



**UNIMORE**

UNIVERSITÀ DEGLI STUDI DI  
MODENA E REGGIO EMILIA

DIPARTIMENTO DI SCIENZE FISICHE, INFORMATICHE E MATEMATICHE

Corso di Laurea Magistrale in Matematica

**Analisi comparativa di metodi  
geometrici e neurali per il  
rilevamento della polvere in dati  
LiDAR**

Candidato:  
Lorenzo Barbieri

Relatrice:  
Prof.ssa Giorgia Franchini

Correlatrice:  
Dott.ssa Micaela Verucchi

Anno Accademico 2024/2025

# Abstract

La presenza di polvere costituisce una delle principali cause di degrado del dato LiDAR in ambienti non strutturati, come cave, miniere e contesti industriali off-road. Le particelle aerodisperse possono infatti generare ritorni spurii o ambigui, alterando la nuvola di punti e riducendo l'affidabilità delle successive fasi della pipeline percettiva, tra cui segmentazione, rilevamento di ostacoli e interpretazione della scena. In questo contesto, il problema affrontato consiste nel distinguere i ritorni prodotti dalla polvere dai punti che descrivono effettivamente la scena osservata. In questa tesi tale problema viene formulato come un compito di segmentazione binaria point-wise, con l'obiettivo di classificare ciascun punto come appartenente alla classe dust oppure non-dust.

Il lavoro confronta approcci di natura differente per la dust detection su dati LiDAR. In particolare, vengono considerati due filtri geometrici basati su regole esplicite, LIDROR e DMNR, e tre modelli neurali di segmentazione: RandLA-Net, WeatherNet e U-Net. La valutazione sperimentale è condotta sul dataset pubblico LiDARDustX, specificamente progettato per scenari polverosi, ed è accompagnata da un'analisi qualitativa su dati reali raccolti in ambito aziendale, in modo tale da discutere il comportamento dei metodi anche in uno scenario operativo reale.

I risultati ottenuti su LiDARDustX mostrano una netta superiorità degli approcci neurali rispetto ai filtri geometrici. In termini di IoU sulla classe dust, WeatherNet raggiunge il valore migliore (0.861), seguito da U-Net (0.833) e RandLA-Net (0.779), mentre tra i metodi geometrici il risultato più elevato è ottenuto da DMNR (0.456), superiore a LIDROR (0.330). Inoltre, dal punto di vista computazionale, U-Net risulta il modello neurale con i tempi di inferenza più contenuti, mostrando il miglior compromesso tra accuratezza ed efficienza inferenziale.

Poiché i dati reali raccolti in ambito aziendale non dispongono di annotazioni point-wise, il relativo confronto può essere svolto solo in termini qualitativi. In questo contesto, emergono tre comportamenti complessivamente più convincenti, seppur con caratteristiche differenti: RandLA-Net appare il metodo più equilibrato,

in quanto segue in modo plausibile le regioni polverose senza compromettere eccessivamente la scena utile; WeatherNet produce segmentazioni più estese e continue, ma tende a essere più aggressivo e a introdurre un numero maggiore di falsi positivi; LIDROR, al contrario, mostra un comportamento più conservativo e, in diversi casi, sottostima l'estensione della polvere. Risultano invece meno convincenti DMNR e U-Net: il primo appare troppo restrittivo nel rilevamento, mentre il secondo risulta troppo aggressivo nella classificazione della polvere.

# Indice

<b>Abstract</b>	<b>1</b>
<b>1 Introduzione</b>	<b>5</b>
1.1 Problema della dust detection con sensori LiDAR . . . . .	5
1.2 Contesto applicativo e motivazioni . . . . .	7
1.3 Obiettivi della tesi e contributi principali . . . . .	8
<b>2 Dati e formulazione del problema</b>	<b>10</b>
2.1 Dataset LiDARDustX . . . . .	10
2.2 Dati reali raccolti in azienda . . . . .	13
2.3 Preprocessing dei dati . . . . .	14
2.4 Rappresentazione dei dati . . . . .	14
2.5 Formulazione del problema . . . . .	15
<b>3 Metodologie</b>	<b>17</b>
3.1 Introduzione ai metodi considerati . . . . .	17
3.1.1 Filtro LIDROR . . . . .	18
3.1.2 Filtro DMNR . . . . .	19
3.1.3 RandLA-Net . . . . .	20
3.1.4 WeatherNet . . . . .	22
3.1.5 U-Net . . . . .	23
3.2 Confronto teorico e considerazioni computazionali . . . . .	25
<b>4 Protocollo sperimentale e metriche di valutazione</b>	<b>27</b>
4.1 Protocollo sperimentale . . . . .	27
4.1.1 Piattaforma di calcolo . . . . .	27
4.1.2 Suddivisione dei dati . . . . .	27
4.2 Metriche di valutazione . . . . .	28
4.3 Misura delle prestazioni computazionali . . . . .	30
4.4 Risultati sul dataset LiDARDustX . . . . .	30

4.4.1	Filtri geometrici . . . . .	30
4.4.2	Modelli neurali . . . . .	32
4.4.3	Analisi quantitativa dei risultati . . . . .	34
4.5	Risultati sui dati reali . . . . .	36
4.5.1	Analisi qualitativa dei risultati . . . . .	37
<b>5</b>	<b>Conclusioni</b>	<b>39</b>
	<b>Bibliografia</b>	<b>41</b>

# Capitolo 1

## Introduzione

### 1.1 Problema della dust detection con sensori LiDAR

Uno degli aspetti più critici per l'impiego di sensori LiDAR in ambienti non strutturati è la presenza di polvere sospesa nell'aria. In scenari quali cave, miniere e contesti off-road, le particelle aerodisperse possono infatti produrre ritorni non riconducibili a superfici stabili della scena, introducendo rumore nella nuvola di punti acquisita. In questo contesto, la dust detection consiste nell'individuare, all'interno della nuvola di punti acquisita dal sensore, i ritorni attribuibili alla polvere, distinguendoli da quelli che descrivono effettivamente l'ambiente osservato [11, 1, 4].

Un esempio visivo del problema è riportato in Figura 1.1. La regione evidenziata in verde corrisponde alla polvere, mentre il resto della nuvola è rappresentato tramite una scala di colore basata sull'intensità del ritorno, in cui i toni più chiari corrispondono a intensità maggiori e quelli più scuri a intensità minori. L'immagine mostra chiaramente come i punti associati alla polvere si sovrappongano alla struttura della scena, rendendo difficile la separazione tra ritorni utili e ritorni spuri.

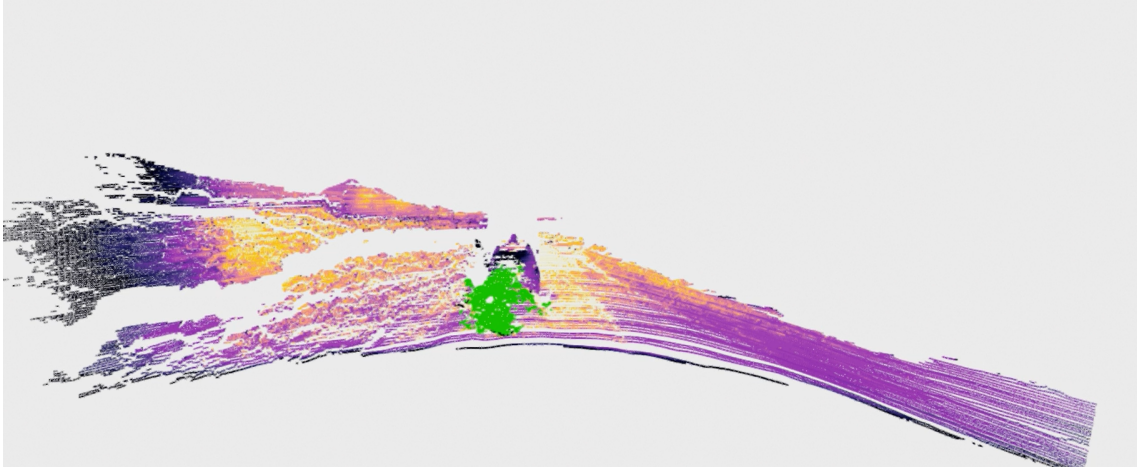


Figura 1.1: Esempio illustrativo del problema della dust detection in una nuvola di punti LiDAR, ottenuto a partire da un frame del dataset LiDARDustX.

La difficoltà della dust detection dipende in primo luogo dalla natura del fenomeno e dalle modalità con cui esso si riflette nel dato acquisito dal sensore. La polvere può infatti presentarsi in forme molto diverse: nubi dense o diffuse, accumuli localizzati, regioni sparse e irregolari, oppure insiemi di punti debolmente strutturati che si sovrappongono a quelli provenienti dal terreno o da altri oggetti presenti nella scena [11, 1, 4]. Inoltre, le misure associate alla polvere possono variare sensibilmente in funzione di diversi fattori, tra cui la distanza dal sensore, l'angolo di osservazione, la densità locale dei punti, l'intensità del segnale e le caratteristiche del dispositivo utilizzato [1, 4]. Di conseguenza, la nuvola di punti può contenere rilevazioni ambigue o indesiderate, difficili da distinguere nettamente da quelle prodotte dalla scena reale.

Questa ambiguità ha conseguenze rilevanti sul piano applicativo: la presenza di punti riconducibili alla polvere può infatti interferire con le successive fasi della pipeline percettiva, influenzando operazioni quali segmentazione, rilevamento di ostacoli, stima della scena e, più in generale, interpretazione dell'ambiente circostante [5]. Per questo motivo, la dust detection non rappresenta soltanto una fase preliminare di pulizia del dato, ma un vero e proprio problema percettivo che incide direttamente sull'affidabilità complessiva dei sistemi basati su sensori LiDAR.

All'interno di questa tesi, il problema sarà trattato come un compito di segmentazione binaria point-wise. Questa scelta, che verrà formalizzata nel Capitolo 2, consente di confrontare in modo omogeneo approcci diversi, mantenendo il focus sul nucleo applicativo del problema.

## 1.2 Contesto applicativo e motivazioni

I sistemi di guida autonoma e, più in generale, le piattaforme robotiche mobili richiedono una percezione affidabile dell'ambiente circostante, così da poter rilevare ostacoli, stimare la geometria della scena e supportare le successive fasi decisionali. Per soddisfare questi requisiti, i veicoli autonomi non si affidano in genere a un singolo sensore, ma integrano tecnologie diverse, tra cui camere, radar e LiDAR, ciascuna caratterizzata da specifici punti di forza e limiti [8, 10]. Un confronto qualitativo tra queste principali tecnologie di percezione è riportato in Tabella 1.1.

<b>Sensore</b>	<b>Informazione principale</b>	<b>Punti di forza</b>	<b>Limiti principali</b>
Camera	Informazione visiva e semantica della scena	Alto contenuto semantico, buona risoluzione dell'immagine, costo contenuto	Sensibile a illuminazione, ombre, abbagliamento e condizioni meteo; non fornisce direttamente la distanza
Radar	Distanza, velocità relativa, presenza di oggetti	Robusto in condizioni ambientali avverse e in scarsa illuminazione; misura direttamente la velocità	Risoluzione spaziale più limitata, minore dettaglio geometrico e semantico
LiDAR	Geometria tridimensionale della scena, distanza, intensità del ritorno	Rappresentazione 3D diretta, buona accuratezza geometrica, utile per segmentazione, localizzazione e rilevamento di ostacoli	Sensibile a pioggia, nebbia, polvere e particelle sospese; costo e complessità generalmente maggiori rispetto alla camera

Tabella 1.1: Confronto qualitativo tra le principali tecnologie di percezione utilizzate in sistemi di guida autonoma e robotica mobile. Tabella costruita a partire da [8, 10].

