



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ  
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Νευρική Απόκριση Αμφιβληστροειδούς και Οπτική Προέξοχη:  
Μοντελοποίηση με Μεθόδους Βαθιάς Μάθησης και  
Συμπεράσματα για την Προσθετική Όραση**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**ΓΕΩΡΓΙΟΣ Α. ΝΑΚΟΠΟΥΛΟΣ**

**Επιβλέπουσα :** Κωνσταντίνα Νικήτα  
Καθηγήτρια Ε.Μ.Π

Αθήνα, Φεβρουάριος 2021





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ  
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Νευρική Απόκριση Αμφιβληστροειδούς και Οπτική Προεξοχή:  
Μοντελοποίηση με Μεθόδους Βαθιάς Μάθησης και  
Συμπεράσματα για την Προσθετική Όραση**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**ΓΕΩΡΓΙΟΣ Α. ΝΑΚΟΠΟΥΛΟΣ**

**Επιβλέπουσα :** Κωνσταντίνα Νικήτα  
Καθηγήτρια Ε.Μ.Π

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 15<sup>η</sup> Φεβρουαρίου 2021.

.....  
Κωνσταντίνα Νικήτα  
Καθηγήτρια Ε.Μ.Π

.....  
Δημήτριος-Διονύσιος  
Κουτσούρης  
Καθηγητής Ε.Μ.Π

.....  
Ανδρέας-Γεώργιος  
Σταφυλοπάτης  
Καθηγητής Ε.Μ.Π

Αθήνα, Φεβρουάριος 2021

.....

**Γεώργιος Α. Νακόπουλος**

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π

Copyright © Γεώργιος Α. Νακόπουλος, 2021.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.



## Περίληψη

Η διαδικασία με την οποία οι άνθρωποι και άλλοι ζωντανοί οργανισμοί αντιλαμβάνονται το περιβάλλον και ερμηνεύουν τα διαρκή οπτικά ερεθίσματα από αυτό ξεκινά στον οφθαλμό. Εκεί, και συγκεκριμένα στον αμφιβληστροειδή πραγματοποιείται μια πολυστρωματική και σύνθετη επεξεργασία. Οι φωτοϋποδοχείς χαρακτηρίζονται από ευαισθησία στο φως και εντοπίζουν πρώτοι τις μεταβολές στο οπτικό πεδίο. Διαφορετικές κατηγορίες κυττάρων επεξεργάζονται το σήμα που τελικά οδηγείται στο επίπεδο εξόδου, τα γαγγλιακά κύτταρα, που αποστέλλουν προς τον εγκέφαλο την κωδικοποιημένη πληροφορία για περαιτέρω επεξεργασία.

Η τεχνολογία της προσθετικής αμφιβληστροειδούς επιδιώκει να καταπολεμήσει ασθένειες που βλάπτουν τους φωτοϋποδοχείς και διαταράσσουν την παραπάνω λειτουργία. Στόχος είναι να παρακαμφθεί ο παθολογικός ιστός και με κατάλληλα ηλεκτρικά σήματα να ενεργοποιηθούν τα γαγγλιακά κύτταρα.

Για έναν άνθρωπο, δεν είναι απαραίτητο να σαρώσει την οπτική σκηνή στην ολότητά της προκειμένου να κατανοήσει τι βλέπει. Ο εγκέφαλος δημιουργεί ένα χάρτη προεξοχής (saliency map), ο οποίος αναπαριστά την ικανότητα κάθε σημείου της εικόνας να ξεχωρίζει από τα γειτονικά του. Με αυτό τον τρόπο ο άνθρωπος, μέσα από συνεχείς κινήσεις των ματιών, εντοπίζει ενδιαφέρουσες περιοχές και εστιάζει σε αυτές.

Στην παρούσα διπλωματική εργασία μελετάμε τη σχέση της οπτικής προεξοχής με την απόκριση του αμφιβληστροειδούς. Αρχικά, αναπτύσσουμε ένα μοντέλο του αμφιβληστροειδούς εκπαιδύοντας ένα συνελκτικό νευρωνικό δίκτυο με πειραματικά βιολογικά δεδομένα καταγραφών της απόκρισης γαγγλιακών κυττάρων σε εικόνες του φυσικού περιβάλλοντος. Αξιοποιώντας την ικανότητα του εκπαιδευμένου μοντέλου να μιμηθεί τη συμπεριφορά του αμφιβληστροειδούς, υπολογίζουμε και απεικονίζουμε τα υποδεκτικά πεδία των νευρώνων του πειράματος μέσω μιας κλασικής τεχνικής που ονομάζεται αντίστροφη συσχέτιση, εκτιμώντας με αυτό τον τρόπο τη στοιχειώδη σχέση μεταξύ διέγερσης και απόκρισης. Ως υποδεκτικό πεδίο ενός γαγγλιακού κυττάρου ορίζεται η περιοχή του οπτικού πεδίου εντός της οποίας η πρόκληση κάποιας αλλαγής θα μεταβάλλει την απόκρισή του. Εξετάσαμε για κάθε κύτταρο τη μορφή του υποδεκτικού του πεδίου σε αντιπαραβολή με την τεκμηριωμένη σε θεωρητικές μελέτες περιγραφή των υποδεκτικών πεδίων. Διακρίναμε μια ξεκάθαρη περιοχή ενδιαφέροντος, αλλά όχι για όλους τους νευρώνες. Μέσα από τη διερεύνηση του χρονικού διαστήματος στο οποίο οι αλλαγές επηρεάζουν τις αποκρίσεις, καταλήξαμε σε συγκεκριμένη τιμή μνήμης για τα κύτταρα και διαπιστώσαμε ότι η υψηλή επίδοση του δικτύου δεν οδήγησε αυτόματα και στα προσδοκώμενα υποδεκτικά πεδία.

Στη συνέχεια υπολογίζουμε το χάρτη προεξοχής για ένα πλήθος από εικόνες μέσω ενός ήδη εκπαιδευμένου για αυτό το πρόβλημα τεχνητού νευρωνικού δικτύου. Παράλληλα, προβάλλουμε τις ίδιες εικόνες στα γαγγλιακά κύτταρα και λαμβάνουμε τις αποκρίσεις τους μέσω του προηγούμενου μοντέλου. Επιχειρούμε να αντιληφθούμε αν η ιδιότητα ενός αντικειμένου να ξεχωρίζει ως προς τα γειτονικά του αντικατοπτρίζεται στη μεταβολή του

ρυθμού πυροδότησης των γαγγλιακών κυττάρων. Με αυτό τον τρόπο θα μπορούσαμε να διεγείρουμε στοχευμένα και αποτελεσματικά τους νευρώνες και ενδεχομένως να βελτιώσουμε τα εμφυτεύματα αμφιβληστροειδούς. Ειδικότερα, συγκρίνουμε τη μέση απόκριση με αυτή στις περισσότερες προεξέχουσες περιοχές, μελετώντας ταυτόχρονα τη συσχέτιση μεταξύ συχνότητας δυναμικών και έντασης προεξοχής. Δεν διαπιστώσαμε συνέπεια στη συμπεριφορά μεταξύ των κυττάρων. Ορισμένα αύξησαν την απόκρισή τους κατά τον εντοπισμό προεξέχοντος αντικειμένου σε περισσότερες περιπτώσεις από άλλα, ενώ είχαν και ισχυρότερους συντελεστές συσχέτισης. Τα κύτταρα που είχαν μέση απόκριση στις εικόνες κάτω από την τιμή ηρεμίας τους έδειξαν μια τάση να αυξήσουν το ρυθμό τους όταν τους προβλήθηκαν περιοχές μέγιστης προεξοχής, έχοντας θετικούς συντελεστές. Αντίθετα τα υπόλοιπα, που επέλεξαν να στέλνουν δυναμικά σε υψηλότερες συχνότητες, μείωναν την απόκρισή τους στις σημαντικές περιοχές, έχοντας αρνητικούς συντελεστές. Οι τιμές στις μετρικές ήταν χαμηλές, κι έτσι δεν μας επιτράπη να ισχυριστούμε πως ο αμφιβληστροειδής συμμετέχει στη διαμόρφωση του χάρτη προεξοχής με έναν τόσο άμεσο τρόπο.

Η εργασία κάνει ένα σημαντικό βήμα στη συσχέτιση ανώτερων λειτουργιών, συγκεκριμένα της οπτικής προσοχής, με την απόκριση του αμφιβληστροειδούς. Περισσότερη διερεύνηση, με ειδικά διαμορφωμένα σύνολα δεδομένων και με καταγραφές μεγαλύτερης χρονικής διάρκειας για πολλά κύτταρα θα είναι σημαντική για την εξαγωγή αποτελεσμάτων που μπορούν να βελτιώσουν την όραση μέσω εμφυτευμάτων.

## **Λέξεις κλειδιά**

μοντέλα αμφιβληστροειδούς, προσθετική αμφιβληστροειδούς, υποδεκτικό πεδίο, αντίστροφη συσχέτιση, συνελκτικά νευρωνικά δίκτυα, προεξοχή, ηλεκτροφυσιολογία, οπτική προσοχή, εμφυτεύματα αμφιβληστροειδούς



## Abstract

The process that enables humans and other living organisms to perceive the environment and interpret the visual stimuli begins in the eye. There, and specifically in the retina, a multilayered and complex processing takes place. Photoreceptors are characterized by sensitivity to light and are the first to detect changes in the field of view. Various cell types process the signal that is finally transmitted to the retinal output, the ganglion cells which send the encoded information to the brain for further processing.

The technology of retinal prosthesis seeks to provide therapy to diseases that hurt the photoreceptors and disrupt this operation. The goal is to bypass the damaged tissue and activate the ganglion cells with the appropriate electrical signals.

For a human, it is not necessary to scan the visual scene in its entirety in order to comprehend what they see. The brain forms a saliency map that represents the property of an image point to stand out from its neighbors. In this way a human, through successive eye movements, identifies important areas and focuses on them.

In this Diploma Thesis we study the relationship between visual saliency and retinal response. We first develop a retina model by training a convolutional neural network in experimental biological data of ganglion cell responses to natural images. By utilizing the trained model's ability to mimic retinal behavior, we estimate and visualize the receptive fields of the neurons through an established technique called reverse correlation, thus estimating the basic relationship between stimulation and response. The receptive field of a retinal ganglion cell is defined as the area of the visual field in which a change will alter its response. We examined, for each cell, the form of its receptive field in comparison to substantiated in theoretical studies analysis of receptive fields. We detected a well-defined area of interest, but not for all neurons. Through the analysis of the time period in which changes affect the responses, we set a specific value for the cells' memory and found that high model performance did not lead to the expected receptive fields.

We then calculate the saliency map for an image dataset through a pre-trained artificial neural network. We also project the same dataset on the ganglion cells, getting their responses through our model. We try to find out if the property of an object to stand out from its neighbors is reflected in the firing rate of ganglion cells. In this way we could stimulate the neurons in a targeted and effective way and potentially improve retinal implants. In particular, we compare the mean response with that in the most salient regions, while studying the correlation between action potential frequency and intensity of salience. We found no consistency in the behavior between the cells. Some increased their response when they detected a salient object in more cases than others, and so they had stronger correlation coefficients. Cells that had an average response below their steady-state value showed a tendency to increase their rate when areas with maximal salience were projected on them, showing positive coefficients. On the contrary, cells that chose to send action potentials at higher frequencies, reduced their response in the important areas, showing negative

coefficients. The values of the metrics were low, so we were not able to claim that the retina is involved in shaping the saliency map in such a direct way.

The thesis takes an important step in correlating superior cognitive functions, namely visual attention, with retinal response. Further research, with specifically designed datasets and longer recordings for many cells will be important to extract results that can improve vision through retinal implants.

### **Key words**

retina models, retinal prosthesis, receptive field, reverse correlation, convolutional neural networks, saliency, electrophysiology, visual attention, retina implants



## Ευχαριστίες

Με την ολοκλήρωση της διπλωματικής μου εργασίας φτάνει στο τέλος της και η περίοδος της προπτυχιακής μου φοίτησης. Θεωρώ απαραίτητο να αναφέρω και να ευχαριστήσω ορισμένους ανθρώπους που με συνόδευαν όλα αυτά τα χρόνια και, ο καθένας με τον τρόπο του, συνέβαλαν ώστε να ολοκληρώσω με επιτυχία αυτό το σημαντικό κεφάλαιο της ζωής μου.

Αρχικά, θα ήθελα να ευχαριστήσω την καθηγήτριά μου κ. Κωνσταντίνα Νικήτα, η οποία με καθοδήγησε στην επιλογή ενός πολύ ενδιαφέροντος και σύγχρονου θέματος. Πολλές και θερμές ευχαριστίες οφείλω, φυσικά, και στον Υ.Δ. Νίκο Μελανίτη-Παρασκευά, για την υπομονή και την ενθάρρυνσή του, την αμεσότητα στην επικοινωνία και την πολύτιμη βοήθεια σε οποιοδήποτε εμπόδιο ανέκυπτε, κατά το διάστημα της συνεργασίας μας. Επίσης, να πω ένα ευχαριστώ στο συμφοιτητή μου Νικόλα Παπαδόπουλο για τις χρήσιμες συζητήσεις μας, σχετικές ή μη με το αντικείμενο των εργασιών μας.

Ύστερα, θα ήθελα να ευχαριστήσω τους φίλους μου, που μαζί περάσαμε όλα αυτά τα χρόνια στη σχολή και ήταν πάντα εκεί να με βοηθάνε να ξεφύγω για λίγο από τις υποχρεώσεις και να μου δίνουν το κίνητρο να προχωράω. Βασίλη, Νίκο, Αναστάση, Ευγένιε, Γιάννη, Ηλιανέ, Αλέξανδρε, σας ευχαριστώ.

Κλείνοντας, δε θα μπορούσα να μην αναφερθώ στην οικογένειά μου. Ένα μεγάλο ευχαριστώ στους γονείς μου Νέλλη και Αριστοτέλη, στην αδελφή μου Αλκμήνη και στο θείο μου Νίκο, που με στηρίζουν αδιάλειπτα και μου παρέχουν απλόχερα οτιδήποτε χρειαστώ. Η εργασία είναι αφιερωμένη και στη γιαγιά μου Αλκμήνη, στον παππού μου Γιώργο και στη γιαγιά μου Ντόλλυ, που δε θα με δουν να αποφοιτώ, με πρόλαβαν όμως ως φοιτητή ενός από τα πιο διακεκριμένα ιδρύματα της χώρας.

Γεώργιος Α. Νακόπουλος,

Αθήνα, 3 Φεβρουαρίου 2021





# Περιεχόμενα

Περίληψη . . . . .	5
Abstract . . . . .	8
<b>1. Εισαγωγή</b>	
1.1 Αντικείμενο της Εργασίας - Κίνητρο . . . . .	16
1.2 Διάρθρωση της Εργασίας . . . . .	17
<b>2. Θεωρητικό Υπόβαθρο</b>	
2.1 Εισαγωγή . . . . .	19
2.2 Οπτικό Σύστημα, Οφθαλμός και Αμφιβληστροειδής χιτώνας . . . . .	20
2.3 Προσθετική Αμφιβληστροειδούς . . . . .	31
2.4 Βιολογικά πειράματα Ηλεκτροφυσιολογίας . . . . .	37
2.5 Οπτική Προσοχή και Προεξοχή . . . . .	38
2.6 Βαθιά Μάθηση και Συνελικτικά Νευρωνικά Δίκτυα . . . . .	44
<b>3. Μέθοδοι</b>	
3.1 Πείραμα καταγραφής των αποκρίσεων γαγγλιακών κυττάρων . . . . .	50
3.2 Περιγραφή Δεδομένων . . . . .	51
3.3 Είσοδος - Έξοδος του δικτύου . . . . .	53
3.4 Μοντέλο εκτίμησης της απόκρισης του αμφιβληστροειδούς . . . . .	56
3.5 Μέθοδος υπολογισμού υποδεκτικού πεδίου . . . . .	62
3.6 Μοντέλο υπολογισμού χάρτη προεξοχής σε εικόνες . . . . .	65
3.7 Συντελεστές συσχέτισης μεταξύ προεξοχής και νευρικής απόκρισης . . . . .	68
<b>4. Αποτελέσματα - Συμπεράσματα</b>	
4.1 Απεικόνιση υποδεκτικών πεδίων και αξιολόγηση του μοντέλου . . . . .	71
4.2 Επιλογή έγκυρων νευρώνων . . . . .	77
4.3 Πρόβλεψη απόκρισης για το σύνολο δεδομένων προεξοχής . . . . .	84
4.4 Οι νευρώνες είναι ευαίσθητοι σε αλλαγές στο χρόνο . . . . .	86
4.5 Σύγκριση έντασης προεξοχής και ρυθμού πυροδότησης των νευρώνων . . . . .	93

<b>5. Επίλογος</b>	
5.1 Σύνοψη και Συμβολή της Εργασίας . . . . .	102
5.2 Μελλοντικές κατευθύνσεις . . . . .	104
<b>Βιβλιογραφία</b> . . . . .	106



# Κεφάλαιο 1

## Εισαγωγή

### 1.1 Αντικείμενο της Εργασίας - Κίνητρο

Οι ταχύτατες εξελίξεις στις περιοχές των υπολογιστικών συστημάτων, της ανάλυσης δεδομένων και της μικροηλεκτρονικής έχουν επιτρέψει στην προσομοίωση φυσιολογικών συστημάτων να γνωρίσει μια αξιοσημείωτη πρόοδο. Ο συγκεκριμένος τομέας επιδιώκει να περιγράψει και να μιμηθεί τη δυναμική συμπεριφορά βιολογικών και φυσιολογικών συστημάτων μέσα από υπολογιστικά μοντέλα, μαθηματικές αναλύσεις και μηχανικά ανάλογα. Ο πρωταρχικός στόχος είναι η κατανόηση της δομής και της λειτουργίας των συστημάτων αυτών η οποία θα προσφέρει νέα εργαλεία στη διάγνωση, πρόληψη και θεραπεία παθήσεων και ασθενειών [1].

Το οπτικό νευρικό σύστημα, στο ευρύτερο πλαίσιο της μελέτης του εγκεφάλου, έχει αποτελέσει πεδίο πολυετούς και εντατικής έρευνας, κυριότερα ο αμφιβληστροειδής. Η πολυεπίπεδη αρχιτεκτονική του και η εντυπωσιακά πολύπλοκη υπολογιστική διαδικασία που επιτελεί έχουν τεκμηριωθεί σε μεγάλο βαθμό [3]. Αυτό οφείλεται μεταξύ άλλων και στις προσπάθειες προσομοίωσης του ανθρώπινου οφθαλμού, με συστήματα που μιμούνται τους διαφορετικούς τύπους κυττάρων και επεξεργάζονται οπτικά σήματα. Την τελευταία δεκαετία, τα βαθιά τεχνητά νευρωνικά δίκτυα επιτρέπουν το συνδυασμό τόσο της προσομοίωσης όσο και της πρόβλεψης της απόκρισης ενός τέτοιου συστήματος με υψηλή ακρίβεια. Επιπλέον, μαθαίνοντας μέσα από φυσικές εικόνες, είναι ικανά να αναπαράγουν βιολογικά φαινόμενα που προηγούμενα μοντέλα αδυνατούσαν, εμπλουτίζοντας περαιτέρω τις γνώσεις μας πάνω στο πρόβλημα [59].

Ωστόσο, αρκετές πτυχές των μηχανισμών με τους οποίους λειτουργεί η οπτική αντίληψη παραμένουν ακόμα άγνωστες. Ως εκ τούτου, η θεραπεία παθήσεων όπως η μελαγχρωστική αμφιβληστροειδοπάθεια (Retinitis Pigmentosa) και η εκφύλιση της ωχράς κηλίδας (Age-related Macular Degeneration) μέσω εμφυτευμάτων αμφιβληστροειδούς αντιμετωπίζει αρκετές δυσκολίες. Τα συστήματα αυτά αποτελούνται από μία κάμερα, μία μονάδα επεξεργασίας και μία μονάδα ενεργοποίησης, συνήθως μικροηλεκτρόδια. Η ιδέα είναι πως οι εικόνες που φτάνουν στην είσοδο υφίστανται κατάλληλη επεξεργασία ώστε να δημιουργηθεί η αλληλουχία παλμών που θα κάνει τον εγκέφαλο να αντιληφθεί σωστά την πληροφορία που δέχεται [22]. Η πλειοψηφία των συσκευών προσφέρει πολύ περιορισμένη όραση και δεν έχει επιτευχθεί κάτι καλύτερο από αναγνώριση φωτεινών σημείων, απλών σχημάτων και ακμών σε υψηλή αντίθεση. Είναι γεγονός πως μπορούν να προσφέρουν μια ευκαιρία για πιο φυσιολογική ζωή σε άτομα με προχωρημένο στάδιο εκφύλισης του αμφιβληστροειδούς, η οπτική αντίληψη που αισθάνονται ωστόσο απέχει πάρα πολύ από αυτή ενός υγιούς ανθρώπου [24].

Ένα πολύ κρίσιμο ζήτημα είναι πως τα συγκεκριμένα συστήματα προς το παρόν αδυνατούν να ενσωματώσουν τον κώδικα λειτουργίας του αμφιβληστροειδούς, δηλαδή να προσομοιώσουν πλήρως τη διαφορετική συμπεριφορά που επιδεικνύουν οι πολυάριθμες ομάδες νευρώνων. Το σύστημα στην ουσία δε γνωρίζει πώς πρέπει να διεγείρει τον κάθε τύπο γαγγλιακών κυττάρων ώστε ο εγκέφαλος να λάβει τα οπτικά σήματα στη σωστή κωδικοποίηση [23], [26]. Έχει βρεθεί πως ο αμφιβληστροειδής δε μεταβιβάζει απλώς την πληροφορία προς τα ανώτερα επίπεδα, αλλά αποδομεί την οπτική σκηνή και εξάγει χαρακτηριστικά μέσα από μη γραμμικούς υπολογισμούς, ιδιότητα που διατηρείται μεταξύ των ειδών [17], [18].

Παράλληλα, ενδιαφέρον έχει αναπτυχθεί από μηχανικούς και νευροεπιστήμονες σχετικά με την ακολουθία σημείων μιας εικόνας στα οποία είναι πιθανότερο ένας άνθρωπος να εστιάσει το βλέμμα του, μια ιδέα που θα μπορούσε να εφαρμοστεί και σε συστήματα προσθετικής μέσα από ένα μοντέλο πρόβλεψης. Ένα σύστημα που, μιμούμενο την ανθρώπινη συμπεριφορά, θα εκτιμά με ακρίβεια τις περιοχές που θεωρούνται πιο σημαντικές από κάποιες άλλες, ή με άλλα λόγια έχουν μεγαλύτερο μέτρο προεξοχής, θα μπορεί να αναλύσει πιο αποτελεσματικά τις εικόνες απορρίπτοντας περιττή πληροφορία. Θα μπορεί να εξάγει τα χαρακτηριστικά που προσδίδουν αυξημένη προεξοχή στις περιοχές και να ενεργοποιήσει πιο στοχευμένα νευρώνες. Η πλειοψηφία των υλοποιημένων μοντέλων πρόβλεψης δημιουργούν χάρτες προεξοχής, προσδίδοντας σε κάθε σημείο μια τιμή που μας λέει το πόσο ξεχωρίζει από τα γειτονικά του και άρα πόσο πιθανό είναι να ελκύσει την προσοχή. Στον άνθρωπο, θεωρείται πως ο χάρτης προεξοχής δημιουργείται στα ανώτερα συστήματα του εγκεφάλου [32].

Συνδυάζοντας αυτές τις δύο κατευθύνσεις, θεωρούμε πως θα μπορούσε να εξεταστεί αν ο αμφιβληστροειδής συνεισφέρει με κάποιο τρόπο στη διαδικασία διαμόρφωσης του χάρτη προεξοχής. Τα υποδεκτικά πεδία των γαγγλιακών κυττάρων καλύπτουν το οπτικό πεδίο σε όλη την έκτασή του, σχηματίζοντας μια διάταξη «μωσαϊκού», με το καθένα να επηρεάζεται από αλλαγές σε πολύ συγκεκριμένο τμήμα της εικόνας [6], [16]. Αν τα γαγγλιακά κύτταρα που σύμφωνα με το υποδεκτικό τους πεδίο ενδιαφέρονται για τις περισσότερες προεξέχουσες περιοχές μιας εικόνας, παρουσιάζουν αυξημένο ρυθμό πυροδότησης σε σχέση με τα υπόλοιπα, τότε μέσω της διάταξης ηλεκτροδίων θα μπορούσαμε να ενεργοποιήσουμε τεχνητά τους αρμόδιους νευρώνες, βοηθώντας τον ασθενή να δει την αντίστοιχη οπτική πληροφορία.

## 1.2 Διάρθρωση της Εργασίας

Στο κεφάλαιο 2 παρουσιάζουμε το απαραίτητο θεωρητικό υπόβαθρο προκειμένου ο αναγνώστης να κατανοήσει τις έννοιες και τις ορολογίες που χρησιμοποιούνται στη συνέχεια της εργασίας, όπως τα γαγγλιακά κύτταρα, το υποδεκτικό πεδίο, ο χάρτης προεξοχής, τα εμφυτεύματα αμφιβληστροειδούς, τα συνελκτικά δίκτυα και οι ηλεκτροφυσιολογικές καταγραφές.

Στο **κεφάλαιο 3** αναλύουμε τα πειραματικά δεδομένα και τις μεθόδους που εφαρμόζουμε. Περιγράφουμε τις τεχνικές προεπεξεργασίας που χρησιμοποιούμε και, αφού κάνουμε μια σύντομη εισαγωγή σχετικά με τις πρώτες προσπάθειες προσομοίωσης του αμφιβληστροειδούς και τα μειονεκτήματά τους, εστιάζουμε στα βαθιά νευρωνικά δίκτυα και στα οφέλη τους. Παρουσιάζουμε την αρχιτεκτονική την οποία επιλέξαμε να μελετήσουμε και παρέχουμε τις λεπτομέρειες της εκπαίδευσης και της αξιολόγησης. Στη συνέχεια εισάγουμε τη θεωρία πίσω από τη μέθοδο εύρεσης του υποδεκτικού πεδίου που χρησιμοποιήσαμε και παρουσιάζουμε το μοντέλο πρόβλεψης προεξοχής που αξιοποιήσαμε για να λάβουμε χάρτες προεξοχής για εικόνες. Τέλος παραθέτουμε τους ορισμούς τριών συντελεστών συσχέτισης που χρησιμοποιούμε για τη σύγκριση του μέτρου προεξοχής με τη νευρική απόκριση.

Στο **κεφάλαιο 4** παρουσιάζουμε και σχολιάζουμε τα αποτελέσματά μας. Ξεκινάμε από την αξιολόγηση της επίδοσης του μοντέλου και τον προσδιορισμό παραμέτρων που ερμηνεύονται από χαρακτηριστικά της απόκρισης του βιολογικού αμφιβληστροειδούς. Απεικονίζουμε τα υποδεκτικά πεδία που υπολογίστηκαν μέσω STA για τον αμφιβληστροειδή και ασχολούμαστε εκτενώς με τη μορφή των χωρικών και χρονικών φίλτρων που τα περιγράφουν. Επιλέγουμε τα αξιόπιστα κύτταρα που έχουν ευκρινείς και σαφώς ορισμένες περιοχές ενδιαφέροντος και προβάλλουμε σε αυτά εικόνες για τις οποίες έχουμε υπολογίσει το χάρτη προεξοχής. Στεκόμαστε ιδιαίτερα στην επιλογή του χρονικού διαστήματος για το οποίο θα προβάλλουμε τις εικόνες, επιδιώκοντας διαφορετικές εισοδοί να προκαλούν αποκρίσεις αυξημένης διακύμανσης. Ελέγχουμε αν τα κύτταρα οδηγούνται σε ενεργοποίηση άνω της μέσης όταν εστιάζουν σε περιοχές που θα ελκύσουν το βλέμμα μας, καθώς και αν υπάρχει κάποια συσχέτιση μεταξύ απόκρισης και έντασης προεξοχής. Καταλήγοντας, κάνουμε μια προσπάθεια να ερμηνεύσουμε τα αποτελέσματα που προέκυψαν από τις συγκρίσεις.

Στο **κεφάλαιο 5** ανακεφαλαιώνουμε τα συμπεράσματα που προέκυψαν από τα πειράματά μας και προτείνουμε ορισμένες βελτιώσεις και μελλοντικές κατευθύνσεις που μπορεί να βοηθήσουν περαιτέρω τον τομέα της προσθετικής αμφιβληστροειδούς.

## Κεφάλαιο 2

### Θεωρητικό Υπόβαθρο

#### 2.1 Εισαγωγή

Σε αυτό το κεφάλαιο παρουσιάζεται το απαραίτητο θεωρητικό υπόβαθρο προκειμένου ο αναγνώστης να κατανοήσει τις έννοιες και τις ορολογίες που χρησιμοποιούνται στη συνέχεια της εργασίας. Αρχικά αναφερόμαστε στη φυσιολογία του ανθρώπινου οφθαλμού και του οπτικού συστήματος και εστιάζουμε στο τμήμα του που μας ενδιαφέρει περισσότερο, τον αμφιβληστροειδή χιτώνα. Παρουσιάζουμε ξεχωριστά τις πέντε βασικές κατηγορίες κυττάρων που τον απαρτίζουν και συμβάλλουν στην οπτική αντίληψη, επικεντρώνοντας το ενδιαφέρον στην τελευταία κατά σειρά, τα γαγγλιακά κύτταρα. Μεταξύ άλλων αναφερόμαστε στη βασική κατηγοριοποίησή τους σε ON, OFF, ON/OFF ανάλογα με το αν είναι ευαίσθητα κυρίως σε αύξηση φωτός, μείωση, ή και στα δύο, αντίστοιχα. Στη συνέχεια εισάγουμε τη θεμελιώδη έννοια του υποδεκτικού πεδίου που είναι ο κεντρικός άξονας του σχολιασμού μας σχετικά με την κυτταρική συμπεριφορά και τέλος θίγουμε ένα ενδιαφέρον δίλημμα, αν ο ρόλος του αμφιβληστροειδούς είναι απλώς να μεταβιβάζει την οπτική πληροφορία προς τον εγκέφαλο, λειτουργώντας ως ένας αισθητήρας έντασης φωτός ή αν μέσα από τα πολύπλοκα και ανεξάρτητα κυκλώματα που τον απαρτίζουν εντοπίζει διαφορετικά χαρακτηριστικά μέσω μιας σύνθετης αποκωδικοποίησης της εικόνας.

Στη δεύτερη ενότητα παρέχουμε πληροφορίες πάνω στον τεχνητό αμφιβληστροειδή οι οποίες θα μας βοηθήσουν να συζητήσουμε πιθανά οφέλη της σύγκρισης μεταξύ κυτταρικής απόκρισης και σημείων εστίασεως του οφθαλμού. Ξεκινάμε με τις δύο ασθένειες τις οποίες στοχεύει να καταπολεμήσει και απαριθμούμε τα βασικά μέρη ενός τυπικού συστήματος προσθετικής, ενώ παράλληλα αναφέρουμε τις τρεις κατηγορίες συστημάτων ανάλογα με το σημείο στο οποίο εμφυτεύονται τα ηλεκτρόδια. Τέλος, ονομάζουμε ορισμένα από τα προβλήματα που αποτρέπουν τα εμφυτεύματα αμφιβληστροειδούς από το να είναι αποτελεσματικά και εστιάζουμε στην αδυναμία ενσωμάτωσης του ακριβούς κώδικα λειτουργίας του αμφιβληστροειδούς, που αποτέλεσε και κίνητρο για την εργασία μας.

Η τρίτη ενότητα είναι αφιερωμένη στην εύρεση των σημείων εστίασεως ενός παρατηρητή σε μια εικόνα με εισαγωγή εννοιών όπως οι σακκαδικές κινήσεις και ο χάρτης προεξοχής. Παρουσιάζονται οι δύο βασικές ομάδες χαρακτηριστικών που θεωρείται πως διαμορφώνουν τη διαδικασία με την οποία επεξεργαζόμαστε μια εικόνα όπως επίσης και υλοποιήσεις που επιδιώκουν να μιμηθούν την ανθρώπινη συμπεριφορά, από τις πρώτες τεχνικές υπολογιστικής όρασης και επεξεργασίας εικόνας μέχρι τις πιο σύγχρονες υλοποιήσεις με συνελκτικά νευρωνικά δίκτυα.

Στις δύο τελευταίες ενότητες του κεφαλαίου γίνεται μια εισαγωγή στη βαθιά μάθηση και τα συνελκτικά νευρωνικά δίκτυα και παρουσιάζονται εν συντομία οι καθιερωμένοι τρόποι

διεξαγωγής βιολογικών πειραμάτων Ηλεκτροφυσιολογίας, μιας και τα κυτταρικά δεδομένα μας προήλθαν από τέτοιου είδους καταγραφές.

## 2.2 Οπτικό Σύστημα, Οφθαλμός και Αμφιβληστροειδής χιτώνας

Το οπτικό σύστημα είναι το τμήμα του κεντρικού νευρικού συστήματος που επιτρέπει στους οργανισμούς να αντιλαμβάνονται τον εξωτερικό κόσμο μέσω επεξεργασίας του φωτός που ανακλάται πάνω στα αντικείμενα. Τα βασικότερα μέρη του είναι ο οφθαλμός, ο πλευρικός γονιδιακός πυρήνας (Lateral Geniculate Nucleus - LGN) και οι πέντε περιοχές του οπτικού φλοιού (visual cortex): ο πρωτοταγής οπτικός φλοιός (primary visual cortex - V1) και οι οπτικές περιοχές 2, 3, 4, 5 (V2, V3, V4, V5). Συνολικά αποτελεί ένα ιδιαίτερα πολύπλοκο δίκτυο από μεμβράνες, κύτταρα και νευρώνες που συνεργάζονται μέσα από βιολογικές, ηλεκτρικές και νευροχημικές διεργασίες.

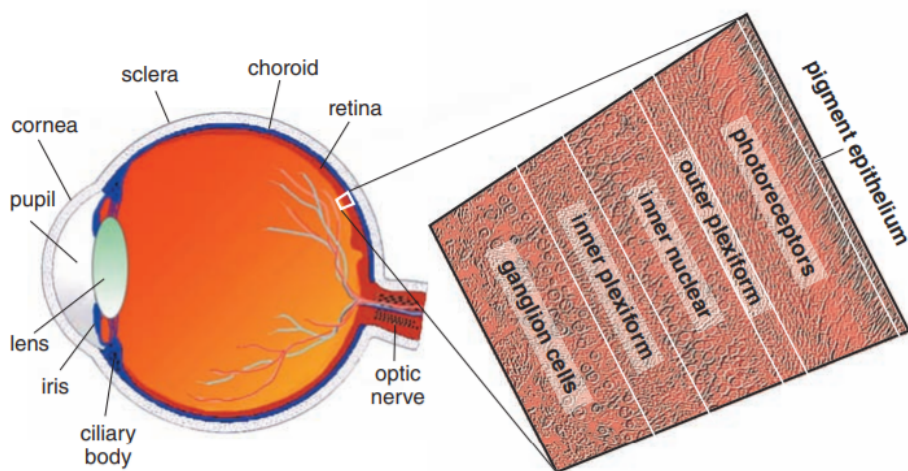
Ο οφθαλμός, ως το πρώτο στάδιο, είναι και ο αισθητήρας του οπτικού συστήματος. Εξωτερικά καλύπτεται από μια σκληρή διάφανη επιφάνεια, τον κερατοειδή χιτώνα που, σε συνεργασία με το φακό, εστιάζει τις φωτεινές ακτίνες στο πίσω μέρος του οφθαλμού και συγκεκριμένα στο βοθρίο. Ο φακός έχει την ικανότητα να μεταβάλλει το σχήμα και την καμπυλότητα του για να διατηρεί τη συνεχόμενη εστίαση της εικόνας στον αμφιβληστροειδή. Το βοθρίο είναι μια συγκεκριμένη μικρή περιοχή του αμφιβληστροειδούς που συγκεντρώνει πολύ μεγαλύτερο αριθμό φωτούποδοχέων και για αυτό είναι υπεύθυνο για την υψηλή ανάλυση που διαθέτει η κεντρική μας όραση. Για να εκμεταλλευόμαστε την ιδιότητα αυτή, το βοθρίο κατευθύνεται πάντα προς το στιγμιαίο σημείο ενδιαφέροντος στο οπτικό πεδίο προκειμένου να πραγματοποιείται η βέλτιστη επεξεργασία του [2].

Ο πλευρικός γονιδιακός πυρήνας βρίσκεται στο θάλαμο του εγκεφάλου και λαμβάνει είσοδο από τον αμφιβληστροειδή. Ο κεντρικός του ρόλος είναι να αναμεταδίδει τα σήματα που δέχεται προς τις ανώτερες περιοχές του οπτικού συστήματος στον εγκέφαλο και συγκεκριμένα στον πρωτοταγή οπτικός φλοιό [2]. Ο πρωτοταγής οπτικός φλοιός είναι η καλύτερα μελετημένη οπτική περιοχή στον εγκέφαλο. Επιτελεί αρκετά σύνθετους υπολογισμούς και εξειδικευμένη επεξεργασία που αφορά κυρίως την κίνηση αντικειμένων στο οπτικό πεδίο.

Η διαδικασία της κωδικοποίησης των οπτικών ερεθισμάτων ξεκινά στον αμφιβληστροειδή και για αυτό αποτελεί ίσως το πιο ενδιαφέρον τμήμα του οφθαλμού. Έχει τη μορφή ενός πολύ λεπτού ιστού που καλύπτει την οπίσθια επιφάνεια του οφθαλμού. Δεδομένου του πάχους και της διαμέτρου του, μόλις μερικών χιλιοστών, είναι ένα πολύπλοκο δίκτυο εντυπωσιακής οργάνωσης που δομείται από πολλά παράλληλα και ανατομικά ισοδύναμα μικροκυκλώματα [3]. Διαθέτει κύτταρα και νευρώνες που ταξινομούνται σε πέντε κύριες κατηγορίες: φωτούποδοχείς, οριζόντια, βραχύινα, δίπολα και γαγγλιακά κύτταρα. Απαρτίζεται από ένα σύνολο από διαχωρίσιμα στρώματα που εντάσσονται σε δύο κατηγορίες, τις κοκκώδεις στοιβάδες που περιλαμβάνουν τα σώματα των κυττάρων και τις συναπτικές στοιβάδες που περιλαμβάνουν τους άξονες και τους



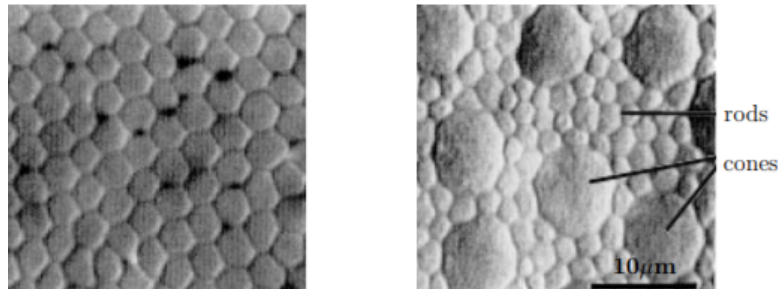
δενδρίτες τους [2]. Ο αμφιβληστροειδής έχει μελετηθεί πολύ εντατικά τα τελευταία 100 χρόνια και έχει συμβάλει καθοριστικά στη μελέτη του τρόπου με τον οποίο πληθυσμοί νευρώνων κωδικοποιούν την πληροφορία συλλογικά. Μέχρι σήμερα έχουν βρεθεί τουλάχιστον 60 διαφορετικοί τύποι κυττάρων στον αμφιβληστροειδή των θηλαστικών, καθένα με διαφορετική δομή, λειτουργία και ρόλο, ενώ υπάρχουν και άλλοι τύποι που δεν έχουν ανακαλυφθεί ακόμη, κυρίως γαγγλιακά κύτταρα. Δεν αποτελούν ωστόσο πληθυσμούς που συνιστούν την πλειοψηφία. Αν και αρκετές λεπτομέρειες παραμένουν άγνωστες, θεωρείται πως σύντομα ο αμφιβληστροειδής θα είναι η πρώτη σύνθετη δομή του κεντρικού νευρικού συστήματος που θα έχει μελετηθεί πλήρως, θα έχουμε δηλαδή μια ολοκληρωμένη εικόνα τόσο της λειτουργίας του όσο και των ξεχωριστών πληθυσμών κυττάρων που τον απαρτίζουν [3], [4].



**Σχήμα 1:** Τομή του ανθρώπινου οφθαλμού μαζί με ένα τμήμα του αμφιβληστροειδούς σε μεγέθυνση που αποκαλύπτει τα επίπεδα του. Λίγο πάνω από το οπτικό νεύρο φαίνεται και το βοθρίο [5].

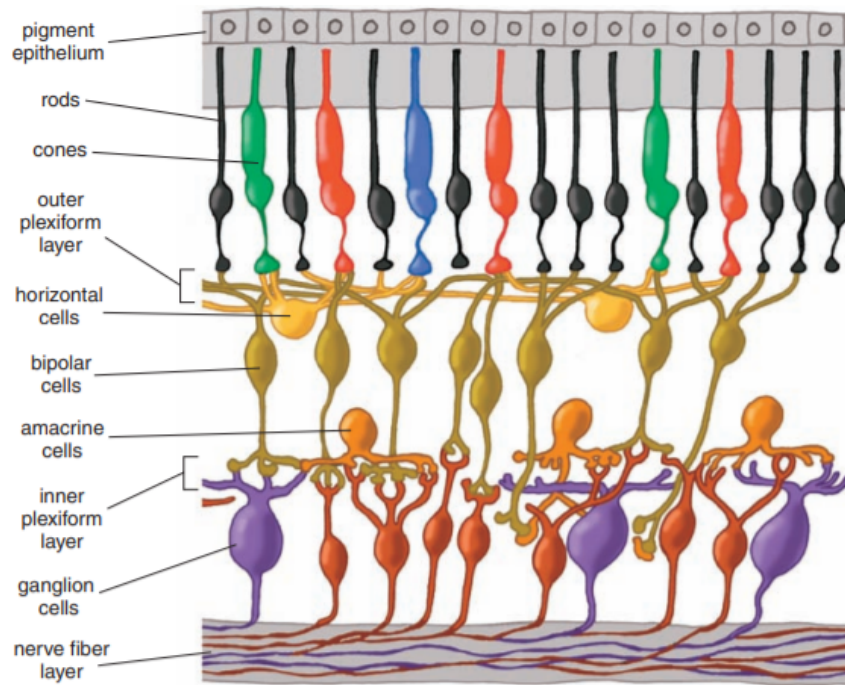
Οι φωτοϋποδοχείς είναι το πρώτο επίπεδο που θα συναντήσουν τα φωτόνια. Χωρίζονται σε δύο κατηγορίες, τα ραβδία και τα κωνία. Τα ραβδία ειδικεύονται στην όραση υπό χαμηλό φωτισμό και όλα περιέχουν μια συγκεκριμένη και ιδιαίτερα φωτοευαίσθητη πρωτεΐνη, τη ροδοψίνη, γεγονός που τα οδηγεί στο να διεγείρονται με μεγάλη ευκολία. Τα κωνία από την άλλη συμβάλλουν στην έγχρωμη όραση. Μπορεί να είναι ευαίσθητα είτε σε μικρό είτε σε μεγάλο μήκος κύματος. Κάθε τύπος διαθέτει μια διαφορετική οψίνη που απορροφά φως σε διαφορετικό φάσμα. Αυτό αποτελεί και τη βάση για την αντίληψη των χρωμάτων στην όραση όπως επίσης και ένα πολύτιμο εργαλείο για την εύκολη ταυτοποίηση των ξεχωριστών τύπων κωνίων. Δεν είναι το ίδιο ευαίσθητα με τα ραβδία. Τα ραβδία και τα κωνία δεν κατανέμονται ομοιόμορφα στον αμφιβληστροειδή. Σε είδη που διαθέτουν βοθρίο (όπως ο άνθρωπος αλλά όχι το ποντίκι), αυτό περιέχει αποκλειστικά κωνία, ενώ η υπόλοιπη επιφάνεια του χιτώνα καταλαμβάνεται κυρίως από ραβδία, τα οποία για αυτό το λόγο

συνεισφέρουν και στην περιφερική όραση. Σε έναν τυπικό αμφιβληστροειδή τα ραβδία είναι περίπου 20 φορές περισσότερα από τα κωνία ενώ έχει βρεθεί ότι περιέχεται ένας μοναδικός τύπος ραβδίων. Και οι δύο τύποι φωτοϋποδοχέων αποκρίνονται με βαθμωτά δυναμικά και συγκεκριμένα υπεροπολώσεις ενώ ο άνθρωπος διαθέτει περίπου εκατό εκατομμύρια φωτοϋποδοχείς [2], [4], [6].



**Σχήμα 2:** Αριστερά: Κωνία στο βοθρίο. Δεξιά: Ραβδία και κωνία στην περιφέρεια του αμφιβληστροειδούς. Είναι ορατή η μεγάλη πυκνότητα του βοθρίου σε κωνία και η διαφορά στο μέγεθος και στο πλήθος μεταξύ ραβδίων και κωνίων στην υπόλοιπη περιοχή [2].

Τα οριζόντια κύτταρα δεν παρουσιάζουν μεγάλη ποικιλομορφία καθώς η πλειοψηφία των θηλαστικών διαθέτει μόνο δύο είδη, με κάποιες εξαιρέσεις όπως τα ποντίκια που διαθέτουν ένα. Αυτό που κάνουν είναι να λαμβάνουν ερεθίσματα στην είσοδο τους από τα ραβδία και τα κωνία, αρκετά σε αριθμό, και να παρέχουν ανατροφοδότηση καταστολής πίσω σε αυτά. Η επικρατέστερη ερμηνεία για τον μηχανισμό αυτό είναι ότι έτσι κρατούν το σήμα εισόδου μέσα στο εύρος τιμών στο οποίο μπορεί να λειτουργήσει το κύκλωμα, εφαρμόζοντας δηλαδή δυναμική προσαρμογή στην ένταση του φωτός. Μετρώντας τοπικά τη μέση τιμή φωτεινότητας στην εικόνα και αφαιρώντας αναλόγως μια τιμή από την έξοδο των φωτοϋποδοχέων, αποτρέπουν τον κορεσμό της απόκρισης που θα προκαλούνταν σε περιοχές που υπάρχει ένα πολύ φωτεινό αντικείμενο. Με άλλα λόγια, τα οριζόντια κύτταρα πραγματοποιούν μείωση της ευαισθησίας, αποκλειστικά στα σημεία που απαιτείται και όχι σε όλη την εικόνα. Επίσης, μια επιπρόσθετη λειτουργία τους, ως επακόλουθο του παραπάνω μηχανισμού, θεωρείται πως είναι η ενίσχυση της αντίθεσης μεταξύ γειτονικών σημείων, δηλαδή η ανίχνευση ακμών και χωρικών λεπτομερειών, φαινόμενο που αποτελεί τη βάση του περίφημου μηχανισμού «κέντρου - περιφέρειας» που συναντάται στα γαγγλιακά κύτταρα και θα εξηγήσουμε [4], [6]. Όπως και οι φωτοϋποδοχείς, τα οριζόντια κύτταρα αποκρίνονται με υπεροπολώσεις [7].



