Developing Deep Neural Network based Brain Computational Models from Psychophysics Data of some Simple Perceptual Phenomena: Visual as well as Auditory

by

Keerthi S Chandran

A thesis presented for the degree of
Doctor of Philosophy
under the supervision of Dr. Kuntal Ghosh

Center for Soft Computing Research,
Indian Statistical Institute,
203 BT Road, Kolkata-700108,
India

January 2026

Abstract

The thesis, consisting of nine chapters, explores the methodology of building testable brain computational models using Deep Neural Networks (DNN), which are trained by psychophysics data. Psychophysics is the quantitative study of perception of physical stimuli. In psychophysics experiments, one or more parameters associated with the stimuli are changed, and the human subject's responses to the stimuli are recorded. This thesis encompasses both experimental psychophysics works, as well as computational models of the phenomena involved.

The contributory chapters of the thesis start in Chapter 2 with the perspective building of the novel methodology followed throughout this research involving psychophysics on one hand, and deep neural network based brain modeling on the other. This chapter also focuses on the possible reasons for the discrepancies that exist between the functioning of existing deep neural networks, and psychological findings.

The present thesis explores three very simple, yet intriguing perceptual phenomena viz. flicker fusion, flicker wheel illusion and sound symbolism. While the first two are purely based on visual perception, the third, i.e. sound symbolism, as the name suggests, involves both visual and auditory perception. The experiments involved in this thesis concerning all these three perceptual phenomena share one common aspect. Not only do they involve very simple stimuli for conducting the psychological experiments, but also the subject response in all the three cases is binary. For the flicker (a flicker stimulus is a visual stimulus with intermittent illumination) fusion experiment, the subject reports whether the stimulus appears flickering or steady; for the flicker wheel illusion, subjects report whether the wheel is perceived as static or flickering (illusory), while in the sound symbolism experiments, the subject assigns a sound stimuli to one of the two visual inputs, provided.

In Chapter 3 of the thesis, a Convolutional Recurrent Neural Network (CRNN) for modeling the flicker fusion phenomenon has been proposed. It is shown that the model is trainable with psychophysics data, and testable with a wide variety of flicker patterns. Next, in Chapter 4, a DNN model that takes into account the microsaccades in the eye is presented for the Flicker Wheel illusion while also building a novel dataset for this illusion. The sound symbolism phenomena is investigated in Chapter 5, for the difference

in words for round and sharp objects across several natural languages. Here again, both behavioral experiments and DNN based modeling are performed.

Thus, by establishing the efficacy of DNN based brain computational models in explaining these psychological phenomena, the present thesis goes on to further investigate the flicker stimulus, already discussed in Chapter 3, to better explore the strengths and limitations of the present brain computational modeling approach.

To this end, first in Chapter 6, the CRNN model is used to probe the relation between the psychophysics and brain electrophysiology involving the flicker stimulus. The work in this chapter, interestingly, demonstrates that many of the reported features of the human electroencephalogram (EEG) response to flicker can actually be explained as being the convolution response to the stimulus, despite the fact that the model is trained with behavioral data only.

Next, to further generate more flicker data, and subsequently put the proposed DNN model to more stringent testing, Chapter 7 of the thesis describes the construction of an indigenous low-cost device that can generate mass psychophysics data on flicker fusion to train and test DNNs. Subsequently, in Chapter 8, the psychophysics data generated from this device was used to train a CRNN. The training yielded symmetric filters, as often found in biological visual systems. The predictions made by the CRNN model on complex flicker patterns were then tested through psychophysics experiments with the device, demonstrating that the model is falsifiable with scopes of further improvement through future research.

Acknowledgments

I thank my supervisor Dr. Kuntal Ghosh for his help in completing the thesis. I thank the professors in RFAC, and PhD/DSc Committee for their help and advice in the completion of the thesis. I am grateful to Prof. Bulusu Uma Sankar for his help in preparing the final presentation of my work. Late Prof. Daniel Dennett had helped me by sending me the reference to a work he mentioned in one of his speeches. I thank Prof. Bijay Bal for his advice on constructing a device, as a part of work of the thesis. I thank Prof. Sankar Kumar Pal, Prof. Ashish Ghosh and Prof. Subhra Sankar Roy as well as other professors in Indian Statistical Institute. I am also grateful to many anonymous reviewers whose comments and criticisms helped me improve my work over the time.

I had worked with Shibsankar Roy for development of the device used in the thesis. Ms. Sumana Ghosh had worked with me in the initial prototype of the device. I am grateful to Sreeja Pal for the work done in collaboration on the topic of sound symbolism. I also thank Ms. Amrita Mukherjee, Avijit Paul, Dr. Swati Banerjee for their helps.

I thank my fellow Research Scholars in Center for Soft Computing Research, Dr. Anjan Chowdhury and Joginder Singh for their helps. I thank all my friends and fellow research scholars in the institute for my help during the thesis. Dr. Umakant Sahoo had been helpful and offered his advices during the work of my thesis. I also thank Dr. Sumit Verma for his helps during my time at Indian Statistical Institute.

I would like to thank Google Cloud Research Credits program for allotting me credits to use machines for training my data via deep learning. I thank those online who helped me get some papers, as well as Sci-Hub for the same. I thank Ms. Sonia Gidla Medidi, Principal of Douglas Memorial Higher Secondary School, Barrackpore for permitting us to do psychophysics experiments on the students in the school. I thank Ms. Sadhana Pal for recording audio for an experiment.

I am also thankful for my friends from Integrated Science Education and Research Center, Santiniketan. I am grateful to Dr. Ragadeepika Pucha for the helps she offered me while completing the thesis. It was refreshing to find my old friends Anindya Pattanayak and Meenal Chauhan in Kolkata. I also thank my Grandmother Ms. K. Saraswati for helping me during the time I spent on this thesis.

I thank all the non-academic staff in the institute for their helps. I thank Mr. Sujit Basak, Ms. Pamli Sengupta, and Ms. Tapashi Sreemani for their helps while working at Center for Soft Computing Research. The electrical maintenance unit had helped in installation of a floodlight used in an experimental setup.

I would like to thank all my friends in Center for Soft Computing Research. I had the pleasure of interacting with Barnini Bhattacharya, Swarup Chattopadhyay, Sanmoy Banerjee, Sukriti, Debarati Bhunia, Chandrani Chatterjee, Romi Banerjee, Soumen Datta, Suman Kundu, Srutiparna Neogi, Jayanta Kumar Pal, Roshni Mandal, Sandipa Roy, Ashish Bakshi, Rajdeep Das and many others during my time at the Center for Soft Computing Research.

List Of Publications

Journal Publications

- Keerthi S. Chandran et al. "Psychophysics may be the game-changer for deep neural networks (DNNs) to imitate the human vision". In: *Behavioral and Brain Sciences* 46 (2023). issn: 1469-1825. doi: http://doi.org/10.1017/S0140525X23001759
- Keerthi S. Chandran and Kuntal Ghosh. "A device for mass generation of psychophysics data to train and test models of flicker fusion". In: *Science Talks* 6 (May 2023), p. 100180. doi: http://doi.org/10.1016/j.sctalk.2023.100180
- Keerthi S. Chandran and Kuntal Ghosh. "A deep learning based cognitive model to probe the relation between psychophysics and electrophysiology of flicker stimulus". In: *Brain Informatics* 11.1 (July 2024). issn: 2198-4026. doi: http://doi.org/10.1186/s40708-024-00231-0

Conference Proceedings

- Keerthi S Chandran, Sreeja Pal, and Kuntal Ghosh. "Deep Learning as tool to distinguish words for sharp and round objects in natural languages". In: *2024 6th International Conference on Natural Language Processing (ICNLP)*. 2024, pp. 372–376. doi: http://doi.org/10.1109/ICNLP60986.2024.10692703
- Kuntal Ghosh and Keerthi S Chandran. "A low-cost device and technique for generating big data in visual psychophysics to train brain models". In: *43rd European Conference on Visual Perception (ECVP) 2021 Online*. Vol. 50. 1 suppl. SAGE Publications, Dec. 2021, pp. 85. doi: http://doi.org/10.1177/03010066211059887
- Keerthi S Chandran and Kuntal Ghosh. "Recurrent Convolutional Neural Networks trained by psychophysics data can predict EEG response to flicker". In: *43rd European Conference on Visual Perception (ECVP) 2021 Online*. Vol. 50. 1 suppl. SAGE Publications, Dec. 2021, pp. 132. doi: http://doi.org/10.1177/03010066211059887

- Keerthi S. Chandran and Kuntal Ghosh "Probing Temporal Filters of Vision via a Falsifiable Model of Flicker Fusion". In: *Brain Informatics. BI 2024. Lecture Notes in Computer Science*, vol 15542. Springer, Singapore, 2025. doi: https://doi.org/10.1007/978-981-96-3297-8_14

Works under revision and in press

- Keerthi S Chandran and Kuntal Ghosh. 'A Deep Learning based Brain Computational Model of Flicker Fusion', Under revision in Cognitive Computation and Systems, published by Wiley
- Keerthi S Chandran and Kuntal Ghosh. "A biologically inspired model for perception of flicker wheel illusion", post-conference (6th Intl. Conf. on Imaging, Vision and Pattern Recognition (IVPR) and 11th Intl. Conf on Informatics, Electronics & Vision (ICIEV)), October 26-29, 2023, London, UK) book chapter of an edited collection published by Taylor and Francis(In press)

Contents

Glossary

Critical Flicker Frequency (CFF)

Critical Flicker Frequency (CFF) is the number of photic pulses per second to eliminate the sensation of flicker. It is a threshold frequency.

Deep Learning

Deep Learning is a subset of machine learning that uses Deep Neural Networks to perform complex operations.

Deep Neural Network(DNN)

Deep Neural Network is an Artificial Neural Network with multiple layers between input and output.

Falsifiability

Falsifiability is standard for scientific theories introduced by the philosopher Karl Popper. A theory or hypothesis is falsifiable if it can be contradicted by an empirical test. Falsifiability makes a model or theory predictable or testable.

Flicker Fusion

When a flicker stimulus appears as steady to a subject, or the subject no longer perceives the flicker, the flicker is said to be fused or that flicker fusion has occurred.

Flicker Stimulus

A flicker stimulus is a visual stimulus with intermittent illumination. The flicker stimuli are kept static is space, but its intensity or luminance varies with time.

Illusory Motion

An illusory motion is a visual stimuli in which a static image appears moving to a human subject.

Psychophysics

Psychophysics is the study of the perception of physical stimuli. In psychophysics experiments, one or more parameters associated with the stimuli are changed, and the subject responses to the stimuli are recorded.

Pulse to Cycle Fraction

It is the duty cycle in a rectangular wave signal.

List of Figures

Chapter 1

Introduction

This work explores the use of deep learning as a tool to model psychological phenomena, which can then be tested via empirical experiments. The psychological phenomena investigated in this thesis fall in the domain of psychophysics. Psychophysics is the quantitative analysis of perceptual processes by studying the effect on a subject's experience or behavior of systematically varying the properties of a stimulus along one or more physical dimensions[1]. Deep Neural Network(DNN)s though originally designed for engineering problems, have been shown as a tool for modeling the biological brain[2]. In spite of the fact that DNNs have been described as the best models of biological vision, DNNs do not account for many findings from psychological research[3]. Bowers et al. has stated that theorists interested in developing biological plausible models of human vision should turn their attention to explaining psychological findings[3]. The current work proposes that psychophysics is an important tool to train and test Brain Computational models based on deep neural networks. This work models three psychological phenomena which are flicker fusion, flicker wheel illusion, and sound symbolism.

1.1 Deep Neural Networks and the Brain

Artificial neural networks were initially inspired by, and also designed to model, the biological nervous system[4]. But they found their use mostly in engineering goal oriented architectures[5]. Convolutional Neural Networks (CNNs) were similarly inspired by the works of Hubel and Wiesel[7]. The DNNs used in object classification were inspired by the functioning and structure of the brain[5]. DNNs have outperformed humans on benchmark datasets like ImageNet, in classification tasks[8]. Many researchers consider DNNs to be the best models of human vision[3]. But, there exists several disparities between the functioning of DNNs, and the functioning of human vision, especially as obtained from psychological research[3]. Yet, there is a strong bias among researchers not only to look for similarities between DNNs and the functioning of brain, but also to

ignore the differences between the two[9]. This thesis will explore new DNN architectures specifically to model psychological phenomena instead of relying upon the existing networks that often successful in accomplishing different engineering goals.

1.2 Psychological phenomena investigated in the thesis

This work has chosen three psychophysics phenomena for investigation. The three phenomena have been chosen because the experimental setup needs to have two binary responses. The stimulus, on the other hand, can be varied by manipulating the physical parameters associated with the stimuli. The phenomena are simple enough that they can even be perceived or experienced by non-human animals[10][11][12].

1.2.1 Flicker Fusion

The flicker stimulus is a visual stimulus with intermittent illumination[10]. The stimulus may appear as steady or flickering to a human subject depending on a number of circumstances[13]. When the flickering is no longer visible, the flicker is then said to have fused or flicker fusion have occurred. The psychophysics experiments on flicker fusion involve presenting a human subject with a visual stimulus of intermittent illumination, and the subject classifying the stimulus as either flickering or steady. When the flicker is no longer visible, the flicker is said to be fused or flicker fusion has occurred. This thesis will explore the phenomena of flicker fusion in multiple chapters.

1.2.2 Flicker from visual illusions

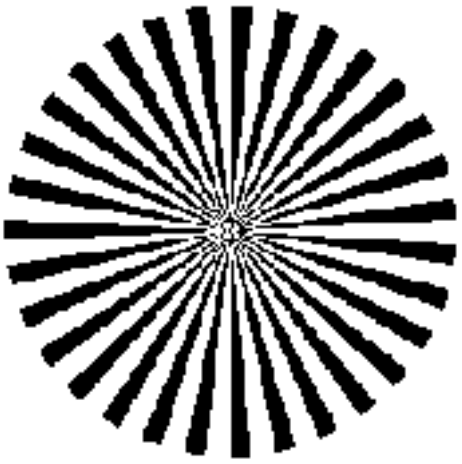


FIGURE 1.1: Flicker Wheel

Visual illusions [14] are like gaps in our visual mechanism through which we can spy, in a non-invasive manner, into the internal mechanisms of the brain. The illusions such as those which are geometric [15], or related to brightness [16],or motion [17] etc., where our perception fails to replicate the environmental reality, are the crucial tools to understand the visual brain's internal processing. Among the class of motion illusions, there exists those where a static image may be perceived as moving e.g. the Peripheral Drift Illusion (PDI) [18]. Flicker wheel illusion which is an illusion consists of a wheel with 30 to 40 spokes as seen in Figure 1.1[19]. The wheel appear to flicker vividly[19]. This thesis proposes a biologically plausible model for mechanism of illusory flicker perception for the flicker wheel illusion.

1.2.3 Sound Symbolism

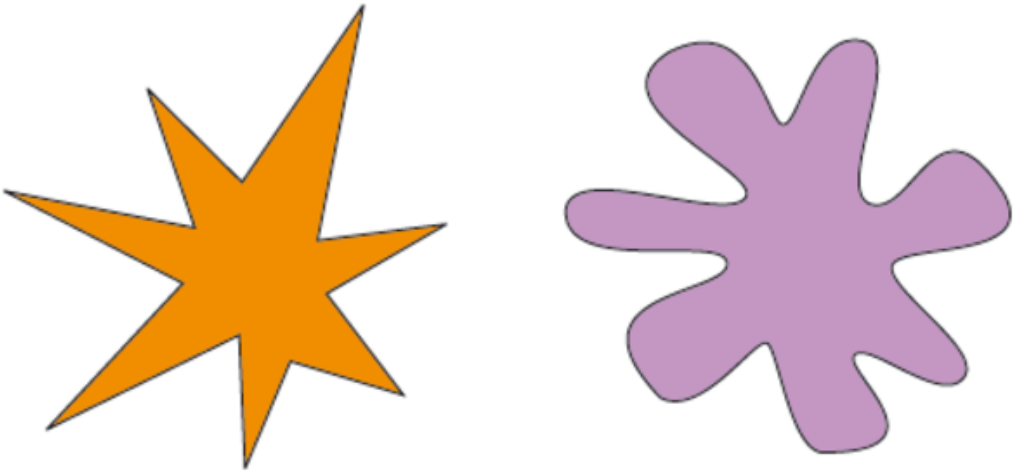


FIGURE 1.2: The images of figures used in Kiki Bouba Experiments. From Wikimedia Commons[20].

The term "Sound Symbolism" refers to the non-arbitrary links between random speech sounds and other meanings[21]. The sound symbolic phenomena exists in natural languages, as human subjects are able to guess the correct word from antonyms like big/small, and round/pointy from unknown languages better than chance[22]. Another sound symbolic effect is the Kiki Bouba phenomena in which, majority of human subjects match the nonce word Kiki to a sharp figure and and Bouba to a round figure for arbitrary shapes as in Figure 1.2[23].

1.3 Background and Related Works

1.3.1 Psychophysics

Psychophysics is the quantitative analysis of perceptual processes by studying the effect on a subject's experience or behavior of systematically varying the properties of a stimulus along one or more physical dimensions[1]. Psychophysics can yield insights into physiology, and mechanisms of human perception. For instance, Thomas Young was able to deduce the trichromatic nature of human vision from the fact that there were three principal colors, which could be mixed together to form other colors[24][25].Much of psychophysics is aimed at identifying the fundamental mechanisms shared by all humans[24]. It examines the most basic, fundamental aspects of human perception common to all normal humans, instead of establishing the more subtle aspects that may fluctuate within or amongst subjects[24]. This assumption is the reason that it is very common to see only a very small number of subjects being used in psychophysics studies[24]. The small number of subjects as low as 2, may surprise scientists from other fields[24]. There are of course psychophysics based studies that aim to quantify the magnitude of a particular parameter within a population, or to compare parameters between different populations[26]. Larger number of subjects will be appropriate for such works[26]. But, there are many other studies that demonstrate new effects without explicitly quantifying population parameters. These studies use a small number of subjects and are reported in the top-tier journals in the domain[26]. The present work, for most part, neither seeks to report a completely new phenomenon in psychophysics, nor does it aim to quantify subjective differences. The thesis, instead, mostly aims to use psychophysics data of known psychological phenomenon to train and test DNNs.

1.3.2 Artificial Neuron and Neural Networks

The artificial neuron and neural networks have been inspired from biology and neuroscience.

Santiago Ramón y Cajal recognized the nerve cell as the fundamental unit of nervous system. The word neuron was coined by Heinrich Wilhelm Waldeyer in 1891. The parts of the neuron consist of a cell body or a soma, dendrites that are projections from the soma and an axon with projections called axonites in it as can be seen in figure 1.3. The dendrites of sensory neurons are connected with sensory receptors. The axonites of motor neurons are connected to muscle cells. The nervous system consists of neurons that are interconnected with each other.

A mathematical model for biological neuron was proposed first by McCulloch and Pitts for understanding the brain[4]. An artificial neuron modelled after a biological neuron can be seen in figure 1.4. An artificial neuron takes in $\boldsymbol{n}$ inputs the same way a

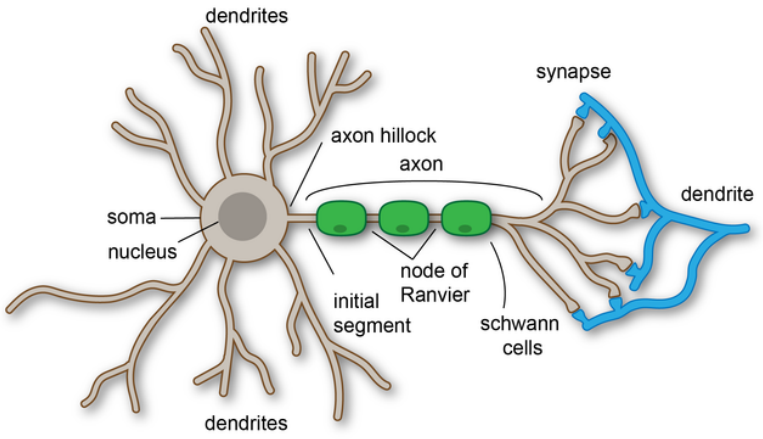


FIGURE 1.3: A biological neuron. From Wikimedia Commons[27]

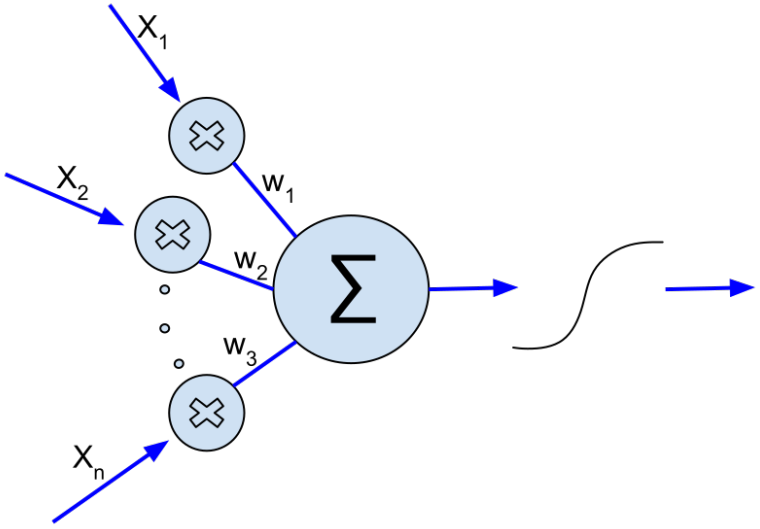


FIGURE 1.4: A McCulloch-Pitts Neuron. From Wikimedia Commons[28]

biological neuron takes in input through various dendrites. The artificial neuron takes in$\boldsymbol{n}$ inputs x_j, where j=1,2,3...n. The neuron computes the weighted sum of inputs as $\sum_{j=0}^{j=n} w_j x_j + b$, where b is a bias contant and w_j is weight associated with input x_j. The neuron gives an output by operating an activation function on the weighted sum. Some examples of activation functions are.

1. ***Sigmoid function***
   $f(x) = \frac{1}{1+e^{-x}}$
2. ***Rectified Linear Unit (relu)***
   $f(x) = 0$ if $x < 0$
   $f(x) = x$ if $x > 0$

1.3.3 Machine Learning and Deep Learning

Machine learning can be described as the ability of the machine to learn pattern from raw data by extracting patterns from the data [5, p. 2]. Machine learning algorithms are classified into supervised learning and unsupervised learning based on what they are allowed to experience during the learning process. In supervised learning, there is a target or label associated with each element in the dataset. Supervised learning learns from experiencing several instances of vectors in the dataset and the associated target and learning to predict the target from the vector. The term supervised learning originates from the fact that the target that needs to be predicted is provided as an instruction. Unsupervised learning is not in the scope of this thesis. Deep Learning is a machine learning technique using artificial neural networks incorporating deeper concepts from neuroscience like convolution, pooling, recurrence, attention etc. where, Artificial Neural Networks can be subjected to supervised learning using an a technique known as backpropagation.

1.3.4 Deep Neural Networks and Biological Cognition

Deep Neural Networks have been developed primarily to solve engineering problems and not for modeling biological cognition[2]. Yet features of biological vision do emerge on DNNs developed for engineering goals and trained on ImageNet. The lower layers of DNNs trained on classification with natural images develop Gabor like features[2]. It has been suggested that visual cortical representation also corresponds to Gabor filtering scheme[29]. The comparison of hidden layer representations of images from DNNs trained with object recognition task, with fMRI and MEG data of human subjects viewing the same images show relation between spatiotemporal cortical dynamics of human visual pathways and DNNs[30]. The outputs from the mid layers of deep neural networks, trained for object recognition tasks, are able to model spiking activities in area V1 of monkeys better than previous models[31]. DNNs trained on tasks like de-noising and de-blurring get deceived by brightness and color illusions the same way a human subject does[32]. DNN models trained on ImageNet show human like responses to some visual illusions and in particular for Müller-Lyer and Hermann grid illusions[33].

1.3.5 Methodology of the thesis

Many of the computational techniques that aid to model the brain from behavioral data were not available a few decades ago[34]. In 1988, Zipser and Andersen used backpropagation as a tool to model the computational activity of the brain for the first time[35]. Backpropagation is a technique used to train artificial neural networks (ANNs) from just the sets of inputs and outputs. The ANN will find the appropriate algorithm through

backpropagation and knowledge of algorithm is not needed to train the network[35]. This work proposes that DNNs can be trained to mimic human behavior obtained using psychophysics experiments conducted under laboratory conditions. Such models could be put to empirical tests from psychophysics data obtained with psychophysics experiments.

1.4 Thesis Structure

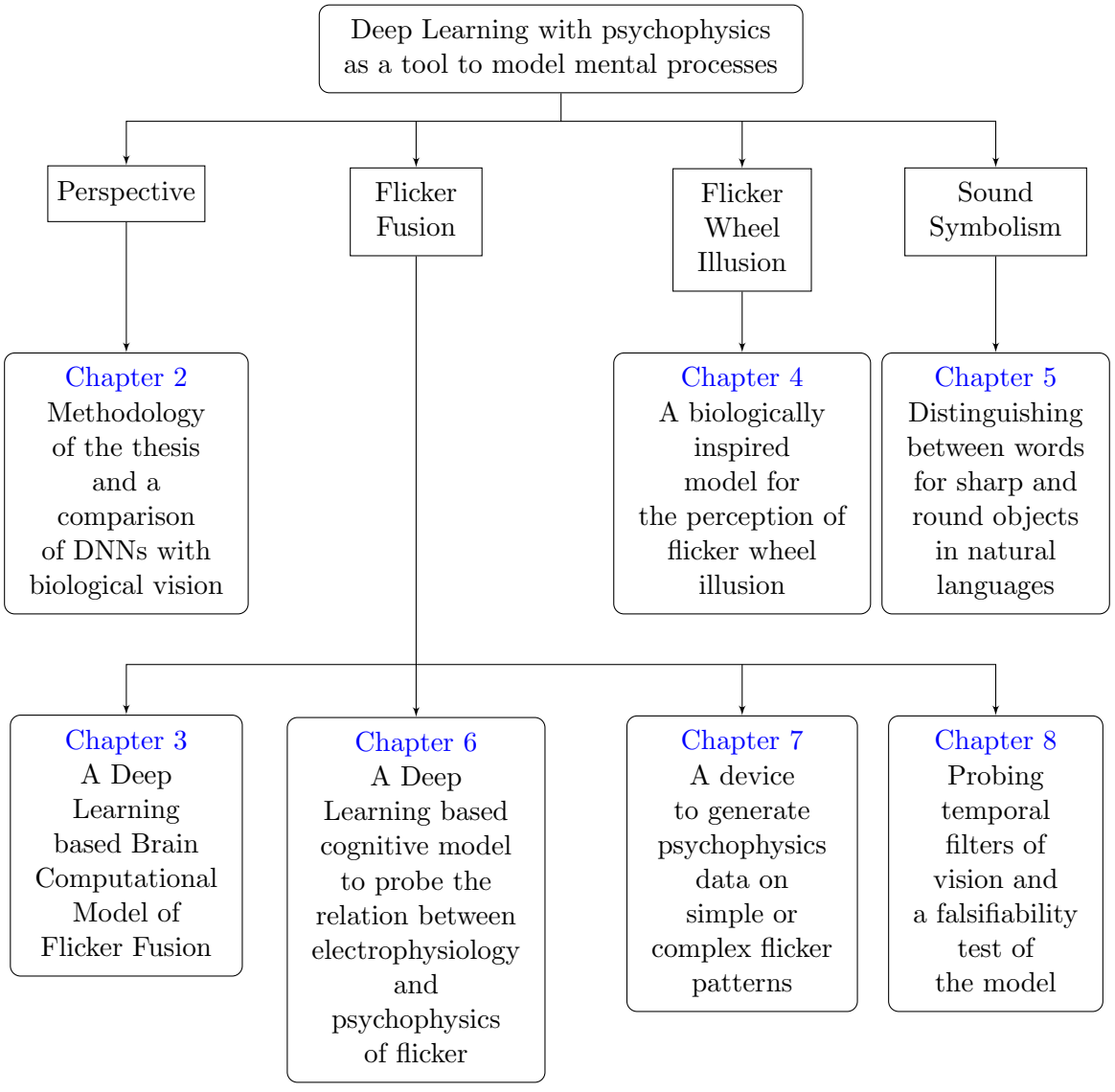


FIGURE 1.5: Thesis Structure

The work done in the thesis are both experimental and theoretical in nature. The theoretical sections contains methodology of the thesis, differences between human and animal vision that might lead to differences between human and DNN vision, and computational models for flicker fusion, flicker wheel illusion and a sound symbolic phenomena.

The experimental work consists of psychophsyics experiments on flicker fusion, flicker wheel illusion, sound symbolism, and construction of an electronic device to generate data on flicker fusion.

1.4.1 Chapter 2

Previous works have compared the internal representations of engineering goal oriented neural networks, with invasive or non-invasive brain signals from human or monkey brains exposed to the same stimulus and found correlation between them. This thesis differs in methodology from those works, by building new DNNs to fit psychophysics data. Such DNNs can make falsifiable predictions. While DNNs are inspired by neuroscience, there exists differences between vision of humans and other animals. This work argues that, many of the reported incongruencies between human psychophysics and DNNs arise because the functioning of DNNs maybe more similar to brains of non-human animals.

1.4.2 Chapter 3

This chapter proposes a DNN based model of flicker fusion that can be trained and tested on psychophysics data of a human subject. The model is based on Convolutional Recurrent Neural Network(CRNN). The chapter shows that the model pretrained on psychophysics data of a subject can be used to train psychophysics data of another subject faster than the original subject, showing that the model captures mechanisms of flicker perception, common across human subjects. The model was trained with simple rectangular wave flicker stimuli and was tested on other types of flicker stimuli that it was not trained with, and subjective comparisons were made with previously published psychophysics data. The work also shows that internal representations of the CRNN can be probed for correlations with brain signals like the EEG associated with flicker stimulus.

1.4.3 Chapter 4

This chapter does a study on flicker wheel illusion, which is a motion illusion. Flicker wheel illusion is a visual illusion of a wheel with spokes which appear to flicker. The perception of flicker wheel illusion is affected by physical parameters of the stimulus like number of spokes and contrast. This work generates a novel dataset of multiple variants of flicker wheel by varying physical parameters associated with the stimulus. The stimuli are then presented on a computer screen and is labeled by a human subject into two classes on basis of whether the subject perceives flicker in the stimulus or not. DNNs are trained to classify the images the same way a human subject does. Many existing

CNN Networks are shown to be able to classify the data. The work also shows that VGG-16 and VGG-19 are unable to train on the generated data. The work proposes a bio-inspired Deep Neural Network that takes into account the microsaccades in the human eye.

1.4.4 Chapter 5

This work attempts deep learning in the well attested psychological phenomena of sound symbolism. Sound symbolic effects, while having been studied by psychologists, have never been modeled before using machine learning. This works performs both human psychophysics and machine learning experiments on words for round and sharp objects in natural language. We go onto show that both humans and machine do better than chance in classifying the words for round and sharp objects in natural languages.

1.4.5 Chapter 6

This work makes comparison between the outputs of intermediate layers of a CRNN trained with psychophysics data of flicker fusion, and the reported features of human EEG response to flicker. We claim that many of the reported features of electrophysiology of flicker stimulus, including the presence of fundamental and harmonics of the stimulus, can be explained as the result of a temporal convolution operation on the flicker stimulus. We further show that the convolution layer output of a CRNN trained with psychophysics data is more responsive to specific frequencies as in human EEG response to flicker, and the convolution layer of a trained CRNN can give a nearly sinusoidal output for 10 hertz flicker stimulus as reported for some human subjects.

1.4.6 Chapter 7

This work describes the construction of a device that can be used to obtain data to train and test DNN models for flicker fusion. The work describe an instrument that can generate simple as well as complex flicker patterns and record the subject classification via button presses. The experimental system also includes a mechanism to prevent dark adaptation by automatic lighting of a high intensity lamp after the button press for a certain time period. The device is controlled by an Arduino. The work also shows that there does not exist a sharp cut-off between regions of flicker and fusion in terms of frequency and outliers can exist outside threshold regions.

1.4.7 Chapter 8

This works describes the training of psychophysics data obtained by device on the proposed computational model for flicker fusion. The trained DNN is tested with psychophyics data. The training yielded symmetric filters, as often found in biological visual systems. The predictions made by the CRNN model on complex flicker patterns, tested through psychophysics experiments with the device, demonstrate that the model is falsifiable with scopes of further improvement through future research.

1.4.8 Conclusion

The last chapters gathers in conclusions drawn from all the works done in the thesis, and points to the possible future directions of research emerging out of these.

Chapter 2

Methodology of the thesis and a comparison of DNNs with biological vision

Publications

- Keerthi S. Chandran et al. "Psychophysics may be the game-changer for deep neural networks (DNNs) to imitate the human vision". In: *Behavioral and Brain Sciences* 46 (2023). issn: 1469-1825. doi: http://doi.org/10.1017/S0140525X23001759

Chapter summary:

Previous works have compared the internal representations of engineering goal oriented neural networks, and invasive or non invasive brain signals from human or monkey brains exposed to the same stimulus and found correlation between them. This thesis differs in methodology from those works, by building new DNNs to fit psychophysics data. Such DNNs can make falsifiable predictions. While DNNs are inspired by neuroscience, there exists differences between vision of humans and other animals. This chapter argues that, many of the reported incongruencies between human psychophysics and DNNs arise because current existing DNNs might be working similar to vision of non-human animals.

2.1 Introduction

Deep Neural Networks have been inspired by neuroscience research[5]. The neuroscience research that had inspired DNNs was invasive brain research conducted on non-human animals like cats. While DNNs were mostly employed to solve engineering problems, Deep Neural Networks were also put forward as the best models for functional architecture of the brain[2]. The correlations between the outputs of a layer to DNN for various inputs and a multidimensional brain signal for the same set of stimulus can be probed by a technique called Representational Similarity Analysis[36]. The dimension of brain signals can be different voxels for fMRI, different sensors for MEG or EEG or different set of internal electrodes in the brain[30, 37, 38, 39]. Although representational similarity analysis has yielded correlations between brain signals and DNN representations, there is much less correlation between reported DNN behaviors and psychological findings[3]. Instead, this thesis proposes that psychophysics combined with deep learning is a tool to model functional activity in the brain. This chapter also offers the perspective that human vision differs from animal vision and DNNs could be closer to animal vision than human vision.

2.2 Aim and methodology of the thesis

2.2.1 Aim of the thesis

The advent of computing machinery and programming languages made it possible to build experimental tests for ideas of intelligence in last half on twentieth century[40]. The term artificial intelligence was coined by John McCarthy and others in 1956 CE[41, p 295]. Alan Turing's paper 'Computing Machinery and Intelligence' in 1950 was a seminal paper in the history of Artificial Intelligence. In this Turning addressed the question 'Can machines think?'. Turing replaces the question with another, that is, whether a digital computer can beat a human in an imitation game[42]. To beat the imitation game a computer communicating with a human using typed text should be able to convince a human that its another human being and not a machine[42]. The advent of general purpose computers was followed by the notion that machines would be more useful if they could perform tasks for which they were given no precise methods[43]. This method was called machine learning[44]. This thesis will simulate some mental processes via machine learning.

In keeping with what Turing proposed for the imitation game [42], a good brain-computational model[34] would not be the one that performs a particular task with equal or greater accuracy than a human being, but rather the one which would be indistinguishable from a human being vis-à-vis input and output. For instance computers

excel humans in playing chess but the patterns of moves made by computers in chess is different from that of humans[45].

2.2.2 Psychophysics as a tool

Psychophysics, interestingly, is also about input and output with the brain as black-box in between[24]. Psychophysics is "the analysis of perceptual processes by studying the effect on a subject's experience or behavior of systematically varying the properties of a stimulus along one or more physical dimensions"[1]. The psychophysics stimulus for vision can be an image or video, and DNN, an information-processing system, may model the subject's response to the stimulus using supervised learning. David Marr had proposed that an information processing system should be understood at three levels: computational, algorithmic, and implementation. The psychophysics task describes the computational level problem, a DNN that performs the same task in silica would represent the algorithmic level, and the electrophysiological or fMRI data obtained during the task will be a by-product of the implementation of the algorithm in the biological brain. If the DNN is considered for an equivalent mapping between input and output as in a psychophysics experiment, then the inputs can be represented by a tensor, whether it is an image, video, sound signal, or a spatially invariant visual stimulus like the flicker; the output would also have a numerical representation which, in case of psychophysics experiments, could be some classification, perceived brightness, color, shape, size, motion, intensity at a particular location in the input signal, or a comparison between two of those perceived sensations at different locations of the stimulus, separated by space or time or both. The algorithm used to transform the stimulus input to output will not be evident from psychophysics experiments, but DNNs can construct that algorithm without its exact knowledge for the programmer.

While this thesis will train DNNs via backpropagation, it does not imply that any equivalent computational mechanism in human visual system was similarly trained by backpropagation. The biological brains are instead the product of evolution by natural selection. The psychophysics phenomena investigated in this thesis like flicker fusion experiments are conducted in controlled setting in the laboratory, and many of these stimuli would not have been naturally present in any evolutionary history. The thesis shall instead use psychophysics experiments to gather data to train models while the biological mechanism being modeled came into being by an unrelated mechanism which is evolution via natural selection.

2.2.3 Methodology of the thesis

The dataset can be prepared by manipulating physical parameters associated with the stimulus and getting the subject response for each of the stimuli. There can be some

subjective differences between the psychophysics data of human subjects for the same stimuli[24]. So, it will be a better strategy to train and test a DNN on the psychophysics data of the same subject. Kubota et al. Kubota, Hiyama, and Inami have shown that it is possible to make comparisons between human perception obtained by psychophysics experiments, and the output of DNNs[46]. DNNs may also be tested on a stimulus, completely different from the one it was trained on, if its output layer is of similar representation to that of the new stimulus, as has been proposed with flicker fusion phenomena in Chapter 3 . The intermediate outputs of a DNN can be compared with the brain electrophysiological signals as done by Zipser and Andersen , and as has been proposed in Chapter 6 flicker stimulus[35]. We argue that more testable models can be constructed by training on less computationally intensive tasks than tasks like object classification into thousands of classes. For instance, a convolutional neural network (CNN) trained for low-level visual tasks gets deceived by brightness and color illusions [32]. DNNs have also been put forth to solve tasks used in experimental psychology like Raven's progressive matrices[47]. New network models, different from the engineering goal-oriented image classification DNNs, could be constructed for the purpose as was previously done for finding head-centered coordinates of external objects by monkey brain by Zipser and Andersen[35]. It could be easier to make correlations between outputs of intermediate layers of a neural network with fewer neurons and layers with brain signals than complex networks.

2.3 Comparison and differences between human, animal and DNN vision

2.3.1 Global vs local shape perception

DNNs rely more on texture for image classification, and even when trained to classify based on shapes, rely more on local shapes than global shapes[3]. Many experiments have shown that global shapes precede local shapes in human perception and have termed it Global Precedence Effect(GPE)[48]. Bowers et al. have claimed that this is a discrepancy in claiming DNNs and human vision are similar[3]. We suggest that this discrepancy arise because the mechanisms of DNNs could be more similar to animals other than humans. Humans have an advantage in global shape perception, while chimpanzees and baboons hold advantage in local shape perception under some conditions[48]. Humans are more capable in processing global shapes than chimpanzees[48, 49]. DNNs may rely more on local shapes for classifications, since they could be more optimized for processing local shapes like other animals.

2.3.2 Invariances

Bowers et al. has noted that human object recognition is invariant to transformations like rotation and scaling[3]. While Bowers et al. has mentioned that DNNs can be trained to support invariances, Bowers et al. listed it as a phenomena that requires more study[3]. In pigeons, object recognition after training with a single object image is dependent on a variety of properties irrelevant for object identification[50]. Object recognition in pigeons after trained with a single object view is affected by rotation, and variation in size[50].

2.3.3 Binocular vision

Bowers et al. mentions that DNNs trained on ImageNet do not encode three-dimensional (3D) features of objects or their depth as opposed to human vision. The above mentioned DNNs are trained with datasets prepared from cameras with monocular vision. But the mammalian brain gets information from the two eyes and it is known that human subjects with one eye are not so efficient with depth perception[51]. Cats show better depth discrimination at binocular vision than monocular vision[52]. Cats which have been monoculalry deprived for first thirty five days of life never recover the binocular superiority[52]. Robots with stereo cameras making use of DNNs are able to do tasks like calculating position of detected fruit from stereo cameras[53]. Stereo vision can enable autonomous driving vehicles to do tasks like object detection, 3D information acquisition, and depth perception[54]. The mammalian brain had input from two eyes throughout the course of its evolutionary history. So training DNNs using stereo camera data might be needed to develop the equivalents of many circuits in the brain.

2.4 Conclusion

To conclude, psychophysics with DNNs could be used to construct many of the smaller agents that compose the human mind as proposed by Minsky[55]. Vision agents that compose the mind need to be likewise constructed via DNNs, which may be associated with fundamental activities like brightness perception, motion detection, depth perception, or even less intelligent activities than that, in the parallel visual pathways. Neural networks for more complex tasks can be built with a combination of smaller DNNs using shared layers, or by using output from some layers of a DNN as input for layers of another DNN. Moreover, DNNs have been inspired by neuroscience vision from other animals and many of the neural network architectures may not be optimal models for human vision.

Chapter 3

A Deep Learning based Brain Computational Model of Flicker Fusion

Chapter summary:

This chapter proposes a deep learning based Brain Computational Model(BCM) of flicker fusion that can be trained and tested on psychophysics data of a human subject. Inspired by Fechener's law, the model performs logarithmic transformation of the stimulus before it is fed into a Convolutional Recurrent Neural Network(CRNN), that classifies a flicker stimulus as flickering or steady like a human subject. The model pretrained on psychophysics data of a subject can be used to train psychophysics data of another subject faster than the original subject, showing that the model captures mechanisms of flicker perception, common across human subjects. The model was trained with simple rectangular wave flicker stimuli and tested on other types of flicker stimuli that it was not trained with, and subjective comparisons were made with previously published psychophysics data. This chapter also shows that internal representations of the CRNN can be probed for correlations with brain signals like the EEG associated with flicker stimulus. The presence of fundamental and harmonics of the flicker stimulus in its human EEG response can be explained by the BCM as a convolution operation on the stimulus.

3.1 Introduction

Deep Neural Networks (DNN), mostly used for engineering problems, have been put forward as a tool to model the human visual system[2]. Models so constructed with DNNs can be tested for falsifiable predictions on human behavioral data[56]. Task performing models that can be tested with behavioral data are central to psychophysics and cognitive science[34]. A brain-computational model(BCM) is one that mimics the brain information processing of some task using a certain level of abstraction[34]. Understanding cognition in the brain would require constructing computational models that can perform cognitive tasks and can be tested with behavioral and brain experiments[34]. This work develops a DNN model that could be trained and tested on psychophysics data of flicker fusion.

3.1.1 Flicker Stimulus and Critical Flicker Frequency(CFF)

The flicker stimulus is a visual stimulus with intermittent illumination[13]. The stimulus may appear as steady or flickering to a human subject depending on a number of circumstances[13]. When the flickering is no longer visible, the flicker is then said to have fused or flicker fusion has occurred. The psychophysics experiments on flicker fusion involve presenting a human subject with a visual stimulus of intermittent illumination, and the subject classifying the stimulus as either flickering or steady. For a human subject, the physical parameters associated with the stimulus determine whether the subject classifies the stimulus as flickering or steady[57, 58]. In addition to that, there are subject-wise differences in the perception of flicker[58]. A numerical parameter used in psychophysics measurements of flicker fusion is CFF. CFF is defined as the Critical threshold of flicker fusion or the point where some change in physical stimulus causes the flicker to appear as a steady sensation or vice versa[13]. CFF is also alternately defined as Critical Flicker Frequency or the number of photic pulses per second to eliminate the sensation of flicker[59]. This work will use the second definition of CFF.

3.1.2 Physical parameters associated with flicker stimulus

The physical parameters associated with the stimulus that determines whether the stimulus appears flickering or steady. The parameters include size and shape of the stimulus, as well whether the stimulus falls in foveal or parafoveal region of the eye[57]. The temporal pattern of the stimulus also determines whether the stimulus appears as flickering or fused. The flicker stimulus can be in a sinusoidal pattern as in Figure 3.2a or in the form of a rectangular waveform as in Figure 3.1[60]. The flicker stimulus of rectangular waveforms consists of a pulse of light that stays on for some time followed by a period of darkness. For a flicker stimulus of sinusoidal waveform the intensity $I(t)$ at timepoint

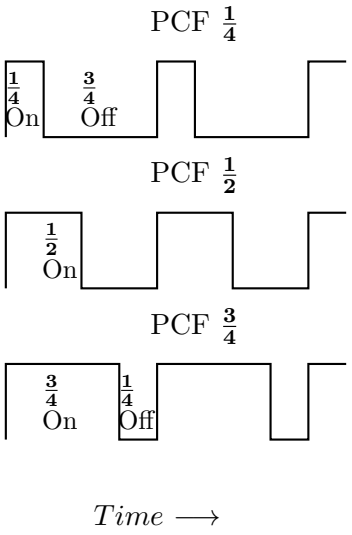
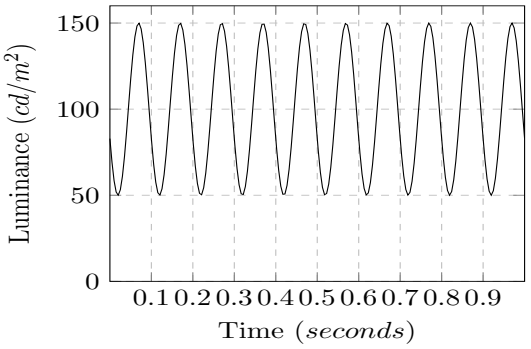


FIGURE 3.1: A diagrammatic representation of three photic pulses as variation of intensity with time. The photic pulses have the same frequency or time periods but three different PCFs which is the same as the duty cycle of a rectangualar wave signal.

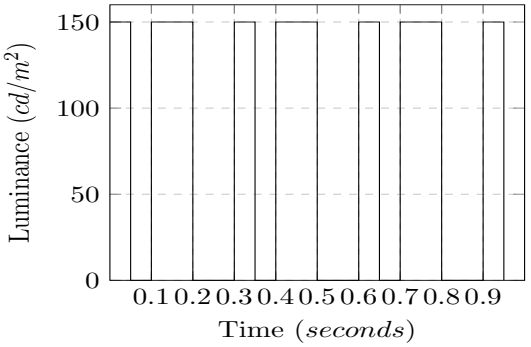
t is given by $I(t) = I(1 + \boldsymbol{m}sin(\omega t))$ where I and $\boldsymbol{m}$ are constants and ω is $2\pi f$, where f is the frequency. Another physical parameter for rectangular waves that determine whether a stimulus appears as flickering or steady is the Pulse to Cycle Fraction(PCF) which is shown in Figure 3.1[59]. Pulse to cycle is the fraction of time duration of the pulse to that of the total cycle[59]. It is same as the duty cycle of a rectangular wave signal. A rectangular wave with PCF $\frac{1}{2}$ is called a square wave signal. Flicker signals with rectangular waveforms can be expressed by three parameters: Intensity, frequency and Pulse to Cycle Fraction. As in many other works, we have used the terms intensity and luminance synonymously.