In questo contesto, il LiDAR (*Light Detection and Ranging*) riveste un ruolo particolarmente importante, poiché fornisce una rappresentazione tridimensionale diretta dell'ambiente sotto forma di nuvola di punti. Ogni punto è descritto da coordinate spaziali e, in molti casi, da un valore di intensità associato al ritorno. Questa caratteristica rende il LiDAR particolarmente utile per compiti di segmentazione, localizzazione, rilevamento di ostacoli e pianificazione del movimento. In particolare, rispetto alle camere, il LiDAR offre una descrizione geometrica più immediata della scena; mentre, rispetto ai radar, fornisce una rappresentazione spaziale generalmente più dettagliata, risultando però più sensibile alla presenza di particelle sospese e a condizioni ambientali difficili [8, 10].

Se in contesti urbani o stradali strutturati la percezione LiDAR è stata ampiamente studiata, in ambienti non strutturati le condizioni operative risultano spesso più complesse. Miniere a cielo aperto, cave di sabbia e, più in generale, contesti estrattivi e off-road rappresentano esempi significativi di scenari nei quali il sensore è esposto a fondo irregolare, forte variabilità geometrica e presenza frequente di polvere [1, 2, 7, 11]. In tali contesti, il degrado della nuvola di punti può compromettere in misura significativa l'affidabilità della percezione.

Questo lavoro di tesi si inserisce nell'ambito del tirocinio svolto presso Hipert SRL, un'azienda informatica del comune di Modena, con esperienza maturata nel settore della robotica e focalizzata sulla capacità di integrare la percezione, la pianificazione e il controllo dell'IA su piattaforme embedded. In questo contesto aziendale, il problema della dust detection è emerso in relazione all'impiego di sensori LiDAR in contesti applicativi reali. La motivazione del lavoro è quindi duplice: da un lato, analizzare un problema ancora aperto nella percezione LiDAR e, dall'altro, valutarne la rilevanza in scenari concreti di interesse industriale.

## 1.3 Obiettivi della tesi e contributi principali

Questa tesi si propone di analizzare e confrontare diversi metodi per la dust detection su dati LiDAR in ambienti polverosi. In particolare, vengono considerate due famiglie di approcci concettualmente differenti: filtri geometrici e metodi neurali. Il confronto tra questi due approcci consente di discutere il problema non solo in termini di capacità di identificare la polvere, ma anche rispetto a efficienza computazionale, interpretabilità e comportamento in scenari differenti. La valutazione sperimentale si sviluppa su due livelli. In primo luogo, i metodi vengono

analizzati sul dataset pubblico LiDARDustX [11], che fornisce un benchmark annotato e permette quindi un confronto sistematico in un contesto controllato. In secondo luogo, gli stessi approcci vengono applicati a dati reali raccolti in ambito aziendale, che non dispongono di etichette di riferimento: in questo caso, pertanto, il confronto non può che essere qualitativo, e mira a verificare se i risultati osservati sul benchmark si mantengano plausibili anche in uno scenario operativo differente.

In vista di un possibile impiego operativo, particolare attenzione è dedicata anche al profilo computazionale dei metodi studiati, poiché una soluzione efficace deve risultare non solo accurata, ma anche compatibile con possibili vincoli di impiego reale.

Il contributo principale del lavoro consiste quindi in una valutazione integrata di approcci geometrici e neurali rispetto al tema della dust detection, condotta su tre piani complementari: prestazioni su benchmark annotato, comportamento su dati reali non etichettati e requisiti computazionali. Pertanto, l'obiettivo della tesi non è soltanto stabilire quale metodo ottenga i risultati migliori, ma analizzare in che modo i diversi approcci si differenzino per prestazioni, robustezza, interpretabilità e sostenibilità operativa.

# Capitolo 2

## Dati e formulazione del problema

### 2.1 Dataset LiDARDustX

Il dataset di riferimento principale per questo lavoro è LiDARDustX, proposto come benchmark specifico per la percezione LiDAR in ambienti non strutturati caratterizzati da una forte presenza di polvere [11]. A differenza dei dataset più diffusi nella letteratura sulla guida autonoma, prevalentemente raccolti in scenari urbani strutturati, LiDARDustX è stato sviluppato per contesti operativi come miniere a cielo aperto e cave di sabbia, nei quali la polvere costituisce un rilevante fattore di degrado della percezione.

Secondo quanto riportato dagli autori, LiDARDustX comprende 30 000 frame annotati, organizzati in 180 sequenze e acquisiti in differenti siti minerari e cave di sabbia situati nelle province cinesi della Mongolia Interna, Shanxi e Gansu [11]. I dati sono stati raccolti in condizioni diurne mediante un camion da miniera TONGLI equipaggiato con più sensori LiDAR, in modo da acquisire scene realistiche in presenza sia di polvere naturale sia di polvere generata da attività operative, quali scavo e movimento dei veicoli. Inoltre, più dell'80% del dataset è costituito da scene con presenza di polvere, caratteristica che rende LiDARDustX particolarmente adatto allo studio del rilevamento della polvere in ambienti estrattivi [11].

Un elemento distintivo del dataset è la varietà della configurazione sensoriale. La Figura 2.1 mostra, infatti, la distribuzione dei frame tra i diversi sensori LiDAR.

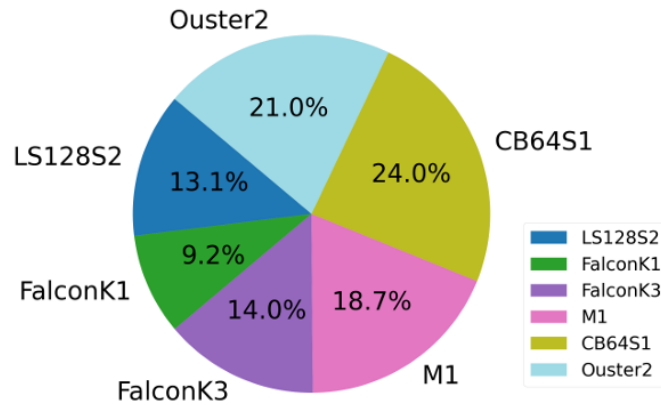


Figura 2.1: Distribuzione dei frame tra i sensori LiDAR. Immagine tratta da [11].

LiDARDustX include dunque sei differenti sensori LiDAR, installati in diverse posizioni del veicolo per acquisire prospettive complementari della scena [11]. In particolare, FalconK3, LS128S2, FalconK1 e RS-LiDAR-M1 sono montati nella parte frontale del camion, mentre Ouster2 e CB64S1 sono collocati lateralmente [11]. Le differenze in termini di campo visivo, portata e accuratezza introducono una significativa eterogeneità nelle acquisizioni, aspetto rilevante per le successive scelte di rappresentazione dei dati e per l'analisi dei metodi considerati. Le principali caratteristiche dei sensori sono riassunte in Tabella 2.1.

Sensore	FOV orizzontale	FOV verticale	Portata max	Accuratezza
LS128S2	120°	da -12.5° a 12.5°	180 m	±2 cm
CB64S1	180°	da -25° a 15°	100 m	±3 cm
FalconK1	120°	25°	250 m	-
FalconK3	120°	20°	300 m	-
RS-LiDAR-M1	120°	25°	200 m	±5 cm
Ouster2	-	da -7.9° a +7.9°	150 m	-

Tabella 2.1: Principali caratteristiche dei sensori LiDAR presenti nel dataset LiDARDustX. Tabella costruita a partire da [11].

Dal punto di vista delle informazioni disponibili, LiDARDustX fornisce, per ciascun frame, la nuvola di punti LiDAR, comprensiva di coordinate spaziali tridimensionali e intensità di ritorno, insieme a due diverse forme di annotazione: bounding box 3D, per il problema di rilevamento degli oggetti, ed etichette semantiche associate

ai singoli punti, per il problema di segmentazione [11]. Questa doppia modalità di annotazione rende il dataset adatto a problemi percettivi differenti e ne amplia l'utilizzabilità come benchmark per l'analisi di metodi sia di rilevamento di oggetti sia di interpretazione semantica della scena.

Per il rilevamento degli oggetti, LiDARDustX fornisce oltre 300 000 bounding box 3D, distribuite su 14 categorie di oggetti. Ogni istanza annotata è rappresentata tramite una bounding box 3D a 7 gradi di libertà, che specifica posizione del centro, dimensioni spaziali e orientamento dell'oggetto.

Le annotazioni semantiche point-wise associano invece a ciascun punto della nuvola LiDAR una specifica etichetta semantica, appartenente a un insieme di 16 classi che descrivono i diversi componenti della scena [11]. La Figura 2.2 mostra la distribuzione del numero di punti per classe e mette in evidenza una forte disomogeneità, resa particolarmente chiara dalla scala logaritmica dell'asse verticale. In questo contesto, la classe *Dust* appare meno dominante rispetto alle classi di sfondo come *Ground* e *Wall*, ma comunque sufficientemente rappresentata da costituire un target significativo per un problema di segmentazione. Questo aspetto conferma che LiDARDustX non si limita a includere episodi polverosi in modo marginale, ma offre una base annotata adeguata per lo studio della dust detection su nuvole di punti.

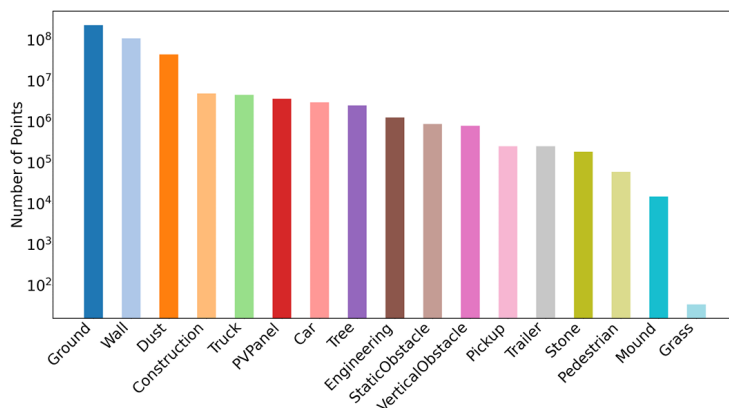


Figura 2.2: Distribuzione del numero di punti nelle diverse classi semantiche del dataset LiDARDustX. Immagine tratta da [11].

Nel presente lavoro, l'analisi è stata focalizzata sul problema della dust detection mediante metodi di segmentazione. Per questo motivo, sono state utilizzate esclusivamente la nuvola di punti e le etichette semantiche point-wise, mentre le annotazioni tramite bounding box 3D non sono state impiegate nelle fasi successive. Inoltre, le analisi non sono state condotte sull'intero dataset descritto

dagli autori, ma sulla porzione effettivamente resa disponibile al momento del download, composta da 7562 frame distribuiti su 174 sequenze. Pur trattandosi di un sottoinsieme del benchmark originario, esso ne conserva le caratteristiche più rilevanti per il problema affrontato, in particolare la varietà sensoriale, la presenza di scene con polvere e la disponibilità di annotazioni semantiche a livello di punto, risultando quindi una base sperimentale adeguata per il confronto dei metodi considerati in questa tesi.

