

SYMBOLSIGHT: Algorithmic Selection of Visually Distinct Symbol Sets for Bionic Vision Systems

Jasmine Lesner and Michael Beyeler
University of California, Santa Barbara

Abstract—People who use retinal implants often struggle to tell letters apart because of limits in electrode density, signal processing, and how the brain interprets stimulation. Rather than modify the implant, we explore choosing visual symbols that stay distinguishable under prosthetic distortions and mapping them to letters. The approach has four steps: (1) generate many candidate symbols; (2) simulate retinal-implant distortions at several levels; (3) estimate confusion probabilities with pre-trained neural networks; and (4) select a letter–symbol mapping that lowers expected errors using language-specific letter transition probabilities. We target the common single-symbol viewing regime and sketch a hybrid approach in which a few very frequent words have dedicated symbols while the letter mapping covers the rest. Across Arabic, Bulgarian, and English, symbols drawn from diverse sources (DCT, Katakana, Braille) are easier to tell apart than standard letters at higher distortion in our simulations. While user studies are needed, the results suggest that carefully chosen symbol sets may help reading with retinal implants without changes to hardware.

Index Terms—Bionic Vision, Retinal Implants, Visual Prosthesis, Symbol Recognition, Accessibility, Human-Computer Interaction, Simulated Vision, Assistive Reading, Biomedical Signal Processing, Visual Rehabilitation

I. INTRODUCTION

People who use retinal implants often struggle to tell letters apart because of limited electrode density, signal processing, and how electrical stimulation is interpreted in the visual system. Rather than change the hardware, we ask a simpler question: can we choose a small set of visual symbols, and map them to letters, so that the symbols stay distinguishable under prosthetic distortions?

In fluent reading with healthy vision, readers benefit from a broad *perceptual span*, *parafoveal preview*, and the *word-superiority effect* [1]–[5]. Current retinal prostheses, however, typically provide a narrow field of view, coarse sampling, and temporal persistence, which together push users toward sequential, one-symbol-at-a-time viewing [6]. In this single-symbol regime, it is natural to focus on symbols that are easy to tell apart even when distorted.

We explore this idea with SYMBOLSIGHT (Fig. 1), a four-step approach: (1) generate many candidate symbols; (2) simulate retinal-implant distortions at several levels; (3) estimate symbol confusions with a pre-trained neural network; and (4) choose a letter–symbol mapping that lowers expected errors using language-specific letter transition probabilities. We also propose a pragmatic hybrid: a few very frequent words receive dedicated symbols, while the letter mapping covers everything else. This keeps the system simple while allowing common words to be read more quickly.

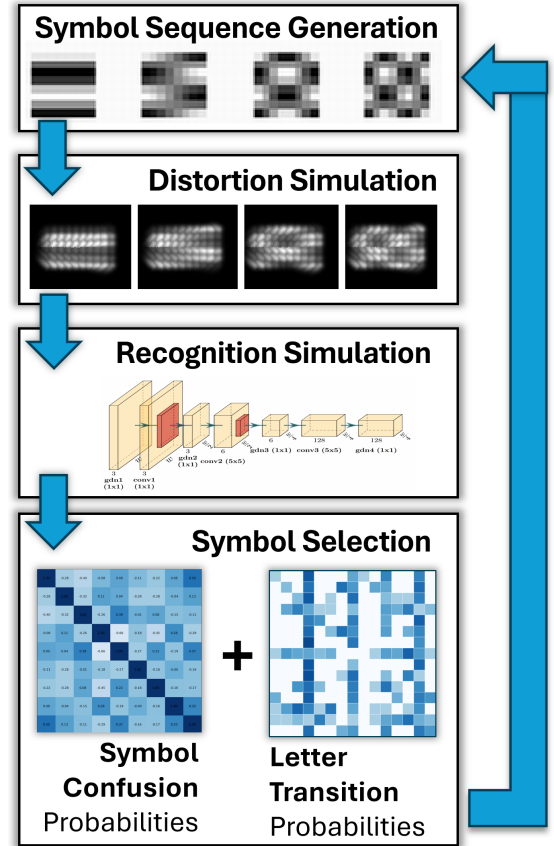


Fig. 1. Our approach has four steps: generate symbols, simulate implant distortions, estimate confusion with a pre-trained neural network, and choose a letter–symbol mapping using letter transition probabilities

In simulations with Arabic, Bulgarian, and English, symbol sets drawn from diverse sources (e.g., DCT, Katakana, Braille) remained more distinct than standard letters as distortion increased. These results are preliminary and depend on models and assumptions; user studies are needed to test whether this approach improves real-world reading with retinal implants.

A. Related Work

Research has explored how visual symbols stay clear in situations with poor viewing conditions. For example, studies on traffic signs have looked at how easy they are to read in different weather, highlighting the need for designs that stay clear amid visual noise and poor visibility [7]–[9]. For example

Clearview typeface was developed specifically to make road signs more readable in low light and bad weather [10], [11].

Similarly, research on airplane cockpits has taught us about symbol recognition in challenging conditions. Cockpit display symbols designed for high-vibration environments must stay readable despite effects like retina image slip and motion blur [12]. Research by Wickens et al. [13] created guidelines for display features that stay distinct even when the human visual system is affected by environmental stress — problems similar to those faced by retina implant users.

