

Chapter 6

Multi layer SNN in Primates' Visual Cortex

6.1 Introduction

The visual cortex is one of the complex biological networks responsible for learning. The visual cortex is continuously fed with natural scenes that later serve as the basis of complex cognitive behavior. Different layers of the visual cortex extract multitude of features from natural scenes. These visual features, collected from different regions of the visual cortex, are routed to the successive layers of ganglion cell networks to generate self-organizing maps responsible for short-term and long-term memory, contour estimation, motion detection, etc. Feature extraction in the visual cortex is computed as early as the striate cortex in the V1 region. These extracted features are then projected to the mutually exclusive parvocellular and magnocellular regions for shaping complex cognitive behavior. Previous state-of-the-art confirms the parvocellular regions being stimulated with projected multi-scale features and magnocellular region being responsive toward coarser features. The parvocellular region and magnocellular region are isolated from each other and responsible for region-specific computation. The parvocellular region stimulated with a multi-scale feature from the V1 region, is responsible for content-based information processing responsible for the scene, object, and face identification type behavior. On the other hand, the magnocellular region, deriving large-scale features, is responsible for boundary estimation, object tracking, contour detection, depth, and motion perception.

Lots have been explored about the global behavior and responses of multi-

layer neural networks. But the local dynamics of deterministic type-specific neuron morphology shaping complex responses are yet to be explored. The role of neuron morphology with localized AIC has not been getting much attention due to the unavailability of sophisticated imaging and measuring techniques. On the contrary, dendritic arbor triggering active AP has been discovered a few decades back and since gained attention. Rigorous studies have been conducted to link the computational power of neurons to the local dynamics.

Significant literature and detailed morphological structures were introduced by Ramon y Cajal [292] in terms of dedicated circuits that are organized in the primate visual cortex. A diverse collection of neuronal morphologies arranged in multiple layers with dedicated functionalities targeted toward enhanced cognition, plasticity, and learning tasks and are achieved via reorganizing connections [293], growth and retraction [294] of dendritic inputs. In 1962, Hubel and Wiesel came forward with the fundamental floor plan of mammalian retina [295, 296] where the localized neurons transduce light signals and encode them into orientation-selective (OS) spiking responses in primary layers whereas scale and location-independent cognition in others. These orientation-selective responses are computed in the primary V1 layers of the visual cortex and seem to be the basis of mammalian visual perception and cognition. The detailed investigation of Hubel and Wiesel [295, 296] discusses the details of a precise organization of RF, antagonistic connectivity of RGC with bipolar cells (BC) [186, 288] and axis of orientation selectivity which were later confirmed in physiological experimentation [123–125, 297]

Based on these initial investigations, a subtle amount of literature has been put forward to replicate localized behavior concerning orientation selectivity [186, 291], edge detection, contour detection, scene segmentation, scale and position-independent cognition and learning [298, 299] that reports findings similar to the landmark works of Hubel and Wiesel. Literature pointed out the importance of neuronal anatomy as well as physiology in functional computation in orientation selectivity but very little has been understood about connectome structure [123–125, 297], the role of unique neuronal morphologies [300, 301] and dynamics of such morphology due to localized ion channels [137, 138, 141, 302]. Recent works [123, 124] confirms precise micro-architecture in mammalian visual cortex organized with single-cell precision whereas literature [137, 138, 302] describes dynamics of neuronal inputs (dendritic arbors) to be dependent on such localized AICs.

This work has been put forward to integrate designed morphologically detailed RGCs to construct a Hubel and Wiesel type feed-forward network of the

6.2. Hubel and Wiesel Type Network

visual cortex. The model tries to emulate simple cell layers $S1$ and $S2$ that are responsible for extracting multiscale-oriented features. These features are forwarded to the complex cell layers.

This chapter emphasizes to explore the role of distinct type-specific neurons in feature extraction and their implication in biological neural networks. Oriented features has been specifically targeted to investigate their probable role in recognition and learning.

