

Εθνικό Μετσοβίο Πολύτεχνειο Σχολή Ηλεκτρολογών Μηχανικών και Μηχανικών Υπολογιστών Τομέας Ηλεκτρικής Ισχυός

## ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ ΕΙΚΟΝΩΝ ΤΟΥ ΟΥΡΑΝΟΥ ΜΕ ΧΡΗΣΗ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΓΙΑ ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΤΗΣ ΦΩΤΟΒΟΛΤΑΪΚΗΣ ΠΑΡΑΓΩΓΗΣ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Στυλιανός Π. Σχίζας

Επίβλεψη: Παύλος Σ. Γεωργιλάκης, Καθηγητής Ε.Μ.Π. Επόπτευση: Μάρκος Κουσουνάδης Κνούσεν, Υποψήφιος Διδάκτορας Ε.Μ.Π.

Αθήνα, Ιούνιος 2025



Εθνικό Μετσοβίο Πολύτεχνειο Σχολή Ηλεκτρολογών Μηχανικών και Μηχανικών Υπολογιστών Τομέας Ηλεκτρικής Ισχυός

## ΜΟΝΤΕΛΟ ΠΡΟΒΛΕΨΗΣ ΕΙΚΟΝΩΝ ΤΟΥ ΟΥΡΑΝΟΥ ΜΕ ΧΡΗΣΗ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΓΙΑ ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΤΗΣ ΦΩΤΟΒΟΛΤΑΪΚΗΣ ΠΑΡΑΓΩΓΗΣ

## ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Στυλιανός Π. Σχίζας

Επίβλεψη: Παύλος Σ. Γεωργιλάκης, Καθηγητής Ε.Μ.Π

Επόπτευση: Μάρκος Κουσουνάδης Κνούσεν, Υποψήφιος Διδάκτορας Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 27 Ιουνίου 2025

..... Παύλος Γεωργιλάκης Καθηγητής Ε.Μ.Π. ..... Άρης Ευάγγελος Δημέας Επίκ. Καθηγητής Ε.Μ.Π. ..... ΒασίλειοςΝικολαΐδης Επίκ. Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2025

.....

Στυλιανός Π. Σχίζας

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Στυλιανός Σχίζας, 2025. 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.

### ΠΕΡΙΛΗΨΗ

Στόχος της παρούσας διπλωματικής εργασίας είναι η υλοποίηση μοντέλου πρόβλεψης εικόνων του ουρανού με χρήση βαθιάς μάθησης για την βραχυπρόθεσμη πρόβλεψη της φωτοβολταϊκής (ΦΒ) παραγωγής. Για τον σκοπό αυτό πραγματοποιήθηκε κατασκευή τεσσάρων αρχιτεκτονικών μοντέλου που βασίζονται σε συνελικτικούς αυτοκωδικοποιητές (Convolutional Auto-Encoders – CAEs), με χρήση 2D και 3D συνελίξεων, ώστε για μια ακολουθία εισόδου εικόνων του ουρανού, το μοντέλο να μπορεί να προβλέψει μια ακολουθία εξόδου (μελλοντικές εικόνες).

Το παραπάνω εγχείρημα πραγματοποιείται με την βοήθεια ενός συνόλου δεδομένων επίγειων εικόνων του ουρανού, που λήφθηκαν από κάμερα παρακολούθησης του ουρανού σε διάστημα δεκατεσσάρων ημερών. Η αξιολόγηση των αποτελεσμάτων γίνεται με ποιοτικά (εποπτική αξιολόγηση), αλλά και ποσοτικά μέσα (μετρητικά αξιολόγησης). Η ανάπτυξη του κώδικα για τη δημιουργία των μοντέλων της παρούσας διπλωματικής εργασίας γίνεται στη γλώσσα προγραμματισμού Python.

Η παρούσα εργασία καινοτομεί με την προσθήκη της τεχνικής της ταξινόμησης των δεδομένων εικόνων σε καταστάσεις του ουρανού (ηλιοφάνεια, μερική συννεφιά και συννεφιά) επεκτείνοντας τις ήδη υπάρχουσες μεθόδους. Παράλληλα, ενισχύει το υπόβαθρο για επερχόμενες ερευνητικές προσπάθειες στον τομέα των προβλέψεων και θέτει τις βάσεις για μελλοντική ερευνητική δραστηριότητα.

### ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Βραχυπρόθεσμη πρόβλεψη, φωτοβολταϊκή παραγωγή, εικόνες του ουρανού, βαθιά μάθηση, συνελικτικοί αυτοκωδικοποιητές, ταξινόμηση δεδομένων εισόδου

## ABSTRACT

The purpose of the present Diploma Thesis is the implementation of a sky image prediction model based on deep learning for short-term photovoltaic (PV) power forecasting. To this end, four model architectures were developed, based on Convolutional Auto-Encoders (CAEs), utilizing 2D and 3D convolutions, so that the model can predict a sequence of future sky images given a sequence of input images.

This effort is supported by a dataset of ground-based sky images captured by a camera over a fourteen-day period. The evaluation of the results is conducted using both qualitative (visual assessment) and quantitative methods (evaluation metrics). The development of the code for the models' building in this Diploma Thesis was carried out using the Python programming language.

This work introduces an innovative approach by incorporating a classification technique that categorizes sky image data into different sky conditions (sunny, cloudy, and overcast), extending existing methodologies. At the same time, it strengthens the foundation for future research efforts in the forecasting field and lays the groundwork for ongoing scientific exploration.

### **KEY WORDS**

Short-term forecasting, photovoltaic power production, sky images, deep learning, convolutional auto-encoders, input data classification

### ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε κατά το ακαδημαϊκό έτος 2024–2025 υπό την επίβλεψη του κ. Παύλου Γεωργιλάκη, καθηγητή της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Ε.Μ.Π. στον οποίο οφείλω ιδιαίτερες ευχαριστίες για την ανάθεσή της, δίνοντάς μου την ευκαιρία να ασχοληθώ με ένα τόσο ενδιαφέρον και δημιουργικό θέμα. Επίσης, θα ήθελα να ευχαριστήσω θερμά τον υποψήφιο διδάκτορα Μάρκο Κουσουνάδη Κνούσεν για την υπομονή, την πολύτιμη βοήθεια και την καθοδήγηση που μου παρείχε σε όλη τη διάρκεια εκπόνησης της εργασίας, καθώς και για τον πολύτιμο χρόνο που μου αφιέρωσε.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου, τους γονείς μου Παναγιώτη και Καλλιόπη, και τις αδερφές μου, για την αμέριστη βοήθειά τους σε όλη τη διάρκεια των σπουδών μου και την συμπαράστασή τους.

## ΠΕΡΙΕΧΟΜΕΝΑ

### ΚΕΦΑΛΑΙΟ 1<sup>0</sup> ΕΙΣΑΓΩΓΗ

1.1	Περιοχή έρευνας	1
1.2	Ανασκόπηση βιβλιογραφίας	2
1.3	Αντικείμενο της εργασίας	3
1.4	Δομή της εργασίας	4

## ΚΕΦΑΛΑΙΟ 2<sup>0</sup> ΒΑΣΙΚΟ ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1	Εισαγο	ωγή	5
2.2	Πρόβλ	εψη φωτοβολταϊκής παραγωγής	6
	2.2.1	Βραχυπρόθεσμη πρόβλεψη φωτοβολταϊκής παραγωγής	7
	2.2.2	Πρόβλεψη εικόνων του ουρανού	9
2.3	Βαθιά	μάθηση	10
2.4	Τεχνη	τά νευρωνικά δίκτυα	11
	2.4.1	Τύποι νευρωνικών δικτύων	12
	2.4.2	Σύντομη ιστορία νευρωνικών δικτύων	13
	2.4.3	Συνελικτικά νευρωνικά δίκτυα	13
		2.4.3.1 Συνέλιξη τετραγωνικών πινάκων	14
		2.4.3.2 Έγχρωμη εικόνα (RGB)	16
		2.4.3.3 Βήμα μετατόπισης (strides) και τεχνική συμπλήρωσης με	
		μηδενικά (padding)	17
		2.4.3.4 Αλγόριθμοι βελτιστοποίησης	18
		2.4.3.5 Συνάρτηση ενεργοποίησης (Activation function)	18
	2.4.4	Συνελικτικοί αυτοκωδικοποιητές	19
		2.4.4.1 Αυτοκωδικοποιητές 2D και 3D συνελίξεων	22
		2.4.4.2 Εκπαίδευση μοντέλων συνελικτικών αυτοκωδικοποιητών -	
		Προβλήματα υπερπροσαρμογής και υποπροσαρμογής	23
		2.4.4.3 Κανονικοποίηση κατά παρτίδες (Batch Normalization)	24

### ΚΕΦΑΛΑΙΟ 3<sup>0</sup> ΑΝΑΛΥΣΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΕΙΚΟΝΩΝ ΤΟΥ ΟΥΡΑΝΟΥ

3.1	Εισαγα	ογή	27
3.2	Ανάλυ	ση δεδομένων	27
3.3	Επεξερ	γασία αρχικών δεδομένων	29
	3.3.1	Απεικόνιση εικόνων στην κλίμακα του γκρι (grayscaling)	29
	3.3.2	Μείωση ανάλυσης εικόνων	30
3.4	Οργάν	ωση των δεδομένων	31
	3.4.1	Ορισμός του συνόλου των εικόνων σε ακολουθίες	31

3.4.2	Διαμοιρασμός των ακολουθιών σε σύνολα εκπαίδευσης, επικύρωσης και	
	αξιολόγησης	33

## ΚΕΦΑΛΑΙΟ 4<sup>0</sup> ΜΕΘΟΔΟΛΟΓΙΑ-ΥΛΟΠΟΙΗΣΗ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΤΟΥ ΜΟΝΤΕΛΟΥ

Περιγραφή μοντέλου πρόβλεψης ακολουθίας εικόνων					35
Περιγραφή	εξεταζόμενων	αρχιτεκτονικών	μοντέλων	συνελικτικών	
αυτοκωδικ	οποιητών				37
4.2.1 Ap	οχιτεκτονικές αυτοκω	δικοποιητών 2D και	3D συνελίξεω	w	37
4.2	2.1.1 Τριών επιπέδων	,			37
4.2	2.1.2 Πέντε επιπέδων	·			38
Περιβάλλον	ν υλοποίησης του μον	ντέλου			39
Εκτέλεση τ	ου μοντέλου				39
	Περιγραφή Περιγραφή αυτοκωδικ 4.2.1 Αρ 4.2 4.2 Περιβάλλον Εκτέλεση τ	Περιγραφή μοντέλου πρόβλεψη Περιγραφή εξεταζόμενων αυτοκωδικοποιητών 4.2.1 Αρχιτεκτονικές αυτοκω 4.2.1.1 Τριών επιπέδων 4.2.1.2 Πέντε επιπέδων Περιβάλλον υλοποίησης του μον	Περιγραφή μοντέλου πρόβλεψης ακολουθίας εικόνα Περιγραφή εξεταζόμενων αρχιτεκτονικών αυτοκωδικοποιητών	Περιγραφή μοντέλου πρόβλεψης ακολουθίας εικόνων Περιγραφή εξεταζόμενων αρχιτεκτονικών μοντέλων αυτοκωδικοποιητών 4.2.1 Αρχιτεκτονικές αυτοκωδικοποιητών 2D και 3D συνελίξεω 4.2.1.1 Τριών επιπέδων 4.2.1.2 Πέντε επιπέδων Περιβάλλον υλοποίησης του μοντέλου Εκτέλεση του μοντέλου	Περιγραφή μοντέλου πρόβλεψης ακολουθίας εικόνων Περιγραφή εξεταζόμενων αρχιτεκτονικών μοντέλων συνελικτικών αυτοκωδικοποιητών 4.2.1 Αρχιτεκτονικές αυτοκωδικοποιητών 2D και 3D συνελίξεων 4.2.1.1 Τριών επιπέδων 4.2.1.2 Πέντε επιπέδων Περιβάλλον υλοποίησης του μοντέλου Εκτέλεση του μοντέλου.

## ΚΕΦΑΛΑΙΟ 5<sup>0</sup> ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1	Μετρη	τικά αξιολόγησης	41				
	5.1.1	MSE	41				
	5.1.2	SSIM	42				
5.2	Μέθοδ	Μέθοδος πρόβλεψης Persistence.					
5.3	Πρόβλ	εψη ακολουθίας εικόνων του ουρανού χωρίς ταξινόμηση εισόδου	43				
	5.3.1	Ποιοτική αξιολόγηση αποτελεσμάτων	43				
	5.3.2	Ποσοτική αξιολόγηση αποτελεσμάτων	45				
5.4	Μελέτ	η περίπτωσης σε σετ δεδομένων χαμηλής μεταβλητότητας	46				
5.5	Πρόβλεψη ακολουθίας εικόνων του ουρανού με ταξινόμηση εισόδου						
	5.5.1	Μέθοδος ταξινόμησης εικόνων του ουρανού	48				
	5.5.2	Ταξινόμηση των δεδομένων σε δύο κλάσεις	49				
		5.5.2.1 Ποιοτική αξιολόγηση αποτελεσμάτων	50				
		5.5.2.2 Ποσοτική αξιολόγηση αποτελεσμάτων	50				
	5.5.3	Ταξινόμηση των δεδομένων σε τρεις κλάσεις	51				
		5.5.3.1 Ποιοτική αξιολόγηση αποτελεσμάτων	52				
		5.5.3.2 Ποσοτική αξιολόγηση αποτελεσμάτων	52				
5.6	Συγκεν	γτρωτικά αποτελέσματα	53				
5.7	Χρόνοι εκπαίδευσης						

## ΚΕΦΑΛΑΙΟ 6<sup>0</sup> ΣΥΜΠΕΡΑΣΜΑΤΑ

UDA	A A 10 70	
6.2	Πιθανές μελλοντικές επεκτάσεις της εργασίας	58
6.1	Σύνοψη και συμπεράσματα της εργασίας	57

KEΦΑΛΑΙΟ $7^{\circ}$
ΒΙΒΛΙΟΓΡΑΦΙΑ

## ΕΙΣΑΓΩΓΗ

### 1.1 ΠΕΡΙΟΧΗ ΕΡΕΥΝΑΣ

Τα τελευταία χρόνια παρατηρείται δυναμική διείσδυση των ανανεώσιμων πηγών ενέργειας (AΠΕ) στα συστήματα ηλεκτρικής ενέργειας, επειδή οι AΠΕ αποτελούν ανεξάντλητες πηγές ενέργειας και είναι φιλικές προς το περιβάλλον. Ένα μεγάλο μέρος των AΠΕ συνιστά η ΦΒ παραγωγή, καθώς με την ραγδαία αύξηση της τεχνολογίας, το κόστος ανά μονάδα εγκατεστημένης ΦΒ ισχύος είναι αρκετά χαμηλό, οι συλλέκτες ηλιακής ενέργειας παρουσιάζουν ολοένα και πιο ικανοποιητικές τιμές μέσης απόδοσης και υπάρχει αρκετή ευελιξία στην τοποθέτηση [1], [2]. Ωστόσο, η ΦΒ παραγωγή χαρακτηρίζεται από έντονη μεταβλητότητα, που οφείλεται κυρίως στην εγγενή στοχαστικότητα που παρουσιάζει η κατανομή των νεφών. Η μεταβλητότητα αυτή καθιστά επιτακτική την ανάγκη εύρεσης μεθόδων πρόβλεψης ΦΒ παραγωγής, με απώτερο σκοπό την ενίσχυση της ευστάθειας και της ομαλότητας στην λειτουργία του ηλεκτρικού δικτύου. Οφέλη από την πρόβλεψη της παραγωγής μπορούν να καρπωθούν τόσο οι διαχειριστές του συστήματος ηλεκτρικής ενέργειας (ΣΗΕ) και οι έμποροι ηλεκτρικής ενέργειας, όσο και οι ίδιοι οι καταναλωτές που ταυτόχρονα είναι και παραγωγό (prosumers) [3].

Ανάλογα με τον χρονικό ορίζοντα, οι προβλέψεις διακρίνονται σε μακροπρόθεσμες, μεσοπρόθεσμες και βραχυπρόθεσμες. Η παρούσα διπλωματική εργασία πραγματεύεται την βραχυπρόθεσμη πρόβλεψη, καθώς είναι εξαιρετικά σημαντική για την ασφαλή και οικονομικά αποδοτική λειτουργία του ΣΗΕ και διαδραματίζει καίριο ρόλο σε εφαρμογές όπως ο έλεγχος του ρυθμού μεταβολής ισχύος, η βέλτιστη διαχείριση των μέσων αποθήκευσης και η απόκριση ζήτησης σε πραγματικό χρόνο [3].

Ένας πολύ διαδεδομένος τρόπος πρόβλεψης είναι με χρήση επίγειων εικόνων του ουρανού. Οι εικόνες σε σχέση με τα αριθμητικά δεδομένα δίνουν πολύ καλύτερη πληροφορία για την ύπαρξη και κίνηση των νεφώσεων. Η πρόβλεψη παραγωγής με χρήση εικόνων του ουρανού, συνοψίζεται στα ακόλουθα βήματα. Αρχικά, πραγματοποιείται η λήψη των φωτογραφιών (εικόνων του ουρανού) με σκοπό να γίνει η μοντελοποίηση της κίνησης των νεφώσεων και η πρόβλεψη των μελλοντικών εικόνων του ουρανού. Στην συνέχεια, πραγματοποιείται μετατροπή κάθε μελλοντικών εικόνας σε ηλιακή ακτινοβολία και τελικά γίνεται η πρόβλεψη της ΦΒ παραγωγής [4]. Μέχρι τώρα στην βραχυπρόθεσμη πρόβλεψη παραδοσιακά εφαρμοζόταν η τεχνολογία ψηφιακής επεξεργασίας εικόνας κατά την οποία χρησιμοποιούνταν δύο διαδοχικές εικόνες του ουρανού για εξαγωγή ενός διανύσματος κίνησης νεφών, από το οποίο θα προέκυπτε η μελλοντική εικόνα. Η μέθοδος αυτή, ωστόσο, απαιτεί δεδομένα υψηλής χρονικής ανάλυσης και υποθέτει γραμμικότητα στην συμπεριφορά των νεφών, εισάγοντας σφάλμα από την αρχή [5]. Έτσι, τα τελευταία χρόνια γίνεται μια διερεύνηση με μεθόδους τεχνητής νοημοσύνης οι οποίες είναι αρκετά ικανές στην διαχείριση μεγάλου όγκου πληροφοριών και την εκπαίδευση για εκμάθηση μη γραμμικών σχέσεων μεταξύ εισόδου και εξόδου [6].

### 1.2 ΑΝΑΣΚΟΠΗΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Στο πλαίσιο εκπόνησης της διπλωματικής εργασίας, αρχικά πραγματοποιήθηκε έρευνα στην βιβλιογραφία σε θέματα που αφορούν στην βραχυπρόθεσμη πρόβλεψη της ΦΒ παραγωγής. Παρακάτω παρουσιάζονται ενδεικτικά ορισμένες τέτοιες μελέτες.

Στην εργασία των Viv Bone, John Pidgeon, Michael Kearney, Ananthanarayanan Veeraragavan [7], οι συγγραφείς ανέπτυξαν ένα σύστημα πρόβλεψης εντός ώρας, της άμεσης κανονικής ηλιακής ακτινοβολίας (DNI) που βασίζεται σε μοντέλα καθαρού ουρανού και παρακολούθηση της κίνησης των νεφών μέσω εικόνων από κάμερα ουρανού. Για το μοντέλο πρόβλεψης νεφοκάλυψης, παρουσιάζεται μία βελτιωμένη μέθοδος "sector-ladder" για τη βελτίωση της πρόβλεψης της κίνησης των νεφών. Το σύστημα αυτό μπορεί να προβλέπει με ακρίβεια τις διακυμάνσεις στην ηλιακή ακτινοβολία και να προειδοποιεί έγκαιρα για πιθανή διαλείπουσα παραγωγή, καθιστώντας το ιδιαίτερα χρήσιμο για τη βελτιστοποίηση της λειτουργίας υβριδικών CSP-θερμικών μονάδων χωρίς την ανάγκη δαπανηρών συστημάτων αποθήκευσης.

Στην εργασία των Zhao Zhen, Shuaijie Pang, Fei Wang, Kangping Li, Zhigang Li, Hui Ren, Miadreza Shafie-khah and João P. S. Catalão [8], οι συγγραφείς πρότειναν την μέθοδος PCPOW (Pattern Classification and PSO Optimal Weights) για τον υπολογισμό της ταχύτητας κίνησης των νεφών με σκοπό τη βελτίωση της πρόβλεψης της ισχύος φωτοβολταϊκών συστημάτων σε χρονική κλίμακα λεπτών. Η μέθοδος αξιοποιεί ταξινόμηση νεφών μέσω k-means clustering και χαρακτηριστικά υφής βασισμένα στον Πίνακα Συνύπαρξης Επιπέδων του Γκρι (Gray-Level Co-occurrence Matrix – GLCM), ενώ στη συνέχεια εφαρμόζει ένα συνδυαστικό υπολογιστικό μοντέλο με βέλτιστα βάρη που έχουν προκύψει από αλγόριθμο PSO για την εκτίμηση της ταχύτητας μετακίνησης των νεφών.

