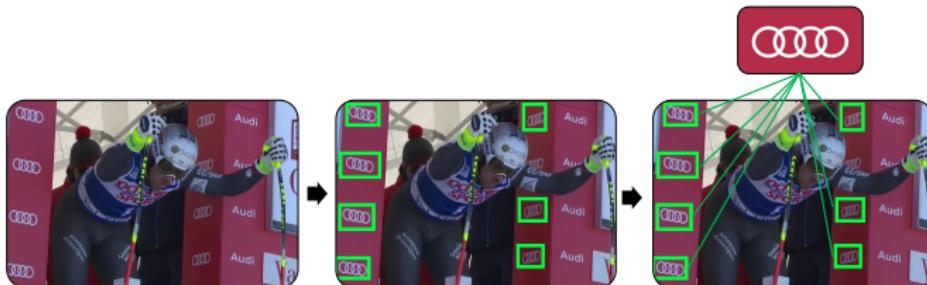


# Logo Retrieval in Massendaten mittels Deep-Learning



Andras Tüzkö

Masterarbeit Abschlusspräsentation am 23. 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

# Überblick

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

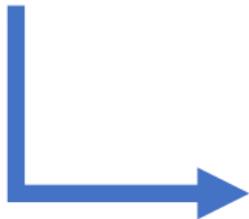
- Die Gesamtfläche eines Logos während einer Sendung
- Zeit, solange das Logo zu sehen ist

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

# Aufgabenstellung

Anfragebilder in den Einzelbildern von Sportvideos zu suchen

Query Set



Video

# Herausforderungen

Logos haben oft schlechte Qualität



Teilsichtbarkeit

Unschärfe

Perspektivische  
Transformation,  
Rotation

Ambiente  
Beleuchtungsänderung

Aussehenvielfalt  
innerhalb  
Firmen

Die Herausforderungen wandeln das Problem in ein Open-Set  
Wiedererkennungsproblem um

# Logo Retrieval

## Lösungsalternativen

### Sliding window method

- Langsam
- Nicht skalierungs invariant

### SIFT<sup>0</sup> oder HOG<sup>1</sup> Feature Extraktion

- Seit den letzten Jahren CNNs haben in vielen Problemen bessere Performance

### CNN-basierter globaler Deskriptor

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

[0] Distinctive Image Features from Scale-Invariant Keypoints [Lowe2004]

[1] Histograms of Oriented Gradients for Human Detection [Dalal2005]

# Logo Retrieval

## Gewählte Lösung

### Proposal-basierte CNNs

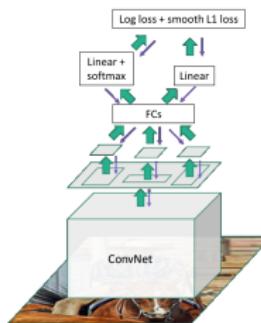
- State-of-the-Art Lösung<sup>2</sup> für Closed-Set Logo Retrieval

- Fast R-CNN<sup>3</sup>

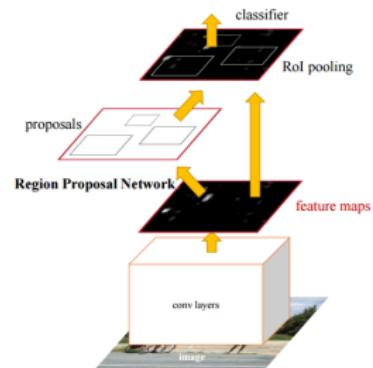
- Objektpositionen von externem System
  - FCN
  - RoIPolling

- Faster R-CNN<sup>4</sup>

- Objektpositionen von Teilnetz
  - End-to-End



Fast R-CNN



Faster R-CNN

[2] Fast R-CNN [Girshick2015]

[3] Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [Girshick2015]

[4] Region-based CNN for Logo Detection [Bao2016]

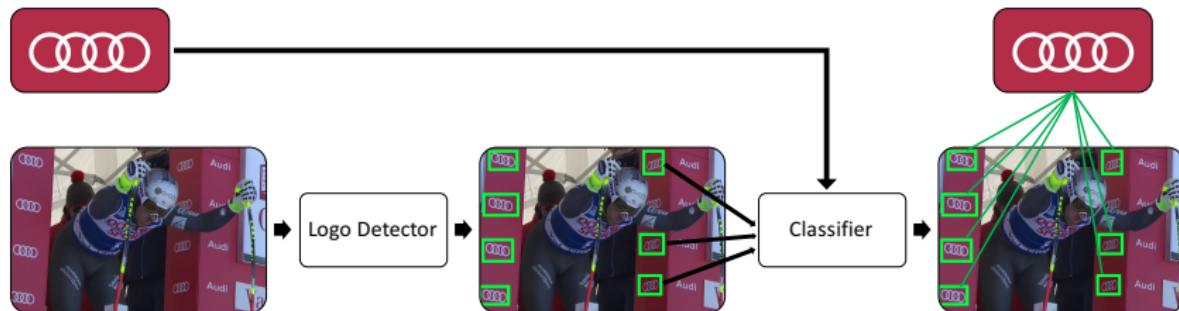
# Logo Retrieval System

## Das Problem ist anders

- Open-Set, weil verschiedene Logos im Training-Set als im Test-Set

## Die Lösung ist in zwei Teile aufgeteilt

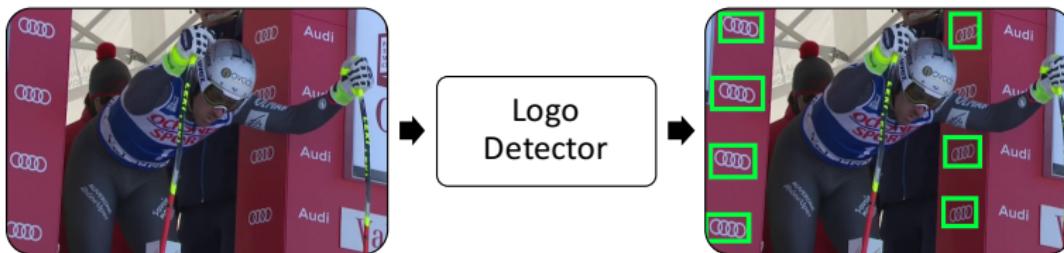
- Logo Detektion
  - ▶ Um alle Arten von Logo-Bildern zu erkennen
  - ▶ Funktioniert ohne Anfragebilder
- Logo Vergleich
  - ▶ Die Feature Vektoren von Regionen werden extrahiert
  - ▶ Vergleich miteinander



# Logo Retrieval System

## Logo Detektion

- Neun Architekturen trainiert und evaluiert
- Effekt von unterschiedlichen Training-Datensätzen untersucht
- Zwei Typen ausgewählt und vorgestellt

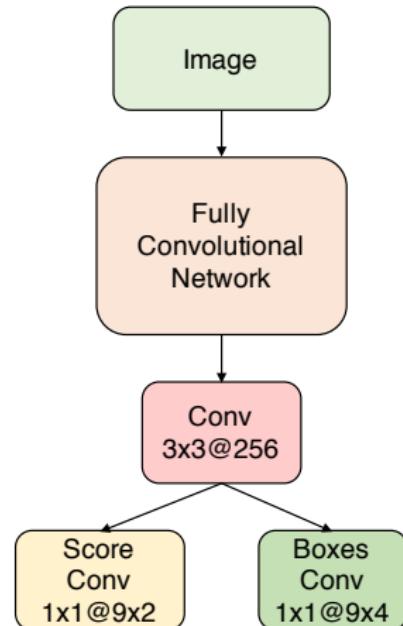


# Logo Retrieval System

## Logo Detektion

### Region Proposal Network Logo Detektor

- Wird gleichzeitig mit dem Faster R-CNN Netz trainiert
- Extrahierbar aus einem schon trainierten Netz
- Sucht auf der Feature Map in Sliding-Window-Fashion
- Vordefinierte Anzahl von Anchor-Boxen
  - ▶ Offset wird trainiert, angewendet auf den Anchor-Boxen

