

**ΔΙΕΘΝΕΣ ΠΑΝΕΠΙΣΤΗΜΙΟ ΤΗΣ ΕΛΛΑΔΟΣ (ΔΙ.ΠΑ.Ε.)  
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ  
Π.Π.Σ.ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ Τ.Ε.**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**Μετασχηματισμός των EEG σημάτων σε εικόνες με τη χρήση  
μοντέλων βαθιάς μάθησης**

**Νικόλαος Γ.Δαμιανός**



**Επιβλέπων:** **Γεώργιος Παπακώστας, Καθηγητής**

**ΚΑΒΑΛΑ  
ΝΟΕΜΒΡΙΟΣ 2022**

# Περίληψη

Η παρούσα εργασία πραγματοποιήθηκε στο πλαίσιο του Προγράμματος Προπτυχιακών Σπουδών του Τμήματος Πληροφορικής, της σχολής Θετικών Επιστημών του Διεθνούς Πανεπιστημίου της Ελλάδος (Δι.Πα.Ε).

Στη πτυχιακή αυτή μελετήθηκε το πρόβλημα της κατανόησης του περιεχομένου των EEG σημάτων μέσω GAN (generative adversarial network), δηλαδή ο μετασχηματισμός τους σε εικόνες. Θα χρησιμοποιηθεί η Python ως γλώσσα προγραμματισμού και η βιβλιοθήκη Pytorch για την υλοποίηση του Νευρωνικού δικτιού.

Σε αυτή την εργασία το dataset πάρθηκε από <http://www.perceivelab.com/datasets> (EEG Data for Visual Classification).

**Λέξης Κλειδιά :** νευρωνικά δίκτυα, βαθιά μάθηση, Ηλεκτροεγκεφαλογράφημα (EEG), Παραγωγικό αντιπαραθετικό δίκτυο (GAN).

## ABSTRACT

In this thesis the problem of understanding the contents of the EEG signals with the implementation of GAN (Generative Adversarial Network) is investigated. It will be applied (GAN) so it will create images by ‘understanding’ the input EEG signals. Python will be the programming language used and the pytorch library (package).

The dataset was taken from <http://www.perceivelab.com/datasets> (EEG Data for Visual Classification).

**Keywords:** neural network, deep learning, Electroencephalography (EEG), Generative Adversarial Network (GAN)



## Συντομογραφίες

EEG: Ηλεκτροεγκεφαλογράφημα

GAN: Generative Adversarial Network (Παραγωγικό αντιπαραθετικό δίκτυο)

TNΔ: Τεχνητά Νευρωνικά Δίκτυα

MLP: Multilayer perceptron (πολύ επίπεδο perceptron)

BP: Backpropagation

STG: Stochastic Gradient Descent

LSTM: Long short-term memory

ReLU: Rectified Linear Unit

RNN: Recurrent Neural Networks

NLP: natural-language processing

FIR : Finite impulse response ( Πεπερασμένης Κρουστικής Απόκρισης)

IIR : Infinite Impulse Response (μη πεπερασμένη Κρουστικής Απόκρισης)

SNR : SIGNAL TO NOISE RATIO

FFT: Fast Fourier Transform

PSD: Power spectral density

SPWV : Smooth pseudo-Wigner-Ville

## Περιεχόμενα

Περίληψη .....	2
Συντομογραφίες .....	5
Πρόλογος .....	8
Κεφάλαιο 1 : EEG.....	9
1.1 Τι είναι τα EEG .....	9
1.2 Γιατι χρησιμοποιούνται τα EEG .....	10
1.3 Ιστορία των EEG.....	11
1.4 Πλεονεκτήματα των EEG .....	12
1.5 Μειονεκτήματα των EEG.....	12
1.6 Ζώνες EEG (EEG Bands) .....	13
1.7 Φιλτράρισμα EEG.....	16
1.7.1 Εισαγωγή .....	16
1.7.2 Βασικά στοιχεία για το ψηφιακό φιλτράρισμα.....	16
1.7.3 Είδη φίλτρων.....	16
2.1 Μέθοδοι εξαγωγής χαρακτηριστικών (feature extraction).....	20
1.8.1 Fast Fourier Transform (FFT).....	20
1.8.2 Wavelet Transform .....	20
1.8.3 Eigenvectors.....	22
1.8.4 Time-Frequency Distributions .....	23
1.8.5 Autoregressive Method.....	24
1.8.6 LSTM.....	24
Κεφάλαιο 2 : Generative Adversarial Networks.....	25
2.1 Τεχνητά Νευρωνικά δίκτυα .....	25
2.1.1 Τι είναι τα Τεχνητά Νευρωνικά δίκτυα .....	25
2.1.2 Πώς λειτουργεί ένα ΤΝΔ .....	26
2.1.3 Τι είναι ο BP .....	29
2.1.4 Είδη ΤΝΔ και βαθιάς μάθησης .....	29
2.2 GAN.....	33
2.2.1 Εισαγωγή .....	33
2.2.2 Ιστορία .....	34
2.2.3 Αρχιτεκτονική.....	34
2.2.4 Εκπαίδευση GAN .....	35

2.2.5 Παραλλαγές GAN.....	35
2.2.6 Που Χρησιμοποιούνται.....	37
Κεφάλαιο 3 : Προτεινόμενη μελέτη .....	39
3.1 Προσέγγιση .....	39
3.1.1 Feature extractor .....	40
3.1.2 Conditional GAN .....	42
3.1.3 Μετρήσεις .....	43
Κεφάλαιο 4 Πειράματα.....	46
4.1 Datasets .....	46
4.2 Αρχιτεκτονική του μοντέλου .....	49
4.2.1 LSTM.....	49
4.2.2 Conditional GAN .....	50
4.3 Κώδικας και δομή αρχείων .....	51
4.4 Αποτελέσματα και Μετρήσεις.....	54
4.4.1 Αποτελέσματα.....	54
4.4.2 Μετρήσεις .....	57
4.5 Συμπεράσματά .....	61
Βιβλιογραφία .....	63

## Πρόλογος

Στην εργασία αυτή θα αναλυθεί το πρόβλημα του μετασχηματισμού των EEG σημάτων σε εικόνες με την χρήση GAN.

Στο πρώτο κεφάλαιο θα αναλυθούν τα σήματα EEG. Δηλαδή ποια η μορφή των EEG, γιατί χρησιμοποιούνται, τα πλεονεκτήματα και μειονεκτήματα τους καθώς και διάφορες εφαρμογές τους.

Στο δεύτερο κεφάλαιο θα αναλυθούν τα GAN. Μια ιστορική αναδρομή, το πως λειτουργούν, το πως εκπαιδεύονται, διάφορες εφαρμογές τους.

Στο τρίτο κεφάλαιο θα συζητηθεί η προσέγγισή που εφαρμόστηκε στο πλαίσιο της πτυχιακής εργασίας, προς το πρόβλημα.

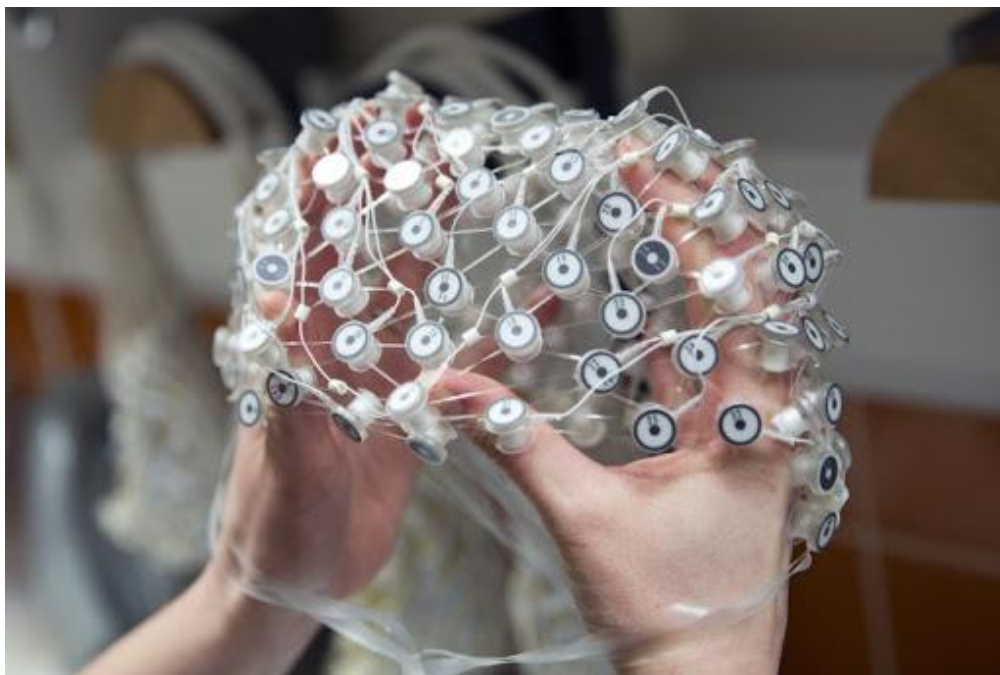
Και στο τέταρτο και τελευταίο κεφάλαιο θα συζητηθούν τα πειράματα, το σύνολο δεδομένων που χρησιμοποιήθηκε, τα αποτελέσματα και τα συμπεράσματα.



# Κεφάλαιο 1 : EEG

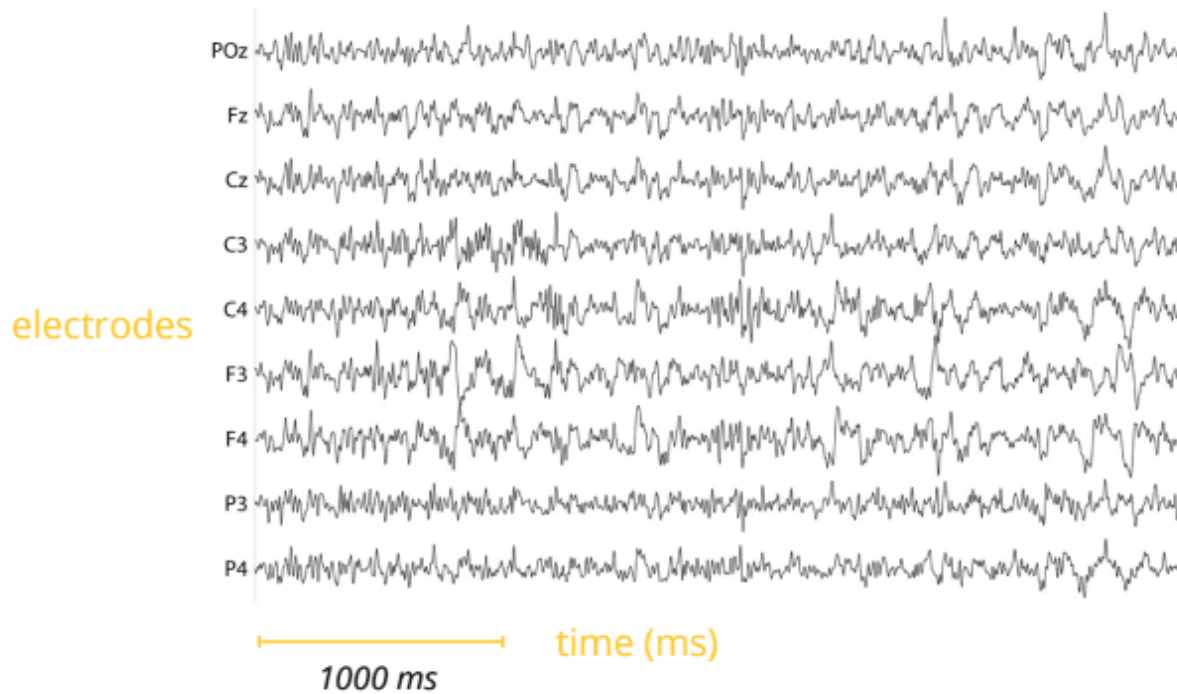
## 1.1 Τι είναι τα EEG

Το Ηλεκτροεγκεφαλογράφημα είναι ένα τεστ στο οποίο ανιχνεύει την ηλεκτρική δραστηριότητα του εγκεφάλου [1]. Δεν χρειάζεται επέμβαση και η ανίχνευση γίνεται με ηλεκτρόδια τοποθετημένα στο δέρμα της κεφαλής [2] (Εικόνα 1). Τα EEG ουσιαστικά μετράνε της διακυμάνσεις τάσης που παράγονται από τους νευρώνες του εγκεφάλου σε μία χρονική περίοδο από διάφορες περιοχές του εγκεφάλου.



*Εικόνα 1: Τα ηλεκτρόδια με τα οποία μετράμε της διακυμάνσεις τάσης*

Από τα ηλεκτρόδια λαμβάνονται οι κυματομορφές (Εικόνα 2) και μπορούν να εξεταστούν έτσι ώστε να μπορεί να διαγνωστεί αν υπάρχει κάποια ασθένεια ή επιληψία.



*Εικόνα 2: Ένα παράδειγμα κυματομορφών EEG*

## 1.2 Γιατι χρησιμοποιούνται τα EEG

Τα EEG βοηθάνε στο να γίνει διάγνωση για τις παρακάτω ασθένειες :

- Επιληψία
- Διαταραχές ύπνου
- Καρκίνο του εγκεφάλου (βοηθάει να βρεθούν όγκοι).
- Να βρεθούν τραυματισμοί στο κεφάλι
- Να βρεθούν δυσλειτουργίες του εγκεφάλου από διάφορες αιτίες
- Φλεγμονές του εγκεφάλου
- Εγκεφαλικά
- Τη νόσο του Alzheimer
- Επίσης στην επιβεβαίωση κάποιου που είναι σε κώμα.

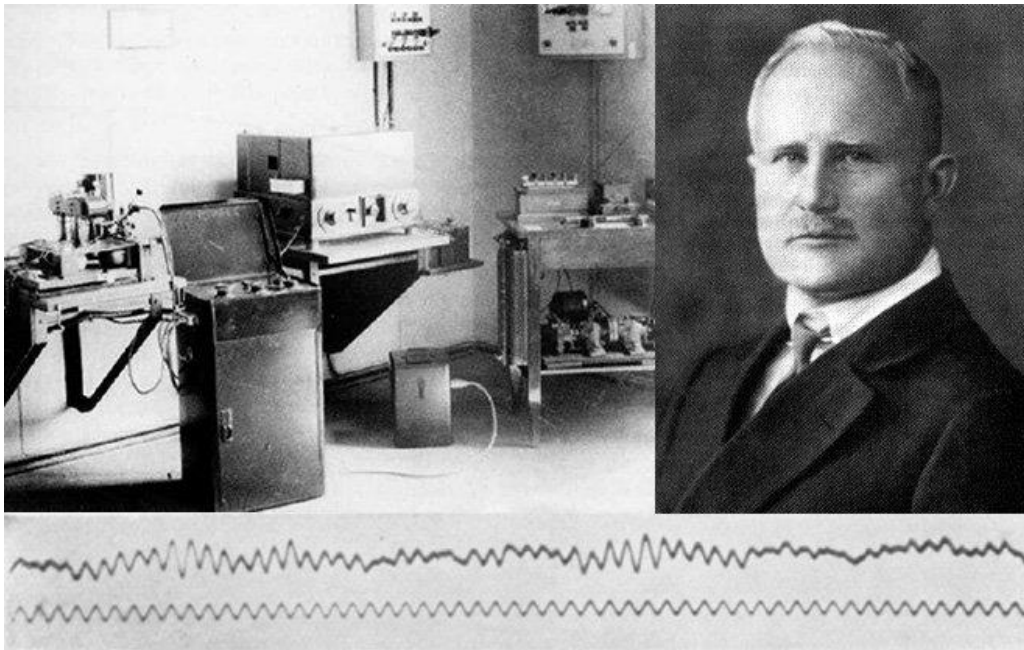
Επίσης έρευνες [3] έχουν δείξει ότι τα EEG μπορούν να χρησιμοποιηθούν για ταξινόμηση νοητικών καταστάσεων [4], ταξινόμηση κίνησης [5], ταξινόμηση συναισθημάτων [6] και με μικρή ακρίβεια αντικείμενα [7].

### 1.3 Ιστορία των EEG

Το 1875 ο Άγγλος γιατρός Ρίτσαρντ Κατον βρήκε στα ημισφαίρια του εγκεφάλου λαγών και πιθήκων ότι παράγονται ηλεκτρικά φαινόμενα. Μετά το 1890 ο Πολωνός γιατρός Αδόλφος Μπεκ δημοσίευσε μια έρευνα για αυθόρμητη ηλεκτρική δραστηριότητα του εγκεφάλου σκύλων και κουνελιών που περιλάμβανε ρυθμικές ταλαντώσεις που μεταβαλλόταν με το φως. Ο Μπεκ ξεκίνησε πειράματα σχετικά με την ηλεκτρική δραστηριότητα των εγκεφάλων ζώων. Τοποθέτησε ηλεκτρόδια απευθείας στην επιφάνεια του εγκεφάλου για να δοκιμάσει την αισθητηριακή διέγερση και βρήκε ότι παράγονται εγκεφαλικά κύματα.

Το 1914 ο Πολωνός φυσιολόγος Ναπολέων Κμπούλσκι έβγαλε φωτογραφία το πρώτο EEG πειραματικής επιληψίας.

Το 1924 το πρώτο ανθρώπινο EEG καταγράφηκε από τον Γερμανό ψυχίατρο και φυσιολόγο Χανς Μπέργκερ. Ο Χανς Μπέργκερ (Εικόνα 3) μετά έφτιαξε τον πρώτο ηλεκτροεγκεφαλογράφο τον οποίο ονόμασε αυτός. Και το περιέγραψε ως μία από τις πιο εκπληκτικές, αξιοσημείωτες και σημαντικές εξελίξεις στην ιστορία της Νευρολογίας.



*Εικόνα 3 Ο κύριος Hans Berger, Ο πρώτος Ηλεκτροεγκεφαλογράφος και το πρώτο ανθρώπινο EEG σήμα..*

Το 1947 δημιουργήθηκε το αμερικάνικο σωματείο για τα EEG και το πρώτο διεθνές συνέδριο για EEG πραγματοποιήθηκε. Η εταιρία Beckman Instruments κατασκεύασε σύστημα ηλεκτροεγκεφαλογραφίας που χρησιμοποιήθηκε στο Project Gemini για την παρακολούθηση των εγκεφαλικών κυμάτων των αστροναυτών κατά την πτήση. Το 1988 οι Stevo Bozinovski, Mihail Sestakov και Liljana Bozinovska δημοσίευσαν μία αναφορά σχετικά με τον έλεγχο ενός robot [8]. Τον Οκτώβριο του 2018 επιστήμονες σύνδεσαν τους εγκεφάλους τριών ανθρώπων για να δούνε άμα γίνεται ανταλλαγή σκέψεων ανάμεσα τους και το ποσοστό επιτυχίας ήταν 81 %.

## 1.4 Πλεονεκτήματα των EEG

Υπάρχουν και άλλες μέθοδοι μελέτης του εγκεφάλου όπως οι γνωστές fMRI(Μαγνητικός τομογράφος) και NMR(τομογράφος πυρηνικού μαγνητικού συντονισμού),κ.α.Τα EEG έχουν πολλά πλεονεκτήματα σε σχέση με μερικές από τις άλλες τεχνικές.

- Ένα από τα πιο σημαντικά πλεονεκτήματα των EEG είναι ότι ο εξοπλισμός τους είναι πολύ χαμηλότερος από της άλλες μεθόδους.Οι τιμές τους κυμαίνονται από 100\$ [9] έως 25000+\$.
- Ο εξοπλισμός των EEG είναι γενικά ‘μικρός’ οπότε μπορεί να χρησιμοποιηθεί μέσα ή έξω από το σπίτι.
- Τα EEG έχουν πολύ υψηλούς ρυθμούς δειγματοληψίας από 250 έως και 20000 Hz.Οι μόνες άλλες μέθοδοι,που δεν χρειάζεται επέμβαση,και μπορούν να φτάσουν στους ρυθμούς δειγματοληψίας είναι οι MEG και EROS αυτό βοηθάει στο να γίνει πιο καλή ανάλυση σήματος.
- Επειδή ο εξοπλισμός έχει μικρό μέγεθος,είναι πολύ ανεκτικό σε κίνησης μέχρι και περπάτημα.Υπάρχουν συσκευές που χρησιμοποιούν Bluetooth.
- Τα EEG αποτρέπουν την ‘περιορισμένη διαθεσιμότητα’ εξοπλισμού που μπορεί να έχουν οι άλλες μέθοδοι.(Επειδή είναι μικρότερα και φθηνότερα.)
- Δεν κάνουν θόρυβο οπότε δεν αποσπάται η προσοχή του ασθενή.
- Δεν εκθέτουν τον ασθενή σε μαγνητικά πεδία υψηλής έντασης(fMRI) ή σε ραδιοσυνδέτες (PET) .
- Στα EEG υπάρχει καλύτερη κατανόηση του ποιο σήμα μετριέται.
- Τέλος το EEG είναι ένα ισχυρό εργαλείο παρακολούθησης αλλαγών του εγκεφάλου σε διάφορες φάσεις της ζωής του.Η ανάλυση ύπνου με EEG μπορεί σημαντικές πτυχές του χρόνου ανάπτυξης του εγκεφάλου συμπεριλαμβανομένης της αξιολόγησης της ωρίμανσης του εφηβικού εγκεφάλου.

