

# OPTIPLAN

**Ein innovatives KI-basiertes Planungstool zur Verkehrserfassung  
für nachhaltiges Flächenmanagement mittels Kamera- und  
Satellitenbilddaten (Verbundvorhaben)**

## **Schlussbericht**

Sachbericht Teil II: Eingehende Darstellung

Zuwendungsempfänger:

Teilvorhaben 1: Ingenieurbüro Behringer & Partner mbB  
(Förderkennzeichen 033LK007A)

Teilvorhaben 2: Deutsches Zentrum für Luft- und Raumfahrt (DLR) e.V.  
(Förderkennzeichen 033LK007B)

Laufzeit des Vorhabens 01.10.2023 – 30.09.2025

Berichtszeitraum 01.10.2023 – 30.09.2025

Berichtersteller: Stefan Gamperer (Ingenieurbüro Behringer & Partner)  
Dr. Tobias Leichtle (DLR)

***Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.***

## **Inhaltsverzeichnis**

<b>Inhaltsverzeichnis</b>	<b>1</b>
<b>Abbildungsverzeichnis</b>	<b>2</b>
<b>Tabellenverzeichnis</b>	<b>3</b>
<b>1 Kurzbericht</b>	<b>4</b>
1.1 Aufgabenstellung sowie wissenschaftlicher und technischer Stand	4
1.2 Ablauf des Vorhabens	4
1.3 Wesentliche Ergebnisse und Zusammenarbeit mit anderen Einrichtungen	5
<b>2 Ausführliche Darstellung der durchgeführten Arbeiten</b>	<b>6</b>
2.1 Projektmanagement (AP1.0)	8
2.2 Praxisanforderungen und Datenvorbereitung (AP2.0)	8
2.2.1 Bedarfsanalyse und Anforderungskatalog (Task 2.1)	8
2.2.2 Akquise & Vorprozessierung der Kamerabilddaten (Task 2.2)	9
2.2.3 Aufbau der Referenzdatenbasis (Task 2.3)	11
2.2.4 Akquise von Flächendaten (u.a. Satellitenbilder, amtliche Flächendaten) (Task 2.4)	13
2.3 KI-Methodenentwicklung zur automatisierten Auswertung von Kamerabilddaten (AP3.0)	14
2.3.1 Konzeptualisierung der KI-Methodenentwicklung (Task 3.1)	14
2.3.2 Entwicklung, Training & Validierung des KI-Verfahrens (Task 3.2)	15
2.3.3 Testung und Übertragbarkeit des KI-Verfahrens (Task 3.3)	18
2.4 Integration der Methodik für ein nachhaltiges Flächenmanagement (AP4.0)	20
2.4.1 Erfassung der Landbedeckung und Landnutzung mittels Erdbeobachtungsdaten (Task 4.1)	20
2.4.2 Evaluierung verschiedener Planungsvarianten zur Quantifizierung des Flächenverbrauchs (Task 4.2)	21
2.4.3 Prototypische Entwicklung eines Planungstools (Task 4.3)	24
2.5 Verbreitung und Praxistest (AP5.0)	26
2.5.1 Workshop mit externen Experten und Anwendern (Task 5.1)	26
2.5.2 Praxistest (Task 5.2)	27
2.6 Referenzen	30
<b>3 Wichtigste Positionen des zahlenmäßigen Nachweises</b>	<b>31</b>
<b>4 Notwendigkeit und Angemessenheit der geleisteten Arbeit</b>	<b>31</b>
<b>5 Voraussichtlicher Nutzen und Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans</b>	<b>32</b>
5.1 Voraussichtlicher Nutzen für das Ingenieurbüro Behringer & Partner	32
5.2 Voraussichtlicher Nutzen für das DLR	32
<b>6 Während der Durchführung des Vorhabens bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen</b>	<b>33</b>
<b>7 Erfolgte und geplante Veröffentlichungen der Ergebnisse</b>	<b>33</b>

## Abbildungsverzeichnis

Abbildung 1: Schematische Übersicht des experimentellen Designs. Die drei Experimente untersuchen komplementäre Einflussfaktoren auf die KI-gestützte Auswertung von Kamerabilddaten: (Experiment 1) den Einfluss von Wetter- und Beleuchtungsbedingungen, (Experiment 2) die Auswirkungen unterschiedlicher semantischer Klassifikationstiefen der Verkehrsteilnehmenden sowie (Experiment 3) den Vergleich verschiedener Modellarchitekturen innerhalb der YOLO-Familie. Die zentrale KI-Auswertung bildet den gemeinsamen methodischen Kern aller Experimente und ermöglicht eine vergleichende Bewertung unter praxisnahen Rahmenbedingungen.....	7
Abbildung 2: Hierarchisches Klassifikationsschema. Über drei Ebenen verteilt werden Verkehrsteilnehmer zunehmend spezifischer differenziert.....	8
Abbildung 3: Übersicht der beiden Untersuchungsgebiete mit den Standorten der kamerabasierten Verkehrserhebungen: links der Münchner Nordosten, rechts die Gemeinde Planegg.....	10
Abbildung 4: Übersicht der verfügbaren Bildausschnitte.....	11
Abbildung 5: Optimierter Prozess zum Aufbau der Referenzdatenbasis.....	11
Abbildung 6: Darstellung der kontextbasierten Fusion von Objektklassen. Links zeigen farbige Bounding Boxes (Rechtecke zur Lokalisierung erkannter Objekte) eine Person (blau) und ein Fahrrad (rot). Rechts ist das Ergebnis der Fusion als einzelner Verkehrsteilnehmer Radler (violett) dargestellt.....	13
Abbildung 7: WorldView-3 Aufnahme vom 04.07.2019. Überblick über das Stadtgebiet Münchens (a) und Detailausschnitt des Untersuchungsgebiets "Münchner Nordosten" (b)	14
Abbildung 8: Drei beispielhafte Aufnahmen illustrieren die unterschiedlichen Bildkompositionen an Standorten. Außerdem zeigen sie unterschiedliche Beleuchtungssituationen: a) sonnig, b) bedeckt und c) dunkel.....	15
Abbildung 9: Normalisierte Konfusionsmatrizen für drei semantische Klassifikationsebenen: a) generische Klassifikation (Ebene 1), b) FGSV-nahe Klassifikation (Ebene 2) und c) projektspezifische Klassifikation (Ebene 3). Die Diagonale zeigt korrekte Zuordnungen, Einträge außerhalb der Diagonale kennzeichnen systematische Fehlklassifikationen zwischen Klassen. Die Farbskala repräsentiert spalten-normalisierte Anteile im Wertebereich von 0 bis 1.....	17
Abbildung 10: Übersicht des integrierten Workflows zur Auswahl der praxisrelevanten KI-Methode. Dargestellt sind aussichtsreichsten Aufnahmebedingungen, ein einfaches semantisches Klassifikationsschema und die leistungsstärkste Modellarchitektur, sowie die darauf aufbauenden Trainings-, Validierungs- und Evaluationsschritte. Das optimierte Modell wird anschließend auf unbekannte Videodaten angewendet.....	18
Abbildung 11: Vergleich der zeitlich aggregierten Verkehrszählungen im Zeitfenster von 5:00 bis 6:00 Uhr. Dargestellt sind die manuelle Zählung des fließenden Verkehrs (schwarz) und die KI-gestützte Auswertung auf Basis frameweiser Detektionen (orange). Trotz abweichender absoluter Zählwerte zeigt sich eine hohe Übereinstimmung der zeitlichen Verkehrsdynamik.....	19
Abbildung 12: Detektionsergebnisse für vier aufeinanderfolgende Frames. Dargestellt sind die Originalaufnahmen (Graustufen) mit überlagerten, farbigen Bounding Boxes der KI-	

---

Detektion. (a–b) Ein Bus hält an der Ampel, (c) lässt eine Person die Straße überqueren und (d) setzt seine Fahrt fort.....	20
Abbildung 13: Ergebnis der Landbedeckungsklassifikation auf Basis höchstaufgelöster Erdbeobachtungsdaten. Überblick über das Stadtgebiet Münchens (a) und Detailausschnitt des Untersuchungsgebiets "Münchner Nordosten" (b).....	21
Abbildung 14: Lage des Untersuchungsgebiets "Münchener Nordosten" im Stadtgebiet Münchens (a) und Planungsentwurf des Gebiets (b).....	22
Abbildung 15: Lage der Vergleichsflächen im Stadtgebiet Münchens.....	23
Abbildung 16: Drei beispielhafte Neubaugebiete in München.....	24
Abbildung 17: Potentielle Landbedeckung für Grünzüge, Hüllgraben und das Aktivitätenband.....	25
Abbildung 18: Potentielle Änderungen der Landbedeckung auf Siedlungsflächen in Abhängigkeit der drei beispielhaften Neubaugebiete.....	26

## Tabellenverzeichnis

Tabelle 1: Klassenweise Detektionsleistung für zwei unterschiedlich komplexe Modellarchitekturen bei identischer Datengrundlage und generischer Klassifikation (Ebene 1). Angegeben sind die mittleren Genauigkeiten pro Klasse sowie die Differenz (v8x – v8n).....	18
--	----

## **1 Kurzbericht**

### **1.1 Aufgabenstellung sowie wissenschaftlicher und technischer Stand**

Ziel des Vorhabens OptiPlan war die Entwicklung eines KI-basierten Ansatzes zur automatisierten Auswertung von Kamerabilddaten sowie deren Verknüpfung mit flächenbezogenen Analysen zur Unterstützung einer nachhaltigen Verkehrs- und Infrastrukturplanung. Hintergrund ist der steigende Bedarf an belastbaren, datenbasierten Entscheidungsgrundlagen, während klassische Erhebungsmethoden wie manuelle Verkehrszählungen mit hohem personellem und zeitlichem Aufwand verbunden sind.

Der Stand von Wissenschaft und Technik ist geprägt durch leistungsfähige Deep-Learning-Verfahren im Bereich Computer Vision sowie durch etablierte Methoden der Erdbeobachtung zur Analyse von Landbedeckung und Flächennutzung. Während beide Bereiche für sich genommen weit entwickelt sind, fehlt bislang eine integrierte Anwendung, die automatisierte Verkehrserfassung und flächenbezogene Bewertung systematisch miteinander verknüpft. Zudem bestehen Herausforderungen hinsichtlich der Übertragbarkeit vortrainierter Modelle auf reale Verkehrsdaten, insbesondere unter variierenden Aufnahmebedingungen und bei unterschiedlichen Klassifikationsanforderungen.

Vor diesem Hintergrund zielte das Projekt darauf ab, eine praxisnahe Datenbasis aufzubauen, geeignete KI-Methoden systematisch zu evaluieren und in ein integratives Konzept zur Unterstützung von Planungsprozessen zu überführen.

### **1.2 Ablauf des Vorhabens**

Das Vorhaben gliedert sich in fünf aufeinander aufbauende Arbeitspakete.

AP1 beinhaltete die Projektkoordination sowie die Abstimmung zwischen den beteiligten Partnern.

AP2 umfasste die Analyse der Praxisanforderungen sowie die Akquise, Aufbereitung und Annotation umfangreicher Kamerabilddaten. Auf Basis von rund 1.200 Stunden Videomaterial wurde eine strukturierte Referenzdatenbasis mit über 17.000 annotierten Verkehrsteilnehmenden aufgebaut. Parallel wurde ein hierarchisches Klassifikationsschema entwickelt, das internationale Datensätze, nationale Standards und projektspezifische Anforderungen integriert.

In AP3 wurde die KI-basierte Methodenentwicklung durchgeführt. Hierbei wurden insbesondere Objektdetektoren der YOLO-Familie (You Only Look Once) implementiert und unter variierenden Bedingungen systematisch getestet. Der Fokus lag auf der Analyse von Einflussfaktoren wie Beleuchtung, Kameraperspektive, Datenqualität und Klassifikationsgranularität.

AP4 diente der Integration der Verkehrsdaten in flächenbezogene Informationen aus der Erdbeobachtung. Ziel war es, den Flächenverbrauch unterschiedlicher Planungsvarianten quantitativ zu erfassen und in einem prototypischen Planungstool zusammenzuführen.

In AP5 wurden die Ergebnisse im Rahmen eines Workshops mit externen Experten sowie im Kontext realer Planungsprojekte bewertet.

### **1.3 Wesentliche Ergebnisse und Zusammenarbeit mit anderen Einrichtungen**

Das Projekt zeigt, dass KI-gestützte Verfahren zur automatisierten Verkehrserfassung unter geeigneten Rahmenbedingungen ein hohes Potenzial besitzen. Insbesondere für häufig vorkommende Verkehrsteilnehmer wie PKW konnten zuverlässige Ergebnisse erzielt werden. Gleichzeitig wurde deutlich, dass die Modelleleistung stark von Faktoren wie Lichtverhältnissen, Kameraposition und Datenqualität abhängt und bei zunehmender Differenzierung der Klassifikation abnimmt.

Ein wesentliches Ergebnis ist der Aufbau einer umfangreichen Referenzdatenbasis sowie die Entwicklung eines hierarchischen Klassifikationsschemas, das eine systematische Analyse unterschiedlicher Klassifikationsebenen ermöglicht. Darüber hinaus konnte durch eine entwickelte Bewertungsmethodik gezeigt werden, wie sich Aufnahmebedingungen auf die Qualität automatisierter Auswertungen auswirken.

Die Integration der Verkehrsdaten in flächenbezogenen Informationen stellt einen weiteren zentralen Beitrag dar. Dadurch wird eine quantitative Bewertung des Flächenverbrauchs unterschiedlicher Planungsvarianten ermöglicht, was einen direkten Mehrwert für eine nachhaltige Verkehrs- und Flächenplanung bietet.

Die Zusammenarbeit zwischen dem Ingenieurbüro Behringer & Partner und dem DLR erwies sich als wesentlicher Erfolgsfaktor, da praxisnahe Anforderungen mit wissenschaftlicher Methodenentwicklung eng verzahnt werden konnten. Der Austausch mit externen Experten und Anwendern, wie z. B. Obermeyer Infrastruktur GmbH & Co. KG, zeigte zudem, dass ein hohes Interesse an automatisierten Verfahren besteht, gleichzeitig jedoch Anforderungen an Robustheit, Transparenz und Praxistauglichkeit weiterhin bestehen. Die Ergebnisse wurden auf der Konferenz REAL CORP 2026 in Wien präsentiert.

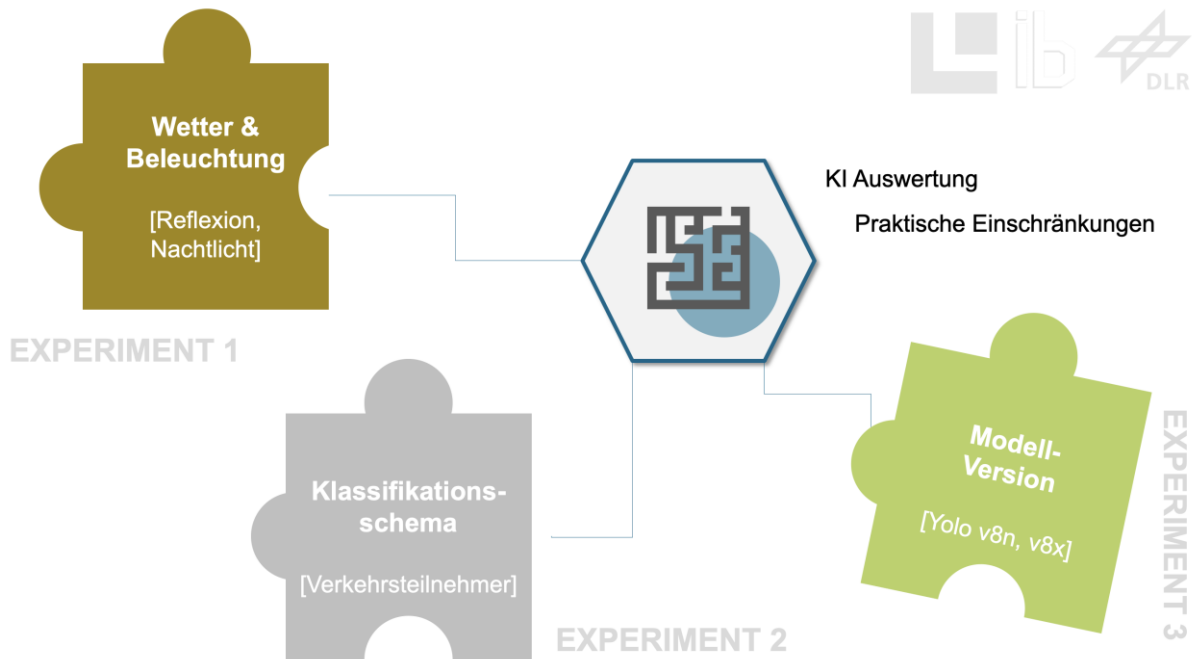
Insgesamt liefert das Projekt eine fundierte Grundlage für die Weiterentwicklung und Anwendung KI-gestützter Methoden in der Verkehrsplanung und trägt zur stärkeren Integration datenbasierter Ansätze in die Planungspraxis bei.

## 2 Ausführliche Darstellung der durchgeführten Arbeiten

Die zuverlässige Erfassung und Differenzierung von Verkehrsströmen ist eine zentrale Grundlage moderner Verkehrs-, Stadt- und Umweltplanung (Matthias et al., 2020; Schulthoff et al., 2022). Insbesondere in städtischen Gebieten mit hohem Verkehrsaufkommen (Al-qaness et al., 2021). Mit der zunehmenden Vielfalt an Verkehrsteilnehmenden stoßen klassische Erhebungsmethoden wie manuelle Zählungen, Induktionsschleifen oder punktuelle Sensorsysteme jedoch an organisatorische, wirtschaftliche und methodische Grenzen. Kamerabasierte Verfahren in Kombination mit Methoden des maschinellen Sehens versprechen hier ein hohes Automatisierungspotenzial (Al-qaness et al., 2021; Galich et al., 2025; Gour & Kanskar, 2019; Stiller et al., 2026; Zhang et al., 2022). Gleichzeitig bestehen erhebliche Unsicherheiten hinsichtlich ihrer praktischen Einsetzbarkeit, insbesondere im Spannungsfeld zwischen international etablierten KI-Modellen, nationalen Planungsstandards und realen Aufnahmebedingungen. Darüber hinaus ermöglichen moderne Daten und Methoden der Erdbeobachtung eine kleinräumig differenzierte Betrachtung und Analyse der Landoberfläche und erlauben es damit, den Flächenverbrauch und etwaige daraus resultierende Landnutzungskonflikte, insbesondere in dicht bebauten Stadtgebieten und deren Umgebung, abzubilden und zu quantifizieren (Shahab et al. 2021, Taubenböck et al. 2024). Ziel des Projekts war es daher, systematisch zu untersuchen, unter welchen Voraussetzungen KI-gestützte Auswerteverfahren einen belastbaren Beitrag zur Verkehrszählung leisten können und wo ihre technischen und methodischen Grenzen liegen.

Im Arbeitspaket AP2 lag der Fokus auf der Schaffung einer belastbaren empirischen Grundlage für die nachfolgende Methodenentwicklung. Hierzu wurden praxisnahe Kamerabilddaten aus realen Verkehrserhebungen akquiriert, inventarisiert und hinsichtlich ihrer Qualität systematisch bewertet. Auf dieser Basis wurde eine umfangreiche Referenzdatenbank aufgebaut, in der insgesamt 17.172 Verkehrsteilnehmende manuell annotiert und gemäß eines hierarchischen Klassifikationsschemas eingeordnet wurden. Dieses Schema verbindet internationale Objektklassen mit den Anforderungen deutscher Verkehrserhebungsstandards sowie projektspezifischen Differenzierungen. AP2 schuf damit nicht nur die notwendige Datenbasis für das Training und die Evaluation von KI-Modellen, sondern machte zugleich zentrale Einflussfaktoren wie Kameraperspektive, Lichtverhältnisse, Bildqualität und Klassenungleichgewichte explizit sichtbar.

Auf dieser Referenzdatenbasis aufbauend, wurde in Arbeitspaket AP3 die Leistungsfähigkeit kamerabasierter KI-Verfahren systematisch untersucht. Im Mittelpunkt standen dabei vortrainierte Objektdetektoren mit YOLO- (You Only Look Once) Architektur (Redmon & Farhadi, 2016; Xia et al., 2024; Zhang et al., 2022), deren Eignung für unterschiedliche Klassifikationstiefen (Forschungsgesellschaft für Straßen- und Verkehrswesen, 2012; Lin et al., 2014) und Aufnahmebedingungen analysiert wurde. Durch ein kontrolliertes experimentelles Design (Abbildung 1) konnten der Einfluss von Beleuchtung, semantischer Differenzierung und Modellarchitektur getrennt bewertet werden. Die Ergebnisse zeigen einen klaren Zielkonflikt zwischen zunehmender semantischer Tiefe und erreichbarer Detektionsgenauigkeit. Gleichzeitig verdeutlichen sie, unter welchen Bedingungen robuste und praxisnahe Ergebnisse erzielt werden können. Unter Idealbedingungen wurde eine mittlere Genauigkeit von 87,8% erreicht. Abschließend wurde die automatisierte Auswertung exemplarisch mit manuellen Verkehrszählungen verglichen, um die Übertragbarkeit in bestehende Planungsprozesse einzuordnen.



**Abbildung 1: Schematische Übersicht des experimentellen Designs. Die drei Experimente untersuchen komplementäre Einflussfaktoren auf die KI-gestützte Auswertung von Kamerabildern: (Experiment 1) den Einfluss von Wetter- und Beleuchtungsbedingungen, (Experiment 2) die Auswirkungen unterschiedlicher semantischer Klassifikationstiefen der Verkehrsteilnehmenden sowie (Experiment 3) den Vergleich verschiedener Modellarchitekturen innerhalb der YOLO-Familie. Die zentrale KI-Auswertung bildet den gemeinsamen methodischen Kern aller Experimente und ermöglicht eine vergleichende Bewertung unter praxisnahen Rahmenbedingungen.**

In Arbeitspaket AP4 wurden räumlich höchst aufgelöste Satellitenbilddaten sowie weitere Geodaten verwendet, um die aktuelle Landbedeckung in einem Planungsgebiet für ein neues Stadtquartier abzuleiten. Auf dieser Grundlage kann der Flächenverbrauch verschiedener Planungsvarianten quantifiziert und auch die Auswirkungen unterschiedlicher Realisierungen der Planung evaluiert werden. Insbesondere hinsichtlich der Wohnbebauung zeigte sich, dass unterschiedliche Arten der Bebauung deutliche Unterschiede hinsichtlich des Flächenverbrauchs aufweisen und dass die im Rahmen des Vorhabens entwickelten Ansätze sehr gut geeignet sind, Landnutzungskonflikte aufzuzeigen und die Diskussion bereits in einer frühen Planungsphase zu bereichern.

Insgesamt zeigt das Projekt, dass KI-gestützte, kamerabasierte Verkehrszählungen weder als schlüsselfertiger Ersatz noch als rein experimentelles Werkzeug zu verstehen sind, sondern als eigenständiges Instrument mit spezifischen Stärken und Grenzen. Durch die enge Verzahnung von Praxisanforderungen, Datenanalyse und methodischer Evaluation konnte ein tiefes Verständnis für die Wechselwirkungen zwischen Datenqualität, Klassifikationsschema und Modellleistung aufgebaut werden. Darüber hinaus konnte der Nutzen von Erdbeobachtungsdaten und daraus abgeleiteter Information zur Abschätzung des Flächenverbrauchs gezeigt werden, um zu einem nachhaltigen Flächenmanagement beizutragen. Das Projekt leistet damit einen wichtigen Beitrag zum Wissenstransfer zwischen Forschung und Planungspraxis und schafft eine fundierte Grundlage für den zukünftigen, verantwortungsvollen Einsatz KI-basierter Verfahren in der Verkehrsplanung.



- **Ebene 1** unterscheidet sieben, aus dem MS-COCO-Datensatz ausgewählte Klassen: *person*, *bicycle*, *motorcycle*, *car*, *truck* und *train*. Durch den hohen Standardisierungsgrad des MS-COCO-Datensatzes lassen sich die erzielten Ergebnisse später in den internationalen Forschungskontext einbetten und vergleichen.
- **Ebene 2** ist den "Empfehlungen für Verkehrserhebungen" (EVE) der FGSV (2012) entnommen. Die Einteilung umfasst neun Grundklassen und wurde um Fußgänger und Fahrradfahrer ergänzt, um den Anforderungen der Verkehrszählung besser gerecht zu werden.
- **Ebene 3** basiert auf den erweiterten, mit den Praxispartnern abgestimmten idealen Anforderungen, um eine präzise und umfassende Verkehrszählung in unterschiedlichen Kontexten zu ermöglichen.

Das hierarchische Klassifikationsschema, das die verschiedenen Ebenen von (1) MS-COCO-Daten, (2) EVE und (3) erweiterte Praxisanforderungen integriert, bildet die Grundlage für Experimente mit maschinellen Lernverfahren. Die Struktur erlaubt es, die Ergebnisse effizient zu vergleichen. Neben einer hohen Praxisrelevanz, trägt sie so zur wissenschaftlichen Entwicklung robuster und anpassungsfähiger Verkehrszählungssysteme bei.

### 2.2.2 Akquise & Vorprozessierung der Kamerabilddaten (Task 2.2)

Für die Entwicklung und Evaluation des KI-basierten Auswerteverfahrens wurden praxisnahe Kamerabilddaten aus zwei vorangegangenen Verkehrsplanungsprojekten der Obermeyer Infrastruktur GmbH & Co. KG herangezogen. In beiden Untersuchungsgebieten stehen umfangreiche infrastrukturelle Maßnahmen an, die voraussichtlich zu einer deutlichen Veränderung des Verkehrsaufkommens führen werden. Entsprechend bestand ein hoher Bedarf an belastbaren Verkehrszählungen, um zukünftige Verkehrsentwicklungen quantifizieren und mit dem Ist-Zustand vergleichen zu können.

Die Untersuchungsgebiete (**Abbildung 3**) umfassen zum einen den Münchner Nordosten, der im Rahmen einer großräumigen Stadterweiterung Platz für bis zu 30.000 neue Einwohnerinnen und Einwohner sowie begleitende Infrastrukturprojekte bieten soll, und zum anderen die Gemeinde Planegg, in der unter anderem der Ausbau des Bahnhofplatzes sowie eine Verlängerung der U-Bahn geplant sind. Über mehrere Tage hinweg wurden in beiden Gebieten insgesamt 51 Verkehrserhebungen durchgeführt. Diese beinhalten sowohl Quer- als auch Längsschnittzählungen an 34 unterschiedlichen Standorten. Neben den Kameradaten liegen für ausgewählte Zeiträume auch manuelle Verkehrszählungen vor, die in der Planungspraxis üblich sind und auf 15-Minuten-Intervalle aggregiert wurden.