3.1.3 Complex Flicker waveforms

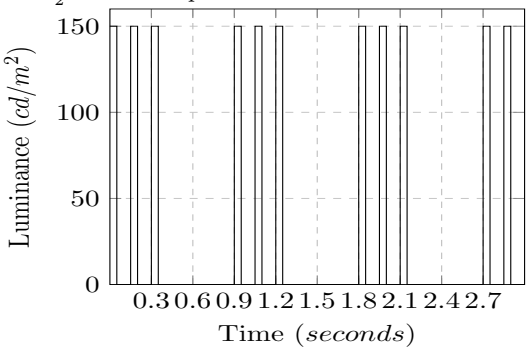
In addition to simple rectangular waveforms and sinusoidal waveforms described above, the psychophysics of flicker fusion have been examined by more complex stimuli. These include waveforms in which alternating cycles have different time duration as in Figure 3.2b [61]. Experiments have also been done with a finite number of pulse trains with a gap between them as in Figure 3.2c. In the above experiment, the finite pulse trains were fused instead of the whole pulse train[62]. The device described in Chapter 7, can generate more complex signals with pulses of different colors and intensities in the same train as in Figure 7.9 and gather the corresponding psychophysics data. The stimulus in flicker fusion experiments can be represented as a video of the target surface. But if the size and shape of stimulus are invariant in the experiment, with just luminance changing with time, the stimulus can be represented as a one dimensional time series of intensity. The temporal array that serves as the representation of the stimulus can be fed into a deep recurrent neural network that could give a binary classification of either fused or



(A) A sinusoidal flicker stimulus where intensity It at time t is $I(t) = I(1 + \boldsymbol{m}sin(\omega t))$.



(B) Stimulus similar to those used by Forsyth and Brown[61]. The alternate pulses both with PCFs $\frac{1}{2}$ have time periods 0.1 and 0.2 milliseconds.



(C) Stimulus similar to those used by Nelson, Bartley, and Harper[62]. Three flashes of PCF $\frac{1}{3}$ are followed by a gap with time duration of three cycles.

FIGURE 3.2: Flicker stimuli different from simple rectangular wave stimulus

flickering. There can be a wide range of input representations consisting of rectangular, sinusoidal or complex waveforms but with just a binary output for the network. This wide range of input stimulus patterns, with just two outputs, offer a chance to build a testable and falsifiable model of flicker fusion.

3.1.4 Artificial Neural Network (ANN) as a tool to model visual perception

Backpropagation is a technique used to train artificial neural networks (ANNs) from just the sets of inputs and outputs. The ANN will find the appropriate algorithm through

backpropagation and knowledge of the algorithm is not needed to train the network[35]. The present work proposes that a model for flicker fusion could be similarly trained on psychophysics data on rectangular waveforms as in Figure 3.1 and could then be tested with psychophysics data of other waveforms as in figures 3.2 and 7.9. Flicker stimulus is a sequential data. This work uses recurrent neural networks which are a family of networks for processing sequential data [5] to model flicker fusion.

3.1.5 Relation of brain signals with computational models of cognition

The importance of such modeling is that, in addition to using DNNs to build models that could predict behavioral data, it would be possible to use such trained deep neural networks to predict the corresponding electrophysiological and fMRI signals associated with some particular behavioral activity. In monkeys, the position of objects in the external world are represented in the brain as head centered coordinates[35]. The monkey brain has to calculate these coordinates from the position of the object image in the retina and the slope of the eyes. The intermediate neurons of an artificial neural network trained to derive such coordinates from a representation of position of object image in the retina and the slope of the eye, generates similar representations as those obtained from electrophysiological recordings from parietal neurons of monkeys[35]. The present work will describe a DNN whose intermediate layers will yield time series data. While the final output layer of the DNN proposed in this work, gives an equivalent for behavioral data as subject's classification of the stimulus, correlations can be sought between the outputs of intermediate layers of the DNN with EEG and fMRI response to flicker stimulus.

3.1.6 Contributions

The main contribution of this work is to develop a biologically plausible Brain Computational Model that could be trained with psychophysics data of flicker fusion. Since there is a lack of such data acquired with the intention of training deep neural networks, we have trained our model by making some assumptions on some previously published data. We have made subjective comparisons of the outputs of the model to a different set of flicker stimuli and their previously published psychophysics results. The methodology followed in this chapter is shown in Figure 3.3. Though the comparisons did not yield exact similarities with the published psychophysics results, we have succeeded in developing a model that could fit the psychophysics data on flicker fusion as well as simulate a wide range of experiments in the domain. The model based on a Convolutional Recurrent Neural Network(CRNN) takes inspiration from two psychophysics laws, the Fechner's and the Ferry-Porter laws, to subject the input stimulus to a logarithmic transformation.

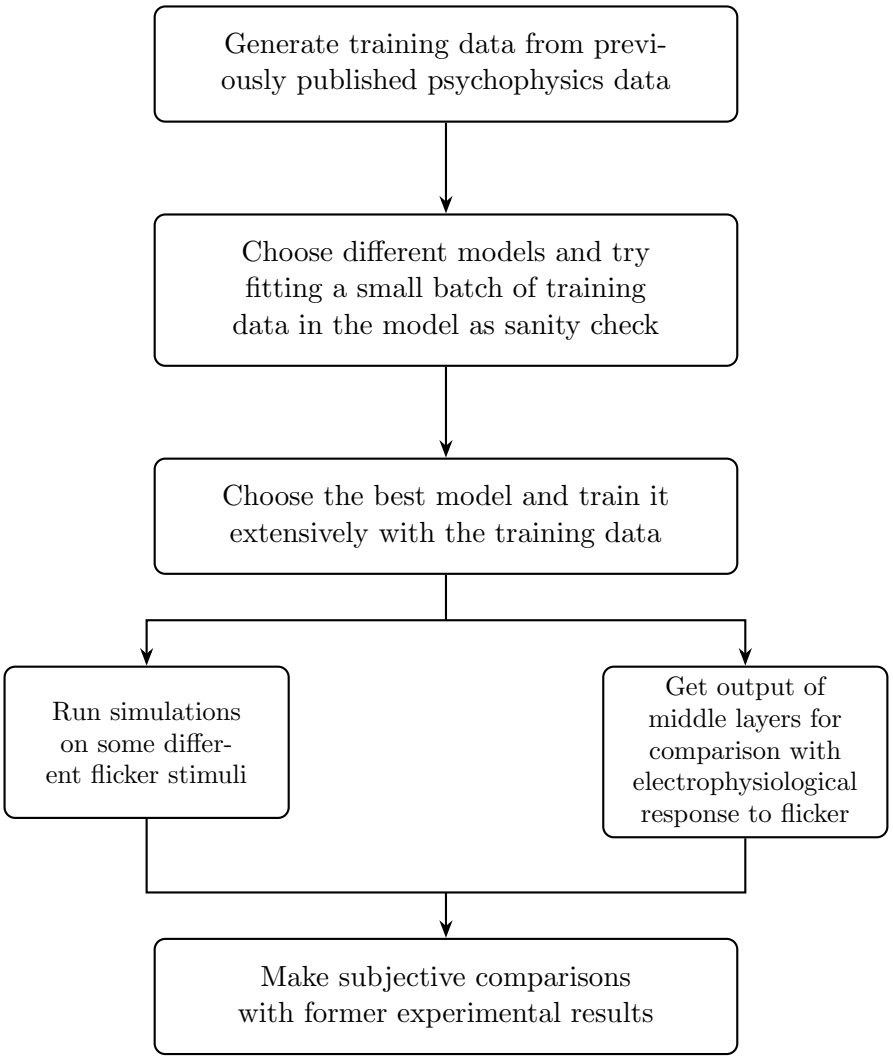


FIGURE 3.3: Block diagram of methodology

3.2 Related Works

3.2.1 Neuroscience and psychophysics of flicker fusion

The human perception of flicker is associated with activity of higher cortical regions[10]. It has also been suggested that the same cells underlie the perceptions of both flicker and motion directionality[63]. CFF could be reliably measured with high repeatability for all three psychophysics methods[64]. The measurement of CFF is also resistant to learning effects[10].

3.2.2 Brain signals associated with flicker stimulus

Another domain of study of flicker perception are the brain signals of the cortical activity induced by the stimulus. Flicker stimulus is known to evoke frequencies in human EEG

which are of the same frequency as the stimulus for frequencies up to 90 hertz[65]. The oscillations so invoked in the EEG also show oscillations around 10, 20, 40 and 80 hertz[65]. The oscillations are invoked in the EEG even when the flicker is not perceived[65]. The stimulus evoked the fundamental frequency, harmonics as well as as the first subharmonic of the stimulus frequency, in EEG response to flicker[65]. Studies using magnetoencephalography (MEG) on subjects exposed to sinusoidal flicker stimulus have revealed that the oscillations in the human visual cortex evoked by the flicker stimulus and gamma oscillations in the brain are evoked in different regions of the human visual cortex and that the flicker stimulus-induced oscillations do not entrain endogenous gamma oscillations[66]. For a flicker stimulus with frequency near the CFF, fMRI data shows that greater activity in frontal and parietal cortex on perception of flicker and greater activity in the occipital extrastriate cortex when the flicker appears fused[67].

3.2.3 Chromatic flicker stimulus

Flicker fusion experiments have also been done with chromatic flickers which are flicker stimuli with alternating colors in the train. When chromatic flickers occur above a threshold, the flicker is no longer perceivable to subjects, but it has been experimentally observed that unperceivable chromatic flicker evokes oscillations in recording from color opponent cells in V1 layers of monkeys[68]. It has also been demonstrated by fMRI that a chromatic flicker, not perceivable to humans, evokes higher responses in many visual cortical areas compared to a steady control stimulus[69].

3.2.4 Computational models of flicker perception

The development of a computational model for flicker perception could help in understanding and modeling of other related psychophysics phenomena as well as other related information processing in the brain. Flicker detection has also been modeled much earlier with analog neural network with resistances and capacitors[70]. Levinson had opined that mind is not a radio and avoided implication of electrical components like resistors and capacitors in the visual system while proposing an n stage filter model for flicker fusion[71]. Many other models have also been put forward in the past, but none of them have seemed durable[60]. Amari, Lieblich, and Karshmer proposed a neural model and opined that the computational mechanisms of flicker perception is also related with another psychophysics phenomena that is metacontrast[72]. More recent works have shown that artificial neural networks can model ElectroRetinogram (ERG) response to flicker stimulus using feedforward networks[73]. The harmonics, first subharmonic and fundamental of the flicker stimulus can be generated by feeding the stimulus representation to a neural mass model consisting of inhibitory and excitatory neurons[74].

3.2.5 Clinical Implications of Flicker Fusion

There are significant differences in CFFs amongst patients with middle and early stage Alzheimer's disease as well as healthy control groups[75]. The CFF is positively correlated with mini mental state examination scores[75]. CFF may play an important role in diagnosis of Alzheimer's disease. The present work offers a computational model that can also prove useful for tracking the progression of disease.

3.3 Materials and Methods

The generation of dataset for training and testing are done using python libraries like scipy and numpy. Keras was used to train and test the models.

3.3.1 Materials

TABLE 3.1: The CFF values for a subject *Coo* for 28 combinations of PCFs and intensities as reported by Bartley and Nelson [76]

PCF Intensity	$\frac{1}{30}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{7}{8}$	$\frac{29}{30}$
1164 cd/ft^2	44.3	31.2	49.0	49.5	39.9	36.0	29.5
116.4 cd/ft^2	35.9	34.2	44.8	44.0	39.8	34.5	30.0
7 cd/ft^2	28.9	33.0	39.3	40.0	33.3	34.4	27.7
1.28 cd/ft^2	21.0	30.7	33.0	34.0	25.5	27.8	20.9

TABLE 3.2: The Spearman Correlation of CFFs between the subjects reported in [76]

subject	*Coo*	*Ran*	*DeB*	*Bal*	*Kra*	*Nel*	*Bar*
Coo	1.0	0.906	0.515	0.688	0.888	0.798	0.786
Ran	0.906	1.0	0.581	0.579	0.861	0.774	0.691
DeB	0.515	0.581	1.0	0.613	0.606	0.612	0.598
Bal	0.688	0.579	0.613	1.0	0.67	0.636	0.769
Kra	0.888	0.861	0.606	0.67	1.0	0.693	0.711
Nel	0.798	0.774	0.612	0.636	0.693	1.0	0.872
Bar	0.786	0.691	0.598	0.769	0.711	0.872	1.0

There is not much available data on CFF with quantifiable values of associated with both luminance and PCFs, to be used to train deep learning models. The psychophysics data previously published by Bartley and Nelson have been used to train the DNN[76]. The target stimulus was a truncated wedge-shaped area which was maximum $\frac{9}{16}$ inches high and $\frac{11}{16}$ inches wide. The subjects viewed the target 30 inches away from the target and the target subtended a visual angle of $1°4'$. The data contained the critical flicker frequency for seven different PCFs and four intensities for seven subjects. There were 28 CFFs recorded for a subject. The CFF values for a subject *Coo* published in the report

can be seen in Table 3.1. In order to select the subject whose data was to be chosen for deep learning, the Spearman correlation between all 28 CFFs for the subjects were calculated , which can be seen in Table 3.2. Then the correlation values were added up. The CFF values for the subject *Coo* gave the maximum sum in this manner, and chosen for training the neural network.

3.3.2 Methods

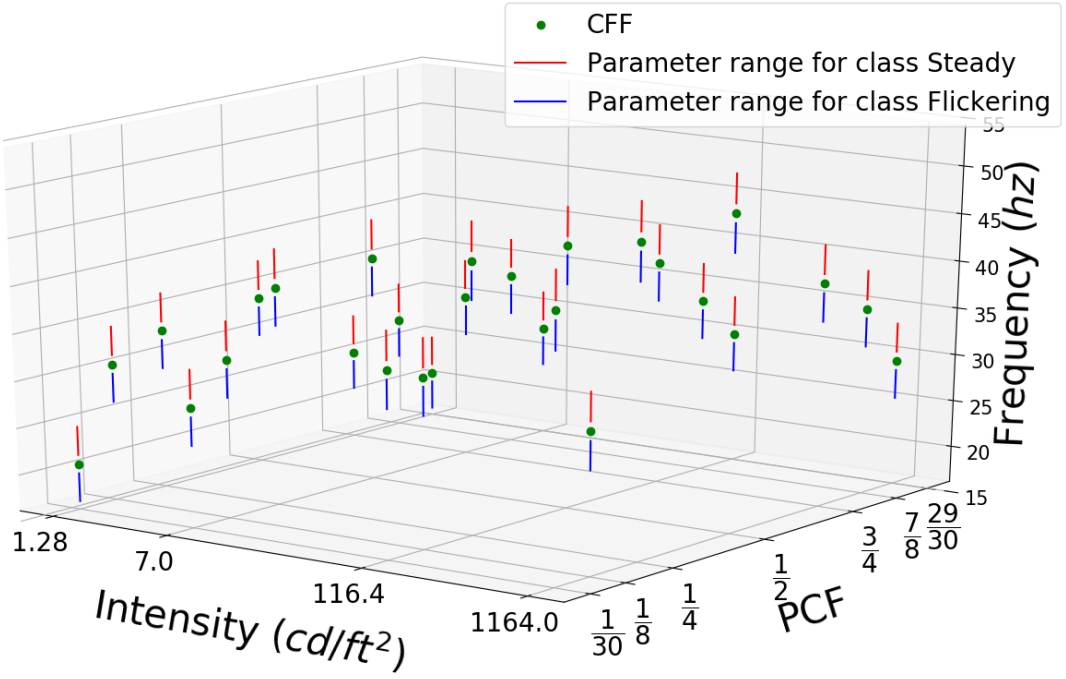


FIGURE 3.4: The training ranges used for generating one dimensional waves with parameters as intensity, frequency and PCF

Since CFF is a threshold frequency, we have made the assumption that for a given luminance and PCF, any wave with the frequency between CFF+1 and CFF+4 will appear as steady as the photic pulses with frequency above CFF appear as steady. Similarly for a given luminance and PCF, the waves with frequency between CFF-1 and CFF-4 were assumed as flickering. The set of parameter combinations used to train each class is seen in Figure 3.4. A photic pulse can be represented by a one dimensional array. One dimensional representation of photic pulses can be constructed from the three parameters, viz. intensity, PCF and frequency. This work has generated photic pulses with a duration of ten seconds for training. The array was sampled at 0.001 seconds for Experiment 1 and 0.0001 seconds for Experiments 2 and 3. So each array had 10000 elements in Experiment 1 and 100000 elements in Experiment 2 and 3. Now for the purpose of training and testing, more than one time series representation could be constructed with a given PCF, intensity and frequency.

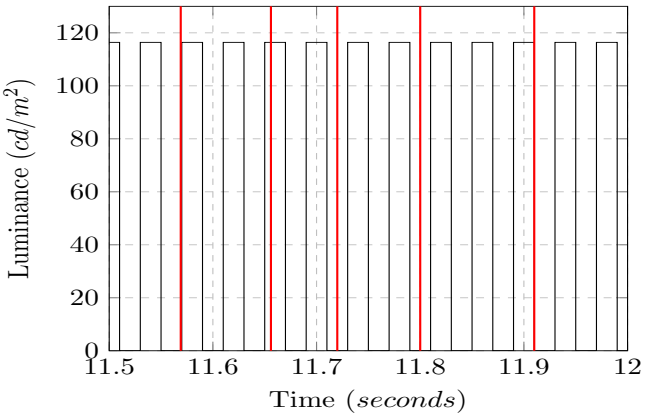
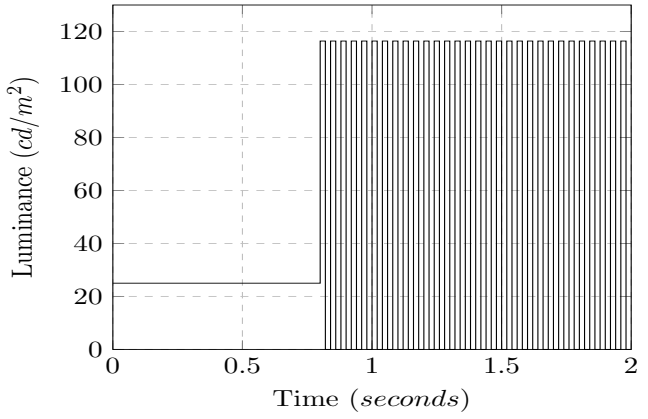
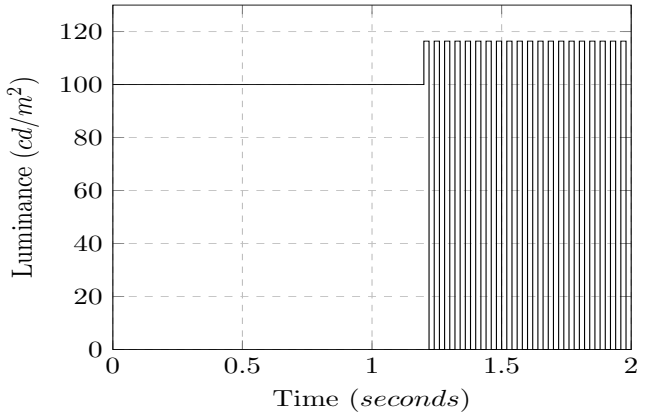


FIGURE 3.5: Five random points chosen in a photic pulse frequency 25 hertz, PCF $\frac{1}{2}$ and duration between 11.5 and 12 seconds. The classification made by a human subject at any of these points for a photic pulse will be the same.



(A) A photic pulse train with a constant amplitude perturbation of 25 cd/ft^2 for first 0.8 seconds



(B) A photic pulse train with a constant amplitude perturbation of 100 cd/ft^2 for first 1.2 seconds

FIGURE 3.6: The first two seconds of a ten second photic pulse with frequency $25Hz$, PCF $\frac{1}{2}$ and amplitude $116.4cd/ft^2$. The final classification made by the subject, after observing the stimulus for 10 seconds, would be independent of the perturbations in the initial two seconds.

3.3.2.1 Time series data point generation procedure

- Initially a rectangular wave was generated with duration of 12 seconds for a PCF, intensity and frequency using signal.square function in python scipy library. From this, a ten second wavelet starting anywhere within the first 2 seconds of the 12 second representation, was selected at random. This was done to randomize the phase shift of the rectangular wave. Some possible endpoints for wavelets of duration 10 seconds, when the starting point of the wavelet was chosen arbitrarily from 1.5 to 2 seconds of the initial 12 second wave, can be be seen in Figure 3.5. The final classification made by a human subject at any of the five endpoints as seen in Figure 3.5 will be the same. It is because, when a subject has observed a periodic photic pulse for a sufficient duration, his classification into steady or flickering will not change if he has observed the stimulus for a few more seconds. For a particular time period, there will be Timeperiod*samplerate datapoints which is equivalent to samplerate/frequency. So for a photic pulse with frequency 40 hertz and a particular PCF, it will correspond to 25 datapoints in Experiment 1 and 250 datapoints for Experiment 2.
- For the ten second wave thus selected, a constant amplitude was chosen at random from between intensity of the wave and zero. A random portion of the beginning of the wave, less than two seconds, was filled with this constant amplitude. Since the subject views the photic pulse for some duration before making the classification, we have assumed that the classification would be independent of the final phase of the wave and some random constant intensity at the beginning of the stimulus. Two such perturbations added to a photic pulse representation of frequency $25Hz$, PCF $\frac{1}{2}$ and amplitude $116.4cd/ft^2$ can be seen in Figure 3.6.

3.3.2.2 ANN library

Python based keras was used to train the model. The models were trained using Nadam optimizer with default values in the library.

3.4 Experiment 1 and its Results

This experiment used only a section of training data. Photic pulse representations with luminance 7 cd/ft^2 and PCFs $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$were used to train and validate the model. The photic pulses were sampled at 0.001 seconds and 10 second representations were generated for a particular set of parameters in the way described in the methodology section. For each PCF, four frequencies above and below the CFF from the parameter range in Figure 3.7 were chosen for a mini batch. The minibatches had 8 photic pulses for a particular CFF and each minibatch had 24 datapoints. New mini batches were

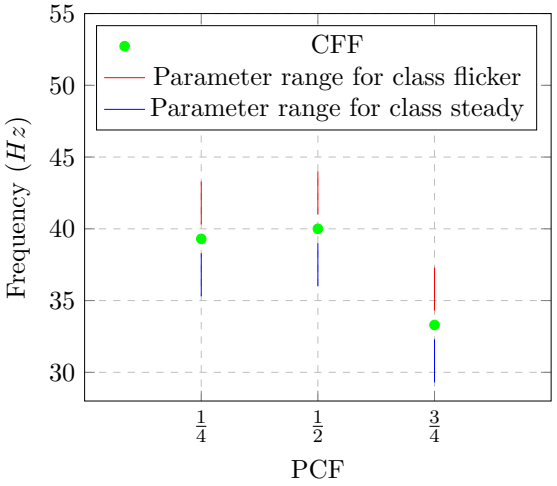


FIGURE 3.7: The parameter ranges of training data points used in Experiment 1. Only photic pulses of intensity $7cd/ft^2$ and three PCFs were used in training for the experiment. The corresponding frequency ranges for the pulses are shown in the diagram.

generated for training and validation in each iteration. Five different neural network architectures, including the proposed Convolutional Recurrent Neural Network (CRNN), were used to train the data.

3.4.1 Networks used

3.4.1.1 Simple RNN 1

In this neural network with three layers and 89 parameters, the input signal, was passed onto a simple RNN layer with 6 neurons. The final state of the recurrent layer was passed onto a dense layer with 5 neurons which was passed onto the final output layer with a single layer. Sigmoid activation function was used for all layers.

3.4.1.2 Simple RNN 2

In this neural network with three layers and 9036021 parameters, the input signal was passed onto a simple RNN layer with 3000 neurons. The final state of the recurrent layer was passed onto a dense layer with 10 neurons which was passed onto the final output layer with a single layer. Sigmoid activation function was used for all layers.

3.4.1.3 LSTM 1

In this neural network with three layers and 233 parameters, the input signal passed onto an LSTM layer with 6 neurons. The final state of the LSTM layer was passed onto a dense layer with 6 neurons which was passed onto the final output layer with a single

layer. Sigmoid activation function was used for the dense layer and tanh activation function was used for the LSTM layer.

3.4.1.4 LSTM 2

In this neural network with three layers and 36054021 parameters, the input signal was passed onto an LSTM layer with 3000 neurons. The final state of the LSTM layer was passed onto a dense layer with 10 neurons which was passed onto the final output layer with a single layer. Sigmoid activation function was used for the dense layer and tanh activation function was used for the LSTM layer.

3.4.1.5 Convolutional Recurrent Neural network

Convolutional Recurrent Neural Networks (CRNN) are artificial neural networks with convolution layers, a recurrent layer and fully connected layers[77]. The CRNN model used in the experiment had 1419 parameters. The input signal was passed onto convolution layer with 8 filters. The convolution layer had 160 elements in each filter. The output of the convolution layer was passed onto a recurrent layer with six neurons with sigmoid activation function. The final state of the recurrent layer output was passed onto a dense layer with five neurons and then to the output layer with a single neuron. Sigmoid activation functions were used for dense layers.

3.4.2 Results

The accuracies and mean squared errors for the five neural networks can be seen in Figure 3.8. The accuracies and mean squared errors in the figure have been Gaussian filtered with a sigma value of 10. The convolutional recurrent neural network architecture was found to be doing best work in training the data and so it was used for training entire dataset in Experiment 2 with further modifications inspired by a biological law as explained in the next section.

3.5 Experiment 2 and its Results

3.5.1 The Proposed Convolutional Recurrent Neural Network Model

The input representation was of dimension (100000, 1). The amplitude values in the input wave were transformed using dense networks as follows. Each timepoint in the network was initially passed onto a dense layer with 16 neurons. We have tested two variants of this architecture that differ in the activation function of this layer. The first

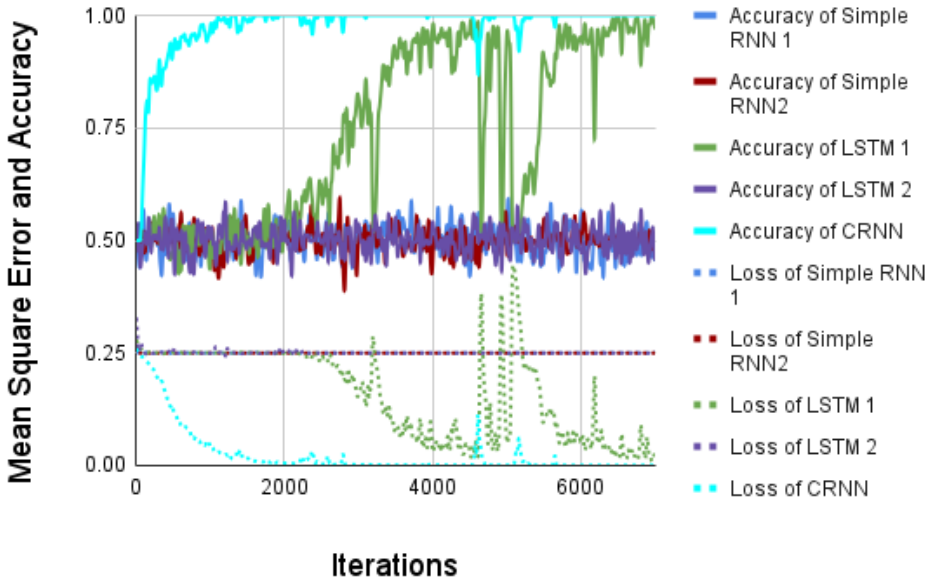


FIGURE 3.8: The accuracies and mean squared errors for each neural network over the iterations in experiment 1. The mean square errors are depicted with the same color as accuracy for a network but with dotted lines. The iterations have been Gaussian filtered with sigma value 10. As can be seen from the diagram, the Convolutional Recurrent Neural Networks (CRNN) perform the best in learning from data.

variant used sigmoid activation function and second variant used logarithmic activation. The DNN model with logarithmic activation is in Figure 3.9. In the variant with logarithmic activation, the weights and bias were constrained to be non negative and were initialized with value 1 for the training process. This logarithmic transformation operation was done taking inspiration from Ferry-Porter law and Fechner's law. While Ferry-Porter law states that $CFF = k(\log(L) - \log(L_o))$ where L is the luminous intensity of the photic pulse and L_o is the threshold intensity[78], the Fechner's law states that magnitude of sensation is not proportional to absolute value of stimulus but to the logarithm of the magnitude of the stimulus[79]. In this DNN, we have made the transformed stimulus passing onto higher layers, logarithm-like by expressing the transformed stimulus as a weighted sum of 16 different logarithmic functions, instead of making the transformation a strict mathematical logarithmic function.

The transformations were then added up separately into four different neurons. This was done as there are four types of photoreceptors, which are the rods and the three types of cones in the human eye. The weights were constrained to be non-negative for these neurons. The human visual system performs low pass and band operations on flicker stimulus[80]. Mammalian retina can also generate transient on and off responses to change in luminance[81]. These operations can be mathematically achieved by a convolution filter. Six different convolutions on the time domain were done on the output of all the four modeled receptor outputs. The convolution filters had a dimension of 1600 elements, six channels and used rectified linear (ReLU) activation function. Valid

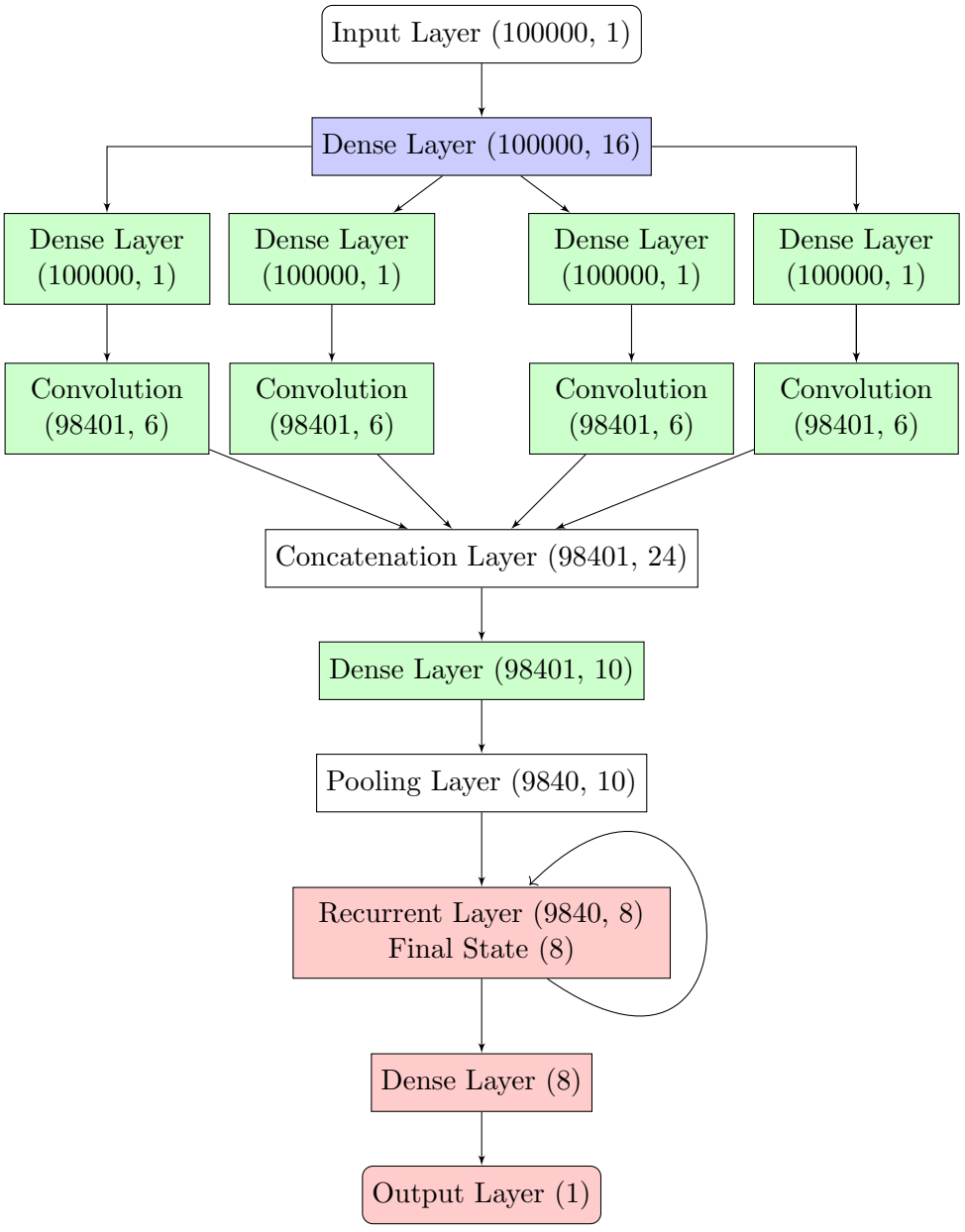


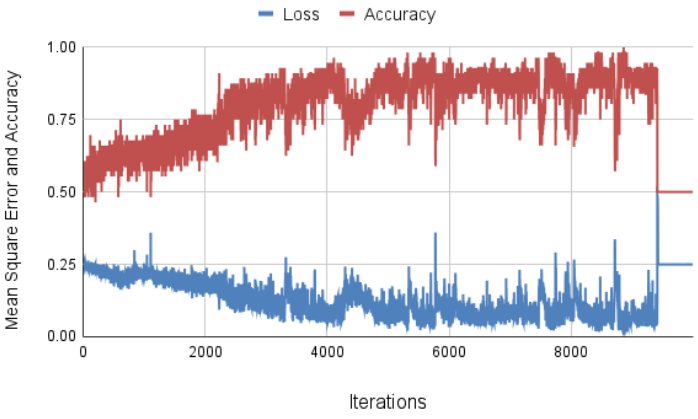
FIGURE 3.9: The model of the neural network used to train psychophysics data. The dimensions of the layers are given in brackets. The nodes colored red, green and blue have sigmoid, ReLU and logarithm as activation functions. The input intensity at a particular time was subjected to logarithmic transformations inspired by Ferry-Porter and Fechner's laws.

convolution operation gave time series representations with 98401 elements. The outputs of four such convolutions with six filters each, were concatenated back to produce a time series representation with 24 channels. The concatenated output was passed onto another dense layer with 10 neurons and rectified linear activation function. This was done to model any summations of various temporal responses that can occur in the visual system. The resulting output was average pooled with a pool size of 10. It reduced the dimension of time series data from (98401, 10) to (9840, 10). The resulting time series was fed to a recurrent layer with 8 neurons and sigmoid activation function. The final state of the recurrent layer was passed onto a dense layer with six neurons and then to the output layer with a single neuron. Sigmoid activation function was used for the final two layers. The output for a photic pulse that appeared flickering to the subject was set as one and that which appeared steady was set as zero. The network had 39007 trainable parameters.

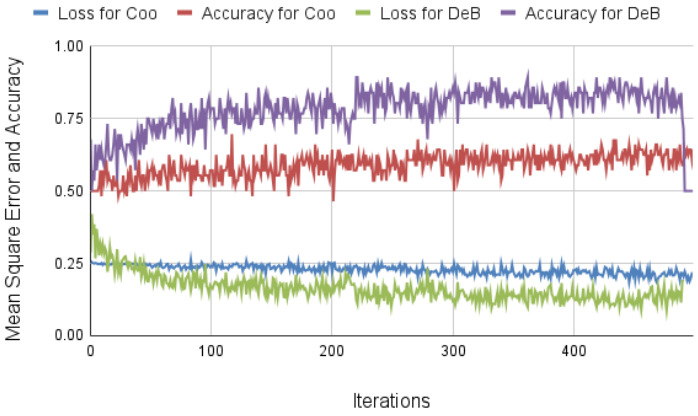
3.5.2 Training Process

The neural network was trained initially on the CFFs of the subject Coo. New mini batches were created for training and validation for each iteration. For each pair of intensity and PCF, four different frequencies were randomly chosen from the corresponding frequency range as in Figure 3.4 for both training and validation. So each mini batch had 224 datapoints. The loss and accuracy from validation minibatches were used to plot the learning curves of this model. The variant with sigmoid activation function in the first dense layer failed to fit the data. On the other hand, the variant with logarithmic activation function in the same layer was able to fit the data, and the changes in loss and accuracy over the iterations can be seen in Figure 3.10a. The subject whose CFF distribution was closest to *Coo* is *Ran* with Spearman's Correlation 0.906 and whose CFF distribution was farthest from that of *Coo* is *DeB* with Spearman's correlation 0.515. The neural network was then preloaded with weights of iteration 7655 on training with CFF data for Coo. It had a validation accuracy 0.9643 with loss of 0.0405. The weights corresponding to that iteration was chosen since it gave the highest accuracy after all the accuracies were Gaussian filtered with a sigma value of 10. The results for training with CFF data for *Ran* can be seen in Figure 3.10c and for *DeB* in Figure 3.10b. The model could fit CFF data for *Ran* with accuracy 0.899 and those of *DeB* with accuracy 0.869. The model was training faster with fewer iterations for *DeB* and *Ran* with preloaded with trained weights of *Coo* as can be seen in Figures 3.10b and 3.10c.

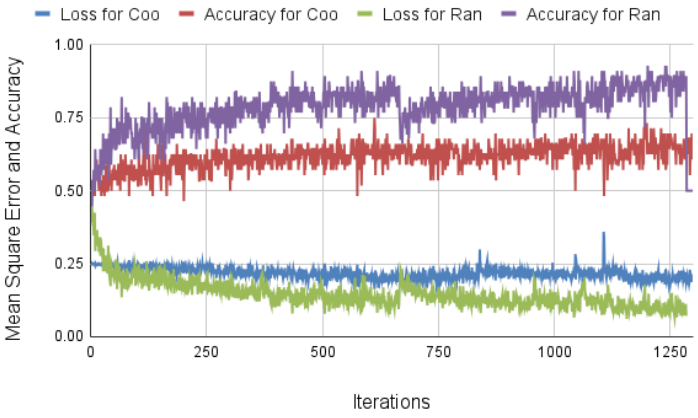
Psychophysics has been mostly devoted to fundamental aspects of perception common across humans than subtle aspects that vary between subjects[24]. Though the subjects differ in the psychophysics data of flicker perception, the faster training of the subjects data on the DNN pretrained with another subject shows that the DNN captures features common to the flicker perception mechanism in the three subjects.



(A) The changes in loss and accuracy of the neural network trained on CFF data of subject *Coo* over the iterations.



(B) The changes in loss and accuracy of neural network trained on CFF data of subject *DeB* over the iterations. The corresponding loss and accuracy for Subject *Coo* at the same iteration is also shown.



(C) The changes in loss and accuracy of the neural network trained on CFF data of subject *Ran* over the iterations. The corresponding loss and accuracy for Subject *Coo* at the same iteration is also shown.

FIGURE 3.10: The change in accuracy and mean square error for psychophysics data of three subjects *Coo*, *DeB* and *Ran* over the iterations for neural network with logarithmic activation in the first dense layer. The neural network was trained beginning with random weights for *Coo* and the pre-trained weights of *Coo* for *Ran* and DeB

3.5.3 Test for the model

To test the model, the classifications at all points were calculated from the recurrent layer output, while only the final layer output was used for training. It was done by passing on all the recurrent layer outputs to the next layer instead of just the final layer as was used for training. Since the output of a stimulus perceived as flicker was set to 1 and that perceived as steady to 0 during training, the output of any stimulus perceived as steady should be close to zero. A constant stimulus without any changing luminance, will be perceived as steady by a subject. The model was tested with constant stimulus representations of 1, 10, 100 and 1000 cd/ft^2. The model gave an output ~ 0.64 for a non intermittent stimulus. Though the DNN was shown to fit the data, The DNN failed in classifying a constant stimulus as steady. We tried to fix this limitation in Experiment 3.

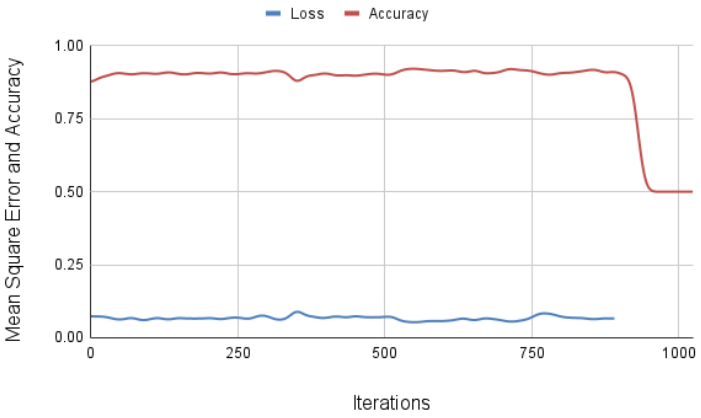
3.6 Experiment 3

Since, the network classified a constant non-intermittent stimulus as flickering in the previous experiment, the condition that a non-intermittent stimulus should be classified steady was introduced during the training process. After choosing a ten second wavelet from 12 second second wave as described in the methodology section, each representation in a mini batch with the label steady was manipulated as follows.

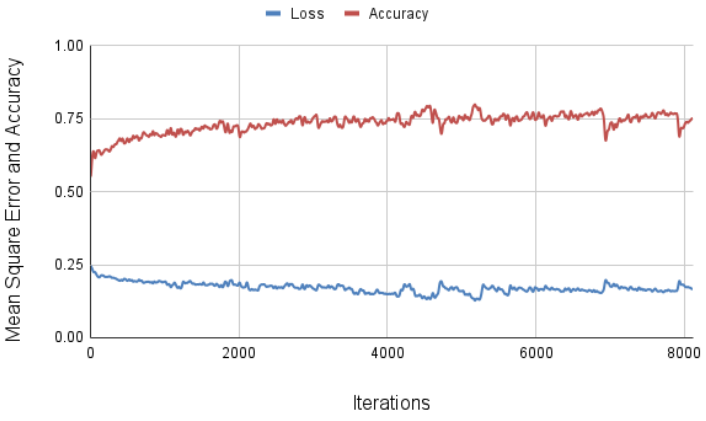
- there was a probability 0.25 that the entire representation was filled in with zeros.
- there was a probability 0.125 that the entire representation was filled with a frequency between 0 and intensity of the pulse.
- there was a probability 0.125 that a timepoint was chosen at random from the first nine seconds of the representation. The remainder of those timepoint was filled with a random intensity between 0 and intensity of the photic pulses.
- There was a probability of 0.5 that the representation would be left as it is.
- A perturbation was added to the beginning of this photic pulses as described in the methodology section.

The network was trained from pre-trained weights from Experiment 2. The results of the training can be seen in Figure 3.11a. The network, after training, was fed with constant stimulus of intensities 10, 1000 and 1000 cd/ft^2, and gave a neuronal output of ~ 0.37.

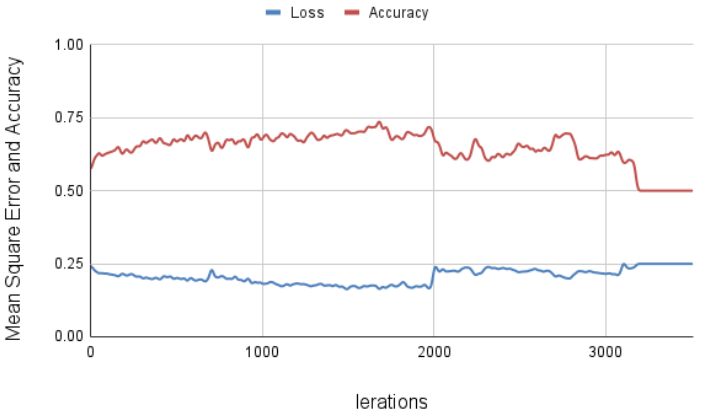
The network could not be trained with new methodology from scratch as can be seen in Figure 3.11b. Since a non intermittent stimulus should be classified as steady,



(A) Results of training with pretrained weights from Experiment 2



(B) The results of training the network from scratch.



(C) Training results of network with frequency range of flicker stimulus from CFF-1 to CFF-2

FIGURE 3.11: The changes in loss and accuracy of the network with methodology of Experiment 3. The plots have been Gaussian filtered with sigma value of 10.

a photic pulse stimulus with very low frequency should also be classified as steady by the network except when there is a change of luminance. Since we do not know this cutoff for low frequency, the network was trained with a reduced frequency range for flickering. The frequency range for flicker stimulus was changed from CFF-1 to CFF-4, to CFF-1 to CFF-2. The results of training the network with new range can be seen in Figure 3.11c. The network could not be trained from scratch even in the new training set range.

3.7 Possible Tests for the model

The DNN trained in Experiment 3 with the pretrained weights of Experiment 2 was chosen to test the model. Experiment 3 weights were chosen because they met the condition that they classified a steady state stimulus as steady. The classifications for all time points were generated by the DNN for testing instead of just the final recurrent state as was used for training.

3.7.1 Output for 1 hertz stimulus

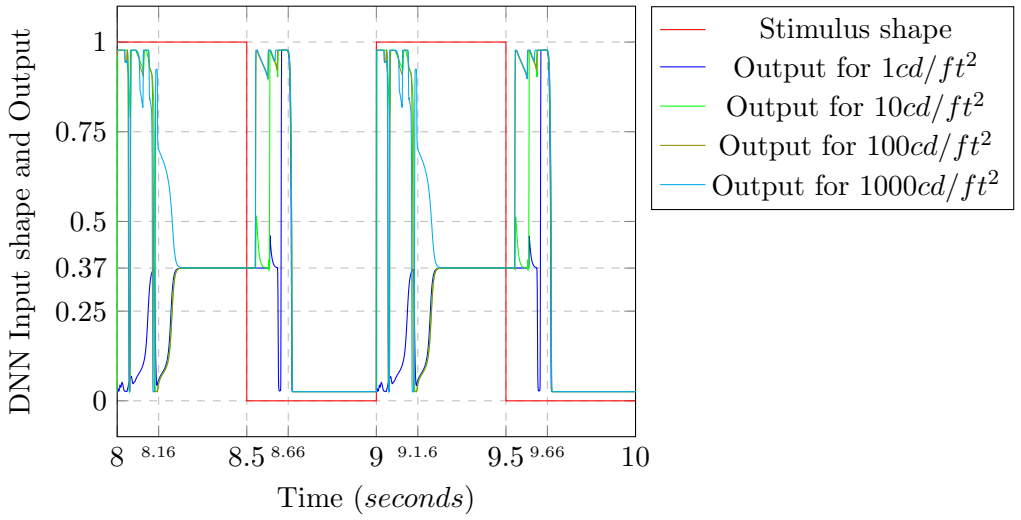


FIGURE 3.12: The DNN output for two seconds duration against pulses of frequency 1 hertz, PCF $\frac{1}{2}$ and luminances of 1 , 10, 100 and 1000 cd/ft^2. A constant stimulus input gives an output 0.37 and a 0 cd/ft^2 input gives as output 0. The output close to 1 indicates classification as flicker and that close to zero indicates the classification as steady. The DNN output rises close to 1 when there is a sudden change of intensity to or from 0 cd/ft^2. As can be seen from the figure, after a sudden change in intensity, it takes more than 160 milliseconds for the DNN to output to fall to 0.37 or 0.