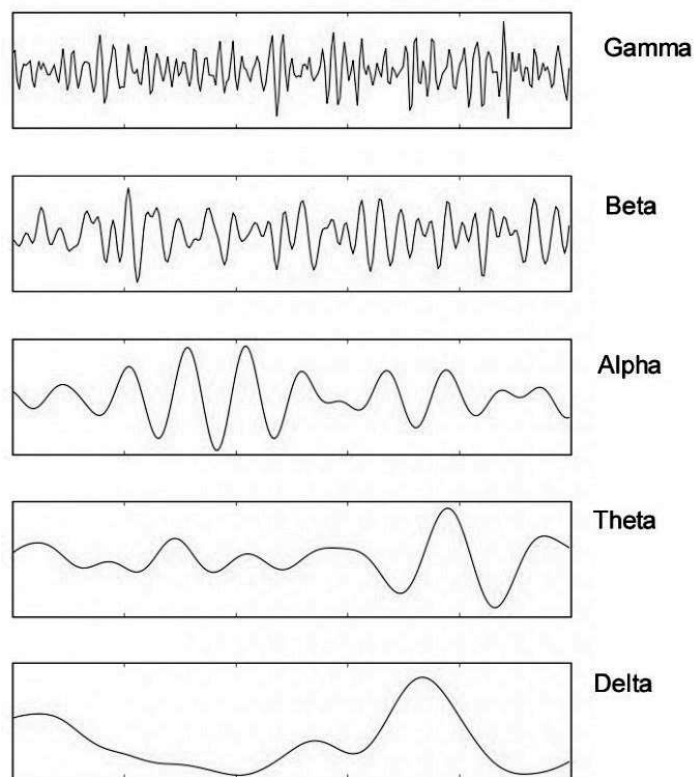
## 1.5 Μειονεκτήματα των EEG

- Η αναλογία σήματος προς θορύβου είναι πολύ χαμηλή,οπότε χρειάζεται πολύπλοκη ανάλυση δεδομένων και μεγάλο αριθμό από ασθενής ή εθελοντές για να παραχθούν χρήσιμες πληροφορίες.
- Το σημαντικότερο μειονέκτημα των EEG είναι η χαμηλή χωρική ανάλυση στο δέρμα της κεφαλής.Δηλαδή τα σήματα που παίρνονται είναι το άθροισμα του ηλεκτρικού πεδίου,στην κατεύθυνση του ηλεκτροδίου στο δέρμα της κεφαλής ,που παράγεται από πολλούς νευρώνες.Η χωρική ανάλυση ενός ηλεκτροδίου είναι σε τάξης μεγέθους του ενός εκατοστού όπου υπάρχουν χιλιάδες ή εκατομμύρια νευρώνες.Οπότε δυνατή ηλεκτρική δραστηριότητα μπορεί να διαβαστεί και από άλλα ηλεκτρόδια.Ουσιαστικά με τα EEG δεν μπορεί να εντοπιστεί η ακριβής πηγή της δραστηριότητας και δεν βοηθάει τους ερευνητές να καταλάβουν από που προέρχονται αυτές η δραστηριότητες για της οποίες ψάχνουν.
- Τα EEG δεν μπορούν να μετρήσουν κάτω από τα ανώτερα στρώματα του εγκεφάλου (των φλοιό).

## 1.6 Ζώνες EEG (EEG Bands)

Ένα EEG έχει διαφορετικές 'ζώνες' που ορίζονται από την συχνότητα των κυμάτων (Εικόνα 4). Αυτές οι ζώνες είναι 5 στον αριθμό και είναι οι εξής :

- Δέλτα : είναι αργά και η συχνότητα τους είναι  $< 4\text{Hz}$ .
- Θήτα : είναι ανάμεσα στις συχνότητες 4 έως 8 Hz.
- Άλφα : είναι ανάμεσα στις συχνότητες 8 έως 12 Hz.
- Βήτα : η οποία κυμαίνεται από 14 έως 30Hz.
- Γάμμα : οπου κυμαίνονται από 30 έως 80Hz.



*Εικόνα 4 Οι ζώνες EEG*

Κάθε μία από αυτές της ζώνες σχετίζεται με διαφορετικές καταστάσεις εγκεφαλικής δραστηριότητας. [10] (Εικόνα 5).

Η ζώνη άλφα συνήθως βρίσκονται σε ανθρώπους που είναι ξύπνιοι αλλά έχουν κλειστά τα μάτια τους και είναι σε κατάσταση ηρεμίας.

Η ζώνη θήτα σχετίζεται με την μνήμη και συναισθήματα.

Η δέλτα παρατηρούνται όταν το άτομο είναι σε βάθη ύπνο ή σε κόμμα.

Η βήτα όταν το άτομο είναι ικανό να επεξεργαστή πληροφορίες ή είναι σε εγρήγορση.

Η γάμμα σχετίζεται με την σκέψη και την συνείδηση δηλαδή τη δημιουργία συνδέσεων μεταξύ διαφόρων τμημάτων του εγκεφάλου προκειμένου να σχηματισουν έννοιες.



## Brainwave Frequencies

## Associated Mental Activity

### Gamma +40 Hz



- Whole-brain activity
- "Super-learning"
- Sudden Insight

### Beta 13-40 Hz



- Cognition
- Focus
- Analytical Thinking
- Stress/Anxiety (at higher frequencies)
- Fight or Flight Response (at higher frequencies)

### Alpha 8-12.9 Hz



- Relaxation
- Creativity
- Light Meditative/Trance State
- Increased Serotonin Production
- Threshold to Unconscious Mind

### Theta 4-7.9 Hz



- REM Sleep
- Increased Production of Catecholamines
- Access to Unconscious Mind
- Healing of Trauma and Addiction
- Deep Meditation
- Transcendent Mind States
- Lucid Dreaming

### Delta 0.2-3.9 Hz



- Sleep/Therapy for Sleep Disorders
- Accelerated Physical Healing
- Deep Relaxation

*Εικόνα 5 Οι ζώνες και οι δραστηριότητες του εγκεφάλου*

## 1.7 Φιλτράρισμα EEG

### 1.7.1 Εισαγωγή

Το φιλτράρισμα των EEG σημάτων είναι μια συνηθισμένη τακτική έτσι ώστε να αφαιρεθεί ο θόρυβος από αυτά. Τα EEG σήματα είναι συνήθως πολύ θορυβώδεις επειδή το SNR είναι πολύ χαμηλό. Αυτό οφείλεται στο ότι ο εγκέφαλος κάνει πολλές εργασίες και όταν μετριοούνται τα EEG σήματα τότε μετριοούνται και τις άλλες διεργασίες και όχι μόνο αυτές που είναι ενδιαφέρουσες για τη μετρήσει τους για παράδειγμα ανοιγοκλείσιμο των ματιών ή κινήσει κάποιοι μυ. Συνήθως χρησιμοποιούνται δύο φίλτρα ένα υψηλής διέλευσης (high-pass filter), για να φιλτραριστούν οι μικρές συχνότητες για παράδειγμα 2Hz και κάτω, και ένα φίλτρο χαμηλής διέλευσης (low-pass filter) το οποίο φιλτράρει της μεγάλες συχνότητες πχ πάνω από 60 Hz. Τα φίλτρα έχουν και μειονεκτήματα το πιο σημαντικό από αυτά είναι ό,τι μπορούν να παραμορφώσουν το σχήμα και την δομή του σήματος. [11].

### 1.7.2 Βασικά στοιχεία για το ψηφιακό φιλτράρισμα

Τα φίλτρα σχεδιάζονται σε FIR ή σε IIR. Η κρουστική απόκριση είναι ο τρόπος με τον οποίο το φίλτρο χειρίζεται τα σήματα στον τομέα του χρόνου δηλαδή το τι κάνει στην δομή του σήματος και ο μετασχηματισμός Fourier είναι γνωστός και ως απόκριση συχνότητας. Η κρουστική απόκριση είναι σημαντική στο να μπορέσει να χαρακτηριστεί η συμπεριφορά του φίλτρου.

Σε ένα φίλτρο FIR η κρουστική απόκριση είναι για πεπερασμένο χρόνο, μετά από αυτόν τον χρόνο η έξοδος γίνεται μηδέν και παράγει ίσες καθυστερήσεις σε όλες της συχνότητες. Αντιθέτως ένα IIR φίλτρο, που είναι γνωστά και ως αναδρομικά φίλτρα, έχουν άπειρη κρουστική απόκριση όπου ένα μέρος της εξόδου χρησιμοποιείται στην είσοδο. Αυτό παράγει άνισες καθυστερήσεις σε διαφορετικές συχνότητες. Αυτό σημαίνει ότι το σήμα εξόδου μετατοπίζεται στο χρόνο σε σχέση με το σήμα εισόδου, με μερικές συχνότητες να μετακινούνται περισσότερο από άλλες. Το μεγαλύτερο πλεονέκτημα των IIR είναι ότι είναι υπολογιστικά πιο αποτελεσματικά.

Ένα άλλο μέρος της σχεδίασης ενός φίλτρου είναι η κατεύθυνση του σήματος. Αυτά λέγονται αιτιατά φίλτρο (casual filter) και μη-αιτιατά φίλτρα (a-casual ή non casual filter).

### 1.7.3 Είδη φίλτρων

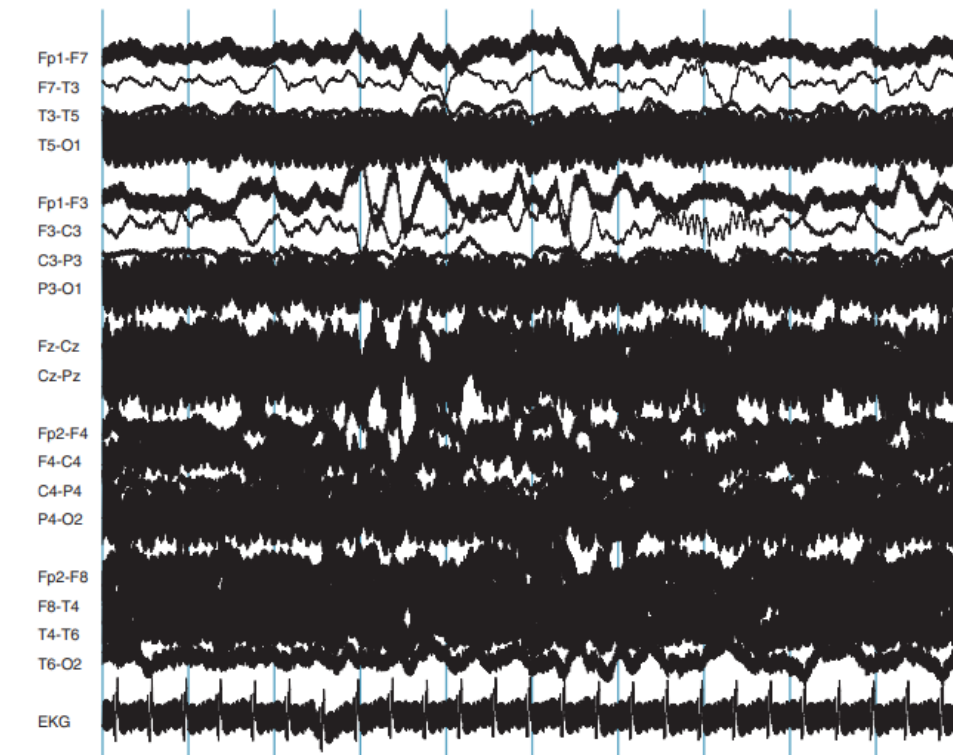
Όπως αναφέρθηκε παραπάνω υπάρχουν τα low pass φίλτρα που αφαιρούν της υψηλές συχνότητες, δηλαδή άμα παρθεί ένα low pass φίλτρο με 11Hz θα περνάνε συχνότητες από 0 έως 10Hz, Τα high pass που είναι ακριβώς τα αντιθετα από τα low pass και κόβουν της μικρές συχνότητες και τα φίλτρα εγκοπής τα οποία αποκλείουν το καθορισμένο εύρος συχνότητας. Στα EEG χρησιμοποιούνται για να αφαιρέσουν διάφορους θορύβους.



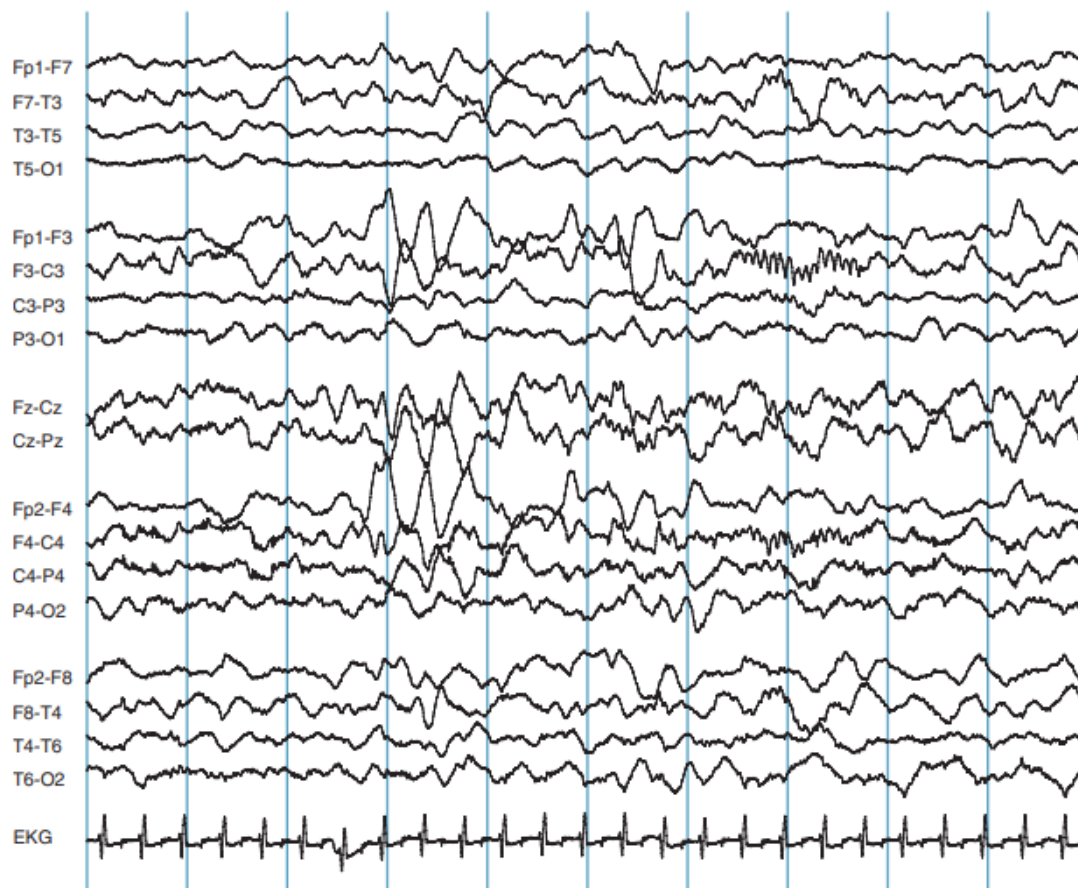
### 1.7.3.1 Notch φίλτρο

Το φίλτρο Notch ή στα ελληνικά φίλτρο εγκοπής είναι ίσως το πιο σημαντικό διότι χωρίς αυτό η ανάγνωση των EEG είναι αδύνατη (Εικόνα 6). Η δουλειά του είναι να αφαιρεί μια συγκεκριμένη συχνότητα. Επειδή το εναλλασσόμενο ρεύμα που παράγεται από της ηλεκτρικές πρίζες της ΗΠΑ είναι 60Hz παράγονται ηλεκτρικά πεδία τα οποία μολύνουν τα EEG έτσι χρησιμοποιείτε ένα φίλτρο εγκοπής τον 60 Hz για να αφαιρεθεί αυτό το ανεπιθύμητο σήμα (Εικόνα 7). Σε χώρες που τα παραγόμενα Hz είναι διαφορετικά χρησιμοποιείτε το αντιστοιχο φίλτρο εγκοπής. [12]

Μια ατέλεια που έχουν τα φίλτρα εγκοπής (η οποία δεν είναι και πολύ σημαντική ) όταν χρησιμοποιείτε είναι ότι αφαιρεί και της διπλανές συχνότητες για παράδειγμα 49Hz και 51Hz. Οι συγκεκριμένες συχνότητες δεν είναι πολύ σημαντικές διότι είναι πολύ γρήγορες.



Εικόνα 6 EEG σήματα χωρίς να περάσουν από φίλτρο εγκοπής



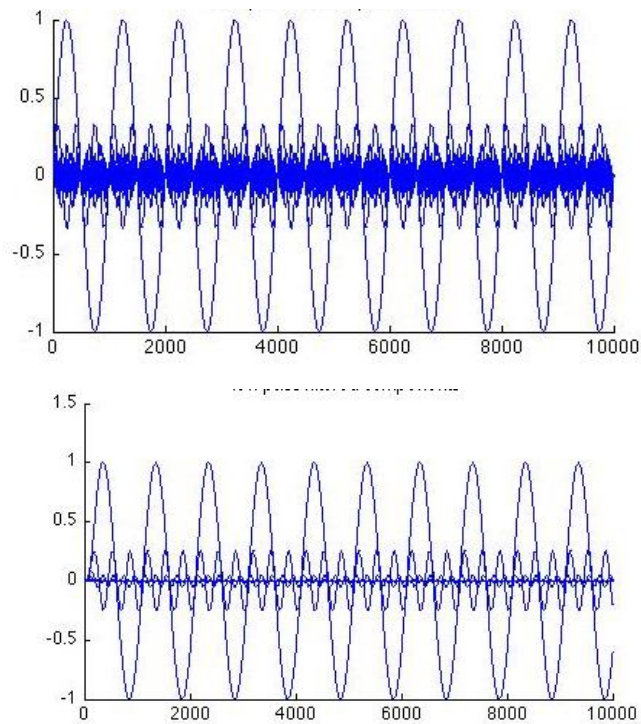
*Εικόνα 7 EEG σήματα μετά το πέρασμα τους από φίλτρο εγκοπής*

### 1.7.3.2 Η επίδραση των Low pass φίλτρων

Βοηθάνε στο να μειωθούν η μεγάλες συχνότητες οι οποίες συνήθως είναι θόρυβος ή σήματα τα οποία έχουν καταγραφή αλλά δεν παράχθηκαν από τον εγκέφαλο, αυτά τα σήματα ονομάζονται **artifact**. Μερικά artifacts μπορούν να μιμηθούν επιληπτικές κρίσεις ή πραγματικές επιληπτικές ανωμαλίες. Για τη αναγνώριση artifacts από εγκεφαλικά κύματα είναι σημαντικό να είναι γνωστό το λογικό τοπογραφικό πεδίο κατανομής των αληθινών EEG.

Ένα αρνητικό είναι το ότι κάνει τους θορύβους με υψηλή συχνότητα, όπως για παράδειγμα artifacts μυών, να μοιάζουν με ημιτονοειδή κύματα (δηλαδή σαν κύματα του εγκεφάλου)

Η σύνηθες συχνότητα η οποία χρησιμοποιείτε στα Low pass φίλτρα είναι τα 70Hz (Εικόνα 8).



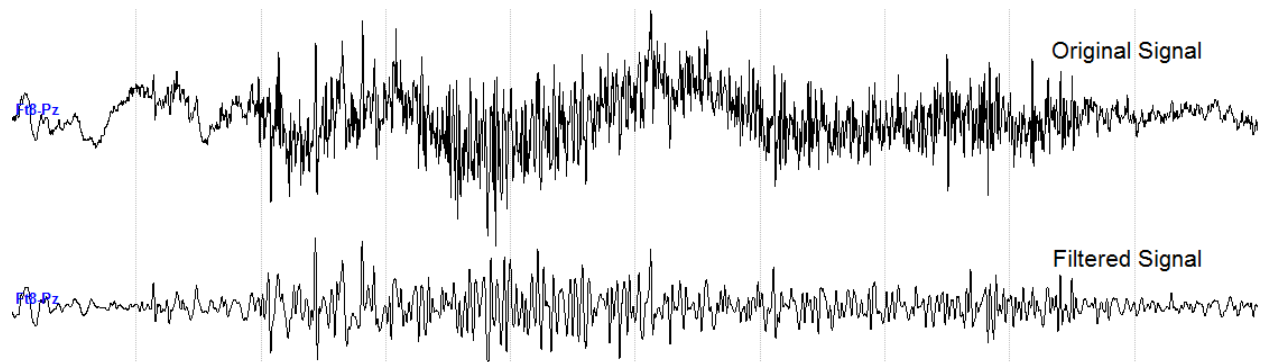
*Εικόνα 8 EEG χωρίς φιλτραρισμα (η πάνω φωτογραφία ) και με φιλτραρισμα Low pass (η κάτω φωτογραφία)*

### 1.7.3.3 Η επίδραση των High pass φίλτρων

Βοηθάνε στο να κρατηθούν τα κύματα κοντά στην βάση επίσης μειώνουν της λειτουργίες με πολύ μικρή συχνότητα που συνήθως είναι artifacts (Εικόνα 9).

Ένα αρνητικό είναι ότι μειώνει την ικανότητα να εκτιμηθούν αργά κύματα και ασυμμετρίες.

Η σύνηθες συχνότητα η οποία χρησιμοποιείτε στα High pass φίλτρα είναι τα 1Hz.



*Εικόνα 9 High pass φίλτρο*

## 2.1 Μέθοδοι εξαγωγής χαρακτηριστικών (feature extraction)

Υπάρχουν μας μέθοδοι για την εξαγωγή χαρακτηριστικών από EEG και η κάθε μία έχει τα πλεονεκτήματα και τα μειονεκτήματά της παρακάτω θα αναλυθούν.

### 1.8.1 Fast Fourier Transform (FFT)

Ο FFT είναι μία μέθοδος που χρησιμοποιεί μαθηματικά μέσα ή εργαλεία για να αναλύσει τα δεδομένα των EEG. Τα χαρακτηριστικά των ληφθέντων σημάτων EEG που πρόκειται να αναλυθούν υπολογίζονται από PSD (power spectral density) έτσι ώστε να αναπαραχθούν επιλεκτικά τα EEG σήματα. Ωστόσο μόνο τέσσερις ζώνες συχνοτήτων πμασέχουν της χαρακτηριστικές κυματομορφές των EEG.

Το PSD υπολογίζεται από μετασχηματισμας Fourier της εκτιμώμενης ακολουθίας αυτοσυσχέτισης που βρίσκεται με μη παραμετρικές μεθόδους μας από αυτές τις μεθόδους είναι η μέθοδος του Welch. Η ακολουθία των δεδομένων εφαρμόζεται σε data windowing και έτσι παράγει τροποποιημένα περιοδογραφήματα. Η ακολουθία πληροφοριών  $X_i(n)$  εκφράζεται ως εξής:

$$x_i(n) = x(n + iD), \quad n = 0, 1, 2, \dots, M-1$$
$$\text{while } i = 0, 1, 2, \dots, L-1;$$

Το  $iD$  είναι τμασημείο έναρξης της  $i$ th ακολουθίας. Το  $L$  μεγέθους  $2M$  αντιπροσωπεύει την ακολουθία των δεδομένων που σχηματιστηκαν. Η υμασογιζόμενη έξοδος μας δίνει :

$$\hat{P}_{xx}^{(i)}(f) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_i(n) w(n) e^{-j2\pi f n} \right|^2.$$

Ημαςνάρτηση παραθύρου  $U$  μας δίνει τον κανμασκοποιημένο συντελεστή της ισχύος ως εξής :

$$U = \frac{1}{M} \sum_{n=0}^{M-1} w^2(n),$$

Όπου το  $w(n)$  είναι η συνάρτηση παραθύρου. Ο μέσος όρος αυτών των τροποποιημένων περιοδογραφημάτων δίνει το φάσμα ισχύος του Welch :

$$P_{xx}^W = \frac{1}{L} \sum_{i=0}^{L-1} \hat{P}_{xx}^{(i)}(f).$$

### 1.8.2 Wavelet Transform

Ο μετασχηματισμός wavelet παίζει σημαντικό ρόλο στον τομέα αναγνώρισης και διάγνωσης: συμπιέζει το χρονικά μεταβαλλόμενο βιοϊατρικό σήμα, το οποίο περιλαμβάνει μασσημεία δεδομένων, σε μερικές μικρές παραμέτρους που αντιπροσωπεύουν το σήμα.

Καθώς το σήμα EEG είναι μη στάσιμο ο καταλληλότερος τρόπος για εξαγωγή χαρακτηριστικών από τα ακατέργαστα δεδομένα είναι η χρήση τμασεθόδων πεδίου χρόνου-συχνότητας όπως ο μετασχηματισμός wavelet που είναι μια τεχνική φασματικής εκτιμησης στην οποία μπορεί να γίνει

οποιαδήποτε γενική συνάρτηση εκφράζεται ως μια άπειρη σειρά κυμάτων. Δεδομένου ότι ο μετασχηματισμός wavelet επιτρέπει τη χρήση παραθύρων μεταβλητού μεγέθους, παρέχει έναν πιο ευέλικτο τρόπο αναπαράστασης χρόνου-συχνότητας ενός σήματος. Προκειμένου να επιτευχθεί καλύτερη ανάλυση χαμηλής συχνότητας, χρησιμοποιούνται wavelet παράθυρα μεγάλης διάρκειας. Αντιθέτως για τη λήψη πληροφοριών υψηλής συχνότητας, χρησιμοποιούνται παράθυρα σύντομου χρόνου. Επιπλέον, ο μετασχηματισμός wavelet περιλαμβάνει μόνο δομή πολλαπλών κλιμάκων και όχι μίας κλίμακας. Αυτή η μέθοδος είναι απλώς η συνέχεια της παραδοσιακής μεθόδου μετασχηματισμού Fourier. Επιπλέον, προορίζεται για την επίλυση προβλημάτων μη στάσιμων σημάτων όπως τα EEG.

Στη μέθοδο μετασχηματισμού wavelet, το αρχικό σήμα EEG αντιπροσωπεύεται από ασφαλή και απλά δομικά στοιχεία γνωστά ως wavelet. Το μητρικό wavelet δημιουργεί τα wavelets ως μέρος παράγωγων συναρτήσεων μέσω της μετάφρασης και της διαστολής, δηλαδή των λειτουργιών (μετατόπισης) και (συμπίεσης και τανύσεις) κατά μήκος του άξονα του χρόνου, αντιστοίχα. Υπάρχουν δύο κατηγορίες για μετασχηματισμού wavelet, το πρώτο είναι συνεχές ενώ το άλλο είναι διακριτό.

### 1.8.2.1 Continuous Wavelet Transform

Ο συνεχής μετασχηματισμός wavelet μπορεί να εκφραστεί ως εξής :

$$\text{CWT}(a, b) = \int_{-\infty}^{\infty} x(t) \psi_{a,b}^*(t) dt,$$

Όπου το  $x(t)$  είναι τα μη επεξεργασμένα EEG, το  $a$  είναι η καθυστέρηση, το  $b$  τον παράγοντα μετάφρασης. Το  $\Psi$  δηλώνει το μιγαδικό συζυγές και μπορεί να υπολογιστεί με:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right),$$

Όπου το  $\psi(t)$  είναι wavelet. Ωστόσο, η κύρια αδυναμία του είναι ότι η παράμετρος κλίμακωσης και η παράμετρος μετάφρασης του συνεχούς μετασχηματισμού wavelet αλλάζουν συνεχώς. Έτσι, οι συντελεστές του wavelet για όλες τις διαθέσιμες κλίμακες με μια υπολογισμό θα καταναλώσουν πολλή προσπάθεια και θα δώσουν πολλές αχρησιμοποίητες πληροφορίες.

### 1.8.2.2 Discrete Wavelet Transform

Προκειμένου να αντιμετωπιστεί η αδυναμία του συνεχούς μετασχηματισμού wavelet, ο διακριτός μετασχηματισμός wavelet έχει οριστεί στη βάση της αναπαράστασης χαρακτηριστικών σε πολλαπλές κλίμακες. Κάθε κλίμακα που εξετάζεται αντιπροσωπεύει ένα μοναδικό πάχος του σήματος EEG. Κάθε βήμα περιέχει δύο ψηφιακά φίλτρα,  $g(n)$  και  $h(n)$ , και δύο downsamplers κατά 2. Το διακριτό μητρικό wavelet  $g(n)$  κατά βάση είναι high pass ενώ το  $h(n)$  είναι low pass. Ο αριθμός των επιπέδων στα οποία αποσυντίθεται το wavelet επιλέγεται ανάλογα με τη συνιστώσα

των δεδομένων EEG με κυρίαρχη συχνότητα. Η σχέση μεταξύ μετασχηματισμός wavelet και φίλτρου  $h$  δηλαδή low pass, μπορεί να αναπαρασταθεί ως εξής:

$$H(z)H(z^{-1}) + H(-z)H(-z^{-1}) = 1.$$

Εδώ, το  $H(z)$  αντιπροσωπεύει τον μετασχηματισμό  $z$  του φίλτρου  $h$ . Ο συμπληρωματικός μετασχηματισμός  $z$  του high pass φίλτρου εκφράζεται ως :

$$G(z) = zH(-z^{-1}).$$

Περιγράφοντας επακριβώς τα χαρακτηριστικά του τμήματος σήματος σε έναν καθορισμένο τομέα συχνότητας και τις ιδιότητες εντοπισμένου τομέα χρόνου, υπάρχουν πιασά πλεονεκτήματα που επισκιάζουν τις υψηλές απαιτήσεις υπολογισμού και μνήμης της συμβατικής υλοποίησης του διακριτού μετασχηματισμού wavelet που βασίζεται σε συνέλιξη.

### 1.8.3 Eigenvectors

Οι μέθοδοι ιδιοδιανυσμάτων ή Eigenvectors χρησιμοποιούνται για την εκτίμηση των συχνοτήτων και των δυνάμεων των σημάτων από μετρήσεις που έχουν καταστραφεί από θόρυβο. Μερικές από αυτές είναι η μέθοδος του Pisarenko, τη μέθοδο MUSIC (multiple signal classification) και η μέθοδος Minimum-Norm.

#### 1.8.3.1 Η μέθοδος Pisarenko

Η μέθοδος Pisarenko είναι μία μέθοδος που χρησιμοποιείται για την αξιολόγηση του PSD. Για να υπολογιστεί το PSD η μαθηματική έκφραση  $A(f)$  που θα χρησιμοποιηθεί δίνεται ως :

$$A(f) = \sum_{k=0}^m a_k^{-j2\pi f k}$$

Οπού το  $a_k$  είναι ο συντελεστής της καθορισμένης εξίσωσης και το  $m$  είναι η σειρά των ιδιοφίλτρων (eigenfilters). Για να υπολογιστεί το PSD :

$$P_{Pisarenko} = \frac{1}{|A(f)|^2}$$

#### 1.8.3.2 Η μέθοδος MUSIC

Η μέθοδος MUSIC χρησιμοποιείται για να δώξει ό,τι προβλήματα μπορεί να δημιουργήσουν το λάθος μηδενικά και υπολογίζει το PSD με την παρακάτω εξίσωση :

$$P_{MUSIC}(f) = \frac{1}{\frac{1}{K} \sum_{i=0}^{K-1} |A(f)|^2}$$



## 1.8.4 Time-Frequency Distributions

Στης TFD μεθόδους τα σήματα χρειάζεται να είναι χωρίς θόρυβο για να έχουν καλά αποτελέσματα, οπότε χρειάζεται προ επεξεργασία για να μην υπάρχουν artifacts. Επειδή είναι μέθοδοι **χρονικής** συχνότητας, ασχολούνται με τη στατική αρχή και για αυτών τον λόγο απαιτητέ διαδικασία **widowing** όταν γίνετε η προ επεξεργασία. Ο ορισμός του TFD για ένα σήμα γενικεύτηκε από τον Cohen ως :

$$P(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A(\theta, \tau) \Phi(\theta, \tau) e^{-j\theta t - j\omega \tau} d\theta d\tau,$$

Όπου το  $A(\theta, \tau)$ :

$$A(\theta, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x\left(u + \frac{\tau}{2}\right) x^*\left(u - \frac{\tau}{2}\right) e^{j\theta u} du.$$

Το  $A(\theta, \tau)$  είναι πιο γνωστό ως συνάμαση ασάφειας (ambiguity function) και το  $\omega, \tau$  είναι αναφέρεται στον πυρήνα της διανομής.

Το SPWV είναι μια παραλλαγή της μεθόδου που ενσωματώνει εξομάλυνση από ανεξάρτητα παράθυρα σε χρόνο και συχνότητα, συγκεκριμένα το  $w_{\omega}(\tau)$  και  $w_t(t)$ :

$$\begin{aligned} \text{SPWV}(t, \omega) &= \int_{-\infty}^{\infty} W_{\omega}(\tau) \left[ \int_{-\infty}^{\infty} W_t(u - t) x\left(u + \frac{\tau}{2}\right) \right. \\ &\quad \left. \times x^*\left(u - \frac{\tau}{2}\right) du \right] e^{-j\omega \tau} d\tau. \end{aligned}$$

Η εξαγωγή χαρακτηριστικών με TFD βασίζεται στην ενέργεια, τη συχνότητα και το μήκος του κύριου track. Πρώτα χωρίζεται το EEG σήμα σε  $k$  τμήματα. Κάθε τμήμα επιστρέφει την ενέργεια  $E_k$ , τη συχνότητα  $F_k$  και το μέγεθος  $L_k$ .

Η ενέργεια για κάθε  $k$  είναι:

$$E_k = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \theta_k(t, f) dt df,$$

Όπου το  $\theta_k(t, f)$  αντιπροσωπεύει την αναπαράσταση χρόνου-συχνότητας του τμήματος. Και η συχνότητα  $F_k$  είναι:

$$F_k = \int_{-\infty}^{\infty} \theta_k(t, f) dt.$$

### 1.8.5 Autoregressive Method

Οι Autoregressive (AR) μέθοδοι εκτιμούν την PSD του EEG χρησιμοποιώντας μια παραμετρική προσέγγιση. Επομένως, οι μέθοδοι AR δεν έχουν πρόβλημα φασματικής διαρροής και έτσι αποδίδουν καλύτερη ανάλυση συχνότητας σε αντιθεση με τη μη παραμετρική προσέγγιση.

Η εκτίμηση του PSD επιτυγχάνεται με τον υπολογισμό των συντελεστών, δηλαδή των παραμέτρων του υπό εξέταση γραμμικού μαζατος.

Δύο τέτοιες μέθοδοι είναι η Yule-Walker και η Burg.

### 1.8.6 LSTM

Ένας άλλος τρόπος για την εξαγωγής χαρακτηριστικών από EEG είναι με το να παρθούν οι έξοδοι των σημάτων από ένα στρώμα εκπαιδευόμενου LSTM (δηλ. της εξόδους από ένα hidden layer ).

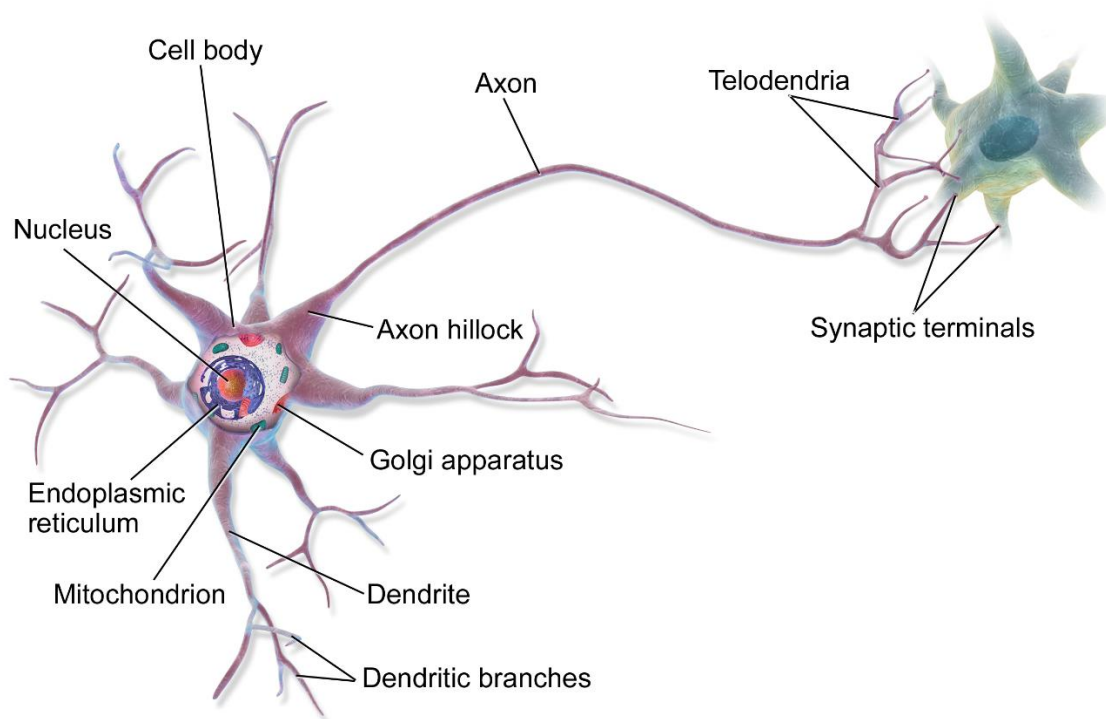


## Κεφάλαιο 2 : Generative Adversarial Networks

### 2.1 Τεχνητά Νευρωνικά δίκτυα

#### 2.1.1 Τι είναι τα Τεχνητά Νευρωνικά δίκτυα

Τα τεχνητά νευρωνικά δίκτυα είναι μασινάκια που προσπαθούν να ‘αντιγράψουν’ την λειτουργία του εγκεφάλου. Ο εγκέφαλος μας αποτελείται από δισεκατομμύρια νευρώνες (Εικόνα 10) οι οποίοι επικοινωνούν μεταξύ τους μέσω δενδριτών. Οι νευρώνες στέλνουν ένα ηλεκτρικό σήμα σε άλλους νευρώνες και έτσι γίνεται η σκέψη ή η εκτέλεση διάφορων δραστηριοτήτων για παράδειγμα περπάτημα ή η όραση και όλες της άλλες δραστηριότητες που μπορεί να κάνει το ανθρώπινο σώμα.

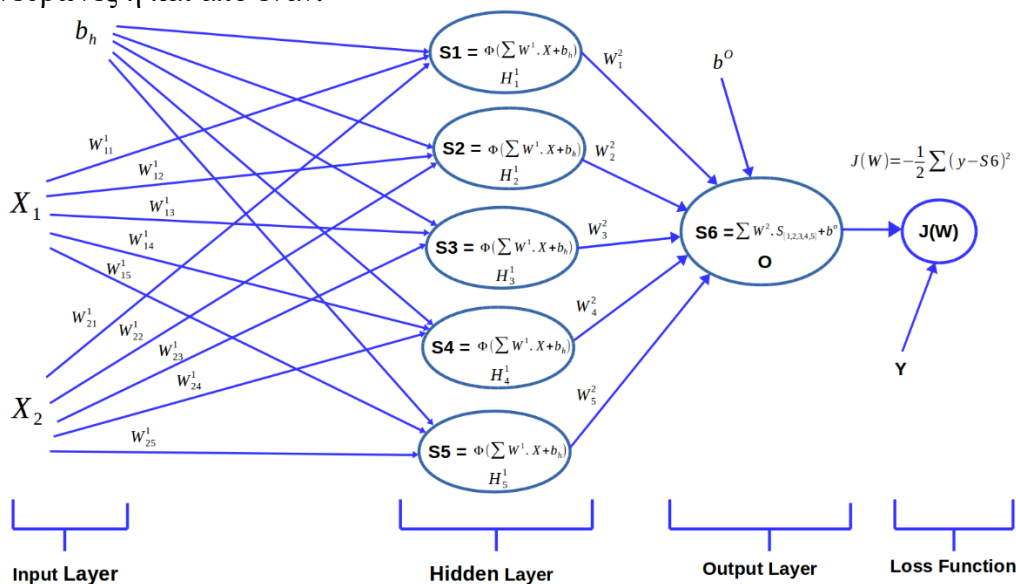


*Εικόνα 10 Ένας Νευρώνας*

Οπότε ένα νευρωνικό δίκτυο είναι ένας πολύ μεγάλος παράλληλος επεξεργαστής ο οποίος αποτελείται από πολλούς νευρώνες που είναι ουσιαστικά απλές μονάδες επεξεργασίας που έχει την δυνατότητα να αποθηκεύει γνώση και να την χρησιμοποιεί και άλλες φορές. Αυτό είναι ένα ΤΝΔ, και ουσιαστικά μιμείται την λειτουργία του εγκεφάλου. Και τα δύο δέχονται δεδομένα από το περιβάλλον τους τα οποία τα επεξεργάζονται και παράγουν γνώση. Η γνώση στα ΤΝΔ είναι τα βάρη του κάθε κόμβου.

### 2.1.2 Πώς λειτουργεί ένα ΤΝΔ

Ο νευρώνας παίρνει της εισόδους του και της πολλαπλασιάζει με ένα βάρος,σε μερικές περιπτώσεις προσθέτει και ένα bias,και μετά όλες οι είσοδοι περνάνε από ένα αθροιστή ο οποίος της αθροίζει και το αποτέλεσμα τους περνάει από μία συνάρτηση ενεργοποίησης (activation function) για παράδειγμα μια rectified linear unit ή αλλιώς πιο γνωστή ως ReLU.Αυτή είναι η βασική δομή ενός ‘νευρώνα’ σε ένα ΤΝΔ (Εικόνα 11).Το ΤΝΔ μπορεί να αποτελείτε από πάρα πολλούς νευρώνες ή και από έναν.



Εικόνα 11 : Ένα παράδειγμα ΤΝΔ(MLP)

Η πιο απλή μορφή ενός ΤΝΔ είναι ένα νευρωνικό perceptron το οποίο αποτελείται από ένα νευρώνα και από μία έξοδο το οποίο δεν είναι δυνατό να λύση πάρα πολλά προβλήματα.Οπότε χρησιμοποιούνται τα MLP τα οποία αποτελούνται το λιγότερο από τρία επίπεδα το επίπεδο της εισαγωγής (Input layer),το κρυφό επίπεδο(hidden layer ) και το επίπεδο εξόδου.Μπορεί να έχει όσους νευρώνες χρειάζεται και να έχει όσα κρυφά επίπεδα χρειάζεται επίσης.Κάτι που δεν ειπώθηκε προηγουμένως είναι ότι το perceptron και το MLP είναι νευρωνικά επιβλεπόμενης μάθησης (supervised learning) αυτό σημαίνει ότι,για το νευρωνικό,στη διαδικασία εκπαίδευσης πρέπει να υπάρχουν δεδομένα εισόδου και το τι έξοδο πρέπει να βγάλουν.

Η δημιουργία ενός νευρωνικού δικτύου μπορεί να χωριστεί σε τρεις φάσης : 1) την φάση του Data,2) την φάση του training και φάση της αξιολόγησης.

Η φάση τον “data” είναι ίσως η πιο σημαντική,σε αυτήν την φάση γίνετε ή φτιάχνετε το dataset, μετά γίνεται η προ επεξεργασία (pre processing ).

Το pre processing μπορεί να γίνει για πολλούς λόγους όπου μερική από αυτούς θα αναπτυχθούν στην συνέχεια :

#### 1) Χειρισμός των μηδενικών(NULL) τιμών ή NaN τιμών :

Πολλές φορές σε ένα dataset μπορεί να υπάρχουν κενές τιμές αυτό θα προκαλέσει σοβαρά προβλήματα στο νευρωνικό μας δίκτυο,διότι κανένα δίκτυο δεν μπορεί να τα διαχειριστή,οπότε πρέπει να γίνει κάτι,είτε να διαγράψουν γραμμές από το dataset οι

οποίες περιέχουν τα Null ή τα NaN, είτε να τα αντικατασταθούν. Αμα το ποσοστό των Null/NaN είναι μικρό τότε συνηθίζετε το πρώτο δηλαδή η διαγραφή, άμα όμως το ποσοστό είναι μεγάλο πχ 30% τότε θα χαθούν πολλά δεδομένα και για αυτό προτιμάτε η αντικατάστασή τους.

## 2) Κανονικοποίηση :

Η κανονικοποίηση είναι μία διαδικασία στην οποία οι τιμές μετατρέπονται στο να είναι ανάμεσα στο 0 και 1. Αυτό γίνεται για να είναι όλες η τιμές στο ίδιο εύρος, για παράδειγμα άμα παρθεί ένα dataset με πλοία στο οποίο αναφέρονται τα κιλά του και την χρονολογία κατασκευής του πχ. `uss new jersey | 120000 | 1942`

Όπως παρατηρείτε τα κιλά με την χρονολογία έχουν διαφορετικές κλίμακες έτσι το νευρωνικό θα νομίζει ότι τα κιλά είναι σημαντικότερα από την χρονολογία το οποίο στο συγκεκριμένο παράδειγμα θα ήταν λάθος.

Οπότε γίνονται οι παρακάτω πράξεις :

$$y = \frac{x_i - \mu}{\sigma} \quad \mu = \frac{\sum(x)}{\text{len}(x)} \quad \sigma = \sqrt{\frac{\sum_{i=0}^{\text{len}(x)} (x_i - \mu)^2}{\text{len}(x)}}$$

## 3) Κατηγορικές μεταβλητές :

Οι κατηγορικές μεταβλητές είναι οι μεταβλητές που είναι διακριτές και όχι συνεχείς. Για παράδειγμα είδος μουσικής (κλασική, ροκ, κτλ.) είναι διακριτή μεταβλητή ενώ η τιμή ενός προϊόντος είναι συνεχείς μεταβλητή.

Τα νευρωνικά δίκτυα διαβάζουν αριθμούς και όχι αλφαριθμητικά έτσι θα πρέπει να γίνει κάτι για να μετατραπουν.

Υπάρχουν 2 είδη διακριτών μεταβλητών : η ταξικές, η οποίες μπορούν να ταξινομηθούν (πχ μέγεθος ρούχων L, XL, XXL), και οι ονομαστικές η οποίες είναι αλφαριθμητικές (πχ είδος μουσικής ροκ, κλασική κτλ.).

Ένας τρόπος για να λυθεί αυτό το πρόβλημα είναι να γίνει χαρτογράφηση (map) των μεταβλητών για παράδειγμα κλασική : 1, ροκ : 2 ...

Αυτό όμως υποδηλώνει ότι η κλασική είναι μικρότερη από την ροκ κάτι το οποίο δεν βγάζει νόημα, το νευρωνικό θα κάνει τη συσχέτιση αλλά αυτή η μέθοδος δεν προτιμάται.

Η πιο καλή μέθοδος είναι η one-hot encoding. Ο one-hot encoder είναι ένας αλγόριθμος ο οποίος φτιάχνει ένα πίνακα με n στήλες όπου το n είναι οι ξεχωριστές κατηγορίες, και βάζει 1 στην κατηγορία όπου είναι, και 0 στις άλλες, για παράδειγμα στο παρακάτω dataset :

Id	Χρώμα	Τιμή
1	Μπλε	10.67
2	Μαύρο	11.8
3	Κίτρινο	12.9

4	Μπλε	11.3
5	Μαύρο	12.1

Μετά τον one-hot encoder θα γίνει :

ID	Μπλε	Μαύρο	Κίτρινο	Τιμή
1	1	0	0	10.67
2	0	1	0	11.8
3	0	0	1	12.9
4	1	0	0	11.3
5	0	1	0	12.1

#### 4) Αλφαριθμητικά :

Σε μερικά προβλήματα τα δεδομένα είναι αλφαριθμητικά (πχ για ένα νευρωνικό δίκτυο που βρίσκει τα spam emails) έτσι θα πρέπει να υλοποιηθούν μερικά βήματα τα οποία είναι τα εξής :  
1) tokenization, 2) vectorization.

Το tokenization ουσιαστικά σπάει το κείμενο σε κομμάτια, δηλαδή σε λέξεις, για παράδειγμα στην πρόταση «Το καναρίνι κελαηδάει, η δεκοχτούρα τι κάνει ;». Μετά τον tokenization, αυτή η πρόταση θα γίνει : ['Το', 'καναρίνι', 'κελαηδάει', ',', 'η', 'δεκοχτούρα', 'τι', 'κάνει', ';'].

Το vectorization είναι μια κλασική προσέγγιση μετατροπής δεδομένων εισόδου από κείμενο σε διανύσματα (vectors) πραγματικών αριθμών να μπορούν να υποστηριχθούν από τα νευρωνικά δίκτυα. Μερικοί αλγόριθμοι vectorization είναι ο bag-of-words, ο word2vec της google (ο καλύτερος) και ο TF-IDF.

### Training:

Η εκπαίδευση ενός νευρωνικού δικτιού είναι πάρα πολύ σημαντική φάση της ανάπτυξης του. Άμα είναι επιβλεπόμενης μάθησης τότε τα βήματα είναι τα εξής :

1. Χωρίζονται τα δεδομένα σε δεδομένα για εκπαίδευση (train) και σε δεδομένα για τεστ (test). Συνήθως τα δεδομένα χωρίζονται 80% (train) και 20% (test).
2. Ορίζονται ο αριθμός των εποχών που θα γίνει η εκπαίδευση. (Σε κάθε εποχή παίρνονται όλα τα δεδομένα (train) από το νευρωνικό μετά τρέχει ο αλγόριθμος BP. Και μετά παίρνονται τα δεδομένα τεστ έτσι ώστε να παρθεί η ακρίβεια του νευρωνικού.)
3. Αφού έχει οριστεί ο αριθμός των εποχών, ξεκινάει η εκπαίδευση για N εποχές.
4. Περνιούνται τα δεδομένα (train) στο νευρωνικό.
5. Λαμβάνονται η έξοδοι του νευρωνικού.
6. Τρέχει ο αλγόριθμος BP. Βρίσκεται το error, το οποίο ισούται με  $error = actual - prediction$  δηλαδή αυτό που έπρεπε να βγάλει, μείον αυτό που έβγαλε. Ο αλγόριθμος BP ανανεώνει τα βάρη.
7. Περνιούνται τα δεδομένα τεστ και παρατηρείτε την ακρίβεια τους.

8. Και ξανά από την αρχή μέχρι να τελειώσουν η εποχές.

### 2.1.3 Τι είναι ο BP

Ο BP είναι μία μέθοδος για να υπολογίζονται παράγωγοι σε ένα feedforward νευρωνικό δίκτυο. Ο BP είναι ένα από τα πιο σημαντικά κομμάτια στην εκπαίδευση ενός αλγόριθμου επιβλεπόμενης μάθησης. Ένα από τα μοντέλα που τον χρησιμοποιούν είναι ο STG, ο οποίος είναι μία μέθοδος για να βρίσκονται οι καταλληλότερες παράμετροι για αλγορίθμους μηχανικής μάθησης. Αυτό που κάνει ο STG είναι επαναληπτικά κάνει μικρές προσαρμογές στο νευρωνικό δίκτυο έτσι ώστε να κατεβάσει το error (loss). Όταν εκπαιδεύεται ένα νευρωνικό με SGD, υπολογίζεται μία συνάρτηση απώλειας (loss), που δείχνει πόσο μακριά από τα πραγματικά labels είναι τα προβλεπόμενα του νευρωνικού. Ο BP μας επιτρέπει να υπολογιστεί η κλίση της συνάρτησης απώλειας για κάθε βάρος του δικτιού. Αυτό δίνει την δυνατότητα να ανανεωθούν τα βάρη ατομικά και έτσι να κατεβαίνει η απώλεια μετά από πολλές επανάληψης. Ο BP δουλεύει ανάποδα δηλαδή από το τελευταίο layer προς το πρώτο για να βρει την κλίση (παράγωγο).

Για να βρει την κλίση σε ένα συγκεκριμένο layer ενώνει όλες της προηγούμενες κλίσεις με τον κανόνα της αλυσίδας του απειροστικού λογισμού (Εικόνα 12). Λόγο της υπολογιστικής δύναμης τα τελευταία χρόνια η βαθιά μάθηση έγινε πολύ δημοφιλής και ο BP είναι πάρα πολύ σημαντικός στην βαθιά μάθηση.

$$\frac{\partial \mathcal{E}(n)}{\partial w_{ji}(n)} = \frac{\partial \mathcal{E}(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)}$$

*Εικόνα 12: Αυτός είναι ο κανόνας αλυσίδας του διαφορικού λογισμού που χρησιμοποιεί ο BP*

### 2.1.4 Είδη ΤΝΔ και βαθιάς μάθησης

Βαθιά μάθηση είναι νευρωνικά δίκτυα τα οποία χρησιμοποιούν πολλά layers.

Υπάρχουν 3 είδη νευρωνικών δικτύων τα με επίβλεψη, με μη επίβλεψη και generative (τα GAN και τα autoencoders)

Υπάρχουν πολύ τύποι βαθιάς μάθησης κάποιοι από τους πιο διάσημους είναι οι εξής :

#### ➤ CNN

Τα CNN είναι νευρωνικά δίκτυα τα οποία χρησιμοποιούνται για ανίχνευση αντικειμένων, επεξεργασία εικόνας, ταξινόμηση εικόνων και βίντεο (Εικόνα 13).

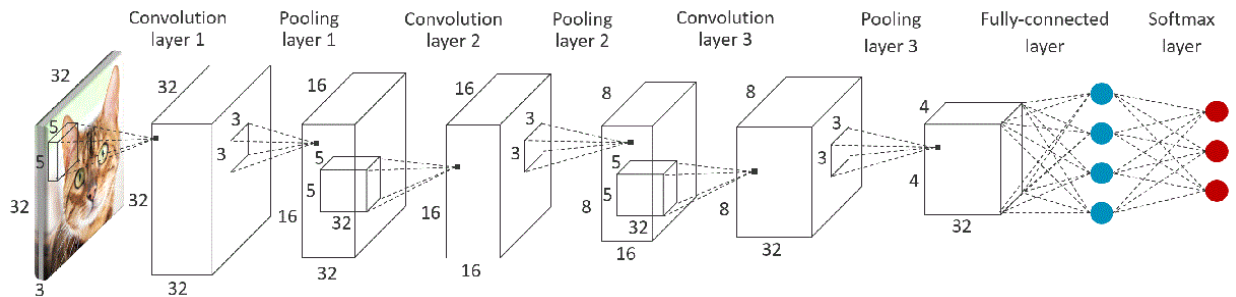
Τα CNN χωρίζονται σε δύο κομμάτια το πρώτο κομμάτι κάνει feature extraction (εξαγωγή χαρακτηριστικών) και το δεύτερο το classification(ταξινόμηση).Το πρώτο κομμάτι χρησιμοποιεί πολλά layers από τα παρακάτω με την ίδια σειρά :

**Convolutional layer** το οποίο εκτελεί διάφορα φίλτρα για να κάνει το convolution.

**ReLU** το οποίο είναι συνάρτηση ενεργοποίησης και είναι ο τύπος  $\max(0,z)$  οπου το z είναι η είσοδος της που στα CNN είναι το convolutional layer.

**Pooling Layer** μετά το ReLU τρέφεται στο pooling layer το feature map το οποίο κάνει δειγματοληψία και μειώνει τις διαστάσεις του feature map μέχρι να φτάσει στα δύο όπου τα κάνει flatten και τα περνάει στο δεύτερο κομμάτι.

Το δεύτερο κομμάτι είναι ένα πλήρες ενωμένο layer το οποίο κάνει την ταξινόμηση.



Εικόνα 13: Μοντέλο CNN

Τα δίκτυα αυτά είναι με επίβλεψη.

## ➤ LSTM

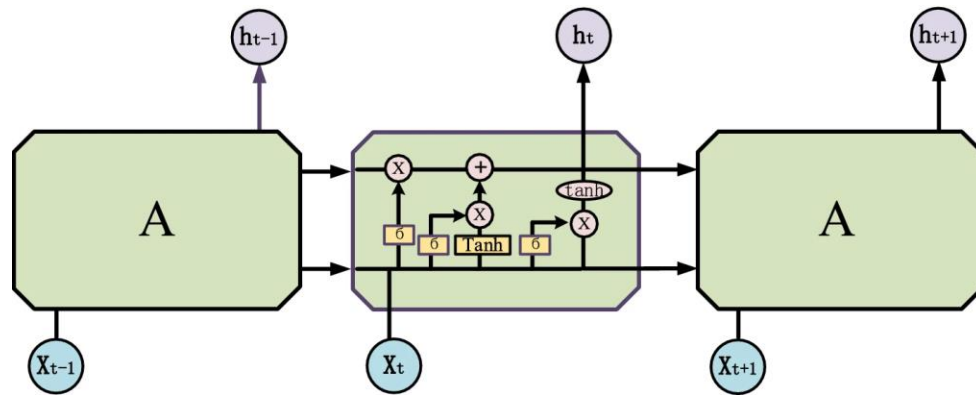
Τα LSTM είναι ένας RNN τύπος δικτύου που μπορεί να απομνημονεύσει μακροπρόθεσμες εξαρτήσεις Σε αντιθεση με τα τυπικά νευρωνικά δίκτυα που είναι feedforward (εμπροσθοδοτούμενα ) τα LSTM διαθέτουν συνδέσμους ανατροφοδότησης που τα επιτρέπουν να επεξεργάζονται δεδομένα ως σειρές.Τα LSTM κρατάνε της πληροφορίες.Είναι πολύ χρήσιμα όταν χρησιμοποιούνται με κυματομορφές και σήματα επειδή θυμούνται προηγούμενες εισόδους.Επίσης χρησιμοποιούνται και σε αναγνώριση φωνής,παραγωγή φαρμάκων και δημιουργία μουσικής.

Τα LSTM λειτουργούν ως εξής :

Πρώτα ξεχνάνε ασήμαντες πληροφορίες από προηγούμενες καταστάσεις

Μετά επιλεκτικά αναβαθμίζουν της τιμές κυτταρικής(cell , Εικόνα 14) κατάστασης

Και τέλος βγάζουν σαν έξοδο ορισμένα κομμάτια της κατάστασης.

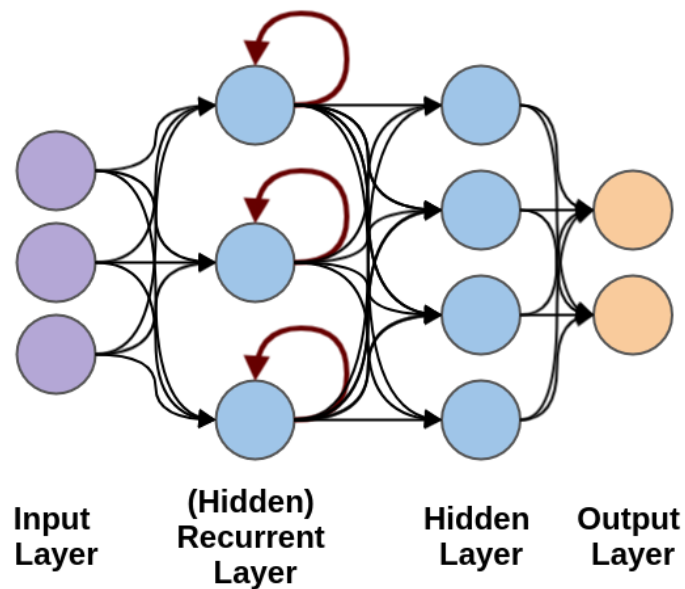


Εικόνα 14: LSTM cells

## ➤ RNN

Τα RNN έχουν συνδέσεις οι οποίες κάνουν κύκλο το οποίο επιτρέπει στο LSTM να πάρει ως είσοδο του την έξοδο της προηγούμενης κατάστασης (Εικόνα 15).

