

Politecnico di Milano

Corso di laurea in
Computer Science and Engineering



Riconoscimento di Anomalie con Ganomaly: patch-wise analysis e transfer learning

Autore:

Daniele Moltisanti

Relatore:

Prof. Giacomo Boracchi

Correlatore:

Luca Frittoli



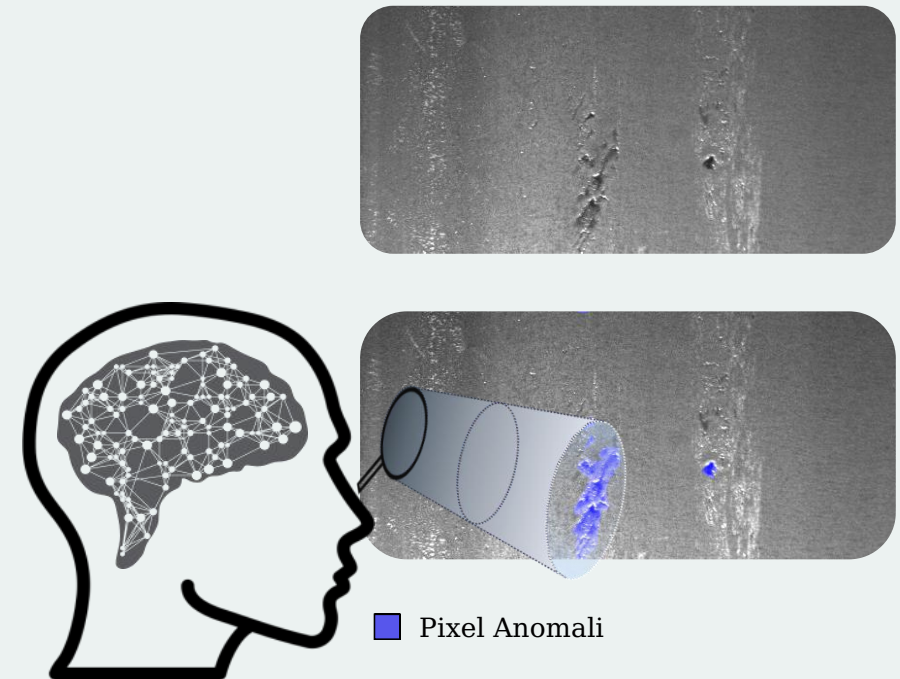
1

Riconoscimento di Anomalie

- Descrizione del problema
- Soluzioni soluzioni al problema

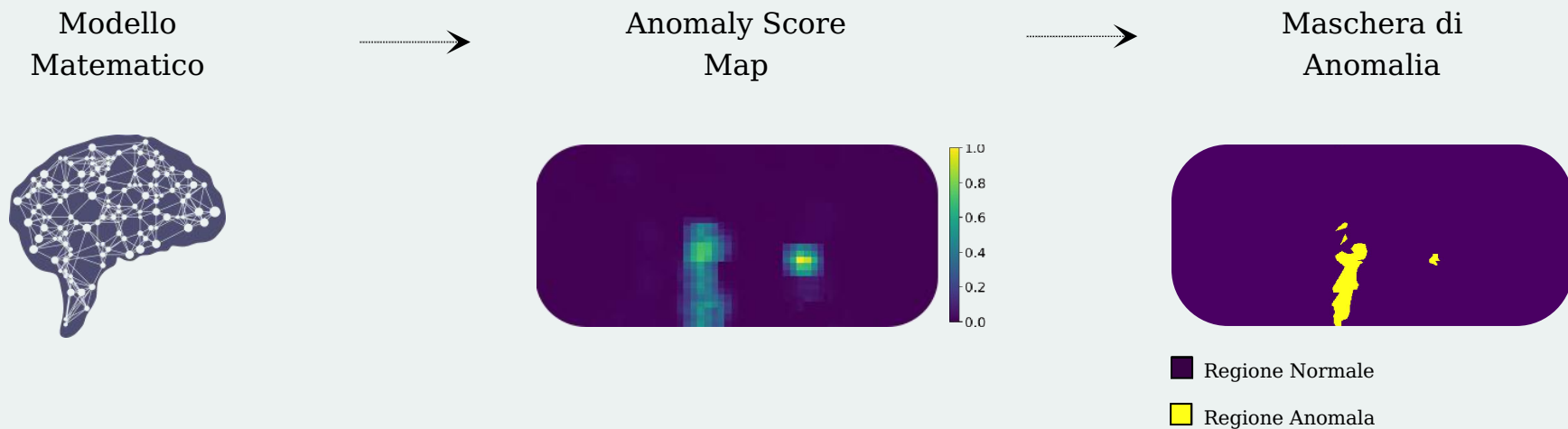
Descrizione del problema

- **Dominio:** immagini relative a lastre di acciaio
- **Obiettivo:** identificare i **difetti (anomalie)** presenti su tale dominio
- **Detector:** strumento che rileva anomalie all'interno di immagini



Obiettivo

Si vuole stimare un **modello data-driven** che misuri il grado di anomalia di ogni pixel



- > L'**anomaly score map** associa ad ogni pixel un valore di anomalia (anomaly score)
- > Pixel con un **alto valore di anomalia** vengono riconosciuti anomali

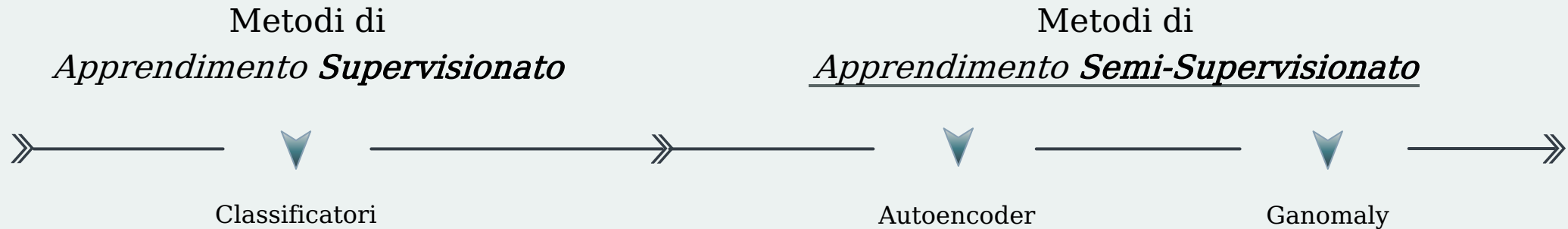


1

Riconoscimento di Anomalie

- Descrizione del problema
- Soluzioni esistenti al problema

Soluzioni esistenti al problema

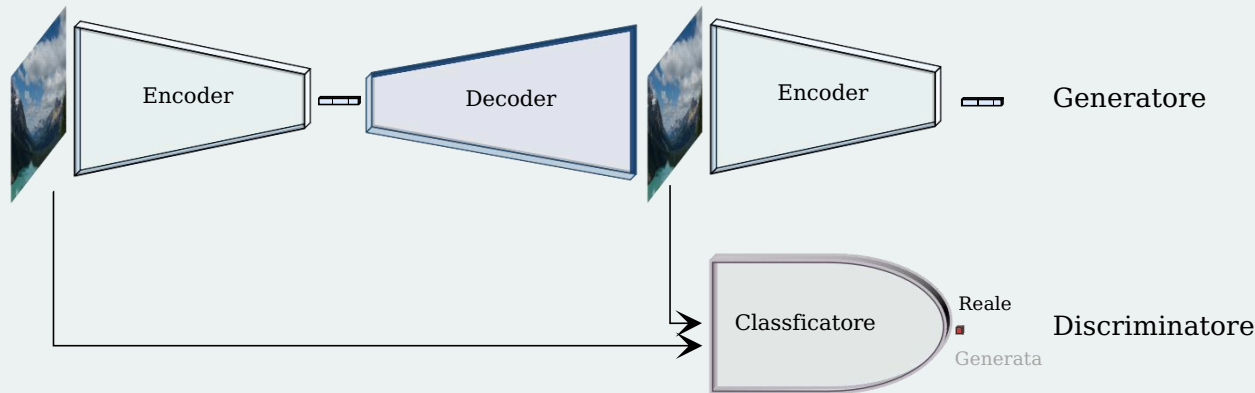
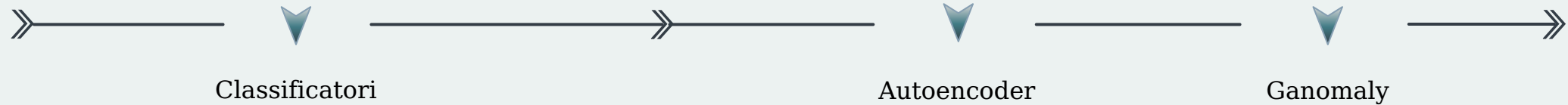


- > I modelli **semi-supervisionati** apprendono i pattern delle **sole** immagini **normali**
- > Calcolano l'**anomaly score** per un'immagine
- > Sfruttano il valore di anomaly score per definire la **presenza di anomalie**
- > Un **alto** valore anomaly score identifica un'**immagine anomala**

Soluzioni esistenti al problema

Metodi di
Apprendimento Supervisionato

Metodi di
Apprendimento Semi-Supervisionato



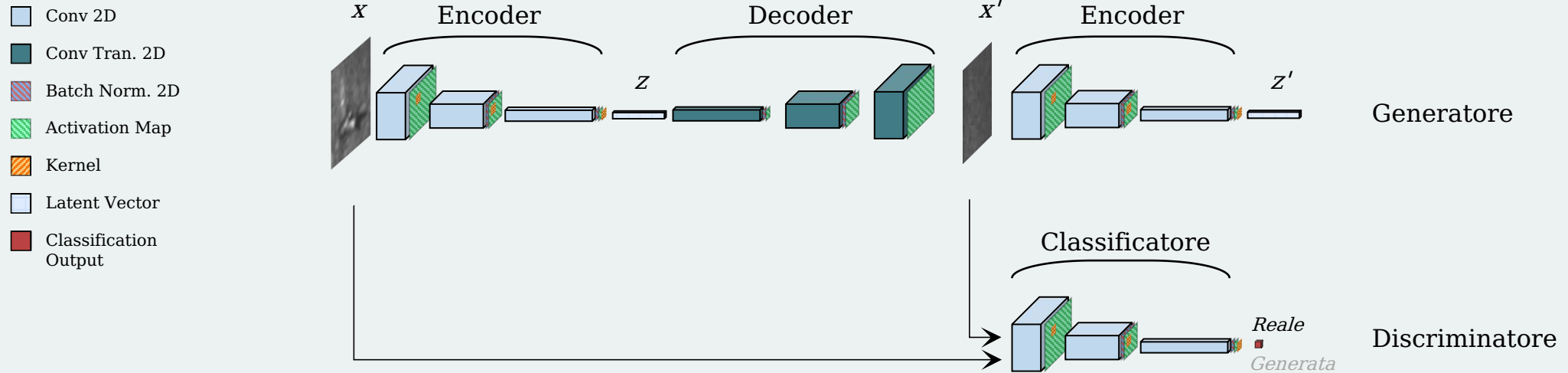
- > Applicato su immagini di dimensioni **ridotte** (32 x 32)
- > Riconosce se un' **intera immagine** sia normale o anomala



Soluzioni esistenti al problema:
Ganomaly

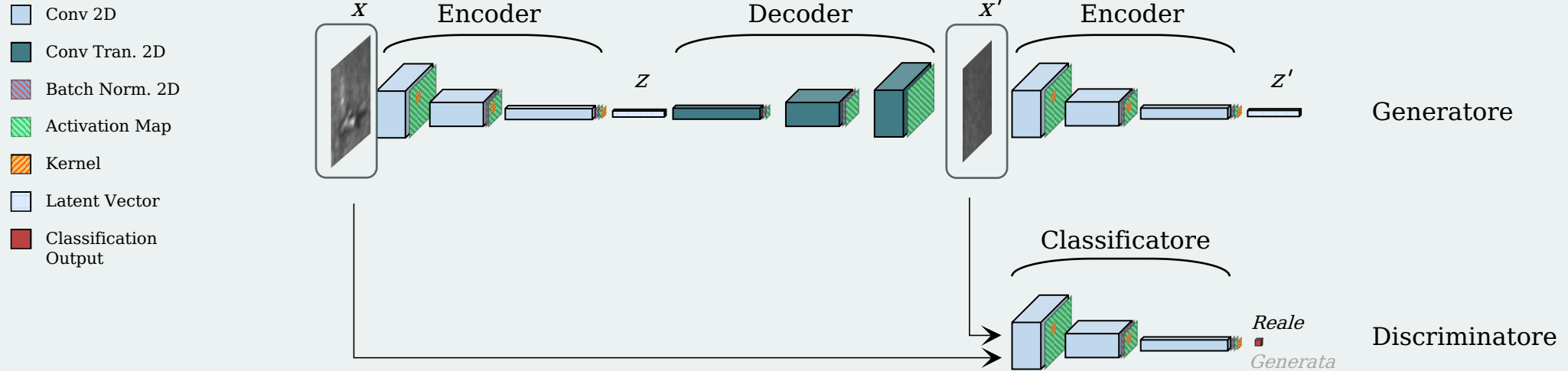


Ganomaly



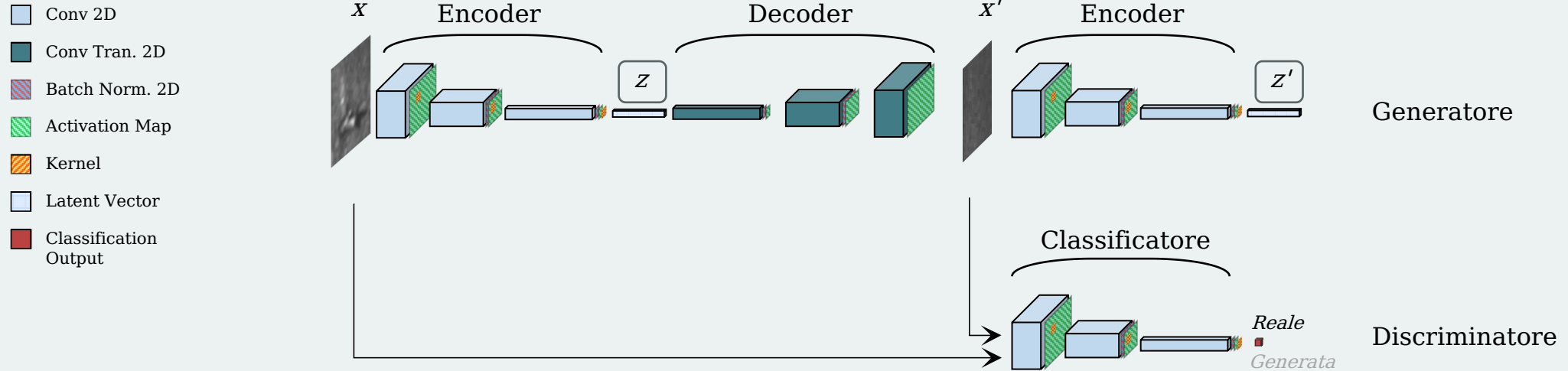
- > Il modello è addestrato per generare immagini normali
- > Un'immagine anomala tende ad essere ricostruita come un'immagine normale
- > L'anomaly score è definito sulla base dell'errore degli encoding
- > Un'immagine anomala determina un alto errore degli encoding

Ganomaly



- > Il modello è addestrato per generare immagini normali
- > Un'immagine anomala tende ad essere **ricostruita** come un'immagine normale
- > L'anomaly score è definito sulla base dell'errore degli encoding
- > Un'immagine anomala determina un **alto** errore degli encoding

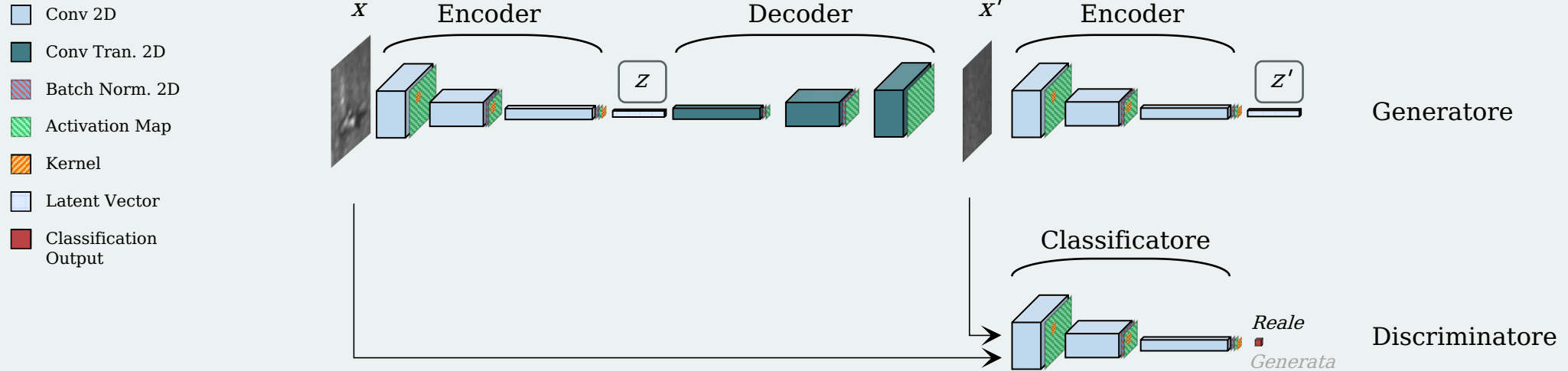
Ganomaly



- > Il modello è addestrato per generare immagini normali
- > Un'immagine anomala tende ad essere **ricostruita** come un'immagine normale
- > L'**anomaly score** è definito sulla base dell'errore degli encoding
- > Un'immagine anomala determina un **alto** errore degli encoding

$$A(x) = \| z - z' \|_2$$

Ganomaly



- > Il modello è addestrato per generare immagini normali
- > Un'immagine anomala tende ad essere **ricostruita** come un'immagine normale
- > L'**anomaly score** è definito sulla base dell'errore degli encoding
- > Un'immagine anomala determina un **alto** errore degli encoding

$$A(x) = \| z - z' \|_2$$



2

Approccio Proposto

- Patch-wise Training e Patch-Ganomaly
- TL-Ganomaly
- Post-Processing: apprendimento di filtri

Obiettivi:

- Identificare la **regione anomala** all'interno di un'immagine
- Soluzione applicabile ad immagini di dimensione **arbitraria**

Soluzioni:

Patch-Ganomaly: estensione di Ganomaly mediante l'utilizzo della tecnica di addestramento 'Patch-wise Training'



Patch-wise Training



Patch-wise Training

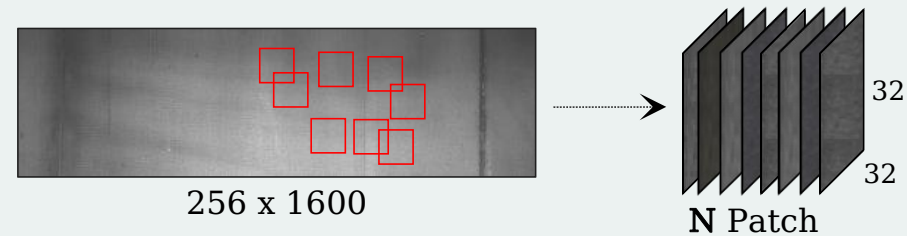
Da immagini a patch:

- > Tecnica di apprendimento basata su porzioni di immagini (**patch**)
- > Da ogni immagine viene estratto un numero N di patch
- > Il training set di immagini viene trasformato in un **patch set**

