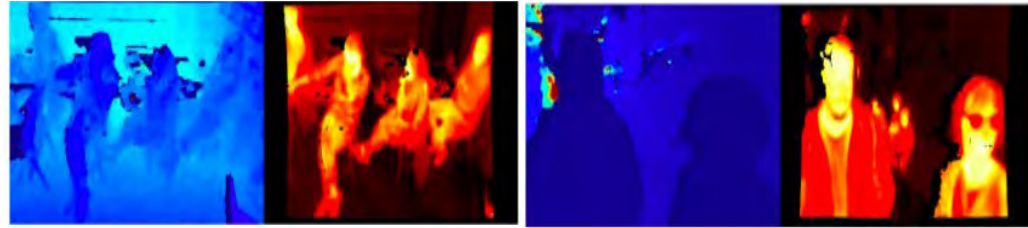


**Ai** App  **informatics**



# Methoden der künstlichen Intelligenz im Strafvollzug

Favoritenstr. 9/193-1, 4. th floor

A-1040 Vienna, AUSTRIA

Phone: +43-1-58801-193164

[www.cvl.tuwien.ac.at/](http://www.cvl.tuwien.ac.at/)

Prof. Martin Kampel

App Informatics ZT GmbH  
Technische Universität Wien

# Überblick

## Einführung

- Technologien im Strafvollzug
- Projekt KIIS – Künstliche Intelligenz im Strafvollzug

## Daten

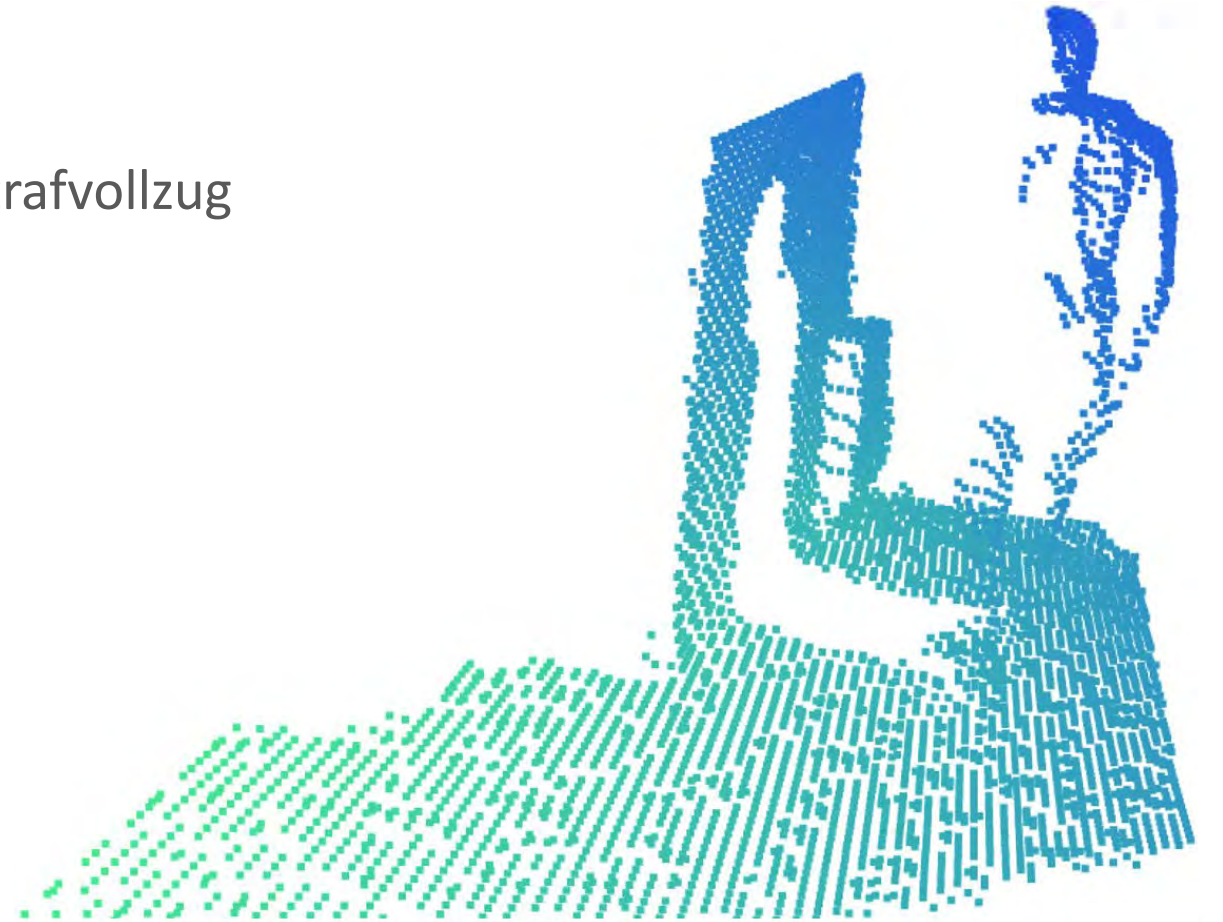
- Sensoren
- Datenaufnahme und Annotation

## Methodik

- Daten Representation
- Workflow
- AI Komponenten

## Begleitende Tätigkeiten

- Sozio Technische Analyse



# Technologien im Strafvollzug – Digitalisierung im Strafvollzug

- **Sicherheitstechnologien**

Gewährleistung der Sicherheit und Kontrolle: Schutz der Insassen und des Personals, aber auch der Gesellschaft

- **Resozialisierungstechnologien**

Bildungs- und Ausbildungsprogramme, Digitale Inklusion, Wiedereingliederung, Prävention von Rückfällen, Kommunikation und sozial Integration

- **Verwaltung und Digitalisierung**

Prozesse laufen effizienter ab, Entlastung von Justizpersonal

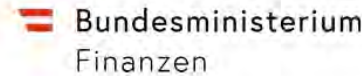
# Sicherheitstechnologien (1)

- **Video Überwachungssystem**  
Videoanalyse und Prävention von gefährlichem Verhalten
- **Zugangskontrollsystem (biometrische Systeme)**  
Auch Selbst - (Schleusensysteme) für weniger Begleitbewegungen
- **Drohnenabwehr**  
Lokalisation und Abwehr unautorisierter Drohnen
- **Mobilfunkblockaden**  
Signalaufbau eines unautorisierten Mobiltelefon unterbinden
- **Körperscanner**  
Einschleusen von unautorisierten Gegenständen unterbinden
- **Bodycams**  
Transparenz und Verantwortlichkeit erhöhen.

# Resozialisierungstechnologien (2)

- **Bildungs- und Lernplattformen**  
E-Learning-Systeme, virtuelle Bibliothek
- **Virtual Reality**  
Jobtraining und Simulation  
Therapeutische Anwendungen: Stressabbau, „Training“ Übergang von Haftzeit und Freiheit
- **Online Kommunikationssysteme**  
Telefonie, Videotelefonie, Arbeitsmarktservice online
- **Elektronische Fußfessel**  
Bedingungen der Freilassung überwachen

- KIIS – Künstliche Intelligenz im Strafvollzug



- Konsortium



- Interdisziplinäres Kooperationsprojekt

Technische Universität - BMJ – KMU – Sicherheitstechnik – Kriminalsoziologie

- Vor Ort Tests in verschiedenen österreichischen Justizanstalten

Finanziert durch KIRAS Sicherheitsforschung





## Problemstellung und Bedarf

- Hohe Belagszahlen und beschränkte Personalressourcen
- Unzureichende technische Hilfsmittel zur frühzeitigen Erkennung von Gefahrensituationen

## Hauptziele

- Bestmögliche Unterstützung der Bediensteten zur Erreichung der Zwecke des Strafvollzuges.
- Frühzeitige Erkennung von sicherheitskritischen Verhaltensmustern
- Schaffung eines Angst- und stressfreien Raumes für Inhaftierte
- Gewinn von Personalressourcen für unmittelbaren Kontakt mit den Inhaftierten



# Szenarien

## Die definierten Szenarien können aus technischer Sicht in drei Stufen unterteilt werden.

### Stufe 1

- Erfassung von Informationen zu einzelnen Personen ohne zeitlichen Kontext.
  - Mögliche Anwendungen: Erkennung des Aufenthaltsortes, Sturzerkennung, Messung von Vitalfunktionen

### Stufe 2

- Detektion und Klassifizierung von Interaktionen zwischen zwei oder mehreren Personen.

