

ZAFH Projekt:

Mikro Sens

In Kooperation mit:



ulm university universität
uulm

Hochschule Ulm



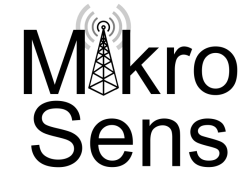
HOCHSCHULE PFORZHEIM 

› Deep Learning zur Erkennung von Kochprozessen

M.Eng. Marco Altmann, Hochschule Heilbronn

MATLAB Expo, München, 21.6.2018

ZAFH Projekt MikroSens finanziert von
EU, EFRE und Baden-Württemberg

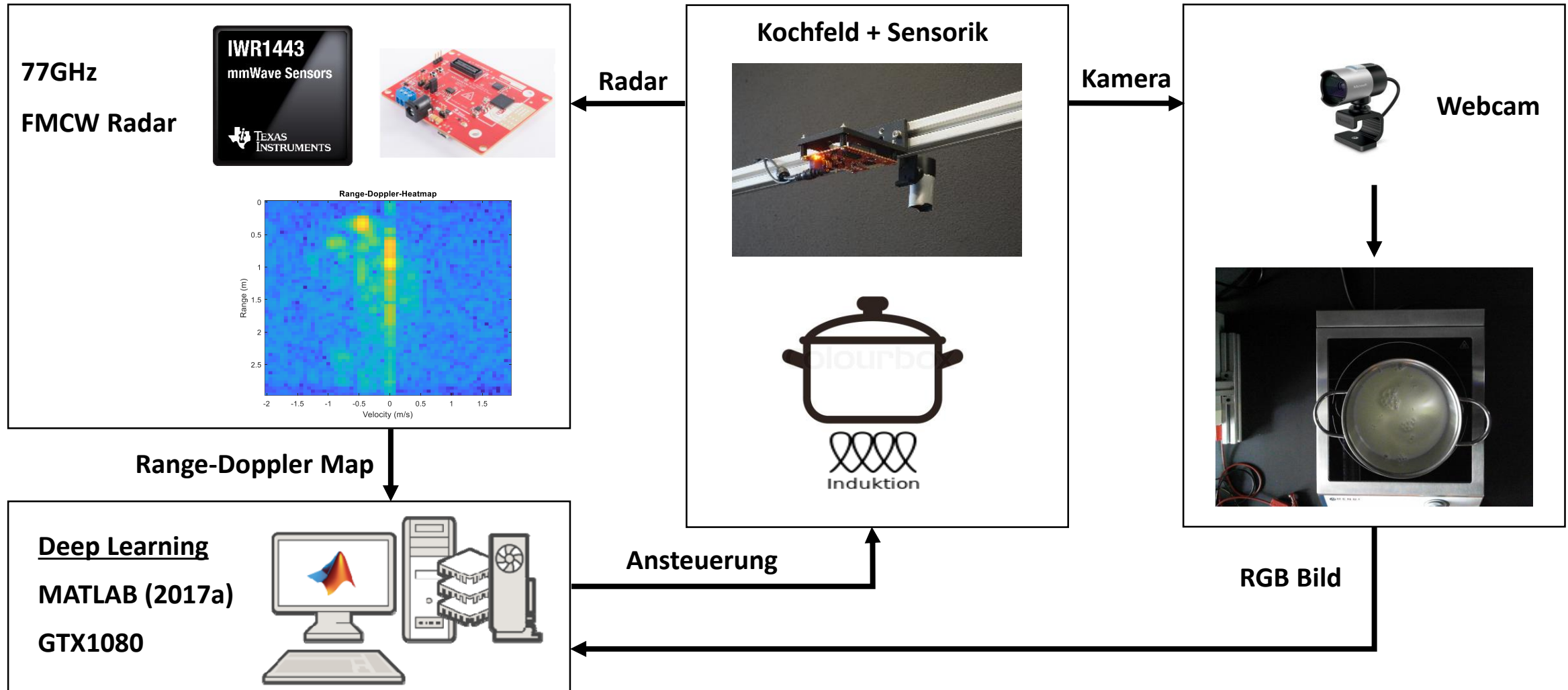


› Deep Learning zur Erkennung von Kochprozessen

M.Eng. Marco Altmann, Hochschule Heilbronn

MATLAB Expo, München, 21.6.2018

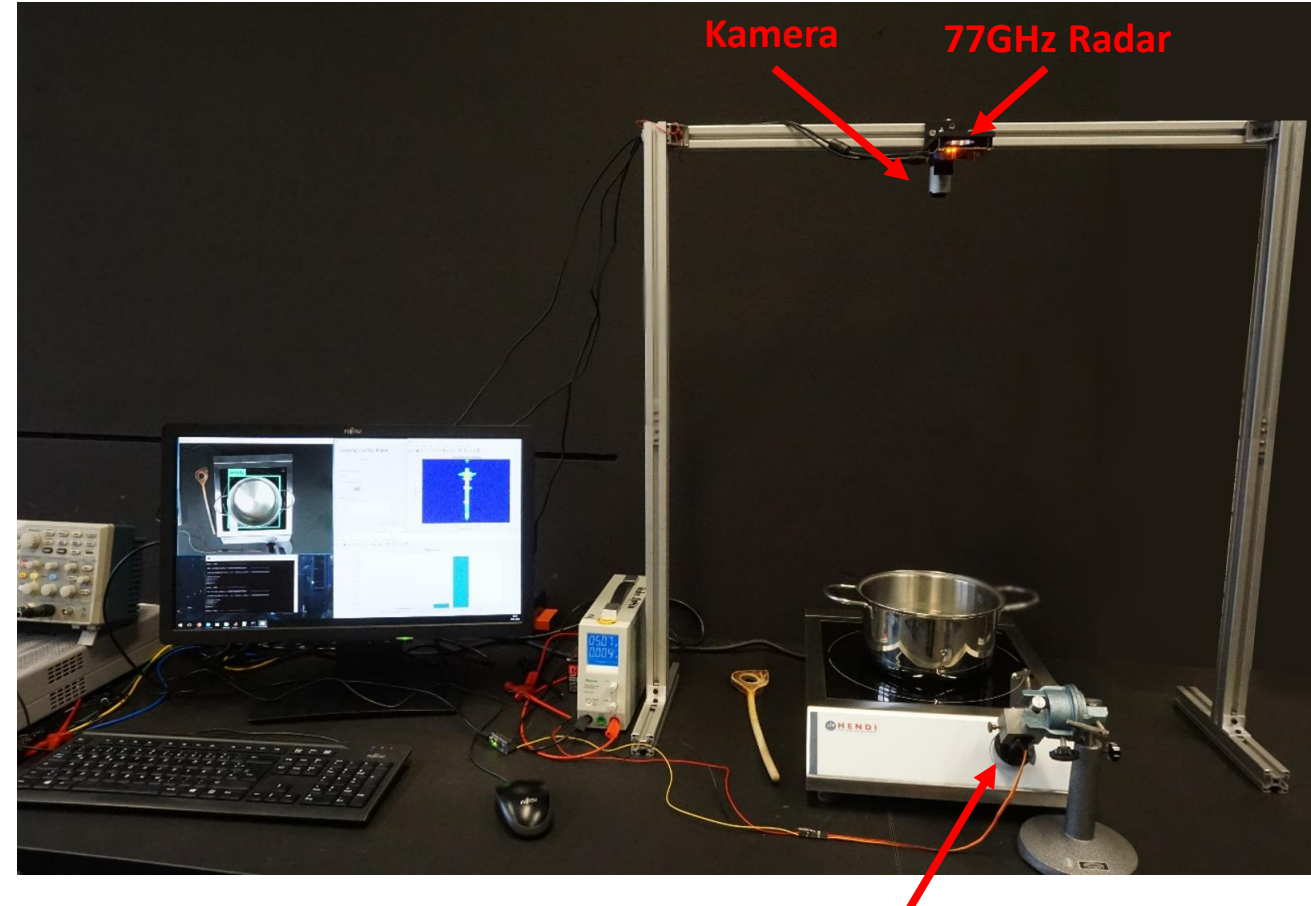
Übersicht Kochsensor



Übersicht Kochsensor

Laboraufbau

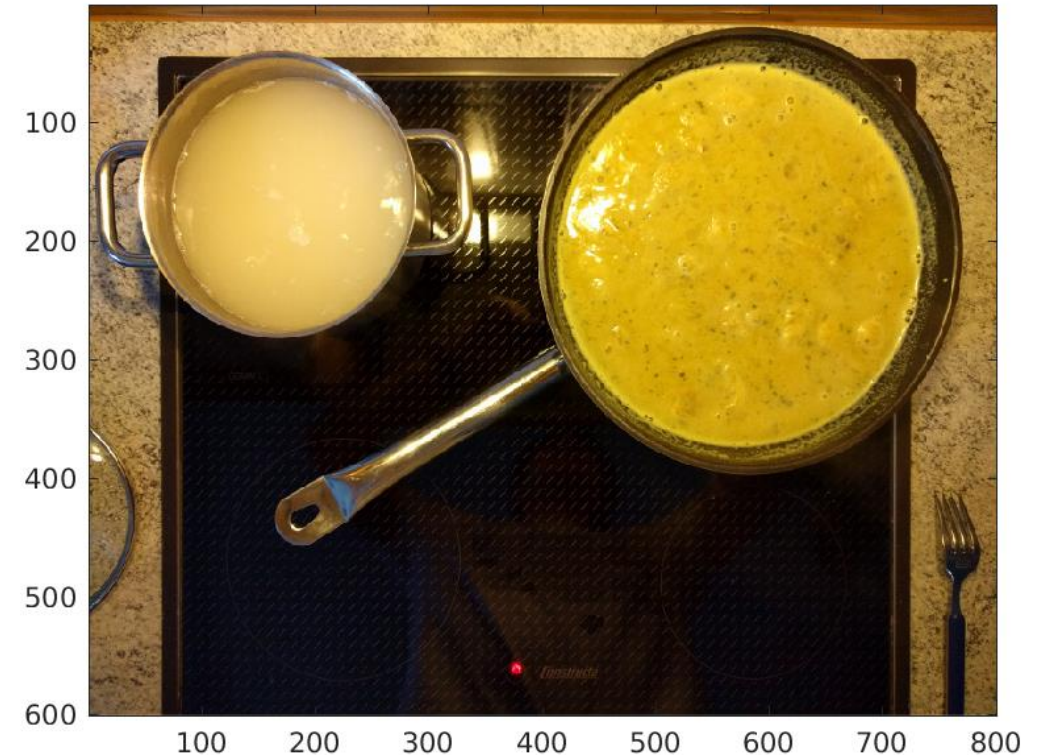