**Σχήμα 3:** Τα κυτταρικά επίπεδα του αμφιβληστροειδούς [5].

Τα δίπολα κύτταρα δρουν με στόχο να μεταφέρουν τα σήματα από τους φωτούποδοχείς στα γαγγλιακά κύτταρα. Υπάρχουν περίπου 12 τύποι κυττάρων με τη συντριπτική πλειοψηφία τους να λαμβάνει είσοδο πρωτίστως από κωνία. Δεν έχουν γίνει ακόμα όλοι οι τύποι κατανοητοί. Κάθε δίπολο κύτταρο λαμβάνει σήμα από έναν μικρό αριθμό κωνίων, μερικές εκατοντάδες στον αριθμό. Μπορούν να είναι ενεργά (ON) ή ανενεργά (OFF) ανάλογα με το αν ενεργοποιούνται (αποπόλωση) ή καταστέλλονται (υπερπόλωση) σε αύξηση της έντασης του φωτός. Η διχοτομία αυτή οφείλεται στο γεγονός ότι στις συνάψεις κωνίου - δίπολου κυττάρου μπορεί να υπάρχουν δύο τύποι νευροδιαβιβαστών. Έχει βρεθεί πως οι δύο αυτές κλάσεις κυττάρων έχουν τους δένδριτες τους σε ξεχωριστά τμήματα της Έσω Συναπτικής Στοιβάδας (Inner Plexiform Layer - IPL), τα ON στο εσωτερικό και τα OFF στο εξωτερικό τμήμα. Μία επιπλέον ταξινόμηση γίνεται βάσει της ταχύτητας με την οποία αδρανοποιούνται μετά από τη διέγερση, παροδικά ή παρατεταμένα. Παρουσιάζουν οργάνωση «κέντρου - περιφέρειας» την οποία οφείλουν στα οριζόντια κύτταρα. Οι συνάψεις που δημιουργούν τα κωνία με τα δίπολα κύτταρα διαμορφώνουν παράλληλα κανάλια πληροφορίας τα οποία λειτουργούν ανεξάρτητα μεταξύ τους. Μέσα από τα κανάλια αυτά κάθε ένας τύπος κυττάρου που επικοινωνεί με το ίδιο κωνίο, μεταβιβάζει προς το εσωτερικό του αμφιβληστροειδούς ένα διαφορετικό στοιχείο ή αλλιώς μια διαφορετική «όψη» της εξόδου αυτού του κωνίου. Ο τρόπος που επιτυγχάνεται αυτό είναι οι διαφορές στη φυσιολογία που επιδεικνύουν οι τύποι των δίπολων κυττάρων προκειμένου να διεγείρονται από ξεχωριστά χαρακτηριστικά του σήματος από τα κωνία [2], [4], [6].

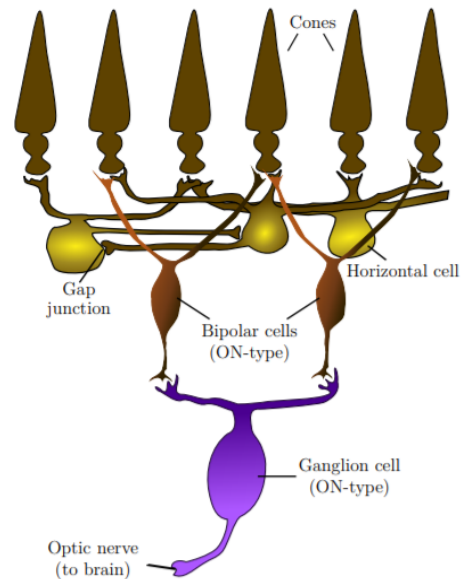
Τα βραχύνια κύτταρα αποτελούν τον πιο ποικιλόμορφο πληθυσμό στον αμφιβληστροειδή, ενώ και οι λειτουργίες τους είναι το ίδιο ιδιαίτερες και σχετικά άγνωστες. Τουλάχιστον 25 διαφορετικά είδη έχουν εντοπιστεί. Είναι ένα ενδιάμεσο στάδιο

επεξεργασίας του σήματος όπως τα οριζόντια κύτταρα με πιο εξειδικευμένες αρμοδιότητες. Έπονται των δίπολων κυττάρων και λαμβάνουν είσοδο από αυτά, ενώ είναι και το τελευταίο επίπεδο πριν τα γαγγλιακά. Βελτιώνουν το σήμα των δίπολων κυττάρων μέσω ανατροφοδότησης πίσω σε αυτά ενώ επικοινωνούν απευθείας με τα γαγγλιακά κύτταρα διαμορφώνοντας τον μηχανισμό «κέντρου-περιφέρειας» τους. Επιπλέον λειτουργίες που επιτελούν είναι η δημιουργία σύνδεσης μεταξύ των ραβδίων και των γαγγλιακών κυττάρων κατά μήκος του αμφιβληστροειδούς, ενώ έχει βρεθεί πως φτιάχνουν συνάψεις με συγκεκριμένα γαγγλιακά κύτταρα που είναι ευαίσθητα στην κίνηση αντικειμένων συνεισφέροντας στην επιλεκτικότητα κατεύθυνσης [3], [4]. Τα βραχύινα κύτταρα μπορούν να προκαλέσουν τόσο υπεροπολώσεις όσο και δυναμικά ενεργείας [7].

Τα γαγγλιακά κύτταρα είναι η μοναδική ομάδα νευρώνων του αμφιβληστροειδούς που επικοινωνεί απευθείας με τον εγκέφαλο. Ο στόχος τους είναι ένας και μοναδικός: να μεταφέρουν την πληροφορία σε κατανοητή μορφή κωδικοποίησης προς τα ανώτερα στρώματα του οπτικού συστήματος για επιπλέον επεξεργασία. Η μορφή αυτή είναι τα δυναμικά ενεργείας (action potentials). Σε συνθήκες ηρεμίας, η διαφορά δυναμικού μεταξύ του εσωτερικού ενός κυττάρου και του χώρου στο εξωτερικό του είναι περίπου  $-70$  mV, με το κύτταρο να ορίζεται ως πολωμένο. Η τιμή αυτή, που ονομάζεται δυναμικό μεμβράνης, ελέγχεται από τα κανάλια ιόντων  $Na^+$  και  $K^+$  που επιτρέπουν ή αποτρέπουν τη ροή ιόντων στο εσωτερικό και το εξωτερικό του κυττάρου ως απόκριση σε μεταβολές δυναμικού και εξωτερικά σήματα. Αν θετικά φορτισμένα ιόντα, υπό τη μορφή ηλεκτρικού ρεύματος εξέλθουν από το κύτταρο μέσω των καναλιών ή αρνητικά φορτισμένα ιόντα εισέλθουν στο εσωτερικό του, το δυναμικό μεμβράνης γίνεται πιο αρνητικό και έχουμε το φαινόμενο της υπερπόλωσης. Κατά την αντίστροφη διαδικασία, το δυναμικό γίνεται λιγότερο αρνητικό και μπορεί να φτάσει και σε θετικές τιμές. Αυτό ονομάζεται αποπόλωση. Αν το κύτταρο αποπολωθεί επαρκώς και το δυναμικό μεμβράνης ξεπεράσει κάποια τιμή κατωφλίου, προκαλείται δυναμικό ενεργείας. Οι νευρώνες είναι μια κλασική κατηγορία κυττάρων που αντιδρούν με δυναμικά ενεργείας, συνεπώς στη συνέχεια της εργασίας που αναφερόμαστε σχεδόν αποκλειστικά σε γαγγλιακά κύτταρα, οι όροι «κύτταρο» και «νευρώνας» είναι εναλλάξιμοι. Η σημαντικότερη ιδιότητα των δυναμικών ενεργείας είναι πως διαδίδονται σε μεγάλες αποστάσεις χωρίς απόσβεση, αποτελώντας τον πιο αξιόπιστο τρόπο μετάδοσης πληροφορίας στο νευρικό σύστημα [8].

Οι περίπου ένα εκατομμύριο άξονες των γαγγλιακών κυττάρων που περιέχουν τα σήματα εξόδου είναι μεγάλοι σε μήκος και συγκεντρώνονται όλοι μαζί διαμορφώνοντας το οπτικό νεύρο, το οποίο εκτείνεται προς τον εγκέφαλο. Αν συγκρίνουμε το πλήθος αυτό με το συνολικό αριθμό των φωτοϋποδοχέων, προκύπτει το συμπέρασμα πως στον αμφιβληστροειδή λαμβάνει χώρα μια ιδιαίτερα σύνθετη επεξεργαστική διαδικασία. Το σημείο συγκέντρωσης ονομάζεται και «τυφλό σημείο» αφού δεν υπάρχουν φωτοϋποδοχείς εκεί. Τα γαγγλιακά κύτταρα λαμβάνουν είσοδο από τα βραχύινα μέσω πλευρικών διασυνδέσεων αλλά κυρίως από τα δίπολα. Ο βασικός διαχωρισμός τους γίνεται ανάλογα με το τμήμα της IPL στο οποίο εντοπίζονται οι δενδρίτες τους, ή αλλιώς με το αν λαμβάνουν είσοδο από ON ή OFF δίπολα κύτταρα. Ορίζονται αντίστοιχα ως ON και OFF γαγγλιακά κύτταρα και παρέχουν ανεξάρτητες μεταξύ τους καλύψεις του οπτικού πεδίου. Έχουν βρεθεί

και κύτταρα που συνδέονται τόσο με ON όσο και με OFF περιοχές. Ορίζονται ως ON-OFF και διαφέρουν ιδιαίτερα στη λειτουργία τους από τα προηγούμενα. Κατηγοριοποίηση γίνεται επιπλέον με βάση το μέγεθος του σώματος, τις συνδέσεις και τις ιδιαιτερότητες στη μορφή της απόκρισής τους σε εξωτερικά ερεθίσματα [2], [9]. Εφόσον τα γαγγλιακά κύτταρα παρουσιάζουν δυναμικά ενεργείας, η ένταση της δραστηριότητάς τους αποτυπώνεται στη συχνότητα με την οποία στέλνουν νευρικούς παλμούς.

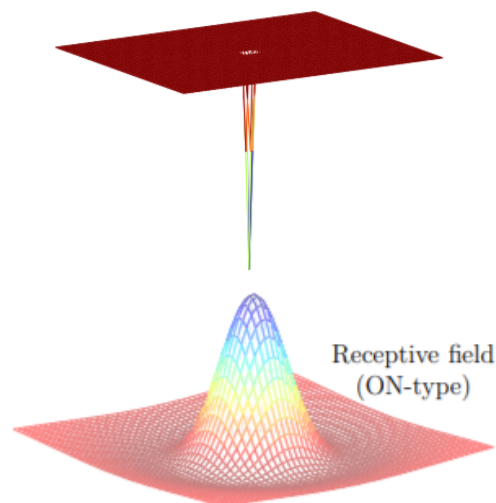


**Σχήμα 4:** Νευρικές διασυνδέσεις ενός ON γαγγλιακού κυττάρου [2].

Η έννοια του υποδεκτικού πεδίου (receptive field) θεωρείται θεμελιώδης για τη νευροεπιστήμη και ο υπολογισμός του είναι η καθιερωμένη μέθοδος όταν μελετάται η σχέση μεταξύ της απόκρισης ενός κυττάρου και της διέγερσης που την προκαλεί. Στο οπτικό σύστημα, υποδεκτικό πεδίο ενός κυττάρου ορίζεται ως η περιοχή του οπτικού πεδίου στην οποία αν εντοπιστεί κάποια μεταβολή ή ερέθισμα, το κύτταρο θα διεγερθεί. Αλλαγές και ερεθίσματα μακριά από το υποδεκτικό πεδίο δεν μπορούν να ενεργοποιήσουν τον νευρώνα απευθείας, μπορούν ωστόσο να επηρεάσουν την απόκρισή του σε διέγερση εντός του πεδίου. Το σχήμα και το μέγεθος του υποδεκτικού πεδίου παρουσιάζουν διαφορές στα διακριτά στάδια του οπτικού συστήματος. Το ίδιο συμβαίνει όμως και εντός του αμφιβληστροειδούς, αφού στο βοθρίο το υποδεκτικό πεδίο των δίπολων και των γαγγλιακών κυττάρων είναι πολύ μικρότερο από ότι στις γύρω περιοχές. Όσο προχωράμε ιεραρχικά, το μέγεθος του υποδεκτικού πεδίου αυξάνει αφού τα κύτταρα ενός επιπέδου λαμβάνουν είσοδο από κύτταρα του προηγούμενου επιπέδου. Στους φωτούποδοχείς το υποδεκτικό πεδίο μπορεί να θεωρηθεί ένα σημείο στο χώρο αφού έρχονται σε άμεση επαφή με τα φωτόνια. Τα δίπολα κύτταρα έχουν μεσαίου μεγέθους υποδεκτικό πεδίο αφού συνδέονται με λίγα κωνία και τα οριζόντια κύτταρα έχουν μεγαλύτερο αφού ενεργοποιούνται από πολλά κωνία [2], [8].

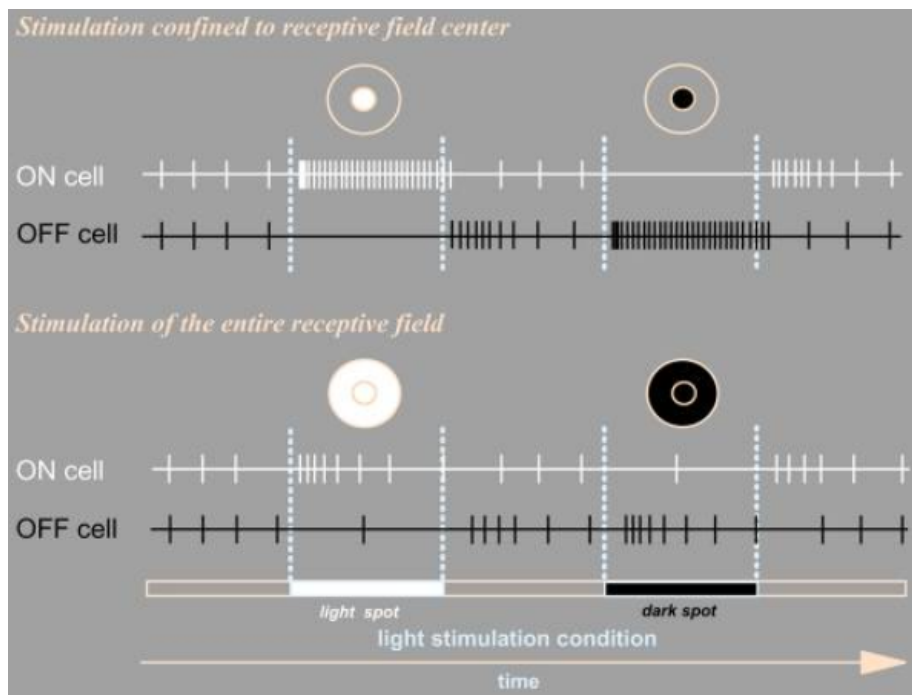


Το υποδεκτικό πεδίο των γαγγλιακών κυττάρων μπορεί εναλλακτικά να ορισθεί ως η περιοχή στο χώρο στην οποία συγκεντρώνονται οι φωτοϋποδοχείς με τους οποίους, έμμεσα, επικοινωνεί. Το υποδεκτικό πεδίο τους είναι μεγάλο εφόσον επικοινωνούν με αρκετά δίπολα κύτταρα και έχει σχήμα δίσκου. Το 1953 ο S. Kuffler [10] ανακάλυψε την οργάνωση «κέντρου-περιφέρειας» που πλέον έχει επιβεβαιωθεί πως χαρακτηρίζει τα γαγγλιακά κύτταρα στην πλειοψηφία τους. Μέσα από προβολή φωτεινών σημάτων στη γάτα, παρατήρησε πως το πεδίο αποτελείται από δύο ομόκεντρους δίσκους, το κέντρο (excitatory region) και την ανασταλτική περιοχή (inhibitory region) που το περιβάλλει. Οι δύο αυτές περιοχές παρουσιάζουν ανταγωνισμό μεταξύ τους, δηλαδή δεν διεγείρονται με τον ίδιο τρόπο. Αυτό σημαίνει πως ένα ON κύτταρο θα έχει μια OFF περιφέρεια και ένα OFF κύτταρο θα έχει μια ON περιφέρεια. Στην πρώτη περίπτωση για να αυξηθεί η συχνότητα πυροδότησης πρέπει να αυξηθεί η ένταση της φωτεινότητας (σε σχέση με μια μέση τιμή στην οποία προσαρμόζεται δυναμικά το σύστημα) αυστηρά και μόνο στο κέντρο. Αν παρουσιαστεί αυξημένη φωτεινότητα μόνο στην περιφέρεια ο νευρώνας θα κατασταλεί. Στη δεύτερη περίπτωση υπάρχει πλήρης αντιστοιχία, για να αυξηθεί η συχνότητα πυροδότησης πρέπει να αυξηθεί η φωτεινότητα αποκλειστικά στην περιφέρεια. Αν παρουσιαστεί αυξημένη φωτεινότητα μόνο στο κέντρο ο νευρώνας θα κατασταλεί. Και στις δύο περιπτώσεις, όταν δεν προβάλλεται διαφορετικό ερέθισμα στο κέντρο και την περιφέρεια το κύτταρο πυροδοτεί σε ένα στοιχειώδη ρυθμό, απόρροια της αλληλοεξουδετέρωσης των δύο περιοχών. Ο νευρώνας λοιπόν αποκρίνεται βέλτιστα σε διέγερση περιορισμένη στο κέντρο του υποδεκτικού πεδίου του. Το φαινόμενο αυτό ονομάζεται πλευρική αναστολή (lateral inhibition) και οφείλεται κυρίως στα οριζόντια και τα βραχύινα κύτταρα. Όπως αναφέρθηκε παραπάνω, παρόμοια δομή έχουν τα δίπολα κύτταρα αλλά και οι νευρώνες στον LGN. Στον V1 διατηρείται ο μηχανισμός «κέντρου-περιφέρειας», όχι όμως σε ομόκεντρους δίσκους αλλά σε επιμηκυμένες ωοειδείς επιφάνειες η μία δίπλα στην άλλη, οργάνωση που μαρτυρά ευαισθησία σε προσανατολισμό ή σε κίνηση προς συγκεκριμένη κατεύθυνση [11].



**Σχήμα 5:** Γραφική αναπαράσταση του υποδεκτικού πεδίου ενός κωνίου και ενός ON γαγγλιακού κυττάρου [2].

Εξίσου σημαντική με την ανακάλυψη της συγκεκριμένης οργάνωσης είναι και η διαπίστωση του λόγου ύπαρξης της. Ο ανταγωνισμός των δύο περιοχών επιτρέπει στους νευρώνες να εντοπίζουν διαφορές στη φωτεινότητα μεταξύ σημείων, δηλαδή μεταβολές στην αντίθεση, λειτουργία που παραπέμπει σε ανίχνευση ακμών σε εικόνες [12]. Ο μηχανισμός αυτός είναι ο πυρήνας της αποτελεσματικής κωδικοποίησης του αμφιβληστροειδούς και της μείωσης της χωρικής συσχέτισης στην πληροφορία. Τα δυναμικά ενεργείας είναι μεταβολικά ακριβά και ο αμφιβληστροειδής, μέσα στον καταγισμό οπτικών ερεθισμάτων που λαμβάνει, έχει βρει έναν αποδοτικό τρόπο να μεταβιβάζει στον εγκέφαλο μόνο την απαραίτητη πληροφορία, η οποία βρίσκεται στις μεταβολές. Εδώ κρύβεται και ένα μέρος της απάντησης στο πώς αυτό το πολύπλοκο νευρωνικό κύκλωμα πηγαίνει από τους εκατό εκατομμύρια φωτοϋποδοχείς στα ένα εκατομμύριο γαγγλιακά κύτταρα. Να σημειωθεί πως έχει βρεθεί πως στην αποδοτική κωδικοποίηση πληροφορίας συμβάλλει, πλην του υποδεκτικού πεδίου, και η μη γραμμική επεξεργασία που πραγματοποιούν πολλά γαγγλιακά κύτταρα [13]. Σε αυτό θα αναφερθούμε παρακάτω.

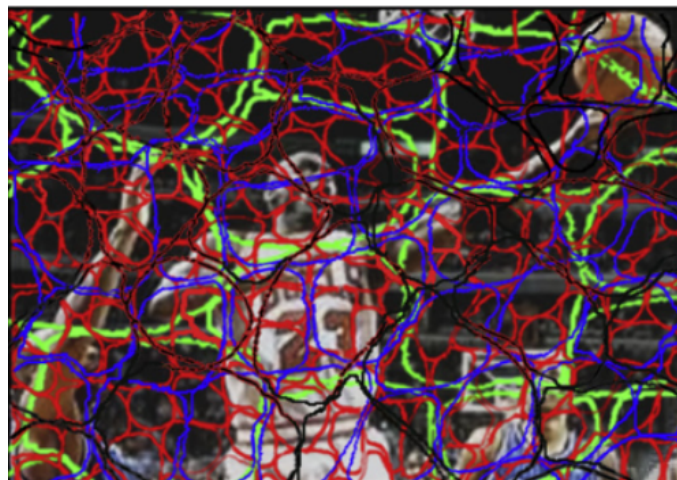


**Σχήμα 6:** Επίδραση του ανταγωνισμού μεταξύ κέντρου και περιφέρειας για ON και OFF γαγγλιακά κύτταρα. Πάνω διεγείρεται μόνο το κέντρο ή μόνο η περιφέρεια ενώ κάτω και οι δύο περιοχές διεγείρονται με τον ίδιο τρόπο. Οι κάθετες γραμμές είναι τα δυναμικά ενεργείας [12].

Το υποδεκτικό πεδίο δεν είναι μια στατική και αμετάβλητη οντότητα. Εξαρτάται από τα οπτικά ερεθίσματα και προσαρμόζεται διαρκώς στις ιδιαιτερότητες τους. Τα γαγγλιακά κύτταρα αρέσκονται στο να παρατηρούν αλλαγές στο χώρο αλλά και στο χρόνο και να ρυθμίζουν δυναμικά την απόκρισή τους. Η προσαρμογή του οπτικού συστήματος

πραγματοποιείται στις διακυμάνσεις της έντασης από μια μέση τιμή, δηλαδή στις αντιθέσεις και μπορεί να συμβαίνει άμεσα, σε χρόνο μικρότερο από 0.1 s ή πιο αργά, σε βάθος μερικών δευτερολέπτων. Κλασικές ενέργειες ταχείας προσαρμογής είναι η αλλαγή στην ευαισθησία ή την ταχύτητα απόκρισης. Σε περιβάλλον υψηλής αντίθεσης τα γαγγλιακά κύτταρα μειώνουν την ευαισθησία τους και αποκρίνονται ταχύτερα μέσω, μεταξύ άλλων, της ανατροφοδότησης καταστολής που αναφέραμε πώς διενεργούν τα οριζόντια κύτταρα και της λειτουργίας των κωνίων που έχουν χαμηλότερο χρόνο απόκρισης από τα ραβδία. Αυτό είναι αναγκαίο διότι οι αντιθέσεις των εικόνων στην καθημερινότητα μπορούν να έχουν εύρος τιμών έως δέκα τάξεις μεγέθους ενώ τα γαγγλιακά κύτταρα μπορούν να πυροδοτήσουν με μέγιστο ρυθμό περίπου 400 Hz [12], [14].

Η δομική αρχιτεκτονική του αμφιβληστροειδούς διατηρείται ίδια και με μικρές διαφορές μεταξύ των ειδών. Αυτό είναι πολύ σημαντικό εύρημα διότι τα πειράματα και οι μελέτες πραγματοποιούνται σε είδη όπως ποντίκια, σαλαμάνδρες και πρωτεύοντα αλλά με κύριο στόχο την κατανόηση των μηχανισμών της όρασης στον άνθρωπο. Μέχρι σήμερα έχουν βρεθεί πάνω από 20 διαφορετικοί τύποι γαγγλιακών κυττάρων στον αμφιβληστροειδή των θηλαστικών, καθένας με ξεχωριστή ανατομική δομή, φυσιολογία και λειτουργία [3]. Στο ποντίκι μόνο υπάρχουν τουλάχιστον 32 τύποι γαγγλιακών κυττάρων [15]. Νέα κύτταρα, άγνωστα στους ερευνητές, τα οποία δεν εντάσσονται στις ήδη καθιερωμένες κατηγορίες ανακαλύπτονται ακόμα. Τα γαγγλιακά κύτταρα κάθε πληθυσμού εκτείνονται σε όλο το οπτικό πεδίο ώστε τα υποδεκτικά τους πεδία να καλύπτουν όλη την εικόνα, σχηματίζοντας μια διάταξη «μωσαϊκού». Κύτταρα ίδιου τύπου αποφεύγουν το ένα το άλλο ενώ αδιαφορούν για κύτταρα άλλων τύπων, διαμορφώνοντας έτσι παράλληλα κανάλια που εξάγουν και μεταφέρουν διαφορετικά χαρακτηριστικά της εικόνας [16]. Η ποικιλομορφία αυτή ωστόσο προκαλεί ένα κρίσιμο ερώτημα που απασχολεί τη νευροφυσιολογία του οφθαλμού εδώ και δεκαετίες: Ποιος είναι τελικά ο ρόλος του αμφιβληστροειδούς, τι είδους επεξεργασία διενεργεί και τελικά σε τι μορφή στέλνει την πληροφορία προς τον εγκέφαλο [17], [18], [19];



**Σχήμα 7:** Κάθε τύπος γαγγλιακών κυττάρων λειτουργεί ανεξάρτητα και καλύπτει όλη την επιφάνεια του οπτικού πεδίου. Οι διαφορετικοί τύποι αναπαρίστανται από διαφορετικά χρώματα [6].



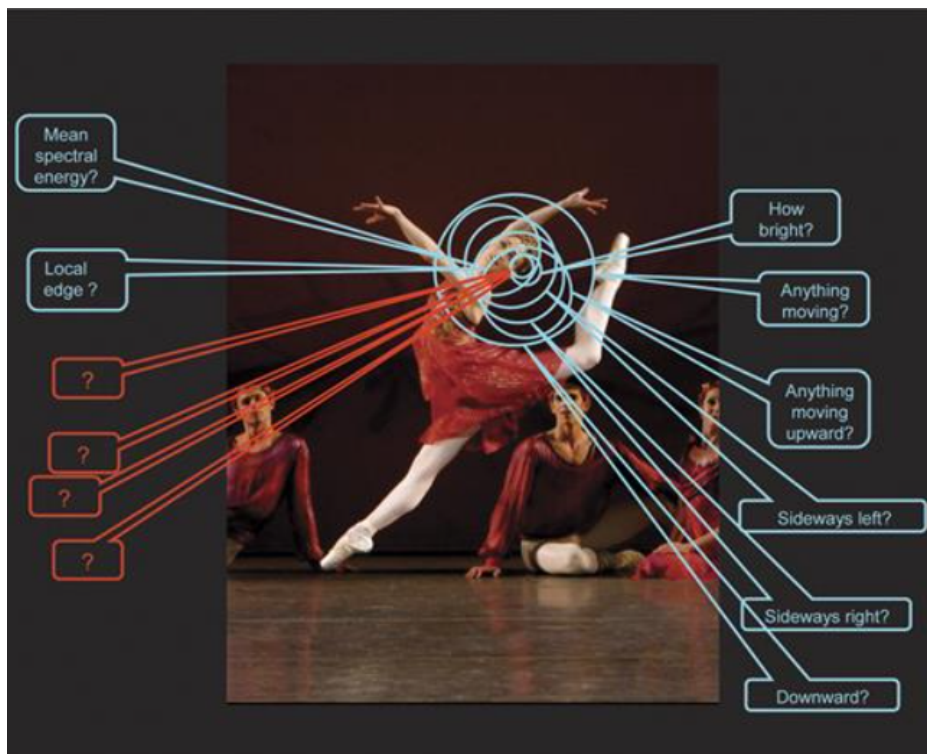
Η άποψη που κυριαρχούσε αρχικά είναι ότι ο αμφιβληστροειδής λειτουργεί αντίστοιχα με τις φωτοδιόδους σε μια ψηφιακή κάμερα. Δεν εξάγει καθόλου χαρακτηριστικά από την εικόνα και έχει μοναδικό στόχο να τη μεταβιβάσει προς τα ανώτερα επίπεδα, αφήνοντας σε αυτά όλη την επεξεργασία. Υιοθετείται ακόμα, έστω και σιωπηρά από ερευνητές που ασχολούνται με τον εγκέφαλο. Υπό αυτό το πρίσμα, ο νευρώνας εκτελεί αποκλειστικά έναν γραμμικό υπολογισμό, μετρώντας την ένταση της φωτεινότητας μέσα στο υποδεκτικό του πεδίο, που λειτουργεί ως ένας πίνακας βαρών. Έπειτα τροποποιεί την απόκρισή του αναλόγως, και τη στέλνει στον εγκέφαλο. Η περιττή πληροφορία που συναντάται στις φυσικές εικόνες απομακρύνεται αποκλειστικά χάρη στην πλευρική καταστολή. Συνεπώς, προκειμένου να κατανοηθεί πλήρως η σχέση διέγερσης - απόκρισης ενός νευρώνα, το καθοριστικό χαρακτηριστικό είναι το υποδεκτικό πεδίο. Η λειτουργία αυτή μπορεί εύκολα να προσομοιωθεί με ένα γραμμικό κωδικοποιητή (quasi-linear pixel encoder) όπως είναι το Γραμμικό - Μη Γραμμικό μοντέλο (Linear Non Linear Models - LN), το οποίο είναι ένα γραμμικό φίλτρο που ακολουθείται από μια στατική μη γραμμική συνάρτηση [20]. Μαθηματικά, διατυπώνεται από τον τύπο

$$r(t) = N(\int \int s(x, t')F(x, t - t')dt'dx), \quad (1)$$

όπου  $s(x, t)$  είναι η ένταση της εικόνας στο σημείο  $x$  τη χρονική στιγμή  $t$  και  $F(x, t)$  είναι το υποδεκτικό πεδίο του νευρώνα στο χώρο και το χρόνο. Το υποδεκτικό πεδίο λειτουργεί σε μια συνάρτηση βάρους η οποία εφαρμόζεται στη συνάρτηση έντασης μέσω συνέλιξης και αποτυπώνει την επιλεκτικότητα του κυττάρου.  $N()$  είναι μια μη γραμμική συνάρτηση που σχετίζεται με την προσαρμογή του κυττάρου σε ακραίες τιμές εισόδου και ενσωματώνει τις βιολογικές ιδιαιτερότητες που ελέγχουν το πώς πυροδοτεί. Δεν αλλάζει το «τι» αναφέρει ο νευρώνας σχετικά με την είσοδο, αλλά το «πώς» το αναφέρει. Τόσο το  $F()$  όσο και η  $N()$  υπολογίζονται εύκολα μέσα από τις αποκρίσεις. Αν το γαγγλιακό κύτταρο ακολουθεί το LN μοντέλο σε όλα τα οπτικά ερεθίσματα τότε είναι απλά ένας γραμμικός κωδικοποιητής. Κύτταρα που προσεγγίζουν αυτή τη λειτουργία υπάρχουν, για παράδειγμα, στον άνθρωπο και ονομάζονται P κύτταρα. Μάλιστα απαρτίζουν το 95% του πληθυσμού στο βοθρίο και το 50% του συνολικού πληθυσμού [17], [18], [19].

Αν αποδεχθούμε εξ ολοκλήρου την παραπάνω θεωρία, προκύπτει η αδυναμία αιτιολόγησης της ύπαρξης 55 διαφορετικών τύπων κυττάρων για μια λειτουργία τόσο απλή. Εναλλακτικά λοιπόν, η πλέον υπερισχύουσα άποψη είναι πως ο αμφιβληστροειδής δεν αφήνει όλη την επεξεργασία στον εγκέφαλο. Αντιθέτως, μέσω των διαφορετικών τύπων κυττάρων εξάγει συγκεκριμένα και σύνθετα χαρακτηριστικά τα οποία προωθεί μέσα από παράλληλα νευρικά κανάλια. Τα κύτταρα αυτά επιτελούν μη γραμμικούς και εξειδικευμένους υπολογισμούς ενώ απορρίπτουν ένα μεγάλο μέρος της οπτικής σκηνής διότι είναι συντονισμένα σε πολύ συγκεκριμένα ερεθίσματα. Το αποτέλεσμα είναι να παραδίδουν στον εγκέφαλο μια έντονα επεξεργασμένη εκδοχή της εικόνας πάνω στην οποία θα μπορεί να ενεργήσει πιο άμεσα. Η αναγνώριση υφής και η ευαισθησία σε κίνηση ενός αντικειμένου ή ακόμη και σε συγκεκριμένη κατεύθυνση κίνησης είναι τέτοιες λειτουργίες. Ανάμεσα στα

γαγγλιακά κύτταρα που ανακαλύφθηκαν ή μελετήθηκαν καλύτερα τα τελευταία χρόνια, τα περισσότερα προσεγγίζουν αυτή τη λειτουργία [17], [18], [19]. Στο ποντίκι το οποίο μελετάμε, ο πιο πολυάριθμος τύπος νευρώνων είναι τα W3 γαγγλιακά κύτταρα. Επιδεικνύουν μη γραμμική ON-OFF συμπεριφορά με πολύ ισχυρό μηχανισμό «κέντρου-περιφέρειας» και μικρό υποδεκτικό πεδίο. Αυτό τα καθιστά εξαιρετικά επιλεκτικά στο πότε θα ενεργοποιηθούν. Πιο συγκεκριμένα, εντοπίζουν αποκλειστικά κινήσεις που συμβαίνουν στο βάθος και σε μακρινή απόσταση, διευκολύνοντας το ζώο να εντοπίσει έγκαιρα εναέρια αρπακτικά. Γενικότερα στο ποντίκι εντοπίζονται πολλοί ON-OFF τύποι κυττάρων που είναι ευαίσθητοι σε συγκεκριμένα χαρακτηριστικά και δε συμμετέχουν καθόλου στην κωδικοποίηση γενικών οπτικών σκηνών. Παρόμοιοι τύποι υπάρχουν και σε άλλα είδη όπως η γάτα ή το κουνέλι [16], [21].



**Σχήμα 8:** Σε κάθε σημείο της εικόνας οι διαφορετικοί τύποι γαγγλιακών κυττάρων ανιχνεύουν ένα πλήθος χαρακτηριστικών τα οποία και αναφέρουν στον εγκέφαλο μέσα από παράλληλα κανάλια επεξεργασίας. Πολλοί τύποι κυττάρων δε γνωρίζουμε ακόμη τι ακριβώς κάνουν [4].

Συμπερασματικά, όλες οι ενδείξεις τείνουν να μας απομακρύνουν από το δόγμα πως ο αμφιβληστροειδής δεν πραγματοποιεί πολύπλοκους μη γραμμικούς μετασχηματισμούς και δεν αποδομεί τις εικόνες του εξωτερικού περιβάλλοντος μέσα από εξαγωγή χαρακτηριστικών και αναπαραστάσεων. Ακόμα και στον ανθρώπινο αμφιβληστροειδή έχουν βρεθεί νευρώνες που χαρακτηρίζονται από μη γραμμική και ιδιαίτερη συμπεριφορά, γεγονός που ενισχύει τον

ισχυρισμό πως όλα τα είδη περιέχουν γραμμικά και μη γραμμικά κύτταρα. Το ζήτημα είναι τι ποσοστό καταλαμβάνουν. Μια υπόθεση είναι πως σε είδη που οι υπολογιστικοί πόροι του εγκεφάλου είναι περιορισμένοι και για λόγους επιβίωσης, όπως ποντίκια και σαλαμάνδρες, ο αμφιβληστροειδής δίνει προτεραιότητα στην εξαγωγή σύνθετων χαρακτηριστικών, ενώ σε είδη που διαθέτουν ισχυρό ανώτερο οπτικό σύστημα ο αμφιβληστροειδής επικεντρώνεται στη γραμμική και αποδοτική μετάδοση της πληροφορίας και δεν προβαίνει σε κάποιου είδους επεξεργασία. Το σενάριο αυτό έχει επιβεβαιωθεί και με προσομοίωση του κυκλώματος «retina - V1» μέσω νευρωνικών δικτύων βαθιάς μάθησης [19]. Εντούτοις, χρειάζονται περισσότερα δεδομένα, σχετιζόμενα κυρίως με φυσικές εικόνες, καθώς δεν υπάρχει ακόμα μια καθολική θεωρία κωδικοποίησης που να εξηγεί ολοκληρωμένα πώς τελικά ο αμφιβληστροειδής συνεισφέρει στην οπτική αντίληψη [17], [18], [19].

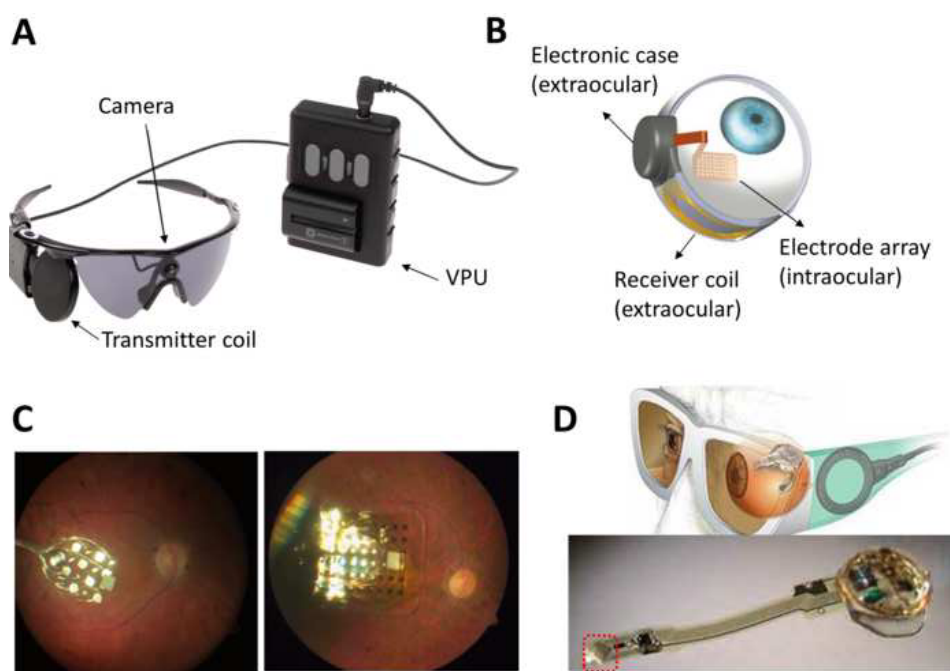
### 2.3 Προσθετική Αμφιβληστροειδούς

Σε έναν υγιή αμφιβληστροειδή, το φως ανιχνεύεται από τους φωτοϋποδοχείς και μέσω μιας αλληλουχίας νευροχημικών διεργασιών φτάνει στο τελευταίο στάδιο, τα γαγγλιακά κύτταρα. Από εκεί, και μέσω του οπτικού νεύρου, η πληροφορία μεταβιβάζεται στα οπτικά κέντρα του εγκεφάλου, σε μορφή δυναμικών ενεργείας. Πολλές ασθένειες μπορούν να διαταράξουν αυτή τη διαδικασία και να οδηγήσουν σε μερική ή ολική απώλεια όρασης. Σε περιπτώσεις βλάβης του οπτικού νεύρου η κατάσταση είναι προς το παρόν μη αναστρέψιμη. Ασθένειες όπως η μελαγχρωστική αμφιβληστροειδοπάθεια (Retinitis Pigmentosa) και η εκφύλιση της ωχράς κηλίδας (Age-related Macular Degeneration) προσβάλλουν τους φωτοϋποδοχείς, αφήνοντας ωστόσο άθικτα τα υπόλοιπα επίπεδα, όπως τα δίπολα και τα γαγγλιακά κύτταρα. Σε αυτό βασίζεται και η τεχνολογία της προσθετικής αμφιβληστροειδούς, η οποία προσπαθεί να παρακάμψει τον ζημιωμένο ιστό και να ενεργοποιήσει με μεθόδους ηλεκτρικής διέγερσης όπως μικροηλεκτρόδια τα γαγγλιακά κύτταρα. Ισχυρό κίνητρο που, τις τελευταίες δεκαετίες, οδήγησε τους ερευνητές στο να εστιάσουν στη συγκεκριμένη τεχνολογία είναι η διαπίστωση πως, ακόμα και σε χρόνια εκφυλισμένους αμφιβληστροειδείς, η ηλεκτρική διέγερση των νευρώνων και η επιτυχής ενεργοποίηση του οπτικού νεύρου μπορούν να οδηγήσουν τους πάσχοντες στο να αντιληφθούν ίχνη φωτός και εικόνας στο οπτικό πεδίο τους [22], [23].



**Σχήμα 9:** Επίδραση των εκφυλιστικών ασθενειών στην αίσθηση της όρασης. Αριστερά: Φυσιολογικό οπτικό πεδίο. Κέντρο: Απώλεια κεντρικής όρασης από εκφύλιση ωχράς κηλίδας. Δεξιά: Απώλεια περιφερειακής όρασης (tunnel vision) λόγω μελαγχρωστικής αμφιβληστροειδοπάθειας [23].

Η μελαγχρωστική αμφιβληστροειδοπάθεια ξεκινά με σταδιακή εκφύλιση των ραβδίων στην περιφέρεια του αμφιβληστροειδούς, προκαλώντας απώλεια νυχτερινής και περιφερειακής όρασης. Αν η βλάβη προχωρήσει και σε αλλοίωση στα κωνία, μπορεί να προκληθεί ολική τύφλωση. Εκτιμάται πως προσβάλλει 1.5 εκατομμύρια ανθρώπους παγκοσμίως. Κατά την εκφύλιση της ωχράς κηλίδας, το πρόβλημα εντοπίζεται κυρίως στα κωνία στην ωχρά κηλίδα, προκαλώντας θολή κεντρική όραση. Όσο η ασθένεια προχωράει, η θολή περιοχή μεγαλώνει και τα τυφλά σημεία στην κεντρική όραση πληθαίνουν. Εκτιμάται πως 30 - 50 εκατομμύρια άνθρωποι προσβάλλονται από εκφύλιση της ωχράς κηλίδας. Αθροιστικά οι δύο ασθένειες συνεισφέρουν έντονα στις συνηθέστερες αιτίες απώλειας όρασης. Όπως αναφέραμε, και στις δύο περιπτώσεις τα επίπεδα που έπονται των φωτοϋποδοχέων διατηρούν την φυσιολογική λειτουργία τους [23].

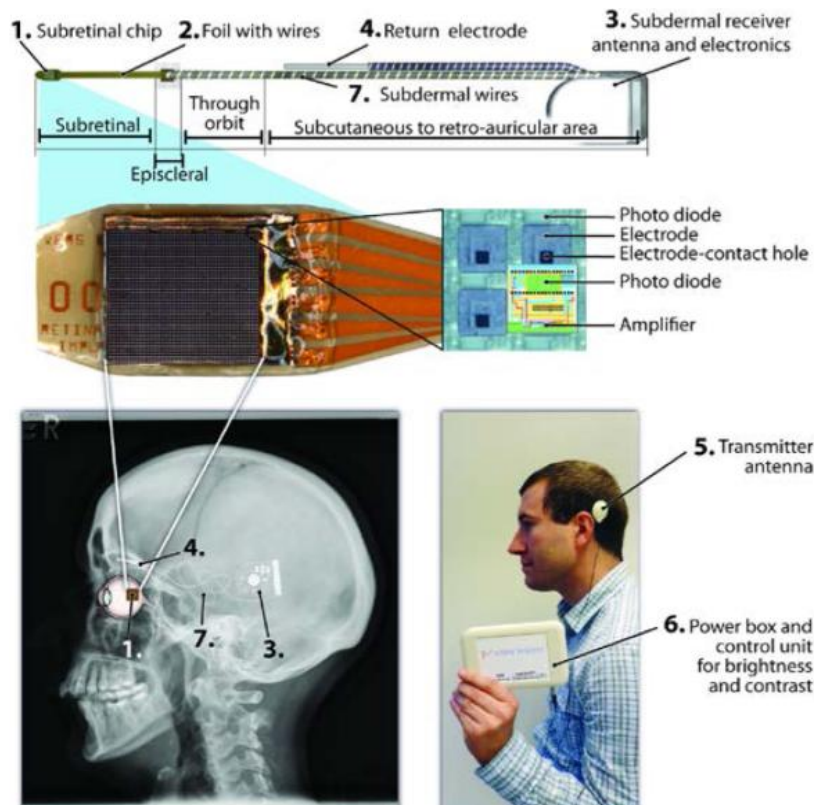


**Σχήμα 10:** Επιαμφιβληστροειδικό εμφύτευμα [23].

Ένα εμφύτευμα αμφιβληστροειδούς λειτουργεί σαν ένα ολοκληρωμένο σύστημα αποτελούμενο από συγκεκριμένα στοιχεία. Τυπικά, αυτά είναι μια συσκευή απόκτησης των δεδομένων όπως μια κάμερα, η κεντρική μονάδα επεξεργασίας που θα προσομοιώσει τα κύτταρα που δυσλειτουργούν αναλαμβάνοντας την επεξεργασία των εικόνων στην είσοδο και τη δημιουργία ηλεκτρικών παλμών και μια διάταξη ενεργοποίησης, δηλαδή ένα πλέγμα μικροηλεκτροδίων.

Η κάμερα μπορεί να είναι εξωτερική, προσαρμοσμένη για παράδειγμα σε γυαλιά, ή εσωτερική, δηλαδή εμφυτευμένη μέσα στον οφθαλμό. Η δεύτερη περίπτωση έχει το πλεονέκτημα ότι ο χρήστης, ελέγχοντας το οπτικό πεδίο της κάμερας μέσω κινήσεων των βολβών των ματιών και όχι του κεφαλιού, εκτελεί τις καθημερινές λειτουργίες με πιο φυσικό τρόπο. Ο συνδυασμός και των δύο, δηλαδή μια εξωτερική κάμερα που ανιχνεύει σε πραγματικό χρόνο την κατεύθυνση του βλέμματος μπορεί να δώσει θετικά αποτελέσματα.

Στη μονάδα επεξεργασίας, οι μελέτες εστιάζονται κυριότερα σε δύο σημεία. Χαρακτηριστικά όπως η μορφή, το πλάτος και η συχνότητα του σήματος που οδηγείται στα μικροηλεκτρόδια έχει αποδειχθεί πως επηρεάζουν σημαντικά την αποτελεσματικότητα και την ασφάλεια της διαδικασίας και δυσκολεύουν την επιλεκτική και αποδοτική διέγερση συγκεκριμένων τύπων γαγγλιακών κυττάρων. Επιπλέον, τεχνικές βελτίωσης της ποιότητας της εικόνας και αλγόριθμοι όρασης υπολογιστών και μηχανικής μάθησης για εξαγωγή σημαντικών σημείων της εικόνας ανάλογα το έργο που επιτελείται, με ταυτόχρονη απόρριψη των υπολοίπων, θεωρείται ότι θα αναβαθμίσουν τις προσπάθειες που γίνονται στον τομέα.

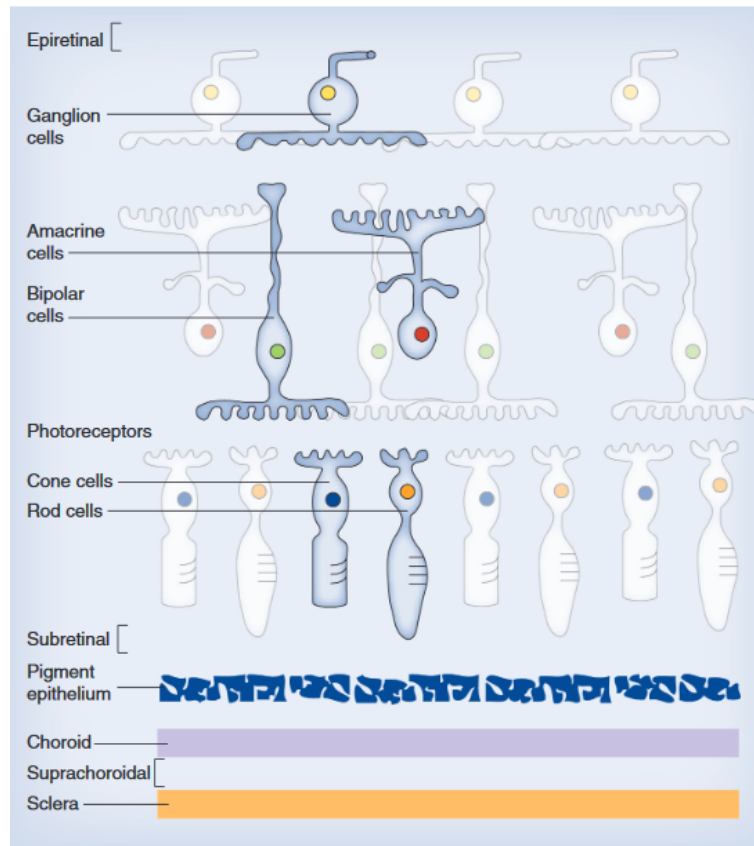


Σχήμα 11: Υποαμφιβληστροειδικό εμφύτευμα [22].

Τα μικροηλεκτρόδια, καθοδηγούμενα από τα σήματα που λαμβάνουν από τη συσκευή επεξεργασίας, διεγείρουν τα κύτταρα ενδιαφέροντος προκαλώντας τους αποπόλωση, δηλαδή μετατόπιση του δυναμικού μεμβράνης προς πιο θετικές τιμές, μέσω ηλεκτρικού ρεύματος. Όταν η αποπόλωση ξεπεράσει μια τιμή κατωφλίου, το κύτταρο θα στείλει ένα δυναμικό ενεργείας. Το υπόστρωμα της διάταξης ενεργοποίησης αποτελεί κρίσιμο παράγοντα καθώς αποτελεί το τμήμα του συστήματος που έρχεται σε άμεση επαφή με τον αμφιβληστροειδή. Οφείλει να είναι εύκαμπτο προκειμένου να μην προκαλέσει ζημιά στο βιολογικό ιστό και βιοσυμβατό ώστε το εμφύτευμα να μην απορριφθεί [22], [23].

Όπως φαίνεται και στην Εικόνα 12, τα εμφυτεύματα αμφιβληστροειδούς διακρίνονται σε τρεις κατηγορίες ανάλογα με το σημείο στο οποίο εμφυτεύονται τα ηλεκτρόδια και, παρά τις όποιες ομοιότητες παρουσιάζουν στην αρχιτεκτονική τους, διαφέρουν στις λεπτομέρειες σχεδίασης τους ανάλογα την κατηγορία στην οποία εντάσσονται. Πρώτον, έχουμε τα επιαμφιβληστροειδικά (epiretinal) εμφυτεύματα που τοποθετούνται πάνω στο πιο εσωτερικό σημείο του αμφιβληστροειδούς, γειτονικά με τα γαγγλιακά κύτταρα, και για αυτό το λόγο μπορούν να τα διεγείρουν απευθείας. Η πλειοψηφία αυτών των συστημάτων χρησιμοποιεί εξωτερική κάμερα και ασύρματη τροφοδοσία και μετάδοση δεδομένων. Το Argus II, η μία από τις μόλις δύο συσκευές που, μέχρι σήμερα, έχει εγκριθεί για εμπορική χρήση και χρησιμοποιείται από άτομα με παθήσεις του αμφιβληστροειδούς, εμπίπτει στη συγκεκριμένη κατηγορία. Έπειτα, έχουμε τα υποαμφιβληστροειδικά (subretinal) εμφυτεύματα που τοποθετούνται κάτω από τον αμφιβληστροειδή στο σημείο που βρίσκονται οι εκφυλισμένοι φωτοϋποδοχείς. Έχουν το πλεονέκτημα ότι μειώνουν τις παραμορφώσεις στην οπτική αντίληψη αφού δεν διεγείρουν απευθείας τους άξονες στο τελευταίο επίπεδο του αμφιβληστροειδούς αλλά σε κάποιο προηγούμενο όπως τα δίπολα κύτταρα, προσπαθώντας να εκμεταλλευτούν όση φυσιολογική λειτουργία έχει απομείνει στα κύτταρα αυτά, με την αποτελεσματικότητά τους να εξαρτάται από την πρόοδο της ασθένειας. Η δεύτερη συσκευή που έχει λάβει έγκριση για χρήση, το Alpha-IMS, ανήκει σε αυτή την κατηγορία. Τέλος, υπάρχουν τα διαχωριοειδικά εμφυτεύματα, όπου τα ηλεκτρόδια τοποθετούνται σχετικά μακριά από τον αμφιβληστροειδή, ανάμεσα στον σκληρό χιτώνα και στο χοριοειδές. Αυτό από τη μία μειώνει το ρίσκο πρόκλησης βλάβης στον αμφιβληστροειδή από την επέμβαση ή το εμφύτευμα, από την άλλη ωστόσο υποχρεώνει το ηλεκτρικό σήμα να ταξιδέψει μεγαλύτερη απόσταση για να φτάσει από το ηλεκτρόδιο στο νευρικό ιστό, με κίνδυνο την παραμόρφωσή του και τη μείωση της χωρικής ανάλυσης στην εικόνα [22], [23], [24].





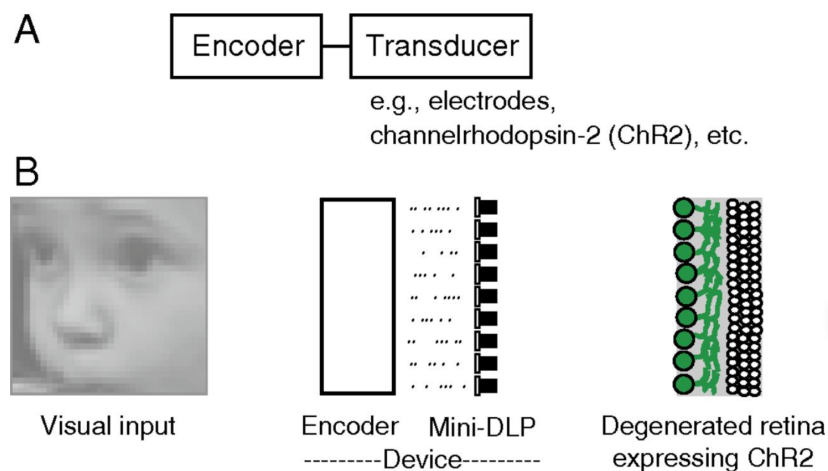
**Σχήμα 12:** Τα σημεία τοποθέτησης των εμφυτευμάτων για τις τρεις κατηγορίες [24].

Παρά τον ενθουσιασμό που εξαρχής δημιούργησε η προσθετική αμφιβληστροειδούς, τα αποτελέσματα δεν είναι ακόμα σε καμία περίπτωση τα επιθυμητά. Η πλειοψηφία των συσκευών προσφέρει πολύ περιορισμένη όραση και δεν έχει επιτευχθεί κάτι καλύτερο από αναγνώριση φωτεινών σημείων, απλών σχημάτων και ακμών σε υψηλή αντίθεση. Η αλήθεια είναι πως και σε αυτό το στάδιο μπορούν να προσφέρουν μια ευκαιρία για πιο φυσιολογική ζωή σε άτομα σε προχωρημένο στάδιο εκφύλισης του αμφιβληστροειδούς, η οπτική αντίληψη που αισθάνονται ωστόσο απέχει πάρα πολύ από αυτή ενός υγιούς ανθρώπου. Οι ερευνητές προσπαθούν ακόμα να κατανοήσουν τι είναι στην πραγματικότητα αυτό που βλέπει ο χρήστης εμφυτεύματος και το βαθμό στον οποίο ο εγκέφαλος προσαρμόζεται σε δεδομένα από μια τέτοια συσκευή [22], [24], [25].

Ένα πολύ σοβαρό πρόβλημα αφορά τον τρόπο που το σύστημα θα επεξεργαστεί τα δεδομένα προκειμένου να παραγάγει έξοδο όσο πιο κοντά γίνεται σε αυτή των γαγγλιακών κυττάρων. Ο αμφιβληστροειδής είναι ένα σύστημα με πολλούς τύπους κυττάρων που κωδικοποιούν με ξεχωριστό τρόπο την πληροφορία, εντοπίζουν διαφορετικά χαρακτηριστικά στις εικόνες και διαφοροποιούνται στη μορφή της απόκρισής τους. Οι αρχικές υλοποιήσεις δεν είχαν ενσωματώσει το στοιχείο αυτό και λειτουργούσαν σαν γραμμικοί κωδικοποιητές, διεγείροντας όλα τα κύτταρα με το ίδιο σήμα, ανάλογο της έντασης της φωτεινότητας στην είσοδο. Για παράδειγμα, αύξηση στη φωτεινότητα οδηγούσε σε διέγερση τα ON κύτταρα αλλά και τα OFF κύτταρα. Προφανώς το αποτέλεσμα ήταν ο εγκέφαλος να αδυνατεί να

ερμηνεύσει τα σήματα που λαμβάνει [23], [26]. Προκύπτει επομένως το συμπέρασμα ότι τα εμφυτεύματα δεν έχουν επιτύχει ακόμα το στόχο τους όχι μόνο λόγω της χαμηλής ανάλυσης, αλλά και του ζητήματος της κωδικοποίησης πληροφορίας.

Μια ιδέα για την επίλυση του προβλήματος είναι η επεξεργασία εικόνων και η εξαγωγή χαρακτηριστικών μέσα από μοντελοποίηση της λειτουργίας που επιτελούν οι διαφορετικοί τύποι γαγγλιακών κυττάρων με μεθόδους όρασης υπολογιστών [81]. Μια ενδιαφέρουσα, διαφορετική προσέγγιση χρησιμοποιεί έναν κωδικοποιητή και έναν μετατροπέα για να προσομοιώσει τη συμπεριφορά του αμφιβληστροειδούς, αυξάνοντας αξιοσημείωτα την απόδοση. Ο κωδικοποιητής είναι ένα υπολογιστικό μοντέλο που μαθαίνει τους μετασχηματισμούς που γίνονται στον αμφιβληστροειδή ώστε να προβλέπει τις αποκρίσεις γαγγλιακών κυττάρων σε μορφή δυναμικών ενεργείας. Τα δυναμικά ενεργείας μετατρέπονται σε φωτεινούς παλμούς μέσω ενός ειδικά διαμορφωμένου ψηφιακού προβολέα φωτός (mini DLP). Ο μετατροπέας είναι μια φωτοευαίσθητη πρωτεΐνη (ChR2) που οδηγεί τα γαγγλιακά κύτταρα στο να πυροδοτήσουν με τον τρόπο που υποδεικνύει η έξοδος του κωδικοποιητή [27]. Το μοντέλο απεικονίζεται σχηματικά στην Εικόνα 13.



**Σχήμα 13:** Σχηματική αναπαράσταση εμφυτεύματος που προσομοιώνει τη συμπεριφορά του αμφιβληστροειδούς [27].

Ιδιαίτερη προσπάθεια αφιερώνεται στη βελτίωση της χωρικής ανάλυσης στην εικόνα με αύξηση τόσο του αριθμού των ηλεκτροδίων διέγερσης όσο και της πυκνότητάς τους, όπως επίσης και στην εξάλειψη παραμορφώσεων στην εικόνα που προκύπτουν όταν το ηλεκτρόδιο δεν διεγείρει μόνο τα γειτονικά του κύτταρα. Παράγοντες στον τομέα της μηχανικής και της μικροηλεκτρονικής όπως η αυτονομία, η πρακτικότητα, η ασφάλεια, η διάρκεια ζωής, η τεχνική συσκευασίας και η επιλογή των στοιχείων του συστήματος οφείλουν να μελετηθούν και να βελτιωθούν, ενώ είναι βέβαιο πως η υποβοήθηση του συστήματος μέσω εκμάθησης από το χρήστη αξιοποιώντας αλγορίθμους όρασης υπολογιστών και μηχανικής μάθησης είναι ένα βήμα προς τη σωστή κατεύθυνση [22], [23], [26].



Πολλά εργαστήρια και εταιρείες αναζητούν τρόπους να βοηθήσουν ασθενείς με προβλήματα όρασης και, παρόλο που δεν υπάρχει η αναμενόμενη πρόοδος και οι δυσκολίες δείχνουν πολλές και αποθαρρυντικές, θεωρείται πως στο άμεσο μέλλον θα υπάρξουν περισσότερα συστήματα προσθετικής αμφιβληστροειδούς που θα περάσουν τις κλινικές δοκιμές και θα βγουν στην αγορά έτοιμα για χρήση. Με τη διαρκή πρόοδο της επιστήμης των υλικών, της επιστήμης των υπολογιστών και της μικροηλεκτρονικής, σε συνδυασμό με την βαθύτερη αντίληψή μας για τους μηχανισμούς με τους οποίους ο αμφιβληστροειδής επεξεργάζεται και κωδικοποιεί την πληροφορία, αναπόφευκτο είναι να φτάσουμε στο σημείο που τα εμφυτεύματα θα μπορούν να μιμηθούν τη φυσιολογική όραση με υψηλή ακρίβεια [23], [24].

## 2.4 Βιολογικά πειράματα Ηλεκτροφυσιολογίας

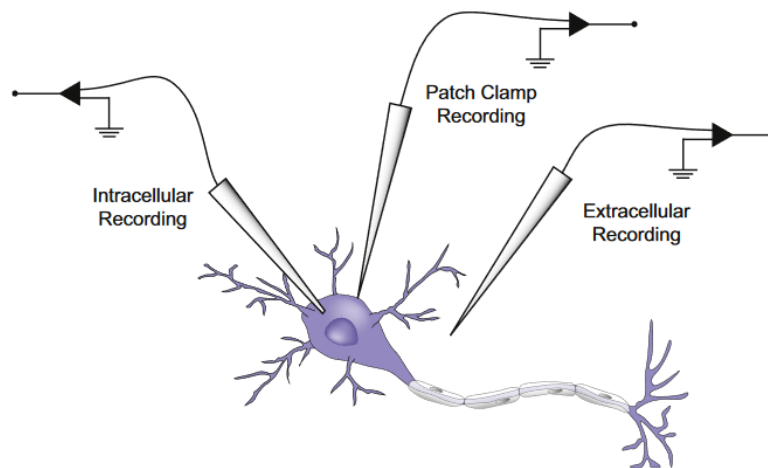
Ηλεκτροφυσιολογία είναι ο κλάδος της Νευροεπιστήμης που ερευνά τις ηλεκτρικές ιδιότητες βιολογικών κυττάρων και νευρώνων. Περιλαμβάνει καταγραφές διαφοράς δυναμικού ή ηλεκτρικής έντασης μεταξύ διαφορετικών κυττάρων ή νευρώνων σε ένα σύστημα με στόχο την κατανόηση της επικοινωνίας και των αλληλεπιδράσεων μεταξύ τους, ή μεταξύ διαφορετικών σημείων στο ίδιο κύτταρο ή νευρώνα, ώστε να μελετηθούν τα φαινόμενα που τους προσδίδουν τα φυσιολογικά τους χαρακτηριστικά [48].

Οι δύο πιο κλασικές μέθοδοι για ηλεκτροφυσιολογικές καταγραφές είναι η ενδοκυτταρική και η εξωκυτταρική. Κατά την ενδοκυτταρική μέθοδο καταγραφής, ένα πολύ μικρό ηλεκτρόδιο εισέρχεται εντός του κυττάρου και μετρά με ακρίβεια τη διαφορά δυναμικού μεταξύ του σημείου εισαγωγής και κάποιου σημείου αναφοράς στην εξωκυτταρική περιοχή. Αποτελεί ένα αξιόπιστο εργαλείο και προτιμάται από τους νευροεπιστήμονες σε πληθώρα πειραμάτων καθώς, σε αντίθεση με την εξωκυτταρική μέθοδο καταγραφής, που μπορεί να εντοπίσει ταχέα ηλεκτρικά φαινόμενα όπως δυναμικά ενεργείας, η τεχνική αυτή μπορεί και συλλαμβάνει φαινόμενα όπως το δυναμικό σε συνάψεις, γύρω από τη μεμβράνη ή στους υποδοχείς, δηλαδή βαθμωτές τάσεις [49]. Στον αμφιβληστροειδή, τέτοιου είδους αποκρίσεις συναντάμε μεταξύ άλλων στα δίπολα κύτταρα, σε αντίθεση με τα γαγγλιακά κύτταρα τα οποία παράγουν δυναμικά ενεργείας [7].

Σε εξωκυτταρικές καταγραφές ηλεκτρικής δραστηριότητας το ηλεκτρόδιο δεν τοποθετείται στο εσωτερικό του κυττάρου αλλά στο εξωκυτταρικό υγρό σε κοντινή απόσταση από το υπό μελέτη κύτταρο, μετρώντας τη διαφορά δυναμικού μεταξύ αυτού του σημείου και ενός δυναμικού αναφοράς σε ένα πιο μακρινό σημείο, πάλι εξωκυτταρικά. Αν η επιφάνεια των ηλεκτροδίων είναι αρκετά μικρή (κοντά στο  $1\mu\text{m}$ ) είναι δυνατή η λήψη του ηλεκτρικού δυναμικού ενός μόνο νευρώνα. Εδώ, το δυναμικό ενεργείας που λαμβάνεται από το ηλεκτρόδιο είναι παρόμοιο με αυτό που θα έδινε η ενδοκυτταρική καταγραφή, αλλά κάποιες τάξεις μεγέθους χαμηλότερης έντασης ( $\mu\text{V}$  έναντι  $\text{mV}$ ). Αν χρησιμοποιηθούν ηλεκτρόδια ελαφρώς μεγαλύτερης επιφάνειας, μπορούν ταυτόχρονα να καταγραφούν οι αποκρίσεις πολλαπλών γειτονικών κυττάρων. Στη δεύτερη περίπτωση είναι αναγκαίος ο διαχωρισμός των αποκρίσεων των κυττάρων μέσα από το συνολικό σήμα που λαμβάνουμε,

ο οποίος γίνεται με αλγορίθμους ταξινόμησης που χρησιμοποιούν τις κυματομορφές που δίνουν τα ηλεκτρόδια από μια μέτρηση προκειμένου να ταυτοποιήσουν ποιο ηλεκτρικό σήμα έρχεται από ποιο κύτταρο. Οι τεχνικές αυτές μπορούν επίσης να δώσουν και τον αριθμό των κυττάρων που βρίσκονται κοντά στο ηλεκτρόδιο, ενώ είναι πιο αποτελεσματικές σε περιοχές με γνωστές και επαρκώς μελετημένες κατηγορίες κυττάρων. Όταν λοιπόν το ερώτημα είναι πώς ένας νευρώνας κωδικοποιεί πληροφορίες μέσω δυναμικών ενεργείας ή απαιτείται η μελέτη ενός νευρωνικού δικτύου και όχι απομονωμένων νευρώνων, η συγκεκριμένη μέθοδος είναι αυτή που προτιμάται [48], [50].

Υπάρχει επίσης και μια τρίτη εναλλακτική μέθοδος κατά την οποία το ηλεκτρόδιο τοποθετείται πάνω στη μεμβράνη του κυττάρου (patch clamp). Εφαρμόζεται σε μελέτες που αφορούν τα κανάλια ιόντων στα κύτταρα και έχει αντικαταστήσει την ενδοκυτταρική μέθοδο σε αρκετά πειράματα λόγω της αξιοπιστίας και της ποιότητας των αποτελεσμάτων της [48].



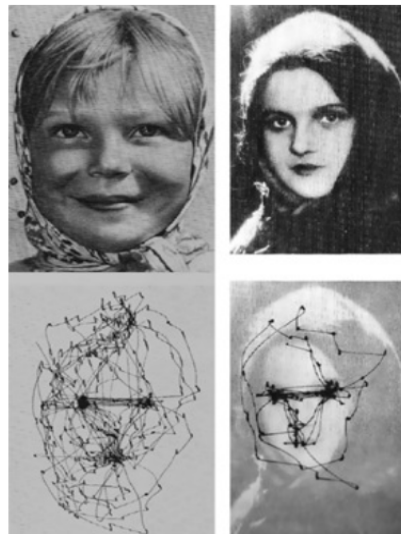
**Σχήμα 14:** Οι τρεις μέθοδοι ηλεκτροφυσιολογικών καταγραφών ανάλογα το σημείο που τοποθετείται το ηλεκτρόδιο: εξωτερικά του νευρώνα (εξωκυτταρική καταγραφή), εσωτερικά του νευρώνα (ενδοκυτταρική), στη μεμβράνη (καταγραφή patch clamp) [48].

## 2.5 Οπτική Προσοχή και Προεξοχή

Όταν ο άνθρωπος παρατηρεί μια εικόνα, δεν επεξεργάζεται κάθε σημείο της με την ίδια προσοχή και στον ίδιο βαθμό. Συγκεκριμένες περιοχές της εικόνας κεντρίζουν το ενδιαφέρον του, τις οποίες ωστόσο ο εγκέφαλος, όσο βιολογικά εξελιγμένος και υπολογιστικά αποτελεσματικός είναι, αδυνατεί να επεξεργαστεί παράλληλα. Μια λύση που έχουν υιοθετήσει τόσο οι άνθρωποι όσο και πολλά είδη ζώων είναι κάθε χρονική στιγμή οι υπολογιστικοί πόροι του οπτικού συστήματος να αφιερώνονται σε μία μοναδική περιοχή της εικόνας, αποσκοπώντας στην πιο λεπτομερή ανάλυση της. Η σειριακή αυτή μετατόπιση του ενδιαφέροντος σε διαφορετικές περιοχές μιας εικόνας ονομάζεται αλληλουχία εστιάσεων και

πραγματοποιείται μέσω των σακκαδικών κινήσεων. Σακκαδικές είναι οι γρήγορες και ταυτόχρονες κινήσεις των οφθαλμών μέσω αποστολής εντολών από τον εγκέφαλο στους οφθαλμικούς μύες, οι οποίες ανακατευθύνουν το βοθρίο που είναι υπεύθυνο για την υψηλή ανάλυση της κεντρικής όρασης σε μια συγκεκριμένη περιοχή ενδιαφέροντος. Οι σακκαδικές κινήσεις μπορεί να είναι αντανάκλαστικές, όταν υφίσταται κάποια οπτική διέγερση, εκούσιες, όταν επιδιώκεται η σάρωση του οπτικού πεδίου, προβλεπόμενες, όταν η σειρά εστίασης βασίζεται στην παρατήρηση κάποιας κίνησης ή οδηγούμενες από μνήμη, όταν τα μάτια εστιάζονται σε κάποιο σημείο χωρίς την παρουσία οπτικού ερεθίσματος. Υπάρχουν επίσης οι αντισακκαδικές κινήσεις στις οποίες συμβαίνει απομάκρυνση της εστίασης από το οπτικό ερέθισμα και οι μικροσακκαδικές κινήσεις που ορίζονται ως οι πολύ μικρές, συνεχείς και ακανόνιστες κινήσεις των οφθαλμών που πραγματοποιούνται όταν εστιάζουμε σε ένα αντικείμενο ώστε να μην εξασθενήσει η οπτική μας αντίληψη.

Κατά την φυσιολογική όραση, η διαδικασία εναλλάσσεται μεταξύ των σακκαδικών κινήσεων που ρυθμίζουν την κατεύθυνση του οπτικού άξονα και στιγμών κατά τις οποίες οι οφθαλμοί μένουν σταθεροί στη νέα τους θέση. Αποκλειστικά σε αυτά τα χρονικά διαστήματα, που είναι μεγαλύτερα σε διάρκεια, συγκεντρώνεται και όλη η χρήσιμη πληροφορία από το οπτικό σύστημα [28], [29].

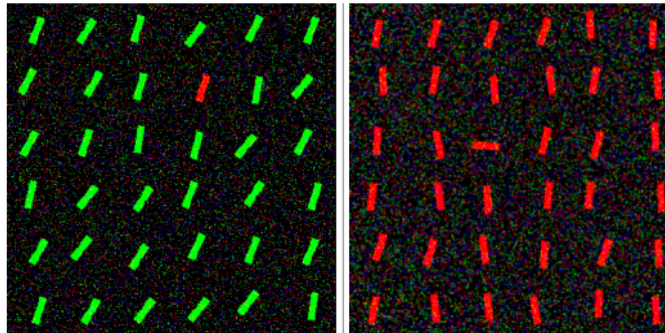


**Σχήμα 15:** Σακκαδικές κινήσεις [30].

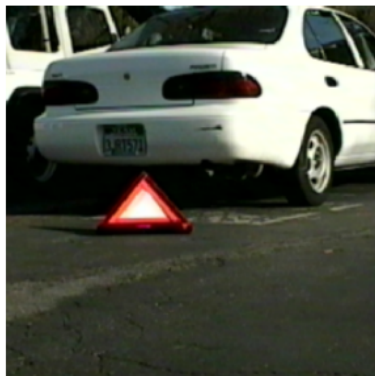
Εδώ και αρκετές δεκαετίες η επιστημονική κοινότητα στα πεδία της νευροεπιστήμης, της ψυχολογίας και της βιολογίας προσπαθεί εντατικά να δώσει απαντήσεις σε συγκεκριμένα ερωτήματα που γεννά η παραπάνω μέθοδος οπτικής δειγματοληψίας και να ερμηνεύσει τους μηχανισμούς με τους οποίους λειτουργεί. Εφόσον κάθε φορά εστιάζουμε σε ένα μεμονωμένο σημείο ή αντικείμενο, με ποιο κριτήριο γίνεται η επιλογή αυτή; Γιατί για τον εγκέφαλό μας μια περιοχή θεωρείται σημαντική ενώ άλλες απορρίπτονται; Πώς ορίζεται η σειρά με την οποία θα επεξεργαστούμε τις σημαντικές περιοχές και πόση ώρα χρειάζεται να εστιάσουμε

τους οφθαλμούς μας στην καθεμία; Η θεωρία που έχει καθιερωθεί πλέον είναι πως η οπτική προσοχή διαμορφώνεται από τον συνδυασμό δύο ομάδων χαρακτηριστικών, αυτά από την βάση προς την κορυφή και αυτά από την κορυφή προς τη βάση [29], [31].

Τα χαρακτηριστικά από τη βάση προς την κορυφή είναι ένας γρήγορος, πρωτόγονος και αυτόματος μηχανισμός που προτρέπει τον παρατηρητή να εστιάσει την προσοχή του σε σημεία της εικόνας που έχουν αυξημένο μέτρο προεξοχής. Ως προεξοχή (saliency) ορίζεται η ικανότητα ενός αντικειμένου να ξεχωρίζει από τα γειτονικά του. Δεν αποτελεί φυσική ιδιότητα ενός οπτικού ερεθίσματος αλλά υφίσταται αποκλειστικά μέσω της αλληλεπίδρασης του με άλλα ερεθίσματα. Η προσέγγιση από τη βάση προς την κορυφή υπολογίζει την προεξοχή μέσα από θεμελιώδη και χαμηλού επιπέδου χαρακτηριστικά της εικόνας όπως το χρώμα, οι αντιθέσεις, ο προσανατολισμός, οι υφές και οι συμμετρίες. Για παράδειγμα, ένα κόκκινο αντικείμενο μέσα σε ένα πράσινο φόντο ή μια οριζόντια γραμμή ανάμεσα σε ένα σύνολο από κάθετες γραμμές θα κεντρίσουν αμέσως το ενδιαφέρον μας για τον απλό λόγο ότι ξεχωρίζουν από το περιβάλλον τους. Η συγκεκριμένη προσέγγιση είναι προφανές ότι θα δώσει παρόμοια αποτελέσματα για όλους τους παρατηρητές της ίδιας διέγερσης, αφού στηρίζεται μόνο σε αντικειμενικά χαρακτηριστικά του ίδιου του ερεθίσματος.



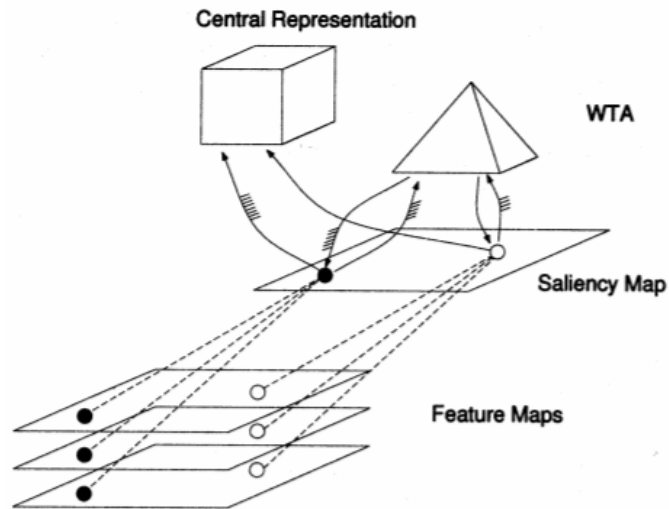
**Σχήμα 16:** Παραδείγματα περιπτώσεων που η προεξοχή ενός αντικειμένου έλκει αμέσως την προσοχή μας [29].



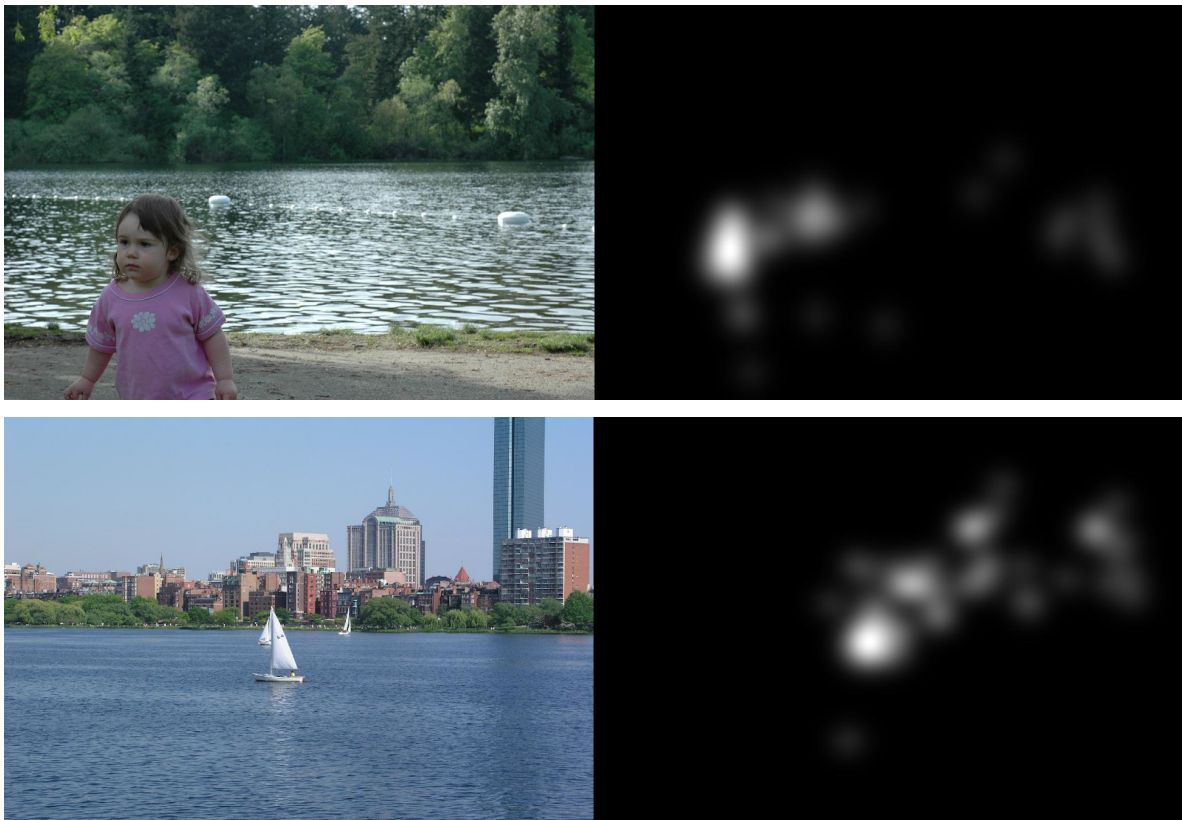
**Σχήμα 17:** Μια απλή εφαρμογή της οπτικής προεξοχής στην καθημερινότητα. Το τρίγωνο έκτακτης ανάγκης θα τραβήξει αμέσως την προσοχή των οδηγών, ακόμη και από μακρινές αποστάσεις, χάρη στη δυνατότητά του να ξεχωρίζει σε σχέση με το περιβάλλον που βρίσκεται [29].

Η επιρροή της δεύτερης συνιστώσας στην οπτική αντίληψη, τα αποκαλούμενα από την κορυφή προς τη βάση χαρακτηριστικά, αναδείχθηκε πρώτα και τεκμηριώθηκε κυρίως από τον A. Yarbus [30]. Στα πειράματά του μελέτησε τις κινήσεις των ματιών προβάλλοντας εικόνες σε παρατηρητές και ταυτόχρονα τους έκανε ορισμένες ερωτήσεις και υποδείξεις σχετικά με την οπτική σκηνή. Διαπίστωσε πως οι αλληλουχίες εστιάσεων και ο τρόπος μελέτης της εικόνας διέφεραν έντονα τόσο μεταξύ των παρατηρητών όσο και για κάθε παρατηρητή ανάλογα με την υπόδειξη που είχε λάβει. Πολλές μελέτες από τότε ενίσχυσαν την ιδέα πως την οπτική αντίληψη επηρεάζει ισχυρά και ένας πιο αργός και υπολογιστικά ακριβός μηχανισμός, δυσκολότερος στη σύλληψη και την κατανόηση, με ποικίλα κριτήρια επιλογής που ορίζεται υποκειμενικά για τον κάθε παρατηρητή οδηγούμενος από γνωστικές και συμπεριφορικές παραμέτρους. Αν δείξουμε σε κάποιον παρατηρητή ένα πλήθος από γραμμές ίδιου χρώματος με διαφορετική κλίση η καθεμία, δε θα εστιάσει πουθενά το ενδιαφέρον του αφού καμία γραμμή δεν ξεχωρίζει από τις γειτονικές της. Αν του πούμε να απομονώσει μόνο τις κάθετες γραμμές, αυτές θα αρχίσουν να προεξέχουν ως επακόλουθο της επιθυμίας του να τις εντοπίσει. Άλλο παράδειγμα είναι η αναγνώριση ενός γνώριμου προσώπου ανάμεσα σε πολλά πρόσωπα. Οι διεργασίες που σχετίζονται με την οπτική προσοχή από την κορυφή προς τη βάση θεωρείται πως σχετίζονται με ενεργοποίηση νευρώνων στις ανώτερες περιοχές του εγκεφάλου, σε αντίθεση με την προηγούμενη προσέγγιση που ξεκινά στα χαμηλότερα επίπεδα του οπτικού συστήματος [29], [31], [32].

Η τεχνική που έχει επικρατήσει ώστε να υπολογίζεται η προεξοχή αντικειμένων στο οπτικό περιβάλλον είναι ο χάρτης προεξοχής. Ορίζεται ως ένας σαφής χάρτης δύο διαστάσεων που κωδικοποιεί και προβάλλει το πόσο ευκρινή είναι τα διαφορετικά σημεία μιας εικόνας. Σαν ιδέα διατυπώθηκε πρώτα από τους C. Koch και S. Ullmann [33] που προσπάθησαν να εξηγήσουν τους βιολογικούς μηχανισμούς πίσω από την οπτική προσοχή μέσα από την εξόρυξη χαρακτηριστικών χαμηλού επιπέδου από εικόνες. Αρχικά εξάγεται ένα πλήθος από βασικά χαρακτηριστικά όπως ακμές, φωτεινότητα, προσανατολισμός και χρώμα, ταυτόχρονα για όλα τα σημεία της εικόνας και σε πολλαπλές χωρικές και χρονικές κλίμακες, σχηματίζοντας χάρτες χαρακτηριστικών που αναλύονται παράλληλα και ο καθένας αναπαριστά την «ποσότητα» ενός συγκεκριμένου χαρακτηριστικού σε κάθε σημείο της εικόνας. Οι χάρτες αυτοί συνδυάζονται σε έναν μοναδικό χάρτη προεξοχής που δίνει τα σημεία τα οποία, σύμφωνα με τα οπτικά χαρακτηριστικά που μελετήθηκαν, ξεχωρίζουν από τα γειτονικά τους. Ο ανταγωνισμός μεταξύ νευρώνων, συναρτήσεων δηλαδή που ενεργοποιούνται όταν ξεπεράσουν μια συγκεκριμένη τιμή κατωφλίου, θα αναδείξει την περισσότερο προεξέχουσα περιοχή που θα αποτελέσει και τον επόμενο οπτικό στόχο, μέσα από τη διαδικασία «ο νικητής τα παίρνει όλα». Όταν η περιοχή αυτή αναχαιτιστεί, το σύστημα μετατοπίζεται αυτόματα στην αμέσως επόμενη προεξέχουσα θέση, προσομοιώνοντας τη διαδικασία αναζήτησης στόχων που διενεργεί το οπτικό σύστημα [29], [31].

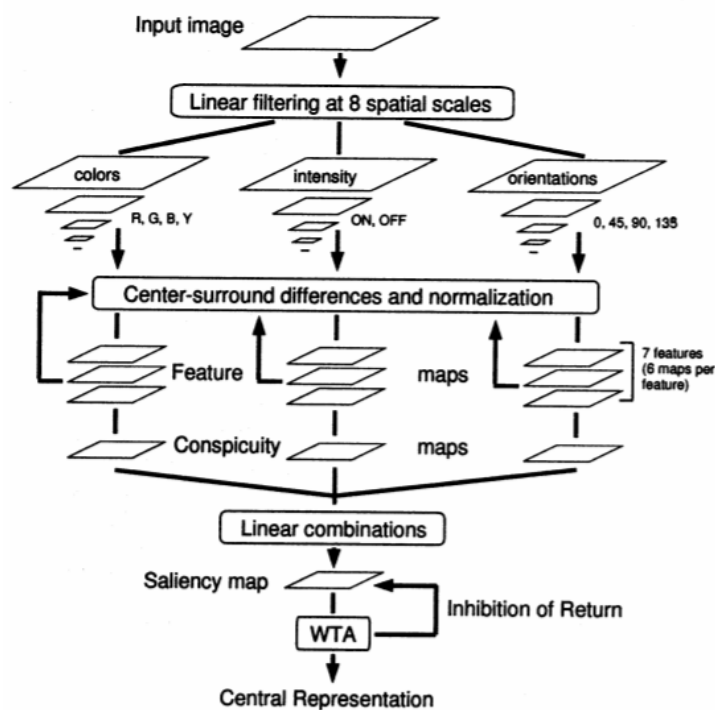


**Σχήμα 18:** Πρωτότυπο διάγραμμα μοντέλου προεξοχής. Central representation: Οπτικό ερέθισμα - εικόνα. Feature Maps: Χάρτες χαρακτηριστικών. Saliency Map: Χάρτης προεξοχής. WTA: "Winner takes it all" (ο νικητής τα παίρνει όλα) [31].



**Σχήμα 19:** Χάρτες προεξοχής [34].

Η πρώτη υλοποίηση σε μορφή ενός σαφούς και δομημένου υπολογιστικού μοντέλου, η οποία θεωρείται θεμελιώδης και έχει επηρεάσει την πλειοψηφία των μοντέλων από τότε, αναπτύχθηκε από τους L. Itti και C. Koch [31], οι οποίοι βασίστηκαν στην παραπάνω ιδέα εξάγοντας τρία χαρακτηριστικά: ένταση, προσανατολισμό και χρώμα. Από τότε, ένα μεγάλο σύνολο υπολογιστικών και βιολογικών μοντέλων έχει αναπτυχθεί με κίνητρο την κατανόηση των μηχανισμών που διαμορφώνουν την οπτική προσοχή στον άνθρωπο. Επιπλέον, η πρόβλεψη των σημείων που ένας άνθρωπος είναι πιθανότερο να εστιάσει είναι χρήσιμο σε πολλές εφαρμογές όπως συμπίεση εικόνων, ρομποτική, προσθετική αμφιβληστροειδούς, περικοπή εικόνων γύρω από σημαντικά σημεία, ιατρικές διαγνώσεις, αυτόματο εντοπισμό αντικειμένων και αλληλεπίδραση ανθρώπου-υπολογιστή [29], [35], [36].



**Σχήμα 20:** Αναλυτικό διάγραμμα του μοντέλου των Itti και Koch [31].

Οι αρχικές υλοποιήσεις βασίζονταν αποκλειστικά στην προσέγγιση από τη βάση προς την κορυφή διότι πιο πολύπλοκα χαρακτηριστικά ήταν δύσκολο να ενσωματωθούν σε ένα μοντέλο πρόβλεψης. Ωστόσο ένας χάρτης προεξοχής, προκειμένου να είναι πλήρης και να προσομοιώνει αποτελεσματικά την ανθρώπινη συμπεριφορά οφείλει να συνδυάζει πληροφορίες από τη βάση προς την κορυφή και από την κορυφή προς τη βάση. Με γνώμονα αυτό, η προσπάθεια εστίασε στη δημιουργία νέων μοντέλων που θα βελτιώναν τα ήδη υπάρχοντα, ενσωματώνοντας νέα, υψηλότερου επιπέδου χαρακτηριστικά και τεχνικές. Τα τελευταία χρόνια, το ενδιαφέρον μετατοπίστηκε σε μοντέλα πρόβλεψης που εκπαιδεύονται αμιγώς μέσα από σύνολα δεδομένων, αρχικά με τεχνικές όρασης υπολογιστών [82] και πλέον με μεθόδους βαθιάς μάθησης και πιο συγκεκριμένα, μιας και το επίκεντρο είναι



εικόνες, με συνελκτικά νευρωνικά δίκτυα. Οι υλοποιήσεις με βαθιά νευρωνικά δίκτυα επιδεικνύουν εντυπωσιακές επιδόσεις σε σχέση με τα κλασικά μοντέλα, αφού η ιεραρχική και πολύπλοκη δομή και λειτουργία τους, τους επιτρέπει να εντοπίζουν περισσότερα χαρακτηριστικά και αναπαραστάσεις που ενδέχεται να προσελκύσουν την οπτική προσοχή, ενσωματώνοντας το μηχανισμό από την κορυφή προς τη βάση. Πολλές φορές είναι δύσκολο κάποιος να ξεχωρίσει την εκτίμησή τους και τον πραγματικό ανθρώπινο χάρτη εστιάσεων. Στις περισσότερες περιπτώσεις χρησιμοποιούνται δίκτυα που είναι ήδη εκπαιδευμένα για κάποιο πρόβλημα ανάλυσης σκηνής όπως ταξινόμηση και τροποποιούνται κατάλληλα για να μάθουν να προβλέπουν τις πιθανές περιοχές εστίασης. Συχνά ωστόσο συστήνονται και νέες αρχιτεκτονικές που εκπαιδεύονται από την αρχή, αντιμετωπίζοντας την πρόβλεψη προεξοχής σαν ένα πρόβλημα ταξινόμησης, εκμεταλλευόμενα τα πλέον επαρκώς μεγάλα διαθέσιμα σύνολα δεδομένων με πειραματικές καταγραφές εστιάσεων από παρατηρητές. Το eDN [37], που αποτελεί την πρώτη προσπάθεια πρόβλεψης της προεξοχής με συνελκτικά δίκτυα, τα Deep Gaze I [38] και Deep Gaze II [39] που αυξάνουν τον αριθμό των χρησιμοποιούμενων επιπέδων, το MrCNN [40] που συνδυάζει τρία δίκτυα που εκπαιδεύονται σε διαφορετικές κλίμακες των εικόνων, το DeepFix [41] που εκμεταλλεύτηκε πρώτο τα αμιγώς συνελκτικά νευρωνικά δίκτυα και τα SAM Nets [42] που αξιοποιούν ένα Επαναλαμβανόμενο Νευρωνικό Δίκτυο (Long Short-Term Memory) [43] για δυναμική βελτίωση του αποτελέσματος είναι ορισμένες αξιοσημείωτες επιτυχίες της βαθιάς μάθησης στην πρόβλεψη των σημείων εστίασης. Τα Δημιουργικά Ανταγωνιστικά Δίκτυα (Generative Adversarial Networks) συμμετέχουν ολοένα και περισσότερο στις εξελίξεις ενώ τα Επαναλαμβανόμενα Νευρωνικά Δίκτυα (Recurrent Neural Networks) επεκτείνουν τη δυνατότητα πρόβλεψης προεξοχής σε δυναμικά και χρονικά μεταβαλλόμενα δεδομένα όπως εικονοσειρές [32], [35], [36].

Η τάση προς τέτοιες προσεγγίσεις είναι αλήθεια πως μας παρέχει πολύ ισχυρά υπολογιστικά εργαλεία και πολλές φορές και σε πολλές εφαρμογές η κορυφαία απόδοση είναι ο βασικός στόχος. Ωστόσο, τα βαθιά νευρωνικά δίκτυα συχνά αντιμετωπίζονται σαν «μαύρα κουτιά». Με άλλα λόγια, δεν μπορούμε να ορίσουμε σαφώς το θεωρητικό υπόβαθρο που εξηγεί τη σύγκλισή τους, αφού βασίζονται περισσότερο σε εμπειρικές μεθόδους «δοκιμής και σφάλματος» και λιγότερο σε τεκμηριωμένες γνώσεις πάνω στο εκάστοτε πρόβλημα. Προκύπτει λοιπόν το ζήτημα του κατά πόσο οι προσεγγίσεις βαθιάς μάθησης βοηθούν στην καλύτερη κατανόησή μας και στην επίλυση θεμελιωδών ερωτημάτων σχετικά με το μηχανισμό της ανθρώπινης οπτικής προσοχής, υπό το πρίσμα της βιολογικής έρευνας και της γνωστικής νευροεπιστήμης [32].

## 2.6 Βαθιά Μάθηση και Συνελκτικά Νευρωνικά Δίκτυα

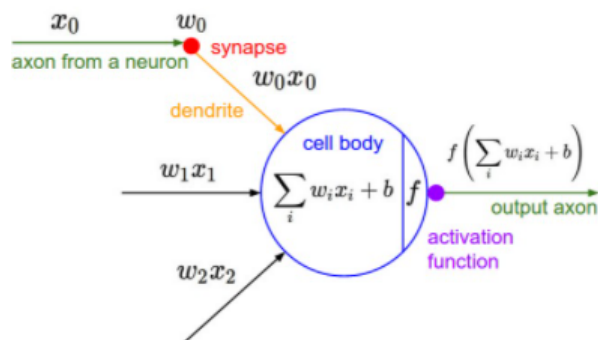
Είναι γνωστό και διαδεδομένο πλέον πως η τεχνητή νοημοσύνη έχει εισχωρήσει σε πολλές πτυχές της μοντέρνας κοινωνίας. Αποτελεί ίσως το ταχύτερα ακμάζον επιστημονικό πεδίο σήμερα με πληθώρα ενεργών εφαρμογών και ανοιχτών ερευνητικών περιοχών. Η αυτοματοποίηση εργασιών, τα συστήματα προτάσεων, η υποβοήθηση ιατρικών διαγνώσεων, η υποστήριξη της επιστημονικής έρευνας και τα μοντέλα προβλέψεων είναι μόνο λίγες από



τις επιτυχίες της. Τις τελευταίες δεκαετίες, ένας συγκεκριμένος τομέας τεχνητής νοημοσύνης συνέβαλε τα μέγιστα στην εντυπωσιακή πρόοδο της, η βαθιά μάθηση (deep learning). Η βαθιά μάθηση εντάσσεται στην ευρύτερη οικογένεια της μηχανικής μάθησης, των μεθόδων δηλαδή που μελετούν την ικανότητα μοντέλων και συστημάτων να αντιμετωπίσουν πραγματικά και πολύπλοκα προβλήματα εξάγοντας πρότυπα από δεδομένα, με στόχο την απόκτησης γνώσης και τη λήψη αποφάσεων που μοιάζουν υποκειμενικές.

Οι αλγόριθμοι βαθιάς μάθησης επιλύουν προβλήματα τα οποία ένας άνθρωπος εκτελεί διαισθητικά, δηλαδή φτάνει εύκολα στη λύση αλλά δυσκολεύεται να τα περιγράψει ως μια σειρά από αυστηρά καθορισμένα βήματα και μαθηματικούς κανόνες, όπως η αναγνώριση προσώπων σε μια εικόνα ή η μετατροπή ακουστικού λόγου σε γραπτό. Αποτελούν την τεχνική που επιτρέπει σε υπολογιστικά μοντέλα να αποκτούν εμπειρία και να βελτιώνονται μέσα από δεδομένα. Βασίζονται στα τεχνητά νευρωνικά δίκτυα και απαρτίζονται από ένα μεγάλο πλήθος επεξεργαστικών επιπέδων (εξού και ο όρος “βαθιά”) που διασπών το κεντρικό πρόβλημα σε απλούστερα και ευκολότερα. Αυτό τους προσδίδει την ικανότητα να μπορούν να διεκπεραιώνουν πολύπλοκες εργασίες με έναν τρόπο που παρομοιάζει την αφαιρετική σκέψη. Η δύναμη των συστημάτων αυτών εντοπίζεται στην ιδέα με την οποία λειτουργούν. Δε διαμορφώνουν απλώς μια σχέση μεταξύ κάποιων ήδη γνωστών χαρακτηριστικών και της εξόδου, αλλά μαθαίνουν από μόνα τους, από την αρχή και μέσα από μια επαναληπτική διαδικασία εκπαίδευσης και βελτιστοποίησης, τα ίδια τα χαρακτηριστικά ώστε να πετύχουν όσο γίνεται καλύτερα το στόχο τους. Με αυτόν τον τρόπο προσαρμόζονται εύκολα σε διαφορετικές εργασίες, απαιτείται μειωμένη ανθρώπινη παρέμβαση και παράλληλα επιτυγχάνεται υψηλή απόδοση.

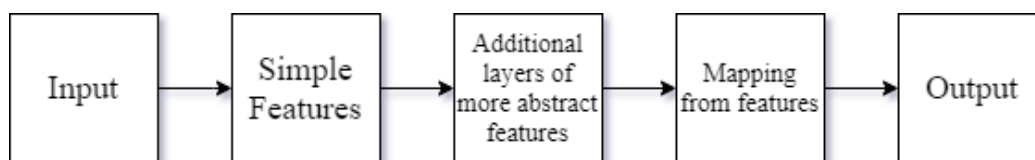
Ένα νευρωνικό δίκτυο βαθιάς μάθησης αποτελείται από μια αλληλουχία διαδοχικών επιπέδων που θα συνεργαστούν ώστε το δίκτυο μέσα από πολύπλοκους μαθηματικούς μετασχηματισμούς να υπολογίσει μια συνάρτηση μεταξύ της εισόδου, όπως μια εικόνα, και της εξόδου, όπως η ταξινόμηση της εικόνας σε κάποια κατηγορία. Για να πάμε από το ένα επίπεδο στο επόμενο, ένα σύνολο από νευρώνες υπολογίζουν το σταθμισμένο άθροισμα των εισόδων που δέχονται από το προηγούμενο επίπεδο και εισάγουν το αποτέλεσμα σε μια μη γραμμική συνάρτηση. Κατά τη διαδικασία αυτή ένας αλγόριθμος εκπαίδευσης, όπως της απότομης κατάβασης (gradient descent) με υπολογισμό της παραγώγου μέσω οπίσθιας διάδοσης (backpropagation), θα υποδείξει στους νευρώνες τον πιο αποτελεσματικό τρόπο μεταβολής των παραμέτρων τους.



Σχήμα 21: Τεχνητός νευρώνας.<sup>1</sup>

Η ιδέα αυτή δε διαφέρει πολύ από ένα κλασικό τεχνητό νευρωνικό δίκτυο. Πράγματι, η βαθιά μάθηση δεν αποτελεί ξαφνική ανακάλυψη των ημερών μας, αλλά χρονολογείται περίπου 80 χρόνια πίσω, γνωρίζοντας αρχικά περιορισμένη δημοφιλία. Η απότομη άνοδος και η αναγνώριση οφείλεται στην εξέλιξη της επιστήμης των υπολογιστών που απολαμβάνουμε σήμερα. Πλέον έχουμε τους υπολογιστικούς πόρους και τη γνώση να κατασκευάζουμε δίκτυα με πάρα πολλά επίπεδα που επιλύουν απαιτητικά προβλήματα πιο αποτελεσματικά. Εκπαιδεύουμε τα μοντέλα με μεγαλύτερα σύνολα δεδομένων επιτρέποντας τους να γενικεύσουν τις γνώσεις τους σε άγνωστα δεδομένα με μεγαλύτερη αξιοπιστία και να προσεγγίσουν ακόμα περισσότερο ή ακόμη και να ξεπεράσουν τις ανθρώπινες επιδόσεις, ενώ διαθέτουμε το υλικό που θα επιταχύνει ραγδαία διαδικασίες που κάποτε χρειαζόνταν μήνες για να ολοκληρωθούν.

Συνεχώς ανακαλύπτουμε προβλήματα στα οποία αλγόριθμοι βαθιάς μάθησης είναι αυτοί που δίνουν τη βέλτιστη λύση. Τα πλεονεκτήματα που διαθέτουν οδηγούν με μαθηματική βεβαιότητα στο συμπέρασμα πως το πεδίο αυτό θα εξακολουθήσει να παρουσιάζει βελτίωση και ευκαιρίες και για τα χρόνια που ακολουθούν [44], [45].



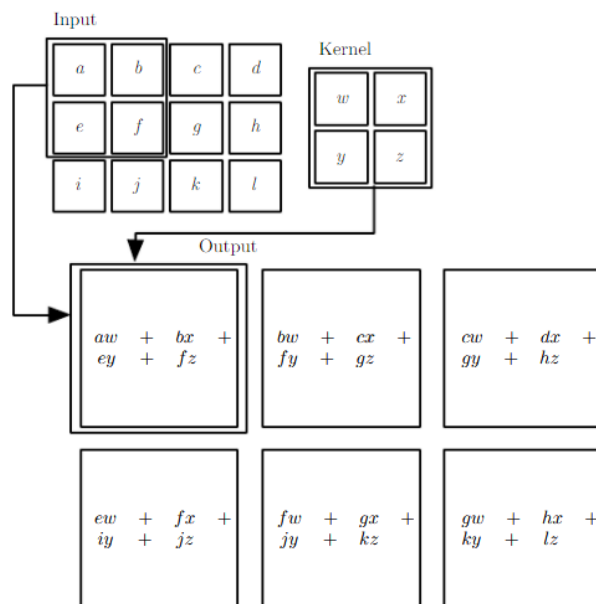
Σχήμα 22: Βαθιά Μάθηση [44].

Ως συνελκτικά ορίζονται τα τεχνητά νευρωνικά δίκτυα που χρησιμοποιούν την πράξη της συνέλιξης σε τουλάχιστον ένα από τα επίπεδα τους. Είναι ιδιαίτερα δημοφιλή σε ένα εκτεταμένο σύνολο εφαρμογών και μία από τις ευρύτερα εφαρμοσμένες αρχιτεκτονικές νευρωνικών δικτύων. Ειδικεύονται στην επεξεργασία δεδομένων που είναι σε μορφή πολυδιάστατου πίνακα (tensor), όπως μια έγχρωμη εικόνα που αποτελείται από τρεις πίνακες

<sup>1</sup> πηγή: Google Image Search

δύο διαστάσεων με τιμές έντασης για τα τρία χρωματικά κανάλια ή μια ακολουθία από εικόνες που λειτουργούν σαν δείγματα σε διακριτά χρονικά διαστήματα.

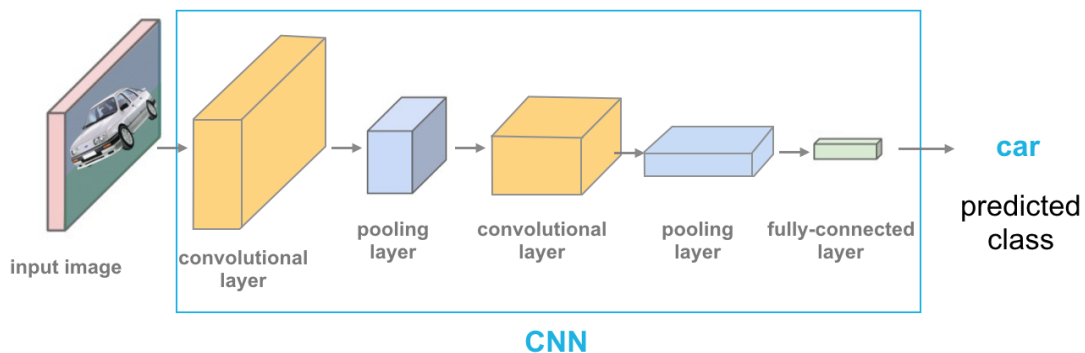
Ένα τυπικό συνελκτικό νευρωνικό δίκτυο δομείται από, κατά κανόνα, αρκετά επίπεδα. Τα πρώτα επίπεδα είναι συνελκτικά και, στη συντριπτική πλειοψηφία των υλοποιήσεων, επίπεδα υποδειγματοληψίας ή ομαδοποίησης (pooling layer). Η διακριτή συνέλιξη πραγματοποιείται μεταξύ της εισόδου και ενός συνόλου από πολυδιάστατους πίνακες στα διαφορετικά επίπεδα που ονομάζονται πυρήνες ή φίλτρα, με παραμέτρους που μεταβάλλονται μέσα από τη διαδικασία εκπαίδευσης. Οι πυρήνες στο κάθε επίπεδο εφαρμόζουν παράλληλες συνελίξεις σε όλα τα κανάλια της εισόδου από το προηγούμενο επίπεδο μέσω της γραμμικής πράξης του πολλαπλασιασμού πινάκων μεταξύ των συντελεστών τους και των τιμών στην είσοδο. Ο στόχος τους είναι να εντοπίσουν ένα πλήθος διαφορετικών χαρακτηριστικών, τα οποία όσο απομακρυνόμαστε από το πρώτο επίπεδο γίνονται πιο σύνθετα. Ουσιαστικά υλοποιούν την κεντρική ιδέα της βαθιάς μάθησης που αναφέραμε παραπάνω, τη διάσπαση δηλαδή του κεντρικού προβλήματος σε απλούστερα. Το αποτέλεσμα περνά από το στάδιο ανίχνευσης, μια μη γραμμική συνάρτηση (συνήθως ένας μη γραμμικός ανορθωτής (ReLU)) και είναι ένα σύνολο φιλτραρισμένων εκδοχών της εισόδου στο εκάστοτε επίπεδο. Για παράδειγμα, αν έχουμε στην είσοδο μια εικόνα την οποία επιθυμούμε να εντάξουμε σε μια κατηγορία ανάλογα με το τι απεικονίζει, το πρώτο επίπεδο θα ψάξει για ακμές, το δεύτερο για γωνίες και περιγράμματα και το τρίτο θα αναζητήσει ολόκληρα αντικείμενα όπως κάποιο πρόσωπο.



**Σχήμα 23:** Συνέλιξη δύο διαστάσεων. Στην απλή μορφή της, για διαστάσεις εισόδου  $H_1 \times W_1$  και πυρήνα  $H_2 \times W_2$  η έξοδος έχει διαστάσεις  $(H_1 - H_2 + 1) \times (W_1 - W_2 + 1)$  [44].

Για να τροποποιήσουμε επιπλέον το αποτέλεσμα του κάθε επιπέδου συνηθίζεται να εφαρμόζεται μια συνάρτηση υποδειγματοληψίας. Η έξοδος του δικτύου σε κάθε σημείο αντικαθίσταται από κάποια στατιστική ποσότητα της γειτονιάς του σημείου, όπως το μέγιστο ή ο μέσος όρος, συγχωνεύοντας σημασιολογικά κοντινά χαρακτηριστικά σε ένα. Για παράδειγμα η συνάρτηση max pooling θα αντικαταστήσει ένα πλέγμα  $2 \times 2$  με το μέγιστο των τεσσάρων τιμών. Η τεχνική αυτή βοηθά την αναπαράσταση να είναι λιγότερο ευαίσθητη σε μικρές αλλαγές της εισόδου, και είναι απαραίτητη σε περιπτώσεις που μας ενδιαφέρει το αν υπάρχει ένα χαρακτηριστικό στην εικόνα αλλά όχι η θέση του. Αν το μοντέλο προσπαθεί να αναγνωρίσει ένα χειρόγραφο «7», θέλουμε να το αναγνωρίσει σε όποια κλίση και να είναι στην εικόνα. Επιπλέον μειώνει τις παραμέτρους του δικτύου και συνεπώς τους υπολογιστικούς πόρους που απαιτούνται, αφού μειώνει τις διαστάσεις της εξόδου των φίλτρων.

Τα συνελκτικά νευρωνικά δίκτυα εκμεταλλεύονται δύο βασικές ιδιότητες. Χρησιμοποιούν πυρήνες μικρότερων διαστάσεων από της εισόδου, δηλαδή κάθε νευρώνας ενός πυρήνα επεξεργάζεται τοπικά και μια φορά ένα συγκεκριμένο τμήμα της εισόδου, που ονομάζεται υποδεκτικό πεδίο, εμπνευσμένο από τις νευροεπιστημονικές μελέτες. Επίσης, οι νευρώνες ενός πυρήνα έχουν ίδιες τιμές στα βάρη μεταξύ τους και ο πυρήνας μπορεί έτσι να εντοπίσει το χαρακτηριστικό που τον ενδιαφέρει σε όποιο σημείο της εικόνας και αν βρίσκεται.



**Σχήμα 24:** Τυπική αρχιτεκτονική συνελκτικού δικτύου.<sup>2</sup>

Η θεμελιώδης ιδέα πίσω από τον τρόπο με τον οποίο τα μοντέλα αυτά δομούνται και λειτουργούν έχει σημαντική επιρροή από τις βιολογικές μελέτες του οπτικού συστήματος και ιδιαίτερα από τα ευρήματα των D. Hubel και T. Wiesel [46], [47], οι οποίοι τιμήθηκαν και με το βραβείο Nobel. Διεγείροντας τον πρωτοταγή οπτικό φλοιό της γάτας, εντόπισαν κύτταρα που αποκρίνονται επιλεκτικά σε διεγέρσεις με πολύ συγκεκριμένα χαρακτηριστικά, όπως φωτεινές γραμμές σε συγκεκριμένη κλίση ή σε κίνηση προς συγκεκριμένη κατεύθυνση. Ταξιόνησαν τα κύτταρα αυτά σε δύο κατηγορίες. Στην πρώτη ενέταξαν αυτά που ενδιαφέρονται για αλλαγές αποκλειστικά στο σαφώς ορισμένο υποδεκτικό τους πεδίο

<sup>2</sup> πηγή: Google Image Search

(simple cells). Τα συνελκτικά επίπεδα μιμούνται τη συμπεριφορά αυτών των κυττάρων. Στη δεύτερη ενέταξαν τα κύτταρα που λαμβάνουν είσοδο από ένα πλήθος κυττάρων της πρώτης κατηγορίας, εντοπίζουν πρότυπα ανεξαρτήτως της θέσης τους στο οπτικό πεδίο ενώ δεν παρουσιάζουν ένα αυστηρά δομημένο υποδεκτικό πεδίο (complex cells). Αποτέλεσαν την έμπνευση για τις μονάδες ομαδοποίησης στα συνελκτικά δίκτυα. Οι μελέτες αυτές έθεσαν τις βάσεις για τη διατύπωση της βασικής αρχής ότι το οπτικό σύστημα διασπά το πολύπλοκο πρόβλημα της ερμίνευσης της οπτικής πληροφορίας σε απλούστερα χαρακτηριστικά και αναπαραστάσεις.

Τα συνελκτικά νευρωνικά δίκτυα αξιοποιούνται με μεγάλη επιτυχία σε προβλήματα που αφορούν επεξεργασία εικόνας όπως αναγνώριση προτύπων και αντικειμένων, ταξινόμηση σε κλάσεις και κατάτμηση. Τα τελευταία χρόνια εφαρμόζονται και σε προβλήματα πραγματικού χρόνου όπως η αναγνώριση προσώπου και η αυτόνομη οδήγηση [44], [45].

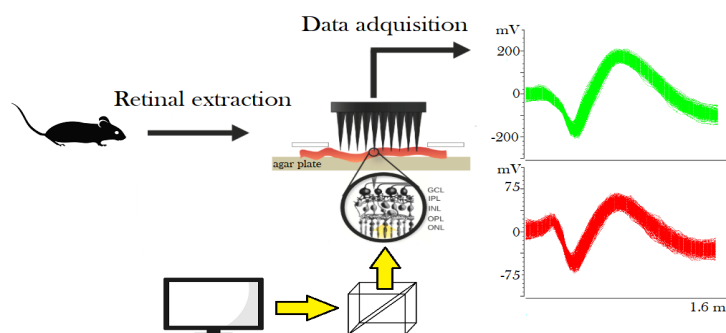
## Κεφάλαιο 3

### Μέθοδοι

#### 3.1 Πείραμα καταγραφής των αποκρίσεων γαγγλιακών κυττάρων

Τα δεδομένα που χρησιμοποιήσαμε για την εκπαίδευση του μοντέλου προέρχονται από το βιολογικό πείραμα του Antonio Lozano και της ερευνητικής του ομάδας [51]. Προκειμένου να καταγράψουν *in vitro* την ηλεκτρική δραστηριότητα πολλαπλών γαγγλιακών κυττάρων, σε πρώτη φάση προσάρμοσαν τα ποντίκια σε συνθήκες σκοταδιού για μία ώρα προκειμένου να αυξηθεί η ευαισθησία του αμφιβληστροειδούς σε μεταβολές φωτεινότητας [14]. Μετά από χορήγηση αναισθησίας, πραγματοποιήθηκε εξάρθρωση του αυχένα, μια τυπική μέθοδος ευθανασίας μικρών ζώων σε εργαστήρια. Αφού αφαιρέθηκαν τα μάτια χωρίς να πειραχθούν οι ιστοί και τα νεύρα, αφαιρέθηκαν και οι αμφιβληστροειδείς, οι οποίοι τοποθετήθηκαν σε ειδικό δισκίο βιολογικής καλλιέργειας με το τελευταίο επίπεδο, αυτό των γαγγλιακών κυττάρων, να κοιτά προς τα πάνω. Τέλος, ο βιολογικός ιστός προετοιμάστηκε κατάλληλα ώστε να τοποθετηθούν τα ηλεκτρόδια και να ξεκινήσει η καταγραφή. Η όλη διαδικασία πραγματοποιήθηκε κάτω από χαμηλό κόκκινο φωτισμό.

Η καταγραφή της δραστηριότητας των κυττάρων έγινε εξωκυτταρικά. Χρησιμοποιήθηκε πλέγμα πολλαπλών ηλεκτροδίων (multi-electrode array) διαστάσεων  $10 \times 10$  με απόσταση  $400 \mu\text{m}$  μεταξύ τους, ενώ η ψηφιοποίηση των δεδομένων από κάθε ηλεκτρόδιο έγινε σε 16-bit ανάλυση με βαθμό δειγματοληψίας 30 kHz. Η ανάλυση και η επεξεργασία των αποτελεσμάτων όπως και η ταυτοποίηση των αποκρίσεων των διαφορετικών νευρώνων έγινε με ειδικό λογισμικό ανοιχτού κώδικα [51].

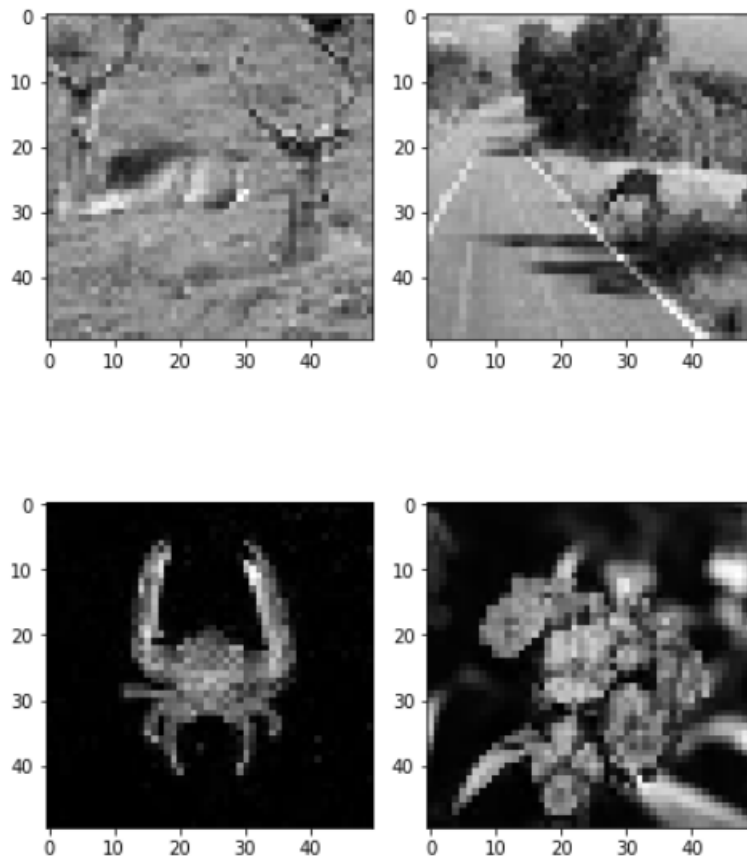


Σχήμα 25 : Πείραμα απόκτησης των δεδομένων [51].

### 3.2 Περιγραφή Δεδομένων

Το σύνολο δεδομένων που χρησιμοποιήθηκε κατά το πείραμα κατασκευάστηκε από 4890 εικόνες διαστάσεων  $50 \times 50$  pixels. Στην εργασία μας δεν ενδιαφερόμαστε για την επίδραση των χρωμάτων, άρα όλες οι εικόνες είναι σε κλίμακα του γκρι (grayscale). Αποτελείται από φωτογραφίες του φυσικού περιβάλλοντος όπως ζώα και στοιχεία της φύσης (λουλούδια, δέντρα), δεν περιλαμβάνει δηλαδή τεχνητά κατασκευασμένες εικόνες, όπως αποκαλούμε για παράδειγμα ένα σύνολο εικόνων λευκού γκαουσιανού θορύβου.

Κάθε καρτέ είχε διάρκεια 10 ms και κάθε εικόνα προβλήθηκε στα γαγγλιακά κύτταρα του αμφιβληστροειδούς για 50 ms. Συνεπώς για κάθε εικόνα έχουμε πέντε διαδοχικά αντίγραφα, με την τελική βάση δεδομένων να αποτελείται από 24450 εικόνες με συνολική διάρκεια 244.5 δευτερόλεπτα. Επιπλέον έχει γίνει κανονικοποίηση ώστε κάθε εικόνα να έχει εύρος στο διάστημα  $[0, 1]$ .



Σχήμα 26: Εικόνες από το σύνολο δεδομένων.

Εφόσον μελετάμε γαγγλιακά κύτταρα του αμφιβληστροειδούς το πείραμα αφορά δυναμικά ενεργείας. Αν σε ένα γαγγλιακό κύτταρο παρουσιαστεί πολλές φορές ακριβώς το ίδιο ερέθισμα, θα παρουσιάσει μεταβλητότητα στις αποκρίσεις του. Για το λόγο αυτό, οι

νευρώνες αντιμετωπίζονται στοχαστικά και το ενδιαφέρον δεν εστιάζεται στο ακριβές χρονικό σημείο πυροδότησης αλλά σε έννοιες όπως το πόσες φορές πυροδοτήσε, ο ρυθμός πυροδότησης ή η πιθανότητα να πυροδοτήσει, μέσα σε κάποιο χρονικό διάστημα  $[t, t + \Delta t]$  [8].

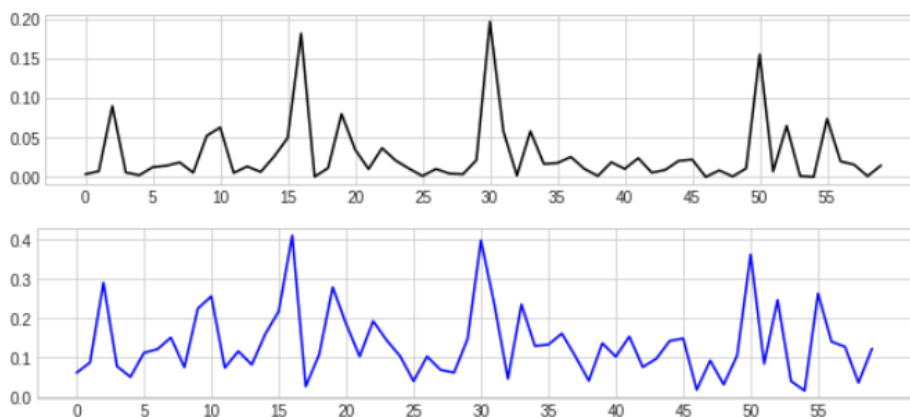
Στη διάθεσή μας είχαμε δεδομένα για 60 κύτταρα. Αρχικά προσδιορίστηκαν πειραματικά τα χρονικά σημεία που το κύτταρο πυροδοτήσε παλμούς ως απάντηση στις εικόνες που του παρουσιάζονταν. Έπειτα, ο άξονας του χρόνου διαιρέθηκε σε μικρά χρονικά παράθυρα διάρκειας  $\Delta t = 10$  ms, όσος και ο χρόνος προβολής ενός καρέ κατά τη διέγερση και υπολογίστηκε το πλήθος των δυναμικών ενεργείας (spikes) που πραγματοποιήθηκαν εντός του κάθε διαστήματος  $[t, t + 0.01]$ , για μια ολοκληρωμένη προβολή όλων των εικόνων και για κάθε κύτταρο ξεχωριστά. Αυτή ήταν και η αρχική μορφή των δεδομένων μας.

Αν διαιρέσουμε επιπλέον και με τη διάρκεια  $\Delta t$  προκύπτει υπό τη μορφή διακριτών τιμών σε μορφή ιστογράμματος μια εκτίμηση του ρυθμού πυροδότησης (firing rate) σε κάθε χρονικό διάστημα ή αλλιώς της συνάρτησης πυκνότητας πιθανότητας εμφάνισης δυναμικού ενεργείας. Εναλλακτικά, το φιλτράρισμα με κάποια συνάρτηση παραθύρου όπως η γκαουσιανή υπολογίζει μια προσέγγιση του ίδιου μεγέθους, συνεχή στο χρόνο, με μονάδα μέτρησης Hz (spikes/second). Για πιο ασφαλή αποτελέσματα και λόγω της στοχαστικότητας που αναφέρθηκε παραπάνω, σε πειράματα που αξιολογείται η σχέση νευρωνικής δραστηριότητας και του εξωτερικού ερεθίσματος που την προκαλεί, συνηθίζεται η διεξαγωγή πολλαπλών πειραμάτων με ακριβώς την ίδια διέγερση [8], [52].



Σχήμα 27: Ενδεικτικά διαγράμματα αποκρίσεων στις εικόνες για 10 s.





Σχήμα 28: Μέση τιμή και τυπική απόκλιση των αποκρίσεων των νευρώνων.

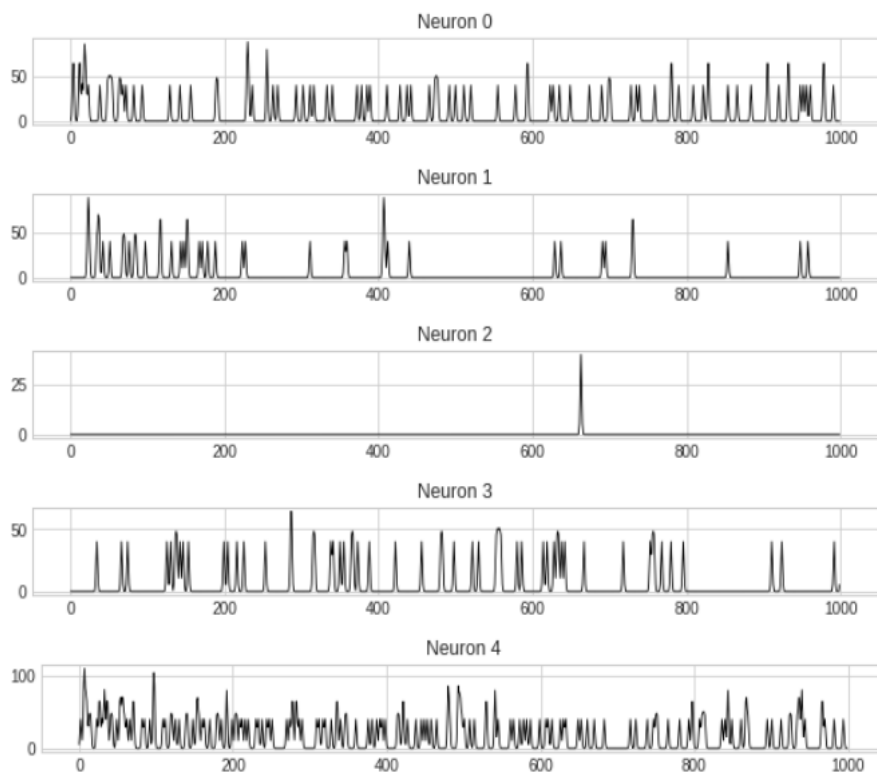
### 3.3 Είσοδος - Έξοδος του δικτύου

Ο αμφιβληστροειδής είναι ένα δυναμικό μοντέλο που προσαρμόζεται στις ιδιαιτερότητες του εξωτερικού περιβάλλοντος όπως τα επίπεδα φωτεινότητας και οι αντιθέσεις ώστε να λειτουργεί αποδοτικά. Η απόκριση ενός γαγγλιακού κυττάρου δεν εξαρτάται μόνο από το ερέθισμα που δέχεται εκείνη τη στιγμή αλλά και από διεγέρσεις που προηγούνται αυτού, έχει δηλαδή μνήμη. Ως μνήμη ορίζεται το χρονικό διάστημα στο οποίο η ύπαρξη διέγερσης μπορεί να επηρεάσει την απόκριση του νευρώνα, και μπορεί να είναι φαινόμενα από 0.1 έως και 10 s πίσω στο χρόνο. Στοχαστικά μοντέλα όπως η διεργασία Poisson αγνοούν τη μνήμη του νευρώνα για λόγους απλοποίησης, χάνουν ωστόσο γεγονότα τα οποία είναι σημαντικά για την κατανόηση της συμπεριφοράς του συστήματος. Γενικότερα, σε υπολογιστικά μοντέλα πρόβλεψης νευρωνικής δραστηριότητας στον αμφιβληστροειδή το ενδιαφέρον εστιάζεται στη βραχυπρόθεσμη μνήμη (μερικές εκατοντάδες milliseconds), η οποία καθορίζεται εμπειρικά [14], [20], [53]. Κατά τη μελέτη μας παρατηρήσαμε μεγάλη ευαισθησία στη συγκεκριμένη παράμετρο. Όπως θα αναλύσουμε καλύτερα και στο κεφάλαιο των αποτελεσμάτων στο στάδιο του υπολογισμού των υποδεκτικών πεδίων, η αναζήτηση κατάλληλης τιμής για τη μνήμη των κυττάρων λειτούργησε ως μια επιπλέον μέθοδος επιλογής και αξιολόγησης του μοντέλου μας. Αποτέλεσε μια κρίσιμη υπερπαράμετρο του δικτύου που έδειχνε όχι μόνο το πόσο καλά εκπαιδεύεται το σύστημα στα δεδομένα, αλλά και το πόσο αποτελεσματικά έχει κατανοήσει και μπορεί να αναπαράγει τα βιολογικά φαινόμενα πίσω από τα δεδομένα αυτά.

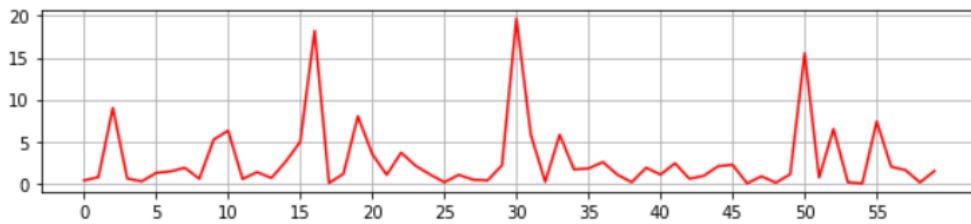
Προκειμένου λοιπόν το μοντέλο να μπορέσει να εντοπίσει πρότυπα και να εξάγει πληροφορίες όχι μόνο στη διάσταση του χώρου, αλλά και του χρόνου, με απώτερο στόχο να προβλέψει με μεγαλύτερη αποτελεσματικότητα την κυτταρική δραστηριότητα, η βάση δεδομένων πέρασε από μια συγκεκριμένη προεπεξεργασία. Κάθε δείγμα εκπαίδευσης δεν αποτελείται πλέον από μία μόνο εικόνα μαζί με την απόκριση σε αυτή, αλλά από μία εικόνα μαζί με ένα σταθερά ορισμένο πλήθος εικόνων που προηγούνται αυτής. Για τον αριθμό αυτό επιλέχθηκε η τιμή 40 που αντιστοιχεί σε χρονική διάρκεια 400 ms. Αυτό σημαίνει πως

μελετήθηκαν φαινόμενα που επηρεάζουν τη συμπεριφορά του νευρώνα σε ένα χρονικό όριο 400 ms πριν την απόκριση. Στα δεδομένα δηλαδή έχει προστεθεί μια επιπλέον διάσταση στο πεδίο του χρόνου η οποία προσδίδει μια αναπαράσταση εικονοσειράς, επιτρέποντας στο δίκτυο να εφαρμόσει τους κατάλληλους υπολογισμούς για τη βέλτιστη πρόβλεψη της επιθυμητής εξόδου. Στην παραπάνω προεπεξεργασία έχει ληφθεί υπόψη και το γεγονός ότι στην αρχή του πειράματος και μέχρι να φτάσει η πρώτη εικόνα, ο αμφιβληστροειδής βλέπει μόνο σκοτάδι.

Το δεύτερο στάδιο της προεπεξεργασίας αφορά τις αποκρίσεις. Η επιθυμητή έξοδος του μοντέλου είναι μια συνεχής συνάρτηση που ορίζεται ως ο στιγμιαίος ρυθμός πυροδότησης του νευρώνα (firing rate (Hz)), την οποία δημιουργήσαμε από την εξομάλυνση μέσω συνέλιξης των αρχικών δεδομένων με μια γκαουσιανή συνάρτηση τυπικής απόκλισης 10 ms, όσος δηλαδή ο χρόνος δειγματοληψίας κατά το πείραμα. Με τον τρόπο αυτό διευκολύνουμε την εκπαίδευση, με το πρόβλημα που καλείται να αντιμετωπίσει το νευρωνικό δίκτυο να μετατρέπεται από ταξινόμηση σε διακριτές κλάσεις (classification) σε πρόβλεψη συνεχών τιμών στην έξοδο (regression) [51]



**Σχήμα 29:** Νευρωνική δραστηριότητα γαγγλιακών κυττάρων μετά από εξομάλυνση.



**Σχήμα 30:** Μέση τιμή του ρυθμού πυροδότησης των νευρώνων.

Το αποτέλεσμα είναι να έχουμε επιβλεπόμενη διαδικασία μάθησης, με κάθε δείγμα με το οποίο θα τροφοδοτήσουμε το δίκτυο να απαρτίζεται από μία νέα εικόνα ακολουθούμενη από άλλες 39 που υποδηλώνουν το ιστορικό της, μαζί με το διάνυσμα με τις τιμές απόκρισης των νευρώνων για την πρώτη εικόνα, τις οποίες το δίκτυο θα επιχειρήσει να εκτιμήσει. Να σημειωθεί πως αυτό το τέχνασμα δε μας ανάγκασε να χρησιμοποιήσουμε συνελίξεις τριών διαστάσεων στο νευρωνικό δίκτυο. Προκειμένου να μπορέσουμε να ενσωματώσουμε τους απαραίτητους υπολογισμούς όχι μόνο στο πεδίο του χώρου αλλά και του χρόνου, το πρώτο συνελκτικό επίπεδο θα αντιμετωπίσει τη χρονική διάσταση του δείγματος όπως θα αντιμετώπιζε τα τρία χρωματικά κανάλια RGB σε μια εικόνα με χρώμα.



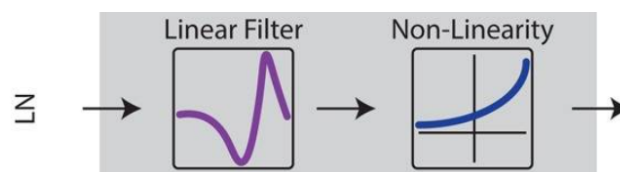
**Σχήμα 31:** Αναπαράσταση του πρώτου δείγματος εκπαίδευσης. Αποτελείται από μία εικόνα και το διάνυσμα με τις αποκρίσεις σε αυτήν όλων των κυττάρων μαζί με 39 μαύρα καρέ. Μετατοπίζοντας το κυλιόμενο παράθυρο προς τα δεξιά, δημιουργούμε τα υπόλοιπα δείγματα.

Μια επιπλέον τεχνική προεπεξεργασίας που εφαρμόζεται σε μελέτες πρόβλεψης αποκρίσεων στον αμφιβληστροειδή ή και στον εγκέφαλο είναι η απόρριψη ορισμένων νευρώνων πριν την εκπαίδευση, με βάση κάποιο κριτήριο. Αυτό μπορεί να είναι η χαμηλή απόδοσή τους σε κάποιο απλούστερο μοντέλο πρόβλεψης, όπως ένα Γενικευμένο Γραμμικό μοντέλο (Generalized Linear Model) [54] ή η χαμηλή συσχέτιση μεταξύ των αποκρίσεων

του νευρώνα για πολλαπλές επαναλήψεις της ίδιας ακολουθίας εικόνων, που τον ορίζει ως αναξιόπιστο για το πείραμα [55]. Πιο συχνά, ως ένδειξη απόρριψης πριν το πείραμα χρησιμοποιούνται περιπτώσεις κατά τις οποίες τεχνικές όπως της αντίστροφης συσχέτισης (Spike-Triggered Average) αδυνατούν να υπολογίσουν το κέντρο του υποδεκτικού πεδίου του νευρώνα, εξέλιξη που σημαίνει πως ο νευρώνας αποκρίθηκε ελάχιστα στην εικόνα. Αυτό συμβαίνει όταν κατά τη διαδικασία προβολής και καταγραφής, το υποδεκτικό πεδίο είναι εκτοπισμένο εκτός του πεδίου της εικόνας, γεγονός που πιθανότατα θα αποτρέψει το πλέγμα ηλεκτροδίων από το να συλλάβει βέλτιστα τα σήματα αυτού του νευρώνα [51], [56]. Στην εργασία μας επιλέξαμε να μην περιορίσουμε τη λίστα των νευρώνων, θέλοντας με αυτό τον τρόπο να δούμε την επίδραση της επιλογής αυτής στην ικανότητα του νευρωνικού δικτύου να μάθει, και αργότερα στον υπολογισμό των υποδεκτικών πεδίων.

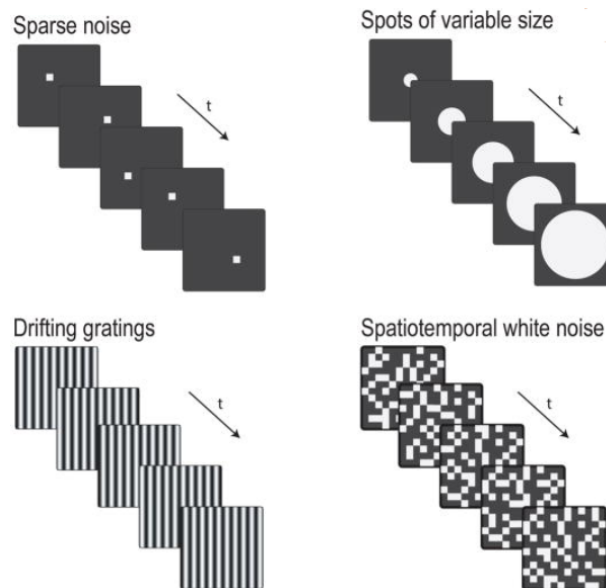
### 3.4 Μοντέλο εκτίμησης της απόκρισης του αμφιβληστροειδούς

Μέσα από τις προσπάθειες για κατανόηση τις βιολογικής λειτουργίας των κυκλωμάτων του αμφιβληστροειδούς έχουν δημιουργηθεί πολλά μαθηματικά και υπολογιστικά μοντέλα που εξηγούν και προβλέπουν την κυτταρική δραστηριότητα σε ποικιλία διαφορετικών διεγέρσεων. Το πιο διαδεδομένο μοντέλο είναι το Γραμμικό - Μη Γραμμικό (LN) μοντέλο που περιγράφει τη σχέση εισόδου - εξόδου μέσα από ένα γραμμικό φίλτρο (υποδεκτικό πεδίο του κυττάρου) που προσομοιώνει τη λειτουργία των φωτοϋποδοχέων και μια στατική μη γραμμική συνάρτηση που μετασχηματίζει τους παλμούς στην έξοδο σε βιολογικά πραγματικές τιμές. Ό,τι ευρήματα έχουμε σχετικά με το οπτικό σύστημα έχουν προκύψει από πειράματα με τεχνητά κατασκευασμένες εικόνες, όπως λευκό γκαουσιανό θόρυβο ή μπάρες σε διάφορους προσανατολισμούς που χρησιμοποιούνται για να αναδείξουν επιλεκτικότητα στην κατεύθυνση του φωτός. Το LN μοντέλο, παρά την απλότητα στη δομή του, είναι αρκετά αποτελεσματικό στην πρόβλεψη σε τέτοιας μορφής δεδομένα. Όπως αναφέραμε και στη θεωρία, η μη γραμμικότητα δεν αλλάζει το «τι» αναφέρει ο νευρώνας σχετικά με την είσοδο, αλλά το «πώς» το αναφέρει. Επομένως το LN μοντέλο υποθέτει πως η επεξεργασία του αμφιβληστροειδούς είναι κυρίως γραμμική. Κάτι τέτοιο όμως δεν ισχύει, αφού συνεχώς ανακαλύπτονται νέοι τύποι γαγγλιακών κυττάρων με ιδιόμορφη και μη γραμμική λειτουργία στο υποδεκτικό πεδίο τους, όπως είναι τα ON-OFF κύτταρα, γεγονός που περιορίζει το LN μοντέλο μόνο σε «απλά» κύτταρα, καθιστώντας το ανεπαρκές [54].



Σχήμα 32: Γραμμικό - Μη Γραμμικό μοντέλο [57].

Επιπλέον, το πρόβλημα με εικόνες που δεν είναι φυσικές είναι πως δεν είναι αρκετά πολύπλοκες ώστε να μπορούν να περιγράψουν πώς συμπεριφέρεται ο αμφιβληστροειδής στη ζωή ενός οργανισμού. Αδυνατούν να αναδείξουν μαζικά ιδιομορφίες της συμπεριφοράς και μη γραμμικούς υπολογισμούς των νευρώνων, κάτι το οποίο θεωρείται πως επιτυγχάνουν εικόνες του φυσικού περιβάλλοντος. Πολλές μελέτες υποστηρίζουν ότι το οπτικό σύστημα είναι εξελιγμένο με στόχο τη βέλτιστη αποκωδικοποίηση κυρίως φυσικών εικόνων. Τα LN μοντέλα, όπως και οι βελτιώσεις που έχουν προταθεί με βάση αυτά, δε γενικεύουν καλά σε φυσικές εικόνες παρουσιάζοντας πολύ χαμηλή απόδοση στην πρόβλεψη αποκρίσεων, ακόμη και αν έχουν εκπαιδευτεί σε τέτοια δεδομένα. Είναι απαραίτητα συνεπώς υπολογιστικά μοντέλα που θα είναι πιο κοντά σε έναν πραγματικό αμφιβληστροειδή και θα προβλέπουν αποτελεσματικά αποκρίσεις και σε φυσικές εικόνες, γενικεύοντας παράλληλα τις γνώσεις του σε ποικιλία διαφορετικών δεδομένων [54], [55].



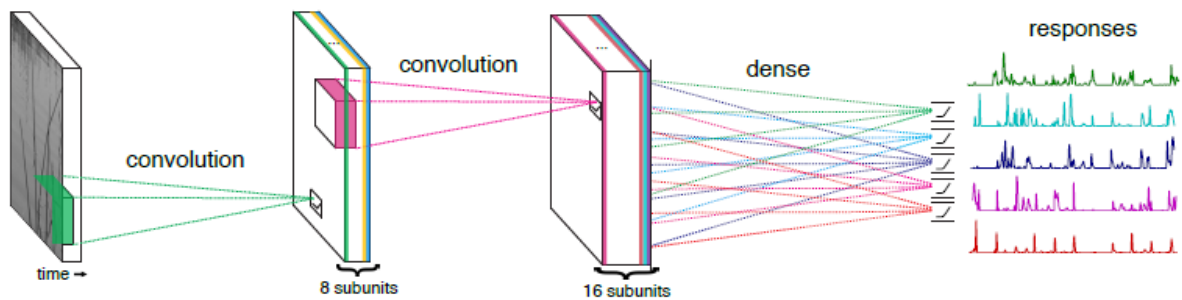
**Σχήμα 33:** Τεχνητά κατασκευασμένες εικόνες [57].

Τα βαθιά συνελκτικά νευρωνικά δίκτυα έρχονται να καλύψουν τις αδυναμίες των προηγούμενων μοντέλων. Έχουν αξιοποιηθεί επιτυχώς στην προσομοίωση περιοχών του οπτικού συστήματος στον εγκέφαλο όπως ο V1, ο V4 και ο κατώτερος κροταφικός φλοιός (IT), και τα τελευταία χρόνια και του αμφιβληστροειδούς. Αυτό που κάνουν είναι ότι δεν υποθέτουν από την αρχή κάποια συμπεριφορά για τους νευρώνες. Μαθαίνουν μέσα από τα δεδομένα και χάρη στην πολυεπίπεδη δομή τους είναι ικανά να εξηγήσουν σε μεγάλο ποσοστό βιολογικές αποκρίσεις, με τις δυνατότητες τους να μην περιορίζονται σε συγκεκριμένους τύπους κυττάρων. Παρουσιάζουν πολύ υψηλές επιδόσεις, μεταξύ άλλων και σε φυσικές εικόνες και μπορούν να κάνουν εκτιμήσεις για πολλούς νευρώνες ταυτόχρονα. Σημαντικό είναι επίσης πως όταν προσομοιώνουν κάποιο βιολογικό σύστημα όπως τον αμφιβληστροειδή, δεν είναι απλώς ένα «μαύρο κουτί» που επιτελεί υπολογισμούς. Έχει

βρεθεί πως αποτελούν ένα ισχυρό αυτοματοποιημένο εργαλείο που μπορεί, εκτός από καλές προβλέψεις, να αναπαράγει έγκυρες επιστημονικές υποθέσεις για τους υπολογιστικούς μηχανισμούς του συστήματος. Έτσι, ο ερευνητής που μελετά τον αμφιβληστροειδή δε χρειάζεται πλέον να σχεδιάζει ειδικά ερεθίσματα και μοντέλα για να προκαλέσει συγκεκριμένες κυτταρικές συμπεριφορές. Εκπαιδευοντας ένα σύνθετο νευρωνικό δίκτυο σε φυσικές εικόνες μπορεί να μελετήσει τη λειτουργία και τη συνεισφορά των διαφορετικών μη γραμμικών κυκλωμάτων του συστήματος ταυτόχρονα, όπως και τη συνεισφορά τους στη συνολική διαδικασία επεξεργασίας της εικόνας [55], [58], [59].

Το συνελκτικό νευρωνικό δίκτυο που επιλέξαμε να χρησιμοποιήσουμε είναι δουλειά των Niru Maheswaranathan, Lane McIntosh και της ερευνητικής τους ομάδας στο Πανεπιστήμιο Stanford [58]. Όπως φαίνεται και στην Εικόνα 34, αποτελεί ένα δίκτυο τριών επιπέδων, το οποίο έρχεται σε αντιστοιχία με το ότι ο αμφιβληστροειδής των σπονδυλωτών οργανισμών (άνθρωπος, πρωτεύοντα, ποντίκια, κτλ) περιλαμβάνει τρεις βασικές στρώσεις κυττάρων (φωτοϋποδοχείς, δίπολα κύτταρα, γαγγλιακά κύτταρα). Αυτό βοηθά στο να μπορεί κάθε επίπεδο να μελετηθεί ξεχωριστά και παράλληλα με το βιολογικό του ανάλογο. Οι λόγοι που επιλέχθηκε το συγκεκριμένο μοντέλο είναι οι εξής:

- Προβλέπει αποκρίσεις σε φυσικές εικόνες ταυτόχρονα για πολλαπλούς νευρώνες με μεγάλη ακρίβεια. Σύμφωνα με τους συγγραφείς, ο συντελεστής συσχέτισης σε εικόνες που το μοντέλο δεν έχει δει κατά την εκπαίδευση φτάνει στο 0.7.
- Είναι αρκετά απλό σε δομή και αρχιτεκτονική και συνεπώς πιο κατανοητό.
- Εκπαιδευμένο σε φυσικές εικόνες αναπαράγει ένα ευρύ φάσμα βιολογικών φαινομένων που συναντώνται στον αμφιβληστροειδή όπως η ταχεία προσαρμογή στις αντιθέσεις φωτεινότητας [14] και η αντιστροφή πολικότητας του υποδεκτικού πεδίου [60], αποκαλύπτοντας έτσι τους πολύπλοκους μηχανισμούς που εφαρμόζει ένα γαγγλιακό κύτταρο κατά την αποκωδικοποίηση φυσικών εικόνων [59].
- Μπορεί και παρουσιάζει ικανοποιητικά αποτελέσματα για τεχνητά κατασκευασμένες εικόνες ενώ έχει εκπαιδευτεί σε φυσικές εικόνες. Αυτό μας βοήθησε στο επόμενο στάδιο της εργασίας μας όταν επιχειρήσαμε να υπολογίσουμε τα υποδεκτικά πεδία των κυττάρων.



