

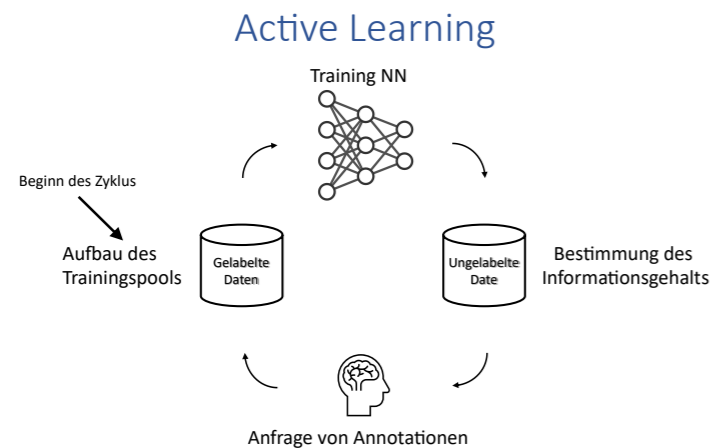
# Active Learning für die 3D-Objekterkennung in Punktwolken

Michael Trei

Referent: Prof. Dr. Andreas Weinmann | Korreferentin: Prof. Dr. Elke Hergenröther

## Einleitung

In der Objekterkennung sind Deep Learning Modelle zum Standard geworden. Durch das Lernen anhand von Beispielen können diese eine beeindruckende Leistungsfähigkeit erzielen. Der größte Nachteil des Deep Learning besteht darin, dass eine große Menge an annotierten Daten benötigt wird. Um die Menge an annotierten Daten zu reduzieren, wurde das Active Learning entwickelt. Das Ziel des AL ist es, die Daten mit dem größten Lerneffekt zu identifizieren und nur diese mit Annotationen zu versehen. Während das AL in der Objekterkennung in Bildern bereits Gegenstand vieler wissenschaftlicher Veröffentlichungen ist, besteht für die Anwendung von AL in der dreidimensionalen Objekterkennung noch Forschungsbedarf. Die Effektivität von Active Learning hängt signifikant von der eingesetzten Modellarchitektur sowie der Charakteristik des zugrundeliegenden Datensatzes ab. Infolgedessen führt eine Erweiterung des methodischen Repertoires zu einer verbesserten Adaptivität von AL auf neue Datensätze und Modellarchitekturen. Die vorliegende Arbeit befasst sich mit der Anwendung von Active Learning Methoden im Kontext der dreidimensionalen Objekterkennung in Punktwolken. Hierbei steht die Übertragbarkeit von AL- Algorithmen aus dem zweidimensionalen in den dreidimensionalen Bereich im Fokus.



Im Active Learning wird zunächst ein Modell auf einem initial annotierten Datensatz trainiert. Anschließend erfolgt die Quantifizierung des Informationsgehalts nicht annotierter Dateninstanzen, woraufhin für die informativsten Samples eine Annotation durch das Orakel angefordert wird. Die neu annotierten Daten werden dem Trainingsdatensatz hinzugefügt.

Dieser Prozess wird iterativ fortgeführt, bis entweder die gewünschte Modellperformanz erreicht, das Annotationsbudget ausgeschöpft oder der Pool nicht annotierter Daten vollständig verwendet wurde.

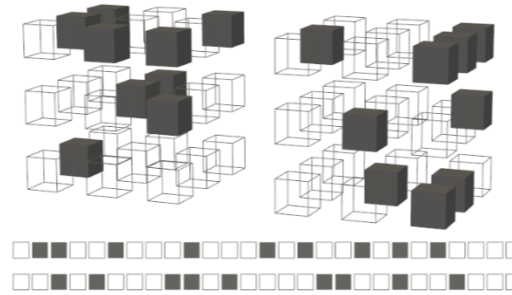
## Bestimmung des Informationsgehalts

Die Wahrscheinlichkeitsverteilung einer Klassifikation zu nutzen ist die einfachste und eine am häufigsten verwendete Möglichkeit den Informationsgehalt zu bestimmen. Durch die Verwendung der Entropie werden die Datenpunkte ausgewählt, welche nahe der Entscheidungsgrenze liegen. Hierdurch kommt es zu einer Verfeinerung der Klassifikationsgrenze, was zur Steigerung der Leistungsfähigkeit des Modells für ebendiese Fälle führt.