These studies help us understand how to design visual symbols that work well despite various distortions. Research in bionic vision has mainly focused on restoring natural vision of symbols by either working with these known distortions [14]–[16] or finding ways to avoid them by improving how the stimulus is encoded [17]–[19].

What’s new about our approach is that instead of trying to adapt or work around the distortions in retina implants, we focus on using the existing implant interface as it is to maximize information transfer. Rather than trying to recreate natural vision, our approach accepts the limits of current retina implant technology and finds sets of symbols (for the letter patterns of a given language) that remain clearly distinguishable despite these limits.

B. Organization

The rest of paper is organized as follows. We first show how different symbol sets are affected by retina implant distortions. We then fine tune neural networks to recognize symbol sets affected by various levels of retina implant distortion to measure the resulting confusion matrices. These serve as a stand-in for how retina implant distortions are likely to cause symbol confusion for users. With these measured confusion matrices, we then demonstrate an algorithm that assigns letters to symbols in a way that minimizes confusion based on letter transition patterns in a given language. The algorithm considers which letter sequences appear most often in a language and assigns them the symbols least likely to be confused by a simulated visual system. This assignment is done for all letters and symbols at once to minimize confusion for an entire language. We conclude by testing our algorithm on three languages: Arabic, Bulgarian and English.

II. METHODOLOGY

When simulating prosthetic vision, we show visual scenes with white on a black background. This increases contrast and makes things easier to see, which is important because current prosthetic vision systems have limited brightness range and resolution. Due to the limits of current retina implants, symbols are shown one at a time, and different implant users see various shapes ranging from small spots to long streaks rather than uniform patterns [6].

This variation happens for several reasons: how much the retina has degenerated affects which cells survive, the distance between electrodes and the retina varies, and passing nerve fibers can be activated. The parameters ρ and λ in their axon map model represent key aspects of how electrical current

spreads: ρ controls how the current spreads perpendicular to nerve fibers (affecting how wide the perceived shape is), while λ determines the spread along nerve pathways (affecting how elongated the shape appears).

A. Symbol Distortion

To make our simulations show a range of possible visual experiences, we used pulse2percept [20] for three levels of distortion (shown top to bottom on the left side of Figure 2) based on Hans et al. methodology [21]:

- **Low Distortion:** no axonal stimulation ($\rho = 100\mu m, \lambda = 0\mu m$),
- **Medium Distortion:** medium sized phosphenes with intermediate axonal stimulation ($\rho = 300\mu m, \lambda = 1000\mu m$),
- **High Distortion:** large phosphenes with strong axonal stimulation ($\rho = 500\mu m, \lambda = 5000\mu m$).

Our simulations use a custom 16x16 electrode grid (with 256 total electrodes) as a middle ground between older retina implants that still have users like Argus II [22] and newer implants that are available or being developed like PRIMA [23].

The right side of Figure 2 shows what happens to the symbols on the left when they are shown rapidly one after another and the leftover image of the previous symbol causes more distortion. This happens because a faded afterimage of the previously viewed symbol remains partly visible. This distortion can greatly affect symbol recognition by creating overlapping visual effects.

B. Symbol Recognition

To measure how these overlapping visual effects impact the ability to recognize symbols distorted by retina implants, we used a neural network (NN) already trained for vision as our simulated vision system.

After testing several pre-trained NNs, we chose MobileNetV3 [24] because it was computationally efficient and shares features with biological vision systems¹.

Examining how distortions affected the symbol sets in Figure 2, we selected three for recognition testing:

- **Katakana Symbols:** These have a simple, attractive appearance and stay recognizable even when blurred.
- **Braille Symbols:** These patterns of dots are used by blind people for reading with their fingers.
- **DCT Symbols:** These are used by image compression algorithms because they can represent key visual features efficiently [27].

To perform recognition testing we modified MobileNetV3 for our needs by replacing its ImageNet classification top layers with layers suitable for our symbol set and only allowing the weights in its top four layers to be changed during our

¹Neuroscience studies show similarities between neural network structures and how biological vision works; for example, certain types of convolutional filters naturally develop center-surround patterns similar to mammalian retina structures [25], and features like shallow paths and skip connections mimic biological neural pathways [26].

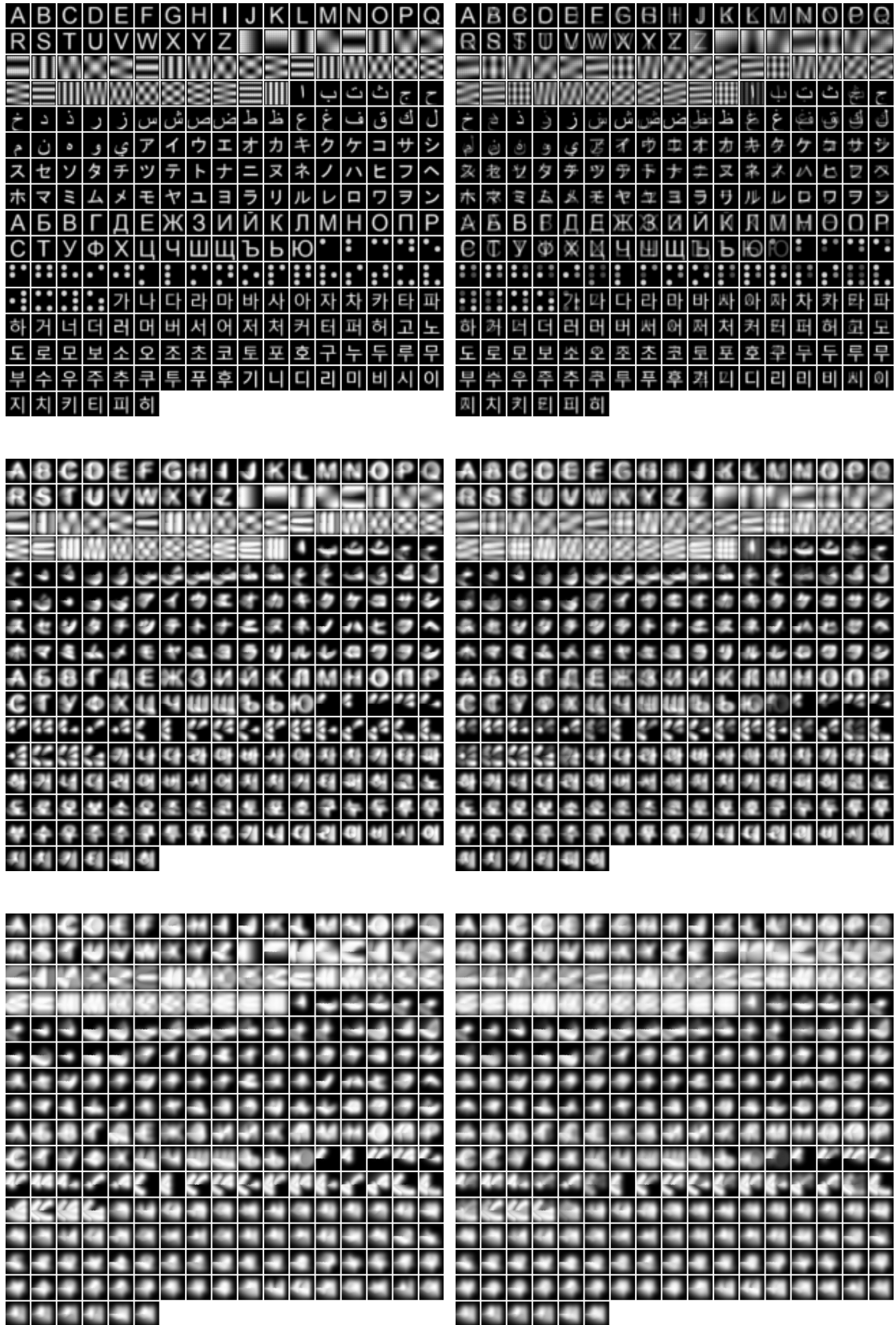


Fig. 2. **Left side:** Seven symbol sets (Latin, DCT, Arabic, Japanese Katakana, Cyrillic, Braille, Korean Hangul) subjected to low, medium and high levels of pulse2percept distortion. **Right side:** Simulation of what happens when symbols are presented sequentially and the perceptual residue of the previous symbol causes further distortion.

fine-tuning. The trained weights in the lower layers were kept fixed to preserve the already learned general feature extraction abilities.

All fine-tuning was limited to seven epochs because our goal was not to train the best possible classifier but to measure which symbols pose the most confusion during learning. We used the Adam optimizer with a small learning rate (0.0001) and gradient clipping (clip value = 0.5), and categorical cross-entropy loss. To prepare symbol images for training, we applied hundred-fold augmentation with ‘mixup’ image augmentation as a way to simulate how images persist in vision [28].

For each of our distortion levels (Section II-A), we fine-tuned a separate MobileNetV3 model keeping everything else the same. For each model, we measured how likely one symbol was to be confused with another. These confusion measurements were then used to guide our algorithm to select the best symbol set.

C. Symbol Selection

To choose which symbols are good to assign to letters, we can use measured confusion probabilities to avoid symbols that are easily confused. And we improve this further by accounting for the letter-to-letter transition probabilities (for given language). Assigning letters that frequently follow each other to symbols least likely to be confused helps reduce reading errors. This requires analyzing both symbol confusion probabilities and letter transition probabilities for a given language.

The optimization problem is: given a letter co-occurrence matrix C ($L \times L$, where L is the number of letters) and a symbol confusion matrix F ($S \times S$, where S is the number of symbols), find a mapping function π that assigns each letter i to a symbol $\pi(i)$ to minimize the cost function:

$$\text{Cost} = \sum_{i=0}^{L-1} \sum_{\substack{j=0 \\ j \neq i}}^{L-1} C_{i,j} \times F_{\pi(i), \pi(j)}$$

This function measures confusion based on two factors: how often letter pairs appear together (e.g., “th” in English) and how likely their assigned symbols are to be mistaken for each other due to retina implant distortions. Multiplying these probabilities for all letter pairs gives a total confusion measure. Our goal is to assign letters to symbols to minimize this confusion.

To solve this problem we developed an adaptive approach that uses different strategies based on the input size:

- For small problems ($L \leq 36, S \leq 100$), we use the Hungarian Algorithm (Kuhn-Munkres algorithm) [29], which guarantees the best possible solution. This approach creates a complete cost matrix for all possible letter-symbol pairings and finds the best overall assignment, though it works best for smaller problem sizes because of its time complexity is $O(L^3)$.
- For medium-sized problems ($L \leq 36, S \leq 500$), we use a local search strategy. This starts by assigning the clearest

symbols to the most frequent letters, then gradually improves the solution by swapping pairs of assignments that reduce the overall confusion cost. This method is faster but may get stuck in locally good solutions rather than finding the best possible solution. The time complexity is $O(L^2 \cdot I \cdot S)$ because each iteration needs to evaluate $O(L^2)$ potential swaps, cost calculation for each swap is $O(S)$ in the worst case, and with I iterations, the total is $O(L^2 \cdot I \cdot S)$.

- For larger problems, it is possible to use a hybrid approach. This first reduces the options by selecting a subset of the most distinct symbols, applies local search to this smaller problem, and further improves the solution using simulated annealing to avoid getting stuck in locally good solutions. The time complexity is $O(L^2 \cdot I \cdot S + L \cdot T \cdot S)$ because symbol subset selection is $O(S \log S)$, local search on the reduced problem is $O(L^2 \cdot I \cdot S)$ and simulated annealing is $O(L \cdot T \cdot S)$ where T = iterations_per_temp * temperature steps.

III. RESULTS

We fine-tuned three different MobileNetV3 neural networks: one for low distortion symbols, one for medium distortion symbols, and one for high distortion symbols. Figure 3 shows some of the symbols we used for training and the confusion probabilities of each trained network. We only show a small sample of the 75,000+ symbol images we used, since how a symbol looks depends on its place in randomly created symbol sequences.

The heatmaps on the left in Figure 3 show how often symbols are confused with each other. The middle diagonal line shows correct identification. Offset diagonal lines appear because similar symbols (mirrored or flipped) have similar features. The faint vertical patterns show which actual symbols get misidentified as others. The box patterns show that DCT symbols are confused with other DCT symbols, Katakana with other Katakana symbols, and Braille with other Braille symbols. This effect is strongest with low distortion and weakest with high distortion when the key visual features become harder to see. Symbols from different sets are less likely to be confused with each other than symbols within the same set, which suggests using symbols from multiple sets together works better.

With one confusion matrix for each distortion level, we tested our symbol selection algorithm on Arabic, English, and Bulgarian languages. We chose these three languages because they are very different from each other, and each has enough text online (from Project Gutenberg + OSCAR [30], [31]) for us to measure letter transition probabilities. Figures 4, 5, and 6 show our measured letter transition probabilities visually. In English, many letter combinations never or rarely happen (shown as light or white areas in the heatmap of Figure 6). These rules about which letters can follow others help English readers understand unclear symbols based on context. Arabic has fewer of these restrictions (fewer light areas in the heatmap of Figure 4), which likely makes it harder to guess unclear symbols from context. Bulgarian falls somewhere in between,