A subject should not perceive flicker for a constant stimulus of any intensity. The flicker should be sensed only when there is a sudden transition of intensity. The DNN was tested for this with a square wave stimulus of time period of 1 second and PCF $\frac{1}{2}$, for luminances 1, 10, 100, 1000 cd/ft^2. The shape of photic pulse and corresponding

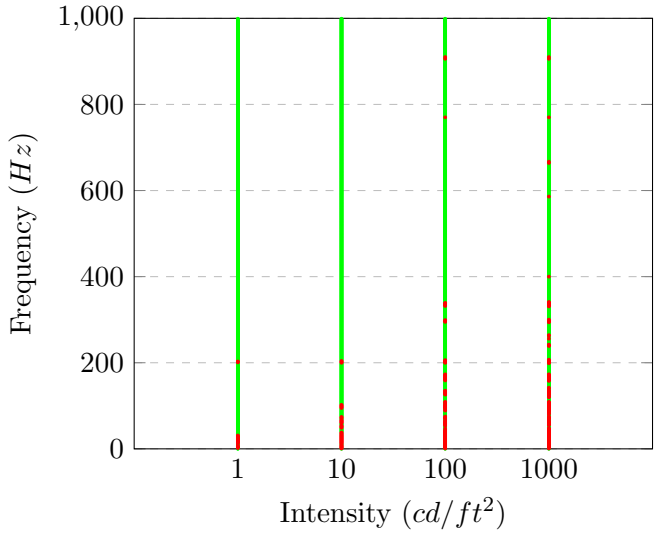
outputs can be seen in Figure 3.12. The DNN had been trained to give an output of 1 for perception of flicker and 0 for steady. The DNN gives output ~ 0 when there is a steady input of 0 cd/ft^2 and an output ~ 0.37 when there is a constant amplitude stimulus. It can be observed in Figure 3.12 that the classification of the stimulus changes to flickering when there is a sudden change in amplitude of the stimulus. After a sudden fall in intensity to 0 cd/ft^2 or a rise from it, the classification comes back to steady after more than 160 milliseconds which is the time duration of the convolution filter.

3.7.2 Square wave stimulus

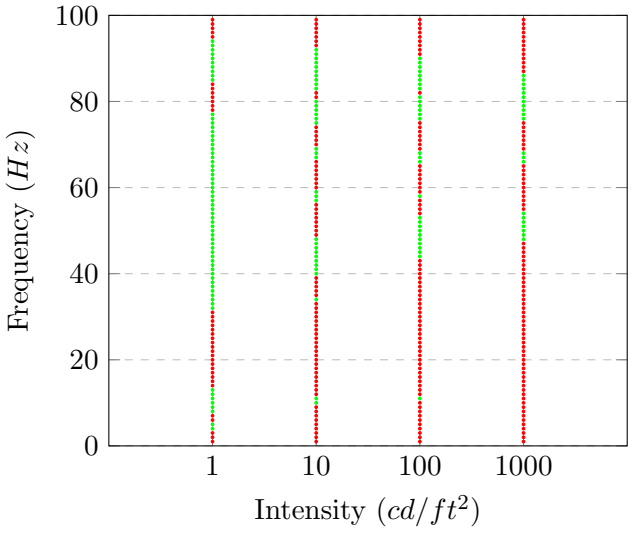
Flicker stimuli of intensities 1, 10, 100 and 1000 cd/ft^2, with PCF $\frac{1}{2}$, duration of ten seconds and frequencies from 1 to 1000 hertz were fed into the input layer. The mean of the last 1 second output with a cutoff of 0.5 was used for classification. Photic pulses with very high frequencies should appear as steady for a subject. It can be seen in Figure 3.13a that the DNN satisfies this criteria for most of the frequencies for all the four intensities. However,the network classifies some frequencies above CFF as flickering. Moreover, it can be seen from Figure 3.13b, that the DNN classifies frequencies as high as 100 hertz as flickering which is not perceivable to a human subject .

3.7.3 Sinusoidal Stimulus

The stimulus was tested with waveforms of the form $I(t) = I(1 + \boldsymbol{m}sin(2\pi ft))$ as in Figure 3.2a. The stimulus had luminances I with values 1, 10, 100 and 1000 cd/ft^2 for stimulus duration of 10 seconds. The means of the output for the last 1 second with threshold value 0.5 have been used to generate the contours for transition between flicker and fusion in Figure 3.14a. The transition contours for a subject for various retinal illuminations in trolands have been plotted in Figure 3.14b from psychophysics data published by Kelly[82]. The transition data for four subjects for 850 trolands retinal illumination obtained by Kelly[83] has been plotted in Figure 3.14c. In Figure 3.14c two subjects show two peaks in the transition curve instead of a single peak unlike the other two subjects in Figure 3.14c and the subject in Figure 3.14b. The boundaries obtained by DNN in Figure 3.14a show several peaks. The model detects sinusoidal flicker at higher frequencies while failing to detect flicker at very low modulation values as in real psychophysics experiments plotted in Figures 3.14b and 3.14c. The psychophysics readings provided for comparison in Figure 3.14 were taken for a stimulus with a larger visual angle than that was used to train the network[82]. The psychophysics readings with smaller visual angles than the ones plotted for comparison also have a similar shape without multiple peaks[60].

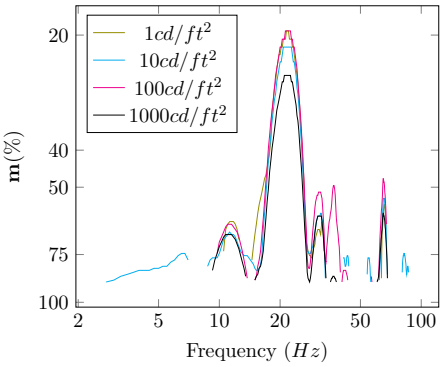


(A) The figure shows classification done by the DNN for photic pulses of frequencies upto 1000 hertz. The DNN for most part classifies photic pulses with very high frequencies as steady, which hold true for humans.

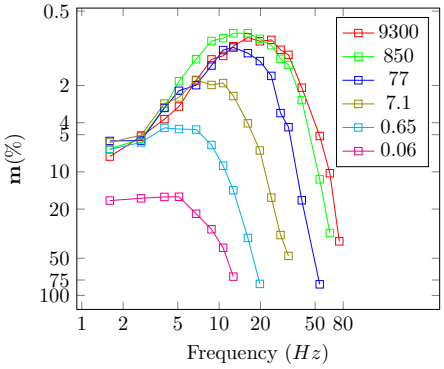


(B) This shows the classification done by the DNN on photic pulses of frequencies from 1 to 100 seconds. The DNN classifies the photic pulse of frequencies as high as 100 hertz as flickering, which is not perceivable to a human subject.

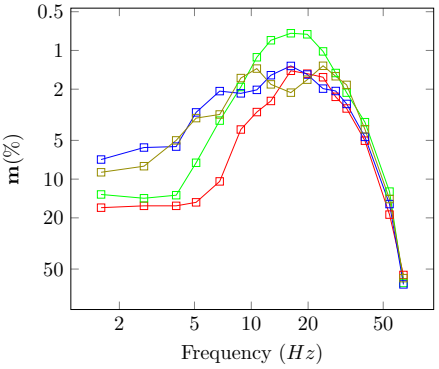
FIGURE 3.13: Classifications made by DNN for square wave stimuli of duration 10 seconds with PCF $\frac{1}{2}$ for four intensities. The classifications were made for photic pulses of each frequency by taking mean of the last 1 second of the output with 0.5 as cutoff. Red points indicate classification as flicker and green points indicate classification as steady.



(A) The boundaries generated in the predicted output of the DNN for four different luminances.



(B) For different retinal illuminances in trolands, for a subject published by Kelly[82].

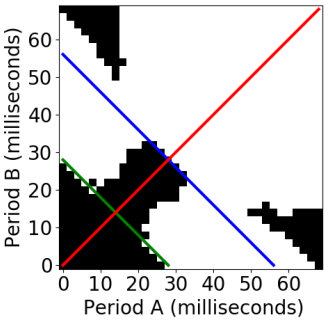


(C) Psychophysics data of four subjects for retinal illuminance 850 trolands published by Kelly[83]. The psychophysics data of two subjects (olive and blue) show two peaks instead of a single peak like the other two.

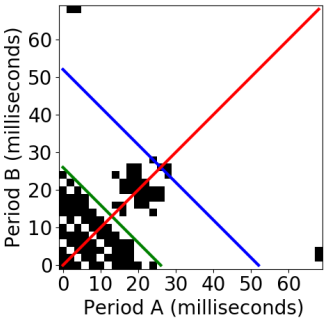
FIGURE 3.14: The boundaries between regions of flicker and fusion for sinusoidal stimulus of waveform $I(t) = I(1 + \boldsymbol{m}sin(2\pi ft))$ as in Figure 3.2a. The parameter **m** has been plotted against frequency of the pulses.

3.7.4 Waveforms with pulses of alternate time period

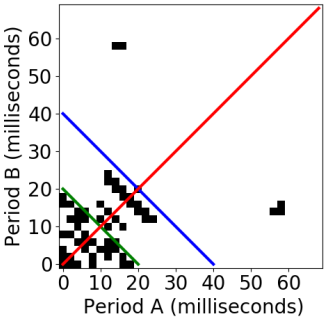
The DNN was tested with stimulus of waveforms in Figure 3.2b for luminances 1, 10, 100 and 1000 cd/ft^2 and ten second duration. The mean of the last 1 second output with a threshold value of 0.5 was used for classification of the stimulus. The classifications of combination of timeperiods as flickering and steady, generated by the DNN for the four luminances are in Figure 3.15. The shape of fusion contour for the stimulus determined via previous experiments have been plotted in Figure 3.15e. The previous works had determined that the boundary conditions for the fusion contour is determined by the lines $x + y = P$ and $x + y = 2P$, where P is the period associated with CFF of a simple stimulus of PCF $\frac{1}{2}$[61]. The lines $x = y$, $x + y = P$ and $x + y = 2P$ have been plotted for predictions made by DNN in Figure 3.15.



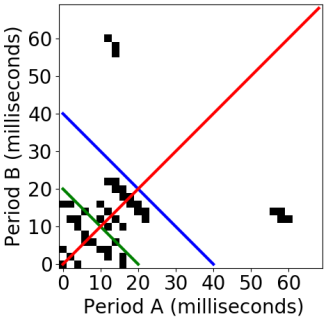
(A) Luminance 1 cd/ft^2



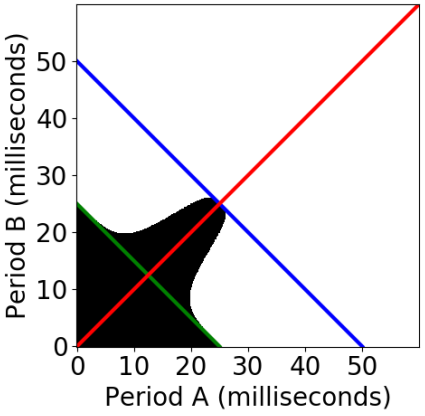
(B) Luminance 10 cd/ft^2



(C) Luminance 100 cd/ft^2



(D) Luminance 1000 cd/ft^2



(E) The shape of fusion contours for subjects obtained from previous psychophysics experiments[61, 84]

FIGURE 3.15: The first four subfigures show the predictions made by the network for photic pulse inputs of the waveform as in Figure 3.2b for four different luminances. The white regions represents the region classified as flickering by the DNN and the dark region represents the region classified as steady. The shape of fusion contour from previous psychophysics experiments have been plotted in the last subfigure. Previous psychophysics experiments have determined that the shape of the figure is determined by equations $x = y$, $x + y = P$ and $x + y = 2P$. P is the period in which a simple flicker of PCF $\frac{1}{2}$ changes from flickering to fused[61]. The equations $x = y$, $x + y = P$ and $x + y = 2P$ have been plotted in figures red, green and blue lines.

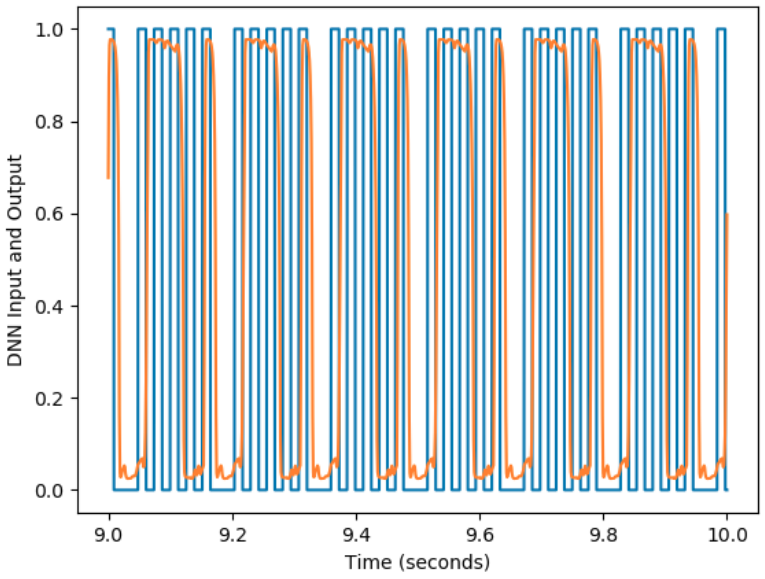


FIGURE 3.16: Nelson, Bartley, and Harper had generated short pulse trains with gaps proportional to the timeperiod of the pulse. The psychophysics experiment consisted of fusing the short pulse trains instead of the whole train[62]. Here the DNN has been used to simulate the perception of a similar stimulus with 5 pulses of PCF $\frac{1}{2}$ with a gap of 1 pulse and luminance $1cd/ft^2$. If the short pulse train is classified as fused by a subject, the output should be close to 1 only at the beginning and end of the short pulse train, similar to a very long pulse as in Figure 3.12. As of now, the proposed model has not been able to find cutoff frequencies below which short pulse trains show such behavior. For the stimulus with timeperiod of 26 milliseconds, the output of the DNN shows an output closer to the desired output when the short pulse trains are fused.

3.7.5 Fusion of short pulse trains

The proposed model can be used to test fusion of short pulse trains with a gap proportional to the time period of a single pulse as in Figure 3.2c. Nelson, Bartley, and Harper had measured CFF of short pulses for combinations of various numbers of pulses, PCFs and gaps for two subjects and five different intensities. In this experiment frequencies were adjusted to fuse the short pulse trains instead of the whole photic pulse train[62]. This experiment is different from the binary classification experiment on which the network was trained on, and the mean of the output or the final element of the output cannot be used to predict the fusion of short pulse trains. The proposed model can still be tested with the fusion of short pulse trains. The testing can be made with the assumption that when the short pulse trains are fused, the fused short pulse trains will be perceived as a single big pulse and the output for the pulses will be the same as that of long pulses in Figure 3.12. So for a short pulse train with frequency above its CFF, the DNN should give an output close to 1 only at the beginning and at the end of a short pulse train. At present, the network was not able to detect such a behavior with a cutoff frequency for short pulse trains. The output of a stimulus of luminance 1 cd/ft^2 with five short pulse trains followed by a gap of 1 pulse, and time period of 26 milliseconds along with the input stimulus can be seen in Figure 3.16. For this particular input, the

output of the DNN has been close to 1 only near the beginning and end of the short pulse train, which should be the case for an ideal network when short pulse train appear as fused.

3.7.6 Similarities between the intermediate layers of the DNN and the EEG response to flicker.

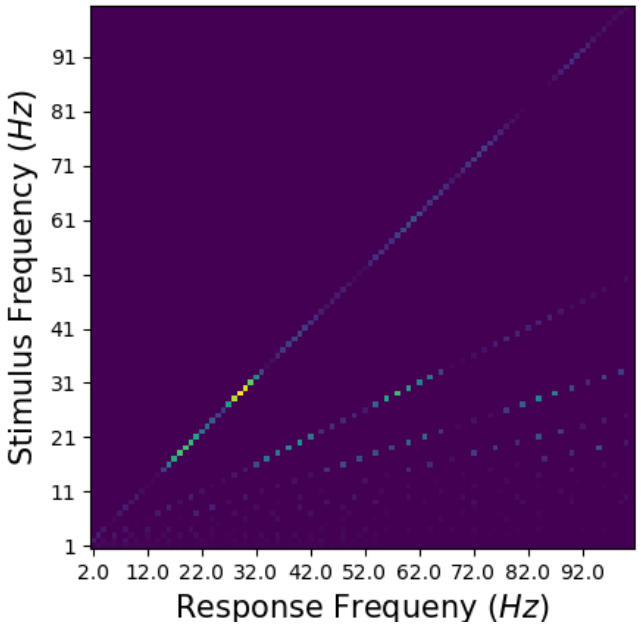


FIGURE 3.17: The response frequencies in the output of the first neuron in the third dense layer of the model have been plotted as a function of the input frequencies of the stimulus. Simple square wave photic pulses with luminance $1cd/ft^2$ and PCF $\frac{1}{2}$ were used as input. The response frequencies were plotted on the X axis in the same fashion as previously plotted by Herrmann on human EEG response to flicker stimulus[65] and flicker response in cat cortex by Rager and Singer[85]. Fundamental frequency, harmonics as well as first subharmonics are present in the human EEG response to flicker stimulus of various frequencies[65]. The plotted diagram shows stimulus frequencies generating fundamental frequency and harmonics in the response which is the same as response in area 17 and 18 of cat cortex to flicker stimus as plotted by Rager and Singer[85].

A related domain in the study of flicker perception is the electrophysiological response to flicker stimulus. So far, there does not seem to be much of a correlation between EEG response to flicker and the psychophysics of flicker fusion[86]. We propose here that correlations between psychophysics of flicker fusion and electrophysiology of flicker stimulus can be made by comparing the output of the intermediate layers of DNN trained with psychophysics data of flicker fusion with the recorded electrophysiological response to the same stimulus. Harmonics and sub-harmonic oscillations in addition to fundamental frequency of flicker stimulus have been detected in the human EEG response to flicker stimulus[65]. The sub-harmonic oscillations were reported to come from the parietal electrodes instead of occipital electrodes unlike in the case of harmonics[74]. No

clear evidence of sub-harmonic oscillations were reported in local field potentials and multi unit activities from areas 17 and 18 of cat visual cortex stimulated with flicker[85]. The harmonics as well as the fundamental frequency are present in the output of a convolution response to flicker stimulus as can be seen in Figure 3.17. Similar to the output in Figure 3.17, fundamentals and harmonics are evoked in the EEG response to flicker stimulus even if the subject is not able to perceive flicker[65]. Harmonics upto fourth harmonic are evoked in human EEG response to flicker [65] which is also present in the output of midlayers of DNN as can be seen in Figure 3.17.

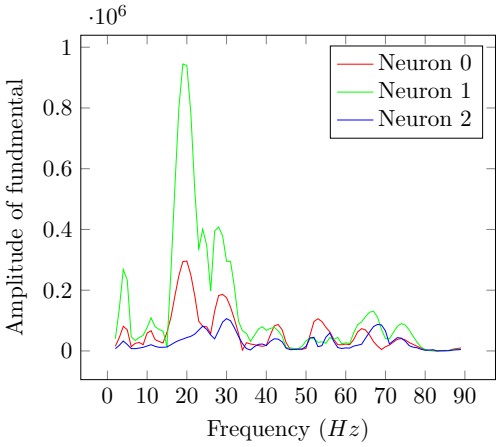


FIGURE 3.18: The amplitude associated with fundamental frequencies for three neurons in a convolution operation done by the DNN. Square photic pulses with PCF $\frac{1}{2}$ and luminance 100 cd/ft^2 were fed as input. The amplitudes of fundamentals show distinct peaks. The amplitudes of fundamentals from EEG response to flicker had shown distinct peaks [65] but with fewer in number than the current simulation.

The human EEG response to flicker stimulus shows resonance at frequencies around 10, 20, 40 and 80 hertz[65]. The average of fundamental amplitude for 10 subjects measured by Herrmann, had shown strong resonance at 10 hertz and weaker peaks at 10-10 and 35-35 hertz range[65]. The profile of the fundamental for three neurons in the first convolution operation of the network, plotted in Figure 3.18 shows distinct peaks at certain frequencies. However the Figure shows much more distinct peaks than have been observed in human EEG response to flicker. The resonance behavior at certain frequencies in human EGG response to flicker stimulus could be the result of a convolution operation.

3.8 Discussion

3.8.1 Contributions

We have proposed a Brain Computational Model for flicker fusion and conducted a host of computer simulation experiments on the perception of flicker. The proposed model

can fit the psychophysics data for a human subject the assumed parameter range. We also show that a model once trained on a particular subject, could fit faster for the psychophysics data of another subject whose CFF distribution is different from that of the first subject, implying the DNN learns features of flicker perception mechanism, common across subjects. The network may be trained with rectangular wave photic pulses as in Figure 3.1 and tested with other waveforms as in figures 3.2 and 7.9. The proposed BCM also gave a methodology to bridge the gap between psychophysics and brain signals associated with flicker by suggesting that human EEG response to flicker stimulus can be the result of a convolution operation on the stimulus.

3.8.2 Comparison with previous models

The present work proposes that a top down model of flicker fusion can be built from psychophysics data using backpropagation technique. Further, signals associated with flicker like EEG or ERG are not needed in building and training the model. This work proposes a new top down approach to model flicker fusion and electrophysiological response to flicker which is distinct from the earlier models[70, 71, 72, 73, 74].

3.8.3 Limitations

The model was trained with the assumption that for a given intensity and PCF, photic pulses with frequencies between CFF+1 and CFF+4 will appear as fused and those with frequencies with CFF-1 and CFF-4 will appear as flickering. This assumption might not always hold in reality as can be seen in some previously published psychophysics results as demonstrated in Chapter 7. We have only provided photic pulses with frequencies upto CFF+4 for the label steady. So it can happen that network will learn to classify regions with frequencies CFF+1 to CFF+4 as steady, but at the same time photic pulses with frequencies greater than CFF+4 can get labeled as flickering which happened to the network as can be seen in Figure 3.13b.

We have plotted results from previous psychophysics works for subjective comparisons to test the model. The subjective comparisons showed only rough similarities with the published psychophysics data. Extra peaks not seen in the data from human subjects have been produced by the proposed network in Figure 3.14a. Extra peaks have also been generated by the network in the profile of fundamental frequency in convolution operation on the stimulus as compared to the observed data in human EEG response to flicker(Figure 3.18). Both of these can be the outcome of the convolution layer being non-optimal. Though we have shown that being pretrained with a subjects psychophysics data, can make the training of the same DNN on other subjects faster as can be seen in Figures 3.10b and 3.10c, the network failed to fully fit the parameter ranges associated with the new subjects with near 100% accuracy as was possible with the first subject.

The comparisons for other flicker stimuli, have not been made on the psychophysics data of the subject used to train the DNN. There exists subjective variation in perception of flicker. This could be overcome by gathering training data and test data from the same subject. The hyperparameters used in the model like the dimension of the convolution filter may not have been the optimum. This could have prevented the model from getting fit to give the most psychophysically faithful results.

3.8.4 Future Works

The proposed neural network, in future, can be trained and tested on psychophysics data from a single subject under similar conditions. This work trained DNN with an achromatic flicker stimulus. DNNs can be trained with psychophysics data of chromatic stimuli as in Figure 7.9b by using different channels for each color in the input representation of the stimulus. Further correlations could be probed between internal representations of a DNN trained on the psychophysics data of a subject as well as fMRI and EEG data acquired from the same subject.

3.9 Conclusions

We conclude that a Convolutional Recurrent Neural Network with the time series representation of intensities subjected to a logarithmic transformation can be used to construct a trainable and testable Brain Computational Model of flicker perception.

- The work proposes a deep neural network model that could take in real time inputs like human visual system and make a classification of flicker fusion the same way a human subject does.
- The model uses a logarithmic transform of the stimulus intensity before it is fed into the classifier. This was in accordance with Fechner's law that says that magnitude of sensation is not proportional to the perceived value of the stimulus but to its logarithm. The model failed to fit the data when logarithmic transformation was replaced by sigmoid transformation.
- The work shows that Convolutional Recurrent Neural Network (CRNN) is better to train time series flicker data than a simple RNN or LSTM.
- The work trains a CRNN making assumptions that for a given PCF and luminance, waveforms with frequencies above the CFF will classify the stimulus as steady and those below it as flickering. The real psychophysics data can have outliers as demonstrated in Chapter 7.

- The work has simulated psychophysics experiments of flicker pattern different from the one used to train the network. Though the simulated output for testing is not entirely in agreement with the reported ground truth, it shows that the network can be used to test on totally unrelated stimulus patterns.
- The proposed model can be used to explain the correlations between electrophysiological responses to flicker stimuli and the psychophysics of flicker perception. The presence of fundamental and harmonics of the pulse frequency in electrophysiological response to flicker can be explained as the action of a temporal convolution operation on the stimulus.

Chapter 4

A biologically inspired model for the perception of flicker wheel illusion

Chapter summary:

Flicker Wheel illusion is a visual illusion of a wheel with spokes which appear to flicker. The perception of flicker wheel illusion is affected by physical parameters of the stimulus like number of spokes and contrast. This work generates multiple variants of flicker wheel by varying physical parameters associated with the stimulus. The stimuli are then presented on a computer screen and is labeled by a human subject into two classes on basis of whether the subject perceives flicker in the stimulus or not. DNNs are trained to classify the images the same way a human subject does. Many existing CNN Networks are shown to be able to classify the data. However, the best performance comes from the proposed bio-inspired Deep Neural Network that takes into account the microsaccades in the human eye. The proposed architecture yields accuracies close to the consistency of human subject in the psychophysics experiment.

4.1 Introduction

Visual illusions [14] are like gaps in our visual mechanism through which we can spy, in a non-invasive manner, into the internal mechanisms of the brain. The illusions such as those which are geometric [15], or related to brightness [16],or motion [17] etc., where our perception fails to replicate the environmental reality, are the crucial tools to

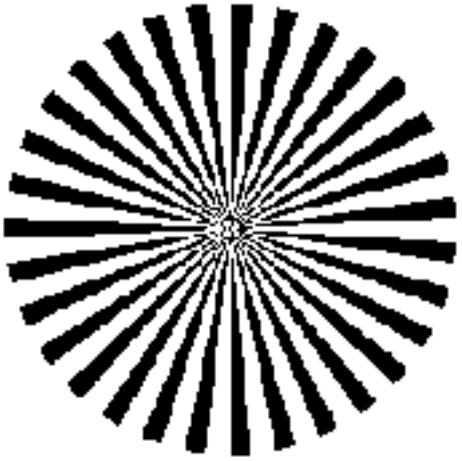


FIGURE 4.1: The Flicker Wheel Illusion

understand the visual brain's internal processing. Among the class of motion illusions, there exists those where a static image may be perceived as moving e.g. the Peripheral Drift Illusion (PDI) [18]. A deep neural network (DNN) is an Artificial Neural Network with multiple hidden layers between input and output layers which perform linear and non linear transformations[87]. Biologically inspired Deep Neural Network(DNN)s like PredNet have been able to replicate the perceived motion in optical illusions like rotating snake illusion[88].

This chapter attempts to model simpler illusions in which there is no perceived direction of rotation of motion. To this end, there exists a well-known phenomenon called flicker fusion, as explained in detail in the previous chapter, where the opposite happens, viz. a subject is unable to perceive a source of light as flickering beyond a frequency threshold of the flicker[60]. Here, instead of modeling with real physical flickering sources as in the previous chapter, we have taken recourse to a standard and well-known visual illusion based study. This is the simple image of a wheel consisting of several spokes, which under suitable conditions, appears as a flickering source on the paper or computer monitor [19]. Data on flickering and still images have been generated by varying physical parameters associated with the stimulus and having them labeled as flickering and still by human subjects. We then go on to propose a deep neural network architecture that may be able to incorporate the microsaccadic eye movement.

4.2 Background and Related Works

The flicker wheel illusion consists of a wheel with 30 to 40 spokes as shown in Figure 4.1. It appears to flicker vividly and regularly[19]. The sensation of flicker so induced is influenced by the contrast in the image as well as the number of spokes[19]. Understanding the human cognition for such a task will require building of models that can

perform cognitive task, and test them with behavioral and brain experiments[34]. Deep Learning has been put forward as a tool for modeling the brain[2]. Although, as already mentioned, biologically inspired deep neural networks like PredNet have been able to replicate the perceived motion in optical illusions like the rotating snake illusion[88], Funke et al. have argued that the mechanism of perception of rotating snakes illusion should be different from that of PredNet, since the network does not emulate such biological processes as microsaccades which are necessary for the perception of rotating snakes illusion[89].

Physiological phenomena like eye movement and pupil measures are day by day assuming greater importance for the computer vision community [90]. Saccades are eye movements that occur when the eye is not fixated on an object. Fixational eye movements are the smaller movements that occur when the eye is fixated upon and object. The fixational eye movements are of three types drift, microsaccades and tremor. Drifts are slow moments that occur between inter saccadic intervals and are similar to random walk problem. Microsaccades are fast eye movements that cause small shift in position of the eye that happens a couple of times a second when the eye is fixated on an object. Microsaccades are much faster than drift[91]. By tracking the eye movements of subjects looking at the rotating snakes illusion and comparing it with their perception of movement in the illusion it has been shown that microsaccades are necessary for the perception of motion in rotating snakes illusion[92]. Microsaccades have also been shown to be a factor behind perceived motion in enigma illusion[93]. Although, the microsaccades are not absolutely necessary for perception of Flicker Wheel illusion since the flicker illusion can also be observed in the afterimage on the retina, but the perception of flicker in flicker wheel illusion is strongest in the presence of the small eye movements[19]. Hence incorporation of this physical phenomenon in any neurocomputational modeling of flicker wheel perception needs to be considered as an important plank.

Some attempts have been made to model eye movements on responses of retinal ganglion cells [94, 95], but DNNs have till now not been related to such models.

4.3 Dataset Generation

The images were generated with the shape 224*224*3. Multiple images were generated by varying the physical parameters associated with the stimulus. The images consisted of a spokes generated around a point. In the original flicker wheel illusion the spokes were enclosed in a circle [19]. In the present work the spokes were enclosed in an ellipse instead of a circle and another small elliptical portion of the wheel was left blank without any spoke. We transitioned from a circular to an elliptical stimulus to better reflect the anisotropy of human vision, as our perception and sensitivity often differ

TABLE 4.1: The numerical ranges of parameters associated with variants of flicker wheel

Parameter	**Value range**
RGB values of the wheel	*0 to 255*
RGB values of the background	*0 to 255*
Lengths of semi major and semi minor axes of inner ellipse	*5 to 30 pixels*
Lengths of semi major and semi minor axes of outer ellipse	*50 to 100 pixels*
Number of Spokes	*3 to 80*
Angular thickness of spoke	0.2 to 0.8 times angle subtended by the spoke
Angular rotation of the flicker wheel	0 to angle subtended by the spoke.

along the horizontal and vertical axes[96]. An ellipse was chosen because it is the most parsimonious geometric extension of a circle that allows us to account for these axial variations. The major and minor axis of both ellipses were embedded along the X-Y axes of the image. The RGB values of the wheel and background were chosen at random. The parameter angle of a spoke, associated with the stimulus, is $\frac{2\pi}{Numberofspokes}$. Another parameter that was varied was the angular thickness of a spoke. After the wheel was created using these parameters, the wheel was subjected to rotation by an angle. The range of physical parameters associated with the stimulus are provided in Table 4.1. The generated images were saved as .png files to a folder. Two such images generated by the method can be can see in Figure 4.2.

The labeled dataset were created via a psychophysics experiment. PsychoPy package of python was used to design the experiment. The images were displayed on an Eizo ColorEdge CG247 Monitor. The image was displayed on the middle of the screen and the entire screen was filled with same color as background of the flicker wheel image. The subject was seated comfortably before the monitor in a dark room with head approximately 80 cm away from the monitor. The subject entered the responses via key presses in a keyboard. The subject pressed the button Y in case motion was perceived in the image. In case the subject did not perceive motion in the image, the subject pressed the key N. Subject could skip classification of any image by pressing S. Immediately after the key-press, the subject response and reaction time was tabulated in a text file and the next image in the folder was loaded onto the screen. The subject stopped the experiment, whenever he wanted, by pressing the key Q. The classifications were done over a period of many days.

In case of psychophysics experiments there will be some difference in subject response to the same stimuli in between subjects, as well as some inconsistency in response of

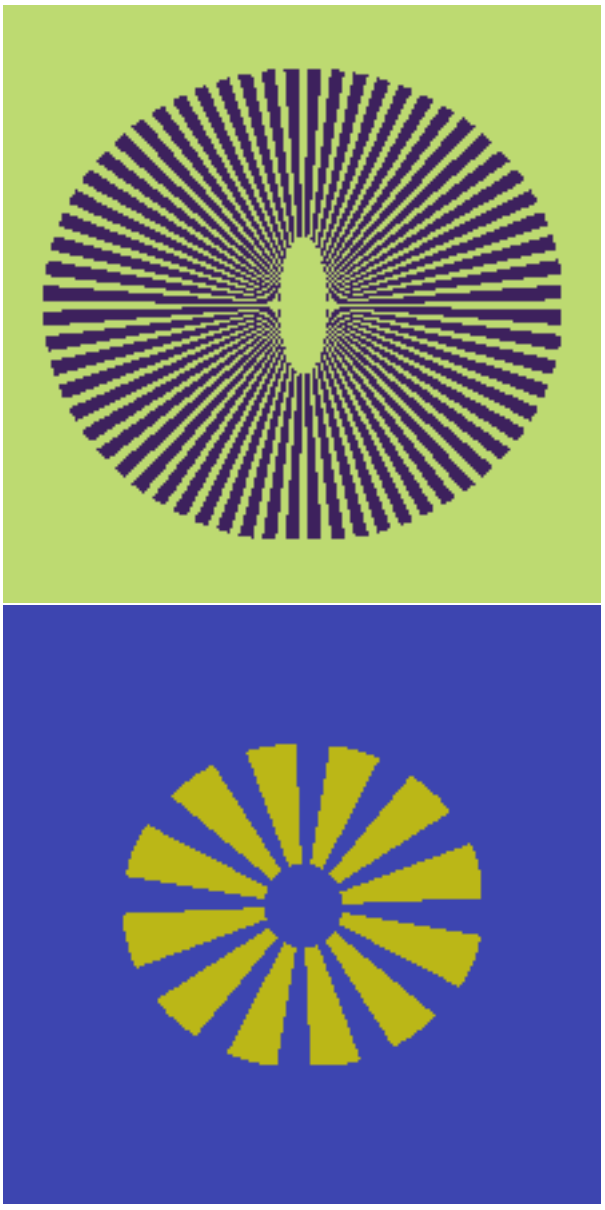


FIGURE 4.2: Two of the images generated for psychophysics experiments. The images might be perceived by a human subject to be flickering, or a still one, based on the physical characteristics associated with the stimulus.

the same subject to the stimulus[24]. A subject who once detected flicker in an image may not classify the same image as flickering when repeated later. If the response of the subject is recorded by key presses, there is also a chance that the subject will give wrong output by mistake. However we can overcome the problem of inconsistency by increasing the size of the dataset. The labeling were done by two subjects. The first subject was a 31 year old male with normal vision and the second subject was a 24 year old female with corrected vision.

TABLE 4.2: The classification of first 119 images done by the male subject repeated twice. The total number of classifications into a label for each set of readings are there in the last row and columns.

		Classification made on repetition		
		Flicker	Still	
Classifications in initial experiment	Flicker	23	11	34
	Still	5	80	85
		28	91	

4.3.1 Labeling by the male subject

The experiment was done on 2569 images of which the subject classified 912 images as flickering and 1657 images as still. The median time taken per image for classification was 2.065 seconds. In order to check the consistency of subjects readings, the experiment was repeated again for the first 119 images. The results of the repeated experiment and initial classifications can be seen in Table 4.2. If we regard the data of initial experiment as the true values, the repetition will have a balanced accuracy of 85.03%. By taking into account the inconsistency in subject classifications, a machine learning model for the perception should also have a maximum possible balanced accuracy of around 85%.

4.3.2 Labeling by the female subject

TABLE 4.3: The classification of first 88 images done by the female subject repeated twice. The total number of classifications into a label for each set of readings are there in the last row and columns.

		Classification made on initial experiment		
		Flicker	Still	
Classifications in repetition	Flicker	27	15	42
	Still	14	32	46
		41	47	

The experiment was done on 1362 images of which the subject classified 547 images as flickering and 815 images as still. The median time taken per image for classification was 3.341 seconds. In order to check the consistency of subjects readings, the experiment was repeated again for the first 88 images. The results of the repeated experiment and initial classifications can be seen in Table 4.3. If we regard the data of initial experiment as the true values, the repetition will have a balanced accuracy of 66.96%. By taking into account the inconsistency in subject classifications, a machine learning model for the perception should also have a maximum possible balanced accuracy of around 67%.

4.4 Proposed Model

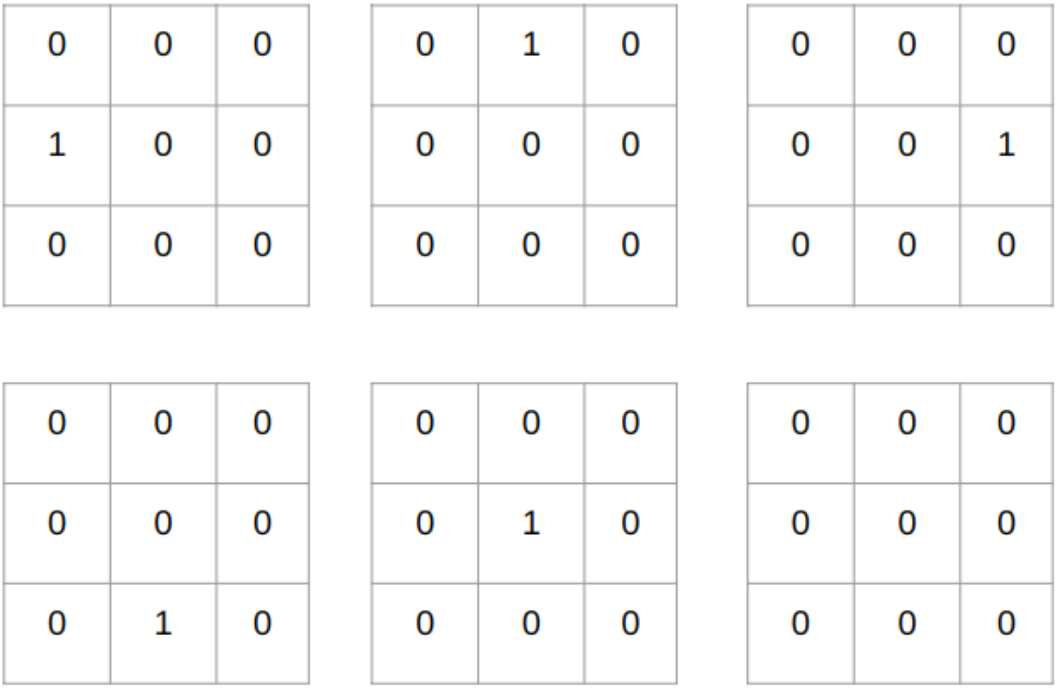


FIGURE 4.3: Filters capable of shifting the image one pixel to left (upper left filter), right (upper right filter), up (upper center filter) or down (lower left filter) as well as one that passes on the same image without doing anything (lower center filter). The bottom right filter is equivalent of a closed eye associated with blinking.

When saccadic movements occur, the retinal image abruptly shifts on the retina, the information about the edges in the object could be transmitted as the differences between successive images in the retina[97]. Schmittwilken and Maertens had proposed a model of edge detection that incorporated fixational eye movements[98]. The model mimicked eye movements by shifting images over time[98]. Convolutional filters are also capable of shifting the images over time as can be seen by filters in Figure 4.3. This work uses a DNN to create a stack of images over time from a single image. An image in this set of images could either be containing the edges in the image detected by edge filter or could have been subjected to an operation as in Figure 4.3.

We propose a neural network model that takes an image with dimensions $height_1$, $width_1$,$channels_1$ as the input. It can be transformed into another tensor with dimensions $height_2$,$width_2$,$channels_2$ by applying a convolution operation over it. The network takes images with dimensions 224*224*3 as input and apply a valid convolution filter which was generated from the image, where each frame with dimension $height_2$,$width_2$,$channels_2$ was obtained by performing a valid convolution on the original image. Using permute layer in keras we change the dimensions of this layer to $channels_2$,$height_2$,$width_2$. The layer is then reshaped to an array on images stacked

over time with dimensions $frames$,$height_3$,$width_3$,$channels_3$ adding an extra dimension $frames$ in the layer. The stack of images are again subjected to a 3D convolution filter of kernel size [6, 5, 5] with strides of 4 along height and width dimensions, to reduce the number of neurons transmitting information to upper layers and hence the number of parameters. The individual images in these stacked images after 3D convolution are all pooled and flattened. Each of these flattened images at a timepoint are fed onto a recurrent layer with a single neuron. The final state of the recurrent neuron is set as the output layer. The architecture of the proposed model with 64 frames is provided in Table 4.4.

TABLE 4.4: Network architecture for proposed model with 64 frames

Layer	Activation	Number of parameters	Dimension
Input			224, 224, 3
2D Convolution	relu	4864	220, 220, 64
Permute			64, 220, 220
Reshape			64,220, 220,1
3D Convolution	relu	151	59, 54, 54, 1
3D Average Pooling			59, 13, 13, 1
Reshape			59, 169
Dropout	70% dropped during training		59, 169
Recurrent Layer (Final State is output)	sigmoid	171	59, 1 1

4.5 Experiment

For conducting the experiment the images in both classes were randomly split into training and validation sets in the ration 7:3. From these separate mini batches of size 64 images with 32 images of each class were selected at random for each iteration for both training and validation.

4.5.1 Convolutional Neural Networks and Vision Transformers

Convolutional Neural Network(CNN)s are specialized kind of network for processing data that has a grid like topology[5]. The networks were trained with sparse categorical cross-entropy loss function with two classes in output layer and Nadam optimizer. Twelve different CNN architectures which were ResNet101V2 ,VGG16 , VGG19, ResNet50, ResNet50V2, ResNet101, ResNet152, ResNet152V2, MobileNet, MobileNetV2, DenseNet121, and NASNetMobile were used to train the data. Vision transformers are another architecture of neural networks that can be used for image classification[99]. In vision transformers, patches of images are turned to input and fed into attention layers[99]. This work used three pre-existing models of vision transformer ViT_base, ViT_small and ViT_tiny[99]. The networks were not preloaded with any weight and were trained for 8000 iterations.

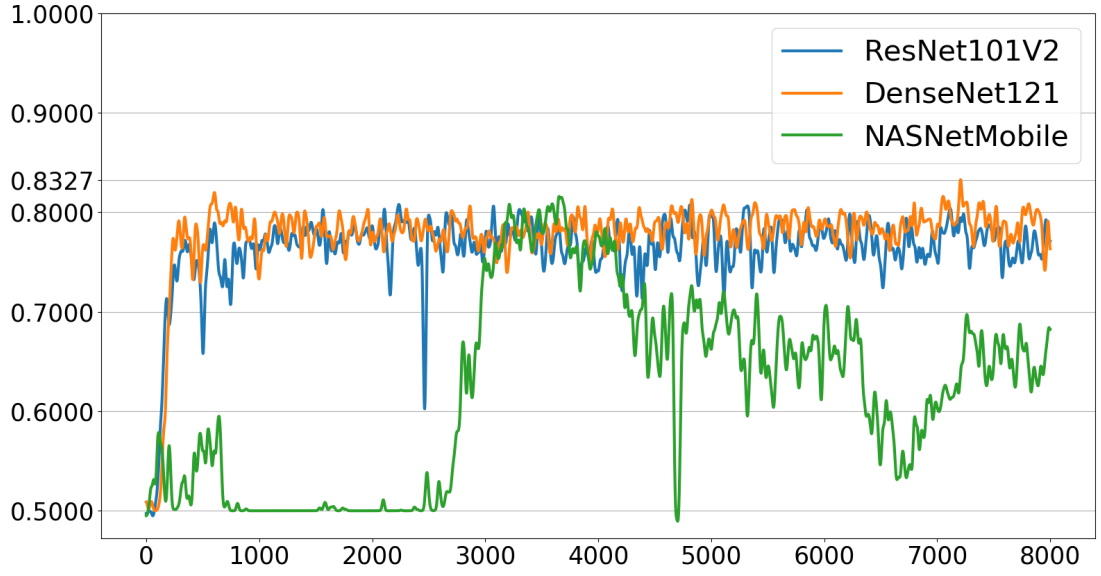
The accuracies over iterations were Gaussian filtered with a sigma value of 10. The accuracies for three CNNs over the iterations for the data from the male and female subjects are plotted in Fig. 4.4 and 4.5. For the male subject, DenseNet121 gave the highest validation accuracy of 83.33%. NASNetMobile gave the highest accuracy of 0.687% for the female subject. VGG-16 and VGG-19 could not learn from the data of both subjects and their accuracies remained at 50% throughout the training as can be seen in Tables 4.5 and 4.6. The vision transformers performed similar to CNNs.

4.5.2 Proposed architecture

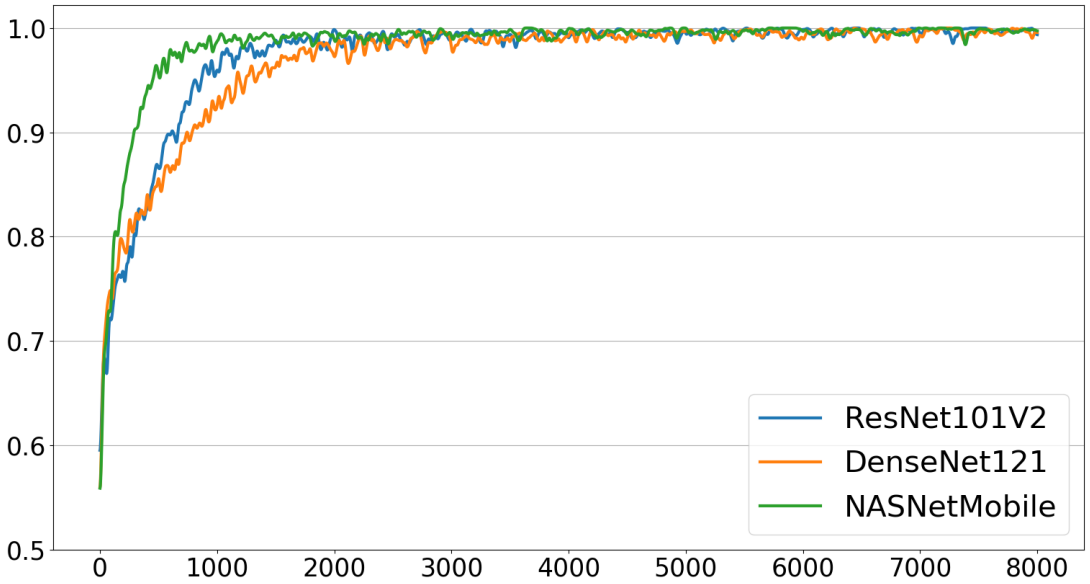
Four different versions of the proposed architecture were used to train the model in which one hyper parameter, namely the number of frames were varied as 8, 16, 32 and 64. The model was trained with loss function Mean Squared Error and Nadam optimizer. The training and validation accuracies for the proposed networks can be seen in Figures 4.6 and 4.7. The proposed model with 64 frames did the best in classifying the data for male subject. Its classification accuracy of 84.9%was close to the consistency value of 85%, that arise because of subjective inconsistencies in the response. The proposed model with 32 frames did the best in classifying the data for the female subject. Its classification accuracy of 72.76%was close to the consistency value of 67%, that arise because of subjective inconsistencies in the response. It can also be seen that the model is less prone to over-fitting. The training over the variants of proposed model that gave maximum validation accuracy for male and female subjects can be seen in Figure 4.8.

