

Rakstzīmju stila pārvešana ar roku rakstītiem burtiem, izmantojot dziļās mašīnmācīšanās metodes

Maģistra darbs

Autors: Stefans Dayneko

Zinātniskais vadītājs: Agris Nikitenko

Zinātniskais konsultants: Ēvalds Urtāns

19.06.2021

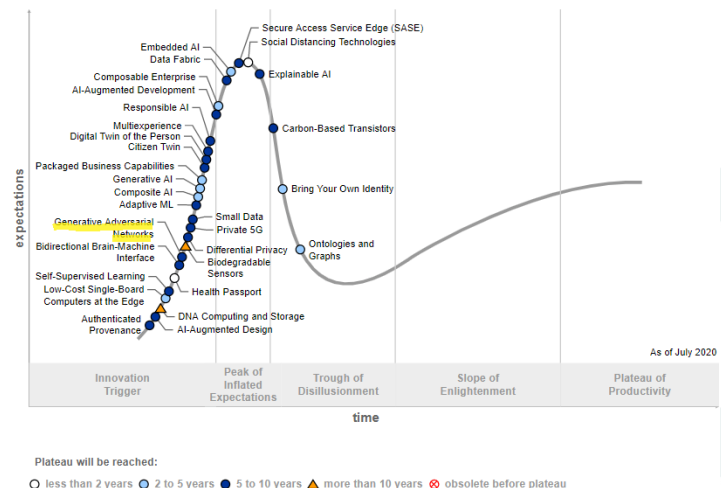


RTU

**DATORZINĀTNES UN
INFORMĀCIJAS
TEHNOLOĢIJAS FAKULTĀTE**

Tēmas aktualitāte

- Saskaņā ar Gartnera pētījumiem, 2020. gadā, GAN tipa tīkli atrodas Inovācijas posmā - posms, kurā GAN tehnoloģija izlaužas.
- Balstoties uz Gartner datiem:
 - GAN tipa tīkli pašlaik nav dziļi pētīti, un joprojām skaitās kā «melnā kaste», kur nav pieejamu plašu novērtēšanas kritēriju, izņemot subjektīvo novērtējumu.
 - Izmantojot GAN tīklus, ir iespējams rādīt simulācijās, kur jaunus datus ir grūti iegūt (datu kopas mašīnmācīšanai), vai rada privātuma risku (medicinās attēli), vai ir dārgi ražošanā (fona attēli spēlēm). Esošais darbs rada pienesumu, ļaujot cilvēkam anonimizēt savu rokrakstu.



Darba mērķis un uzdevumi

- Darba mērķis ir izstrādāt dziļās mašīnmācīšanās modeli, balstoties uz dziļo konvolūciju neironu tīkla arhitektūru, kas spēj pārveidot rokraksta burtus par burtiem rakstzīmju stilos, izmantojot stila pārnešanas metodi.
- Mērķa sasniegšanai tika izvirzīti sekojošie uzdevumi:
 1. Izpētīt Konvolūciju neironu tīklu darbības principus un to veidus,
 2. Izpētīt Ģeneratīvos sacīkšu tīklus (GAN), tās kļūdas funkcijas un novērtēšanas metrikas,
 3. Izpētīt populārākus mašīnmācīšanās algoritmus un to darbības principus,
 4. Definēt mērķa risinājuma arhitektūru,
 5. Izstrādāt AE-DCGAN dziļo neironu tīklu, kas spētu pārnest rakstzīmju stilu ar roku rakstītiem burtiem,
 6. Pārmeklēt izstrādātā dziļā neironu tīkla hiperparametrus,
 7. Novērtēt iegūtos rezultātus izmantojot attēla un iezīmju novērtēšanas metrikās.

Darba hipotēze

- Izmantojot stila pārneses GAN ar semantisko īpašību filtriem var panākt rakstītu zīmju pārnesi uz kaligrāfiskiem fontiem precīzāk kā ar stila pārnesi, bez semantisko īpašību filtriem.

