Long-Wave Infrared Polarimetry for License Plate Recognition

Mattias Larsson, Max Modig

Master's thesis
2024:E29

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematics

CENTRUM SCIENTIARUM MATHEMATICARUM

Long-Wave Infrared Polarimetry for License Plate Recognition

Mattias Larsson and Max Modig

Supervisors:
Anders Heyden[LU]
Thomas Winzell
Jimmy Heimdal

Examiner:
Niels Christian Overgaard[LU]

Abstract

With a growing global demand for improved surveillance and security, the thermal camera has seen its use cases and market share expand. Unlike the visual camera, the thermal camera does not require external illumination. Instead, it relies on the fact that all matter with a temperature above absolute zero emits thermal radiation. The thermal camera measures intensity, but another characteristic of electromagnetic radiation is its polarization state. The polarization state can provide additional information about the scene, such as the orientation and shape of objects and their material composition. By utilizing the features of the polarized domain, polarimetry has been used to solve a wide range of problems in a number of different fields. This project has revealed another application of polarimetric imaging in the area of license plate recognition. A full system, including a polarized thermal camera and required software solutions, has been constructed for this purpose. The limitations and restrictions of polarimetric imaging and how it specifically relates to the problem of license plate recognition have also been investigated. The observed polarization effect in the long-wave infrared domain is mainly influenced by the temperature difference between the object and its optical background, the incidence angle of the radiation, and the refraction index of the object. These three factors are, in turn, influenced by other, more or less controllable aspects. Through visual evaluation, it is clear that the polarimetric images provide additional and relevant information that is not present in a pure thermal image. Further evaluation of the implemented system, given that proper conditions are met, confirms an empirical advantage. The implemented system serves as a proof of concept and requires more development before any real-world implementations can be considered.

Keywords: LWIR polarimetry, Stokes parameters, license plate recognition, computer vision, machine learning, U-Net, OCR

Acknowledgements

Firstly, we would like to thank our supervisors, Thomas Winzell and Jimmy Heimdal, for the opportunity and the invaluable guidance throughout this project. They and the rest of the team have been a huge asset and were essential in realizing this project. Moreover, we want to thank Nicolás Rudervall for helping us with the design of the polarized camera system. Without his help, the process of capturing and deriving polarimetric images would truly have been a nuisance.

Finally, we would like to thank our supervisor, Anders Heyden, for his advice and help, and Niels Christian Overgaard for taking on the role as examiner.

Mattias Larsson and Max Modig

Table of Contents

1 Introduction

Unlike the visual camera, the thermal camera does not require external illumination. Instead, it relies on the fact that all matter, with a temperature above absolute zero, emits thermal radiation. The emitted radiation, at reasonable temperatures, is focused in the long-wave infrared spectrum, which in turn has beneficial transmission properties in atmospheric environments [1]. This contributes to improved vision and longer detection ranges compared to the visual camera, which makes thermal cameras useful in surveillance and security applications. While the thermal camera can distinguish temperatures effectively, it is considerably worse at capturing fine details, which are easily obtained with a visual camera. This makes object recognition with the thermal camera, in general, more difficult than with the visual camera.

The thermal camera measures the intensity of the incoming thermal radiation. Besides its intensity, another characteristic of electromagnetic radiation is its polarization state. The feature space of the polarized domain contains information about shape, orientation, and material composition [2][3]. By measuring the polarization signature of the radiation, more information can be extracted from the scene. This gives rise to new possibilities for thermal imaging in a variety of applications.

One potential application of thermal imaging is license plate recognition. In most cases, monitoring of cars is done using a visual camera. When external illumination is poor or undesirable, this approach may not be sufficient. In such situations, alternative approaches are necessary, and thermal polarimetry may be a plausible solution.

1.1 Problem Description

The goal of this project is, besides investigating the potential benefits of polarimetric imaging, to construct a polarized thermal camera system for the purpose of license plate recognition. The aim is to investigate if there is an advantage to using polarimetric imaging over pure thermal imaging for the given problem. Additionally, the limitations of polarimetric imaging and how it specifically relates to the problem of license plate recognition are also investigated.

As of this time, thermal cameras are not used for license plate recognition. In most cases, thermal imaging lacks the contrast to distinguish the characters from the background of the license plate. This is due to the fact that the temperature and emmisitivity properties are, in general, consistent across the plate. The additional information from the polarized domain may enhance the differences between the characters and the background and hence make the thermal camera a viable option for license plate recognition.

1.2 Related Work

Long-wave infrared polarimetry is a topic that has been active since the early 1990s and has been proposed as a solution to a number of different problems. Material classification is an important application. It has been shown that the polarization signature of radiation reflected off an object can give an insight into its material composition and can help distinguish between dielectric materials and metals [3]. In the field of astronomy, polarimetric imaging is used in the analysis of terrestrial objects, where it can provide information about structure and composition [4]. In both applications, the concept of polarization due to reflection is used. This is a widely known phenomenon that has been studied in depth [5]. Reflections can, in some cases, be a nuisance on the topic of object tracking and detection. It has therefore been proposed that the polarization signature of incoming radiation can be used in an image processing system to eliminate reflections [6]. Polarimetric imaging has also shown improvements in the application of facial recognition compared to standard thermal imaging. The main reason is the added information from the polarimetric domain about the texture and shape of the face [7]. Autonomous driving is also an area where this concept has shown promise. It has, for example, been suggested as a solution to the problem of road detection, where the polarization signature of the radiation reflected by the road is used to draw conclusions about the scene [8].

2 Background

In this section, the relevant physics concerning electromagnetic radiation and polarization is presented. An overview of artificial neural networks and projective transformations is also included.

2.1 Electromagnetic Radiation

Electromagnetic radiation consists of electromagnetic waves. Electromagnetic waves are, in turn, composed of two transverse components: an electric field and a magnetic field. The electric field and the magnetic field oscillate orthogonally both to each other and to the direction of propagation. Electromagnetic radiation interacting with a medium can either be reflected, transmitted, or absorbed. The outcome depends on the wavelength and the material that the wave is interacting with [9].

A black body absorbs radiation of all wavelengths and from all incident angles. If the black body is in thermal equilibrium with its surroundings, then the amount of emitted energy from the object equals the amount of absorbed energy. Ideal black bodies do not exist in nature. Instead, naturally occurring matter also reflects and transmits part of the inclined radiation [10]. The amount of energy that is emitted is measured by the emissivity of the matter. Similarly, absorptivity measures how much energy is absorbed [9]. The amount of emitted energy is directly proportional to the amount of absorbed energy. This can be explained by Kirchoff's law of thermal radiation, which states that if the matter is in thermal equilibrium, then the emissivity equals the absorptivity [11]. This is a direct result of the law of conservation of energy.

All matter with a temperature above 0K radiates electromagnetic radiation. The spectral density of the electromagnetic radiation emitted by a black body in thermal equilibrium is described by Planck's radiation law

$$E_\lambda(T) = \frac{2\pi hc^2}{\lambda^5} \cdot \frac{1}{e^{hc/k_B T\lambda} - 1} \tag{2.1}$$

where λ is the wavelength, T the temperature, h Plank's constant, c the speed of light in vacuum, and k_B Boltzmann's constant [10]. Figure 2.1 shows spectral densities of emitted radiation as a function of wavelength for different temperatures. For the given temperatures, it is clear that the majority of the emitted power is concentrated in the long-wave infrared spectrum.

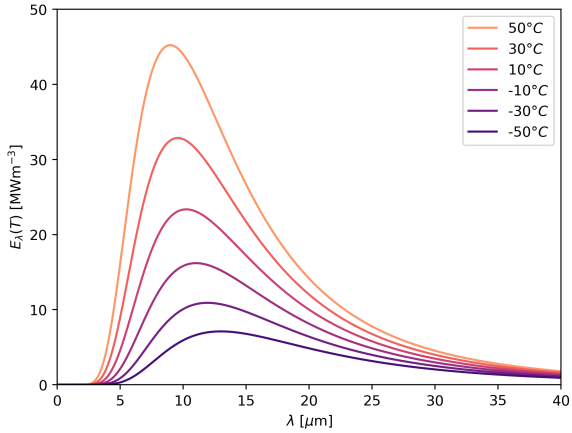


Figure 2.1: Spectral densities of emitted radiation for different temperatures.

2.1.1 Long-Wave Infrared Radiation

The LWIR (Long-wave infrared) spectrum is a subset of the infrared band of the electromagnetic spectrum. It covers the wavelengths of 8μm to 14μm. The transmittance, i.e., the ratio of the transmitted radiation to the total radiation incident upon the medium, of air as a function of the wavelength can be seen in Figure 2.2. This particular spectrum is measured at sea level over a distance of one nautical mile (1852m). Note that the transmission is dependent on the length and characteristics of the path that the radiation travels [1]. This spectrum indicates that LWIR radiation has better transmission properties in atmospheric environments compared to other wavelengths.

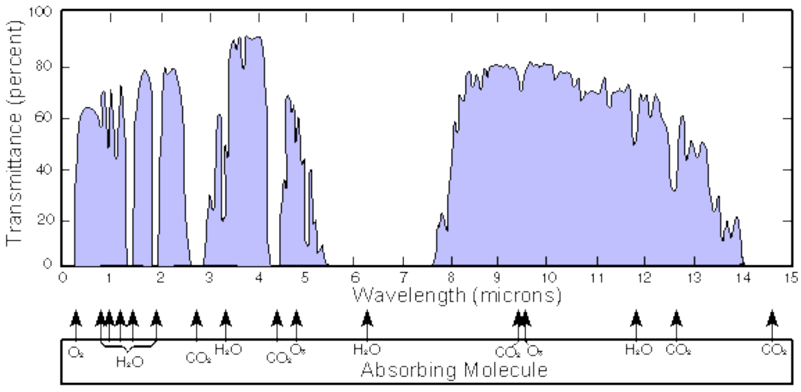


Figure 2.2: Transmittance of atmosphere [1].

