



UNIVERSITÀ DEGLI STUDI DI
MILANO-BICOCCA



TETRIS PUZZLE

Image Processing Project

Sofia Amarù, s.amaru@campus.unimib.it
Sofia Damaso, s.damaso@campus.unimib.it
Gioele De Pianto, g.depianto@campus.unimib.it
Aniello Di Polito, a.dipolito@campus.unimib.it

INDICE

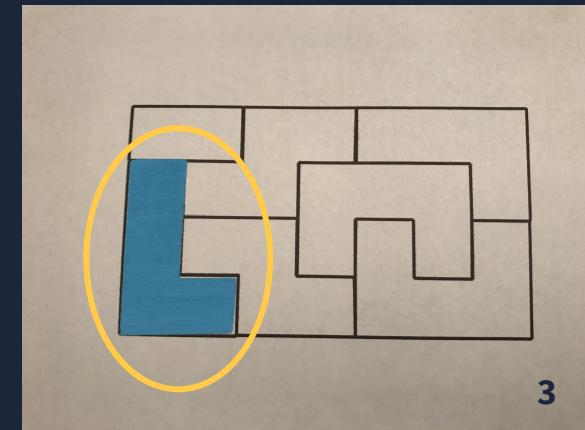
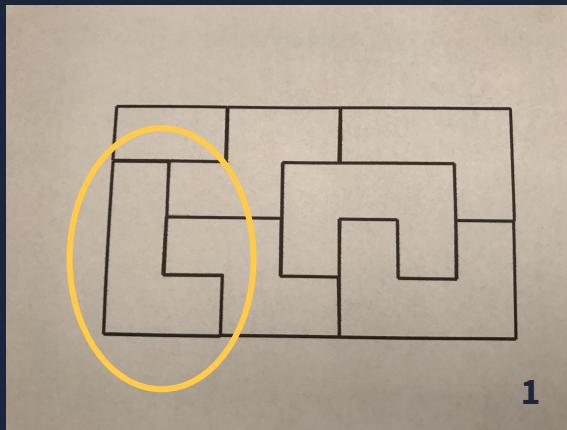
| | |
|---|----|
| Introduzione | 3 |
| Descrizione e analisi dei dati | 5 |
| Assunzioni | 9 |
| Pipeline | 10 |
| Segmentazione | 11 |
| Segmentazione della scena | |
| Segmentazione dello schema | |
| Signature | 15 |
| Calcolo | |
| Training | |
| Classificazione | |
| Matching | |
| Posizionamento | 19 |
| Extras | 22 |
| Analisi dei risultati | 24 |
| Possibili miglioramenti | 26 |
| Contributi | 27 |

INTRODUZIONE

3

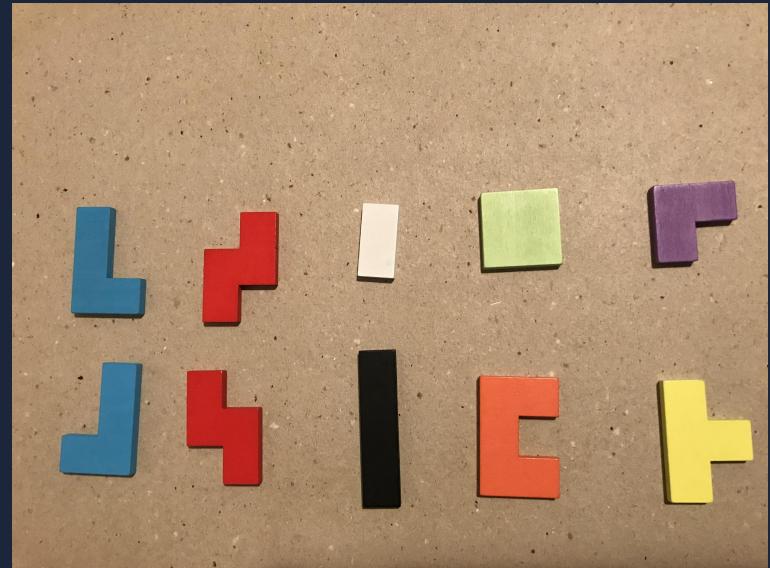
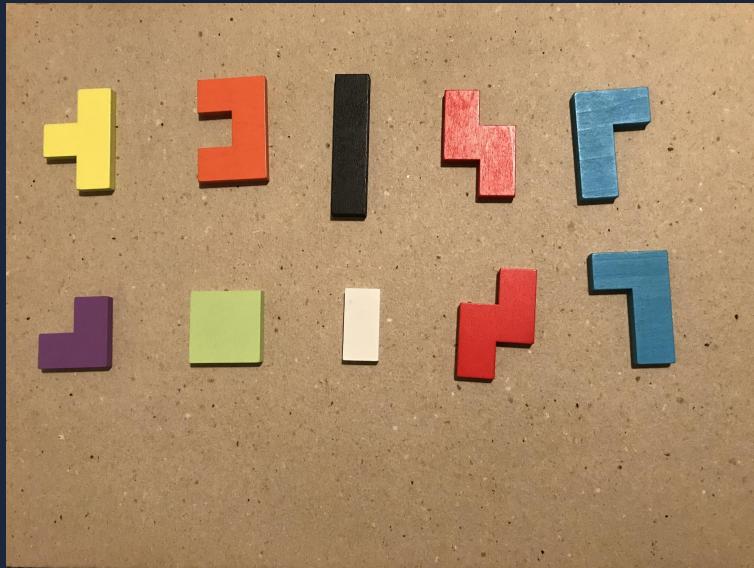
Nel gioco Puzzle Tetris vi sono diverse figure geometriche chiamate tetramini che possono essere combinate fra loro. L'obiettivo del gioco è comporre uno schema senza buchi.

Il nostro scopo è stato sviluppare un'applicazione che, prese in input due immagini - una di schema (1) e una di scena (2) - posizioni sull'immagine dello schema i tetramini corrispondenti (3).



DESCRIZIONE E ANALISI DEI DATI

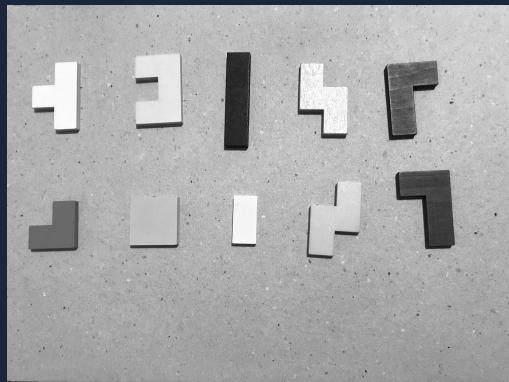
Il **training set** è composto dalle due seguenti immagini di scena:



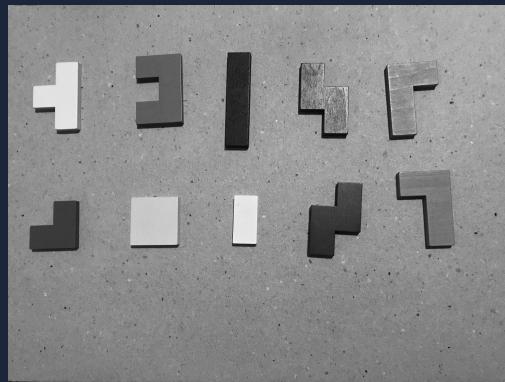
Di seguito, le immagine di training nei canali RGB, YCbCr e HSV.