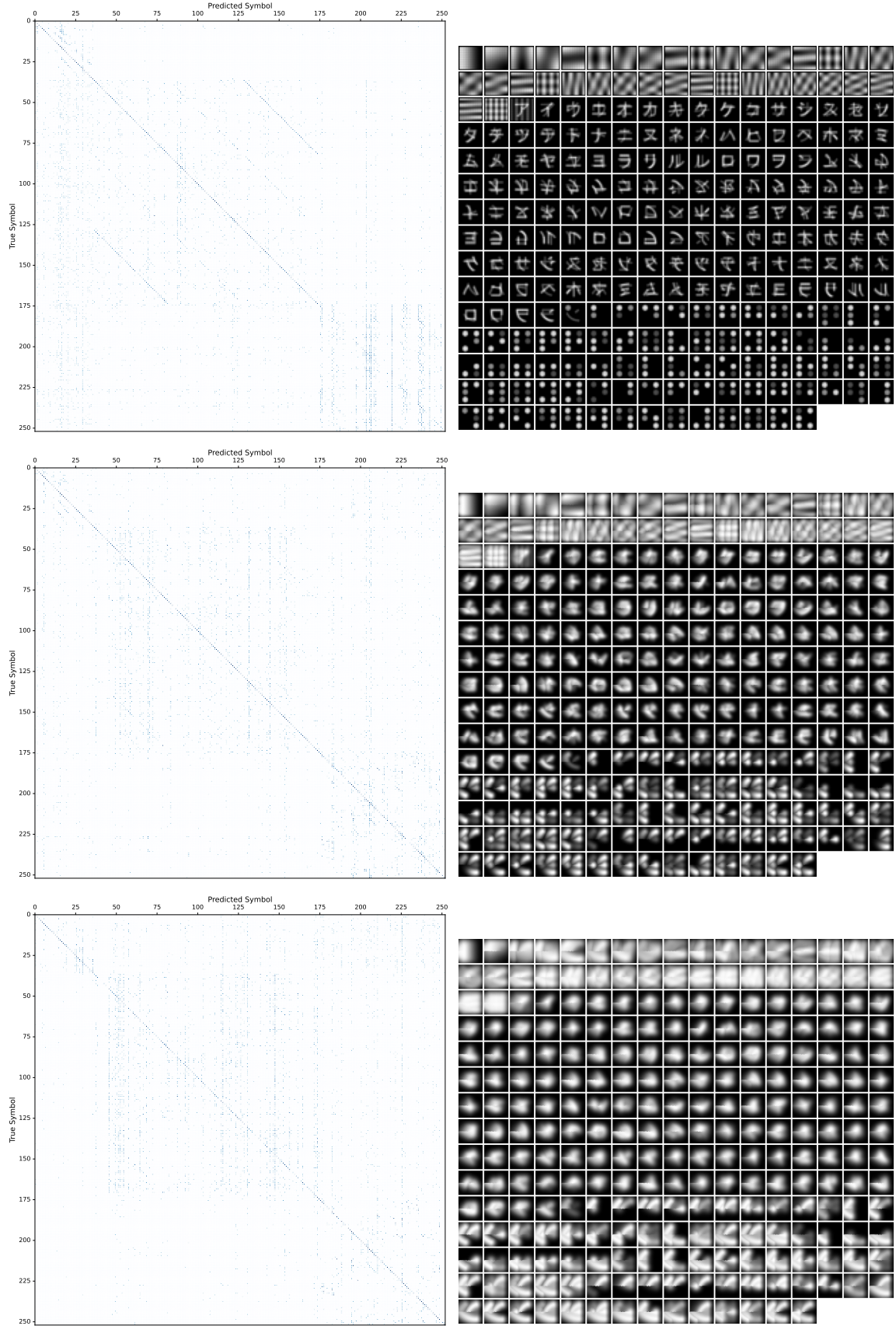


Fig. 3. **Right:** DCT, Katakana and Braille symbol sets (with some mirrored and flipped symbols added) selected for recognition testing. These symbol grids appear from top to bottom at low, medium and high levels of distortion. **Left:** Heatmaps show confusion probabilities of a pre-trained neural network fine-tuned to recognize the symbols to their right.

with some letter restrictions (some light areas in the heatmap of Figure 5) but not as many as English.

Static images don't show what happens when symbols appear quickly one after another (in the typical order of a given language) and the leftover image of the previous symbol causes more distortion. Despite this limitation, by comparing the third row to the sixth row in Figures 7, 8, and 9, we can see that with higher distortion, the new symbols chosen by our algorithm are easier to tell apart for all tested languages. We expect the biggest benefit for Arabic, which has letter symbols that are easily distorted (rows two and three in Figure 7) and few letter sequence rules that could help with unclear symbols.

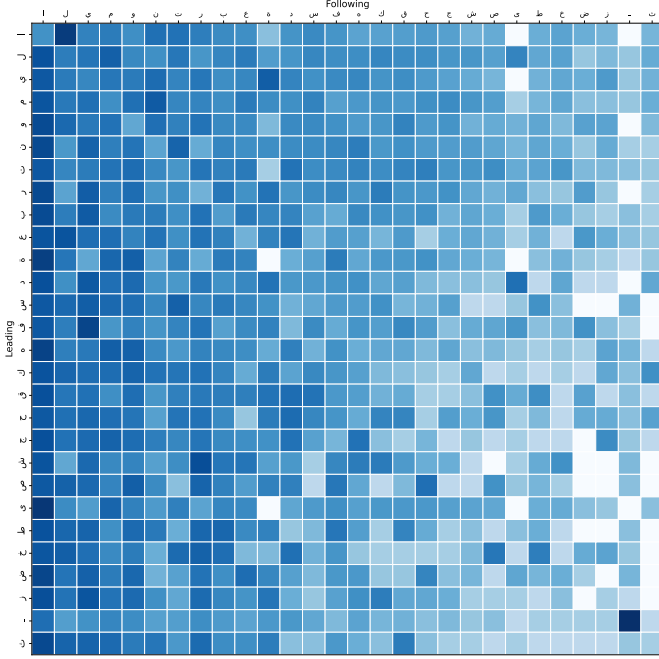


Fig. 4. Heatmap showing Arabic language letter transition probabilities.

IV. DISCUSSION

This work is an early exploration. Our results come from simulations and a neural network stand-in for human perception; we did not run user studies. The models do not capture the full variety of percepts seen with retinal implants, so any claimed benefits should be treated as hypotheses to test, not conclusions.

Symbol design. We sampled a wide but still limited set of symbols. Future work should try a broader pool (e.g., flags, emoji, simple face icons) and also learn new symbols directly with contrastive objectives that push symbols apart under distortion. Mixing sources seemed to reduce confusions; a more systematic search could make this effect clearer.

Distortion modeling. Our simulations covered three distortion levels but left out important realities. Next steps include:

- Varying electrode-grid density to span current and emerging devices,
- Simulating dead electrodes and other “blind spots,”
- Extending temporal effects beyond two symbols to capture longer persistence,

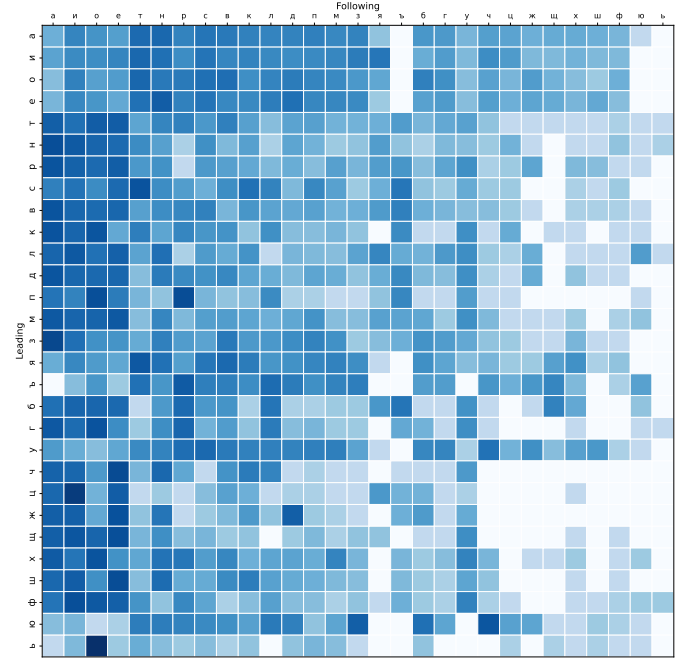


Fig. 5. Heatmap showing Bulgarian language letter transition probabilities.