4.5.3 Comparison of results

The accuracies for validation and training dataset for the first 8000 iterations for both subjects are shown in Table 4.5 and 4.6. The CNNs used in this study are much more



(A) Validation accuracies



(B) Training accuracies

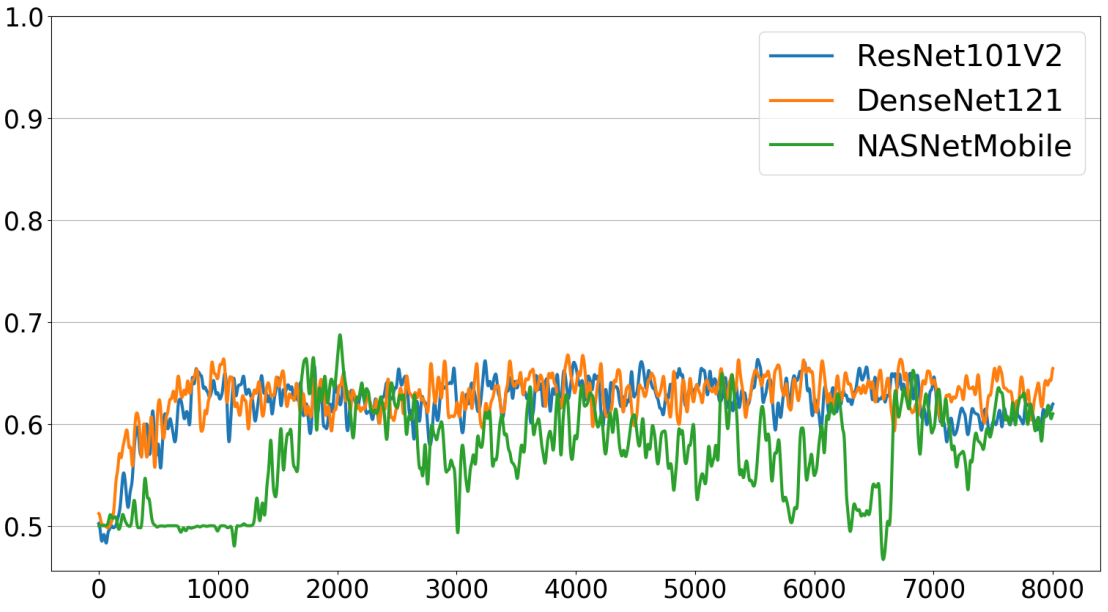
FIGURE 4.4: The accuracies for training and validation for three CNNs for the male subject. The accuracies have been Gaussian filtered with a sigma value of 10.

prone to overfitting when compared with the proposed model.

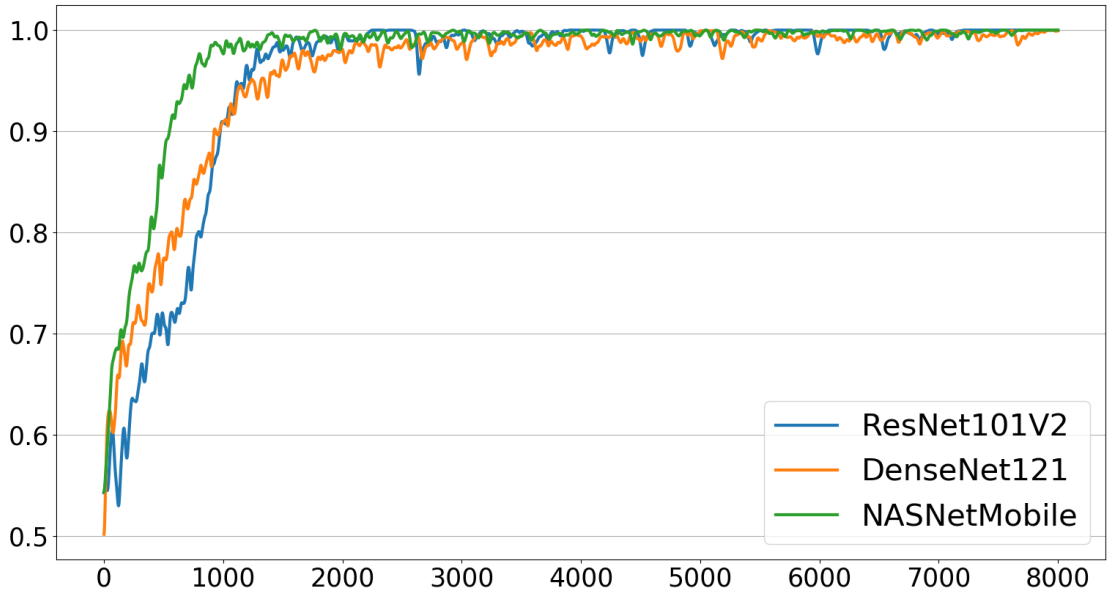
4.6 Discussion

4.6.1 Failure of VGG 16 and VGG 19

A reason for failure of VGG-16 and VGG-19 to train the data (with the same learning rates and optimizer as other CNNs), can be the lack of skip connections in them. An



(A) Validation accuracies



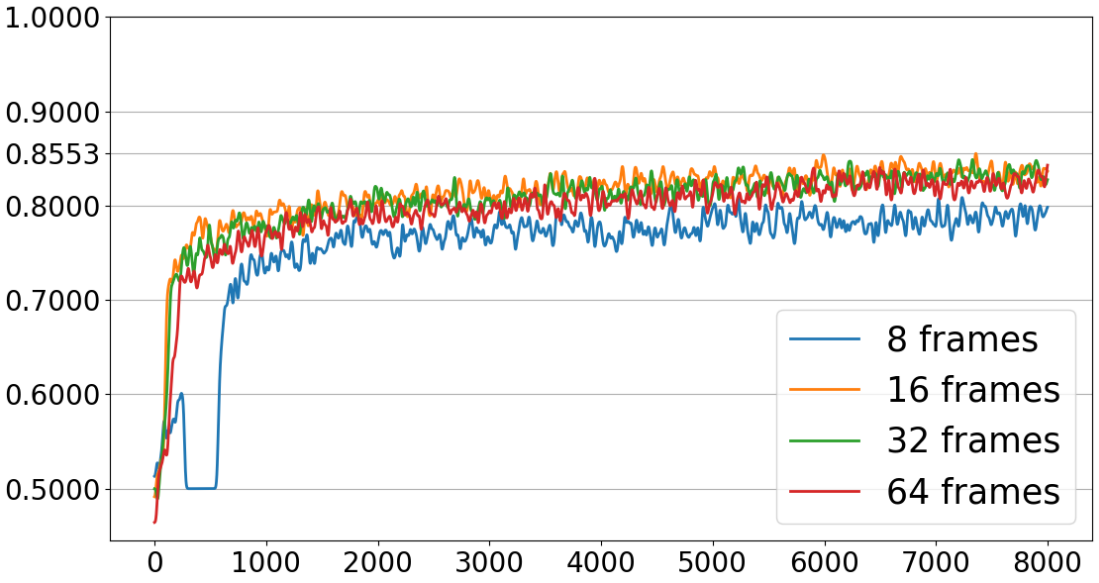
(B) Training accuracies

FIGURE 4.5: The accuracies for training and validation for three CNNs for the female subject. The accuracies have been gaussian filtered with a sigma value of 10.

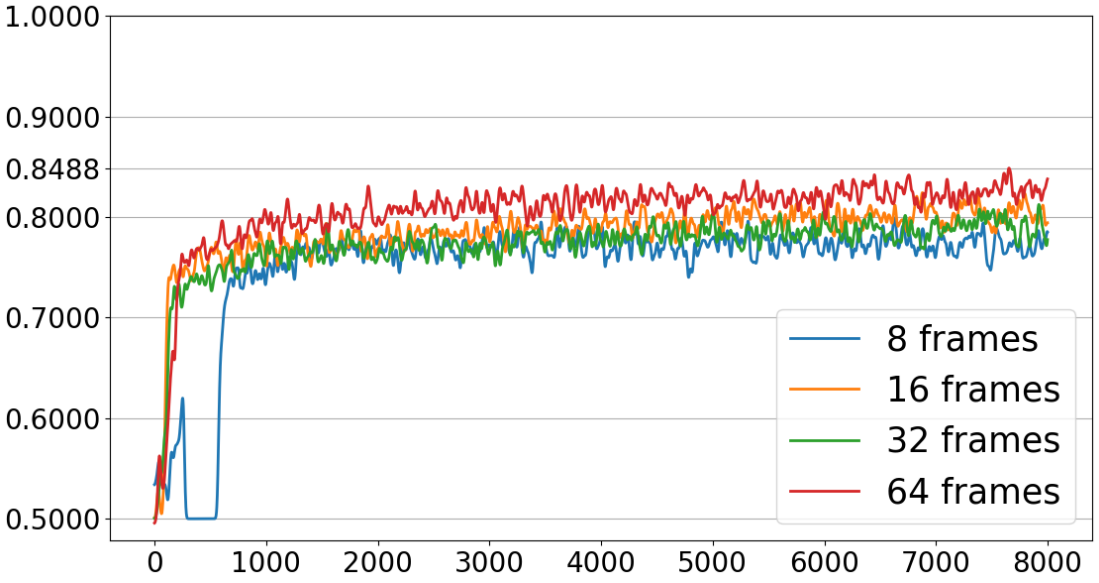
RNN can be unrolled over time to a feedforward neural network. In the proposed DNN , the initial input to the RNN at all timepoints is from a single static image. The unrolled RNN will also have skip connection.

4.6.2 Limitations

The study was done on the variant of a particular motion illusion called flicker wheel illusion. The labeling by subject only involved a binary classification. The strength of

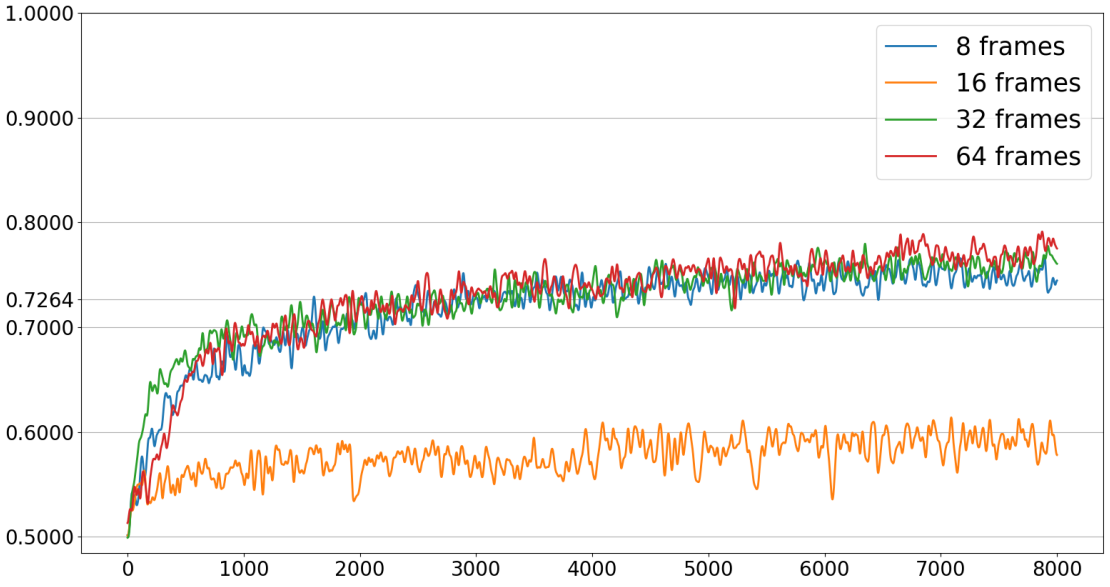


(A) Training accuracies

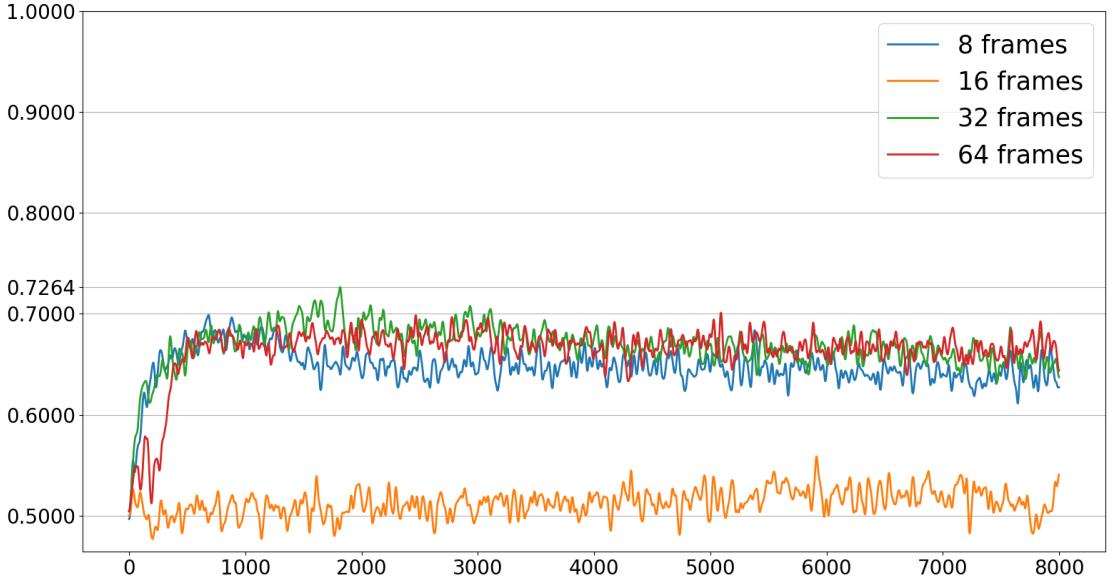


(B) Validation accuracies

FIGURE 4.6: The accuracies for training and validation for the proposed model with variation in the hyperparameter, viz. number of frames for the male subject. The accuracies have been Gaussian filtered with a sigma value of 10.

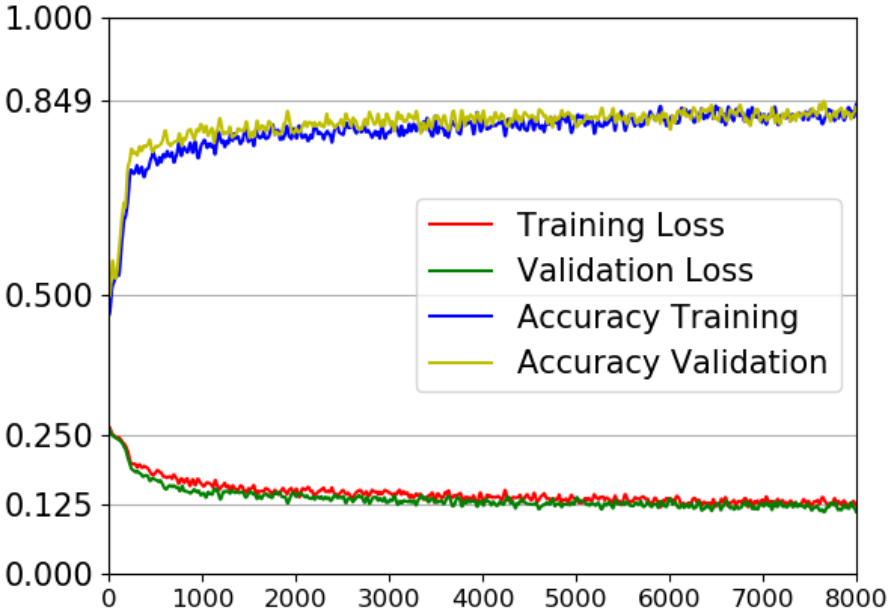


(A) Training accuracies

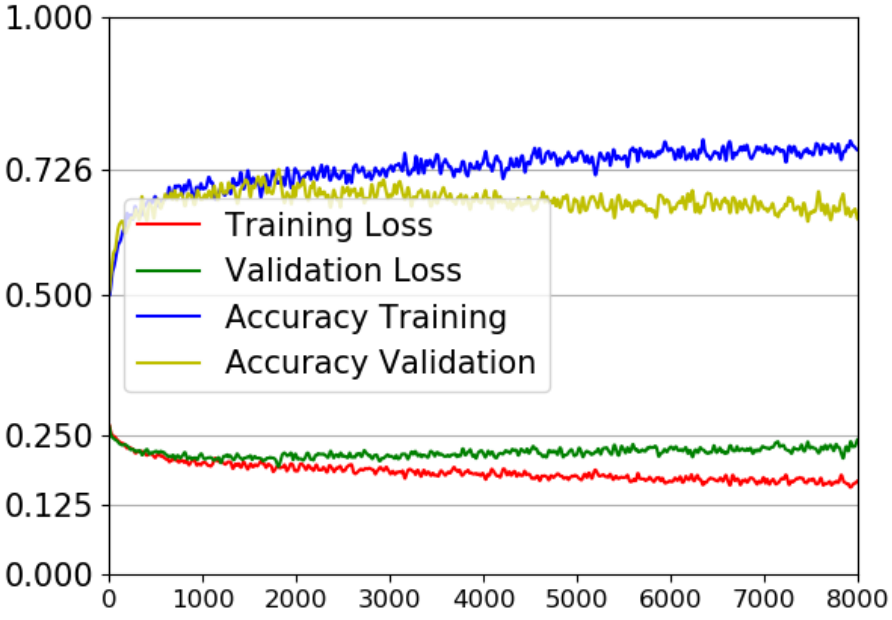


(B) Validation accuracies

FIGURE 4.7: The accuracies for training and validation for the proposed model with variation in the hyperparameter, viz. number of frames for the female subject. The accuracies have been Gaussian filtered with a sigma value of 10.



(A) For male subject



(B) For female subject

FIGURE 4.8: The losses and accuracy for proposed model for 64 frames for male subject and 32 frames for the female subject. The numerical values have been filtered with a Gaussian filter of sigma value 10.

TABLE 4.5: The maximum training and validation accuracies over the first 8000 iterations over training filtered with a Gaussian filter of sigma 10 for the psychophysics data of the male subject.

Model	Validation	Training
Proposed Model 8 frames	0.799	0.809
Proposed Model 16 frames	0.824	0.855
Proposed Model 32 frames	0.813	0.849
Proposed Model 64 frames	**0.849**	0.843
ResNet101V2	0.808	1.0
VGG16	**0.5**	0.497
VGG19	**0.5**	0.495
ResNet50	0.78	1.0
ResNet50V2	0.805	1.0
ResNet101	0.781	1.0
ResNet152	0.792	1.0
ResNet152V2	0.778	1.0
MobileNet	0.796	1.0
MobileNetV2	0.783	1.0
DenseNet121	0.833	1.0
NASNetMobile	0.816	1.0
ViT_Tiny	0.756	1.0
ViT_Small	0.759	1.0
ViT_Base	0.764	1.0

the illusion can be another measurable parameter.

4.6.3 Future Work

The psychophysics work involved in this paper only involved binary classification into flicker and still images and a subject viewed most images at most once. There could strong flicker illusory images that appears to flicker no matter how many times one looks at them as well as weak flicker illusory images that sometimes appear to flicker and sometimes not. There could similarly be strong and weak still images. These possibilities were not considered for training and testing of the present model. A test for the model with this strategy will be to see if the output of the neural network for an image is correlated with the strong or weak flicker/still category. The DNN output of images with a weak still or weak flicker could be closer to 0.5 than being near 0 or 1 for strong flicker/still category images.

The present work models a very simple kind of motion which is flicker. In future, attempt should be made to explain more complex stimuli including other illusions, that

TABLE 4.6: The maximum training and validation accuracies over the first 8000 iterations over training filtered with a gaussian filter of sigma 10 for the pschophysics data of the female subject.

Model	Validation	Training
Proposed Model 8 frames	0.699	0.767
Proposed Model 16 frames	0.559	0.614
Proposed Model 32 frames	**0.726**	0.78
Proposed Model 64 frames	0.701	0.791
ResNet101V2	0.663	1.0
VGG16	**0.5**	0.497
VGG19	**0.5**	0.498
ResNet50	0.671	1.0
ResNet50V2	0.672	1.0
ResNet101	0.66	1.0
ResNet152	0.647	1.0
ResNet152V2	0.7	1.0
MobileNet	0.671	1.0
MobileNetV2	0.685	1.0
DenseNet121	0.668	1.0
NASNetMobile	0.687	1.0
ViT_Tiny	0.66	1.0
ViT_Small	0.66	1.0
ViT_Base	0.651	1.0

evoke the perception of motion. The methodology used in this work could be replicated for other illusions by varying the physical parameters in the illusions and obtaining psychophysics data to train neural networks. The model might also be used to generate more visual illusions or manipulate an image to flicker by using it with generative neural networks. Similar works have been done on static illusions[100]. Microsaccades and blinks are necessary for perception of the visual illusions like Enigma illusion and Rotating Snakes illusion. The present work has used DNN to create a model for the effects of eye movement over time on the image. The pretrained lower layers of the proposed DNN may prove useful for development of DNN models for more complex illusions via transfer learning. The labeled dataset allows researchers who work on more complex visual illusions to get weights associated with lower layers for the model. With computer vision techniques beginning to get used for modeling biological vision, the proposed methodology and data set may prove useful for vision research.

4.7 Conclusion

This work shows that its possible for deep neural networks to learn to classify images labeled on the basis of psychophysics data, the same way a human subject does. The maximum accuracies obtained by the proposed network is close to the accuracy if the labeling is again done by the same subject. The low number of subjects used in this study is a limitation of the current work. This can be improved in future works. The work has also proposed a model where a time invariant image is taken as an input, and is transformed to a series of images, or edge information stacked over time. The time stack images could then classified using a recurrent neural network.

Chapter 5

Distinguishing between words for sharp and round objects in natural languages

Publications

- Keerthi S Chandran, Sreeja Pal, and Kuntal Ghosh. "Deep Learning as tool to distinguish words for sharp and round objects in natural languages". In: *2024 6th International Conference on Natural Language Processing (ICNLP)*. 2024, pp. 372–376. doi: http://doi.org/10.1109/ICNLP60986.2024.10692703

Chapter summary:

Sound Symbolism is a well studied psychological phenomena of the relation between sound and meaning in natural language. Although the phenomenon has been studied by psychologists and linguists, the phenomenon has not been modeled by machine learning. In this chapter, we select words for round and sharp objects from various natural languages. We tried to see if a machine learning algorithm could perform better than chance in distinguishing words for round and sharp objects in natural languages. We performed a psychophysics experiment to see if human subjects will associate words for sharp objects with a round object and a round object with a sharp figure. We show that human subjects are more likely than chance to associate words for sharp objects with sharp figure and vice versa. We propose that algorithms can be improved by using training sets consisting of words whose sound symbolic properties are labeled by psychophysics experiments.

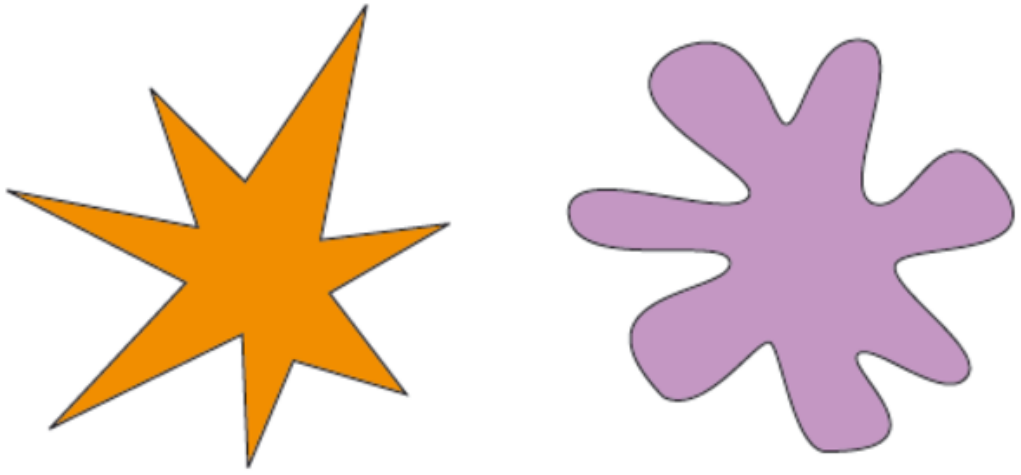


FIGURE 5.1: The Kiki(left) and Bouba(right) images used in the experiment. From Wikimedia Commons[20].

5.1 Introduction

Sound Symbolism refers to non-arbitrary links between random speech sounds and other meanings[21]. A well-known illustration of this phenomenon is the relationship between pseudo-words and abstract visual forms[101]. Köhler describes a well-known demonstration, the maluma-takete effect[102]. The round-sounding pseudo word maluma fits better to describe an abstract circular figure in the maluma-takete example, but the 'sharp'-sounding pseudo word takate fits better to describe an abstract edgy figure[103]. Combinations of some speech sounds are better suited to communicate the visual quality of a round or an angular shape in this sensible sound-shape mapping[103]. The 'Maluma-takete' mapping, and a similar 'bouba-kiki' mapping have been recorded by native speakers of various languages and cultures[23]. The nonce word bouba is associated with a round shape, while the word kiki is associated with a spiky shape[23]. This association between speech sounds and visual characteristics has the potential to have significant effects on the evolution of spoken language that we are habituated to use in everyday life[104, 105, 106, 107, 108]. That is why experiments have been performed in this work with the same shapes for different spoken languages. The strength of the effect across cultures and the impact of orthography, however, are up for debate[109]. An online study that examined the Kiki-Bouba effect among 25 speakers of various languages and nine different language families and ten different writing systems[109]. Bouba elicited more congruent responses than kiki, and overall strong evidence for the effect was found across languages[109]. A study of the orthographic shapes of the words in various scripts revealed that the effect was not stronger for scripts that use rounder forms for bouba and spikier forms for kiki[109]. In the study, participants who spoke languages with Roman scripts were only slightly more likely to exhibit the effect[109].

The results provided the strongest evidence to date that the Kiki-Bouba effect is robust across cultures and writing systems[109]. It was also confirmed that the Kiki-Bouba phenomenon is rooted in a cross-modal correspondence between aspects of the voice and visual shape, which is largely independent of orthography[109].

Numerous explanations have been put up to explain how the articulatory and auditory characteristics of phonemes (From the Kiki-Bouba effect, phonemes, like /b/, /o/, and /u/, to roundness and others, like /k/ and /i/, to sharpness) come to be connected to other inputs[110]. The co-occurrence of phoneme characteristics and associated stimuli in the real world may be the cause of the relationships[110]. Another idea is that phoneme characteristics and the stimuli they are associated with some sort of perceptual, intellectual, emotive, or linguistic trait in common[110]. The links might also result from aspects of the brain's architecture, evolution, or patterns gleaned from language[110]. Although there is a great deal of experimental data supporting the effects of symbolic sound associations, the methods by which they might be produced have received much less attention[110].

Machine learning is the study of intelligence for the design and development of algorithms that can learn from data, to learn from experience, and to improve their learning behavior over time[111]. Finding meaningful structural and/or temporal patterns (knowledge) in data can be difficult because they are frequently buried in arbitrarily high dimensional spaces that are inaccessible to humans[111]. Deep Learning which is a machine learning algorithms has the advantage of producing new features from limited set of features present in the dataset[112]. Given that no knowledge of features involved are needed to model the phenomena, deep learning methodologies are an ideal tool for machines to learn sound symbolism. This study took the inspiration from the study done by Sidhu et al. where, using a large sample of words, they examined whether the maluma/takete effect is attested in English language[113]. They discovered evidence that words referring to round objects tend to have more phonemes associated with roundness, and words referring to spiky objects tend to have more phonemes associated with spikiness which illustrates the fact that human language exhibits iconicity, which thus makes it non-arbitrary[113]. Human subjects are able to guess the correct word from antonyms like big/small, and round/pointy from unknown languages better than chance[22]. In the present study, words for naturally sharp objects such as- "Axe", "Blade", "Pin", "Scissor", etc. and round objects – such as "Apple", "Ball", "Bubble", "Coins" etc. - were taken from 22 languages and put into study using Machine Learning algorithms. We also performed a psychophysics experiment in which human subjects were asked to match sounds for sharp and round objects from foreign languages to the sharp and round figures used in Kiki-Bouba experiment.

5.2 Materials and Methods

5.2.1 Materials

A list of objects which could be sharp or round were prepared. The list consisted of 35 words for sharp ones like scissors, spears, knife, etc. and 49 round objects like sun, moon, lemon, oval, etc. Complex linguistic terms resembling words such as "Golf Ball", "Coconut", "Sea Urchin" and "Porcupine" were omitted from the list [8]. Sounds for the words were downloaded in 26 languages from Collins Dictionary[9], leading to a collection of 1913 sound files (1172 files for words resembling roundness and 741 files for words resembling sharpness). The languages include American-English, Arabic, Brazilian-Portuguese, British-English, Chinese, Croatian, Czech, Danish, Dutch, European-Portuguese, European-Spanish, Finnish, French, German, Greek, Italian, Japanese, Korean, Latin American-Spanish, Norwegian, Polish, Russian, Swedish, Thai, Turkish, and Vietnamese. Given that the languages, such as English—American and British; Spanish—European and Latin American; and Portuguese—European and Brazilian, were basically the same language pronounced in different accents by different speakers, only one of the languages was utilized during training. For instance, of the two available varieties of English—British and American—only, American English was used during training. Only one language among Portuguese and Spanish were considered as the words were similar. The sound files for all 26 languages could be obtained for 25 sharp objects and 43 round objects from Collins dictionary website. In the remaining ten sharp objects and six round objects, the words could not be obtained in all languages. The sounds has a sampling frequency of 44100 Hz with 1 channel.

5.2.2 Psychophysics

The psychophysics experiment was done on 199 subjects from Indian state of West Bengal. A selection of 35 words for round and 35 words for sharp objects were chosen from the list of objects whose sounds could not be obtained for all languages. English language words were excluded from the psychophysics experiment as the subjects were aware of the language. The subjects were shown the program seen in Figure 5.2 on a laptop screen. The subjects could play a sound by pressing the blue play button. The sounds could be played more than once. The subjects were instructed to click the figure they felt corresponded to the sound. The clicks of the subjects were recorded and the subject could proceed onto listening to a new sound. All the seventy sounds were presented to the subjects in a random order. The aim of the psychophysics experiment was to check if there exists any phonetic property in words for sharp and round objects in natural languages, that will cause human subjects to match the word for a sharp or round object in an unknown natural language to the sharp and round figures used in the Kiki-Bouba experiment.

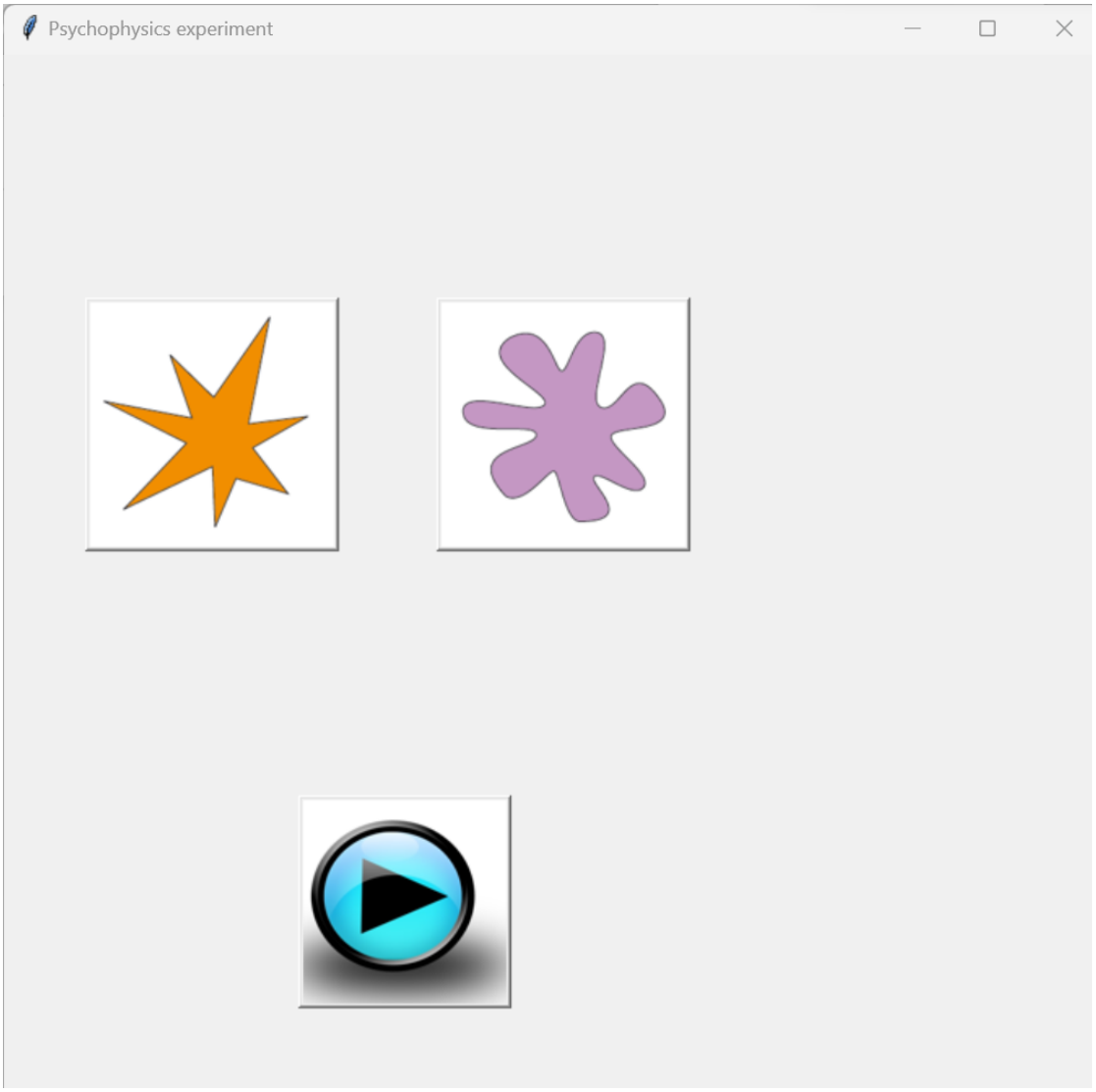


FIGURE 5.2: Screenshot of program displayed to the subjects during psychophysics experiment

5.2.3 Machine Learning

The machine learning was done using 25 words for sharp objects and 43 words for round objects in 22 languages. The words chosen for training and validating the neural network were not used in the psychophysics experiment. The machine learning architecture took inspiration from the previous work by Passi and Arun, which noted that the Kiki Bouba effect was correlated with mean frequency of the sounds. The input vector corresponding to a sound file was made as follows.

i The array was subjected to fast Fourier transform and the amplitudes were converted to their absolute value.

ii The absolute value of amplitudes from 0 to 15000 Hz in increments of 1 Hz were calculated by linear interpolation.

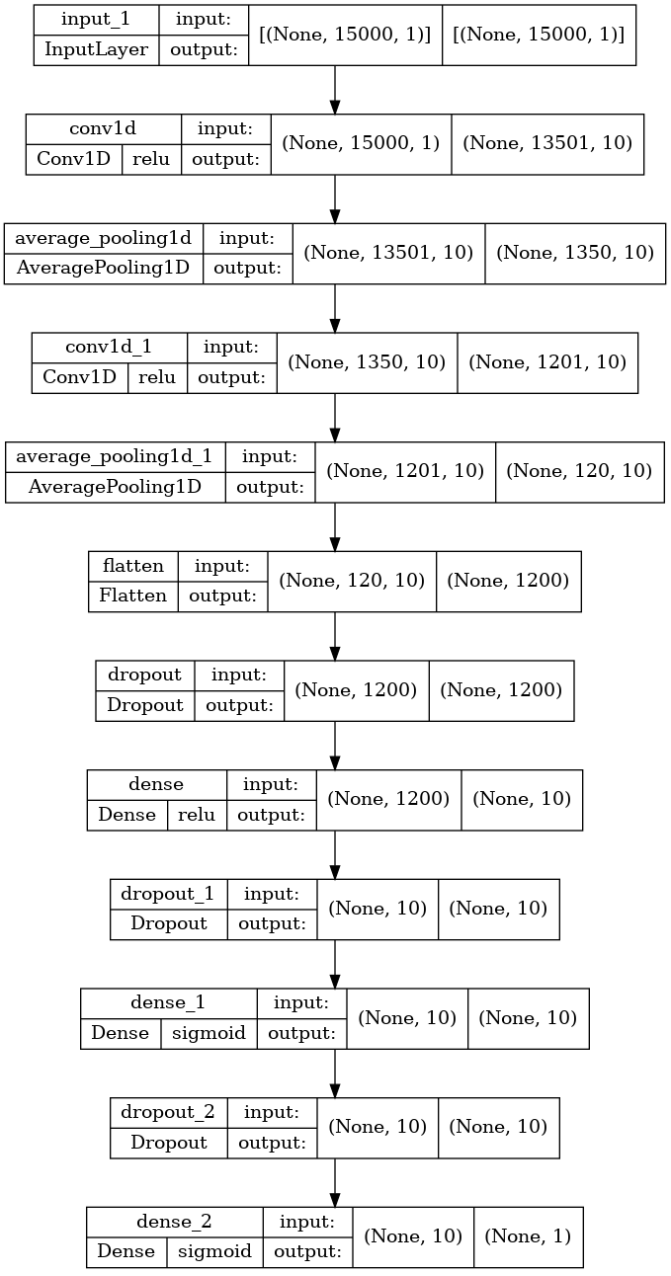


FIGURE 5.3: The model of the DNN used for classification

iii A new vector was dividing the previous vector by its mean.

iv This new normalized vector served as input vector to the machine learning model.

The sound files were used to train the DNN shown in Figure5.3. The DNN used mean square error as loss function with SGD optimizer for training. The model was subjected to Monte Carlo cross validation with 25 repetitions. The mini batch for training and validation of the 25 words for sharp and 43 words for round. Of this a fixed number of words from each class ranging from 10 to 15 were were randomly chosen out for validation in each repetition of the cross validation process. The words in all languages were used

TABLE 5.1: The words labeled round by most subjects(n=199)

Object Classification	Word (language)	Written word	Sharp labeling	p-value
round	Foam (German)	Schaum	63	$2.467 * 10^{-07}$
round	Sphere (Italian)	sfera	79	$4.455 * 10^{-3}$
round	Bloom (French)	fleur	74	$3.681 * 10^{-4}$
round	Droplet (French)	gouttelette	78	$2.815 * 10^{-3}$
round	Ripple (German)	kleine Welle	81	$01.053 * 10^{-2}$
round	Bloom (European Portuguese)	flor	65	$1.131 * 10^{-06}$
round	Foam (Brazilian Portuguese)	espuma	70	$3.477 * 10^{-05}$
round	Droplet (Brazilian Portuguese)	gotícula	73	$2.105 * 10^{-4}$
round	Bloom (German)	Blüte	72	$1.179 * 10^{-4}$
round	Sphere (European Portuguese)	esfera	75	$6.308 * 10^{-4}$
round	Bloom (Italian)	fiore	73	$2.105 * 10^{-4}$
round	Foam (European Portuguese)	espuma	80	$6.915 * 10^{-3}$
round	Sphere (German)	Kugel	83	$2.306 * 10^{-2}$
round	Bloom (Brazilian Portuguese)	flor	62	$1.114 * 10^{-07}$
round	Bloom (Latin American Spanish)	floración	71	$6.469 * 10^{-05}$
round	Ripples (European Portuguese)	ondulação	74	$3.682 * 10^{-4}$

TABLE 5.2: The words labeled sharp by most subjects(n=199)

Object Classification	Word (language)	Written word	Sharp labeling	p-value
sharp	Bayonet(French)	baïonnette	123	$01.059 * 10^{-3}$
sharp	Shears(Italian)	cesoie	115	$3.319 * 10^{-2}$
sharp	Tusk(Brazilian Portuguese)	presa	114	$4.689 * 10^{-2}$
round	Droplet(European Spanish)	gotita	115	$03.319 * 10-2$
sharp	Icicles(European Spanish)	carámbano	132	$4.751 * 10-6$
sharp	Bayonet(European Spanish)	bayoneta	119	$06.915 * 10^{-3}$

for training and validation. The mini batches had 20 sounds for both round and sharp objects for training, 80 sounds for both sharp and round objects for validation. The list of words and languages used for training and validation is in Appendix A.

5.3 Results

5.3.1 Psychophysics

The psychophysics experiment was done on 199 subjects aged between 9 and 54. Of this 172 subjects were school students between ages of 11 and 15. Of the set of 35 words for round objects and 35 words for sharp objects, the number of times each word got labeled to the sharp figure instead of the round figure was tabulated. The full list of tabulations is provided in Appendix A. The t test for the number of labeling for words in both the tests were calculated. A t test statistic of 5.958 with p-value of $1.007 * 10^{-7}$

TABLE 5.3: The results of cross validation of the model. The values were filtered with a gaussian filter of sigma 5 before calculating the maximum and minimum values.

Iteration	Validation words for a class	Training Loss (lowest)	Training Accuracy (highest)	Validation Loss (lowest)	Validation Accuracy (lowest)	Validation Accuracy (highest)
0	8	0.241	0.593	0.23	0.5	0.676
1	12	0.216	0.672	0.237	0.5	0.609
2	11	0.243	0.574	0.242	0.492	0.647
3	9	0.236	0.627	0.238	0.519	0.647
4	8	0.224	0.647	0.25	0.446	0.539
5	9	0.22	0.672	0.239	0.5	0.602
6	8	0.238	0.601	0.238	0.5	0.632
7	12	0.221	0.655	0.231	0.517	0.629
8	8	0.226	0.635	0.242	0.496	0.586
9	11	0.218	0.672	0.24	0.5	0.6
10	8	0.212	0.688	0.253	0.434	0.503
11	10	0.228	0.651	0.242	0.499	0.603
12	9	0.23	0.651	0.244	0.49	0.577
13	10	0.217	0.649	0.235	0.5	0.626
14	10	0.223	0.653	0.24	0.5	0.623
15	9	0.226	0.645	0.242	0.5	0.599
16	12	0.19	0.724	0.231	0.496	0.634
17	9	0.23	0.638	0.233	0.499	0.62
18	8	0.225	0.66	0.24	0.5	0.624
19	11	0.222	0.657	0.234	0.499	0.62
20	11	0.18	0.755	0.236	0.516	0.628
21	9	0.224	0.664	0.239	0.5	0.613
22	10	0.197	0.74	0.24	0.499	0.607
23	10	0.227	0.661	0.243	0.497	0.579
24	9	0.216	0.69	0.242	0.5	0.588
Mean	9.64	0.221	0.659	0.239	0.496	0.608
Median	9.0	0.224	0.655	0.24	0.5	0.613

was obtained. A subject was able to assign sounds to figures with an accuracy as high as 85.7%. The words so labeled were all from the category of chosen round objects. The words labeled mostly as sharp and round can be seen in Tables 5.1 and 5.2 and these words were matched to either of the figures better than by chance.

5.3.2 Machine Learning

The cross validation was done by Monte Carlo Cross validation method. The dataset was trained separately 25 times and the results of training with maximum and minimum accuracies filtered over iterations with a Gaussian filter of sigma 5 can be seen in Table 5.3. It can be seen that the proposed method yields an accuracy higher than chance of $\sim$ 60%. The input data which was not divided by the mean of values did not fit the network. The machine learning algorithm was tested with the 70 words used in

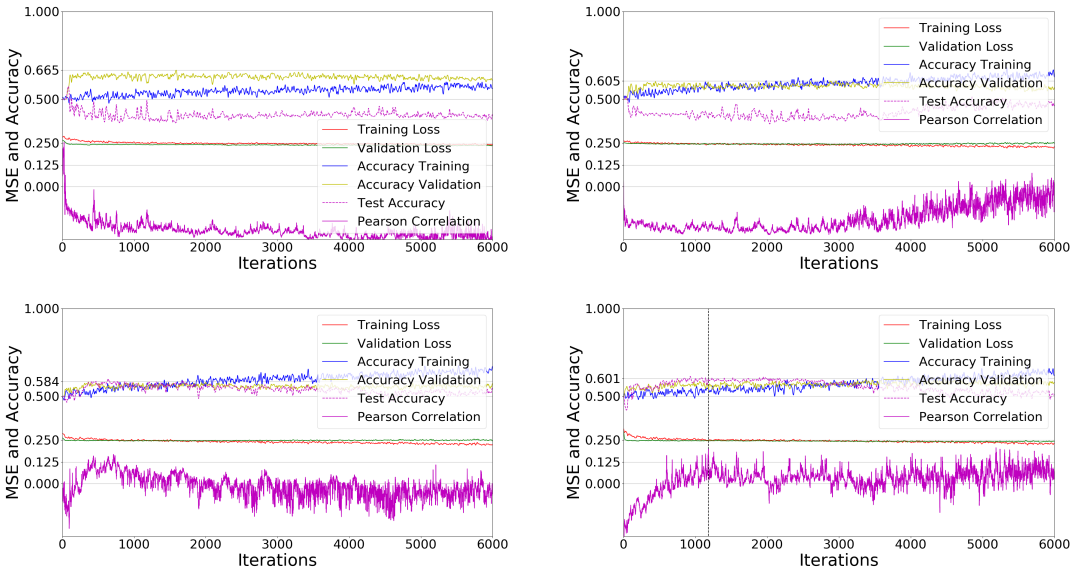
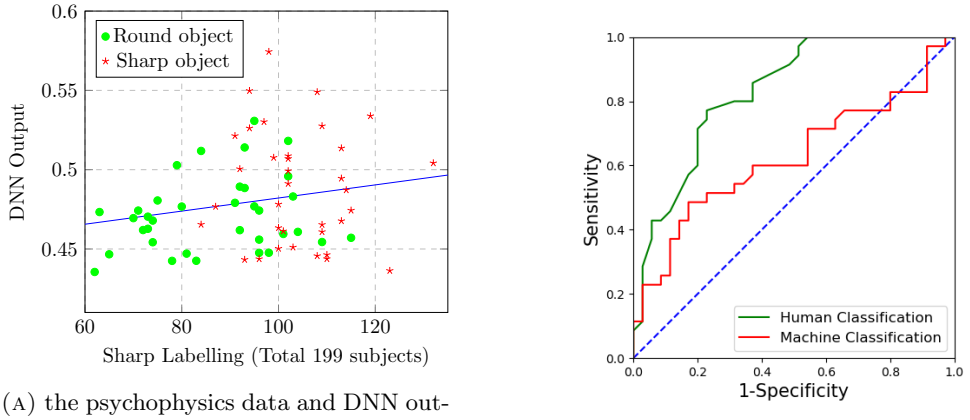


FIGURE 5.4: Four trainings done with the words used in psychophysics experiments used as test data. The losses and accuracies have been Gaussian filtered with a sigma of 5. The Person correlation between the output for the test set as well as number of sharp labeling made by the subjects have also been plotted. The weights corresponding to a selected iteration has been marked by the black vertical line.

psychophysics experiments. The training curves for five different trainings are plotted in 5.3.2. The classification by the machine corresponding to the weights that gave highest Pearson Correlation with psychophysics readings on the same word has been plotted in Figure 5.5.



(A) the psychophysics data and DNN output

(B) ROC

FIGURE 5.5: The ROC curve for the human classifcation based on probability of sharp labellig by humans and a machine classifcation that gave similar classification

5.4 Conclusion

The psychophysics experiments shows that humans have a tendency to associate words for sharp and objects in natural languages to sharp and round figures used in the Kiki Bouba experiment. The experiment also show that there could be outliers in classifications made by human subjects. Many words will also be labeled randomly by people with no affinity to either round or sharp figures. The machine learning algorithm was able to distinguish between the round and sharp words better than chance. Many words in the dataset could be assigned by a human to sharp or round figure independent of the fact whether the corresponding object is sharp or round. This problem can be overcome in future by performing psychophysics test on sounds in the dataset with multiple human subjects. The sounds that get labeled to a particular image better than chance, can be used to train and test models. In future, better algorithms to make use of the phenomena of sound symbolism for natural language processing could be developed via machine learning algorithms.

