Data Intrapreneurship

Project Final Report

WoonVision

Florentijn Ramaekers	*2076872*
Friso Veldhorst	*2047908*
Eugen Falca	*2150472*
Mathijs van der Schoot	*2147586*
Yan Blankenvoorde	*2146090*
JM0120-M-6	*December 20, 2024*

Contents

1 Introduction

With ever increasing maintenance costs for social housing, social housing corporations (SHCs) are looking at different ways to cut costs. To help in this endeavor, Kleurrijk Wonen, one of those SHCs, has asked our team to look at different ways to do this and potentially look at implementing them.

The team consists of 5 first-years JADS students with backgrounds ranging from data science to computer science and from artificial intelligence to business & entrepreneurship, from different universities. This diversity in backgrounds allows us to have multiple different points of view from which to tackle the problem.

One of the problems posed by Kleurrijk Wonen is the following. As it stands, due to the decreasing number of handymen, contracting them is more costly. Right now, contractors are sent to a house for an inspection, after which a specialist is send there to repair the damage. This is further exacerbated by the aging Dutch labor force and the declining interest in technical education. In van der Vliet's words "these factors necessitate a reassessment of our working methods and an exploration of how modern technology can help us maintain high service levels for our tenants" (Van der Vliet, n.d.). Replacing the contractor inspection with drones paired with damage recognition software could cut out the contractor, allowing for a potential reduction in costs.

In the first phase of this project, we explored multiple avenues that could reduce the costs of Kleurrijk Wonen. In particular, we kept in mind the fact that Kleurrijk Wonen saw potential in drones and so mostly stayed within that realm. We did this by applying the Market Opportunity Navigator Framework (MON) to figure out problems within the market that need a data-driven solution. Using these MON's, we looked at the challenge versus the potential that the opportunity had. The challenge consists of what could be done within the scope of Data Intrapreneurship in Action, looking at the timeframe and skill level of the team. On the other hand, the potential was considered by looking at the potential value our solution has and the viability for it to be turned in a business, which will be further discussed in the part Business Context.

2 Market Opportunity Navigators

2.1 AI-powered Maintenance Scheduling Application

This application aims to reduce unnecessary or premature contractor visits by using AI to analyse photos of maintenance issues submitted by tenants.

Compelling reason to buy	High	Implementation Obstacles	Low
Market Volume	Mid	Time To Revenue	High
Economic Viability	Mid	External Risks	Mid
Potential	Mid	**Challenge**	Mid

Table 1: Attractiveness Evaluation of Market Opportunity 1

2.1.1 Potential

The earlier mentioned issues of long travel times and an aging labor force drive the need for modern technological solutions to maintain service levels. While the social housing market in the Netherlands has shown limited growth of 1.8% over ten years, the high demand (with waiting times exceeding seven years in some municipalities (De Jong & De Jong, 2021)) highlights a pressing need for innovation. However, limited profitability in the sector (approximately €1,500 per building (NOS, 2018)) and rising maintenance costs pose constraints.

Despite these challenges, the potential for this solution remains high due to the urgent need for cost-efficient alternatives. However, the exact development costs and margins are uncertain, primarily due to a lack of data and rising construction prices. Overall, the high need for modernization underscores the opportunity, though uncertainties around financial feasibility temper its potential rating.

2.1.2 Challenges

The implementation challenges for this application are minimal, as the development can be handled by a small student team with the necessary data provided by Kleurrijk Wonen. The product is designed to be easily installable on tenants' phones and employees' computers, reducing distribution obstacles and funding requirements. Development is estimated to take 4-12 months depending on whether a prebuilt or custom AI model is used, though limited time availability may extend this timeline.

However, widespread adoption will take longer due to behavioural factors and the need to achieve economies of scale, which are critical for cost optimization. The product's uniqueness and potential lock-in effect for customers reduce competitive threats and external risks Viard, 2007. Dependencies on reliable third-party infrastructure, such as cloud services and app platforms, further mitigate technical risks, making this a low-challenge opportunity overall.

2.2 Drone-based Inspection Servies

This opportunity involves the use of drones to perform inspections on hard-to-reach infrastructure, reducing reliance on manual inspections by engineers.

Compelling reason to buy	Mid	Implementation Obstacles	High
Market Volume	High	Time To Revenue	Super High
Economic Viability	Mid	External Risks	Mid
Potential	Mid	**Challenge**	High

Table 2: Attractiveness Evaluation of Market Opportunity 2

2.2.1 Potential

Drone-based inspection services present a compelling market opportunity due to their compelling benefits, such as efficiency in accessing difficult areas, cost reduction, improved safety, and the ability to conduct non-invasive inspections. Target customers include social housing organizations, infrastructure maintenance companies, public infrastructure agencies, and cultural heritage preservation entities. The advantages are supported by substantial market growth forecasts, such as the infrastructure monitoring market's CAGR of 11% between 2023 and 2030 (Research, 2023). Public infrastructure agencies also represent a significant growth area, with the global drone inspection market expected to grow from $11.6 billion in 2022 to $23.0 billion by 2027, at a CAGR of 14.6% (Markets & Markets, 2023).

2.2.2 Challenges

The implementation of drone-based inspection services faces several problems. Regulatory issues are a significant barrier, as stringent and varying regulations - such as no-fly zones, altitude restrictions, and mandatory certifications - can complicate operations, especially in urban or sensitive areas (European Union, 2019). Initial investments, including specialized software, staff training, and adjustments to existing protocols, also pose challenges, even when outsourcing services (Deloitte, 2022). Additionally, organizational resistance to change, particularly from personnel accustomed to traditional methods, may hinder adoption and require substantial change management efforts (Kotter, 1995)

Time to revenue varies depending on factors such as implementation speed and operational scale. Outsourcing drone services can provide quicker financial benefits by bypassing the need for extensive in-house preparation. However, pilot programs to mitigate risk can delay revenue generation.

External risks include the rapidly evolving regulatory environment, which could impose new restrictions or compliance costs, and advancements in competing technologies, such as robotics or AI-driven inspection methods, which could reduce demand for drones. Public perception and privacy concerns, particularly in densely populated areas, could also lead to resistance or stricter local regulations (PricewaterhouseCoopers, 2016), affecting adoption rates and market growth. Despite these obstacles, these risks are manageable and unlikely to entirely derail adoption.

2.3 Data Science Consultancy for Social Housing

This consultancy firm would leverage data science to optimize tenant selection processes, implement predictive/preventive maintenance strategies, improve revenue/cost forecasting, develop rent pricing strategy and to manage the asset lifecycle for social housing organizations.

Compelling reason to buy	High	Implementation Obstacles	High
Market Volume	Mid	Time To Revenue	Low
Economic Viability	Mid	External Risks	Mid
Potential	Mid	**Challenge**	Mid

Table 3: Attractiveness Evaluation of Market Opportunity 3

2.3.1 Potential

A consultancy firm specializing in data-driven optimization for social housing has high potential due to the pressing challenges faced by SHCs, including rising housing costs, decreased income limits, vandalism, and sustainability obligations (Teo et al., 2024). The firm's ability to provide tailored insights, particularly in maintenance optimization and tenant selection, addresses an unmet need. While many SHCs have internal business intelligence (BI) teams, this consultancy's industry-wide data and specialized expertise offer deeper insights, making its services highly valuable.

The market volume is substantial, with approximately 2.5 million social housing units in the Netherlands managed by 275 housing corporations (Aedes, 2022). While the number of clients is limited, the high-value nature of consultancy services offsets this constraint. Additionally, housing corporations' spending is projected to double from 2022 to 2026, further supporting market growth (Aedes, 2023). Although data gathering from SHCs may be challenging, the firm's expertise in data cleaning and modeling ensures that even smaller datasets can provide actionable insights. Combining anonymized data from multiple clients will enhance the accuracy of models over time.

Operational costs are low, but convincing SHCs to adopt the service may require initial low pricing strategies due to their tight budgets. While consultancy is generally profitable, limited clients and the potential for SHCs to outgrow the need for external services present long-term challenges.

2.3.2 Challenges

The challenges for a social housing consultancy primarily revolve around implementation, time to revenue, and external risks.

Establishing market presence is challenging due to the need for housing market expertise and difficulties in acquiring initial clients without existing contacts. Funding is necessary to hire experts, which adds to startup costs. Additionally, the quality and availability of client-provided data are critical for meaningful analysis, creating uncertainty, especially with the first client. Continuous refinement of methods and outperforming existing BI departments are essential but require significant effort and expertise.

While the market is currently favorable, the limited capacity to handle multiple clients and the lengthy sales cycle may delay profitability. A single project could take several months, constraining short-term revenue generation despite minimal preparation requirements for launching services.

The consultancy is vulnerable to changes in government subsidies and legislation, as the social housing market heavily relies on public funding. Although support for social housing is likely to continue due to the housing crisis, any funding reductions could impact the willingness of housing corporations to invest in consultancy services. Additionally, while there is no direct competition yet, success in the field may attract competitors over time.

2.4 Damage detection for property maintenance

This opportunity proposes using drone images and AI-based image detection models to identify and localize building facade damages, reducing the time and effort required for manual inspections. We ended up going with this MON as it seemed the most viable option for us to do and after talking with Kleurrijk Wonen they seemed to be the most enthusiastic about this MON too.

Compelling reason to buy	Mid	Implementation Obstacles	Mid
Market Volume	High	Time To Revenue	High
Economic Viability	High	External Risks	Super High
Potential	High	**Challenge**	High

Table 4: Attractiveness Evaluation of Market Opportunity 4

2.4.1 Potential

The solution addresses process optimization by automating damage detection, reducing costs, and improving efficiency compared to manual inspections. Using a model like YOLO offers speed, scalability, and cost-effectiveness (Redmon et al., 2016). However, its accuracy is dependent on data quality, and it doesn't yet solve an unmet need. Manual inspection remains adaptive and reliable but is slower, costlier, and harder to scale. While automation can significantly benefit companies like KleurrijkWonen, the solution's effectiveness is uncertain due to variability in the number of images analyzed daily.

The housing maintenance market in the Netherlands is highly attractive, valued at €36 billion in 2023, representing 4.5% of GDP (MaintWorld, 2023). Though the solution targets only a portion of this market (housing maintenance), the scale and relevance of the problem ensure a substantial market volume.

The proposed pay-per-use business model, delivered via an API, ensures recurring revenue. With the Residential Real Estate Leases market generating $57 billion in 2024 and staffing costs forming a significant expense for real estate agencies, the solution has strong appeal for cost reduction. Additionally, once a company adopts this system, switching to a competitor involves process changes, creating client stickiness.

2.4.2 Challenges

The development of the solution is not particularly difficult. The team has relevant knowledge in model development, and using a pre-trained object detection model like YOLO reduces the amount of data and time required for training (Redmon et al., 2016). However, annotating images with bounding boxes is a manual process and whilst it does not require advanced expertise, it is still a considerable task. Computational resources for model training can be obtained via cloud platforms like Google Colab, and deploying the model via an API is straightforward. The main difficulty lies in accessing the market, as it requires changes to existing maintenance processes within companies, particularly in terms of how images are collected and processed.

The model can be ready for clients within about two months, provided there is a consistent set of relevant images. A key milestone is ensuring that the model can accurately detect damages without overlooking areas that require maintenance. However, market readiness poses a significant challenge, as companies are accustomed to traditional methods, and transitioning to a new process may take several months or even years. Convincing entire teams, rather than individuals, to adopt the solution adds complexity, and the dependency on external image collection further delays the process.

There are significant external risks, including competition from companies like DroneDeploy and Flyability, which already offer end-to-end solutions for image collection and damage analysis. The quality of the images collected, which is outside the team's control, is also a critical factor. Moreover, the need to integrate the model into existing systems and processes presents substantial adoption barriers.

2.5 AI-Powered Rent Payment Prediction

This MON aims to leverage artificial intelligence to predict whether tenants will pay their rent on time. This initiative aligns with their mission to provide affordable housing by minimizing financial risks and ensuring efficient resource allocation.

Compelling reason to buy	High	Implementation Obstacles	High
Market Volume	High	Time To Revenue	High
Economic Viability	Mid	External Risks	Super High
Potential	High	**Challenge**	Super High

Table 5: Attractiveness Evaluation of Market Opportunity 5

2.5.1 Potential

Housing corporations face significant challenges in ensuring timely rent payments, which strain financial predictability and long-term planning (May, 2024). Current methods like manual assessments and basic credit checks are inefficient and reactive (So, 2022). An AI model trained on large datasets offers a proactive, accurate, and scalable solution to predict payment delays (Brynjolfsson & Mcafee, 2017), enabling preemptive actions such as early tenant interventions. This improves operational efficiency and tenant relationships, making the solution highly compelling.

The Netherlands has a substantial housing market. These corporations represent a significant segment of the market, and as housing shortages grow, the rental sector is expected to expand further, with plans to build 900,000 new homes by 2030 (van Algemene Zaken, 2023).

The AI model adds value by reducing delays in rent payments, improving financial predictability, and cutting administrative and legal costs. However, housing corporations' social mission limits their ability to pay for the model, as their rent income is capped by government regulations (van Algemene Zaken, 2024). While the solution's benefits may justify the costs over time, this financial constraint lowers economic viability.

2.5.2 Challenges

The AI model's success relies heavily on access to comprehensive and relevant data (Ideology, 2024), which may be lacking, particularly for smaller housing corporations. Many housing corporations lack the infrastructure to manage large datasets, and convincing them of the model's value could be challenging due to financial constraints and skepticism about AI's "black box" nature (Yasar & Wigmore, 2023). This means securing funding to build and distribute the model further is considerably complex.

To develop the model, most likely we would first have to collect and prepare the data, which could delay the process. Even after development, housing corporations may conduct internal testing and require training for employees, extending the timeline. Additionally, tenant associations in the Netherlands, which must legally be consulted, may delay decisions due to concerns about privacy (Bock, 2024).

The market faces intense competitive threats from tech giants, in-house IT teams of large housing corporations, and other companies that may produce more accurate and cost-effective models. There are also dependencies on third-party cloud providers, which could increase costs, and on housing corporations for data availability. Barriers to adoption, including resistance from tenant associations and integration challenges with existing software, further compound the risks.

3 First iteration - Damage Detection

After carefully considering which project to pick from our 5 ideas, we settled on Eugen's idea: AI damage detection for property maintenance. We felt that this showed the most potential, and was actually doable in the time frame that we had. Things turned out differently, but with the help of the lean startup method, we successfully pivoted to a similar but improved idea. The next sections will discuss our process in detail.

Kleurrijk Wonen has divided their maintenance activities into 3 categories: reparational, planned and mutational. Reparational is the most straightforward type of maintenance, it encompasses small repairs that occur when Kleurrijk Wonen is notified of damage by a tenant. This is quite expensive because they often first send someone to check what has to be repaired, and only afterwards send the actual repairman. Kleurrijk Wonen is looking for solutions that reduce the number of 'movements' this takes; reparational maintenance currently takes 4 movements, and they wish to halve it. The movement count is defined as the number of individual actions that must be done before the damage is fixed. Reducing the movement count is a big cost saver, and so we looked into it. A model for damage detection could reduce the movement count, by having the tenant send a picture of the damage. This would not work for all types of damage, but if we could do it for some, it would be a successful project.

A model for damage detection is also applicable to the next type of maintenance: planned. This category of maintenance encompasses all large maintenance operations such as the building roof, foundation or utilities such as gas and electricity. This type is the biggest expense for Kleurrijk Wonen and occupies about 30% of their turnover. We came up with the idea to optimize the planned maintenance through intelligent scheduling. Kleurrijk Wonen wants to bother their tenants as little as possible, so they try to plan multiple maintenance operations at the same time. For example, when maintenance for the roof is required every 7 years but only 6 have passed, you could do it now rather than in a year, because you are doing the gas pipes maintenance now anyway. This is a simplified example, and the reality is much more complicated. Our idea was to build a scheduling algorithm/model to optimize this. However, we scrapped this idea since we did not have any experience with scheduling algorithms. Furthermore, a model for damage detection can also greatly help. Detecting damages on drone images of roofs and building facades gives Kleurrijk Wonen an insight into the state of the building, and whether maintenance is required. Further insight into the building condition is always of great help, because it is information that they can use for many different purposes. The final type of maintenance is mutational maintenance, which is defined as building modernization. It is commonly done when a tenant leaves, and Kleurrijk

Wonen wants to make the house attractive again with architectural and amenities adaptations. This is an area of maintenance that can not be optimized or improved with data.

In summary, there are 2 types of maintenance that can be optimized using a damage detection model. However, the reparational maintenance could be helped with damage detection of an image taken by a tenant, while the planned maintenance could be helped with damage detection of drone images. The image types are vastly different, so creating a model to do damage detection on both would be incredibly challenging. There are certainly methods to deal with varying building types and structures (Lee et al., 2023), but none deal with the fact that many of our images are taken from a whole different perspective. Therefore, we decided to focus on damage detection from drone images. This would keep the project more feasible, with regards to our expertise and time constraint.

3.1 Value Proposition

3.1.1 Customer jobs

Companies that manage properties need efficient, accurate and cost effective solutions for maintaining their buildings. Their task is to assess damages, planning maintenance and ensuring that the tenants are happy. They aim to reduce operational inefficiencies, minimize tenant disruptions and to allocate their resources effectively whilst maintaining financial stability.

3.1.2 Pains

Some of the key challenges faced by property management companies are as follows: The primary issue is operational inefficiency, where high costs and time delays can happen due to the need for multiple site visits for maintenance. Additionally obtaining a comprehensive and up to date assessment of building conditions is difficult. This lack of insight complicates prioritization and prevents a proactive management of maintenance tasks. Other major pains include tenant disruption, where tenants might be negatively affected by frequent and poorly coordinated maintenance operations. Finally inefficient resource allocation often leads to wasted resources, escalated costs and missed opportunities for optimization.

3.1.3 Gains

Some of the key benefits that this solution will offer to property management companies includes cost reduction by reducing site visits for maintenance and optimizing the scheduling of this maintenance to reduce unnecessary expenses. This improved efficiency saves time and resources due to the faster detection of damage and smooth flow of work. Another benefits is that there will be an improvement in decision making, since access to proper and relevant data allows for better long term planning and strategic management of the properties. And finally it contributes to greater tenant satisfaction because disruptions can be minimized by the better scheduling of maintenance activities, creating trust and improving relationships between tenants and landlords.

3.1.4 Pains and Gains diagram

The pains and gains diagram below 1 illustrates the pain points such as operational inefficiencies, lack of insights, tenant disruption and resource allocation issues on one side. On the other side you can see the gains, including cost reduction, improved efficiency, enhanced decision making and tenant satisfaction. The customer job can also be seen which is assessing damage efficiently and accurately, allocating their resources effectively and ensuring a high level of tenant satisfaction.

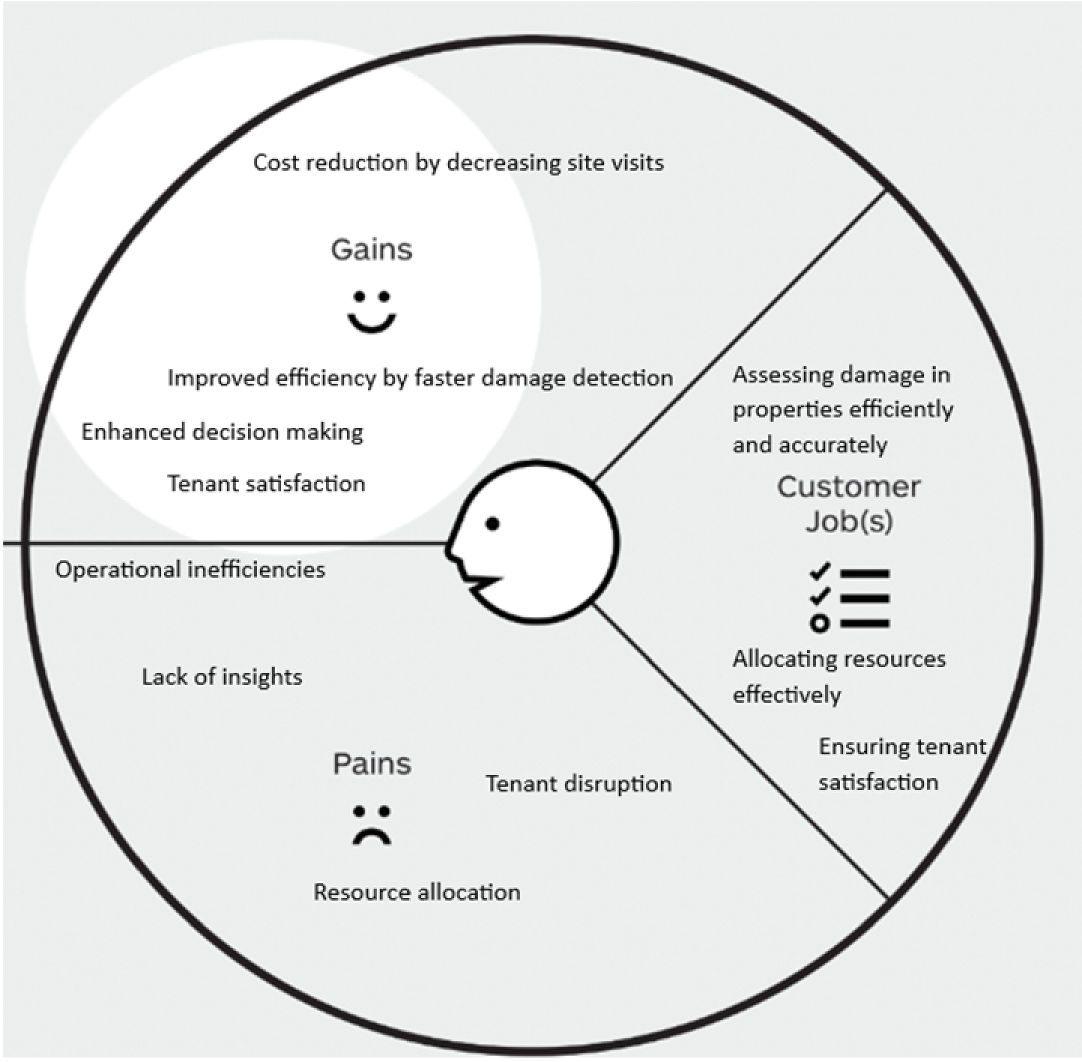

Figure 1: The pain and gain diagram.