DESCRIZIONE E ANALISI DEI DATI

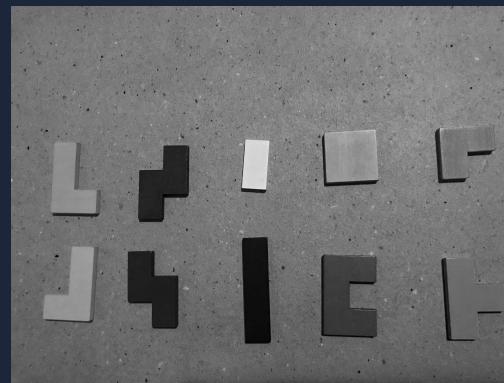
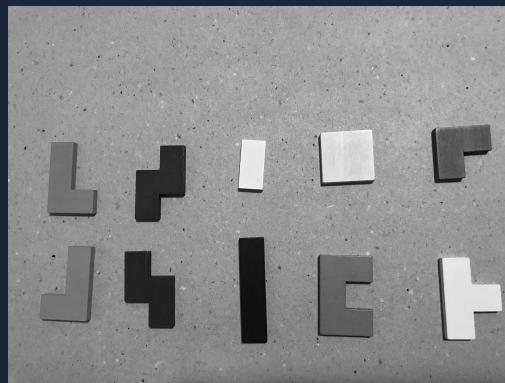
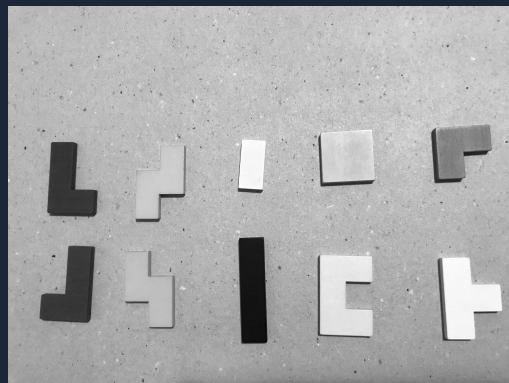
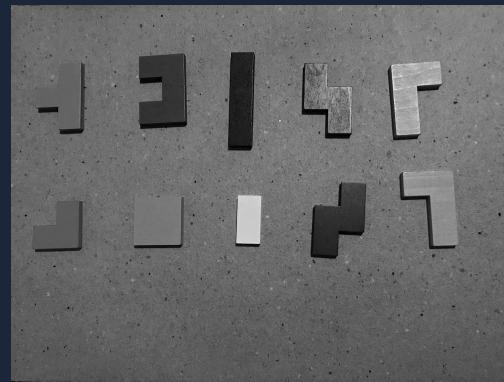
R



G

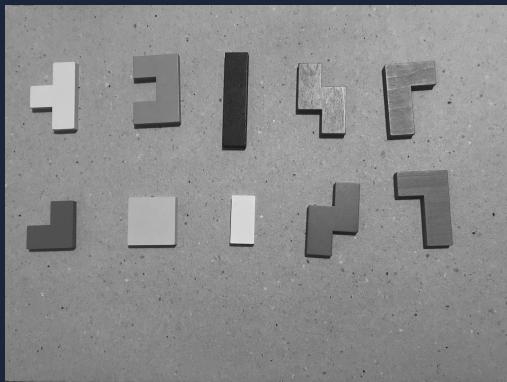


B



DESCRIZIONE E ANALISI DEI DATI

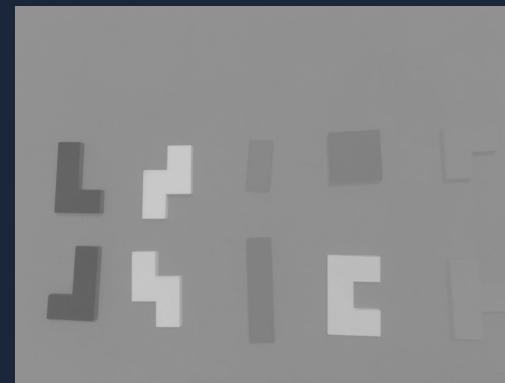
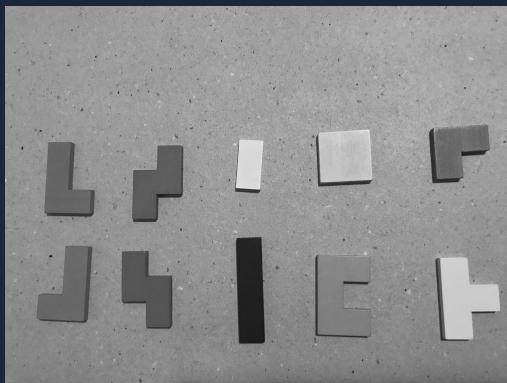
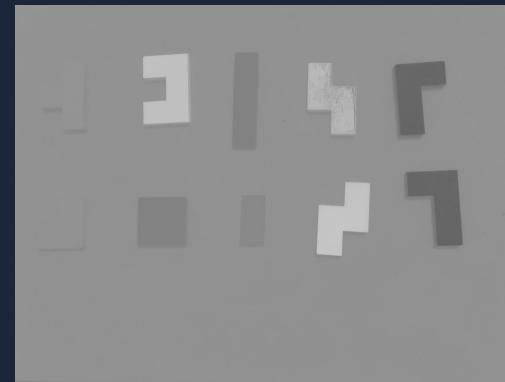
Y



Cb

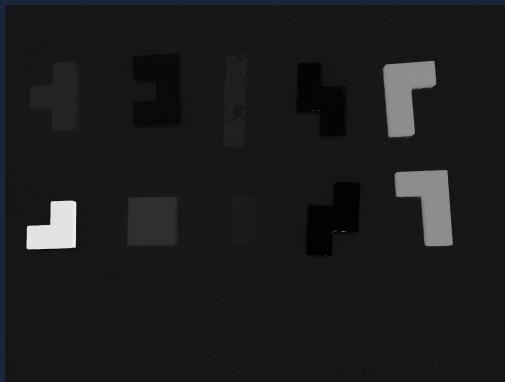


Cr

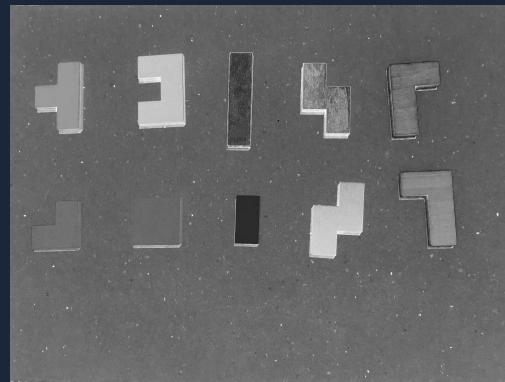


DESCRIZIONE E ANALISI DEI DATI

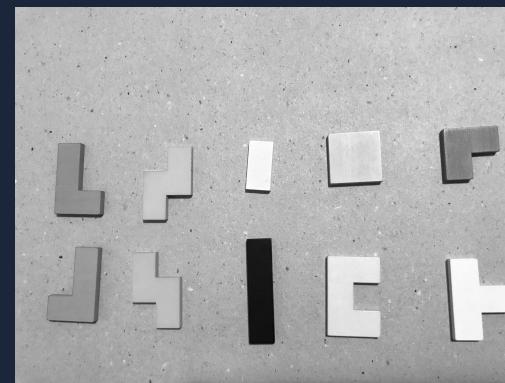
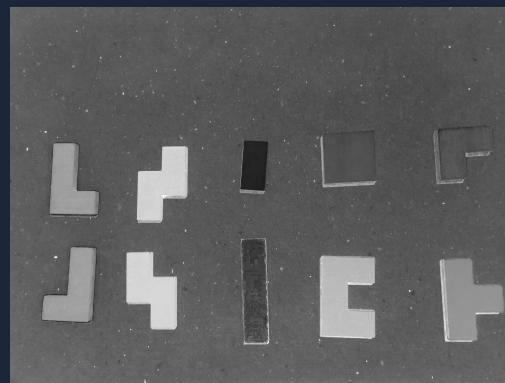
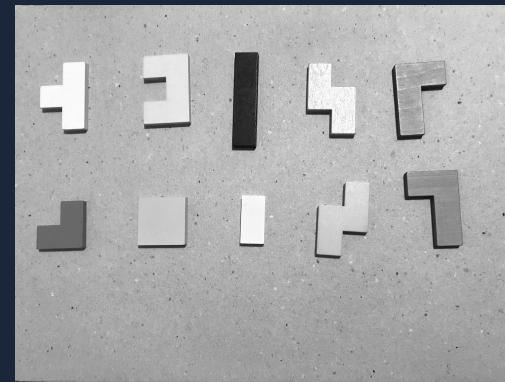
H



