



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI INFORMATICA, SCIENZE E INGEGNERIA

CORSO DI LAUREA IN  
INFORMATICA

# INTELLIGENZA ARTIFICIALE PER CREARE FUMETTI: ESPERIMENTI E LIMITAZIONI

**Relatore**

**Prof. Angelo Di Iorio**

**Presentata da**

**Martina Ritrovato**

---

**Sessione I**

**Anno Accademico 2023/2024**

*Hear now, lonely child*

*Whatever grief has crossed your way*

*Soon the sun burns out*

*Another turn, another page*

- MELLIE METEORS, "LULLABY FOR THE LOST"

# SOMMARIO

<b>1. INTRODUZIONE</b>	<b>5</b>
<b>2. STATO DELL'ARTE</b>	<b>7</b>
<b>2.1 Progetti editoriali</b>	<b>7</b>
2.1.1 Black Jack	7
2.1.2 Fastwalkers	9
2.1.3 The Abolition of Man	10
<b>2.2 Strumenti per generazione di fumetti con IA</b>	<b>13</b>
2.2.1 Plugger.ai	13
2.2.2 Fantoons	14
2.2.3 Storywizard.ai	15
<b>2.3 Ricerche accademiche</b>	<b>16</b>
2.3.1 Colorazione di manga	16
2.3.2 Simulazione del processo creativo di un manga	17
2.3.3 Traduzione di fotografie in disegni in stile manga	18
<b>3. ESPERIMENTO: OBIETTIVO E STRUMENTI UTILIZZATI</b>	<b>19</b>
<b>3.1 AI Comic Factory</b>	<b>19</b>
<b>3.2 Mage Space</b>	<b>21</b>
<b>3.3 AI Comic Generator</b>	<b>24</b>
<b>4. ESPERIMENTO: METODOLOGIA E CRITERI DI VALUTAZIONE</b>	<b>27</b>
<b>4.1 Metodologia seguita nel corso dell'esperimento</b>	<b>27</b>
<b>4.2 Criteri di valutazione</b>	<b>28</b>

<b>4.3</b>	<b>Elaborazione trama della tavola</b>	<b>29</b>
<b>5.</b>	<b>ESPERIMENTO: ANALISI DEI RISULTATI E DISCUSSIONE</b>	<b>31</b>
<b>5.1</b>	<b>AI Comic Factory</b>	<b>31</b>
<b>5.2</b>	<b>Mage Space</b>	<b>42</b>
<b>5.3</b>	<b>Valutazione finale</b>	<b>48</b>
<b>7.</b>	<b>CONCLUSIONI</b>	<b>50</b>
<b>8.</b>	<b>FONTI BIBLIOGRAFICHE E SITOGRAFIA</b>	<b>52</b>
	<b>RINGRAZIAMENTI</b>	<b>55</b>

# 1. INTRODUZIONE

L'avvento dell'intelligenza artificiale (IA) ha aperto nuove prospettive e avuto notevoli impatti in molti campi, rivoluzionando numerosi settori, compreso quello artistico. Le tecnologie basate sull'IA offrono nuove prospettive e impatti significativi nel campo artistico, aprendo la strada a nuove forme di espressione e creazione. In questa tesi di laurea esploreremo l'impatto delle intelligenze artificiali nella generazione di tavole di fumetti. Analizzeremo le opportunità che queste tecnologie offrono per la narrativa visiva e le sfide legate alla creazione di personaggi e ambientazioni coerenti, in particolare secondo lo stile di fumetto giapponese, il manga.

Da quando è andata in voga la generazione di immagini con l'uso dell'intelligenza artificiale, tra i vari stili di disegno emulati troviamo anche stili fumettistici, giapponesi e occidentali. Di conseguenza potremmo chiederci se ci sono state delle persone che hanno provato a coinvolgere l'IA nel processo di realizzazione di un fumetto, e la risposta sarebbe affermativa: le tecnologie basate sull'IA attualmente presenti consentono di automatizzare il processo creativo, dalla creazione di personaggi e ambientazioni al layout delle vignette, ognuno di questi da realizzarsi con diversi strumenti IA specializzati uniti a un lavoro manuale di assemblaggio. Troviamo quindi delle opere realizzate, ad esempio, con l'ausilio di ChatGPT nella scrittura dei personaggi, delle ambientazioni, come supporto allo sviluppo delle trame da raccontare, nella suddivisione delle varie scene e dialoghi che le compongono, attività successivamente accompagnate da strumenti di generazione di immagini per raffigurare le idee precedentemente elaborate, realizzando una vignetta alla volta, che verranno infine composte ad hoc da un operatore umano con quindi arrangiamento delle immagini nei layout delle pagine, e sovrapposizione dei balloon testuali contenenti i dialoghi generati prima.

Al momento, questo appena descritto è l'unico modo in cui sono stati creati dei fumetti con l'uso di IA. Questa ricerca invece concentra la sua attenzione nell'analisi di strumenti quanto più completi ed efficaci, a cui dare inizialmente un input testuale, e che siano in grado di automatizzare gli step descritti sopra, in grado di sopperire al lavoro manuale e in alcuni casi tedioso, come può essere quello di copiare l'output di una scena

immaginata da ChatGPT e incollarlo nello strumento di generazione di immagini, oppure comporre le varie immagini in un layout di pagina.

Tuttavia, questa eventuale integrazione solleva sfide significative. Da un lato l'IA offrirebbe nuove opportunità per la narrativa visiva, consentendo di esplorare stili artistici diversi e di generare tavole in autonomia, senza interpellare figure specializzate del mestiere. Dall'altro, emergerebbero interrogativi sulla perdita di originalità e sulla capacità dell'IA di catturare l'essenza dell'arte umana.

Questa tesi si propone di esaminare in dettaglio l'impatto dell'IA nella creazione di tavole di fumetti, analizzando gli strumenti ad oggi presenti sul mercato in grado di svolgere da soli tutti gli step descritti sopra, necessari per la generazione di un fumetto con IA, fornendo in input un singolo prompt, le loro limitazioni e implicazioni artistiche. Per raggiungere questo obiettivo, adotteremo un approccio che combinerà analisi degli strumenti, studi di tentativi effettuati in altre ricerche, e sperimentazione degli strumenti in prima persona con analisi e confronto dei risultati.

La tesi sarà strutturata in quattro parti principali.

Nella prima parte presentiamo il contesto e i risultati già raggiunti in materia, e le metodologie seguite per ottenerli. Verrà anche presentata una descrizione di strumenti dedicati alla generazione di tavole fumettistiche.

La seconda parte sarà dedicata a un'approfondita revisione degli algoritmi e delle tecniche adottate dai principali strumenti che si sono rivelati adeguati per lo svolgimento di questo esperimento.

Nella terza parte descriveremo la metodologia seguita in fase di sperimentazione, spiegando come abbiamo progettato i nostri casi di studio e le indagini empiriche. Discuteremo i metodi utilizzati per raccogliere e analizzare i dati, nonché le misure adottate per garantire la validità e l'affidabilità della nostra ricerca.

Infine, nella quarta sezione presenteremo i risultati dei nostri studi e delle indagini, analizzandoli alla luce del nostro problema di ricerca. Esamineremo i potenziali benefici dell'IA in termini di miglioramento delle prestazioni dei suoi utilizzatori, adattamento alle esigenze individuali, feedback personalizzato e così via.

## 2. STATO DELL'ARTE

### 2.1 Progetti editoriali

Nel corso degli ultimi due anni hanno visto la luce diversi progetti editoriali sotto forma di fumetto occidentale e manga, realizzati con l'ausilio dell'intelligenza artificiale. Tra questi il più importante è un nuovo episodio di "Black Jack", celebre manga del defunto Osamu Tezuka; altre opere rilevanti sono "Fastwalkers" di Ilan Manouach e "The Abolition of Men" di Carson Grubaugh. Di seguito un approfondimento sulle opere appena citate.

#### 2.1.1 Black Jack

In occasione del 50esimo anniversario della serie di Osamu Tezuka intitolata "Black Jack", celebratosi nel 2023, l'industria dei manga ha deciso di esplorare il potenziale rivoluzionario dell'intelligenza artificiale (IA) nel processo creativo, cimentandosi nella creazione di un episodio inedito della serie citata. Questo progetto è stato realizzato dal team TEZUKA2023, composto da:

- Lo studio Tezuka Productions e il suo direttore Makoto Tezuka, figlio del defunto Osamu;
- Il prof. Satoshi Kurihara, docente di ingegneria alla Keio University ed esperto di IA;
- Il regista e sceneggiatore Kaizo Hayashi;
- Il mangaka di "Mega Man" Shigeto Ikehara;
- Il character designer Tsunogai.

L'IA ha studiato dati testuali da 200 episodi di "Black Jack" e immagini facciali di 20.000 personaggi tratti dai manga di Tezuka. Attraverso dialoghi tra l'IA e gli esseri umani, è stata concepita una trama e nuovi personaggi sono stati progettati basandosi su immagini generate dall'IA. Tuttavia, tutti i disegni sono stati realizzati dallo staff di Tezuka Productions e da creatori esterni, senza l'utilizzo di immagini generate dall'IA nel manga finale.



FIGURA 1 - PAGINE DAL NUOVO EPISODIO DI "BLACK JACK" CREATO CON L'AIUTO DI IA

I fan di “Black Jack” hanno però avuto reazioni contrastanti riguardo all’uso dell’intelligenza artificiale (IA) per creare un nuovo capitolo del manga: alcuni di loro sono scettici riguardo al coinvolgimento dell’IA in una serie così consolidata, seguendo il pensiero che utilizzare il supporto dell’IA vada contro i temi espressi da Tezuka nella serie, che enfatizzava spesso l’importanza della presenza umana ad affiancare i miglioramenti repentini che la tecnologia vede ad oggi; un altro grande timore dei fan è che una IA possa essere stata sviluppata e allenata per il solo scopo di emulare lo stile leggendario di Tezuka quando ad oggi diversi capitoli della serie son stati scritti e disegnati da numerosi artisti in grado di emulare lo stile di disegno del maestro.

Vediamo invece contrapporsi a questi pensieri l’opinione di Makoto Tezuka che, nonostante riconosca la controversia che avvolge il tema dell’IA generativa, le potenziali violazioni di diritti e riduzione della domanda di artisti umani, è comunque rimasto sorpreso dai dettagli inclusi nell’episodio generato dall’IA: egli ritiene che l’intelligenza artificiale sia riuscita a catturare l’essenza delle opere di suo padre, in merito ai toni e sfumature utilizzate nella narrativa e soprattutto alle tematiche trattate, il ché può apparire contraddittorio essendo queste ultime un’enfatizzazione della dignità degli esseri viventi – elemento chiave della narrativa di Tezuka.

L'integrazione dell'IA solleva quindi dubbi e preoccupazioni. Alcuni temono la perdita di posti di lavoro per gli artisti e la banalizzazione dell'arte: attualmente, le leggi giapponesi consentono l'uso di contenuti diversi per l'addestramento dell'IA senza il consenso esplicito dei creatori, suscitando dibattiti sulla necessità di regole più rigorose. Seppur sia vero che l'uso dell'IA ha del potenziale anche dal punto di vista dell'efficienza dell'intero processo creativo ed è capace di rendere tale arte accessibile a un bacino sempre più largo di persone, tale comportamento finirebbe per rendere tale arte sempre meno speciale, divenendo alla portata di tutti.

Nonostante ciò, c'è una visione positiva del ruolo dell'IA nella creazione di manga. L'IA potrebbe aprire nuove strade per la narrazione e l'espressione artistica, collaborando con gli esseri umani anziché sostituirli completamente.

Il team TEZUKA2023 sembra essere ottimista riguardo al futuro dell'IA nell'arte e nei manga. Makoto Tezuka ha perfino dichiarato che se suo padre fosse ancora vivo, avrebbe sicuramente utilizzato l'IA. Questo progetto dimostra come l'IA possa supportare la creazione di manga senza compromettere la qualità delle opere originali. L'obiettivo è esplorare la possibilità di una collaborazione congiunta tra l'IA e gli esseri umani su un livello più elevato, aprendo nuove prospettive creative.

### 2.1.2 Fastwalkers

Fastwalkers (Manouach, 2022) è il primo fumetto sintetico co-creato con l'ausilio dell'intelligenza artificiale emergente. Questo progetto rappresenta una meditazione non lineare sul deep learning, celebrando la poesia inaspettata della computazione generativa e esplorando il suo potenziale nel formare nuove sensibilità nei lettori. In questo libro psichedelico e vibrante, tutte le immagini e il testo sono stati prodotti con machine learning.

Il fumetto è stato sviluppato attraverso un lavoro di squadra interdisciplinare all'interno del programma Nvidia Inception Accelerator. Il team ha utilizzato gli ultimi algoritmi di IA generativa, disponibili allora nel 2022, come GAN e GPT-3, addestrandoli con immagini di anime (stile di animazione giapponese) e hentai (fumetto erotico giapponese), e l'opera nell'insieme rappresenta frutto di sperimentazione e ibridazione tra diverse tecnologie.

