

NewsImages: The Role of Images in Online News

Data Lab Project central:

<https://datalab12.github.io/work/mediaEval2021.html>

Yuxiao Zhou

Parisa Tabassum

Andres Gonzalez

Jelena Tešić

Image-Text Re-Matching

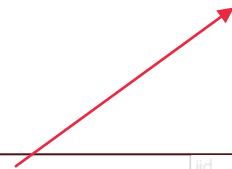
In this project, we will be gathering data to establish a link between images and text in the hopes of correctly matching a cover image to its associated news story.

News articles online have long presented with headlines, text, and an associated image to capture the emotion of a news story. This project aims to study the potentially complex, multimodal relationship between the article's textual content and its presented image.



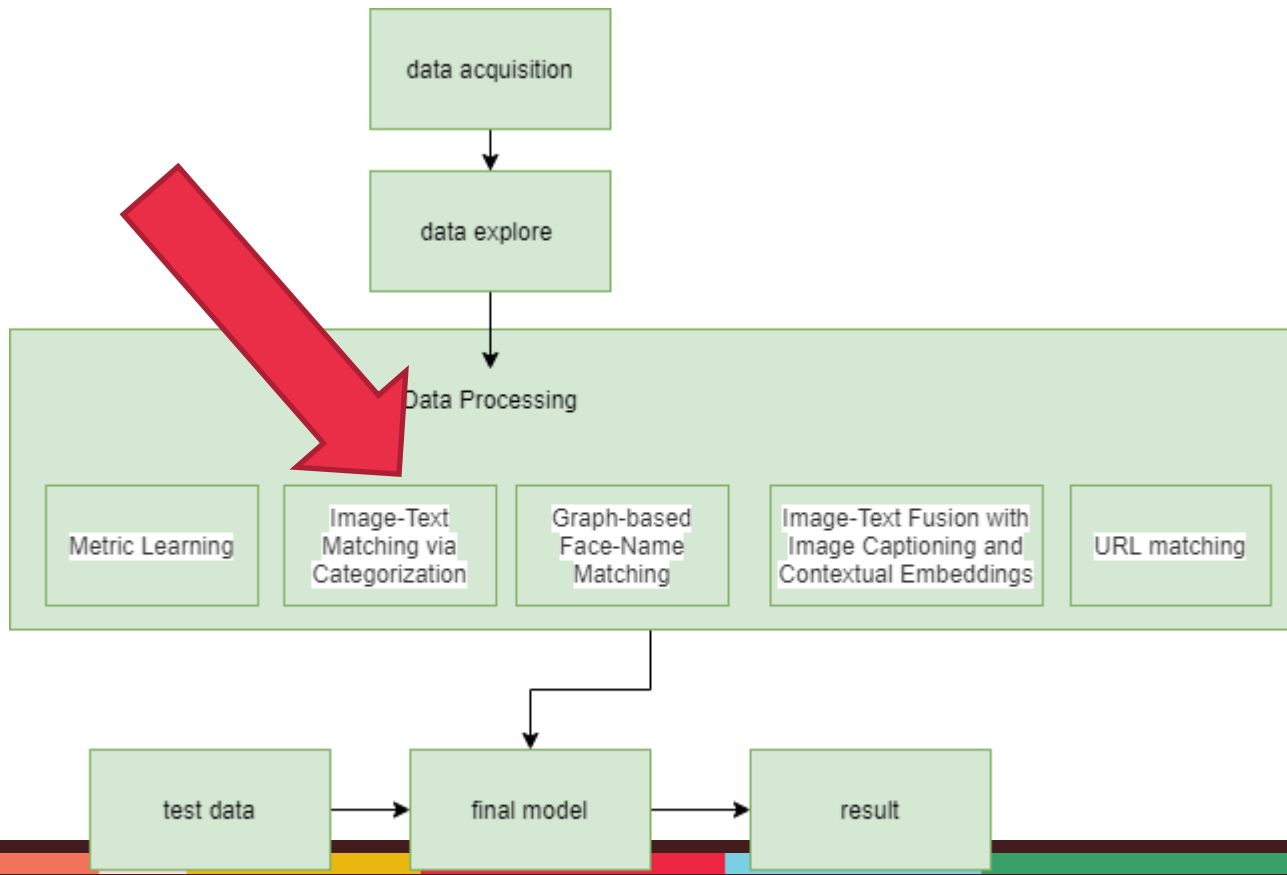
Data Features Used

- Article
- AID
- url
- Img
- IID
- Hashvalue
- Title
- Text
- nImpressions
- nRecs
- nClicks
- imgFile



