



Faculteit Bedrijf en Organisatie

Het gebruik van Computer Vision API's voor de beschrijving van cultureel-erfgoedcollecties

Nastasia Vanderperren

Scriptie voorgedragen tot het bekomen van de graad van
professionele bachelor in de toegepaste informatica

Promotor:
Koen Mertens
Co-promotor:
Henk Vanstappen

Instelling: Huis van Alijn

Academiejaar: 2018-2019

Derde examenperiode

Faculteit Bedrijf en Organisatie

Het gebruik van Computer Vision API's voor de beschrijving van cultureel-erfgoedcollecties

Nastasia Vanderperren

Scriptie voorgedragen tot het bekomen van de graad van
professionele bachelor in de toegepaste informatica

Promotor:
Koen Mertens
Co-promotor:
Henk Vanstappen

Instelling: Huis van Alijn

Academiejaar: 2018-2019

Derde examenperiode

Woord vooraf

Samenvatting

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus.

Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Inhoudsopgave

1	Inleiding	13
1.1	Probleemstelling	13
1.2	Onderzoeksvraag	14
1.3	Onderzoeksdoelstelling	14
1.4	Opzet van deze bachelorproef	15
2	Stand van zaken	17
2.1	Computer vision	17
2.2	Computer Vision API	19
2.3	Computer vision voor cultureel erfgoed	20
2.3.1	Uitdaging	20
2.3.2	Resultaten	21
2.4	Experimenten met beeldherkenning in musea	22

3	Methodologie	25
4	Conclusie	27
A	Onderzoeksvoorstel	29
A.1	Introductie	29
A.2	Stand van zaken	30
A.2.1	Computer Vision en (cultureel) erfgoed	30
A.2.2	Theoretisch onderzoek	31
A.2.3	Computer Vision API's	32
A.3	Methodologie	33
A.4	Verwachte resultaten	33
A.5	Verwachte conclusies	34
	Bibliografie	37

Lijst van figuren

- A.1 Het overzicht van de velden van de basisregistratie en het aantal keer dat ze aanwezig zijn in de collectiedata van MSK Gent. 35
- A.2 Vrouw en man met boek worden door Microsoft Computer Vision geïdentificeerd als man en vrouw die een selfie nemen (Roddiss, 2018). 36

Lijst van tabellen

1. Inleiding

In deze bachelorproef wordt onderzocht of Computer Vision API's zoals Clarifai, Google Cloud Vision of Microsoft Computer Vision gebruikt kunnen worden om collectiemedewerkers te ondersteunen bij het beschrijven van cultureel-erfgoedobjecten. De digitale fotocollectie van het Huis van Alijn werd hiervoor als testcase gebruikt.

Dit onderzoek zal als case mee opgenomen worden in het cultureel-erfgoedproject 'Beeldherkenning in de registratiepraktijk' van FOMU. In dit project wil het museum, samen met Datable en PACKED, onderzoek voeren naar de toepassing van Visual Recognition Services (of Computer Vision API's) voor de basisregistratie en iconografische ontsluiting van erfgoedobjecten. Hiervoor maken ze gebruik van de eigen collectie, maar leveren ook Netwerk Oorlogsbronnen, MoMu en Erfgoedcel Brugge beelden aan (Derveaux, 2019).

1.1 Probleemstelling

De Vlaamse musea lijden aan een achterstand m.b.t. het registreren van de eigen collectie. Het registreren van collectiestukken bestaat o.m. uit het systematisch beschrijven van de historische context en de kenmerken waarmee je het object kan identificeren, zoals de titel, een korte beschrijving, het soort object (schilderij, foto, standbeeld, stoel), afmetingen, gebruikte materiaal,... Dit is tijdrovend werk dat door domeinexperten gedaan wordt. Vooral formele en administratieve gegevens worden geregistreerd. Voor het beschrijven van inhoudelijke informatie zoals afgebeelde personen of objecten, emoties en sfeer ontbreekt het de musea aan tijd en personeel, terwijl dat net de informatie is die interessant is voor ontsluiting en onderzoek.

Daarnaast is er ook een digitalisering aan de gang waarbij steeds meer collectiestukken

gedigitaliseerd worden. Ook deze digitale beelden moeten voorzien worden van metadata of tags. Het doorzoeken of vinden van die digitale beelden is immers moeilijk als je bij de zoekactie geen gebruik kunt maken van trefwoorden. Dit verschilt van digitale tekstbestanden, waarbij op basis van *full text search* bestanden teruggevonden kunnen worden. Ook voor dit werk ontbreekt het de musea aan tijd en mankracht. Het gevolg is dat musea over steeds meer beelden beschikken die niet ontsloten of gebruikt kunnen worden.

Daarom willen we in deze bachelorproef onderzoeken of artificiële intelligentie (AI) de collectiemedewerker kan bijstaan in het beschrijven van de cultureel-erfgoedcollecties. Beeldherkenningssoftware is er de laatste jaren enorm op vooruitgegaan en wordt ook steeds eenvoudiger om te gebruiken.

1.2 Onderzoeksvraag

De centrale vraag in dit onderzoek is of Computer Vision API's gebruikt kunnen worden voor het (inhoudelijk) beschrijven van cultureel-erfgoedcollecties. Aan de hand van één gekozen API zal nagegaan worden of het ingebouwde model voldoende is voor beschrijving van de beelden en of het eenvoudig is om de beeldherkenningssoftware te trainen indien de beschrijving van het ingebouwde model niet voldoet.

We willen hierbij ook te weten komen of de software eenvoudig in gebruik is zodat museummedewerkers zelf de API's kunnen gebruiken om hun beelden te *taggen*. Gebruiksgemak zal daarom een belangrijk aspect zijn voor de keuze van de API. Een ander belangrijk aspect is de aanwezigheid van goede documentatie en tutorials.

1.3 Onderzoeksdoelstelling

Het onderzoek zal uitgevoerd worden op een deel van de fotocollectie van het Huis van Alijn. Het Huis van Alijn is het museum van het dagelijkse leven en beschikt over een grote fotocollectie die het dagelijkse leven in de twintigste eeuw documenteert. Samen met Huis van Alijn zullen er use cases geformuleerd worden waarvoor beeldherkenning ingezet kan worden.

We focussen ons in de eerste plaats op fotocollecties omdat we vermoeden dat de Computer Vision API's hier het best op scoren. Een *proof of concept* zal opgezet worden om de beelden te laten taggen en de beeldherkenningssoftware te trainen. Vervolgens zullen we een vergelijking maken tussen de metadata van de beelden die aangeleverd werden door de API en de metadata van de beelden zoals ze door het museummedewerker aangeleverd werden.

Met dit onderzoek wordt de haalbaarheid van het gebruik van Computer Vision API's in de erfgoedsector onderzocht. Tevens zijn de use cases gekend waarvoor beeldherkenningssoftware een goede partner is.

1.4 Opzet van deze bachelorproef

De rest van deze bachelorproef is als volgt opgebouwd:

In Hoofdstuk 2 wordt een overzicht gegeven van de stand van zaken binnen het onderzoeksdomein, op basis van een literatuurstudie.

In Hoofdstuk 3 wordt de methodologie toegelicht en worden de gebruikte onderzoekstechnieken besproken om een antwoord te kunnen formuleren op de onderzoeksvragen.

In Hoofdstuk 4, tenslotte, wordt de conclusie gegeven en een antwoord geformuleerd op de onderzoeksvragen. Daarbij wordt ook een aanzet gegeven voor toekomstig onderzoek binnen dit domein.

2. Stand van zaken

In dit hoofdstuk wordt uitgelegd wat Computer Vision en Computer Vision API's zijn en inhouden. Tevens wordt een stand van zaken gegeven van het onderzoek naar beeldherkenning voor cultureel-erfgoedcollecties, de uitdagingen hierbij en experimenten met beeldherkenningstoepassingen in musea.

2.1 Computer vision

Computer Vision (CV) is het onderzoeksveld waarin technieken ontwikkeld worden om computers te helpen bij het zien en het begrijpen van de inhoud van digitale beelden zoals foto's en video's. Het is een deelgebied van Artificiële Intelligentie en Machine Learning (Wikipedia contributors, 2019a). Het doel van deze techniek is om computers te gebruiken voor taken inzake digitale beelden of video's die normaal door mensen gedaan zouden worden. Computer Vision wordt momenteel met succes toegepast voor een breed scala van uitdagingen zoals inspectie van machines, geautomatiseerde kassa-afrekeningen in de retail, motion capture, medische beeldvorming, bewaking en politieonderzoek, het herkennen van vingerafdrukken en het assisteren van mensen bij het identificeren van de inhoud van een foto of video (Brownlee, 2019a).

De bovenvermelde uitdagingen gebruiken verschillende CV-taken voor het verkrijgen en analyseren van informatie uit de beelden, zoals beeldherkenning en bewegingsanalyse. Voor het onderzoek van deze bachelorproef is beeldherkenning de belangrijkste taak. Enkele typische voorbeelden hiervan zijn (Wikipedia contributors, 2019a):

- OCR: het identificeren van tekens in beelden van teksten;
- object classification: het classificeren van het beeld in een bepaalde categorie;

- object recognition: het identificeren van objecten in het beeld;
- object detection: het herkennen van een object of entiteit (zoals een persoon), maar tevens aanduiden op het beeld waar het zich bevindt;
- face detection en recognition: het herkennen van personen en aanduiden waar ze zich op het beeld bevinden;
- content-based image retrieval: het vinden van beelden in een dataset met een bepaalde inhoud, bv. gelijkaardige beelden vinden.

Momenteel zijn algoritmes die gebaseerd zijn op Convolutional Neural Networks (CNN) het best om deze taken uit te voeren. CNN voor beeldherkenning werd voor het eerst met succes geïntroduceerd in 2012. Sindsdien kent het veld van beeldherkenning een enorme groei. Artificiële neurale netwerken, zoals CNN, leren taken uit te voeren op basis van een trainingsset. Als je wil dat het netwerk een bus herkent, dan bestaat de trainingsset uit verschillende voorbeelden van een bus, maar ook van wat een bus niet is. Wanneer, na de training, het model nieuwe beelden ziet, zou het moeten kunnen zeggen of het afgebeelde object een bus is of niet. Als het netwerk nog een andere taak moet uitvoeren, dan moet het netwerk nieuwe trainingsdata krijgen, bijvoorbeeld om te leren wat een fiets is (Pokharna, 2016).

ImageNet is de dataset die doorgaans gebruikt wordt voor het ontwikkelen van toepassingen voor Computer Vision. ImageNet is een visuele databank die bestaat uit meer dan veertien miljoen beelden die annotaties hebben over wat er op de beelden staat. De databank voorziet 20.000 concepten, zoals ballon of aardbei. Elk van deze concepten bestaat uit enkele honderden beelden (Wikipedia contributors, 2019b, p. 2; Brownlee, 2019b). Sinds 2010 vindt jaarlijks de ImageNet Large Scale Visual Recognition Challenge (ILSVRC) plaats. Hierbij wordt een beperkte set van ImageNet (één miljoen beelden voor duizend concepten) gebruikt om de prestatie van de verschillende algoritmes te testen. Het valt hierbij op dat de resultaten jaarlijks enorm stijgen. Terwijl er in 2011 een foutenmarge was van 25% voor de uitdaging van image classification, was dit tegen 2017 al teruggevallen tot 5%.

Russakovsky e.a. (2014) stellen vast dat domeinexperten beter blijven dan de CV-modellen, maar dat de CV-modellen significant beter scoren dan mensen die geen expert zijn. Mensen zijn bijvoorbeeld goed in het onderscheiden van een hond en een kat, maar een CV-model kan dan weer veel beter de honden indelen in het juiste hondenras. De modellen waren in het algemeen erg goed in het herkennen van verschillende soorten zoogdieren en hun rassen, maar hadden het veel moeilijker met metalen en doorzichtige voorwerpen. Ze hebben ook moeilijkheden met beelden die in meer dan vijf verschillende concepten te classificeren zijn (de menselijke classifier had het hier ook moeilijk mee) en met het identificeren van kleine objecten. Daarnaast hebben de classifiers ook problemen met foto's die vervormd zijn door het gebruik van kleurenfilters of met abstracte voorstellingen van voorwerpen zoals 3D rendered beelden, schilderijen, tekeningen en schetsen. Aangezien schilderijen, tekeningen, schetsen en objecten algemeen voorkomen in erfgoedcollecties, kan dit een probleem vormen.

2.2 Computer Vision API

CV API's zijn REST API's die het developers mogelijk maken om machine learning toe te passen zonder hier zelf een expert in te zijn. De services zijn zelf al getraind, waardoor het niet nodig is om zelf voor datasets te zorgen en de modellen te trainen. De meeste services bieden ook de mogelijkheid aan om eigen modellen te creëren. De diensten zijn zo toegankelijk dat iedere ontwikkelaar zonder machine learning expertise modellen kan bouwen. Volgens Lardinois (2018) zijn de services zo eenvoudig dat zelfs mensen zonder enige programmeerervaring de diensten kunnen gebruiken om beelden te laten taggen en categoriseren. De bekendste services zijn Clarifai, IBM Visual Recognition, Microsoft Computer Vision, Amazon Rekognition en Google Cloud Vision.

Welke datasets gebruikt zijn om de CV API te trainen is niet duidelijk. Wel hebben zowel Clarifai, Google als Microsoft deelgenomen aan de ImageNet Challenge. In 2013 was ZFNet (Clarifai) de winnaar van de beeldclassificatietask, in 2014 behaalde Google topresultaten voor objectdetectie met hun GoogLeNet-model en in 2015 behaalde Microsoft Research topresultaten met het ResNet-model. Op basis hiervan kunnen we veronderstellen dat de services minstens ImageNet gebruikt hebben (Brownlee, 2019b).

Het gebruik van de CV API's is zeer toegankelijk. De meeste van deze CV API's beschikken over een webinterface waarmee de service uitgetest kan worden en waarmee custom modellen gebouwd kunnen worden. Voor intensiever gebruik is er een REST API waarmee over HTTP de service aangeroepen wordt en een antwoord in JSON ontvangen wordt. Daarnaast hebben al die services ook API clients in verschillende programmeertalen (o.a. Java, Python, Javascript/Node.js, C#, PHP) die het de developer toelaat om de services te gebruiken in zelf ontwikkelde applicaties of scripts. Beeldmateriaal kan op verschillende manieren aan de services bezorgd worden:

- via een URL naar het beeld;
- in de vorm van base64;
- opgeslagen in het ecosysteem van de aanbieder (bv. S3 buckets voor Amazon Rekognition, of Google Cloud Storage buckets voor Google Cloud Vision).

De services kunnen verschillende aspecten van een beeld beschrijven:

- OCR: het omzetten van tekens in beelden naar tekst;
- object detection: beschrijven van op het beeld aanwezige objecten of personen;
- algemene beschrijving van het soort beeld;
- de aard van het beeld (foto, tekening, schilderij)
- gezichtsherkenning: identificatie van bekende personen;
- landmarks: identificatie van bekende plaatsen;
- aboutness: emoties, sfeer en thema van het beeld;
- gebruikte kleuren, zowel hoofdkleuren als accentkleuren;
- NSFW-score: score op de aanwezigheid van content waaraan aanstoot genomen kan worden zoals erotiek, seksisme, racisme en geweld.

De resultaten worden meestal in de vorm van tags gegeven, maar ook taglines (een

korte zin die het beeld beschrijft) of codes (kleurcodes) zijn mogelijk, net als URI's (voor gelijkaardige beelden die op het web gevonden zijn). Bij de resultaten wordt een waarschijnlijkheidsscore gegeven (Vanstappen, 2019).

Na een vergelijking van de top vijf van CV API's was Oberoi (2016) onder de indruk van de resultaten. De services waren voldoende om de kern van een beeld te krijgen op een erg snelle en relatief goedkope manier.

2.3 Computer vision voor cultureel erfgoed

Er is al veel onderzoek gebeurd naar het gebruik van computer vision voor cultureel erfgoed. Omwille van de digitalisering blijven de beeldcollecties in musea en cultureel-erfgoedinstellingen groeien waardoor er een enorme registratie-achterstand is. Van CV wordt verwacht dat het het werk van de registratoren verlicht en de achterstanden wegwerkt. Computers worden hier met andere woorden opgeleid als domeinexperten. Ze moeten van beelden van kunstwerken of erfgoedobjecten kunnen bepalen:

- wie de kunstenaar of maker is;
- uit welk materiaal de werken gemaakt zijn;
- in welke periode of jaar het werk gemaakt is;
- welk soort object het werk is;
- en tot welke kunststroming het werk behoort.

In zulke projecten trainen onderzoekers een eigen model of classifier om beelden van erfgoedobjecten te classificeren of - in een minderheid - te labelen. Een voorbeeld van zo'n onderzoek is het in 2017 gestarte Belgische INSIGHT-project, waarin men wil onderzoeken of AI gebruikt kan worden voor het beschrijven van museale collecties. De digitale collecties van de Koninklijke Musea voor Schone Kunsten van België en de Koninklijke Musea voor Kunst & Geschiedenis worden als testcase gebruikt. Het einddoel van dit project is de ontwikkeling en release van een reeks praktische machine learning tools om digitale collecties te beheren (Universiteit Antwerpen, 2017a).

Naast deze onderzoeken wordt ook onderzocht of CV gebruikt kan worden voor het herkennen van vervalsingen (Dickson, 2018), of om te verifiëren of een kunstenaar een bepaald werk gemaakt heeft op basis van penseelstreken en potloodstrepen (ook om vervalsingen te ontdekken) (Elgammal, 2017). Daarnaast wordt AI gebruikt om nieuwe kunstwerken te maken. Dickson 2019 ING, Microsoft, TU Delft en Mauritshuis hebben in 2016 bijvoorbeeld zo een nieuw kunstwerk van Rembrandt gemaakt op basis van een databank van zijn schilderijen (ING, 2016).

2.3.1 Uitdaging

Een probleem bij het gebruik van computer vision voor erfgoedobjecten zijn de veel kleinere datasets:

- Blessings en Wen (2013) creëerden een dataset van 1.400 beelden via Google Images van de kunstenaars Cézanne, Dali, Dürer, Monet, Picasso, Rembrandt en Van Gogh (tweehonderd beelden per kunstenaar);
- Mensink en van Gemert (2014) ontwikkelden een dataset van meer dan 112.000 beelden die het Rijksmuseum via een open licentie en API ter beschikking stelden. De set bestond uit verschillende soorten objecten die een periode overspannen van het begin van onze jaartelling tot de late 19^e eeuw. Ze wilden hiermee een dataset creëren die representatief is voor een museumcollectie. De dataset bestond uit kunstwerken van meer dan 6.000 kunstenaars, die op te delen zijn in 1.824 objecttypes en waarvoor in totaal 406 verschillende soorten materialen gebruikt waren. Het ging onder meer over beelden van schilderijen, foto's, keramiek, en meubels. De onderzoekers stelden hun geannoteerde datasets ter beschikking voor verder onderzoek.
- Sabatteli, Kestemont, Daelemans en Geurts (2018) gebruikten die dataset van Mensink en Gemert, maar vulden dit aan met een kleinere dataset uit het DAMS Antwerpen dat de collecties omvat van de belangrijkste GLAM-instellingen van Antwerpen. In totaal bestond de dataset uit ongeveer 100.000 beelden van kunstwerken die op te delen zijn in 2.016 materialen van kunstwerken, 135.809 beelden van kunstwerken die op te delen zijn in 1.677 objectsoorten en 90.674 beelden van kunstwerken van 2.099 kunstenaars.
- Elgammal, Mazzone, Liu, Kim en Elhoseiny (2018) gebruikten een dataset van 76.921 beelden van schilderijen afkomstig van 1.119 kunstenaars uit de 15^e eeuw tot de hedendaagse tijd. De dataset werd gehaald van WikiArt, een online kunstencyclopedie.

Dit stelt weinig voor als je dit vergelijkt met de veertien miljoen beelden van ImageNet. Die beperking heeft te maken met copyright issues op beelden van kunstwerken en de afwezigheid van goede metadata bij de wel beschikbare beelden.

Om dit probleem aan te pakken, verkenden zowel Sabatteli e.a. (2018) als Elgammal e.a. (2018) het veld van Transfer Learning. Hierbij worden geen neurale netwerken from scratch gebouwd, maar worden reeds ontwikkelde neurale netwerken verder gefinetuned. De onderzoekers gebruikten hiervoor neurale netwerken die getraind waren op ImageNet en die op de ImageNet Challenge state of the art waren. De artificiële netwerken werden verder gefinetuned met de testdata. Beide onderzoeken toonden aan dat het beter was om de artificiële netwerken te finetunen dan from scratch zelf een model te bouwen. Netwerken die verder getraind waren op kunstcollecties leverden wel nog betere resultaten op dan de neurale netwerken die enkel op ImageNet getraind waren (Sabatteli e.a., 2018). Elgammal e.a. (2018) vermoeden dat betere resultaten mogelijk zijn met from scratch netwerken als er grotere datasets ter grootte van ImageNet voor erfgoedobjecten ter beschikking zouden zijn.

2.3.2 Resultaten

Voor het classificeren van kunstwerken per kunstenaar behaalden Blessings en Wen (2013) goede resultaten. Ze behaalden een resultaat van 85% correctheid en vermoedden dat

dit nog beter kon indien ze een grotere dataset hadden. Hun onderzoek concentreerde zich wel maar op zeven kunstenaars. Bij de onderzoeken met meer kunstenaars, zoals *The Rijksmuseum Challenge: Museum-Centered Visual Recognition* van Mensink en van Gemert (2014) (6.629 kunstenaars, maar het model werd enkel getraind met de 374 kunstenaars die minstens tien werken in de dataset hebben) en Sabatteli e.a. (2018) (2.096 kunstenaars) waren de resultaten lager. Bij die laatste werd een maximumscore van 69% behaald.

Het classificeren van werken per materiaal werd beschouwd als de eenvoudigste classificatie, onder meer omdat het aantal klassen vijf keer kleiner is dan bij de kunstenaarsclassificatie en typeclassificatie. Sabatteli e.a. (2018) vermoedden dat deze uitdaging ook het meest overeenstemt met de uitdagingen voor de ImageNet Challenge. Ze haalden hier een resultaat van 92,95% op. Mensink en van Gemert (2014) behaalden een resultaat van 94%. Ook voor het classificeren van kunstwerken per objecttype werden in beide onderzoeken gelijkaardige resultaten behaald.

Tot slot was ook het classificeren op basis van kunststroming succesvol. De uitdaging bestond uit het classificeren van 76.921 beelden in twintig kunststromingen. Er werd een score behaald van maximum 73%. Ondanks dat het CV-model geen kennis heeft van kunstenaars, creatiedatum of geschiedenis van stijlen, worden de schilderijen uit dezelfde periode bij elkaar geplaatst en is er een geleidelijke overgang van stijlen te zien vanaf de Renaissance (15^e eeuw) tot de abstracte kunst (20^e eeuw) (Elgammal e.a., 2018).

2.4 Experimenten met beeldherkenning in musea

Verschillende musea zijn reeds aan de slag gegaan met beeldherkenning en Computer Vision API's. De use cases zijn anders dan bij het (academisch) onderzoek. Musea willen de bezoekers van hun website nieuwe ervaringen aanbieden om hun beeldcollectie te ontdekken, maar willen tevens hun beelden beter doorzoekbaar maken. Uit onderzoek blijkt namelijk dat gebruikers de behoefte hebben om beelden te kunnen zoeken op basis van inhoudelijke kenmerken. Dit kan gaan over identificeerbare objecten (Eiffeltoren), generieke objecten (stoel) of op basis van iconologische thema's (Het Laatste Avondmaal) en abstracte begrippen (geluk, jeugd) (Vanstappen, 2019). Om deze ervaringen aan te bieden, beginnen musea meer en meer te experimenteren met beeldherkenning.

The Museum of Modern Art (MoMA) gebruikt AI-diensten van Google om historische foto's van afgelopen tentoonstellingen te koppelen aan de beelden uit de kunstcollectie die te zien zijn op die tentoonstellingszichten. Het CV-model analyseerde hiervoor alle foto's van tentoonstellingen. Wanneer het model een kunstwerk op de foto herkende, dan legde het een koppeling met het beeld van het kunstwerk. MoMA stelde hierbij vast dat de CV goed scoorde op tweedimensionale, statische afbeeldingen (zoals een schilderij), maar dat het slechter scoort op 3D-objecten (zoals een sculptuur) of bewegende beelden (**MOMA2018**).

Voor het Noorse Nasjonalmuseet werd het deep learning framework Caffe gebruikt om

composicionele gelijkenissen te zoeken tussen de kunstwerken. De kunstwerken werden eveneens geassocieerd op basis van de Iconclass. Dit resulteerde in een vernieuwde publiekstoegang waarbij kunstwerken op basis van gelijkenissen gevisualiseerd werden. Hoe meer gelijkenissen een kunstwerk heeft, hoe dichter de kunstwerken bij elkaar staan (Nasjonalmuseet2017).

Wellcome Collection heeft 120.000 beelden die beschikbaar zijn via een API en nog eens veertig miljoen beelden die via een open licentie beschikbaar zijn voor het publiek. Ondertussen blijft het digitaliseringsteam de rest van de collectie digitaliseren waardoor er dagelijks duizenden nieuwe beelden bijkomen. Het is onmogelijk om deze beelden manueel te gaan beschrijven en ontsluiten, terwijl het zonder metadata onmogelijk is om beelden te vinden. Daarom wordt machine learning gebruikt om de collectie meer toegankelijk te maken. Wellcome Collection heeft zelf een model getraind om de beelden te categoriseren en om gelijkaardige beelden te clusteren. Dit wordt voornamelijk intern gebruikt om ongewenste beelden te verwijderen uit de collectiewebsite en de registratie-workflow te verbeteren. In de toekomst wil men het mogelijk maken dat bezoekers van de collectiewebsite gelijkaardige beelden vinden op basis van een beeld (Pim, 2017).

Het Britse webbedrijf CogApp liet drie CV API's (Clarifai, Google Cloud Vision en Microsoft Computer Vision) los op tweeduizend beelden van schilderijen van het Zweedse Nationalmuseum. Ze wilden hiermee de collectie beter doorzoekbaar maken op basis van de inhoudelijke kenmerken. Dit resulteerde in een visuele zoekmachine waarin een selectie van beelden op basis van filters verkregen kan worden. Iedere tag die een CV API gaf aan een beeld werd gebruikt als filter, zoals Renaissance, snor, cape, baby, etc. CogApp concludeerde hieruit dat de CV API's eenvoudig zijn in gebruik en accurate beschrijvingen kunnen geven van beelden. Ze vermoedden dat foutieve beschrijvingen een gevolg zijn van het trainen van de CV API met hedendaagse beelden, terwijl de beelden van het Nationalmuseum historisch zijn (Hindle, 2017).

Ook Auckland Museum heeft een grote collectie: zeven miljoen objecten die gaan van kunst tot archieven, culturele collecties, natuurwetenschappelijke specimen, oorlogscollecties en een onderzoeksbibliotheek. Door digitalisering van de collectie komen er maandelijks tweeduizend nieuwe beelden bij. Er werd ingeschat dat het decennia zou duren eer de volledige gedigitaliseerde collectie geregistreerd zal zijn. Daarom verkent ook Auckland Museum CV API's om de collectie automatisch te laten taggen en een basisrecord per beeld te creëren. Microsoft Computer Vision werd gebruikt om een test te doen met tweeduizend beelden. Tags met een lagere waarschijnlijkheidsscore dan 60% werden verwijderd om te vermijden dat er beschamende of misleidende records gepubliceerd worden. Moriarty (2018) concludeert dat CV API nuttig zijn om snel basisrecords te creëren voor beelden waarmee zowel interne als externe gebruikers zich een weg kunnen zoeken doorheen de collectie, maar er is nog werk om dit te gaan implementeren. Hij heeft nog vragen hoe al die tags gereviewed moeten worden en hoe (en of) je aan de gebruiker moet laten weten dat de tags door een AI-systeem gecreëerd werden.

Voor de Sarjeant Gallery werd een nieuwe collectiewebsite ontwikkeld. De collectie kan op de nieuwe website doorzocht worden op basis van kleur, beeldoriëntatie en tags. Die tags werden gegenereerd via de Google Vision API. Het originele plan was om die

tags enkel te gebruiken voor intern gebruik zodat collectiemedewerkers de tags konden gebruiken om sets van beelden rond een bepaald onderwerp te maken. Men vond de tags echter zo goed dat besloten werd om ze ook op de website te publiceren. Doordat veel kunstwerken geen onderwerpbeschrijving hadden, konden de tags gebruikt worden om gerelateerde kunstwerken te vinden. Foutieve tags worden verborgen door de collectiemedewerkers (Rowe, 2017). Ook The Swedish National Heritage Board heeft een webinterface gecreëerd waarbij de collectie doorzoekbaar gemaakt wordt met onder meer Google Cloud Vision (Haskiya, 2019).

Ook in Vlaanderen wordt geëxperimenteerd met CV API. Vanstappen (2019) heeft voor MoMu onderzocht in welke mate off-the-shelf CV API zonder training bruikbaar zijn voor erfgoedcollecties. Het was een vergelijkend onderzoek waarbij verschillende modellen van Clarifai, Google Cloud Vision en Microsoft Computer Vision vergeleken werden op een set van 164 beelden van objecten (inclusief close-ups en detailfoto's), catwalkbeelden, eventfoto's en scenografie. Vanstappen (2019) concludeerde dat de verkregen tags voor een dergelijk gespecialiseerd onderwerp geen hoge kwaliteit hadden, maar dat ze wel bruikbare resultaten gaven voor een globale beschrijving van de erfgoedcollecties. Het sterke punt is vooral de performantie. Een API is veel sneller en goedkoper dan een menselijke registrator. Bovendien geeft de beeldherkenningssoftware ook onderwerpen die buiten de traditionele collectiebeschrijving vallen, zoals kleur en sfeer.

3. Methodologie

Etiam pede massa, dapibus vitae, rhoncus in, placerat posuere, odio. Vestibulum luctus commodo lacus. Morbi lacus dui, tempor sed, euismod eget, condimentum at, tortor. Phasellus aliquet odio ac lacus tempor faucibus. Praesent sed sem. Praesent iaculis. Cras rhoncus tellus sed justo ullamcorper sagittis. Donec quis orci. Sed ut tortor quis tellus euismod tincidunt. Suspendisse congue nisl eu elit. Aliquam tortor diam, tempus id, tristique eget, sodales vel, nulla. Praesent tellus mi, condimentum sed, viverra at, consectetur quis, lectus. In auctor vehicula orci. Sed pede sapien, euismod in, suscipit in, pharetra placerat, metus. Vivamus commodo dui non odio. Donec et felis.

Etiam suscipit aliquam arcu. Aliquam sit amet est ac purus bibendum congue. Sed in eros. Morbi non orci. Pellentesque mattis lacinia elit. Fusce molestie velit in ligula. Nullam et orci vitae nibh vulputate auctor. Aliquam eget purus. Nulla auctor wisi sed ipsum. Morbi porttitor tellus ac enim. Fusce ornare. Proin ipsum enim, tincidunt in, ornare venenatis, molestie a, augue. Donec vel pede in lacus sagittis porta. Sed hendrerit ipsum quis nisl. Suspendisse quis massa ac nibh pretium cursus. Sed sodales. Nam eu neque quis pede dignissim ornare. Maecenas eu purus ac urna tincidunt congue.

Donec et nisl id sapien blandit mattis. Aenean dictum odio sit amet risus. Morbi purus. Nulla a est sit amet purus venenatis iaculis. Vivamus viverra purus vel magna. Donec in justo sed odio malesuada dapibus. Nunc ultrices aliquam nunc. Vivamus facilisis pellentesque velit. Nulla nunc velit, vulputate dapibus, vulputate id, mattis ac, justo. Nam mattis elit dapibus purus. Quisque enim risus, congue non, elementum ut, mattis quis, sem. Quisque elit.

Maecenas non massa. Vestibulum pharetra nulla at lorem. Duis quis quam id lacus dapibus interdum. Nulla lorem. Donec ut ante quis dolor bibendum condimentum. Etiam egestas

tortor vitae lacus. Praesent cursus. Mauris bibendum pede at elit. Morbi et felis a lectus interdum facilisis. Sed suscipit gravida turpis. Nulla at lectus. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Praesent nonummy luctus nibh. Proin turpis nunc, congue eu, egestas ut, fringilla at, tellus. In hac habitasse platea dictumst.

Vivamus eu tellus sed tellus consequat suscipit. Nam orci orci, malesuada id, gravida nec, ultricies vitae, erat. Donec risus turpis, luctus sit amet, interdum quis, porta sed, ipsum. Suspendisse condimentum, tortor at egestas posuere, neque metus tempor orci, et tincidunt urna nunc a purus. Sed facilisis blandit tellus. Nunc risus sem, suscipit nec, eleifend quis, cursus quis, libero. Curabitur et dolor. Sed vitae sem. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Maecenas ante. Duis ullamcorper enim. Donec tristique enim eu leo. Nullam molestie elit eu dolor. Nullam bibendum, turpis vitae tristique gravida, quam sapien tempor lectus, quis pretium tellus purus ac quam. Nulla facilisi.

4. Conclusie

Curabitur nunc magna, posuere eget, venenatis eu, vehicula ac, velit. Aenean ornare, massa a accumsan pulvinar, quam lorem laoreet purus, eu sodales magna risus molestie lorem. Nunc erat velit, hendrerit quis, malesuada ut, aliquam vitae, wisi. Sed posuere. Suspendisse ipsum arcu, scelerisque nec, aliquam eu, molestie tincidunt, justo. Phasellus iaculis. Sed posuere lorem non ipsum. Pellentesque dapibus. Suspendisse quam libero, laoreet a, tincidunt eget, consequat at, est. Nullam ut lectus non enim consequat facilisis. Mauris leo. Quisque pede ligula, auctor vel, pellentesque vel, posuere id, turpis. Cras ipsum sem, cursus et, facilisis ut, tempus euismod, quam. Suspendisse tristique dolor eu orci. Mauris mattis. Aenean semper. Vivamus tortor magna, facilisis id, varius mattis, hendrerit in, justo. Integer purus.

Vivamus adipiscing. Curabitur imperdiet tempus turpis. Vivamus sapien dolor, congue venenatis, euismod eget, porta rhoncus, magna. Proin condimentum pretium enim. Fusce fringilla, libero et venenatis facilisis, eros enim cursus arcu, vitae facilisis odio augue vitae orci. Aliquam varius nibh ut odio. Sed condimentum condimentum nunc. Pellentesque eget massa. Pellentesque quis mauris. Donec ut ligula ac pede pulvinar lobortis. Pellentesque euismod. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent elit. Ut laoreet ornare est. Phasellus gravida vulputate nulla. Donec sit amet arcu ut sem tempor malesuada. Praesent hendrerit augue in urna. Proin enim ante, ornare vel, consequat ut, blandit in, justo. Donec felis elit, dignissim sed, sagittis ut, ullamcorper a, nulla. Aenean pharetra vulputate odio.

Quisque enim. Proin velit neque, tristique eu, eleifend eget, vestibulum nec, lacus. Vivamus odio. Duis odio urna, vehicula in, elementum aliquam, aliquet laoreet, tellus. Sed velit. Sed vel mi ac elit aliquet interdum. Etiam sapien neque, convallis et, aliquet vel, auctor non, arcu. Aliquam suscipit aliquam lectus. Proin tincidunt magna sed wisi. Integer blandit

lacus ut lorem. Sed luctus justo sed enim.

Morbi malesuada hendrerit dui. Nunc mauris leo, dapibus sit amet, vestibulum et, commodo id, est. Pellentesque purus. Pellentesque tristique, nunc ac pulvinar adipiscing, justo eros consequat lectus, sit amet posuere lectus neque vel augue. Cras consectetur libero ac eros. Ut eget massa. Fusce sit amet enim eleifend sem dictum auctor. In eget risus luctus wisi convallis pulvinar. Vivamus sapien risus, tempor in, viverra in, aliquet pellentesque, eros. Aliquam euismod libero a sem.

Nunc velit augue, scelerisque dignissim, lobortis et, aliquam in, risus. In eu eros. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Curabitur vulputate elit viverra augue. Mauris fringilla, tortor sit amet malesuada mollis, sapien mi dapibus odio, ac imperdiet ligula enim eget nisl. Quisque vitae pede a pede aliquet suscipit. Phasellus tellus pede, viverra vestibulum, gravida id, laoreet in, justo. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Integer commodo luctus lectus. Mauris justo. Duis varius eros. Sed quam. Cras lacus eros, rutrum eget, varius quis, convallis iaculis, velit. Mauris imperdiet, metus at tristique venenatis, purus neque pellentesque mauris, a ultrices elit lacus nec tortor. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent malesuada. Nam lacus lectus, auctor sit amet, malesuada vel, elementum eget, metus. Duis neque pede, facilisis eget, egestas elementum, nonummy id, neque.

A. Onderzoeksvoorstel

Het onderwerp van deze bachelorproef is gebaseerd op een onderzoeksvoorstel dat vooraf werd beoordeeld door de promotor. Dat voorstel is opgenomen in deze bijlage.

A.1 Introductie

De Vlaamse Overheid investeert in het Vlaamse cultureel erfgoed door collectiebeherende instellingen te ondersteunen en kwaliteitslabels uit te reiken. In ruil voor die steun verwacht de Vlaamse Overheid dat die instellingen een aantal taken uitvoeren, waaronder het registreren, inventariseren en metadateren van cultureel-erfgoedobjecten op een gestandaardiseerde manier, het onderzoeken (en faciliteren van onderzoek) en het presenteren van de collectie. (Agentschap Kunsten en Erfgoed, 2014; Gatz, 2016)

De collectiebeherende instellingen lijden aan een historische achterstand m.b.t. de registratie van de eigen collectie (Gatz, 2016). In 2018 werd daarom een nieuwe subsidielijlijn opgestart om de digitale collectieregistratie weg te werken. Deze subsidielijlijn werd opgestart vanuit de vaststelling dat de competenties en strategieën ontbreken om een inhaalbeweging te realiseren (Departement Cultuur, Jeugd en Media, 2018).

In de bachelorproef willen we onderzoeken in welke mate Computer Vision API's (vanaf nu afgekort als CVA), zoals Google Cloud Vision¹ of Microsoft Computer Vision API², ingezet kunnen worden om dit registratieproces te versnellen en als strategie gebruikt kunnen worden om een inhaalbeweging te realiseren. Momenteel gebeuren registraties

¹<https://cloud.google.com/vision/>

²<https://azure.microsoft.com/nl-nl/services/cognitive-services/computer-vision/>

door domeinexperten. Dit is tijdrovend werk. Met behulp van artificiële intelligentie (AI) kan dit proces deels geautomatiseerd worden. Dit geeft de collectieregistrator de mogelijkheid om zich met minder basaal werk bezig te houden en geeft de musea de kans hun collectie sneller te ontsluiten.

Het VKC Datahub Dashboard³ geeft een goed beeld van de registratieachterstand in Vlaanderen. Dit dashboard analyseert de collectieregistratie van de Vlaamse musea voor Schone (en in de toekomst ook Hedendaagse) Kunsten die aangesloten zijn bij de Vlaamse KunstCollectie (VKC)⁴, waaronder het aantal records ingevuld volgens de minimale registratie en het aantal records ingevuld volgens de basisregistratie.⁵ Dit zijn cijfers van de grootste kunstmusea van Vlaanderen. Uit de cijfers blijkt dat vooral formele en administratieve gegevens geregistreerd worden; inhoudelijke informatie, zoals afgebeelde persoon of afgebeeld concept, die interessant is voor ontsluiting en onderzoek, ontbreekt hoofdzakelijk.

In het verleden werd reeds (theoretisch) onderzoek verricht naar het gebruik van Computer Vision voor cultureel erfgoed en kunst.⁶ De resultaten van dit onderzoek waren veelbelovend. Nieuwe ontwikkelingen zorgen ervoor dat Computer Vision steeds meer accuraat wordt en steeds eenvoudiger in gebruik, waardoor het meer en meer een instrument wordt waarmee ook developers aan de slag kunnen (Hindle, 2017).

A.2 Stand van zaken

A.2.1 Computer Vision en (cultureel) erfgoed

Een aantal instellingen maken al gebruik van Computer Vision voor de ontsluiting van hun collectie:

- **The Museum of Modern Art (MoMA)**⁷ gebruikte diensten van Google via *Google Arts Culture Lab* om aan de historische foto's van afgelopen tentoonstellingen in MOMA foto's uit de kunstcollectie te koppelen. Het algoritme analyseerde hiervoor alle foto's van tentoonstellingszichten. Als het een kunstwerk op de foto's kon herkennen, dan werd er een koppeling gelegd met de afbeelding van dit kunstwerk in de collectie van MOMA.⁸ MOMA stelde hierbij vast dat het algoritme vooral goed scoort op het vlak van tweedimensionele, statische afbeeldingen (bv. een afbeelding van een kunstwerk), maar dat het minder goede resultaten geeft met afbeeldingen

³<https://dashboard.vlaamsekunstcollectie.be>

⁴<http://vlaamsekunstcollectie.be/>

⁵De minimale registratie omvat de elementen en velden die in overeenstemming met de ICOM Code of Ethics minimaal gedocumenteerd moeten worden wanneer een object een museum binnenkomt: d.i. bewaarinstelling, objectnummer, titel, korte beschrijving, objectnaam, verwervingsmethode, verwervingsbron en verwervingsdatum. De velden voor de basisregistratie bevatten de acht velden voor minimale registratie en voegen hier nog een tiental andere velden aan toe. Voor meer info, zie het Invalboek Objecten op CEST

⁶Zie infra, in het deel *Stand van zaken*.

⁷<https://www.moma.org/>

⁸Bekijk een voorbeeld over een tentoonstelling over Cézanne, Gauguin, Seurat en Van Gogh uit 1929

van 3D-objecten (bv. een standbeeld) of bewegend beeld (MOMA, 2018).

- Het **Smithsonian**⁹ gebruikt AI-technologie om stalen van planten te classificeren. Het Smithsonian is gestart met het systematisch digitaliseren van de collectie voor wetenschappers en in functie van online ontsluiting. De AI-technologie slaagde erin om via deze afbeeldingen twee gelijkaardige planten te onderscheiden met een succesgraad van meer dan 90% (Smith, 2017).
- Het **Nasjonalmuseet**¹⁰ was het onderwerp van het *Principal Components* onderzoek. Hierbij werd met het deep learning framework Caffe¹¹ gezocht naar compositio-nale gelijkenissen tussen kunstwerken en werden ze geclassificeerd op basis van de Iconclass-termen¹². Dit resulteerde in een vernieuwende publiekstoegang tot de collectie waarbij de kunstwerken op basis van gelijkenissen gevisualiseerd werden. Hoe meer gelijkenissen, hoe dichter de kunstwerken bij elkaar staan.¹³ (Nasjonalmuseet, 2017; Westvang, 2017)

A.2.2 Theoretisch onderzoek

Daarnaast is eerder theoretisch onderzoek gedaan naar het gebruik van Computer Vision binnen een erfgoedcontext. In onderstaande lijst worden de onderzoeken vermeld die zich focussen op museumcollecties:

- **Using Machine Learning for Identification of Art Paintings (2013):** In dit onderzoek werd machine learning gebruikt voor het classificeren van kunstwerken van zeven kunstenaars: Cézanne, Dali, Dürer, Monet, Picasso, Rembrandt en Van Gogh. Per kunstenaar werden er tweehonderd kunstwerken gezocht. In 87,13% van de gevallen was de computer correct. De onderzoekers vermoedden dat dit resultaat nog beter kan zijn als de training set groter is. (Blessings & Wen, 2013)
- **The Rijksmuseum Challenge: Museum-Centered Visual Recognition (2014):** Beelden die publiek beschikbaar zijn via de API van Het Rijksmuseum¹⁴ werd gebruikt voor het oplossen van vier challenges:
 - voorspel de kunstenaar van het afgebeelde object;
 - voorspel het materiaal dat gebruikt werd voor het afgebeelde object;
 - voorspel het jaar waarin het object gemaakt werd;
 - voorspel het soort object (schilderij, tekening, standbeeld...) dat afgebeeld wordt.

Het ging om objecten die afkomstig waren van de oudheid tot de late 19e eeuw en om een veelheid aan objecten: schilderijen, foto's, keramiek, meubels, etc. Ook in dit onderzoek waren de resultaten veelbelovend. (Mensink & van Gemert, 2014)
- **INSIGHT (2017-2020):** Hier wil men onderzoeken hoe AI kan gebruikt worden om collecties uit de cultureel-erfgoedsector van beschrijvende metadata te voorzien.

⁹<https://www.smithsonianmag.com/>

¹⁰<http://www.nasjonalmuseet.no/en/>

¹¹Voor meer info, zie: <http://caffe.berkeleyvision.org/>

¹²Iconclass is een gespecialiseerd kunsthistorisch classificatiesysteem, <https://nl.wikipedia.org/wiki/Iconclass>

¹³Bekijk bijvoorbeeld schilderijen op basis van hun motief.

¹⁴<https://www.rijksmuseum.nl/en/api>

De collecties van de Koninklijke Musea voor Schone Kunsten van België en de Koninklijke Musea voor Kunst & Geschiedenis worden als testcase gebruikt. De focus ligt op het vrijgeven van die data als open datasets. (Universiteit Antwerpen, 2017b) Uit de paper *Deep Transfer Learning for Art Classification Problems* van Sabatelli, Kestemont, Daelemans en Geurts (2018) kwamen veelbelovende resultaten naar boven voor het voorspellen van materiaal, objecttype en kunstenaar.

- **Automated Image Analysis with IIIF (2017):** Dit onderzoek is eerder praktisch gericht en werd uitgevoerd buiten de academische wereld. Het werd uitgevoerd door CogApp, een bedrijf dat software ontwikkelt voor online archieven en musea.¹⁵ Zij hebben verschillende testen gedaan met machine learning. Het meest interessant voor dit voorstel is het onderzoek waarbij drie Computer Vision API's (Google Cloud Vision, Microsoft Computer Vision en Clarifai¹⁶) getest werden om beelden te voorzien van extra tags om de doorzoekbaarheid te verbeteren. Naast de grappige resultaten (Roddis, 2018), leidde dit tot goede resultaten waarmee de beeldencollectie op een interessante manier doorzocht kan worden: Je kan er een selectie van beelden verkrijgen door te filteren op verschillende tags, bv. *ik wil een kunstwerk uit de Renaissance van iemand met een snor en een cape*.¹⁷ De onderzoekers concludeerden dat de CVA accurate beschrijvingen geven en dat Computer Vision steeds eenvoudiger in gebruik wordt. De foutieve beschrijvingen, zoals in Figuur 2, zijn vermoedelijk een gevolg van het trainen van CVA met hedendaagse beelden die gemaakt werden met een smartphone. (Hindle, 2017)

A.2.3 Computer Vision API's

Tot slot lichten we kort CVA toe. CVA, ook wel image/visual recognition API's genoemd, zijn API's die afbeeldingen automatisch kunnen taggen (metadateren), organiseren en zoeken via machine learning. Het biedt de mogelijkheid om machine learning toe te passen zonder dat je hier zelf een expert in moet zijn. Je kan er zelf modellen mee creëren om de API's te trainen naar je eigen behoefte. Zo leerde Matt Fraser Google Cloud Vision verschillende spinnen te herkennen door de dienst te trainen met honderd afbeeldingen per spin (Fraser, 2018). De bekendste diensten zijn Clarifai, IBM Visual Recognition¹⁸, Microsoft Computer Vision en Google Cloud Vision. De diensten zijn zo toegankelijk dat iedere ontwikkelaar zonder machine learning kennis modellen kan bouwen, maar dat ook, volgens TechCrunch, iedereen zonder programmeerervaring de diensten kan gebruiken om afbeeldingen te taggen en categoriseren (Lardinois, 2018). Ook uit het onderzoek van CogApp bleken de diensten eenvoudig in gebruik en precies te trainen (Hindle, 2017).

¹⁵<https://www.cogapp.com/about/>

¹⁶<https://www.clarifai.com/>

¹⁷Probeer het zelf op: <http://labs.cogapp.com/iiif-ml/>

¹⁸<https://www.ibm.com/watson/services/visual-recognition/>

A.3 Methodologie

In dit onderzoek zal bestudeerd worden of CVA helpen bij het inhoudelijk ontsluiten van erfgoedcollecties. Kunnen de ingebouwde modellen van de API's gebruikt worden, of is er nood aan (doorgedreven) training? Dit wordt onderzocht aan de hand van een prototype waarmee verschillende beelden geanalyseerd worden. Hiervoor wordt een CVA gekozen die het toelaat om eigen modellen te trainen via een webinterface, voorzien is van goede documentatie en tutorials en beschikt over API clients in een programmeertaal die gekend is door de onderzoeker (o.a. Java, Javascript, C#). Achtereenvolgens wordt een vergelijking gemaakt tussen een set beelden die manueel beschreven werden, een set beelden die beschreven werden door de ongetrainde CVA en een set beelden die beschreven werden door een getrainde CVA.

We gebruiken hiervoor beelden van het Huis van Alijn. Het Huis van Alijn is het museum van het dagelijkse leven¹⁹ Het museum beschikt over een grote collectie beelden die het dagelijkse leven uit de 20e eeuw documenteren. In 2011 organiseerde het museum een crowdsourcingproject waarbij foto's uit de collectie 'Anonieme snapshots' getagged konden worden om ze te beschrijven en beter toegankelijk te maken (Wiericx, 2011). Het zou interessant zijn om deze beelden te vergelijken met de resultaten van de CVA. Collectiemedewerkers van Huis van Alijn zullen mee de resultaten van de proefopstellingen analyseren en bepalen welke set van beelden voor hen over de meest bruikbare tags of metadata beschikken: de manueel beschreven beelden, de beelden getagged door de ongetrainde CVA of de beelden getagged door de getrainde CVA.

A.4 Verwachte resultaten

Van dit onderzoek worden drie resultaten verwachten.

1. een methodologie die door musea gebruikt kan worden om CVA in te zetten voor collectiebeschrijving. In een rapport zal beschreven worden hoe een museum CVA kan gebruiken en hoe modellen getraind kunnen worden via de webinterface en API clients. Als de CVA zo eenvoudig in gebruik zijn (Lardinois, 2018), dan zouden collectieregistratoren zelf in staat moeten zijn om de software te trainen om hen bij te staan in het ontsluiten van de collectie. Het rapport van de methodiek zal zo opgesteld worden dat de musea er zelf mee aan de slag kunnen.
2. concrete use cases die aansluiten bij de registratiepraktijk en behoeftes van een museum.
3. musea worden bewust gemaakt over het bestaan van computer vision en hoe die technologie kan bijdragen in hun werking.

Afgaand op de resultaten uit reeds eerder gevoerd onderzoek, vermoeden we dat de CVA goed scoren op het herkennen van materiaal, het type object en de voorgestelde objecten en figuren op het object. De kwaliteit van de resultaten zal afhankelijk zijn van de specificiteit

¹⁹<http://www.huisvanalijn.be>

die verwacht wordt. Hoe specifiek de resultaten van de CVA moeten zijn, hoe meer de modellen getraind moeten worden.

A.5 Verwachte conclusies

CVA kan de metadata van een erfgoedobject verrijken met inhoudelijke informatie die nu niet voorkomt in de basisregistratie, zoals de sfeer en het gevoel dat op een kunstwerk weergegeven wordt en de gebruikte kleuren. Mogelijk is het zelfs in staat om bekende personen op de afbeelding te herkennen, zoals bij Hindle (2017). Deze informatie wordt nu niet opgenomen in collectiebeheersystemen, maar is wel interessant voor onderzoekers en het publiek. We verwachten dat de meerwaarde van de CVA vooral ligt in het bijstaan van de registrator bij het inhoudelijk beschrijven. Voor collecties van het museum die momenteel niet beschreven worden door een gebrek aan personeel, tijd of budget kunnen CVA een eenvoudige manier om deze collecties toch van inhoudelijke metadata te voorzien en ze doorzoekbaar te maken.

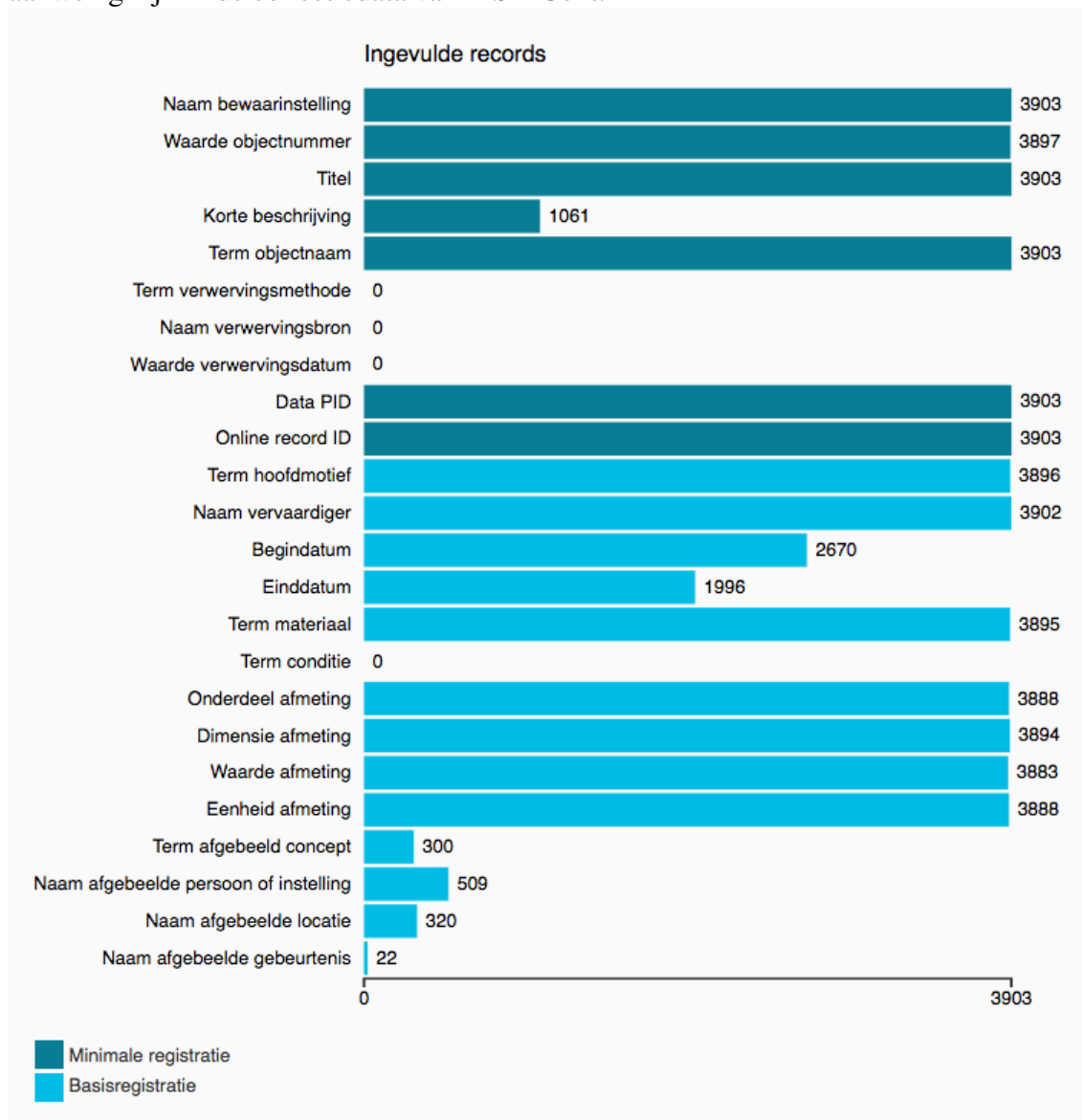
CVA kunnen een hulp zijn bij het realiseren van een inhaalbeweging voor registratieachterstand. Er wordt verwacht dat ze de registrator kunnen ondersteunen in hun taak en informatie kunnen bezorgen die zowel voor het museum, de onderzoeker als de erfgoedgeïnteresseerde interessant is. De CVA zal de registrator niet vervangen, maar kan een meerwaarde bieden aan diens werk en de registrator bijstaan bij het beschrijven van de collectie. In dit onderzoek willen we daarom aantonen wat de meerwaarde van CVA voor collectieregistratoren zijn en dat de technologie een hulpmiddel kan zijn voor collecties die niet beschreven kunnen worden.

Deze technologie zal meer en meer ingebouwd worden in DAM-systemen²⁰. DAM-software met Computer Vision is al op de markt: Gezichtsherkenning is ingebouwd in verschillende DAM-systemen.²¹ In dit onderzoek willen we aantonen voor welke use cases deze software gebruikt kan worden in musea.

²⁰Digital Asset Management Systemen, een systeem voor het opslaan, beheren en distribueren van digitale assets, zoals afbeeldingen.

²¹Zoals bij ResourceSpace: <https://www.resourcespace.com/knowledge-base/user/facial-recognition>

Figuur A.1: Het overzicht van de velden van de basisregistratie en het aantal keer dat ze aanwezig zijn in de collectiedata van MSK Gent.



Figuur A.2: Vrouw en man met boek worden door Microsoft Computer Vision geïdentificeerd als man en vrouw die een selfie nemen (Roddiss, 2018).



A man and a woman taking a selfie

Bibliografie

- Agentschap Kunsten en Erfgoed (Red.). (2014). *Handleiding bij het Cultureel-Erfgoeddecreet: Het Kwaliteitslabel*.
- Blessings, A. & Wen, K. (2013). *Using Machine Learning for Identification of Art Paintings*. Stanford University.
- Brownlee, J. (2019a, maart 19). A Gentle Introduction to Computer Vision. Verkregen van <https://machinelearningmastery.com/what-is-computer-vision/>
- Brownlee, J. (2019b, mei 1). A Gentle Introduction to the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Verkregen van <https://machinelearningmastery.com/introduction-to-the-imagenet-large-scale-visual-recognition-challenge-ilsvrc/>
- Departement Cultuur, Jeugd en Media. (2018, juli 12). Inhaalbeweging voor digitale collectieregistratie. Verkregen 8 december 2018, van <http://www.kunstenenerfgoed.be/nl/nieuws/inhaalbeweging-voor-digitale-collectieregistratie>
- Derveaux, A. (2019). Beeldherkenning in de registratiepraktijk. Verkregen van https://www.fotomuseum.be/collectie/onderzoek0/projecten/Beeldherkenning_in_de_registratiepraktijk.html
- Dickson, A. (2018, juli 20). A.I. and the Art of Spotting Fakes. Verkregen van <https://medium.com/s/story/a-i-and-the-art-of-spotting-fakes-6a674b0bdfef>
- Elgammal, A. (2017, november 16). Picasso, Matisse, or a Fake? A.I. for Attribution and Authentication of Art at the Stroke Level. Verkregen van https://medium.com/@ahmed_elgammal/picasso-matisse-or-a-fake-a-i-for-attribution-and-authentication-of-art-at-the-stroke-level-f4ec329c8c26
- Elgammal, A., Mazzone, M., Liu, B., Kim, D. & Elhoseiny, M. (2018). *The Shape of Art History in the Eyes of the Machine*. Rutgers University en College of Charleston.

- Fraser, M. (2018, maart 14). Using Google Cloud AutoML to classify poisonous Australian spiders. Verkregen 9 december 2018, van <https://shinesolutions.com/2018/03/14/using-google-cloud-automl-vision-to-classify-poisonous-australian-spiders/>
- Gatz, S. (2016, maart). Conceptnota aan de Vlaamse Regering. Naar een duurzame cultureel-erfgoedwerking in Vlaanderen. Een langetermijnvisie voor cultureel erfgoed en cultureel-erfgoedwerking in Vlaanderen. Vlaamse Regering.
- Haskiya, D. (2019, april 9). How to set up a generous interface prototype in less than a day. Verkregen van <https://pro.europeana.eu/post/how-to-set-up-a-generous-interface-prototype-in-less-than-a-day>
- Hindle, A. (2017, juni 20). Automated image analysis with IIIF: Using Artificial Intelligence for bulk image analysis. Verkregen 8 december 2018, van <https://blog.cogapp.com/automated-image-analysis-with-iiif-6594ff5b2b32>
- ING. (2016). The Next Rembrandt. Verkregen van <https://www.nextrembrandt.com/>
- Lardinois, F. (2018, januari 17). Google's AutoML lets you train custom machine learning models without having to code. Verkregen van <https://techcrunch.com/2018/01/17/googles-automl-lets-you-train-custom-machine-learning-models-without-having-to-code>
- Mensink, T. E. J. & van Gemert, J. C. (2014). The Rijksmuseum Challenge: Museum-Centered Visual Recognition. In *ACM International Conference on Multimedia Retrieval*. Verkregen van <https://ivi.fnwi.uva.nl/isis/publications/2014/MensinkICMIR2014>
- MOMA. (2018). Identifying art through machine learning: A project with Google Arts & Culture Lab. Verkregen 8 december 2018, van <https://www.moma.org/calendar/exhibitions/history/identifying-art>
- Moriarty, A. (2018, juli 23). AI and Museum Collections. Verkregen van <https://medium.com/@adamrmor/ai-and-museum-collections-c74bdb724c07>
- Nasjonalmuseet. (2017). Project: "Principal Components". Verkregen 8 december 2018, van http://www.nasjonalmuseet.no/en/collections_and_research/collection_management/digital_collection_management/Project%3A+%C2%ABPrincipal+Components%C2%BB.b7C_wJjU4L.ips
- Oberoi, G. (2016, juli 11). Comparing the Top Five Computer Vision APIs. Verkregen 9 december 2018, van <https://goberoi.com/comparing-the-top-five-computer-vision-apis-98e3e3d7c647>
- Pim, H. (2017, oktober 17). Exploring Wellcome Collection with computer vision. Verkregen van <https://stacks.wellcomecollection.org/exploring-wellcome-collection-with-computer-vision-7513dff8126d>
- Pokharna, H. (2016, juli 26). For Dummies - The Introduction to Neural Networks we all need ! (Part 1). *TechnologyMadeEasy*. Verkregen van <https://medium.com/technologymadeeasy/for-dummies-the-introduction-to-neural-networks-we-all-need-c50f6012d5eb>
- Roddis, T. (2018, mei 16). When automated analysis goes wrong. Verkregen 8 december 2018, van <https://www.slideshare.net/Europeana/when-automated-analysis-goes-wrong-by-tristan-rod-dis-europeana-tech-conference-2018>
- Rowe, P. (2017, oktober 31). Looking at Sarjeant Gallery's collection through robot eyes. Verkregen van https://medium.com/@armchair_caver/looking-at-sarjeant-gallery-s-collection-through-robot-eyes-c7fd0281814e

- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., . . . Fei-Fie, L. (2014). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision - September 2014*.
- Sabatelli, M., Kestemont, M., Daelemans, W. & Geurts, P. (2018). *Deep Transfer Learning for Art Classification Problems*. Universiteit Antwerpen en Université de Liège.
- Sabatelli, M., Kestemont, M., Daelemans, W. & Geurts, P. (2018). *Deep Transfer Learning for Art Classification Problems*. Université de Liège en Universiteit Antwerpen.
- Smith, R. P. (2017, november 3). How Artificial Intelligence Could Revolutionize Archival Museum Research. Verkregen 8 december 2018, van <https://www.smithsonianmag.com/smithsonian-institution/how-artificial-intelligence-could-revolutionize-museum-research-180967065/>
- INSIGHT. (2017a). Verkregen 8 december 2018, van <http://uahost.uantwerpen.be/insight/index.php/about/>
- INSIGHT. (2017b). Verkregen 8 december 2018, van <http://uahost.uantwerpen.be/insight/index.php/about/>
- Vanstappen, H. (2019). *VR4CH: Visual recognition voor erfgoedcollecties. Eindrapport*. Datable en MOMU.
- Westvang, E. (2017). Principal Components: Machine learning in search of the uncanny. Verkregen 8 december 2018, van <http://bengler.no/principalcomponents>
- Wiericx, B. (2011, augustus 9). Crowdsourcing in het Huis van Alijn. Verkregen van <https://faro.be/blogs/bram-wiericx/crowdsourcing-in-het-huis-van-alijn>
- Wikipedia contributors. (2019a). Computer vision — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Computer_vision&oldid=896283892. [Online; accessed 15-June-2019].
- Wikipedia contributors. (2019b). ImageNet — Wikipedia, The Free Encyclopedia. <https://en.wikipedia.org/w/index.php?title=ImageNet&oldid=900080629>. [Online; accessed 15-June-2019].