

UNIVERSITÀ DEGLI STUDI DI NAPOLI “FEDERICO II”

**DIPARTIMENTO DI INGEGNERIA ELETTRICA
E TECNOLOGIE DELL'INFORMAZIONE**



CORSO DI LAUREA TRIENNALE IN INFORMATICA

Titolo

Sviluppo di una pipeline generativa per la creazione di immagini per il settore della moda

Candidato

Gianluigi Mercurio

N86003834

Relatore

Prof. Porfirio Tramontana

Anno Accademico 2023/2024

Introduzione

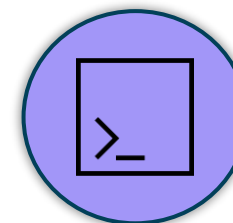
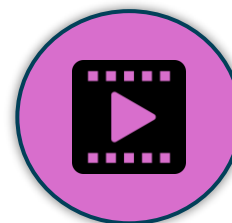
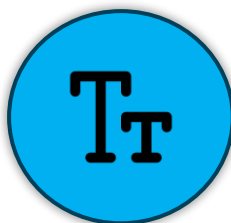


Negli ultimi anni, l'**Intelligenza Artificiale** ha registrato progressi significativi rendendo le macchine capaci di simulare processi cognitivi umani, riconoscere schemi complessi e prendere decisioni autonome analizzando grandi quantità di dati.

I fattori chiave che ne hanno favorito l'espansione sono:

- **Applicabilità in moltissimi settori**
- **Utilizzo come strumento di assistenza**
- **Nuove frontiere di creatività**
- **Innovazione del pensiero e del ragionamento**

Tra i numerosi approcci all'Intelligenza Artificiale troviamo l'**IA Generativa**, un ramo avanzato che si differenzia dagli altri in quanto **focalizzato sulla creazione di contenuti nuovi e originali**, piuttosto che sull'analisi e la classificazione dei dati.



IA nel settore della moda

Tra i settori in cui l'IA generativa trova applicazione, non poteva mancare quello della moda, storicamente all'avanguardia nell' adottare nuove tecnologie per innovare il design e la produzione.

L'attività di tirocinio svolto presso l'azienda **Kineton** mi ha permesso di lavorare ad un progetto innovativo che ne indagasse le applicazioni.

Tale attività ha portato alla **realizzazione di una pipeline generativa** per la creazione di immagini contenenti specifici capi d'abbigliamento utilizzando modelli di Intelligenza Artificiale.



kineton
High-Tech Humanity



Pipeline generative

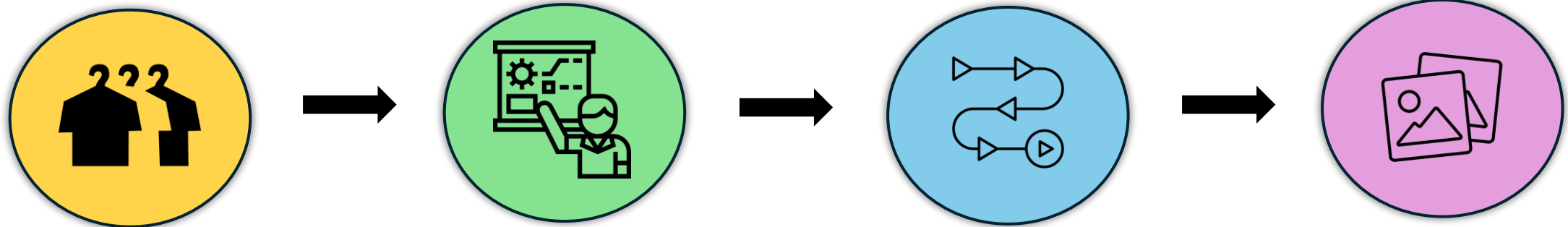
Cosa sono?

Le pipeline generative sono processi strutturati che permettono di creare nuovi contenuti come immagini e video, a partire dalla richiesta di un utente e sfruttando modelli addestrati di IA che ne determinano le capacità generative.

Come si sviluppano?



1. Raccolta dati dei capi selezionati
2. Addestramento modelli generativi
3. Definizione struttura della pipeline
4. Generazione delle immagini.



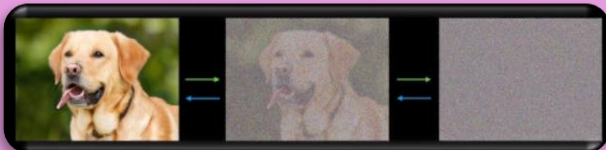
Pipeline generative

Cosa sono i modelli? Cosa si intende per addestramento?

