
Archive content annotation and search with Taylor

Research project TailoredMedia

Christoph Bauer, Werner Bailer, Peter Judmaier



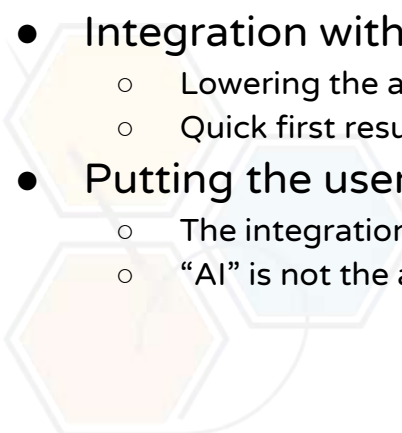
TailoredMedia: Tailored and Agile enrichment and Linking FOR sEmantic Description of multiMedia

- Austrian research project (ICT for the Future programme)
- 24 months, started Nov. 2020
- two media creators/archives
 - ORF, Austrian Mediatheque
- three research/technology partners
 - JOANNEUM RESEARCH
 - Redlink
 - Sankt Pölten University of Applied Sciences
 - including their student TV channel



Motivation



- Leveraging use of maturing AI tools
 - Overcoming some of the main obstacles
 - Increase the trust-level
 - Obtaining relevant data that fits both documentation and search needs
 - Preventing a data-avalanche of irrelevant information gained by AI mining processes
 - Teaching AI about relevance...
 - Integration with existing workflows
 - Lowering the acceptance-barrier
 - Quick first results and positive impact by aiming at major workloads
 - Putting the user back into the focus of development
 - The integration of AI tools is not a showcase for IT-development
 - “AI” is not the adversary but a relevant support and tool
- 

Challenges

- **Taming the Beast (data-avalanche)**
 - with as little human intervention as possible
 - finding new ways of data-presentation
 - improving the dialogue between user and machine/tool
- **Improving Data-Quality deriving from AI-based Data-mining-processes**
 - lowering the amount of “false positives”
 - putting in use pre-existing data
 - clever combining of mining-processes
- **Lowering the (pre-)investment of manpower**
 - introducing integrated, simplified and reduced AI-Training efforts
- **Integrating the user in development and decision-making**
 - open POCs, early test-integrations, collaboration-efforts beyond the own company

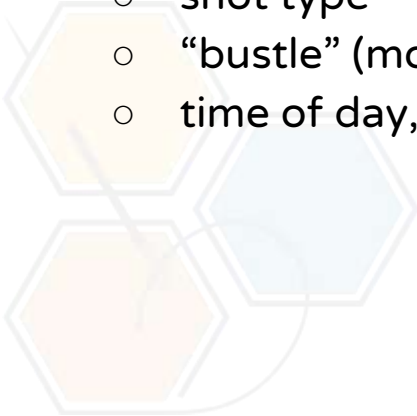
Use cases, examples

- **Annotation & Documentation**
 - Current Situation: Documentalist acts as Data Hub and Filter for incoming Information
 - Future Workflow: As little human intervention as possible
- **Searching**
 - Current Situation: Highly “prepared” data structure (see above)
 - Future Workflow: More unstructured data? → Data-avalanche !?!?
- **Use-Cases for the prototype**
 - Search for specific content
 - Support of Query-Construction (How to tell my machine)
 - Make results comprehensible (You have got this because...)
 - Annotation-task (Control your AI)
 - Automatic guidance (Bring important things to the front)
 - Customizable (There are still individuals in front of the screen...)
 - Support of different target groups
 - Introducing TAILOR (A new star is born...)