article	aid	url	img	id	hashvalue
509005634	116171	https://www.ksta.de/panorama/selbstgebauer-knaller--mann-	https://www.ksta.de/image/31811446/2x1/300/150/c816ac031a5	116805	
509006281	116172	https://www.ksta.de/wirtschaft/neuanfang-im-job-wie-der-koe	https://www.ksta.de/image/31808308/2x1/300/150/85ee043d58	116806	
509006327	116173	https://www.ksta.de/wirtschaft/interview-ueber-jobw-echsel--w	https://www.ksta.de/image/31808358/2x1/300/150/806096e3cf	116807	
509006329	116174	https://www.ksta.de/panorama/panne-bei-feuerwerk-sydneys	https://www.ksta.de/image/31811494/2x1/300/150/cf18c73a2fa3	116808	
509007066	116175	https://www.ksta.de/politik/us-sanktionen-kim-jong-un-droht-i	https://www.ksta.de/image/31811554/2x1/300/150/29c50be85cf	116809	

title	text	nImpressions	nRecs	nClicks	imgFile
Mann wird durch BÄÄtler schwer a	In der NÄ*he von Celle ist ein 38-jÄ*hriger Mann am Montagabend durch einen Knall	14	0	0	31811446c816ac031a5a0b3add7d47a3813aec62rC.jpg
Wie der KÄÄtler Daniel Opoku sei	Daniel Opoku, Outdoor-Kleidung, schlanke Statur, fester HÄÄndedruck, offenes LÄÄ	1249	0	0	3180830885ee043d58cd779d33cca38888e6b57vZ.jpg
ÄÄtler auf Dauer im Beruf unglÄÄt	Herr Conen, gibt es so etw as wie ein ÄÄtler den Neuanfang?Horst Conen: Ja	36	0	0	31808358806096e3c65a13dcfd59cde588e7eb0auc.jpg
Sydneys Feuerwerks-Meister unti	Australiens Metropole Sydney hat das neue Jahr wieder mit einem spektakulÄÄren Fe	35	0	0	31811494cf18c73a2fa30f16a949044c88b9be2x.jpg
Kim Jong Un droht mit Abkehr von	Nordkoreas Machthaber Kim Jong Un droht im Streit um das Atomw affenprogramm se	17	0	0	3181155429c50be85cf894763325d201ab241bc6oS.jpg




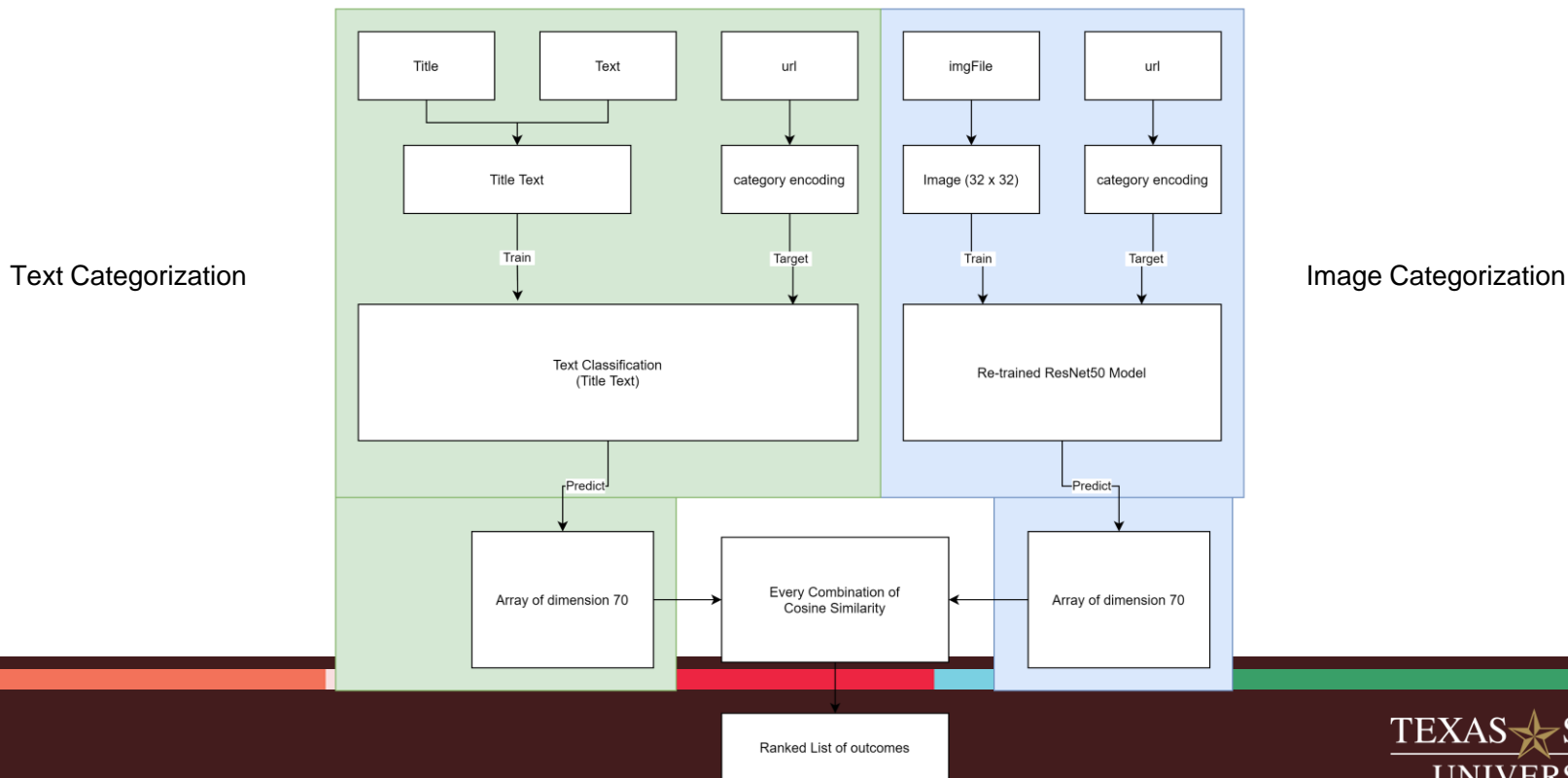
 Complete

Image-Text Matching Via Categorization



{Article} URL to Category

<https://www.ksta.de/panorama/selbstgebauter-knaller--mann-wird-durch-boeller-schwer-an-der-hand-verletzt-31811448>



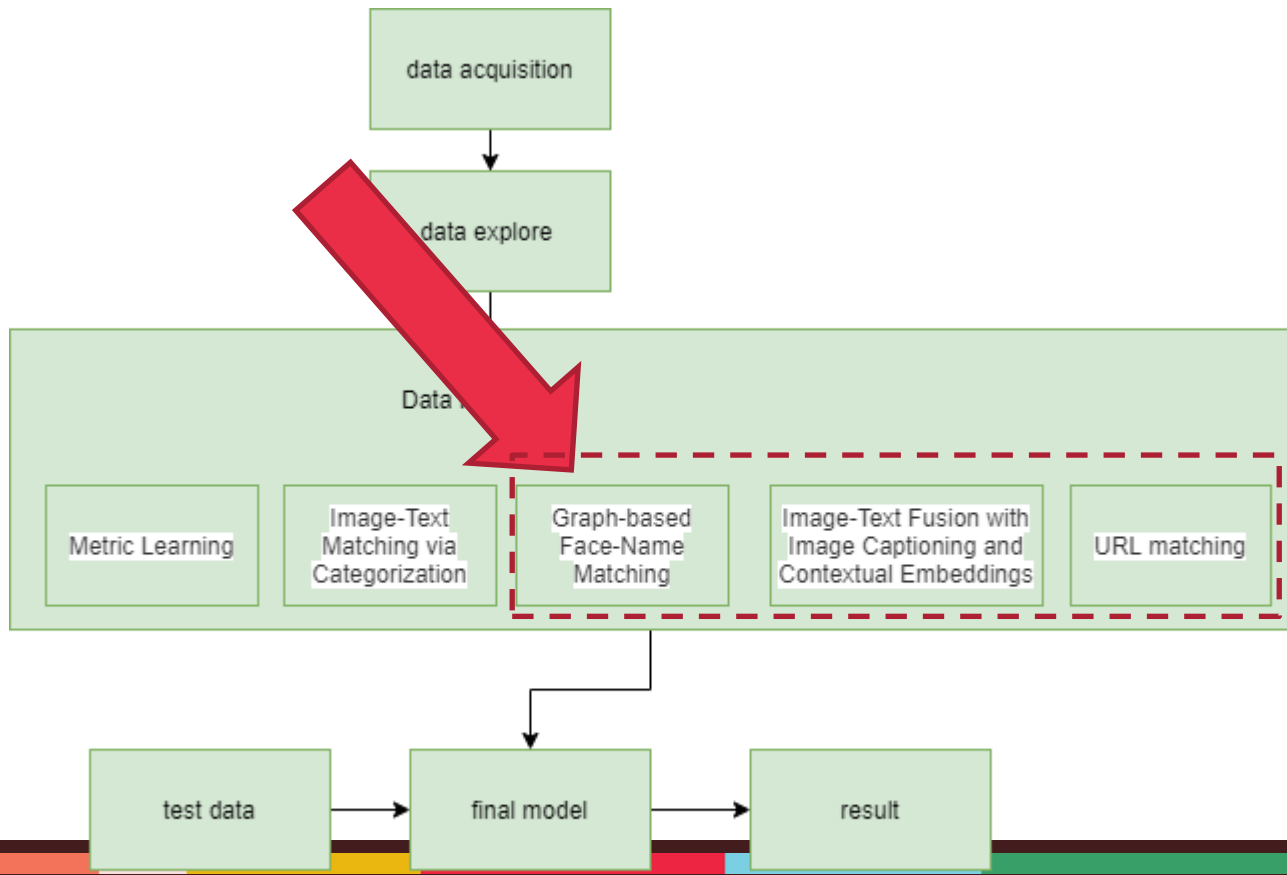
<https://www.ksta.de/panorama/selbstgebauter-knaller--mann-wird-durch-boeller-schwer-an-der-hand-verletzt-31811448>


Text Classification

- Load and process the data
- Encode the target
- Construct vectorization layer
- Construct the model
- Fit the model
- Use the model to predict

Image Classification

- Load and process data
- Load images
- Construct a re-trained ResNet50
- Fit the model
- Use the model to predict



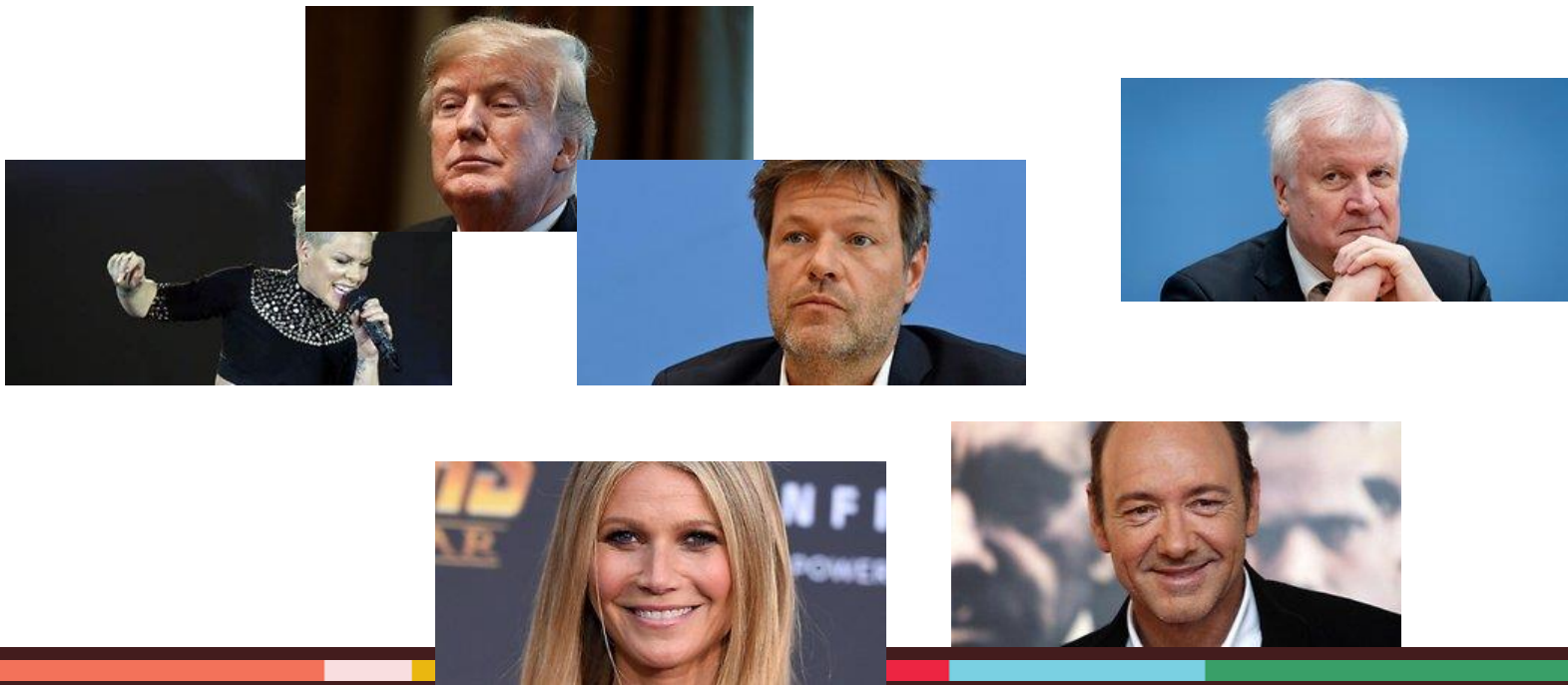
 Complete

Face-Name Matching

In many image-article pairs, the publisher uses a portrait of person which mentioned in the article as the news image. Therefore, we may re-match images and texts by matching the names within the texts and the faces within the images.

- Text Translation
- Name Extraction
- Name Face Matching
- Face encoding

Face Name Matching



Face-Name Matching

- The Stanford Named Entity Recognizer (NER) is adopted for the named entity recognition. The NER provided named entity recognizer particularly for the 3 classes (PERSON, ORGANIZATION, LOCATION). We mainly focus on the person's name extraction.
- By performing the NER on the 7530 given news article title, it is found that at least 24% of them include person's name.

Title	EnglishTitle	PersonName
Vorstandschef Rummenigge: Beckenbauer „thront ...	Board Chief Rummenigge: Beckenbauer "Thront ov...	Rummenigge,Beckenbauer
Spahn: Krebs ist in absehbarer Zeit besiegbar	Spahn: Cancer is defeated in the foreseeable f...	Spahn
Bergisch Gladbacher Ordnungsamt künftig bis 22...	Bergisch Gladbacher Ordnungsamt in the future ...	Bergisch Gladbacher,Ordnungsamt
Grundsteuer-Durchbruch? Spitzentreffen bei Scholz	Property tax breakthrough?Top meeting at Scholz	Scholz
Venezuelas Gegenpräsident Guaidó will nicht me...	Venezuela's president Guaidó does not want to ...	Guaidó
...
Bücherei in Paffrath ab sofort geschlossen	Library in Paffrath now closed	Paffrath

Face-Name Matching

- Article titles include a person name, also use the portrait of the mentioned person as image for that article. In the other words, there should be a matching relationship between an image and the recognized person's name from the corresponding new article title.