## 2.2 Dati reali raccolti in azienda

Oltre al dataset pubblico LiDARDustX, in questa tesi sono stati considerati anche dati reali raccolti in ambito aziendale. Tali acquisizioni provengono da una cava a cielo aperto e sono state ottenute nell'ambito di un progetto finalizzato allo sviluppo di un camion a guida autonoma per ambienti estrattivi. Si tratta quindi di dati acquisiti in un contesto operativo reale, distinto da quello del benchmark pubblico ma strettamente coerente con lo scenario applicativo di interesse del presente lavoro.

Il dataset messo a disposizione dall'azienda comprende complessivamente circa 17000 frame costituiti da nuvole di punti LiDAR. Dal punto di vista della rappresentazione, ciascun frame deriva dalla fusione di più sensori LiDAR, producendo nuvole di punti più estese rispetto a quelle considerate nel dataset pubblico. Le acquisizioni sono state analizzate considerando principalmente le coordinate spaziali dei punti e il corrispondente valore di intensità. Nel complesso, queste caratteristiche introducono una maggiore complessità geometrica e una diversa distribuzione spaziale dei punti rispetto a LiDARDustX.

Nel presente lavoro non è stato utilizzato l'intero archivio, ma soltanto un sottoinsieme di circa 300 frame ricavato da una sequenza temporale più ampia, selezionata manualmente in quanto caratterizzata da contesti operativi nei quali la presenza di polvere risultava plausibile, come nel caso di veicoli in movimento. Poiché la frequenza di acquisizione era elevata e molti frame temporalmente adiacenti risultavano quasi identici, è stato successivamente adottato un sottocampionamento selezionando un frame ogni dieci, così da ridurre la ridondanza senza perdere l'evoluzione della scena nel tempo. La selezione è stata effettuata mediante analisi congiunta delle nuvole di punti LiDAR e delle corrispondenti immagini RGB acquisite da una camera frontale, così da facilitare l'interpretazione delle scene osservate.

A differenza del dataset pubblico, i dati aziendali non disponevano di annotazioni semantiche point-wise e, pertanto, non sono stati impiegati per una valutazione

quantitativa supervisionata. Essi sono stati invece utilizzati per un'analisi qualitativa dei risultati prodotti dai diversi approcci, osservando quali regioni della nuvola di punti venissero identificate come polvere e confrontandole con l'evidenza visiva disponibile. Questa fase ha permesso di avvicinare l'analisi al problema reale di interesse aziendale e di osservare la capacità dei metodi sviluppati sul benchmark pubblico di generalizzare a uno scenario operativo concreto.

## 2.3 Preprocessing dei dati

Prima dell'applicazione dei metodi, che saranno analizzati nel dettaglio nel Capitolo 3 della presente tesi, i dati LiDAR sono stati sottoposti a un preprocessing essenziale, limitato alle sole operazioni necessarie per garantirne un utilizzo coerente nelle analisi successive.

La prima operazione ha riguardato la rimozione dei punti coincidenti con la posizione del sensore, ossia dei punti con distanza euclidea nulla, considerati non informativi ai fini dell'analisi. In questo modo, sono stati esclusi i punti privi di reale contenuto geometrico, mantenendo soltanto quelli associati a misure valide.

Un secondo intervento ha riguardato il valore di intensità. Nel dataset LiDARDustX, l'intensità è stata normalizzata nell'intervallo  $[0, 1]$  mediante una riscalatura dipendente dal sensore, poiché le acquisizioni provengono da dispositivi differenti, ciascuno caratterizzato da una propria scala di intensità. Questa operazione ha permesso di rendere più omogenei i valori di intensità tra frame acquisiti con sensori diversi. Nei dati reali aziendali, invece, l'intensità risultava già espressa nell'intervallo  $[0, 1]$  e quindi non sono state apportate ulteriori modifiche.

Nel complesso, il preprocessing adottato è stato limitato alle sole operazioni comuni a tutti i metodi considerati, finalizzate a rimuovere misure non significative e a rendere confrontabili le caratteristiche utilizzate. Le trasformazioni successive dell'input, richieste solo da specifici approcci e legate alla particolare rappresentazione del dato, verranno descritte nel paragrafo seguente e approfondite nel Capitolo 3.

## 2.4 Rappresentazione dei dati

Dopo le operazioni di preprocessing, le acquisizioni LiDAR vengono rappresentate come insiemi non strutturati di punti tridimensionali, in cui ciascun punto è descritto dalle coordinate spaziali  $(x, y, z)$  e dal corrispondente valore di intensità ( $I$ ). Questa struttura, che costituisce la forma base del dato, riflette la natura intrinsecamente

sparsa e irregolare delle nuvole di punti, che non sono quindi organizzate su una griglia regolare come accade, ad esempio, nel caso delle immagini.

Nel caso dei metodi geometrici, i dati vengono elaborati direttamente in questa forma originaria, ovvero come nuvole di punti 3D sulle quali è possibile calcolare proprietà legate alla disposizione spaziale dei punti, alla struttura locale della scena e all'intensità del segnale. In questo senso, la rappresentazione del dato coincide sostanzialmente con la sua forma nativa.

Per gli approcci neurali, invece, la rappresentazione dipende dalla struttura del modello considerato. Nel presente lavoro sono state adottate due modalità principali. La prima è una rappresentazione point-based, nella quale il frame viene trattato come insieme di punti 3D con intensità associata, ed è stata utilizzata nel modello RandLA-Net che opera direttamente sulla nuvola di punti. La seconda modalità, invece, è una rappresentazione di tipo *range-view* multilivello, ottenuta proiettando la nuvola di punti su una griglia bidimensionale dipendente dal sensore e mantenendo, per ciascuna cella, i primi  $K$  punti associati, ordinati secondo la loro distanza dal sensore. Quest'ultima rappresentazione è stata utilizzata nei modelli WeatherNet e U-Net, che richiedono dati organizzati in forma regolare per poter essere elaborati da architetture convoluzionali. Tali rappresentazioni consentono quindi di confrontare approcci che lavorano direttamente sul dato tridimensionale originario con metodi che richiedono invece una riorganizzazione preliminare della nuvola di punti. I dettagli specifici delle trasformazioni adottate per ciascun metodo, così come le scelte implementative relative alla costruzione degli input, verranno discussi nel Capitolo 3, in relazione alle singole architetture.

## 2.5 Formulazione del problema

Nel presente lavoro, il tema della dust detection viene affrontato come un problema di segmentazione binaria point-wise su nuvole di punti LiDAR. Sia

$$P = \{p_i \mid i = 1, \dots, N\}$$

l'insieme dei punti che costituiscono un singolo frame LiDAR, dove  $N$  può variare da frame a frame. Ogni punto

$$p_i = (x_i, y_i, z_i, I_i) \in \mathbb{R}^4$$

è descritto dalle coordinate spaziali  $(x_i, y_i, z_i)$  e dal corrispondente valore di intensità  $I_i$ . L'obiettivo del problema è associare a ciascun punto  $p_i$  un'etichetta binaria

$y_i \in \{0, 1\}$ , dove

$$y_i = \begin{cases} 1 & \text{se il punto appartiene alla classe } dust, \\ 0 & \text{se il punto appartiene alla classe } non-dust. \end{cases}$$

Di conseguenza, a ogni frame è associato un vettore di etichette

$$\mathbf{y} = (y_1, \dots, y_N) \in \{0, 1\}^N.$$

Per ciascun punto viene quindi calcolato uno score  $s_i \in \mathbb{R}$ , che misura il grado di appartenenza del punto alla classe *dust*. L'etichetta predetta  $\hat{y}_i$  viene ottenuta confrontando tale score con una soglia decisionale  $\tau$ , secondo la regola

$$\hat{y}_i = \begin{cases} 1 & \text{se } s_i \geq \tau, \\ 0 & \text{se } s_i < \tau. \end{cases}$$

Nel caso dei metodi geometrici, lo score è ottenuto tramite regole esplicite basate su proprietà locali della nuvola di punti, come geometria, distanza dal sensore, intensità o densità dei punti. Nel caso dei metodi neurali, invece, lo score è prodotto dal modello a partire dai dati di input e può essere interpretato come una misura di confidenza o come una probabilità di appartenenza alla classe *dust*.

La scelta di una formulazione binaria è coerente con l'obiettivo principale della tesi, che non consiste nella segmentazione completa della scena, ma nell'individuazione della polvere come fattore di degrado della percezione LiDAR. A tal fine, nel dataset LiDARDustX le annotazioni originali sono state ricondotte a due sole classi: *dust* e *non-dust*. Questa scelta semplifica il problema rispetto a una segmentazione multiclasse, poiché richiede di distinguere unicamente la polvere dal resto dei punti della scena, mantenendo l'attenzione sul reale obiettivo applicativo del lavoro.

Dal punto di vista teorico, il problema presenta diverse criticità. La polvere può infatti manifestarsi come insiemi di punti sparsi, irregolari e spesso poco densi, con caratteristiche che dipendono dalla distanza, dall'angolo di osservazione, dal sensore utilizzato e dalle condizioni ambientali. Inoltre, il passaggio da un dataset pubblico annotato a dati reali acquisiti in ambito aziendale introduce un ulteriore fattore di complessità legato al cambio di dominio: ciò rende particolarmente rilevante valutare non solo la capacità dei metodi di riconoscere correttamente la polvere nel contesto di addestramento, ma anche il loro comportamento in scenari differenti e meno controllati.

# Capitolo 3

## Metodologie

### 3.1 Introduzione ai metodi considerati

Nel presente lavoro vengono confrontati cinque metodi per il problema della dust detection su dati LiDAR, riconducibili a due famiglie principali: filtri geometrici e modelli neurali. La scelta di includere entrambe le categorie nasce dall'esigenza di analizzare approcci differenti non solo in termini di accuratezza, ma anche di interpretabilità, costo computazionale e robustezza rispetto al cambio di dominio.

I filtri geometrici si basano su regole esplicite costruite a partire da proprietà osservabili della nuvola di punti, quali intensità del ritorno, distanza dal sensore, posizione nello spazio e densità locale. In questa tesi vengono considerati due metodi di questo tipo, LIDROR e DMNR, che affrontano il problema della rimozione della polvere tramite criteri adattivi ma fondati su ipotesi differenti.

Accanto a questi metodi, sono stati analizzati tre approcci neurali: RandLA-Net, WeatherNet e U-Net. A differenza dei filtri geometrici, tali modelli non definiscono esplicitamente le regole di separazione tra punti dust e non-dust, ma apprendono dai dati le caratteristiche più utili alla discriminazione. I tre modelli differiscono inoltre per la rappresentazione del dato in ingresso: RandLA-Net opera direttamente sulla nuvola di punti, mentre WeatherNet e U-Net lavorano su una rappresentazione *range-view* multilivello ottenuta proiettando il frame LiDAR su una griglia regolare.

Il confronto tra questi approcci è particolarmente interessante perché permette di analizzare due strategie profondamente diverse: da un lato, i filtri geometrici sono più semplici e più leggibili dal punto di vista interpretativo; dall'altro, i modelli neurali sono in grado di catturare relazioni più complesse tra geometria locale, intensità e contesto spaziale, ma richiedono una fase di addestramento e una maggiore disponibilità di risorse computazionali. Nelle sezioni seguenti, ciascun metodo viene descritto mettendone in evidenza il principio di funzionamento, le

ipotesi di base e gli aspetti più rilevanti ai fini del confronto sperimentale. Le specifiche configurazioni adottate negli esperimenti, così come la scelta dei parametri e degli iperparametri, saranno discusse nel Capitolo 4.

### 3.1.1 Filtro LIDROR

Il primo approccio geometrico considerato in questa tesi è LIDROR (*Low-Intensity Dynamic Radius Outlier Removal*), introdotto in [1]. L'idea di base del metodo è che molti punti appartenenti alla classe *dust* presentano valori di intensità bassi e compaiono spesso come ritorni isolati rispetto alla distribuzione locale dei punti.

LIDROR può essere descritto come un filtro in due stadi. Nel primo stadio viene applicato un criterio basato sull'intensità: per ogni punto  $p_i$ , descritto da coordinate spaziali  $(x_i, y_i, z_i)$  e intensità  $I_i$ , si verifica la condizione

$$I_i \leq \tau_I,$$

dove  $\tau_I$  è una soglia di intensità. Solo i punti che soddisfano questa condizione vengono considerati candidati alla classe *dust*, mentre gli altri vengono direttamente classificati come *non-dust*.

Nel secondo stadio, per ciascun punto candidato viene valutata la presenza di un sufficiente supporto locale. Il raggio di ricerca non è fisso, ma dipende dalla distanza orizzontale del punto dal sensore. In particolare, si considera una relazione del tipo

$$R_i = \max(r_{\min}, k\alpha r_{xy,i}), \quad r_{xy,i} = \sqrt{x_i^2 + y_i^2},$$

dove  $r_{xy,i}$  denota la distanza del punto dal sensore sul piano orizzontale,  $r_{\min}$  è un raggio minimo,  $k$  è un coefficiente di scala e  $\alpha$  rappresenta la risoluzione angolare del sensore.

Questa scelta permette di tenere conto del fatto che la densità della nuvola di punti diminuisce con l'aumentare della distanza dal sensore. Di conseguenza, un criterio basato su un raggio costante risulterebbe meno adatto nelle regioni più lontane.

Nella versione adottata in questa tesi, il controllo sul vicinato non è stato formulato contando esplicitamente il numero di punti contenuti all'interno del raggio dinamico; si è invece confrontata la distanza del punto dal vicino di rango  $M$  con il raggio  $R_i$ . In questo modo, un punto candidato viene classificato come *non-dust* se tale distanza è minore o uguale a  $R_i$ , mentre viene classificato come *dust* nel caso contrario. I parametri principali del filtro sono dunque  $\tau_I$ , che controlla la selezione

iniziale dei candidati,  $r_{\min}$  e  $k$ , che determinano il raggio dinamico, e  $M$ , che regola il supporto locale minimo richiesto.

Nel complesso, LIDROR combina due criteri semplici e complementari: da un lato, seleziona i punti con intensità ridotta; dall'altro, verifica se tali punti risultano sufficientemente supportati dal vicinato locale. In questo modo, il metodo evita di classificare come *dust* tutti i ritorni deboli, limitandosi a quelli che appaiono anche isolati rispetto al contesto circostante. Per questa ragione, LIDROR rappresenta una baseline geometrica semplice, interpretabile e utile come termine di confronto con approcci più complessi.

### 3.1.2 Filtro DMNR

Il secondo filtro geometrico considerato è DMNR (*Dynamic Multi-threshold Noise Removal*), proposto in [3]. A differenza di LIDROR, DMNR basa la classificazione dei punti su una combinazione di più fattori, che dipendono dalla geometria del frame, dalla distanza dal sensore, dall'intensità e dalla densità locale.

Anche DMNR può essere descritto come un metodo in due stadi. Nel primo stadio viene introdotta una soglia di riferimento  $H_i$  che dipende dalla distanza del punto dal sensore. Indicando con

$$d_i = \sqrt{x_i^2 + y_i^2 + z_i^2}$$

la distanza tridimensionale del punto e con  $z_i$  la sua coordinata verticale, si considera la quantità

$$H_i = \frac{h_1}{\max(d_i, \varepsilon)} + h_2,$$

dove  $\varepsilon$  è una costante positiva molto piccola introdotta per evitare divisioni per zero. I parametri  $h_1$  e  $h_2$ , invece, dipendono dalle caratteristiche globali del frame; in particolare,

$$h_1 = \frac{m_x}{2}, \quad h_2 = \min_j z_j - 1,$$

dove  $m_x = \max_j d_j$  rappresenta la massima distanza dei punti dal sensore nel frame considerato. I punti che soddisfano la condizione

$$z_i > H_i$$

vengono classificati direttamente come *non-dust* e non passano allo stadio successivo. Questo primo passaggio serve a escludere subito i punti che, per posizione geometrica, risultano compatibili con la scena osservata.

Nel secondo stadio, invece, per ciascun punto candidato viene calcolata la distanza media dai suoi  $k$  vicini più prossimi, indicata con  $\bar{d}_i^{(k)}$ . A partire da questa quantità, il metodo definisce una soglia dinamica del tipo

$$T_i = \mu (k_1 e^{k_2 d_i} + k_3 I_i) d_i,$$

dove  $I_i$  è l'intensità normalizzata del punto,  $\mu$  rappresenta la media, calcolata sull'intero frame, delle distanze medie ai  $k$  vicini più prossimi, e  $k_1$ ,  $k_2$ ,  $k_3$  sono coefficienti di scala. Il punto viene classificato come *non-dust* se

$$\bar{d}_i^{(k)} < T_i,$$

mentre viene classificato come *dust* in caso contrario.

L'idea alla base di questo secondo passaggio è che la classificazione non dipende da una soglia costante, ma da una regola adattiva. Per ciascun punto candidato, la quantità  $\bar{d}_i^{(k)}$  misura il grado di isolamento locale: valori più elevati indicano infatti punti meno supportati dal contesto circostante. La soglia con cui questa quantità viene confrontata dipende però anche da  $\mu$ , cioè dalla media, calcolata sull'intero frame, delle distanze medie ai  $k$  vicini più prossimi. In questo modo, la decisione finale tiene conto sia della configurazione locale del singolo punto sia delle caratteristiche complessive della nuvola nel frame considerato. Inoltre, la presenza di  $d_i$  e  $I_i$  nella definizione di  $T_i$  permette di adattare ulteriormente il criterio decisionale alla distanza dal sensore e all'intensità del ritorno.

Nel complesso, DMNR introduce un criterio geometrico più flessibile rispetto a LIDROR, grazie all'impiego di una soglia dinamica che tiene conto contemporaneamente di proprietà locali e caratteristiche globali del frame. In questa formulazione, il parametro  $k$  controlla il numero di vicini utilizzati sia nel calcolo della quantità locale  $\bar{d}_i^{(k)}$  sia nella media globale  $\mu$ , mentre i parametri  $k_1$ ,  $k_2$  e  $k_3$  regolano il comportamento della soglia dinamica, determinando rispettivamente il peso del termine dipendente dalla distanza, la sua variazione al crescere della distanza dal sensore e il contributo dell'intensità alla soglia finale.

### 3.1.3 RandLA-Net

Tra gli approcci neurali considerati in questa tesi, RandLA-Net costituisce il modello point-based di riferimento [6]. Il modello originale è stato proposto

per la segmentazione semantica di nuvole di punti di grandi dimensioni, con l'obiettivo di mantenere un buon compromesso tra capacità descrittiva ed efficienza computazionale. L'idea di base è ridurre progressivamente il numero di punti elaborati mediante *random sampling*, compensando la possibile perdita di informazione con moduli di aggregazione locale capaci di sfruttare in modo esplicito la struttura geometrica del vicinato.

Per ogni punto in ingresso, la rete utilizza come features iniziali le coordinate spaziali e l'intensità normalizzata. Queste informazioni vengono prima proiettate in uno spazio di features di dimensione maggiore e successivamente aggiornate tramite una sequenza di *Dilated Residual Block*. In ciascun blocco, il punto non viene elaborato in modo isolato, ma con un insieme di vicini selezionati nello spazio tridimensionale mediante una ricerca  $k$ -NN dilatata. Le nuove features vengono così costruite combinando le proprietà del punto con le informazioni provenienti dal suo intorno locale.

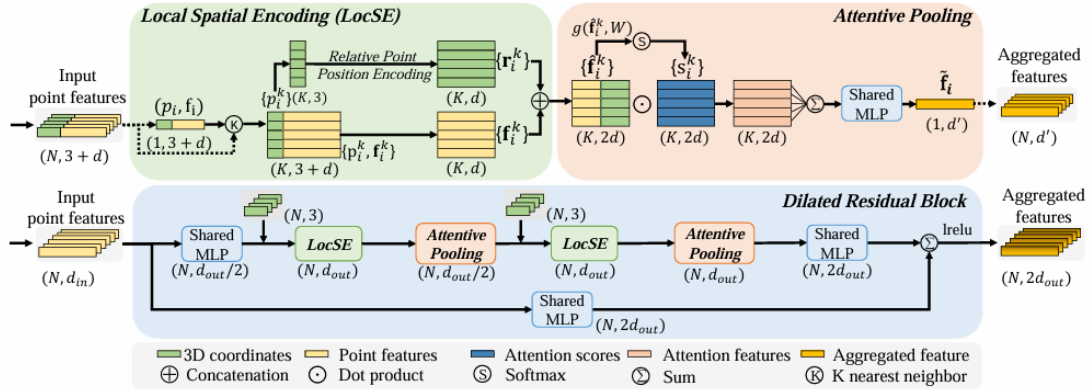


Figura 3.1: Schema del blocco di aggregazione locale e del blocco residuo dilatato in RandLA-Net. Immagine tratta da [6].

Come mostrato in Figura 3.1, ciascun Dilated Residual Block combina una connessione residua con un ramo principale dedicato all'aggregazione delle informazioni locali: in particolare, la connessione residua permette di preservare l'informazione in ingresso e di facilitare l'addestramento della rete; il ramo principale, invece, contiene due fasi successive di *Local Feature Aggregation*, separate da una proiezione intermedia delle features.

All'interno del modulo di *Local Feature Aggregation*, la geometria del vicinato viene esplicitata tramite *Local Spatial Encoding*, che combina le coordinate del punto centrale, quelle dei punti vicini, la loro differenza vettoriale e la loro distanza euclidea. Le features così ottenute vengono poi fuse mediante *Attentive Pooling*,

che assegna un peso diverso ai contributi dei vari vicini prima di produrre una rappresentazione locale del punto. In questo modo, il modello non si limita a raccogliere informazioni dal vicinato, ma impara anche a stabilire quali siano le più rilevanti.

Dal punto di vista architetturale, la rete segue una struttura encoder-decoder. L'architettura implementata è composta da quattro livelli di encoder, un livello centrale e quattro livelli di decoder. Nell'encoder, la nuvola di punti viene progressivamente sottocampionata mediante *random sampling*, mentre le features vengono arricchite attraverso i *Dilated Residual Block*. Il livello centrale applica un'ulteriore trasformazione alle features prodotte dall'encoder, prima della fase di ricostruzione. Infine il decoder riporta le features alla risoluzione iniziale mediante operazioni di *upsampling*, basate sul vicino più prossimo, e *skip connection*, così da combinare informazione globale e dettaglio locale.

Nel presente lavoro, RandLA-Net è stato adattato al problema della segmentazione binaria point-wise della polvere. L'output del modello consiste quindi in un logit per punto, utilizzato per discriminare tra classe *dust* e classe *non-dust*.

### 3.1.4 WeatherNet

Tra gli approcci neurali basati su una rappresentazione regolare su griglia e considerati in questo lavoro, il primo è WeatherNet, ispirato al modello proposto in [5]. Il lavoro originale introduce WeatherNet come rete convoluzionale per il *de-noising* di point cloud LiDAR in presenza di condizioni atmosferiche avverse. Nel presente lavoro, il modello è stato adattato al problema della segmentazione binaria della polvere.

Come anticipato nel Capitolo 2, l'input della rete è costituito dalla rappresentazione *range-view* multilivello, costruita a partire dai canali di distanza e intensità.

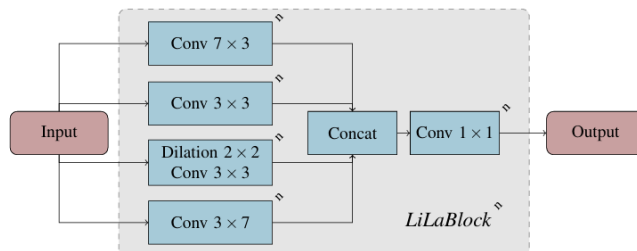


Figura 3.2: Schema del blocco *LiLaBlock* utilizzato in WeatherNet. Immagine tratta da [5]

Come mostrato in Figura 3.2, ciascun *LiLaBlock* è formato da più rami convoluzionali paralleli, caratterizzati da kernel differenti. Questa scelta permette di estrarre features complementari dalla stessa rappresentazione, cogliendo relazioni locali con estensioni diverse lungo le due dimensioni della griglia. Le mappe prodotte dai vari rami vengono quindi concatenate e, successivamente, ricondotte a un nuovo insieme di features tramite una convoluzione  $1 \times 1$ .

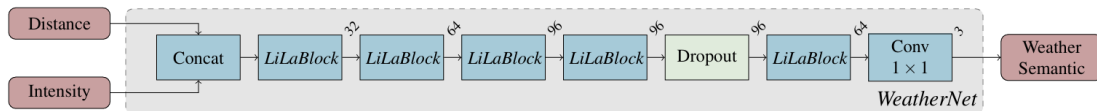


Figura 3.3: Schema dell'architettura WeatherNet. Immagine tratta da [5].

La Figura 3.3 mostra la struttura complessiva della rete. Il modello è composto da quattro *LiLaBlock*, seguiti da un livello di *dropout*, da un quinto *LiLaBlock* e da una convoluzione finale  $1 \times 1$ , che produce l'uscita del modello. I blocchi condividono la stessa struttura, ma non la stessa dimensionalità delle features: il numero di canali cresce nei primi livelli, raggiunge il valore massimo nella parte centrale della rete, per poi ridursi nell'ultimo blocco, secondo la sequenza 32, 64, 96, 96, 64. La rete elabora quindi la rappresentazione su griglia mantenendo invariata la sua risoluzione spaziale.

L'output del modello è una mappa di predizione associata alla rappresentazione su griglia, successivamente ricondotta ai punti del frame per ottenere la classificazione finale.

### 3.1.5 U-Net

Come nel caso di WeatherNet, anche U-Net, proposta in [9], opera sulla rappresentazione *range-view* multilivello introdotta nel Capitolo 2. La scelta di questa architettura si ispira a lavori nei quali la classificazione point-wise di particelle aerodisperse in dati LiDAR viene affrontata a partire da una proiezione bidimensionale della nuvola di punti.

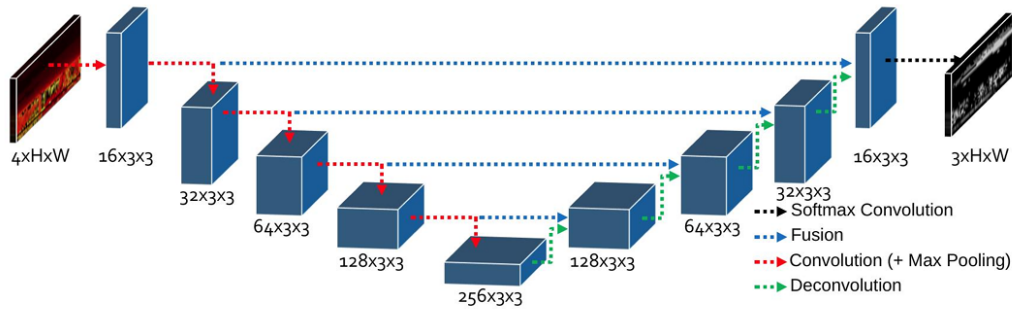


Figura 3.4: Schema dell'architettura U-Net utilizzata per la segmentazione della rappresentazione su griglia. Immagine tratta da [9]

Dal punto di vista architetturale, la rete segue una struttura encoder-decoder articolata in quattro livelli di encoder, un livello centrale e quattro livelli di decoder. Come mostrato in Figura 3.4, ciascun livello dell'encoder è costituito da una convoluzione  $3 \times 3$ , seguita da *batch normalization* e funzione di attivazione. Dopo ogni blocco viene applicato un *max pooling* con finestra  $(1, 2)$ , che riduce la risoluzione soltanto lungo la dimensione orizzontale della griglia, mantenendo invece invariata quella verticale.

Il livello centrale applica un'ulteriore convoluzione  $3 \times 3$  che porta le features a 256 canali. A questo livello, viene inoltre applicato un *dropout*, con l'obiettivo di ridurre il rischio di overfitting prima dell'inizio della fase di ricostruzione.

Nel decoder, la risoluzione viene ripristinata progressivamente mediante operazioni di *upsampling* con fattore  $(1, 2)$ . A ogni livello, le features ottenute vengono concatenate con quelle provenienti dal corrispondente livello dell'encoder tramite *skip connection*; il risultato viene poi elaborato da una nuova convoluzione  $3 \times 3$ , seguita da *batch normalization* e attivazione. In questo modo, il modello combina l'informazione semantica più profonda con il dettaglio locale preservato nei livelli superficiali della rete.

La dimensionalità delle features evolve lungo la rete secondo una struttura simmetrica: 16, 32, 64, 128 canali nell'encoder, 256 nel livello centrale e, quindi, 128, 64, 32, 16 canali nel decoder. L'output finale è ottenuto tramite una convoluzione  $1 \times 1$ , seguita da una funzione sigmoide, che produce la predizione associata alla rappresentazione su griglia e successivamente ricondotta ai punti del frame.

## 3.2 Confronto teorico e considerazioni computazionali

I cinque metodi considerati in questa tesi rispondono al problema della dust detection con strategie differenti, che portano a compromessi diversi tra accuratezza, interpretabilità e costo computazionale.

I filtri geometrici, rappresentati in questo lavoro da LIDROR e DMNR, presentano il vantaggio di una maggiore leggibilità del criterio decisionale. Le regole che portano a classificare un punto come *dust* oppure *non-dust* sono infatti esplicitamente definite tramite soglie e relazioni geometriche. Questo rende tali approcci più facilmente interpretabili e, in linea di principio, più semplici da controllare e adattare. Inoltre, non richiedono una fase di addestramento supervisionato e possono quindi risultare utili in contesti nei quali la disponibilità di dati annotati è limitata. Tuttavia, proprio perché fondati su regole fisse, questi metodi tendono a mostrare una minore flessibilità quando la polvere assume configurazioni molto variabili o si sovrappone in modo ambiguo alla scena utile. In particolare, la sola combinazione di intensità, distanza e densità locale può non essere sufficiente a descrivere pattern più complessi, specialmente in presenza di forti variazioni tra sensori o tra domini acquisitivi differenti.

Gli approcci neurali, al contrario, apprendono dai dati i criteri di separazione tra dust e non-dust. Questo permette loro di modellare relazioni più articolate tra proprietà geometriche locali, intensità e contesto, ma introduce una maggiore dipendenza dalla qualità del dataset di addestramento e dalla capacità del modello di generalizzare a condizioni non osservate durante il training.

All'interno degli approcci neurali, è utile distinguere tra RandLA-Net, che opera direttamente sulla nuvola di punti, e WeatherNet e U-Net, che lavorano invece su una rappresentazione ottenuta proiettando il dato su griglia. Nel primo caso viene mantenuta la struttura tridimensionale originaria del frame, ma l'elaborazione richiede di gestire in modo esplicito le relazioni di vicinato nello spazio 3D, con un costo computazionale generalmente più elevato. Nel secondo caso, invece, la proiezione su griglia consente di sfruttare l'efficienza delle convoluzioni bidimensionali, al prezzo di una riorganizzazione preliminare del dato e di una possibile perdita di informazione; tale perdita è però mitigata, nella rappresentazione adottata in questa tesi, dal mantenimento di più punti per cella ordinati in base alla distanza dal sensore.

Nel complesso, il confronto tra questi metodi non riguarda soltanto la qualità della segmentazione, ma anche il compromesso tra semplicità, capacità espressiva, costo computazionale e robustezza nell'applicazione a dati reali.

# Capitolo 4

## Protocollo sperimentale e metriche di valutazione

In questo capitolo viene descritto il protocollo sperimentale utilizzato per confrontare i metodi presentati nel Capitolo 3. Dopo aver specificato la piattaforma di calcolo e la costruzione degli insiemi di training, validation e test, si introducono le metriche utilizzate per la valutazione quantitativa sul benchmark annotato e le quantità considerate nell'analisi del costo computazionale. Infine, si presentano e discutono i risultati ottenuti sul dataset LiDARDustX e sui dati reali non annotati.

### 4.1 Protocollo sperimentale

#### 4.1.1 Piattaforma di calcolo

Gli esperimenti descritti in questa tesi sono stati eseguiti su una piattaforma di calcolo NVIDIA DGX, in ambiente Ubuntu, utilizzando una GPU NVIDIA GB10 e CUDA 13.0. L'intera pipeline sperimentale è stata implementata in Python: i modelli neurali sono stati sviluppati e addestrati mediante la libreria PyTorch, mentre i metodi geometrici sono stati valutati tramite una pipeline dedicata all'elaborazione delle nuvole di punti.

#### 4.1.2 Suddivisione dei dati

La costruzione degli insiemi di training, validation e test è stata effettuata sulla porzione di LiDARDustX effettivamente disponibile al momento del download, costituita da 174 sequenze per un totale di 7562 frame. Lo split è stato costruito separatamente per ciascun sensore e, all'interno di ogni sensore, a livello di sequenza,

così da preservare la varietà sensoriale del dataset in tutti e tre gli insiemi e mantenere la coerenza temporale delle acquisizioni. La scelta di operare a livello di sequenza ha inoltre evitato che frame consecutivi, o comunque fortemente correlati, fossero assegnati a insiemi diversi, riducendo il rischio di una sovrastima delle prestazioni dovuta alla presenza di campioni quasi sovrapponibili in insiemi distinti.

Lo split è stato quindi definito cercando di mantenere nei tre insiemi una distribuzione il più possibile equilibrata sia rispetto al numero di frame sia rispetto alla proporzione di punti appartenenti alla classe *dust*. La suddivisione finale ottenuta è risultata pari a 5041 frame per il training, 1007 frame per la validation e 1514 frame per il test.