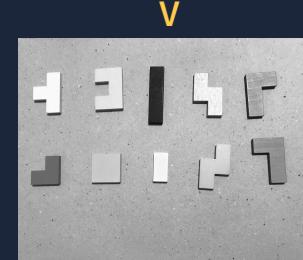
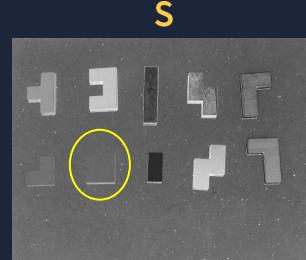
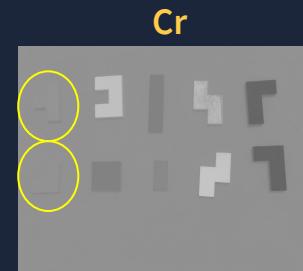
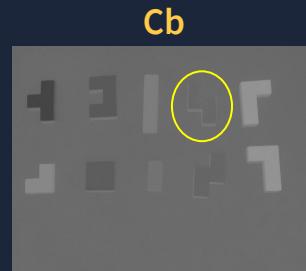
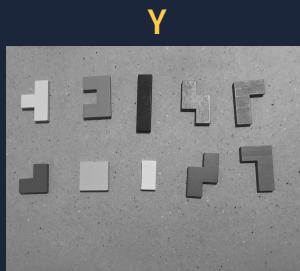
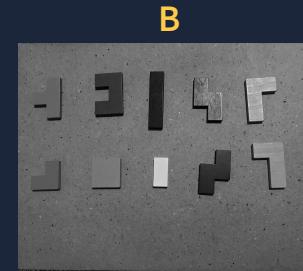
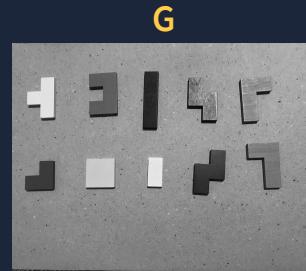
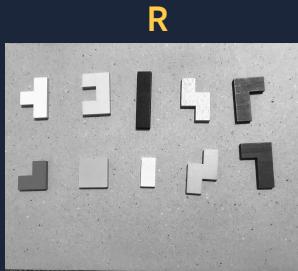
S



V



DESCRIZIONE E ANALISI DEI DATI

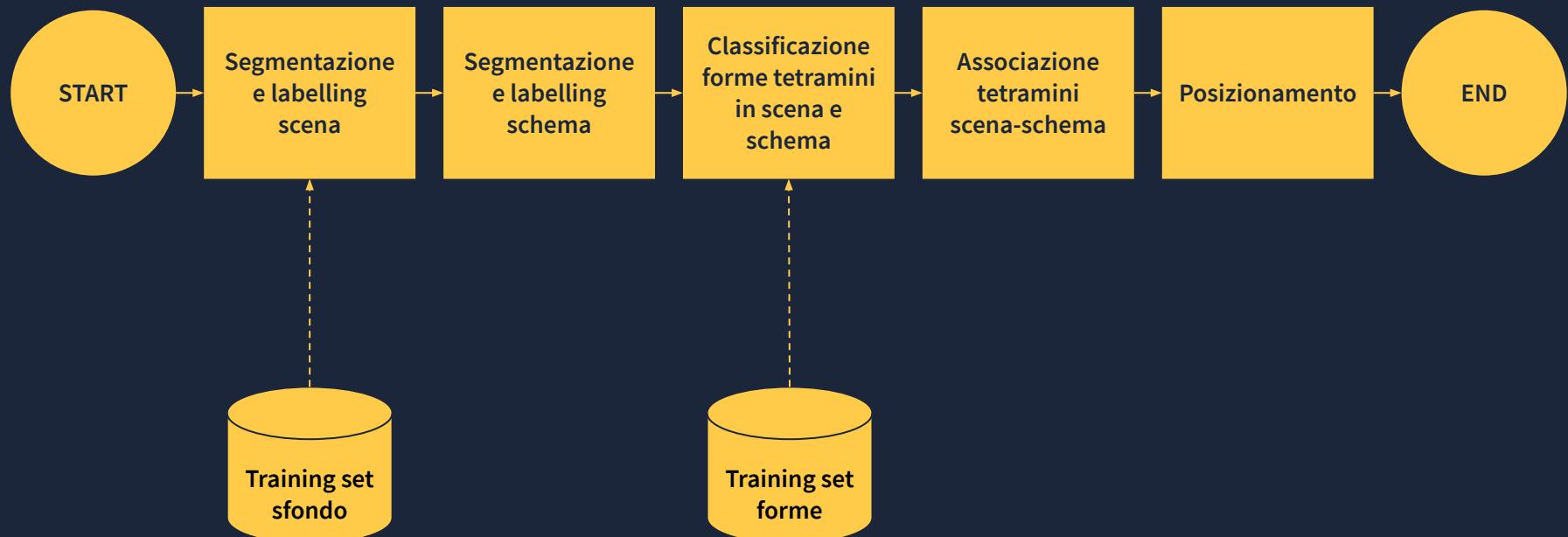


Per un corretto funzionamento dell'applicazione, assumiamo che:

- le immagini di input abbiano la stessa dimensione
- lo sfondo delle immagini di scena sia lo stesso delle immagini di training e i tetramini non assumano lo stesso colore di sfondo
- nessun tetramino esca fuori dall'immagine di scena o sia tagliato
- nessuno schema esca fuori dall'immagine o sia tagliato
- i tetramini delle immagini di scena non si sovrappongano
- le condizioni di luminosità e rumore della scena siano simili a quelle delle immagini di training
- l'illuminazione dello schema sia pressochè uniforme, senza zone particolarmente sottoesposte/sovraesposte, e non sia presente un forte rumore
- gli schemi abbiano bordi neri ben visibili e non interrotti

L'APPLICAZIONE - PIPELINE

10



SEGMENTAZIONE DELLA SCENA

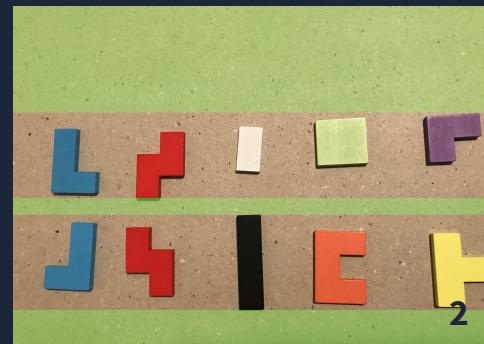
11

La segmentazione avviene tramite **classificazione pixel-based**.

A partire dalle immagini di training (1,2) abbiamo creato una nuova immagine che rappresenta lo sfondo (3):



1



2



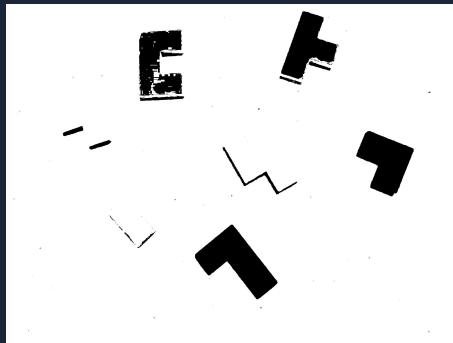
3

Questa immagine viene usata per addestrare il classificatore affinché riconosca lo sfondo nelle immagini di scena. Otteniamo così la maschera binaria dello sfondo; il suo negativo è la maschera dei tetramini. Per migliorarne la qualità, adoperiamo la **morfologia matematica** e usiamo **sauvola** per eliminare le ombre.