Chapter 6

A deep learning based cognitive model to probe the relation between psychophysics and electrophysiology of flicker stimulus

Publications

- Keerthi S Chandran and Kuntal Ghosh. "Recurrent Convolutional Neural Networks trained by psychophysics data can predict EEG response to flicker". In: *43rd European Conference on Visual Perception (ECVP) 2021 Online*. Vol. 50. 1 suppl. SAGE Publications, Dec. 2021, pp. 132. doi: http://doi.org/10.1177/03010066211059887
- Keerthi S. Chandran and Kuntal Ghosh. "A deep learning based cognitive model to probe the relation between psychophysics and electrophysiology of flicker stimulus". In: *Brain Informatics* 11.1 (July 2024). issn: 2198-4026. doi: http://doi.org/10.1186/s40708-024-00231-0

Chapter summary:

The flicker stimulus is a visual stimulus of intermittent illumination. A flicker stimulus can appear flickering or steady to a human subject, depending on the physical parameters associated with the stimulus. When the flickering light appears steady, flicker fusion is said to have occurred. This work aims

to bridge the gap between the psychophysics of flicker fusion and the electrophysiology associated with flicker stimulus, through a Deep Learning based computational model of flicker perception. Convolutional Recurrent Neural Networks (CRNNs) were trained with psychophysics data of flicker stimulus obtained from a human subject. We claim that many of the reported features of electrophysiology of the flicker stimulus, including the presence of the fundamental and harmonics of the stimulus, can be explained as the result of a temporal convolution operation on the flicker stimulus. We further show that the convolution layer output of a CRNN trained with psychophysics data is more responsive to specific frequencies as in human EEG response to flicker, and the convolution layer of a trained CRNN can give a nearly sinusoidal output for 10 hertz flicker stimulus as reported for some human subjects.

6.1 Introduction

Human visual psychophysics is a field of research that employs specialized methods generating several established findings[115]. The flicker stimulus, which is a stimulus with intermittent illumination is one such stimulus used in visual psychophysics. Under a given circumstance, a flicker stimulus may appear steady or flickering to a human subject depending on a number of parameters like size, shape, luminance, color composition, and the temporal waveform of the stimulus[57]. When a flickering stimulus no longer appears flickering but appears steady, the flicker is said to be fused, or flicker fusion has occurred. The psychophysics of flicker perception has been studied for the last two and a half centuries[60]. A related domain in the study of flicker perception is the electrophysiological signals of the cortical activities induced by the flicker stimulus. The electrophysiological signals can be the scalp EEG of humans or invasive recordings from the cortex of animals like cats[65][85]. Although psychophysics and EEG use completely different methodologies to explain the processing of signals within the brain, there have not been many attempts to integrate, combine or compliment these two behavioral (macro) and electrophysiological (micro) viewpoints in order to better understand brain signal processing. So far, there does not seem to be much of relationship between the electrophysiology and psychophysics of flicker stimuli[86]. The present work attempts to bridge this gap between the two domains through a DNN based computational model of flicker fusion that provides an explanation for some important features of the electrophysiological response to flicker stimuli. We have shown that the intermediate layers of the DNN may show features of the electrophysiological response to the stimulus.

6.1.1 Electrophysiology of flicker perception

The electrophysiological response to flicker stimuli at areas 17 and 18 of the cat cortex has been recorded as Multi Unit Activities(MUAs) and Local Field Potentials (LFP) by Rager and Singer[85]. The anesthetized cat's retinas were exposed to flicker stimuli in screen in front of them[85]. Both LFPs and MUAs, thus recorded for a flicker stimulus of a particular frequency, show the fundamental amplitude associated with the flicker stimulus as well as its harmonics[85]. In contrast to multiunit activities components of lower frequencies were present in the LFPs though there is no evidence that they could be subharmonics[85]. Herrmann did a study with ten subjects in which the subjects were exposed to flicker stimulus using specially designed goggles, and their EEG responses to flicker were measured. The diagrammatic representation of the light pulse by Herrmann showed square waves. The human EEG response to flicker stimulus with ten subjects shows the fundamental, harmonics as well as the first subharmonic of the flicker stimulus[65]. The EEG response also shows resonances for stimuli around 10, 20, 40 and 80 hertz[65]. The oscillations are evoked up to 90 hertz and are evoked even when there is no conscious perception of flicker[65]. The average of the fundamental frequency of the evoked EEG response in ten subjects showed a clear resonance for 10 hertz flicker stimulus[65]. Tweel had detected a clear sinusoidal response near 10 hertz stimulus in some subjects[86]. In another study with ten subjects that used both square and sinusoidal flicker, the subharmonic frequencies were hardly detectable in the occipital electrodes but were instead detected in the parietal electrodes[74]. In the same study, subharmonics were detected only in eight of the ten subjects[74].

6.1.2 Prevailing explanations for EEG response to flicker stimulus

There are two hypotheses on the origin of Steady State Visually Evoked Potentials (SSVEPs) to flicker stimulus: the entrainment of brain oscillators and the superposition of Event Related Potentials(ERPs)[116, 117, 118]. It has been assumed by many authors that the harmonics and subharmonics in electrophysiological response to flicker are generated by nonlinearities in the visual system[74]. The source of these nonlinearities has never been demonstrated conclusively[74]. Previous modeling of EEG response to human stimulus involved feeding periodic signals to models of cortex like the corticothalamic model and the neural mass model[119, 74, 120, 121]. Labecki et al. have shown that harmonics and subharmonic responses can be generated by feeding the representation of flicker stimulus to a neural mass model with inhibitory and excitatory neurons[74].

6.1.3 Aim of the current work

This work aims to connect the relation between psychophysics and electrophysiology of a human subject through a Deep Neural Network (DNN) based computational model

that will be trained with psychophysics data of flicker fusion of a human subject, and no characteristics of EEG data obtained from human subjects will be taken into consideration while training the network. While the network will be trained by backpropagation, it does not imply that any equivalent computational mechanism in the human visual system was similarly trained by backpropagation. The biological brains are instead the product of evolution by natural selection. The flicker stimuli with controlled waveforms used in laboratories are not something that occur naturally, but are instead artificial stimuli used to understand the human visual system which is a product of biological evolution.

6.1.4 The methodology of the work

The first work to model the computational activity in the brain using Artificial Neural Networks (ANN) via backpropagation was by Zipser and Andersen[35], who modeled the neuronal activity in the parietal cortex of monkey brain, used for calculating the head-centered coordinates of the external objects from the position of their images in the retina and slope of the eyes[35]. The outputs of the internal hidden layers of an ANN trained to mimic the same task via supervised learning, gave outputs which were visually similar to the electrophysiological readings obtained from parietal neurons of the monkeys performing the same task. This was not the case with untrained ANNs with random weights[35]. The position of an external object in the world for the purpose of calculating its head-centered coordinates is static, and the task is a problem in static vision. But in experiments of flicker fusion the stimulus is kept fixed at a particular location, and the stimulus intensity varies with time. In a flicker fusion experiment, where the size, shape, and position of the stimulus are invariant with time and the only variant parameter is the intensity, the stimulus can be represented as a timeseries of intensities. Since a human subject classifies the stimulus as flickering or steady, the output can be represented as a binary classification. The input representation will be in the form of a sequence of intensities sampled over time. Recurrent neural networks, which are a family of neural networks for processing sequential data[5], will be used to model flicker phenomena. The human visual system performs low pass and band operations on flicker stimulus[80]. The mammalian retina can also generate transient on and off responses to change in luminance[81]. Receptive field based convolution filters can also detect sudden changes in intensity as is their use for edge detection[122]. Low pass and band pass operations can be mathematically achieved using a linear filter through convolution sum[123]. This work will train Convolutional Recurrent Neural Networks (CRNNs), which consist of convolution layer followed by a recurrent and dense layer[77]. The output of convolution layer of the trained CRNN could be subjectively compared to the features of the EEG response to flicker stimulus.

6.1.5 Contributions

- The work shows that the fundamental and harmonics of a flicker stimulus can be elicited by a temporal convolution operation on the stimulus.
- The convolution layer of a CRNN trained on psychophysics data will be more sensitive to particular frequencies, similar to the human EEG response to flicker.
- A pure sinusoidal output can be elicited for a 10 Hz flicker stimulus from the convolution layer of CRNN trained on psychophysics data.

6.2 Related Works

6.2.1 Entrainment

Entrainment is a phenomenon where two oscillators interact with each other to synchronize their oscillations[124]. These synchronizations can enhance or negate each other's effects as well as synergize with each other to achieve amplitudes greater than the sum of amplitudes of both the oscillators[124]. The frequency of one or both the oscillating systems can be altered so that they become phase locked or the phase difference between the oscillating systems remains invariant in time, and robust to perturbations[124]. The interaction between the oscillators can be linear or non-linear, and oscillations are merely superimposed with each other. Kirschfeld found that phase of alpha wave oscillations is not affected by flash evoked potentials, and the flash evoked potentials are superimposed on alpha waves without resetting their phase[125]. Schwab et al. found evidence of entrainment to flicker stimulus in both EEG and magnetoencephalography (MEG) with stronger frequency entrainment in MEG when compared to EEG[116]. Notbohm, Kurths, and Herrmann found evidence of entrainment in EEG response to flicker stimulus but not to the superposition of event related potentials[118]. Gulbinaite et al. showed that flicker stimulus with frequencies near the alpha oscillation can impair stimulus processing in a selective attention task[126]. MEG study by Duecker et al. found that gamma oscillations in the human brain and EEG response to flicker stimulus are evoked in different areas of the human visual cortex [66].

6.2.2 Role of attention

Flicker stimulus elicits human EEG response in two separate cortical networks depending on attention and temporal frequency[127]. The effect of attention on resonance to flicker stimulus is negative for a flicker stimulus with frequency in alpha band and positive for s stimulus with frequency in gamma band[128].

6.3 Materials and Methods

6.3.1 Materials

The psychophysics data previously published in [62], but never used in training any neural network or any other brain computational model to the best of our knowledge, was used in training our proposed model. In this work [62], the CFF was established by the method of limits. The published data was for an observer S. H. seen in Table 6.1, was used for training the neural networks. The flickers in [62] were produced on an opal glass surface, and the subject classified the stimulus as either flickering or a steady after observing it for ten seconds. The stimulus target subtended a visual angle of 2° 5' with the vertical and 4° 5' with the horizontal. The CFFs were obtained for PCFs $\frac{1}{6}$, $\frac{2}{3}$, $\frac{1}{2}$, $\frac{2}{3}$, $\frac{5}{6}$. The CFFs for five different PCFs were obtained for intensities 5340 cd/ft^2 , 534 cd/ft^2, 53.4 cd/ft^2 , 5.34 cd/ft^2 and 0.53 cd/ft^2 [62]. The CFFs can be seen in Table 6.1.

Intensity (cd/ft^2)	PCF $\frac{1}{6}$	PCF $\frac{1}{3}$	PCF $\frac{1}{2}$	PCF $\frac{2}{3}$	PCF $\frac{5}{6}$
5340	58.17	51.83	59.17	52.5	47.00
534	48.0	50.0	50.83	48.7	46.17
53.4	34.67	36.17	36.33	35.2	32.67
5.34	23.17	27.	25.67	25.2	20.83
0.53	17.17	19.0	17.83	19.0	15.67

TABLE 6.1: The CFFs for photic pulses for five different intensities and five PCFs for an observer S.H. (in *hertz*) published by Nelson, Bartley, and Harper[62]

6.3.2 Methods

6.3.2.1 Model

A photic pulse can be represented as a timeseries array of intensities. The retina can have on pathways, off pathways as well as transient on and transient off cells[81]. The on pathways are activated when there is light falling on them, and off pathways are activated when there is no light falling on them. In addition to these, there are transient on and off pathways which are activated when there is a sudden change in the intensity of light falling on them. The transient on and off operations can be mathematically represented by linear filtering operations. The machine could thus mimic the transient on and off pathways by performing a convolution operation on the signal. While the machine stores the whole representation of the photic pulse in memory while doing the convolution operation, that will not be the case in reality. Only the intensities in the immediate past will be necessary for the human visual system to perform transient on

and off operations. The output of the convolution layer are fed into a recurrent layer. Recurrent neural networks are a family of neural networks for processing sequential data[5]. The final state of the recurrent layer is fed into a Multi Layer Perceptron for classification.

Python based Keras library was used to train and run the model. Loss function sparse_categorical_crossentropy with Adam optimizer was used to train the model. Cross entropy is a loss function which is used in classification problems. The cross entropy H for a datapoint is given by equation $H(p,q) = -\sum_{x \in classes} p(x) log(q(x))$ where $p(x)$ is the true probability distribution of the classes and $q(x)$ is the probability distribution of the classes predicted by the model[129]. Sparse categorical cross entropy was used to provide labels as integers. Time series representation of flicker stimulus with desired PCF and intensity can be generated as one dimensional arrays. These could be assigned the labels flickering or fused based on psychophysics data.

6.3.2.2 Minibatch Generation

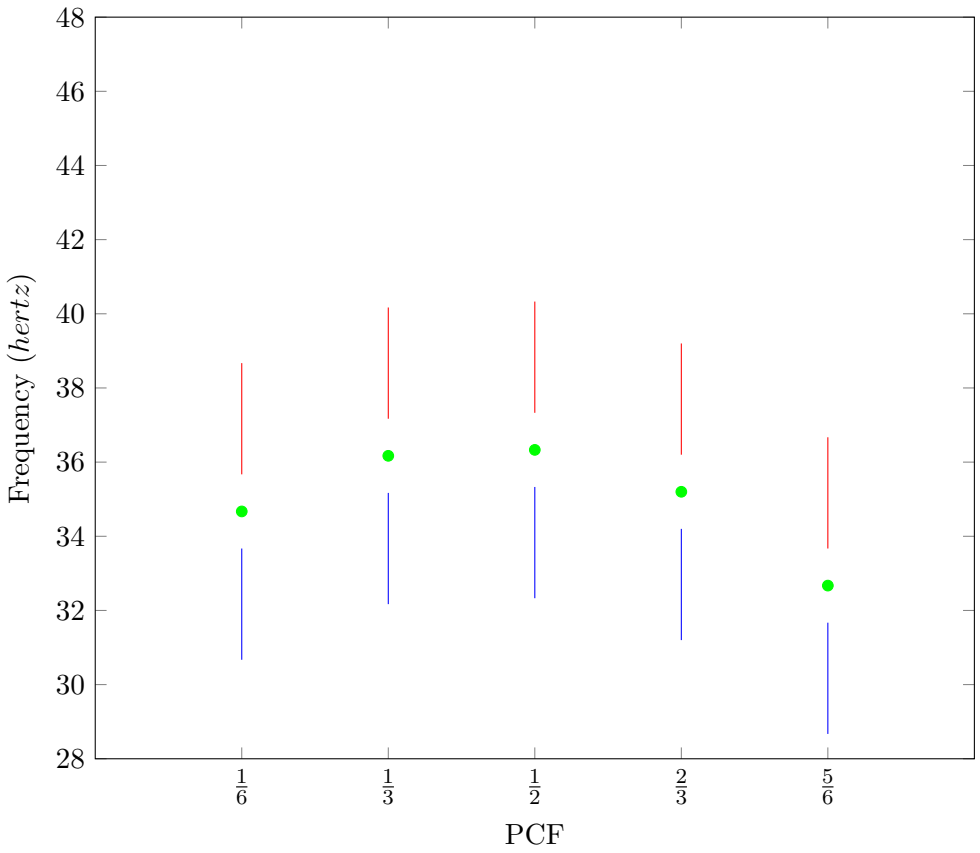


FIGURE 6.1: The frequency ranges used to select the train flickering and fused datapoints for photic pulses with intensity 53.4 cd/ft^2, for training the network. The photic pulses with a particular frequency and PCF lying in the blue line were assigned the label 'flickering'. Those in the red line were assigned the label 'fused'. The corresponding CFFs for the five PCFs have been marked in green.

The time series data with a desired PCF and frequency were generated with the Python library function *scipy.signal.square*(). The square waves so generated will have two values -1 and 1 in them. By adding the value 1 to that wave, and then multiplying by 0.5, followed by further multiplication by the intensity value of the wave we get the sampled representation of a wave with desired intensity.

	Label Fused	Label Flickering
Validation Data (n=50)	n waves with frequency between CFF+1.3 to 100 Hz	n waves with frequency between CFF−2.5 to CFF−1.3 Hz
Training Data (n=10)	50% probability of waves with frequency CFF+1.3 to CFF+10 Hz 25% probability of a continuous signal with with value between intensity of wave and zero 25% probability of a continuous signal with value 0	n waves with frequency between CFF−2.5 to CFF-1.3Hz

TABLE 6.2: Frequency ranges or constant amplitudes associated with the photic pulses in a minibatch for training and validation for a particular range and PCF.

We have interpolated the data based on the assumption that for a frequency above CFF for a particular PCF, the photic pulse will appear as fused to the subject, and for frequencies below that CFF it will appear as flickering. Waves with frequencies just below CFF were assigned the label of flickering stimulus. Training data and validation data were generated from scratch in each iteration during the training process. Waves with frequencies between CFF−1.3 and CFF−2.5 were assigned the label flickering for each PCF. The frequency interval in which photic pulses were assigned the label flickering was kept small, as a photic pulse of long duration is indistinguishable from a steady source of light. We had no way to estimate that duration. Waves with frequencies between CFF+1.3 and CFF+10 were assigned the label non-flickering. This was done as any stimulus above the CFF will be perceived as non-flickering or fused. A constant non-flickering stimulus is indistinguishable from a fused flicker stimulus. So constant stimuli with amplitude between intensity of the flicker stimulus and 0 cd/ft^2, as well

as a complete dark stimuli with 0 cd/ft^2 throughout it were used in training dataset with the label fused. A diagrammatic representation of data ranges used for training for intensity 53.4 cd/ft^2 is seen in figure 6.1.

For validation, waves with frequencies of CFF+1.3 to CFF+100 were assigned the label fused. This was because photic pulses with frequencies above the CFF should be labeled as fused or steady by a subject. Waves with frequencies between CFF−1.3 and CFF−2.5 were used in with label flickering for validation as well. The trained neural network should detect a time series of constant amplitudes as fused. Since we did not know the minimum time period for which the neural network will classify a constant amplitude series as fused, the frequency range of photic pulses used in validating the neural network was kept the same as that was used for training the neural network. For a training minibatch, ten photic pulses each for both flicker and fused were generated for all PCFs. For five PCFs and two labels it provided 100 samples. Similarly, 50 photic pulses with label flicker and fused were generated for each PCF in validation minibatch. It provided 500 photic pulses for validation minibatch in each iteration. New minibatches were generated for each iteration by randomly sampling from the frequency range in both Experiment 1 and Experiment 2 described in the next sections. The training and validation data for each iteration is shown in table 6.2. In Table 6.2, n is the number of photic pulses generated for a class with a particular PCF in a minibatch. The total number of samples in the minibatch will be 10 times n, as there are 5 PCFs and two classes.

6.4 Experiments and Results

6.4.1 Experiment 1

Only the photic pulses with the middlemost intensity of 53.4 cd/ft^2 were used for training and testing in this experiment. Photic pulses of desired frequency and PCF with intensity 53.4 cd/ft^2 were generated with the method mentioned before for 12 seconds duration and sampled at 1 milliseconds. From this array, a smaller array of 7168 elements was selected from one of the first 3000 elements chosen at random. This was done to make the last element in photic pulse irrelevant during the training process. The classification of a photic pulse is independent of the fact whether the last element is of intensity 0 cd/ft^2 or some other intensity. From this representation of a photic pulse with 7168 elements, one of the first 500 elements was chosen at random. The elements in the array from index 0 to that of chosen element were filled with random intensity between 0 and the intensity of the photic pulse which is 53.4 cd/ft^2. This was done under the assumption that for a photic pulse of a finite enough time duration, some small perturbations at the start of the photic pulse will not have any effect on the final classification after a prolonged time period.

The neural network model used in the experiment is shown in Table 6.3. Eight convolution filters of 120 elements were each used in the layer 1. Biases were used in each layer. Activation function Leaky ReLU with default parameters (alpha=0.3) from the tensorflow library was used for convolution operation (layer 1). Valid convolution was performed in convolution layer.

Layer	Dimension	Dropout	Activation Function	Number of Parameters
Layer 0 (Input Layer)	[None, 7168, 1]			
Layer 1 (1 dimensional convolution)	[None, 7049, 8]		leaky ReLU	968
Layer 2 (RNN Basic Cell Final State)	[None, 8]	0.5	sigmoid	136
Layer 3 (Dense Layer) (Dense Layer)	[None, 8]	0.5	sigmoid	72
Layer 4 (output Layer) (output Layer)	[None, 2]		softmax	18

TABLE 6.3: Structure of neural network used in Experiment 1

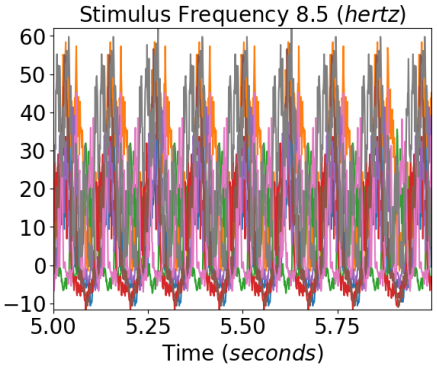
6.4.2 Results of Experiment 1

For testing, photic pulses of desired frequency, intensity 53.4 cd/ft^2 and PCF $\frac{1}{2}$ generated for a duration of 7.168 seconds, sampled at a frequency of 1 millisecond, were generated by scipy.signal.square function from the Python library. They were fed to the input layer (Layer 0). For a neural network with random weights, no noticeable differences could be observed for convolution outputs for 10 hertz or 20 hertz stimuli and stimuli of other frequencies. The set of weights for iteration with the lowest loss was chosen. The weights after iteration 1582, with a loss of 0.1153 and an accuracy of 0.972 for validation data were loaded into the network.

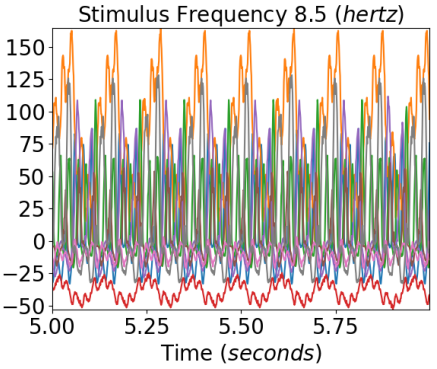
6.4.2.1 Output for 10 hertz signal

The output of a convolution layer for an untrained neural network can be seen in Figures 6.2a, 6.2c, and 6.2e for input stimuli of frequencies 8.5, 10, and 11.5 *hertz*. No particular resonance can be seen for a stimulus of frequency 10 *hertz*.

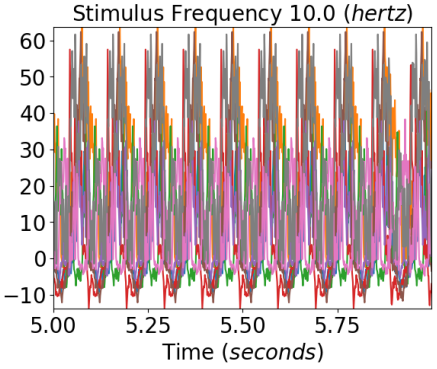
The outputs of a convolution layer for a trained neural network can be seen in Figure 6.2b, Figure 6.2d, and 6.2f for input stimuli of frequencies 8.5, 10, and 11.5 *hertz*. Sinusoidal ouputs can be observed for 10 *hertz* input in Figure 6.2d which cannot be seen for 8.5 *hertz* input and 11.5 *hertz* input in Figure 6.2b and Figure 6.2f. The outputs of neurons that gave sinusoidal outputs at ten hertz have been plotted with red



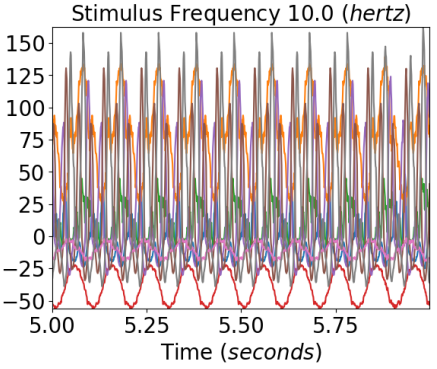
(A) For an untrained neural network



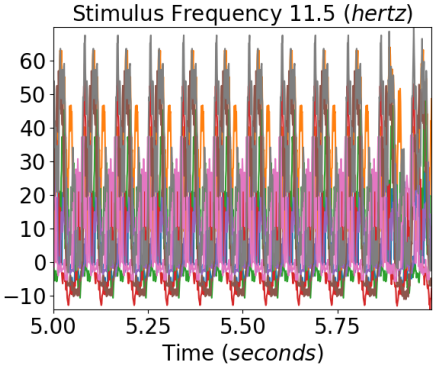
(B) For a trained neural network



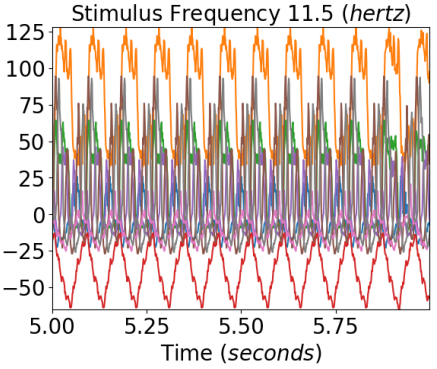
(C) For an untrained neural network



(D) For a trained neural network



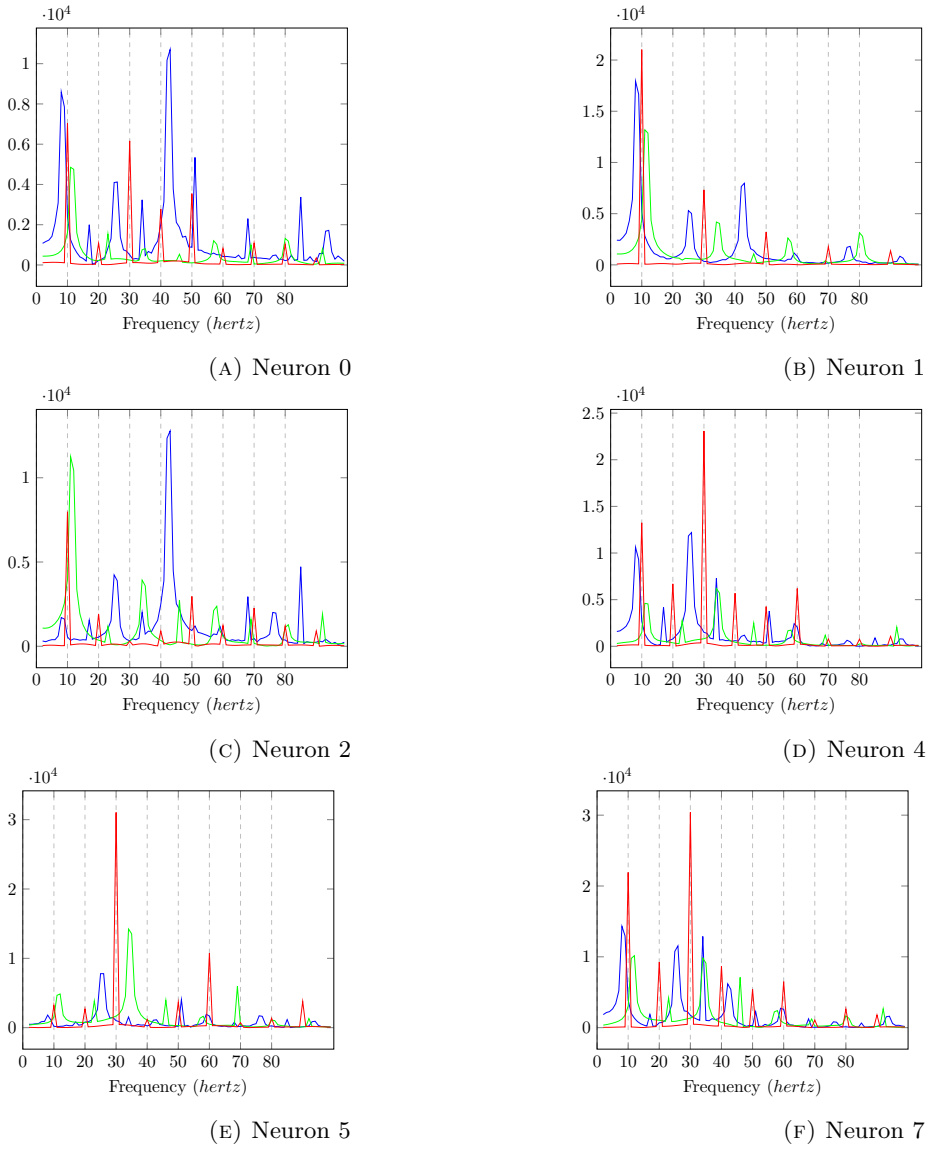
(E) For an untrained neural network



(F) For a trained neural network

FIGURE 6.2: Convolution layer outputs of the eight different neurons for input stimulus representations for 8.5, 10, and 11.5 hertz square waves are shown in the figure. The left side images show the output of an untrained neural network and the right side images show the output for the trained neural network. Sinusoidal outputs can be seen in the output to 10 hertz square wave inputs for the trained network for neurons 3 and 6, marked by red and green colors.

and green lines. For a pure sinusoidal signal, its fourier transform will have amplitudes at the frequency of the signal alone. It can be seen that for a ten hertz signal, the six neurons whose Fourier amplitudes have been plotted in 6.3 are not sinusoidal. For the two neurons that showed sinusoidal output at ten hertz stimulus, the neurons showed a prominent peak at 10 hertz in fourier transform compared to subharmonics as can be



(A) Neuron 0

(B) Neuron 1

(C) Neuron 2

(D) Neuron 4

(E) Neuron 5

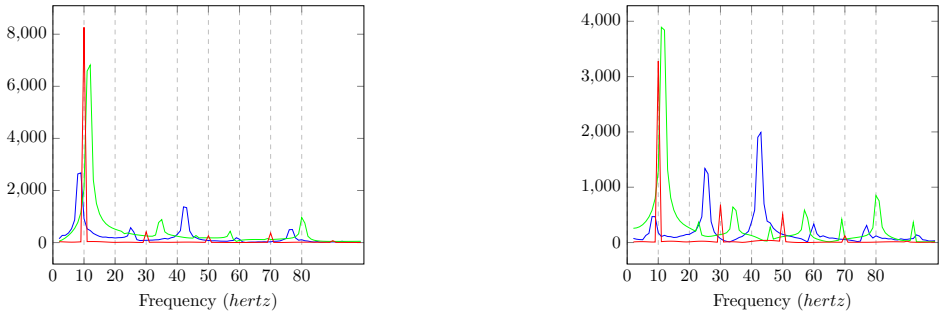
(F) Neuron 7

FIGURE 6.3: The amplitudes of the Fourier transform of the convolution output of the trained neural network for the six neurons whose convolution outputs did not show clear sinusoidal patterns for the ten hertz signal. The Fourier amplitudes for 8.5, 10, and 11.5 hertz have been plotted in blue, red, and green, respectively. We can see that multiple peaks can be observed for the 10 hertz signal at subharmonics of 10 hertz.

seen in Figure 6.4. The The output of two neurons for which the sinusoidal response at 10 hertz is prominent can be observed from Figures 6.5a, 6.5b, 6.5c.

6.4.2.2 Profile of fundamental frequency

The amplitude associated with the fundamental frequency of a subject as plotted by Herrmann showed clear resonance peaks around 10, 20 and 40 hertz[65]. The amplitude of fundamentals associated with neurons 3 and 6, which showed sinusoidal responses for ten hertz can be seen Figure 6.6. The profile of the trained network shown in Figure



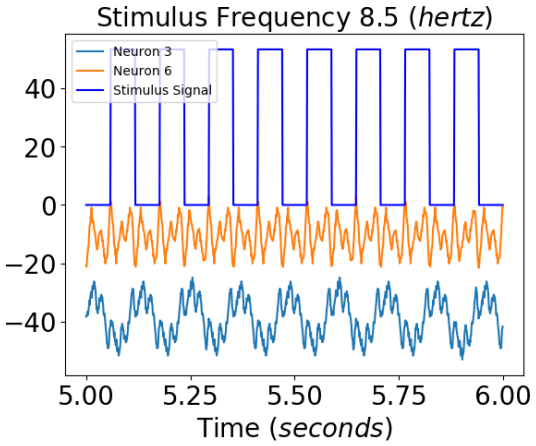
(A) Neuron 3 (B) Neuron 6

FIGURE 6.4: The amplitudes of the Fourier transform of the convolution output of the trained neural network for the six neurons whose convolution outputs showed sinusoidal patterns for the ten hertz signal. The Fourier amplitudes for 8.5, 10, and 11.5 hertz have been plotted in blue, red, and green. The peaks at subharmonics for ten hertz are much lower than the fundamental frequency for the output of a ten hertz signal.

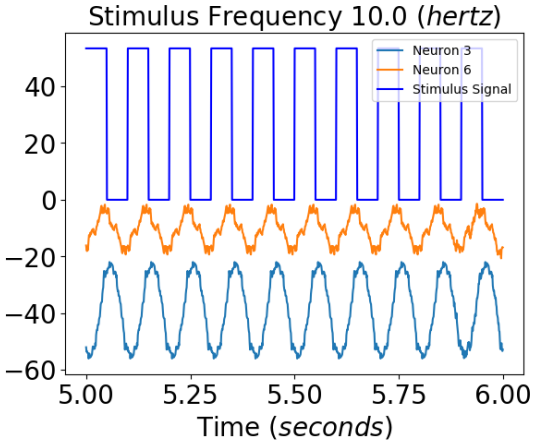
6.6b shows four distinct peaks around 10, 20, 40 ,and 80 hertz for the neurons. This is in contrast with the profile for an untrained network in figure 6.6a, which shows numerous peaks but no distinct resonance peaks. Previous works on human EEG response to flicker stimulus have shown that the average of fundamental frequency across 10 subjects exhibited strong resonance peaks around 10 hertz and weaker peaks in the 20-30 hertz and 40-50 hertz range[65]. The data for a single subject consisted of only one peak in the mentioned three ranges[65].

6.4.2.3 Output for 80 hertz signal

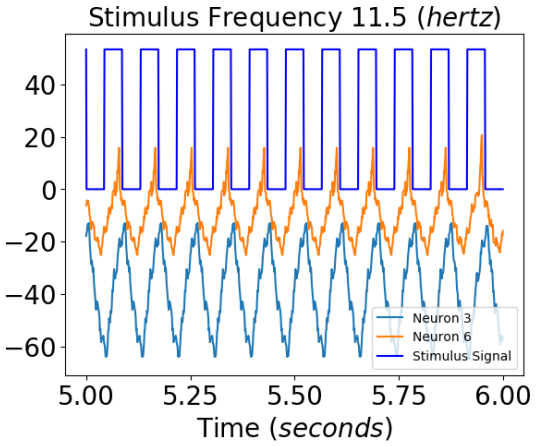
An 80 hertz flicker stimulus is known to invoke clear ten hertz response in human EEG which is not present at adjacent frequencies[65]. The output of the convolution layer for an untrained neural network for frequencies 79, 80 and 81 hertz can be seen in Figures 6.7a, 6.7c and 6.7e. The output of the convolution layer for an trained neural network for frequencies 79, 80 and 81 hertz can be seen in Figures 6.7b, 6.7d and 6.7f. We could not obtain a sinusoidal 10 hertz response to 80 hertz stimulus as reported in literature. But an envelope which is not observed at 79 and 81 hertz stimulus could be observed for 80 hertz stimulus for some neurons of a trained neural network as can be seen in Figure 6.8. The envelopes of a signal can be obtained by extrema sampling followed by signal reconstruction using sampled extrema with methods like Cubic Spline[130]. The envelopes were constructed by the method of sampling the maxima with a sample window of 25 points and signal reconstruction using cubic spline. The amplitudes of the evoked frequencies for stimulus of 79, 80 and 81 hertz for both trained and untrained neural networks can be seen in Figure 6.9. A clear subharmonic of 40 hertz can be seen in output of trained network for 80 hertz as can be seen in Figure 6.9d while the subharmonic is absent for output of stimulus of 79 and 81 hertz of the same network as can be seen in Figures 6.9b and 6.9f.



(A) Convolution layer output of two selected neurons for 8.5 *hertz* input signal

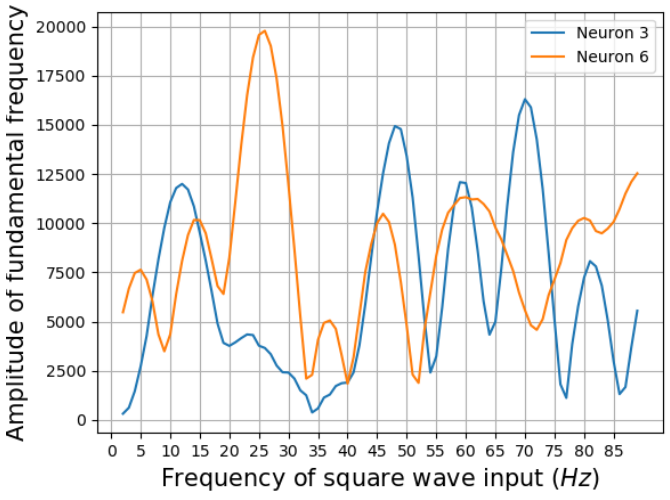


(B) Convolution layer output of two selected neurons for 10 *hertz* input signal

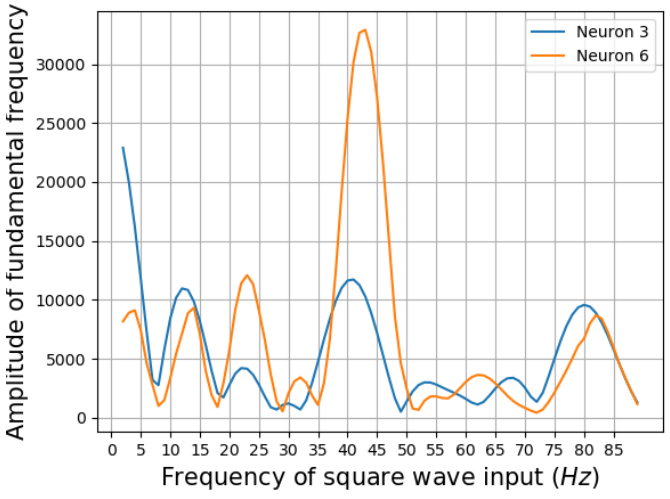


(C) Convolution layer output of two selected neurons for 11.5 *hertz* input signal

FIGURE 6.5: The convolution layer output of the trained neural network used in Experiment 1 for two selected neurons for 8.5, 10, and 11.5 hertz. A clear sinusoidal output could be observed for 10 hertz signal which is not the case with 8.5 or 11.5 hertz signals.



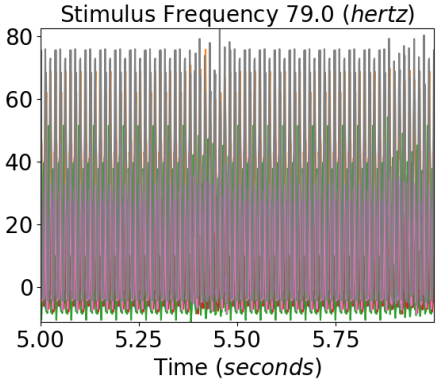
(A) Profile of untrained network



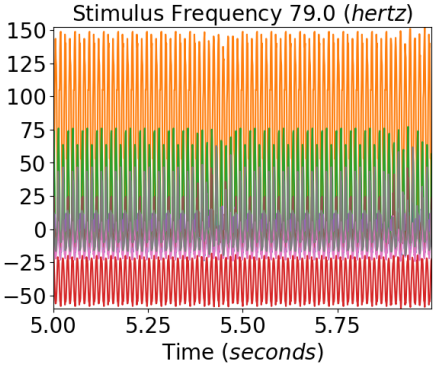
(B) Profile of a trained network

FIGURE 6.6: Profile of the fundamental for the two neurons that gave sinusoidal output for 10 hertz stimulus. Previous studies have shown that the human EEG output of individual subjects gave three distinct peaks around 10 hertz and in 20-30 hertz and 40-50 hertz ranges[65]. The subjective comparisons show that the profile of fundamental frequency for a trained network is closer to that of human EEG when compared to that of an untrained network, which is having a lesser number of peaks

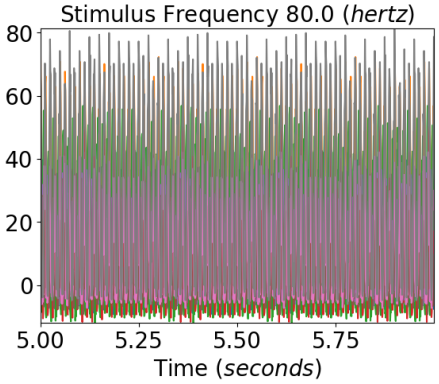
.



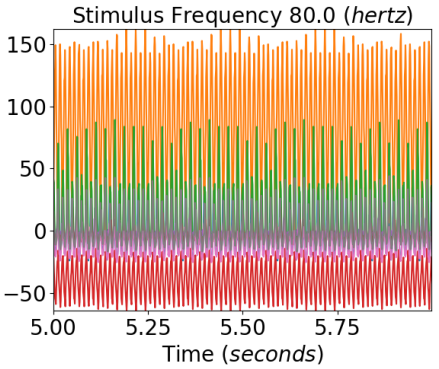
(A) For an untrained neural network



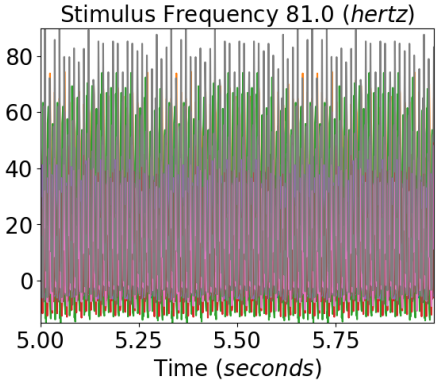
(B) For a trained neural network



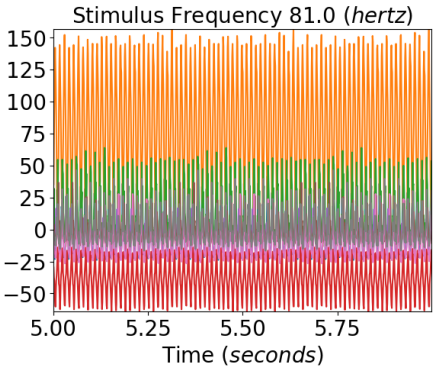
(C) For an untrained neural network



(D) For a trained neural network

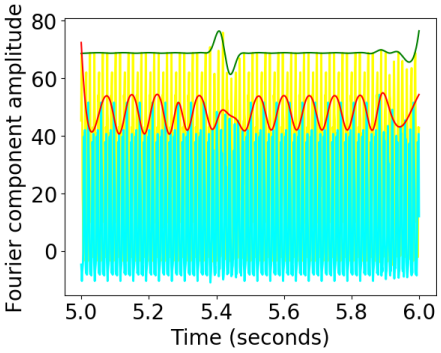


(E) For an untrained neural network

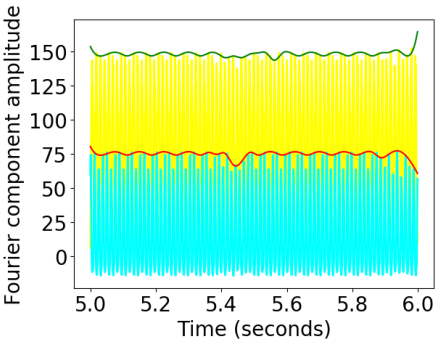


(F) For a trained neural network

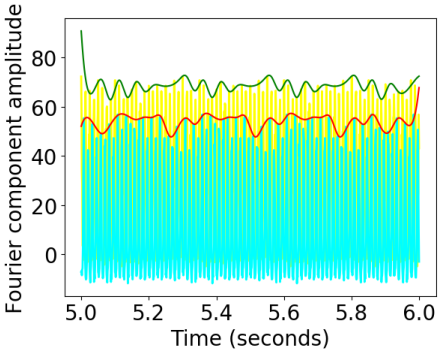
FIGURE 6.7: Convolution layer outputs for input stimulus representations for 79, 80 and 81 hertz. Previous studies have found a 10 hertz component in human EEG response to 80 hertz flicker[65]. Two envelopes with a low frequency can be seen in output of the trained network for 80 hertz stimulus.



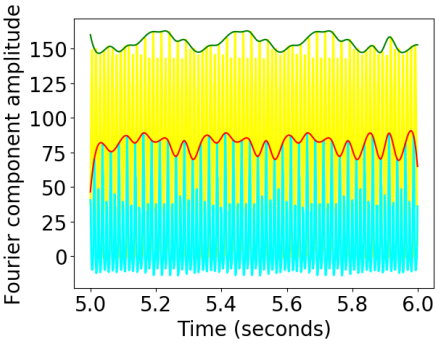
(A) For an untrained neural network



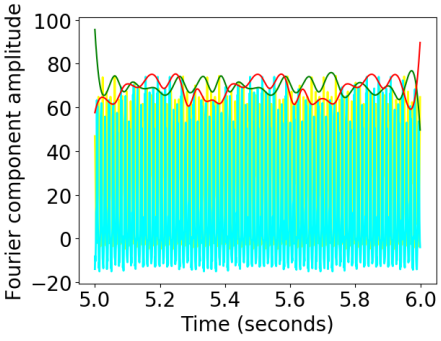
(B) For a trained neural network



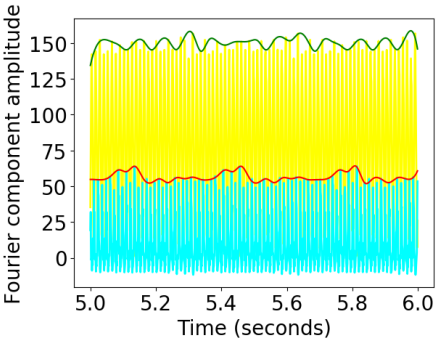
(C) For an untrained neural network



(D) For a trained neural network

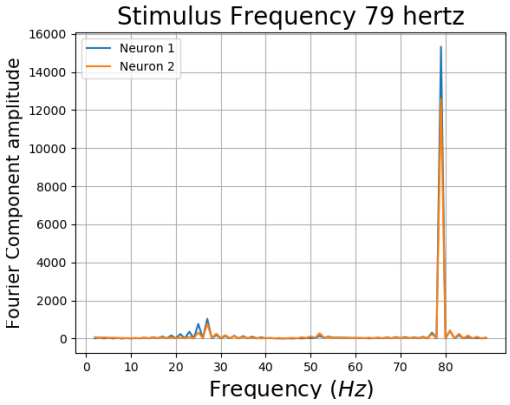


(E) For an untrained neural network

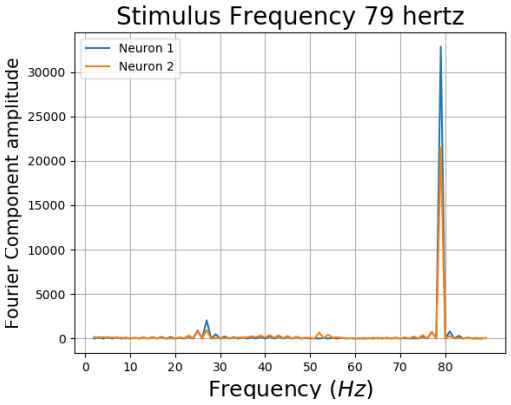


(F) For a trained neural network

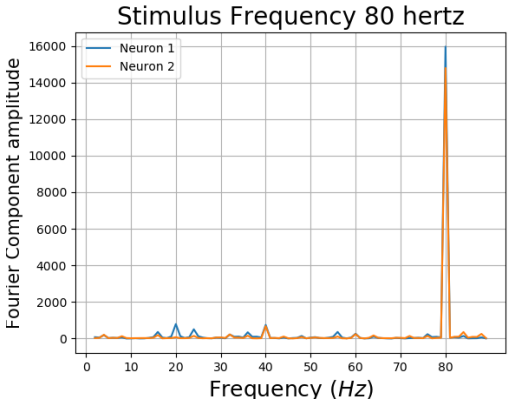
FIGURE 6.8: The convolution layer output of the trained neural network used in Experiment 1 for two selected neurons 1 and 2 have been plotted in cyan and yellow for 79, 80 and 81 hertz. An envelope with low frequency can be seen in the convolution output for 80 hertz signal. The envelops have been constructed using maxima sampling and reconstruction of signal using cubic spline interpolation.