Patch-wise Training

Da immagini a patch:

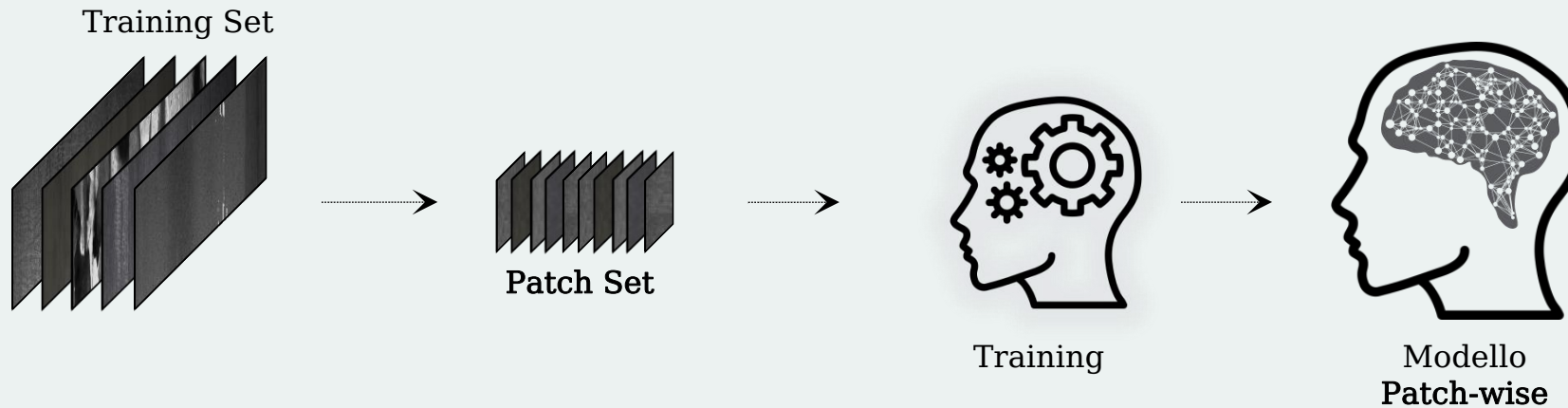
- > Tecnica di apprendimento basata su porzioni di immagini (**patch**)
- > Da ogni immagine viene estratto un numero **N** di patch
- > Il training set di immagini viene trasformato in un **patch set**



Patch-wise Training

Da immagini a patch:

- > Tecnica di apprendimento basata su porzioni di immagini (**patch**)
- > Da ogni immagine viene estratto un numero **N** di patch
- > Il training set di immagini viene trasformato in un **patch set**



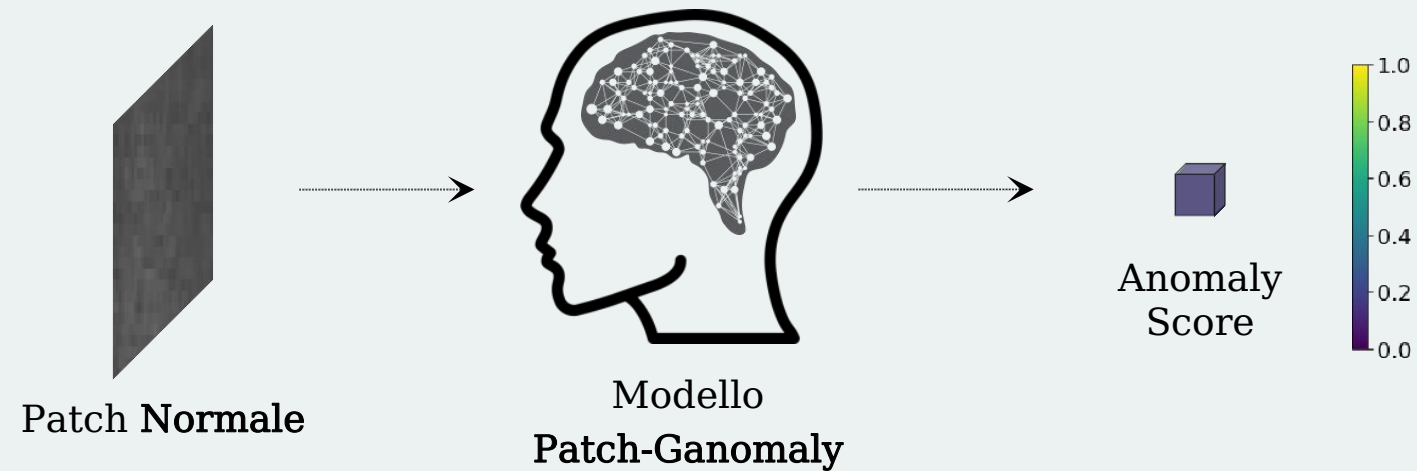


Patch-Ganomaly



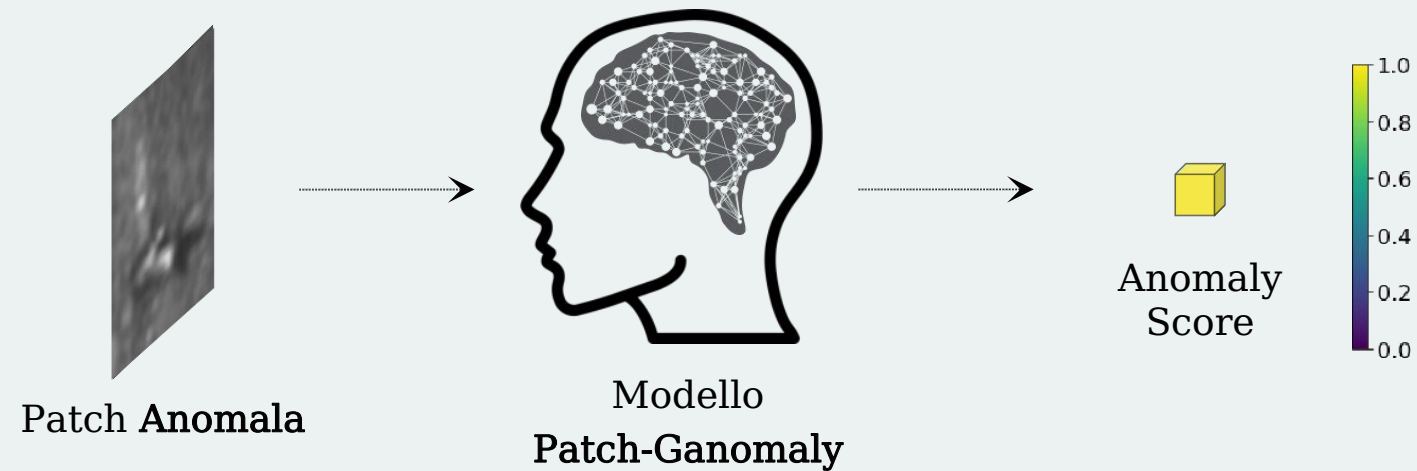
Inferenza


A partire da una patch, **Patch-Ganomaly** calcola l'anomaly score ad esso relativo



Inferenza


A partire da una patch, **Patch-Ganomaly** calcola l'anomaly score ad esso relativo





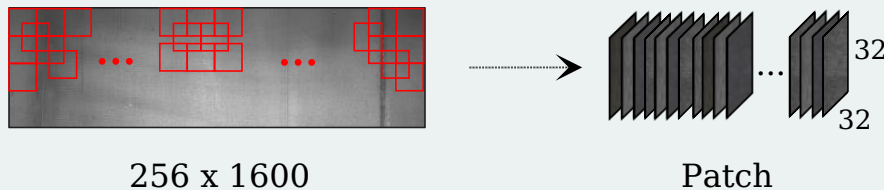
Patch-Ganomaly:

Segmentazione di un'immagine



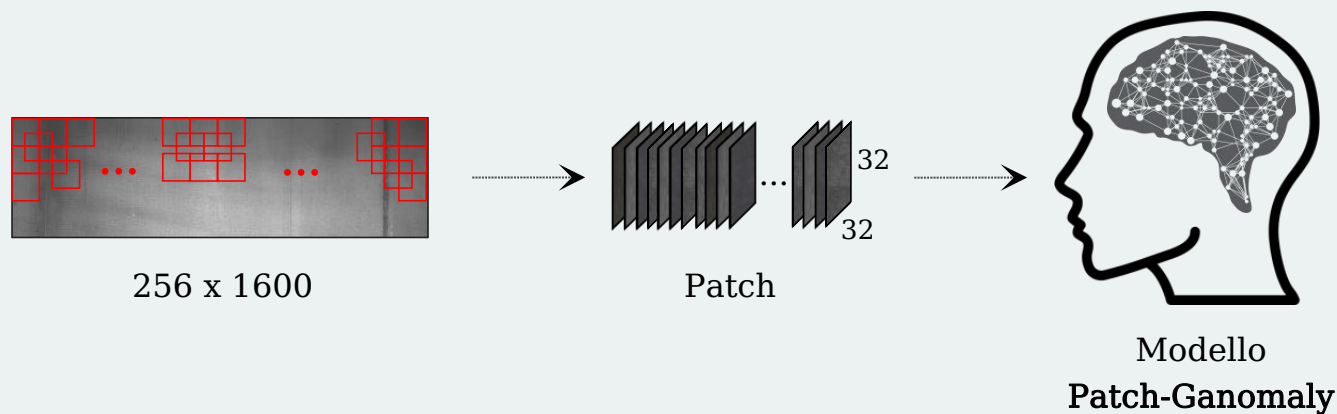
Segmentazione di un'immagine

- > Estrazione di **tutte** le patch che compongono un'immagine
- > Inferenza del modello **Patch-Ganomaly** sulle patch estratte
- > Gli anomaly score delle patch formano l'**anomaly score map** associata all'immagine
- > Stabilendo una **soglia di decisione** si possono discriminare gli anomaly score provenienti da pixel anomali



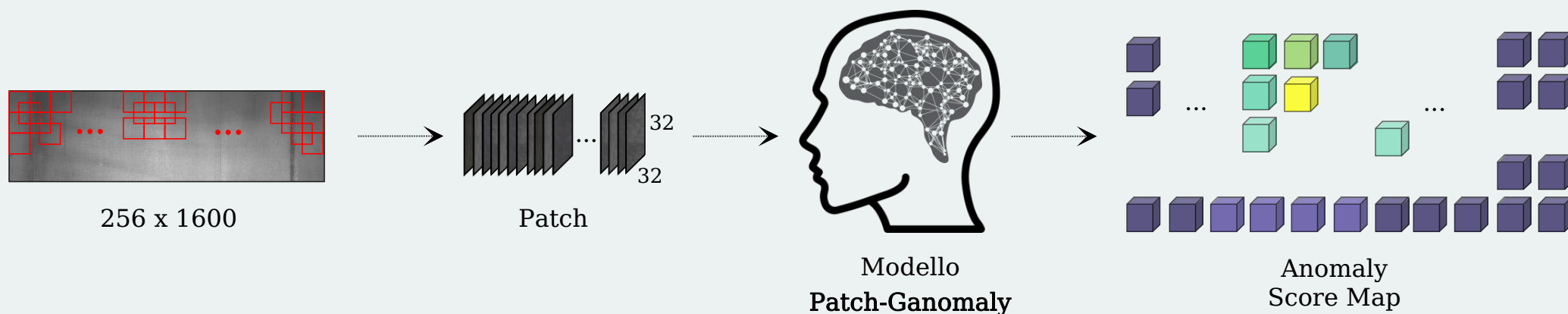
Segmentazione di un'immagine

- > Estrazione di **tutte** le patch che compongono un'immagine
- > Inferenza del modello **Patch-Ganomaly** sulle patch estratte
- > Gli anomaly score delle patch formano l'**anomaly score map** associata all'immagine
- > Stabilendo una **soglia di decisione** si possono discriminare gli anomaly score provenienti da pixel anomali



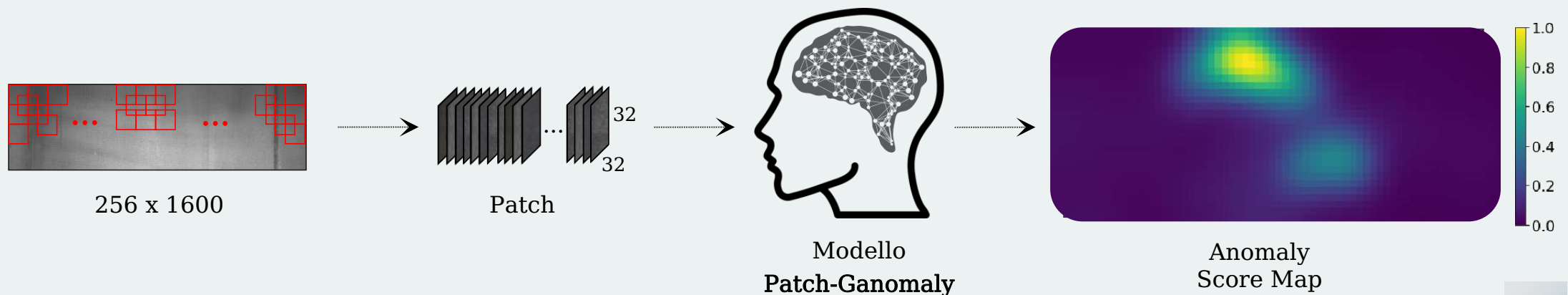
Segmentazione di un'immagine

- > Estrazione di **tutte** le patch che compongono un'immagine
- > Inferenza del modello **Patch-Ganomaly** sulle patch estratte
- > Gli anomaly score delle patch formano l'**anomaly score map** associata all'immagine
- > Stabilendo una **soglia di decisione** si possono discriminare gli anomaly score provenienti da pixel anomali



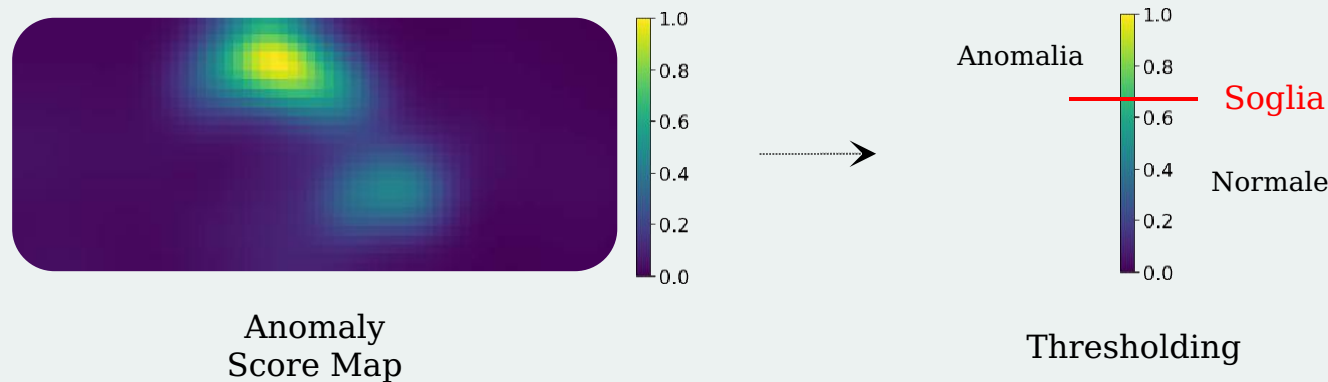
Segmentazione di un'immagine

- > Estrazione di **tutte** le patch che compongono un'immagine
- > Inferenza del modello **Patch-Ganomaly** sulle patch estratte
- > Gli anomaly score delle patch formano l'**anomaly score map** associata all'immagine
- > Stabilendo una **soglia di decisione** si possono discriminare gli anomaly score provenienti da pixel anomali



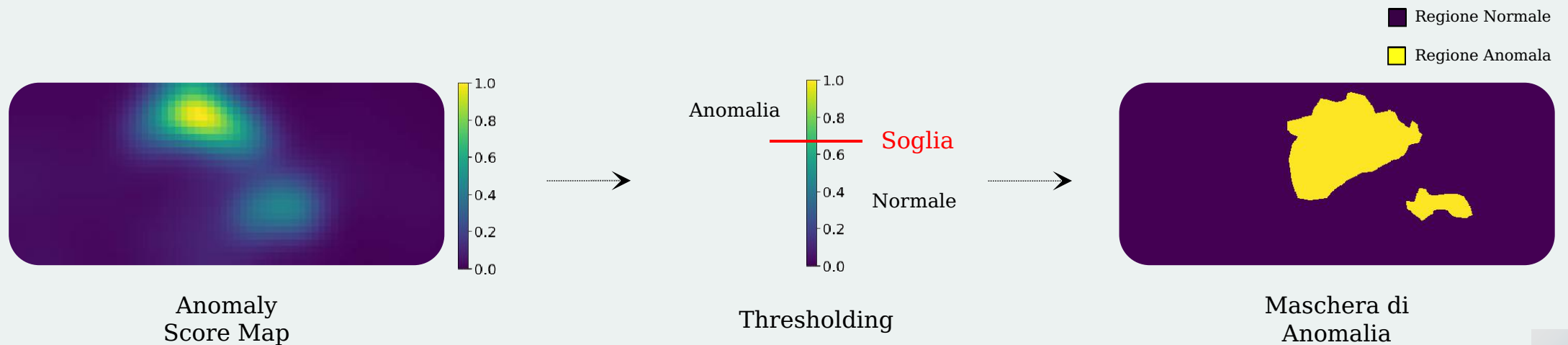
Segmentazione di un'immagine

- > Estrazione di **tutte** le patch che compongono un'immagine
- > Inferenza del modello **Patch-Ganomaly** sulle patch estratte
- > Gli anomaly score delle patch formano l'**anomaly score map** associata all'immagine
- > Stabilendo una **soglia di decisione** si possono discriminare gli anomaly score provenienti da pixel anomali



Segmentazione di un'immagine

- > Estrazione di **tutte** le patch che compongono un'immagine
- > Inferenza del modello **Patch-Ganomaly** sulle patch estratte
- > Gli anomaly score delle patch formano l'**anomaly score map** associata all'immagine
- > Stabilendo una **soglia di decisione** si possono discriminare gli anomaly score provenienti da pixel anomali





2

Approccio Proposto

- Patch-wise Training e Patch-Ganomaly
- TL-Ganomaly
- Post-Processing: apprendimento di filtri



Obiettivi:

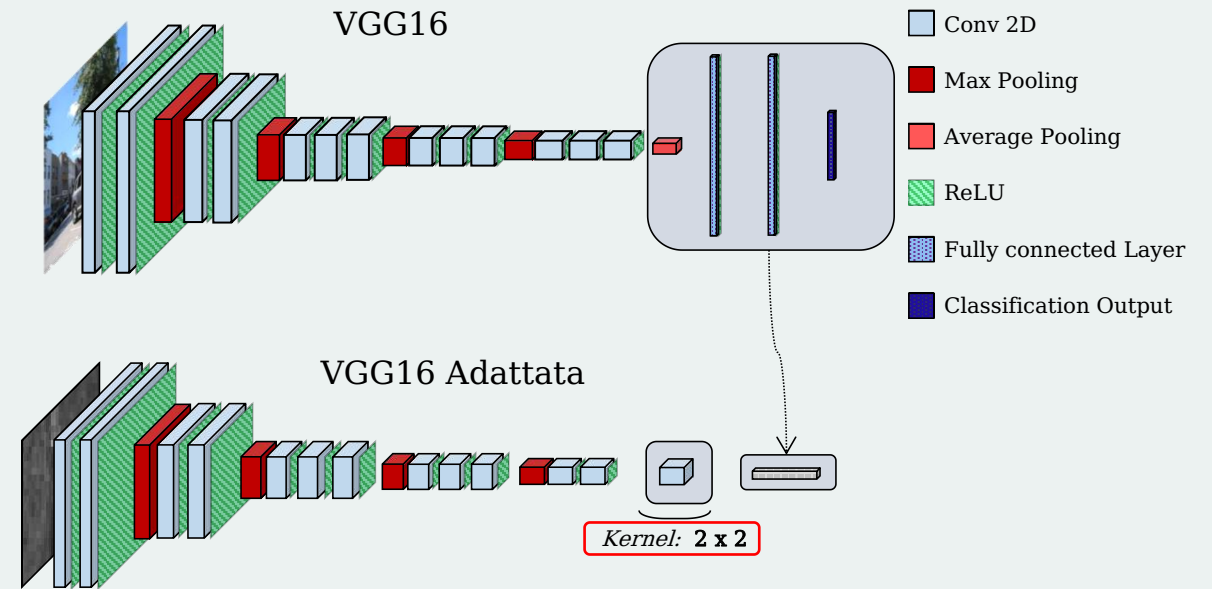
- Migliorare la **precisione** del modello Patch-Ganomaly
- Valutare il contributo apportato dall'**integrazione** di **reti pre-addestrate**

Soluzioni:

- **TL-Ganomaly**: patch-wise training e transfer learning su Ganomaly
- 

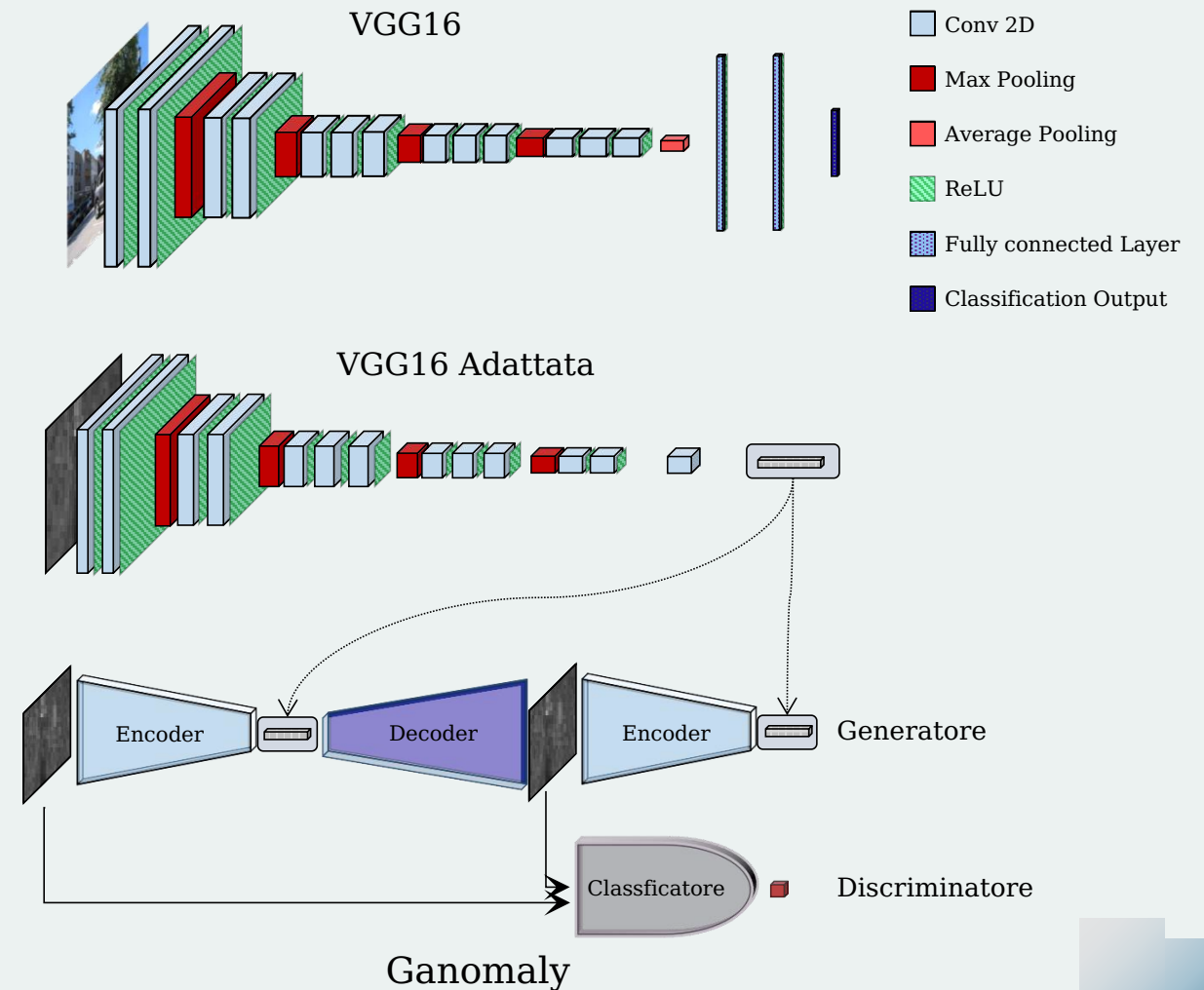
TL-Ganomaly

- > Rete pre-addestrata: VGG16
- > Riduzione dei layer della VGG16
- > Estrazione della rappresentazione latente



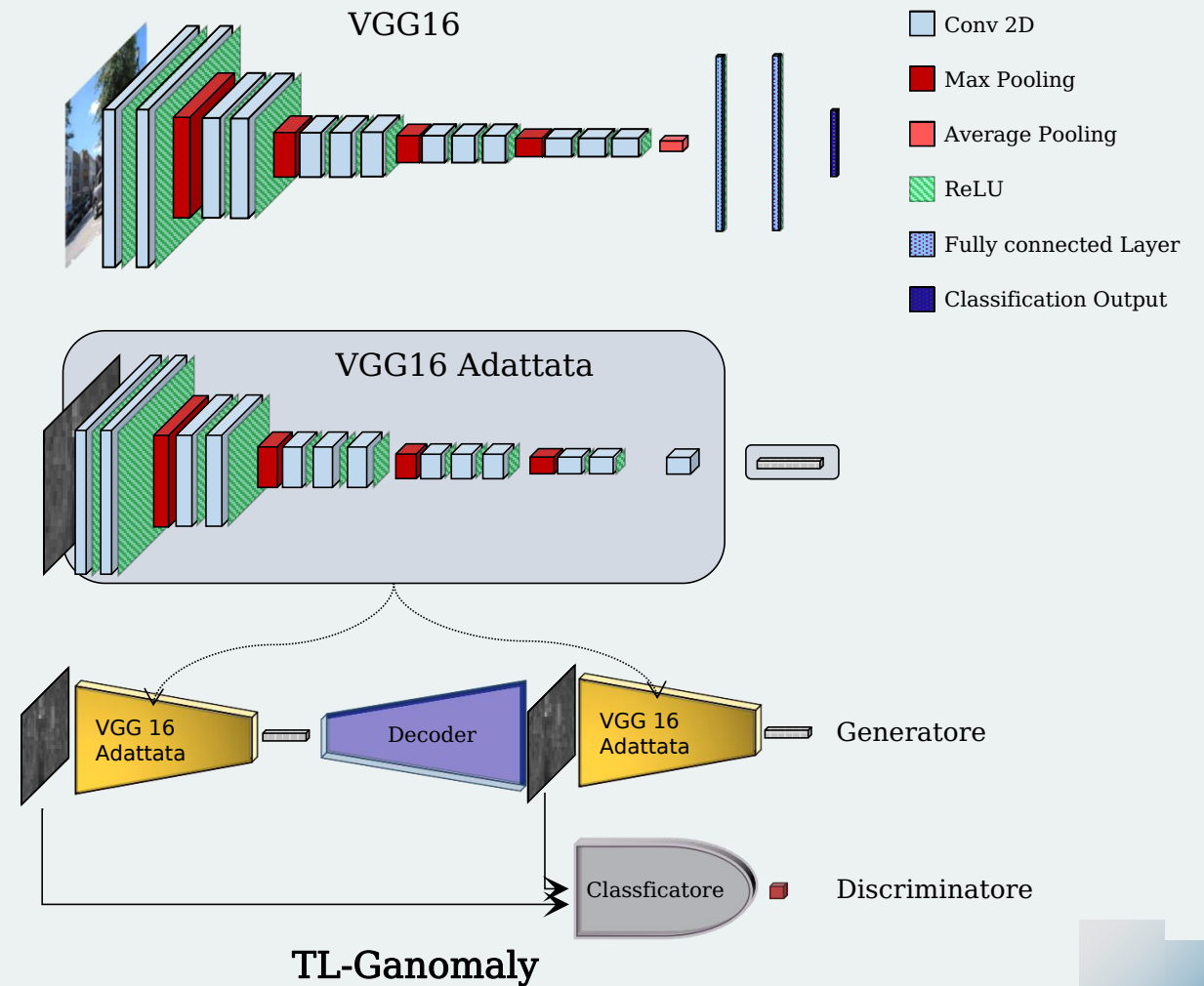
TL-Ganomaly

- > Rete pre-addestrata: VGG16
- > Riduzione dei layer della VGG16
- > Estrazione della rappresentazione latente
- > Adattamento della **rappresentazione latente** della VGG16 al vettore latente di Ganomaly

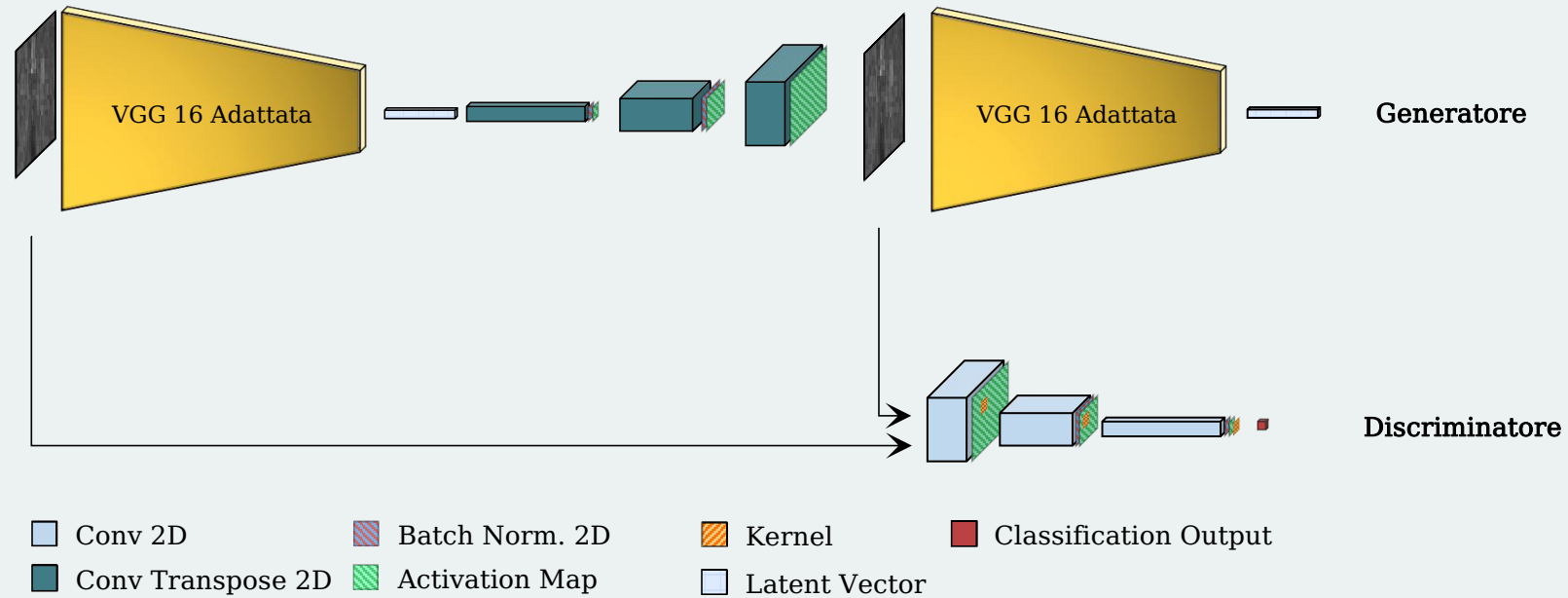


TL-Ganomaly

- > Rete pre-addestrata: VGG16
- > Riduzione dei layer della VGG16
- > Estrazione della rappresentazione latente
- > Adattamento della **rappresentazione latente** della VGG16 al vettore latente di Ganomaly
- > L'**encoder** del modello TL-Ganomaly diventa la **VGG16 adattata**



TL-Ganomaly





2

Approccio Proposto


- Patch-wise Training e Patch-Ganomaly
- TL-Ganomaly
- Post-Processing: apprendimento di filtri



Obiettivi:

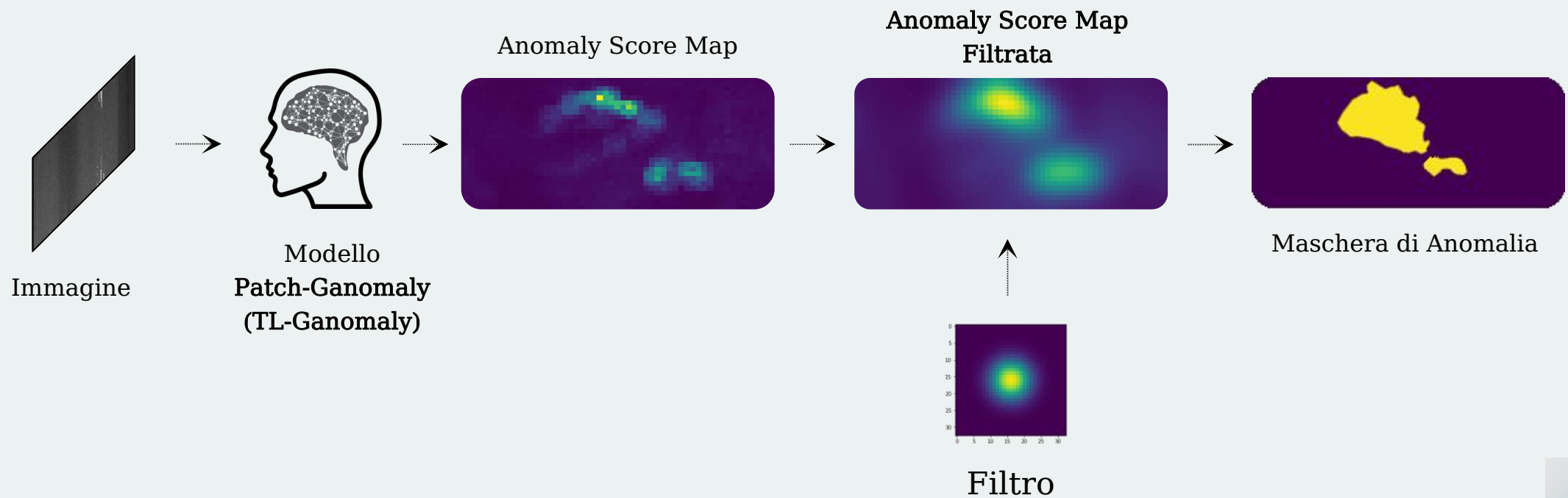
- Migliorare l'**efficacia** dei modelli di segmentazione proposti applicando un filtro agli **anomaly score map** generati dai modelli
- Ideare un post-processing **ad-hoc** per il dataset utilizzato

Soluzioni:

- **Conv-Processing**: apprendimento di filtri per la fase di post-processing
- 

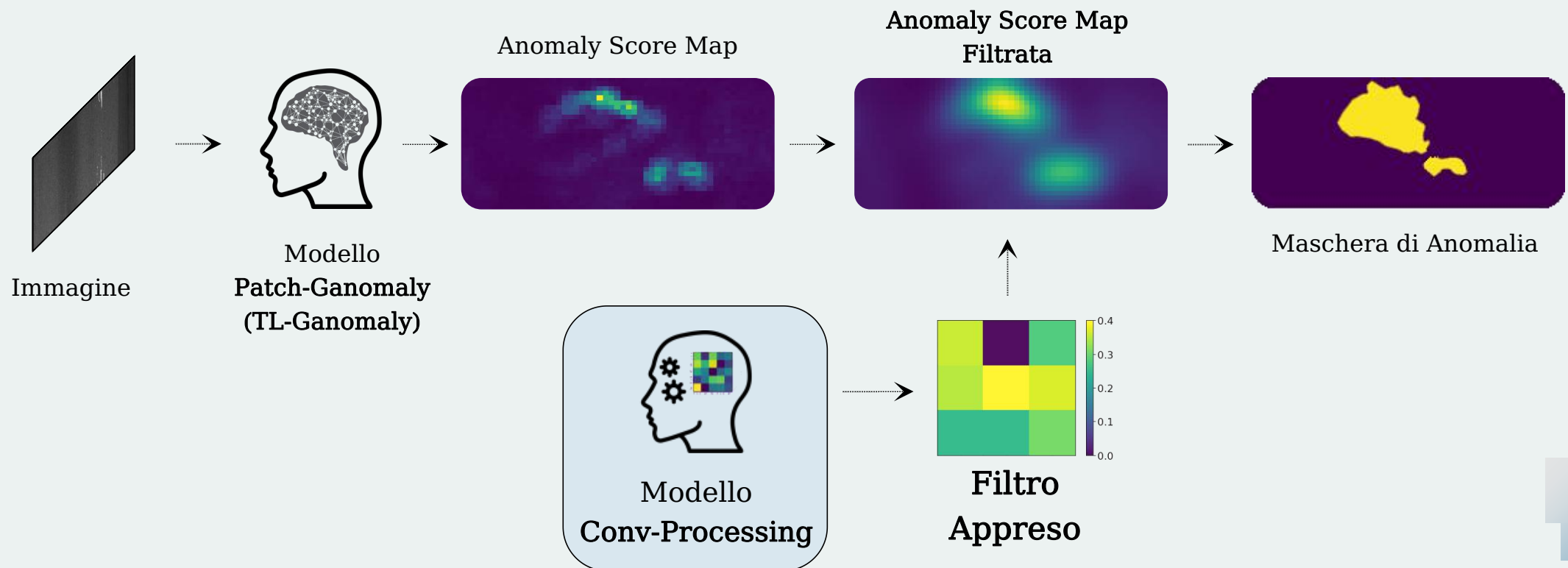
Post-Processing

- > Migliorativo per i problemi di **segmentazione**
- > Applicazione di un filtro all'**anomaly score map**