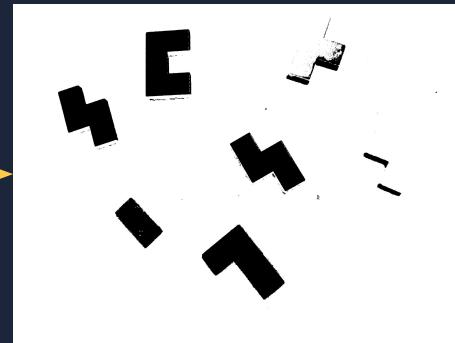
COME FUNZIONA IL CLASSIFICATORE Possiamo descrivere la distribuzione dei canali dei pixel di sfondo con due gaussiane, una per il canale Cb e una per il canale Cr. Dato un generico pixel nei canali YCbCr, sappiamo che fitta la distribuzione dei pixel di sfondo se e solo se è compreso contemporaneamente nelle due gaussiane. Il valore **k = 6.5** è stato scelto sulla base di valutazioni euristiche.

SEGMENTAZIONE DELLA SCENA

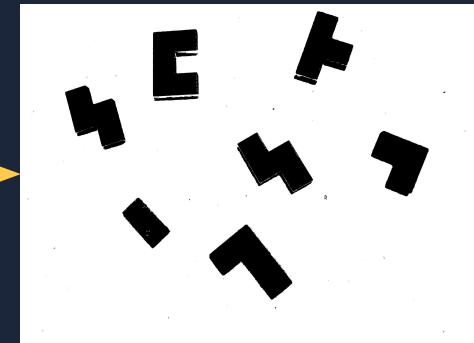
maschera del canale Cb



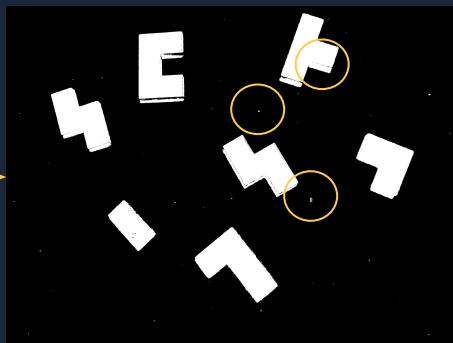
maschera del canale Cr



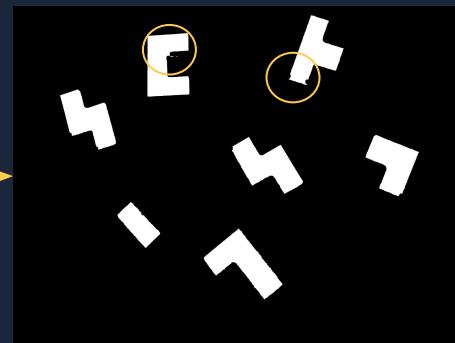
maschera Cb & maschera Cr



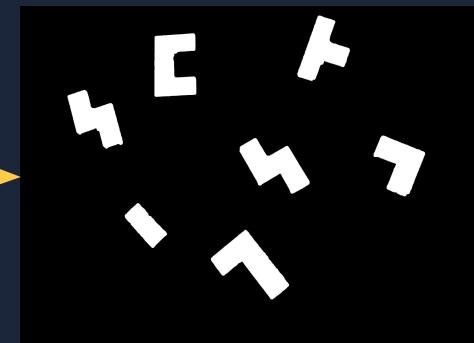
\neg (maschera Cb & maschera Cr)



dopo morfologia matematica



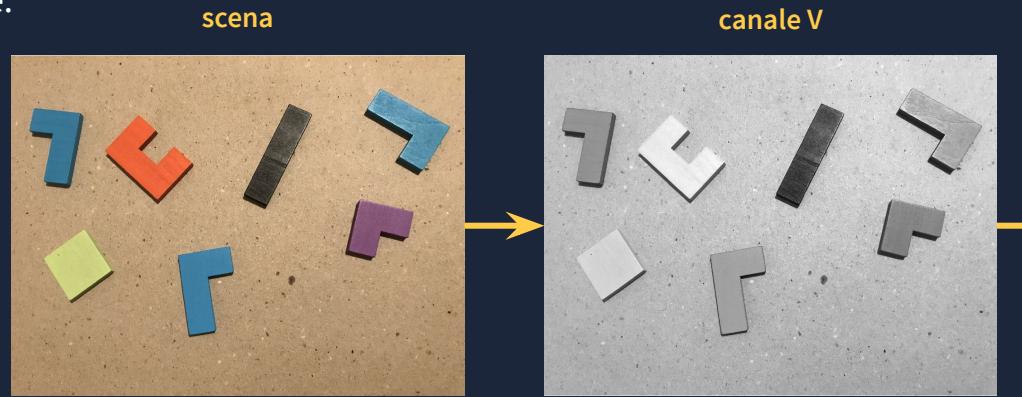
dopo aver eliminato le ombre



SEGMENTAZIONE DELLA SCENA

13

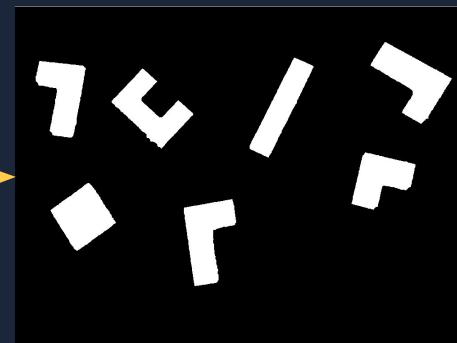
Eliminazione delle ombre:



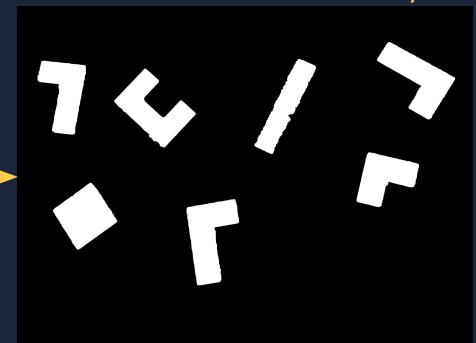
dopo sauvala (default $k = 0.34$)



\neg (maschera Cb & maschera Cr)



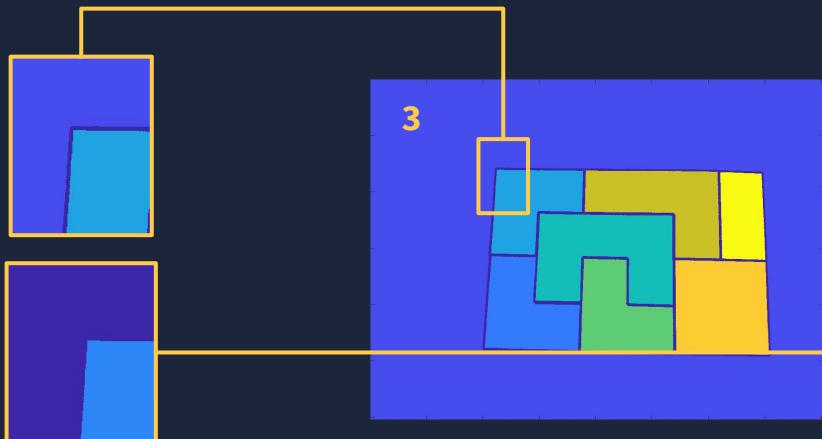
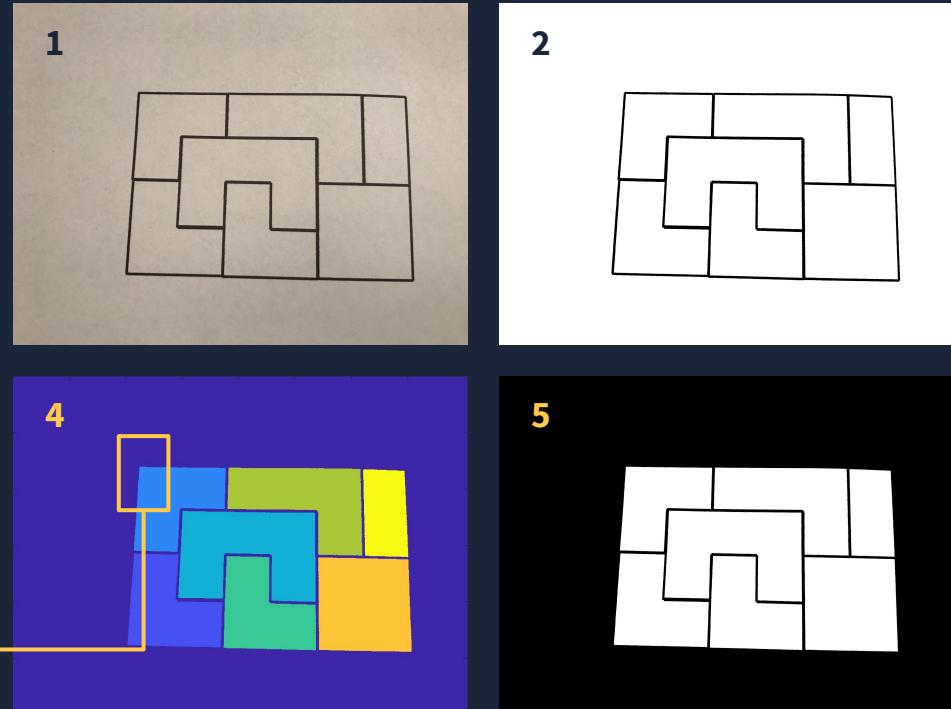
maschera finale (maschera ombre & maschera classificazione)



