ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Stäger, Elena

Article

Neuronale Netze in der Baustatistik: Automatisiertes Erkennen von Baustellen anhand von Luftbildern

WISTA - Wirtschaft und Statistik

Provided in Cooperation with:
Statistisches Bundesamt (Destatis), Wiesbaden

Suggested Citation: Stäger, Elena (2025) : Neuronale Netze in der Baustatistik: Automatisiertes Erkennen von Baustellen anhand von Luftbildern, WISTA - Wirtschaft und Statistik, ISSN 1619-2907, Statistisches Bundesamt (Destatis), Wiesbaden, Vol. 77, Iss. 2, pp. 136-147

This Version is available at:
https://hdl.handle.net/10419/316331

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Elena Stäger

hat Statistik an der Technischen Universität Dortmund studiert und arbeitet seit 2021 bei IT.NRW – Statistisches Landesamt Nordrhein-Westfalen im Service „Mathematisch-statistische Methoden und experimentelle Statistik“. Für ihre Masterarbeit „Neuronale Netze in der Baustatistik: Automatisiertes Erkennen von Baustellen anhand von Luftbildern“ wurde sie im Jahr 2022 mit dem Wissenschaftlichen Nachwuchspreis „Statistical Science for the Society“ des Statistischen Bundesamtes in der Kategorie „Master/Bachelorarbeiten“ geehrt. Sie stellt diese von Prof. Dr. Markus Pauly (Technische Universität Dortmund) und Dr. Johannes Rohde (IT.NRW) betreute Arbeit in diesem Beitrag vor.

NEURONALE NETZE IN DER BAUSTATISTIK: AUTOMATISIERTES ERKENNEN VON BAUSTELLEN ANHAND VON LUFTBILDERN

Elena Stäger

↘ **Schlüsselwörter:** Bilderkennung – Maschinelles Lernen – Orthophotos – räumliche Strukturen – Testregionen

ZUSAMMENFASSUNG

Können Methoden der Künstlichen Intelligenz helfen, Baustellenaktivitäten in Nordrhein-Westfalen zu erfassen und somit mehr Informationen über das dortige Baugeschehen zu erhalten? Der Artikel untersucht ob es möglich ist, einen Datensatz zu erstellen, um einen Algorithmus zu trainieren, der Baustellen in unbekannten Bildern erkennen und die Vollständigkeit der gemeldeten Bauvorhaben überprüfen kann. Nach einer kurzen Einordnung der Baustatistik beschreibt der Beitrag die Erfassung und Verarbeitung von Bilddaten sowie die Implementierung von Convolutional Neural Networks zur Bildanalyse.

↘ ***Keywords:*** *image recognition – machine learning – orthophotos – spatial structures – test regions*

ABSTRACT

Can artificial intelligence methods be used to capture construction site activities in Nordrhein-Westfalen (NRW) and help gather more information on building activities in the state? This article explores whether it is possible to create a dataset to train an algorithm that can identify construction sites in unknown images and check whether all construction projects have been reported. After a brief explanation of construction statistics in NRW, the article goes on to describe the steps involved in capturing and processing image data and the implementation of convolutional neural networks for image analysis.

1

Einleitung

In Deutschland gibt es in jedem der 16 Bundesländer eine Landesbauordnung; diese basieren auf der Musterbauordnung[1], unterscheiden sich jedoch in Details. In Nordrhein-Westfalen müssen nach der Bauordnung für das Land Nordrhein-Westfalen alle Bauvorhaben an Gebäuden gemeldet werden, bestimmte Kenngrößen daraus werden in der Baugenehmigungsstatistik veröffentlicht. Dies ist wichtig, um Aussagen über die aktuelle Wirtschaftsentwicklung zu treffen, die in der Bautätigkeitsstatistik sichtbar werden, und dient beispielsweise zur Entscheidungsfindung in der Städteplanung. Es ist allerdings unklar, ob wirklich alle Bauvorhaben gemeldet werden.

Ziel der diesem Beitrag zugrunde liegenden Masterarbeit war herauszufinden, ob Methoden der Künstlichen Intelligenz angewendet werden können, um mehr Informationen über das Baugeschehen in Nordrhein-Westfalen zu erhalten. Satellitenbilder sind flächendeckend und in regelmäßigen Zeitabständen für die ganze Welt verfügbar. Diese Bilder haben allerdings eine geringe Auflösung. Für Nordrhein-Westfalen liefert die Bezirksregierung Köln von Flugzeugen aus aufgenommene Luftbilder mit einer höheren Auflösung, die in geringeren Zeitabständen als die Satellitenbilder verfügbar sind. Nun soll mithilfe dieser Daten ein Datensatz erstellt werden, der benutzt wird, um einen Algorithmus zu trainieren. Dieser soll an unbekannten Bildern erkennen können, wo sich Baustellen befinden. Diese Erkenntnisse können mit den gemeldeten Bauvorhaben verglichen werden.

Zur Bildverarbeitung stehen unterschiedliche Algorithmen zur Verfügung, die in der Literatur etabliert sind. Aufgrund ihrer Architektur, die räumliche Strukturen berücksichtigt, haben Neuronale Netze in diesem Zusammenhang besondere Bedeutung. Insbesondere zur Lokalisierung von Gebäuden und zur Schadenserkennung an ihnen haben in den vergangenen Jahren Convolutional Neural Networks (Castelluccio und andere, 2015; Gardner/Nichols, 2017) Anwendung gefunden und vielversprechende Ergebnisse erzielt.

Nach einer kurzen Einführung in die Baustatistik in Kapitel 2 erläutert Kapitel 3, welche Daten verfügbar sind. Kapitel 4 enthält Beschreibungen der verwendeten Methoden, während sich Kapitel 5 der Analyse der erstellten Datensätze widmet. Diese Ergebnisse werden in Kapitel 6 diskutiert und abschließend wird ein Ausblick auf weitere Entwicklungsmöglichkeiten gegeben.