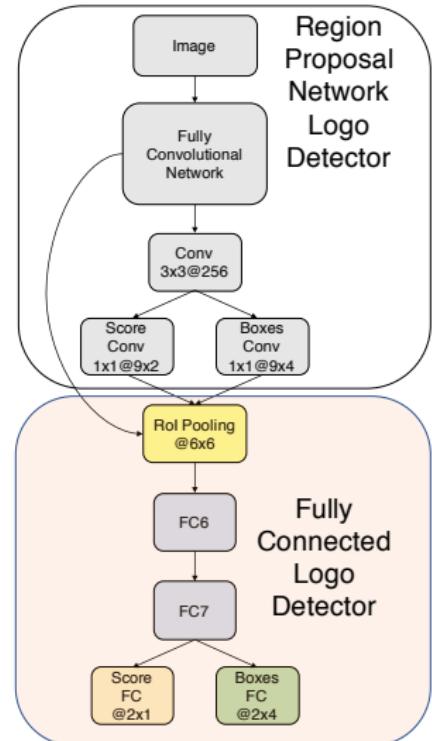


# Logo Retrieval System

## Logo Detektion

### Faster R-CNN Logo Detektor

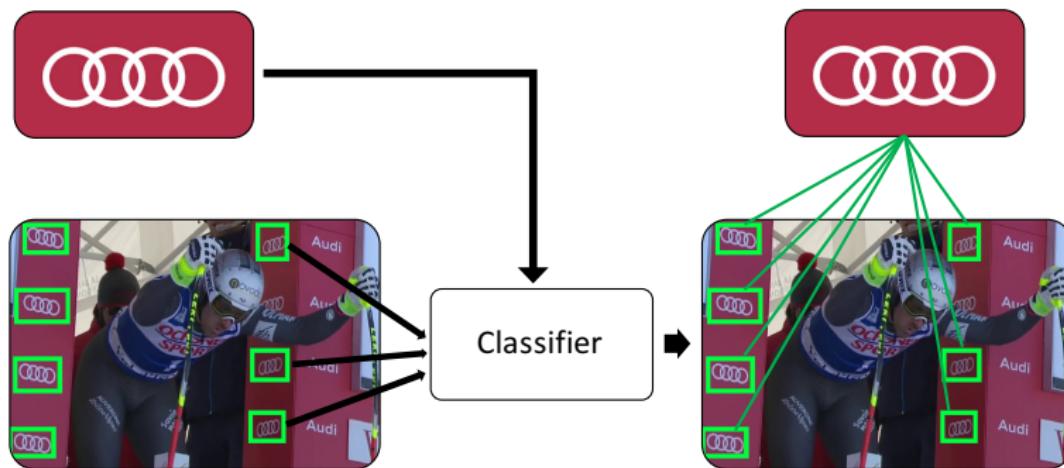
- Trainiert für zwei Klassen
- RPN Teil kann als schwacher Klassifikator betrachtet werden
- Zusammen ergeben eine Kaskade von Detektoren



# Logo Retrieval System

## Logo Vergleich

- Zwölf verschiedene Architekturen trainiert und evaluiert
- Effekt von unterschiedlichen Training-Datensätzen untersucht
- Vier Typen ausgewählt und vorgestellt

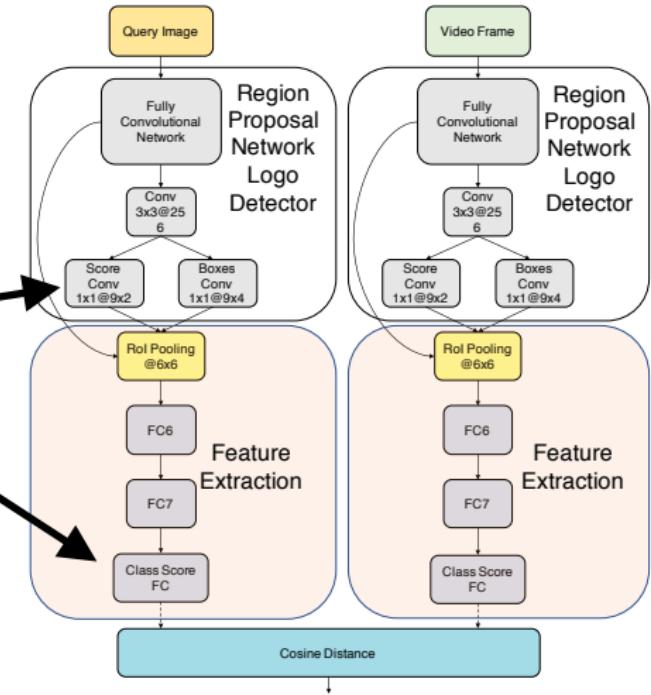


# Logo Retrieval System

## Logo Vergleich

### Faster-Logos - Baseline

- State-of-the-Art in Closed-Set Logo Retrieval
- Grundlage: Faster R-CNN
- Angepasst für Open-Set
- Die Score Ausgabe von RPN wird für Detektion benutzt
- Klassenwahrscheinlichkeiten werden als Features benutzt
- Für Anfragebild, die Bounding-Box mit der höchsten Detektionsscore

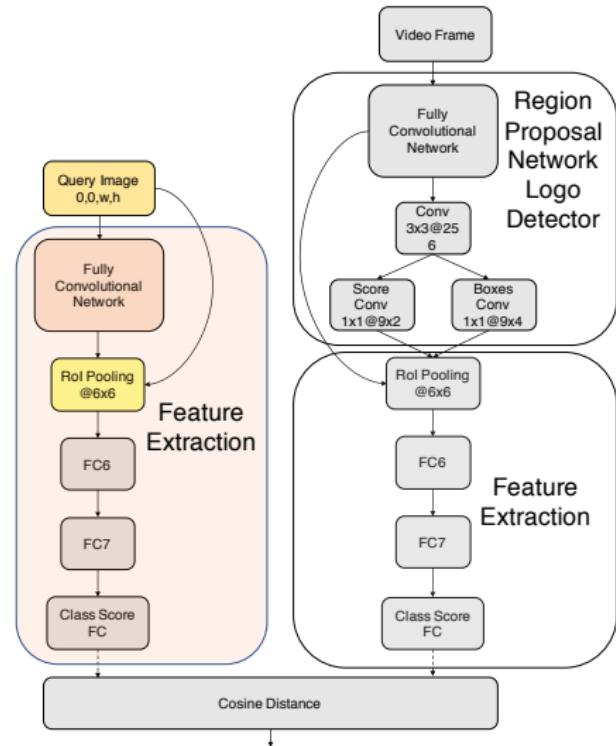


# Logo Retrieval System

## Logo Vergleich

### Komplettanfragebild

- Oft falsche Bounding-Box-Vorhersage von Anfragebild
- Entspricht für Fast R-CNN
- Nachteil: die Logos sollen gut zugeschnitten werden

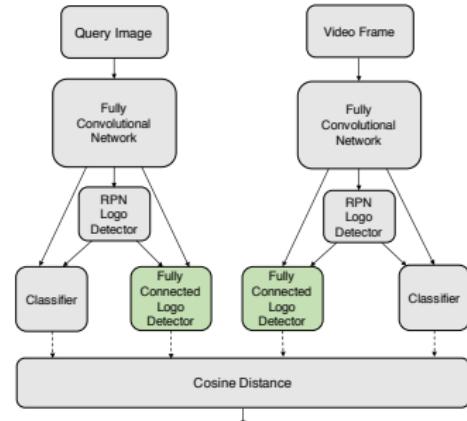


# Logo Retrieval System

## Logo Vergleich

### Siam-Logos

- Siamesisches Netz<sup>6</sup>, 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 spezifisches Brand-Label gibt



Test Phase

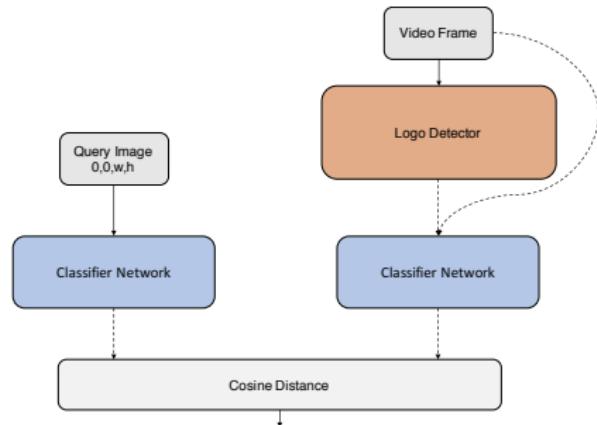
[6] Signature verification using a „Siamese“ time delay neural network [Bromley1993]

# Logo Retrieval System

## Logo Vergleich

### Separater Detektor und Klassifikator

- Probleme
  - ▶ Merkmale von Feature Map zwischen Detektion und Merkmalsextraktion geteilt
  - ▶ Keine Spezialisierung auf die eigentliche Aufgabe
- Idee: Trennung von Detektor und Klassifikator
- Vorteile
  - ▶ Alle state-of-the-art Netze verwendbar für Feature Extraktion
  - ▶ Flexibler in Bezug auf vortrainierten Netze als mit Fast-Faster R-CNN



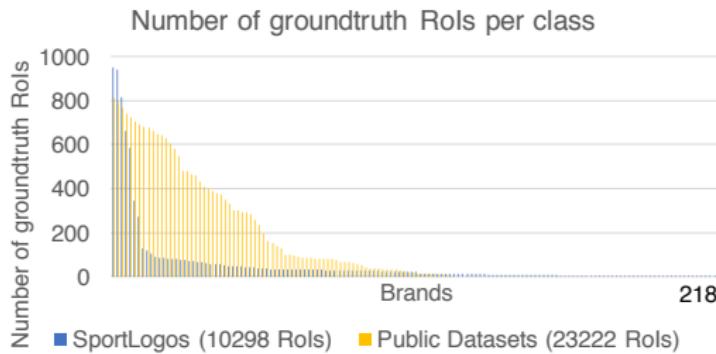
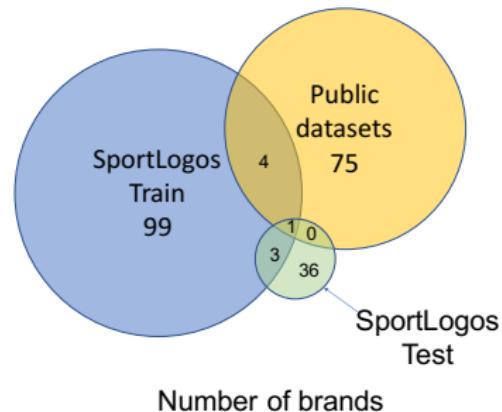
# Logo Datensätze

