technische universität dortmund

fakultät für informatik

# Wordspotting in historischen Dokumenten

Gernot A. Fink

TU Dortmund, Fakultät für Informatik

Herne, 18. Februar 2017

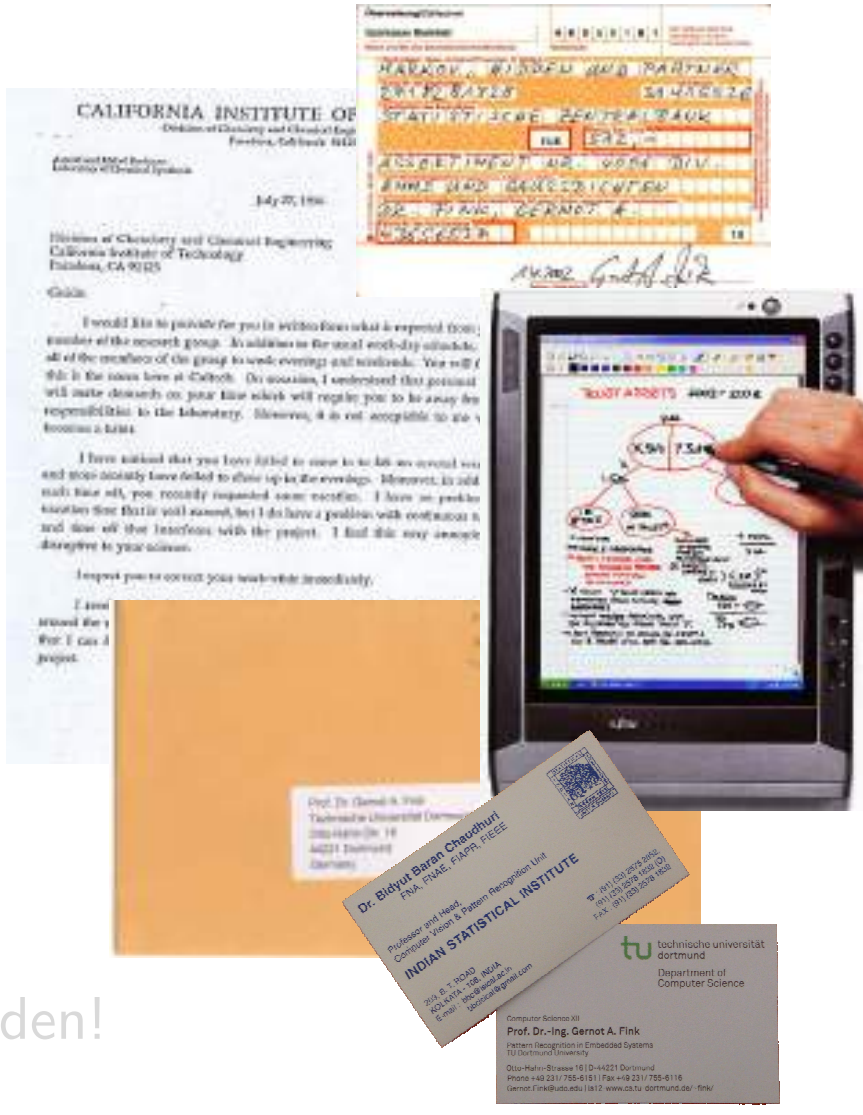
- Einführung — *Warum Wordspotting?*
- Wordspotting — *Grundlagen und Methoden*
- Tiefe neuronale Netze — *... der Stand-der-Kunst*
- Zusammenfassung — *... und ein Anwendungsbeispiel*

Mit Beiträgen von: *Leonard Rothacker, Sebastian Sudholt*

# Einführung: Systeme zum Maschinellen Lesen

## Stand des Maschinellen Lesens:

- Eines der ersten Anwendungsgebiete der Informatik
- Ex. OCR-Systeme für maschinell gedruckte Texte (inkl. Fraktur)
- Ex. *online* Handschrifterkennung (auf SmartPhones, Tablets, ...)
- *Offline* Handschrifterkenung: *offenes Forschungsproblem*



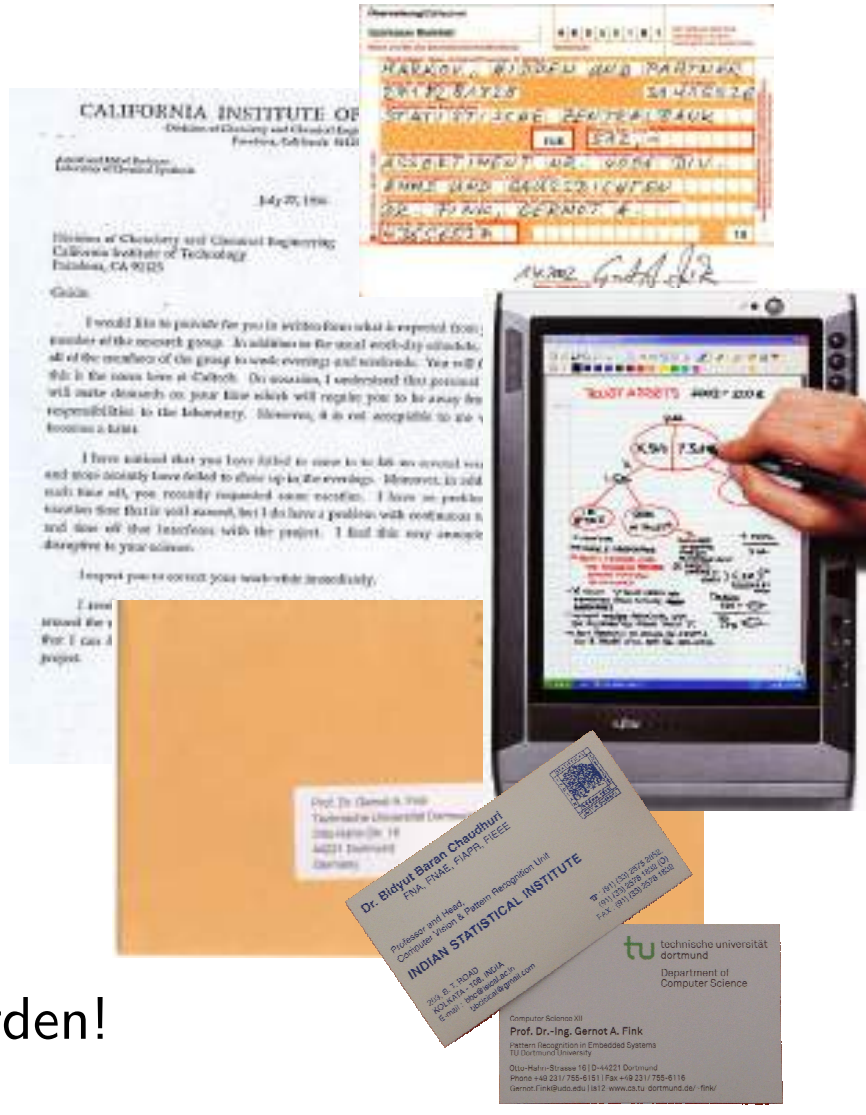

## Generelle Methodik/Problematik:

Transkriptionsmodelle müssen auf *beträchtlichen* Mengen von *annotierten* Beispieldaten trainiert werden!

# Einführung: Systeme zum Maschinellen Lesen

## Stand des Maschinellen Lesens:

- Eines der ersten Anwendungsgebiete der Informatik
- Ex. OCR-Systeme für maschinell gedruckte Texte (inkl. Fraktur)
- Ex. *online* Handschrifterkennung (auf SmartPhones, Tablets, ...)
- *Offline* Handschrifterkenung: *offenes Forschungsproblem*




## Generelle Methodik/Problematik:

Transkriptionsmodelle müssen auf *beträchtlichen* Mengen von *annotierten* Beispieldaten trainiert werden!