3.1.5 Products and services

Our product entails an AI powered model for damage detection specifically trained in processing drone images. Such a model will offer property management firms substantial capacity to detect any underlying damages with much increased efficiency. In particular to assure accessibility and ease of use the solution is integrated through two main channels: an API for seamless batch processing of images and a user-friendly web portal where smaller companies can easily upload and analyze their data. Moreover we will continuously support organizations in the process of integrating the model into existing workflows, making the implementation smooth and effective.

3.1.6 Pain Relievers

The AI powered damage detection model will address some critical challenges that property management companies face. Through remote damage detection we remove the need for preliminary inspections which will reduce the amount of site visits and this will reduce the operational inefficiencies significantly. The drone image analysis will provide a comprehensive view of the health of the building. This will help property managers understand their assets better, which will allow them to make more proactive maintenance strategies. This will enhance the overall efficiency and effectiveness of the maintenance.

3.1.7 Gain Creators

The AI powered model of damage detection creates huge benefits for real estate management firms. Costs are drastically reduced because many site visits will no longer be needed, and overall resources are optimally deployed. This will allow better budget utilization by the companies. Along with that, the operational efficiency related to quicker identification of damages, thus maintaining a smooth work cycle, also improves the overall productivity. The solution also provides access to reliable and detailed building data, enabling property managers to make informed, strategic decisions for long term planning and asset management. Additionally, the model minimizes disruption to tenants by facilitating better-coordinated maintenance schedules, ensuring greater comfort and satisfaction of tenants.

3.1.8 Gain Creators and Pain Relievers diagram

The diagram below 2 represents the relationship between the pains, such as operational inefficiencies and tenant disruption, and the gains, such as cost reduction and enhanced decision-making. It highlights how addressing these pains directly leads to the realization of these gains.

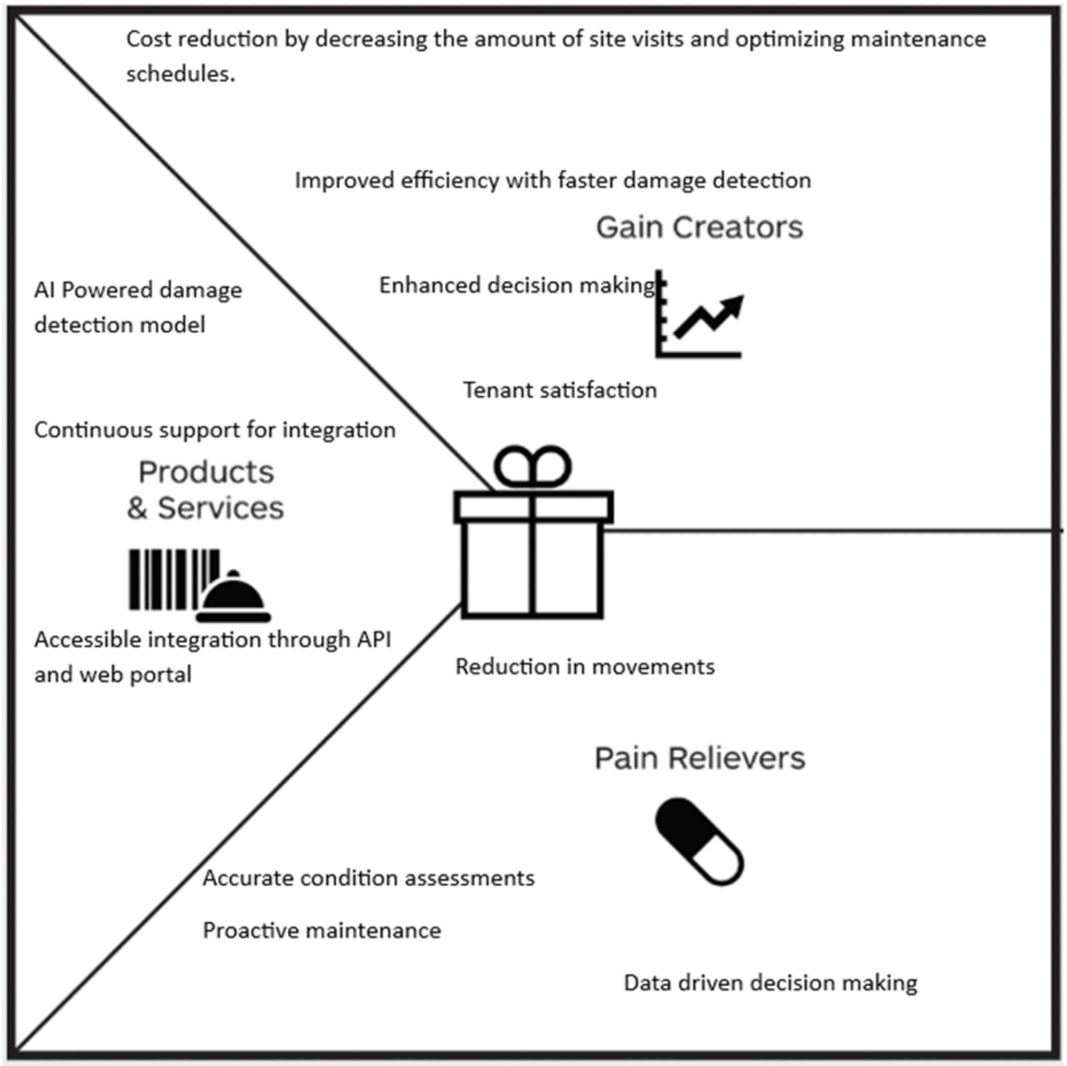

Figure 2: The gain creators and pain relievers diagram.

3.2 Business Model Canvas

Figure 3 gives the business model canvas for this idea. The business model canvas is a compact visual representation of the business model, proposed by Oswalder in his PhD Thesis (2004). We can use it to sum up and evaluate the business model.

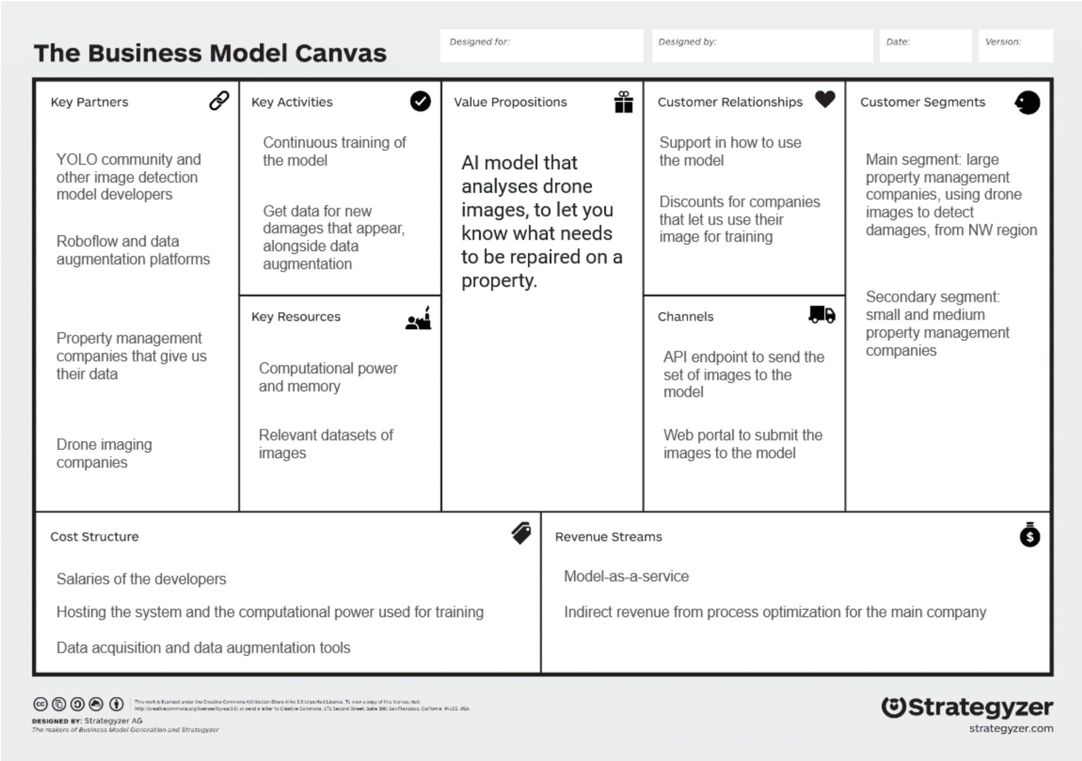

Figure 3: The business model canvas, for a drone image damage detection model.

This canvas describes the key components of a potential company that would provide the damage detection product. Starting with the Key Partners. We aim to collaborate with the YOLO community and other developers of image detection models, using their expertise in image detection technologies. This community can also serve as a platform to hire people, because we want employees with this specific expertise. Platforms like Roboflow and other data augmentation services are also important. We can use these platforms to improve and accelerate the development cycles, because they provide data and tools for computer vision projects such as ours. Another key partnership lies with property management companies, who serve as our clients during our projects with them. Finally, drone imaging companies play a fundamental role, as they supply the high-quality aerial images of properties that form the foundation of the damage detection process.

The primary activities focus on continuous improvement and adaptation of the AI model. This includes training the model regularly with new data and implementing data augmentation techniques to simulate various damage scenarios. The model's utility depends on its ability to detect emerging types of damage, necessitating a dynamic and iterative development approach. Furthermore, buildings come in different shapes and sizes, which also necessitates continuous development.

Considering the key resources, this operation hinges on computational power and memory to train and deploy the AI model effectively. Additionally, access to relevant and comprehensive datasets of property images is indispensable. These datasets must be curated to ensure the inclusion of diverse property types and damage scenarios to maximize the model's applicability. An important note to make is that our data set increases over time. We have to add labeled data for every client, to tailor the model to their needs. If the client agrees, this labeled data will also be used for subsequent projects.

The core value proposition centers on an AI model that can analyze drone images to identify damages. This innovation can be used by property management businesses to quickly and inexpensively assess the state of their buildings. Such insights are business intelligence that can be utilized to improve maintenance- and financial planning.

To ensure adoption and customer satisfaction, we offer support to help companies integrate the model into their workflows effectively. The product is subscription-based, so this service will be accounted for in the price. Additionally, companies that contribute their image data for training will be eligible for discounts, fostering a collaborative ecosystem.
The model will be accessible through two main channels: an API endpoint for developers and larger property management firms to send batches of images, and a user-friendly web portal for smaller companies to upload and analyze images directly.

The target market consists of large property management companies in the randstad region, who utilize drone imaging extensively for maintenance operations. This section of the housing market has vast amounts of large buildings, and can therefore benefit from economies of scale. A secondary segment includes smaller property management firms that may not have this direct advantage, but can still benefit from our product.

The primary costs involve salaries for developers working on the AI model, hosting and computational resources required for model training and inference, and expenses related to data acquisition and augmentation tools. Especially the costs for computation will be large, because sophisticated AI models tend to be computationally expensive. It certainly depends on the sophistication of the model, but we obviously want to have a very well-performing and refined model, so we will have to pay for that.

The revenue will come with a model-as-a-service pricing structure, where companies pay for access to the AI model's capabilities. When we are still a start-up, providing this service will be quite difficult due to lacking data and computer vision expertise. However, we will improve as time goes on, because we gain experience along the way. We classify this as a revenue stream in the business model canvas. An accountant will not agree with the classification, but we believe the increase in skill is an incredibly valuable resource, and will significantly optimize our workflows. As proclaimed by Kleurrijk Wonen, damage detection is extremely difficult, so expertise in this area is priceless.

This Business Model Canvas demonstrates the feasibility and scalability of the proposed solution. By focusing on a well-defined target audience and leveraging strategic partnerships, we can deliver a solution that adds value across multiple dimensions. The decision to narrow the project scope to drone image analysis was instrumental in aligning our objectives with our expertise and available resources.

3.3 Development

3.3.1 Learn

The next stage of the project consists of the development of the minimal viable product (MVP), which is the first usable version of a product, meant for testing and a better understanding of the market. The methodology used for this process is the Lean Startup , which is divided into 3 main components, these being learn, build and measure (Ries, 2011). Firstly, it is focused on a short time-to-market. This not only decreases the time until the users have access to the product but also the development stage is happening while constantly thinking about solving the actual needs of the clients while collecting feedback. This customer-centric mentality prevents the creation of a product that reaches advanced improvement stages, while not presenting a good market fit. In this way, the teams can save resources and pivot to a new direction after learning more about the clients.

3.3.2 Task understanding

Damage detection is considered a challenging task since it requires specific data because damage can consist of small representations, such as cracks or spelling and a clear definition. Because of this, similar research is not using a single set of images to detect the images, but rather multiple sets from different times. The usage of multi-temporal imagery is seen as an improvement towards damage detection, presenting an improvement in accuracy of 22% over the mono-temporal ones (Duarte et al., 2019). This methodology implies the usage of an initial images dataset and of a different dataset of building facades, that are taken before and after an event that can deteriorate the building, such as an earthquake. In this way, they use a complex convolutional neural network, consisting of a combination of CNN and batch normalisation layers, with a RELU activation function.

Furthermore, we are aware of existing pre-trained models, relevant to our task. This method is called transfer learning and represents the usage of a model initially trained on a source domain, which is then fine-tuned on the target domain. Moreover, this method is expected to reduce data requirements and improve performance in low-resource scenarios (Zhuang et al, 2022). The next step was looking into pre-trained models for object detection, a popular one being You Only Look Once (YOLO). The model consists of a simplified architecture, having only one CNN that divides the image into a grid. Consequently, the model treats the task as a regression problem, predicting the bounding boxes and class probabilities directly, while performing only one evaluation (Redmon et al., 2016).

Figure 4: Caption for the first image

Figure 5: Caption for the second image

3.3.3 Data

In terms of image data, Kleurrijk Wonen provided numerous sources. Firstly, it was provided a Megafile with 3315 images and access to Spotr and Cyclomedia. The last 2 represent apps that have the purpose of processing large-scale systematic visualizations of physical environments, used for remote property inspection and monitoring, using images from various public sources. Some sources are Google Street View, where the majority of images are made from a height view level, or Google Earth for images from an aerial plan. Since Kleurrijk Wonen wants in the future to focus drone images, we decided to focus on these images, which were provided in the Megafile.

The 3315 images from the Megafile consist of a mix of pictures from a view height level and aerial perspective. The majority of the images have a 4:3 format, while all images present a high quality, with 11664x8750 pixels. The methodology that was used to analyse these images was to look over them and to understand some insights. We conclude that there are around 10 buildings presented within the dataset, but none of the images presented damaged components. Another observation regarding the dataset is the noise that is added in the picture, because of the vegetation from the front of the house, which partially or fully covers the elements of the mansion. Lastly, the images are taken from different perspectives, where the view height images are either perpendicular or oblique towards the facade of the building, while the aerial images are perpendicular or oblique towards a building, or perpendicular towards the ground. Some examples of this dataset can be seen in Figure 4 and 5.

Since the existing data did not provide examples of images for damage detection, we decided to test with a dataset from Roboflow. This dataset has 104 images, where there are presented clear cases of damage. There are 3 labelled classes, 'peel', 'slight' and 'wide' and the dataset has synthetic images that were created using data augmentation methods, such as mosaic, vertical and horizontal flipping, or rotation, meant to improve the robustness of a model trained on this dataset. 2 images of this dataset can be observed in Figure 6 and 7.

Figure 6: Damage dataset - 1

Figure 7: Damage dataset - 2

3.3.4 Build

The creation of our solution represents the development of a model that detects damage on a building facade. The initial YOLO paper describes its second and third versions, whereas we will use in the project the YOLOv8 model. The development environment is Google Colab and the training is done on a T4 GPU. The YOLOv8 provides models of different sizes, whereas we use the YOLOv8 extra large, which is a base model that was pre-trained on the Common Objects in Context dataset. YOLO framework offers the option to train the model with commands in the terminal, but in our case, we use Python and the ultralytics library to The model was trained on 100 epochs, on a batch size of 16, with an image size of 640x640, where no other flags for hyperparameters of data augmentation were changed.

3.3.5 Measure

To test the results of our model, we used the testing dataset, which consists of 8 pictures. Here, there are 28 class instances of 'peel', 18 of class 'slight' and 23 of class 'wide'. The results can be seen in the confusion matrix from Figure **??**. The results are calculated on a 50% intersection over union (IoU), which represents that the 50% of the surface of the bounding box of the prediction is overlapping the bounding box of the ground truth. From these results, we observe that the class 'peel' is detected 60% of the time, whereas for the other 2 classes, the predictions are less than 30%. Nonetheless, the model has the tendency to detect the class 'slight' and 'wide' when they are actually part of the background.

3.3.6 Why do we pivot?

After a loop of the Lean Startup methodology, we understood the assignment better, from a development and business perspective. The next step represented actually translating the acquired knowledge towards a solution that can be used on the data from Kleurrijk Wonen. For this, we needed 2 important components, a set of images from which we can see examples of damage and a set of rules from which we can define what damage is, like if there are multiple layers of importance etc. In order to use this data, it needs to be labelled. Since there are numerous tools that allow data labelling, in the case in which the dataset and the set of rules were defined, we could've done the labelling on our own.
After the second update presentation, we had a meeting with the company, to present the elements that we need, in order to proceed with this task. Unfortunately, after this meeting, we concluded that there is no existing dataset that covers our requirements. Moreover, most of the literature related to this task is focused on damage detection in situations where damage is clearly defined, like after natural disasters, and not about instances like broken windows or cracked bricks. Moreover, the development of the set of rules that defines damages would've been potentially redundant, since we had no guarantee to find the relevant data and to build around it. Because of the lack of data and the set of rules, we decided to pivot towards a similar task in object detection, relevant to the use cases of Kleurrijk Wonen. This is object detection for windows and doors, where the goal is the ability to count how many instances are present within a picture. The clear advantage of this task was not only the availability of appropriate datasets but also the aspect that it is part of a field better developed from a research perspective.