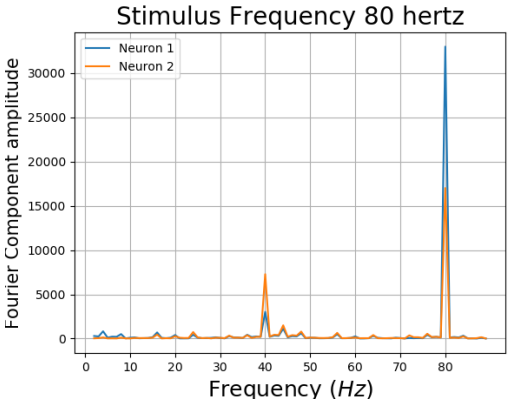
(A) For an untrained network



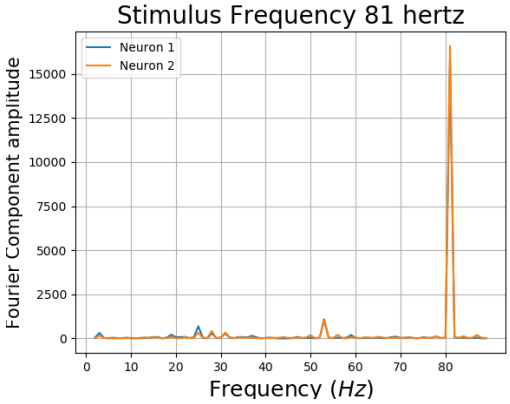
(B) For a trained network



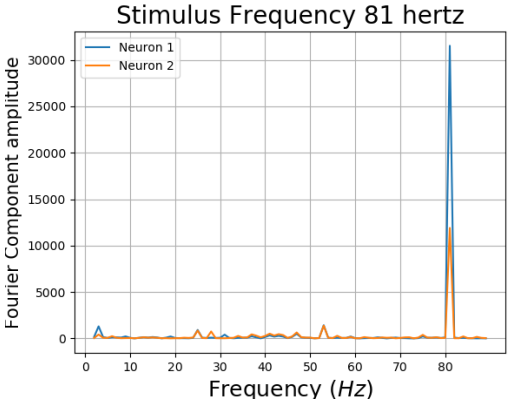
(C) For an untrained network



(D) For a trained network



(E) For an untrained network



(F) For a trained network

FIGURE 6.9: The amplitudes associated with various frequencies in the convolution output for square wave stimulus of frequencies 79, 80 and 81 hertz for both trained and untrained neural network. A clear subharmonic with frequency 40 hertz can be seen in output of 80 hertz stimulus for the trained neural network.

6.4.3 Experiment 2 and Results

The data was trained with a different neural network, having ReLU activation function in the convolution layer. The outputs of convolution layer were added together in a dense layer before going to the recurrent layer. It was done under the assumption that different cells in human visual system can have different temporal responses to a flickering stimulus and they might be present in the same biological layers in the visual system. In human retina the midget and parasol ganglion cells, have different temporal responses[131].

The photic pulse trains were sampled at 0.5 milliseconds to create a representation of waveforms. Square Wave arrays of 22384 elements with desired frequencies were generated. From these an array of length 16384 elements were selected, starting with one of the first 3000 elements chosen at random. The initial elements of this array were chosen at random from a number between 0 and 700. The initial elements of photic pulse representation so generated, was filled with a random intensity between 0 and intensity of the wave. The reason for this randomization was the same as in Experiment 1. These perturbations augment the training data. The whole train represented a photic pulse of duration 8.192 seconds.

Layer	Dimension	Dropout	Kernel constraint	Activation Function	Number of Paramenters
Layer 0 (Input Layer)	[None, 16384, 1]				
Layer 1 (Convolution Layer)	[None, 16245, 8]			ReLU	2248
Layer 2 (Dense Layer)	[None, 16245, 4]		ReLU	ReLU	36
Layer 3 (RNN Basic Cell Final State)	[None, 5]	0.5		sigmoid	50
Layer 4 (Dense Layer)	[None, 5]	0.5		sigmoid	30
Layer 5 (Output Layer)	[None, 2]			softmax	12

TABLE 6.4: Structure of neural network used in Experiment 2

The structure of neural network used in the experiment can be seen in Table 6.4. Bias were used for both dense layers, convolution layer and RNN basic cell. In the dense layer after convolution (layer 2), kernel constraint Rectified Linear Output was applied to the kernel parameters. A convolution layer with a convolution operator length of 280 weights was used in the first layer. It corresponded with a time period of 140 ms.

The loss and accuracy for each iteration was tabulated after training. The losses and accuracies for validation set over iterations were filtered with a Gaussian filter of sigma 11. The selected iterations have been listed in Table 6.5.

Intensity cd/ft^2	Total iterations	Least Loss	Selected Iteration	Accuracy
5340	4000	0.649	1893	0.6040
534	4000	0.3298	3411	0.89
53.4	4000	0.1065	2128	0.968
5.34	2000	0.0448	1938	0.998
0.53	4000	0.043	3953	0.994

TABLE 6.5: The iteration whose corresponding weights were chosen for neural network.

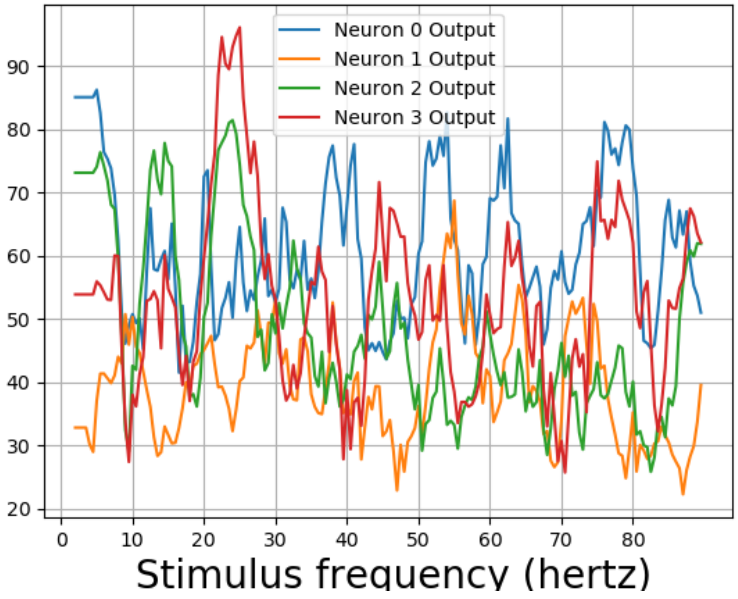


FIGURE 6.10: The differences in maximum and minimum amplitudes in a 1 second interval from the output of Layer 2 of an untrained neural network used in experiment 2.

For testing, the model waves of intensities 5340 cd/ft^2 and 534 cd/ft^2 were not considered due to low training accuracies. The difference in maximum and minimum values from layer 2, for one second interval for an untrained neural network for various frequencies for photic pulse of intensity 53.4 cd/ft^2, can be be seen in Figure 6.10. The differences from layer 2 output of neural network trained twice, with 53.4 *hertz* stimuli can be seen in figure 6.11 and 6.12. A peak can be observed at 10 hertz in both Figures. Peaks could also be seen near 20 hertz and 40 hertz. Some extra peaks including those at 30 hertz were obtained in both the Figures. The output for a neural network trained with 5.34 cd/ft^2 for same operation can be seen in Figure 6.13. The output did not show peaks at 10, 20, 40 or 80 hertz.

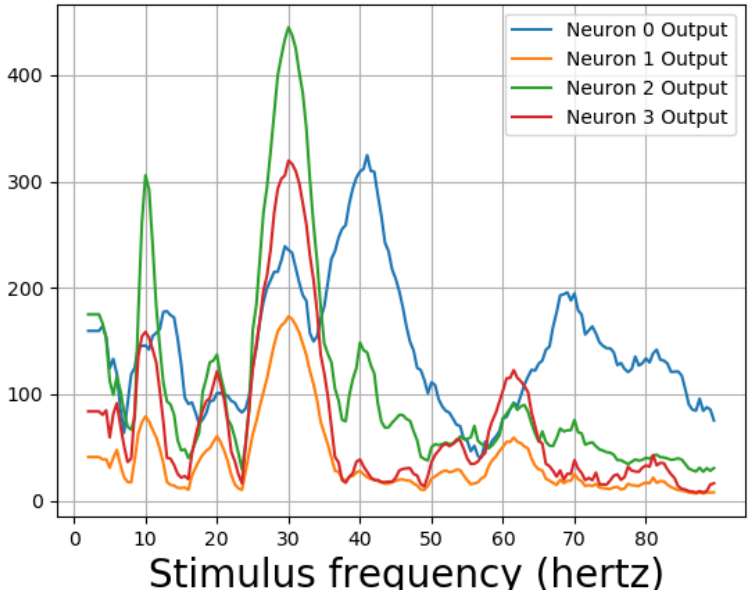


FIGURE 6.11: The differences in maximum and minimum amplitudes from the output of layer 2 for 1 second interval for a neural network trained and tested with 53.4 cd/ft^2 photic pulses.

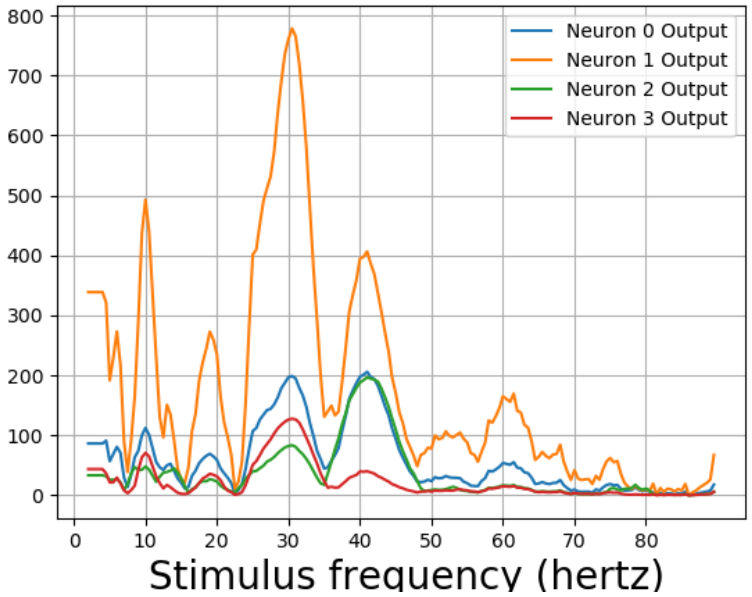


FIGURE 6.12: The differences in maximum and minimum amplitudes from the output of layer 2 for 1 second interval of a neural network trained and tested with 53.4 cd/ft^2 photic pulses. The neural network was trained with same method for a second time.

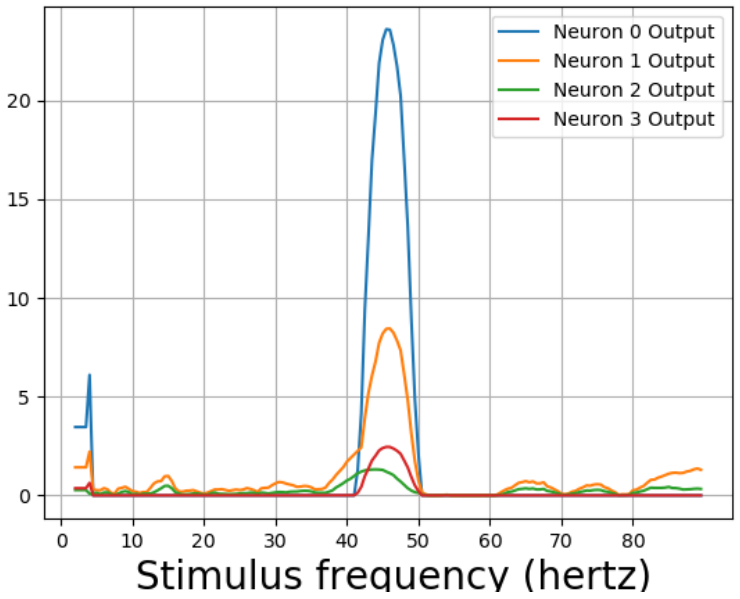
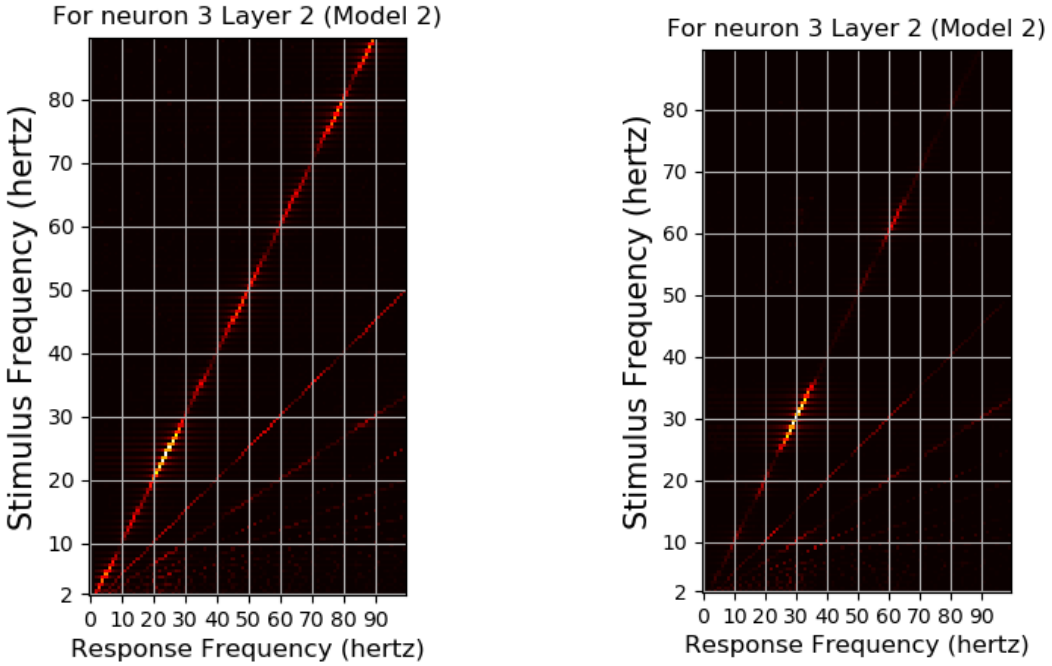


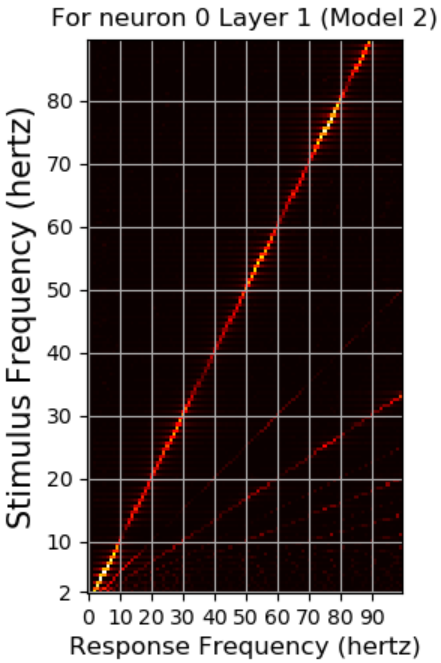
FIGURE 6.13: The differences in maximum and minimum amplitudes from the output of layer 2 for 1 second interval for a neural network trained and tested with 5.34 cd/ft^2 photic pulses.



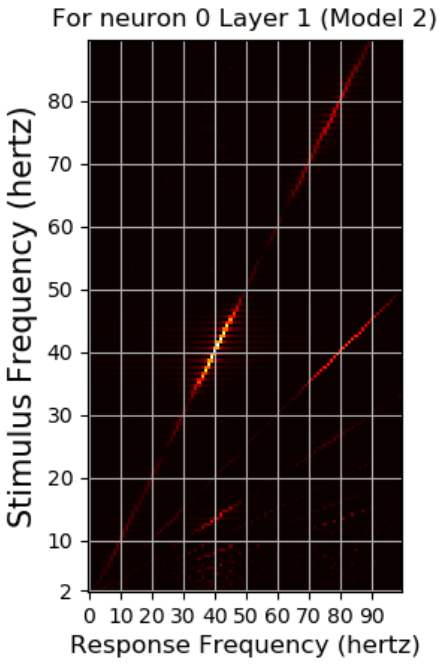
(A) For a untrained neural network

(B) For a trained neural network

FIGURE 6.14: The response frequencies vs stimulus frequencies in output of neural network in Neuron 3 for Layer2



(A) For an untrained neural network. This shows that harmonics of fundamental frequencies of the square waves will be present in the output of a random convolution operation on it.



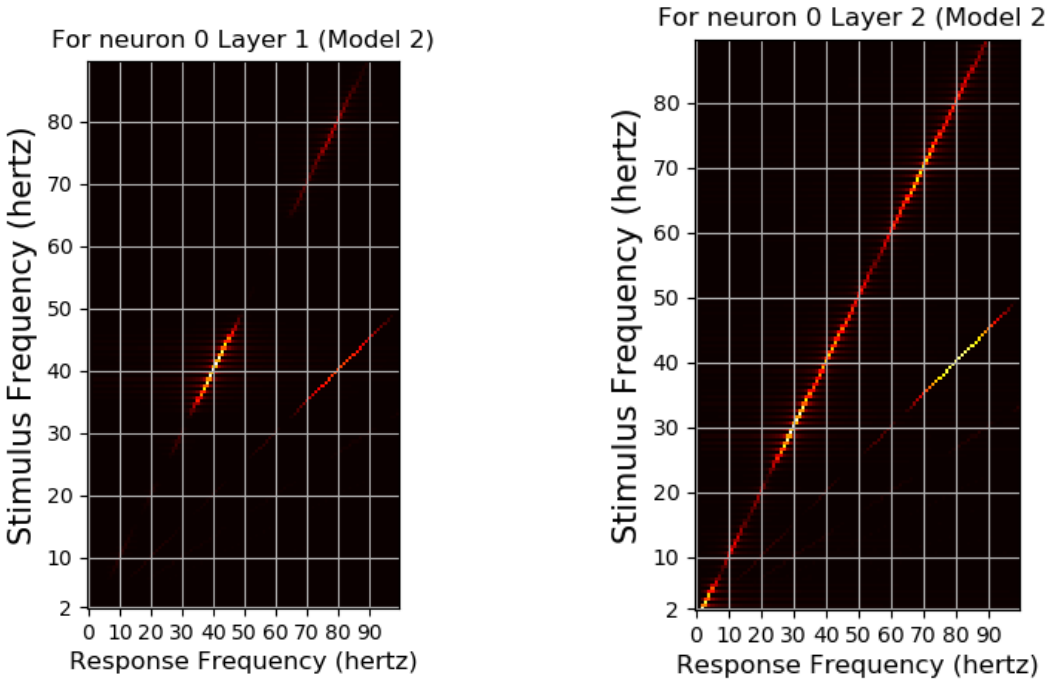
(B) For a trained neural network. It can be seen that response frequencies are more for some specific frequencies unlike a random convolution operation.

FIGURE 6.15: The response frequencies vs stimulus frequencies in output of Neuron 0 of Layer1 of neural network used in Experiment 2.

The stimulus frequencies for square waves that invokes responses for frequencies from 2 to 100 hertz for this neural network for photic pulse of intensity 53.4 cd/ft^2 have been plotted in figures 6.14a, 6.14b, 6.15a and 6.15b. The present model corresponding to Experiment 2 (referred to as Model 2) predicts fundamental frequencies as well as harmonics in the output of the neural network as can be seen in figures 6.14b and 6.15b. Even a convolution filter with random weight can generate harmonic frequencies in output of a periodic stimulus as is evident from 6.15a. It can be seen from the figures that the convolution output of a trained neural network is more sensitive to specific frequencies than an untrained neural network. The output of a trained neural network is similar to human EEG response to flicker in being sensitive to certain frequencies. Further as can be seen in Figure 6.16 the fundamental as well as harmonics are evoked even in the sinusoidal response to flicker stimulus.

6.5 Discussion

The findings from the present work indicate that CRNNs can be used to train psychophysics data of photic pulses for a particular intensity that varies in PCF and frequency. The present work shows that it is possible to obtain a clear sinusoidal response



(A) For neuron 0 in Layer 1

(B) For neuron 0 in Layer 2

FIGURE 6.16: The response frequencies vs stimulus frequencies for a sinusodial stimulus of waveform $I = I_o(1 + \mathbf{m}sin(\omega t))$ for the neurons for a neural network of Model 2 trained with photic pulse representations of $54.3cd/ft^2$. Here $Io = 54.3$, $\mathbf{m} = 0.1$ and $\omega = 2\pi f$ where f is the the frequency of the stimulus. The subharmonics are present in the output of the sinusoidal stimulus.

at 10 hertz from a neural network trained with psychophysics data, in spite of the fact that no 10 hertz photic pulse signal or EEG data were used in training the neural network. Clear sinusoidal responses have been previously obtained near 10 hertz for EEG response to flicker[86]. It also shows that the convolution layers of such a CRNN are more likely to show resonant output at some particular frequencies, similar to the human EEG response to flicker. The similarities in output between the convolution layer of the CRNN and the EEG response to the stimulus can be explained by the assumption that a small region in a cortical layer with thousands of neurons in it, acts in a manner similar to that of an artificial neuron in an ANN. When neurons present in a cortical layer fire in unison, the electric fields generated by the neurons can be measured outside the brain via EEG. The model put forth in this work that human EEG response to flicker stimulus is a convolution operation on the stimulus is in line with the hypothesis that the response is the superposition of event related potentials as opposed to the entrainment hypothesis.

The main sources of light in the natural world, viz., sun and moon, emit steady light with no sudden change in intensity. The only way the intensity of light emitted from any point in the natural world would change is when there is some motion at that point. The cell assemblies that lead to characteristic response in EEG, to flicker stimuli may have evolved in order to detect motion at particular points.

Flicker fusion can happen at various stages of information processing, beginning with the retina. Here, only one stage of differentiation in the brain was considered. This can be the reason why only one intensity was able to produce an EEG-like response in the intermediate layers of CRNN, trained to imitate the human brain. It could also be a reason why the neural network could not be trained with photic pulses of intensities greater than 53.4 cd/ft^2. The intermediate outputs of neural networks trained with intensities lesser than 53.4 cd/ft^2 did not show resonances near the reported frequencies where human EEG shows resonant response to flicker stimulus. Also, harmonics and sub-harmonic oscillations, in addition to fundamental frequency of flicker stimulus have been detected in human EEG response to flicker stimulus. The sub-harmonic oscillations, unlike harmonic oscillations, were reported to come from the parietal electrodes instead of the occipital electrodes[74]. No clear evidence of sub-harmonic oscillations were reported in local field potentials and multi unit activities of cat visual cortex stimulated with flicker stimulus[85]. The present work has been unable to detect sub-harmonic oscillations in the intermediate convolution layers. Digital circuits with flip flops are can be used as frequency doublers to generate a pulse with doubled frequency of the clock pulse[132]. A similar mechanism might be happening in the generation of the sub-harmonics of the fundamental frequencies of the flicker stimulus, with the flicker stimulus acting as a clock signal.

6.5.1 Limitation and future work

The present work was trained only on the data of a single subject, which is a limitation of this study. But the work does provides support for the methodology that can be tested with psychophysics and EEG data acquired for the purpose in future. The amplitudes of the fundamental and the first and second harmonics in SSVEPs are not stable over time in some human subjects[133]. The present model has been unable to explain this phenomenon. Moreover, we generated training and validation data on the assumption that, for a particular PCF and intensity, the human visual system will perceive all photic pulses with frequencies above the CFF as flickering. However outliers exist above and below CFF as demonstrated in Chapter 7 of this thesis. Also, in the present work, no EEG recordings were made at the time of acquisition of psychophysics data used in training the neural network. A future experiment with both psychophysics data and EEG data collected from the same subject in identical circumstances will possibly be able to ascertain more correlations between the two. In this case, direct mathematical correlations can be carried out instead of relying on subjective comparisons alone. The classification for data points represented by intensity, PCF and frequency can be measured for the subject by the method of constant stimulus. A large set of data points for a particular intensity with two parameters, viz. frequency and PCF, labeled into two classes, could be used to train a similar neural network. In spite of some of these limitations mentioned above, the present work provides the first such computational

framework that involves training a deep neural network with psychophysics data to predict brain activity at electrophysiological level.

6.6 Conclusion

The present work used a recurrent neural network based framework to model flicker, a time dependent psychophysics data. We have shown that the intermediate layers of the network could show features of the electrophysiological response to the stimulus. Clear sinusoidal responses could be obtained from intermediate layers of the network for a ten hertz stimulus input although no electrophysiological data was used to train the model. The presence of the fundamental frequency of the flicker stimulus as well as the harmonics can be explained as a temporal convolution operation of the stimulus. We have further shown that the output from convolution layer of a CRNN trained with psychophysics data will be more responsive at particular frequencies, similar to the human EEG response to flicker. The proposed CRNN model could be used to test the relationship between the psychophysics of flicker fusion and electrophysiology of flicker from a same subject.

Chapter 7

A device to generate mass psychophysics data on flicker fusion

Publications

- Kuntal Ghosh and Keerthi S Chandran. "A low-cost device and technique for generating big data in visual psychophysics to train brain models". In: *43rd European Conference on Visual Perception (ECVP) 2021 Online*. Vol. 50. 1 suppl. SAGE Publications, Dec. 2021, pp. 85. doi: http://doi.org/10.1177/03010066211059887
- Keerthi S. Chandran and Kuntal Ghosh. "A device for mass generation of psychophysics data to train and test models of flicker fusion". In: *Science Talks* 6 (May 2023), p. 100180. doi: http://doi.org/10.1016/j.sctalk.2023.100180

Chapter summary:

This chapter describes an instrument that can generate simple as well as complex flicker patterns, and record subject classification of the stimulus as either flickering or steady by button press. The experimental system also includes a mechanism to prevent dark adaptation by automatic lighting of a high intensity lamp after the button press for a certain time period. The next set of photic trains to be tested can be generated immediately after. The device is controlled by an Arduino, and the experiment can be done without the need of of human assistance to change temporal waveform and record the readings. We also describe a mechanism to measure the intensity of the

photic pulse train from a voltage reading generated from the instrument. The instrument is powered by a constant voltage source, and the intensity of pulse can be controlled by changing a variable resistance in series with the circuit. The instrument can generate photic pulses with different intensities and color in the same pulse train and allows the generation of data for simpler pulses which may be used for constructing a brain computational model as shown in Chapter 3 of this thesis. The model can be tested on the psychophysics results of the complex pulses.

7.1 Introduction

One of the important challenges in the measurement of human perception is its instrumentation and methods. The instrument-oriented research involves the measurement of external physical stimuli that gives rises to psychological responses. It also involves measurement of perception of the stimulus as well as development of sensors that can mimic human perception[134]. This thesis had proposed a DNN based brain computational model model for Flicker Fusion phenomena in Chapter 3. This chapter describes the construction and calibration of an instrument, for generating data to train and test such models. This work describes an instrument that may generate huge volume of psychophysics data on simple square flicker waves that can be described by four parameters color, luminance, PCF and frequency. This data, in turn, can be used to train Brain Computational Models via machine learning techniques. A central challenge of machine learning is that an algorithm should perform well on previously unseen input. This ability to perform well on previously unobserved inputs is called generalization. The generalization error of a machine learning model is tested on a test set of inputs that have been collected separately from the data used to train the model[5]. Since our device can generate complex waveforms where the alternate pulses might be of different intensity or color, the model may be validated using such psychophysics data on complex stimuli.

7.2 Disadvantages of computer screens

It would seem that computers could be programmed to produce many of the former experiments done on psychophysics of flicker perception on the screen. However, computer monitors are ill suited to conduct psychophysics experiment on flicker stimulus[135]. Cathode Ray tubes have fixed framerates[136]. It will impose limits on temporal sampling of flicker stimulus[136]. If the framerate is not an integer multiple of the flicker frequency, it will lead to temporal aliasing[136]. CRT monitors time to go off is dependent on phosphor decay times that can range from microseconds to milliseconds[136].

It severely limits higher frequency flickers whose psychophysics properties can be studied with a CRT monitor [136]. Computer monitors with Liquid Crystal Displays have slower response times than CRT monitors [135], making them ill suited for experiments on flicker perception.

7.3 The Instrument

FIGURE 7.1: The Light Emitting diodes soldered together in the device

Light Emitting Diodes(LEDs) have been used for generating flicker stimulus for measuring CFF[137]. These are semiconductor devices that can emit photons, with typical spectrum line width between 5 and 20 nanometers. The light is nearly monochromatic [138] and it allows us to study flicker perceptions and pathways in human visual systems involved with the particular wavelengths of light. Both the rise time of an LED (or the time it takes to reach the maximum value of its output in photons per second) and the fall time of the LED (or the time it takes to reach the minimum value of its output from maximum) are in the order of nanoseconds[139]. As the LEDs work very fast when compared to the human visual system, these constitute an ideal light source for the study of flicker perception.

Twenty eight Common Anode RGB LEDs together were used as the source of light. Each anode was connected to the positive terminal of a constant voltage source through a 100Ω resistance. The cathodes corresponding to Red, Green and Blue LEDs were shorted together. The soldered together LEDs can be seen in Figure 7.1. This made it possible to light up all the LEDs of a particular color together. Wires were connected and taken out from either side of a 100Ω resistance connected to an anode. The intensity

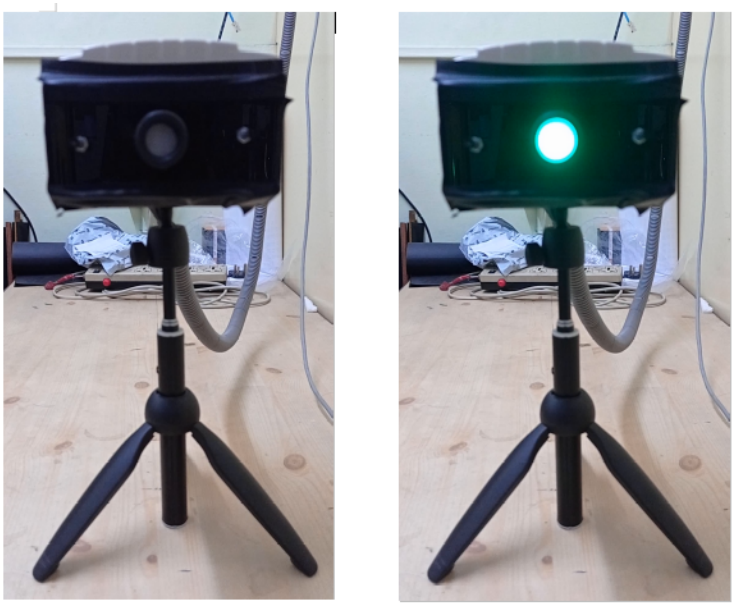


FIGURE 7.2: The device with no lights on (in the left-hand picture) and with green lights on (in the right-hand picture)

of light emitted by an LED will be a function of the current flowing through it. By measuring the potential difference across the 100Ω resistance, it is possible to calculate the intensity of the light. The Light Emitting Diodes were then placed in a box and mounted as shown in Figure 7.2 in two modes of operation. The light from the LEDs fall on a diffusion plate and the plate can be seen through a circular aperture of diameter 1.45 cm. The diffusion plate is used to create an even source of illumination from the light of LEDs. The light emitted out of the device will be uniform throughout the diffusion plate. The aperture blocks like from coming out of the entire diffusion plate and the light comes out of a source with circular shape.

The signals for photic pulses were generated using an Arduino UNO R3. Arduinos are an open source platform used for construction and programming in electronics in high level language[140]. The high level languages codes used for experiments have been provided in Appendix B. The setup includes a keypad with two buttons for subject response. Arduino is programmed to generate a new waveform after each button press as well as to print the subject response and physical parameters of the stimulus on a desktop computer. The signal from the Arduino was used to control flow of current from the cathode to the ground of a constant voltage supply with the use of transistors. Multiple transistors were connected in parallel, as the current through the collector could get higher than the rated current of the transistor. The emitters of transistors were connected in two different paths to the ground in series with variable resistances. By changing the resistance, it would be possible to vary the flow of current through the LEDs and hence the intensity of light. The current from the cathodes could flow to the ground of the power supply through two pathways that could be switched on and off via transistors.

Moreover, if an observer stays in a dark room for a long amount of time, the eyes will get dark adapted. It will change the values associated with the photic pulses in our psychophysics experiments[57]. In order to prevent it, a floodlight was switched on between the readings for a few seconds. The floodlight was powered by external AC power supply. A BT136 Triac controlled by an MOC3021 optocoupler was used to turn the floodlight on and off. The MOC3021 was controlled with signal from the Arduino.

7.4 Calibration

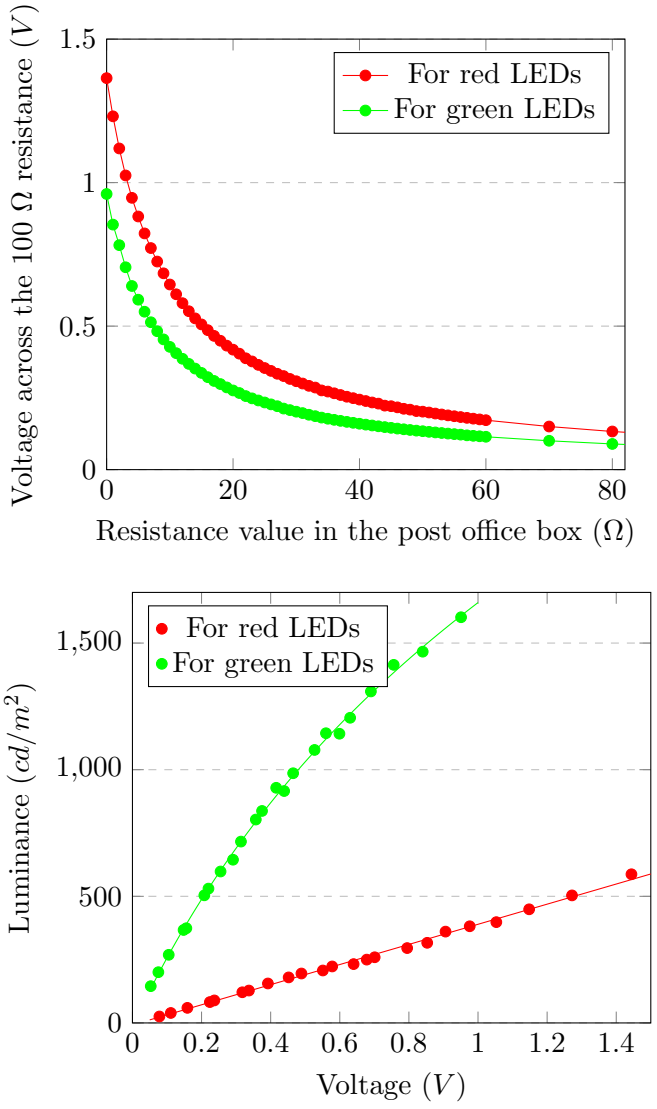


FIGURE 7.3: The top plot shows the variance of potential difference across the 100 Ω resistance with the resistance value in post office box for LEDs. The bottom graph shows variance of luminance with voltage across that resistance..

For calibration purpose, the circuits for LEDs of three colors were separately switched on without flicker. Calibrations were made between the potential difference across the

100 Ω resistance connected in series with an anode and resistance of a post office box used as a variable resistance. The luminance at the position of the eye was measured with a Gossen Mavospot 2 Luminance meter. It was calibrated with the voltage. It would then be possible to know the luminance of the device from the potential difference across the 100 Ω resistance. The device could be set to a desired luminance by changing the external resistance. The calibrations for LEDs of two colors can be seen in Figure 7.3.

7.5 Measurement Procedure and techniques

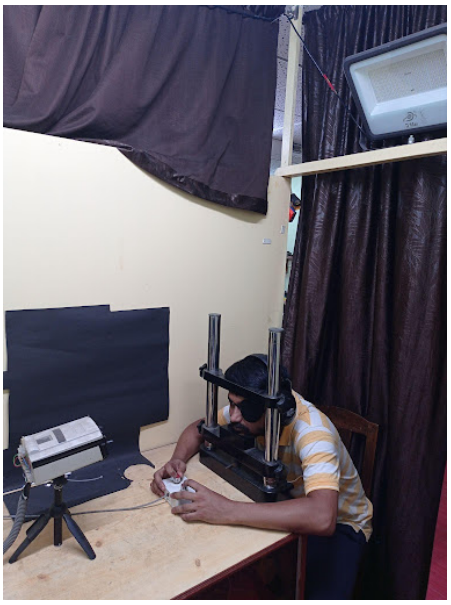


FIGURE 7.4: A subject seated infront of the device

The measurements were made in a dark room. The subject's head was kept fixed by a chin rest. The device was kept at a distance from subject to make a viewing angle of 2°. A subject viewed the aperture monocularly with right eye and left eye was left closed with an eye patch. A subject with the keypad and the floodlight to prevent dark adaptation can be seen in Figure 7.4.

7.5.1 Binary Search Method

In computer science, binary search is a searching algorithm that finds a position of a target value within a sorted array. Binary search compares the target value to an element in the middle of the array. If they are not equal, the half in which the target cannot lie is eliminated and the search continues on the remaining half, again taking the middle element to compare with the target value, and repeating this until the target value is found[141]. Our device makes use of binary search algorithm with the assumption that there is a threshold time period below which the flicker will appear as fused. The

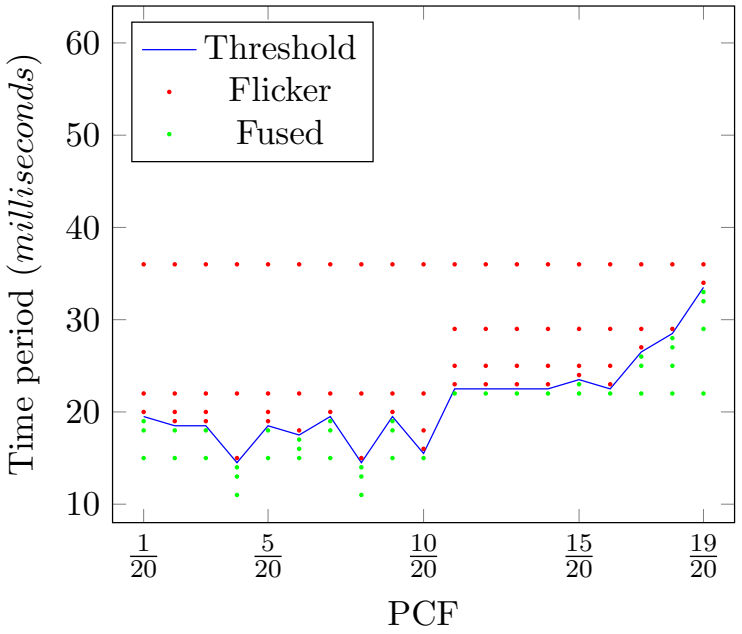


FIGURE 7.5: Data generated by binary search method

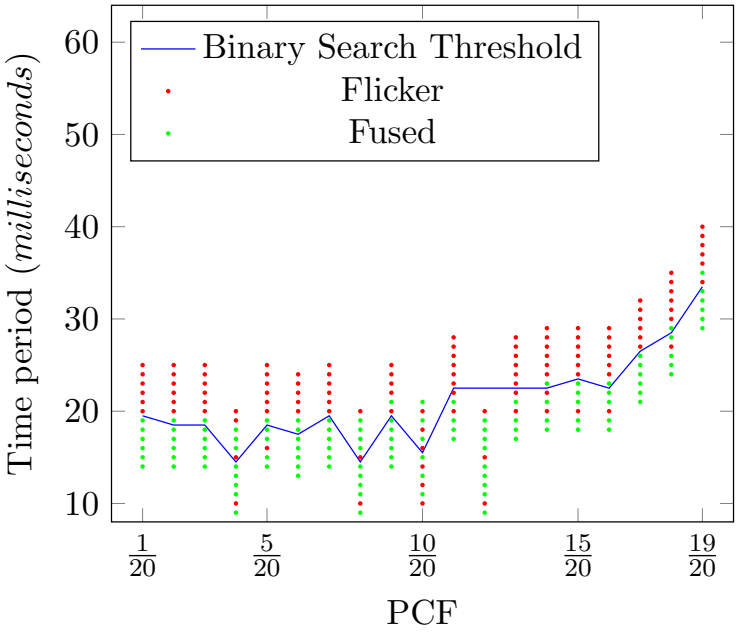
Arduino was programmed to employ a binary search algorithm to find this threshold time period. For a particular PCF, the device searches for this threshold time period for time period intervals between 64 and 8 milliseconds. The time period was measured for 19 PCFs between $\frac{1}{20}$ and $\frac{19}{20}$. The PCFs were presented in a random order.

7.5.2 Method of Constant Stimuli

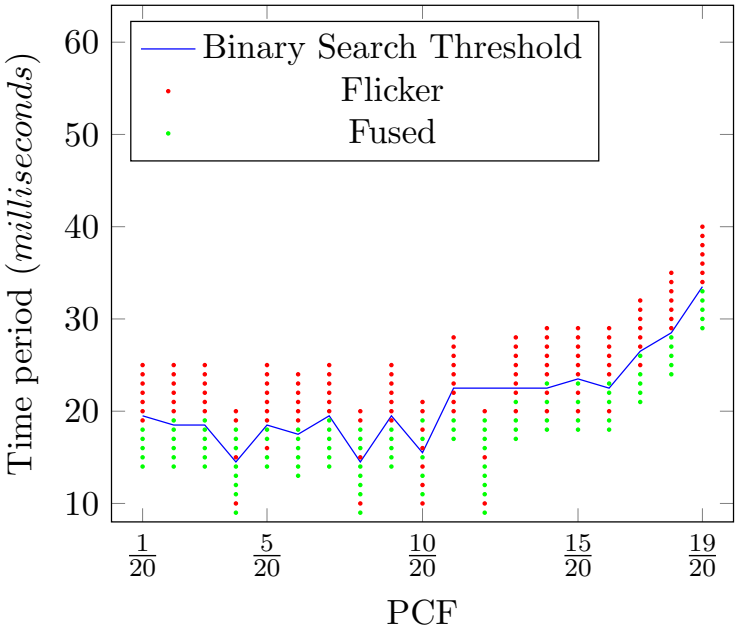
Method of constant stimuli is a psychophysics procedure in which each of a fixed set of stimuli (ranging near the threshold) is presented repeatedly in random order[142]. The threshold values detected by binary search method were used as thresholds. Twelve time period values near threshold values were presented for a PCF. The waveforms were presented to the subject in a total random order. The subject took readings by the method of constant stimulus twice on two different days. The readings obtained from method of constant stimulus can be seen in Figure 7.6.

7.5.3 Intra subject variation in measurement

The datapoints which were labeled as either flickering or steady in both readings, as well as datapoints that labeled differently in both readings have been plotted in Figure 7.7. The subject classified 113 datapoints as flicker and 113 datapoints as steady on both the days. The subject classified 13 datapoints differently.



(A) First Set of readings



(B) Second Set of readings

FIGURE 7.6: Data generated by the method of constant stimuli.

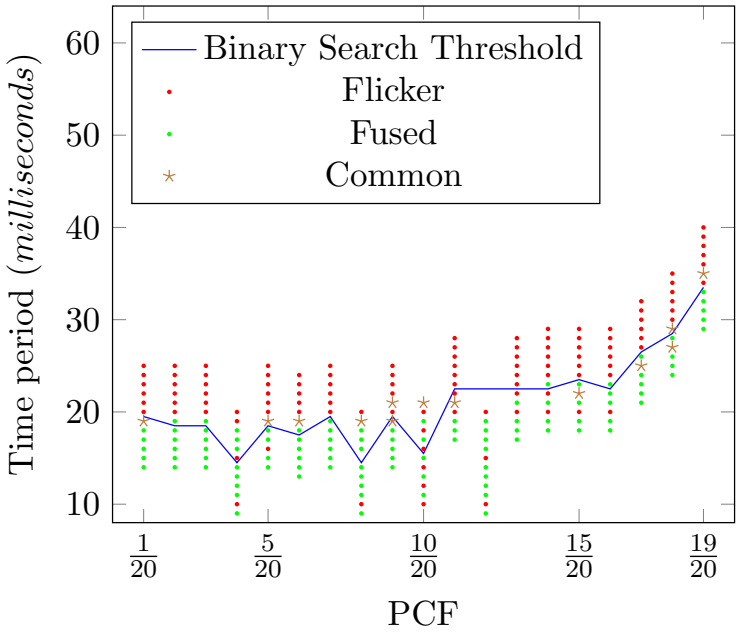


FIGURE 7.7: The datapoints labeled as flickering and steady by the subject in both set of readings as well as points marked flicker in one reading and steady in another reading

7.5.4 Uncertainty in measurement

- To calculate the uncertainty in such response, we can begin with the null hypothesis that the subject is giving completely random responses.
- In that case the possible responses of subject for two days for same data point is {(flicker, flicker), (flicker, steady), (steady, flicker), (steady, steady)}.
- So for a particular data point, there is a probability $p = \frac{1}{2}$ that subject gave same response on both the days and a probability $q = \frac{1}{2}$ that subject the subject did not.
- From experimental data, the subject gave same responses for 215 data points and different responses for 13 data points.
- Assuming null hypothesis is true, the probability value that there will be events with same or lower probabilities than the event in which 13 data points had different responses is, $p - value = 2\sum_{i=0}^{i=13} \frac{^nC_i}{p^i q^{n-i}}$ where n=228, which is total number of datapoints and i varying from 0 to 13.
- Calculating the above, we get $p - value = 2.514897 * 10^{-48}$
- So there is probability of $2.514897 * 10^{-48}$ that null hypothesis is true.
- So we can reject null hypothesis that the subject gave random responses and conclude that subject did not give random responses.