6.1.1 Contribution

The proposed model suggests the probable role of different RGC structures linked to individual functional aspects that plays an important role in feature extraction, object detection, and recognition. Preliminary test in face recognition has been conducted by incorporating the morphologically detailed RGC neurons in the $S1$ and $S2$ layers of hmax model inspired by Hubel and Wiesel. The model performs very well considering firing rate based encoding in individual neuron. The model performance is expected to improve the learning capability when incorporated with multiple RF sizes. The model also opens scope of future research in the field incorporating spike-time dependent encoding and similar spike base encoding techniques.

6.2 Hubel and Wiesel Type Network

Experiments performed on an anesthetized cat, macaque, and spider monkey and simultaneous recordings from the layer between the Lateral Gyrus (LG) and post-Lateral Gyrus (PSG) by Hubel and Wiesel explores details about functional anatomy and physiology. It categorizes the research into distinct sections including the RF organization, ocular dominance, and functional architecture. In the RF organization, they grouped the majority of the cells into two major groups called ‘simple cells’ and ‘complex cells’ corresponding to their responses to stimuli. These cells are organized into excitatory and inhibitory regions in their respective RFs that shape their natural responses. These simple cells are responsive to oriented features and resonate at maximum spiking frequency when the stimulated oriented feature matches the preferred orientation of the cell. On the other hand, complex cells don’t have any clear separation of the RF into excitatory and inhibitory regions. These complex cells respond to broad slits and dark bars rather than

normal slits and bars. This particular features in complex cells has been termed as max-pooling type activity [298, 303].

6.2.1 Simple Cell Layer

Simple cells in the designed visual cortex network are replaced with designed morphologically detailed RGCs discussed in the previous chapters. These RGCs are arranged in precise repetitive patterns with single-cell precision. The connectivity arrangement of the RGCs is configured to the connectome specificity of the OS-RGC network responsible for oriented feature extraction as discussed in [295, 296, 298]. Very similar to the feedforward network discussed in [298, 303], orientation selectivity has been achieved by the connectivity configuration of the RGC, RF size is controlled by the dendritic spread of the RGC, and bandwidth of the RGC is controlled by the specific combination of localized AICs. Specificity of the Hubel and Wiesel [295, 296] network has strictly adhered while designing the simple cell layer as well as the complex cell layers. Details of the RGC morphology, bandwidth selectivity, and RF sizes have been discussed in the section follows.

Simple cell layers in the designed model have been configured to extract oriented features from the inputs. Oriented feature selectivity has been configured using the connectome specificity of the OS-RGC network and has been discussed in section 4.4. The first layer of simple cell is responsible for simple oriented features at 0° , 45° , 90° , and 135° . The second layer of simple cell is responsible for extracting complex features due to the intersection of oriented features.

6.2.1.1 RGC Morphology and Bandwidth

The model framework is designed incorporating Hubel and Wiesel type network [295, 296] as the base. In vivo findings reported in the literature [125, 300, 301] has also been integrated to the proposed framework to replicate local RGC dynamics [137, 138, 142]. RGC with detailed morphology has been designed considering dendrite arborization and dendritic spread of midget and parasol RGCs and briefed in section 4.3.2.1, Figure4-4 and section 5.3.1, Figure5-1. Distal dendrites of the RGC are responsible for collecting inputs from the connection site. Sub-cellular distribution of localized AICs renews the overall incoming signals in terms of active APs for feed-forward propagation. Passive fibers in the dendrites helps to propagate the signal with attenuation proportional to the length of the passive fiber. The nodes/ dendritic junctions act as summing nodes responsible for the

6.2. Hubel and Wiesel Type Network

temporal summation of the incoming signals. The type of local AIC distribution controls the bandwidth selectivity. Tuning of the bandwidth has been achieved by controlling the total amount of propagating current reaching the nodes/ dendritic junctions.

