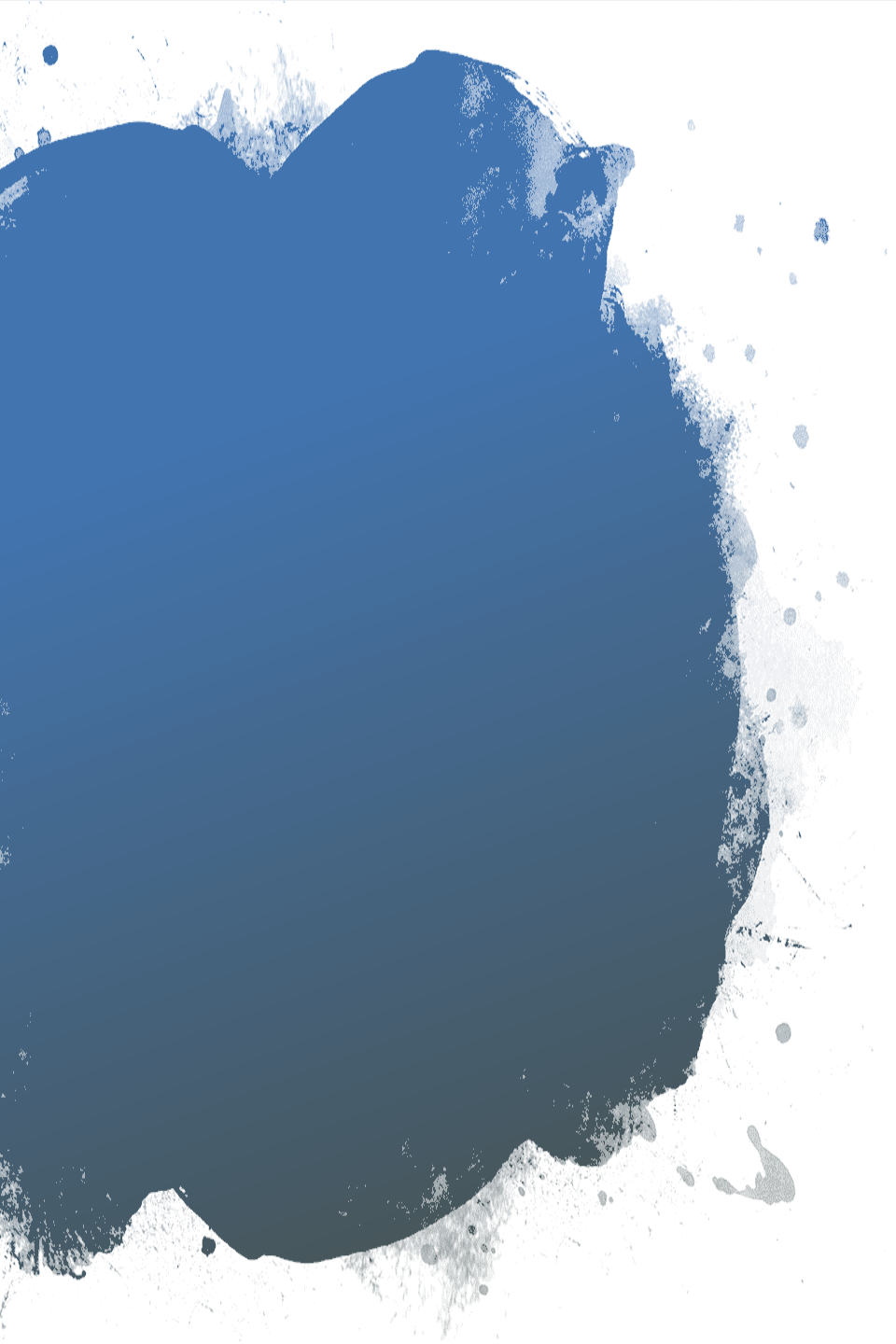
Engaging Content
Engaging People

Social-media Storytelling Linking

Hao Wu
Seamus Lawless
Gareth Jones
Francois Pitie

Ireland's European Structural and Investment Funds Programmes 2014-2020
Co-funded by the Irish Government and the European Union
European Union
European Regional Development Fund
SFI Science Foundation Ireland

The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.



- Task definition
- Challenges & Solutions
- Training
- Searching
- Result

Mercy

PRESENTS

BITES+BREWS+BANDS

FARMERS MARKET PAVILION

Tour France

Story segment 1	Story segment 2	Story segment 3
The Edinbrugh Festival is home to one of the most unique celebrations of arts.	Street circus is a popular attraction at Edfest with peculiar performances such as unicle juggling.	Street circus is full of colorful artists.

Segment 1 illustration	Segment 2 illustration	Segment 2 illustration

www.adaptcentre.ie

Challenges & Solutions

Lack of training data

Video can’t be concluded by only one sentences.

Challenges

Pre-train + Fine tuning

Video segmentation

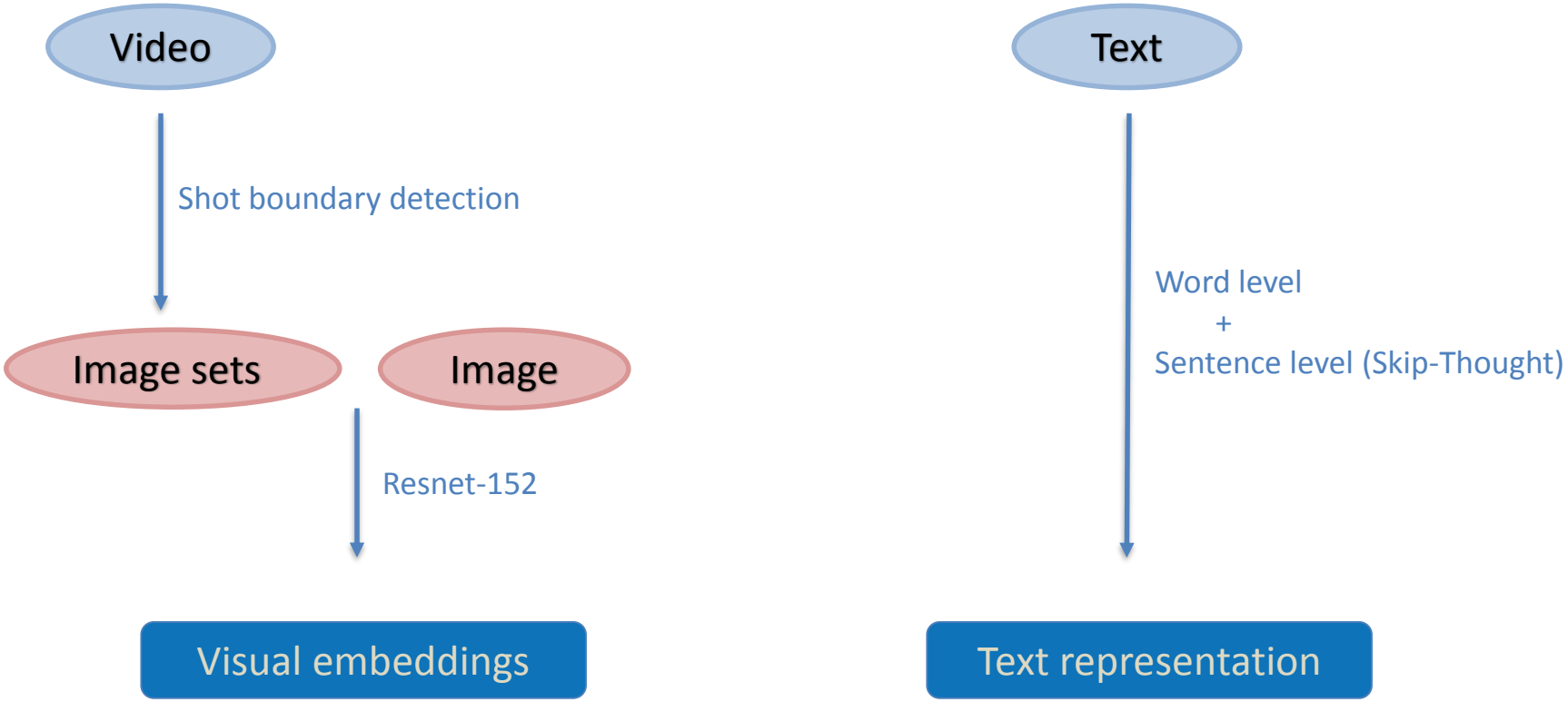
+

Length normalization

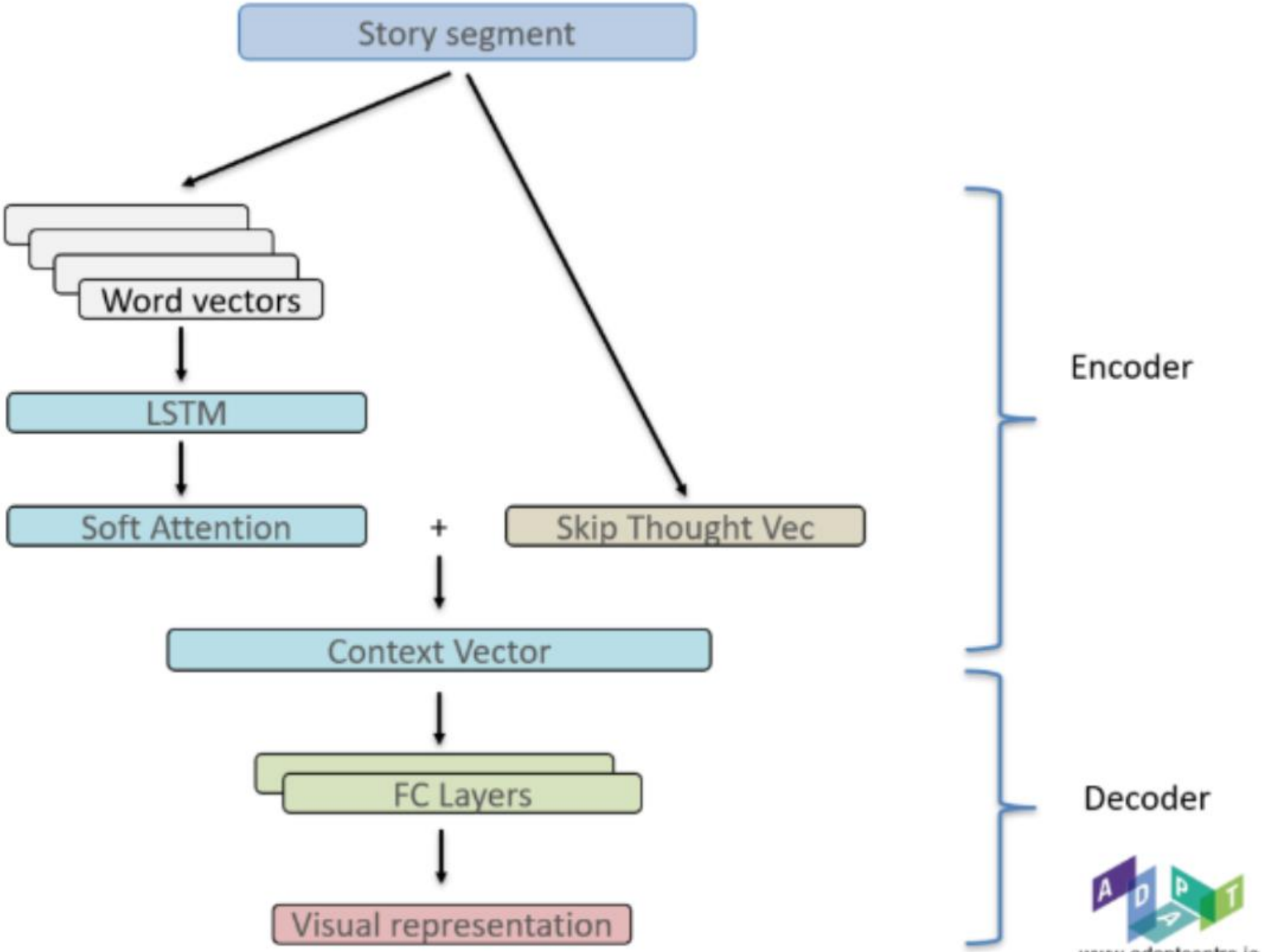
Solutions

Data
pre-processing

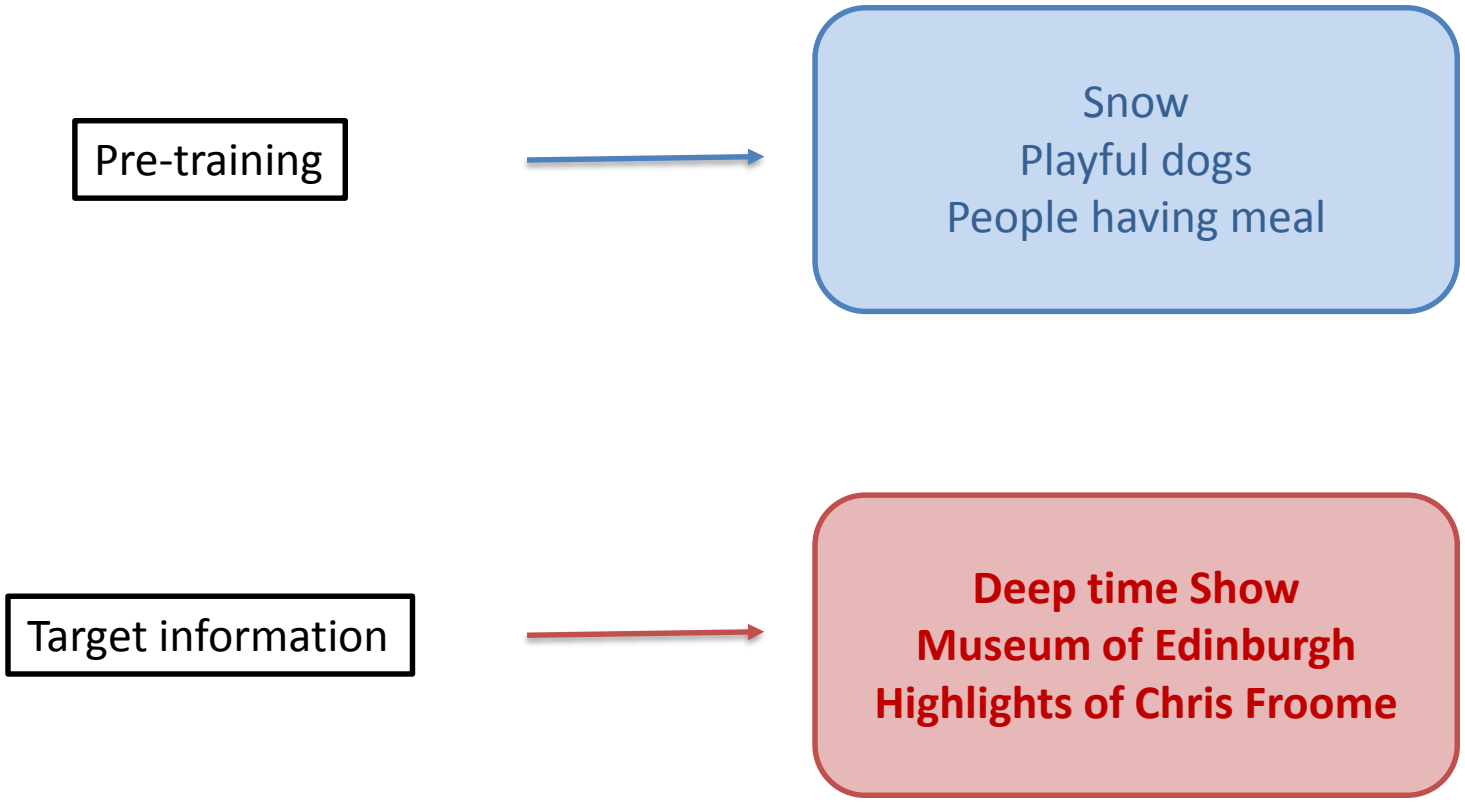
	Images	Videos	Queries
Edinburgh Festival	32k	6.2k	60
Le Tour de France	66k	19k	58


Video
Shot boundary detection
Image sets
Image
Resnet-152
Visual embeddings
Text
Word level
+
Sentence level (Skip-Thought)
Text representation

Model overview


Story segment
Word vectors
LSTM
Soft Attention
+
Skip Thought Vec
Context Vector
FC Layers
Visual representation
Encoder
Decoder
www.adaptcentre.ie

Training


Pre-training
Snow
Playful dogs
People having meal
Target information
Deep time Show
Museum of Edinburgh
Highlights of Chris Froome

Examples

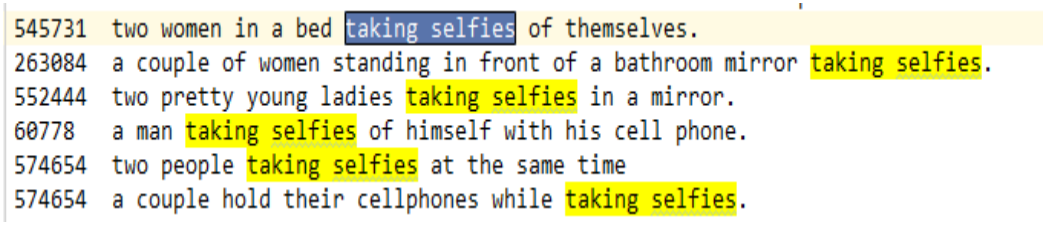
A boy in a dark shirt is reading a book while sitting on a piano bench

Pre-training

Introducing **Flickr30k** (High quality “image”- “text” pairs)

Collecting from source domain:

- Identify keywords from query file.
- Match keywords with data in the source.



E.g. Keyword: **taking selfies**.

Collecting from search engine:

- Collect labels from online image search engine (Google and Bing) using story segments + event name as query.

Model

Target information collecting

Chris Froome pedaling

Snow

Searching

Trade-off between consistency and accuracy

$$R_t = 0.2 * R_{t-1} + 0.8 * M_t$$

(M is the model raw output, R is the modified output)

Search

There are 5 runs submitted. The main difference is the value of λ:

Conf	Run1	Run2	Run3	Run4	Run5
λ	3	5	12	20	50
Source	Google+ Bing	Google	Google	Google	Google

$$\bar{S}_c = S_c * Sig(\frac{\lambda}{\log_2 L})$$

λ used in penalizing long videos;
L denotes number of segments;
Sig() is sigmoid function.

$$Sig(x) = \frac{e^x}{e^x + 1}$$

Search

Results

Summary Quality
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Run1
Run2
Run3
Run4
Run5
Edfest
Tourfrance

Conclusion & Future Work

Target specific information are crucial.

Improve video representations by applying key frame selection (or building sequence model).

Build a classifier to filter crawled images to make this process automatic.

Engaging Content
Engaging People

Thanks for listening.

Ireland's European Structural and Investment Funds Programmes 2014-2020
Co-funded by the Irish Government and the European Union
European Union
European Regional Development Fund
Science Foundation Ireland

The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.