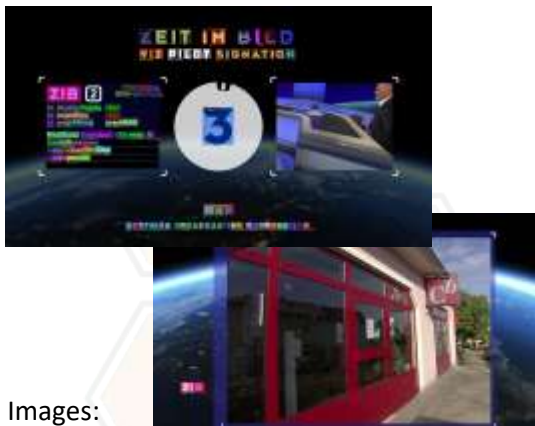
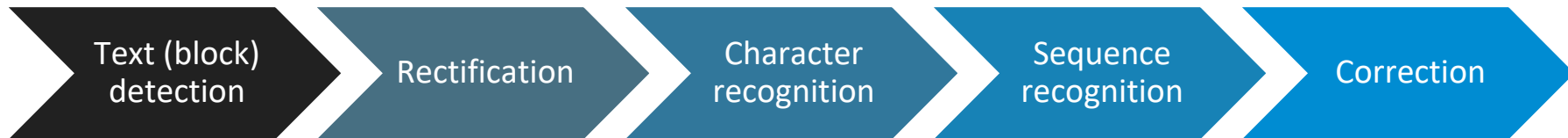


Selected technical topics



- Scene text detection & recognition (“text in the wild”)
 - Few-shot learning
 - Efficient classification of shots
 - place type (coarse and fine)
 - shot type
 - “bustle” (more or less populated)
 - time of day, season (under development)
- 

Scene text detection and classification



Images:
ORF

Headlines

BILD

cb

ORF

Headllnes

B I L D

C b

O R F

Headlines

BILD

Cb

ORF

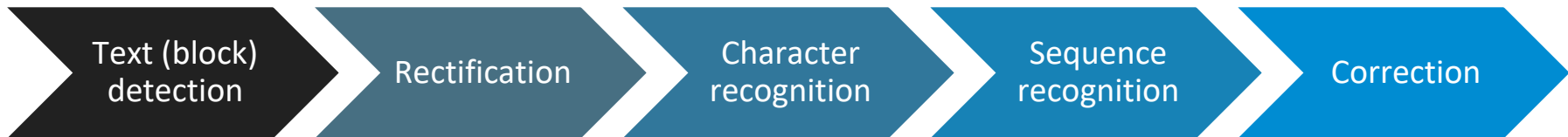
Headlines

BILD

CH

ORF

Scene text detection and classification



Script



Image:
<https://vcar.github.io/ASAYAR/>

**Script +
language
specific
characters**

**probabilities
of character
co-
occurrence**

(variabel)