SEGMENTAZIONE DELLO SCHEMA

Per ottenere la maschera dello schema (1), viene eseguita una **segmentazione a soglia fissa sul canale V** dell'immagine. La soglia ($t=0.3$) è stata scelta analizzando l'istogramma dell'immagine. Viene poi effettuata una closing sul nero per rimuovere eventuali piccoli buchi. Otteniamo l'immagine 2.

Vengono etichettate le componenti connesse (3). Assumiamo che lo sfondo sia l'oggetto con il maggior perimetro. Eliminando la label dello sfondo otteniamo la maschera contenente i soli oggetti interessati (4,5).

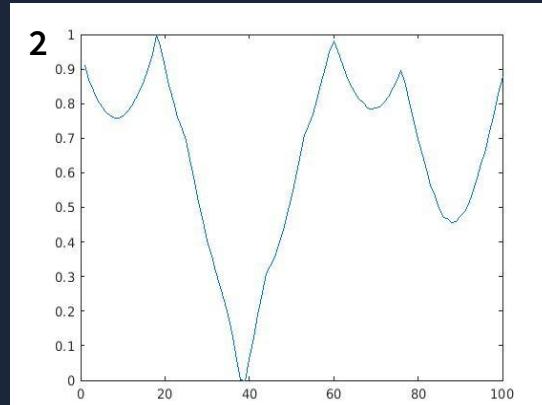
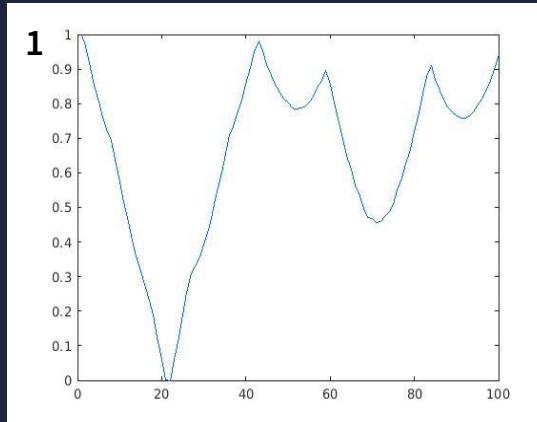


CALCOLO DELLE SIGNATURE

15

Le **signature** di un pezzo (sia scena che schema) sono calcolate attraverso i seguenti passaggi:

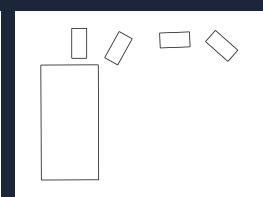
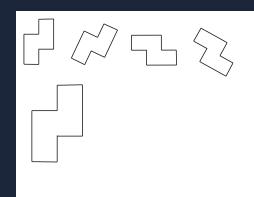
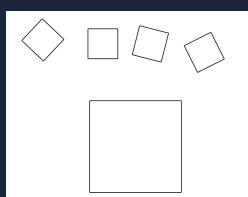
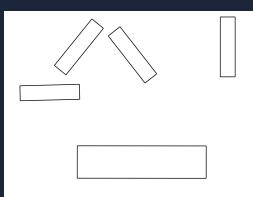
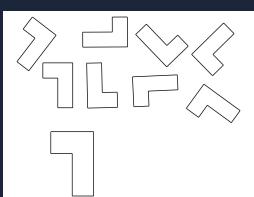
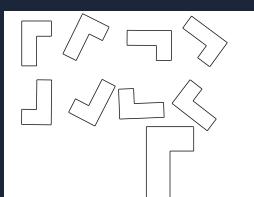
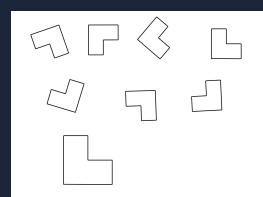
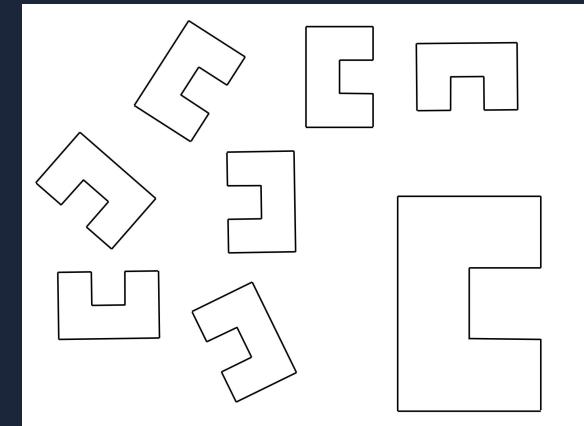
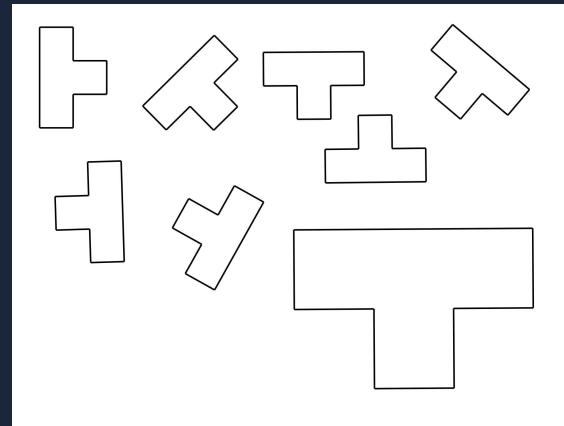
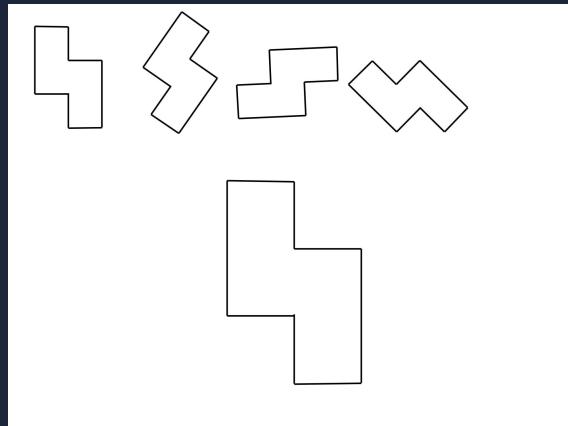
1. calcolo della distanza tra ogni punto del contorno ed il centroide* del pezzo;
2. viene sottocampionato il vettore delle signature a 100 punti utilizzando l'interpolazione lineare;
3. si calcolano tutti i massimi locali che hanno valore maggiore del 70% del massimo globale e distanti fra loro di almeno 5 unità (1);
4. per ogni massimo trovato si crea una nuova signature (2) traslando quella originale, ponendo il massimo come punto iniziale (invarianza alla rotazione);
5. si normalizzano le signature tra 0 e 1 (invarianza alla scala).