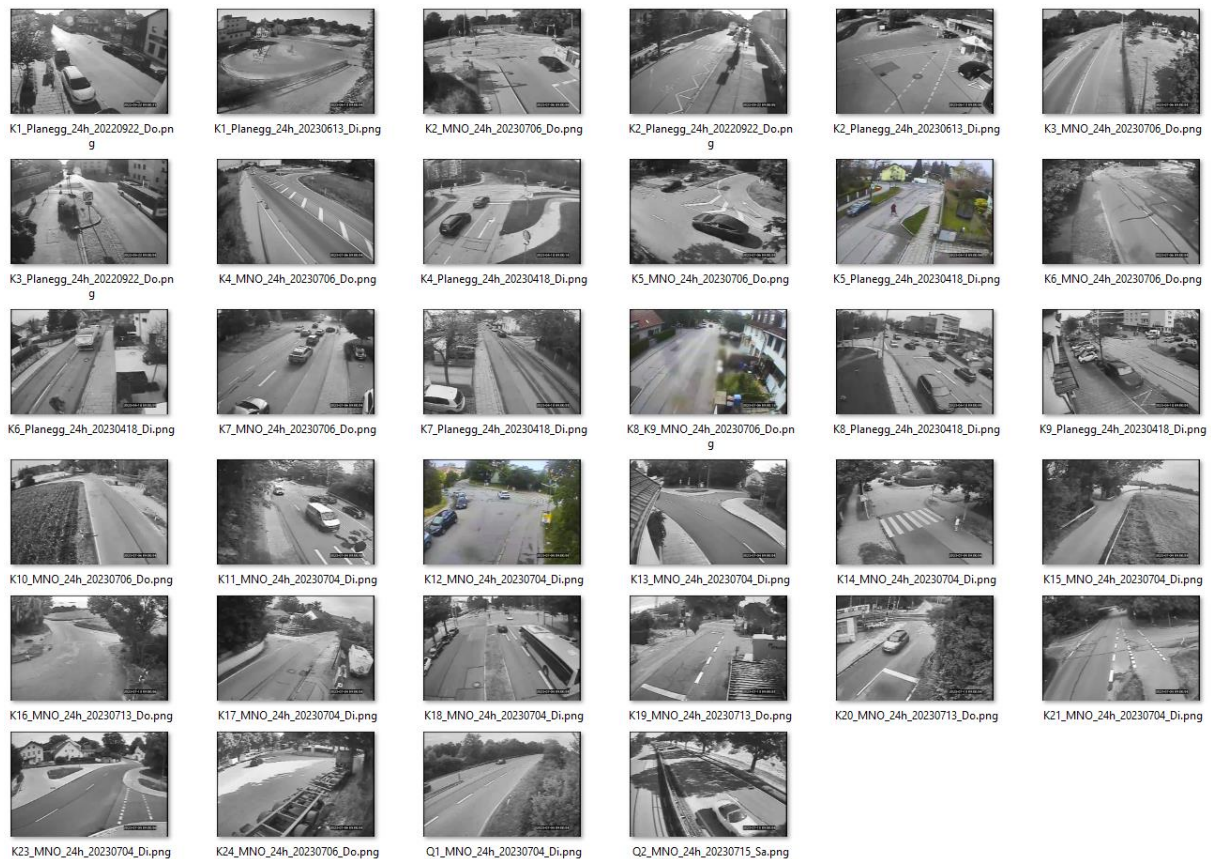


Abbildung 4: Übersicht der verfügbaren Bildausschnitte.

### 2.2.3 Aufbau der Referenzdatenbasis (Task 2.3)

Sowohl für die ad-hoc Evaluation als auch für die Weiterentwicklung der Methodik ist es essenziell, die einzelnen Objekte innerhalb der Aufnahmen präzise zu lokalisieren und zu annotieren. Die Annotation, das sogenannte „Label“, beinhaltet also sowohl die Klassifikation als auch die Lage der Verkehrsteilnehmer im Bild. Ein wichtiger Aspekt ist dabei, sicherzustellen, dass für jede Zielklasse ausreichend Labels vorhanden sind, um eine hohe Klassifikationsgenauigkeit zu gewährleisten. Da das manuelle Labeln hinreichend genau, jedoch sehr zeitaufwendig ist, haben wir den Vorgang teilweise automatisiert. Der Aufbau der Referenzdatenbasis besteht entsprechend aus drei Arbeitsschritten (Abbildung 5).

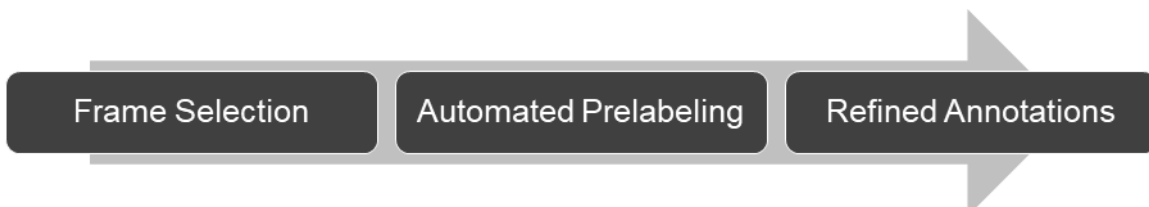


Abbildung 5: Optimierter Prozess zum Aufbau der Referenzdatenbasis.

## **Frame Selection**

Der erste Schritt im Aufbau der Referenzdatenbasis ist die Auswahl relevanter Frames aus den gesammelten Videos. Hierzu wurde aus den ursprünglichen Videos alle 10 Sekunden ein Frame extrahiert. Dieses Maß wurde gewählt, um Duplikate im fließenden Verkehr zu vermeiden und gleichwohl eine repräsentative Stichprobe der Verkehrssituation zu erhalten.

Durch diese Vorgehensweise entstanden 360 Frames pro Stunde und Video. Um einer ungleichen Verteilung der verschiedenen Verkehrsteilnehmerklassen (Stichwort „Class Imbalance“) entgegenzuwirken, wurde die Stichprobenziehung im Verlauf des Projekts zunehmend stratifiziert. Es wurden gezielt Frames ausgewählt, die unterrepräsentierte Klassen – wie bspw. Lastenfahräder und Spezialfahrzeuge - enthalten.

Durch diese präzise und strategische Auswahl an Kamerabilddaten wird eine ausgewogene und repräsentative Datenbasis geschaffen, die sowohl die Vielfalt als auch die Häufigkeit der verschiedenen Verkehrsteilnehmerklassen widerspiegelt.

## **Automated Prelabeling**

Der zweite Schritt im Aufbau der Referenzdatenbasis ist das automatisierte Vor-Labeln der ausgewählten Frames. Dieser Prozess umfasst bereits den ersten Einsatz von Computer-Vision-Algorithmen, um Verkehrsteilnehmer in den ausgewählten Frames zu erkennen und vorläufig zu lokalisieren. Dabei wurde ein vortrainiertes Objekterkennungsmodell (YOLOv8.2) verwendet. Die vorläufigen Labels (entsprechen Ebene 1 aus Task 2.1) wurden in die Datenbasis integriert und dienten als Grundlage für die nachfolgende manuelle Verfeinerung.

Obwohl an dieser Stelle noch keine quantitative Genauigkeitsabschätzung durchgeführt werden konnte, erlaubt dieser voll automatisierte Arbeitsschritt erste Erfahrungen zu sammeln, die nachfolgend bei der Konzeptualisierung der KI-Methodenentwicklung (Task 3.1) eingebracht werden können.

## **Refined Annotations**

Der dritte und abschließende Schritt im Aufbau der Referenzdatenbasis ist die manuelle Überprüfung der automatisierten Lokalisation im Allgemeinen, sowie die händische Annotation der Verkehrsteilnehmer auf der hierarchischen Klassifikationsebene 3 (vgl. Task 2.1) im Speziellen. Dieser Prozess stellt sicher, dass die Labels sowohl in ihrer Lage als auch in ihrer Art eine hohe Genauigkeit aufweisen. Insbesondere bei komplexen Szenen oder schwer erkennbaren Objekten waren Korrektur und Verfeinerungen notwendig.

Die studentischen Annotatoren wurden intensiv geschult, um einheitliche Standards bei der Label-Erstellung zu gewährleisten. Dazu gehörte die klare Definition der zu labelnden Klassen sowie spezifische Richtlinien zur Positionierung der Labels im Bild. Um eine hohe Qualität und Konsistenz der Annotationen zu gewährleisten, fanden stichprobenartige Überprüfungen durch zusätzliche Annotatoren statt. Ein spezielles Versionierungswerkzeug erlaubt es außerdem, Änderungen im Referenzdatensatz nachzuvollziehen und die Datenintegrität zu jedem Zeitpunkt sicherzustellen.

## **Management der Referenzdatenbasis**

Um die gesammelten Daten möglichst effizient zu verwalten, wird auf das open-source Werkzeug FiftyOne zurückgegriffen. Es erlaubt, alle 3.924 annotierten Bilder in einer Bibliothek zusammenzufassen und zu indexieren. Mithilfe der Filterfunktionen, können sowohl gezielt Klassenannotationen inspiziert werden, als auch auf Grundlage der Metadaten, wie Standort, Aufnahmezeitpunkt, Wetter und technischer Einschränkungen, Referenzdaten für die darauf aufbauenden Experimente in Task 3.2 kuratierter werden.

Darüber hinaus erlaubt die offene Python-Bibliothek eigene Funktionen zu integrieren, beispielsweise um Klassen über das hierarchische Klassifikationsschema hinweg zu transponieren („Mapping“). Ein besonderer methodischer Sonderfall ergibt sich bei Verkehrsteilnehmenden, die visuell aus mehreren Objekten bestehen, beispielsweise Radfahrenden. Wie Abbildung 6 demonstriert, erfolgt die Zusammenführung im Zuge eines kontextualisierten Mappings hier regelbasiert, indem die räumliche Überlappung der Bounding Boxes (Intersection over Union) zwischen Person und Fahrrad berechnet und bei Überschreiten eines definierten Schwellenwerts eine Fusion zu einer gemeinsamen Verkehrsteilnehmerklasse vorgenommen wird.



**Abbildung 6:** Darstellung der kontextbasierten Fusion von Objektklassen. Links zeigen farbige Bounding Boxes (Rechtecke zur Lokalisierung erkannter Objekte) eine Person (blau) und ein Fahrrad (rot). Rechts ist das Ergebnis der Fusion als einzelner Verkehrsteilnehmer Radler (violett) dargestellt.

#### **2.2.4 Akquise von Flächendaten (u.a. Satellitenbilder, amtliche Flächendaten) (Task 2.4)**

Im Rahmen des Vorhabens wurden verschiedene Satellitendaten sowie weitere Geodaten beschafft. Die zentrale Datengrundlage stellt eine räumlich höchstauflöste Satellitenbilddaufnahme des Sensors WorldView-3 mit 30 cm Bodenauflösung dar (Abbildung 7). Die Szene wurde am 04.07.2019 unter wolkenfreien Bedingungen mit acht Spektralkanälen aufgenommen. Höheninformationen werden von der Stadt München durch ein normalisiertes digitales Oberflächenmodell (nDSM) mit einer räumlichen Auflösung von 50 cm bereitgestellt. Zusätzlich werden Daten aus Open Street Map (OSM) verwendet, die denselben Zeitraum abdecken. Gebäudegeometrien werden aus dem Open Data Portal des Freistaats Bayern bezogen.



**Abbildung 7: WorldView-3 Aufnahme vom 04.07.2019. Überblick über das Stadtgebiet Münchens (a) und Detailausschnitt des Untersuchungsgebiets "Münchner Nordosten" (b)**

## **2.3 KI-Methodenentwicklung zur automatisierten Auswertung von Kamerabildern (AP3.0)**

### **2.3.1 Konzeptualisierung der KI-Methodenentwicklung (Task 3.1)**

Im großen Forschungsfeld zu Künstlicher Intelligenz grenzt Computer Vision speziell den Bereich ab, der sich mit der automatisierten Interpretation visueller Informationen befasst. Objekterkennung und -klassifikation haben sich Convolutional Neural Networks (CNNs) etabliert (LeCun et al., 2010). Für kamerabasierte Verkehrserhebungen sind dabei insbesondere die Modellarchitekturen relevant, die sowohl eine präzise Lokalisierung als auch eine robuste Klassifikation von Verkehrsteilnehmenden ermöglichen. Vor diesem Hintergrund wurde im Projekt auf Modelle der YOLO-Familie zurückgegriffen (Redmon et al., 2016), welche sich bereits im automatisierten Pre-Labeling als hinreichend robust erwiesen hatten und grundsätzlich auch für Echtzeitanwendungen geeignet sind.

Ein zentrales konzeptionelles Ziel war die realitätsnahe Bewertung von KI-Methoden. In der Praxis sind vortrainierte Modelle üblich, welche sich sowohl out-of-the-box anwenden als auch individuell anpassen lassen. Wir griffen auf die zu Projektbeginn aktuellste Version zu. Die Referenzdatenbasis aus AP2 wurde für alle Experimente eine einheitliche Datenaufteilung von 50 % Training, 20 % Validierung und 30 % Test definiert.

Zu der realitätsnahen Bewertung gehören außerdem, die beiden in AP2 herausgearbeiteten Einflussfaktoren Bildqualität und Semantische Tiefe, sowie die Ressourceneinsatz und Echtzeittauglichkeit. Letztere hängt von der verwendeten Modellgröße ab. Um kausale Zusammenhänge zwischen Modellkonfiguration, Datencharakteristika und Detektionsleistung nachvollziehbar zu machen, haben wir ein Studiendesign gewählt, in dem die drei Parameter isoliert variiert werden.

### 2.3.2 Entwicklung, Training & Validierung des KI-Verfahrens (Task 3.2)

#### Experiment 1: Variable Lichtverhältnisse

Optische Systeme sind in der Planungspraxis zwangsläufig mit wechselnden Licht- und Sichtverhältnissen konfrontiert, die sich der direkten Kontrolle entziehen und die visuelle Erkennbarkeit von Verkehrsteilnehmenden erheblich beeinflussen können (Staab et al., 2021; Xia et al., 2024; Zhang et al., 2022).

Zur isolierten Analyse dieses Einflusses wurde ein auf der semantischen Ebene 1 vortrainiertes YOLOv8n-Modell eingesetzt. Modellarchitektur und Klassifikationsschema wurden konstant gehalten, sodass beobachtete Leistungsunterschiede ausschließlich auf variierende Aufnahmebedingungen zurückzuführen sind. Als Stratifizierungsmerkmal dienten die bereits in Task 2.2 systematisch erfassten Beleuchtungskategorien sonnig, bedeckt und dunkel (Abbildung 8). Die Modellleistung wurde anhand von Precision, Recall und dem daraus abgeleiteten F1-Score bewertet, der sich insbesondere bei unausgeglichene Klassenverteilungen als robuste Kenngröße etabliert hat (Powers, 2011; Sokolova & Lapalme, 2009).



**Abbildung 8: Drei beispielhafte Aufnahmen illustrieren die unterschiedlichen Bildkompositionen an Standorten. Außerdem zeigen sie unterschiedliche Beleuchtungssituationen: a) sonnig, b) bedeckt und c) dunkel.**

Über alle 3.924 Bilder hinweg wurde out-of-the-box ein gewichteter F1-Score von 0,48 erreicht, resultierend aus einer hohen Precision von 0,79 bei gleichzeitig niedriger Sensitivität von 0,35. Unter guten Lichtverhältnissen (sonnig) erreicht das Modell einen etwas höheren F1-Score von 0,51. Bei moderaten Beleuchtungsbedingungen (bedeckt) sinkt die Erkennungsleistung wieder auf 0,48. Deutlich stärker fällt der Leistungsabfall bei dunklen Aufnahmebedingungen aus, bei denen lediglich ein F1-Score von 0,30 erzielt wird.

