Automatisierung des Erfassungsprozesses von topographischen Vektordaten mit freiverfügbaren Geodaten und KI

Erläuterung der Problemstellung, die mit der Idee gelöst werden soll:

Hochgenaue und aktuelle topographische Vektordaten bilden die Grundlage für digitale Systeme: als taktische und operative Führungskarten, oder in Logistiksystemen (Routing). Dabei vermitteln sie einen raschen Überblick über Bebauungsstrukturen, Landnutzung und Geländeformen oder dienen der Navigation.

Die Erfassung detaillierter Vektordaten als Grundlage hochgenauer topographischer Karten erfolgt derzeit fast ausschließlich manuell. Insbesondere die sehr hohen Anforderungen der Sicherheitsbehörden und des Militärs bzgl. Qualität, geometrischer und thematischer Genauigkeit und Aktualität, sowie der Umfang der zu berücksichtigenden Objektklassen, haben bisher eine Automatisierung nicht zugelassen.

Methoden der klassischen automatischen Bildanalyse und Bildklassifikation werden diesen hohen Anforderungen nicht gerecht. Aufwände in der Nachbereitung der Ergebnisse und die eingeschränkte Übertragbarkeit der Modelle und Regelwerke dieser Methoden auf andere Bilddaten und geographische Gegebenheiten sprachen bislang gegen eine weitgehende Automatisierung.

Die Verfügbarkeit freier Geoinformationen steigt stetig an. Mit OpenStreetMap beispielsweise existiert ein täglich wachsender weltweiter Vektordatensatz. Die heterogene Qualität und Vollständigkeit sprachen bisher gegen eine Nutzung der freien Daten. Zu berücksichtigen sind auch die derzeit inkompatiblen Lizenzbedingungen von OpenStreetMap und der Geheimhaltungsgrad militärischer Daten.

Der hohe Zeit- und Arbeitsaufwand für die manuelle Erzeugung und Aktualisierung großer Datenbestände verursacht hohe Kosten und lange Beschaffungszeiträume. Um den sich schnell ändernden geopolitischen Situationen und der Vielzahl an Interessens- und Einsatzgebieten der Bundeswehr gerecht zu werden, muss die Automatisierung der Vektordatenerfassung vorangetrieben werden. Dies betrifft sowohl die Neuerfassung als auch die Aktualisierung der stetig wachsenden Bestandsdaten.

Beschreibung der Idee und wie sie das Problem lösen soll:

Grundidee

Die Kombination aus Methoden der KI, wie Deep Learning, und frei verfügbaren Vektordaten bietet ein großes Potenzial, die zurzeit vollständige manuelle Erfassung und Aktualisierung der Vektordaten weitgehend zu automatisieren.

Am Beispiel des Straßennetzes konnte gezeigt werden, dass der Einsatz von Deep Learning zu einer erheblichen Effizienzsteigerung in der Vektordatenproduktion führen kann. Mit klassischen Methoden des maschinellen Lernens ist dies nicht erreicht worden. Hervorzuheben ist

an dieser Stelle die Unabhängigkeit von der geographischen Situation oder des Aufnahmezeitpunktes. Voraussetzung dafür ist neben einer adäquaten Netzarchitektur ein hochgenauer Trainingsdatensatz und ein auf die Objektart (hier Straßennetz) abgestimmtes Post-Processing. Neben der Extraktion der reinen Geometrie können zusätzlich Attribute wie Straßenbreite und Belag automatisch bestimmt werden.

Die Automatisierung des gesamten Erfassungsprozesses sollte aufgrund der großen Anzahl an Objektarten nicht in einem Schritt, sondern sukzessiv für einzelne Objektarten (z. B. Straßen, Gewässer, Pylone) erfolgen. Nur so kann auch in Zukunft das hohe Qualitätsniveau der Datengewinnung sichergestellt werden. Es sollte deshalb eine Priorisierung unter folgenden Gesichtspunkten vorgenommen werden: Häufigkeit bzw. Verbreitung der Objektart, Einsparungspotential gegenüber der manuellen Erfassung, Erfolgsaussicht der technischen Umsetzung.

Freiverfügbare Vektordaten, wie OpenStreetMap, weisen in vielen Teilen der Erde für einzelne Objektarten eine hohe räumliche Genauigkeit und Vollständigkeit auf. Die Verwendung dieser Daten würde die Effizienz der Vektordatenerfassung oder -aktualisierung steigern. Voraussetzung für die Nutzbarmachung der freien Vektordaten ist der Einsatz von Conflation Werkzeugen mit denen unterschiedliche Datenquellen zusammengeführt werden können.

Insbesondere die Kombination des Einsatzes von KI und der Nutzung freier Vektordaten erhöhen die Effektivität der Vektordatenproduktion. Das Vorgehen ist entsprechend der Qualität der automatischen Vektordatenerfassung mittels KI, sowie der Qualität und Verfügbarkeit freiverfügbarer Vektordaten zu bestimmen. Der Workflow sollte u. a. folgende Eigenschaften aufweisen: Anreicherung der automatisch erfassten Vektoren mit Attributdaten aus freien Datenquellen, Lagekorrektur der freiverfügbaren Daten mittels automatisch abgeleiteter Vektordaten, Zusammenführen von freiverfügbaren Vektordaten mit automatisch abgeleiteten Vektordaten, Aktualisierung von Bestandsdaten mittels automatisch abgeleiteter und/oder freiverfügbarer Vektordaten.

Sowohl die mittels KI erfassten als auch die freiverfügbaren Vektordaten stimmen in der Regel nicht mit den geforderten Generalisierungsvorschriften und dem semantischen Datenmodell überein. Geeignete Algorithmen und Workflows sind dafür zu entwickeln.

Die Menge und Diversität von Trainingsdaten zum Anlernen der Deep Learning Modelle bestimmen wesentlich die Güte der Ergebnisse. Die Rückführung der extrahierten Objekte in den Trainingsprozess sollte deshalb ein wichtiger Bestandteil der Vektordatenerfassung sein.

Technisch und Organisatorische Umsetzung

Die sukzessive Vorgehensweise bei der Automatisierung erlaubt die Verwendung bewährter Strukturen des etablierten Produktionsprozesses und stellt somit die Einhaltung der hohen Qualitätskriterien auch in Zukunft sicher. Der modulare Aufbau des Workflows bietet die Möglichkeit, die Erfassung einzelner Objektarten zu automatisieren. Routinen der Qualitätssicherung können dabei direkt aus dem bestehenden Workflow übernommen werden. Am Beispiel des Straßennetzes konnte dieses Vorgehen bereits in der Praxis erfolgreich erprobt werden.
Für die technologische Umsetzung ist die bestehende Umgebung zur Vektordatenerfassung um folgende Hardware-Komponenten zu erweitern:

Eine GPU-Workstation zum Trainieren der Neuronalen Netze. Alternativ kann dafür auch ein (privater) cloud-processing Dienst verwendet werden.
Ein oder mehrere leistungsstarke Workstations mit mindestens einer GPU zur Objektextraktion/-detektion, zur automatischen Nachbearbeitung der Vektorobjekte (Postprocessing) und zur Integration von OpenStreetMap Daten (Conflation).
Die software-technische Umsetzung basiert auf vorhandene FOSS (Conflation-Software, KI-Frameworks). Die Implementierung des Workflows erfolgt durch die Entwicklung von Skripten.

Vorliegende technische Komponenten

Die für die Umsetzung der Lösung notwendigen Hardware-Komponenten sind kommerziell erhältlich. Prinzipiell ist die Nutzung leistungsstarker Consumer-Technologien möglich. Bei hohen Anforderungen an den Durchsatz im Produktionsprozess können rechenintensive Anteile in die Cloud ausgelagert werden.

Die Software-Komponenten der Lösung basieren auf etablierte FOSS-Bibliotheken oder Komponenten (z. B. Tensorflow für Deep Learning und freie Conflation Software).

Die automatische Extraktion von Straßendaten wurde als Prototyp implementiert. Die Integration in den Produktionsprozess wird derzeit realisiert. Die Straßenextraktion bildet somit die Vorlage für die Automatisierung der Produktion weiterer Objektarten.

Zusammenarbeit von Akteuren

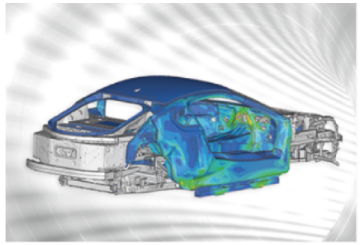
Für die erfolgreiche Umsetzung der Lösung müssen verschiedene Akteure ihren Beitrag leisten und mit den jeweils anderen beteiligten Akteuren im Austausch stehen. Wichtige Beiträge liefern u.a.:

Militär/Sicherheitsagenturen: Bereitstellen von Trainingsdaten (Satellitenbilddaten, Vektordaten als Labels), Ausschreibung von Projekten zur Umsetzung der Lösung (Finanzierung)
Privatwirtschaft: Erzeugung von Labels aus verfügbaren Vektordaten, Umsetzung/Implementierung der Lösung, Integration in den etablierten Produktionsprozess

OpenStreetMap-Community: Klärung bzgl. Lizenzbedingungen und Nutzbarkeit der OSM-Daten

Konvergenz zu bestehenden und anderen Systemen

Die vorgeschlagene sukzessive Automatisierung der Vektordatenproduktion integriert sich nahtlos in den bestehenden Produktionsprozess. Negative Auswirkungen auf Qualitätsstandards und die Konformität zu den existierenden Datenschemata sind nicht zu erwarten.



AUTOMOTIVE

INFOKOM

MOBILITÄT, ENERGIE & UMWELT

LUFTFAHRT

RAUMFAHRT

VERTEIDIGUNG & SICHERHEIT

Kartenerstellung mit KI

Automatisierung des Erfassungsprozesses von topographischen Vektordaten mit frei verfügbaren Geodaten und KI

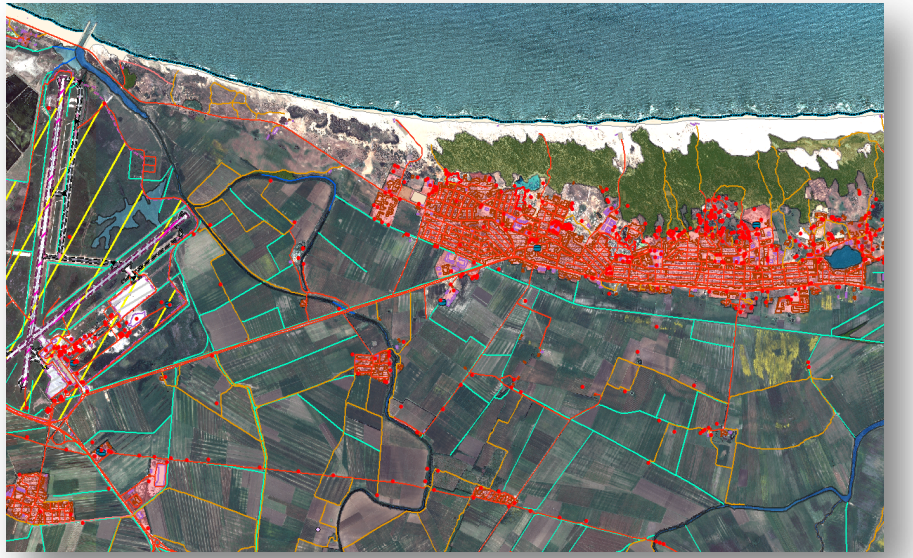
Iris Wieser

München, 10.07.2019

IABG

Topographische Vektordaten

Fernerkundungsdaten
(Satellitenbild, Luftbild, UAV-Daten)



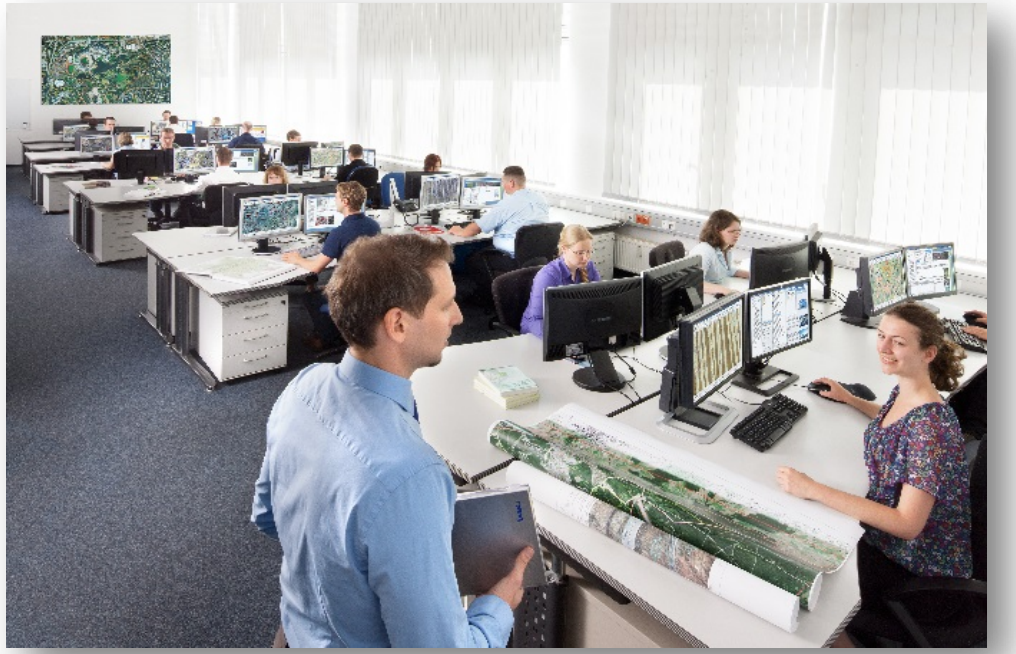
Erfassung von Objekten & Attributen

Taktische und operative Führungskarten, Grundlage für Logistiksysteme,..

Status quo

- Im Wesentlichen **manuell**
- Gründe:
  - Hoher Anspruch an Qualität
  - Viele verschiedene Objektarten

➢ Hoher Aufwand
➢ Lange Produktionszeiträume



Die Idee

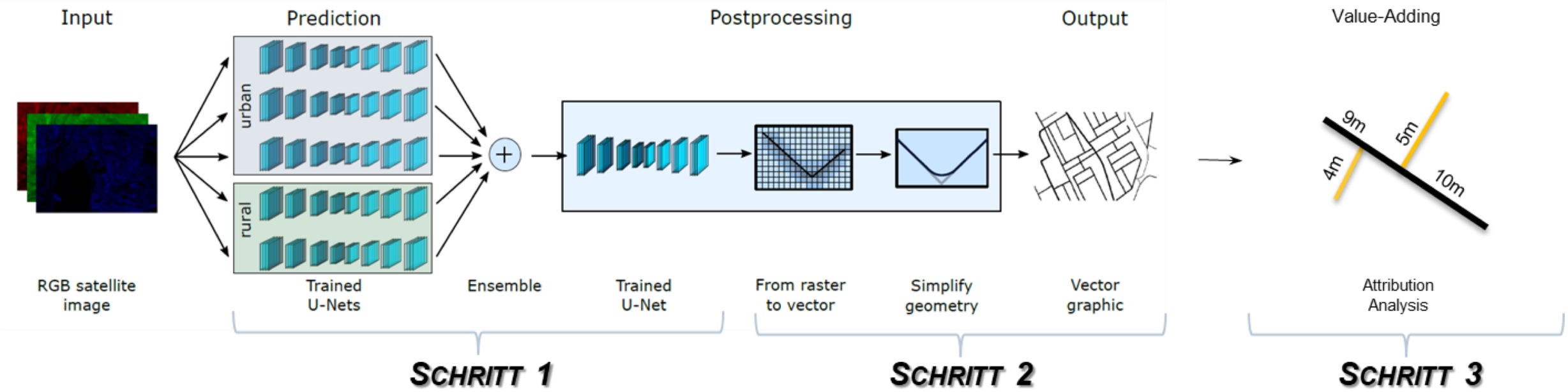
1. **Neuronale Netze**
   - Sehr effizient bei Bilderkennung
   - Hohe Übertragbarkeit
   - Lernen aus Erfahrung

   → Detektion von **Objekten** (z.B. Brunnen, Straßen)
   → Ableitung von **Attributen** (z.B. Straßenbelag)

2. **Frei verfügbare Geodaten**
   - Geometrische Verbesserung
   - Vervollständigung der Objekte und Attribute

© IABG 2018

Neuronale Netze zur Straßenerkennung



© IABG 2018

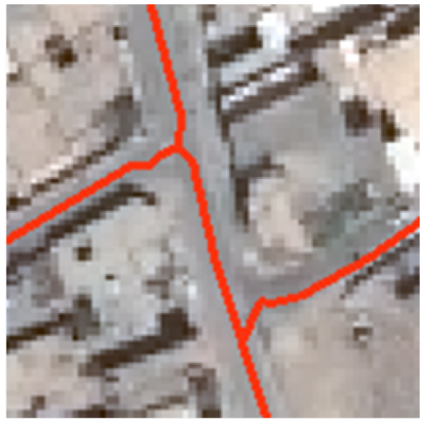
Beispiel Straßennetz

1. Schritt: Neuronale Netze

© IABG 2018

Beispiel Straßennetz

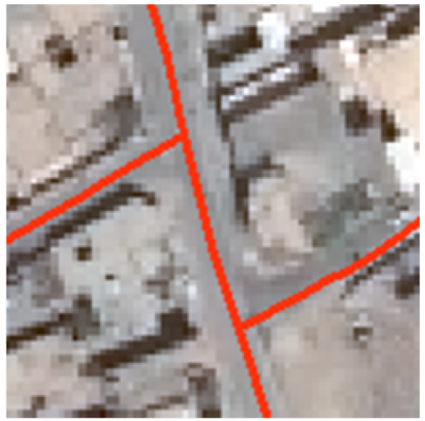
2. Schritt: Vektorisierung



© IABG 2018

Beispiel Straßennetz

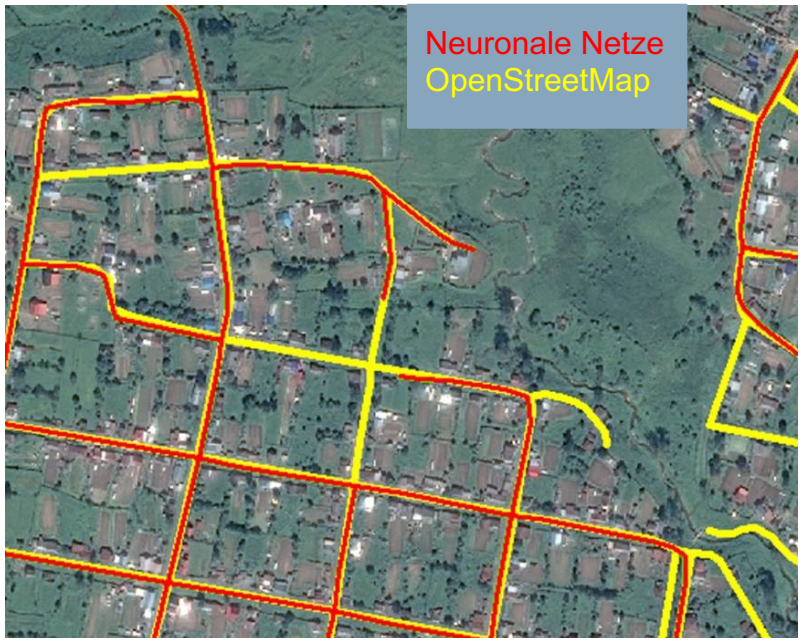
2. Schritt: Generalisierung



© IABG 2018

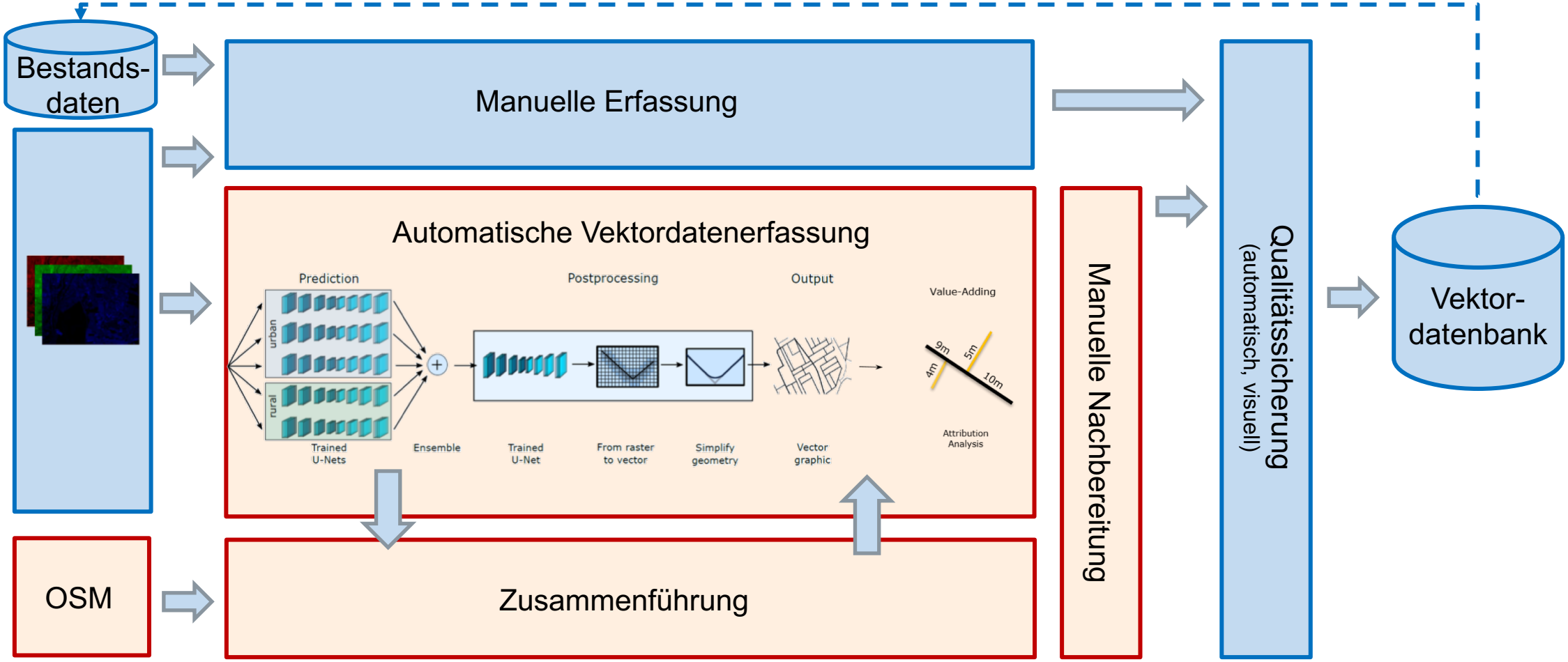
Integration frei verfügbarer Daten

- Weltweiter Datensatz (z. B. OpenStreetMap)
- Schwierigkeiten:
  - Heterogene Qualität und Vollständigkeit
  - Lizenzbedingungen von OpenStreetMap



Zusammen-führung

Pipeline


Bestands-daten
Manuelle Erfassung
Automatische Vektordatenerfassung
Prediction
Postprocessing
Output
Value-Adding
urban
rural
Trained U-Nets
Ensemble
Trained U-Net
From raster to vector
Simplify geometry
Vector graphic
9m
5m
4m
10m
Attribution Analysis
OSM
Zusammenführung
Manuelle Nachbereitung
Qualitätssicherung (automatisch, visuell)
Vektor-datenbank

Fazit und Ausblick

Zusammenfassung

- Dynamische geopolitische Situationen
- Manuelle Erfassung
  → *große Beschaffungszeiträume*

Unser Vorschlag

- Neuronale Netze erlauben eine effizientere Produktion
- Verbesserung durch frei verfügbare Daten
- Erhöhung der Effizienz
  → *kleinere Beschaffungszeiträume*

Nächste Schritte

- Model für alle Objektarten
- Nutzung frei verfügbarer Vektordaten

Ihr Ansprechpartner

IABG mbH
Innovationszentrum
Iris Wieser
Einsteinstraße 20
85521 Ottobrunn
Telefon +49 89 6088-3812
Fax +49 89 6088-xxxx
wieser@iabg.de
www.iabg.de

Danke