

Institute for optical systems

Annual report 2021, 2022, 2023

Preface

Since its founding in 2008, the Institute for Optical Systems (IOS) at the HTWG Konstanz has been dealing with a representative cross-section of current optical technologies. The focus is on the system idea, which requires thematic coverage of all areas and components that are important in the industrial environment. Our goal is to be a competent contact point for all system components. Applications outside of optical systems are also considered and the following subject areas are covered:

- lighting technology,
- optical measurement technology,
- image and geometry processing,
- computer graphics,
- computer vision,
- machine learning and AI,
- mathematical modeling and simulation.

The research focus of the IOS has emerged at the HTWG Konstanz as one of three focal points in the area of research and teaching across the faculties of computer science, mechanical engineering, civil engineering and electrical engineering and information technology.

Contents

Preface	2
Institute Profile	4
IOS Building and Location Plan	5
Institute Members	6
IOS Staff	8
External Funding and Grants	9
Cooperation with Research Institutions and Industry	10
Publications 2020-2023	12
Theses and Student Projects	15
Research Activities	19
BiLeSA: Mit digitalen Bildern Mathe besser verstehen	20
Cman: Crowdmanagement in der Lehre	21
Softwaretool zur Berechnung von Personenflüssen und Infektionskrankheiten	22
Validierte Simulation von Personenströmen und Infektionsrisiken	23
Energieeinsparpotential durch intelligente Thermostate	24
Environmental Impact Assessment of IoT Devices: A Graph-based Approach	26
Entwicklung eines neuronalen Netzes zur Vorhersage von Simulationsergebnissen eines Personenstrommodells	28
Untersuchung von Methoden zur multivariaten Anomaliedetektion in HVAC-Signalen	30
Novel AI-Based Algorithm for the Automated Computation of Coronal Parameters	41
Using Gaussian Process for Robust Temperature Forecast	43
Accelerating BNN's in the context of active learning	45
Geometrische Kalibrierung zwischen IMU/GPS und LiDAR	47
Lageerkennung der Wasseroberfläche durch semantische Segmentierung und Stereo Vision	49
Transferlernen für Objektdetektion auf Binnengewässern	51
Farbkalibrierung mit spektrometrischen Referenzdaten für den Inline-Digitaldruck	52
Analyzing keypoint detectors and descriptors learned on RGB-D images	54
Eine Untersuchung über den Einsatz von Ensemble-Methoden in der Anomalieerkennung	56
Intelligente Bildverarbeitung in der Materialmikroskopie durch maschinelles Lernen zur Objekt- und Fehlererkennung	57
Multifunktional-skalierbare generische Inlinelnspektion für flexible Fertigungsprozesse	59
Punktwolkenklassifikation maritimer Daten	61
Analyse von latenten Gaußschen Mischmodellen zur Registrierung von Punktwolken	63
Kamera-Lidar-Kalibrierung mit Partikel-Schwarm-Optimierung	65
Untersuchung des Multi-View 3D-Rekonstruktionspotentials von Plenoxels und FastNeRF	67
Detektion von Kalibrierobjekten in Punktwolken und Kamerabildern	69
Registrierung von Fuß-Scans mit Raumdeformationsverfahren	71
Bildbasierte Geolokation für Augmented Reality in Gebirgen	72
Optisches Multisensorsystem für autonome Anwendungen auf Binnengewässern	74
Organizational Activities	76
Mitglieder Liste	77
IOS Leitung	77
Mitarbeiter*innen und Doktorand*innen	77
IOS Sitzungen	78
Wissenschaftliche Selbstverwaltung	78
Ausblick	79

Institute Profile

IOS BUILDING AND LOCATION PLAN



Institute for Optical Systems

Physical address:

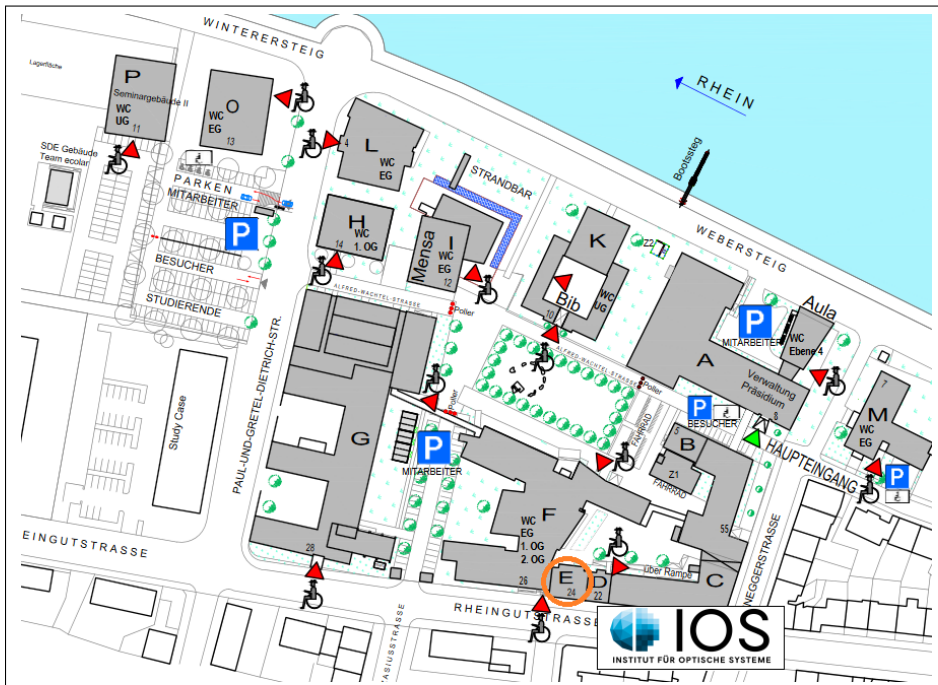
Rheingutstraße 24
Building E
3rd and 4th floor
78462 Konstanz

Postal address:

Alfred-Wachtel-Straße 8
HTWG Konstanz
D-78462 Konstanz
Germany

Web address:

www.ios.htwg-konstanz.de



INSTITUTE MEMBERS

Prof. Dr. Rebekka Axthelm



Diploma in mathematics at the Albert-Ludwigs University of Freiburg, doctorate in mathematics in 2007 funded by a DFG scholarship. Research topics were theory and numerics of partial differential equations and finite element methods. For three years she worked in the development department of a manufacturer of semiconductor chips for HD television in the field of mathematical methods for image processing. She worked for eight years as a lecturer at the Zurich University of Applied Sciences in various engineering degree programs. Since 2017, she has been a professor of mathematics in the Faculty of Computer Science in Konstanz. Her research interests include the modeling and numerical calculation of differential equations. The main area of application is the macroscopic modeling of pedestrian movements, as well as the numerical calculation and simulation thereof and the spread of infectious diseases based on this.

Prof. Dr. Doris Bohnet



M.A. in philosophy and Diploma in mathematics from TU Berlin and University of Hamburg, doctoral degree in mathematics in 2012. Research stay at University of Evanston, USA, and postdoctoral studies at Université de Bourgogne, France, both funded by scholarships. Research topics were the classification of smooth dynamical systems. During 5 years, she worked as researcher and project leader in the automobile industry in the area of modeling and simulating of thermodynamical systems, mostly cooling of engines, thermal validation and exhaust system. Since 2018, she is professor, first at University of Applied Sciences Ruhr West, then, since 2020, at Konstanz. She is interested in analyzing, modeling and optimizing thermodynamical systems and related systems and in grey or black box models for such systems combining ML approaches.

Prof. Dr. Oliver Dürr



Ph.D. in theoretical and computational physics from the University of Konstanz, supported by research fellowships. Specialized in computational and theoretical methodologies for analyzing complex systems. Prior to joining the University of Applied Sciences, Konstanz (HTWG) as a professor of data science in 2018, held positions as a researcher Genedata, Basel, and as a lecturer at Zurich University of Applied Sciences. Presently, actively engaged in research that delves into the intersection of deep learning and statistics, investigating their symbiotic relationship across various data analysis scenarios.

Prof. Dr. Matthias Franz



M.Sc. in Atmospheric Sciences from SUNY at Stony Brook, Diploma in physics from the Eberhard-Karls-Universität, Tübingen and doctoral degree in 1998. Thesis research in visual insect and robot navigation at the MPI for Biological Cybernetics and as a PostDoc at the Australian National University in Canberra. In industry he worked on various aspects of autonomous vision systems. 2002, he returned to the MPI as a group leader in the area of machine learning and computer vision. Since 2007 professor at the University of Applied Sciences in Konstanz. Main research activities in the development of automatically generated vision systems, optimization and probabilistic modeling, with applications in industrial machine vision, texture analysis and steganalysis.

Prof. Dr. Hartmut Gimpel



Physics degree at University of Heidelberg (2002) on quantum experiments with ultra-cold atoms. Academic year at UIUC, Illinois, USA (2000) on superconductor physics. PhD at Max-Planck Institute in Heidelberg (2006) on generation of ultra-short laser pulses with stable offset phase for experiments in atomic physics. An eighth century as optics and system engineer at Sick AG near Freiburg. Design and product development of a multitude of optical sensors: cameras, laser line cameras, laser illuminations, time of flight sensors for distance measurement, classical 'lidar' laser scanners, multi-layer lidars and novel concepts for lidars. Since 2019 professor for physics and sensors at the University of Applied Sciences in Konstanz. Main research interests: the interplay of physical sensing hardware with subsequent data analysis, optical sensors, especially time-of-flight lidar, commodity sensors as solution for industrial challenges, image processing.

Prof. Dr. Christian Hettich



Diploma in physics and doctoral degree (2002) at the University of Konstanz in the fields of quantum-optics and nano-optics. Post-doc at the Eidgenössische Technische Hochschule (ETH) Zürich and at the Niels Bohr Institut in Copenhagen. Systems engineer for illuminators of semiconductor lithography systems at Carl Zeiss SMT GmbH in Oberkochen. Systems engineer for time-of-flight cameras at ifm automotive GmbH in Kressbronn. Project leader for automated microscopes at Sensovation AG in Radolfzell. Since 2014 professor for physics, feedback control systems and metrology at the department of mechanical engineering of the University of Applied Sciences Konstanz. Member of the 'Institut für Naturwissenschaften und Mathematik' (INM) and the 'Institute for Optical Systems' (IOS). Main research interests: optical 3D measurements, spectroscopy, interferometry, holography, microscopy.

Prof. Dr. Bernd Jödicke



Study of physics at the University of Karlsruhe. Doctoral degree at Technical University Hamburg and University Karlsruhe in the field of high frequency technology. After that, industrial work at ABB Baden, Switzerland, as executive director for R&D. Since 1992 professor for applied physics at University of Applied Sciences Konstanz. Member of 'Institute for Applied Research (IAF)', 'Institut für Naturwissenschaften und Mathematik (INM)', 'Institute for Optical Systems (IOS) Konstanz' and 'Deutsche Lichttechnische Gesellschaft'. Head of laboratory for light engineering at HTWG. Main research activities in color and light measurements and color camera systems.

Prof. Dr. Burkhard Lehner



Diploma in computer science from University of Kaiserslautern, 2004. Doctoral degree in computer science from University of Kaiserslautern, 2008. Software development at Sirona Dental System GmbH, Bensheim, 2008-2013. Since 2013 professor for computer science at University of Applied Sciences Konstanz. Since 2014 member of the 'Institute for Optical Systems (IOS)'. Main interests: software development, computational geometry, optical 3D measurement (especially in dental CAD/CAM).

Prof. Dr. Georg Umlauf



Diploma in computer science from University of Karlsruhe, 1996. Doctoral degree in computer science from University of Karlsruhe, 1999. PostDoc at University of Karlsruhe and University of Florida, Gainesville, USA, 1999-2000. Software development and senior researcher at Tebis AG, Hamburg, 2000-2002. Assistant professor for geometric algorithms at University of Kaiserslautern, 2002-2009. Interim professor for computer graphics at University of Karlsruhe, 2009. Since 2009 professor for computer graphics at University of Applied Sciences Konstanz and head of the computer graphics lab. Since 2010 member of the 'Institute for Optical Systems (IOS)' and 'Institute for Applied Research (IAF)'. Main research interests: Computer graphics, geometric modeling (splines, subdivision), reverse engineering, physical simulations.

IOS STAFF

Professors

Rebekka Axthelm	vice-director IOS
Doris Bohnet	
Oliver Dürr	
Matthias Franz	
Hartmut Gimpel	
Christian Hettich	
Bernd Jödicke	
Burkhard Lehner	
Georg Umlauf	director IOS

Academic Staff

Tobias Birkle	until 2021
Daniel Dold	
Fabian Freiberg	until 2021
Dennis Grießer	
Matthias Herrmann	
Andreas Löffler	
Frieder Luginsland	
Mohamed Ramadane	since 01/08/2023
Carsten Röhricht	
Robert Straubmüller	
Benjamin Wagner	

PhD Students

Daniel Dold	
Dennis Grießer	
Martin Schall	until 2022
Matthias Hermann	
Mohamed Ramadane	

EXTERNAL FUNDING AND GRANTS

- Interreg, IBH-Lab Seamless Learning, BiLeSA und Cman, 2017-2020.
- Interreg "Alpenrhein-Bodensee-Hochrhein", eFlow - Softwaretool zur Berechnung von Personeneinflüssen und der damit verbundenen Ausbreitung von Infektionskrankheiten, 2021-2023.
- Interreg "Alpenrhein-Bodensee-Hochrhein", PInG - Validierte Simulation von Personenströmen und Infektionsrisiken zur Gestaltung sicherer Veranstaltungen und Räume, 2023-2026.
- Interreg, IoT Sustainability Lab, 2023-2027.
- Baden-Württemberg Stiftung, Ausschreibung 'Autonome Robotik', Forschungsprojekt 'Open SeeGround', 2023-2025.
- Small Digitalization Project, Erweiterung des cfdOF-Plugin in CAD-Software FreeCAD, cooperation with Peter Stein, Department of Mechanical Engineering, 09/2022-02/2023.
- Small Internationalization Projects (HAW International), German-French student projects on HVAC-Sensordatenanalyse, cooperation with Gioia Vago and Alain Jacquemard (ESIREM, Dijon), 09/2022-02/2024.
- MultiSenseLakePerceptor, Optisches Multisensorsystem für autonome Anwendungen auf Binnengewässern, January 2019 - September 2022, Baden-Württemberg Stiftung, cooperation with Prof. Reuter (ISD).
- MultiFlexInspect, Multifunktional-skalierbare generische InlineInspektion für flexible Fertigungsprozesse in vernetzten Produktionsanlagen, September 2017 – August 2021, BMBF.
- KLEVER, Intelligente Bildverarbeitung in der Materialmikroskopie durch maschinelles Lernen zur Objekt- und Fehlererkennung, September 2018 – August 2021, BMBF, cooperation with HAW Aalen.
- SfMSizeApp, Entwicklung einer erstmalig diskriminierungsfreien SizeRecommendation-App auf Basis neuartiger KI-basierten SfM-Verfahren, April 2023- December 2024, invest bw.
- DeepCarbPlanner Decarbonization of energy-intensive industries through smart sector coupling using AI-based probabilistic forecasts and operational management, May 2023 - April 2026, Carl-Zeiss-Stiftung, cooperation with Prof. Schubert (IAF).
- DeepDoubt Advancement of uncertainty measures to enhance transparency and traceability in deep learning, applied to optical systems, April 2020 - March 2023, BMBF.

Academic and Institutional Cooperation

- University of Queensland, Brisbane
- Humboldt-Universität zu Berlin
- ZARM (drop tower), Center of Applied Space Technology and Microgravity, Bremen
- DLR Institut für Raumfahrtssysteme Bremen
- University of Tübingen
- Max-Planck-Institute for Biological Cybernetics, Tübingen
- German Federal Office for Information Security (BSI), Bonn
- Universität Konstanz
- Zurich University of Applied Sciences, Winterthur
- FHV - Vorarlberg University of Applied Sciences, Dornbirn
- Statistical Learning and Data Science, LMU Munich
- Munich Center of Machine Learning (MCML)
- Heidelberg University, Interdisciplinary Center for Scientific Computing
- Department of Orthopaedics and Traumatology, Freiburg University Hospital

Industry Cooperations

- GFT Integrated Systems GmbH, Konstanz
- Bosch Sicherheitssysteme GmbH, Grasbrunn
- TWT GmbH, Stuttgart
- Zentrale Stelle für Informationstechnik im Sicherheitsbereich (ZITiS), München
- doubleSlash Net-Business GmbH, Friedrichshafen
- Johner Institut GmbH, Konstanz
- Data Migration International AG, Kreuzlingen
- Siemens Postal, Parcel & Airport Logistics GmbH, Konstanz
- Körber, Konstanz
- Formigas, Konstanz
- ShoeFitter bzw. Sooley, Konstanz
- OpenDress, Konstanz
- UGD, Konstanz

- Chromasens GmbH, Konstanz
- Baumer Inspection GmbH, Konstanz
- Subsequent, Konstanz
- FONDIIUM, Singen
- Sick AG, Waldkirch
- Holiday Check, Bottighofen
- RAYLYTIC GmbH, Germany
- Xanevo, Konstanz
- SÜDKURIER, Konstanz

Books

1. Dürr, O., Sick B., & Murina E. (2020). Probabilistic deep learning: With python, keras and tensorflow probability.

Journal Papers (Peer-Reviewed)

1. Dürr, O., Fan P-Y., & Yin Z-X. (2023). Bayesian Calibration of MEMS Accelerometers. IEEE Sensors Journal.
2. Herzog, L., Kook L., Götschi A., Petermann K., Hänsel M., Hamann J., et al. (2023). Deep transformation models for functional outcome prediction after acute ischemic stroke. Biometrical Journal. 65.
3. Berlin, C., Adomeit S., Grover P., Dreischarf M., Halm H., Dürr O., et al. (2023). Novel AI-Based Algorithm for the Automated Computation of Coronal Parameters in Adolescent Idiopathic Scoliosis Patients: A Validation Study on 100 Preoperative Full Spine X-Rays. Global Spine Journal.
4. Arpogaus, M., Voss M., Sick B., Nigge-Uricher M., & Dürr O. (2023). Short-term density forecasting of low-voltage load using Bernstein-polynomial normalizing flows. IEEE Transactions on Smart Grid.
5. Berlin, C., Adomeit S., Grover P., Dreischarf M., Dürr O., & Obid P. (2022). 140. Automated measurement technique for coronal parameters using a novel artificial intelligence algorithm: an independent validation study on 100 preoperative AP spine X-rays. The Spine Journal. 22, S74.
6. Kibii, J. E., Dreher A., Wormser P. L., & Gimpel H. (2022). Design and Calibration of Plane Mirror Setups for Mobile Robots with a 2D-Lidar. Sensors 22, no. 20: 7830.
7. Kook, L., Herzog L., Hothorn T., Dürr O., & Sick B. (2022). Deep and interpretable regression models for ordinal outcomes. Pattern Recognition. 122.
8. Hermann, M., Umlauf G., Goldlücke B., & Franz M. O. (2022). Fast and efficient image novelty detection based on mean-shifts. Sensors — Unusual Behavior Detection Based on Machine Learning.
9. Adomeit, S., Berlin C., Grover P., Dreischarf M., Halm H., Dürr O., et al. (2022). Validation study of an algorithm based on artificial intelligence for automated computation of coronal parameters on preoperative AP X-rays. Brain and Spine.
10. Pearse, G. D., Tan A. Y. S., Watt M. S., Franz M. O., & Dash J. P. (2020). Detecting and mapping tree seedlings in UAV imagery using convolutional neural networks and field-verified data. ISPRS Journal of Photogrammetry and Remote Sensing. 168, 156 - 169.
11. Herzog, L., Murina E., Dürr O., Wegener S., & Sick B. (2020). Integrating uncertainty in deep neural networks for MRI based stroke analysis. Medical Image Analysis. 65.

Book Chapters

1. Axthelm, R., Luppold S., & Moroff M. (2022). Crowd Management in der Lehre. Seamless Learning, Grenz- und kontextübergreifendes Lehren und Lernen in der Bodenseeregion. 123-132.
2. Axthelm, R. (2022). Mathematik mit digitalen Bildern sichtbar machen. Seamless Learning, Grenz- und kontextübergreifendes Lehren und Lernen in der Bodenseeregion. 133-145.

Conference Proceedings (Peer-Reviewed)

1. Bohnet, D., Meyer, S., Ramadane, M. (2023). Environmental Impact Assessment of IoT Devices: A Graph-based Approach. *Procedia Computer Science*. ISGTA 2023.
2. Bodai, T., Lembo V., Lembo V., Lee S-S., Ishizu M., & Franz M. (2023). Development and application of a climate emulator. EGU23.
3. Hoher, P., Reuter J., Dold D., Griesser D., Govaers F., & Koch W. (2023). Extended Target Tracking With a Lidar Sensor Using Random Matrices and a Gaussian Processes Regression Model. *International Conference on Information Fusion (FUSION)*. 1-8.
4. Hermann, M., Umlauf G., Goldlücke B., & Franz M. O. (2023). Incremental one-class learning using regularized null-space training for industrial defect detection. *16th International Conference on Machine Vision (ICMV)*.
5. Griesser, D., Umlauf G., & Franz M. O. (2023). Visual Pitch and Roll Estimation For Inland Water Vessels. *IEEE International Conference on Robotics and Automation (ICRA)*. 1961-1967.
6. Hermann, M., Umlauf G., & Franz M. O. (2022). Fast and memory-efficient independent component analysis using Lie group techniques. *International Conference on Curves and Surfaces*.
7. Hermann, M., Goldlücke B., & Franz M. O. (2022). Image novelty detection based on mean-shift and typical set size. *21th International Conference on Image Analysis and Processing, ICIAP*.
8. Hermann, M., Umlauf G., & Franz M. O. (2022). Large-scale independent component analysis by speeding up Lie group techniques. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP*.
9. Hermann, M., Griesser D., Gundel B., Dold D., Umlauf G., & Franz M. O. (2022). Target-less Lidar-camera registration using patch-wise mutual information. *International Conference on Information Fusion*.
10. Scharpf, P., Hong C. Lap, & Duerr O. (2021). Accelerating Active Learning Image Labeling Through Bulk Shift Recommendations. *2021 International Conference on Data Mining Workshops (ICDMW)*. 398-404.
11. Grunwald, M., Hermann M., Freiberg F., & Franz M. O. (2021). Biologically-inspired vs. CNN texture representations in novelty detection. *Applications of Machine Learning 2021*.
12. Sick, B., Hothorn T., & Dürr O. (2021). Deep transformation models: Tackling complex regression problems with neural network based transformation models. *Accepted for Proceedings of the 25th International Conference on Pattern Recognition (ICPR), Milan/Online, 2021*.
13. Arpogaus, M., Voß M., Sick B., Nigge-Uricher M., & Dürr O. (2021). Probabilistic Short-Term Low-Voltage Load Forecasting using Bernstein-Polynomial Normalizing Flows. *ICML 2021, Workshop Tackling Climate Change with Machine Learning, June 26, 2021*.
14. Hörtling, S., Dold D., Dürr O., & Sick B. (2021). Transformation models for flexible posteriors in variational bayes. *arXiv preprint*.
15. Griesser, D., Dold D., Umlauf G., & Franz M. O. (2020). CNN-Based Monocular 3D Ship Detection Using Inverse Perspective. *Global Oceans*.
16. Hake, F., Hermann M., Alkhatib H., Hesse C., Holste K., Umlauf G., et al. (2020). Damage Detection for Port Infrastructure by Means of Machine-Learning-Algorithms. *FIG Working Week 2020*.
17. Brach, K., Sick B., & Dürr O. (2020). Single Shot MC Dropout Approximation. *ICML Workshop on Uncertainty and Robustness in Deep Learning*.

Posters

1. Dold, D., Arpogaus M., & Dürr O. (2023). Deep probabilistic modelling for energy forecasting.
2. Baur, T., Albrecht, M., & Griesser, D. (2022). MultiSenseLakePerceptor, Forschungstag Baden-Württemberg Stiftung.
3. Hermann, M., Dold D., Umlauf G., & Dürr O. (2022). DeepDoubt - Improving uncertainty measures in machine learning to improve explainability and transparency. 2022 All-Hands-Meeting of the BMBF-funded AI Research Projects at Munich Center for Machine Learning.

Preprints

1. Bohnet, D., Vartziotis, D. (2023). An Angular Transformation of Triangles. arXiv:2307.14007.
2. Brach, K., Sick, B., & Dürr, O. (2023). Single-shot Bayesian approximation for neural networks. arXiv preprint arXiv:2308.12785.
3. Kook, L., Baumann, P. F., Dürr, O., Sick, B., & Rügamer, D. (2022). Estimating conditional distributions with neural networks using r package deeptrafo. arXiv preprint arXiv:2211.13665.
4. Dürr, O., Hörling, S., Dold, D., Kovylov, I., & Sick, B. (2022). Bernstein flows for flexible posteriors in variational bayes. arXiv preprint arXiv:2202.05650.

PhD Theses

1. Martin Schall, Multi-Dimensional Connectionist Classification: Segmentation-Free Handwriting Recognition, 2022.

Master Theses

1. Caroline Hess: Reproducing state-of-the-art Machine Learning-based schema matching algorithms, 2023.
2. Niklas Polster: Deep Learning für eine effektive und effiziente Text Detection in der Gepäck-Vision-Anwendung der Flughafenlogistik, 2023.
3. Dominim Vllasa: Training eines Deep Learning Algorithmus zur Texterkennung auf Basis künstlich erzeugter Zeilenbilder, 2023.
4. Matthias Asche: Rekonstruktion von Fassaden aus 3D-Punktwolken, 2023.
5. Frieder Luginsland: Bildbasierte Geolokation für Augmented Reality in Gebirgen, 2023.
6. Till Reitlinger: Training neuronaler Netze auf Basis eines synthetisch generierten Punktwolkendatensatzes zur Berechnung von Wandabmessungen, 2022.
7. Thimo Schaub: Punktwolkenklassifikation auf maritimen Daten, 2022.
8. Leonardo Labanca: Using Gaussian Process for Robust Temperature Forecast, 2022.
9. Tobias Pfrommer: Point Cloud Registrierung mithilfe von Fully Convolutional Geometric Features zur Rekonstruktion eines 3D Fußmodelles im Kontext einer Footwear E-Commerce Applikation, 2021.
10. Felix Hinderer: Supervised 3D semantic segmentation on Human-based point clouds with increased focus on realism, 2021.
11. Lukas Luschin: Analyse von latenten Gaußschen Mischmodellen zur Registrierung von Punktwolken, 2021.
12. Konstantin Christ: Lageerkennung der Wasseroberfläche durch semantische Segmentierung und Stereo Vision, 2021.
13. David Wiessner: Eine Untersuchung über den Einsatz von Ensemble-Methoden in der Anomalieerkennung, 2021.
14. Daniel Bencic: Analyzing keypoint detectors and descriptors learned on RGB-D images, 2021.
15. Erkan Emre: Modellierung von Energieprognosen mit Unsicherheiten durch maschinelles Lernen, 2023
16. Sonja Adomeit: Automatic Determination of Coronal Parameters for Spine AP X-ray Images in the Setting of Adolescent Idiopathic Scoliosis, 2022
17. Lukas Rohrer: Multimodal Anomaly Detection: Fusion of Deep and Shallow Models to Detect Anomalies on Sensor-Based Industrial Manufacturing Data, 2022
18. Fabio Brückle: Entwicklung und Implementierung eines Konzepts zur automatisierten Fehler-Ursachen-Analyse für Erprobungsfahrzeuge anhand ihrer Loggerdaten, 2022

19. Maksim Sandybekov: Accelerating Bayesian Neural Networks in the context of active learning, 2021
20. Ivonne Kovylov: Prediction of melanoma types using semi-structured Bayesian deep learning models, 2021
21. Leon Sick: Active Learning for Grocery Object Recognition in the Context of Smart Cooling Devices, 2021
22. Barbara Opitz: Erstellung eines Prognosemodells des Energiebedarfs einer energietechnischen Anlage im Produktionswerk, 2021
23. Tim Hardenacke: Usage of reinforcement learning for the generation of counter-UAV trajectories, 2021
24. Stephan Hörting: Modeling of flexible Posterior Distributions for Bayesian Neural Networks
25. Steffen Kaus: Unsicherheitsquantifizierung in Deep Neural Networks unter Verwendung von Deep Ensembles und Teacher-Student Verfahren, 2021
26. Marcel Arpogaus: Probabilistic short-term load forecasting, 2021
27. Simon Christofzik: Modellierung von Schwarmverhalten - Untersuchung von Methoden zur Parameter Approximation, 2020

Bachelor Theses

1. Elena Totzki: Mögliche Methoden zur Automatisierung der Verschlagwortung von regulatorischen Dokumenten, 2023.
2. Tim Geißendörfer: Entwicklung eines Konzepts für eine cloudbasierte Data Factory in aws, 2023.
3. Corinna Kurz: Datenanalyse und Big Data Mechanismen anhand einer PV-Anlage unter der Einbeziehung von Meteodaten, 2023.
4. Max Schroeter: Untersuchung der Reproduzierbarkeit und Vorhersagbarkeit von Simulationsergebnissen eines Personenstrommodells mittels Machine Learning- Methoden (Teil 1), 2023.
5. Maren Mutter: Untersuchung der Reproduzierbarkeit und Vorhersehbarkeit von Simulationsergebnissen eines Personenstrommodells mittels Machine-Learning Methoden (Teil 2), 2023.
6. Jan-Niclas Serr: Automated data gathering and analysis for autonomous driving, 2023.
7. Soniesen Ratnasingam: Automatisierte Prädiktion von Budgetreports, 2023.
8. Niklas Strittmatter, Untersuchung des Multi-View 3D-Rekonstruktionspotentials von Plenoxels und FastNeRF, 2023.
9. Simon Warmers: Transfer learning on inland waterway data, 2023.
10. Niclas Graf: Registrierung von Fuß-Scans mit Raumdeformationsverfahren, 2023.
11. Bernhard Gundel: Implementierung und Analyse von Partikel-Schwarm-Optimierung im Rahmen von Kamera-Lidar-Kalibrierung, 2022.
12. Ismail Zeybek: Datenanalyse im Kontext der Gebäudeautomation unter Verwendung von Anomaliedetektionsalgorithmen, 2022.
13. Florian Halter: Multimodale Registrierung von 3D-Fuß-Scans auf Basis der aktuellen iPhone-Sensorik, 2022.

14. Cesim Keskin: Evaluation and Implementation of a lossless dataformat for 2,5-D data using HEIF, 2022.
15. Julia Hansi: Automatisches Erkennen von Polygonzügen auf Grundrissbildern, 2022.
16. Yvonne Wehrlein, Implementierung von Novelty Detection auf Basis von Komponentenanalyse unter Verwendung von vortrainierten neuronalen Netzen, 2022.
17. Andreas Ly: Comparing non-parametric and parametric anomaly detection methods for industrial defect detection, 2022.
18. Caroline Wenig: Euklidische kürzeste Wege für die Büroplanung, 2021.
19. Matthias Deberling: Detektion von Kalibrierobjekten in Punktwolken und Kamerabildern, 2021.
20. Felix Mayer, Geometrische Kalibrierung zwischen IMU/GPS und LiDAR, 2021.
21. Carsten Röhricht: Farbkalibrierung mit spektrometrischen Referenzdaten für den Inline-Digitaldruck, 2021.
22. Yue Zhang: Entwicklung einer webbasierten Umfrage von typischen Personenbewegungen in öffentlichen Räumen am Beispiel Supermarkt zur Erzeugung entsprechender Wahrscheinlichkeitsverteilungen, 2022
23. Simon Awißus: Planung und Programmierung einer physischen Zwischeninstanz im Druckprotokoll und Verarbeitung der Bilddateien mit Sensordaten, 2022
24. Nik Steinbrügge: Entwicklung einer Software zur 2D-Modellierung von Räumen und Berechnung einer Delaunay-Triangulierung, 2021
25. Markus Eckhardt: Simulation und Validierung verschiedener Modellansätze für Migrationswege von Fischlarven, 2021
26. Aaron Moser, Simon Rufener: Physical Modelling and Simulation of Pedestrian Flow Using the Finite Element Method, 2020
27. Fabian Beck: Anwendung von Data Analytics Methoden im modernen Presswerk, 2022
28. Mark Freibauer: Feature Selection und Model-Testing für eine Machine-Learning basierte Churn-Prediction, 2021
29. Dominim Villasa: Entwicklung und Evaluation eines Algorithmus unter der Verwendung von neuronalen Netzen, zur Vorschlagsgenerierung beim Referenzieren von Daten, 2020
30. Ma Junyi: Evaluation of Different Uncertainty Measures in Deep Learning for Self-driving Car Applications, 2020
31. Christoph Hänsel: Potenzial von Mustererkennung bei vergangener Vakanzprüfung zur Preistransparenz für Kunden, 2020
32. Daniel Dietze: Accelerating catalog data integration via Candidate Recommendation, 2020
33. James E. Kibii: Modelling and Control of a Small Vehicle with LiDAR & Ground Measurements, 2022

Student Projects

1. Björn Johannisson, Dennis Todoric, Kevin Bauer, Reif Dervisbegovic: pFlow - Entwicklung einer Personenzähler-App, 2020
2. Anna Bernauer, Caroline Weng, Robin Rudolf, , Nik Steinbruegge, Julia Burri, Helen Nganga: Entwicklung einer Verleihplattform, 2020
3. Andre-Christian Goehring, Karsten Huber, Marc Freibauer: BiLeSA - Lernapperweiterung mit dem Thema Komplexe Zahlen, 2021
4. Daniel Oezyurt, Jan-Niclas Serr, Simon Warmers, Alina Goettig, Rebecca Braun: pFlow-Visualisierung mit Virtual Reality, 2022
5. Julian Dietsche, Markus Beez, Julian Heggenberger, Sonja Adomeit, Sebastian Rätzer: Pandemiemodellierung, 2021
6. Miles Müller, Marco Traber, Fabian Fritzen, Nico Kriessler, Patrick Dec: Untersuchung Heizeinsparpotential eines intelligenten Thermostaten am Beispiel eines Hörsaals, 2023.
7. Alex Jauch-Walser, Elias Pampuch, Timm Schimmel, Philipp Schneider, Nils Herhaus, Amanda Schaal, Michelle Müller: Entwicklung eines WIN-Chatbots, 2023.
8. Mirjam Abele, Aida Vehapi, Elida Vehapi, Viktoria Stiem: Fourieranalyse HVAC-Sensordaten, german-french project (funded by DAAD), 2023.
9. Benjamin Bruenau, Joel Merath, Jan Schmidt, Lilian Krieg Carreño, Khadidja Kebaili: Ein Chatbot für den Prüfungsausschuss, 2023.
10. Silas Brosi, Kai Gaiser, Karoly Hegyi, Daniel Sturm, Sebastian Unger, Human: Action Recognition in Mannschaftssportarten, 2023.
11. Tim Geißendörfer, Jens Ullrich, Julian Post: Multivariate HVAC-Sensordatenanalyse, german-french project (funded by DAAD), 2022.
12. Elena Totzki, Sarah Blase, Simon Oberholzner, Julius Felber, Stefan Lam: Erweiterung des openFOAM-plugin cfdOF in CAD-Freeware, in cooperation with Department of Mechanical Engineering, 2022.
13. Thomas Hummler, Lukas Peter, Magnus Mink, Sven Gemeinhardt: Entwicklung von Webtutorials für FE-Netzoptimierer, 2022.
14. Alexander Geis, Dennis Zipfel: Entwicklung und Realisierung eines akustischen Getriebe-Prüfstands, 2023

Research Activities_____

BiLeSA: Mit digitalen Bildern Mathe besser verstehen

Benjamin Dums, Stefan Ankele, Rebekka Axthelm

Das Projekt BiLeSA¹ ist Teilprojekt des IBH-Labs "Seamless Learning", welches sich folgender Forschungsfrage widmet:

"Wie muss Seamless Learning für die Bodenseeregion mit ihren Spezifika in Bezug auf Didaktik, Technik und Lerner / Lehrer konzeptualisiert werden, um nahtloses lebenslanges Lernen in Aus- und Weiterbildung zu ermöglichen?"

BiLeSA_LIN und BiLeSA_PDE

Im Projekt BiLeSA wird sich dem Bruch "Theorie-Praxis" bei der Lehre der Mathematik gewidmet, mit dem Ziel, die Brücke zwischen Theorie und Praxis so zu gestalten, dass die Neugierde auf das Relevante nicht verloren geht. "Mathematik sehen", "mathematische Prozessbeschreibung verstehen", "Lernen durch Erfahren" sind die Zielsetzungen von BiLeSA. Wir machen uns dabei die Möglichkeiten der Bildverarbeitung zunutze, um Methoden der Mathematik sichtbar zu machen. Zu diesem Zweck wurde eigens eine App entwickelt. Der didaktische Mehrwert besteht darin, dass die betrachteten Prozesse nicht einfach per Klick in sinnvoller Weise ausgeführt werden können, sondern dass zur Bedienung zunächst mathematische Kenntnisse erarbeitet werden müssen.

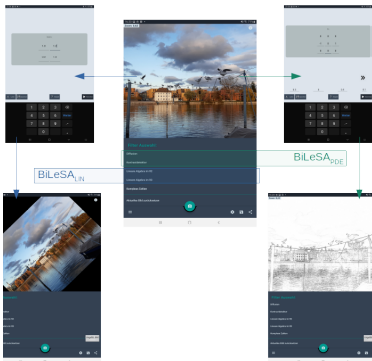


Figure 1: Übersicht der BiLeSA-App

Das führt einerseits zu einer gesteigerten intrinsischen Motivation der Lernenden, sich mit den Lehrinhalten zu befassen und andererseits bietet es die Möglichkeit, in einen von vielen Anwendungsbereichen, nämlich der Bildverarbeitung, Einblick zu gewähren.

Im Projekt wurden zwei Themenbereiche aufgegriffen, nämlich lineare und affin lineare Abbildungen (BiLeSA_LIN) und partielle Differentialgleichungen am Beispiel von isotroper und anisotroper Diffusion (BiLeSA_PDE). Abbildung 1 zeigt das Hauptmenü der App.

Ergebnisse und Fazit

Es wurden zwei Evaluationen in je zwei Schritten durchgeführt: Jeweils eine Pre- und eine Post-Befragung vor und nach der Durchführung dieser Lehrinheit. Die erste Evaluation fand im Sommersemester 2019, die zweite im Sommersemester 2020 statt. Die Evaluationen erlauben einen Einblick, sowohl in die Entwicklung jeweils einer Studierendengruppe (Pre- und Posttest), als auch in die Entwicklung des Unterrichtskonzepts selbst.

Wir sehen die Verwendung von BiLeSA im Lehrkonzept und die damit verbundene Möglichkeit, die Inhalte auf vielfältige Weise erarbeiten zu können, als ein innovatives Konzept mit großem Erfolgspotenzial an, um Theorie und Praxis von der ersten Stunde an näher zusammenzubringen. Das größte Problem bei diesem Lehrkonzept war der Konzeptwechsel im Laufe der Gesamtveranstaltung. In der Tat empfehlen wir die komplette Lehrveranstaltung in dieser Art zu konzipieren, da sonst ein neuer Bruch entsteht, nämlich der Bruch des Lehrkonzepts innerhalb einer Vorlesung.

Literatur

[1] Axthelm, R. (2022). Mathematik mit digitalen Bildern sichtbar machen. Seamless Learning, Grenz- und kontextübergreifendes Lehren und Lernen in der Bodenseeregion. 133-145.

¹<https://www.ios.htwg-konstanz.de/modelling-and-simulation-BiLeSA>

Cman: Crowdmanagement in der Lehre

Sebastian Rätzer, Rebekka Axthelm

“Chaos beim Crowd-Management” (TAZ),
“Tödliche Massenpanik in Schanghai” (ka-news),
“Als ein Fußballstadion zur Todesfalle wurde”
(Die Welt)

Solche und ähnliche Schlagzeilen erfahren wir fast täglich aus den Nachrichten. Jeder kennt das unbehagliche Gefühl, wenn man dichtgedrängt vor der Konzerttribüne steht oder in einer Großstadt-U-Bahn in der rush-hour unterwegs ist und kaum Platz zum Atmen hat. Je größer eine Veranstaltung ist desto größer ist die Bedeutung von gutem Crowd Management. Crowd Management ist “die systematische Planung für, und die Lenkung von, geordneten Bewegungen und Ansammlungen von Menschen”²

Crowd Management ist Bestandteil des Curriculums im Studiengang “BWL-, Messe-, Kongress- und Eventmanagement” der DHBW-Ravensburg. Bislang stand eine Software, die Massenbewegungen im Raum simuliert für die Lehre nicht zur Verfügung. Dies war die Motivation, eine entsprechende Simulationssoftware zu entwickeln. Die Software basiert auf einem neuen Modellansatz altbekannter Gleichungen, die dann aber noch für die speziell anvisierten Anwendungsfelder erweitert und angepasst wurden. Sie ist webbasiert aufgebaut und steht jederzeit und kostenfrei unter [2] zur Verfügung. Das Lernobjekt hat eine Brücke zwischen Theorie und Praxis geschlagen, da es in der Praxis durchaus nicht unüblich ist, mit einer Simulationssoftware zu arbeiten, bzw. entsprechende Berechnungen bei Anbietern anzufordern.

Die numerischen Berechnungen, die die Software durchführt, basiert auf mathematischen Gleichungen, wie sie im Grundlagenstudium an der ZHAW gelehrt werden. Die Aufbereitung der Software als Lernobjekt steht im direkten Zusam-

menhang zu den Unterrichtseinheiten, die sich auf Softwareentwicklung beziehen, wie sie an der HTWG unterrichtet werden. Hierin sind die thematischen Anknüpfungspunkte zu den drei beteiligten Hochschulen ersichtlich. Im Sinne des grenzenlosen Lernens, wie es im IBH-Lab Seamless Learning definiert ist, werden in diesem Projekt die Übergänge in den Brüchen “Theorie - Praxis”, “Theorie - Realität” und “zwischen Taxonomiestufen” adressiert.

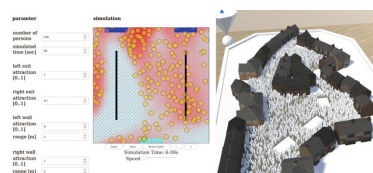


Figure 2: links: Computersimulation von Personenströmen, rechts: Visualisierung der Berechnungen in einer Virtual Reality (Erarbeitet im studentischen Teamprojekt an der HTWG)

Abbildung 2 zeigt ein Beispiel wie sich bevorzugte Ausgänge und Bereiche im Raum auf die Dichteverteilung bei einer Evakuierung auswirken. Studierende können theoretische Überlegungen zu entsprechenden Raumplanungen durch simulierte Werte erweitern.

Literatur

[1] Axthelm, R., Luppold, S., & Moroff, M. (2022). Crowd Management in der Lehre. In B. Dilger, J. Erlemann, C. Müller, & C. Rapp (Hrsg.), Seamless Learning, Grenz- und kontextübergreifendes Lehren und Lernen in der Bodenseeregion (S. 123-132). Springer.

[2] <http://www-home.htwg-konstanz.de/~raxthelm/Cman/>

² John F. Fruin, Quelle: <https://www.siflux.com/was-ist-crowd-management>

eFlow: Softwaretool zur Berechnung von Personenflüssen und der damit verbundenen Ausbreitung von Infektionskrankheiten

Benjamin Wagner, Robert Straubmüller, Doris Bohnet, Rebekka Axthelm

Ziel des Projektes eFlow ist die Entwicklung eines Softwaretools, mit dessen Hilfe Personenströme in Räumen und gleichzeitig die damit verbundene Ansteckungswahrscheinlichkeit simuliert werden können. Das Projekt wurde von der Interreg gefördert mit der Auflage, den Fokus auf Praxistransfer zu legen. Das Ergebnis ist ein öffentlich zugänglicher, kostenfreier und webbasierter Prototyp, der Anwender*innen die Möglichkeit gibt, verschiedene Raumgestaltungen auf Vorkommnisse hoher Personendichten oder Ansteckungsraten zu prüfen.

Mathematische Modelle

Die zugrundeliegende mathematische Beschreibung der Personenströme in Zeit $T \subset \mathbb{R}_0^+$ und Raum $\Omega \subset \mathbb{R}^2$ erfolgt über die Dichteverteilung $\varrho : T \times \Omega \rightarrow [0, 1]$ und das Geschwindigkeitsfeld $u : T \times \Omega \rightarrow \mathbb{R}^2$. Die Berechnung dieser beiden Größen beruht auf der Masseerhaltung, die durch die Kontinuitätsgleichung

$$\varrho_t + \nabla \cdot (\varrho u) = 0$$

gewährleistet wird und dem Geschwindigkeitsfeld

$$u = -f(\varrho) \frac{\nabla \Phi}{|\nabla \Phi|},$$

der am schnellsten zum Ziel führt und durch die Lösung der Eikonalgleichung

$$|\nabla \Phi| = \frac{1}{f(\varrho)}$$

berechnet wird. Dazu kommen passende Randbedingungen an Wänden, Ausgängen und gegebenenfalls Eingängen des betrachteten Raumes (siehe [1]). Abbildung 3 zeigt ein Beispiel von Dichteverteilung und Geschwindigkeitsfeld in einem Raum mit zwei Ausgängen und einer Säule als Hindernis im Raum.

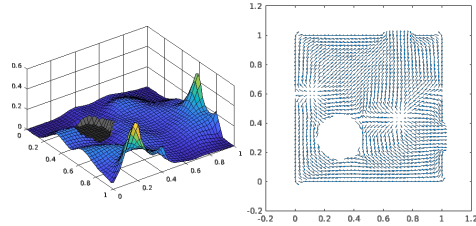


Figure 3: links: Dichte ϱ , rechts: Geschwindigkeitsfeld u

Die Infektionsausbreitung wird agentenbasiert beschrieben, wobei Agenten durch $x_i(t) \in \Omega$, $t \in T$ beschrieben werden mit der Bewegungsvorschrift $\dot{x} = u$. Wie infizierbare Agenten (S) durch infektiöse (I) zu infizierten, noch nicht infektiösen (E) angesteckt werden basiert auf dem klassischen SEIR-Modell

$$\begin{aligned} S' &= -\beta S I \\ E' &= \beta S I - \alpha E \\ I' &= \alpha E - \gamma I \end{aligned}$$

Dabei beschreibt α die Übergangsrate, β die Transmissionsrate (Ansteckungswahrscheinlichkeit) und γ die Erholungsrate (Kehrwert der mittleren Latenzzeit). Abbildung 4 zeigt ein Beispiel der Berechnungen für ein SEIR-Modell, bei dem es eine Unterscheidung zwischen Infektiösen und Infizierbaren gibt.

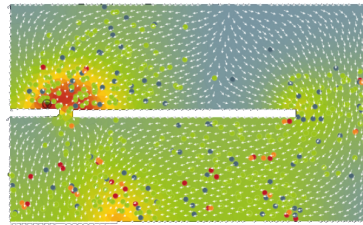


Figure 4: Agenten basierend auf der berechneten Dichteverteilung. rot: infektiös, orange: infiziert, grün: infizierbar, blau: nicht infizierbar

Ergebnis

Das Softwaretool flowculate enthält integrierte Settings von verschiedenen Räumungsszenarien zu demonstrativen Zwecken und bietet die Möglichkeit, ohne großen Einarbeitungsaufwand, eigene Szenarien aufzusetzen. Die Simulation der speziellen Situation lässt sich mit Hilfe verschiedener Parameter steuern: Anzahl Personen, Anteil wahrscheinlich erkrankter Personen, Kennzahlen der Infektionskrankheit wie Infektionsrate, Übertragungsart, Größe des Ansteckungsbereichs, etc. Als Ergebnis liefert die Simulation die statis-

tisch geschätzte Anzahl der Neuinfektionen unter diesen Bedingungen während eines bestimmten Zeitabschnitts und eine Simulation der Ansteckungsausbreitung.

Literatur & Quellennachweis

[1] Axthelm, R. (2016) Finite Element Simulation of a Macroscopic Model for Pedestrian Flow. In V.L. Knoop, & W. Daamen (Hrsg.), Traffic and Granular Flow '15 (pp. 233–240).

[2] Softwareseite: <http://www.flowculate.eu>

PlnG: Validierte Simulation von Personenströmen und Infektionsrisiken zur Gestaltung sicherer Veranstaltungen und Räume

Benjamin Wagner, Robert Straubmüller, Doris Bohnet, Rebekka Axthelm

PlnG ist die Fortsetzung des Interreg-Projekts eFlow und wird ebenfalls von Interreg gefördert. Das Ziel des Projekts PlnG ist die Weiterentwicklung der Software flowculate in den Kategorien "Verbesserung der User Experience", "Erweiterung der zugrunde liegenden mathematischen Modelle" und "Validierung und Verifikation der Berechnungen".

User Experience

Abbildung 5 zeigt einen Ausschnitt der Bedienelemente von flowculate. Die App bietet noch nicht alle wählbaren Parameter. Welche für User relevant und nützlich sind wird in mehreren Workshops mit unseren Praxispartnern, die aus verschiedenen Bereichen kommen, evaluiert und die Software entsprechend erweitert.

Erweiterung der mathematischen Modelle

Die Berechnung von Infektionsausbreitungen erfolgt in flowculate durch ein agentenbasiertes Modell. Ein Kontinuummodell, das durch Projektpartner der ZHAW³ erstellt wurde soll nun integriert und validiert werden.

Des weiteren wird untersucht ob und inwiefern sich die Belüftung eines Raumes auf das Infek-

tionsgeschehen auswirken. Gleichzeitig soll untersucht werden ob und inwiefern sich die Bewegung von Personen im Raum auf Luftbewegungen auswirken.

Verifikation und Validierung

Um eine Software mit stabilen Berechnungen zu erzeugen, sind Verifikationen der numerischen Berechnungen und theoretische Untersuchungen dazu unumgänglich. Validierung mit Realdaten belegen die Güte der Modellierung.



Figure 5: Screenshot der Software flowculate

³Zürcher Hochschule für Angewandte Wissenschaften, Winterthur

Energieeinsparpotential durch intelligente Thermostate

Miles Müller, Patrick Deckert, Marco Traber, Fabian Fritzen, Nico Kriessler, Sonja Meyer, Doris Bohnet

Dieses studentische Projekt hat untersucht, in welcher Weise durch Nutzung von intelligenten Thermostaten Heizenergie in Bestandsgebäuden eingespart werden kann. Motiviert war dieses Projekt durch den Werbeslogan der Bosch Smart Home, die eine Einsparung von 20% der Heizenergie durch Nutzung ihrer Thermostate verspricht. Entsprechend wurden uns von Bosch Smart Home Thermostate zur Verfügung gestellt. Da nur eine genauere Analyse des Fraunhofer Instituts für Einfamilienhäuser und Geschosswohnungen bekannt ist, sollte die Anwendung in einem Hochschulgebäude untersucht werden. Als Beispielraum diente ein Hörsaal im A-Gebäude der HTWG, das seine Ursprünge in den 1950-60er Jahren hat und durch eine relativ schlechte Isolierung charakterisiert ist.

Die übergeordnete Fragestellung dieses Projektes war, ob das Einsparpotential die Investition in die intelligenten Thermostate lohnt, für die zusätzlich ein Gateway, Internet Router und ein mobiles Endgerät notwendig sind und die wartungsintensiver als herkömmliche Thermostate sind. Diese Frage wird weiter im Detail und zusätzlichen Fallstudien im Rahmen des Forschungsprojekts IoT Sustainability Lab verfolgt.

Im ersten Schritt wurden die Funktionalitäten und externen Steuerungsmöglichkeiten des Thermostats untersucht und anschließend verschiedene Heizszenarien definiert. Das Bosch Produkt Radiator Thermostat II in Kombination mit dem Smart Home Controller II bietet nicht vollumfänglich die Möglichkeit zur intelligenten Steuerung, da insbesondere kein IFTTT unterstützt wird sowie Geofencing, Integration von Wetterdaten und Lernfähigkeit fehlen. Als Referenzszenario diente die von dem Facility Management der HTWG implementierte Heizstrategie, die sich bereits um eine große Einsparung auch unter Komforteinbußen bemüht.

Es wurde ein thermisches Modell in Simulink für den Raum erstellt, das mit den Raumdaten des Hörsaals, der technischen Spezifikation der Heizung sowie den Wetterdaten des Deutschen

Wetterdiensts versorgt wurde und mit dem anschließend verschiedene Szenarien simuliert werden konnten. Das Modell wurde anhand eines vorhandenen thermischen Modells in der Fakultät Maschinenbau plausibilisiert. Leider standen keine längeren Messreihen zur Validierung zur Verfügung. Insbesondere die Trägheit der Thermostatregelung musste geschätzt werden.

Es wurden folgende Szenarien für die Heizperiode Oktober 2022 - März 2023 ausgewertet:

1. Referenzszenario: aktuelle Heizregelung der HTWG. Abkühlung auf 15 Grad über Nacht, Ausschaltung am Wochenende, sonst von 6 bis 18 Uhr 19 Grad.
2. 1 Grad Szenario: Solltemperatur von 19 auf 18 Grad abgesenkt, sonst wie Referenzszenario. Keine smarte Thermostate notwendig.
3. Smart Fenster-Szenario: Intelligentes Thermostat bemerkt die Öffnung eines Fensters und schaltet Heizung runter. In der Simulation wurde angenommen, dass je 5 Minuten pro Stunde während Vorlesungszeiten gelüftet wird.
4. Smart Anwesenheitsszenario: Kopplung des Thermostaten an einen Anwesenheitssensor. Heizung wird nur bei detektierter Anwesenheit eingeschaltet.
5. Anwesenheitsszenario: Thermostat wird entsprechend der normalen Vorlesungszeiten gesteuert. Da die tatsächliche Belegung stark vom Vorlesungsplan LSF abwich, mussten relativ großzügig eine Belegung angenommen werden. Für dieses Szenario ist kein intelligentes Thermostat notwendig.
6. Wetterszenario: Heizregelung berücksichtigt die Wetterprognose und schaltet nicht an bzw. aus, wenn warmes Wetter vorhergesagt ist.

Ergebnisse

Insgesamt hat sich gezeigt, dass durch das Abschalten der Heizung bei Abwesenheit oder länger geöffneten Fenstern hohe Einsparpotentiale von 12% bzw. 16% bezogen auf das Referenzszenario erzielen lassen. Im Vergleich hierzu bringt eine pauschale Absenkung um 1 Grad eine Ersparnis von 7%. Das smarte Fenster-szenario geht allerdings auch mit einem Komfortverlust einher, da der Raum bei längerem Lüften abkühlt und Zeit braucht, um anschließend wieder die Solltemperatur zu erreichen. Dies ist in Fig. 6 und Fig. 7 an den deutlichen Piks nach Fensteröffnung oder bei Beginn der Belegung ersichtlich. Da Hörsäle anders als private Räumlichkeiten unregelmäßig und im Mittel deutlich weniger genutzt werden, erscheint der Einbau von intelligenten Thermostaten, die an Fenster- und Anwesenheitssensoren gekoppelt sind, vielversprechend. Erfahrungsberichte sprechen teils von hohem Wartungsaufwand von Smart Home-Komponenten, den man vor einer Umsetzung in der Fläche zunächst im Rahmen eines Pilotversuchs abschätzen und testen müsste. Wenn die elektronischen Vorlesungspläne mit der tatsächlichen Belegung der Räume übereinstimmen würden, wäre eine optimierte Heizregelung auch ohne intelligente Thermostate und ihrer zusätzlichen IT-Infrastruktur möglich. Noch besser wäre selbstverständlich, wenn Lehrende beim Verlassen von Räumen eigenständig die Heizung abschalteten oder zumindest die Fenster schlossen. Dies wäre sicherlich

der kosteneffizienteste Weg, der aber erfahrungsgemäß auch der am schwierigsten umzusetzende ist.

Dieser Analyse des Einsparpotentials werden im weiteren Verlauf des IoT Sustainability Labs die Kosten der IoT-Lösung aus Herstellung, Installation, Wartung, Betrieb und Entsorgung gegenübergestellt, um eine Bewertung des ökologischen Nutzens entlang des Lebenszyklus zu erhalten.

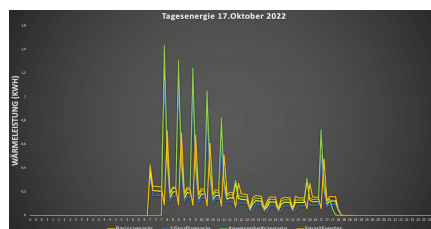


Figure 6: Ergebnisse.

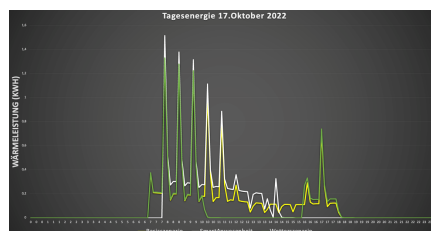


Figure 7: Ergebnisse.

Environmental Impact Assessment of IoT Devices: A Graph-based Approach

Mohamed Ramadane, Sonja Meyer, Doris Bohnet

This project is part of the IoT Sustainability Lab which aims to assess the positive and negative environmental impact of IoT solutions in buildings. As first step, we are trying to define a flexible model of IoT solutions which allow a faster assessment using ML methods in the near future.

The proliferation of the Internet of Things (IoT) has enriched modern life, but their increasing ubiquity raises concerns about environmental impact. To address this, comprehensive Life Cycle Assessments (LCAs) of IoT products, which have historically been manual, costly, and time-consuming, are vital. Noting the recurring nature of core components in IoT devices, such as CPUs and sensors, we propose to use graphs and machine learning to simplify and scale LCA estimations for IoT products. This paper introduces a novel approach to representing IoT devices as graphs with specific component characteristics and interconnections. Applied to a preliminary dataset of smart home IoT devices, the methodology unveils insights into structural similarities using a composite kernel approach. This initial phase lays the groundwork for the machine learning component. The integration of machine learning planned as part of ongoing research, provides a pathway for efficient and timely ecological assessments, ensuring that the rapid growth of IoT aligns with sustainable practices.

Results

In this preliminary analysis, we applied both the Weisfeiler-Lehman (WL) subtree kernel and the Graphlet kernel to analyze the structural characteristics of various devices, represented as graphs. Following the kernel computations, multidimensional scaling (MDS) was employed to visualize the similarities between the devices based on the computed kernel matrices. The MDS value represents the dissimilarity between two devices, calculated as one minus the kernel value. The Graphlet

kernel analysis revealed the most significant similarity between the Temperature+Humidity Tracker and the Water Leak Detector, with a score of 0.72937, suggesting a substantial overlap in their subgraph structures. This high score indicates that these devices might share similar components or design elements. The second highest similarity in the Graphlet kernel was observed between the pH/Temp/Light/Water Level Sensor and the Temperature Wearable Tracker, with a score of 0.61699, pointing to a notable similarity in their subgraph patterns. Visual inspection of those graphs suggests that they share many two stars/path graphs and one 3-cycle graph connecting the double compute units with the battery.

In the Weisfeiler-Lehman (WL) subtree kernel, after five iterations, the highest similarity was again between the Temperature+Humidity Tracker and the Water Leak Detector, but with a lower score of 0.43. This indicates a strong resemblance in their overall architecture, suggesting that these devices share similar connectivity patterns or core structural designs. The pH/Temp/Light/Water Level Sensor and the Temperature Wearable Tracker also showed a high similarity in the WL kernel with a score of 0.43, reinforcing the notion that they share similar overall architectural designs in addition to their subgraph structures.

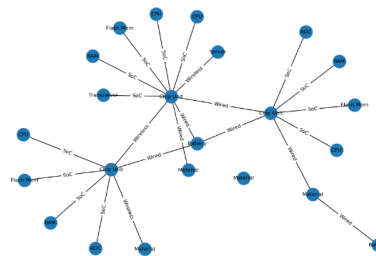


Figure 8: Graph representation of a smart plant watering system.

Future Work

Using a set of Life Cycle Impact Assessment (LCIA) indicators, our planned approach incorporates applying machine learning models to a graph-based representation of IoT devices. The emphasis remains on developing suitable graph kernels and optimizing a Multiple Kernel Learning (MKL) model for precise LCIA indicator prediction. This initiative is directed towards assessing the environmental footprint of IoT devices, contributing to sustainable design and manufacturing practices. Integral to this approach will be the it-

erative refinement of our final graph model based on the accuracy of the models. This refinement process is key to enhancing our predictions, ensuring they are robust and reliable.

Literatur

[1] Ramadane, M., Meyer, S., Bohnet, D., Environmental Impact Assessment of IoT Devices: A Graph-based Approach. *Procedia of Computer Science*, Elsevier, 2023.

Entwicklung eines neuronalen Netzes zur Vorhersage von Simulationsergebnissen eines Personenstrommodells

Yvonne Wehrlein, Doris Bohnet

Als Teil des Interreg-Projekts Ping untersuchen wir Methoden für datengetriebene Ersatzmodelle des Personenstrommodells, das auf einem System von zwei gekoppelten, partiellen Differentialgleichungen beruht. Ziel ist einerseits die Kopplung des Ersatzmodells an Echtzeitdaten von Besucherfrequentierung von Räumlichkeiten (z.B. eines Museums), aus denen dann die wahrscheinliche zeitliche Entwicklung der Besucherdichte berechnet wird, und andererseits die Optimierung mittels des Ersatzmodells von Raumgeometrien (z.B. Platzierung von Hindernissen oder Türen) für verschiedene Ziele des Besuchermanagements (z.B. schnellst mögliche Evakuierung oder best mögliche Verteilung).

Problemstellung

Gegeben eine zweidimensionale Raumgeometrie, die Wände, Aus- und Eingänge sowie Hindernisse kennt, sowie eine anfängliche Verteilung von Personen im Raum und deren Ziel (z.B. auf schnellstem Weg den Ausgang zu erreichen) soll die zeitliche Entwicklung der Anzahl Personen pro Quadratmeter, kurz Dichte, sowie der Geschwindigkeit der Personen berechnet werden. Die aktuelle Simulationssoftware flowculate, die in verschiedenen Projekten entstanden ist, basiert - wie oben erwähnt - auf der numerischen Berechnung der Lösung zweier gekoppelter, partieller Differentialgleichungen, die zum einen die Geschwindigkeit, die von der Personendichte abhängt, andererseits die zeitliche und räumliche Entwicklung der Dichte entsprechend der Geschwindigkeit berechnen. Eine Besonderheit dieser mathematischen Modellierung ist, dass die Personen nicht einzeln, sondern die Gesamtheit der Personen ähnlich einer Flüssigkeit mit einer gewissen Dichte und Geschwindigkeit modelliert werden.

Das neuronale Netz soll als Ersatzmodell dieses mathematisch-physikalischen Modells dienen und entsprechend anhand desselben Inputs (Raum-

geometrie, Dichte zum Zeitpunkt 0, funktionaler Zusammenhang Dichte-Geschwindigkeit) eine Zeitreihe der räumlichen Dichte- und Geschwindigkeitsfelder als Output zurückgeben.

Modell

Die In- und Outputs lassen sich mathematisch als Tensoren $\mathbb{R}^{T \times n \times m}$ beschreiben, wobei n, m die Raumdimensionen und T die Anzahl an betrachteten Zeitschritten bezeichnen. Da das Netz gleich große Bilder (oder Matrizen) als Input benötigt, werden die 2D-Raumgeometrien zunächst bzgl. ihrer längeren Seite skaliert und mit einem $n \times m$ -Gitter versehen. Das Dichtefeld $\rho(t, \cdot, \cdot)$ wird zu jedem Zeitschritt $t \in [0, T]$ auf den Gitterpunkten ausgewertet und entsprechend als $n \times m$ -Matrix übergeben. Dabei wird die Dichte mittels der maximal erlaubten Dichte auf $[0, 1]$ skaliert, so dass es sich als Graubild darstellt. Im ersten Ansatz wurde nur die diskrete Zeitreihe $\rho_t, t \in [0, T]$ übergeben, um zu untersuchen, inwieweit das Netz die Raumgeometrie aus der zeitlichen Dichteentwicklung ableiten kann. Im zweiten Ansatz soll die Raumgeometrie, meist gegeben als ein polygonales Gebiet, mittels einer *signed distance function* an das Netz übergeben werden, d.h. jedem Gitterpunkt (x, y) wird der kleinste signierte Abstand zum Ausgang zugeordnet, wobei das Vorzeichen negativ ist, falls sich die Koordinate außerhalb des Raums befindet.

Als Grundlage für das Modell wurde das sogenannte *unet-lstm*-Netz von Taylor et al. [1], das auf dem ConvLSTM-Modell basiert, das von Shi et al. [2] entwickelt wurde. Es handelt sich im Prinzip um ein LSTM, das aber dank Faltungsschichten als Input Tensoren, also räumlich-zeitliche Strukturen zulässt. Es wurde bereits erfolgreich zur Vorhersage von Satellitenbildern wie Niederschlagsvorhersage oder Temperaturentwicklungen verwendet. Es wird zum Training eine diskrete Zeitreihe verwendet. Anschließend kann das Modell auf Basis einer vorgegebenen Anzahl

vergangener Zeitschritte zukünftige Zeitschritte vorhersagen.

Ergebnisse

Es konnten im ersten Zeitraum (Oktober-Dezember 2023) große Trainingsdaten aus Simulationsdaten automatisiert erzeugt werden. Dabei wurde eine Methode entwickelt, unterschiedliche Raumgeometrien mit Innenwänden und Ausgängen zufällig zu erzeugen. Die aktuellen kleineren Versuche liefern qualitativ ähnliche Dichteentwicklungen wie die Simulation und bestätigen die prinzipielle Machbarkeit des

Ansatzes. Der Ansatz soll im Rahmen der Masterarbeit von Frau Wehrlein in SoSe 2024 nun systematisch untersucht und optimiert werden.

Literatur

- [1] Taylor, J. et al., Data-driven global weather predictions at high resolutions, The International Journal of High Performance Computing Applications, Volume 36, Issue 2, March 2022, Pages 130-140.
- [2] Shi, X. et al., Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, arxiv, 2015.

Untersuchung von Methoden zur multivariaten Anomaliedetektion in HVAC-Signalen

Nolane Delille, Pierre Desbruns, Joachim Druhet, Mahaut Galice, Tim Geißendörfer, Julian Post, Adrian Roy, Viktoria Stiem, Jens Ullrich, Aida Vehapi, Elida Vehapi, Doris Bohnet

Im Rahmen von zwei *Small Internationalization Projects* finanziert vom DAAD haben insgesamt 12 Studierende, je 6 von der HTWG und ESIREM (Dijon, Frankreich), im WiSe 2022/23 sowie WiSe 23/24 sich mit Methoden zur multivariaten Anomalieerkennung in HVAC-Signalen beschäftigt.

Hintergrund

Von Bosch Sicherheitssysteme GmbH wurden Sensordaten des Jahres 2021 von einer HVAC (heating-ventilation-air conditioning)-Automatisierung eines größeren Bürogebäudes zur Verfügung gestellt. Dieses Bürogebäude verfügt über eine sehr moderne, vollständig automatisierte Heizung- und Klimatisierungsanlage. Im Betrieb hat sich gezeigt, dass die Regelung nicht immer fehlerfrei ist: Zeitweise wurde beispielsweise trotz warmer Außentemperaturen geheizt, so dass die Zimmertemperatur viel zu hoch wurde. Übergeordnetes Ziel ist es, fehlerhafte Regelungen oder Sensordaten automatisch zu detektieren und dem diensthabenden Techniker im Gebäude zu melden, der dann den Fehler beheben kann. Die Anzahl der Zeitreihen ist so groß und das Zusammenspiel der Heiz-/Kühlkomponenten so komplex, dass es für den Techniker schwierig ist, Fehler selbst an den Daten zu erkennen. Deswegen wird ein stabiler Algorithmus gesucht, der Fehler in den Zeitreihen findet.

Im SoSe 2022 hatte sich bereits eine Bachelorarbeit bei Bosch mit diesem Thema beschäftigt und dabei einen Algorithmus aus Microsoft Azure verwendet, der essentiell auf zwei Graph Attention Networks basiert, aber keine zufriedenstellenden Ergebnisse lieferte. Aus diesem Grund sollten im Rahmen der studentischen Projekte auch klassische Methoden wie die Fourieranalyse für univariate Anomaliedetektion, klassisch-statistische Methoden sowie physikalische Modelle für die multivariate Zeitreihenanalyse untersucht werden. Hauptproblem bei beiden Projekte war die

Fehlerhaftigkeit der vorliegenden Daten, die keine gelabelten Anomalien besitzen, so dass detektierte Anomalien nicht verifiziert, sondern nur qualitativ plausibilisiert werden konnten.

Ergebnisse

Zunächst wurden die vorliegenden Zeitreihen analysiert und 14 ausgewählt, die im Zusammenhang mit der Regelung der Raumtemperatur eines Raums des Bürogebäudes stehen. Anschließend wurden verschiedene Methoden untersucht:

Fourier-Analyse: Für Zeitreihen mit empirisch indizierter Periodizität und ohne Interdependenzen zu anderen Zeitreihen wie die Außentemperatur oder Sonnenscheindauer wurde die klassische Fourieranalyse verwendet. Hierbei wurden hochfrequente Störungen herausgefiltert und das mittels inverser Fouriertransformation erhaltene Modell für die Vorhersage der Zeitreihe verwendet. Abweichungen größer als 3σ von gemessenen zu vorhergesagten Werten wurden als Anomalien markiert, s. Fig. 9.

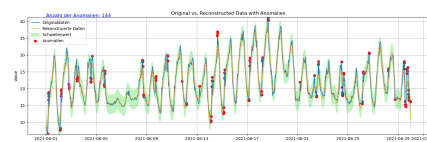


Figure 9: Anomalien detektiert durch Fourier-Analyse für die Außentemperatur.

Physikalische Modelle: Vorlauf-, Rücklauf-temperaturen, Volumenstrom und Heizleistung folgen einer bekannten, physikalischen Gesetzmäßigkeit, die mathematisch als Gleichung $P = C\dot{V}\Delta T$ formuliert werden kann. Es wurde angenommen, dass alle Fehler in den Messdaten durch Wahl des Parameters C abgebildet werden kann und normalverteilt sind. Entsprechend wurde C mit Hilfe der Messdaten gefittet und die Gleichung als Modell zur Vorhersage der jeweili-

gen Zeitreihen verwendet. Entsprechende Abweichungen wurden als Anomalien gekennzeichnet, s. Fig. 10. Diese Methode ist selbstverständlich nur anwendbar, wenn entsprechende physikalische Zusammenhänge zwischen Signalen bekannt sind.

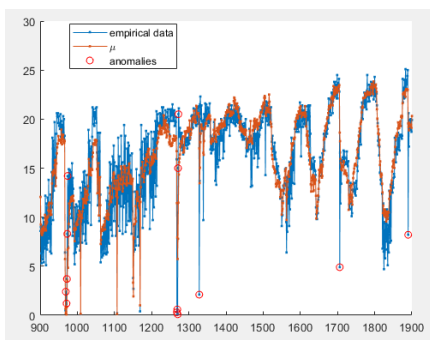


Figure 10: Anomalien detektiert durch physikalisches Modell für die Heizleistung.

Principal Component Analysis: Auf die Gesamtheit der Zeitreihen $X \in \mathbb{R}^{n \times T}$, wobei n Anzahl gemessene Komponenten, T Anzahl Zeitschritte, wurde PCA angewendet und damit mittels der gewählten Basis für einen geeigneten Untervektorraum, das hinreichend die Varianz aller Signale erklärt, ein Modell für Interdependenz der Signale gefunden. Es wurde angenommen, dass die Fehler normalverteilt sind. Abweichungen von mehr als 3σ wurden als Anomalien gekennzeichnet. Problematisch war dieser Ansatz zum einen, weil die zeitliche Abhängigkeit der Signale verloren gegangen ist. Zum anderen wurde kein Test durchgeführt, um die statistische Verteilung der Fehler zu validieren. **Mahalanobis-Distanz:** Die Mahalanobis-Distanz erlaubt uns, den statistischen Zusammenhang zwischen den Zeitreihen zu messen. Wie die PCA beruht die Mahalanobis-Distanz auf der Korrela-

tion zwischen den verschiedenen Daten und ist wie folgt definiert: Sei $X = (X_1, \dots, X_n)$, n Zufallsvariablen (=gemessene Komponenten), und $\bar{X} = (\bar{X}_1, \dots, \bar{X}_n)$ die empirischen Mittelwerte über alle Stichproben, C die Kovarianzmatrix:

$$D_M(X) = \sqrt{(X - \bar{X}) C^{-1} (X - \bar{X})^T} \quad (1)$$

Wenn die n Zufallsvariablen gaußverteilt sind, dann ist $D_M^2(X)$ $\chi^2(n)$ -verteilt. Entsprechend kann ein Wert mit einem Risiko von $\alpha \in]0, 1[$ als Ausreißer betrachtet werden, wenn er außerhalb des $1 - \alpha$ -ten Quantil der $\chi^2(n)$ -Verteilung liegt, s. Fig. 11. Auch bei dieser Methode wird leider die zeitliche Dimension der Zeitreihen nicht berücksichtigt. Da die zu unterschiedlichen Zeitpunkten gemessenen Werte als Stichproben interpretiert wurden, ist nicht unbedingt von einer Gauß-Verteilung auszugehen.

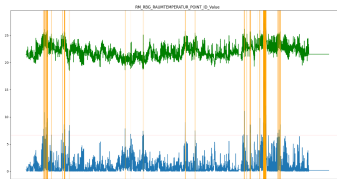


Figure 11: Anomalien detektiert mittels Mahalanobis-Distanz (5%-Risiko).

Fazit:

Auch wenn mit den relativ einfachen Verfahren keine durchweg zufriedenstellenden Ergebnisse erzielt werden konnten, hat die detaillierte Arbeit mit den Zeitreihen mit Verfahren, die die Studierenden verstehen konnten, zu einem guten Verständnis der Schwierigkeiten und Prinzipien der Signalverarbeitung beigetragen.

BERNSTEIN POLYNOMIAL NORMALIZING FLOWS

A valuable way to describe a probability distribution is by considering it as a transformation from a basic distribution to a complex distribution. See Figure 12 for the basic principle.

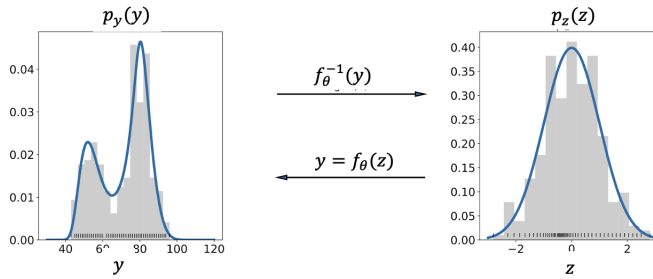


Figure 12: Transformation function f_θ between a simple distribution on the right and a complex distribution on the left. The transformation function is parameterized by θ , which we control by neural network. Using a neural network it is easy to condition θ on external variables like images.

In a series of our research papers, we model the transformation function using neural networks, a concept recognized as normalizing flows within the machine/deep learning community. A distinctive feature of our approach across these papers is the utilization of Bernstein Polynomials. We introduced the method in **Deep Transformation Models for Probabilistic Forecasting** and applied it in various applications, such as probabilistic forecasting ("**Enhancing Low-Voltage Load Forecasting Through Probabilistic Modeling with Bernstein-Polynomial Normalizing Flows**") or in a multivariate extension as a variational distribution in variational Bayes ("**Flexible Posteriors in Variational Bayes with Bernstein Polynomial Normalizing Flows**").

Deep Transformation Models for Probabilistic Forecasting

Beate Sick, Torsten Hothorn, Oliver Dürr

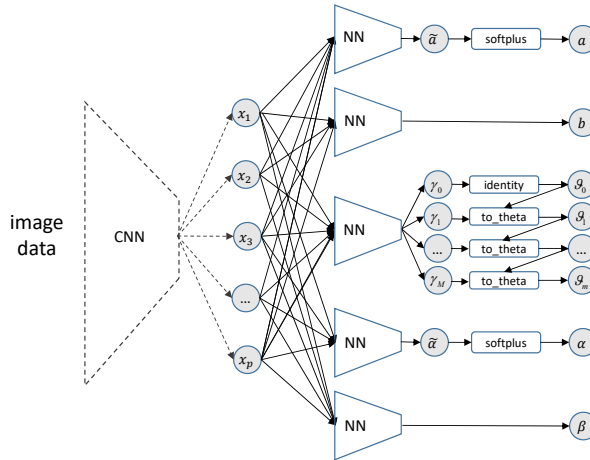


Figure 13: The network architecture of our model which is trained end-to-end starting with the input x (such as images) yielding the conditional parameters of the transformation functions $\theta|x$ from figure 12. Important is the fact that the network controls the parameters $\theta = \{a, b, \vartheta_0, \dots, \vartheta_M, \alpha, \beta\}$ of the transformation function f .

Deep learning has revolutionized predictive modeling with its unparalleled accuracy in complex data scenarios. However, its application in regression tasks has traditionally been limited to point estimates, ignoring the inherent probabilistic nature of many real-world phenomena. The Deep Transformation Model (DTM) addresses this limitation by providing a framework for probabilistic regression, estimating entire conditional probability distributions (CPD) for given inputs. This advancement allows for a more comprehensive understanding and quantification of predictive uncertainty, which is crucial in decision-making processes, especially in domains like healthcare.

Methodology

At the heart of the DTM approach is the combination of ideas from statistical transformation models, specifically Most Likely Transformation with Bernstein polynomials controlled by a neural network see Figure 12. The method leverages a parameterized transformation function trained via maximum likelihood estimation to model complex outcome distributions. This function is adept at transforming a simple base distribution into a more sophisticated target distribution, capturing the full range of potential outcomes and their probabilities. By incorporating neural networks, the model parameters are conditioned on different input data, allowing the DTM to directly handle high-dimensional inputs like images or other modalities. This project showcases the potent combination of deep learning flexibility and statistical modeling rigor. An example application is

shown in figure 14 where we use a convolutional neural network (CNN) to predict the probability distribution of a person's age given an image x as input.

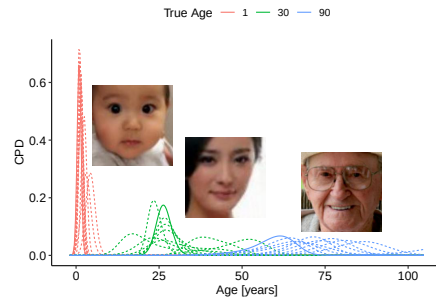


Figure 14: Conditional probability distributions (CPD) for different ages. The images are shown as 10 randomly picked test set results for different true ages (1, 30, and 90 years), and the images correspond to the solid lines. The model results meet our expectations, yielding narrow CPDs for infants, while for older people, the spread of the CPD broadens with increasing age.

Applications and Results

The DTM's capability is demonstrated through a series of experiments, ranging from one-dimensional toy data to real-world applications like age estimation from facial images. These examples underline the model's ability to capture complex, non-Gaussian distributions and provide meaningful uncertainty estimates across various contexts. Notably, the model outperforms existing

probabilistic regression methods in several benchmark datasets, highlighting its potential as a robust tool for predictive modeling.

Conclusion and further references

The Deep Transformation Model represents a significant leap forward in probabilistic forecasting, marrying the depth of neural networks with the principled approach of transformation models. It opens up new avenues for research and application

in fields that benefit from a nuanced understanding of uncertainty in predictions. More information about this innovative approach can be found in the paper:

Sick, B., Hothorn, T., & Dürr, O. (2021). Deep transformation models: Tackling complex regression problems with neural network based transformation models. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 2476-2481). IEEE. Available at <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9413177>

Enhancing Low-Voltage Load Forecasting Through Probabilistic Modeling with Bernstein-Polynomial Normalizing Flows

Marcel Arpogaus, Marcus Voss, Beate Sick, Mark Nigge-Uricher, Oliver Dürr

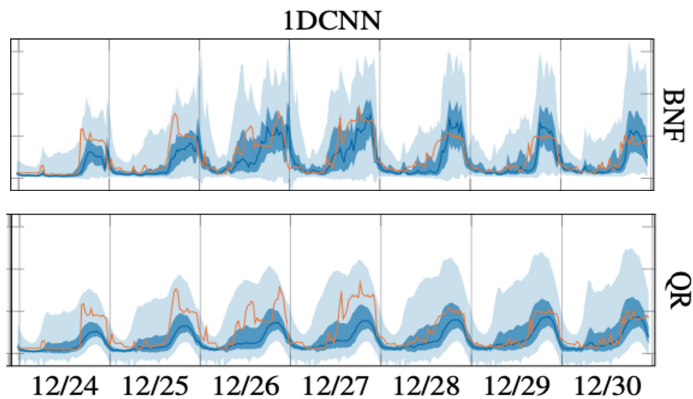


Figure 15: Comparison of the day ahead power consumption. Shown is our method, the Bernstein Polynomial Flow (BNF) against a quantile regression forecast (QR). In orange is the predicted data, and in blue are the predicted quantiles. In the QR code often too narrow predictions are given.

The evolving dynamics of renewable energy grids demand advanced probabilistic forecasting models that can effectively manage the inherent uncertainties in low-voltage demand predictions. Conventional forecasting methodologies primarily offer point estimates, falling short of addressing the complexities introduced by high fluctuations in the low voltage domain. In response to these challenges, this paper employs a novel approach, rooted in the concept that a valuable way to describe a probability distribution is by considering it as a transformation from a basic distribution to a complex distribution. Utilizing this founda-

tion, we apply the methods introduced in "Deep Transformation Models for Probabilistic Forecasting," which model the transformation function using neural networks, to the energy grid load forecasting domain. Applying Bernstein Polynomial Normalizing Flows within a neural network framework provides a flexible density forecasting solution, significantly enhancing predictive precision for short-term (24-hour ahead) load forecasting. This approach adeptly quantifies the uncertainties, as demonstrated in an empirical analysis of 3639 smart meter customers, showcasing its potential in various applications.

Method

The core of our methodological innovation lies in applying Bernstein polynomial normalizing flows for conditional density forecasting, emphasizing the probabilistic nature of the forecast. This approach offers a detailed landscape of possible outcomes, each quantified with respective probabilities, unlike traditional models that present a single expected value. Our probabilistic model is validated using the publicly available dataset from the CER Smart Metering Project, ensuring a robust foundation for the empirical study.

Results

Our findings illuminate the substantial benefits of adopting a probabilistic forecasting approach, particularly in contexts fraught with high variability and uncertainty. The Bernstein polynomial normalizing flows method outperforms standard Gaussian and Gaussian mixture models by delivering more accurate and comprehensive density forecasts. These probabilistic predictions offer a

clearer understanding of potential load variations and could significantly improve decision-making processes for energy management and grid operations in the renewable energy era.

Conclusion

This study underscores the importance of probabilistic forecasting in navigating the uncertainties of low-voltage load predictions. The proposed Bernstein polynomial normalizing flows model emerges as a superior tool, offering detailed and reliable forecasts essential for the efficient and resilient operation of future energy grids. Our work marks a step towards embracing the full complexity of demand forecasting, facilitating a smoother transition to renewable energy sources.

For more details, please refer to the paper:

Arpogaus, M., Voss, M., Sick, B., Nigge-Uricher, M., & Dürr, O. (2023). Short-term density forecasting of low-voltage load using bernstein-polynomial normalizing flows. *IEEE Transactions on Smart Grid*. Available at: <https://arxiv.org/abs/2204.13939>

Flexible Posteriors in Variational Bayes with Bernstein Polynomial Normalizing Flows

Oliver Dürr, Stefan Hörtling, Daniel Dold, Ivonne Kovylov, Beate Sick

In Bayesian inference, the development of precise and adaptable probabilistic models has taken a significant step forward with the introduction of Bernstein Polynomial Normalizing Flows (BP-NFs). Originating from the master's thesis projects of Stephan Hörtling and Ivonne Kovylov, this method contributes to advancing the field of Variational Inference (VI). Our research demonstrates the potential of BP-NFs to accurately and flexibly approximate complex posterior distributions.

Method

The Bernstein Flow Variational Inference (BF-VI) method exploits the Bernstein polynomials' capabilities for a highly adaptive approximation of posterior distributions in Bayesian frameworks. BF-VI adeptly handles the intricate dependencies present in multidimensional datasets by employing invertible transformations from a simplistic latent distribution to a multifaceted posterior. This technique

is remarkably versatile and suitable for a wide array of applications ranging from straightforward one-dimensional problems to the complex, semi-structured data often found in deep learning endeavors.

Results

Through rigorous benchmarks, including the challenging 8-schools example and a pioneering application to melanoma diagnosis using semi-structured data, BF-VI's superiority is highlighted. It outperforms both traditional Gaussian VI and other normalizing flow-based approaches, demonstrating unparalleled proficiency in capturing the subtleties of posterior distributions. The melanoma diagnosis case, in particular, exemplifies BF-VI's strength in integrating deep learning with probabilistic modeling, significantly improving predictive outcomes.

Conclusion and further links

BF-VI represents an advancement in the field of probabilistic modeling, offering a powerful and flexible tool for Bayesian analysis across diverse scenarios. The groundwork established through the master's thesis projects of Stephan Hörling and Ivonne Kovylov has led to a novel methodol-

ogy that not only propels forward the domain of Bayesian deep learning but also paves the way for future explorations in probabilistic forecasting.

For further details on the methodology and its applications, readers are encouraged to consult the full paper on arXiv (arXiv:2202.05650), accepted for publication in AStA Advances in Statistical Analysis.

While deep learning is exceptional at identifying complex patterns, statistical models offer superior interpretability. Our research aims to merge these two worlds, creating powerful yet understandable models. In **Ordinal Neural Network Transformation Models: Bridging Deep Learning and Interpretable Regression for Ordinal Outcomes**, we develop an interpretable method combining deep learning and statistical modeling for ordinal data. In **Deep transformation models for functional outcome prediction after acute ischemic stroke** we apply the method to predict the outcome after stroke using MRI (with deep learning) and patient information (statistical modeling part) to predict the functional outcome after stroke.

Ordinal Neural Network Transformation Models: Bridging Deep Learning and Interpretable Regression for Ordinal Outcomes

Lucas Kook, Lisa Herzog, Torsten Hothorn, Oliver Dürr, Beate Sick

Abstract

Ordinal outcomes are prevalent in predictive tasks across various domains, often accompanied by complex data forms such as images or tabular predictors. While Deep Learning (DL) excels in tasks like image classification, it typically overlooks the natural order in outcomes and lacks interpretability. Classical ordinal regression models, on the other hand, respect outcome order and offer interpretable predictions but struggle with complex non-tabular data. This paper introduces Ordinal Neural Network Transformation Models (ONTRAMs), a novel methodology that unifies DL's flexibility with ordinal regression's interpretability. This approach not only enhances predictive performance but also preserves the model's interpretability, offering a powerful tool for applications requiring detailed uncertainty estimates and insight into the predictive process.

Introduction and Contributions

ONTRAMs represent a significant advancement in modeling ordinal outcomes, especially when dealing with mixed data types. Traditional methods either sacrifice the natural order of outcomes for the sake of model complexity (DL) or cannot adequately handle non-tabular data (ordinal regression). By transforming the ordinal outcome to a continuous variable using a monotonically increasing transformation function—facilitated. The key contributions of this work include:

A flexible yet interpretable model framework that incorporates both image and tabular data. The use of a maximum likelihood-based approach

for parameter estimation, enhancing training efficiency and model interpretability. Demonstrated equivalence in prediction performance to traditional DL models, with added benefits of interpretability and efficient training.

Methodology

At the core of ONTRAMs is the transformation function, which is additively decomposed into components for handling different types of input data. This decomposition allows for the modeling of complex data interactions while retaining interpretability through the probabilistic modeling of ordinal outcomes. Bernstein Polynomial Normalizing Flows play a crucial role in this process, enabling the transformation of simple distributions into complex, data-driven outcome distributions. Through modular architecture, ONTRAMs can accommodate various data types and complexities, making them highly versatile and applicable across different domains.

Applications and Results

To validate the effectiveness and interpretability of ONTRAMs, we conduct extensive experiments across multiple datasets, including UTKFace for age classification from images and tabular data. The results showcase ONTRAMs' ability to outperform or match traditional DL models in prediction accuracy while providing significant advantages in terms of interpretability and insight into the effects of predictors on ordinal outcomes.

Conclusion

ONTRAMs mark a pivotal development in the intersection of DL and statistical modeling for ordinal outcomes. By leveraging the strengths of both domains, ONTRAMs offer a powerful solution for tasks requiring nuanced understanding and pre-

diction of ordinal data. For more information see:

Kook, L., Herzog, L., Hothorn, T., Dürr, O., & Sick, B. (2022). Deep and interpretable regression models for ordinal outcomes. *Pattern Recognition*, 122, 108263. Available also at <https://arxiv.org/abs/2010.08376>

Deep transformation models for functional outcome prediction after acute ischemic stroke

Lisa Herzog, Martin Hänsel, Beate Sick, Lucas Kook, Andrea Götschi, Katrin Petermann, Janne Hamann, Oliver Dürr, Susanne Wegener

This research focuses on applying Deep Transformation Models (DTMs) for predicting functional outcomes after acute ischemic stroke, leveraging a dataset of 407 patients. It highlights DTMs' capability to handle semi-structured data, including tabular and image data, for interpretable distributional regression predictions.

Introduction

The study addresses the critical need for rapid and informed decision-making in stroke treatment, which remains a leading cause of disability and death worldwide. By employing DTMs for distributional regression, the research aims to balance high prediction performance with interpretability, a notable advancement over traditional methods that might provide only point predictions or fail to accommodate complex data forms.

Methodology

The methodology centers on applying DTMs to semi-structured data, combining tabular patient information with brain imaging results (CT or MRI), to predict ordinal functional outcomes measured by the modified Rankin Scale (mRS). This approach emphasizes the adaptability of DTMs to varied research questions, facilitated by specifying neural networks for different data inputs, thereby combining deep learning's predictive power with the interpretability of traditional statistical models.

Results and Discussion

The analysis compares various DTMs, including baseline-adjusted models and those incorporating

semi-structured data, focusing on ordinal outcome prediction. Key findings include the significance of both tabular clinical data and brain imaging in outcome prediction, with models based solely on tabular data outperforming those based only on imaging data. Additionally, the combination of data modalities did not significantly enhance prediction accuracy, underlining the pivotal role of tabular data.

Conclusion and further references

The study underlines DTMs as a powerful approach for analyzing semi-structured data, with significant implications for supporting clinical decision-making through a blend of interpretability and advanced prediction performance. It advocates for a balanced approach in transformation model-building, emphasizing the combination of interpretability and flexibility to facilitate nuanced clinical assessments.

This research propels the use of deep learning in medical prognosis, integrating diverse data forms into a coherent, interpretable framework for stroke outcome prediction, and highlights the importance of statistical principles in developing effective predictive models. More details can be found in the paper:

Herzog, Lisa, Lucas Kook, Andrea Götschi, Katrin Petermann, Martin Hänsel, Janne Hamann, Oliver Dürr, Susanne Wegener, and Beate Sick. "Deep transformation models for functional outcome prediction after acute ischemic stroke." *Biometrical Journal* 65, no. 6 (2023): 2100379. Also available at <https://arxiv.org/abs/2206.13302>

Simplifying MEMS Accelerometer Calibration Using Bayesian Methods

Oliver Dürr, Po-Yu Fan, Zong-Xian Yin

In recent years, micro-electro-mechanical systems (MEMS) accelerometers have become integral to various applications, necessitating precise calibration to correct errors and improve accuracy. Traditional calibration techniques, while effective, often do not account for the inherent uncertainties in parameter estimation, which can lead to sub-optimal calibration and performance. This study introduces a Bayesian approach to MEMS accelerometer calibration that naturally incorporates uncertainty and prior knowledge into the calibration process, offering a more comprehensive and accurate calibration method.

Introduction

MEMS accelerometers, crucial for tasks ranging from activity monitoring to gesture recognition, require calibration to mitigate errors from various sources, such as sensor misalignments and signal drifts. Conventional calibration methods focus on estimating error-correcting function parameters without adequately addressing the uncertainties in these estimations. By applying Bayesian methods, we propose a calibration process that improves precision and quantifies uncertainty, enhancing the overall reliability of the sensor.

Methodology

The Bayesian calibration process treats model parameters as random variables to simplify incorporating prior knowledge and uncertainty into the calibration model. This study leverages recent advances in probabilistic programming frameworks, such as Stan and Pyro, to implement a straightforward Bayesian calibration method for MEMS accelerometer data. The approach begins by defining priors for the parameters of interest, followed by updating these priors with observed data to obtain posterior distributions that reflect updated beliefs about the parameters, including uncertainty estimates.

Our approach is based on the principle that when at rest, the magnitude of acceleration ($|g|$) remains constant across all orientations. A handful

of recordings at rest are sufficient for calibration, simplifying the process. This insight significantly enhances the method's practicality in real-world applications. In Figure 16, the glove with 6 accelerometers, for which we tested the method, is shown.



Figure 16: The glove used to demonstrate the calibration of the accelerometer data: Five IMUs are placed on the fingers, approximately at the position of the fingernails, and one IMU is located on the palm. In addition, a single IMU unit is displayed near the thumb for reference.

Results and Applications

Simulation studies demonstrate the effectiveness of the Bayesian approach in capturing both aleatoric and epistemic uncertainties, providing more informative calibration results than traditional methods. The application of this method to real-world accelerometer data from wearable devices showcases its practical utility and the advantages of incorporating uncertainty into the calibration process. The Bayesian method not only yields accurate calibration parameters but also provides credibility intervals, offering insight into the reliability of the calibration.

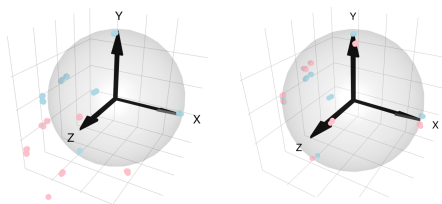


Figure 17: Acceleration data for the IMUs 5 (pink) and 2 (turquoise). The 44 raw data points are shown on the left side. After calibration (right side), the data lies closely on the sphere with radius $|g|$ (shown in gray).

Conclusion

The proposed Bayesian calibration method for MEMS accelerometers represents a significant ad-

vancement in sensor calibration techniques. By fully embracing the probabilistic nature of parameter estimation and the uncertainties involved, this approach allows for more accurate and reliable calibration. The simplicity of implementation, coupled with the quantification of uncertainty, makes Bayesian calibration a compelling alternative to traditional calibration methods, promising to enhance the performance of MEMS accelerometers across a wide range of applications.

For more information, please refer to: Dürr, O., Fan, P. Y., & Yin, Z. X. (2023). Bayesian Calibration of MEMS Accelerometers. *IEEE Sensors Journal*. Also available at <https://arxiv.org/abs/2306.06144>

Novel AI-Based Algorithm for the Automated Computation of Coronal Parameters in Adolescent Idiopathic Scoliosis Patients: A Validation Study on 100 Preoperative Full Spine X-Rays

Clara Berlin, Sonja Adomeit, Priyanka Grover, Marcel Dreischarf, Henry Halm, Oliver Dürr, Peter Obid

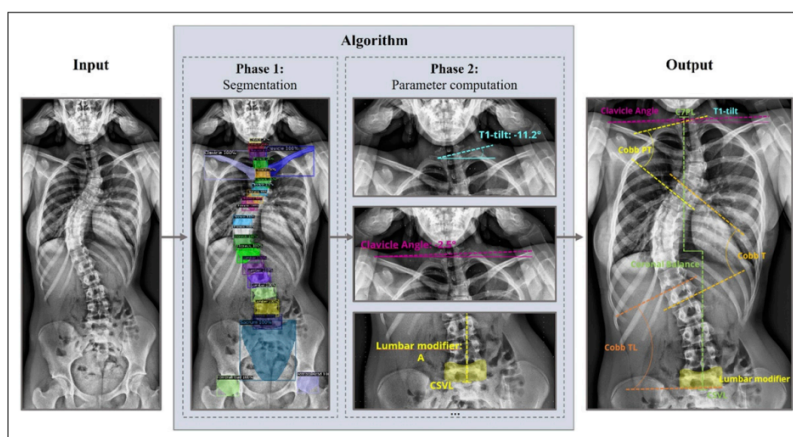


Figure 18: Schematic representation of the pipeline of the presented algorithm showing an AP full spine X-ray as input (left); segmentation of anatomical structures of interest (middle left); computation of parameters based on the segmentation results (middle right); visualization of the coronal parameters (right).

This validation study based on the Master Thesis of Sonja Adomeit introduces a novel AI-based algorithm for the automated computation of coronal parameters in Adolescent Idiopathic Scoliosis (AIS) patients, validated against measurements from experienced physicians on 100 preoperative full spine X-rays. The AI algorithm demonstrates excellent reliability and agreement with human raters, suggesting its potential to significantly enhance efficiency in clinical practice and research.

Introduction

Adolescent idiopathic scoliosis, a prevalent spinal deformity, necessitates regular radiological examinations for management. Existing manual methods for radiological image analysis are time-consuming and subject to observer experience. This study proposes an AI-based algorithm to overcome these limitations, aiming for an efficient

and objective assessment of coronal radiographic parameters.

Methods

The study developed an AI algorithm capable of detecting anatomical structures of interest and automatically calculating essential radiographic parameters in AP spine X-rays. The algorithm was validated using measurements from two experienced physicians on 100 preoperative AP full spine X-rays of AIS patients, serving as the ground truth for evaluating inter-rater and intra-rater reliability.

AI-Algorithm

The AI algorithm consists of two phases: segmentation, where anatomical structures are identified and classified, and parameter determination, where radiographic parameters are computed

based on the segmentation results. The algorithm processes images to automatically compute coronal parameters, including T1-tilt, clavicle angle, coronal balance, lumbar modifier, and Cobb angles, with high accuracy and reliability.

Results

The AI algorithm achieved excellent ICC values (.78-.98) across all evaluated parameters, showing remarkable reliability and agreement with human raters. The ability to determine all parameters in 100 percent of the images underscores the algorithm's robustness and potential to streamline the assessment of scoliosis in clinical settings.

Discussion

The study emphasizes the AI algorithm's capability to provide rapid, reliable, and objective assessments of coronal radiographic parameters, which is essential for efficiently managing AIS patients. This technological advancement could revolutionize the evaluation of spinal deformities, offering a valuable tool for clinical practice and large-scale research studies.

For more details, please refer to the paper: Berlin, C., Adomeit, S., Grover, P., Dreischarf, M., Halm, H., Dürr, O., & Obid, P. (2023). Novel AI-based algorithm for the automated computation of coronal parameters in adolescent idiopathic scoliosis patients: A validation study on 100 pre-operative full spine X-rays. *Global Spine Journal*.

Using Gaussian Process for Robust Temperature Forecast

Leonardo Labanca, Daniel Dold, Oliver Dürr

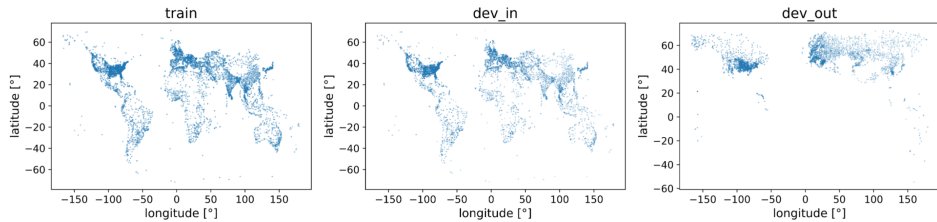


Figure 19: Locations of the train (left), dev_in (middle), and dev_out (right) data. The x-axis shows the longitude coordinate in degree and the y-axis represents the latitude coordinate in degree.

In today's highly dynamic world the requirements for algorithms are continuously changing. Especially in the area of machine learning and probabilistic modeling, new fields of application emerge every day. Gaussian processes are playing an increasingly important role and gain extensive attention in the state-of-the-art literature [1]. Similarly, artificial neural networks have been an essential part of probabilistic deep learning for quite some time and have proved their reliability and applicability in a wide variety of use cases. In a data-driven world, where more data usually implies more information and better performance, the handling of large-scale datasets is also receiving increasing interest. It became obvious a long time ago that the handling of Big Data in the economic and industrial sector is an inevitable must-have. In this field of application in particular, the treatment of data from the real world is essential. Such data is often affected by distributional shifts, resulting in mismatches between training and test data as well as in their relation to each other. In most cases we differentiate between three different types of dataset shifts: covariate shift, prior probability shift, and concept drift [2]. This work focuses mainly on the effects of covariate shifts. Such distributional shifts and their effects, in turn, can have a negative impact on the performance of a machine learning model and therefore require special attention. This is why the 35th Neural Information Processing Systems Conference 2021 also dealt with this topic and published the Shifts

Challenge in this context. The challenge addresses the robustness and uncertainty estimates of predictive models regarding real-world datasets with distributional shifts [3].

Method

This paper provides an approach for robust and interpretative modeling using an extensive real-world industrial application example - a temperature forecast based on temporal, spatial, and meteorological data (See Figure 19 Gaussian processes are particularly suited to deal with temporal and spatial correlation structures, hence these features will be considered using a Stochastic Variational Gaussian Process, which is scaled sufficiently by specific modifications in order to achieve an acceptable runtime even for very large datasets [4, 5]. Probabilistic neural networks have often shown their capability in applications with multidimensional data as well as non-linear optimization problems [6]. In the temperature forecast model, both methods a Gaussian process and a probabilistic neural network are combined and adapted to a large-scale dataset. The data used to develop and evaluate the forecast model is provided as part of the Shifts Challenge and can be downloaded from their GitHub directory [7].

Results

In several experiments the performance of the model is analyzed in terms of predictive accuracy

as well as robustness. Likewise, the significance of the uncertainty estimates will be evaluated. In this context the influence of the Gaussian process and the neural network will be considered in two combination approaches. Subsequently, the stand-alone Gaussian process model will be analyzed and individual components will be examined in more detail. The final results provide an approach for robust modeling on data with distributional shifts including predictions and interpretive uncertainty estimates. This work confirms the benefits of Gaussian processes when modeling temporal and spatial structures, but also reveals their problems concerning extrapolation. The results also underline the importance of appropriate priors for time-series modelling and the main disadvantages of stationary kernel functions.

Literatur

- [1] Rasmussen, C. & Williams, C. Gaussian Processes for Machine Learning. (MIT Press,2006)
- [2] Quiñero-Candela, J., Sugiyama, M., Schwaighofer, A. & Lawrence, N. Dataset shift in machine learning. (MIT Press,2010)
- [3] Foundation, Neural Information Processing Systems Competition Track 2021.
- [4] Tran, D., Ranganath, R. & Blei, D. The Variational Gaussian Process.
- [5] Hensman, J., G. de G. Matthews, A. & Ghahramani, Z. Scalable Variational Gaussian Process Classification. Proceedings Of The 18th International Conference On Artificial Intelligence And Statistics (AISTATS). (2014)
- [6] Nielsen, M. Neural Networks and Deep Learning. (Determination Press,2015)
- [7] Malinin, A., Band, N., Ganshin, Alexander, Chesnokov, G., Gal, Y., Gales, M., Noskov, A., Ploskonosov, A., Prokhorenkova, L., Provilkov, I., Raina, V., Raina, V., Roginskiy, Denis, Shmatova, M., Tigas, P. & Yangel, B. Shifts: A Dataset of Real Distributional Shift Across Multiple Large-Scale Tasks.

Accelerating BNN's in the context of active learning

Maksim Sandybekov, Daniel Dold, Oliver Dürr

In this work, we apply a sampling free-approximation approach called Moment Propagation [1] in an active learning environment. By running different experiments and comparing the outcome to a widespread approximation technique, we show that adopting a sampling-free approach greatly reduces the runtime of the active learning loop while keeping a comparable performance.

Introduction

The practical application of deep learning requires a massive amount of data. The needed data is not always existent and needs to be acquired first, a task that is time-consuming and expensive. An approach called active learning tackles this issue. In active learning, we leverage an unlabeled dataset to identify the most valuable data points for a given neural network architecture. For this purpose, we create two different pools. One pool for unlabeled datapoints (D_U) and another one that keeps the labeled datapoints (D_L). To initialize D_L we randomly select a specific number of data points from D_U and label them. The labels are requested from a so-called oracle which can either be another neural net or a human. In an experimental setting, a dataset with known labels is used to pseudo-label the data points. Iteratively we perform the following steps. First, we train a given Bayesian neural network (BNN) on the data in D_L . Afterward, we rank the data points in D_U , select a specific number by ranking, label, and move them into D_L . For the ranking mechanism, an acquisition function is used. This function leverages the previously trained model to calculate a specific criterion to rank the datapoints in D_U by.

However, Bayesian neural networks are slower than their regular neural network counterparts during test time. The reason is that currently, widespread approximation techniques make use of a sampling approach. Meaning, that during test time a value is sampled n -times from the BNN for a given data point. A very common sampling-based approach is MC Dropout [2], which uses

dropout not only during training but also during test time. For each forward pass, another network realization is used. This approach therefore can also be interpreted as using an ensemble of n neural networks.

The heavy use of Bayesian neural networks limits the application of active learning. By applying a faster approximation technique, we can speed up the active learning loop and decrease the iteration time. Over the last couple of years, sampling-free approximation approaches emerged. One such approach is called Moment Propagation [1]. Like MC Dropout, this approach uses dropout to induce uncertainty. Contrary to other approximation approaches, the variance and expectation of the approximated distribution are passed directly from layer to layer.

Results

For our experiments, we use the MNIST dataset split into training-, test- and validation sets. The training set encompasses 59.900 samples and represents our initial unlabeled pool D_U .

We perform the same experiments on each approximation technique (Moment Propagation and MC Dropout) using a set of single-criterion acquisition functions outlined in [3]. The labeled pool D_L is initialized with 20 randomly selected data points. Each active learning round we select 10 new datapoints from D_U for D_L . To validate our results we use 10 different seed configurations.

To get a more detailed view of the difference between Moment Propagation and MC Dropout we calculate the labeling efficiency using the test accuracy of both models. This metric represents a relative measure of the test accuracy reached by a specific acquisition function and the baseline for a given approximation. The baseline in our case is the random selection of data points from D_U .

Considering Figure 20 we can see that both models follow a similar performance profile for the Max. Entropy acquisition function. This specific acquisition function ranks data points by calculating the entropy and selecting data points

that reach the highest entropy values. In the early stages, the acquisition function underperforms because there is a lack of information in the labeled pool. In later stages, a higher performance is reached compared to random selection. Every other acquisition function we use shows the same behavior.

By comparing the runtimes of both approaches in Figure 21, we can see that Moment Propagation greatly reduces the total runtime of the active learning loop. Which in turn increases the speed with which experiments can be carried out.



Figure 20: Labeling efficiency of Max. Entropy acquisition function between MC Dropout and Moment Propagation on MNIST. On inference, the MC dropout uses 25 samples for the posterior approximation

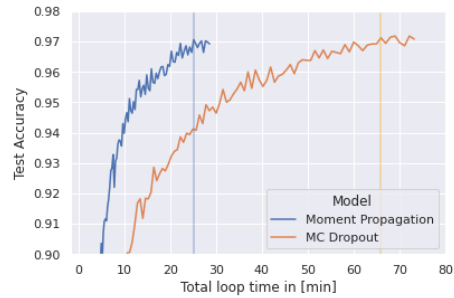


Figure 21: Active learning runtime in minutes [min] to reach a specific test accuracy. With total query time and test accuracy meaned over the acquisition functions. Reaching 97% test accuracy is marked by a vertical line.

Literatur

- [1] Brach, K., Sick, B., and Dürr, O. (2020). Single Shot MC Dropout Approximation.
- [2] Gal, Y. and Ghahramani, Z. (2016). Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. 33rd International Conference on Machine Learning, ICML 2016, 3:1651–1660.
- [3] Gal, Y., Islam, R., and Ghahramani, Z. (2017). Deep Bayesian active learning with image data. 34th International Conference on Machine Learning, ICML 2017, 3:1923–1932.

Geometrische Kalibrierung zwischen IMU/GPS und LiDAR

Felix Mayer, Dennis Griebner, Matthias Franz

Mit dem Aufkommen kostengünstiger und kompakter LiDAR- und IMU-Sensoren finden diese in SLAM- sowie VIO-Verfahren vermehrt Anwendung. Um die Daten der Sensoren in einem gemeinsamen Raum darstellen zu können, ist eine Kalibrierung der räumlichen Beziehung der Sensoreinheiten erforderlich. Das Problem der Bestimmung der räumlichen Beziehung zweier Koordinatensysteme ist in der Robotik unter dem Namen Hand-Auge-Kalibrierung bekannt. Die relative Transformation wird anhand der simultanen Bewegungen der beiden Einheiten durch den Raum bestimmt.

In dieser Arbeit werden zwei Verfahren aus der Literatur zur Lösung des Problems implementiert und die Anwendbarkeit dieser für eine Kalibrierung eines LiDAR-Sensors zu einer IMU/GPS-Einheit getestet. Das erste Verfahren wurde von Roger Y. Tsai und Reimar K. Lenz entwickelt und 1989 im IEEE Transactions on Automation Science and Engineering publiziert [1]. Das zweite Verfahren stammt von den Autoren Yacong Wang und Long Zhao [2] und ist 2017 in "Lecture Notes in Electrical Engineering Vol. 459" erschienen.

Für eine realistische Evaluierung der Anwendbarkeit wird die vorgesehene Sensorik in einer Gazebo/ROS-Simulation modelliert. Die Sensorik wird in der Simulation mit einem mobilen Roboter kombiniert, wie in 22 dargestellt. Die Sensorik wird mit verschiedenen starken Störeinflüssen konfiguriert, wobei die Konfigurationen den Einsatz unterschiedlich hochwertiger Sensoreinheiten repräsentieren. Für die Bestimmung der Trajektorie des LiDAR-Sensors wird das, von den Autoren Tixiao Shan und Brendan Englot im Jahr 2018 in "LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain" vorgestellte Verfahren eingesetzt [3]. Die Trajektorie der IMU/GPS-Einheit wird durch einen Kalman-Filter der Autoren Thomas Moore und Daniel Stouch für ROS bestimmt [4].

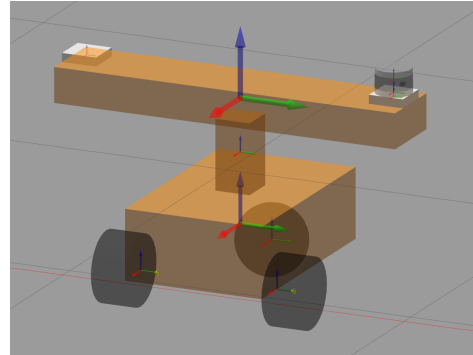


Figure 22: Mobiler Roboter modelliert in Gazebo mit Sensorik.

23 und 24 zeigen die durch LeGO-LOAM und Kalman-Filter geschätzten Trajektorien. Das Sensorrauschen beträgt in dieser Konfiguration 3 Zentimeter Standardabweichung für die Entfernungswerte des LiDAR-Sensors. Die Standardabweichungen des Rauschens für die IMU betragen 1.0 Grad für Roll- und Nickwinkel sowie 1.5 Grad für den Gierwinkel. Für das GPS-System wurden 10 und 50 Zentimeter Standardabweichung für den horizontalen sowie vertikalen Lokalisierungsfehler gewählt. Neben dieser Konfiguration werden drei weitere Konfigurationen mit stärkerem und schwächerem Sensorrauschen getestet.

Das Verfahren der Autoren Wang & Zhao ist in dieser Konfiguration Stande, die relative Orientierung der beiden Sensoreinheiten mit einem Fehler von 0.539 Grad zu schätzen. Der Fehler der relativen Translation beträgt 14.1 Zentimeter und damit etwa 14% des Betrags des tatsächlichen Translationsvektors.

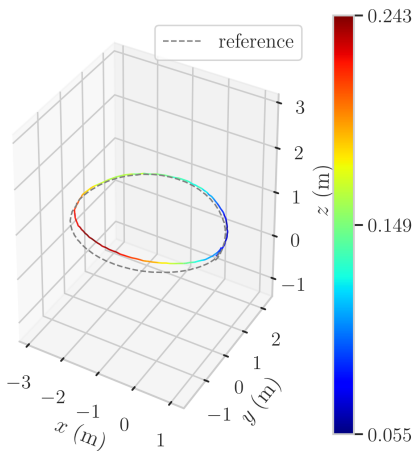


Figure 23: APE bezüglich Translation in Meter mit LeGO-LOAM.

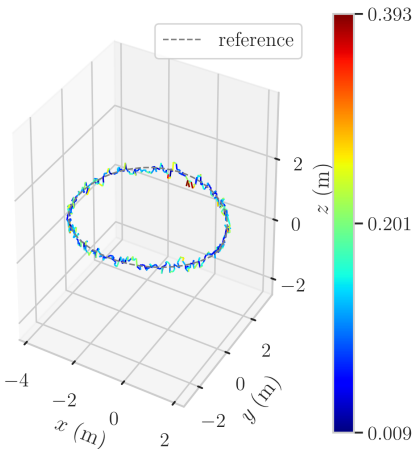


Figure 24: APE bezüglich Translation in Meter mit Kalman-Filter.

Die Ergebnisse zeigen, dass die Verfahren der Autoren Wang & Zhao und Tsai & Lenz, welche das Schätzen der relativen Transformation zwecks Linearität in Rotation und Translation aufteilen, zwar im Stande sind die relative Orientierung zu schätzen, sich jedoch nicht für die Bestimmung der relativen Translation eignen. Diese Tatsache wurde bereits von Wang & Zhao beschrieben [2]. Für das Schätzen der relativen Translation sollte auf ein manuelles Ausmessen des Aufbaus der Sensorik oder nichtlineare Schätzverfahren, wie etwa jenes der Autoren Strobl & Hirzinger [5], zurückgegriffen werden.

Literatur

- [1] R. K. Lenz and R. Y. Tsai, "A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration," *IEEE Transactions on Robotics and Automation*, vol. 5, no. 3, pp. 345–358, 1989.
- [2] Y. Wang and L. Zhao, "Hand-Eye Calibration of IMU and Camera Without External Equipments," in *Lecture Notes in Electrical Engineering 360 the 2015 Chinese Intelligent Systems Conference*, vol. 2, 2015.
- [3] T. Shan and B. Englot, "LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain," *IEEE International Conference on Intelligent Robots and Systems*, pp. 4758–4765, 2018.
- [4] T. Moore and D. Stouch, "A generalized extended Kalman filter implementation for the robot operating system," in *Advances in Intelligent Systems and Computing*, vol. 302, pp. 335–348, Springer Verlag, 2016.
- [5] K. H. Strobl, G. Hirzinger, and K. S. De, "Optimal Hand-Eye Calibration," in *International Conference on Intelligent Robots and Systems*, pp. 4647–4653, 2006.

Lageerkennung der Wasseroberfläche durch semantische Segmentierung und Stereo Vision

Konstantin Christ, Dennis Grießer, Matthias Franz

Die Datenauswertung eines auf einem Boot installierten Sensorsystems wird durch kontinuierliche Informationen über die Lage im Wasser deutlich vereinfacht. In vielen Fällen wird dafür auf inertielle Messeinheiten zurückgegriffen. Diese sind jedoch mit einem Sensordrift behaftet und müssen zur Erkennung und Korrektur dieser Abweichung mit anderen Sensoren wie GPS oder Kamerasystemen kombiniert werden. Diese Arbeit beschäftigt sich mit der Entwicklung einer Lageerkennung durch Stereo Vision für das bestehende Sensorcluster des Forschungsboots *Solgenia*. Im Zuge dessen werden drei Netzarchitekturen zur semantischen Segmentierung der Wasseroberfläche evaluiert. Die zur Regression der Wasseroberfläche nötige Punktwolke wird durch Stereo Matching Verfahren erzeugt. In diesem Kontext wird die Leistungsfähigkeit einer angepassten Version des *Semi-Global (Block) Matching* Algorithmus [1] mit einem vortrainierten *AANet* [2] Disparitätsnetz verglichen. Darüber hinaus werden einige Anwendungsmöglichkeiten für die Lageschätzung in Kombination mit anderen Sensoren, wie Lidar oder Trägheitsnavigationssystem, aufgezeigt und prototypisch als *Robot Operating System* Knoten implementiert.

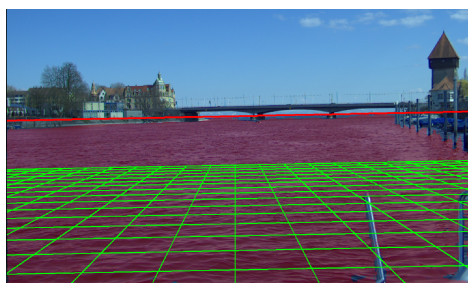


Figure 25: Detektion der Wasseroberfläche und des künstlichen Horizonts.

Hinsichtlich der semantischen Segmentierung

der Wasseroberfläche werden die Netzarchitekturen *FCN-AlexNet* [3], *U-Net* [4] sowie ein *VGG16 basierter Encoder-Decoder*⁴ verglichen. Als Trainings- und Validierungsdaten fungieren der *Tampere* [5] und *WaterNet* [6] Datensatz. Die Beurteilung erfolgt anhand der Vorhersagegenauigkeit für 30 Bilder des Seerheins und Bodensees, welche durch die Kameras des Zielsensorsystems aufgenommen und im Rahmen dieser Arbeit annotiert wurden. Die beiden Verfahren zur Disparitätsbestimmung werden zunächst direkt in Bezug auf den Disparitätsfehler der Vorhersage beurteilt. Die Ground-Truth-Disparitätskarten werden durch eine Projektion von Lidarmessungen in das Kamerabild erstellt. Mithilfe der semantischen Segmentierung und der nachfolgenden Disparitätsermittlung kann eine Punktwolke der Wasseroberfläche gewonnen werden, welche wiederum durch RANSAC in Verbindung mit linearer Regression durch eine Ebene angenähert werden kann. Im Kontext der somit realisierten Lageerkennung werden *Semi-Global Block Matching* und *AANet* erneut evaluiert. Zur Beurteilung der Lageschätzungen werden die Messungen des im Sensorcluster enthaltenen Trägheitsnavigationssystems herangezogen.

Es hat sich gezeigt, dass alle der verwendeten Architekturen das vorliegende Segmentierungsproblem hinreichend gut lösen können. Mit dem *WaterNet* Datensatz trainierte Netze, welcher unter anderem von den Testdaten stark abweichende Bilder enthält (etwa von Wasserfällen oder Schwimmbädern), erzielten deutlich bessere Ergebnisse als Modelle, die anhand des *Tampere* Datensatzes gelernt wurden. Dessen Bilder ähneln in großen Teilen dem Test-Szenario, sind aber zu einheitlich, um eine gute Generalisierung zu ermöglichen. Das *U-Net* eignet sich aufgrund der zu hohen Inferenzzeit nicht für den gegebenen Einsatzzweck und liefert trotz seiner komplexeren Struktur

⁴Das Netz wird vom Ersteller ungenauerweise als FCN8s bezeichnet: <https://github.com/pochih/FCN-pytorch>

keine besseren Ergebnisse als FCN-AlexNet und VGG16 Encoder-Decoder. Obwohl die Qualität der Disparitätsschätzungen des AANets im Allgemeinen vielversprechend ausfällt, wurde deutlich, dass sich zur dreidimensionalen Erfassung der Wasseroberfläche mit anschließender Regression ein entsprechend konfigurierter Semi-Global Block Matching Algorithmus besser eignet. Dazu ist anzumerken, dass ein auf Bilder des Straßenverkehrs (KITTI 2012) trainiertes Netz verwendet wurde und die Verwendung von geeigneterem Trainingsmaterial mutmaßlich mit einer signifikanten Verbesserung der Ergebnisse einhergehen würde.

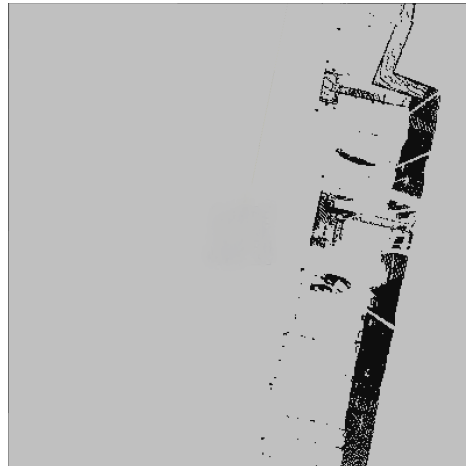


Figure 26: Belegtheitsgitter für Roboternavigation.

Die auf der semantischen Segmentierung und der Schätzung von Stereokorrespondenzen aufbauende Lageerkennung erzielt unter optimalen Bedingungen eine ähnliche Genauigkeit wie die Referenzmessungen des Trägheitsnavigationssystems. Liegen nicht ausreichend eindeutig zuordenbare Strukturen in der Wassertextur vor oder treten störende Reflexionen auf, fällt die Qualität der Lageschätzungen massiv ab. Da unter den meisten Bedingungen eine zuverlässige Lageerkennung möglich ist, können im Rahmen dieser Arbeit einige Anwendungsmöglichkeiten, wie das Erstellen von Belegtheitsgittern zur Hinderniserkennung unter Verwendung der Lidar-Punktwolke (26) oder die Kalibrierung von inertialen Messeinheiten, demonstriert werden.

Literatur

- [1] Hirschmüller, H., Accurate and efficient stereo processing by semi-global matching and mutual information, 2005.
- [2] Xu, Haoifei and Zhang, Juyong, AANet: Adaptive Aggregation Network for Efficient Stereo Matching, 2020.
- [3] Jonathan Long and Evan Shelhamer and Trevor Darrell, Fully Convolutional Networks for Semantic Segmentation, 2015.
- [4] Ronneberger, Olaf and Fischer, Philipp and Brox, Thomas, U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015.
- [5] Taipalmaa, Jussi and Passalis, Nikolaos and Zhang, Honglei and Gabbouj, Moncef and Raitoharju, Jenni, High-Resolution Water Segmentation for Autonomous Unmanned Surface Vehicles: a Novel Dataset and Evaluation, 2019.
- [6] Liang, Yongqing and Jafari, Navid and Luo, Xing and Chen, Qin and Cao, Yanpeng and Li, Xin, WaterNet: An adaptive matching pipeline for segmenting water with volatile appearance, 2020.

Transferlernen für Objektdetektion auf Binnengewässern

Simon Warmers, Dennis Griebner, Matthias Franz

Transfer Learning ist eine Methode, um Modelle mit eigenen Daten feinzutunen. In dieser Arbeit haben wir uns auf die Objekterkennung verschiedener Bootsklassen auf dem Bodensee konzentriert. Diese umfassen: Ruderboot, Motorboot, Segelboot mit Segel, Segelboot ohne Segel, Stand-Up-Paddle (SUP), Katamaran, Fähre, Tretboot, Motorschiff und Anlegestellen. In unserer Reihe von Experimenten haben wir ein vortrainiertes Faster Region-based Convolutional Neural Network [1] verwendet, das mit dem MS COCO-Datensatz [2] trainiert wurde. Dieser Datensatz umfasst über 200.000 annotierte Bilder mit 81 verschiedenen Klassen, von Menschen bis hin zu Flugzeugen. Dieser Datensatz enthält auch die Klasse Boot, unterscheidet jedoch nicht zwischen den verschiedenen Bootstypen. Aus diesem Grund wollten wir unser eigenes Modell trainieren, um die zehn verschiedenen Bootstypen zu klassifizieren. Die Methoden, die zur Verbesserung der Modelle verwendet wurden, waren Trainingsdauer, Augmentierung und Lernratenplanung.



Figure 27: Beispielhafte Detektionen.

Wir begannen damit, eine initiale Methode mit

den Rohbildern aus dem Datensatz zu erstellen. Basierend auf dem ersten Experiment haben wir die Lernzeit von 20 Epochen auf bis zu 60 Epochen erhöht. Dies führte zu einer insgesamt nicht signifikanten Verbesserung der Leistung basierend auf der IoU-Metrik. Doch in einer qualitativen Analyse auf neuen, nicht beschrifteten Daten wurde die Leistung mit weniger Abweichungen und einer sehr guten Vorhersagerate verbessert. Ein Beispiel für Vorhersagen dieses Modells ist in 27 zu sehen. Anschließend experimentierten wir mit der Verdopplung und Vervierfachung des Datensatzes durch Bildaugmentierungen. Dies führte zu einem guten Netzwerk. Die Generalisierungsleistung wurde durch Änderungen der Sättigung des Trainingsdatensatzes verbessert. Darüber hinaus haben wir Experimente mit der Planung der Lernrate durchgeführt, um die Leistung des Netzwerks zu verbessern.

Der Vergleich zeigt, dass mit der Transfer-Learning-Methode erstaunliche Ergebnisse für verschiedene Daten erzielt werden können, selbst ohne fortschrittliche Bildaugmentierungsmethoden. Mit 60 trainierten Epochen sahen wir den besten Leistungsanstieg. Wir schlossen die Forschung mit einem Plan für zukünftige Trainingsansätze ab, die sogar verwendet werden können, um das Modell mit einem kleinen Datensatz zu verbessern.

Literatur

- [1] Ren, S., He, K., Girshick, R., and Sun, J., Faster r-cnn: Towards real-time object detection with region proposal networks, 2015.
- [2] Lin, Tsung-Yi, et al., Microsoft coco: Common objects in context, 2014.

Farbkalibrierung mit spektrometrischen Referenzdaten für den Inline-Digitaldruck

Carsten Röhrich, Tobias Birkle, Matthias Franz

Einleitung

Im Rahmen dieser Arbeit werden vier verschiedene Regressionsmethoden zur Farbkalibrierung von RGB-Kameras getestet: die lineare Regression sowie die polynomiale Regression zweiten und dritten Grades. Diese werden mit einem Deep-Learning-Ansatz zur Regression verglichen, um festzustellen, ob dieser geeignet ist, die Kameras so genau zu kalibrieren, dass diese zur Qualitätskontrolle im Inline-Digitaldruck verwendet werden können.

Experimente

Um die Farbkalibrierung der RGB-Kameras vorzunehmen, wird ein empirisches Verfahren angewendet. Dabei werden mit den Kameras verschiedene Farbpatches aufgenommen, um die Kamera dann so zu kalibrieren, dass sie diese Farben korrekt wahrnimmt. Dafür wurden zwei mit einer Kalibrierungslampe gleichmäßig beleuchtete Colourchecker mit einem Spektroradiometer aufgenommen. Aus deren Spektraldaten wurden $L^*a^*b^*$ -Werte errechnet, deren euklidische Distanz ΔE der Wahrnehmung des Menschen für Farbunterschiede entspricht. Mit dem einen Colourchecker wurden dann die Modelle zur Kalibrierung erstellt bzw. trainiert, mit dem anderen wurde anschließend die Kalibrierung getestet. Für die Kalibrierung wurde eine Referenzkamera gewählt. Die RGB-Ausgangswerte der anderen Kameras wurden dann auf die der Referenzkamera kalibriert. Für die Referenzkamera wurde eine Kalibrierung in $L^*a^*b^*$ mit den vorher erstellten Labeldaten vorgenommen. Die Daten aller Kameras wurden dann mit diesem Modell in $L^*a^*b^*$ umgewandelt. Bei den anderen Kameras als der Referenzkamera wurden dafür vorab die RGB-Werte mit der Kalibrierung auf die Referenzkamera angepasst. Für jedes trainierte Modell wurde eine Kreuzvalidierung durchgeführt, um den bestmöglichen Trainingsdatensatz zu finden. Außerdem wurden in einer Gridsearch diverse Hy-

perparameter getestet, um ein Modell zu finden, das möglichst gut generalisiert.

Ergebnisse

Die maximale Abweichung ΔE der in $L^*a^*b^*$ kalibrierten Referenzkamera ist mit einem Wert von 1.29180 als gut zu bewerten, da sie dem Menschen kaum auffällt. Außerdem handelt es sich bei dieser Abweichung um eine Ausnahme, wie man in Abbildung 28 erkennt, welche die Abweichungen in ΔE farblich darstellt. Der Durchschnittswert ist mit 0.04590 noch wesentlich besser, denn Abweichungen dieser Größe sind für Menschen unmöglich zu erkennen.

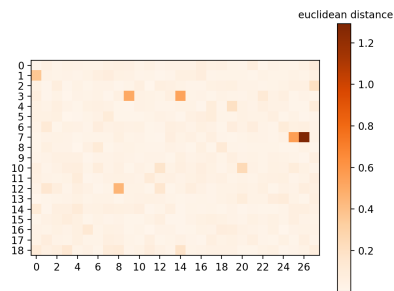


Figure 28: Referenzkamera 2 RGB zu $L^*a^*b^*$ Regression mit DL (Hyperparameter: Aktivierungsfunktion = Sigmoid, Fehlerfunktion = MSELoss, Lernrate = 0.01 und Zyklen, bis Senkung der Lernrate = 10), Validierung mit dem zweiten Fünftel des Datensatzes.

Mit Deep-Learning-Modellen können auch Artefakte im Bild einzelner Kameras gut beseitigt werden. Eine der in den Experimenten verwendeten Kameras besitzt eine vignettierte Linse (siehe Abbildung 29). Nach der Kalibrierung auf die Referenzkamera konnte Deep-Learning als einziges Kalibrierungsverfahren diesen Fehler beseitigen (siehe Abbildung 30).

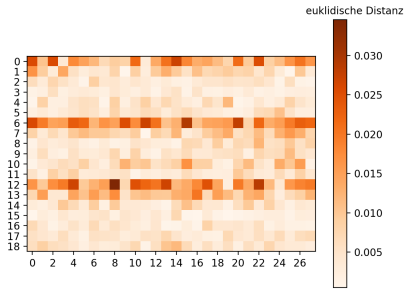


Figure 29: Vergleich der RGB-Werte einer Kamera mit vignettierter Linse zu einer anderen Kamera ohne vignettierter Linse.

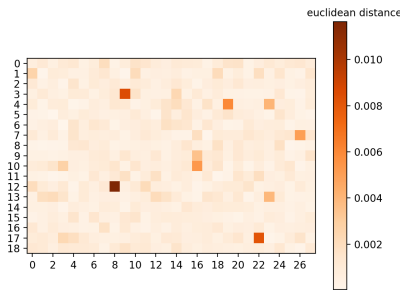


Figure 30: Aufnahmen des ersten Colourcheckers mit Kamera 1 und Referenzkamera 2. Validiert wurde mit dem ersten Fünftel des Datensatzes. Kamera 1 wurde mit einem DL-Modell auf Referenzkamera 2 kalibriert.

Für die Kalibrierung zur Referenzkamera hat sich die Aktivierungsfunktion ReLU, für die in $L^*a^*b^*$ hingegen Sigmoid als besonders geeignet erwiesen.

Fazit

Deep-Learning ist ein sehr gutes Verfahren zur Kalibrierung von Kameras für die Qualitätskontrolle. Die dabei erreichte Genauigkeit ist so hoch, dass Menschen Abweichungen nicht erkennen können. Besonders wichtig zum Erstellen eines guten Kalibrierungsmodells ist die Qualität der Labels. Außerdem ist es für bessere Ergebnisse von Vorteil, vorher zu testen, welche Teile des Datensatzes für das Training und welche für die Validierung besonders gut geeignet sind. Die Kreuzvalidierung hat sich hierfür als geeignet erwiesen. Mit der Grid-search kann das Kalibrierungsmodell dann weiter optimiert werden.

Analyzing keypoint detectors and descriptors learned on RGB-D images

Daniel Bencic, Matthias Hermann, Matthias Franz

Introduction

Keypoints are needed for many algorithms in computer vision and robotics. They determine important locations in images and can be used for algorithms like camera calibration, image stitching, object recognition or visual SLAM. With increasing amounts of data and computing power available, convolutional neural networks (CNNs) and deep learning are used for keypoint detection and description. Due to sparse connectivity and weight sharing they can learn spatially local features and are well suited to solve this problem. This thesis analyzes existing learned keypoint detectors and descriptors and proposes a learned detector for occlusion edges. It is learned on multimodal RGB-D images and based on handcrafted keypoint detectors. In addition to proposing a keypoint detector, a working fully convolutional keypoint descriptor for 3D pointclouds is transferred to 2D RGB-D images and evaluated.

Approach

In this thesis the problem of learning keypoint response maps is formulated as binary image segmentation. The proposed descriptor is therefore based on the fully convolutional U-Net [1] architecture. To speed up optimizer convergence, a second architecture is built by combining the U-Net architecture with residual learning proposed by He et. al. [2]. The ground truth used for training these architectures is calculated by the handcrafted keypoint detectors Harris Corner Detector and Canny Edge Detector. The networks are trained in a unimodal setting, where response maps are learned for grayscale and depth images separately and in a multimodal setting, where the network should predict response maps calculated on the depth image when given a grayscale image. The descriptor is adapted from Choy et. al. [3] and uses deep metric learning to cluster spatially local features between RGB channels and depth channel to extract features present in both modalities.

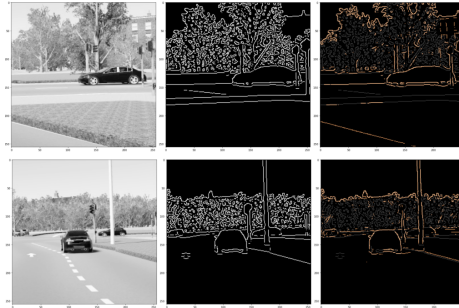


Figure 31: Example of a network generalizing the concept of occlusion edges. (Top) Network correctly predicts an occlusion edge along the curb, that is not present in the ground truth. (Bottom) Network correctly doesn't predict an occlusion edge around the traffic arrow on the street.

Findings

Performing Harris Corner Detection and Canny Edge Detection on grayscale and depth images separately is a problem that can be learned well by the fully convolutional neural networks with Intersection over Union (IoU) metrics above 70% for 256×256 pixel images. The multimodal detection of corner keypoints yields no usable results. Occlusion edge detection however yields good response maps predicted by the networks (31). The from Choy et. al. [3] adapted keypoint descriptor is optimized with a "hardest contrastive" loss and this loss assumes features not being independently and identically distributed. This is the case for 3D pointclouds but not for uniformly distributed pixels of RGB-D images. Therefore it does not produce good results.

Conclusion

The conducted experiments in this thesis showed, that learning traditional gradient-based handcrafted keypoint detectors on unimodal images with CNNs is a solvable problem. CNNs are also able to benefit from multisensor RGB-D data to learn occlusion edge keypoints that are invariant

to translation and rotation. It could also be shown, that the hardest contrastive loss proposed by Choy et. al [3] indeed relies on the assumption that the distribution of features is locally correlated.

Literatur

- [1] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation, 2015.
- [2] K. He, X. Zhang, S. Ren, and J. Sun, Deep residual learning for image recognition, 2016.
- [3] C. Choy, J. Park, and V. Koltun, Fully convolutional geometric features, 2019.

Eine Untersuchung über den Einsatz von Ensemble-Methoden in der Anomalieerkennung

David Wiessner, Matthias Hermann, Matthias Franz

Einführung

In diskriminativen Modellen des maschinellen Lernens hat sich die Berechnung und Nutzung der Modellunsicherheit durch Verfahren wie Variational Inference oder MC-Dropout bewährt. Ein weiterer vielversprechender Ansatz ist die Nutzung von Modell-Ensembles. So werden in Deep-Ensembles Modelle unterschiedlich initialisiert und anschließend trainiert (vgl. [1]). Die Ergebnisse konnten hinsichtlich der Erkennungsleistung überzeugen und konkurrierende Ansätze sogar übertreffen. Betrachtet man aktuelle Methoden in der Anomalieerkennung bzw. Novelty-Detection, scheinen diese Ansätze kaum eine Rolle zu spielen. Da in diesem Bereich Ergebnisse für gewöhnlich in Form des ROC-AUC-Scores angegeben werden, sind Unsicherheiten nicht sichtbar. Im Rahmen dieser Masterarbeit soll untersucht werden, ob Ensemble-Methoden in der Novelty-Detection zur Ermittlung von Modellunsicherheiten genutzt werden können, um Verbesserungen bei der Erkennungsrate zu erzielen oder praktische Anwendungsszenarien robuster zu gestalten.

Projektbeschreibung

Die Untersuchungen werden auf den Modellen Slow-PCA, ICA, Tiled-ICA und VAE, unter Verwendung der Datensätze MNIST und CIFAR10, durchgeführt. Im Rahmen der Masterarbeit wird die allgemeine Theorie von Latent-Space Modellen sowie die Ansätze der verwendeten Modelle erarbeitet. Zudem werden Ensemble-Ansätze erschlossen und implementiert. Die Modell-Unsicherheit wird durch die Varianz der Novelty-Scores über die Ensemble-Modelle erfasst. Durch unterschiedliche Experimente wie die Nutzung des Ensemble-Wissens und Verrechnung der Score-Varianzen wird gezeigt, dass eine direkte Verbesserung des ROC-AUC-Scores im Vergleich zu gut konditionierten (einzelnen) Modellen nicht möglich ist. Aufgrund der ausgebliebenen

Verbesserungen der Scores werden mögliche Ursachen unter Beachtung des aktuellen Forschungsstandes untersucht und diskutiert. Da in einem praktischen Anwendungsszenario konkrete Grenzwerte für Novelty-Detection Modelle benötigt werden, wird in einem Experiment die Einführung eines Unsicherheitsgrenzwertes untersucht. Dieser soll dazu dienen Inspektionsaufgaben robuster zu gestalten, indem Datenpunkte mit stark variierenden Scores erkannt und gefiltert werden. Hierbei handelt es sich um einen vielversprechenden Ansatz für die Nutzung von Ensembles.

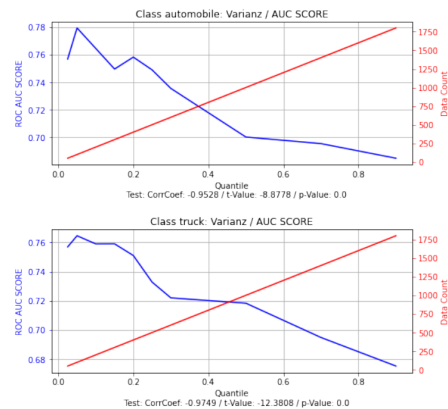


Figure 32: RA-Scores nach Filterung - Slow PCA - Norm z - Score.

In 32 ist ein Ausblick über die Verbesserungen der ROC-AUC-Scores auf Basis des Norm z-Scores für die Klassen Automobile und Truck durch Filterung unsicherer Datenpunkte erkennbar.

Literatur

- [1] Fort, S. and Lakshminarayanan, B. Deep Ensembles: A Loss Landscape Perspective, 2019.
- [2] Pimentel, M. A.F. and Clifton, A. and Clifton, L. and Tarassenko, L., A review of novelty detection, 2014.

Intelligente Bildverarbeitung in der Materialmikroskopie durch maschinelles Lernen zur Objekt- und Fehlererkennung ⁵

Fabian Freiberg, Matthias Franz

Ziel und Inhalt

Die materialmikroskopische Charakterisierung ist weit verbreitet, um quantitative Zusammenhänge zwischen Mikrostruktur (Phasenanteile, Kornstrukturen, Texturen, Gefügefehler) und Eigenschaften für Entwicklungs- und Qualitätssicherungsaufgaben zu ermitteln. Viele Fragestellungen z.B. Fehler in Gefügen und geometrische Abweichungen in Strukturen sind ohne Automatisierung nicht mehr umsetzbar. Vorhandene schwellwert- und objektbasierte Algorithmen für die Ermittlung von Strukturmerkmalen stoßen aufgrund Rechenaufwands, zunehmender Komplexität der Bilder und Datenaufkommen an ihre Grenzen. Neuentwicklungen bei Licht-, Rasterelektronen- und 3D-Röntgenmikroskopie/-tomographie stellen diese Strukturen zwar immer besser dar, treiben aber gleichzeitig die Datenmenge und Detailtiefe in den Bildern in die Höhe - vorhandene Algorithmen zur Auswertung halten mit dieser Entwicklung nicht Schritt. In anderen Bereichen sind Mustererkennung und lernende Algorithmen zur Detektion von Auffälligkeiten in Bildern bereits weiterentwickelt. Diese Ansätze müssen auf die Materialmikroskopie übertragen, d.h. industriell tauglich, robust und präzise implementiert werden. Zielstellung von KLEVER ist deshalb im Kontext einer automatisierten und intelligenten Materialmikroskopie durch Kopplung der Bereiche eine robuste automatisierte Bildakquisition bei 2D und 3D-Materialmikroskopie, und eine intelligente Bildverarbeitung mit Ansätzen aus Bereichen des Machine Learning einschließlich Deep Learning bereitzustellen. Im Zentrum stand dabei durch Anwendung von Large Data Processing, eine solidere technologische Basis zur Objekt-, Struktur- und Fehlererkennung (z.B. Schichtdickenschwankungen, geometrische Abweichungen, Poren, Risse, Fremdeinschlüsse) zu schaffen. Darüber hinaus wurden vorhandene Konzepte und Algorithmen der objekt- und musterbasierten Ver-

fahren bezüglich Präzision und Robustheit zum Suchen und Finden von Strukturen und Abweichungen bewertet und angewandt.

Wesentliche Ergebnisse

Im Rahmen des Projekts KLEVER konnte auf Seiten der HTWG Konstanz und im speziellen des Instituts für Optische Systeme die angestrebte intelligente Bildverarbeitung zur Objekt-, Struktur- und Fehlererkennung umgesetzt werden. Die Fehlererkennung der implementierten Methode, basiert auf einer Multiobjekterkennung auf Basis tiefer neuronaler Netze und ist in der Funktionsweise an die Faster-RCNN-Methode angelehnt. Des Weiteren wurden verschiedene Merkmalsextraktoren und die daraus resultierenden Merkmale verglichen, um eine mögliche effiziente Vorabrechnung der Merkmalsextraktion zu erwirken und somit den Trainingsprozess der künstlichen neuronalen Netze zu beschleunigen und für große Datenmengen zu optimieren. Insbesondere Techniken aus dem Bereich optimaler Signalkodierung stellten sich hierfür als geeignet dar. Durch den Einsatz neuronaler Netze und deren effizienter Implementierung auf Mehrkernprozessoren und Grafikeinheiten (GPUs) konnte darüber hinaus die technologische Basis für Large Data Processing gelegt, im Rahmen der Gefügeanalyse validiert, und im Sinne der Interdisziplinarität mit der Materialforschung am IMFAA in Aalen geteilt werden.

Nutzen, Anwendungsmöglichkeiten und Notwendigkeit

Bei dem Projekt KLEVER handelte es sich um eine vorwettbewerbliche Grundlagenaktivität mit breitem Nutzen für die Branchen der Materialwissenschaften. Gleichzeitig ist es interdisziplinär angelegt und versucht Synergien zwischen Maschinenbau, Physik und Informatik zu generieren. Durch die im Projekt entstanden

⁵Kurzbericht eingereicht bei BMBF.

Softwareentwicklungspipeline konnten so effektiv neue Algorithmen aus dem Bereich Bildverarbeitung und Machine Learning in Richtung Anwendung gebracht werden. Die in KLEVER entwickelten Lösungen bieten enormes Potential zur Effizienzsteigerung etwa in der mikroskopischen Qualitätssicherung fertiger Unternehmen. Hier sind häufig Bildanalyse-Algorithmen noch nicht GPU- oder mehrkernprozessorenfähig, was oftmals die Analyse hochauflöster Bilder erschwert. Die im Rahmen des Projekts entwickelte Methode auf Basis der Faster-RCNN-Methode zur Identifikation von Gefügefehlern in Li-Ion-Akkus ist sowohl ressourceneffizient als auch zuverlässig hinsichtlich Erkennungsrate. Hieraus ergeben sich große Potentiale im Bereich Automation. Durch den datengetriebenen Ansatz (Machine Learning) ist darüber hinaus die Übertragbarkeit auf andere Fragestellungen gegeben und oftmals nur eine Frage der Datenakquise. Für die Qualität der Objekt- und Fehlererkennung ist im Wesentlichen die die Merkmalsextraktion verantwortlich. Hier konnte im Rahmen des Projekts festgestellt werden, dass transferierte Neu-

ronale Netze große Gemeinsamkeiten mit der Theorie optimaler Kodierung aufweisen und mittels Independent Component Analysis (ICA) angenähert werden. Diese Methoden bieten darüber hinaus durch ihre Interpretierbarkeit sehr viel Potential die Auswirkungen von durch den Anwender gewählten Parametern auf die Analyseergebnisse zu reduzieren und somit die Aussagekraft der Untersuchungen zu erhöhen. Zudem konnte der Mehrwert von Deep Learning gezeigt werden. Dies lässt sich im Projekt beispielsweise durch den Einsatz des Grad-CAM sehen, der es erlaubt Analyseergebnisse für den Anwender nachvollziehbar und transparent darzustellen. Das Risiko der Nicht-Zielerreichung in Verbindung mit dem vorwettbewerblichen Charakter des Projekts machte eine öffentliche Förderung des Vorhabens erforderlich, da gerade die Erprobung neuartiger Algorithmen wie das Wasserstein-GAN sehr aufwändig ist, aber oftmals nicht zum erwünschten Erfolg führt. Außerdem stehen Unternehmen häufig keine Mitarbeiter mit nötigen Informatik-Fachwissen zur Verfügung, wodurch fachübergreifende Kooperationen notwendig sind.

Multifunktional-skalierbare generische Inlinelnspektion für flexible Fertigungsprozesse in vernetzten Produktionsanlagen ⁶

Matthias Hermann, Matthias Franz

Dieses Teilvorhaben bezieht sich auf die Bekanntmachung Photonik für die flexible, vernetzte Produktion – Optische Sensorik im Rahmen des Förderprogramms Photonik Forschung Deutschland. Das Teilvorhaben Generische Texturmodelle unter Einbeziehung psychophysischer Aspekte wurde im Verbundprojekt Multifunktional-skalierbare generische Inlinelnspektion für flexible Fertigungsprozesse in vernetzten Produktionsanlagen (MultiFlexInspect) zusammen mit den Partnern Baumer Inspection GmbH, Chromasens GmbH, Silicon Software GmbH (heute Basler AG), sowie dem assoziierten Unternehmen Classen Industries GmbH durchgeführt. Im Kontext des Gesamtverbunds, liefert das Teilvorhaben die technologische Basis zur Inspektion dynamisch variierte Prüflinge (Losgröße 1) ohne aufwendigen Einlernprozess. Konkrete Aufgabe im Teilvorhaben war es intelligente Algorithmen aus dem Bereich maschinelles Lernen und künstliche Intelligenz einzusetzen und die in der Ausschreibung geforderte Flexibilität und Autonomie aktueller Produktionsmaschinen zu erhöhen. Die photonische Erfassung komplexer Zustände, im Projekt sichtbarer bzw. menschlich nicht wahrnehmbarer Defekte in 2D- Textur sowie 3D-Relief, gepaart mit intelligenten Erkennungsmethoden, sollten im Rahmen des Projekts die Forderung nach einer schnellen und verlässlichen Qualitätskontrolle auch kleinster Stückzahlen (Losgröße 1) ohne zusätzliche manuelle Anpassung lösen können. Zudem sollte das Vorhaben mit seinem Fokus auf Defektdetektion sowie Defektklassifikation direkt auf die Forderung nach exakten Kundenvorgaben abzielen. Konkretes Ziel des Teilvorhabens war es, die Entwicklung von Algorithmen zur intelligenten, an der menschlichen Wahrnehmung orientierten Inspektion für farbig texturierte 3D-Oberflächen voranzutreiben. Im Kontext des Verbundprojektes sollte damit die technologische Basis zur Detektion von allgemeinen Oberflächenanomalien (Novelty Detec-

tion) in 2D und 3D sowie deren Klassifikation mit der besonderen Berücksichtigung psychophysischer Aspekte geschaffen werden. Besondere Schwierigkeit bestehen hierbei bei Oberflächen wie bspw. Laminatböden, da hier sowohl Farbtextur als auch Relief den Qualitätsansprüchen der Kunden genügen muss. Hinzu kommt das Problem der Übersortierung: Viele in Druckanlagen entstehende Fehler sind für den menschlichen Betrachter nicht sichtbar.

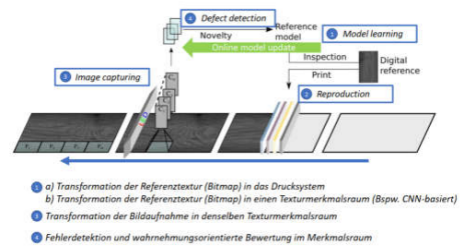


Figure 33: Schematische Darstellung des entwickelten intelligenten Systems zur Inline-Inspektion auf Basis von Novelty Detection.

