Artificial Intelligence Project
Interim Presentation

2020095178 YoonSeon Choi (최윤선)
2021029443 YeEun Jeon (전예은)

Dating apps

I want my future partner to be:
- 27-29 years old
- School above Hanyang University
- Etc..

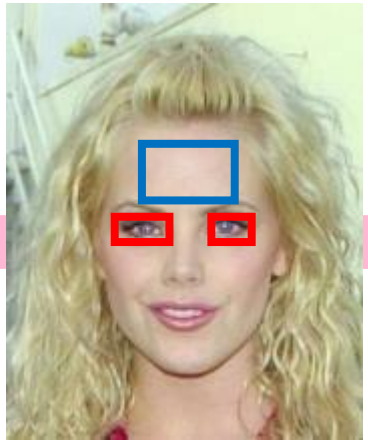
Hmm.. I am afraid to show my photo, but I am curious about the appearance of the man that will show up in blind date.

“Get to know the image in a situation where we can’t see the image.”

How?

Samantha

User uploads photo



Skin color:

✓bright ✗normal ✗dark

Eye size:

✗small ✓normal ✗big

Classify facial attributes by standard

Bright, normal

Dark, big

Bright, big

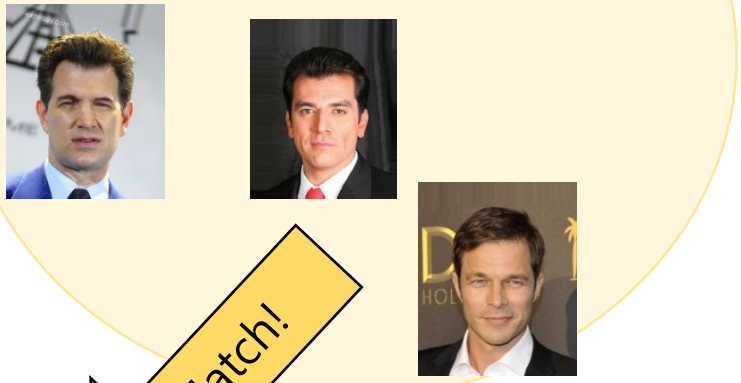
Get huge number of user data with labels

Bright, normal

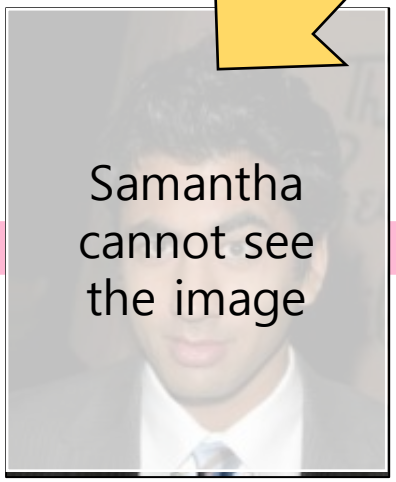
I want to meet a man with

✗bright ✗normal ✓dark skin and

✗small ✗normal ✓big eyes



Match!



Dark, big

Instead, Samantha see:

A man with big nose and curly dark black hair is wearing a suit.

User's ideal preference information

User cannot see the photo

Instead, user sees the word description of photo

What? Data Preprocessing (completed)

What? Data Preprocessing (completed)

<Total 200,000 facial images>

- Classify gender
- Criteria of image selection
  - Front view photos
  - Remove the photos which have sunglasses, unusual lightning, big smile, etc...
- Final selected images
  - Men: 9,429
  - Women: 10,000