## Öffentliche Datensätze

- BelgaLogos-32<sup>7</sup>, FlickrBelgaLogos<sup>8</sup>, Flickr Logos 27<sup>9</sup>, FlickrLogos-32<sup>10</sup>, Logos-32Plus<sup>11</sup>, TopLogo-10<sup>12</sup>

## Selbst-annotierte Datensätze - SportLogos

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



# Logo Datensätze



BelgaLogos



Flickr Logos 27



FlickrLogos-32



Logos-32Plus



TopLogo-10



SportLogos

## Verwendeter Test Set

- SportLogos: Football-2

## Anfragebilder

- Von dem Video ausgeschnitten

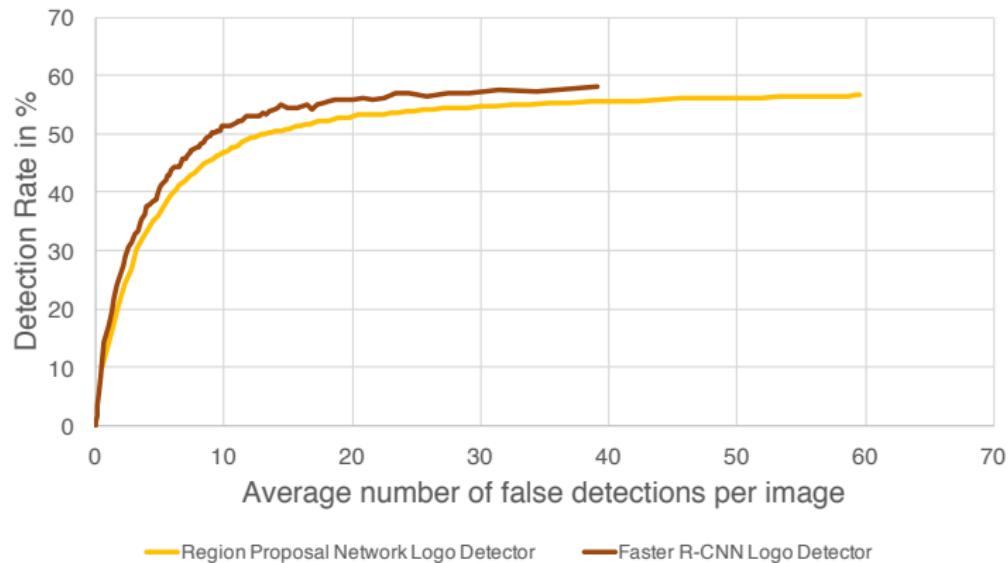
## Detection Rate

- Dargestellt über die Durchschnittsanzahl von falschen Detektionen
- Gleich mit Recall, True Positive Rate

## Detection Identification Rate

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

## Logo Detektion Vergleich

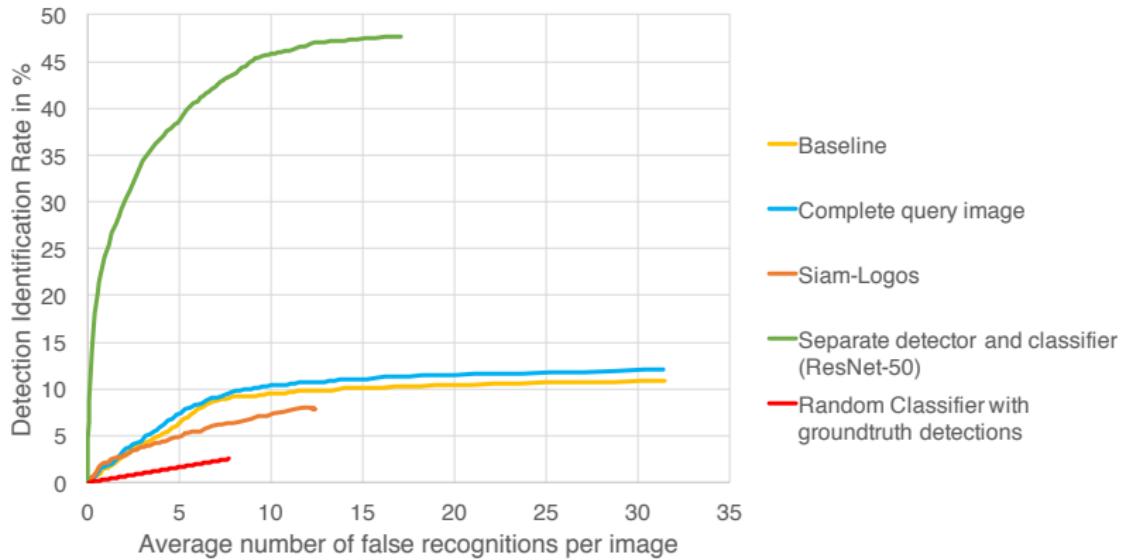


# Evaluation

## Logo Detektion Beispiel



## Logo Retrieval Vergleich



# Evaluation

## Logo Retrieval Beispiel



# Zusammenfassung und Ausblick

## Zusammenfassung

- Logo System vorgestellt für Open Set Retrieval
- Erreicht 47% auf dem Test Datensatz
- Schwierigkeiten
  - Schlechte Qualität
  - Aussehenvielfalt



## Ausblick

- 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 ist zu klein für Deep Learning

**Vielen Dank für Ihre Aufmerksamkeit.**

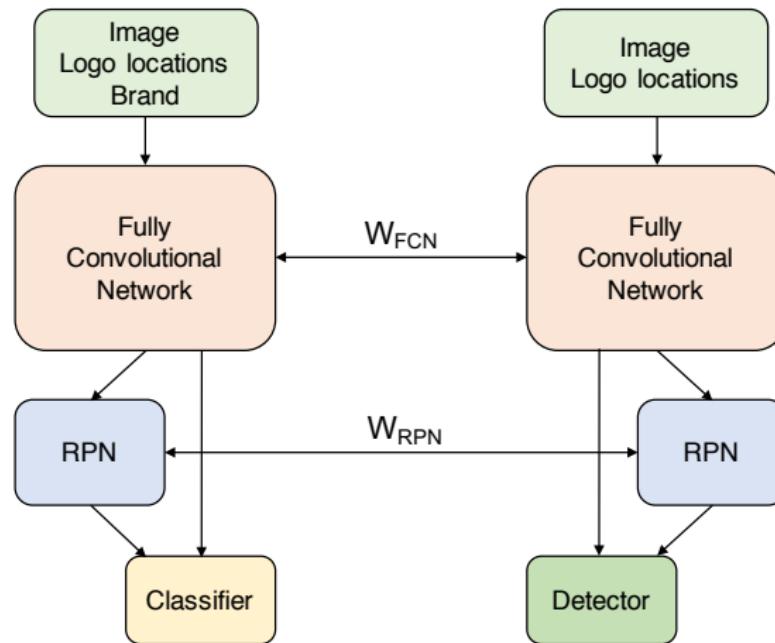


## Weitere Bibliografie

- [7] Logo retrieval with a contrario visual query expansion [Alexis2009]
- [8] Scalable mining of small visual objects [Alexis2012]
- [9] Scalable Triangulation-based Logo Recognition [Kalantidis2011]
- [10] Scalable logo recognition in real-world images [Romberg2011]
- [11] Deep learning for logo recognition [Bianco2017]
- [12] Deep Learning Logo Detection with Data Expansion by Synthesising Context [Su2016]

# Anhang

# Siam-Logos Test Phase



# Ähnlichkeitsscores

Query Image



Search Image



Query Images

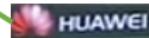
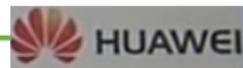


90.5%

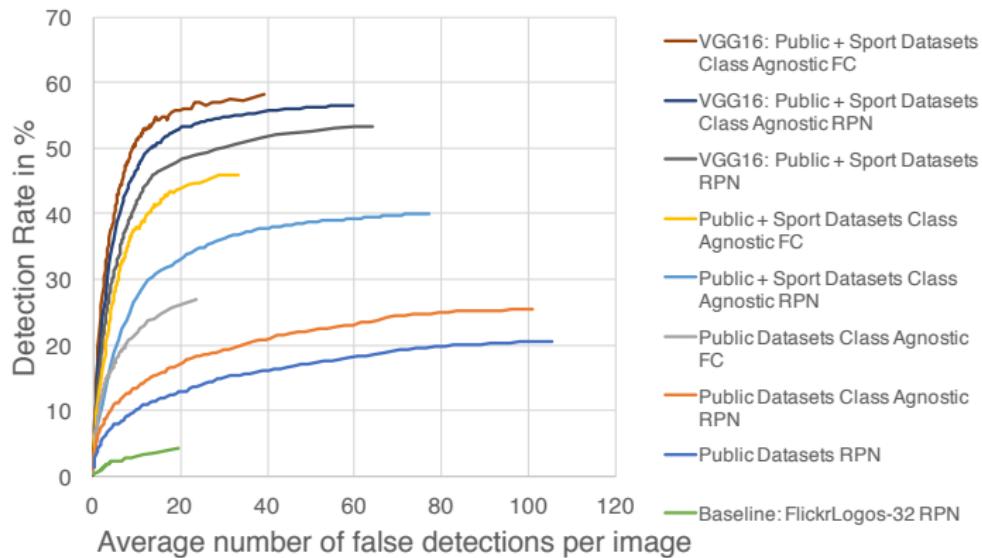
94.8%

65.4%

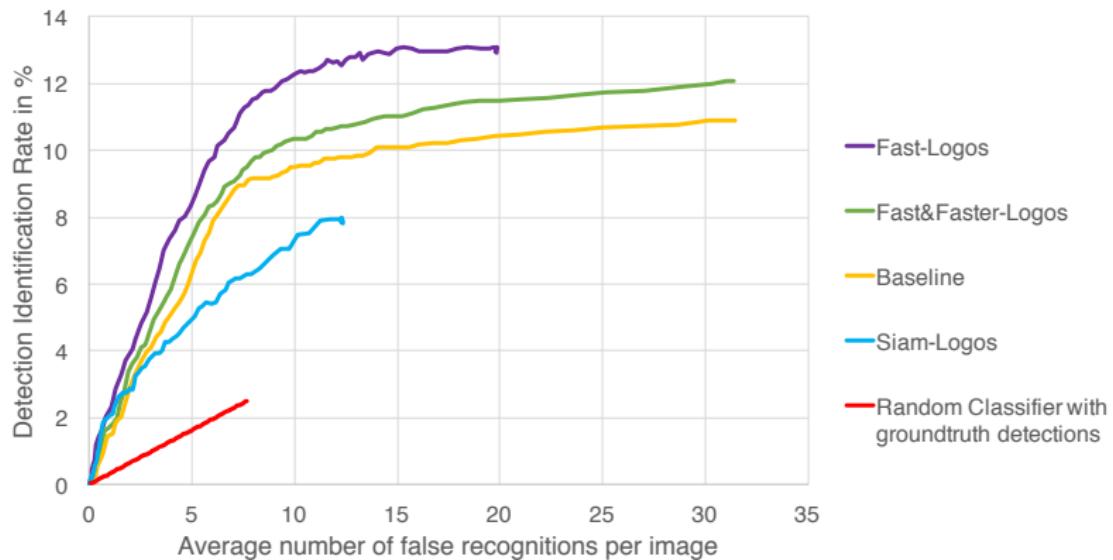
89.8%



## Logo Detektion

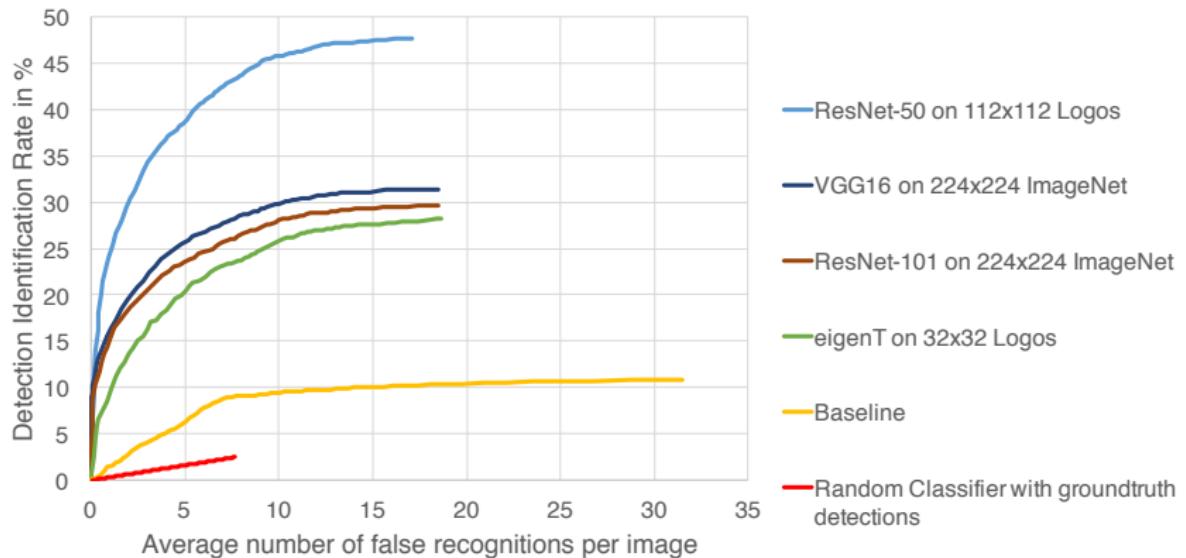


## Fast-Logos, Fast &Faster-Logos vs. Baseline



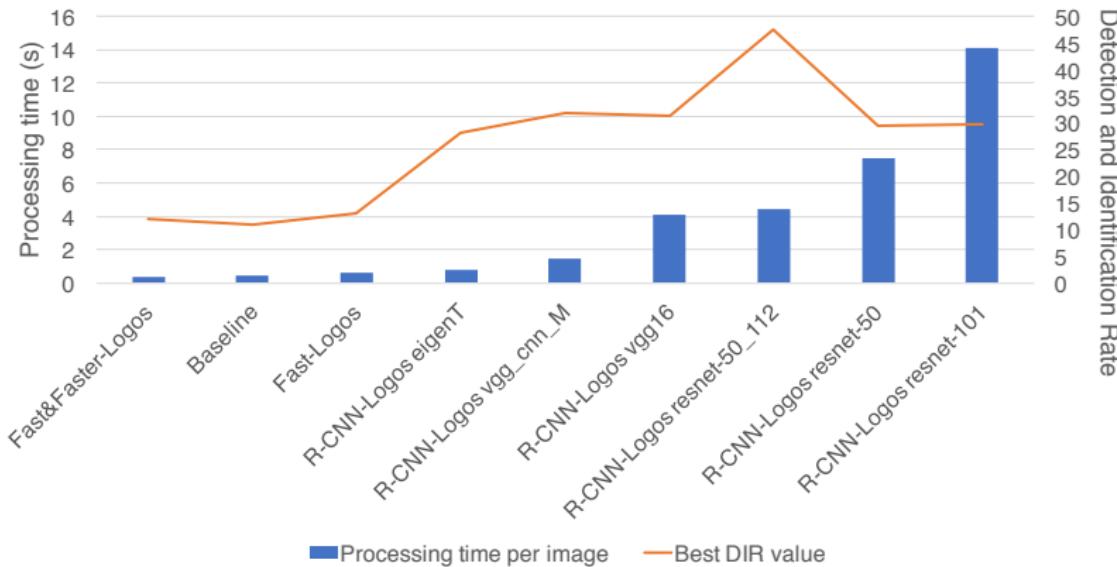
# Evaluation

## R-CNN-Logos vs. Baseline

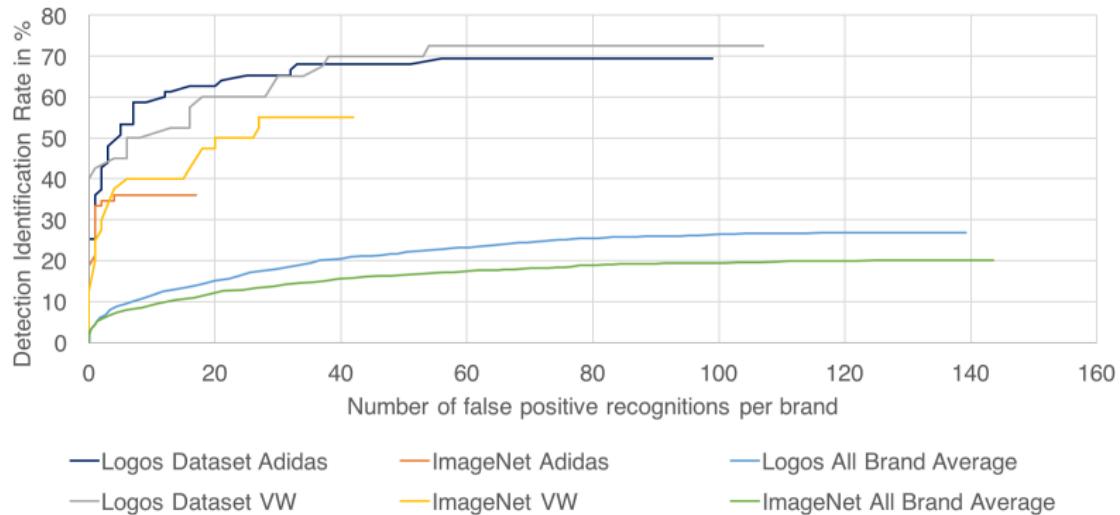


# Evaluation

## Verarbeitungszeit pro Bilder und beste DIR Werte



## Performance von Klassen, die in dem Training-Set vorkommen



## Logo Vergleich