A partire da questa suddivisione, il protocollo sperimentale è stato applicato in modo parzialmente diverso ai due gruppi di metodi. Per i metodi geometrici non è stato impiegato il validation set: la selezione dei parametri è stata effettuata sul training set, mentre il test set è stato riservato esclusivamente alla valutazione finale. Per i modelli neurali, invece, i tre insiemi di training, validation e test sono stati utilizzati rispettivamente per l'addestramento, la selezione del modello e la valutazione finale.

## 4.2 Metriche di valutazione

Coerentemente con la formulazione binaria del problema introdotta nel Capitolo 2, la valutazione quantitativa dei metodi è stata condotta mediante metriche derivate dalla matrice di confusione. Nel caso della dust detection, il ricorso a più metriche risulta particolarmente utile, poiché i diversi approcci possono privilegiare comportamenti differenti: alcuni tendono a rilevare la polvere in modo più esteso, aumentando il rischio di falsi positivi, mentre altri adottano criteri più restrittivi, con il rischio opposto di perdere parte dei punti appartenenti alla classe di interesse. Per questo motivo, il confronto non è stato affidato a un solo indicatore, ma a un insieme di misure complementari, capaci di descrivere sia l'affidabilità delle predizioni positive sia la capacità di recuperare i punti *dust*.

La matrice di confusione è ottenuta confrontando, per ciascun punto della nuvola, l'etichetta predetta con l'etichetta di riferimento. In particolare, si denotano con:

- *TP* (*true positive*) il numero di punti *dust* correttamente classificati;
- *FP* (*false positive*) il numero di punti *non-dust* erroneamente classificati come *dust*;

- $FN$  (*false negative*) il numero di punti *dust* erroneamente classificati come *non-dust*;
- $TN$  (*true negative*) il numero di punti *non-dust* correttamente classificati.

La *precision* misura l'affidabilità delle predizioni positive, cioè la frazione dei punti classificati come *dust* che appartengono effettivamente a tale classe. Essa penalizza quindi i *false positives* ed è definita da

$$\text{Precision} = \frac{TP}{TP + FP}.$$

Il *recall*, invece, misura la capacità del metodo di individuare i punti effettivamente appartenenti alla classe *dust*, penalizzando quindi i *false negatives*. Esso è definito da

$$\text{Recall} = \frac{TP}{TP + FN}.$$

Per disporre di una misura sintetica del compromesso tra questi due aspetti è stato considerato anche l'*F1-score*, definito come

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2TP}{2TP + FP + FN}.$$

Tale metrica risulta utile quando si vuole riassumere in un unico indicatore sia la qualità delle predizioni positive sia la capacità di recuperare la classe di interesse.

Come metrica principale per il confronto tra i metodi è stata tuttavia adottata la *Intersection over Union* della classe *dust*, definita da

$$\text{IoU}_{\text{dust}} = \frac{TP}{TP + FP + FN}.$$

Questa metrica confronta direttamente l'insieme dei punti predetti come *dust* con l'insieme dei punti effettivamente appartenenti a tale classe, penalizzando simultaneamente sia i *false positives* sia i *false negatives*. Rispetto all'*F1-score*, la *IoU* fornisce inoltre una valutazione più severa degli errori, poiché non attribuisce un peso doppio ai *true positives*. Per questa ragione è stata assunta come riferimento principale nella selezione dei parametri dei filtri geometrici, nella scelta dei checkpoint dei modelli neurali e nel confronto finale sul test set.

## 4.3 Misura delle prestazioni computazionali

Oltre alla qualità della segmentazione, nel confronto tra i metodi è stato considerato anche il costo computazionale, così da valutare l'efficienza dei diversi approcci in vista di un possibile impiego operativo.

Per i filtri geometrici, la quantità principale considerata è stata il tempo medio di elaborazione per frame, utilizzato come indicatore diretto del costo del metodo in fase di utilizzo.

Nel caso dei modelli neurali, il confronto computazionale è stato riferito principalmente alla fase di test, considerando il tempo medio di inferenza per frame come indicatore del costo di utilizzo del modello una volta completato l'addestramento. Durante il training sono stati inoltre osservati l'andamento della *loss* e delle metriche sul training set e sul validation set, non come misura diretta del costo computazionale in senso stretto, ma come strumento per analizzare la regolarità della convergenza e la stabilità dell'apprendimento.

## 4.4 Risultati sul dataset LiDARDustX

In questa sezione vengono presentati i risultati ottenuti sul dataset LiDARDustX, assunto come benchmark principale per il confronto tra i metodi considerati. Nelle sottosezioni seguenti si discutono dapprima i risultati dei filtri geometrici, poi quelli dei modelli neurali, per concludere con una sintesi quantitativa complessiva del confronto.

### 4.4.1 Filtri geometrici

Per i filtri geometrici è stato esplorato un insieme di configurazioni ottenute variando i principali parametri di ciascun metodo. In entrambi i casi, la configurazione finale è stata selezionata sul training set assumendo come metrica di riferimento principale la  $\text{IoU}_{\text{dust}}$ .

Per DMNR è stata considerata una griglia di configurazioni ottenuta variando il numero di vicini  $k$  e i coefficienti  $k_1$ ,  $k_2$  e  $k_3$  che intervengono nella definizione della soglia dinamica. In particolare, sono stati esplorati i valori

$$k \in \{3, 4, 10, 20\}, \quad k_1 \in \{0.001, 0.005, 0.01, 0.05, 0.1, 0.5\},$$

$$k_2 \in \{0.001, 0.05, 0.1, 0.3, 0.5\}, \quad k_3 \in \{0.001, 0.005, 0.01, 0.05, 0.1, 0.5\}.$$

La configurazione finale selezionata corrisponde a

$$k = 4, \quad k_1 = 0.05, \quad k_2 = 0.3, \quad k_3 = 0.01.$$

L'analisi della griglia suggerisce che, nel caso di DMNR, il comportamento del metodo sia influenzato soprattutto dai parametri  $k_1$  e  $k_2$ , mentre l'effetto di  $k_3$  risulta meno evidente. In particolare, le configurazioni più favorevoli si concentrano per valori moderati di  $k \in \{3, 4\}$ , con  $k_1$  intorno a 0.05 e  $k_2$  intorno a 0.3, il che suggerisce che il tuning del metodo dipenda soprattutto dalla calibrazione congiunta dei termini principali della soglia dinamica.

Per LIDROR è stata invece esplorata una griglia di configurazioni ottenuta variando la soglia di intensità  $I$ , il numero minimo di punti  $M$ , il raggio minimo  $r_{\min}$  e il coefficiente  $k$  del raggio dinamico. In particolare, sono stati considerati i valori

$$I \in \{0.01, 0.015, 0.02, 0.025, 0.03, 0.035, 0.04, 0.045, 0.05\},$$

$$M \in \{3, 5, 10, 20, 50, 100\}, \quad r_{\min} \in \{0.001, 0.005, 0.01\},$$

$$k \in \{0.001, 0.005, 0.01\}.$$

La configurazione finale selezionata corrisponde a

$$I = 0.025, \quad M = 100, \quad r_{\min} = 0.01, \quad k = 0.001.$$

L'analisi della griglia mostra che il comportamento del metodo dipende in misura preponderante dalla soglia di intensità  $I$ . In particolare, le configurazioni migliori si ottengono sistematicamente con  $M = 100$  e  $r_{\min} = 0.01$ , mentre la variazione di  $k$  nell'intervallo considerato produce effetti trascurabili sulle metriche finali. Il miglior compromesso si osserva per valori intermedi della soglia, in particolare intorno a  $I = 0.025$ , che corrisponde anche al miglior valore di  $\text{IoU}_{\text{dust}}$ .

Nel complesso, l'analisi del tuning suggerisce che, sul benchmark LiDARDustX, il comportamento di LIDROR sia determinato quasi interamente dal primo stadio basato sulla soglia di intensità, mentre il secondo stadio di raffinamento geometrico fornisce un contributo molto limitato. Per verificare più direttamente questo punto, la Tabella 4.1 confronta, per ciascun valore di soglia  $I$  comune ai due esperimenti, i risultati ottenuti dal solo primo stadio e quelli della migliore configurazione completa di LIDROR corrispondente.

$I$	F1-score		IoU <sub>dust</sub>	
	Primo stadio	LIDROR	Primo stadio	LIDROR
0.010	0.389 14	0.390 28	0.241 57	0.242 45
0.015	0.444 61	0.446 18	0.285 85	0.287 15
0.020	0.491 50	0.492 59	0.325 82	0.326 78
0.025	0.498 10	0.498 66	0.331 65	0.332 15
0.030	0.489 23	0.489 33	0.323 83	0.323 92
0.035	0.454 22	0.454 12	0.293 84	0.293 76
0.040	0.386 05	0.385 93	0.239 20	0.239 11
0.050	0.360 98	0.360 81	0.220 24	0.220 12

Tabella 4.1: Confronto tra il primo stadio di LIDROR e la migliore configurazione completa di LIDROR per ciascun valore della soglia di intensità  $I$ .

La tabella mostra che le differenze tra i due approcci sono estremamente contenute: per tutti i valori di intensità  $I$  considerati, gli scarti osservati in termini di  $F1$ -score e IoU<sub>dust</sub> rimangono dell'ordine di  $10^{-3}$ . Questo costituisce un risultato rilevante dell'analisi sperimentale, perché indica che, sul benchmark LiDARDustX, il secondo stadio di LIDROR fornisce un contributo molto limitato e che il metodo si comporta in modo quasi coincidente con il proprio primo stadio basato sulla sola intensità.

## 4.4.2 Modelli neurali

Per i modelli neurali, il confronto è stato condotto tenendo conto sia delle prestazioni finali sul test set sia del comportamento osservato durante l'addestramento. In tutti i casi, il checkpoint finale è stato selezionato sulla base della IoU<sub>dust</sub> osservata sul validation set. Durante l'addestramento sono state inoltre registrate, epoca per epoca, le curve di  $loss$  e di IoU<sub>dust</sub> su training set e validation set, successivamente utilizzate per verificare la regolarità dell'apprendimento e la coerenza tra andamento in validazione e risultati finali sul test set.

Dal punto di vista della rappresentazione del dato, U-Net e WeatherNet operano sulla *range-view* multilivello introdotta nei capitoli precedenti. In particolare, ogni frame è stato proiettato su una griglia di dimensione  $128 \times 1024$ , mantenendo per ciascuna cella i primi 10 punti ordinati rispetto alla distanza dal sensore. L'input del modello è stato costruito associando, a ciascuno dei 10 livelli della griglia, due canali distinti: uno contenente la distanza dal sensore e l'altro contenente il valore di intensità. In questo modo si ottiene un tensore di input a 20 canali, mentre l'output

atteso è stato mantenuto su 10 livelli. La scelta di mantenere 10 punti per cella è stata adottata in modo da preservare tutti i punti del frame nella rappresentazione utilizzata negli esperimenti, evitando perdita di informazione nella proiezione su griglia e mantenendo il calcolo delle metriche coerente con la nuvola originale.