6.2.1.2 Bipolar Connectivity, Receptive Fields and Receptive Field Sizes

RGC connectivity with the BCs forms the base of RFs. RFs are the regions over the retina that influences the firing frequency of the connected neuron. Temporal signals from the BCs are fed to the RGCs and the RGCs respond correspondingly. Details of the BC signal stimulation and intensity to Spatio-temporal signal generation have been briefed in section 4.3.1.2. RGCs are connected to BCs with connectome specificity discussed in section 4.3.2.2 and section 5.3.2 with single-cell precision [123–125] and well-defined excitatory and inhibitory regions. Balanced excitatory and inhibitory connectivity of RGCs with BCs defines the orientation specificity of the RGC. The connected RGC starts resonating when the stimulated feature matches the preferred orientation and vice versa. Feature orientation specificity of the RF is defined by the axis following through the excitatory and inhibitory connections. Connectome neighborhood results in the formation of RFs of different sizes giving the RGCs multiscale feature extraction capabilities.

6.2.2 Complex Cell Layer

The complex cell layer used in the model is similar to complex cells discussed in Hubel and Wiesel [295, 296]. Pooling functions are conducted by complex cells within larger RF sizes than simple cells. Pooling operations are carried out in the neighborhood of simple cell responses with multiple RF sizes, which provide location-independent behavior within their RF, which is an intrinsic characteristic of complex cells. Similarly, the effect of scale pooling is a broadening of the frequency bandwidth from simple to complex units, which is consistent with physiology.

6.3 Modeling and Simulation

Shown in Figure 6-1 is the visual cortex-inspired learning system, which has been designed by adaptation of the well-acknowledged Hubel and Wiesel architecture [295, 296]. At the top of the model is the **Input** which is the natural stimulus fed to the visual cortex model which is being sensed by the retina. Spatial information from the retina is fed to the BCs, where the signals are transduced and converted to Spatio-temporal sequences depending on the exposure time of the input. The process of spatial to Spatio-temporal signal conversion has been elaborated in section 4.3.1.2 considering the response time of primates' vision. OS-RGCs are connected to BCs with connection specificity discussed in section 5.3.2 with single-cell precision. Four OS-RGC layers are stacked over the BCs to extract oriented features at 0° , 45° , 90° and 135° configured to extract oriented features at three scales, primarily 3×3 , 5×5 and 7×7 , and 2 opposite phases. The two opposite phase-oriented features are computed by reversing the connectome polarity of the connected RGCs which in turn gives the **ON phase OS filter** and **OFF phase OS filter** responses of the simple cell layer **S1**.

