



**Universiteit  
Leiden**  
The Netherlands

## **Machine learning and computer vision for urban drainage inspections**

Meijer, D.W.J.

### **Citation**

Meijer, D. W. J. (2023, November 7). *Machine learning and computer vision for urban drainage inspections*. Retrieved from <https://hdl.handle.net/1887/3656056>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3656056>

**Note:** To cite this publication please use the final published version (if applicable).

# NEDERLANDSE SAMENVATTING

Rioolbuizen vormen een essentiële infrastructuur in de moderne samenleving, het goed functioneren ervan is van belang voor de volksgezondheid. Rioolbuizen worden met regelmaat geïnspecteerd op defecten om ze zoveel mogelijk operationeel te houden. Zulke inspecties maken het mogelijk onderhoud in te plannen in plaats van reparaties uit te voeren als het probleem kritiek geworden is. Op deze manier kunnen de gebruikers van een riool rekening houden met de onderbreking; daarnaast is het hierdoor minder waarschijnlijk dat problemen langdurig onopgemerkt blijven.

Rioolinspecties worden over het algemeen visueel uitgevoerd met behulp van een “pipe inspection gadget”, afgekort PIG. De PIG is een op afstand bestuurd voertuig met camera’s en mogelijk andere sensoren. De PIG wordt door een rioolput naar beneden gebracht om een deel van de riool te inspecteren, waarna het weer omhoog gebracht wordt. Een speciaal getrainde inspecteur beoordeelt de camerabeelden, vaak tegelijkertijd met het besturen van de PIG vanuit een voertuig op straatniveau.

Inspectierapportages worden gedaan volgens een Europese classificatienorm. Deze norm groepeerde defecten in typen en heeft richtlijnen voor wanneer gradaties van 1 (“geen interventie nodig”) tot 5 (“onmiddellijke interventie noodzakelijk”) aan de orde zijn. Een probleem is echter dat deze richtlijnen geen rekening houden met externe factoren. Een scheur in de rioolwand bijvoorbeeld, krijgt volgens de richtlijnen een gradatie van 1 tot 5 afhankelijk van de afmetingen van de scheur. De werkelijke gevolgen van een scheur hangen af van veel andere factoren, zoals of de riool zich boven of onder het grondwaterpeil bevindt, of de omgeving een woonwijk of industriewijk is, etc. Inspecteurs waarderen de gradatie daarom vaak niet volgens de richtlijnen, maar naar hun intuïtieve inschatting van de ernst. Dit heeft als gevolg dat de gradatie die een defect in een rapportage ontvangt veel kan verschillen tussen verschillende inspecteurs, en zelfs tussen verschillende rapportages van dezelfde inspecteur

Dit maakt rioolinspecties aantrekkelijk om te automatiseren. Automatisering belooft in het algemeen een potentiële verbetering in kwaliteit en efficiëntie; in dit geval zou het ook de problematische variabiliteit van de rapportages verminderen. Naast de redenen voor automatisering zijn de mogelijkheden voor automatisering ook aantrekkelijk: er is veel visuele data verzameld in de loop van decennia die gebruikt kan worden om algoritmen te trainen.

Dit proefschrift beschrijft het resultaat van vijf jaar onderzoek naar mogelijke automatisering van rioolinspecties met behulp van *machine learning* en *computer vision* technieken. Drie verschillende maar complementaire aanpakken van automatisering van rioolinspecties worden behandeld.

Hoofdstuk 3, *Image-Based Unsupervised Anomaly Detection*, beschrijft een aanpak met als kern het detecteren van afwijkingen in afbeeldingen. In deze fase van het onderzoek bestond de voor ons beschikbare data uit afbeeldingen van inspecties uit twee Nederlandse gemeenten. De rapportages over de inspecties waren niet beschikbaar, wat betekende dat het onduidelijk was welke afbeeldingen wel en geen defecten lieten zien. Hoewel meer data in een later stadium beschikbaar zou worden, hebben we besloten de afbeeldingen die we hadden toch te gebruiken.

De structuur van de verschillende afbeeldingen is soortgelijk: de buizen waren gefotografeerd met dezelfde apparatuur, en buizen uit een gemeente worden veelal geïnstalleerd in hetzelfde jaar, zijn afkomstig van dezelfde fabrikant, en worden gebruikt onder grotendeels dezelfde omstandigheden. Dit leidde tot een tweedeling van de afbeeldingen, één set van afbeelding met buizen van glad beton, één set van afbeeldingen van buizen van ruwer granulaat. Binnen een van deze twee sets zien de afbeeldingen er grotendeels hetzelfde uit, waardoor afwijkingen—zowel verwachte (zoals aansluitingen) als onverwachte (zoals defecten)—opvallen.

We gebruiken *principal component analysis* (factoranalyse) op zowel de afbeeldingen zelf als op geëxtraheerde kenmerken uit de afbeeldingen, om gemeenschappelijke factoren te herkennen. Wanneer we een afbeelding uitdrukken in de meest voorkomende van deze gemeenschappelijke factoren, blijft een afbeelding zonder afwijking getrouw aan het origineel, terwijl een afbeelding met afwijkingen minder getrouw aan het origineel zal zijn. Gebruikmakend van dit feit vergelijken we een gereconstrueerde afbeelding met het origineel om de waarschijnlijkheid dat het origineel een afwijking bevat in te schatten.

Daarnaast hebben we ook een convolutionele autoencoder, een type neurale netwerk, getraind om een soortgelijke functie als de factoranalyse uit te voeren, zonder de beperking van factoranalyse dat de factoren een lineaire relatie beschrijven.