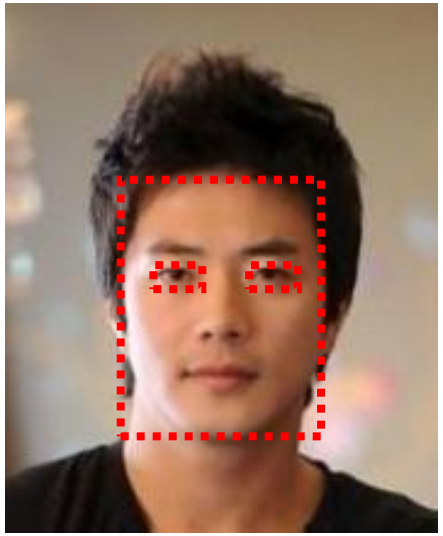
<selected>

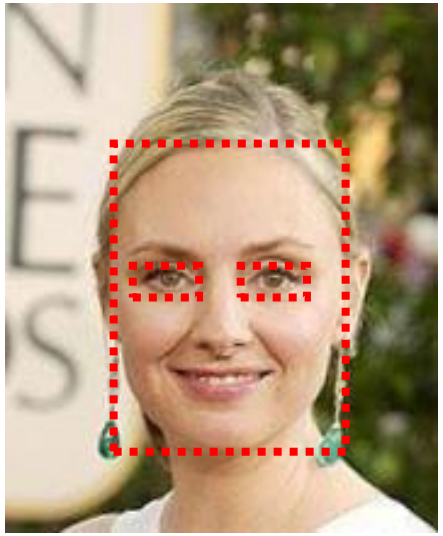
<removed>

What?