118876



119004



119228



119238



119610



119798



119827



120653



120898



121015



121502



121544



121659



121873



122436



124762



124825



125020



125044



126492



126494



126760



126835



126841



127004



127007



127396



127808



128121



128710



128740



128770



129467



129919



130640



131504



132012



132542



133159



133689



133987

Image Captioning

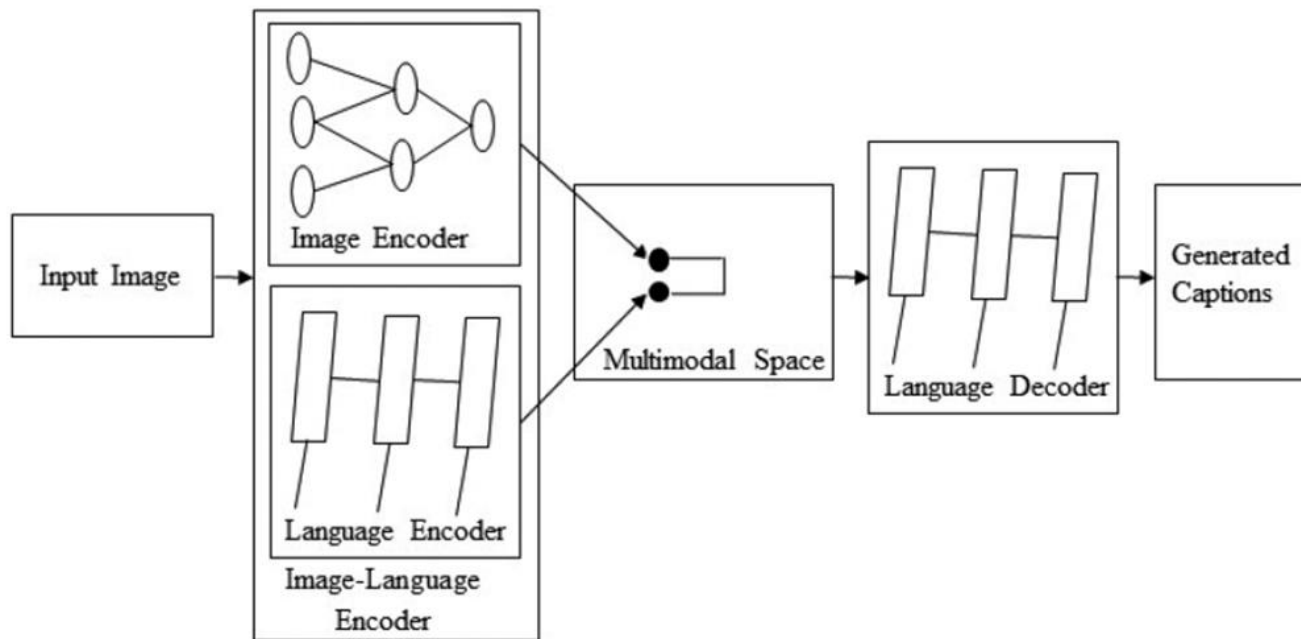


Image Captioning

We first adopt an image captioning model pre-trained with COCO dataset for image caption generation



134147

134147.jpg



134149

134149.jpg



134150

134150.jpg



134151

134151.jpg



134154

134154.jpg



134155

134155.jpg

a group of men standing next to a building with a cake

a group of people standing next to each other

a red bus is driving down a street

a group of scissors sitting on top of a table

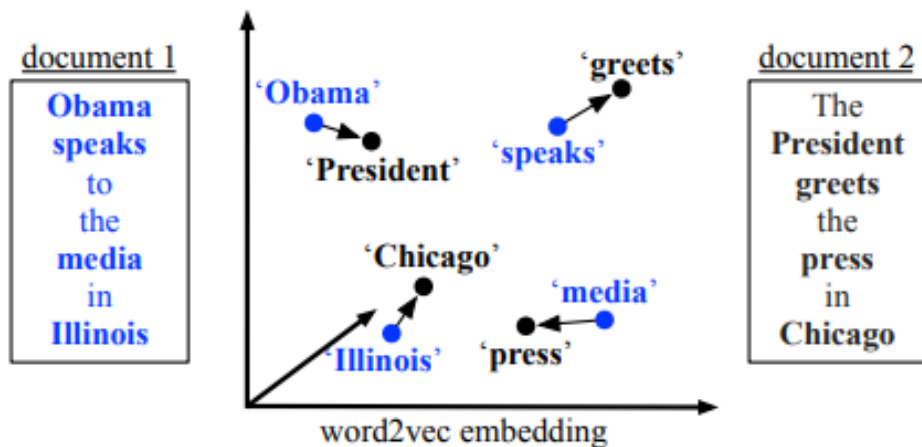
a man and a woman sitting on a table with a laptop

