

Generatiiviset mallit: GAN ja VAE – sovellukset yhdessä ja erikseen

TURUN YLIOPISTO
Tietotekniikan laitos
LuK
Tietojenkäsittelytiede
Lokakuu 2025
Markus Koponen
Ohjaaja: Jari Björne

TURUN YLIOPISTO
Tietotekniikan laitos

MARKUS KOPONEN Ohjaaja: Jari Björne : Generatiiviset mallit: GAN ja VAE –
sovellukset yhdessä ja erikseen

LuK, 32 s.
Tietojenkäsittelytiede
Lokakuu 2025

Generatiiviset mallit ovat keskeinen tekoälyn osa-alue, ja niitä on monenlaisia. Tässä tutkielmassa keskitytään GAN- ja VAE-malleihin sekä niitä hyödyntäviin hybridi-malleihin. Generatiiviset vastakkaisverkot (GAN) ovat laajasti käytettyjä visuaalisen datan generoinnissa. Näiden avulla voidaan tuottaa realistisia ja yksityiskohtaisia tuloksia. Variaatiollista automaattikoodainta (VAE) käyttävät mallit puolestaan soveltuvat hyvin datan pakkaamiseen ja poikkeamien tunnistukseen, koska niiden avulla voidaan luoda rakenteellisesti johdonmukaisia malleja. Hybridimallit yhdistävät VAE:n ja GAN:n lähestymistavat.

Variaatiollinen automaattikoodain koostuu kahdesta neuroverkosta: dekooderista ja enkooderista. Malli hyödyntää todennäköisyyspohjaista lähestymistapaa. Generatiiviset vastakkaisverkot puolestaan rakentuvat kahdesta neuroverkosta. Nämä kaksi neuroverkkoa ovat generaattori ja diskriminaattori, näiden neuroverkkojen välinen kilpailu johtaa tasapainotilaan — tilanteeseen, jossa on saavutettu haluttu tulos. Hybridimallit sisältävät kolme pääosaa: enkooderin, generaattorin ja diskriminaattorin. Niiden avulla voidaan kehittää sovelluksia, joissa vältetään yksittäisten mallien heikkoudet. Tilanteisiin sopivien mallien valinta on ratkaisevaa näiden käytössä.

Asiasanat: GAN, VAE, Generatiiviset mallit, hybridimallit, GAI

Sisällys

1	Johdanto	1
2	Generatiiviset mallit	7
2.1	Yleiskatsaus	7
2.2	Rajaus GAN ja VAE	9
3	Generatiiviset vastakkaisverkot (GAN)	11
3.1	Toiminta	11
3.2	Sovellukset	14
4	Variaatiollinen automaattikoodain (VAE)	17
4.1	Toiminta	17
4.2	Sovellukset	19
5	GAN ja VAE - hybridimallit	23
6	Analyysi	28
7	Yhteenveto	30
7.1	Pohdinta ja havainnot	30
7.2	Tutkimuskysymyksiä vastaukset	31
	Lähdeluettelo	33

Kuvat

1.1	Hakumenetelmän kuvaus	2
3.1	GAN toimintaperiaate	12
3.2	GAN toiminta	13
4.1	VAE toimintaperiaate	18
5.1	Hybridi-mallin toimintaperiaate	24
6.1	Mallien vertailutaulukko	28

Taulukot

1.1	Taulukko valituista generatiivisia malleja käsittelevistä artikkeleista. .	5
-----	--	---

1 Johdanto

Generatiiviset mallit ovat olleet keskeisessä asemassa viime vuosien tekoälykehityksessä. Tämän hetken trendinä on yhä parempien generatiivisten mallien tuottaminen, jota tarkastellaan tässä (8/2025 asti ulottuvassa) tutkielmassa. Generatiiviset mallit eroavat merkittävästi perinteisistä ennustavista malleista, sillä pelkän ennustamisen sijaan ne tuottavat myös uutta, koulutusaineistoa muistuttavaa dataa [1]. Tässä tutkielmassa keskitytään generatiivisiin vastakkaisverkkoihin (GAN) sekä variaatiollisiin automaattikoodereihin (VAE) ja tarkastellaan niiden hyödyntämistä yhdessä ja erikseen erilaisissa generointitehtävissä.

GAN-malli esiteltiin vuonna 2014 ja VAE-malli vuonna 2013. GAN ja VAE ovat edelleen keskeisiä lähestymistapoja generatiivisissa malleissa ja edustavat generatiivisen mallinnuksen eri suuntauksia: kilpailullista ja probabilistista lähestymistapaa.[2]

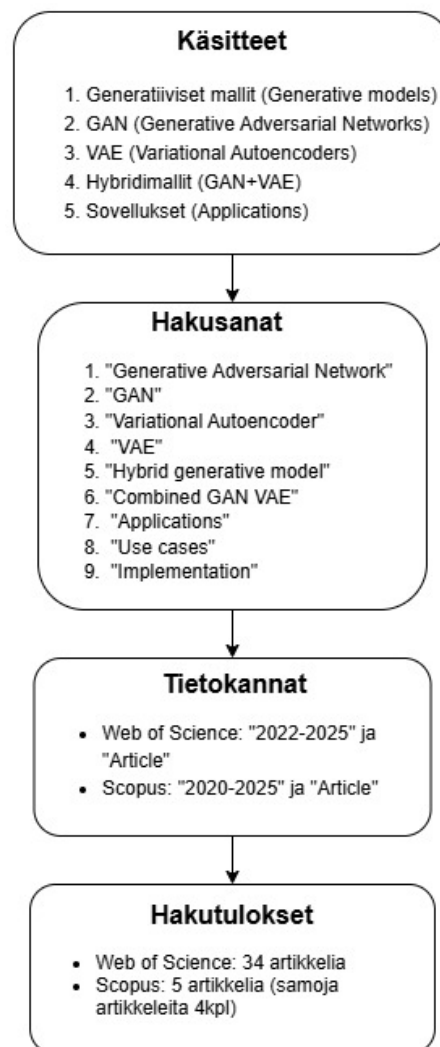
[1] Työn tavoitteena on koota kirjallisuuskatsaukseen generatiivisten vastakkaisverkkojen (GAN) ja variaatiollisten automaattikoodereiden (VAE) toiminta paitsi erillisissä sovelluksissa myös sellaisissa sovelluksissa, joissa nämä kaksi mallia on yhdistetty. Tutkielmassa selvitetään näiden mallien toimintaperiaatteet, kartoitetaan sovelluskohteita ja tarkastellaan, mitä hyötyä hybridimalleista saadaan. Näitä teemoja käsitellään seuraavien tutkimuskysymysten kautta:

TK1: Miten GAN- ja VAE-mallit eroavat toimintaperiaatteeltaan ja rakenteeltaan?

TK2: Mihin sovelluksiin hyödynnetään GAN- ja VAE-malleja?

TK3: Mitä hyötyjä voidaan saavuttaa yhdistämällä GAN- ja VAE-mallit?

Seuraavaksi tarkastelellaan tutkielman tiedonhakuprosessia. Alla olevassa kuvassa 1.1 on kuvattu tiedonhakuprosessi.



Kuva 1.1: Hakumenetelmän kuvaus

Tutkielman tiedonhaussa on käytetty kahta eri tietokantaa, Web of Sciencea ja Scopusta. Hakulause oli: ("Generative Adversarial Network" OR GAN) AND

("Variational Autoencoder"OR VAE) AND ("hybrid model"OR "combined model") AND (application OR "use case)". Web of Sciencesta löytyi hakuehdoillani 34 artikkelia ja Scopusesta 5 artikkelia. Viidestä artikkelista, jotka löytyivät Scopusen kautta, neljä oli samoja kuin Web of Sciencesta löydetty. Suurin osa artikkeleista käsitteli kaikkia haluttuja teemoja, mutta selkeyttääkseen tiedonhaun onnistumista artikkelit on lajiteltu. Tutkielmaa tehdessä haettiin myös tietoa yksittäisistä GAN-tai VAE-malleista hakulauseilla: ("variational autoencoder") AND ("representation learning"OR "latent space"OR "unsupervised learning") NOT ("generative adversarial network"OR "GAN") ja ("generative adversarial network") AND ("image synthesis"OR "data generation"OR "adversarial training") NOT ("variational autoencoder"OR "VAE"). VAE:hen liittyviä artikkeleita löytyi rajauksella vuosi 2025, 101 tulosta. GAN:iin liittyviä artikkeleita löytyi 96 tulosta. Näistä hauista saatuja artikkeleita käytettiin työssä tukemaan alkuperäisen haun tuloksia. Alla olevassa taulukossa 1.1 on tutkielman tiedonhakemisen tulokset. Tässä taulukossa artikkelit on lajiteltu teemoittain.

#	Artikkelin nimi	GAN	VAE	Hybridi (GAN+VAE)	Sovellukset
0	Large-Scale Completion of Ionospheric TEC Maps Using Machine Learning Models With Constraints Conditions (Li, QF; Fang, HX; (...); Ren, GM, 2025)				X
1	EvaGoNet: An integrated network of variational autoencoder and Wasserstein generative adversarial network with gradient penalty for binary classification tasks (Luo, C; Xu, Y; Shao, Y; Wang, Z; Hu, J; Yuan, J; Liu, Y; Duan, M; Huang, L; Zhou, F, 2023)	X	X	X	X
2	Collaborative-GAN: An Approach for Stabilizing the Training Process of Generative Adversarial Network (Megahed, M; Mohammed, A, 2024)	X			
3	RAMT-GAN: Realistic and accurate makeup transfer with generative adversarial network (Yuan, Q-L; Zhang, H-L, 2022)	X			X
4	Generative adversarial network based synthetic data training model for lightweight convolutional neural networks (Rather, IH; Kumar, S, 2023)	X			X
5	Data Generation Based on Generative Adversarial Network with Spatial Features (Sun, L; Yang, Y; Mao, X; Wang, X; Li, J, 2023)	X			X
6	PConv: simple yet effective convolutional layer for generative adversarial network (Park, S; Yeo, YJ; Shin, YG, 2022)	X			
7	A data balancing approach based on generative adversarial network (Jamoos, M; Mora, A M; AlKhanafseh, M; Surakhi, O, 2023)	X			X
8	Current status, application, and challenges of the interpretability of generative adversarial network models	X			X
9	Data Augmentation Based on Generative Adversarial Network with Mixed Attention Mechanism (Yang, Y; Sun, L; (...); Zhao, M, 2022)	X			X
10	Triple Generative Adversarial Networks (Li, C; Xu, K; Liu, J; Zhu, J; Zhang, B, 2019)	X			
11	Effective shortcut technique for generative adversarial networks (Park, S; Yoo, CH; Shin, YG, 2023)	X			
12	Robust generative adversarial network (Zhang, SF; Qian, Z; (...); Lu, CY, 2023)	X			
13	Benefits of using blended generative adversarial network images to augment classification model training data sets (McCloskey, BJ, 2025)	X			X
14	Mutual learning generative adversarial network (Mao, L; Wang, M; Yang, D; Zhang, R, 2024)	X			
15	SCGAN: stacking-based generative adversarial networks for multi-fidelity surrogate modeling (Zhang, C; Liu, L; Wang, H; Song, X; Tao, D, 2022)	X			X
16	A deep data augmentation framework based on generative adversarial networks (Wang, QP; Luo, L; (...); Zhang, DT, 2022)	X			X
17	Synthesizing Realistic Knee MRI Images: A VAE-GAN Approach for Enhanced Medical Data Augmentation (Revathi, S A; Sathish Babu, B, 2024)	X	X	X	X
18	Multi-adversarial autoencoders: Stable, faster and self-adaptive representation learning (Wu, X; Jang, H, 2024)		X		

19	Frechet AutoEncoder Distance: A new approach for evaluation of Generative Adversarial Networks (Buzuti, LF; Thomaz, CE, 2023)	X	X		
20	ab-GAN: Robust generative adversarial networks (Zhang, SF; Qian, Z; (...); Lu, CY, 2023)	X			
21	Training generative adversarial networks by auxiliary adversarial example regulator (Gan, Y; Ye, M; Liu, D; Liu, Y, 2023)	X			
22	Augmenting data with generative adversarial networks: An overview (Creswell, A; White, T; (...); Bharath, AA, 2018)	X			X
23	Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers (Bengesi, S; El-Sayed, H; (...); Oladunni, T, 2024)	X	X		X
24	Conditional Generative Adversarial Networks and Deep Learning Data Augmentation: A Multi-Perspective Data-Driven Survey Across Multiple Application Fields and Classification Architectures (Ribas, LC; Casaca, W; Fares, RT, 2025)	X			X
25	The Evolution of Generative AI: Trends and Applications (Trigka, M; Dritsas, E, 2025)	X	X	X	X
26	Bi-EvoGAN: Bi-level Evolutionary Approach for Generative Adversarial Networks (Nouri, HE; Ghandri, A; Belkahla Driss, O; Ghedira, K, 2023)	X			
27	ARGAN: Adversarially Robust Generative Adversarial Networks for Deep Neural Networks Against Adversarial Examples (Choi, SH; Shin, JM; (...); Choi, YH, 2022)	X			X
28	Deep adversarial transition learning using cross-grafted generative stacks (Hou, JY; Ding, XJ; (...); Cranefield, S, 2022)	X	X	X	
29	Optimized Generative Adversarial Networks for Adversarial Sample Generation (Alghazzawi, DM; Hasan, SH; Bhatia, S, 2022)	X			X
30	A Pilot Study on the Use of Generative Adversarial Networks for Data Augmentation of Time Series (Forest, F; Lohr, C; (...); Korner, M, 2020)	X			X
31	Topological magnetic structure generation using VAE-GAN hybrid model and discriminator-driven latent sampling (Park, SM; Yoon, HG; (...); Won, C, 2023)	X	X	X	X
32	SUGAN: A Stable U-Net Based Generative Adversarial Network (Cheng, SJ; Wang, LF; (...); Meng, Y, 2023)	X			
33	Generative Adversarial Networks (GANs): Challenges, Solutions, and Future Directions (Saxena, D; Cao, JN, 2022)	X			X
34	A Hybrid Approach for Deep Generative Handwritten Arabic Text Recognition (Lamtougui, H; El Moubtahij, H; Fouadi, H; Satori, K, 2024)	X	X	X	X
35	Slice-to-voxel stochastic reconstructions on porous media with hybrid deep generative model (Zhang, F; Teng, Q; Chen, H; He, X; Dong, X, 2021)	X	X	X	X

Taulukko 1.1: Taulukko valituista generatiivisia malleja käsittelevistä artikkeleista.

Luvussa kaksi tullaan käsittelemään yleisesti, mitä ovat generatiiviset mallit ja miksi tämän kirjallisuuskatsauksen mallien käsittely on rajattu GAN- ja VAE-malleihin. Kolmannessa luvussa perehdytään tarkemmin siihen, mitä ovat generatiiviset vastakkaisverkot (GAN) [3]. Tässä luvussa kerrotaan myös tämän generatiivisen mallin toiminnasta ja sovelluksista. Neljännessä luvussa kerrotaan variaatiolisesta automaattikoodaimesta (VAE). Luvussa käydään läpi myös sen toimintaperiaatteet ja sovellukset. Viidennessä luvussa käsitellään valituista malleista tehtyjä hybridimalleja. Tässä luvussa käydään läpi, miten ne toimivat, sekä viime vuosien merkittävimpiä sovelluksia. Lisäksi luvussa kerrotaan myös mallien yhdistämisen hyödyistä. Kuudes luku on yhteenveto, jossa on pohdintaa aiheesta, havaintoja ja tutkimuskysymykseen vastaamista. Tämän lisäksi luvussa tarkastellaan tutkielmassa edellä mainittujen mallien eroja.

2 Generatiiviset mallit

2.1 Yleiskatsaus

Generatiiviset mallit ovat osa GAI:ta (Generative Artificial Intelligence), ja generatiivinen tekoäly käsitteenä sisältää kaikki mallit, jotka kykenevät luomaan dataa koulutukseen käytetyn datan perusteella. Generatiivisia malleja on lukuisia erilaisia, ja niiden lähestymistavat voivat poiketa paljon toisistaan [4]. Seuraavaksi käsitellään, mitä generatiiviset mallit yleisesti ovat ja miten ne toimivat. Kun puhutaan generatiivisista malleista, puhutaan yleensä koneoppimismalleista. Koneoppimismalli on algoritmi, joka kykenee tekemään datasta havaintoja automaattisesti. Näitä havaintoja ovat datan sisäiset rakenteet ja riippuvuudet. Koneoppimismallien toiminta perustuu suurten datamäärien läpikäymiseen ja niistä oppimiseen, tunnistuen datan piirteitä, rakenteita ja sääntöjä [1].

Yleisimmin koneoppiminen jaetaan kahteen pääluokkaan: ohjattuun ja ohjaamattomaan oppimiseen. Ohjatussa oppimisessa koneoppimismalli opetetaan tunnistamaan alkuperäisestä opetusmateriaalista ihmisen määrittämät ”oikeat” vastaukset, eli opetusmateriaali sisältää sekä piirteet että vastaukset. Ohjaamattomassa oppimisessä mallille taas ei anneta oikeita vastauksia, vaan mallin täytyy itse löytää datan rakenteet. Ohjaamattomien mallien hyöty nähdään varsinkin sovelluksissa, joissa mallin tulee toimia ilman ennakkokäsityksiä.

Generatiiviset mallit ovat menetelmiä, joissa malli oppii alkuperäisestä datasta ja

kykenee luomaan uutta dataa. Generatiiviset mallit ovat syväoppimismalleja, mikä on eräs koneoppimismallien tyyppi. Nämä mallit hyödyntävät neuroverkon kerroksia, jotta mallit oppivat datan piirteet ja rakenteet [1]. Generatiivinen tekoäly (GAI) on yksi tekoälyn osa-alue. GAI-mallit oppivat datan taustalla olevan todennäköisyysjakauman ja luovat sen pohjalta alkuperäistä dataa muistuttavia tuloksia [2]. Generatiivisen tekoälyn sopeutumiskyky mahdollistaa sen käytön useissa eri sovelluksissa. GAI:ta käytetään esimerkiksi tuottamaan tekstiä, kuvia, ääniä ja rakenteellisia tietoja. Taustalla oleva todennäköisyysjakauma tekee edellä mainitut mahdolliseksi [5]. GAI:n nopea kehitys ja nousu alkoi, kun vuonna 2014 Goodfellow esitteli GAN-mallin [2].

Ennen nykyaikaisia generatiivisia malleja on ollut jo generatiivisia malleja, jotka kuitenkin ovat hieman poikenneet nykyaikaisista. Näistä malleista esimerkkeinä ovat mallit, kuten Gaussian Mixture -malli ja Hidden Markov -malli. Aiemmissä malleissa oli merkittäviä rajoitteita, kuten laskennallinen monimutkaisuus ja suurten datamäärien käsittelyn ongelmat. Edellä mainitut rajoitteet aiheuttivat siirtymisen kohti syviä generatiivisia malleja. Syvät generatiiviset mallit käyttävät monikerroksisia neuroverkkoja ja takapropagaatiota saavuttaakseen tehokkaan oppimisen. Monikerroksisuudella tarkoitetaan, että malli koostuu kerroksista, jotka oppivat datasta eri tasoisia asioita. Alimmilla kerroksilla käsitellään yksinkertaiset rakenteet ja ylimmillä kerroksilla käsitellään abstraktit ominaisuudet [1] [2]. Malli kykenee monikerroksisuuden avulla oppimaan datan riippuvuuksia paremmin, jolloin tuotettu data on realistisempaa. Monikerroksisuuden kouluttaminen perustuu vastavirta-algoritmin toimintaan. Vastavirta-algoritmi on oppimisalgoritmi, jossa algoritmin malli oppii virheistään kuljettamalla virheen takaisinpäin verkossa ja säätämällä painotusta virheiden perusteella [1]. Neuroverkot ovat tärkeässä roolissa ja muodostavat tämän hetken tekoälyn perustan. Ne ovat mahdollistaneet jatkuvasti uusia läpimurtoja erityisesti kuvantunnistuksessa ja luonnollisten kielten käsittelyssä [2].

Generatiivisten mallien etuna on se, kuinka ne pystyvät oppimaan monimutkaisia jakaumia. Näiden jakaumien pohjalta pystytään tuottamaan korkean tason dataa. Kaikki generatiiviset mallit eivät toimi samalla tavalla. Kaksi tunnetuinta generatiivista mallia ovat variaatiollinen automaattikoodain (VAE) ja generatiiviset vastakkaisverkot (GAN). Näiden mallien toiminta on jo lähtökohtaisesti erilainen, vaikka molemmat ovat generatiivisia malleja. Variaatiollinen automaattikoodain pohjautuu todennäköisyysmallinnukseen, ja generatiivinen vastakkainasettelumalli perustuu kahden neuroverkon väliseen kilpailuun. Nämä kaksi mallia muodostavat tämän hetken generatiivisten mallien perustan. Työssä syvennyttään näiden mallien sovelluksiin irrallisina toisistaan ja yhdessä toistensa kanssa. GAN- ja VAE-malleja hyödyntäviä algoritmeja ja sovelluksia on tullut runsaasti viime vuosina. Viime vuosien aikana tutkimukset ovat keskittyneet myös Transformer-arkkitehtuureihin perustuvien GAN:ien kehittämiseen. Näiden menetelmien etuna on se, että ne pystyvät käsittelemään pitkän kantaman riippuvuuksia ja globaaleja rakenteita tarkemmin kuin aikaisemmat mallit. Malleissa on pysynyt pääpiirteittäin samat tavoitteet: halutaan tuottaa entistä realistisempia ja yksityiskohtaisempia ratkaisuja. Generatiiviset mallit ovat tärkeässä roolissa tekoälyn kehityksessä, datan analysoinnissa ja luomisessa [1] [2].

2.2 Rajaus GAN ja VAE

Tässä kirjallisuuskatsauksessa perehdytään tarkemmin generatiivisiin malleihin GAN ja VAE. GAN ja VAE edustavat molemmat merkittäviä menetelmiä syväoppimisessa. Näillä menetelmillä voidaan tuottaa dataa, mutta kuitenkin näiden menetelmien lähestymistavat generatiiviseen prosessiin ovat erilaiset. [1] [2]

Variaatiollinen automaattikoodain (VAE) on generatiivinen malli, jonka pohja on todennäköisydessä. Tämä malli yhdistää automaattikoodaimen rakenteen ja bayesilaisen päättelyn. Bayesilainen päättely on yksi suurimmista tilastotieteen

ajattelutapojen haaroista, ja se on tilastollinen lähestymistapa. Tässä lähestymistavassa alustavat uskomukset yhdistetään havaintoaineistoon Bayesin kaavan avulla. Lopputulokseksi muodostuu päivitetty uskomus tuntemattomasta suureesta [6]. Auto-maattikoodaimen rakenne koostuu kahdesta osasta, jotka ovat enkooderi ja dekodeeri. Enkooderi pakkaa datan pienempään latenttivektoriin. Tämän jälkeen dekodeeri rakentaa datan luodusta tiivistetystä datasta. Generatiivisen mallin tavoitteena on luoda dataa, joka on mahdollisimman lähellä alkuperäistä. Tämä menetelmä mahdollistaa vakuuttavien näytteiden tuottamisen [1].

Generatiivinen vastakkaisverkko (GAN) on generatiivinen malli, joka perustuu kahden neuroverkon väliseen kilpailuun. Nämä kaksi neuroverkkoa ovat generaattori ja diskriminaattori [4]. Generaattori on neuroverkko, joka pyrkii tuottamaan realistisia näytteitä. Diskriminaattori yrittää erottaa, ovatko näytteet generaattorin muodostamia vai aitoja. Tämä prosessi on toistuva. Tässä yritetään jatkuvasti parantaa generaattorin tulosta, tavoitteena huijata diskriminaattoria. Tavoite on luoda mahdollisimman realistista ja monipuolista dataa [2] [7] [8]. VAE:n painotus on tilastollisessa johdonmukaisuudessa, ja GAN:n keskittyminen enemmän yksityiskohtiin [5]. Näiden generatiivisten mallien piirteiden takia voidaan nähdä, että VAE-mallit soveltuvat enemmän datan sisäisen rakenteen tutkimiseen ja GAN-mallit kuvien ja videoiden tuottamiseen [2].

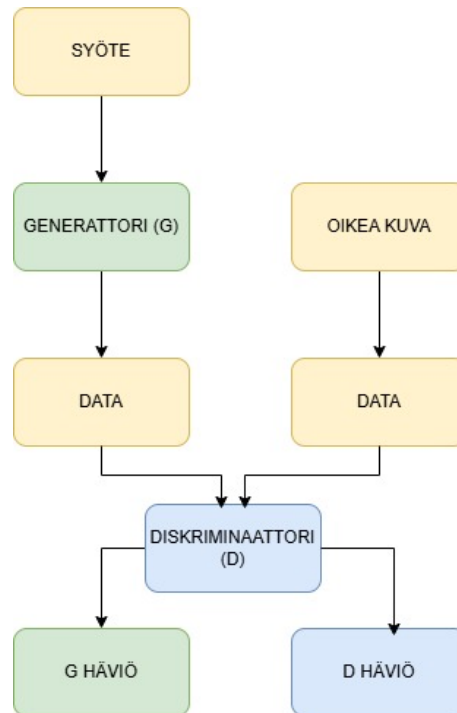
Viime vuosina tutkimuksissa on yhdistetty näitä malleja. Näitä GAN- ja VAE-yhdistelmämallia kutsutaan hybridimalleiksi, ja niistä on saatu vakuuttavia tuloksia viime vuosina. Nämä mallit ovat suoriutuneet yksittäisiä malleja paremmin useissa sovelluksissa. Hybridimalleilla pyritään tuottamaan realistista, tarkkaa, rakenteellisesti ja johdonmukaisesti olevaa dataa [1] [5]. Työ rajataan elokuuhun 2025 mennessä julkaistuihin artikkeleihin. Näiden kahden generatiivisen mallin yhdistäminen hybridimalleiksi avaa uusia mahdollisuuksia monelle eri alalle.

3 Generatiiviset vastakkaisverkot (GAN)

3.1 Toiminta

Generatiiviset vastakkaisverkot (Generative Adversarial Network) ovat generatiivinen malli [3]. Mallin esitteli Ian Goodfellow vuonna 2014 [4]. Tästä mallista on tullut merkittävä tekoälyn malli, ja se on tunnettu sekä käytetty ympäri maailmaa. Toiminta perustuu kahden keskenään kilpailevan neuroverkon toimintaan. Näitä neuroverkkoja kutsutaan nimellä generaattori (G) ja diskriminaattori (D) [8]. Nämä kaksi neuroverkkoa kommunikoivat toistensa kanssa. Kaksi neuroverkkoa toteuttaa *minimax*-pelin. GAN:ssa generaattori yrittää luoda dataa, jota diskriminaattori ei pysty erottamaan aidosta datasta. Tästä voidaan puhua koulutuksena; koulutus jatkuu, kunnes saavutetaan tasapainotila. Tasapainotila tässä tapauksessa tarkoittaa tilannetta, jossa diskriminaattori ei luotettavasti pysty erottamaan aitoa dataa ja generaattorin luomaa dataa [9] [10]. Yleiskielessä generaattoria voitaisiin ajatella väärennöksen tekijäksi ja diskriminaattoria asiantuntijaksi, joka yrittää paljastaa väärennöksen.

GAN:n toimintaperiaatteen tavoitteena on oppia monimutkainen todellisen datan jakauma $P_r(x)$. Tätä jakaumaa hyödynnetään sen jälkeen siten, että pyritään tuottamaan jakauman mukaisia näytteitä. Käydään järjestyksessä toiminta läpi ja



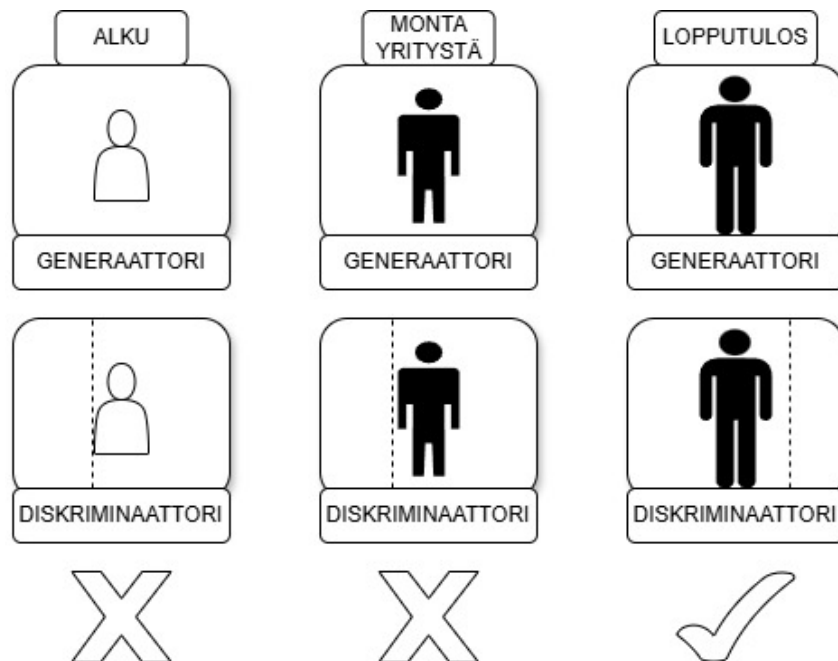
Kuva 3.1: GAN toimintaperiaate

aloitetaan siitä, että diskriminaattori (D) saa dataa jakaumasta. Diskriminaattori pyrkii tämän jälkeen antamaan mahdollisimman korkean todennäköisyyden, jossa $D(x) \rightarrow 1$ kaikille näytteille, jotka ovat aitoja. Seuraavaksi diskriminaattori saa generaattorilta näytteen $G(z)$. Diskriminaattorin tehtävä on pyrkiä antamaan mahdollisimman matalan todennäköisyyden $D(G(z)) \rightarrow 0$. Tämä tarkoittaa pyrkimystä paljastaa generaattorin luoma väärennös. Generaattori (G):n toiminta on päinvastainen. Generaattori pyrkii mahdollisimman lähelle tulosta $D(G(z)) \rightarrow 1$. Yhdessä nämä neuroverkot muodostavat tasapainotilan. Tällä tarkoitetaan tilaa, jossa saavutettua tulosta ei voi parantaa, muuten toisen tulos heikkenee [9] [10] [11].

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log (1 - D(G(z)))]$$

Tässä yläpuolella oleva lauseke on GAN-mallien matemaattisen toiminnan perusta [8]. Tämän generatiivisen mallin toiminnan ymmärtämisen kannalta on tärkeää

ymmärtää tämä lauseke. Tämän takia seuraavaksi käydään läpi tämä lauseke osissa, vasemmalta oikealle. Ensimmäinen osa on $\min_G \max_D V(D, G)$. Tässä osassa käytännössä kerrotaan, että kyseessä on ns. minimax-peli. Diskriminaattorin D tavoitteena on maksimoida hyvä tulos, ja generaattorin G pyrkimys on yrittää minimoida sen tulos. Yksinkertaisesti voidaan sanoa, että diskriminaattori yrittää erottaa aidon datan siitä datasta, mitä generaattori on luonut [12]. Seuraava läpi käytävä osa on $\mathbb{E}x \sim p_{\text{data}}(x)[\log D(x)]$. Tämä osa kertoo, kuinka hyvin diskriminaattori D pystyy tunnistamaan aidon datan. Kun $D(x)$ antaa tuloksen lähelle yhtä, malli pitää dataa todennäköisesti aitona [13], [14], [15]. Viimeinen osa tästä lausekkeesta on $\mathbb{E}z \sim p_z(z)[\log(1 - D(G(z)))]$. Tässä osassa kuvataan, kuinka G ja D toimivat yhdessä [13]. Lopputulos tässä tulee olemaan tasapainotila. Tässä tasapainotilassa G tai D ei voi enää parantaa suoritusta huonontamatta toisen suoritusta [12]. Alapuolella on havainnollistava kuva, kuinka GAN-mallin toiminta etenee.



Kuva 3.2: GAN toiminta

3.2 Sovellukset

Sovelluksia, joissa käytetään GAN-mallia tai jotka pohjautuvat GAN-malliin, on runsaasti. Tässä korostuu se, että GAN-malli on ollut esillä vuodesta 2014 asti [4]. Joka vuosi löydetään uusia tapoja hyödyntää tätä generatiivista mallia. Tämä osoittaa, kuinka merkittävä generatiivinen malli GAN:sta on tullut viime vuosien aikana. Nykyajassa teknologia kehittyy kovaa vauhtia eteenpäin, mutta tästä huolimatta GAN-mallit ovat pysyneet tärkeässä roolissa. GAN-mallin lähestymistapa datan tuottamiseen on syy siihen, miksi tämä malli on pysynyt relevanttina [7]. Sovelluksia, jotka pohjautuvat GAN-malliin, on tehty useille eri aloille. Seuraavaksi tarkastellaan eri aloja, joissa hyödynnetään GAN-mallia.

Lääketieteessä ja terveydenhuollossa on käytetty GAN-malleja useita vuosia. Tällaisia sovelluksia on esimerkiksi CT-angiografiakuvausten parantaminen [11]. Toinen terveydenhuoltoalan merkittävä sovellus on identiteetin tunnistukseen liittyvä malli, jolla voidaan suojata yksilön yksityisyyttä. Sotilasalalla GAN-mallia on hyödynnetty aaltopituuksien valintaan sekä signaalien ja piirteiden merkittävyysarviointiin. Rikosten ennustamisessa on saatu lupaavia tuloksia. Poikkeuksien havaitseminen reaaliajassa on viime vuosina ollut myös suosittu tutkimuskohde [16]. Tällaisessa myös ennustavassa mallissa muun muassa hyödynnetään alueiden tapahtumien vaikutussuhteita.

Tähtitieteessä on huomattu GAN-mallin hyöty galaksikuvien erottelussa. Kuva- ja multimodaaliset sovellukset ovat yksi merkittävä osa, josta löytyy sovelluksia, jotka hyödyntävät GAN-mallia. Näitä sovelluksia ovat esimerkiksi kuvien tuottaminen, tekstin muuttaminen kuviksi ja kasvojen ikääntymisen simulointi. Tietoturva-alalla on huomattu, miten GAN-malleja voidaan hyödyntää datan anonymisoinnissa ja tietosuojan parantamisessa [17].

GAN-malleista on kehitetty useita malleja ja versioita. Yksi näistä on GL-GAN. Tämän mallin nimi tulee sanoista Adaptive Global And Local Bilevel Optimization

GAN. GL-GAN:n tavoitteena on saada kuvien generoinnista parempilaatuista ja tehostaa koulutusta. Merkittävimmät uudet keksinnöt tässä mallissa ovat ”Feature map” ja ”ADA-OP”. Feature map on diskriminaattorin tekemä ns. kartta, joka kertoo, mitkä alueet kuvassa ovat hyviä ja mitkä huonoja. Tämän avulla generaattori saa korjattua huonot kohdat kuvassa muuttamatta koko kuvaa. ADA-OP eli Adaptive Global and Local Optimization on menetelmä, joka tekee päätökset, milloin korjataan huonot kohdat ja milloin kokonaiskuva. Yhdessä nämä menetelmät tekevät tuotetuista kuvista tasalaatuisempia, ja koulutus on tehokkaampaa [15].

GAN-malleja on valtava määrä. Yksi uuden lähestymistavan tuonut malli on SUGAN (Stable U-Net GAN). Tutkimuksissa on esitelty, kuinka tällä mallilla saadaan parempi vakaus kaatumista ajatellen ja nostettua generoitujen kuvien laatua. Tutkimuksessa tätä uutta lähestymistapaa verrattiin lukuisiin GAN-malleihin. Näitä malleja, joihin verrattiin, olivat nimeltä DCGAN, LSGAN, WGAN, WGAN-GP, SNGAN, GNGAN, U-Net GAN, HRGAN ja UWGAN-GP. SUGAN ei keksi GAN:n uutta rakennetta, vaan parantaa olemassa olevaa U-Net GAN:ia. Kaikkiin verrattuihin malleihin SUGAN suoriutui parhaiten yksityiskohdissa, laadussa ja vakaudessa.

Seuraavaksi käydään läpi lyhyesti U-Net-rakenne, johon tämä SUGAN perustuu. Tämä rakenne koostuu enkooderista ja dekooderista, mutta ne on yhdistetty skip-yhteydellä. Tämä skip-yhteys tarkoittaa neuroverkossa olevaa suoraa yhteyttä aikaisempien ja myöhempien kerrosten välillä. Tämä mahdollistaa datan hyödyntämisen välissä olevien kerrosten kautta. Suurin ero SUGAN-mallissa ja U-Net GAN:ssa on koulutuksen aikana käytettävä Gradient normalization. Tämä tarkoittaa sitä, että koulutuksen aikana arvoissa ei tapahdu suuria heilahduksia. Tämä varmistaa laadukkaamman oppimisen ja vakauden [18].

Koska GAN-malleja on runsaasti, on mahdotonta käsitellä kaikkia eri sovelluksia tässä tutkielmassa. Kuitenkin tässä tutkielmassa keskitytään viime vuosien merkittäviin lähestymistapoihin, jotka ovat edistäneet GAN-sovellusten kehitystä.

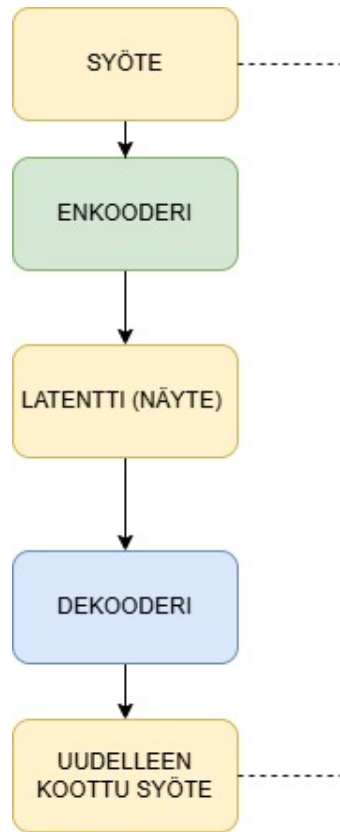
Äsken esiteltyt mallit ovat tuoneet parannuksia GAN-mallin tuottamaan laatuun, vakauteen ja tehokkuuteen. Esiteltyt mallit eivät kuvaa kokonaisuudessaan kaikkia, mutta havainnollistavat kuitenkin, miten GAN-malleissa on huomioitu haasteet ja pyritty kehittämään niitä. Tätä generatiivista mallia on hyödynnetty runsaasti eri aloilla. Merkittäviä esimerkkejä näistä aloista ovat lääketiede, tähtitiede, kuvantaminen ja tietoturva. Näitä sovelluksia on myös käsitelty antaakseen oikean kuvan GAN-sovellusten käyttömahdollisuuksista.

4 Variaatiollinen automaattikoodain (VAE)

4.1 Toiminta

Toinen malleista, johon keskitytään tässä työssä, on variaatiollinen automaattikoodain (Variational Autoencoder) eli VAE. VAE-mallin esittelivät Diederik P. Kingma ja Max Welling vuonna 2013 [19]. Esittelystä asti VAE on ollut keskeisessä roolissa generatiivisten mallien kehityksessä. VAE yhdistää syväoppimisen ja todennäköisyysmallin periaatteet. Kun puhutaan mallin toiminnasta, on keskeistä ymmärtää, mitä tarkoitetaan latenttitilalla, syväoppimisella, monikerroksisilla neuroverkoilla ja todennäköisyysmalleilla. Latenttitila on tila, jossa data on tiivistetty niin, että se sisältää vain datan tärkeimmät piirteet. Syväoppimisella tarkoitetaan monikerroksisten neuroverkkojen hyödyntämistä. Monikerroksiset neuroverkot pystyvät oppimaan monimutkaisiakin piirteitä ja suhteita annetusta datasta [13] [12]. Mainittujen todennäköisyysmallien hyöty on se, että saadaan perusta epävarmuuksien ja satunnaisuuksien osoittamiseen. Näiden kahden yhdistämisen ansiosta saadaan malli, joka tunnistaa, tiivistää ja generoi kontrolloidusti dataa [13] [14].

VAE-mallin tavoitteena on oppia todellisen datan $x \sim P_r(x)$ jakauma ja sen esitys z jatkuvassa latenttitilassa [13] [14]. VAE:n rakenne on kaksiosainen. Se koostuu kahdesta neuroverkosta, jotka ovat nimeltään enkooderi ja dekooderi. Molemmilla



Kuva 4.1: VAE toimintaperiaate

osilla on merkittävä tehtävä toiminnan kokonaisuudessa. Enkooderi $q_\phi(z|x)$ muuntaa syötedatan x latenttivektoriksi z , ja dekooderi $p_\theta(x|z)$ yrittää uudelleen rakentaa alkuperäisen datan latenttitilasta. Jokaisella kerralla, vaikka alkuperäinen data olisi sama, VAE tuottaa hieman erilaisen tuloksen [20]. VAE:n toiminta perustuu todennäköisyysmalliin. Tämän mallin tavoite on maksimoida datan todennäköisyys ottamalla huomioon kaikki mahdolliset muuttujan arvot. Tässä alapuolella oleva lauseke tunnetaan nimellä Evidence Lower Bound eli ”ELBO”. ELBO-lauseke on VAE-mallin matemaattisen toiminnan perusta, joten on merkittävää ymmärtää, mitä tämä lauseke tarkoittaa [21].

$$\log p_\theta(x) \geq \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - D_{KL}(q_\phi(z|x)||p(z))$$

Aloitetaan käymään läpi yläpuolella olevaa ”ELBO-lauseketta vasemmalta oikealle.

Vasemmalla reunalla oleva osa $\log p_\theta(x)$ tarkoittaa, että havaintomme x logaritmoitua todennäköisyyttä mallimme p_θ mukaan. Tässä halutaan, että malli tuottaa dataa, joka muistuttaa x :ää mahdollisimman hyvin [13] [14]. Tavoitteena on maksimoida log-todennäköisyys.

Jatketaan lausekkeessa oikealle päin ja siirrytään ensimmäiseen termiin oikealla puolella lauseketta. Tämä tarkoittaa odotusarvoa latenttimuuttujan z jakaumasta $q_\phi(z|x)$. Termi kertoo, kuinka hyvin mallimme $p_\theta(x|z)$ kykenee uudelleen rakentamaan alkuperäisen havainnon x . Voidaan ajatella, että tämä uudelleen kokoamisosa palkitsee mallia, jos se pystyy selittämään datan mahdollisimman hyvin muuttujien avulla [12].

Viimeisenä käsitellään lausekkeen oikean puolen seuraavaa osaa. Tämä osa kertoo, kuinka paljon posteriorijakauma $q_\phi(z|x)$ poikkeaa priorijakaumasta $p(z)$. Voidaan sanoa, että tämän toiminnan tarkoitus on toimia säännöllistäjänä. Säännöllistäjä estää mallia oppimasta liian yllättäviä latenttimuuttujia [13].

4.2 Sovellukset

VAE on ollut käytössä yli vuosikymmenen, mutta se on yhä laajasti käytössä oleva malli [19]. Sovelluksia on kertynyt paljon, ja koko ajan julkaistaan lisää. Tässä luvussa käydään läpi eri aloilla olevia sovelluksia. Käydään myös läpi, miten näitä sovelluksia voidaan hyödyntää.

Teollisuudessa on sovelluksia, joissa VAE mallintaa sensoridatan jakaumaa ja piirteitä. Tätä hyödynnetään kyberturvallisuuden parantamisessa. Poikkeamien tunnistaminen on tärkeää, esimerkiksi vesilaitoksissa tai energiaverkoissa, joissa vääränlainen arvo merkitsee usein virhetilaa [22]. Vuoden 2025 artikkelissa eriteltiin VAE-malliin perustuva malli, joka tunnistaa lämpövoimalaitosten turbiinien toiminnan poikkeavuudet. Tämä malli yhdistää neljä osaa. Nämä osat ovat Gaussian Mixture Modelin (GMM), Deep Autoencoder Local Outlier Factor (DAE-LOF), Long Short-Term

Memory Variational Autoencoder (LSTM-VAE) ja Deep Advanced Features (DAF). Nämä osat luovat Enhanced Long Short-Term Memory Variational Autoencoder with Deep Advanced Feature and Gaussian Mixture Modelin (ELSTMVAE-DAF-GMM). Tämä malli ei vaadi ulkoista valvontaa. LSTM-VAE:n tehtävä tässä sovelluksessa on tiivistää aikasarjadata. DAE-LOF on algoritmi, joka saa mallin erottelemaan paremmin oikeanlaiset poikkeamat. DAF:n tehtävä sovelluksessa on yhdistää latenttidatan ja koontamisvirheet; tämän avulla saadaan selkeämpi esitys datasta. GMM:n tehtävä on arvioida datan jakaumaa ja tehdä havaintoja normaalista poikkeavista arvoista. Tulokset osoittavat, että tämä sovellus tunnistaa hienovaraisetkin poikkeamat. Saatut tarkkuus oli 94,6 prosenttia, ja vain 5,43 prosenttia oli vääriä hälytyksiä. Tämä malli on todettu sopivaksi sovellukseksi teollisuuteen [23].

VAE-pohjaisia sovelluksia voidaan myös käyttää satelliittitutkimuksessa ja maanmittauksessa. Yksi sellainen sovellus on tutkateknologia InSAR. Tämä tulee sanoista ”Interferometric Synthetic Aperture Radar”. Tässä teknologiassa on hyödynnetty VAE-mallia kohinan ja signaalien puhdistuksessa. Sovelluksessa on hyödynnetty InSAR-teknologiaa; tällä on saatu parannettua tarkkuutta, skaalautuvuutta ja mahdollistettua reaaliaikaisuus [24].

Pankkialalla pystytään hyödyntämään merkittävästi VAE-mallia. Tällä alalla VAE:ta käytetään luottoriskin arvioinnissa. Sovelluksen tavoitteena on parantaa maksuhäiriöiden tunnistamista. VAE-mallin rooli tässä sovelluksessa on ominaisuusvektorien tiivistys ja latenttien edustuksen oppiminen. Neuroverkko luokittelee pankin asiakkaat maksukykyisiin ja maksukyvyttömiin asiakkaihin. Tässä käytetään VAE:n tekemää latenttitilaa [25].

Seuraavaksi käsitellään lähestymistapaa, joka parantaa VAE:n suorituskykyä käyttäen adversaarista oppimista. Tätä menetelmää kuvataan nimellä SAVAE (Semi-Supervised Adversarial Variational Autoencoder). Pohjimmiltaan tämä on VAE-sovellus, vaikkakin tähän menetelmään on yhdistetty GAN-tapainen osa. Edellä

mainitun takia tutkielmassa tämä sovellus on laitettu VAE:n sovellusosioon. GAN:n rooli on tuoda SAVAE:hen kilpaileva komponentti, jotta saadaan tarkempi ja realistisempi lopputulos. Käytännössä tämä tarkoittaa tilannetta, jossa dekooderin tarvitsee voittaa diskriminaattori [26].

Yksi sovelluksista, joka antaa hyvän kuvan monipuolisuudesta, mitä VAE-sovelluksilla voidaan luoda, on vuonna 2024 esitelty sovellus. Tämä sovellus yhdistää usean muotoista dataa ja dynaamisen VAE:n. Tätä mallia kutsutaan Multimodal Dynamical VAE:ksi (MDVAE). MDVAE-mallin etu on, kuinka se käsittelee ajallista dataa ja usean muotoista dataa samanaikaisesti. MDVAE:ta on testattu sovelluksissa, joissa latenttitila jakaa datan pysyviin ja muuttuviin ominaisuuksiin. Pysyviä ominaisuuksia on esimerkiksi puhuja itse, ja muuttuvia ominaisuuksia on esimerkiksi kasvonliikkeet. Tulokset saatiin audiovisuaalisen datan käsittelystä. Tulokset osoittavat, että MDVAE toimii paremmin kuin yksittäiset mallit [27].

VAE:ta on hyödynnetty kyberturvallisuusosalalla. Vuonna 2025 esiteltiin FedSAGE-niminen tunkeutumisen havaitsemisjärjestelmä. Tämä järjestelmä luotiin tunnistamaan DGA (Domain Generation Algorithm) -pohjaiset haittaohjelmat. Järjestelmässä hyödynnetään VAE:ta niin, että se tuottaa jokaisen käyttäjän verkkodatasta latenttivektorin. Tämän jälkeen järjestelmä käyttää latenttivektoreita vertailussa toisiinsa. Vertailun jälkeen samankaltaiset verkkodatasta saadut latenttivektorit lajitellaan samoihin ryhmiin. Tämä helpottaa tunnistamista, eikä käyttäjien tarvitse jakaa raakadataansa. Tutkimuksen tulokset osoittavat, että tällä mallilla saadaan yli kymmenyksen paremmalla tarkkuudella tuloksia, ja energiankulutus on huomattavasti pienempi kuin vanhoilla malleilla [28] [29].

Generatiivisia malleja, joissa käytetään VAE-mallia pohjalla, on tehty jo yli vuosikymmenen ajan. Vuosikymmenen saatosta voidaan huomata, kuinka sovellusalue on laajentunut. Sovellusten valtavan määrän takia tutkielmaan on valittu esimerkkejä VAE-sovelluksista, jotka pyrkivät antamaan kuvan tämän hetken tilanteesta.

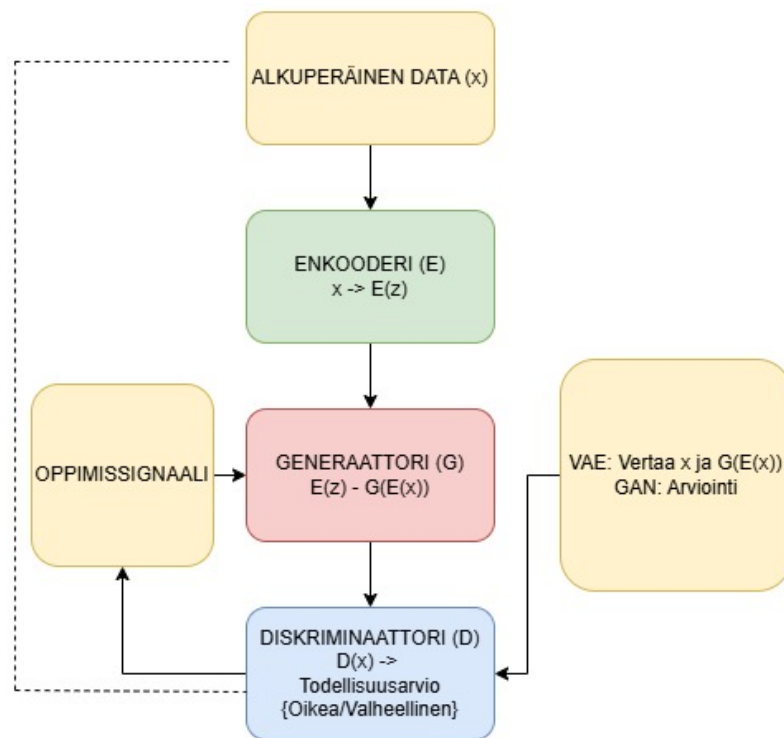
Teollisuus ja kyberturvallisuus ovat merkittävät alat, joilla päästään hyödyntämään VAE:n perustuvia malleja. Teollisuudessa ja kyberturvallisuudessa hyödynnetään generatiivisen mallin kykyä hallita sensoreista saatua dataa, poikkeuksien tunnistamista ja piirteiden mallintamista. Uusimmissa esitetyissä sovelluksissa on löydetty VAE:n hyöty myös pankkisektorilla [25]. Latenttitila ja sen tulkittavuus ovat keskeisiä VAE:ssa. Tämä näkyy myös uusissa sovelluksissa, joiden tutkimussuunta on esimerkiksi biologisessa datassa [30]. Annetut esimerkit kuvaavat, kuinka VAE-malleja voidaan jatkossakin hyödyntää monipuolisuutensa vuoksi.

5 GAN ja VAE - hybridimallit

Tässä luvussa tutustutaan siihen, mitä ovat GAN- ja VAE-hybridimallit. Sen jälkeen tarkastellaan, mikä on näiden mallien yhdistämisen tarkoitus. Seuraavaksi aloitetaan kartoittamalla, miten tällaiset mallit toimivat. Kartoittamisen jälkeen käsitellään, mitä hybridimalleista hyödytään ja mitkä ovat GAN:n ja VAE:n roolit näissä malleissa. Yleisellä tasolla hybridimallien hyödyt näkyvät siinä, että niillä voidaan tuottaa tarkempia ja monipuolisempia tuloksia. Hybridimallilla on saatu paremmat tulokset tekstin generoinnissa kuin yksittäisillä malleilla tai niiden sovelluksilla [31]. Hybridimalleilla on myös pystytty luomaan laadukkaampia ja monipuolisempia kuvia. Tähän merkittävä erikoistunut generatiivinen malli on nimeltään Conditional Variational Autoencoder GAN (CAPGAN). CAPGAN-mallilla on saatu karsittua aiempien mallien ongelmia. Aiemmissä sovelluksissa on ollut haasteita epävakaudessa ja tulosten vinoutumisessa. Tulosten vinoutuminen on useimmiten ollut ongelma, kun tuotetaan kuvia sellaisista kohteista, mistä on vain vähän havaintoja [32]. Samalla on tärkeää huomioida myös haasteet, joita mallien yhdistämisestä seuraa, kuten koulutuksen vaativuus. Seuraavissa kappaleissa käsitellään näitä asioita tarkemmin.

GAN- ja VAE-malleja on yhdistetty useamman vuoden ajan. Molemmissa järjestelmissä on rajoitteita yksinään, joten on luotu hybridimalleja. VAE:n puolelta saadaan laaja kattavuus ja monipuolisuus. GAN:n puolelta saadaan realismi ja vahva uskottavuus. Seuraavaksi käsitellään tarkemmin hybridimallin sisäistä rakennetta. Hybridimalli koostuu kolmesta eri hermoverkosta. Nämä hermoverkot ovat

enkooderi, generaattori ja diskriminaattori. Enkooderi E vastaanottaa annetun datan ja laittaa sen tiivistettyyn muotoon, eli datasta x muuttuu data z . Tämän jälkeen generaattori G uudelleen rakentaa näytteen koodista z , jolloin muodostuu $x = G(z)$ -data. Seuraavaksi, kun data on muodostettu, diskriminaattori D arvioi tuotetun datan ja tunnistaa, onko kyseessä realistinen vai keinotekoinen data. Hybridimallin koulutusprosessi on yhdistelmä GAN:n ja VAE:n koulutusprosesseista. Tällä hybridimallilla pyritään tuottamaan rakenteellisesti ja ominaisuuksiltaan oikeita tuloksia [33].



Kuva 5.1: Hybridimallin toimintaperiaate

Seuraavaksi tarkastellaan sovelluksia, joita on tehty hybridimalleista. Aloitetaan tarkastelemalla, kuinka Multi-Adversarial Autoencoder (MAAE) toimii. Tämän mallin perusta on samankaltainen kuin edellisessä kappaleessa käsitelty, mutta tässä mallia laajennetaan ottamalla käyttöön useita diskriminaattoreita D yhden sijaan. MAAE on todettu sopivaksi muun muassa datan augmentointiin. Tällaisia sovelluksia on esimerkiksi lääketieteellinen kuvantaminen. Koska tässä mallissa on useita

diskriminaattoreita D , näistä jokainen toimii itsenäisesti. Näiden erilliset arviot yhdistetään menetelmällä, jossa malli oppii itse, kuinka paljon millekin D :lle annetaan. Tätä kutsutaan soft-ensemble-menetelmäksi. Alla oleva kaava kertoo, kuinka kokonaisarvio saadaan painottamalla jokaisen D :n antama arvio [34].

$$L_{\text{ensemble}} = \sum_i \text{SoftMax}_i(\alpha) L_{D_i}(z, z_{\text{prior}})$$

Koulutus etenee kahdessa vaiheessa. Vaiheet ovat rekonstruktio- ja adversaarivaihe. Rekonstruktio-vaiheessa E ja D opetetaan toimimaan niin, että data vastaa mahdollisimman hyvin alkuperäistä dataa. Tätä tehdään pienentämällä rekonstruktiotappiota. Rekonstruktiotappio kuvaa alkuperäisen ja rekonstruktioidun datan eroa. Seuraava vaihe on adversaarinen vaihe; tässä vaiheessa E yrittää huijata D :tä. MAAE:n innovaatio on se, kuinka se osaa etsiä automaattisesti tasapainon datan tiedon ja jäljittelyn laadun välille. Rajoitteita tässä uudessa mallissa ovat toiminnasta johtuva kasvanut laskennallinen kuormitus sekä koulutuksen pidempi kesto. Nämä ovat haittapuolet, jotka syntyvät siitä, että rinnakkaiskäytössä on useampia D :tä [34].

Seuraavaksi tarkastellaan hybridimallia nimeltä CGVAE-label. Tämä tulee sanoista "Completion GAN Variational Autoencoder - label". Tämän mallin tarkoitus on täydentää puuttuvat tiedot laajamittaisessa ionosfäärin TEC-kartassa. TEC-kartta on sellainen, joka näyttää elektronien määrän ionosfäärissä globaalisti. Puuttuvia tietoja syntyy, kun ei ole riittävästi satelliitteja eikä vastaanottimia mittaamaan alueita. Voidaan todeta, että tässä mallissa hyödynnetään hybridimallia täysin eri tavalla kuin MAAE:ssa [35].

Seuraavassa tutkielmaan valikoidussa tutkimuksessa esitellään malli, joka tuottaa kolmiulotteisen materiaalirakenteen vain yhdestä kaksiulotteisesta kuvasta. Toimiesaan tämä malli kykenee korvaamaan kalliita ja hitaita röntgen- tai CT-kuvauksia. Malli rakentuu kolmesta osasta, joita aiemmassa kappaleessa käsiteltiin. Tässä mallissa hyödynnetään molempien mallien parhaita puolia. VAE:n etuna on, että saadaan

luotua hyvä latentitiedustus datasta. GAN-mallista saadaan hyöty vastakkainasettelun avulla, kun saadaan tuotettua parempaa kuvanlaatua ja realistisempia tuloksia. Tämän tutkimuksen mallia koulutettiin ja testattiin kahdella eri datalla. Näitä dataa olivat isotrooppiset ja anisotrooppiset. Isotrooppisella tässä tapauksessa tarkoitetaan materiaalia, jonka ominaisuudet ovat joka suuntaan samat. Anisotrooppisella tarkoitetaan tässä tapauksessa sitä, että materiaalin ominaisuudet vaihtelivat eri suuntien mukaan. Tutkimuksessa todettiin, että tämä malli on hyvin tehokas ja uudelleenkäytettävä. Haaste, joka vaatii jatkokehitystä tulevaisuudessa, on suurten kolmiulotteisten kuvien tuottaminen. Tässä tulee haasteiksi laskennallinen vaativuus ja GPU-muistin tuottamat rajoitukset [36].

Seuraava hybridimalli, jota käsitellään, on EvaGoNet. Tämä on piirteidensuunnittelukehys (feature engineering framework), joka on tehty taulukkomuotoiselle datalle. Taulukossa data on jaettu havaintoihin ja piirteisiin. Piirreinsinöörikehyksellä tarkoitetaan sellaista prosessia, jossa alkuperäisestä datasta muodostetaan paremmin ennustamiseen sopivia piirteitä. Tämä malli on myös hybridimalli, jossa yhdistyy VAE ja GAN. Kuitenkin ero moniin muihin malleihin tulee siinä, mitä versiota GAN-mallista käytetään. Tässä mallissa käytetään ns. parannettua versiota GAN-mallista, eli WGANgp:ää. WGAN on julkaistu vuonna 2017, ja tämän jälkeen edelleen kehitetty malli on WGANqp. Tämä malli eroaa siten, että alkuperäinen häviöfunktio on korvattu Wasserstein-etäisyydellä. Wasserstein-etäisyys on matemaattinen mitta, joka kuvaa kahden todennäköisyysjakauman välistä eroa. Tässä mallissa on myös ”gp”-osa. Tämä osa tulee sanoista ”gradient penalty”, eli gradientti-hyvyys. WGANqp:n hyöty alkuperäiseen versioon verrattuna on siinä, että se on vakaampi ja tuottaa laadukkaampia piirteitä. Tutkimuksessa esitelty kehys EvaGoNet osoittaa, että syväoppimiseen perustuvalla piirteiden muodostuksella voidaan täydentää tai korvata asiantuntijoiden tekemiä piirrevalintoja. Tämä luo edellytykset tehokkaammalle mallien kehittämiseksi. [37]

Hybridimalleja on viimeisen viiden vuoden aikana tullut valtavasti esille. Yksi merkittävä osa-alue on myös erikseen liitettävät kytkennät. Kytkentöjä liittämällä alkuperäisestä generatiivisesta mallista saadaan tehokkaampi. Vuonna 2025 julkaistiin tutkimus, jossa esiteltiin PluGeN-mallia. PluGeN-malli voidaan liittää ennalta koulutettuihin malleihin. Tämä malli ei muuta alkuperäisen mallin, kuten hybridimallin, VAE:n tai GAN:n, arvoja. PluGeN liitetään näiden mallien latenttitilaan. Liittämisen jälkeen kyetään muuttamaan yksittäisiä ominaisuuksia tai piirteitä kokonaan erikseen. Tämä lähestymistapa ratkaisee alkuperäisten mallien ongelman antamalla täyden kontrollin toiminnasta. Tutkimuksessa tätä mallia ehdotetaan biolääketieteen sovelluksiin, kasvontunnistukseen ja kemiallisiin suunnitteluihin sopivaksi generatiiviseksi malliksi. [38]

Edellä mainitut mallit osoittavat, kuinka eri aloille voidaan kehittää hyödynnettäviä malleja. Esiteltyt hybridimallit osoittavat, miten ajankohtaiset VAE- ja GAN-mallit ovat vieläkin. Sovellukset, joita tutkielmassa on esitelty, eivät edusta kokonaan kaikkia sovelluksia, joita on kehitetty. Kuitenkin näissä tuodut uudet parannukset antavat selkeän kuvan hybridimallien tulevaisuuden suunnasta. Esiteltyt parannukset, kuten useiden D hyödyntäminen, kilpailulliset optimointimenetelmät ja Wasserstein-etäisyys gradienttihyväksynnällä, ovat hyviä esimerkkejä sovelluksista, jotka mahdollistavat entistä parempia tuloksia. Uudet tutkimukset osoittavat, kuinka yhdistämällä nämä mallit saadaan tuotettua laadukkaampia tuloksia [39]. Tulosten parantaminen hybridimalleilla on merkittävä osa generatiivisten mallien tulevaisuutta. Näitä malleja ei ole hylätty, vaan niiden kehitystä on jatkettu eteenpäin uusien innovaatioiden myötä. Hybridimallien laskennallinen kuormitus, muistin määrä ja koulutuksen mahdollinen monimutkaisuus huomioon ottaen on selvää, että tulevaisuudessa tarvitaan jatkotutkimuksia, joiden tavoitteena on optimointi.

6 Analyysi

Tässä tutkielman osiossa käydään läpi kokonaisuutta näiden generatiivisten mallien sovellusten osalta. Tarkoituksena tässä on tuoda esille, mitä etuja ja haasteita hybridimallit tuovat sekä missä osa-alueissa erilliset mallit ovat yhä hyödyllisiä. Hybridimallien koulutus on haastavampaa, muistin määrä, mitä vaaditaan, on suurempi, ja niiden laskennallinen kuormitus on suurempi [36]. Tämän takia on tärkeää tunnistaa sovellusten heikkoudet ja vahvuudet eri toimintaympäristöissä. Alapuolella on koottu taulukko, joka antaa selkeämmän kuvan eri malleista, hyödyistä, haasteista ja sopivuuksista.

MALLI	SOVELLUKSET	HYÖDYT	HAASTEET	KÄYTTÖKOhteET
GAN	<ul style="list-style-type: none"> • Kuvagenerointi ja muokkaus • Videoiden parantaminen • Lääketieteellinen kuvantaminen • Signaalien ja aaltopitoisuuden optimointi 	<ul style="list-style-type: none"> • Tuottaa realistisia ja yksityiskohtaisia tuloksia • Soveltuu monimutkaisen datan mallintamiseen 	<ul style="list-style-type: none"> • Koulutus voi olla epävakaa • Saattaa tuottaa epärealistisia tuloksia 	<ul style="list-style-type: none"> • Visuaaliset ja multimodaaliset sovellukset • Realismia ja yksityiskohtia vaativat sovellukset
VAE	<ul style="list-style-type: none"> • Datan pakkaus ja tiivistys • Poikkeamien tunnistus • Pankkisovellukset • Kyberturvallisuus • Latentin esityksen oppiminen 	<ul style="list-style-type: none"> • Laaja latenttilata • Hyvä datarakenteiden mallinnus • Mahdollistaa kontrolloidun generoinnin 	<ul style="list-style-type: none"> • Tuottaa usein sumeaa dataa • Vähemmän realistinen • Parametrien tarkka valinta 	<ul style="list-style-type: none"> • Sovellukset, joissa tarvitaan datan sisäisten suhteiden mallintamista ja hallintaa
HYBRIDIMALLIT	<ul style="list-style-type: none"> • Kaikki edellä mainitut • Sopii monimutkaisiin ja kriittisiin sovelluksiin 	<ul style="list-style-type: none"> • Yhdistää GAN:in visuaalisuuden ja VAE:n hallittavuuden 	<ul style="list-style-type: none"> • Suuri laskennallinen kuormitus • Korkeat koulutusvaatimukset 	<ul style="list-style-type: none"> • Sovellukset, joissa tarvitaan sekä realismia että tarkkuutta

Kuva 6.1: Mallien vertailutaulukko

Yllä olevasta kuvasta nähdään, kuinka kaikilla malleilla on selkeästi omat vahvuutensa. GAN-mallit erottuvat edukseen selkeästi yksityiskohtien ja realististen tulosten

tuottamisessa. Voidaan olettaa, että visuaaliset sovellukset ovat tämän mallin vahvin osa-alue. VAE-mallien vahvuus näkyy datan rakenteiden ymmärtämisessä. Hybridimallit omaavat parhaat puolet molemmista, kuitenkin yhdistämisestä johtuvat haasteet ovat tärkeitä muistaa. Voidaan todeta, että generatiivinen malli tulee valita kohteen mukaan.

Käydään seuraavaksi muutamia konkreettisia esimerkkejä läpi, jotta saadaan luotua selkeä kuva tilanteesta. Jos tarvitaan esimerkiksi nopeaa ja yksinkertaista suorituskykyä, kuten suorissa lähetyksissä, niin kannattanee valita GAN-malli. Hybridi ei välttämättä olisi tarpeeksi nopea, eikä tämä toimenpide tarvitse niin monimutkaista toimintaa. VAE loistaa datan poikkeamien havainnoinnissa; tällaisissa sovelluksissa voidaan nähdä, että ei ole tarvetta hybridisovelluksille. Kuitenkin hybridimallien hyöty ja tärkeys tulee esille, kun tarvitaan tarkkaa toimintaa. Tällaisia sovelluksia voi olla esimerkiksi lääketieteen kuvantaminen. Hybridimallien etu on siinä, että esimerkiksi visuaalisissa sovelluksissa ei tarvitse tinkiä laadusta.

7 Yhteenveto

7.1 Pohdinta ja havainnot

Tutkielma käsittelee tekoälyn osa-aluetta GAI:ta. Kirjallisuuskatsauksessa tarkasteltiin generatiivisia malleja GAN ja VAE sekä niiden sovelluksia erikseen ja yhdessä. Tavoitteena oli selvittää näistä malleista tehtyjen sovellusten tilannetta, rajoittaen materiaalin julkaisupäivän tutkielman tekoon elokuuhun 2025. GAI on viime vuosien suurten kielimallien takia noussut jokaisen ihmisen tietoisuuteen. Ihmiset eivät välttämättä sitä itse tunnista, koska tuntevat vain suuret sovellukset nimeltä, kuten ChatGPT:n ja Copilotin. Näiden tulo markkinoille on ollut myös avain suurempiin rahoituksiin ja luonut mahdollisuudet kehittää entistä pidemmälle generatiivisia malleja. Tässä huomataan hyvin yhteiskunnan vaikutus tälle sektorille. Kirjallisuuskatsauksen tulokset osoittavat, kuinka tärkeässä roolissa GAN- ja VAE-generatiiviset mallit ovat vieläkin. Tässä tutkielmassa ei käsitelty diffuusiomalleja tai transformer-pohjaisia malleja. Näiden mallien suosio on kuitenkin viime vuosina nostanut ne keskeiseen rooliin GAN- ja VAE-mallien rinnalle. Viime vuosina on noussut myös esille diffuusiomallien vakaas ja suorituskyky verrattuna alkuperäiseen GAN-malliin [40]. Selkeän vahvan perustansa ansiosta, varsinkin hybridimallit, ovat varmasti myös tulevaisuudessa tärkeitä. Koko ajan löydetään uusia käyttötarkoituksia tai saadaan kehitettyä ”vanhaa” tukien tehokkaammaksi ja paremmaksi. Kun puhutaan generatiivisista malleista ja yleisestikin ihmisen tekemistä sovelluksista ja koneista,

niiltä ei hyväksytä heikkouksia tai virheitä. Tässä nähdään hybridimallien tarjoavan jatkuvasti parempia vaihtoehtoja.

Tutkielmasta voidaan todeta, että generatiiviset mallit pääsevät loistamaan parhaiten osana isompia kokonaisuuksia. Niiden tarkoitus ei ole tulevaisuudessakaan olla yksinään ratkaisuja. Malleilta vaaditaan kuitenkin laatua, tehokkuutta, varmuutta ja kykyä mukautua useaan eri tilanteeseen. Yhteiskunnalliset vaikutukset tulevat lisääntymään tulevina vuosina, kun generatiivisten ja varsinkin hybridimallien sovellukset tulevat käyttöön yhä useampaan paikkaan sekä tarkoitukseen. Tutkielmassa ei tarkasteltu eettisiä kysymyksiä, ja jatkotutkimuksen kannalta datan oikeellisuuden tarkistus on asia, jonka merkitys yhteiskunnassa lisääntyy. Tulevaisuuden katsauksissa kannattanee myös keskittyä GAN- ja VAE:n yhdistämisen uusiin malleihin, kuten transformer-pohjaisiin malleihin.

7.2 Tutkimuskysymyksiä vastaukset

TK1-kysymyksessä käsiteltiin valittujen mallien eroja toimintaperiaatteeltaan ja rakenteeltaan. Tutkielman luvut 3.1 ja 4.1 antoivat selkeän kuvan näiden mallien toiminnasta. GAN- ja VAE-mallien erot ovat rakenteelliset. VAE-malli koostuu kahdesta osasta, jotka ovat enkooderi ja dekooderi. VAE:n toiminta perustuu datan jakamiseen todennäköisyyspohjalta [20]. Tässä mallissa saadut tulokset saavutetaan tasapainottelemalla datan tarkkuuden ja latenttitilan säännöllisyyden välillä. GAN-malli puolestaan koostuu generaattorista ja diskriminaattorista. Nämä kaksi osaa kilpailevat keskenään [2]. Tällä kilpailutilanteella saadaan tuotettua terävämpiä ja realistisempia tuloksia kuin VAE-mallissa. VAE:n tulokset ovat usein pehmeämpiä ja sumeampia, mikä johtuu datan tasoittamisesta. GAN:n tulokset puolestaan ovat usein realistisempia. Kuitenkin siinä, missä VAE:n koulutus on selkeämpää, GAN:n koulutus saattaa olla epävakampaa.

TK2-kysymyksessä käsiteltiin näiden mallien sovelluksia erikseen. Tähän kysy-

mykseen vastattiin luvuissa 3.2 ja 4.2. Näitä sovelluksia on tullut valtava määrä vuosikymmenen aikana. Niissä sovelluksissa myös huomataan, kuinka alkuperäistä mallia on kehitetty eteenpäin, jotta saadaan paremmin sopiva ja parempaa tulosta antava malli. Ei ole vain yhtä alaa, jossa hyödynnetään GAN- ja VAE-mallien sovelluksia, vaan niitä on useita. Kirjallisuuskatsauksen perusteella voidaan todeta, että VAE- ja GAN-malleja hyödynnetään painottuen eri tehtäviin. Tyypillisesti VAE-malleja käytetään enemmän sovelluksissa, joihin liittyy datan jakaumien mallintaminen tai datan pakkaaminen. GAN-malleja puolestaan hyödynnetään enemmän visuaalisissa sovelluksissa. Tällaisia sovelluksia ovat esimerkiksi kuvien generointi ja kuvien korjaus.

TK3-tutkimuskysymyksessä käsiteltiin hybridimallien hyötyjä ja mahdollisuuksia. Tätä aihetta käsittelee kokonaisuudessaan tutkielman luku viisi. Hybridimallien tarkastelussa korostui niiden sovellusten mahdollisuuksien laajuus. Yhdistämällä kaksi eri mallia pystytään tuottamaan asioita ja sovelluksia, joita yksinään ei pystytä. Hybridimallien ideana on, että saadaan hyödynnettyä molempien generatiivisten mallien parhaat puolet ja vahvistettua osa-alueita, joissa ne ovat heikompia [34]. Sovellusesimerkeistä juuri lääketieteellinen kuvantaminen on hyvä, koska siinä tarvitaan realistista generointia ja enemmän tarkkuutta. Hybridimallit tuovat monimutkaisuudessaan laskennallista kuormitusta, joten näiden mallien tasapainottaminen voi olla hankalampaa kuin yksittäisten mallien [36].

Lähdeluettelo

- [1] R. Mehmood, R. Bashir ja K. J. Giri, ”Deep Generative Models: A Review”, *Indian Journal of Science and Technology*, vol. 16, nro 7, s. 460–467, 2023. DOI: 10.17485/IJST/v16i7.2296. url: <https://doi.org/10.17485/IJST/v16i7.2296>.
- [2] S. Bengesi, H. El-Sayed ja T. Oladunni, ”Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers”, *IEEE Access*, vol. 12, s. 69 812–69 837, 2024. url: <https://ieeexplore-ieee-org.ezproxy.utu.fi:2443/stamp/stamp.jsp?tp=&arnumber=10521640>.
- [3] Microsoft, *What is Generative AI?*, Viitattu: 21.10.2025, 2025. url: <https://www.microsoft.com/fi-fi/ai/ai-101/what-is-generative-ai>.
- [4] I. J. Goodfellow et al., ”Generative Adversarial Networks”, *arXiv preprint arXiv:1406.2661*, 2014. url: <https://doi.org/10.48550/arXiv.1406.2661>.
- [5] M. Trigka ja E. Dritsas, ”The Evolution of Generative AI: Trends and Applications”, *IEEE Access*, vol. 13, s. 98 504–98 529, 2025. url: <https://ieeexplore-ieee-org.ezproxy.utu.fi:2443/stamp/stamp.jsp?tp=&arnumber=11016906>.
- [6] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari ja D. B. Rubin, *Bayesian Data Analysis*, 3rd. Chapman ja Hall/CRC, 2013. url: <https://sites.stat.columbia.edu/gelman/book/BDA3.pdf>.

- [7] H. Yu, Y. Gu, M. Li, S. Deng, W. Lu ja Y. Heng, "Recovering Missing Passenger Flow Data in Subway Stations via an Enhanced Generative Adversarial Network", *IET Intelligent Transport Systems*, vol. 19, nro 1, 2025. DOI: 10.1049/itr2.70005.
- [8] Y. Yang, L. Sun, X. Mao ja M. Zhao, "Data Augmentation Based on Generative Adversarial Network with Mixed Attention Mechanism", *Electronics*, vol. 11, nro 11, s. 1718, 2022, Published: 20 June 2022. DOI: 10.3390/electronics11111718. url: <https://www.mdpi.com/2079-9292/11/11/1718>.
- [9] A. T. Gnanha, W. Cao, X. Mao, S. Wu, H.-S. Wong ja Q. Li, " $\alpha\beta$ -GAN: Robust generative adversarial networks", *Information Sciences*, vol. 593, s. 177–200, 2022. DOI: 10.1016/j.ins.2022.177200. url: <https://www.sciencedirect.com/science/article/pii/S0020025522001128?via%3Dihub>.
- [10] S. Zhang et al., "Robust Generative Adversarial Network", *Machine Learning*, vol. 112, s. 5135–5161, 2023. DOI: 10.1007/s10994-023-06367-0. url: <https://link.springer.com/article/10.1007/s10994-023-06367-0>.
- [11] S. Islam et al., "Generative Adversarial Networks (GANs) in Medical Imaging: Advancements, Applications, and Challenges", *IEEE Access*, vol. 12, s. 35 728–35 753, 2024. DOI: 10.1109/ACCESS.2024.3370848.
- [12] M. Megahed ja A. Mohammed, "Collaborative-GAN: An Approach for Stabilizing the Training Process of Generative Adversarial Network", *IEEE Access*, vol. 12, s. 138 716–138 735, 2024. DOI: 10.1109/WSC.2010.5678901. url: <https://ieeexplore.ieee.org/document/5678901>.
- [13] S. M. Park, H. G. Yoon ja C. Won, "Topological magnetic structure generation using VAE-GAN hybrid model and discriminator-driven latent sampling",

- Scientific Reports*, vol. 13, nro 1, 2023. DOI: 10.1038/s41598-023-45678-9.
url: <https://www.nature.com/articles/s41598-023-45678-9>.
- [14] Q. F. Li, H. X. Fang ja G. M. Ren, "Large-Scale Completion of Ionospheric TEC Maps Using Machine Learning Models With Constraint Conditions", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, s. 17092–17104, 2025. DOI: 10.1109/JSTARS.2025.1234567.
url: <https://ieeexplore.ieee.org/document/1234567>.
- [15] Y. Liu, H. Fan, X. Yuan ja J. Xiang, "GL-GAN: Adaptive global and local bilevel optimization for generative adversarial network", *Pattern Recognition*, vol. 114, s. 108375, 2021. DOI: 10.1016/j.patcog.2021.108375. url: <https://doi.org/10.1016/j.patcog.2021.108375>.
- [16] D. M. Alghazzawi, S. H. Hasan ja S. Bhatia, "Optimized Generative Adversarial Networks for Adversarial Sample Generation", *Computers, Materials & Continua*, vol. 72, nro 2, s. 3877–3897, 2022, Received: 24 October 2021; Accepted: 15 February 2022; Issue published: 29 March 2022. DOI: 10.32604/cmc.2022.024613. url: <https://doi.org/10.32604/cmc.2022.024613>.
- [17] S. Wang, C. Zhao, L. Huang, Y. Li ja R. Li, "Current status, application, and challenges of the interpretability of generative adversarial network models", *Concurrency and Computation: Practice and Experience*, 2022, First published: 12 December 2022. DOI: 10.1002/coin.12564. url: <https://doi-org.ezproxy.utu.fi:2443/10.1111/coin.12564>.
- [18] S. Cheng, L. Wang, M. Zhang, C. Zeng ja Y. Meng, "SUGAN: A Stable U-Net Based Generative Adversarial Network", *Sensors*, vol. 23, nro 17, s. 7338, 2023, Submission received: 26 July 2023 / Revised: 16 August 2023 / Accepted: 18 August 2023 / Published: 23 August 2023. DOI: 10.3390/s23177338. url: <https://www.mdpi.com/1424-8220/23/17/7338>.

- [19] D. P. Kingma ja M. Welling, "Auto-Encoding Variational Bayes", *arXiv preprint arXiv:1312.6114*, 2013, Submitted on 20 Dec 2013 (v1), last revised 10 Dec 2022 (v11). DOI: 10.48550/arXiv.1312.6114. url: <https://arxiv.org/abs/1312.6114>.
- [20] P. Cristovão, H. Nakada, Y. Tanimura ja H. Asoh, "Generating In-Between Images Through Learned Latent Space Representation Using Variational Autoencoders", *IEEE Access*, vol. 8, s. 149 456–149 467, 2020. DOI: 10.1109/ACCESS.2020.3016313.
- [21] C. Guo, J. Zhou, H. Chen, N. Ying, J. Zhang ja D. Zhou, "Variational Autoencoder With Optimizing Gaussian Mixture Model Priors", *IEEE Access*, vol. 8, s. 43 992–44 005, 2020. DOI: 10.1109/ACCESS.2020.2977671.
- [22] L. Xu, Z. Yang, D. Zhao, F. Yu, Y. Zhou ja H. Zhang, "G-VAE: Variational autoencoder-based adversarial attacks and defenses in industrial control systems", *Computers & Electrical Engineering*, 2025, Received: 6 September 2024; Revised: 5 March 2025; Accepted: 18 March 2025; Available online: 7 April 2025; Version of Record: 7 April 2025. DOI: 10.1016/j.compeleceng.2025.108680. url: <https://www.sciencedirect.com/science/article/pii/S0045790625000581>.
- [23] W. Xu ja P. Zhang, "Steam turbine anomaly detection: an unsupervised learning approach using enhanced long short-term memory variational autoencoder", *Applied Thermal Engineering*, vol. 278, s. 127 138, marraskuu 2025. DOI: 10.1016/j.applthermaleng.2025.127138.
- [24] B. Ghosh, M. Motagh, M. A. Anvari ja S. Maghsudi, "Improving Atmospheric Noise Correction from InSAR Time Series Using Variational Autoencoder with Clustering (VAE-Clustering) Method", *Remote Sensing*, vol. 17, nro 18, s. 3189, 2025, Submission received: 5 June 2025; Revised: 16 August 2025; Accepted:

- 22 August 2025; Published: 15 September 2025. DOI: 10.3390/rs17183189.
url: <https://doi.org/10.3390/rs17183189>.
- [25] D. Atif, "VAE-INN: Variational Autoencoder with Integrated Neural Network Classifier for Imbalanced Credit Scoring, Utilizing Weighted Loss for Improved Accuracy", *Computational Economics*, 2025, Received: 4 August 2024 / Accepted: 18 August 2025. DOI: 10.1007/s10614-025-11094-w. url: <https://link-springer-com.ezproxy.utu.fi:2443/article/10.1007/s10614-025-11094-w>.
- [26] R. Zemouri, "Semi-Supervised Adversarial Variational Autoencoder", *Machine Learning and Knowledge Extraction*, vol. 2, nro 3, s. 361–378, 2020, Published: 6 September 2020. DOI: 10.3390/make2030020. url: <https://doi.org/10.3390/make2030020>.
- [27] S. Sadok, S. Leglaive, L. Girin, X. Alameda-Pineda ja R. Ségurier, "A multimodal dynamical variational autoencoder for audiovisual speech representation learning", *Neural Networks*, vol. 172, s. 106–120, huhtikuu 2024, Open access under Creative Commons license. DOI: 10.1016/j.neunet.2024.106120.
- [28] V. D. Ma, M. D. Pham, T. P. Tran, D. T. Truong, H. Vu ja H. T. Nguyen, "Detecting Emerging DGA Malware in Federated Environments via Variational Autoencoder-Based Clustering and Resource-Aware Client Selection", *Future Internet*, vol. 17, nro 7, s. 299, 2025, This article belongs to the Special Issue: Security of Computer System and Network. DOI: 10.3390/fi17070299. url: <https://doi.org/10.3390/fi17070299>.
- [29] J. C. Tu et al., "Deep learning-based embedding of functional connectivity profiles for precision functional mapping", *Imaging Neuroscience*, vol. 3, IMAG.a.129, syyskuu 2025, Open Access. DOI: 10.1162/IMAG.a.129.

- [30] K.-L. Hsieh et al., "iGTP: learning interpretable cellular embedding for inferring biological mechanisms underlying single-cell transcriptomics", *Briefings in Bioinformatics*, vol. 26, nro 3, bbaf296, 2025. DOI: 10.1093/bib/bbaf296. url: <https://doi.org/10.1093/bib/bbaf296>.
- [31] H. Wang, Z. Qin ja T. Wan, "Text Generation Based on Generative Adversarial Nets with Latent Variables", teoksessa *Advances in Knowledge Discovery and Data Mining, PAKDD 2018, Part II*, D. Phung, V. S. Tseng, G. I. Webb, B. Ho, M. Ganji ja L. Rashidi, toim., sarja Lecture Notes in Artificial Intelligence, Indexed 2018-09-12, vol. 10938, Springer, 2018, s. 92–103. DOI: 10.1007/978-3-319-93037-4_8.
- [32] Y. Yao et al., "Conditional Variational Autoencoder with Balanced Pre-training for Generative Adversarial Networks", teoksessa *Proceedings of the 2022 IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA)*, J. Z. Huang et al., toim., Indexed 2023-05-06, IEEE, 2022, s. 156–165. DOI: 10.1109/DSAA54385.2022.10032367.
- [33] S. M. Park, H. G. Yoon ja C. Won, "Topological magnetic structure generation using VAE–GAN hybrid model and discriminator-driven latent sampling", *Scientific Reports*, vol. 13, nro 1, 2023, Demonstrates a VAE–GAN hybrid model for generating 2D magnetic topological structures, combining GAN's realism and VAE's diversity. DOI: 10.1038/s41598-023-46678-z. url: <https://doi.org/10.1038/s41598-023-46678-z>.
- [34] X. Wu ja H. Jang, "Multi-adversarial autoencoders: Stable, faster and self-adaptive representation learning", *Expert Systems with Applications*, vol. 262, s. 125554, 2025. DOI: 10.1016/j.eswa.2024.125554. url: <https://doi.org/10.1016/j.eswa.2024.125554>.

- [35] Q. Li, H. Fang, C. Xiao, D. Duan, H. Huang ja G. Ren, "Large-Scale Completion of Ionospheric TEC Maps Using Machine Learning Models With Constraints Conditions", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, s. 17 092–17 104, heinäkuu 2025, Open Access, under a Creative Commons License (CC BY 4.0). DOI: 10.1109/JSTARS.2025.3585566.
- [36] F. Zhang, Q. Teng, H. Chen, X. He ja X. Dong, "Slice-to-voxel stochastic reconstructions on porous media with hybrid deep generative model", *Computational Materials Science*, vol. 183, s. 110 018, 2020, Received 6 May 2020, Revised 25 July 2020, Accepted 23 August 2020, Available online 12 September 2020, Version of Record 12 September 2020. DOI: 10.1016/j.commatsci.2020.110018. url: <https://doi-org.ezproxy.utu.fi:2443/10.1016/j.commatsci.2020.110018>.
- [37] C. Luo et al., "EvaGoNet: An integrated network of variational autoencoder and Wasserstein generative adversarial network with gradient penalty for binary classification tasks", *Information Sciences*, vol. 629, s. 109–122, kesäkuu 2023, Early Access February 2023, Indexed March 10, 2023. DOI: 10.1016/j.ins.2023.01.133.
- [38] M. Proszewska, M. Wołczyk, M. Zieba, P. Wielopolski, Ł. Maziarka ja M. Śmieja, "Multi-Label Conditional Generation From Pre-Trained Models", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, nro 9, s. 6185–6198, maaliskuu 2024. DOI: 10.1109/TPAMI.2024.3382008.
- [39] F. Ye ja A. G. Bors, "Lifelong Generative Adversarial Autoencoder", *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, nro 10, s. 14 684–14 698, lokakuu 2024, ISSN: 2162-237X. DOI: 10.1109/TNNLS.2023.3281091. url: <https://doi.org/10.1109/TNNLS.2023.3281091>.

-
- [40] T. Chakraborty, U. R. K S, S. M. Naik, M. Panja ja B. Manvitha, "Ten years of generative adversarial nets (GANs): a survey of the state-of-the-art", *Machine Learning: Science and Technology*, vol. 5, nro 1, s. 011 001, 2024, Open access; Focus on Generative AI in Science. DOI: 10.1088/2632-2153/ad1f77.