<Facial Image Detection>

- Faster R-CNN
- MTCNN





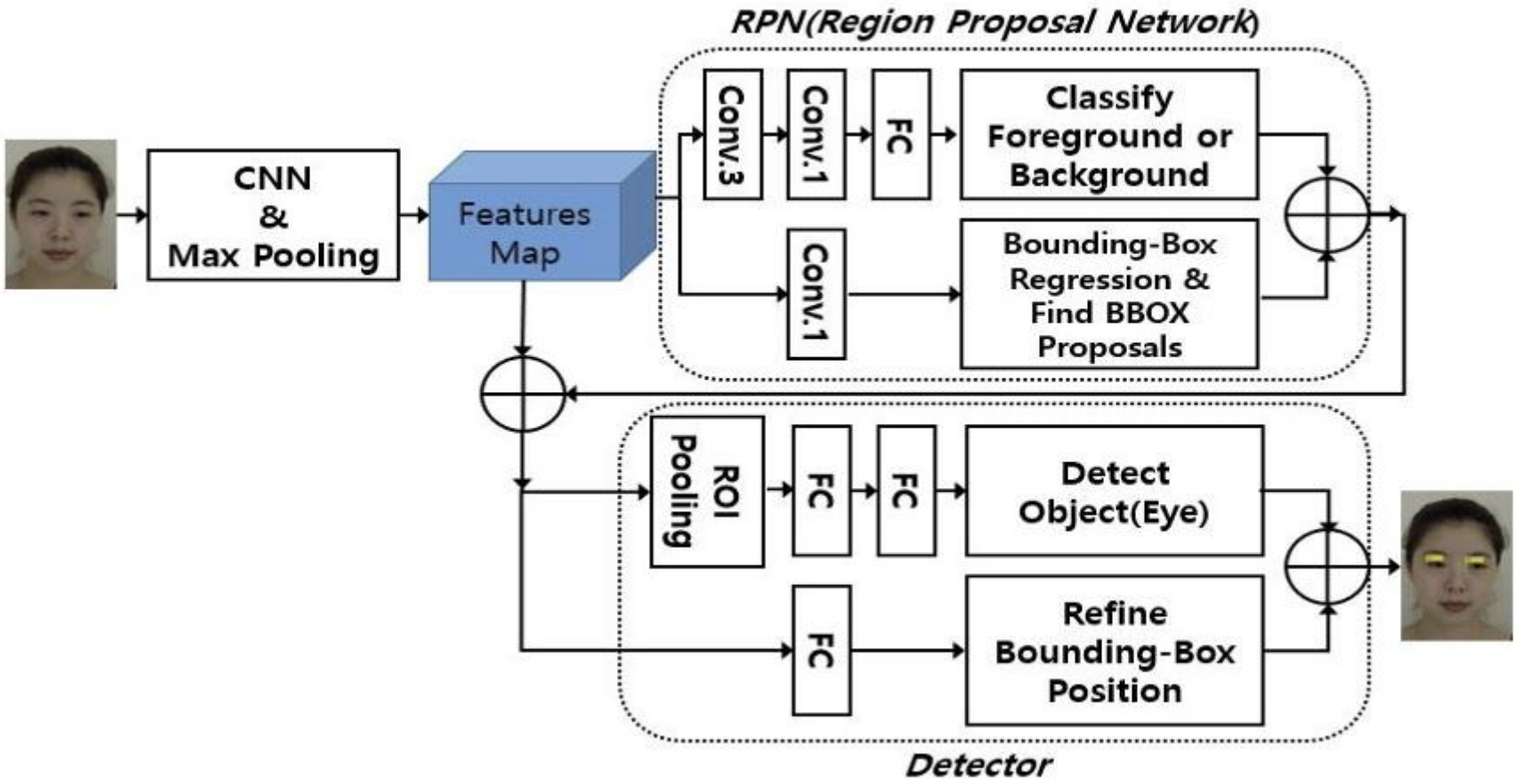
<Image Captioning>

- ResNet50-LSTM
- Transformer-based

He is wearing glasses and has beard...

What? for Facial Image Detection (in progress))

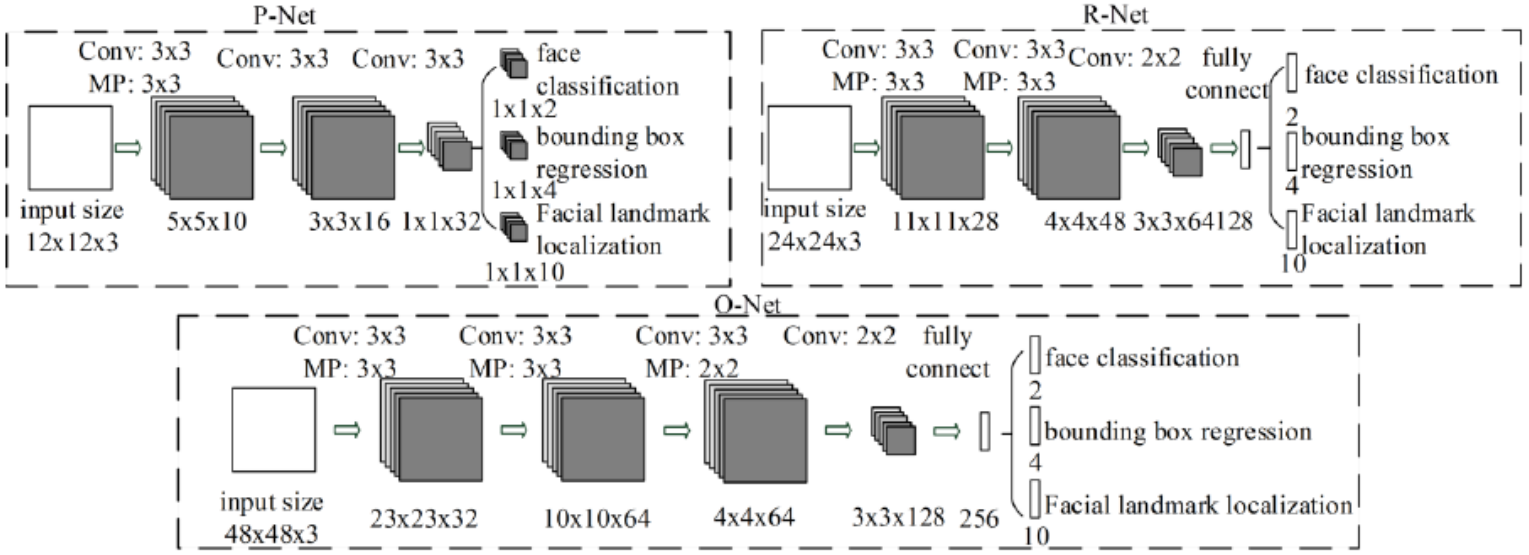
- **Faster R-CNN**

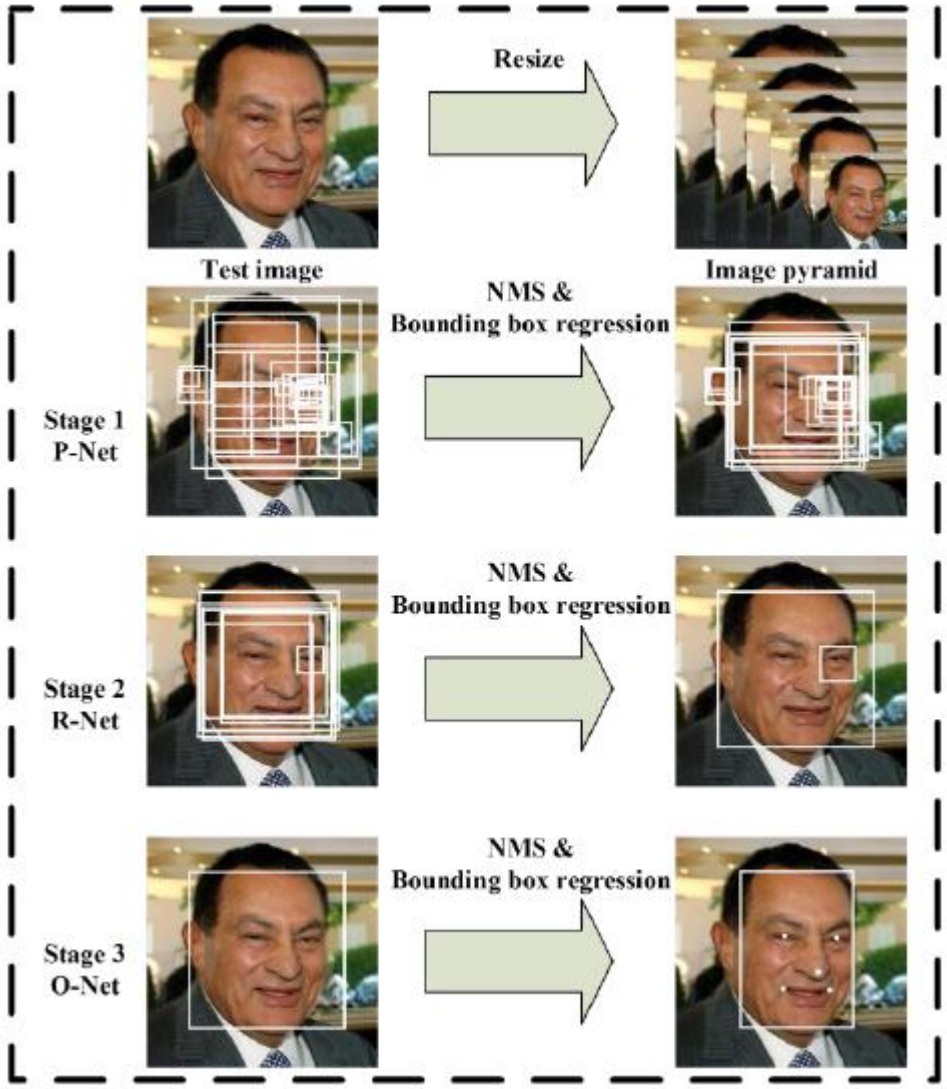


Journal of the Korea Convergence Society, <A Method of Eye and Lip Region Detection using Faster R-CNN in Face Image>
Jeong-Hwan Lee Department of Electronic Engineering, Andong University

What? for Facial Image Detection (in progress))

- **MTCNN**





IEEE <Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks>
Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, *Senior Member, IEEE*, and Yu Qiao, *Senior Member*

What? for Image Captioning

- **ResNet50-LSTM**

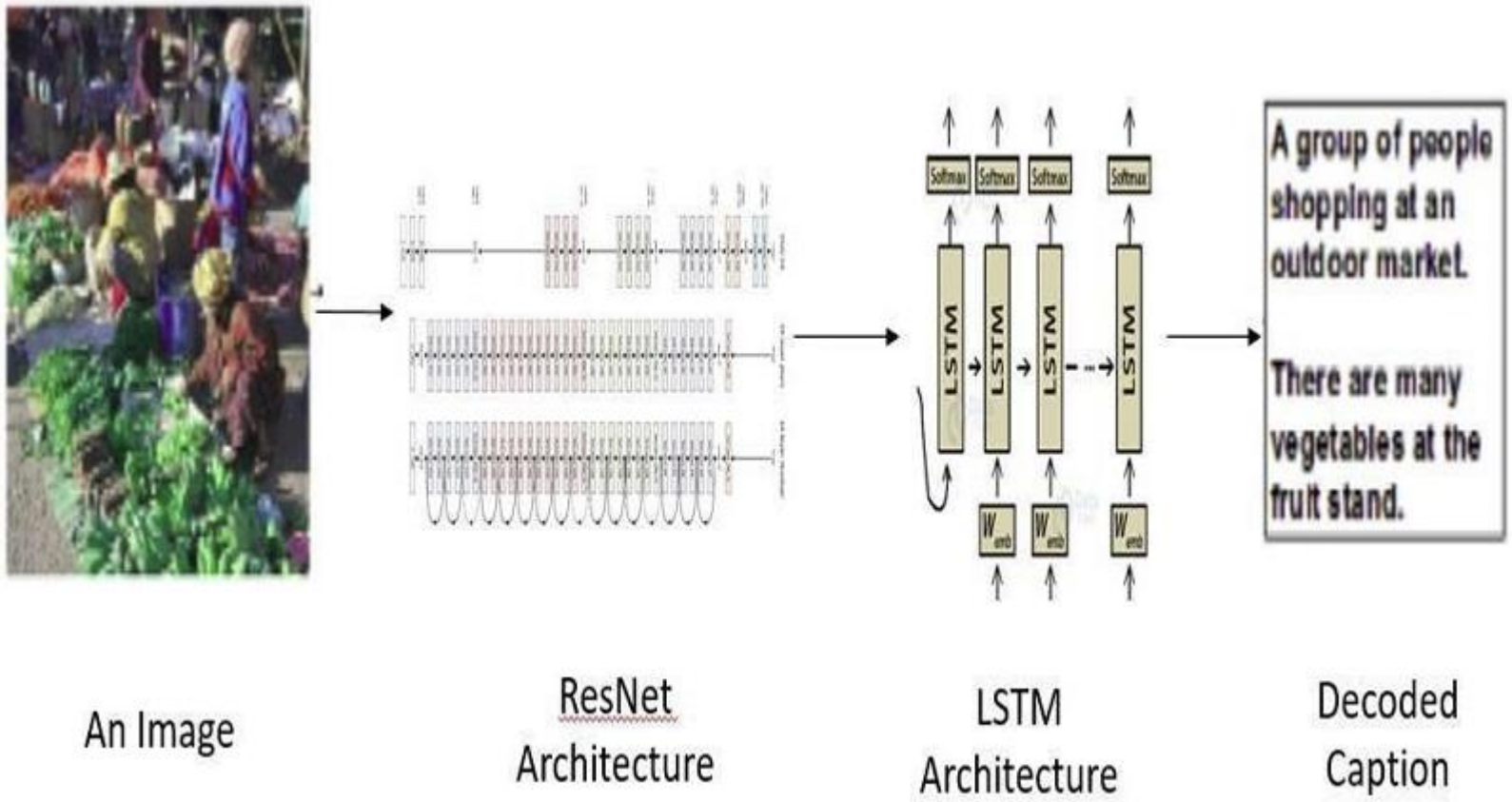
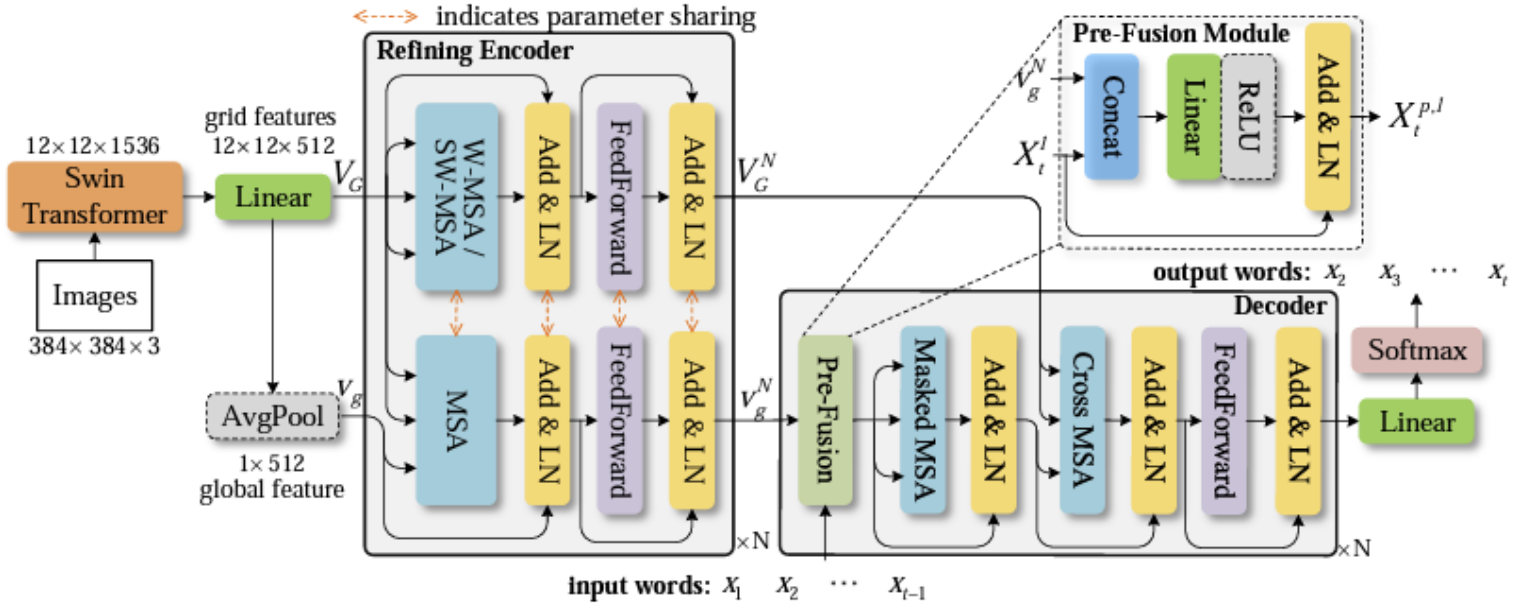


