Dulce Karina Pimentel Hurlburt
Deciphering Cognitive Complexity: Investigating Neural Networks and Cognitive Processes
05/8/2024
COGS 123
Professor Dave Noelle

Deciphering Cognitive Complexity: Investigating Neural Networks and Cognitive Processes

Abstract

This study investigates cognitive complexity using a cortical network model, ABAC, which simulates paired associate learning tasks and exhibits catastrophic interference. This research explores how neural networks can mimic biological systems to understand cognitive processes. Previous studies have demonstrated the versatility of neural networks in linguistic tasks (Chikuri, 2021) and working memory implementation (Hoerzer, 2012), and addressed challenges such as perceptual illusions like the McGurk effect (Iqbal, 2023). The McGurk Effect is a perceptual phenomenon discovered by Harry McGurk & John MacDonald in 1970 where they observed that when participants heard the sound 'ba-ba' but saw lips gesturing 'ga-ga' that their brain would fuse these inputs together to created a perceived input of 'da-da' which is different from either input alone (McGurk, 1976). Building upon these findings, this study focuses on understanding the impact of visual and auditory processing overload on recognition accuracy. Results show that when stimuli were presented individually, auditory stimuli exhibit higher initial accuracy than visual stimuli, but visual stimuli show potential for improvement over time. However, when stimuli combined without prior training, the error rates initially decrease but begin to increase over time. These findings mirror real-world scenarios, such as communication through video calls, where auditory cues play a dominant role in facilitating accurate speech comprehension, while visual cues alone appear insufficient. The study highlights the brain's reliance on both auditory and visual cues in understanding speech, exemplified by phenomena like the McGurk effect. Findings contribute to bridging gaps in understanding neural network simulations of human brain processes during complex cognitive tasks and can inform therapeutic approaches for speech and language disorders. However, limitations exist in cross-cultural studies, emphasizing the need for further research to address these complexities.

Introduction

The literature on neural networks in pattern recognition tasks showcases the interdisciplinary field between neural networks, artificial intelligence and cognitive science. While these tasks have been separate fields for quite some time, this literature explores how neural networks interact with biological systems and their contributions to understanding cognitive processes and artificial intelligence. Previous studies explored natural language processing using the Legendre Memory Unit architecture which exhibits neural networks versatility with linguistics (Chikuri & Eliasmith, 2021). Other neural network architectures such as Hebbian learning, demonstrate how biological neural networks implement working memory and pattern generalization (Hoerzer, Legenstein, & Maass, 2012). These architectures are important to tackle perceptual illusions such as the McGurk effect which showcases the difficulty with contextual influences shaping human perception (Iqbal, 2023).

Processing perception with the difficulty of contextual influences is shown not to be specific to biological neural networks due to shown limitations of autonomous car systems that imply difficulties in processing while out in complex urban environments which increases error in pattern recognition (Nugraha, 2017). Nevertheless, neural network frameworks like the Caffee architecture have demonstrated enhanced accuracy in managing pattern recognition tasks despite ambiguity, closely resembling cognitive pattern recognition processes in the human brain using visual patterns from preprocessed information (Sladojevic, 2016).

While there is a discourse about which neural network architect is most efficient, Sladojevic's high accuracy achieved in recognizing patterns in plant diseases presents limitations other architects face when tasked with different contexts and challenges for pattern recognition tasks (Nugraha, 2017). To effectively address the challenges posed by ambiguity in neural networks, a larger training dataset is proposed to mitigate the risk of overgeneralization and to enhance the capability of the network to recognize diverse patterns (Sladojevic, 2016)**.** Iqbal's research provides insight on how computational models can integrate auditory and visual information similar to how humans handle multisensory integration tasks in noisy environments (Iqbal, 2023). Neural networks provide a unique avenue for delving into the intricate processes of pattern recognition in the human brain, leveraging reinforcement learning and perceptual processing as powerful tools for exploration.

These insights are especially important in informing neural networks' ability to handle ambiguity in biological neural networks and uncertain inputs in artificial intelligence. Overall, there is not much discourse on neural networks' ability to handle pattern recognition tasks as demonstrated by the studies on plant disease pattern recognition, language processing, and other neural architectures (Sladojevic, 2016). These researchers have explored and aided the research of how these mimic biological processes with sensory information.

