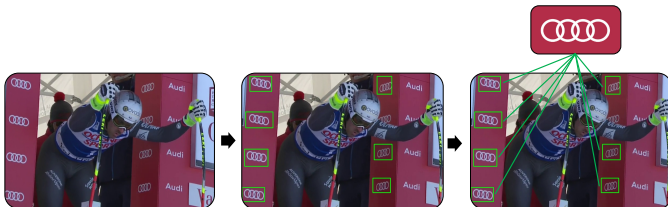


# Logo Suche in Massendaten mit Deep-Learning



**Andras Tüzkö**

**Masterarbeit Abschlusspräsentation am 20. Juni 2017**

Karlsruher Institut für Technologie (KIT)  
Institut für Anthropomatik  
Lehrstuhl für Interaktive Echtzeitsysteme (IES)

in Kooperation mit

Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung (IOSB)

**Betreuer: Dipl.-Inform. Christian Herrmann, Dipl.-Inform. Daniel Manger**

- ① Motivation und Aufgabenstellung
- ② Logo Retrieval
- ③ Das vorgestellte Logo Retrieval System
- ④ Logo Datensätze
- ⑤ Evaluation
- ⑥ Ausblick

**Statische Werbungen sind einer der wichtigsten Werbemethoden im Sport Bereich**

- Sponsoring von Teams
- Kauf von Werbeflächen

**Die Werbeflächen bedeuten große Ausgaben für die Firmen.**

**Messung der Effektivität ist gewünscht:**

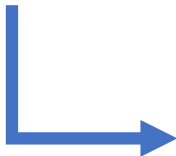
- Die Gesamtfläche eines Logos von einer Firma während einer Sendung,
- Zeit, das Logo zu sehen ist.

**Verwendung der gemessenen Daten um die Kosteneffizienz zu beurteilen.**

# Aufgabenstellung

Anfragebilder in den Einzelbildern von Sport Videos zu suchen

Query Set



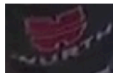
Search Set

# Herausforderungen

## Logos haben oft schlechtes Aussehen

- Teilsichtbarkeit
- Unschärfe
- Perspektivische Transformation, Rotation
- Unschärfe

## Aussehenvielfalt innerhalb Firmen



**Die Herausforderungen umwandeln das Problem in einer Open-Set Wiedererkennungsproblem**

### **Sliding window method**

- Langsam
- Nicht skalierungsinvariant

### **SIFT oder HOG Feature Extraktion**

- Keine Detektion von unbekannten Logos nur Wiedererkennung

### **CNN-basierter globaler Deskriptor**

- Von dem ganzen Bild z.B. Szene Wiedererkennung
- Ungeeignet für kleinere Objekte

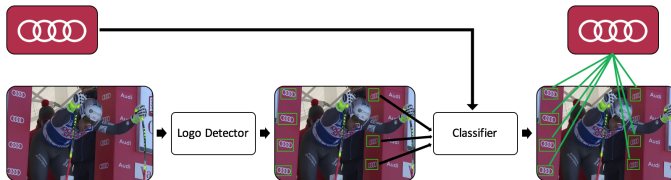
### Entwicklung der Proposal-basierten CNNs

- Regionen mit CNN
  - ▶ Vorschläge für Objektpositionen von externen Systemen
  - ▶ Berechnung der Deskriptoren von Regionen mit CNN individuell voneinander
- Fast R-CNN
  - ▶ Berechnung der Deskriptoren von Regionen gemeinsam mit Fully Convolutional Netzwerk
  - ▶ Ausschneiden von Regionen aus dem Feature-Map mit RoIPooling Schicht
- Faster R-CNN
  - ▶ Vorschläge für Objektpositionen von netzinternen Mini-Netz (RPN)
  - ▶ Das ganze System kann End-to-End trainiert werden