Στην εργασία των Yuwei Fu, Hua Chai, Zhao Zhen, Fei Wang, Xunjian Xu, Kangping Li, Miadreza Shafie-khah, Payman Dehghanian, João P. S. Catalão [6] οι συγγραφείς πρότειναν την χρήση συνελικτικών αποκωδικοποιητών (CAE) για την πρόβλεψη εικόνων ουρανού με σκοπό τη βελτίωση της πρόβλεψης της παραγωγής φωτοβολταϊκής ενέργειας. Η μέθοδος αυτή αποδεικνύεται ανώτερη συγκριτικά με τις παραδοσιακές μεθόδους, όπως τις PIV και FPCT.

Στην εργασία των Jingxuan Liu, Haixiang Zang, Lilin Cheng, Tao Ding, Zhinong Wei, και Guoqiang Sun [9], οι συγγραφείς αξιοποίησαν πολυτροπικά δεδομένα —ιστορικά GHI και εικόνες του ουρανού— μέσω τεχνικών βαθιάς μάθησης. Συγκεκριμένα, χρησιμοποίησαν το δίκτυο Informer για την επεξεργασία των χρονοσειρών, ενώ οι εικόνες ουρανού μετατράπηκαν σε χάρτες οπτικής ροής και αναλύθηκαν με Vision Transformer. Η ενσωμάτωση των δύο πηγών πληροφορίας επιτεύχθηκε με μηχανισμό cross-modality attention, και η πρόβλεψη υλοποιήθηκε με γενετικό αποκωδικοποιητή.

Στην εργασία των Yuhao Niea, Eric Zelikmanc, Andea Scotta, Quentin Palettad, και Adam Brandt [10] οι συγγραφείς παρουσίασαν το SkyGPT, ένα στοχαστικό μοντέλο πρόβλεψης βίντεο, το οποίο εισήγαγαν ερευνητές αξιοποιώντας παραγωγική τεχνητή νοημοσύνη για τη δημιουργία πιθανών μελλοντικών εικόνων του ουρανού βάσει ιστορικών δεδομένων. Το μοντέλο επέδειξε υψηλή ικανότητα στην πρόβλεψη της δυναμικής των νεφών, γεγονός που συνέβαλε σημαντικά στη βελτίωση της ακρίβειας πρόβλεψης της φωτοβολταϊκής παραγωγής 15 λεπτών μπροστά. Σε συνδυασμό με ένα U-Net-based μοντέλο πρόβλεψης ισχύος, το SkyGPT ξεπέρασε παραδοσιακές μεθόδους, όπως τα κλασικά συνελικτικά νευρωνικά δίκτυα και τα έξυπνα μοντέλα επιμονής (smart persistence models).

### **1.3** ANTIKEIMENO TH $\Sigma$ EPFA $\Sigma$ IA $\Sigma$

Το αντικείμενο της παρούσας διπλωματικής εργασίας είναι η υλοποίηση μοντέλου πρόβλεψης εικόνων του ουρανού με χρήση βαθιάς μάθησης για την βραχυπρόθεσμη πρόβλεψη της φωτοβολταϊκής (ΦΒ) παραγωγής. Στόχος του προβλήματος είναι για μια ακολουθία εισόδου εικόνων του ουρανού, το μοντέλο να μπορεί να προβλέψει μια ακολουθία εξόδου (μελλοντικές εικόνες του ουρανού). Η αντικειμενική συνάρτηση στο πρόβλημα πρόβλεψης εικόνων είναι η ελαχιστοποίηση του σφάλματος μεταξύ των πραγματικών και των προβλεπόμενων εικόνων.

Η εργασία χρησιμοποίησε επίγειες εικόνες του ουρανού από ΦΒ σύστημα εγκατεστημένης ισχύος 1.2 kW, στην Βοιωτία μέσω κάμερας παρακολούθησης ASI-16.

Στην συνέχεια, για την υλοποίηση του μοντέλου πρόβλεψης πραγματοποιήθηκε κατασκευή τεσσάρων αρχιτεκτονικών που βασίζονται σε συνελικτικούς αυτοκωδικοποιητές (Convolutional Auto-Encoders – CAEs), με χρήση 2D και 3D συνελίξεων. Η επιλογή βαθιάς μάθησης γίνεται για να αντιμετωπιστούν τα προβλήματα μη γραμμικότητας των παραδοσιακών μεθόδων (όπως αναφέρεται στην Ενότητα 1.1). Ταυτόχρονα, η επιλογή των CAE έναντι της παραγωγικής τεχνητής νοημοσύνης (generative artificial intelligence) γίνεται για λόγους αποφυγής υψηλής πολυπλοκότητας και περιορισμούς στο hardware.

Στο πλαίσιο των δοκιμών πραγματοποιήθηκε συνδυασμός μοντέλων πρόβλεψης με ταξινόμηση εισόδου για τη μείωση της μεταβλητότητας του συνόλου εκπαίδευσης και την αύξηση της ικανότητας γενίκευσης, τεχνική που δεν συναντήθηκε κατά την βιβλιογραφική έρευνα.

Τα αποτελέσματα αναδεικνύουν την σημασία της χρήσης της τεχνικής της ταξινόμησης εισόδου στην πρόβλεψη ακολουθίας εικόνων του ουρανού, μειώνοντας το μέσο τετραγωνικό σφάλμα έως και 53.2%, χωρίς να επιβαρύνει το υπολογιστικό κόστος της συνολικής διαδικασίας. Κατά αυτόν τον τρόπο αναδεικνύεται πως με τον συνδυασμό μοντέλων βαθιάς μάθησης και τεχνικών ανάλυσης και προεπεξεργασίας δεδομένων μπορεί να επιτευχθεί ικανοποιητική απόδοση χωρίς την αύξηση των υπολογιστικών απαιτήσεων.

### 1.4 ΔΟΜΗ ΤΗΣ ΕΡΓΑΣΙΑΣ

Η παρούσα εργασία οργανώνεται σε επτά κεφάλαια:

- Στο Κεφάλαιο 2 επισημαίνεται η ανάγκη για πρόβλεψη της φωτοβολταϊκής παραγωγής και παρουσιάζεται το βασικό θεωρητικό υπόβαθρο σχετικά με τα συνελικτικά νευρωνικά δίκτυα και τους συνελικτικούς αυτοκωδικοποιητές.
- Στο Κεφάλαιο 3 αναλύεται η διαδικασία συλλογής των δεδομένων και στην συνέχεια η επεξεργασία και η οργάνωσή τους για την είσοδό τους στο μοντέλο.

- Στο Κεφάλαιο 4 αναλύεται η μεθοδολογία και η υλοποίηση των τεσσάρων αρχιτεκτονικών του μοντέλου πρόβλεψης ακολουθίας εικόνων του ουρανού.
- Στο Κεφάλαιο 5 παρουσιάζονται και σχολιάζονται τα αποτελέσματα του μοντέλου πρόβλεψης ακολουθίας εικόνων του ουρανού.
- Στο Κεφάλαιο 6 γίνεται μια σύνοψη της εργασίας, καταγράφονται τα συμπεράσματά της και προτείνονται τρόποι βελτίωσης και επέκτασής της.
- Στο Κεφάλαιο 7 παρουσιάζεται η βιβλιογραφία που χρησιμοποιήθηκε σε όλη την έκταση της διπλωματικής εργασίας.

## КЕФАЛАЮ 2

## ΒΑΣΙΚΟ ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

### 2.1 ΕΙΣΑΓΩΓΗ

Τις τελευταίες δεκαετίες, οι ανανεώσιμες πηγές ενέργειας (ΑΠΕ), έχουν διεισδύσει δυναμικά στο Σύστημα Ηλεκτρικής Ενέργειας (ΣΗΕ), καθώς παρέχουν πληθώρα προτερημάτων στο ενεργειακό τοπίο. Όπως είναι γνωστό, οι ΑΠΕ είναι ανεξάντλητες πηγές ενέργειας, συμβάλλουν στην μείωση της εξάρτησης από συμβατικές μονάδες παραγωγής και είναι φιλικές προς το περιβάλλον. Συγκεκριμένα, η ηλιακή ενέργεια είναι η πιο άφθονη από όλες τις ενεργειακές πηγές, πρακτικά ανεξάντλητη και διαθέσιμη για χρήση σε όλες τις περιοχές του κόσμου. Η ποσότητα της ηλιακής ενέργειας που προσπίπτει στη Γη είναι περίπου 10.000 φορές μεγαλύτερη από την ενέργεια που καταναλώνει η ανθρωπότητα [1]. Μία βασική εφαρμογή της ηλιακής ενέργειας είναι τα φωτοβολταϊκά (Φ/Β) συστήματα παραγωγής ενέργειας, τα οποία χρησιμοποιούν Φ/Β κύτταρα για να μετατρέψουν την ηλιακή ενέργεια σε ηλεκτρική.

Τα Φ/Β συστήματα είναι αρκετά φερέγγυα όσον αφορά στην ενσωμάτωση στο ΣΗΕ. Αρχικά, η Φ/Β παραγωγή παρουσιάζει ημερήσια περιοδικότητα, γεγονός που προσδίδει στην αξιοπιστία της, λόγω της δυνατότητας ασφαλέστερης πρόγνωσης σε σχέση με άλλες τεχνολογίες, όπως για παράδειγμα οι ανεμογεννήτριες. Επιπλέον, τα Φ/Β συστήματα παρουσιάζουν αρκετή ευελιξία στην τοποθέτηση, καθώς μπορούν να εγκατασταθούν σε μικρή κλίμακα και στο αστικό περιβάλλον. Ταυτόχρονα, το κόστος των Φ/Β τεχνολογιών μειώνεται ραγδαία με την πάροδο του χρόνου (26% μείωση κόστους ανά μονάδα ονομαστικής ισχύος τη τελευταία δεκαετία και 90% από το 2000) [2]. Αξιοσημείωτο είναι, επίσης, ότι η μέση απόδοση, παρόλο που είναι ακόμα σχετικά χαμηλή, την τελευταία δεκαετία έχει αυξηθεί με εισαγωγή υλικών όπως ο περοβσκίτης [11] και τεχνολογίες όπως τα Φ/Β αλληλοδιάδοχης διάταξης (tandem photovoltaics) [12].

Ορισμένες μορφές ΑΠΕ, παρά τα αξιοσημείωτα προτερήματα που προσφέρουν στην παραγωγή ισχύος, χαρακτηρίζονται από έντονη μεταβλητότητα, γεγονός που πλήττει την αξιοπιστία τους. Συγκεκριμένα, στο πλαίσιο της ηλιακής ενέργειας, η σημαντικότερη πηγή μεταβλητότητας οφείλεται στην κίνηση των νεφών, αφού η σχετική θέση του ηλίου είναι αρκετά εύκολα υπολογίσιμη, λόγω της περιοδικότητάς της. Συνεπώς, προς την κατεύθυνση της αύξησης της ενσωμάτωσης της Φ/Β παραγωγής στο ΣΗΕ και την ταυτόχρονη διατήρηση της αξιοπιστίας του ΣΗΕ, το ενδιαφέρον έχει στραφεί στην πρόβλεψη Φ/Β παραγωγής ισχύος.

### 2.2 ΠΡΟΒΛΕΨΗ ΦΩΤΟΒΟΛΤΑΪΚΗΣ ΠΑΡΑΓΩΓΗΣ

Η μεταβλητότητα της Φ/Β παραγωγής καθιστά επιτακτική την ανάγκη εύρεσης μεθόδων πρόβλεψης της παραγωγής. Οι προβλέψεις Φ/Β παραγωγής χρησιμοποιούνται σε πολυάριθμα πλαίσια, με ποικίλες χωροχρονικές κλίμακες, και η ακρίβειά τους μπορεί να επηρεάσει σημαντικά την αξιοπιστία του ΣΗΕ. Παρακάτω παρουσιάζονται κάποιες χρήσεις και οφέλη της πρόβλεψης Φ/Β παραγωγής [3]:

- Εξισορρόπηση παραγωγής-ζήτησης και άρα διατήρηση της συχνότητας των ΣΗΕ
- Εξομάλυνση έντονων διακυμάνσεων παραγωγής (πχ, μέσω τοποθέτησης και βέλτιστης χρήσης ηλεκτρονικών ισχύος) και συνεπώς την βελτίωση της ποιότητας ισχύος
- Προγραμματισμός συστημάτων αποθήκευσης και εφεδρειών και άρα εξασφάλιση αποδοτικότερης λειτουργίας, μείωση περίσσειας ενέργειας, και τελικά τον περιορισμό των εκπομπών CO<sub>2</sub>.

Η πρόβλεψη της Φ/Β παραγωγής παρέχει οφέλη σε όλους όσους εμπλέκονται στην αγορά ηλεκτρικής ενέργειας. Συγκεκριμένα:

- Στους διαχειριστές του ΣΗΕ (αποσυμφόρηση, εξαγωγή ευελιξίας)
- Στους παραγωγούς (συμμετοχή στην αγορά ηλεκτρικής ενέργειας, στην αγορά εξισορρόπησης, στην αποφυγή κυρώσεων)
- Στους ίδιους τους καταναλωτές που ταυτόχρονα είναι και παραγωγοί (prosumers) (βέλτιστη διαχείριση οικιακού φορτίου)

Ανάλογα με την εφαρμογή, τον χρονικό ορίζοντα πρόβλεψης και τις απαιτήσεις χωροχρονικής ανάλυσης, χρησιμοποιούνται διαφορετικά δεδομένα εισόδου και μέθοδοι πρόβλεψης [3]:

- Μοντέλα χρονοσειρών βάσει τοπικών μετρήσεων: Απαιτούν επιτόπιες παρατηρήσεις της ηλιακής ακτινοβολίας και της παραγόμενης ισχύος από Φ/Β συστήματα, καθώς και ενδεχομένως άλλων μετεωρολογικών μεταβλητών, οι οποίες υπόκεινται επεξεργασία με στατιστικές μεθόδους ή αλγόριθμους τεχνητής νοημοσύνης (Artificial Intelligence – AI) και μηχανικής μάθησης (Machine Learning – ML), όπως τα τεχνητά νευρωνικά δίκτυα. Παρέχουν συνήθως αξιόπιστες προβλέψεις έως και λίγες ώρες μπροστά υπό σχετικά σταθερές συνθήκες ουρανού· ωστόσο, σπάνια αποδίδουν καλά σε μεταβλητές συνθήκες λόγω της χαοτικής συμπεριφοράς των νεφών και των περιορισμένων πληροφοριών από τις επιτόπιες σημειακές παρατηρήσεις.
- 2. Προβλέψεις με βάση εικόνες από κάμερες ουρανού (All-Sky Imager ASI): Οι προβλέψεις τέτοιων δεδομένων βασίζονται συνήθως σε τεχνικές διανυσμάτων κίνησης νεφών (Cloud Motion Vector CMV) για εξαγωγή της μελλοντικής θέσης των νεφών, όπως και μεθόδους μηχανικής μάθησης. Αξιοποιώντας πληροφορίες για την τοπική κατανομή νεφών από μία ή περισσότερες επιτόπιες ASI, η ακρίβεια πρόγνωσης μπορεί να βελτιωθεί σε σχέση με τα μοντέλα χρονοσειρών βασισμένα

μόνο σε τοπικές μετρήσεις. Αυτές οι πληροφορίες είναι καθοριστικές για προβλέψεις ηλιακής ακτινοβολίας με χρονική ανάλυση δευτερολέπτων έως λεπτών και χωρική ανάλυση 10–100 μέτρων σε μια περιοχή λίγων τετραγωνικών χιλιομέτρων γύρω από την ASI. Ο συνήθης χρονικός ορίζοντας πρόγνωσης για μία κάμερα είναι 10–20 λεπτά (εάν τοποθετηθούν παραπάνω κάμερες, ο χρονικός ορίζοντας μπορεί να αυξηθεί), ανάλογα με το ύψος και την ταχύτητα των νεφών.

- 3. Δορυφορικά δεδομένα γεωστατικών δορυφόρων: Οι προβλέψεις που προκύπτουν από τέτοια δεδομένα βασίζονται συνήθως σε τεχνικές CMV. Οι προβλέψεις έχουν ορίζοντα αρκετές ώρες μπροστά, αξιοποιώντας ευρείας κλίμακας παρατηρήσεις νεφικών πεδίων. Η τυπική χωρική ανάλυση είναι 1–5 km<sup>2</sup> για τους σύγχρονους γεωστατικούς δορυφόρους, με ενημέρωση κάθε 10–30 λεπτά, και χρονικό ορίζοντα 4–6 ώρες.
- 4. Αριθμητικά μοντέλα πρόγνωσης καιρού (Numerical Weather Prediction NWP): Αποτελεί την κύρια προσέγγιση για ορίζοντες πρόγνωσης μεγαλύτερους από μερικές ώρες και έως αρκετές ημέρες ή εβδομάδες. Τα μοντέλα NWP προσομοιώνουν τη δυναμική της ατμόσφαιρας, όπως τη δημιουργία, μεταφορά και διάλυση νεφών, λύνοντας και παραμετροποιώντας τα αντίστοιχα φυσικά συστήματα εξισώσεων. Τα παγκόσμια μοντέλα έχουν χωρική ανάλυση ~0.1°-0.5° και χρονική ανάλυση 1-3 ώρες· τα περιφερειακά μοντέλα έχουν χωρική ανάλυση λίγων χιλιομέτρων και χρονική ανάλυση ανά ώρα.
- 5. Μετα-επεξεργασία και συνδυασμός μοντέλων (model blending): Συνδυασμός μεθόδων και δεδομένων για δημιουργία υβριδικών μοντέλων. Τα υβριδικά μοντέλα μπορούν να συνδυάσουν τα πλεονεκτήματα των επιμέρους μεθόδων για την βελτίωση των προβλέψεων ή την διεύρυνση του πεδίου χρήσης, συνήθως σε κόστος της υπολογιστικής πολυπλοκότητας.

Η μετατροπή της προβλεπόμενης ακτινοβολίας σε προβλεπόμενη ισχύ απαιτεί μοντελοποίηση του Φ/Β συστήματος είτε με παραμετρικά μοντέλα είτε με μεθόδους μηχανικής μάθησης. Οι προσεγγίσεις διακρίνονται σε:

- Φυσικά και εμπειρικά μοντέλα, βασισμένα σε φυσικούς νόμους ή προσεγγίσεις (πχ μοντέλα καθαρού ουρανού, CMV)
- Στατιστικά μοντέλα και μοντέλα μηχανικής μάθησης, που μαθαίνουν σχέσεις εισόδου-εξόδου από ιστορικά δεδομένα [3].

### 2.2.1 Βραχυπρόθεσμη πρόβλεψη φωτοβολταϊκής παραγωγής

Οι προβλέψεις Φ/Β παραγωγής μπορούν να κατηγοριοποιηθούν ανάλογα με τον χρονικό ορίζοντα πρόβλεψης (πχ μακροπρόθεσμες, μεσοπρόθεσμες, βραχυπρόθεσμες). Η βραχυπρόθεσμη πρόβλεψη, παρόλο που διαθέτει τον μικρότερο χρονικό ορίζοντα (πρόβλεψη μερικά λεπτά μπροστά), είναι εξαιρετικά σημαντική για την ασφαλή και οικονομικά αποδοτική λειτουργία του ΣΗΕ. Ο ρόλος της είναι καίριος σε εφαρμογές όπως:

- ο έλεγχος του ρυθμού μεταβολής ισχύος (ramp rate control)
- η βελτιστοποίηση στην διαχείριση των μέσων αποθήκευσης
- η απόκριση ζήτησης σε πραγματικό χρόνο

Ωστόσο, η παροχή ακριβούς βραχυπρόθεσμης πρόβλεψης είναι δύσκολη. Η ακρίβεια της πρόβλεψης επηρεάζεται σημαντικά από τη δυναμική μεταβολή της νεφοκάλυψης. Η επιφανειακή ηλιακή ακτινοβολία και η παραγόμενη ηλιακή ενέργεια μπορεί να αυξηθούν ή να μειωθούν ραγδαία μέσα σε λίγα λεπτά, ακολουθώντας τις μεταβολές των νεφών στον ουρανό.

Όπως δηλώνει η σχέση (2.1):

 $PV_{OUTPUT} = f(H\lambdaιακή ακτινοβολία (Κατανομή των νεφών)) + ε$  (2.1)

η Φ/Β παραγωγή είναι κατά βάση συνάρτηση της προσπίπτουσας ηλιακής ακτινοβολίας στην επιφάνεια των Φ/Β, η οποία εξαρτάται πρωτίστως από την κατανομή των νεφών. Προφανώς η Φ/Β παραγωγή εξαρτάται και από άλλες παραμέτρους (τυχαία μεταβλητή ε), όπως τη γωνία πρόσπτωσης της ακτινοβολίας – η οποία ωστόσο μπορεί εύκολα να μοντελοποιηθεί γιατί οφείλεται σε περιοδικά φαινόμενα με ελάχιστη στοχαστικότητα – αλλά και παράγοντες όπως η κατανομή της θερμότητας πάνω στο πάνελ – που είναι πολύ πιο δύσκολο να μοντελοποιηθεί, αλλά δεν επηρεάζει σημαντικά την παραγωγή. Επομένως, ο πιο σημαντικός παράγοντας που λαμβάνεται υπόψιν στην βραχυπρόθεσμη πρόβλεψη Φ/Β παραγωγής είναι τα σύννεφα, γιατί:

- επηρεάζουν άμεσα την έξοδο του Φ/Β και
- 2) είναι διαδικασία αρκετά δύσκολη στην μοντελοποίηση

Παρόλο που ο χρονικός ορίζοντας πρόβλεψης είναι μικρός και άρα η αβεβαιότητα μικρή, οι εφαρμογές που βασίζονται στην βραχυπρόθεσμη πρόβλεψη απαιτούν συνήθως υψηλή χρονική ανάλυση. Ωστόσο, όσο μεγαλύτερη η χρονική ανάλυση, τόσο πιο πολύπλοκη (λιγότερο εξομαλυμένη) είναι η καμπύλη που χρειάζεται να προβλεφθεί. Μέσα σε λίγα δευτερόλεπτα η πτώση στην παραγωγή μπορεί να είναι της τάξης του 70% [13].

Μέχρι στιγμής έχουν προταθεί πολυάριθμες μέθοδοι βραχυπρόθεσμης πρόβλεψης, οι οποίες μπορούν να ταξινομηθούν με βάση τα δεδομένα εισόδου και εξόδου τους. Από την πλευρά της εξόδου του μοντέλου, υπάρχουν δύο ομάδες μεθόδων:

- πρόβλεψη της ηλιακής ακτινοβολίας και έπειτα μετατροπή της σε παραγόμενη Φ/Β ισχύ.
- 2) άμεση πρόβλεψη της παραγόμενης  $\Phi/B$  ισχύος.

Από την πλευρά της εισόδου, οι μέθοδοι βραχυπρόθεσμης πρόβλεψης διακρίνονται σε δύο κατηγορίες:

- μεθόδους που βασίζονται σε εικόνες, οι οποίες λαμβάνουν ως είσοδο εικόνες (συνήθως νεφών).
- μεθόδους βασισμένες σε αριθμητικά δεδομένα, οι οποίες δεν χρησιμοποιούν εικόνες ως είσοδο.

### 2.2.2 Πρόβλεψη εικόνων του ουρανού

Σε σύγκριση με τις μεθόδους που βασίζονται σε αριθμητικά δεδομένα, οι μέθοδοι που βασίζονται σε εικόνες του ουρανού έχουν ένα φυσικό πλεονέκτημα, καθώς περιέχουν καλύτερη πληροφορία σχετικά με την ύπαρξη, μορφή, και κίνηση των νεφών. Έτσι, μέσω της ανάλυσης εικόνων νεφών, η πρόβλεψη Φ/Β παραγωγής μπορεί να γίνει ευκολότερη και πιο ακριβής.

Στο πλαίσιο των βραχυπρόθεσμων προβλέψεων, εφαρμόζεται η μέθοδος που βασίζεται σε επίγειες εικόνες του ουρανού (Ground-based Sky Images – GSIs). Ο εξοπλισμός που χρησιμοποιείται είναι ένα αυτόματο σύστημα κάμερας υψηλής ανάλυσης με πεδίο θέασης 180° και ευρυγώνιο φακό. Η κάμερα τοποθετείται στο έδαφος και οι λήψεις της αρχειοθετούνται σε μονάδα αποθήκευσης με την μορφή αρχείου φωτογραφίας. Παράδειγμα λήψης επίγειας εικόνας του ουρανού παρουσιάζεται στο Σχήμα 2.1.



Σχήμα 2.1: Παράδειγμα λήψης επίγειας εικόνας του ουρανού

Η όλη διαδικασία από την λήψη των φωτογραφιών μέχρι την τελική πρόβλεψη της Φ/Β παραγωγής συνοψίζεται στα ακόλουθα βήματα [4]:

- 1. Λήψη φωτογραφιών
- 2. Μοντελοποίηση κίνησης των νεφώσεων
- 3. Πρόβλεψη μελλοντικής εικόνας
- 4. Μετατροπή μελλοντικής εικόνας σε ηλιακή ακτινοβολία
- 5. Πρόβλεψη παραγωγής

Η παρούσα διπλωματική ασχολείται με το δεύτερο και το τρίτο βήμα, δηλαδή την μοντελοποίηση κίνησης των νεφώσεων και την πρόβλεψη μελλοντικής εικόνας. Στο Σχήμα 2.2 εμφανίζονται διαγραμματικά τα παραπάνω πέντε βήματα.



## Σχήμα 2.2: Μεθοδολογία πρόβλεψης φωτοβολταϊκής παραγωγής με χρήση εικόνων του ουρανού

Ως προς την μέθοδο πρόβλεψης εικόνας, μέχρι τώρα η παραδοσιακή τεχνολογία ψηφιακής επεξεργασίας εικόνας (Digital Image Processing Technology – DIPT) χρησιμοποιήθηκε ευρέως για την βραχυπρόθεσμη πρόβλεψη επίγειων εικόνων του ουρανού. Σε αυτήν την μέθοδο χρησιμοποιούνται δύο διαδοχικές εικόνες για την εξαγωγή του CMV και εφαρμόζεται γραμμική προέκταση για τη δημιουργία της επόμενης εικόνας [5]. Οι μέθοδοι διακρίνονται σε δύο κατηγορίες: αυτές που βασίζονται σε φωτεινότητα εικόνας (πχ μετασχηματισμός χαρακτηριστικών ανεξάρτητης κλίμακας [14], οπτική ροή [15], μέτρηση ταχύτητας σωματιδίων μέσω εικόνας (Particle Image Velocimetry – PIV [16]) και αυτές που βασίζονται στον χωρικό μετασχηματισμό συχνότητας (π.χ. Fourier phase correlation theory – FPCT [17]). Ωστόσο, η DIPT παρουσιάζει κάποιες βασικές αδυναμίες. Αρχικά, οι παραδοσιακές μέθοδοι απαιτούν πολύ υψηλή χρονική ανάλυση – συχνότητα δειγματοληψίας των εικόνων για ικανοποιητική ακρίβεια και σταθερότητα. Επιπλέον, η γραμμική εξαγωγή της μελλοντικής εικόνας εισάγει σφάλματα πρόβλεψης από την αρχή, καθώς η κίνηση των νεφών δεν είναι γραμμική, αλλά και επειδή αλλάζει η μορφή των νεφών [6][18].

Για την αποφυγή αυτών των περιορισμών, η προσοχή έχει στραφεί σε μεθόδους τεχνητής νοημοσύνης και συγκεκριμένα στον τομέα της βαθιάς μάθησης. Με την ραγδαία ανάπτυξη των υπολογιστικών συστημάτων τα τελευταία χρόνια, έχει καταστεί δυνατή η χρήση μοντέλων με εκατομμύρια παραμέτρους, όπως είναι τα μοντέλα βαθιάς μάθησης. Τα μοντέλα βαθιάς μάθησης έχουν την ικανότητα να χρησιμοποιούν μεγάλο όγκο πληροφοριών, να εκπαιδεύονται και να μαθαίνουν μη γραμμικές σχέσεις εισόδου-εξόδου. Συγκεκριμένα στο πλαίσιο των εικόνων, χρησιμοποιούνται τεχνητά νευρωνικά δίκτυα (ΤΝΔ). Τα ΤΝΔ είναι αρμόδια στο να προσομοιώνουν την λειτουργία του ανθρώπινου εγκεφάλου, καθώς αποτελούνται από ένα σύστημα διασυνδεδεμένων τεχνητών νευρώνων. Τα ΤΝΔ είναι κατάλληλα για εξαγωγή χαρακτηριστικών και μοτίβων από εικόνες και άρα μπορούν να στρατευτούν στην διαδικασία της πρόβλεψης εικόνων του ουρανού [6].

### 2.3 ΒΑΘΙΑ ΜΑΘΗΣΗ

Η βαθιά μάθηση αποτελεί υποσύνολο της μηχανικής μάθησης και χρησιμοποιεί ΤΝΔ πολλαπλών υπολογιστικών επιπέδων, γνωστά ως βαθιά νευρωνικά δίκτυα, για να προσομοιώσει την πολύπλοκη ικανότητα λήψης αποφάσεων του ανθρώπινου εγκεφάλου.

Πίσω από τις περισσότερες σύγχρονες εφαρμογές τεχνητής νοημοσύνης βρίσκεται κάποια μορφή βαθιάς μάθησης.

Η βασική διαφορά μεταξύ βαθιάς μάθησης και παραδοσιακής μηχανικής μάθησης είναι η δομή της υποκείμενης αρχιτεκτονικής του ΤΝΔ. Τα παραδοσιακά μοντέλα μηχανικής μάθησης χρησιμοποιούν απλά ΤΝΔ με ένα ή δύο υπολογιστικά επίπεδα, ενώ τα μοντέλα βαθιάς μάθησης χρησιμοποιούν από τρία έως και μερικές χιλιάδες υπολογιστικά επίπεδα.

Η βαθιά μάθηση αποτελεί κλάδο της επιστήμης των δεδομένων (data science) που υποστηρίζει πολλές εφαρμογές και υπηρεσίες αυτοματοποίησης, εκτελώντας αναλυτικές και φυσικές εργασίες χωρίς ανθρώπινη παρέμβαση. Έτσι, καθίστανται δυνατά πολλά προϊόντα και υπηρεσίες της καθημερινής ζωής, όπως οι ψηφιακοί βοηθοί, τα τηλεχειριστήρια με φωνητική ενεργοποίηση, η ανίχνευση απάτης σε συναλλαγές με κάρτες, τα αυτοκινούμενα οχήματα και η γενετική τεχνητή νοημοσύνη.

Η βαθιά μάθηση απαιτεί τεράστια υπολογιστική ισχύ. Οι μονάδες γραφικών υψηλής απόδοσης (GPUs) είναι ιδανικές, καθώς μπορούν να διαχειριστούν μεγάλο όγκο υπολογισμών σε πολλαπλούς πυρήνες με άφθονη διαθέσιμη μνήμη. Αντίστοιχα, η κατανεμημένη υπολογιστική ισχύς στο υπολογιστικό νέφος (cloud computing) μπορεί επίσης να βοηθήσει. Αυτό το επίπεδο υπολογιστικής ισχύος είναι απαραίτητο για την εκπαίδευση αλγορίθμων βαθιάς μάθησης. Ωστόσο, η διαχείριση πολλών GPUs σε τοπικό επίπεδο μπορεί να επιβαρύνει σημαντικά τους εσωτερικούς πόρους και να είναι ιδιαίτερα δαπανηρή ως προς την κλιμάκωση (scaling). Όσον αφορά τις απαιτήσεις λογισμικού, οι περισσότερες εφαρμογές βαθιάς μάθησης προγραμματίζονται με ένα από τα τρία βασικά πλαίσια μάθησης: JAX, PyTorch ή TensorFlow [19].

### 2.4 ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks – ANNs) αποκαλούνται επίσης προσομοιωμένα νευρωνικά δίκτυα (Simulated Neural Networks – SNNs). Αποτελούν υποσύνολο της μηχανικής μάθησης και βρίσκονται στον πυρήνα των μοντέλων βαθιάς μάθησης (deep learning).

Ένα TNΔ είναι ένα μοντέλο μηχανικής μάθησης που λαμβάνει αποφάσεις με τρόπο παρόμοιο με τον ανθρώπινο εγκέφαλο, χρησιμοποιώντας διεργασίες που προσομοιώνουν τον τρόπο με τον οποίο συνεργάζονται οι βιολογικοί νευρώνες για να εντοπίσουν φαινόμενα, να σταθμίσουν επιλογές και να καταλήξουν σε συμπεράσματα. Κάθε TNΔ αποτελείται από στρώματα κόμβων ή τεχνητών νευρώνων: ένα στρώμα εισόδου, ένα ή περισσότερα (ενδιάμεσα) στρώματα και ένα στρώμα εξόδου. Στο Σχήμα 2.3 απεικονίζεται ένα τυπικό νευρωνικό δίκτυο.



Σχήμα 2.3: Τυπικό νευρωνικό δίκτυο

Οι κόμβοι των ΤΝΔ συνδέονται μεταξύ τους με συνάψεις – βάρη. Κατά την διαδικασία της εκπαίδευσης, οι κόμβοι προσομοιώνουν τη σχέση εισόδου-εξόδου μέσω της προσαρμογής αυτών των βαρών. Για την επιτυχή εκπαίδευση είναι απαραίτητη η ύπαρξη ενός συνόλου δεδομένων εκπαίδευσης, δηλαδή ενός συνόλου ιστορικών δεδομένων εισόδουεξόδου. Η εκπαίδευση είναι μια επαναληπτική διαδικασία που βασίζεται σε κάποιον αλγόριθμο βελτιστοποίησης με σκοπό τη βελτιστοποίηση μιας αντικειμενικής συνάρτησης. Οι πιο κοινοί αλγόριθμοι βελτιστοποίησης που χρησιμοποιούνται για την εκπαίδευση ΤΝΔ είναι οι αλγόριθμοι που βασίζονται στη μείωση της κλίσης (gradient descent optimization) [20]. Η εφαρμογή των αλγορίθμων gradient descent στα ΤΝΔ γίνεται μέσω της μεθόδου οπίσθιας διάδοσης (backpropagation) του σφάλματος [21].

Εργασίες όπως η αναγνώριση φωνής ή εικόνας, που θα απαιτούσαν ώρες με ανθρώπινη παρέμβαση, μπορούν να εκτελούνται σε λίγα λεπτά. Ένα από τα πιο γνωστά παραδείγματα TNΔ είναι ο αλγόριθμος αναζήτησης της Google [22].

#### 2.4.1 Τύποι νευρωνικών δικτύων

Τα ΤΝΔ μπορούν να ταξινομηθούν σε διάφορους τύπους, καθένας από τους οποίους χρησιμοποιείται για διαφορετικούς σκοπούς. Υπάρχουν δύο μεγάλες κατηγορίες ΤΝΔ: [22]:

- Εμπρόσθιας τροφοδότησης: Η πληροφορία ρέει προς μία κατεύθυνση μόνο, διαδοχικά από το στρώμα εισόδου έως το στρώμα εξόδου. Σε αυτήν την κατηγορία ανήκουν τα MLP, τα CNNs, καθώς και άλλα TNΔ, όπως τα Extreme Learning Machines, κτλ.
- Αναδρομικά ΤΝΔ: Εδώ η πληροφορία μπορεί να ανατροφοδοτηθεί σε νευρώνες του ίδιου στρώματος ή προηγούμενων στρωμάτων. Παραδείγματα αποτελούν τα LSTM, NARX.

### 2.4.2 Σύντομη ιστορία τεχνητών νευρωνικών δικτύων

Η ιστορία των ΤΝΔ είναι πιο μακρά από όσο φαντάζονται οι περισσότεροι. Αν και η ιδέα μιας «σκεπτόμενης μηχανής» ανάγεται ήδη στους αρχαίους Έλληνες, εδώ παρουσιάζονται ορισμένα βασικά γεγονότα που διαμόρφωσαν την πορεία και την εξέλιξη των ΤΝΔ, των οποίων η δημοτικότητα έχει αυξηθεί και μειωθεί κατά καιρούς:

- 1943: Οι Warren S. McCulloch και Walter Pitts δημοσιεύουν το άρθρο «A Logical Calculus of the Ideas Immanent in Nervous Activity». Η έρευνά τους είχε ως στόχο την κατανόηση του τρόπου με τον οποίο ο ανθρώπινος εγκέφαλος μπορεί να παράγει πολύπλοκα πρότυπα μέσω συνδεδεμένων εγκεφαλικών κυττάρων, δηλαδή νευρώνων. Μία από τις βασικές ιδέες του έργου τους ήταν η αντιστοίχιση των νευρώνων με δυαδικά κατώφλια προς τη λογική Boolean (δηλαδή δηλώσεις αληθές/ψευδές ή 1/0) [23].
- 1958: Ο Frank Rosenblatt θεωρείται ο δημιουργός του perceptron, όπως περιγράφεται στην έρευνά του «The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain». Βασισμένος στο έργο των McCulloch και Pitts, εισήγαγε την έννοια των βαρών. Με τη βοήθεια του υπολογιστή IBM 704, κατάφερε να εκπαιδεύσει ένα σύστημα ώστε να διακρίνει κάρτες με σημάδια αριστερά και δεξιά [24].
- 1974: Παρότι πολλοί ερευνητές συνέβαλαν στην ανάπτυξη της μεθόδου της οπίσθιας διάδοσης (backpropagation) του σφάλματος, ο Paul Werbos ήταν ο πρώτος που περιέγραψε τη χρήση της σε TNΔ, στην διδακτορική του διατριβή στις ΗΠΑ [25].
- 1989: Ο Yann LeCun δημοσίευσε άρθρο που εξηγούσε πώς οι περιορισμοί στην οπίσθια διάδοση του σφάλματος και η ενσωμάτωσή της στη δομή των TNΔ μπορούν να χρησιμοποιηθούν για την εκπαίδευση αλγορίθμων. Η έρευνα αυτή αξιοποίησε επιτυχώς ένα TNΔ για την αναγνώριση χειρόγραφων ψηφίων ταχυδρομικών κωδικών που παρείχε η Ταχυδρομική Υπηρεσία των ΗΠΑ [26].

### 2.4.3 Συνελικτικά νευρωνικά δίκτυα

Τα συνελικτικά νευρωνικά δίκτυα (Convolutional Neural Networks – CNNs) χρησιμοποιούνται κυρίως στην όραση των υπολογιστών και σε εφαρμογές ταξινόμησης εικόνας. Έχουν τη δυνατότητα να εντοπίζουν χαρακτηριστικά και πρότυπα μέσα σε εικόνες και βίντεο, επιτρέποντας εργασίες όπως η ανίχνευση αντικειμένων, η αναγνώριση εικόνας, η αναγνώριση προτύπων και η αναγνώριση προσώπων. Τα CNNs αξιοποιούν αρχές της γραμμικής άλγεβρας, ιδιαίτερα τον πολλαπλασιασμό πινάκων, για να εντοπίσουν μοτίβα μέσα σε μια εικόνα.

Τα CNNs αποτελούνται από συνελικτικά στρώματα. Σε πολύπλοκες εφαρμογές, ένα CNN μπορεί να περιέχει έως και χιλιάδες κρυφά στρώματα. Μέσω της συνέλιξης, που επεξεργάζεται επανειλημμένα τα αρχικά δεδομένα εισόδου, μπορούν να αποκαλυφθούν λεπτομερή πρότυπα. Με κάθε επιπλέον στρώμα, ένα CNN γίνεται όλο και πιο πολύπλοκο, εντοπίζοντας όλο και μεγαλύτερα τμήματα της εικόνας. Τα πρώτα στρώματα επικεντρώνονται σε απλά χαρακτηριστικά, όπως χρώματα και άκρα. Καθώς τα δεδομένα της εικόνας προχωρούν μέσα από τα στρώματα του CNN, το δίκτυο αρχίζει να αναγνωρίζει μεγαλύτερα στοιχεία ή σχήματα του αντικειμένου, μέχρι να εντοπίσει πλήρως το ζητούμενο αντικείμενο.

Τα CNNs ξεχωρίζουν από άλλα TNΔ χάρη στην ανώτερη απόδοσή τους όταν η είσοδός τους αποτελείται από εικόνες. Πριν από την εμφάνιση των CNNs, η αναγνώριση αντικειμένων σε εικόνες γινόταν με μη αυτόματες και χρονοβόρες μεθόδους εξαγωγής χαρακτηριστικών. Πλέον, τα CNNs προσφέρουν μια πιο επεκτάσιμη και αυτοματοποιημένη προσέγγιση στην ταξινόμηση εικόνων και στην αναγνώριση αντικειμένων, και μπορούν να διαχειριστούν αποτελεσματικότερα δεδομένα υψηλών διαστάσεων. Υπάρχουν, ωστόσο, και μειονεκτήματα. Τα CNNs είναι υπολογιστικά απαιτητικά, με σημαντικό κόστος σε χρόνο και προϋπολογισμό, καθώς απαιτούν πολλές μονάδες γραφικών (GPUs) [22].

Η βάση των ΤΝΔ είναι οι γραμμικοί μετασχηματισμοί: ένα διάνυσμα λαμβάνεται ως είσοδος και πολλαπλασιάζεται με έναν πίνακα για να παραχθεί η έξοδος (στην οποία συνήθως προστίθεται και ένα διάνυσμα πόλωσης – bias – πριν περάσει το αποτέλεσμα από μια μη γραμμική συνάρτηση ενεργοποίησης). Αυτό είναι εφαρμόσιμο σε κάθε τύπο εισόδου, είτε πρόκειται για εικόνα, ηχητικό απόσπασμα, είτε για μια μη ταξινομημένη συλλογή χαρακτηριστικών: ανεξαρτήτως της διάστασής τους, η αναπαράστασή τους μπορεί πάντα να ισοπεδωθεί (flattened) σε ένα διάνυσμα πριν τον μετασχηματισμό.

Οι εικόνες αναπαρίστανται από τρεις διαστάσεις:

- το ύψος της εικόνας
- το μήκος της εικόνας
- το πλήθος των χρωματικών καναλιών

Τα CNNs, σε αντίθεση με τα MLPs ή τα LSTM, μπορούν να διατηρούν αυτήν την τοπολογική πληροφορία. Αυτό συμβαίνει καθώς μπορούν να «σκανάρουν» κατευθείαν την εικόνα, χωρίς να πρέπει να γίνει η διαδικασία του flatten [27].

#### 2.4.3.1 Συνέλιξη Τετραγωνικών Πινάκων

Στην ενότητα αυτή θα περιγραφεί η διαδικασία της συνέλιξης μεταξύ τετραγωνικών πινάκων. Στο Σχήμα 2.4 φαίνεται ένα παράδειγμα του αποτελέσματος της συνέλιξης μεταξύ δύο τετραγωνικών πινάκων. Ο πρώτος πίνακας (A) αντιπροσωπεύει την εικόνα ενώ ο δεύτερος (B) το φίλτρο. Τα συνελικτικά στρώματα των CNN ουσιαστικά επιτελούν συνέλιξη μεταξύ της (τροποποιημένης) εικόνας και ενός φίλτρου.

Τα φίλτρα σε μία συνέλιξη είναι αρμόδια στο να ανιχνεύουν συγκεκριμένα χαρακτηριστικά. Διαφορετικά φίλτρα ανιχνεύουν διαφορετικά edges (κάθετες γραμμές, οριζόντιες γραμμές, κύκλους, λοξές γραμμές κλπ). Κατά την διαδικασία ανίχνευσης αντικειμένων, εκτελείται η πράξη της συνέλιξης για *c* διαφορετικά φίλτρα, που το καθένα ανιχνεύει κάτι διαφορετικό, οπότε θα λαμβάνεται ως έξοδος *c* στο πλήθος διαφορετικά output frames. Σύμφωνα με μελέτες, το ιδανικό μέγεθος φίλτρου στην αναγνώριση αντικειμένων με CNN, είναι το 3×3 [28].



Σχήμα 2.4: Συνέλιξη τετραγωνικών πινάκων

Όπου:

- *n*: η διάσταση της εικόνας
- f: η διάσταση του φίλτρου
- (*n-f*+1): η διάσταση του παραχθέντος τετραγωνικού πίνακα (αποτέλεσμα συνέλιξης)

Συνέλιξη μεταξύ του frame A και του φίλτρου B δίνει το αποτέλεσμα συνέλιξης (output frame) C.

$$\begin{bmatrix} a_{00} & \cdots & a_{0n} \\ \vdots & \ddots & \vdots \\ a_{n0} & \cdots & a_{nn} \end{bmatrix} * \begin{bmatrix} b_{00} & \cdots & b_{0m} \\ \vdots & \ddots & \vdots \\ b_{m0} & \cdots & b_{mm} \end{bmatrix} = \begin{bmatrix} c_{00} & \cdots & c_{0k} \\ \vdots & \ddots & \vdots \\ c_{k0} & \cdots & c_{kk} \end{bmatrix}, m \le n \quad (2.2)$$

Στη συνέχεια, περιγράφεται αναλυτικά η διαδικασία υπολογισμού του πίνακα C. Για τον υπολογισμό του  $c_{0j}$ , απομονώνεται το κομμάτι του πίνακα A για  $ij \in [00,mm]$  και πολλαπλασιάζονται στοιχείο προς στοιχείο οι πίνακες A (μαύρος πίνακας) και B (κόκκινος πίνακας). Το άθροισμα των παραπάνω γινομένων, δίνει το στοιχείο  $c_{00}$ . Στο Σχήμα 2.5 φαίνεται σχηματικά το πως εφαρμόζεται το φίλτρο B στον πίνακα A ώστε να προκύψει το  $c_{00}$ .

Σχήμα 2.5: Υπολογισμός c<sub>00</sub>

Έπειτα, πραγματοποιείται μετατόπιση του *B* κατά ένα στοιχείο δεξιά και επαναλαμβάνεται η ίδια διαδικασία για τον υπολογισμό των *c*<sub>0j</sub>, όπως φαίνεται και στο Σχήμα 2.6.



Σχήμα 2.6: Υπολογισμός c<sub>01</sub>

Για τον υπολογισμό των c<sub>1j</sub> ο πίνακας *B* επιστρέφει στην αρχική του θέση, μετατοπίζεται ένα στοιχείο προς τα κάτω και επαναλαμβάνεται η προαναφερθείσα διαδικασία (Σχήμα 2.7).



**Σχήμα 2.7:** Υπολογισμός  $c_{10}$ 

Η συνέλιξη είναι υπεύθυνη για την αναγνώριση χαρακτηριστικών (features) και ακμών (edges) στις εικόνες.

### 2.4.3.2 Έγχρωμη εικόνα (RGB)

Όταν χρειάζεται να γίνει εξαγωγή στοιχείων από έγχρωμη εικόνα με RGB κανάλια, εφαρμόζεται η συνέλιξη με το φίλτρο σε κάθε ένα κανάλι ξεχωριστά (δηλαδή τρεις πίνακες: A1 (κόκκινος), A2 (πράσινος), A3 (μπλε)), όπως φαίνεται και στο Σχήμα 2.8.



Σχήμα 2.8: Συνέλιξη φίλτρου με έγχρωμη εικόνα

# 2.4.3.3 Βήμα μετατόπισης (Strides) και τεχνική συμπλήρωσης με μηδενικά (Padding)

### • Strides

Το μέγεθος της οριζόντιας/κάθετης μετατόπισης που υφίσταται το φίλτρο μετά από κάθε συνέλιξη.

Ισχύει ότι:

- Stride =  $1 \rightarrow \lambda$ επτομερής ανάλυση (χωρίς απώλεια διαστάσεων)
- Stride > 1  $\rightarrow$  αραίωση του χάρτη εξόδου (υποδειγματοληψία)
- Padding

Επειδή κατά την συνέλιξη ενός πίνακα A με ένα φίλτρο B τα ακριανά στοιχεία δεν περιλαμβάνονται στις πράξεις αρκετές φορές (υποχρησιμοποίηση πληροφορίας), εφαρμόζεται η τεχνική της συμπλήρωσης με μηδενικά (padding) στον πίνακα A και μετασχηματίζεται στον πίνακα A'.



A′							
0	0	0	0	0	0	0	0
0							0
0							0
0							0
0							0
0							0
0							0
0	0	0	0	0	0	0	0

Σχήμα 2.9: Τεχνική συμπλήρωσης με μηδενικά

Παρατήρηση: μετά τον ορισμό των παραμέτρων strides και padding το αποτέλεσμα της συνέλιξης έχει τις διαστάσεις που φαίνονται στο Σχήμα 2.10.



Σχήμα 2.10: Συνέλιξη με χρήση strides και padding

- *n*: Διαστάσεις της εικόνας
- f: Διαστάσεις του φίλτρου

- *p*: Padding
- s: Strides
- [ ]: Συνάρτηση ακεραίου μέρους (πχ [2,7]=2)

### 2.4.3.4 Αλγόριθμοι βελτιστοποίησης

Σκοπός ενός αλγορίθμου βελτιστοποίησης είναι να βελτιστοποιήσει μια αντικειμενική συνάρτηση (συνήθως ονομάζεται συνάρτηση απώλειας – loss function), η οποία αντιπροσωπεύει τη μέση απόκλιση μεταξύ των προβλεπόμενων και των αναμενόμενων τιμών. Η βελτιστοποίηση αυτής της συνάρτησης σημαίνει την εύρεση του συνόλου παραμέτρων (βάρη και πολώσεις στην περίπτωση των TNΔ) που οδηγεί στα καλύτερα αποτελέσματα για το εκάστοτε πρόβλημα. Οι πιο κοινοί αλγόριθμοι βελτιστοποίησης που χρησιμοποιούνται για την εκπαίδευση TNΔ είναι οι αλγόριθμοι που βασίζονται στη μείωση της κλίσης. Τυπικά παραδείγματα τέτοιων αλγορίθμων αποτελούν οι Adam, AdaDelta, και RMSProp [29]. Οι αλγόριθμοι αυτοί συμπεριλαμβάνουν και στοχαστικές παραμέτρους προκειμένου να μειώσουν την πιθανότητα πρόωρης σύγκλισης σε κάποιο τοπικό βέλτιστο. Η εφαρμογή των αλγορίθμων gradient descent στα TNΔ γίνεται μέσω της μεθόδου οπίσθιας διάδοσης (backpropagation) του σφάλματος.

### 2.4.3.5 Συνάρτηση ενεργοποίησης (Activation function)

Στα ΤΝΔ, οι συναρτήσεις ενεργοποίησης είναι εξαιρετικά σημαντικές, καθώς βοηθούν στη μάθηση και στην κατανόηση μη γραμμικών και πολύπλοκων αντιστοιχιών μεταξύ εισόδων και εξόδων. Οι συναρτήσεις ενεργοποίησης χρησιμοποιούνται για να μετασχηματίσουν ένα σήμα εισόδου σε ένα σήμα εξόδου, το οποίο με τη σειρά του τροφοδοτείται ως είσοδος στο επόμενο στρώμα. Σε ένα ΤΝΔ, υπολογίζεται το άθροισμα των γινομένων των εισόδων με τα αντίστοιχα βάρη και εφαρμόζεται μια συνάρτηση ενεργοποίησης ώστε να ληφθεί η έξοδος του συγκεκριμένου στρώματος, η οποία στη συνέχεια δίνεται ως είσοδος στο επόμενο στρώμα. Η μαθηματική μοντελοποίηση φαίνεται στην Εξίσωση (2.3):

$$Y_{i} = f\left(\sum_{i} X_{i} * W_{ij} + B_{j}\right)$$
(2.3)

όπου:

- W<sub>ij</sub>: δηλώνει την i-οστή σύναψη βάρους στο j-οστό στρώμα
- Χ<sub>i</sub>: αντιπροσωπεύει την είσοδο που αντιστοιχεί στον νευρώνα
- B<sub>i</sub>: είναι το διάνυσμα πόλωσης (bias) του j-οστού στρώματος
- Υ<sub>i</sub>: είναι ο χάρτης χαρακτηριστικών εξόδου του j-οστού στρώματος
- *f*: είναι μια συνάρτηση ενεργοποίησης

Αν δεν χρησιμοποιηθεί συνάρτηση ενεργοποίησης, τότε το σήμα εξόδου θα είναι απλώς μια γραμμική συνάρτηση — δηλαδή ένα πολυώνυμο πρώτου βαθμού. Παρόλο που μια γραμμική εξίσωση είναι απλή και εύκολη στη λύση, η υπολογιστική της ικανότητα είναι περιορισμένη και δεν μπορεί να μάθει ή να αναγνωρίσει πολύπλοκες απεικονίσεις. Ένα ΤΝΔ χωρίς συνάρτηση ενεργοποίησης λειτουργεί όπως ένα γραμμικό μοντέλο παλινδρόμησης, με περιορισμένη απόδοση και υπολογιστική ισχύ. Επομένως, είναι επιθυμητό ένα ΤΝΔ να μην μαθαίνει και να υπολογίζει απλά μια γραμμική συνάρτηση, αλλά να εκτελεί πιο περίπλοκες εργασίες, όπως η μοντελοποίηση πολύπλοκων δεδομένων (εικόνες, βίντεο, ήχος, ομιλία, κείμενο, κλπ). Αυτός είναι ο λόγος που χρησιμοποιούνται συναρτήσεις ενεργοποίησης και τεχνικές ΤΝΔ, όπως η βαθιά μάθηση (Deep Learning), οι οποίες είναι κατάλληλες για πολύπλοκα, υψηλής διάστασης και μη γραμμικά δεδομένω. Σε αυτά, το μοντέλο διαθέτει πολλαπλά κρυφά στρώματα και σύνθετη αρχιτεκτονική ώστε να εξάγει γνώση — που είναι και ο απώτερος στόχος.

Οι βασικοί τύποι συναρτήσεων ενεργοποίησης είναι οι εξής [30]:

- 1. Δυαδική Συνάρτηση Βήματος (Binary Step Function)
- 2. Γραμμική (Linear)
- 3. Σιγμοειδής (Sigmoid)
- 4. Υπερβολική Εφαπτομένη (Tanh)
- 5. ReLU (Rectified Linear Unit)
- 6. Leaky ReLU
- 7. Παραμετροποιημένη ReLU (Parametrized ReLU)
- 8. Εκθετική Γραμμική Μονάδα (Exponential Linear Unit ELU)
- 9. Swish
- 10. SoftMax

#### 2.4.4 Συνελικτικοί αυτοκωδικοποιητές

Οι αυτοκωδικοποιητές (Autoencoders) είναι είδη βαθιών νευρωνικών δικτύων, σχεδιασμένα να μαθαίνουν την κωδικοποίηση των εισερχόμενων δεδομένων. Τα δεδομένα αυτά μπορεί να είναι μονοδιάστατα (1D), δισδιάστατα (2D) ή και πολυδιάστατα, όπως για παράδειγμα χρονοσειρές, εικόνες ή σήματα επιτάχυνσης. Ο βασικός στόχος των αυτοκωδικοποιητών είναι να συμπιέσουν τις πληροφορίες εισόδου σε ένα μικρότερο διάνυσμα χαρακτηριστικών (feature vector), μέσω ενός πρώτου νευρωνικού υποσυστήματος που λειτουργεί ως κωδικοποιητής και στη συνέχεια, χρησιμοποιώντας αυτό το διάνυσμα, να αναπαραγάγουν την αρχική είσοδο μέσω ενός δεύτερου νευρωνικού υποσυστήματος που λειτουργεί ως αποκωδικοποιητής (decoder). Το διάνυσμα αυτό τοποθετείται σε ένα ενδιάμεσο σημείο του δικτύου που ονομάζεται "στένωση" (bottleneck), και η συνολική στρατηγική τόσο της κωδικοποίησης όσο και της αποκωδικοποίησης αποτυπώνεται άμεσα στην αρχιτεκτονική του αυτοκωδικοποιητή. Η αρχιτεκτονική αυτή αποσκοπεί στη μείωση των διαστάσεων των δεδομένων σε λιγότερα χαρακτηριστικά, με τέτοιον τρόπο ώστε να είναι εφικτή η αξιόπιστη ανακατασκευή τους. Αυτή η ιδιότητα είναι ιδιαιτέρως χρήσιμη σε εφαρμογές όπως η συμπίεση δεδομένων ή η σύγκριση εικόνων με βάση χαρακτηριστικά πέραν του επιπέδου των μεμονωμένων εικονοστοιχείων (pixels). Στο Σχήμα 2.12 απεικονίζεται διαγραμματικά η δομή του αυτοκωδικοποιητή.



Σχήμα 2.12: Δομή του αυτοκωδικοποιητή

### Κωδικοποιητής και Αποκωδικοποιητής

Ένας αυτοκωδικοποιητής περιλαμβάνει δύο βασικά νευρωνικά υποσυστήματα:

- Ένα δίκτυο κωδικοποίησης (encoder), το οποίο λαμβάνει τα δεδομένα εισόδου x ∈ ℝ<sup>n</sup> και τα μετασχηματίζει σε μία κωδικοποιημένη αναπαράσταση z ∈ ℝ<sup>m</sup>, m < n (code embedding) στο bottleneck, μέσω της εφαρμογής μιας μη γραμμικής συνάρτησης z, χαρακτηριστικής των TNΔ εμπρόσθιας τροφοδότησης.
- Ένα δίκτυο αποκωδικοποίησης (decoder), το οποίο λαμβάνει ως είσοδο το διάνυσμα από το επίπεδο στένωσης και επιχειρεί να το αναπροσαρμόσει στον αρχικό χώρο των εισόδων R<sup>n</sup>.

Η έξοδος  $\hat{x}$  έχει παρόμοιες διαστάσεις με την αρχική είσοδο x, γεγονός που επιτρέπει τη σύγκρισή τους και την εκτίμηση της απόκλισης μέσω μιας κατάλληλης αντικειμενικής συνάρτησης. Ο στόχος είναι η έξοδος να αναπαράγει με τη μέγιστη δυνατή ακρίβεια την είσοδο. Συνηθισμένες μετρικές περιλαμβάνουν το Μέσο Τετραγωνικό Σφάλμα (Mean Squared Error – MSE) ή τη διασταυρούμενη εντροπία (cross-entropy), ανάλογα με τη συνάρτηση ενεργοποίησης.

Παρακάτω φαίνονται οι μαθηματικές σχέσεις που περιγράφουν έναν αυτοκωδικοποιητή:

$$x \xrightarrow{Encoder} z \xrightarrow{Decoder} \hat{x}$$
(2.4)

$$z = f_{CE}(x) \tag{2.5}$$

$$\hat{x} = f_{CD}(z) \tag{2.6}$$

όπου:

-  $f_{CE}$ : Η διαδικασία κωδικοποίησης

*f*<sub>CD</sub>: Η διαδικασία αποκωδικοποίησης

Ένας αυτοκωδικοποιητής εκπαιδεύεται για να μπορεί να ανακατασκευάζει την είσοδό του. Επομένως, ιδανικά η έξοδος ενός αυτοκωδικοποιητή πρέπει να ταυτίζεται με την είσοδό του. Όσον αφορά στους αυτοκωδικοποιητές, για κάθε είσοδο, η έξοδος είναι εκ των προτέρων γνωστή, άρα ανήκουν στην κατηγορία της μη επιτηρούμενης μάθησης.

#### • Στένωση

Κομβικό στοιχείο των αποκωδικοποιητών είναι το ενδιάμεσο στρώμα στένωσης (bottleneck) – επίσης γνωστό ως "code layer" ή "embedding" – το οποίο διαθέτει λιγότερους νευρώνες σε σχέση με την είσοδο. Η έξοδος, από την άλλη, έχει ίδιες διαστάσεις με τα αρχικά δεδομένα ώστε να καθίσταται δυνατή η ποιοτική ανακατασκευή μέσω σύγκρισης της εισόδου και της εξόδου.

Η στένωση αντιπροσωπεύει την κωδικοποιημένη πληροφορία. Ο αριθμός νευρώνων στο στρώμα στένωσης αντιστοιχεί στη διάσταση της κωδικοποιημένης πληροφορίας. Η αύξηση των νευρώνων του στρώματος στένωσης μπορεί να οδηγεί σε μικρότερα σφάλματα ανακατασκευής, ωστόσο μειώνει τα οφέλη της κωδικοποίησης [31].

Όλα τα επιμέρους στοιχεία που συνιστούν έναν συνελικτικό αυτοκωδικοποιητή, παρουσιάζονται στο Σχήμα 2.13.



Σχήμα 2.13: Δομή συνελικτικού αυτοκωδικοποιητή

### 2.4.4.1 Αυτοκωδικοποιητές 2D και 3D συνελίξεων

Ανάλογα με την δομή της πληροφορίας εξόδου, διακρίνονται δύο είδη συνελικτικών αυτοκωδικοποιητών: 2D συνέλιξη και 3D συνέλιξη.

### 2D-Συνέλιξη

Εφαρμόζεται συνέλιξη σε μία ακολουθία εικόνων. Το κάθε φίλτρο δύο διαστάσεων εφαρμόζεται ταυτόχρονα σε κάθε εικόνα σύμφωνα με την διαδικασία της Ενότητας 2.4.3.1. Το αποτέλεσμα είναι ένας χάρτης χαρακτηριστικών δύο διαστάσεων (Σχήμα 2.14).