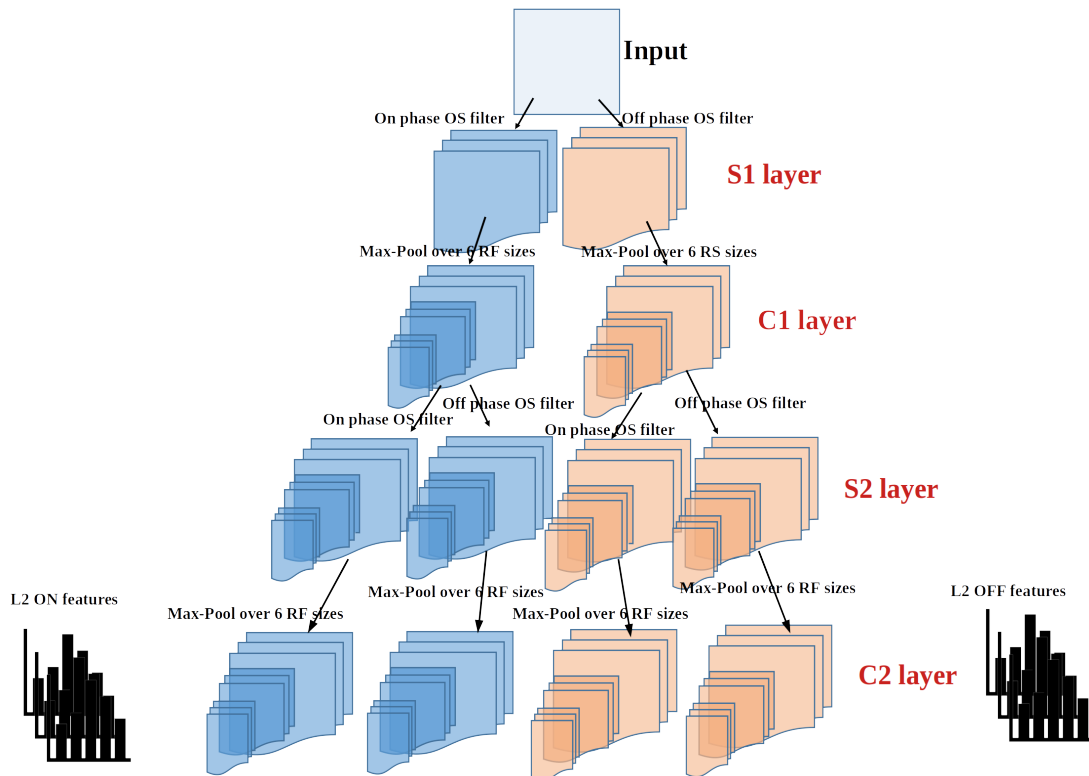


Figure 6-1: Visual cortex inspired learning model

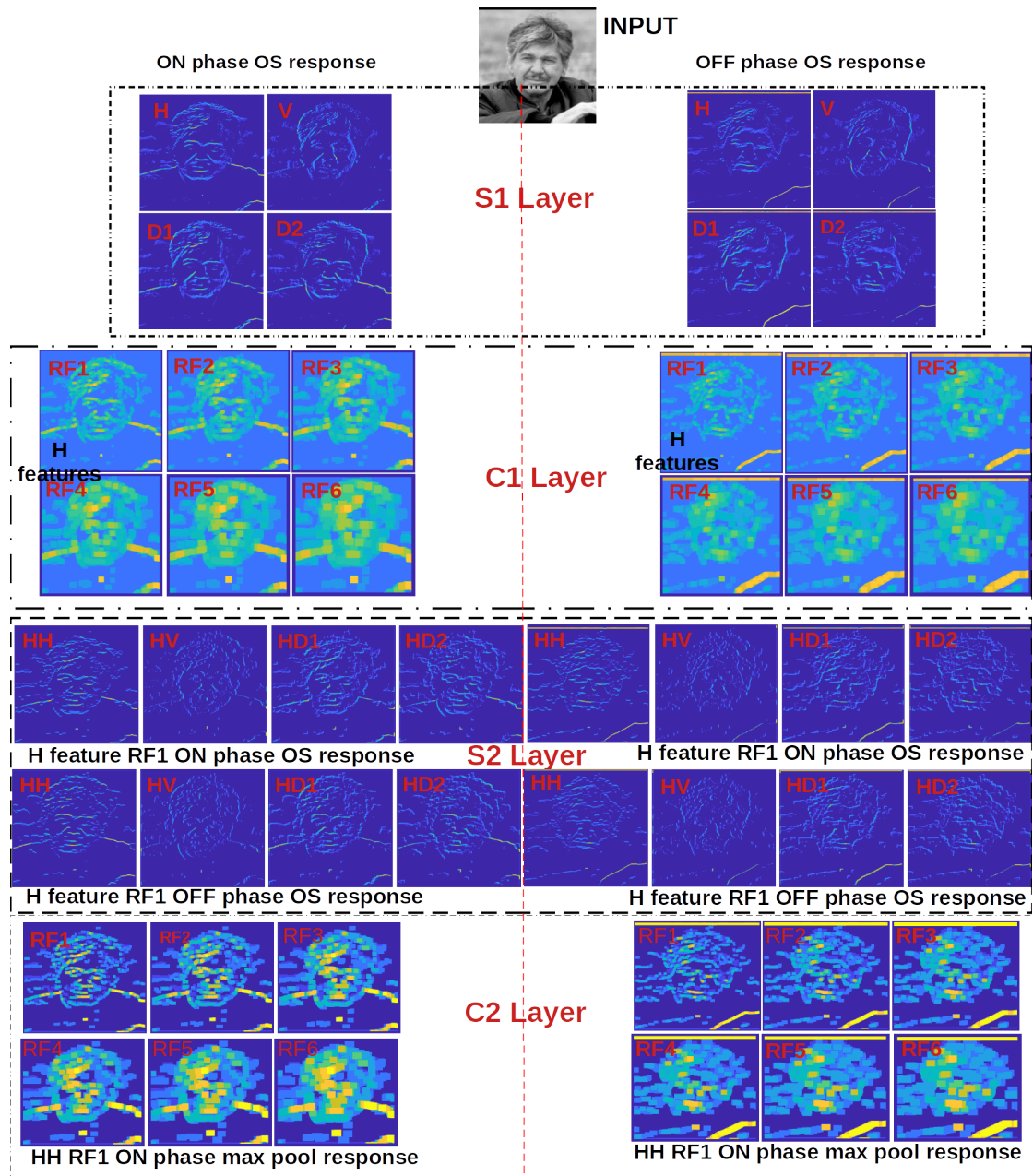
These simple cell responses from **S1** layer are feed-forwarded to the complex cell layer **C1** where max-pooling operations are performed on oriented feature

6.3. Modeling and Simulation

responses [298, 299]. Max-pool operation on the oriented features helps in broadening the frequency bandwidth. Six RFs different than the ‘S1’ layer RFs sizes have been used in the ‘C1’ layer. This pooling operation with sizes different than the ‘S1’ RF gives location-independent behavior. These complex cell responses from the ‘C1’ layer are again forwarded to simple cell layer ‘S2’, responsible for complex oriented feature extraction. Simple cell layer ‘S2’ being fed with a broader band-oriented feature, extracts complex features such as the intersection of two or more oriented features and contours. From these oriented features at complex cell layer ‘C2’, feature vectors are created. Since population response in terms of spiking frequency has been considered, the population spiking histogram has been taken as a feature vector at the two complex cell layers for the learning model.

Shown in Figure6-2 is a sample of input images being processed by the proposed model inspired by the visual cortex. ‘S1 Layer’ and ‘S2 Layer’ are the simple cell layers of OS-RGCs responsible for the extraction of oriented features from the inputs whereas ‘C1 Layer’ and ‘C2 Layer’ are layers of complex cells responsible for performing the max-pooling operation on the oriented responses.

As shown in Figure6-2, when inputs are fed to the visual cortex, spatial information is converted to Spatio-temporal signals via the BCs and forwarded to the first simple cell layer ‘S1 Layer’. The ‘S1 Layer’, being configured for preferred oriented features, extracts orientation information in 4 orientation and 3 scales, namely 0° , 45° , 90° , 135° and RF sizes of 3×3 , 5×5 and 7×7 cell neighborhood, for ON and OFF phases. The On and OFF phase-oriented features are complementary to each other and the same has been computed for each RF size and the responses are forwarded to the ‘C1 Layer’. In Figure6-2, ‘H’, ‘V’, ‘D1’ and ‘D2’ at the left are the ON phase-oriented features and at the right are the OFF phase-oriented features for a RF size of 3×3 RGC neighborhood for the ‘Input’ shown. In the ‘C1 Layer’, max-pooling operation is performed with 6 different RF sizes, namely ‘RF1’, ‘RF2’, ‘RF3’, ‘RF4’, ‘RF5’ and ‘RF6’ corresponding to 5×5 , 7×7 , 9×9 , 11×11 , 13×13 and 15×15 RGC neighborhood respectively for the two phases. Responses from ‘C1 Layer’ are feed-forwarded to the second simple cell layer ‘S2 Layer’ where oriented features are extracted. Successive operations of oriented feature extraction results in contour detection and complex feature detection such as a combination of oriented features. In the proposed framework shown in Figure6-2, ‘HH’, ‘HV’, ‘HD1’ and ‘HD2’ are some of the complex oriented features extracted that are combinations of two oriented features, 0° - 0° (‘HH’), 0° - 90° (‘HV’), 0° - 45° (‘HD1’) and 0° - 135° (‘HD2’). These complex features are forwarded to the second complex cell layer ‘C2 Layer’,