Examining various neural network architectures and their correlation with biological systems aids in uncovering insights of cognitive processes using artificial intelligence. By using neural networks and exploring phenomena such as the McGurk effect, challenges faced by autonomous car systems, and pattern recognition models, the complexities of contextual influences on perception and language processing are highlighted. Past studies have examined the efficiency of neural network frameworks such as the Caffee architecture in managing ambiguity and enhancing pattern recognition accuracy. For example, research on plant disease pattern recognition and multisensory integration tasks demonstrate neural networks' capabilities in mimicking biological processes and handling uncertain inputs (Jiménez, 2023).

To further our understanding, in this study, I will focus on the ABAC neural cortical network which explores the classic paired associate learning task. The ABAC network offers a unique insight into how visual and audio processing overload impacts recognition accuracy through catastrophic levels of interference. By using the ABAC network, valuable insights into complex cognitive processes will be gained, essential to bridging gaps in our understanding of how these processes arise in the brain. Additionally, I will discuss the significance of neural networks in investigating the complex processes of pattern recognition in the human brain, leveraging perceptual processing as a powerful tool for exploration. The first hypothesis is that auditory stimuli will have greater accuracy over visual stimuli. The second hypothesis is that when I attempt to replicate the McGurk Effect with the combined stimuli list, it will lead to degradation in network performance on the AB list as it learns associations from the AC list, reflecting interference effects.

Methods

This model employs a cortical-like network that contains input, hidden, and output layers. This model used distributed representations of random bit patterns to encode stimuli. The hidden layer is where pattern associations are formed and the output layer corresponds to the stimuli. The learning algorithm this model uses is a combination of inhibition and Hebbian learning dynamics which mimics competitive learning. The default parameters refer to the learning rate, initial random weight values, and hidden layer inhibition. For this experiment, the parameters this model embodies will remain at their defaults as it is not our main focus. This model utilizes paired associates learning task data, which I will refer to as the AB-AC pairs where A is associated with two different responses (B and C) to form associations between them. This approach enables the investigation of how conflicting visual information influences the perception of auditory speech sounds and allows for the exploration of strategies to mitigate interference to improve accuracy in speech perception.

The associated responses list (B and C) will be altered to correspond with sample stimuli. The B list refers to the stimuli sound and will contain activations in the two leftmost columns totalling one activation in each B pattern. These patterns are changed within the ‘input’ layer.

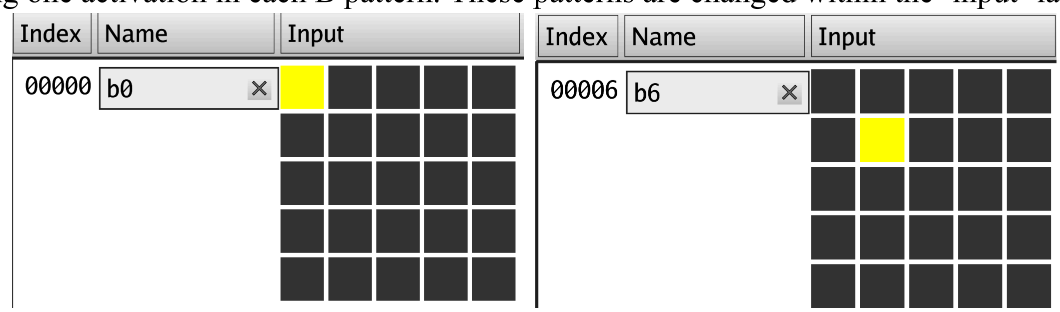

Figure 1: Showing example patterns in the input layer for B list in b0 and b6

The C list refers to the stimuli sight and contains activations in the two rightmost columns totalling one activation in each C pattern for a total of ten patterns in each list. These patterns are changed within the ‘input’ layer.

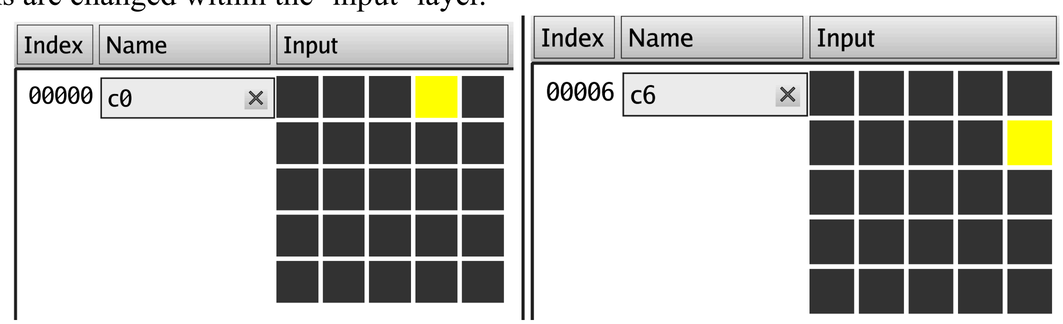

Figure 2: Showing example patterns in the input layer for the C list in c0 and c6

The stimulus materials represent visual and auditory stimuli that human participants engage with when handling the McGurk Effect. Sight patterns represent video clips showing a person speaking different phonemes along with image sets depicting lip movements, and sound patterns represent audio clips of spoken words of spoken phonemes. These stimulus materials are represented in separate response lists either B or C to observe the model's accuracy when it is trained and run on one stimulus at a time. In our third trial, sound list B will be manipulated to include the C list patterns within the B patterns to introduce different levels of ambiguity to the visual information.

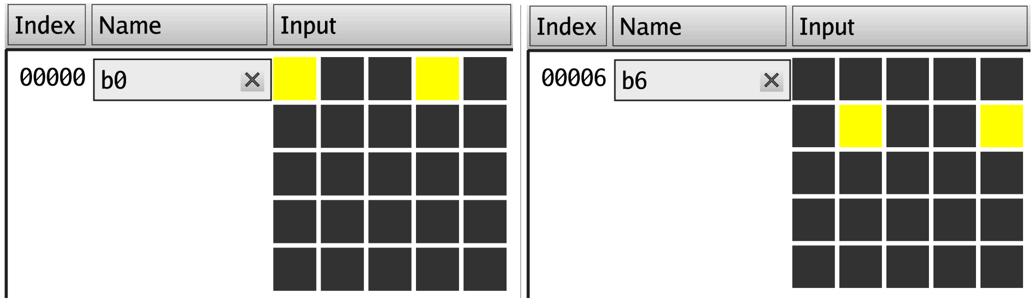

Figure 3: Showing example patterns of b0 and b6 in the input layer for AB/AC list together within the B list

The stimulus materials contain incongruent visual and auditory components to simulate difficulties in processing ambiguity as seen in the Mcgurk Effect. The data collection procedures involve training and testing the model based on the stimulus materials in both the B and C lists. The data collection for the combined stimulus list will involve testing the model without any prior training.

The model will first be evaluated based on its ability to correctly produce the associated response (B or C) given stimulus A. The second evaluation will be based on the model's accuracy based on the incongruent patterns when patterns from C were introduced into pattern B and then run. The key metric for evaluation will be error rate which focuses on the interference between the AB and AC list as well as its ability to learn associations despite interference. The data analysis is conducted utilizing analysis of variance (ANOVA) which is used to determine the difference in error rate in the varying separate and combined list. Using these values, I will depict my results in the form of bar and line graphs using R Studio to visualize the data from the downloaded csv files.

Results

The experiment used a cortical-like network model called ABAC consisting of input, hidden, and output layers. The model utilized a distributed representation of patterns for encoding stimuli. These materials were organized into two response lists (B and C) to observe model accuracy when trained and run on individual stimuli. Additionally, a third trial was introduced by manipulating the B list to include patterns from the C list. The key metric for evaluation was error rate, focusing on the model's ability to learn associations despite interference.

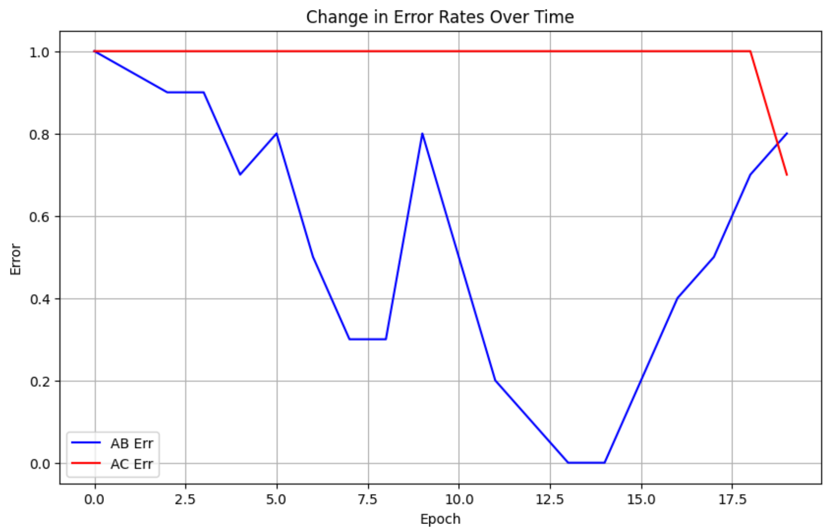

Figure #4 Depicting Change in Error Rates Over Time For Both AB and AC List Ran Individually

Results from the data analysis between AB list on its individual test run and AC list reveal AB Err fluctuates over time whereas in AC list accuracy is consistently incorrect until epoch 15 where it then begins to slowly improve its error rate.