a man is holding a glass of wine in a table

Image Captioning

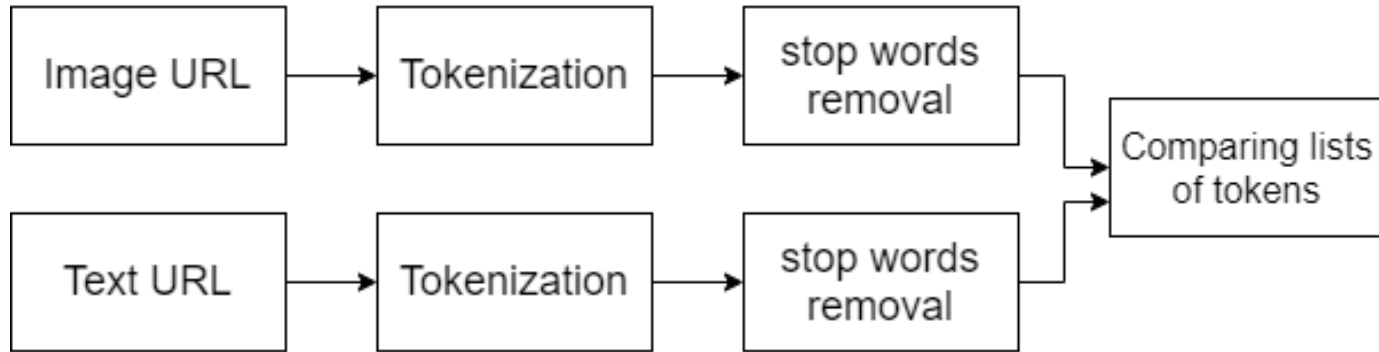
Then we calculate the similarity score between the generated image captions and given news headlines.

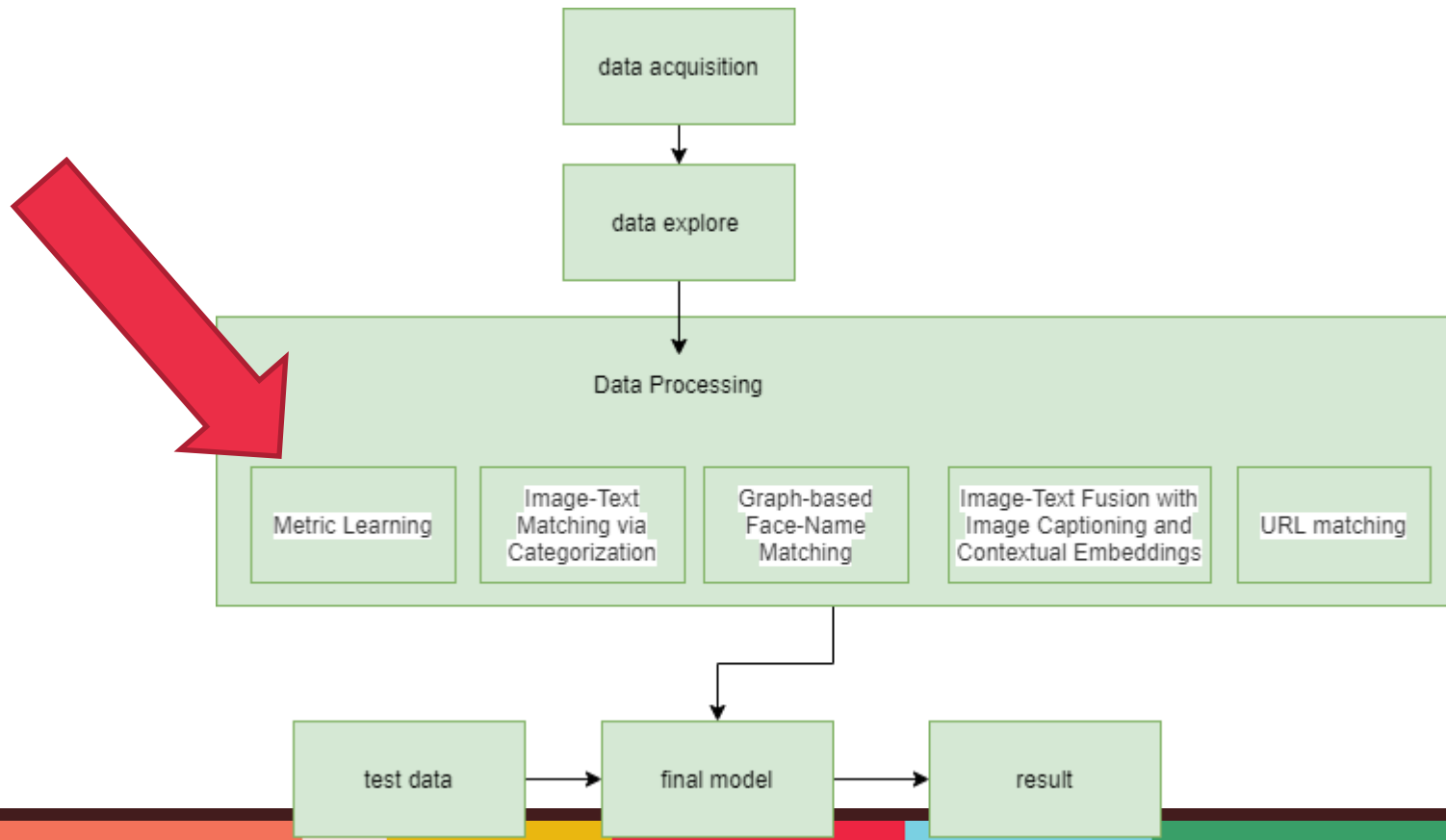
We use the Word Mover's Distance (WMD) to compare the similarity between image captions and articles title. The WMD algorithm uses normalized Bag-of-Words and word embeddings to calculate the distance between documents/