Σχήμα 2.14: 2D Συνέλιξη

όπου:

- *H*<sub>1</sub>: Πρώτη διάσταση εικόνας εισόδου
- W<sub>1</sub>: Δεύτερη διάσταση εικόνας εισόδου
- L: Πλήθος ακολουθίας εικόνων
- f: Μέγεθος φίλτρου
- Η<sub>2</sub>: Πρώτη διάσταση εικόνας εξόδου
- W<sub>2</sub>: Δεύτερη διάσταση εικόνας εξόδου

Αν αυτή η διαδικασία πραγματοποιείται για *K* στο πλήθος φίλτρα (Ενότητα 2.4.3.2), προκύπτει ότι:

Input 
$$(H_1 x W_1 x L) \rightarrow Output (H_2 x W_2 x K)$$
 (2.7)

### 3D-Συνέλιξη

Στην 3D συνέλιξη, κάθε φίλτρο τριών διαστάσεων εφαρμόζεται ταυτόχρονα σε dεικόνες (στην περίπτωση του Σχήματος 2.14 έγινε η επιλογή d=f) ξεκινώντας από την πρώτη εικόνα και παράγει έναν χάρτη χαρακτηριστικών σύμφωνα με την διαδικασία της Ενότητας 2.4.3.1 (δηλαδή εφαρμογή στις εικόνες 1 ως f). Στην συνέχεια πραγματοποιείται η ίδια διαδικασία ξεκινώντας από την δεύτερη εικόνα (δηλαδή εφαρμογή στις εικόνες 2 ως f+1). Η διαδικασία τερματίζεται όταν το φίλτρο έχει εφαρμοστεί σε όλον τον άξονα του χρόνου (δηλαδή μέχρι και τις εικόνες L-f+1 ως L). Το αποτέλεσμα είναι ένας όγκος χαρακτηριστικών τριών διαστάσεων (Σχήμα 2.15).



Σχήμα 2.15: 3D Συνέλιξη

Όπου:

- *H*<sub>1</sub>: Πρώτη διάσταση εικόνας εισόδου
- W<sub>1</sub>: Δεύτερη διάσταση εικόνας εισόδου
- L: Πλήθος ακολουθίας εικόνων εισόδου
- f: Μέγεθος φίλτρου
- *H*<sub>2</sub>: Πρώτη διάσταση εικόνας εξόδου
- W<sub>2:</sub> Δεύτερη διάσταση εικόνας εξόδου
- L'(=L-d+1): Πλήθος ακολουθίας εικόνων εξόδου

Αν αυτή η διαδικασία πραγματοποιείται για *K* στο πλήθος φίλτρα (Ενότητα 2.4.3.2), προκύπτει ότι:

Input 
$$(H_1 \times W_1 \times L) \rightarrow \text{Output} (H_2 \times W_2 \times (L - f + 1) \times K)$$
 (2.8)

# 2.4.4.2 Εκπαίδευση μοντέλων συνελικτικών αυτοκωδικοποιητών – Προβλήματα υπερπροσαρμογής και υποπροσαρμογής

Ένα από τα πιο κοινά προβλήματα κατά την εκπαίδευση TNΔ είναι η υπερπροσαρμογή (overfitting) ή η υποπροσαρμογή (underfitting) στο σύνολο δεδομένων εκπαίδευσης.

### • Υπερπροσαρμογή:

Υπερπροσαρμογή είναι η εκμάθηση τυχαίων παραμέτρων και ιδιαιτεροτήτων (θόρυβος) πέρα από την θεμελιώδη σχέση που συνδέει την είσοδο και την έξοδο ή αλλιώς η υπερβολική προσαρμογή στο σύνολο δεδομένων εκπαίδευσης. Η υπερπροσαρμογή δεν είναι επιθυμητή, διότι επηρεάζει την ικανότητα γενίκευσης του μοντέλου, δηλαδή την ικανότητά του να ανταποκριθεί σε καινούρια δεδομένα. Ένας τρόπος για να εντοπιστεί ένα υπερπροσαρμοσμένο μοντέλο είναι η εξέταση του εάν υπάρχει αναντιστοιχία μεταξύ του τελικού μέσου σφάλματος κατά την εκπαίδευση και του μέσου σφάλματος κατά την εκπαίδευση και του μέσου σφάλματος κατά την εκπαίδευση του μοντέλου.

### • Υποπροσαρμογή:

Η υποπροσαρμογή είναι το αντίθετο της υπερπροσαρμογής. Ως υποπροσαρμογή ορίζεται η αδυναμία του μοντέλου να μάθει ικανοποιητικά τη σχέση εισόδου-εξόδου από το σύνολο δεδομένων εκπαίδευσης

### Αποφυγή υπερπροσαρμογής και υποπροσαρμογής με χρήση πρόωρης διακοπής

Στην παρούσα διπλωματική χρησιμοποιείται η μέθοδος της πρώιμης διακοπής της εκπαίδευσης (early stopping). Για να εφαρμοστεί η τεχνική αυτή, εκτός από το σύνολο εκπαίδευσης, κατά την εκπαίδευση απαιτείται να υπάρχει και το σύνολο επικύρωσης (validation), ώστε να καθορίζονται τα κριτήρια διακοπής της εκπαίδευσης. Έτσι, τα δεδομένα χωρίζονται σε τρία μέρη [32]:

- 1. Σύνολο για εκπαίδευση (αναγνώριση χαρακτηριστικών και μοτίβων)
- 2. Σύνολο για επικύρωση (αξιολόγηση δυνατότητας γενίκευσης)
- 3. Σύνολο για δοκιμή (αξιολόγηση μοντέλου)

Η πρώιμη διακοπή πραγματοποιείται ως εξής:

- 1) Εκκίνηση εκπαίδευσης
- Σε κάθε επανάληψη, υπολογισμός αντικειμενικής συνάρτησης για σύνολο εκπαίδευσης και σύνολο επικύρωσης
- Εάν η αντικειμενική συνάρτηση, υπολογισμένη για το σύνολο επικύρωσης, δεν βελτιωθεί για έναν προκαθορισμένο αριθμό επαναλήψεων, διακόπτεται η εκπαίδευση
- Ως βέλτιστος συνδυασμός βαρών κρατείται αυτός για τον οποίο η αντικειμενική συνάρτηση εμφάνισε την καλύτερη τιμή για το σύνολο επικύρωσης

Η παραπάνω διαδικασία, παρά το γεγονός ότι είναι ζωτικής σημασίας για την έγκυρη εξαγωγή του μοντέλου, παρουσιάζει ορισμένες προκλήσεις, όπως:

- Δυσκολία στον προσδιορισμό του προκαθορισμένου αριθμού επαναλήψεων «υπομονής» (patience)
- Αύξηση υπολογιστικής πολυπλοκότητας, καθώς σε κάθε επανάληψη πρέπει να γίνει ο υπολογισμός της αντικειμενικής συνάρτησης δύο φορές αντί για μία, αλλά και επειδή πρέπει να αποθηκεύεται η καλύτερη κατάσταση

### 2.4.4.3 Κανονικοποίηση κατά παρτίδες (Batch Normalization)

Η κανονικοποίηση κατά παρτίδες (Batch Normalization – BN) είναι μια τεχνική κανονικοποίησης των συναρτήσεων ενεργοποίησης στα κρυφά στρώματα ενός TNΔ. Η τάση της να βελτιώνει την ακρίβεια και να επιταχύνει την εκπαίδευση, την έχει καθιερώσει ως μία από τις πιο δημοφιλείς τεχνικές στη βαθιά μάθηση. Η BN πρωτίστως επιτρέπει την εκπαίδευση με μεγαλύτερους ρυθμούς μάθησης, κάτι που οδηγεί σε ταχύτερη σύγκλιση και καλύτερη γενίκευση. Η BN αποφεύγει αυτό το πρόβλημα, κανονικοποιώντας συνεχώς τις
συναρτήσεις ενεργοποίησης ώστε να έχουν μέση τιμή μηδέν και τυπική απόκλιση τη μονάδα. Με την άνοδο της βαθιάς μάθησης, η BN επεκτείνει αυτή την ιδέα στα κρυφά στρώματα, αν και για λόγους ταχύτητας η κανονικοποίηση εφαρμόζεται σε μικρές παρτίδες και όχι σε ολόκληρο το σύνολο εκπαίδευσης [33].

# ΚΕΦΑΛΑΙΟ 3

# ΑΝΑΛΥΣΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΕΙΚΟΝΩΝ ΤΟΥ ΟΥΡΑΝΟΥ

#### 3.1 ΕΙΣΑΓΩΓΗ

Για την πρόβλεψη ακολουθίας εικόνων του ουρανού με χρήση CNNs, χρειάζεται ένα επαρκές πλήθος ιστορικών δεδομένων εικόνων. Αυτό θα χρησιμοποιηθεί στην εκπαίδευση του μοντέλου, ώστε να γίνει η εκμάθηση των μοτίβων της διασποράς των νεφών και τελικά να γίνει η εξαγωγή της κατάλληλης σχέσης εισόδου-εξόδου. Από την στιγμή της συλλογής των εικόνων από το αποθηκευτικό σύστημα της κάμερας, μέχρι την διαδικασία εισαγωγής τους για την δημιουργία μοντέλου πρόβλεψης, λαμβάνει χώρα η ανάλυση και επεξεργασία τους.

#### 3.2 ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ

Στην παρούσα διπλωματική εργασία χρησιμοποιήθηκαν επίγειες εικόνες του ουρανού που λήφθηκαν σε διάστημα δεκατεσσάρων ημερών (από 16/11/2023 έως 29/11/2023), από την ανατολή μέχρι την δύση του ηλίου, με συχνότητα δειγματοληψίας ένα καρέ ανά λεπτό. Συνολικά συγκεντρώθηκαν 8.168 εικόνες, μέσω κάμερας παρακολούθησης ASI-16 κατασκευασμένη από την εταιρεία CMS Ing. Dr. Schreder GmbH. Η συγκεκριμένη κάμερα διαθέτει πεδίο θέασης 180°, ανάλυση 5MP και ευρυγώνιο φακό τύπου "fisheye", προστατευμένο με αντανακλαστικό-ανθεκτικό θόλο από χαλαζία, με κυκλοφορία του αέρα για αποφυγή θολώματος. Οι εικόνες που λήφθηκαν ήταν έγχρωμες (RGB) και αποθηκεύτηκαν σε μορφή JPG με μέσο μέγεθος 76,52 KB.

Οι τρεις κύριες καταστάσεις του ουρανού που παρατηρούνται είναι η ηλιοφάνεια, η μερική νεφοκάλυψη και η πλήρης νεφοκάλυψη σε ποσοστά 32%, 22%, και 46% αντίστοιχα. Στα Σχήματα 3.1, 3.2, και 3.3 παρουσιάζονται παραδείγματα επίγειων εικόνων του ουρανού για τις τρεις αυτές καταστάσεις. Για την εξαγωγή όσο το δυνατόν πιο έγκυρων αποτελεσμάτων, επιλέχθηκε τοποθεσία λήψης τέτοια ώστε να ελαχιστοποιείται ο θόρυβος από εξωτερικούς παράγοντες (όπως κτήρια, δέντρα), σε ένα Φ/Β σύστημα εγκατεστημένης ισχύος 1.2 kW στην περιοχή της Βοιωτίας. Ωστόσο, καθώς η κάμερα διαθέτει ευρυγώνιο φακό με πεδίο θέασης 180°, είναι αναπόφευκτο να παρατηρούνται και στοιχεία πέρα από τον ήλιο και τα σύννεφα, όπως βουνά ή έντομα πάνω στον φακό της κάμερας (Σχήμα 3.4) [34].



Σχήμα 3.1: Παράδειγμα εικόνας του ουρανού για συνθήκες ηλιοφάνειας



Σχήμα 3.2: Παράδειγμα εικόνας του ουρανού για συνθήκες μερικής νεφοκάλυψης



Σχήμα 3.3: Παράδειγμα εικόνας του ουρανού για συνθήκες πλήρους νεφοκάλυψης



Σχήμα 3.4: Παράδειγμα εικόνας του ουρανού με έντομο πάνω στον φακό της κάμερας

#### 3.3 ΕΠΕΞΕΡΓΑΣΙΑ ΑΡΧΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Οι εικόνες του ουρανού που λαμβάνονται από την κάμερα είναι έγχρωμες με ανάλυση 1536×1536 εικονοστοιχεία (pixels). Η επεξεργασία των εικόνων του ουρανού γίνεται με χρήση της γλώσσας προγραμματισμού python. Οι εικόνες φορτώνονται στην python ως τρισδιάστατοι πίνακες (ύψος × μήκος × χρωματικό κανάλι), όπου κάθε στοιχείο του πίνακα αντιστοιχεί στην τιμή  $\alpha \in [0, 255]$  που έχει το αντίστοιχο εικονοστοιχείο σε αυτό το κανάλι. Όπως αναφέρθηκε στο Κεφάλαιο 2, τα CNN επιτελούν συνέλιξη εφαρμόζοντας φίλτρα πάνω στις εικόνες, ανά ομάδα εικονοστοιχείων. Είναι φανερό λοιπόν πως η διατήρηση των αρχικών διαστάσεων των εικόνων είναι υπολογιστικά μη βιώσιμη. Για αυτόν τον λόγο, το σύνολο των εικόνων υφίσταται διαδοχικά απεικόνιση στην κλίμακα του γκρι και μείωση της ανάλυσης.

#### 3.3.1 Απεικόνιση εικόνων στην κλίμακα του γκρι (grayscaling)

Το πρώτο στάδιο επεξεργασίας των εικόνων του ουρανού περιλαμβάνει την απεικόνιση στην κλίμακα του γκρι, μειώνοντας τα κανάλια της κάθε εικόνας από τρία (Red-Green-Blue) σε ένα. Με αυτόν τον τρόπο περιορίζεται το πλήθος των πράξεων, χωρίς να χάνεται πληροφορία που αφορά στην αναγνώριση αντικειμένων όπως είναι ο ήλιος και η κατανομή των νεφών, τα οποία και χρειάζονται για την πρόβλεψη μελλοντικών εικόνων του ουρανού.

Η απεικόνιση των εικόνων στην κλίμακα του γκρι στην python πραγματοποιείται με την βιβλιοθήκη OpenCV, σύμφωνα με την σχέση 3.1 [35].

$$Y_{\rm i} = 0.299R_{\rm i} + 0.587G_{\rm i} + 0.114B_{\rm i} \tag{3.1}$$

όπου:

- R<sub>i</sub>: Η τιμή του εικονοστοιχείου *i* του κόκκινου καναλιού της αρχικής εικόνας
- G<sub>i</sub>: Η τιμή του εικονοστοιχείου *i* του πράσινου καναλιού της αρχικής εικόνας

- B<sub>i</sub>: Η τιμή του εικονοστοιχείου i του μπλε καναλιού της αρχικής εικόνας
- Υ<sub>i</sub>: Η τιμή του εικονοστοιχείου i της νέας εικόνας

Παράδειγμα απεικόνισης μιας επίγειας εικόνας του ουρανού στην κλίμακα του γκρι φαίνεται στο Σχήμα 3.5.



Σχήμα 3.5: Απεικόνιση εικόνας στην κλίμακα του γκρι (grayscaling)

Στο νέο κανάλι που προκύπτει μετά την επεξεργασία, κάθε εικονοστοιχείο έχει μία ακέραιη τιμή στο διάστημα [0, 255]. Η τιμή 0 αντιστοιχεί στο μαύρο χρώμα, η τιμή 255 στο λευκό, ενώ οι ενδιάμεσες τιμές αντιστοιχούν στις ενδιάμεσες αποχρώσεις του γκρι (Σχήμα 3.6).



Σχήμα 3.6: Εννιά χρωματικά δείγματα από την κλίμακα του γκρι

#### 3.3.2 Μείωση ανάλυσης εικόνων

Μετά την απεικόνιση των δεδομένων στην κλίμακα του γκρι, το σύνολο των εικόνων υφίσταται μείωση της ανάλυσης από 1536×1536 εικονοστοιχεία σε 32×32 εικονοστοιχεία. Η επεξεργασία αυτή κρίνεται απαραίτητη, καθώς μειώνεται σημαντικά το πλήθος των συνολικών εικονοστοιχείων της κάθε εικόνας (από 1536<sup>2</sup>=2359296 σε 32<sup>2</sup>=1024) και άρα το πλήθος των συνελίξεων και ο χρόνος εκπαίδευσης. Όπως φαίνεται στο Σχήμα 3.7, η μείωση της ανάλυσης δεν έρχεται σε βάρος της αναγνώρισης των επιθυμητών στοιχείων (σύννεφα, ήλιος).

Η μείωση της ανάλυσης των εικόνων στην python πραγματοποιείται με την βιβλιοθήκη OpenCV [36]. Υπολογίζεται ο μέσος όρος των εικονοστοιχείων που περιλαμβάνονται μέσα στην περιοχή του νέου εικονοστοιχείου. Για κάθε νέο εικονοστοιχείο στην εικόνα, ισχύει:

$$P' = \frac{1}{N} \sum_{i=0}^{N} P_i$$
 (3.2)

όπου:

- P': είναι η τιμή του νέου εικονοστοιχείου,
- *P<sub>i</sub>*: είναι η τιμή του κάθε παλιού εικονοστοιχείου που εμπίπτει στην περιοχή που αντιστοιχεί στο νέο εικονοστοιχείο
- Ν: είναι ο αριθμός των εικονοστοιχείων που εμπίπτουν σε αυτή την περιοχή.

Το μέσο μέγεθος των επεξεργασμένων εικόνων είναι 913,57 bytes.

Παράδειγμα μείωσης της ανάλυσης μιας επίγειας εικόνας του ουρανού φαίνεται στο Σχήμα 3.7.



Σχήμα 3.7: Μείωση της ανάλυσης της εικόνας

#### 3.4 ΟΡΓΑΝΩΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ

Η παρούσα διπλωματική εργασία πραγματεύεται την πρόβλεψη ακολουθίας εικόνων με σκοπό την βραχυπρόθεσμη πρόβλεψη της Φ/Β παραγωγής. Το σύνολο των δεδομένων θα πρέπει να οργανωθεί κατάλληλα, με σκοπό την αρτιότερη υλοποίηση του επιθυμητού μοντέλου.

#### 3.4.1 Ορισμός του συνόλου των εικόνων σε ακολουθίες

Όπως αναλύθηκε στο Κεφάλαιο 2, τα CAE χρησιμοποιούνται για ανακατασκευή εικόνων. Στην παρούσα διπλωματική εργασία, όμως, σκοπός είναι η πρόβλεψη εικόνας. Έτσι, επειδή για δεδομένη είσοδο παράγεται διαφορετική έξοδος, το εγχείρημα εμπίπτει στην περίπτωση της επιτηρούμενης εκπαίδευσης. Σε αυτήν την περίπτωση, θα πρέπει τα δεδομένα να οργανωθούν σε ζεύγη εισόδου-εξόδου. Η πιο απλή περίπτωση είναι η είσοδος να περιλαμβάνει την τρέχουσα εικόνα, και η έξοδος την εικόνα που αντιστοιχεί στον ορίζοντα πρόβλεψης. Ωστόσο, είναι γνωστό από τη βιβλιογραφία [6] πως η χρήση ακολουθίας εικόνων στην είσοδο έχει σημαντικά οφέλη, γιατί προσφέρει χρονική πληροφορία σχετικά με την κίνηση των νεφών. Παράλληλα, στις περισσότερες εφαρμογές, η πρόβλεψη Φ/Β παραγωγής απαιτεί πρόβλεψη πολλαπλών βημάτων (multi-step ahead forecasting), δηλαδή κάθε χρονική στιγμή να παράγεται μια ακολουθία προβλέψεων που εκτείνεται μέχρι τον ορίζοντα πρόβλεψης. Επομένως, είναι ωφέλιμο τόσο η είσοδος, όσο και η έξοδος να είναι ακολουθίες. Η μαθηματική μοντελοποίηση των παραπάνω φαίνεται στην Εξίσωση (3.3):

 $[SI_{t_0+t_{out}}, \dots, SI_{t_0+(h-1)t_{out}}, SI_{t_0+ht_{out}}] = f_{CAE} ([SI_{t_0+(n-1)t_{in}}, \dots, SI_{t_0-t_{in}}, SI_{t_0}])$ (3.3)

όπου:

- SI: Εικόνα του ουρανού
- tout: Η χρονική ανάλυση της ακολουθίας εικόνων εξόδου
- t<sub>in</sub>: Η χρονική ανάλυση της ακολουθίας εικόνων εισόδου
- t<sub>0</sub>: Η χρονική στιγμή παραγωγής της πρόβλεψης
- f<sub>CAE</sub>: Το μοντέλο συνελικτικού αυτοκωδικοποιητή
- *n*: Μήκος ακολουθίας εισόδου
- h: Ο χρονικός ορίζοντας της πρόβλεψης (σε πλήθος βημάτων)

Στην παρούσα εργασία θα εξεταστεί η περίπτωση με 10 εικόνες εισόδου και 10 εικόνες εξόδου με χρονική ανάλυση το 1 λεπτό (άρα  $t_{out} = t_{in} = 1$  min, n = h = 10). Ο συγκεκριμένος συνδυασμός επιλέχθηκε μετά από μελέτη της βιβλιογραφίας σχετικής με την εκπαίδευση μοντέλων πρόβλεψης εικόνας του ουρανού με χρήση CAE [6].