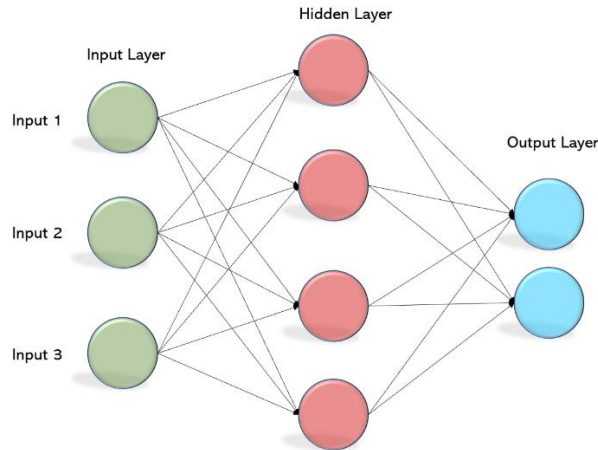
Η έξοδος από το LSTM γίνεται είσοδος στην επόμενη φάση και μπορεί να απομνημονεύσει λόγω της εσωτερικής τους μνήμης συνήθως χρησιμοποιούνται για NLP, μετάφραση, αναγνώριση γραφικών χαρακτήρων και ανάλυση χρόνου σειρών.



Εικόνα 15: RNN

### ➤ MLP

Όπως ειπώθηκε και πιο πάνω τα MLP είναι perceptron με πολλά layers είναι feedforward (Εικόνα 16). Ένα MLP απαρτίζεται από τουλάχιστον 3 στρώματα : το στρώμα εισόδου, το κρυφό στρώμα και το στρώμα εξόδου, το κρυφό στρώμα μπορεί να είναι παραπάνω από 1. Είναι με επίβλεψη και χρησιμοποιούν τον αλγόριθμο BP. Χρησιμοποιούνται στο fitness approximation το οποίο προσπαθεί να μιμηθεί μαθηματικές πράξεις, στην ταξινόμηση, στην αναγνώριση εικόνων και φωνής και σε μεταφραστές (πχ google translate).

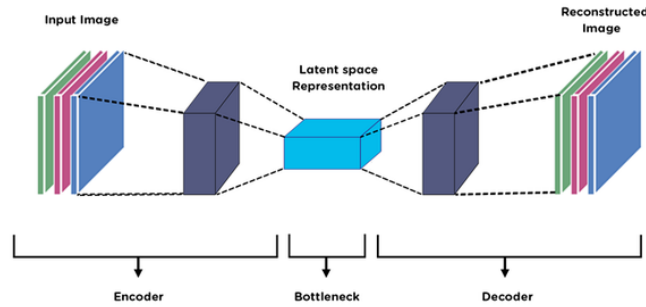


*Εικόνα 16 Ένα απλό MLP*

### ➤ Autoencoders

Οι autoencoders (Εικόνα 17) είναι νευρωνικά τα οποία η είσοδος τους είναι ίδια με την έξοδο τους χρησιμοποιούνται για κρυπτογραφία [13], για παραγωγή εικόνας, για συμπίεση και άλλες πολλές χρήσεις. Είναι δίκτυα χωρίς επίβλεψη. Υπάρχουν πολλά είδη autoencoder μερικά είναι τα εξής : autoencoder αφαίρεσης θορύβου (DAE), συμβατικός (contractive) autoencoder (CAE), τα concrete autoencoders και τα variational autoencoder τα οποία χρησιμοποιούνται και σαν GAN. Επίσης μπορούν να χρησιμοποιηθούν και σε μείωση διαστάσεων (dimensionality reduction, πχ από 3d σε 2d), ανάκτηση πληροφοριών, ανίχνευση ανωμαλιών (anomaly detection), επεξεργασία εικόνας, ανακάλυψη φαρμάκων [14] και πρόβλεψη δημοτικότητας .



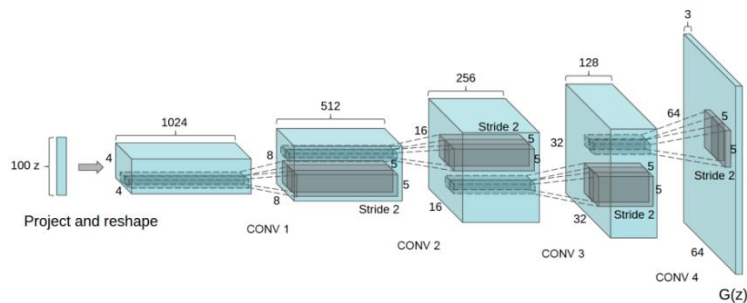


Εικόνα 17: Autoencoder

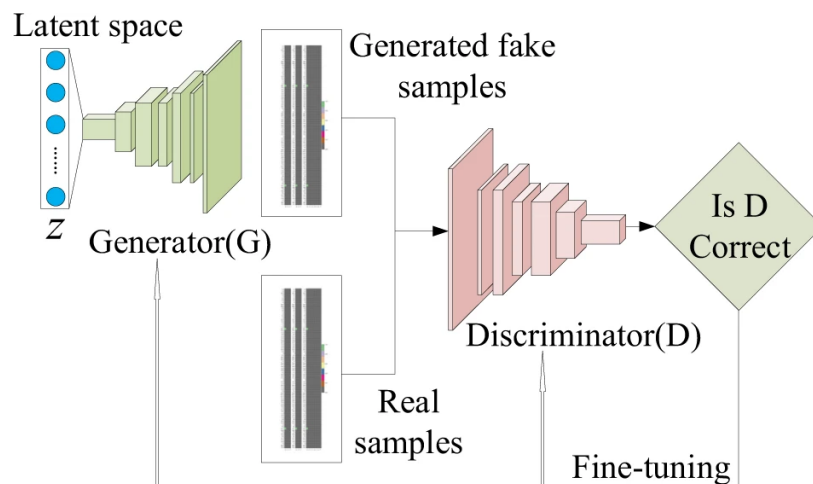
## 2.2 GAN

### 2.2.1 Εισαγωγή

Τα GAN (Εικόνα 19) είναι σχετικά πολύ νέα τεχνολογία/τύπος νευρωνικού δημιουργήθηκαν το 2014 από τον Ian Goodfellow [15]. Τα GAN είναι δύο νευρωνικά τα οποία το ένα προσπαθεί να 'ξεγελάσει' το άλλο, και το άλλο να 'πιάσει' το άλλο σαν ένα παιχνίδι κλέφτες και αστυνόμοι. Το ένα δίκτυο είναι ο generator (Εικόνα 18) ο οποίος παράγει φωτογραφίες, σαν είσοδο παίρνει έναν θόρυβο  $z$  και δημιουργεί μια φωτογραφία. Το άλλο δίκτυο ο Discriminator ελέγχει αν η φωτογραφία είναι αληθινή. Έτσι καταλήγει ο Generator να παράγει φωτογραφίες που ο Discriminator δεν μπορεί να της καταλάβει ότι είναι 'ψεύτικες'. Τα GAN γενικά είναι πολύ δύσκολα στην εκπαίδευση γιατί έχουν δύο δίκτυα τα οποία εκπαιδεύονται 'ταυτόχρονα' και το ένα βασίζεται στο άλλο. Ο Generator λειτουργεί σαν ανάποδο CNN ενώ ο Discriminator είναι ένα CNN.



Εικόνα 18: Generator



Εικόνα 19: GAN

### 2.2.2 Ιστορία

Η ιδέα της τοποθέτησης δύο αλγορίθμων, ενός εναντίον του άλλου προήρθε από έναν ερευνητή της IBM τον Artur Samuel στον οποίο αποδίδεται η διάδοση του όρου ‘Machine Learning’ (μηχανική μάθηση). Ο Arthur Samuel έφτιαξε ένα παιχνίδι ντάμας που το ονόμασε ‘The Samuel Checkers’ το οποίο ήταν ένα από τα πρώτα επιτυγχές self-learn πρόγραμμα, υπολογίζοντας την πιθανότητα νίκης κάθε πλευράς σε μια δεδομένη θέση.

Αν ο Arthur Samuel είναι ο “παππούς” των GAN τότε ο Ian Goodfellow είναι ο “πατέρας” των GAN. Ο Ian Goodfellow είναι ερευνητής και διευθυντής μηχανικής μάθησης στο τμήμα Special Project Group στην Apple ενώ προηγουμένως ήταν ερευνητής στο Google Brain της Google. Το 2014 δημοσίευσε μία από τις σημαντικότερες ερευνητικές εργασίες για μηχανική μάθηση με τίτλο “Generative Adversarial Nets” [15] και είναι η πρώτη λειτουργική εφαρμογή ενός generative μοντέλου που βασίζεται σε δίκτυα αντιπάλων. Ο Goodfellow δεν ήταν ο μόνος που επιδίωξε τον σχεδιασμό ενός μοντέλου αντιπάλων για τεχνητή νοημοσύνη, ο Juergen Schmidhuber συνδιευθυντής του ινστιτούτου Dalle Molle για τεχνητή νοημοσύνη, υποστήριξε την ελαχιστοποίηση προβλεψιμότητας, μια τεχνική που μοντελοποιεί τις διανομές μέσω ενός encoder που μεγιστοποιεί την αντικειμενική συνάρτηση (τη συνάρτηση που καθορίζει το πρόβλημα που πρέπει να λυθεί από το σύστημα) που ελαχιστοποιείται από έναν predictor. Χρησιμοποιεί τον κανόνα απόφασης minimax, όπου η πιθανή απώλεια για το χειρότερο σενάριο ελαχιστοποιείται όσο το δυνατόν περισσότερο.

### 2.2.3 Αρχιτεκτονική

Όπως ειπώθηκε πιο πάνω τα GAN είναι 2 δίκτυα ένας Generator και ένας Discriminator. Ο Discriminator είναι ένας δυαδικός ταξινομητής ο οποίος μας επιστρέφει μια τιμή ανάμεσα στο 0 με 1, το 0 είναι το ότι η είσοδος είναι ψεύτικη ενώ το 1 ότι η είσοδος αληθινή. Ψεύτικη είσοδος είναι τα δεδομένα που παράχθηκαν από τον generator ενώ αληθινά είναι του dataset. Ο generator είναι ένα ανάποδο CNN η είσοδος του είναι random θόρυβος (έτσι ώστε να παράγει τυχαία αποτελέσματα) και η έξοδος του είναι παραγόμενη πληροφορία για παράδειγμα ο πιο

συνηθισμένος τύπος δεδομένων που μπορούν να παραχθούν είναι εικόνες αλλά μπορούν να παραχθούν όλα τα είδη δεδομένων.

## 2.2.4 Εκπαίδευση GAN

Η εκπαίδευση γίνεται ως εξής :

1. Χωρίζονται τα δεδομένα σε μικρές ομάδες (Batches)
2. Παράγεται τυχαίος θόρυβος συνήθως μεγέθους 100 για κάθε ένα στοιχείο στο batch για παράδειγμα άμα το μέγεθος του batch είναι 16 τότε θα παραχθεί ένας 16X100 2d πίνακας.
3. Μετά περνιέται ο θόρυβος στον Generator και παράγονται οι ψεύτικες φωτογραφίες.
4. Δημιουργείται ένας πίνακας με 1 ο οποίος θα είναι τα label των αληθινών δεδομένων.(Ο discriminator ουσιαστικά επιστρέφει το πόσο από μηδέν έως ένα το πόσο αληθινή είναι η εικόνες οπότε το ιδανικό είναι να φτάσει στο να δείχνει 0,5)
5. Δημιουργείται ένας πίνακας με 0 ο οποίος θα είναι τα label των ψεύτικων δεδομένων.
6. Περνιέται το αληθινό batch στον Discriminator και με labels τα 1 και λαμβάνεται το error.
7. Μετά περνιέται τα ψεύτικα δεδομένα στον Discriminator με labels τα 0 και λαμβάνεται το error.
8. Γίνεται πρόσθεση των 2 errors και περνιούνται από BP για να γίνει η εκπαίδευση.
9. Μετά περνιούνται οι ψεύτικες φωτογραφίες στον Discriminator και με labels τα 1 και μετά τρέχει ο BP για τον Generator για να εκπαιδευτεί.
10. Τέλος επαναλαμβάνεται μέχρι να τρέξουν όλα τα Batches για όσες εποχές έχουν οριστεί.

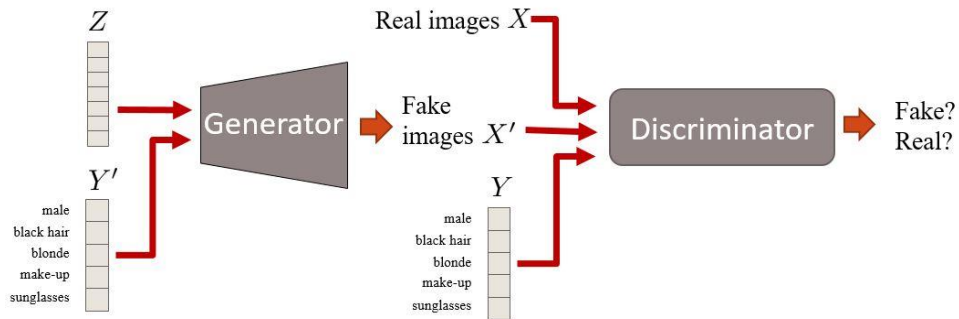
## 2.2.5 Παραλλαγές GAN

Μερικές από της πιο διάσημες παραλλαγές τον GAN είναι οι παρακάτω:

### 2.2.5.1 Conditional GAN

Μία από της πιο διάσημες παραλλαγές GAN είναι το conditional GAN[16](αυτό που χρησιμοποιήθηκε στην εργασία,Εικόνα 20) .Η διαφορά του από ένα κανονικό GAN είναι ότι στο conditional GAN υπάρχει και ένα condition (ένα vector) για παράδειγμα z είναι ο θόρυβος και το condition είναι ένα vector με features από αυτοκίνητα (τύπος,χρώμα,κυβικά) αυτά ενώνονται και περνιούνται ως είσοδο στον generator έτσι θα πρέπει ο generator να παράξει ένα αυτοκίνητο με της προϋποθέσεις που έχουν βαλθεί,στον Discriminator περνιέται και σε αυτόν το condition.

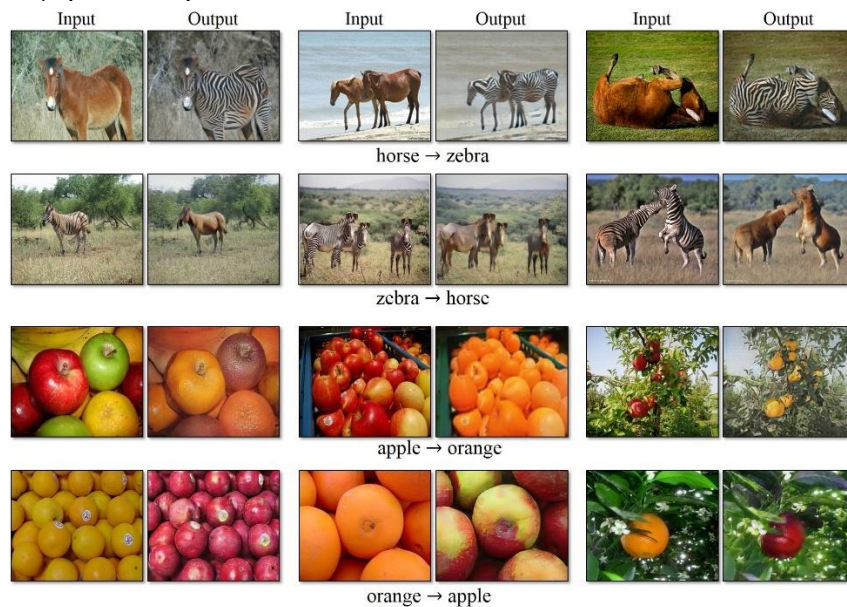
Ουσιαστικά από δίκτυο **χωρίς** επίβλεψη που είναι ο generator σε ένα κανονικό GAN στο conditional GAN ο generator γίνεται **με** επίβλεψη.



Εικόνα 20: Conditional GAN

### 2.2.5.2 Cycle GAN

Τα Cycle GAN [17] μαθαίνουν να “μεταφράζουν” μία εικόνα σε μία άλλη δηλαδή για παράδειγμα να μετατρέψει τα μήλα σε πορτοκάλια (Εικόνα 21).



Εικόνα 21 Cycle GAN αποτελέσματα

### 2.2.5.3 SRGAN (Super-Resolution)

Τα SRGAN [18] όπως λέει και το όνομα τους μεγαλώνουν την ανάλυση μιας εικόνας.(Εικόνα 22)



Εικόνα 22 SRGAN

### 2.2.5.4 pix2pix

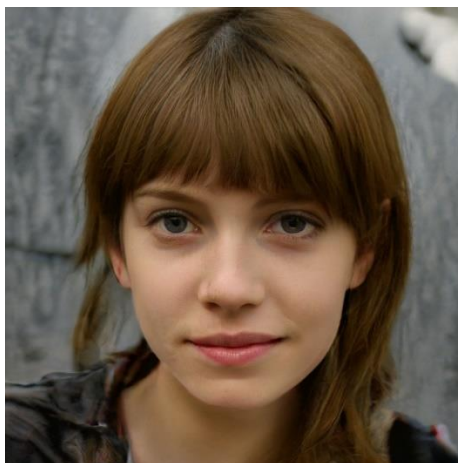
Το pix2pix [19] κάνει την ίδια “δουλειά” με το cycle GAN η κύρια διαφορά μεταξύ τους είναι ότι σε αντίθεση με το Pix2Pix GAN που αποτελείται από μόνο δύο δίκτυα (Discriminator και Generator), το Cycle GAN αποτελείται από τέσσερα δίκτυα (δύο Discriminators και δύο Generators).

### 2.2.6 Που Χρησιμοποιούνται

Τα GAN έχουν πάρα πολλές χρήσεις παρακάτω θα αναφερθούν μερικές από της πιο σημαντικές χρήσεις :

#### 1. Σύνθεση εικόνας και βίντεο

Η πιο διάσημη χρήση του GAN είναι η δημιουργία ή σύνθεση εικόνων. Για παράδειγμα ένα μοντέλο της Nvidia το StyleGAN δημιουργεί ρεαλιστικά πρόσωπα “ψεύτικων” ανθρώπων (Εικόνα 23).



*Εικόνα 23 Εικόνα από StyleGAN*

## **2. “Πρόβλεψη” βίντεο**

Η DeepMind δημοσίευσε ένα paper [20] στο οποίο δημιούργησαν ένα GAN το οποίο χρησιμοποιεί 2 discriminator το DVD-GAN το οποίο μπορεί να παράγει βίντεο τον 256 x 256 pixel με πρόβλεψη μέχρι και 48 frames.

Το 2018 το πανεπιστήμιο του Cambridge δημοσίευσαν ένα μοντέλο το οποίο ονομάζεται DeepRay [21] το οποίο δημιουργεί frames σε παραμορφωμένα βίντεο.

## **3. Μουσική**

Γενικά τα GAN είναι πολύ καλά στο να παράγουν πολυμέσα οπότε και μουσική 2 τέτοια δίκτυα είναι το deepComposer της amazon και ένα Conditional LSTM-GAN [22]

## **4. Τεχνητή φωνή**

Η Google με το Imperial college του Λονδίνου δημιούργησαν ένα GAN-TTS [23] το οποίο παράγει αληθοφανές φωνές.

## **5. Φάρμακα**

Στον ιατρικό τομέα έχουν χρησιμοποιηθεί τα GAN για να παράγουν δεδομένα όπου άλλα μοντέλα θα χρησιμοποιήσουν, δηλαδή datasets ή εισόδους για τα νευρωνικά δίκτυα. Αλλά έχουν ερευνηθεί GAN και για άλλες λειτουργίες ένα από αυτά είναι το Snore-GAN το οποίο αναπτύχθηκε από τα Αυτοκρατορικό Κολέγιο του Λονδίνου (το Imperial College), το πανεπιστήμιο του Άουγκσμπουργκ και το Πολυτεχνείο του Μονάχου και το οποίο συνθέτει δεδομένα για να καλύψει κενά σε αληθινά δεδομένα. Επίσης η Nvidia μαζί με την κλινική Mayo και το κέντρο κλινικών επιστημονικών δεδομένων MGH και BWH δημιούργησαν ένα νευρωνικό δίκτυο το οποίο παράγει εικόνες MRI από εγκεφάλους με καρκινικούς όγκους [24]. Μια εταιρεία η οποία λέγεται Insilico Medicine χρησιμοποίησε GAN για να φτιάξει μοριακές δομές σε με γνωστό σύμπλοκο (ένα σύνθετο βιομόριο) αλλά χωρίς στόχο (μια πρωτεΐνη που συσχετίζεται με μία διαδικασία ασθένειας). Η ομάδα των ερευνητών της εργάζεται ενεργά σε προγράμματα ανακάλυψης φαρμάκων στον καρκίνο, τις δερματολογικές παθήσεις, την ίνωση, τη νόσο του Πάρκινσον, τη νόσο Αλτσχάιμερ, το ALS, τον διαβήτη, τη σαρκοπενία και τη γήρανση. Το GAN αυτό



ονομάζεται generative tensorial reinforcement learning ή GENTRL για συντομία, και κατάφερε να βρει έξι πολλά υποσχόμενες θεραπείες για την ίνωση μέσα σε 21 ημέρες, εκ των οποίων μία από αυτές είχε καλά αποτελέσματα σε πειράματα με ποντίκια. [25]

## 6. Deepfakes

Τα deepfakes είναι βίντεο στα οποία τα πρόσωπα ή το σώμα των ανθρώπων αλλάζουν με άλλα. Αυτή η χρήση των GAN είναι και η πιο αμφιλεγόμενη διότι έχει χρησιμοποιηθεί για κακούς σκοπούς.

## 7. Ανίχνευση deepfakes

Όπως αναφέρθηκε παραπάνω τα GAN μπορούν να παράγουν deepfakes, αυτό δημιουργεί προβλήματα. Έτσι χρειάστηκε να ερευνηθούν τρόποι ανίχνευσης των deepfakes. Η έρευνα δείχνει ότι τα GAN θα μπορούν να ξεριζώσουν τα deepfakes εξίσου αποτελεσματικά. Ένα έγγραφο που δημοσιεύτηκε στο Arxiv.org περιγράφει το spamGAN [26] το οποίο μπορεί να ανιχνεύει τα deepfakes.

## 8. Κείμενο σε εικόνα

Ένα από τα πιο απαιτητικά προβλήματα στον κόσμο της μηχανικής όρασης είναι η σύνθεση εικόνων από περιγραφές κειμένου. Τα GAN έχουν παρατηρηθεί ότι έχουν πολύ καλές επιδόσεις. [27]

# Κεφάλαιο 3 : Προτεινόμενη μελέτη

Ο Μετασχηματισμός των EEG σημάτων σε εικόνες με GAN είναι κάτι που δεν έχει μελετηθεί σε μεγάλο βαθμό. Υπάρχουν τρεις έρευνες για αυτό το θέμα από το ίδιο ερευνητικό κέντρο το PeRCeiVe Lab οι οποίες έχουν τους τίτλους «Brain2Image: Converting Brain Signals into Images» [28], «Generative Adversarial Networks Conditioned by Brain Signals» [29] και «Decoding brain representations by multimodal learning of neural activity and visual features» [30].

Σε αυτό το κεφάλαιο θα μελετηθούν οι παραπάνω ερευνητικές εργασίες και θα συζητηθεί η προσέγγιση της παρούσας πτυχιακής εργασίας.

## 3.1 Προσέγγιση

Τα βήματα που θα πραγματοποιηθούν είναι τα εξής :

- 1) Θα δημιουργηθεί ένας encoder ο οποίος ουσιαστικά θα είναι ένας feature extractor για τα EEG σήματα.
- 2) Θα δημιουργηθεί ένα conditional GAN το οποίο θα παίρνει τα features από τον encoder και θα τα ενώνει με ένα θόρυβο, ο οποίος μετά την έξοδο θα αντιστοιχηθεί στο latent

space. Το GAN θα εκπαιδευτεί και θα έχει ως έξοδο εικόνες που αντιστοιχούν στην κλάση η οποία αντιστοιχεί το ‘condition’.

Στη συνέχεια θα γίνουν μετρήσεις για τα αποτελέσματα του GAN.

### **Τι είναι το latent space;**

Το latent space επίσης λέγεται και latent feature space και χώρος ενσωμάτωσης (Εικόνα 24), είναι μία ενσωμάτωση ενός συνόλου στοιχείων μέσα σε μία πολυπλοκότητα (manifold) όπου τα στοιχεία που μοιάζουν μεταξύ τους είναι πιο κοντά το ένα στο άλλο μέσα στο latent space. Η θέση στο latent space μπορεί να ειπωθεί ότι ορίζεται από ένα σύνολο latent μεταβλητών που προκύπτουν από της ομοιότητες των στοιχείων.



*Εικόνα 24 Ένα παράδειγμα latent spaces*

Υπάρχουν αλγόριθμοι η οποίοι δημιουργούν ενσωματώσεις latent space δεδομένου ενός συνόλου δεδομένων και μίας συνάρτησης ομοιότητας.

Τα GAN αυτό που κάνουν ουσιαστικά είναι να αντιστοιχούν το latent space σε εικόνα. Όταν ένα GAN παράγει μία εικόνα X και μία Y, πρακτικά ‘περπατάμε’ την γραμμική διαδρομή μεταξύ δύο latent spaces που δημιούργησαν της εικόνες και μπορούν να παραχθούν οι ενδιάμεσες εικόνες.

Το latent space για τα GAN είναι μια κατανομή Gauss με μέσο όρο 0 και τυπική απόκλιση 1 που σημαίνει ότι οι αριθμοί είναι ανάμεσα στο -1 και το 1.

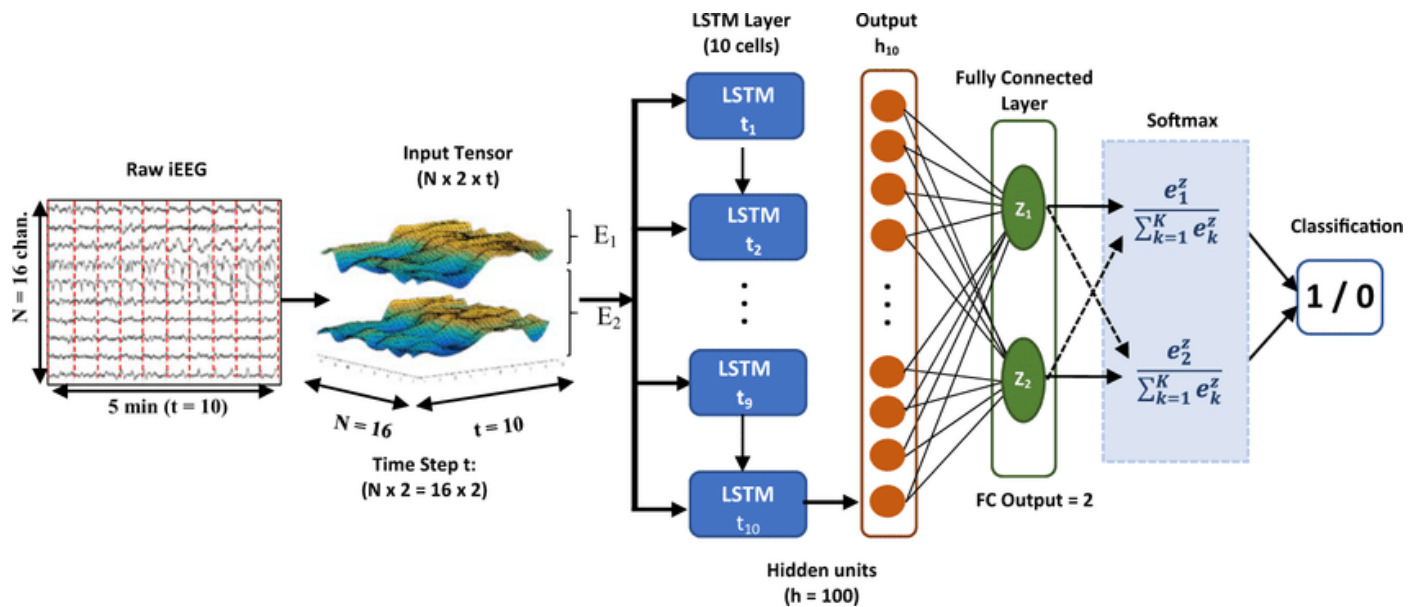
### **3.1.1 Feature extractor**

Το πιο σημαντικό είναι να βρεθεί ή να δημιουργηθεί ένα dataset, μόλις γίνει αυτό το επόμενο βήμα είναι να γίνει το feature extraction.



Όπως είναι γνωστό τα νευρωνικά δίκτυα παίρνουν εισόδους και παράγουν εξόδους που περιγράφουν κλάσεις είτε ό,τι άλλο ζητηθεί, αυτό σημαίνει ότι οι τιμές στα hidden layer είναι features τα οποία μετά την εκπαίδευση μπορεί το νευρωνικό να τα αναγνωρίσει και να κατανοήσει. Όπως αναφέρθηκε στο κεφάλαιο 1 και [ενότητα 8](#) υπάρχουν πολλοί τρόποι για feature extraction σε σήματα EEG τα νευρωνικά δίκτυα έχουν αποδειχθεί ότι είναι ίσως και τα καλύτερα συστήματα για να γίνει classification, έτσι για feature extraction είναι πολύ καλά, τα δύο πιο συνηθισμένα νευρωνικά δίκτυα που χρησιμοποιούνται για classification σημάτων EEG είναι τα CNN και τα LSTM (και RNN). Σε ένα CNN μετατρέπονται τα σήματα σε εικόνες και μετά γίνεται η εκπαίδευση, είναι καλά και αξιόπιστα αλλά τα καλύτερα είναι τα LSTM.

Σε αυτή την εργασία θα χρησιμοποιηθεί ένα LSTM με ένα layer και 128 νευρώνες το οποίο layer θα συνδέεται με ένα layer ReLU εδώ θα παρθούν τα features αλλά θα έχει και ένα softmax layer (Εικόνα 25).



Εικόνα 25 Ένα παράδειγμα classification EEG

Το Softmax θα χρειαστεί διότι θα γίνει το classification, μετά από την εκπαίδευση το softmax layer δεν θα χρειαστεί.

Τα βήματα που θα γίνουν είναι :

- 1) Δημιουργία του LSTM
- 2) Εκπαίδευση του LSTM
- 3) Περνιούνται τα EEG από το LSTM και λαμβάνονται τα features
- 4) Συνέχεια στο Conditional GAN

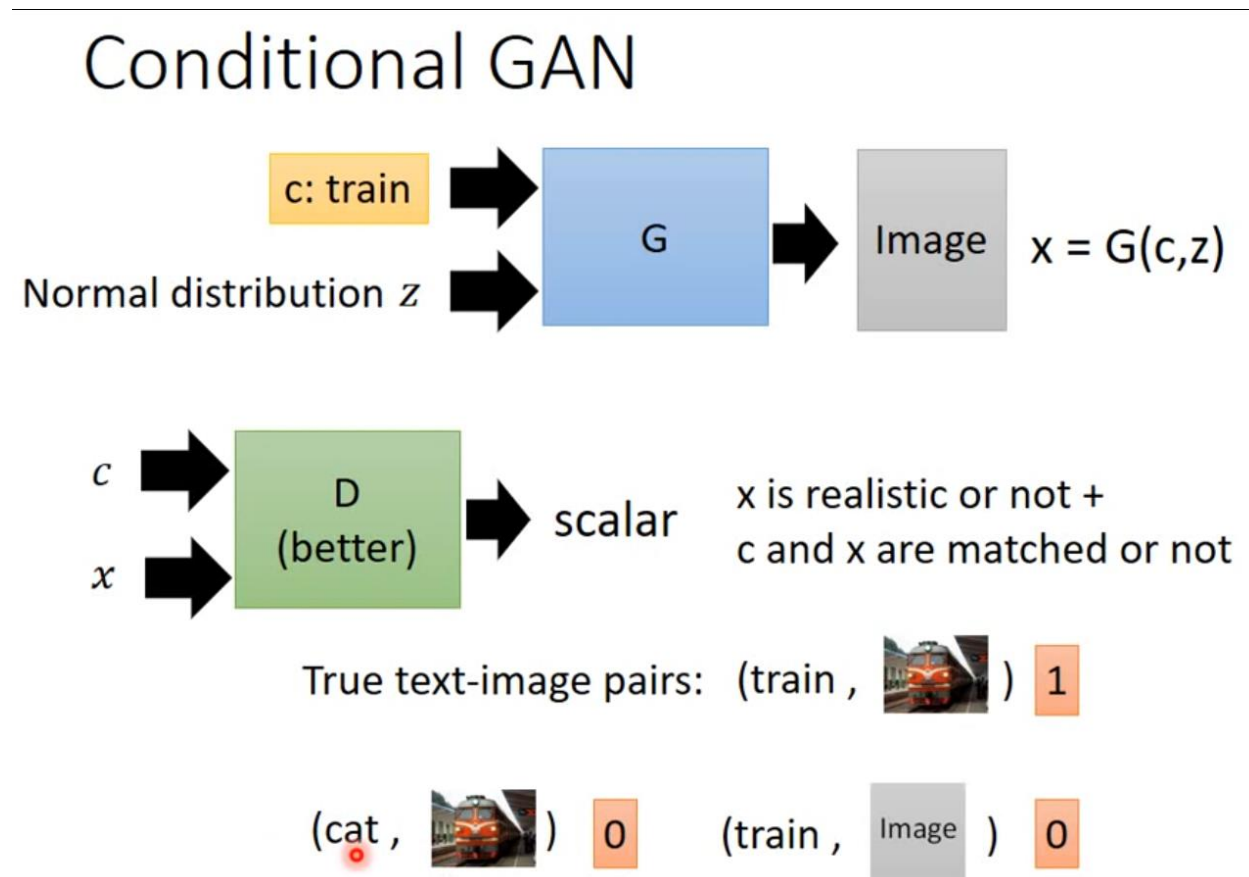
### 3.1.2 Conditional GAN

Στο προηγούμενο κεφάλαιο αναφέρεται ότι τα GAN έχουν πολλά είδη ένα από αυτά είναι και το Conditional GAN (Εικόνα 26).

Στο conditional GAN ο Generator είναι ίδιος με ενός κανονικού GAN με διαφορά την είσοδο η οποία είναι ένας θόρυβος που ενώνετε με τα features των δεδομένων.

Ο Discriminator παίρνει για είσοδο και την εικόνα αλλά και τα features και η 'δουλειά' του είναι να: 1) βρει αν είναι αληθοφανείς η εικόνα και 2) αν η εικόνα αντιστοιχεί στο condition (τα features).

Η εκπαίδευση τους είναι εξαιρετικά δύσκολη όπως και σε ένα κανονικό GAN απλώς είναι λίγο πιο περίπλοκη και πιο χρονοβόρα (η εκπαίδευση τον GAN είναι γενικά πολύ αργή αλλά η έξτρα παράμετροι των features προσθέτουν παραπάνω χρόνο ).



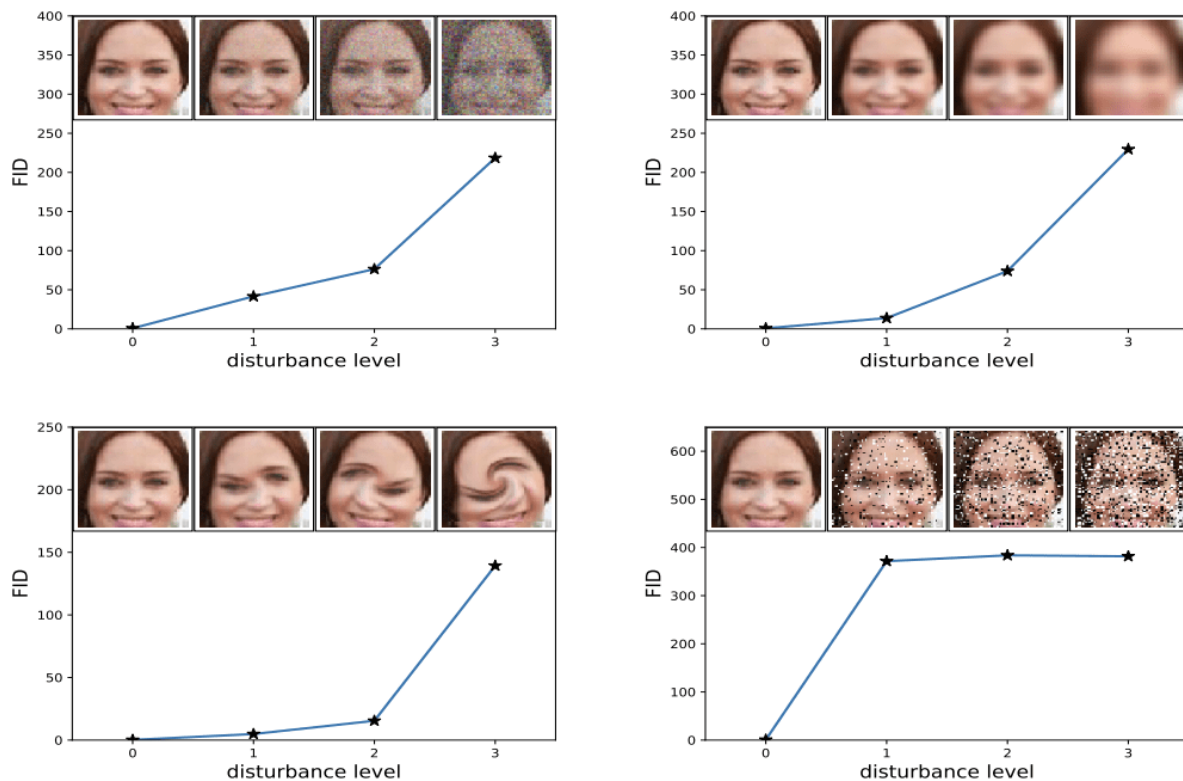
Εικόνα 26 Παράδειγμα Conditional GAN

### 3.1.3 Μετρήσεις

Σημαντικό είναι να δειχθούν και οι μέτρησις για να μπορεί να γίνει κρίση της αποτελεσματικότητας του GAN, στην εργασία αυτή θα υπάρχουν τρεις μετρήσεις το FiD το IS και το SSIM.

#### **Fid:**

Το FiD [31] ή με το ολοκληρωμένο του όνομα Fréchet inception Distance μετράει την απόσταση δύο dataset εικόνων (Εικόνα 27). Αυτό σημαίνει ότι όσο πιο μικρή η απόσταση τόσο πιο ‘ίδιες’ είναι οι εικόνες. Άμα το score είναι πάρα πολύ μικρό τότε υπάρχει ο κίνδυνος για overfitting (overfitting είναι όταν το νευρωνικό ‘μαθαίνει’ πάρα πολύ καλά τα δεδομένα της εκπαίδευσης και μόνο αυτά. Αυτό συμβαίνει όταν ένα μοντέλο μαθαίνει τις λεπτομέρειες και τον θόρυβο στα δεδομένα εκπαίδευσης σε τέτοιο βαθμό που επηρεάζει αρνητικά την απόδοση του μοντέλου σε νέα δεδομένα.)



*Εικόνα 27 Frechet Inception Distance*

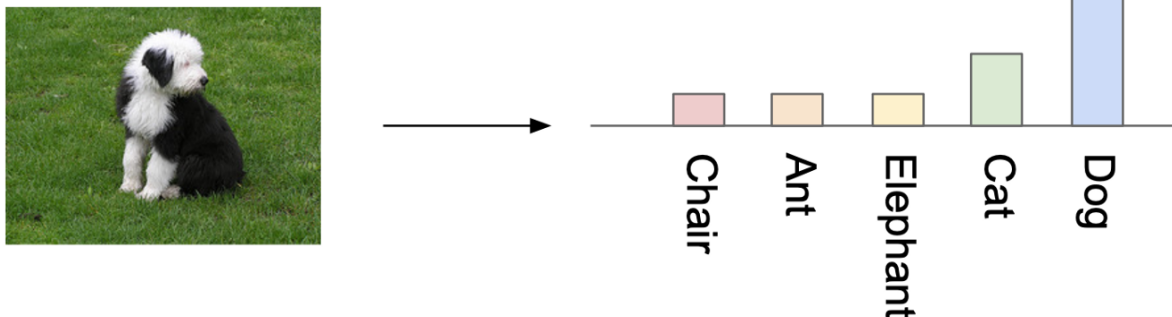
## IS:

Το IS [32] ή Inception Score μετράει το πόσο ρεαλιστική είναι η έξοδος ενός GAN. Το IS μετράει δύο πράγματα, πρώτων οι εικόνες να έχουν ποικιλία και δεύτερων οι εικόνες να μοιάζουν με κάτι πχ ένα σκύλο. Άμα ισχύουν και τα δύο τότε το score θα είναι μεγάλο αν δεν ισχύουν το score θα είναι μικρό ουσιαστικά είναι ‘αντίθετο’ του FiD στο θέμα το ότι το FiD όσο πιο μικρό τόσο πιο καλό.

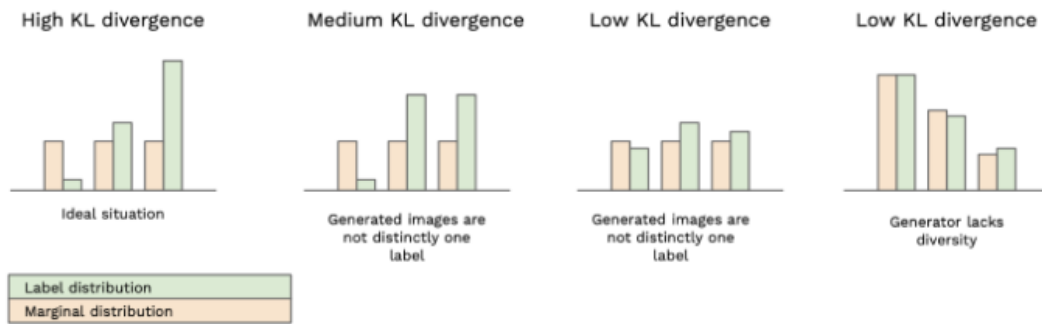
Το inception score και το FiD υπολογίζονται από ένα προ-εκπαιδευμένο μοντέλο που ονομάζεται Inception. Στο FiD χρησιμοποιούνται οι ενεργοποιήσεις του τελευταίου επιπέδου πριν της εξόδους του Inception v3. Το επίπεδο αυτό έχει 2048 ενεργοποιήσεις αυτό ονομάζεται coding vector ή feature vector. πρώτα περνάει της αληθινές φωτογραφίες και μετά της ‘ψεύτικες’ και μετά χρησιμοποιείτε η παρακάτω συνάρτηση:

$$d^2 = \|\mu_1 - \mu_2\|^2 + \text{Tr}(C_1 + C_2 - 2\sqrt{C_1 C_2})$$

Το IS από την άλλη ‘βλέπει’ την έξοδο του inception και αν μία από της εξόδους του έχει μεγάλη πιθανότητα τότε σημαίνει ότι η εικόνα μοιάζει σε κάτι και ότι μπορεί να το διακρίνει.



Μετά ‘βλέπει’ και άλλες φωτογραφίες και βλέπει της πιθανότητες τους. Άμα δεν υπάρχουν κορυφές και οι περισσότερες είναι στο ίδιο επίπεδο τότε σημαίνει ότι το δίκτυο παράγει ποικίλα δεδομένα. Οπότε αν ενωθούν αυτά τα δύο τότε μπορεί να παρθεί το IS. Η συνάρτηση για τον υπολογισμό του IS είναι ο Kullback–Leibler divergence (Εικόνα 28).

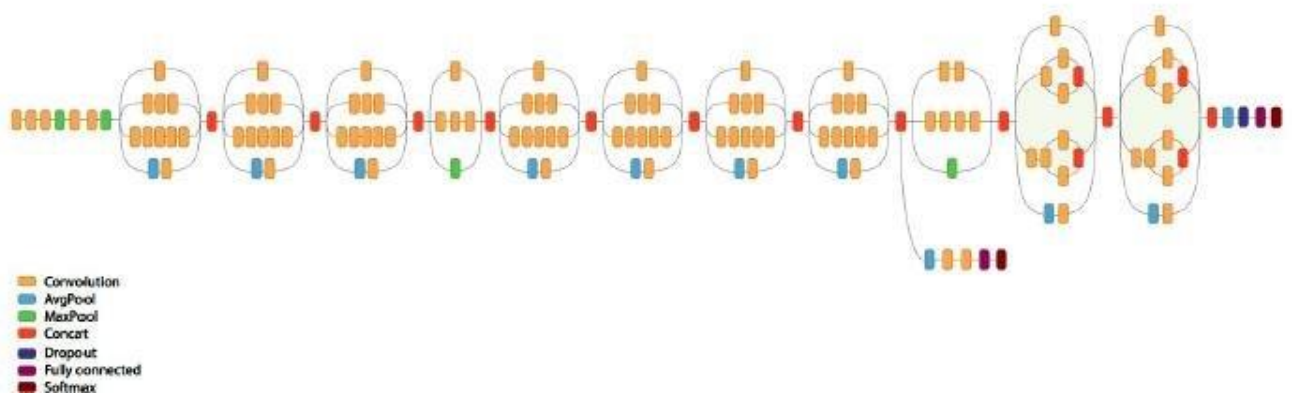


Εικόνα 28 Το Ιδανικό IS

Το IS είναι ο μέσος όρος των αποκλίσεων KL(KL divergence) όλων των φωτογραφιών μετά το πέρασμα κάθε μίας χωριστά από τον KL.

### Τι είναι το Inception v3 :

Το Inception v3 (Εικόνα 29) είναι η τρίτη έκδοση του μοντέλου inception ενός νευρωνικού δικτύου που εκπαιδεύτηκε στο να κάνει ταξινόμηση στο ImageNet (το οποίο έχει 1000 κλάσεις ) το έχει δημιουργήσει η Google και η αρχιτεκτονική του είναι :



Εικόνα 29 Inception V3

## Κεφάλαιο 4 Πειράματα

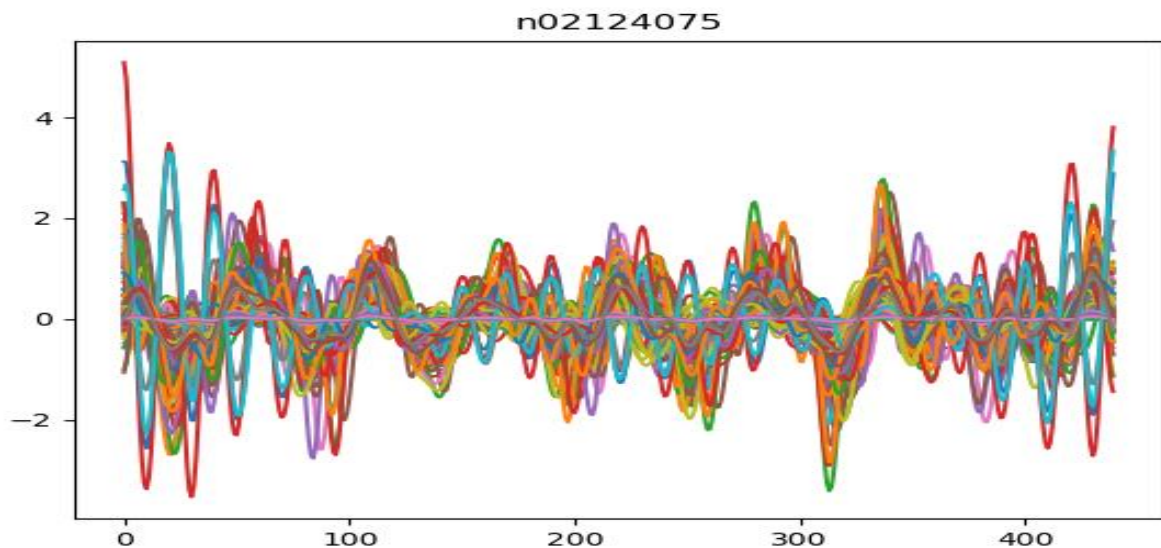
### 4.1 Datasets

Για να μπορέσει να γίνει ο μετασχηματισμός EEG σε εικόνα πρέπει πρώτα να βρεθούν δεδομένα EEG και εικόνες. Το dataset για τα EEG που χρησιμοποιήθηκε είναι το «EEG Data for Visual Classification» του PeRCeiVe Lab [30] [33]. Σε αυτό το dataset υπάρχουν δεδομένα EEG από 6 άτομα. Οι εγγραφές για κάθε άτομο αποτελούνται από 40 κλάσεις και για κάθε κλάση υπάρχουν 50 εικόνες, οπότε 2000 εικόνες. Οι εικόνες εμφανιζόταν διαδοχικά σε μία μόνο ακολουθία. Η κάθε εικόνα εμφανιζόταν για 0.5 δευτερόλεπτα. Μετά από κάθε κλάση εμφανιζόταν μαύρη οθόνη στην οποία η εγγραφή συνεχιζόταν. Κάθε EEG αποτελείται από 128 κανάλια καταγεγραμμένα για 0.5 δευτερόλεπτα και 1 kHz ρυθμό δειγματοληψίας, οπότε για κάθε κανάλι υπάρχουν 500 δείγματα, δηλαδή το κάθε EEG είναι 128x500. Τα δεδομένα είναι φιλτραρισμένα σε τρεις περιοχές συχνοτήτων : 14-70Hz, 5-95Hz και 55-95Hz.

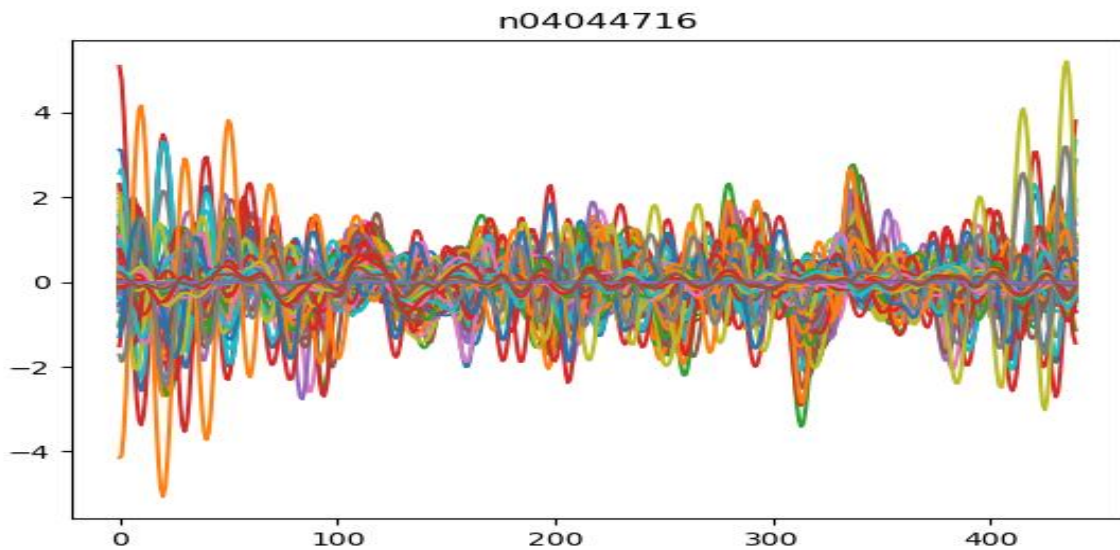
Κωδικός	Κλάση
n02106662	German shepherd
n02124075	Egyptian cat
n02281787	Lycaenid butterfly
n02389026	sorrel horse
n02492035	Capuchin
n02504458	Elephant
n02510455	Panda
n02607072	Anemone fish
n02690373	Airliner
n02906734	Broom
n02951358	Canoe
n02992529	Cellphone
n03063599	Mug
n03100240	Convertible
n03180011	Desktop PC
n03197337	Digital watch
n03272010	Electric guitar
n03272562	Electric locomotive
n03297495	Espresso maker
n03376595	Folding chair
n03445777	Golf ball
n03452741	Piano
n03584829	Iron

n03590841	Jack-o'-lantern
n03709823	Mailbag
n03773504	Missile
n03775071	Mitten
n03792782	Mountain bike
n03792972	Mountain tent
n03877472	Pyjama
n03888257	Parachute
n03982430	Pool table
n04044716	Radio telescope
n04069434	Reflex camera
n04086273	Revolver
n04120489	Running shoe
n07753592	Banana
n07873807	Pizza
n11939491	Daisy
n13054560	Bolete

Οι κλάσεις με τους κωδικούς τους.







*Εικόνα 30 Δύο EEG με διαφορετικές κλάσεις*

**Επιλέχτηκε αυτό το Dataset διότι :**

**1) δεν υπάρχουν πολλά dataset EEG για classification (ταξινόμηση) (διότι είναι δύσκολο να 'φτιαχτούν').**

**2) αυτό το dataset έχει πιο πολλά δεδομένα από άλλα παρόμοια.**

**3) τα δεδομένα είναι φιλτραρισμένα.**

**4) έχουν δεδομένα από 6 διαφορετικούς ανθρώπους.**

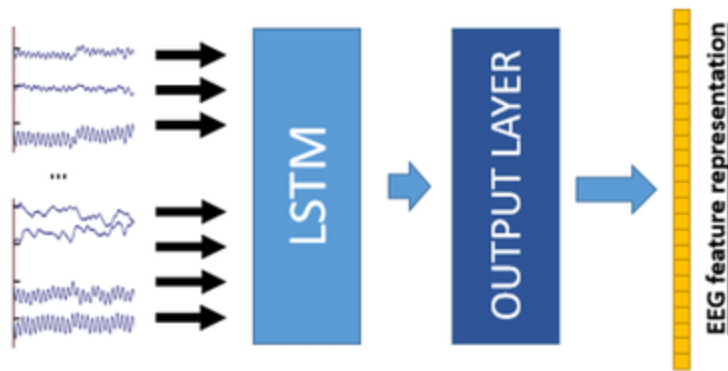
**5) έχουν πολλές κλάσεις**

**6) έχει δημιουργηθεί από ερευνητικό κέντρο (το PeRCeiVe Lab)**

Για της εικόνες,παρθήκαν από το IMAGENET και έγινε και augmentation για να είναι όλες οι εικόνες ίδιες σε αριθμό.Για κάθε κλάση 1546 εικόνες . Μερικά δήγματα φένονται στην εικόνα 31.







*Εικόνα 32 Η αρχιτεκτονική LSTM*

Αυτό το δίκτυο εκπαιδεύτηκε σε όλες της φιλτραρισμένες συχνότητες σαν ένας ταξινομητής (classifier) στις 200 εποχές μετά αποθηκεύτηκαν τα accuracies στον φάκελο 'encoder\_model/metrics' και το καλύτερο εύρος ήταν το 14-70Hz με accuracy validation : 20, accuracy test : 22, accuracy train : 96. Μετά περάστικαν όλα τα δεδομένα των EEG από το LSTM και αποθηκεύτικαν οι εξόδοι από την ReLU στο αρχείο 'Relu.csv' και τα Labels στο αρχείο 'Labels.csv'.

Το LSTM δημιουργήθηκε και εκπαιδεύτηκε στο αρχείο 'encoder.py'.

Τα features παρθήκαν από το αρχείο 'features\_from\_EEG.py'.

#### 4.2.2 Conditional GAN

Η αρχιτεκτονική για το GAN είναι :

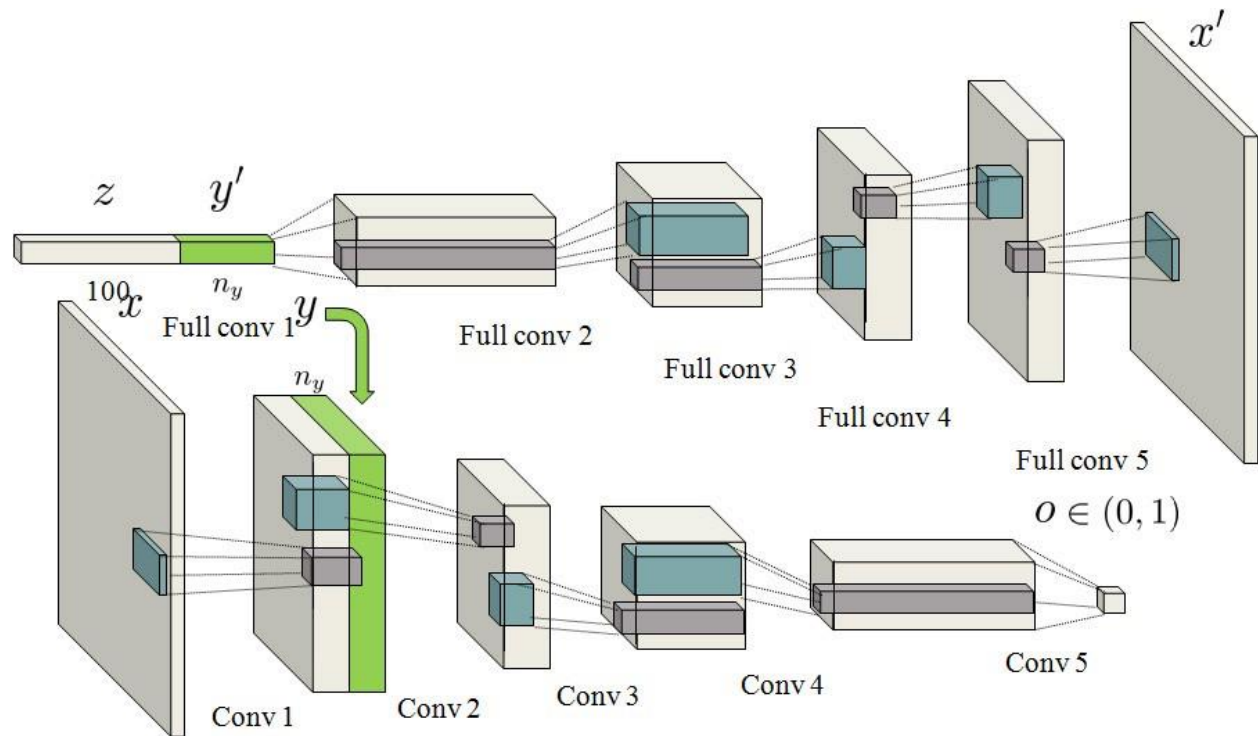
Generator:

```
(ConvTransposed2d
BatchNorm2d
ReLU) * 4
ConvTransposed2d
Tanh
```

Discriminator:

```
Conv2d
Leaky ReLU (0.2)
(Conv2D
BatchNorm2D
Leaky ReLU (0.2)) *3
Conv2d
Flatten
Linear
```

Αξίζει να σημειωθεί ότι αυτή η αρχιτεκτονική του GAN έχει αποδειχθεί ότι είναι η πιο αποτελεσματική [34] (εικόνα 33).



Εικόνα 33 Η αρχιτεκτονική του Conditional GAN

Το GAN δημιουργήθηκε και εκπαιδεύτηκε στο αρχείο 'EEG\_GAN\_pytorch.py'. Και τα αποτελέσματα της εκπαίδευσης είναι στον φάκελο 'results/false'.

### 4.3 Κώδικας και δομή αρχείων

Τα αρχεία του κώδικα είναι τα εξής:

- 'Dataset40\_class.py'
- 'encoder.py'
- 'features\_from\_EEG.py'
- 'img\_dataset.py'
- 'FiD\_Is.py'
- 'Similarities.py'
- 'EEG\_GAN\_pytorch.py'

Τα αρχεία του dataset:  
'EEG\_14\_70\_std.pth'  
'splits\_by\_image.pth'  
'Relu.csv'  
'Labels.csv'  
'unique.csv'

Οι φωτογραφίες είναι στον φάκελο 'data' > 'images' και '64\_64\_2'

## **Dataset40\_class.py**

Αυτό το αρχείο περιέχει βοηθητικές κλάσεις και συναρτήσεις για το dataset

Κλάση **EEGDataset** : διαβάζει το dataset και επιστρέφει τα δεδομένα, τα labels τα targets και τα EEG.

Κλάση **Splitter** : είναι για να διαβάζει ένα από τα αρχεία του dataset το 'splits\_by\_image.pth' (τελικά δεν χρησιμοποιήθηκε ).

Συνάρτηση **get\_all\_labels()** : επιστρέφει τα labels

Συνάρτηση **get\_labels\_unique()** : επιστρέφει τα ξεχωριστά labels

Κλάση **EEGtoImageDataset**: (δεν χρησιμοποιήθηκε) είναι ίδια με την EEGDataset απλώς επιστρέφει τα EEG σαν εικόνες.

## **Encoder.py**

Κλάση **Model** : Το LSTM μοντέλο

Μετά σε αυτό το αρχείο εκπαιδεύεται το LSTM με όλες της φιλτραρισμένες συχνότητες και αποθηκεύει ανά 5 εποχές τα βάρη του LSTM επίσης αποθηκεύει τα accuracies και losses.

## **features\_from\_EEG.py**

Σε αυτό το αρχείο 'τρέχει' το LSTM με όλα τα δεδομένα και τα αποθηκεύει σε csv.  
Τα csv 'LSTM.csv', 'Relu.csv', 'Label.csv'.

## **img\_dataset.py**

Αυτό το αρχείο είναι για τη δημιουργία του dataset των εικόνων δηλαδή το κατέβασμα των εικόνων από το ImageNet και το augmentation των εικόνων.

