



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Text-to-Point Cloud Image Synthesis using deep learning

ΕΜΜΑΝΟΥΗΛ ΒΟΤΣΑΡΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΥΠΕΥΘΥΝΟΣ

ΚΟΖΥΡΗ ΜΑΡΙΑ

Αναπληρώτρια Καθηγήτρια (Associate Professor)

Λαμία ...24/02... έτος 2025



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

Text-to-Point Cloud Image Synthesis using deep learning

ΕΜΜΑΝΟΥΗΛ ΒΟΤΣΑΡΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

ΥΠΕΥΘΥΝΟΣ

ΚΟΖΥΡΗ ΜΑΡΙΑ

Αναπληρώτρια Καθηγήτρια (Associate Professor)

Λαμία ...24/02... έτος 2025



UNIVERSITY OF
THESSALY

SCHOOL OF SCIENCE

DEPARTMENT OF COMPUTER SCIENCE & TELECOMMUNICATIONS

Text-to-Point Cloud Image Synthesis using deep learning

EMMANOUIL VOTSARIS

FINAL THESIS

ADVISOR

KOZIRI MARIA
Associate Professor

Lamia ...24/02... year 2025

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις ⁽¹⁾, που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί **χωρίς να τα περικλείω σε εισαγωγικά** και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάσθηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.
2. Δέχομαι ότι η αυτολεξεί **παράθεση χωρίς εισαγωγικά**, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.
3. Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια
4. Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.

Ημερομηνία: 24/02/2025

Ο Δηλ.

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.»

ΠΕΡΙΛΗΨΗ

Σε αυτή την πτυχιακή εργασία θα πραγματοποιηθεί μια ανάλυση της τεχνολογίας του Point Cloud και των οφελών που αυτή παρουσιάζει στην καθημερινότητά μας. Έπειτα θα δούμε κώδικες σε Python και θα τους αναλύσουμε έτσι ώστε να γίνει κατανοητή η λειτουργία τους στον χώρο του image synthesis χρησιμοποιώντας τα δεδομένα του Point Cloud και μετά θα κάνουμε ένα βήμα παρακάτω και θα προσπαθήσουμε να εντοπίσουμε αντικείμενα χρησιμοποιώντας το Deep Learning.

Table of Contents

ΠΕΡΙΛΗΨΗ	7
<u>ΚΕΦΑΛΑΙΟ 1 ΕΙΣΑΓΩΓΗ.....</u>	10
ΤΕΧΝΟΛΟΓΙΑ 3D	10
Η ΤΕΧΝΟΛΟΓΙΑ ΤΟΥ LIDAR ΚΑΙ Η ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ ΜΕΣΩ ΑΥΤΗΣ.....	10
ΦΩΤΟΓΡΑΜΜΕΤΡΙΑ ΚΑΙ Η ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ ΜΕΣΩ ΑΥΤΗΣ.....	12
<u>ΚΕΦΑΛΑΙΟ 2 ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ.....</u>	13
2.1 ΤΕΧΝΟΛΟΓΙΑ ΝΕΦΩΝ ΣΗΜΕΙΩΝ (POINT CLOUDS).....	13
2.2 ΣΥΝΘΕΣΗ ΕΙΚΟΝΑΣ ΚΑΙ ΝΕΦΗ ΣΗΜΕΙΩΝ	14
2.3 ΒΑΘΙΑ ΜΑΘΗΣΗ ΚΑΙ ΝΕΦΗ ΣΗΜΕΙΩΝ.....	14
2.4 ΜΕΤΑΤΡΟΠΗ ΚΕΙΜΕΝΟΥ ΣΕ ΝΕΦΗ ΣΗΜΕΙΩΝ	16
2.5 ΕΡΓΑΛΕΙΑ ΚΑΙ ΤΕΧΝΟΛΟΓΙΕΣ.....	16
2.6 ΣΥΝΟΛΑ ΔΕΔΟΜΕΝΩΝ ΓΙΑ ΝΕΦΗ ΣΗΜΕΙΩΝ.....	18
2.7 ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΚΛΗΣΕΙΣ.....	18
<u>ΚΕΦΑΛΑΙΟ 3.....</u>	20
<u>POINT CLOUD</u>	20
3.1 POINT CLOUD: ΟΡΙΣΜΟΣ ΚΑΙ ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΚΑΙ ΔΗΜΙΟΥΡΓΙΑ.....	20
3.2 ΔΟΜΗ ΚΑΙ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΩΝ POINT CLOUDS.....	21
3.3 ΕΦΑΡΜΟΓΕΣ POINT CLOUDS	22
3.4 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΚΑΙ ΠΡΟΚΛΗΣΕΙΣ ΤΩΝ POINT CLOUDS.....	23
3.5 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ POINT CLOUDS.....	23
3.6 ΜΕΛΛΟΝΤΙΚΕΣ ΤΑΣΕΙΣ ΣΤΗΝ ΤΕΧΝΟΛΟΓΙΑ POINT CLOUDS.....	24
<u>ΚΕΦΑΛΑΙΟ 4.....</u>	26
4.1 ΕΙΣΑΓΩΓΗ ΣΤΗ ΜΕΤΑΤΡΟΠΗ ΚΕΙΜΕΝΟΥ ΣΕ ΝΕΦΗ ΣΗΜΕΙΩΝ	26
4.2 ΜΕΘΟΔΟΛΟΓΙΑ.....	27
4.3 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΜΟΝΤΕΛΩΝ	28
4.4 ΑΞΙΟΛΟΓΗΣΗ ΚΑΙ ΜΕΤΡΙΚΕΣ ΑΠΟΔΟΣΗΣ	29
4.5 ΕΦΑΡΜΟΓΕΣ.....	31
<u>ΚΕΦΑΛΑΙΟ 5 ΕΜΠΕΙΡΙΚΟ ΜΕΡΟΣ.....</u>	33
5.1 ΚΟΡΜΟΣ ΤΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ.....	33
5.2 ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ	34
5.3 ΠΡΩΤΟΚΟΛΛΟ ΕΚΠΑΙΔΕΥΣΗΣ.....	34
5.4 ΑΞΙΟΛΟΓΗΣΗ ΚΑΙ ΟΠΤΙΚΟΠΟΙΗΣΗ.....	35
<u>ΚΕΦΑΛΑΙΟ 6 ΣΥΜΠΕΡΑΣΜΑΤΑ.....</u>	41

<u>ΒΙΒΛΙΟΓΡΑΦΙΑ</u>	<u>43</u>
----------------------------------	------------------

<u>ΠΑΡΑΡΤΗΜΑ.....</u>	<u>48</u>
------------------------------	------------------

ΚΕΦΑΛΑΙΟ 1 Εισαγωγή

Τεχνολογία 3D

Η τεχνολογία 3D υπάρχει πολλά περισσότερα χρόνια από όσα μπορεί να φανταστεί κάποιος. Ξεκίνησε με την δημιουργία της στερεοσκοπικής φωτογραφίας στα τέλη του 19^{ου} αιώνα. Η εξέλιξη της όμως δεν ήταν ραγδαία κατά τον 20^ο αιώνα λόγω των εργαλείων που ήταν διαθέσιμα εκείνα τα χρόνια. Στα τέλη του 20^{ου} και στον 21^ο αιώνα, λόγω των αλμάτων που πραγματοποίησε η τεχνολογία και η υπολογιστική δύναμη που υπάρχει στα χέρια της ανθρωπότητας δημιουργήθηκαν οι προϋποθέσεις για την αναζωπύρωση του ενδιαφέροντος γύρω από τον χώρο του 3D με πολλά περισσότερα εργαλεία επεξεργασίας να δημιουργούνται και να εντάσσονται στην εργαλειοθήκη των επιστημόνων.

Το πιο σημαντικό εργαλείο που χρησιμοποιούμε για την αποτύπωση των τρισδιάστατων δεδομένων είναι το 3D scanner το οποίο είναι μία συσκευή που αναλύει ένα αντικείμενο ή περιβάλλον έτσι ώστε να συλλέξει δεδομένα για το σχήμα, την υφή και την εμφάνιση.

Η ανάπτυξη της τρισδιάστατης απεικόνισης με Laser πραγματοποιήθηκε κατά τα μέσα προς τέλη του 20^{ου} αιώνα. Η αρχική τεχνολογία χρησιμοποιούσε ακτίνες φωτός σε συνδυασμό με κάμερες και προβολείς για την παραγωγή τρισδιάστατων δεδομένων. Εξαιτίας όμως των περιορισμένων δυνατοτήτων που είχε αυτός ο εξοπλισμός ήταν πολύ χρονοβόρα και ήθελε πολύ προσπάθεια για να παραχθούν σωστά και ακριβή δεδομένα κατά την ανάλυση ενός αντικειμένου ή περιβάλλοντος.

Στις μέρες μας υπάρχουν πολλές επιλογές και τύποι 3D laser σαρωτών. Τα δεδομένα που συλλέγουμε από αυτούς τους σαρωτές χρησιμοποιούνται ευρέως σε τομείς όπως τα αυτόνομα οχήματα, την ρομποτική, την εικονική πραγματικότητα, την χαρτογράφηση, την μοντελοποίηση καταστροφών, στον χώρο του θεάματος και στα video games.

Μια από τις πιο ευρέως διαδεδομένες και σημαντικές τεχνολογίες που χρησιμοποιούμε για την συλλογή τρισδιάστατων δεδομένων και δεδομένων Point cloud είναι το LiDAR το οποίο προέρχεται από τα αρχικά του Light detection and ranging.

Η τεχνολογία του LiDAR και η συλλογή δεδομένων μέσω αυτής

Η τεχνολογία του LiDAR ήρθε για πρώτη φορά στην επιφάνεια το 1961 και από τότε έχει εξελιχθεί με τον χρόνο έτσι ώστε από την τοπογραφία και την γενική χρήση της στις οικολογικές αναλύσεις να εισέλθει στον βιομηχανικό χώρο και στην καθημερινή ζωή λειτουργώντας ως τα μάτια των μηχανών όπως τα αυτόνομα οχήματα, τα UAV(unmanned aerial vehicles) και των ρομπότ.

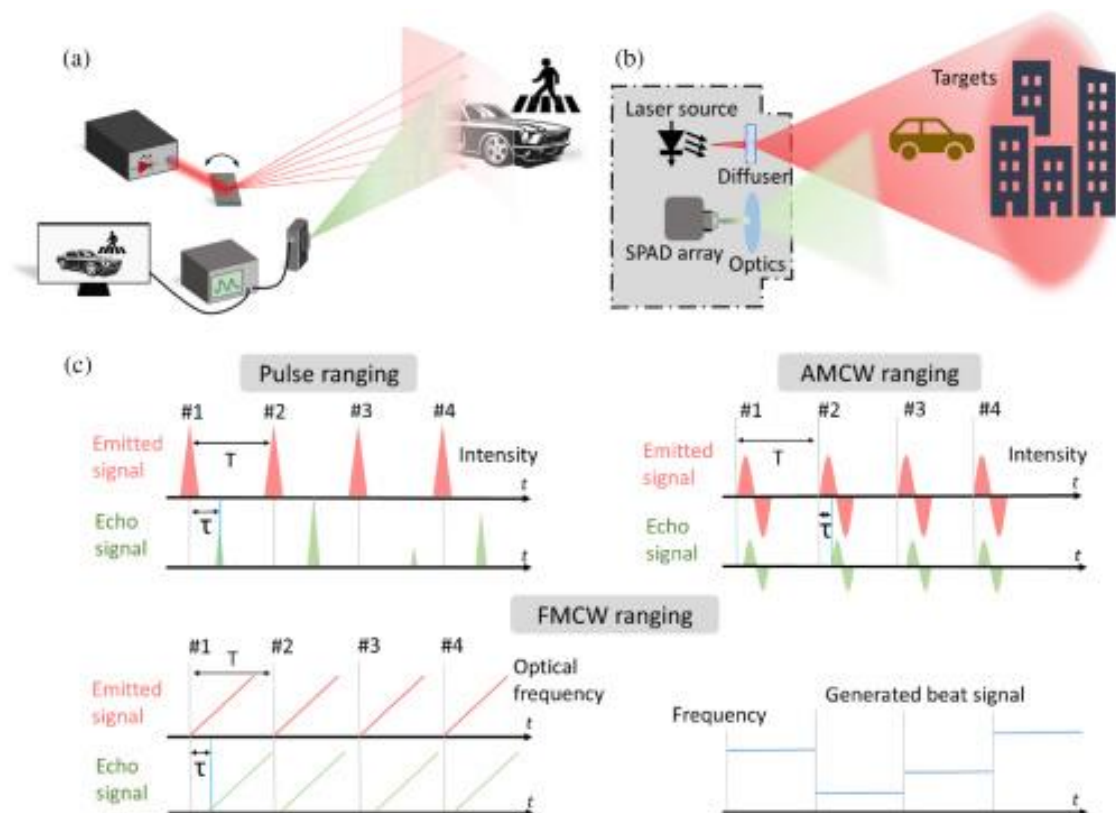
Το LiDAR λειτουργεί σαν ένα κλασικό radar το οποίο περιλαμβάνει έναν πομπό και έναν δέκτη. Η απόσταση μετριέται με το πόσο θα καθυστερήσει το φως που εκπέμπει ο πομπός να ανακλαστεί από κάποιο αντικείμενο στον χώρο και να επιστρέψει πίσω στον δέκτη. Η πιο διαδεδομένη δέσμη φωτός που χρησιμοποιείται στο LiDAR είναι η δέσμη laser διότι το laser έχει πολύ καλές ιδιότητες τις οποίες χρειαζόμαστε.

Για να δημιουργήσουμε τώρα μία 3D εικόνα, οι ακτίνες του φωτός-laser θα πρέπει να πάνε προς όλες τις κατευθύνσεις στο πεδίο που έχουμε να αναλύσουμε. Τα εργαλεία που χρησιμοποιεί για να επιτευχθεί αυτός ο σκοπός είναι το Pulsed LiDAR, το οποίο παρέχει μία ανεκτή ακρίβεια στα δεδομένα που λαμβάνει και λειτουργεί στέλνοντας πολύ μικρούς χρονικά παλμούς φωτός. Αυτοί οι παλμοί αφού ανακλαστούν πάνω στο αντικείμενο ή χώρο ενδιαφέροντος, λαμβάνονται από το LiDAR από τους Single-photon avalanche detectors, οι οποίοι είναι ουσιαστικά φωτοδιοδοί χιονοστιβάδας και λειτουργούν σε reversed biased mode. Αυτό σημαίνει ότι είναι συνδεδεμένοι με αντίθετη πολικότητα και το οποίο συνεισφέρει

στο να αφήνουν πολύ λιγότερο ρεύμα να διαρρέεται αλλά το ηλεκτρικό πεδίο να είναι πολύ ισχυρό έτσι ώστε να γίνεται η ανίχνευση των φωτονίων που φτάνουν στον δέκτη. (Behroozpour et al., October 2017)

Μία ακόμα μέθοδος η οποία χρησιμοποιείται είναι η Amplitude-Modulated Continuous Wave (AMCW) που πετυχαίνει ακρίβεια ίδια με την μέθοδο του Pulsed LiDAR αλλά έχει μικρότερη εμβέλεια. Προτιμάται για μικρές αποστάσεις λόγω του κόστους και την απλότητα της κατασκευής τους και λειτουργεί αυξάνοντας και μειώνοντας την δύναμη του φωτός που εκπέμπει. Στη συνέχεια, το φως που γυρνάει στον δέκτη περνάει από δύο πύλες που κινούνται και αναλόγως τον χρόνο που κάνει για να φτάσει το επιστρεφόμενο φως περνάει περισσότερο από την πρώτη πύλη εάν είναι κοντά και περισσότερο από την δεύτερη πύλη εάν είναι μακριά. (Behroozpour et al., October 2017)

Στις παραπάνω μεθόδους, οι δέκτες τους αντιμετωπίζουν τα φωτόνια σαν στοιχεία των οποίων η πληροφορία μεταφέρεται και λαμβάνεται από τους χρόνους άφιξης στους λήπτες. Η μέθοδος Frequency-Modulated Continuous Wave (FMCW) λειτουργεί μετρώντας την συχνότητα και το μήκος των κυμάτων φωτός και την διαφορά τους από την εκπομπή μέχρι την λήψη τους. Λόγω των χαρακτηριστικών της, με την μέθοδο FMCW έχουμε την δυνατότητα να επιτύχουμε ακρίβεια κάτω του μικρόμετρου στα δεδομένα που αποκτάμε κατά την σάρωση ενός χώρου ή αντικειμένου. (Behroozpour et al., October 2017)



Η βασική αρχιτεκτονική του LiDAR. Όπου τ η διαφορά χρόνου εκπομπής με χρόνο λήψης του σήματος.

Φωτογραμμετρία και η συλλογή δεδομένων μέσω αυτής

Ένας άλλος διαδεδομένος τρόπος συλλογής δεδομένων Point Cloud είναι η μέθοδος της φωτογραμμετρίας. Τι είναι όμως η φωτογραμμετρία και πώς λειτουργεί ;

Ο όρος φωτογραμμετρία προέρχεται από τις λέξεις φως, γραφή και μέτρο και ο ορισμός της είναι η επιστήμη του να λαμβάνουμε αξιόπιστες πληροφορίες σχετικά με τις ιδιότητες των επιφανειών χωρίς να υπάρξει φυσική επαφή με τα αντικείμενα. Η τεχνολογία αυτή βασίζεται στην στερεοσκοπία η οποία είναι η διαδικασία που μας επιτρέπει να βλέπουμε και να καταλαβαίνουμε τις τρεις διαστάσεις. Περιλαμβάνει την λήψη δύο εικόνων από μικρές διαφορετικές γωνίες έτσι ώστε να αποτυπωθεί στα δεδομένα η αίσθηση του βάθους, των τριών διαστάσεων και την καταγραφή της ηλεκτρομαγνητικής ακτινοβολίας που επιστρέφει από τα αντικείμενα στον χώρο και καταγράφεται με την βοήθεια των φωτογραφιών.

Εφόσον γίνει η λήψη των δύο φωτογραφιών τα επόμενα βήματα που έχουμε να κάνουμε είναι:

α) Να ταυτίσουμε τα σημεία που βρίσκονται σε αυτές τις δύο διαδοχικές φωτογραφίες και να υπολογίσουμε την απόσταση μεταξύ αυτών των σημείων και την θέση στην οποία λήφθηκε η φωτογραφία.

β) Αφότου γίνει ταύτιση πολλών σημείων σε αυτές, καταλήγουμε σε μία μαθηματική ανάλυση για κάθε μία από τις θέσεις φωτογράφισης. Με αυτόν τον τρόπο μπορούν να υπολογιστούν οι συντεταγμένες της θέσης (x,y,z), ο προσανατολισμός της κάμερας, η εστιακή απόσταση του φακού και οι σχετικές θέσεις των κοινών σημείων στην επιφάνεια της γης.

γ) Έτσι δημιουργούμε ένα πυκνό νέφος σημείων (Point Cloud) από θέσεις επάνω στην φωτογραφισμένη επιφάνεια χρησιμοποιώντας τις παραμέτρους της κάθε θέσης φωτογράφισης.

δ) Τέλος μεταφέρουμε τις συντεταγμένες από ένα σχετικό σύστημα σε ένα σύστημα γεωγραφικών συντεταγμένων. Το παραπάνω βήμα αφορά κατά κύριο λόγο τα δεδομένα που λαμβάνονται για τοπογραφικούς λόγους. (Bao et al., 2024).

Η καταγραφή τώρα των δεδομένων με την φωτογραμμετρία μας δίνει την σημαντική πληροφορία των γεωμετρικών ιδιοτήτων του χώρου και των αντικειμένων που βρίσκονται σε αυτόν. Επίσης καταγράφονται πολλές από τις φυσικές ιδιότητες όπως εάν κάποιο από τα αντικείμενα εκπέμπουν ακτινοβολία. Τέλος με την χρήση της φωτογραμμετρίας σε διάστημα δύο διαφορετικών χρονικών περιόδων μπορούμε να λάβουμε πληροφορίες σχετικά με το πόσο έχει μεταβληθεί το περιβάλλον στο οποίο έχει χρησιμοποιηθεί αυτή η μέθοδος.

Στα επόμενα κεφάλαια θα αναλυθούν σε βάθος ο όρος Νέφη σημείων(Point Cloud) μαζί με την δομή τους, τα χαρακτηριστικά τους και τον τρόπο δημιουργίας τους και ανάλυσης των δεδομένων τους. Έπειτα θα προβούμε στην ανάλυση της μεθοδολογίας της μετατροπής κειμένου σε νέφη σημείων (Text-to-Point-Cloud). Τέλος στο εμπειρικό μέρος θα υπάρχει κώδικας σε γλώσσα προγραμματισμού Python που χρησιμοποιεί μηχανική μάθηση και ο οποίος θα θέτει σε εφαρμογή τις πληροφορίες που θα σας έχω παρουσιάσει έτσι ώστε να γίνει πιο κατανοητός ο σκοπός της παρούσας πτυχιακής.

Στο κεφάλαιο 2 τώρα, θα γίνει μία βιβλιογραφική επισκόπηση με τις πηγές και τα επιστημονικά άρθρα που χρησιμοποιήθηκαν για να αντληθούν πληροφορίες.

ΚΕΦΑΛΑΙΟ 2

Βιβλιογραφική Επισκόπηση

2.1 Τεχνολογία Νεφών Σημείων (Point Clouds)

Τα νέφη σημείων (Point Clouds) αποτελούν μια κρίσιμη τεχνολογία στην τρισδιάστατη αποτύπωση και ανάλυση αντικειμένων και περιβαλλόντων. Χρησιμοποιούνται ευρέως σε τομείς όπως η αρχιτεκτονική, η ρομποτική, η χαρτογράφηση και η ανάπτυξη αυτόνομων οχημάτων.

Ένα νέφος σημείων αποτελείται από μια συλλογή διακριτών σημείων, καθένα από τα οποία ορίζεται από συντεταγμένες σε ένα τρισδιάστατο σύστημα (X, Y, Z). Επιπλέον, τα σημεία αυτά μπορούν να περιλαμβάνουν πρόσθετα χαρακτηριστικά, όπως χρώμα, ένταση και ανακλαστικότητα (Grilli, Menna, & Remondino, 2017). Αυτή η μη δομημένη μορφή δεδομένων παρέχει τη δυνατότητα αποτύπωσης σύνθετων γεωμετρικών λεπτομερειών, καθιστώντας τα νέφη σημείων αναντικατάστατα για εφαρμογές τρισδιάστατης μοντελοποίησης και ανάλυσης (Hackel, Wegner, & Schindler, 2016).

Η συλλογή δεδομένων νέφους σημείων πραγματοποιείται μέσω προηγμένων τεχνολογιών, όπως το LiDAR (Light Detection and Ranging) και η φωτογραμμετρία, που αποτελούν θεμέλιο για τη δημιουργία υψηλής ακρίβειας τρισδιάστατων δεδομένων. Το LiDAR χρησιμοποιεί παλμούς λέιζερ για τη μέτρηση αποστάσεων προς επιφάνειες, δημιουργώντας νέφη σημείων υψηλής ακρίβειας. Αυτή η τεχνολογία είναι ιδιαίτερα κατάλληλη για τοπογραφικές εφαρμογές, περιβαλλοντική χαρτογράφηση και αυτόνομη οδήγηση, καθώς μπορεί να διεισδύει στη βλάστηση και να καταγράφει λεπτομερή γεωμετρικά χαρακτηριστικά (Dong et al., 2020). Οι πρόσφατες εξελίξεις στους αισθητήρες LiDAR έχουν βελτιώσει την ακρίβεια, μειώνοντας παράλληλα το κόστος, γεγονός που την καθιστά πιο προσιτή και ευρέως χρησιμοποιούμενη (Zhuang et al., 2024). Από την άλλη, η φωτογραμμετρία βασίζεται στη χρήση επικαλυπτόμενων φωτογραφιών από διαφορετικές γωνίες. Μέσω αλγορίθμων υπολογιστικής όρασης, τα σημεία αναγνωρίζονται και τριγωνοποιούνται, αναπαριστώντας τρισδιάστατες δομές. Αυτή η μέθοδος είναι ιδιαίτερα χρήσιμη για την αποτύπωση υφών και βρίσκει εφαρμογή σε περιπτώσεις όπως οι γεωλογικές μελέτες (Granshaw et al., 2017).

Η επεξεργασία των δεδομένων νέφους σημείων είναι μια διαδικασία πολλαπλών σταδίων που εξασφαλίζει τη χρηστικότητα και την ποιότητα των δεδομένων. Πρώτο βήμα αποτελεί ο καθαρισμός των δεδομένων, που περιλαμβάνει την απομάκρυνση θορύβου και ανωμαλιών μέσω τεχνικών όπως το φιλτράρισμα βάσει στατιστικών μοντέλων (Guo et al., 2020). Στη συνέχεια, η τμηματοποίηση διαδραματίζει κεντρικό ρόλο, καθώς ομαδοποιεί τα σημεία με κοινά χαρακτηριστικά για την αναγνώριση αντικειμένων ή περιοχών ενδιαφέροντος (Hackel et al., 2016). Το τελικό στάδιο είναι η ευθυγράμμιση, όπου δεδομένα από πολλαπλές σαρώσεις ενοποιούνται για τη δημιουργία συνεκτικών και ολοκληρωμένων τρισδιάστατων μοντέλων (Dong et al., 2020).

Τα χαρακτηριστικά των δεδομένων νέφους σημείων, όπως η πυκνότητα, η ποιότητα και η ακρίβεια, καθορίζονται σε μεγάλο βαθμό από τις τεχνικές συλλογής και επεξεργασίας. Για παράδειγμα, το LiDAR δημιουργεί εξαιρετικά πυκνά και ακριβή νέφη σημείων, ενώ η φωτογραμμετρία προσφέρει δεδομένα πλούσια σε υφή (Weinmann, Jutzi, & Mallet, 2017). Η επιλογή της κατάλληλης τεχνολογίας και των μεθόδων επεξεργασίας εξαρτάται από τις απαιτήσεις της εκάστοτε εφαρμογής, με στόχο την επίτευξη της μέγιστης ποιότητας και ακρίβειας στα παραγόμενα δεδομένα.

2.2 Σύνθεση Εικόνας και Νέφη Σημείων

Η σύνθεση εικόνας από δεδομένα νέφους σημείων αποτελεί έναν ταχέως εξελισσόμενο τομέα που συνδυάζει τεχνικές επεξεργασίας τρισδιάστατων δεδομένων με μοντέρνες μεθόδους βαθιάς μάθησης. Αυτή η διαδικασία έχει εφαρμογές σε ποικίλα πεδία, όπως η αρχιτεκτονική, η χαρτογράφηση και η επαυξημένη πραγματικότητα.

Η σύνθεση εικόνας από δεδομένα νεφών σημείων αποτελεί σημαντικό πεδίο έρευνας και μπορεί να επιτευχθεί μέσω διαφόρων μεθόδων, που διακρίνονται σε γεωμετρικές τεχνικές, μεθόδους βαθιάς μάθησης και συνδυασμένες προσεγγίσεις.

Οι γεωμετρικές τεχνικές βασίζονται στη γεωμετρική ανάλυση των νεφών σημείων για τη δημιουργία πολυγωνικών πλεγμάτων ή τρισδιάστατων μοντέλων, τα οποία στη συνέχεια αποδίδονται σε εικόνες. Παραδοσιακές μέθοδοι, όπως η τριγωνοποίηση Delaunay και η ανακατασκευή επιφανειών Poisson, είναι ιδιαίτερα δημοφιλείς (Kazhdan & Hoppe, 2013). Αυτές οι τεχνικές είναι γνωστές για την αξιοπιστία τους στη δημιουργία ρεαλιστικών εικόνων. Ωστόσο, συχνά απαιτούν υψηλή υπολογιστική ισχύ και εντατική χειροκίνητη παρέμβαση, καθιστώντας τις λιγότερο πρακτικές για μεγάλης κλίμακας εφαρμογές.

Αντίθετα, οι μέθοδοι βαθιάς μάθησης έχουν φέρει επανάσταση στη σύνθεση εικόνας από δεδομένα νεφών σημείων. Μοντέλα όπως τα PointNet και PointNet++ επιτρέπουν την αποτελεσματική επεξεργασία νεφών σημείων, ενώ τεχνικές που βασίζονται σε Generative Adversarial Networks (GANs) χρησιμοποιούνται για τη δημιουργία φωτορεαλιστικών εικόνων από τρισδιάστατα δεδομένα (Qi et al., 2017). Οι μέθοδοι βαθιάς μάθησης ξεχωρίζουν για την ικανότητά τους να εξάγουν αυτόματα χαρακτηριστικά από δεδομένα και να συνθέτουν εικόνες με υψηλή ακρίβεια, μειώνοντας την ανάγκη για χειροκίνητη παρέμβαση.

Επιπλέον, οι συνδυασμένες προσεγγίσεις που ενσωματώνουν γεωμετρικές τεχνικές με μεθόδους βαθιάς μάθησης έχουν αποδειχθεί ιδιαίτερα αποδοτικές. Για παράδειγμα, η χρήση γεωμετρικών μοντέλων για την προεπεξεργασία των δεδομένων πριν από την εφαρμογή τεχνικών βαθιάς μάθησης βελτιώνει τόσο την ακρίβεια όσο και την αποτελεσματικότητα της σύνθεσης εικόνας (Zhang et al., 2023). Με αυτόν τον τρόπο, επιτυγχάνεται ένας συνδυασμός της αξιοπιστίας των γεωμετρικών τεχνικών με την αυτοματοποιημένη ευελιξία των μεθόδων βαθιάς μάθησης.

Οι παραδοσιακές προσεγγίσεις, όπως οι γεωμετρικές μέθοδοι, παρέχουν σταθερότητα και υψηλή ακρίβεια στη σύνθεση εικόνας, αλλά συχνά περιορίζονται από την υπολογιστική τους πολυπλοκότητα και την ανάγκη για προσεκτική χειροκίνητη ρύθμιση (Kazhdan & Hoppe, 2013). Αντίθετα, οι σύγχρονες μέθοδοι βαθιάς μάθησης, όπως τα PointNet και GANs, είναι πιο ευέλικτες, αυτοματοποιημένες και ικανές να επεξεργάζονται μεγάλα σύνολα δεδομένων. Παρά τη μεγαλύτερη απαίτηση σε δεδομένα εκπαίδευσης, οι μέθοδοι αυτές προσφέρουν υψηλότερη αποδοτικότητα και μπορούν να προσαρμοστούν σε ποικίλες εφαρμογές.

2.3 Βαθιά Μάθηση και Νέφη Σημείων

Η βαθιά μάθηση έχει φέρει επανάσταση στη διαχείριση και ανάλυση νεφών σημείων, παρέχοντας καινοτόμες λύσεις για την αυτόματη επεξεργασία μεγάλων και σύνθετων δεδομένων. Η δυνατότητα των μοντέλων βαθιάς μάθησης να εξάγουν χαρακτηριστικά απευθείας από τα δεδομένα καθιστά την τεχνολογία αυτή απαραίτητη για εφαρμογές σε ποικίλους τομείς.

Η εφαρμογή της βαθιάς μάθησης σε νέφη σημείων εστιάζει σε προβλήματα όπως η ταξινόμηση, η τμηματοποίηση και η αναγνώριση αντικειμένων. Αντί να απαιτούνται

σύνθετες διαδικασίες προεπεξεργασίας, τα μοντέλα βαθιάς μάθησης μπορούν να λειτουργούν απευθείας σε μη δομημένα δεδομένα, χρησιμοποιώντας αλγόριθμους που προσαρμόζονται στις ιδιαιτερότητες των νεφών σημείων (Guo et al., 2020). Αυτή η προσέγγιση προσφέρει αυξημένη ακρίβεια και αποτελεσματικότητα, ειδικά σε εφαρμογές που απαιτούν ανάλυση σε πραγματικό χρόνο.

Η εξέλιξη των μοντέλων βαθιάς μάθησης έχει συμβάλει καθοριστικά στην αξιοποίηση των νεφών σημείων, επιτρέποντας την εξαγωγή πολύπλοκων χαρακτηριστικών και τη βελτιστοποίηση εφαρμογών σε ποικίλους τομείς. Μερικά από τα πιο σημαντικά μοντέλα περιλαμβάνουν:

- ✚ Τα Convolutional Neural Networks (CNNs) αποτέλεσαν μια από τις πρώτες προσπάθειες επεξεργασίας τρισδιάστατων δεδομένων μέσω της προβολής τους σε 2D επίπεδο. Αυτή η μέθοδος επιτυγχάνει τη μετατροπή νεφών σημείων σε εικόνες, στις οποίες εφαρμόζονται παραδοσιακές τεχνικές CNN. Αν και αυτή η προσέγγιση υστερεί στην αποτύπωση πλήρων χωρικών πληροφοριών, παραμένει αξιόπιστη και ευρέως χρησιμοποιούμενη στις αρχικές εφαρμογές επεξεργασίας νεφών σημείων (Qi et al., 2017).
- ✚ Τα PointNet και PointNet++ έφεραν επανάσταση στην ανάλυση νεφών σημείων, επιτρέποντας την επεξεργασία τους απευθείας, χωρίς να απαιτείται προηγούμενη μετατροπή σε άλλα μορφότυπα. Αυτά τα μοντέλα βασίζονται σε συμμετρικές συναρτήσεις που εξασφαλίζουν την αμεταβλητότητα στη σειρά των σημείων και μπορούν να εξάγουν χαρακτηριστικά απευθείας από τις συντεταγμένες τους. Το PointNet++ επεκτείνει τη λειτουργικότητα του PointNet με ιεραρχική προσέγγιση, ενσωματώνοντας πληροφορίες τοπικής γειννίας (Qi et al., 2017).
- ✚ Τα Transformers αποτελούν μια πιο πρόσφατη εξέλιξη, με εξαιρετική ικανότητα μοντελοποίησης μη γραμμικών σχέσεων μεταξύ των σημείων. Το Point Transformer χρησιμοποιεί μηχανισμούς προσοχής για την εξαγωγή πληροφοριών τόσο χωρικών όσο και θεματικών, καθιστώντας το κατάλληλο για σύνθετες εφαρμογές, όπως η κατανόηση περιβάλλοντος και η κατηγοριοποίηση αντικειμένων (Zhao et al., 2021).

Η χρήση βαθιάς μάθησης σε συνδυασμό με δεδομένα νέφους σημείων έχει σημαντικές εφαρμογές, μερικές από τις οποίες παρατίθενται στη συνέχεια:

- ✚ Αυτόνομα Συστήματα: Οι αισθητήρες LiDAR, σε συνδυασμό με μοντέλα βαθιάς μάθησης, είναι απαραίτητοι για την πλοήγηση αυτόνομων οχημάτων. Ειδικότερα, χρησιμοποιούνται για την ανίχνευση εμποδίων, τη χαρτογράφηση του περιβάλλοντος και την πρόβλεψη της κίνησης άλλων οχημάτων και πεζών (Li et al., 2020).
- ✚ Ρομποτική: Στον τομέα της ρομποτικής, τα νέφη σημείων παρέχουν κρίσιμες πληροφορίες για την πλοήγηση και την αλληλεπίδραση με αντικείμενα σε δυναμικά περιβάλλοντα. Η βαθιά μάθηση ενισχύει την ικανότητα των ρομπότ να αντιλαμβάνονται το περιβάλλον και να προσαρμόζονται σε πολύπλοκες συνθήκες (Guo et al., 2020).
- ✚ Χαρτογράφηση και Εικονική Πραγματικότητα: Τα νέφη σημείων είναι ιδανικά για την ακριβή αποτύπωση περιβαλλόντων, ενώ η εφαρμογή βαθιάς μάθησης ενισχύει τη δυνατότητα ανάλυσης των δεδομένων αυτών. Το αποτέλεσμα είναι η δημιουργία καθηλωτικών εμπειριών εικονικής πραγματικότητας, που χρησιμοποιούνται σε τομείς όπως η αρχιτεκτονική, η πολιτιστική κληρονομιά και η ψυχαγωγία (Zhuang et al., 2024).

Η πρόοδος στα μοντέλα βαθιάς μάθησης και η ενσωμάτωσή τους σε εφαρμογές που βασίζονται σε νέφη σημείων έχουν ανοίξει νέους δρόμους για την καινοτομία, καθιστώντας δυνατές εφαρμογές υψηλής ακρίβειας και απόδοσης σε ποικίλα πεδία.

2.4 Μετατροπή Κειμένου σε Νέφη Σημείων

Η μετατροπή κειμένου σε νέφη σημείων (Text-to-Point-Cloud) είναι ένα αναδυόμενο πεδίο που συνδυάζει τεχνικές φυσικής γλώσσας (Natural Language Processing - NLP) με τεχνολογίες τρισδιάστατης αναπαράστασης. Η δυνατότητα μετάφρασης περιγραφών κειμένου σε τρισδιάστατα δεδομένα ανοίγει νέους ορίζοντες σε τομείς όπως η αρχιτεκτονική, η εικονική πραγματικότητα και η ρομποτική.

Οι πρώτες προσπάθειες για τη μετατροπή κειμένου σε τρισδιάστατα δεδομένα επικεντρώθηκαν στο πεδίο του Text-to-3D, όπου περιγραφές κειμένου μετατρέπονταν σε μοντέλα τρισδιάστατων πλεγμάτων. Η πρόσφατη έρευνα εστιάζει σε νέφη σημείων, αξιοποιώντας την ευελιξία και τη λεπτομέρεια που προσφέρουν αυτά τα δεδομένα. Για παράδειγμα, το μοντέλο Point-E της OpenAI χρησιμοποιεί συνδυασμό διάχυσης και νευρωνικών δικτύων για τη δημιουργία νεφών σημείων από περιγραφές κειμένου (Nichol et al., 2022).

Οι μεθοδολογίες και προσεγγίσεις στη σύνθεση νεφών σημείων από περιγραφές κειμένου βασίζονται σε συνδυασμούς προηγμένων αρχιτεκτονικών και καινοτόμων τεχνικών που αξιοποιούν τη δύναμη της φυσικής γλώσσας και της τρισδιάστατης αναπαράστασης. Οι αρχιτεκτονικές μοντέλων, όπως οι Transformers, συνδυάζονται με τεχνικές τρισδιάστατης αναπαράστασης για να επιτύχουν εντυπωσιακά αποτελέσματα. Ένα χαρακτηριστικό παράδειγμα είναι το Text2Mesh, το οποίο μετατρέπει περιγραφές κειμένου σε λεπτομερείς τρισδιάστατες γεωμετρίες, παρέχοντας έτσι νέες δυνατότητες στην προσαρμοσμένη δημιουργία 3D μοντέλων (Xiao et al., 2024).

Παράλληλα, τα Generative Adversarial Networks (GANs) προσαρμόζονται για τη σύνθεση νεφών σημείων με βάση σύνολα δεδομένων τρισδιάστατων αντικειμένων και περιγραφών. Τα GANs εκπαιδεύονται για να παράγουν νέφη σημείων που ανταποκρίνονται στις περιγραφικές εισόδους, διευρύνοντας τις δυνατότητες δημιουργίας και προσαρμογής γεωμετριών (Wu et al., 2021). Επιπλέον, οι διαδικασίες διάχυσης (diffusion processes) παρέχουν μια άλλη ισχυρή προσέγγιση, καθώς επιτρέπουν τη σταδιακή ανακατασκευή τρισδιάστατων δεδομένων από θόρυβο με την καθοδήγηση περιγραφών κειμένου. Αυτή η μέθοδος έχει αποδειχθεί εξαιρετικά αποτελεσματική για τη δημιουργία σύνθετων και λεπτομερών γεωμετρικών σχημάτων (Nichol et al., 2022).

Ωστόσο, το πεδίο αυτό δεν στερείται περιορισμών και προκλήσεων. Ένας από τους κύριους περιορισμούς είναι τα περιορισμένα σύνολα δεδομένων που συνδέουν περιγραφές κειμένου με νέφη σημείων. Η έλλειψη ποικίλων και μεγάλων συνόλων δεδομένων περιορίζει την ικανότητα των μοντέλων να γενικεύουν και να παράγουν αξιόπιστα αποτελέσματα (Kolmet et al., 2022). Επιπλέον, οι ασαφείς ή ανεπαρκείς περιγραφές κειμένου δυσχεραίνουν τη δημιουργία ακριβών νεφών σημείων, ενώ η υπολογιστική πολυπλοκότητα που συνοδεύει γενετικές μεθόδους, όπως αυτές που βασίζονται στη διάχυση ή σε GANs, αποτελεί εμπόδιο για την ευρεία υιοθέτησή τους σε πρακτικές εφαρμογές (Wu et al., 2021).

Παρά τα εμπόδια αυτά, η πρόοδος στο πεδίο υπόσχεται να ενισχύσει τη σύνδεση μεταξύ φυσικής γλώσσας και τρισδιάστατων δεδομένων. Η εξέλιξη αυτών των τεχνολογιών ανοίγει τον δρόμο για τη διευκόλυνση της κατανόησης και της αλληλεπίδρασης με σύνθετες χωρικές πληροφορίες, καθιστώντας δυνατές νέες και καινοτόμες εφαρμογές.

2.5 Εργαλεία και Τεχνολογίες

Η δημιουργία και επεξεργασία νεφών σημείων αποτελεί κρίσιμη διαδικασία που βασίζεται στη χρήση προηγμένων εργαλείων λογισμικού και τεχνολογιών, παρέχοντας ακρίβεια, ευελιξία και αυτοματοποίηση. Οι τεχνολογίες αυτές χρησιμοποιούνται ευρέως σε στάδια όπως η συλλογή δεδομένων, η τμηματοποίηση, η μοντελοποίηση και η ανάλυση, επιτρέποντας την παραγωγή υψηλής ποιότητας αποτελεσμάτων σε διάφορους τομείς εφαρμογής.

Ένα από τα πιο διαδεδομένα εργαλεία για την επεξεργασία νεφών σημείων είναι το CloudCompare, ένα λογισμικό ανοιχτού κώδικα που χρησιμοποιείται για τη σύγκριση δεδομένων, την τμηματοποίηση και την ανάλυση τρισδιάστατων επιφανειών. Το CloudCompare ξεχωρίζει για την ικανότητά του να διαχειρίζεται μεγάλα σύνολα δεδομένων και την εξειδίκευσή του στη διαχείριση δεδομένων LiDAR (Girardeau-Montaut, 2015).

Το Autodesk ReCap αποτελεί ένα εμπορικό εργαλείο που εστιάζει στην εισαγωγή, επεξεργασία και οπτικοποίηση δεδομένων νέφους σημείων. Είναι ιδιαίτερα δημοφιλές στην αρχιτεκτονική και την κατασκευαστική βιομηχανία, καθώς προσφέρει συμβατότητα με το Building Information Modeling (BIM), διευκολύνοντας τον σχεδιασμό και την υλοποίηση έργων (Pro, n.d.).

Το MeshLab, ένα ακόμη εργαλείο ανοιχτού κώδικα, εστιάζει στη δημιουργία πλεγμάτων από δεδομένα νέφους σημείων. Αποτελεί ιδανική επιλογή για την τρισδιάστατη μοντελοποίηση, την ανακατασκευή επιφανειών και την προετοιμασία δεδομένων για 3D εκτύπωση (Cignoni et al., 2008).

Τέλος, το Potree είναι μια βιβλιοθήκη σχεδιασμένη για την απεικόνιση νεφών σημείων στον ιστό. Χάρη στη διαδραστική διεπαφή του, το Potree είναι ιδανικό για την παρουσίαση και εξερεύνηση μεγάλων συνόλων δεδομένων μέσω διαδικτύου, διευκολύνοντας την πρόσβαση και την ανάλυση (Schütz, 2015).

Οι σύγχρονες τεχνολογίες συλλογής και επεξεργασίας δεδομένων νεφών σημείων περιλαμβάνουν:

- **LiDAR:** Οι σαρωτές LiDAR παρέχουν υψηλής ακρίβειας δεδομένα νέφους σημείων, καθιστώντας τους απαραίτητους για εφαρμογές όπως η χαρτογράφηση και η ανάλυση τοπογραφικών δεδομένων. Νέας γενιάς συστήματα, όπως αυτά της εταιρείας Velodyne, προσφέρουν μεγαλύτερη πυκνότητα σημείων και βελτιωμένη απόδοση, ενισχύοντας τη λειτουργικότητά τους σε ποικίλα περιβάλλοντα (Cui et al., 2024).
- **Φωτογραμμετρία:** Εργαλεία όπως το Pix4D και το Agisoft Metashape χρησιμοποιούνται για τη δημιουργία νεφών σημείων από εικόνες, αξιοποιώντας αλγορίθμους υπολογιστικής όρασης για την τρισδιάστατη ανακατασκευή. Αυτές οι τεχνολογίες παρέχουν ευελιξία και υψηλή ακρίβεια, ιδιαίτερα σε περιπτώσεις όπου οι σαρωτές LiDAR δεν είναι διαθέσιμοι (Pix4D, 2022).
- **GPU Acceleration:** Οι σύγχρονες κάρτες γραφικών, όπως οι NVIDIA RTX, χρησιμοποιούνται για την επιτάχυνση της επεξεργασίας νεφών σημείων. Ειδικά στις εφαρμογές βαθιάς μάθησης και τμηματοποίησης, η χρήση GPU μειώνει σημαντικά τον απαιτούμενο χρόνο επεξεργασίας και αυξάνει την αποδοτικότητα (NVIDIA, 2023).

Η συνδυασμένη χρήση προηγμένων εργαλείων λογισμικού και τεχνολογιών συλλογής δεδομένων καθιστά τη δημιουργία και την επεξεργασία νεφών σημείων μια ισχυρή διαδικασία για την επίλυση σύνθετων προβλημάτων σε τομείς όπως η αρχιτεκτονική, η χαρτογράφηση, η ρομποτική και τα αυτόνομα συστήματα. Η επιλογή του κατάλληλου εργαλείου εξαρτάται από την εφαρμογή και τις απαιτήσεις. Εργαλεία όπως το CloudCompare είναι κατάλληλα για την ανάλυση ακατέργαστων δεδομένων, ενώ το Autodesk ReCap ενσωματώνεται άψογα σε επαγγελματικά περιβάλλοντα BIM. Από την άλλη, η MeshLab προσφέρει εξαιρετικές δυνατότητες για μοντελοποίηση, ενώ το Potree είναι ιδανικό για την απεικόνιση δεδομένων στον ιστό.

2.6 Σύνολα Δεδομένων για Νέφη Σημείων

Ο τομέας της παρουσίασης των δεδομένων των οποίων η συλλογή έχει γίνει μέσα από τις μεθόδους όπως το LiDAR είναι μία διαρκής πρόκληση για τους επιστήμονες που ασχολούνται με αυτό το αντικείμενο. Επομένως αυτό είχε ως αποτέλεσμα να σημειωθούν ποικίλες προσεγγίσεις και επιτεύγματα στον τομέα της επεξεργασίας και ανάλυσης νεφών σημείων. Ένα χαρακτηριστικό έργο είναι το ShapeNet, που παρουσιάστηκε από τους Chang et al. (2015). Το ShapeNet αποτελεί μια μεγάλη βάση δεδομένων τρισδιάστατων μοντέλων με αναλυτικές περιγραφές και ταξινομήσεις. Χρησιμοποιείται ευρέως για την εκπαίδευση και την αξιολόγηση μοντέλων βαθιάς μάθησης που διαχειρίζονται νέφη σημείων. Η μελέτη ανέδειξε τη σημασία του συνδυασμού γεωμετρικών δεδομένων με περιγραφές κειμένου, παρέχοντας νέες δυνατότητες για την κατανόηση σύνθετων τρισδιάστατων δομών.

Μια άλλη σημαντική συνεισφορά είναι το Stanford Large-Scale 3D Indoor Spaces Dataset (S3DIS), που παρουσιάστηκε από τον Armeni et al. (2016). Το S3DIS προσφέρει δεδομένα που επικεντρώνονται στην τμηματοποίηση και ανάλυση εσωτερικών χώρων. Η συγκεκριμένη βάση δεδομένων έθεσε τα θεμέλια για την ανίχνευση αντικειμένων και την κατασκευή ψηφιακών μοντέλων κτιρίων, παρέχοντας δεδομένα υψηλής ανάλυσης που είναι απαραίτητα για την κατανόηση και αναπαράσταση εσωτερικών χώρων.

Μια ακόμα πρωτοποριακή εργασία είναι το Deep Global Registration, που παρουσιάστηκε από τους Choy et al. (2020). Σε αυτήν, προτάθηκε ένα σύστημα βαθιάς μάθησης για την ευθυγράμμιση νεφών σημείων μεγάλης κλίμακας. Η μέθοδος συνδύασε τρισδιάστατες πληροφορίες με τεχνικές μετασχηματισμού, επιτυγχάνοντας βελτιωμένη ακρίβεια στην επεξεργασία δεδομένων, ειδικά σε εφαρμογές μεγάλης κλίμακας.

Η ανάλυση και σύγκριση των παραπάνω μελετών καταδεικνύει τις διαφορετικές εφαρμογές και εστιασμούς τους. Το ShapeNet ξεχωρίζει για τη δημιουργία πλούσιων βάσεων δεδομένων που διευκολύνουν τη σύνθεση εικόνων και τη μοντελοποίηση τρισδιάστατων δεδομένων. Το S3DIS εστιάζει στην ανάλυση εσωτερικών χώρων, προσφέροντας δεδομένα υψηλής ανάλυσης για την τμηματοποίηση σημείων. Τέλος, το Deep Global Registration συμβάλλει στη βελτίωση της ακρίβειας στην ευθυγράμμιση νεφών σημείων μέσω καινοτόμων μεθόδων βαθιάς μάθησης. Οι μελέτες αυτές, αν και διαφορετικές στις προσεγγίσεις και τους στόχους τους, συνδυαστικά προάγουν το πεδίο της ανάλυσης και επεξεργασίας νεφών σημείων, ανοίγοντας νέους δρόμους για την εφαρμογή τους σε ποικίλα πεδία.

Παρά τις διαφορετικές προσεγγίσεις, και οι τρεις μελέτες υπογραμμίζουν τη σημασία της συνδυαστικής χρήσης δεδομένων και αλγορίθμων βαθιάς μάθησης για την ανάλυση και επεξεργασία νεφών σημείων. Εντούτοις, η έλλειψη επαρκούς γενίκευσης σε συγκεκριμένες περιπτώσεις παραμένει μια κοινή πρόκληση.

2.7 Μελλοντικές Προκλίσεις

Παρά την αξιοσημείωτη πρόοδο στη χρήση νεφών σημείων και στις τεχνολογίες βαθιάς μάθησης, εξακολουθούν να υφίστανται σημεία στη βιβλιογραφία που απαιτούν περαιτέρω διερεύνηση. Τα κενά αυτά αφορούν τόσο τη θεωρητική κατανόηση όσο και τις πρακτικές εφαρμογές σε σύνθετα και εξειδικευμένα περιβάλλοντα, περιορίζοντας την πλήρη αξιοποίηση των δυνατοτήτων που προσφέρει η σύγχρονη τεχνολογία.

Ένα από τα πιο σημαντικά ζητήματα είναι η έλλειψη δεδομένων για εξειδικευμένες εφαρμογές. Ενώ βάσεις δεδομένων όπως το ShapeNet και το S3DIS έχουν θέσει τα θεμέλια για γενικές εφαρμογές, υπάρχει ανάγκη για σύνολα δεδομένων που να συνδέουν περιγραφές κειμένου με νέφη σημείων για εξειδικευμένες χρήσεις, όπως η αυτόνομη πλοήγηση σε αδόμητα περιβάλλοντα (Xu et al., 2023). Παράλληλα, οι υπάρχουσες μέθοδοι παρουσιάζουν περιορισμένη ακρίβεια στη μετατροπή κειμένου σε νέφη σημείων, με τα τρέχοντα μοντέλα, όπως το Point-E, να υποφέρουν από απώλεια λεπτομέρειας και ακρίβειας, ειδικά κατά τη σύνθεση σύνθετων τρισδιάστατων δομών (Jiang, 2024).

Επιπλέον, οι υψηλές υπολογιστικές απαιτήσεις των περισσότερων μεθόδων αποτελούν σοβαρό εμπόδιο για την πρακτική εφαρμογή τους σε πραγματικό χρόνο. Η έλλειψη αποδοτικών τεχνικών υπολογιστικής βελτιστοποίησης περιορίζει τη χρήση των μεθόδων αυτών σε περιβάλλοντα όπου απαιτείται άμεση επεξεργασία (Kivi et al., 2022). Τέλος, τα μοντέλα βαθιάς μάθησης δυσκολεύονται να γενικεύσουν σε νέα και άγνωστα δεδομένα, καθώς συχνά εξαρτώνται σε μεγάλο βαθμό από τα δεδομένα εκπαίδευσης, δυσκολεύοντας την εφαρμογή τους σε περιβάλλοντα που διαφέρουν από αυτά (Hazer & Yildirim, 2025).

ΚΕΦΑΛΑΙΟ 3

Point Cloud

3.1 Point Cloud: Ορισμός, Βασικές Αρχές και Δημιουργία σημείων νεφών.

Ένα νέφος σημείων αντιπροσωπεύει μια συλλογή δεδομένων που ορίζονται μέσα σε ένα τρισδιάστατο σύστημα συντεταγμένων. Αυτά τα σημεία περιγράφουν συλλογικά τη γεωμετρία αντικειμένων ή περιβαλλόντων, προσφέροντας μια ψηφιακή αναπαράσταση των εξωτερικών τους επιφανειών. Κάθε σημείο χαρακτηρίζεται από τις Καρτεσιανές συντεταγμένες του (X, Y, Z) και μπορεί επίσης να περιλαμβάνει πρόσθετα χαρακτηριστικά, όπως ένταση, χρώμα και ανακλαστικότητα, ανάλογα με την τεχνολογία που χρησιμοποιείται για την απόκτηση των δεδομένων (Grilli et al., 2017).

Η εφαρμογή των νεφών σημείων εκτείνεται σε διάφορους τομείς. Στην αρχιτεκτονική και την κατασκευή, τα νέφη σημείων χρησιμοποιούνται για τη δημιουργία μοντέλων πληροφορίας κτιρίων (BIM), επιτρέποντας ακριβή σχεδιασμό και παρακολούθηση προόδου (Grilli et al., 2017). Είναι επίσης αναπόσπαστα για αυτόνομα συστήματα, όπως τα αυτόνομα οχήματα, όπου τα δεδομένα νέφους σημείων σε πραγματικό χρόνο βοηθούν στον εντοπισμό αντικειμένων και την πλοήγηση σε περιβάλλοντα (Lesniak & Tucker, 2018). Επιπλέον, η δημιουργία δεδομένων νέφους σημείων είναι κρίσιμη για την καταγραφή της τρισδιάστατης γεωμετρίας αντικειμένων και περιβαλλόντων, τη διατήρηση πολιτιστικής κληρονομιάς και τις καθηλωτικές εμπειρίες εικονικής πραγματικότητας. Οι δύο κύριες μέθοδοι για τη δημιουργία δεδομένων νέφους σημείων είναι το LiDAR (Light Detection and Ranging) και η φωτογραμμετρία, καθεμία από τις οποίες προσφέρει μοναδικά πλεονεκτήματα.

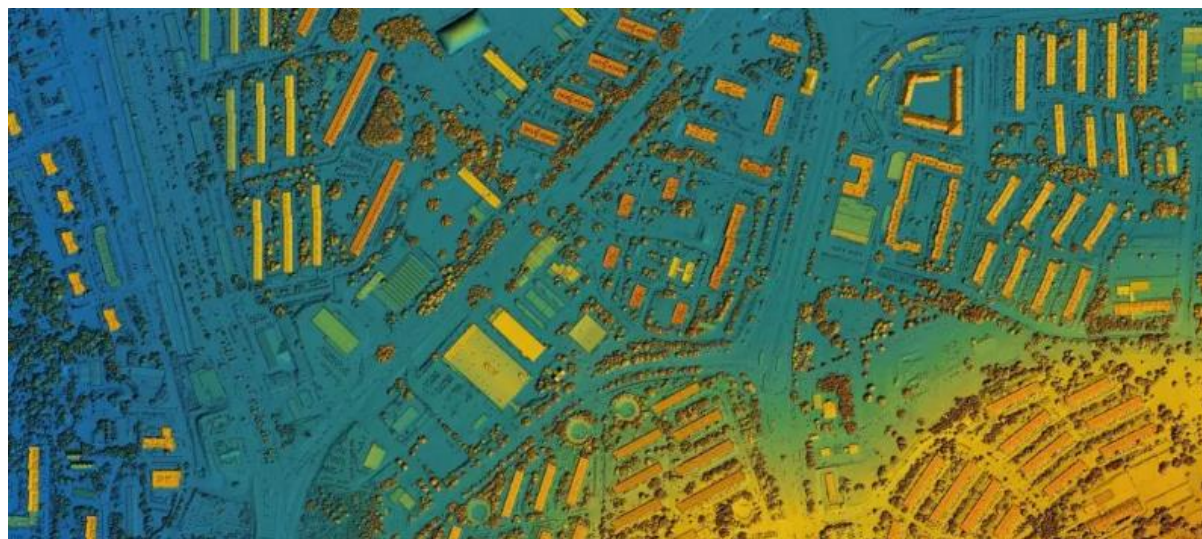
Το LiDAR χρησιμοποιεί παλμούς λέιζερ για τη μέτρηση αποστάσεων προς επιφανειακούς στόχους, επιτρέποντας τη δημιουργία ακριβών τρισδιάστατων αναπαραστάσεων. Αυτή η μέθοδος είναι ιδιαίτερα αποτελεσματική για εφαρμογές που απαιτούν υψηλής ανάλυσης χωρικά δεδομένα, όπως η τοπογραφική χαρτογράφηση και η περιβαλλοντική παρακολούθηση. Οι εξελίξεις στην τεχνολογία LiDAR έχουν βελτιώσει την ακρίβεια και την αποδοτικότητά της, καθιστώντας την αξιόπιστη επιλογή για μεγάλης κλίμακας συλλογή δεδομένων (Xu et al., 2023). Η ενσωμάτωση του LiDAR με άλλα σύνολα δεδομένων, όπως η φωτογραμμετρία, εμπλουτίζει περαιτέρω τις εφαρμογές του, βελτιώνοντας την πληρότητα και την ακρίβεια των μοντέλων (Buckley et al., 2008).

Η φωτογραμμετρία ανακατασκευάζει τρισδιάστατα μοντέλα από επικαλυπτόμενες φωτογραφίες που λαμβάνονται από διαφορετικές γωνίες. Με τη χρήση αλγορίθμων υπολογιστικής όρασης, η μέθοδος αυτή εντοπίζει κοινά χαρακτηριστικά μεταξύ των εικόνων και υπολογίζει τις χωρικές τους θέσεις μέσω τριγωνοποίησης. Είναι ιδιαίτερα χρήσιμη για την καταγραφή λεπτομερών υφών και έχει εφαρμοστεί αποτελεσματικά σε γεωλογικές μελέτες και την τεκμηρίωση πολιτιστικής κληρονομιάς (Granshaw et al., 2017). Η φωτογραμμετρία συμπληρώνει το LiDAR, παρέχοντας πρόσθετες οπτικές λεπτομέρειες, ειδικά σε περιπτώσεις όπου είναι απαραίτητα τόσο τα γεωμετρικά όσο και τα υφικά δεδομένα.

Τα δεδομένα νέφους σημείων που προκύπτουν από αυτές τις μεθόδους συχνά απαιτούν προεπεξεργασία, όπως η απομάκρυνση θορύβου και η ευθυγράμμιση. Τεχνικές ευθυγράμμισης (registration) συνδυάζουν σύνολα δεδομένων από διάφορες πηγές, διευκολύνοντας τη δημιουργία συνεκτικών και ακριβών μοντέλων (Xu et al., 2023). Οι

εργασίες μετά την επεξεργασία, όπως η τμηματοποίηση και η ανακατασκευή επιφανειών, μετατρέπουν τα ακατέργαστα νέφη σημείων σε λειτουργικά τρισδιάστατα μοντέλα.

Οι πρόσφατες εξελίξεις στη μηχανική μάθηση έχουν εισαγάγει μοντέλα βαθιάς μάθησης ικανά να δημιουργούν συνθετικά νέφη σημείων LiDAR. Αυτές οι καινοτομίες αντιμετωπίζουν προκλήσεις όπως η έλλειψη δεδομένων και είναι ιδιαίτερα χρήσιμες για την εκπαίδευση αυτόνομων συστημάτων (Zyrianov et al., 2022). Η δημιουργία δεδομένων νέφους σημείων, καθοδηγούμενη από το LiDAR και τη φωτογραμμετρία, συνεχίζει να εξελίσσεται, με τις καινοτομίες στις μεθόδους συλλογής και επεξεργασίας δεδομένων να διευρύνουν τις εφαρμογές της σε πολυάριθμους τομείς.



Στην παραπάνω φωτογραφία βλέπουμε μία πόλη μετά από σκανάρισμα του με το σύστημα LiDAR και την ψηφιακή αποθήκευση σε Point Cloud.

3.2 Δομή και Χαρακτηριστικά των Point Clouds

Ένα νέφος σημείων είναι μια αναπαράσταση χωρικών δεδομένων, αποτελούμενη από διακριτά σημεία που ορίζονται σε ένα τρισδιάστατο σύστημα συντεταγμένων. Κάθε σημείο χαρακτηρίζεται από Καρτεσιανές συντεταγμένες (X , Y , Z) και μπορεί να περιλαμβάνει πρόσθετα χαρακτηριστικά, όπως χρώμα, ένταση και διανύσματα κανονικών επιφανειών, τα οποία είναι απαραίτητα για την ακριβή γεωμετρική και υλική αναπαράσταση των σαρωμένων αντικειμένων (Hackel et al., 2016).

Τα νέφη σημείων είναι εγγενώς μη δομημένα σύνολα δεδομένων, πράγμα που σημαίνει ότι δεν περιέχουν πληροφορίες συνδεσιμότητας μεταξύ των σημείων. Αυτή η ανομοιογένεια δημιουργεί προκλήσεις για διαδικασίες όπως η τμηματοποίηση και η ταξινόμηση. Αλγόριθμοι όπως οι k -πλησιέστεροι γείτονες (k -NN) και η διαμέριση βασισμένη σε voxels χρησιμοποιούνται συνήθως για την επιβολή δομής και τη διευκόλυνση της ανάλυσης (Guo et al., 2020). Η πυκνότητα των σημείων είναι ένα κρίσιμο χαρακτηριστικό που επηρεάζει την ανάλυση και τη λεπτομέρεια των δεδομένων. Υψηλότερες πυκνότητες επιτρέπουν λεπτομερείς αναπαραστάσεις, αλλά απαιτούν σημαντικό χώρο αποθήκευσης και επεξεργαστική ισχύ (Dong et al., 2020).

Η ποιότητα των δεδομένων νέφους σημείων επηρεάζεται από τις μεθόδους συλλογής. Για παράδειγμα, το LiDAR παράγει πυκνά και ακριβή νέφη σημείων, ενώ η φωτογραμμετρία συχνά δημιουργεί λιγότερο πυκνά αλλά οπτικά πλούσια σύνολα δεδομένων (Weinmann et al., 2017). Ο θόρυβος είναι ένα άλλο σημαντικό χαρακτηριστικό, που εισάγεται συνήθως από περιβαλλοντικούς παράγοντες ή περιορισμούς αισθητήρων. Τεχνικές προεπεξεργασίας,

όπως η απομάκρυνση ανωμαλιών και τα φίλτρα εξομάλυνσης, είναι κρίσιμες για τη μείωση του θορύβου και τη βελτίωση της χρηστικότητας των δεδομένων.

Οι εφαρμογές των νεφών σημείων, όπως η τρισδιάστατη μοντελοποίηση και η χαρτογράφηση αστικών περιοχών, απαιτούν συχνά τη μετατροπή των ακατέργαστων δεδομένων σε πολυγωνικά ή τριγωνικά πλέγματα. Μέθοδοι ανακατασκευής επιφανειών, όπως η τριγωνοποίηση Delaunay και η ανακατασκευή επιφανειών Poisson, χρησιμοποιούνται ευρέως για αυτόν τον σκοπό (Kazhdan & Hoppe, 2013).

3.3 Εφαρμογές Point Clouds

Τα νέφη σημείων έχουν βρει εκτεταμένες εφαρμογές σε διάφορους τομείς, χάρη στην ευελιξία τους και την ικανότητά τους να καταγράφουν λεπτομερή χωρικά δεδομένα. Η τρισδιάστατη φύση τους τα καθιστά απαραίτητα σε πεδία όπως η κατασκευή, η ρομποτική, η πολιτιστική κληρονομιά και οι περιβαλλοντικές επιστήμες.

Τα νέφη σημείων χρησιμοποιούνται εκτενώς στη γεωχωρική ανάλυση για τη δημιουργία λεπτομερών χαρτών και μοντέλων αστικών και αγροτικών περιοχών. Τα νέφη σημείων που δημιουργούνται από LiDAR επιτρέπουν την ακριβή τοπογραφική χαρτογράφηση, κρίσιμη για τον αστικό σχεδιασμό, τη μοντελοποίηση πλημμυρών και την ανάπτυξη υποδομών (Lague et al., 2013). Η δυνατότητα του LiDAR να διεισδύει στη βλάστηση το καθιστά προτιμώμενη επιλογή για τη χαρτογράφηση εδαφών σε πυκνά δάση, συμβάλλοντας στη διαχείριση γης και τις μελέτες βιοποικιλότητας (Lewis & Hancock, 2007).

Η διατήρηση πολιτιστικών μνημείων και αντικειμένων έχει επωφεληθεί σημαντικά από τα νέφη σημείων. Η υψηλής ανάλυσης τρισδιάστατη σάρωση επιτρέπει την αναλυτική ψηφιακή τεκμηρίωση ιστορικών μνημείων, διευκολύνοντας την αποκατάσταση και την εικονική εξερεύνηση. Αυτή η προσέγγιση έχει εφαρμοστεί στην τεκμηρίωση αρχαίων μνημείων, εξασφαλίζοντας τη διατήρησή τους παρά τις περιβαλλοντικές και ανθρωπογενείς απειλές (Remondino, 2011).

Στον τομέα των κατασκευών, τα νέφη σημείων είναι αναπόσπαστο μέρος της Μοντελοποίησης Πληροφοριών Κτιρίων (BIM), παρέχοντας ακριβή μοντέλα όπως-κατασκευάστηκαν για τον σχεδιασμό και τη διαχείριση έργων. Επιτρέπουν τη σύγκριση μεταξύ σχεδιασμένων και πραγματικών κατασκευών, ενισχύοντας τον ποιοτικό έλεγχο και μειώνοντας τα κόστη (Tang et al., 2010). Επιπλέον, τα νέφη σημείων διευκολύνουν τις ανακαινίσεις, καταγράφοντας τις υπάρχουσες συνθήκες και ελαχιστοποιώντας τις αποκλίσεις κατά την κατασκευή.

Τα αυτόνομα οχήματα και η ρομποτική βασίζονται σε μεγάλο βαθμό στα νέφη σημείων για πλοήγηση και ανίχνευση αντικειμένων. Οι αισθητήρες LiDAR παράγουν νέφη σημείων σε πραγματικό χρόνο, επιτρέποντας στα ρομπότ να εντοπίζουν εμπόδια και να πλοηγούνται με ασφάλεια σε περίπλοκα περιβάλλοντα (Li et al., 2020). Στη γεωργική ρομποτική, για παράδειγμα, τα νέφη σημείων χρησιμοποιούνται για την παρακολούθηση καλλιέργειών και την εκτίμηση παραγωγής, βελτιώνοντας την παραγωγικότητα και τη βιωσιμότητα.

Τα νέφη σημείων είναι απαραίτητα για την παρακολούθηση περιβαλλοντικών αλλαγών με την πάροδο του χρόνου. Αναλύοντας διαφορές σε δεδομένα που συλλέγονται σε διαφορετικά χρονικά διαστήματα, οι ερευνητές μπορούν να παρακολουθούν την υποχώρηση παγετώνων, την αποψίλωση δασών και τη διάβρωση των ακτών (Rosser et al., 2005). Αυτή η ικανότητα υποστηρίζει μελέτες για την κλιματική αλλαγή και ενημερώνει στρατηγικές διατήρησης.

Στη βιομηχανία ψυχαγωγίας, τα νέφη σημείων διευκολύνουν τη δημιουργία καθηλωτικών τρισδιάστατων περιβαλλόντων. Τεχνολογίες όπως η φωτογραμμετρία μετατρέπουν πραγματικά αντικείμενα σε υψηλής ανάλυσης νέφη σημείων, τα οποία στη συνέχεια μετασχηματίζονται σε εικονικά στοιχεία για ταινίες και βιντεοπαιχνίδια (Stamos et al., 2008). Αυτή η μέθοδος παρέχει απaráμιλλο ρεαλισμό στα οπτικά εφέ και τον σχεδιασμό παιχνιδιών.

Σε βιομηχανίες όπως η αεροδιαστημική και η αυτοκινητοβιομηχανία, τα νέφη σημείων χρησιμοποιούνται για την εξασφάλιση ποιότητας και την αντίστροφη μηχανική. Σαρώσεις υψηλής ακρίβειας επιτρέπουν στους κατασκευαστές να εντοπίζουν ελαττώματα, να εξασφαλίζουν συμμόρφωση με τις προδιαγραφές σχεδιασμού και να βελτιστοποιούν τις διαδικασίες παραγωγής (El-Hakim et al., 2004).

3.4 Πλεονεκτήματα και Προκλήσεις των Point Clouds

Τα νέφη σημείων, ως συλλογές διακριτών σημείων δεδομένων σε τρισδιάστατο χώρο, αποτελούν ένα ισχυρό εργαλείο με πολλές εφαρμογές και πλεονεκτήματα, ενώ παράλληλα συνοδεύονται από προκλήσεις που απαιτούν προσεκτική διαχείριση. Η κύρια δύναμη των νεφών σημείων έγκειται στη δυνατότητά τους να παρέχουν υψηλής ανάλυσης αναπαραστάσεις επιφανειών αντικειμένων και περιβαλλόντων. Καταγράφοντας περίπλοκες γεωμετρικές λεπτομέρειες, καθίστανται ανεκτίμητα σε τομείς όπως η τρισδιάστατη μοντελοποίηση και η τεκμηρίωση πολιτιστικής κληρονομιάς, εξασφαλίζοντας λεπτομερείς απεικονίσεις που διευκολύνουν την ανάλυση και τη συντήρηση (Han et al., 2023). Επιπλέον, η ευελιξία στη συλλογή δεδομένων τους, μέσω μεθόδων όπως η σάρωση λέιζερ και η φωτογραμμετρία, τους επιτρέπει να προσαρμόζονται στις απαιτήσεις διαφορετικών έργων και διαθέσιμων πόρων (Tychola et al., 2024). Παράλληλα, τα νέφη σημείων αποτελούν τη βάση για προηγμένες τεχνολογίες, όπως η εικονική πραγματικότητα, η αυτόνομη πλοήγηση και η γεωργία ακρίβειας, διευρύνοντας το εύρος των εφαρμογών τους και επιτρέποντας την ακριβή μοντελοποίηση και ανάλυση περιβαλλόντων (Han et al., 2023).

Ωστόσο, τα νέφη σημείων συνοδεύονται και από σημαντικές προκλήσεις, που σχετίζονται κυρίως με την πολυπλοκότητα και το μέγεθος των δεδομένων τους. Η υψηλή πυκνότητά τους οδηγεί σε εξαιρετικά μεγάλα σύνολα δεδομένων, τα οποία παρουσιάζουν δυσκολίες στην αποθήκευση, την επεξεργασία και τη μετάδοσή τους, απαιτώντας αποτελεσματικές στρατηγικές διαχείρισης για την αντιμετώπιση αυτών των ζητημάτων (Zhuang et al., 2024). Επίσης, τα νέφη σημείων χαρακτηρίζονται από την έλλειψη δομής, καθώς δεν διαθέτουν πληροφορίες συνδεσιμότητας μεταξύ των σημείων, γεγονός που περιπλέκει διαδικασίες όπως η τμηματοποίηση, η ταξινόμηση και η αναγνώριση αντικειμένων. Για την ερμηνεία αυτών των δεδομένων, απαιτούνται προηγμένοι αλγόριθμοι που να μπορούν να διαχειριστούν αυτήν τη μη δομημένη φύση τους (Han et al., 2023).

Επιπλέον, τα δεδομένα νεφών σημείων συχνά έχουν ατέλειες και θόρυβο, που προκύπτουν από κρυφές περιοχές κατά τη διαδικασία σάρωσης ή από περιορισμούς στους αισθητήρες. Η αντιμετώπιση αυτών των ζητημάτων είναι κρίσιμη για την ακριβή τρισδιάστατη ανακατασκευή και ανάλυση, καθώς οι θόρυβοι και οι ελλείψεις μπορούν να επηρεάσουν τη συνολική αξιοπιστία των αποτελεσμάτων (Han et al., 2023).

Τέλος, η επεξεργασία νεφών σημείων μεγάλης κλίμακας απαιτεί σημαντικούς υπολογιστικούς πόρους, δημιουργώντας περιορισμούς, ιδιαίτερα σε εφαρμογές που βασίζονται σε δεδομένα σε πραγματικό χρόνο. Η συνεχής έρευνα επικεντρώνεται στη βελτιστοποίηση αλγορίθμων ώστε να είναι αποδοτικοί χωρίς να θυσιάζεται η ακρίβεια, εξασφαλίζοντας ότι τα νέφη σημείων μπορούν να αξιοποιηθούν σε ένα ευρύτερο φάσμα εφαρμογών (Zhuang et al., 2024). Η δυναμική των νεφών σημείων, σε συνδυασμό με τις εξελισσόμενες τεχνολογίες, εγγυάται την περαιτέρω αξιοποίησή τους σε ποικίλους τομείς, ενώ οι λύσεις στις προκλήσεις που τα συνοδεύουν ανοίγουν το δρόμο για ακόμα πιο καινοτόμες εφαρμογές.

3.5 Επεξεργασία Δεδομένων Point Clouds

Η επεξεργασία δεδομένων νέφους σημείων περιλαμβάνει μια σειρά βημάτων που αποσκοπούν στη μετατροπή ακατέργαστων χωρικών δεδομένων σε χρήσιμη πληροφορία για διάφορες εφαρμογές. Τα βασικά στάδια περιλαμβάνουν τη συλλογή δεδομένων, την

προεπεξεργασία, την ευθυγράμμιση, την τμηματοποίηση και την εξαγωγή χαρακτηριστικών:

❖ Συλλογή Δεδομένων

Τα νέφη σημείων αποκτώνται συνήθως μέσω τεχνολογιών όπως το LiDAR και η φωτογραμμετρία. Η επιλογή της μεθόδου συλλογής επηρεάζει την πυκνότητα και την ακρίβεια των δεδομένων νέφους σημείων.

❖ Προεπεξεργασία

Η προεπεξεργασία περιλαμβάνει τον καθαρισμό των ακατέργαστων δεδομένων νέφους σημείων για την απομάκρυνση θορύβου και ανωμαλιών. Τεχνικές όπως η στατιστική ανάλυση και τα φίλτρα εφαρμόζονται για τη βελτίωση της ποιότητας των δεδομένων. Για παράδειγμα, οι μέθοδοι απομάκρυνσης στατιστικών ανωμαλιών αξιολογούν την κατανομή των αποστάσεων μεταξύ των σημείων για την ανίχνευση και εξάλειψη ανωμαλιών, βελτιώνοντας την αξιοπιστία της ανάλυσης.

❖ Ευθυγράμμιση

Η ευθυγράμμιση (registration) ευθυγραμμίζει πολλά νέφη σημείων σε ένα ενιαίο συνεκτικό σύστημα συντεταγμένων. Αυτό είναι απαραίτητο κατά τον συνδυασμό δεδομένων από διαφορετικές σαρώσεις ή αισθητήρες. Αλγόριθμοι όπως ο Iterative Closest Point (ICP) χρησιμοποιούνται συχνά για αυτόν τον σκοπό, μειώνοντας επαναληπτικά την απόσταση μεταξύ αντιστοιχών σημείων για την επίτευξη βέλτιστης ευθυγράμμισης.

❖ Τμηματοποίηση

Η τμηματοποίηση χωρίζει το νέφος σημείων σε διακριτές περιοχές ή αντικείμενα βάσει γεωμετρικών ή χωρικών ιδιοτήτων. Αυτό το βήμα είναι κρίσιμο για την απομόνωση χαρακτηριστικών ενδιαφέροντος μέσα σε περίπλοκα σύνολα δεδομένων. Μέθοδοι όπως η ανάπτυξη περιοχών (region growing) και η ομαδοποίηση (clustering) αναλύουν την εγγύτητα και την ομοιότητα των σημείων για την ομαδοποίησή τους.

❖ Εξαγωγή Χαρακτηριστικών

Η εξαγωγή χαρακτηριστικών εντοπίζει και ποσοτικοποιεί συγκεκριμένα χαρακτηριστικά μέσα στο νέφος σημείων, όπως άκρα, επιφάνειες ή υφές. Αυτά τα χαρακτηριστικά είναι καθοριστικά για εφαρμογές όπως η αναγνώριση αντικειμένων και η κατανόηση σκηνών. Προηγμένοι αλγόριθμοι αναλύουν τη χωρική διάταξη και τις σχέσεις μεταξύ των σημείων για την ανίχνευση σημαντικών προτύπων και δομών. (Mirzaei et al. 2022)

3.6 Μελλοντικές Τάσεις στην Τεχνολογία Point Clouds

Η τεχνολογία νεφών σημείων εξελίσσεται ραγδαία, με αρκετές αναδυόμενες τάσεις να ενισχύουν τις δυνατότητες και τις εφαρμογές της. Οι πιο σημαντικές εξελίξεις περιλαμβάνουν την ενσωμάτωση αλγορίθμων βαθιάς μάθησης, τη βελτίωση των μεθόδων συλλογής δεδομένων και την ανάπτυξη αποδοτικών αλγορίθμων για επεξεργασία σε πραγματικό χρόνο.

Μία από τις σημαντικότερες εξελίξεις είναι η ενσωμάτωση αλγορίθμων βαθιάς μάθησης για την ταξινόμηση τρισδιάστατων νεφών σημείων. Τα μοντέλα βαθιάς μάθησης έχουν αποδείξει ανώτερη απόδοση στην επεξεργασία μη δομημένων δεδομένων νεφών σημείων, οδηγώντας σε πιο ακριβή αναγνώριση αντικειμένων και κατανόηση σκηνών. Αυτή η εξέλιξη διευκολύνει προηγμένες εφαρμογές, όπως τα αυτόνομα οχήματα και η επαυξημένη πραγματικότητα. (Zhang et al., 2023).

Η πρόοδος στις τεχνολογίες LiDAR και φωτογραμμετρίας επιτρέπει την απόκτηση υψηλότερης ανάλυσης νεφών σημείων. Οι καινοτομίες αυτές οδηγούν στη συλλογή πιο λεπτομερών χωρικών πληροφοριών, οι οποίες είναι κρίσιμες για εφαρμογές όπως η αυτόνομη οδήγηση, η εικονική πραγματικότητα και η χαρτογράφηση ακριβείας. Η αύξηση της ανάλυσης συμβάλλει στη δημιουργία ακριβέστερων μοντέλων και στη βελτίωση της ανάλυσης πολύπλοκων περιβαλλόντων. (Behroozpour et.al. 2017)

Ένας τομέας έντονου ερευνητικού ενδιαφέροντος είναι η ανάπτυξη αποδοτικών αλγορίθμων για την επεξεργασία μεγάλων νεφών σημείων σε πραγματικό χρόνο. Αυτή η τάση είναι ιδιαίτερα σημαντική για τη ρομποτική και τις διαδραστικές εφαρμογές, όπου η έγκαιρη ερμηνεία των δεδομένων είναι κρίσιμη. Η βελτιστοποίηση των διαδικασιών επεξεργασίας επιτρέπει την ταχύτερη λήψη αποφάσεων και τη βελτίωση της λειτουργικότητας σε πραγματικό χρόνο.

4.1 Εισαγωγή στη Μετατροπή Κειμένου σε Νέφη Σημείων

Η μετατροπή κειμένου σε νέφη σημείων (Text-to-Point-Cloud) είναι μια καινοτόμος τεχνολογία που επιτρέπει τη δημιουργία τρισδιάστατων αναπαραστάσεων βασισμένων σε περιγραφές φυσικής γλώσσας. Αυτή η διαδικασία γεφυρώνει το χάσμα μεταξύ της κείμενης περιγραφής και της τρισδιάστατης απεικόνισης, αξιοποιώντας προηγμένες τεχνικές μηχανικής μάθησης και βαθιάς μάθησης.

Η διαδικασία Text-to-Point-Cloud αναφέρεται στη δημιουργία τρισδιάστατων νεφών σημείων από δομημένες ή μη περιγραφές κειμένου. Ένα νέφος σημείων αποτελείται από διακριτά σημεία σε έναν τρισδιάστατο χώρο που περιγράφουν τη γεωμετρία ενός αντικειμένου ή ενός περιβάλλοντος. Με τη χρήση μοντέλων βαθιάς μάθησης, οι περιγραφές κειμένου μπορούν να μετατραπούν σε ακριβείς τρισδιάστατες αναπαραστάσεις, διευκολύνοντας εφαρμογές που απαιτούν σύνθετη χωρική ανάλυση (Zhu et al., 2024).

Η τεχνολογία Text-to-Point-Cloud έχει σχεδιαστεί για να απλοποιήσει τη δημιουργία και επεξεργασία τρισδιάστατων δεδομένων, εξαλείφοντας την ανάγκη χρήσης παραδοσιακών εργαλείων, όπως οι σαρωτές LiDAR ή η φωτογραμμετρία. Πρόκειται για μια καινοτόμο προσέγγιση που προσφέρει σημαντικά πλεονεκτήματα, όπως η αυξημένη αποδοτικότητα, η δυνατότητα δημιουργίας συνθετικών δεδομένων και η ευελιξία στη διαχείριση ποικίλων περιγραφών κειμένου, καθιστώντας την ιδανική για πληθώρα εφαρμογών.

Μια βασική εφαρμογή της τεχνολογίας είναι στα αυτόνομα συστήματα και τη ρομποτική, όπου χρησιμοποιείται για την κατανόηση του περιβάλλοντος και την πλοήγηση. Με τη μετατροπή περιγραφών κειμένου σε τρισδιάστατα μοντέλα, η τεχνολογία επιτρέπει τη λήψη αποφάσεων σε πραγματικό χρόνο, προσφέροντας στα αυτόνομα συστήματα τη δυνατότητα να προσαρμόζονται δυναμικά στις ανάγκες του περιβάλλοντος (Feng et al., 2024).

Στον τομέα της αρχιτεκτονικής και της εικονικής πραγματικότητας, η τεχνολογία Text-to-Point-Cloud συμβάλλει σημαντικά στην απλοποίηση της δημιουργίας τρισδιάστατων αναπαραστάσεων. Οι περιγραφές σχεδίων μετατρέπονται σε λεπτομερείς τρισδιάστατες δομές, διευκολύνοντας την οπτικοποίηση, την ανάλυση και την προσαρμογή σχεδίων. Στην εικονική πραγματικότητα, η τεχνολογία χρησιμοποιείται για τη δημιουργία καθηλωτικών περιβαλλόντων, ενισχύοντας τη ρεαλιστική απεικόνιση και τη διαδραστικότητα (Shang et al., 2024).

Ένας άλλος σημαντικός τομέας εφαρμογής είναι η εκπαίδευση μοντέλων βαθιάς μάθησης, όπου η τεχνολογία χρησιμοποιείται για τη σύνθεση συνθετικών δεδομένων. Αυτή η δυνατότητα είναι ιδιαίτερα σημαντική, καθώς επιτρέπει τη δημιουργία μεγάλων συνόλων δεδομένων που δεν περιορίζονται από φυσικές μετρήσεις, ενισχύοντας τη διαδικασία εκπαίδευσης και βελτιώνοντας την απόδοση των μοντέλων (Shang et al., 2024).

Παρά τη σημαντική πρόοδο, η τεχνολογία Text-to-Point-Cloud βρίσκεται ακόμα σε φάση εξέλιξης. Οι κύριες προκλήσεις περιλαμβάνουν τη βελτίωση της γενίκευσης των μοντέλων, ώστε να μπορούν να ανταποκρίνονται σε διαφορετικές και απρόβλεπτες περιγραφές, καθώς και την αύξηση της ακρίβειας κατά τη μετατροπή των δεδομένων. Παρόλα αυτά, η συνεχιζόμενη έρευνα και οι τεχνολογικές εξελίξεις αναμένεται να ξεπεράσουν τα εμπόδια αυτά, προσφέροντας νέες δυνατότητες και εφαρμογές που θα αλλάξουν τον τρόπο με τον οποίο διαχειριζόμαστε τα τρισδιάστατα δεδομένα.

4.2 Μεθοδολογία

Η δημιουργία νεφών σημείων από περιγραφές κειμένου (Text-to-Point-Cloud) απαιτεί μια συστηματική προσέγγιση που περιλαμβάνει πολλαπλά βήματα και τη χρήση εξειδικευμένων αλγορίθμων και μοντέλων βαθιάς μάθησης για τη σύνθεση τρισδιάστατων γεωμετρικών δεδομένων. Η διαδικασία μετατροπής κειμένου σε νέφη σημείων περιλαμβάνει τρία κύρια στάδια, καθένα από τα οποία συμβάλλει στη δημιουργία τρισδιάστατων δεδομένων υψηλής ποιότητας.

Το πρώτο στάδιο είναι η ανάλυση κειμένου (Text Parsing), όπου οι περιγραφές κειμένου εισάγονται στο σύστημα και αναλύονται μέσω τεχνικών Επεξεργασίας Φυσικής Γλώσσας (Natural Language Processing - NLP). Στο στάδιο αυτό, χρησιμοποιούνται εργαλεία όπως το BERT (Bidirectional Encoder Representations from Transformers), το οποίο εκπαιδεύεται σε μεγάλα σύνολα δεδομένων και εξάγει εννοιολογικά χαρακτηριστικά από το κείμενο. Αυτά τα χαρακτηριστικά μπορεί να περιλαμβάνουν γεωμετρικές πληροφορίες, χρώματα, σχέσεις και χωρικές ιδιότητες που αποτελούν τη βάση για τη δημιουργία τρισδιάστατων δομών (Radford et al., 2021). Η δυνατότητα του BERT να κατανοεί τη σύνταξη και τη σημασιολογία είναι κρίσιμη για τη σύνδεση του περιεχομένου του κειμένου με τις γεωμετρικές δομές που θα παραχθούν (Devlin, 2018).

Το δεύτερο στάδιο είναι η γενετική μοντελοποίηση (Generative Modeling). Τα χαρακτηριστικά που εξάγονται από το κείμενο τροφοδοτούνται σε μοντέλα βαθιάς μάθησης, όπως τα Variational Autoencoders (VAEs) και τα Generative Adversarial Networks (GANs). Τα GANs δημιουργούν νέφη σημείων με υψηλή λεπτομέρεια, χρησιμοποιώντας ένα γενετικό δίκτυο για την παραγωγή δεδομένων και ένα διακριτικό δίκτυο για την αξιολόγηση της ποιότητάς τους (Goodfellow et al., 2014). Τα VAEs, από την άλλη, επικεντρώνονται στη δημιουργία των βασικών γεωμετρικών χαρακτηριστικών, ελαχιστοποιώντας τα σφάλματα κατά τη διαδικασία σύνθεσης (Kingma, 2013). Οι μέθοδοι αυτές επιτρέπουν τη διατήρηση των χαρακτηριστικών που περιγράφονται στο κείμενο, ενώ παρέχουν ένα σημείο εκκίνησης για τη δημιουργία σύνθετων δομών (Nichol & Dhariwal, 2021).

Το τρίτο και τελικό στάδιο είναι δημιουργία και η βελτιστοποίηση νέφους σημείων (Point Cloud Refinement). Εδώ, η αρχική δομή που δημιουργείται από τα μοντέλα δημιουργίας περνά από μια διαδικασία βελτίωσης μία διαδικασία φιλτραρίσματος δηλαδή και κανονικοποίησης για τη βελτίωση της ακρίβειας και της ποιότητας. Τεχνικές όπως το Point Refinement εφαρμόζονται για την απομάκρυνση θορύβου, τη μείωση γεωμετρικών σφαλμάτων και την κανονικοποίηση κλιμάκων, ώστε τα δεδομένα να είναι συνεκτικά και υψηλής ποιότητας (Nath et al., 2024). Εργαλεία όπως το PointNet++ χρησιμοποιούνται για τη διαχείριση των νεφών σημείων, βελτιώνοντας τη γεωμετρία και την ακρίβεια της τελικής αναπαράστασης (Hu & Xie, 2024).

Η διαδικασία περιλαμβάνει και σημαντικές τεχνικές προεπεξεργασίας και κανονικοποίησης δεδομένων, οι οποίες διασφαλίζουν την ακρίβεια και την ποιότητα των παραγόμενων δεδομένων. Το φιλτράρισμα θορύβου εφαρμόζεται για την απομάκρυνση ανεπιθύμητων δεδομένων που μπορεί να υποβαθμίσουν την ακρίβεια του μοντέλου. Επιπλέον, η κανονικοποίησης κλιμάκων διασφαλίζει ότι όλα τα σημεία βρίσκονται εντός ενός συγκεκριμένου χώρου, όπως το εύρος $[-1, 1]$, διευκολύνοντας τη διαχείριση και επεξεργασία τους. Η ανάπτυξη συμμετρίας είναι επίσης κρίσιμη για την αποφυγή ασυμμετριών ή ανακρίβειών που μπορεί να προκύψουν από τις μετατροπές, διασφαλίζοντας τη συνοχή της τελικής αναπαράστασης (Xue et al., 2023). Με τη συνδυασμένη χρήση αυτών των σταδίων, η μετατροπή κειμένου σε νέφη σημείων επιτυγχάνει υψηλή ακρίβεια, λεπτομέρεια και ποιότητα, δημιουργώντας ευέλικτα τρισδιάστατα δεδομένα για ποικίλες εφαρμογές.

Η εκπαίδευση μοντέλων Text-to-Point-Cloud απαιτεί εξειδικευμένα σύνολα δεδομένων που συνδέουν περιγραφές κειμένου με αντίστοιχα νέφη σημείων. Για παράδειγμα, το Text2Shape είναι ένα σύνολο δεδομένων που περιέχει συνδέσεις μεταξύ περιγραφών και τριοδιάστατων γεωμετρικών σχημάτων, σχεδιασμένο για την ανάπτυξη μοντέλων σύνθεσης (Chen et al., 2019). Το ShapeCaption3D περιλαμβάνει λεπτομερείς περιγραφές τριοδιάστατων αντικειμένων, επιτρέποντας την καλύτερη σύνδεση μεταξύ κειμένου και γεωμετρίας (Leberl et al., 2010). Επιπλέον, το PartNet χρησιμοποιείται για την κατανόηση και την αναπαραγωγή σύνθετων γεωμετρικών σχημάτων, βοηθώντας στη δημιουργία νέφους σημείων που αντικατοπτρίζουν πιστά τις περιγραφές (Mo et al., 2019).

Η εφαρμογή του Text2Shape έχει δείξει εντυπωσιακά αποτελέσματα. Περιγραφές όπως "καρέκλα με τέσσερα πόδια και ψηλή πλάτη" οδηγούν στη δημιουργία ρεαλιστικών νεφών σημείων που αναπαριστούν το περιγραφόμενο αντικείμενο, με γεωμετρική ακρίβεια που φτάνει το 93% (Amaduzzi et al., 2023). Με τη χρήση Diffusion Models, σύνθετες περιγραφές όπως "διώροφο κτίριο με στρογγυλές κολόνες" αποδίδουν νέφη σημείων που περιλαμβάνουν λεπτομερείς και πιστές γεωμετρικές δομές (Nichol & Dhariwal, 2021). Επιπλέον, η βελτιστοποίηση μέσω Point Refinement βελτιώνει την ποιότητα των δεδομένων κατά 20%, μειώνοντας θόρυβο και γεωμετρικά σφάλματα (Luo et al., 2023).

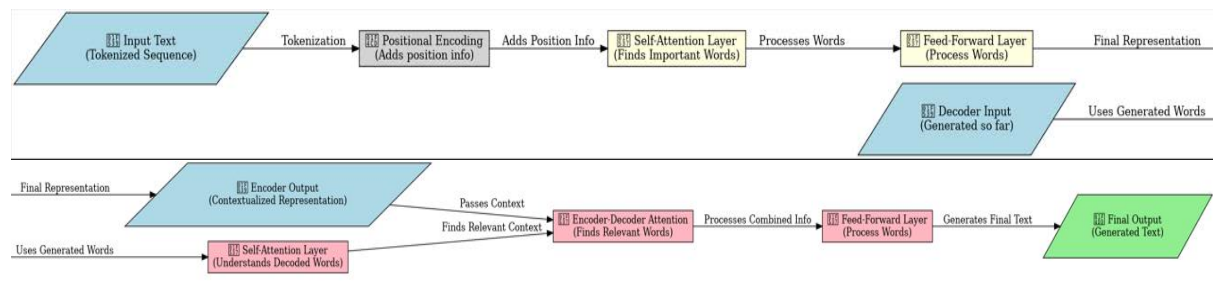
Η συνολική διαδικασία, από την ανάλυση κειμένου έως τη βελτιστοποίηση, υποστηρίζεται από προηγμένα μοντέλα και εξειδικευμένα σύνολα δεδομένων, διασφαλίζοντας ότι τα παραγόμενα νέφη σημείων είναι ακριβή, λεπτομερή και κατάλληλα για ποικίλες εφαρμογές.

4.3 Αρχιτεκτονικές Μοντέλων

Η ανάπτυξη αποτελεσματικών μοντέλων βαθιάς μάθησης είναι πολύ σημαντική στη διαδικασία μετατροπής κειμένου σε νέφη σημείων, καθώς αυτά τα μοντέλα εξασφαλίζουν την εξαγωγή εννοιολογικών χαρακτηριστικών από το κείμενο και τη δημιουργία τριοδιάστατων αναπαραστάσεων. Βασισμένα σε σύγχρονες αρχιτεκτονικές, όπως οι Transformers, τα Generative Adversarial Networks (GANs) και τα Diffusion Models, τα μοντέλα αυτά αποτελούν την αιχμή της τεχνολογίας για τη σύνθεση νεφών σημείων από περιγραφές κειμένου.

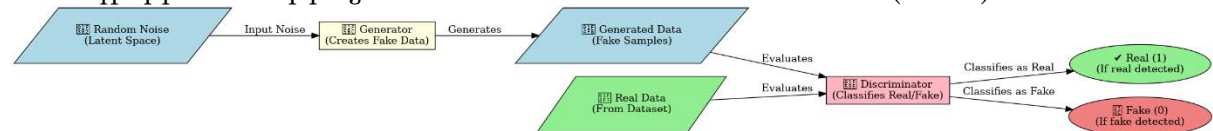
Οι Transformers είναι αρχιτεκτονικές που βασίζονται σε μηχανισμούς προσοχής και επιτρέπουν την κατανόηση των συσχετίσεων μεταξύ λέξεων, μοντελοποιώντας σύνθετες σχέσεις. Στην εφαρμογή Text-to-Point-Cloud, μοντέλα όπως το Text-to-3D Transformer αξιοποιούν ενσωματωμένες πληροφορίες από το κείμενο για τη δημιουργία νεφών σημείων με υψηλή ακρίβεια και λεπτομέρεια (Vaswani, 2017). Οι Transformers διακρίνονται για την ικανότητά τους να επεξεργάζονται μεγάλα σετ δεδομένων και να προσαρμόζονται σε διαφορετικούς τύπους εισόδου και εξόδου, καθιστώντας τα ιδανικά για σύνθετες εφαρμογές. Ωστόσο, παρουσιάζουν περιορισμούς, όπως οι μεγάλες απαιτήσεις σε υπολογιστική ισχύ και δεδομένα εκπαίδευσης. Οι Transformers αποτελούνται από δύο βασικά μέρη, Τον Encoder και τον Decoder. Ο Encoder λαμβάνει την είσοδο και δημιουργεί μία ενδιάμεση αναπαράσταση και ο Decoder παίρνει τα ενδιάμεσα δεδομένα και παράγει την τελική έξοδο.

Το διάγραμμα λειτουργίας των Transformers:



Τα Generative Adversarial Networks (GANs) περιλαμβάνουν δύο νευρωνικά δίκτυα: ένα γενετικό δίκτυο που δημιουργεί δεδομένα και ένα διακριτικό που τα αξιολογεί. Στην περίπτωση της μετατροπής κειμένου σε νέφη σημείων, τα GANs έχουν αποδείξει τη δυνατότητά τους να δημιουργούν γεωμετρικές δομές που αντιστοιχούν στις περιγραφές κειμένου (Feng et al., 2024). Η δυνατότητα των GANs να παράγουν δεδομένα υψηλής ποιότητας και να αντιμετωπίζουν την ανεπάρκεια δεδομένων μέσω εξομίωσης τα καθιστά ιδιαίτερα χρήσιμα. Παρ' όλα αυτά, τα GANs είναι ευαίσθητα σε αστάθειες κατά τη διάρκεια της εκπαίδευσης και δυσκολεύονται να γενικεύσουν για πιο σύνθετες περιγραφές.

Το διάγραμμα λειτουργίας των Generative Adversarial Networks (GANs):



Τα Diffusion Models είναι μια πρόσφατη προσέγγιση που χρησιμοποιεί στοχαστικές διεργασίες για τη δημιουργία τρισδιάστατων δεδομένων από κείμενο. Αυτά τα μοντέλα λειτουργούν με σταδιακή προσέγγιση, δημιουργώντας σύνθετα νέφη σημείων μέσω κατευθυνόμενης βελτίωσης (Ho et al., 2020). Διακρίνονται για την εξαιρετική ακρίβεια και λεπτομέρεια που επιτυγχάνουν, ενώ απαιτούν μικρότερη εξάρτηση από παραδοσιακά αρχιτεκτονικά χαρακτηριστικά. Ωστόσο, οι Diffusion Models έχουν αυξημένες απαιτήσεις χρόνου εκπαίδευσης, γεγονός που περιορίζει τη χρήση τους σε εφαρμογές όπου η ταχύτητα είναι κρίσιμη.

Η επιλογή της κατάλληλης αρχιτεκτονικής εξαρτάται από τις απαιτήσεις της εφαρμογής. Οι Transformers και τα GANs είναι κατάλληλα για περιπτώσεις που απαιτούν υψηλή ευελιξία και προσαρμοστικότητα, ενώ τα diffusion models προσφέρουν μεγαλύτερη ακρίβεια για πολύπλοκες δομές. Παρόλα αυτά, η βελτιστοποίηση της υπολογιστικής αποδοτικότητας παραμένει μια πρόκληση για όλες τις αρχιτεκτονικές.

4.4 Αξιολόγηση και Μετρικές Απόδοσης

Η αξιολόγηση της ποιότητας των παραγόμενων νεφών σημείων είναι κρίσιμη για την αποτίμηση της απόδοσης και της ακρίβειας των μοντέλων Text-to-Point-Cloud. Οι μεθοδολογίες αξιολόγησης και οι μετρικές που χρησιμοποιούνται παρέχουν ένα ποσοτικό πλαίσιο για τη σύγκριση και την ανάλυση των αποτελεσμάτων.

Η αξιολόγηση της ποιότητας των παραγόμενων νεφών σημείων αποτελεί βασικό στοιχείο της ανάπτυξης και βελτίωσης μοντέλων σύνθεσης. Η διαδικασία περιλαμβάνει ποσοτικές και ποιοτικές μεθοδολογίες που στοχεύουν στη μέτρηση της ακρίβειας, της πληρότητας, της πυκνότητας, αλλά και της αντιληπτής ποιότητας των παραγόμενων δεδομένων.

Μία από τις κύριες προσεγγίσεις είναι η σύγκριση με χειροκίνητα επισημασμένα δεδομένα ή υφιστάμενα σύνολα υψηλής ακρίβειας. Αυτή η μέθοδος παρέχει μια άμεση εκτίμηση της ακρίβειας και της πληρότητας των παραγόμενων νεφών σημείων, καθώς τα δεδομένα συγκρίνονται με βάση αναφορές που θεωρούνται το "Golden Prototype" (Kolmet et al., 2022).

Η χρήση μετρικών ομοιότητας, όπως το Chamfer Distance και το Earth Mover's Distance (EMD), προσφέρει μια πιο τυποποιημένη ποσοτική αξιολόγηση. Το Chamfer Distance υπολογίζει τη μέση απόσταση μεταξύ των σημείων του παραγόμενου και του πραγματικού νέφους, μετρώντας την ακρίβεια σε τοπικό επίπεδο. Αντίθετα, το Earth Mover's Distance αξιολογεί τη συνολική "εργασία" που απαιτείται για να μετατραπεί το παραγόμενο νέφος στο πραγματικό, παρέχοντας μια ολιστική εκτίμηση της γεωμετρικής ομοιότητας (Fan et al., 2017).

Η ανάλυση πληρότητας και πυκνότητας εξετάζει την κάλυψη του παραγόμενου νέφους σημείων και την ομοιομορφία της κατανομής των σημείων. Η πληρότητα μετρά εάν το παραγόμενο νέφος καλύπτει πλήρως τη γεωμετρία του αντικειμένου, ενώ η πυκνότητα αξιολογεί αν τα σημεία είναι ομοιόμορφα κατανεμημένα, κάτι που είναι κρίσιμο για εφαρμογές όπως η ρομποτική και η αρχιτεκτονική (Fei et al., 2022).

Σε περιπτώσεις όπου η ποσοτική αξιολόγηση δεν είναι αρκετή, χρησιμοποιείται η υποκειμενική αξιολόγηση. Παρατηρητές αξιολογούν την ποιότητα των παραγόμενων νεφών σημείων με βάση την αντιληπτή ακρίβεια και τη λεπτομέρεια, παρέχοντας πολύτιμα δεδομένα για την προσαρμογή των μοντέλων (Yu et al., 2024). Η απόδοση των μοντέλων συγκρίνεται συνήθως με βάση τέσσερις βασικές μετρικές:

- Ακρίβεια (Precision): Αναφέρεται στο ποσοστό των σημείων του παραγόμενου νέφους που ταιριάζουν με το πραγματικό. Υψηλή ακρίβεια δείχνει ότι το μοντέλο αποφεύγει την παραγωγή θορύβου.
- Πληρότητα (Recall): Μετρά την ικανότητα του μοντέλου να αναπαράγει όλα τα σημεία του πραγματικού νέφους. Υψηλή πληρότητα υποδηλώνει πλήρη κάλυψη του αντικειμένου από το παραγόμενο νέφος.
- F1-Score: Συνδυάζει την ακρίβεια και την πληρότητα για μια συνολική εκτίμηση της απόδοσης, εξισορροπώντας την αναπαραγωγή δεδομένων και την αποφυγή θορύβου (Yang et al., 2024).
- Χρόνος Επεξεργασίας: Η ταχύτητα δημιουργίας των νεφών σημείων είναι κρίσιμη, ιδιαίτερα για εφαρμογές πραγματικού χρόνου. Τα μοντέλα που επιτυγχάνουν χαμηλότερους χρόνους επεξεργασίας είναι πιο κατάλληλα για πρακτική χρήση.

Η συγκριτική ανάλυση δείχνει ότι τα diffusion models υπερέχουν σε γεωμετρική ακρίβεια (Fan et al., 2017), ενώ τα GANs προσφέρουν καλύτερη πυκνότητα και πληρότητα (Shu et al., 2019). Οι Transformers παρέχουν ισορροπία μεταξύ ακρίβειας και ταχύτητας, καθιστώντας τα κατάλληλα για εφαρμογές πραγματικού χρόνου. Παρά τη σημαντική πρόοδο στον τομέα της μετατροπής κειμένου σε νέφη σημείων (Text-to-Point-Cloud), εξακολουθούν να υπάρχουν αρκετοί περιορισμοί και προκλήσεις που επηρεάζουν την απόδοση, την ακρίβεια και τη γενίκευση των υφιστάμενων μεθοδολογιών.

Ένα από τα κύρια ζητήματα είναι η έλλειψη μεγάλων και ποικίλων συνόλων δεδομένων που συνδέουν περιγραφές κειμένου με νέφη σημείων. Τα υφιστάμενα δεδομένα είναι συχνά περιορισμένα σε μέγεθος και ποικιλία, γεγονός που δυσχεραίνει τη δυνατότητα των μοντέλων να γενικεύουν σε νέες εφαρμογές και αντικείμενα. Η ανάπτυξη μεγαλύτερων και πιο αντιπροσωπευτικών συνόλων δεδομένων είναι κρίσιμη για τη βελτίωση της απόδοσης των μοντέλων (Ma et al., 2023).

Η υπολογιστική πολυπλοκότητα αποτελεί επίσης σημαντική πρόκληση. Οι μέθοδοι που βασίζονται σε γενετικά μοντέλα, όπως τα GANs και τα diffusion models, απαιτούν υψηλή υπολογιστική ισχύ τόσο κατά την εκπαίδευση όσο και κατά την παραγωγή. Αυτή η ανάγκη περιορίζει τη δυνατότητα χρήσης των μοντέλων σε πραγματικό χρόνο, ειδικά σε εφαρμογές

που απαιτούν άμεση απόκριση (Kolmet et al., 2022). Παράλληλα, παρατηρείται ελλιπής χωρική και γεωμετρική ανάλυση, με πολλά μοντέλα να αποτυγχάνουν να αποδώσουν σύνθετες γεωμετρικές λεπτομέρειες, όπως καμπύλες ή πολύπλοκες τοπολογίες. Αυτό αποτελεί σημαντικό περιορισμό για εφαρμογές υψηλής ακρίβειας, όπως η ρομποτική και η αρχιτεκτονική (Chen et al., 2024).

Η δυσκολία γενίκευσης σε νέα δεδομένα είναι ένας ακόμα κρίσιμος περιορισμός. Τα περισσότερα μοντέλα Text-to-Point-Cloud εκπαιδεύονται σε περιορισμένα και συγκεκριμένα δεδομένα, γεγονός που δυσχεραίνει την απόδοσή τους σε νέους τύπους αντικειμένων ή περιγραφών. Η έλλειψη ποικιλίας στα δεδομένα εκπαίδευσης εντείνει αυτό το πρόβλημα (Bao et al., 2024). Παράλληλα, οι περιγραφές κειμένου συχνά περιλαμβάνουν ασάφειες ή ελλείψεις πληροφοριών, καθιστώντας δύσκολη τη σύνθεση ακριβών νεφών σημείων. Οι υπάρχουσες μεθοδολογίες απαιτούν καλύτερους μηχανισμούς κατανόησης της φυσικής γλώσσας, ώστε να ανταποκρίνονται πιο αποτελεσματικά σε τέτοιου είδους ασαφή δεδομένα (Martinez-Carranza et al., 2023).

Παρότι εφαρμόζονται τεχνικές κανονικοποίησης και βελτιστοποίησης, η ακρίβεια και η λεπτομέρεια των παραγόμενων νεφών σημείων παραμένουν προβληματικές. Η γεωμετρική αναπαράσταση συχνά παρουσιάζει σφάλματα, περιορίζοντας τη χρησιμότητα των δεδομένων σε εφαρμογές που απαιτούν υψηλή πιστότητα (Tanksale, 2023).

Η αντιμετώπιση αυτών των προκλήσεων απαιτεί πολυεπίπεδη προσέγγιση. Η ανάπτυξη μεγαλύτερων και πιο ποικίλων συνόλων δεδομένων μπορεί να ενισχύσει τη γενίκευση των μοντέλων, ενώ η βελτιστοποίηση των αλγορίθμων για χαμηλότερη υπολογιστική πολυπλοκότητα θα διευκολύνει την εφαρμογή τους σε πραγματικό χρόνο. Επιπλέον, η ενσωμάτωση πολυτροπικών προσεγγίσεων, όπως ο συνδυασμός εικόνων και κειμένου, μπορεί να συμβάλει στη βελτίωση της ακρίβειας και της αξιοπιστίας των παραγόμενων νεφών σημείων, καθιστώντας την τεχνολογία πιο πρακτική και αποδοτική.

4.5 Εφαρμογές

Η τεχνολογία μετατροπής κειμένου σε νέφη σημείων (Text-to-Point-Cloud) βρίσκεται εφαρμογή σε πολλούς τομείς, αξιοποιώντας τη δυνατότητα δημιουργίας τρισδιάστατων γεωμετρικών αναπαραστάσεων από περιγραφές κειμένου και παρέχοντας νέες δυνατότητες για τη δημιουργία τρισδιάστατων δεδομένων από περιγραφές φυσικής γλώσσας. Αυτή η καινοτομία επιτρέπει την αυτοματοποίηση και την αναβάθμιση διεργασιών σε τομείς όπως η αρχιτεκτονική, η ρομποτική και η εικονική πραγματικότητα.

Στον τομέα της αρχιτεκτονικής και του σχεδιασμού, η τεχνολογία αυτή επιτρέπει στους αρχιτέκτονες να μετατρέπουν περιγραφές σχεδίων σε λεπτομερή τρισδιάστατα μοντέλα. Με αυτόν τον τρόπο, διευκολύνεται η ταχεία δημιουργία πρωτοτύπων και η ανάλυση σχεδίων, εξοικονομώντας χρόνο και πόρους. Για παράδειγμα, μια περιγραφή όπως "κτίριο με πέντε ορόφους και στρογγυλά παράθυρα" μπορεί να μετατραπεί άμεσα σε μοντέλο υψηλής ακρίβειας, επιταχύνοντας σημαντικά τη διαδικασία σχεδιασμού (Bono, 2024).

Στον τομέα της ρομποτικής, τα παραγόμενα νέφη σημείων χρησιμοποιούνται για την πλοήγηση και την αντίληψη περιβάλλοντος από αυτόνομα συστήματα. Ρομπότ που βασίζονται σε περιγραφές κειμένου μπορούν να δημιουργούν χάρτες σε πραγματικό χρόνο, επιτρέποντας την αποφυγή εμποδίων και τη λήψη αποφάσεων με ακρίβεια. Αυτή η δυνατότητα είναι ιδιαίτερα χρήσιμη σε περιβάλλοντα χωρίς προηγούμενη χαρτογράφηση ή σε δυναμικές συνθήκες, όπου οι χάρτες πρέπει να ενημερώνονται συνεχώς (Τσαγκάρης, 2014).

Στον χώρο της εικονικής και επαυξημένης πραγματικότητας (VR/AR), η τεχνολογία Text-to-Point-Cloud επιτρέπει τη δημιουργία καθηλωτικών περιβαλλόντων που βασίζονται σε φυσικές περιγραφές. Οι χρήστες μπορούν να περιηγούνται και να αλληλεπιδρούν με προσομοιωμένα τρισδιάστατα περιβάλλοντα που παράγονται από περιγραφές όπως "δάσος με

μεγάλα δέντρα και ποτάμι". Αυτή η εφαρμογή είναι εξαιρετικά πολύτιμη για την ψυχαγωγία, την εκπαίδευση και την εκπόνηση προσομοιώσεων (Michel et al., 2022).

Η χρήση της τεχνολογίας επεκτείνεται επίσης στην εκπαίδευση και τις προσομοιώσεις, όπου τα παραγόμενα νέφη σημείων χρησιμεύουν για την εκπαίδευση συστημάτων βαθιάς μάθησης και τη δημιουργία συνθετικών δεδομένων. Αυτά τα δεδομένα χρησιμοποιούνται σε τομείς όπως η γεωγραφία, η αυτόνομη οδήγηση και άλλες εφαρμογές που απαιτούν μεγάλη ποικιλία προσομοιωμένων σεναρίων (Wang et al., 2023).

Η τεχνολογία Text-to-Point-Cloud αναδεικνύεται ως ισχυρό εργαλείο που ενισχύει τη δημιουργικότητα, την αποδοτικότητα και την ακρίβεια σε πολλούς τομείς, μετατρέποντας περιγραφές φυσικής γλώσσας σε τρισδιάστατες αναπαραστάσεις με ευρύ φάσμα εφαρμογών.

ΚΕΦΑΛΑΙΟ 5 Εμπειρικό Μέρος

5.1 Κορμός του προγράμματος

Η δημιουργία Νέφους Σημείων με κατανόηση σχήματος είναι ένα νευρωνικό δίκτυο σχεδιασμένο να συνθέτει τριδιάστατα νέφη σημείων ενώ ταυτόχρονα ταξινομεί το αντίστοιχο σχήμα. Το μοντέλο αποτελείται από τρεις κύριες ενότητες: τον κωδικοποιητή κειμένου, τον ταξινομητή σχημάτων και τη γεννήτρια σημείων.