Oriented Features
 H = Horizontal V = Vertical D1 = 45 degree D2 = 135 degree
Combinations of Oriented Feature
 HV = Horizontal-Vertical HH = Horizontal-Horizontal
 HD1 = Horizontal-45 degree HD2 = Horizontal-135 degree
 RF1, RF2, RF3, RF4, RF5, RF6 different receptive field sizes for max pooling in Complex cell layer

Figure 6-2: Sample response to an input image at S1, C1 and S2 layers for different RF size of ‘S1’ and ‘S2’ layer and three pooling operations in of the complex cell layer C1 and C2 with RF sizes of 5, 7, 9, 11, 13, 15 RGC neighborhood is shown for the designed network.

where max-pooling operations are being performed for 6 RF sizes.

6.4 Population Feature

As shown in Figure6-1 and Fig6-2, ‘Input’ has been fed to the network, where 4 oriented features are extracted in the ‘S1’ layer at 3 different RF sizes of 2 opposite phases which are forwarded to ‘C1’ layer. These oriented features are max-pooled with 6 RF sizes and feed-forwarded to the successive simple cell layer ‘S2’ and then to the ‘C2’ layer to extract the max activation information. This max-activation information of the neuron population in the ‘C2’ layer is used as feature space for performance estimation in face recognition. Since the spiking frequency of the neuron population has been emphasized and shows very promising results in edge map reconstruction, a neuron population spiking histogram has been used for the identification of faces. Population spiking histogram has been extracted from each RF response of the ‘C2’ layer and later used for identification of face. For identification of faces, ‘C2’ layers features are computed and population spiking histograms are extracted and compared against the already existing histogram features of the learned faces. Euclidean distance between the corresponding RF histogram features is calculated, which is found to be minimum in most of the matching cases whereas the Euclidean distance measure for dissimilar faces is larger.

6.5 Result and Discussion

To investigate the feasibility of the proposed model in learning and object recognition, it has been integrated into the hmax model discussed in [298, 299] with minor changes in the number of filters. The hmax model is inspired from Hubel and Wiesel [295, 296] RF architecture. The presented approach replaces the simple cell layers ‘S1’ and ‘S2’ in the hmax model to extract directional features, which are then fed to populations of complex cell layers C1 and C2’s responsible for pooling maximum activation information. In the hmax model, the ‘S1’ unit has 2 phases, 4 orientations, and 17 sizes of RFs. However, due to the computational complexity of the proposed model, the ‘S1’ unit is accommodated with 2 phases, 4 orientations, and 3 sizes of RFs. The model’s two phases are due to the morphological connectivity matrices discussed in section 4.3.2.2, but with the polarity of the connections reversed. The ‘S1’ layer discussed in the hmax

model implements a Gabor filter responsible for the extraction of local directional feature information, and parameters such as θ , γ , σ , ϕ , and λ control RF properties such as orientation, aspect ratio, effective width, phase, and wavelength. The proposed morphologically defined neuron model, very similar to the hmax model, works as a local directional texture information extractor from the input image. The connectivity matrix controls orientation in the proposed model, the spread of neurons controls the aspect ratio, wavelength tuning is controlled by the type of localized ion channels at a specific site, and phase is controlled by the polarity of connectivity matrices. The sensitivity of neuron morphology, on the other hand, can be controlled by varying the amount of input current from BCs. In the proposed model, the neuron's bandwidth is controlled by a combination of spiking activity at the junctions and the neuron's cell body. Bursting and chattering type of neuron spiking activity have been considered for optimal model behavior and considering types of spiking activity seen in the visual cortex. In the hmax model, the 'C1' neuron layer is responsible pooling the maximum activation of neurons within a specific RF.