33 zeigt das im Rahmen des Projekts entwickelte intelligente Inspektionssystem, das auf Basis von Novelty Detection arbeitet. Eine solche Inspektionsanlage ist flexibel und kann eigenständig auf Parameteranpassungen der Produktionsanlagen reagieren. Hierdurch kann eine kosteneffiziente Produktion auch kleinster Stückzahlen gewährleistet werden. Die implementierten Algorithmen sind humanperzeptorisch motiviert und erlauben aufgrund Ihrer Einfachheit eine echtzeitfähige Inline-Inspektion von 2D/3D-Farboberflächen. Zentrales Element des Systems ist die Wahl des Texturmodells, das auf vorinitialisierten Convolutional Neural Networks (CNN) basiert. Ein solches Modell erlaubt sowohl eine hochgenaue Detektionsrate und ist parallelisierbar. Durch die Wahl eines biologisch

⁶Schlussbericht eingereicht bei BMBF.

motivierten Texturmodells, kann darüber hinaus die menschliche Wahrnehmung in den Inspektionsprozess einbezogen werden. Mittels Kombination der beiden Texturmodelle können so Fehler im Produktionsprozess effektiv reduziert werden. Die Bedeutung der Ergebnisse für das betreffende Technologiefeld konzentriert sich im Wesentlichen auf die Wahl des Texturmodells und die verwendeten Fehlerbewertungsfunktionen. Gerade in Bezug auf das 3D-Oberflächen-Relief im Verbund mit Farbtexturen war hier intensive Forschungsarbeit zu leisten, da Ansätze aus der

Literatur zwar bekannt, aber nicht ausreichend erforscht und bisher nicht im geplanten Kontext eingesetzt wurden. Weiter wurden zum ersten Mal psychophysische Modelle für eine humanperzeptorisch motivierte Merkmalsextraktion und Qualitätsbewertung im Rahmen von Inline-Inspektion angewandt. Um die für die praktische Anwendung nötige Flexibilität zu gewährleisten, wurden die entwickelten Modelle – Texturmodell und Psychophysik – stets als neuronale Netze formuliert, was den flexiblen Einsatz in generischen Inlineinspektionsprozessen vereinfacht.

Punktwolkenklassifikation maritimer Daten

Thimo Schaub, Dennis Grießer, Georg Umlauf

Das Projekt beschäftigt sich mit der Klassifikation von maritimen Daten. Hierbei liegt der Fokus, über der Klassifikation hinaus, auf der Schätzung der Bounding Boxen sowie der Unterteilung der Daten, die von einem LiDAR Sensor aufgenommen wurde.

Versuche

Diese drei Aspekte wurden nacheinander anhand der Trainingsdaten evaluiert und abschließend als Gesamtanwendung ausgewertet. Für die Klassifikation und als Grundlage für das neuronale Netz, welches die Bounding Box für die Objekte erstellen soll, wurde hierfür das von Qi et al.[1] vorgestellte PointNet verwendet. Die vorgestellte Architektur zeichnet sich besonders dadurch aus, dass sie Punktwolken als Eingaben verwendet und somit keine aufwendige Vorverarbeitung oder enormen Speicheraufwand benötigen. Das Ergebnis hiervon ist eine zu 60 Prozent genaue Klassifikation von sechs Klassen, welche sich aufteilen in Motorschiff, Motorboot, Markierungen, Kataran und Segelboote mit und ohne Segel. Hierbei ist jedoch anzuprechen, dass die Klassifikation von Segelbooten mit und ohne Segel und die der Markierungen bei über 90 Prozent liegt. Lediglich die motorgetriebenen Schiffe lassen sich mit dem gelernten Model schlecht unterscheiden. Aufbauend auf diesem Netzwerk werden dann die Parameter der Bounding Box der einzelnen Objekte geschätzt. Da jedoch die PointNet Architektur vor der Klassifikation BatchNorm-Layers verwendet, wurde hierfür das Netzwerk angepasst, um diese zu umgehen und folglich nur die Featurevektoren von PointNet zu verwenden. Es lässt sich zeigen, dass das Netzwerk eine Überschneidung der Bounding Boxen von 29 Prozent erreicht. Es ist jedoch an dieser Stelle zu erwähnen, dass die Rotation nicht mithilfe der von PointNet [1] vorgestellten Architektur durchführbar ist, aber mit der PCA ausreichend gut berechnet werden kann. Die letzte Aufgabe, die in diesem Zusammenhang bearbeitet werden muss, ist die Aufteilung der Objekte aus der

Gesamtaufnahme. Hierfür wurde der Clusteralgorithmus DB-Scan verwendet. Dieser zeichnet sich dadurch aus, dass er nicht konvexe Cluster anhand der Nachbarschaft erkennt. Mithilfe dieses Algorithmus kann eine Überschneidung von 50 Prozent zwischen den Punktwolken der manuellen Annotation und den berechneten Punktwolken erreicht werden.

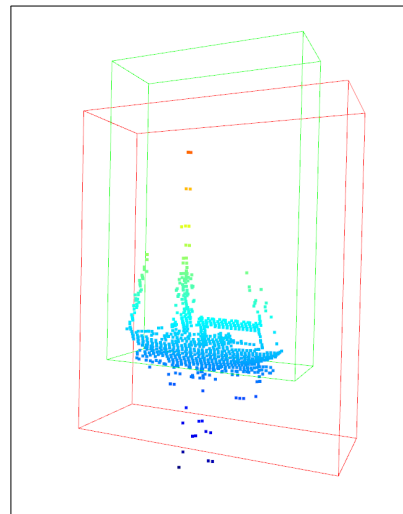


Figure 34: Beispiel der Bounding Box Regression anhand der von DB-Scan berechneten Punktwolke.

Ergebnisse und Fazit

Die aus diesen Bestandteilen zusammengesetzte Lösung erkennt mit einer Frequenz von maximal zwei Hertz die Klasse des Objekts zu 60 Prozent (Figure 35) und die Bounding Box mit einer Genauigkeit von 29 Prozent, Figure 34.

Die gezeigten Beispiele lassen sich exemplarisch für die Auswertung der Arbeit betrachten. Es ist zu sehen, dass die Klassifikation der motorgetriebenen Schiffe nicht ausreichend erfolgt. Ebenso lässt sich erkennen, dass die Bounding Box nur anhand der Punkte geschätzt wird und nicht wie erwartet die Form des gesamten Ob-

jekts erkennt. Des Weiteren zeigt sich, dass DB-Scan die Punktwolke zwar gut erkennt aber Reflexionen oder Wellen um das Objekt mitberechnet. Somit lässt sich abschließend sagen, dass das Netz wichtige Bestandteile erkennt, jedoch weitere Untersuchungen nötig sind, bevor hieraus eine Anwendung entstehen kann.

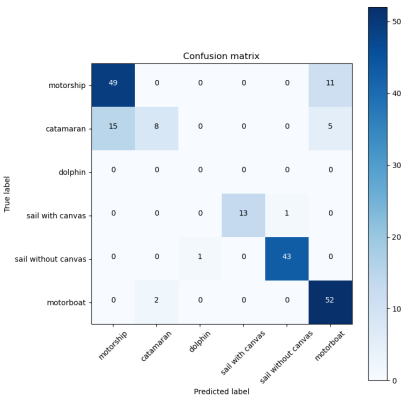


Figure 35: Ergebnisse der Klassifikation.

Literatur

[1] C.R. Qi, H. Su, K. Mo, L.J. Guibas. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation, CVPR, 2017.

Analyse von latenten Gaußschen Mischmodellen zur Registrierung von Punktwolken

Lukas Luschin, Matthias Hermann, Georg Umlauf

Punktwolkenregistrierung wurde durch Anwendung in der Robotik, im maschinellen Sehen und der dreidimensionalen Computergrafik zu einem fundamentalen Problem. Sowohl iterative Algorithmen als auch Ansätze mit neuronalen Netzen arbeiten beim Lösen dieser Problemstellung auf Punkt-zu-Punkt-, Merkmal-zu-Merkmal- oder Verteilung-zu-Verteilung-Korrespondenzen.

In dieser Arbeit soll der neue Ansatz DeepGMR [3] zur Punktwolkenregistrierung analysiert und hinsichtlich dessen Einsatz im Bereich des autonomen Fahrens getestet werden. Er basiert auf einer neuronalen Netzarchitektur, welche die Parameter mehrerer Gaußverteilungen zur Clusterung der Punktwolke konsistent lernt, um so eine punktweise Klassifizierung der Daten zu erreichen. Die resultierenden Cluster werden dann mithilfe einer Erweiterung des Rigid-Body-Transform Verfahrens zueinander ausgerichtet, um so eine globale Transformation zu schätzen.

Methodik

Für die Tests werden Punktwolken aus verschiedenen Datensätzen mit unterschiedlichen Eigenschaften bezogen. Der ModelNet Datensatz [2] beinhaltet mit CAD modellierte Objekte und dessen Repräsentation als Punktwolke. Er zeichnet sich durch eine gleichmäßige, rauschfreie Datenwolke aus und wird für das grundlegende Training des Modells verwendet. Die Generalisierungseigenschaften und die Robustheit des Netzes werden dann mit dem ShapeAim Datensatz, welcher auch mit einem Laserscanner aufgezeichnete Objekte beinhaltet, validiert. Als drittes werden LIDAR-Scans des Kitti Datensatzes [1] auf das Eingabeformat des Korrespondenznetzes des DeepGMRs angepasst, um so die Tauglichkeit im Straßenverkehr zu testen. Zum Schluss werden potenzielle Verbesserungsmöglichkeiten aufgestellt und implementiert, um eine erweiterte Genauigkeit des Systems zu erzielen.

Für die genaue Analyse des Registrierungsprozesses werden Evaluations- und Loggerskripte entwickelt, welche einen genaueren Einblick in die Zwischenschritte des Modells ermöglichen.

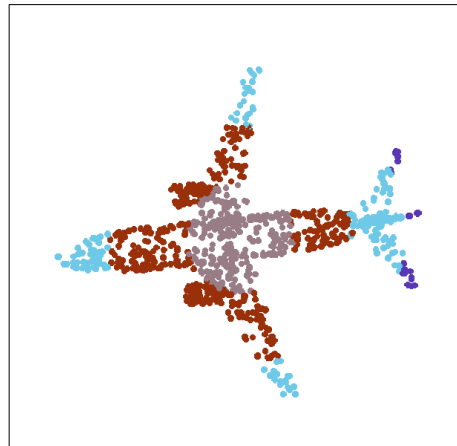


Figure 36: Ergebnis der Parameterschätzung des Korrespondenznetzes als GMM Plot.

Ergebnisse und Zusammenfassung

Die Ergebnisse, die in dieser Arbeit erzielt wurden, zeigen, dass das DeepGMR Modell eine präzise und schnelle Registrierung auf objektrepräsentierenden Punktwolken, welche in ModelNet und ShapeAim vorkommen, liefert. Auch bietet die Repräsentation der Datenwolke als GMM Verteilung eine ausreichende Robustheit gegenüber Permutationen und Samplingunterschieden, da die resultierenden Verteilungsparameter die Merkmale des Objektes immernoch stabil zusammenfasst.

Jedoch ist die Information der reduzierten Punktwolken der LIDAR-Scans im Kitti Datensatz nicht ausreichend, um konsistente GMM Parameter zu generieren. Die Darstellung der vom Korrespondenznetz erstellten Verteilungen zeigt, dass die Klassifizierung der einzelnen Punkte in

Cluster unter der geringen Anzahl an Punkten unzureichend und ungenau ist.

Abschließend führten diese Tests zu dem Ergebnis, dass durch aktuelle Hardwaregrenzen, eine Registrierung der Kitti LIDAR-Scans nicht möglich ist.

Literatur

[1] A. Geiger, P. Lenz, C. Stiller & R. Urtasun. Vision meets robotics: The KITTI dataset. Inter-

national Journal of Robotics Research, 2013.

[2] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang & J. Xiao. 3d ShapeNets: A deep representation for volumetric shapes. CVPR, 2015.

[3] W. Yuan, B. Eckart, K. Kim, V. Jampani, D. Fox, & J. Kautz. DeepGMR: Learning latent Gaussian mixture models for registration. ECCV, pp. 733–750, 2020.

Kamera-Lidar-Kalibrierung mit Partikel-Schwarm-Optimierung

Bernhard Gundel, Matthias Hermann, Georg Umlauf

Das computergestützte Sehen spielt in der mobilen Robotik eine immer größere Rolle. Durch die Erkennung, Klassifizierung und das Tracking von Objekten sind autonome Fahrzeuge in der Lage, Umgebungen interpretieren zu können. Die Voraussetzung für diese Fähigkeit ist das Zusammenspiel verschiedenster Sensoren. Radar-, Sonar- und Lidar-Sensoren werden eingesetzt, um Entfernungen zu messen und die Tiefeninformation einer Umgebung interpretieren zu können. Im Zusammenspiel mit RGB-Kameras lassen sich Objekte einfacher klassifizieren und identifizieren. Die automatische extrinsische Kalibrierung der Sensoren ist die Herausforderung, mit der sich dieses Projekt beschäftigt. Wie ein Lidar-Sensor und eine Kamera räumlich zueinander ausgerichtet sind, lässt sich mit Hilfe einer Transformationsmatrix beschreiben. Diese Transformationsmatrix wird durch die Maximierung der Normalized Mutual Information zwischen Kamerabild und projizierten Lidar-Punkten mit verschiedenen Methoden geschätzt. Die zu vergleichenden Optimierungsmethoden beinhalten eine Partikel-Schwarm-Optimierung, einen Gradientenaufstieg und ein künstliches neuronales Netz. Die Partikel-Schwarm-Optimierung wird implementiert, verbessert und beschleunigt, um eine möglichst genaue Schätzung der Transformationsmatrix zur extrinsischen Kalibrierung zu erzielen. Die zu optimierende Normalized Mutual Information ist definiert als [1]

$$\text{NMI}(X, Y) = \frac{H(X) + H(Y)}{H(X, Y)},$$

wobei $H(X)$ und $H(Y)$ jeweils die Entropie und $H(X, Y)$ die Joint-Entropie von X und Y ist.

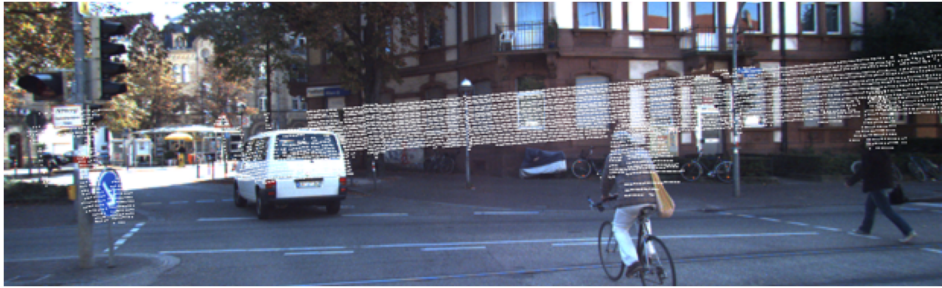
Vorgehen

Im Vorfeld wird die Partikel-Schwarm-Optimierung mit dem Ziel der Maximierung der Normalized Mutual Information mit Python implementiert. Experimente sollen zeigen, welche Parametrisierungen und welche Lidar-Features am

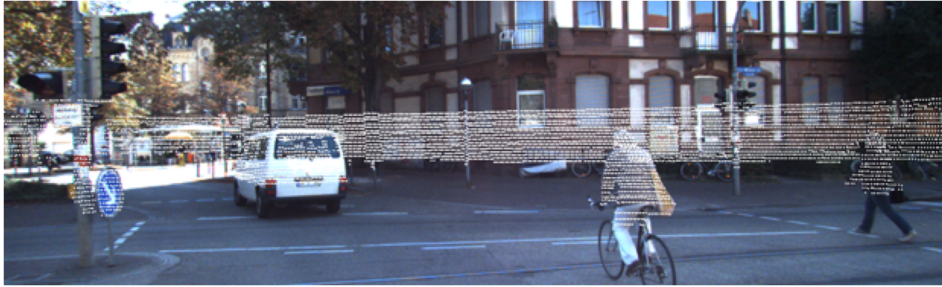
geeignetsten sind. Die Experimente werden mit der Nutzung drei unterschiedlicher Datensätze durchgeführt: Dem KITTI-Datensatz, der am Karlsruher Institut für Technologie als Benchmark für computergestütztes Sehen entwickelt wurde und zwei Multisense-Datensätzen, die mit ähnlichen, aber höher aufgelösten Kamerabildern und Punktwolken an der HTWG Konstanz entwickelt wurden. Die Experimente werden anschließend für die anderen beiden Methoden wiederholt. Zur Bewertung der Genauigkeit der jeweiligen Methode wird die Ground-Truth, also die angenommene wahre Lösung des Optimums sowohl in der Rotation, als auch in der Translation gleichermaßen verzerrt. Die Differenz zwischen geschätzter und wahrer Lösung, sowie die Standardabweichungen, Mittelwerte und Mediane über mehrere Runs werden als Performancemetriken für die Genauigkeit der jeweiligen Methoden eingesetzt.

Erkenntnisse

Es zeigt sich, dass die Wahl der Parameter für die Partikel-Schwarm-Optimierung, sowie die Wahl des Lidar-Features ausschlaggebend für die Genauigkeit dieser Methode sind. Die Schätzung der Normalenvektoren als Features stellt sich wegen der Spärlichkeit der Punktwolken als ungeeignet heraus. Die Reflektanz hingegen liefert relativ zuverlässige Ergebnisse mit geringen Differenzen zur Ground Truth und Standardabweichungen. Außerdem lässt sich aufzeigen, dass die Methoden auf den Datensätzen unterschiedlich gut performen. Unterschiedliche Lidar-Intensitäten der Datensätze sind verantwortlich für die Divergenz der Methode des Gradientenaufstiegs und des neuronalen Netzes. Da das neuronale Netz auf den KITTI-Datensatz trainiert wurde, erreicht dieses für die Multisense-Daten für die gesuchte optimale Translationsmatrix sehr ungenaue Schätzungen mit hohen Abweichungen über mehrere Runs. Falls das neuronale Netz jedoch mit den passenden Daten trainiert wird, ist es den beiden anderen Methoden sowohl in der



(a) Resultierendes Bild mit verzerrter Ground Truth vor der Optimierung.



(b) Resultierendes Bild mit verzerrter Ground Truth nach der Optimierung.

Figure 37: Projektion der Lidar-Punkte auf das Kamerabild vor und nach der Optimierung.

Geschwindigkeit, als auch in der Genauigkeit deutlich überlegen. Abbildung 37 zeigt die Projektion der Lidar-Punkte auf das Kamerabild vor und nach der Schätzung der Transformation mit der Partikel-Schwarm-Optimierung.

Literatur

[1] Z. Taylor, J. Nieto. A Mutual Information Approach to Automatic Calibration of Camera and Lidar in Natural Environments, ACRA 2012.

Untersuchung des Multi-View 3D-Rekonstruktionspotentials von Plenoxels und FastNeRF

Niklas Strittmatter, Matthias Hermann, Georg Umlauf

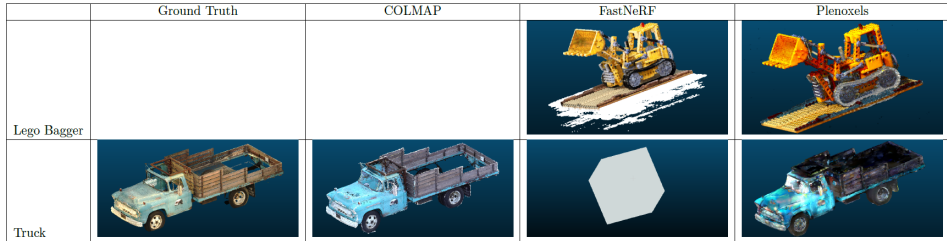


Figure 38: Vergleich der Technologien anhand von Bildschirmaufnahmen der erzeugten Rekonstruktionen. Dabei wurde die Plenoxels Punktwolken noch mit einem Filter bearbeitet, um Ausreißer zu filtern.

Das Thema des Projektes ist die Untersuchung und Evaluierung zweier Methoden zu Novel View Synthesis auf ihr Potential zur Multi-View 3D-Rekonstruktion: Plenoxels von Yu et al [1] und FastNeRF von Garbin et al [2]. 3D-Rekonstruktion aus Bildern gewinnt immer mehr an Bedeutung, da sie in Bereichen wie Robotik, Virtual Reality und Augmented Reality eingesetzt wird, welche immer präsenter werden. Im ersten Experiment wird der NeRF-Synthetic Lego Datensatz von Mildenhall et al [3] qualitativ untersucht. Dabei ergibt sich, wie in Tabelle 1 zu sehen ist, dass beide Methoden qualitativ gute Ergebnisse liefern. Da in beiden Methoden allerdings die Farbe der Voxel von der Blickrichtung abhängt, ist die Farbtreue nicht sehr genau. Das bestätigt sich auch in den qualitativen Ergebnissen der anderen Experimente. Die Zeitmessungen der beiden Methoden zeigt auch deutlich, dass Plenoxels mit ca. 10 Minuten Trainingszeit einen deutlichen Vorteil gegenüber FastNeRF hat, welches für die Rekonstruktion ca. 6 Stunden benötigt. Das zweite Experiment soll die eine qualitative Auswertung noch um einen quantitativen Aspekt erweitern. Dafür nutzen wir den Datensatz Tanks & Temples Truck von Knapitsch et al [4], mit von Liu et al [5] entfernten Hintergründen. Wir nutzen diesen Datensatz, weil er eine Ground-Truth Punktwolke mitliefert und wir diese durch die entfernten Hintergründe besser damit vergleichen können. Zum

Vergleich nutzen wir die Hausdorff-Distanz. Die Hausdorff-Distanz ist eine

Metrik zur Distanzmessung zweier Mengen, hier Punktwolken. Zuerst fällt hier auf, dass die Rekonstruktion von FastNeRF keine verwertbare Punktwolke liefert. Es wurde die Implementierung von Vandegar [6] genutzt, welche im Rahmen der Papers in 100 Lines of Code Reihe entwickelt wurde. Diese Implementierung ist sehr kurz gehalten und wird mit einem vorverarbeitetem Datensatz zu NeRF-Synthetic Lego ausgeliefert. Der Prozess zur Anpassung ist leider auch nicht transparent dokumentiert, wodurch möglicherweise ein Fehler in der Vorbereitung des Datensatzes zu den unverwertbaren Ergebnissen geführt haben könnte. Durch die lange Trainingsdauer des Modells von fast 4,5 Tagen ist die Fehlersuche und Fehlerbehebung nicht im Rahmen der Bachelorarbeit machbar gewesen. In diesem Experiment wird die Farbtreue noch deutlicher, da keine Details wie Logos oder Aufschriften mehr zu erkennen sind. Die quantitative Auswertung bestätigt aber die solide Rekonstruktion, mit einer Abweichung von wenigen Prozenten zur Ground-Truth Punktwolke. Das dritte und letzte Experiment beschäftigt sich mit der Frage, wie gut die Rekonstruktion im Vergleich zu einer gängigen Rekonstruktions-Software ist. Dafür nehmen wir COLMAP, was auf Multi-View-Stereo [7] und Structurefrom-Motion-Verfahren

[8] beruht und ziehen Vergleiche der zwischen Plenoxels und COLMAP auf qualitativer und quantitativer Ebene. Hier bestätigen sich die Auswertungen aus den ersten beiden Experimenten, dass Plenoxels eine gute Rekonstruktion liefert, dabei aber vor allem die schnelle Trainingsdauer ein Vorteil ist. Sobald die Kriterien Farbtreue und sehr hohe Präzision sind, sollte auf andere Rekonstruktionsverfahren, wie eben COLMAP zurückgegriffen werden. Schließlich liefert Plenoxels hauptsächlich in zeitkritischen Anwendungen einen Vorteil und kann so zum Beispiel in der Robotik genutzt werden, um komplexe Objekte zu rekonstruieren, welche von Robotern gehandhabt werden müssen. Für eine Ausweitung auf etwa die Erzeugung von Karten zu Lokalisierung oder Navigation bedarf es noch weiterer Forschung. Auch könnte die Farbtreue möglicherweise durch eine Spezifizierung der Blickrichtung, zum Beispiel anhand von IV Normalen, verbessert werden. Aber auch hier müsste noch weiter experimentiert werden.

Literatur

- [1] S. Fridovich-Keil, A. Yu, M. Tancik, Q. Chen, B. Recht, A. Kanazawa. Plenoxels: Radiance Fields without Neural Networks. In CVPR, 2022.
- [2] S.J. Garbin, M. Kowalski, M. Johnson, J. Shotton, J. Valentin. FastNeRF: High-Fidelity Neural Rendering at 200FPS. arXiv preprint arXiv:2103.10380, 2021.
- [3] B. Mildenhall, P.P. Srinivasan, M. Tancik, J.T. Barron, R. Ramamoorthi, R. Ng. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In ECCV, 2020.
- [4] A. Knapitsch, J. Park, Q.-Y. Zhou, and V. Koltun. Tanks and Temples: Benchmarking Large-Scale Scene Reconstruction. In ACM Transactions on Graphics, volume 36, number 4, 2017.
- [5] L. Liu, J. Gu, K.Z. Lin, T.-S. Chua, C. Theobalt. Neural Sparse Voxel Fields. In NeurIPS, 2020.
- [6] M. Vandegar. Papers in 100 Lines of Code, FastNeRF: High-Fidelity Neural Rendering at 200FPS. GitHub repository, 2023. <https://github.com/MaximeVandegar/Papers-in-100-Lines-of-Code/>
- [7] J.L. Schönberger, E. Zheng, M. Pollefeys, J.-M. Frahm. Pixelwise View Selection for Unstructured Multi-View Stereo. In European Conference on Computer Vision (ECCV), 2016.
- [8] J.L. Schönberger and J.-M. Frahm. Structure-from-Motion Revisited. In Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

Detektion von Kalibrierobjekten in Punktwolken und Kamerabildern

Matthias Deberling, Dennis Griebler, Georg Umlauf

Die Arbeit mit LiDAR-Sensoren ist heutzutage immer weiter verbreitet. Genaue Erkennung und Identifizierung der Umgebung, kleinere Modelle und kostengünstigere Bauweisen machen sie ideal für viele Bereiche bei denen es um differenzierte Betrachtung geht. So ist beispielsweise im Bereich des autonomen Fahrens das Erkennen und Identifizieren der Umgebung eine essenzielle Aufgabe. Unterschiedliche Systeme, wie LiDARs, RADARs, GPS oder optische Kameras werden dafür gemeinsam eingesetzt. Kameras haben den Vorteil, dass sie kostengünstig und hochauflösend sind sowie Farbinformationen erfassen können. Außerdem ist die Bildverarbeitung ein gut erforschtes Gebiet, weshalb viele Algorithmen und Methoden zur Bilderkennung existieren und angewendet werden.

Im Rahmen eines mehrteiligen Projekts soll für den Katamaran ein autonomes Erkennungssystem für Hindernisse entwickelt werden. Dieses soll aus einem LiDAR- und einem Stereokamera-System bestehen. Damit die beiden Systeme zusammenarbeiten, müssen sie aufeinander kalibriert werden.

Für die Kalibrierung werden zwei aufeinander geklebte Pakete verwendet. Es werden mit beiden Systemen jeweils zwei Aufnahmen gemacht, eine mit Paket und eine ohne. Anschließend werden die Stereokameradaten in Punktwolken transformiert und in das Koordinatensystem des LiDAR überführt. Für die Verarbeitung wird die Bibliothek *open3d* verwendet [1]. Sobald die Pakete aus der Szene extrahiert und eventuelle Ausreißer entfernt wurden, können die Seiten des Pakets mittels k-Means Algorithmus [2,3] separiert werden. Danach werden mit Hilfe des RANSAC Algorithmus [4] die Ebenengleichungen der drei Seiten geschätzt. Mit diesen geschätzten Ebenengleichungen können dann insgesamt sieben Schnittpunkte berechnet werden. Jeder Schnittpunkt liegt dabei auf einer Kante des Pakets. Abschließend wird mit den Schnittpunkten und dem Iterative Closest Point Algorithmus [5,6] die Translation und Rotation berechnet, um die gewünschte Konvergenz zu erreichen.

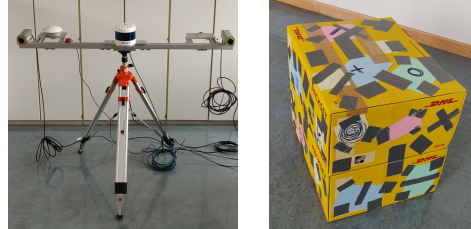


Figure 39: Sensorsystem und Kalibrierobjekt mit Textur.

Die vorgestellte Methode ist eine einfache und schnell durchführbare Möglichkeit für die Kalibrierung eines LiDAR-Sensors und Stereokamera-Systemen. Sie setzt auf bekannte und robuste Algorithmen und liefert gute Ergebnisse. Die Ergebnisse unterliegen jedoch einer gewissen Varianz. Dieser kann entgegengewirkt werden, indem der RANSAC für jede Seite häufiger durchgeführt wird und dann der Durchschnitt der Ebenengleichungen gebildet wird. Auch hat sich gezeigt, dass bessere Ergebnisse durch die Verwendung von mehreren Positionen bei gleicher Entfernung erzielt werden.

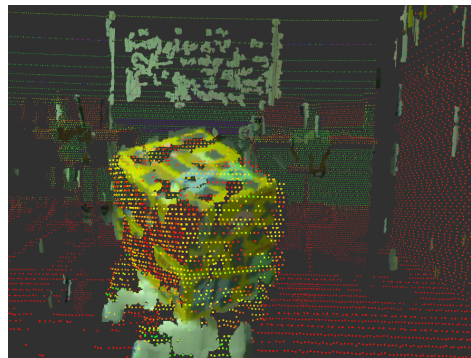


Figure 40: Registrierte Punktwolken.

Literatur

- [1] Open3D, <http://www.open3d.org/>.
- [2] Stuart LLOYD. Least squares quantization in

- pcm. IEEE Transactions on Information Theory, 28(2):129-137, 1982.
- [3] Edward Forgy. Cluster analysis of multivariate data: Efficiency versus interpretability of classification. Biometrics, 21(3):768-769, 1965.
- [4] Martin A. Fischler, Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM, 24(6):381-395, 1981.
- [5] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. Proceedings. 1991 IEEE International Conference on Robotics and Automation, 1991.
- [6] P.J. Besl and Neil D. McKay. A method for registration of 3-d shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2):239-256, 1992.

Registrierung von Fuß-Scans mit Raumdeformationsverfahren

Niclas Graf, Dennis Griebner, Georg Umlauf

Die neue Generation von Smartphones ermöglicht es durch LIDAR-Scanner, 3D-Scans einem weiten Personenkreis zugänglich zu machen. Dadurch können 3D-Scans dann auch ohne spezialisierte Hardware erstellt werden. Dies liefert neue Anwendungsmöglichkeiten für das Smartphone, sowohl in der Vermessung, als auch in der Visualisierung von Objekten. Da ein einziger Blickwinkel für die Erfassung eines Objekts meist nicht ausreicht, müssen auch Problemstellungen behandelt werden, die sich mit dem Zusammenfügen mehrerer Scans befassen. So gibt es auch die Idee, mithilfe seines Smartphones, 3D-Modelle von Füßen zu erstellen. Dies soll ermöglichen, anhand der Form des Fußes, einen passenden Schuh zu finden. Dafür muss der Fuß aus mehreren Perspektiven gescannt werden, um alle Details des Fußes erfassen zu können. Um nun den ganzen Fuß zu erhalten, müssen diese einzelnen Scans sinnvoll zusammengesetzt werden. Bei bisherigen ICP basierten Versuchen kam es dabei aber zu einer Lücke an der Ferse des Fußes. Diese Arbeit soll hierbei eine Grundlage liefern, um diese Lücke zu schließen.

Um die ICP-Registrierung zu verbessern, soll ein Modell gefunden werden, an welchem die Registrierung durchgeführt werden kann. Dieses Modell sollte möglichst große Ähnlichkeiten zum tatsächlichen Objekt haben. Außerdem sollte es keine Lücke an der Ferse haben. Nur dann könnte die Lücke bei einer erneuten ICP-Registrierung geschlossen werden. Um dieses Modell zu erhalten, wurde ein Space Deformation-Ansatz gewählt. Es soll dabei eine Raumdeformation gefunden werden, die einen Modellfuß möglichst gut auf ein Zielobjekt abbildet. Das Zielobjekt wird dabei das Resultat der 1. ICP-Registrierung sein, das zwar eine Lücke an der Ferse aufweist, aber ansonsten den Fuß gut abbildet. Als Modell-

fuß kann eine beliebige Punktwolke eines Fußes gewählt werden. Als Grundlage für den gewählten Ansatz wurde das Verfahren von M. Rouhani [1] ausgewählt. Dieses nutzt die Raumdeformation, um das Modell iterativ an das Zielobjekt anzupassen. Um sinnvolle Parameter für die Raumdeformation zu finden, werden dabei Oberflächeninformationen, wie die Krümmung und die Normale, verwendet.

Der gewählte Ansatz passt das Modell gut an das Zielobjekt an. Es passt sich jedoch auch an die Verschiebungen an der Ferse des Zielobjekts an. Deshalb bildet sich an der Ferse des Modells ein Wulst. Dies zeigt, dass eine nicht starre Registrierung mithilfe der Raumdeformierung möglich ist. Weitere Versuche müssen zeigen, ob dieses Modell die Anforderungen für eine genauere 2. ICP-Registrierung erfüllen kann.

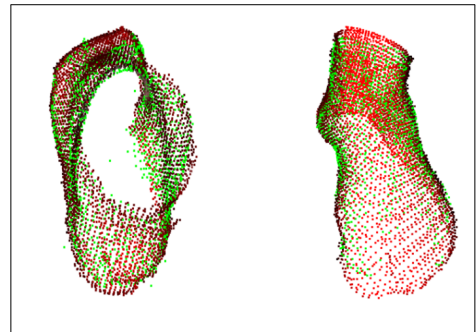


Figure 41: Registrierungsergebnis: grün der Modellfuß, rot die Zieldaten.

Literatur

[1] Rouhani, M., and Sappa, A. Non-rigid shape registration: A single linear least squares framework, 2012.

Bildbasierte Geolokation für Augmented Reality in Gebirgen

Frieder Luginsland, Dennis Griebler, Georg Umlauf

Die Wegfindung und Routenplanung in den Bergen, insbesondere beim alpinen Mehrseillängenklettern, stellen eine große Herausforderung dar. Kletterer greifen üblicherweise auf Kartenmaterial und grafische Darstellungen von Kletterrouten (Topografien) zurück. Es ergeben sich bei der Verwendung von Bild-Topografien Probleme, denn meist nimmt der Betrachter eine andere Perspektive zum Berg ein, als die Topografie. Die perspektivische Verzerrung erschwert die Lokalisierung der Kletterroute. Ziel dieser Arbeit besteht darin, Kletterrouten perspektivisch korrekt auf Smartphone-Aufnahmen zu projizieren. Dabei umfasst der Inhalt sowohl die Generierung dreidimensionaler Kletterrouten, als auch deren Projektion in die zweidimensionale Bildebene. Ein präziser Standort und die Ausrichtung der Kamera sind dabei grundlegend für die Anwendung. Die Problemstellung fällt unter den Begriff der Geolokalisierung von Bilddateien, einem Forschungsgebiet, in welchem bereits viele Veröffentlichungen publiziert wurden. Grundsätzlich handelt es sich um ein Optimierungsproblem mit sechs Freiheitsgraden. Drei für die Position (x, y, z) und drei für die Ausrichtung (Yaw, Pitch, Roll). Der vorgestellte Algorithmus für die Geolokation von Bilddateien setzt sich dabei aus mehreren Schritten zusammen.

Zuerst werden aus den aufgenommenen Bildern der Horizont in Bildkoordinaten, durch Anwendung von Algorithmen aus der Graphentheorie, extrahiert. Aus einem Datensatz, der die Oberfläche der Schweiz abbildet, werden digitalen Horizonte in Abhängigkeit des Projektionszentrums gerendert. Durch die Projektion auf eine Kugel kann eine Rundumsicht des gesamten Horizontes erreicht werden. Für einen gegebenen Standort und eine initiale Schätzung der Ausrichtung, bestimmt durch den Lagesensor des Smartphones, kann der digitale Horizont auf der Kugeloberfläche in Bildkoordinaten umgerechnet werden. Es kann so nach den optimalen Eulerwinkel gesucht werden die analogen und digitalen Horizont bestmöglich registrieren. Dabei wird eine adaptierte Erweiterung des gängigen

ICP Algorithmus verwendet. Die 3D Rotation der Kamera verursacht eine nicht-lineare Transformation der Bildpunkte. Die grundlegende Idee besteht darin, die gefundene euklidische Transformation als 3D-Rotation zu approximieren. Dafür muss die optimale Verschiebung auf der i -Achse in eine Yaw-Winkeländerung umgerechnet werden. Für Verschiebungen auf der j -Achse wird eine Winkeländerung des Pitch-Winkels gesucht. Eine 3D-Rotation um den Roll-Winkel kann als 2D-Rotation beschrieben werden. Der vorgeschlagene Algorithmus ermöglicht es, eine 3D-Rotation robust und mit minimalem numerischen Aufwand zu bestimmen. Durch Variation des Standortes, welcher sich in der Umgebung der geschätzten Lage befindet, kann eine Position gefunden werden, die eine bessere Übereinstimmung der Horizonte erzeugt. Bei der Variation der Position muss die Optimierung der Eulerwinkel jedes Mal durchgeführt werden. Durch die Suche des besten Projektionszentrums kann der Fehler global gesehen minimiert werden. Für die Optimierung wird das Nelder-Mead-Verfahren sowie eine Bayes'sche Optimierung getestet. Es zeigt sich das die Bayes'sche Optimierung besser für die verrauschten Zielfunktion geeignet ist.

Nachdem die Position und Ausrichtung der Kamera gefunden ist können Bildpunkte in Weltpunkte und umgekehrt transformieren werden. Die Ausgangslage für die Rückprojektion besteht aus einem georeferenzierten Bild, den Bildpunkten und einem Mesh, das den abgebildeten Berg approximiert. Für jeden Bildpunkt wird ein Strahl konstruiert, der ausgehend vom Projektionszentrum durch den Bildpunkt verläuft. Falls ein Linsenmodell bekannt ist, können die Bildpunkte vorab so transformiert werden, dass die Annahme einer Lochkamera zutrifft. Für jeden Strahl kann der Schnittpunkt mit dem Mesh, das die Erdoberfläche approximiert, bestimmt werden. Der Schnittpunkt gibt die Lage des zugehörigen Weltpunktes für einen Bildpunkt an.

Für die Umwandlung von Weltpunkten in Bildpunkte kann die klassische Projektion verwendet werden. Dabei wird mithilfe der Projektion-

smatrix direkt der Schnittpunkt des Weltpunktes mit der Bildebene bestimmt. Dabei wird der Weltpunkt unter Berücksichtigung der Geometrie einer Lochkamera auf die Bildebene projiziert. Falls ein Linsenmodell bekannt ist, können die Bildpunkte verzerrt werden, um die Einflüsse der Linse zu approximieren. In der Grafik 42 wird die Rückprojektion und Projektion visualisiert.

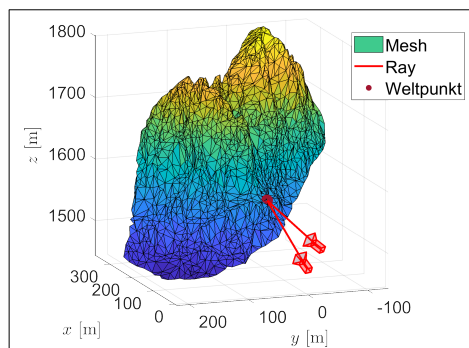


Figure 42: Projektion und Rückprojektion von Weltpunkten bzw. Bildpunkten.

Der Routenersteller benötigt eine Aufnahme der Felswand, auf der die gesamte Kletterroute sichtbar ist. Diese Aufnahme wird mithilfe der zuvor beschriebenen Methode georeferenziert. In dieser Aufnahme wird die Kletterroute manuell eingezeichnet. Zusätzlich können beliebige weitere Informationen, wie zum Beispiel Stände für Mehrseillängen oder Abseilpisten, eingezeichnet werden. Für die Rückprojektion wird ein Mesh der aufgenommenen Felswand benötigt. Ausgehend von einer Punktwolke werden alle Punkte innerhalb eines Polygons entnommen, welches die Felswand repräsentiert. Der Benutzer nimmt ein Bild von einer Felswand auf, auf welcher der Verlauf der Kletterroute vermutet wird. Das Bild wird wie beschrieben geolokalisiert. Anschließend werden die abgespeicherten Weltpunkte in die Aufnahme projiziert. Der Nutzer benötigt eine kalibrierte Kamera.



Figure 43: Projektion und Rückprojektion von 3D Kletterroute anhand von Beispiel.

In der oberen Spalte der Abbildung 43 ist die Aufnahmen des Routenerstellers zu sehen, auf denen die Kletterrouten und Ständen eingezeichnet und in Weltpunkte überführt werden. In der unteren Spalte sind weitere Aufnahmen derselben Felswände abgebildet, die von einem Applikationsnutzer aufgenommen wurden. Die Weltpunkte wurden perspektivisch korrekt in die Aufnahme eingefügt.

Die Validierung der Geolokation ist aufgrund fehlender Referenzwerte nicht trivial. Die gegebenen Sensorwerte für Lage und Ausrichtung der Kamera dienen als Initialwerte für den Algorithmus und sollen verbessert werden. Es existiert kein Datensatz mit genaueren Referenzwerten, da die Position üblicherweise mit GPS-Messungen bestimmt wird. Um die Güte der Geolokation abzuschätzen, werden daher mehrere eigene Fehlermaße eingeführt. Anhand einer Fallstudie und der eingeführten Metriken wird die Geolokation validiert. Es konnte gezeigt werden, dass die erreichte Genauigkeit der Geolokation für die meisten Kletter-Szenarien ausreichend ist.

Optisches Multisensorsystem für autonome Anwendungen auf Binnengewässern⁷

Dennis Griebler, Tim Baur, Matthias Albrecht, Johannes Reuter, Matthias Franz, Georg Umlauf

Im Rahmen der Ausschreibung Optische 3D-Sensorsysteme für mobile Anwendungen der Baden-Württemberg Stiftung gGmbH forschten das Institut für Optische Systeme (IOS) und das Institut für Systemdynamik (ISD) der HTWG Konstanz an einem optischen Multisensorsystem, basierend auf Kameras und Lidar, zur Umfelderkennung auf Binnengewässern. Während der Projektlaufzeit entwickelte das Projektteam Detektions- und Trackingalgorithmen, mit denen ein autonomes Anlegen auf dem Forschungsboot der Hochschule realisiert wurde.



Figure 44: Optisches Sensorsystem montiert auf dem Forschungsboot *Solgenia*.

Damit die Detektions- und Trackingalgorithmen entwickelt werden konnte, wurde ein Multisensorsystem, bestehend aus einem rotierenden 360° Lidar, einem Stereosystem, einer inertialen Messeinheit (IMU) und einem globalen Positionierungssystem (GPS), entwickelt. Mithilfe des Sensorsystems konnte dann die Datenakquise auf zwei Versuchsträgern durchgeführt werden. Zum einen konnten Daten mit dem Forschungsboot *Solgenia* (siehe Abbildung 44) der HTWG Konstanz und zum anderen mit dem Katamaran *Constanze* der Bodensee-Schiffsbetriebe GmbH aufgezeichnet werden. Dabei wurden verschiedene Szenarien bei unterschiedlichen Witterungsverhältnissen aufgezeichnet. Um die Datenannotation und Datenfusion der einzelnen Sen-

soren einfacher zu gestalten, wurden die Kameras geometrisch kalibriert sowie das Stereosystem und der Lidar-Sensor zueinander registriert.

Da es aufgrund von Wellenschlag zu einer starken Eigenbewegung des Versuchsträgers kommen kann, muss die Position und Orientierung des Sensorsystems in Bezug zur Wasseroberfläche zu jedem Zeitschritt bestimmt werden. Deshalb wurde im Rahmen des Projekts eine optische Lagebestimmung entwickelt. Zur visuellen Lagebestimmung wurde zuerst ein neuronales Netzwerk eingesetzt, das in der Lage ist, die Wasseroberfläche innerhalb eines RGB-Bildes zu segmentieren. Anschließend wurde die dreidimensionale Wasseroberfläche mithilfe des Stereosystems rekonstruiert. Durch die Approximation der Wasseroberfläche als Ebene konnte dann die Höhe der Kamera sowie die Nick- und Rollbewegung der Kamera in Bezug zur Wasseroberfläche bestimmt werden. Zusätzlich zur Lage sollen Wasserfahrzeuge, die mögliche Kollisionspartner darstellen, auf der Wasseroberfläche lokalisiert und klassifiziert werden. Hierzu wurde ein neuronales Netzwerk entwickelt, das Stereo-Bildpaare aufnimmt und daraus dreidimensionale Detektionen in Form von Bounding-Boxen berechnet. Damit das neuronale Netzwerk trainiert werden konnte, wurde ein Datensatz für Wasserfahrzeuge auf Binnengewässern entwickelt. Dabei wurde die Position und Ausdehnung der Objekte mit Bounding-Boxen markiert. Jedem Objekt wurde zusätzlich noch eine Kategorie zugeordnet, sodass klar war, um was für ein Wasserfahrzeug es sich handelt. Der Datensatz umfasst insgesamt 1167 annotierte Stereobildpaare und Lidar-Punktwolken mit den Kategorien: Motorboot, Segelboot mit und ohne Segel, Stand-Up Paddle, Katamaran, Fähre, Tretboot, Personenschiff und Dalbe. Mit dem annotierten Datensatz wurde anschließend das neuronale Netzwerk trainiert, validiert und getestet. Qualitative Ergebnisse des Lageerkennungssystems in Kombination

⁷Text veröffentlicht über Photonics BW.

mit dem Detektionsnetzwerk sind in Abbildung 45 dargestellt.



Figure 45: Detektierte Bounding-Boxen und Objektklassen (rot) sowie die zugehörigen Annotationen (grün).

Um die dynamischen Objekte, welche durch das Stereosystem detektiert werden, über die Zeit verfolgen zu können wurden nichtlineare Varianten des Kalman Filters eingesetzt. Innerhalb dieser Filter müssen Bewegungsmodelle des zu verfolgenden Objekts und Messmodelle des Sensors angegeben werden.

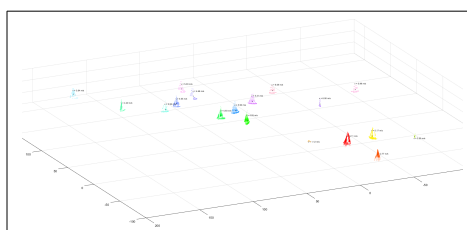


Figure 46: Trackingergebnis eines Regatta-Szenarios.

Demzufolge wurden zu Beginn des Projekts entsprechende Modelle entwickelt und auf Realdaten ausgewertet. Da der Lidar-Sensor eine Vielzahl an Messungen pro Objekt und Zeitschritt liefert, wurden im Bereich der Messmodelle spezielle Verfahren entwickelt, welche zusätzlich zum kinematischen Zustand eine Filterung der Ausdehnung des Objekts ermöglichen. In einem weiteren Schritt wurden anschließend Algorithmen für die Verfolgung mehrerer Objekte entwickelt und ebenfalls anhand von Realdaten evaluiert. Aufgrund der Echtzeitproblematik wurde dabei ein vereinfachtes Verfahren zur Datenassoziation von Messung zu Objekt eingesetzt. Außerdem wurden die Objekte innerhalb dieses Algorithmus als mathematische Punkte ohne Ausdehnung modelliert. Das Trackingergebnis eines Regatta-Szenarios, bei dem der Katamaran zwischen den Segelbooten hindurchfahren musste, ist in Abbildung 46 zu sehen. Dabei sind die Messungen

eines jeden Segelbootes segmentiert und in unterschiedlichen Farben dargestellt. Ein entsprechender Segmentierungsalgorithmus wurde ebenfalls innerhalb des Projekts entwickelt und untersucht. Weiterhin ist zu erkennen, dass jedes Segelboot einzeln verfolgt und die entsprechende Geschwindigkeit geschätzt werden kann. In einem weiteren Schritt wurde der Trackingalgorithmus mit den Ergebnissen des Detektionsnetzwerks kombiniert. Dabei wurden die Detektionen und die entsprechenden Bounding-Boxen genutzt, um die Messungen von dynamischen Objekten aus der Lidar-Punktwolke zu filtern, um dadurch lediglich, für die Kollisionsvermeidung entscheidende, dynamische Objekte verfolgen zu können. Insbesondere im Uferbereich bietet dies einen entscheidenden Vorteil, da dort viele Messungen von statischen Objekten aufgezeichnet werden.

In einem finalen Arbeitsschritt wurden die erzielten Ergebnisse für die Entwicklung eines Demonstrators zur automatisierten Anlegestellensuche und zur Automatisierung des Anlegevorgangs zusammengeführt. Dabei wurde die Anlegestelle innerhalb der Lidar-Punktwolke mithilfe eines neuronalen Netzwerks automatisiert detektiert und eine Bounding-Box um den Anlegesteg berechnet (siehe Abbildung 47). Diese Bounding-Box wurde dann für die Bestimmung der Anlegestelle genutzt.

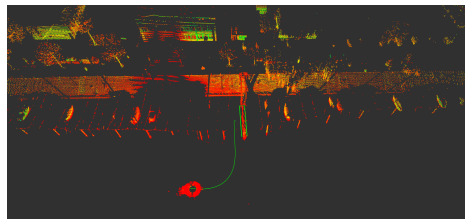


Figure 47: Darstellung des automatisierten Anlegevorgangs innerhalb der Lidar-Punktwolke. Berechnete Trajektorie als grüne Linie, detektierte Anlegestelle als grüne Box.

Mit bereits bestehenden Algorithmen zur Regelung der Pose des Forschungsboots *Solgenia*, wurde dieses dann entlang einer berechneten Trajektorie automatisiert an die Anlegestelle manövriert. Die Pose des Forschungsboots für die Rückführung innerhalb der Regelung wurde dabei über einen SLAM-Algorithmus mit den Lidar-Messungen bestimmt.

Institut für optische Systeme
Hochschule Konstanz
Alfred-Wachtel-Straße 8
D-78462 Konstanz
www.ios.htwg-konstanz.de