

Logo 1



Logo2

Goethe-Universität Frankfurt

Fachbereich 12 - Informatik und Mathematik

# Bachelorarbeit

zur Erlangung des akademischen Grades  
Bachelor of Science

**Thema:** Deep Neural Networks for occluded Image Recognition

**Autor:** Julius Taylor <s8423760@stud.uni-frankfurt.de>  
MatNr. 5210444

**Version vom:** 9. November 2017

**Betreuer:** Prof. Dr. Jochen Triesch

## **Sperrvermerk**

Die vorliegende Arbeit beinhaltet interne und vertrauliche Informationen der Firma <Firmenname>. Die Weitergabe des Inhalts der Arbeit im Gesamten oder in Teilen sowie das Anfertigen von Kopien oder Abschriften - auch in digitaler Form - sind grundsätzlich untersagt. Ausnahmen bedürfen der schriftlichen Genehmigung der Firma <Firmenname>.

## Zusammenfassung

Im Rahmen dieser Bachelorarbeit wurde ein Convolutional Neural Net zur Interpretation von Bildinformationen im anspruchsvollen Szenario der teilweisen Okklusion entwickelt und trainiert. Den Ausgangspunkt dieser Arbeit bietet die Arbeit von Spoerer und Kriegeskorte, welche das Konzept rekursiver Verbindungen in Neuronalen Netzen ausnutzen, um die Performance bei der Erkennung von Bildern mit teilweise fehlenden Informationen zu verbessern. Grundlage ist die Beobachtung, dass im ventralen Visuellen System laterale und rückwärtsgewandte Verbindungen und dadurch rekurrente Dynamiken zum Einsatz kommen, welche in der Bilderkennung bisher kaum oder gar nicht genutzt werden. Spoerer und Kriegeskorte vermuten, dass diese Rekurrenzen die Performance bei teilweise verdeckten Stimuli verbessern können. Um dies zu verifizieren, wurde ein simples Stimulus-Set generiert, welches Ziffern enthält, die wahlweise mit zufällig zerschnittenen Teilen von verschiedenen Zahlen verdeckt werden können. Das zugrundeliegende Problem der Klassifikation ist ein simples und wohl erforschtes. Daher kann sich in der Untersuchung allein auf die zusätzliche Schwierigkeit, die durch das Verdecken besagter Ziffern entsteht, konzentriert werden. Diese Arbeit reproduziert die Ergebnisse, die zeigen, dass rekurrente Dynamik die Klassifikationsrate bei sowohl unokkludierten als auch okkludierten Ziffern verbessert. Dabei werden die rekurrenten Modelle einerseits mit reinen feed forward Architekturen verglichen, die ungefähr in der Anzahl der Parameter mit den rekurrenten vergleichbar sind. Darüber hinaus, stelle ich sie alternativen Modellen gegenüber, die zusätzliche Konvolutions-Layer besitzen, um die Anzahl der Konvolutionen vergleichbar zu machen.

## Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>5</b>
<b>Tabellenverzeichnis</b>	<b>5</b>
<b>Listingverzeichnis</b>	<b>5</b>
<b>Abkürzungsverzeichnis</b>	<b>5</b>
<b>1 Einleitung</b>	<b>9</b>
<b>2 Materialien und Methoden</b>	<b>9</b>
2.1 Generatives Modell für Stimuli . . . . .	9
2.2 Modelle . . . . .	9
2.2.1 Implementierung . . . . .	9
<b>3 Ergebnisse</b>	<b>9</b>
<b>4 Diskussion</b>	<b>9</b>
<b>5 Danksagung</b>	<b>9</b>
<b>Literaturverzeichnis</b>	<b>12</b>
<b>Anhang</b>	<b>13</b>
<b>Eidesstattliche Erklärung</b>	<b>13</b>

## Abbildungsverzeichnis

1	Beispiel einer Bildbeschreibung . . . . .	10
2	Beschreibung . . . . .	10

## Tabellenverzeichnis

## Listingverzeichnis

1	Die Datei <code>data-config.xml</code> dient als Beispiel für XML Quellcode . . .	10
2	Das Listing zeigt Java Quellcode . . . . .	10

## Abkürzungsverzeichnis

CMS . . . . .	
---------------	--

CSS .....

ERM .....

GNU .....

GPL .....

GUI .....

HTML .....

IM .....

JS .....

JSON .....

KPI .....

LGPL .....

OCR .....

RSS .....

SQL .....

TDD .....

UGC .....

WWW .....

XMPP .....



# 1 Einleitung

## 2 Materialien und Methoden

### 2.1 Generatives Modell für Stimuli

Um die Effekte von Rekursion auf das klassifizieren von okkludierten Objekten zu untersuchen, wurde ein möglichst einfaches Grundproblem betrachtet. Die Klassifizierung von Ziffern ist ein gut untersuchtes Machine Learning Problem, in dem übermenschliche Performance erreicht werden kann. Allgemein gilt das sehr ähnliche MNIST-Datenset gemeinhin als das "Hello world!" des Machine Learning. Durch die Einfachheit der Aufgabe, können die Auswirkungen von Rekursion isoliert von anderen Herausforderungen betrachtet werden.

### 2.2 Modelle

Um die Wirkungskraft von Rekursion zu untersuchen, wurden verschiedene CNNs mit variierenden Graden an Rekursion benutzt und systematisch miteinander verglichen. Die Nomenklatur richtet sich am Paper von Spoerer und Kriegeskorte aus. Als Grundlage dient eine Standard feed-forward Architektur (B) mit reinen 'bottom-up' Verbindungen. Da diese jedoch in der Anzahl der Parameter und der Anzahl der durchgeführten Konvolutionen gegenüber seinen Rekursiven Varianten unterlegen ist, wurden zum Vergleich zusätzlich Modelle entwickelt, die in den entsprechenden Domänen angepasst wurden. Einerseits wurde die Größe der Konvolutionskernel angepasst und somit die Größe der erlernbaren Features. Andererseits wurde die Anzahl der Konvolutionen erhöht, die Gewichte in den zusätzlichen Konvolutionen jedoch mit den anderen Konvolutionen geteilt. Somit kann die Anzahl der Faltungsoperationen vergleichbar gemacht werden, ohne die freien Parameter zu erhöhen. Die Architekturen werden im Folgenden BK respektive BKC genannt.

#### 2.2.1 Implementierung

## 3 Ergebnisse

## 4 Diskussion

## 5 Danksagung

Abbildung 2 [S.10]

---

<sup>1</sup>Bildquelle: Beispielquelle

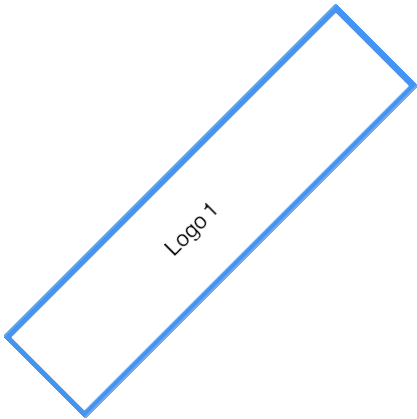


Abbildung 1: Beispiel einer Bildbeschreibung<sup>1</sup>



Abbildung 2: Beschreibung

Überschrift 1	Überschrift 2
Info 1	Info 2
Info 3	Info 4

```
1 <dataConfig>
2   <dataSource type="JdbcDataSource"
3       driver="com.mysql.jdbc.Driver"
4       url="jdbc:mysql://localhost/bms_db"
5       user="root"
6       password="" />
7   <document>
8     <entity name="id"
9       query="select id, htmlBody, sentDate, sentFrom, subject, textBody
10      from mail">
11     <field column="id" name="id" />
12     <field column="htmlBody" name="text" />
13     <field column="sentDate" name="sentDate" />
14     <field column="sentFrom" name="sentFrom" />
15     <field column="subject" name="subject" />
16     <field column="textBody" name="text" />
17   </entity>
18 </document>
19 </dataConfig>
```

Listing 1: Die Datei data-config.xml dient als Beispiel für XML Quellcode

```
1 /* generate TagCloud */
2 Cloud cloud = new Cloud();
3 cloud.setMaxWeight(_maxSizeOfText);
```

```

4 cloud.setMinWeight(_minSizeOfText);
5 cloud.setTagCase(Case.LOWER);
6
7 /* evaluate context and find additional stopwords */
8 String query = getContextQuery(_context);
9 List<String> contextStoplist = new ArrayList<String>();
10 contextStoplist = getStopwordsFromDB(query);
11
12 /* append context stoplist */
13 while(contextStoplist != null && !contextStoplist.isEmpty())
14     _stoplist.add(contextStoplist.remove(0));
15
16 /* add cloud filters */
17 if (_stoplist != null) {
18     DictionaryFilter df = new DictionaryFilter(_stoplist);
19     cloud.addInputFilter(df);
20 }
21 /* remove empty tags */
22 NonNullFilter<Tag> nnf = new NonNullFilter<Tag>();
23 cloud.addInputFilter(nnf);
24
25 /* set minimum tag length */
26 MinLengthFilter mlf = new MinLengthFilter(_minTagLength);
27 cloud.addInputFilter(mlf);
28
29 /* add taglist to tagcloud */
30 cloud.addText(_taglist);
31
32 /* set number of shown tags */
33 cloud.setMaxTagsToDisplay(_tagsToDisplay);

```

Listing 2: Das Listing zeigt Java Quellcode

Die Zuordnung aller möglichen Werte, welche eine Zufallsvariable annehmen kann nennt man *Verteilungsfunktion* von  $X$ .

Die Funktion  $F: \mathbb{R} \rightarrow [0,1]$  mit  $F(t) = P(X \leq t)$  heißt Verteilungsfunktion von  $X$ .<sup>2</sup>

Für eine stetige Zufallsvariable  $X: \Omega \rightarrow \mathbb{R}$  heißt eine integrierbare, nicht-negative reelle Funktion  $w: \mathbb{R} \rightarrow \mathbb{R}$  mit  $F(x) = P(X \leq x) = \int_{-\infty}^x w(t)dt$  die *Dichte* oder *Wahrscheinlichkeitsdichte* der Zufallsvariablen  $X$ .<sup>3</sup>

---

<sup>2</sup>Konen, vgl. [?] [S.55]

<sup>3</sup>Konen, vgl. [?] [S.56]

## **Literaturverzeichnis**

## **Anhang**

## Eidesstattliche Erklärung

### Eidesstattliche Erklärung zur <-Arbeit>

Ich versichere, die von mir vorgelegte Arbeit selbstständig verfasst zu haben. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer entnommen sind, habe ich als entnommen kenntlich gemacht. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit benutzt habe, sind angegeben. Die Arbeit hat mit gleichem Inhalt bzw. in wesentlichen Teilen noch keiner anderen Prüfungsbehörde vorgelegen.

*Unterschrift :*

*Ort, Datum :*

