

Facoltà di Ingegneria dell'Informazione, Informatica e Statistica
Corso di Laurea in Informatica (L-31)

Anno Accademico 2021-2022



SAPIENZA
UNIVERSITÀ DI ROMA

Relazione di tirocinio

Studente: William De Vena
e-mail: william98wdv@gmail.com

Responsabile: Prof. Danilo Avola
e-mail: avola@di.uniroma1.it

Corresponsabili: Prof. Luigi Cinque
e-mail: cinque@di.uniroma1.it

Dott. Alessio Fagioli
e-mail: fagioli@di.uniroma1.it



Approccio basato su Deep Learning per la Segmentazione Semantica di Immagini Aeree



Dominio di applicazione



Dominio di applicazione: motivazione

- Disastri naturali sempre **più frequenti e più gravi**
- **Le alluvioni** compongono circa il **43%** del totale degli eventi
- **157 mila** vittime e **2.3 miliardi** di persone colpite (1995-2015)
- **1 miliardo** di dollari di danni (USA nel 2020)
- La fornitura di precise, tempestive e comprensibili informazioni è **fondamentale** nella gestione di questi eventi





Dominio di applicazione: uso degli UAV

- Gli UAV (Unmanned Aerial Vehicle) possono accedere **rapidamente** alle aree colpite
- Possono raggiungere aree altrimenti **non raggiungibili dagli umani**
- Forniscono rapidamente immagini a bassa altitudine e ad **alta risoluzione**



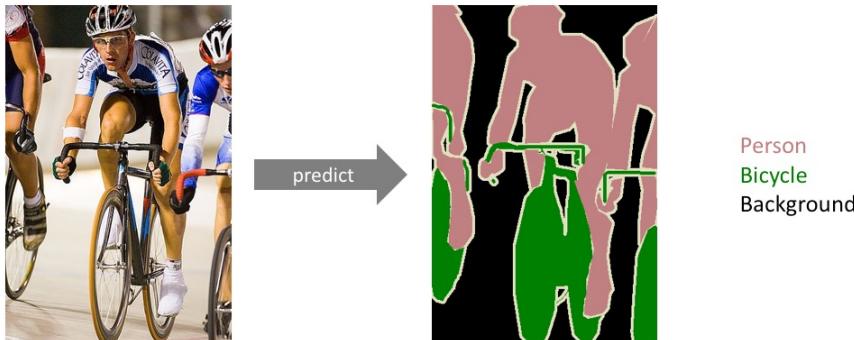


Segmentazione semantica



Segmentazione semantica: overview

- Possiamo definirla come la **classificazione di ogni pixel** dell'immagine in una determinata classe
- Esistono sia metodi tradizionali sia **metodi basati su architetture di Deep Learning**



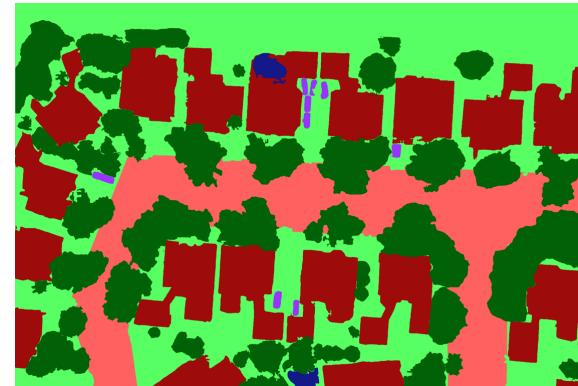


Dataset FloodNet



FloodNet: overview

- Pubblicato a seguito della FloodNet Challenge del workshop **EARTHVISION 2021**
- Catturate tra il 30 Agosto e il 4 Settembre del 2017 in Texas (USA) immediatamente dopo il disastro causato dall'**Uragano Harvey** con un DJI Mavic Pro
- **2343 immagini** catturate a 200 piedi di altitudine e con una risoluzione di 1.5 cm per pixel
- **9 classi**: edificio allagato, edificio non allagato, strada allagata, strada non allagata, acqua, albero, veicolo, piscina, prato





FloodNet: principali difficoltà e soluzioni

- | | | |
|---|--|----------------------------|
| Presenza di errori nelle maschere |  | Data cleaning |
| Forte sbilanciamento delle classi |  | Data augmentation offline |
| Oggetti di diversa scala |  | Architettura context-based |
| Difficoltà intrinseca di alcune classi |  | Architettura context-based |



Approccio proposto



Approccio proposto

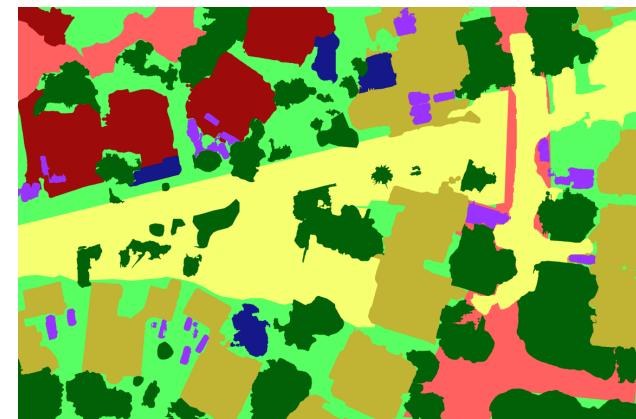
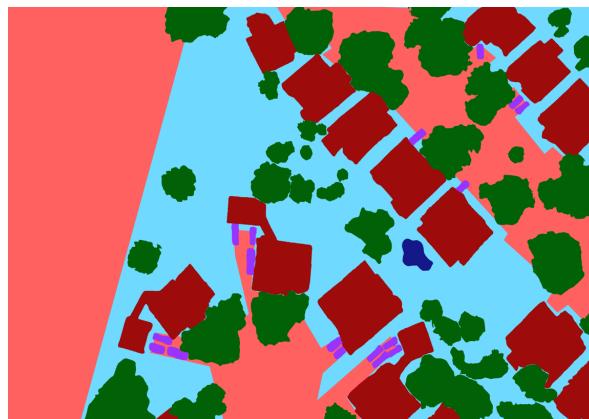
- Basato su tre principali parti:
 - **Data cleaning**
 - **Data augmentation offline**
 - **Architettura context-based**





Approccio proposto: Data cleaning

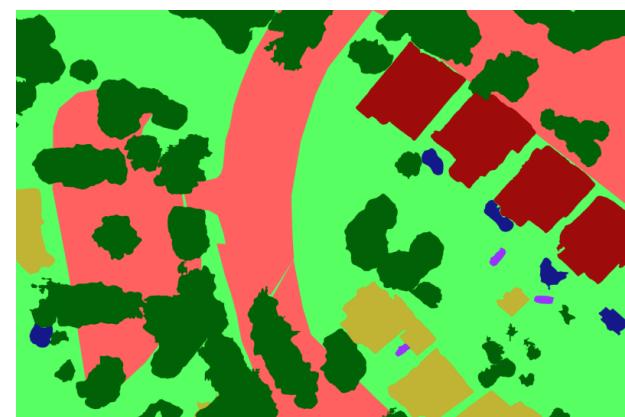
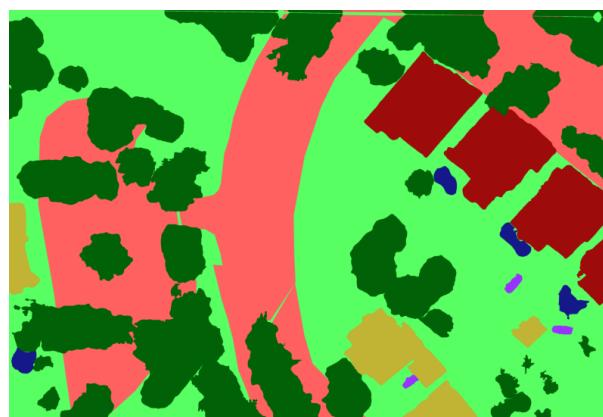
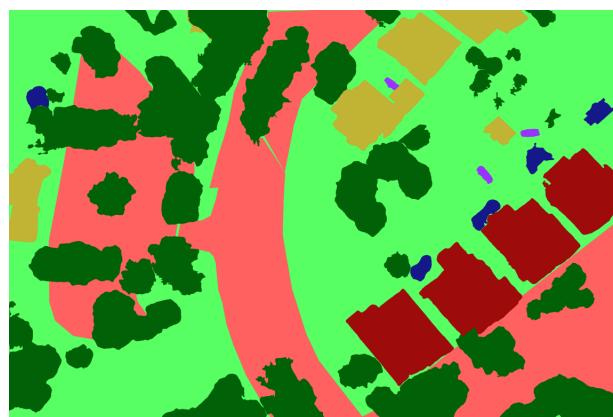
- Fase di **pulizia e correzione** del dataset
- Scansione **manuale** di tutte le immagini e le corrispettive maschere
- **182** maschere trovate con errori
- Tre **principali tipologie** di errori trovate:
 - classificazione erronea di pixel
 - mancata occorrenza di oggetti nelle maschere
 - presenza di incoerenza e confusione