- > 77GHz FMCW-Radar (misst Abstände/Füllstände & Geschwindigkeiten/Bewegungen)
- > Kamera (Detektion der Kochstelle & Klassifikation des Kochguts)
- > Induktionsplatte
- > Ansteuerung der Herdplatte
- > MATLAB 2017a
- > GTX1080, CUDA8.0, Caffe



Servo-Motor zur Ansteuerung

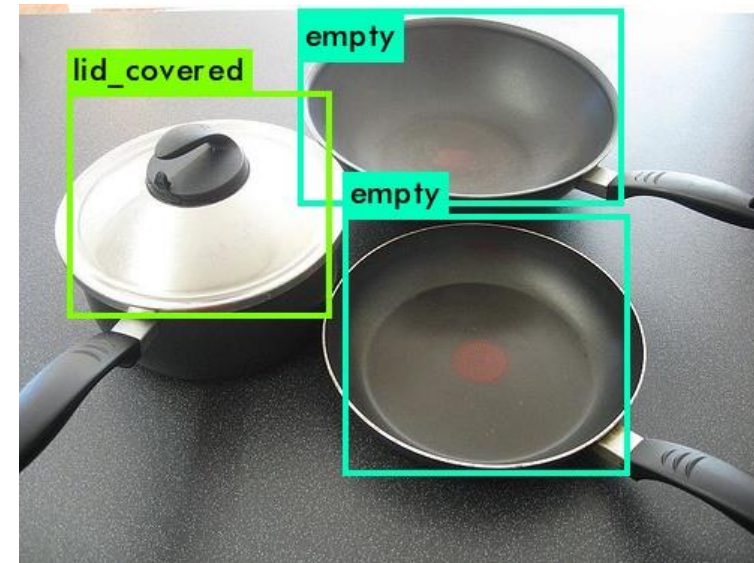
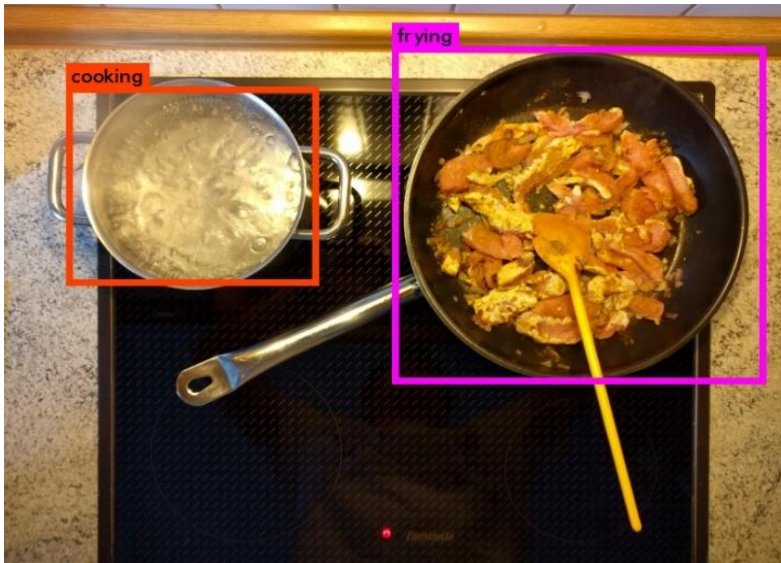
Kamera

- > Welche Kochstelle wird verwendet?
(Detektion)
- > Was wird gekocht?
(Klassifikation)



Kamera - Detektion

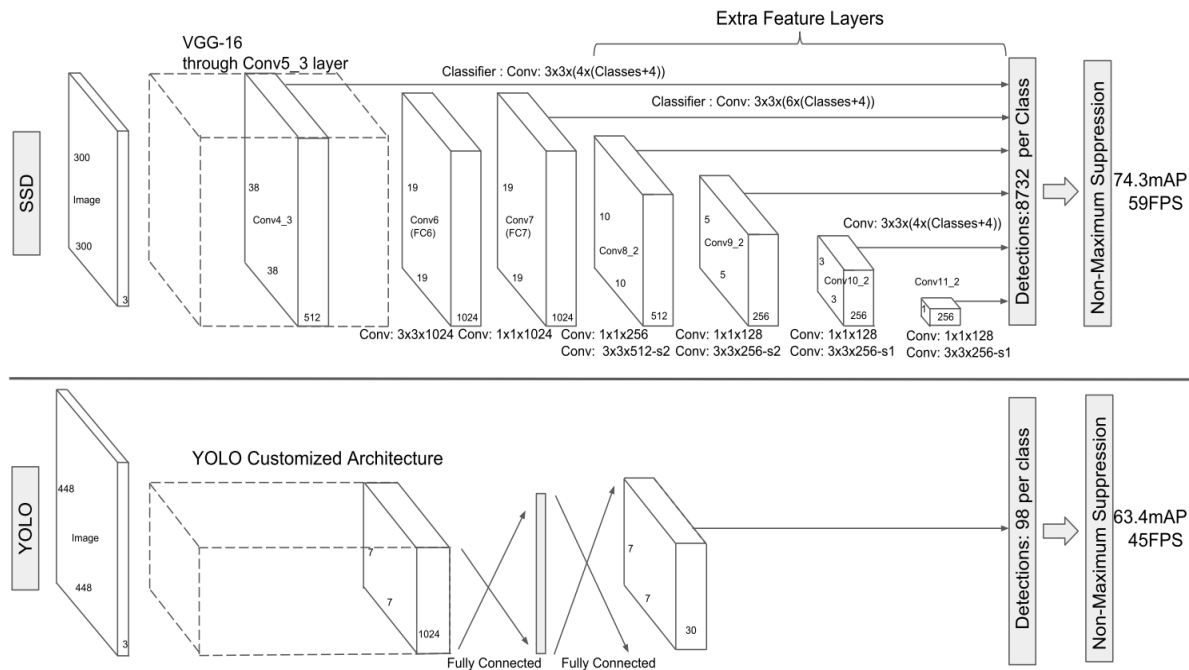
- > Detektion (BoundingBox + Vorklassifizierung) mit YOLO [1] / SSD [2]
- > Datensatz mit ~1200 Bildern
- > Vorklassifizierung in 4 Klassen: frying, cooking, lid_covered, empty



