A Framework to Develop Handwriting Neural Networks for Biological Investigation

Sungjae Cho, Taegon Kim

Brain Science Institute, Korea Institute of Science and Technology
Seoul, Republic of Korea

neuromatch conference

KIST Brain Science Institute

Korea Institute of Science and Technology

NMC5
September 27, 2022

1. Introduction

- We aimed to understand **biological motor control** through computational modeling.

- Target motor task: **Handwriting**
  - Implemented by the essential components of sensorimotor integration
  - Handy collection of data by digital tablets

- **Top-down** approach to handwriting modeling

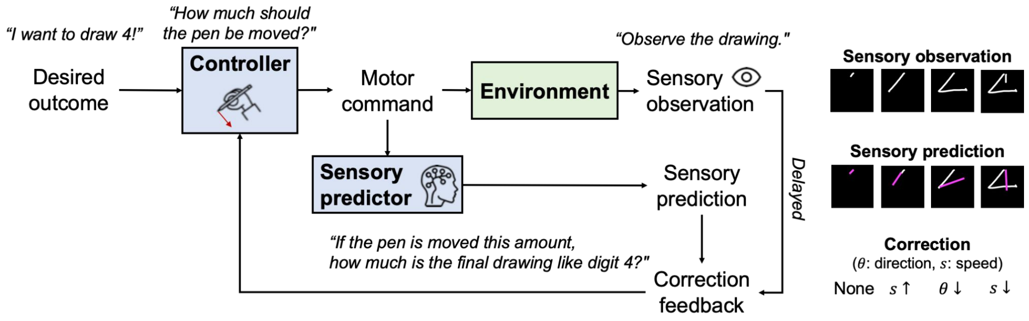


Fig. 1: The roles of internal models in sensorimotor control [McNamee & Wolpert, 2019].

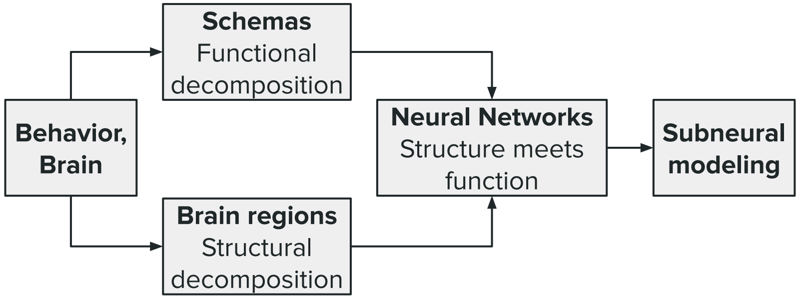


Fig. 2: Levels of analysis behavior and brain [Arbib, 2002].

1. Introduction

- State-of-the-art ANN models for handwriting generation [Graves, 2013; Ha & Eck, 2018]
    - Generated **realistically-shaped** handwriting
    - Overlooked the **realistic behavioral dynamics** of handwriting

- Therefore, we contrived a **framework** to develop neural networks that generate handwriting within human behavioral spatiotemporal scales.
    - This framework uses **metrics** measuring **behavioral difference** from human handwriting.

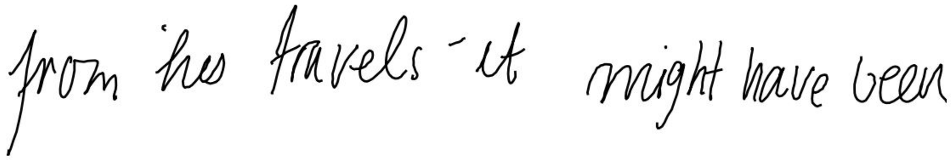


Fig. 3: Handwriting generation by recurrent neural networks [Graves, 2013].

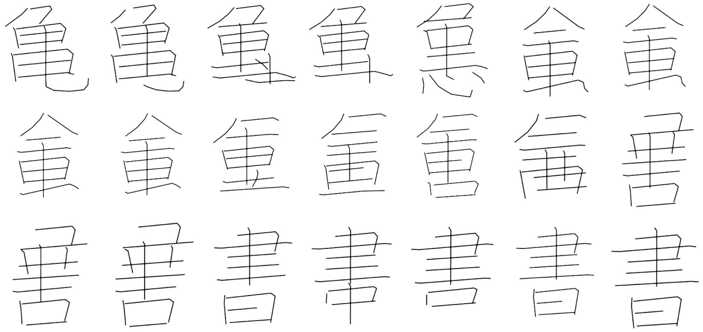


Fig. 4: Complex handwriting generation by recurrent neural networks [Ha & Eck, 2018].

2. Methods

Data collection

- Wacom tablet (PTH-460)
- Min resolution: 1px, 1ms
- 10 digits
- Whole set = 1,000/digit (from one person)
- Sampling rate: 20ms
- Image size: 256x256
- Rescale pixel space by STD: dx /= STD(dx), dy /= STD(dy)

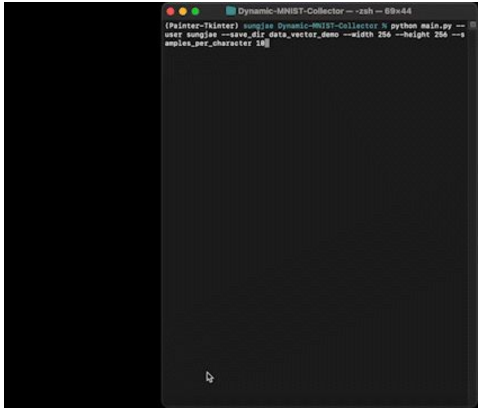


Fig. 5: Handwriting collector.