This characteristic, together with the fact that the emitted energy at reasonable temperatures mainly consists of wavelengths from the LWIR spectrum, makes the LWIR spectrum valuable to observe.

2.2 Polarization

In general, electromagnetic radiation is unpolarized. This means that the direction of the electric field, and therefore also the magnetic field, does not follow any particular pattern. For polarized radiation, the electric field component instead oscillates in a non-random pattern. The polarization state of the radiation can be modeled in different ways. One way is with the polarization ellipse.

2.2.1 The Polarization Ellipse

For a monochromatic wave, i.e., a wave with only one wavelength, propagating in the z-direction, the x- and y-components of the electric field can be described by

$$\begin{cases} E_x(z,t) = E_{0x}\cos(\omega t - kz + \delta_x) \\ E_y(z,t) = E_{0y}\cos(\omega t - kz + \delta_y) \end{cases} \tag{2.2}$$

where E_{0x} and E_{0y} are the maximal amplitudes, ω the angular frequency, $k = \frac{2\pi}{\lambda}$ the wavenumber, and δ_x and δ_y the phases. By eliminating ωt from (2.2), it can be shown that

$$\frac{E_x^2(z,t)}{E_{0x}^2} + \frac{E_y^2(z,t)}{E_{0y}^2} - \frac{2E_x(z,t)E_y(z,t)}{E_{0x}E_{0y}}\cos\delta = \sin 2\delta \tag{2.3}$$

where $\delta = \delta_y - \delta_x$ [12]. Equation (2.3) is the equation of an ellipse and is called the polarization ellipse. This means that as a polarized electromagnetic wave propagates through space, the electric field traces out an ellipse, or some special case of an ellipse, such as a circle or straight line, in the plane orthogonal to the propagation direction. The polarization ellipse is depicted in Figure 2.3. With the polarization ellipse, the complete polarization signature of the electromagnetic wave can be described.

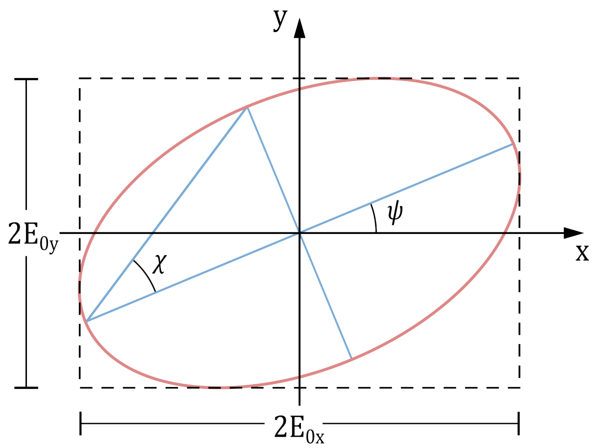


Figure 2.3: The polarization ellipse.

The angle of rotation ψ and ellipticity angle χ in Figure 2.3 are described by

$$\tan 2\psi = \frac{2E_{0x}E_{0y}\cos\delta}{E_{0x}^2 - E_{0y}^2}, \quad 0 \leq \psi \leq \pi \tag{2.4}$$

and

$$\sin 2\chi = \frac{2E_{0x}E_{0y}\sin\delta}{E_{0x}^2 + E_{0y}^2}, \quad -\frac{\pi}{4} \leq \chi \leq \frac{\pi}{4} \ . \tag{2.5}$$

These angles measure the orientation of the polarization and how linear or circular the polarization is, respectively [12].

2.2.2 Stokes Polarization Parameters

One limitation with the polarization ellipse is that it can only describe completely polarized radiation and not unpolarized nor partially polarized. To overcome this issue, the Stokes polarization parameters, which are four measurable quantities that can completely describe any state of polarized radiation, can be considered [12].

By letting $z = 0$ without losing generality, the polarization ellipse in (2.3) can be written as

$$\frac{E_x^2(t)}{E_{0x}^2} + \frac{E_y^2(t)}{E_{0y}^2} - \frac{2E_x(t)E_y(t)}{E_{0x}E_{0y}} \cos\delta = \sin 2\delta \ . \tag{2.6}$$

In order to represent (2.6) to measurements of the radiation, an average over the time of observation has to be considered. Let the time average of $E_i(t)E_j(t)$ be defined according to

$$\langle E_i(t)E_j(t)\rangle = \lim_{T\to\infty} \frac{1}{T}\int_0^T E_i(t)E_j(t)\,dt, \quad i,j = x,y \tag{2.7}$$

where T is the time of observation. It can be shown that applying (2.7) to (2.6) gives that

$$S_0^2 = S_1^2 + S_2^2 + S_3^2 \tag{2.8}$$

where

$$\begin{cases} S_0 = E_{0x}^2 + E_{0y}^2 \\ S_1 = E_{0x}^2 - E_{0y}^2 \\ S_2 = 2E_{0x}E_{0y}\cos\delta \\ S_3 = 2E_{0x}E_{0y}\sin\delta \end{cases} \tag{2.9}$$

and these parameters are called the Stokes polarization parameters. The first parameter, S_0, describes the total intensity, S_1 the amount of horizontal or vertical polarized radiation, S_2 the amount of $+45°$ or $-45°$ polarized radiation, and S_3 the amount of right or left circular polarized radiation [12].

Stokes parameters are also related to ψ and χ in the polarization ellipse according to

$$\tan 2\psi = \frac{S_2}{S_1} \tag{2.10}$$

and

$$\sin 2\chi = \frac{S_3}{S_0} \ . \tag{2.11}$$

The relation in (2.8) is only valid for completely polarized radiation. Hence, any state of polarization, including unpolarized and partially polarized radiation, fulfills the inequality

$$S_0^2 \geq S_1^2 + S_2^2 + S_3^2 \ . \tag{2.12}$$

A measurement that follows from the inequality in (2.12) is the degree of polarization, which is defined according to

$$p = \frac{I_{polarized}}{I_{polarized} + I_{unpolarized}} = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \tag{2.13}$$

where $I_{polarized}$ and $I_{unpolarized}$ represent the intensities of the polarized and unpolarized part of the radiation, respectively [12].

2.2.3 Stokes Polarization Parameters in Practice

Stokes polarization parameters can be measured in several ways. Let I_θ represent the intensity of the radiation oriented θ degrees from the horizontal level. Given the four intensities I_0, I_{45}, I_{90}, and I_{135}, a reduced and only linear version of Stokes parameters can be derived according to

$$\begin{cases} S_0 = \frac{1}{2}(I_0 + I_{45} + I_{90} + I_{135}) \\ S_1 = I_0 - I_{90} \\ S_2 = I_{45} - I_{135} \end{cases} \tag{2.14}$$

where S_3, which measures circular polarization, is excluded. The orientation angle given in (2.10) and the degree of polarization given in (2.13) are defined in a linear fashion as the Angle of Linear Polarization (AoLP) and the Degree of Linear Polarization (DoLP). These are denoted A and D, respectively, and are defined according to

$$A = \frac{1}{2} \arctan \frac{S_2}{S_1} + k\frac{\pi}{2}, \quad k \in \mathbb{Z} \tag{2.15}$$

and

$$D = \frac{\sqrt{S_1^2 + S_2^2}}{S_0} \tag{2.16}$$

where the value of A is chosen in the range $[0, \pi)$ for which

$$\operatorname{sign}(\cos 2A) = \operatorname{sign}(S_1) \tag{2.17}$$

is satisfied [8][13].

2.2.4 Polarization due to Reflection

Unpolarized radiation can become polarized as a result of a number of different physical processes. In general, the polarization of radiation occurs due to an interaction between the electromagnetic wave and a medium. One example of this is polarization due to reflection off a surface.

Radiation reflected off a surface can be separated into two distinct components: a diffuse component and a specular component. The diffuse component is created by radiation penetrating the surface of the material and undergoing multiple reflections and refractions before emerging from the material. In general, the diffuse component is unpolarized. The specular component is created on the surface of the material and has the same angle relative to the normal of the surface as the inclined radiation. The specular component can be polarized [14].

Assume that all inbound radiation upon an object is either reflected or absorbed and that the absorptivity is equal to the emissivity in each polarization state. This gives that

$$\begin{cases} \varepsilon_s(\theta) + r_s(\theta) = 1 \\ \varepsilon_p(\theta) + r_p(\theta) = 1 \end{cases} \tag{2.18}$$

is satisfied [5]. Here, ε is the emissivity of the object, r is the reflectance of the object, θ is the incident angle, and s and p represent the radiation perpendicular and parallel to the plane of incidence as depicted in Figure 2.4. The optical background refers to the source of the inbound radiation.

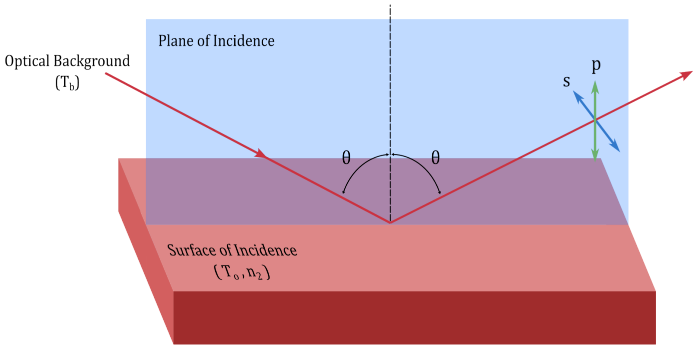


Figure 2.4: Polarization due to reflection.

If the temperature of the object and its optical background are denoted T_o and T_b, respectively, the total s- and p-polarized spectral radiance leaving the object, $L_\lambda^{(s)}$ and $L_\lambda^{(p)}$, are given by

$$\begin{cases} L_\lambda^{(s)}(\theta) = \varepsilon_s(\theta) E_\lambda(T_o) + r_s(\theta) E_\lambda(T_b) \\ L_\lambda^{(p)}(\theta) = \varepsilon_p(\theta) E_\lambda(T_o) + r_p(\theta) E_\lambda(T_b) \end{cases} \tag{2.19}$$

where $E_\lambda(T)$ is Plank's black body radiation curve at temperature T. It is here assumed that the emissivity and reflectivity are constant over the wavelengths, which is a reasonable assumption in the LWIR spectrum [5]. Using (2.18), (2.19) can be rewritten as

$$\begin{cases} L_\lambda^{(s)}(\theta) = E_\lambda(T_o) + r_s(\theta)(E_\lambda(T_b) - E_\lambda(T_o)) \\ L_\lambda^{(p)}(\theta) = E_\lambda(T_o) + r_p(\theta)(E_\lambda(T_b) - E_\lambda(T_o)) \end{cases} \tag{2.20}$$

which implies that if $T_o = T_b$ then $L_\lambda^{(s)}(\theta) = L_\lambda^{(p)}(\theta)$ and the radiation is unpolarized. If instead $T_o \neq T_b$ and $r_s(\theta) \neq r_p(\theta)$, then the radiation leaving the object is polarized. If $T_o \ll T_b$, the polarization is mostly due to the reflection, and if $T_o \gg T_b$, it is mostly due to emission. The values of $r_s(\theta)$ and $r_p(\theta)$ for a given angle θ are given by the Fresnel equations

$$\begin{cases} r_s(\theta) = \left| \frac{n_1 \cos\theta - n_2 \sqrt{1 - (\frac{n_1}{n_2}\sin\theta)^2}}{n_1 \cos\theta + n_2 \sqrt{1 - (\frac{n_1}{n_2}\sin\theta)^2}} \right|^2 \\ r_p(\theta) = \left| \frac{n_1 \sqrt{1 - (\frac{n_1}{n_2}\sin\theta)^2} - n_2 \cos\theta}{n_1 \sqrt{1 - (\frac{n_1}{n_2}\sin\theta)^2} + n_2 \cos\theta} \right|^2 \end{cases} \tag{2.21}$$

where n_1 and n_2 are refracting indices for the first and second mediums, respectively [5]. Figure 2.5 shows the difference between r_s and r_p for the angles $\theta \in [0°, 90°]$ and different values of n_2. The value of n_1 is set to 1.

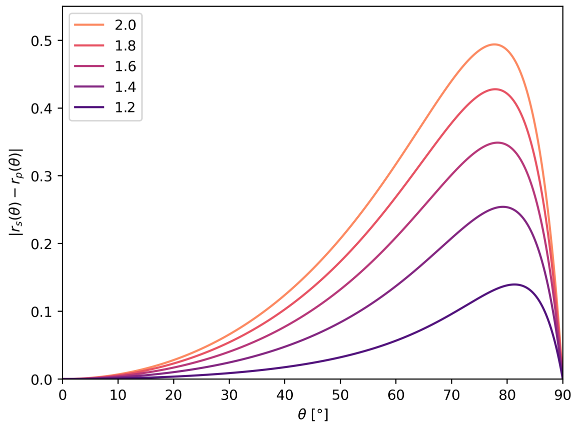


Figure 2.5: Difference in reflectance as a function of θ for different values of n_2.

Figure 2.5 indicates that the difference between r_s and r_p increases with the angle up to its maxima but also with the refraction index of the surface. If the difference between r_s and r_p increases, so does the difference between $L_\lambda^{(s)}$ and $L_\lambda^{(p)}$, and hence the degree of polarization.

2.2.5 Polarizing Filters

Polarizing filters affect the polarization signature of incoming radiation through selective absorption. Radiation passed through a polarizing filter loses intensity if the angle of its polarization signature does not match the transmission direction of the filter. The field transmitted through the filter equals the component aligned with the filter. For a linearly polarized field U with an angle α relative to the filter, the transmitted field appears as

$$u = U \cos \alpha \tag{2.22}$$

with the intensity

$$I = |u|^2 = U^2 \cos^2 \alpha \ . \tag{2.23}$$

This means that if the inclined wave oscillates in the transmission direction of the filter, then the wave does not lose any intensity. For all other angles, the field loses intensity, and when the polarization angle is orthogonal to the transmission direction of the filter, the field is completely eliminated. Incident unpolarized radiation that is transmitted through the filter becomes polarized, and the intensity is halved [15].

2.3 Artificial Neural Networks

An artificial neural network (ANN) is a machine learning model that is inspired by the structure and functionality of the human brain. It is composed of artificial neurons, or nodes, that are connected to each other in a layered structure. The process of employing an artificial neural network is to both design its structure and train it to perform well on a given task. The training requires mainly three things: training data, a loss function, and an optimization method. Neural networks are mostly used in a supervised setting, which means that the training data is a set of paired input data and expected output data. The loss function is a function that takes a true output and a predicted output and computes a loss that measures how well the network models the training data. The training is then executed through iterative minimization of the loss function given the training data and with respect to the parameters of the network.

2.3.1 The Perceptron

The perceptron is one of the simplest examples of an artificial neural network. It consists of a single node that, given a set of inputs x_i, $i = 1, ..., N$, produces the output $\hat{y}$ according to

$$\hat{y} = \varphi\left(w_0 + \sum_{i=1}^{N} w_i x_i\right) \tag{2.24}$$

where φ is an activation function, w_i weights, and w_0 a bias. The activation function models the behavior of a neuron, where a large input gives a large output and the triggering of the neuron follows a threshold-like behavior. Common activation functions are the sigmoid function and the ReLU function, which are defined as

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{2.25}$$

and

$$\mathrm{ReLU}(x) = \max(0, x) \tag{2.26}$$

[16].

2.3.2 Multilayer Perceptrons

A multilayer perceptron (MLP) is a fully connected feedforward artificial neural network that is composed of several perceptrons connected to each other in a layered structure. There are three kinds of layers: the input layer, the hidden layers, and the output layer. The depth of the network and the size of each layer can vary for different architectures. By increasing the complexity of the network, more complex functions can be approximated [16].

2.3.3 Convolutional Neural Networks

Convolutional neural networks (CNNs) are a subset of neural networks that are particularly applicable for two- and three-dimensional data, such as images. The foundation of the CNN is the use of convolutional layers. For each pair of input channels $x(\cdot, \cdot, l)$ and output channels $y(\cdot, \cdot, k)$, there is a kernel $w(\cdot, \cdot, l, k)$. Each output channel is computed as a bias $w_0(k)$ plus the sum of convolutioning each input channel with the corresponding kernel according to

$$y(i, j, k) = w_0(k) + \sum_u \sum_v \sum_l x(i - u, j - v, l) w(u, v, l, k) \ . \tag{2.27}$$

Another common building block in CNNs is the pooling layer. The purpose of the pooling layer is to decrease the dimensions of the feature map while maintaining important information. A common type of pooling is max-pooling. For max-pooling, a window of a certain size is shifted across an image, and the maximal value within this region will be the output.

In many instances, the CNN also contains one or more fully connected layers and non-linear activation functions, i.e., an MLP. A common architecture is to begin the CNN with sequential blocks of a convolutional layer, followed by an activation function and a pooling layer. These blocks are used to learn features. By then flattening the features to one dimension, they can be passed to a fully connected network [16].

2.3.4 The U-Net Architecture

The network architecture of the U-Net was proposed by Ronneberger et al. [17]. It is a model that has shown much success in the segmentation of images. Figure 2.6 shows the architecture of a U-Net. It resembles the structure of an encoder-decoder model. In the U-Net architecture, the encoder and the decoder are denoted as the contracting path and the expansive path, respectively [17].

The contracting path follows the design of a standard CNN. The convolutional blocks consist of two successive unpadded 3×3 convolutions, each followed by a ReLU, and a 2×2 max-pooling layer with a stride of two.

The expansive path is similar in structure to the contracting path. One difference lies in the introduction of up-convolutions, which increase the dimensions of the feature maps. The feature maps are also, in every step, concatenated with the corresponding feature maps from the contracting path. This process is called skip connections.

The contracting and expansive paths are connected by two successive convolution and ReLU operations. A 1×1 convolution is also added as a final layer to map the 64 feature maps to an arbitrary number of segmentation classes [17].

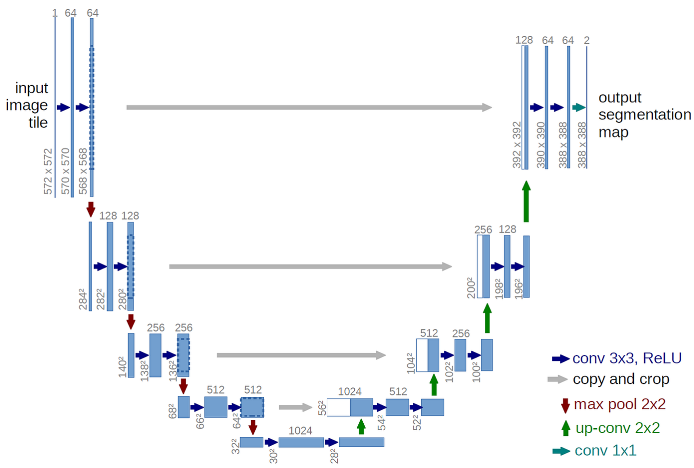


Figure 2.6: Network architecture of U-Net [17].

The skip connections have the added benefit that they help preserve information that would otherwise be lost while down-sampling. This effectively allows contextual information to be propagated throughout the network [17].

2.4 Projective Transformations in 2D

Two different images of a plane in 3D space are, assuming a pinhole camera model, related by a projective transformation, also known as a homography. The scenario is depicted in Figure 2.7.

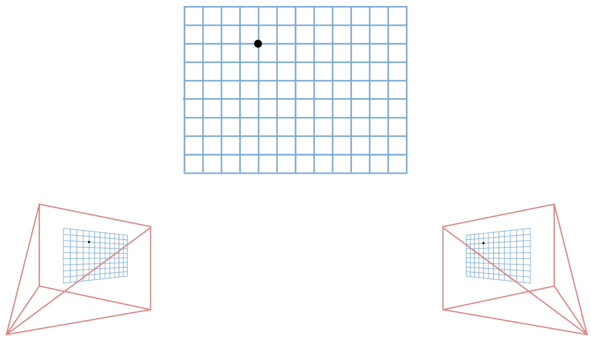


Figure 2.7: Two projected images of a planar surface.

More precisely, a point in the first image, $\boldsymbol{x}_i \in \mathbb{R}^2$, and the corresponding point in the second image, $\boldsymbol{y}_i \in \mathbb{R}^2$, are related according to

$$\lambda_i \begin{bmatrix} \boldsymbol{y}_i \\ 1 \end{bmatrix} = H \begin{bmatrix} \boldsymbol{x}_i \\ 1 \end{bmatrix} \tag{2.28}$$

where $\lambda_i \neq 0$ and H is an invertible 3×3 matrix. Since λ_i can take any value except zero, the scale of H is arbitrary. Because scale does not matter, H, which has nine elements, has eight degrees of freedom [16]. For every pair of corresponding points, (2.28) gives three equations and one additional unknown (λ_i). To determine H, at least as many equations as unknowns are needed. With n pairs of corresponding points, there are $3n$ equations and $8+n$ unknowns, and the number of needed correspondences is given by

$$3n \geq 8 + n \iff n \geq 4 \ . \tag{2.29}$$

This means that at least four pairs of points are sufficient to find the homography H.

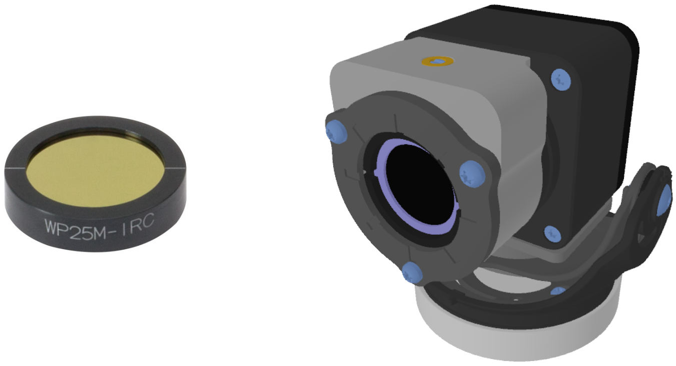
3 Method

The premise of this project, i.e., license plate recognition using polarimetric imaging, is only possible because of material differences on the Swedish license plate. The characters on the license plate are coated in matte black paint, and testing has shown that it tends to polarize thermal radiation less than the reflective background. This section explains the final system utilizing this fact for license plate recognition. The full system is divided into a polarimetric imaging system, a segmentation system, and an optical character recognition (OCR) system. The polarimetric imaging system involves taking and deriving polarimetric images. After that, the segmentation system takes a polarimetic image and produces a rectified and cropped image of the license plate. Finally, the OCR system takes the image of the rectified license plate and segments and classifies each individual character.

3.1 Polarimetric Imaging System

The main components of the thermal camera system are a thermal camera and a linear polarization filter. The thermal camera has a resolution of 288×384. The linear polarization filter that is used is shown in Figure 3.1a. It is designed for the wavelengths 7–15μm. The polarization extinction ratio, i.e., the ratio of the power in the transmission direction to the power orthogonal to the transmission direction after propagation through the filter, is 1:10000. It also has a transmittance of over 68% for the radiation components parallel to the transmission direction.

The linear polarization filter is attached in front of the thermal camera using a 3D-printed mold. This allows the filter to be rotated in different directions. The filter is placed as close to the lens of the camera as possible in order to reduce noise caused by the air pocket between the filter and the lens. The mold is attached to the camera, and the camera is in turn placed on a 3D-printed base, making it possible to tilt and rotate the camera. The whole ensemble is shown in Figure 3.1b.



(a) filter **(b)** camera system

Figure 3.1: Polarization filter and polarimetric thermal camera.

Much emphasis was put into the placement of the camera relative to the license plate. From theory, see Section 2.2.4, it is well known that in order to observe any polarization effect, there must be a temperature difference between the object and its optical background. The sky is considerably cold and easily accessible in an outdoor environment, which makes it suitable as an optical background for polarimetric purposes. The camera is therefore positioned below the level of the license plate and aimed upwards, which enables it to capture the reflection of the sky on the plate.

Early experiments suggest that there is a strong relationship between the angle of the camera relative to the license plate and the observed polarization effect. This also agrees with the theory behind polarization due to reflection given in Section 2.2.4. When the angle is increased, so is the observed degree of polarization. A too-large angle can, however, be problematic for two reasons. One issue is that if the angle is too large, the relative difference in observed polarization between the characters and the background starts to decrease. The second problem is that the effective resolution of the plate also decreases with an increased angle. Therefore, the angle of the camera relative to the center of the license plate is chosen in the range of $40° - 60°$, as this gives a good balance between the perceived polarization effect and geometric distortion.

The distance between the camera and the license plate is chosen in the range of 1–2 m, and the reason behind this decision is twofold. Firstly, the resolution of the thermal camera sets an upper limit for the distance since the license plate in the final image needs to be composed of enough pixels. A too large distance also makes it more difficult to capture the reflection of the sky on the plate since buildings and other obstacles may get in the way.

During each experiment, four images of the scene, denoted I_0, I_{45}, I_{90}, and I_{135}, are captured with the filter rotated to the angles $0°$, $45°$, $90°$, and $135°$ relative to the horizontal level of the camera coordinate system. From the four images, Stokes parameters, S_0, S_1, and S_2, are calculated according to (2.14), and from Stokes parameters, the AoLP and DoLP images are calculated according to (2.15) and (2.16). When deriving the AoLP image, there are special cases that have to be considered. For pixels where S_1 is zero, the AoLP image is undefined. The value in this situation is chosen according to the sign of S_2, as $45°$ if S_2 is positive and $135°$ if it is negative. If S_2 is also zero, i.e., if the radiation is unpolarized, then the value is chosen randomly in the range $[0°, 180°)$.

In total, besides the four images of the scene, five images, S_0, S_1, S_2, AoLP, and

DoLP, could be of interest for segmentation and classification of the license plate. The S_0 image describes the total intensity in the scene and would, in theory, if noise from the optics and imperfect rotation of the filter were ignored, represent the scene as if it were photographed with a pure thermal camera. Hence, it does not contain information about the polarization state. Both S_1 and S_2 contain polarization information. If the difference between horizontal and vertical polarization is large, then S_1 would contain interesting information. A similar reasoning can be done for S_2.

The AoLP and DoLP images both have interesting physical interpretations and can provide information regardless of the angle of the polarization. The AoLP image describes the angle of the polarization. In general, surfaces that are oriented differently relative to the camera polarize the radiation to different angles. This concept may be useful in several applications, as it gives insight into the shape and orientation of an object. Regarding license plates, since the characters and the background of the plate belong to parallel planes, the angle of the polarization across the whole plate appears homogeneous. This makes it difficult to distinguish the characters from the background in the AoLP image. The entire license plate may, however, be distinguishable from its surroundings, and the image may therefore be beneficial for segmentation of the plate but not for classification of the characters. The shape and orientation of objects also influence the degree of polarization. The reason for this is that the observed polarization effect is influenced by the incident angle of the radiation, which, in turn, depends on the shape and orientation of the object. Beyond this, the degree of polarization is also influenced by the refraction index of the surface. Since the characters and the background of the plate are often made of different surface materials, they tend to polarize the radiation to different degrees. This yields a contrast that makes it plausible to distinguish the characters from the background. Thus, the DoLP image may be useful for both segmentation and classification.

In the final system, the DoLP image is used for both segmentation of the plate and classification of the characters. The reason for this is that it provides the most relevant information and gives the best representation of the scene.

3.2 Segmentation System

The segmentation system takes the DoLP image as an input and produces a segmented and rectified image of the license plate. The DoLP image is first processed and then passed to a segmentation model. This model produces a binary segmentation of the plate that is used to mask out and rectify the plate.

The pre-processing of the input image includes histogram equalization and normalization. Histogram equalization increases the contrast in the image. This can enhance details in parts of the DoLP image where the difference in polarization is small. Normalization is applied to compensate for the varying degree of polarization in each experiment.

The segmentation model closely follows the design of the U-Net described in Section 2.3.4 and is shown in Figure 3.2. One difference is the introduction of padding in the convolutional layers. As a consequence, the dimensions of the feature maps do not change after the convolutional layers. Since the original dimensions (288×384) are four times divisible by two, the skip connections do not require any cropping or resizing. Another difference is the use of batch normalization in each convolutional block. The reason for this is that batch normalization has shown great success in increasing the convergence rate during training.

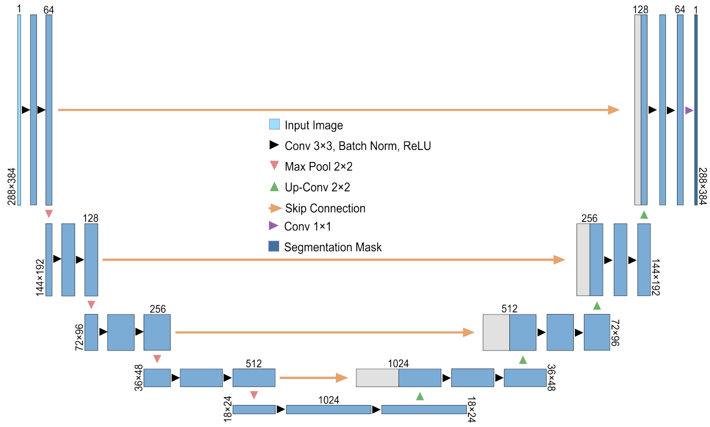


Figure 3.2: Network architecture of segmentation model.

The binary mask from the U-Net is used to find the corners of the license plate. This is done by approximating a polygon from the mask and extracting its corners. These four points are in turn used to estimate a homography. By cropping and applying the transformation to the DoLP image, a rectified image of the license plate is produced.

3.2.1 Dataset

The U-Net model is trained on pre-processed DoLP images and manually annotated segmentation masks. The training dataset is composed of 220 images taken during a variety of weather conditions. By considering different weather conditions, the robustness of the model can be improved. Five scenes are captured for each car in the dataset, varying the position and angle of the camera in between each experiment. The limited number of available cars makes it necessary to perform multiple experiments on the same car in order to gather enough data.

The testing dataset contains 60 images. These images are further divided into three categories. The first set of images were taken on a clear, cloud-free day. The second set includes images taken during a day with mild clouds, enough to partially obscure the sky. For the third set, images are taken during a day with a substantial amount of cloud cover. Each set is composed of two different images from ten different cars. No car used in any of the test sets is also present in the training set. Also, license plates with defects such as dirt and damage are excluded from the test sets. This is done in order to clearly understand the impact that the amount of cloud cover has on the system and also to see how well it can perform when proper conditions are met.

3.2.2 Training

The U-Net model is trained for a total of 100 epochs with a batch size of 16 and a fixed learning rate of 10^{-4}. The loss function used is the binary cross-entropy function with a built-in sigmoid function, and the Adam optimization method is used as an optimizer. At every epoch, each image in the training set is augmented by a random affine transformation and with Gaussian noise.

3.3 OCR System

The input to the OCR system is a rectified image of the license plate. Pre-processing is applied to the image before the characters are segmented. The segmented characters are then passed to a CNN for classification.

The pre-processing of the input image includes median blurring, adaptive thresholding, and image opening. The median blur reduces salt-and-pepper noise and other inconsistencies in the image. Adaptive thresholding is preferred over global thresholding since it can compensate for varying degrees of polarization across the license plate. After the thresholding, the characters can, at times, be connected to each other and unwanted regions in the background. This is the motivation behind the use of image opening. Besides the image opening, parts of the image, including edges and areas where no characters should be present, are also cut. The individual characters are segmented from the binary image by analyzing connected components and their area and center of mass. This produces six binary images of one character each.

The character classifier is a typical CNN composed of three blocks that contain a convolutional layer, a batch normalization, a ReLU, and a max-pooling layer. The convolutional blocks are followed by three fully connected layers with ReLU activation functions. A dropout layer is also present in between the last convolutional block and the fully connected network. This reduces overfitting and generalizes the model. The network architecture of the character classifier is shown in Figure 3.3. The relative simplicity of the input data allows the classifier to remain relatively simple, but a certain amount of complexity is still required because of the large number of classes. In total, there are 33 possible characters on a Swedish license plate. This includes the numbers *0–9* and the letters *A–Z*, excluding *I*, *Q*, and *V*. It is known that the first three characters on the license plate can only be letters, that the two following characters can only be numbers, and that the last character can be either, excluding the letter *O*. This prior knowledge is utilized to restrict the final prediction appropriately.

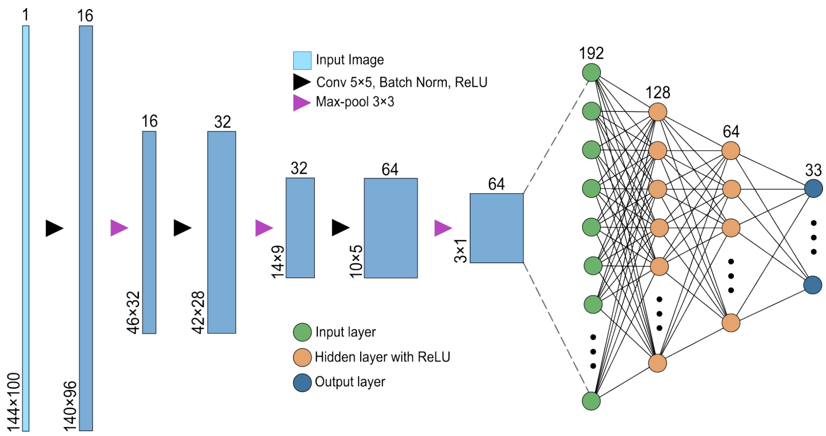


Figure 3.3: Network architecture of classification model.

3.3.1 Dataset

Due to the limited scope of the project and the limited number of available cars, gathering training data for the character classifier by photographing cars is not feasible. The dataset used to train the U-Net does not provide an even distribution of characters. The classifier is therefore trained on images of letters and numbers from two different font styles. The first font was used on license plates issued on Swedish cars between the years 2000 and 2014. The second font is the font currently used on Swedish license plates as of this writing. There are minor differences between the two font styles, and an example can be seen in Figure 3.4. By training the model on both font styles, it is able to make accurate predictions on both modern and slightly older cars.

(a) old font

(b) new font

Figure 3.4: Example of old and current font style used on Swedish license plates.

3.3.2 Training

The model is trained over a total of 25 epochs. The dataset is duplicated four times so that every character occurs a total of 32 times. With a larger set of data, the batch size can remain relatively large, and in this case, a batch size of 32 is used. The cross-entropy loss function and the Adam optimizer with a fixed learning rate of 10^{-3} are used. Augmentation is performed on the data during each epoch and includes a dilation, a random crop, and a random affine transformation. The purpose of the dilation is to mimic the increased boldness that the segmented characters possess as a result of the median blurring.

4 Results

This section contains results regarding the polarimetric imaging system, segmentation system, and OCR system. An evaluation of the full system is also presented. Regarding the segmentation system, in order to draw conclusions about the benefit of polarimetric imaging for this application, the system is also trained and evaluated on the S_0 image. Because of the interesting physical interpretation of the AoLP image, the system is trained and evaluated on this image as well. In order to enhance the difference in informational value between these images, different levels of complexity in the U-Net are also investigated.

4.1 Polarimetric Imaging System

The foundation of all the polarimetric images is given by the four taken images: I_0, I_{45}, I_{90}, and I_{135}. An example of how the images can look is shown in Figure 4.1.

(a) I_0

(b) I_{45}

(c) I_{90}

(d) I_{135}

Figure 4.1: Four taken images.

The Stokes polarization parameters, S_0, S_1, and S_2, that are computed from the four images can be seen in Figure 4.2.

(a) S_0

(b) S_1

(c) S_2

Figure 4.2: Stokes parameters.

The AoLP and DoLP images that are derived from Stokes parameters can be seen in Figure 4.3.

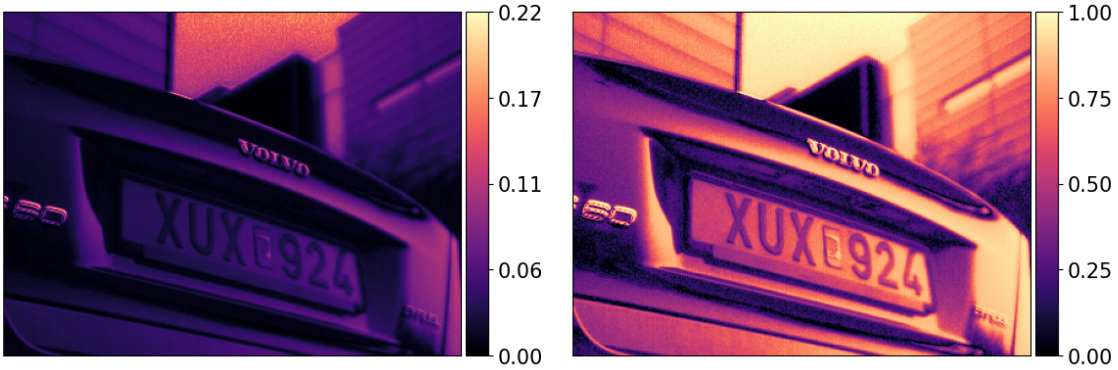
(a) AoLP

(b) DoLP

Figure 4.3: AoLP and DoLP images.

4.2 Segmentation System

Before the DoLP image is passed to the U-Net for segmentation, it is pre-processed. The DoLP image and the processed DoLP image are shown in Figure 4.4.



(a) DoLP **(b)** processed DoLP

Figure 4.4: DoLP and processed DoLP.

The segmentation results for the example scene presented above can be seen in Figure 4.5. The dice score between the predicted mask and the ground truth is 0.98.

(a) predicted mask **(b)** ground truth

Figure 4.5: Predicted mask and ground truth mask.

The resulting dice scores for different U-Net models, both for the train set and the test sets, are shown in Table 4.1. The models included are trained on S_0, AoLP, and DoLP images, respectively. The different architectures represent the depth of the U-Net, i.e., the depth of the contracting and the expanding paths.

Table 4.1: Dice scores for predicted masks.

architecture	input image	train dataset	test dataset		
			clear sky	mild clouds	clouds
U-Net 4	S_0	0.99	0.95	0.97	0.95
	AoLP	0.97	0.96	0.96	0.95
	DoLP	0.99	0.98	0.98	0.96
U-Net 3	S_0	0.98	0.93	0.95	0.82
	AoLP	0.97	0.96	0.96	0.91
	DoLP	0.99	0.97	0.97	0.93
U-Net 2	S_0	0.93	0.87	0.88	0.74
	AoLP	0.91	0.89	0.88	0.53
	DoLP	0.95	0.94	0.95	0.76

The corners of the masks are used to estimate homographies. Figure 4.6 shows the rectified and cropped license plate from the example above.

Figure 4.6: Rectified and cropped plate.

A subjective evaluation of the segmentation system is presented in Table 4.2. Each rectified and cropped plate is gauged as either acceptable or not based on the appearance of the image. A plate is considered acceptable if it contains all six characters with a minimal amount of geometric distortion. The plate in Figure 4.6 is considered a correctly rectified plate.

Table 4.2: Accuracies for correctly rectified plates.

architecture	input image	test dataset		
		clear sky	mild clouds	clouds
U-Net 4	S_0	0.90	0.95	0.90
	AoLP	0.80	0.90	0.65
	DoLP	1.00	1.00	0.95
U-Net 3	S_0	0.85	0.85	0.50
	AoLP	0.75	0.80	0.25
	DoLP	1.00	1.00	0.85
U-Net 2	S_0	0.35	0.25	0.35
	AoLP	0.20	0.30	0.10
	DoLP	0.80	0.65	0.25

The root mean squared errors (RMSEs) of the estimated corners are presented in Table 4.3. The RMSE of the estimated corners in Figure 4.5a is 2.45.

Table 4.3: RMSEs for estimated corners.

architecture	input image	test dataset		
		clear sky	mild clouds	clouds
U-Net 4	S_0	22.99	11.38	20.77
	AoLP	34.31	22.00	37.17
	DoLP	2.88	3.17	12.01
U-Net 3	S_0	35.99	59.49	89.85
	AoLP	11.61	47.80	73.96
	DoLP	3.89	3.42	36.16
U-Net 2	S_0	113.60	133.16	131.17
	AoLP	118.22	127.54	140.37
	DoLP	65.04	106.52	139.80

4.3 OCR System

The rectified image of the license plate is processed through a sequential set of steps before character classification. Figure 4.7 presents the major steps.

(a) input image

(b) after median blur

(c) after adaptive threshold

(d) after edge cut and image open

Figure 4.7: DoLP image after each processing step.

The largest connected components are segmented from the processed image. The segmented characters for the example are shown in Figure 4.8.

Figure 4.8: Segmented characters.

The segmented characters in Figure 4.8 are passed to the character classifier for a prediction. In this case, the model predicted the plate correctly. The accuracies for the OCR system on the train and test datasets can be seen in Table 4.4. In order to evaluate the OCR system independently of the segmentation system, the plate is rectified using manually annotated corners.

Table 4.4: Accuracies for the OCR system.

train dataset	test dataset		
	clear sky	mild clouds	clouds
1.00	0.98	0.96	0.85

4.4 Full System

The accuracies for correctly classified characters and correctly classified license plates for different weather conditions are presented in Table 4.5. A license plate is considered correctly classified if all six characters are correctly classified. The system is evaluated using the DoLP image as input and the U-Net 4 architecture.

Table 4.5: Accuracies for full system.

	clear sky	mild clouds	clouds
characters	0.98	0.97	0.79
license plates	0.90	0.90	0.55

5 Discussion

In this section, the results are discussed. The polarimetric images, the segmentation, and the character classification are all evaluated and discussed individually, followed by a complete discussion of the full system. Beyond the discussion of the results, limiting factors and future improvements are also presented.

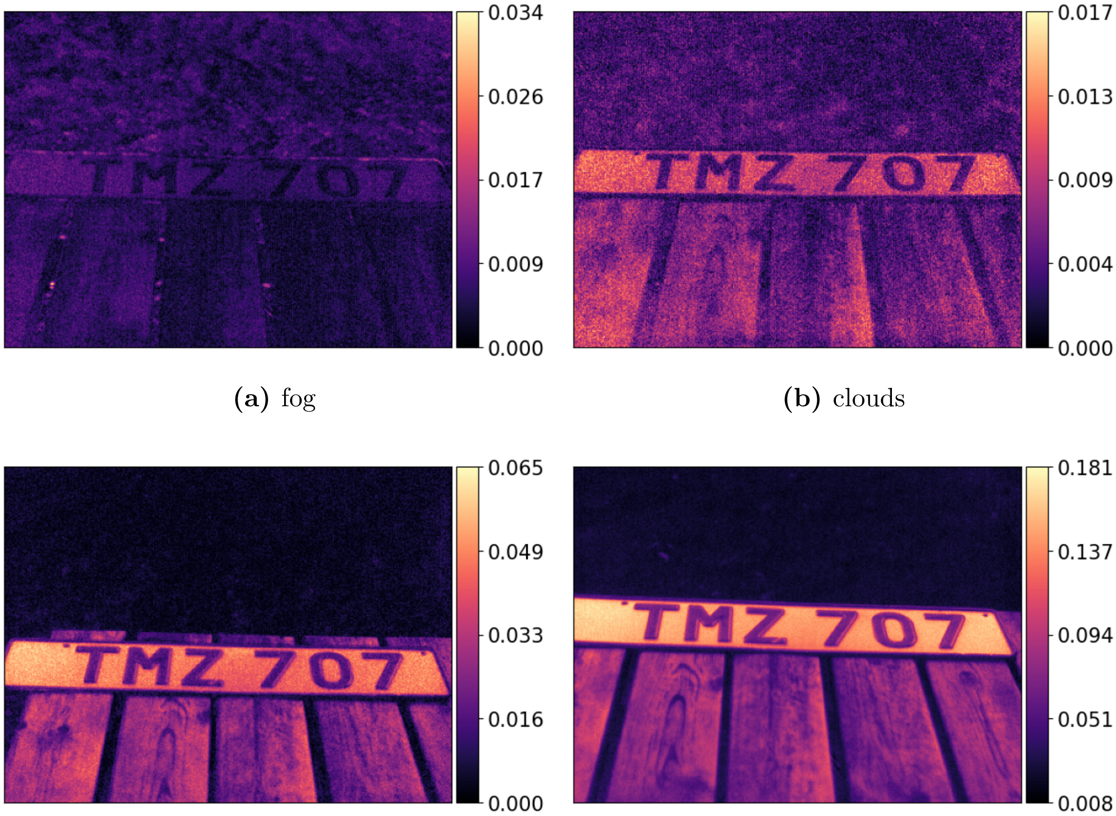
5.1 Polarimetric Imaging System

From experiments and evaluation of the system, it can be concluded that the observed polarization effect is mainly influenced by three factors. That is, the temperature difference between the object and its optical background, the incidence angle of the radiation, and the refraction index of the object, which agrees with the theory in Section 2.2.4. These three factors are, in turn, influenced by other, more or less controllable aspects.

The position of the camera has a large impact on the observed polarization effect. For pure polarimetric purposes, it is generally beneficial to use a large angle to increase this effect. Regarding license plate recognition, increasing the angle not only increases the polarization effect on the background but also on the characters of the plate. If the angle is large enough, then the difference in degree of polarization between the background and the characters decreases. This makes it harder to distinguish them, which in turn interferes with segmentation and classification. On the other hand, if the angle is too small, then there is not a large enough difference in the s- and p-components of the spectral radiance to observe any substantial polarization effect. Another disadvantage with a large angle is the decreased resolution of the license plate. The resolution of the license plate is not only affected by the angle but also by the resolution of the camera and the distance to the plate.

The distance between the camera and the license plate is chosen in the range 1–2m. For most practical settings, this distance is too short and may be problematic. If the car is moving at high speed, then a short distance makes the timing critical. The distance could be expanded as long as the sky remains the optical background. This demand also limits the camera position to always be below the level of the license plate. The height of the license plate relative to the ground varies from car to car, and in some situations, the camera may even have to be placed below ground level to capture the reflection of the sky.

Besides the placement of the camera, the observed polarization is also affected by the weather. Clouds and fog effectively mask the cold sky, which decreases the temperature difference and therefore the observed polarization effect. The influence of the amount of cloud cover is shown in Figure 5.1. Note that most images, in general, contain more noise and less contrast between the characters and the background compared to these images. The clear weather dependence means that the system is sensitive to changes in the scene that are beyond control, which is a clear drawback.



(a) fog

(b) clouds

(c) mild clouds

(d) clear sky

Figure 5.1: DoLP images for four different weather conditions.

Using the sky as the optical background has a practical appeal since it is almost always accessible in an outdoor environment. An alternative solution is to add a source of thermal radiation to the scene. The source could be controlled remotely and positioned appropriately. This would give more freedom in the placement of the camera regarding distance, angle, and level relative to the license plate. Early testing showed that using a thermal source with a temperature that differs by 5–10°C relative to the license plate gives enough polarizing effect to distinguish the characters. This is, however, a subjective observation, and more rigorous testing is required if this approach is to be implemented. Another possibility that would relax the restrictions on the positioning of the camera is to introduce a reflector. The reflector could be used to reflect the radiation from the sky onto the license plate.

The use of a rotating filter greatly limits the possible applications for this system since it requires the scene to be static. This means that any situation in which the car is moving cannot be considered. Because the filter is rotated manually, it is never placed exactly at the intended angle. Consequently, a certain amount of error is added to each of the taken images. This error propagates through the derivation of the other polarimetric images as well. Besides this, the filter also adds a certain amount of noise. The noise is created by the optics of the filter and is most notable in situations where illumination is poor. The rotating filter could be replaced by a grid of filters attached to each pixel in the sensor, i.e., a micro-polarizer array, as has been suggested by previous work [8][5]. This would allow the polarimetric images to be constructed from interpolations of a single image. Another approach could be to use a filter wheel that is rotated at a high speed. Such a method might be able to handle scenes that are not completely static and may reduce the error from imperfect placement of the filter.

In Section 4.1, the results concerning the polarimetric imaging system are shown. Figure 4.1 displays what the four images I_0, I_{45}, I_{90}, and I_{135} can look like. The images closely resemble the appearance of pure thermal images. There is a lack of sharpness in the images, and relevant polarization information can be difficult to observe. The images appear identical, as only minor differences are present. These differences can be hard to perceive but are made apparent when deriving Stokes parameters. Figure 4.2 shows Stokes parameters. The first image, S_0, does not contain any additional information compared to a pure thermal image. This is to be expected since S_0 represents the total intensity in the scene and does not include polarization information. More relevant information can be observed in S_1 and S_2. By comparing S_1 and the four taken images, it is clear that the contrast between character and background has increased. This is the case for S_2 as well. The AoLP and DoLP images for the example are shown in Figure 4.3. The AoLP image contains information about the structure and orientation of objects present in the scene. From the image, it is clear that surfaces oriented in the same way have similar values, and vice versa. The DoLP image is the sharpest out of all the images and also features the most relevant information. Comparing the AoLP and DoLP images with S_0 demonstrates the real advantage of polarimetric imaging. More features are present, and conclusions can be drawn about the scene that are impossible with just S_0. This difference is made clear in Figure 5.2.