**Σχήμα 34:** Αρχιτεκτονική του συνελκτικού νευρωνικού δικτύου [55].

Δεν τροποποιήσαμε καθόλου την αρχιτεκτονική των επιπέδων του δικτύου. Το πρώτο επίπεδο αποτελείται από ένα σύνολο φίλτρων συνέλιξης που επιχειρούν να εντοπίσουν παράλληλα και μέσω γραμμικών υπολογισμών ένα πλήθος διαφορετικών χαρακτηριστικών από τις εικόνες εισόδου στα πεδία του χώρου και του χρόνου. Κάθε πυρήνας συνέλιξης ελέγχει όλα τα κανάλια εισόδου, 40 στο πλήθος, εφαρμόζοντας διαδοχικές συνέλιξεις μέσω πολλαπλασιασμού πινάκων μεταξύ των συντελεστών του, που προσαρμόζονται κατά την εκπαίδευση, και των τιμών της εικόνας. Αποτέλεσμα είναι ένα σύνολο φιλτραρισμένων εκδοχών της εισόδου. Στη συνέχεια εφαρμόζεται η τεχνική του Batch Normalization [61] που επιταχύνει τη σύγκλιση και σταθεροποιεί την εκπαίδευση, επιλύοντας το πρόβλημα της εσωτερικής μετατόπισης της συνδιακύμανσης (internal covariance shift), δηλαδή τον κορεσμό του ρυθμού μάθησης που οφείλεται στις συνεχείς αλλαγές της κατανομής των εισόδων στα επίπεδα του δικτύου κατά τη διάρκεια της εκπαίδευσης. Τέλος, η έξοδος του επιπέδου περνά από μια μη γραμμική συνάρτηση ενεργοποίησης, το γραμμικό ανορθωτή (ReLU).

Το δεύτερο επίπεδο διατηρεί την αρχιτεκτονική του πρώτου, με μόνη διαφορά το πλήθος και το μέγεθος των συνελκτικών φίλτρων. Στο τελευταίο στάδιο, το οποίο προσομοιώνει τα γαγγλιακά κύτταρα έχουμε ένα επίπεδο, οι νευρώνες του οποίου λαμβάνουν είσοδο από όλους τους νευρώνες του προηγούμενου επιπέδου (Fully-Connected Layer) και αντικαθιστά τα συνελκτικά φίλτρα. Επίσης αλλάζει και η συνάρτηση ενεργοποίησης, αφού έχει επιλεγεί ο εξομαλυμένος γραμμικός ανορθωτής (Softplus). Τελικά το αποτέλεσμα είναι να έχουμε 60 νευρώνες στην έξοδο του μοντέλου, που ο καθένας παράγει μια εκτίμηση για την εικόνα στην είσοδο.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 50, 50, 40)	0
conv2d_1 (Conv2D)	(None, 41, 41, 8)	32008
flatten_1 (Flatten)	(None, 13448)	0
batch_normalization_1 (Batch Normalization)	(None, 13448)	53792
reshape_1 (Reshape)	(None, 41, 41, 8)	0
gaussian_noise_1 (GaussianNoise)	(None, 41, 41, 8)	0
activation_1 (Activation)	(None, 41, 41, 8)	0
conv2d_2 (Conv2D)	(None, 37, 37, 16)	3216
flatten_2 (Flatten)	(None, 21904)	0
batch_normalization_2 (Batch Normalization)	(None, 21904)	87616
reshape_2 (Reshape)	(None, 37, 37, 16)	0
gaussian_noise_2 (GaussianNoise)	(None, 37, 37, 16)	0
activation_2 (Activation)	(None, 37, 37, 16)	0
flatten_3 (Flatten)	(None, 21904)	0
dense_1 (Dense)	(None, 60)	1314240
batch_normalization_3 (Batch Normalization)	(None, 60)	240
activation_3 (Activation)	(None, 60)	0

Total params: 1,491,112  
 Trainable params: 1,420,288  
 Non-trainable params: 70,824

**Σχήμα 35:** Σύνοψη των σταδίων επεξεργασίας των δειγμάτων από το δίκτυο.

Εκπαιδύσαμε το συνελκτικό νευρωνικό δίκτυο με όλες τις διαθέσιμες εικόνες και για όλους τους νευρώνες ταυτόχρονα. Οι παράμετροι, δηλαδή τα βάρη του δικτύου, βελτιστοποιήθηκαν προσπαθώντας να ελαχιστοποιήσουν τη συνάρτηση απώλειας Poisson,

$$L(y_t, \hat{y}_t) = \frac{1}{N} \sum_{t=0}^N (\hat{y}_t - y_t \log(\hat{y}_t)), \quad (2)$$

όπου  $y_t$  είναι η πραγματική απόκριση και  $\hat{y}_t$  η πρόβλεψη στο χρονικό σημείο  $t$ , αντίστοιχα ενώ  $N$  είναι το πλήθος των δειγμάτων εκπαίδευσης ανά ενημέρωση της συνάρτησης απώλειας (batch size). Η στοχαστική διεργασία Poisson χρησιμοποιείται εκτενώς για να μοντελοποιηθεί νευρωνική δραστηριότητα σε μορφή δυναμικών ενεργείας, υπολογίζοντας την πιθανότητα πυροδότησης ενός νευρώνα σε ένα πολύ μικρό χρονικό διάστημα  $[t, t + \Delta t]$  από τη σχέση

$$R * \Delta t, \quad (3)$$



με  $R$  τον ρυθμό πυροδότησης. Αποτελεί το απλούστερο στοχαστικό εργαλείο για τέτοιου είδους αναλύσεις καθώς υποθέτει στατιστική ανεξαρτησία μεταξύ των μεμονωμένων γεγονότων της ακολουθίας. Τα δεδομένα προς πρόβλεψη αφορούν πλήθος και ρυθμό εμφάνισης ενός γεγονότος μέσα σε ένα ορισμένο χρονικό παράθυρο και συνεπώς η συγκεκριμένη επιλογή συνάρτησης απώλειας έχει μια ισχυρή βάση [8], [53].

Η βελτιστοποίηση πραγματοποιήθηκε με τη βοήθεια του αλγορίθμου Adam [62], μια παραλλαγή του κλασικού αλγορίθμου απότομης κατάβασης (gradient descent). Η υλοποίηση και η εκπαίδευση έγινε με τις βιβλιοθήκες μηχανικής μάθησης Tensorflow [63] και Keras [64] σε Python.

Χρησιμοποιήσαμε το 70% των δεδομένων για την εκμάθηση του δικτύου, με το υπόλοιπο 30% να προορίζεται για επαλήθευση και αξιολόγηση. Για τις υπερπαραμέτρους, ξεκινώντας από τις τιμές που είχε ήδη το μοντέλο, πειραματιστήκαμε σε μια προσπάθεια να βρούμε τον συνδυασμό τιμών που να οδηγεί στην υψηλότερη επίδοση στα δεδομένα επαλήθευσης, επιδιώκοντας με αυτό τον τρόπο την όσο το δυνατό μέγιστη δυνατότητα γενίκευσης. Υπερπαραμέτροι είναι οι παράμετροι που, σε αντίθεση με τα βάρη, οι τιμές τους δε μεταβάλλονται κατά την εκπαίδευση και ο ρόλος τους είναι να ελέγχουν και να ρυθμίζουν τη διαδικασία. Έτσι, για το ρυθμό μάθησης (learning rate) διατηρήσαμε την τιμή  $10^{-2}$  και για το batch size τα 5000 δείγματα (50 s διέγερσης). Καθώς η συγκεκριμένη επιλογή υπερβαίνει κατά πολύ τις τυπικές τιμές που χρησιμοποιούνται, δοκιμάσαμε μικρότερες τιμές και είδαμε σημαντική πτώση στη δυνατότητα του μοντέλου να μάθει χαρακτηριστικά της συμπεριφοράς των γαγγλιακών κυττάρων. Αναφορικά με το πλήθος και το μέγεθος των φίλτρων συνέλιξης, εξετάσαμε διαφορετικούς συνδυασμούς και συμβιβαστήκαμε με 8 φίλτρα στο πρώτο επίπεδο και 16 στο δεύτερο, μεγέθους  $10 \times 10$  και  $5 \times 5$  αντίστοιχα.

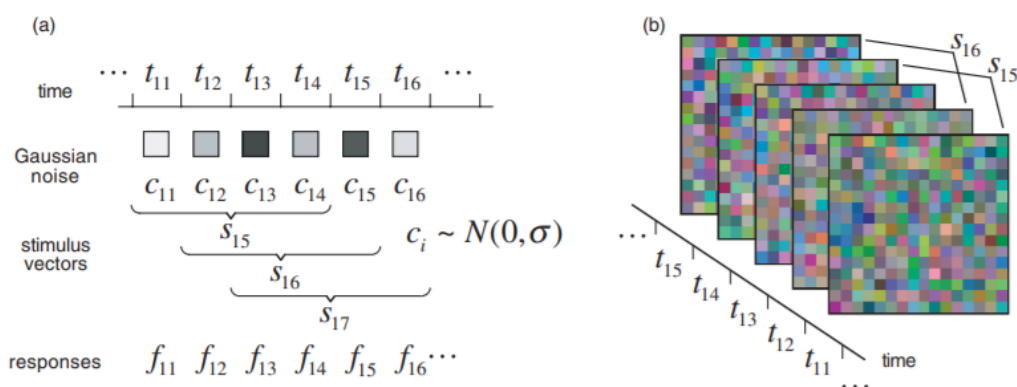
Για την παρακολούθηση της πορείας της επίδοσης χρησιμοποιήθηκαν, πλην της συνάρτησης απώλειας, τρεις μετρικές: ο συντελεστής συσχέτισης Pearson (Pearson correlation coefficient - cc) [65] που δίνει το βαθμό ικανότητας μιας γραμμικής συνάρτησης να περιγράψει τη σχέση μεταξύ πραγματικής απόκρισης και εκτίμησης, το ποσοστό της διακύμανσης της πραγματικής απόκρισης που η πρόβλεψη καταφέρνει να προβλέψει (Fraction of explained variance) [66] και η ρίζα του μέσου τετραγωνικού σφάλματος.

Ορίσαμε στο 500 το μέγιστο αριθμό εποχών κατά την εκπαίδευση. Μετά την ολοκλήρωσή της επιλεγόταν η εποχή με τη χαμηλότερη τιμή συνάρτησης απώλειας στα δεδομένα επαλήθευσης. Αν η επίδοση δε βελτιωνόταν για 10 εποχές ο ρυθμός μάθησης μειωνόταν κατά πέντε φορές και αν δε βελτιωνόταν για 20 εποχές η διαδικασία σταματούσε. Όσο αυξάναμε την τιμή της χωρικής συνιστώσας στην είσοδο τόσο περισσότερο χρόνο απαιτούσε το δίκτυο για να εκπαιδευτεί. Επίσης οι απαιτήσεις μνήμης ήταν αρκετά μεγάλες. Για όλες τις εικόνες και για τιμή χρόνου 40 χρειάστηκαν κατά μέσο όρο 3.5 - 4 ώρες για την εκπαίδευση.

### 3.5 Μέθοδος υπολογισμού υποδεκτικού πεδίου

Προκειμένου να υπολογιστούν τα υποδεκτικά πεδία των νευρώνων του πειράματος χρησιμοποιήθηκε η μέθοδος της αντίστροφης συσχέτισης (reverse correlation) ή, με την πιο δημοφιλή ονομασία, Spike-triggered Average (STA). Η μέθοδος αυτή προσδιορίζει μια γραμμική σχέση διέγερσης - απόκρισης. Προσπαθεί να εκτιμήσει την αιτία πρόκλησης ενός γεγονότος και πιο συγκεκριμένα, κοιτώντας πίσω στο χρόνο, τι είδους διέγερση κατά μέσο όρο οδήγησε ένα νευρώνα να πυροδοτήσει.

Η τεχνική STA είναι ένας αλγόριθμος δύο βημάτων. Το πρώτο βήμα είναι να παρουσιαστεί στον υπό μελέτη νευρώνα ένα διάνυσμα διέγερσης  $s_t$  με διάσταση  $k$ . Οι καταχωρήσεις στο διάνυσμα αυτό είναι τιμές έντασης συναρτήσεως του χρόνου αμέσως πριν το χρονικό σημείο  $t$  και του χώρου. Η διάρκεια του διανύσματος διέγερσης οφείλει να ξεπερνά τη μνήμη του νευρώνα ώστε να μη χαθούν φαινόμενα που μπορεί να επηρεάζουν την απόκρισή του. Έτσι, η απόκριση  $f_t$  του νευρώνα στο χρόνο  $t$  εξαρτάται μόνο από το  $s_t$ . Η  $f_t$  είναι συνήθως το πλήθος των δυναμικών ενεργείας που εντοπίζονται στο μικρό χρονικό παράθυρο  $t$  διάρκειας  $\Delta t$ . Για παράδειγμα, έστω ότι προβάλλεται μια ακολουθία μονοχρωματικών εικόνων, διαστάσεων  $50 \times 50$ , με μεταβλητή τιμή έντασης σε κάθε pixel, σε ρυθμό ανανέωσης 100 Hz, δηλαδή  $\Delta t = 10$  ms ανά εικόνα. Έστω επίσης πως η μνήμη του νευρώνα είναι 400 ms ( $40\Delta t$ ). Τότε κάθε διάνυσμα  $s_t$  έχει διάσταση  $k = 50 \times 50 \times 40 = 100,000$  καταχωρήσεις, με την καθεμία να είναι ένα pixel σε μια απόχρωση του γκρι [20].



**Σχήμα 36:** (a) Δείγματα μιας χωρικά ομοιόμορφης ακολουθίας μονοχρωματικών εικόνων λευκού θορύβου υπό κανονική κατανομή, για νευρώνα με μνήμη  $4\Delta t$ .  $c_t$  είναι η τιμή αντίθεσης της εικόνας και  $f_t$  η απόκριση, στο χρονικό σημείο  $t$ . Το διάνυσμα διέγερσης  $s_t$  έχει διάσταση  $k = 4$  και περιλαμβάνει τα δεδομένα που προβλήθηκαν στα χρονικά σημεία  $t - 1$  έως  $t - 4$ , δηλαδή κοιτάμε φαινόμενα αυστηρά εντός των τεσσάρων καρέ ακριβώς πριν την απόκριση. (b) Εναλλακτικά, οι εικόνες μπορεί να μην είναι χωρικά ομοιόμορφες ούτε μονοχρωματικές. Κάθε pixel θα έχει μια διαφορετική χρωματική απόχρωση στο χώρο και το χρόνο [20].

Η μέθοδος λειτουργεί βέλτιστα και υλοποιείται σχεδόν αποκλειστικά με στοχαστική είσοδο, και πιο συγκεκριμένα με λευκό γκαουσιανό θόρυβο. Κάθε διάνυσμα  $s_t$  δημιουργείται από  $k$  ανεξάρτητες μεταξύ τους λήψεις από μια κανονική κατανομή. Τότε, όλα τα  $s_t$  διαμορφώνουν μαζί μια πολυδιάστατη κανονική κατανομή που ικανοποιεί την υπόθεση της ακτινικής συμμετρίας, που σημαίνει πως οποιαδήποτε δύο διανύσματα διέγερσης ίδιου μήκους έχουν ίδια πιθανότητα να επιλεγθούν. Η ιδιότητα αυτή καθιστά τη διέγερση πλήρως στοχαστική ενώ πλέον οι καταχωρήσεις δεν είναι τιμές φωτεινότητας, δηλαδή μόνο θετικές, αλλά τιμές αντιθέσεων, δηλαδή αποκλίσεις από τη μέση ένταση. Επίσης, τα  $s_t$  παρουσιάζονται διαδοχικά στο χρόνο και χωρίς παύσεις, συνεπώς υπάρχει επικάλυψη μεταξύ τους και δεν είναι στατιστικά ανεξάρτητα [20].

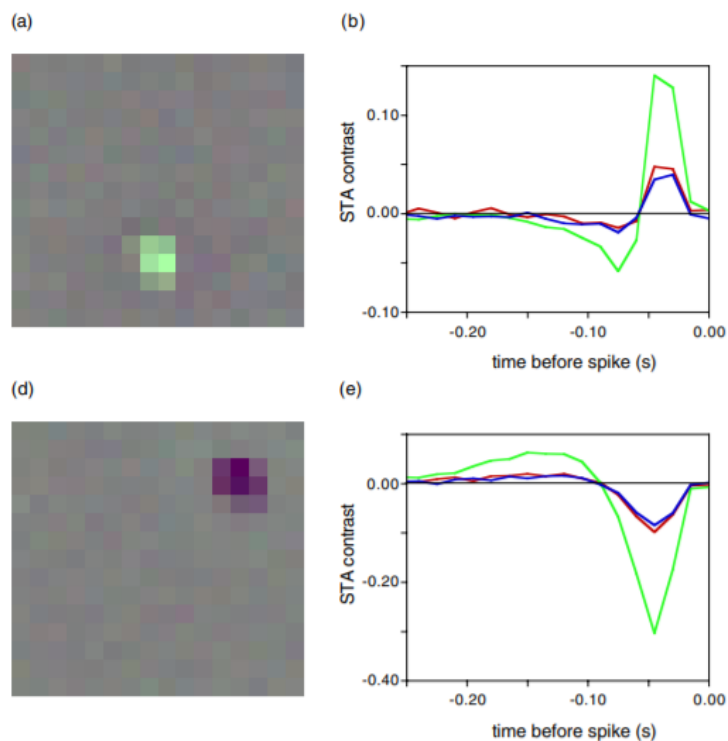
Το δεύτερο βήμα είναι ο υπολογισμός του STA μέσω του λόγου:

$$a = \frac{\sum_{t=1}^T s_t f_t}{\sum_{t=1}^T f_t}, \quad (4)$$

όπου  $T$  είναι η περίοδος καταγραφής του πειράματος. Το  $a$  θα είναι ένα διάνυσμα ίδιας διάστασης με το  $s_t$  που δείχνει τι ερέθισμα, κατά μέσο όρο, υποχρέωσε τον νευρώνα να εμφανίσει δραστηριότητα. Το κρίσιμο σημείο που καθιστά πολύ χρήσιμη τη συγκεκριμένη μέθοδο είναι πως το  $a$  είναι ευθέως ανάλογο στο γραμμικό φίλτρο που περιλαμβάνουν μοντέλα προσέγγισης όπως το LN που αναφέραμε. Αν αναλύσουμε το  $a$  σε τρεις συνιστώσες, μια χωρική, μια χρονική και μια χρωματική, μπορούμε να υπολογίσουμε, αντίστοιχα, το υποδεκτικό πεδίο του νευρώνα σε μορφή εικόνας, τη μέση μορφή διέγερσης που προηγήθηκε της απόκρισης και την επίδραση καθενός από τα τρία βασικά χρώματα: κόκκινο, πράσινο, μπλε. Με την STA, έχοντας απλώς τις αποκρίσεις και τίποτα άλλο, είναι πλέον εύκολη η εύρεση του γραμμικού φίλτρου  $F$  της (1) που περιέχει όλη την απαραίτητη πληροφορία για το πώς ο νευρώνας επεξεργάζεται την είσοδο. Απαραίτητη προϋπόθεση ωστόσο είναι η ύπαρξη ακτινικής συμμετρίας στην είσοδο, η οποία καθιστά την παραπάνω προσέγγιση αμερόληπτη. Σε περιπτώσεις που η διέγερση δεν είναι λευκός θόρυβος, δηλαδή εντοπίζονται μη μηδενικές συσχετίσεις στο χώρο και το χρόνο, η προσέγγιση είναι μακριά από το γραμμικό υποδεκτικό πεδίο [20].

Για να γίνει πιο κατανοητή η διαδικασία ακολουθεί ένα παράδειγμα. Στην Εικόνα 37 απεικονίζονται οι συνιστώσες του STA για ένα ON και ένα OFF γαγγλιακό κύτταρο από τον αμφιβληστροειδή ενός πρωτεύοντος. Για τον υπολογισμό χρησιμοποιήθηκε ακολουθία πολυχρωματικών εικόνων με τυχαία και ανεξάρτητη επιλογή τιμών pixels από κανονική κατανομή. Το STA είναι ένα διάνυσμα με τιμές απόκλισης από τη μέση τιμή εισόδου, για κάθε χρώμα και σε κάθε σημείο στο χώρο, σε μορφή χρονικής ακολουθίας πριν από την απόκριση. Στο σχήμα 37a φαίνεται το υποδεκτικό πεδίο του ON κυττάρου. Μπορούμε να καταλάβουμε πως αυτό το κύτταρο διεγείρεται, κατά μέσο όρο, μετά από αύξηση της φωτεινότητας στην (χρωματισμένη με πράσινο χρώμα) περιοχή της εικόνας που το STA έχει πολύ θετικές τιμές. Η περιοχή αυτή είναι και το κέντρο του υποδεκτικού πεδίου και, στις

περισσότερες περιπτώσεις, περιβάλλεται από μια ανταγωνιστική περιφέρεια. Μακριά από αυτή την περιοχή οι τιμές του STA τείνουν στο 0, υποδεικνύοντας πως μεταβολές μακριά από το κέντρο του πεδίου δεν επηρεάζουν τις αποκρίσεις του κυττάρου. Στο σχήμα 37b φαίνονται, ξεχωριστά για τις τρεις χρωματικές συνιστώσες, οι μεταβολές στην ένταση που κατά μέσο όρο οδήγησαν το κύτταρο να πυροδοτήσει, σε βάθος χρόνου έως 250 ms πριν από την πυροδότηση. Όσο πηγαίνουμε προς τα αριστερά μελετάμε φαινόμενα όλο και πιο πίσω στο χρόνο. Εδώ μπορούμε να διακρίνουμε μια διφασική κυματομορφή με πολύ μεγαλύτερη θετική συνιστώσα, μιας και το κύτταρο είναι ON, δηλαδή επιλεκτικό σε αύξηση της έντασης. Οι τιμές της κυματομορφής προκύπτουν μελετώντας ξεχωριστά σε κάθε χρονικό σημείο τι είδους μεταβολή παρουσιάστηκε στην ένταση και έπειτα υπολογίζοντας το μέσο όρο για όλες τις αποκρίσεις. Η σύγκλιση προς το 0 περίπου 200 ms πριν από την απόκριση μαρτυρά πως γεγονότα πριν από αυτό το σημείο δε σχετίζονται με τη συμπεριφορά του νευρώνα. Με αυτό τον τρόπο μπορούμε εμπειρικά να ορίσουμε τη μνήμη του νευρώνα. Να σημειωθεί πως η χρονική συνιστώσα του STA αφορά το κέντρο του υποδεκτικού πεδίου, δηλαδή το pixel με τη μέγιστη απόλυτη τιμή. Παρόμοια συμπεράσματα μπορούν να εξαχθούν και για το OFF κύτταρο (σχήματα 37d, 37e) [20].



**Σχήμα 37:** Χωρική και χρονική συνιστώσα STA για ένα ON (πάνω) και ένα OFF (κάτω) γαγγλιακό κύτταρο [20].

Λόγω της απλότητας και της αξιοπιστίας της, η STA είναι η τεχνική που χρησιμοποιείται περισσότερο στη μελέτη και χαρτογράφηση του υποδεκτικού πεδίου σε γαγγλιακά κύτταρα.

Μπορεί να υπολογίσει αποτελεσματικά και ολοκληρωμένα μια γραμμική προσέγγιση της συμπεριφοράς του κυττάρου και να βοηθήσει τον ερευνητή, μέσα από τη μορφή του χρονικού πυρήνα του STA να ταξινομήσει τα κύτταρα σε ON ή OFF. Για το λόγο αυτό είναι και η κυρίαρχη μέθοδος προσδιορισμού του γραμμικού φίλτρου στα LN μοντέλα. Επιπρόσθετα, η χωρική ανομοιομορφία μεταξύ γειτονικών pixels στο λευκό θόρυβο επιτρέπει τόσο την ξεχωριστή διέγερση του κέντρου και της περιφέρειας του πεδίου, προκειμένου να αναδειχθεί επαρκώς ο ανταγωνισμός των δύο περιοχών, όσο και την καταγραφή και τον εντοπισμό του υποδεκτικού πεδίου πολλαπλών κυττάρων ταυτόχρονα [20], [57], [67].

### 3.6 Μοντέλο υπολογισμού χάρτη προεξοχής σε εικόνες

Η επιλογή της υλοποίησης η οποία θα μας παρέχει τις προβλέψεις της οπτικής προσοχής έγινε αναζητώντας ένα δίκτυο το οποίο θα φτάνει σε υψηλές επιδόσεις στις μετρικές που αξιοποιούνται στο πρόβλημα και για ένα σύνολο από διαφορετικές βάσεις δεδομένων που χρησιμοποιούνται για την εκπαίδευση και αποτίμηση τέτοιων μοντέλων. Στόχος είναι κατά τη σύγκριση του ρυθμού ενεργοποίησης του αμφιβληστροειδούς με την προεξοχή μιας εικόνας, οι εκτιμήσεις να είναι και για τις δύο πλευρές όσο πλησιέστερα γίνεται στο πραγματικό. Μελετήσαμε ένα πλήθος από διαφορετικές προσεγγίσεις που στηρίζονται σε συνελκτικά νευρωνικά δίκτυα και καταλήξαμε στο Saliency Attentive Model ή SAM [42]. Η επιλογή έγινε για δύο λόγους. Αρχικά, η αρχιτεκτονική της συγκεκριμένης υλοποίησης παρουσιάζει ιδιαίτερο ενδιαφέρον. Στηρίζεται σε τέσσερις άξονες:

- Αξιοποιεί ένα Επαναλαμβανόμενο Νευρωνικό Δίκτυο για δυναμική βελτίωση του αποτελέσματος και πιο συγκεκριμένα την αρχιτεκτονική Long Short-Term Memory (LSTM) [43]. Τα LSTM αναπτύχθηκαν ώστε να αντιμετωπίσουν το πρόβλημα των εξαφανιζόμενων κλίσεων (vanishing gradient problem) το οποίο περιορίζει νευρωνικά δίκτυα πολλών επιπέδων. Τα Επαναλαμβανόμενα Νευρωνικά Δίκτυα χρησιμοποιούνται για επεξεργασία ακολουθιών δεδομένων με χρονικές εξαρτήσεις, όπως φυσική γλώσσα. Εδώ, αξιοποιείται αυτή η ιδιότητα των LSTM για επαναληπτική επεξεργασία χαρακτηριστικών στο χώρο, αντικαθιστώντας τα εσωτερικά γινόμενα με πράξεις συνέλιξης. Το δίκτυο δέχεται στην είσοδο ένα σύνολο από χαρακτηριστικά που έχουν εξαχθεί από το προηγούμενο στάδιο, εστιάζει στις περισσότερο προεξέχουσες περιοχές της εικόνας και επαναληπτικά βελτιώνει την πρόβλεψή του, υπολογίζοντας τελικά έναν ανανεωμένο χάρτη χαρακτηριστικών που προσεγγίζει σε κάθε χρονικό βήμα ολοένα και περισσότερο τον πραγματικό.
- Ενσωματώνει την τάση ενός παρατηρητή να εστιάζει πρωτίστως το βλέμμα του στο κέντρο όταν παρατηρεί μια εικόνα. Η μοντελοποίηση γίνεται μέσα από ένα σύνολο 16 δισδιάστατων γκαουσιανών συναρτήσεων που λειτουργούν σαν χάρτες και ενσωματώνονται στους ήδη υπολογισμένους χάρτες χαρακτηριστικών του

προαναφερθέντος επιπέδου. Το δίκτυο αφήνεται μόνο του να μάθει τη μέση τιμή και την απόκλιση κάθε συνάρτησης αποκλειστικά από τα δεδομένα, δίχως να βασίζεται σε υποθέσεις από βιολογικές μελέτες και μέσω της σχέσης

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\left(\frac{(x-\mu_x)^2}{2\sigma_x^2} + \frac{(y-\mu_y)^2}{2\sigma_y^2}\right)\right). \quad (5)$$

- Καταπολεμά την αναπόφευκτη μείωση των διαστάσεων της εισόδου που συμβαίνει σε υλοποιήσεις με βαθιά συνελκτικά δίκτυα τα οποία προβλέπουν την προεξοχή μέσω εξαγωγής χαρακτηριστικών στα πολυάριθμα διαδοχικά επίπεδα τους. Αντί για κλασικές, χρησιμοποιεί διεσταλμένες συνελίξεις, διατηρώντας ακέραια την κλίμακα στην οποία λειτουργεί το κάθε συνελκτικό επίπεδο και υψηλή την ανάλυση στην έξοδο, ενώ παράλληλα το πλήθος των παραμέτρων του δικτύου παραμένει σε χαμηλά επίπεδα. Μέσω των μεγάλων υποδεκτικών πεδίων των πυρήνων, βοηθά το μοντέλο να υπολογίσει την προεξοχή ενός σημείου λαμβάνοντας υπόψη τι συμβαίνει σε μια ευρεία περιοχή γύρω από αυτό, βελτιώνοντας περαιτέρω την ακρίβεια. Η τροποποίηση αυτή εφαρμόζεται στο ResNet-50 [68], ένα συνελκτικό νευρωνικό δίκτυο που αποτελείται από 50 επίπεδα που ομαδοποιούνται σε πέντε στρώματα, και ένα fully-connected επίπεδο στο τέλος. Το πρώτο στρώμα περιλαμβάνει ένα συνελκτικό και ένα max pooling επίπεδο ενώ τα υπόλοιπα είναι αμιγώς συνελκτικά. Διεσταλμένες συνελίξεις εφαρμόζονται στα δύο τελευταία στρώματα ενώ το fully-connected επίπεδο αγνοείται προκειμένου να μη χαθεί η χωρική πληροφορία και η τοπογραφική συνέπεια της εξόδου μιας και αποζητάμε την εξαγωγή ενός χάρτη σε μορφή δισδιάστατης εικόνας. Το ResNet-50 ανήκει στην οικογένεια των Υπολειμματικών Νευρωνικών Δικτύων (Residual Neural Networks). Στα συγκεκριμένα δίκτυα η διαφοροποίηση είναι πως επιπλέον με την τυπική επεξεργασία που γίνεται στο κάθε επίπεδο, η είσοδος ενός επιπέδου προστίθεται στην έξοδό του. Αυτές οι συνδέσεις εξαλείφουν το πρόβλημα της υποβάθμισης της απόδοσης που εντείνεται όσο αυξάνεται το βάθος ενός δικτύου.
- Χρησιμοποιεί ένα συνδυασμό από μετρικές ως συνάρτηση κόστους κατά την εκπαίδευση, λαμβάνοντας υπόψη διαφορετικούς παράγοντες ποιότητας που οι εκτιμήσεις πρέπει να βελτιστοποιήσουν. Το κόστος υπολογίζεται από τη γραμμική σχέση

$$L(y^{est}, y^{fix}, y^{den}) = \alpha L_1(y^{est}, y^{fix}) + \beta L_2(y^{est}, y^{den}) + \gamma L_3(y^{est}, y^{den}), \quad (6)$$

όπου  $y^{est}, y^{fix}, y^{den}$  είναι αντίστοιχα η εκτίμηση του χάρτη οπτικής προεξοχής, ο πραγματικός δυαδικός χάρτης εστιάσεων που περιέχει 1 στα σημεία που ο παρατηρητής εστίασε το βλέμμα του και 0 στα υπόλοιπα, και ο πραγματικός χάρτης κατανομής πιθανότητας που προκύπτει μέσα από εξομάλυνση του  $y^{fix}$  με μια δισδιάστατη γκαουσιανή συνάρτηση. Οι  $\alpha, \beta, \gamma$  είναι τρεις πραγματικοί συντελεστές

που σταθμίζουν το άθροισμα και ορίστηκαν από τους συγγραφείς στις τιμές -1, -2, 10 ενώ  $L_1, L_2, L_3$  είναι τρεις καθιερωμένες μετρικές αξιολόγησης πρόβλεψης προεξοχής, η Normalized Scanpath Saliency (NSS), ο γραμμικός συντελεστής συσχέτισης (CC) και η απόκλιση Kullback-Leibler (KL-Div).

Η NSS δημιουργήθηκε αποκλειστικά για μοντέλα οπτικής προσοχής. Υπολογίζεται ως η μέση κανονικοποιημένη προεξοχή στα σημεία εστίασης από τη σχέση

$$L_1(y^{est}, y^{fix}) = \frac{1}{N} \sum_i \frac{y_i^{est} - \mu(y^{est})}{\sigma(y^{est})} \cdot y_i^{fix}, \quad (7)$$

με  $i$  τον αύξοντα αριθμό του pixel και  $N = \sum_i y_i^{fix}$  ο συνολικός αριθμός των pixels που έχουν τιμή 1. Ο  $y^{est}$  κανονικοποιείται ώστε να έχει μέση τιμή 0 και τυπική απόκλιση 1.

Ο CC αντιμετωπίζει τα  $y^{est}, y^{den}$  σαν τυχαίες μεταβλητές και απλώς υπολογίζει τη γραμμική σχέση μεταξύ τους. Δίνεται από τη σχέση

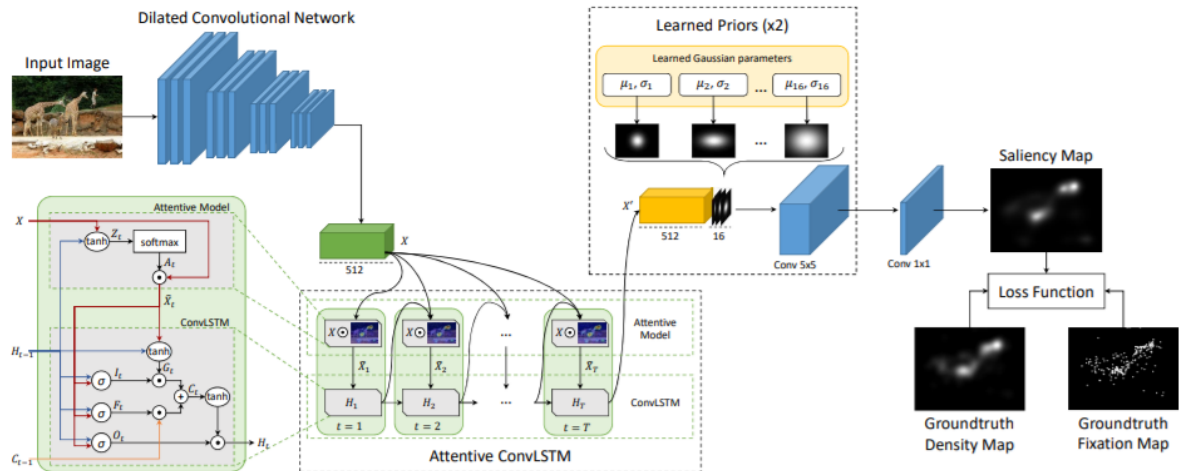
$$L_2(y^{est}, y^{den}) = \frac{\sigma(y^{est}, y^{den})}{\sigma(y^{est}) \cdot \sigma(y^{den})}, \quad (8)$$

με  $\sigma()$  τη συνδιακύμανση των δύο χαρτών.

Η απόκλιση Kullback-Leibler εκτιμά την απώλεια πληροφορίας όταν ο  $y^{den}$  προσεγγίζεται από τον  $y^{est}$ , συνεπώς αντιμετωπίζει με μια πιθανοτική ερμηνεία τους χάρτες. Υπολογίζεται από τη σχέση

$$L_3(y^{est}, y^{den}) = \sum_i y_i^{den} \log\left(\frac{y_i^{den}}{y_i^{est} + \epsilon} + \epsilon\right), \quad (9)$$

με  $i$  τον αύξοντα αριθμό του pixel και  $\epsilon$  μια σταθερά κανονικοποίησης. Είναι μετρική ανομοιομορφίας, συνεπώς επιδιώκεται η χαμηλότερη δυνατή τιμή και όχι η υψηλότερη.



**Σχήμα 38:** Σχηματικό διάγραμμα του μοντέλου πρόβλεψης προεξοχής SAM [42].

Ο δεύτερος και βασικός λόγος είναι πως με το συνδυασμό των παραπάνω στοιχείων επιδεικνύει εξαιρετικά αποτελέσματα στις μετρικές αξιολόγησης, ισοφαρίζοντας και στις περισσότερες περιπτώσεις ξεπερνώντας τις ήδη υπάρχουσες υλοποιήσεις. Σε ορισμένες περιπτώσεις το μοντέλο δύναται να αντιληφθεί και τη σχετική σημασία διαφορετικών παρατηρητών στην ίδια σκηνή, μια ανθρώπινη συμπεριφορά που τα μοντέλα προεξοχής εξακολουθούν να δυσκολεύονται να αναπαράγουν. Οι πίνακες των αποτελεσμάτων για τα MIT1003, MIT300, SALICON [69] και CAT2000, τα τέσσερα πιο κλασικά σύνολα δεδομένων καθώς και οι λεπτομέρειες της διαδικασίας εκμάθησης βρίσκονται στη [42]. Εμείς χρησιμοποιήσαμε τα βάρη ενός ήδη εκπαιδευμένου δικτύου στο SALICON, το μεγαλύτερο διαθέσιμο σύνολο εικόνων και η σύγκριση αργότερα στην εργασία έγινε στο Toronto [70], μια μικρή βάση 120 εικόνων. Οι μετρικές ήταν κι εδώ σε υψηλές τιμές και παρέχονται στη [71].

### 3.7 Συντελεστές συσχέτισης μεταξύ προεξοχής και νευρικής απόκρισης

Στο τελευταίο στάδιο της εργασίας χρησιμοποιήσαμε τρεις συντελεστές συσχέτισης για τη σύγκριση μεταξύ της απόκρισης και του μέτρου προεξοχής στις εικόνες. Δεδομένου πως δεν έχουμε κάποια υποψία για την κατανομή που ακολουθούν τα δεδομένα, χρησιμοποιήσαμε ένα συντελεστή για την αξιολόγηση γραμμικής σχέσης και δύο συντελεστές συσχέτισης κατάταξης.

Ο πρώτος συντελεστής είναι ο ευρύτερα χρησιμοποιούμενος συντελεστής συσχέτισης συνδιακύμανσης Pearson [65]. Δίνεται από τη σχέση



$$r_{xy} = \frac{\sum_{i=1}^N (x_i - m_x)(y_i - m_y)}{\sqrt{\sum_{i=1}^N (x_i - m_x)^2 \sum_{i=1}^N (y_i - m_y)^2}}, \quad (10)$$

όπου  $N$  το πλήθος των δειγμάτων ενός συνόλου,  $(x_i, y_i)$  το μεμονωμένο ζεύγος τιμών, στην περίπτωση μας ένα ζεύγος τιμών (απόκριση, προεξοχή) και  $m_x, m_y$  η μέση τιμή για τα δύο μεγέθη. Η μετρική αυτή δείχνει τη γραμμική σχέση μεταξύ δύο συνόλων δεδομένων. Αν απεικονίσουμε τα ζεύγη τιμών των δεδομένων ως σημεία σε ένα διάγραμμα με δύο άξονες, έναν για κάθε μέγεθος, ο  $r$  μας λέει κατά πόσο τα σημεία αυτά μπορούν να προσεγγιστούν από μια ευθεία γραμμή. Μπορεί να λάβει τιμές μεταξύ  $-1$  και  $1$ . Οι δύο ακραίες τιμές υπονοούν ότι τα δύο σύνολα μπορούν να περιγραφούν πλήρως από μια γραμμική συνάρτηση και όλα τα σημεία βρίσκονται πάνω στην ευθεία. Για  $r = 1$  το ένα μέγεθος αυξάνει όσο αυξάνει και το άλλο ενώ για  $r = -1$  οι πορείες των δύο μεγεθών ακολουθούν αυστηρά αντίστροφες κατευθύνσεις. Η τιμή  $0$  υποδηλώνει πως δεν εντοπίζεται κάποια γραμμική σχέση ανάμεσα στα σύνολα. Με άλλα λόγια, υψηλή απόλυτη τιμή του συντελεστή φανερώνει ισχυρή γραμμική συσχέτιση και χαμηλή απόλυτη τιμή φανερώνει ασθενή γραμμική συσχέτιση.

Ο δεύτερος συντελεστής ονομάζεται συντελεστής συσχέτισης κατάταξης του Spearman και συμβολίζεται με  $\rho$  [77]. Αξιολογεί το πόσο καλά μπορεί να περιγραφεί η σχέση μεταξύ των δύο μεγεθών ενός συνόλου παρατηρήσεων από μια μονότονη συνάρτηση. Διαφέρει από τον συντελεστή Pearson στο ότι δεν αναζητά αποκλειστικά γραμμική σχέση αλλά μονοτονία γενικότερα. Στην ουσία είναι ο συντελεστής Pearson μεταξύ του βαθμού των διακριτών σημείων μέσα στο σύνολο και όχι των τιμών τους και υπολογίζεται από τη (10) αφού γίνει η συγκεκριμένη μετατροπή. Κατά αντιστοιχία, τιμή  $1$  ή  $-1$  υποδηλώνει απόλυτα μονότονη σχέση και τιμή  $0$  το αντίθετο, ενώ το πρόσημο μαρτυρά την κατεύθυνση της μονοτονίας. Όπως φαίνεται, μπορεί να δώσει τιμή τέλει συσχέτισης για οποιαδήποτε μονότονη συνάρτηση, σε αντίθεση με τον Pearson που θα δώσει τη μέγιστη κατά απόλυτη τιμή μόνο σε γραμμική σχέση, συνεπώς είναι και λιγότερο ευαίσθητος σε μη κανονικές κατανομές.

Ο τρίτος συντελεστής ανήκει, όπως του Spearman, στους συντελεστές συσχέτισης κατάταξης και ονομάζεται Kendall [78]. Έστω ότι έχουμε δύο ζεύγη παρατηρήσεων  $(x_i, y_i)$  και  $(x_j, y_j)$  χωρίς καμία περίπτωση ισότητας τιμών στα δύο μεγέθη και  $i < j$ . Τα ζεύγη ορίζονται ως:

- σύμφωνα όταν ισχύει  $(x_i > x_j \text{ και } y_i > y_j)$  ή  $(x_i < x_j \text{ και } y_i < y_j)$ ,
- ασύμφωνα όταν ισχύει  $(x_i < x_j \text{ και } y_i > y_j)$  ή  $(x_i > x_j \text{ και } y_i < y_j)$ .

Ο συντελεστής Kendall συγκρίνει το πλήθος σύμφωνων και ασύμφωνων ζευγών, έστω  $P, Q$  αντίστοιχα. Υπολογίζεται από τη σχέση

$$\tau = \frac{P-Q}{\frac{n(n-1)}{2}}, \quad (11)$$

όπου ο παρονομαστής είναι ο διωνυμικός συντελεστής για το πλήθος των δυνατών τρόπων που μπορούμε να επιλέξουμε ένα μη διατεταγμένο υποσύνολο δύο στοιχείων μέσα από ένα σύνολο  $n$  στοιχείων, στην περίπτωση μας τα δυνατά ζεύγη παρατηρήσεων που μπορούμε να συγκρίνουμε. Ισχύει  $-1 \leq \tau \leq 1$ . Τιμή 1 ισοδυναμεί με πλήρη συμφωνία μεταξύ προεξοχής και απόκρισης, δηλαδή για όλα τα δυνατά ζευγάρια αύξηση της πρώτης συνοδεύεται από αύξηση της δεύτερης, και αντίστοιχα για τη μείωση. Τιμή -1 φανερώνει πλήρη διαφωνία και τιμή 0 ανεξαρτησία ανάμεσα στα παρατηρούμενα μεγέθη.

## Κεφάλαιο 4

### Αποτελέσματα - Συμπεράσματα

#### 4.1 Απεικόνιση υποδεκτικών πεδίων και αξιολόγηση του μοντέλου

Το πρώτο μέρος του πειράματος αφορά την επιλογή του δικτύου με το οποίο θα πορευτούμε στη συνέχεια. Η αξιολόγηση έγινε σε ένα πρώτο στάδιο μέσω των μετρικών, κυρίως για τις υπερπαραμέτρους του δικτύου και έπειτα, πιο ποιοτικά και εμπειρικά, μέσα από τα αποτελέσματα που μας έδωσε η μέθοδος STA. Όπως τονίσαμε και παραπάνω, το περισσότερο ενδιαφέρον εμφανίστηκε στη μνήμη (έστω *temporal\_filter*), δηλαδή το χρονικό διάστημα πριν από μια πυροδότηση στο οποίο μελετάμε τι είδους αλλαγές, κατά μέσο όρο, οδήγησαν το νευρώνα στο να διεγερθεί. Στον Πίνακα 1 παραθέτουμε το συντελεστή συσχέτισης για τρεις τιμές της μεταβλητής *temporal\_filter* σε φυσικές εικόνες από το σύνολο εκπαίδευσης και αποτίμησης.

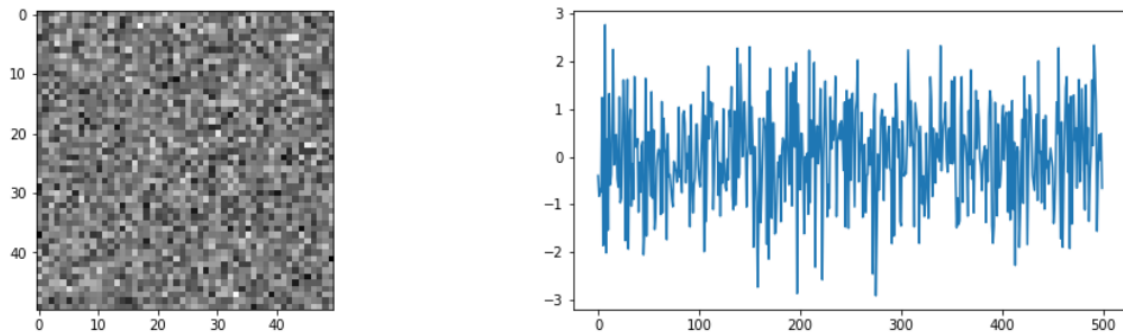
**Πίνακας 1:** Επίδοση του μοντέλου για διαφορετικές τιμές στη χρονική διάσταση των δειγμάτων.