RandLA-Net, invece, opera direttamente sulla nuvola di punti, senza richiedere una proiezione su griglia. In fase di addestramento e validazione, ogni frame è stato ricampionato casualmente a 16 384 punti per epoca, così da mantenere fisso il numero di punti elaborati in ciascun passo e, al tempo stesso, esporre il modello a porzioni diverse della nuvola. In questo modo, pur non utilizzando tutti i punti del frame in ogni singola epoca, il modello osserva nel corso dell’addestramento una parte molto ampia della nuvola, grazie al carattere casuale del ricampionamento. In fase di test, invece, il modello è stato applicato a tutti i punti del frame, così da produrre una classificazione sull’intera scena e preservare la coerenza della valutazione con la nuvola originale. Nella specifica implementazione adottata in questa tesi, la costruzione del vicinato è stata realizzata mediante il calcolo esplicito delle distanze *pairwise*, per ragioni di compatibilità con l’ambiente di esecuzione utilizzato.

Poiché il problema presenta uno sbilanciamento tra la classe *dust* e la classe *non-dust*, l’addestramento è stato condotto mediante una *binary cross-entropy* pesata, in modo da attribuire maggiore importanza agli errori commessi sulla classe di interesse. Indicando con  $y \in \{0, 1\}$  l’etichetta target e con  $\hat{p} \in [0, 1]$  la probabilità predetta, la *loss* elementare può essere scritta nella forma

$$\mathcal{L}_{\text{BCE}}(y, \hat{p}) = -w^+ y \log(\hat{p}) - (1 - y) \log(1 - \hat{p}),$$

dove  $w^+$  rappresenta il peso assegnato alla classe positiva.

Sia i pesi della *loss* sia la soglia decisionale  $\tau$ , utilizzata in fase di inferenza per ottenere la classificazione binaria finale, sono stati scelti sulla base di prove preliminari sul validation set, assumendo come riferimento principale la  $\text{IoU}_{\text{dust}}$ . Nel caso di U-Net e WeatherNet, che operano sulla rappresentazione *range-view*, la *loss* è stata calcolata solo sulle posizioni valide della griglia, ottenendo così una *masked binary cross-entropy*, ed è stato utilizzato  $w^+ = 4.0$  per entrambi i modelli. Per RandLA-Net, invece, è stata impiegata una *binary cross-entropy with logits*, adottando un peso per la classe positiva pari a  $w^+ = 8.0$ . Le soglie decisionali finali selezionate corrispondono a  $\tau = 0.7$  per U-Net,  $\tau = 0.8$  per WeatherNet e  $\tau = 0.8$  per RandLA-Net.

Per valutare la stabilità dell’addestramento e la coerenza tra comportamento in validazione e prestazioni finali sul test set, si riportano nelle Figure 4.1 e 4.2 le curve di  $\text{IoU}_{\text{dust}}$  e della *loss* dei tre modelli neurali.

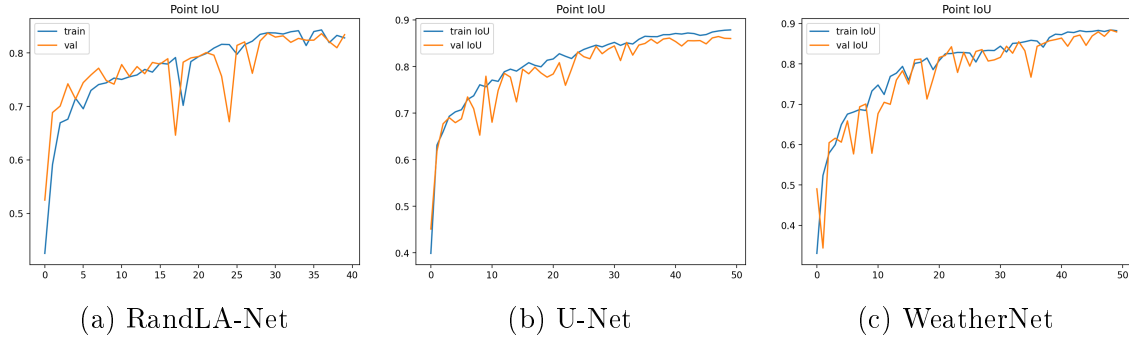


Figura 4.1: Andamento della  $\text{IoU}_{\text{dust}}$  su training e validation durante l'addestramento dei tre modelli neurali.

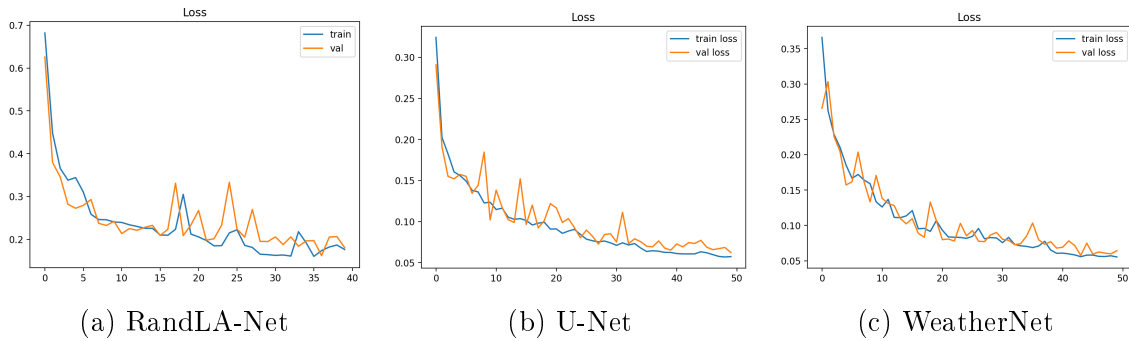


Figura 4.2: Andamento della loss su training e validation durante l'addestramento dei tre modelli neurali.

Dall'osservazione delle Figure 4.1 e 4.2 emerge un apprendimento complessivamente regolare per tutti e tre i modelli. In particolare, l'andamento della  $\text{IoU}_{\text{dust}}$  sul validation set conferma una chiara superiorità di WeatherNet, seguito da U-Net e RandLA-Net, in coerenza con quanto osservato successivamente sul test set. Allo stesso tempo, le curve della *loss* non mostrano instabilità marcate o divergenze tra training e validation tali da far pensare a un comportamento anomalo in fase di addestramento. Nel complesso, le figure suggeriscono quindi che i checkpoint selezionati non corrispondano a fluttuazioni accidentali nelle ultime epoche, ma riflettano un comportamento già chiaramente visibile durante il processo di apprendimento.

### 4.4.3 Analisi quantitativa dei risultati

Per sintetizzare il confronto quantitativo tra tutti i metodi considerati, la Tabella 4.2 riporta i principali risultati ottenuti sul test set di LiDARDustX, includendo sia le metriche di segmentazione sia il tempo medio di elaborazione per frame.

Metodo	IoU	Precision	Recall	F1	Tempo per frame
LIDROR	0.3301	0.4828	0.5107	0.4964	241.1 ms
DMNR	0.4564	0.8670	0.4908	0.6268	17.1 ms
RandLA-Net	0.7789	0.9084	0.8453	0.8757	514.0 ms
U-Net	0.8331	0.8710	0.9504	0.9090	15.2 ms
WeatherNet	0.8609	0.9103	0.9407	0.9253	74.6 ms

Tabella 4.2: Confronto quantitativo dei metodi sul test set di LiDARDustX.

Come mostrato in Tabella 4.2, il confronto quantitativo evidenzia in modo netto la superiorità degli approcci neurali rispetto ai filtri geometrici. Anche il meno performante tra i modelli neurali, RandLA-Net, raggiunge infatti una  $\text{IoU}_{\text{dust}} = 0.7789$ , valore ampiamente superiore a quello del migliore filtro geometrico, DMNR, che si ferma a 0.4564. Questo divario suggerisce che, nel setup considerato, regole esplicite basate su intensità, distanza e densità locale non siano sufficienti a descrivere in modo accurato la variabilità del fenomeno, mentre i modelli neurali risultano più efficaci nel cogliere relazioni più complesse tra geometria della nuvola, intensità del ritorno e contesto spaziale.

Il confronto tra i due filtri geometrici mostra inoltre che DMNR risulta preferibile a LIDROR sia sul piano prestazionale sia su quello computazionale. In termini di  $\text{IoU}_{\text{dust}}$ , DMNR è sensibilmente migliore rispetto a LIDROR, passando da 0.3301 a 0.4564. Le metriche elementari mostrano però due comportamenti differenti: LIDROR presenta valori di precision e recall relativamente vicini (0.4828 e 0.5107), mentre DMNR adotta una strategia molto più conservativa, con precision elevata (0.8670) ma recall più contenuta (0.4908). Ciò indica che DMNR tende a classificare come polvere un insieme più ristretto di punti, ma con una quota di falsi positivi molto più bassa. A questo si aggiunge un vantaggio computazionale marcato, dal momento che DMNR richiede in media 17.1 ms per frame contro i 241.1 ms di LIDROR.

All'interno dei modelli neurali, WeatherNet ottiene il miglior risultato complessivo, con  $\text{IoU}_{\text{dust}} = 0.8609$ , seguito da U-Net con 0.8331 e da RandLA-Net con 0.7789. Le metriche di precision e recall confermano sostanzialmente questo ordinamento, ma mettono anche in evidenza differenze nel profilo di classificazione: WeatherNet mantiene un equilibrio molto favorevole tra precision (0.9103) e recall (0.9407); U-Net mostra una recall ancora più alta (0.9504) a fronte di una precision

leggermente inferiore (0.8710), suggerendo una segmentazione più estesa della classe *dust*; RandLA-Net, invece, pur restando nettamente superiore ai filtri geometrici, risulta il meno efficace tra i modelli neurali, soprattutto per una recall più contenuta (0.8453). Anche in questo caso, dal punto di vista computazionale, le differenze tra i modelli neurali sono molto marcate. U-Net risulta complessivamente il più rapido, con un tempo medio di inferenza di 15.2 ms per frame, inferiore non solo a quello di WeatherNet (74.6 ms), ma anche a quello di entrambi i filtri geometrici. WeatherNet presenta un costo intermedio, che appare giustificato dal miglior risultato quantitativo complessivo. RandLA-Net, invece, a causa del calcolo esplicito delle distanze pairwise richiede 514.0 ms per frame, risultando di gran lunga il metodo più oneroso. Nel presente lavoro, questi risultati indicano quindi che gli approcci neurali basati su rappresentazione *range-view* multilivello offrono il compromesso più favorevole tra accuratezza della segmentazione ed efficienza inferenziale.

## 4.5 Risultati sui dati reali

Accanto alla valutazione quantitativa sul benchmark pubblico, i metodi considerati sono stati applicati anche ai dati reali raccolti in ambito aziendale. Poiché tali acquisizioni non disponevano di annotazioni point-wise della polvere, il confronto non è stato condotto in forma quantitativa, ma mediante un'analisi qualitativa del comportamento dei diversi approcci.

Per rendere tale confronto il più possibile omogeneo, i risultati dei cinque metodi sono stati osservati sugli stessi frame e rappresentati secondo una medesima vista frontale, così da evidenziare in modo diretto le differenze nel rilevamento della polvere. I frame selezionati per questa analisi contenevano scene in cui la presenza di polvere risultava visibile o comunque plausibile sulla base dell'ispezione delle nuvole di punti e delle immagini camera associate.