Approccio proposto: Data augmentation offline

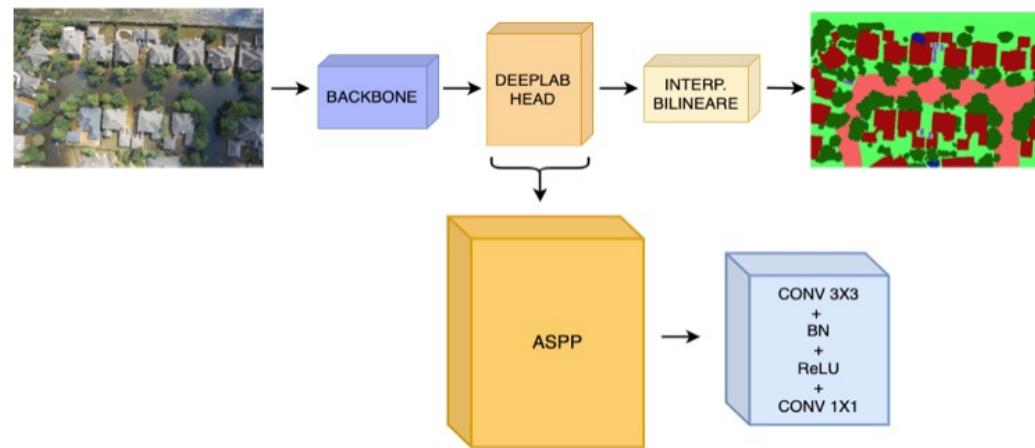
- **Scopo:** aumentare in modo **mirato** il numero di immagini per far fronte allo sbilanciamento delle classi
- 4 tipi di **trasformazione:** Rotazione, Flip orizzontale, Flip verticale e variazione della luminosità e del contrasto
- Da ognuna delle **140 immagini selezionate** sono state prodotte altre tre immagini con la corrispettiva maschera (**+420 immagini**)





Approccio proposto: Architettura

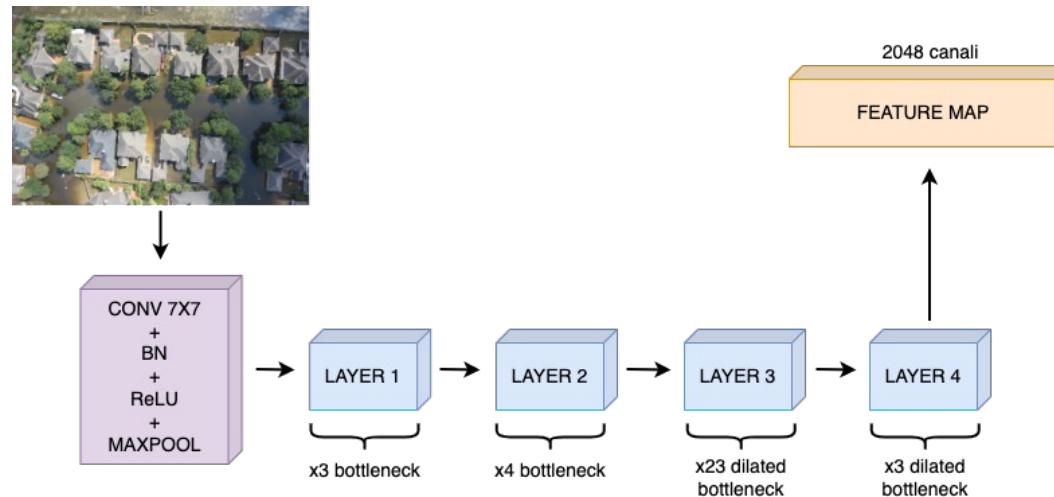
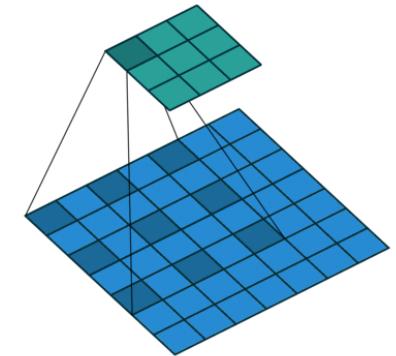
- Ispirata alla **DeepLabV3**, architettura proposta nel 2017
- Composta da 3 parti:
 - **Backbone**
 - **DeepLabHead (ASPP)**
 - **Interpolazione bilineare**