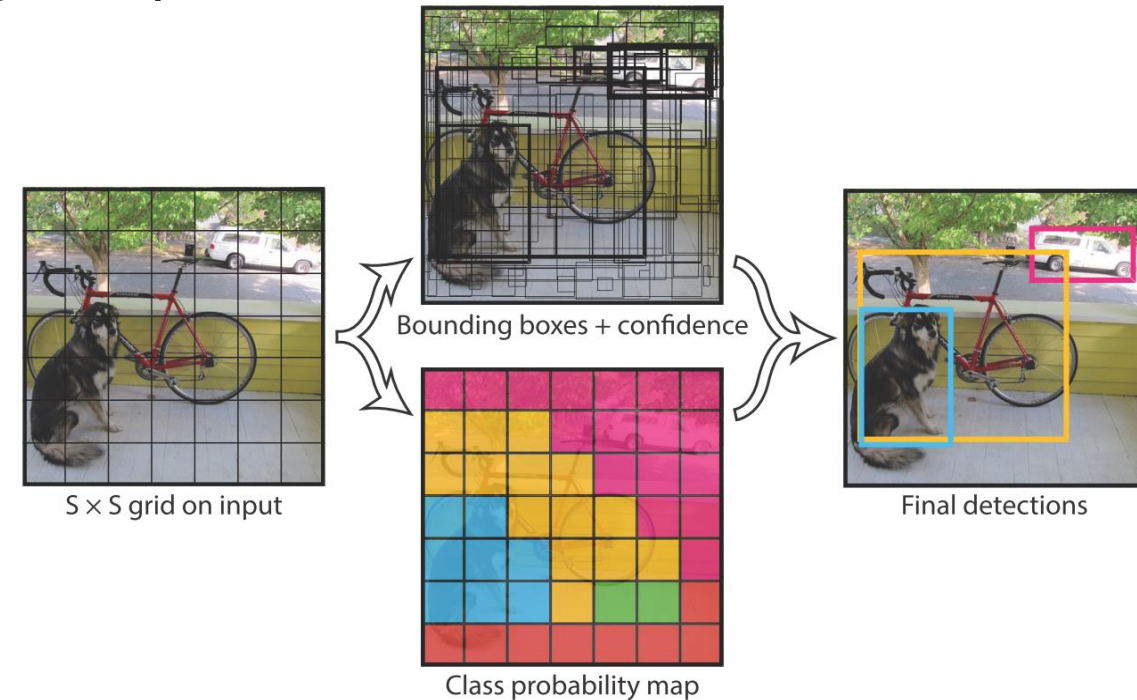
YOLO / SSD - Detektion

Single Shot Detectors: YOLO und SSD

- > Das Netzwerk wird nur einmalig durchlaufen (im Ggs. zu R-CNN Ansätzen)
- schnell, end-to-end Training (Backpropagation)



[Quelle 2: Wei Liu - SSD]



[Quelle 1: Redmond - YOLO]

Kamera - Detektion

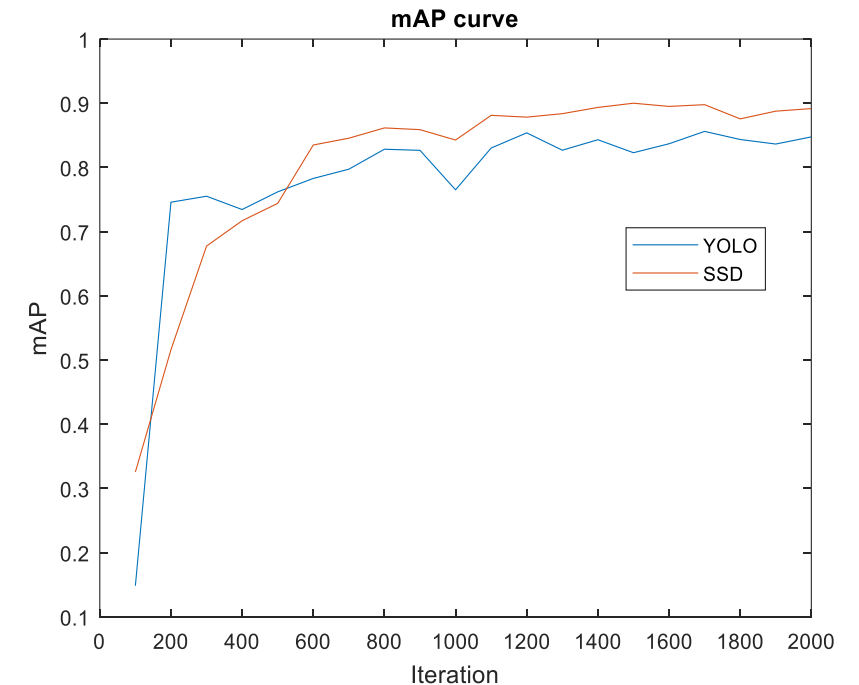
Ergebnisse Detektion

- > YOLO: darknet framework, C-Code (Aufruf mittels system(command))
- > SSD: caffe framework, C++, Python/MATLAB interface (MEX-File)

Klasse	YOLO	SSD
frying (AP)	89,69%	91,77%
cooking (AP)	78,83%	81,46%
lid_covered (AP)	80,75%	89,53%
empty (AP)	93,24%	97,28%
all classes (mAP)	85,62%	90,01%

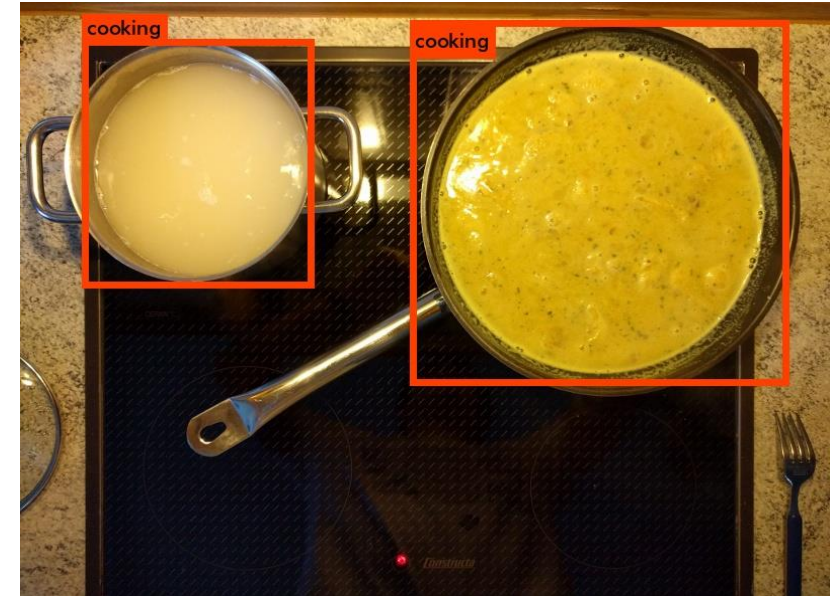
Nach PASCAL VOC2012 Metrik

- AP: Average Precision (Flächeninhalt der Precision-Recall-Kurve, IoU=0.5)
- mAP: mean Average Precision



Kamera - Klassifikation

- > Detektierte Bounding Box aus dem Bild ausschneiden
- > Klassifikation des Kochguts mit separatem Netzwerk (Transfer Learning mit AlexNet, VGG16/19, GoogLeNet, ResNet)
- > Datensatz von Essen mit ~62.000 Bildern und 16 Klassen

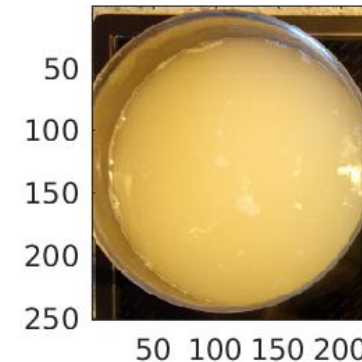
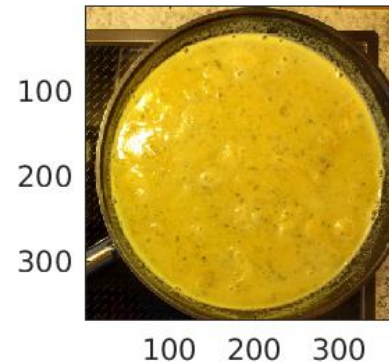


- > Ergebnis:

im_num: 1

label: chicken_curry

conf: 0.9601



im_num: 2

label: rice

conf: 0.8375

Kamera - Klassifikation

> Beispielbilder der 16 Klassen



Kamera - Klassifikation

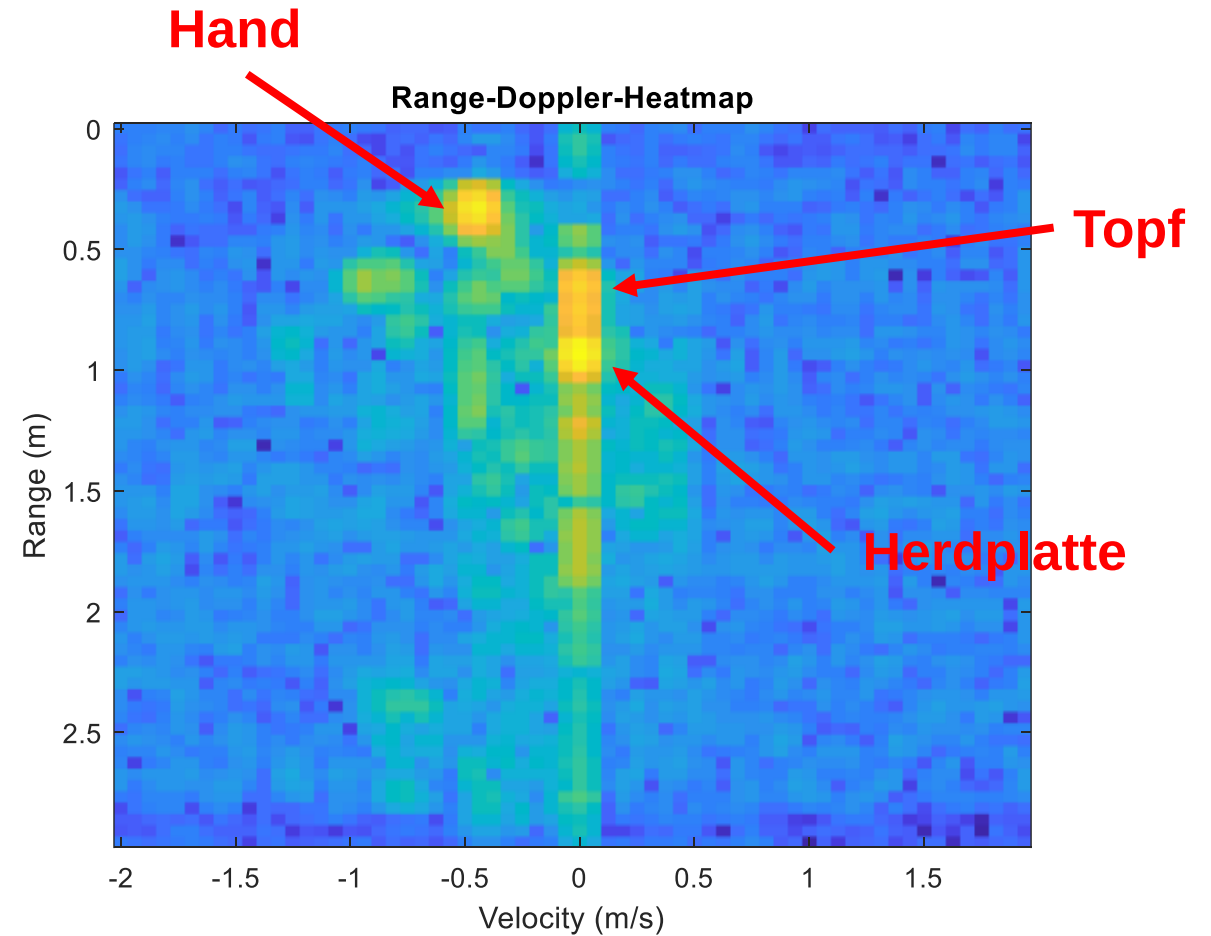
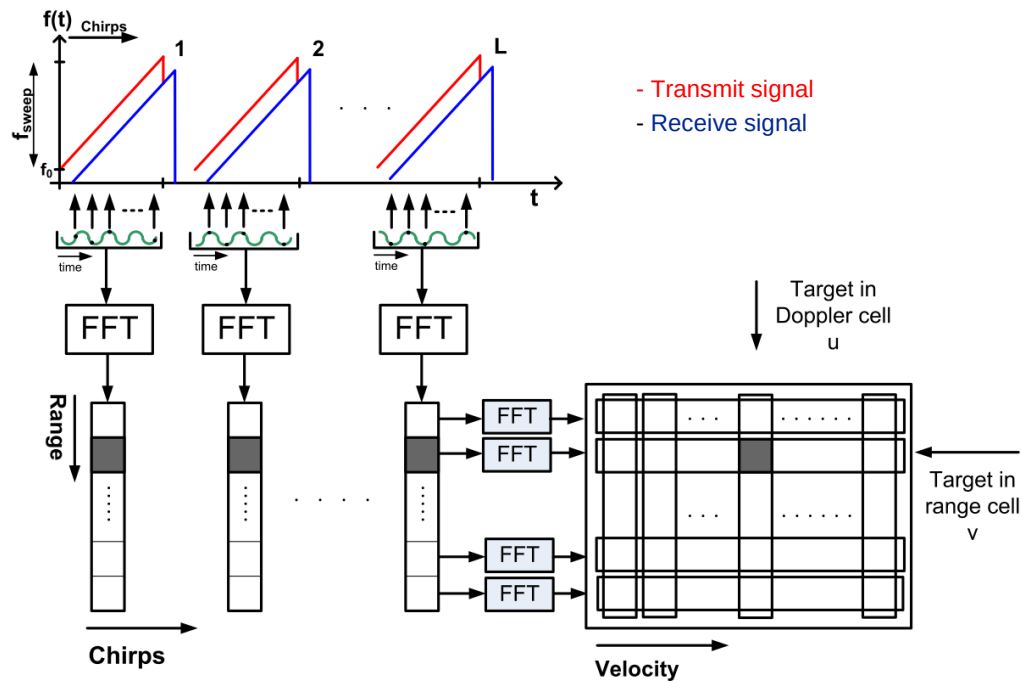
Ergebnisse Transfer Learning

- > Vortrainierte Netzwerke: AlexNet, VGG-16, VGG-19, GoogLeNet, ResNet
- > food_classification-Datensatz: ~62.000 Bilder, 16 Klassen

Netzwerk	Genauigkeit	Anzahl Conv/FC Layer	Modellgröße
AlexNet	72,5%	5 / 3	207MB
VGG-16	78,9%	13 / 3	488MB
VGG-19	82,6%	16 / 3	508MB
GoogLeNet	79,7%	Inception Modules	22MB
ResNet50	80,7%	Residual Network	86MB