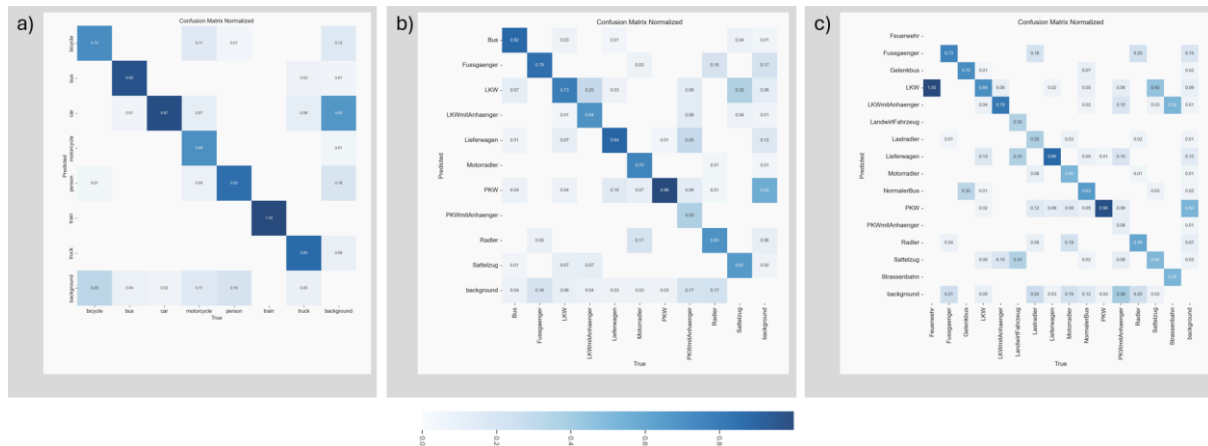
Das Experiment zeigt, dass insbesondere dunkle Aufnahmebedingungen zu erheblichen Verzerrungen von Verkehrszählungen führen können. Für Einsatzszenarien in der dunklen Jahreszeit deutet dies auf die Notwendigkeit zusätzlicher Maßnahmen hin, etwa durch infrarot- oder thermalsensitive Bildgebung oder durch die Kombination mit aktiven Sensorsystemen.

## **Experiment 2: Semantische Klassifikationstiefe**

In der Verkehrserhebungspraxis besteht häufig der Wunsch nach einer möglichst feingranularen Differenzierung von Verkehrsteilnehmenden, da diese eine detailliertere Analyse des Verkehrsaufkommens und eine gezieltere Planung ermöglicht. Gleichzeitig steigen mit zunehmender semantischer Tiefe die Anforderungen an Datenbasis, Annotation und Modellrobustheit, sodass ein Zielkonflikt zwischen inhaltlicher Differenzierung und verlässlicher automatisierter Erkennung entsteht.

Zur systematischen Untersuchung dieses Zielkonflikts wurde die Modelleistung für unterschiedliche semantische Klassifikationsebenen evaluiert. Hierzu wurde ein YOLO-Modell projektspezifisch angepasst, wobei über alle drei hierarchischen Klassifikationsebenen aus Task 2.1 die respektiven Trainingsparameter konstant gehalten wurden. Auf diese Weise lassen sich Unterschiede in der Detektionsleistung eindeutig auf die semantische Granularität zurückführen.

Die Analyse zeigt einen klaren und systematischen Zielkonflikt zwischen semantischer Differenzierung und Detektionsleistung. Bei konstanter Modellarchitektur (YOLOv8n), identischer Datengrundlage und unveränderten Trainingsparametern nimmt die Modellgüte mit zunehmender semantischer Tiefe deutlich ab. Für die generische Klassifikationsebene (Ebene 1, sieben Klassen) wird mit einem mAP@50 von 0,863 eine hohe Detektionsleistung erreicht, was sich in einer stark ausgeprägten Diagonalen der Konfusionsmatrix und nur vereinzelt Fehlklassifikationen widerspiegelt (Abbildung 9a). Bei der FGSV-nahen Klassifikation (Ebene 2, zehn Klassen) sinkt der mAP@50 auf 0,780. Gleichzeitig treten vermehrt systematische Verwechslungen zwischen semantisch und visuell ähnlichen Klassen auf – vor allem zwischen Fahrzeugen mit und ohne Anhänger oder zwischen unterschiedlichen Nutzfahrzeugtypen (Abbildung 9b). Trotz dieses Leistungsrückgangs bleibt die überwiegende Mehrzahl der Zuordnungen korrekt, was auf eine grundsätzlich stabile Modelleistung auch bei erhöhter semantischer Komplexität hinweist. Die Klassifikation mit der höchsten Differenzierung (Ebene 3, sechzehn vorkommende Klassen) führt zu einem weiteren deutlichen Leistungsabfall auf einen mAP@50 von 0,647. Die Streuung in der Konfusionsmatrix (Abbildung 9c) weist Fehlzugeordnungen insbesondere zwischen strukturell ähnlichen Klassen aus. Einzelne oder visuell atypische Fahrzeugtypen, wie etwa ein Feuerwehrfahrzeug, werden in diesen Fällen häufig allgemeineren Klassen wie LKW zugeordnet.



**Abbildung 9: Normalisierte Konfusionsmatrizen für drei semantische Klassifikationsebenen: a) generische Klassifikation (Ebene 1), b) FGSV-nahe Klassifikation (Ebene 2) und c) projektspezifische Klassifikation (Ebene 3). Die Diagonale zeigt korrekte Zuordnungen, Einträge außerhalb der Diagonale kennzeichnen systematische Fehlklassifikationen zwischen Klassen. Die Farbskala repräsentiert spalten-normalisierte Anteile im Wertebereich von 0 bis 1.**

Das Experiment verdeutlicht, dass eine hohe semantische Differenzierung aus planerischer Sicht zwar wünschenswert ist, jedoch erhebliche Anforderungen an Umfang, Ausgewogenheit und Qualität der Referenzdaten stellt. Einschränkend ist hierbei zu berücksichtigen, dass drei der im Zielkatalog vorgesehenen Klassen im verwendeten Referenzdatensatz nicht vorkamen, wodurch sich die effektive Klassenanzahl reduzierte und bestehende Klassenungleichgewichte weiter verstärkt wurden. Diese Class-Imbalance-Effekte tragen zusätzlich zur beobachteten Abnahme der Modellgüte bei zunehmender semantischer Tiefe bei. Für praxisnahe Anwendungen erscheint es daher sinnvoll, die Klassifikationstiefe kontextabhängig zu wählen und nicht pauschal eine maximale semantische Differenzierung anzustreben.

### Experiment 3: Modell Varianten

Für den Einsatz kamerabasierter Verkehrszählungen im Feld ist eine möglichst vollständige Erfassung aller Verkehrsteilnehmenden wünschenswert. Vor diesem Hintergrund stellt sich die Frage, ob sich die in den vorherigen Experimenten beobachteten Genauigkeitsverluste technisch durch den Einsatz leistungsfähigerer Modellarchitekturen abbildern lassen.

Zur Untersuchung dieses Aspekts wurden zwei Modellvarianten derselben Modellgeneration miteinander verglichen: das kompakte, für Echtzeitanwendungen optimierte YOLOv8n-Modell mit rund 3 Millionen trainierbaren Parametern sowie das deutlich komplexere YOLOv8x-Modell mit etwa 68 Millionen Parametern. Die Trainingsdauer wurde architektur-spezifisch gewählt, um jeweils eine stabile Konvergenz zu gewährleisten. Datengrundlage und semantisches Klassifikationsschema (Ebene 1) wurden über beide Trainingsläufe hinweg konstant gehalten, sodass Leistungsunterschiede eindeutig auf die Modellarchitektur zurückzuführen sind. Die Bewertung erfolgte anhand einer klassenweisen Analyse der Detektionsleistung (mAP@50), sowie dem Delta zwischen den Modellvarianten.

Der Vergleich (Tabelle 1) zeigt, dass der Einsatz einer größeren Modellarchitektur zu einer leichten, systematischen Verbesserung der aggregierten Detektionsleistung führt. Der mittlere mAP@50 steigt von 0,863 (YOLOv8n) auf 0,878 (YOLOv8x), was einem Zuwachs von +0,015 entspricht. Eine differenzierte Betrachtung der Klassen macht deutlich, dass insbesondere

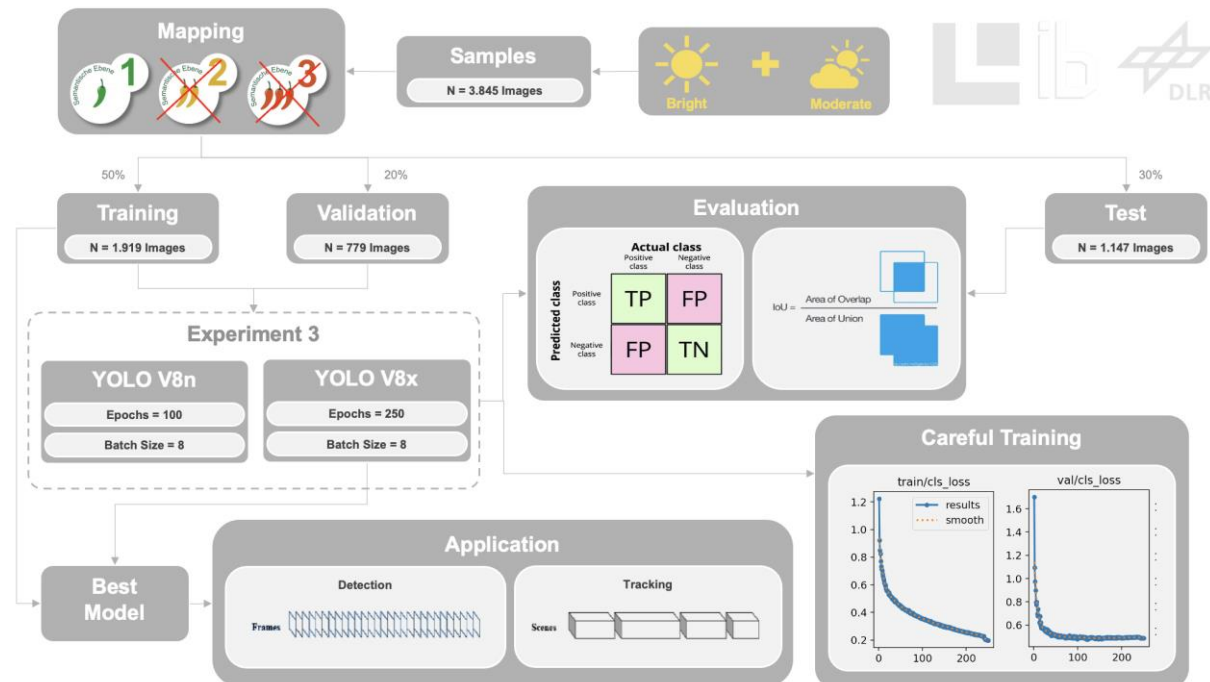
kleinere oder strukturell komplexere Objekte wie motorcycle, bicycle und person von der höheren Modellkapazität profitieren. Für größere und visuell klar abgegrenzte Klassen wie car, bus und truck lassen sich hingegen keine konsistenten Leistungssteigerungen beobachten; in einzelnen Fällen fällt die Detektionsleistung des größeren Modells sogar geringfügig niedriger aus.

**Tabelle 1: Klassenweise Detektionsleistung für zwei unterschiedlich komplexe Modellarchitekturen bei identischer Datengrundlage und generischer Klassifikation (Ebene 1). Angegeben sind die mittleren Genauigkeiten pro Klasse sowie die Differenz (v8x – v8n).**

Semantische Klasse	mAP@50		Δ
	YOLO v8n	YOLO v8x	
train	0.995	0.995	±0.000
car	0.983	0.978	-0.005
bus	0.943	0.928	-0.015
truck	0.912	0.870	-0.042
person	0.853	0.868	+0.015
motorcycle	0.713	0.832	+0.119
bicycle	0.643	0.675	+0.032
<b>arithmetisches Mittel</b>	<b>0.863</b>	<b>0.878</b>	<b>+0.015</b>

### 2.3.3 Testung und Übertragbarkeit des KI-Verfahrens (Task 3.3)

Auf Basis der vorangegangenen Experimente wurden die aussichtsreichste Modellkonfiguration in eine praxisnahe Anwendung überführt. Einen vollständigen Überblick über den integrierten Evaluations- und Entscheidungsworkflow, der die in Task 3.2 gewonnenen Erkenntnisse systematisch zusammenführt, gibt Abbildung 10.



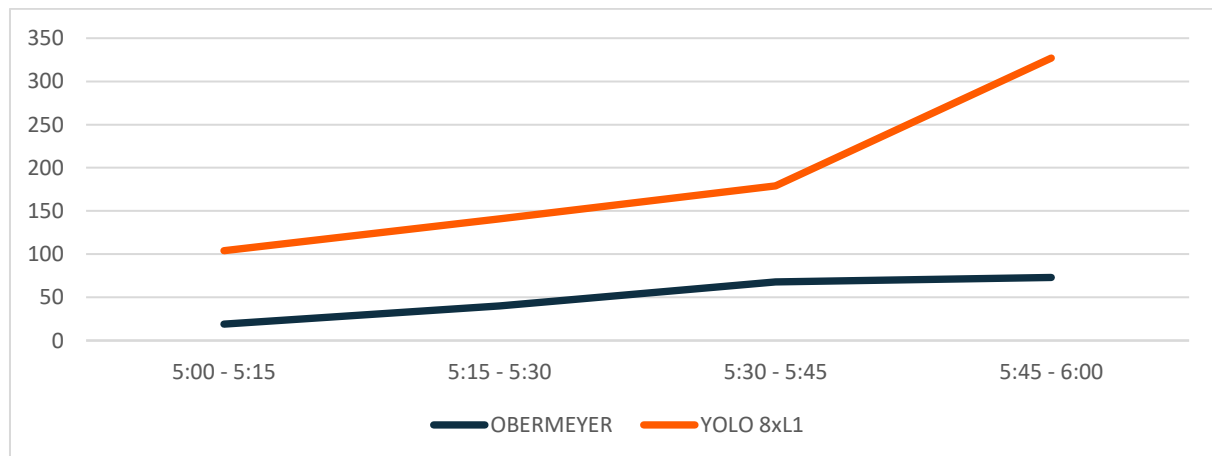
**Abbildung 10: Übersicht des integrierten Workflows zur Auswahl der praxisrelevanten KI-Methode. Dargestellt sind aussichtsreichsten Aufnahmebedingungen, ein einfaches semantisches Klassifikationsschema und die leistungstärkste Modellarchitektur, sowie die darauf aufbauenden Trainings-,**

**Validierungs- und Evaluationsschritte. Das optimierte Modell wird anschließend auf unbekannte Videodaten angewendet.**

Für die abschließende Bewertung ist entscheidend, wie belastbar die KI-gestützten Ergebnisse im Vergleich zu etablierten manuellen Verkehrszählungen sind. Als Referenz dient eine in der Planungspraxis übliche manuelle Zählung, bei der geschulte Mitarbeitende vier aufeinanderfolgende 15-minütige Videosequenzen auswerten und alle Fahrzeuge im fließenden Verkehr erfassen. Im Fokus steht ein einstündiges Zeitfenster von 5:00 bis 6:00 Uhr morgens, für das sich unabhängig von Fahrtrichtung und Fahrzeugtyp eine Gesamtzahl von 200 Fahrzeugen ergibt.

Da in der vorliegenden Anwendung keine explizite Objektverfolgung eingesetzt wird, basiert die KI-gestützte Auswertung auf der Detektion von Verkehrsteilnehmenden in einzelnen Bildframes. Um Mehrfachersfassungen zu begrenzen und eine zeitlich konsistente Aggregation zu ermöglichen, wurden aus den vier Videosequenzen in einem festen Intervall von zehn Sekunden Einzelbilder extrahiert. Insgesamt wurden 585 Fahrzeuge detektiert.

Obgleich die absoluten Werte divergieren, zeigt der Vergleich der resultierenden Zeitreihen (Abbildung 11), dass sich beide Verfahren hinsichtlich der zeitlichen Verkehrsdynamik grundsätzlich ähnlich verhalten. Ab etwa 5:00 Uhr steigt die Anzahl der erfassten Fahrzeuge in beiden Ansätzen kontinuierlich bis 6:00 Uhr an. Während die Verläufe in den frühen Zeitintervallen nahezu parallel verlaufen, weichen die automatisierten Detektionen in der letzten Viertelstunde deutlich nach oben ab. Dieser Effekt ist darauf zurückzuführen, dass die manuelle Zählung ausschließlich den fließenden Verkehr berücksichtigt, während die KI-basierte Methode auch ruhende sowie an Lichtsignalen wartende Verkehrsteilnehmende in konsekutiven Einzelbildern mehrfach erfasst. Abbildung 12 illustriert diesem Effekt exemplarisch.



**Abbildung 11: Vergleich der zeitlich aggregierten Verkehrszählungen im Zeitfenster von 5:00 bis 6:00 Uhr. Dargestellt sind die manuelle Zählung des fließenden Verkehrs (schwarz) und die KI-gestützte Auswertung auf Basis frameweiser Detektionen (orange). Trotz abweichender absoluter Zählwerte zeigt sich eine hohe Übereinstimmung der zeitlichen Verkehrsdynamik.**



**Abbildung 12: Detektionsergebnisse für vier aufeinanderfolgende Frames. Dargestellt sind die Originalaufnahmen (Graustufen) mit überlagerten, farbigen Bounding Boxes der KI-Detektion. (a–b) Ein Bus hält an der Ampel, (c) lässt eine Person die Straße überqueren und (d) setzt seine Fahrt fort.**

Die beobachteten Abweichungen spiegeln damit weniger eine Fehlleistung des KI-Modells wider als vielmehr grundlegende, systemimmanente Unterschiede in der Zähllogik beider Verfahren. Trotz der Unterschiede auf absoluter Ebene zeigt der berechnete Korrelationskoeffizient von  $r = 0,82$  eine hohe Übereinstimmung der relativen zeitlichen Dynamik. Dies unterstreicht, dass die KI-gestützte Auswertung insbesondere für die Analyse zeitlicher Muster und relativer Veränderungen des Verkehrsaufkommens geeignet ist, auch wenn absolute Zählwerte aufgrund konzeptioneller Unterschiede nur eingeschränkt vergleichbar sind.

Insgesamt zeigen die Ergebnisse, dass kamerabasierte KI-Verfahren unter einen belastbaren Beitrag zur Verkehrszählung leisten können, sofern ihre spezifischen Eigenschaften und Grenzen berücksichtigt werden. Die Auswertung folgt dabei einer anderen Logik als manuelle Zählungen, erlaubt jedoch eine konsistente Analyse zeitlicher Dynamiken und relativer Veränderungen des Verkehrsaufkommens. Die Ergebnisse unterstreichen zugleich die hohen Anforderungen an Datenqualität, Klassifikationskonzepte und methodische Transparenz. Richtig eingesetzt ermöglichen KI-gestützte Verfahren eine kosteneffiziente, großflächige und langfristige Erfassung von Verkehrsdaten und stellen damit eine sinnvolle Ergänzung etablierter Erhebungsmethoden dar.

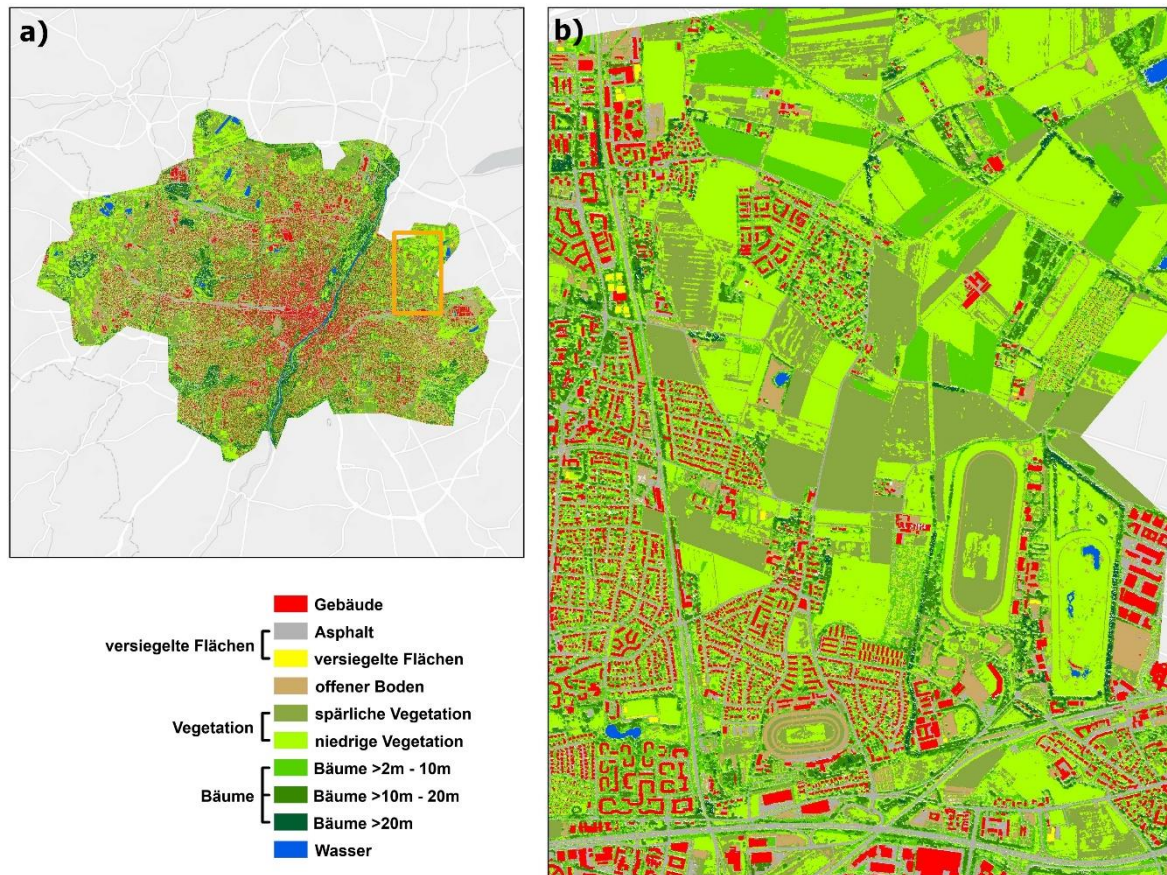
Zukünftige Studien sollten insbesondere die Integration expliziter Objektverfolgung untersuchen, um Mehrfacherfassungen zu reduzieren und absolute Verkehrskennwerte robuster ableiten zu können. Ergänzend bietet die Kombination kamerabasierter Verfahren mit weiteren Sensorsystemen sowie der Einsatz echtzeitfähiger Modellarchitekturen (wie YOLO v8n) im Feldeinsatz Potenzial, die Erfassungsrobustheit insbesondere für kleine oder weit entfernte Verkehrsteilnehmende weiter zu verbessern.

## **2.4 Integration der Methodik für ein nachhaltiges Flächenmanagement (AP4.0)**

### **2.4.1 Erfassung der Landbedeckung und Landnutzung mittels Erdbeobachtungsdaten (Task 4.1)**

Der objektbasierte Klassifikationsansatz für dieses Vorhaben wird nach dem Prinzip der Arbeiten von Taubenböck et al. (2010) und Wurm et al. (2011) umgesetzt. Dabei werden die WorldView-3 Bilddaten, das nDSM, OSM-Daten sowie Gebäudegeometrien in einem wissens- und regelbasierten Ansatz kombiniert, um folgende Klassen abzuleiten: Gebäude, versiegelte Flächen, offener Boden, Bäume, niedrige Vegetation und Wasser. Abbildung 13 zeigt das Ergebnis der objektbasierten Klassifizierung der Landbedeckung auf Basis höchstauflösender Erdbeobachtungsdaten. Neben der am 04.07.2019 aufgenommenen WorldView-3 Szene

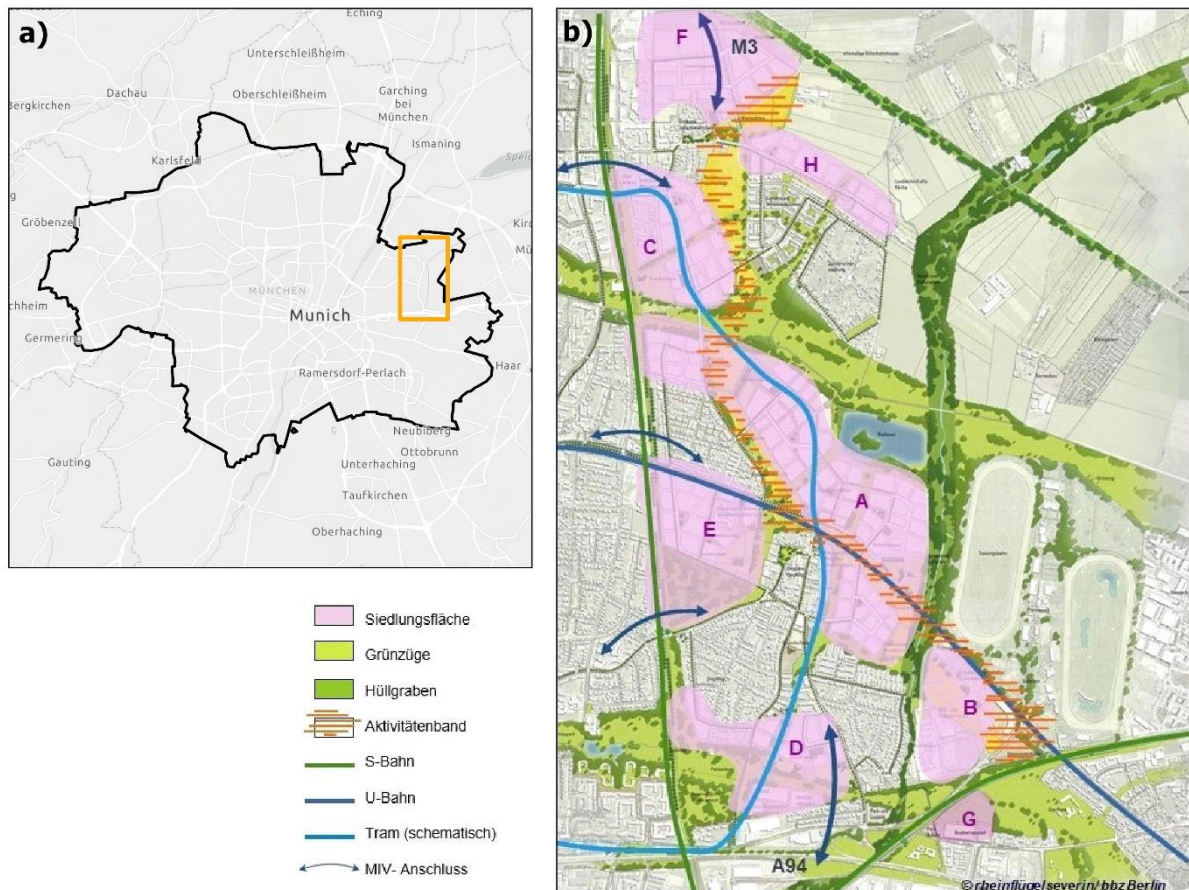
werden bei diesem Ansatz auch das höchst aufgelöste nDSM, OSM-Daten und Gebäudegeometrien berücksichtigt. Die Genauigkeitsabschätzung auf Grundlage einer visuellen Interpretation zufällig verteilter Punkte innerhalb des Untersuchungsgebiets ergab eine Gesamtgenauigkeit von 88 %. Für die weitere Analyse in dieser Arbeit werden die ursprünglichen Landbedeckungsklassen wie in Abbildung 13 dargestellt aggregiert.