*centroide calcolato attraverso i momenti

TRAINING DELLE SIGNATURE

Per effettuare il training delle signature abbiamo creato delle **immagini di training**, una per ogni classe di oggetto:



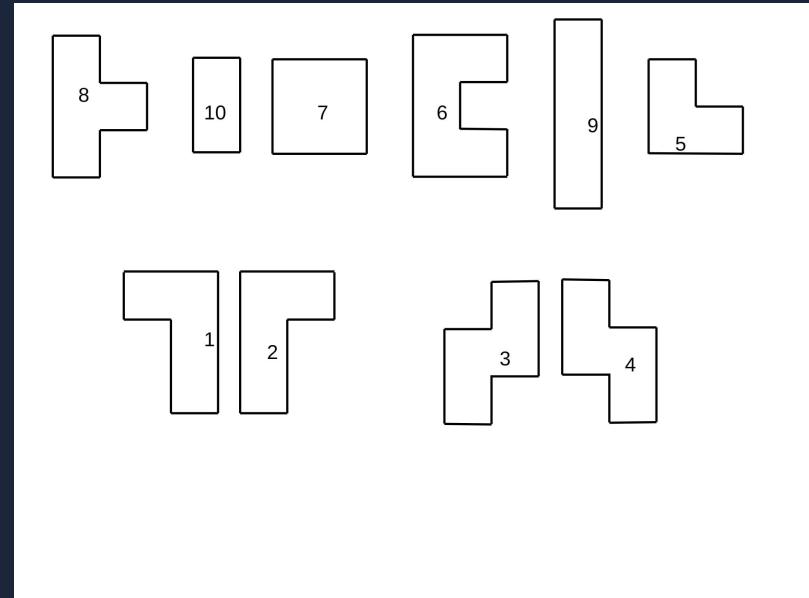
CLASSIFICAZIONE DELLE SIGNATURE

La classificazione dei pezzi (sia di scena che di schema) viene effettuata tramite un **classificatore**. Il classificatore compara le signature del singolo pezzo con quelle delle classi di training e associa il pezzo a quella con distanza minima. Ogni classe di training è composta da un insieme di signature.

La funzione distanza tra un pezzo ed un classe di training è così calcolata:

1. per ogni combinazione signature pezzo e signature classe si calcola la **distanza euclidea** tra di esse;
2. la distanza è il minimo valore trovato al punto 1

Abbiamo notato che mediamente la distanza tra un pezzo e la sua classe di appartenenza è a circa 0.5 tuttavia in alcuni casi si supera leggermente il valore 1. Abbiamo quindi deciso di introdurre una **classe di rigetto** (identificata da -1) per i casi in cui la distanza sia maggiore di 2.



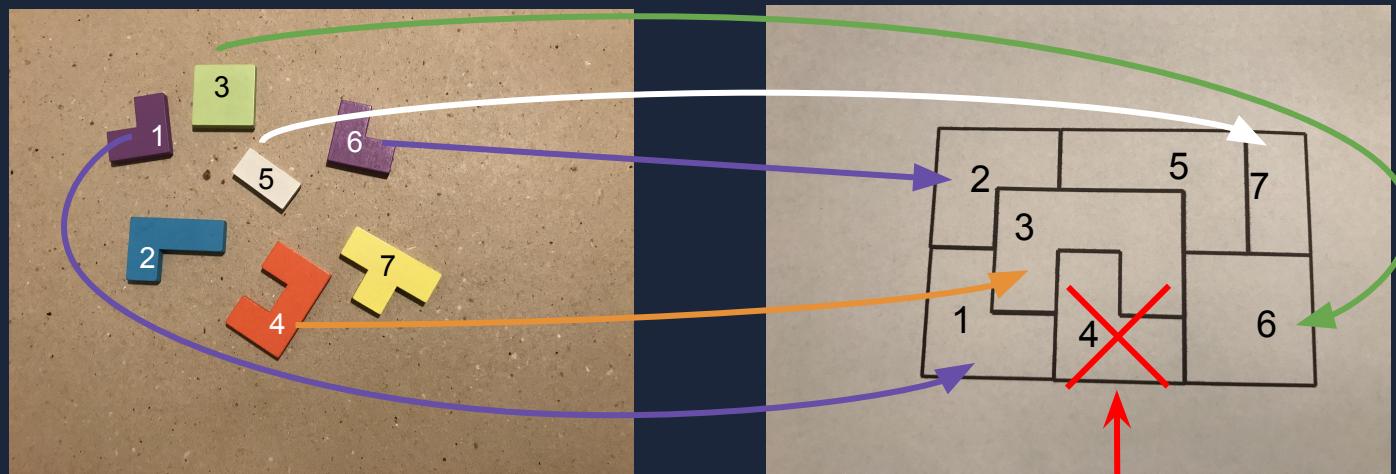
Classi associate ai tetramini

MATCHING DELLE SIGNATURE

Una volta effettuata la classificazione si procede al matching tra pezzi dello schema e pezzi della scena.

Verranno matchati i pezzi a cui è stata associata la stessa classe.

Un singolo pezzo della scena non può essere inserito più volte nello schema.



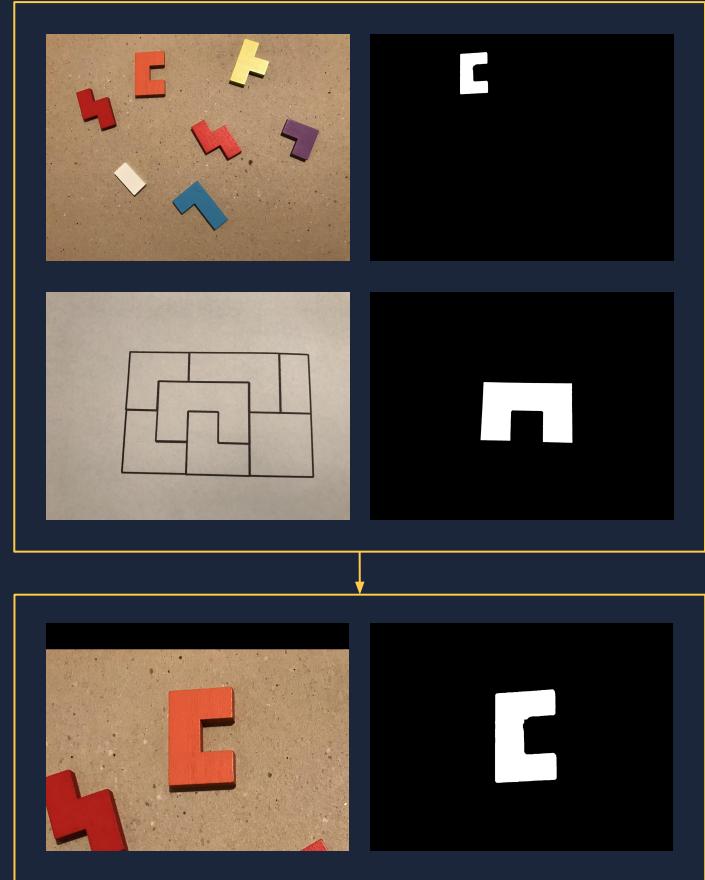
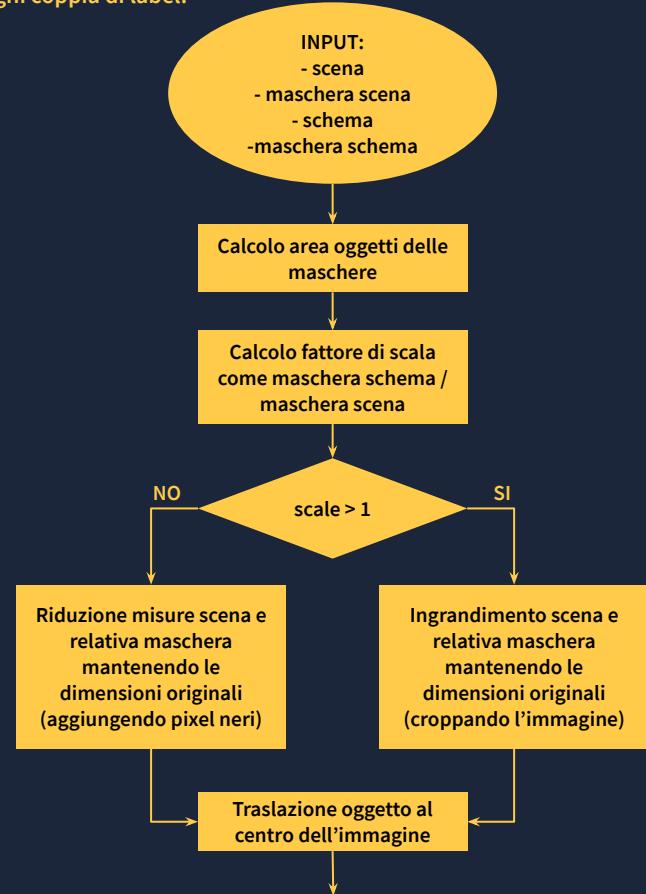
| ↓Pezzo schema / Pezzo scena→ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|------------------------------|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

Il pezzo 4 non ottiene match poiché i tetramini compatibili 1 e 6 vengono collocati sui pezzi 2 e 1 di schema.

