

Annotationsrichtlinien: Identifikation von Ziel-Objekten in Bildern

Simeon Junker
simeon.junker@uni-bielefeld.de

29. Januar 2024

Übersicht

Bei den Daten, die annotiert werden sollen, handelt es sich um Beschreibungen für Objekte in Fotografien, die von unterschiedlichen Systemen generiert wurden.

Die Ziel-Objekte für die Beschreibungen sind jeweils mit einer Bounding Box markiert. Annotiert werden soll, ob das jeweilige Ziel-Objekt vom System **korrekt identifiziert** wurde, also ob in der Beschreibung eine **valide Bezeichung des Typs des Ziel-Objekts** enthalten ist.

Typ-Bezeichnungen sind in den meisten Fällen als Nomen realisiert. Beispiel: Wenn die Bounding Box im Bild einen Hund markiert, würde eine Beschreibung wie „the dog on the right“ mit *dog* eine valide Bezeichnung umfassen, da das Ziel-Objekt vom Typ *Hund* ist. Mögliche Alternativen wären z.B. *animal* oder *pet*, da der abgebildete Hund auch in diese Kategorien fällt. „The cat“ müsste als falsch annotiert werden. Beschreibungen wie „it is on the right“, „upper left“, „the grey one“ oder „the thing“ beinhalten keine Typ-Bezeichnungen für das Ziel-Objekt und müssten entsprechend annotiert werden.

Für die Annotation sind ausschließlich die Typ-Bezeichnungen von Relevanz.

Es kann ignoriert werden, ob die Beschreibungen weiterer Eigenschaften (z.B. Farbe, Position im Bild, ...) wahrheitsgetreu sind. Der Satz „the dog on the left“ müsste hinsichtlich der Typ-Bezeichnung also auch dann als korrekt annotiert werden, wenn das markierte Objekt ein Hund, aber auf der rechten Seite des Bildes ist. Ebenso wäre „the red suitcase“ auch dann als korrekt zu annotieren, wenn das markierte Objekt ein Koffer, aber tatsächlich blau wäre.

Es sollen die Outputs mehrerer Systeme annotiert werden. Für jedes System existiert eine HTML- und eine dazugehörige CSV-Datei, jeweils mit gleichem Dateinamen (abgesehen von der Dateiendung). Die HTML-Datei beinhaltet Fotografien mit markierten Ziel-Objekten, eine ID, sowie eine generierte Beschreibung pro Objekt. Die CSV-Datei beinhaltet dieselben IDs und Beschreibungen in derselben Reihenfolge, sowie eine leere Spalte, in die die Annotation eingetragen werden soll.

Kategorien

Jedem Item soll **genau eine der folgenden vier Kategorien** zugewiesen werden:

- **A (Adequate)**: Die Beschreibung enthält eine **valide Typ-Bezeichnung** für das Target.
- **O (Omission)**: Die Beschreibung enthält **keine Typ-Bezeichnung** für das Target (z.B. aufgrund von Pronominalisierung, allgemeinen Nomen wie *thing*, oder der Beschränkung auf einzelne Eigenschaften wie *blue; large* oder der Position im Bild, z.B. *bottom left corner; on the right; top left*).
- **M (Misaligned)**: Die Beschreibung enthält eine **Typ-Bezeichnung für ein anderes Objekt** als das intendierte Target, welches jedoch auch (teilweise) von der Bounding Box erfasst ist.
- **F (False)**: Die Beschreibung enthält eine **falsche Typ-Bezeichnung**, die nicht auf das Target zutrifft (und sich auch auf kein anderes Objekt in der Bounding Box bezieht).

Beispiele für die einzelnen Kategorien sind am Ende dieser Datei. Als vollständige Beispiele bitte die Dateien *demo_file.csv* und *demo_file.html* anschauen.

Vorsicht: Wie bereits beschrieben sind bei der Kategorisierung **ausschließlich die Typ-Beschreibungen** zu beachten. Es ist unerheblich, ob andere Attribute oder Aspekte wahrheitsgetreu beschrieben sind (z.B. ob *the left dog* tatsächlich links im Bild ist, oder ob *the blue suitcase* tatsächlich blau ist). Ebenso soll ignoriert werden, ob sich eine Beschreibung auch auf andere Objekte im Bild beziehen könnte (wie *the dog*, wenn mehrere Hunde abgebildet sind).

Bei Grenzfällen in den Daten kann die Kategorie verwendet werden, die intuitiv am ehesten zutrifft. In seltenen Fällen kann es strittig sein, was im Satz die Objekt-Bezeichnung darstellt. Hier sollten generell Bezeichnungen von (a) konkreten Dingen genommen werden, die (b) keine Attribute oder Präzisierungen für andere Objekt-Bezeichnungen darstellen. Bei „bowl of apples“ wäre demnach „bowl“ ausschlaggebend (da „bowl“ ein konkretes Ding bezeichnet und „of apples“ eine Präzisierung); bei „bunch of apples“ dagegen „apples“ (da „bunch“ kein konkretes Ding bezeichnet).

Formalien

Bitte folgendes bei der Annotation beachten:

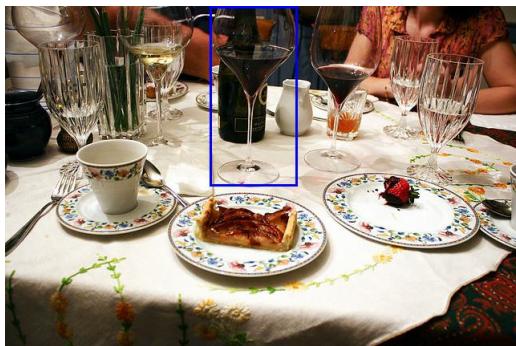
1. Die Kategorien für die Items bitte ausschließlich in der Spalte „annotation“ in der dazugehörigen CSV-Datei annotieren.
2. Die Kategorien sollen unter Verwendung der jeweiligen Buchstaben (also A, O, M oder F) annotiert und nicht ausgeschrieben werden. Bitte Großbuchstaben verwenden (also z.B. mit Capslock arbeiten).
3. **Ergänzt die Dateinamen der bearbeiteten CSV-Dateien bitte um euren Vornamen** (abgetrennt durch einen Unterstrich), damit die Resultate der einzelnen Annotator:innen später zusammengetragen werden können. Beispiel: Den Namen der von mir annotierten Datei „eine_datei.csv“ würde ich vor der Abgabe in „eine_datei_simeon.csv“ ändern. **Abgesehen davon bitte die Dateinamen unter keinen Umständen verändern, da sie codiert sind und ansonsten nicht mehr zugeordnet werden können!**
4. Die ausgefüllten und entsprechend umbenannten CSV-Dateien sollen bitte in folgenden File-Drop-Ordner abgelegt werden:
<https://uni-bielefeld.sciebo.de/s/0NzCs2OClb7HGVf>
Bitte darauf achten, dass der Vorname dem Dateinamen hinzugefügt wurde!

Sonstige Hinweise

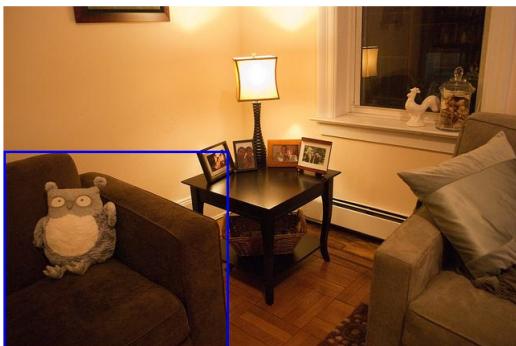
- Eigener Erfahrung nach arbeitet es sich gut im Split-Screen, wenn die HTML-Datei auf der einen und die CSV-Datei auf der anderen Hälfte des Bildschirms geöffnet ist.
- Bei Rückfragen: Bitte bei Simeon melden
 - Mail: simeon.junker@uni-bielefeld.de
 - Matrix: @sschuez:uni-bielefeld.de

Beispiele

Kategorie A



index: 9
ann_id: 667576
expression: middle wine glass

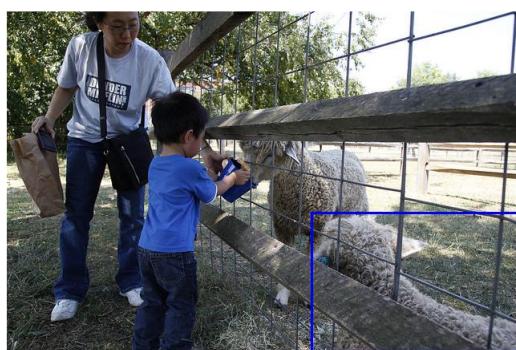


index: 69
ann_id: 117421
expression: left couch

Kategorie O



index: 15
ann_id: 322419
expression: bottom left

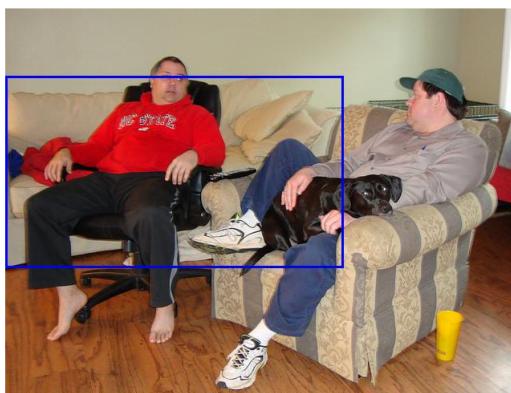


index: 24
ann_id: 65175
expression: right side of screen

Kategorie M

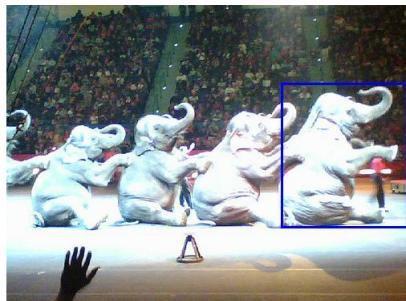


index: 5
ann_id: 97835
expression: chair in middle

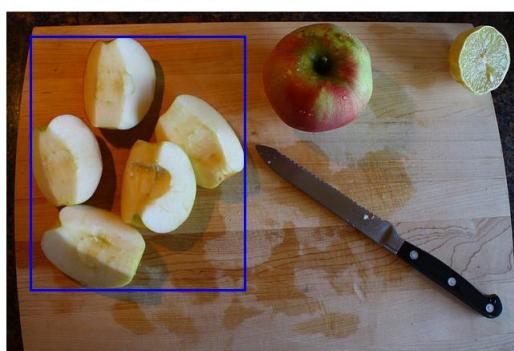


index: 82
ann_id: 111364
expression: red shirt

Kategorie F



index: 10
ann_id: 580317
expression: white dog



index: 30
ann_id: 1048809
expression: banana slices on left



index: 96
ann_id: 1118003
expression: white book