Visualization of 3D latent space of My CVAE (Conditional VAE)

Ayumu Manabe (215x120x)

※Japanese version : https://aymp.github.io/InfoVis2021/FinalTask/report_ja.pdf

1. Introduction

We are currently studying the application of a deep generative model called CVAE (Conditional VAE) in the field of biometrics, and although there are several possible model structures in CVAE depending on the dependency between input, label, and latent variables, the model used in my research has the structure shown in Fig. 1.

Each arrow in Fig. 1 corresponds to a neural network, such as q_ξ (y|x), which is capable of inferring the label y (the label here is to identify which person) from the input data x. Initially, we thought of using this trained classifier to predict which person the input data belonged to for biometric authentication. In fact, the classifier was able to correctly classify the subject from the input data with 96% accuracy.

However, authentication and classification are not the same thing. In this research, given the data of N people beforehand, it is necessary to perform 1:N authentication to determine which of the N people the image is or is not when data is input that does not identify the individual. In other words, it is necessary to distinguish between a previously learned person and an unlearned unknown person.

In this research paper, we consider a method to determine whether the input data belongs to a learned person or not, based on the distribution of the latent variable z obtained by q_φ (z|x,y) in Fig. 1. Specifically, we consider whether it is possible to classify data plotted close to the distribution of the latent space at the time of learning as the person himself, and data plotted far away as others.

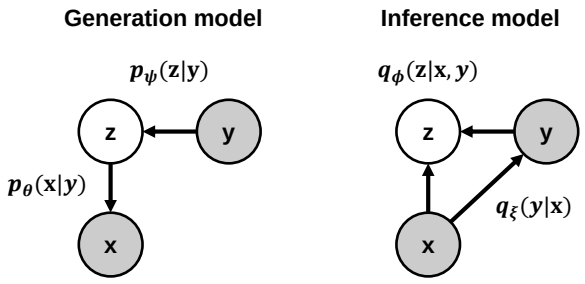


Fig 1 Graphical models of the proposal CVAE

2. Method

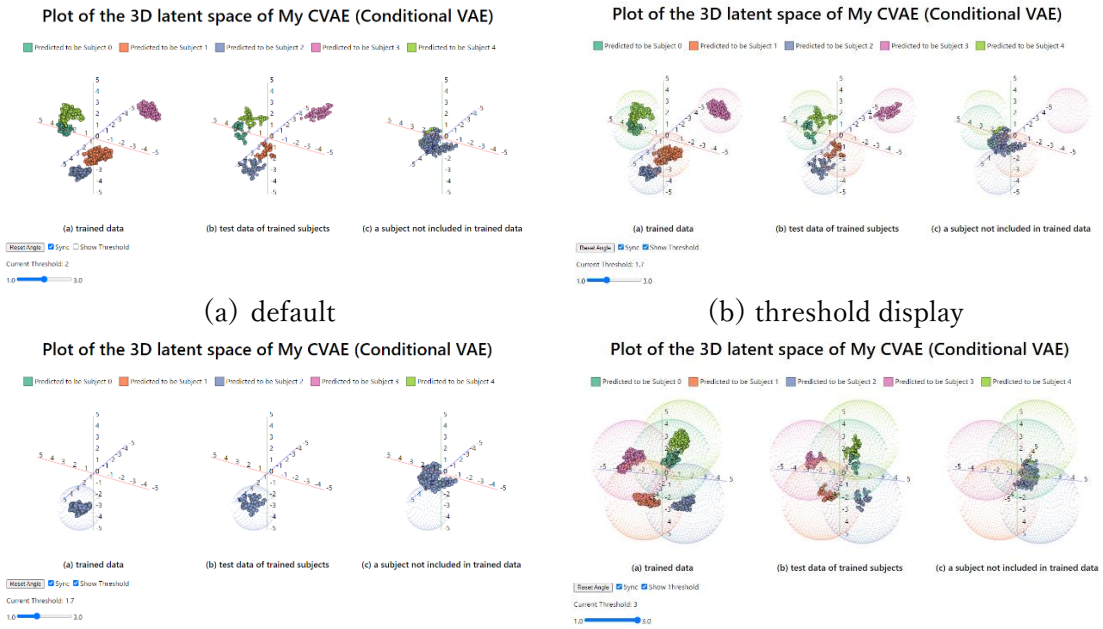
We plot three latent spaces: the first is the latent space of the model trained with the training data of five users; the second is the latent space of the same five test data inputs, which

corresponds to the case where a legitimate user tries to authenticate in the actual authentication; and the third is the latent space of the sixth user input, which is not in the training data, and corresponds to the case where an attacker tries to authenticate. The third is the latent space for the case where the attacker tries to authenticate by entering the data of the sixth person who is not in the training data.

In addition, the plot has the following features:

- Rotation by dragging (3 screens synchronized/independent)
- Highlighting by mouse over (erase subjects other than the one you want to focus on)
- Display of a sphere centered on the center of gravity coordinate of each person in the latent space when training data is input.(The sphere represents the distance from the reference coordinate, so you can check if you can separate the person from the unknown person.

Examples of each function are shown in Fig. 2.



(a) default (b) threshold display

(c) hovered to refine (d) change angles and thresholds

Fig 2 Examples of interactive operations

3. Result

All unknown persons were classified by the classifier as subjects 2 and 4. The plots of the data classified as subjects 2 and 4 when the threshold value (corresponding to the radius of the sphere in the figure) was set to 1.3 and 2.0 are shown in Fig. 3 - Fig. 6.

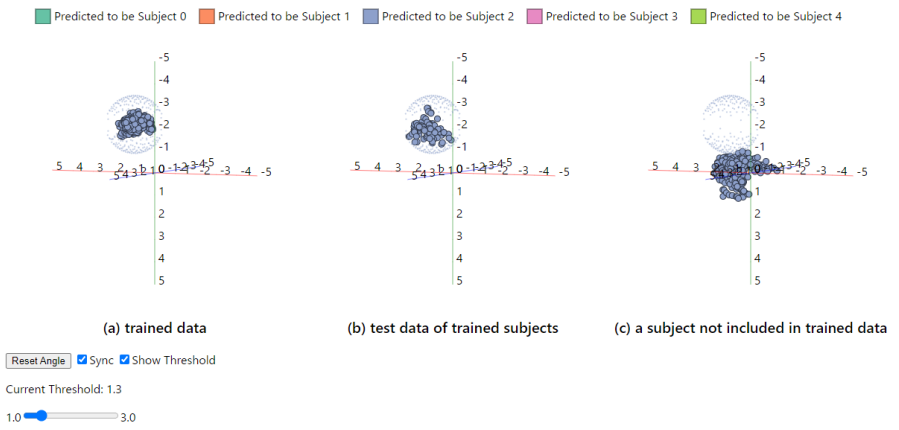


Fig 3 Data predicted to be subject 2 (threshold=1.3)

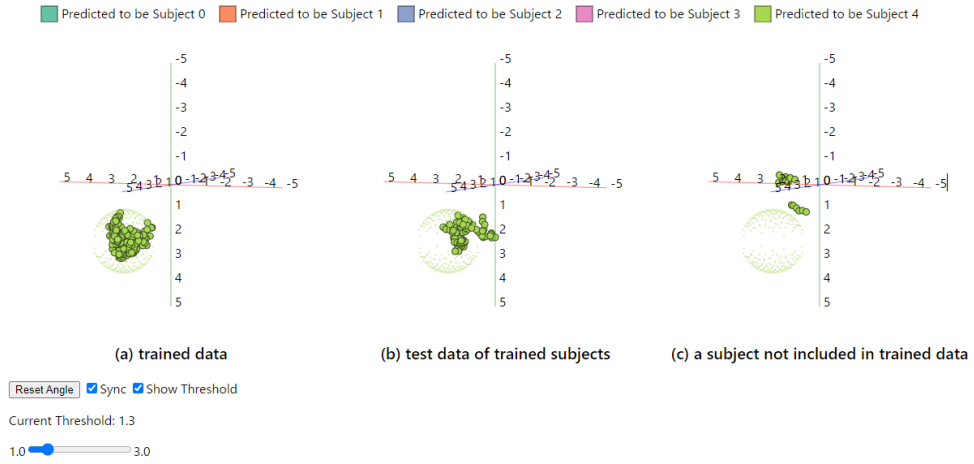


Fig 4 Data predicted to be subject 4 (threshold=1.3)

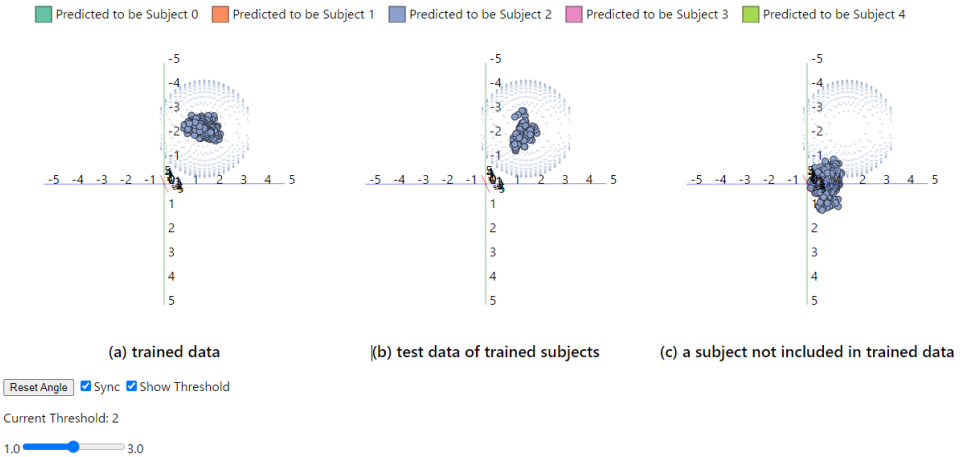


Fig 5 Data predicted to be subject 2 (threshold=2.0)

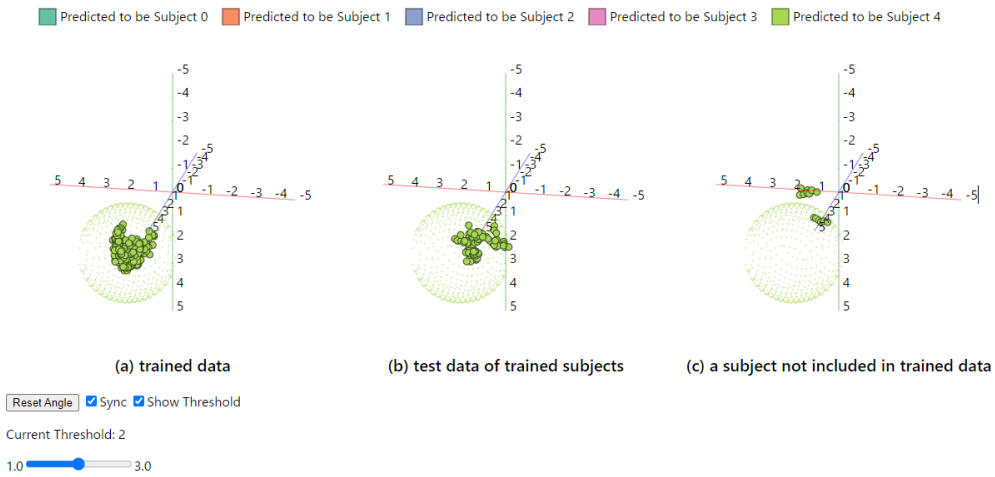


Fig 6　Data predicted to be subject 4 (threshold=2.0)

4. Discussion

From Fig. 3 - Fig. 6, it can be confirmed that the latent variables tend to be biased and distributed near the origin in the order of (a) training data, (b) test data, and (c) data of unknown persons. This indicates that the latent variables of the data with large differences from the training data have the property of being distributed near the origin. In other words, there is a possibility that the distance from the center of gravity of the latent variable in the training data can be used to determine whether the person is the subject or not.

Next, let's see if there is a threshold value at which 100% separation is possible. 1.3 from the reference coordinate is the threshold value in Fig. 3. Therefore, if the threshold is set to 1.3, it is possible to classify subject 2 as a person or a stranger. However, when we look at Fig. 4 (b), we can see that there are several subjects whose data are plotted outside the sphere, even though they are their own data.

When the threshold is changed to 2, the plots are shown in Fig. 5 and Fig. 6. For both subjects, their own data is distributed inside the sphere, but some of other people's data is also distributed inside the sphere. In other words, there is no threshold value that can completely separate the person from the other person in this experiment.

5. Conclusion

In this research paper, I examined how to apply the latent space of CVAE, which I use in my research, to biometric authentication. As a result, I confirmed that it is possible to separate whether the data is the person or not at a certain rate depending on the distribution of latent variables, but it is difficult to achieve 100% accuracy.

In future research, we would like to calculate the "person rejection rate," which is the rate

at which the person is judged to be a stranger, and the "person acceptance rate," which is the rate at which the person is judged to be a stranger, using multiple thresholds, and examine the threshold at which the balance between these indicators in a trade-off relationship is optimal. It is also necessary to understand mathematically why the distribution of the latent variables is affected by the size of the difference from the training data.

6. Refference

- d3-3d(https://github.com/Niekes/d3-3d)
- 3D scatter plot(https://bl.ocks.org/Niekes/1c15016ae5b5f11508f92852057136b5)
- Interactive grouped scatterplot in d3.js(Mouse over to highlight) (https://www.d3-graph-gallery.com/graph/scatter_grouped_highlight.html)
- Categorical legend: square (https://www.d3-graph-gallery.com/graph/custom_legend.html#cat3)