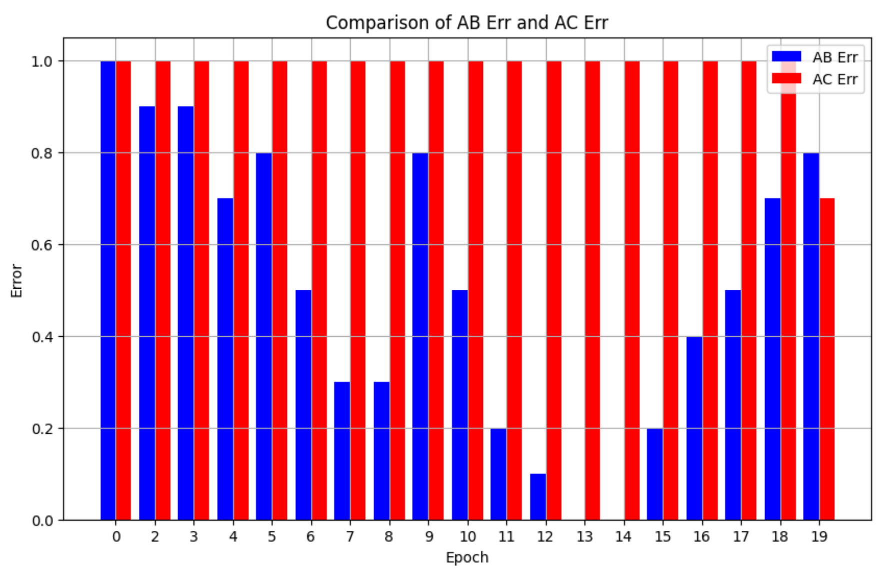

Figure #5 Depicting Comparison of AB Err and AC Err Individual Runs Over Time

Showing the comparison between AB Err which depicts sound pattern and AC Err which exhibits sight pattern, AC list had higher error rates over time having an error value of 1 until epoch 19 whereas AB Err decreased over time until it eventually had 0 error then began having error again.

The third trial was conducted by manipulating the B sound pattern to include the patterns from the C sight list to see how accuracy changed when both stimuli were being processed together. However, it is important to note that this combined pattern list was not trained and it was only run using the test all button. When the stimuli patterns that were originally ran individually were then run together, error rate over time was at a decline and eventually bottomed out at epoch 10 where it then began to have increased error as time went on.

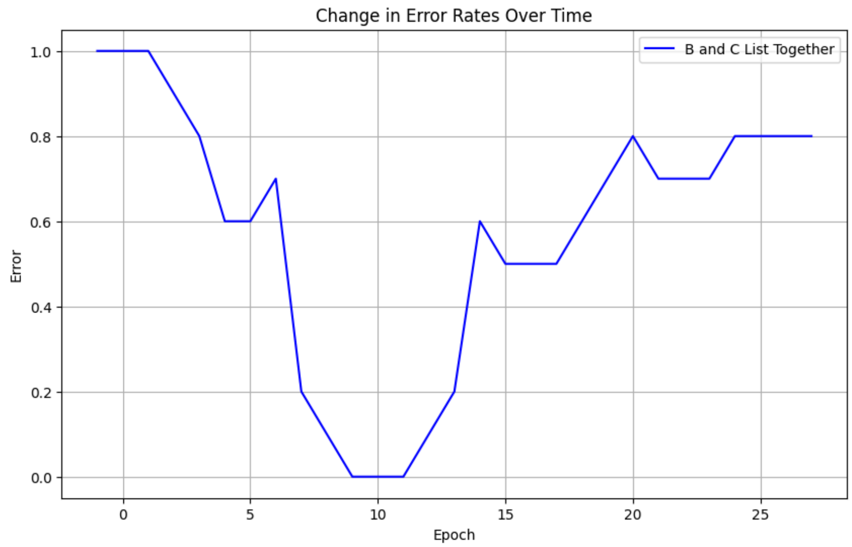

Figure #6 Depicting Change in Error Rate Over Time for AB/AC List Together

Based on the three trials, AB sound list contained the lowest average error rate and AC sight list contained the highest average error rate with AB/AC Err rate being the median average error rate seeing a decline in accuracy performance from AB list but increased accuracy performance from AC list.