Zur Bestimmung des Informationsgehalts von Regressionsparametern lässt sich die Stabilität der Regressionsparameter durch die Analyse der Ausgabekonsistenz bei transformierten Eingabedaten quantifizieren. Das zugrundeliegende Prinzip basiert auf der Annahme, dass konstante Prädiktionen über verschiedene Datentransformationen hinweg auf eine robuste interne Repräsentation hinweisen. Signifikante Varianzen in den Vorhersagen deuten hingegen auf eine noch unzureichende Erfassung der inhärenten Struktur der Daten hin, was auf einen großen Informationsgehalt schließen lässt.

Diversitätsbasierte Verfahren konzentrieren sich auf die strukturelle Beschaffenheit der Trainingsdaten. Das Ziel dieser Verfahren ist es, anhand der Struktur der Daten eine repräsentative Teilmenge auszuwählen, die den Eingaberaum möglichst vollständig abdeckt. Wenn das Modell mit möglichst diversen Daten trainiert wird, führt dies zu einer gesteigerten Generalisierbarkeit, da weniger in unbekannte Bereiche extrapoliert werden muss. Hierfür wird häufig der Feature-Raum verwendet. Dieser bietet den Vorteil, dass dort abstrakte Repräsentationen vorliegen, bei welchen semantisch ähnliche Objekte nahe beieinander liegen. Mithilfe eines Distanzmaßes werden Datenpunkte ausgewählt, die sich von den bereits annotierten Daten maximal unterscheiden.

## Problematik bei der Anwendung von Diversitätsmaßen auf Punktwolken



Die Anwendbarkeit diversitätsbasierter Verfahren wird durch die charakteristische Dünnschichtigkeit der Daten signifikant eingeschränkt. Da

ein Großteil des Raums leer ist und folglich nur an wenigen Positionen Datenpunkte vorliegen, treten Features nur selten an identischen Positionen auf, wodurch die Berechnung paarweiser Ähnlichkeitsmaße erheblich erschwert wird.

## Lösungsansätze der Problematik

### Modellarchitektur-basiertes Vorgehen

Der Ursprung der Problematik liegt im Datenformat und in den dünnbesetzten Features. Der Einsatz einer adäquaten Modellarchitektur, welche durch ihren intrinsischen Aufbau das Auftreten der beschriebenen Problematik bereits auf struktureller Ebene verhindert, ist somit als vielversprechende Lösungsstrategie zu bewerten. Die Voraussetzung an die Modellarchitektur ist hierbei die Transformation der dünnbesetzten Eingabedaten in dichtbesetzte Feature-Repräsentationen. Diese Konvertierung ermöglicht die Anwendung derselben diversitätsbasierten Verfahren, wie sie bei zweidimensionalen Daten verwendet werden.

### Gradienten-basiertes Vorgehen

Ein weiterer Lösungsansatz liegt im Wechsel des betrachteten Merkmals. Die Verwendung von Gradienten als alternative Merkmalsrepräsentation bietet sich hierfür an. Die Implementierung gradientenbasierter Methoden bietet dabei den signifikanten Vorteil, dass für sämtliche Modellparameter Gradienten generiert werden können, die von der Anzahl der Eingabepunkte und den dünnbesetzten Daten unabhängig sind.

### Feature-Aggregation

Durch das Aggregieren der Features kann die Problematik beseitigt werden. Hierbei lassen sich zwei fundamentale Strategien differenzieren: die globale sowie die lokale Aggregation.

Die globale Aggregationsmethode zeichnet sich durch die Zusammenfassung sämtlicher Features zu einem einzelnen Feature-Vektor aus. Es ist jedoch anzumerken, dass diese Methodik erhebliche Einschränkungen aufweist. Insbesondere ist ein zunehmender Informationsverlust bei steigender Feature-Anzahl festzustellen, wodurch die Effektivität von Ähnlichkeitsmaßen reduziert wird.

Im Gegensatz hierzu steht die lokale Aggregationsstrategie, bei der die Feature-Aggregation auf räumlich begrenzte Bereiche beschränkt wird. Die Effektivität dieses Ansatzes ergibt sich aus der Aggregation einer geringeren Anzahl von Features in den jeweiligen lokalen Bereichen. Es ist jedoch zu berücksichtigen, dass die Implementation dieser Methodik die Entwicklung spezifischer Strategien zur Ähnlichkeitsbestimmung erfordert, da die variierende Anzahl lokaler Regionen die direkte Anwendung etablierter Ähnlichkeitsmaße limitiert.