Neural network

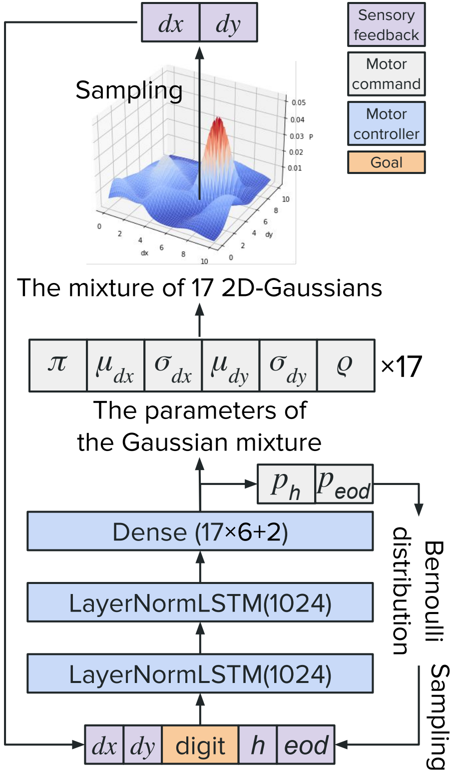


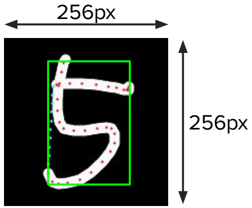
Fig. 6: Our handwriting neural network.

Evaluation

- Strategy: Using statistical difference from human handwriting data wrt the following variables

Spatiotemporal variables

- Duration
- Trajectory length
- Temporal movement
- Width & height



Nearest centroid classification accuracy

- Dynamic time warping (DTW) is a method to align different sequences and compute their difference.
- **Soft-DTW** [Cuturi & Blondel, 2017] is used b/c it yields more realistic motor programs than DTW.

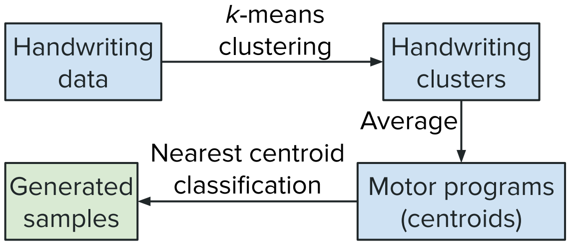


Fig. 7: A workflow of nearest centroid classification for handwriting.

2. Methods | Soft-DTW nearest centroid classifier

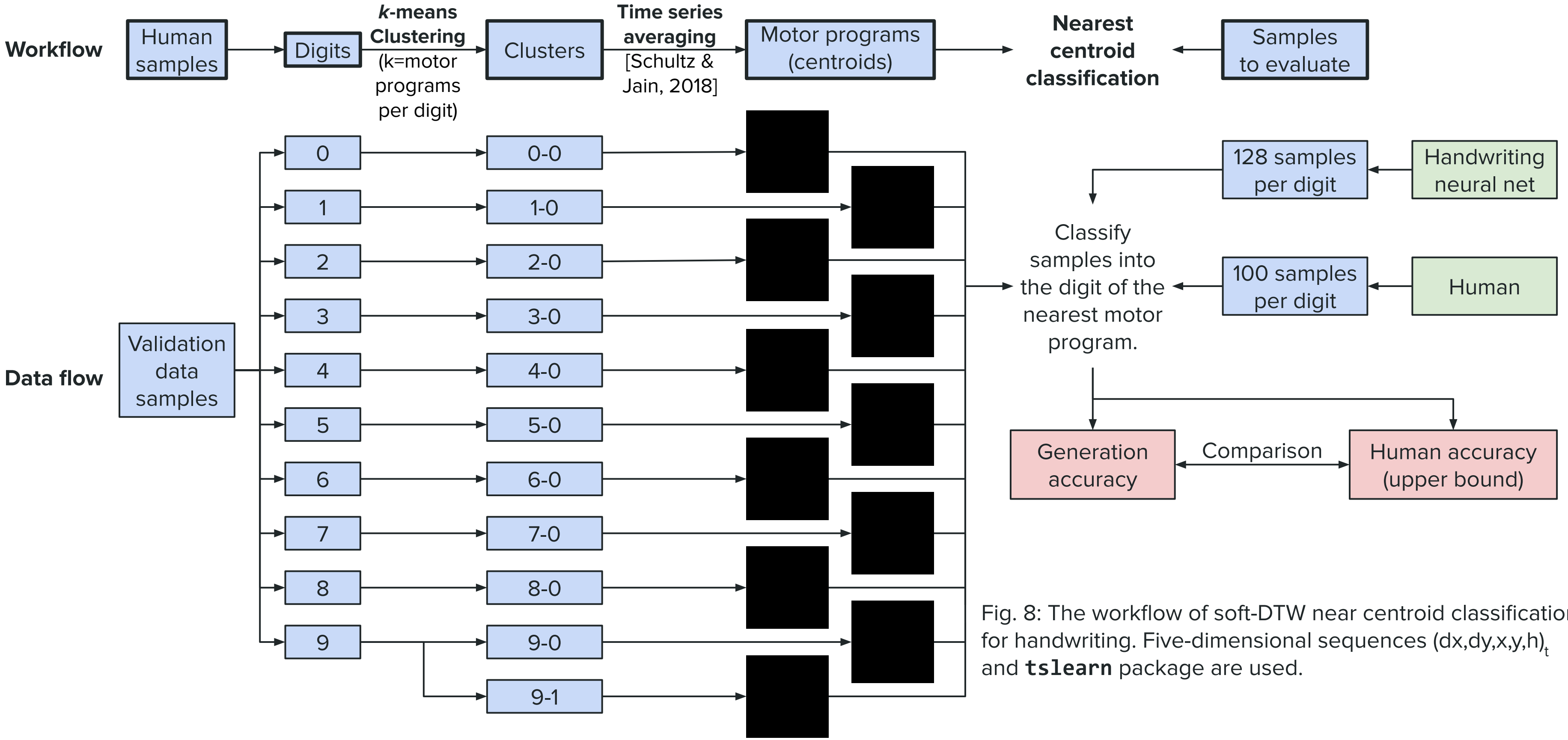


Fig. 8: The workflow of soft-DTW near centroid classification for handwriting. Five-dimensional sequences $(dx,dy,x,y,h)_t$ and `tslearn` package are used.

3. Framework

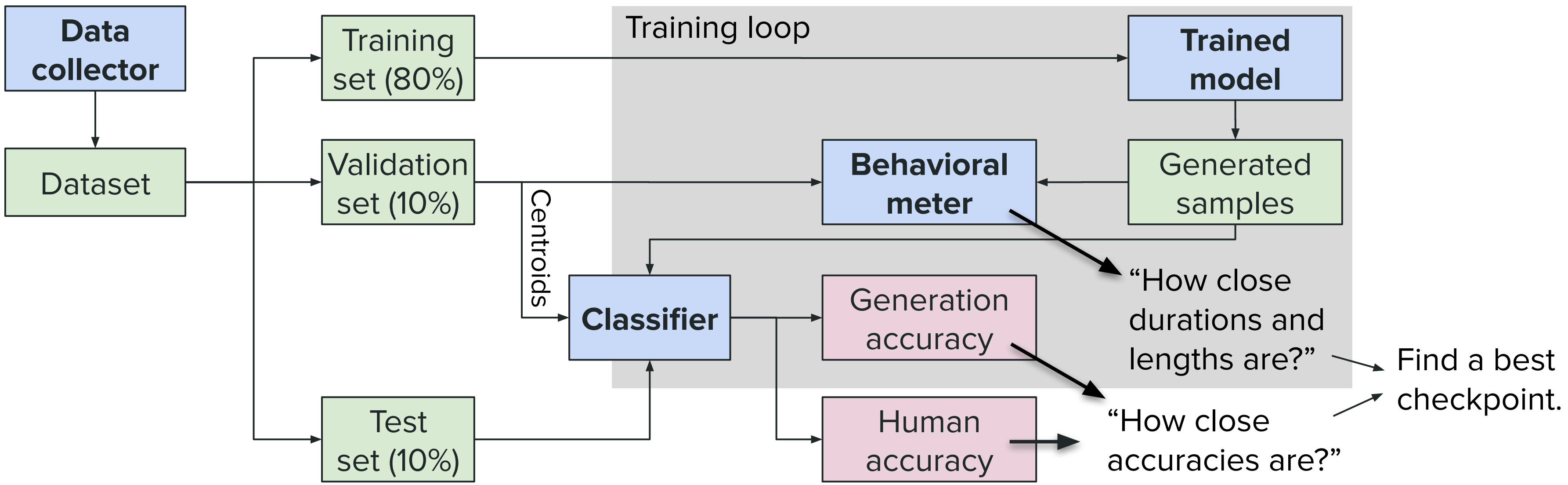


Fig.9: Our framework to develop handwriting neural networks for biological investigation. The suggested methods are combined to the framework that helps finding a best checkpoint, which generates biological plausible handwriting behavior.

4. Results | 4.1. Averaged human motor programs

Fig.10: The averaged human motor programs by time series averaging [Schultz & Jain, 2018] of the validation set.