URL Matching

The image URL and text URL pairs demonstrate some explicit relationships between them. More specifically, an image and a text may be matched if their URLs contain one or more common tokens.





Complete

Initial Model

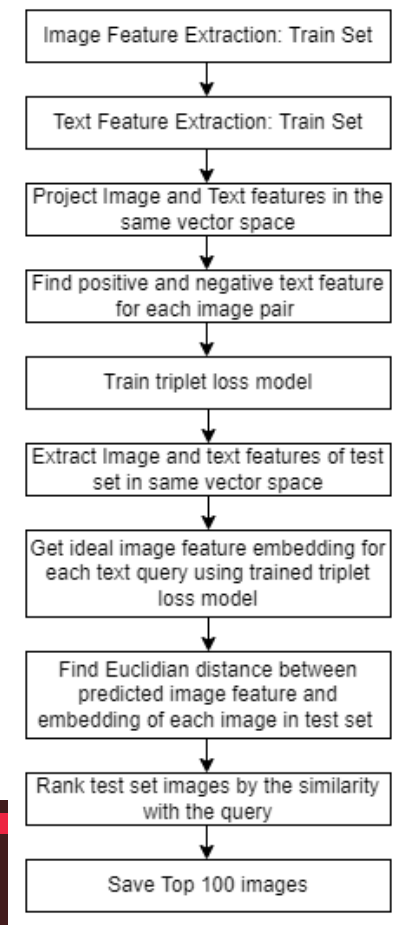
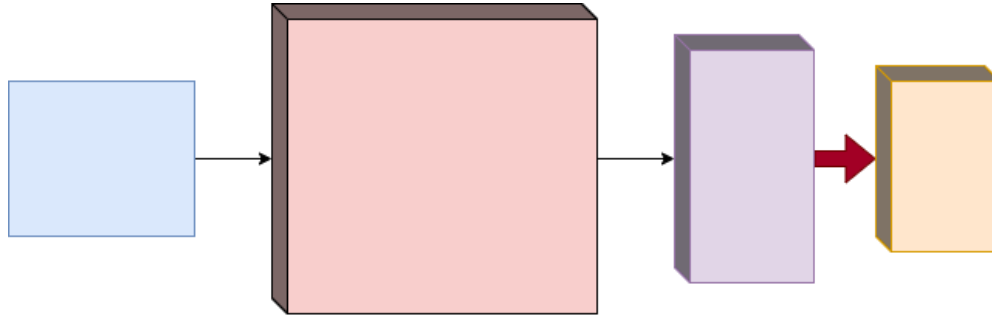