7.5.5 Binary search threshold as a classifier

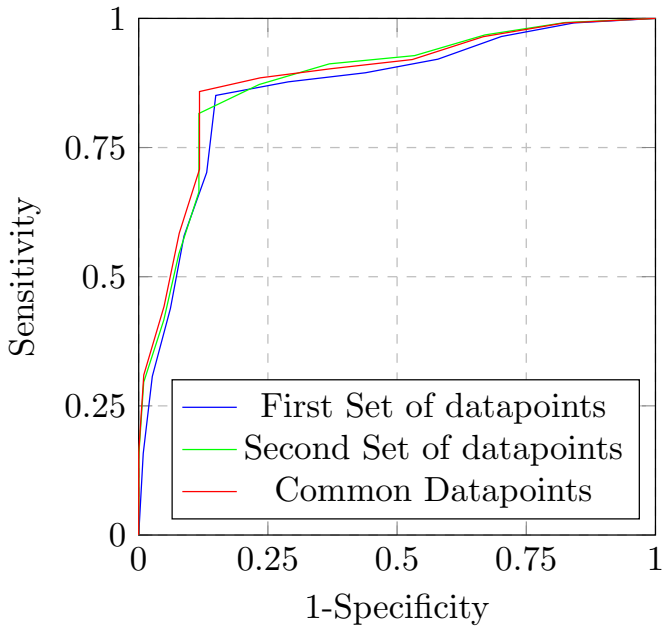
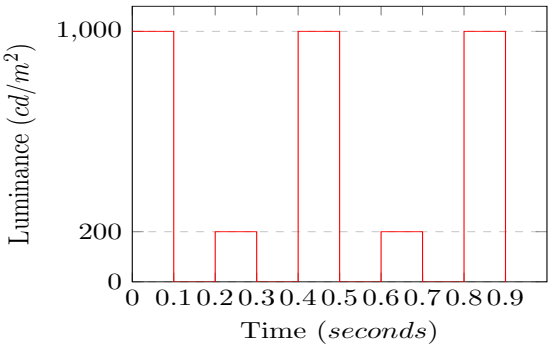


FIGURE 7.8: The ROC curve for the threshold obtained by binary search method. Its performance was tested on the two sets of readings and their common data points. The threshold does a better work in classifying the datapoints that were common on both days, as it has a slightly bigger area under the curve.

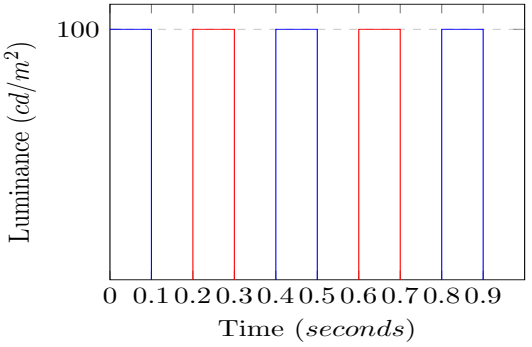
The efficacy of the threshold obtained by binary search method can be seen in the ROC curve plotted in Figure 7.8. The label flicker has been treated as positive and steady as negative. The binary search threshold line obtained in Figure 7.5 has been moved vertically along time period in Figures 7.6 and Figure 7.7 to obtain the ROC curve. As we can see that there does not exist a perfect threshold time period to classify a photic pulse as flickering or fused. There are outliers.

7.6 Conclusion

The device requires no human assistance to change the frequency or PCF while taking the readings. The generation of psychophysics data in large volume may enable it to be used in the construction of brain computational models using deep learning techniques. The photic pulses generated by the machine could be represented as a one dimensional temporal signal with multiple channels as seen in Figure 7.9b. The model thus trained by the device can be fitted with more complex flicker patterns where the alternate photic pulses can have different time periods, intensities, or colors as in Figure 7.9. Further the psychophysics study showed the presence of outliers in human flicker fusion data, that there is no strict time period or frequency separating steady from label flicker.



(A) A complex flicker pattern (achromatic) with pulses of different luminance in the same train



(B) A chromatic flicker pattern with photic pulses of different colors in the same train.

FIGURE 7.9: Examples of two complex photic pulse patterns that can be generated by the instrument. In the top image the alternate photic pulses are of green light but have different intensities. In the bottom image, the alternate photic pulses are of of two different colors, red and blue.

Chapter 8

Probing temporal filters of vision via a falsifiable model of flicker fusion

Publications

- Keerthi S. Chandran and Kuntal Ghosh "Probing Temporal Filters of Vision via a Falsifiable Model of Flicker Fusion". In: *Brain Informatics. BI 2024. Lecture Notes in Computer Science*, vol 15542. Springer, Singapore, 2025. doi: https://doi.org/10.1007/978-981-96-3297-8_14

Chapter summary:

This work trains a Deep Neural Network (DNN) based model of flicker fusion with human psychophysics data. The convolution filters of DNNs trained on natural images acquire the features of Gabor filters. This work shows that in a similar fashion, the convolution filters of the DNN trained with temporal psychophysics data acquired symmetrical features. Derivatives of Gaussian and Gabor functions found in human visual systems are often symmetric. The predictions made by the DNN on a complex flicker stimulus was tested with psychophysics experiment. The model is shown to be falsifiable and can be improved with further training.

8.1 Introduction

Many studies have suggested similarities between Deep Neural Networks(DNNs) trained for engineering tasks and the cognition in biological brains[9]. Many of the conclusions on similarities between DNNs and human cognition rely on the concept of emergence, which is training DNNs to one task leading to another known phenomena[9]. For instance, training DNNs on object recognition can lead to emergence of shape bias in neural networks[9]. Another emergent property of DNNs trained in object recognition is the emergence of Gabor like filters in convolution layer of DNNs trained on object recognition[2]. There have been criticisms that similarities between DNNs and biological or human vision are overstated, and the claims of similarities between DNNs and biological vision have been put to severe testing[9]. As per philosopher Karl Popper, falsifiability is a necessity for a theory to be considered scientific. Bowers et al. emphasizes the need to incorporate falsification tests in models comparing DNNs with human cognition[9]. This chapter uses data from psychophysics experiments done on the device described in previous chapter to model flicker fusion. An emergent phenomena that could be investigated is the presence of temporal filters in visual system. The spatiotemporal filters are are used to detect motion in models of computer and human vision[143]. Previous works have shown the existence of spatiotemporal filters in mammalian visual system[144][145]. This work examines only the temporal aspect of the visual system. In addition to it, the model is put to a falsification test via an empirical psychophysics experiment. This work shows that its possible to build falsifiable models of psychological phenomena, and that falsification tests can show that models could be further improved to be closer to biological reality.

8.1.1 Linear filters in the biological system and DNNs

The convolution filter of DNNs trained on object recognition acquire the features of Gabor filters, which are found in V1 neurons of biological brains[2]. Gabor wavelets were also learned as a result of training with sparse coding on natural images[146]. Filters in the shape of difference of Gaussian, and Laplacian of Gaussian are also found in the retinal receptive fields of cats and primates[147]. Previous works have modeled spatiotemporal filters in the visual system, which turned out to be derivatives of Gaussian[144][145]. Spatiotemporal filters are used as motion detectors in computer and human vision[143]. This work will probe the possibility of existence of filters similar to Gabor or derivatives of Gaussian in the human visual system via the CRNN model.

8.1.2 Contributions

There are two contributions in this work. The first is to show that, it is possible to model the temporal filters in visual system via data obtained from human psychophysics experiments. We go on to show that the filters so modeled have symmetric features. Second, we show that the model is falsifiable with human psychophysics experiments and the model can be further improved by fitting data to the model more rigorously.

8.2 Proposed Methodology

8.2.1 Artificial Neural Network Model

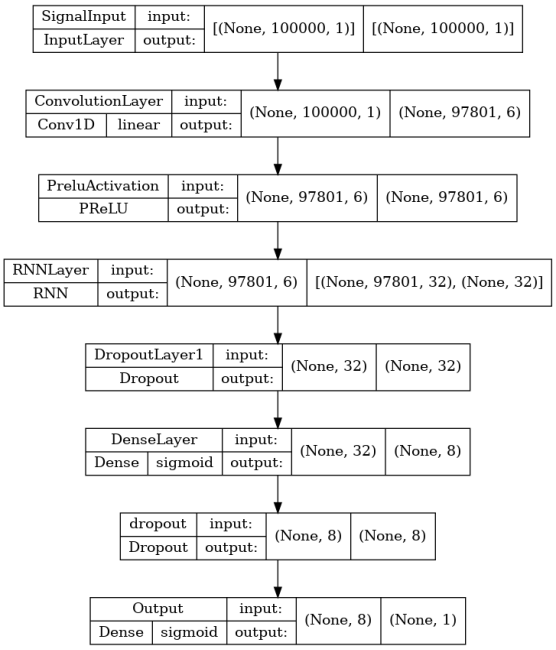


FIGURE 8.1: Model of DNN

The model used for training the psychophysics data of flicker fusion is a Convolutional Recurrent Neural Network, as done in previous chapters. The network used to train the model can be seen in Figure 8.1. The final predictions are made by the single neuron with a sigmoid activation function in the output layer. The network is trained to give an output of 1 if there is flicker and 0 if the output is steady. While training the network the output layer made predictions from the final state of recurrent layer. After training the network, the DNN structure can be changed to predict the output at all time points by connecting the dense layer above recurrent layer to outputs of all states instead of just the final state.

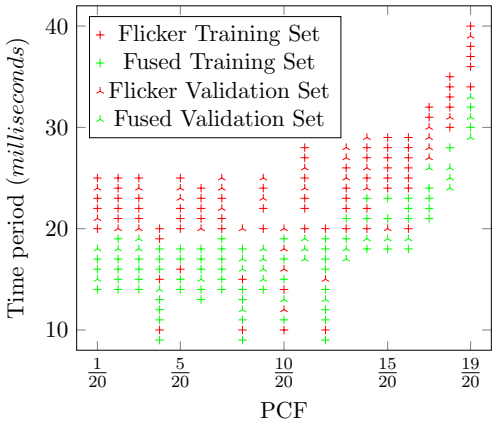


FIGURE 8.2: Data generated by the method of constant stimuli.

8.2.2 Training Data

The psychophysics data from a male subject generated from the device described in Chapter 7 was used to train the network. The psychophysics experiments were done with green light of luminance $\sim 1000 cd/m^2$. The dataset can be can be found at https://github.com/keerthischandran/flickerfusiondata . The dataset had classifications made by a subject on photic pulses with 19 PCFs from $\frac{1}{20}$ to $\frac{19}{20}$ with an increment of $\frac{1}{20}$. The measurements were made twice by the method of constant stimulus and the photic pulses that were classified same, both times, were added to the dataset. There were 113 photic pulses with label flicker and 102 photic pulses with label steady. The datapoints in both labels were split in the ratio 7:3 for training and validation. There are 71 and and 79 datapoints for training with labels steady and flicker. There are 31 and and 34 datapoints for validation with labels steady and flicker. The datapoints along with their sets have been plotted in Figure 8.2.

8.2.3 Signal Generation

- **Waveform Construction:** A binary wave representation was created for each combination of time period and Pulse Cycle Fraction (PCF). Time points where the green light was active were assigned a value of 1, while periods of inactivity were assigned 0.
- **Pulse Synthesis:** Photic pulses were synthesized by concatenating these binary sequences to match the desired PCF and time period characteristics.
- **Temporal Extension:** The wave representation was repeated to create a 10-second signal consisting of 200,000 datapoints.
- **Subsampling:** From the initial 10-second signal, a 5-second segment (100,000 datapoints) was extracted at a random starting position. This was done to randomize how the signal terminates.

- **Initial Perturbation:** To ensure the final classification remained robust against onset artifacts, a random value between 0 and 1 was assigned to the initial 0.5 to 0.75 seconds of every waveform. This introduction of noise provides a controlled perturbation that does not influence the final classification.

8.2.4 Minibatch Selection

New minibatches were generated for both training and validation in each iteration. Each minibatch had a selection of 64 datapoints for both flicker and steady with repetitions allowed in selection. In total a minibatch had 128 datapoints in it.

8.3 Results

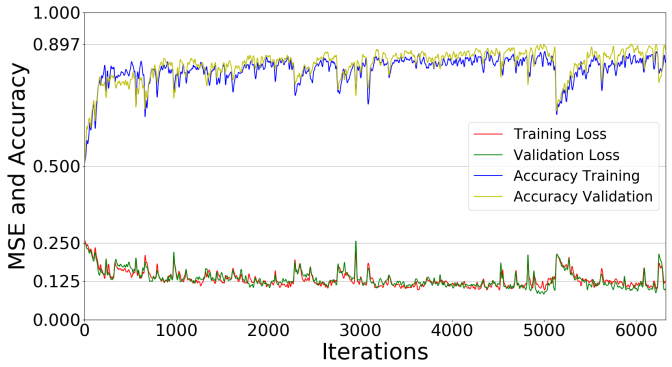
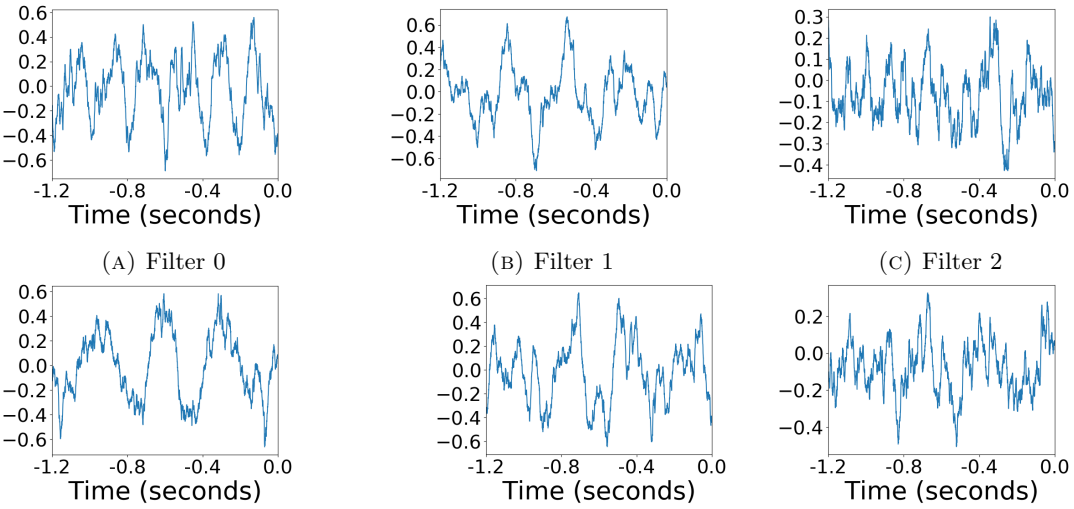


FIGURE 8.3: The training curve. The accuracies and losses have been Gaussian filtered with a sigma 5.

The training curve for the network for the dataset is in Figure 8.3. Of it the weight corresponding to iteration 5954 was chosen as it had highest accuracy in the array accuracies filtered by Gaussian filter of sigma 5.

8.3.1 Symmetric filters

The filters for the trained network have been plotted in Figures 8.4. The filters of Convolutional Neural Networks (CNNs) trained on engineering tasks like object classification acquire the features of Gabor filters[2]. An elementary Gabor filter can have an odd and even component[148]. The even component will be symmetric. Similar derivatives of Gaussians known to exist in the visual system can be symmetric. In order to find out if these are symmetrical filters, the filter arrays of dimension 2200 were flipped. The Pearson correlation were calculated for all possible wavelets of dimension 2100 from the filter and its flipped array. The filters which are symmetric or are even will have a high correlation. Three filters which showed symmetric features along with their shifted flipped

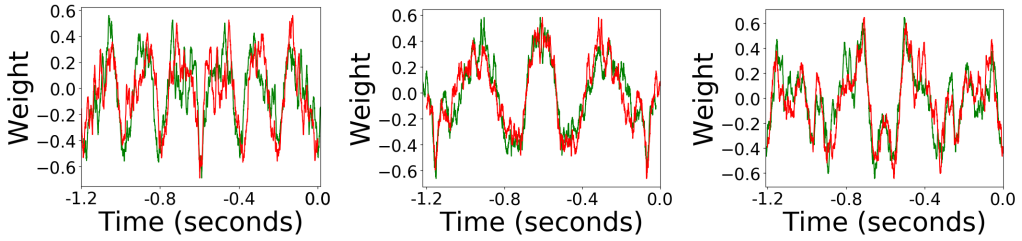


(A) Filter 0 (B) Filter 1 (C) Filter 2

(D) Filter 3 (E) Filter 4 (F) Filter 5

FIGURE 8.4: The convolution filters of the trained neural network

images have been plotted in Figure 8.5. The training of CRNN with psychophysics data led to formation of symmetric functions in the convolution filters.



(A) For neuron 0 with Pearson correlation of 0.665

(B) For neuron 3 with Pearson correlation of 0.902

(C) For neuron 4 with Pearson correlation of 0.840

FIGURE 8.5: The filters which showed symmetry have been plotted along with their time shifted reversed images. The filters have been plotted in red and their mirror images in green.

8.4 Falsifiability tests for the model

Bowers et al. have argued that many of the previous conclusions between DNNs and human brain arose because of lack of severe testing of hypothesis[9]. A theory is falsifiable if there are are basic statements that contradict it[149]. ANNs can be understood as computational languages in which falsifiable hypothesis can be expressed[150]. The model used in this work is shown to be falsifiable via two empirical experiments. For testing the model the output at all timepoints can be predicted by connecting the classifier over the recurrent layer to all recurrent states, instead of just being connected to final state as used for training the network.

8.4.1 Constant stimulus

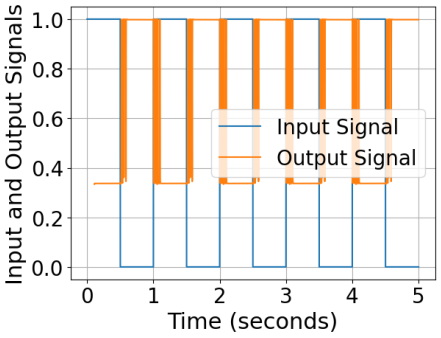


FIGURE 8.6: The output for stimulus of PCF $\frac{1}{2}$ and frequency 1 hertz.

A stimulus with constant amplitude should be classified as steady. It will give a classification of flicker only when there is a sudden change of luminance. As we can see in Figure 8.6, the network failed to satisfy this criteria.

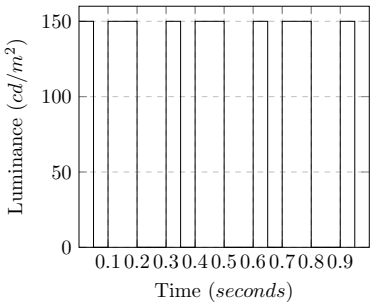
8.4.2 A Complex flicker pattern

Psychophsyics experiments have been conducted on photic pulses in which alternate pulses have different time periods[61][84] as in Figure 8.7a. The shape of the plot showing fused and flickering areas for various stimuli obtained from experiments have been plotted in Figure 8.7b. Former experimental works had determined that the boundary conditions for the fusion contour is determined by the lines $x + y = P$ and $x + y = 2P$, where P is the period associated with CFF of a simple stimulus of PCF $\frac{1}{2}$[61].

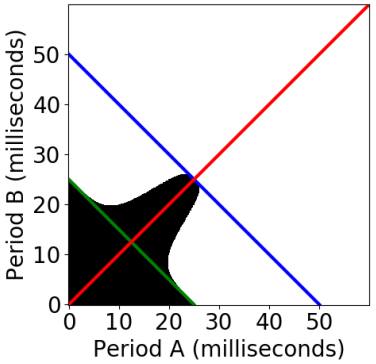
The experiment was simulated with the trained DNN. The mean of last one second outputs for a trained and untrained network have been plotted in Figures 8.7c and 8.7d. It can be seen that the fusion curves for DNN predictions are also enclosed by the lines $x + y = P$ and $x + y = 2P$.

8.4.3 An empirical test with the complex waveform

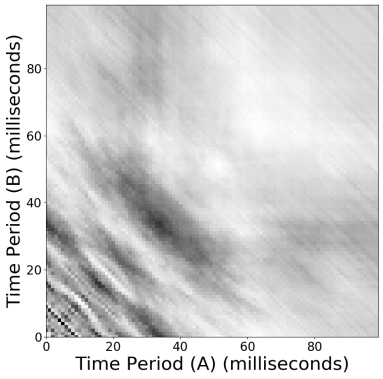
To identify the boundary between the "fusion" and "steady" regions predicted by the neural network in Figure 8.7d, we located datapoints where the value differed from its neighbors by an absolute magnitude greater than 0.5. From this intersection, a test dataset of 100 datapoints was randomly selected; each selected point was located within a maximum distance of five datapoints (horizontally or vertically) from the boundary. These points are visualized in Figure 8.8a.Psychophysics experiments were conducted using the device described in Chapter 7, under the same conditions and with the same subject used for the training data. The subject's classifications, along with their mirror



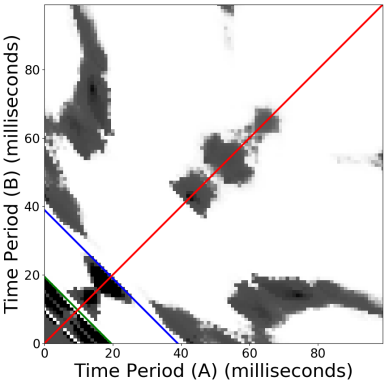
(A) A photic pulse stimulus with different time periods in alternate cycles. The alternate pulses both with PCFs $\frac{1}{2}$ have time periods 0.1 and 0.2 milliseconds.



(B) The shape of fusion curve from previous experiments. The dark regions denotes the classification as steady and white region denoted classification as flickering.

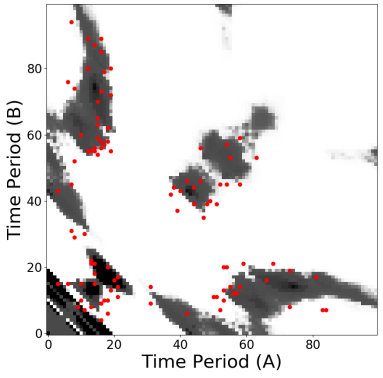


(C) Predictions by an untrained network

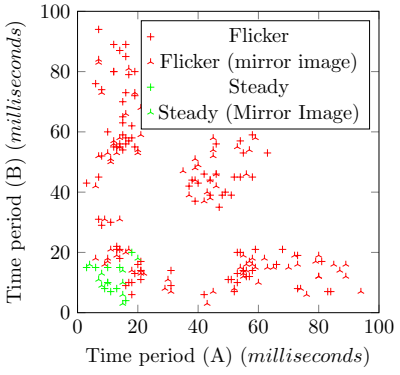


(D) Predictions by a trained network.

FIGURE 8.7: A complex flicker pattern and its psychophysics results from experiments done by Forsyth and Brown[61][84] The red, green and blue lines denote the lines $x = y$, $x + y = P$ and $x + y = 2P$



(A) Selected random points near edges



(B) The results of psychophysics tests on the selected dataponts by he human subject.

FIGURE 8.8: The psychophysics and simulation data for photic pulses with alternate timeperiods. The dark regions denotes the classification as steady and white region denoted classification as flickering.

images across the line $x = y$, are plotted in Figure 8.8b. As illustrated, the subject's psychophysics data aligns more closely with the patterns observed in previous experiments

(Figure 8.7b) than with the model predictions shown in Figure 8.7d.

8.5 Revising the Model

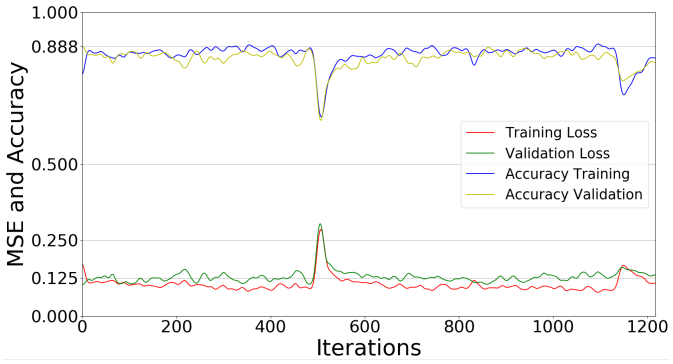
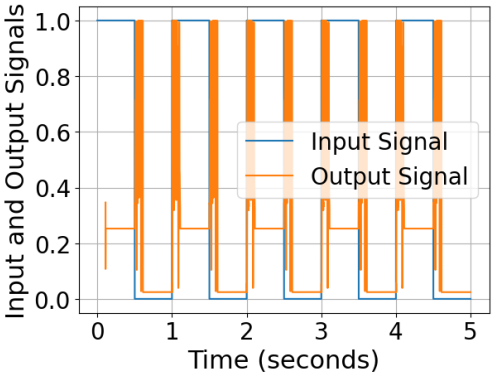


FIGURE 8.9: Training curve of new training with pretrained data

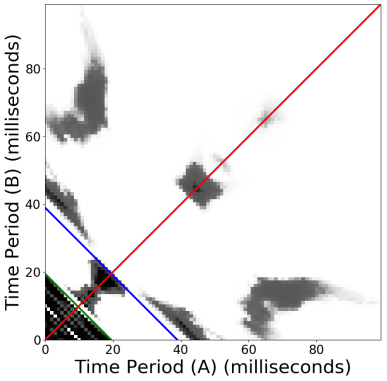
When a scientific theory is falsified, scientists can reject or revise the theory or maintain the theory as is and changing an auxiliary hypothesis[151]. Here the model is improved on further training the pretrained model with a few more conditions. In the new condition, the process of signal generation was the same. However in every minibatch, there was a probability $\frac{1}{2}$ that every wave with label fused was replaced by a new wave. The new wave had equal probability of being from the steady or flicker dataset. The last 0.5 to 0.25 seconds of the signals were all made zero in half of those signals, and a random value between 0 and 1 in the other half. This change was only done in the training set and not validation set. This was done to force the condition that constant stimulus should be classified as steady onto the DNN. The training curve on the pretrained dataset is seen in Figure 8.9 and weights of iteration 958 were loaded to test the model.

8.5.1 Results

The convolution filters were similar to the pretrained model. But now, the output for constant stimulus are closer to zero as seen in 8.10a which was not the case with the previous weights as can be seen in 8.6. The predictions after training are closer to the actual psychophysics results than DNN with previous weights as can be seen in comparison of Figure 8.7b with Figures 8.7d and 8.10b. The network with new weights classified lesser number of photic pulses in flickering region as steady.



(A) Constant Stimulus Output



(B) Complex Stimulus Output

FIGURE 8.10: Training curve and output of revised model

8.6 Conclusion

The present work trained psychophysics data of flicker fusion on a proposed DNN model. The training yielded symmetric filters as found in biological visual systems. The model was put to falsifiability test via an empirical experiment. The psychophysics results did not match exactly with the predictions of the model. But it could be seen that model trained with a dataset forcing the condition that a constant stimulus should be classified steady did a better job of predicting psychophysics results. The model can be improved by bigger training datasets and changing hyperparameters associated with the network. The datasets used for training and testing are in https://github.com/keerthischandran/flickerfusiondata.

Chapter 9

Conclusion

This thesis explored the use of deep learning as a tool to model psychophysics phenomena. Psychophysics experiments can be used to obtain a black box model of the brain. The psychophysics experiments provide a clear relation between a set of input and output signals, where the input signals are the structured representations of stimuli and the output signals are the corresponding behavior responses. The conclusions drawn from this thesis are listed in the next section.

9.1 Inferences drawn from the work

9.1.1 New architectures should be constructed to fit psychophysics data

While many of the existing neural network architectures have shown correlations with brain signals obtained from humans, they are not the perfect models of biological vision. New models, different from engineering goal oriented DNNs need to be built to fit psychophysics data. As has been shown in Chapter 4, two existing neural network architectures VGG16 and VGG19 failed to classify the psychophysics data generated for flicker wheel illusion.

9.1.2 Psychophysics knowledge is needed to design DNNs

This thesis proposed biologically inspired models for flicker fusion, flicker wheel illusion and sound symbolism. The convolution operation in CRNNs proposed for flicker fusion, were inspired by transient on off operations in retina which is mathematically a convolution operation. The appropriate activation function in the lowest layer of DNN for modeling flicker fusion turned out to be the logarithmic function. The logarithmic

activation function was inspired by Fechner's law. The knowledge of microsaccades and the role it plays in perception of various illusory motions, aided in the construction of a proposed model for flicker wheel illusion in Chapter 4. Previous literature had shown existence of phonetic differences in words of English language for objects perceived as sharp or round. This along with the fact that previous psychophysics work showed differences in Fourier spectrum in words labeled to sharp and round Figures in Kiki Bouba experiment, paved a path for designing an experiment to model sound symbolism phenomena through machine learning.

9.1.3 Understanding psychological phenomena can have practical applications

Sound symbolic phenomena is considered to be an aid in the learning of languages in humans. However machine perception of languages do not incorporate sound symbolic process. Building computational models of sound symbolic phenomenon, as done in Chapter 5, can aid in machines understanding and learning natural languages better.

9.1.4 Building DNN models explain other biological phenomena

In Chapter 6, a relation between psychophysics of flicker fusion and EEG response to flicker stimulus was investigated through CRNN based model of flicker fusion. Similarly in chapter 8, we demonstrated that training such models with psychophysics data, can yield insights like the temporal properties of biological filters acting in the temporal dimension.

9.1.5 Structured experimental setups facilitate the acquisition of high-volume data required for DNN training

The thesis mentioned the construction of a device to generate mass data to train and test DNN models in Chapter 7. Visual angle was kept constant throughout the experiment. Intra-subject variations in psychophysics data, that can arise from dark adaptation of the eye was avoided by programming the device to turn on a floodlight in between readings of individual pulses. The results thus obtained showed that psychophysics experiments on flicker fusion are repeatable with greater accuracy without much uncertainty due to intra-subject variation. Such data can be used to reconstruct the accurate computational mechanism of the psychophysics phenomena through a deep learning model. In case of experiments with high intra-subject variation as that on flicker wheel illusion as described in chapter Chapter 4, DNNs were able to predict with similar accuracies that can be obtained with the repetition of the experiment.

9.1.6 DNN based models of psychophysics are falsifiable

A genuine scientific theory must be falsifiable. As shown in Chapter 3, the proposed DNN based model of flicker fusion is falsifiable in many ways. The model can be trained with rectangular waveforms and then be tested with sinusoidal or complex waveforms. The falsification test has been done with data specifically obtained for training Deep Neural Network in Chapter 8. The DNN was be trained with psychophysics data of rectangular waveforms and the data could nearly fit the model. The model made a prediction on a complex waveform which was falsified by doing psychophysics experiment on the prediction.

9.1.7 Corollary

The comprehensive corollary of this research is that the human visual system can be effectively modeled as a hierarchy of discrete, specialized agents, aligning with Minsky's 'Society of Mind' framework. This work demonstrates that DNNs, when biologically constrained, serve as the ideal building blocks for these agents—representing fundamental processes such as cross modal correspondences, flicker perception or motion illusion. Crucially, this implies that artificial general intelligence (AGI) in vision should not be pursued through a single, monolithic architecture, but through the integration of smaller, psychophysically-grounded DNNs. By showing that standard engineering-oriented architectures are non-optimal for human-specific phenomena, this thesis establishes a new design requirement: complex cognitive models must be constructed by stacking or sharing layers between these specialized, 'low-intelligence' agents to truly replicate human perception. Further many of the present DNN architectures could be more similar to animal brains than human brains.

9.2 Future Directions

The methodology explored in the work and articulated in Chapter 2, opens the scope of further research. The proposed CRNN architecture applied on flicker fusion in Chapter 3, can be trained and tested with more psychophysics data obtained from subjects. The chapter Chapter 7 had described the construction of device to gather more psychophysics data on flicker fusion. The correlation of brain signals like EEG, fMRI or MEG with the mid layers of the model can be probed by method described in Chapter 6, or by other techniques like Representational Similarity Analysis. The data gathered for flicker wheel illusion as well as proposed modeling technique as described in Chapter 4 can be applied for modeling more complex visual illusions. Other sound symbolic phenomena unrelated to sharpness and roundness like big and small, can be modeled via the methodology explored in Chapter 5.

9.2.1 Some Preclusive Insights

The findings of this thesis provide several preclusive insights that constrain and guide the direction of future computational psychophysics:

- **Inadequacy of Generic Architectures:** The failure of standard architectures like VGG16 and VGG19 to classify flicker wheel illusions (Chapter 4) precludes the use of generic, task-oriented DNNs for modeling biological vision. Future models *must* prioritize the integration of oculomotor data, such as microsaccades, into the architecture to account for illusory motion.
- **Non-linearity in Temporal Dimension:** The falsification of the flicker fusion model when tested with complex waveforms (Chapter 8 and Chapter 3) precludes the reliance on simple linear temporal integration. This indicates that future models must incorporate non-linear feedback or adaptive biological filters to maintain predictive accuracy across diverse stimulus profiles.
- **Linguistic Perception:** Incorporating sound-symbolic process may help computational models of language acquire human like depth as per the results in Chapter 5. This precludes purely statistical approaches to natural language understanding when the goal is to replicate human-centric learning biases.

Bibliography

[1] Mark A. Georgeson Vicki Bruce Patrick R. Green. *Visual Perception: Physiology, Psychology, and Ecology*. Psychology press, 1996, p. 418.

[2] Nikolaus Kriegeskorte. "Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing". In: *Annual Review of Vision Science* 1.1 (2015). PMID: 28532370, pp. 417–446. DOI: 10.1146/annurev-vision-082114-035447.

[3] Jeffrey S. Bowers et al. "Deep Problems with Neural Network Models of Human Vision". In: *Behavioral and Brain Sciences* (Dec. 2022), pp. 1–74. DOI: 10.1017/s0140525x22002813.

[4] Warren S. McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: *The Bulletin of Mathematical Biophysics* 5.4 (Dec. 1943), pp. 115–133. DOI: 10.1007/bf02478259.

[5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.

[6] D. H. Hubel and T. N. Wiesel. "Receptive fields of single neurones in the cat's striate cortex". In: *The Journal of Physiology* 148.3 (Oct. 1959), pp. 574–591. ISSN: 1469-7793. DOI: 10.1113/jphysiol.1959.sp006308. URL: http://dx.doi.org/10.1113/jphysiol.1959.sp006308.

[7] Guangyu Robert Yang and Xiao-Jing Wang. "Artificial Neural Networks for Neuroscientists: A Primer". In: *Neuron* 107.6 (Sept. 2020), pp. 1048–1070. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2020.09.005.

[8] Kaiming He et al. "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification". In: *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE, Dec. 2015. DOI: 10.1109/iccv.2015.123. URL: http://dx.doi.org/10.1109/ICCV.2015.123.

[9] Jeffrey S. Bowers et al. "On the importance of severely testing deep learning models of cognition". In: *Cognitive Systems Research* 82 (Dec. 2023), p. 101158. ISSN: 1389-0417. DOI: 10.1016/j.cogsys.2023.101158. URL: http://dx.doi.org/10.1016/j.cogsys.2023.101158.

[10] Natalia D. Mankowska et al. "Critical Flicker Fusion Frequency: A Narrative Review". In: *Medicina* 57.10 (2021). ISSN: 1648-9144. DOI: 10.3390/medicina57101096.

[11] Maria Loconsole, Lucia Regolin, and Giorgio Vallortigara. "Matching sounds to shapes: Evidence of the Bouba-Kiki effect in naïve baby chicks". In: *bioRxiv* (2024). DOI: 10.1101/2024.05.17.594640.

[12] Rasmus Bååth, Takeharu Seno, and Akiyoshi Kitaoka. "Cats and Illusory Motion". In: *Psychology* 5.9 (2014), pp. 1131–1134. DOI: 10.4236/psych.2014.59125.

[13] Carney Landis. "Something about Flicker-Fusion". In: *The Scientific Monthly* 73 (1951). URL: https://www.jstor.org/stable/20441.

[14] R. L. Gregory. "Visual illusions classified". In: *Trends in cognitive sciences* 1.5 (1997), pp. 190–194. DOI: 10.1016/S1364-6613(97)01060-7.

[15] C. Q. Howe and D. Purves. *Perceiving geometry: Geometrical illusions explained by natural scene statistics*. Vol. 1. Germany: Springer Science & Business Media, 2005. DOI: 10.1007/b135453.

[16] A. G. Shapiro and D. Todorovic. *The Oxford compendium of visual illusions*. London: Oxford University Press, 2016. DOI: 10.1093/acprof:oso/9780199794607.001.0001.

[17] Y. S. Bonneh, A. Cooperman, and Sagi D. "Motion-induced blindness in normal observers". In: *Nature* 411.6839 (2001), pp. 798–801. DOI: 10.1038/35081073.

[18] Jocelyn Faubert and Andrew M Herbert. "The peripheral drift illusion: A motion illusion in the visual periphery". In: *Perception* 28.5 (1999), pp. 617–621. DOI: 10.1068/p2825.

[19] R. Sokoliuk and R. VanRullen. "The Flickering Wheel Illusion: When Rhythms Make a Static Wheel Flicker". In: *Journal of Neuroscience* 33.33 (Aug. 2013), pp. 13498–13504. DOI: 10.1523/jneurosci.5647-12.2013. URL: https://doi.org/10.1523/jneurosci.5647-12.2013.

[20] Wikimedia Commons. *File:BoobaKiki.png — Wikimedia Commons, the free media repository*. [Online; accessed 7-August-2024]. 2023. URL: https://commons.wikimedia.org/w/index.php?title=File:BoobaKiki.png&oldid=828523312.

[21] John J Ohala, Leanne Hinton, and Johanna Nichols. "Sound symbolism". In: *Proc. 4th Seoul International Conference on Linguistics [SICOL]*. 1997, pp. 98–103.

[22] Kate Pirog Revill et al. "Cross-linguistic sound symbolism and crossmodal correspondence: Evidence from fMRI and DTI". In: *Brain and Language* 128.1 (Jan. 2014), pp. 18–24. ISSN: 0093-934X. DOI: 10.1016/j.bandl.2013.11.002. URL: http://dx.doi.org/10.1016/j.bandl.2013.11.002.

[23] Vilayanur S Ramachandran and Edward M Hubbard. "Synaesthesia–a window into perception, thought and language". In: *Journal of consciousness studies* 8.12 (2001), pp. 3–34.

[24] J.C.A. Read. "The place of human psychophysics in modern neuroscience". In: *Neuroscience* 296 (June 2015), pp. 116–129. DOI: 10.1016/j.neuroscience.2014.05.036.

[25] Thomas Young. "The Bakerian Lecture. On the theory of light and colours". In: *Philosophical Transactions of the Royal Society of London* 92 (Dec. 1802), pp. 12–48. ISSN: 2053-9223. DOI: 10.1098/rstl.1802.0004.

[26] Andrew John Anderson and Algis Jonas Vingrys. "Small Samples: Does Size Matter". In: *Investigative Ophthalmology & Visual Science* 42.7 (June 2001), pp. 1411–1413. ISSN: 1552-5783. eprint: https://arvojournals.org/arvo/content_public/journal/iovs/933588/7g060101411.pdf.

[27] Wikimedia Commons. *File:Anatomy of neuron.png — Wikimedia Commons, the free media repository*. [Online; accessed 7-August-2024]. 2023. URL: https://commons.wikimedia.org/w/index.php?title=File:Anatomy_of_neuron.png&oldid=791957473.

[28] Wikimedia Commons. *File:Artificial Neuron.svg — Wikimedia Commons, the free media repository*. [Online; accessed 7-August-2024]. 2023. URL: https://commons.wikimedia.org/w/index.php?title=File:Artificial_Neuron.svg&oldid=804418691.

[29] S. Marĉelja. "Mathematical description of the responses of simple cortical cells*". In: *Journal of the Optical Society of America* 70.11 (Nov. 1980), p. 1297. ISSN: 0030-3941. DOI: 10.1364/josa.70.001297. URL: http://dx.doi.org/10.1364/JOSA.70.001297.

[30] Radoslaw Martin Cichy et al. "Comparison of deep neural networks to spatio-temporal cortical dynamics of human visual object recognition reveals hierarchical correspondence". In: *Scientific Reports* 6.1 (June 2016). ISSN: 2045-2322. DOI: 10.1038/srep27755. URL: http://dx.doi.org/10.1038/srep27755.

[31] Santiago A. Cadena et al. "Deep convolutional models improve predictions of macaque V1 responses to natural images". In: *PLOS Computational Biology* 15.4 (Apr. 2019), pp. 1–27. DOI: 10.1371/journal.pcbi.1006897.

[32] A. Gomez-Villa et al. "Color illusions also 12eive CNNs for low-level vision tasks: Analysis and implications". In: *Vision Research* 176 (Nov. 2020), pp. 156–174. DOI: 10.1016/j.visres.2020.07.010. URL: https://doi.org/10.1016/j.visres.2020.07.010.

[33] Hongtao Zhang and Shinichi Yoshida. "Exploring Deep Neural Networks in Simulating Human Vision through Five Optical Illusions". In: *Applied Sciences* 14.8 (Apr. 2024), p. 3429. ISSN: 2076-3417. DOI: 10.3390/app14083429. URL: http://dx.doi.org/10.3390/app14083429.

[34] Nikolaus Kriegeskorte and Pamela K. Douglas. "Cognitive computational neuroscience". In: *Nature Neuroscience* 21.9 (Aug. 2018), pp. 1148–1160. DOI: 10.1038/s41593-018-0210-5. URL: https://doi.org/10.1038/s41593-018-0210-5.

[35] D. Zipser and R. A. Andersen. "A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons". In: *Nature* 331.6158 (Feb. 1988), pp. 679–684. ISSN: 0028-0836. DOI: 10.1038/331679a0.

[36] Jörn Diedrichsen and Nikolaus Kriegeskorte. "Representational models: A common framework for understanding encoding, pattern-component, and representational-similarity analysis". In: *PLOS Computational Biology* 13.4 (Apr. 2017). Ed. by Radoslaw Cichy, e1005508. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1005508. URL: http://dx.doi.org/10.1371/journal.pcbi.1005508.

[37] Lin Wang et al. "Neural Evidence for the Prediction of Animacy Features during Language Comprehension: Evidence from MEG and EEG Representational Similarity Analysis". In: *The Journal of Neuroscience* 40.16 (Mar. 2020), pp. 3278–3291. ISSN: 1529-2401. DOI: 10.1523/jneurosci.1733-19.2020. URL: http://dx.doi.org/10.1523/JNEUROSCI.1733-19.2020.

[38] "Comparison of neuronal responses in primate inferior temporal cortex and feed forward deep neural network model with regard to information processing of faces". In: *J Comput Neuroscience* (2021). DOI: https://doi.org/10.1007/s10827-021-00778-5.

[39] Seyed-Mahdi Khaligh-Razavi et al. "Fixed versus mixed RSA: Explaining visual representations by fixed and mixed feature sets from shallow and deep computational models". In: *Journal of Mathematical Psychology* 76 (Feb. 2017), pp. 184–197. DOI: 10.1016/j.jmp.2016.10.007. URL: https://doi.org/10.1016/j.jmp.2016.10.007.

[40] Bruce G. Buchanan. "A (Very) Brief History of Artificial Intelligence". In: *AI Magazine* 26.4 (Dec. 2005), p. 53. DOI: 10.1609/aimag.v26i4.1848. URL: https://ojs.aaai.org/index.php/aimagazine/article/view/1848.

[41] Gerard O'Regan. *A Brief History of Computing*. Springer International Publishing, 2021. DOI: 10.1007/978-3-030-66599-9. URL: https://doi.org/10.1007/978-3-030-66599-9.

[42] A. M. TURING. "I.—COMPUTING MACHINERY AND INTELLIGENCE". In: *Mind* LIX.236 (Oct. 1950), pp. 433–460. DOI: 10.1093/mind/lix.236.433. URL: https://doi.org/10.1093/mind/lix.236.433.

[43] R. M. Friedberg. "A Learning Machine: Part I". In: *IBM Journal of Research and Development* 2.1 (Jan. 1958), pp. 2–13. DOI: 10.1147/rd.21.0002. URL: https://doi.org/10.1147/rd.21.0002.

[44] Howard Campaigne. "Some experiments in machine learning". In: *Papers presented at the the March 3-5, 1959, western joint computer conference on XX - IRE-AIEE-ACM '59 (Western)*. ACM Press, 1959. DOI: 10.1145/1457838.1457868. URL: https://doi.org/10.1145/1457838.1457868.

[45] Hanan Rosemarin and Ariel Rosenfeld. "Playing Chess at a Human Desired Level and Style". In: *Proceedings of the 7th International Conference on Human-Agent Interaction*. HAI '19. ACM, Sept. 2019. DOI: 10.1145/3349537.3351904. URL: http://dx.doi.org/10.1145/3349537.3351904.

[46] Yuki Kubota, Atsushi Hiyama, and Masahiko Inami. "A Machine Learning Model Perceiving Brightness Optical Illusions: Quantitative Evaluation with Psychophysical Data". In: *Augmented Humans Conference 2021*. AHs '21. ACM, Feb. 2021. DOI: 10.1145/3458709.3458952. URL: http://dx.doi.org/10.1145/3458709.3458952.

[47] Marius Jahrens and Thomas Martinetz. "Solving Raven's Progressive Matrices with Multi-Layer Relation Networks". In: *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, July 2020. DOI: 10.1109/ijcnn48605.2020.9207319. URL: http://dx.doi.org/10.1109/ijcnn48605.2020.9207319.

[48] Joël Fagot and Masaki Tomonaga. "Global and local processing in humans (Homo sapiens) and chimpanzees (Pan troglodytes): Use of a visual search task with compound stimuli." In: *Journal of Comparative Psychology* 113.1 (Mar. 1999), pp. 3–12. ISSN: 0735-7036. DOI: 10.1037/0735-7036.113.1.3. URL: http://dx.doi.org/10.1037/0735-7036.113.1.3.

[49] Tomoko Imura and Masaki Tomonaga. "Differences between chimpanzees and humans in visual temporal integration". In: *Scientific Reports* 3.1 (Nov. 2013). ISSN: 2045-2322. DOI: 10.1038/srep03256. URL: http://dx.doi.org/10.1038/srep03256.

[50] Fabian A. Soto and Edward A. Wasserman. "Mechanisms of object recognition: what we have learned from pigeons". In: *Frontiers in Neural Circuits* 8 (Oct. 2014). ISSN: 1662-5110. DOI: 10.3389/fncir.2014.00122. URL: http://dx.doi.org/10.3389/fncir.2014.00122.

[51] W. WESTLAKE. "Is a one eyed racing driver safe to compete? Formula one (eye) or two?" In: *British Journal of Ophthalmology* 85.5 (May 2001), pp. 619–624. ISSN: 0007-1161. DOI: 10.1136/bjo.85.5.619. URL: http://dx.doi.org/10.1136/bjo.85.5.619.

[52] Brian Timney. "The effects of early and late monocular deprivation on binocular depth perception in cats". In: *Developmental Brain Research* 7.2–3 (Apr. 1983), pp. 235–243. ISSN: 0165-3806. DOI: 10.1016/0165-3806(83)90180-3. URL: http://dx.doi.org/10.1016/0165-3806(83)90180-3.

[53] Yuki Onishi et al. "An automated fruit harvesting robot by using deep learning". In: *ROBOMECH Journal* 6.1 (Nov. 2019). ISSN: 2197-4225. DOI: 10.1186/s40648-019-0141-2. URL: http://dx.doi.org/10.1186/s40648-019-0141-2.

[54] Rui Fan et al. "Computer Stereo Vision for Autonomous Driving: Theory and Algorithms". In: *Recent Advances in Computer Vision Applications Using Parallel Processing*. Springer International Publishing, 2023, pp. 41–70. ISBN: 9783031187353. DOI: 10.1007/978-3-031-18735-3_3. URL: http://dx.doi.org/10.1007/978-3-031-18735-3_3.

[55] Marvin L Minsky. *Socity of Mind*. Simon & Schuster, Feb. 1987.

