



UNIVERSITÄT ZU LÜBECK
INSTITUTE FOR SOFTWARE ENGINEERING
AND PROGRAMMING LANGUAGES

isp

Fußgängererkennung beim autonomen Fahren

Wie ich dahin kam und was ich dort forsche

Dr. Gesina Schwalbe

Institut für Softwaretechnik und Programmiersprachen, Universität zu Lübeck

25.04.2024

Girls' Day Uni Bamberg

Gliederung

Wer bin ich

Wo arbeite ich jetzt

Was mache ich aktuell – am Beispiel Fußgängererkennung

Was gehört dazu

Sicherheit: DNNs erklären und verifizieren

Zusammenfassung

Gliederung

Wer bin ich

Wo arbeite ich jetzt

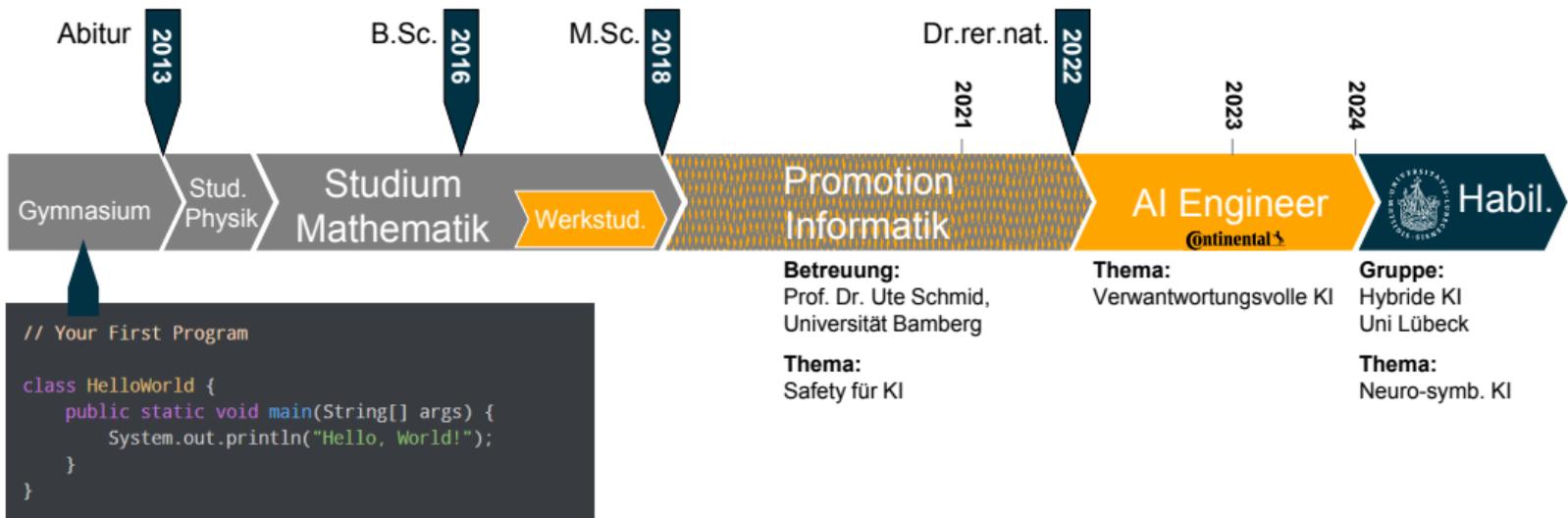
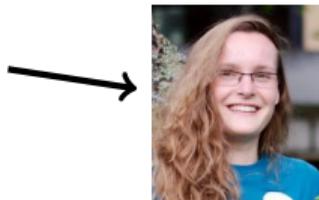
Was mache ich aktuell – am Beispiel Fußgängererkennung

Was gehört dazu

Sicherheit: DNNs erklären und verifizieren

Zusammenfassung

Wer bin ich



Gliederung

Wer bin ich

Wo arbeite ich jetzt

Was mache ich aktuell – am Beispiel Fußgängererkennung

Was gehört dazu

Sicherheit: DNNs erklären und verifizieren

Zusammenfassung

Als Postdoc an der Universität zu Lübeck



UNIVERSITÄT ZU LÜBECK



Im Focus das Leben.
An der Schnittstelle von

- ▶ Medizin
- ▶ Naturwissenschaft
- ▶ Technik

Gliederung

Wer bin ich

Wo arbeite ich jetzt

Was mache ich aktuell – am Beispiel Fußgängererkennung

Was gehört dazu

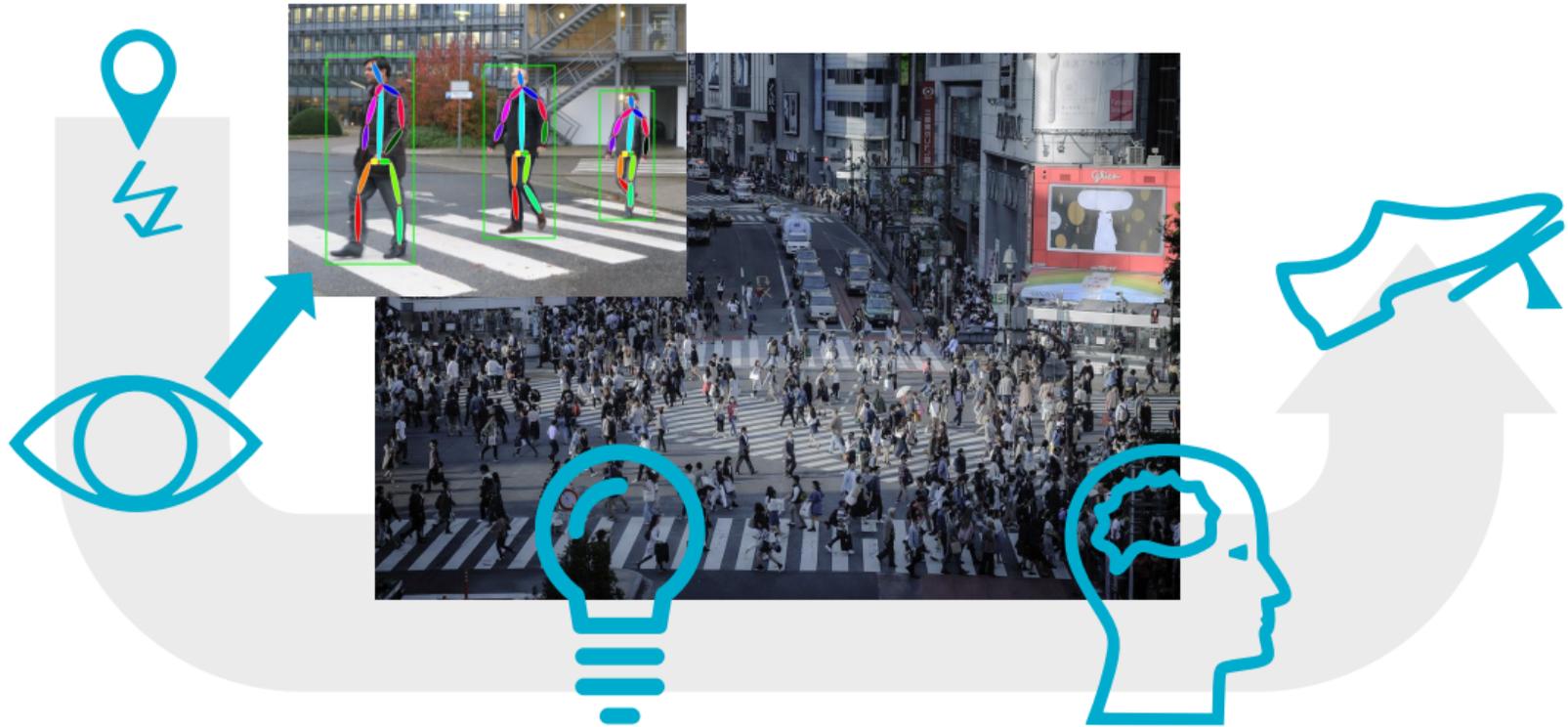
Sicherheit: DNNs erklären und verifizieren

Zusammenfassung

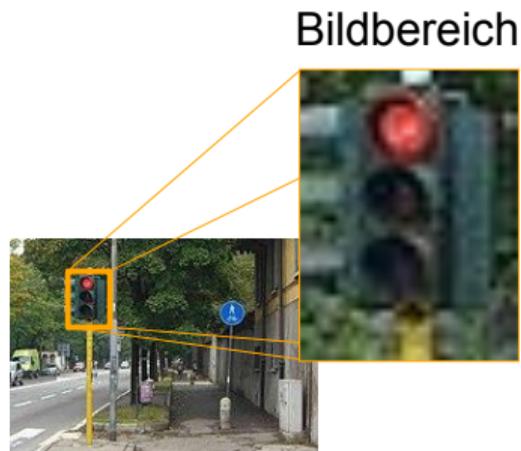
Automatisiertes Fahren: Ein paar Problemstellungen



Automatisiertes Fahren: Ein paar Problemstellungen



Neuronale Netze für Objekterkennung

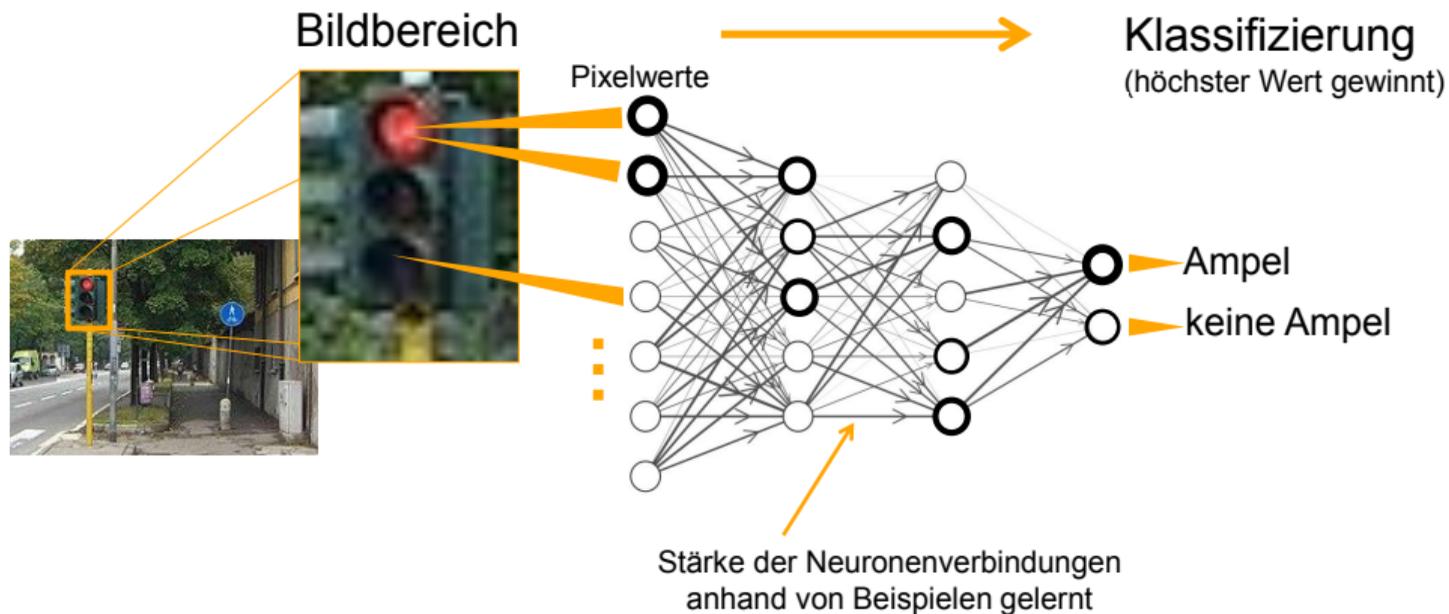


Klassifizierung
(höchster Wert gewinnt)

Ampel

keine Ampel

Neuronale Netze für Objekterkennung



Überwachtes Maschinelles Lernen: Gradientenabstieg

Vorraussetzungen:

- ▶ Funktion f mit zu optimierenden **Parameter**(n)
- ▶ **Differenzierbarer Loss**: Für jede Funktionsausgabe ist bekannt, in welche Richtung Parameterwerte verändert werden müssen für „besseres“ Resultat
- ▶ **Beispiele**

Finde das Tal ...



Überwachtes Maschinelles Lernen: Gradientenabstieg

Vorraussetzungen:

- ▶ Funktion f mit zu optimierenden **Parameter**(n)
- ▶ **Differenzierbarer Loss**: Für jede Funktionsausgabe ist bekannt, in welche Richtung Parameterwerte verändert werden müssen für „besseres“ Resultat
- ▶ **Beispiele**

Finde das Tal ...



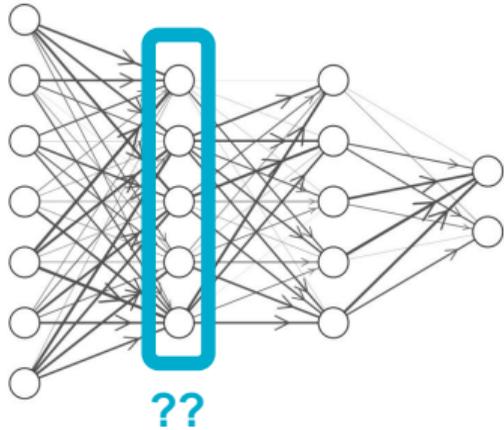
Wiederhole

Für jedes Beispiel

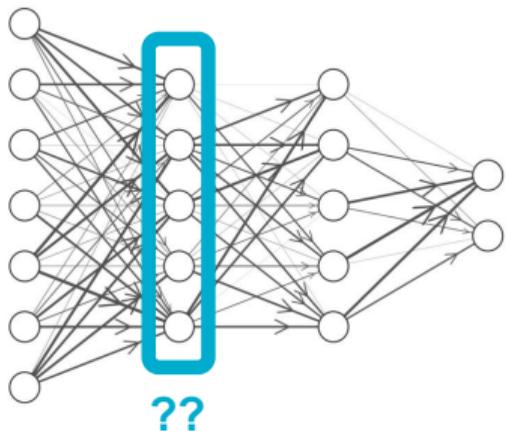
1. Werte Funktion aus
2. Finde beste Änderungsrichtung
3. Ändere Parameter (ein wenig) in diese Richtung

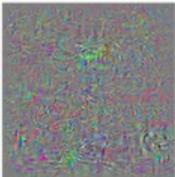
bis Funktion gut genug

Mein Thema: Was hat das DNN gelernt?



Mein Thema: Was hat das DNN gelernt?

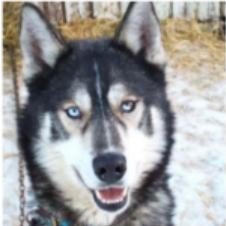


Eingabe:  +  = 

Ausgabe: "bus" + "ostrich" ???

Eingabe: 

Ausgabe: "30km/h" ?!!!

Eingabe: 

Ausgabe: "Wolf" ??

Relevante Bildteile für Entscheidung: 

????!!!

Was hat das DNN gelernt? – Feature Visualization

Frage: Was kodiert ein ausgewähltes Neuron im DNN?

Verfahren zur Visualisierung:

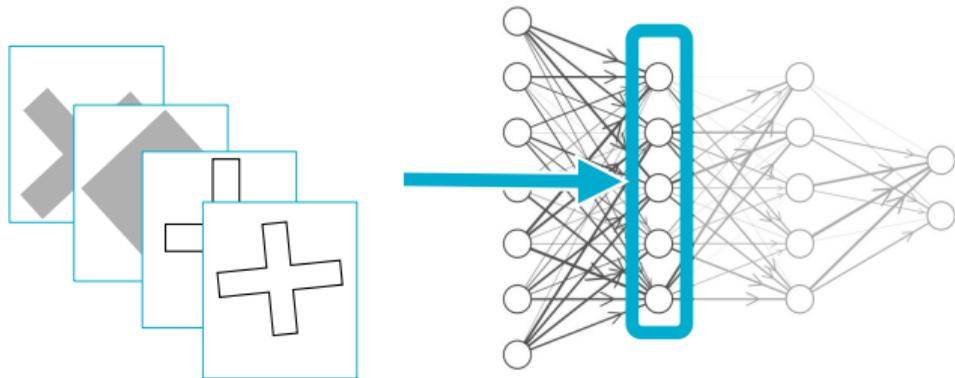
Beispiele

bei denen das Neuron
stark aktiviert



DeepDream
Prototypen

Was hat das DNN gelernt? – Wurden Konzepte erlernt?



Können in der Zwischenausgaben Eingaben mit und ohne dem Konzept (hier: Füllung) unterschieden werden?

Wenn ja, wie?

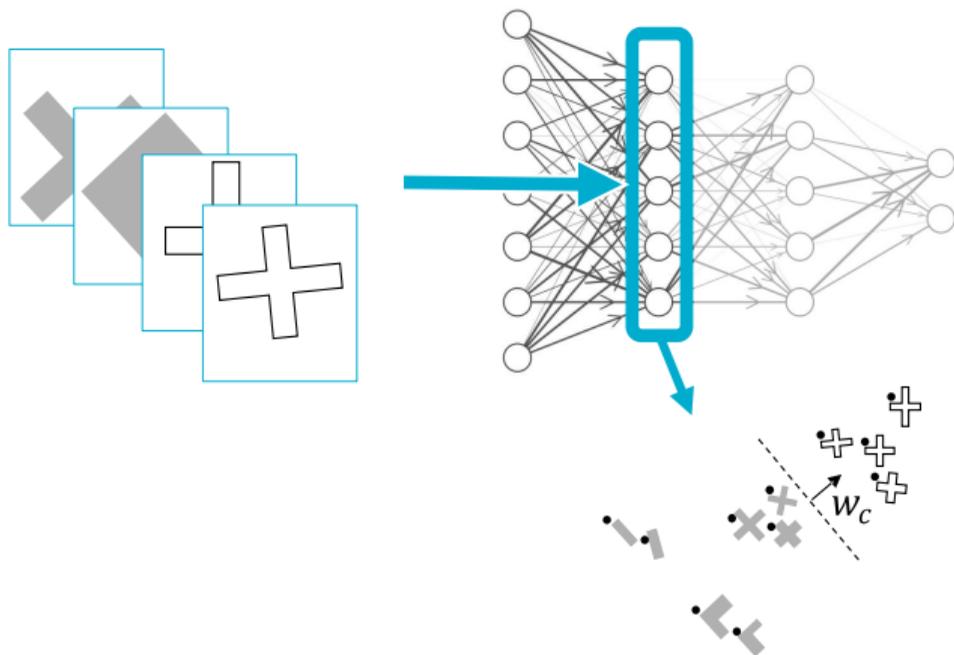
Beispiele für andere Konzepte

Objekt(teile)



Texturen

Was hat das DNN gelernt? – Wurden Konzepte erlernt?



Können in der Zwischenausgaben Eingaben mit und ohne dem Konzept (hier: Füllung) unterschieden werden?

Wenn ja, wie?

Beispiele für andere Konzepte

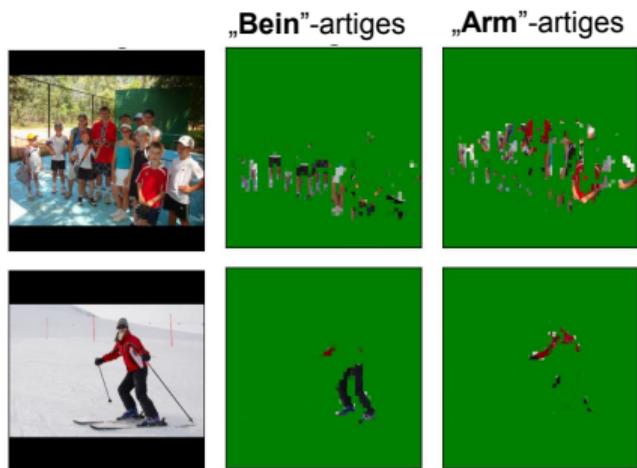
Objekt(teile)



Texturen

Verwendung gefundener Konzepte für Verifikation

Wurde das Konzept erkannt?



Werden Konzepte richtig verwendet?



Gliederung

Wer bin ich

Wo arbeite ich jetzt

Was mache ich aktuell – am Beispiel Fußgängererkennung

Was gehört dazu

Sicherheit: DNNs erklären und verifizieren

Zusammenfassung

Zusammenfassung

- ▶ Zu mir:
 - ▶ *Was mich zur Informatik gebracht hat:* Neugier und Praxisrelevanz
 - ▶ *Was mich daran fasziniert:* Breite der Themen und Anwendungen

- ▶ Zu automatisiertem Fahren:
 - ▶ Besteht aus vielen Teilproblemen
 - ▶ Umgebungswahrnehmung kann mithilfe von DNNs und ML gelöst werden;
DNNs können aber fehlerhaft sein!

⇒ **Meine Forschung:** Einblicke in erlerntes Wissen im DNN bekommen

Zusammenfassung

- ▶ Zu mir:
 - ▶ *Was mich zur Informatik gebracht hat:* Neugier und Praxisrelevanz
 - ▶ *Was mich daran fasziniert:* Breite der Themen und Anwendungen

- ▶ Zu automatisiertem Fahren:
 - ▶ Besteht aus vielen Teilproblemen
 - ▶ Umgebungswahrnehmung kann mithilfe von DNNs und ML gelöst werden;
DNNs können aber fehlerhaft sein!

⇒ **Meine Forschung:** Einblicke in erlerntes Wissen im DNN bekommen

Fragen? 😊

Vielen Dank fürs Zuhören :-)

`gesina.schwalbe@uni-luebeck.de`

`https://gesina.github.io`