constrained to be between 0 and 1, and the sum of all coefficients was always set to 1. We performed 70 iterations of Bayesian optimization and searched for the combination of coefficients that maximized the micro F-score for syllabic component classification on the development data. It is shown in Table 5, where the optimal coefficient values obtained by Bayesian optimization and the corresponding micro F-scores are shown in contrast to the micro F-scores obtained by simply summing the losses. After three evaluations, the micro F-score for the handshape component showed a slight improvement, although the micro F-scores for the location and movement components showed a slight decrease. However, these score changes were within the margin of error, indicating no significant difference resulted from simply summing the losses for each syllabic component. Therefore, we evaluated the test data using a simple sum of losses with equal weights.

| | Location |
|---|---|
| pre-trained VKNet w/ neutral space | 80.75 (± 1.02) |
| pre-trained VKNet w/o neutral space | 41.67 (± 4.54) |

Table 3: Results of location component classification with and without neutral space instances. Neutral space instances constitute a large portion of the dataset. The performance is measured using the micro F-score (%), with the reported values showing the average and standard deviation over three evaluation runs.

|  |  | Dev |  |  | Test |  |  |
|---|---|---|---|---|---|---|---|
| Method |  | Location | Movement | Handshape | Location | Movement | Handshape |
| Multitask | VKNet | 82.40 (± 1.27) | 34.06 (± 0.52) | 39.75 (± 1.83) | 80.33 (± 1.06) | 38.55 (± 1.25) | 35.20 (± 2.55) |
|  | + Pre-training | 82.20 (± 1.17) | 39.94 (± 2.85) | 47.41 (± 2.29) | 81.99 (± 0.00) | 45.76 (± 0.82) | 42.23 (± 1.34) |
| Singletask | VKNet | 83.85 (± 1.01) | 34.57 (± 0.39) | 43.89 (± 0.29) | 80.75 (± 1.02) | 38.29 (± 2.54) | 39.54 (± 1.05) |
|  | + Pre-training | 83.44 (± 0.77) | 44.98 (± 1.06) | 47.82 (± 1.02) | 81.16 (± 2.05) | 52.41 (± 0.86) | 44.72 (± 3.55) |

Table 4: Results of syllabic component classification with and without pertaining and with and without multitask learning. The evaluation metric is the micro F-score (%). The mean and standard deviation of the three evaluations are shown.

|  | Dev |  |  |
|---|---|---|---|
| hyperparameter | Location | Movement | Handshape |
| alpha = 0.095704<br>beta = 0.597839<br>gamma = 0.306457 | 78.46 (± 1.17) | 38.70 (± 1.86) | 48.24 (± 1.63) |
| alpha = beta = gamma | 82.20 (± 1.17) | 39.94 (± 2.85) | 47.41 (± 2.29) |

Table 5: Micro F-score (%) of syllabic component classification using the optimized hyperparameters obtained from Bayesian optimization and an equal weight baseline. Coefficients for location, movement, and handshape are denoted as alpha, beta, and gamma, respectively.