77GHz FMCW Radar

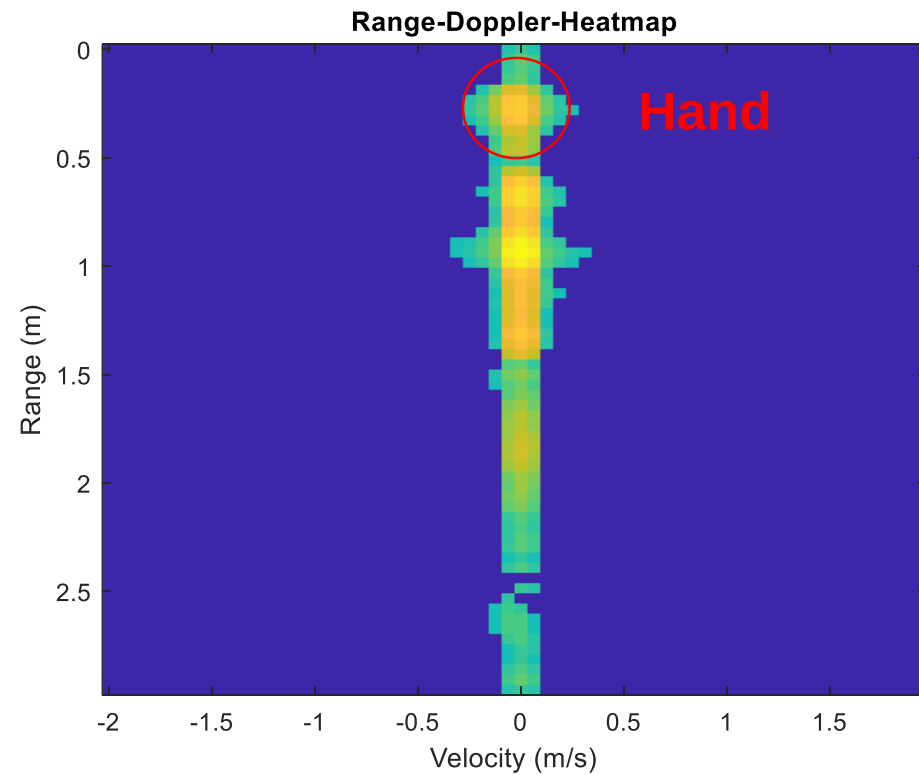
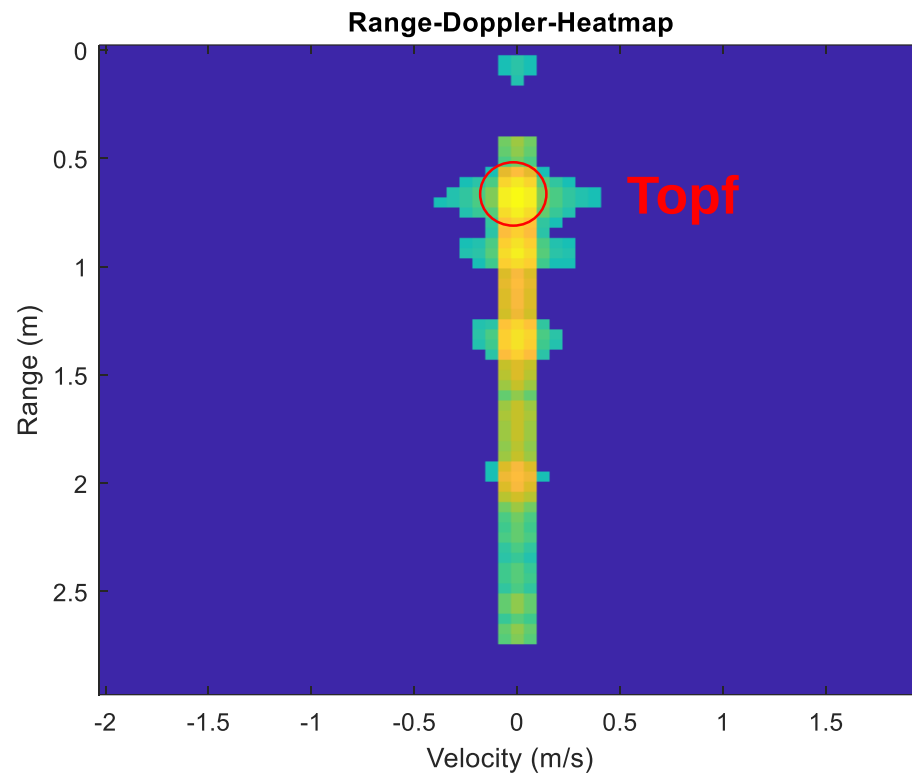
- > FMCW: Frequency Modulated Continuous Wave
- > Frequenz „Rampen“ (77-81 GHz)
- > Range-Doppler Maps (64x64 Pixel)



[Quelle 3: Schroeder, Rohling - X-Band FMCW Radar System]

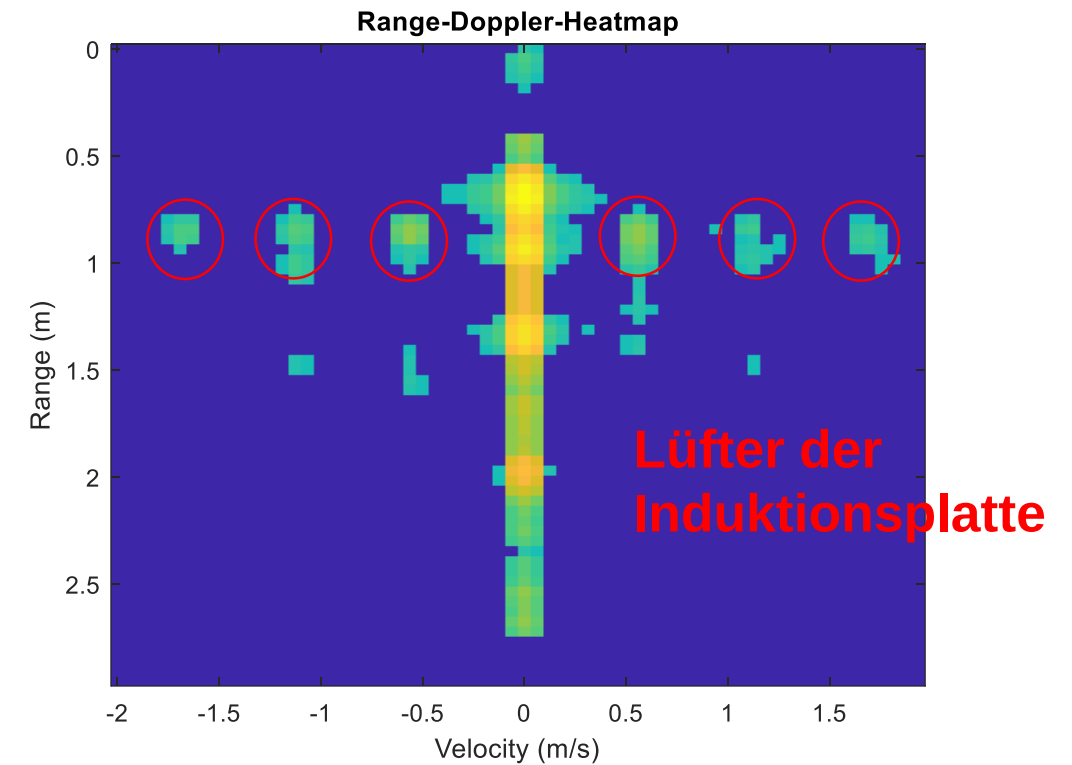
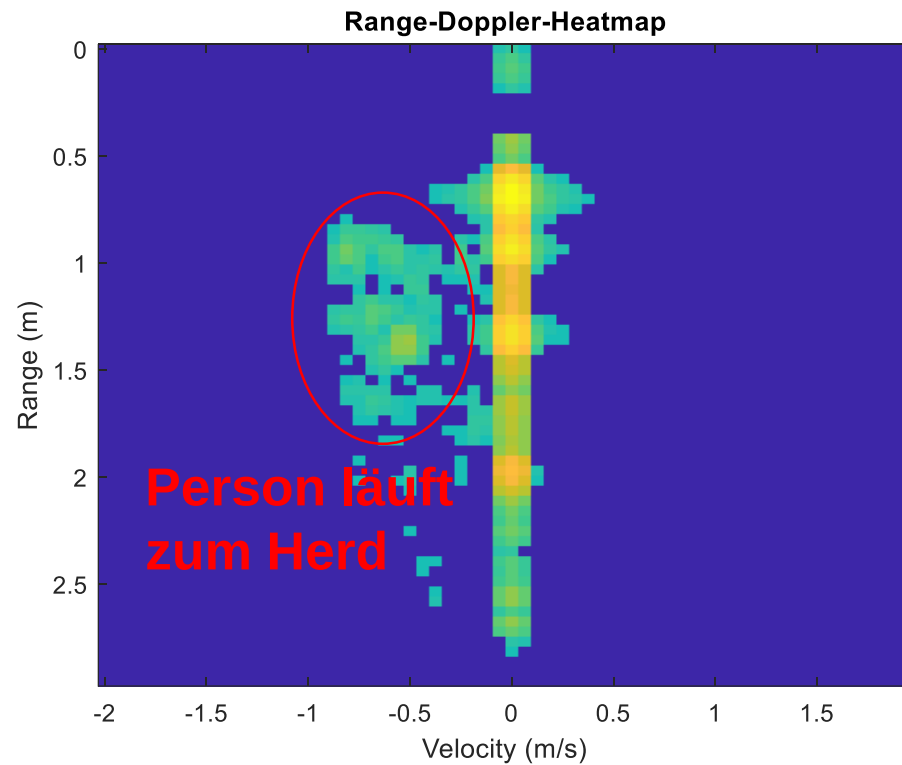
Radar

> Beispiele mit Rauschfilter



Radar

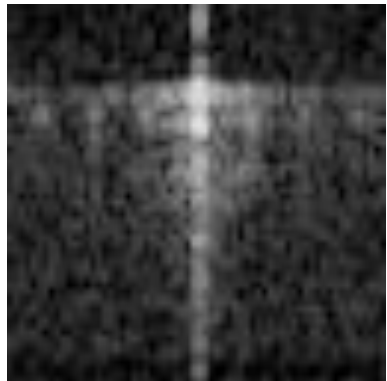
> Beispiele mit Rauschfilter



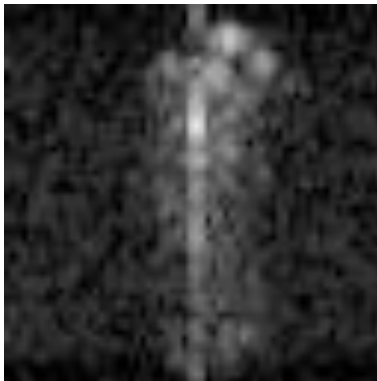
Radar - Klassen

- > 6 Klassen
- > je 3000 Range-Doppler Maps (18000 in Summe)

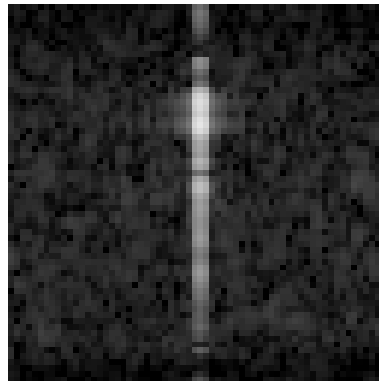
Boiling



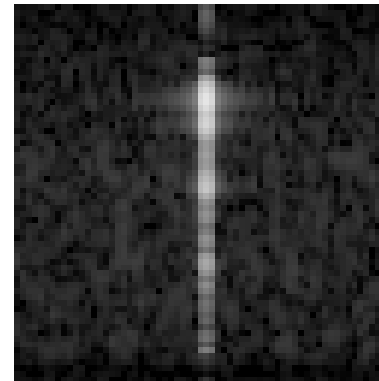
Hand-movement



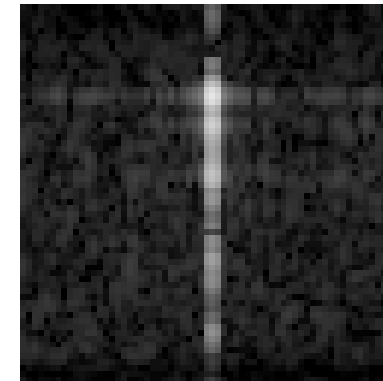
No-pot



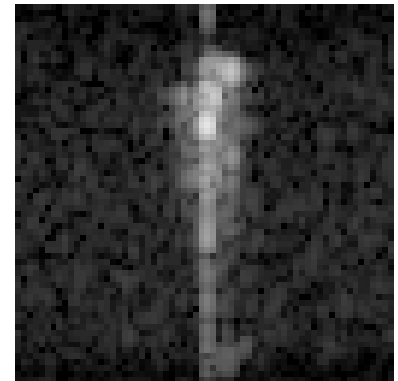
Pot-empty



Pot-filled



Stirring



Radar - CNN

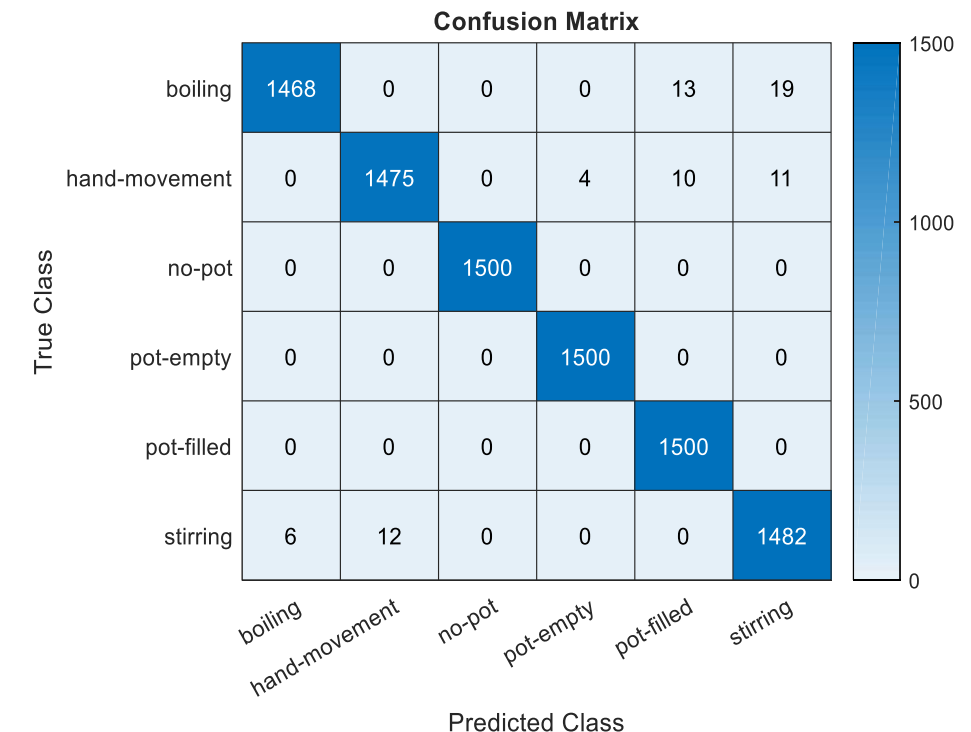
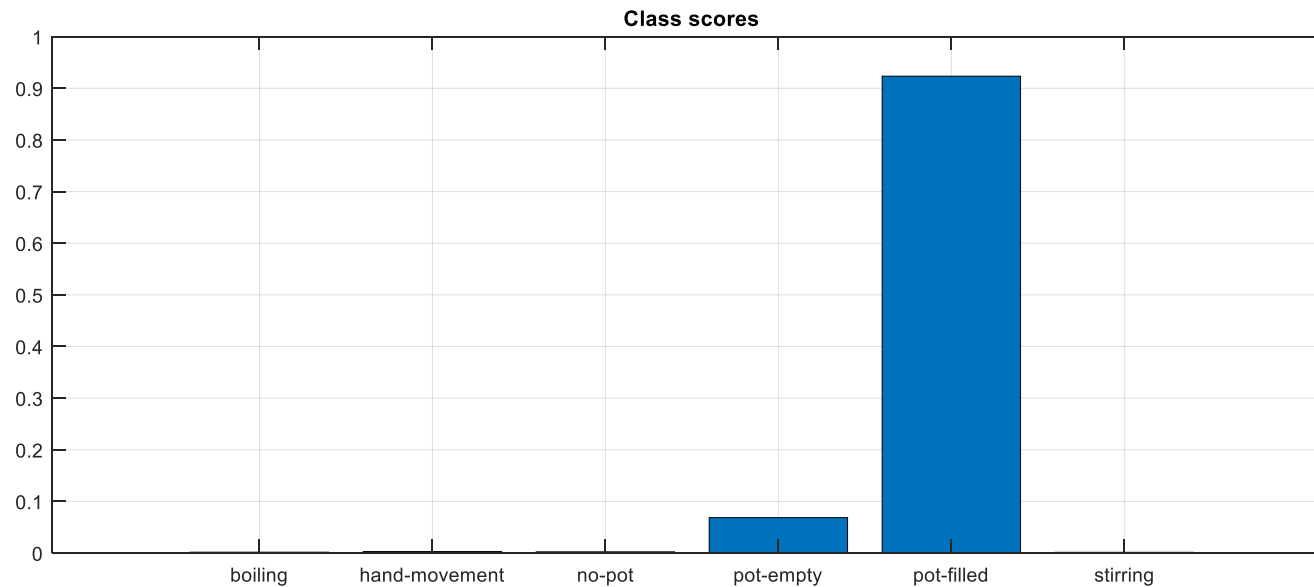
> Range-Doppler Maps: Auswertung mit Convolutional Neural Network (CNN)

