

MĀKOŅU KOGNITĪVO PAKALPOJUMU SNIEDZĒJU SALĪDZINOŠĀ ANALĪZE, IZMANTOJOT MAŠĪNREDZES OBJEKTU ATPAZĪŠANU COMPARATIVE ANALYSIS OF CLOUD COGNITIVE SERVICE PROVIDERS BY USING COMPUTER VISION'S OBJECT DETECTION

Autori: **Jāzeps IVULIS**, e-pasts: jazeps.ivulis@gmail.com,

Kārlis ŠTEKELS, e-pasts: k.stekels@gmail.com

Zinātniskā darba vadītājs: **Artis TEILĀNS**, Dr.sc.ing. prof., e-pasts: Artis.Teilans@rta.lv

Rēzeknes Tehnoloģiju akadēmija,
Rēzekne, Atbrīvošanas aleja 115

Abstract. *The given paper's "Comparative analysis of cloud cognitive service providers by using computer vision's object detection" goal is to analyze prediction precision of objects in digital image, using different cognitive services what are available in the market. During the research process there were made a several trials using different computer vision services to analyze the same number of images. Every separate image was under detailed analysis for prediction accuracy of each object. The results show that there is significant difference between different vision services.*

Keywords: *cognitive services, computer vision, object detection.*

Ievads

Ar katru gadu tehnoloģijas attīstās arvien straujāk un pieaug pieprasījums pēc mākslīgā intelekta šajās tehnoloģijās. Mūsdienų pasaulē ir daudz problēmu, kuras varētu risināt ar mākslīgā intelekta palīdzību, lai atvieglotu un uzlabotu mūsu ikdienas dzīvi. Mākslīgā intelekta potenciāls ir neierobežots un to var pielietot jebkurā darbības jomā. [1]

Viena no visizplatītākajām tehnoloģijā, kas saistīta ar mākslīgo intelektu ir attēlā redzamo objektu atpazīšana. Objektu atpazīšana ir tehnoloģija, kura saistīta ar mašīnredzi un attēlu apstrādi, kas nodarbojas ar noteiktu, semantisko objektu klašu gadījumu atklāšanu digitālajos attēlos un video. Šai tehnoloģijai nav ierobežojumu, un to var pielietot gan vienkāršās sistēmās, gan sarežģītajās sistēmās ar savu specifiku. Mūsdienų industrijā lielu lomu spēlē lēmumu pieņemšanas ātrums un precizitāte, kas cilvēkam nav pa spēkam, bet skaitļošanas mašīna gan to var nodrošināt efektīvi, tāpēc tagad aktuāli ir automatizēt procesus ar mākslīgā intelekta palīdzību, kas spēs aizstāt cilvēku un veikt darbu daudz ātrāk un ar lielāku precizitāti. [2][3]

Pēdējo gadu laikā, lielākās kompānijas, tādas kā "Amazon", "Google", "Microsoft" un "Facebook" investē miljardus dolāru mašīnredzes produkta izpētē un izstrādē. [4] Mašīnredzei pielietojumu var atrast, piemēram, automašīnbūvē, rūpniecībā, medicīnā, lauksaimniecībā, banku sektorā, drošības sistēmas, gudrās pilsētās u.c. Ar vien lielāku popularitāti gūst autonomie auto, kuri izmanto mašīnredzes tehnoloģiju. Ar šo tehnoloģiju, autonomie transporta līdzekļi spēj redzēt apkārtējo vidi un orientēties tajā, droši pārvietojoties no viena punkta, uz citu, bez jebkādas cilvēka iejaukšanās. Medicīnā mašīnredzi var izmantot lai spētu ātrāk un precīzāk noteikt pacientam diagnozi. Rūpniecībā ar mašīnredzi var uzraudzīt saražotās produkcijas kvalitāti. [5][6] Šādus piemērus, kur var izmantot mašīnredzi, ir ļoti daudz, bet, šī tehnoloģija vēl ir tikai attīstības līmenī, tāpēc ir svarīgi analizēt iegūtos datus, pirms tos izmatot reālajā dzīvē uz reāliem objektiem. Dotajā brīdī ir daudz servisu, kas piedāvā savus kognitīvos pakalpojumus ar mašīnredzi, visi viņi strādā pietiekami augstā līmenī, bet servisu mašīnredzes tehnoloģijas precizitāte būtiski atšķiras, tāpēc ir svarīgi zināt, kurš kognitīvais serviss ir vispiemērotākais veicamajam uzdevumam.

Materiāli un metodes

Galvenais darba mērķis:

- salīdzināt vadošos objektu atpazīšanas kognitīvos servisos un noteikt servisu, kurš skaitliski spēj atpazīt visvairāk objektu ar visaugstāko uzticamības līmeni. Lai sasniegtu darba mērķi, darba autori noteica vairākus darba uzdevumus:
- izvēlēties 10 attēlus, kuros ir attēloti autobusi, automašīnas, cilvēki, motocikli, velosipēdi;
- izvēlēties 4 vadošos kognitīvo servisu sniedzējus, kuri piedāvā objektu atpazīšanas servisos;
- izvēlētajos 10 attēlos, izmantojot 4 objektu atpazīšanas kognitīvos servisos, veikt objektu atpazīšanas eksperimentus pie 3 dažādiem minimālajiem uzticamības līmeņiem.

Lai izanalizētu mašīnredzes objektu atpazīšanas tehnoloģijas precizitāti, autori izvēlējās četrus vadošos kognitīvos servisu provaiderus: "Google Cloud Vision", "Microsoft Azure Computer Vision", "IBM Watson Visual Recognition" un "Amazon Rekognition". Ar šo servisu palīdzību tika veikta objektu atpazīšana 10 attēliem, kuros bija attēloti autobusi, automašīnas, cilvēki, motocikli un velosipēdi.

Visi kognitīvie servisi izmantoto padziļināto mašīnas apmācību, kas spēj attēlā atpazīt dažāda veida objektus, ainas, sejas, kā arī, nolasīt tekstus, atpazīt slavenības un noteikt neatbilstošu saturu attēlos. Visi kognitīvie servisi ir balstīti uz jau pierādītas un pielāgojamas padziļinātās mašīnmācības tehnoloģijas. [2][3][7][8]

Lai to izdarītu, tika praktiski izmantots katrs no četriem kognitīvajiem servisiem. Lai būtu iespējams tos izmantot, katram kognitīvajam servisam bija nepieciešams izveidot bezmaksas kontus. Katrs no četriem kognitīvo servisu sniedzējiem uz bezmaksas kontiem piedāvā neierobežotu attēlu apstrādes funkciju (teksta, sejas, objekta atpazīšana u.c.) izmantošanu līdz 1000 attēliem mēnesī. Tas pilnībā apmierina nepieciešamo attēlu apstrādes skaitu šī darba ietvaros nekādīgi neietekmēs eksperimentu rezultātu objektivitāti. Pašu kognitīvo servisu mašīnmācības modeļu apmācība netika veikta, tika izmantoti gatavi risinājumi – iepriekš apmācīti modeļi, kurus jau piedāvā katrs objektu atpazīšanas servisa sniedzējs.

Rezultāti un to izvērtējums

Visu eksperimentu rezultātu dati tika ievākti un apkopoti 3 tabulās. Šajā nodaļā tiks uzrādīti, izanalizēti un paskaidroti eksperimentu rezultātu datu tabulās.

Atpazīto objektu skaits pie minimālā uzticamības līmeņa 50% atkarībā no izmantotā objektu atpazīšanas servisa tiek parādīts 1. tabulā.

1. tabula

Atpazīto objektu skaits pie minimālā uzticamības līmeņa 50% atkarībā no izmantotā objektu atpazīšanas servisa

Objekts	Amazon Rekognition	Google Cloud Vision	IBM Watson Visual Recognition	Microsoft Azure Computer Vision
Autobuss	10	7	4	7
Automašīna	73	22	35	36
Cilvēks	82	47	36	21
Motocikls	11	9	6	8
Velosipēds	15	5	3	5
Kopā	191	90	84	77

Pēc 1. tabulas datiem ir skaidri redzams objektu atpazīšanas servisa “Amazon Rekognition” pārākums pār pārējiem objektu atpazīšanas servisiem, kuru darbība tika eksperimentāli pārbaudīta un salīdzināta šī darba ietvaros. Izanalizējot eksperimenta rezultātus, pie minimālā uzticamības līmeņa 50%, “Amazon Rekognition” serviss atpazīst vidēji par 127% vairāk objektu nekā pārējie servisi.

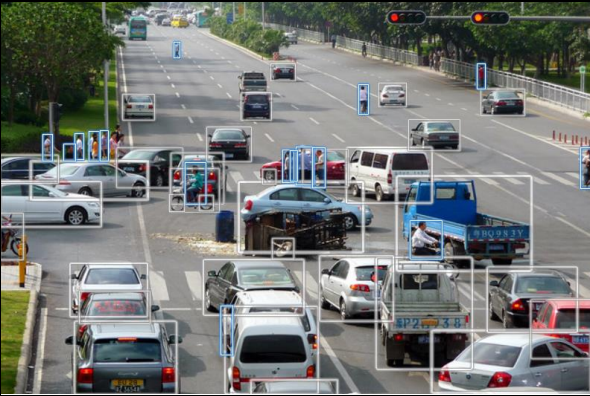
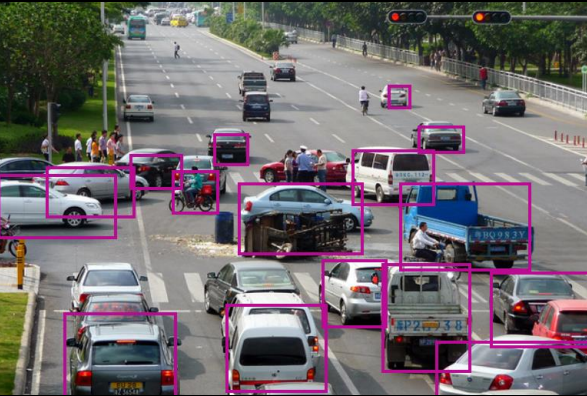
Atsevišķā gadījumā, salīdzinot “Amazon Rekognition” un “Microsoft Azure Computer Vision” servissus, pie objekta “Cilvēks”, “Amazon Rekognition” atpazīst pat par 290% vairāk objektu nekā konkurents.

Iepriekš konstatēto anomāliju ar servisa “Microsoft Azure Computer Vision” zemajiem, objekta “Cilvēks” atpazīšanas rādītājiem var izskaidrot ar 2 ierobežojumiem, kurus pats pakalpojuma sniedzējs “Amazon” ir uzrādījis un kurus ir jāņem vērā izmantojot šo servisu:

- objekti parasti netiek atpazīti, ja tie ir mazi (mazāk nekā 5% no attēla);
- objekti parasti netiek atpazīti, ja tie ir izvietoti cieši kopā (piemēram, šķīvju kaudze, cilvēku pūlis). [9]

Vizuālu piemēru, salīdzinot abu iepriekš minēto servisu attēla apstrādi priekš objektu atpazīšanas, var aplūkot 2. tabulā, kurā redzams, ka serviss “Amazon Rekognition” attēlā bez grūtībām atpazīst individuālus objektus “Cilvēks”. Turpretim serviss “Microsoft Azure Computer Vision”, tajā pašā attēlā neatpazīst nevienu objektu “Cilvēks.”

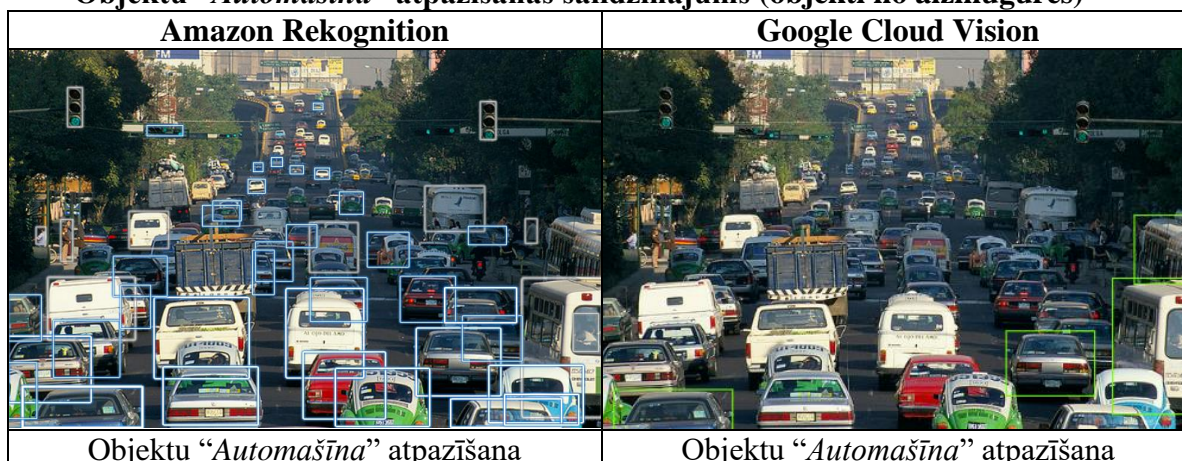
2. tabula

Objektu “Cilvēks” atpazīšanas salīdzinājums	
Amazon Rekognition	Microsoft Azure Computer Vision
	
Objekti “Cilvēks” tiek atpazīti	Objekti “Cilvēks” netiek atpazīti

Izanalizējot 1. tabulas datus, var konstatēt vēl vienu anomāliju – serviss “Google Cloud Vision” atpazīst ievērojami mazāk objektus “Automašīnā” nekā tā konkurenti. Serviss “Google Cloud Vision” atšķirībā no servisa “Microsoft Azure Computer Vision” publiski neatklāj attēlu apstrādes ierobežojumus, kuri var ietekmēt objektu atpazīšanas rezultātus, bet pēc vairākiem veiktajiem objektu atpazīšanas eksperimentiem ir iespējams noteikt potenciālos iemeslus.

Iepriekš minētā anomālija ir izveidojusies, jo servisam “Google Cloud Vision” bija grūtības atpazīt objektus “Automašīna” vienā no eksperimentā apstrādātajiem attēliem. Objektu “Automašīna” atpazīšanas vizuālu salīdzinājumu šajā attēlā starp servisiem “Amazon Rekognition” un “Google Cloud Vision” var aplūkot 3. tabulā.

Objektu “Automašīna” atpazīšanas salīdzinājums (objekti no aizmugures)

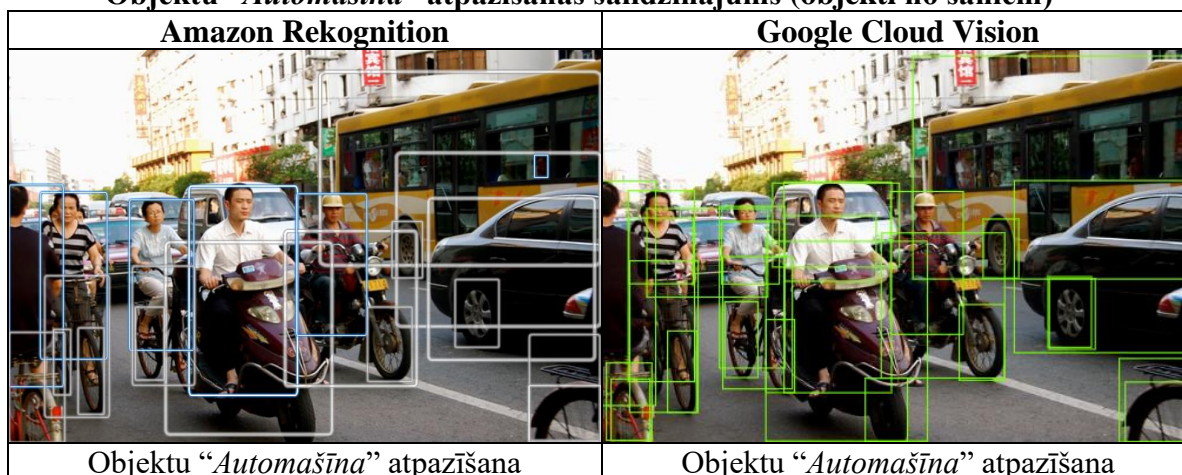


3. tabulā ir redzams, ka servisam “Google Cloud Vision” ir grūtības konkrētajā attēlā atpazīt objektus “Automašīna”. Salīdzinot konkrēto attēlu ar pārējiem eksperimentā izmantotajiem attēliem, var pieņemt, ka servisa “Google Cloud Vision” objektu atpazīšanas rezultātus varētu ietekmēt:

- objektu pozicionēšanās attēlā (dotajā attēlā objekti “Automašīna” ir ar “muguru” pret novērotāju);
- objektu novērošanas leņķis (dotajā attēlā leņķis ir mazs, tādā priekšā esošie objekti daļēji aizsedz aizmugurē esošos objektus);
- objektu pārblīvējums attēlā (dotajā attēlā redzams sauszemes transportlīdzekļu sastrēgums).

Lai apstiprinātu iepriekš izvirzīto teoriju, tiek apskatīts gadījums, kad serviss “Google Cloud Vision” attēlā atpazīst vairāk objektu “Automašīna” nekā serviss “Amazon Rekognition”. Šo konkrēto gadījumu var vizuāli aplūkot 4. tabulā.

Objektu “Automašīna” atpazīšanas salīdzinājums (objekti no sāniem)



4. tabulā ir redzams, ka serviss “Google Cloud Vision” atpazīst par 2 objektiem “Automašīna” vairāk nekā serviss “Amazon Rekognition”. Galvenā atšķirība starp 3. un 4. tabulas attēliem ir objektu pozicionēšanās – 4. tabulas attēlā objekti “Automašīna” ir pozicionēti ar priekšpusi vai sāniem pret novērotāju. Kaut gan objektu novērošanas leņķis ir

mazs un objektu pārblīvējums ir liels, iepriekš izvirzītā hipotēze par objektu pozicionēšanos attēlā ietekmi uz objektu atpazīšanas rezultātiem tika apstiprināta.

Atpazīto objektu skaitu pie minimālā uzticamības līmeņa 70% atkarībā no izmantotā objektu atpazīšanas servisa var aplūkot 5. tabulā.

5. tabula

**Atpazīto objektu skaits pie minimālā uzticamības līmeņa 70%
atkarībā no izmantotā objektu atpazīšanas servisa**

Objekts	Amazon Rekognition	Google Cloud Vision	IBM Watson Visual Recognition	Microsoft Azure Computer Vision
Autobuss	7	4	3	4
Automašīna	65	19	16	14
Cilvēks	75	42	19	13
Motocikls	8	7	5	5
Velosipēds	11	4	2	0
Kopā	166	76	45	36

Principā eksperimenta rezultāti pie minimālā uzticamības līmeņa 70% ir ļoti līdzīgi eksperimenta rezultātiem pie minimālā uzticamības līmeņa 50% - joprojām neapstrīdams līderis atpazīto objektu skaita ziņā ir “Amazon Rekognition”. Izanalizējot šī eksperimenta rezultātu datus, pie minimālā uzticamības līmeņa 70% serviss “Amazon Rekognition” atpazīst vidēji par 219% vairāk objektu nekā tā konkurenti.

Pie minimālā uzticamības līmeņa 70% dramatiski ir samazinājies atpazīto objektu skaits servisam “Microsoft Azure Computer Vision” – samazinājums par 114%.

Atpazīto objektu skaitu pie minimālā uzticamības līmeņa 90% atkarībā no izmantotā objektu atpazīšanas servisa var aplūkot 6. tabulā.

6. tabula

**Atpazīto objektu skaits pie minimālā uzticamības līmeņa 90%
atkarībā no izmantotā objektu atpazīšanas servisa**

Objekts	Amazon Rekognition	Google Cloud Vision	IBM Watson Visual Recognition	Microsoft Azure Computer Vision
Autobuss	5	2	1	0
Automašīna	39	7	1	0
Cilvēks	51	7	2	0
Motocikls	7	2	4	1
Velosipēds	6	0	0	0
Kopā	108	18	8	1

Izanalizējot pēdējā eksperimenta rezultātu datus, kurus var aplūkot 6. tabulā, var secināt, ka objektu atpazīšanas serviss ar visaugstāko atpazīto objektu varbūtību ir “Amazon Rekognition”.

Ļoti dramatisku atpazīto objektu skaita samazinājumu pie minimālā uzticamības līmeņa 90% ir piedzīvojuši visi trīs servisa “Amazon Rekognition” konkurenti. Šis ir kritisks punkts objekta atpazīšanas servisa izvēlē, ja ir nepieciešams serviss, kurš spēj apmierināt klientu augstās prasības.

Objektu atpazīšanas serviss “Amazon Rekognition” visu eksperimentu laikā uzrādīja augstus un stabilus rezultātus un noteikti ir viens no labākajiem, ja ne pats labākais, objektu atpazīšanas attēlā servisiem, ja no objektu atpazīšanas modeļa tiek sagaidīta augsta atpazīto

objektu kvantitāte, varbūtība un tiek plānots izmantot iepriekš apmācītu modeli, kuru jau piedāvā pakalpojuma devējs.

Secinājumi

- Dažādu kognitīvo servisu attēlu atpazīšana, izmantojot gatavus mašīnmācības modeļus parāda dažādus rezultātus.
- Objektu atpazīšanas kognitīvais serviss “*Amazon Rekognition*”, izvēlētajos 10 attēlos atpazīna skaitliski visvairāk objektu pie visiem uzticamības līmeņiem.
- Rezultāti liecina par to, ka daži kognitīvo servisu iepriekš izveidotie modeļi nav pietiekami apmācīti, lai varētu noteikt objektus ar lielāku uzticamības līmeni.
- Objektu atpazīšanas kognitīvais serviss “*Microsoft Azure Vision*”, izvēlētajos 10 attēlos atpazīna skaitliski vismazāk objektu, un to uzticamības līmenis bija zem 90%.

Summary

Different cognitive services for computer vision to detect objects has showed different results.

Amazon Recognition cognitive service detected more objects from 10 presented images with highest confidence scores.

The results is showing that some of cognitive service models for machine learning is not enough trained to detect objects with highest confidence score.

Microsoft Azure vision cognitive service detected less objects from 10 presented images and with lowest confidence scores.

Literatūra

- [1] “New Artificial Neurons Can `Think` Faster Than You”, 2018. [Tiešsaiste]. Pieejams: <https://interestingengineering.com/new-artificial-neurons-can-think-faster-than-you>. [Piekļuve: 20.03.2020.].
- [2] “AI & Machine Learning Products - Vision AI”, 2020. [Tiešsaiste]. Pieejams: <https://cloud.google.com/vision>. [Piekļuve: 20.03.2020.].
- [3] “What is Computer Vision?”, 2020. [Tiešsaiste]. Pieejams: <https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/home>. [Piekļuve: 20.03.2020.].
- [4] “The 10 tech companies that have invested the most money in AI”, 2018. [Tiešsaiste]. Pieejams: <https://www.techrepublic.com/article/the-10-tech-companies-that-have-invested-the-most-money-in-ai/>. [Piekļuve: 20.03.2020.].
- [5] “Computer Vision Applications in 10 Industries”, 2020. [Tiešsaiste]. Pieejams: https://algorithmxlab.com/blog/computer-vision/#What_is_Computer_Vision?. [Piekļuve: 20.03.2020.].
- [6] “Object Detection: Current and Future Directions”, 2015. [Tiešsaiste]. Pieejams: <https://www.frontiersin.org/articles/10.3389/frobt.2015.00029/full>. [Piekļuve: 20.03.2020.].
- [7] “Amazon Rekognition Image”, 2020. [Tiešsaiste]. Pieejams: <https://aws.amazon.com/rekognition/image-features/?nc=sn&loc=3&dn=2>. [Piekļuve: 21.03.2020.].
- [8] “Watson Visual Recognition”, 2020. [Tiešsaiste]. Pieejams: <https://www.ibm.com/cloud/watson-visual-recognition>. [Piekļuve: 21.03.2020.].
- [9] “Detect common objects in images”, 2020. [Tiešsaiste]. Pieejams: <https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/concept-object-detection>. [Piekļuve: 22.03.2020.].