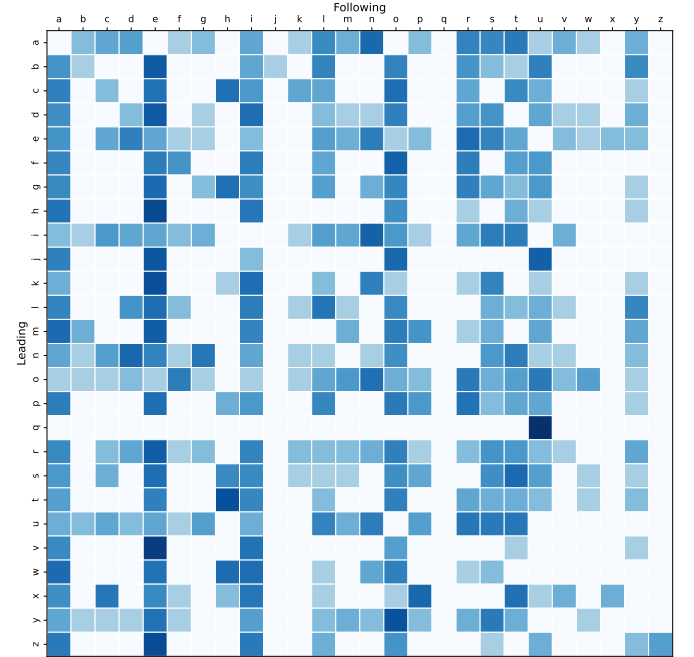


Fig. 6. Heatmap showing English language letter transition probabilities.

- Using pulse2percept's temporal features on symbol-sequence videos.

Recognition model. We fine-tuned MobileNetV3 for convenience, not because it best matches human vision. Architectures that better align with known visual pathways could yield more realistic confusion matrices and help us understand when symbol differences actually matter to users [32].

Letters vs. words. Reading with prosthetic vision is often serial due to narrow field of view, low sampling, and persis-

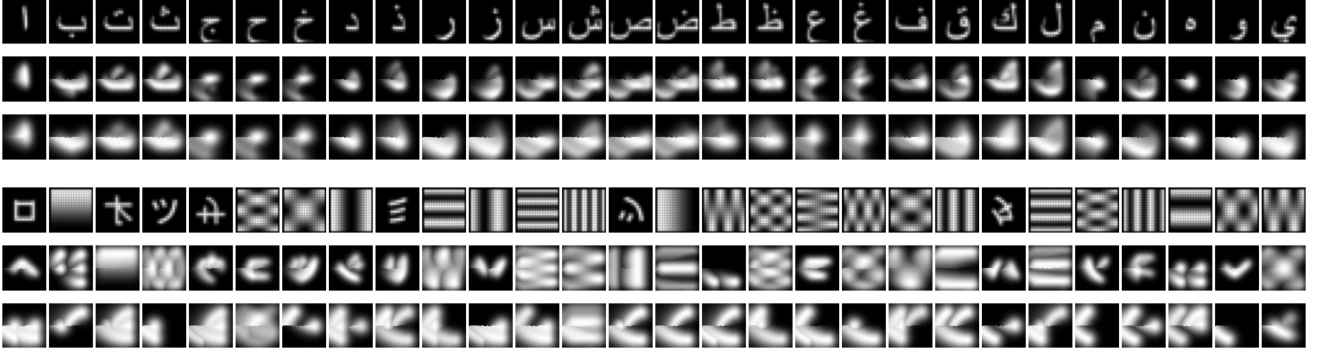


Fig. 7. The first three rows show Arabic letter symbols at low, medium and high distortion. The bottom three rows show the matching output of our symbol selection algorithm.

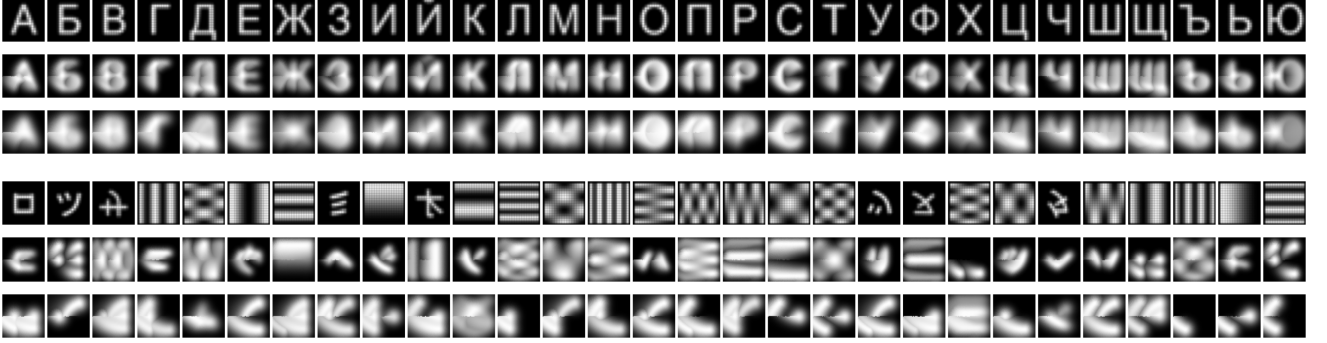


Fig. 8. The first three rows show Bulgarian letter symbols at low, medium and high distortion. The bottom three rows show the matching output of our symbol selection algorithm.