**State-of-the-Art Lösungen verwenden Proposal-basierten CNN Systeme für Closed-Set Logo Retrieval**

## Die Lösung ist in zwei Teilen aufgeteilt

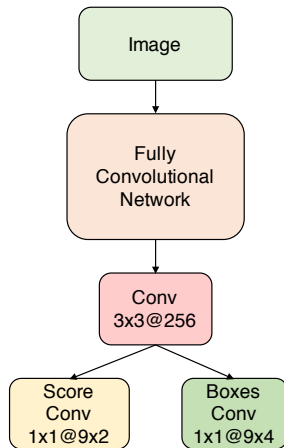
- Logo Detektion
  - Der Detektor ist trainiert alle Arten von Logo-Bilder zu erkennen
  - Funktioniert ohne Anfragebilder
- Logo Vergleich
  - Die Feature Vektoren von Reigonen werden extrahiert
  - Vergleich miteinander





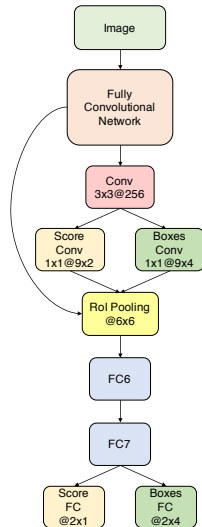
### RPN Logo Detektion

- Wird gleichzeitig mit dem Faster R-CNN Netz trainiert
- Aus einem schon trainierten Netz extrahierbar
- Sucht auf dem Feature Map wie die Sliding-Window-Methode
- Vordefiniere Anzahl von Anchor-Boxen
- Die sind mit Regression finegetuned



### Klassenagnostischer Logo Detektor

- Faster R-CNN, trainiert für zwei Klassen
- RPN Teil kann als schwacher Klassifikator betrachtet werden
- Zusammen ergeben eine Kaskade von Detektoren

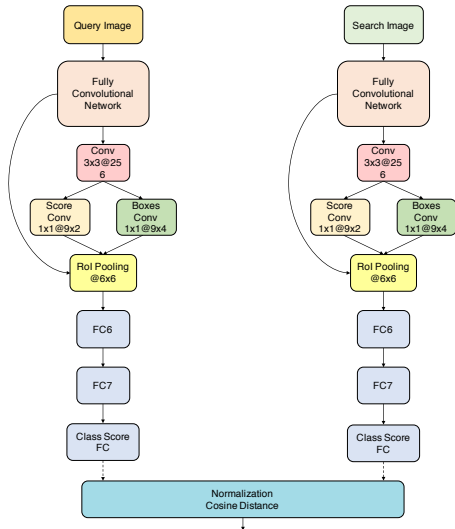


# Logo Retrieval System

## Logo Vergleich

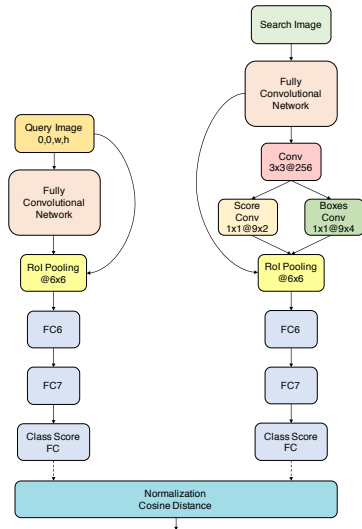
### Faster-Logos - Baseline

- Faster R-CNN Feature Extraktion sowohl von der Anfrage- als auch von der Suchbildern
- Klassen-wahrscheinlichkeiten werden nur für Features benutzt
- Die Wahrscheinlichkeiten sind nicht geeignet um Logo auszusuchen
- Die Score Ausgabe von RPN wird dafür benutzt
- Für Anfragebild der Bounding-Box mit der höchsten Wahrscheinlichkeit



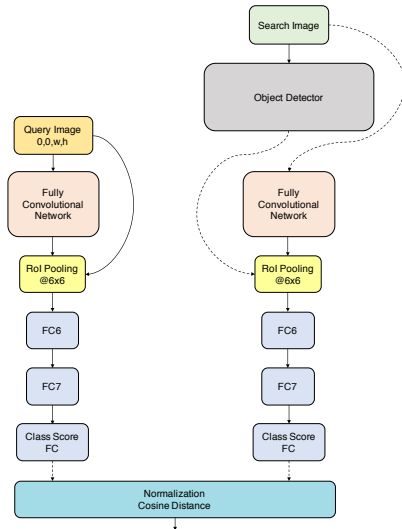
### Fast&Faster-Logos

- Oft falsche Bounding-Box-Vorhersage von Anfragebild
- Fast R-CNN Inferenz auf das gesamte Anfragebild



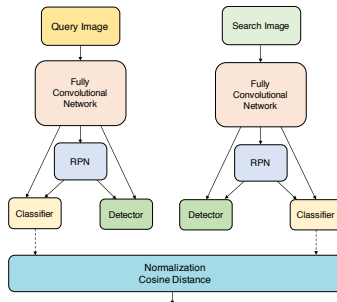
### Fast-Logos

- Vorherige Lösungen benutzten nur RPN Logo Detektor
- Verwenden des besten vorgeschlagenen Logo-Detektors für externe Generierung von Objektpositionsvorschlägen



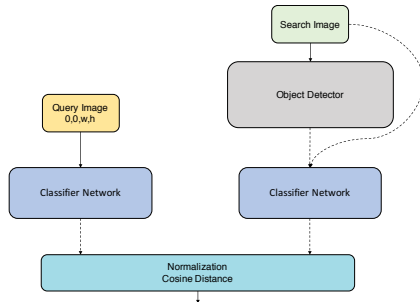
### Siam-Logos

- Siamesisches Netz, trainiert gemeinsam sowohl für Detektion als auch für Klassifikation
- Gewichte von FCN und RPN sind geteilt zwischen die Äste des Netzes
- Die Ausgabe des Detektors ist für Objektness Score benutzt
- Das Netz kann weitertrainiert werden, wenn es noch zusätzliche Logo-Daten ohne spezifische Brand Label gibt



### R-CNN-Logos

- Feature Extraktion mit Klassifikator Netzwerk
- Alle Arten von CNNs einsetzbar



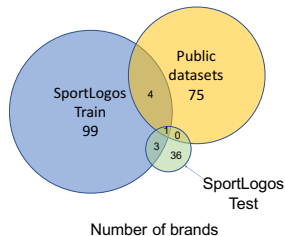
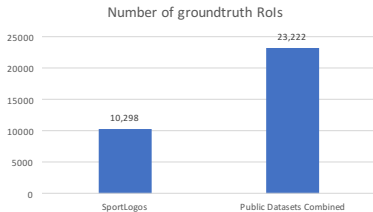
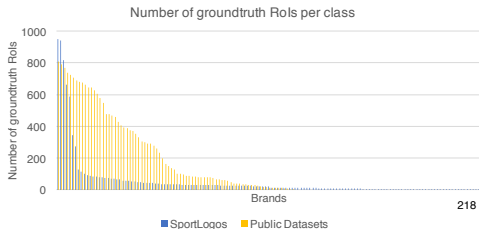
# Logo Datensätze

## Öffentliche Datensätze

- BelgaLogos-32, FlickrBelgaLogos, Flickr Logos 27, FlickrLogos-32, Logos-32Plus, TopLogo-10