Temporal Filter	CC	
	Train	Test
200 ms	0.874	0.626
300 ms	0.901	0.622
400 ms	0.618	0.364

Αρχικά, παρατηρούμε πως για χρονική διάσταση δείγματος 20 και 30 καρέ το δίκτυο φτάνει σε αρκετά υψηλά και κοντινά επίπεδα απόδοσης. Σε δεδομένα αξιολόγησης εμφανίζει, κατά μέσο όρο για τους 60 νευρώνες,  $cc = 0.626$  και  $cc = 0.622$  αντίστοιχα. Εκπαιδεύεται αποτελεσματικά σε φυσικές εικόνες και χρησιμοποιεί τις γνώσεις του προκειμένου να κάνει καλές εκτιμήσεις. Αυτό ωστόσο δε σημαίνει πως το μοντέλο με τις υψηλότερες επιδόσεις έχει κατανοήσει και μπορεί να αναπαράγει τα βιολογικά φαινόμενα πίσω από τα δεδομένα που επεξεργάζεται, για αυτό και δεν επιλέγουμε κατευθείαν την τιμή 200 ms. Το μοντέλο εφαρμόζει μετασχηματισμούς και εξάγει χαρακτηριστικά χωρίς να γνωρίζει ότι τα δεδομένα αυτά είναι κυτταρικές αποκρίσεις. Για διάρκεια 400 ms φαίνεται σημαντική πτώση στην επίδοση. Ο συντελεστής συσχέτισης δεδομένων εκπαίδευσης έφτασε

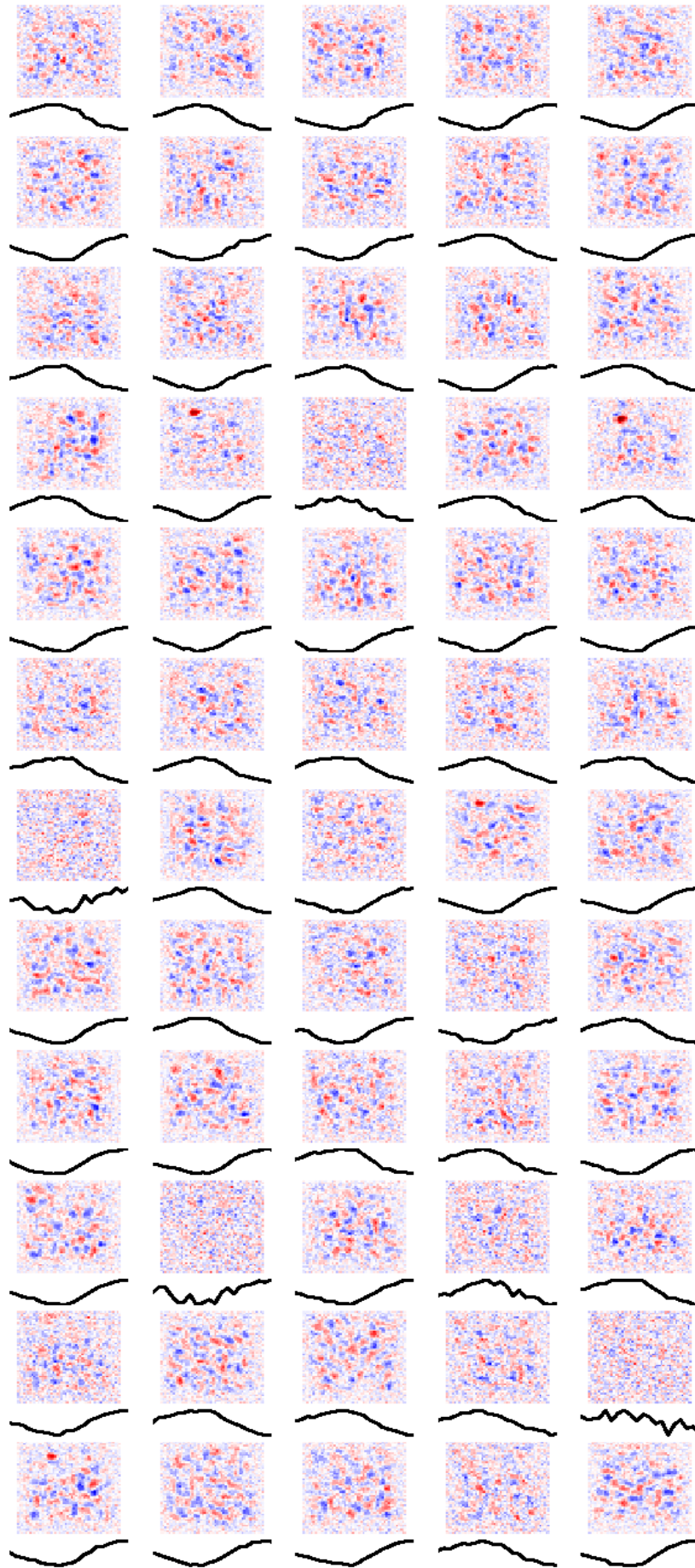
μέχρι την τιμή 0.712, ωστόσο η συνάρτηση απώλειας κατά την επαλήθευση εμφάνιζε άνοδο, αποτέλεσμα πιθανότατα κάποιας υπερπροσαρμογής στο σύνολο εκπαίδευσης.

Προκειμένου να επιλέξουμε την καταλληλότερη τιμή για τη μνήμη των νευρώνων, προχωρήσαμε στον υπολογισμό και την απεικόνιση των υποδεκτικών πεδίων (RFs) μέσω STA. Το σκεπτικό μας ήταν πως αν, για μια τιμή μνήμης, τα RFs πλησιάζουν ποιοτικά τη μορφή που αναμένουμε να έχουν σύμφωνα με τη θεωρία, τότε αυτό αποτελεί μια καλή ένδειξη πως η τιμή αυτή είναι κοντά στην πραγματική και συνεπώς το μοντέλο έχει μάθει πιο ικανοποιητικά τη βιολογική σχέση διέγερσης-απόκρισης. Το αναμενόμενο αποτέλεσμα για τους νευρώνες είναι αυτό που φαίνεται και στην Εικόνα 37: η χωρική συνιστώσα περιέχει μια οπτικά ξεκάθαρη και προεξέχουσα περιοχή ενδιαφέροντος που θα είναι το κέντρο του πεδίου, περιβαλλόμενη από μια ανταγωνιστική περιφέρεια και η χρονική συνιστώσα μια διφασική ασύμμετρη κυματομορφή που θα υποδεικνύει και τον τύπο του νευρώνα (ON ή OFF).



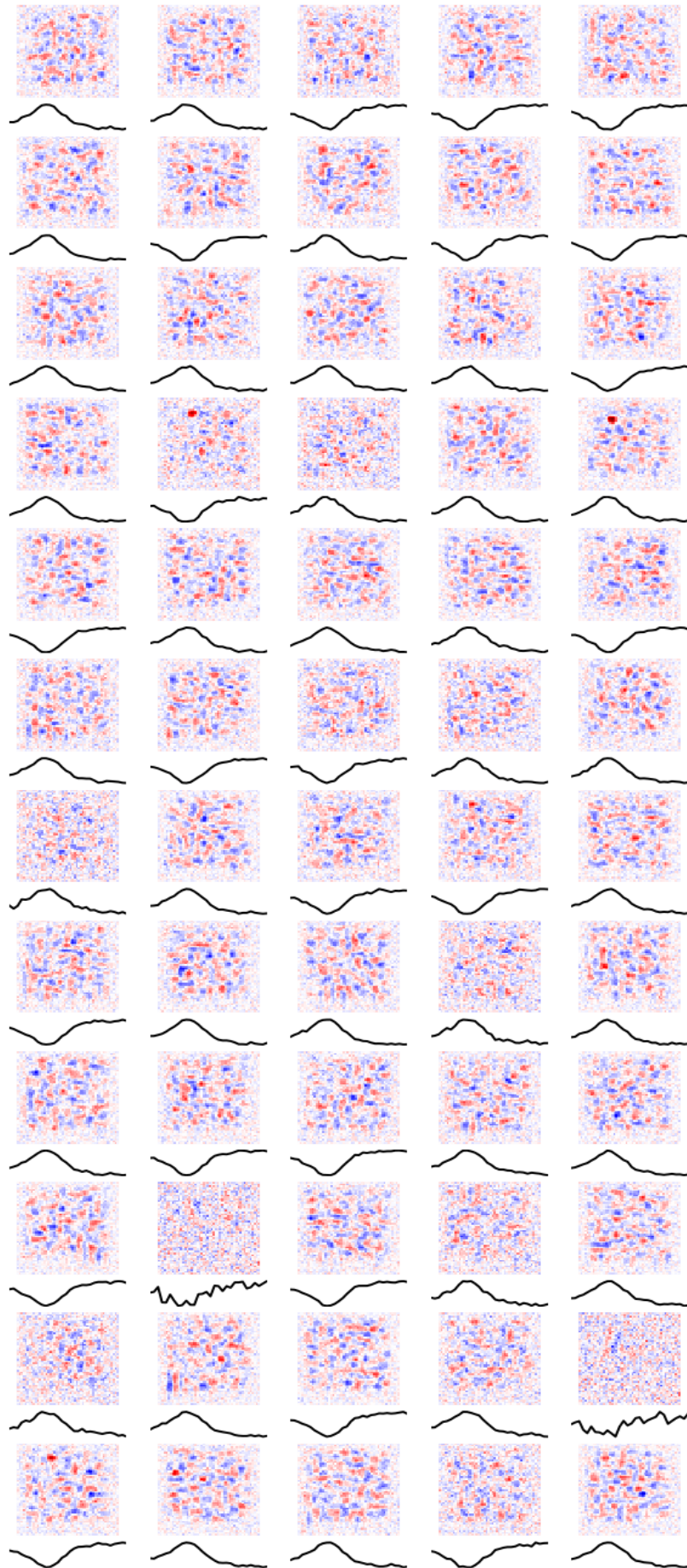
**Σχήμα 39:** Λευκός θόρυβος υπό τυποποιημένη κανονική κατανομή.

Χρησιμοποιήσαμε τις προβλέψεις που υπολογίζει το νευρωνικό δίκτυο σε χωρικά ανομοιομορφο λευκό θόρυβο με δείγματα από τυποποιημένη κανονική κατανομή ( $N(0,1)$ ). Εδώ φάνηκε και στην πράξη η ικανότητα του μοντέλου να γενικεύει και σε εικόνες που δεν είναι φυσικές. Δεν ενδιαφερόμαστε για τις χρωματικές επιδράσεις οπότε χρησιμοποιήσαμε μονοχρωματική είσοδο. Φορτώσαμε τα βάρη και την αρχιτεκτονική του επιλεγμένου μοντέλου, δημιουργήσαμε λευκό γκαουσιανό θόρυβο στην είσοδο (Εικόνα 39) και υπολογίσαμε το STA, διάστασης (40, 50, 50) για κάθε νευρώνα, ταυτόχρονα για όλους τους νευρώνες. Έπειτα, με τη βοήθεια της βιβλιοθήκης ανοικτού κώδικα pyret [72] για ανάλυση, επεξεργασία και απεικόνιση πειραματικών δεδομένων αμφιβληστροειδούς, προχωρήσαμε στη διάσπαση του STA σε ένα χρονικό και χωρικό πυρήνα, διάστασης 40 και (50, 50) αντίστοιχα, πάλι για κάθε νευρώνα. Έπειτα και από μια εξομάλυνση με ένα πολυδιάστατο γκαουσιανό φίλτρο προέκυψαν τα αποτελέσματα που φαίνονται στις Εικόνες 40, 41, 42 για τις τρεις τιμές του χρονικού φίλτρου, αντίστοιχα.

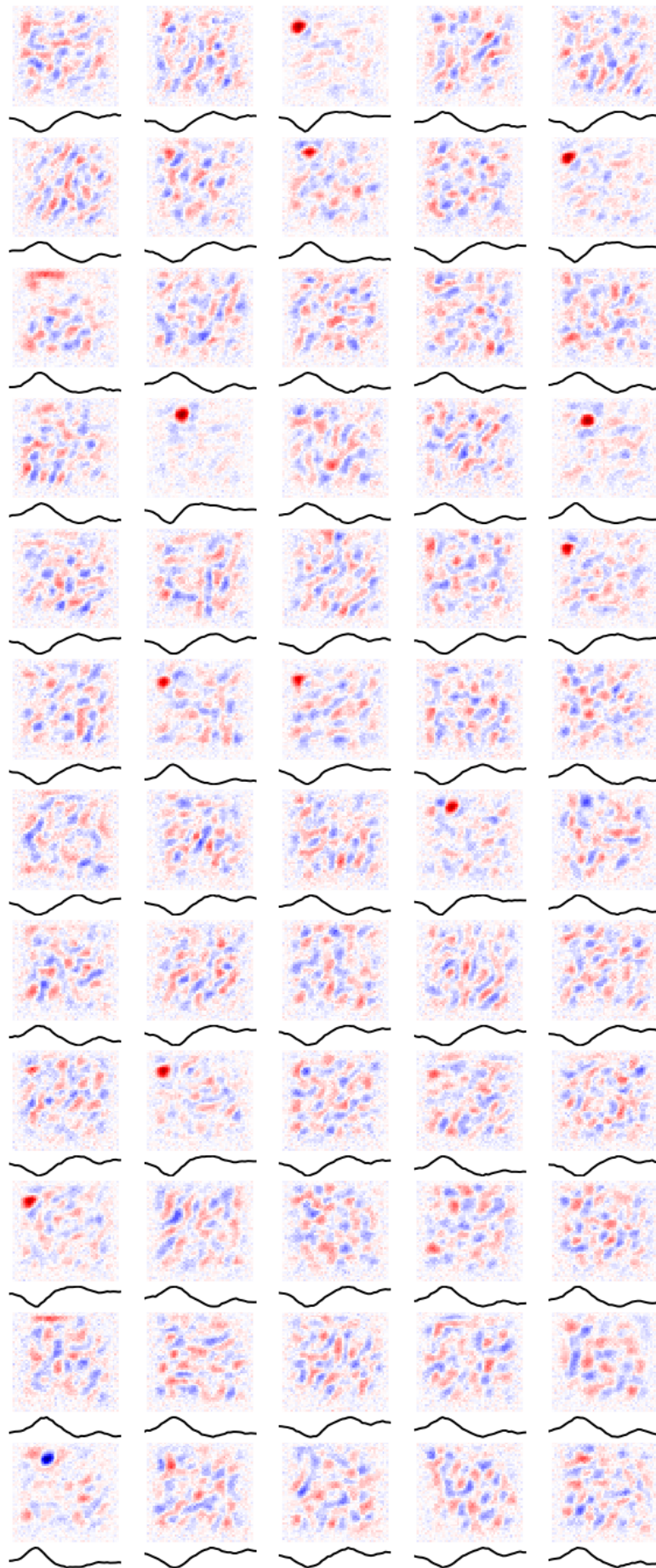


Σχήμα 40: Υποδεκτικά πεδία για τους 60 νευρώνες για χρονικό φίλτρο 200 ms.





Σχήμα 41: Υποδεκτικά πεδία για τους 60 νευρώνες για χρονικό φίλτρο 300 ms.



Σχήμα 42: Υποδεκτικά πεδία για τους 60 νευρώνες για χρονικό φίλτρο 400 ms. 75

Μπορούμε αρχικά να παρατηρήσουμε πως στις δύο πρώτες περιπτώσεις δεν είναι ορατή, για κανένα γαγγλιακό κύτταρο, κάποια περιοχή της εικόνας που να παρομοιάζει με το κέντρο ενός υποδεκτικού πεδίου. Οι νευρώνες, σύμφωνα με τη σχέση διέγερσης-απόκρισης που υπολογίστηκε, δεν έδειξαν προτίμηση σε κάποιο συγκεκριμένο σημείο ενδιαφέροντος, όπως προστάζουν τα βιολογικά τους χαρακτηριστικά και βλέπουμε πως η απόκριση είναι συσχετισμένη εξίσου με αλλαγές σε όλη την έκταση του πεδίου στο χώρο. Όσον αφορά τις χρονικές συνιστώσες, βλέπουμε πως τα αποτελέσματα του πρώτου πειράματος δε μοιάζουν ποιοτικά με την προσδοκώμενη συνάρτηση του χρονικού STA και δε θυμίζουν διφασική κυματομορφή, φαίνονται ίσως σαν να μην είναι ολοκληρωμένες. Σε λίγες περιπτώσεις επίσης ο χρονικός πυρήνας του STA είναι απλώς θόρυβος. Στο δεύτερο πείραμα οι κυματομορφές φαίνονται πιο κοντά στο αναμενόμενο αφού τείνουν να διαμορφώσουν μια πιο προφανή προτίμηση προς κάποια πολικότητα. Το νευρωνικό μοντέλο κατά την εκπαίδευση έλαβε ως πληροφορία πως η μνήμη των νευρώνων εκτείνεται μέχρι 300 ms πίσω στο χρόνο και προχωρά σε προβλέψεις υποθέτοντας ακριβώς αυτό, δηλαδή ότι όλες οι αλλαγές που διαμορφώνουν τη συμπεριφορά των κυττάρων εντοπίζονται εντός αυτού του διαστήματος. Το STA ωστόσο μαρτυρά πως υπάρχουν φαινόμενα που επηρεάζουν την κυτταρική δραστηριότητα σε χρονικό βάθος μεγαλύτερο και από τα 300 ms πριν από μια πυροδότηση. Συμπεραίνουμε λοιπόν πως οι υψηλές τιμές στις μετρικές δεν οδήγησαν και στα προσδοκώμενα υποδεκτικά πεδία και πως η τιμή μνήμης ίσως δεν είναι αρκετά μεγάλη.

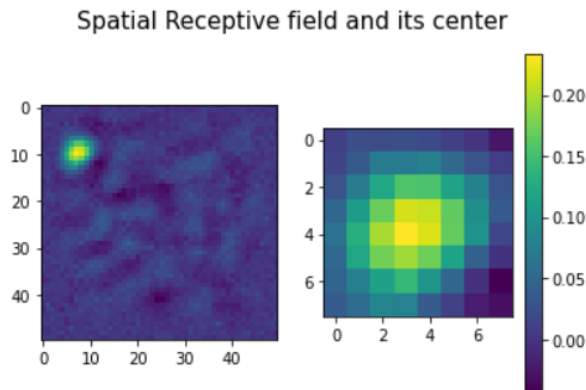
Πράγματι, στο πείραμα με τιμή χρονικού φίλτρου 0.4 s σημειώνεται σημαντική βελτίωση. Για ορισμένα γαγγλιακά κύτταρα έχει διαμορφωθεί μια πιο έντονη χρωματικά, ευκρινής περιοχή ενδιαφέροντος, συγκεκριμένης πολικότητας και σε σχήμα δίσκου, η οποία απουσίαζε στις δύο προηγούμενες απεικονίσεις. Παράλληλα στα ίδια κύτταρα, αλλού σε μεγαλύτερο και αλλού σε μικρότερο βαθμό, παρατηρείται μείωση της συσχέτισης μεταξύ διέγερσης και απόκρισης σε περιοχές της εικόνας μακριά από το κέντρο του υποδεκτικού πεδίου. Οι χρονικοί πυρήνες μοιάζουν με αυτούς στα 300 ms αλλά με πιο έντονα διαμορφωμένη διφασική μορφή.

Είναι φανερό από τις απεικονίσεις πως οι νευρώνες είναι ευαίσθητοι και σε αλλαγές στην είσοδο οι οποίες πραγματοποιούνται σε χρόνο μεγαλύτερο των 300 ms πίσω από την απόκριση και θεωρούμε πως η επιλογή 400 ms ίσως βρίσκεται πλησιέστερα στη βιολογική τιμή μνήμης. Λόγω της εμπειρικής προσέγγισης του ζητήματος, θεωρούμε πως αξίζει και η μελέτη μεγαλύτερων τιμών, πάνω από 400 ms. Αυτό ωστόσο, για ένα επαρκές μεγάλο σύνολο δεδομένων εκπαίδευσης απαιτεί αυξημένους υπολογιστικούς πόρους μνήμης, και για αυτό αποφασίσαμε να μη μελετήσουμε περαιτέρω, αφού και η τιμή φίλτρου 40 έδινε αποδεκτά αποτελέσματα στα πλαίσια της εργασίας μας.

Το συγκεκριμένο μοντέλο λοιπόν, παρά το γεγονός πως δεν υπερισχύει σε απόδοση σε σχέση με τα άλλα, δείχνει να έχει μάθει πιο ικανοποιητικά και ολοκληρωμένα τη βιολογική συμπεριφορά των νευρώνων, τουλάχιστον στο βασικό επίπεδο που αφορά τον προσδιορισμό του υποδεκτικού πεδίου. Επομένως, μιας και ενδιαφερόμαστε για: (α) τον εντοπισμό και τον ποιοτικό σχολιασμό των πεδίων και (β) μια πρώτη προσπάθεια συσχέτισης μεταξύ απόκρισης του αμφιβληστροειδούς και προεξοχής, θεωρήσαμε καταλληλότερη την επιλογή αυτής της υλοποίησης. Προσπαθήσαμε, για χρονικό φίλτρο 400 ms, να πειραματιστούμε με



τις υπόλοιπες υπερπαραμέτρους του δικτύου, όπως το μέγεθος και το πλήθος των συνελκτικών φίλτρων και το batch size, με στόχο να έχουμε ένα μοντέλο που οι εκτιμήσεις του υπολογίζουν σωστά τα υποδεκτικά πεδία και ταυτόχρονα επιδεικνύει και υψηλές επιδόσεις, ωστόσο το αποτέλεσμα του Πίνακα 1 ήταν το βέλτιστο που είδαμε.



**Σχήμα 43:** Παράδειγμα υποδεκτικού πεδίου από το μοντέλο μας, μαζί με το κέντρο του.

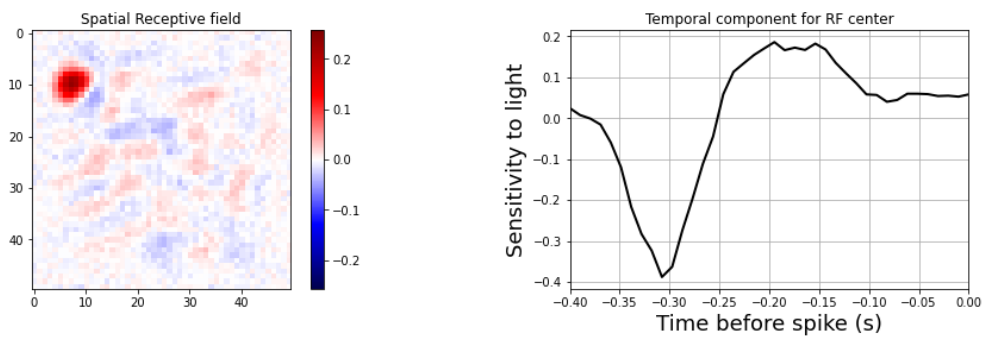
## 4.2 Επιλογή έγκυρων νευρώνων

Ένα στοιχείο που διακρίνουμε είναι πως δεν έχουν όλοι οι νευρώνες παρατηρήσιμο υποδεκτικό πεδίο. Το STA έχει καταφέρει να υπολογίσει το πεδίο για συνολικά 12 από τους 60 νευρώνες, αυτούς με αύξοντα αριθμό 2, 7, 9, 16, 19, 24, 26, 27, 33, 41, 45, 55. Οι υπόλοιποι νευρώνες δε σημείωσαν χαμηλότερες επιδόσεις στις μετρικές, συνεπώς το παραπάνω φαινόμενο δεν οφείλεται στο μοντέλο. Θεωρούμε πως συνδέεται με αυτό που αναφέραμε στην Ενότητα 3.3 σχετικά με την απόρριψη νευρώνων κατά το πείραμα καταγραφής. Πιθανότατα οι φυσικές εικόνες δεν εστιάστηκαν ακριβώς στο πεδίο ενδιαφέροντος των υπόλοιπων νευρώνων, κι έτσι διεγέρθηκαν ελάχιστα από αυτές. Το υποδεκτικό πεδίο αυτών των νευρώνων βρίσκεται εκτός της εικόνας και το σήμα που λαμβάνουν τα μικροηλεκτρόδια δεν είναι αντιπροσωπευτικό και χρήσιμο για μελέτη. Επομένως, απορρίπτουμε αυτούς τους νευρώνες και κρατάμε τους υπόλοιπους 12. Δοκιμάσαμε να εκπαιδεύσουμε ακριβώς το ίδιο δίκτυο, με 12 εξόδους αντί για 60, μία για κάθε έγκυρο νευρώνα. Η απόδοση δε βελτιώθηκε, αντιθέτως ήταν ελαφρώς κατώτερη, συνεπώς κρατήσαμε το μοντέλο με τους 60.

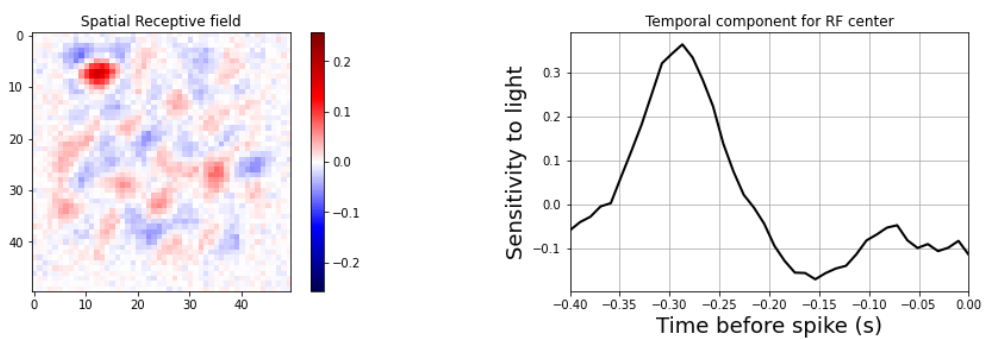
Σαν επόμενο βήμα, πάμε να σχολιάσουμε πιο αναλυτικά τη μορφή των πεδίων για τους έγκυρους νευρώνες. Με μια πρώτη ματιά φαίνεται πως το μοντέλο μας, παρά το γεγονός πως δεν επιτυγχάνει πολύ υψηλά επίπεδα απόδοσης σε φυσικές εικόνες, προβλέπει αποκρίσεις σε λευκό θόρυβο οι οποίες διαμορφώνουν αρκετά ικανοποιητικά τα υποδεκτικά πεδία των γαγγλιακών κυττάρων. Έχοντας εκπαιδευτεί σε ένα είδος διέγερσης αρκετά απαιτητικό, με ιδιομορφίες που μπορούν να αναδείξουν τους πολύπλοκους μετασχηματισμούς του αμφιβληστροειδούς, φαίνεται πως του είναι εύκολο να προχωρήσει σε ακριβείς εκτιμήσεις

στο θόρυβο. Στην Εικόνα 44 φαίνεται για τους νευρώνες αυτούς το μέσο χωρικό υποδεκτικό πεδίο καθώς και η χρονική πορεία της ευαισθησίας στο φως για το κεντρικό ριχέλ του πεδίου, όπως υπολογίστηκαν από την STA.

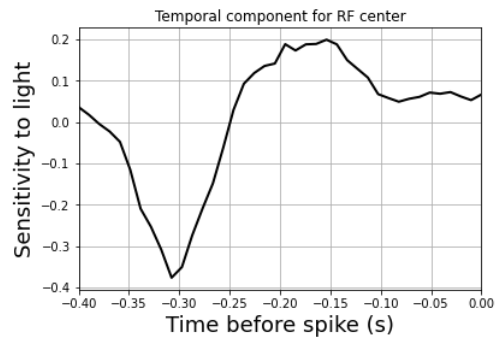
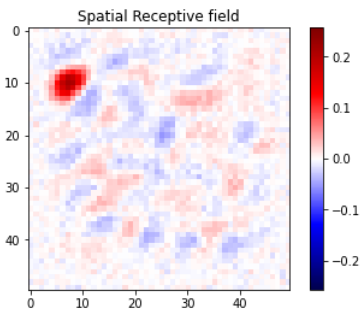
Cell 2



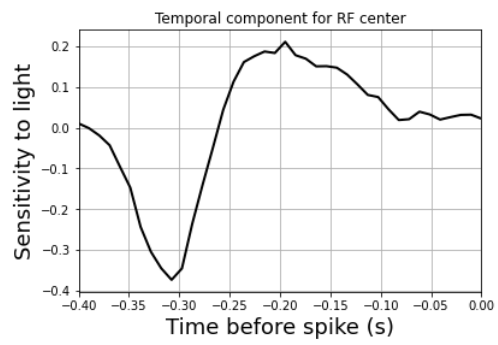
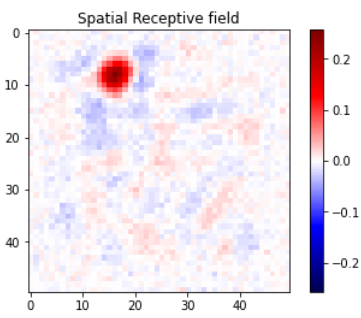
Cell 7



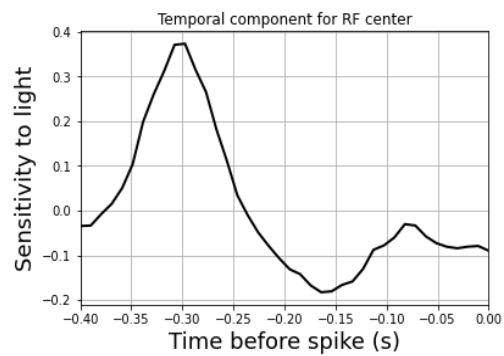
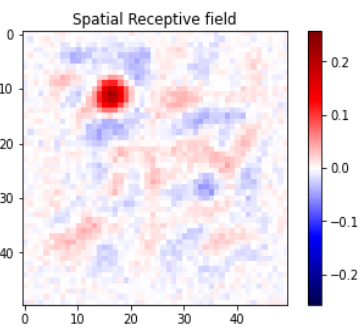
Cell 9



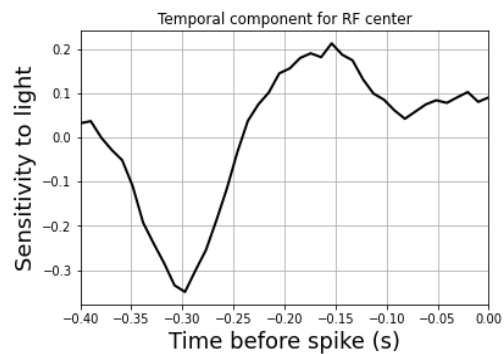
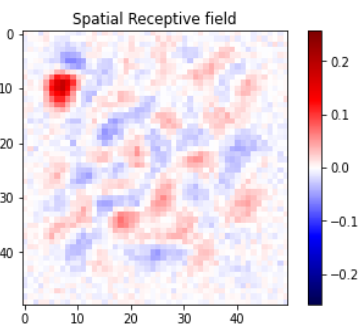
Cell 16



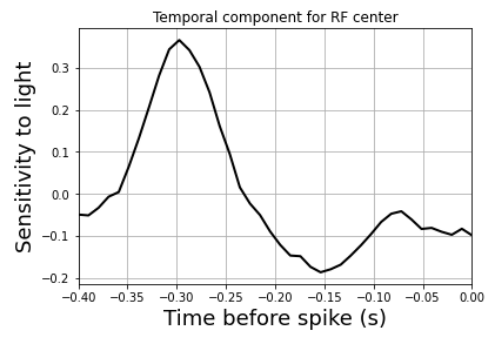
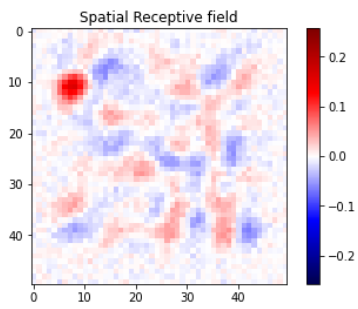
Cell 19



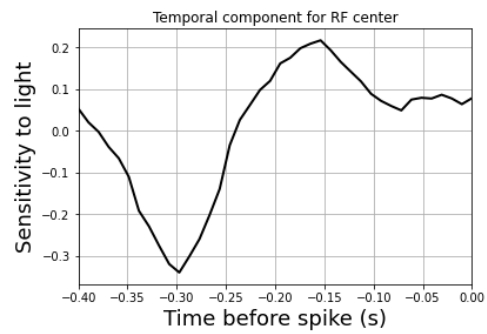
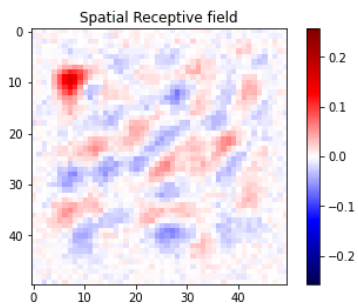
Cell 24



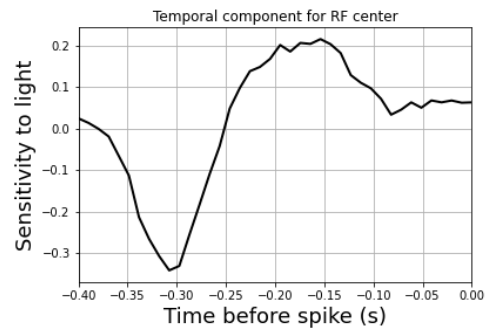
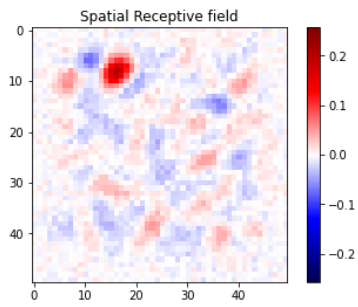
Cell 26



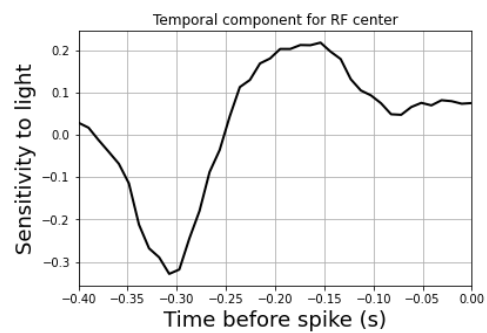
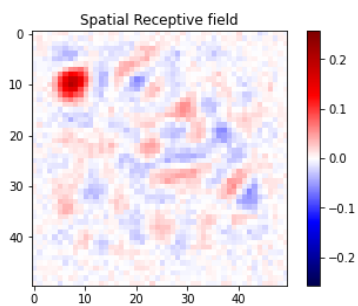
Cell 27



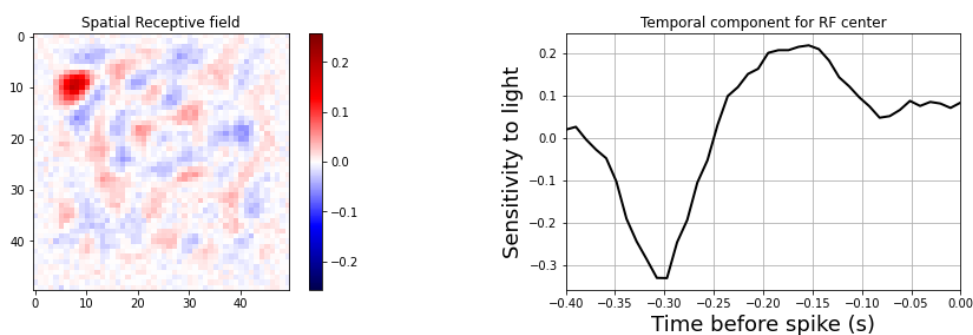
Cell 33



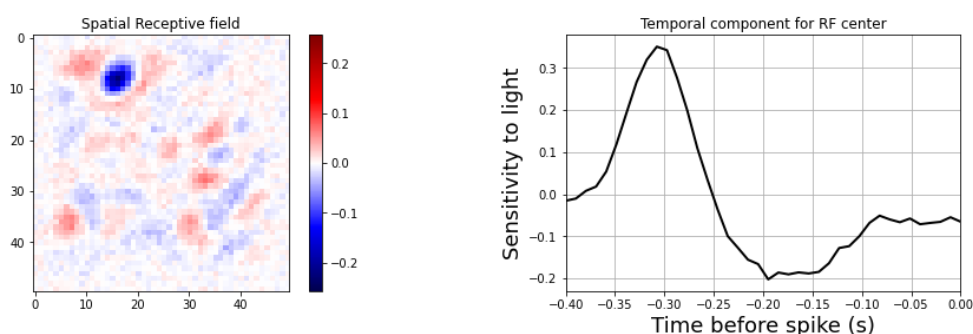
Cell 41



Cell 45



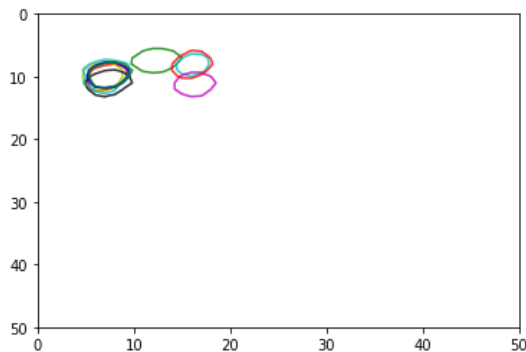
Cell 55



**Σχήμα 44:** Απεικόνιση του STA για τους 12 έγκυρους νευρώνες.

Αρχικά, παρατηρούμε πως τα κέντρα των πεδίων παρουσιάζουν επικάλυψη μεταξύ τους. Δεν είναι διασκορπισμένα σε διαφορετικά σημεία καλύπτοντας όλη την εικόνα, αλλά συγκεντρώνονται σε μια περιοχή πάνω και αριστερά στο οπτικό πεδίο. Αυτό φαίνεται και στην Εικόνα 45 με τα περιγράμματα των κέντρων. Η εξήγηση για αυτό το φαινόμενο ίσως εντοπίζεται σε μια βασική ιδιότητα του V1 που ονομάζεται αμφιβληστροειδοτοπικότητα (retinotopy) και ορίζεται ως η αντιστοιχία μεταξύ της τοπογραφικής θέσης των νευρώνων και των σημείων στο χώρο τα οποία τους ενεργοποιούν. Πιο συγκεκριμένα, γειτονικοί νευρώνες ενδιαφέρονται για γειτονικές μεταξύ τους περιοχές στο οπτικό πεδίο, η πληροφορία δηλαδή διατηρεί τη χωρική συνοχή της. Αν δηλαδή απεικονίσουμε χρωματικά την ενεργοποίηση των νευρώνων του V1 κατά την προβολή π.χ ενός σχήματος στους οφθαλμούς, τότε θα δούμε το σχήμα αυτό αποτυπωμένο πάνω στον V1. Θεωρείται πως, αφού το φαινόμενο αυτό εντοπίζεται στον V1, για λειτουργικούς λόγους θα εμφανίζεται και στον αμφιβληστροειδή, με το οπτικό νεύρο να διατηρεί την τοπογραφική δομή της πληροφορίας από τους νευρώνες του οφθαλμού [73]. Αν αποδεχθούμε ότι ισχύει αυτή η υπόθεση, τότε νευρώνες που οι αποκρίσεις τους καταγράφηκαν ταυτόχρονα κατά την εξωκυτταρική μέθοδο που χρησιμοποιήθηκε στο πείραμα, θα βρίσκονται κοντά στο ηλεκτρόδιο και άρα κοντά μεταξύ τους. Επομένως, λόγω της τοπογραφικής συνέπειας στον αμφιβληστροειδή, ενδιαφέρονται για γειτονικά χωρικά σημεία, και τα υποδεκτικά τους πεδία παρουσιάζουν επικάλυψη. Επίσης οι νευρώνες αυτοί λογικά λαμβάνουν είσοδο και από ίδια

κύτταρα των προηγούμενων σταδίων και άρα υπάρχει επικάλυψη στην πληροφορία που μεταβιβάζουν προς τα επόμενα στάδια.



**Σχήμα 45:** Επικάλυψη των κέντρων των υποδεκτικών πεδίων.

Οι μεταβολές, όπως έχουμε πει, δεν αφορούν τιμές φωτεινότητας αλλά αντιθέσεις, αποκλίσεις δηλαδή από τη μέση τιμή της εισόδου, που εδώ είναι 0. Στο κέντρο των πεδίων που, όπως αναμέναμε, έχει ένα ελλειψοειδές σχήμα που προσεγγίζει κυκλικό δίσκο, βλέπουμε ON συμπεριφορά για όλα τα κύτταρα εκτός του 55, το οποίο παρουσιάζει OFF χαρακτήρα. Αυτό ωστόσο δε μας λέει αμέσως πως το κύτταρο είναι ON ή OFF, αφού η χρονική συνιστώσα, που είναι αυτή που βοηθά στον διαχωρισμό των δύο τύπων, παρουσιάζει έντονα διφασική μορφή, που σημαίνει ότι το κύτταρο έχει ευαισθησία και σε αύξηση και σε μείωση στην ένταση χωρίς καθαρή προτίμηση. Το υποδεκτικό πεδίο που φαίνεται στην απεικόνιση είναι εξάλλου ένας μέσος όρος των 40 εγγραφών του STA, μία για κάθε χρονικό σημείο μελέτης, για κάθε χωρικό ρίχνελ και για όλες τις καταγεγραμμένες πυροδοτήσεις στο λευκό θόρυβο. Η ακριβής πορεία των αυξομειώσεων στην αντίθεση για το χρόνο που προηγείται της απόκρισης, καθώς και οι μεταβολές πολικότητας που υφίσταται το πεδίο που, όπως είπαμε, δεν αποτελεί μια στατική οντότητα, φαίνονται από το χρονικό πυρήνα του STA.

Ένα επιπλέον στοιχείο είναι η υποψία ανταγωνιστικής συμπεριφοράς μεταξύ κέντρου και περιφέρειας που γνωρίζουμε πως έχουν τα περισσότερα γαγγλιακά κύτταρα στο ποντίκι [74]. Δεν είναι ιδιαίτερα ορατή και φαίνεται πιο έντονα στα κύτταρα 7, 16, 19, 55. Επιπρόσθετα, σε ορισμένους νευρώνες όπως οι 7, 26, 27 παρατηρείται πως υπάρχει μια αυξημένη συσχέτιση μεταξύ απόκρισης και σημείων που είναι μακριά από το υποδεκτικό πεδίο, δηλαδή ο νευρώνας επηρεάζεται μάλλον αρκετά από αλλαγές και σε άλλες, απομακρυσμένες περιοχές της εικόνας. Θεωρούμε πως για τη βελτίωση των δύο παραπάνω, δηλαδή ο ανταγωνισμός «κέντρου - περιφέρειας» να είναι πιο ξεκάθαρος και το STA να έχει τιμές κοντά στο 0 μακριά από το υποδεκτικό πεδίο όπως οφείλει, χρειάζονται καταγραφές για περισσότερες εικόνες και για περισσότερες από μια επαναλήψεις του ίδιου συνόλου εικόνων, για πιο εύρωστα βιολογικά δεδομένα. Τέλος, διαπιστώσαμε πως ο νευρώνας 55 ήταν ο

μοναδικός που άλλαζε το μέσο πρόσημο του κέντρου του από θετικό (ON) σε αρνητικό (OFF) και αντιστρόφως μεταξύ διαδοχικών υπολογισμών STA.

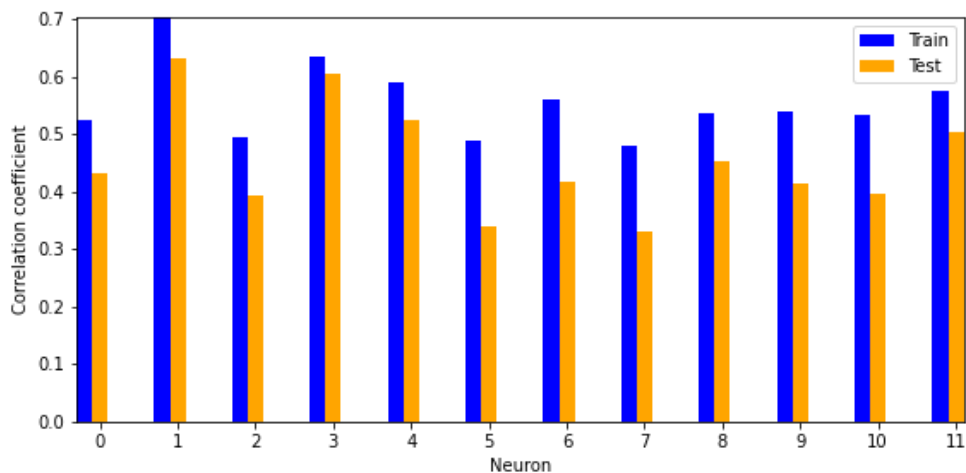
Αναφορικά με τη χρονική συνάρτηση, είναι διφασική στη μορφή της, όπως στα περισσότερα τυπικά γαγγλιακά κύτταρα. Στα 0.4 s πριν από την πυροδότηση, για όλους τους νευρώνες η τιμή της κλίνει προς το 0 και συμπεραίνουμε πως σωστά πράξαμε κατά την επιλογή της μνήμης του νευρώνα. Φαίνεται ωστόσο πως δεν υπάρχει μια ξεκάθαρη προτίμηση πολικότητας, όπως στην Εικόνα 37, προκειμένου να εξαχθεί με ασφάλεια ένα συμπέρασμα για τη συμπεριφορά του κυττάρου, αν δηλαδή είναι ON ή OFF. Οι υπό μελέτη νευρώνες διεγείρονται σχεδόν εξίσου από αυξήσεις και από μειώσεις. Αυτό μας προβλημάτισε, και μας έκανε να αντιληφθούμε ένα σοβαρό μειονέκτημα της μεθόδου STA. Η συγκεκριμένη τεχνική είναι αρκετά αποτελεσματική στο να διαχωρίζει ON από OFF κύτταρα με βάση το υποδεκτικό πεδίο, αποτυγχάνει ωστόσο σε κύτταρα που είναι ON-OFF, δηλαδή μη γραμμικά, που εξαρτώνται έντονα από τα πειραματικά δεδομένα που χρησιμοποιούνται για τον προσδιορισμό της απόκρισής τους. Τα κύτταρα αυτά έχουν τουλάχιστον ένα ON και ένα OFF πεδίο ενδιαφέροντος και ένα μοναδικό γραμμικό φίλτρο αδυνατεί να εξηγήσει τα διαφορετικά κανάλια επεξεργασίας που τα χαρακτηρίζουν [57], [67], [75]. Για να γίνει πιο κατανοητό τι ακριβώς συμβαίνει, επειδή η STA προσδιορίζει τα αίτια που προκάλεσαν δραστηριότητα στους νευρώνες, κατά μέσο όρο για όλες τις αποκρίσεις, αν, για παράδειγμα, στη χρονική στιγμή  $-t$  πριν από ένα δυναμικό έχουμε αύξηση στην ένταση και για κάποιο άλλο δυναμικό στην ίδια χρονική στιγμή  $-t$  έχουμε μείωση, το οποίο μπορεί να συμβεί σε ένα ON-OFF κύτταρο, η συμπεριφορά αυτή θα χαθεί με την STA, αφού αυτή "εγκλωβίζει" την ON και OFF επεξεργασία σε ένα μοναδικό, γραμμικό φίλτρο. Όταν λοιπόν δε γνωρίζουμε εξαρχής ποιος είναι ο τύπος κυττάρων που μελετάμε, όπως στην περίπτωση μας, αν η STA υπολογίσει διφασική χρονική κυματομορφή με ασύμμετρες ON-OFF συνιστώσες, ο νευρώνας δεν είναι απαραίτητα ON ή OFF, ανάλογα με τη συνιστώσα που υπερισχύει, αλλά μπορεί να είναι ON-OFF [67]. Αυτό είναι ένα ιδιαίτερο πρόβλημα στο ποντίκι, στο οποίο ο πιο κοινός τύπος γαγγλιακού κυττάρου είναι ON-OFF [75].

Ορίσαμε λοιπόν τους νευρώνες ως "κατεξοχήν-ON" (αντί για ON) όταν η μέγιστη θετική τιμή είναι πιο κοντά στη χρονική στιγμή 0 της χρονικής συνάρτησης από τη μέγιστη αρνητική τιμή, ενώ στην αντίθετη περίπτωση ορίστηκαν ως "κατεξοχήν-OFF" (αντί για OFF). Αυτό έγινε με το σκεπτικό πως η πολικότητα που βρίσκεται πλησιέστερα στη χρονική στιγμή πυροδότησης είναι και αυτή που επηρεάζει περισσότερο στη διαμόρφωσή της [67]. Εδώ, τα κύτταρα 2, 9, 16, 24, 27, 33, 41, 45 είναι "κατεξοχήν-ON" και τα 7, 19, 26, 55 είναι "κατεξοχήν-OFF". Η εναλλαγή συμπεριφοράς που αναφέραμε πως εμφανίζε το κύτταρο 55 ίσως να οφείλεται στο ότι είναι ON-OFF και συνεπώς καμία πολικότητα δεν κυριαρχεί.

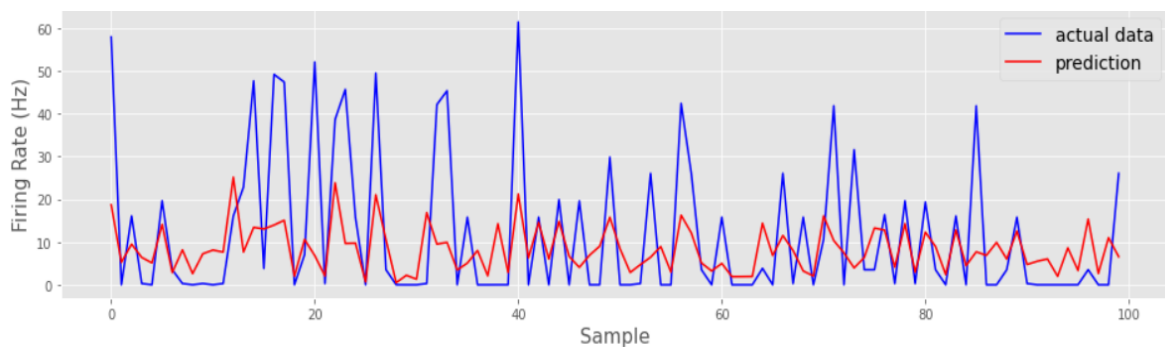
Τον περιορισμό της STA μπορούν να μετριάσουν πιο ιδιαίτερες και ειδικές μέθοδοι όπως η STC (Spike-Triggered Covariance) που, βασιζόμενη στη διακύμανση, υπολογίζει πολλαπλά γραμμικά φίλτρα παρέχοντας μια πιο αναλυτική προσέγγιση [67], [76]. Δεν ασχοληθήκαμε με αυτές τις τεχνικές καθώς η STA παραμένει αξιόπιστη στο να εντοπίζει τη θέση ενδιαφέροντος του πεδίου ακόμη και σε ON-OFF κύτταρα καθώς και να παρέχει

κάποια βασικά χαρακτηριστικά που αρκούν για το επίπεδο σχολιασμού που σκοπεύαμε να κάνουμε.