### Stufe 3

- Analyse von Verhaltensweisen und Handlungsmustern von Personen und Gruppen über längere Zeiträume.

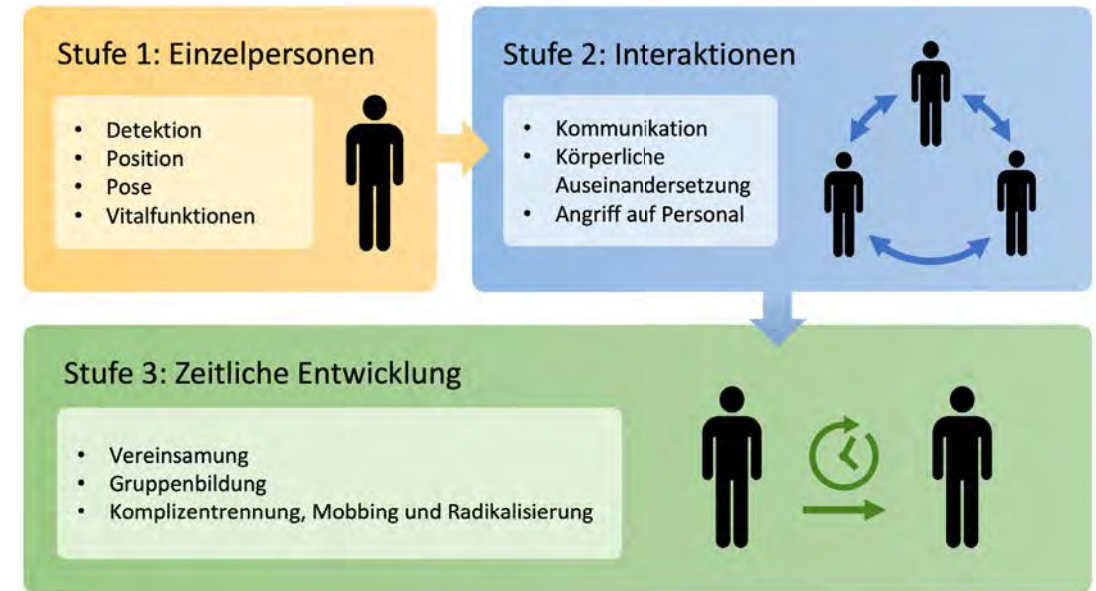


Abbildung 1: Angestrebte Szenarien



# Technischer Ansatz

- Multimodalität (lokal)

- Wearables (Positionsbestimmung, Vitalfunktionen)
- Bildgebende Sensoren
  - 3D
  - Thermal
  - RGB (kontextsensitiv)
- Weitere optionale Datenquellen (z.B. Audio)

- Fusion und XAI (zentral)

- Zusammenführung und verständliche Interpretation der erfassten Ereignisse.

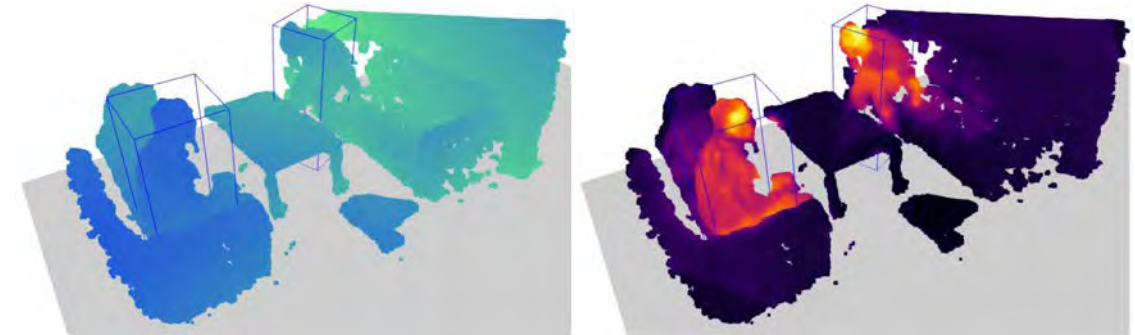


Abbildung 2: Fusion von 3D- und Wärmebilddaten

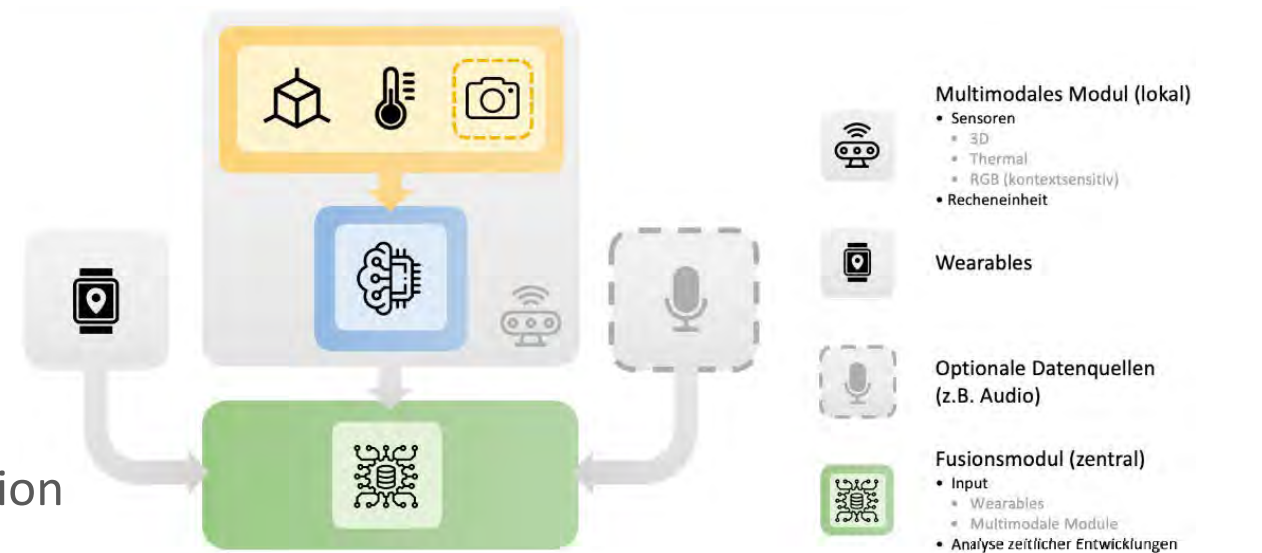


Abbildung 3: Informationsfluss und Interaktionen zwischen Modulen und Sensoren

**Ai** App   
informatics

- Mittels dezidierter Hardware Indoor-Tracking
- Autonomem Gang von A nach B von Insassen\*innen (keine Begleitbewegungen)
- Neugewonnene Selbstständigkeit trägt zu Normalisierung bei und spart zeitliche Ressourcen bei Justizpersonal

Daten – Sensoren, Aufnahme, Annotation

# Daten – Tiefensensoren

## Vorteile

- Anonymisierung – Akzeptanz – Privacy by Design
- Funktioniert unabhängig von Lichtverhältnissen
- Ermöglicht einfache Lokalisierung in 3D

## Nachteile

- Wenige öffentlich verfügbare Datensätze. Die meisten RGB-D-Datensätze enthalten keine Personen
- Besonders in statischen Szenen kann der Mangel an Farbe und Textur die Objekterkennung erschweren – auch für Menschen
- Verschiedene Tiefensensoren haben unterschiedliche Rauscheigenschaften → Erkennungsmodelle, die auf einem Sensortyp (z.B. Microsoft Kinect) trainiert wurden, funktionieren nicht unbedingt gut mit Daten von einem anderen Sensor (z.B. Orbbec Astra)



Orbbec Astra Sensor



Tiefenbild – Farbe zeigt Entfernung an

# Eine kostengünstige Tri-Modal-Kameraeinheit (LTCAT) für RGBDT-Vision



# Tri-Modal = RGB + Tiefenbild + Wärmebild



RGB / Video



Tiefenbild/3D



Wärmebild



# Datenaufnahme in Justizanstalten 1/3

Ab 2020 wurden Daten in österreichischen Justizanstalten gesammelt. Insgesamt wurden 9 Sensoreinheiten installiert, die alle in der Lage sind, mehrere Wochen lang rund um die Uhr unbeaufsichtigt aufzuzeichnen.