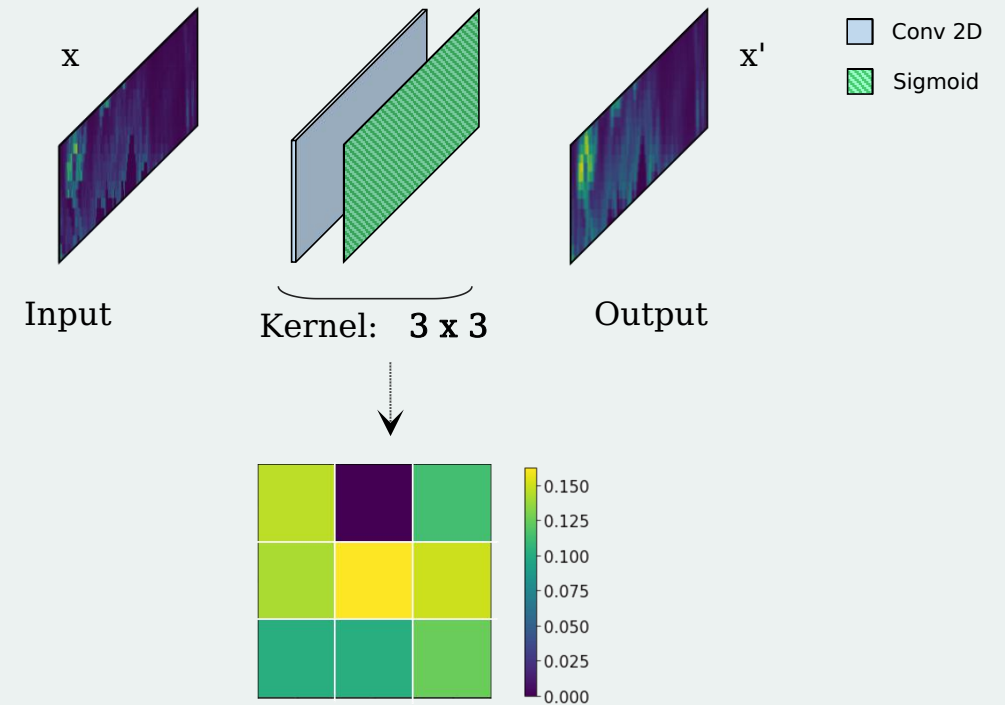
Post-Processing

- > Modello Supervisionato
- > Training set formato da una **piccola** parte di immagini etichettate
- > Fornisce il **filtro convoluzionale**



Conv-Processing

- > Rete Neurale: un layer **Convoluzionale**
- > Training set etichettato: **anomaly score map e maschere di anomalia**
- > Apprende i parametri del **kernel convoluzionale**





3

Esperimenti

- Dataset e Modelli di riferimento
- Patch-Ganomaly e TL-Ganomaly
- Post-Processing: apprendimento di filtri

Dataset

- > Dataset fornito da **Kaggle**
- > Competizione '**Steel Defect Detection**'
(<https://www.kaggle.com/c/severstal-steel-defect-detection>)
- > ~**15.000** immagini (256 x 1600)

kaggle





Esperimenti:

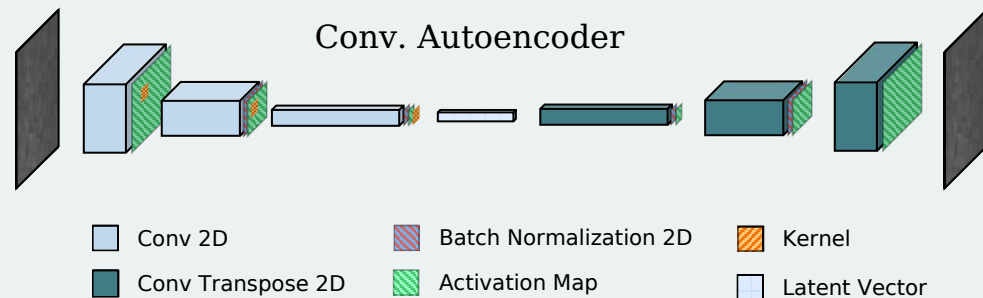
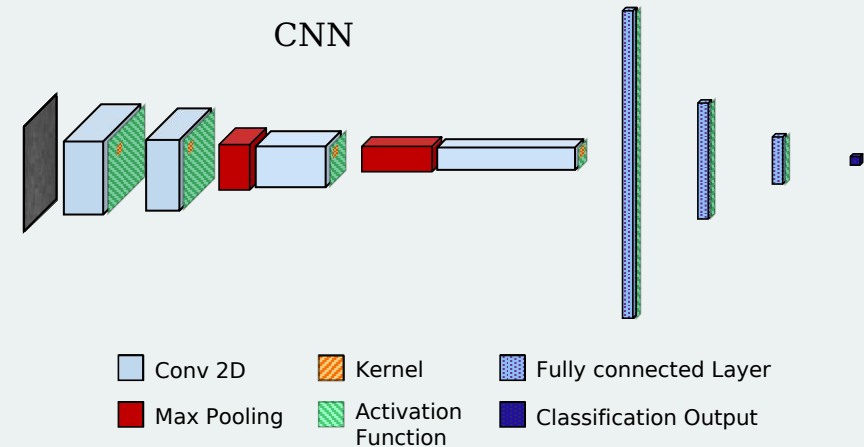
Modelli di Riferimento



Modelli di Riferimento

Il modelli proposti, **Patch-Ganomaly** e **TL-Ganomaly**, vengono confrontati con due **modelli di riferimento**:

- > Modello **Supervionato**: Rete Neurale Convoluzionale (CNN) fully-connected
- > Modello **Semi-Supervisionato**: Convolutional Autoencoder
- > Addestrati sfruttando la tecnica del **patch-wise training**







3

Esperimenti

- Dataset e Modelli di riferimento
- Patch-Ganomaly e TL-Ganomaly
- Post-Processing: apprendimento di filtri

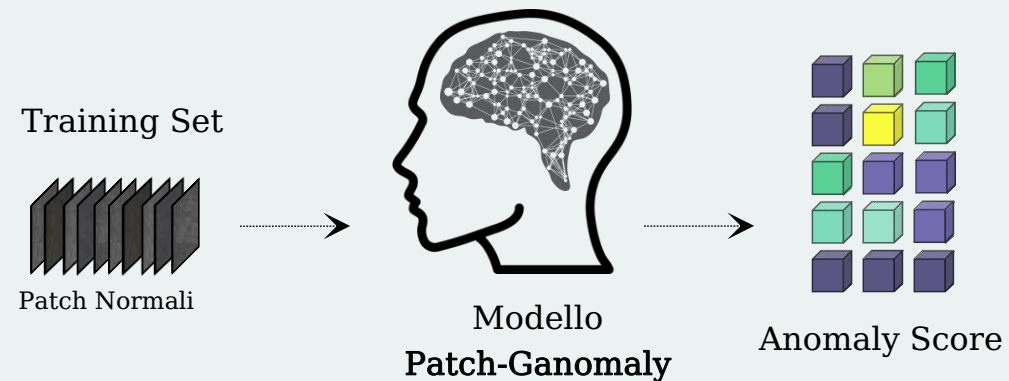
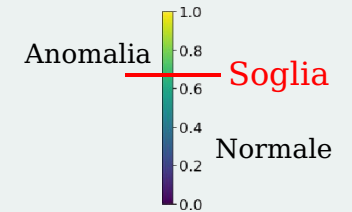


Esperimenti: Thresholding



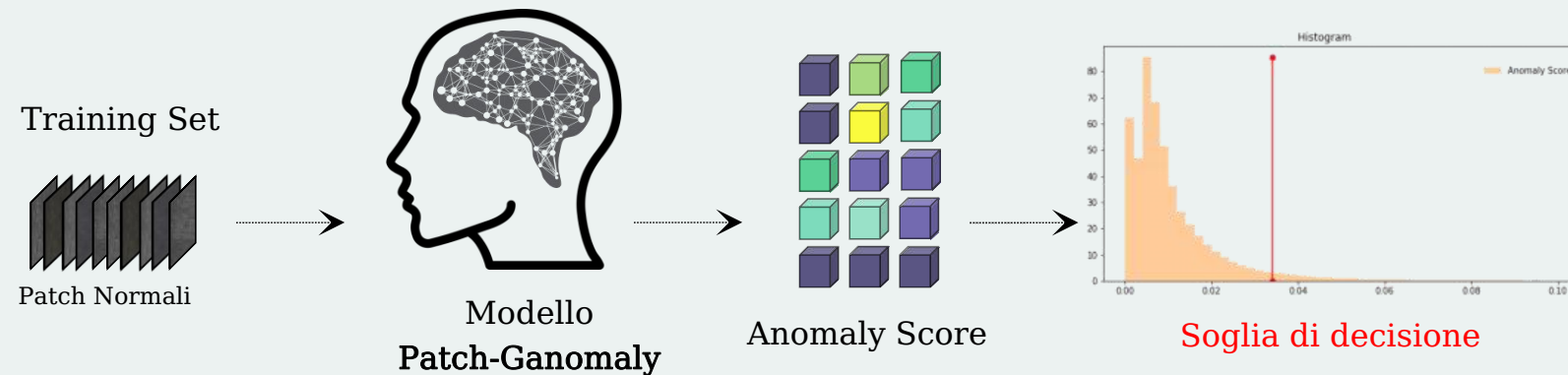
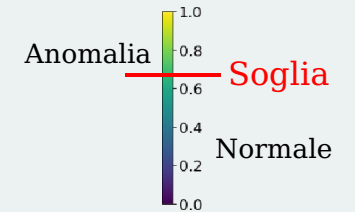
Thresholding

- > Patch normali del training set usato per l'addestramento
- > Calcolo degli anomaly score
- > Il quantile al 95% degli anomaly score rappresenta la soglia di decisione
- > Tolleranza del 5% di falsi positivi



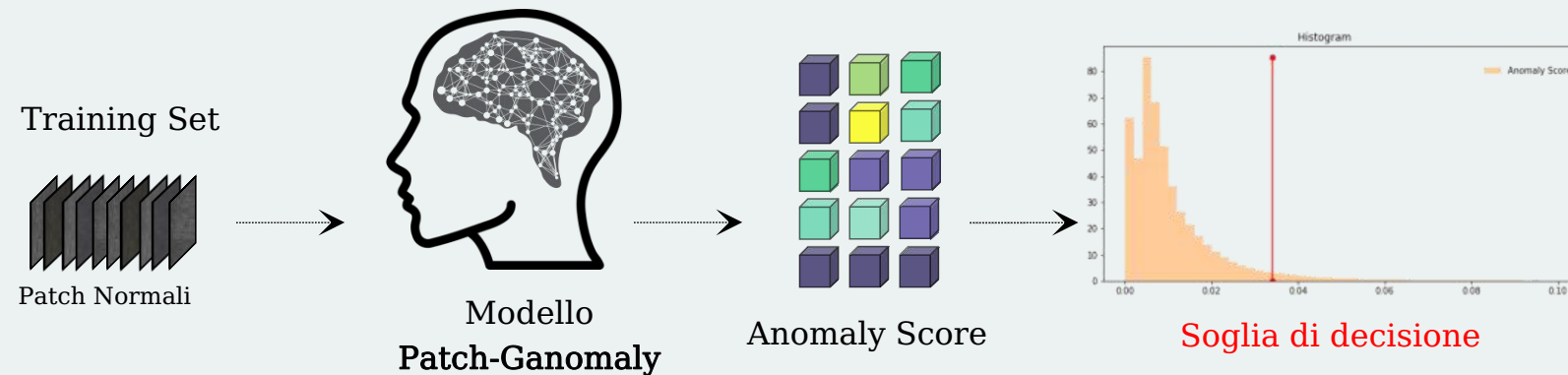
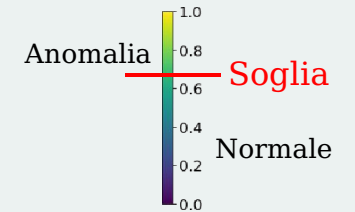
Thresholding

- > Patch normali del training set usato per l'addestramento
- > Calcolo degli anomaly score
- > Il **quantile al 95%** degli anomaly score rappresenta la **soglia di decisione**
- > Tolleranza del 5% di falsi positivi



Thresholding

- > Patch normali del training set usato per l'addestramento
- > Calcolo degli anomaly score
- > Il **quantile al 95%** degli anomaly score rappresenta la **soglia di decisione**
- > Tolleranza del 5% di **falsi positivi**





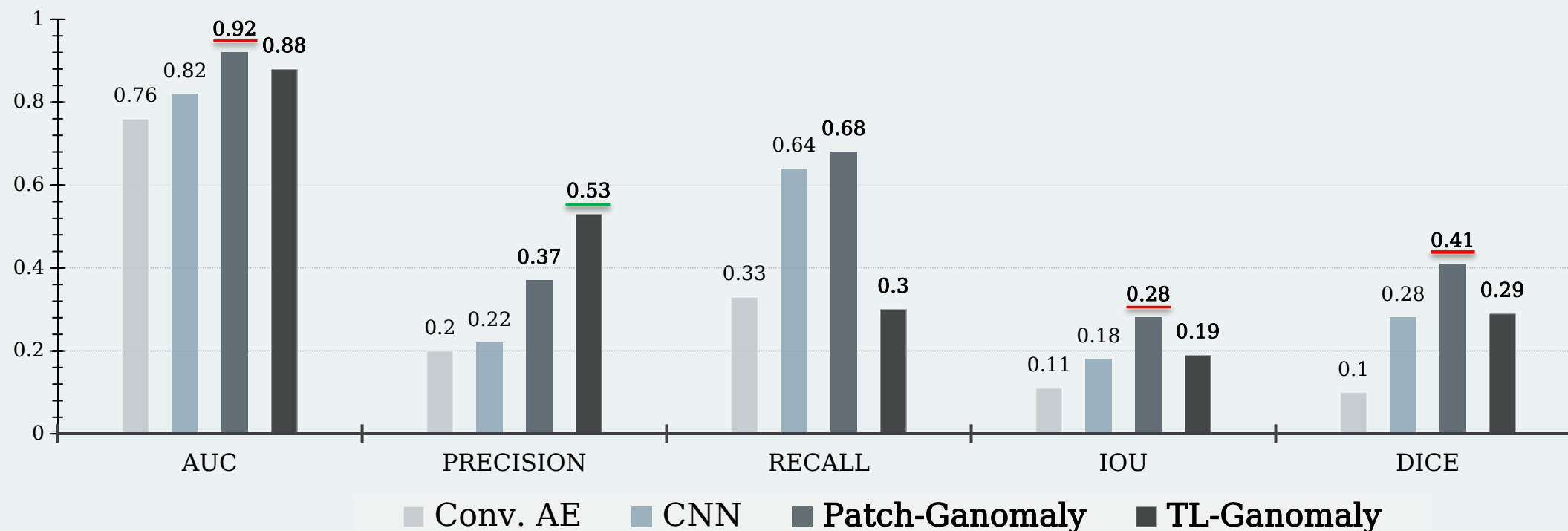
Esperimenti:

Risultati

Risultati

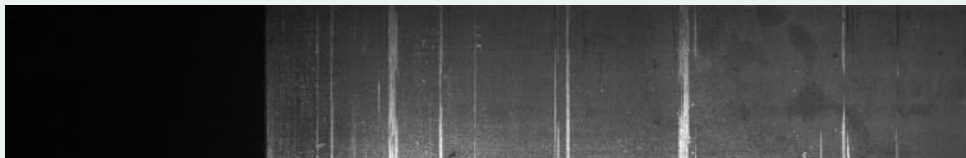
Su un test set di 120 immagini:

- > Segmentazione migliore: Patch-Ganomaly
- > Migliore Precisione: TL-Ganomaly



Immagini Test

Immagine Originale

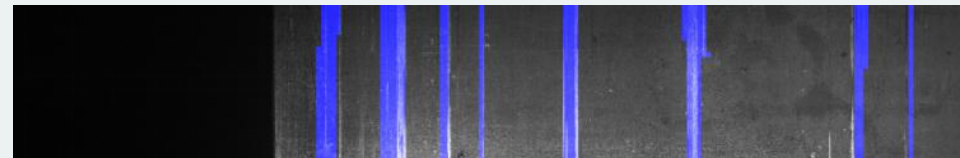


Maschera di Anomalia

Difetto di tipo 1



Difetto di tipo 2



Difetto di tipo 3



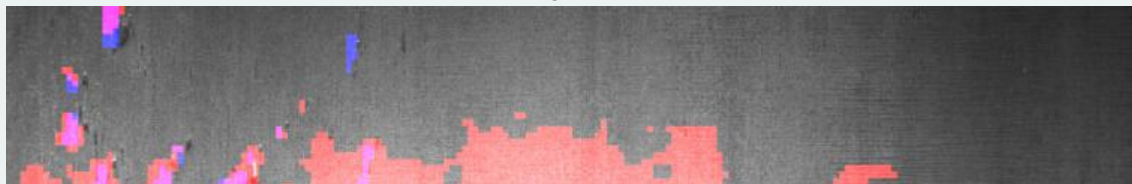
 Pixel Anomali

Difetti di tipo 1

Convolutional Autoencoder



CNN Fully-connected



Patch-Ganomaly

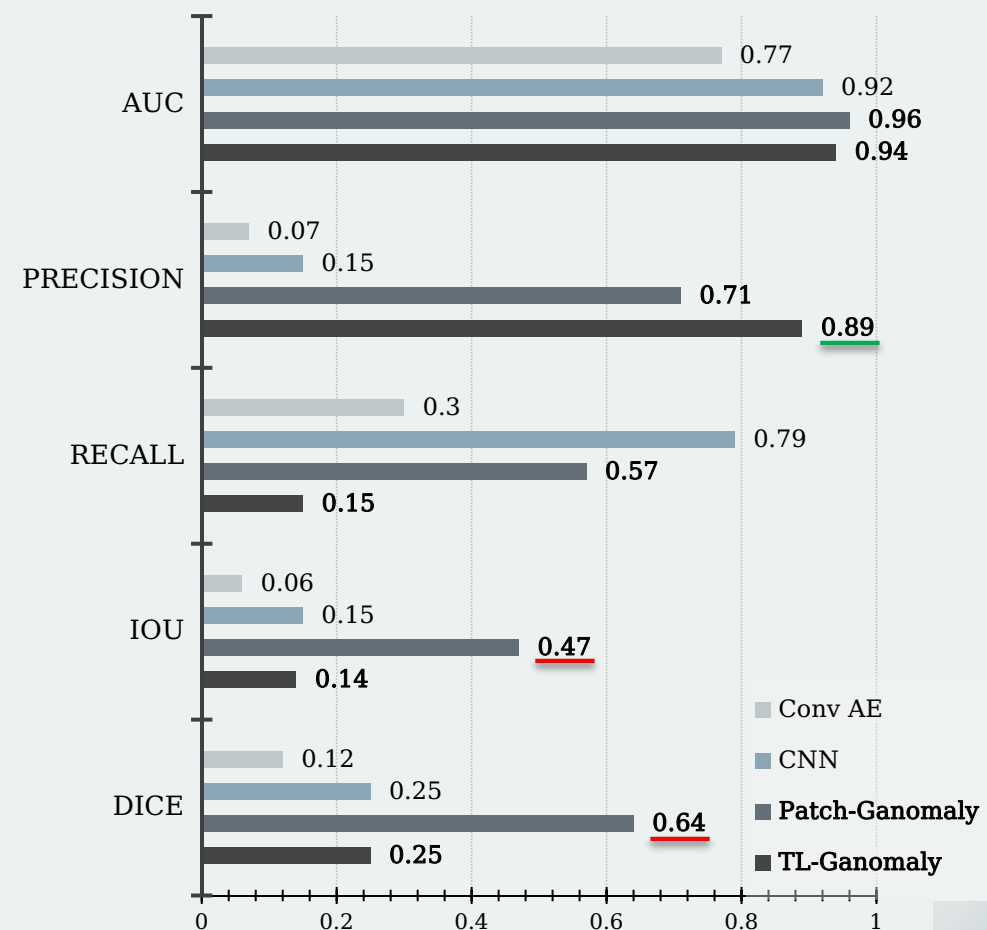


TL-Ganomaly



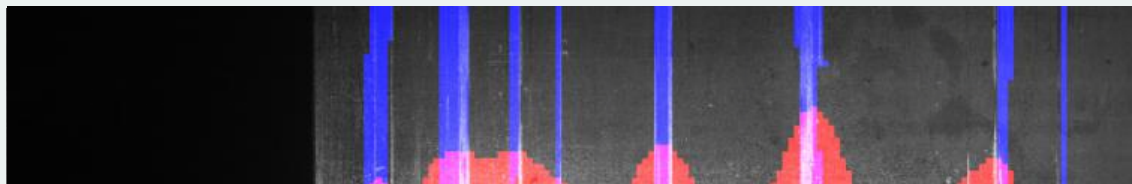
Pixel Anomali Pixel Riconosciuti Anomali

