

BILDER GENERERADE MED HJÄLP AV

ARTIFICIELL INTELLIGENS

Alexander Winberg

Kandidatavhandling i datavetenskap

Handledare: Annamari Soini

Fakulteten för naturvetenskaper och teknik

Åbo Akademi

ABSTRAKT

AI-program som genererar bilder har exploderat i popularitet sedan de blev allmänt tillgängliga 2022. För att förstå hur dessa program kan generera bilder är det viktigt att förstå den underliggande tekniken och hur programmen är uppbyggda. Det är också synnerligen viktigt att förstå hur den data som används för att träna programmen är insamlad och hur den används. AI-program genererar i princip vad som helst, det är upp till användaren att bestämma hurdan indata som ges, detta behandlas i avhandlingen samt också denna teknologis potentiellt destruktiva resultat.

Sökord: AI, artificiell intelligens, bildgenerering, Midjourney, Dall-E 2, NightCafé

ABSTRAKT	1
1. INLEDNING	3
2. MASKININLÄRNING	4
2.1. ÖVERVAKAD INLÄRNING	4
2.2. OÖVERVAKAD INLÄRNING	5
2.3. FÖRSTÄRKNINGSINLÄRNING	5
2.4. ARTIFICIELLA NEURONNÄTVERK	6
2.5. DJUPINLÄRNING	8
2.6. DATA	8
2.7. TRÄNING	9
2.8. BRUS	12
2.9. SPRÅKTEKNIK	15
3. TILLÄMPNING I PRAKTIKEN	17
3.1. MIDJOURNEY	18
3.2. DALL-E 2	19
3.3. STABLE DIFFUSION / NIGHTCAFÈ	20
3.4. BILDGENERERING	22
3.4.1. KIRBY OCH SONIC I MIDJOURNEY	23
3.4.2. KIRBY OCH SONIC I DALL-E 2	24
3.4.3. KIRBY OCH SONIC I NIGHTCAFÉ	25
3.4.4. SLOTT I MIDJOURNEY	27
3.4.5. SLOTT I DALL-E 2	28
3.4.6. SLOTT I NIGHTCAFÉ	29
4. AI GENERERADE BILDER OCH SAMHÄLLET	31
4.1. AI GENERERAD PORNOGRAFI	33
4.2. AI GENERERAD PROPAGANDA	35
5. SLUTSATSER	36
KÄLLFÖRTECKNING	37

1. INLEDNING



Bild 1 – no titel (Nake 1967)

Bild 1 föreställer verket no titel av Frieder Nake, en av de första bilder som genererats med hjälp av ett datorprogram. Syftet med avhandlingen är att undersöka hur AI-program genererar unika bilder idag. AI-program har sedan tidigare kunnat utföra många olika uppgifter, som till exempel översättning och bildigenkänning. För detta används i regel maskininläring som är ett delområde inom artificiell intelligens. Maskininläring ligger också till grund för artificiella neuronätverk, det är dessa nätverk som används av AI-program för just bildigenkänning men också att generera bilder. Vad jag eftersträvar är att granska hur maskininläring och artificiella neuronätverk byggs upp och hur de tillämpas.

Teorin i avhandlingen avgränsar sig främst till just maskininläring och den data och de metoder som behövs för att framgångsrikt implementera maskininläring.

I den praktiska delen av arbetet granskar jag tre olika AI-program och de bilder som dessa program kan generera och jämför dem dels med varandra, dels med vad man kan förvänta sig av bilderna.

Slutligen vill jag begrunda hur denna teknologi påverkar samhället och de risker den för med sig.

2. MASKININLÄRNING

Syftet med maskininläring som ett vetenskapligt delområde inom artificiell intelligens är att konstruera mjukvara som automatiskt kan förbättra sitt resultat med hjälp av erfarenhet (Mitchell 1997). I artikeln *Some Studies in Machine Learning Using the Game of Checkers*, publicerad 1959, presenterade Arthur Samuel hur ett datorprogram skulle kunna lära sig spela dam och för varje spel förbättra sin prestation (Samuel 1959). Samuel beskrev hur ett program kunde lära sig att hitta vinnande strategier för spelet utgående endast från spelets regler, rudimentära förhållningsorder och en bristfällig lista över direktiv för spelet. För att uppnå fungerande maskininläring ansåg Samuel att följande kriterier måste uppfyllas.

1. Aktiviteten kan inte vara deterministisk.
2. Ett klart mål måste vara definierat.
3. Klara regler måste finnas och de måste vara kända.
4. Det bör finnas användbara data som kan användas vid testande av inlärningsprocessen.
5. Aktiviteten bör vara känd så att dess resultat kan mätas.

Dessa enkla regler har sedan dess utvecklats, men de utgör ännu kärnan inom maskininläring.

Idag delas maskininläring vanligen in i följande tre huvudkategorier; övervakad inläring, oövervakad inläring och förstärkningsinläring (Theobald 2018).

2.1. ÖVERVAKAD INLÄRNING

Theobald beskriver i *Machine Learning for Absolute Beginners* att övervakad inläring innebär att programmet tränas på kategoriserad källdata och måldata, syftet är att finna kopplingar mellan dessa två (Theobald 2018). På detta sätt bygger programmet upp regler för hur ett givet mål kan nås utgående från den indata som ges. Som exempel kan programmet tränas på data som föreställer rektanglar, cirklar, trianglar och hexagoner. Data som används är klart markerad så att varje form har det korrekta namnet, då lär sig programmet känna igen dem korrekt. Om den indata som sedan ges föreställer hexagoner och cirklar kommer programmet att kunna separera och identifiera dem korrekt.

Övervakad inläring tillämpas bland annat för att träna program som använder sig av datorseende att känna igen olika objekt (IBM n.d.). Ett annat användningsområde är artificiella neuronätverk, som tas upp i stycke 2.4.

2.2. ÖÖVERVAKAD INLÄRNING

Öövervakad inläring skiljer sig från öövervakad inläring då varken källdata eller måladata är kategoriserad och de potentiella förhållandena mellan källdata och måladata är okända (Theobald 2018). Målet för programmet är att hitta mönster i datamängderna och på så sätt kunna skapa strukturer.

Denna metod kan till exempel användas i kundsegmentering då ett företag vill dela in sina kunder i olika grupper. Om företaget vill sända ut reklam om ett erbjudande om billigare energidrycker kan det vara mera effektivt att först dela upp kunderna utgående från deras intresse i den sorts erbjudande. Då kan företaget vara säkert på att bara de som kan tänkas köpa energidrycken får erbjudandet och att effekten då blir större samt att det kan vara billigare än att sända ut erbjudandet till alla kunder.

2.3. FÖRSTÄRKNINGSINLÄRNING

Kaelbling, Littman och Moore (1996) beskriver förstärkningsinläring som en metod som vare sig kräver att källdata och måladata är kategoriserad, eller att deras förhållanden är etablerade. Målet är i stället för programmet att via utforskning utnyttja den kunskap som programmet besitter för att nå det bästa tänkbara resultatet. I takt med att programmets kunskap ökar kommer också dess förståelse att öka och resultatet blir gradvis bättre.

Förstärkningsinläring kan till exempel användas inom spel, exempelvis programmet MuZero från DeepMind tillämpar det (Schrittwieser 2020). MuZero spelar bland annat go, schack och shogi och har lärt sig att bemästra spelen utan att ursprungligen känt till reglerna. I stället har programmet provat sig fram och lärt sig vilka strategier i varje given situation som mest sannolikt är vinnande.

2.4. ARTIFICIELLA NEURONNÄTVERK

Precis som maskininläring är artificiella neuronnätverk ett gammalt koncept, redan 1944 beskrev Warren McCullough och Walter Pitts grunderna för hur ett artificiellt neuronnätverk kunde fungera (Hardesty 2017). Ett artificiellt neuronnätverk är i grund och botten ett försök att konstruera en kopia av de biologiska neuronnätverk som ligger som grund för hjärnans verksamhet.

Artificiella neuronnätverk använder övervakad inlärning, programmet vet med andra ord hur det önskvärda resultatet ser ut men inte hur målet skall nås.

Artificiella neuronnätverk består av ett visst antal noder, antalet kan variera mellan ett fåtal och miljarder, som är uppbyggda i lager ovanpå varandra. Varje nod har ett visst värde och då en input sänds ut i nätverket, som en signal, förstärks eller försvagas dess värde beroende på hur noderna påverkar signalen (Hardesty 2017). På detta sätt lär sig sedan programmet att känna igen mönster, och också att hitta nya mönster som inte varit bekanta för användaren. Den kanske vanligaste formen av artificiella neuronnätverk är bakåtpropageringsnätverk, vilket innebär att programmet inte enbart skickar signalen framåt. Om det behövs kan signalen skickas bakåt för att korrigera inställningarna inuti nätverket och på så vis nå det önskade resultatet.