## Aufzeichnungssetup

- Sensor gesteuert durch Raspberry Pi
- Elektrische Boxen mit großen Festplatten, verschiedenen Adaptern, Netzsteckern usw. Eine Einheit mit gepanzertem Gehäuse

## Ergebnis

- 2.6 TB Daten
- 3 Monate Tiefensequenzen



Boxen mit Festplatten und verschiedenen Adaptern



Gepanzerte Einheit – Sensor (oben) und Raspberry Pi (unten)

# Datenaufnahme in Justizanstalten 2/3

## Arten von Sensorinstallationen

- 3x öffentliche Bereiche
- 2x private Räume – freiwillige Teilnahme der Insass\*innen

## Justizanstalt Göllersdorf

1. Gemeinschaftsraum – Geringe Aktivität rund um die Uhr (60m<sup>2</sup>)
2. Einzelhafttraum – Einige gewalttätige Verhaltensweisen, gepanzerte Sensoreinheit (12m<sup>2</sup>)

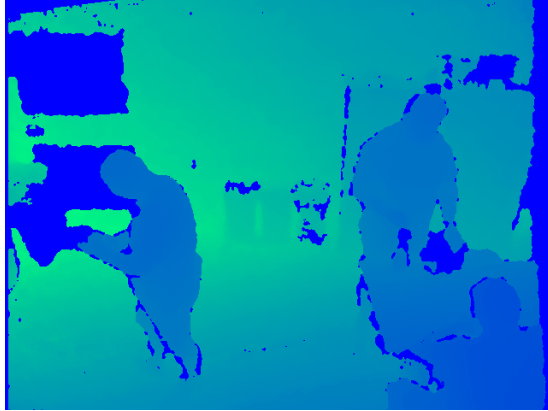
## Justizanstalt Simmering:

3. Doppelhafttraum – Soziale Interaktionen, kreatives Training (15m<sup>2</sup>)
4. Werkstattbereich – Regelmäßige Aktivität vor Mittag (80m<sup>2</sup>)
5. Raucherraum – Viel Aktivität und soziale Interaktionen (18m<sup>2</sup>)

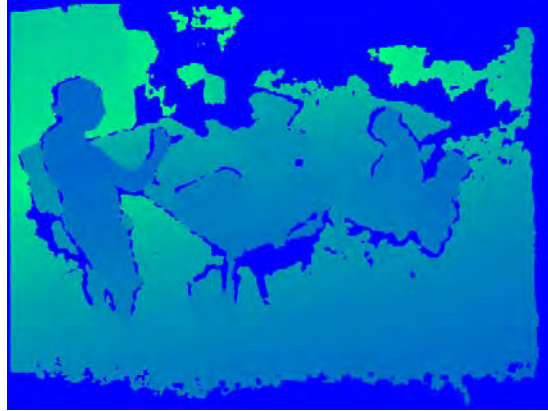


Sensorinstallationen in Strafanstalten  
(2020)