Performance

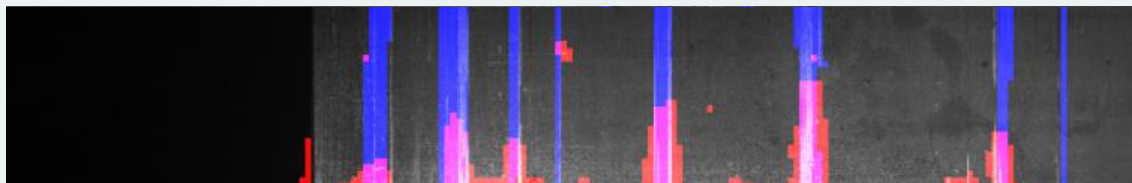


Difetti di tipo 2

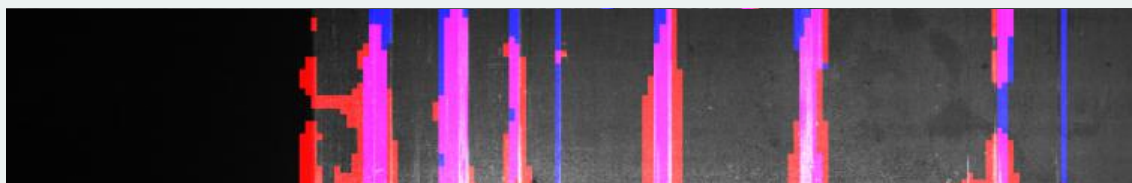
Convolutional Autoencoder



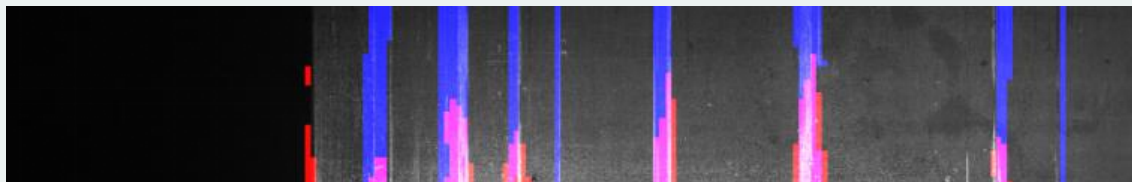
CNN Fully-connected



Patch-Ganomaly



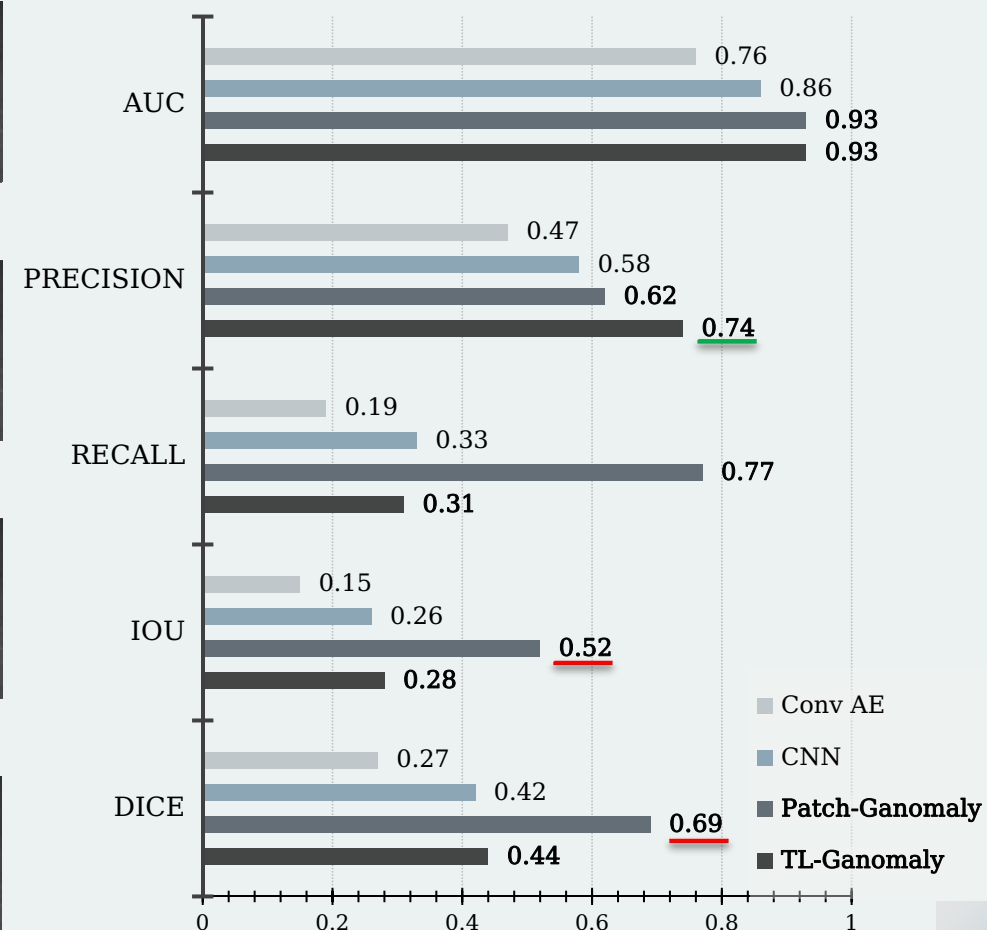
TL-Ganomaly



Pixel Anomali

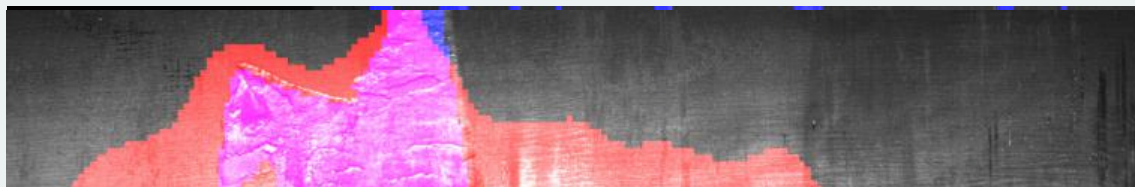
Pixel Riconosciuti Anomali

Performance

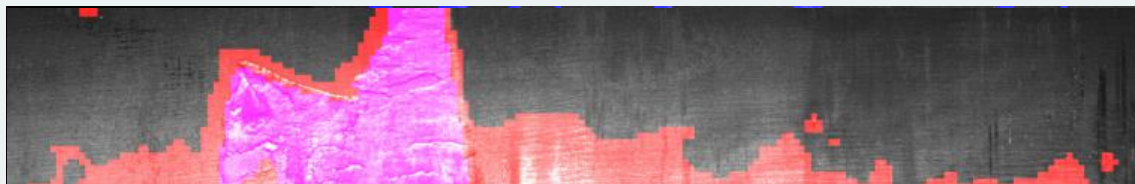


Difetti di tipo 3

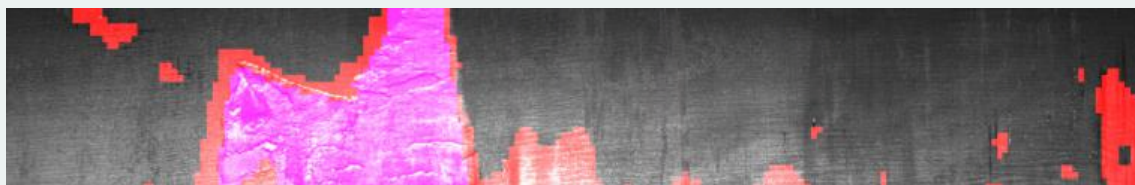
Convolutional Autoencoder



CNN Fully-connected



Patch-Ganomaly

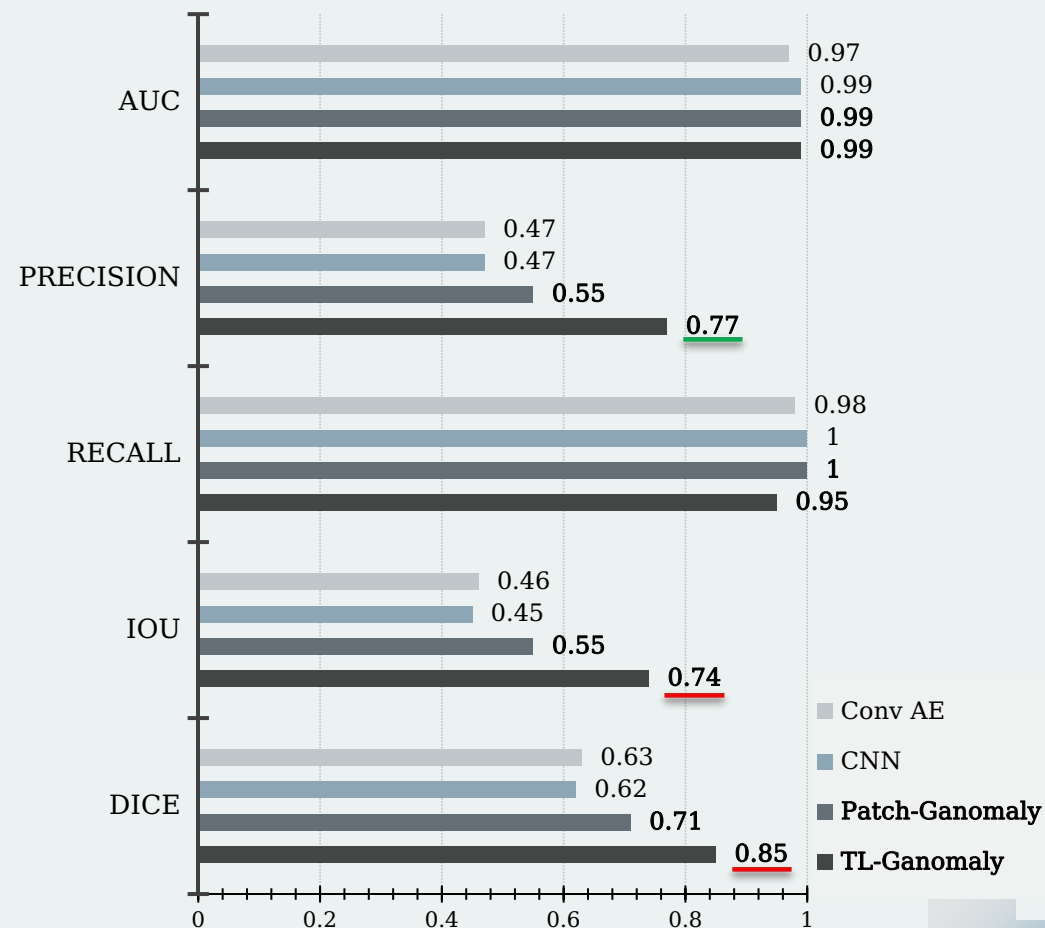


TL-Ganomaly



Pixel Anomali Pixel Riconosciuti Anomali

Performance





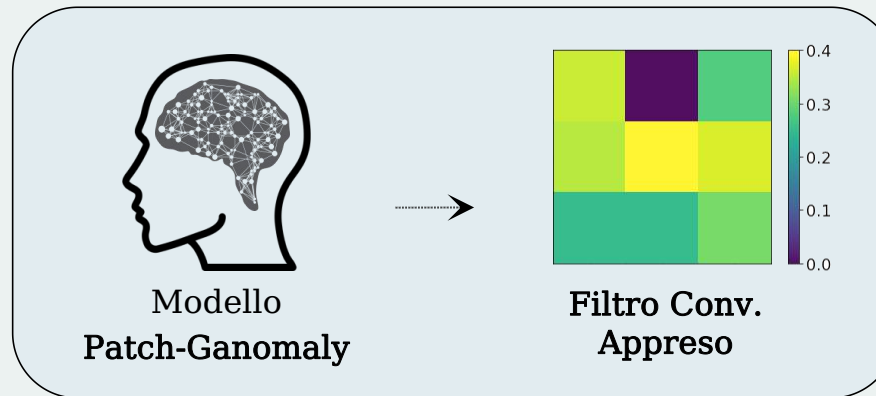
3

Esperimenti

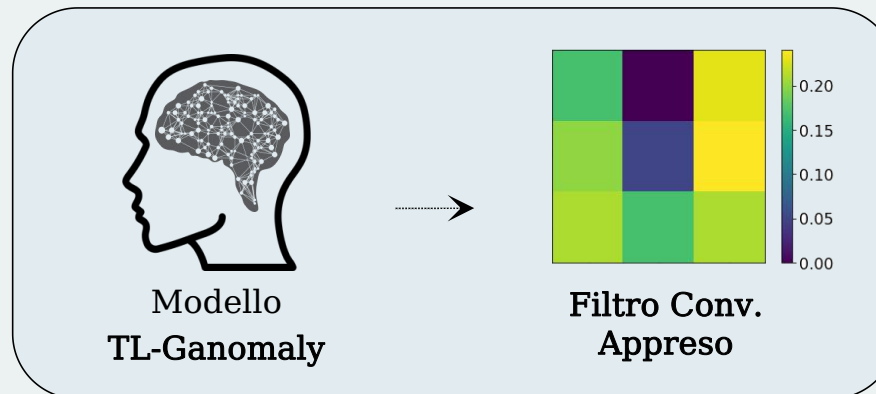
- Dataset e Modelli di riferimento
- Patch-Ganomaly e TL-Ganomaly
- Post-Processing: apprendimento di filtri

Post-Processing

PG-Detector



TL-Detector





Conv-Processing:

PG-Detector



Risultati

Su un test set di 120 immagini:

- > Segmentazione migliore: PG-Detector
- > Migliore Precisione: PG-Detector
- > Migliore Precision/Recall: PG-Detector

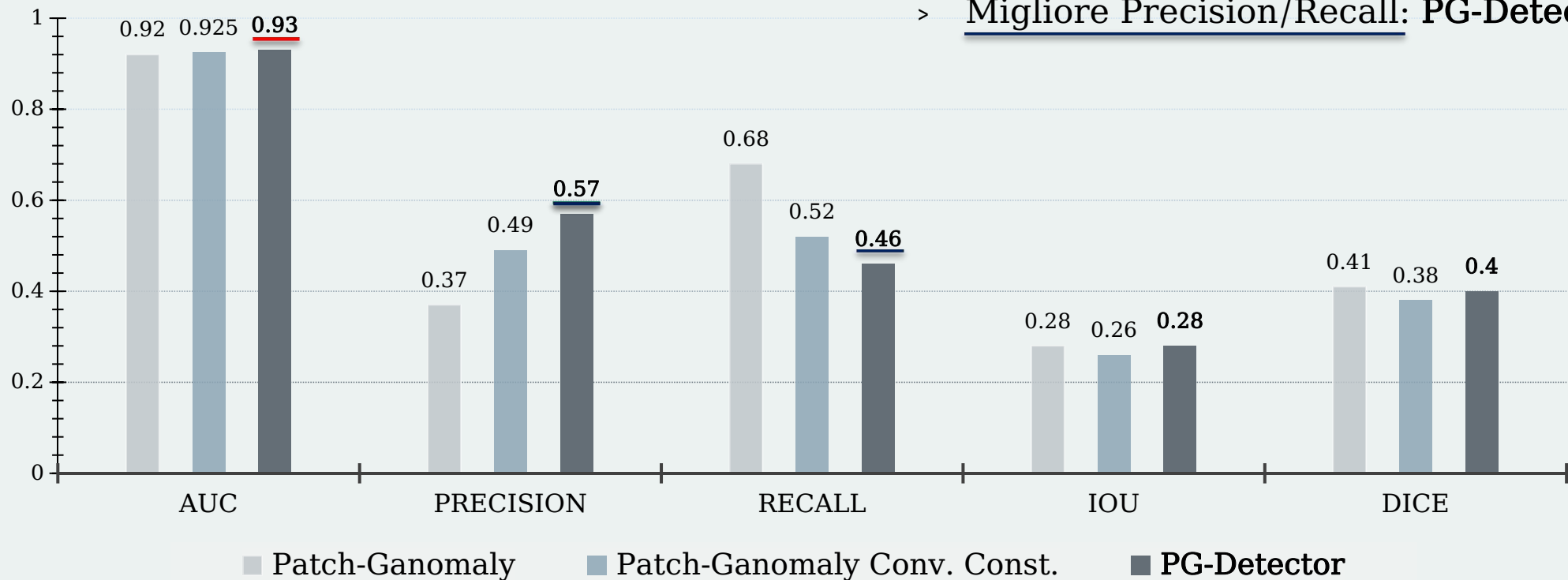
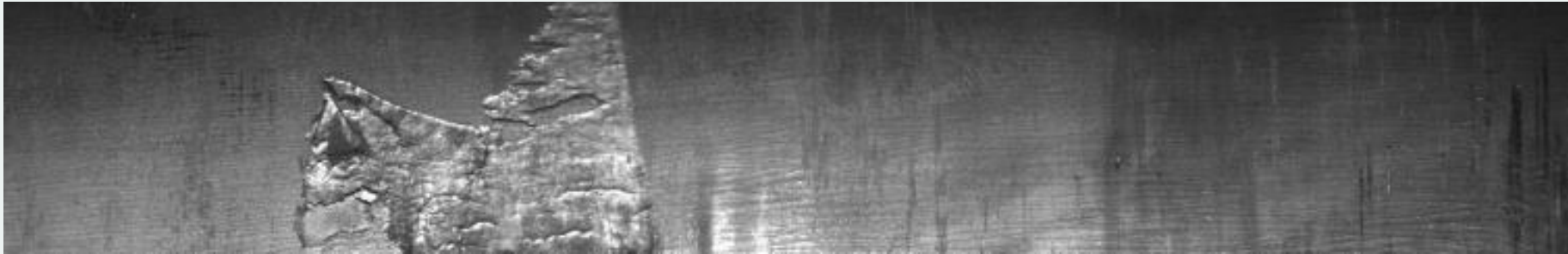



Immagine Test

Immagine Originale



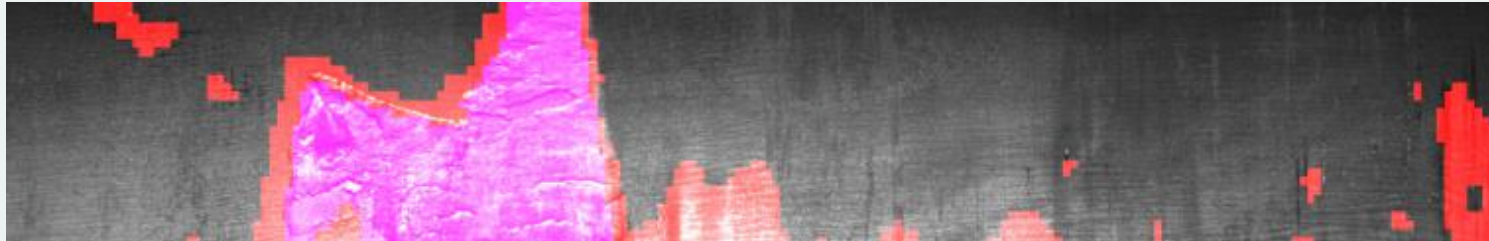
Maschera di Anomalia



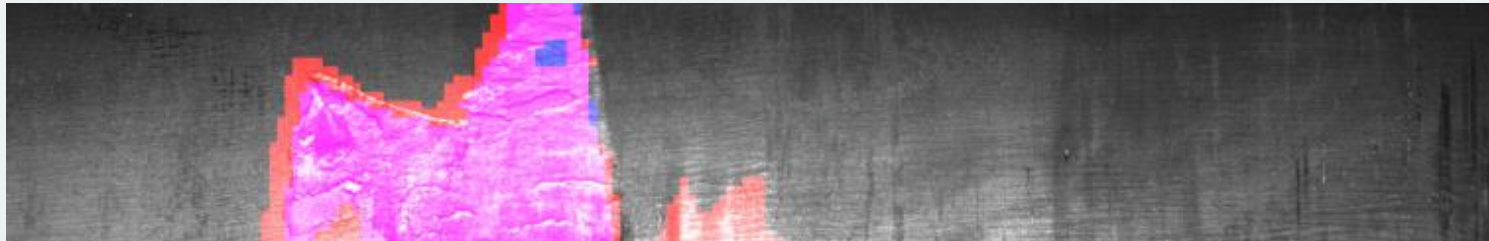
 Pixel Anomali

Post-Processing

Patch-Ganomaly

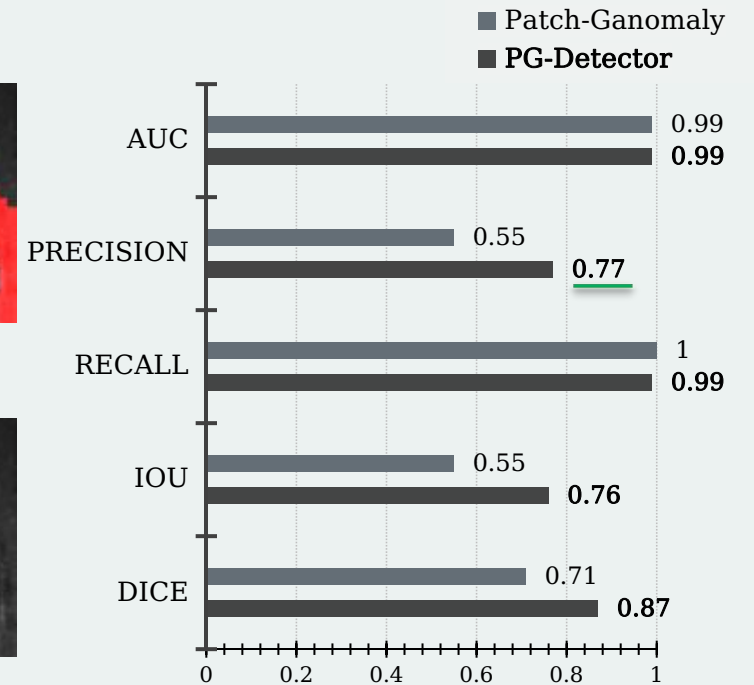



PG-Detector




- Pixel Anomali
- Pixel Riconosciuti Anomali

Performance





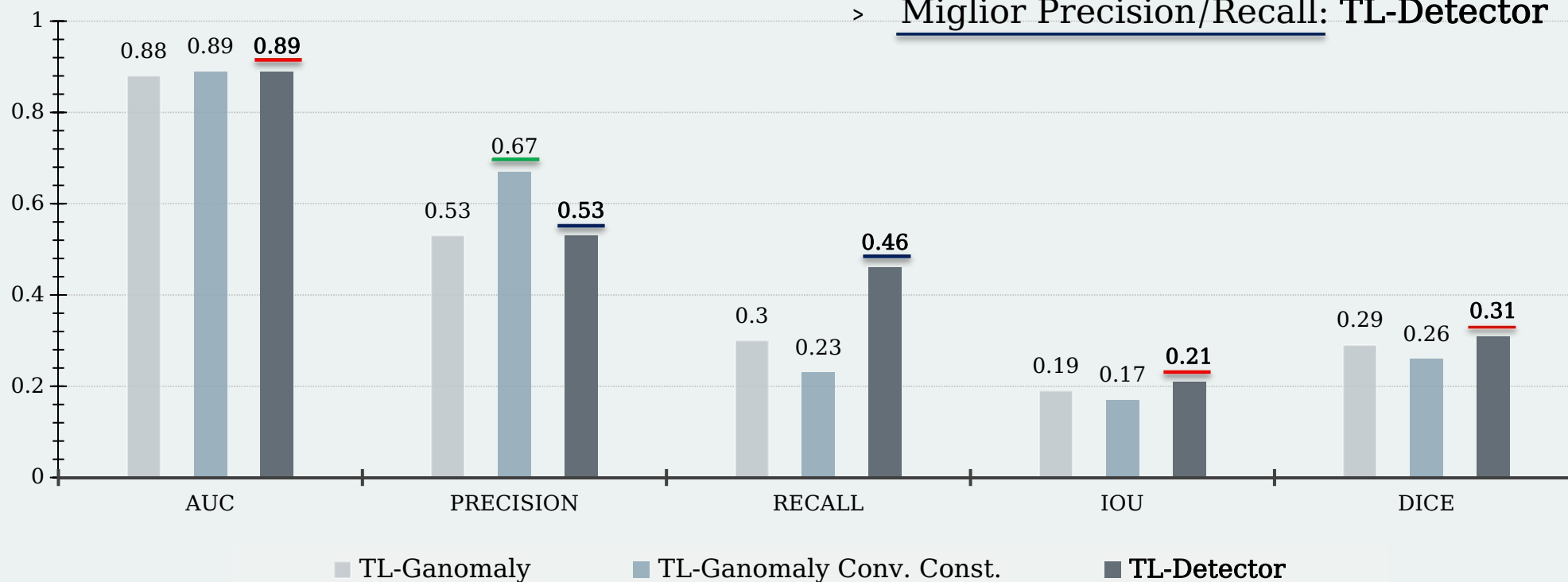
Conv-Processing: TL-Detector



Risultati

Su un test set di 120 immagini:

- > Segmentazione migliore: TL-Detector
- > Migliore Precisione: TL-Ganomaly Conv.Costante
- > Miglior Precision/Recall: TL-Detector





4

Conclusioni

Conclusioni

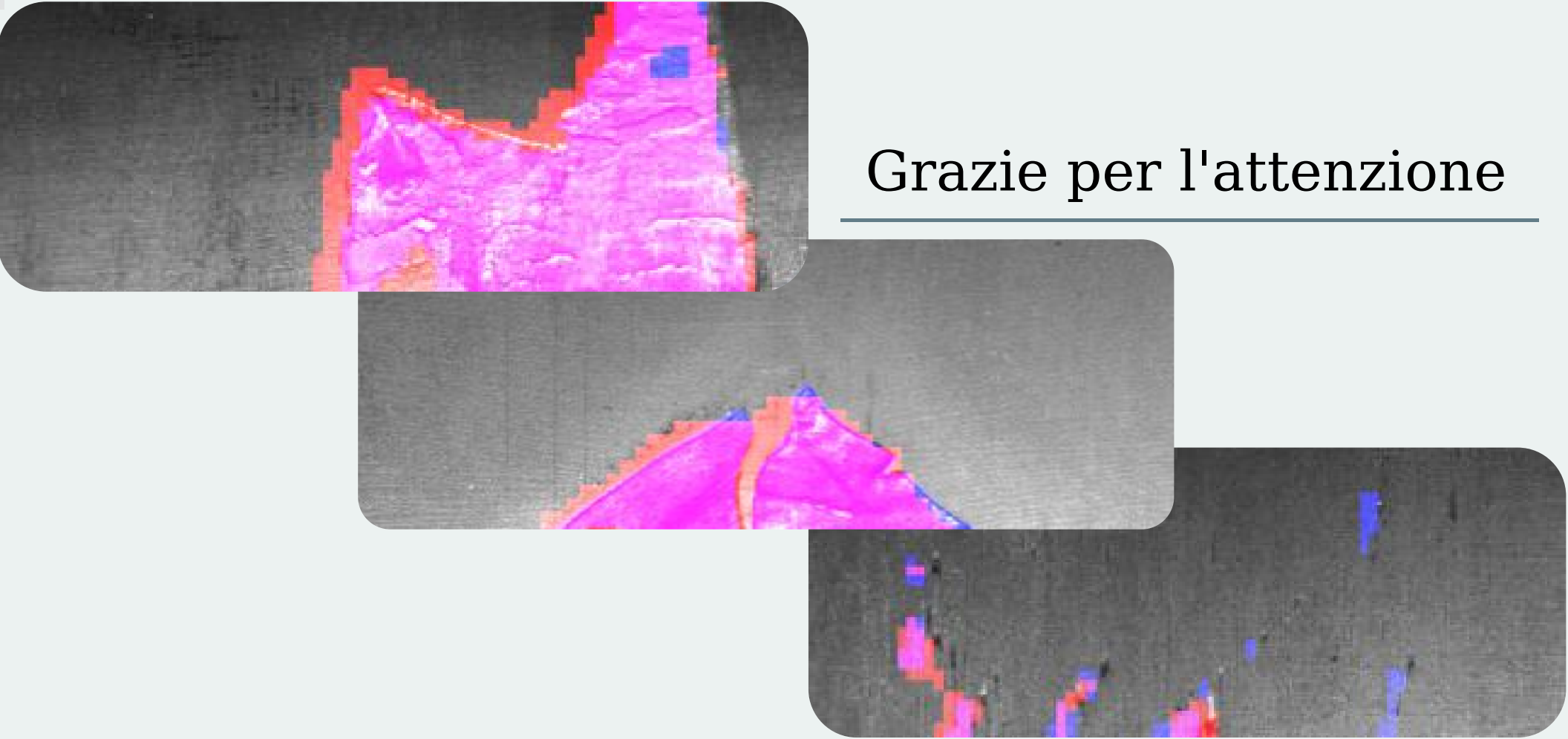


Contributi:

- > Mediante l'utilizzo del **patch-wise training** e il **transfer learning** abbiamo creato i modelli **Patch-Ganomaly** e **TL-Ganomaly**
- > Riconoscono la **regione anomala** all'interno di un'immagine, risolvendo il problema di **segmentazione**
- > Sono applicabili ad immagini di **dimensione arbitraria**
- > **Conv-Processing** fornisce il filtro **ad hoc** per le **anomaly score map** generate dai modelli proposti

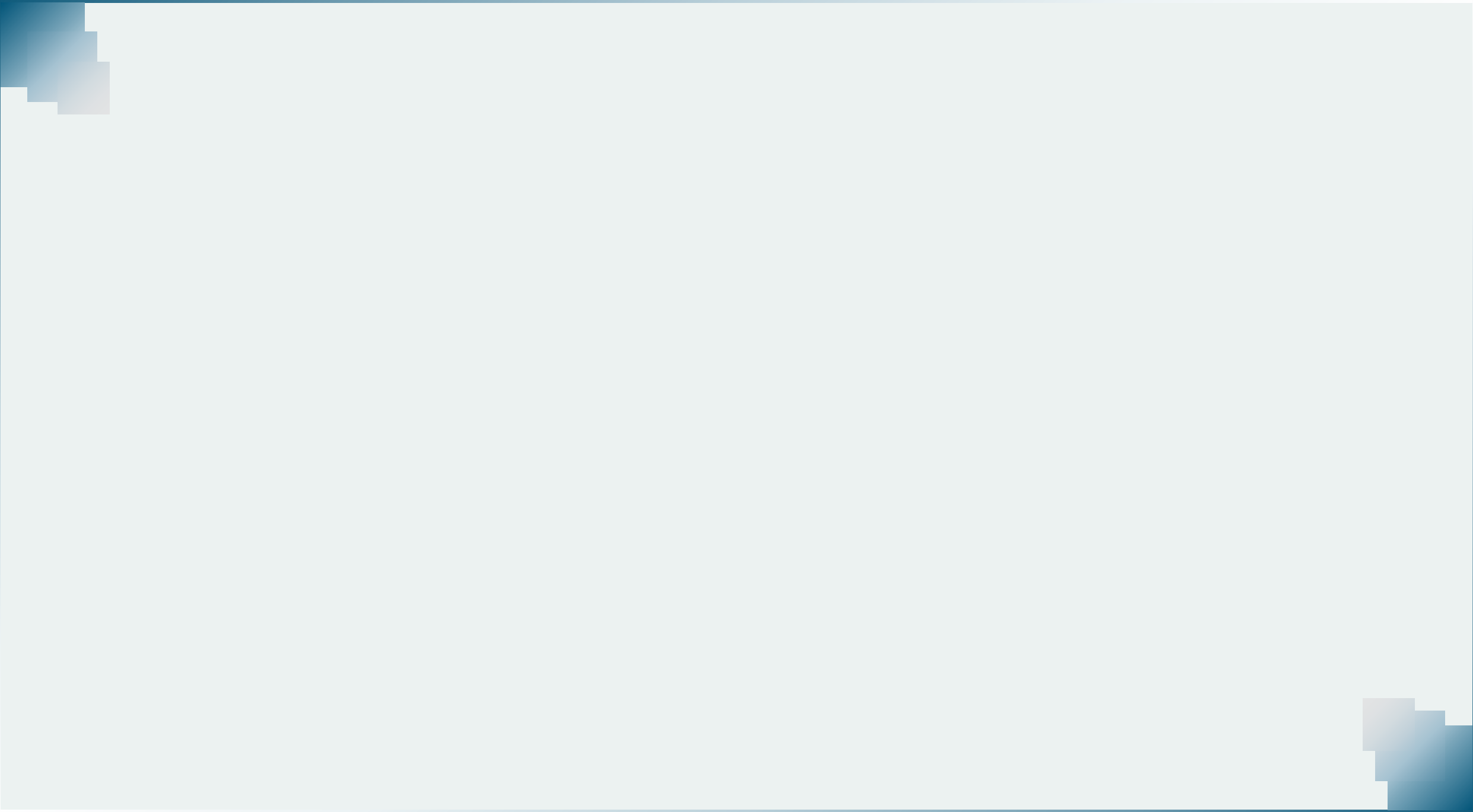
Conclusioni

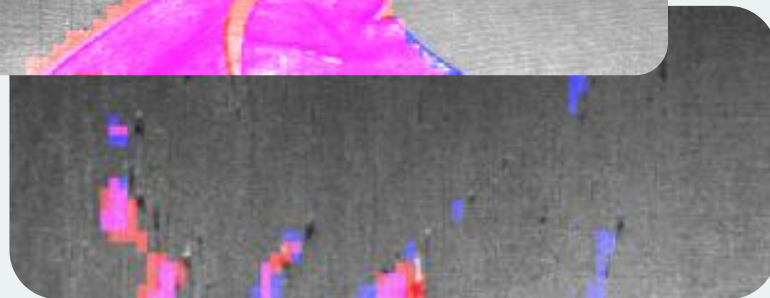
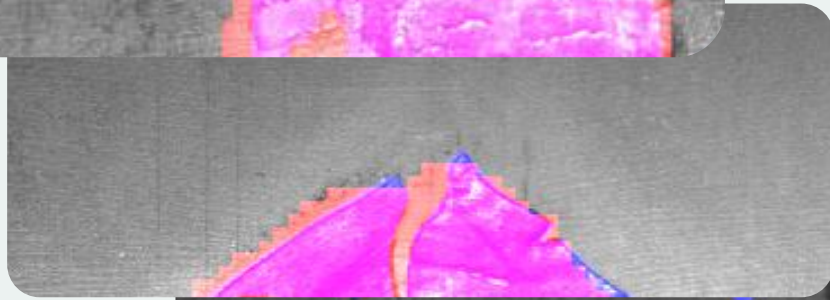
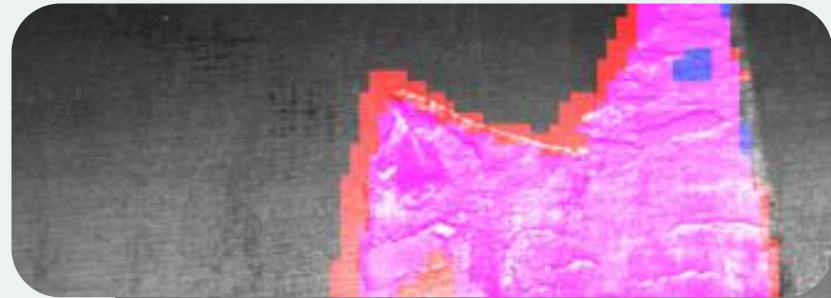
Risultati:

- > **Patch-Ganomaly** e **TL-Ganomaly** risolvono il problema di segmentazione in maniera più **efficace** rispetto ai modelli di riferimento
- > **TL-Ganomaly** si contraddistingue per la sua **precisione**
- > **Patch-Ganomaly** fornisce una **migliore segmentazione** della regione anomala
- > Il post-processing apportato con il modello **Conv-Processing** migliora le performance dei modelli proposti

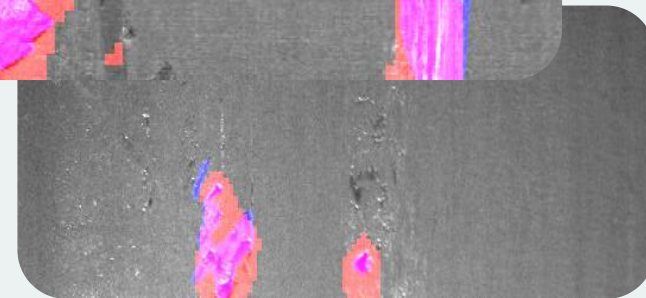
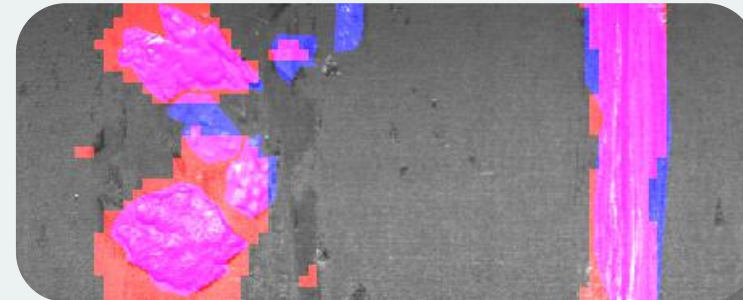


Grazie per l'attenzione





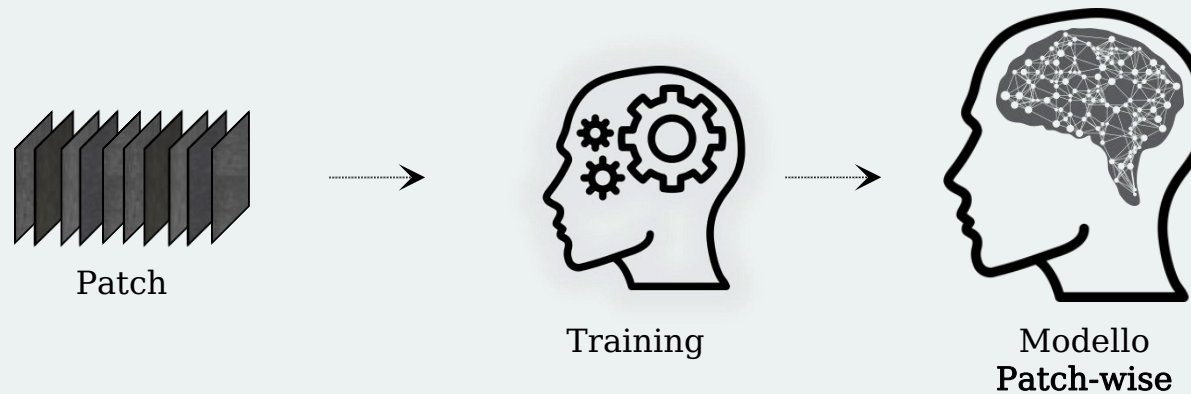
Grazie per l'attenzione



Patch-wise Training

Addestramento:

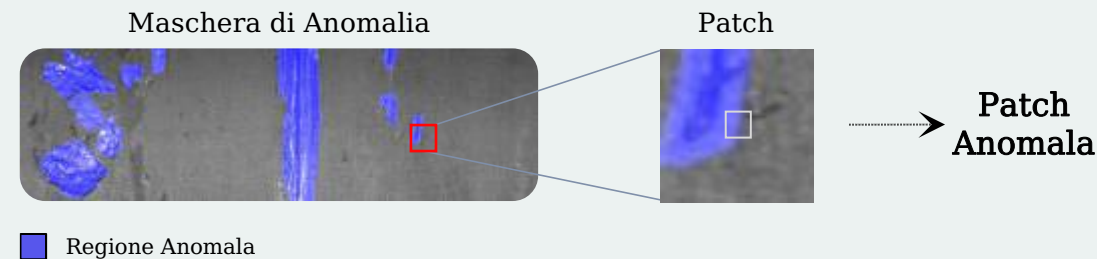
- > Apprendimento basato su un training set di **patch**
- > La **patch** è definita **anomala** se il suo centro fa parte della regione anomala
- > La **patch normale** presenta il suo centro privo di anomalie



Patch-wise Training

Addestramento:

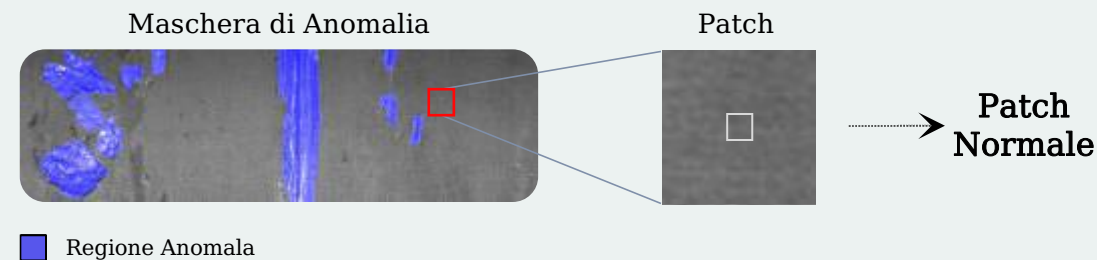
- > Apprendimento basato su un training set di **patch**
- > La **patch** è definita **anomala** se il suo centro fa parte della regione anomala
- > La **patch normale** presenta il suo centro privo di anomalie




Patch-wise Training

Addestramento:


- > Apprendimento basato su un training set di **patch**
- > La **patch** è definita **anomala** se il suo centro fa parte della regione anomala
- > La **patch normale** presenta il suo centro privo di anomalie



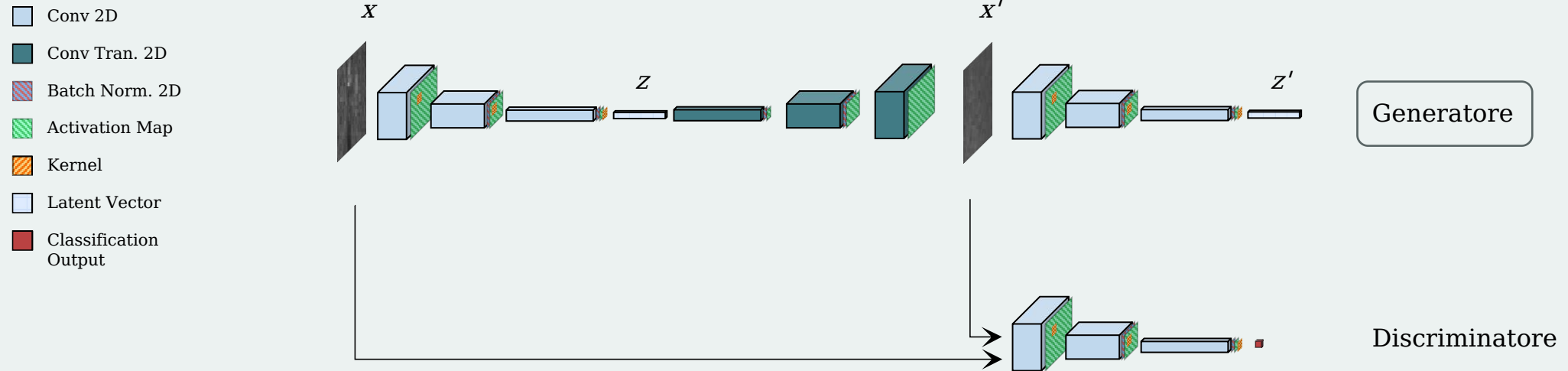


Ganomaly:

Addestramento

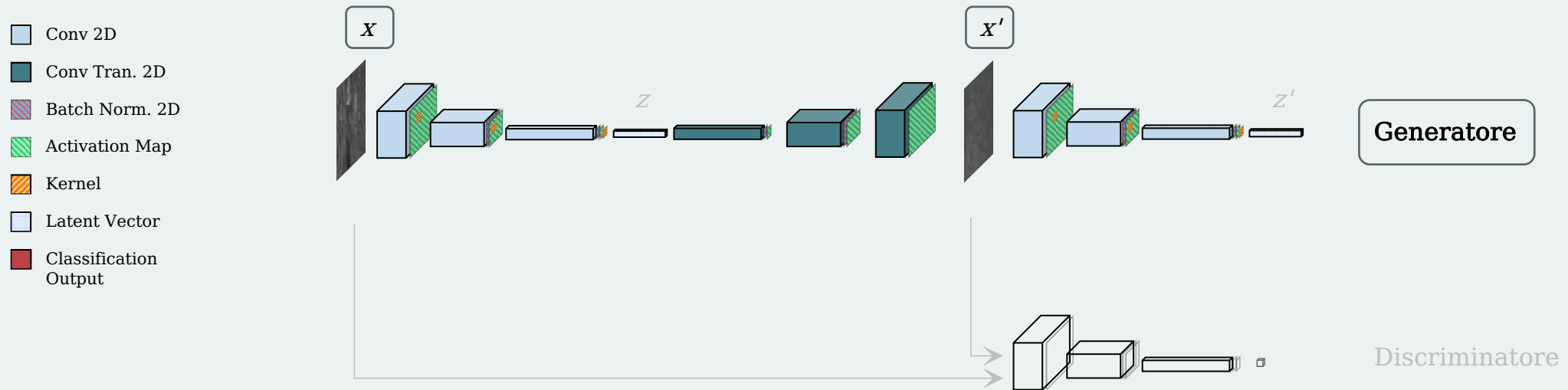


Addestramento



- > La **funzione obiettivo** (objective function) del **Generatore** tiene conto di tre **funzioni di perdita** (loss functions):
 - Errore di Ricostruzione
 - Errore degli Encoding
 - Errore delle Feature

Addestramento










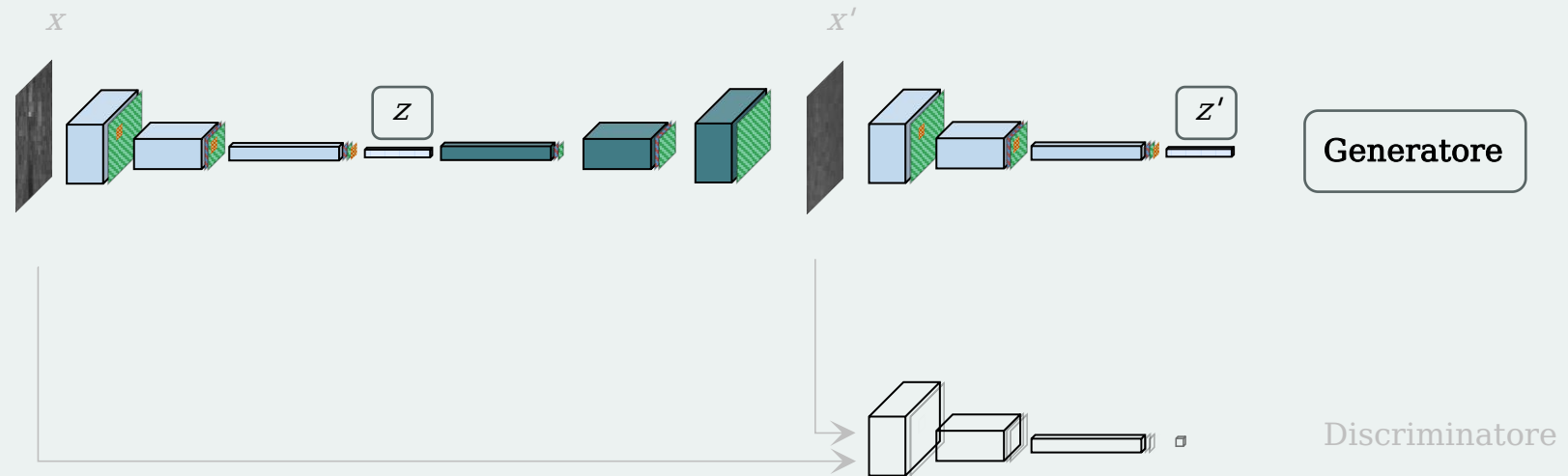
- > La **funzione obiettivo** (objective function) del **Generatore** tiene conto di tre **funzioni di perdita** (loss functions):

- Errore di Ricostruzione
- Errore degli Encoding
- Errore delle Feature

$$L_{con} = \| x - x' \|_1$$

Addestramento

-  Conv 2D
-  Conv Tran. 2D
-  Batch Norm. 2D
-  Activation Map
-  Kernel
-  Latent Vector
-  Classification Output










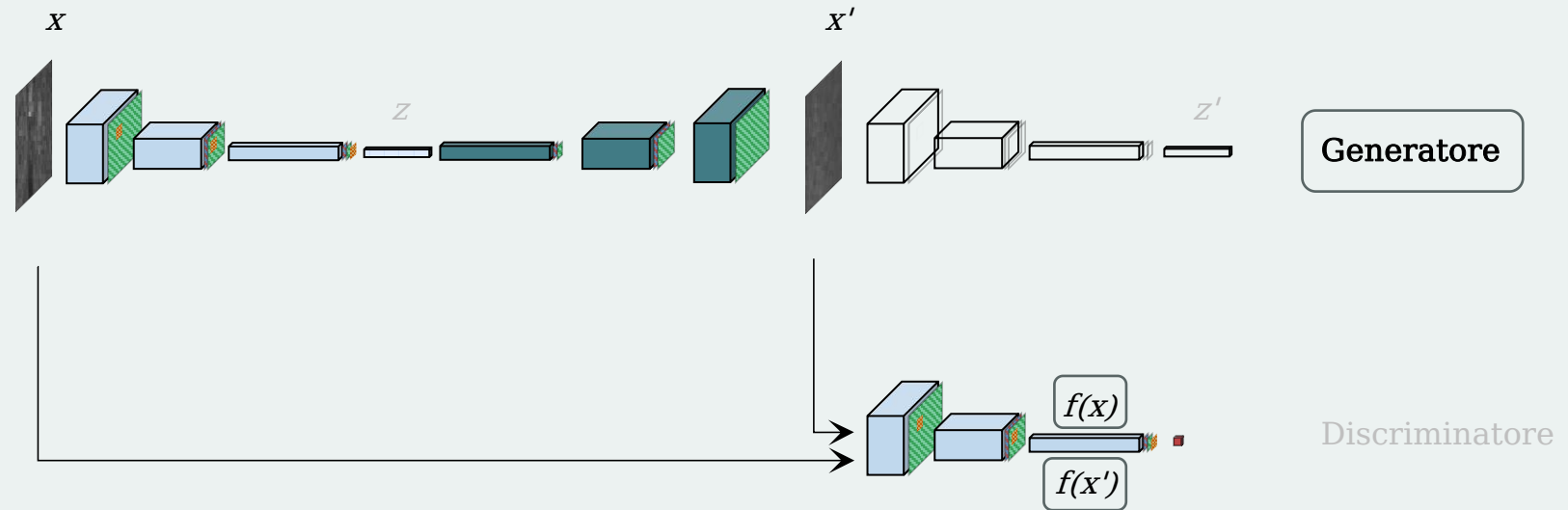
- > La **funzione obiettivo** (objective function) del **Generatore** tiene conto di tre **funzioni di perdita** (loss functions):

- Errore di Ricostruzione
- Errore degli Encoding
- Errore delle Feature

$$L_{enc} = \| z - z' \|_2$$

Addestramento

-  Conv 2D
-  Conv Tran. 2D
-  Batch Norm. 2D
-  Activation Map
-  Kernel
-  Latent Vector
-  Classification Output

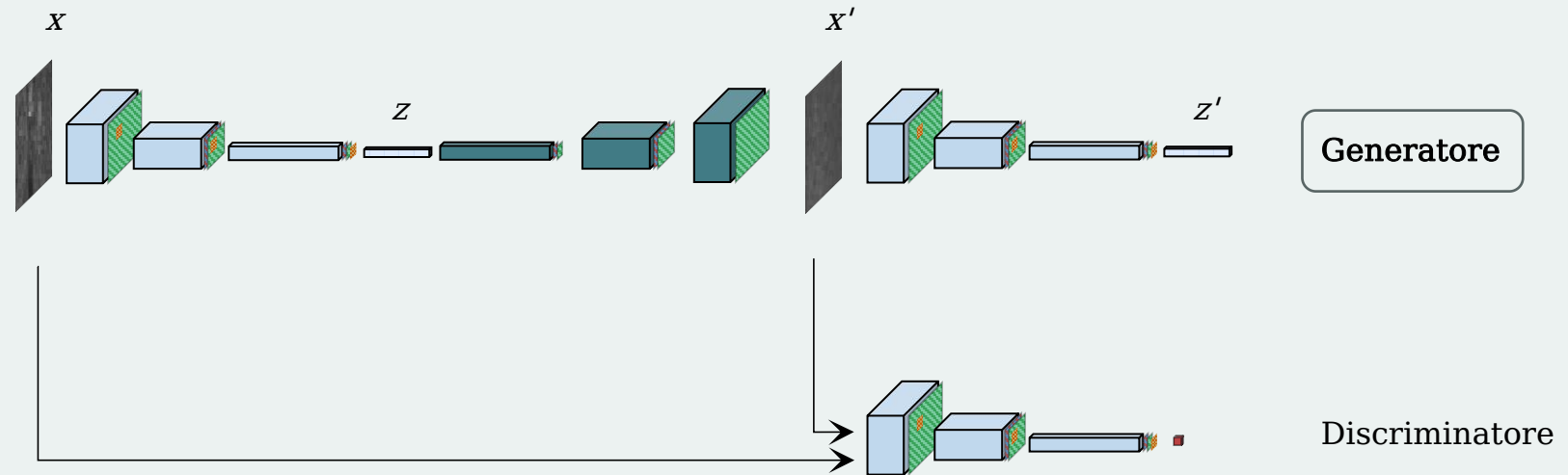
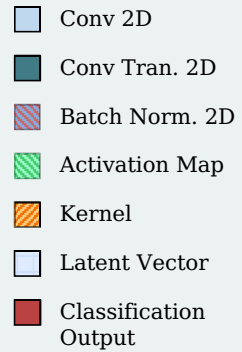


> La **funzione obiettivo** (objective function) del **Generatore** tiene conto di tre **funzioni di perdita** (loss functions):

- Errore di Ricostruzione
- Errore degli Encoding
- Errore delle Feature

$$L_{adv} = \|f(x) - f(x')\|_2$$

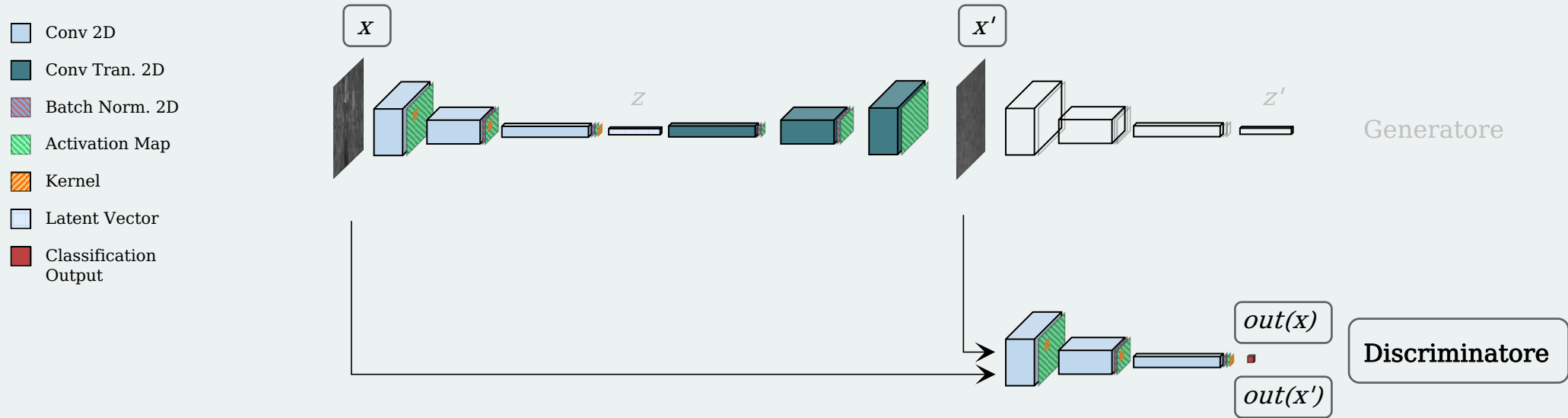
Addestramento



- > La **funzione obiettivo** (objective function) del **Generatore** è la **somma pesata** delle tre loss functions:

$$L_{gen} = w_{con}L_{con} + w_{enc}L_{enc} + w_{adv}L_{adv}$$

Addestramento

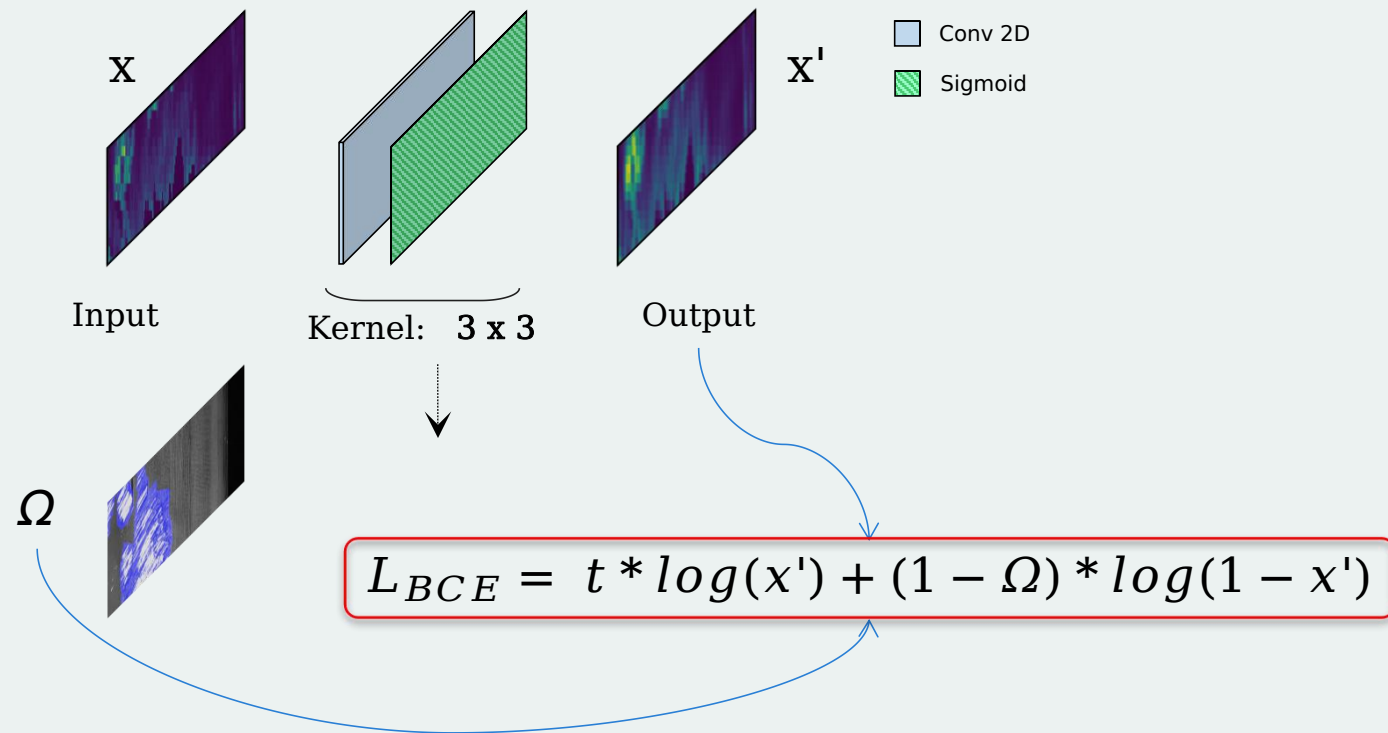


- > La **funzione obiettivo** (objective function) del **Discriminatore** tiene conto della seguente **funzione di perdita** (loss function):

- Binary Cross Entropy

$$L_{Dis} = \frac{L_{BCE}(1, out(x)) + L_{BCE}(0, out(x'))}{2}$$

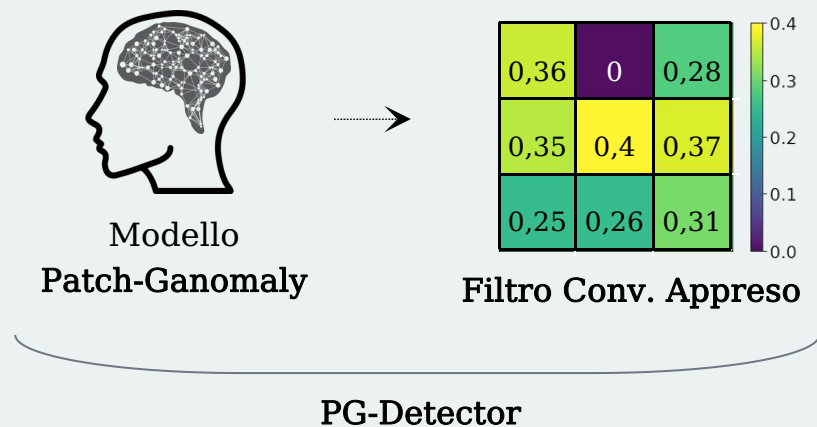
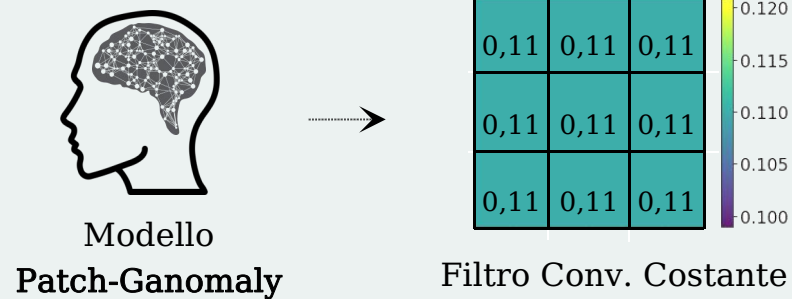
Conv-Processing



Post-Processing

Primo confronto tra:

- > Patch-Ganomaly con **filtro convoluzionale costante** in post-processing
- > Patch-Ganomaly con **filtro convoluzionale** fornito dal modello **Conv-Processing (PG-Detector)**



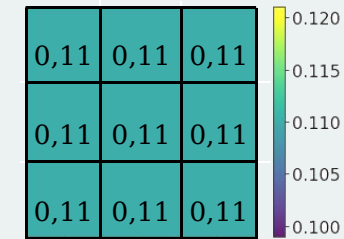
Post-Processing

Secondo confronto tra:

- > TL-Ganomaly con **filtro convoluzionale costante** in post-processing
- > TL-Ganomaly con **filtro convoluzionale** fornito dal modello **Conv-Processing (TL-Detector)**



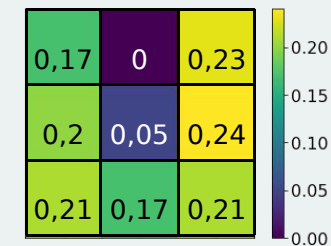
Modello
TL-Ganomaly



Filtro Conv. Costante



Modello
TL-Ganomaly



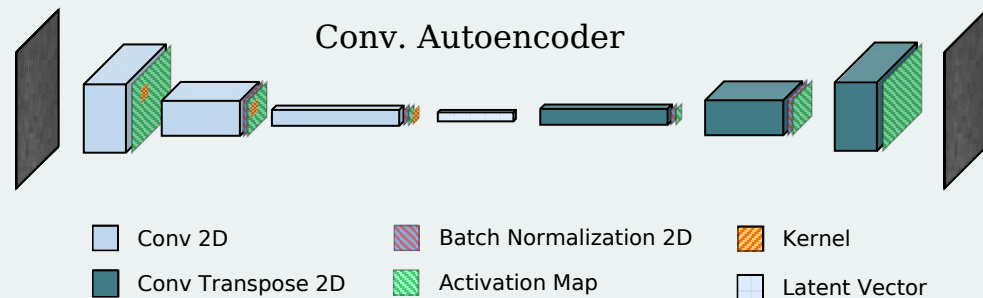
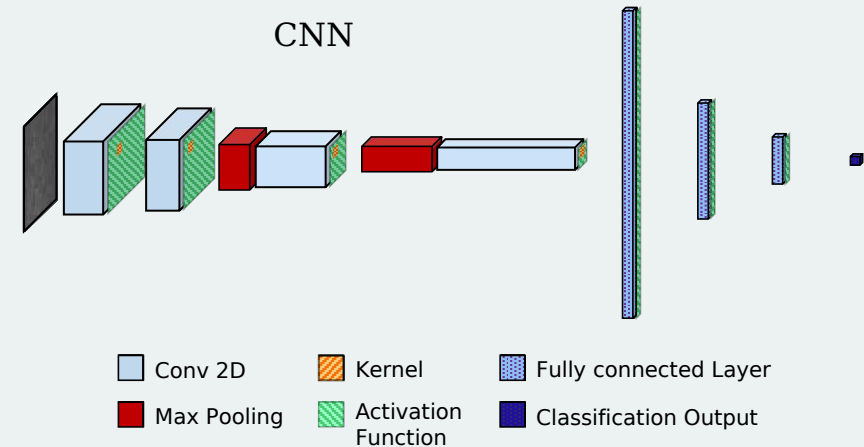
Filtro Conv. Appreso

TL-Detector

Modelli di Riferimento

Il modelli proposti, **Patch-Ganomaly** e **TL-Ganomaly**, vengono confrontati con due modelli di riferimento:

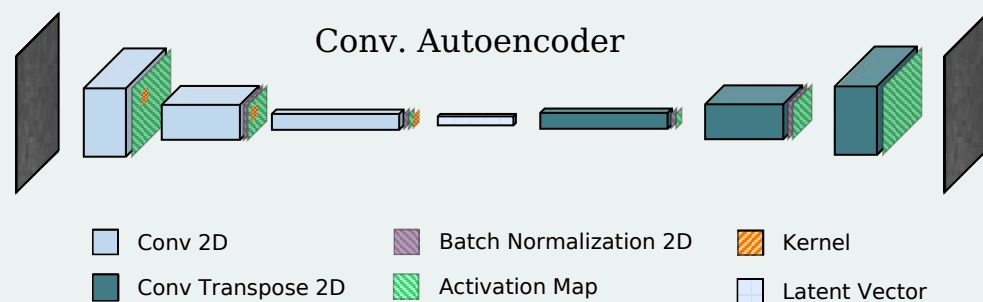
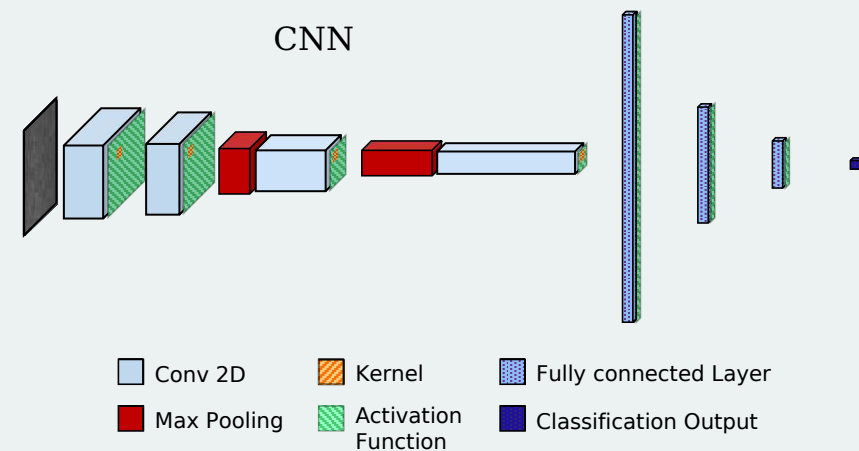
- > Modello **Supervionato**: Rete Neurale Convoluzionale (CNN) fully-connected
- > Modello **Semi-Supervisionato**: Convolutional Autoencoder
- > Addestrati sfruttando la tecnica del **patch-wise training**
- > Training set di 30.000 patch



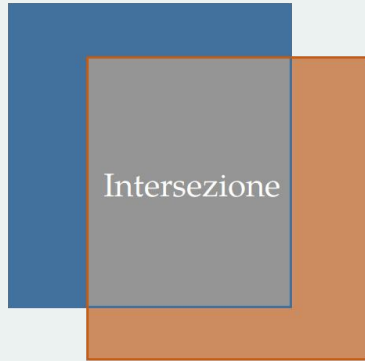
Modelli di Riferimento

Anomaly Score:

- > CNN: **binary cross entropy**
- > Autoencoder: **errore di ricostruzione**

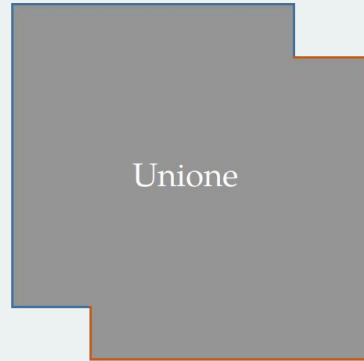


Area Anomala Predetta



Area Anomala Reale

Area Anomala Predetta



Area Anomala Reale

$$\mathbf{IoU} = \frac{P \cap R}{P \cup R} = \frac{TP}{TP + FP + FN}$$

Più peso ai **Veri Positivi**



$$\mathbf{Dice} = 2 \frac{Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN}$$

Immagini Test

Immagine Originale

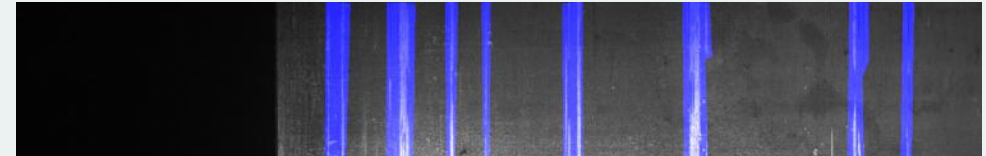


Maschera di Anomalia

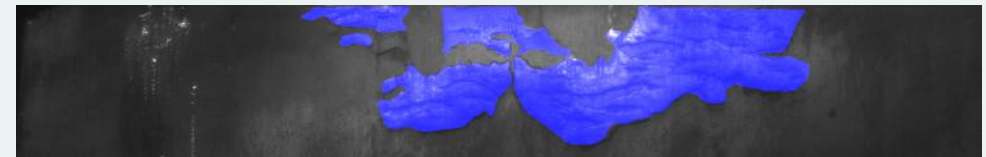
Difetto di tipo 1



Difetto di tipo 2



Difetto di tipo 3



■ Pixel Anomali



Difetti di tipo 1

TL-Ganomaly Conv. Costante

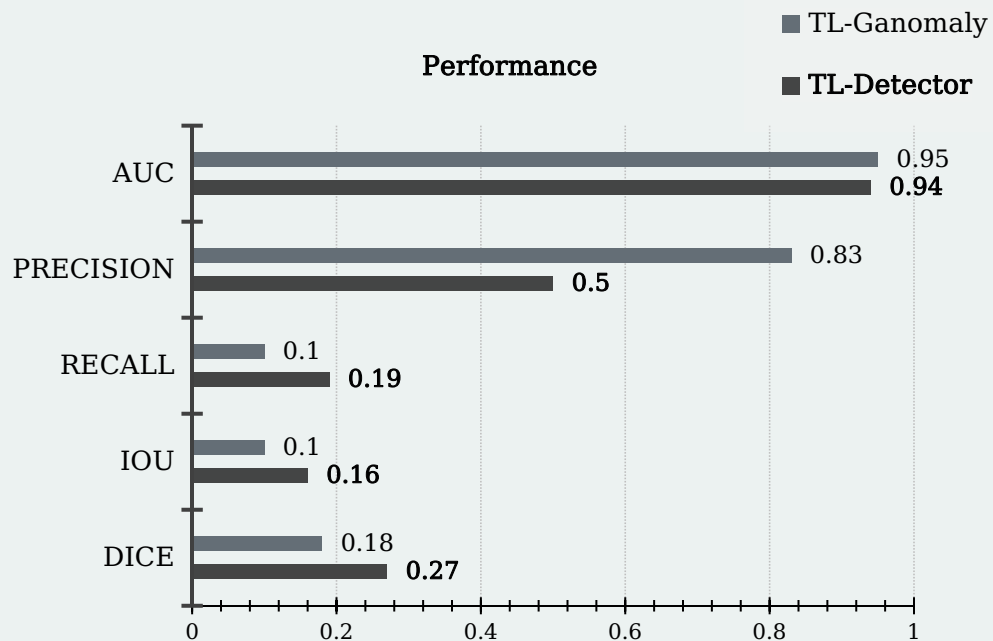


TL-Detector



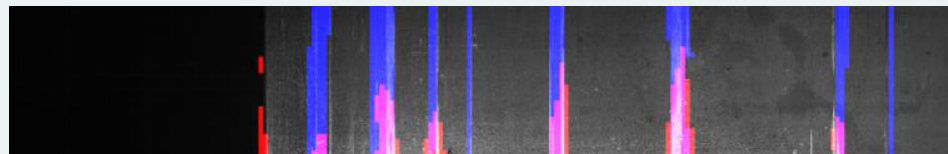
-  Pixel Anomali
-  Pixel Riconosciuti Anomali

Performance

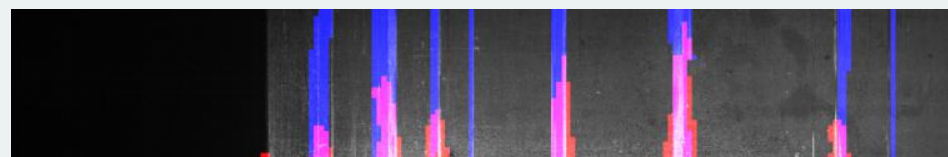


Difetti di tipo 2

TL-Ganomaly Conv. Costante

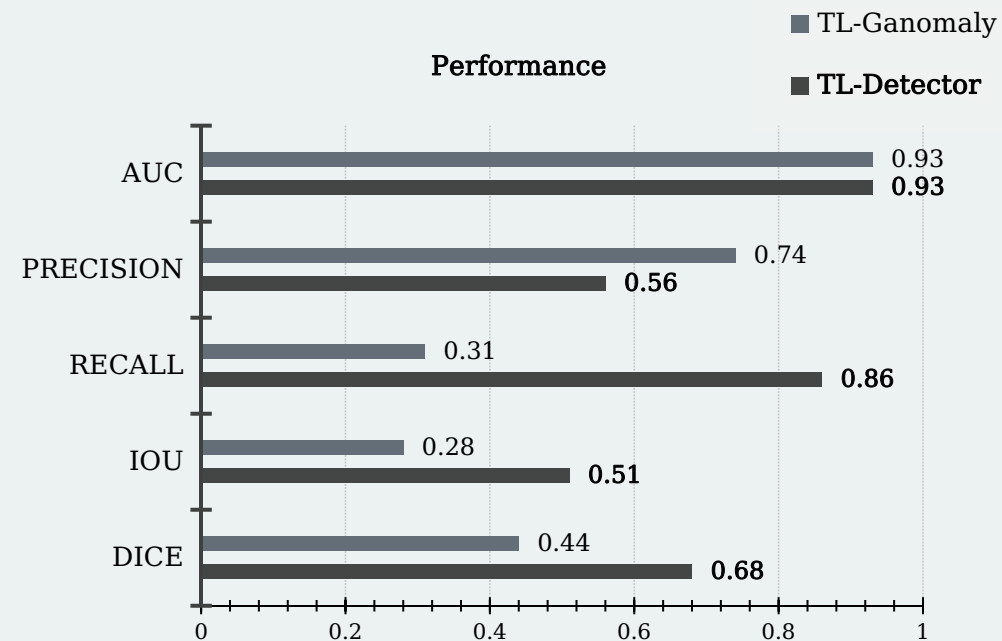


TL-Detector



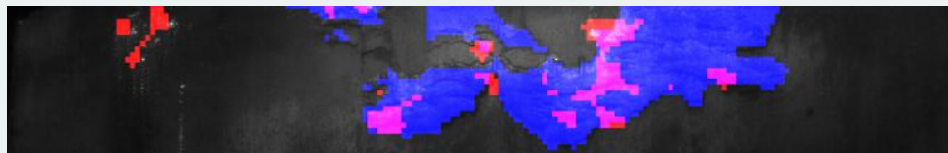
- Pixel Anomali
- Pixel Riconosciuti Anomali

Performance

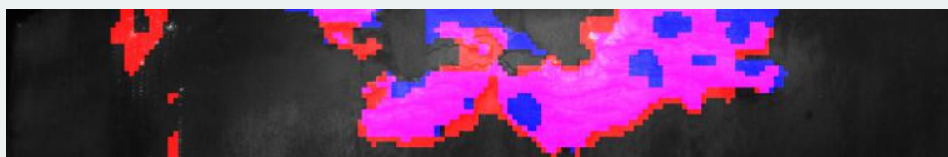


Difetti di tipo 3

TL-Ganomaly Conv. Costante



TL-Detector



- Pixel Anomali
- Pixel Riconosciuti Anomali

Performance