Backbone

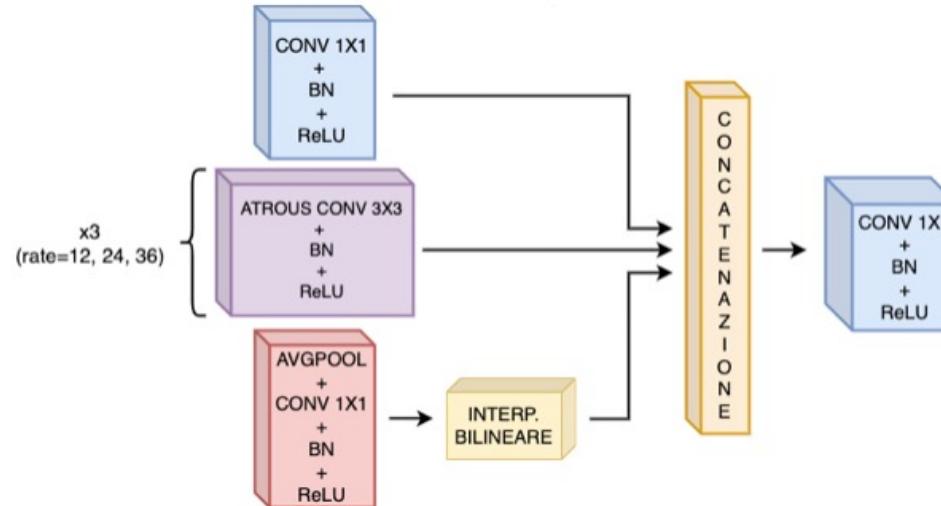
- Responsabile della **feature extraction**
- Rete convoluzionale composta da circa **101 strati** (ispirata alla **ResNet101**)
- Utilizza le **convoluzioni dilatate**





DeepLabHead e l'ASPP

- Composta principalmente dall'**ASPP (Atrous Spatial Pyramid Pooling)**
- Ispirato al SPP (Spatial Pyramid Pooling)
- L'idea principale è catturare **contesti a diverse scale** utilizzando diverse convoluzioni **in parallelo**





Risorse hardware



Limitazioni hardware

- Le risorse hardware hanno rappresentato un grosso **ostacolo** per il lavoro
- E' stata utilizzata la piattaforma Google Colab
- Disponibilità molto **limitata**: 3/5 ore giornaliere





Esperimenti e Risultati



Esperimenti

- **Iperparametri** in comune tra tutti gli esperimenti:
 - Adam
 - Batch size = 2
 - Learning Rate = 0.01 (tranne per l'ultimo esperimento)
 - Dataset split: 60% per il training, 20% per la validation e 20% per il testing
- **1[^] exp:** baseline (600*800)

Classe	mIoU								
1	2	3	4	5	6	7	8	9	
0.0018	0.087	0.0002	0.28	0.14	0.339	0.006	0.0	0.309	0.129

- **2[^] exp:** + Data augmentation online (600*800)

Classe	mIoU								
1	2	3	4	5	6	7	8	9	
0.0003	0.12	0.001	0.27	0.23	0.38	0.0	0.03	0.52	0.175



Esperimenti

- **3[^] exp:** + Data cleaning (750*1000)

Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe	mIoU
1	2	3	4	5	6	7	8	9	
0.14	0.47	0.06	0.48	0.46	0.55	0.34	0.26	0.81	0.402

- **4[^] exp:** + Data augmentation offline (750*1000)

Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe	mIoU
1	2	3	4	5	6	7	8	9	
0.32	0.51	0.24	0.55	0.55	0.6	0.4	0.44	0.84	0.5

- **5[^] exp:** + Learning rate dinamico

Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe	Classe	mIoU
1	2	3	4	5	6	7	8	9	
0.41	0.60	0.32	0.6	0.57	0.65	0.49	0.52	0.86	0.564



Risultati e confronto con altri lavori dello Stato dell'Arte

- **Confronto:**
 - la loro versione del dataset ha **857** immagini in più (**+36%**)
 - Disponibilità continua di risorse computazionali

Modello	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6	Classe 7	Classe 8	Classe 9	mIoU
ENet	0.069	0.473	0.124	0.484	0.489	0.683	0.322	0.424	0.762	0.426
DeepLabV3+	0.327	0.728	0.52	0.7	0.75	0.77	0.42	0.47	0.84	0.61
Approccio proposto	0.41	0.60	0.32	0.6	0.57	0.65	0.49	0.52	0.86	0.564

Facoltà di Ingegneria dell'Informazione, Informatica e Statistica
Corso di Laurea in Informatica (L-31)

Anno Accademico 2021-2022

Grazie per la Vostra Attenzione!



SAPIENZA
UNIVERSITÀ DI ROMA

Relazione di tirocinio

Studente: William De Vena
e-mail: william98wdv@gmail.com

Responsabile: Prof. Danilo Avola
e-mail: avola@di.uniroma1.it

Corresponsabili: Prof. Luigi Cinque
e-mail: cinque@di.uniroma1.it

Dott. Alessio Fagioli
e-mail: fagioli@di.uniroma1.it