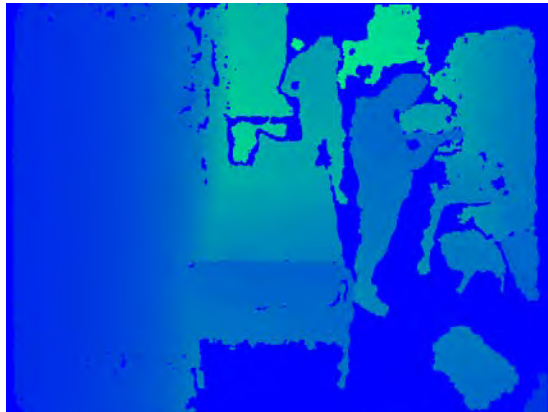
# Datenaufnahme in Justizanstalten 3/3



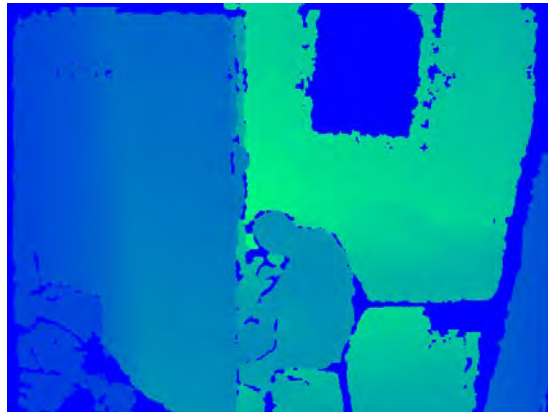
Gemeinschaftsraum



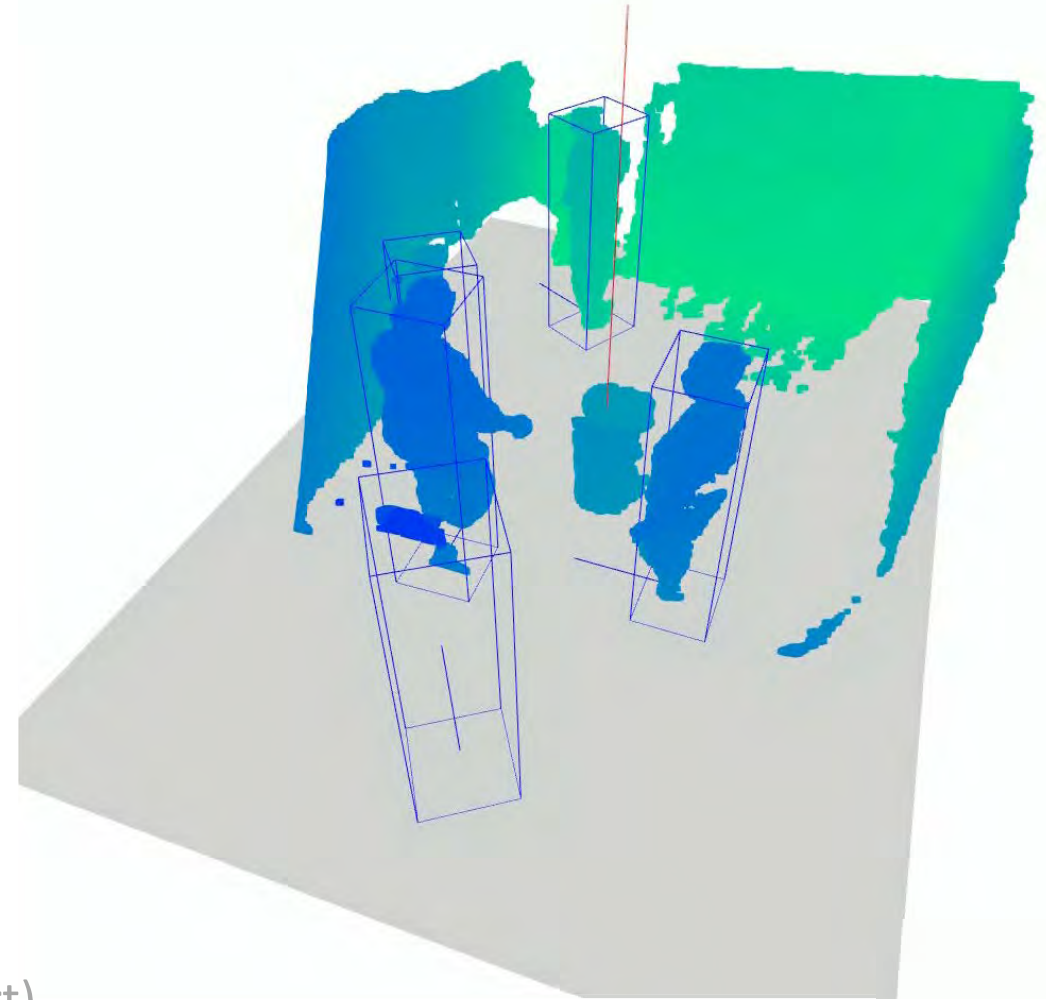
Werkstattbereich



Doppelhafttraum



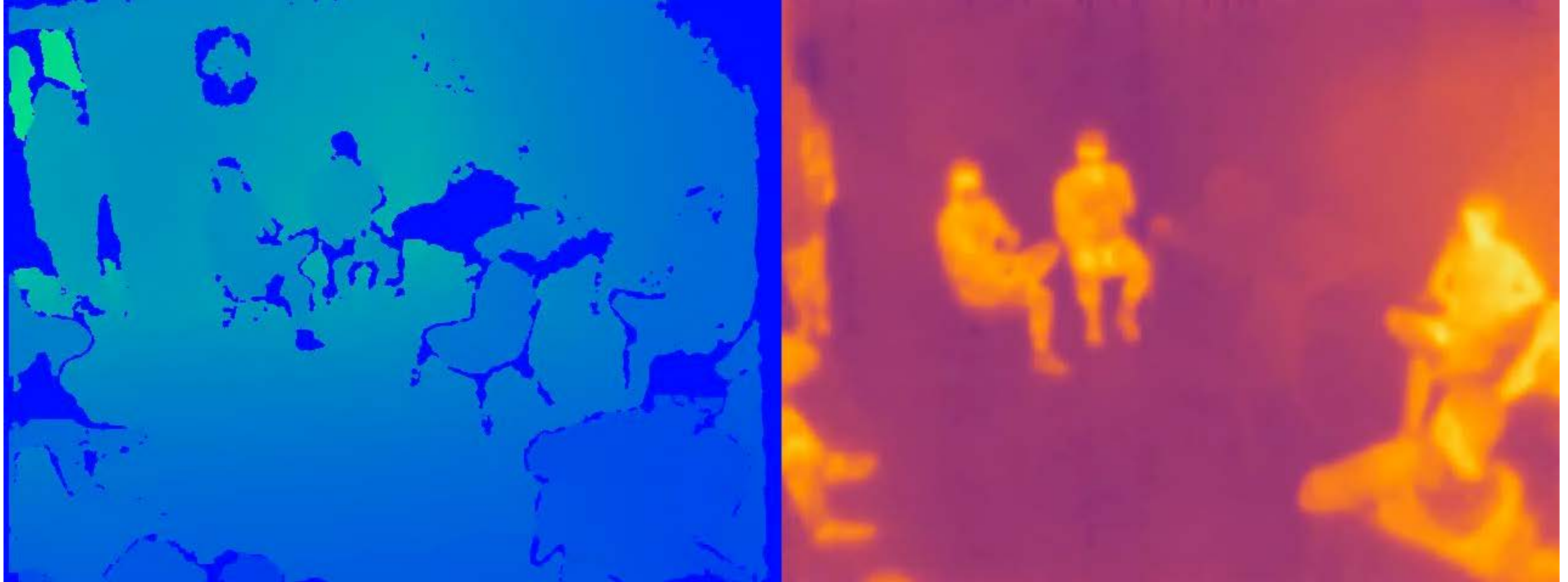
Einzelhafttraum (Sensor gepanzert)



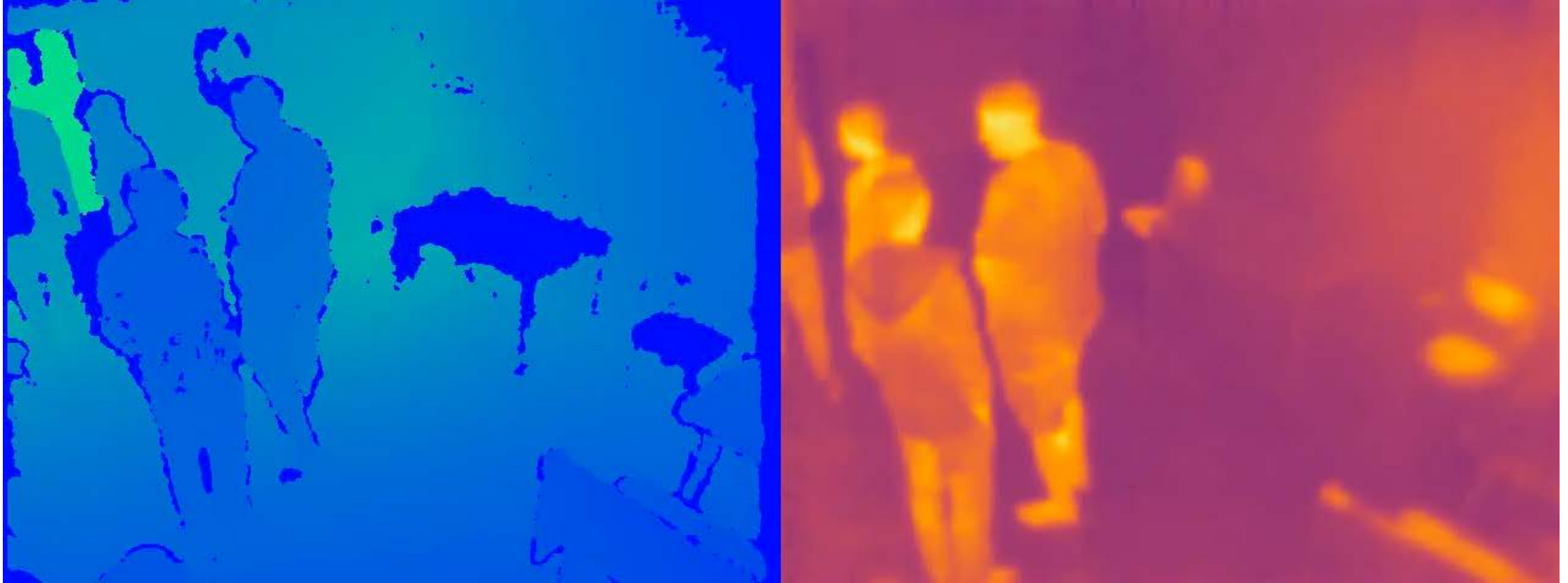
Video: Vollständige 3D-Rekonstruktion der Szene



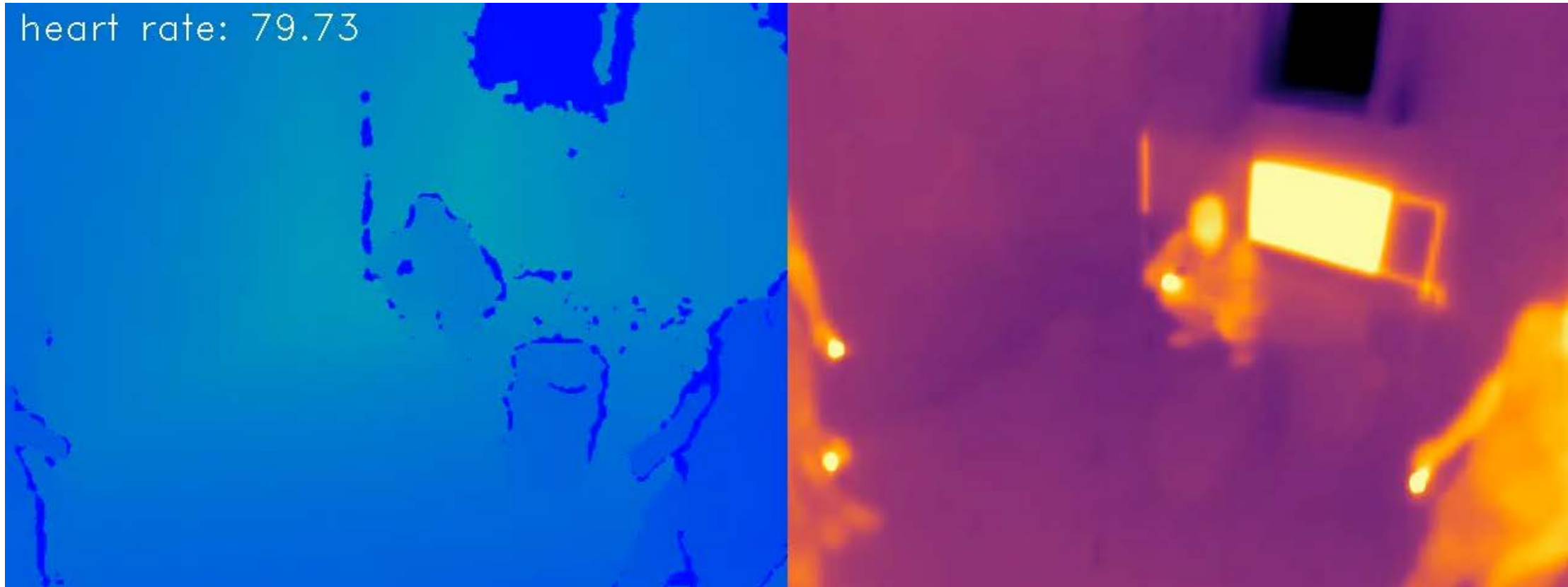
# Datenaufnahme Gruppendiskussion Gerasdorf 1/2



# Datenaufnahme Interaktion Gerasdorf 2/2



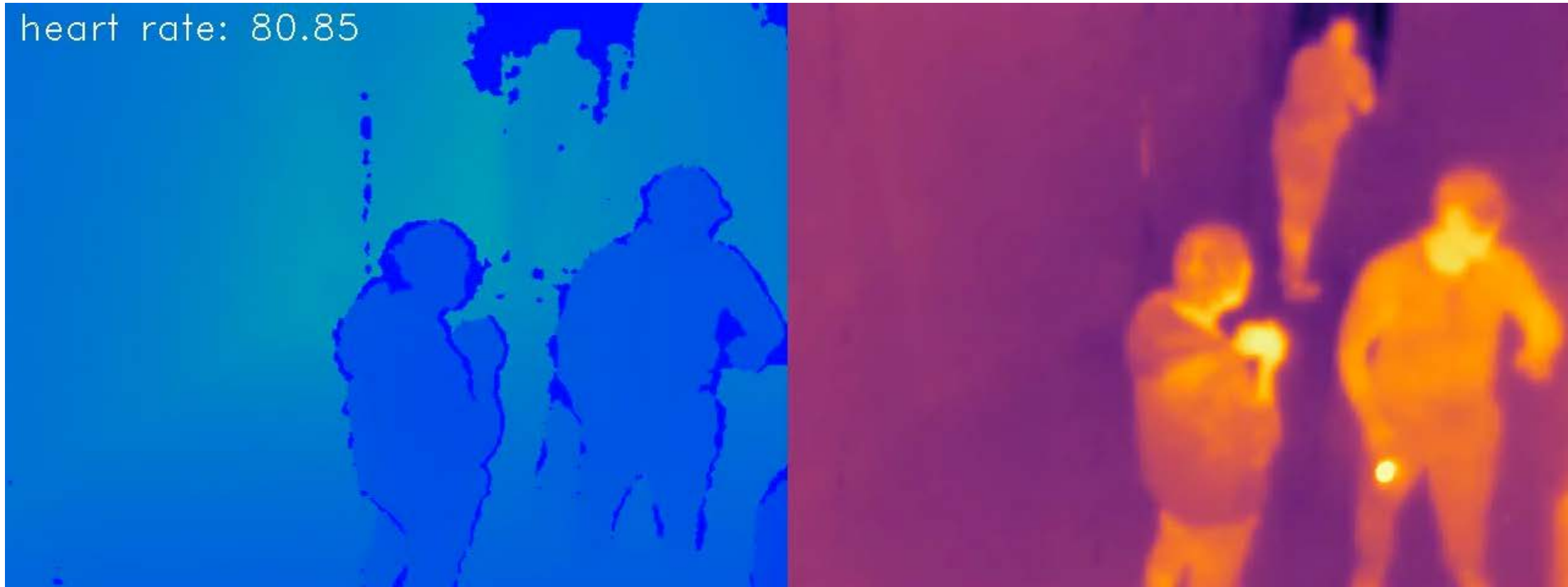
# Datenfusion Wearable – Visuals Wien-Simmering 1/2



Datenfusion: Herzfrequenz von einem tragbaren Gerät synchronisiert mit den Bildern



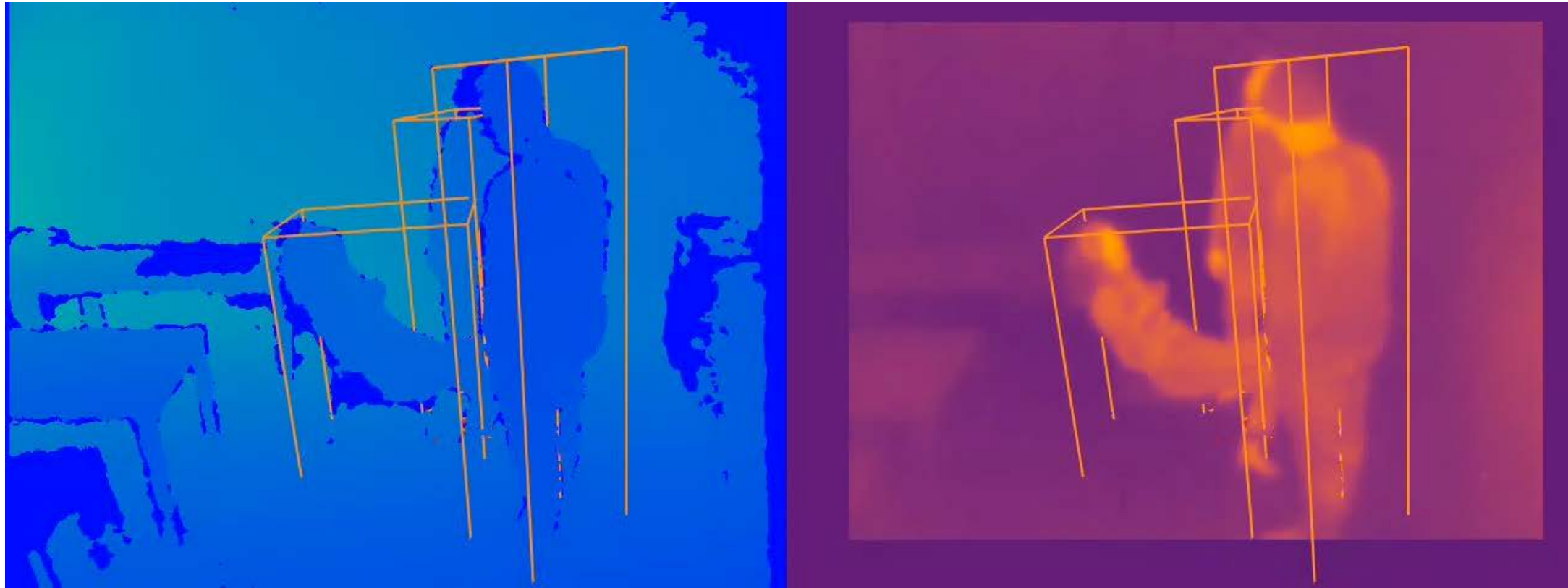
# Datenfusion Wien-Simmering 2/2



Diskussion

Datenfusion: Herzfrequenz von einem tragbaren Gerät synchronisiert mit den Bildern

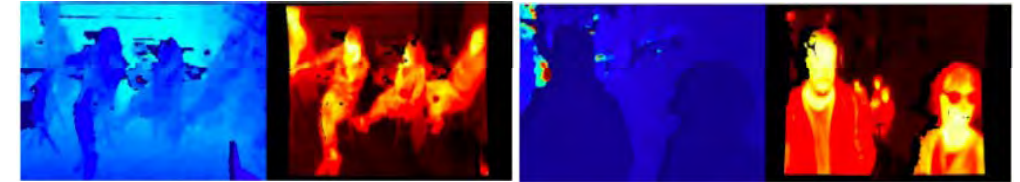
# Datenaufnahme „Gewalt“ Workshop



Simulierte Gewalt / Raubüberfall

# Daten – Öffentliche Datensätze

- [1] IPHD – Identitätserhaltende Menschenerkennung
  - Zeitlich und räumlich abgestimmte Tiefen-/Wärmebilder
  - Keine Sequenzen, 2D-Begrenzungsrahmenannotation
- [2] AAU VAP Trimodaler Personen-Segmentierungsdatensatz
  - Ausgerichtete RGB-DT-Videosequenzen
  - 2D-Segmentierungsannotation
  - Klein (3 Sequenzen)
- [3] RGB-D Personen-Datensatz
  - Ausgerichtete RGB-D-Videos
  - 2D-Begrenzungsrahmenannotation
  - Viele Personen in der Szene, aber nur eine einzelne Sequenz
- Häufige Schwächen bestehender Datensätze:
  - Keine 3D-Annotationen
  - Nur Bilder, keine Sequenzen
  - Kleine Größe



IPHD Datensatz



AAU VAP Trimodaler Personen-Segmentierungsdatensatz



RGB-D Personen-Datensatz

[1] Clapés, Albert, Julio CS Jacques Junior, Carla Morral, and Sergio Escalera. "ChaLearn LAP 2020 Challenge on Identity-preserved Human Detection: Dataset and Results." In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pp. 801-808. IEEE, 2020.

[2] Palmero, Cristina, Albert Clapés, Chris Bahnsen, Andreas Møgelmo, Thomas B. Moeslund, and Sergio Escalera. "Multi-modal rgb–depth–thermal human body segmentation." *International Journal of Computer Vision* 118, no. 2 (2016): 217-239.

[3] Spinello, Luciano, and Kai O. Arras. "People detection in RGB-D data." In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3838-3843. IEEE, 2011.

# Data – CVL Databases

## Aufgenommene Datensätze des Computer Vision Labs:

- [1] IPT (Identity Preserved Tracking) Dataset (2020)
  - Video Sequenzen mit statischem Sensor
  - Fixe Anzahl von Personen in der Szene
  - 3D Annotationen der Personen Verortung
  - >70k Frames
- Unpublished:
  - IPT Update: Extension to full 3D bounding boxes (2021)
  - Real world data collection in penal institutions (2020)



IPT Dataset Samples

[1] Heitzinger T. and Kampel M. “IPT: A Dataset for Identity Preserved Tracking in Closed Domains”, in Proceedings of the 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, January 2021

# Detektionsmodelle – Einschränkungen, Pipeline, Komponenten



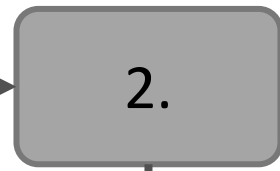
# Detectionmodel – Übersicht & Pipeline

## 1. Überlegungen basierend auf Einschränkungen:

- Hardware
- Datenrepräsentation



Model Input



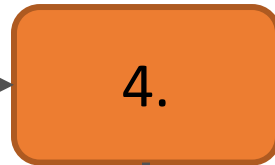
Vorverarbeitung:

- Normalisierung, Erweiterung
- Hintergrundmodell



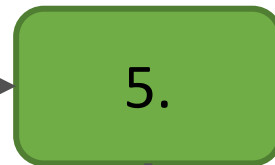
Objekterkennung

- Grundmodell
- Zielzuweisung



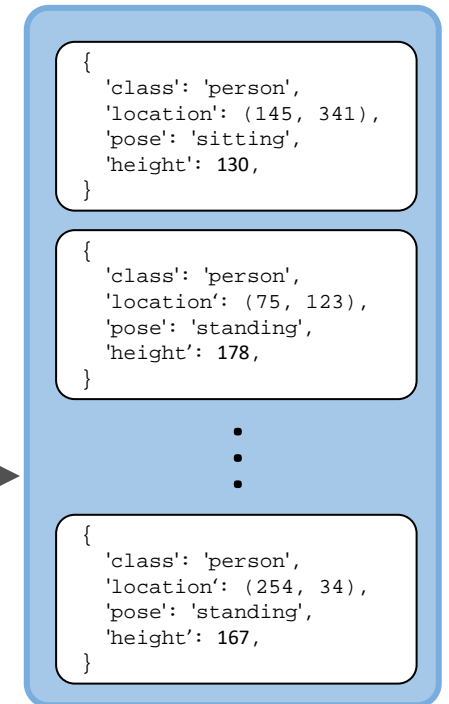
Nachverarbeitung

- Unterdrückung von Nicht-Maxima
- Erkennungsclustering



Tracking

DeepSORT



Model Output



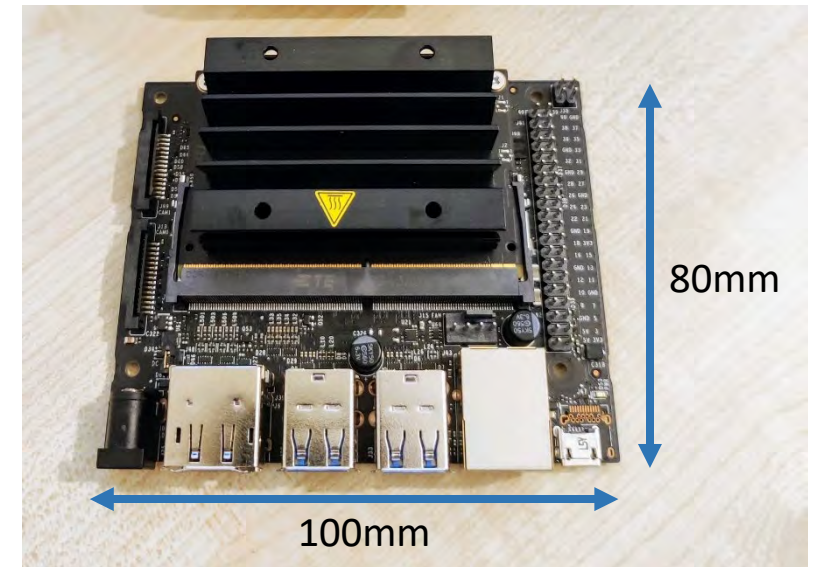
# Detectionmodel – Hardware

## Gründe für Edge Computing

- Datenerfassung und –verarbeitung auf demselben Gerät
- Potenziell sensible Daten verlassen das Gerät nicht

## Herausforderungen

- 3D-Erkennungsmodelle neigen dazu, langsam zu sein – im Vergleich zu 2D-Erkennung muss eine zusätzliche Dimension in der Modellarchitektur eingeführt werden
- Neuronale Netze sind groß (Millionen von Parametern)
- Strikte Anforderung: Verhaltensanalyse in Echtzeit (idealerweise  $> 10\text{fps}$ )



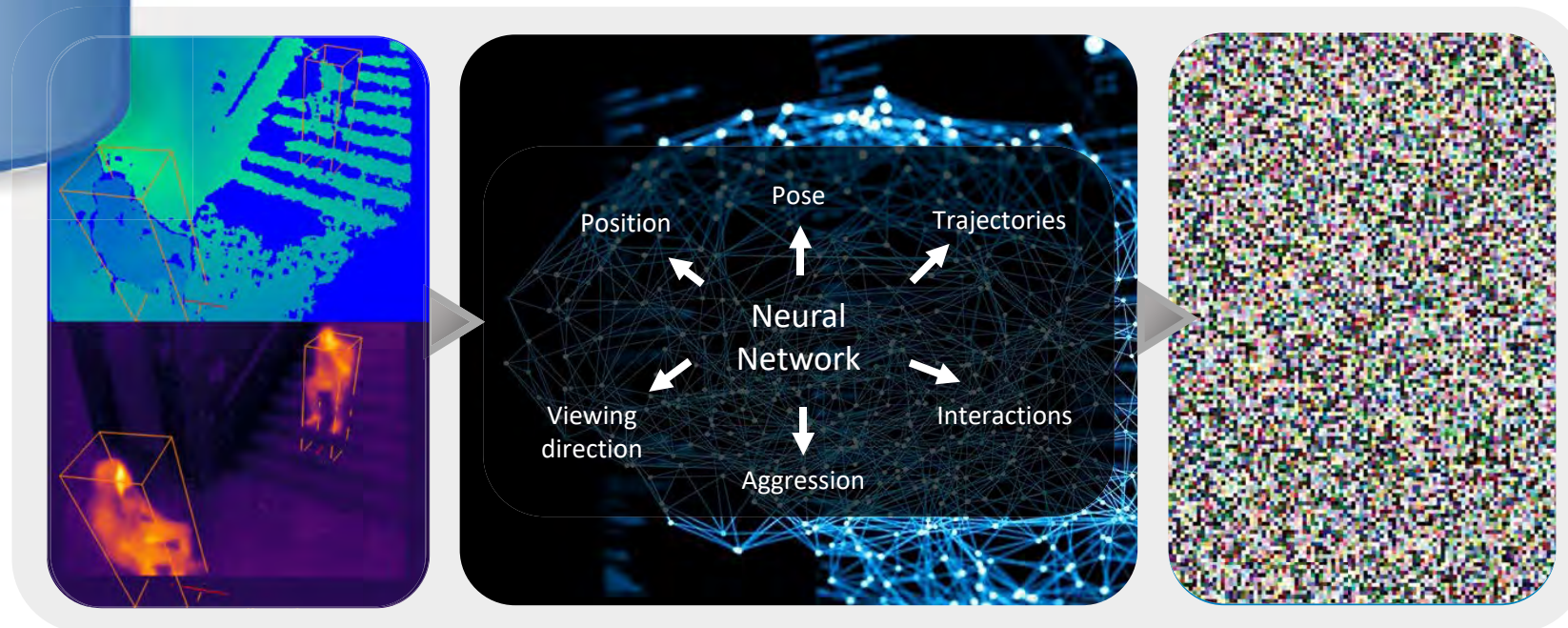
Derzeit verwendetes Gerät – NVIDIA Jetson Nano

# Neuronales Netzwerk | Lernphase



**Lernphase:** Ein anfänglich zufällig initialisiertes neuronales Netzwerk lernt Muster, die in einer annotierten Datenbank bereitgestellt werden (überwachtes Lernen).

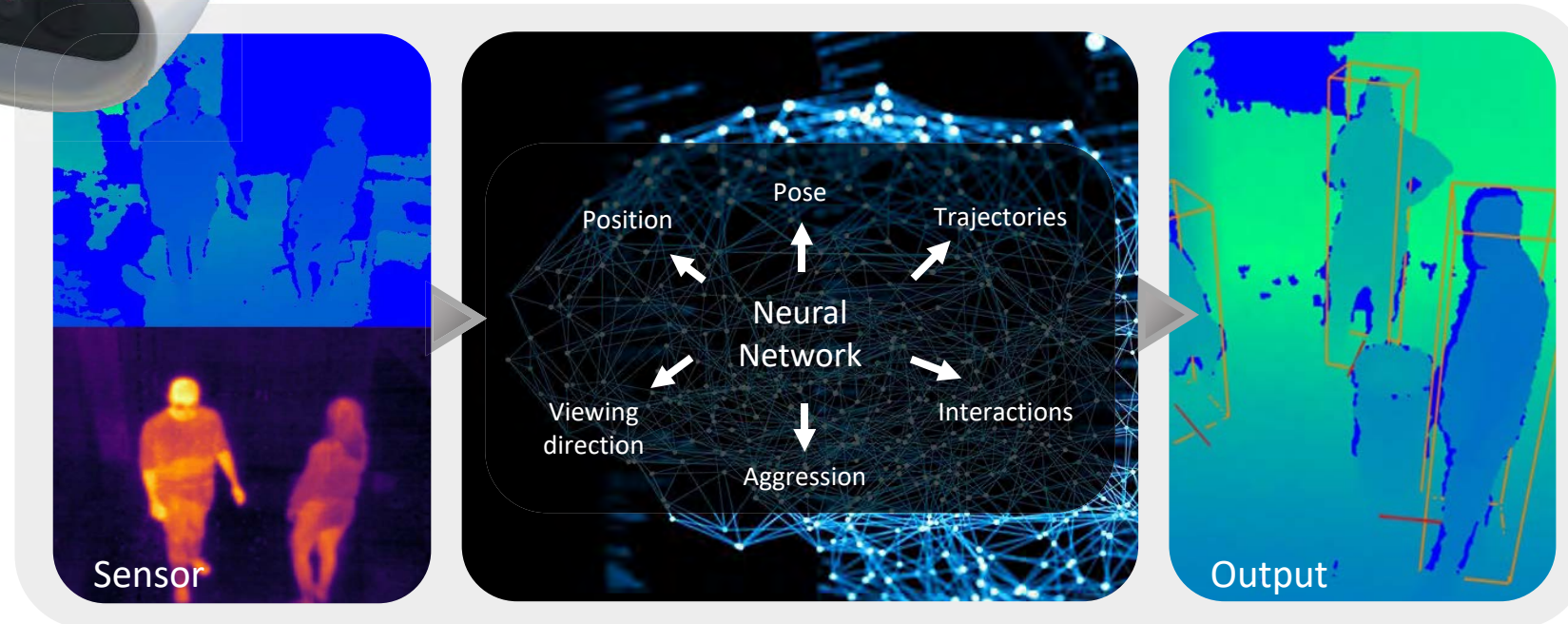
**Datenquelle:** Große Menge annotierter Bildsequenzen



# Neuronales Netzwerk | Anwendungsphase

**Anwendungsphase:** Das zuvor trainierte neuronale Netzwerk erkennt die gelernten Muster in bisher unbekanntem Daten.

**Datenquelle:** Multimodaler Sensor (Live-Daten)

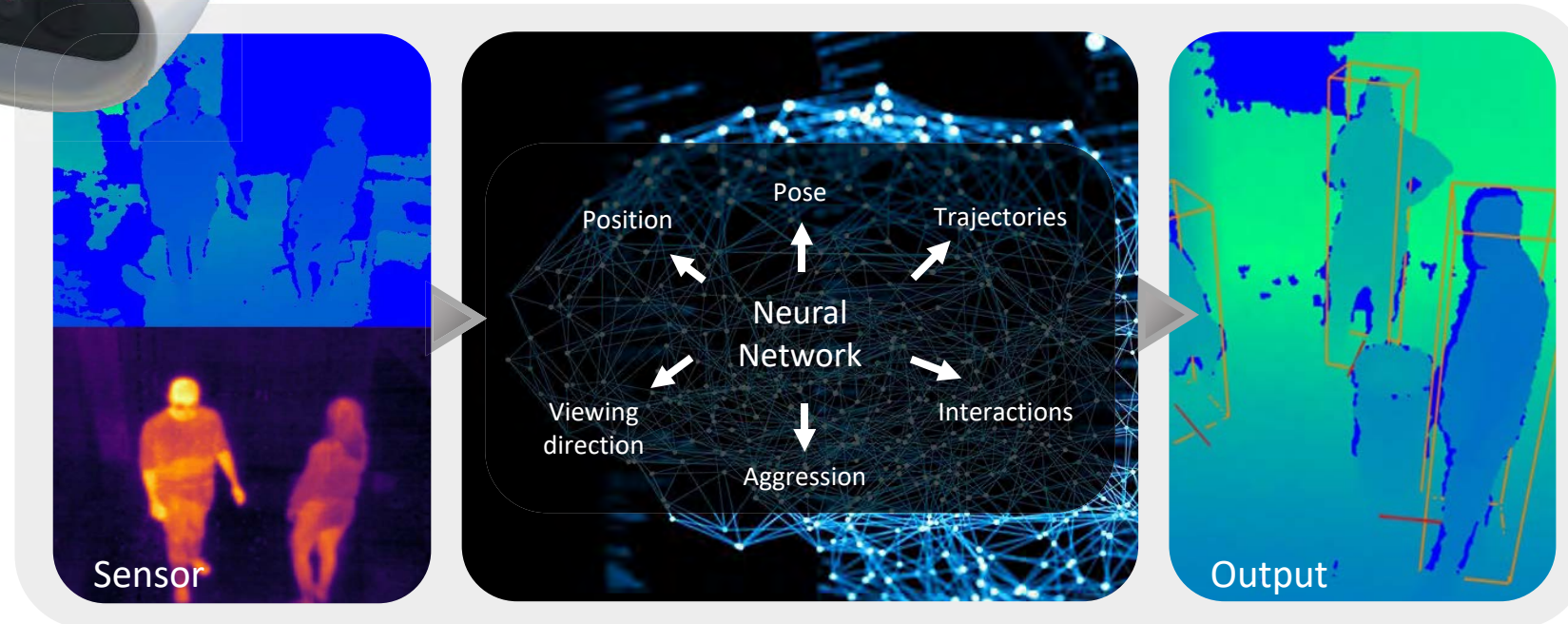
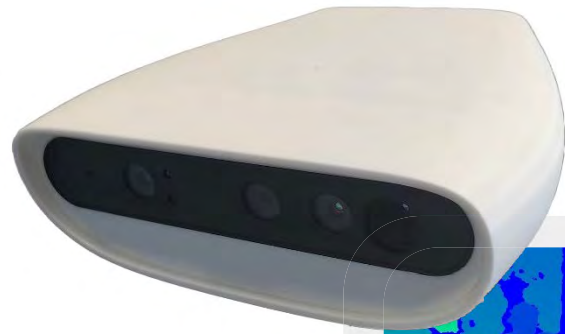




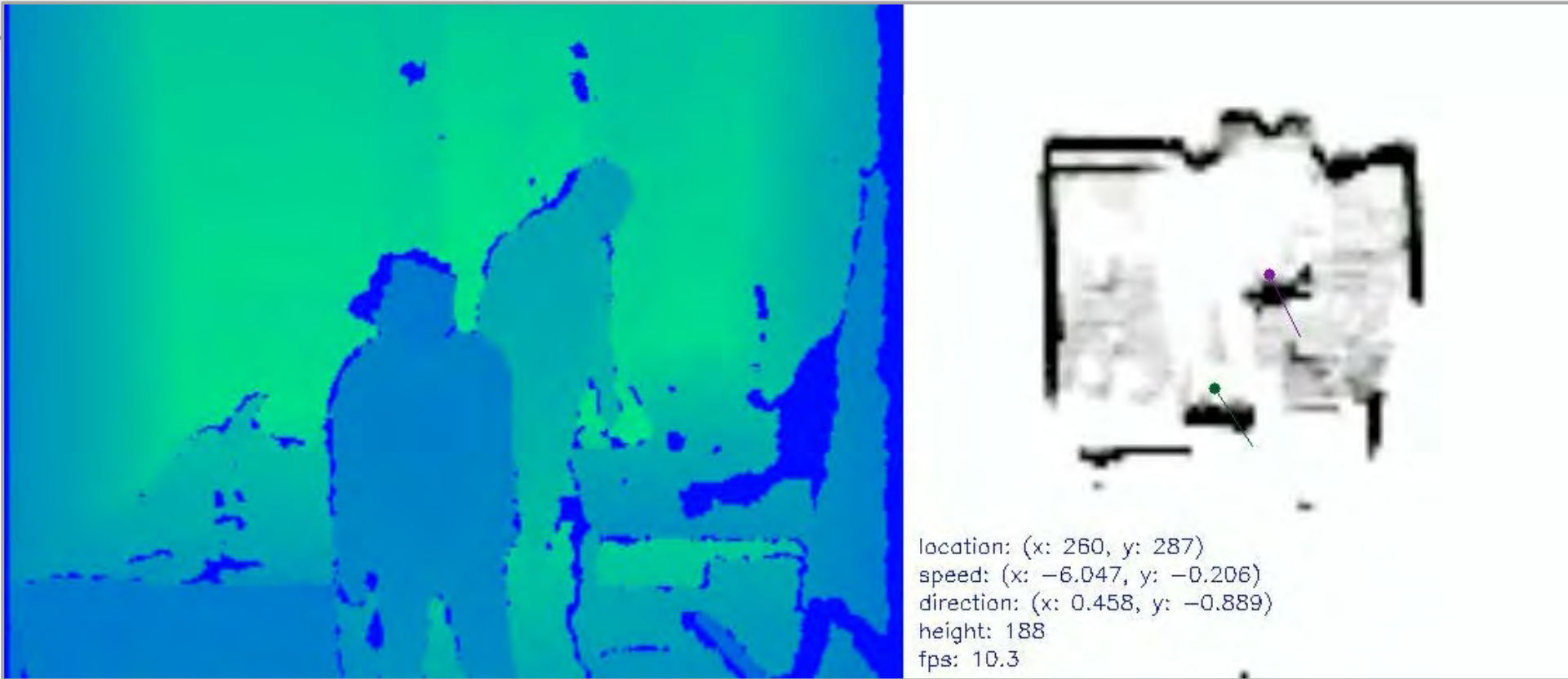
# Neuronales Netzwerk | Risiken

## Risiken

- Das neuronale Netzwerk kann nur Muster lernen, die während der Lernphase bereitgestellt wurden.
- Muster, die während der Lernphase nur selten gezeigt wurden, werden auch nur selten erkannt.



# Outcome



## Echtzeit-Erkennung und –Verfolgung auf dem Jetson Nano

Links: Netzwerkeingabe – Sequenz von Tiefenbildern

Rechts: Vogelperspektive derselben Szene

## Erkennung in der Vogelperspektive dargestellt:

- Punkte (Position) und Linien (Orientierung) für jede erkannte Person
- Farbwechsel deuten auf einen (falschen) Identitätswechsel hin

# Implementation & Ethics

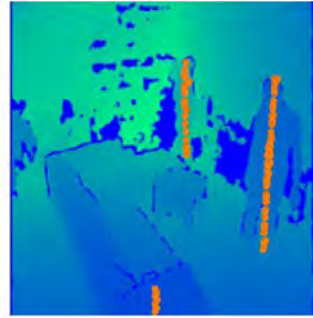
1. Entwicklung von Ethikrichtlinien „Maschinelles Lernen zur Verhaltensanalyse im Strafvollzug“
2. Begutachtung des Projekts durch die Forschungsethikkommission der TU Wien „Responsible Research Practices“.
3. Spezifische Richtlinien werden berücksichtigt:
  - Guidance document: Informed Consent: Gute Praxis
  - Europäische Kommission: Ethik in den Sozial- und Geisteswissenschaften, Ethik und Datenschutz, Identifizierung schwerwiegender und komplexer ethischer Fragen in der EU-finanzierten Forschung
  - Universität von Malta: Ethische Leitlinien
4. Rücksprache mit dem Datenschutzbeauftragten des Bundesministeriums der Justiz
5. Information an betroffene Personen: Namen und die Kontaktdaten der Verantwortlichen, Zweck der Verarbeitung, Kontaktdaten der Datenschutzbeauftragten, Speicherdauer der Daten und das Bestehen eines Auskunfts- und Beschwerderechts der betroffenen Person
6. Informationsveranstaltung in der JA: jeweils f. Insassen und Gefängnispersonal (Psychologen, Haustechniker, Anstaltsleitung)



# Zusammenfassung

## Hauptziel – Unterstützung des Personals in Justizanstalten

- Ortserkennung von Insassen
- Erkennung von Notfällen (Stürze, Streitereien)
- Erkennung von Handlungen / Interaktionen



## Technologische Ziele & Einschränkungen

- Aufzeichnung und Annotation geeigneter Datensätze
- Entwicklung spezialisierter Erkennungsmodelle für alternative Bildgebungsmethoden
- Echtzeitinferenz auf Edge-Geräten

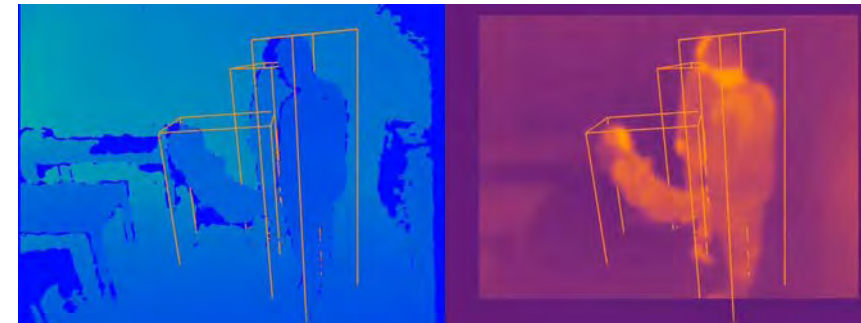
## Aufgezeichnete Daten

- Human Action Dataset: 20 GB
- Streaming-Daten: 100 TB
- Multimodal Identity Preserved: 7 GB
- Trimodale Segmentierung: 20 GB



## Sozialwissenschaftliche Analyse – laufend

- Gewinnung von Humanressourcen
- Auswirkung auf die Beziehung zwischen Personal und Insassen
- Gewaltprävention
- Erhöhte Sicherheit und verbessertem Schutz für Personal und Insassen



**Vielen Dank für Ihre Aufmerksamkeit**



Martin Kappel  
kappel@app-informatics.at

Favoritenstr. 9/193-1, 4. th floor


A-1040 Vienna, AUSTRIA

Phone: +43-1-58801-19364

Fax: +43-1-58801-18399

[www.cvl.tuwien.ac.at](http://www.cvl.tuwien.ac.at)



 Bundesministerium  
Finanzen