Thick Blue = Barycenter by soft-DTW

Thin lines = Validation samples

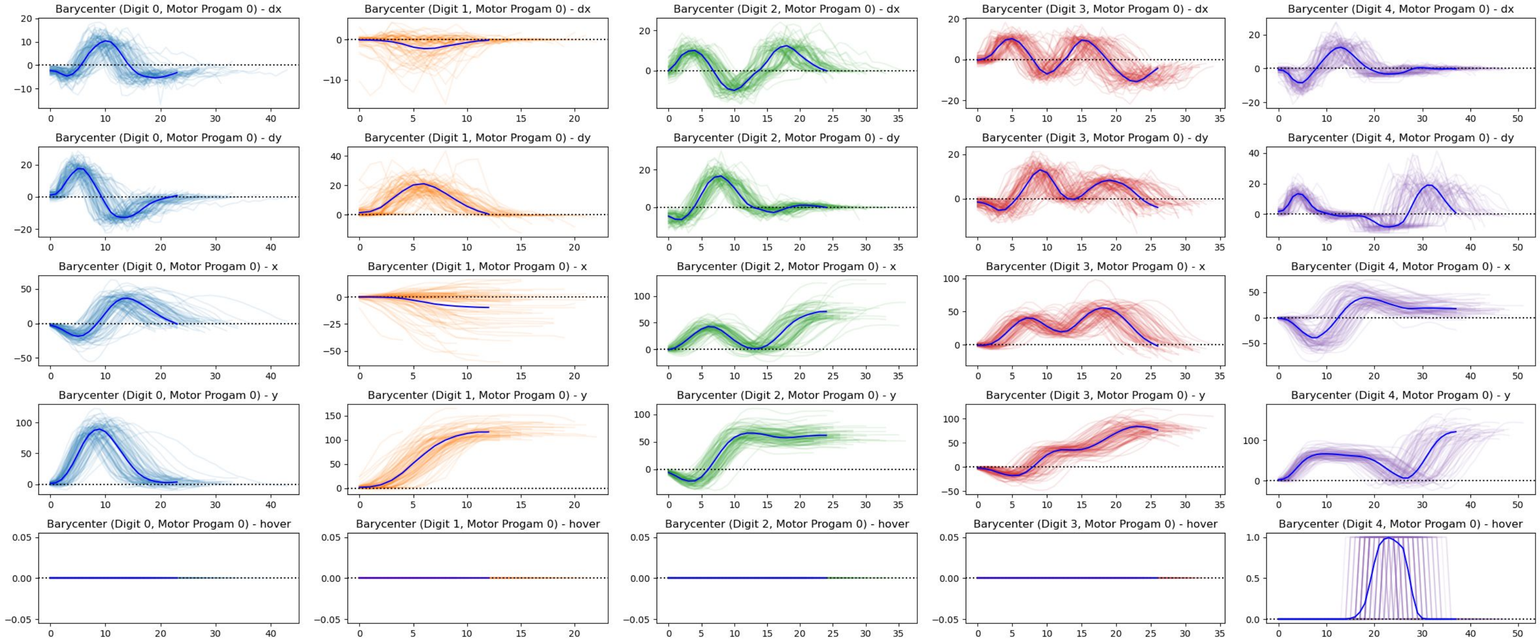
dx

dy

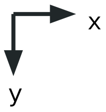
x

y

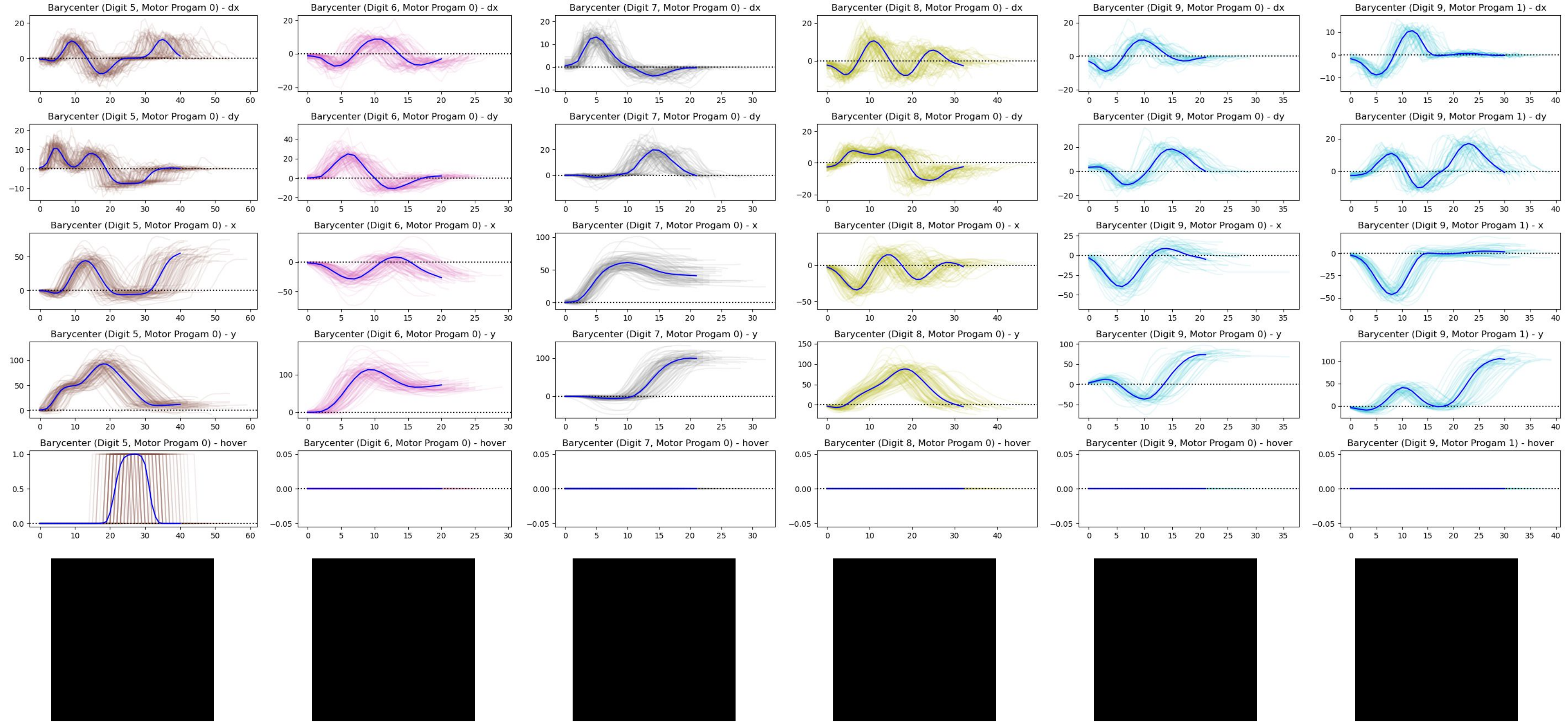
hover



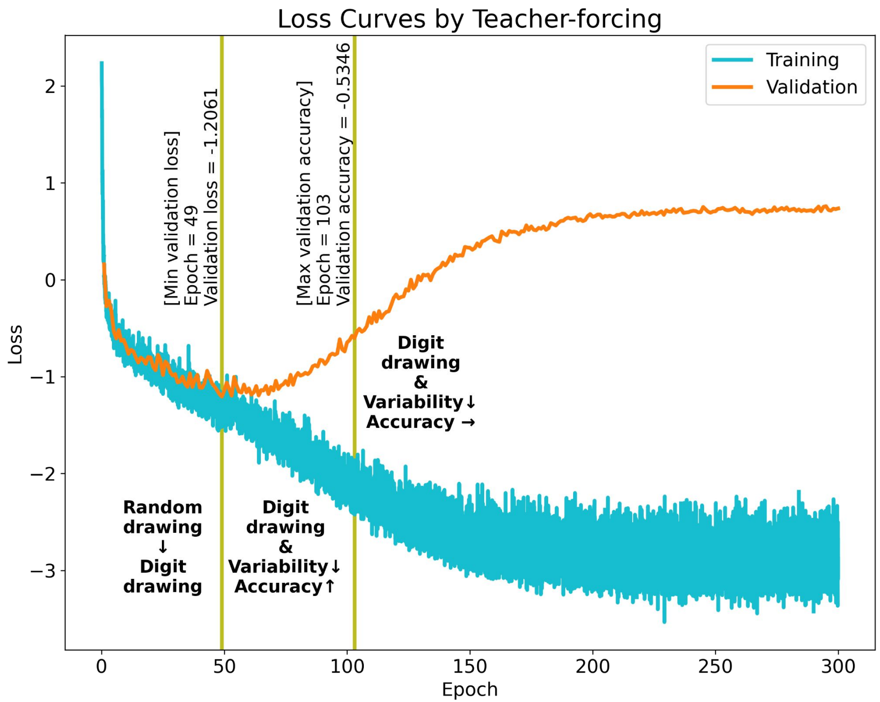
Averaged motor programs

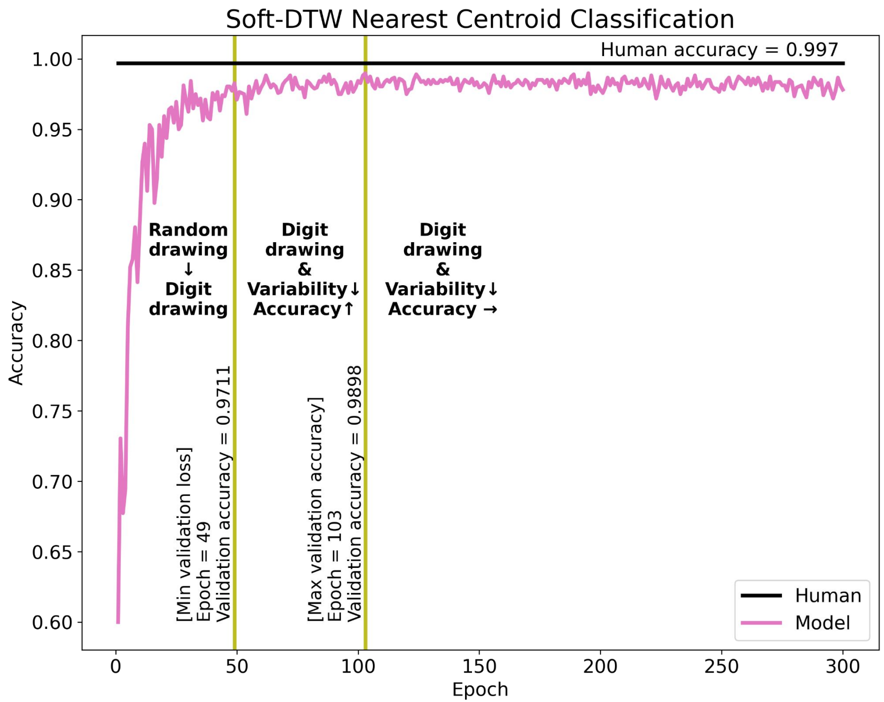


4. Results | 4.1. Averaged human motor programs



4. Results | 4.2. Learning curves and model selection



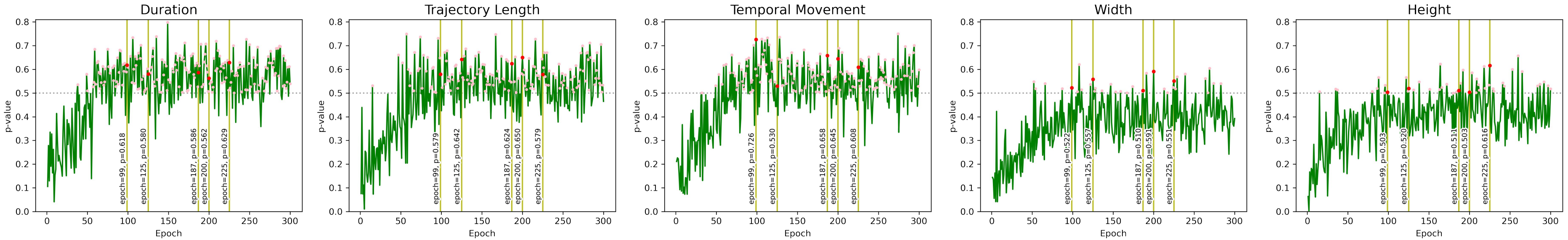


Digit	0	1	2	3	4	5	6	7	8	9	Total
Human Accuracy	1	1	1	.98	.99	1	1	1	1	1	.997
Max Generation Accuracy	1	.99	1	.98	.93	1	1	.99	1	1	.990

4. Results | 4.2. Learning curves and model selection