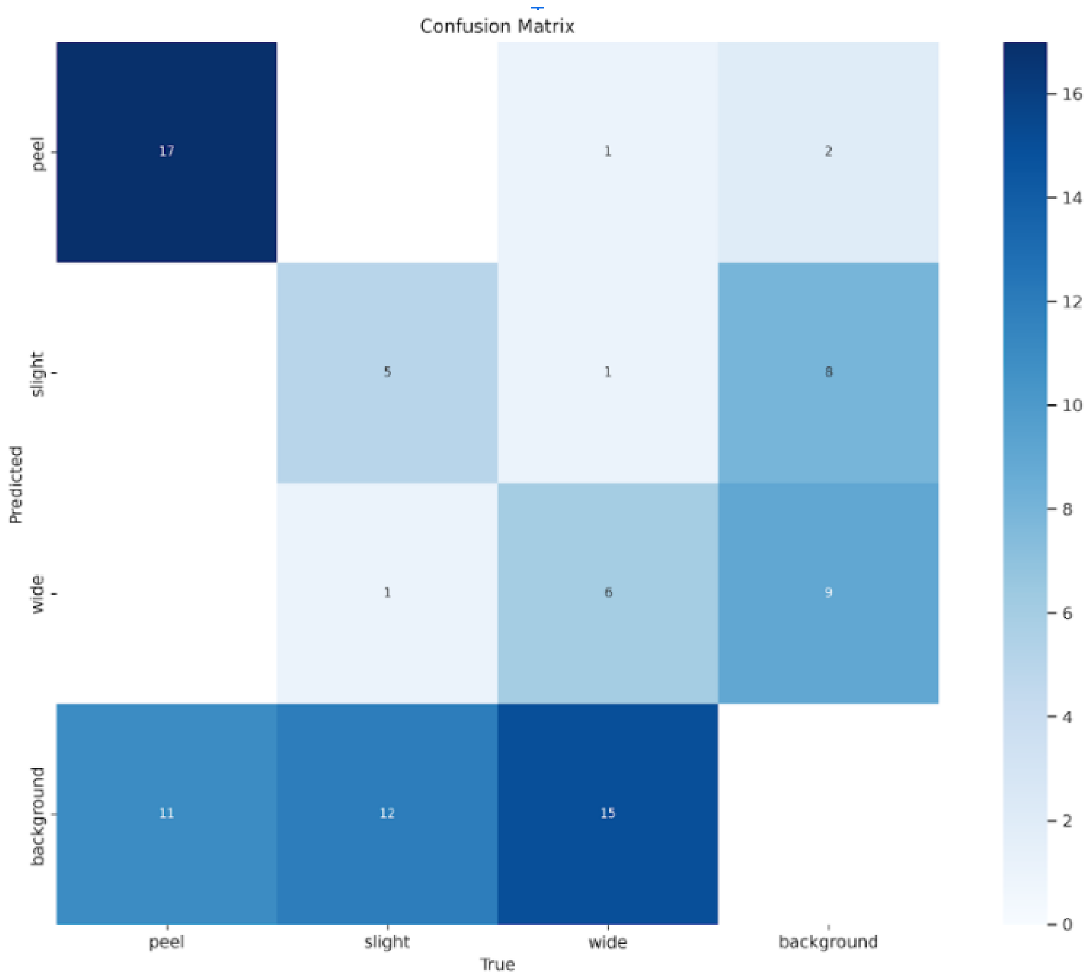

Figure 8: Confusion matrix

4 Second iteration - Object Detection

4.1 Value Proposition

4.1.1 Customer jobs

Social housing companies, municipalities, and real estate firms all manage large property portfolios. They all have to effectively and efficiently maintain these properties without disturbing the tenants. To be able to do this they need to allocate resources and make informed decisions about renovation, compliance, and future planning. Social housing companies struggle to balance customer satisfaction with their operational budget and planning.

Municipalities and city planners need tools that will help them to easily assess and monitor compliance, zoning and other factors related to urban planning. Their job is to stay informed on the state of public spaces and public buildings and make sure that all of these spaces meet regulations while balancing the needs of the people living in their jurisdiction.

Real estate firms on the other hand have a very different set of jobs since they focus on effectively marketing properties and closing deals. To be able to do this they need accurate and detailed information about all buildings in their property portfolio. Through this increased amount of information, real estate firms can be more transparent about their buildings and therefore be more appealing to new customers and build trust with their current renters. Across all three the customer jobs include: reducing manual labour, maintaining accuracy in building assessments, and responding quickly to emerging needs, whether for maintenance, compliance, or marketability.

4.1.2 Pains

Many organizations that manage extensive property portfolios face a similar set of challenges that decrease their operational efficiency. The first of these challenges is the manual nature of property assessments, these assessments are labour-intensive, time-consuming, and prone to human error, particularly when dealing with large portfolios and more so when many of the buildings are old and lack specifications. This creates inefficiencies in resource and budget allocation that lead to delays in addressing maintenance needs that will ultimately lower tenant satisfaction.

For municipalities, a lack of an up-to-date overview of building and public spaces assessment makes compliance monitoring and urban planning more reactive, which often causes civilians in the area to be dissatisfied with their municipality and local government. Real estate firms sometimes struggle with incomplete or inaccurate details about their properties which could lead to distrust among buyers and renters and therefore could possibly delay transactions and prevent sales from happening at all.

4.1.3 Gains

Customers desire tools that will help simplify property management. A major gain is the automated process of property analysis and assessment, which also provides accurate data and reduces the need for manual inspections. This will help companies manage their properties smarter by optimizing resource allocation, prioritizing the correct long-term maintenance tasks, and improving tenant satisfaction in the case of social housing corporations and real estate firms.

Municipalities want to benefit from these automated property assessments through a more proactive approach to the urban planning of public spaces. A comprehensive overview of compliance and zoning will help them to make informed decisions on design plans before inhabitants of their jurisdiction get dissatisfied with them.

Real estate firms can enhance their property listings by using automated property analysis. This will allow them to gain a competitive advantage over other companies in their competitive space. Through an increased amount of trust as a result of ultimate transparency real estate firms will be able to get more customers and make deals happen faster than ever before. For all types of companies in the customer space, the ability to integrate WoonVision into their existing systems and workflows is another key gain, this will ensure a smooth adaptation process and almost endless scalability.

4.1.4 Pains and Gains diagram

In the diagram below 9 the pains and gains of the customers are described. This diagram allows us to get a quick overview of who our customers are and what problem the customers are facing. For these problems it is also shown what there is to gain with our solution.

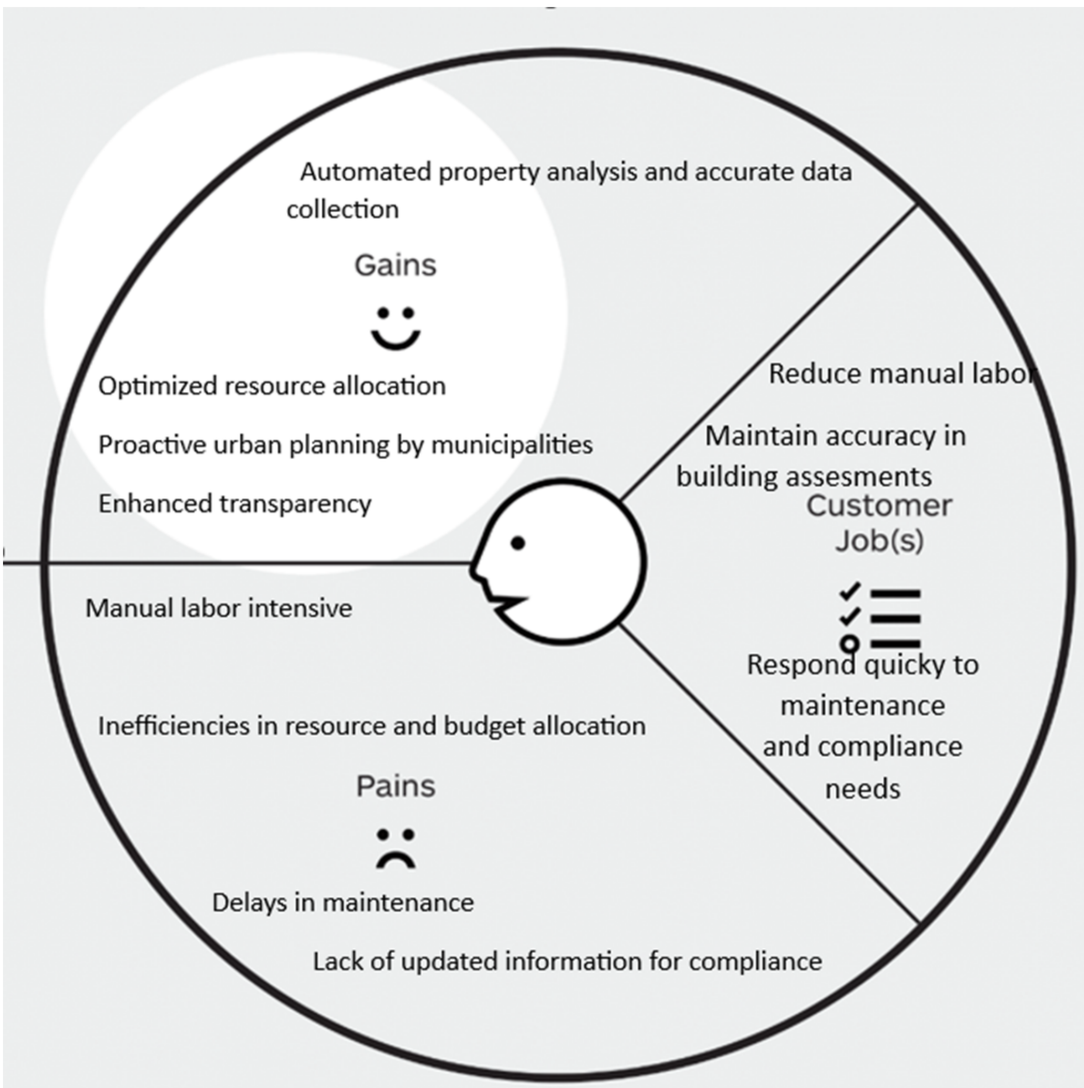

Figure 9: The pain and gain diagram.

4.1.5 Products and services

WoonVision is an AI driven computer vision platform which is intended to change how property portfolios are managed. It automatically detects and quantifies building elements like windows, doors, and balconies giving a full and accurate overview of the assets. Additionally, WoonVision offers a cloud based property management tool that fits need of the clients, from beginners in the use of technology to expert users. API and microservices integration allow for a seamless adaptation into existing workflows for advanced users to customize the software in their specific needs.

4.1.6 Pain Relievers

WoonVision covers a wide range of customer challenges. WoonVision will automate the labor intensive and error prone manual property assessments which will be reducing inefficiencies and ensuring precise results. By offering accurate and up-to-date data WoonVision minimizes errors and delays thus enabling better decision making and resource allocation. For municipalities, this includes improved compliance monitoring and proactive urban planning to meet regulatory requirements and better serve their constituents. Real estate firms have increased their transparency, building trust among buyers and tenants. Besides, the smooth integration of WoonVision with existing systems will ensure continuity with minimal disruption in the operations of users.

4.1.7 Gain Creators

WoonVision enhances decision making by providing actionable insights which will allow organizations to better plan renovations, maintenance, and resource allocation. Social housing companies and municipalities are able to improve the satisfaction of tenants and citizens simply by better addressing maintenance needs. Real estate firms gain a competitive advantage in offering transparent and detailed property information and municipalities benefit from the tool since it allows for proactive urban planning and compliance. WoonVision's user friendly interface makes property assessment easy for a wide range of users, from technical experts down to non-technical teams.

4.1.8 Gain Creators and Pain Relievers diagram

In the diagram below 10 the product and services, gain creators and pain relievers are described. This diagram allows us to get a quick overview of what solution we are creating and what kind of pain this relieves and what the customers are gaining from using our product.

