

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

Onderzoeksvorstel Bachelorproef 2018-2019

Nastasia Vanderperren¹

Samenvatting

Om de registratieachterstand bij collectiebeherende instellingen in Vlaanderen weg te werken, dient een inhaalbeweging gemaakt te worden. We stellen in dit onderzoek voor om hiervoor Computer Vision API's in te schakelen. Via een prototype zal nagegaan worden hoe Computer Vision API's kunnen bijdragen aan de beschrijving van cultureel-erfgoedcollecties. Met dit onderzoek wordt een methodiek geambieerd waarmee musea aan de slag kunnen om hun registratieachterstand weg te werken. We willen weten of deze software in de toekomst gebruikt kan worden om het werk van collectiemedewerkers te verlichten.

Sleutelwoorden

Onderzoeksdomein. Machine learning en kunstmatige intelligentie — Computer Vision — Digital humanities

Co-promotor

Henk Vanstappen² (Datable)

Contact: ¹ nastasia.vanderperren.v8632@student.hogent.be; ² henk@datable.be;

Inhoudsopgave

1	Introductie	1
2	Stand van zaken	2
2.1	Computer Vision en (cultureel) erfgoed	2
2.2	Theoretisch onderzoek	2
2.3	Computer Vision API's	3
3	Methodologie	3
4	Verwachte resultaten	4
5	Verwachte conclusies	4
	Referenties	4

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 subsidielij n opgestart om de digitale collectieregistratie weg te werken. Deze subsidielij n 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 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

¹ <https://cloud.google.com/vision/>

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

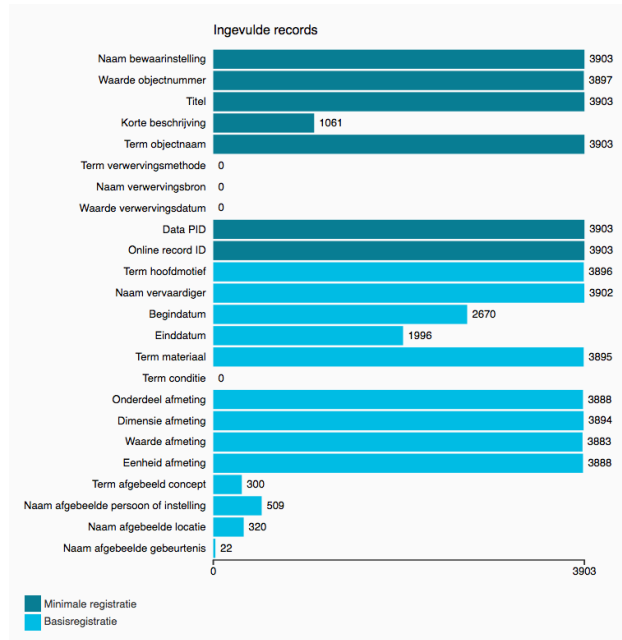
³ <https://dashboard.vlaamsekunstcollectie.be>

⁴ <http://vlaamsekunstcollectie.be/>

⁵ De 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

informatie, zoals afgebeelde persoon of afgebeeld concept, die interessant is voor ontsluiting en onderzoek, ontbreekt hoofdzakelijk.

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



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).

2. Stand van zaken

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

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 compositionele gelijkenissen tussen kunstwerken en werden ze geclassificeerd op basis van de Iconclass-terminen¹². 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)

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 veelbelo-

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

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>

vend. (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. De collecties van de Koninklijke Musea voor Schone Kunsten van België en de Koninklijke Musea voor Kunst & Geschiedenis worden als test-case gebruikt. De focus ligt op het vrijgeven van die data als open datasets. (Universiteit Antwerpen, 2017) 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)

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 iedereen zonder programmeerervaring de diensten kan gebruiken om afbeeldingen te taggen en categoriseren (Lardi-

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



nois, 2018). Ook uit het onderzoek van CogApp bleken de diensten eenvoudig in gebruik en precies te trainen (Hindle, 2017).

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.

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/>

¹⁹<http://www.huisvanalijn.be>

4. Verwachte resultaten

Van dit onderzoek worden drie resultaten verwacht.

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

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.

Referenties

- 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.
- 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>
- 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.
- 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>
- Lardinois, F. (2018). Google's AutoML lets you train custom machine learning models without having to code. Verkregen 9 december 2018, van <https://techcrunch.com/2018/01/17/googles-automl-lets-you-train-custom-machine-learning-models-without-having-to-code>
- Mensink, T. & van Gemert, J. (2014). *The Rijksmuseum Challenge: Museum-Centered Visual Recognition*. ISLA Lab - Informatics Institute University of Amsterdam.
- 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>
- 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
- Roddiss, 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-roddiss-europeanatech-conference-2018>

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

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

- Sabatelli, M., Kestemont, M., Daelemans, W. & Geurts, P. (2018). *Deep Transfer Learning for Art Classification Problems*. Universiteit Antwerpen en Université de Liège.
- 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. (2017). Verkregen 8 december 2018, van <http://uahost.uantwerpen.be/insight/index.php/about/>
- 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>