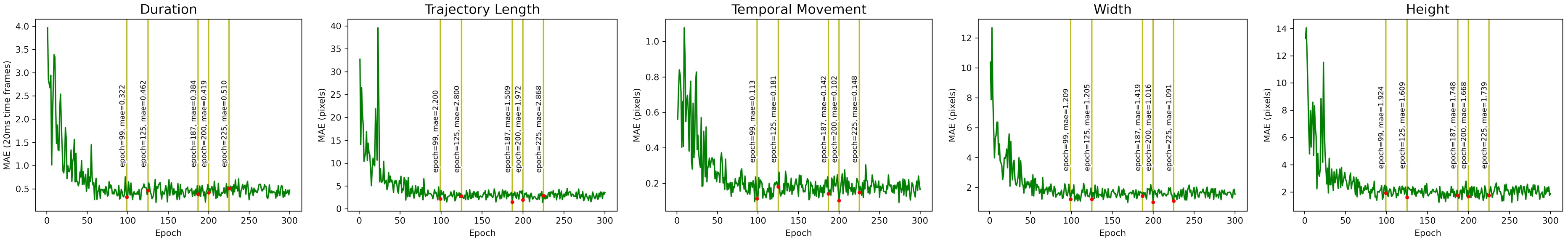
Mann-Whitney U test: Null hypothesis = “Two distributions are identical.”

- Find checkpoints where all p-values > 0.5. ⇒ Epochs: 99, 125, 187, 200, 225.



Mean absolute error (MAE) = $|E[A] - E[B]|$

- For model at each epoch, check MAEs are acceptable. All MAEs are acceptable.



Duration	Trajectory Length	Temporal Movement	Width	Height
max(MAE)=0.51=10.2ms [acceptable]	max(MAE)=2.9 pixels [acceptable]	max(MAE)=0.18 pixels [acceptable]	max(MAE)=1.4 pixels [acceptable]	max(MAE)=1.9 pixels [acceptable]

4. Results | 4.2. Learning curves and model selection

- Pick the checkpoint that results in the highest classification performance.
  - The checkpoint at epoch 225

		Accuracy for each digit set											MAE metrics				
	Epoch	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Total**	**Dur**	**Tra**	**Tem**	**Wid**	**Hei**
Human	-	1.	1.	1.	.98	.99	1.	1.	1.	1.	1.	**.997**	-	-	-	-	-
Generation	99	1.	.99	1.	.98	.88	.99	.99	.98	.98	1	.980	.32	2.2	.11	1.2	1.9
	125	1.	1.	1.	1.	.92	1.	.97	.98	1.	1.	.987	.46	2.8	.18	1.2	1.6
	187	1.	1.	1.	.98	.88	1.	.98	.99	1.	1.	.983	.38	1.5	.14	1.4	1.7
	200	1.	.99	1.	.96	.88	.99	.98	.98	.98	1.	.978	.42	2.0	.10	1.0	1.7
	225	1.	.99	1.	.99	.93	.99	.98	.98	1.	1.	**.988**	.51	2.9	.15	1.1	1.7