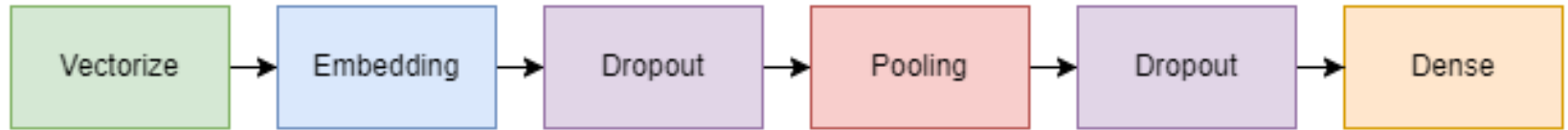


Image Feature Extraction

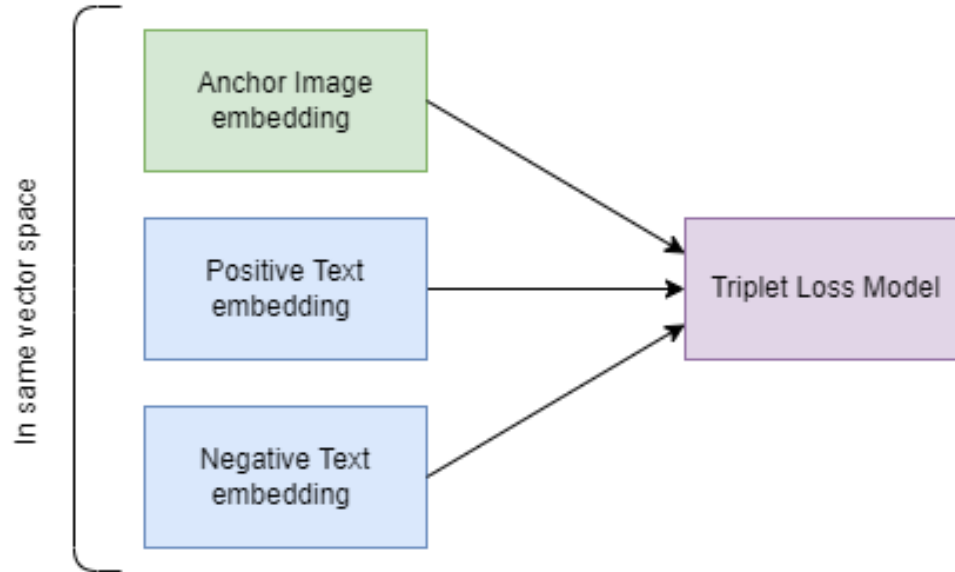


-  Input Image (Prepressed for EfficientNet)
-  EfficientNet without fully-connected layer at the top
-  Global Average Pooling
-  Dropout Regularizer
-  Dense Layer + Softmax

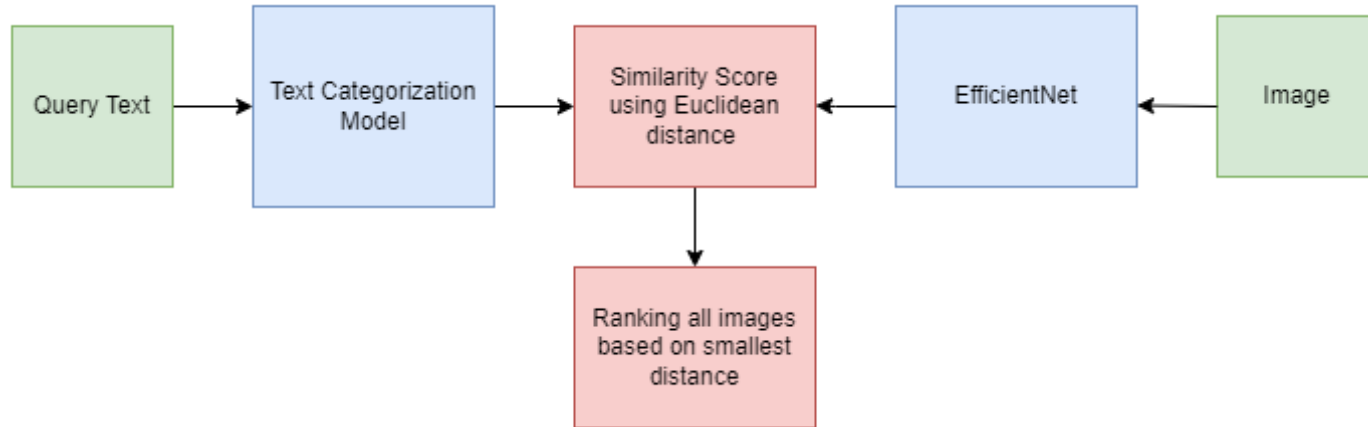
Text Feature Extraction



Triplet Loss Model



Final Model



Submissions

Our proposed approach uses an ensemble design, so our submissions are combined results from three or four models.

- Run1 combines three different methods, equal weights are assigned to categorization-based method, and combination of face-name matching and image captioning-based method. The ranking of a candidate image in Run1 is:

$$R_{Run1} = 0.5R_{Categorization} + 0.5(R_{Face} + R_{caption})$$

- Run2 combines all proposed methods. The first three models are ensembled using the same approach as in Run1. This ensembled model is used for creating the initial top 100 image list. Then we append the result, which generated from URL matching based method, to the end of the top 100 image list.
- Run3 is similar to Run2. The only difference is that we append the result of last method to the head of the top 100 image list.

Result and Conclusion

Ru n	Recall@10 0	MR@5	MR@1 0	MR@5 0	MR@10 0
1	0,00668	0,0083 6	0,0109 7	0,0297 7	0,05274
2	0,01147	0,0083 6	0,0109 7	0,0302 9	0,49347
3	0,28788	0,3718	0,4094	0,4668 4	0,49347

image-text relation-based model can be used for news image re-matching prediction, but it seemed poor performance, while the usage of text feature, the image URL, can improve the performance a lot. In future, we plan to incorporate metric learning in our model. We also plan to conduct the image-text matching experiment with improved features like news image caption with embedded named entities or metadata. Further, the ensemble may be extended by applications of techniques such as bagging, boosting and stacking.

Thank you!

Q and A

More Info:

<https://mediaflo.txstate.edu/Playlist/NewsImages>

<https://github.com/DataLab12/newsimages>