Στην Εικόνα 46 έχουμε εμφανίσει τις τιμές του συντελεστή συσχέτισης για τους νευρώνες που επιλέχθηκαν να προσομοιώσουν την έξοδο του αμφιβληστροειδούς στο πείραμα. Στα δεδομένα αξιολόγησης διακρίνεται μια διακύμανση στις αποδόσεις, με τους νευρώνες 7 και 16 να έχουν σχεδόν διπλάσιο συντελεστή συγκριτικά με τους χειρότερους νευρώνες, το οποίο, αν και δεν είναι όσο έντονα αισθητό θα θέλαμε, αποτυπώνεται και στην ποιοτική απεικόνιση των υποδεκτικών πεδίων τους.



**Σχήμα 46:** Συντελεστής συσχέτισης για τους 12 επιλεγμένους νευρώνες.



**Σχήμα 47:** Διάγραμμα πραγματικής απόκρισης και πρόβλεψης για νευρώνα με  $cc_{test} = 0.606$ .

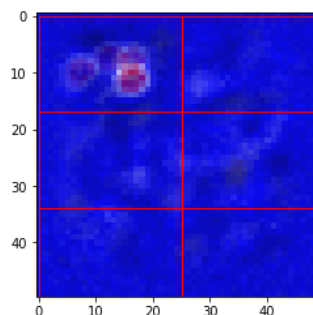
### 4.3 Πρόβλεψη απόκρισης για το σύνολο δεδομένων προεξοχής

Το τελευταίο μέρος του πειράματος είναι η σύγκριση της προεξοχής με την απόκριση του αμφιβληστροειδούς σε φυσικές εικόνες. Το σκεπτικό είναι να μελετηθεί αν περιοχές της εικόνας με μεγαλύτερη τιμή προεξοχής ωθούν τα κύτταρα σε πιο έντονη ενεργοποίηση και αντίστοιχα αν περιοχές της εικόνας για τις οποίες ο άνθρωπος δεν ενδιαφέρεται σε πρώτο

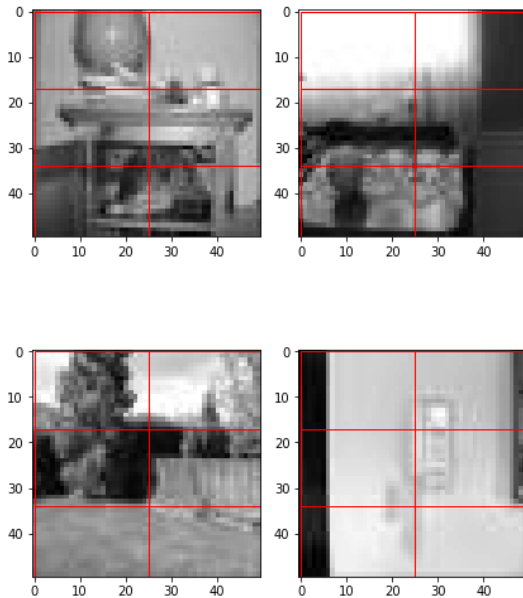


βαθμό προκαλούν χαμηλότερη απόκριση. Χρησιμοποιήσαμε ένα κλασικό σύνολο δεδομένων προεξοχής, το Toronto [70] που αποτελείται από 120 έγχρωμες εικόνες εσωτερικού και εξωτερικού χώρου με διαστάσεις  $681 \times 511$  pixels και καταγραφές εστιάσεων από 20 παρατηρητές. Προκειμένου το νευρωνικό δίκτυο να εκτιμήσει τις αποκρίσεις στο σύνολο, προηγήθηκε μια διαδικασία προεπεξεργασίας μέσω Matlab όπου οι εικόνες μετατράπηκαν από έγχρωμες σε μονοχρωματικές, κανονικοποιήθηκαν στο εύρος  $[0,1]$  και οι διαστάσεις τους άλλαξαν σε  $50 \times 50$ .

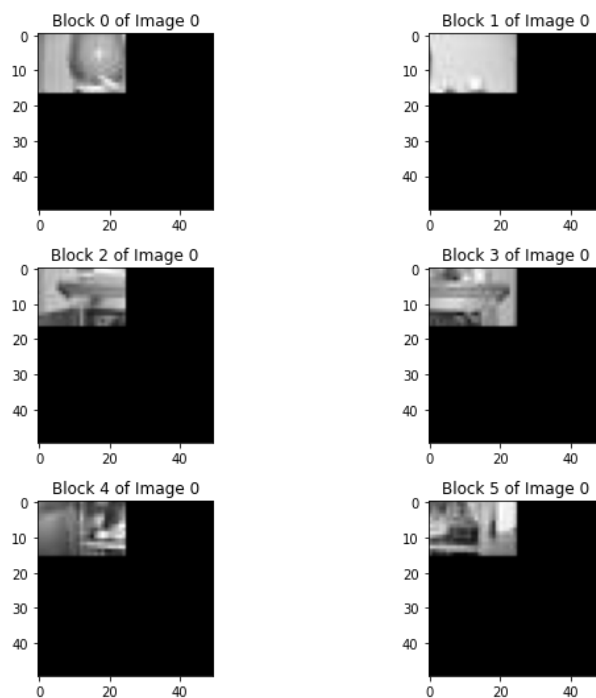
Ιδανικά αυτό που θα θέλαμε να δούμε είναι τα υποδεκτικά πεδία των 12 νευρώνων του πειράματος να σχηματίζουν ένα πλέγμα στην εικόνα καλύπτοντας διαφορετικές περιοχές του οπτικού πεδίου. Με αυτό τον τρόπο θα μπορούσαμε να μελετήσουμε ευκολότερα τη σχέση προεξοχής-απόκρισης, συγκρίνοντας την ένταση της απόκρισης του νευρώνα που ενδιαφέρεται για το πιθανότερο σημείο εστίασης με την ένταση της απόκρισης των υπολοίπων νευρώνων, των οποίων το ενδιαφέρον επικεντρώνεται μακριά από την περιοχή εστίασης. Αντ'αυτού τα υποδεκτικά πεδία του πειράματος μας επικαλύπτονται. Για αυτό το λόγο προχωρήσαμε σε μια επιπλέον παρέμβαση. Αντί να τροφοδοτήσουμε το δίκτυο με ολόκληρη την εικόνα σε ένα δείγμα, τη διασπάμε σε έξι παραλληλόγραμμα τμήματα ίσων διαστάσεων  $17 \times 25$ , τα οποία δείχνουμε διαδοχικά το ένα μετά το άλλο και ένα ανά δείγμα στους νευρώνες. Το κάθε τμήμα θα καταλαμβάνει την πάνω αριστερά περιοχή της εικόνας, εκεί δηλαδή όπου συγκεντρώνονται τα υποδεκτικά πεδία και η ανταγωνιστική περιοχή της περιφέρειας, ενώ τα υπόλοιπα pixels θα είναι μαύρα με στόχο το ενδιαφέρον να απομονώνεται αποκλειστικά εκεί όπου οι νευρώνες έχουν δείξει ότι είναι περισσότερο ευαίσθητοι σε αλλαγές. Το σύνολο δεδομένων συνεπώς θα έχει  $120 \times 6 = 720$  εικόνες. Με τη συγκεκριμένη τροποποίηση δυστυχώς χάνεται η ευκαιρία να συγκρίνουμε ταυτόχρονα νευρώνες για την ίδια διέγερση αφού ουσιαστικά όλοι ενδιαφέρονται για την ίδια περιοχή και δεν έχει κάποιο ιδιαίτερο νόημα. Έτσι, αυτό που μελετάμε είναι απλώς αν προεξέχουσες περιοχές προκαλούν αύξηση στο ρυθμό πυροδότησης, χωριστά για κάθε νευρώνα.



**Σχήμα 48:** Διάσπαση της εικόνας σε τμήματα και θέση των κέντρων των υποδεκτικών πεδίων.



**Σχήμα 49:** Εικόνες από το Toronto μετά από επεξεργασία, μαζί με τα τμήματα στα οποία θα χωριστούν.



**Σχήμα 50:** Δείγματα προς πρόβλεψη.

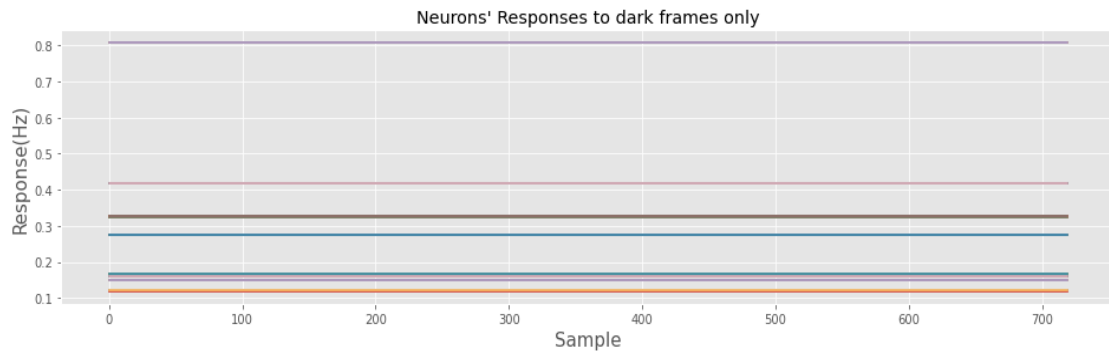
#### 4.4 Οι νευρώνες είναι ευαίσθητοι σε αλλαγές στο χρόνο

Ιδιαίτερο ενδιαφέρον παρουσιάστηκε στη διαμόρφωση της χρονικής διάστασης των δειγμάτων. Παραμένει προφανώς στην τιμή 40, ωστόσο προτιμήσαμε να εξαλείψουμε την

εξάρτηση από άλλες εικόνες κατά την πρόβλεψη κι έτσι διαφοροποιηθήκαμε ελαφρώς από την τεχνική που εφαρμόσαμε κατά την εκπαίδευση, όπου κάθε δείγμα περιελάμβανε μια εικόνα μαζί με τις 39 που προηγήθηκαν αυτής κατά το πείραμα. Το δίκτυο έχει εκπαιδευτεί έτσι ώστε να αντιλαμβάνεται πως σε ένα δείγμα οι εικόνες που έπονται της πρώτης είναι αυτές που έχουν προηγηθεί χρονικά αυτής, και αφήνεται σε εμάς να επιλέξουμε τι θα προβάλλουμε στους νευρώνες και με ποιο τρόπο. Τώρα κάθε δείγμα θα περιέχει μόνο την εικόνα προς πρόβλεψη σε διαδοχικά  $N$  αντίγραφα, ακολουθούμενα από  $40 - N$  σκοτεινά καρέ, με το  $N$  να είναι σταθερό για όλα τα δείγματα. Τα δείγματα είναι ανεξάρτητα μεταξύ τους που σημαίνει πως δεν έχει διαφορά η σειρά με την οποία θα δείξουμε τα τμήματα της εικόνας στα κύτταρα.

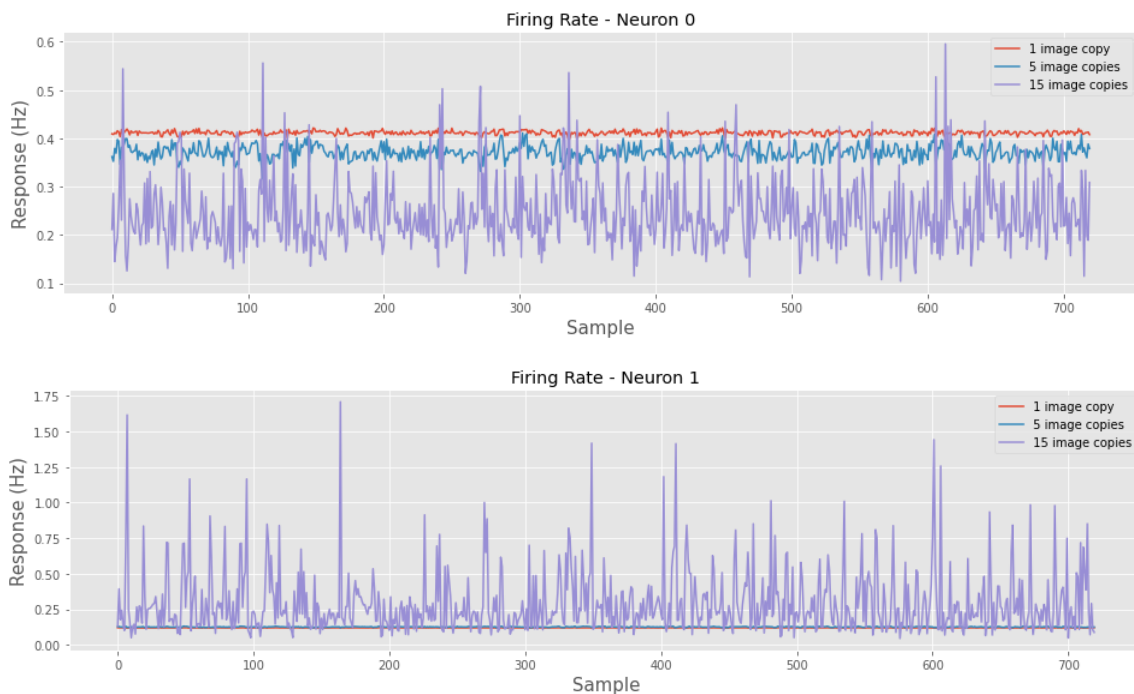
Για την επιλογή του κατάλληλου  $N$  το ζητούμενο είναι η μεγάλη διακύμανση, δηλαδή διαφορετικές διεγέρσεις να προκαλούν διαφορετικές αποκρίσεις. Για  $N = 1$ , που είναι σαν να κοιτάμε την απόκριση αμέσως μόλις δείξουμε την εικόνα για 10 ms, παρατηρήσαμε πως δεν έχουμε διακύμανση στη απόκριση, δηλαδή όλες οι εικόνες προκαλούν παρόμοια απόκριση. Μελετήσαμε λοιπόν τιμές μεγαλύτερες του 1. Με αυτό τον τρόπο, αντί να κοιτάζουμε την απόκριση του αμφιβληστροειδούς μόλις δείξουμε την εικόνα, προβάλλουμε την εικόνα για διάστημα  $N \times 10$  ms με  $N > 1$  και στην συνέχεια κοιτάμε την απόκριση, λαμβάνοντας πιο αντιπροσωπευτικά αποτελέσματα. Για παράδειγμα, για  $N = 9$  το δείγμα έχει εννέα αντίγραφα της ίδιας εικόνας και μετά 31 καρέ σκοταδιού. Αυτό προσομοιώνει την προβολή σκοταδιού για 310 ms και αμέσως μετά την προβολή της εικόνας για 90 ms.

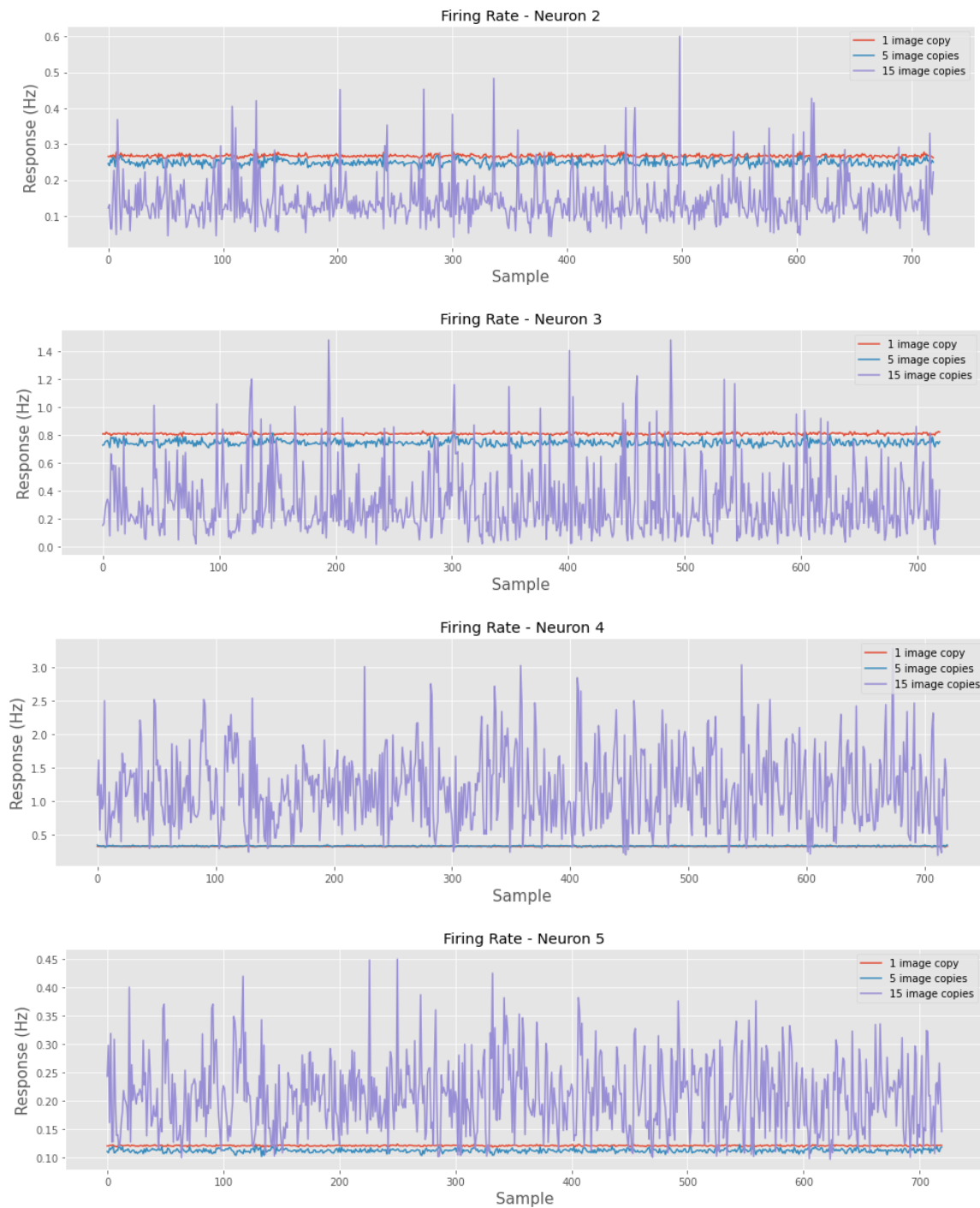
Στην Εικόνα 51 φαίνεται η δραστηριότητα των 12 κυττάρων για δείγματα που περιέχουν αποκλειστικά σκοτάδι ( $N = 0$ ). Βλέπουμε πως όλοι οι νευρώνες έχουν σταθερή απόκριση σε μια χαμηλή τιμή ηρεμίας. Αρχίσαμε να αυξάνουμε τον αριθμό των αντιγράφων της κάθε εικόνας από 1 έως 40. Για  $N = 1$ , όπως αναφέρθηκε, έχουμε πάρα πολύ χαμηλή δραστηριότητα από τους νευρώνες που κυμαίνεται γύρω από την τιμή ηρεμίας και σε ένα πολύ μικρό εύρος τιμών. Όσο αυξάνεται το  $N$ , δηλαδή ο χρόνος που ο νευρώνας βλέπει την εικόνα, παρατηρήσαμε πως αυξάνεται και η διακύμανση στις αποκρίσεις, δηλαδή έχουμε εντονότερη δραστηριότητα και πυροδοτήσεις προς πιο ακραίες τιμές. Αυτή η συμπεριφορά σημειώνεται μέχρι κάποια τιμή κορυφής, ύστερα από την οποία η απόκλιση εμφάνισε μια πτωτική τάση. Αυτό το φαινόμενο μπορεί να μας πει πως ο νευρώνας ενδιαφέρεται για μεταβολές στο χρόνο. Δείχνοντας του διαρκώς το ίδιο μοτίβο, είτε αυτό είναι σκοτάδι είτε είναι η ίδια εικόνα, μειώνει το ρυθμό ενεργοποίησής του. Προσαρμόζεται στη σταθερή διέγερση που δέχεται και επιλέγει να αντιδράσει μονάχα σε χρονικές αλλαγές. Αυτό αποτελεί έναν ιδιαίτερα αποδοτικό τρόπο κωδικοποίησης της πληροφορίας. Εναλλακτικά, μπορούμε να ισχυριστούμε ότι όσο λιγότερα σκοτάδια έχουμε (τιμή 0 στα pixels), τόσο περισσότερα βάρη του νευρωνικού δικτύου συμμετέχουν στη διαμόρφωση της εξόδου  $out = f[w \cdot x + b]$ , άρα ξεφεύγουμε και περισσότερο από την τιμή ηρεμίας του νευρώνα.



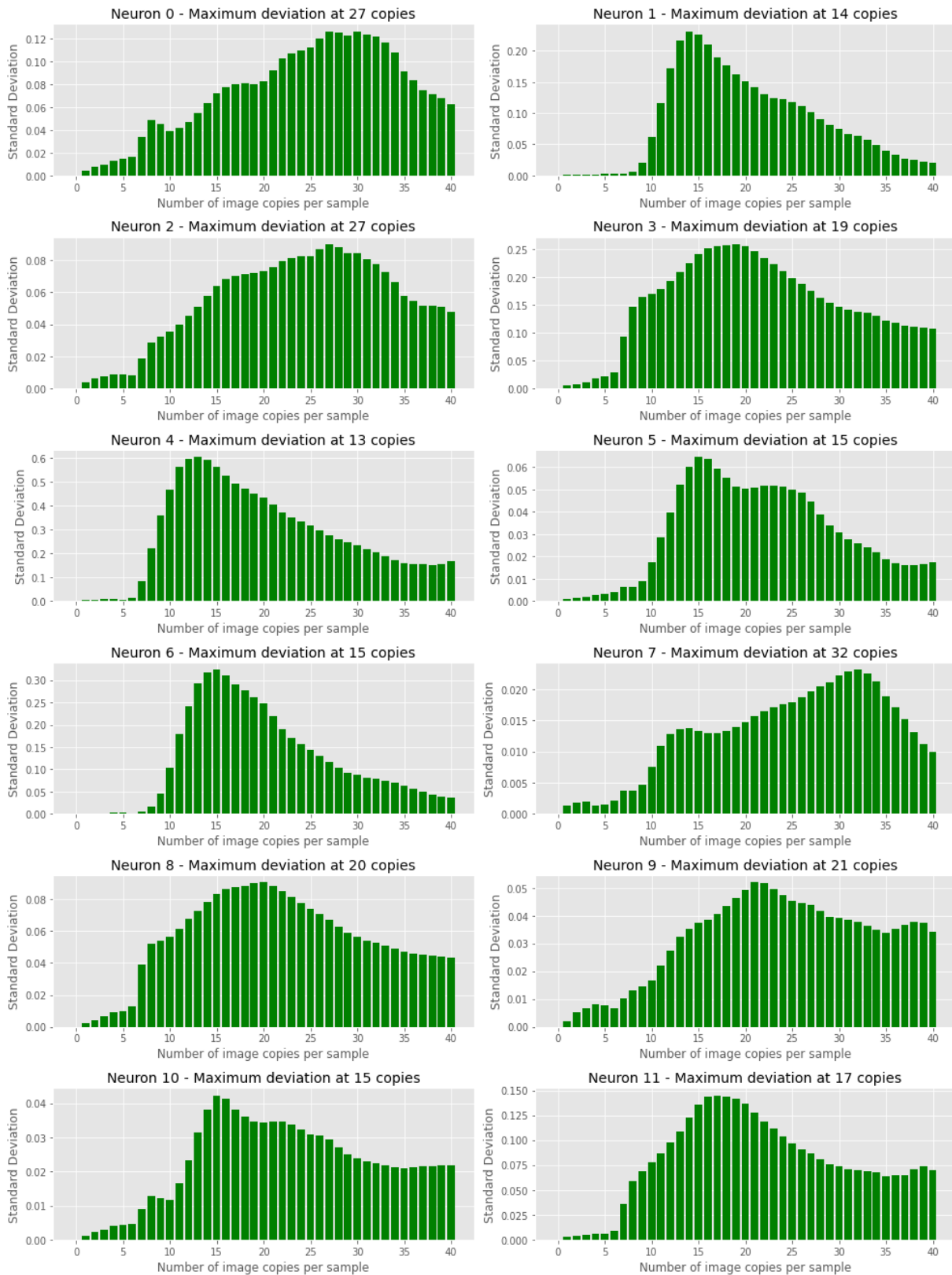
**Σχήμα 51:** Οι αποκρίσεις των νευρώνων για  $N = 0$ .

Στην Εικόνα 52 φαίνεται η συμπεριφορά που περιγράψαμε, ενδεικτικά για κάποια κύτταρα και για τρεις τιμές του  $N$  για το σύνολο δεδομένων. Στην Εικόνα 53 απεικονίζεται σε μορφή ιστογράμματος η πορεία της τυπικής απόκλισης για όλες τις δυνατές τιμές του  $N$  και για όλα τα κύτταρα, ενώ αναγράφεται και το  $N$  για το οποίο η απόκλιση μεγιστοποιείται. Τέλος, στην Εικόνα 54 μπορούμε να δούμε το πώς μεταβάλλεται η μέση τιμή των αποκρίσεων σε σχέση με την τιμή ηρεμίας.

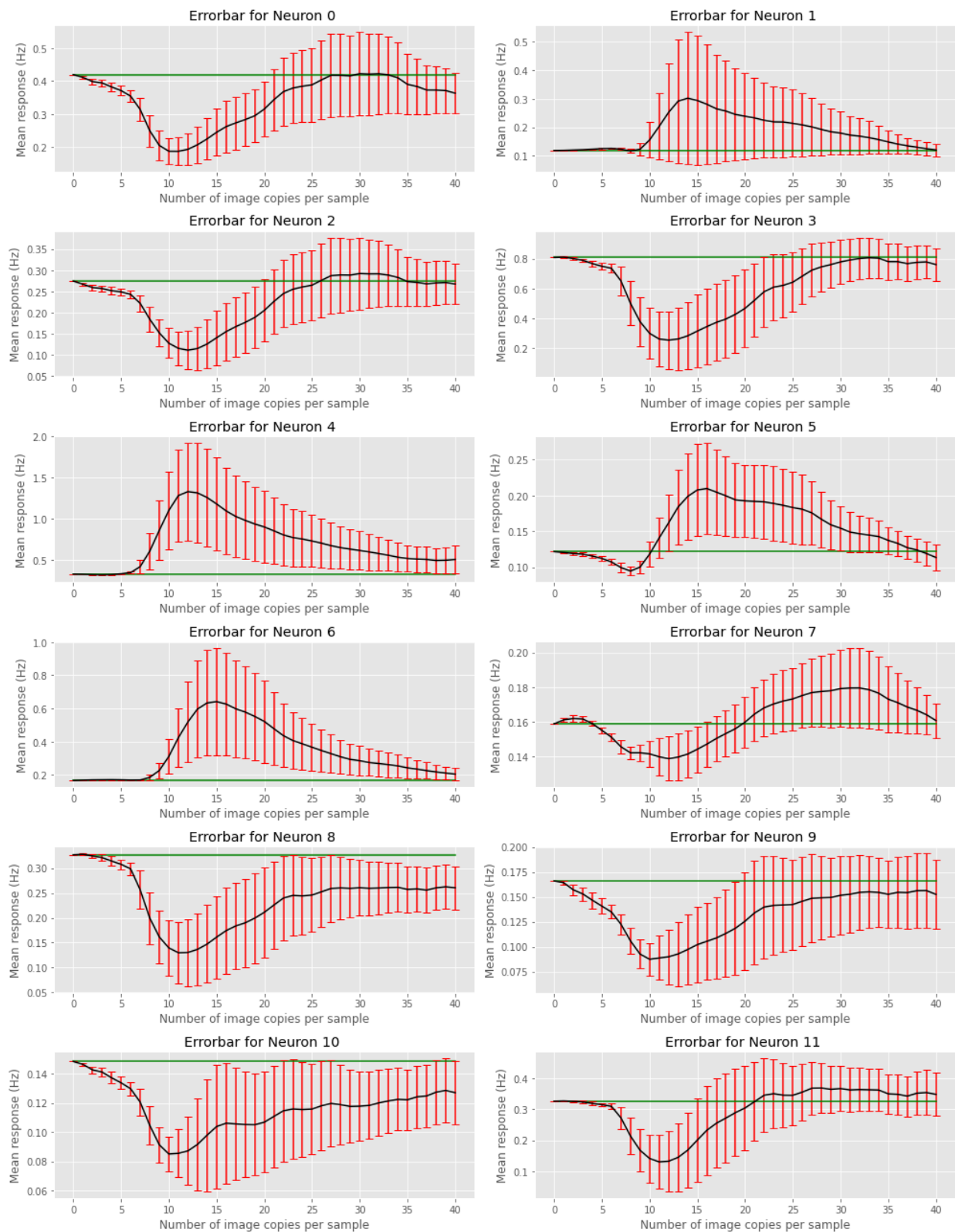




**Σχήμα 52:** Κυματομορφές της απόκρισης για  $N = 1, 5, 15$ . Είναι εμφανής η εντονότερη δραστηριότητα όσο το  $N$  αυξάνει.

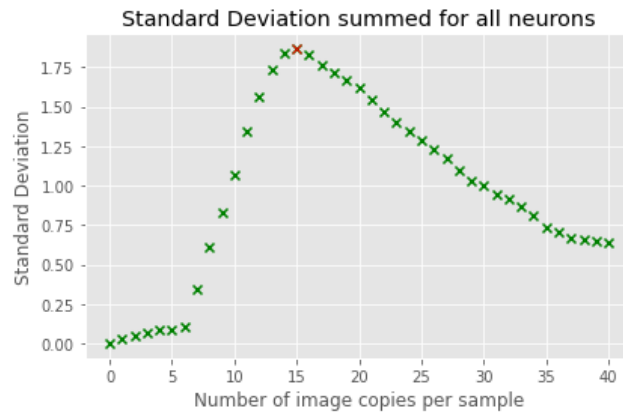


**Σχήμα 53:** Ιστογράμματα τυπικής απόκλισης για μεταβαλλόμενο  $N$ . Διακρίνουμε την αυξητική τάση μέχρι κάποια τιμή, μετά από την οποία σημειώνεται πτώση.

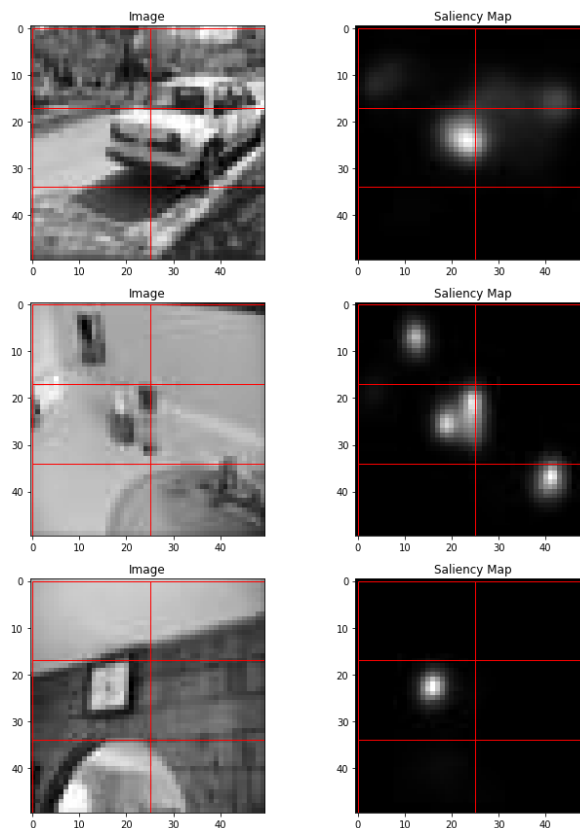


**Σχήμα 54:** Μεταβολή της συνολικής μέσης απόκρισης για κάθε νευρώνα συγκριτικά με την τιμή ηρεμίας (ευθεία γραμμή).

Επιλέγουμε την τιμή  $N = 15$  αφού αυτή επιδεικνύει αθροιστικά για τους 12 νευρώνες τη μέγιστη διακύμανση (Εικόνα 55).



Σχήμα 55: Τυπική απόκλιση συνολικά για τους 12 νευρώνες για όλες τις τιμές του  $N$ .



Σχήμα 56: Χάρτες προεξοχής, όπως υπολογίστηκαν από το SAM.



## 4.5 Σύγκριση έντασης προεξοχής και ρυθμού πυροδότησης των νευρώνων

Αφού έχουμε υπολογίσει μια εκτίμηση για κάθε κομμάτι των εικόνων της βάσης δεδομένων για τις 12 εξόδους του αμφιβληστροειδούς, βρίσκουμε για κάθε εικόνα το κομμάτι που έχει την υψηλότερη προεξοχή, υπολογίζοντας τη μέση τιμή έντασης για το καθένα στο χάρτη προεξοχής που έχει δημιουργήσει το δίκτυο SAM. Στη συντριπτική πλειοψηφία των περιπτώσεων, το κομμάτι με τη μεγαλύτερη μέση τιμή έντασης στο χάρτη είναι και αυτό στο οποίο εντοπίζεται το διακριτό σημείο της συνολικής εικόνας με τη μέγιστη τιμή προεξοχής, το σημείο δηλαδή στο οποίο το νευρωνικό δίκτυο δίνει τη μεγαλύτερη πιθανότητα ένας άνθρωπος να εστιάσει σε πρώτο χρόνο. Υπολογίζουμε το λόγο

$$\lambda = \frac{\text{firing rate}^*}{\langle \text{firing rate} \rangle}, \quad (12)$$

όπου  $\text{firing rate}^*$  είναι ένας πίνακας διαστάσεων (εικόνες, νευρώνες) = (120, 12) με τις αποκρίσεις των νευρώνων για το ένα και μοναδικό τμήμα κάθε εικόνας με τη μέγιστη προεξοχή και  $\langle \text{firing rate} \rangle$  είναι δύο ξεχωριστοί πίνακες που περιλαμβάνουν:

- (120, 12) καταχωρήσεις με τη μέση απόκριση κάθε νευρώνα για κάθε εικόνα ξεχωριστά, δηλαδή κάθε καταχώρηση υπολογίζεται αποκλειστικά για τα έξι κομμάτια της εκάστοτε εικόνας, και
- 12 καταχωρήσεις με τη μέση απόκριση κάθε νευρώνα για τα 720 δείγματα εισόδου.

Στην πρώτη περίπτωση μπορούμε να μελετήσουμε αν και κατά πόσο ισχύει πως, μέσα στα πλαίσια μιας εικόνας, το τμήμα με τη μεγαλύτερη τιμή προεξοχής προκαλεί και εντονότερη δραστηριότητα σε σχέση με τα υπόλοιπα. Στη δεύτερη περίπτωση θα μπορούσαμε να δούμε πού βρίσκεται η απόκριση για το επιλεγμένο κομμάτι συγκριτικά με το μέσο ρυθμό ενεργοποίησης του νευρώνα συνολικά για όλα τα δείγματα. Και στις δύο δοκιμές το αποτέλεσμα είναι ένας πίνακας διαστάσεων (120, 12) με λόγους της μορφής (12). Εξετάζουμε ξεχωριστά κάθε νευρώνα και εξάγουμε:

(i) μέγιστη τιμή, ελάχιστη τιμή, μέση τιμή και τυπική απόκλιση,

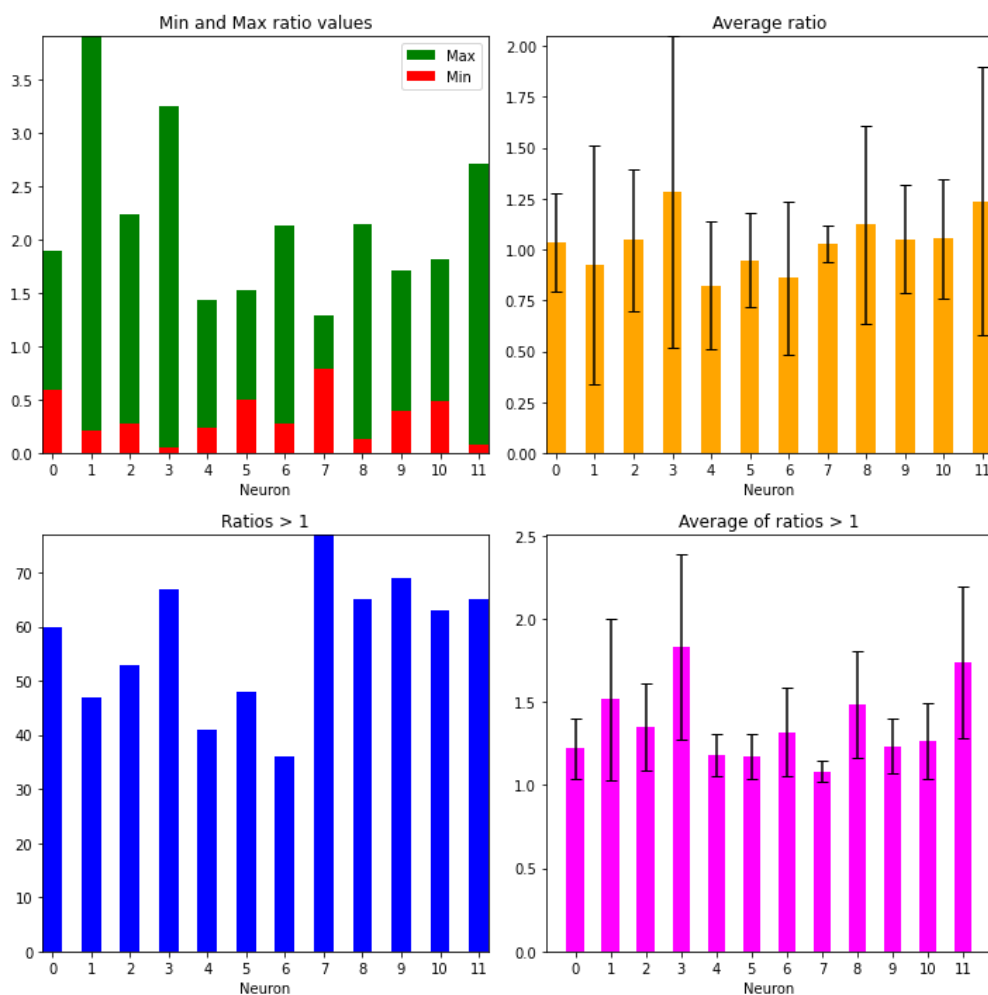
(ii) πλήθος λόγων που είναι μεγαλύτεροι του 1, δηλαδή περιπτώσεις που το πιο προεξέχον τμήμα προκαλεί ρυθμό απόκρισης άνω του μέσου,

(iii) μέση τιμή και τυπική απόκλιση μόνο για τους λόγους που υπάγονται στο (ii).

Τα αποτελέσματα για το πρώτο πείραμα φαίνονται στον Πίνακα 2 και στην Εικόνα 57 και για το δεύτερο πείραμα στον Πίνακα 3 και στην Εικόνα 58.

**Πίνακας 2:** Στατιστικά στοιχεία για τη σύγκριση του *firing rate* \* με τη μέση απόκριση για κάθε εικόνα χωριστά.

Στοιχεία για το σύνολο του πίνακα	
Ελάχιστη τιμή	0.059
Μέγιστη τιμή	3.900
μ (σ)	1.034 (0.450)
Πλήθος λόγων > 1	691/1440
μ (σ) για πλήθος λόγων > 1	1.372 (0.390)



**Σχήμα 57:** Στατιστικά στοιχεία για τη σύγκριση του *firing rate* \* με τη μέση απόκριση για κάθε εικόνα χωριστά για τους 12 νευρώνες.

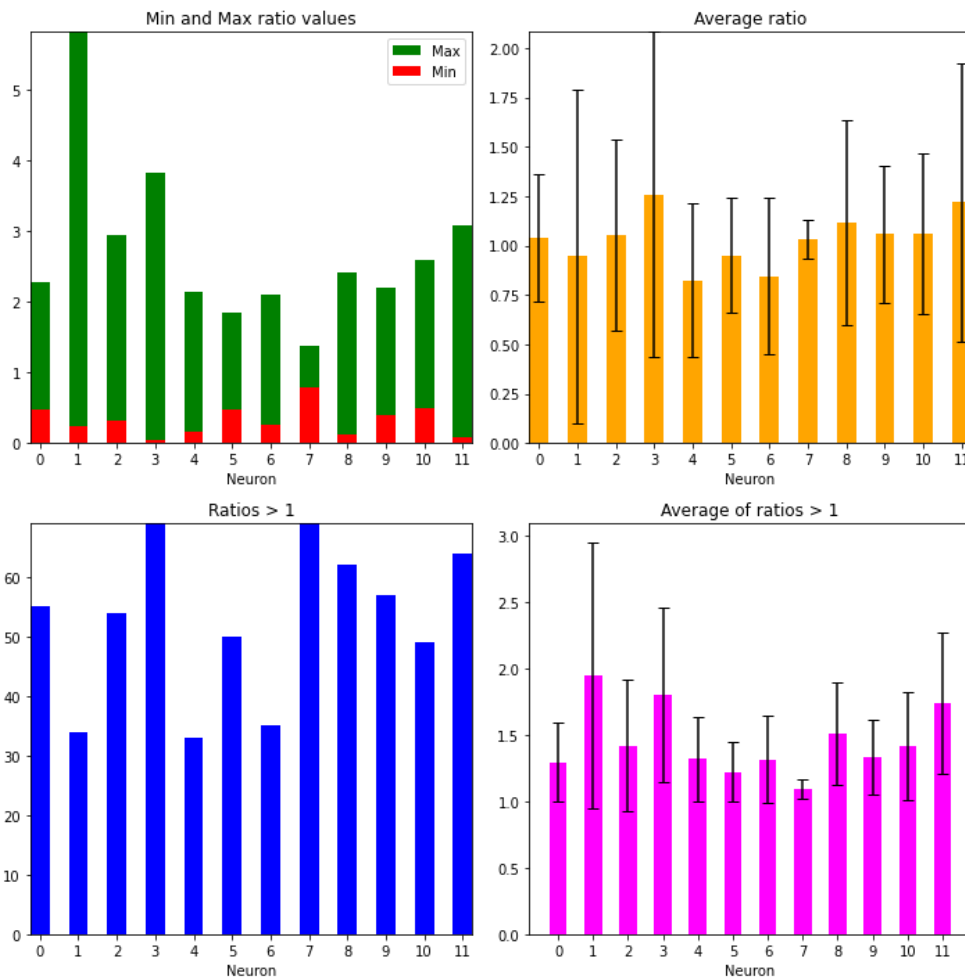
Στην πρώτη περίπτωση, αρχικά για το σύνολο του πίνακα (Πίνακας 2) βλέπουμε πως το  $\lambda$  έχει ελάχιστη τιμή πολύ χαμηλότερη του 1 και μέγιστη αρκετά υψηλότερη. Έχουμε μέση τιμή λίγο πάνω από το 1 συνεπώς οι λόγοι κυμαίνονται γύρω από αυτή την τιμή, ωστόσο βλέπουμε υψηλή τυπική απόκλιση το οποίο μας κάνει να υποπτευθούμε πως το  $\lambda$  θα κυμαίνεται σε ένα μεγάλο εύρος τιμών. Από όλα τα δείγματα, στο 47.9% το περισσότερο προεξέχον κομμάτι οδήγησε και στη μεγαλύτερη απόκριση στα πλαίσια της ίδιας εικόνας, ωστόσο κι εδώ έχουμε σχετικά χαμηλή μέση τιμή, άρα οι αποκρίσεις δεν ξεφεύγουν ιδιαίτερα από τη μέση δραστηριότητα.

Τα πρώτα αποτελέσματα δείχνουν πως δε μπορεί να εξαχθεί ακόμη κάποιο συμπέρασμα για το αν παρουσιάζεται η υπό μελέτη ιδιότητα, για αυτό προχωράμε στη μελέτη για κάθε νευρώνα μεμονωμένα. Στην Εικόνα 57, το πρώτο που διακρίνουμε είναι διαφορετική συμπεριφορά μεταξύ των νευρώνων, ιδιαίτερα στη μέγιστη και ελάχιστη τιμή και στο πλήθος  $\lambda > 1$ . Οι νευρώνες 1, 3 και 11 φτάνουν σε υψηλές τιμές ενώ οι 4, 5, 7 μετά βίας ξεπερνούν τη μονάδα. Στο νευρώνα 7 σε αρκετές εικόνες, 77 στο σύνολο, το μέγιστο προεξέχον τμήμα προκαλεί μεγαλύτερο ρυθμό πυροδότησης από το μέσο και από τους υπόλοιπους αρκετοί είναι αυτοί που πλησιάζουν αυτή την τιμή, είναι και άλλοι ωστόσο (1, 4, 6) με υποδιπλάσιο πλήθος. Στη μέση τιμή τόσο όλων των λόγων όσο κι αυτών που ξεπερνούν τη μονάδα η κατάσταση είναι πιο ισορροπημένη. Στην πρώτη περίπτωση, αναμενόμενα από τον προηγούμενο πίνακα, οι τιμές κυμαίνονται γύρω από το 1 και στη δεύτερη περίπτωση δε σημειώνονται τιμές υψηλότερες του 2. Επιπλέον οι περισσότεροι νευρώνες, ιδιαίτερα οι 1, 3 και 11, σημειώνουν αυξημένη τυπική απόκλιση, μια επιπλέον ένδειξη της απουσίας σταθερής συμπεριφοράς.

Από τις παραπάνω παρατηρήσεις αντιλαμβανόμαστε ότι δε μπορούμε να δηλώσουμε πως κάποιος από τους νευρώνες παρουσιάζει κάποιο μοτίβο στη συμπεριφορά του, δηλαδή για ένα αρκετά υψηλό ποσοστό εικόνων το τμήμα μέγιστης προεξοχής να τον διεγείρει με εντονότερο ρυθμό, συγκριτικά με τα υπόλοιπα τμήματα της ίδιας εικόνας. Οι 77/120 (64%) περιπτώσεις του νευρώνα 7 θεωρούμε πως δεν επαρκούν ώστε να προκύψει ένα ασφαλές συμπέρασμα. Για τον ίδιο νευρώνα, τα υπόλοιπα στατιστικά στοιχεία μαρτυρούν πως οι λόγοι βρίσκονται αρκετά κοντά και γύρω από το 1, οπότε φαίνεται ότι πιο πολύ τυχαίο είναι ποιοι λόγοι είναι πάνω και ποιοι κάτω από αυτή την κρίσιμη τιμή.

**Πίνακας 3:** Στατιστικά στοιχεία για τη σύγκριση του *firing rate* \* με τη μέση ενεργοποίηση του νευρώνα για όλα τα δείγματα.

Στοιχεία για το σύνολο του πίνακα	
Ελάχιστη τιμή	0.051
Μέγιστη τιμή	5.812
μ (σ)	1.032 (0.530)
Πλήθος λόγων > 1	631/1440
μ (σ) για πλήθος λόγων > 1	1.451 (0.518)



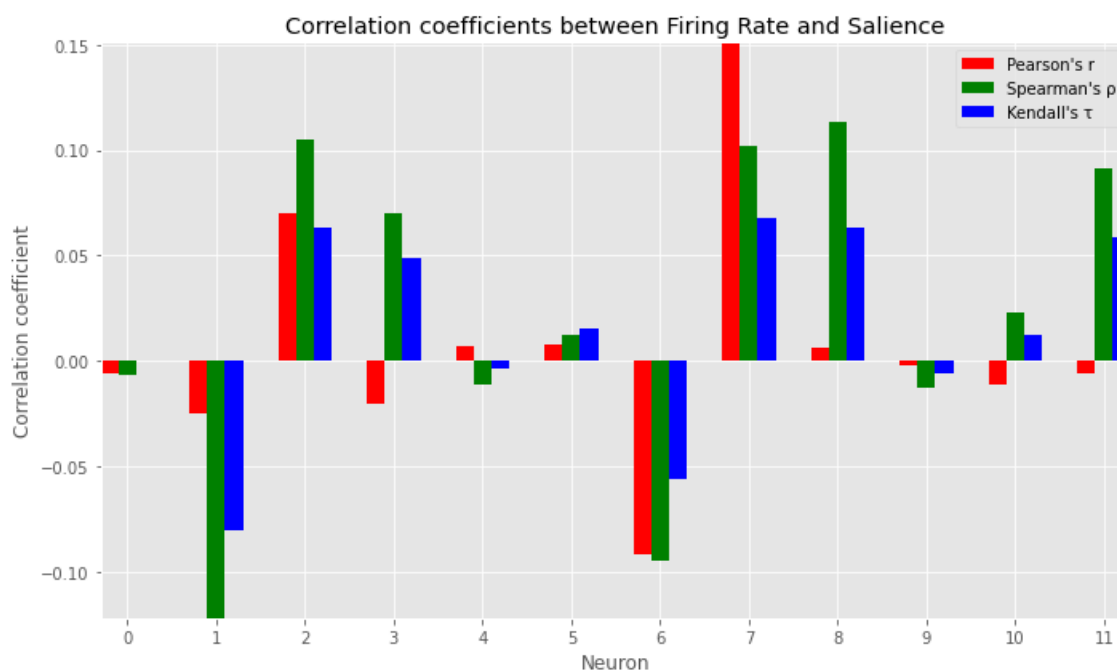
**Σχήμα 58:** Στατιστικά στοιχεία για τη σύγκριση του *firing rate* \* με τη μέση ενεργοποίηση του νευρώνα για όλα τα δείγματα, χωριστά για τους νευρώνες.

Στη δεύτερη σύγκριση η κατάσταση δε μεταβάλλεται σημαντικά, όπως μαρτυρά και η μορφή των διαγραμμάτων. Εδώ μελετάμε πού βρίσκεται η απόκριση *firing rate* \* συγκριτικά με τη μέση απόκριση του νευρώνα συνολικά για όλα τα δείγματα. Ο Πίνακας 3 μας λέει πως στο 43.8% των περιπτώσεων ισχύει  $\lambda > 1$ , ποσοστό για το οποίο σημειώνεται και μια μικρή άνοδος στη μέση τιμή σε σχέση με πριν. Στην Εικόνα 58 οι νευρώνες με τις υψηλότερες τιμές πριν, εξακολουθούν να υπερισχύουν και τώρα. Ούτε σε αυτή την περίπτωση τα αποτελέσματα είναι τέτοια ώστε να θεωρήσουμε πως ο αμφιβληστροειδής επηρεάζει με κάποιο προφανή τρόπο τη διαμόρφωση των χαρτών προεξοχής στο πείραμά μας.

Για να έχουμε μια πιο ολοκληρωμένη εικόνα, συγκρίναμε επιπλέον για κάθε κύτταρο χωριστά το διάνυσμα της απόκρισής του με το διάνυσμα των τιμών προεξοχής για τα δείγματα, μέσα από τρεις συντελεστές συσχέτισης. Αυτό έγινε αρχικά για τα 720 δείγματα εισόδου και αργότερα μόνο για τα 120 δείγματα με την υψηλότερη ένταση προεξοχής (Εικόνες 59, 60).



**Σχήμα 59:** Συντελεστές συσχέτισης μεταξύ διανυσμάτων απόκρισης και προεξοχής για όλα τα δείγματα στο σύνολο ελέγχου.



**Σχήμα 60:** Συντελεστές συσχέτισης μεταξύ διανυσμάτων απόκρισης και προεξοχής για το τμήμα κάθε εικόνας με τη μέγιστη τιμή προεξοχής.

Στην Εικόνα 59 φαίνονται οι τρεις συντελεστές για όλα τα δείγματα ελέγχου. Οι τιμές δεν ξεπερνούν το 0.25. Τη μέγιστη τιμή έχει ο νευρώνας 7 ο οποίος είχε και τα καλύτερα αποτελέσματα στην προηγούμενη σύγκριση. Η πλειοψηφία των νευρώνων έχει θετικούς συντελεστές δηλαδή η απόκρισή τους δείχνει μια τάση μεταβολής στην ίδια κατεύθυνση με την προεξοχή. Βλέπουμε όμως και αρνητικούς συντελεστές. Αυτό μας λέει πως κάποιιοι νευρώνες μπορεί να δείχνουν τα σημαντικά σημεία με μείωση της απόκρισης. Σύμφωνα με την Εικόνα 54 έχουμε 8 νευρώνες, τους 0, 2, 3, 7, 8, 9, 10, 11, που καταστέλλονται, δηλαδή για όλα τα  $N > 0$  η μέση απόκριση βρίσκεται κάτω από την εκάστοτε τιμή ηρεμίας. Για  $N = 15$  αυτή η μέση τιμή είναι χαμηλότερη από την αντίστοιχη των νευρώνων που δεν καταστέλλονται. Αυτό δεν επηρεάζει κάπως τη διακύμανση η οποία παραμένει αυξημένη, απλώς οι αποκρίσεις κυμαίνονται γύρω από μια χαμηλότερη τιμή. Οι ίδιοι ακριβώς νευρώνες είναι αυτοί που έχουν τους θετικούς συντελεστές. Επίσης είναι αυτοί που έχουν το μεγαλύτερο πλήθος  $\lambda > 1$  σε σχέση με τους υπόλοιπους. Φαίνεται πως αυτοί οι νευρώνες επιλέγουν να αυξήσουν το ρυθμό τους όταν βλέπουν προεξέχουσες περιοχές αλλά παράλληλα έχουν χαμηλώσει το συνολικό επίπεδο της δραστηριότητάς τους. Ο ρυθμός ενεργοποίησης των υπολοίπων ακολουθεί αντίστροφη πορεία με την προεξοχή, αλλά συνολικά η μέση δραστηριότητά τους είναι εντονότερη. Φυσικά οι συντελεστές είναι χαμηλοί για να ισχυριστούμε κάτι τέτοιο με βεβαιότητα, ωστόσο προέκυψε μια συμπεριφορά που αξίζει να αναφερθεί.

Στη δεύτερη δοκιμή που αφορά λιγότερα δείγματα (Εικόνα 60), οι συντελεστές είναι ακόμα χαμηλότεροι. Αναμενόμενα, νευρώνες με αρνητικές τιμές είτε παραμένουν με αρνητικές τιμές ή κυμαίνονται γύρω από το 0. Και στα δύο διαγράμματα ο συντελεστής  $\rho$

είναι μεγαλύτερος από τον  $\tau$ , μια λογική εξέλιξη αφού ο  $\tau$  είναι πιο ισχυρός και αυστηρός μιας και μετρά τη διαφορά μεταξύ σύμφωνων και ασύμφωνων ζευγών ενώ ο  $\rho$  εντοπίζει γενικότερα μονοτονική συμπεριφορά. Στις περισσότερες περιπτώσεις είναι μεγαλύτερος και από τον  $r$ , που είναι επίσης λογικό αφού ο  $r$  υποθέτει πως τα μεγέθη ακολουθούν κανονική κατανομή, κάτι το οποίο δε γνωρίζουμε εκ των προτέρων αν ισχύει η όχι. Ο νευρώνας 7 αποτελεί την εξαίρεση αφού, κυρίως στη δεύτερη δοκιμή, φαίνεται να προσεγγίζεται καλύτερα από γραμμική συνάρτηση.

Όπως και να χει, οι συντελεστές συσχέτισης δεν ανέδειξαν κάτι μη αναμενόμενο αλλά ενίσχυσαν ακόμα περισσότερο το συμπέρασμά μας. Τα γαγγλιακά κύτταρα δε φαίνεται να επηρεάζονται ιδιαίτερα από το μέτρο της προεξοχής και συνεπώς, στο πείραμά μας πάντα, ο αμφιβληστροειδής δε φαίνεται να συμμετέχει στη διαμόρφωση του χάρτη προεξοχής, τουλάχιστον με τον άμεσο και προφανή τρόπο της μεταβολής του ρυθμού απόκρισης προς την ίδια κατεύθυνση με την πιθανότητα ενός αντικειμένου να ελκύσει την προσοχή. Ο τρόπος με τον οποίο ο άνθρωπος θα επεξεργαστεί μια εικόνα, η επιλογή των σημείων στα οποία θα εστιάσει το βλέμμα του είναι συνδυασμός συμπεριφορικών στοιχείων υψηλού επιπέδου και χαρακτηριστικών χαμηλού επιπέδου όπως χρώμα, αντιθέσεις, προσανατολισμός, ένας μηχανισμός που υλοποιείται στα ανώτερα τμήματα του οπτικού συστήματος στον εγκέφαλο. Οι επιστήμονες δεν έχουν αποφανθεί ακόμη για το πού ακριβώς διαμορφώνεται η οπτική προσοχή και αν η εγκεφαλική περιοχή προέλευσής της είναι ίδια για όλα τα είδη. Μια άποψη είναι πως στα πρωτεύοντα, ο V1 είναι αυτός που δημιουργεί πρώτος ένα χάρτη προεξοχής αλλά όχι σε σπονδυλωτά όπως το ποντίκι. Μια από τις διαπιστώσεις που ενισχύουν το διαχωρισμό αυτό είναι πως η πλειοψηφία των γαγγλιακών κυττάρων στα πρωτεύοντα, που έχουν γραμμική συμπεριφορά και απλώς μεταβιβάζουν την πληροφορία, επικοινωνούν μόνο με τον LGN και συνεπώς με τον V1. Αντίθετα, σε άλλα είδη, στα οποία η πλειοψηφία γαγγλιακών κυττάρων είναι ανιχνευτές χαρακτηριστικών, αυτά επικοινωνούν με άλλη, ανώτερη περιοχή (superior colliculus). Συνεπώς, εφόσον η οπτική ανάλυση έχει μεταναστεύσει μέσα από την εξέλιξη σε διαφορετικές περιοχές, το ίδιο θα ισχύει και για τον υπολογισμό του χάρτη προεξοχής. Αυτό πιθανώς να έχει γίνει λόγω των διαφορετικών επεξεργαστικών αναγκών που έχουν τα διάφορα είδη, αναλογικά με τους υπολογιστικούς πόρους που έχουν διαθέσιμους. Γενικότερα, θεωρείται απίθανο να υπάρχουν σπονδυλωτά που διαμορφώνουν το χάρτη προεξοχής τους στον αμφιβληστροειδή. Στα πρωτεύοντα η ανάλυση που γίνεται εκεί δεν επαρκεί και στα κατώτερα σπονδυλωτά, παρόλο που ο αμφιβληστροειδής επιτελεί έντονη επεξεργασία, είναι αναγκαία επιπλέον ανάλυση από τον εγκέφαλο. Ίσως λοιπόν να είναι εξίσου απίθανο να συμβάλλει στη διαδικασία οπτικής προσοχής με έναν τόσο άμεσο και επιδραστικό τρόπο [79], [80].

Το μοντέλο υπολογίζει την προεξοχή σε φυσικές εικόνες, οι οποίες είναι αρκετά σύνθετες και μια γρήγορη εκτίμηση για το πού θα εστιάσει ένας παρατηρητής δεν είναι εύκολη, όπως θα ήταν π.χ. στην αναγνώριση μιας κόκκινης μπάλας μέσα σε ένα πλήθος από λευκές μπάλες. Στην πραγματικότητα, οι ερευνητές τώρα μαθαίνουν τη συμπεριφορά των γαγγλιακών κυττάρων σε εικόνες φυσικού περιβάλλοντος, μέσω τεχνικών βαθιάς μάθησης, συνεπώς πολλές πτυχές είναι παντελώς άγνωστες. Οι νευρώνες είναι ευαίσθητοι σε πολύ συγκεκριμένα χαρακτηριστικά, από απλά όπως η χωρική διαφορά της έντασης και ο

εντοπισμός ακμών σε πιο σύνθετα όπως κίνηση προς συγκεκριμένη κατεύθυνση. Εμείς, δε γνωρίζουμε ακριβώς ποια ερεθίσματα διεγείρουν τους νευρώνες του πειράματος και χρειάζονται πιο εξειδικευμένες μελέτες για να το μάθουμε. Χρειάζεται λοιπόν σύγκριση των χαρακτηριστικών που ορίζουν μια περιοχή ως πιθανότερη να τραβήξει την προσοχή με τη συμπεριφορά των αρμόδιων για την περιοχή αυτή νευρώνων στα χαρακτηριστικά αυτά. Σε αυτό σίγουρα θα βοηθήσει μια ολοκληρωμένη τεκμηρίωση πάνω στο πρόβλημα του ρόλου που τελικά επιτελεί ο αμφιβληστροειδής στην οπτική αντίληψη, όπως αναφέραμε στη θεωρία. Αν ο αμφιβληστροειδής είναι τελικά όσο πολύπλοκος θεωρείται πως είναι, εξίσου πολύπλοκος μπορεί να είναι και ο τρόπος με τον οποίο το γαγγλιακό κύτταρο ενημερώνει τον εγκέφαλο ότι εντόπισε πιθανό σημείο εστίασεως. Με άλλα λόγια, μπορεί ένας νευρώνας να καταστέλλεται στην παρουσία ενός χαρακτηριστικού που ελκύει το βλέμμα μας, δηλαδή να ενημερώνει τον εγκέφαλο με άλλο τρόπο και όχι με άνοδο στη συχνότητα απόκρισης. Δεδομένης και της δυσκολίας των φυσικών εικόνων, ίσως να είναι πιο διαφωτιστικό ένα πείραμα σε απλές τεχνητές εικόνες.

Πιθανές αιτίες εντοπίζονται και σε επιλογές και παραδοχές στην υλοποίησή μας, αφού ό,τι συμπεράσμα προέκυψε σχετικά με τη σύγκριση, προέκυψε μέσα στα πλαίσια του δικού μας πειράματος. Η σχετικά χαμηλή ακρίβεια του επιλεγμένου μοντέλου αμφιβληστροειδούς είναι πιθανό να επηρέασε τη διαδικασία. Ναι μεν αυτό δεν έδειξε να μας εμποδίζει από το να υπολογίσουμε τα υποδεκτικά πεδία, αλλά εδώ το πρόβλημα ίσως είναι πιο απαιτητικό για τα κύτταρα, αναδεικνύοντας το πόσο κομβικό είναι να έχουμε ένα μοντέλο που μιμείται αποτελεσματικά τον αμφιβληστροειδή. Μια επιπλέον παρατήρηση είναι πως στη διαμόρφωση του χάρτη προεξοχής διαδραματίζει σημαντικό ρόλο το χρώμα. Πολλές είναι οι φορές που ένα αντικείμενο ελκύει το βλέμμα μας απλώς επειδή ξεχωρίζει χρωματικά από τα γειτονικά του. Στο μοντέλο του αμφιβληστροειδούς ο παράγοντας αυτός δε λαμβάνεται υπόψη, οι νευρώνες εκπαιδεύονται σε μονοχρωματικά δεδομένα. Θεωρούμε πως αυτό αξίζει να μελετηθεί περαιτέρω.

Μια επιπλέον σκέψη που κάναμε είναι η προέλευση των δεδομένων μας. Οι αποκρίσεις με τις οποίες εκπαιδεύτηκε το μοντέλο αμφιβληστροειδούς προήλθαν από ποντικό και τα δεδομένα προεξοχής από ανθρώπινους παρατηρητές. Γνωρίζουμε πως η δομή του αμφιβληστροειδούς μεταξύ ανθρώπου και ποντικού διατηρείται χωρίς μεγάλες διαφορές, λόγος για τον οποίο το ποντίκι χρησιμοποιείται διαχρονικά στις εργαστηριακές μελέτες. Δε μπορούμε να πούμε το ίδιο ωστόσο και για τη λειτουργία του. Όπως είπαμε και στη θεωρία, οι επιστήμονες εξετάζουν αν σε είδη που οι υπολογιστικοί πόροι του εγκεφάλου είναι περιορισμένοι, όπως τα ποντίκια, και για λόγους επιβίωσης, ο αμφιβληστροειδής δίνει προτεραιότητα στην εξαγωγή σύνθετων χαρακτηριστικών, ενώ σε είδη που διαθέτουν ισχυρό ανώτερο οπτικό σύστημα (άνθρωπος) ο αμφιβληστροειδής επικεντρώνεται στη γραμμική και αποδοτική μετάδοση της πληροφορίας και δεν προβαίνει σε κάποιου είδους επεξεργασία. Η διαφοροποίηση αυτή μπορεί να συναντάται και στο ποσοστό στο οποίο τα διάφορα τμήματα του εγκεφάλου που είναι υπεύθυνα για την οπτική αντίληψη εξαρτώνται από σήματα που προέρχονται από τον οφθαλμό. Διαφορές στο πώς τα δύο είδη αντιλαμβάνονται τα ερεθίσματα και στο πώς οι νευρώνες τους αποκρίνονται σε αυτά θα επηρεάσουν το αποτέλεσμα. Παράλληλα, η διαφορετική λειτουργία ενδέχεται να επηρεάζει και τη



διαδικασία της οπτικής προσοχής. Η λογική λέει πως μηχανισμούς οπτικής προσοχής διαθέτουν ακόμη και τα χαμηλότερα σπονδυλωτά, καθώς είναι ο βασικός τρόπος η περιορισμένη επεξεργαστική ισχύς του εγκεφάλου να αφιερωθεί αποκλειστικά σε συγκεκριμένες, σημαντικές περιοχές. Παρόλα αυτά, λίγα πράγματα γνωρίζουμε για τον τρόπο με τον οποίο τα ποντίκια σαρώνουν το οπτικό πεδίο και τα κριτήρια με τα οποία δίνουν προτεραιότητα σε συγκεκριμένα αντικείμενα έναντι άλλων.

Παρατηρήσαμε πως αρκετές από τις εικόνες του Toronto δεν περιλαμβάνουν συγκεκριμένες περιοχές ενδιαφέροντος. Επαναλάβαμε λοιπόν το πείραμα στο MIT1003, ένα μεγαλύτερο σύνολο δεδομένων. Αποτελείται από 1003 έγχρωμες εικόνες, 779 από τις οποίες είναι τοπία και 228 πορträίτα. Η μέγιστη διάσταση κάθε εικόνας είναι 1024 pixels, ενώ η άλλη διάσταση ποικίλει μεταξύ 405 και 1024 pixels, με την πλειοψηφία των εικόνων να έχει διάσταση 768 pixels. Για τη λήψη των πραγματικών δεδομένων επιστρατεύτηκαν 15 παρατηρητές. Χρησιμοποιήσαμε ακριβώς την ίδια εκπαίδευση στο δίκτυο και οι μετρικές αξιολόγησης του SAM για το MIT1003 φαίνονται στην [72]. Δε σημειώθηκε κάποια αξιοσημείωτη μεταβολή στα αποτελέσματα συγκριτικά με τα πειράματα για το Toronto.

## Κεφάλαιο 5

### Επίλογος

#### 5.1 Σύνοψη και Συμβολή της Εργασίας

Στην παρούσα διπλωματική εργασία χρησιμοποιήσαμε δεδομένα ηλεκτροφυσιολογικών καταγραφών αποκρίσεων γαγγλιακών κυττάρων του ποντικού για να εκπαιδύσουμε ένα συνελκτικό νευρωνικό δίκτυο που μοντελοποιεί τον αμφιβληστροειδή. Οι καταγραφές αφορούσαν 24450 εικόνες από το φυσικό περιβάλλον και 60 κύτταρα. Μέσω του μοντέλου και της τεχνικής STA εντοπίσαμε τα υποδεκτικά πεδία, αφενός ως ένα επιπλέον εργαλείο αξιολόγησης των διαφορετικών υλοποιήσεων του μοντέλου και αφετέρου για να προσδιορίσουμε με έναν αξιόπιστο τρόπο την περιοχή ενδιαφέροντος των κυττάρων. Μέσα από αυτό το βήμα μας δόθηκε η ευκαιρία να μελετήσουμε εκτενώς τη μορφή των αποτελεσμάτων. Οι περισσότεροι νευρώνες δεν είχαν ευκρινές υποδεκτικό πεδίο και απορρίφθηκαν. Οι υπόλοιποι εμφάνισαν επικάλυψη μεταξύ τους, διφασική χρονική κυματομορφή και υποψία ανταγωνιστικής περιφέρειας γύρω από το κέντρο, ενώ φάνηκε να ανήκουν στην κατηγορία ON-OFF. Συμπεράναμε πως οι έγκυροι νευρώνες επηρεάζονται, κατά μέσο όρο, από μεταβολές στην είσοδο που συμβαίνουν τουλάχιστον 400 ms προτού πυροδοτήσουν.

Στη συνέχεια, προχωρήσαμε στη σύγκριση μεταξύ της απόκρισης των κυττάρων και του βαθμού προεξοχής σε άγνωστες εικόνες. Αφού καταλήξαμε στο χρονικό διάστημα για το οποίο θα προβάλλουμε το κάθε δείγμα, προχωρήσαμε σε κατάτμηση των εικόνων και εστίαση των τμημάτων τους πάνω στα υποδεκτικά πεδία. Οι χάρτες προεξοχής για το σύνολο των δεδομένων υπολογίστηκαν από ένα νευρωνικό δίκτυο, ήδη εκπαιδευμένο σε αυτό το πρόβλημα. Μελετήσαμε αν η τιμή προεξοχής μιας περιοχής επηρεάζει την ενεργοποίηση ενός νευρώνα που εστιάζει στην περιοχή αυτή. Συγκρίναμε το ρυθμό πυροδότησης για το τμήμα με τη μέγιστη προεξοχή με τη μέση δραστηριότητα κάθε νευρώνα μέσω του λόγου τους  $\lambda$  και εξετάσαμε αν υπάρχει κάποια συσχέτιση ανάμεσα στους τρόπους που μεταβάλλονται τα δύο μεγέθη, επιστρατεύοντας το συντελεστή συσχέτισης συνδιακύμανσης Pearson και τους συντελεστές συσχέτισης κατάταξης Spearman και Kendall. Οι λόγοι  $\lambda$  διαφοροποιήθηκαν μεταξύ των κυττάρων αλλά συνολικά παρέμειναν σε σχετικά χαμηλές τιμές, όπως και οι συντελεστές συσχέτισης, οι οποίοι ήταν αλλού θετικοί και αλλού αρνητικοί. Διαπιστώσαμε πως οι νευρώνες που επιλέγουν να αυξήσουν το ρυθμό τους όταν βλέπουν προεξέχουσες περιοχές χαμηλώνουν τη μέση δραστηριότητά τους κάτω από την τιμή στην οποία προσαρμόζονται στην κατάσταση ηρεμίας. Οι υπόλοιποι, για τους οποίους ο ρυθμός ενεργοποίησης ακολουθεί αντίστροφη πορεία από την ένταση προεξοχής, παρουσιάζουν κατά μέσο όρο υψηλότερες τιμές συχνότητας. Οι χαμηλοί συντελεστές

συσχέτισης, όπως και οι λόγοι, δεν μας επέτρεψαν να αποδείξουμε ότι το συμπέρασμα αυτό είναι αρκετά ισχυρό ώστε να γενικευθεί για την πλειοψηφία των δειγμάτων ελέγχου.

Τα πειράματα έγιναν στο πλαίσιο της προσθετικής αμφιβληστροειδούς. Η συμβολή της εργασίας εντοπίζεται κυρίως σε δύο σημεία: στην ποιοτική αξιολόγηση ενός μοντέλου μέσα από την ικανότητά του να αναπαράγει βιολογικά φαινόμενα (εντοπισμός κυττάρων στο χώρο) χωρίς πρότερη γνώση σχετικά με τα πειραματικά δεδομένα και στην προσπάθεια σύγκρισης του σήματος που παράγει ο αμφιβληστροειδής με την πιθανότητα ένα αντικείμενο να τραβήξει την προσοχή μας, αμιγώς μέσω δικτύων βαθιάς μάθησης.

Καταλήγοντας, ανακεφαλαιώνουμε τα κύρια αποτελέσματα και συμπεράσματα που προέκυψαν μέσα από τα πειράματα που πραγματοποιήσαμε:

- Προκειμένου τα συνελκτικά νευρωνικά δίκτυα να ενσωματωθούν σε εμφυτεύματα αμφιβληστροειδούς δεν αρκεί να αξιολογούνται μόνο σε μετρικές αξιολόγησης της σύγκλισής τους. Είναι απαραίτητο να εξεταστεί η ικανότητά τους να αναπαράγουν βιολογικά φαινόμενα όπως ο υπολογισμός του υποδεκτικού πεδίου. Αυτό θα αποτελέσει ισχυρή ένδειξη ότι προσομοιώνουν πιστά την απόκριση του αμφιβληστροειδούς. Διαπιστώσαμε πως η κρίσιμη παράμετρος ήταν το βραχυπρόθεσμο χρονικό διάστημα στο παρελθόν στο οποίο θα αναζητήσουμε τι είδους αλλαγές διεγείρουν τα κύτταρα. Για τιμές 0.2 και 0.3 το δίκτυο αδυνατούσε να ερμηνεύσει ικανοποιητικά τη σχέση διέγερσης-απόκρισης, παρόλο που έφτανε σε υψηλές αποδόσεις. Για τιμή 0.4 καταφέραμε να αναδείξουμε τα υποδεκτικά πεδία, καταλήγοντας στο συμπέρασμα πως η μνήμη των νευρώνων βρίσκεται πλησιέστερα στην τιμή αυτή, με την οποία το δίκτυο μιμείται καλύτερα τη λειτουργία για την οποία το δημιουργήσαμε.
- Το επιλεγμένο μοντέλο δεν έφτανε στα επιθυμητά επίπεδα απόδοσης στις μετρικές. Ωστόσο, η εκπαίδευση σε φυσικές εικόνες του επέτρεψε να προχωρήσει σε ικανοποιητικές εκτιμήσεις στο λευκό θόρυβο. Το αποτέλεσμα ήταν τα υποδεκτικά πεδία να προσεγγίζουν επαρκώς την αναμενόμενη μορφή, δεδομένου μάλιστα πως είχαμε καταγραφές από ένα μόνο πέρασμα των εικόνων.
- Για την πλειοψηφία των κυττάρων δεν ήταν δυνατός ο υπολογισμός του υποδεκτικού τους πεδίου. Ακόμη και για αυτούς που ήταν, τα πεδία παρουσίασαν επικάλυψη μεταξύ τους. Αυτά θεωρήσαμε πως ήταν θέματα που οφείλονται στην πειραματική διαδικασία λήψης των δεδομένων.
- Ο νευρώνας ενδιαφέρεται και για μεταβολές στο χρόνο. Δείχνοντας του διαρκώς το ίδιο μοτίβο, είτε αυτό είναι σκοτάδι είτε είναι η ίδια εικόνα, μειώνει το ρυθμό ενεργοποίησής του. Προσαρμόζεται στη σταθερή διέγερση που δέχεται και επιλέγει να αντιδράσει μονάχα όταν είναι απαραίτητο. Με αυτό τον τρόπο κωδικοποιεί πιο αποδοτικά την πληροφορία.

- Η αξιοποίηση τεχνικών βαθιάς μάθησης είναι ο πλέον αποτελεσματικός τρόπος ενσωμάτωσης χαρακτηριστικών από την κορυφή προς τη βάση στη διαμόρφωση του χάρτη οπτικής προσοχής.
- Τα αποτελέσματα του τελευταίου σταδίου της εργασίας δεν ήταν τέτοια ώστε να εξάγουμε κάποιο ασφαλές συμπέρασμα σχετικά με το αν η προεξοχή ενός αντικειμένου επιδρά στο ρυθμό απόκρισης ενός νευρώνα. Δεν παρουσιάστηκε κάποια συμπεριφορά που να μαρτυρά ξεκάθαρα πως το σημείο στο οποίο είναι πιθανότερο να εστιάσουμε την προσοχή μας θα ενεργοποιήσει εντονότερα το γαγγλιακό κύτταρο που εστιάζει σε αυτό, σε σχέση με τα υπόλοιπα σημεία.

## 5.2 Μελλοντικές κατευθύνσεις

Τα εμπορικά εμφυτεύματα αμφιβληστροειδούς δε βρίσκονται ακόμη στο επιθυμητό στάδιο αποτελεσματικότητας. Οι λόγοι είναι πολλοί, από αναπάντητα ερωτήματα σχετικά με τη λειτουργία του οπτικού συστήματος μέχρι εμπόδια μηχανικής και μικροηλεκτρονικής φύσεως. Στο πλαίσιο αυτό, αφορμώμενοι από τα αποτελέσματα της εργασίας μας, προτείνουμε ορισμένες μελλοντικές κατευθύνσεις που θεωρούμε πως θα βοηθήσουν τον τομέα της προσθετικής αμφιβληστροειδούς και μπορούν να ληφθούν υπόψιν.

Αρχικά, η χρήση τεχνικών βαθιάς μάθησης μόνο θετικά αποτελέσματα θα επιφέρει. Τα συνελκτικά νευρωνικά δίκτυα μπορούν να μιμηθούν και να προβλέψουν με υψηλή ακρίβεια την απόκριση του αμφιβληστροειδούς και συνεπώς μπορούν να ενσωματωθούν σε συστήματα προσομοίωσης της οφθαλμικής λειτουργίας. Η εκπαίδευση σε φυσικές εικόνες μπορεί να αποκαλύψει πολύ καλύτερα το πώς τα γαγγλιακά κύτταρα λειτουργούν στην πραγματική ζωή. Προς αυτή την κατεύθυνση, θεωρούμε αναγκαία τη συλλογή δεδομένων με καταγραφές αποκρίσεων σε πολύ μεγάλο όγκο φυσικών εικόνων, για πολλές δοκιμές, από διάφορα είδη οργανισμών και για πολυπληθείς ομάδες νευρώνων, δεδομένα που οφείλουν να είναι διαθέσιμα για όλους τους ερευνητές. Στο ίδιο πλαίσιο, μπορούν να αξιοποιηθούν και νευρωνικά δίκτυα αναγνώρισης αντικειμένων που θα ενσωματώσουν χαρακτηριστικά από την κορυφή προς τη βάση στον υπολογισμό της προεξοχής.

Μια επιπλέον μελλοντική κατεύθυνση περιλαμβάνει την εφαρμογή ανάλυσης STA και STC για τον προσδιορισμό του βαθμού αλλαγής στη μορφή των υποδεκτικών πεδίων του εκφυλισμένου αμφιβληστροειδούς σε σύγκριση με τον υγιή αμφιβληστροειδή, καθώς και του τρόπου με τον οποίο οι οπτικές πληροφορίες που κωδικοποιούνται από τα γαγγλιακά κύτταρα μεταβάλλονται ανάλογα με το στάδιο εκφύλισης [76].

Αναφορικά με τη σύγκριση προεξοχής και απόκρισης του αμφιβληστροειδούς, η προσπάθεια θα μπορούσε να εστιαστεί στον τρόπο με τον οποίο οι νευρώνες μεταδίδουν την πληροφορία. Θα μπορούσε να μελετηθεί το ενδεχόμενο ένας νευρώνας να δείχνει ότι ενδιαφέρεται για σημείο με υψηλή πιθανότητα να τραβήξει την προσοχή μας πυροδοτώντας νωρίτερα από τους υπόλοιπους και όχι αυξάνοντας ή μειώνοντας τη συχνότητα των δυναμικών που στέλνει. Χρήσιμο θα ήταν τα δεδομένα να είναι από γαγγλιακά κύτταρα,

πολλά στον αριθμό, που τα υποδεκτικά τους πεδία σχηματίζουν ένα πλέγμα πάνω στο οπτικό πεδίο προκειμένου να μπορούν να συγκριθούν οι αποκρίσεις τους πάνω στην ίδια εικόνα και να αναλυθεί η αλληλεπίδραση μεταξύ τους, ως μονάδες ενός συνόλου. Η προεξοχή είναι μια ιδιότητα που ένα αντικείμενο έχει μόνο αν συγκριθεί με τα γειτονικά του και με το συνολικό περιεχόμενο του περιβάλλοντος που βρίσκεται, που σημαίνει πως απαιτείται σύγκριση διαφορετικών οπτικών περιοχών και διαφορετικών νευρώνων. Να πούμε επίσης ότι καταγραφές από πρωτεύοντα, που είναι πιο κοντά στον άνθρωπο, ίσως είναι καταλληλότερες στο συγκεκριμένο πρόβλημα. Όπως και να χει, η επέκταση της σύγκρισης που δοκιμάσαμε θα μπορούσε δυνητικά να συμβάλλει στην εξαγωγή ισχυρών συμπερασμάτων και στη βελτίωση των αποτελεσμάτων.

Τέλος, μπορούν να γίνουν και επιπλέον πειράματα στα δεδομένα που χρησιμοποιήσαμε εμείς. Η μέθοδος STC που αναφέραμε και παραπάνω θα εμφανίσει περισσότερες λεπτομέρειες για τη λειτουργία των νευρώνων, όπως και η επέκταση του χρονικού ορίου στο οποίο κοιτάμε πίσω στο χρόνο πέρα από τα 400 ms. Φυσικά, οι επιλογές αυτές απαιτούν και πολύ αυξημένο υπολογιστικό κόστος.

## Βιβλιογραφία

- [1] Κωνσταντίνα Σπ. Νικήτα, *Προσομοίωση Φυσιολογικών Συστημάτων*. Αθήνα: Εκδόσεις ΤΖΙΟΛΑ, 2010.
- [2] J.C. Martins and L.A. Sousa. *Bioelectronic Vision: Retina Models, Evaluation Metrics, and System Design*. Series on bioengineering and biomedical engineering. World Scientific, 2009.
- [3] R. H. Masland, “The fundamental plan of the retina,” *Nature Neuroscience*, vol. 4, no. 9, pp. 877-886, Sep. 2001.
- [4] R. H. Masland, “Cell populations of the retina: The proctor lecture,” *Investigative Ophthalmology and Visual Science*, vol. 52, no. 7, pp. 4581–4591, Jun. 2011.
- [5] H. Kolb, E. Fernandez, R. Nelson *et al*, *Webvision: The Organization of the Retina and Visual System*. Salt Lake City, UT: University of Utah Health Sciences Center, 1995.
- [6] R. H. Masland, “The Neuronal Organization of the Retina,” *Neuron*, vol. 76, no. 2. Cell Press, pp. 266–280, 18-Oct-2012.
- [7] Peter Schiller, and M. Brown. *9.04 Sensory Systems*. Fall 2013. Massachusetts Institute of Technology: MIT OpenCourseWare, <https://ocw.mit.edu>. License: Creative Commons BY-NC-SA.
- [8] L. F. Abbott and P. Dayan, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA, USA: The MIT Press, 2005.
- [9] “Ganglion Cell - an overview | ScienceDirect Topics.” [Online]. Available: <https://www.sciencedirect.com/topics/neuroscience/ganglion-cell>. [Accessed: 26-Dec-2020].
- [10] S. W. Kuffler, “Discharge patterns and functional organization of mammalian retina,” *Journal of Neurophysiology*, vol. 16, no. 1, pp. 37–68, Jan. 1953.
- [11] L. Spillmann, “Receptive fields of visual neurons: The early years,” *Perception*, vol. 43, no. 11, pp. 1145–1176, 2014.
- [12] P. H. Schiller and E. J. Tehovnik, “Visual prosthesis,” *Perception*, vol. 37, no. 10, pp. 1529–1559, 2008.

- [13] X. Pitkow and M. Meister, “Decorrelation and efficient coding by retinal ganglion cells,” *Nature Neuroscience*, vol. 15, no. 4, pp. 628–635, Apr. 2012.
- [14] S. A. Baccus and M. Meister, “Fast and slow contrast adaptation in retinal circuitry,” *Neuron*, vol. 36, no. 5, pp. 909–919, Dec. 2002.
- [15] T. Baden, P. Berens, K. Franke, M. Román Rosón, M. Bethge, and T. Euler, “The functional diversity of retinal ganglion cells in the mouse,” *Nature*, vol. 529, no. 7586, pp. 345–350, Jan. 2016.
- [16] J. R. Sanes and R. H. Masland, “The Types of Retinal Ganglion Cells: Current Status and Implications for Neuronal Classification,” *Annual Review of Neuroscience*, vol. 38, pp. 221–246, Jul. 2015.
- [17] B. Roska and M. Meister, “The Retina Dissects the Visual Scene into Distinct Features,” in *The New Visual Neurosciences*, John S. Werner and Leo M. Chalupa, Eds. MIT Press, 2014, ch. 13, pp. 163-182.
- [18] T. Gollisch and M. Meister, “Eye Smarter than Scientists Believed: Neural Computations in Circuits of the Retina,” *Neuron*, vol. 65, no. 2. NIH Public Access, pp. 150–164, 28-Jan-2010.
- [19] J. Lindsey, S. A. Ocko, S. Ganguli, and S. Deny, “A Unified Theory of Early Visual Representations from Retina to Cortex through Anatomically Constrained Deep CNNs,” *arXiv: 1901.00945*, Jan. 2019.
- [20] E. J. Chichilnisky, “A simple white noise analysis of neuronal light responses,” *Network Computation in Neural Systems*, vol. 12, no. 2, pp. 199–213, 2001.
- [21] Y. Zhang, I. J. Kim, J. R. Sanes, and M. Meister, “The most numerous ganglion cell type of the mouse retina is a selective feature detector,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 36, pp. E2391–E2398, Aug. 2012.
- [22] J. D. Weiland and M. S. Humayun, “Retinal prosthesis,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1412–1424, 2014.
- [23] L. Yue, J. D. Weiland, B. Roska, and M. S. Humayun, “Retinal stimulation strategies to restore vision: Fundamentals and systems,” *Progress in Retinal and Eye Research*, vol. 53. Elsevier Ltd, pp. 21–47, 01-Jul-2016.
- [24] T. Guenther, N. H. Lovell, and G. J. Suaning, “Bionic vision: System architectures - A review,” *Expert Review of Medical Devices*, vol. 9, no. 1, pp. 33–48, Jan. 2012.

- [25] J. R. Golden *et al.*, “Simulation of visual perception and learning with a retinal prosthesis,” *bioRxiv*, p. 206409, Jan. 2018.
- [26] I. Fine and G. M. Boynton, “Pulse trains to percepts: The challenge of creating a perceptually intelligible world with sight recovery technologies,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 370, no. 1677, Sep. 2015.
- [27] S. Nirenberg and C. Pandarinath, “Retinal prosthetic strategy with the capacity to restore normal vision,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 37, pp. 15012 LP – 15017, Sep. 2012.
- [28] J. Walker, “Human saccadic eye movements,” *Scholarpedia*, vol. 7, no. 7, p. 5095, 2012.
- [29] L. Itti, “Visual salience,” *Scholarpedia*, vol. 2, no. 9, p. 3327, 2007.
- [30] B. W. Tatler, N. J. Wade, H. Kwan, J. M. Findlay, and B. M. Velichkovsky, “Yarbus, eye movements, and vision,” *i-Perception*, vol. 1, no. 1, pp. 7–27, 2010.
- [31] L. Itti and C. Koch, “A saliency-based search mechanism for overt and covert shifts of visual attention,” in *Vision Research*, 2000, vol. 40, no. 10–12, pp. 1489–1506.
- [32] S. Krasovskaya and W. J. Macinnes, “Saliency models: A computational cognitive neuroscience review,” *Vision (Switzerland)*, vol. 3, no. 4, p. 56, Dec. 2019.
- [33] C. Koch and S. Ullman, *Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry*, Dordrecht: Springer Netherlands, 1987, pp 115–141.
- [34] J. Tilke, K. Ehinger, F. Durand, and A. Torralba, “Learning to predict where humans look,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2009, pp. 2106–2113.
- [35] A. Borji, “Saliency Prediction in the Deep Learning Era: Successes, Limitations, and Future Challenges,” *arXiv: 1810.03716*, 08-Oct-2018.
- [36] J. Tsotsos and A. Rothenstein, “Computational models of visual attention,” *Scholarpedia*, vol. 6, no. 1, p. 6201, 2011.
- [37] E. Vig, M. Dorr, and D. Cox, “Large-scale optimization of hierarchical features for saliency prediction in natural images,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2798–2805.



- [38] M. Kümmerer, L. Theis, and M. Bethge, "Deep Gaze I: Boosting Saliency Prediction with Feature Maps Trained on ImageNet," *3rd International Conference on Learning Representations, ICLR 2015 - Workshop Track Proceedings*, Nov. 2014.
- [39] M. Kümmerer, T. S. A. Wallis, L. A. Gatys and M. Bethge, "Understanding Low- and High-Level Contributions to Fixation Prediction," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, 2017, pp. 4799-4808
- [40] Nian Liu, J. Han, D. Zhang, Shifeng Wen and T. Liu, "Predicting eye fixations using convolutional neural networks," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015 pp. 362-370.
- [41] S. S. S. Kruthiventi, K. Ayush, and R. V. Babu, "DeepFix: A Fully Convolutional Neural Network for predicting Human Eye Fixations," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4446–4456, Oct. 2015.
- [42] M. Cornia, L. Baraldi, G. Serra, and R. Cucchiara, "Predicting Human Eye Fixations via an LSTM-based Saliency Attentive Model," *IEEE Transactions on Image Processing*, vol. 27, no. 10, pp. 5142–5154, Nov. 2016.
- [43] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
- [44] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning*. The MIT Press, 2016.
- [45] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [46] D. H. Hubel and T. N. Wiesel, "Receptive fields of single neurones in the cat's striate cortex," *The Journal of Physiology*, vol. 148, no. 3, pp. 574–591, Oct. 1959.
- [47] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *The Journal of Physiology*, vol. 160, no. 1, pp. 106–154, Jan. 1962.
- [48] Matt Carter, Jennifer Shieh, "Chapter 4 - Electrophysiology," in *Guide to Research Techniques in Neuroscience*, 2nd ed, Academic Press, 2015, ch. 4, pp. 89-115.
- [49] M. Weckstrom, "Intracellular recording," *Scholarpedia*, vol. 5, no. 8, p. 2224, 2010.
- [50] "Electrophysiology - Wikipedia." [Online][Accessed: 05-Dec-2020].

- [51] A. Lozano, C. Soto-Sánchez, J. Garrigós, J. J. Martínez, J. M. Ferrández, and E. Fernández, “A 3D Convolutional Neural Network to Model Retinal Ganglion Cell’s Responses to Light Patterns in Mice,” *International Journal of Neural Systems*, vol. 28, no. 10, Dec. 2018.
- [52] W. Gerstner, W. M. Kistler, R. Naud, and L. Paninski, *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition*. Cambridge: Cambridge University Press, 2014.
- [53] C. van Vreeswijk, “Stochastic Models of Spike Trains,” in *Analysis of Parallel Spike Trains*, Springer US, 2010, pp. 3–20.
- [54] A. Heitman, N. Brackbill, M. Greschner, A. Sher, A. Litke, and E. J. Chichilnisky, “Testing pseudo-linear models of responses to natural scenes in primate retina,” *bioRxiv*, p. 045336, Dec. 2016.
- [55] L. T. McIntosh, N. Maheswaranathan, A. Nayebi, S. Ganguli, and S. A. Baccus, “Deep learning models of the retinal response to natural scenes,” in *Advances in Neural Information Processing Systems*, 2016, vol. 29, pp. 1369–1377.
- [56] E. Batty *et al.*, “Multilayer Recurrent Network Models of Primate Retinal Ganglion Cell Responses,” Nov. 2016.
- [57] S. Wienbar and G. W. Schwartz, “The dynamic receptive fields of retinal ganglion cells,” *Progress in Retinal and Eye Research*, vol. 67. Elsevier Ltd, pp. 102–117, 01-Nov-2018.
- [58] N. Maheswaranathan *et al.*, “The dynamic neural code of the retina for natural scenes,” *bioRxiv*, p. 340943, Jan. 2019.
- [59] H. Tanaka, A. Nayebi, N. Maheswaranathan, L. McIntosh, S. A. Baccus, and S. Ganguli, “From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction,” *arXiv: 1912.06207*, Dec. 2019.
- [60] M. N. Geffen, S. E. J. de Vries, and M. Meister, “Retinal Ganglion Cells Can Rapidly Change Polarity from Off to On,” *PLoS Biology*, vol. 5, no. 3, p. e65, Mar. 2007.
- [61] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *32nd International Conference on Machine Learning, ICML 2015*, 2015, vol. 1, pp. 448–456.

- [62] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.
- [63] M. Abadi *et al.*, “TensorFlow: A system for large-scale machine learning,” *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016*, pp. 265–283, May 2016.
- [64] “GitHub - keras-team/keras: Deep Learning for humans.” [Online]. Available: <https://github.com/keras-team/keras>. [Accessed: 06-Dec-2020].
- [65] “Pearson correlation coefficient - Wikipedia.” [Online][Accessed: 07-Dec-2020].
- [66] “Fraction of variance unexplained - Wikipedia.” [Online][Accessed: 07-Dec-2020].
- [67] D. R. Cantrell, J. Cang, J. B. Troy, and X. Liu, “Non-Centered Spike-Triggered Covariance Analysis Reveals Neurotrophin-3 as a Developmental Regulator of Receptive Field Properties of ON-OFF Retinal Ganglion Cells,” *PLoS Computational Biology*, vol. 6, no. 10, p. e1000967, Oct. 2010.
- [68] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778, December 2015.
- [69] M. Jiang, S. Huang, J. Duan, and Q. Zhao, “SALICON: Saliency in Context,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1072–1080.
- [70] N. Bruce and J. Tsotsos, “Attention based on information maximization,” *Journal of Vision*, vol. 7, no. 9, pp. 950–950, Jun. 2007.
- [71] M. Cornia, L. Baraldi, G. Serra, and R. Cucchiara, “SAM: Pushing the limits of saliency prediction models,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2018 pp. 1971–1973.
- [72] “pyret: retinal data analysis in python — pyret 0.6.0 documentation.” [Online]. Available: <https://pyret.readthedocs.io/en/master/index.html>. [Accessed: 05-Dec-2020].
- [73] MIT Course 9.11: The Human Brain, Nancy Kanwisher (Spring 2018). Accessed: Jan 8, 2021. [Online Video]. Available: <https://www.youtube.com/playlist?list=PLyGKBDfknk-iAQx4Kw9JeVqspsbg77sfAK0>

- [74] M. Van Wyk, H. Wässle, and W. R. Taylor, “Receptive field properties of ON- and OFF-ganglion cells in the mouse retina,” *Visual Neuroscience*, vol. 26, no. 3, pp. 297–308, May 2009.
- [75] Q. Shi, P. Gupta, A. K. Boukhvalova, J. H. Singer, and D. A. Butts, “Functional characterization of retinal ganglion cells using tailored nonlinear modeling,” *Scientific Reports*, vol. 9, no. 1, pp. 1–12, Dec. 2019.
- [76] J. Ahn, B. Rueckauer, Y. Yoo, and Y. S. Goo, “New features of receptive fields in mouse retina through spike-triggered covariance,” *Experimental Neurobiology*, vol. 29, no. 1, pp. 38–49, 2020.
- [77] “Spearman's rank correlation coefficient - Wikipedia.” [Online][Accessed: 18-Jan-2021].
- [78] “Kendall rank correlation coefficient - Wikipedia.” [Online][Accessed: 18-Jan-2021].
- [79] L. Zhaoping, “From the optic tectum to the primary visual cortex: migration through evolution of the saliency map for exogenous attentional guidance,” *Current Opinion in Neurobiology*, vol. 40, pp. 94–102, Oct. 2016.
- [80] S. Treue, “Visual attention: the where, what, how and why of saliency,” *Current Opinion in Neurobiology*, vol. 13, no. 4, pp. 428–432, Aug. 2003.
- [81] N. Melanitis and K. S. Nikita, “Biologically-inspired image processing in computational retina models,” *Computers in Biology and Medicine*, vol. 113, p. 103399, Oct. 2019.
- [82] A. Alevizaki, N. Melanitis, and K. Nikita, “Predicting eye fixations using computer vision techniques,” in *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*, Oct 2019, pp. 309–315.