Modelli di Diffusione

I Modelli di Diffusione sono una tipologia di modelli di IA che apprendono il processo di conversione di un'immagine in rumore visivo [**diffusione diretta**] .

Tali modelli sono in grado di invertire il processo per generare un'immagine, partendo dal solo rumore e raffinandolo fino a ottenere un'immagine realistica [**diffusione inversa**] .



Addestramento

Poiché addestrati su enormi volumi di dati, i modelli sono in grado di svolgere compiti generici ed essere applicati a molteplici contesti.

Per rendere tali modelli in grado di svolgere compiti più specifici, si ricorre a dei processi di **perfezionamento** che danno vita ad un nuovo modello più raffinato, addestrato su un insieme di dati più specifici.

Low-Rank Adaption (LoRA)

I modelli *LoRA* sono il **risultato dell'attività di addestramento** applicata ai Modelli di Diffusione.

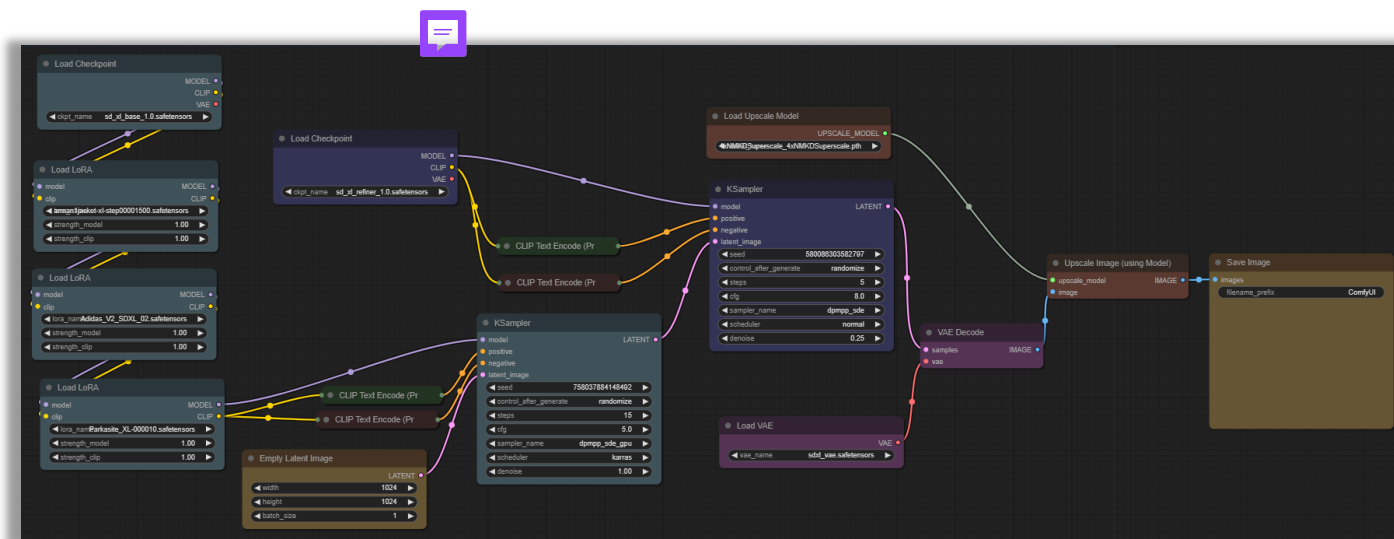
Tali modelli sono estremamente efficienti nei casi in cui si voglia generare immagini che includono stili artistici particolari e contenuti realmente esistenti come i volti di persone o particolari capi d'abbigliamento.

Pipeline sviluppata

Nella pipeline generativa sviluppata con **ComfyUI**, sono state combinate le capacità generative dei modelli di diffusione di base e le conoscenze approfondite dei modelli *LoRA* da essi ottenuti, per creare immagini contenenti elementi specifici, come i capi d'abbigliamento nei loro dettagli, in contesti generici, simulando veri e propri shooting.