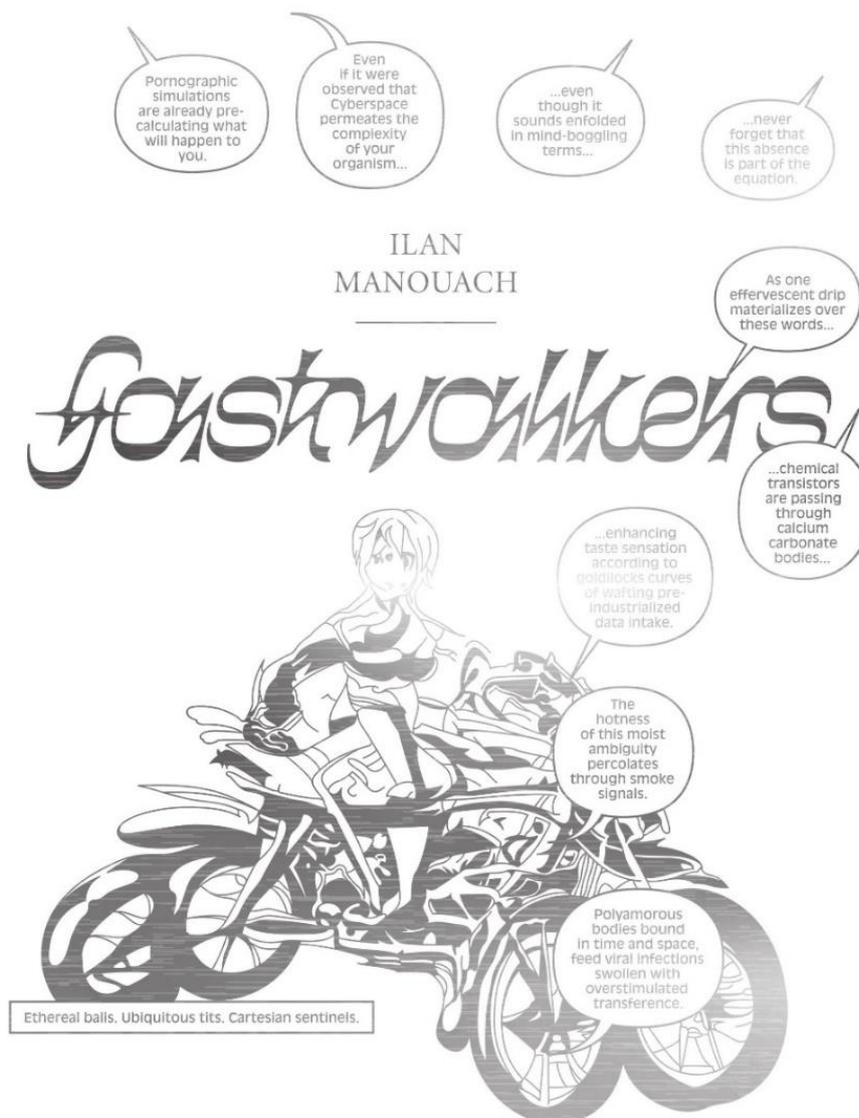


FIGURA 2 - COPERTINA DEL FUMETTO SINTETICO "FASTWALKERS"

Il risultato che si ottiene con quest'opera è comunque molto distante dall'idea comune di fumetto, nonostante la presenza dei balloon per i dialoghi testuali le immagini trasformano linee distorte (e quindi il problema principale della generazione di immagini con IA) in elementi caratteristici e "punti di forza" dello stile utilizzato in Fastwalkers.

### 2.1.3 The Abolition of Man

The Abolition of Man (Grubaugh, 2022) è una serie di fumetti innovativa e pionieristica, creata dall'artista e fumettista Carson Grubaugh, illustratore "analogico" pluripremiato e nominato agli Eisner Awards, gli oscar del fumetto negli Stati Uniti. Questa serie è stata una delle prime, nella storia dei fumetti, a essere completamente illustrata da

un'intelligenza artificiale (IA). Grubaugh ha utilizzato il sistema di generazione di immagini MidJourney AI per interpretare un saggio di Cleve Staples Lewis (scrittore, saggista e teologo britannico, autore de "Le Cronache di Narnia"), immaginare un futuro satirico e distopico e produrre grafiche per un saggio del professore di filosofia ed etica dell'informazione presso l'Oxford Internet Institute dell'Università di Oxford Luciano Floridi. L'edizione deluxe raccoglie tutti e cinque i numeri della serie, inclusa una ristampa di un'edizione alternativa estremamente rara, generata con Dall-E 2, del primo numero, in cui l'interno è stato completamente ridisegnato. Il progetto include anche discussioni approfondite sull'etica e l'impatto dell'arte IA con ospiti speciali come Dave McKean (illustratore), NFN Kalyan (pittore), John Mahoney (concept designer e filmmaker) e Luciano Floridi.

Inizialmente, Grubaugh era scettico nei confronti dell'intelligenza artificiale, convinto che fosse impossibile riuscire a trovare qualcosa di genuinamente nuovo e originale, ma al contempo ne rimase incuriosito, per cui intraprende un percorso di ricerca; questo lo porta ad accorgersi della creatività dei nuovi modelli di generazione di immagini e della potenziale velocizzazione che questo metodo di creazione di immagini porterebbe all'intero processo produttivo, e una volta addentratosi in questa ricerca decide di contribuire all'evoluzione di questa intelligenza artificiale, e di utilizzarla a suo vantaggio: è così che nasce questo fumetto. Le immagini che compongono le tavole di quest'opera seguono i due stili esemplificati nelle immagini a seguito (Figura 3), riuscendo a mantenere consistenza e coerenza stilistica.

The Abolition of Man resta comunque un risultato ottenuto con la supervisione di un artista, che fonde il contributo dell'intelligenza artificiale con la filosofia, utilizzando il fumetto come veicolo.



FIGURA 3 - DUE TAVOLE RAFFIGURANTI I DIVERSI STILI PRESENTI IN "THE ABOLITION OF MAN" DI GRUBAUGH

Le opere discusse in questo capitolo sono quindi state generate col supporto dell'intelligenza artificiale mediante un accurato lavoro suddiviso in diverse fasi: prima di tutto assistenza nella generazione di una storia mediante l'uso di chatbot come ChatGPT, poi allenamento di un algoritmo con immagini e idee coerenti con quanto si vuole generare, a seguire generazione dei diversi personaggi che compariranno nel prodotto sia dal punto di vista della loro identità personale che da quello del character design, infine creazione dei vari disegni che andranno a comporre il fumetto, e relativo assemblaggio con vignette contenenti i dialoghi. Ognuno di questi step richiede l'approvazione di un supervisore umano prima di poter proseguire nel processo creativo dell'opera, e taluni non sono automatizzati (ad esempio l'assemblaggio).

Diversamente da questo approccio, con questa tesi si vogliono esplorare i vari strumenti presenti sul mercato in grado di semplificare e comprimere tutte queste fasi, per capire se è realistico pensare che un'intelligenza artificiale possa, ad oggi, creare da sola (o

meglio, con il minimo necessario intervento umano in caso di aggiustamenti e approvazione finale del prodotto) una storia a fumetti.

## 2.2 Strumenti per generazione di fumetti con IA

A seguito di ricerche approfondite, sono stati individuati una serie di strumenti e piattaforme che provano a creare un prodotto completo partendo da un unico prompt testuale, che da un principio di trama fornito in input, e descrizione breve del contesto e dei personaggi, riescano a sviluppare una storia che abbia un suo flusso narrativo, suddividerla in scene e successivamente generare il prodotto finale, composto da disegni collocati in maniera appropriata in dei layout di pagina e corredati da dialoghi testuali in balloon. Di seguito un elenco:

- **AI Comic Factory** (<https://huggingface.co/spaces/jbilcke-hf/ai-comic-factory>);
- **Mage Space** (<https://www.mage.space/build>);
- **AI Comic Generator** (<https://aicomicgenerator.net/playground>);
- Plugger.ai (<https://www.plugger.ai>);
- Fantoons (<https://www.fantoons.xyz/>);
- Storywizard.ai (<https://www.storywizard.ai/>).

I primi tre strumenti, evidenziati in grassetto, verranno approfonditi nella sezione successiva poiché utilizzati per condurre l'esperimento. Per quanto riguarda gli altri tre, verranno descritti in breve di seguito, sottolineando i motivi per cui ho deciso di escluderli dall'esperimento.

### 2.2.1 Plugger.ai

Plugger.ai (<https://www.plugger.ai>) è una piattaforma che raccoglie strumenti per marketing e design basati su intelligenza artificiale. Si presenta come un'assistente virtuale che offre una vasta gamma di servizi: spazia dai servizi di business quali creazioni di grafiche per social media e fotografie per e-commerce, a servizi incentrati sulla fotografia come foto stock e supporto nell'editing fotografico, con rimozione di

background e strumenti di miglioramento della qualità delle immagini, arrivando infine a una selezione di servizi mirati al design, tra cui troviamo anche un AI Comic Generator.

Quest'ultimo strumento, come tanti altri del suo genere, promette di prendere in input un prompt testuale e di trasformarlo in strisce di fumetti dinamiche, complete di personaggi, sfondi e dialoghi. Supporta un ampio range di stili, dall'inchiostro all'illustrazione per graphic novels. Tuttavia non si riescono a trovare online fonti ed esempi di generazione utilizzando tale strumento, per cui essendo a pagamento e sembrando non troppo affidabile, ho deciso di non utilizzarlo nel corso della mia sperimentazione; tale opinione su questo strumento è rinforzata dall'unico video trovato su YouTube in cui un ragazzo lo menziona (link video in sitografia), raccontando come nonostante avesse sottoscritto un abbonamento a pagamento, lo strumento a ogni prompt inserito continua a chiedere soldi e non dà risultati.

### 2.2.2 Fantoons

Fantoons (<https://www.fantoons.xyz/>) è un generatore di immagini basato su IA, presentato come creatore di vignette a fumetti. Accedendo al sito si presentano una serie di opzioni tra cui scegliere, che includono un universo narrativo (Harry Potter o BTS), la scelta del personaggio che si vuole rappresentare, lo stile del disegno e l'ambientazione. Purtroppo queste selezioni sono gli unici driver per la generazione della vignetta, per cui non si riesce a generare una precisa scena che si ha in mente con personaggi che presentano determinate caratteristiche. Dopo la generazione, lo strumento ti permette di aggiungere un balloon a scelta tra diversi design e di scrivere il testo che andrà al suo interno, per cui questo passo è completamente lasciato all'utente. A causa di questi motivi ho deciso di non tenere in considerazione questo strumento per l'esperimento successivo.



FIGURA 4 - IMMAGINE GENERATA CON FANTOONS RAFFIGURANTE HERMIONE GRANGER, UN PERSONAGGIO DELLA SAGA DI HARRY POTTER

### 2.2.3 Storywizard.ai

Un altro risultato che mi è apparso durante la ricerca di strumenti per la generazione di fumetti è Storywizard.ai (<https://www.storywizard.ai/>), un sito dedicato alla generazione di percorsi educativi personalizzati sotto forma di storie illustrate. Assegnato un prompt, lo strumento genera una storia di moderata lunghezza ma comunque semplice basata su quanto indicato, e suddivisa in diverse pagine. L'organizzazione del racconto generato è la seguente, a sinistra viene mostrata la narrazione generata dallo strumento e sulla destra l'immagine associata al frammento di storia.

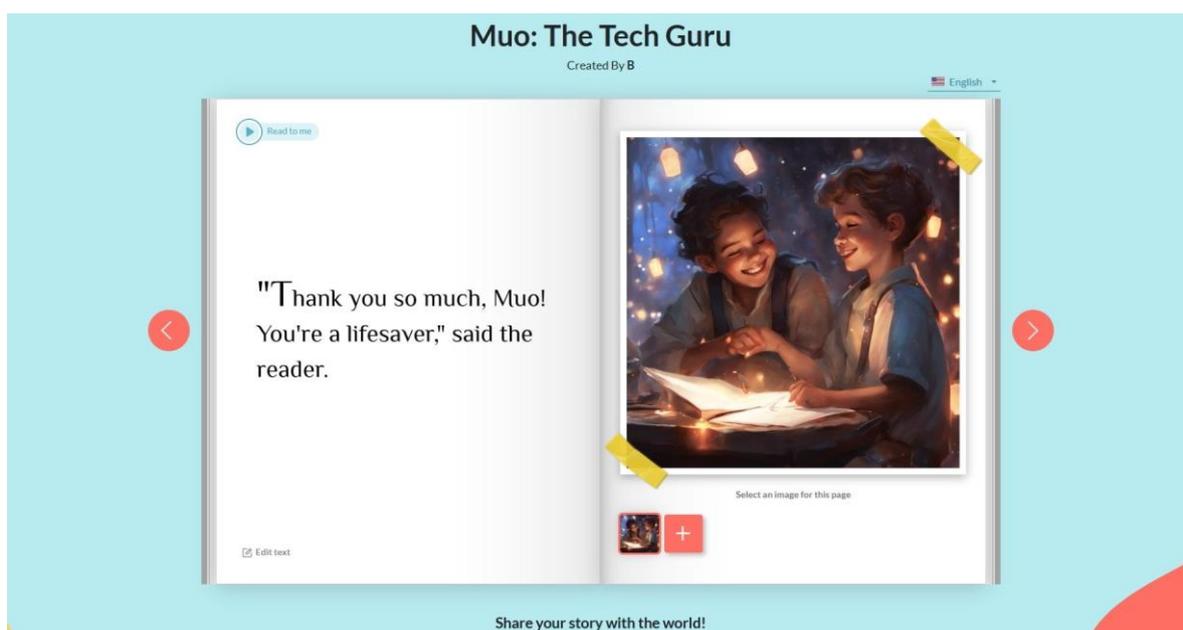


FIGURA 5 - CATTURA SCHERMO DEL SITO STORYWIZARD.AI CON ESEMPIO DI RISULTATO

Siccome il risultato non rispecchia comunque la struttura di un fumetto, composto da pagine suddivise in pannelli illustrati e corredati da balloons testuali, ho deciso di escludere questo strumento dalla mia sperimentazione.

## 2.3 Ricerche accademiche

In questa sezione verranno raccontate alcune delle ricerche accademiche condotte sul tema del coinvolgimento dell'intelligenza artificiale generativa nella produzione di immagini in stile fumettistico giapponese, il manga, stile su cui ci concentreremo anche nel corso dell'esperimento.

### 2.3.1 Colorazione di manga

Il titolo di questa ricerca è "*Automatic Manga Colorization with Color Style by Generative Adversarial Nets*" (Kataoka, 2017).

In questo documento si affronta il problema della colorizzazione automatica di immagini di fumetti, prendendo in input dall'utente un'immagine colorata da cui trarre ispirazione.

L'obiettivo che si pone è dunque quello di, dato uno schizzo in bianco e nero e delle preferenze sulla colorazione, generare delle immagini di manga a colori.

Per ottenere questo risultato viene utilizzata una combinazione di Deep Convolutional Neural Networks (CNN o DCNN), reti neurali convoluzionali profonde utilizzate per il riconoscimento di pattern a partire da un'immagine, e Conditional Generative Adversarial Networks (CGAN), che consentono la generazione e discriminazione più precisa di immagini, condizionandole su dati specifici.

La sfida principale presentata in questo studio è posta dal fatto che i disegni a linee (come sono gli schizzi) contengono meno informazioni rispetto alle immagini in scala di grigi, per cui necessitano di molti suggerimenti sui colori forniti dall'utente per migliorare la qualità delle immagini generate. Di seguito un esempio di immagine estratta dal documento accademico.

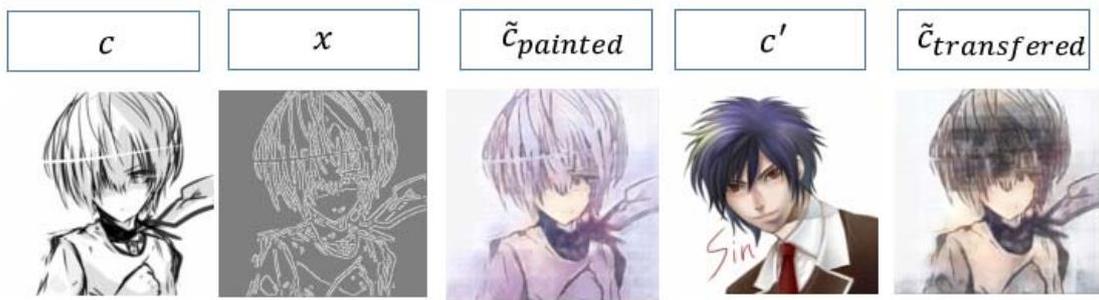


FIGURA 6 - COLORIZZAZIONE AUTOMATICA DI UNO SCHIZZO BASATA SU PREFERENZE DI COLORE FORNITE DALL'UTENTE

### 2.3.2 Simulazione del processo creativo di un manga

Il titolo di questa ricerca è “*Generating Manga from Illustrations via Mimicking Manga Creation Workflow*” (Zhang, 2021).

Questo documento presenta un framework per generare manga a partire da illustrazioni digitali. Negli studi professionali di manga, il workflow seguito nella creazione comprende tre fasi chiave:

1. Lineart dei disegni: gli artisti utilizzano disegni composti da linee semplici per delineare gli schemi strutturali nei storyboard dei manga;
2. Applicazione retini regolari: gli artisti applicano diversi tipi di retini regolari per creare ombreggiature, occlusioni e texture dei diversi materiali di cui sono fatti i vari soggetti della scena;
3. Applicazione retini irregolari: gli artisti incollano selettivamente retini irregolari sulla tela per ottenere vari layout di sfondo o effetti speciali.

Motivati da questo flusso di lavoro, i ricercatori propongono un framework che attinge da un dataset contenente annotazioni su questi tre step fornite da artisti umani, per convertire un'illustrazione digitale nelle tre componenti descritte prima: la lineart del manga, retinatura regolare e irregolare.

Per raggiungere il risultato mostrato di seguito, anche qui, sono state allenate delle Convolutional Neural Networks (CNN).

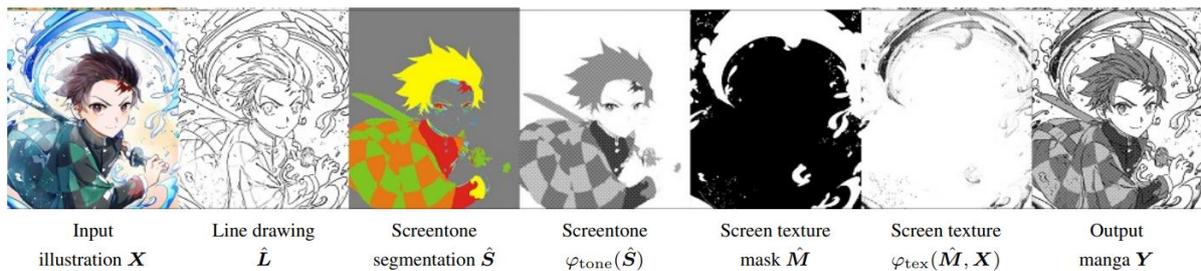


FIGURA 7 – CONVERSIONE DI UN’ILLUSTRAZIONE DIGITALE IN DISEGNO STILE MANGA, E LE VARIE FASI DEL PROCESSO

### 2.3.3 Traduzione di fotografie in disegni in stile manga

Il titolo di questa ricerca è “*MangaGAN: Unpaired Photo-to-Manga Translation Based on the Methodology of Manga Drawing*” (Su, 2021).

Questo documento propone un metodo innovativo per la traduzione automatica di foto in stile manga, preservando sia la somiglianza facciale che lo stile manga.

Come intuibile dal nome, MangaGAN una Generative Adversarial Network (GAN), appositamente progettata e allenata, per generare le caratteristiche geometriche dei volti in pieno stile manga. Successivamente, traduce ciascuna regione facciale nel dominio manga attraverso un’architettura multi-GANs personalizzata. Questa GAN è stata addestrata con caratteristiche facciali, corpi, tratti distintivi presi da una serie di manga popolare, *Bleach*. Di seguito troviamo un esempio, raffigurante una tavola del manga appena citato i cui volti sono stati sostituiti da quelli generati con MangaGAN.



FIGURA 8 – GENERAZIONE DI VOLTI IN STILE MANGA A PARTIRE DA FOTOGRAFIE DI PERSONE REALI

### **3. ESPERIMENTO: OBIETTIVO E STRUMENTI UTILIZZATI**

L'esperimento che verrà tenuto consiste nella generazione, con i due strumenti AI Comic Factory e Mage Space (descritti qui di seguito), di una intera tavola di un fumetto in stile manga a partire da un singolo prompt, con successiva analisi dei risultati, e valutazione degli stessi. L'obiettivo è quello di valutare se tali strumenti riescono a elaborare, a partire da un input testuale, le varie fasi che compongono il processo creativo di un fumetto: sviluppo della trama, suddivisione in vignette distinte, rappresentazione delle immagini e composizione delle stesse in un layout verosimile corredate da balloon testuali per i dialoghi.

#### **3.1 AI Comic Factory**

AI Comic Factory (<https://huggingface.co/spaces/jbilcke-hf/ai-comic-factory>) è uno strumento opensource ospitato su Hugging Face Spaces e sviluppato dal data scientist Julian Bilcke. Questo strumento sfrutta modelli di intelligenza artificiale all'avanguardia per generare fumetti a 4 pannelli basati su prompt di testo forniti dall'utente. Utilizza modelli di elaborazione del linguaggio naturale (NLP) come Llama per creare didascalie per i fumetti. Inoltre, l'AI Comic Factory è un'applicazione NextJS, distribuita tramite Docker, e si basa su un approccio client-server, richiedendo due API per funzionare: un API di LLM (Large Language Model, tecnologia IA avanzata incentrata sulla comprensione e sull'analisi del testo) e un API di rendering per la generazione delle immagini.

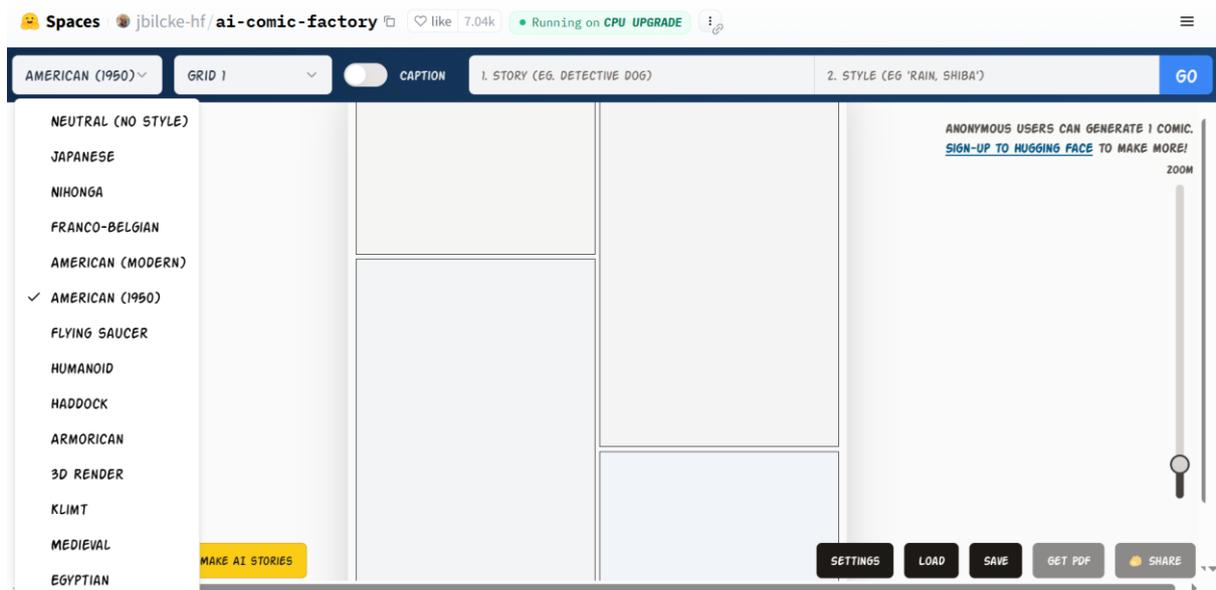


FIGURA 9 - INTERFACCIA DI AI COMIC FACTORY SULLA PIATTAFORMA HUGGING FACE

L'applicazione web presenta diversi campi da compilare per la generazione:

- Stile: stile fumettistico in cui si desidera generare la tavola;
- Layout: scelta tra 4 diversi layout di tavola, con i quattro pannelli disposti in proporzioni differenti;
- Caption: flag che specifica se generare anche dei testi didascalici ad accompagnare i pannelli;
- Prompt "Story": prompt che va a definire la scena da rappresentare nel corso dei pannelli;
- Prompt "Style": secondo prompt che permette di specificare ulteriormente le caratteristiche desiderate nell'estetica del disegno.

Il modello IA per la generazione di storie utilizzato di default nella versione corrente (aggiornata l'ultima volta a maggio 2024) è Zephyr 7B  $\beta$ . Secondo modello della serie Zephyr, sviluppata da Hugging Face, si tratta di una versione ottimizzata del modello Mistral-7B-v0.1, un modello GPT-like a 7 miliardi di parametri (7B) addestrato su una combinazione di dataset pubblici e sintetici utilizzando tecniche come Direct Preference Optimization (DPO), paradigma di ottimizzazione del modello basato sulle preferenze umane.

Per quanto riguarda il modello per la generazione delle immagini utilizzato, parliamo di Stable Diffusion XL, uno dei modelli di immagini ad accesso libero sviluppati da Stability

AI. Questo modello è costruito su una nuova architettura, composta da un modello di base con 3,5 miliardi di parametri e da un refiner con 6,6 miliardi di parametri. Le immagini vengono generate a partire dal prompt testuale frutto dell'output di Zephyr, combinato con le indicazioni fornite dall'utente in input. Si tratta di un Latent Diffusion Model che utilizza due encoder di testo pre-addestrati: OpenCLIP-ViT/G e CLIP-ViT/L, responsabili della traduzione del prompt testuale in rappresentazione visiva. Stable Diffusion suddivide la generazione dell'immagine in due fasi: prima di tutto, viene usato il modello base per generare una rappresentazione latente delle dimensioni desiderate; successivamente, viene usato un modello specializzato ad alta risoluzione e una tecnica chiamata "img2img" che prende in input l'immagine latente generata prima, il prompt testuale di partenza, e genera una nuova immagine basata sull'originale, impiegando più tempo per ottenere un risultato più fedele al prompt dato.

La versione gratuita di questo applicativo permette di generare i fumetti all'interno della web app con i modelli impostati di default e raccontati prima ma, se si è disposti a pagare, l'applicativo è compatibile con altri modelli, tra cui per il lato di generazione di immagini possiamo trovare istanze di SDXL e LoRA custom, allenate a proprio gusto, e DALL-E 3 by OpenAI, mentre tra i modelli di generazione di storie abbiamo compatibilità per Claude by Anthropic, ChatGPT by OpenAI e modelli opensource disponibili su Groq.

Delle features interessanti di questo applicativo, features che Mage Space non presenta, sono: la possibilità di ridisegnare i singoli pannelli che compongono la tavola, e quella di visualizzare il testo alla base di ogni disegno, testo che è una rielaborazione dei prompt testuali inseriti dall'utente e integrata dallo sviluppo della storia che si sta cercando di rappresentare, quest'ultima elaborata dal modello di generazione di storie usato.

### **3.2 Mage Space**

Mage Space (<https://www.mage.space>) è uno strumento di generazione di immagini, gratuito e accessibile online, basato su intelligenza artificiale che ti consente di creare arte digitale a partire da prompt o descrizioni; Utilizza una combinazione di tecnologie all'avanguardia per generare immagini a partire da testo descrittivo: principalmente, si basa su reti neurali generative, che sono modelli di deep learning addestrati su grandi

dataset di immagini e testo; questi modelli apprendono a tradurre le descrizioni testuali in rappresentazioni visive. Tuttavia, i dettagli specifici dell'implementazione sono proprietari e non sono disponibili pubblicamente.

Tra i modelli di generazione di immagini usati da Mage Space troviamo Stable Diffusion XL, Stable Diffusion v1.5 e Stable Diffusion v2.1. Da ciò si evince che questo strumento è basato su Stable Diffusion, ma a meno che non si acceda alla versione a pagamento non si riesce ad avere pieno controllo su quali modelli vengono utilizzati nella generazione. Inoltre, a seconda del tipo di abbonamento che si sottoscrive, si possono ottenere sempre più vantaggi tra cui la possibilità di utilizzare dei modelli di tipo LoRA allenati personalmente, come anche su AI Comic Factory.



FIGURA 10 - HOME PAGE DI MAGE.SPACE

Per accedere allo strumento è necessario un account, sottoscrivibile anche con Google. Nella homepage di questo sito troviamo le immagini consultate di recente e una serie di immagini generate dalla community e rese pubbliche sulla piattaforma, mentre nella sezione di "Search" si possono ricercare immagini e profili di utenti registrati.

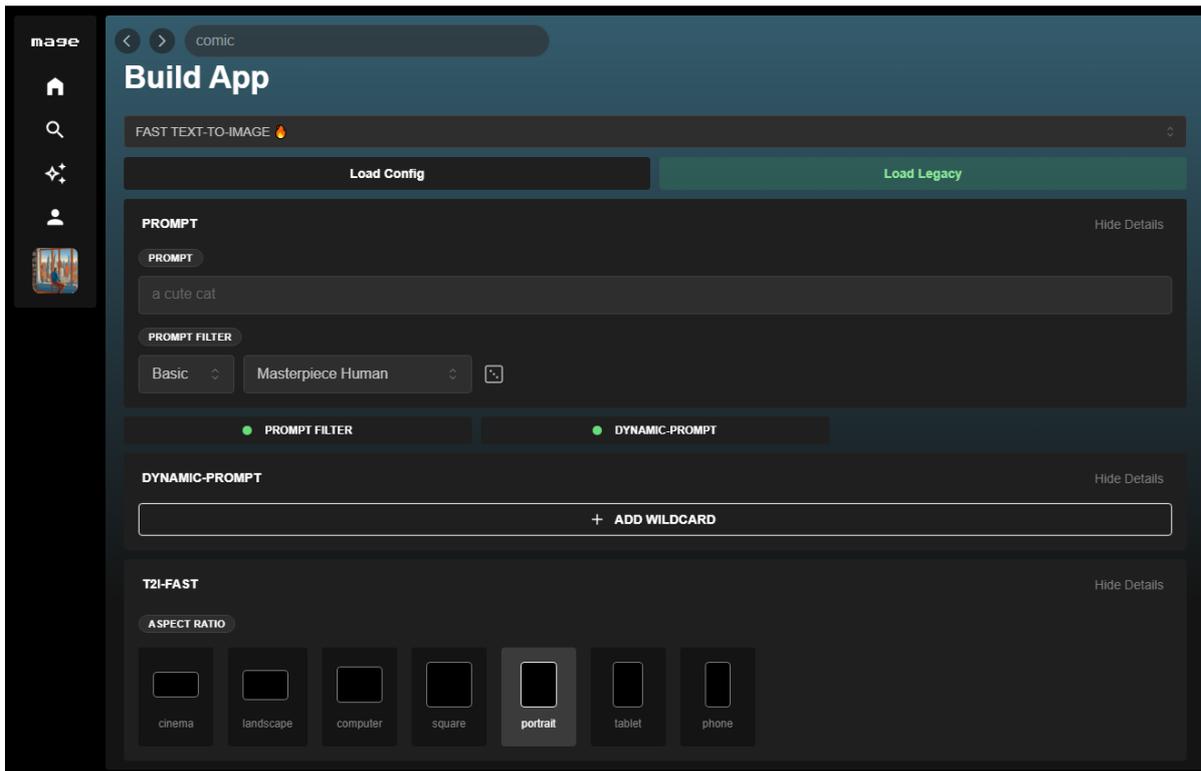


FIGURA 11 - SEZIONE BUILD APP DI MAGE.SPACE

La sezione di nostro interesse di questo sito è la sezione “Build App”, dedicata alla generazione di nuove immagini. I parametri che troviamo in questa sezione sono:

- Tipologia di applicazione: le uniche disponibili per l’uso gratuito sono FAST TEXT-TO-IMAGE e FAST SELFIE, in base a cosa si seleziona cambiano gli altri parametri da configurare come input della generazione; quelli descritti in questa sezione saranno relativi a FAST TEXT-TO-IMAGE poiché di interesse per la tesi;
- Load Config: questo pulsante consente di caricare dei preset, esistenti di default o salvati dall’utente, di configurazioni secondo cui generare l’immagine;
- Prompt: la sezione prompt permette di inserire il testo secondo cui generare l’immagine e di selezionare, tramite i due menu a tendina presenti, lo stile desiderato; tendenzialmente gli stili presenti nei menu a tendina sono tutti differenti anche nel nome, ma sono comunque presenti dei casi di due diverse entry dello stesso stile, per cui suppongo che le elaborazioni effettuate (e probabilmente anche i modelli utilizzati) differiscono da una entry all’altra (presumibilmente unico modo in modalità gratuita di “selezionare” il modello generativo usato);

- Dynamic-Prompt: questa sezione permette di dichiarare delle wildcards, delle variabili che specificano elementi dell'immagine che vogliamo vengano generati in maniera più creativa e interpretativa dal modello, senza dover essere particolarmente specifici nella descrizione di questi;
- Aspect Ratio: infine l'ultimo parametro da specificare è il rapporto in cui si vuole generare l'immagine.

Mage Space, una volta impostati i parametri e generata un'immagine, offre la possibilità di utilizzare il seed di quella immagine per ritentare l'ottenimento di un risultato coerente al precedente.

Quindi come analizzato finora, entrambi questi strumenti basano la loro generazione di immagini su modelli di Stable Diffusion, la differenza principale consta nel fatto che AI Comic Factory non solo genera l'immagine ma elabora ulteriormente il prompt con un modello dedicato alla generazione di storie per poterlo sviluppare in un concetto rappresentabile da 4 diverse immagini sequenziali.

### **3.3 AI Comic Generator**

Per la conduzione di questo esperimento, era previsto il coinvolgimento di un ulteriore strumento, AI Comic Generator (<https://aicomicgenerator.net/playground>). Questo strumento, così come AI Comic Factory, nasce su Hugging Face, e infatti non a caso hanno dei nomi simili: Hugging Face è una piattaforma orientata allo sviluppo di applicazioni che utilizzano AI, e offre agli utenti uno spazio e l'utilizzo dei vari modelli che ospita, per poter eseguire queste applicazioni, i cui nomi quindi rispecchiano lo scopo e le intenzioni dell'applicativo sviluppato (sulla piattaforma sono disponibili diverse applicazioni con questi nomi).

Le somiglianze con AI Comic Factory si evincono agilmente dallo screenshot inserito a seguire: l'applicazione sembra completamente identica se non per il fatto che AI Comic Generator si impone di generare di default due pagine, e inoltre rispetto all'altro strumento sono assenti dei pulsanti in basso quali Settings, Load e Save.

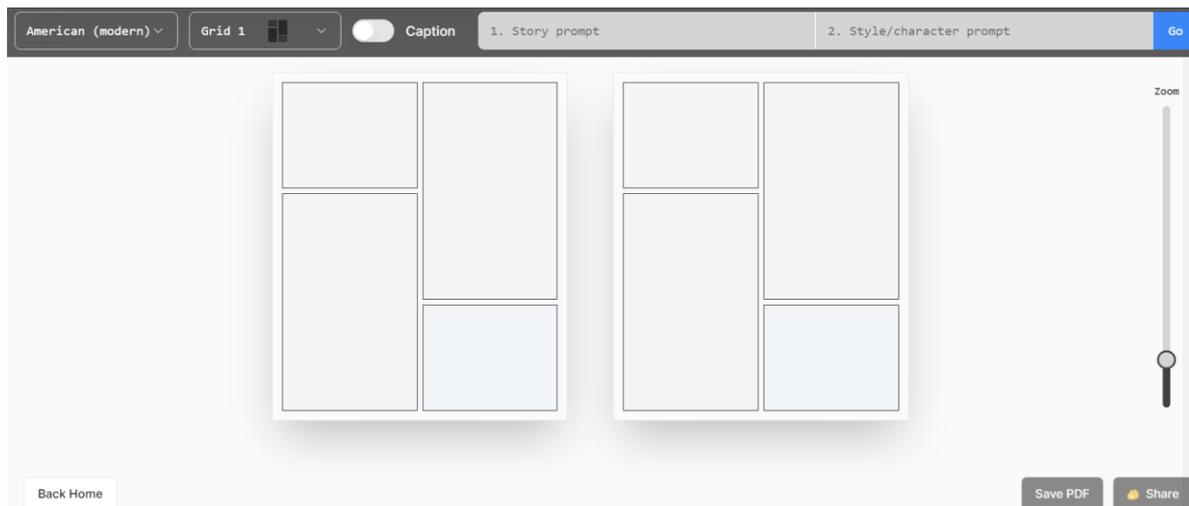


FIGURA 12 - AI COMIC GENERATOR PLAYGROUND

Il motivo per cui non si è rivelato possibile utilizzare questo strumento in fase di sperimentazione è uno: nel momento in cui ho tentato di fare le generazioni previste per questo esperimento lo strumento, una volta aver premuto il pulsante Go, fa comparire in ogni pannello delle rotelle di caricamento che però, una volta completate, invece di far apparire l'immagine generata restano immobili, e dunque non si riesce a giungere alla conclusione di questa generazione. Questo problema suppongo sia emerso negli ultimi tempi poiché, quando ad Aprile ho iniziato le ricerche riguardo questi strumenti, avevo effettuato un test di generazione e ero riuscita ad ottenere un risultato, mostrato di seguito.



FIGURA 13 - TEST DI GENERAZIONE CON AI COMIC GENERATOR RISALENTE AL 21 APRILE 2024

In base al risultato di questa generazione ritengo che i modelli utilizzati alla base fossero non solo diversi da quelli adottati da AI Comic Factory, ma per certi aspetti più soddisfacenti, ma ne discuterò meglio una volta mostrati i risultati dei due strumenti discussi precedentemente.

## 4. ESPERIMENTO: METODOLOGIA E CRITERI DI VALUTAZIONE

Dopo aver analizzato i due strumenti che verranno usati nel corso di questo esperimento, procedo ad illustrare la metodologia che ho deciso di seguire per condurre l'esperimento stesso, metodologia che verrà applicata equamente ad entrambi i sistemi di generazione.

Dopo aver ideato la trama che cercherò di rappresentare con entrambi gli strumenti, procederò a fare tentativi di generazione, iterando il processo di valutazione dell'immagine generata, correzione del prompt per cercare di far convergere il risultato in un prodotto accettabile, e seguente generazione della nuova immagine. Verrà utilizzato un timer per misurare il tempo trascorso sperimentando con i singoli strumenti, in modo da poter esprimere un giudizio relativo anche al tempo impiegato per il raggiungimento del risultato finale. Alla fine, dopo aver terminato le generazioni con entrambi gli strumenti, i risultati intermedi e finali verranno valutati in base ai criteri descritti nel Capitolo 4.2.

Di seguito una descrizione più dettagliata del workflow seguito per condurre l'esperimento.

### 4.1 Metodologia seguita nel corso dell'esperimento

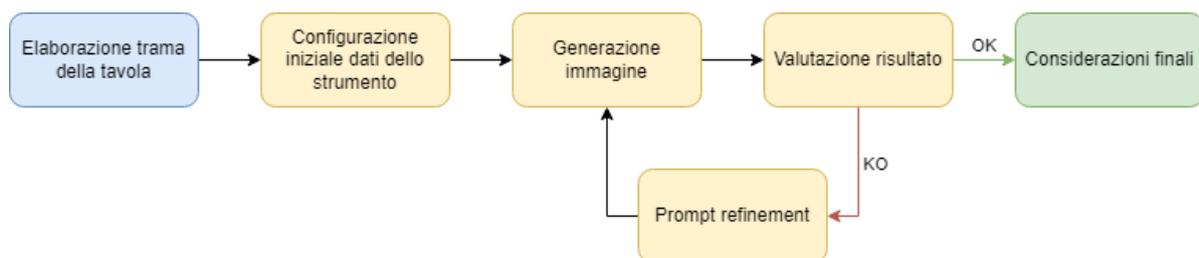


FIGURA 14 - WORKFLOW SEGUITO IN FASE DI SPERIMENTAZIONE

Il primo step, evidenziato in azzurro, è propedeutico all'intero esperimento; gli step evidenziati in giallo saranno applicati a ognuno dei due strumenti; l'ultimo step, evidenziato in verde, segna la fine dell'esperimento, con relativa analisi dei risultati e considerazioni sul comportamento degli strumenti.

Qui di seguito sono spiegati uno ad uno i vari step illustrati sopra:

1. **Elaborazione trama della tavola:** questo step viene eseguito una volta sola nel corso dell'intero esperimento e consiste nell'ideare una storia di partenza da dare in input ad entrambi gli strumenti che stiamo testando;
2. **Configurazione iniziale dati dello strumento:** in questo step viene fatto partire un timer e vengono impostati tutti i parametri in input dello strumento che stiamo testando, tra cui il prompt e stile desiderato per la generazione;
3. **Generazione immagine:** vengono dati in pasto agli algoritmi i dati di configurazione impostati, si attende la generazione dell'immagine, e a seguito di ogni generazione verrà registrato il tempo impiegato dall'inizio del timer nello step 2;
4. **Valutazione risultato:** viene analizzato il risultato della generazione e valutato se accettabile o meno, in base alla fedeltà al prompt e alla qualità dell'immagine prodotta; nel caso in cui il risultato non sia soddisfacente, il test sarà ko e passeremo allo step 5, e nel caso il test darà risultati accettabili allora verrà considerato ok e si interromperà il timer per poi passare allo step 6 (nel caso questo non dovesse accadere, dopo un numero di tentativi a mia discrezione riterrò l'esperimento con tale strumento concluso);
5. **Prompt refinement:** a questo punto verranno adattati i dati forniti in input. E quindi modificato il prompt, in base alla reazione del modello al prompt precedente, in modo da poter manipolare il risultato verso i desiderata, prima di passare di nuovo alla generazione dell'immagine (step 3);
6. **Considerazioni finali:** i risultati parziali e finali ottenuti, con i relativi tempi, verranno usati per costruire una tabella che illustra il risultato della sperimentazione, e verranno esposti ad un campione di utenti, esperti e non in campo di fumetti, a cui verrà chiesta una valutazione secondo criteri specifici che verranno illustrati più avanti.

## 4.2 Criteri di valutazione

I criteri che comporranno la scheda di valutazione, che verrà poi applicata al frutto dell'intero processo creativo, tenendo conto dei primi risultati, dei progressi intermedi e del prodotto finale, sono i seguenti:

- **Tempo impiegato:** con questo si intende il quantitativo di tempo intercorso tra lo step 2 in cui viene fatto partire un timer, fino allo step 6 ossia il momento in cui l'esperimento con lo strumento in oggetto sarà reputato finito;
- **Necessità di rielaborazioni:** verrà giudicato il numero di rielaborazioni che sono state necessarie per sistemare e rifinire l'immagine generata, e giungere quindi a un risultato accettabile;
- **Creatività:** verrà valutato quanta libertà nell'interpretazione e nel generare la situazione a partire dal prompt abbastanza vago inserito come parametro;
- **Fedeltà al prompt:** valutazione della rappresentazione dei soggetti e delle scene in base a quanto dichiarato nel prompt fornito in input;
- **Presenza di distorsioni:** qui si intende la presenza di elementi non realistici, indice di limitazioni dei modelli, come ad esempio dita in più, posizionamenti sbagliati, proporzioni irrealistiche e significativamente mal proporzionate;
- **Gradimento personale:** quanto il risultato sia esteticamente appagante/godibile.

### 4.3 Elaborazione trama della tavola

Come spiegato prima, il primo step dello schema è di preparazione e propedeutico all'intero esperimento, non dipendente dai singoli strumenti: consiste nella definizione di una trama riassumibile nella realizzazione di 4 pannelli, da dare successivamente in pasto in egual modo a entrambi gli strumenti. Tuttavia, date le differenze nell'organizzazione dei dati in input dei due strumenti, nonostante la sostanza del prompt rimane la stessa, la forma e organizzazione delle frasi da inserire sarà purtroppo differente.

In italiano, la scena che cercherò di far disegnare ai due strumenti nel corso dell'esperimento sarà la seguente:

*Due ragazzi stanno per cominciare a combattere in un magazzino. Il primo ragazzo, di nome MAHITO, indossa abiti ordinari, ha capelli lunghi e la sua faccia è coperta da cicatrici. Il secondo ragazzo, di nome CHOSO, indossa abiti tradizionali giapponesi, ha delle occhiaie, i suoi capelli sono raccolti e combatte utilizzando una lama fatta del suo stesso sangue.*

Nella costruzione di questa trama, sono stati assegnati dei nomi ai personaggi per aiutare a gestire la distinzione degli elementi descrittivi dei due nella scrittura dei prompt, ma anche perché il primo dei due strumenti, in grado di generare del testo comprensibile, avrebbe altrimenti assegnato dei nomi di sua invenzione. I nomi usati sono nomi giapponesi, per coerenza con lo stile con cui andremo a generare il fumetto ossia lo stile manga, e traggono ispirazione da un'opera di fantasia, a cui attinge anche l'ambientazione stessa del prompt. Inizialmente ho avuto il timore che, scegliendo una scena assimilabile ad un'opera esistente e abbastanza popolare ad oggi (ed usando soprattutto i nomi tratti da tale opera), avrei agevolato il lavoro ai modelli di generazione di immagini dato l'enorme quantitativo di immagini, ufficiali e non, già esistenti a loro nome, ma subito dopo aver tentato la generazione sono stata smentita.

## 5. ESPERIMENTO: ANALISI DEI RISULTATI E DISCUSSIONE

### 5.1 AI Comic Factory

Il primo passo nel condurre questo test è quello di impostare i dati di input nello strumento:

- Stile: *japanese*;
- Layout: *grid 1*, quella impostata di default, ho ritenuto fosse simile al layout utilizzato nei manga, gli altri tipi di layout mi ricordavano di più il fumetto occidentale (menzione speciale per la *grid 0*, che divide la pagina in 4 sezioni uguali, poiché ricorda un formato particolare di manga, chiamato *yonkoma manga*, incentrato su storie brevi raffigurate in una sola pagina);
- Caption: flag attivato;
- Prompt “Story”: *two boys starting a fight in a storehouse*;
- Prompt “Style”: *the first boy is dressed in casual clothes, has long hair, and his face is covered in stitches. The second boy is dressed in traditional japanese clothes, has bags under his eyes, his hair is tied up and fights with a blade made of his own blood.*

La suddivisione del prompt tra “Story” e “Style” di questo primo tentativo, rispetto a quanto raccontato prima in fase di descrizione degli strumenti, è evidentemente errata: questo problema nasce dai testi placeholder inseriti nei due input testuali, visibili anche dallo screenshot di Figura 6, che suggerivano di inserire nel prompt relativo alla storia una descrizione generale dell’azione e soggetti da rappresentare, mentre nel secondo prompt dovessero venir inserite le specifiche stilistiche non solo relative al tipo di illustrazione da rappresentare ma anche alle caratterizzazioni estetiche di personaggi e ambientazioni.

Una ulteriore osservazione che faccio è, diversamente da quanto detto nel paragrafo precedente relativo all’ideazione della trama e soprattutto all’uso dei nomi propri per riferirsi ai personaggi, in questo primo tentativo ho generalizzato indicandoli come “Il primo ragazzo” e “Il secondo ragazzo”, timorosa di poter generare confusione, e soprattutto non immaginando che nelle caption delle vignette venissero scritti dei nomi.



FIGURA 15 - PRIMA GENERAZIONE CON AI COMIC FACTORY; TEMPO 00:10

Dopo quindi 10 secondi, tempo necessario principalmente all’inserimento dei dati nell’interfaccia, otteniamo il primo risultato. Da questo possiamo già notare quanto sia veloce l’elaborazione del prompt e successiva generazione dell’immagine, immagino frutto di una limitazione di risorse in termini di tempo imposta al modello di default.

Le immagini inserite come risultati dell’esperimento in questo capitolo sono un po’ carenti di qualità in quanto screenshot; infatti la funzionalità di “Save” presente in interfaccia scarica un file con estensione .clap, estensione letta dall’applicativo tramite il pulsante “Load” che carica l’intera configurazione con immagine generata, mentre il pulsante “Get PDF” esporta sì l’immagine in formato .pdf ma priva di caption, elementi presenti solo in applicativo.

Possiamo fare diverse osservazioni su questo risultato:

- I personaggi coinvolti nella pagina sono ben più dei due ragazzi designati nel prompt in input: ne possiamo contare sette, tutti molto simili tra di loro in termini di capigliatura e abbigliamento;

- Fatta eccezione per la prima vignetta, i vestiti appaiono eccessivamente e innaturalmente stropicciati, causando nella maglia del personaggio in basso a sinistra degli evidenti errori di generazione;
- Non si riescono a distinguere i due personaggi principali descritti nel prompt, tranne l'ultimo in basso a destra, contraddistinto dall'abbigliamento diverso rispetto agli altri; nonostante questo però, i restanti elementi di caratterizzazione estetica sono stati mescolati e non attribuiti al corretto personaggio (il ragazzo con abiti tradizionali avrebbe dovuto avere capelli raccolti e una lama per combattere, eppure egli è raffigurato a mani nude e con capelli lunghi), e alcuni elementi sono addirittura completamente assenti (cicatrici in volto);
- Altri errori di generazione particolarmente evidenti sono i polsi e le mani, di quasi ogni generazione, con articolazioni innaturali e delle masse indistinguibili in alcuni casi al posto delle mani;
- Uno dei ragazzi di sfondo nella vignetta in alto a destra è compenetrato con il tavolo, dalla posa deduco che il motore volesse raffigurarlo seduto, con gli avambracci poggiati sulla superficie e le gambe che spuntano da sotto.

Nonostante uno dei problemi noti di Stable Diffusion sia la generazione di volti mal formati, in quest'immagine tutti i personaggi hanno dei volti costruiti correttamente a livello di proporzioni e collocamenti di occhi, naso e bocca; l'unico personaggio che mi destabilizza leggermente è quello che dovrebbe essere un personaggio femminile nella prima vignetta, con una leggera sproporzione nell'occhio sinistro e un'ombra sul mento che ricorda un pizzetto.

Utilizzando lo strumento che mi permette di leggere il testo scritto dai modelli sotto ognuna delle quattro immagini mi sono resa conto dell'errore fatto nella suddivisione del prompt nei due campi di input: ciascuno dei quattro testi sotto le vignette riportava in maniera letterale, e sottolineo letterale, il testo del prompt testuale di "Style", con in aggiunta le indicazioni su colorazione e stile di illustrazione e davvero poco di rielaborazione. Ho quindi realizzato che il prompt "Story" veniva usato per creare le differenti scene, mentre il prompt stilistico viene applicato a prescindere e in egual modo a tutti i disegni.

Ho quindi modificato i due prompt in:

- Prompt “Story”: *MAHITO and CHOSO are starting a fight in a storehouse. MAHITO is dressed in casual clothes, has long hair, and his face is covered in stitches. CHOSO is dressed in traditional japanese clothes, has bags under his eyes, his hair is tied up and fights with a blade made of his own blood. ;*
- Prompt “Style”: *manga.*

Sul prompt “Style” ho deciso quindi di mantenermi sul generico poiché non avevo idee su particolari scelte stilistiche da applicare all’intera pagina.

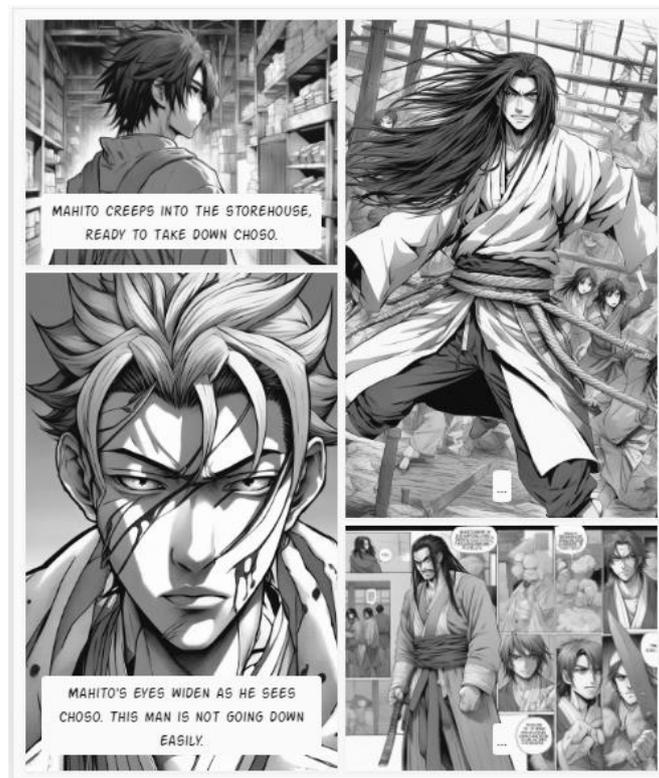


FIGURA 16 – SECONDA GENERAZIONE; TEMPO 04:00

Da questa seconda immagine generata notiamo come l’intera idea dietro alla pagina sia stata stravolta; i testi presenti nelle caption sono diversi dai precedenti e riportano i nomi inseriti nel prompt. Tuttavia, ritroviamo gli stessi problemi di prima e non solo:

- I personaggi presenti sono in numero maggiore a quello desiderato ma almeno presentano differenze più evidenti;
- Le caratteristiche dei personaggi sono state ancora una volta mescolate, stavolta non è presente neanche un personaggio vestito con abiti ordinari, unica nota che faccio è che il personaggio in basso a sinistra con capelli corti sembra avere effettivamente delle occhiaie come descritto dal prompt;

- Gli errori di generazione si concentrano nell'ultima vignetta, unica in cui sono presenti delle mani (nelle altre vignette troviamo dei close up con estremità degli arti fuori campo). Tra questi errori purtroppo troviamo oltre alle mani, errori nella generazione di ogni singola faccia, e molti elementi completamente indistinguibili, oltre a una assolutamente non richiesta suddivisione ulteriore in parti.

A questo punto ho provveduto a controllare il testo presente sotto ogni vignetta, valutarne la creatività, la fedeltà rispetto a quanto raffigurato, per poi provare a rigenerare singolarmente le vignette in base a quanto scritto sotto (e dove necessario apportando correzioni al prompt).



FIGURA 17 – PRIMA (SINISTRA) E DOPO RIDISEGNAMENTO VIGNETTA 1 (DESTRA); TEMPO 13:20

Il testo presente sotto questa immagine è:

*grayscale, detailed drawing, japanese manga, manga, dimly lit storehouse, with Mahito's back facing the camera as he enters. Mahito is a male character in his early 20s with a pale complexion and stitches covering his face. He is dressed in casual clothes and has long, disheveled hair. The storehouse is dark, with only a faint flicker of light coming from a nearby torch. Try to use soft, muted tones to create a suspenseful atmosphere.*

Il ripetersi della parola manga è dato da quanto inserito da me nel prompt “Style” e da quanto indicato come stile nel menù a tendina con “Japanese”; il modello di generazione della storia sembra aver recepito correttamente le caratteristiche associate ai singoli personaggi e vuole rappresentare in questa immagine iniziale unicamente il primo, con tutte le caratteristiche indicate da me nel prompt iniziale. Tuttavia nell'immagine

generata inizialmente dal primo tentativo, non viene rispettata questa descrizione dell'estetica del personaggio, a partire dall'abbigliamento che con il redraw è stato sistemato correttamente, fino alla lunghezza di capelli e l'assenza delle cicatrici desiderate sul volto. Sia posa che ambientazione sono invece sia fedeli alla descrizione testuale che coerenti in entrambi i tentativi.

Quindi ho provveduto a modificare questo testo, cercando di specificare meglio queste due caratteristiche che non sono state rappresentate nemmeno dal redraw, le modifiche segnate in grassetto:

*grayscale, detailed drawing, japanese manga, manga, dimly lit storehouse, with Mahito's back facing the camera as he enters. Mahito is a male character in his early 20s with a pale complexion and stitches covering his face, **running vertically across his right eye**. He is dressed in casual clothes and has **very long, straight hair**. The storehouse is dark, with only a faint flicker of light coming from a nearby torch. Try to use soft, muted tones to create a suspenseful atmosphere.*



FIGURA 18 - PRIMA (SINISTRA) E DOPO ULTERIORE RIDISEGNAMENTO VIGNETTA 1 (DESTRA), CON FIX AL PROMPT;  
TEMPO 17:00

Questa volta invece vediamo modificata completamente la posa del personaggio non rispettando il prompt poiché non più di spalle, probabilmente poiché le ulteriori specifiche sulle cicatrici in volto facevano esplicito riferimento agli occhi del personaggio. Queste cicatrici continuano comunque ad essere assenti, e i capelli risultano più corti della versione precedente nonostante l'enfasi apposta sulla loro lunghezza.

I testi delle caption presenti in queste rigenerazioni appaiono piccoli poiché non sono disegnati ma sovrapposti alla grafica, e quindi non si sono ingranditi con lo zoom effettuato direttamente sul sito, utilizzando lo slider di ingrandimento presente sulla destra in interfaccia.



FIGURA 19 - PRIMA (SINISTRA) E DOPO RIDISEGNAMENTO VIGNETTA 2 (DESTRA); TEMPO 25:00

Il redraw in questo caso è stato fatto senza modificare il testo sotto, trascritto di seguito: *grayscale, detailed drawing, japanese manga, manga, Close-up of Mahito's face as he spots Choso in the shadows. Choso is a male character in his late 30s with a rugged appearance and bags under his eyes. His hair is pulled back tightly in a bun, and he wields a long blade made of his own blood. The storehouse is dimly lit, with rays of light casting eerie shadows on the walls.*

Nonostante il modello sia riuscito a comprendere a livello sintattico/semantico le differenze tra i due personaggi, e si riferisca agli stessi tramite i loro nomi, troviamo incoerenze già solo nell'elaborazione del testo cui sopra: ci viene preannunciato che l'immagine contiene una ripresa da vicino del personaggio dal nome Mahito, e successivamente procede a descrivere invece Choso. Logicamente, questa descrizione finisce per essere quello che viene rappresentato, ossia un close-up di un personaggio con capelli legati come si nota (a fatica) nell'immagine di destra, sono presenti delle zone d'ombra più accentuate rispetto agli altri personaggi e che quindi interpreto come occhiaie, e nell'immagine di destra si percepisce anche il gonfiore delle stesse. Nelle

immagini non sono presenti né la lama di cui si parla nella descrizione, né il magazzino dello sfondo; invece troviamo sul volto del disegno a sinistra delle tracce di sangue, elemento di cui doveva essere composta la lama, mentre nel disegno a destra si notano dei solchi attorno all'occhio: non sono riuscita a comprendere se questi sono una reminiscenza del sangue presente nell'immagine precedente al redraw, oppure avendo citato il nome di Mahito nel testo ho supposto potessero essere le cicatrici che "scorrono verticalmente attraversando il suo occhio destro", descritte in questo modo nel testo della vignetta analizzata prima. Parlando dell'abbigliamento, mentre nell'immagine sulla sinistra sembri ricordare un kimono dalla stratificazione di tessuti e presenza di pieghe sugli stessi, nella versione a sinistra somiglia più al colletto di un impermeabile, ma il testo da cui è stata generata l'immagine non ha riportato le specifiche sull'abbigliamento tradizionale.

Nel complesso ritengo che il disegno a destra sembri rappresentare Wolverine, con le ombreggiature sulle guance, piene di striature che ricordano presenza di peluria, e personalmente non comprendo e non apprezzo la presenza di tali peli sul collo verso l'interno della giacca.



FIGURA 20 - PRIMA (SINISTRA) E DOPO RIDISEGNAMENTO VIGNETTA 3 (DESTRA); TEMPO 29:00

Il testo presente sotto questa immagine purtroppo presenta l'intero prompt trascritto in maniera letterale, senza rielaborazioni da parte del modello:

*grayscale, detailed drawing, japanese manga, manga, MAHITO and CHOSO are starting a fight in a storehouse. MAHITO is dressed in casual clothes, has long hair, and his face is covered in stitches. CHOSO is dressed in traditional japanese clothes, has bags under his eyes, his hair is tied up and fights with a blade made of his own blood.*

Si nota che non è stato sottoposto in alcun modo a rielaborazioni dal fatto che i nomi sono stati riportati in caps lock così come da me inseriti. Nonostante questo prompt descriva l'intera scena, il modello di generazione di immagine decide di rappresentare un solo ragazzo e mantiene anche una certa congruenza nella posa e ambientazione tra i due disegni qua sopra, pre e post redraw. Questa rappresentazione presenta quindi una fusione di tutte le caratteristiche descritte: il ragazzo ha capelli lunghi ma nell'immagine a destra sono anche legati, ha delle occhiaie, indossa indumenti tradizionali e impugna una spada. Dal braccio che si intravede impugnare la spada si notano deformazioni in termini di proporzioni e articolazione del polso. Nell'immagine a destra non si capisce

cosa è quella zona più scura presente dietro alla gamba sinistra, sproporzionata per essere una parte del pantalone in ombra ma al contempo segue in maniera troppo precisa le forme del pantalone per poter essere un lembo di tessuto svolazzante come nell'immagine a sinistra. Anche il piede nudo che si intravede nel disegno presenta deformazioni anatomiche nella posizione in cui si trova, estremamente ruotato. Nello sfondo si intuisce la presenza di persone, ma osservando attentamente non si riescono a distinguere i tratti e i soggetti.

Ho tentato di modificare il testo sotto al prompt, sostituendo alla descrizione di Mahito la rielaborazione del modello stesso, presa dal testo della vignetta 1:

*grayscale, detailed drawing, japanese manga, manga, mahito and choso are starting a fight in a storehouse. **Mahito is a male character in his early 20s with a pale complexion and stitches covering his face, running vertically across his right eye. He is dressed in casual clothes and has very long, straight hair.** CHOSO is dressed in traditional japanese clothes, has bags under his eyes, his hair is tied up and fights with a blade made of his own blood.*

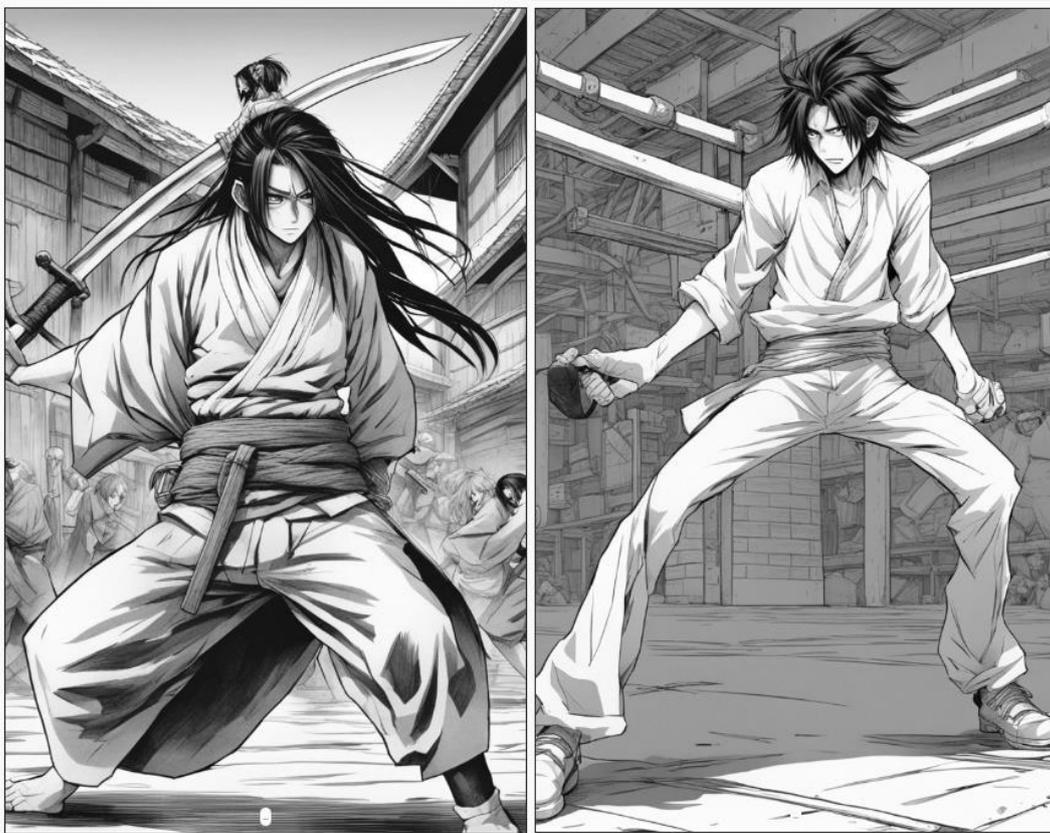


FIGURA 21 - PRIMA (SINISTRA) E DOPO ULTERIORE RIDISEGNAMENTO VIGNETTA 3 (DESTRA), CON FIX AL PROMPT;  
TEMPO 32:00

Nonostante nel testo sia stata sostituita solo la descrizione di uno dei due personaggi citati, il disegno ha subito un cambiamento drastico: da un'ambientazione esterna passiamo ad una interna, il personaggio non ha più capelli lunghi, l'abbigliamento ricorda un po' il precedente ma fuso con una camicia nell'area del colletto e delle maniche, mentre i pantaloni ricordano più uno stile casual. Le mani si riconfermano essere una massa indistinguibile di tratti confusionari, e dove prima c'era una spada ora troviamo qualcosa che immagino voglia essere l'impugnatura di un fioretto ma sembra più quella di un guinzaglio. La posa è rimasta invariata, con l'unica differenza che è stata specchiata e il personaggio è rivolto col corpo verso destra ma con lo sguardo che punta a sinistra. Le cicatrici sul volto continuano a mancare.



FIGURA 22 - PRIMA (SINISTRA) E DOPO RIDISEGNAMENTO VIGNETTA 4 (DESTRA); TEMPO 35:00

Da questo zoom dell'immagine originale sulla sinistra si vede come i balloon con testo presenti sono completamente illeggibili, noto limite di Stable Diffusion.

Il testo presente sotto questa vignetta è identico alla precedente:

*grayscale, detailed drawing, japanese manga, manga, MAHITO and CHOSO are starting a fight in a storehouse. MAHITO is dressed in casual clothes, has long hair, and his face is covered in stitches. CHOSO is dressed in traditional japanese clothes, has bags under his eyes, his hair is tied up and fights with a blade made of his own blood.*

La prima cosa che salta all'occhio è la rimozione, nel redraw, delle ulteriori suddivisioni in vignette nello sfondo, e con esse la maggior parte dei problemi di malformazioni presenti. La molteplicità delle persone raffigurate invece persiste, tutte vestite in egual maniera, e nelle capigliature vediamo rappresentati sia i capelli lunghi che quelli raccolti,

ma ancora una volta non vengono raffigurate le cicatrici sul volto del soggetto. Le deformazioni presenti rimangono in tutte le mani, alcune completamente assenti, altre sembrano reggere degli oggetti (presumibilmente else di spade); osservando attentamente i volti ci accorgiamo che purtroppo anche in questa immagine, nonostante non particolarmente densa come la versione precedente, troviamo distorsioni in prossimità degli occhi di ciascun individuo.

Infine, dopo aver esplorato e sperimentato con questo strumento, ho notato che il testo inserito nelle captions non cambia nel momento in cui viene chiesto di ridisegnare l'intera tavola o le singole vignette, ma resta identico; l'unico caso in cui viene rielaborato e riscritto è quello in cui viene modificato il prompt "Story".

## 5.2 Mage Space

Cambiato strumento, azzeriamo il timer e ritorniamo alla fase di impostazione dei parametri iniziali del nuovo strumento, Mage Space:

- Tipologia di applicazione: FAST TEXT-TO-IMAGE;
- Prompt: *comics in japanese manga style, grayscale, with four panels about two guys fighting in a storehouse. Their names are MAHITO and CHOSO. MAHITO is dressed in casual clothes, has long hair, and his face is covered in stitches. CHOSO is dressed in traditional japanese clothes, has bags under his eyes, his hair is tied up and fights with a blade made of his own blood. ;*
- Prompt Filter: *Pack 3 – Manga;*
- Aspect Ratio: Portrait.

Cercando consistenza nelle varie generazioni che andrò a fare, ho deciso di non indicare nessuna wildcard tra i prompt dinamici. Inoltre il prompt, precedentemente suddiviso in due campi diversi con AI Comic Factory, qui viene presentato in un unico testo, che comprende anche indicazioni stilistiche.



FIGURA 23 - PRIMA GENERAZIONE CON MAGE SPACE; TEMPO 01:00

Rispetto al prompt sono stati rispettati i criteri di divisione della tavola in quattro, nonostante ci siano dei punti in cui questa divisione diventa labile e si fonde con lo sfondo. Per il resto, i personaggi generati in realtà raffigurano la stessa persona ma specchiata e vista da inquadrature diverse: questa persona presenta alcuni dei criteri richiesti sui due protagonisti della scena e fusi, ha capelli sia lunghi che legati, indossa abiti tradizionali e in alcuni punti sembra avere delle occhiaie; non c'è invece traccia della lama e delle cicatrici.

In quanto ai difetti di generazione, sono stati inseriti dei balloon con del testo illeggibile, che in alcuni punti tende verso la scrittura cinese/giapponese e in altri verso quella romana, e nonostante anche questo strumento si basi su Stable Diffusion notiamo che le mani sono gestite decisamente meglio.

Detto questo, il risultato non sembra rappresentare una scena di combattimento, è alquanto noioso, statico e poco creativo.

Diversamente dallo strumento precedente, con questo non si possono fare modifiche relative alle singole vignette, per cui il processo di affinamento è stato condotto generando

diverse immagini dallo stesso prompt e vedendo come si comporta il motore di generazione; in alcuni casi è stata sfruttata la possibilità di generare una nuova immagine a partire dal seed di una generata precedentemente.



FIGURA 24 - SECONDA GENERAZIONE CON MAGE SPACE; TEMPO 02:30

Questa immagine è stata ottenuta semplicemente cliccando di nuovo sul pulsante genera, senza apportare modifiche. Possiamo notare che nelle due vignette superiori, il design è rimasto consistente con la generazione precedente, ma è più presente l'intento di rappresentare una scena di combattimento; inoltre il personaggio in alto a destra ha dei segni in volto che potrebbero essere riconducibili sia a tracce di sangue che alle cicatrici desiderate.

I problemi più pesanti sono invece presenti nel resto dell'immagine: vediamo teste tagliate dovute probabilmente al layout che il modello ha cercato di generare per la suddivisione della tavola, diverse deformazioni facciali in termini di proporzioni delle forme del viso e delle dimensioni degli occhi, e infine nell'ultima vignetta una figura indistinta sullo sfondo, che ricorda un essere umano con indosso dei vestiti più casual che tradizionali.



FIGURA 25 - TERZA GENERAZIONE UTILIZZANDO IL SEED DELLA PRIMA; TEMPO 04:50

Dato che la seconda generazione aveva presentato decisamente più problemi, ho provato a prendere il seed della prima ed effettuare una nuova, con sempre lo stesso prompt, ma basata su quella. Dal risultato si vede l'ispirazione presa dalla prima: i pannelli sono diminuiti, da quattro a due, ma sono nuovamente presenti personaggi identici ma specchiati nella posa, che differiscono solo per la colorazione dell'abbigliamento.

Diversamente dagli altri tentativi, stavolta non è stato fatto alcun tentativo verso la generazione di balloon testuali per i dialoghi, mentre le altre problematiche restano le stesse in merito a elementi mancanti, assenza di creatività, fedeltà al prompt. I volti sono stati disegnati correttamente, e le mani che tendono ad avere problemi sono fuori inquadratura.

Il prossimo risultato è stato ottenuto cambiando l'impostazione delle due tendine di Prompt Filter in: *Pack 5 – Manga*.



FIGURA 26 - QUARTA GENERAZIONE MODIFICANDO IL PROMPT FILTER; TEMPO 07:00

Avendo trovato un'altra scelta Manga tra questi stili, ho immaginato potesse usare una diversa versione dei modelli alla base dell'elaborazione del prompt e generazione dell'immagine, quindi ho voluto fare una prova. È abbastanza evidente che sia cambiato qualcosa: vi è un'elaborazione e sviluppo del prompt in maniera più articolata, i testi presenti nei balloon presentano caratteri più verosimili e vicini all'alfabeto romano, e in generale lo stile con cui vengono rappresentati i volti mi sembra diverso.

In quanto agli errori, ve ne sono di meno relativi alla suddivisione della tavola in diverse sezioni, con separazioni evidenti e definite, quasi mai labili se non in alto a destra; anche le mani, dove sono state disegnate, presentano problemi di dita aggiuntive, gonfiori, assenza di dettagli e indistinguibilità dallo sfondo.

Ma i problemi a parer mio maggiori, e a tratti disturbanti, li troviamo nei volti: dalle solite mal proporzioni e deformazioni a parti mancanti, fino ad arrivare all'intero volto assente e "rotto" nel ragazzo in basso al centro.



FIGURA 27 - ULTIMO TENTATIVO, CAMBIANDO IL SEED E IL TESTO NEL PROMPT; TEMPO 09:00

Una volta testati i limiti e i comportamenti dello strumento, ho tentato un'ultima generazione, impostando il seed dell'immagine precedente per cercare di trarre ispirazione nella suddivisione del layout in diverse immagini, e per sopperire alle mancanze tra le caratteristiche richieste nei personaggi ho modificato il prompt testuale in:

*comics in japanese manga style, grayscale, with four panels about two guys fighting in a storehouse. Their names are MAHITO and CHOSO. MAHITO is dressed in a **t-shirt and jeans**, has long **straight** hair, and his face is covered in stitches **across his eyes**. CHOSO is dressed in traditional japanese clothes, has **dark circles** under his eyes, his hair is tied up in a **pair of buns** and fights with a **sword** made of his own blood.*

In grassetto nel prompt qui sopra ho evidenziato le aggiunte. Stavolta sembrano apparire in alcuni close-up dei segni riconducibili a cicatrici, inoltre l'abbigliamento stavolta è

unicamente casual. Rimane il fatto che il modello non riesca a comprendere il fatto che le due persone rappresentate sono distinte e differenti, e finisce per concepire un singolo design fondendo vari elementi richiesti, e lo riproduce molteplici volte con inquadrature differenti.

Gli errori più visibili sono le solite deformazioni alle proporzioni dei volti, problemi nella generazione delle mani, e nelle immagini più in alto sembra che i personaggi abbiano un post-it attaccato in fronte (dettaglio non richiesto ma che mi ha fatto sorridere).

### 5.3 Valutazione finale

Le valutazioni nella tabella seguente saranno date in forma numerica su scala da 1 a 10, dove 6 è la sufficienza; questi voti sono stati dati grazie ad un sondaggio condotto su un campione di cinque persone, tutte accomunate dalla passione per il fumetto e che hanno quindi giudicato i risultati con occhio critico.

Parametro di valutazione	Ai comic factory	Mage space
Tempo impiegato	2	N.Q.
Necessità di rielaborazioni	3	2
Creatività	3	3
Fedeltà al prompt	4	1
Presenza di distorsioni	3	4
Gradimento personale	2	1

Questo test non è stato superato da nessuno degli strumenti utilizzati: nel corso dell'esposizione e analisi dei risultati ci siamo resi conto di come non si sia mai riusciti a raggiungere un risultato propriamente accettabile, neanche se inteso come disegno di un fumetto in stile manga fatto da qualcuno che non sa scrivere storie e illustrarle, poiché sono stati mancati diversi elementi descritti chiaramente nei prompt; erano presenti molte distorsioni in punti in comune tra entrambi gli strumenti, anche se il Mage Space sembrava affrontarle leggermente meglio laddove AI Comic Factory distorceva ad ogni occasione presentata; nel complesso, se anche provassimo a immaginarci un disegno

brutto, realizzato da qualcuno totalmente incapace nel mestiere come potrei essere io stessa, comunque ne risulterebbe qualcosa di integro, senza moncherini e fusioni di elementi distanti sia concettualmente che geograficamente.

Al Comic Factory risulta essere unico nel suo genere come strumento, l'unico che provi a fondere insieme diversi modelli per generare a partire da un prompt sia una storia che una sua esposizione illustrata, ma purtroppo non riesce a coniugare bene nella realizzazione della cosa, nonostante la storia in alcuni punti venga rielaborata, espansa, interpretata, riformulata, il modello di generazione di immagini non riesce a rappresentarla correttamente. Sinceramente mi aspettavo che, essendo il prompt dato al modello che genera la parte grafica una rielaborazione del prompt utente fatto dall'applicativo stesso e per l'applicativo, che questo prompt venisse costruito in maniera mirata, in modo da risultare il più comprensibile possibile a Stable Diffusion; quindi mi sarei immaginata più margine di errore nel comprendere e rielaborare il testo inserito dall'utente, piuttosto che in fase di traduzione testo-immagine.

Per quanto riguarda Mage Space, nonostante mi fosse stato proposto da varie ricerche mirate a trovare strumenti per la generazione di fumetti, ritengo che questo sia un semplice strumento di generazione immagini, che non ha nessuna particolare aggiunta o integrazione che lo renda diverso, speciale, più adatto alla generazione di un fumetto, ma invece presenta delle limitazioni ad interventi più mirati rispetto ad altri applicativi del suo genere come Playground.ai (permette di evidenziare con un lazo alcune zone e rigenerarne il contenuto, per correggere le distorsioni principalmente, ma ha pressoché gli stessi problemi nella generazione di una tavola a fumetti degli altri strumenti).

La cosa che mi ha sorpreso più di tutti, e in un certo senso deluso, è il fatto che non sono riuscita ad includere nell'esperimento lo strumento AI Comic Generator, di cui alla Figura 10, poiché aveva uno stile ancora differente dagli altri, e percepivo un qualche tipo di sforzo creativo e immaginativo maggiore, dei testi più presenti e articolati, e nonostante l'incongruenza delle immagini tra loro stesse, si poteva comunque notare una certa varietà nell'espressione; per questi motivi avrei gradito provare a condurre l'esperimento con anche questo strumento.

## 7. CONCLUSIONI

In conclusione, non è ancora possibile generare un fumetto in stile manga con intelligenza artificiale, senza un lavoro di postproduzione sia nell'elaborazione della storia che nella correzione dei disegni, in grado di rispondere agli standard qualitativi del mercato attuale (anche considerando il fumetto in commercio più brutto esteticamente per magari scarsità di esperienza dell'artista). Nell'industria è ancora necessaria la presenza di artisti umani, e tale presenza da valorizzare e non dare per scontata, poiché ritengo quella del fumetto essere un'arte in cui vada enfatizzata, e data la priorità giusta e necessaria, alla qualità piuttosto che la quantità. Qualità che l'intelligenza artificiale non è ancora riuscita a raggiungere.

Dubito che l'esito di questo esperimento sia stato negativo a causa del fatto che gli strumenti testati siano stati strumenti gratuiti, poiché anche le pubblicazioni analizzate nel Capitolo 2.1 vedevano il coinvolgimento dell'intelligenza artificiale nella scrittura e nell'ideazione dei design, più che nel disegno effettivo dell'opera: nel caso di "Black Jack", nonostante il modello utilizzato sia stato allenato con i design dell'autore originale, hanno comunque ritenuto opportuno assegnare il lavoro di disegno a degli artisti umani; "Fastwalkers" e "The Abolition of Man" vedono invece pubblicati dei disegni generati con IA ma che non tentano di simulare uno stile ben preciso e netto, pulito, come quello del manga, ma nel primo troviamo uno stile in bianco e nero, molto sopra le righe, volutamente distorto e confusionario, nel secondo invece dei disegni più pittoreschi, piacevoli se osservati da lontano ma non esenti da stranezze per occhi più attenti. Va tenuto di conto che anche dove si è usata l'intelligenza artificiale per illustrare, le immagini sono state generate singolarmente e assemblate da mano umana in una fase successiva, altro step che l'IA non è in grado di automatizzare.

In questa tesi non è stato trattato il punto di vista etico sulla generazione di immagini con intelligenza artificiale poiché ho preferito dare più spazio alla parte tecnica e oggettiva, ma personalmente mi trovo in disaccordo con la direzione che questi strumenti stanno prendendo, non tanto per l'IA in sé ma per l'uso che se ne fa.

Nonostante ad oggi alcuni di questi strumenti sono ancora acerbi e non possono interamente sostituire l'uomo in questo settore, credo che nel futuro prossimo si

continuerà a insistere e investire sul coinvolgimento dell'IA nell'industria manga e anime e, considerando i risultati delle ricerche accademiche presentate sopra (Capitolo 2.3), questo futuro potrebbe essere più vicino di quanto pensiamo. I motivi che mi spingono a credere che il mercato cavalcherà questa scia sono due: il primo è dato dal fatto che l'industria manga e anime sta prendendo sempre più piede in occidente, facendo salire la domanda e rendendo quindi sempre più difficile stare al passo per gli artisti che ci lavorano, finendo per trovarsi in condizioni deprecabili; il secondo consta nei guadagni che gli strumenti di generazione di immagini possono fruttare davanti a una società che vuole sempre di più, sempre più in fretta, sempre al miglior prezzo sul mercato.

## 8. FONTI BIBLIOGRAFICHE E SITOGRAFIA

AI Comic Factory webapp [<https://huggingface.co/spaces/jbilcke-hf/ai-comic-factory>]

Ai DIY, *Boost Your Instagram with AI Art: A Warning About Plugger.ai | AI DIY Tutorial*, canale “Ai DIY”, 24 Marzo 2024 [<https://www.youtube.com/watch?v=xDgNFU7BvjA>]

Angelique G., *AI in Manga: Transforming Japanese Comics*, in culture.org, 7 Marzo 2024 [<https://culture.org/art-and-culture/literature/ai-in-manga/>]

Babb T., *The son of manga legend Osamu Tezuka is using AI (and human creators) to continue Tezuka's classic manga Black Jack*, in “POPVERSE”, 12 Giugno 2023 [<https://www.thepopverse.com/osamu-tezuka-productions-ai-human-collaboration-manga-black-jack>]

Balaji Y., Nah S., Huang X., Vahdat A., Song J., Zhang Q., Kreis K., Aittala M., Aila T., Laine S., Catanzaro B., Karras T., Liu M., *eDiff-I: Text-to-Image Diffusion Models with an Ensemble of Expert Denoisers*, arXiv:2211.01324, data ultima modifica 14 Marzo 2023

Bilcke J., “AI Comic Factory” Github repository [<https://github.com/jbilcke-hf/ai-comic-factory>]

Bilcke J., *Deploying the AI Comic Factory using the Inference API*, in “Hugging Face”, 2 Ottobre 2023 [<https://huggingface.co/blog/ai-comic-factory>]

Catacchio O., *L’AI ora sa creare i fumetti: che ne sarà dei comics dopo “The Abolition of Man”*, in “AgendaDigitale.eu”, 20 Luglio 2022 [<https://www.agendadigitale.eu/cultura-digitale/lai-ora-sa-creare-i-fumetti-che-ne-sara-dei-comics-dopo-the-abolition-of-man/>]

Dheda G., *AI Comic Factory by Hugging Face: Review and Tips for Better Consistency*, in “Open AI Master”, 16 Novembre 2023 [<https://openaimaster.com/ai-comic-factory-by-hugging-face-review/>]

Gose L. P., *Black Jack Unveils First Official AI-Generated Manga – To the Horror of Tezuka Fans*, in CBR.com, 21 Novembre 2023 [<https://www.cbr.com/black-jack-osamu-tezuka-ai-generated-manga-fan-reaction/>]

Grubaugh C., pagina prodotto “*The Abolition of Man*”, dal sito dell’editore “*Living the Line*”, 14 Giugno 2022 [<https://www.livingthelinebooks.com/books/abolition-of-man>]

Henry J., *The Abolition of Man is the First Comic From A.I.*, in “*The Nerd Stash*”, 14 Giugno 2022 [<https://thenerdstash.com/the-abolition-of-man-is-the-first-comic-from-a-i/>]

Johnston R., *Abolition Of Man, First Comic Book Entirely Drawn By A.I. Algorithm*, in “*Bleeding Cool*”, 14 Giugno 2022 [<https://bleedingcool.com/comics/abolition-of-man-first-comic-book-entirely-drawn-by-a-i-algorithm/>]

Kataoka Y., Matsubara T., Uehara K., *Automatic Manga Colorization with Color Style by Generative Adversarial Nets*, pubblicato in “*2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*”, pp. 495-499, DOI: 10.1109/SNPD.2017.8022768, 26-28 Giugno 2017

Kuroda K., *AI helps author new episode for Tezuka’s ‘Black Jack’ manga*, in “*The Asahi Shinbun*”, 21 Novembre 2023 [<https://www.asahi.com/ajw/articles/15062886>]

Liu S., *Mega Man Manga Artist Worked on AI Black Jack Chapter*, in “*Siliconera*”, 21 Novembre 2023 [<https://www.siliconera.com/mega-man-manga-artist-worked-on-ai-black-jack-chapter/>]

Lo spazio bianco, *D Editore pubblica “Fastwalker”, il fumetto del conceptual comics Ilan Manouach*, in “*Lo spazio bianco*”, 28 Novembre 2022 [<https://www.lospaziobianco.it/d-editore-pubblica-fastwalker-il-fumetto-del-conceptual-comics-ilan-manouach/>]

Manouach I., pagina prodotto “*Fastwalkers*”, dal sito dell’editore “*D Editore*” [<https://deditore.com/prodotto/fastwalkers/>]

Meng C., He Y., Song Y., Song J., Wu J., Zhu J., Ermon S., *SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations*, arXiv:2108.01073, data ultima modifica 5 Gennaio 2022

Playground AI [<https://playground.com/>]

Podell D., English Z., Lacey K., Blattmann A., Dockhorn T., Müller J., Penna J., Rombach R., *SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis*, arXiv:2307.01952, 4 Luglio 2023

Rafailov R., Sharma A., Mitchell E., Ermon S., Manning C. D., Finn C., *Direct Preference Optimization: Your Language Model is Secretly a Reward Model*, arXiv:2305.18290, data ultima modifica 13 Dicembre 2023

Robinson S.M., *Abolition of Man, World's First AI Illustrated Comic Book, Returns in Signing and Talk*, in "Patch", 29 Settembre 2023 [<https://patch.com/connecticut/newhaven/calendar/event/20230929/f842a088-b9b1-4225-a870-42953d4013a5/abolition-of-man-worlds-first-ai-illustrated-comic-book-returns-in-signing-and-talk>]

Stable Diffusion XL 1.0 API [<https://huggingface.co/stabilityai/stable-diffusion-xl-base-1.0>]

Su H., Niu J., Liu X., Li Q., Cui J., Wan J., *MangaGAN: Unpaired Photo-to-Manga Translation Based on The Methodology of Manga Drawing*, pubblicato in "Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 35 No. 3: AAAI-21 Technical Tracks 3", pp. 2611-2619, DOI: <https://doi.org/10.1609/aaai.v35i3.16364>, 18 Maggio 2021

The Japan News, *AI Version of Osamu Tezuka's 'Black Jack' Manga Series Planned*, in "The Japan News", 13 Giugno 2023 [<https://japannews.yomiuri.co.jp/culture/manga-anime/20230613-115940/>]

The Japan News, *AI Helps Create New 'Black Jack' Manga Episode for Osamu Tezuka's Iconic Maverick Doctor*, in "The Japan News", 21 Novembre 2023 [<https://japannews.yomiuri.co.jp/culture/manga-anime/20231121-150901/>]

Zephyr 7B  $\beta$  API [<https://huggingface.co/HuggingFaceH4/zephyr-7b-beta>]

Zhang L., Wang X., Fan Q., Ji Y., Liu C., *Generating Manga From Illustrations via Mimicking Manga Creation Workflow*, pubblicato in "2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)", pp. 5638-5647, DOI: 10.1109/CVPR46437.2021.00559, 20-25 Giugno 2021

## RINGRAZIAMENTI

*Se sono arrivata alla fine del mio percorso di studi è solo grazie alle varie persone che mi hanno accompagnato in questo viaggio.*

*Grazie al Professore Angelo Di Iorio, che ha accolto con entusiasmo la mia proposta di tesi e mi ha compresa, rassicurata e supportata verso il raggiungimento di questo risultato.*

*Grazie a Davide, Mauro, Matteo, Erasmo e Alessandro, per il prezioso contributo fornito in quanto campione prezioso in fase valutativa dei risultati dell'esperimento.*

*Grazie alla mia famiglia, per avermi sopportata e supportata in questi lunghi anni di dubbi e indecisioni, ma anche persistenza e passione.*

*Grazie ai miei amici, di Bologna e del web, in particolare Davide, Letizia e Choosy, che hanno dovuto gestirmi e rallegrarmi nei miei momenti più bassi, e che sono riusciti a farmi rinsavire ogni volta che pensavo di aver toccato il fondo.*

*Grazie ai miei colleghi di Next, per sopportarmi nella mia instabilità emotiva anche sul posto di lavoro, e per avermi fornito dei punti di vista sul percorso che ho affrontato, da persone che ci sono passate prima di me.*

*Grazie alla mia psicologa, la Dott.ssa Marika Tedesco, per avermi aiutato a riflettere su me stessa negli ultimi tre anni, e a capire come aggirare gli ostacoli che ho trovato lungo la strada per questo traguardo.*

*Grazie a Mellie, alla sua musica e alla sua persona, in grado di infondermi speranza e emozioni per tutto quello che ancora non so della vita, e che non vedo l'ora di scoprire.*

*Grazie a tutte le persone che non ci sono più, senza le quali non sarei diventata la persona che sono oggi.*

*Grazie alle mie passioni, che mi hanno aiutato a respirare tra uno sforzo e l'altro, diventate anche protagoniste di questa tesi.*

*L'ultimo grazie va al Futuro, per esistere, per avermi acceso la voglia di viverlo.*