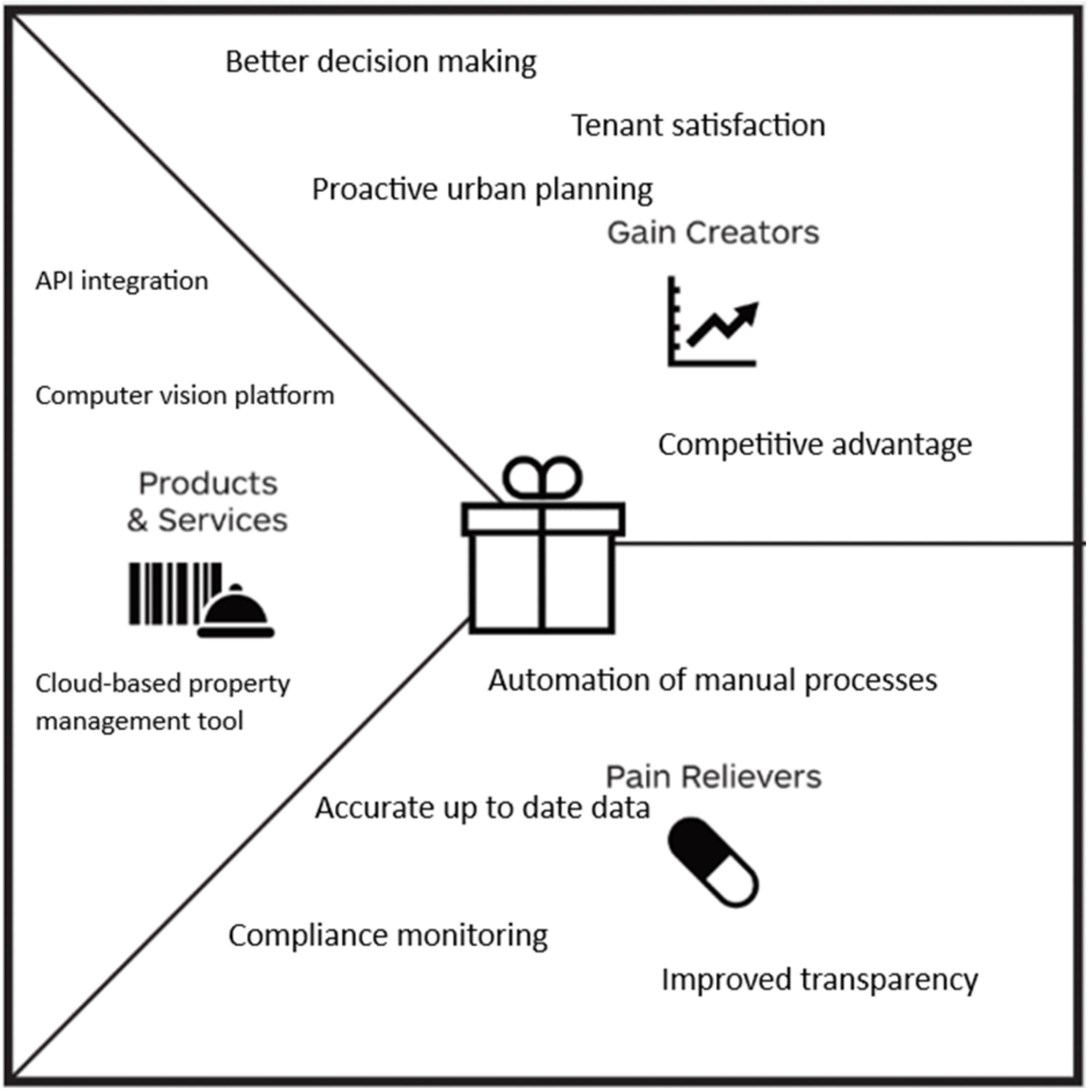

Figure 10: The gain creators and pain relievers diagram.

4.2 Business Model Canvas

In this section the business model canvas is described of the second iteration. In this iteration WoonVision will do automated identification and quantification of building elements. The filled in business model canvas can be seen in figure 11

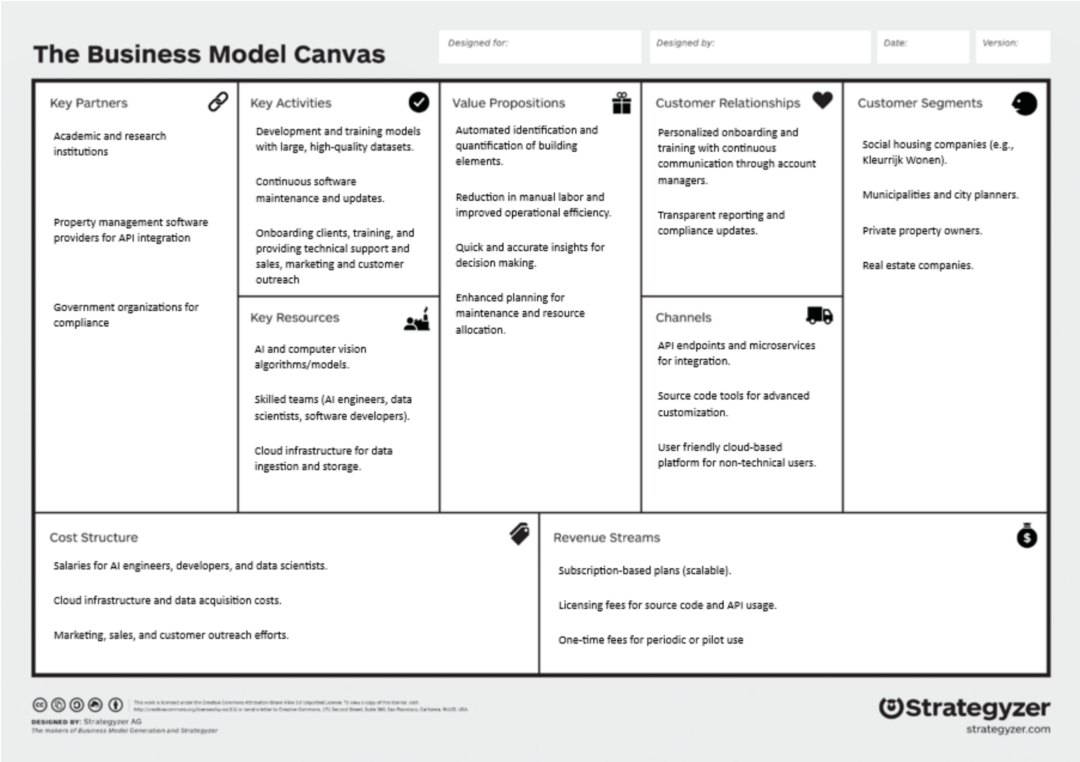

Figure 11: The business model canvas, automated identification and quantification of building elements.

4.2.1 Value Proposition

WoonVision aims to transform property management through AI-powered computer-vision that detects and counts building features. Through the automation of the identification and of the summation of building elements such as: windows, doors, roof tiles, balconies, etc., our platform is able to provide our clients with a clear and concise overview of their assets. The use of WoonVision drastically reduces manual labor, improves operational efficiency, enables informed decision-making and it gives our customers an upside in their negotiations with their contractors.

The value of our product lies in WoonVision's ability to analyse large amounts of image-data, while quickly delivering accurate insights to the users. By enabling users to quickly assess and understand the composition (counts of different building elements) of their entire product portfolio, WoonVision also empowers customers to proactively manage their resources and improve tenant convenience. The software is a game-changer for companies that manage extensive property portfolios, especially those counting older buildings of which specifications have never been noted down, by providing a high-tech and cost-effective solution as an alternative to manual inspections.

4.2.2 Customer Segments

While WoonVision was initially developed for Kleurrijk Wonen, a social housing company in the middle of the Netherlands, it can also cater to other organizations that are tasked with managing extensive property portfolios such as other social housing companies, private property owners but also real estate companies and municipalities. Social housing companies like Kleurrijk Wonen benefit greatly from WoonVision's ability to automate the identification and quantification of building elements as it allows them to manage their assets smarter and allocate their resources more effectively and efficiently. It will also help them to improve the planning of their long-term maintenance projects.

Municipalities and city planners can use WoonVision to oversee public and private buildings in their areas of jurisdiction. Through leveraging the insights gained from WoonVision, this type of customer can maintain an up-to-date overview of building inventories, track compliance with all different kinds of sets of regulations. WoonVision's ability to generate a comprehensive set of actionable insights could make it a crucial tool for city planning and development.

Real estate companies could use WoonVision to take their property listings to the next level by providing buyers or renters with precise details about the properties building elements and features. This highest level of transparency improves the customer experience through an increase in confidence and the acceleration of the sales process. By addressing these different but related customer segments WoonVision cements its viability as a property management solutions that addresses both operational and strategic challenges.

4.2.3 Channels

To effectively deliver the product to its customers WoonVision aims to offer a flexible set of channels tailored to the customer's needs and technological expertise. For companies with experienced development teams WoonVision will be available through a set of API endpoints and microservices, which will allow clients to integrate the product into the property management tool they are already using. For teams with advanced developers and highly-skilled technical experts WoonVision will also offer source code-level tools that will allow the customer to further customize or expand the tool's features to suit their own specific needs if they want to. For the companies that do not have the required expertise either of these options WoonVision will also be available through a cloud-based platform, this version will be developed with the goal of being as user-friendly as possible while making sure none of the core features are lost. These different channels ensure that WoonVision can reach each and every potential regardless of their technical expertise.

4.2.4 Customer Relationships

WoonVision aims to build strong and long-term relationships with its customers through personalized support. Each and every client starts their journey with an onboarding process that includes product training and assistance in the setup of the software tailored to the specific needs of the customer. This step is taken to ensure a seamless introduction of our product into the already large amount of software that is necessary to manage property effectively, this will ultimately enhance the customer experience and build a strong relationship between WoonVision and its customers.

Account managers will maintain customer relationships through continuous communication, ensuring that customers receive updates and assistance as their portfolio's change. These check-ins will also allow the development team of WoonVision to understand the ever changing needs and requirements of customers, which contributes to the continuous development of WoonVision and enhances the customized solutions we are able to offer.

Another step that WoonVision takes to improve their Customer relationships is the offer of continuous technical support, which ensures that issues will be resolved quickly. To make it as easy as possible for the customer's this technical support will be reachable through live chat, email and phone. Customers will also have access to an extensive knowledge base that could serve as self-service support for the more experienced development teams amongst the customers.

To increase trust in our AI-powered technology, WoonVision will try to be as transparent as possible through detailed reports and update logs. Software updates will not only make sure that new (personalized) features get added, but also that the model is compliant with the ever changing regulations that surround artificial intelligence.

4.2.5 Revenue Streams

WoonVision generates revenue through a mixture of subscription-based services, one-time fees, and tailored solutions. There will also be options for clients that seek periodic or pilot use, these clients are offered per-use fees for one time property assessments. This allows flexibility without immediate long-term commitments, this will be especially important in the early stages of the company, to generate some cash flow, as at this point there will not be as many strong customer relationships who want to commit to the longer-term options.

Subscription plans are the foundation of WoonVision's revenue model. The subscription plans are perfect for customers that require continuous use of the product for them to be able to effectively manage their property portfolio. These subscription plans will also be scalable, which is perfect for customers with evolving property portfolios and needs.

For the more advanced users and larger corporations, WoonVision provides options for licensing the source-code and API endpoints. These clients will pay licensing fees for longer terms than those of the subscription plans. As these terms will be longer term and the prices will be higher we will create a lock-in effect which increases the chance of the clients staying with us after their first term in comparison to switching to an alternative.

4.2.6 Key Resources

The success of the WoonVision product is built on a combination of technological, human and operational resources. The primary assets are the AI and computer vision algorithms and models which are designed to conveniently process high-resolution images and deliver concise and actionable insights into building elements and components. The nature of these technological resources ensure that they are not easily imitable which ensures a competitive edge and it allows for continuous development and improvement to cater more to the customer's specific needs.

WoonVision also depends on skilled teams consisting of AI engineers, data scientists, and software developers, who are tasked with the crucial task of the maintenance and enhancement of the product's capabilities. Furthermore, the customer service and tech support teams play a vital role in keeping client satisfaction at a high standard, handling the onboarding process and providing technical support where needed.

The most important operational resource is the data pipeline, which enables the efficient ingestion, processing and storage of the large amount of image data. This resource requires a well setup cloud infrastructure that will allow customers to upload an amount of image data from anywhere in the world.

4.2.7 Key Activities

The core of WoonVision key activities revolve around the process of continuous improvement. The development and improvement of AI computer-vision models are central to operations of WoonVision, improving the models will ultimately increase the customer satisfaction. To develop and improve these models they have to be trained on large datasets containing high-quality datasets, the collection of this data is an inherent key activity to the training of the models. The regular updates and the addition of new features to the software that will result from this training and development will mean nothing if the software does not work for one or more of the clients which leads to the key activity of keeping the software working at all times.

Another key activity is the customer engagement, this includes: onboarding clients, offering training, and providing personalized support to ensure smooth adoption. These steps are all taken to make sure that people working in property management can smoothly integrate WoonVision into their already complicated workflow. These activities will help build and foster long-term partnerships with our clients from which we could gather feedback used for the continuous improvement of the product.

WoonVision also needs to focus on sales and marketing of the product, because however nice the product is but nobody uses it, it is basically useless. The efforts made towards this include organizing webinars and workshops, attending trade shows and targeted customer outreach to expand the client base.

4.2.8 Key Partnerships

One of the key drivers of WoonVision's success formula are the strategic partnerships, these partnerships enable WoonVision to deliver value and to expand its reach. The first and most fundamental partnership is with academic and research institutions such as JADS, not only was the software developed through an assignment of this university. Working together with young and bright minds makes sure that WoonVision stays at the forefront of AI and computer vision advancements, this further enables the company to continuously improve its detection and analysis capabilities.

WoonVision also aims to collaborate with other property management software providers as much as possible. This will help enable seamless integration of WoonVision into the already existing workflow without overcomplicating the already difficult process. By offering API-based integration WoonVision becomes an extension of existing property management tools.

Last but not least, we have partnerships with government organizations and other regulatory bodies, collaborating with these entities helps to align WoonVision with the all different sets of rules and regulations it has to comply with such as GDPR.

4.2.9 Cost Structure

WoonVision's cost structure mainly revolves around the key activities that involve maintaining, scaling and improving the AI-driven computer-vision software. A large proportion of the costs are associated with salaries, especially those for the teams of AI engineers, data scientists, and software developers who are tasked with continuously developing and improving the software functionalities. The spending of these resources is crucial as these people ensure the technology is accurate and reliable.

Another set of costs associated with maintaining and improving the software are the cloud infrastructure and other computer power, which also includes the acquiring of high-quality datasets to train the models on. Especially since WoonVision deals with these large datasets of high-quality images there are significant costs related with processing of these images, these costs cannot however not be reduced as they are essential for the business operations.

The last large contributor to the cost structure are the marketing and outreach activities, these costs are essential as the company will not be able to return any profit if they do have any customers. To expand its customer base, significant efforts have to be made towards building customer relationships. Building these relationships will require WoonVision to attend industry events and perform targeted marketing campaigns.

4.3 Development

4.3.1 Data - Train set

There are no changes from the previous iteration regarding the sources of data offered by Kleurrijk Wonen. A new addition represents the data from the Sezen et al paper since it presents data already labelled for doors and windows. A disadvantage of this dataset is that it consists of buildings from Turkey, which present a different architecture than the one from the Netherlands, which we expect to negatively affect the results. This dataset is obtained from the "View Street" option of Google Maps and it has 727 images, with around 1000 instances of doors and 8000 instances of windows. They present a uniform distance in terms of view height and angle towards the building. In terms of labels, they are almost evenly distributed within the frame of the picture. Furthermore, from Figure X, we can observe the tendency of the labels to have a width longer than the height, because most of the windows are labelled as one when there are multiple windows next to each, so their frame is not necessarily used as a bounding box. To be mentioned that these pictures have a square format, with 640x640 pixels. This dataset will be used as a training dataset for our model. Some examples of this dataset can be seen in Figure 12 and Figure 13.

Figure 12: Turkey dataset - Example 1

Figure 13: Turkey dataset - Example 2

4.3.2 Data - Test set

The goal of our project was to test our model on a test set as close as possible to a potential real case within Kleurrijk Wonen. Because of this, we decided to label 100 images from the Megafile, for doors and windows, to be used in the testing phase. For this, we used Label Studio, which is an open-source data labelling and annotation platform. The set of rules for labelling was that if an element is cropped or hidden by any noise, such as vegetation, for more than 50% of its surface, it won't be labelled as an element that we want to detect. Furthermore, the selection of the images was made so that no multiple similar images are added in the test set and to have a perpendicular angle towards the facade of the buildings. Even if the Megafile has 3315, we had difficulties in finding these 100 images, since there are numerous images with a high degree of similarity. In these 100 images, there are 36 instances of doors and 636 instances of windows. The pictures are from the eye-level and aerial perspective and are non-uniform in terms of distance to the facade of the building.

4.3.3 Build

To start with, after the discussion with our client, we understood that an easy way to interact with our model is necessary. Because of this, a Flask web app was developed, where people can upload images and the model will return the image with the corresponding predicted bounding boxes and the number of door and window instances added to the image.

For model development, we use a similar development environment as in the first iteration, the only difference is that we used the L4 GPU from the Google Colab, which decreased almost 6 times the time for training and testing. Moreover, numerous tries and directions were made to improve the model performance, such as dealing with class imbalance, adding more pictures in the train set, or improving the model hyperparameters and data augmentation flags.

Firstly, there is a significant class imbalance towards the window class. Because of this, we checked if we could give a different weight towards the loss per class. Unfortunately, YOLOv8 doesn't have a built-in option to perform this, so a custom loss function needed to be defined, but we didn't look further into this. Another option to decrease the class imbalance was to use a model like Stable Diffusion to generate a matrix of Dutch doors. We dropped this idea after we observed that the generated images did not present the expected representation of the doors and we would risk introducing noise in the model.

Secondly, we used a single dataset with images from Turkey. What we tried is using other similar datasets, from Roboflow. Some datasets consisted of images taken during the night, while other datasets had only 1 relevant class for us labelled, which was the class window. In order to use them, we tried to use a multi-stage approach, where we were training the model only for class windows, since the doors were not labelled, while for the images taken during the night, we tried to illuminate them using the data augmentation flags. These approaches didn't show any improvement, so we dropped the idea of adding new datasets, that are not similar to our test set.

Thirdly, we tried to improve the model by modifying hyperparameters and data augmentation flags. Modifying the hyperparameters by hand didn't show an improvement, while performing hyperparameter tuning using a method like Grid-Search would be unfeasible, considering the computational power and time necessary to perform this technique. In terms of data augmentation flags, the default values cover a general case of object detection, like ours, so they already have the values that are useful to improve the performance and robustness of the model.

In the end, the final methodology applied was a multi-stage approach. In the first stage, we trained the YOLOv8 large model for 10 epochs while freezing the core of the model, which is the first 10 layers. Freezing the core decreases the training time, prevents overfitting while improving the generalisation capability of the model and uses the knowledge obtained in the pre-training phase. The image size is 640x640 and the batch size is 32. The second stage consists of taking the last obtained weights from the first stage and continuing to train the model on the same set of hyperparameters, without freezing any layers, for 14 epochs.

4.3.4 Measure

In our previous experiments, when we trained and tested a model on the Turkish dataset, we obtained good results, comparable to the ones from the Sezen et al, 2022 paper. At the moment when we created this train-test split, we expected to have worse results on the Dutch dataset, since it's hard for the model to map a function from the train set to the test set that presents similarities, but is not identical. Another option was to decrease the test set and move a set of images from the Dutch dataset in the training set, but we wanted an extensive test set of at least 100 images, to test the capabilities of the model.

The results of our model can be seen in Figure 14. The model has a 42% accuracy and 33.7% precision for class doors and 63% accuracy and 61.1% precision for class windows. As it can be seen, the model has difficulties in detecting the class door, whereas the model has the tendency to identify windows, that normally would represent the background. Moreover, our target is to count the number of windows and doors within the facades of the building. This means that we deal with a regression task. Because of this, we calculate the MAE by performing the difference between the predicted number of instances per class per image, and we calculate the absolute difference with the ground truth. This results in an MAE of 0.57 for class doors and 1.71 MAE for class windows.

A set of images with the predictions of our model can be seen in Figure 15. Considering that the train set does not have images similar to the test set, we can say that the results were expected, but this doesn't mean that the model is ready for a production environment. It still has difficulties in detecting both classes, but we need to understand that the task for window detection was not fully defined. At this moment, the model detects groups of windows, but we don't know if Kleurrijk Wonen prefers this method or a different one. The conclusion of this MVP is that it shows that this task is solvable, being confident that with a relevant dataset and an improved methodology, it can be created a model ready for production.

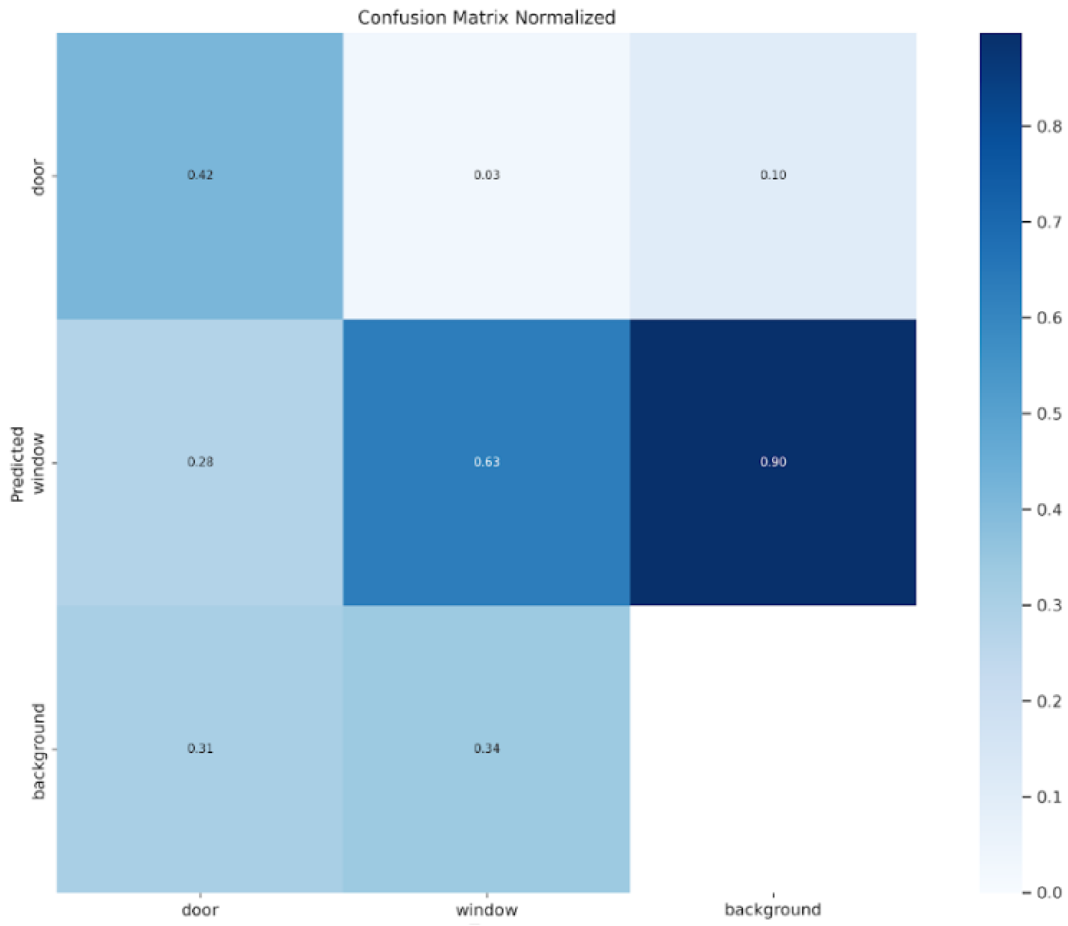

Figure 14: Confusion matrix

Figure 15: Predicitons of the model

4.4 Feasibility

4.4.1 GDPR Compliance

The images taken by a drone, by their very nature, include sensitive personal data, such as recognizable faces and license plates, which are classified under the definition of personal data within the GDPR. Currently, Kleurrijk Wonen manually blurs such recognizable details before further processing in order to ensure that nobody can be identified from the images. While effective, this manual process is labor-intensive, costly, and prone to human error. It's just both practical and beneficial to apply AI-driven anonymization tools, automating this task to bring down operational costs and improve uniformity. The integration into the workflow of systems developed for the identification and obscuration of sensitive components is possible, such as faces or license plates. Technologies could be used that utilize already developed object detection frameworks like YOLO or open-source resources like OpenCV to drive large-scale data anonymization.

Moreover, it should have in its procedural framework the principles of GDPR—data minimization and purpose limitation. For example, images that do not bring value to the model are removed after preprocessing and thus assure that there is no storage of data that is unnecessary. What's more, if external cloud service providers for data processing and storage—like AWS or Google Cloud—are involved in collaborative endeavors, then Data Processing Agreements (DPAs) are needed to provide for compliance with GDPR regarding appropriate data protection.

Tenant associations in the Netherlands, where these are part of the decision-making processes of housing corporations, may express their concerns regarding data privacy Bock, 2024. If Kleurrijk Wonen has not already taken action on this, clear policies need to be developed stating practices of data collection, anonymization, and retention. Distribution of these policies to the tenant associations will help trust and reduce delays in making decisions.

4.4.2 Intellectual Property (IP) Considerations

The reliance on open-source software is, therefore, a strategic one that aligns with the government's affinity for transparency and cost-effective solutions. Open-source applications like YOLO for object detection and TensorFlow for building artificial intelligence models significantly bring down development costs while encouraging innovation and adaptability. However, certain issues of intellectual property will have to be tackled.

First of all, open-source tools are governed by a variety of licenses, including MIT, Apache 2.0, and GPL, each coming with different terms of use. It's important to ensure that licenses of the tools chosen permit commercial use and the development of derivative works. This would be more relevant if the model is to be expanded or used in other contexts beyond Kleurrijk Wonen operations. Also, it's important to follow attribution requirements when modifying or extending open-source code. Non-compliance with these provisions could lead to potential legal liabilities and a damaged reputation.

To address these challenges, Kleurrijk Wonen should maintain an inventory of all open-source components used, their licenses, and their attribution requirements. It will ensure, through periodic review of the software stack, that the licensing terms are complied with to save the organization from unintended violations.

4.4.3 Privacy and Security Risks

Another important consideration is data security: the images being processed contain sensitive information. Even in an anonymized form, drone images can still contain information related to private property, potentially raising concern among tenants. Strong security measures should be put in place, including end-to-end encryption of data in transit and secure storage solutions with role-based access control, to prevent unauthorized access to sensitive information.

Equally, success also relies on the trust of tenants. To allay the fears of tenants, we should have proactive communication with the tenant associations, and there should be clear documentation about data handling and protection policies in place. This involves instituting strict internal policies that prevent the use of identifiable data beyond the intended purpose of damage classification.

4.5 Team roles distribution

For this project, Eugen was the person in charge and responsible for the development of the models. Yan helped in both, model development and the business side. Mathijs was responsible for data labelling, developing the web app and the business side. Friso took care of finding relevant datasets and improving the business side and Florentijn was the main business developer of this project.

References

Aedes. (2022). Aedes datacentrum - algemeen corporatiesector.

Aedes. (2023). Aedes datacentrum - aedes-forecast - financiën.

Bock. (2024). Huurdersorganisaties – behartigen belang per verhuurder. https://www.wooninfo.nl/bewonersinvloed/huurderskoepel/

Brynjolfsson, E., & Mcafee, A. (2017). Artificial intelligence, for real. *Harvard Business Review*, *1*, 1–31.

De Jong, L. K. E. W., & De Jong, L. K. E. W. (2021). Sociale huurwoning? in zeker een kwart van de gemeenten wacht je meer dan 7 jaar. *NOS*. https://nos.nl/op3/artikel/2377995-socialehuurwoning-in-zeker-een-kwart-van-de-gemeenten-wacht-je-meer-dan-7-jaar

Deloitte. (2022). Drone eam by infor and deloitte: How drones can accelerate the maintenance of industrial assets. https://www.deloitte.com/de/de/Industries/industrial-construction/perspectives/drone-assisted-asset-management.html

Duarte, D., Nex, F., Kerle, N., & Vosselman, G. (2019). Damage detection on building façades using multi-temporal aerial oblique imagery. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV-2/W5*, 29–36. https://doi.org/10.5194/isprs-annals-IV-2-W5-29-2019

European Union. (2019). Commission implementing regulation (eu) 2019/947. https://eur-lex.europa.eu/eli/reg_impl/2019/947/oj

Ideology, D. (2024, April). Understanding ai and data dependency. https://www.dataideology.com/understanding-ai-and-data-dependency/

Kotter, J. P. (1995). Leading change: Why transformation efforts fail. *Harvard Business Review*. https://hbr.org/1995/05/leading-change-why-transformation-efforts-fail-2

Lee, Y., Kim, H., Min, S., & Yoon, H. (2023). Structural damage detection using deep learning and FE model updating techniques. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-46141-9

MaintWorld. (2023). Euromaintenance comes to the netherlands.

Markets & Markets. (2023). *Drone inspection and monitoring market - global forecast to 2027*.

May, A. (2024, April). What are the risks of late rent payments? https://speedhome.com/blog/what-are-the-risks-of-late-rent-payments/

NOS. (2018, March). Woningcorporaties maken ruime winst op sociale huurhuizen. https://nos.nl/artikel/2219978-woningcorporaties-maken-ruime-winst-op-sociale-huurhuizen

Osterwalder, A. (2004). The business model ontology – a proposition in a design science approach.

PricewaterhouseCoopers. (2016). *Clarity from above: Pwc global report on the commercial applications of drone technology*.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. https://doi.org/10.1109/CVPR.2016.91

Research, G. V. (2023). *Infrastructure monitoring market size, share & trends analaysis report 2023-2030*.

Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Crown Business.

So, W. (2022). Which information matters? measuring landlord assessment of tenant screening reports. *Housing Policy Debate*, *33*(6), 1484–1510. https://doi.org/10.1080/10511482.2022.2113815

Teo, J. S., Mohd Aini, A., & Zyed, Z. A. (2024). Elevating property management in public housing: A systematic literature review. *International Journal of Strategic Property Management*, *28*(4), 234–247. https://doi.org/10.3846/ijspm.2024.21856

Van der Vliet, K. (n.d.). The application of drone imaging within the built environment.

van Algemene Zaken, M. (2023, July). 900.000 nieuwe woningen om aan groeiende vraag te voldoen. https://www.rijksoverheid.nl/onderwerpen/volkshuisvesting/nieuwe-woningen

van Algemene Zaken, M. (2024, September). Hoeveel huur betaal ik maximaal voor mijn woning? https://www.rijksoverheid.nl/onderwerpen/woning-huren/vraag-en-antwoord/hoeveel-huur-betaal-ik-maximaal-voor-mijn-woning

Viard, V. B. (2007). Do switching costs make markets more or less competitive? the case of 800-number portability. *The RAND Journal of Economics*, *38*(1), 146–163. https://doi.org/10.1111/j.1756-2171.2007.tb00049.x

Yasar, K., & Wigmore, I. (2023, March). Black box ai. https://www.techtarget.com/whatis/definition/black-box-AI