## Selbst-annotierte Datensätze - SportLogos

- Football-1, Football-2, Ski, IceHockey





## Verwendeter Test Set

- SportLogos: Football-2

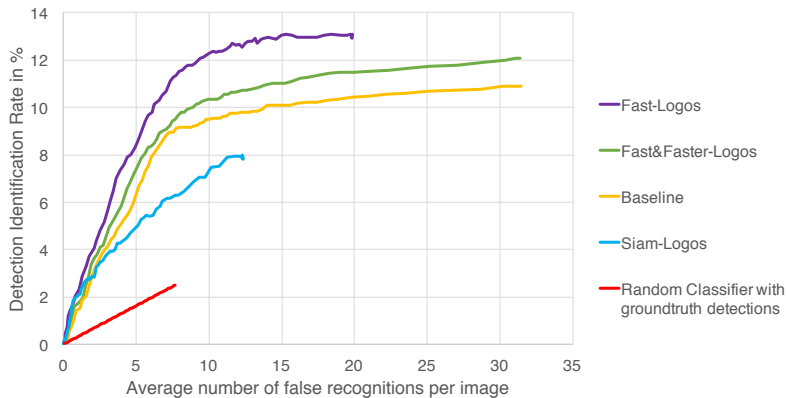
## Anfragebilder

- Von dem Video ausgeschnitten

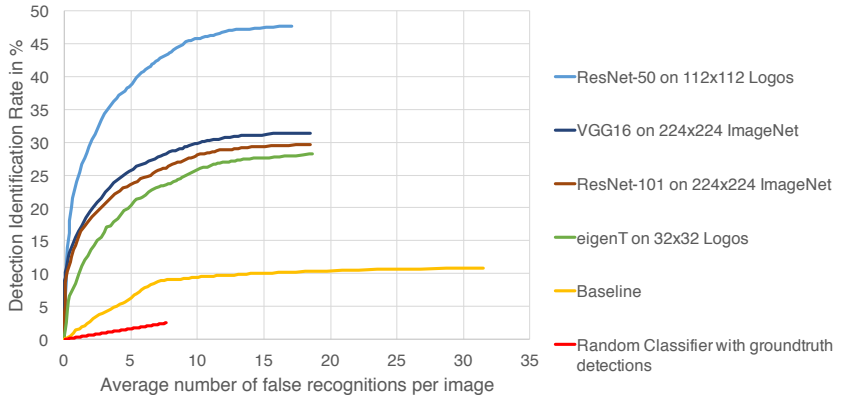
## Detektion und Identifikation Rate Kurve

- Gibt einen holistischen Überblick über Leistung des gesamten Retrieval-Systems
- Dargestellt über die Durchschnittsanzahl von falschen Klassifikationen

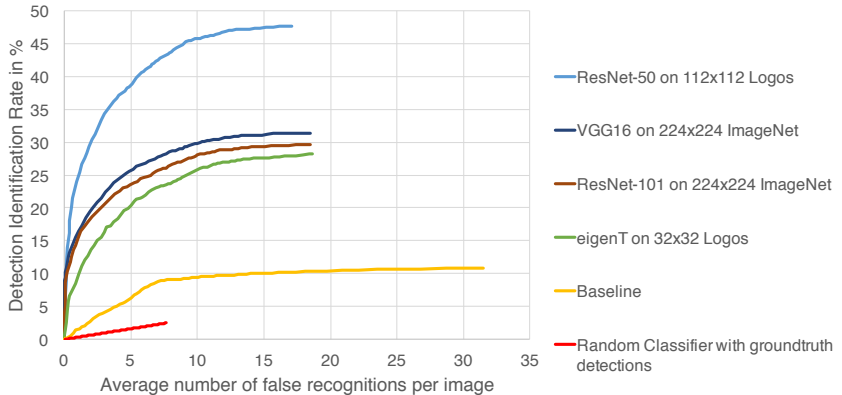
## Lösungen vs. Baseline



## R-CNN-Logos vs. Baseline



## Verarbeitungszeit pro Bilder und beste DIR Werte



## Textbasierte Logos

- Überwiegende Mehrheit der Logos ist textbasiert
- Erweiterung des Systems mit Texterkennung Subsystem

## Logo Tracking

- Ergänzung bei fehlender Detektion
- Reduzierung einmaliger Fehlklassifikation der getrackten Objekte

## Mehr Daten

- Die Größe der fusionierten Logo Datensätzen zu klein für Deep Learning

## MAC - Maximum Activation of Convolutions

- Gängige Feature Extraction Methode für Objekt- und Szene Retrieval
- Gesucht wird der maximum Wert von allen Channels, der Vektor wird als Feature benutzt

Ende



# Anhang

# Siam-Logos Train Phase