Έτσι, η διαδικασία του ορισμού των ακολουθιών είναι η ακόλουθη:

- Κάθε μια από τις 14 ημέρες αρχειοθετείται ξεχωριστά (αφού οι ακολουθίες πρέπει να περιλαμβάνουν διαδοχικές εικόνες).
- Στην συνέχεια, επιλέγονται ακολουθίες (εικοσάδες) διαδοχικών εικόνων με βήμα 1 σε όλο το διάστημα της ημέρας (αν μια μέρα έχει N εικόνες, τότε θα μπορούν να εξαχθούν N-19 ακολουθίες). Από αυτές, οι 10 πρώτες θα έχουν ρόλο εισόδου και οι 10 επόμενες ρόλο εξόδου.



Σχήμα 3.8: Σε κάθε 20 διαδοχικές εικόνες ορίζονται οι 10 πρώτες ως είσοδος και οι 10 επόμενες ως έξοδος

Ημέρα	Πρώτη Λήψη	Τελευταία λήψη	Σύνολο φωτογραφιών ανά ημέρα	Σύνολο ακολουθιών ανά ημέρα	Μέσο Μ Εικόνας αν (KI	έγεθος νά Ημέρα 3)
					Αρχικά	Τελικά
16/11/2023	16:12	18:41	151	132	76.69	0.919
17/11/2023	07:37	18:40	664	645	76.24	0.907
18/11/2023	07:38	18:40	663	644	80.52	0.932
19/11/2023	07:39	18:39	661	642	69.26	0.886
20/11/2023	07:40	18:38	659	640	67.19	0.883
21/11/2023	07:41	18:38	658	639	75.55	0.918
22/11/2023	07:42	18:37	656	637	71.71	0.896
23/11/2023	07:43	18:37	655	636	81.21	0.941
24/11/2023	07:44	18:36	653	634	68.50	0.905
25/11/2023	07:45	18:36	652	633	86.97	0.935
26/11/2023	07:46	18:36	651	632	86.20	0.944
27/11/2023	07:47	18:35	649	630	69.62	0.892
28/11/2023	07:48	18:35	648	629	73.92	0.914
29/11/2023	07:49	10:17	148	129	87.70	0.918
		Σύνολο	8168	7902		

Πίνακας 3.1: Χαρακτηριστικά δεδομένων εισόδου.

Το Σχήμα 3.8 παρουσιάζει ένα παράδειγμα ακολουθίας είκοσι διαδοχικών εικόνων με τις δέκα πρώτες να αναπαριστούν την είσοδο και τις δέκα τελευταίες την έξοδο. Ο Πίνακας 3.1 παρουσιάζει αναλυτικά τα χαρακτηριστικά των δεδομένων εισόδου του μοντέλου.

### 3.4.2 Διαμοιρασμός των ακολουθιών σε σύνολα εκπαίδευσης, επικύρωσης και αξιολόγησης

Η διαδικασία της υλοποίησης του μοντέλου πρόβλεψης με χρήση CAE περιλαμβάνει τρία βασικά στάδια (εκπαίδευση, επικύρωση και δοκιμή), όπως αναφέρθηκε στην ενότητα 2.4.4.2. Αντίστοιχα, το σύνολο των ακολουθιών χωρίζεται σε 3 μέρη: εκπαίδευσης (train set), επικύρωσης (validation set) και δοκιμής (test set), σε ποσοστά 70%, 15%, και 15%, αντίστοιχα. Τα παραπάνω ποσοστά επιλέχθηκαν βάσει βιβλιογραφίας, σχετικής με εκπαίδευση μοντέλων CAE [6]. Ο Πίνακας 3.2 παρουσιάζει το πλήθος ακολουθιών που αντιστοιχούν σε κάθε σύνολο.

Σύνολο	Ποσοστό (%)	Πλήθος ακολουθιών	
Εκπαίδευσης	70	5532	
Επικύρωσης	15	1185	
Δοκιμής	15	1185	

Πίνακας 3.2: Διαμοιρασμός ακολουθιών σε σύνολα εκπαίδευσης, επικύρωσης, και δοκιμής

Ο διαμοιρασμός των ακολουθιών στα σύνολα εκπαίδευσης, επικύρωσης, και δοκιμής γίνεται με τυχαίο τρόπο, διότι είναι σκόπιμο να μειωθεί η αρνητική επίδραση που μπορεί να έχει η ανισορροπία του συνόλου δεδομένων. Για παράδειγμα, αν ο διαμοιρασμός γινόταν σειριακά, θα ήταν πολύ πιθανό η εκπαίδευση να περιλάμβανε κυρίως ηλιοφάνεια και η δοκιμή συννεφιά ή το αντίστροφο. Και στις δύο περιπτώσεις το CAE δεν θα μπορούσε να γενικεύσει με ικανοποιητική απόδοση. Για τον τυχαίο διαμοιρασμό των ακολουθιών στα σετ χρησιμοποιήθηκε η συνάρτηση train\_test\_split της βιβλιοθήκης scikit-learn της python [37].

# ΚΕΦΑΛΑΙΟ 4

# ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΥΛΟΠΟΙΗΣΗ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΤΟΥ ΜΟΝΤΕΛΟΥ

#### 4.1 ΠΕΡΙΓΡΑΦΗ ΜΟΝΤΕΛΟΥ ΠΡΟΒΛΕΨΗΣ ΑΚΟΛΟΥΘΙΑΣ ΕΙΚΟΝΩΝ

Το μοντέλο πρόβλεψης ακολουθίας εικόνων του ουρανού της παρούσας διπλωματικής εργασίας υλοποιήθηκε με χρήση συνελικτικών αυτοκωδικοποιητών. Όπως αναφέρθηκε στην Ενότητα 2.4.4, ένας συνελικτικός αυτοκωδικοποιητής περιλαμβάνει δύο βασικά νευρωνικά υποσυστήματα. Το πρώτο είναι ο κωδικοποιητής, ο οποίος λαμβάνει τα δεδομένα εισόδου και τα μετασχηματίζει σε μια κωδικοποιημένη αναπαράσταση στο επίπεδο στένωσης, χρησιμοποιώντας συνελικτικά στρώματα (convolutional layers). Το δεύτερο είναι ο αποκωδικοποιητής ο οποίος λαμβάνει την κωδικοποιημένη αναπαράσταση και δίνει τα δεδομένα εξόδου, χρησιμοποιώντας ανεστραμμένα συνελικτικά στρώματα (transposed convolutional layers).

Οι συνελικτικοί αυτοκωδικοποιητές χρησιμοποιούν ορισμένες τεχνικές που βοηθούν στην εξαγωγή του κατάλληλου αποτελέσματος. Παραδείγματα τέτοιων τεχνικών είναι ο αλγόριθμος βελτιστοποίησης (Ενότητα 2.4.3.4), η κανονικοποίηση κατά παρτίδες (Ενότητα 2.4.4.3) και η συνάρτηση ενεργοποίησης (Ενότητα 2.4.3.5). Παρακάτω παρουσιάζονται οι λεπτομέρειες των τεχνικών των μοντέλων συνελικτικών αυτοκωδικοποιητών που αναπτύχθηκαν στην παρούσα διπλωματική εργασία με σκοπό την πρόβλεψη ακολουθίας εικόνων του ουρανού.

#### Συνάρτηση ενεργοποίησης

Στην παρούσα διπλωματική εργασία εφαρμόζεται η συνάρτηση Leaky ReLU. Η συνάρτηση Leaky ReLU αποτελεί βελτιωμένη εκδοχή της συνάρτησης ReLU (Ενότητα 2.4.3.5), η οποία για τις αρνητικές τιμές του x επιστρέφει ένα πολύ μικρό γραμμικό πολλαπλάσιο του x αντί για μηδέν. Αυτό βοηθά στην αποφυγή του προβλήματος των "νεκρών νευρώνων", όπου πολλαπλοί νευρώνες με συνάρτηση ενεργοποίησης ReLU παραμένουν ανενεργοί [38].

Μαθηματικά, η Leaky ReLU ορίζεται ως:

$$f(x) = \begin{cases} x, & x \ge 0\\ \alpha x, & x < 0 \end{cases}$$
(4.1)

και απεικονίζεται στο Σχήμα 4.1 για την τιμή που χρησιμοποιήθηκε στην παρούσα εργασία  $\alpha = 0.2$ .



Σχήμα 4.1: Γραφική παράσταση της συνάρτησης ενεργοποίησης Leaky ReLu

#### • Αλγόριθμος βελτιστοποίησης

Επιλέχθηκε ο βελτιστοποιητής Adam με ρυθμό εκμάθησης ίσο με 0.001.

Η κανονικοποίηση κατά παρτίδες και η συνάρτηση ενεργοποίησης εφαρμόζονται μετά από κάθε συνελικτικό επίπεδο στην έξοδο των κρυφών στρωμάτων.

Στο Σχήμα 4.2 παρουσιάζονται σχηματικά οι υπολογισμοί που συμβαίνουν σε κάθε στρώμα.



Σχήμα 4.2: Υπολογισμοί που συμβαίνουν σε κάθε στρώμα.

## 4.2 ΠΕΡΙΓΡΑΦΗ ΕΞΕΤΑΖΟΜΕΝΩΝ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΜΟΝΤΕΛΩΝ ΣΥΝΕΛΙΚΤΙΚΩΝ ΑΥΤΟΚΩΔΙΚΟΠΟΙΗΤΩΝ

Στο πλαίσιο της παρούσας διπλωματικής εργασίας εξετάστηκαν αρχιτεκτονικές 2D και 3D συνελίξεων, τριών και πέντε επιπέδων ( $\{2D, 3D\} \times \{3L, 5L\}$ ).

#### 4.3.1 Αρχιτεκτονικές αυτοκωδικοποιητών 2D και 3D συνελίξεων

Όπως αναφέρθηκε στην Ενότητα 2.4.4.1, οι υλοποιημένες αρχιτεκτονικές αυτοκωδικοποιητών χρησιμοποιούν 2D και 3D συνελίξεις, δηλαδή η συνέλιξη εφαρμόζεται στην τρισδιάστατη πληροφορία (μήκος × πλάτος × μέγεθος ακολουθίας) της εισόδου και για κάθε φίλτρο, επιστρέφει έναν τένσορα δύο και τριών διαστάσεων, αντίστοιχα.

#### 4.2.1.1 Τριών επιπέδων

Στην περίπτωση των τριών επιπέδων, ο κωδικοποιητής, το επίπεδο στένωσης, και ο αποκωδικοποιητής περιέχουν από ένα στρώμα και οι μεταβολές που υφίσταται η ακολουθία εισόδου μετά από κάθε επίπεδο, είναι οι εξής:

#### 1) Κωδικοποιητής:

- Οι διαστάσεις της κάθε εικόνας μειώνονται από 32×32 σε 16×16
- Χρησιμοποιούνται 64 φίλτρα
- Strides = 2

#### 2) Επίπεδο στένωσης:

- Οι διαστάσεις της κάθε εικόνας παραμένουν 16×16
- Χρησιμοποιούνται 96 φίλτρα
- Strides = 1

#### 3) Αποκωδικοποιητής:

- Οι διαστάσεις της κάθε εικόνας αυξάνονται από 16×16 σε 32×32
- Χρησιμοποιούνται 64 φίλτρα
- Strides = 2

Σε όλα τα στρώματα, τα φίλτρα που χρησιμοποιήθηκαν ήταν μεγέθους 3×3 και επιλέχθηκε padding='same', το οποίο προσθέτει αρκετά μηδενικά ώστε η έξοδος να μην μικρύνει παραπάνω από το επιθυμητό.

Στο Σχήμα 4.3 φαίνεται σχηματικά η αρχιτεκτονική 2D-3L και 3D-3L. Το τμήμα EN αντιστοιχεί στον κωδικοποιητή (ENCODER), το τμήμα BL στο επίπεδο στένωσης (Bottleneck Layer) και το τμήμα DE στον αποκωδικοποιητή (DECODER). Στο πάνω μέρος του κάθε στρώματος αναγράφεται η διάσταση του αποτελέσματος της συνέλιξης του συγκεκριμένου στρώματος, ενώ στο κάτω μέρος το πλήθος των φίλτρων που χρησιμοποιήθηκαν.



Σχήμα 4.3: Αρχιτεκτονικές 3 στρωμάτων

#### 4.2.1.2 Πέντε επιπέδων

Στην περίπτωση των πέντε επιπέδων, ο κωδικοποιητής και ο αποκωδικοποιητής περιέχουν δύο στρώματα, ενώ το επίπεδο στένωσης ένα. Οι μεταβολές που υφίσταται η ακολουθία εισόδου μετά από κάθε στρώμα, είναι οι εξής:

#### 1) Κωδικοποιητής:

#### 1° Επίπεδο:

- Οι διαστάσεις της κάθε εικόνας παραμένουν 32×32
- Χρησιμοποιούνται 64 φίλτρα
- Strides = 1

#### 2° Επίπεδο:

- Οι διαστάσεις της κάθε εικόνας μειώνονται από 32×32 σε 16×16
- Χρησιμοποιούνται 96 φίλτρα
- Strides = 2

#### 2) Επίπεδο στένωσης:

- Οι διαστάσεις της κάθε εικόνας παραμένουν 16×16
- Χρησιμοποιούνται 128 φίλτρα
- Strides = 1

#### 3) <u>Αποκωδικοποιητής:</u>

#### 1° Επίπεδο:

- Οι διαστάσεις της κάθε εικόνας αυξάνονται από 16×16 σε 32×32

- Χρησιμοποιούνται 96 φίλτρα
- Strides = 2

#### **2° Επίπεδο:**

- Οι διαστάσεις της κάθε εικόνας παραμένουν 32×32
- Χρησιμοποιούνται 64 φίλτρα
- Strides = 1

Σε όλα τα στρώματα, τα φίλτρα που χρησιμοποιήθηκαν ήταν μεγέθους 3×3×3 και επιλέχθηκε padding='same', το οποίο προσθέτει αρκετά μηδενικά ώστε η έξοδος να μην μικρύνει παραπάνω από το επιθυμητό.

Στο Σχήμα 4.4 φαίνεται σχηματικά οι αρχιτεκτονικές 2D-5L και 3D-5L. Αντίστοιχα με το Σχήμα 4.3, το τμήμα EN αντιστοιχεί στον κωδικοποιητή (ENCODER), το τμήμα BL στο επίπεδο στένωσης (Bottleneck Layer) και το τμήμα DE στον αποκωδικοποιητή (DECODER). Στο πάνω μέρος του κάθε στρώματος αναγράφεται η διάσταση του αποτελέσματος της συνέλιξης του συγκεκριμένου στρώματος, ενώ στο κάτω μέρος το πλήθος των φίλτρων που χρησιμοποιήθηκαν.



Σχήμα 4.4: Αρχιτεκτονικές 5 στρωμάτων

#### 4.3 Περιβάλλον υλοποίησης του μοντέλου

Όλα τα μοντέλα υλοποιήθηκαν στη γλώσσα προγραμματισμού python, στο περιβάλλον της spyder, με την βοήθεια της βιβλιοθήκης tensorflow – keras [39].

Το υπολογιστικό σύστημα που χρησιμοποιήθηκε ήταν 11th Gen Intel (R) Core (TM) i5-1135G7 @ 2.40GHz 2.42 GHz, με 8,00 GB RAM.

#### 4.4 Εκπαίδευση των μοντέλων

Η αντικειμενική συνάρτηση στο πρόβλημα πρόβλεψης εικόνων είναι η ελαχιστοποίηση του σφάλματος μεταξύ της πραγματικής και της προβλεπόμενης εικόνας.

Κατά την εκτέλεση του προγράμματος, πραγματοποιείται η εκπαίδευση των μοντέλων συνελικτικών αυτοκωδικοποιητών για την εξαγωγή της σχέσης εισόδου-εξόδου. Αυτή η διαδικασία λαμβάνει χώρα μέσα από επαναληπτικές προσπελάσεις όλων των δεδομένων του συνόλου εκπαίδευσης (epochs). Στο τέλος του κάθε epoch γίνεται η αξιολόγηση από το σύνολο

επικύρωσης και καταγράφεται η μέση απόκλιση των ακολουθιών που προέβλεψε το μοντέλο, από τις επιθυμητές ακολουθίες (σφάλμα επικύρωσης).

Το πρόγραμμα περιλαμβάνει τον μηχανισμό πρώιμης διακοπής της εκπαίδευσης (Early Stopping) που διακόπτει την εκπαίδευση αν δεν παρατηρείται βελτίωση στο σφάλμα επικύρωσης για N (παράμετρος patience) συνεχόμενα epochs. Στην παρούσα διπλωματική εργασία, οι παράμετροι επιλέχθηκαν οι τιμές: epochs = 100 και patience = 10, εμπειρικά μετά από σειρά δοκιμών, με στόχο τη διατήρηση της ακρίβειας σε συνδυασμό με τον περιορισμό του χρόνου εκπαίδευσης και της υπολογιστικής επιβάρυνσης.

Στο Σχήμα 4.5 φαίνονται διαγραμματικά οι διαδικασίες που επιτελεί το τελικό πρόγραμμα.



Σχήμα 4.2: Δομή του προγράμματος

# ΚΕΦΑΛΑΙΟ 5

# ΑΠΟΤΕΛΕΣΜΑΤΑ

#### 5.1 ΜΕΤΡΗΤΙΚΑ ΑΞΙΟΛΟΓΗΣΗΣ

Για την αξιολόγηση της απόδοσης και την εκτίμηση του σφάλματος ενός μοντέλου πρόβλεψης εικόνων, απαιτείται σύγκριση μεταξύ των εικόνων που προβλέφθηκαν και των επιθυμητών εικόνων. Η σύγκριση αυτή επιτυγχάνεται με ποσοτικούς δείκτες αξιολόγησης (μετρητικά αξιολόγησης). Στην περίπτωση των εικόνων, συχνά χρησιμοποιούνται οι δείκτες μέσου τετραγωνικού σφάλματος (Mean Squared Error – MSE) και μέτρησης δομικής ομοιότητας (Structural Similarity Index Measure – SSIM) [6].

#### 5.1.1 MSE

Πριν την είσοδό τους στο μοντέλο, οι εικόνες έχουν απεικονιστεί στην κλίμακα του γκρι και τα εικονοστοιχεία τους έχουν αποκτήσει μια κανονικοποιημένη τιμή στο διάστημα [0,1] (το 0 αντιστοιχεί στο μαύρο και το 1 στο άσπρο). Το MSE συγκρίνει τις δύο εικόνες εικονοστοιχείο προς εικονοστοιχείο ως προς την τιμή που έχουν. Παρακάτω αναλύεται η μαθηματική σχέση που εκφράζει το MSE [40]:

$$MSE(x,y) = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [x(i,j) - y(i,j)]^2$$
(5.1)

όπου:

- x: η εικόνα που προβλέφθηκε
- y: η εικόνα που αναμενόταν
- (*i*, *j*): συντεταγμένες του εικονοστοιχείου
- *M*, *N*: οι διαστάσεις της εικόνας (στην παρούσα εργασία M = N = 32)

Πριν την είσοδό τους στο μοντέλο, οι εικόνες έχουν απεικονιστεί στην κλίμακα του γκρι και τα εικονοστοιχεία Στην συνέχεια μπορεί να οριστεί και το μέσο σφάλμα ακολουθίας, ως ο μέσος όρος των MSEs και των δέκα εικόνων.

Μέσο Σφάλμα Ακολουθίας = 
$$\frac{1}{10} \sum_{\nu=0}^{10} MSE_{\nu}$$
 (5.2)

• *MSE<sub>v</sub>*: το MSE της ν-οστής εικόνας της ακολουθίας

Όσο πιο μικρό είναι το MSE, τόσο πιο όμοιες είναι δύο εικόνες.

#### 5.1.2 SSIM

Ο SSIM δηλώνει το μέγεθος της ομοιότητας ανάμεσα σε δύο εικόνες. Εξαρτάται από τους εξής τρεις παράγοντες [41]:

- Φωτεινότητα (Luminance): Μέτρηση της φωτεινότητας των δύο εικόνων.
- Αντίθεση (Contrast): Σύγκριση της αντίθεσης (διαφορά μεταξύ των φωτεινών και σκοτεινών περιοχών στην ίδια την εικόνα) ανάμεσα στις δύο εικόνες.
- Δομή (Structure): Εκτίμηση της χωρικής διάταξης των φωτεινών στοιχείων στις εικόνες.