[56] Andrew Saxe, Stephanie Nelli, and Christopher Summerfield. "If deep learning is the answer, what is the question?" In: *Nature Neuroscience* 22.1 (Nov. 2020), pp. 55–67. DOI: 10.1038/s41583-020-00395-8.

[57] Carney Landis. "Determinants of the Critical Flicker-Fusion Threshold". In: *Physiological Reviews* 34.2 (Apr. 1954), pp. 259–286. DOI: 10.1152/physrev.1954.34.2.259. URL: https://doi.org/10.1152/physrev.1954.34.2.259.

[58] Min xia Liu et al. "The Research and Analysis of Factors Affecting Critical Flicker Frequency". In: *Procedia Manufacturing* 3 (2015), pp. 4279–4286. DOI: 10.1016/j.promfg.2015.07.417.

[59] S. Howard Bartley. "Some Factors Influencing Critical Flicker Frequency". In: *The Journal of Psychology* 46.1 (1958), pp. 107–115. DOI: 10.1080/00223980.1958.9916274.

[60] D. H. Kelly. "Flicker". In: *Jameson D., Hurvich L.M. (eds) Visual Psychophysics. Handbook of Sensory Physiology*, (1972). DOI: https://doi.org/10.1007/978-3-642-88658-4_11.

[61] D. M. Forsyth and Charles R. Brown. "Flicker Contours for Intermittent Photic Stimuli of Alternating Duration∗†". In: *J. Opt. Soc. Am.* 49.8 (Aug. 1959), pp. 760–763. DOI: 10.1364/JOSA.49.000760.

[62] Thomas M. Nelson, S. Howard Bartley, and Earl S. Harper. "CFF for Short Trains of Photic Stimulation having Various Temporal Distributions and Separations". In: *The Journal of Psychology* 58.2 (1964), pp. 333–341. DOI: 10.1080/00223980.1964.9916753.

[63] Aaron R. Seitz et al. "Perceptual Learning of Motion Leads to Faster Flicker Perception". In: *PLoS ONE* 1.1 (Dec. 2006). Ed. by Martin Giurfa, e28. DOI: 10.1371/journal.pone.0000028.

[64] Auria Eisen-Enosh et al. "Evaluation of Critical Flicker-Fusion Frequency Measurement Methods for the Investigation of Visual Temporal Resolution". In: *Scientific Reports* 7.1 (Nov. 2017). DOI: 10.1038/s41598-017-15034-z.

[65] Christoph S. Herrmann. "Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena". In: *Experimental Brain Research* 137 (2001), pp. 346–353. DOI: 10.1007/s002210100682.

[66] Katharina Duecker et al. "No Evidence for Entrainment: Endogenous Gamma Oscillations and Rhythmic Flicker Responses Coexist in Visual Cortex". In: *Journal of Neuroscience* 41.31 (2021), pp. 6684–6698. ISSN: 0270-6474. DOI: 10.1523/JNEUROSCI.3134-20.2021.

[67] David Carmel, Nilli Lavie, and Geraint Rees. "Conscious Awareness of Flicker in Humans Involves Frontal and Parietal Cortex". In: *Current Biology* 16.9 (May 2006), pp. 907–911. DOI: 10.1016/j.cub.2006.03.055. URL: https://doi.org/10.1016/j.cub.2006.03.055.

[68] Moshe Gur and D. M. Snodderly. "A Dissociation Between Brain Activity and Perception: Chromatically Opponent Cortical Neurons Signal Chromatic Flicker that is not Perceived". In: *Vision Research* 37.4 (Feb. 1997), pp. 377–382. DOI: 10.1016/s0042-6989(96)00183-6.

[69] Yi Jiang, Ke Zhou, and Sheng He. "Human visual cortex responds to invisible chromatic flicker". In: *Nature Neuroscience* 10.5 (Apr. 2007), pp. 657–662. DOI: 10.1038/nn1879.

[70] George Sperling and Man Mohan Sondhi. "Model for Visual Luminance Discrimination and Flicker Detection". In: *Journal of the Optical Society of America* 58.8 (Aug. 1968), p. 1133. DOI: 10.1364/josa.58.001133.

[71] John Z. Levinson. "Flicker Fusion Phenomena". In: *Science* 160.3823 (1968), pp. 21–28. ISSN: 00368075, 10959203. DOI: 10.1126/science.160.3823.21.

[72] Shun-ichi Amari, Israel Lieblich, and Arthur I. Karshmer. "A Neural Model for the Handling of Phenomena Associated with Trains of Light Stimuli: An Updated Version to Fit Fusion Data". In: *Systems Neuroscience*. Elsevier, 1977, pp. 55–66. DOI: 10.1016/b978-0-12-491850-4.50007-6.

[73] Razieh Falahian et al. "Artificial neural network-based modeling of brain response to flicker light". In: *Nonlinear Dynamics* 81.4 (May 2015), pp. 1951–1967. DOI: 10.1007/s11071-015-2118-x.

[74] Maciej Labecki et al. "Nonlinear Origin of SSVEP Spectra—A Combined Experimental and Modeling Study". In: *Frontiers in Computational Neuroscience* 10 (2016), p. 129. ISSN: 1662-5188. DOI: 10.3389/fncom.2016.00129.

[75] Azar Abiyev, Funda Datlı Yakaryılmaz, and Zeynel Abidin Öztürk. "A new diagnostic approach in Alzheimer's disease: The critical flicker fusion threshold". In: *Dement Neuropsychol* 16.1 (Mar. 2022), pp. 89–96. ISSN: 1980-5764. DOI: 10.1590/1980-5764-dn-2021-0054. URL: http://dx.doi.org/10.1590/1980-5764-DN-2021-0054.

[76] S. Howard Bartley and Thomas M. Nelson. "Equivalence of Various Pulse-to-Cycle Fractions in Producing Critical Flicker Frequency∗". In: *Journal of the Optical Society of America* 50.3 (Mar. 1960), p. 241. DOI: 10.1364/josa.50.000241.

[77] Zhen Zuo et al. "Convolutional recurrent neural networks: Learning spatial dependencies for image representation". In: *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2015, pp. 18–26. DOI: 10.1109/CVPRW.2015.7301268.

[78] Christopher W. Tyler and Russell D. Hamer. "Analysis of visual modulation sensitivity IV Validity of the Ferry–Porter law". In: *Journal of the Optical Society of America A* 7.4 (Apr. 1990), p. 743. DOI: 10.1364/josaa.7.000743.

[79] Gustav Theodor Fechner. *Elements of psychophysics*. Trans. by Herbert Sidney Langfeld. Vol. 1. Holt, Rinehart and Winston New York, 1912. URL: https://lru.praxis.dk/Lru/microsites/hvadermatematik/hem2download/kap4_projekt_4_9_ekstra_Fechners_originale_vaerk.pdf.

[80] Bernard Moulden, Judy Renshaw, and George Mather. "Two Channels for Flicker in the Human Visual System". In: *Perception* 13.4 (Aug. 1984), pp. 387–400. DOI: 10.1068/p130387.

[81] Michiel Van Wyk, Heinz Wässle, and W. Rowland Taylor. "Receptive field properties of ON- and OFF- ganglion cells in the mouse retina". In: *Visual Neuroscience* 26.3 (May 2009), pp. 297–308. DOI: 10.1017/s0952523809990137.

[82] D. H. Kelly. "Visual Responses to Time-Dependent Stimuli∗ I Amplitude Sensitivity Measurements†". In: *Journal of the Optical Society of America* 51.4 (Apr. 1961), p. 422. DOI: 10.1364/josa.51.000422.

[83] D. H. Kelly. "Visual Responses to Time-Dependent Stimuli∗ III Individual Variations†". In: *Journal of the Optical Society of America* 52.1 (Jan. 1962), p. 89. DOI: 10.1364/josa.52.000089.

[84] Charles R. Brown and D. M. Forsyth. "Fusion Contour for Intermittent Photic Stimuli of Alternating Duration". In: *Science* 129.3346 (Feb. 1959), pp. 390–391. DOI: 10.1126/science.129.3346.390.

[85] Günter Rager and Wolf Singer. "The response of cat visual cortex to flicker stimuli of variable frequency". In: *European Journal of Neuroscience* 10.5 (May 1998), pp. 1856–1877. DOI: 10.1046/j.1460-9568.1998.00197.x.

[86] L. H Van Der Tweel. "Relation Between Psychophysics and Electrophysiology of Flicker". In: *Doc Ophthalmol* 18 (1964), pp. 287–304. DOI: 10.1007/BF00160581.

[87] Giuseppe Casalino. "[INVITED] Computational intelligence for smart laser materials processing". In: *Optics & Laser Technology* 100 (Mar. 2018), pp. 165–175. DOI: 10.1016/j.optlastec.2017.10.011. URL: https://doi.org/10.1016/j.optlastec.2017.10.011.

[88] Eiji Watanabe et al. "Illusory Motion Reproduced by Deep Neural Networks Trained for Prediction". In: *Frontiers in Psychology* 9 (Mar. 2018). DOI: 10.3389/fpsyg.2018.00345. URL: https://doi.org/10.3389/fpsyg.2018.00345.

[89] Christina M. Funke et al. "Five points to check when comparing visual perception in humans and machines". In: *Journal of Vision* 21.3 (Mar. 2021), p. 16. DOI: 10.1167/jov.21.3.16. URL: https://doi.org/10.1167/jov.21.3.16.

[90] Bhanuka Mahanama et al. "Eye movement and pupil measures: A review". In: *Frontiers in Computer Science* 3 (2022), p. 733531. DOI: 10.3389/fcomp.2021.733531.

[91] Martin Rolfs. "Microsaccades: Small steps on a long way". In: *Vision Research* 49.20 (Oct. 2009), pp. 2415–2441. DOI: 10.1016/j.visres.2009.08.010. URL: https://doi.org/10.1016/j.visres.2009.08.010.

[92] J. Otero-Millan, S. L. Macknik, and S. Martinez-Conde. "Microsaccades and Blinks Trigger Illusory Rotation in the "Rotating Snakes" Illusion". In: *Journal of Neuroscience* 32.17 (Apr. 2012), pp. 6043–6051. DOI: 10.1523/jneurosci.5823-11.2012. URL: https://doi.org/10.1523/jneurosci.5823-11.2012.

[93] Xoana G. Troncoso et al. "Microsaccades drive illusory motion in the iEnigma/i illusion". In: *Proceedings of the National Academy of Sciences* 105.41 (Oct. 2008), pp. 16033–16038. DOI: 10.1073/pnas.0709389105. URL: https://doi.org/10.1073/pnas.0709389105.

[94] Matthias H Hennig and Florentin Wörgötter. "Effects of fixational eye movements on retinal ganglion cell responses: A modelling study". In: *Frontiers in Computational Neuroscience* (2007), p. 2. DOI: 10.3389/neuro.10.002.2007.

[95] Kristian Donner and Simo Hemilä. "Modelling the effect of microsaccades on retinal responses to stationary contrast patterns". In: *Vision research* 47.9 (2007), pp. 1166–1177. DOI: 10.1016/j.visres.2006.11.024.

[96] Bruce C. Hansen and Edward A. Essock. "A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes". In: *Journal of Vision* 4.12 (Dec. 2004), p. 5. ISSN: 1534-7362. DOI: 10.1167/4.12.5. URL: http://dx.doi.org/10.1167/4.12.5.

[97] Kenneth Gaarder. "Transmission of Edge Information in the Human Visual System". In: *Nature* 212.5059 (Oct. 1966), pp. 321–323. DOI: 10.1038/212321a0. URL: https://doi.org/10.1038/212321a0.

[98] Lynn Schmittwilken and Marianne Maertens. "Fixational eye movements enable robust edge detection". In: *Journal of Vision* 22.8 (July 2022), p. 5. DOI: 10.1167/jov.22.8.5. URL: https://doi.org/10.1167/jov.22.8.5.

[99] Jurn-Gyu Park et al. "Patch and Model Size Characterization for On-Device Efficient-ViTs on Small Datasets Using 12 Quantitative Metrics". In: *IEEE Access* 13 (2025), pp. 25704–25722. DOI: 10.1109/ACCESS.2025.3536471.

[100] Alex Gomez-Villa et al. "On the synthesis of visual illusions using deep generative models". In: *Journal of Vision* 22.8 (July 2022), p. 2. DOI: 10.1167/jov.22.8.2. URL: https://doi.org/10.1167/jov.22.8.2.

[101] Konstantina Margiotoudi and Friedemann Pulvermüller. "Action sound–shape congruencies explain sound symbolism". In: *Scientific reports* 10.1 (2020), p. 12706. DOI: 10.1038/s41598-020-69528-4.

[102] Wolfgang Köhler. "Gestalt psychology". In: *Psychologische Forschung* 31.1 (1967), pp. XVIII–XXX.

[103] Konstantina Margiotoudi et al. "Bo-NO-bouba-kiki: picture-word mapping but no spontaneous sound symbolic speech-shape mapping in a language trained bonobo". In: *Proceedings of the Royal Society B: Biological Sciences* 289.1968 (Feb. 2022). DOI: 10.1098/rspb.2021.1717. URL: https://doi.org/10.1098/rspb.2021.1717.

[104] Aleksandra Ćwiek et al. "Novel vocalizations are understood across cultures". In: *Scientific Reports* 11.1 (2021), p. 10108. DOI: 10.1038/s41598-021-89445-4.

[105] Marcus Perlman, Rick Dale, and Gary Lupyan. "Iconicity can ground the creation of vocal symbols". In: *Royal Society open science* 2.8 (2015), p. 150152. DOI: 10.1098/rsos.150152.

[106] Johannes Dellert et al. "Preferred sound groups of vocal iconicity reflect evolutionary mechanisms of sound stability and first language acquisition: evidence from Eurasia". In: *Philosophical Transactions of the Royal Society B* 376.1824 (2021), p. 20200190. DOI: 10.1098/rstb.2020.0190.

[107] Niklas Erben Johansson, Jon W Carr, and Simon Kirby. "Cultural evolution leads to vocal iconicity in an experimental iterated learning task". In: *Journal of Language Evolution* 6.1 (2021), pp. 1–25. DOI: 10.1093/jole/lzab001.

[108] David Vinson et al. "Iconicity emerges and is maintained in spoken language." In: *Journal of Experimental Psychology: General* 150.11 (2021), p. 2293. DOI: 10.1037/xge0001024.

[109] Aleksandra Ćwiek et al. "The bouba/kiki effect is robust across cultures and writing systems". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 377.1841 (Nov. 2021). DOI: 10.1098/rstb.2020.0390. URL: https://doi.org/10.1098/rstb.2020.0390.

[110] David M. Sidhu and Penny M. Pexman. "Five mechanisms of sound symbolic association". In: *Psychonomic Bulletin & Review* 25.5 (Aug. 2017), pp. 1619–1643. DOI: 10.3758/s13423-017-1361-1. URL: https://doi.org/10.3758/s13423-017-1361-1.

[111] Andreas Holzinger. "Introduction to MAchine Learning & Knowledge Extraction (MAKE)". In: *Machine Learning and Knowledge Extraction* 1.1 (July 2017), pp. 1–20. DOI: 10.3390/make1010001. URL: https://doi.org/10.3390/make1010001.

[112] Shams Forruque Ahmed et al. "Deep learning modelling techniques: current progress, applications, advantages, and challenges". In: *Artificial Intelligence Review* 56.11 (Apr. 2023), pp. 13521–13617. ISSN: 1573-7462. DOI: 10.1007/s10462-023-10466-8. URL: http://dx.doi.org/10.1007/s10462-023-10466-8.

[113] David M. Sidhu et al. "Sound symbolism shapes the English language: The maluma/takete effect in English nouns". In: *Psychonomic Bulletin & Review* 28.4 (Apr. 2021), pp. 1390–1398. DOI: 10.3758/s13423-021-01883-3. URL: https://doi.org/10.3758/s13423-021-01883-3.

[114] Ananya Passi and S. P. Arun. "The Bouba–Kiki effect is predicted by sound properties but not speech properties". In: *Attention, Perception, & Psychophysics* 86.3 (Dec. 2022), pp. 976–990. ISSN: 1943-393X. DOI: 10.3758/s13414-022-02619-8. URL: http://dx.doi.org/10.3758/s13414-022-02619-8.

[115] Yesenia Taveras-Cruz, Jingyi He, and Rhea T. Eskew. "Visual psychophysics: Luminance and color". In: *Progress in Brain Research*. Elsevier, 2022, pp. 231–256. DOI: 10.1016/bs.pbr.2022.04.004. URL: https://doi.org/10.1016/bs.pbr.2022.04.004.

[116] Karin Schwab et al. "Alpha entrainment in human electroencephalogram and magnetoencephalogram recordings". In: *NeuroReport* 17.17 (Nov. 2006), pp. 1829–1833. DOI: 10.1097/01.wnr.0000246326.89308.ec. URL: https://doi.org/10.1097/01.wnr.0000246326.89308.ec.

[117] Almudena Capilla et al. "Steady-State Visual Evoked Potentials Can Be Explained by Temporal Superposition of Transient Event-Related Responses". In: *PLoS ONE* 6.1 (Jan. 2011). Ed. by Pedro Antonio Valdes-Sosa, e14543. DOI: 10.1371/journal.pone.0014543. URL: https://doi.org/10.1371/journal.pone.0014543.

[118] Annika Notbohm, Jürgen Kurths, and Christoph S. Herrmann. "Modification of Brain Oscillations via Rhythmic Light Stimulation Provides Evidence for Entrainment but Not for Superposition of Event-Related Responses". In: *Frontiers in Human Neuroscience* 10 (Feb. 2016). DOI: 10.3389/fnhum.2016.00010. URL: https://doi.org/10.3389/fnhum.2016.00010.

[119] F H Lopes da Silva et al. "Model of brain rhythmic activity". In: *Kybernetik* 15 (1974), pp. 27–37. DOI: 10.1007/BF00270757.

[120] P A Robinson et al. "Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials". In: *Phys. Rev. E* 63 (2001). DOI: 10.1103/PhysRevE.63.041909.

[121] J.A. Roberts and P.A. Robinson. "Quantitative theory of driven nonlinear brain dynamics". In: *NeuroImage* 62.3 (2012), pp. 1947–1955. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2012.05.054.

[122] "Theory of edge detection". In: *Proceedings of the Royal Society of London. Series B. Biological Sciences* 207.1167 (Feb. 1980), pp. 187–217. DOI: 10.1098/rspb.1980.0020. URL: https://doi.org/10.1098/rspb.1980.0020.

[123] Eduardo A.B. da Silva and Gelson V. Mendonça. "Digital Image Processing". In: *The Electrical Engineering Handbook*. Elsevier, 2005, pp. 891–910. DOI: 10.1016/b978-012170960-0/50064-5. URL: https://doi.org/10.1016/b978-012170960-0/50064-5.

[124] Alba Jiménez et al. "Principles, mechanisms and functions of entrainment in biological oscillators". In: *Interface Focus* 12.3 (Apr. 2022). DOI: 10.1098/rsfs.2021.0088. URL: https://doi.org/10.1098/rsfs.2021.0088.

[125] Kuno Kirschfeld. "The physical basis of alpha waves in the electroencephalogram and the origin of the ?Berger effect?" In: *Biological Cybernetics* 92.3 (Feb. 2005), pp. 177–185. DOI: 10.1007/s00422-005-0547-1. URL: https://doi.org/10.1007/s00422-005-0547-1.

[126] Rasa Gulbinaite et al. "Individual Alpha Peak Frequency Predicts 10 Hz Flicker Effects on Selective Attention". In: *The Journal of Neuroscience* 37.42 (Sept. 2017), pp. 10173–10184. DOI: 10.1523/jneurosci.1163-17.2017.

[127] J. Ding, R. Srinivasan, and G. Sperling. "Flicker elicits eeg responses in two distinct cortical networks depending on attention and flicker frequency". In: *Journal of Vision* 6.6 (Mar. 2006), pp. 515–515. DOI: 10.1167/6.6.515.

[128] Rasa Gulbinaite, Diane H.M. Roozendaal, and Rufin VanRullen. "Attention differentially modulates the amplitude of resonance frequencies in the visual cortex". In: *NeuroImage* 203 (2019), p. 116146. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2019.116146.

[129] Aurelien Geron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. 2nd. O'Reilly Media, Inc., 2019. ISBN: 1492032646.

[130] Yanli Yang. "A Signal Theoretic Approach for Envelope Analysis of Real-Valued Signals". In: *IEEE Access* 5 (2017), pp. 5623–5630. ISSN: 2169-3536. DOI: 10.1109/access.2017.2688467. URL: http://dx.doi.org/10.1109/ACCESS.2017.2688467.

[131] Florentina Soto et al. "Efficient Coding by Midget and Parasol Ganglion Cells in the Human Retina". In: *Neuron* 107.4 (Aug. 2020), 656–666.e5. DOI: 10.1016/j.neuron.2020.05.030.

[132] Ronald J. Tocci, Neal S. Widmer, and Gregory L. Moss. *Digital systems: Principles and applications*. Pearson Education, 2007, pp. 250–251. ISBN: 0-13-173969-7.

[133] Maciej Labecki, Maria Malgorzata Nowicka, and Piotr Suffczynski. "Temporal Modulation of Steady-State Visual Evoked Potentials". In: *International Journal of Neural Systems* 29.03 (Mar. 2019), p. 1850050. DOI: 10.1142/s0129065718500508.

[134] Giovanni B. Rossi and Birgitta Berglund. "Measurement involving human perception and interpretation". In: *Measurement* 44.5 (June 2011), pp. 815–822. DOI: 10.1016/j.measurement.2011.01.016. URL: https://doi.org/10.1016/j.measurement.2011.01.016.

[135] Billy R. Wooten et al. "A practical method of measuring the human temporal contrast sensitivity function". In: *Biomedical Optical Express* 1.1 (Aug. 2010), pp. 47–58. DOI: 10.1364/BOE.1.000047. URL: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-1-1-47.

[136] Michael Bach, Thomas Meigen, and Hans Strasburger. "Raster-scan cathode-ray tubes for vision research-limits of resolution in space, time and intensity, and some solutions". In: *Spatial Vision* 10.4 (1997), pp. 403–414. ISSN: 1568-5683. DOI: 10.1163/156856897x00311. URL: http://dx.doi.org/10.1163/156856897x00311.

[137] Auria Eisen-Enosh et al. "Evaluation of Critical Flicker-Fusion Frequency Measurement Methods for the Investigation of Visual Temporal Resolution". In: *Scientific Reports* 7 (2017). DOI: https://doi.org/10.1038/s41598-017-15034-z.

[138] Kwok K. Ng S. M. Sze. "Physics of Semiconductor Devices". In: John Wiley and sons, 2007, p. 601.

[139] T. P. Lee. "Effect of junction capacitance on the rise time of led's and on the turn-on delay of injection lasers". In: *The Bell System Technical Journal* 54.1 (1975), pp. 53–68. DOI: 10.1002/j.1538-7305.1975.tb02825.x.

[140] Yusuf Abdullahi Badamasi. "The working principle of an Arduino". In: *2014 11th International Conference on Electronics, Computer and Computation (ICECCO)*. 2014, pp. 1–4. DOI: 10.1109/ICECCO.2014.6997578.

[141] Anthony Lin. "Binary search algorithm". In: *WikiJournal of Science* 2.1 (2019), p. 5. DOI: 10.15347/wjs/2019.005. URL: https://doi.org/10.15347/wjs/2019.005.

[142] M.D. Binder, N. Hirokawa, and N. Windhorst. "Method of Constant Stimuli". In: *Encyclopedia of Neuroscience*. Berlin, Heidelberg: Springer, 2009. DOI: 10.1007/978-3-540-29678-2_3464.

[143] Yibin Tian et al. "Biological Basis and Computer Vision Applications of Image Phase Congruency: A Comprehensive Survey". In: *Biomimetics* 9.7 (July 2024), p. 422. ISSN: 2313-7673. DOI: 10.3390/biomimetics9070422. URL: http://dx.doi.org/10.3390/biomimetics9070422.

[144] Richard Young, Ronald Lesperance, and W. Weston Meyer. "The Gaussian Derivative model for spatial-temporal vision: I. Cortical model". In: *Spatial Vision* 14.3–4 (2001), pp. 261–319. ISSN: 1568-5683. DOI: 10.1163/156856801753253582. URL: http://dx.doi.org/10.1163/156856801753253582.

[145] Ronald Lesperance and Richard Young. "The Gaussian Derivative model for spatial-temporal vision: II. Cortical data". In: *Spatial Vision* 14.3–4 (2001), pp. 321–389. ISSN: 1568-5683. DOI: 10.1163/156856801753253591. URL: http://dx.doi.org/10.1163/156856801753253591.

[146] Bruno A. Olshausen and David J. Field. "Sparse coding with an overcomplete basis set: A strategy employed by V1?" In: *Vision Research* 37.23 (Dec. 1997), pp. 3311–3325. ISSN: 0042-6989. DOI: 10.1016/s0042-6989(97)00169-7. URL: http://dx.doi.org/10.1016/S0042-6989(97)00169-7.

[147] Richard A. Young. "Gaussian derivative model for machine vision". In: *Annual Meeting Optical Society of America*. OAM. Optica Publishing Group, 1985. DOI: 10.1364/oam.1985.wj40. URL: http://dx.doi.org/10.1364/OAM.1985.WJ40.

[148] R. Mehrotra, K.R. Namuduri, and N. Ranganathan. "Gabor filter-based edge detection". In: *Pattern Recognition* 25.12 (Dec. 1992), pp. 1479–1494. ISSN: 0031-3203. DOI: 10.1016/0031-3203(92)90121-x. URL: http://dx.doi.org/10.1016/0031-3203(92)90121-X.

[149] Karl Popper. *The logic of scientific discovery*. Routledge Classics (Hardcover). London, England: Routledge, Feb. 2002.

[150] Tal Golan et al. "Deep neural networks are not a single hypothesis but a language for expressing computational hypotheses". In: *Behavioral and Brain Sciences* 46 (2023). ISSN: 1469-1825. DOI: 10.1017/s0140525x23001553. URL: http://dx.doi.org/10.1017/S0140525X23001553.

[151] Brendan Shea. "Karl Popper". In: *Internet Encyclopedia of Philosophy*. URL: https://iep.utm.edu/pop-sci/.

Appendix A

Appendix to the Sound Symbolism work

This list of words used to model the sound symbolism in Chapter 5 is presented below. The mp3 files of the words were downloaded from the Collins Dictionary website (https://www.collinsdictionary.com/dictionary/).

A.1 Languages and objects used for machine learning

A.1.1 Languages

Arabic, American English, Brazilian Portuguese, Chinese, Croatian, Czech, Danish, Dutch, Finnish, French, German, Greek, Italian, Japanese, Korean, Norwegian, Polish, Russian, Swedish, Thai, Turkish, Vietnamese

A.1.2 Sharp Objects

Scissors, Knife, Saw, Sword, Axe, Pin, Blade, Tweezers, Nail, Fork, Hook, Dart, Needle, Pine, Beak, Claws, Screw, Sting, Shovel, Razor, Chisel, Drill, Spine, Tooth, Cactus

A.1.3 Round Objects

Moon, Sun, Bubble, Melon, Ball, Poppy, Pepper, Doughnut, Pea, Onion, Rings, Egg, Eyes, Orange, Dandelion, Tyre, Button, Wheel, Beads, Cookies, Coins, Medal, Grapes, Tomato, Plate, Cabbage, Disk, Globe, Hole, Clock, Bulb, Lemon, Clouds, Bracelet, Apple, Potato, Circle, Oval, Blossom, Daisy, Bun, Wool, Flower

A.2 Psychophysics results

All words were presented in the same random order to subjects.

TABLE A.1: The psychophysics results for words in category sharp. (n=199)

Presentation Order	Object	Language	Word	Sharp Labelings
0	Horn	German	Horn	84
2	Icicle	French	glaçon	94
4	Bayonet	French	baïonnette	123
5	Spike	Latin American Spanish	punta	103
7	Bayonet	Brazilian Portuguese	baioneta	109
10	Dagger	Italian	pugnale	93
11	Shears	Brazilian Portuguese	podadeira	100
13	Shears	European Portuguese	podadeira	110
14	Shears	Italian	cesoie	115
15	Dagger	German	Dolch	87
16	Bayonet	European Portuguese	baioneta	101
18	Tusk	Brazilian Portuguese	presa	114
21	Spike	European Spanish	punta	96
22	Dagger	Brazilian Portuguese	punhal	102
25	Spike	German	Spitze	99
26	Tusk	Latin American Spanish	colmillo	109
27	Dagger	Latin American Spanish	puñal	100
29	Sabre	Latin American Spanish	sable	97
30	Sabre	European Spanish	sable	94
32	Scalpel	Italian	bisturi	102
36	Tusk	French	défense	92
40	Icicle	European Portuguese	sincelo	91
44	Sabre	French	sabre	108
45	Sabre	German	Säbel	100
46	Spear	Latin American Spanish	sable	98
48	Scalpel	French	scalpel	113
51	Dagger	French	poignard	108
53	Icicle	European Spanish	carámbano	132
54	Bayonet	European Spanish	bayoneta	119
56	Horn	Italian	corno	102
58	Horn	Brazilian Portuguese	chifre	109
59	Icicle	Latin American Spanish	carámbano	113
60	Shears	French	cisaille	113
64	Spike	European Portuguese	ponta	110
67	Horn	French	corne	102

TABLE A.2: The psychophysics results for words in category round. (n=199)

Presentation Order	Object	Language	Word	Sharp Labelings
1	Ripple	Latin American Spanish	onda	98
3	Foam	German	Schaum	63
6	Sphere	Italian	sfera	79
8	Bloom	French	fleur	74
9	Droplet	European Portuguese	gotícula	92
12	Droplet	French	gouttelette	78
17	Droplet	Italian	gocciolina	101
19	Sphere	Brazilian Portuguese	esfera	91
20	Ripple	Brazilian Portuguese	ondulação	104
23	Sphere	Latin American Spanish	esfera	102
24	Droplet	European Spanish	gotita	115
28	Ripple	German	kliene Welle	81
31	Bloom	European Portuguese	flor	65
33	Foam	Latin American Spanish	espuma	96
34	Foam	Brazilian Portugese	espuma	70
35	Foam	French	écume	96
37	Sphere	French	sphère	84
38	Ripple	European Spanish	onda	102
39	Ripple	Italian	increspatura	109
41	Droplet	Brazilian Portuguese	gotícula	73
42	Bloom	German	Blüte	72
43	Sphere	European Portuguese	esfera	75
47	Ripple	French	ondulation	103
49	Foam	European Spanish	espuma	93
50	Bloom	Italian	fiore	73
52	Foam	European Portuguese	espuma	80
55	Droplet	German	Tröpfchen	95
57	Sphere	German	Kugel	83
61	Bloom	European Spanish	floración	92
62	Bloom	Brazilian Portuegese	flor	62
63	Foam	Italian	schiuma	93
65	Bloom	Latin American Spanish	floración	71
66	Sphere	European Spanish	esfera	95
68	Ripple	European Portuguese	ondulação	74
69	Droplet	Latin American Spanish	gotita	96

Appendix B

Codes used for the device

This is the code for Arduino for the device described in Chapter 7. The randomizations have been done from a desktop computer before being written down in Arduino code. The datapoints in constant stimulus program has been chosen from the results of a binary search program as described in Chapter 7. The binary search program and the constant stimulus method program were used in Chapter 7. The complex stimulus program was used in Chapter 8.

B.1 Binary Search Program

```
int SignalPin=7;
int TriacPin=6;
int SmallLEDPin=5;

int FlickerPin=8;
int FusedPin=9;

int PValues[]={16, 6, 4, 5, 11, 3, 14, 2, 9, 10, 15, 18, 8, 12, 7, 1, 17, 13, 19};
int p=0;
int P=0;
int UpperTimePeriod=64;
int LowerTimePeriod=8;
int TimePeriod=0;
int time=0;
int flickers=0;
/*int flickerCount=2500; */

int OnMilliSeconds=0;
```

```
19 int OffMilliSeconds=0;
20 int OnMicroSeconds=0;
21 int OffMicroSeconds=0;
22 int incomingByte = 0;
23
24 void setup()
25 {
26   pinMode(SignalPin, OUTPUT);
27   pinMode(TriacPin, OUTPUT);
28   pinMode(SmallLEDPin, OUTPUT);
29   pinMode(FlickerPin, INPUT);
30   pinMode(FusedPin, INPUT);
31   updateValues();
32   Serial.begin(9600);
33 }
34 void loop()
35 {
36   ++flickers;
37   digitalWrite(SignalPin, HIGH);
38   delayMicroseconds(OnMicroSeconds);
39   delay(OnMilliSeconds);
40   digitalWrite(SignalPin, LOW);
41   delayMicroseconds(OffMicroSeconds);
42   delay(OffMilliSeconds);
43   if(digitalRead(FlickerPin)==HIGH)
44   {
45     Serial.print("Flicker\t");
46     PrintDetails();
47     if((TimePeriod-LowerTimePeriod)<=1)
48       newP();
49     else
50       UpperTimePeriod=TimePeriod;
51     updateValues();
52   }
53     if(digitalRead(FusedPin)==HIGH)
54   {
55     Serial.print("Fused\t");
56     PrintDetails();
57     if((UpperTimePeriod-TimePeriod)<=1)
58       newP();
59     else
60       LowerTimePeriod=TimePeriod;
```

```
61      updateValues();
62    }
63    if (Serial.available() > 0)
64    {
65      // read the incoming byte:
66      incomingByte = Serial.read();
67
68      UpperTimePeriod=64;
69      LowerTimePeriod=8;
70      if(incomingByte==48) // 0 is pressed
71        ++p;
72      else
73          if(incomingByte==49) // 1 is pressed
74            p=p+2;
75
76      else
77          if(incomingByte==50) // 2 is pressed
78            p=p+4;
79
80      else
81          if(incomingByte==51) // 3 is pressed
82            p=p+8;
83
84      else
85          if(incomingByte==52) // 4 is pressed
86            p=p+16;
87      Serial.print("p value set to ");
88      Serial.println(p);
89      updateValues();
90    }
91  }
92  void newP()
93  {
94    ++p;
95      if (p>=20) // if all 19 PCFs value have been tested
96    while(1>0)
97    {
98      digitalWrite(TriacPin, HIGH); // keep  LED constantly on
99      digitalWrite(SignalPin, HIGH); // keep LED constantly on
100   }
101   UpperTimePeriod=64;
102   LowerTimePeriod=8;
```

```
103 }
104
105
106 void PrintDetails()
107 {
108   digitalWrite(TriacPin, HIGH);
109
110   Serial.print("Time:\t");
111   Serial.print(TimePeriod);
112   Serial.print("\tP\t");
113   Serial.print(P);
114   Serial.print("\tFlickers\t");
115   Serial.print(flickers);
116   Serial.print("\tMilliMicroValues\t");
117   Serial.print(OnMilliSeconds);
118   Serial.print("\t");
119   Serial.print(OnMicroSeconds);
120   Serial.print("\t");
121   Serial.print(OffMilliSeconds);
122   Serial.print("\t");
123   Serial.print(OffMicroSeconds);
124   Serial.print("\n");
125   delay(4000);
126   digitalWrite(TriacPin, LOW);
127   delay(2000);
128 }
129 void updateValues()
130 {
131   TimePeriod=(UpperTimePeriod+LowerTimePeriod)/2;
132   /*flickerCount=10000/TimePeriod;*/
133
134   flickers=0;
135   P=PValues[p];
136   time=P*TimePeriod;
137   OnMicroSeconds=50*(time%20);
138   OnMilliSeconds=time/20;
139   time=(20-P)*TimePeriod;
140   OffMicroSeconds=50*(time%20);
141   OffMilliSeconds=time/20;
142 }
143
144
```

B.2 Constant Stimulus Method Program

```
int SignalPin=7;
int TriacPin=6;
int SmallLEDPin=5;

int FlickerPin=8;
int FusedPin=9;

int currenttimePoint=0;
int timepoint=0;
int timepoints[]={90, 75, 3, 5, 125, 86, 227, 149, 33, 144, 152, 63, 38, 180, 160,
                  113, 98, 168, 114, 207, 37, 43, 146, 186, 12, 215, 172, 104, 65,
                  46, 107, 109, 61, 155, 158, 70, 77, 91, 163, 176, 195, 196, 60,
                  26, 0, 159, 187, 28, 111, 137, 115, 205, 198, 27, 56, 218, 131,
                  138, 92, 153, 179, 185, 224, 47, 71, 17, 140, 221, 174, 100, 139,
                  105, 78, 68, 22, 121, 119, 194, 44, 164, 97, 192, 19,
                  202, 81, 24, 94, 8, 30, 210, 42, 143, 20, 136, 130,
                  49, 40, 112, 141, 154, 50, 18, 177, 89, 142,
                  102, 122, 184, 25, 101, 167, 110, 203, 148, 54, 58,
                  128, 85, 166, 211, 64, 15, 161, 96, 134, 35, 11,
                  72, 84, 223, 188, 197, 208, 59, 36, 171, 74, 23,
                  14, 145, 150, 129, 151, 175, 133, 108, 118, 76, 182,
                  80, 135, 4, 10, 124, 178, 169, 212, 87, 116, 117, 147,
                  173, 51, 193, 191, 48, 170, 53, 31, 52, 219, 2,
                  204, 69, 67, 62, 39, 127, 216, 95, 126, 41, 156, 1, 32,
                  132, 66, 226, 201, 165, 190, 57, 34, 79, 123,
                  45, 99, 82, 209, 225, 29, 13, 55, 181,
                  220, 222, 88, 7, 157, 9, 6, 189, 199, 120,
                  206, 214, 213, 103, 217, 93, 21, 83, 183, 16, 106, 73, 162, 200};
int TimePeriods[]={19  ,  19  ,  19  ,  14  ,  19  ,  18  ,  19  ,  14  ,  19  ,  15
                   22  ,  14  ,  22  ,  23  ,  23  ,  23  ,  26  ,  29  ,  34};
int p=0;
int P=0;
int TimePeriod=0;
int time=0;
int extratime=0;
int flickers=0;

int OnMilliSeconds=0;
int OffMilliSeconds=0;
int OnMicroSeconds=0;
```

```
41 int OffMicroSeconds=0;
42 int incomingByte = 0;
43
44 void setup()
45 {
46   pinMode(SignalPin, OUTPUT);
47   pinMode(TriacPin, OUTPUT);
48   pinMode(SmallLEDPin, OUTPUT);
49   pinMode(FlickerPin, INPUT);
50   pinMode(FusedPin, INPUT);
51   updateValues();
52   Serial.begin(9600);
53 }
54 void loop()
55 {
56   ++flickers;
57   digitalWrite(SignalPin, HIGH);
58   delayMicroseconds(OnMicroSeconds);
59   delay(OnMilliSeconds);
60   digitalWrite(SignalPin, LOW);
61   delayMicroseconds(OffMicroSeconds);
62   delay(OffMilliSeconds);
63   if(digitalRead(FlickerPin)==HIGH)
64   {
65     Serial.print("Flicker\t");
66     PrintDetails();
67     updateValues();
68   }
69     if(digitalRead(FusedPin)==HIGH)
70   {
71     Serial.print("Fused\t");
72     PrintDetails();
73     updateValues();
74   }
75   if (Serial.available() > 0)
76   {
77     // read the incoming byte:
78     incomingByte = Serial.read();
79
80
81     if(incomingByte==48) //0
82       ++currenttimePoint;
```

```
83      else
84          if(incomingByte==49) //1
85            currenttimePoint=currenttimePoint+2;
86
87      else
88          if(incomingByte==50) //2
89            currenttimePoint=currenttimePoint+4;
90
91      else
92          if(incomingByte==51) //3
93            currenttimePoint=currenttimePoint+8;
94
95      else
96          if(incomingByte==52) //4
97            currenttimePoint=currenttimePoint+16;
98
99
100     else
101         if(incomingByte==53) //5
102           currenttimePoint=currenttimePoint+32;
103
104     else
105         if(incomingByte==54) //6
106           currenttimePoint=currenttimePoint+64;
107
108     else
109         if(incomingByte==55) //7
110           currenttimePoint=currenttimePoint+128;
111
112     Serial.print("Current time point set to ");
113     Serial.println(currenttimePoint);
114     time=0;
115     updateValues();
116   }
117 }
118
119 void PrintDetails()
120 {
121   digitalWrite(SmallLEDPin, LOW);
122   digitalWrite(TriacPin, HIGH);
123
124   Serial.print(currenttimePoint);
```

```
125
126   Serial.print("\ttimepoint:\t");
127   Serial.print(timepoint);
128
129   Serial.print("\tTime:\t");
130   Serial.print(TimePeriod);
131   Serial.print("\tP\t");
132   Serial.print(p);
133   Serial.print("\tFlickers\t");
134   Serial.print(flickers);
135   Serial.print("\tMilliMicroValues\t");
136   Serial.print(OnMilliSeconds);
137   Serial.print("\t");
138   Serial.print(OnMicroSeconds);
139   Serial.print("\t");
140   Serial.print(OffMilliSeconds);
141   Serial.print("\t");
142   Serial.println(OffMicroSeconds);
143   delay(4000);
144   digitalWrite(TriacPin, LOW);
145   delay(2000);
146   ++currenttimePoint;
147 }
148 void updateValues()
149 {
150   timepoint=timepoints[currenttimePoint];
151   if (currenttimePoint>=228)
152   while(1>0)
153   {
154     digitalWrite(TriacPin, HIGH);
155     digitalWrite(SignalPin, HIGH);
156   }
157   p=timepoint/12;
158   extratime=timepoint%12;
159
160   flickers=0;
161   TimePeriod=TimePeriods[p]+extratime-5;
162   ++p;
163   time=p*TimePeriod;
164   OnMicroSeconds=50*(time%20);
165   OnMilliSeconds=time/20;
166   time=(20-p)*TimePeriod;
```

```
167   OffMicroSeconds=50*(time%20);
168   OffMilliSeconds=time/20;
169 }
170
```

B.3 Complex Stimulus Program

```
1  int SignalPin=7;
2  int TriacPin=6;
3  int SmallLEDPin=5;
4
5  int FlickerPin=8;
6  int FusedPin=9;
7
8  int currenttimePoint=0;
9
10 int m1seconds;
11 int m2seconds;
12
13 int m1microseconds;
14 int m2microseconds;
15
16 int m1[]={ 6 , 12 , 7 , 54 , 58 , 49 , 19 , 16 , 6 , 53 , 8 , 39 , 12 , 53 , 7 , 47 ,
17 50 , 15 , 58 , 14 , 11 , 42 , 44 , 7 , 46 , 16 , 10 , 14 , 15 , 21 , 58 , 19 , 12 , 3
18 37 , 14 , 46 , 16 , 13 , 21 , 16 , 17 , 54 , 16 , 18 , 54 , 52 , 16 , 15 , 83 , 19 ,
19 18 , 51 , 55 , 17 , 13 , 68 , 40 , 81 , 14 , 18 , 52 , 20 , 19 , 8 , 12 , 55 , 38 , 1
20 13 , 9 , 44 , 63 , 14 , 53 , 15 , 17 , 73 , 48 , 11 , 31 , 10 , 8 , 14 , 51 };
21 int m2[]={ 76 , 80 , 31 , 20 , 14 , 40 , 55 , 4 , 15 , 20 , 52 , 37 , 80 , 13 , 45 , 3
22 65 , 59 , 18 , 30 , 46 , 39 , 94 , 46 , 73 , 10 , 21 , 54 , 14 , 45 , 80 , 89 , 43 , 1
23 55 , 56 , 56 , 55 , 17 , 9 , 10 , 45 , 89 , 10 , 57 , 7 , 58 , 59 , 7 , 13 , 70 , 19
24 57 , 21 , 21 , 43 , 17 , 87 , 20 , 45 , 16 , 72 , 74 , 55 , 14 , 44 , 62 , 21 , 12 , 6
25 53 , 15 , 10 , 63 , 79 , 8 , 39 , 7 , 9 , 60 , 29 , 56 , 11 };
26 int flickers=0;
27 int incomingByte = 0;
28
29 void setup()
30 {
31   pinMode(SignalPin, OUTPUT);
32   pinMode(TriacPin, OUTPUT);
33   pinMode(SmallLEDPin, OUTPUT);
34   pinMode(FlickerPin, INPUT);
```

```
35   pinMode(FusedPin, INPUT);
36   updateValues();
37   Serial.begin(9600);
38 }
39
40 void loop()
41 {
42   ++flickers;
43   digitalWrite(SignalPin, HIGH);
44   delayMicroseconds(m1microseconds);
45   delay(m1seconds);
46   digitalWrite(SignalPin, LOW);
47   delayMicroseconds(m1microseconds);
48   delay(m1seconds);
49
50   digitalWrite(SignalPin, HIGH);
51   delayMicroseconds(m2microseconds);
52   delay(m2seconds);
53   digitalWrite(SignalPin, LOW);
54   delayMicroseconds(m2microseconds);
55   delay(m2seconds);
56
57
58   if(digitalRead(FlickerPin)==HIGH)
59   {
60     Serial.print("Flicker\t");
61     PrintDetails();
62     updateValues();
63   }
64     if(digitalRead(FusedPin)==HIGH)
65   {
66     Serial.print("Fused\t");
67     PrintDetails();
68     updateValues();
69   }
70   if (Serial.available() > 0)
71   {
72     // read the incoming byte:
73     incomingByte = Serial.read();
74
75
76     if(incomingByte==48) //0
```

```
77        ++currenttimePoint;
78      else
79          if(incomingByte==49) //1
80            currenttimePoint=currenttimePoint+2;
81
82      else
83          if(incomingByte==50) //2
84            currenttimePoint=currenttimePoint+4;
85
86      else
87          if(incomingByte==51) //3
88            currenttimePoint=currenttimePoint+8;
89
90      else
91          if(incomingByte==52) //4
92            currenttimePoint=currenttimePoint+16;
93
94
95      else
96          if(incomingByte==53) //5
97            currenttimePoint=currenttimePoint+32;
98
99      else
100         if(incomingByte==54) //6
101           currenttimePoint=currenttimePoint+64;
102
103     Serial.print("Current time point set to ");
104     Serial.println(currenttimePoint);
105     updateValues();
106   }
107 }
108
109
110 void PrintDetails()
111 {
112   digitalWrite(SmallLEDPin, LOW);
113   digitalWrite(TriacPin, HIGH);
114
115   Serial.print(currenttimePoint);
116     Serial.print("\tm1:\t");
117     Serial.print(m1[currenttimePoint]);
118         Serial.print("\tm2:\t");
```

```
119
120   Serial.print(m2[currenttimePoint]);
121
122     Serial.print("\tm1seconds:\t");
123   Serial.print(m1seconds);
124
125     Serial.print("\tm1microseconds:\t");
126   Serial.print(m1microseconds);
127
128     Serial.print("\tm2seconds:\t");
129   Serial.print(m2seconds);
130
131     Serial.print("\tm2microseconds:\t");
132   Serial.print(m2microseconds);
133     Serial.print("\tFlickers\t");
134   Serial.println(flickers);
135   delay(4000);
136   digitalWrite(TriacPin, LOW);
137   delay(2000);
138   ++currenttimePoint;
139 }
140 void updateValues()
141 {
142     if (currenttimePoint>=100)
143   while(1>0)
144   {
145     digitalWrite(TriacPin, HIGH);
146     digitalWrite(SignalPin, HIGH);
147   }
148   m1seconds=m1[currenttimePoint]/2;
149   if (m1[currenttimePoint]%2==0)
150   m1microseconds=0;
151   else
152   m1microseconds=500;
153
154   m2seconds=m2[currenttimePoint]/2;
155   if (m2[currenttimePoint]%2==0)
156   m2microseconds=0;
157   else
158   m2microseconds=500;
159   flickers=0;
160
```

```
161 }
```