2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Machine Learning Approaches as An Alternative to Human Psychophysical Tests of Prosthetic Vision

Na Min An, Hyeonhee Roh, Soomin Jung, Eun Ju Kim, and Maesoon Im, *Member, IEEE*

***Abstract*— Microelectronic retinal prostheses can evoke artificial visual percepts to blind individuals. Among several factors, expression of original images using limited resources (i.e., small number of pixels and low gray scales) is critical for improved recognition of artificial vision. To assess effectiveness of new image processing techniques, prosthetic researchers have performed psychophysical tests with normally-sighted subjects. Here, we evaluated facial recognition task performances of two machine learning (ML) models for phosphene images in various conditions, representing several electrical stimulation cases. Both ML models showed correct response ratios with trends expected from psychophysical studies with human subjects. Our results suggest ML approaches would be useful for evaluating newly-developed artificial vision image processing techniques. In particular, the use of ML approaches is expected to substantially reduce time and cost needed for psychophysical testing of phosphene images.**

***Clinical Relevance*— Our results suggest machine learning approaches may replace psychophysical tests of human subjects in verifying new image processing methods for artificial vision.**

I. Introduction

Several retinal prostheses have demonstrated electrical stimulation can evoke pixelated artificial vision in blind individuals. To further improve the quality of prosthetic vision, previous studies have proposed more effective image processing methods [1-3] to represent our complex visual worlds using a limited number of pixels and gray scales, and have performed psychophysical tests with human subjects. Here, we propose the use of machine learning (ML) models to evaluate recognition performances of phosphene images.

II. Methods

We used two representative ML models: Principal Component Analysis and Logistic Regression (PCA+LR) and Convolutional Neural Networks (CNN). The two models were trained with original high-resolution images obtained from AI Hub K-faces (https://aihub.or.kr/) and then were tested with low-resolution phosphene images [4] (Fig. 1A). Test sets of phosphenes were generated by edge-detecting, gray-scaling, downsizing, and re-upsizing original images [2, 4]. We have trained the models with two different numbers of classes: 1) 210 times (10 combination 4) with four faces (classes) out of ten faces to align with human experiment [3], and 2) one time with ten classes to prove expandability of our approach [4].

III. Results

As expected, the correct response rate increased as the phosphene image resolution increased (Fig. 1B). Interestingly, it seems the number of grayscales does not significantly affect the recognition accuracy (compare bars in each color, Fig. 1B). In particular, correct response values saturated earlier for the larger grid sizes with increasing grayscale numbers (see purple bars, Fig. 1B). Between our results with 10 and 4 classes, we observed strong correlation levels (Pearson's r) for both ML models. This suggests performance of a smaller number of classes can be estimated from a larger number of classes (Fig. 1C), removing computational time for creating combinations.

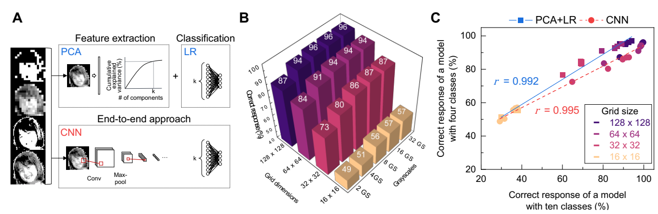

Figure 1. (A) Diagram of the two machine learning models. (B) Performance of the CNN model for various conditions we tested. (C) Scatter plot of correct response ratios between models with four classes and ten classes.

IV. Discussion & Conclusion

Our ML models have showed similar trends in the face recognition performances, which would be expected from human psychophysical tests, suggesting a possibility of replacing psychophysical tests with ML approaches. To further accurately verify this, we will perform recognition tests with volunteers and then compare their performances with those of ML for each phosphene image conditions.

*Research supported by KIST (2E31251), National Research Foundation of Korea (2020R1C1C1006065, 2021M3F3A2A01037366).

N. M. An is with University of Seoul, Seoul 02592, South Korea. H. Roh is with Korea University, Seoul 02841, South Korea. S. Jung is with Ewha Womans University, Seoul 03760, South Korea. E. J. Kim is with Seoul National University of Science and Technology, Seoul 01811, South Korea. M. Im is with University of Science and Technology, Seoul 02792, South Korea. All authors are also with Brain Science Institute, Korea Institute of Science and Technology (KIST), Seoul 02792, South Korea. (correspondence e-mail: maesoon.im@kist.re.kr)

References

[1] M. H. Chang, H. S. Kim, J. H. Shin, and K. S. Park, "Facial identification in very low-resolution images simulating prosthetic vision," *J. Neural Eng.*, vol. 9, no. 4, Jul. 2012, Art. no. 046012.

[2] E. Ho, J. Boffa, and D. Palanker, "Performance of complex visual tasks using simulated prosthetic vision via augmented-reality glasses," *J. Vision*, vol. 19, no. 13, Nov. 2019, Art. no. 22.

[3] M. Vurro, G. Baselli, F. Orabona, and G. Sandini, "Simulation and assessment of bioinspired visual processing system for epi-retinal prostheses," *Proc. 28th Int. Conf. of the IEEE Eng. Med. Biol. Soc. (EMBC),* Aug. 2006, pp. 3278-3281.

[4] S. Dodge and L. Karam, "A study and comparison of human and deep learning recognition performance under visual distortions," *Proc. 26th Int. Conf. Com. Commun. Net. (ICCCN)*, Jul. 2017, pp. 1-7.