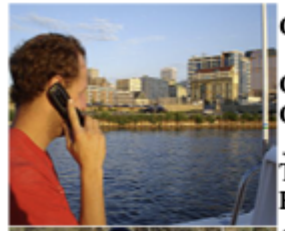
Figure 2: Image Caption Generation Architecture

IJRES <Image Caption Generator Using RESNET-LSTM>, Shobiya L

What? for Image Captioning

- **Transformer-based**





GT1: a man talking on a cell phone on a boat with a city in the background
GT2: a man is on a boat using his cell phone
GT3: a man is standing by the water and talking on the phone
$\mathcal{M}^2$: a man talking on a cell phone next to the water
Transformer: a man talking on a cell phone on a boat
PureT: a man talking on a cell phone on a boat in the water

GT1: a person sits on top of a motorcycle with a stuffed toy
GT2: a person riding a motorcycle with a stuffed animal on the back
GT3: a person on a motorcycle with a stuffed animal on back
$\mathcal{M}^2$: a man riding a motorcycle on a street
Transformer: a man sitting on a motorcycle with a teddy bear
PureT: a man riding a motorcycle with a stuffed animal on it

GT1: a giraffe stands with several birds resting on it's neck
GT2: a giraffe that has some birds perched on it
GT3: this is a close up picture of a giraffe that is standing.
$\mathcal{M}^2$: a close up of a giraffe with a bird
Transformer: a giraffe standing with a group of birds on it
PureT: a close up of a giraffe with birds on its neck

GT1: a gray day at a park with a stone bench
GT2: a sidewalk sitting near a green next to a body of water
GT3: a tree that is sitting in the grass
$\mathcal{M}^2$: a bench on the side of a dirt road
Transformer: a park bench sitting next to a body of water
PureT: a stone bench sitting next to a tree in a park

arXiv <End-to-End Transformer Based Model for Image Captioning>
Yiyu Wang, Jungang Xu, Yingfei Sun

When? Gantt Chart

Who	Task	Date	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
Together	Data Preprocessing (including midterm period)	April												Data Preprocessing																			
전예은	Object Detection Modeling (Faster RNN)	May	Object Detection Modeling																														
최윤선	Object Detection Modeling (MTCNN)		Object Detection Modeling																														
Together	Compare two object detection models & Choose one																Compare & Choose																
전예은	Image Captioning Modeling (Transformer-based)																			Image Captioning Modeling												
최윤선	Image Captioning Modeling (Resnet-LSTM)																			Image Captioning Modeling												
Together	Compare two image captioning models & Choose one	June	Compare & Choose																														
Together	Programming to show the results					Programming																											
Together	Prepare Final Presentation											Prepare Presentation																					

Thank you