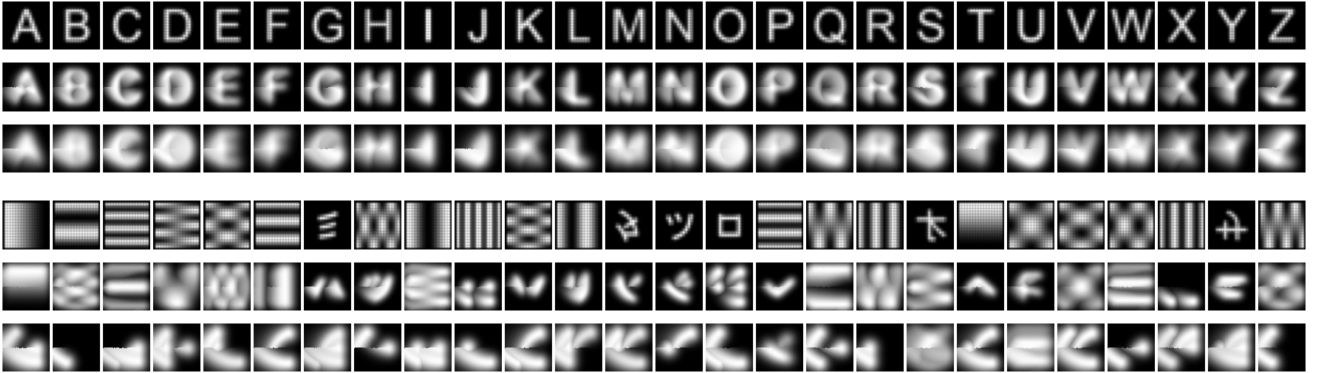


Fig. 9. The first three rows show English letter symbols at low, medium and high distortion. The bottom three rows show the matching output of our symbol selection algorithm.

tence [6]. Our letter-level optimization targets this one-symbol-at-a-time regime while keeping key terms like *perceptual span*, *parafoveal preview*, and the *word-superiority effect* in mind [1]–[5]. A practical extension is a *hybrid set*: give a small number of very frequent words their own symbols and use the letter mapping for everything else. Because word frequencies are Zipf-like, a small word set could cover much of everyday text without overwhelming learning.

Objective function. We optimized using letter bigrams. A direct word-level objective (e.g., making whole words harder to confuse) might better predict reading performance. Comparing letter-based and word-based optimization under

matched distortion would show how much is gained by moving to sequences.

Overall, we see this as a starting point. The next step is careful user testing, richer distortion models, broader symbol searches, and head-to-head comparisons of letter-only versus hybrid symbol sets on speed, accuracy, and learning curves.

V. CONCLUSION

SYMBOLSIGHT takes a simple approach: instead of changing hardware, pick symbols that stay distinct when seen through retinal implants. We follow four steps—generate symbols, simulate implant distortions, estimate confusion with

a pre-trained neural network, and choose a letter–symbol mapping using letter transition probabilities. In simulation, mixing diverse symbols (e.g., DCT, Katakana, Braille) helped Arabic, Bulgarian, and English stay more readable at higher distortion than standard letters.

This is an early result. Our simulations do not capture the full variety of percepts, and neural networks are only proxies for people. We focused on the common single-symbol, serial reading setting and did not run user studies. Next, we should test with implant users, expand the distortion models, and explore more symbol types. It may also help to try a small set of whole-word symbols for frequent words alongside the letter mapping and directly compare speed, accuracy, and learning.

Even with these limits, the findings suggest that carefully chosen symbol sets could make reading with retinal implants easier—without changing the hardware.