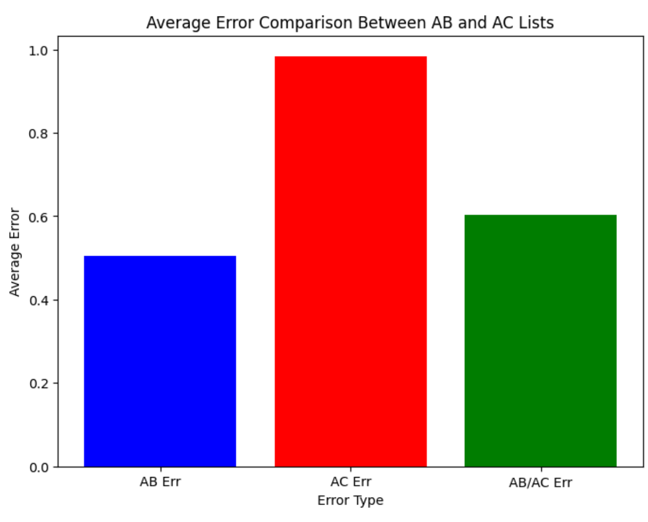

Figure #7 Depicting Average Error Comparison Between The Three Trials.

Discussion

This present study aimed to investigate cognitive complexity through the lens of the ABAC simulation which explores the classic paired associate learning task in a cortical network that exhibits catastrophic levels of interference. By employing this model, I sought out to gain insights into how the brain processes incongruent stimuli, with a particular focus on visual and auditory processing overload and its impact on accuracy.

The results of the study provide insights into the cognitive processes involved in cognitive processing tasks. Findings indicate that when the stimuli were presented individually, the error rates fluctuate over time, with the sound patterns (AB) exhibiting a decreasing trend in error rates, while the sight patterns (AC) showed consistently high error rates until a gradual improvement was observed after epoch 15. This suggests that ABAC may have different effectiveness in processing auditory versus visual stimuli, with auditory stimuli showing greater initial accuracy but visual stimuli showing potential for improvement over time if given the ability to train longer.

This translates to real-world scenarios, such as communication through video calls; the interplay between auditory and visual stimuli influences our ability to comprehend speech. For example, when engaging in a video call where the speaker's face is not visible, but their voice is audible, individuals often rely on auditory cues to accurately interpret spoken words. This phenomenon underscores the durability of auditory processing in facilitating comprehension despite the absence of visual input. However, if the situation were reversed, wherein only visual cues such as lip movements are available without accompanying auditory information, comprehension becomes significantly impaired. For example, when observing a muted speaker's lip movements during a video call, individuals often struggle to discern the spoken words, leading to a high error rate in understanding the message. This disparity in comprehension between auditory and visual stimuli highlights the differential contributions of senses to speech perception. While auditory cues play a dominant role in facilitating accurate speech comprehension, visual cues alone appear to be insufficient for effective communication.

In the McGurk Illusion, McGurk & MacDonald observed that when participants visually saw lip gesturing ‘ga-ga’ but heard ‘ba-ba’ their perception of the combined stimuli resulted in ‘da-da’ (McGurk, 1976). In this experiment, when stimuli from both the sound (B) and sight (C) lists were combined without prior training, the error rates initially decreased but began to increase again after epoch 10. This suggests that while combining auditory and visual stimuli may initially enhance recognition accuracy, there may be a threshold of processing both stimuli that leads to increased errors. These results go against previous studies that studied accuracy regarding the McGurk effect overtime where researchers previously saw that as participants were repeatedly exposed to the McGurk Effect overtime their brains begin to ‘fuse’ the auditory and visual stimuli information which led to more consistent perception as their brain adapted to a combined sensory input (Magnotti, 2024).

Although my hypothesis that the list containing both stimuli would have less accuracy was not correct, my findings are consistent with previous research highlighting the challenges associated with processing ambiguous stimuli, such as in the McGurk effect. The observed fluctuations in error rates and the differential accuracy between auditory and visual stimuli mirror the complexities of perception documented in previous studies (Iqbal, 2023; Nugraha, 2017).