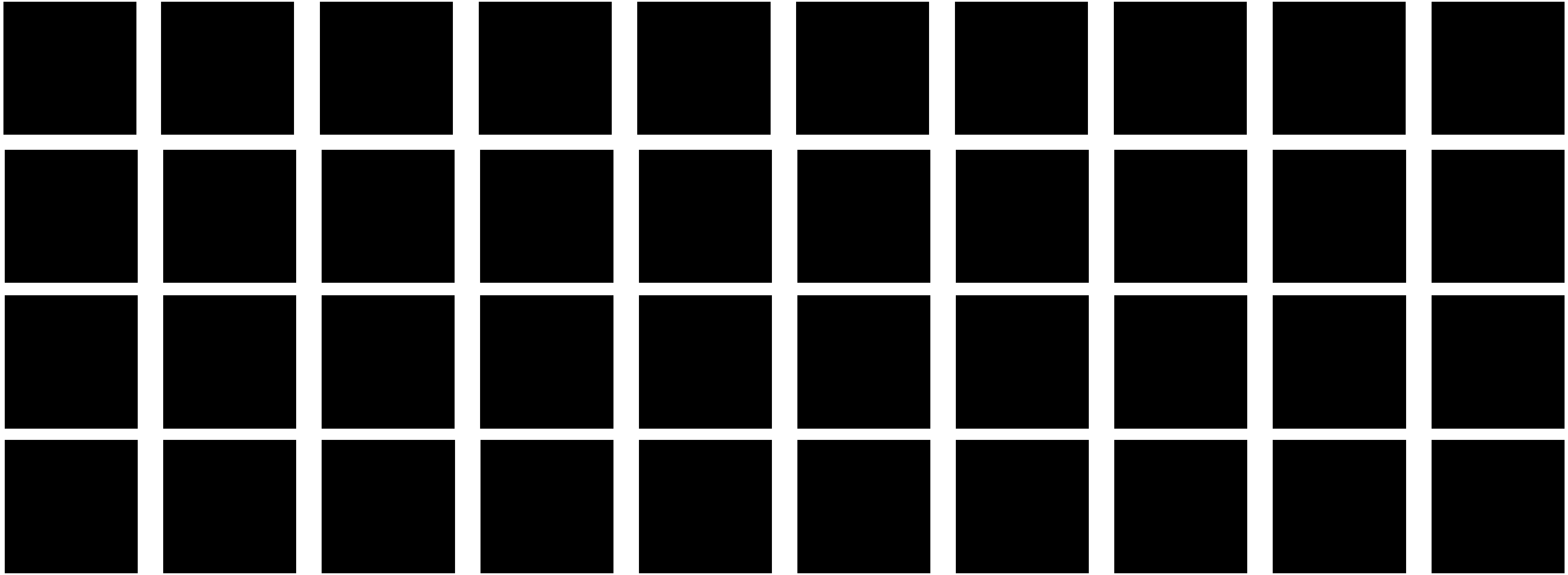
4. Results | 4.3. Generated sequences

Slow

Fast

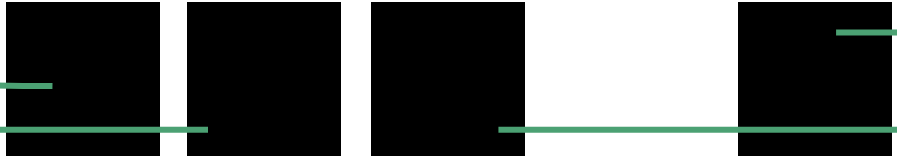
Small

Large



Bad

It rarely happens from the Bernoulli distribution of EOD even though p_{eod} is small.