ComfyUI
sviluppato da *comfyanonymous*

- Open source
- Interfaccia Grafica intuitiva
- Struttura a nodi
- Ampia personalizzazione
- API

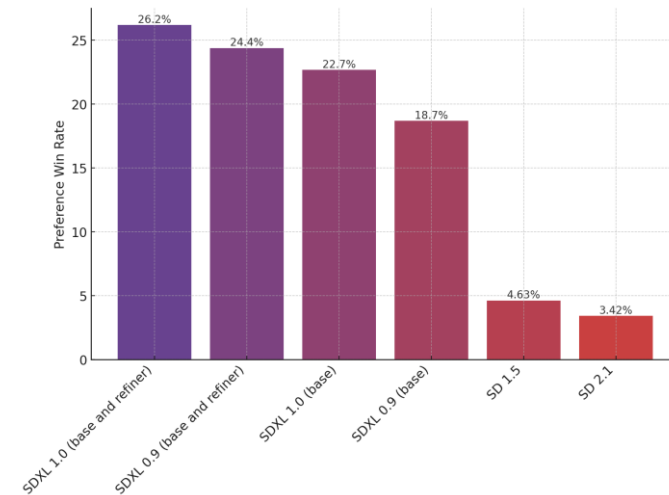


Struttura della pipeline sviluppata

Modelli Utilizzati

La capacità generativa della pipeline è stata determinata dai modelli di:

- *Stable Diffusion XL 1.0 base*, come modello di base
- *Stable Diffusion XL 1.0 refiner*, in combinazione con la versione base
- *LoRA arman1jacket_XL*, per generare immagini contenenti giacche armani
- *LoRA Parkasite_XL*, per generare immagini contenenti giubbotti parkasite
- *LoRA Adidas_V2_XL*, per generare immagini contenenti tute adidas
- *4xNMKD_UPSCALER* per produrre immagini in risoluzione 4k



Arman1jacket



Parkasite



Adidas_V2

Genai software

È stato sviluppato un **applicativo software** dotato di interfaccia grafica che consenta agli utenti di interagire efficacemente con la pipeline sviluppata.

Il software è stato sviluppato utilizzando:

- Il linguaggio **Python**
- La libreria **Gradio**
- L'**API** fornita da *ComfyUI*

Tale applicativo consente all'utente di:

1. **Inserire i prompt positivi e negativi**
2. **Visualizzare i capi d'abbigliamento**
3. **Visualizzare e scaricare il risultato**
4. **Visualizzare degli esempi di generazione**
5. **Personalizzare le impostazioni**

Generator

Settings

StableDiffusion - LoRA image generator.

Compile both positive and negative prompt to tell AI Model what you want to generate.

Selecting an item from gallery will automatically insert the correct token in positive prompt.

Click "Generate Image" button to perform the generation request.

Select Item

Positive prompt

GIORGIO ARMANI, full body photo of ginger male model wearing dark gray jacket, arman1jacket:1.2, realistic face, extremely high quality RAW photograph, detailed background, intricate, Exquisite details and textures, sharp focus, highly detailed, ultra detailed photograph, warm lighting, 4k, high resolution, detailed skin, detailed eyes, 8k UHD, DSLR, high quality, film grain, Fujifilm XT3, private party on yacht, summer days

Negative prompt

(worst quality, greyscale), ac_neg2, zip2d_neg, ziprealism_neg, watermark, username, signature, text, bad anatomy, bad hands, text, error, missing fingers, extra digit, fewer digits, cropped, jpeg artifacts, bad feet, extra fingers, mutated hands, poorly drawn hands, bad proportions, extra limbs, disfigured, bad anatomy, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, mutated hands, fused fingers, too many fingers, long neck

Generated image:



Examples:

Generated image:	Positive prompt	Negative prompt
	GIORGIO ARMANI, full body photo of ginger male model wearing dark gray jacket, arman1jacket:1.2, realistic face, extremely high quality RAW photograph, detailed background, intricate, Exquisite details and textures, sharp focus, highly detailed, ultra detailed photograph, warm lighting, 4k, high resolution, detailed skin, detailed eyes, 8k UHD, DSLR, high quality, film grain, Fujifilm XT3, private party on yacht, summer days	(worst quality, greyscale), ac_neg2, zip2d_neg, ziprealism_neg, watermark, username, signature, text, bad anatomy, bad hands, text, error, missing fingers, extra digit, fewer digits, cropped, jpeg artifacts, bad feet, extra fingers, mutated hands, poorly drawn hands, bad proportions, extra limbs, disfigured, bad anatomy, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, mutated hands, fused fingers, too many fingers, long neck

Pages: 1 2 3 4

Generate image

Clear

Genai dataset generator software

È stato sviluppato un **applicativo software** dotato di interfaccia grafica che consenta di sviluppare *dataset* in maniera automatizzata.

Il software è stato sviluppato utilizzando:

- Il linguaggio **Python**
- La libreria **Gradio**
- Il modello di IA **BlipImageCaptioning**, sviluppato da *Salesforce*

Tale applicativo offre all'utente le seguenti funzioni:

1. **Caricare una cartella di immagini**
2. **Ridimensionare le immagini**
3. **Specificare il termine identificativo** detto *token*
4. **Etichettatura manuale**
5. **Etichettatura automatizzata** sfruttando il modello IA
6. **Acquisire i risultati** elaborati e tabulati

Dataset generator

Upload your data directory and select the resize options.

Generate and download the output folder containing the processed images.

Select the folder containing images to be processed

Drop File Here

- Or -

Click to Upload

Resize images

Width

Select width in pixels for resize

0

Height

Select height in pixels for resize

0

Set necessary token to perform dataset generation

Use a keyword.

Turn on to generate a text file for each dataset image with token label. Check 'Image Captioning' section for more. Turn off to deactivate.

☒ Enable Labelling

Image Captioning.

*** READ BEFORE CONTINUING ***

Set this options only if 'Labelling' is enable.

Activate to use AI Model to automate image captioning.

☒ Use AI for feature extraction

Caption for every image

Type the caption to set to every image. Leave blank to not use.

Generate dataset

Generated dataset

Download file

Validazione pipeline

Si è scelto di validare la pipeline generativa sviluppata attraverso una validazione qualitativa per garantire che le immagini prodotte soddisfino i criteri di qualità e coerenza definiti.

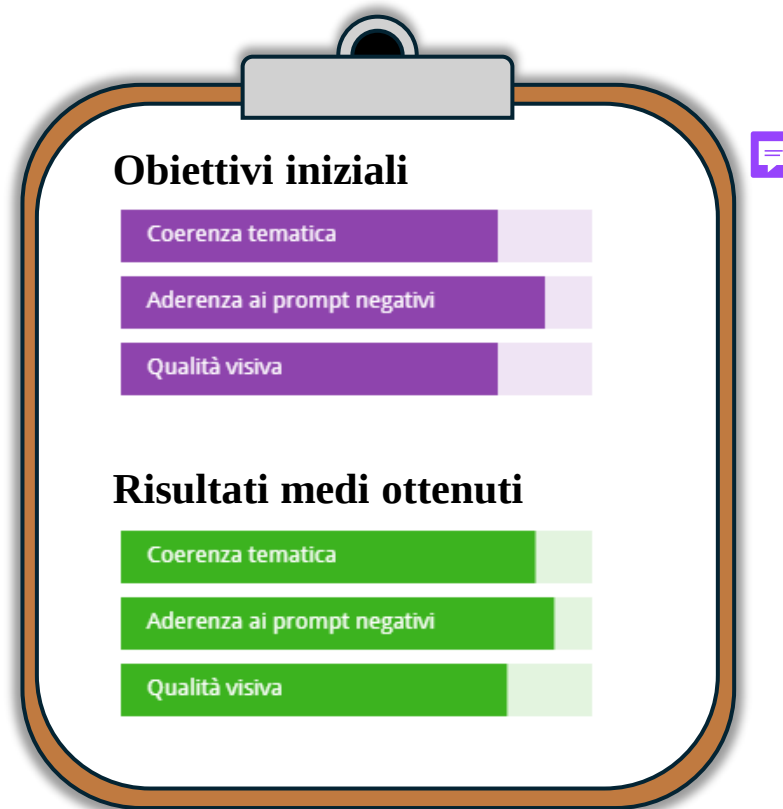
Obiettivi della Validazione

- Valutare la coerenza tematica
- Verificare l'aderenza ai prompt negativi
- Misurare la qualità visiva



Metodologia

1. Selezione del campione
2. Valutazione da parte degli utenti
3. Raccolta dei punteggi
4. Calcolo dei risultati
5. Analisi dei risultati



Obiettivi raggiunti



L'elaborato di tesi ha esplorato:

- La **creazione e la validazione di una *pipeline* generativa** capace di produrre immagini coerenti e di buona qualità.
- Lo **sviluppo di un applicativo software** che fornisce un'interfaccia grafica agli utenti per interagire con la *pipeline* sviluppata;
- Lo **sviluppo di un applicativo software** con cui automatizzare la generazione di dataset sfruttando un modello di IA dedicato.

L'adozione di un approccio basato su modelli di Intelligenza Artificiale ha permesso di sviluppare strumenti robusti e versatili a supporto dei designer, mostrando come il loro utilizzo possa rivoluzionare i processi creativo legato al settore della moda.

Esplorando tali implicazioni risulta chiaro che strumenti di questo tipo possono trovare spazio in tantissimi ambiti e settori industriali, aprendo l'orizzonte a nuove frontiere di creatività.

Grazie per l'attenzione.