Conclusion

Through the ABAC model, I explored the dynamics of paired associate learning task and the results cognitive overload holds on recognition accuracy. My findings shed light on the differential processing of auditory and visual stimuli, mirroring real-world scenarios such as speech perception during video calls or silent movies.

The insights gained from my research offer possible implications for both theoretical and practical understandings. By analyzing how closely this simulation mimics human behavior and processes, it gives researchers the capability for studying human processes without human subjects.By bridging gaps in our understanding of how neural networks simulate human brain processes, my study contributed to both artificial and neuroscience fields. Through the exploration of the McGurk effect, valuable insights are gained on the complexities of speech perception with potential applications for diagnostic and therapeutic interventions for individuals with speech and language disorders.

However, my study is not without its limitations. For example, the complexity of cross-cultural studies shows the need for further research to explore the differences of cognitive processes across diverse populations. Studies show that Japanese speakers experience a less pronounced McGurk effect over Spanish and English heritage speakers due to factors such as cultural face avoidance, and the tone and syllabic structures of the Japanese language (Tiippana, 2023). While these factors can be studied with real-world human participants, these factors might not be easily researched using this simulation.

Acknowledgements

Thank you Liza Oh for all of your help regarding the simulation and break down of the topics essential to this paper. Your assistance in figuring out the simulation and reading over my paper gave me feedback crucial to my success. Thank you Professor Noelle for your valuable insights on how to best conduct this model and how to implement my ideas into reality. Your feedback regarding my oral presentation and your aiding my simulation troubles allowed me to make sure my paper and simulation modifications properly reflected my research interest. Thank you Professor Munoz for reading over my draft of results for this paper and giving me feedback on how to better convey my findings.

References

By, Perera, A., on, U., 7, S., & Perera Researcher B.A, A. (2023, September 7). The McGurk Effect: Audio-visual speech perception illusion. Simply Psychology. https://www.simplypsychology.org/mcgurk-effect.html

Chilkuri, N. R., & Eliasmith, C. (2021, July). Parallelizing legendre memory unit training. In the International *Conference on Machine Learning* (pp. 1898-1907). PMLR.

Hale, J. T., Campanelli, L., Li, J., Bhattasali, S., Pallier, C., & Brennan, J. R. (2022). Neurocomputational models of language processing. *Annual Review of Linguistics*, *8*, 427-446.

Hoerzer, G. M., Legenstein, R., & Maass, W. (2014). Emergence of complex computational structures from chaotic neural networks through reward-modulated Hebbian learning. *Cerebral cortex*, *24*(3), 677-690.

Iqbal, Z. J., Shahin, A. J., Bortfeld, H., & Backer, K. C. (2023). The McGurk Illusion: A Default Mechanism of the Auditory System. *Brain sciences*, *13*(3), 510.

Jiménez, M., Avedillo, M. J., Linares-Barranco, B., & Núñez, J. (2023). Learning algorithms for oscillatory neural networks as associative memory for pattern recognition. *Frontiers in Neuroscience*, *17*, 1257611.

McGurk, H., MacDonald, J. Hearing lips and seeing voices. *Nature* 264, 746–748 (1976). https://doi.org/10.1038/264746a0

Magnotti, J. F., Lado, A., Zhang, Y., Maasø, A., Nath, A., & Beauchamp, M. S. (2024, April 3). *Repeatedly experiencing the McGurk effect induces long-lasting changes in auditory speech perception*. Nature News. https://www.nature.com/articles/s44271-024-00073-w

Nugraha, B. T., & Su, S. F. (2017, October). Towards self-driving cars using convolutional neural networks and road lane detectors. In *2017 2nd international conference on automation, cognitive science, optics, micro electro-mechanical system, and information technology (ICACOMIT)* (pp. 65-69). IEEE.

O'Reilly, R. C., Munakata, Y., Frank, M. J., Hazy, T. E., and Contributors (2012). Computational Cognitive Neuroscience. Wiki Book, 4th Edition (2020). URL: https://github.com/CompCogNeuro/ed4

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep neural networks based recognition of plant diseases by leaf image classification. *Computational intelligence and neuroscience*, *2016*.

Tiippana, K., Ujiie, Y., Peromaa, T., & Takahashi, K. (2023, August 13). *Investigation of cross-language and stimulus-dependent effects on the McGurk effect with Finnish and Japanese speakers and listeners*. MDPI. https://www.mdpi.com/2076-3425/13/8/1198