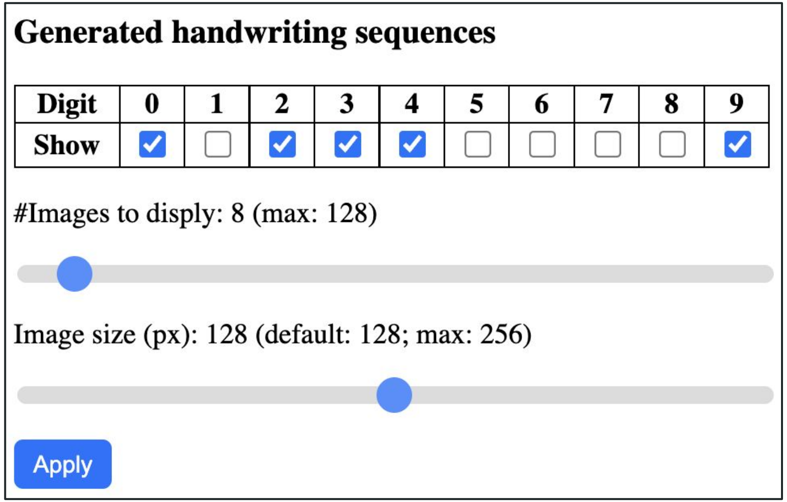


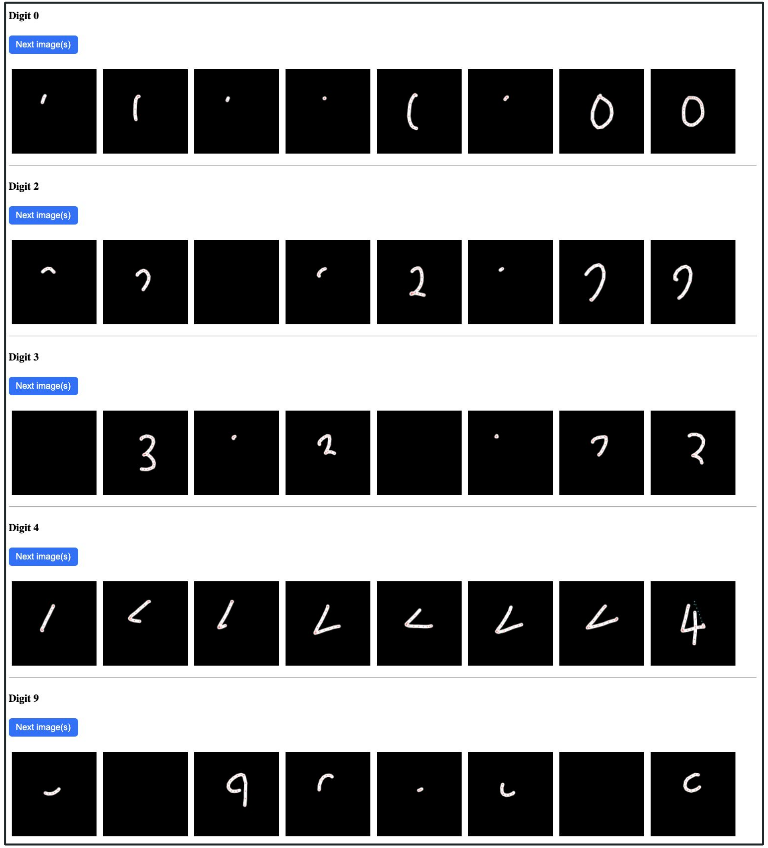
If (dx,dy) is sampled a rare direction, it happens.

4. Results | 4.3. Generated sequences

Visit this **demo page**:
https://sungjae-cho.github.io/nmc2022-handwriting

- More generated handwritings
- 128 generation images per digit
- These samples are not cherry-picked.



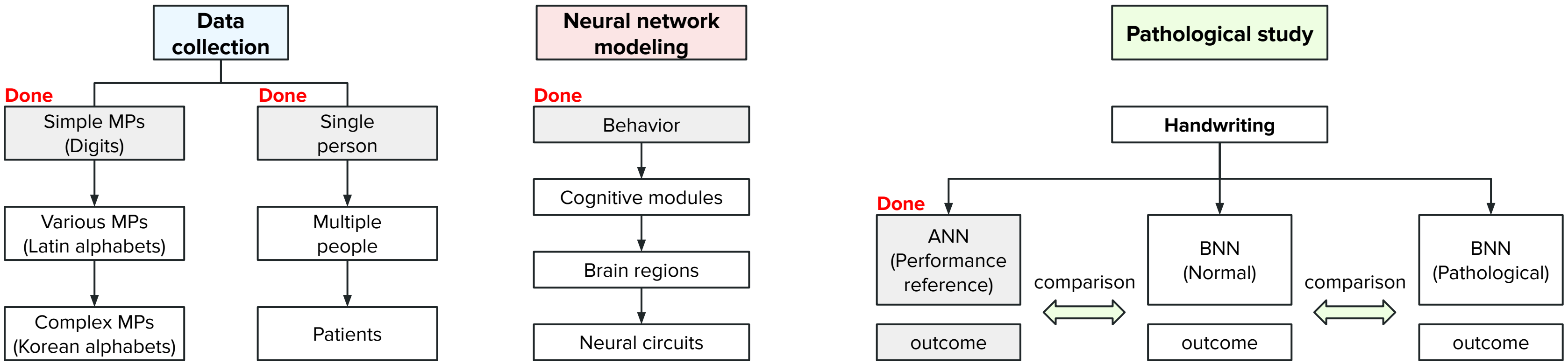


5. Conclusion

Summary

- We developed a **framework** to mimic handwriting with realistic behavioral dynamics.
- The framework includes
  1. **Data collection**: Dynamic handwriting strokes
  2. **Neural network**: RNN composed of LSTM and Gaussian mixture
  3. **Evaluation**: Spatiotemporal behavioral metrics & soft-DTW nearest centroid classifier
- The framework yielded a neural network generating the handwriting motor programs for 10 digits within human behavioral spatiotemporal scales.

Perspective



References

- [LevelAnalysis] Arbib, M. A. (2002). *The Handbook of Brain Theory and Neural Networks*. MIT press.
- [MotorControl] McNamee, D., & Wolpert, D. M. (2019). Internal models in biological control. *Annual Review of Control, Robotics, and Autonomous Systems, 2*.
- [LSTM-GMM] Graves, A. (2013). Generating sequences with recurrent neural networks. *arXiv:1308.0850*.
- [Sketch-RNN] Ha, D., & Eck, D. (2018). A Neural Representation of Sketch Drawings. *International Conference on Learning Representations*.
- [Soft-DTW] Cuturi, M., & Blondel, M. (2017). Soft-DTW: a differentiable loss function for time-series. *International Conference on Machine Learning*.
- [TimeSeriesAvg] Schultz, D., & Jain, B. (2018). Nonsmooth analysis and subgradient methods for averaging in dynamic time warping spaces. *Pattern Recognition, 74*, 340-358.
- [DTW library: tslearn] tslearn.readthedocs.io

End

neuromatch
conference