# Einführung: Warum Wordspotting?

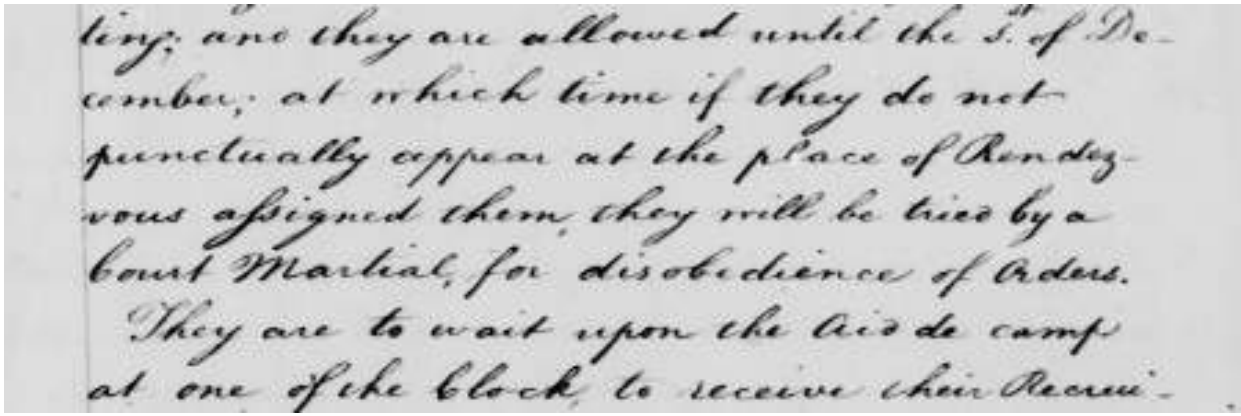
Was wenn die automatische Transkription von Handschrift nicht länger möglich ist?


Bild einer Feldpostkarte aus der *Privatsammlung Dr. Britta Bley*, Dortmund

# Einführung: Warum Wordspotting?

Was wenn die automatische Transkription von Handschrift nicht länger möglich ist?

... lassen Sie es uns etwas einfacher machen!




Alternative: Suche ("Retrieval") einzelner Wörter im Gegensatz zu Transkription ("query-by-example")

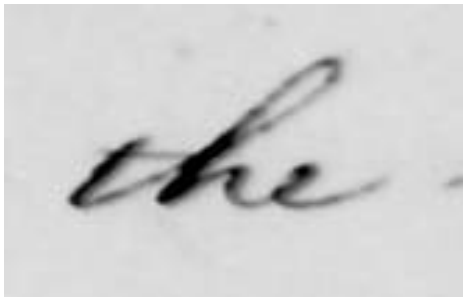
Bilder aus *The George Washington Papers at the Library of Congress, 1741-1799*

# Einführung: Warum Wordspotting?

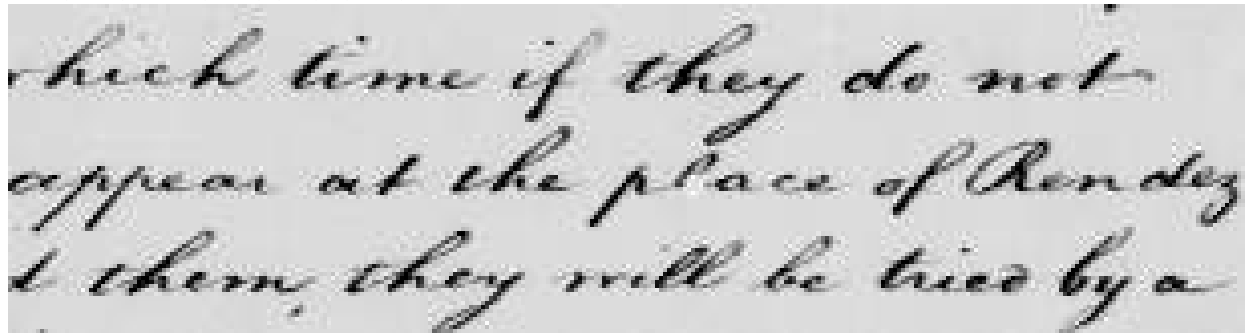
Was wenn die automatische Transkription von Handschrift nicht länger möglich ist?

Alternative: Suche ("Retrieval") einzelner Wörter im Gegensatz zu Transkription ("query-by-example")

Anfrage-Wortabbild



Dokumentabbild



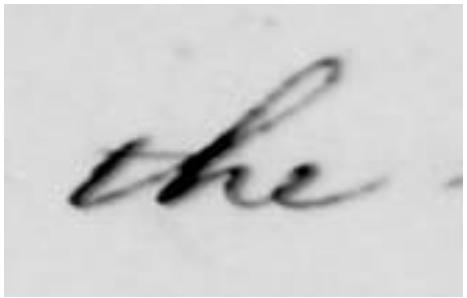
Bilder aus *The George Washington Papers at the Library of Congress, 1741-1799*

# Einführung: Warum Wordspotting?

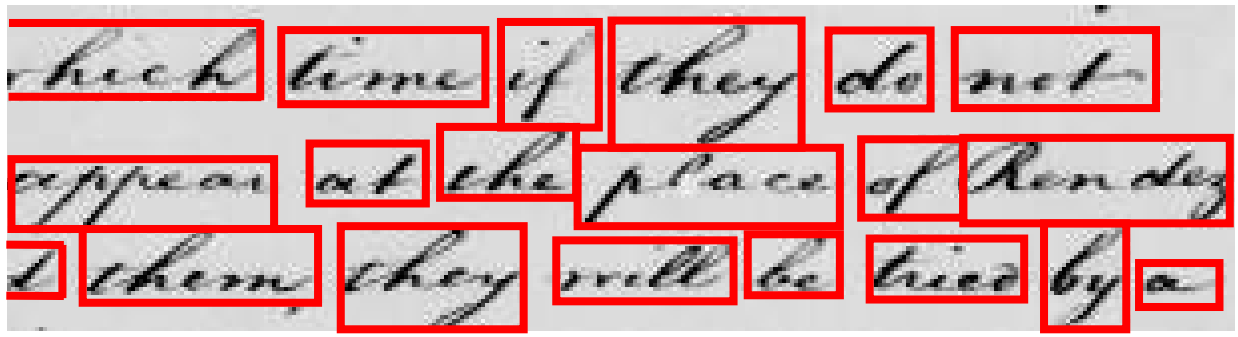
Was wenn die automatische Transkription von Handschrift nicht länger möglich ist?

Alternative: Suche ("Retrieval") einzelner Wörter im Gegensatz zu Transkription ("query-by-example")