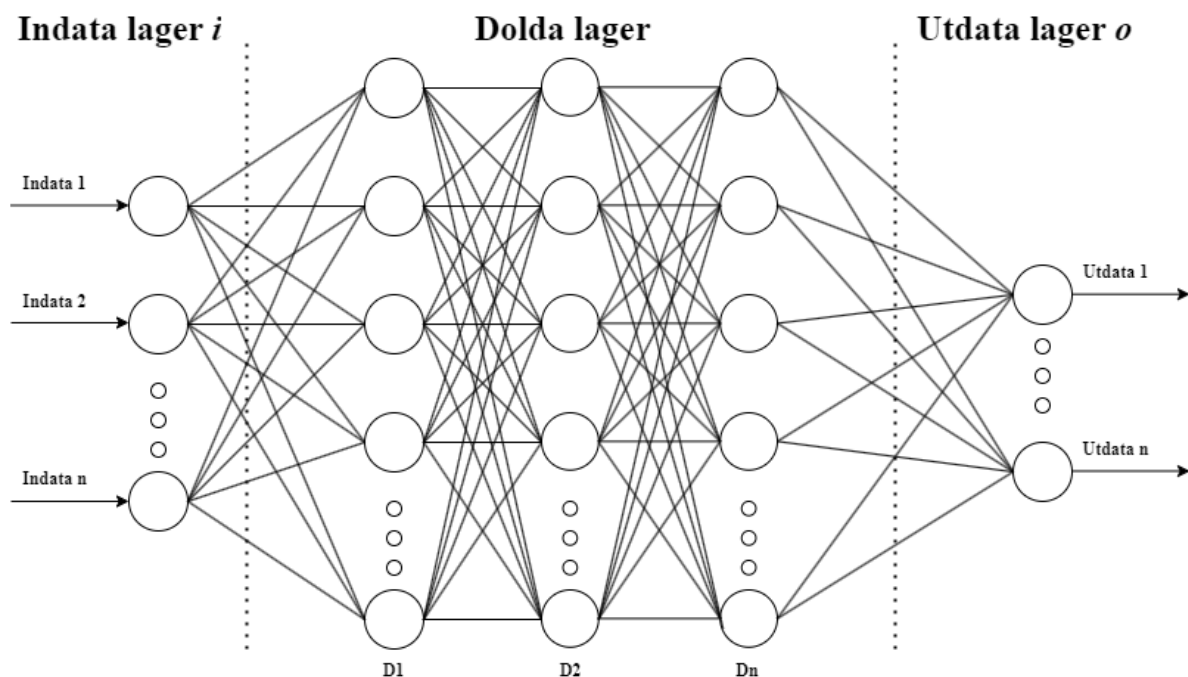


Bild 2 - artificiellt neuronnätverk

Neuronnätverket i bild 2 består av ett lager för indata, ett antal dolda lager och slutligen ett utdatalager. I indatalagret tas den givna inputen emot, det kan då vara data i olika former som exempelvis ljud, bilder eller text. Efter att inputdata har behandlats enligt angivna krav för programmet sänds den vidare till de dolda lagren. De dolda lagren kan vara så få som ett, men vanligen består de av tre eller flera lager. I dessa lager tas sedan inputen emot, behandlas och skickas vidare till nästa lager, eller ifall det är ett bakåtpropageringsnätverk kan data

skickas tillbaka ett eller flera lager. Slutligen når data utdatalagret, som kan bestå av en eller flera noder beroende på programmets syfte (AWS n.d.).

Artificiella neuronätverk används bland annat för att känna igen ansikten, tolka handskrivna texter men också för att till exempel förhindra kreditkortsbedrägeri. Ett nätverk lär sig hur en kreditkortsinnehavare vanligen använder sitt kort och kan då alarmera om kortet används på ett ovanligt sätt, till exempel om stora uttag görs fort efter varandra.

Ett annat exempel är inom medicinsk forskning, 2019 upptäckte forskare vid Massachusetts Institute of Technology att molekylen halicin hade hittills upptäckta användningar som antibiotika. Denna upptäckt gjordes med hjälp av en tillämpningsmetod av artificiella neuronätverk som kallas djupinlärning (MIT News 2020).

2.5. DJUPINLÄRNING

Om artificiella neuronätverk är ett ämne inom maskininlärning så är djupinlärning ett ämne inom artificiella neuronätverk (Kavlakoglu 2020). Det som främst skiljer dem åt är att i djupinlärning läggs ett större fokus på de dolda lagren i nätverket. I bild 2 har nätverket minst tre lager, D1, D2 och Dn som kan antas vara ekvivalent med D3, om då mängden lager ökade skulle nätverket klassas som en form av djupinlärning.

Enligt LeCun et al. (2015) används djupinlärning inom bland annat bildigenkänning, översättning och medicinsk forskning.

2.6. DATA

För att träna ett fungerande artificiellt neuronät krävs data. Då det artificiella nätverkets målsättning är att träna ett AI-program att generera bilder krävs data i form av bilder med text som beskriver bilden. Det är oklart exakt hur mycket data som ingår i databaserna de olika AI-programmen tränas på, till exempel OpenAI har meddelat att programmet DALL-E 2 är tränat på hundratals miljoner bilder utan att gå in på exakt hur dessa bilder har valts ut (Baio 2022). Däremot har Stability AI hävdats (Stability.ai n.d.) att deras programvara, Stable Diffusion, först tränades på databasen LAION-Aesthetic som är en filtrerad version av LAION 5B, som i sin tur består av över 5.85 miljarder par av bilder och texter. Eftersom LAION-Aesthetic är en öppen databas är det då möjligt för användare att närmare bekanta sig med innehållet (Baio 2022).

En av organisationerna som skapat dessa databaser är Common Crawl, som är en ideell organisation vars syfte är att just tillhandhålla öppna data, däribland just den data som används i LAION-Aesthetic, sammanställd av LAION. LAION är en tysk ideell organisation som sammanställer databaser (Common Crawl n.d.). Common Crawl samlar in data genom att skrapa (eng. *scrape*) webbsidor i jakt på bilder, i detta syfte använder Common Crawl programmet CCBot (Common Crawl n.d.) som är en spindel (eng. *web crawler*) baserad på mjukvaran Nutch. På detta sätt kan Common Crawl skapa databaser som innehåller bilder kombinerade med den text de ursprungligen hörde ihop med. Som exempel se bild 3 som är tagen från databasen LAION-Aesthetic, bredvid bilden ses dess ursprungliga namn och en kort länk till dess ursprung (LAION-Aesthetic n.d.).

För att göra databaserna mera användarvänliga filtreras bilderna enligt specifika krav. LAION-Aesthetic klassificerar bilderna först utgående från språk, och sedan separeras bilderna enligt deras resolution, sannolikheten för att de är märkta med en digital vattenstämpling och bildens sannolika estetiska poäng (Baio 2022). Detta leder till att just LAION-Aesthetic till stor del består av akvarellbilder av landskap och kvinnor då främst dessa får höga estetiska poäng på grund av de krav som ursprungligen ställts. Andra databaser kan ha ett annat fokus.

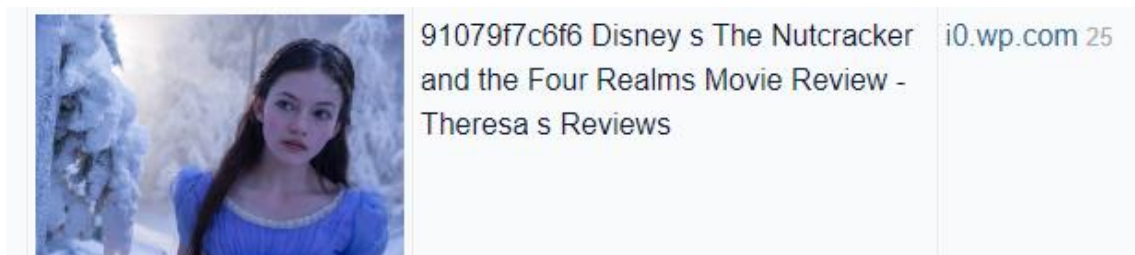


Bild 3 - exempel på LAION-Aesthetic (LAION-Aesthetic n.d.)

Det är vid skrivandet av denna avhandling (våren 2023) ännu oklart om det är lagligt i USA och EU för företag att samla in data på detta sätt (Vincent 2022). Common Crawl uppger på sin hemsida att det är möjligt för webbsidor att blockera CCBots åtkomst till sidan, men det bör påpekas att ansvaret i detta fall ligger på webbsidans ägare i stället för på Common Crawl. Utan att i någon större grad spekulera om de juridiska implikationerna kan det ändå konstateras att då EU:s direktiv vad gäller användningen av webbkakor specificerar att samtycke bör ges till deras användning (EUR-Lex 2020), är det rimligt att anta att samtycke också bör ges till skrapning.

2.7. TRÄNING

Tidigare modeller av text till bild genererande AI program använde sig i regel av generativa motståndsnätverk (GAN), vilket är en form av djupinlärning (Hypotenuse AI n.d.). Däremot använder sig nyare program som Stable Diffusion av diffusionsmodeller, också kallad probabilistisk diffusionsmodell, som är en klass av latentvariablamodeller.

Som Zylinska beskriver i AI Art bygger ett generativt motståndsnätverk på att två neurala nätverk tävlar med varandra (Zylinska 2020). Det ena nätverket, det generativa, har som mål att skapa en övertygande bild, utgående från given input, och det andra nätverket, diskriminatoren, har som syfte är att avgöra om bilden är övertygande enligt ett falskt-sant kriterium. Detta innebär att nätverken, bildligt talat, bollar bilden fram och tillbaka mellan sig tills bilden blir godkänd. Medan detta pågår lär sig också de båda nätverken att bli bättre på att dels generera bilder, dels på att avgöra vad som förväntas av bilden i fråga. En nackdel med generativa motståndsnätverk är att de inte lämpar sig för jämförelser, vilket gör det svårt att utvärdera dem (Song and Ermon 2019). Det finns också en risk för så kallad lägeskollaps, vilket innebär att nätverket i värsta fall kontinuerligt returnerar samma resultat, oavsett vilka krav som ställs (Srivastava 2017). En fördel är dock att de är relativt snabba.

Bild 4 är ett exempel på en bild skapad av Mario Klingemann med hjälp av GAN (Vincent 2019).



Bild 4 - Memories of Passersby I (Klingemann 2018)

Diffusionsmodeller skiljer sig från generativa motståndsnätverk genom att använda sig av en radikalt annorlunda metod (Hypotenuse AI n.d.). Diffusionsmodeller har egentligen först blivit allmänt populära 2021, främst tack vare deras tillämpning för just generativa AI-program (Croitoru, et al. 2022). Författarna beskriver tillämpningen av diffusionsmodeller enligt följande kriterier; programmet lär sig först att känna igen potentiella relationer mellan bilden och texten. Genom att sedan stegvis lägga till brus (eng. *noise*) dekonstruerar programmet bilden och lär sig då vilka element, som träd och hus, bilden är uppbyggd av. Slutligen lär sig programmet att från endast brus konstruera en bild som uppfyller de kriterier som ges av användaren.

För att AI-programmet skall kunna hitta relationer mellan bilden och texten är det då viktigt att databasen består av båda delarna, och som CCBot visade är detta en del av vad dessa spindlar skrapar. Första steget för programmet är att hitta relationerna och lära sig av dem, exempelvis bild 5 visar en bild med texten ”*Mexican woman selling oranges*”.



Mexican woman selling oranges

static.wixstatic.com 89

Bild 5 - exempel på LAION-Aesthetic (LAION-Aesthetic n.d.)

I all enkelhet kan man konstatera att programmet lär sig att det är en kvinna i bilden. Därför är en klart beskrivande text till fördel i detta stadium, och vilket också klargör varför bilderna i ett tidigare stadium har filterats utgående från språket i texten. Även om det är möjligt att ha ett multilingvistiskt artificiellt neuronät skulle detta sannolikt kunna innebära att nätverkets komplexitet mångdubblas då antalet neuroner skulle öka.

2.8. BRUS

Brus uppstår i en digital bild då någonting har gått fel i bilden, med andra ord om en pixel inte har det korrekta värdet. Brus uppstår i en bild på grund av fel då bilden har tagits, till exempel på grund av dålig belysning, i överföringen eller så introduceras det då signalen förvärvas (Chen, Hung and Zou 2017). Vad det i praktiken innebär är att bilden då ser ut att innehålla fel, som i bild 6 (Hilary 2022).



Bild 6 - exempel på brus (Hilary 2022)

Brus kan vara av flera olika typer, impulsbrus som är slumpmässigt utspritt över bilden och leder till att pixlarna får missvisande färgvärden eller som leder till vad som kallas salt och peppar-brus då pixlarna slumpmässigt ges värden 255 (vitt) eller 0 (svart). Impulsbrus kan vidare vara av två olika typer: bestämt värderat impulsbrus eller slumpmässigt värderat impulsbrus.

En annan form av brus kan uträknas genom att tillämpas en Gaussiskt distribuerad kurva på bruset.

$$P_G(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

Gaussiskt brus är då sannolikheten för att en enskild pixel skall påverkas ofördelaktigt av dess gränsande pixlar.

För att ett program som tillämpar en probabilistisk diffusionsmodell skall kunna skapa nya bilder behövs Gaussiskt brus. Efter att programmet lärt sig av relationerna mellan bild och text introducerar det gradvis mer och mer Gaussiskt brus in i varje bild, en så kallad framåtriktad diffusionsprocess. Enligt följande formel kan vi utläsa att X_0 förlorar sina utmärkande särdrag i takt med att t växer, vilket slutligen innebär att då T närmar sig oändligheten blir X_T i det närmaste ekvivalent med isotropisk Gaussisk distribuering (Weng 2021). Detta innebär att variansen är enhetligt distribuerad över alla dimensioner, vilket kan beskrivas som totalt brus. Detta kan ses i bild 7 då x närmar sig T .

$$q(X_{1:T} | X_0) = \prod_{t=1}^T q(X_t | X_{t-1})$$

Formeln kan också visualiseras på följande sätt där vi ser hur bilden förändras från X_0 till X_T (Karagiannakos and Adaloglou 2022).

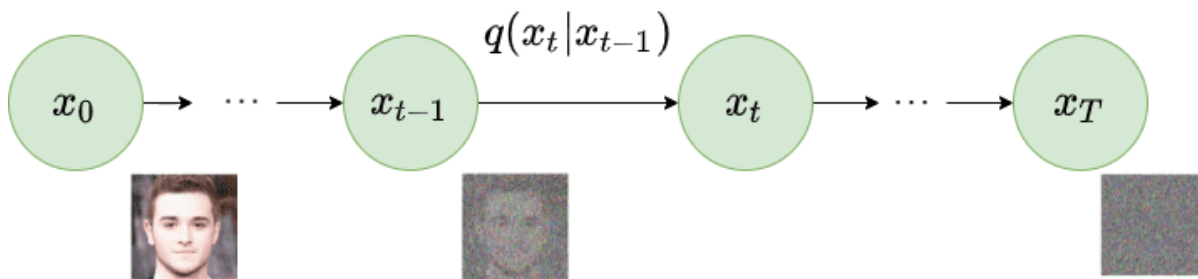


Bild 7 - exempel på brus (Karagiannakos and Adaloglou 2022)

Med hjälp av den framåtriktade diffusionsprocessen lär sig programmet vad varje enskild bild i dess databas består av på en extremt låg nivå. Nackdelen är dock att modellen är långsam eftersom den använder en Markov-kedja i uträkningen.

$$p_0(X_{0:T}) = p_0(X_T) \prod_{t=1}^T p_0(X_{t-1} | X_t)$$

Följande steg för programmet är att lära sig återskapa bilder från total isotropisk Gaussisk distribution, som är en omvänd diffusionsprocess. Om den matematiska formeln vänds för samtliga steg $p_0(X_{0:T})$ är det möjligt att nå en punkt där programmet vid varje enskilt steg t kan förutse följande steg.

I följande bilder (bild 8 till bild 12) genererade av mig i NightCafé ser man hur processen går till, från en bild av endast brus nås ett slutligt alternativ.

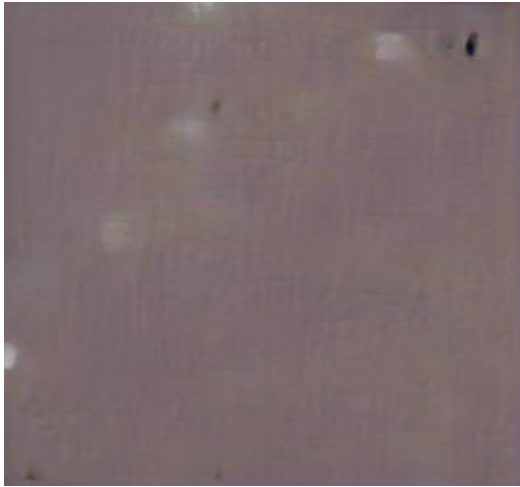


Bild 8 - steg ett



Bild 9 - steg två



Bild 10 - steg tre



Bild 11 - steg fyra



Bild 12 - steg fem

På detta sätt skapas en modell som låter AI-programmet lära sig relationer, vad bilden innehåller genom tillämpningen av brus och hur den skapar en bild från brus. Det kan konstateras att det är extremt viktigt att programmet använder sig av en stor databas som grund, för ett acceptabelt resultat är det viktigt att databasen antingen har ett klart fokus som LAION-Aesthetic, eller har ett tillräckligt brett urval för att kunna täcka önskemål som i princip kan vara näst intill oändliga i sin variation.

2.9. SPRÅKTEKNIK

För att ett AI-program skall kunna generera en bild måste det finnas en input i form av ord som beskriver det önskade resultatet. Mängden av rekommenderade ord varierar mellan programmen och också typen av ord. Eftersom genererande AI-program inte är deterministiska finns det inget som garanterar att samma ord kommer att ge samma resultat, snarare motsatsen. Följande bilder (bild 13 och 14) var båda genererade av mig i NightCafé, utgående från Stable Diffusions modell och med följande text som input; ”*sunshine on a snow covered field, mountains on the horizon*”.



Bild 13

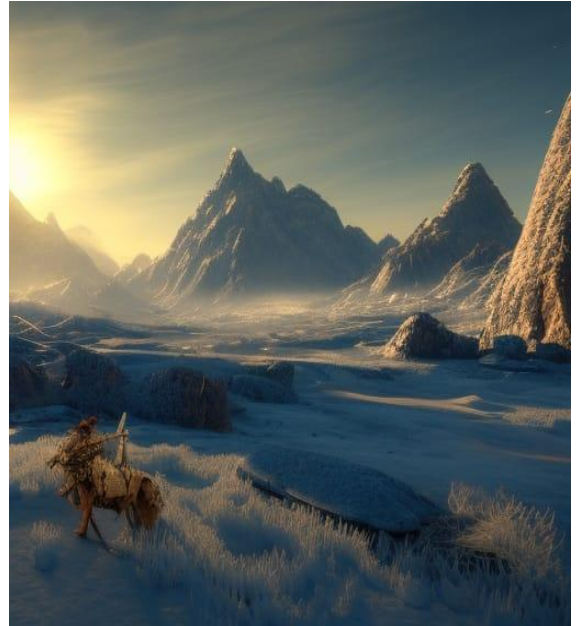


Bild 14

För att ett AI-program skall kunna behandla de ord som ges som input måste orden först omformateras och för att detta skall kunna göras effektivt används i regel språkteknik (eng. *Natural Language Processing*). Det första som måste göras för att behandla den givna inputtexten är förbearbetning, som delas in i två processer: dokument triage och textsegmentering (Indurkha and Damerau 2010).

I dokument triage behandlas den givna texten för att göra den läsbar för en maskin, detta kan bestå av identifiering av vilka tecken som används, vilket språk som tillämpas och utrensning av icke-önskvärda aspekter som till exempel bilder och hyperlänkar. Denna del tycks inte tillämpas i någon större utsträckning i till exempel NightCafé då dess användargränssnitt endast tillåter ASCII-tecken, man kan inte lägga till bilder som textinput utan det görs skilt i en annan process och den fungerar endast på engelska.

Vid textsegmentering behandlas sedan den bearbetade texten vidare i ett antal processer. Stoppord som till exempel ”är”, ”den” och ”i” kan komma att raderas i detta skede, men det beror på hur programmet värderar dem eftersom det är möjligt att dessa ord innehåller relevant information för meningens struktur. Beroende på kraven som ställs på AI-programmet kan också ord som anses vara olämpliga tas bort. Till exempel Stable Diffusion tar bort ord som ”nazi”, ”gore” och diverse sexuella ord (Vincent 2022). Kommersiella produkter har i de flesta fall diverse inbyggda spärrar som sällar ut ord som kunde komma att vara problematiska för företaget som tillverkade programmet. Exempel inkluderar extremt våld, rasism, sexuellt våld och pornografi. Vilka ord som sällas ut som förbjudna kan också påverkas av nationell lagstiftning, till exempel är det i Tyskland förbjudet av visa symboler anknytta till nazism och i ett flertal arabländer är det förbjudet att visualisera profeten Muhammed (McManu 2015). Detta är något som just NightCafé har tagit i beaktning i en viss mån då det inte går att skapa bilder på Adolf Hitler och Josef Stalin, dock är det ännu i mars 2023 möjligt att skapa bilder på profeten Muhammed.

Slanguttryck orsakar också delvis problem då AI-programmet kan vara okunnig om ordets betydelse, eller kan vilseledas om ordets användning är förändrad. Men det bör påpekas att lösningar till detta problem har presenterats (Pei, Sun and Xu 2019).

Varje ord som återstår formateras sedan om till sin grundform och blir på så sätt lättare för AI-programmet att behandla. Dessa ord kan sedan jämföras med de ord som programmet tar från dess databas, som består av bilder och text. Den slutliga bearbetningen av den angivna

texten är meningssegmentering. Detta innebär att kopplingen mellan orden betonas och deras innebörd klargörs och sker vanligen i tandem med textsegmentering.

När programmets språkteknik är uppbyggd kan programmet tekniskt generera en bild. Det artificiella neuronnätverket är korrekt uppbyggt och har använts för att träna AI-programmet. Som grund för träningen används den data som har skrapats från webben och språktekniken lär AI-programmet att tolka en textbaserad input.

3. TILLÄMPNING I PRAKTIKEN

År 2022 släpptes följande AI-program; Midjourney från Midjourney Inc, Dall-E 2 från OpenAI och Stable Diffusion från CompVis och Runway.

Det finns också en uppsjö av mer eller mindre effektiva program och kopior, men för denna avhandling har jag valt att begränsa mig till dessa tre på grund av deras popularitet och hur lättillgängliga de är. Alla tre program skapar relativt små bilder i låg resolution, som man mot betalning kan förbättra. Det är högst antagligen en kostnadsfråga som ligger bakom detta val då högre resolution skulle kräva mera av AI-programmen och det skulle innebära större kostnader för företagen.

3.1. MIDJOURNEY

Midjourney använder sig främst av en kanal på forumapplikationen Discord, med över 13 miljoner registrerade användare. För att använda Midjourney måste man gå med i Discord-kanalen och man kan där generera bilder via input, dessa bilder kan sedan manipuleras till en viss grad. Det är värt att notera att Discord främst fungerar som ett öppet forum, vilket innebär att alla medlemmar kan se alla bilder som genereras och texten som använts. Vidare kan också andra användare fortsätta arbeta med alla bilder som genereras. En klar fördel med detta är att alla användare kan lära sig hur verk som de uppskattar har genererats. Eftersom AI-programmen inte är deterministiska kan man inte återskapa någon annans verk, men ändå lära sig av dem. I bild 15 är en bild från Midjourney som en okänd användare har genererat med input: ”*cartoon animals rabbit, fox, squirrel, and bird in a forest cartoonish*”.



Bild 15 - Midjourny

3.2. DALL-E 2

Till skillnad från Midjourney använder OpenAi en webbsida för AI-programmet Dall-E 2, www.labs.openai.com, där man kan använda AI-programmet. Detta ger användaren en viss avskildhet och integritetsskydd, men det gör också att man inte kan se hur andra genererar bilder. Dall-E 2 tillåter att användaren laddar upp en bild och kan sedan fortsätta utveckla den, med varierande resultat. Bild 16 har jag genererat i Dall-E 2 med följande input: *”3D render of a cute tropical fish in an aquarium on a dark blue background, digital art”*.



Bild 16 - Dall-E 2

3.3. STABLE DIFFUSION / NIGHTCAFÈ

För Stable Diffusion har jag valt att använda webbsidan NightCafé som integrerar Stable Diffusion i sin applikation. NightCafé skiljer sig så till vida från de andra två att den tillåter användaren att välja vilken algoritm som skall användas i genereringen. Alternativen är VQGAN+CLIP, CLIP-Guided Diffusion, Dall-E 2 och Stable Diffusion. Man kan också välja att ladda upp en bild och utveckla den med hjälp av olika filter. Bild 17 visar att dessa filter klart är baserade på kända konstnärer som Hokusai, Pablo Picasso, Vincent van Gogh och Henri Matisse.

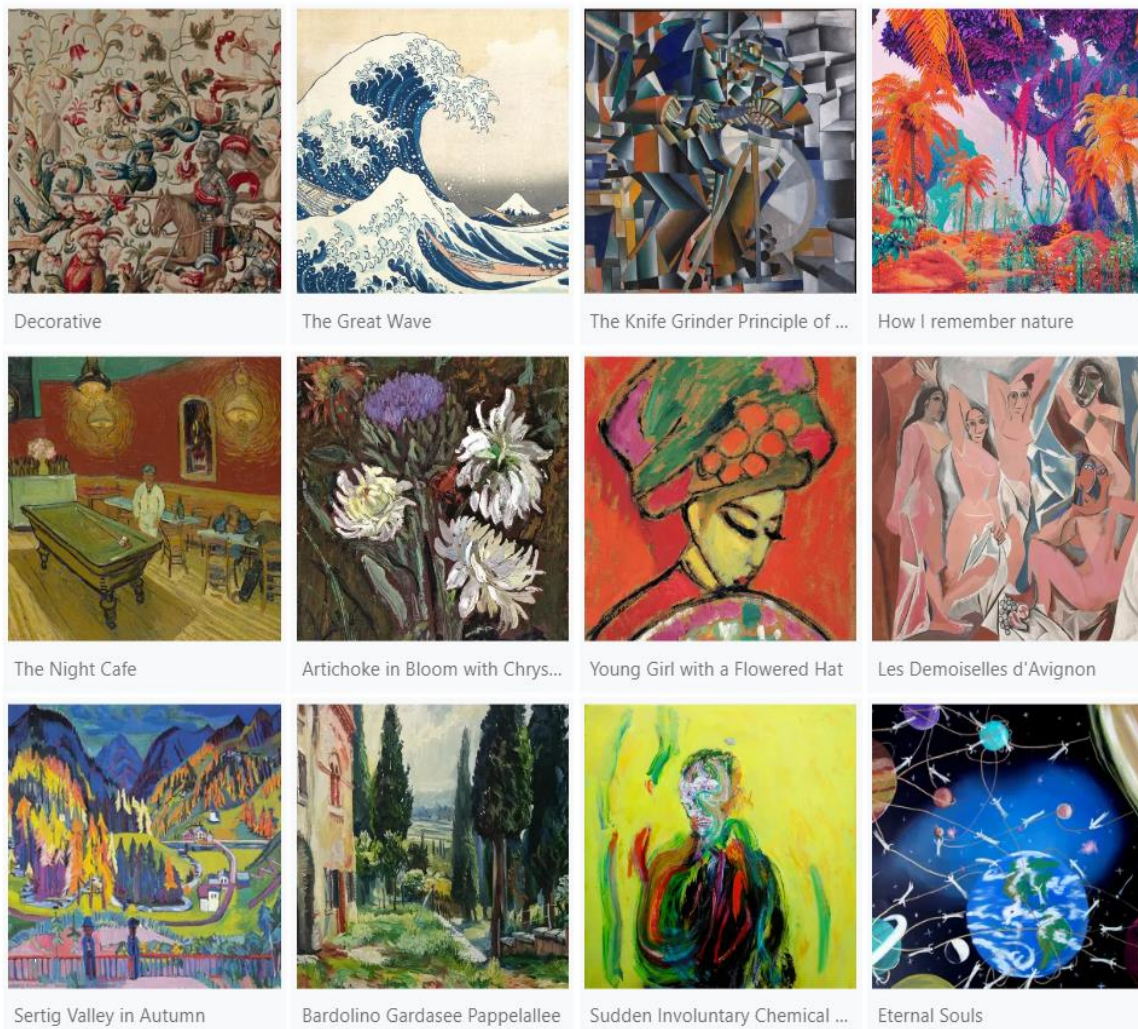


Bild 17 - NightCafé filter

Om man väljer att generera en bild från text kan användaren inte bara välja algoritm utan också stil, dessa filter varierar från *3D Game* och *Dark Fantasy* till *Cubist* och *Photo*. Dessa filter kan sedan användas om och om igen för att förändra bilden. Bild 18 genererade jag från input: *"a great wave crashing in on a beach"* med filtret NightCafe.



Bild 18 - NightCafé Stable Diffusion

3.4. BILDGENERERING

För att bättre förstå dessa tre AI-program och deras skillnader har jag valt att använda samma textinput i samtliga och jämföra resultaten.

Det första försöket hade följande input; ”*Kirby eating dinner in the Eiffel Tower with Sonic the Hedgehog*”. Kirby är en populär karaktär från Nintendo (Nintendo n.d.) och Sonic the Hedgehog var länge maskoten för Sega (Sega n.d.).



Bild 20 – Kirby (Nintendo n.d.)



Bild 19 – Sonic the Hedgehog (Sega n.d.)

Mitt syfte var att se om AI-programmen skulle kunna återskapa dessa karaktärer som båda är relativt populära, *Sonic the Hedgehog* har medverkat i 36 olika videospel (Sega n.d.) och Kirby i 38 (NintendoLife n.d.). Båda karaktärerna har tämligen mjuka drag och en begränsad färgpalett vilket jag hoppades skulle göra det lätt att utvärdera om programmet lyckades skapa dem. Eiffeltornet har en väldigt distinkt och känd siluett och torde därför vara tämligen lätt att återskapa.

3.4.1. KIRBY OCH SONIC I MIDJOURNEY

Det första programmet var Midjourney och som man kan se har Sonic och Kirby i alla fyra bilderna kombinerats till ett slags hybrid, vilket framkommer tydligt i övre bilden till höger där Sonic har fått Kirbys kännetecknande fläckar på kinderna. Det är bara i nedre bilden till vänster som separata karaktärer har skapats, men i totalt originella skepnader.

AI-programmet klarade dock av att tämligen väl generera Eiffeltornet, även om karaktärerna inte är inne i tornet. Detta kan mycket väl bero på hur texten har behandlats, om "in" var ett stopword har meningen förändrats. Vad exakt det är som är serverat är tämligen oklart, men trots det förstår man genast vad som händer i bilden.

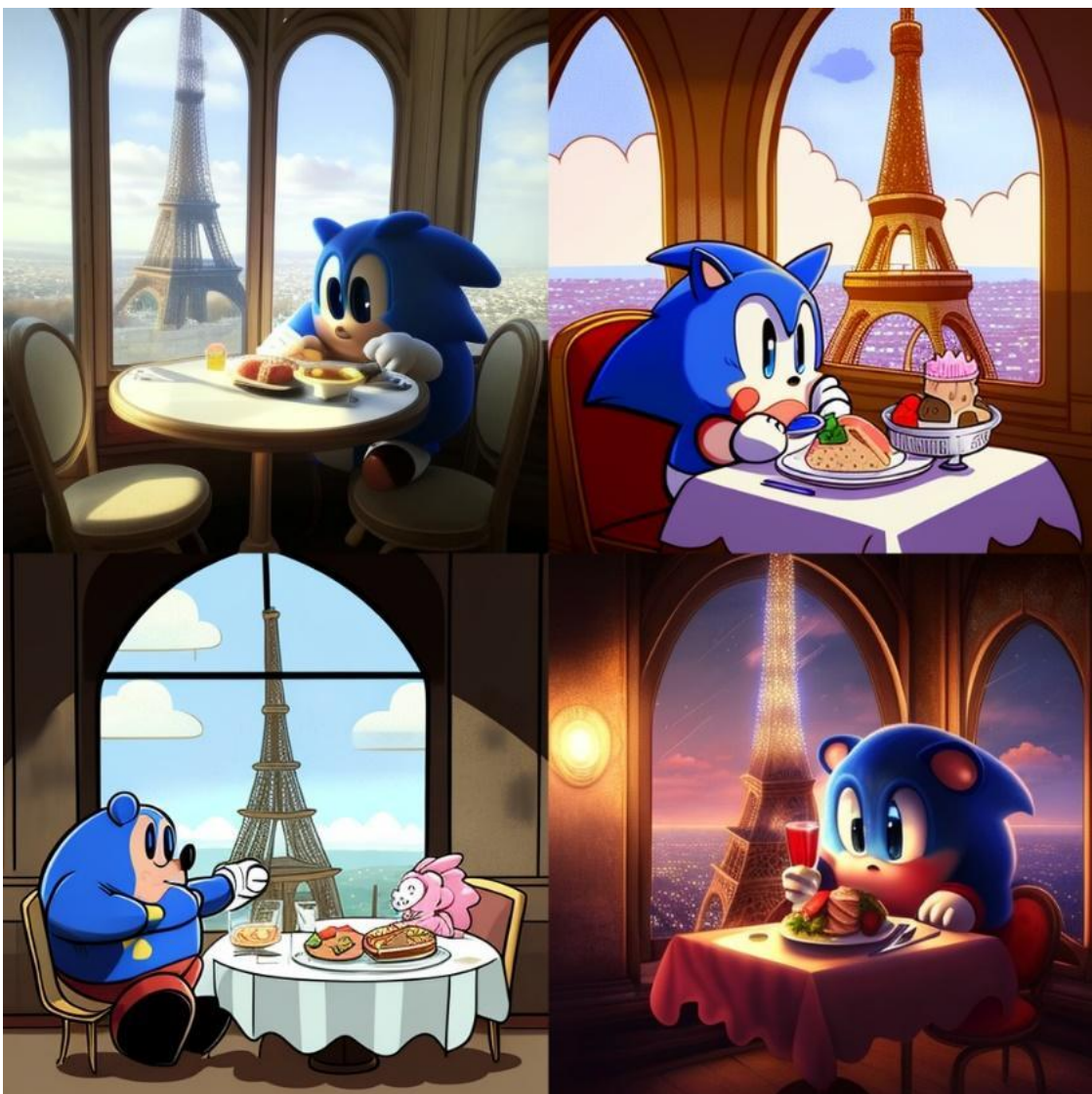


Bild 21 - Midjourney

3.4.2. KIRBY OCH SONIC I DALL-E 2

Om då Midjourney i en viss mån lyckades generera en bild med åtminstone drag av input önskemålet så var Dall-E 2:s resultat betydligt sämre. Här har programmet valt en extremt annorlunda still från vad Midjourneys bilder. Från resultatet kan man utläsa att programmet främst fäst sig vid tre saker, nämligen Eiffeltornet, äter och igelkott (eng. *hedgehog*). De två vänstra bilderna har lyckats göra igelkotten blå, och det är möjligt att den övre vänstra bildens rosa dräkt är en referens till Kirby. I övrigt har AI-programmet inte kunnat skapa det önskade resultatet.



Bild 22 - Dall-E 2

3.4.3. KIRBY OCH SONIC I NIGHTCAFÉ

Slutligen genererade NightCafé följande fyra bilder utgående från samma input med filtret NightCafé. Igen har programmet valt att fokusera på delar av den givna texten, främst Eiffeltornet och Sonic. I den övre vänstra bilden har Sonic och Kirby klart kombinerats till en karaktär, en originell karaktär har lagts till i bilden, men middagen finns med. I resten av bilderna saknas middagen och bara den nedre till höger innehåller två karaktärer. Även om NightCafés resultat kan anses innehålla mera detaljer än vad Dall-E 2 presterade har programmet inte lyckats ta med de nyckelord som angavs.



Bild 23 - NightCafé

Ett andra försök gjordes med följande input; ” *Medieval German castle, surrounded by mountains, high fantasy, epic, digital art*”. Eftersom texten fokuserar på ett landskap torde det vara lättare för samtliga AI-program att tolka den och återskapa de två viktigaste elementen, nämligen slottet och bergen. Stilarna högfantasi och episk är vanliga inom fantasigenren som bild 25 och 26 visar. Båda illustrationerna är av John Howe och båda föreställer staden Gondolin från J.R.R. Tolkiens bok *Silmarilion* (Howe n.d.). Bilderna valdes på grund av deras sammanställning, slottet är klart presenterat i båda bilderna, det omges av berg. Dessutom används ljus och mörker för att fokusera betraktarens blick och för att öka dramatiken.



Bild 24 - *Morgoth's Forces before Gondolin* (Howe n.d.)



Bild 25 - *The Fall of Gondolin* (Howe n.d.)

3.4.4. SLOTT I MIDJOURNEY

Midjourney lyckades få med både slottet och bergen, stilen stämmer överens med vad som kan förväntas av högfantasi och epik. Arkitekturen på slottet är orealistisk med väggar som inte tycks sitta ihop och tornen dyker delvis upp från ingenstans. Men som snabbt producerad konceptkonst fungerar bilderna och man förstår vad de föreställer. En detalj Midjourney konstant höll fast vid var att placera slottet på en kulle eller ett mindre berg, en sjö var också ett återkommande element. Dessa element kvarstod i bilderna även när jag tog bort ”surrounded by mountains” från input texten. Det är mycket möjligt att en stor del av den data AI-programmet tränats på har förstärkt sambandet mellan slott och kullar och detta leder till det resultatet. I bild 27 är det ofta ljusare element i himlen och mörkare element närmare de nedre kanterna på bilderna, något som också tydligt figurerar i verken av John Howe.



Bild 26 - Midjourney

3.4.5. SLOTT I DALL-E 2

Dall-E 2 lyckades betydligt bättre med att skapa slott än spelkaraktärer. Samma kopplingar mellan slott och höjder finns här, precis som i Midjourneys verk och orsakerna är högst antagligen de samma. Bilderna saknar däremot samma känsla av digitalkonst som Midjourney lyckades återskapa, faktum är att bilderna är betydligt grövre och klumpigare med en högre pixelering. Bilderna är ändå klara i vad de föreställer och elementen är på plats.

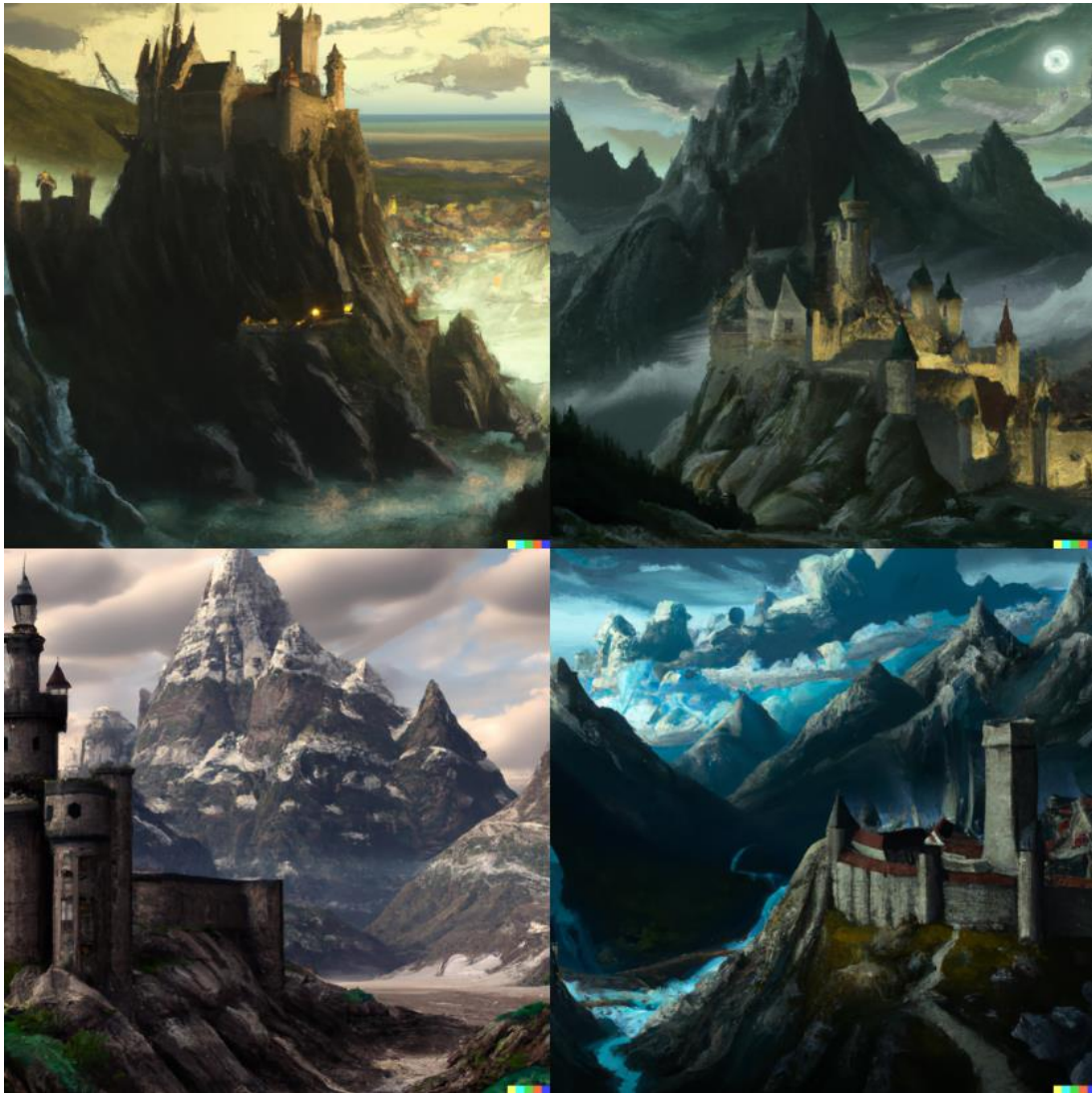


Bild 27 - Dall-E 2

3.4.6. SLOTT I NIGHTCAFÉ

I NightCafé använde jag igen samma inställningar som tidigare, algoritmen var Stable och stilen NightCafé. Bekanta element som sjöar och kullar är igen med, men i NightCafé återfinns större skillnader i arkitekturen. Slotten är inte lika klart ett befäst centrum i bilderna utan är istället utspritt över en större yta och i bilden nere till höger är det närmast frågan om ett övervuxet hus med torn. Om man jämför bilderna med dem från Midjourney ser man också att den klara användningen av ljus och mörker som ett sätt att skapa fokus saknas.



Bild 28 - NightCafé

Från dessa två jämförelser av AI-programmen kan det konstateras att det är lättare för samtliga program att generera bilder som föreställer landskap än det är att generera populärkulturella karaktärer. Detta kan möjligtvis bero på att det ursprungliga träningsmaterialet innehåller mera bilder på slott och att dessa bilder är av en typ som ges större tyngd i träningen.

För att kontrollera detta jämförde jag i Google Trends söktermerna ”*castle*”, ”*Kirby*” och ”*Sonic*”, med följande specifikationer; Över hela världen, 2004 till i dag, Alla kategorier och Webbsökning. Undersökningen berättar också hur många träffar varje term fick. Syftet är att se om det finns ett samband mellan termernas popularitet och hurdana resultat som generades av AI-programmen.

”*Castle*” ligger stadigt mellan 75 och 100, med en viss sänkning från 2015 och 2021 då Google förbättrade sin insamling av data. Sökningen gav ungefär 1,77 miljarder resultat.

”*Sonic*” kretsar för det mesta kring 50, med motsvarande sänkning 2015, men med en skarp, men kort, höjning 2022 då den andra *Sonic the Hedgehog* filmen släpptes. Sökningen gav ungefär 758 miljoner resultat.

”*Kirby*” ligger också för det mesta kring 50, men med små ökningar då nya spel har släppts. Sökningen gav ungefär 204 miljoner resultat.

Utgående från detta är det möjligt att se att slott är vanligare söktermer och ger mera resultat, då är det också sannolikt att det finns mera material relaterat till slott i de databaser som har använts för träning. Detta skulle i så fall leda till att AI-programmen kunde vara bättre på att generera slott, något som också de bilder som genererades tyder på.

4. AI GENERERADE BILDER OCH SAMHÄLLET

När Dall-E 2, Midjourney och Stable Diffusion släpptes kom mycket av diskussionen kring dessa AI-program att fokusera på deras implikationer för konstnärer (Shaffi 2023). Frågan var om AI genererade bilder skulle kunna ersätta konstnärer och också om hur den träningsdata som använts samlats in. Problemet för många var att AI-programmen använde sig av träningsdata som samlats in utan tillstånd av de som innehar upphovsrätten till verken, något som Midjourneys vd, David Holz, medgav i en intervju med Forbes (Salkowitz 2022). Enligt Holz finns det inte klara regler för vad som gäller när det kommer till dataskrapning och konstnärers rätt att vägra tillåta att deras verk används i dataträning. Det dröjde dock inte länge innan Stability AI, företaget bakom LAION och Midjourney alla stämades för upphovsrättsbrott (Butterick 2023).

Vad gäller upphovsrätten för bilder genererade av AI-program är det enligt amerikansk och EU-lagstiftning inte möjligt att få upphovsrätt för dylika verk. Det beror främst på att ett verk måste vara skapat av en människa för att kunna skyddas, och ett verk genererat av ett AI-program inte har tillräckligt med mänskligt deltagande i skapandet (De Grauwe and Gryspeerdt 2022). Dock har de amerikanska myndigheterna, *U.S. Copyright Office*, den 16 mars 2023 meddelat att ett verk kan anses kvalificera för upphovsrättsskydd trots att verket är AI genererat om en människa sedan har förändrat verket i en tillräckligt stor utsträckning (U.S. Copyright Office 2023)

Att AI genererade bilder kan komma att ersätta till exempel konceptkonstnärer är mycket sannolikt, även om det också är möjligt att konstnärer kan komma att utnyttja AI-program som verktyg i visuellt skapande (Salkowitz 2022). Effektiv konceptkonst tar kring 20 minuter att skapa av en tränad artist, ett AI-program behöver bara ett par sekunder. Kostnadsfrågan kan då tyvärr komma att avgöra ifall en eller flera artister behövs, eller om ett program kan komma att helt eller delvis ersätta dem. Som exempel kan visas illustrationen *Fountain Plaza* från spelet *Dark Souls 3* (bild 30) och en bild genererad i Midjourney från den enkla inputen ”*concept art for dark souls video game*” (bild 31). *Fountain Plaza* är rent tekniskt betydligt bättre gjord och saknar de problem som vanligen associeras med AI genererade bilder, som till exempel groteska händer, linjer som sticker ut och förvridna proportioner. Dock är bilden från Midjourney inte oäven och den innehåller samma känsla som *Fountain Plaza*, vilket kan anses vara det främsta syftet för konceptkonst.



Bild 29 - Fountain Plaza (Fenlon 2015)



Bild 30 - Midjourney

Trots dessa utmaningar för konstnärer och det potentiellt existentiella hotet mot deras yrke hävdar jag att de största utmaningar som generativa AI-program för med sig är på ett bredare samhällsligt perspektiv.

4.1. AI GENERERAD PORNOGRAFI

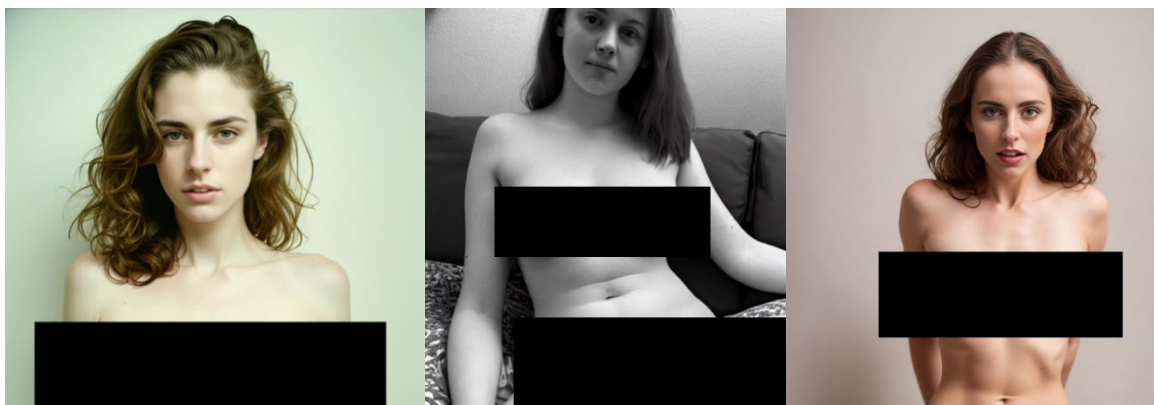


Bild 31 - Unstable Diffusion

Som tidigare nämnts har de kommersiella versionerna av AI-programmen inbyggda begränsningar som, i teorin, hindrar dem från att bland annat skapa pornografiska bilder (Midjourney n.d.). Men detta gäller endast kommersiella produkter och eftersom Stable Diffusion finns tillgänglig som öppen källkod tog det inte länge förrän AI-program designade för pornografi blev tillgängliga (Cole 2022). Dessa program gör det möjligt att skapa oändliga mängder hentai, anime inspirerad tecknad pornografi, och än mera oroande, djupfejk pornografi av människor.

Djupfejk är en form av bildmanipulation där man med hjälp av djupinlärning tränar ett neuronätverk att producera bilder eller videon av en person (Sample 2020). I rapporten *The State of Deepfakes* från 2019 av DeepTrace är upp till 96% av all djupfejk pornografisk, och de flesta föreställer kända skådespelare (Ajder, et al. 2019). Denna rapport är sammanställd innan bild genererande AI-program slog igenom som en allmänt tillgänglig produkt, men det finns dock ingen orsak att tro att antalet skulle ha förändrats till det bättre sedan dess.

AI-programmet Unstable Diffusion från Unstability AI, en kopia av Stable Diffusion utan samma spärrar, är ett av många program som tillåter användare att generera pornografiskt material, men inte av riktiga personer. Programmet bygger på öppen källkod, vilket innebär

att det inte är svårt för en kunnig användare att ändra på programmets spärrar för att generera djupfejk bilder av människor.

Detta kommer oundvikligen att användas för att skapa ofrivillig pornografi, det är ofrånkomligt eftersom det redan skedde innan AI genererade bilder gjorde det möjligt att generera dylikt material på några sekunder. Förutom att detta rent juridiskt är olagligt inom ett flertal länder är detta fenomen ett hot som kan komma att ha oförutsedda konsekvenser för kvinnor, eftersom i de allra flesta fall dessa bilder är av kvinnor (Ajder, et al. 2019).

Enligt en undersökning utförd av säkerhetsföretaget Sensity (Vincent 2020) var de flesta som svarade främst intresserade av att se sexuellt material av personer de kände, bild 29. Om det sedan rör sig om en vilja att skapa hämndporr eller inte är irrelevant då den slutgiltiga effekten är den samma.

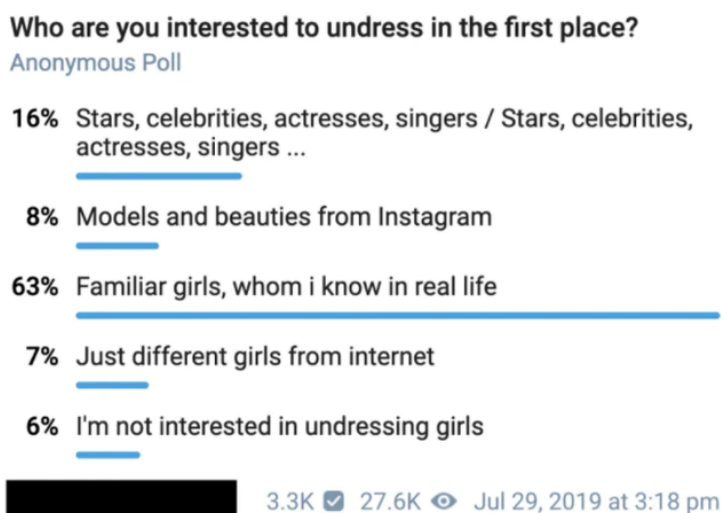


Bild 32 – Sensity (Vincent 2020)

Som redan konstaterats kan AI-program använda en bild som input, bilden kan föreställa vad som helst, till exempel kan det vara en profilbild från Facebook, LinkedIn eller någon annan plattform. Om användaren kan utnyttja en profilbild som input i ett AI-program finns det i princip ingenting offret kan göra för att stoppa detta (O'Halloran 2021). Obegränsade mängder av genererade bilder kan sedan skapas och spridas, till oförutsebar skada för offret då det i princip är omöjligt att rensa bort dylikt material. I värsta fall kan detta komma att plåga offret för resten av hennes liv, speciellt då ansiktsgenkänning gör det möjligt för dylika bilder att ständigt dyka upp i sökmotorer (Hill 2022). Kvinnorna i bild 31 är alla genererade i Unstable Diffusion av en okänd person och även om de inte ännu är fotorealistiska kan det

ändå konstateras att det redan är möjligt att skapa mycket realistiska bilder, också i program som inte ägs av företag.

4.2. AI GENERERAD PROPAGANDA

År 2018 beskrev den politiske rådgivaren Steve Bannon sin strategi på följande sätt: ”*flood the zone with shit*”, fritt översatt som ”översvämma zonen med skit” (Lewis 2018). Meningen är att helt enkelt dränka alla andra åsikter i så mycket innehåll att de segment av populationen man siktar på antingen inte blir hörda, tappar intresse på grund av oljudet eller accepterar det meddelande man vill sprida. Enligt Amnesty Internationals rapport ”*The Social Atrocity*” användes en motsvarande strategi i Myanmar för att piska upp hat mot den rohingy minoriteten i landet, något som ledde till en förödande etnisk resning (Amnety International 2022).

I och med att AI-program gör det lättare att skapa material är det oundvikligt att en del av detta material kommer att användas för att sprida desinformation och i propagandasyfte. Detta sker redan med hjälp av företag som Synthesia som använder AI-program för att skapa digitala avatarer, som i bild 31 (Satariano and Mozur 2023).



Bild 33 - digital avatar (Satariano and Mozur 2023)

5. SLUTSATSER

I denna avhandling har bild genererande AI-program presenterats, både den bakomliggande tekniken som utgörs av avancerad maskininlärning och de för tillfället populäraste programmen. Jämförelserna mellan Midjourney, Dall-E 2 och NighCafé visar att Midjourney i skrivande stund är det program som bäst klarar av att tolka den input som ges i textform. Problemen med AI genererade bilder diskuteras också och de implikationer de för med sig för både individen och samhället.

Vi måste acceptera att bilder genererad av AI-program inte kommer att försvinna, det som en gång har skapats kan inte göras ogjort. Däremot behöver vi inte acceptera att programmen används i vilka syften som helst, vilket innebär att vi som samhälle står inför stora utmaningar då det kommer till genererade bilder. Delvis måste vi finna en lösning på hur de skall kunna användas lagligt och så att de kompenserar dem vars material används i träningssyfte, och delvis måste vi lösa problemet med genererade djupfejk bilder. Lösningen måste göra det möjligt för konstnärer att utöva sitt yrke, och för alla de som vill skapa konst för nöjes skull, och skydda individens integritet och självbestämmanderätt över sin liknelse. Det är i princip lönlöst att förvänta sig att en lösning kan komma från teknologisektorn, detta är ett samhällligt problem och kräver därför en samhälllig lösning som inkluderar lagstiftning, monetärrättvisa och jämställdhet. Ett första steg är att se över lagstiftningen kring hämndporr, ovillig porr och klart förbjudna dess skapande och spridning. Vidare bör också upphovsrättsinnehavarnas ställning snarast möjligt blir klar så att samtliga berörda parter vet vilka regler som gäller kring insamling av material och dess användning.

Då användare kan generera sina egna bilder och snabbt ge dem spridning genom till exempel sociala medier är det fullt möjligt att mer eller minder avsiktligt manipulera åsikter med hjälp av dessa bilder. I takt med att AI genererade bilder blir allt mera fotorealistiska kommer detta problem att växa exponentiellt och i värsta fall leda till att det i många fall blir omöjligt att skilja på verkliga fotografier och förfalskningar. Detta är något som också måste begrundas och tyvärr finns det inga enkla lösningar.

KÄLLFÖRTECKNING

- Ajder, Henry, Giorgio Patrini, Francesco Cavalli, och Laurence Cullen. 2019. *The State of Deepfakes: Landscape, Threats, and Impact*. Deeptrace.
- Amnesty International. 2022. *THE SOCIAL ATROCITY - META AND THE RIGHT TO REMEDY FOR THE ROHINGYA*. Amnesty International .
- AWS. u.d. *What Is A Neural Network?* Använd den 27 mars 2023.
<https://aws.amazon.com/what-is/neural-network/>.
- Baio, Andy. 2022. *Exploring 12 Million of the 2.3 Billion Images Used to Train Stable Diffusion's Image Generator*. Använd den 27 mars 2023.
<https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/>.
- Butterick, Matthew. 2023. "Because AI needs to be fair & ethical for everyone." Stable Diffusion litigation. Använd den 27 mars 2023. <https://stablediffusionlitigation.com/>.
- Chen, Qing-Qiang, Mao-Hsiung Hung, och Fumin Zou. 2017. "Effective and adaptive algorithm for pepper-and-salt noise removal." IET Image Processing 11 (9).
doi:<https://doi.org/10.1049/iet-ipr.2016.0692>.
- Cole, Samantha. 2022. "This AI Tool Is Being Used to Make Freaky, Machine-Generated Porn." Vice. Använd den 27 mars 2023.
<https://www.vice.com/en/article/xgygy4/stable-diffusion-stability-ai-nsfw-ai-generated-porn>.
- Common Crawl. u.d. *Frequently Asked Questions*. Använd den 27 mars 2023.
<https://commoncrawl.org/big-picture/frequently-asked-questions/>.
- Croitoru, Florinel-Alin, Vlad Hondru, Radu Tudor Ionescu, och Mubarak Shah. 2022. "Diffusion Models in Vision: A Survey." doi:10.48550/arXiv.2209.04747.
- De Grauwe, Pieter och Gryspeerdt, Sacha. 2022. "Artificial intelligence (AI): The qualification of AI creations as "works" under EU copyright law". Gevers.
- EUR-Lex. 2020. *Dataskydd inom sektorn för elektronisk kommunikation*. Använd den 27 mars 2023. <https://eur-lex.europa.eu/legal-content/SV/TXT/?uri=LEGISSUM:l24120>.

- Fenlon, Wes. 2015. "Dark Souls 3 preview: a grander sense of scale, but still familiar". PCGamer. Använd den 30 mars 2023.
- Hardesty, Larry. 2017. *Explained: Neural networks*. Använd den 27 mars 2023.
<https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>.
- Hilary. 2022. "How to remove noise from a photo." inPixio. Använd den 27 mars 2023.
<https://www.inpixio.com/blog/tutorials/photo-noise-reduction/>.
- Hill, Kashmir. 2022. "A Face Search Engine Anyone Can Use Is Alarmingly Accurate." New York Times. Använd den 27 mars 2023.
<https://www.nytimes.com/2022/05/26/technology/pimeyes-facial-recognition-search.html>.
- Howe, John. u.d. John-Howe.com. Använd den 27 mars 2023. <https://www.john-howe.com/blog/>.
- Hypotenuse AI. u.d. "AI Image Generators: How They Work And Why They Are Important." Hypotenuse AI. Använd den 27 mars 2023. <https://www.hypotenuse.ai/blog/ai-image-generator>.
- IBM. u.d. *What is Computer Vision?* Använd den 27 mars 2023.
<https://www.ibm.com/topics/computer-vision>.
- Indurkha, Nitin, och Fred J. Damerau. 2010. *Handbook of Natural Language Processing*. New York: Taylor & Francis Group.
- Kaelbling, Leslie Pack, Michael L. Littman, och Andrew W. Moore. 1996. "Reinforcement Learning: A Survey." *Journal of Artificial Intelligence Research* 4.
- Karagiannakos, Sergios, och Nikolas Adaloglou. 2022. "How diffusion models work: the math from scratch." AI SUMMER. Använd den 27 mars 2023.
<https://theaisummer.com/diffusion-models/>.
- Kavlakoglu, Eda. 2020. *AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?* Använd den 2023 mars 27.
<https://www.ibm.com/cloud/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks>.

- Klingemann, Mario. 2018. *Memories of Passersby I*. Använd den 27 3 2023.
<https://www.theverge.com/2019/3/5/18251267/ai-art-gans-mario-klingemann-auction-sothebys-technology>.
- LAION. u.d. *About*. Använd den 27 mars 2023. <https://laion.ai/about/>.
- LAION-Aesthetic. u.d. *images*. Använd den 27 mars 2023. <https://laion-aesthetic.datasette.io/laion-aesthetic-6pls/images>.
- . u.d. *images*. Använd den 27 mars 2023. https://laion-aesthetic.datasette.io/laion-aesthetic-6pls/images?_search=woman&_sort_desc=aesthetic.
- . u.d. *images*. Använd den 26 mars 2023. https://laion-aesthetic.datasette.io/laion-aesthetic-6pls/images?_sort_desc=aesthetic.
- LeCunn, Yann, Yoshua Bengio, och Geoffrey Hinton. 2015. ”*Deep learning*.” Nature (Macmillan Publishers Limited.) 521.
- Lewis, Michael. 2018. ”*Has Anyone Seen the President?*” Bloomberg. Använd den 27 mars 2023. <https://www.bloomberg.com/opinion/articles/2018-02-09/has-anyone-seen-the-president>.
- McManu, John. 2015. ”*Have pictures of Muhammad always been forbidden?*” BBC News. Använd den 27 mars 2023. <https://www.bbc.com/news/magazine-30814555>.
- Midjourney. u.d. ”*Community Guidelines*.” midjourney.com. Använd den 27 mars 2023. <https://docs.midjourney.com/docs/community-guidelines>.
- MIT News. 2020. *Artificial Intelligence Yields New Antibiotic*. Använd den 2023 mars 27. <https://betterworld.mit.edu/artificial-intelligence-yields-new-antibiotic/>.
- Mitchell, Tom M. 1997. *Machine Learning*. McGraw-Hill Science/Engineering/Math.
- Nake, Frieder. 1967. *no title*. Tate. Använd den 27 3 2023.
<https://www.tate.org.uk/art/artworks/nake-no-title-p80806>.
- Nintendo. u.d. *Kirby*. Använd den 27 3 2023. <https://kirby.nintendo.com/>.
- Nintendo Switch. 2022. ”*It’s a smorgasbord of Kirby with games, videos, and art of the brave pink puff!*” nintendo.com. Använd den 27 mars 2023.
<https://www.nintendo.com/whatsnew/its-a-smorgasbord-of-kirby-with-games-videos-and-art-of-the-brave-pink-puff/>.

- NintendoLife. u.d. ”*All Kirby Games.*” nintendolife.com. Använd den 27 mars 2023.
<https://www.nintendolife.com/games/browse?title=series%3Akirby>.
- O’Halloran, Amelia. 2021. ”*The Technical, Legal, and Ethical Landscape of Deepfake Pornography.*” Brown University.
- Pei, Zhengqi, Zhewei Sun, och Yang Xu. 2019. ”*Slang detection and identification.*” Proceedings of the 23rd Conference on Computational Natural Language Learning. Hong Kong: Association for Computational Linguistics.
- Salkowitz, Rob. 2022. ”*AI Is Coming For Commercial Art Jobs. Can It Be Stopped?*” Forbes. Använd den 27 mars 2023. <https://www.forbes.com/sites/robsalkowitz/2022/09/16/ai-is-coming-for-commercial-art-jobs-can-it-be-stopped/>.
- . 2022. ”*Midjourney Founder David Holz On The Impact Of AI On Art, Imagination And The Creative Economy.*” Forbes. Använd den 27 mars 2023. <https://www.forbes.com/sites/robsalkowitz/2022/09/16/midjourney-founder-david-holz-on-the-impact-of-ai-on-art-imagination-and-the-creative-economy/>.
- Sample, Ian. 2020. ”*What are deepfakes – and how can you spot them?*” The Guardian. Använd den 27 mars 2023. <https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them>.
- Samuel, Arthur. 1959. ”*Some Studies in Machine Learning Using the Game of Checkers.*” IBM Journal of Research and Development 3 (3).
- Satariano, Adam, och Paul Mozur. 2023. ”*The People Onscreen Are Fake. The Disinformation Is Real.*” New York Times. Använd den 27 mars 2023. <https://www.nytimes.com/2023/02/07/technology/artificial-intelligence-training-deepfake.html>.
- Schrittwieser, Julian. 2020. *MuZero: Mastering Go, chess, shogi and Atari without rules.* Använd den 27 mars 2023. <https://www.deepmind.com/blog/muzero-mastering-go-chess-shogi-and-atari-without-rules>.
- Sega. u.d. ”*Search Results for Sonic.*” sega.com. Använd den 27 mars 2023. <https://www.sega.com/search-games?title=sonic>.

- Sega. u.d. *Sonic the Hedgehog*. Använd den 27 3 2023.
https://play.google.com/store/apps/details?id=com.sega.sonic1px&hl=en_US&pli=1.
- Sega. u.d. *Sonic the Hedgehog*. Använd den 27 3 2023. <https://www.sega.com/games/sonic-hedgehog>.
- Shaffi, Sarah. 2023. ”*It’s the opposite of art’: why illustrators are furious about AI.*” the Guardian. Använd den 27 mars 2023.
<https://www.theguardian.com/artanddesign/2023/jan/23/its-the-opposite-of-art-why-illustrators-are-furious-about-ai>.
- Song, Yang, och Stefano Ermon. 2019. ”*Generative Modeling by Estimating Gradients of the Data Distribution.*” 33rd Conference on Neural Information Processing Systems.
- Srivastava, Akash. 2017. ”*VEEGAN: Reducing Mode Collapse in GANs using Implicit Variational Learning.*” Published as a conference paper at NIPS.
- Stability.ai. u.d. *Stable Diffusion Launch Announcement*. Använd den 27 mars 2023.
<https://stability.ai/blog/stable-diffusion-announcement>.
- Theobald, Oliver. 2018. *Machine Learning for Absolute Beginners*.
- U.S. Copyright Office. 2023. *Rules and Regulations Vol. 88 No. 51*. Federal Register.
- Vincent, James. 2019. ”*A never-ending stream of AI art goes up for auction.*” The Verge. Använd den 27 mars 2023. <https://www.theverge.com/2019/3/5/18251267/ai-art-gans-mario-klingemann-auction-sothebys-technology>.
- . 2022. ”*Anyone can use this AI art generator — that’s the risk.*” The Verge. Använd den 27 mars 2023. <https://www.theverge.com/2022/9/15/23340673/ai-image-generation-stable-diffusion-explained-ethics-copyright-data>.
- . 2020. ”*Deepfake bots on Telegram make the work of creating fake nudes dangerously easy.*” The Verge. Använd den 27 mars 2023.
<https://www.theverge.com/2020/10/20/21519322/deepfake-fake-nudes-telegram-bot-deepnude-sensity-report>.
- . 2022. ”*The scary truth about AI copyright is nobody knows what will happen next.*” The Verge. Använd den 27 mars 2023. <https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data>.

Weng, Lilian. 2021. "What are Diffusion Models?" Lil'Log. Använd den 27 mars 2023.
<https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>.

Zylinska, Joanna. 2020. *AI Art: Machine Visions and Warped Dreams*. Open Humanities Press CIC.