REFERENCES

- [1] K. Rayner, “Eye movements in reading and information processing: 20 years of research,” *Psychological Bulletin*, vol. 124, no. 3, pp. 372–422, 1998.
- [2] E. D. Reichle, K. Rayner, and A. Pollatsek, “The e-z reader model of eye-movement control in reading: Comparisons to other models,” *Behavioral and Brain Sciences*, vol. 26, no. 4, pp. 445–476, 2003.
- [3] G. M. Reicher, “Perceptual recognition as a function of meaningfulness of stimulus material,” *Journal of Experimental Psychology*, vol. 81, no. 2, pp. 275–280, 1969.
- [4] D. D. Wheeler, “Processes in word recognition,” *Cognitive Psychology*, vol. 1, no. 1, pp. 59–85, 1970.
- [5] E. R. Schotter, B. Angele, and K. Rayner, “Parafoveal processing in reading,” *Attention, Perception, & Psychophysics*, vol. 74, no. 1, pp. 5–35, 2012.
- [6] M. Beyeler, D. Nanduri, J. D. Weiland, A. Rokem, G. M. Boynton, and I. Fine, “A model of ganglion axon pathways accounts for percepts elicited by retinal implants,” *Scientific reports*, vol. 9, no. 1, p. 9199, 2019.
- [7] J. W. Hutchinson and T. A. Pullen, “Performance of signs under dew and frost conditions,” *Transportation research record*, vol. 681, pp. 16–20, 1978.
- [8] C. F. Paulo and P. L. Correia, “Traffic sign recognition based on pictogram contours,” in *2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pp. 67–70, 2008.
- [9] F.-r. Huo, Y.-r. Feng, and F. Fang, “Legibility of variable message signs on foggy highway: Effect of text color and spacing,” *Displays*, vol. 84, p. 102789, 2024.
- [10] P. M. Garvey and M. T. Pietrucha, “Development of a new guide sign alphabet,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 40, pp. 1285–1285, SAGE Publications Sage CA: Los Angeles, CA, 1996.
- [11] P. J. Carlson and A. Holick, “Maximizing legibility of unlit freeway guide signs with clearview font and combinations of retroreflective sheeting materials,” *Transportation research record*, vol. 1918, no. 1, pp. 26–34, 2005.
- [12] B. D. Adelstein, B. R. Beutter, M. K. Kaiser, R. S. McCann, L. S. Stone, M. R. Anderson, F. Renema, and W. H. Paloski, “Influence of combined whole-body vibration plus g-loading on visual performance,” tech. rep., NASA Ames Research Center, 2009.
- [13] C. D. Wickens, W. S. Helton, J. G. Hollands, and S. Banbury, *Engineering Psychology and Human Performance*. Routledge, 5th ed., 2021.
- [14] N. R. Srivastava, P. R. Troyk, and G. Dagnelie, “Detection, eye–hand coordination and virtual mobility performance in simulated vision for a cortical visual prosthesis device,” *Journal of neural engineering*, vol. 6, no. 3, p. 035008, 2009.
- [15] F. I. Kiral-Kornek, C. O. Savage, E. O’Sullivan-Greene, A. N. Burkitt, and D. B. Grayden, “Embracing the irregular: A patient-specific image processing strategy for visual prostheses,” in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3563–3566, IEEE, 2013.
- [16] A. Bruce and M. Beyeler, “Greedy optimization of electrode arrangement for epiretinal prostheses,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 594–603, Springer, 2022.
- [17] J. Granley, L. Relic, and M. Beyeler, “Hybrid neural autoencoders for stimulus encoding in visual and other sensory neuroprostheses,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 22671–22685, 2022.
- [18] J. Granley, T. Fauvel, M. Chalk, and M. Beyeler, “Human-in-the-loop optimization for deep stimulus encoding in visual prostheses,” *Advances in neural information processing systems*, vol. 36, pp. 79376–79398, 2023.
- [19] Y. Wu, J. Wittmann, P. Walter, and J. Stegmaier, “Optimizing retinal prosthetic stimuli with conditional invertible neural networks,” *arXiv preprint arXiv:2403.04884*, 2024.
- [20] M. Beyeler, G. M. Boynton, I. Fine, and A. Rokem, “pulse2percept: A python-based simulation framework for bionic vision,” *BioRxiv*, p. 148015, 2017.
- [21] N. Han, S. Srivastava, A. Xu, D. Klein, and M. Beyeler, “Deep learning–based scene simplification for bionic vision,” in *Proceedings of the Augmented Humans International Conference 2021*, pp. 45–54, 2021.
- [22] E. Strickland, “Their bionic eyes are now obsolete and unsupported,” *IEEE Spectrum*, 2022.
- [23] H. Lorach, G. Goetz, R. Smith, X. Lei, Y. Mandel, T. Kamins, K. Mathieson, P. Huie, J. Harris, A. Sher, *et al.*, “Photovoltaic restoration of sight with high visual acuity,” *Nature medicine*, vol. 21, no. 5, pp. 476–482, 2015.
- [24] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, *et al.*, “Searching for mobilenetv3,” in *Proceedings of the IEEE/CVF Winter Conference on Computer Vision*, pp. 1314–1324, 2019.
- [25] Z. Babaiee, P. M. Kiasari, D. Rus, and R. Grosu, “Neural echos: Depthwise convolutional filters replicate biological receptive fields,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 8216–8225, 2024.
- [26] M. Winding, B. D. Pedigo, C. L. Barnes, H. G. Patsolic, Y. Park, T. Kazimiers, A. Fushiki, I. V. Andrade, A. Khandelwal, J. Valdes-Aleman, *et al.*, “The connectome of an insect brain,” *Science*, vol. 379, no. 6636, p. eadd9330, 2023.
- [27] N. Ahmed, T. Natarajan, and K. R. Rao, “Discrete cosine transform,” *IEEE Transactions on Computers*, vol. 100, no. 1, pp. 90–93, 1974.
- [28] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, “mixup: Beyond empirical risk minimization,” *arXiv preprint arXiv:1710.09412*, 2017.
- [29] H. W. Kuhn, “The hungarian method for the assignment problem,” *Naval research logistics quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.
- [30] Project Gutenberg, “Project Gutenberg,” n.d. Retrieved March 15, 2025, from urlwww.gutenberg.org.
- [31] J. Abadji, P. Ortiz Suarez, L. Romary, and B. Sagot, “Towards a Cleaner Document-Oriented Multilingual Crawled Corpus,” *arXiv e-prints*, p. arXiv:2201.06642, Jan. 2022.
- [32] F. Sarfraz, E. Arani, and B. Zonooz, “A study of biologically plausible neural network: The role and interactions of brain-inspired mechanisms in continual learning,” *arXiv preprint arXiv:2304.06738*, 2023.