DCGAN (Radford, Metz et al., 2016)

- Dziļie Konvolūciju Ģeneratīvie Sacīkšu Tīkli - viena no konvolūciju tīklu arhitektūram, kura stabilizēja GAN tīklu apmācības procesu un padarīja to par izmantojamu attēlu raksturiezīmju nepārraudzītā apmācībā.
- Pamatā ir četras izmaiņas konvolūciju tīklu arhitektūrās:
 1. Apvienošanās slāņu aizstāšana ar konvolūciju slāņiem,
 2. Pilnībā savienotu slāņu aizstāšana ar konvolūciju slāņiem,
 3. Partīciju normalizēšanas ieviešana gan ģeneratorā, gan diskriminātorā,
 4. ReLU un Leaky ReLU aktivizācijas funkcijas ģeneratorā un diskriminātorā.
- Atšķirībā no oriģināla DCGAN, šajā darbā tika ieviestas vēl papildus izmaiņas:
 - Leaky ReLU ir pielietots arī ģeneratorā,
 - Klašu varbūtību nolīdzināšanas metode,
 - Klašu apgriešanas metode,
 - Pieredzes atmiņas metode.

«Baltās Kastēs» Multifimu GAN (Wang & Yu, 2020)

- Pētījums apraksta stila pārņemšanas problēmu, kur seja vai ainava ir attēlotā multfilmas stilā.
- Tiek pielietotas «baltās kastēs» metodes:
 - Virsmas attēlojuma filtrs,
 - Struktūras attēlojuma filtrs,
 - Satura attēlojuma filtrs.
- Izmantots iepriekš apmācīts VGG konvolūciju neironu tīkls, lai iegūtu semantiskas attēla iezīmes un uzlikt telpiskos ierobežojumus starp struktūru, ievaddatiem un izvaddatiem.
- Atšķirībā no BKM GAN, šajā darbā tika ieviestas papildus izmaiņas:
 - Neironu tīkls tika apmācīts, izmantojot kosinusa distances kļūdas funkciju,
 - Iezīmju izgūšanas procesā tika pievienoti jauni «baltās kastēs» filtri,
 - Ģeneratora pamata struktūru nodrošina Auto-Kodētājs,
 - SSIM kļūdas funkcijā kopējās dispersijas (TV) vietā,
 - RAdam optimizācijas algoritms.

Jaunais modelis – AE-DCGAN

- Praktiskā darba daļa ir sadalītā 2 eksperimentos:
 1. Eksperiments – pārbaudīt Auto-Kodētāju, salīdzinot ievaddatus un izvaddatus, novērtējot attēla saturu.
 2. Eksperiments – pārbaudīt Auto-Kodētāju kopā ar GAN tīklu, pielietojot rakstzīmju stila attēlu, salīdzinot ievaddatus un izvaddatus, novērtējot attēla stilu un saturu.
- Secinājums pēc 1. Eksperimenta:
 - Izmantojot semantiskās iezīmes no VGG tīkla, tīkla apmācīšanas ātrums krietni uzlabojās, salīdzinot ar VAE modeli.
- Secinājums pēc 2. Eksperimenta:
 - Izmantojot semantisko īpašību filtrus, attēla kvalitāte ir augstāka nekā VAE modelim, kas tos neizmanto.

Jaunievedums – Semantisko īpašību filtri

Rakstzīmju krāsains piemērs

8	5	2	3	6	7	6	0
6	5	8	9	7	1	0	4
4	7	5	6	4	8	0	7
1	1	9	5	0	6	8	9
7	8	2	9	1	2	0	5
3	1	8	4	7	6	2	4
6	9	1	2	8	4	2	4
3	5	8	4	7	2	1	8
3	2	6	3	0	4	2	3
0	3	4	6	9	3	7	4
8	9	1	0	5	3	2	5
0	8	5	9	2	4	7	1
5	1	3	3	3	8	9	2
0	2	1	3	1	7	0	6
6	9	6	0	7	5	7	0
6	1	5	7	5	4	9	9

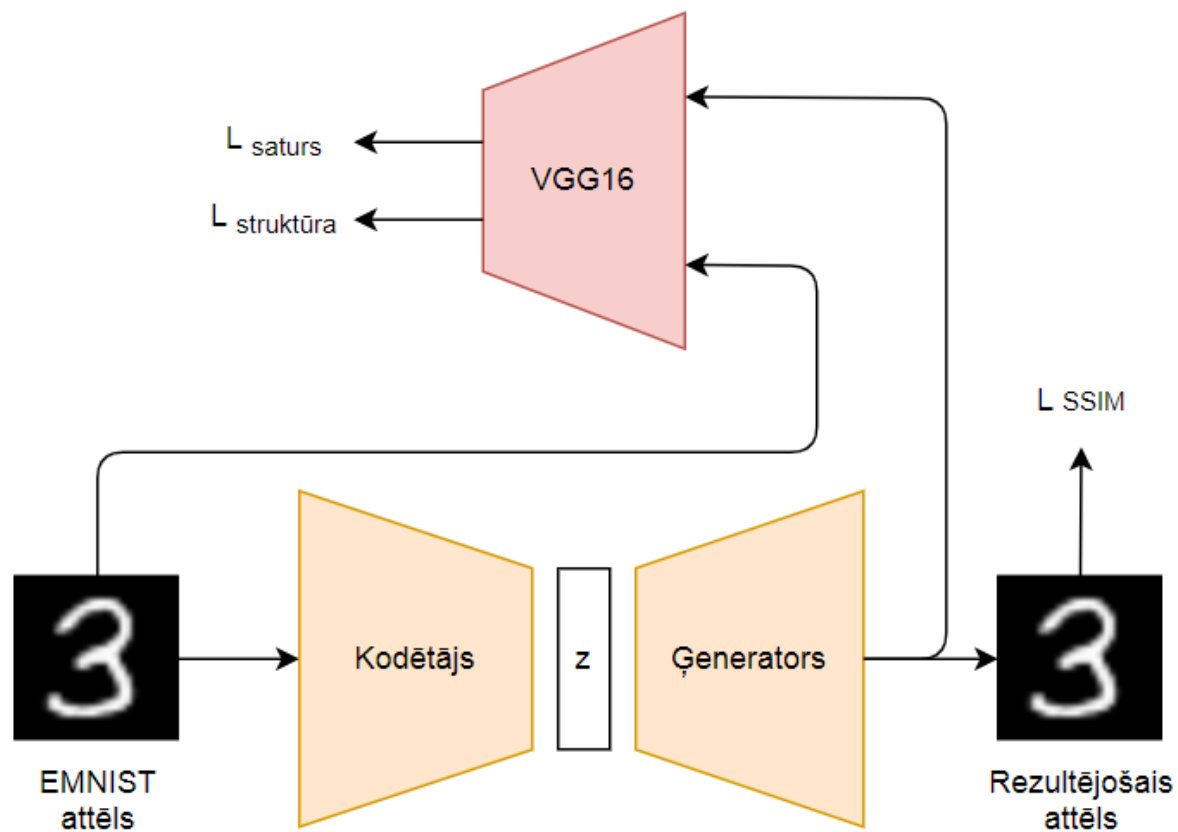
Struktūras filtrs

8	5	2	3	6	7	6	0
6	5	8	9	7	1	0	4
4	7	5	6	4	8	0	7
1	1	9	5	0	6	8	9
7	8	2	9	1	2	0	5
3	1	8	4	7	6	2	4
6	9	1	2	8	4	2	4
3	5	8	4	7	2	1	8
3	2	6	3	0	4	2	3
0	3	4	6	9	3	7	4
8	9	1	0	5	3	2	5
0	8	5	9	2	4	7	1
5	1	3	3	3	8	9	2
0	2	1	3	1	7	0	6
6	9	6	0	7	5	7	0
6	1	5	7	5	4	9	9

Satura attēli

8	5	2	3	6	7	6	0
6	5	8	9	7	1	0	4
4	7	5	6	4	8	0	7
1	1	9	5	0	6	8	9
7	8	2	9	1	2	0	5
3	1	8	4	7	6	2	4
6	9	1	2	8	4	2	4
3	5	8	4	7	2	1	8
3	2	6	3	0	4	2	3
0	3	4	6	9	3	7	4
8	9	1	0	5	3	2	5
0	8	5	9	2	4	7	1
5	1	3	3	3	8	9	2
0	2	1	3	1	7	0	6
6	9	6	0	7	5	7	0
6	1	5	7	5	4	9	9

1. eksperiments



1. eksperimenta rezultāti

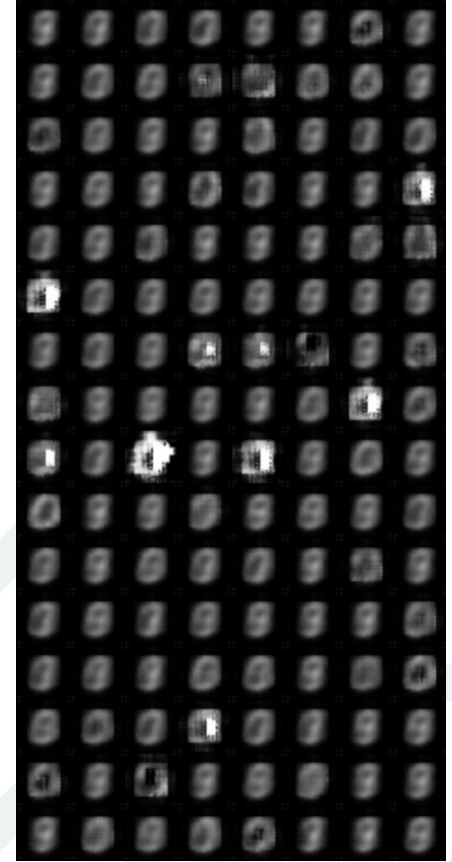
EMNIST datu kopa



AE attēli



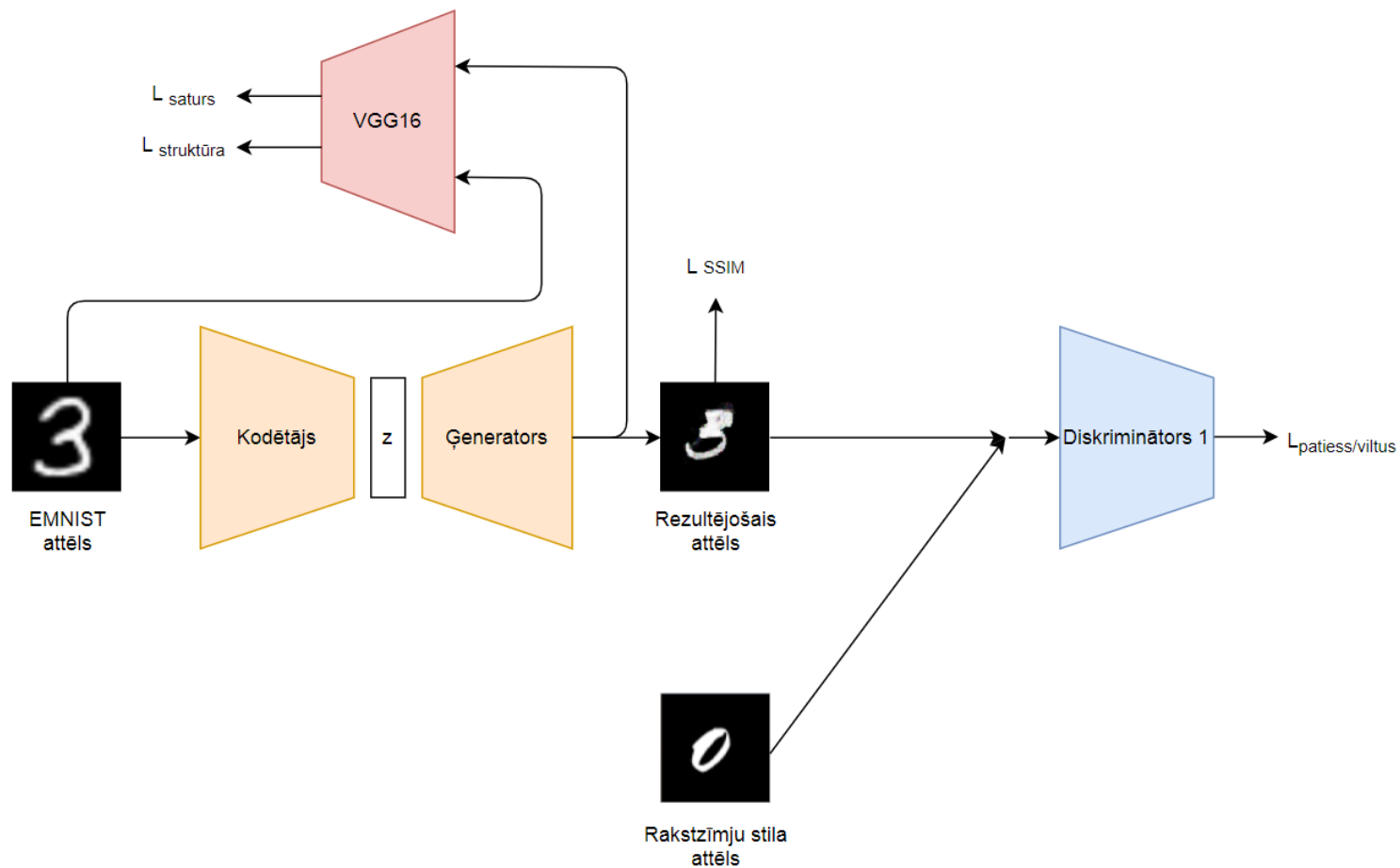
VAE attēli



1. eksperimenta rezultāti

Metrika	Stila attēls / stila attēls	Stila attēls / viltotais attēls	Stila attēls / VAE attēls
Strukturāla līdzība (SSIM)	1	0.9170	0.2281
Iezīmju līdzība (FSIM)	1	0.9305	0.5895
Kopēja dispersija (TV)	11.2897	10.8463	11.4717
Satura līdzība	0	880.9444	4746.3784
Frečeta attālums (FID)	0	12.7831	376.1996

2. eksperiments

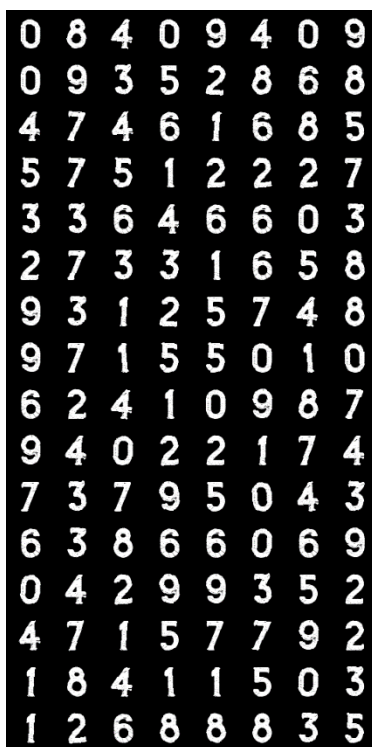


2. eksperimenta rezultāti

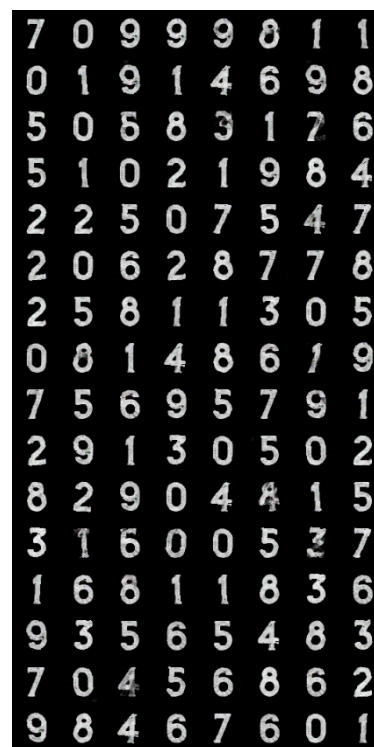
EMNIST datu kopa



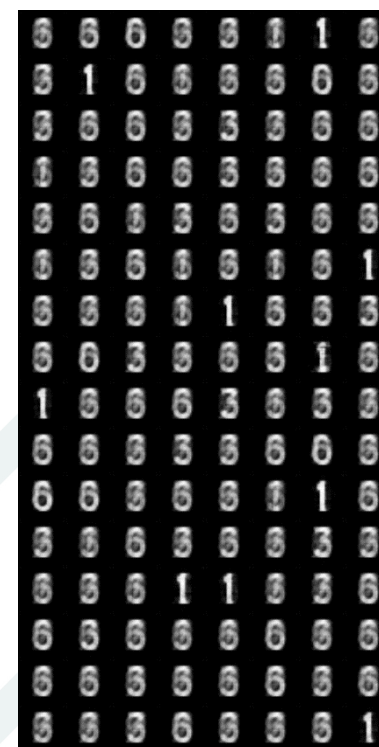
Mērķa stila fonts



AE-DCGAN attēli



VAE attēli



2. eksperimenta rezultāti

Metrika	Stila attēls / stila attēls	Stila attēls / viltotais attēls	Stila attēls / VAE attēls
Strukturāla līdzība (SSIM)	1	0.5901	0.6366
Iezīmju līdzība (FSIM)	1	0.7333	0.7756
Kopēja dispersija (TV)	19.8620	15.8670	11.4273
Satura līdzība	0	5815.6572	3135.9367
Frečeta attālums (FID)	0	48.7738	216.171

Rezultāti ar citiem rakstzīmju stiliem

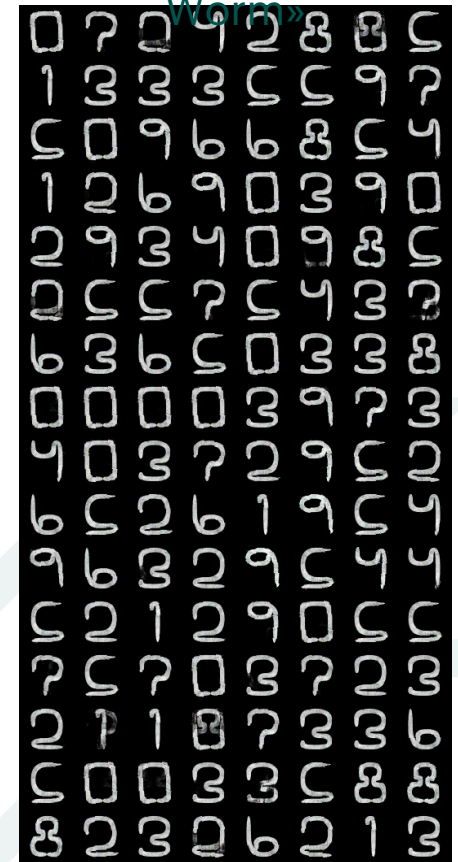
EMNIST datu kopa



Rakstzīmju stils



Rakstzīmju stils «Space



Secinājumi

1. Mērķis sasniegts, implementēts GAN, kurš spēj pārnest rakstzīmju stila uz ar roku rakstītiem burtiem.
2. Auto-Kodētāju apmācot ar kosinusa attāluma kļūdas funkciju, apmācās ātrāk nekā VAE.
3. Frečeta attālums var tikt izmantots kā novērtējuma metrika attēlu kvalitātei.
4. Var izmantot nepārraudzīto apmācību, lai apmācītu modeli saglabāt rakstzīmju semantisko informāciju.

Darba hipotēze pierādīta, ka izmantojot semantiskos īpašību filtrus var iegūt augstākas kvalitātes rezultātus.

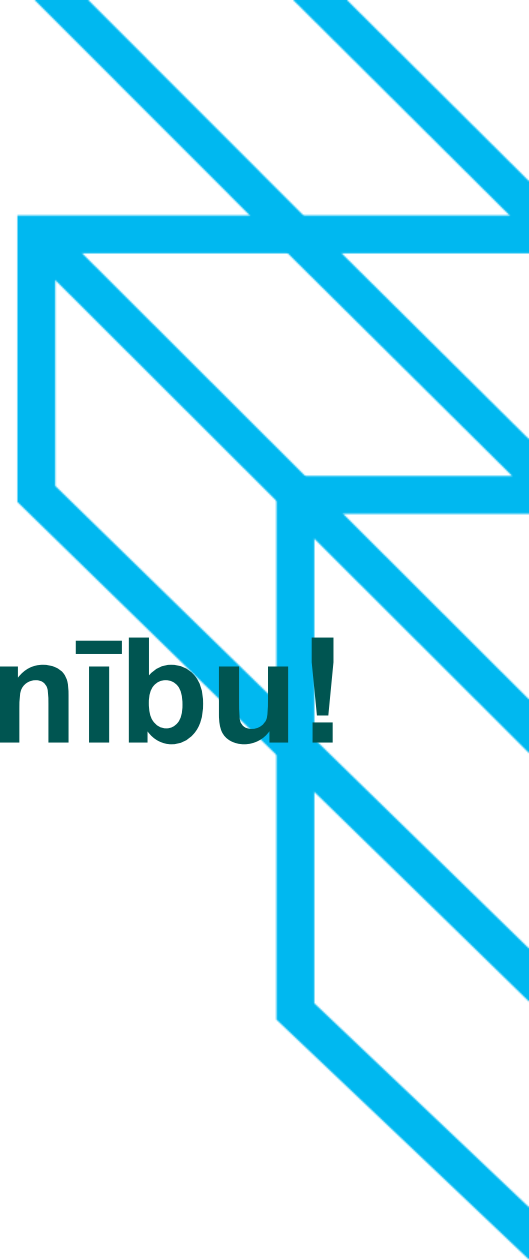
Turpmākie pētījumi

1. Tikla apmācīšanai pielietot daļēji uzraudzīto apmācīšanu.
2. Izmēģināt citas kodētāja un ģeneratora arhitektūras.
3. Pievienot papildus filtrus, kuri izdala papildus attēla iezīmes, piem.: krāsu histogrammas, utt.
4. Izmēģināt citus attāluma kļūdas funkcijas, piem.: Heminga, Levenšteina, Manhetenā, utt.
5. Pielietot doto modeli citai datu kopai, piem.: CelebA, kur ir apkopotas cilvēku sējas.

Pateicības

- Agris Nikitenko
- Ēvalds Urtāns
- RTU akadēmiskais personāls

Paldies par uzmanību!



J&A[1]

- Darbā ir ļoti daudz gramatikas, interpunkcijas un stila kļūdu. Pieļautas neuzmanības kļūdas teorētiskajā daļā. Piemēram, tabulā 1.3 norādīts, ka RAdam nav nepieciešama iesildīšanas fāze, bet tāda nepieciešama Adam. Adam metodes nosaukums kļūdaini iztulkots kā 'Ādams'. MNIST datu kopas nosaukums vietām kļūdaini dots kā EMNIST.

1.3. tabula

Optimizācijas algoritmu salīdzinājums

Tips	SGD	Adam	RAdam
Parādīšanās gads	1960	2014	2020
Konverģences ātrums	Lēns	Ātrs	Ātrs
Pamatā	Gradienta nolaišanās	SGD AdaGrad RMSProp	Adam
Uzlabojumu pamats	Uzlabo klasisko gradienta nolaišanās metodi, mainot iesaistīto datu daudzumu	Uzlabo SGD algoritmu, pielietojot adaptīvo mācīšanās ātrumu, novērtējot pirmās un otrās kārtas gradientus.	Uzlabo Adam, ieviešot iesildīšanas fāzi optimizātorā.
Mācīšanas ātrums	Viens	Adaptīvs	Adaptīvs
Iesildīšanas fāze	Nav nepieciešama	Ir nepieciešama	Nav nepieciešama

J&A[2]

- Vietām netiek doti apzīmējumu skaidrojumi. Piemēram, tabulā 1.4 nav dots apzīmējumu skaidrojums hiperparametru un apmācāmo parametru apzīmējumiem. Tapāt netiek skaidrota saīsinājumā SSIM nozīmē un tas lietots pirms dots tā apraksts (lietots nodaļā 5.3, bet aprakstīts nodaļa 5.5).

1.4. tabula

Normalizēšanas metožu salīdzinājums

Tips	Sērijas normalizēšana	Grupas normalizēšana
Parādīšanās gads	2015	2018
Hiperparametri	-	G
Apmācāmie parametri	γ β	γ β

J&A[3]

- Tekstā netiek dotas atsauces uz izmantotajiem attēliem un formulām.

J&A[4]

- Ir dots shematisks uzmetums izmantotajai arhitektūrai, bet netiek dots pilns parametru apskats izmantotajiem tikliem. Būtu nepieciešams uzskatāmi norādīt kārtu skaitu, to lielumu, filtrus un savienojumus līdzīgi kā pielikuma dotajām arhitektūrām.

J&A[5]

- Lai gan saprotu, ka skaitļošanas rezultāti ir ierobežoti, rezultāti būtu jādod ar vairāk kā vienu stila veidu. Grūti objektīvi novērtēt darbā veikto uzdevumu balstoties uz tikai vienu piemēru.

J&A[6]

- Nevajadzētu tekstā tulkot rakstu nosaukumus. Pietiek ar atsaucēs ievietošanu.