**Abbildung 13: Ergebnis der Landbedeckungsklassifikation auf Basis höchst aufgelöster Erdbeobachtungsdaten. Überblick über das Stadtgebiet Münchens (a) und Detailausschnitt des Untersuchungsgebiets "Münchner Nordosten" (b).**

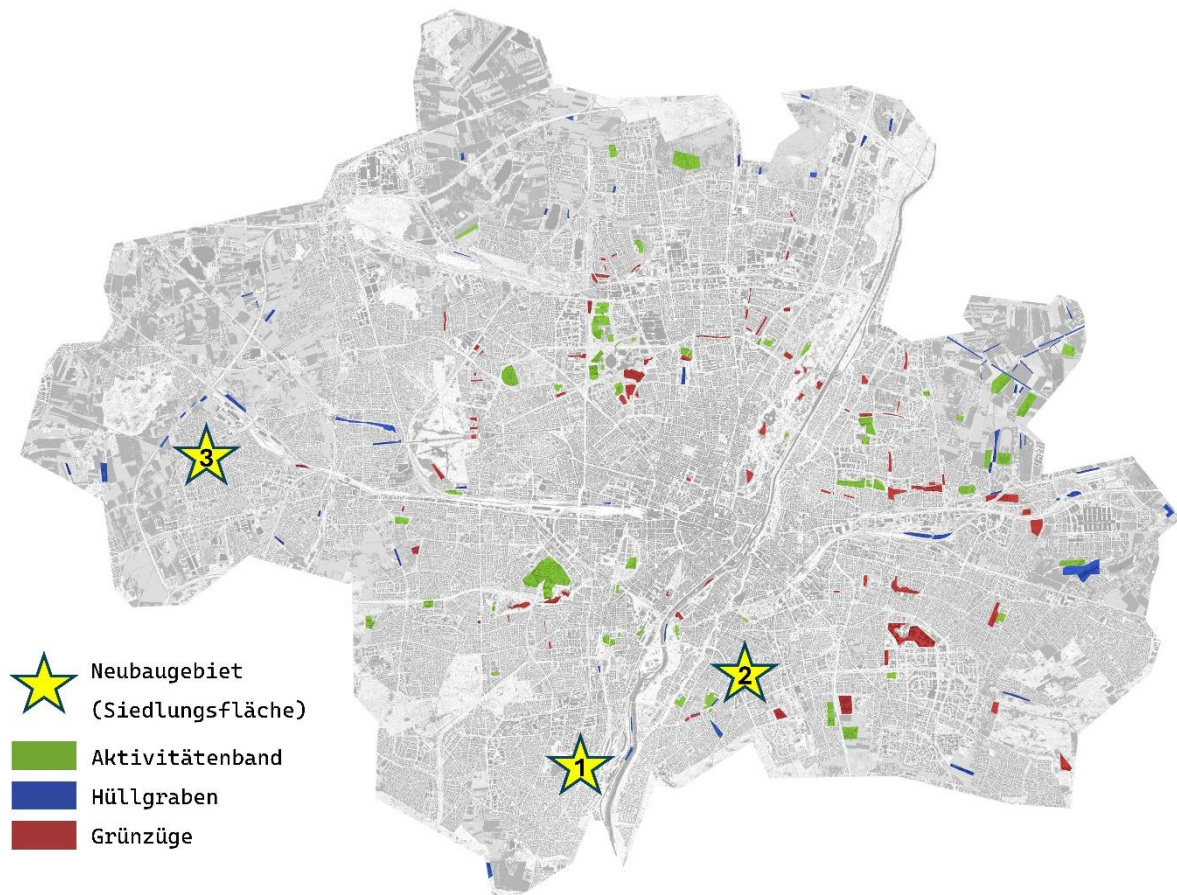
#### **2.4.2 Evaluierung verschiedener Planungsvarianten zur Quantifizierung des Flächenverbrauchs (Task 4.2)**

Zur Evaluierung verschiedener Planungsvarianten wird zunächst der aktuell verfügbare Planungsentwurf des Untersuchungsgebiets „Münchener Nordosten“ herangezogen (Abbildung 14). Dabei handelt es sich um ein großes Stadtentwicklungsprojekt der Stadt München, dessen Ziel es ist, ein völlig neues Stadtviertel für bis zu 30.000 neue Einwohner zu schaffen. Der Entwurf grenzt Siedlungsflächen von Grünstreifen sowie dem Hüllgraben und einem Aktivitätenband ab.



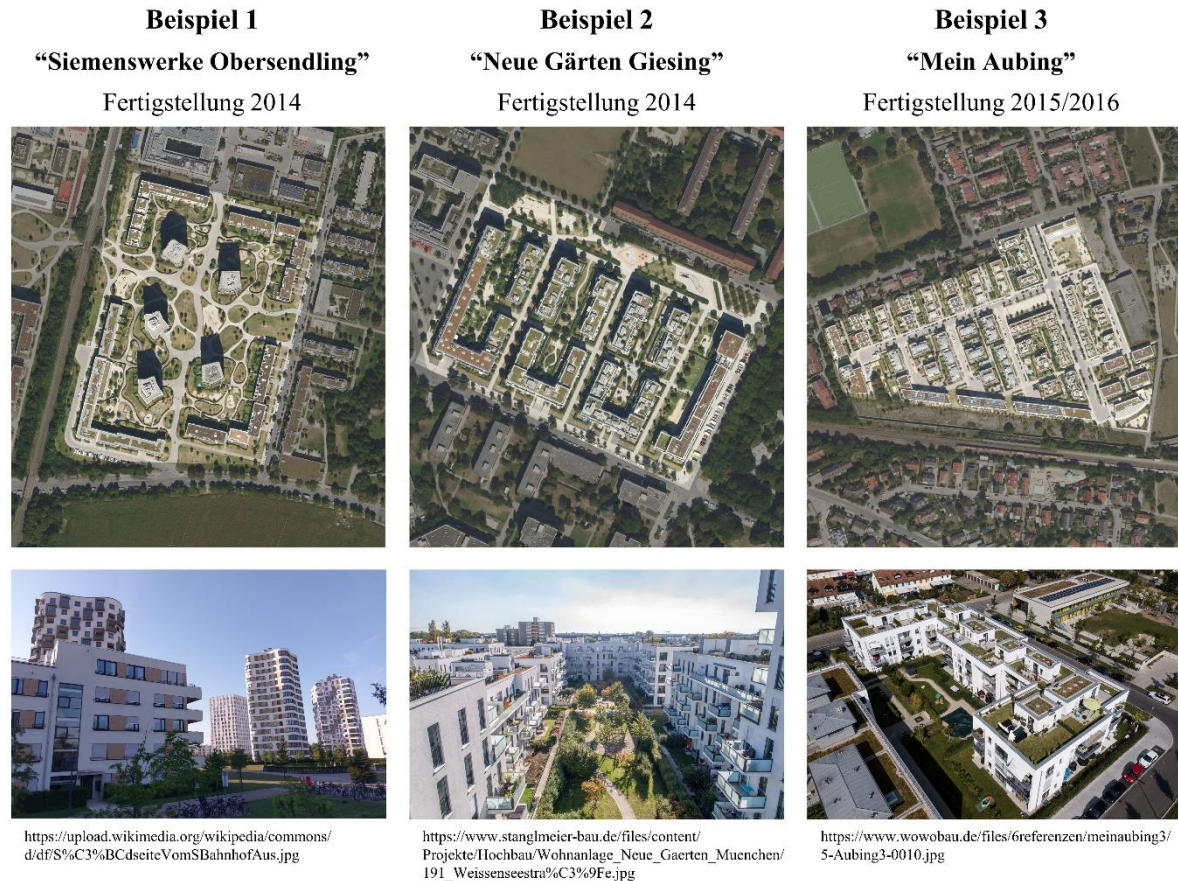
**Abbildung 14: Lage des Untersuchungsgebiets "Münchener Nordosten" im Stadtgebiet Münchens (a) und Planungsentwurf des Gebiets (b)**

Da die konkrete Realisierung der verschiedenen Teilbereiche im Planungsgebiet aktuell noch nicht bekannt bzw. erst ermittelt werden muss, werden möglichst repräsentative Flächen im gesamten Stadtgebiet der Stadt München gesammelt. Für Siedlungsflächen werden drei Neubaugebiete in unterschiedlichen Lagen im Stadtgebiet, die innerhalb der letzten 10 Jahre fertiggestellt worden sind, ausgewählt. Für Freiflächen (Grünzüge, Hüllgraben, Aktivitätenband) werden unabhängig vom Datum der Fertigstellung Vergleichsflächen gesammelt. Die Verteilung der Flächen zeigt Abbildung 15.



**Abbildung 15: Lage der Vergleichsflächen im Stadtgebiet Münchens**

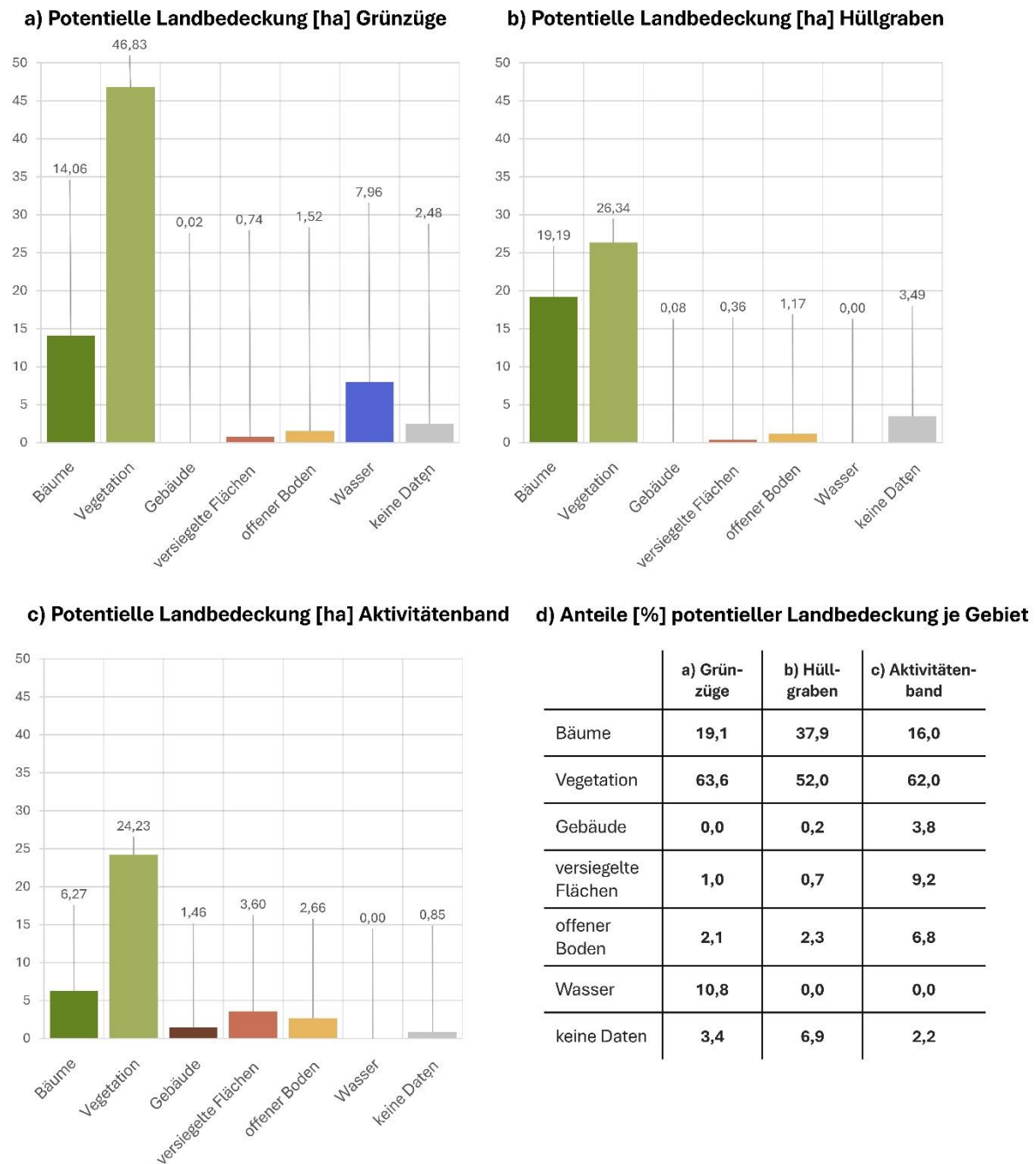
Die drei beispielhaften Neubaugebiete in München sind in Abbildung 16 dargestellt. Das erste Beispiel „Siemenswerke Obersendling“ ist eine neue Wohnsiedlung im Süden Münchens mit fünf Hochhäusern mit jeweils 15 Stockwerken, die von vier- bis fünfstöckigen Randgebäuden umgeben sind. Diese Wohnsiedlung wurde 2014 fertiggestellt und zeichnet sich durch eine autofreie Gestaltung aus, wobei die Zufahrt und die Parkplätze über die umliegenden Straßen erreichbar sind. Das zweite Beispiel „Neue Gärten Giesing“ befindet sich südöstlich des Stadtzentrums und stellt eine relativ zentrale Lage innerhalb Münchens dar, welches 2014 fertiggestellt wurde. Das Gebiet ist autofrei gestaltet und zeichnet sich durch eine dichte Wohnbebauung mit vier bis sieben Stockwerken aus. Das dritte Beispiel „Mein Aubing“ im Westen Münchens ist in konventioneller Bauweise (d.h. Erschließung durch motorisierten Individualverkehr) mit zwei- bis dreistöckigen Gebäuden gestaltet. Die Lage dieser Siedlung kann als traditionelle Vorstadtumgebung charakterisiert werden; ihre Errichtung wurde 2015/16 abgeschlossen.



**Abbildung 16: Drei beispielhafte Neubaugebiete in München**

### 2.4.3 Prototypische Entwicklung eines Planungstools (Task 4.3)

Für die Abschätzung der Auswirkungen der Planung auf Freiflächen wird angenommen, dass hohe (>20 m) und mittelhohe (≥10 m–20 m) Bäume erhalten bleiben, während andere Bereiche entsprechend der Verteilung der Landbedeckung in den Vergleichsflächen verändert werden. Die daraus resultierende potentielle Landbedeckung ist für Grünzüge, Hüllgraben und das Aktivitätenband in Abbildung 17 dargestellt. Insgesamt machen Vegetation und Bäume den größten Teil der Landbedeckung in allen drei Arten von Freiflächen aus. In Grünzügen (Abbildung 17a und d) entfallen 63,6 % bzw. 19,1 % der Bodenbedeckung auf Vegetation und Bäume. Darüber hinaus macht Wasser 10,8 % der Gesamtfläche aus, die durch den Erholungssee in der Mitte des Planungsgebiets (Abbildung 14b) einbezogen wird. Im Hüllgraben (Abbildung 17b und d) sind Vegetation und Bäume mit 52,0 % bzw. 37,9 % gleichmäßiger verteilt. Wie zu erwarten, können diese Gebiete als die natürlichsten bezeichnet werden, mit dem geringsten Anteil an versiegelten Flächen und Gebäuden sowie dem höchsten Anteil an Vegetation. Schließlich weist das Aktivitätenband (Abbildung 17c&d) mit 62,0 % bzw. 16,0 % eine ähnliche Verteilung von Vegetation und Bäumen auf wie Grünzüge. Der deutlichste Unterschied besteht in den Anteilen von Gebäuden (3,8 %) und versiegelten Flächen (9,2 %) sowie unbewachsenem Boden (6,8 %). Dies erklärt sich durch den Neubau von Infrastruktur für Fußgänger- und Fahrradverkehr sowie von Bereichen für Sport- und Freizeitaktivitäten, die je nach Nutzung sowohl Gebäude als auch versiegelte Flächen umfassen. Darüber hinaus umfasst das Aktivitätenband auch Flächen für Spielplätze, Sitzgelegenheiten, Entspannung und Erholung.



**Abbildung 17: Potentielle Landbedeckung für Grünzüge, Hüllgraben und das Aktivitätenband**

Die potentiellen Änderungen der Landbedeckung auf Siedlungsflächen sind für die drei beispielhaften Neubaugebiete in Abbildung 18 veranschaulicht. Einerseits geht die neue Bebauung mit einem Verlust von natürlichen und naturnahen Flächen einher, was sich in allen drei Beispielen durch den Verlust von Bäumen, Vegetation und Gewässern zeigt. Andererseits führt die neue Bebauung zu einer Zunahme von Gebäuden und versiegelten Flächen, wobei das Ausmaß bei den drei Beispielgebieten erheblich variiert. Der geringste Flächenverbrauch ist beim ersten Beispiel (Siemenswerke) zu beobachten, das aufgrund der Bebauung mit Hochhäusern sehr flächensparend ist. Dies geht mit einem hohen Anteil an Freiflächen

zwischen den Gebäuden einher, die in diesem Beispiel mit Vegetation bedeckt sind, was in Summe zu einem relativ geringen Verlust von Vegetation führt. Das zweite Beispiel (Giesing) zeichnet sich durch einen moderaten Flächenverbrauch durch Gebäude und versiegelte Flächen aus, während der Verlust an Bäumen in diesem Beispiel am geringsten ist. Das dritte Beispiel (Aubing) weist die höchste Bebauungsdichte sowie den höchsten Versiegelungsgrad auf, was auf die Einbeziehung des motorisierten Individualverkehrs und den entsprechenden Platzbedarf der damit verbundenen Infrastruktur zurückzuführen ist.

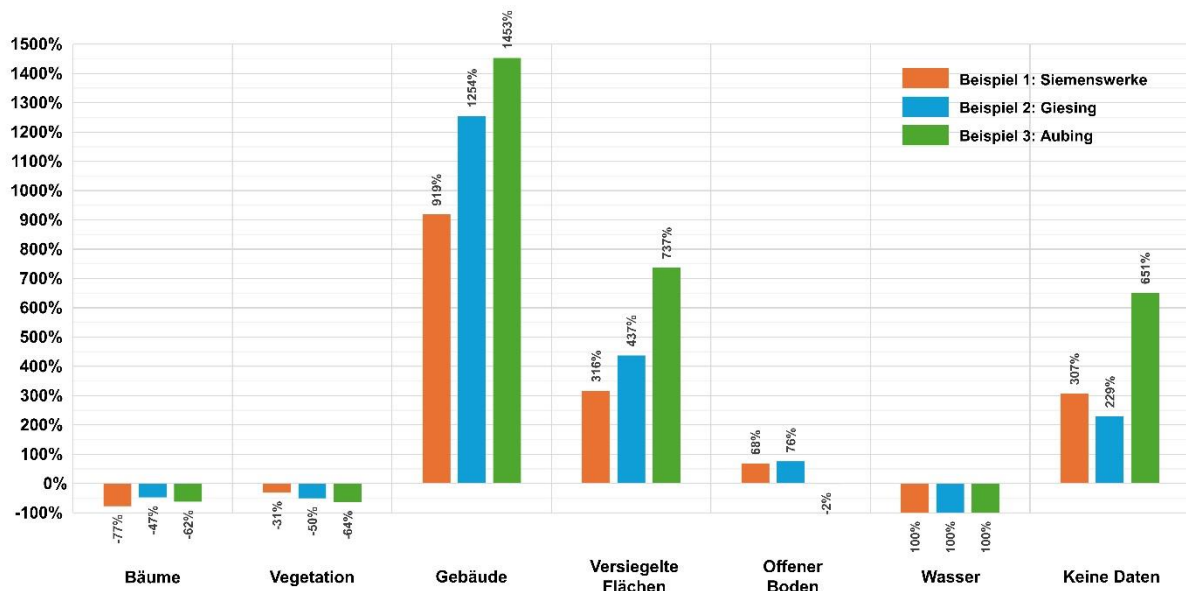


Abbildung 18: Potentielle Änderungen der Landbedeckung auf Siedlungsflächen in Abhängigkeit der drei beispielhaften Neubaugebiete

## 2.5 Verbreitung und Praxistest (AP5.0)

### 2.5.1 Workshop mit externen Experten und Anwendern (Task 5.1)

Am 19. September 2025 fand der Workshop mit externen Experten und Anwendern aus Wissenschaft und Industrie am DLR statt. Am Workshop nahmen folgende Personen teil:

- Robert Behringer (Ingenieurbüro Behringer & Partner)
- Stefan Gamperer (Ingenieurbüro Behringer & Partner)
- Dr. Michael Wurm (DLR)
- Dr. Tobias Leichtle (DLR)
- Lena Huber (Obermeyer Infrastruktur GmbH & Co. KG)
- Helmuth Ammerl (Obermeyer Infrastruktur GmbH & Co. KG)
- Dr. Klaus Martin (Sachverständigenbüro für Luftbildauswertung und Umweltfragen)
- Lukas Glauz (Richard Schulz Tiefbau GmbH & Co. KG)

Großes Interesse wurde außerdem von Christian Schwander (Planungsverband Äußerer Wirtschaftsraum München) und Prof. Ulrike Pröbstl-Haider (Arbeitsgruppe für Landnutzungsplanung GmbH) bekundet, die aus Termingründen nicht am Workshop teilnehmen konnten.

Im Rahmen eines Workshops mit externen Experten und Anwendern wurden die im Projekt entwickelten Ansätze zur KI-gestützten Auswertung von Kamerabilddaten vorgestellt und gemeinsam diskutiert. Ziel war es, die methodischen Ergebnisse hinsichtlich ihrer Praxistauglichkeit zu reflektieren und Anforderungen potenzieller Anwender frühzeitig in die Bewertung einzubeziehen.

Die Rückmeldungen verdeutlichen ein grundsätzlich hohes Interesse an automatisierten Verfahren zur Verkehrserhebung, insbesondere im Hinblick auf Effizienzgewinne und neue Auswertungsmöglichkeiten. Gleichzeitig wurden zentrale Herausforderungen benannt, darunter die Abhängigkeit der Ergebnisse von Aufnahmebedingungen, die eingeschränkte Robustheit bei seltenen Verkehrsteilnehmerklassen sowie Anforderungen an Transparenz und Nachvollziehbarkeit der Ergebnisse.

Aus Sicht der Planungspraxis wurde zudem hervorgehoben, dass kamerabasierte KI-Verfahren derzeit noch nicht als vollwertiger Ersatz etablierter Erhebungsmethoden angesehen werden. So werden in aktuellen Ausschreibungen kommunaler Auftraggeber i.d.R. weiterhin manuelle Verkehrszählungen gefordert. Die entwickelte Methodik ist daher primär als ergänzendes Werkzeug zu verstehen, das insbesondere bei geeigneten Rahmenbedingungen einen Mehrwert bieten kann.

Gleichzeitig leisten die im Projekt erarbeiteten methodischen Grundlagen, die systematische Evaluation sowie die Aufbereitung der Ergebnisse einen wichtigen Beitrag zur weiteren Etablierung solcher Verfahren. Insbesondere durch transparente Analysen, nachvollziehbare Bewertungskriterien und die Veröffentlichung der Ergebnisse kann Vertrauen in die Leistungsfähigkeit KI-gestützter Ansätze aufgebaut und deren zukünftige Integration in die Planungspraxis unterstützt werden.

Der Workshop leistete damit einen wichtigen Beitrag zur Einordnung der entwickelten Methodik im Anwendungskontext und zur Identifikation weiterer Forschungs- und Entwicklungsbedarfe. Insbesondere wurde deutlich, dass für eine breitere Anwendung neben technischen Verbesserungen auch Fragen der Standardisierung, Validierung und Akzeptanz in der Planungspraxis adressiert werden müssen.

### **2.5.2 Praxistest (Task 5.2)**

Im Rahmen von Task 5.2 war ursprünglich die Durchführung eines dedizierten Praxistests vorgesehen, um die praktische Anwendbarkeit der entwickelten Methodik zur automatisierten Auswertung von Kamerabilddaten und zur Integration in ein nachhaltiges Flächenmanagement zu validieren. Aufgrund der bereits umfassenden Anwendung der Methodik in den vorangegangenen Arbeitspaketen AP2.0, AP3.0 und AP4.0 anhand von realen Daten aus zwei realen Planungsprojekten der Obermeyer Infrastruktur GmbH & Co. KG konnte jedoch bereits ein hoher Grad an Praxisrelevanz und Anwendbarkeit demonstriert werden. Dieser Abschnitt fasst die wichtigsten praxisrelevanten Erkenntnisse zusammen und bietet Empfehlungen, wie die im Rahmen des Projekts entwickelte Methodik die Planungspraxis unterstützen kann.

Die verwendeten Praxisbeispiele aus zwei abgeschlossenen Planungsprojekten stellen grundlegend verschiedene Anwendungsfälle dar. Das Projekt im Münchner Nordosten ist ein großes Stadtentwicklungsprojekt der Stadt München mit dem Ziel, ein völlig neues Stadtviertel

für bis zu 30.000 neue Bewohner zu schaffen. Im zweiten Untersuchungsgebiet in der Gemeinde Planegg soll neben dem Ausbau und der Neugestaltung des Bahnhofplatzes eine Verlängerung der U-Bahn vorbereitet werden. Der Fokus von Obermeyer Infrastruktur GmbH & Co. KG liegt in beiden Projekten in der Verkehrsplanung und der Auswertung von eigens zu diesem Zweck erhobenen Kamerabilddaten.

Da beide Projekte unterschiedliche Anwendungsfälle abdecken und die Rahmenbedingungen voneinander abweichen, konnte im Rahmen von OptiPlan die Praxistauglichkeit einer einheitlichen Methodik demonstriert werden. Das in AP3.0 entwickelte Verfahren zur KI-basierten Objekterkennung und automatisierten Auswertung von Kamerabilddaten konnte grundsätzlich in allen Anwendungsszenarien eingesetzt werden. Ein Schwerpunkt der Arbeiten in OptiPlan waren die Experimente hinsichtlich praxisrelevanter Rahmenbedingungen, wie dem Einfluss von Kameraperspektive, Wetter und Beleuchtung, der Rolle des Klassifikationsschlüssels und damit der semantischen Tiefe, sowie der Untersuchung unterschiedlicher Modellarchitekturen. Folgende Schlussfolgerungen und Praxisempfehlungen können abgeleitet werden.

### 1. Einfluss von Kameraperspektive, Wetter und Beleuchtung

Die Genauigkeit der Objekterkennung variiert in Abhängigkeit der Kameraperspektive, Bildqualität, Wetterbedingungen und der Beleuchtungssituation. Insbesondere bei ungünstigen Lichtverhältnissen (Dunkelheit, Gegenlicht) und Wetterverhältnissen die sich negativ auf die Bildqualität auswirken (z.B. Regen, Schnee, Nebel) sinkt die Erkennungsrate. Unterschiedliche Kameraperspektiven (z.B. Vogelperspektive, seitliche Ansicht) erfordern möglicherweise unterschiedliche Modellparameter oder sogar ein erneutes Training der Modelle. Die Ergebnisse verdeutlichen, dass sich extreme Kameraperspektiven grundsätzlich negativ auf die Erkennung auswirken.

#### Praxisempfehlungen:

- **Kameraauswahl:** Bei der Auswahl der Kameras sollten robuste Modelle mit hoher Bildqualität und guter Leistung bei unterschiedlichen Lichtverhältnissen bevorzugt werden.
- **Kameraaufstellung:** Der Aufstellort sollte so gewählt werden, dass die zu erfassenden Objekte möglichst in der Bildmitte mit ausreichender Größe abgebildet werden. Schwierige Lichtverhältnisse (z.B. Gegenlicht) sollten durch eine angepasste Aufstellung (Standort, Blickrichtung) vermieden werden.
- **Bildvorverarbeitung:** Der Einsatz von Bildvorverarbeitungstechniken (z.B. Rauschunterdrückung, Kontrastverstärkung) kann die Bildqualität verbessern und die Erkennungsrate erhöhen.
- **Datenaugmentierung:** Die Verwendung von Datenaugmentierungstechniken (z.B. Rotation, Skalierung, Helligkeitsänderung) während des Trainings kann die Robustheit des Modells gegenüber unterschiedlichen Rahmenbedingungen erhöhen.

### 2. Rolle des Klassifikationsschlüssels und der semantischen Tiefe

Die Wahl des Klassifikationsschlüssels und der semantischen Tiefe haben einen erheblichen Einfluss auf die erzielbare Genauigkeit des Verfahrens. Eine feingliedrige Aufschlüsselung (z.B. Unterscheidung verschiedener PKW- oder LKW-Typen, Erkennung von

Spezialfahrzeugen) kann die Erkennungsrate verringern, während ein zu stark generalisierender Klassifikationsschlüssel nicht für alle Anwendungsfälle ausreicht.

#### **Praxisempfehlungen:**

- **Zielorientierte Auswahl des Klassifikationsschlüssels:** Der Klassifikationsschlüssel sollte individuell an die spezifischen Anforderungen des Planungsprojekts angepasst werden, hierfür ist ein frühzeitiger Austausch mit den Anwendern sinnvoll.
- **Ausgewogene semantische Tiefe:** Die semantische Tiefe sollte so gewählt werden, dass sie möglichst alle notwendigen Informationen liefert, ohne die Erkennungsrate durch zu große Detailtiefe unnötig zu beeinträchtigen. Eine größere semantische Tiefe bedeutet auch einen erhöhten Aufwand bei der Sammlung und Aufbereitung von Referenzdaten, da Spezialfahrzeuge sehr heterogene Erscheinungsbilder aufweisen können, so dass besonders umfangreiche Referenzdaten zu erfassen sind.
- **Iterative Verfeinerung:** Sofern es der Zeitplan erlaubt, können der Klassifikationsschlüssel und die semantische Tiefe iterativ verfeinert werden, um diese auf Basis der Ergebnisse der KI-Auswertung und den Rückmeldungen der Anwender schrittweise anzupassen.

#### **3. Untersuchung unterschiedlicher Modellarchitekturen:**

Der Vergleich verschiedener KI-Modelle mit unterschiedlicher Komplexität hat gezeigt, dass die Wahl der Architektur einen klaren Einfluss auf die Leistung des Systems hat. Während die weniger komplexe Architektur erwartungsgemäß niedrigere Genauigkeiten bietet, ist die Laufzeit entsprechend kürzer bzw. sind auch weniger leistungsstarke Auswertesysteme geeignet. Darüber hinaus konnte gezeigt werden, dass der Genauigkeitsgewinn der komplexeren Architektur vor allem die Erkennung kleiner Objekte verbessert, während große Objekte wie PKW oder LKW kaum profitieren.

- **Architekturauswahl:** Bei der Auswahl der Modellarchitektur sollten die spezifischen Anforderungen des Planungsprojekts berücksichtigt werden (z.B. die Komplexität der Verkehrssituation, die Größe der zu erkennenden Objekte).
- **Systematischer Vergleich:** Sofern möglich, können verschiedene Architekturen systematisch verglichen werden, um die optimale Lösung zu identifizieren.
- **Transfer Learning:** Der Einsatz von Transfer Learning (d.h. die Verwendung von vortrainierten Modellen) kann die Trainingszeit verkürzen und die Leistung des Modells verbessern.

Die Anwendung der entwickelten Methodik in den realen Planungsprojekten zeigte deutliche Vorteile gegenüber herkömmlichen Methoden. Die automatisierte Auswertung von Kamerabildern reduziert den bisher üblichen manuellen Aufwand erheblich und beschleunigt damit den Auswertungs- und Planungsprozess. Die KI-basierte Objekterkennung liefert präzise Daten über das Verkehrsaufkommen und die Verteilung von Verkehrsträgern. Die Integration der Daten in die Flächenanalyse ermöglicht eine fundiertere Bewertung des Flächenverbrauchs und unterstützt die Entwicklung nachhaltiger Planungsvarianten. Die automatisierte Auswertung minimiert subjektive Fehler und sorgt für eine objektive Datengrundlage. Die positiven Rückmeldungen der Experten und Anwender untermauern die Praxistauglichkeit und den Mehrwert der entwickelten Methodik.

## 2.6 Referenzen

- Al-qaness, M. A. A., Abbasi, A. A., Fan, H., Ibrahim, R. A., Alsamhi, S. H., & Hawbani, A. (2021). An improved YOLO-based road traffic monitoring system. *Computing*, 103(2), 211–230. <https://doi.org/10.1007/s00607-020-00869-8>
- Forschungsgesellschaft für Straßen- und Verkehrswesen (Ed.). (2012). *Recommendations for Traffic Surveys—Empfehlungen für Verkehrserhebungen (EVE)* (Ausz. 2012). FGSV-Verl.
- Galich, A., Stiller, D., Wurm, M., & Taubenböck, H. (2025). AI-Based Counting of Traffic Participants: An Explorative Study Using Public Webcams. *Future Transportation*, 5(3), 87. <https://doi.org/10.3390/futuretransp5030087>
- Gour, D., & Kanskar, A. (2019). Automated AI Based Road Traffic Accident Alert System: YOLO Algorithm. 8(08), 5.
- LeCun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional networks and applications in vision. *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium On*, 253–256. <http://ieeexplore.ieee.org/abstract/document/5537907/>
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740–755). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10602-1\\_48](https://doi.org/10.1007/978-3-319-10602-1_48)
- Matthias, V., Bieser, J., Mocanu, T., Pregger, T., Quante, M., Ramacher, M. O. P., Seum, S., & Winkler, C. (2020). Modelling road transport emissions in Germany – Current day situation and scenarios for 2040. *Transportation Research Part D: Transport and Environment*, 87, 102536. <https://doi.org/10.1016/j.trd.2020.102536>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788. [http://www.cv-foundation.org/openaccess/content\\_cvpr\\_2016/html/Redmon\\_You\\_Only\\_Look\\_CVPR\\_2016\\_paper.html](http://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Redmon_You_Only_Look_CVPR_2016_paper.html)
- Redmon, J., & Farhadi, A. (2016). YOLO9000: Better, faster, stronger. *arXiv Preprint arXiv:1612.08242*. <https://arxiv.org/abs/1612.08242>
- Schulthoff, M., Kaltschmitt, M., Balzer, C., Wilbrand, K., & Pomrehn, M. (2022). European road transport policy assessment: A case study for Germany. *Environmental Sciences Europe*, 34(1), 92. <https://doi.org/10.1186/s12302-022-00663-7>
- Shahab, S., Hartmann, T., Jonkman, A. Strategies of municipal land policies: housing development in Germany, Belgium, and Netherlands. *European Planning Studies*, 29(6), 1132–1150, 2021. <https://doi.org/10.1080/09654313.2020.1817867>
- Stiller, D., Wurm, M., Staab, J., Stark, T., Starz, G., Rauh, J., Dech, S., & Taubenböck, H. (2026). Open webcam data for traffic monitoring: YOLOv8 detection of road users before and during COVID-19. *Transportation Research Interdisciplinary Perspectives*, 36, 1058574. <https://doi.org/10.1016/j.trip.2025.101774>
- Taubenböck, H., Esch, T., Wurm, M., Roth, A., Dech, S. Object-based feature extraction using high spatial resolution satellite data of urban areas. *Journal of Spatial Science* 55, 117–132, 2010. <https://doi.org/10.1080/14498596.2010.487854>

Taubenböck, H., Mast, J., Geiß, C., Wurm, M., Esch, T., Seto, K. Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios. *Cities* 151, 105117, 2024. <https://doi.org/10.1016/j.cities.2024.105117>

Wurm, M., Taubenböck, H., Schardt, M., Esch, T., Dech, S. Object-based image information fusion using multisensor earth observation data over urban areas. *International Journal of Image and Data Fusion* 2, 121–147, 2011. <https://doi.org/10.1080/19479832.2010.543934>

Xia, W., Li, P., Huang, H., Li, Q., Yang, T., & Li, Z. (2024). TTD-YOLO: A Real-Time Traffic Target Detection Algorithm Based on YOLOV5. *IEEE Access*, 12, 66419–66431. <https://doi.org/10.1109/ACCESS.2024.3394693>

Zhang, Y., Guo, Z., Wu, J., Tian, Y., Tang, H., & Guo, X. (2022). Real-Time Vehicle Detection Based on Improved YOLO v5. *Sustainability*, 14(19), Article 19. <https://doi.org/10.3390/su141912274>

### **3 Wichtigste Positionen des zahlenmäßigen Nachweises**

Die Projektausgaben des Ingenieurbüros Behringer & Partner entstanden aus den Kosten für die Projektleitung und die Tätigkeiten des eingesetzten Personals.

Die Projektausgaben des DLR umfassen die Kosten für einen wissenschaftlichen Mitarbeiter (KAT I DFD, dies entspricht einem Mitarbeiter der Entgeltgruppe 13).

### **4 Notwendigkeit und Angemessenheit der geleisteten Arbeit**

Die durchgeführten Arbeiten waren zum Erreichen der angestrebten Projektziele notwendig, der Verlauf der Arbeit im Projekt folgte dabei der im Projektantrag formulierten Planung. Alle im Arbeitsplan formulierten Aufgaben wurden erfolgreich bearbeitet, ohne dass zusätzliche Ressourcen zur Durchführung des Vorhabens aufgewendet werden mussten.

Um ein an den Bedürfnissen potentieller Nutzer ausgerichtetes, praxisrelevantes Produkt zu entwickeln, war es zunächst notwendig die Bedarfe der Planungspraxis zu identifizieren und einen spezifischen Anforderungskatalog, insbesondere hinsichtlich der praxisrelevanten Verkehrsteilnehmer, zu entwickeln. Hierfür wurden praxisnahe Kamerabilddaten aus zwei vorangegangenen Verkehrsplanungsprojekten der Obermeyer Infrastruktur GmbH & Co. KG herangezogen. Auf dieser Grundlage wurde eine umfangreiche Referenzdatenbasis aufgebaut, in der rund 17.172 Verkehrsteilnehmer manuell annotiert und in das hierarchische Klassifikationsschema eingeordnet wurden. Zur Evaluierung des Flächenverbrauchs von Infrastrukturmaßnahmen wurden aktuelle Satellitenbilddaten mit sehr hoher räumlicher Auflösung sowie weitere Geodaten akquiriert, um die kleinräumig heterogenen Strukturen der Landbedeckung berücksichtigen und die Auswirkungen von planerischen Entscheidungen abbilden zu können.

Um die die Leistungsfähigkeit kamerabasierter KI-Verfahren in einem praxisnahen Umfeld evaluieren zu können, wurden verschiedene Faktoren systematisch untersucht. Dazu zählen Wetter & Beleuchtung, das verwendete Klassifikationsschema bzw. dessen semantische Tiefe sowie die Modellarchitektur selbst, die in voneinander unabhängigen Experimenten untersucht wurden. Dieses Vorgehen ermöglicht eine Abschätzung, unter welchen Bedingungen robuste und praxisnahe Ergebnisse erzielt werden können. Als KI-Modell wurden vortrainierte

Objektdetektoren der YOLO-Familie verwendet, deren grundlegende Eignung bereits in vorherigen wissenschaftlichen Studien dokumentiert ist. Die Implementierung dieses methodisch komplexen Verfahrens und die systematische Umsetzung der verschiedenen Experimente wären ohne die Förderung durch das BMBF für das Ingenieurbüro Behringer & Partner nicht machbar. Auch für das DLR sind die spezifischen Entwicklungsarbeiten nicht durch die Grundfinanzierung abgedeckt.

Das große Interesse der Obermeyer Infrastruktur GmbH & Co. KG während des Vorhabens zeigt die Notwendigkeit neuer Methoden zur Erfassung von Verkehrsströmen sowie einer Differenzierung von Verkehrsteilnehmern. Vor allem die durchgeführten Experimente zur Einschätzung der Praxistauglichkeit der angewandten KI-Methoden ist für die Planungspraxis von großem Interesse. Auch das positive Feedback und die anregende Diskussion im Verlauf des Workshops heben die Relevanz der durchgeführten Arbeiten im Rahmen von OptiPlan hervor.

## **5 Voraussichtlicher Nutzen und Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans**

### **5.1 Voraussichtlicher Nutzen für das Ingenieurbüro Behringer & Partner**

Für das Ingenieurbüro ergibt sich aus den entwickelten Methoden und Ergebnissen ein Wissensvorsprung in der Anwendung von KI allgemein und speziell bei Verkehrszählungen, aus dem mittelfristig auch ein wirtschaftlicher Mehrwert zu erwarten ist. Auch wenn bisher einschlägige Ausschreibungen eine KI-gestützte Erfassung des Verkehrsaufkommens nicht als Methode nennen, wird das Ingenieurbüro zukünftig Kunden auf diese Erfassungsmethode zumindest als Ergänzung hinweisen. Im Zuge der rasanten Entwicklung und Verbreitung von KI ist eine entsprechende Nachfrage mittelfristig zu erwarten, und das nicht zuletzt angesichts geringerer Kosten, sowie einer effizienteren und objektiveren Erfassung von Verkehr zu Planungszwecken.

### **5.2 Voraussichtlicher Nutzen für das DLR**

Von wissenschaftlicher Seite sind vor allem die Methoden zur Objekterkennung und Differenzierung von Verkehrsteilnehmern sehr innovativ. Dieses Anwendungsfeld ermöglicht dem DLR den Einstieg in ein neues Forschungsfeld. Aufbauend auf die Erkenntnisse aus OptiPlan ist eine Veröffentlichung in einem internationalen „peer-reviewed“ Journal geplant, um die Ergebnisse dieses aktuellen Themas der breiten Forschungsgemeinschaft zugänglich zu machen. Darüber hinaus hat das DLR die methodischen und inhaltlichen Arbeiten auf der internationalen Planungskonferenz REAL CORP 2026 vorgestellt und so die Sichtbarkeit in der wissenschaftlichen Community erhöht. Außerdem plant das DLR über ähnliche Projekte die Methoden und Algorithmen weiter zu verfolgen und zu vertiefen, um die KI-Methodik stetig zu verbessern. Die während dem Projekt entstandenen Kontakte zu Wirtschaftsunternehmen bieten außerdem neue Möglichkeiten für zukünftige Kooperationen und Weiterentwicklungen im Rahmen zukünftiger Ausschreibungen und Forschungsprojekte.

## **6 Während der Durchführung des Vorhabens bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen**

Von dritter Seite sind keine Ergebnisse bzw. Fortschritte bekannt geworden, die für die Durchführung des Vorhabens relevant sind.

## **7 Erfolgte und geplante Veröffentlichungen der Ergebnisse**

Konferenzbeiträge:

Leichtle, T., Kühnl, M., Staab, J., Stiller, D., Gamperer, S., Martin, K., Behringer, R., Wurm, M. & Taubenböck, H. (2026): Earth Observation and open geodata for sustainable urban planning – A practical case study from the city of Munich, Germany. In: Proceedings of the Real CORP, Real CORP 2026, 23.-25.03.2026, Wien, Österreich.

[https://proceedings.corp.at/papers2026/CORP2026\\_146.pdf](https://proceedings.corp.at/papers2026/CORP2026_146.pdf)

Staab, J., Stiller, D., Leichtle, T., Kühnl, M., Huber, L., Gamperer, S., Martin, K., Behringer, R., Ammerl, H., Wurm, M. & Taubenböck, H. (2026): Von MS-COCO zur FGSV: Die Entwicklung und Validierung eines hierarchischen Klassifikationsschemas für die hochspezifische KI-gestützte Verkehrszählung. In: Proceedings of the Real CORP, Real CORP 2026, 23.-25.03.2026, Wien, Österreich.

[https://proceedings.corp.at/papers2026/CORP2026\\_130.pdf](https://proceedings.corp.at/papers2026/CORP2026_130.pdf)