No	Layer	Parameters	Input Data Size
1	Conv1	32, 3x3, Stride 2, Pad 0	64x64x1
2	Conv2	64, 3x3, Stride 2, Pad 0	32x32x32
3	Conv3	128, 3x3, Stride 2, Pad 0	16x16x64
4	FC1	256	8x8x128
5	FC2	Softmax (6 classes)	1x1x256

- > Batch Normalization, ReLU
- > 50% Dropout (nach Conv2/Conv3/FC1)
- > Input: Zerocenter Normalization
- > 2,2 Millionen Parameter, Modellgröße ~ 6,2 MB

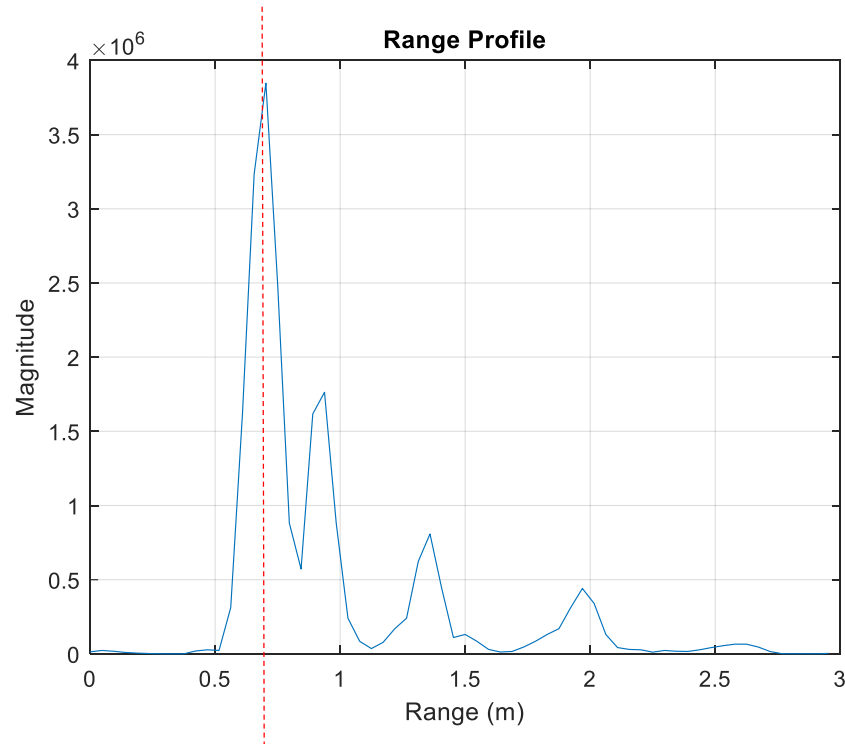
Radar - Ergebnisse

- > 6 Klassen, je 3000 Bilder
- > Ausgabe: Klassen-Wahrscheinlichkeiten
- > Wird Kochen erkannt, kann die Herdplatte abgeschaltet werden
- > Precision: 99,17%



Radar - Füllstandsmessung

- > Füllstandsmessung im Range Profile ($v=0$)
- > Referenzmessung zum Topfboden
- > Differenzmessung zur Flüssigkeitsoberfläche



Bedienoberfläche

Cooking Control Panel

Set Power manually

Power

Set Automatic Control

Off ☒ On

Status Initializing...
Automatic Control: OFF

Level 4.9 cm

Seq Length

Seq Name

Radar – Deep Dream

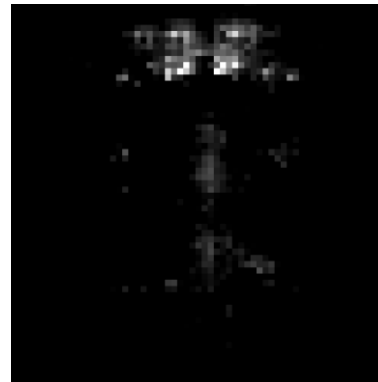
Visualisierung: „Was hat das Netzwerk gelernt?“