De resultaten van deze experimenten waren veelbelovend voor de set met afbeeldingen van glad beton, maar minder succesvol voor de set met afbeeldingen van ruwer granulaat.

Hoofdstuk 4, *Convolutional Neural Network Classification*, beschrijft een aanpak gebaseerd op *supervised learning* (“leren onder toezicht”) met een convolutioneel neurale netwerk.

Convolutionele neurale netwerken zijn neurale netwerken die bijzonder geschikt zijn om afbeeldingen, geluid, en video te classificeren. We waren voorzien van afbeeldingen van rioolbuizen zoals die in hoofdstuk 3, maar in groter volume, en ditmaal inclusief de rapportages zoals toegekend door de menselijke inspecteurs, in een formaat geschikt voor digitale verwerking. In totaal waren 2,2 miljoen afbeeldingen beschikbaar en de bijbehorende classificaties lieten ons inschatten welk type defecten zichtbaar zouden moeten

zijn in elk van deze afbeeldingen. Een enkel neurale netwerk werd getraind om de twaalf meest voorkomende defecten te detecteren.

Het detecteren van defecten in riolen is een ‘ongebalanceerd’ probleem: slechts ongeveer 1 % van de afbeeldingen bevat werkelijk een defect. Het merendeel van de bestaande wetenschappelijke literatuur maakte een inschatting van de prestaties van een model door te kijken naar de *accuracy* (nauwkeurigheid), het deel van de afbeeldingen dat correct geïdentificeerd is op de aanwezigheid of afwezigheid van een defect. Met een realistische dataset is het behalen van een accuracy van 99 % mogelijk door elke afbeelding te classificeren als geen defect bevattend, wat duidelijk niet de intentie van defectdetectie is. Om dit probleem te corrigeren, worden datasets veelal ‘hergebalanceerd’, zodat deze ongeveer 50 % afbeeldingen met defecten bevat. Hoewel dit niet per se een probleem oplevert, is het belangrijk dat deze herbalancering alleen op de trainingsset plaatsvindt zodat de inschatting van de prestaties niet beïnvloedt wordt, wat vaak niet het geval was in de literatuur. Veelal werden foutpositieve resultaten (onterechte detectie van een defect) en foutnegatieve resultaten (onterecht gebrek aan detectie van een defect) gelijk behandeld, terwijl deze zeer verschillende resultaten kunnen opleveren: een foutpositief resultaat kost extra tijd, maar een foutnegatief resultaat kan een gevaar voor de volksgezondheid opleveren.

Een groot deel van de bestaande literatuur verdeelde de afbeeldingen van riolen willekeurig in training en test set, wat kan betekenen dat afbeeldingen van dezelfde riool op nabije locaties zich zowel in de training als test set kunnen bevinden. Dit introduceert een gevaar voor *overfitting*: een goede prestatie op de test set betekent in dit geval mogelijk niet dat de defecten worden herkend, maar dat de riolen zelf worden herkend.

Wij opperden dat om enige betekenis te hebben in de echte wereld, de test set die gebruikt wordt bij het inschatting van de prestaties van het model zo realistisch mogelijk moet zijn, betekenend onder meer dat deze een realistische verhouding van afbeeldingen met en zonder defecten heeft, en dat deze geen afbeeldingen bevat van riolen die ook in de trainingsset aanwezig waren. We hebben het probleem ook meer beschouwd vanuit een context-gevoelig perspectief dan eerder werk, en geopperd dat accuracy geen nuttige maatstaf is in realistische situaties. In plaats daarvan hebben we meer betekenisvolle maten geïntroduceerd die gemakkelijker door mensen geïnterpreteerd kunnen worden, en direct vertaald kunnen worden naar operationele impact.

Hoofdstuk 5, *Stereovision and Geometry Reconstruction*, reikt voorbij het huidige inspectieproces en onderzoekt de toegevoegde waarde van een tweede camera die ons een driedimensioneel beeld van de riool laat reconstrueren. Zoals mensen diepte kunnen zien met beide ogen open laat een tweede camera ons de posities van objecten in verhouding tot het

kijkpunt inschatten.

In samenwerking met de Technische Universiteit Eindhoven hebben wij 26 riolen in verschillende staten van gebruik gefotografeerd met twee zij-aan-zij camera's. Door te bouwen op bestaande stereovisie technieken en deze uit te breiden en aan te passen voor ons unieke doeleinde, kunnen we een drie-dimensionale *point-cloud* reconstrueren van de binnenwand van een riool.

De rioolwand kan nu worden gemodeleerd met een model, onder enkel de aanname dat de camera's correct uitgelijnd zijn en ongeveer middenin de rioldoorsnede gericht staan. Het model is complex genoeg om de afmeting van alle riolen die gefotografeerd zijn te vatten, maar gebaseerd op menselijk begrip van de vorm van een riool, waardoor het resultaat goed te interpreteren is.

Het model wordt toegepast op de point cloud om de oorspronkelijk vorm van de riool in te schatten, zonder de mogelijke gebruikssporen hierin mee te nemen. Hierdoor kunnen de gebieden waar het oppervlak van de rioolwand afwijkt van de te verwachten vorm herkend worden. De gedetecteerde afwijkingen op het oppervlak correleren met de aanwezigheid van defecten in dit kleinschalige experiment. Het eindresultaat is een interpreteerbare beeldverwerkingstechniek die gebruikt kan worden om inspecteurs te assisteren.