DE: -bel
EN: -ble

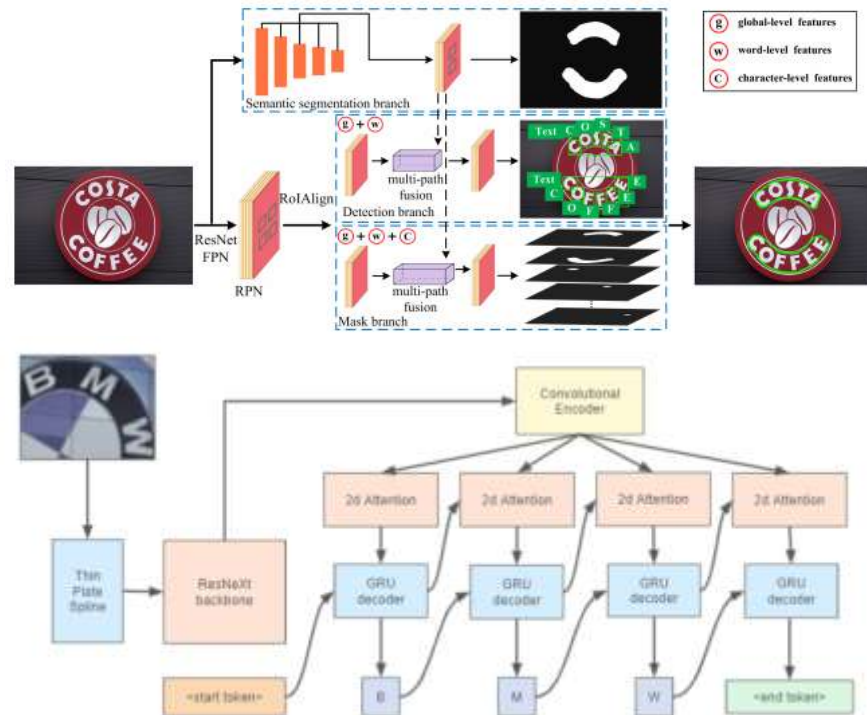
Vocabulary



Commerzbank

Scene text detection and classification

- Detection
 - approach based on TextFuseNet [1]
- Recognition
 - attention-based recognition approach [2]
 - issues with German text
 - retraining on synthetic German samples (using SynthText)



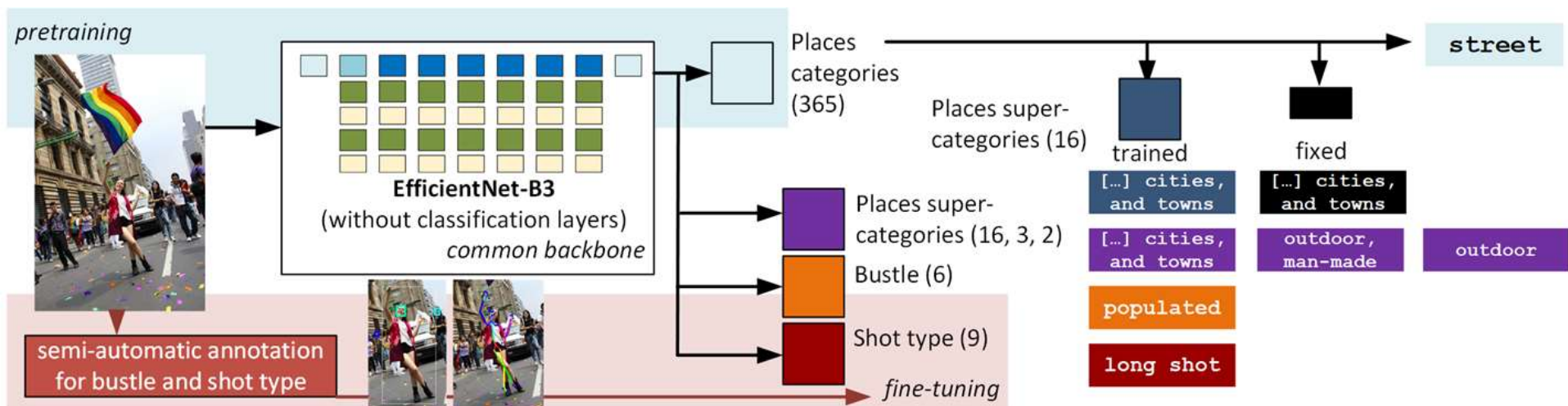
Few-shot learning

- learn new objects from few (~10) annotated samples
- automation of training process: dataset preparation
- running training and validation steps
- framework based on [3]
- integration with MakeSense annotation tool
- experiments with adding 7 shot classes to COCO



Efficient classification of shots

- combine per-shot classifications
- semi-automatically create dataset for shot type and bustle
- classifier based on EfficientNet-B3 [4]



Taylor

- Personalized AI named Taylor
- Connected to popular science fiction stories
- Elements of gamification
- Taylor should be seen as a "person" that can explain
- How good can an AI explain its interpretations/motivations?
- Key for making big "media data treasures" usable/available



Archive Interface - controlling/editing

The screenshot displays the 'Tailored Media' software interface. At the top, there is a navigation bar with 'Annotieren', 'Suche', 'Training', and 'Über' menus, and a user profile for 'MAX MÜSTERMANN' with an 'Ausloggen' button. The main window is titled 'DETAILS' and shows a video player with a leopard in a jungle. To the right of the video is a 'Timeline' view with a vertical playhead and various colored bars representing annotations. Below the video player is a 'Timeline' view with a playhead and a 'Relevanz' slider. The right sidebar contains a list of 'Informationen', 'Speech-to-Text', 'Petroren', and 'Landmark' sections, each with a dropdown menu. A 'Speichern' button is located at the bottom right of the sidebar. On the left side of the slide, there is a decorative graphic of overlapping hexagons in yellow, orange, and blue.

Archive Interface

Non linear view

- tag confidence color coded
- click to jump into linear view
- Taylor explains results and supports with suggestions
- tags based on a tag-tree

Linear view -> video editing interface

- timeline
- clip
- tags overview
- tag categories
- tags
- editing interface



TalendStudio UI

localhost:5080/tasks


TAILORED MEDIA

AS

Logout

OVERVIEW

c-tv 159 > All parts



0:13 / 26:50

Person

Landmark

Term

Object

Speech2Text

ShotType

Person

33 tags in total

Landmark

0 tags in total

Term

0 tags in total

Object

0 tags in total

00:00:00 00:09:00 00:18:00 00:27:00 00:36:00 00:45:00

00:00:00 00:09:00 00:18:00 00:27:00 00:36:00 00:45:00

Search Interface - Journalists

The screenshot shows the Tailored Media search interface. At the top, there are navigation tabs: "Annotieren", "Suche" (selected), "Training", and "Über". On the right, there are language options "DE | EN", a "Hilfe" button, and a user profile for "MAX MUSTERMANN" with the text "Ausloggen".

The main search area contains a search bar with the text "Magnus Brunner (Finanzminister*in) Ministerium (Finanzen)". Below the search bar, there are two rows of filters. The first row is for "Magnus Brunner (Finanzminister*in)" and the second row is for "Ministerium (Finanzen)". Each row has a "spezifischer Eigennama" filter and a "H" filter. The "H" filter for the second row has sub-filters for "Video", "Audio", and "Transkript".

Below the filters, there is a "Sortieren" dropdown menu set to "Reifender - Absteigend".

There are two bar charts. The first chart is titled "Zeitraum" and shows a bar chart for the years 1950 to 2010. The second chart is titled "Clippings" and shows a bar chart for the years 04 to 40.

At the bottom, there are four search results. The first result is "Finanzminister Brunner: Presskonferenz" with a thumbnail showing a man in a suit. The second result is "Finanzminister interview" with a thumbnail showing a man in a suit. The third result is "Minister Brunner" with a thumbnail showing a man in a suit. The fourth result is "Im Gespräch: Finanzminister Brunner" with a thumbnail showing a man in a suit.

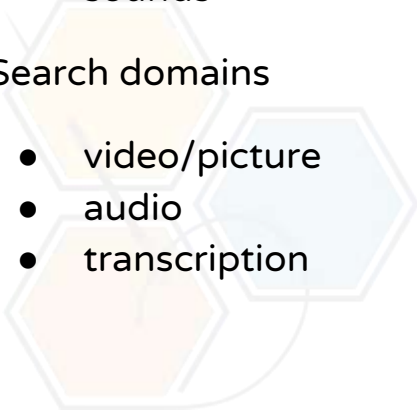
Search Interface



Input field for

- text
- pictures
- videos
- sounds

Search domains

- video/picture
 - audio
 - transcription
- 

Text input

- if text known by AI -> transformed into tag
- if no transformation possible -> search can only performed in transcription
- tags based on same tag-tree as archive

TailoredMedia Ltd

localhost:3000/search



TAILORED MEDIA

Annotate Search About

DE | EN Help Projektraum Logout

TAILORED MEDIA

Enter term...

 **Taylor:**
Hi, what can I search for you?
Please enter your search term in the search bar or upload a media file above ().

Search Interface




Taylor explains

- what was searched
- what was the interpretation of search terms/tags (meaning of tag)
- in which domains was searched
- where the results were found

Meta-communication with Taylor

- does it understand the term?
- does it know the term?
- change tags or search domains
- add new search terms/tags

Q&A



Christoph Bauer

(christoph.bauer@orf.at)

Werner Bailer

(werner.bailer@joanneum.at)

Peter Judmaier

(peter.judmaier@fhstp.ac.at)



The research leading to these results has been funded partially by the program “ICT of the Future” by the Austrian Federal Ministry of Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK).