



UNIVERSITÀ POLITECNICA DELLE MARCHE

FACOLTÀ DI INGEGNERIA

Corso di Laurea triennale in *Ingegneria Informatica e dell'Automazione*

Sviluppo di un modello GAN per la generazione di immagini e relativa segmentazione

Develop of a GAN model for image generation with related segmentations

Relatore:

Prof. Adriano Mancini

Laureando:

Massimiliano Biancucci

Prefazione

Il mio percorso nel campo dell'intelligenza artificiale è iniziato diversi anni fà, alle superiori per l'esattezza, dove sentii per la prima volta parlare di reti neurali, ad un corso pomeridiano voluto dal prof. Roberto Lulli il quale mi ha mostrato per primo questo affascinante campo di ricerca.

Ho svolto durante il mio percorso di studi diversi progetti incentrati su questa tematica, partendo da semplici reti neurali, e confrontandomi con progetti sempre più complessi fino ad arrivare ai modelli generativi basati sull'architettura GAN (Generative Adversarial Network), del quale in questa tesi proporò una variante.

Lo scopo di questa tesi è quello di proporre un rimedio ad uno dei grandi problemi che affligge le aziende che si occupano di addestrare modelli neurali per la segmentazione di immagini, ovvero la difficoltà nel reperire immagini annotate ed gli elevatissimi costi per la realizzazione di un dataset ad hoc per il proprio task, problema con il quale mi sono dovuto confrontare io stesso nell'ultimo anno come sviluppatore presso l'azienda Cloe.ai.

Indice

Elenco delle figure

Capitolo 1

Introduzione

1.1 Il problema

Oggi nello scenario industriale e di ricerca, per l'addestramento di modelli neurali una delle principali difficoltà è la mancanza di dataset adatti al proprio task, i quali per la segmentazione di immagini possono essere particolarmente difficili da reperire o realizzare, in quanto richiedono un elevato numero di immagini annotate. Più il task è complesso più il dataset necessario per addestrare un modello sarà grande, e potenzialmente rendendo ogni singolo esempio più difficile da realizzare, richiedendo una quantità di ore uomo molto elevata.

Uno dei task in cui oggi i modelli neurali trovano un vasto impiego è la segmentazione dei difetti, principalmente in campo industriale, i quali sono utilizzati per verificare automaticamente la qualità di un prodotto o di un semilavorato, per ridurre i costi e i tempi del controllo manuale, il quale è spesso prone ad errori in quanto l'uomo è soggetto a stanchezza o a mancanza di concentrazione, al contrario un modello con un'elevata accuratezza è in grado di riconoscere un difetto con prestazioni costanti.

La problematica di avere un modello con elevata accuratezza per task di questo tipo è relativa alla preparazione dei dati. Molti task richiedono molta concentrazione da parte dell'annotatore in quanto non sempre i difetti sono ben visibili e dunque ogni singolo esempio deve essere controllato attentamente anche da più persone, e questo rende il processo di annotazione molto lungo e costoso. Spesso in oltre anche dopo diversi controlli alcuni difetti possono comunque sfuggire al processo di annotazione.

1.2 L'idea

L'obiettivo di questo progetto è proporre una soluzione a questo problema, o almeno mitigare la sua gravità. Per far ciò si propone di realizzare un architettura in grado di generare un oggetto in un'immagine persistente, specificando la maschera di quest'ultimo.

Tale architettura deve essere in grado di prendere in input un'immagine e una maschera, e generare l'oggetto all'interno della maschera, lasciando il più possibile invariata l'immagine al di

fuori dell'area d'interesse. Il modello dovrebbe essere in grado di generare un oggetto all'interno dell'area obbiettivo seguendo la distribuzione morfologica degli oggetti presenti nel dataset di addestramento.

Tale tecnica è principalmente pensata per tutti quei casi in cui non è semplice o possibile realizzare sinteticamente un oggetto a partire da metodi deterministici, ad esempio per la segmentazione dei difetti, i quali non possono essere copiati e incollati in una nuova immagine senza che sia presente un evidente linea di separazione tra il difetto e il resto dell'immagine, e quindi non possono essere generati con una tecnica di tipo copy paste, senza introdurre artefatti, che potrebbero portare ad un peggioramento delle prestazioni del modello, in quanto questo associerebbe il difetto alla linea di separazione.

In questo progetto il dataset utilizzato è Severstal steel defect detection, il quale presenta campioni di acciaio più tosto simili, ma in generale questa tecnica potrebbe essere utilizzata anche per casi più complessi come il rilevamento di danni su veicoli o edifici, dove la morfologia del difetto è comune, ma la morfologia, il colore e la texture dell'oggetto su cui deve essere applicato sono differenti dal dataset di addestramento, in quest'ultimo scenario infatti non è possibile utilizzare una tecnica di tipo copy paste, in quanto il difetto non potrebbe essere applicato in modo coerente con l'immagine base.

Capitolo 2

Dataset e architetture base utilizzate

2.1 Il Dataset: Severstal steel defect detection

2.2 Lama:

2.3 Stylegan2: