

Optimierungen von Objekterkennungsmodellen in IR und RGB-Bildern

Objekterkennungsmodellenoptimierungen unter
verschiedenen Umweltbedingungen

BACHELORARBEIT

ausgearbeitet von

Ege Çağdaş ALADAĞ

zur Erlangung des akademischen Grades
BACHELOR OF SCIENCE (B.Sc.)

vorgelegt an der

TÜRKISCH-DEUTSCHEN UNIVERSITÄT
FAKULTÄT FÜR INGENIEURWISSENSCHAFTEN

im Studiengang
INFORMATIK

Betreuer/in: Asst. Prof. Dilek Göksel DURU
Türkisch-Deutsche Universität

Istanbul, im August 2024

Inhaltsverzeichnis

1	Einführung und Aufbau	4
2	Grundlagen	6
2.1	Digital Imagery	6
2.1.1	Visible Spectrum(RGB) Imagery	6
2.1.2	Thermal(Infrared, IR) Imagery	6
2.2	Machine Learning (ML)	6
2.2.1	Deep Learning	8
3	Stand Der Technik	9
4	Weitere Kapitel	10
4.1	Fußnoten und Listen	10
4.1.1	Beispiel Unterabschnitte	10
4.1.2	Listen	10
4.2	description-Umgebung	12
	Abbildungsverzeichnis	13
	Tabellenverzeichnis	14
	Literaturverzeichnis	16

Kurzfassung

Fügen Sie hier die Kurzfassung Ihrer Arbeit, welche bestenfalls strukturiert sein sollte, z.B. Einleitung, Hintergrund, Problemstellung, Zielsetzung, Vorgehen/Methode, Ergebnis, Fazit.

Abstract

Hier folgt die Kurzfassung auf Englisch.

1 Einführung und Aufbau

Object detection is a technology related to computer vision and image processing that deals with detecting instances of objects of a certain class (such as humans, cars, animals and drones) in digital images and videos. It has roots dating back to 1990s, although there has been major leaps in techniques and algorithms over the years. In the relatively recent years, deep learning methods have been prevalent in the object detection technologies.

In the rapidly advancing field of cameras and computer vision, the development of robust object recognition models is essential for applications ranging from autonomous systems to surveillance and beyond in many different fields. As technology continues to evolve, the integration of different forms of imaging has become a key focus to enhance the adaptability and reliability of these models. Visible images are affected by environmental and illumination variations such as low lighting and sun glare; meanwhile thermal and infrared images are noisy and have low resolution. [Bustos et al., 2023, p.1] The main advantage of thermal and infrared imagery is that they are not affected by light conditions, thus they can see objects that would otherwise be very difficult or even impossible to see with visible imagery.

The increasing usage and market size of infrared cameras and imagery (see Figure 1.1) and AI-based object detection continuously require better optimized and well performing models, especially in difficult environmental conditions such as rainy, foggy weather and high or low temperatures. Improvements to these detection methods and systems can have benefits extending into fields such as autonomous vehicles, agriculture, smart cities, search and rescue operations, public safety, security and military.

Object detection in the visible spectrum has seen a lot of interest and progress throughout the history of object detection. Deep learning methods have been developed within the past decade, that have continued to bring faster and more accurate detection performances. Some of the most prominent methods and algorithms currently used in object detection can be named as; R-CNN [Girshick et al., 2014] and its variants such as Fast R-CNN [Girshick, 2015] and Faster R-CNN [?], You Only Look Once(YOLO) [Redmon et al., 2016], Single Shot Multibox Detector(SSD) [Liu et al., 2016], RetinaNet [Lin et al., 2018], EfficientDet [Tan et al., 2020].

1 Einführung und Aufbau

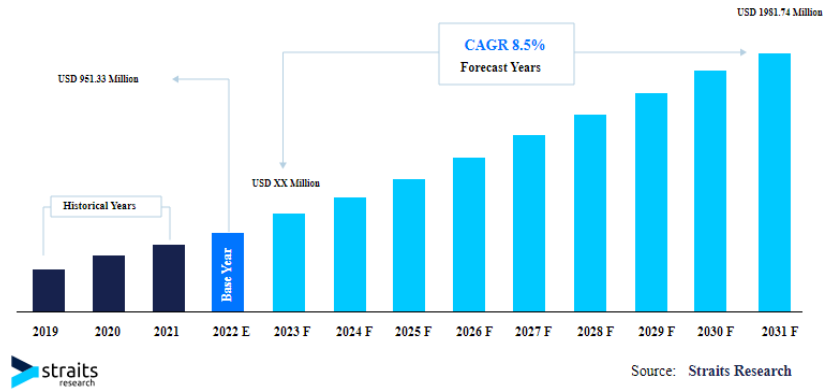


Abbildung 1.1: Infrared Camera Market Size [Straits Research, 2022]

The IR spectrum, on the other hand, is a relatively newer field in the context of object detection. Though it has been explored a lot, there really is no limit to the performance that may be achieved with further development. There has been works so far that have attempted to utilise the fusion of IR-RGB images to achieve better detection performances such as Wang et al. [2022]

2 Grundlagen

2.1 Digital Imagery

Digital imagery refers to visual content in digital form, that can be recognized and displayed by computers. For the purposes of this paper, we need to make a distinction between the following image types.

- Visible Spectrum(RGB) Imagery
- Thermal(Infrared) Imagery

2.1.1 Visible Spectrum(RGB) Imagery

NASA Science Mission Directorate [2010] defines the visible light spectrum as the part of the electromagnetic spectrum visible to the human eye, ranging from approximately 380 to 700 nanometers in wavelength. This range encompasses all the colors perceivable by the human eye, from violet to red.

2.1.2 Thermal(Infrared, IR) Imagery

SPI Corp [2014] defines thermal imaging, or thermography as the detection and measuring of radiation in the infrared spectrum being emitted from an object with the use of thermographic cameras. This type of imagery can collect temperature data from its field of view and display it using a variety of color palettes. Each pixel in a thermal image represents a temperature data point, and these data points are assigned a unique color or shade based on their value, meaning that as the thermal sensor detects changes in heat energy, it will express this change by adjusting the color or shade of a pixel. [Teledyne FLIR, 2021]

2.2 Machine Learning (ML)

“The studies reported here have been concerned with the programming of a digital computer to behave in a way which, if done by human beings or animals, would be described as involving the process of learning.” [Samuel, 1959]

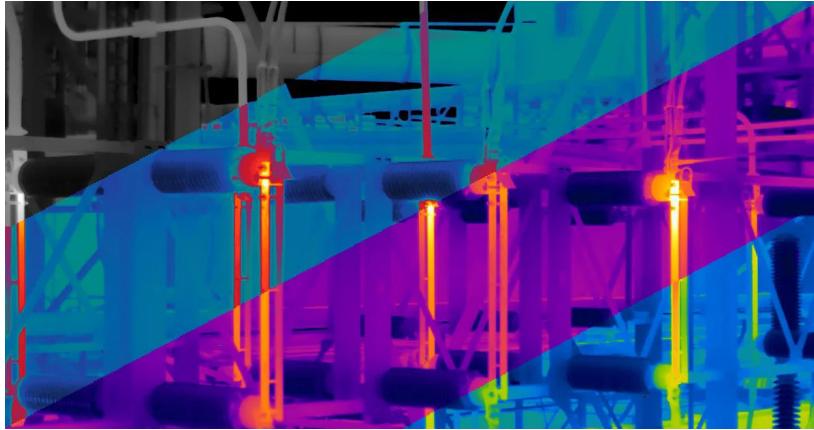


Abbildung 2.1: Thermal Color Palettes [Teledyne FLIR, 2021]

Machine Learning is recognized to be a term originally coined by Arthur L. Samuel in Samuel [1959]. It refers to the behavior of a computer to learn by processing data, improving its performance on certain tasks by working through a dataset and adjusting the algorithms to achieve better results. The idea is to require minimal human input while still performing well over a previously unseen set of data.

For a task to be fit for a machine learning application, a definite goal must exist, and at least one criterion or intermediate goal must exist which has a bearing on the achievement of the final goal and for which the sign should be known. [Samuel, 1959] Today, machine learning systems are used to identify objects in images, transcribe speech into text, match news items, posts or products with users' interests, and select relevant results of search. [LeCun et al., 2015]

Machine learning algorithms essentially work by processing training data and refining certain parameters to increase performance. This happens through iterations on the data and adjusting parameters by increments. The types of machine learning can be defined as so:

Supervised Learning is a very common method of machine learning, where the training data contains the expected output of an ML model. The training algorithm can then compute an objective function that measures the error (or distance) between the output scores and the desired pattern of scores [LeCun et al., 2015] and adjust the feature weights to improve the result by minimizing the error. The goal is to make correct predictions for new, unseen data.

Unsupervised Learning in contrast, does not contain any specific expected output within the training data. Thus the aim of the training is to recognize patterns within the dataset and categorize the data, rather than predicting an accurate label.

Reinforcement Learning algorithms that mimics how humans learn by trial and error.

These algorithms use a reward and punishment paradigm, where the actions that work towards the goal are reinforced and the actions that detract from the goal are punished or ignored. This feedback loop allows the algorithm to reach the goal by learning what is good and what is bad.

2.2.1 Deep Learning

3 Stand Der Technik

4 Weitere Kapitel

4.1 Fußnoten und Listen

Fußnoten können zum Teil sehr nützlich sein. Bitte beachten Sie, dass bei übermäßiger Verwendung von Fußnoten, die Lesbarkeit einschränkt sein kann¹.

4.1.1 Beispiel Unterabschnitte

Sie können in L^AT_EX Unterabschnitte verwenden. Wenn Sie allerdings einen Unterabschnitt einfügen, sollten es mindestens zwei sein. Es ist unüblich, dass man beispielsweise nur einen Abschnitt in einem Kapitel hat, oder nur einen Unterabschnitt in einem Abschnitt. Siehe zum Beispiel folgenden Tipp von Dave Patterson:

“Its strange to have a single subsection (e.g., 5.2.1 in section 5.2). Why do you need to number it if there is only one? Either eliminate the single subsection, or change the part that precedes the subsection into a second subsection” [Patterson, 2013]

4.1.2 Listen

Hier folgen einige Beispiele für Listen. Zunächst eine nicht nummerierte Liste:

- Item 1
- Item 2
- Item 3

Nun eine nummerierte Liste:

1. Item 1
2. Item 2
3. Item 3

¹da der Lesefluss unterbrochen wird!

Man kann auch Symbole verwenden:

→ Item 1

→ Item 2

→ Item 3

Beispiel für Unter-Unterabschnitt

Die gängige Gliederungstiefe von 3 Ebenen (Kapitel, Abschnitt, Unterabschnitt) sollte in der Regel nicht unterschritten werden. Sie können zwar eine weitere Ebene tiefer gehen, da dies ggf. die Lesbarkeit verringert wird diese in L^AT_EX nicht automatisch im Inhaltsverzeichnis aufgeführt. Hier werden beispielhaft Unter-Unterabschnitte verwendet.

Der folgende Text ist lediglich ein Platzhalter: *Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.*

Noch ein Unter-Unterabschnitt

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

Paragraph als Alternative Für diesen Abschnitt wurde nicht der `subsubsection{}` Befehl verwendet, sondern “paragraph”, der auch verwendet werden kann, um einen Unterabschnitt zu generieren. Verglichen zum `subsubsection{}` Befehl beginnt der Text hier nicht in einer neuen Zeile, sondern direkt nach der Überschrift. Wenn Ihr Dokument eher kurz gehalten ist, kann dies auch angemessen sein.

4.2 description-Umgebung

Wenn Sie bestimmte Konzepte beschreiben wollen, ist eine Liste oder ein Unterabschnitt ggf. nicht der beste Weg. Als Alternative gibt es außerdem die *description* Umgebung, die hier nützlich sein kann.

Konzept A Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua. At vero eos et accusam et justo duo dolores et ea rebum. Stet clita kasd gubergren, no sea takimata sanctus est Lorem ipsum dolor sit amet.

Konzept B Lorem ipsum dolor sit amet, consetetur sadipscing elitr, sed diam nonumy eirmod tempor invidunt ut labore et dolore magna aliquyam erat, sed diam voluptua.

Abbildungsverzeichnis

1.1	Infrared Camera Market Size [Straits Research, 2022]	5
2.1	Thermal Color Palettes [Teledyne FLIR, 2021]	7

Tabellenverzeichnis

Literaturverzeichnis

- N. Bustos, M. Mashhadi, S. K. Lai-Yuen, S. Sarkar, and T. K. Das. A systematic literature review on object detection using near infrared and thermal images. *Neurocomputing*, 560:126804, 2023. ISSN 0925-2312. doi: <https://doi.org/10.1016/j.neucom.2023.126804>. URL <https://www.sciencedirect.com/science/article/pii/S092523122300927X>.
- R. Girshick. Fast r-cnn. 2015.
- R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. 2014.
- Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.
- T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection. 2018.
- W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. *SSD: Single Shot MultiBox Detector*, page 21–37. Springer International Publishing, 2016. ISBN 9783319464480. doi: 10.1007/978-3-319-46448-0_2. URL http://dx.doi.org/10.1007/978-3-319-46448-0_2.
- NASA Science Mission Directorate. Visible light, June 2010. URL https://science.nasa.gov/ems/09_visiblelight. Last accessed 14 May 2024.
- D. Patterson. Dave patterson’s writing advice, 2013. URL <http://www.cs.berkeley.edu/~pattrsn/talks/writingtips.html>. Last accessed 16 July 2013.
- J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. 2016.
- A. L. Samuel. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3):210–229, 1959. doi: 10.1147/rd.33.0210.
- SPI Corp. What is thermal imaging?, 2014. URL <https://www.x20.org/knowledgebase/what-is-thermal-imaging/>. Last accessed 15 May 2024.

- Straits Research. Infrared camera market size, share & trends analysis report by technology (cooled ir camera, uncooled ir camera), by end-user (defense and military, industrial, commercial surveillance, automotive, bfsi, healthcare, residential, others) and by region(north america, europe, apac, middle east and africa, latam) forecasts, 2023-2031, 2022. URL <https://straitsresearch.com/report/infrared-camera-market>. Last accessed 30 April 2024.
- M. Tan, R. Pang, and Q. V. Le. Efficientdet: Scalable and efficient object detection. 2020.
- Teledyne FLIR. Picking a thermal color palette, Apr. 2021. URL <https://www.flir.com/discover/industrial/picking-a-thermal-color-palette/>. Last accessed 15 May 2024.
- Q. Wang, Y. Chi, T. Shen, J. Song, Z. Zhang, and Y. Zhu. Improving rgb-infrared object detection by reducing cross-modality redundancy. *Remote Sensing*, 14(9), 2022. ISSN 2072-4292. doi: 10.3390/rs14092020. URL <https://www.mdpi.com/2072-4292/14/9/2020>.

Eidesstattliche Erklärung

Ich versichere, die von mir vorgelegte Arbeit selbständig verfasst zu haben.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer entnommen sind, habe ich als entnommen kenntlich gemacht. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit benutzt habe, sind angegeben.

Die Arbeit hat mit gleichem Inhalt bzw. in wesentlichen Teilen noch keiner anderen Prüfungsbehörde vorgelegen.

Istanbul, 23. Mai 2024

Max Mustermann