Anfrage-Wortabbild



Dokumentabbild

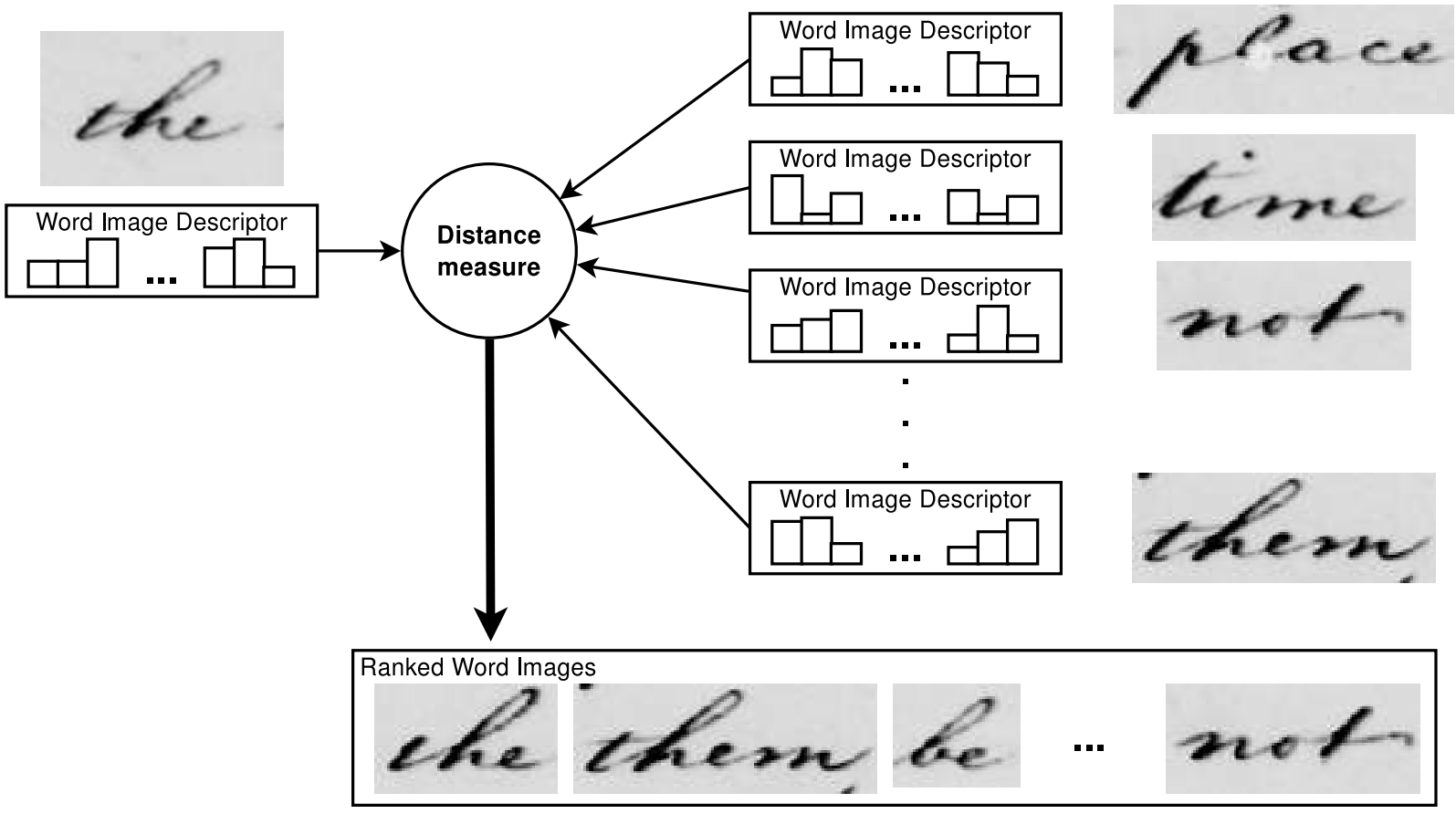


Bilder aus *The George Washington Papers at the Library of Congress, 1741-1799*

# Einführung: Warum Wordspotting?

*Query-by-example* Wordspotting

*Woher kennen wir das?*



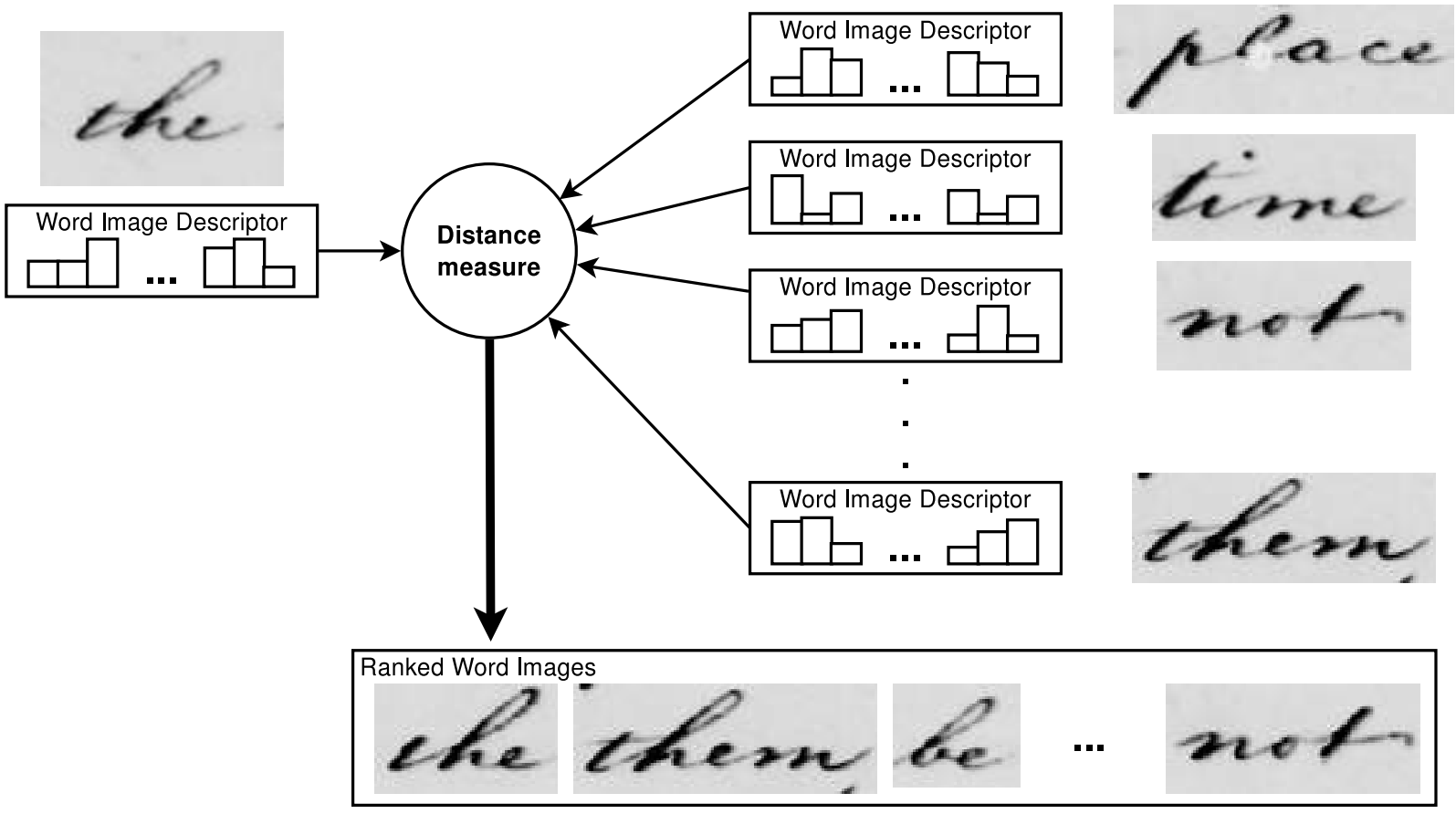

Bilder aus *The George Washington Papers at the Library of Congress, 1741-1799*

Nach [Rath & Manmatha, IJRAR'07]

# Einführung: Warum Wordspotting?

*Query-by-example* Wordspotting

*Woher kennen wir das?*




Bilder aus *The George Washington Papers at the Library of Congress, 1741-1799*

Nach [Rath & Manmatha, IJRAR'07]

# Einführung: Warum Wordspotting?

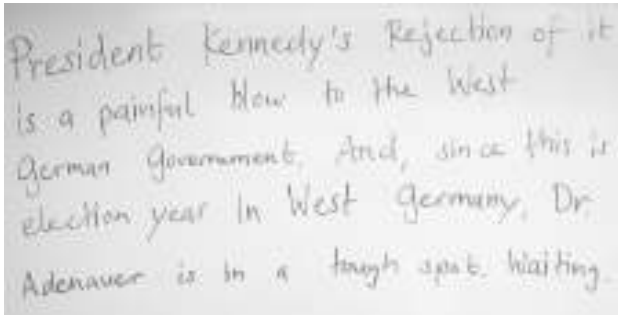
*QbE* Wordspotting ≈ Spezialfall der **Inhaltsbasierten Bildsuche**

(Quelle: The Chinese University of Hong Kong)

# Einführung: Transkription vs. Suche

# Einführung: Transkription vs. Suche

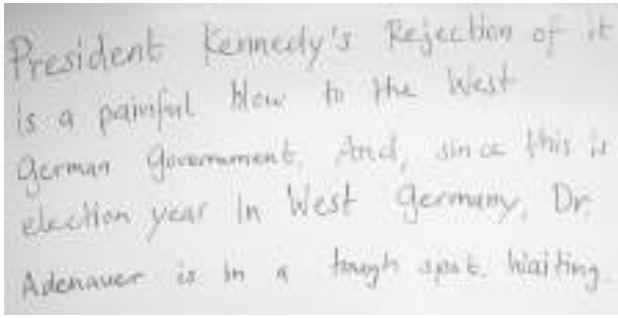
Transkription von Dokumenten (= "klassische" Erkennung)




*President Kennedy is Rejection of it is a painful blow to the West German Government And since this is selection year in West Germany of Adenauer is in a tough spot waiting*

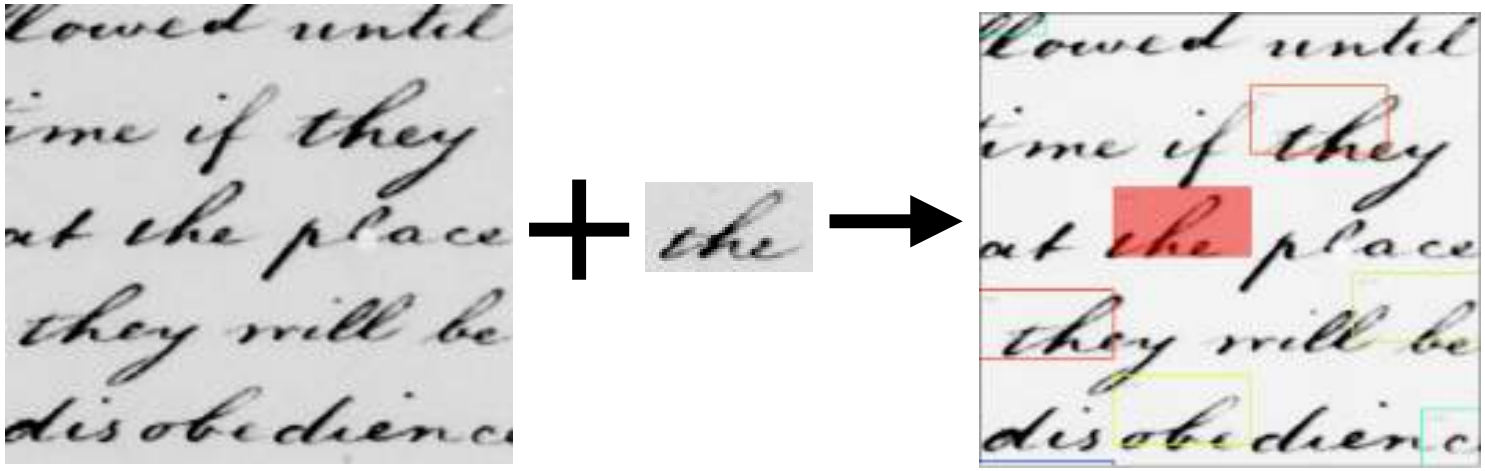
# Einführung: Transkription vs. Suche

## Transkription von Dokumenten (= "klassische" Erkennung)



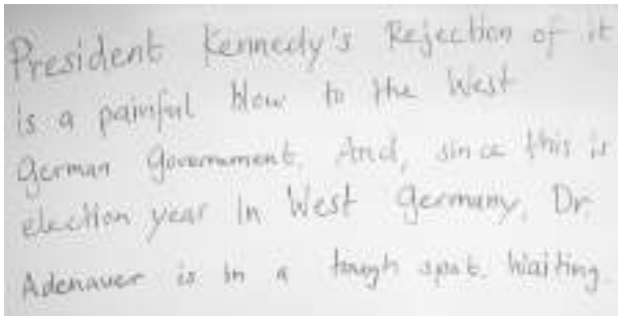

*President Kennedy is Rejection of it is a painful blow to the West German Government And since this is selection year in West Germany of Adenauer is in a tough spot waiting*

## Suche in Dokumenten (aka "Wordspotting")



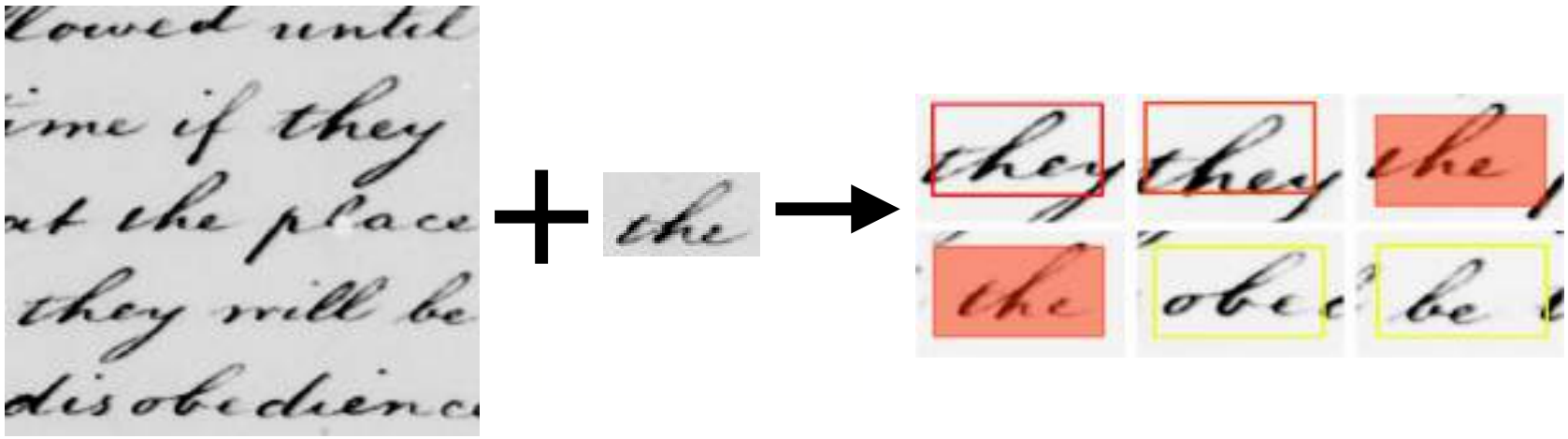
# Einführung: Transkription vs. Suche

## Transkription von Dokumenten (= "klassische" Erkennung)



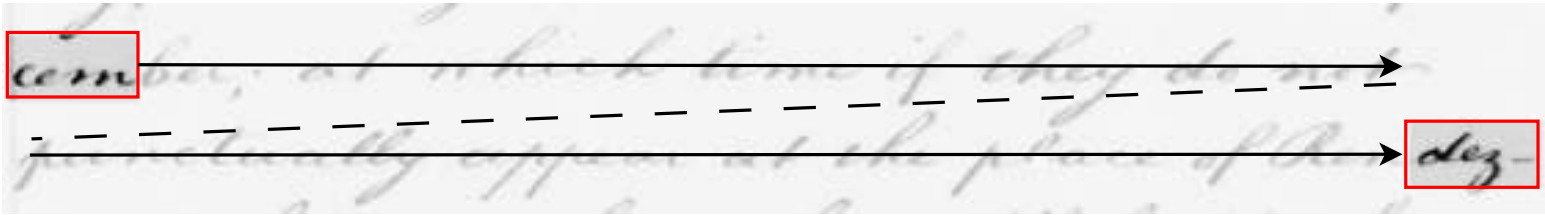

*President Kennedy is Rejection of it is a painful blow to the West German Government And since this is selection year in West Germany of Adenauer is in a tough spot waiting*

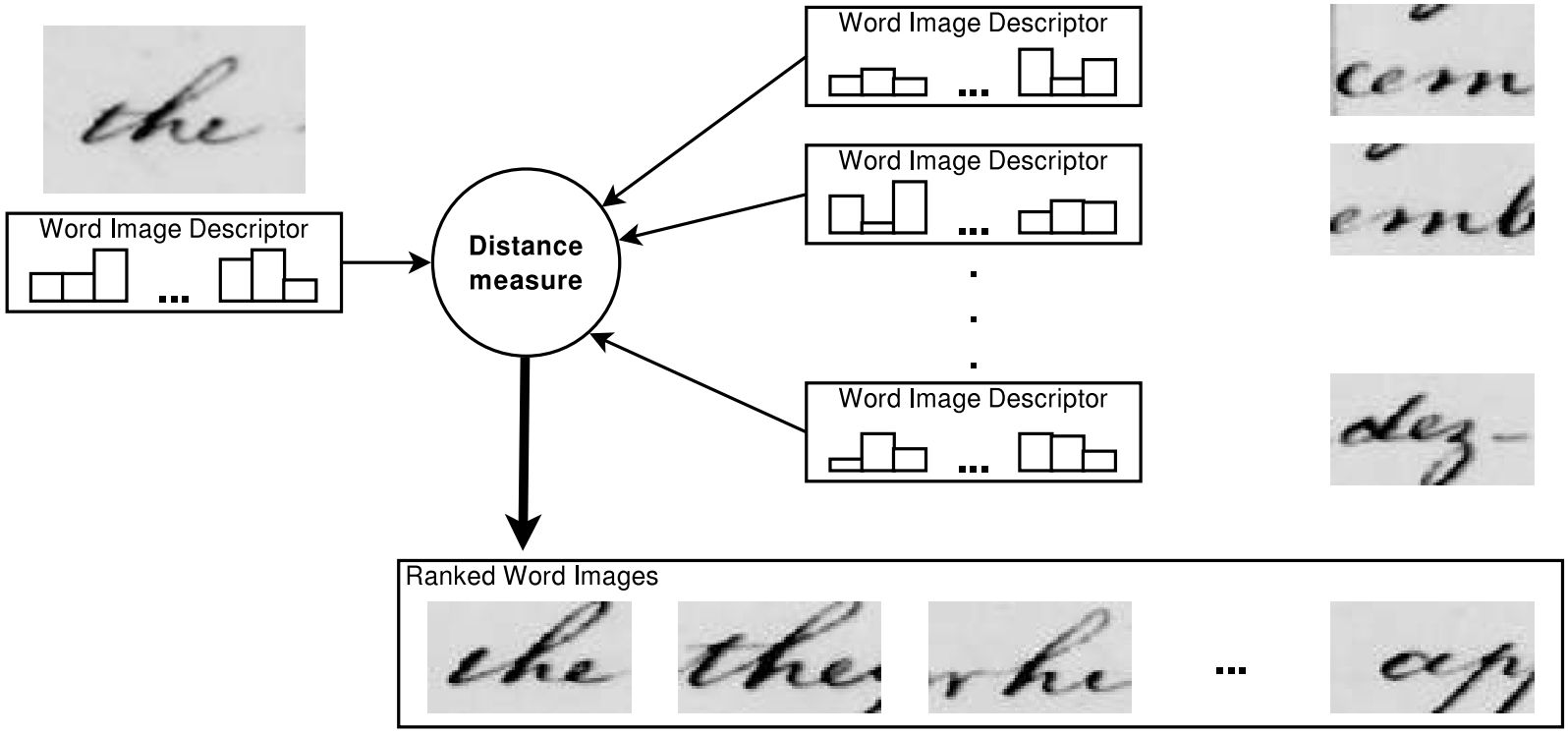
## Suche in Dokumenten (aka "Wordspotting")



# Wordspotting: Segmentierung?

*Segmentierungsfreies Wordspotting* mit gleitendem Suchfenster






(Rusiñol *et al.*: ICDAR, 2011)

# Wordspotting: Schriftsysteme?

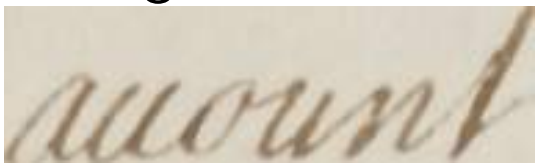
*Prinzipielle Methodik praktisch universell einsetzbar!*

Anfragebild Dokumentabbild

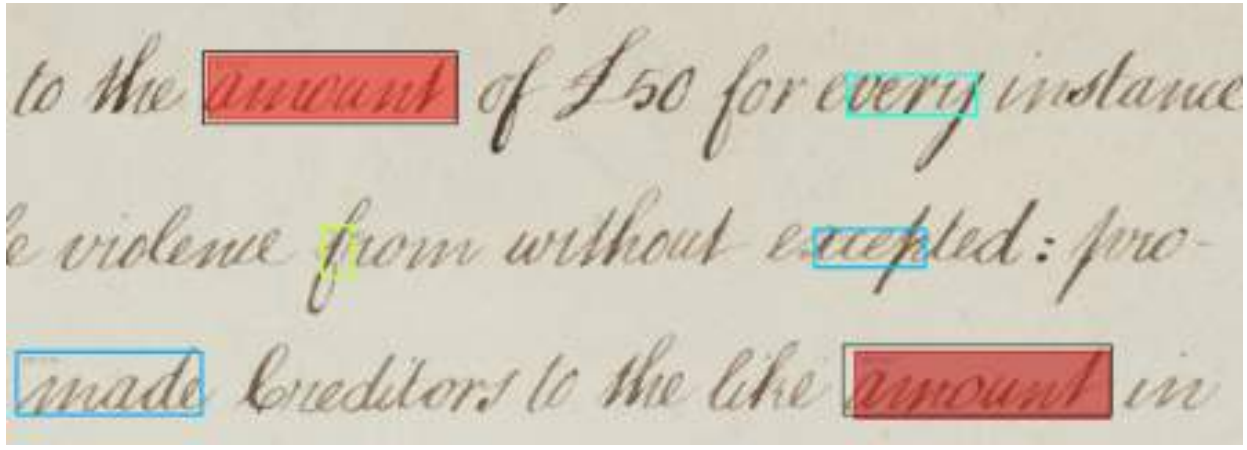
# Wordspotting: Schriftsysteme?

*Prinzipielle Methodik praktisch universell einsetzbar!*

Anfragebild



Dokumentabbild

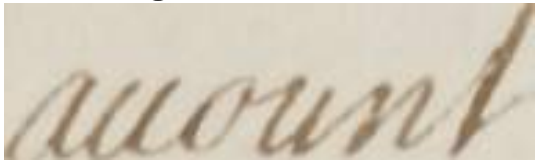



(Quelle: UCL Special Collections, Bentham Papers)

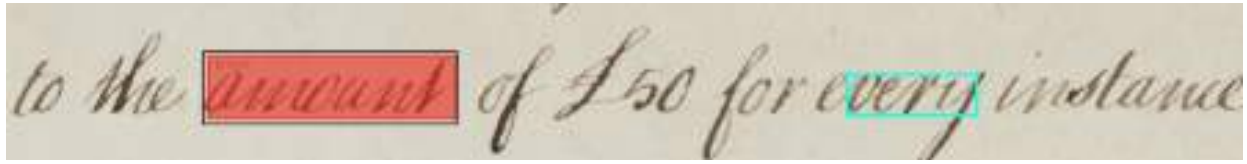
# Wordspotting: Schriftsysteme?

*Prinzipielle Methodik praktisch universell einsetzbar!*

Anfragebild



Dokumentabbild




(Quelle: Feldpostkarte, *Privatsammlung Dr. Britta Bley*, Dortmund)

# Wordspotting: Schriftsysteme?

*Prinzipielle Methodik praktisch universell einsetzbar!*

Anfragebild

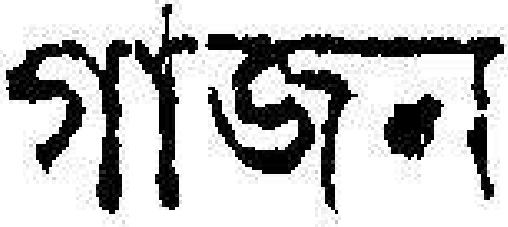
Dokumentabbild

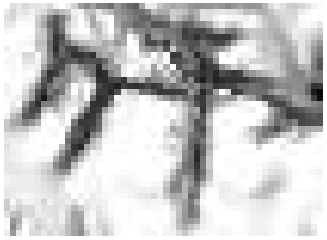
(Quelle: H. Palit. Aadyer Gambhira. Krishnacharan Sarkar, Maldaha, 1913)

# Wordspotting: Schriftsysteme?

*Prinzipielle Methodik praktisch universell einsetzbar!*

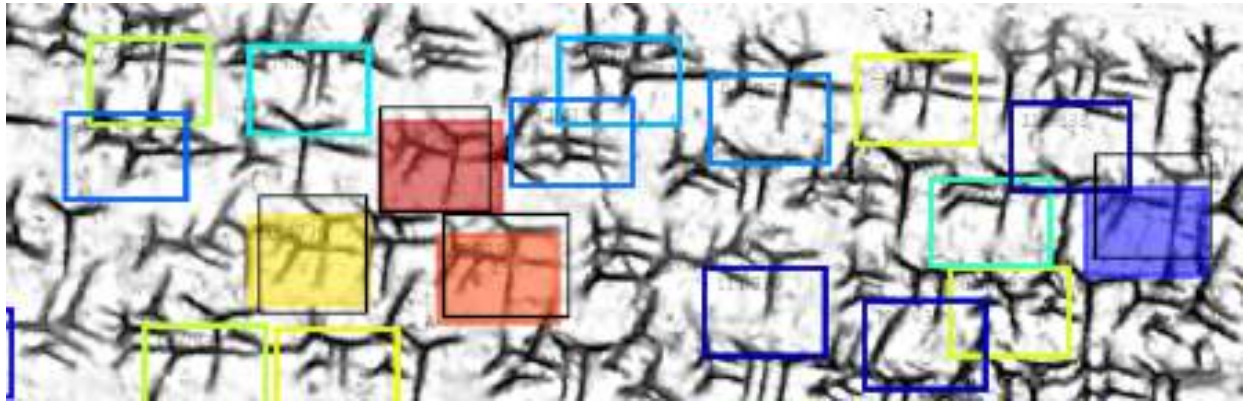
Anfragebild





Dokumentabbild

(Quelle: H. Palit. Aadyer Gambhira. Krishnacharan Sarkar, Maldaha, 1913)



(Keilschrifttafel, Hattusa, 30.000 v. Chr., Quelle: Hethitologie Portal, Mainz)

# Wordspotting: Zwischenfazit

Methode betrachtet bisher: *Query-by-Example Wordspotting*

Wesentlicher Vor- / Nachteil:

- ✓ Erfordert kein Trainingsmaterial / keine Annotationen!
- ↯ Anfragen können nur *ausgewählt* werden.

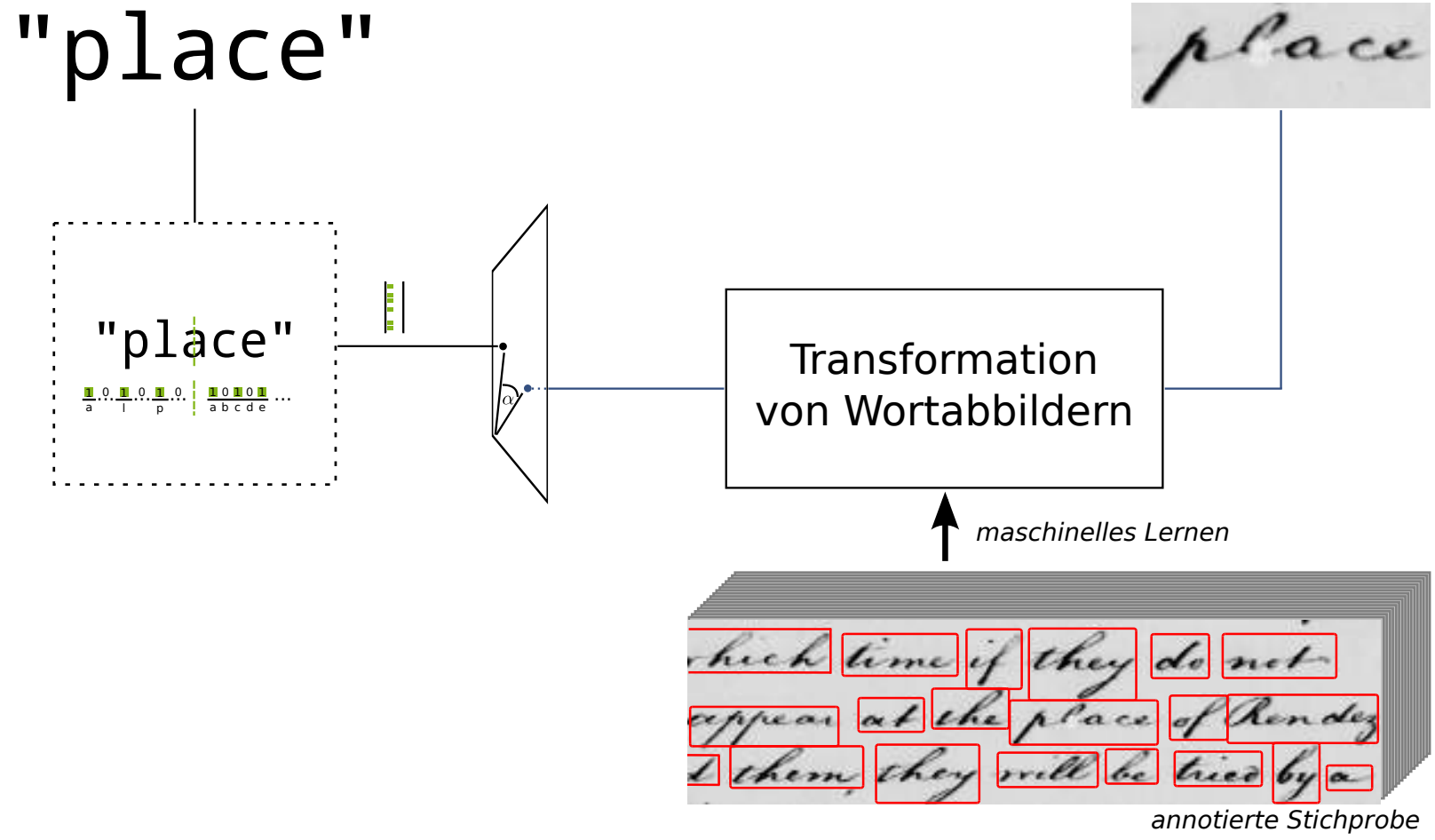
Was ist mit beliebigen Anfragen?

Anforderung: Erzeugung eines Schriftmodells aus einer *textuellen* Eingabe (= "string")

⇒ *Query-by-String Wordspotting*

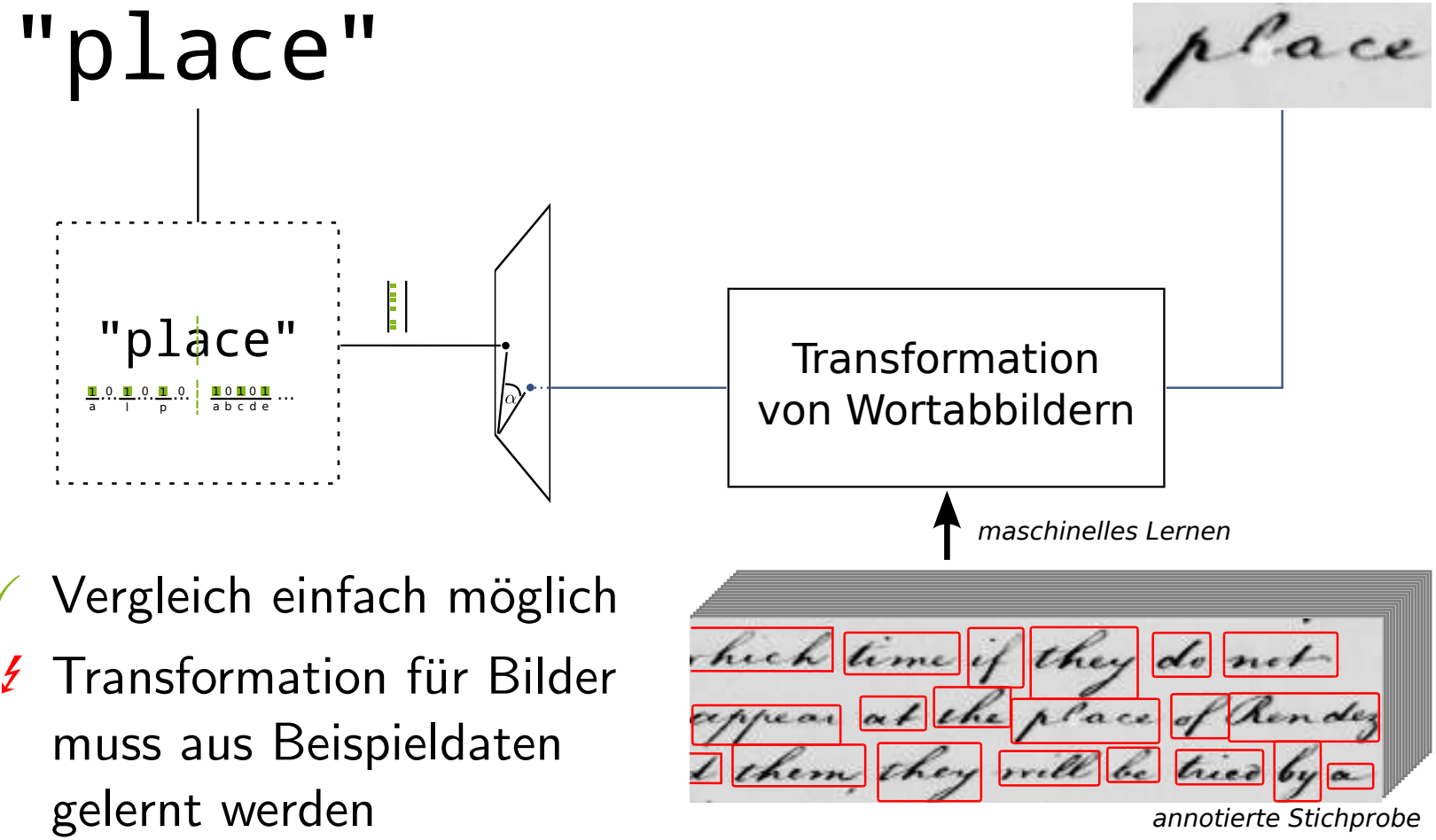
# Wordspotting: Query-by-String

Idee: Gemeinsame Repräsentation für Schrift und Bild



# Wordspotting: Query-by-String

Idee: Gemeinsame Repräsentation für Schrift und Bild



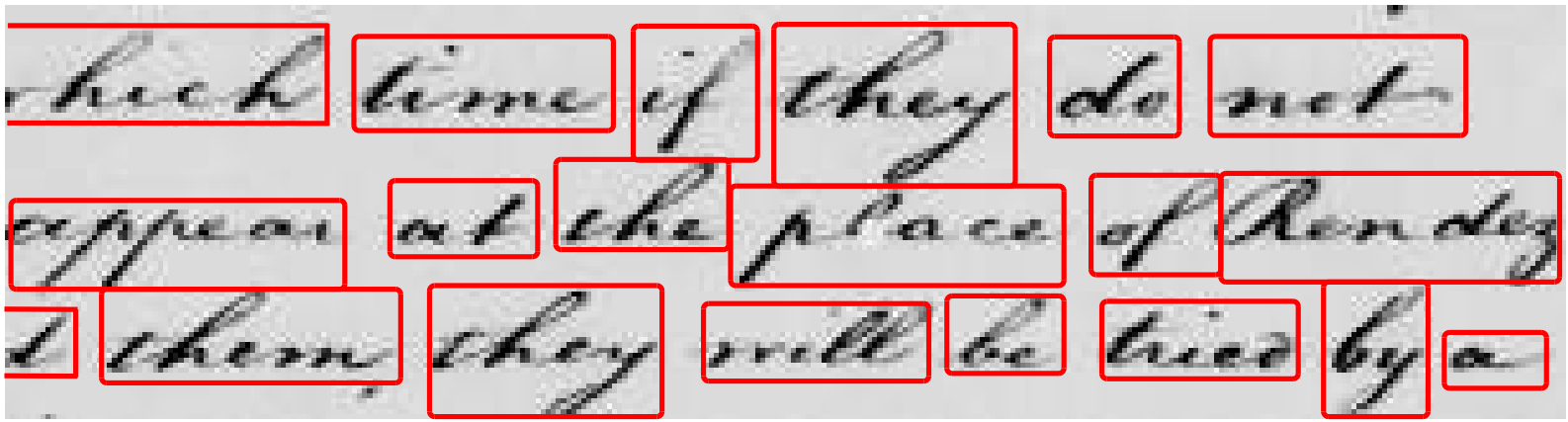

- ✓ Vergleich einfach möglich
- ↯ Transformation für Bilder muss aus Beispieldaten gelernt werden

⇒ *Annotierte* Stichprobe erforderlich!

# Wordspotting: Query-by-String

Idee: Gemeinsame Repräsentation für Schrift und Bild

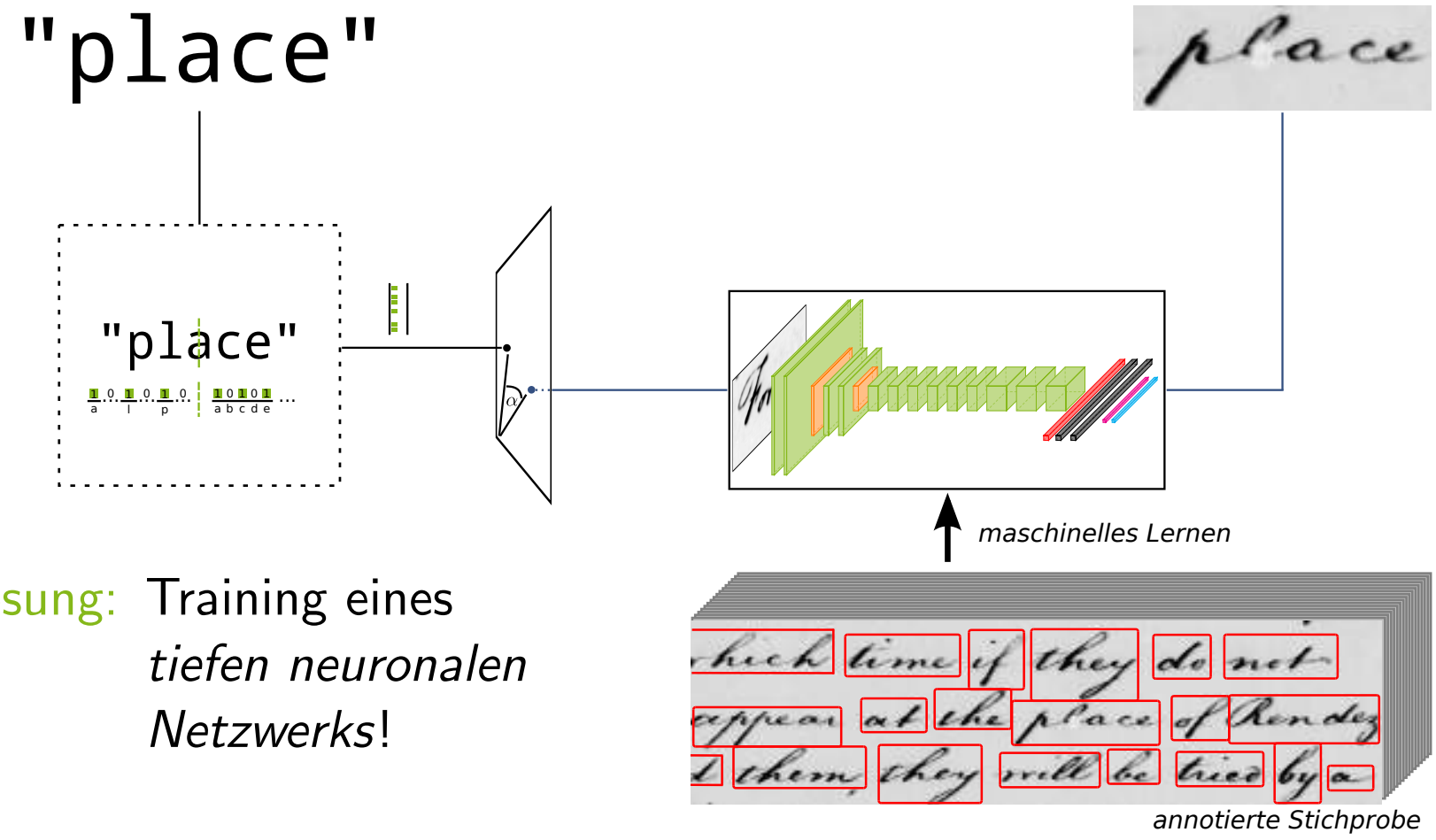
*Beispielhafte* Annotation auf Wortebene



*which time if they do not appear at the place of Rendez them they will be tied by a*

# Wordspotting: Query-by-String

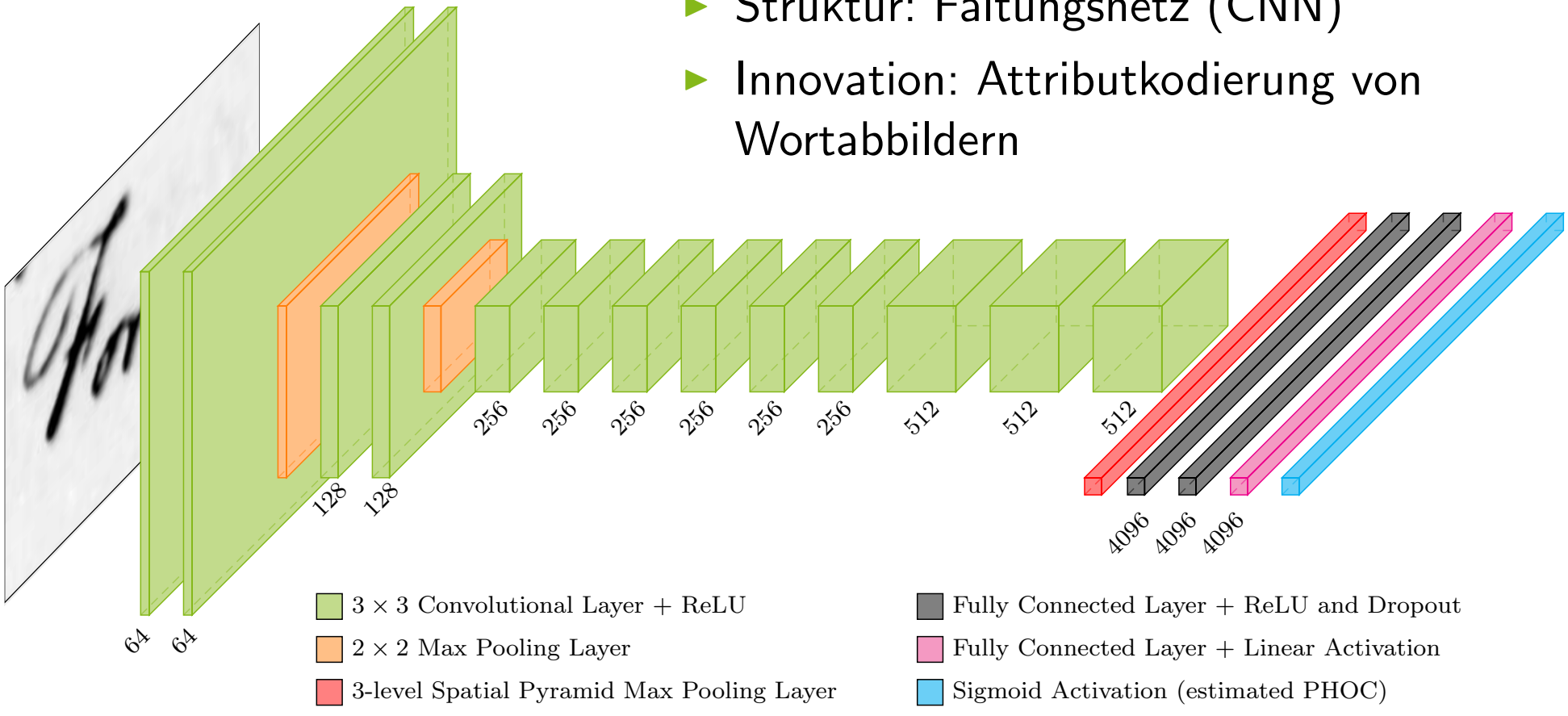
Idee: Gemeinsame Repräsentation für Schrift und Bild




Lösung: Training eines *tiefen neuronalen Netzwerks*!

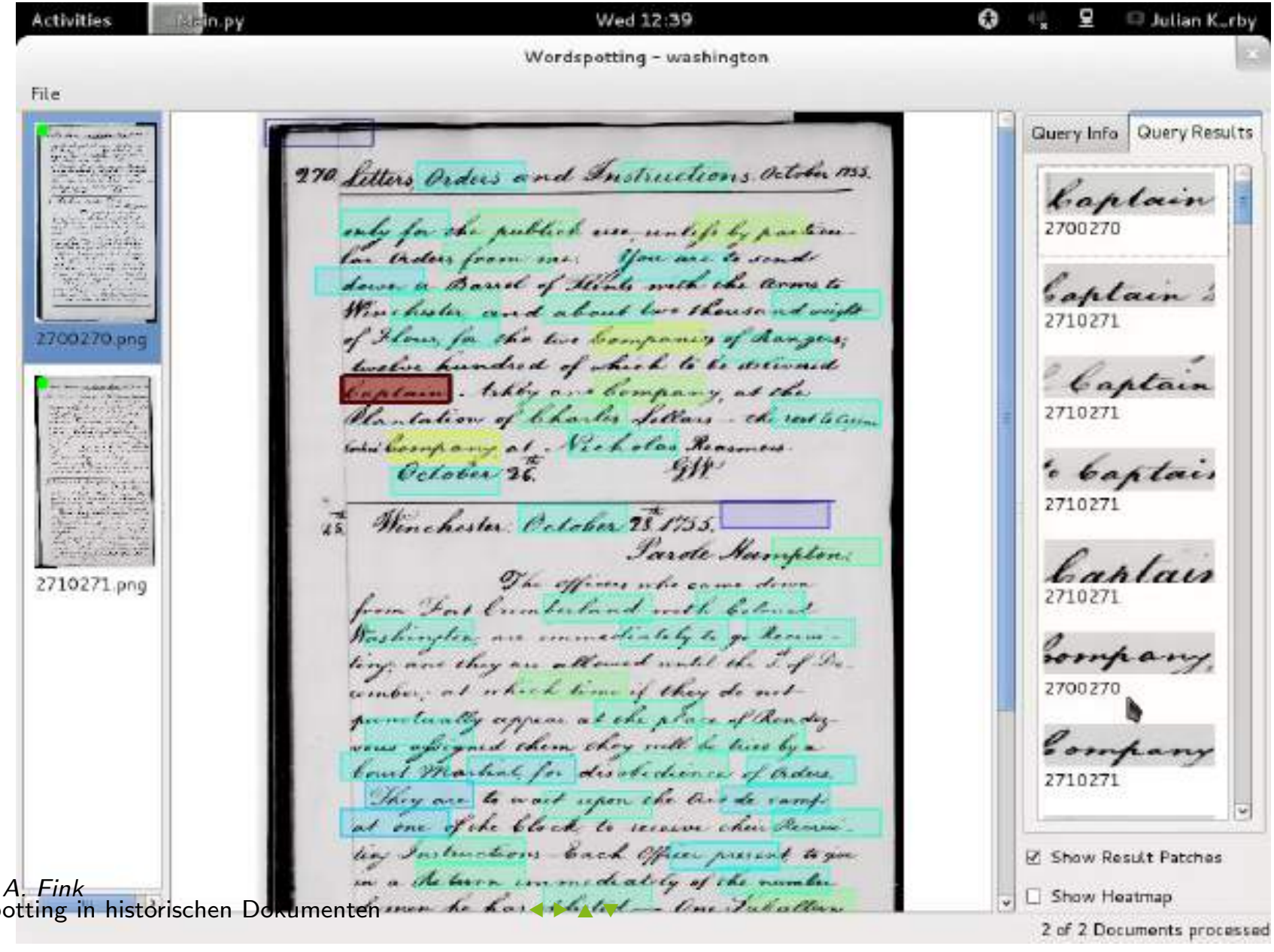
# Wordspotting: Tiefe Neuronale Netze

- Struktur: Faltungsnetz (CNN)
- Innovation: Attributkodierung von Wortabbildern




(Sudholt & Fink: ICFHR, 2016, Gewinner des *Best Paper Award*!)

# Anwendungsbeispiel Wordspotting



# Zusammenfassung

Grundmethode: *Query-by-Example Wordspotting*

- ✓ Gute Leistungsfähigkeit auf isogenen Dokumenten
- ✓ Kein Annotationsaufwand erforderlich!
- ↯ Schlechte Generalisierung auf verschiedene Schriftstile

Alternative: *Query-by-String Wordspotting*

- ✓ Jede (textuelle) Anfrage kann verwendet werden
- STOP Erfordert annotiertes Trainingsmaterial!
- ✓ Stand-der-Technik: *Tiefe Neuronale Netze*

# Zusammenfassung

Grundmethode: *Query-by-Example Wordspotting*

- ✓ Gute Leistungsfähigkeit auf isogenen Dokumenten
- ✓ Kein Annotationsaufwand erforderlich!
- ↯ Schlechte Generalisierung auf verschiedene Schriftstile

Alternative: *Query-by-String Wordspotting*

- ✓ Jede (textuelle) Anfrage kann verwendet werden
- STOP Erfordert annotiertes Trainingsmaterial!
- ✓ Stand-der-Technik: *Tiefe Neuronale Netze*