La Figura 4.3 riporta un esempio rappresentativo del confronto visivo tra i cinque approcci. Nell'immagine camera si osserva un camion in movimento in un contesto di cava, con una nube di polvere visibile nella porzione sinistra della scena. Nelle rappresentazioni LiDAR, i punti classificati come *dust* sono evidenziati in verde, mentre il resto della nuvola è colorato in funzione dell'intensità del ritorno, con toni più scuri per intensità minori e toni più chiari per intensità maggiori.

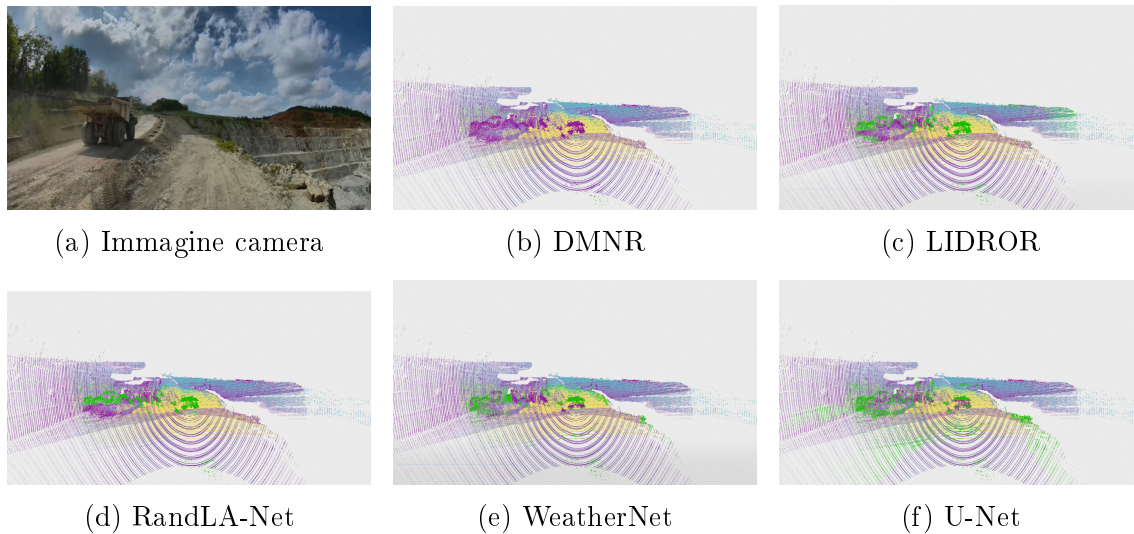


Figura 4.3: Esempio di confronto qualitativo sui dati reali tra immagine camera e segmentazioni prodotte dai cinque metodi considerati.

### 4.5.1 Analisi qualitativa dei risultati

L'analisi qualitativa suggerisce che le differenze tra i metodi non riguardino soltanto la quantità di punti classificati come *dust*, ma soprattutto il tipo di errore che ciascun approccio tende a commettere. La Figura 4.3 consente di cogliere in modo diretto queste differenze.

Tra i filtri geometrici, DMNR appare il meno convincente nel trasferimento ai dati reali: nell'esempio mostrato, la segmentazione positiva è quasi assente e non segue in modo apprezzabile la nube visibile nella scena. LIDROR, pur risultando più attivo, mantiene un comportamento prudente e individua solo porzioni limitate della polvere, soprattutto nella zona sinistra della scena e in alcune regioni centrali. Questa segmentazione appare quindi frammentaria e non riesce a seguire in modo continuo l'estensione della nube.

I modelli neurali, invece, mostrano una maggiore capacità di seguire strutture polverose spazialmente estese, ma con comportamenti differenti. RandLA-Net produce una segmentazione più aderente alla forma della nube e meglio confinata rispetto alla scena utile, risultando il metodo più equilibrato tra quelli osservati. WeatherNet segue la polvere in modo più pieno e continuo, ma tende ad allargare maggiormente la regione classificata, introducendo una quota più elevata di falsi positivi. Infine, U-Net presenta il comportamento più invasivo: nell'esempio considerato, la classificazione positiva si estende in modo marcato anche a porzioni della scena che appaiono visivamente stabili, contaminando in misura più evidente la struttura utile della nuvola di punti. Nel complesso, il confronto qualitativo

mostra quindi che, su dati reali non annotati, il risultato visivamente più convincente non coincide necessariamente con la segmentazione più ampia: in questo contesto diventano infatti centrali non solo la capacità di intercettare la polvere, ma anche il controllo dei falsi positivi, la conservazione della struttura utile della nuvola di punti e la robustezza del metodo quando cambia la geometria complessiva del dato in ingresso.

# Capitolo 5

## Conclusioni

Questa tesi ha affrontato il problema della dust detection su dati LiDAR in ambienti non strutturati, confrontando approcci geometrici basati su regole esplicite e modelli neurali di segmentazione binaria point-wise. L'obiettivo del lavoro non era soltanto individuare il metodo più accurato sul benchmark considerato, ma valutare in modo più ampio il comportamento dei diversi approcci anche in relazione a efficienza computazionale e plausibilità in uno scenario operativo reale.

I risultati ottenuti sul dataset LiDARDustX hanno mostrato una netta superiorità dei modelli neurali rispetto ai filtri geometrici, indicando che la polvere costituisce un fenomeno troppo variabile per essere descritto in modo soddisfacente tramite sole regole locali basate su intensità, distanza e densità dei punti. In particolare, nel setup considerato, gli approcci neurali basati su rappresentazione *range-view* hanno evidenziato il comportamento quantitativo più favorevole, confermando l'efficacia di una rappresentazione regolare del dato per questo specifico problema. Allo stesso tempo, l'analisi qualitativa sui dati reali raccolti in ambito aziendale ha mostrato che la valutazione di un metodo non può essere ricondotta esclusivamente alle metriche ottenute su un benchmark annotato. Nel passaggio a uno scenario operativo differente diventano infatti centrali anche il controllo dei falsi positivi, la capacità di preservare la scena utile e la robustezza del metodo al cambio di dominio. Da questo punto di vista, il lavoro ha evidenziato come il metodo quantitativamente migliore non coincida necessariamente con quello più convincente in termini di comportamento complessivo sui dati reali.

Nel complesso, il contributo principale della tesi consiste quindi nell'aver proposto una valutazione comparativa articolata su più livelli, fondata su confronto quantitativo, analisi qualitativa ed efficienza computazionale. Ne emerge una conclusione generale: nel problema della dust detection, la scelta del metodo più adatto non può basarsi soltanto sulla massimizzazione dell'accuratezza, ma

deve tenere conto dell'equilibrio tra qualità della segmentazione, sostenibilità computazionale e affidabilità in scenari applicativi differenti.

Pur avendo prodotto risultati significativi e coerenti con gli obiettivi iniziali, il lavoro presenta alcuni limiti. Gli esperimenti sul benchmark LiDARDustX sono stati condotti solo su una porzione del dataset, riducendo in parte l'ampiezza del contesto sperimentale. Inoltre, i dati reali raccolti in azienda non disponevano di annotazioni point-wise della polvere e hanno quindi consentito soltanto un confronto qualitativo, senza permettere una verifica quantitativa rigorosa della capacità di generalizzazione dei diversi approcci. Un ulteriore limite riguarda la trasferibilità delle configurazioni selezionate sul benchmark pubblico: il comportamento osservato sui dati reali suggerisce infatti che alcuni metodi, soprattutto quelli geometrici più dipendenti da soglie o quantità statistiche globali, risultino sensibili al cambiamento di dominio e non mantengano automaticamente la stessa efficacia in contesti differenti. Infine, anche per i modelli neurali il confronto è stato svolto su un insieme limitato di architetture e di configurazioni sperimentali, sufficiente per un'analisi comparativa significativa ma non esaustivo rispetto alle possibili alternative metodologiche.

In prospettiva futura, sarebbe quindi utile estendere l'analisi a una porzione più ampia del benchmark, oppure ad altri dataset analoghi, così da rafforzare ulteriormente la solidità del confronto. La costruzione di un insieme di dati reali annotati rappresenta inoltre uno degli sviluppi più rilevanti, perché consentirebbe di valutare quantitativamente i metodi anche in condizioni operative reali e di studiarne con maggiore rigore la robustezza al cambio di dominio. Sarebbe infine interessante approfondire strategie di adattamento al dominio reale, ampliare il confronto ad altre rappresentazioni del dato e ad architetture differenti, e considerare possibili pipeline ibride in cui approcci geometrici e neurali cooperino per migliorare robustezza, interpretabilità e qualità complessiva della segmentazione.

# Bibliografia

- [1] Ali Afzalaghaeinaeini et al. “Design of Dust-Filtering Algorithms for LiDAR Sensors Using Intensity and Range Information in Off-Road Vehicles”. In: *Sensors* 22.11 (2022), p. 4051. DOI: 10.3390/s22114051.
- [2] Bruno Cavieres, Nicolás Cruz e Javier Ruiz-del-Solar. “Dust Filtering in LIDAR Point Clouds Using Deep Learning for Mining Applications”. In: *Sensors* 25.20 (2025), p. 6441. DOI: 10.3390/s25206441.
- [3] Chu Chen et al. “DMNR: Unsupervised De-noising of Point Clouds Corrupted by Airborne Particles”. In: *CoRR* abs/2305.05991 (2023). DOI: 10.48550/arXiv.2305.05991. URL: <https://arxiv.org/abs/2305.05991>.
- [4] Jannis Egelhof, Patrick Wolf e Karsten Berns. “Disturbance and Particle Detection in LiDAR Data”. In: *48th Annual Conference of the IEEE Industrial Electronics Society (IECON 2022)*. Brussels, Belgium, 2022. DOI: 10.1109/IECON49645.2022.9968972.
- [5] Robin Heinzler et al. “CNN-based Lidar Point Cloud De-Noising in Adverse Weather”. In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 2514–2521. DOI: 10.1109/LRA.2020.2972865.
- [6] Qingyong Hu et al. “RandLA-Net: Efficient Semantic Segmentation of Large-Scale Point Clouds”. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 11108–11117. DOI: 10.1109/CVPR42600.2020.01112.
- [7] Xianyao Jiang et al. “Algorithm for Point Cloud Dust Filtering of LiDAR for Autonomous Vehicles in Mining Area”. In: *Sustainability* 16.7 (2024), p. 2827. DOI: 10.3390/su16072827.
- [8] Enrique Martí et al. “A Review of Sensor Technologies for Perception in Automated Driving”. In: *IEEE Intelligent Transportation Systems Magazine* 11.4 (2019), pp. 94–108. DOI: 10.1109/MITS.2019.2907630.

- [9] Leo Stanislas et al. “Airborne Particle Classification in LiDAR Point Clouds Using Deep Learning”. In: *Field and Service Robotics*. A cura di Genya Ishigami e Kazuya Yoshida. Vol. 16. Springer Proceedings in Advanced Robotics. Singapore: Springer, 2021, pp. 395–410. DOI: 10.1007/978-981-15-9460-1\_28.
- [10] Jorge Vargas et al. “An Overview of Autonomous Vehicles Sensors and Their Vulnerability to Weather Conditions”. In: *Sensors* 21.16 (2021), p. 5397. DOI: 10.3390/s21165397.
- [11] Chenfeng Wei et al. “LiDARDustX: A LiDAR Dataset for Dusty Unstructured Road Environments”. In: *2025 IEEE International Conference on Robotics and Automation (ICRA)*. Atlanta, GA, USA, 2025, pp. 12703–12709. DOI: 10.1109/ICRA55743.2025.11127917.