(a) S_0

(b) processed DoLP

Figure 5.2: S_0 and processed DoLP.

In some instances, the polarization effect on the license plate differs from the images shown in Figures 4.1-4.3. At times, the characters polarize the radiation as much as the background of the plate. This is a substantial problem that cannot be controlled.

While no definite explanation for this phenomenon has been discovered, a possibility is that the license plate on some cars is coated in a substance that makes the refraction index across the license plate homogeneous. This is something that should be investigated further. Other factors that can distort the appearance of the plate are defects such as dirt, water, and dents, and that the reflective material has been damaged. These issues are more difficult to overcome and can cause problems even for visual cameras.

5.2 Segmentation System

Figure 4.4 shows the DoLP image before and after processing. It can be observed that the processing enhances contrast and details that were otherwise obscured. The final image is also more appealing to the human eye. This is, however, just a subjective evaluation. Training the segmentation model on both the processed and pure DoLP images shows an empirical benefit of the pre-processing.

The resulting segmentation from passing the processed DoLP to the full U-Net can be seen in Figure 4.5a. While the shape of the predicted mask closely follows the shape of the ground truth, there are differences across the edges and in the corners. The edges are not completely straight, and the corners are not as sharp as in the ground truth. The corners of the masks are critical for this application because they are used to estimate the homography that rectifies the plate. Since the dice score metric considers the entire segmentation mask and not just the corners, it is a reasonable metric for evaluating the U-Net. But it is not an optimal metric for the entire segmentation system. A predicted mask with suboptimal corners can still yield a high dice score, and vice versa. Therefore, evaluations of the estimated corners are used to gauge the performance of the full segmentation system.

Table 4.1 shows the dice scores for three different U-Net architectures, each trained and evaluated on S_0, AoLP, and DoLP images. The AoLP and DoLP images are considered because of the information they provide and the physical attributes they represent. The consideration of S_0 makes it possible to compare polarimetric imaging with pure thermal imaging for the given task. The motivation behind testing different U-Net architectures is to find the level of model complexity that is necessary for each type of polarimetric image. The AoLP models perform worse than the DoLP models for both the train and test datasets. The DoLP models also perform as well as or better than the S_0 models for all datasets. This suggests that there is an advantage to using polarimetric imaging for this specific purpose. It is, however, hard to draw any definite conclusions about the difference between the DoLP and S_0 models using the dice score metric. Their respective values are relatively similar and do not give much insight into the performance of the entire segmentation system.

A subjective evaluation of the segmentation system is presented in Table 4.2. Each rectified and cropped plate is gauged as either acceptable or not based on the appearance of the image. To objectively evaluate the segmentation system, the RMSEs of the estimated corners, as can be seen in Table 4.3, are also considered. Both measures confirm the conclusions drawn from Table 4.1, i.e., that the DoLP models outperform the AoLP models and, more importantly, the S_0 models. These measures are not only more representative of the performance of the entire segmentation system but also more strongly emphasize the superiority of the DoLP models.

It is difficult to conclude from Tables 4.1, 4.2 and 4.3 if the system performs better on the *clear sky* dataset or *mild clouds* dataset. This seems to indicate that a mild cloud cover does not have an adverse effect on the performance of the system. But when the cloud cover is substantial, it is clear that the system starts to struggle. Also worth mentioning is that the DoLP model performs slightly worse than the S_0 model with the U-Net 2 architecture for the *clouds* dataset. This difference is minor, and since both models fail to generalize to this dataset, no definite conclusions can be drawn. Even though the results give an indication of the relative performance of each model, the limited size of the test datasets makes the results sensitive to outliers and poor data entries. With more test data, the credibility of the drawn conclusions would be elevated.

A possibility that was not explored during this project is to concatenate multiple polarimetric images. By combining several images, more information would be available, and this could yield better performance. An interesting observation is that the AoLP, S_0 and DoLP images naturally translate to the channels in the HSV (Hue, Saturation, Value) color space. Figure 5.3 shows an image where the hue, saturation, and value are set by the AoLP, S_0 and DoLP images, respectively.

Figure 5.3: HSV image with H = AoLP, S = S_0, and V = DoLP.

The U-Net architecture has proved to be a valid network structure for the task of segmenting the license plate. As expected, when the complexity of the U-Net architectures decreases, so does the performance and robustness of the system. The less complex models, however, manage to highlight the informational value of each image with respect to the given task. To obtain better results, more training data could be collected. The training data is also slightly biased towards clear sky and mild clouds since it contains a larger percentage of data captured during such conditions. To some extent, this discrepancy is mitigated by the augmentations applied to the training data. Even though this approach is limited, the augmentations could be developed further to better model the effects of varying conditions.

Despite the fact that the U-Net proves to be a viable choice, there are a number of other approaches that may reduce the overall complexity. A potential solution is to construct a CNN that directly predicts the four corners of the license plate. The benefit of this approach is that resources are not wasted classifying each pixel in the image, which can be argued to be redundant work. Beyond this, a classical method may be worth investigating. Such a method could, for instance, be based on detection of parallel lines or corners. This method would most likely require less computation than the cumbersome U-Net.

5.3 OCR System

Figures 4.7 and 4.8 show the process of character segmentation. As mentioned earlier, the processing includes median blurring, thresholding, and image opening. The aim was to choose as few operations as possible but also have a clear motivation behind each step. Many alternative approaches were examined, but this set of operations gave the best results and proved to be adaptable to most images. Median blurring reduces noise, which ultimately contributes to more successful thresholding. One problem with this operation is that the characters get bolder and rounder, which affects fine details such as holes and sharp edges. This means that there is a balance between the reduction of noise and the distortion of the characters that has to be considered. The thresholding method used is an adaptive thresholding method. This method is favored over global tresholding since it performs better on images with varying degrees of polarization across the plate. Due to the nature of how the images are taken, the right side of the plate will observe more of a polarizing effect since it is at a larger angle relative to the camera compared to the left side of the plate. More consideration could be given to utilizing this prior knowledge, especially regarding the shape of the sliding window. At times, after thresholding, the characters are connected to the background or each other, which is undesirable. This can be seen in Figure 4.7c. Image opening and cutting of the edges are useful for separating the characters without distorting their shape and structure. One issue with this approach is that parts of the characters that are connected by fine structures can be disconnected. When the image is processed, the characters are segmented by picking the six largest connected components, and the order is based on the x-value of their centroids. For this task, there is more prior knowledge that could be utilized. Assuming a proper homography, the positions and sizes of the characters are approximately known. This knowledge could be used to further constrain the segmentation.

Applying the segmentation method mentioned above to any but the DoLP image yields poor results. The main reason is that this method relies on an intensity difference between the characters and the background. This intensity difference is, in almost every case, only present in the DoLP image. Several attempts to create a system for segmenting the characters in the other images were made. Different gradient-based methods were, for instance, tested on the S_0 image, but to no avail. This also points to a definite advantage of using polarimetric imaging for license plate recognition in the LWIR spectrum.

The accuracies of the OCR system for the train dataset and the three test datasets are presented in Table 4.4. The system performs well on the *clear sky* and *mild clouds* datasets. The few characters that are predicted incorrectly fail because of poor segmentation and not due to the classification network. The system performs noticeably worse on the *clouds* dataset, which is to be expected since the images contain more noise and have less polarization effect. This makes segmentation of the characters significantly more difficult. But in the general case, if the input image is clear and has a distinct polarization effect across the license plate, then the predictions are made correctly.

Much thought went into deciding on the complexity of the character classification network. The desire was to keep the network simple while maintaining the ability to distinguish between 33 classes. An interesting simplification that could be applied is to merge the classes representing the letter *O* and the number *0*. This is possible because they appear almost identical and never end up in the same position on the plate. Other network architectures and other machine learning approaches may also be worth investigating.

Training the classification network on characters from the Swedish license plate fonts proved to be a valid approach. Comparing the characters in Figures 3.4 and 4.8 shows that there is a difference between the font styles and the look of the segmented characters. The segmented characters have softer edges and are bolder than the font styles. Therefore, it is critical that the augmentation of the training data matches the appearance of the characters after segmentation. Thus, much consideration went into the design choices regarding augmentation. If the scope of the project was extended, a possibility would be to train the classification network on actual images of characters taken with the camera system. This is currently not possible due to the limited amount of collected data.

5.4 Full System

Table 4.5 presents the results for the full system. The DoLP image and the full U-Net architecture are used because of their high performance compared to other polarimetric images and architectures. As already concluded, there is a direct relationship between the performance of the system and the amount of cloud cover. Since the accuracy for the characters is higher than for the plates, especially for the *clouds* dataset, it can be concluded that the weakest link, in the majority of cases, is the OCR system. As discussed in Section 5.3, it is in particular the frailty of the character segmentation that affects the performance of the OCR system. This means that in order to construct a more robust system, more emphasis may be put on the character segmentation or on the aspects that affect the images it receives.

6 Conclusion

While the regular thermal camera measures the intensity of incoming thermal radiation, another characteristic of electromagnetic radiation is its polarization state. The polarization signature can provide additional information about the scene, such as the orientation and shape of objects and their material composition. This information can, in turn, be used to solve new problems and improve existing solutions. The observed polarization effect in the LWIR domain is mainly influenced by three factors: the temperature difference between the object and its optical background, the incidence angle of the radiation, and the refraction index of the object. These three factors are, in turn, influenced by other, more or less controllable aspects. Thus, in order to utilize LWIR polarimetry in real-world applications, a comprehensive understanding of the scene is necessary. In situations where the scene cannot be controlled or predicted, the capability and reliability of this approach may suffer.