**Ο Downloader δεν δουλεύει πλέον διότι άλλαξε το ImageNet (το site τους).Για να κατεβεί το dataset μετά την αλλαγή γίνεται μέσω του Kaggle.**

Οι φωτογραφίες είναι ‘επεξεργασμένες’ σε 64x64 pixels,η επεξεργασία γίνεται ταυτόχρονα με το κατέβασμά τους με την βοήθεια της βιβλιοθήκης Pillow (PIL) και το κατέβασμα έγινε με την βοήθεια της βιβλιοθήκης Beautiful soup.

Η κάθε κλάση έχει 300 φωτογραφίες για την εκπαίδευση του νευρωνικού δικτύου,οπότε συνολικά είναι 12000 φωτογραφίες.

Ενώ κατεβήκαν για κάθε κλάση 1546 φωτογραφίες και διάλεξα 300 τυχαίες.Στης κλάσεις που είχαν λιγότερες από 1546 πραγματοποιήθηκε augmentation έτσι ώστε να έχουν όλες οι κλάσεις ίδιο αριθμό φωτογραφιών.

## **EEG\_GAN\_pytorch.py**

Αυτό το αρχείο είναι το αρχείο στο οποίο δημιουργεί το GAN και το εκπαιδεύει.

Κλάση **EEG\_Image\_Dataset**: διαβάζει και δημιουργεί το dataset των εικόνων.

Συνάρτηση **get\_id\_of\_label(Name)**: επιστρέφει τον ακέραιο αριθμό της κλάσης για παράδειγμα άμα είναι πιτζάμα θα επιστρέψει 10.

Συνάρτηση **weights\_init(m)**: Αυτή η συνάρτηση είναι αναγκαία για την αρχικοποίηση των βαρών του νευρωνικού [32].

Κλάση **Generator** : ο generator.

Κλάση **Discriminator**: ο Discriminator.

## **FiD\_Is.py**

Το αρχείο αυτό είναι το αρχείο με το οποίο γίνονται οι μετρήσεις FID και IS

Είναι όπως το EEG\_GAN\_pytorch.py με την διαφορά τι χρησιμοποίηση των βιβλιοθηκών torchvision και ignite οι οποίες εμπεριέχουν συναρτήσεις για της μετρήσεις.

## **Similarities.py**

Αρχείο για την μέτρηση ssim (Το ssim είναι μια αντιληπτική μέτρηση που ποσοτικοποιεί την υποβάθμιση της ποιότητας της εικόνας που προκαλείται από επεξεργασία όπως η συμπίεση δεδομένων ή από απώλειες στη μετάδοση δεδομένων).Ενώ έχει γίνει η μέτρηση ssim,συνήθως δεν χρησιμοποιείται σε GANs διότι τα GANs παράγουν νέες φωτογραφίες και οι πιο αποδεκτές μετρήσεις είναι οι FID και η IS.Αυτό που έγινε είναι το ότι πάρθηκαν 16 εικόνες από το dataset για κάθε κλάση (είναι αποθηκευμένες στον φάκελο for\_sim/true/κλάση ) και 16 από το GAN

(for\_sim/false/κλάση) και χρησιμοποιήθηκε η συνάρτηση structural\_similarity της βιβλιοθήκης skimage.metrics.

### **EEG\_14\_70\_std.pth**

Το dataset των EEG με φιλτραρισμένες στις συχνότητες 14 έως 70Hz.

### **splits\_by\_image.pth**

Βοηθητικό αρχείο από τον δημιουργό του dataset.

### **Relu.csv**

Αρχείο csv όπου είναι αποθηκευμένα τα feature από τα EEG που έχουν περάσει από το LSTM.

### **Labels.csv**

Τα labels για κάθε EEG όπου είναι με την ίδια σειρά με το Relu.csv (τα δεδομένα στο relu.csv είναι ανακατεμένα ).

### **unique.csv**

Στο αρχείο αυτό είναι η αντιστοίχιση του ονόματος των κλάσεων με το νούμερο τους.

## **4.4 Αποτελέσματα και Μετρήσεις**

### **4.4.1 Αποτελέσματα**

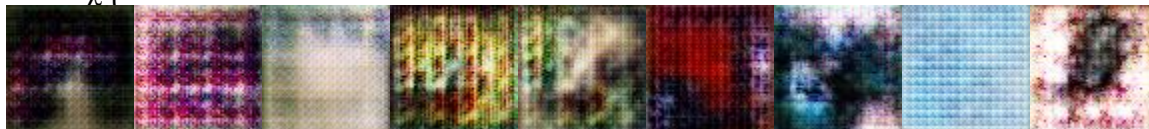
Μερικά από τα αποτελέσματά των εικόνων για της 20 πρώτες εποχές :

1<sup>η</sup> εποχή:





2<sup>η</sup> εποχή:



3<sup>η</sup> εποχή:



4<sup>η</sup> εποχή:



5<sup>η</sup> εποχή:



6<sup>η</sup> εποχή:



7<sup>η</sup> εποχή:



8<sup>η</sup> εποχή:





9<sup>η</sup> εποχή:



10<sup>η</sup> εποχή:



11<sup>η</sup> εποχή:



12<sup>η</sup> εποχή:



13<sup>η</sup> εποχή:



14<sup>η</sup> εποχή:



15<sup>η</sup> εποχή:



16<sup>η</sup> εποχή:



17<sup>η</sup> εποχή:





18<sup>η</sup> εποχή:



19<sup>η</sup> εποχή:



20<sup>η</sup> εποχή:



Μετά την 20<sup>η</sup> εποχή δεν διαφέρουν πολύ η φωτογραφίες από την 20<sup>η</sup> εποχή. Όπως παρατηρούνται οι εικόνες δεν μοιάζουν με τίποτα από της κλάσεις του dataset.

Το νευρωνικό δίκτυο εκπαιδεύτηκε για 50 εποχές σε gpu με το framework της pytorch. Χρησιμοποιήθηκε pytorch επειδή γενικά είναι γρηγορότερο από το keras/tensorflow, ο κώδικας είναι πιο “όμορφος” και επίσης το dataset και το βοηθητικό αρχείο ‘splits\_by\_image’ είναι σε τύπο αρχείου ‘.pth’ το οποίο χρησιμοποιείται και δημιουργείτε συνήθως από το pytorch.

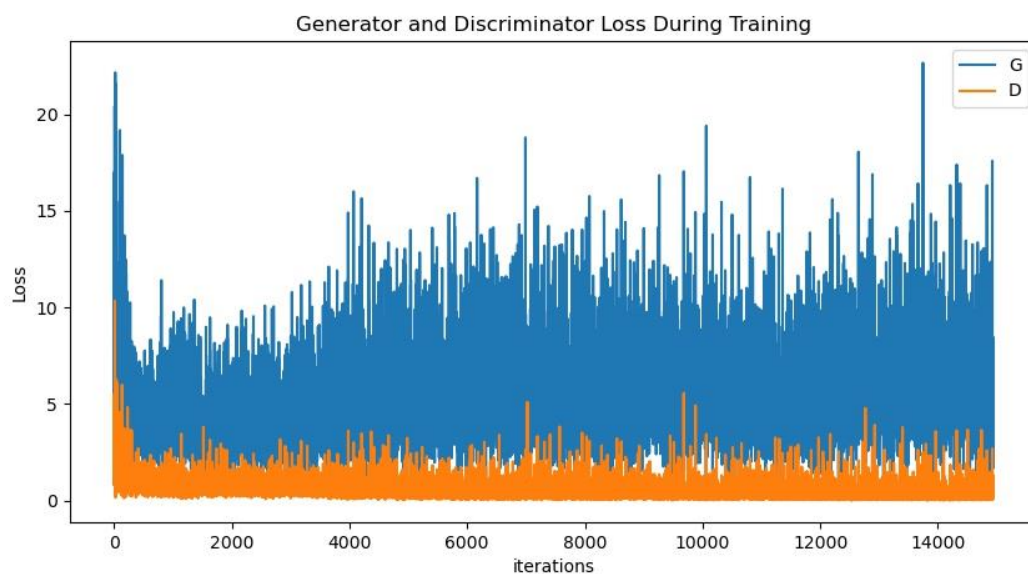
Οι φωτογραφίες που δημιουργεί το GAN είναι 64x64 pixels και φυσικά έγχρωμες.

#### 4.4.2 Μετρήσεις

Για της μετρήσεις των FID και IS ξανά εκπαιδεύτηκε το δίκτυο σε 20 εποχές επίσης χρησιμοποιήκαν της βιβλιοθήκες torchvision και ignite. Το ignite είναι μία βιβλιοθήκη υψηλού επιπέδου, η οποία βοηθάει στο να γίνει η εκπαίδευση αλλά και η αξιολόγηση του νευρικού δικτύου στο pytorch.

Ο παρακάτω πίνακας δείχνει τα αποτελέσματα των μετρήσεων και η Εικόνα 34 δείχνει το loss του generator και του discriminator στην διάρκεια της εκπαίδευσης.

Εποχή	Fréchet inception Distance	Inception Score
1	0.2692114092959352	1.4737281216453628
2	0.2346661602664523	1.742658732747686
3	0.27109605720837193	2.396288839832887
4	0.22428710610622704	2.1062956473618337
5	0.23643820524260808	1.6840346881984583
6	0.22237954841462584	1.8609725611616221
7	0.22136791566885464	1.9279578563923223
8	0.21003039829855702	1.771043750821079
9	0.21653501717164442	1.6539582229838106
10	0.30597010502376093	1.8228390469547457
11	0.3152774073776288	1.979254969077292
12	0.31274793695102027	2.1826749165886032
13	0.21606651787099462	2.3004097972306807
14	0.2096915179406642	1.848682136477969
15	0.2737618745279763	2.3035621254330545
16	0.24110796400568724	2.3161851319968263
17	0.19384937458143053	2.472822678457718
18	0.22155340511794552	2.271502005186409
19	0.2340431697296692	2.363420164244413
20	0.277713578924852	2.2473132244588014



Εικόνα 34 Το loss του Generator (G) και του Discriminator (D) κατά την εκπαίδευση

Για το SSIM :

Κλάση	M.O
0	0.06177580050903608
1	0.19144165325653978
2	0.13087509691342541
3	0.07036469947976873
4	0.07701661645581155
5	0.14051554794148002
6	0.07784813002752183
7	0.08526668186495573
8	0.04891047090330162
9	0.05938991329082893
10	0.06518103864361786
11	0.09674315211015473
12	0.06408686031425845
13	0.06125969776902211
14	0.04411163981846229
15	0.07688016667993403
16	0.13893702209208694
17	0.052214565319168495
18	0.11635875204816504

19	0.06506448579873945
20	0.08010332620359518
21	0.04786292624192514
22	0.10149368706869312
23	0.08222968687104398
24	0.059110512727433825
25	0.06716933976372523
26	0.0934520614820365
27	0.06913179464423347
28	0.09858887740147829
29	0.07505490361250802
30	0.05850161302809121
31	0.11253239393678227
32	0.1093991706234753
33	0.06136607690767009
34	0.07431149195828868
35	0.06011558506123927
36	0.18945473978697155
37	0.053385887645746144
38	0.10521361674115441
39	0.0863941286227827

Το Μ.Ο. είναι ο μέσος όρος των ssim κάθε φωτογραφίας από κάθε κλάση.

Το SSIM δεν είναι ο καλύτερος τρόπος για να βγούν μετρήσεις για ένα GAN. Το SSIM score για το συγκεκριμένο νευρωνικό δεν είναι καλό. Αυτό όπως ειπώθηκε και πιο πάνω συμβαίνει διότι οι εικόνες είναι καινούργιες και όχι οι ίδιες με αυτές του dataset.

## 4.5 Συμπεράσματα

Από της μετρήσεις συμπερένεται ότι στο GAN της συγκεκριμένης εργασίας δεν βγαίνουν ικανοποιητικά αποτελέσματα. Μία υπόθεση για το ότι τα αποτελέσματα δεν είναι ικανοποιητικά είναι ότι τα EEG γενικά δεν είναι και τα πιο αποτελεσματικά για την αναγνώριση αντικειμένων διότι είναι πολύ δύσκολο ένας άνθρωπος να μην σκέφτεται άλλα πράγματα ή να μην κουνιέται οπότε θα υπάρχει θόρυβος, ο οποίος φιλτράρετε αλλά δεν φεύγει τελείως.

Μία άλλη υπόθεση είναι το ότι τα δεδομένα δεν είναι επαρκείς και χρειάζονται και άλλα δεδομένα. Αυτό το dataset ήταν το καλύτερο που βρέθηκε και έχει 12000 EEG σήματα από 128 κανάλια.

Τέλος μία άλλη υπόθεση είναι το ότι το δίκτυο είναι underfitted.

Τα EEG όπως γράφτηκε στο κεφάλαιο 1 έχουν μειονεκτήματα τα οποία δεν τα καθιστούν και ως την καλύτερη επιλογή για ταξινόμηση αντικειμένων ένας καλύτερος τρόπος είναι τα mri αλλά είναι πιο δύσκολο να παρθούν.



## Βιβλιογραφία

- [1] «mayo clinic: EEG (electroencephalogram), about,» 15 Απρίλιος 2020. [Ηλεκτρονικό]. Available: <https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>.
- [2] «Wikipedia : Electroencephalography,» [Ηλεκτρονικό]. Available: <https://en.wikipedia.org/wiki/Electroencephalography>.
- [3] J. J. Bird, L. J. Manso, A. Ekart και D. R. Faria, « A Study on Mental State Classification using EEG-based Brain-Machine Interface,» *Madeira Island, Portugal: 9th international Conference on Intelligent Systems 2018*, 3 Δεκέμβριος 2018.
- [4] L. Vézard, P. Legrand, M. Chavent, F. Faïta-Aïnseba και L. Trujillo, «EEG classification for the detection of mental states,» *Applied Soft Computing ,Volume 32*, pp. 113-131, 20 Μάρτιος 2015.
- [5] J.-S. Huang, Y. Li, B.-Q. Chen, C. Lin και a. B. Yao, «An Intelligent EEG Classification Methodology Based on Sparse Representation Enhanced Deep Learning Networks». Shenzhen Institutes of Advanced Technology;Chinese Academy of Sciences, Shenzhen, China 30 Σεπτέμβριος 2020.
- [6] W. Zheng, J. Zhu, Y. Peng και B. Lu, «EEG-based emotion classification using deep belief networks,» *2014 IEEE International Conference on Multimedia and Expo (ICME), Chengdu, China*, pp. 1-6, 2014.
- [7] A. Khasnobish, A. Konar, D. N. Tibarewala, S. Bhattacharyya και R. Janarthanan, «Object Shape Recognition from EEG Signals during Tactile and Visual Exploration,» σε *Pattern Recognition and Machine Intelligence. PReMI 2013. Lecture Notes in Computer Science*, Berlin, Springer, 2013.
- [8] S. Bozinovski, M. Sestakov και L. Bozinovska, «Using EEG alpha rhythm to control a mobile robot,» σε *G. Harris, C. Walker (eds.) Proc. IEEE Annual Conference of Medical and Biological Society*, New Orleans, IEEE, 1988, pp. 1515-1516.
- [9] «products/NeuroSky MindWave Mobile 2 : sparkfun,» [Ηλεκτρονικό]. Available: <https://www.sparkfun.com/products/14758>.
- [10] «THE DIFFERENT TYPES OF SLEEP : goldensupplement,» 27 Σεπτέμβριος 2018. [Ηλεκτρονικό]. Available: <https://goldensupplement.org/the-different-types-of-sleep/>.
- [11] E. S. Kappenman και S. J. Luck, «The effects of electrode impedance on data quality and statistical significance in ERP recordings,» 2010.
- [12] «electroneurodiagnostics.org,» [Ηλεκτρονικό]. Available: <http://www.electroneurodiagnostics.org/resources/EEG-Material/7-Filters-in-the-Electroencephalogram.pdf>.
- [13] V. Sagar και K. Kumar, «Autoencoder Artificial Neural Network Public Key Cryptography in Unsecure Public channel Communication,» *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 2019.
- [14] A. Zhavoronkov, «Deep learning enables rapid identification of potent DDR1 kinase inhibitors,» *Nature Biotechnology*, pp. 1038-1040, 3 January 2019.

- [15] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville και Y. Bengio, «Generative Adversarial Nets,» Departement d'informatique et de recherche opérationnelle, Montreal, 2014.
- [16] M. Mirza και S. Osindero, «arxiv:1411.1784,» Thursday November 2014. [Ηλεκτρονικό]. Available: <https://arxiv.org/abs/1411.1784>.
- [17] J.-Y. Zhu, T. Park, P. Isola και A. A. Efros, «Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks,» ICCV, Venice, Italy, 2017.
- [18] C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang και W. Shi, «Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network,» 15 September 2016. [Ηλεκτρονικό]. Available: <https://doi.org/10.48550/arXiv.1609.04802>.
- [19] P. Isola, J.-Y. Zhu, T. Zhou και A. A. Efros, «Image-to-Image Translation with Conditional Adversarial Networks,» CVPR, Hawai'i Convention Center, 2016.
- [20] C. Aidan, D. Jeff και S. Karen, *ADVERSARIAL VIDEO GENERATION ON COMPLEX DATASETS*, London, UK: DeepMind, 2019.
- [21] S. Bakirtzis, K. Qiu, J. Zhang και I. Wassell, *DeepRay: Deep Learning Meets Ray-Tracing*, 2022.
- [22] Y. Yu, A. Srivastava και S. Canales, *Conditional LSTM-GAN for Melody Generation from Lyrics*, 2021.
- [23] M. Binkowski, J. Donahue, S. Dieleman, A. Clark, E. Elsen, N. Casagrande, L. C. Cobo και K. Simonyan, *HIGH FIDELITY SPEECH SYNTHESIS WITH ADVERSARIAL NETWORKS*, DeepMind, 2019.
- [24] N. Alarcon, «AI Can Generate Synthetic MRIs to Advance Medical Research,» Nvidia, 16 Σεπτέμβριος 2016. [Ηλεκτρονικό]. Available: <https://developer.nvidia.com/blog/ai-can-generate-synthetic-mris-to-advance-medical-research/>.
- [25] A. Zhavoronkov, Y. A. Ivanenkov, A. Aliper, M. S. Veselov, V. A. Aladinskiy, A. V. Aladinskaya, V. A. Terentiev, D. A. Polykovskiy, M. D. Kuznetsov, A. Asadulaev, Y. Volkov, A. Zholus και R. R. Shayakhmetov, «Deep learning enables rapid identification of potent DDR1 kinase inhibitors,» *Nature Biotechnology*, pp. 1038-1040, 2 Σεπτέμβριος 2019.
- [26] G. Stanton και A. A. Irissappane, *GANs for Semi-Supervised Opinion Spam Detection*, 2019.
- [27] K. Hu, W. Liao, M. Y. Yang και B. Rosenhahn, *Text to Image Generation with Semantic-Spatial Aware GAN*.
- [28] I. Kavasidis, S. Palazzo, C. Spampinato, D. Giordano και M. Shah, *Brain2Image: Converting Brain Signals into Images*, USA: MM: International Multimedia Conference, 2017.
- [29] S. Palazzo, C. Spampinato, I. Kavasidis και D. Giordano, *Generative Adversarial Networks Conditioned by Brain Signals*, Catania - Italy: IEEE, 2019.
- [30] S. Palazzo, C. Spampinato, I. Kavasidis, D. Giordano, J. Schmidt και M. Shah, «Decoding Brain Representations by Multimodal Learning of Neural Activity and Visual Features,» *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 2020.



- [31] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler και S. Hochreiter, *GANs Trained by a Two Time-Scale Update Rule*, arXiv:1706.08500v6, 2018.
- [32] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford και X. Chen, *Improved Techniques for Training GANs*, <https://doi.org/10.48550/arXiv.1606.03498>, 2016.
- [33] C. Spampinato, S. Palazzo, I. Kavasidis, D. Giordano, N. Souly και M. Shah, «Deep Learning Human Mind for Automated Visual Classification, International Conference on Computer Vision and Pattern Recognition,» CVPR , 2017.
- [34] A. Radford, L. Metz και S. Chintala, «UNSUPERVISED REPRESENTATION LEARNING,» International Conference on Learning Representations, 2016.
- [35] M. Warren και W. Pitts, «A Logical Calculus of Ideas Immanent in Nervous Activity,» σε *Bulletin of Mathematical Biophysics*, 1943, p. 115–133.
- [36] D. Ü. Elif, D. Cvetkovic και I. Cosic, «Eigenvector methods for analysis of Human PPG,,» σε *Alterations in sleep EEG activity during the hypopnoea episodes*, 2007.