2

Hintergrund Baustatistik

Die Statistiken der Baugenehmigungen und -fertigstellungen sowie der Bauabgänge werden erhoben, um Einblick in die Entwicklung der Bautätigkeit zu bekommen. Aus ihren Ergebnissen wird berechnet, wie sich der Wohnungsbestand entwickelt; diese Daten stehen für Wirtschaft, Forschung und Städteplanung bereit. Zudem dienen diese Erhebungen als Indikator, um die Wirtschaftsentwicklung im Bausektor zu bewerten, und damit der kurzfristigen Konjunkturbeobachtung in Deutschland. Die Statistik der Baugenehmigungen ist eine monatliche Gesamterhebung der Statistischen Ämter des Bundes und der Länder, die mithilfe der Bauaufsichtsbehörden und der Bauherrinnen sowie Bauherren durchgeführt wird. Die Statistiken der Baufertigstellungen und der Bauabgänge haben eine jährliche Periodizität.

Rechtsgrundlagen für die Erhebungen sind das Hochbaustatistikgesetz (HBauStatG) in Verbindung mit dem Bundesstatistikgesetz (BStatG).

Es besteht Auskunftspflicht (§ 6 Absatz 1 Satz 1 HBauStatG), deren Einhaltung den Bauherrinnen und Bauherren sowie den Bauaufsichtsbehörden obliegt (§ 6 Absatz 2 HBauStatG). Zwar sind die Einzelangaben geheim zu halten (§ 16 BStatG), jedoch dürfen faktisch anonymisierte Daten an Hochschulen und sonstige Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung weitergegeben werden (§ 16 Absatz 6 BStatG).

1 Die von der Bauministerkonferenz verabschiedete Musterbauordnung dient als Grundlage für die Umsetzung in spezifisches Landesrecht. Sie entfaltet somit keine unmittelbare Rechtswirkung. Jedes Land entscheidet, in welchem Umfang die Landesregelung dem Muster folgt. Die aktuelle, zuletzt durch Beschluss der Bauministerkonferenz vom 23./24. November 2023 geänderte Musterbauordnung steht auf deren Website zur Verfügung.

3

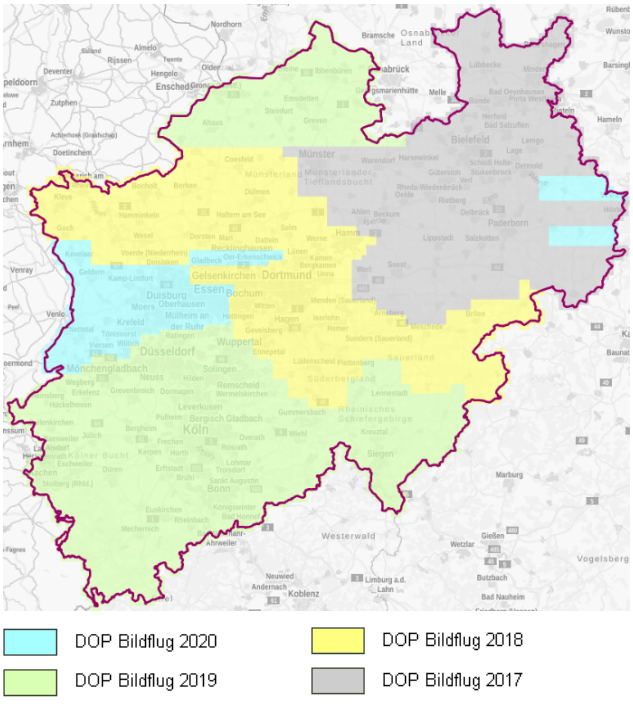
Daten

Die verwendeten Daten zur Analyse sind digitale Orthophotos, die die Bezirksregierung Köln öffentlich zur Verfügung stellt. Die Orthophotos dienen als Planungsgrundlage im Straßenbau, im Umweltschutz und in der Land- und Forstwirtschaft. Zusätzlich können sie bei Schadensfällen, zum Beispiel bei Naturkatastrophen, herangezogen werden.

„Orthophotos sind hochauflösende, verzerrungsfreie, maßstabsgetreue, photographische Abbildungen der Erdoberfläche." (Bezirksregierung Köln, 2020a) Die Grundlage bilden Luftbilder, die von Luftfahrzeugen senkrecht zum Boden aufgenommen werden. Diese Bilder werden anschließend digital mithilfe eines Höhenmodells verarbeitet, um die digitalen Orthophotos zu produzieren. Diese Bilder werden für Nordrhein-Westfalen hergestellt und zusätzlich georeferenziert, das heißt die Bilder werden mit den richtigen Koordinaten versehen. Die Bodenauflösung beträgt zehn Zentimeter je Pixel und die Bilder werden mit vier Kanälen aufgenommen: dem Rot-, dem Grün- und dem Blau-Kanal für die Farbdarstellung sowie in Infrarot. In der diesem Beitrag zugrunde liegenden Arbeit werden nur die drei Farbkanäle verwendet.

Insgesamt wird so das Land Nordrhein-Westfalen alle drei Jahre fotografiert, in jedem Jahr ungefähr ein Drittel.[2] Welche Gebiete in welchen Jahren überflogen und veröffentlicht wurden, zeigt ↘ Grafik 1 zum Stand 10. November 2020. In den folgenden Monaten wurden weitere Daten für 2020 ergänzt, nachdem sie fertig verarbeitet und der Öffentlichkeit zur Verfügung gestellt wurden.

Grafik 1
Bildflugprogramm Nordrhein-Westfalen nach Jahren



Quelle: Bezirksregierung Köln, 2020b.
DOP = digitale Orthophotos

4

Methoden

4.1 Neuronale Netze

(Künstliche) Neuronale Netze basieren auf dem Konzept der synaptischen Verbindungen im Nervensystem. Ein Neuronales Netz besteht aus sogenannten Knoten und Gewichten, die zusammen Neuronen simulieren. Ein solches Netz wird in Schichten organisiert. Es gibt eine Eingangsschicht (Input Layer), eine Ausgangsschicht (Output Layer) und Zwischenschichten (Hidden Layers). Jede Schicht besteht aus mehreren Knoten. Für zwei aufeinanderfolgende Schichten gilt, dass jeder Knoten der Schicht mit jedem Knoten der folgenden Schicht durch ein Gewicht verbunden ist. Knoten einer Schicht sowie Knoten nicht aufeinanderfolgender Schichten sind nicht miteinander verbunden (Goodfellow und andere, 2016).

4.2 Convolutional Neural Network

Ein Convolutional Neural Network (CNN) ist eine besondere Form des Neuronalen Netzes. Es ist besonders für Daten mit räumlichen Strukturen geeignet. Anders als bei Neuronalen Netzen mit vollständig verbundenen aufeinanderfolgenden Schichten benutzen CNNs

2 Im Jahr 2020 wurde auf einen Zwei-Jahres-Rhythmus umgestellt.

ein kleines gleitendes Raster, um lokale Eigenschaften zu extrahieren (Convolution) und dann diese Eigenschaften mit dem sogenannten Pooling zu aggregieren. Damit wird die Zahl der Parameter stark reduziert, es können dennoch effizient räumliche Eigenschaften in unterschiedlichen Regionen des Bildes gefunden werden. Außerdem sind die gleitenden Raster translationsinvariant, das heißt Merkmale können unabhängig von ihrer Position im Bild gefunden werden (Goodfellow und andere, 2016).

4.3 Vortrainierte CNNs

Es hat sich etabliert, für die Klassifikation von Bilddaten auf CNNs zuzugreifen, die schon auf anderen Datensätzen trainiert und optimiert wurden (Castelluccio und andere, 2015; Gardner/Nichols, 2017; Papadomanolaki und andere, 2016; Perez und andere, 2019).

Beispielsweise nutzten Gardner und Nichols (2017) ein CNN-Modell zur Klassifizierung von Satellitenbildern des Amazonas, um die Abholzung nachzuverfolgen. Es lagen 40 000 Bilder aufgeteilt in 17 Klassen vor. Drei vortrainierte CNNs wurden als Basis benutzt: VGG16, Inception-v3 und ResNet50. ResNet50 erbrachte die besten Ergebnisse mit einer Genauigkeit von 90,8 %. Inception-v3 und VGG-16 erreichten eine Accuracy von 90,1 % beziehungsweise 89,7 %.

Aufgrund der guten Ergebnisse mit Daten basierend auf Satellitenbildern wurde entschieden, die CNNs VGG, ResNet50 und Inception für die Analyse zu benutzen. Alle drei liefern gute Ergebnisse in ähnlichen Anwendungsgebieten und sind in Python Tensorflow verfügbar.

4.4 Stand der Forschung

Convolutional Neural Networks (CNNs) wurden bei der Klassifizierung von Satelliten- und Luftbildern von Regionen nach Naturkatastrophen, zum Beispiel Erdbeben, benutzt, um beschädigte Gebäude zu erkennen und Hilfsaktionen lenken zu können (Duarte und andere, 2018; Ji und andere, 2018). In Duarte und andere (2018) konnte basierend auf Daten aus Italien, Ecuador und Haiti schon eine Genauigkeit von über 90 % erreicht werden. Insbesondere führte hier die Kombination von Satelliten- und Luftbildern mit unterschiedlichen Auflösungen zu verbesserten Ergebnissen. Auch in Ji und andere (2018) wurden Luftbilder von Haiti verwendet, und es konnte mit SqueezeNet eine Genauigkeit von 80 % erreicht werden.

Außerdem wurden CNNs verwendet, um Satellitenbilder pixelweise verschiedenen Klassen wie Vegetation, Gebäuden und Straßen zuzuordnen (Albert und andere, 2017; Maltezos und andere, 2017; Postadjiana und andere, 2017).

In Maltezos und andere (2017) wurde die pixelweise Klassifikation von CNNs auch mit zwei Arten von Support Vector Machines (SVMs) verglichen. Die SVMs konnten 73 % beziehungsweise 76 % der Pixel richtig vorhersagen und das CNN 82 %.

Des Weiteren wurden CNNs verwendet, um Straßenkreuzungen in Satellitenbildern zu erkennen (Barsia/Heipke, 2003) und um verschiedene Arten von Bautätigkeiten auf Baustellen zu klassifizieren (Luo und andere, 2018).

Aufgrund der vielversprechenden Ergebnisse mit Convolutional Neural Networks, insbesondere mit Daten, die auf Luftbildern basieren und sich mit Gebäuden beschäftigen, wurde beschlossen, für die Entwicklung des Algorithmus zur Erkennung von Luftbildern CNNs zu verwenden.

5 Analyse

5.1 Datensatz

Die diesem Beitrag zugrunde liegende Arbeit ist ein erster Versuch, Baustellen anhand von Luftbildern zu erkennen. Daher lag noch kein Datensatz mit klassifizierten Bildern vor und Baustellen mussten „per Hand“ gefunden werden. Das Programm Arc-GIS wurde verwendet, um die aktuellen Luftbilddaten der Bezirksregierung Köln zu laden und somit Baustellen ausfindig zu machen. Bei der Durchsicht der Daten wurden 2 912 Baustellen gefunden und deren Koordinaten gesammelt. Es wurde darauf geachtet, möglichst alle Regionen Nordrhein-Westfalens und alle Jahre von 2015 bis 2020 abzubilden sowie möglichst viele unterschiedliche Arten

Grafik 2
Beispielbilder mit Baustellen

von Baustellen (zum Beispiel einzelne Wohnhäuser oder gesamte Neubaugebiete) und die verschiedenen Stadien des Bauvorhabens.

Von den 2 912 Baustellen wurden etwa 800 auf Grundlage der Baugenehmigungsdaten der Jahre 2015 und 2016 gefunden.

Gut 2 000 Areale wurden von der Autorin aufgrund der Informationen, die bestätigten Baustellenbildern entnommen wurden, als wahrscheinliche Baustellen eingestuft.

Da die Luftbilder bis 2020 nur alle drei Jahre erstellt wurden, ist der Anteil der Gebäude in der Baugenehmigungsstatistik, die sich genau am Aufnahmetag in einem Zustand befinden, der sie eindeutig als Baustelle ausweist, eher gering. Weil die zur Verfügung gestellten Adressen aus den Jahren 2015 und 2016 stammen, waren viele dieser Baumaßnahmen vermutlich bereits abgeschlossen, als die Luftbilder in den Jahren 2018 bis 2020 erstellt wurden. Es ist daher zu vermuten, dass die Baustellen, die unabhängig davon gefunden wurden, in den Genehmigungen der folgenden Jahre auftauchen werden. Zudem ist anzumerken, dass es teilweise schwierig war, Baustellen und ländliche Betriebe auseinanderzuhalten. Baustellen weisen häufig Kräne, Bagger, herumliegende Baumaterialien und einen sandigen Boden auf. Diese Merkmale können leicht mit Traktoren und landwirtschaftlichen Materialien sowie ähnlich sandigem Boden verwechselt werden. Es sollte daher darauf geachtet werden, wie gut der trainierte Algorithmus mit diesen Ähnlichkeiten umgeht, da diese für Menschen schwer zu beurteilen sind.

Die Koordinaten für die Bilder ohne Baustellen wurden zufällig generiert, sodass das Bundesland Nordrhein-Westfalen gleichmäßig abdeckt ist. Insgesamt gibt es 6 880 Bilder ohne Baustellen. Mit den 2 912 Bildern, bei denen Baustellen zu erkennen sind, besteht der Datensatz für das Training der CNNs aus 9 792 Bildern – 30 % mit Baustellen und 70 % ohne Baustellen. ↘ Grafik 2 zeigt drei Bilder mit Baustellen in unterschiedlichen Stadien des Bauvorgangs, ↘ Grafik 3 drei Bilder ohne Baustellen, nämlich Wald, Häuser in Städten und Straßen.

Grafik 3
Beispielbilder ohne Baustellen

5.2 Training von vortrainierten CNNs

Im Folgenden werden alle Berechnungen mit Python 3.7.7 (van Rossum/Drake, 2009) durchgeführt. Zusätzlich wurde das Paket TensorFlow 2.3.0 verwendet.

Aufgrund der guten Ergebnisse, die vortrainierte CNNs auf ähnlichen Datensätzen erreichen konnten, wurde entschieden, dies auch für diese Anwendung zu verwenden. Ausgesucht wurden die drei folgenden: VGG-16, Inception-v3 und ResNet50, da diese auch in dem Python-Paket Tensorflow implementiert sind. Diese drei Modelle wurden komplett übernommen einschließlich der trainierten Gewichte. Zusätzlich wurden zwei Fully-Connected-Schichten mit 4096 Neuronen und ReLu-Aktivierungsfunktion angehängt sowie eine Schicht mit einer Softmax-Funktion für die binäre Klassifizierung. Außerdem wurde zwischen diesen Schichten jeweils eine Dropout-Schicht mit einem Dropout von 50 % hinzugefügt. Diese Architektur erwies sich nach ersten Versuchen als am besten geeignet. Beim Trainieren der verschiedenen Modelle wurden die Parameter der vortrainierten CNNs fest angenommen und nur die Gewichte der später hinzugefügten Layer wurden berechnet.

Für alle Modelle werden die Standardeinstellungen von TensorFlow übernommen, das heißt die Hyperparameter sind:

- $\alpha = 0{,}001$
- $ß_1 = 0{,}9$
- $ß_2 = 0{,}999$
- $B = 32$

Zusätzlich wurde die Epochenzahl auf 20 und das Early Stopping auf 10 Epochen ohne Verbesserung gesetzt. Der Datensatz wurde aufgeteilt, sodass 80 % zum Trainieren benutzt wurden und die restlichen 20 % zum Validieren. Insgesamt wurde jedes Modell mit dem gleichen Datensatz dreimal berechnet.

Bezogen auf die durchschnittliche Genauigkeit der drei Berechnungen für die Modelle ergab sich, dass VGG und ResNet mit festen Parametern die besten Modelle waren mit Genauigkeiten von 96,46 % beziehungsweise 98,54 %. Inception blieb mit 89,33 % unter diesen Genauigkeiten.

Insgesamt zeigen diese ersten Berechnungen, dass CNNs für die Erkennung von Baustellen ein sinnvolles Verfahren sind. Insbesondere VGG und ResNet mit festen Parametern liefern die besten Ergebnisse mit einer Accuracy von über 95 %, die schon innerhalb der ersten Epochen erreicht werden kann.

Deswegen wurde entschieden, diese beiden Algorithmen weiter zu verwenden, um die Hyperparameter zu optimieren mit dem Ziel, Verbesserungsmöglichkeiten in der Berechnung der Algorithmen zu finden und die Genauigkeit noch weiter zu erhöhen. Mit den Hyperparametern $\alpha = 0{,}00001$, $ß_1\,1 = 0{,}7$, $ß_2\,2 = 0{,}999$ und Batchgröße $B = 64$, sowie einer Reduzierung der Epochen auf 10 und Early Stopping auf 5 konnte somit die Accuracy für ResNet auf 99,74 % und für VGG auf 99,48 % verbessert werden.

5.3 Testen auf Testregionen

Um zu prüfen, ob die berechneten Algorithmen für unbekannte Regionen ähnlich gute Ergebnisse liefern, wurde ein neuer Datensatz erstellt. Dazu wurden die Stadt Gelsenkirchen und der Hochsauerlandkreis ausgewählt. Beide haben eine Einwohnerzahl von ungefähr 260 000 (Landesdatenbank NRW, 2020). Der Hochsauerlandkreis als ländliche Gegend sollte mit einer städtischen, hier: Gelsenkirchen, verglichen werden. Es wurden zufällige neue Bilder für die beiden Regionen generiert und von der Autorin jeweils rund 50 Baustellen gesucht, sodass ein Datensatz für Gelsenkirchen mit 986 Nicht-Baustellen und 58 Baustellen entstand sowie ein Datensatz für den Hochsauerlandkreis mit 1 034 Nicht-Baustellen und 50 Baustellen. Diese beiden Datensätze wurden dann jeweils mit dem Algorithmus 5 des VGG und des ResNet des vorherigen Abschnitts klassifiziert.

↘ Tabelle 1 zeigt die Accuracy des Testdatensatzes für VGG und ResNet, dargestellt für die Stadt Gelsenkirchen und für den Hochsauerlandkreis. ResNet liefert hier für beide Regionen bessere Ergebnisse (93 % beziehungs-

Tabelle 1
Accuracy (Genauigkeit) der Algorithmen VGG und ResNet

	Gelsenkirchen	Hochsauerlandkreis
	%	
VGG	88,89	97,14
ResNet	92,91	98,15

weise 98 %) als VGG (89 % beziehungsweise 97 %). Die Bilder von Gelsenkirchen werden für beide Algorithmen mit nur 89 % beziehungsweise 93 % Genauigkeit klassifiziert. Der Hochsauerlandkreis hingegen weist eine Genauigkeit von 97 % beziehungsweise 98 % vor. Das beste Ergebnis liegt bei ResNet auf der ländlichen Region vor.

Zu beachten ist hier, dass es jeweils nur rund 50 Bilder mit Baustellen gab und somit der Vergleich der Accuracy zu Bildern mit Baustellen mit um die 1 000 Bildern schwierig ist. Deswegen ist es wichtig, die Sensitivität und die Präzision zu betrachten. ↘ Tabelle 2 zeigt die Sensitivität der zwei Datensätze jeweils für das VGG- und das ResNet-Modell. Der Datensatz für Gelsenkirchen weist hier eine geringere Sensitivität (10 beziehungsweise 14 Prozentpunkte) als der Hochsauerlandkreis auf. ResNet liefert für den Hochsauerlandkreis das beste Ergebnis mit 98 %, das heißt 98 % der tatsächlichen Baustellen werden vom Algorithmus auch erkannt. In Gelsenkirchen werden nur 84 % der Baustellen erkannt.

Tabelle 2
Sensitivität der Algorithmen VGG und ResNet

	Gelsenkirchen	Hochsauerlandkreis
	%	
VGG	82,76	92,00
ResNet	84,48	98,00

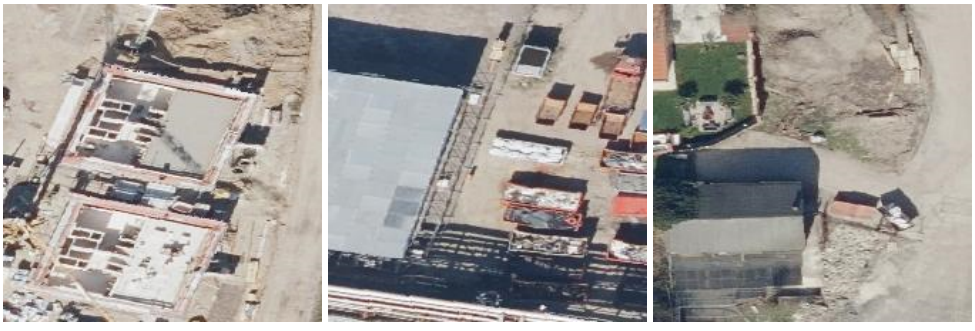
↘ Tabelle 3 stellt die Präzision für das VGG- und das ResNet-Modell auf Basis der Daten von Gelsenkirchen und dem Hochsauerlandkreis dar. Für das ResNet-Modell liegt die Präzision für beide Regionen um 10 Prozentpunkte über der des VGG-Modells. Beim VGG liegt die Präzision für Gelsenkirchen bei 31 % und für den Hochsauerlandkreis bei 63 %. Das heißt in Gelsenkirchen sind nur 31 % der vom Modell als Baustelle klassifizierten Bilder auch wirklich Baustellen. Für ResNet sind dies in Gelsenkirchen 43 %. Für den Hochsauerlandkreis liefert ResNet die größte Präzision mit 72 %. Diese niedrigen Werte sind darauf zurückzuführen, dass in den zwei benutzten Datensätzen 1 000 Bilder ohne Baustellen vorliegen. Selbst wenn diese mit einer Wahrscheinlichkeit von über 90 % richtig klassifiziert werden, bleibt eine hohe Anzahl von Bildern ohne Baustelle, die fälschlicherweise als Bild mit Baustelle klassifiziert werden. Bei dem Verhältnis Bilder von tatsächlichen Baustellen zu Bildern ohne Baustellen von 1 zu 20 führt selbst eine geringe Fehlerquote von 5 % dazu, dass genauso viele Baustellen fälschlicherweise identifiziert werden wie tatsächlich im Datensatz vorhanden sind. Um die Präzision zu erhöhen, könnte der Trainingsdatensatz um mehr Bilder aus städtischen Regionen erweitert werden.

Tabelle 3
Präzision der Algorithmen VGG und ResNet

	Gelsenkirchen	Hochsauerlandkreis
	%	
VGG	31,17	63,01
ResNet	42,98	72,06

↘ Grafik 4 zeigt eine Auswahl von Baustellen in Gelsenkirchen, die von beiden Algorithmen fälschlicherweise als Nicht-Baustelle klassifiziert wurden. Das linke Bild ist für das menschliche Auge recht leicht als Baustelle identifizierbar. Die anderen beiden Bilder könnten auch eine Mülldeponie beziehungsweise Boden ohne Gras darstellen. Da das Urteil des Menschen hier unsicherer als bei anderen Bildern ist, ist es möglich, dass ähnli-

Grafik 4
Fälschlicherweise nicht als Baustellen klassifizierte Baustellen in Gelsenkirchen



Grafik 5
Fälschlicherweise nicht als Baustelle klassifizierte Baustelle im Hochsauerlandkreis

che Bilder seltener im Trainingsdatensatz vorkommen und somit der Algorithmus deren Merkmale schlechter erlernen kann.

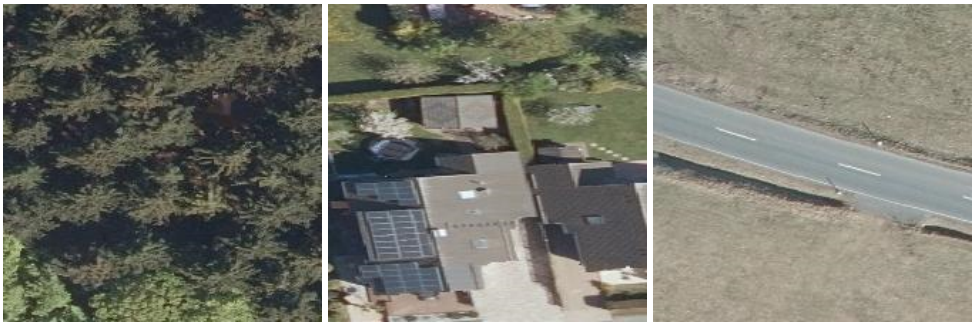
Auch die falsch klassifizierte Baustelle aus dem Datensatz für den Hochsauerlandkreis stellt eine Baustelle dar, die für das menschliche Auge schwer als Baustelle zu erkennen ist. Hier hat die Autorin die Umgebung genauer betrachtet, was zur Einstufung als Baustelle führte. ↘ Grafik 5

↘ Grafik 6 zeigt Bilder ohne Baustellen in Gelsenkirchen, die von beiden Algorithmen als Baustellen klassifiziert wurden. Für menschliche Beobachtende ist schnell offensichtlich, dass es sich hier nicht um Baustellen handelt. Eventuell führten Straßenmarkierungen oder die Schatten von Bäumen zu einer Falschklassifizierung. Das rechte Bild zeigt, dass ein Wohngebiet als Baustelle klassifiziert werden kann.

In ↘ Grafik 7 sind drei falsch klassifizierte Nicht-Baustellen des Hochsauerlandkreises abgebildet. Im linken sowie im rechten Bild ist nur Wald beziehungsweise eine Straße zu sehen, wohingegen im mittleren Bild eventuell der graue Boden als Zeichen für eine Baustelle interpretiert wurde.

Grafik 6
Fälschlicherweise als Baustellen klassifizierte Nicht-Baustellen in Gelsenkirchen

Grafik 7
Fälschlicherweise als Baustellen klassifizierte Nicht-Baustellen im Hochsauerlandkreis



6

Diskussion und Ausblick

Das Ziel der diesem Beitrag zugrunde liegenden Arbeit war zu prüfen, ob Convolutional Neural Networks benutzt werden können, um auf Orthophotos Baustellen zu erkennen. Die Ergebnisse der Auswertung verdeutlichen, dass CNNs sehr gut dafür geeignet sind, und es wurden dafür mehrere Algorithmen trainiert und verglichen.

Zu dieser speziellen Anwendung mit Orthophotos in Nordrhein-Westfalen existierte zuvor noch kein Datensatz, daher musste dieser selbstständig erstellt werden. Es konnten zwar die georeferenzierten Daten der Baugenehmigungsstatistik zu Hilfe genommen werden, aber es stand nur der Datensatz für die Jahre 2015 und 2016 zur Verfügung. Die aktuellsten Orthophotos stammen aus den Jahren 2018 bis 2020 und so war es schwierig, aufgrund dieses Datensatzes Bauprojekte zu finden, da die meisten zu diesem Zeitpunkt schon abgeschlossen waren. Deshalb ist mehr als die Hälfte der Bilder nicht mit Eintragungen in der Baugenehmigungsstatistik verbunden. Außerdem gelten rund 40 % der Baugenehmigungen nicht für Neubauten, das heißt viele Bauvorhaben (innerhalb von Gebäuden) sind nicht auf Orthophotos zu sehen, selbst wenn das Foto zum Zeitpunkt des Bauens aufgenommen wird. Zusätzlich ist zu beachten, dass die Aufnahme der Orthophotos in einem Rhythmus von drei beziehungsweise seit 2020 zwei Jahren erfolgt. Dadurch kann ein Großteil der Bauvorhaben, die zwischen diesen Terminen stattfinden, nicht auf den Orthophotos erfasst werden. Hinzu kommt, dass der Datensatz zum Trainieren der CNNs von einer einzigen Person erstellt wurde, wodurch womöglich eine Verzerrung in den Datensatz eingeflossen ist. Beispielsweise wurden vermehrt Baustellen in den Datensatz aufgenommen, die für das menschliche Auge einfach zu erkennen sind, und bei einer Unsicherheit, ob es sich wirklich um eine Baustelle handelt, oblag die Entscheidung der persönlichen Einschätzung der Autorin. Somit sind im Datensatz der Bilder mit Baustellen möglicherweise Aufnahmen vorhanden, die auf den Orthophotos wie eine Baustelle aussehen, aber in Wirklichkeit nicht genehmigungspflichtig sind. Beispiele finden sich bei landwirtschaftlichen Betrieben, Gartenarbeiten oder unaufgeräumten Abstellplätzen. Anzumerken ist außerdem, dass Bauarbeiten an Straßen oder Windrädern nicht in den Datensatz mit aufgenommen wurden. Es ist aber zu vermuten, dass diese trotzdem als Baustelle klassifiziert werden. Es ist auch möglich, dass deshalb mehr Baustellen aus städtischen Gebieten im Datensatz vorhanden sind, weil diese für eine Person schneller auffindbar sind.

Für die Analyse wurden drei vortrainierte CNNs ausgesucht, die auf ähnlichen Datensätzen vielversprechende Ergebnisse lieferten. Es gibt natürlich noch viele andere CNNs und andere Arten von Neuronalen Netzen, die man hier testen könnte. Zusätzlich wurde die Architektur mit Dropout von 0,5 und zwei Fully-Connected-Schichten nur an einem kleinen Teil des Datensatzes getestet. Hier wäre es theoretisch möglich, noch mehr Zeit zu investieren, um eine bessere Architektur zu finden. Allerdings lieferten die ersten Berechnungen insbesondere mit ResNet Genauigkeiten von über 95 %, sodass selbst mit anderen CNNs und Architekturen der Algorithmen nur geringe Verbesserungen zu erwarten wären.

Für die Optimierung der Hyperparameter wurden die zwei Algorithmen von VGG und ResNet verwendet, die bei den vorherigen Experimenten eine Genauigkeit über 96 % erreichen konnten. Danach wurden die Hyperparameter noch optimiert. Die besten Ergebnisse lieferte die Kombination: $\alpha = 0{,}00001$, $ß_1\,1 = 0{,}7$, $ß_2\,2 = 0{,}999$ und Batchgröße B = 64 sowohl für VGG und ResNet mit Genauigkeiten über 99 %, wobei auch hier ResNet das bessere Ergebnis lieferte. Insgesamt konnte durch die Optimierung der Hyperparameter eine Verbesserung von 3 % für das VGG-Modell und von 1,2 % für das ResNet-Modell erreicht werden.

Der Testdatensatz mit Bildern aus Gelsenkirchen und dem Hochsauerlandkreis lieferte eine Accuracy von über 85 % für Gelsenkirchen und über 90 % für den Hochsauerlandkreis. Es wurde erwartet, dass der Hochsauerlandkreis schlechter abschneiden würde als Gelsenkirchen, da im Trainingsdatensatz viele Baustellen mit Häusern in Neubaugebieten und in Städten vorhanden sind und vermutet wurde, dass im ländlichen Gebiet mehr landwirtschaftliche Betriebe zu finden sind, bei denen die Klassifikation schwierig sein könnte. Insbesondere die Präzision für die Bilder aus Gelsenkirchen ist sehr gering (VGG: 30 %, ResNet: 40 %). Es könnte eine Verbesserung für Gelsenkirchen erreicht werden, wenn der Datensatz um mehr Bilder aus Städten erweitert werden könnte. Eine andere Möglichkeit wäre, den

Bildausschnitt zu vergrößern, damit dem Algorithmus bei schwer zu klassifizierenden Bildern mehr Informationen zur Umgebung zur Verfügung stehen. Zusätzlich könnte ein vergrößerter Bildausschnitt dazu genutzt werden, den Algorithmus das gesamte Land klassifizieren zu lassen und so flächendeckend Baustellen zu erkennen. Dies wäre bei einem kleinen Bildausschnitt sehr viel aufwendiger. Allerdings ist zu beachten, dass dann der gesamte Datensatz, auf dem diese Arbeit aufbaut, geprüft werden müsste, denn es ist möglich, dass bei Bildern ohne Baustellen dann Baustellen im Ausschnitt liegen. Zusätzlich müssten Bilder von Neubaugebieten eventuell aussortiert werden, falls ein Bild mehrere Baustellen zeigt.

Ein weiteres Problem stellt die Verknüpfung mit der offiziellen Baugenehmigungsstatistik dar. Können durch die Ergebnisse dieser Arbeit nun flächendeckend die Baustellen auf den Orthophotos entdeckt werden, ist dennoch zusätzlich zu prüfen, ob diese Baustellen in der Statistik vorhanden sind. Das Bild gibt keine Information dazu, wann genau die Baugenehmigung erfolgt ist, sodass mindestens die Statistiken der drei vorherigen Jahre geprüft werden müssten. Hinzu kommt, dass die Verarbeitung der Baugenehmigungen einige Zeit beansprucht und teilweise die statistischen Erhebungsbogen nachgereicht werden. Das bedeutet, dass bei aktuellen Bildern einige Zeit abzuwarten ist, bis die dazugehörigen Baugenehmigungen vollständig geprüft werden können.

Darüber hinaus sind 22 % der Georeferenzierungen nicht gebäudegenau. Es ist zum Beispiel nicht zu unterscheiden, welche Baustelle genehmigt wurde, wenn für eine Straße mehrere Baustellen entdeckt wurden, aber nur eine offiziell gemeldet ist. Es ist dann auch nicht zu erkennen, ob eine Baustelle gemeldet wurde, die eventuell zum Zeitpunkt der Bildaufnahme gar nicht sichtbar war. Zudem wären solche Genehmigungen zuerst auszusortieren, bei denen es sich nicht um eine Neubebauung handelt und die Bautätigkeit innerhalb des Gebäudes stattfinden könnte – was nicht auf Orthophotos sichtbar sein kann.

Es wäre möglich, den Datensatz mit Satellitenbildern zu kombinieren. In Ji und andere (2018) führte die Kombination von Luft- und Satellitenbildern zu verbesserten Ergebnissen. Die Satellitenbilder liefern zwar eine schlechtere Auflösung, sind aber in kleineren zeitlichen Abständen verfügbar, sodass Baustellen in mehreren Stadien des Bauvorgangs erkennbar sein müssten.

Insgesamt sind die vielversprechendsten Vorschläge, den Trainingsdatensatz um Baustellen aus städtischen Regionen zu erweitern und für eine flächendeckende Analyse des Landes den Bildausschnitt zu vergrößern. Soll der Algorithmus auf anderen Regionen oder Ländern als Nordrhein-Westfalen getestet werden, wäre es sinnvoll, den Datensatz um einen kleinen Datensatz mit Bildern aus der jeweiligen Region zu erweitern. In Xu und andere (2019) konnte damit die Accuracy für neue Regionen um ein paar Prozentpunkte gesteigert werden.

Insgesamt hat die Autorin in ihrer Masterarbeit festgestellt, dass CNNs sehr gut geeignet sind, um Baustellen in Orthophotos zu erkennen.

LITERATURVERZEICHNIS

Albert, Adrian/Kaur, Jasleen/Gonzalez, Marta. *Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale*. 2017. [Zugriff am 4. Februar 2025]. Verfügbar unter: arxiv.org

Barsia, Arpad/Heipke, Christian. *Artificial neural networks for the detection of road junctions in aerial images*. ISPRS Archives. 2003. [Zugriff am 5. Februar 2025]. Verfügbar unter: www.pf.bgu.tum.de

Bezirksregierung Köln. *Digitale Orthophotos*. 2020a. [Zugriff am 20. Februar 2025]. Verfügbar unter: www.bezreg-koeln.nrw.de

Bezirksregierung Köln. *TIM-online*. 2020b. [Zugriff am 20. Februar 2025]. Verfügbar unter: www.bezreg-koeln.nrw.de/geobas

Castelluccio, Marco/Poggi, Giovanni/Sansone, Carlo/Verdoliva, Luisa. *Land Use Classification in Remote Sensing Images by Convolutional Neural Networks*. 2015. [Zugriff am 24. Januar 2025]. Verfügbar unter: arxiv.org

Duarte, Diogo/Nex, Francesco/Kerle, Norman/Vosselman, George. *Satellite Image Classification of Building Damages using Airborne and Satellite Image Samples in a Deep Learning Approach*. In: ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. Ausgabe IV-2/2018, Seite 89 ff.

Gardner, Daniel/Nichols, David. *Multi-Label Classification Of Satellite Images With Deep Learning*. 2017. [Zugriff am 24. Januar 2025]. Verfügbar unter: cs231n.stanford.edu

Goodfellow, Ian/Bengio, Yoshua/Courville, Aaron. *Deep Learning. An MIT Press book*. 2016. [Zugriff am 4. Februar 2025]. Verfügbar unter: www.deeplearningbook.org

Ji, Min/Liu, Lanfa/Buchroithner, Manfred. *Identifying Collapsed Buildings Using Post-Earthquake Satellite Imagery and Convolutional Neural Networks: A Case Study of the 2010 Haiti Earthquake*. In: Remote Sensing. Ausgabe 10(11)2018, Seite 1689. DOI: 10.3390/rs10111689

Landesdatenbank NRW. *Datenlizenz Deutschland – Namensnennung – Version 2.0*. 2020. [Zugriff am 11. März 2025]. Verfügbar unter: www.govdata.de

Luo, Xiaochun/Li, Heng/Cao, Dongping/Dai, Fei/Seo, JoonOh/Lee, SangHyun. *Recognizing Diverse Construction Activities in Site Images via Relevance Networks of Construction-Related Objects Detected by Convolutional Neural Networks*. In: Journal of Computing in Civil Engineering. Jahrgang 32, Ausgabe3/2018. DOI: 10.1061/(ASCE)CP.1943-5487.0000756

Maltezos, Evangelos/Doulamis, Nikolaos/Doulamis, Anastasios/Ioannidis, Charalabos. *Deep convolutional neural networks for building extraction from orthoimages and dense image matching point clouds*. In: Journal of Applied Remote Sensing. Jahrgang 11. Ausgabe 4/2017, Seite 1. DOI: 10.1117/1.JRS.11.042620

LITERATURVERZEICHNIS

Papadomanolaki, Maria/Vakalopoulou, Maria/Zagoruyko, Sergey/Karantzalos, Konstantinos. *Benchmarking deep learning frameworks for the classification of very high-resolution satellite multispectral data*. In: ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. Ausgabe III-7/2016, Seite 83 ff. [Zugriff am 6. Februar 2025]. Verfügbar unter: noa.gwlb.de

Perez, Husein/Tah, Joseph H. M./Mosavi, Amir. *Deep Learning for Detecting Building Defects Using Convolutional Neural Networks*. In: Sensors. Ausgabe 19(16)/2019, Seite 3556. [Zugriff am 6. Februar 2025]. Verfügbar unter: arxiv.org

Postadjiana, Tristan/Le Bris, Arnaud/Sahbi, Hichem/Mallet, Clémont. *Investigating the potential of deep neural networks for large-scale classification of very high-resolution satellite images. ISPRS Annals of the Photogrammetry*. In: Remote Sensing and Spatial Information Sciences. Ausgabe IV-1/W1/2017, Seite 183 ff. [Zugriff am 5. Februar 2025]. Verfügbar unter: isprs-annals.copernicus.org

Van Rossum, Guido/Drake, Fred L. *Python 3 Reference Manual*. Scotts Valley, 2009. [Zugriff am 20. Februar 2025]. Verfügbar unter: dl.acm.org

Xu, Joseph Z./Lu, Wenhan/Li, Zebo/Khaitan, Pranav/Zaytseva, Valeriya. *Building Damage Detection in Satellite Imagery Using Convolutional Neural Networks*. 2019. [Zugriff am 7. Februar 2025]. Verfügbar unter: arxiv.org

RECHTSGRUNDLAGEN

Bauordnung für das Land Nordrhein-Westfalen (BauO NRW) in der Fassung der Bekanntmachung vom 21. Juli 2018, zuletzt geändert durch das Gesetz vom 31. Oktober 2023 (GV. NRW. Ausgabe 2023 Nr. 31 Seite 1167).

Gesetz über die Statistik der Bautätigkeit im Hochbau und die Fortschreibung des Wohnungsbestandes (Hochbaustatistikgesetz – HBauStatG) vom 5. Mai 1998 (BGBl. I Seite 869), das zuletzt durch Artikel 3 des Gesetzes vom 8. August 2020 (BGBl. I Seite 1728) geändert worden ist.

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.

Herausgeber
Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung
Dr. Daniel Vorgrimler
Redaktion: Ellen Römer

Ihr Kontakt zu uns
www.destatis.de/kontakt

Erscheinungsfolge
zweimonatlich, erschienen im April 2025
Ältere Ausgaben finden Sie unter www.destatis.de sowie in der Statistischen Bibliothek.

Artikelnummer: 1010200-25002-4, ISSN 1619-2907

© Statistisches Bundesamt (Destatis), 2025
Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.