Ο κωδικοποιητής κειμένου μετασχηματίζει τις περιγραφές κειμένου σε χαρακτηριστικές αναπαραστάσεις υψηλών διαστάσεων. Χρησιμοποιεί μια σειρά γραμμικών επιπέδων διαχωρισμένων με ενεργοποιήσεις ReLU, μειώνοντας τελικά τη διαστασιμότητα των εισόδων σε έναν χώρο κατάλληλο για περαιτέρω επεξεργασία. Ο ταξινομητής σχημάτων χρησιμοποιεί τα κωδικοποιημένα χαρακτηριστικά για να προβλέψει την κατηγορία του σχήματος, παρέχοντας πιθανότητες για δέκα προκαθορισμένες κατηγορίες σχημάτων, όπως κύβος, σφαίρα, κύλινδρος, κ.ά. Τέλος, η ενότητα γεννήτριας σημείων χρησιμοποιεί τα κωδικοποιημένα χαρακτηριστικά για να παράγει τριδιάστατες συντεταγμένες, επεκτείνοντας τις διαστάσεις των χαρακτηριστικών ώστε να ταιριάζουν με τον αριθμό των σημείων στο επιθυμητό νέφος σημείων.

Αυτή η αρχιτεκτονική διασφαλίζει κοινή αναπαράσταση χαρακτηριστικών μεταξύ της ταξινόμησης σχημάτων και της παραγωγής σημείων, επιτρέποντας στο μοντέλο να μάθει μια ενοποιημένη κατανόηση και των δύο εργασιών.

Παρακάτω θα δούμε το class diagram του προγράμματος και τα βήματα που πραγματοποιεί κατά την εκτέλεσή του:



1. **Text Input:** Ο χρήστης παρέχει μια περιγραφή κειμένου (π.χ., "έναν κύβο").
2. **AutoTokenizer:** Το AutoTokenizer μετατρέπει την περιγραφή κειμένου σε tokens, τα οποία είναι αριθμητικές αναπαραστάσεις των λέξεων.
3. **Text Model (BERT):** Το προ-εκπαιδευμένο μοντέλο BERT επεξεργάζεται τα tokens και παράγει μία ανάλυση των χαρακτηριστικών κειμένου (Text Features).
4. **ShapeAwarePointCloudGenerator:** Το μοντέλο γεννήτριας λαμβάνει τα χαρακτηριστικά κειμένου και παράγει δύο εξόδους: ένα σημείο νέφους (Point Cloud) και λογικές σχήματος (Shape Logits).
5. **MSELoss & CrossEntropyLoss:** Υπολογίζονται οι συναρτήσεις απώλειας για το σημείο νέφους (χρησιμοποιώντας MSELoss) και για την ταξινόμηση σχήματος (χρησιμοποιώντας CrossEntropyLoss).
6. **Total Loss:** Οι δύο απώλειες συνδυάζονται για να δημιουργηθεί η συνολική απώλεια.
7. **Adam Optimizer:** Ο αλγόριθμος βελτιστοποίησης Adam χρησιμοποιείται για την ενημέρωση των παραμέτρων του μοντέλου γεννήτριας με βάση τη συνολική απώλεια.
8. **Visualization (matplotlib):** Το σημείο νέφους απεικονίζεται χρησιμοποιώντας τη βιβλιοθήκη matplotlib.
9. **Shape Classification:** Οι λογικές σχήματος χρησιμοποιούνται για την ταξινόμηση του σχήματος.

5.2 Σύνολο Δεδομένων

Για την αποτελεσματική εκπαίδευση του μοντέλου, δημιουργήθηκε ένα συνθετικό σύνολο δεδομένων που συνδυάζει περιγραφές κειμένου με αντίστοιχα νέφη σημείων και σχετικές ετικέτες σχημάτων. Το σύνολο δεδομένων που ανακτώνται μέσω του PyTorch περιλαμβάνει δέκα υποστηριζόμενα σχήματα: κύβος, σφαίρα, κύλινδρος, πυραμίδα, κώνος, τόρος, αστέρι, δίσκος, πολυέδρο και ελλειψοειδές.

Κάθε σχήμα δημιουργείται χρησιμοποιώντας ένα συγκεκριμένο μαθηματικό μοντέλο. Μια σφαίρα δημιουργείται χρησιμοποιώντας σφαιρικές συντεταγμένες, ενώ ένας τόρος ορίζεται με παραμετρικές εξισώσεις για τη γεωμετρία του. Αυτές οι μαθηματικές διατυπώσεις διασφαλίζουν ποικιλομορφία και ακρίβεια στην αναπαράσταση διαφορετικών σχημάτων.

Για παράδειγμα στον κύβο θεωρούμε για κέντρο το σημείο (0,0,0) και πλευρά μήκους 2. Αυτό σημαίνει ότι οι συντεταγμένες κάθε σημείου (x,y,z) στον κύβο θα βρίσκονται στην περιοχή [-1,1]. Δηλαδή όταν θα θέλουμε να δημιουργήσουμε ένα τυχαίο σημείο και κατά συνέπεια πολλών τυχαίων σημείων για την αναπαράσταση του κύβου αυτά θα έχουν την μορφή x[-1,1], για y[-1,1] και για z[-1,1].

Κάθε νέφος σημείων συνδέεται επίσης με μια περιγραφή κειμένου, όπως "Ένα σχήμα σφαίρας" ή "Ένα σχήμα τόρου." Αυτές οι περιγραφές καθοδηγούν το μοντέλο κατά την εκπαίδευση, επιτρέποντάς του να δημιουργήσει συνδέσεις μεταξύ της εισόδου κειμένου και των τρισδιάστατων γεωμετρικών δεδομένων. Το σύνολο δεδομένων αποτελείται από 1000 δείγματα, με κάθε δείγμα να περιλαμβάνει 2048 σημεία καταναμημένα στην επιφάνεια του σχήματος. Η κατανομή των κλάσεων διατηρείται ομοιόμορφη για την αποφυγή μεροληψίας προς οποιαδήποτε κατηγορία σχημάτων.

5.3 Πρωτόκολλο Εκπαίδευσης

Η διαδικασία εκπαίδευσης ξεκινά με τη φόρτωση δεδομένων, όπου μια προσαρμοσμένη κλάση PyTorch Dataset διαχειρίζεται την προεπεξεργασία και τη διαχείριση των συνθετικών δεδομένων. Αυτή η κλάση προβαίνει στην αναγνώριση λεξικών μονάδων των περιγραφών κειμένου χρησιμοποιώντας τις ιδιότητες της αναγνώρισης λεξικών μονάδων του

λεκτικού μοντέλου BERT και ομαλοποιεί τα δεδομένα των νεφών σημείων για συνοχή. Κάθε δείγμα περιλαμβάνει μια ομαλοποιημένη περιγραφή κειμένου, ένα τρισδιάστατο νέφος σημείων και μια αντίστοιχη ετικέτα σχήματος.

Η εκπαίδευση βελτιστοποιεί δύο στόχους: την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (MSE) μεταξύ των προβλεπόμενων και των στοχευμένων νεφών σημείων, και τη μείωση της απώλειας λόγω της διασταυρούμενης εντροπίας για την ταξινόμηση των σχημάτων. Ο αλγόριθμος Adam χρησιμοποιείται για βελτιστοποίηση λόγω των προσαρμοστικών δυνατοτήτων εκμάθησής του, με ρυθμό εκμάθησης ορισμένο στο $1e-4$ για σταθερή σύγκλιση. Η εκπαίδευση διεξάγεται σε n epochs (όπου n ο αριθμός των επαναλήψεων) με μέγεθος παρτίδας 16 δειγμάτων, επιτυγχάνοντας ισορροπία μεταξύ υπολογιστικής απόδοσης και απόδοσης μοντέλου. (Esraa Hassan. 28 September 2022)

Κατά τη διάρκεια κάθε επανάληψης εκπαίδευσης, οι περιγραφές κειμένου κάνουν tokenization και μετατρέπονται σε Tensor representation, οι οποίοι περνούν από το προεκπαιδευμένο μοντέλο BERT για την εξαγωγή αναπαραστάσεων από το CLS token. Οι αναπαραστάσεις αυτές επεξεργάζονται από τον κωδικοποιητή κειμένου για την παραγωγή κωδικοποιημένων χαρακτηριστικών. Αυτά τα χαρακτηριστικά χρησιμοποιούνται από τη γεννήτρια σημείων για την πρόβλεψη του νέφους σημείων και από τον ταξινομητή σχημάτων για την πρόβλεψη της κατηγορίας του σχήματος. Οι απώλειες και για τις δύο εργασίες υπολογίζονται και οι βαθμίδες επιστρέφονται για την ενημέρωση των βαρών του μοντέλου. Μετρικές όπως η απώλεια σημείων και η απώλεια σχημάτων καταγράφονται για την παρακολούθηση της προόδου της εκπαίδευσης.

5.4 Αξιολόγηση και Οπτικοποίηση

Η διαδικασία αξιολόγησης περιλαμβάνει τον υπολογισμό του μέσου τετραγωνικού σφάλματος (MSE) μεταξύ των παραγόμενων και των στοχευμένων νεφών σημείων, καθώς και τη μέτρηση της ακρίβειας στην ταξινόμηση σχημάτων. Επιπλέον, μια προσαρμοσμένη συνάρτηση οπτικοποίησης εμφανίζει τα παραγόμενα νέφη σημείων σε τρισδιάστατο χώρο, με το βάθος-ύψος να αναπαρίσταται μέσω ενός χρωματικού χάρτη. Αυτή η μέθοδος επιτρέπει ποιοτική αξιολόγηση των παραγόμενων σχημάτων και βοηθά στον εντοπισμό πιθανών προβλημάτων στις προβλέψεις του μοντέλου.

Οι χρήστες μπορούν να εισάγουν μια περιγραφή, όπως "Ένα σχήμα τόρου," η οποία περιγραφή αναγνωρίζεται και χωρίζεται σε λεξικές μονάδες οι οποίες κωδικοποιούνται από το προεκπαιδευμένο μοντέλο BERT. Τα κωδικοποιημένα χαρακτηριστικά περνούν στη συνέχεια από τη γεννήτρια που έχει εκπαιδευτεί, για να παραχθεί ένα νέφος σημείων και να προβλεφθεί το αντίστοιχο σχήμα.

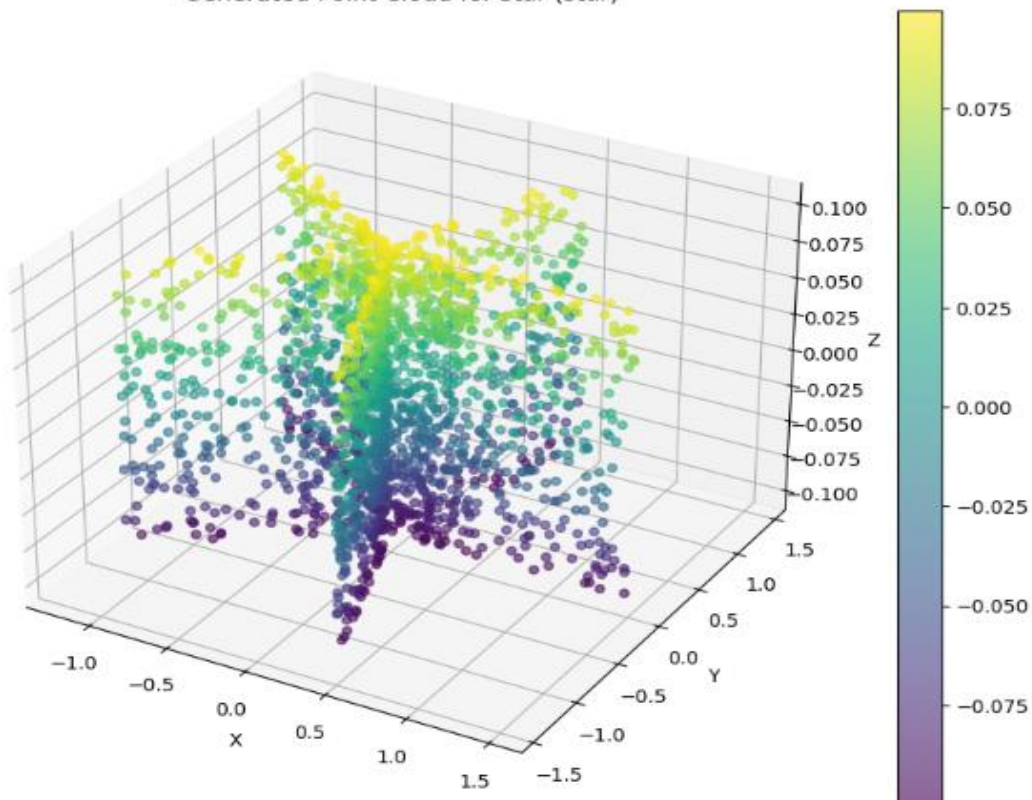
Εάν το προβλεπόμενο σχήμα ευθυγραμμίζεται με την περιγραφή της εισόδου, το μοντέλο δημιουργεί και οπτικοποιεί ένα νέφος σημείων που αντιστοιχεί στο συγκεκριμένο σχήμα. Διαφορετικά, οπτικοποιείται το παραγόμενο νέφος σημείων με μια γενική ετικέτα. Οι διαδραστικές οπτικοποιήσεις επιτρέπουν στους χρήστες να εξερευνήσουν τα παραγόμενα νέφη σημείων με λεπτομέρεια, παρέχοντας έναν διασθητικό τρόπο αξιολόγησης της απόδοσης του μοντέλου.

Παρακάτω θα δούμε τα τελικά αποτελέσματα του προγράμματος:

Epoch 1/10, Point Loss: 0.32912770434031413, Shape Loss: 2.288019687410385
 Epoch 2/10, Point Loss: 0.3156924323430137, Shape Loss: 1.6952827544439406
 Epoch 3/10, Point Loss: 0.2781885282860862, Shape Loss: 0.4056632329547216
 Epoch 4/10, Point Loss: 0.2443300975220544, Shape Loss: 0.07117795349941367
 Epoch 5/10, Point Loss: 0.24303372204303741, Shape Loss: 0.01748931369492932
 Epoch 6/10, Point Loss: 0.24225518272982705, Shape Loss: 0.008205986736963192
 Epoch 7/10, Point Loss: 0.24345392580070194, Shape Loss: 0.004841075884917426
 Epoch 8/10, Point Loss: 0.2424738314889726, Shape Loss: 0.0032183260376018193
 Epoch 9/10, Point Loss: 0.24251265279830447, Shape Loss: 0.002317031850993988
 Epoch 10/10, Point Loss: 0.24204451578950126, Shape Loss: 0.0017347021216881416

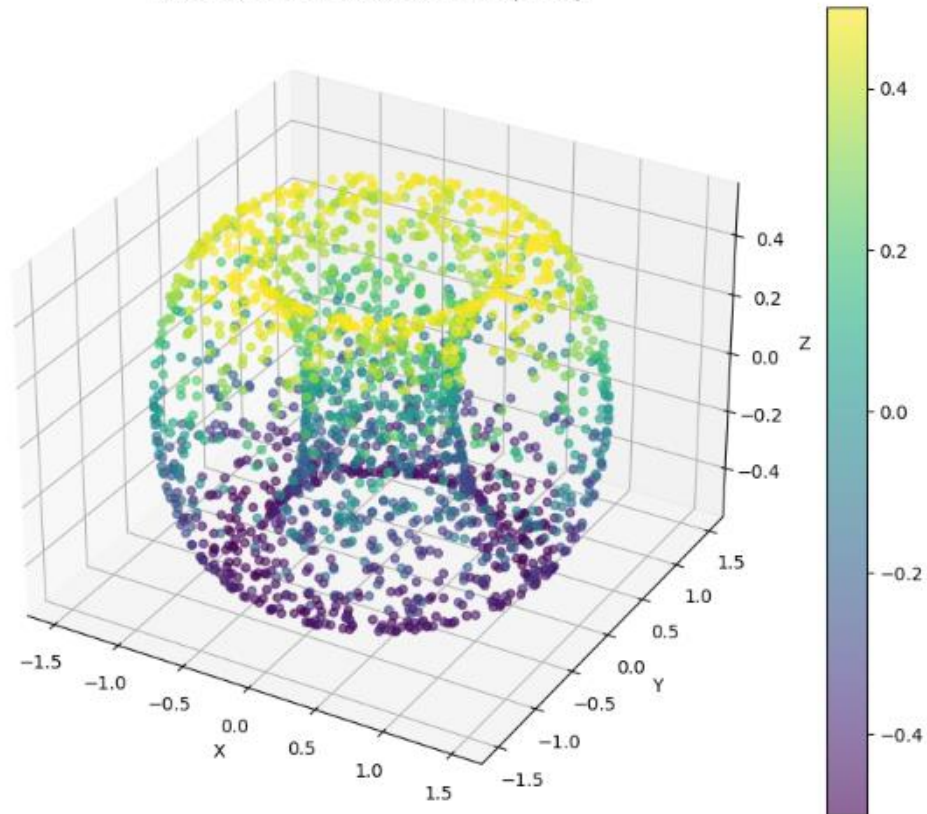
Enter the description or quit for exit: star
 Visualizing: star

Generated Point Cloud for star (star)



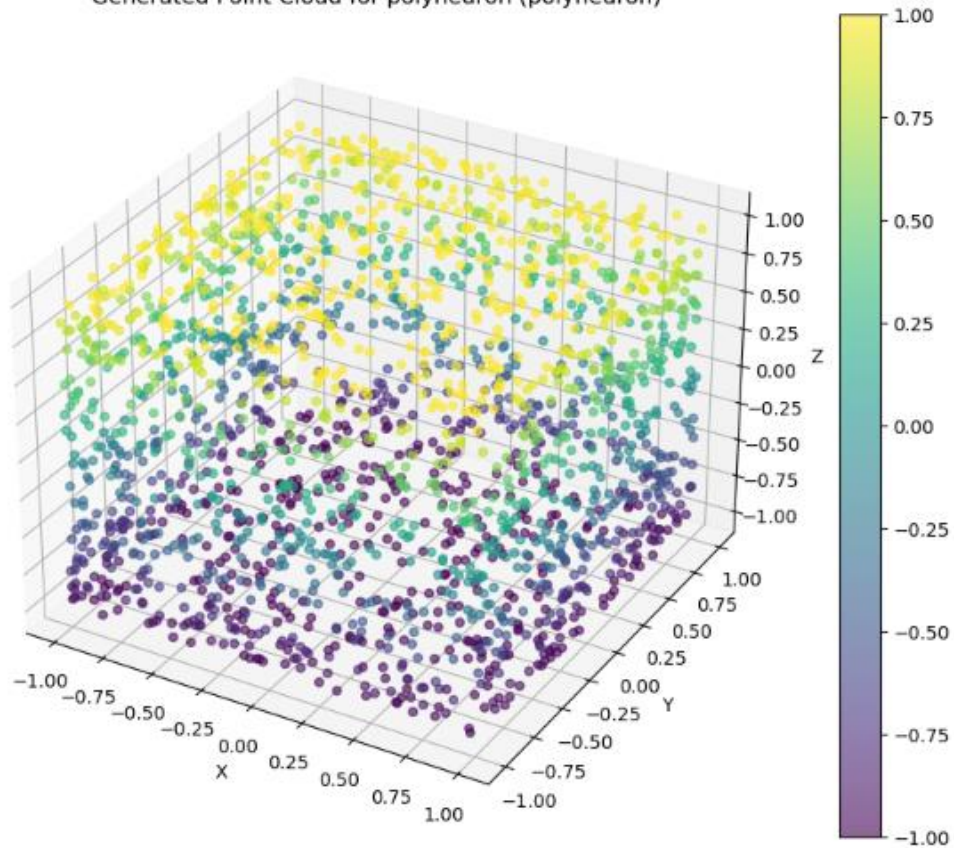
Enter the description or quit for exit: torus
Visualizing: torus

Generated Point Cloud for torus (torus)



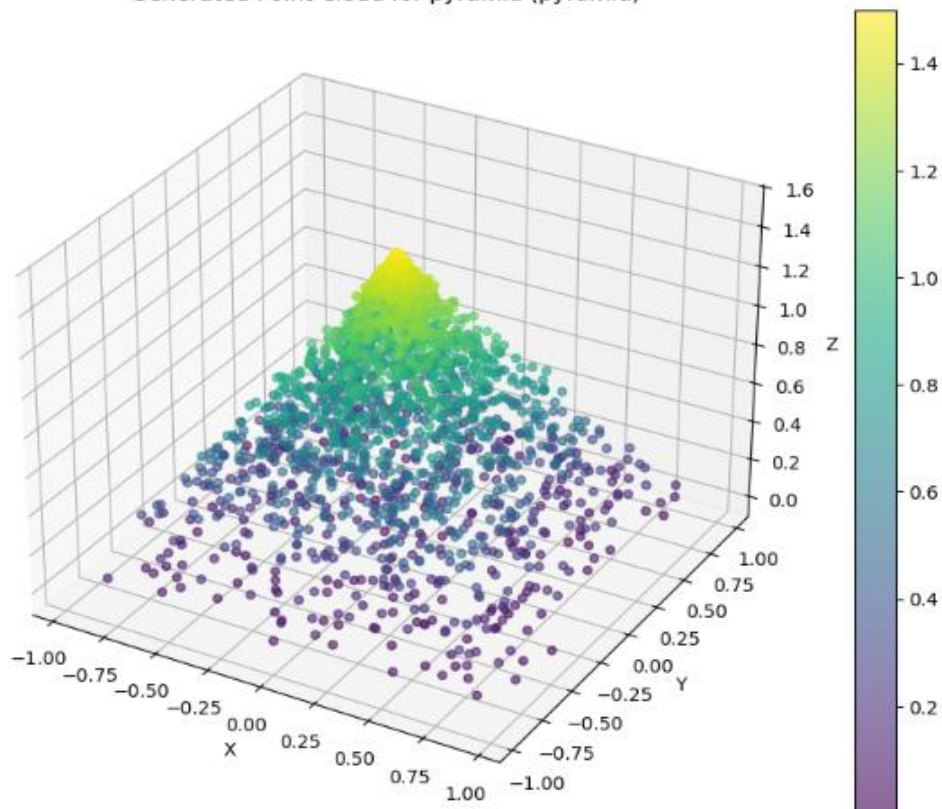
Enter the description or quit for exit: polyhedron
Visualizing: polyhedron

Generated Point Cloud for polyhedron (polyhedron)



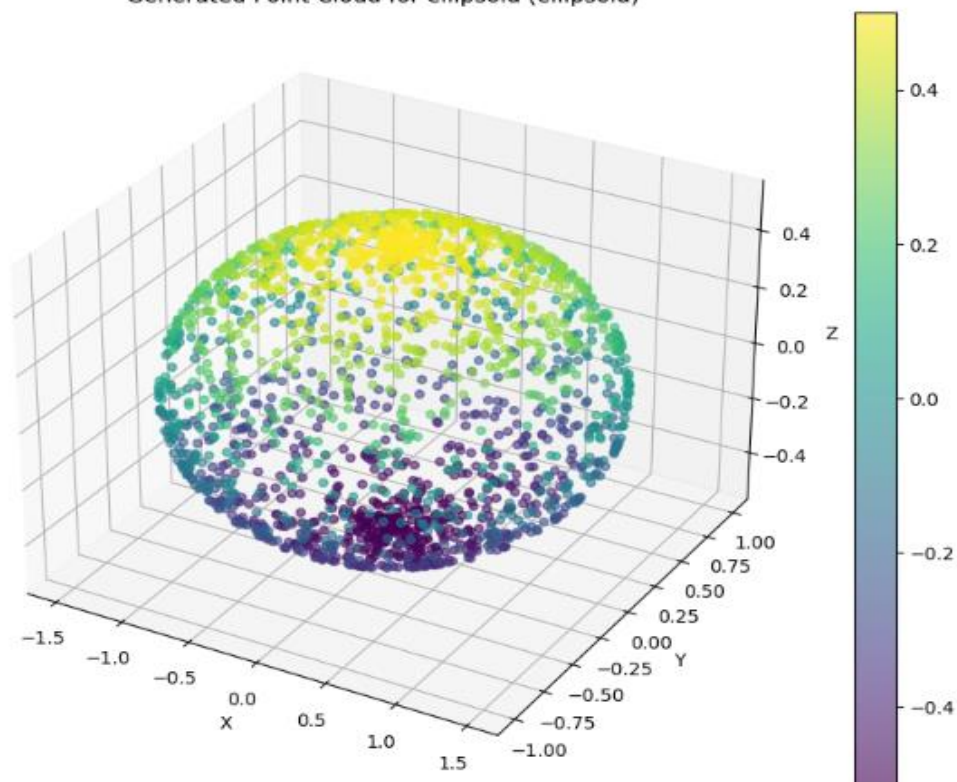
Enter the description or quit for exit: pyramid
Visualizing: pyramid

Generated Point Cloud for pyramid (pyramid)



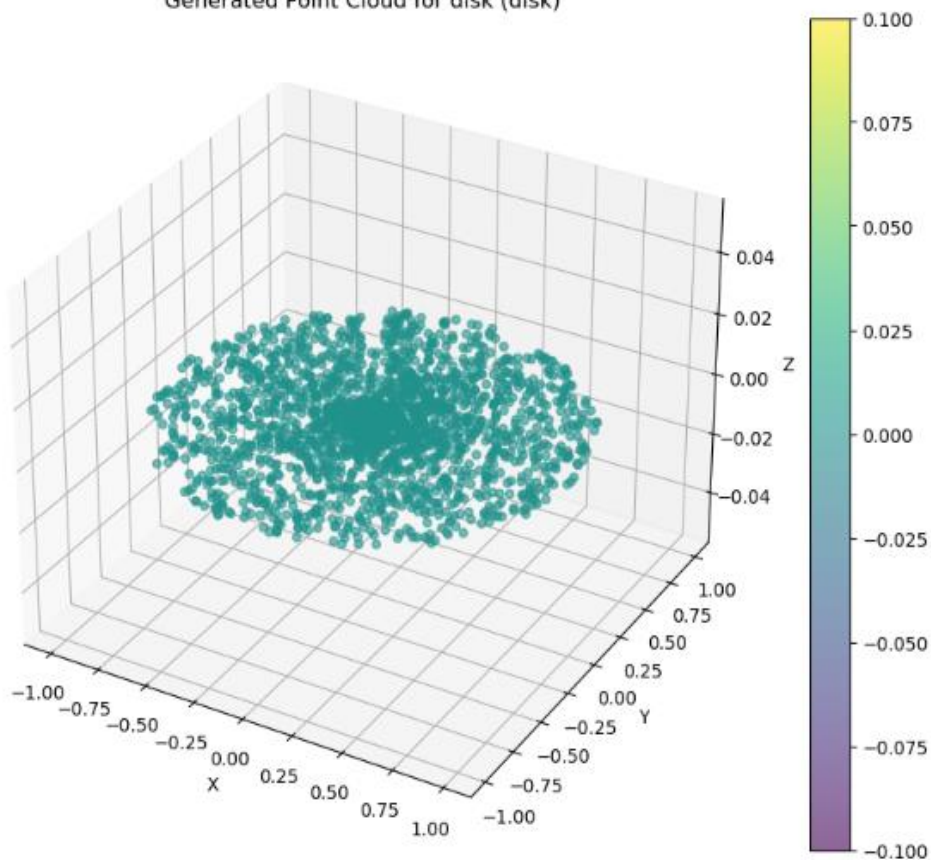
Enter the description or quit for exit: ellipsoid
Visualizing: ellipsoid

Generated Point Cloud for ellipsoid (ellipsoid)



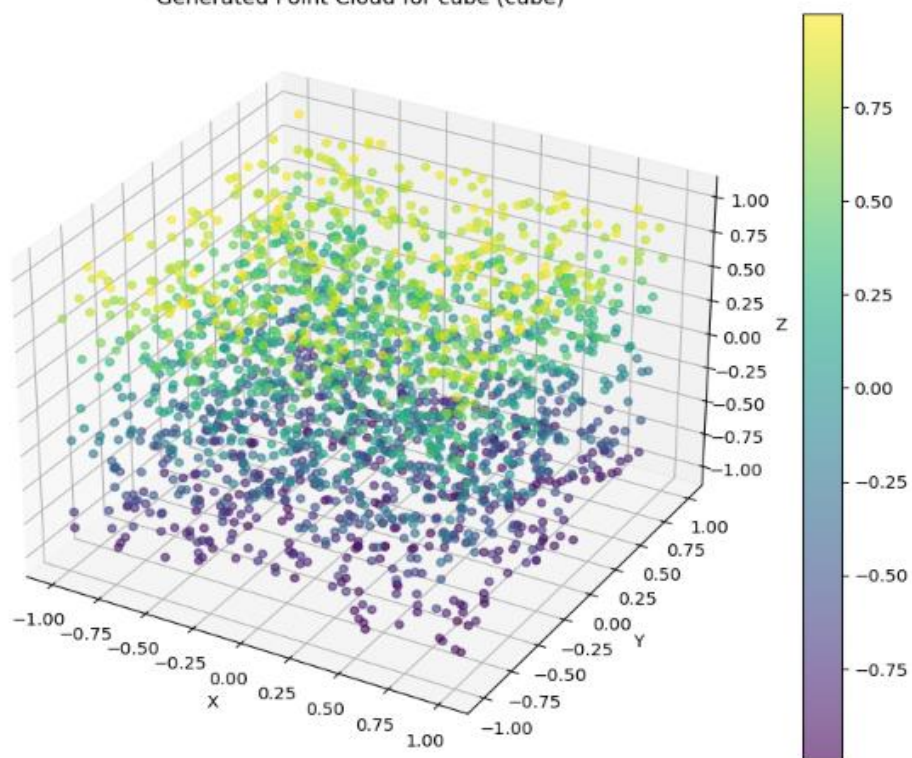
Enter the description or quit for exit: disk
Visualizing: disk

Generated Point Cloud for disk (disk)



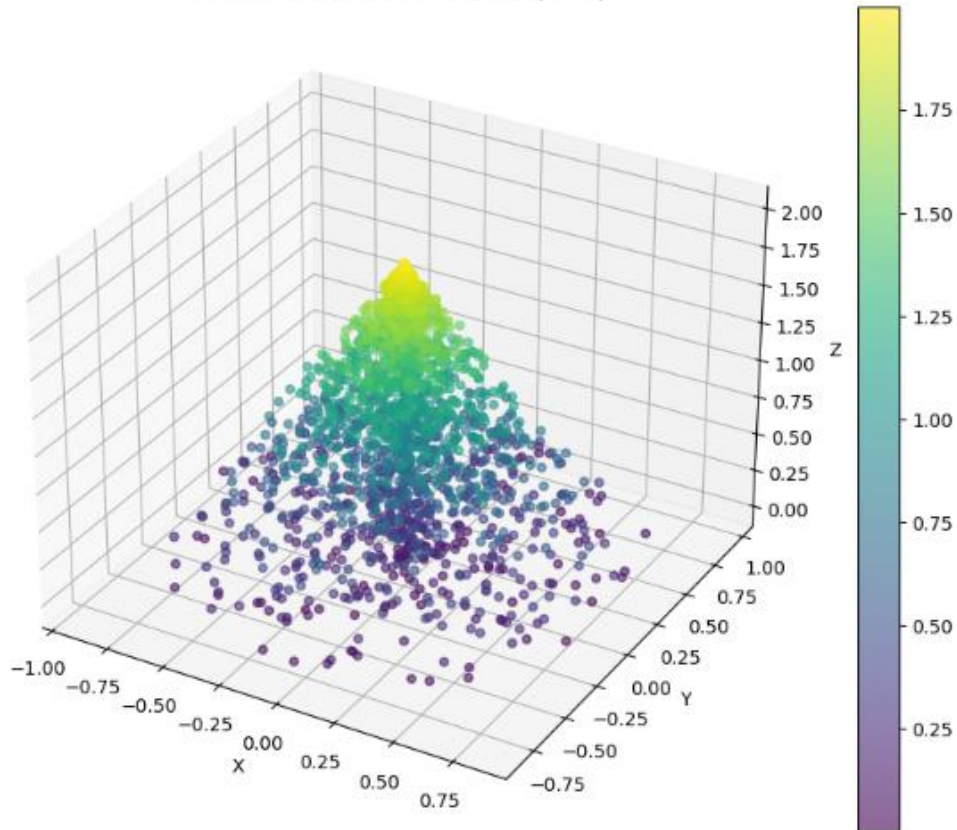
Enter the description or quit for exit: cube
Visualizing: cube

Generated Point Cloud for cube (cube)



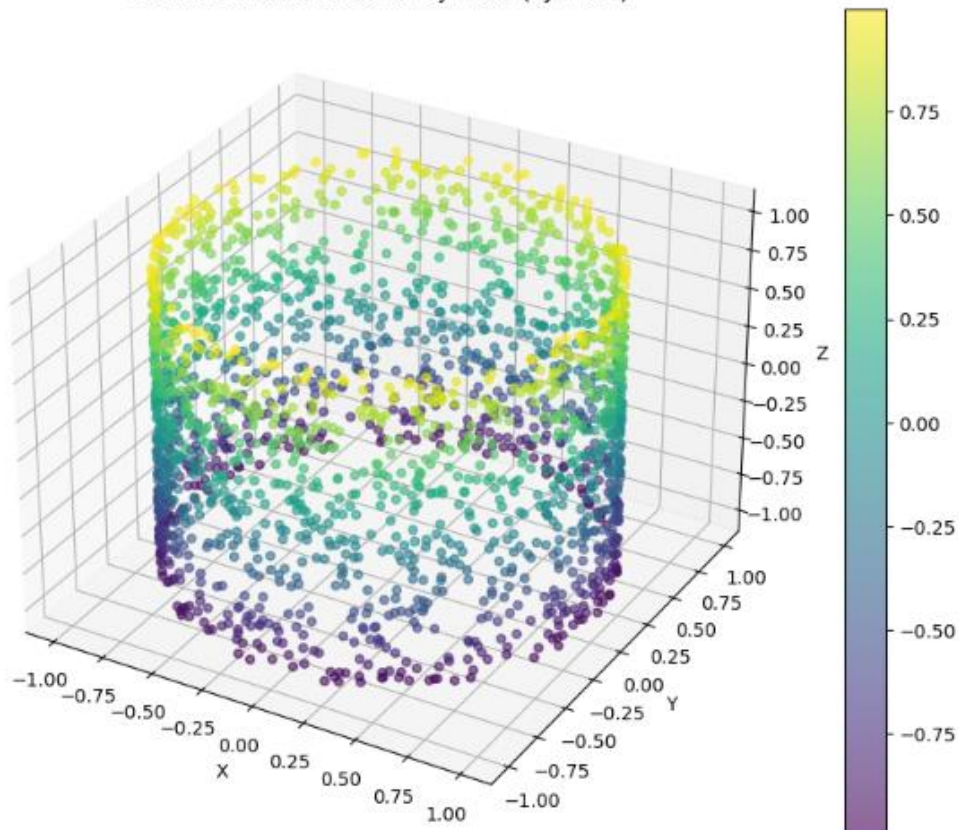
Enter the description or quit for exit: cone
Visualizing: cone

Generated Point Cloud for cone (cone)



Enter the description or quit for exit: cylinder
Visualizing: cylinder

Generated Point Cloud for cylinder (cylinder)



ΚΕΦΑΛΑΙΟ 6 Συμπεράσματα

Η παρούσα διπλωματική εργασία ανέλυσε την ενσωμάτωση τεχνολογιών για την λήψη νέφους σημείων (Point Clouds), όπως το LiDAR και τη φωτογραμμετρία και επίσης τη σύνθεση εικόνων από δεδομένα νέφους σημείων χρησιμοποιώντας τεχνικές βαθιάς μάθησης. Η εργασία αποσκοπούσε να αναδείξει τη σημασία αυτών των τεχνολογιών στη δημιουργία τρισδιάστατων αναπαραστάσεων και να προσφέρει μια ολοκληρωμένη μεθοδολογία για την εφαρμογή τους σε σύνθετες εφαρμογές.

Μέσα από την ανάλυση και τη μελέτη, προέκυψαν ενδιαφέροντα ευρήματα που αποτυπώνουν την εξέλιξη και τη χρήση της τεχνολογίας Point Cloud. Αυτή η τεχνολογία αποτελεί ένα κρίσιμο εργαλείο για την τρισδιάστατη αποτύπωση με εφαρμογές σε τομείς όπως η χαρτογράφηση, η αρχιτεκτονική και η ρομποτική. Η ενσωμάτωση τεχνολογιών LiDAR και φωτογραμμετρίας παρέχει τη δυνατότητα συλλογής ακριβών δεδομένων με υψηλή πυκνότητα και ποιότητα, προσφέροντας σημαντικά πλεονεκτήματα σε μια σειρά εφαρμογών (Grilli et al., 2017; Zhuang et al., 2024).

Παράλληλα, η σύνθεση εικόνας από δεδομένα νέφους σημείων αναδείχθηκε ως ένα ταχέως εξελισσόμενο πεδίο που αξιοποιεί γεωμετρικές μεθόδους σε συνδυασμό με τεχνικές βαθιάς μάθησης. Τα μοντέλα PointNet και Generative Adversarial Networks (GANs) διαδραμάτισαν καθοριστικό ρόλο στη βελτίωση της ακρίβειας και της ταχύτητας αυτής της διαδικασίας, καθιστώντας τη σύνθεση εικόνας πιο αποδοτική και προσαρμόσιμη στις ανάγκες των εφαρμογών (Qi et al., 2017).

Επιπλέον, το πεδίο της μετατροπής κειμένου σε νέφη σημείων (Text-to-Point-Cloud) παρουσιάζει σημαντική πρόοδο. Μέσα από τη χρήση προηγμένων αρχιτεκτονικών, όπως οι Transformers και τα GANs, καθίσταται δυνατή η σύνθεση τρισδιάστατων δεδομένων από περιγραφές κειμένου. Παρά τις προκλήσεις, όπως οι περιορισμένες βάσεις δεδομένων και οι υψηλές υπολογιστικές απαιτήσεις, οι προοπτικές που ανοίγονται είναι ιδιαίτερα ενθαρρυντικές (Nichol et al., 2022; Xiao et al., 2024).

Η εργασία ανέδειξε επίσης τις εφαρμογές και τις προκλήσεις που συνοδεύουν τα νέφη σημείων. Η χρήση τους σε εφαρμογές που απαιτούν ακριβή χαρτογράφηση και ανάλυση είναι ευρέως διαδεδομένη, ενώ ταυτόχρονα επισημάνθηκαν προβλήματα, όπως η έλλειψη δομής και η υψηλή πυκνότητα των δεδομένων, τα οποία χρήζουν αντιμετώπισης μέσω βελτιστοποίησης αλγορίθμων και προεπεξεργασίας δεδομένων (Han et al., 2023; Zhuang et al., 2024).

Η συνεισφορά της εργασίας έγκειται στη διερεύνηση της σύνθεσης εικόνας από δεδομένα νέφους σημείων, εστιάζοντας στη χρήση μοντέρνων τεχνικών βαθιάς μάθησης. Η ολοκληρωμένη προσέγγιση που συνδυάζει γεωμετρικές τεχνικές και μοντέλα μηχανικής μάθησης προσφέρει δυνατότητες για εφαρμογές υψηλής ακρίβειας σε διάφορους τομείς, όπως η αρχιτεκτονική, η ρομποτική και η χαρτογράφηση.

Στο πλαίσιο των μελλοντικών προοπτικών, προτείνεται η ανάπτυξη μεγαλύτερων και πιο διαφοροποιημένων συνόλων δεδομένων που να συνδέουν περιγραφές κειμένου με νέφη σημείων, η βελτίωση των μοντέλων για την αντιμετώπιση προβλημάτων υπολογιστικής πολυπλοκότητας και θορύβου, καθώς και η διεύρυνση των εφαρμογών σε νέους τομείς, όπως η υγειονομική περίθαλψη και η γεωργία ακριβείας.

Η ενσωμάτωση τεχνολογιών νέφους σημείων και βαθιάς μάθησης αποτελεί ένα δυναμικό πεδίο έρευνας με σημαντικές εφαρμογές στην επιστήμη και τη βιομηχανία. Παρά τις υπάρχουσες προκλήσεις, οι προοπτικές για περαιτέρω καινοτομία είναι ενθαρρυντικές, καθιστώντας τα νέα ευρήματα αυτής της εργασίας σημαντική συνεισφορά στο πεδίο της σύνθεσης εικόνας και τρισδιάστατης αναπαράστασης.

Βιβλιογραφία

1. Amaduzzi, A., Lisanti, G., Salti, S., & Di Stefano, L. (2023). Looking at words and points with attention: a benchmark for text-to-shape coherence.
2. Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., & Savarese, S. (2016). 3d semantic parsing of large-scale indoor spaces. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1534-1543).
3. Bao, Y., Ding, T., Huo, J., Liu, Y., Li, Y., Li, W., ... & Luo, J. (2024). 3d gaussian splatting: Survey, technologies, challenges, and opportunities. *arXiv preprint arXiv:2407.17418*.
4. Behnam Behroozpour, Phillip A. M. Sandborn, Ming C. Wu, and Bernhard E. Boser(October 2017) "Lidar System Architectures and Circuits"
5. Benedict O'Neill "Point clouds explained: scanning, processing, 3D models"
6. Bono, G. (2024). Text-to-building: experiments with AI-generated 3D geometry for building design and structure generation. *Architectural Intelligence*, 3(1), 24.
7. Buckley, S. J., Vallet, J., Braathen, A., & Wheeler, W. (2008). Oblique helicopter-based laser scanning for digital terrain modelling and visualisation of geological outcrops. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(Part B4).
8. Chang, A. X., Funkhouser, T., Guibas, L., Hanrahan, P., Huang, Q., Li, Z., ... & Yu, F. (2015). Shapenet: An information-rich 3d model repository. *arXiv preprint arXiv:1512.03012*.
9. Chen, K., Choy, C. B., Savva, M., Chang, A. X., Funkhouser, T., & Savarese, S. (2019). Text2shape: Generating shapes from natural language by learning joint embeddings. In *Computer Vision-ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2-6, 2018, Revised Selected Papers, Part III 14* (pp. 100-116). Springer International Publishing.
10. Chen, Z., Wang, F., Wang, Y., & Liu, H. (2024). Text-to-3d using gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 21401-21412).
11. Choy, C., Dong, W., & Koltun, V. (2020). Deep global registration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2514-2523).
12. Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008, July). Meshlab: an open-source mesh processing tool. In *Eurographics Italian chapter conference* (Vol. 2008, pp. 129-136).
13. Cui, J., He, Y., Niu, J., Ouyang, Z., & Xing, G. (2024, December). αLiDAR: An Adaptive High-Resolution Panoramic LiDAR System. In *Proceedings of the 30th Annual International Conference on Mobile Computing and Networking* (pp. 1515-1529).
14. Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

15. Dong, Z., Liang, F., Yang, B., Xu, Y., Zang, Y., Li, J., ... & Stilla, U. (2020). Registration of large-scale terrestrial laser scanner point clouds: A review and benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163, 327-342.
16. El-Hakim, S. F., Beraldin, J. A., Picard, M., & Godin, G. (2004). Detailed 3D reconstruction of large-scale heritage sites with integrated techniques. *IEEE computer graphics and applications*, 24(3), 21-29.
17. Fan, H., Su, H., & Guibas, L. J. (2017). A point set generation network for 3d object reconstruction from a single image. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 605-613).
18. Fei, B., Yang, W., Chen, W. M., Li, Z., Li, Y., Ma, T., ... & Ma, L. (2022). Comprehensive review of deep learning-based 3d point cloud completion processing and analysis. *IEEE Transactions on Intelligent Transportation Systems*, 23(12), 22862-22883.
19. Feng, Y., Qin, Y., Peng, D., Zhu, H., Peng, X., & Hu, P. (2024). PointCloud-Text Matching: Benchmark Datasets and a Baseline. *arXiv preprint arXiv:2403.19386*.
20. Girardeau-Montaut, D. (2015). CloudCompare: 3D point cloud and mesh processing software. *Open Source Project*, 197.
21. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
22. Granshaw, S. I., Dowman, I., Förstner, W., Hoffmann, C., Leberl, F., & Walker, S. (2017). Photogrammetry and industry. *The Photogrammetric Record*, 32(158), 74-92.
23. Granshaw, S. I., Dowman, I., Förstner, W., Hoffmann, C., Leberl, F., & Walker, S. (2017). Photogrammetry and industry. *The Photogrammetric Record*, 32(158), 74-92.
24. Grilli, E., Menna, F., & Remondino, F. (2017). A review of point clouds segmentation and classification algorithms. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, 339-344.
25. Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., & Bennamoun, M. (2020). Deep learning for 3d point clouds: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 43(12), 4338-4364.
26. Hackel, T., Wegner, J. D., & Schindler, K. (2016). Contour detection in unstructured 3D point clouds. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1610-1618).
27. Han, X. F., Feng, Z. A., Sun, S. J., & Xiao, G. Q. (2023). 3D point cloud descriptors: state-of-the-art. *Artificial Intelligence Review*, 56(10), 12033-12083.
28. Hazer, A., & Yildirim, R. (2025). Hybrid Method for Point Cloud Classification. *IEEE Access*.
29. Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33, 6840-6851.
30. Hu, X., & Xie, X. (2024). PointNet++ Network with Contextual Feature and Mutual Learning for Point Sets.

31. Jiang, C. (2024). A Survey On Text-to-3D Contents Generation In The Wild. *arXiv preprint arXiv:2405.09431*.
32. Kazhdan, M., & Hoppe, H. (2013). Screened poisson surface reconstruction. *ACM Transactions on Graphics (ToG)*, 32(3), 1-13.
33. Kingma, D. P. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
34. Kivi, P. E., Mäkitalo, M. J., Žádník, J., Ikkala, J., Vadakital, V. K. M., & Jääskeläinen, P. O. (2022). Real-time rendering of point clouds with photorealistic effects: a survey. *IEEE Access*, 10, 13151-13173.
35. Kolmet, M., Zhou, Q., Ošep, A., & Leal-Taixé, L. (2022). Text2pos: Text-to-point-cloud cross-modal localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 6687-6696).
36. Lague, D., Brodu, N., & Leroux, J. (2013). Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ). *ISPRS journal of photogrammetry and remote sensing*, 82, 10-26.
37. Lesniak, K., & Tucker, C. S. (2018). Dynamic rendering of remote indoor environments using real-time point cloud data. *Journal of Computing and Information Science in Engineering*, 18(3), 031006.
38. Lewis, P., & Hancock, S. (2007). LiDAR for vegetation applications. *UCL, Gower St, London, UK*.
39. Li, Y., Ma, L., Zhong, Z., Liu, F., Chapman, M. A., Cao, D., & Li, J. (2020). Deep learning for lidar point clouds in autonomous driving: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 32(8), 3412-3432.
40. Luo, C., Li, J., Zhou, C., & Lei, H. (2023). An improved adaptive hp mesh refinement method in solving optimal control problems. *Optimal Control Applications and Methods*, 44(4), 1828-1853.
41. Ma, Y., Zhang, X., Sun, X., Ji, J., Wang, H., Jiang, G., ... & Ji, R. (2023). X-mesh: Towards fast and accurate text-driven 3d stylization via dynamic textual guidance. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 2749-2760).
42. Martinez-Carranza, J., Hernández-Farías, D. I., Vazquez-Meza, V. E., Rojas-Perez, L. O., & Cabrera-Ponce, A. A. (2023). A Study on Generative Models for Visual Recognition of Unknown Scenes Using a Textual Description. *Sensors*, 23(21), 8757.
43. Michel, O., Bar-On, R., Liu, R., Benaim, S., & Hanocka, R. (2022). Text2mesh: Text-driven neural stylization for meshes. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13492-13502).
44. Mirzaei, K., Arashpour, M., Asadi, E., Masoumi, H., Bai, Y., & Behnood, A. (2022). 3D point cloud data processing with machine learning for construction and infrastructure applications: A comprehensive review. *Advanced Engineering Informatics*, 51, 101501.
45. Mo, K., Zhu, S., Chang, A. X., Yi, L., Tripathi, S., Guibas, L. J., & Su, H. (2019). Partnet: A large-scale benchmark for fine-grained and hierarchical part-level 3d object understanding. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 909-918).

46. Nath, U., Goel, R., Jeon, E. S., Kim, C., Min, K., Yang, Y., ... & Turaga, P. (2024). Deep Geometric Moments Promote Shape Consistency in Text-to-3D Generation. *arXiv preprint arXiv:2408.05938*.
47. Nichol, A. Q., & Dhariwal, P. (2021, July). Improved denoising diffusion probabilistic models. In *International conference on machine learning* (pp. 8162-8171). PMLR.
48. Nichol, A., Jun, H., Dhariwal, P., Mishkin, P., & Chen, M. (2022). Point-e: A system for generating 3d point clouds from complex prompts. *arXiv preprint arXiv:2212.08751*.
49. NVIDIA. (2023). NVIDIA RTX GPU: Accelerating point cloud processing. Retrieved from
50. Pix4D. (2022). Pix4Dmapper: Photogrammetry software for professional drone mapping. Retrieved from
51. Pro, A. R.(n.d.). *Reality capture and 3D scanning software*.
52. Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 652-660).
53. Rabbani, T., Dijkman, S., van den Heuvel, F., & Vosselman, G. (2007). An integrated approach for modelling and global registration of point clouds. *ISPRS journal of Photogrammetry and Remote Sensing*, 61(6), 355-370.
54. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., ... & Sutskever, I. (2021, July). Learning transferable visual models from natural language supervision. In *International conference on machine learning* (pp. 8748-8763). PMLR.
55. Remondino, F. (2011). Heritage recording and 3D modeling with photogrammetry and 3D scanning. *Remote sensing*, 3(6), 1104-1138.
56. Rosser, N. J., Petley, D. N., Lim, M., Dunning, S. A., & Allison, R. J. (2005). Terrestrial laser scanning for monitoring the process of hard rock coastal cliff erosion. *Quarterly Journal of Engineering Geology and Hydrogeology*, 38(4), 363-375.
57. Schütz, M. (2015). *Potree: Rendering large point clouds in web browsers* (Doctoral dissertation, Technische Universität Wien).
58. Shang, T., Li, Z., Pei, W., Xu, P., Deng, Z., & Kong, F. (2024). MambaPlace: Text-to-Point-Cloud Cross-Modal Place Recognition with Attention Mamba Mechanisms. *arXiv preprint arXiv:2408.15740*.
59. Shu, D. W., Park, S. W., & Kwon, J. (2019). 3d point cloud generative adversarial network based on tree structured graph convolutions. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 3859-3868).
60. Stamos, I., Liu, L., Chen, C., Wolberg, G., Yu, G., & Zokai, S. (2008). Integrating automated range registration with multiview geometry for the photorealistic modeling of large-scale scenes. *International Journal of Computer Vision*, 78, 237-260.
61. T. Schenk "Introduction to Photogrammetry"
62. Tang, P., Huber, D., Akinci, B., Lipman, R., & Lytle, A. (2010). Automatic reconstruction of as-built building information models from laser-scanned

- point clouds: A review of related techniques. *Automation in construction*, 19(7), 829-843.
63. Tanksale, V. (2023, December). Leveraging Large Language Models for Web3D: Applications, Challenges, and Future Directions. In *2023 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 254-259). IEEE.
 64. Τσαγκάρης, Α. (2014). Αναγνώριση δακτυλικών χειρονομιών σε πραγματικό χρόνο για την αλληλεπίδραση ανθρώπου-ρομπότ: εφαρμογή στον έλεγχο συστημάτων μηχανοκίνητης.
 65. Tychola, K. A., Vrochidou, E., & Papakostas, G. A. (2024). Deep learning based computer vision under the prism of 3D point clouds: a systematic review. *The Visual Computer*, 1-43.
 66. Vaswani, A. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*.
 67. Wang, G., Fan, H., & Kankanhalli, M. (2023, June). Text to point cloud localization with relation-enhanced transformer. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 2, pp. 2501-2509).
 68. Weinmann, M., Jutzi, B., & Mallet, C. (2017). Geometric features and their relevance for 3D point cloud classification. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 4, 157-164.
 69. Wu, S., Wei, R., & Lu, L. (2021, September). Text-Based Point Cloud Generation Adversarial Network. In *2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI)* (pp. 760-764). IEEE.
 70. Xiao, H., Kang, W., Li, Y., & Xu, H. (2024). Text-Free Controllable 3D Point Cloud Generation. *IEEE Transactions on Instrumentation and Measurement*.
 71. Xu, N., Qin, R., & Song, S. (2023). Point cloud registration for LiDAR and photogrammetric data: A critical synthesis and performance analysis on classic and deep learning algorithms. *ISPRS open journal of photogrammetry and remote sensing*, 8, 100032.
 72. Xue, L., Gao, M., Xing, C., Martín-Martín, R., Wu, J., Xiong, C., ... & Savarese, S. (2023). Ulip: Learning a unified representation of language, images, and point clouds for 3d understanding. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 1179-1189).
 73. Yang, R., He, Y., Lu, X., Zhao, Y., Li, Y., Yang, Y., ... & Liu, F. (2024). 3D-based precise evaluation pipeline for maize ear rot using multi-view stereo reconstruction and point cloud semantic segmentation. *Computers and Electronics in Agriculture*, 216, 108512.
 74. Yang, Y., Wang, J., Guo, X., Yang, X., & Qin, W. (2025). Methods for Improving Point Cloud Authenticity in LiDAR Simulation for Autonomous Driving: A Review. *IEEE Access*.
 75. Zhang, H., Wang, C., Tian, S., Lu, B., Zhang, L., Ning, X., & Bai, X. (2023). Deep learning-based 3D point cloud classification: A systematic survey and outlook. *Displays*, 79, 102456.

76. Zhao, H., Jiang, L., Jia, J., Torr, P. H., & Koltun, V. (2021). Point transformer. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 16259-16268).
77. Zhi Li, Yaqi Han, Lican Wu, Zihan Zang, Maolin Dai, Sze Yun Set, Shinji Yamashita, Qian Li, and H. Y. Fu "Towards an ultrafast 3D imaging scanning LiDAR system: a review"
78. Zhu, J., Zhang, J., Zhu, Q., Zuo, L., Liang, C., Chen, X., & Xie, Y. (2024). Virtual geographical scene twin modeling: a combined data-driven and knowledge-driven method with bridge construction as a case study. *International Journal of Digital Earth*, 17(1), 1-23.
79. Zhuang, Z., Zhi, Z., Han, T., Chen, Y., Chen, J., Wang, C., ... & Ma, L. (2024). A Survey of Point Cloud Completion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.
80. Zyrianov, V., Zhu, X., & Wang, S. (2022, October). Learning to generate realistic lidar point clouds. In *European Conference on Computer Vision* (pp. 17-35). Cham: Springer Nature Switzerland.
81. Esraa Hassan & Mahmoud Y. Shams & Noha A. Hikal & Samir Elmougy. (28 September 2022) The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study.

Παράρτημα

```
import torch
import numpy as np
import matplotlib.pyplot as plt
from transformers import AutoTokenizer, AutoModel
from torch import nn
from torch.utils.data import Dataset, DataLoader
from torch.optim import Adam

# Device configuration
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
print(f"Using device: {device}")

# Shape-Aware Point Cloud Generator Model
class ShapeAwarePointCloudGenerator(nn.Module):
    def __init__(self, feature_dim, output_dim):
        super(ShapeAwarePointCloudGenerator, self).__init__()
        self.text_encoder = nn.Sequential(
            nn.Linear(feature_dim, 512),
            nn.ReLU(),
            nn.Linear(512, 1024),
```

```

        nn.ReLU(),
        nn.Linear(1024, 512),
        nn.ReLU(),
        nn.Linear(512, 256),
        nn.ReLU(),
        nn.Linear(256, 128),
        nn.ReLU()
    )
    self.shape_classifier = nn.Sequential(
        nn.Linear(128, 64),
        nn.ReLU(),
        nn.Linear(64, 10) # 10 classes: cube, sphere, cylinder, pyramid,
        cone, torus, star, disk, polyhedron, ellipsoid
    )
    self.point_generator = nn.Sequential(
        nn.Linear(128, 256),
        nn.ReLU(),
        nn.Linear(256, 512),
        nn.ReLU(),
        nn.Linear(512, 1024),
        nn.ReLU(),
        nn.Linear(1024, 512),
        nn.ReLU(),
        nn.Linear(512, output_dim)
    )

    def forward(self, text_features):
        encoded = self.text_encoder(text_features)
        shape_logits = self.shape_classifier(encoded)
        point_cloud = self.point_generator(encoded)
        return point_cloud, shape_logits

# Generate and visualize all shapes
def generate_shape_specific_points(shape, num_points=2048):
    if shape == "cube":
        points = np.random.uniform(-1, 1, (num_points, 3))
    elif shape == "sphere":
        phi = np.random.uniform(0, np.pi * 2, num_points)
        costheta = np.random.uniform(-1, 1, num_points)
        u = np.random.uniform(0, 1, num_points)
        theta = np.arccos(costheta)
        r = u ** (1/3)
        x = r * np.sin(theta) * np.cos(phi)
        y = r * np.sin(theta) * np.sin(phi)
        z = r * np.cos(theta)
        points = np.stack((x, y, z), axis=1)
    elif shape == "cylinder":
        theta = np.random.uniform(0, 2 * np.pi, num_points)

```

```

    z = np.random.uniform(-1, 1, num_points)
    x = np.cos(theta)
    y = np.sin(theta)
    points = np.stack((x, y, z), axis=1)
elif shape == "pyramid":
    base_size = 1.0
    height = 1.5
    base_x = np.random.uniform(-base_size, base_size, num_points)
    base_y = np.random.uniform(-base_size, base_size, num_points)
    z = np.zeros(num_points)
    apex_x, apex_y, apex_z = 0, 0, height
    t = np.random.uniform(0, 1, num_points)
    x = base_x * (1 - t) + apex_x * t
    y = base_y * (1 - t) + apex_y * t
    z = z * (1 - t) + apex_z * t
    points = np.stack((x, y, z), axis=1)
elif shape == "cone":
    height = 2.0
    base_radius = 1.0
    theta = np.random.uniform(0, 2 * np.pi, num_points)
    r = np.random.uniform(0, base_radius, num_points)
    base_x = r * np.cos(theta)
    base_y = r * np.sin(theta)
    base_z = np.zeros(num_points)
    apex_x, apex_y, apex_z = 0, 0, height
    t = np.random.uniform(0, 1, num_points)
    x = base_x * (1 - t) + apex_x * t
    y = base_y * (1 - t) + apex_y * t
    z = base_z * (1 - t) + apex_z * t
    points = np.stack((x, y, z), axis=1)
elif shape == "torus":
    R = 1.0 # Major radius
    r = 0.5 # Minor radius
    theta = np.random.uniform(0, 2 * np.pi, num_points)
    phi = np.random.uniform(0, 2 * np.pi, num_points)
    x = (R + r * np.cos(phi)) * np.cos(theta)
    y = (R + r * np.cos(phi)) * np.sin(theta)
    z = r * np.sin(phi)
    points = np.stack((x, y, z), axis=1)
elif shape == "star":
    num_arms = 5
    arm_length = 1.5
    points = []
    for _ in range(num_points):
        arm = np.random.randint(0, num_arms)
        angle = arm * (2 * np.pi / num_arms)
        r = np.random.uniform(0.5, arm_length) if np.random.rand() > 0.5
    else np.random.uniform(0, 0.5)

```

```

        x = r * np.cos(angle)
        y = r * np.sin(angle)
        z = np.random.uniform(-0.1, 0.1) # Slight 3D depth
        points.append([x, y, z])
    points = np.array(points)
elif shape == "disk":
    r = np.random.uniform(0, 1, num_points)
    theta = np.random.uniform(0, 2 * np.pi, num_points)
    x = r * np.cos(theta)
    y = r * np.sin(theta)
    z = np.zeros(num_points)
    points = np.stack((x, y, z), axis=1)
elif shape == "polyhedron":
    vertices = [
        [1, 1, 1], [-1, 1, 1], [1, -1, 1], [-1, -1, 1],
        [1, 1, -1], [-1, 1, -1], [1, -1, -1], [-1, -1, -1]
    ]
    faces = [
        [0, 1, 2], [1, 3, 2], [4, 5, 6], [5, 7, 6],
        [0, 1, 4], [1, 5, 4], [2, 3, 6], [3, 7, 6],
        [0, 2, 4], [2, 6, 4], [1, 3, 5], [3, 7, 5]
    ]
    points = []
    for face in faces:
        v0, v1, v2 = np.array(vertices[face[0]]), np.array(vertices[face[1]]),
np.array(vertices[face[2]])
        for _ in range(num_points // len(faces)):
            u, v = np.random.rand(), np.random.rand()
            if u + v > 1:
                u, v = 1 - u, 1 - v
            p = v0 + u * (v1 - v0) + v * (v2 - v0)
            points.append(p)
    points = np.array(points)
elif shape == "ellipsoid":
    a, b, c = 1.5, 1.0, 0.5 # Semi-principal axes
    phi = np.random.uniform(0, np.pi * 2, num_points)
    theta = np.random.uniform(0, np.pi, num_points)
    x = a * np.sin(theta) * np.cos(phi)
    y = b * np.sin(theta) * np.sin(phi)
    z = c * np.cos(theta)
    points = np.stack((x, y, z), axis=1)
else:
    points = np.random.uniform(-1, 1, (num_points, 3))

return points

```

Visualization function with enhanced customization

```

def visualize_point_cloud(point_cloud, title="Point Cloud",
shape_name="Unknown"):
    fig = plt.figure(figsize=(10, 8))
    ax = fig.add_subplot(111, projection='3d')
    scatter = ax.scatter(point_cloud[:, 0], point_cloud[:, 1], point_cloud[:, 2],
        c=point_cloud[:, 2], cmap='viridis', alpha=0.6)
    ax.set_title(f'{title} ({shape_name})')
    ax.set_xlabel('X')
    ax.set_ylabel('Y')
    ax.set_zlabel('Z')
    plt.colorbar(scatter)
    plt.show()

# Dataset class with added support for new shapes
class ShapeAwarePointCloudDataset(Dataset):
    def __init__(self, data, tokenizer, num_points=2048):
        self.data = data
        self.tokenizer = tokenizer
        self.num_points = num_points
        self.shape_classes = ["cube", "sphere", "cylinder", "pyramid", "cone",
"torus", "star", "disk", "polyhedron", "ellipsoid"]

    def __len__(self):
        return len(self.data)

    def __getitem__(self, idx):
        sample = self.data[idx]
        text = sample["text"]
        point_cloud = sample["point_cloud"]

        # Padding/truncating point clouds to ensure consistent length
        if len(point_cloud) < self.num_points:
            padding = np.zeros((self.num_points - len(point_cloud), 3))
            point_cloud = np.vstack((point_cloud, padding))
        elif len(point_cloud) > self.num_points:
            point_cloud = point_cloud[:self.num_points]

        shape = sample.get("shape", "unknown")
        shape_label = self.shape_classes.index(shape)
        tokens = self.tokenizer(text, return_tensors="pt", padding=True,
truncation=True)
        point_cloud = torch.tensor(point_cloud, dtype=torch.float32)

        for key in tokens:
            tokens[key] = tokens[key].squeeze()

        # Verification of expected tensor shape

```

```

        if point_cloud.shape[0] != self.num_points or point_cloud.shape[1] !=
3:         print(f"Warning: Unexpected point cloud shape {point_cloud.shape}
for sample {idx}.")

        return tokens, point_cloud, torch.tensor(shape_label)

# Custom collate function for DataLoader
def custom_collate_fn(batch):
    tokens_batch = {k: [] for k in batch[0][0].keys()}
    point_clouds = []
    shape_labels = []

    # Collate and pad
    for tokens, point_cloud, shape_label in batch:
        for k in tokens.keys():
            tokens_batch[k].append(tokens[k])
            point_clouds.append(point_cloud)
            shape_labels.append(shape_label)

    # Convert to tensors with padding
    tokens_batch = {k: torch.nn.utils.rnn.pad_sequence(v, batch_first=True)
for k, v in tokens_batch.items()}
    point_clouds = torch.stack(point_clouds)
    shape_labels = torch.tensor(shape_labels, dtype=torch.long)

    return tokens_batch, point_clouds, shape_labels

# Create synthetic dataset with all supported shapes
def create_synthetic_dataset(num_samples=1000, num_points=2048):
    data = []
    shapes = ["cube", "sphere", "cylinder", "pyramid", "cone", "torus", "star",
"disk", "polyhedron", "ellipsoid"]
    for _ in range(num_samples):
        shape = np.random.choice(shapes)
        points = generate_shape_specific_points(shape, num_points)
        description = f"A {shape} shape"
        data.append({"text": description, "point_cloud": points, "shape":
shape})
    return data

# Generate and visualize all shapes
def visualize_all_shapes():
    shapes = ["cube", "sphere", "cylinder", "pyramid", "cone", "torus", "star",
"disk", "polyhedron", "ellipsoid"]
    for shape in shapes:
        print(f"Visualizing: {shape}")
        point_cloud = generate_shape_specific_points(shape)

```

```

        visualize_point_cloud(point_cloud, title=f"Generated Point Cloud for
{shape}", shape_name=shape)

# Training function with combined loss
def train_model(generator, dataset, tokenizer, text_model,
num_epochs=10, batch_size=16, learning_rate=1e-4):
    data_loader = DataLoader(dataset, batch_size=batch_size,
shuffle=True, collate_fn=custom_collate_fn)
    optimizer = Adam(generator.parameters(), lr=learning_rate)
    point_loss_fn = nn.MSELoss()
    shape_loss_fn = nn.CrossEntropyLoss()

    for epoch in range(num_epochs):
        total_point_loss = 0
        total_shape_loss = 0
        for batch in data_loader:
            tokens, target_point_cloud, shape_labels = batch
            tokens = {k: v.to(device) for k, v in tokens.items()}
            target_point_cloud = target_point_cloud.to(device)
            shape_labels = shape_labels.to(device)
            with torch.no_grad():
                text_features = text_model(**tokens).last_hidden_state[:, 0,
:].to(device)
            predicted_point_cloud, shape_logits = generator(text_features)
            point_loss = point_loss_fn(
                predicted_point_cloud.view(-1, 2048, 3),
                target_point_cloud.view(-1, 2048, 3)
            )
            shape_loss = shape_loss_fn(shape_logits, shape_labels)
            total_loss = point_loss + shape_loss
            optimizer.zero_grad()
            total_loss.backward()
            optimizer.step()
            total_point_loss += point_loss.item()
            total_shape_loss += shape_loss.item()
        print(f"Epoch {epoch + 1}/{num_epochs}, Point Loss: {total_point_loss /
len(data_loader)}, Shape Loss: {total_shape_loss / len(data_loader)}")

def find_specific_words():
    # Ask the user to provide the text and words to search
    text = input("Enter the description or quit for exit: ")

    words_to_find = ["cube", "sphere", "cylinder", "pyramid", "cone", "torus",
"star", "disk", "polyhedron", "ellipsoid", "quit"]

    # Clean up the words to remove extra spaces and normalize to lowercase
    words_to_find = [word.strip().lower() for word in words_to_find]

```

```

# Normalize the text to lowercase and split into words
text_words = text.lower().split()

# Search for occurrences of the target words
matches = [word for word in words_to_find if word in text_words]

if matches == 'null':
    matches == 'random'

return matches

# Main execution
if __name__ == "__main__":
    feature_dim = 768
    output_dim = 2048 * 3
    tokenizer = AutoTokenizer.from_pretrained("bert-base-uncased")
    text_model = AutoModel.from_pretrained("bert-base-uncased")
    text_model.eval()
    text_model = text_model.to(device)
    generator = ShapeAwarePointCloudGenerator(feature_dim,
output_dim)
    generator = generator.to(device)
    data = create_synthetic_dataset(num_samples=1000)
    dataset = ShapeAwarePointCloudDataset(data, tokenizer)
    train_model(generator, dataset, tokenizer, text_model, num_epochs=10)
    torch.save(generator.state_dict(), "shape_point_cloud_generator.pth")

# Visualize all shapes
#print("Visualizing all shapes...")
#visualize_all_shapes()

# Interactive mode
should_continue = True
while should_continue:
    shape = find_specific_words()

    for row in shape:
        if row == 'quit':
            print("Exiting program...")
            should_continue = False
            break

    print(f"Visualizing: {row}")
    point_cloud = generate_shape_specific_points(row)
    visualize_point_cloud(point_cloud, title=f"Generated Point Cloud
for {row}", shape_name=row)

```