Matrice usata per rappresentare i match

POSIZIONAMENTO

Per ogni coppia di label:

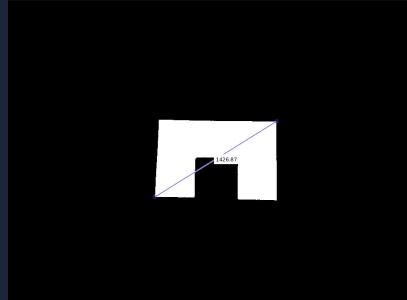


POSIZIONAMENTO

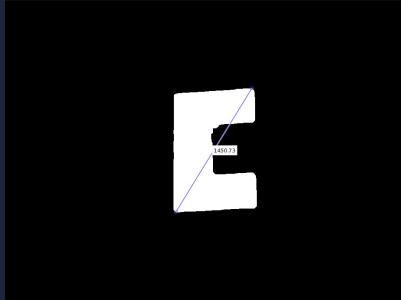
20



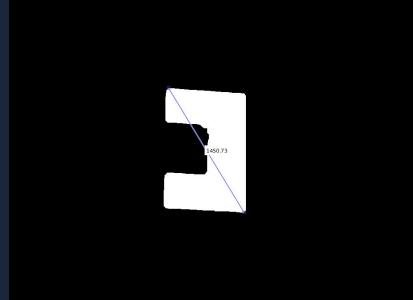
Schema



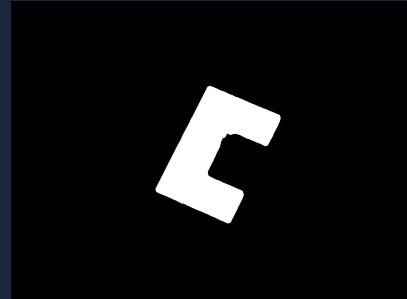
Scena



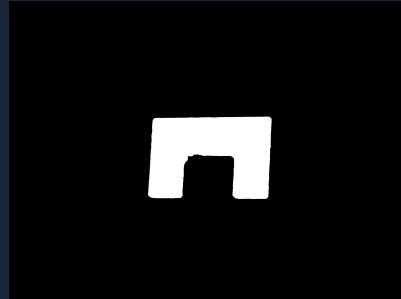
Scena flippata



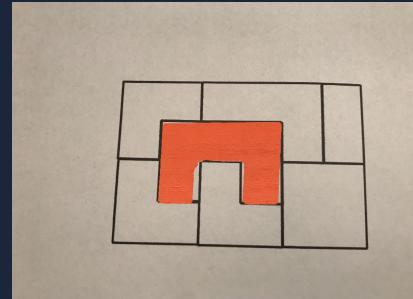
Scena ruotata
(prima rotazione)



Scena flippata ruotata
(prima rotazione)



Risultato finale



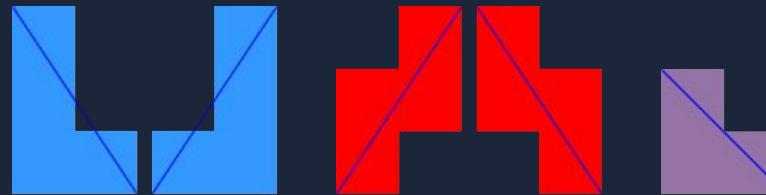
*per la scena, l'angolo di rotazione è -(angolo schema - angolo scena);

**per la scena flippata, l'angolo di rotazione è -(angolo schema - angolo scena flippata);

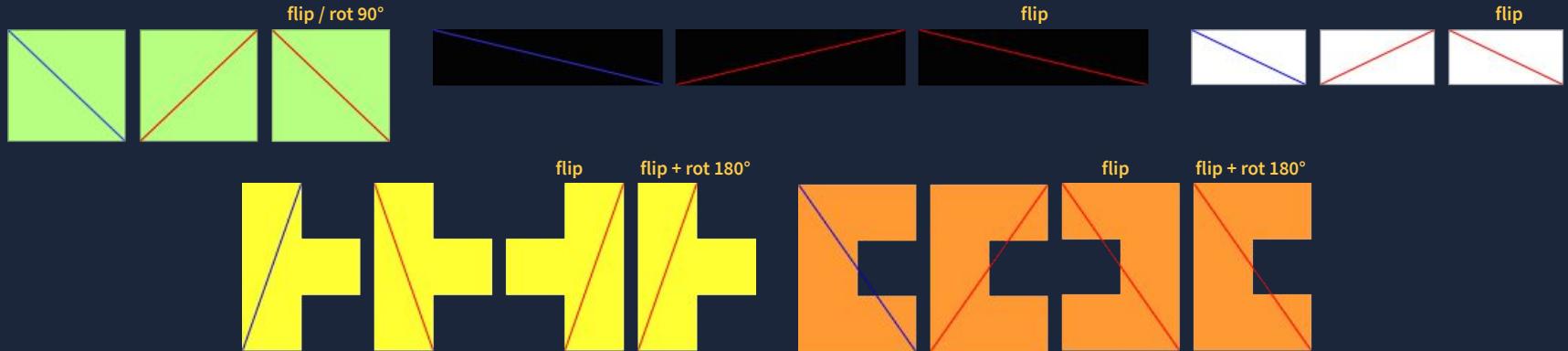
**controlliamo quale tetramino fitta meglio tra quelli di: 1. scena; 2. scena flippata; 3. scena ruotata di 180°; 4. scena flippata ruotata di 180°

POSIZIONAMENTO

Perché flippiamo i tetramini? Alcune classi di tetramini hanno un solo possibile diametro:



Sugli altri tetramini, invece, il diametro può assumere due configurazioni diverse. Si può ovviare al problema
flippando (e/o ruotando) la figura:

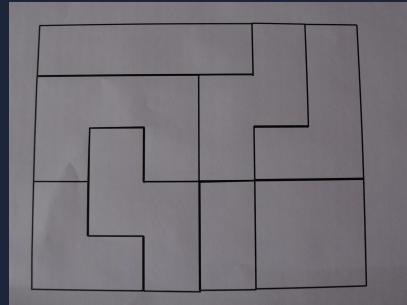


EXTRAS

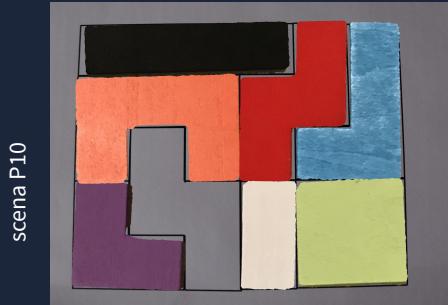
22

Abbiamo svolto dei test anche con immagini di schema e di scena con caratteristiche differenti rispetto agli schemi e alle scene fornite dai docenti e anche in un casi che violano le assunzioni poste:

SCHEMA INGRANDITO



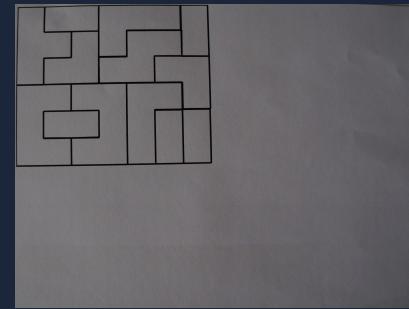
extras S07



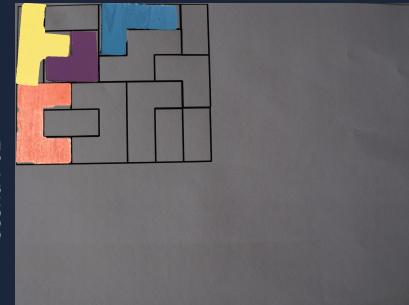
scena P10

✓ Comportamento previsto

SCHEMA NON CENTRATO



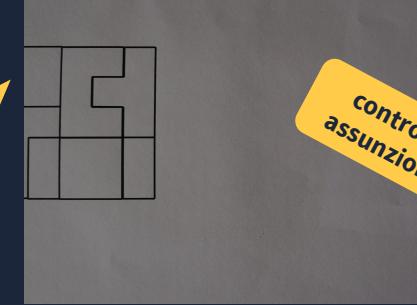
extras S08



scena P01

✓ Comportamento previsto

SCHEMA TAGLIATO



extras S09



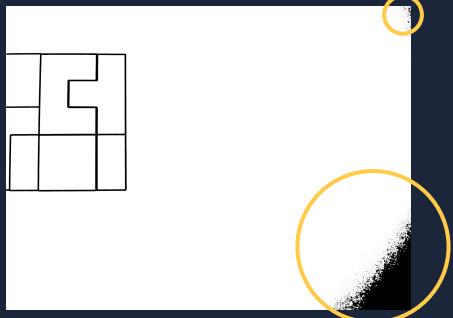
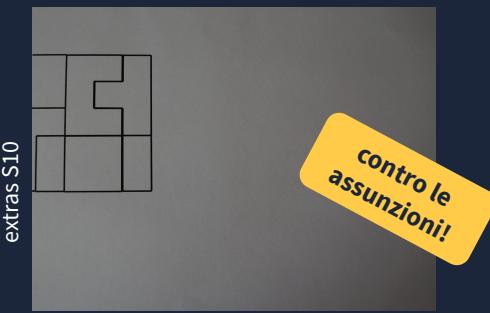
scena P10

contro le
assunzioni!

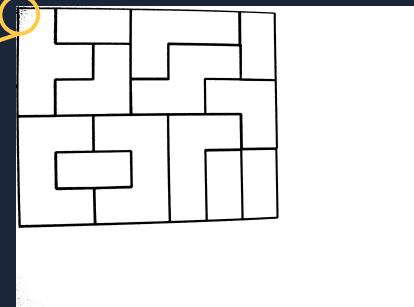
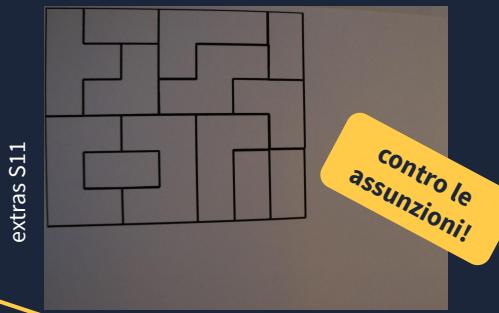
✗ Il quadrato tagliato viene
considerato un rettangolo corto

EXTRAS

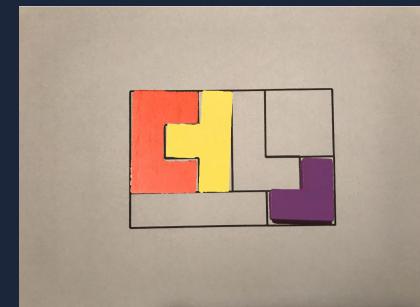
SCHEMA SOTTOESPOSTO



SCHEMA SOTTOESPOSTO



SCENA ZOOMATA



✓ Comportamento previsto

ANALISI DEI RISULTATI - MATRICI DI CONFUSIONE

24

Tramite l'analisi delle matrici di confusione possiamo affermare che il programma ottiene buoni risultati. Abbiamo riscontrato un minimo errore nella segmentazione delle scene e nella segmentazione dello schema, mentre la classificazione e il matching schema-scena hanno ottenuto un'accuracy del 100%.

SEGMENTAZIONE SCENE

| Immagine | Accuracy |
|--------------|---------------|
| P01 | 98,83% |
| P02 | 98,73% |
| P03 | 98,46% |
| P04 | 98,69% |
| P05 | 98,45% |
| P06 | 98,08% |
| P07 | 98,35% |
| P08 | 98,35% |
| P09 | 98,64% |
| P10 | 97,68% |
| Media | 98,43% |

CLASSIFICAZIONE SCENE

| Classe | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------------|----------|----------|----------|----------|-----------|-----------|----------|----------|----------|----------|
| 1 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 13 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 11 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9 |
| Accuracy: 100% | | | | | | | | | | |

ANALISI DEI RISULTATI - MATRICI DI CONFUSIONE

25

SEGMENTAZIONE SCHEMA

| Immagine | Accuracy |
|--------------|---------------|
| S01 | 99,69% |
| S02 | 99,68% |
| S03 | 99,58% |
| S04 | 99,26% |
| S05 | 99,57% |
| S06 | 99,52% |
| Media | 99,55% |

CLASSIFICAZIONE SCHEMI

| Classe | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|--------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| 1 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 7 | 0 | 0 | 0 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |

Accuracy: 100%

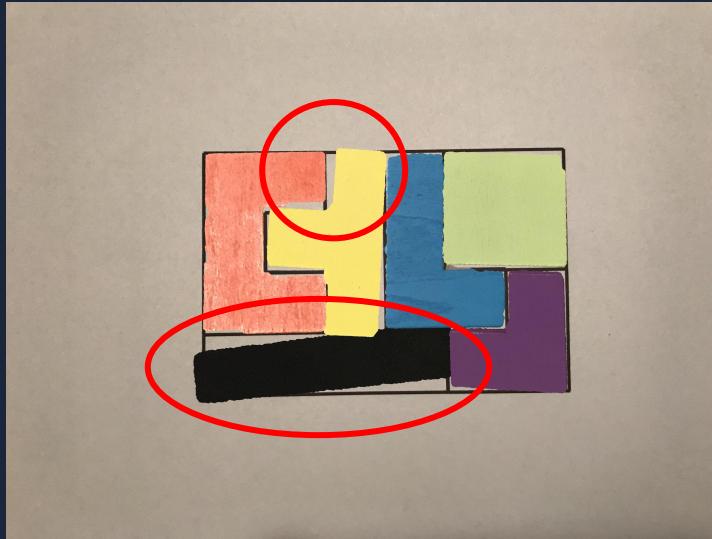
MATCH SCENA-SCHEMA

Accuracy match tra pezzi dello schema e pezzi della scena su un campione di test casuale su 20 combinazioni: **100%**

POSSIBILI MIGLIORAMENTI

26

- 1) Gestione ipotetico overfitting dei dati
- 2) Alcuni tetramini non combaciano perfettamente



Sofia Amarù ~ 25%

- Segmentazione schema soglia fissa
- Placement dei tetramini
- Segmentazione scena tramite classificazione
- Creazione groundtruth segmentazioni scene e schema
- Test extras

Gioele De Pianto ~ 25%

- Labelling dello schema
- Classificazione tetramini
- Calcolo signature
- Creazione groundtruth classificatore
- Analisi risultati classificatore
- Segmentazione distanza euclidea (CbCr,HCrCb)*

Sofia Damaso ~ 25%

- Gestione matching schema-scena
- Creazione matrici di confusione
- Creazione groundtruth matching
- Analisi risultati segmentazione di schemi e scene
- Segmentazione tramite colore*

Aniello Di Polito ~ 25%

- Analisi dataset
- Test extras
- Segmentazione tramite superpixel*
- Segmentazione con RGB*
- Associazione tetramini schema-scena senza classificatore*