Table 6.1: Retrieval performance of the hmax model after incorporating the proposed morphologically detailed RGC neuron model in the S1 and S2 layer.

Face Id	Average Retrieval Rate(in %)
Charles Bronson 1	81.667
Charles Bronson 2	83.333
Colin Prescot	68.333
Dino de Laurentiis	73.333
Emma Thompson	66.667
Francis Ricciardone	71.667
Fuji Cho	63.333
Gloria Macapaga Arroyo 1	63.333
Gloria Macapaga Arroyo 2	71.667
Gwyneth Paltrow	68.333

Because complex cell units are insensitive to the location of the stimulus within their respective RFs, the RFs of 'C1' cells are normally larger than those of simple cells 'S1' and 'S2'. As a result, 6 RFs (5, 7, 9, 11, 13, 15) are considered for 'C1' cell pooling operations. The pooled activity of 'C1' cells is re-projected in 4 directional orientations with 2 opposite phases capable of capturing more complex patterns such as contours and combinations of orientation information in the subsequent simple cell layer 'S2', which in turn are later fed to the 'C2' complex cells. The model's 'C2' cells incorporated in the model are functionally similar to 'C1' complex cells and responsible for max-pooling operations. Initial learning and recognition model incorporating simple detailed morphological and connectomic

structure of neuron to hmax model [298, 299] has been implemented on the face database provided with the official computational hmax model (found in <https://maxlab.neuro.georgetown.edu/hmax.html#updated>). Facial features from the face database were learned in the ‘C2’ layer in terms of population histogram, and testing and validation were performed on 6 shifted or tilted face images, position-independent face images for each subject, and overall recognition and retrieval accuracy is calculated from the best 4 retrieved images. The model’s average face detection and retrieval accuracy are 71.166%, which appears very interesting given that only three RF sizes are incorporated in the ‘S1’ layer and six in the ‘C1’ layer, and it is expected to perform better by increasing the number of RF-dependent filters in the ‘S1’ and ‘S2’ regions. The hmax model, on the other hand, performs better, with a face identification benchmark of 90.4%, which could be attributed to the amount of direction information extracted. Table.6.1 shows the proposed model’s face detection and retrieval accuracy for each subject.

6.6 Summary

The visual cortex is a sophisticated information processor that can process visual input and extract relevant information in parallel to generate accurate and useful information. Such a robust system is outfitted with a variety of RGN morphologies, each with its connectome specificity and dendritic spread organized in modular patterns and segregated into multiple layers. At least 50 different types of RGCs with distinct morphologies can be found, including midjets and parasol cells. The proposed model hypothesizes the role of various RGC structures linked to individual functional aspects in extracting features, object detection, and recognition. A preliminary test to understand its role in face recognition is performed by incorporating Hubel and Weisel inspired morphologically detailed RGC neurons in the S1 and S2 layers of the hmax model. Given the number of filters included in the model, the model performs well in terms of face recognition and retrieval. The proposed model’s object recognition performance is expected to improve as the range of RF widths that can focus on a wide range of spatial frequencies for multi-scale feature extraction expands. These robust networks are capable of extracting not only multi-scale features, shapes, and edge information, but also global information such as depth estimation, motion detection, gaze stabilization, and so on. Combining such local features with global features may result in error-free detection of shapes, edges, and objects, among other things. Multi-scale local feature detectors, in conjunction with multi-scale global feature

extractors, may play a significant role in learning feature optimization for object identification, possibly in terms of self-organization neural nets via dendritic spine extension and retraction. Thus, a thorough examination of the significance of morphology and its related inter-layer and intra-layer connectivity could lead to a new understanding of the dynamic change in dendritic morphology and arbor growth during learning.