> Deep Dream Image: Das neuronale Netz erstellt Range-Doppler Maps

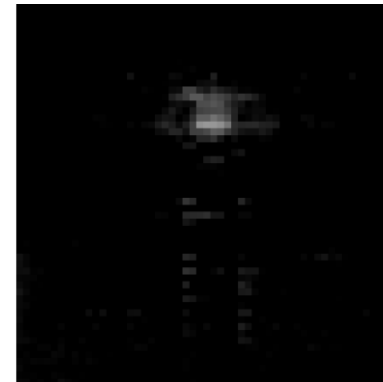
Boiling



Hand-movement



Pot-empty

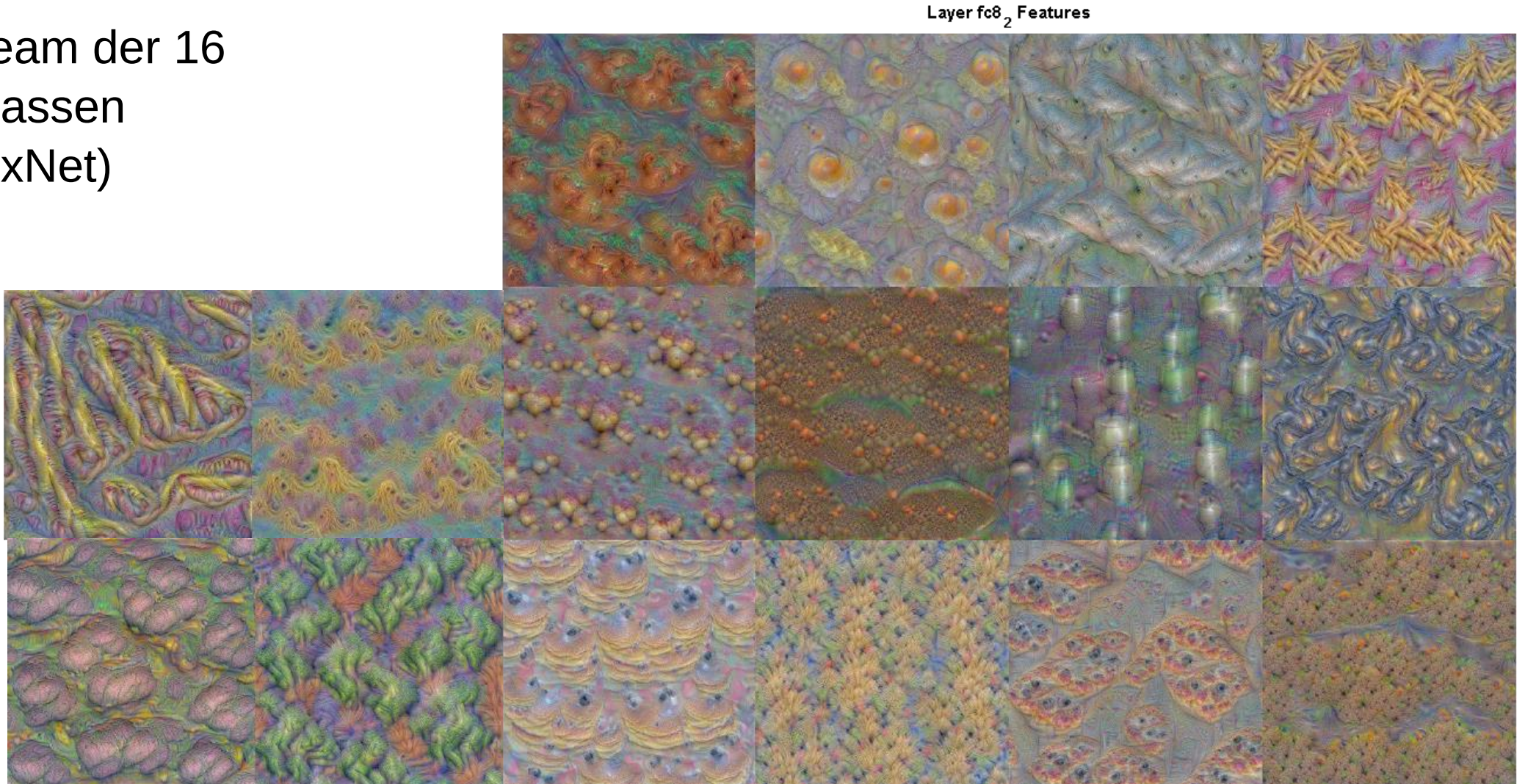


Pot-filled



Kamera – Deep Dream

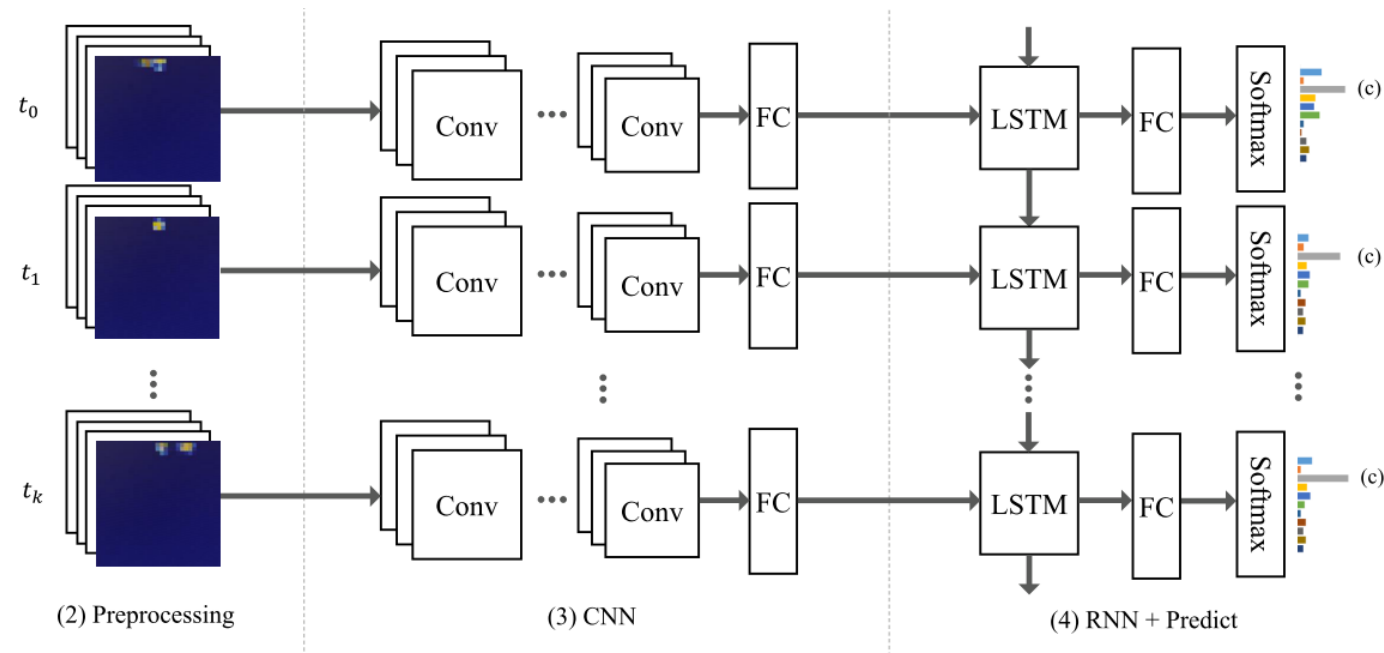
- > Deep Dream der 16 Essensklassen (Bsp. AlexNet)



Radar – Sequenz / LSTM

Ausblick

- > Sequenz von Eingangsdaten verarbeiten → Recurrent Neural Network
- > CNN + RNN (LSTM cells, „Gedächtniszellen“)



[Quelle 4: Google – Interacting with Soli]

Zusammenfassung

Kamera

- > Detektion der Kochstellen: mAP = 90% (SSD)
- > Klassifikation des Kochguts: Precision = 82,6% (VGG-19)

Radar

- > Klassifikation verschiedener Kochzustände
- > Füllstandsmessung
- > Abschaltung der Kochplatte bei drohendem Überkochen/Überschäumen

Ausblick

- > 160GHz/122GHz Sensoren (statt 77GHz) verwenden
 - Höhere Auflösung
- > Sensordatenfusion (Radar + Kameradaten) in gemeinsamen neuronalem Netz verarbeiten
- > Gestensteuerung (per Radar, Studienarbeit läuft)

VIELEN DANK!

Danksagungen:

- Prof. Dr.-Ing. Nicolaj Stache
- Prof. Dr.-Ing. Peter Ott

Bei Fragen kontaktieren Sie bitte:

M.Eng. Marco Altmann

Hochschule Heilbronn | Projekt MikroSens
marco.altmann@hs-heilbronn.de

Quellen

1. Redmon, Joseph ; Divvala, Santosh ; Girshick, Ross ; Farhadi, Ali: *You Only Look Once: Unified, Real-Time Object Detection* (2016). URL <http://pjreddie.com/yolo/>
2. Liu, Wei ; Anguelov, Dragomir ; Erhan, Dumitru ; Szegedy, Christian ; Reed, Scott ; Fu, Cheng-Yang ; Berg, Alexander C.: *SSD: Single Shot MultiBox Detector* (2016). URL <http://arxiv.org/pdf/1512.02325v5>
3. C. Schroeder und H. Rohling : IEEE-Paper: *X-Band FMCW Radar System with Variable Chirp Duration*. Hamburg University of Technology, 2010.
4. Wang ; Song ; Lien ; Poupyrev ; Hilliges: *Interacting with Soli: Exploring Fine-Grained Dynamic Gesture Recognition in the Radio-Frequency Spectrum* : in Proceedings of the 29th Annual Symposium on User Interface Software and Technology, New York UIST 2016.