Η μαθηματική έκφραση του δείκτη SSIM ανάμεσα σε δύο εικόνες παρουσιάζεται μέσω των σχέσεων (5.3) έως (5.7):

$$SSIM(x,y) = [l(x,y)]^{\alpha} [c(x,y)]^{\beta} [s(x,y)]^{\gamma}, SSIM \in [0,1]$$
(5.3)

$$l(x, y) = \frac{2 \,\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{5.4}$$

$$c(x, y) = \frac{2 \sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$
(5.5)

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}$$
(5.6)

$$C_1 = (K_1 L)^2, C_2 = (K_2 L)^2, C_3 = C_2/2$$
 (5.7)

όπου:

- x: η εικόνα που προβλέφθηκε
- y: η εικόνα που αναμενόταν
- *l(x, y)*: συντελεστής φωτεινότητας
- c(x, y): συντελεστής αντίθεσης
- *s*(*x*, *y*): συντελεστής δομής
- α, β, γ > 0: ρυθμίζουν την βαρύτητα των συντελεστών (στην παρούσα εργασία επιλέχτηκε α = β = γ = 1)
- $\mu_x, \mu_y$ : μέση τιμή των εικονοστοιχείων των εικόνων x και y
- $\sigma_x, \sigma_y$ : τυπική απόκλιση των εικονοστοιχείων των εικόνων x και y
- $\sigma_{xy}$ : συνδιακύμανση των εικονοστοιχείων των εικόνων x και y
- C<sub>1</sub>, C<sub>2</sub>, C<sub>3</sub>, : σταθερές που προστίθενται ώστε ο παρονομαστής των συντελεστών να μην πλησιάζει την τιμή μηδέν.

Επιλέχθηκαν οι τιμές  $K_1 = 0.001, K_2 = 0.003$  και L = 255.

Όσο πιο μεγάλο είναι το SSIM τόσο πιο όμοιες είναι δύο εικόνες [6], [42].

#### 5.2 ΜΕΘΟΔΟΣ ΠΡΟΒΛΕΨΗΣ PERSISTENCE

Η μέθοδος persistence είναι μια μέθοδος πρόβλεψης η οποία υποθέτει ότι οι μελλοντικές τιμές μιας μεταβλητής θα είναι ίσες με τις τωρινές (ευκολότερο δυνατό σενάριο).

$$\hat{P}_{t+k} = P_t, \qquad k = 10$$
 (5.7)

όπου:

- $P_t$ : Η τιμή της τυχαίας μεταβλητής την στιγμή t
- $\hat{P}_{t+k}$ : Η πρόβλεψη για την στιγμή t+k

Στην περίπτωση της παρούσας διπλωματικής θα είναι k = 10, αφού γίνονται δέκα προβλέψεις για δέκα εικόνες εισόδου. Η συγκεκριμένη μέθοδος χρησιμοποιείται κυρίως ως σημείο αναφοράς, αφού αποτελεί το ευκολότερο δυνατό σενάριο για μια πρόβλεψη.

### 5.3 ΠΡΟΒΛΕΨΗ ΑΚΟΛΟΥΘΙΑΣ ΕΙΚΟΝΩΝ ΤΟΥ ΟΥΡΑΝΟΥ ΧΩΡΙΣ ΤΑΞΙΝΟΜΗΣΗ ΕΙΣΟΔΟΥ

Αρχικά το μοντέλο εκπαιδεύτηκε στο σύνολο των δεδομένων (7902 ακολουθίες των 20 εικόνων). Η αξιολόγηση των αποτελεσμάτων έγινε ποιοτικά (εποπτική αξιολόγηση), αλλά και ποσοτικά (μετρητικά αξιολόγησης).

#### 5.3.1 Ποιοτική αξιολόγηση αποτελεσμάτων

Στα Σχήματα 5.1, 5.3, 5.5 και 5.7 φαίνεται το αποτέλεσμα της πρόβλεψης που εξήγαγε το μοντέλο σε κατάσταση ηλιοφάνειας, ενώ στα Σχήματα 5.2, 5.4, 5.6 και 5.8 σε κατάσταση συννεφιάς για τις τέσσερις αρχιτεκτονικές ({2D,3D}x{3L,5L}). Σε κάθε Σχήμα διακρίνονται δύο δεκάδες εικόνων. Για την διευκόλυνση της εποπτικής αξιολόγησης, παρατίθεται η δεκάδα φωτογραφιών που έπρεπε να προβλέψει το μοντέλο (Πραγματική) και στην συνέχεια η δεκάδα που προέβλεψε (Πρόβλεψη).



Σχήμα 5.1: Πρόβλεψη μοντέλου 2D-3L σε κατάσταση ηλιοφάνειας







Σχήμα 5.8: Πρόβλεψη μοντέλου 3D-5L σε κατάσταση συννεφιάς

Από την εποπτική αξιολόγηση προκύπτει ότι το μοντέλο ανταποκρίνεται επαρκώς στην κατάσταση ηλιοφάνειας, καθώς εντοπίζει την θέση του ηλίου και σε αρκετά ικανοποιητικό βαθμό την φωτεινότητα των εικόνων. Στην περίπτωση της συννεφιάς, ωστόσο, ενώ εντοπίζει την θέση του ηλίου, αδυνατεί να προβλέψει την κατανομή των νεφών.

#### 5.3.2 Ποσοτική αξιολόγηση αποτελεσμάτων

Στον Πίνακα 5.1 παρουσιάζονται οι επιδόσεις των τεσσάρων αρχιτεκτονικών μοντέλου που υλοποιήθηκαν, ως προς την ικανότητα πρόβλεψης και της βελτίωσης που επιτυγχάνουν σε σχέση με την μέθοδο persistence.

		MSE	SSIM		
Μοντέλο	MSE (p.u.)	Βελτίωση σε σχέση με persistence (%)	SSIM (p.u.)	Βελτίωση σε σχέση με persistence (%)	
Persistence	0.002	0	0.761	0	
2D-3L	0.0014	30	0.7987	4.95	
2D-5L	0.0014	30	0.8018	5.36	
3D-3L	0.0017	15	0.7873	3.46	
3D-5L	0.0016	20	0.7969	4.72	

Πίνακας 5.1: Αποτελέσματα μοντέλου στο σύνολο των δεδομένων για τις τέσσερις αρχιτεκτονικές

Είναι εμφανές ότι και οι τέσσερις αρχιτεκτονικές υπερτερούν σε σχέση με το persistence, με τις καλύτερες επιδόσεις να ανήκουν στην αρχιτεκτονική 2D-5L.

Τα μοντέλα CAE πετυχαίνουν βελτίωση του μέσου MSE που εκτείνεται από 15% έως 30% και αύξηση του μέσου SSIM από 3.46% έως 5.36%.

Επίσης, από τις παραπάνω μεθόδους αξιολόγησης συνάγεται το συμπέρασμα ότι σε περιπτώσεις εικόνων δεν αρκεί η χρήση ποσοτικών δεικτών, αλλά είναι αναγκαία και η παράθεση εικόνων εξόδου. Ενώ τα μετρητικά αξιολόγησης δηλώνουν ικανοποιητική ομοιότητα μεταξύ των εικόνων, η εποπτική αξιολόγηση καθιστά επιτακτική την ανάγκη για καλύτερα αποτελέσματα.

## 5.4 ΜΕΛΕΤΗ ΠΕΡΙΠΤΩΣΗΣ ΣΕ ΣΕΤ ΔΕΔΟΜΕΝΩΝ ΧΑΜΗΛΗΣ ΜΕΤΑΒΛΗΤΟΤΗΤΑΣ

Όπως φαίνεται από την Ενότητα 5.3, τα αποτελέσματα του μοντέλου στην περίπτωση της εκπαίδευσης στο σύνολο των δεδομένων δεν ήταν ικανοποιητικά. Προς την κατεύθυνση της επίλυσης του προβλήματος, πραγματοποιήθηκε εξέταση ενδεχόμενης αδυναμίας ανταπόκρισης του μοντέλου σε δεδομένα με μεγάλη ανομοιομορφία. Για αυτό, επιλέχθηκε μικρό δείγμα από το σύνολο των δεδομένων, με ακολουθίες παρόμοιων καταστάσεων συννεφιάς (αφού εκεί εντοπίστηκε το πρόβλημα).

Στο Σχήμα 5.9 φαίνεται αποτέλεσμα της παραπάνω δοκιμής, στην περίπτωση με:

- Μήκος ακολουθίας = 20 εικόνες και
- Αριθμός εικόνων εισόδου = 10
- Αριθμός εικόνων εξόδου = 10

Στο Σχήμα 5.10 φαίνονται πέντε αποτελέσματα της παραπάνω δοκιμής, στην περίπτωση με:

- Μήκος ακολουθίας = 2 εικόνες
- Αριθμός εικόνων εισόδου = 1
- Αριθμός εικόνων εξόδου = 1 (εξέταση σε ακόμα ευκολότερο πρόβλημα).



Σχήμα 5.10: Πρόβλεψη με εκπαίδευση σε μικρό δείγμα δεδομένων

Όπως φαίνεται από τα Σχήματα 5.9 και 5.10, το μοντέλο ανταποκρίθηκε και σε δύσκολα μοτίβα όταν εκπαιδεύτηκε με πιο ομοιόμορφα δεδομένα. Η εκπαίδευση έγινε με την αρχιτεκτονική 2D-5L, αφού σύμφωνα με τον Πίνακα 5.1 έχει τις καλύτερες επιδόσεις.

## 5.5 ΠΡΟΒΛΕΨΗ ΑΚΟΛΟΥΘΙΑΣ ΕΙΚΟΝΩΝ ΤΟΥ ΟΥΡΑΝΟΥ ΜΕ ΤΑΞΙΝΟΜΗΣΗ ΕΙΣΟΔΟΥ

Με αφορμή τα εμφανώς βελτιωμένα αποτελέσματα της Ενότητας 5.4, πραγματοποιήθηκε η τεχνική της ταξινόμησης των δεδομένων εισόδου. Με αυτήν την τεχνική, το σύνολο των δεδομένων οργανώνεται σε υποομάδες (κλάσεις), οι οποίες περιέχουν ακολουθίες παρόμοιων καταστάσεων του ουρανού (και άρα περισσότερο ομοιόμορφες εικόνες). Στην συνέχεια, εκπαιδεύονται και αποθηκεύονται μοντέλα, ένα για κάθε κλάση. Έτσι, όταν θα πρέπει να γίνει η πρόβλεψη μίας ακολουθίας εικόνων, θα καλείται το κατάλληλο εκπαιδευμένο μοντέλο και θα εφαρμόζει την πρόβλεψη. Στο Σχήμα 5.11 παρουσιάζεται διαγραμματικά η διαδικασία πρόβλεψης με ταξινόμηση εισόδου.



Σχήμα 5.11: Διαδικασία πρόβλεψης με ταξινόμηση εισόδου

#### 5.5.1 Μέθοδος ταξινόμησης εικόνων του ουρανού

Μια από τις μεθόδους ταξινόμησης εικόνων του ουρανού σε κλάσεις ανάλογα με την κατάσταση του ουρανού (πχ ηλιοφάνεια, μερική συννεφιά και συννεφιά), είναι η μέθοδος R-B [7]. Σύμφωνα με αυτήν, αφαιρούνται οι τιμές του B-καναλιού (μπλε κανάλι) της έγχρωμης εικόνας από αυτές του R-καναλιού (κόκκινο κανάλι), εικονοστοιχείο προς εικονοστοιχείο. Όταν σε ένα εικονοστοιχείο ισχύει ότι R~B, τότε το χρώμα του είναι συνήθως κάποια γκρίζα απόχρωση, άρα το εικονοστοιχείο θεωρείται ότι αντιστοιχεί σε νέφος. Αντίστοιχα, όταν B>>R υπερτερεί το μπλε χρώμα (καθαρός ουρανός). Υπολογίζοντας τη μέση τιμή όλων των εικονοστοιχείων, λαμβάνεται μια εικόνα για την συνολική ποσόστωση συννεφιάς. Έτσι, όσο μικρότερος είναι ο μέσος όρος της R-B τιμής τόσο περισσότερη ηλιοφάνεια-καθαρός ουρανός παρατηρείται, ενώ όσο ο μέσος όρος αυξάνεται παρατηρείται περισσότερη συννεφιά. Οι τιμές του μέσου όρου R-B κυμαίνονται στο διάστημα [-148.52, -15.35]. Στα Σχήματα 5.12 έως 5.14 φαίνονται παραδείγματα εικόνων του ουρανού, συνοδευόμενα από τον μέσο όρο των R-B τιμών τους.



Σχήμα 5.12: Κατάσταση ηλιοφάνειας, R-B = -105



Σχήμα 5.13: Κατάσταση μερικής συννεφιάς, R-B = -92



**Σχήμα 5.14:** Κατάσταση συννεφιάς, R-B = -36

#### 5.5.2 Ταξινόμηση των δεδομένων σε δύο κλάσεις

Αρχικά, πραγματοποιήθηκε ταξινόμηση εισόδου με χρήση δύο κλάσεων (ηλιοφάνεια-συννεφιά). Η τιμή του κατωφλίου (threshold) ορίστηκε ίση με –91.9 με οπτική αξιολόγηση του συνόλου των εικόνων. Έτσι, όσες εικόνες έχουν μέσο όρο R-B  $\in$  [–148.52, –91.9] θα ανήκουν στην κατάσταση ηλιοφάνειας (1<sup>η</sup> κλάση με 2445 ακολουθίες), ενώ όσες έχουν μέσο όρο R-B  $\in$  [–91.9, –15.35] θα ανήκουν στην κατάσταση συννεφιάς (2<sup>η</sup> κλάση με 5457 ακολουθίες). Στο Σχήμα 5.15 απεικονίζεται σχηματικά η ταξινόμηση του συνόλου των δεδομένων σε δύο κλάσεις.



Σχήμα 5.15: Διαίρεση των δεδομένων σε 2 κλάσεις ανάλογα με τον μέσο όρο της R-B τιμής

#### 5.5.2.1 Ποιοτική αξιολόγηση αποτελεσμάτων

Η ταξινόμηση εφαρμόστηκε στην αρχιτεκτονική 2D-5L, που όπως δείχνει ο πίνακας 5.1, έχει τα καλύτερα αποτελέσματα. Στα Σχήματα 5.16 και 5.17 παρουσιάζονται τα αποτελέσματα στις καταστάσεις ηλιοφάνειας και συννεφιάς.



Σχήμα 5.17: Πρόβλεψη μοντέλου σε κατάσταση συννεφιάς

Από την εποπτική αξιολόγηση προκύπτει ότι το μοντέλο ανταποκρίνεται επαρκώς στην κατάσταση ηλιοφάνειας, καθώς εντοπίζει την θέση του ηλίου και σε αρκετά ικανοποιητικό βαθμό την φωτεινότητα των εικόνων. Στην περίπτωση της συννεφιάς, αδυνατεί να προβλέψει την κατανομή των νεφών επαρκώς, αλλά παρουσιάζει κάποια βελτίωση σε σχέση με την περίπτωση χωρίς ταξινόμηση των δεδομένων εισόδου.

#### 5.5.2.2 Ποσοτική αξιολόγηση αποτελεσμάτων

Στον Πίνακα 5.2 φαίνονται τα αποτελέσματα για κάθε κλάση ξεχωριστά, ενώ στον Πίνακα 5.3 συνολικά.

Μοντέλο		MSE (p.u.)	SSIM (p.u.)
Persistence		0.002	0.761
	Ηλιοφάνεια	0.0003	0.9464
2D-5L	Συννεφιά	0.0013	0.7763

Πίνακας 5.2: Αποτελέσματα του μοντέλου για τις 2 κλάσεις

	MSE		SSIM	
Μοντέλο	MSE (p.u.)	Βελτίωση σε σχέση με persistence (%)	SSIM (p.u.)	Βελτίωση σε σχέση με persistence (%)
Persistence	0.002	0	0.761	0
2D-5L (2 Κλάσεις)	0.00098	51	0.830732	9.16

Πίνακας 5.3: Αποτελέσματα του μοντέλου στην περίπτωση της ταξινόμησης εισόδου με 2 κλάσεις

Τα μετρητικά αξιολόγησης δείχνουν ότι τα αποτελέσματα του μοντέλου βελτιώθηκαν όταν υπήρξε εκπαίδευση σε ομοιόμορφα δεδομένα. Συγκεκριμένα, η τεχνική της ταξινόμησης των δεδομένων εισόδου με 2 κλάσεις πετυχαίνει βελτίωση του μέσου MSE κατά 51% και αύξηση του μέσου SSIM από 9.16%

#### 5.5.3 Ταξινόμηση των δεδομένων σε τρεις κλάσεις

Στη συνέχεια, πραγματοποιήθηκε ταξινόμηση εισόδου με χρήση τριών κλάσεων: ηλιοφάνεια, μερική συννεφιά, και συννεφιά. Σύμφωνα με την οπτική αξιολόγηση του συνόλου των εικόνων ορίστηκε ότι: όσες εικόνες έχουν μέσο όρο R-B  $\in$  [–148.52, –91.9] θα ανήκουν στην κατάσταση ηλιοφάνειας (1<sup>η</sup> κλάση με 2445 ακολουθίες), όσες έχουν μέσο όρο R-B  $\in$  [–91.9, –74] θα ανήκουν στην κατάσταση μερικής συννεφιάς (2<sup>η</sup> κλάση με 1704 ακολουθίες) και όσες έχουν μέσο όρο R-B  $\in$  [–74, –15.35] θα ανήκουν στην κατάσταση συννεφιάς (3<sup>η</sup> κλάση με 3753 ακολουθίες). Στο Σχήμα 5.18 απεικονίζεται σχηματικά η ταξινόμηση του συνόλου των δεδομένων σε τρεις κλάσεις.



Σχήμα 5.18: Διαίρεση των δεδομένων σε 3 κλάσεις ανάλογα με τον μέσο όρο της R-B τιμής

#### 5.5.3.1 Ποιοτική αξιολόγηση αποτελεσμάτων

Η ταξινόμηση εφαρμόστηκε στην αρχιτεκτονική 2D-5L, που όπως δείχνει ο Πίνακας 5.1, έχει τα καλύτερα αποτελέσματα. Στα Σχήματα 5.19. 5.20 και 5.21 παρουσιάζονται τα αποτελέσματα στις καταστάσεις ηλιοφάνειας, μερικής συννεφιάς και συννεφιάς.



Σχήμα 5.21: Πρόβλεψη μοντέλου σε κατάσταση συννεφιάς

Από την εποπτική αξιολόγηση προκύπτει ότι το μοντέλο ανταποκρίνεται επαρκώς στην κατάσταση ηλιοφάνειας, καθώς εντοπίζει την θέση του ηλίου και σε αρκετά ικανοποιητικό βαθμό την φωτεινότητα των εικόνων. Στην περίπτωση της συννεφιάς, αδυνατεί να προβλέψει την κατανομή των νεφών επαρκώς, αλλά παρουσιάζει κάποια βελτίωση σε σχέση με την περίπτωση χωρίς ταξινόμηση των δεδομένων εισόδου.

#### 5.5.3.2 Ποσοτική αξιολόγηση αποτελεσμάτων

Στον Πίνακα 5.4 φαίνονται τα αποτελέσματα για κάθε κλάση ξεχωριστά, ενώ στον Πίνακα 5.5 συνολικά.

Μοντέλο		MSE (p.u.)	SSIM (p.u.)
Persistence		0.002	0.761
	Ηλιοφάνεια	0.0003	0.9464
2D-5L	Μερική συννεφιά	0.0011	0.8416
	Συννεφιά	0.0013	0.777

Πίνακας 5.4: Αποτελέσματα του μοντέλου για τις 2 κλάσεις

Πίνακας 5.5: Αποτελέσματα του	υ μοντέλου στην	περίπτωση	της ταξινόμησης	εισόδου με 2
	κλάσεια			

		MSE	SSIM		
Μοντέλο	MSE (p.u.)	Βελτίωση σε σχέση με persistence (%)	SSIM (p.u.)	Βελτίωση σε σχέση με persistence (%)	
Persistence	0.002	0	0.761	0	
2D-5L (3 Κλάσεις)	0.00093	53.2	0.8454	11.09	

Τα μετρητικά αξιολόγησης δείχνουν ότι τα αποτελέσματα του μοντέλου βελτιώθηκαν ακόμα περισσότερο, όταν υπήρξε εκπαίδευση σε ομοιόμορφα δεδομένα. Συγκεκριμένα, η τεχνική της ταξινόμησης των δεδομένων εισόδου με 2 κλάσεις πετυχαίνει βελτίωση του μέσου MSE κατά 53.2% και αύξηση του μέσου SSIM κατά 11.09%.

#### 5.6 ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Στον πίνακα 5.6 βρίσκονται συγκεντρωμένα τα μετρητικά αξιολόγησης για την επίδοση του μοντέλου 2D-5L για τις καταστάσεις εκπαίδευσης:

- Χωρίς ταξινόμηση εισόδου
- Ταξινόμηση εισόδου με 2 κλάσεις
- Ταξινόμηση εισόδου με 3 κλάσεις

	Ν	ASE	SSIM		
Μοντέλο	MSE (p.u.)	Βελτίωση σε σχέση με persistence (%)	SSIM (p.u.)	Βελτίωση σε σχέση με persistence (%)	
Persistence	0.002	0	0.761	0	
2D-5L	0.0014	30	0.8018	5.36	
2D-5L (2 Κλάσεις)	0.00098	51	0.830732	9.16	
2D-5L (3 Κλάσεις)	0.00093	53.2	0.8454	11.09	

Πίνακας 5.6: Συγκεντρωτικά αποτελέσματα για την απόδοση του μοντέλου 2D-5L με και χωρίς ταξινόμηση εισόδου

Όπως φαίνεται από τον Πίνακα 5.6, όσο αυξάνεται το πλήθος των κλάσεων τόσο ο δείκτης MSE, όσο και ο SSIM φαίνονται να βελτιώνονται. Από την παραπάνω ενδελεχή έρευνα συνάγεται εύλογα το συμπέρασμα ότι το μοντέλο αποδίδει καλύτερα (καθίσταται ικανό στο να αναγνωρίζει μοτίβα) όταν εκπαιδεύεται σε περισσότερο ομοιόμορφες εικόνες.

#### 5.5 ΧΡΟΝΟΙ ΕΚΠΑΙΔΕΥΣΗΣ

Στον Πίνακα 5.7 παρουσιάζεται ο συνολικός χρόνος εκπαίδευσης του μοντέλου 2D-5L για τις καταστάσεις εκπαίδευσης:

- Χωρίς ταξινόμηση εισόδου (1 κλάση)
- Ταξινόμηση εισόδου με 2 κλάσεις
- Ταξινόμηση εισόδου με 3 κλάσεις

Πίνακας 5.7: Χρόνος εκπαίδευσης του μοντέλου 2D-5L με και χωρίς ταξινόμηση εισόδου

Πλήθος Κλάσεων	Χρόνος Εκτέλεσης Εκπαίδευσης
1	15 hrs 07 min 12 sec
2	13 hrs 31 min 42 sec
3	9 hrs 53 min 56 sec

Όπως φαίνεται από τον Πίνακα 5.7, με την αύξηση του πλήθους των κλάσεων ο συνολικός χρόνος εκτέλεσης μειώνεται σημαντικά. Με την προσθήκη δύο κλάσεων, ο συνολικός χρόνος εκπαίδευσης μειώθηκε κατά 10.52%, ενώ με την προσθήκη τριών κλάσεων, μειώθηκε κατά 34.53%.

Το γεγονός ότι ο συνολικός χρόνος εκπαίδευσης μειώνεται με την αύξηση των κλάσεων μπορεί να οφείλεται στους παρακάτω λόγους:

- Στην περίπτωση της εκπαίδευσης με ταξινόμηση εισόδου, το κάθε μοντέλο εκπαιδεύεται σε υποσύνολο δεδομένων και άρα χρειάζεται λιγότερος χρόνος ανά epoch.
- Η κάθε κλάση περιέχει πιο ομοιόμορφα δεδομένα (ευκολότερο πρόβλημα)
  και άρα ενεργοποιείται νωρίτερα η πρώιμη διακοπή της εκπαίδευσης.
- Επειδή οι πράξεις που γίνονται είναι μη γραμμικές, μπορεί σε κάποιες περιπτώσεις τα βάρη να αποκτάνε τέτοιες τιμές που ο χρόνος ανά υπολογιστική μονάδα να βγαίνει μεγαλύτερος στην περίπτωση χωρίς ταξινόμηση των δεδομένων εισόδου.
- Στον τρόπο που η βιβλιοθήκη tensorflow δεσμεύει χώρο μνήμης και κάνει προσπέλαση στα δεδομένα.

Η μείωση του συνολικού χρόνου εκπαίδευσης με την αύξηση του πλήθους των πράξεων συνιστά σημαντικό πλεονέκτημα, καθώς αποδεικνύει ότι το εγχείρημα για ταξινόμηση των δεδομένων εισόδου είναι εφαρμόσιμο και προσδίδει τόσο στην ακρίβεια των αποτελεσμάτων, όσο και στην αποδοτικότητα του μοντέλου.

# КЕФАЛАЮ 6

# ΣΥΜΠΕΡΑΣΜΑΤΑ

#### 6.1 ΣΥΝΟΨΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ ΤΗΣ ΕΡΓΑΣΙΑΣ

Στην παρούσα διπλωματική εργασία πραγματοποιήθηκε η υλοποίηση μοντέλου πρόβλεψης εικόνων του ουρανού με χρήση βαθιάς μάθησης για την βραχυπρόθεσμη πρόβλεψη της ΦΒ παραγωγής. Το εγχείρημα προσεγγίστηκε μέσα από μεθόδους πρόβλεψης από επίγειες εικόνες του ουρανού, αφού παρέχουν αρκετή πληροφορία αναφορικά με την σχετική θέση των νεφών και του ηλίου. Επιπλέον, προτιμήθηκε μέθοδος με χρήση βαθιάς μάθησης, καθώς έχει αποδειχθεί ικανή στην διαχείριση μεγάλου όγκου πληροφοριών και αποτελεί χρήσιμο εργαλείο στην εκπαίδευση μοντέλων και την εκμάθηση μη γραμμικών σχέσεων μεταξύ εισόδου και εξόδου.

Αρχικά, συγκεντρώθηκαν 8168 εικόνες του ουρανού σε διάστημα 14 ημερών με συχνότητα λήψεων 1 εικόνα ανά λεπτό, από επίγεια κάμερα ευρυγώνιου φακού με πεδίο θέασης 180°. Στην συνέχεια, το πλήθος των δεδομένων υποβλήθηκε σε κατάλληλη επεξεργασία (με απεικόνιση στην κλίμακα του γκρι και μείωση της ανάλυσης) και οργάνωση (ακολουθίες των 20 διαδοχικών εικόνων και χωρισμός σε σύνολα εκπαίδευσης, επικύρωσης και δοκιμής) πριν την είσοδό του στο μοντέλο.

Έπειτα, υλοποιήθηκαν οι τέσσερις αρχιτεκτονικές μοντέλου συνελικτικού αυτοκωδικοποιητή, με χρήση 2D και 3D συνελίξεων, για 3 και 5 στρώματα ({2D,3D}×{3L,5L}). Το κρίσιμο σημείο της υλοποίησης των δομικών μερών του αυτοκωδικοποιητή (κωδικοποιητής, επίπεδο στένωσης και αποκωδικοποιητής) φάνηκε να είναι ο προσδιορισμός των υπερπαραμέτρων (πχ., πλήθος φίλτρων, μέγεθος φίλτρων, strides, padding, συνάρτηση ενεργοποίησης, αλγόριθμος βελτιστοποίησης) που απαρτίζουν το μοντέλο.

Η ανάλυση των αποτελεσμάτων οδηγεί στα εξής συμπεράσματα:

- Οι συνελικτικοί αυτοκωδικοποιητές μειώνουν σημαντικά το μέσο σφάλμα σε σχέση με την μέθοδο persistence.
- Ανάμεσα στις τέσσερις αρχιτεκτονικές που δοκιμάστηκαν, βρέθηκε ότι καλύτερη είναι αυτή που χρησιμοποιεί 2D συνελίξεις των 5 στρωμάτων, σύμφωνα με τα μετρητικά αξιολόγησης.
- Η εκπαίδευση με χρήση της τεχνικής της ταξινόμησης των δεδομένων εισόδου παρουσιάζει εμφανώς βελτιωμένα αποτελέσματα, τόσο ποιοτικά

(μέσω της εποπτικής αξιολόγησης), όσο και ποσοτικά (μετρητικά αξιολόγησης).

- Όσο αυξάνεται το πλήθος των κλάσεων τόσο βελτιώνονται τα αποτελέσματα, αφού το μοντέλο εκπαιδεύεται σε περισσότερο ομοιόμορφα δεδομένα και συνεπώς "μαθαίνει" πιο εύκολα.
- Ο συνολικός χρόνος εκπαίδευσης μικραίνει με την αύξηση του πλήθους των κλάσεων, γεγονός που καθιστά την τεχνική της ταξινόμησης πρακτική και άρα προσδίδει ταυτόχρονα και στην κλιμάκωση του μοντέλου, αλλά και στην ακρίβεια των αποτελεσμάτων.
- Τα μετρητικά αξιολόγησης δεν αρκούν από μόνα τους όσον αφορά στην αξιολόγηση εικόνων. Τα αποτελέσματα θα πρέπει να συνοδεύονται και από οπτικές ενδείξεις (εικόνες εξόδου).

Το βασικό πόρισμα που απορρέει από τα παραπάνω είναι ότι οι συνελικτικοί αυτοκωδικοποιητές, παρόλο που έχουν αποδειχθεί ικανοί στην κωδικοποίηση εικόνων, κρίνονται ανεπαρκείς στην πρόβλεψη ακολουθίας εικόνων από μόνοι τους, λόγω αδυναμίας γενίκευσης. Ο συνδυασμός τους ωστόσο με τεχνικές προεπεξεργασίας δεδομένων, όπως η ταξινόμηση, μπορεί να βελτιώσει αισθητά την ποιότητα των προβλέψεων, καθώς μειώνεται η διακύμανση του συνόλου δεδομένων. Επομένως, κρίνεται απαραίτητος ο συνδυασμός μεθόδων προκειμένου να αντιμετωπιστεί το πρόβλημα πρόβλεψης ακολουθίας εικόνων.

#### 6.2 ΠΙΘΑΝΕΣ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ ΤΗΣ ΕΡΓΑΣΙΑΣ

Παρόλο που τα αποτελέσματα της διπλωματικής δείχνουν να είναι προς την σωστή κατεύθυνση, προφανώς υπάρχουν τομείς στους οποίους μπορούν να υπάρξουν περαιτέρω βελτιώσεις. Κάποιες πιθανές μελλοντικές προεκτάσεις της εργασίες μπορεί να είναι:

- Εξέταση αποτελεσμάτων με αύξηση του πλήθους των κλάσεων κατά την ταξινόμηση των δεδομένων εισόδου. Με αυτόν τον τρόπο, οι επιμέρους εκπαιδεύσεις καθίστανται ευκολότερες (εκπαίδευση σε ομοιόμορφα δεδομένα) και οι προβλέψεις πιο ακριβείς.
- Δοκιμές αρχιτεκτονικών μοντέλου με περισσότερα στρώματα. Στην παρούσα διπλωματική εργασία χρησιμοποιήθηκαν αρχιτεκτονικές τριών και πέντε στρωμάτων. Ενδεχομένως με αύξηση των στρωμάτων το μοντέλο να αποδώσει καλύτερα, αφού με κάθε νέο στρώμα, λόγω των συνελίξεων, δίνεται παραπάνω δυνατότητα για εξαγωγή χαρακτηριστικών.
- Χρήση διαφορετικών διαστάσεων στα φίλτρα που χρησιμοποιούνται στις συνελίξεις. Σε όλες τις αρχιτεκτονικές που υλοποιήθηκαν χρησιμοποιήθηκαν φίλτρα διαστάσεων 3×3 για τις 2D συνελίξεις και 3×3×3 για τις 3D. Ωστόσο, θα μπορούσε να πραγματοποιηθεί και μια διερεύνηση

με φίλτρα μεγαλύτερων διαστάσεων (πχ 5×5, 7×7), γιατί τα 3×3 ίσως είναι μικρά, από την άποψη πως οι συνελίξεις πραγματοποιούνται μόνο τοπικά σε κάθε εικόνα, περιορίζοντας τη δυνατότητα εξαγωγής καθολικής (global) πληροφορίας).

- Δοκιμές διαφορετικών αλγορίθμων βελτιστοποίησης. Στην παρούσα εργασία χρησιμοποιήθηκε ο αλγόριθμος βελτιστοποίησης Adam, μετά από έρευνα στην βιβλιογραφία πάνω σε μοντέλα εκμάθησης μοτίβων από εικόνες. Θα μπορούσε, όμως, να γίνει και δοκιμή με διαφορετικούς αλγορίθμους βελτιστοποίησης, όπως χρήση μεταευρετικών αλγορίθμων για αποφυγή πρόωρης σύγκλισης σε τοπικό βέλτιστο.
- Συνδυασμός CAEs με τις παραδοσιακές μεθόδους DIPT για δημιουργία υβριδικών μοντέλων. Μια ενδιαφέρουσα επέκταση θα ήταν και η εξέταση μοντέλων που χρησιμοποιούν συνελικτικούς αυτοκωδικοποιητές μαζί με τις υπάρχουσες τεχνολογίες ψηφιακής επεξεργασίας εικόνας συνδυάζοντας τα πλεονεκτήματα που προσφέρουν και οι δύο μέθοδοι.
- Αξιοποίηση και αξιολόγηση των αποτελεσμάτων της παρούσας εργασίας σε προβλήματα βραχυπρόθεσμης πρόβλεψης ΦΒ παραγωγής. Η παρούσα εργασία ανέδειξε την σημασία της τεχνικής της ταξινόμησης εισόδου στην βραχυπρόθεσμη πρόβλεψη ακολουθίας εικόνων του ουρανού. Μελλοντικές επεκτάσεις μπορούν να εκτιμήσουν την αξία αυτής της κατηγοριοποίησης των δεδομένων και να την εντάξουν ως αναπόσπαστο κομμάτι της μεθοδολογίας πρόβλεψης εικόνας.
## ΒΙΒΛΙΟΓΡΑΦΙΑ

- Chow, Stanley KH, Eric WM Lee, and Danny HW Li. "Short-term prediction of photovoltaic energy generation by intelligent approach." *Energy and Buildings* 55 (2012): 660-667.
- [2] Solar Panel Prices Over Time, [online] Available: https://www.cladco.co.uk/blog/post/ solar-panel-prices-over-time
- [3] Manajit Sengupta, Aron Habte, Stefan Wilbert, Christian Gueymard, Jan Remund, Elke Lorenz, Wilfried van Sark, and Adam R. Jensen, "Best Practices Handbook for the Collection and Use of Solar Resource Data for Solar Energy Applications: Fourth Edition", 2024.
- [4] Fan Lin, Yao Zhang , Jianxue Wang, "Recent advances in intra-hour solar forecasting: A review of ground-based sky image methods", International Journal of Forecasting 39 (2023) 244–265
- [5] Z. Zhen, Z. Wang, F. Wang, Z. Mi, K. Li, "Research on a cloud image forecasting approach for solar power forecasting," Energy Procedia, vol. 142, pp. 362-368, Dec. 2017.
- [6] Yuwei Fu, Hua Chai, Zhao Zhen, Fei Wang, Xunjian Xu, Kangping Li, Miadreza Shafie-khah, Payman Dehghanian, João P. S. Catalão, "Sky Image Prediction Model Based on Convolutional Auto-encoder for Minutely Solar PV Power Forecasting", *IEEE Transactions on Industry Applications*, 2021.
- [7] Bone, V., Pidgeon, J., Kearney, M., & Veeraragavan, A. (2018). Intra-hour direct normal irradiance forecasting through adaptive clear-sky modelling and cloud tracking. *Solar Energy*, 159, 852-867.
- [8] Zhen, Z., Pang, S., Wang, F., Li, K., Li, Z., Ren, H. & Catalão, J. P. (2019). Pattern classification and PSO optimal weights based sky images cloud motion speed calculation method for solar PV power forecasting. *IEEE Transactions on Industry Applications*, 55(4), 3331-3342.
- [9] Liu, J., Zang, H., Cheng, L., Ding, T., Wei, Z., & Sun, G. (2023). A Transformerbased multimodal-learning framework using sky images for ultra-short-term solar irradiance forecasting. *Applied Energy*, 342, 121160.
- [10] Nie, Y., Zelikman, E., Scott, A., Paletta, Q., & Brandt, A. (2024). Skygpt: Probabilistic ultra-short-term solar forecasting using synthetic sky images from physics-constrained videogpt. *Advances in Applied Energy*, 14, 100172.
- [11] Snaith, H. J. (2018). Present status and future prospects of perovskite photovoltaics. *Nature materials*, 17(5), 372-376

- [12] Martinho, F. (2021). Challenges for the future of tandem photovoltaics on the path to terawatt levels: a technology review. *Energy & Environmental Science*, 14(7), 3840-3871.
- [13] Barbieri, F., Rajakaruna, S., & Ghosh, A. (2017). Very short-term photovoltaic power forecasting with cloud modeling: A review. *Renewable and Sustainable Energy Reviews*, 75, 242-263.
- [14] M. Lourenco, J. P. Barreto, and F. Vasconcelos, "sRD-SIFT: keypoint detection and matching in images with radial distortion," IEEE Trans. Robot., vol. 28, no. 3, pp. 752-760, Jun. 2012.
- [15] S. Dev, F. M. Savoy, Y. H. Lee, and S. Winkler, "Short-term prediction of localized cloud motion using ground-based sky imagers," in 2016 IEEE Region 10 Conference (TENCON), 2016, pp. 2563-2566.
- [16] A. Zaher, S. Thil, J. Nou, A. Traoré, and S. Grieu, "Comparative study of algorithms for cloud motion estimation using sky-imaging data," IFAC-PapersOnLine, vol. 50, no. 1, pp. 5934-5939, Jul. 2017.
- [17] F. Wang, Z. Zhen, C. Liu, Z. Q. Mi, S. M. Hodge, M. Shafie-khah, and J. P. S. Catalão, "Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting," Energy Convers. Manag., vol. 157, pp. 123-135, Feb. 2018.
- [18] Z. Zhen, S. Pang, F. Wang, K. Li, Z. Li, H. Ren, M. Shafie-khah, and J. P.S. Catalão, "Pattern classification and PSO optimal weights based sky images cloud motion speed calculation method for solar PV power forecasting," IEEE Trans. Ind. Appl., vol. 55, no. 4, pp. 3331-3342, Jul.-Aug. 2019.
- [19] Jim Holdsworth, Mark Scapicchio, What is deep learning?, [online] Available: https://www.ibm.com/think/topics/deep-learning
- [20] Ruder, S. (2016). An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*.
- [21] P.S. Georgilakis, Spotlight on Modern Transformer Design. Springer, London, UK, August 2009. ISBN: 978-1 84882-667-0
- [22] What is a neural network?, [online] Available: https://www.ibm.com /think/topics/neural-networks
- [23] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*, 115-133.
- [24] Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.
- [25] Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. *PhD thesis, Committee on Applied Mathematics, Harvard University, Cambridge, MA*.

- [26] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4), 541-551.
- [27] Vincent Dumoulin and Francesco Visin, "A guide to convolution arithmetic for deep learning", 2018
- [28] Danupon Chansong, Siriporn Supratid, "Impacts of Kernel Size on Different Resized Images in Object Recognition Based on Convolutional Neural Network", 2021 International Electrical Engineering Congress (iEECON2021)
- [29] Tato, A., & Nkambou, R. (2018). Improving adam optimizer.
- [30] Sharma, S., Sharma, S., & Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, 6(12), 310-316.
- [31] Κωνσταντίνος Βαϊόπουλος, Αυτοκωδικοποιητής μεταβολών για μη παρεμβατική παρακολούθηση φορτίου, Διπλωματική εργασία, Τμήμα Πληροφορικής, Πρόγραμμα μεταπτυχιακών σπουδών: Τεχνητή νοημοσύνη, Θεσσαλονίκη 2021
- [32] Jabbar, H., & Khan, R. Z. (2015). Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). *Computer science, communication and instrumentation devices*, 70(10.3850), 978-981.
- [33] Bjorck, N., Gomes, C. P., Selman, B., & Weinberger, K. Q. (2018). Understanding batch normalization. *Advances in neural information processing systems*, *31*.
- [34] Κυριάκος Β. Αρμύρας, Εφαρμογη Μεθοδων μη Επιτηρουμενης Ταξινομησης σε Δεδομενα Εικονων του Ουρανου για Βραχυπροθεσμη Προβλεψη Φωτοβολταϊκης Παραγωγης. Μεταπτυχιακή Εργασία, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Φεβρουάριος 2025.
- [35] https://docs.opencv.org/3.4/de/d25/imgproc\_color\_conversions.html?utm\_sou rce=chatgpt.com
- [36] https://docs.opencv.org/4.x/da/d6e/tutorial\_py\_geometric\_transformations.ht ml
- [37] https://scikitlearn.org/stable/modules/generated/sklearn.model\_selection.train\_test\_s plit.html
- [38] Sharma, S., Sharma, S., & Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, 6(12), 310-316.
- [39] https://www.tensorflow.org/guide/keras
- [40] https://www.britannica.com/science/mean-squared-error
- [41] https://medium.com/@akp83540/structural-similarity-index-ssimc5862bb2b520

- [42] Channappayya, S. S., Bovik, A. C., & Heath, R. W. (2008). Rate bounds on SSIM index of quantized images. *IEEE Transactions on Image Processing*, 17(9), 1624-1639.
- [43] Π. Σ. Γεωργιλάκης, Σύγχρονα Συστήματα Μεταφοράς και Διανομής Ηλεκτρικής Ενέργειας. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις, Αθήνα, 2015. DOI: <u>http://dx.doi.org/10.57713/kallipos-808</u>.
- [44] Π. Σ. Γεωργιλάκης, Οικονομική και Αξιόπιστη Λειτουργία Σύγχρονων Συστημάτων Ηλεκτρικής Ενέργειας. Κάλλιπος, Ανοικτές Ακαδημαϊκές Εκδόσεις, Αθήνα, 2023. DOI: <u>http://dx.doi.org/10.57713/kallipos-213</u>.