LWIR polarimetry has been proven to be a possible and viable solution to the problem of license plate recognition. The implemented system is, however, restricted by a number of different factors. The positioning of the camera needs to be considered since it both affects the observed polarization effect and the quality of the image. The temperature difference between the license plate and the sky is the most critical factor that determines the degree of the observed polarization. This means that the weather has a substantial effect on the results. The difference in refraction index between the characters and the background of the plate is also crucial. This difference can, in some cases, be reduced or even eliminated as a consequence of damage to the license plate or due to changes to the surface material of the plate. The performance and robustness of the system could be improved in several ways. One of the most important potential improvements would be to add an external source of thermal radiation to replace the sky as the optical background. Many of the presented limitations and restrictions, e.g., the weather dependence and the placement of the camera, could be relaxed with this approach.

Even though the proposed system imposes a number of different restrictions on the scene and the setup, it performs well under the proper conditions. The implemented system serves as a proof of concept and requires more development before any real-world implementations can be considered.

This project corroborates an additional use case for LWIR polarimetry. The field of polarimetry is, however, not exhausted, and several interesting problems and applications are yet to be explored.

Bibliography

[1] D.O. Defense. *Electronic Warfare And Radar Systems Engineering Handbook*. Jeffrey Frank Jones. URL: https://books.google.se/books?id=RB_r9ccb-60C.

[2] Xuan Li, Zhiqiang Liu, Yudong Cai, Cunying Pan, Jiawei Song, Jinshou Wang, and Xiaopeng Shao. "Polarization 3D imaging technology: a review". In: *Frontiers in Physics* 11 (2023). ISSN: 2296-424X. DOI: 10.3389/fphy.2023.1198457. URL: https://www.frontiersin.org/articles/10.3389/fphy.2023.1198457.

[3] Hua Chen and L.B. Wolff. "Polarization phase-based method for material classification and object recognition in computer vision". In: *Proceedings CVPR IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1996, pp. 128–135. DOI: 10.1109/CVPR.1996.517064.

[4] Kevin Pohl, Jonathan Black, Jonathan Pitt, and Edward Colbert. "Low Earth orbit space object differentiation using long-wave infrared polarimetry". In: *Polarization: Measurement, Analysis, and Remote Sensing XIV*. Ed. by David B. Chenault and Dennis H. Goldstein. Vol. 11412. International Society for Optics and Photonics. SPIE, 2020, p. 1141206. DOI: 10.1117/12.2557901. URL: https://doi.org/10.1117/12.2557901.

[5] J. Scott Tyo, Bradley M. Ratliff, James K. Boger, Wiley T. Black, David L. Bowers, and Matthew P. Fetrow. "The effects of thermal equilibrium and contrast in LWIR polarimetric images". In: *Opt. Express* 15.23 (2007), pp. 15161–15167. DOI: 10.1364/OE.15.015161. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-15-23-15161.

[6] Ning Li, Yongqiang Zhao, Quan Pan, and Seong G. Kong. "Removal of reflections in LWIR image with polarization characteristics". In: *Opt. Express* 26.13 (2018), pp. 16488–16504. DOI: 10.1364/OE.26.016488. URL: https://opg.optica.org/oe/abstract.cfm?URI=oe-26-13-16488.

[7] Kristan P. Gurton, Alex J. Yuffa, and Gorden W. Videen. "Enhanced facial recognition for thermal imagery using polarimetric imaging". In: *Opt. Lett.* 39.13 (2014), pp. 3857–3859. DOI: 10.1364/OL.39.003857. URL: https://opg.optica.org/ol/abstract.cfm?URI=ol-39-13-3857.

[8] Ning Li, Yongqiang Zhao, Quan Pan, Seong G. Kong, and Jonathan Cheung-Wai Chan. "Illumination-invariant road detection and tracking using LWIR polarization characteristics". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 180 (2021), pp. 357–369. ISSN: 0924-2716. DOI: https://doi.org/10.1016/j.isprsjprs.2021.08.022. URL: https://www.sciencedirect.com/science/article/pii/S0924271621002288.

[9] Rajendra Karwa. *Laws of Thermal Radiation*. Singapore: Springer Singapore, 2020, pp. 733–765. ISBN: 978-981-15-3988-6. DOI: 10.1007/978-981-15-3988-6_10. URL: https://doi.org/10.1007/978-981-15-3988-6_10.

[10] S.M. Stewart and R.B. Johnson. *Blackbody Radiation: A History of Thermal Radiation Computational Aids and Numerical Methods*. Optical Sciences and Applications of Light. CRC Press, 2016. ISBN: 9781482263145. URL: https://books.google.se/books?id=MxwNDgAAQBAJ.

[11] D.R.L. Borissova, F. Smarandache, and L. Borissova. *Progress in Physics, vol. 4/2009: The Journal on Advanced Studies in Theoretical and Experimental Physics, including Related Themes from Mathematics*. Infinite Study. URL: https://books.google.se/books?id=GzI7DwAAQBAJ.

[12] D.H. Goldstein. *Polarized Light*. CRC Press, 2017. ISBN: 9781351833677. URL: https://books.google.se/books?id=cAlEDwAAQBAJ.

[13] L. Li, Z. Li, K. Li, L. Blarel, and M. Wendisch. "A method to calculate Stokes parameters and angle of polarization of skylight from polarized CIMEL sun/sky radiometers". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 149 (2014), pp. 334–346. ISSN: 0022-4073. DOI: https://doi.org/10.1016/j.jqsrt.2014.09.003. URL: https://www.sciencedirect.com/science/article/pii/S0022407314003744.

[14] Fang XS. & Boult T Nayar S.K. *Separation of Reflection Components Using Color and Polarization*. International Journal of Computer Vision. Kluwer Academic Publishers, 1997.

[15] K.J. Gåsvik. *Optical Metrology*. Online access: EBSCO Computers & Applied Sciences Complete. Wiley, 2003. ISBN: 9780470846704. URL: https://books.google.se/books?id=u15atbXzADUC.

[16] R. Szeliski. *Computer Vision: Algorithms and Applications*. Texts in Computer Science. Springer International Publishing, 2022. ISBN: 9783030343729. URL: https://books.google.se/books?id=QptXEAAAQBAJ.

[17] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

Master's Theses in Mathematical Sciences 2024:E29
ISSN 1404-6342

LUTFMA-3537-2024

Mathematics
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lth.se/

Populärvetenskaplig sammanfattning

Examensarbete: Long-Wave Infrared Polarimetry for License Plate Recognition
Författare: Mattias Larsson och Max Modig
Handledare: Anders Heyden[LU], Thomas Winzell och Jimmy Heimdal
Examinator: Niels Christian Overgaard[LU]

AVLÄSNING AV REGISTRERINGSSKYLTAR MED POLARISERAD VÄRMESTRÅLNING

På många platser i världen används kameror i olika säkerhets- och övervakningssyften. Vanligtvis finns det ett stort intresse av att kunna identifiera bilar och andra fordon som passerar kamerans vy. Generellt sätt görs detta med en visuell kamera genom avläsning av fordonets registreringsskylt. Detta arbete presenterar en ny metod för avläsning av registreringsskyltar med hjälp av polariserad värmestrålning. Systemet inkluderar en fysisk kamerauppställning med medföljande metodik för framtagning av polarimetriska bilder, och diverse mjukvaru- och maskininlärningsbaserade lösningar för att avläsa skylten.

Det polariserade värmekamerasystemet består av en värmekamera och ett linjärt polariseringsfilter. Genom att placera filtret framför kameran och ta bilder på ett systematiskt sätt kan polariseringen av den infraröda strålningen mätas och synliggöras.

För att observera någon märkbar polarisering används konceptet att elektromagnetisk strålning kan bli polariserad då den reflekteras av en yta. Hur stark den polariserande effekten är beror bland annat på ytans material. Eftersom karaktärerna på skylten och bakgrunden av skylten ofta är gjorda av olika material tenderar de att polarisera strålning olika mycket. Denna skillnad kan mätas med kamerasystemet och är anledningen till varför detta projektet är möjligt. Kamerans position relativt till registreringsskylten är också kritisk för att uppnå en nödvändig polariseringseffekt. Genom att placera kameran lägre än skylten så kan himlens reflektion fångas i skylten. Den höga temperaturskillnaden mellan skylten och den kalla himlen förstärker den polariserande effekten.

Figur 1 visar ett exempel på hur bilderna från kamerasystemet kan se ut. Den vänstra bilden är en klassisk värmekamerabild och visar intensiteten av den infallande värmestrålningen. Bilden till höger beskriver strålningens grad av polarisering. Detta är också bilden som används för avläsning av registreringsskylten. Det implementerade systemet som gör detta består i huvudsak av två neurala nätverk. Det första används för att segmentera ut skylten från bilden och det andra för att klassificera de sex karaktärerna. Utöver de två nätverken innefattar systemet dessutom nödvändig bildbearbetning och estimering av en projektiv transformation. Denna transformation ändrar perspektivet i bilden så att det uppfattas som om den är tagen rakt framför skylten.

Genom utvärdering av det implementerade systemet är det tydligt att termisk polarimetri är att föredra över klassisk termisk bildanalys för just avläsning av registreringsskyltar. Det finns däremot signifikanta begränsningar och mycket som måste tas hänsyn till. Den polariserande effekten beror främst på skylten och dess skick, temperaturskillnaden mellan skylten och himmeln samt kamerans positionering. Det är alltså viktigt att ha en stor förståelse för omgivningen och de faktorer som påverkar polariseringseffekten. När man har detta och rätt villkor är uppfyllda kan termisk polarimetri vara ett framgångsrikt tillvägagångssätt för avläsning av registreringsskyltar.

Figur 1: Värmekamerabild och framtagen polarimerisk bild.