

The Human Impact on Information Security

Paige Brown

Outline

- Introduction
 - Motivation
 - Project Description
- Method
 - Study One
 - Study Two
- Future Work

Motivation

- Improved ability to design systems and train users
- Understanding users can assist with understanding attackers
- Industry gap acknowledged but only anecdotally addressed

Project Description

- Exploring Human Error to understand risk to information security
- Using a case study approach to evaluate risk categories
- Developing recommendations for integrating Human Error into cyber risk assessment

Study One

- User study with 18 participants
- Exploring the problem space
- Three scenarios with personas

Results

- Five major themes
 1. Personally Identifiable Information
 2. Perceptions
 3. Mistakes
 4. Guesswork
 5. Expectations

Study Two

- User study with 6 participants
- In depth focus on Personally Identifiable Information
- Four scenarios with personas

Results

- Four major themes
 1. Optimism Bias
 2. Education
 3. Exposure
 4. Ease

Next Steps

- Further analysis of data from study two
- Developing recommendations

Deep Emotion Recognition

- **Presenter: Surekha Gaikwad**
- **Student ID: u6724013**

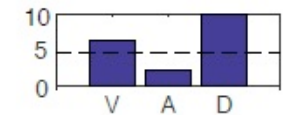


Motivation

- To identify and analyse the complete emotional state of the person.
- To understand the emotions of the person for non-frontal head poses.
- To make model to learn to say “I don’t know”.



Peace
Esteem
Happiness



Yearning
Disquietment
Annoyance



Current state of the art



Majority of models focus on only facial expressions ignoring contextual details.



Recent two models EMOTIC and CAER always considers contextual details and are non adaptive.



All the models follows frequentist approach and weights are fixed.



Do not take in to account the uncertainty present in the model.

Approach



Formulated problem as multi-class classification by including 7 basic emotion categories.

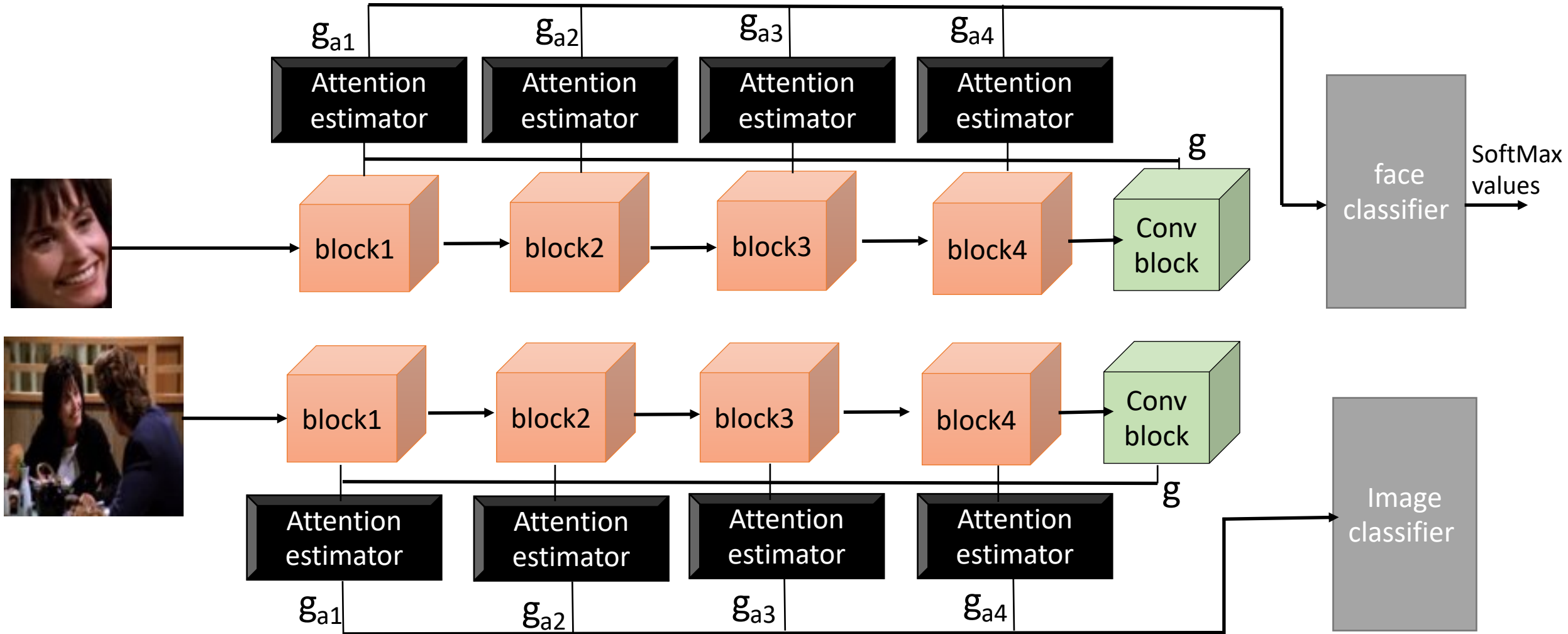


Made model as adaptive to consider when to take contextual details into account.



Used bayesian approach to find uncertainty in the model and also to make it more adaptive at granular level.

Frequentist Approach

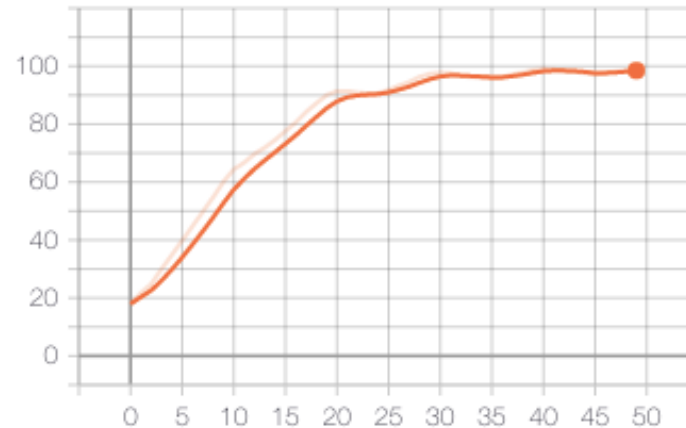


If SoftMax values \geq threshold then only go for Image classifier

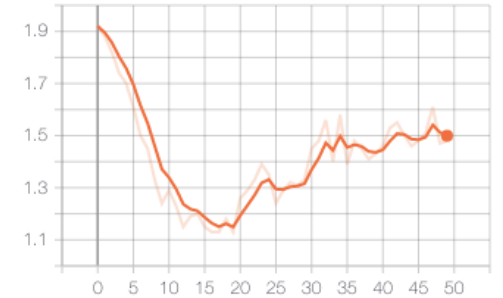
Experiment & Observation

- Hyperparameters used
- Stochastic gradient descent
- no_of_epochs = 50
- batch_size = 64
- weight_decay = 1e-3
- alpha = 0.6 & beta=0.4
- CyclicLR scheduler with step_size = 3065, base_lr=0.07 and max_lr = 0.09

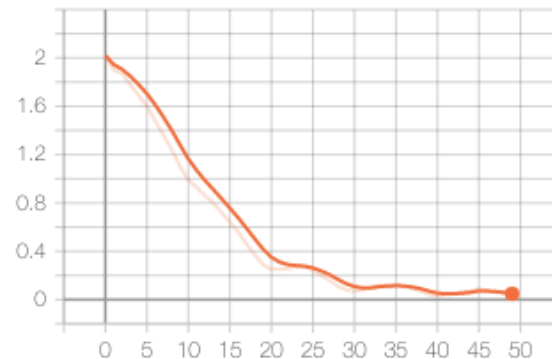
Accuracy
tag: Train/Accuracy



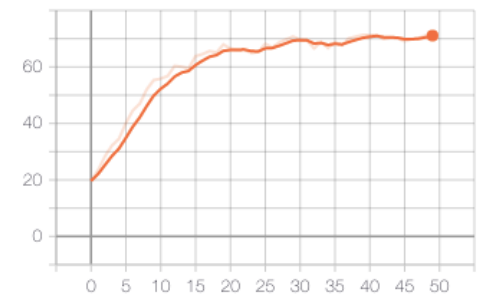
Loss
tag: Val/Loss



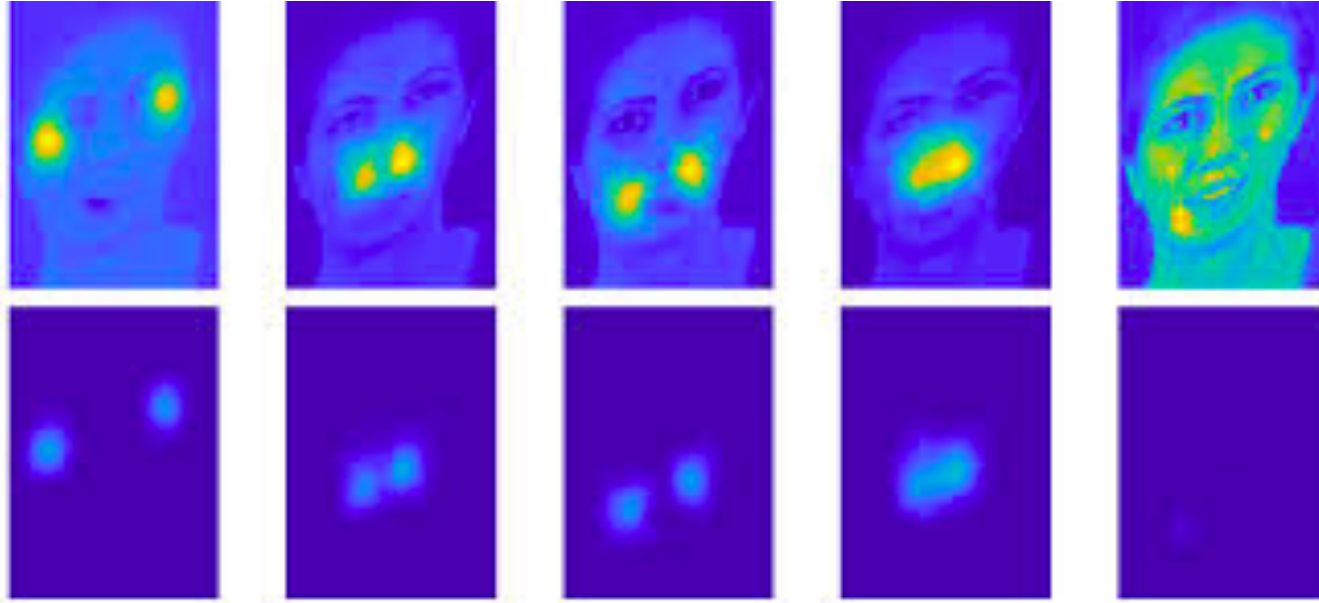
Loss
tag: Train/Loss



Accuracy
tag: Val/Accuracy



Observations



Category	Accuracy (%)
Anger	78.02 %
Disgust	63.12 %
Fear	60.53 %
Happy	75.90 %
Neutral	53.89 %
Sad	67.76 %
Surprise	70.54 %



Model	Accuracy
Context Aware Emotion Recognition (CAER)	73.51 %
Deep Emotion Recognition (DER)	71.52%

Downside of Frequentist Approach

01

Model is prone to overfit and fails to generalize when available data is less.

02

Follows deterministic approach where weights are fixed after training on certain datasets hence may give wrong confidence score on unseen data.

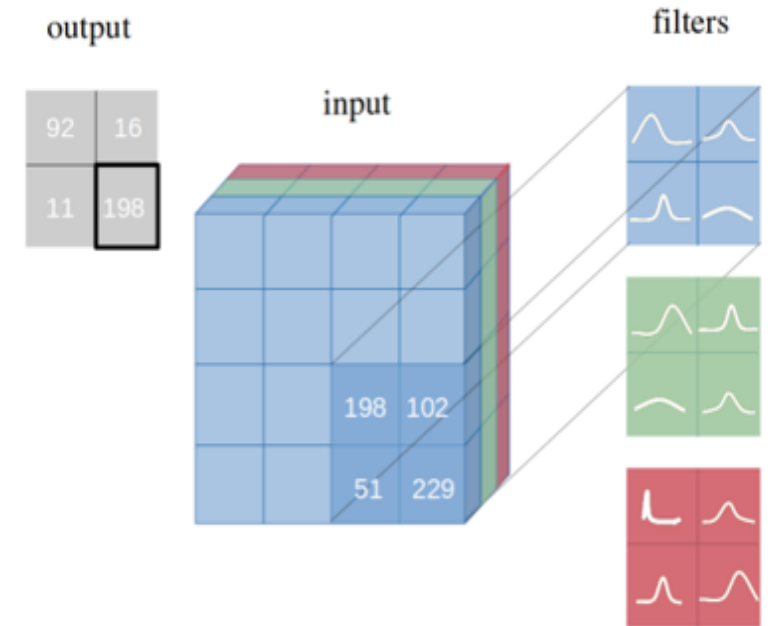
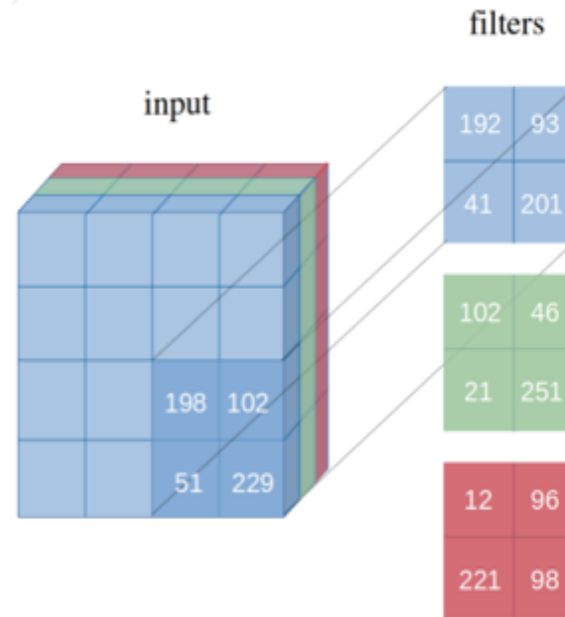
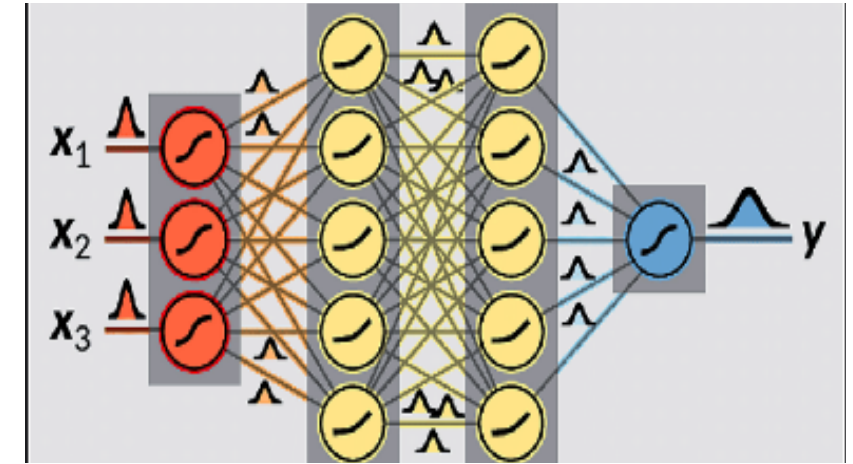
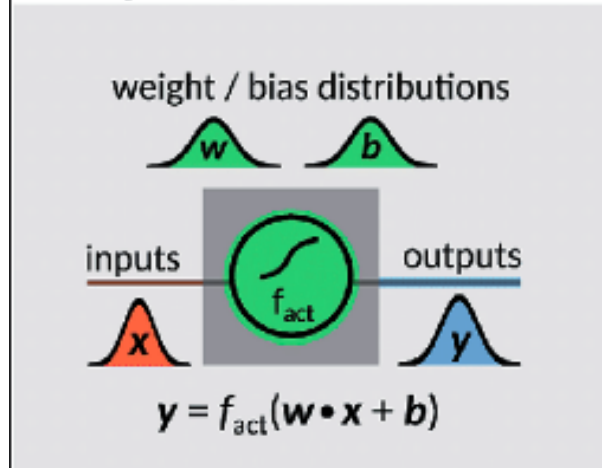
03

Unable to show uncertainty present in the model for unseen data.



Bayesian Approach

A) Single Bayesian neuron



Frequentist
v/s Bayesian
Network

Given dataset $D = \{(x_n, y_n)\}_{n=1}^N$

- Neural Network is trained using Maximum likelihood principle (stochastic gradient descent)

$$w^{MLE} = \operatorname{argmax} \sum_i P(y_i | x_i, w)$$

L2 Regularisation added as by considering prior on weights.

$$w^{MAP} = \operatorname{argmax} \sum_i P(y_i | x_i, w) + \log P(w)$$

With Bayesian Approach

$$P(w | D) = \frac{P(D | w)P(w)}{P(D)} = \frac{P(D | w)P(w)}{\int P(D | w)P(w)dw}$$

$P(w | D)$ is the posterior distribution of weights given data.



Inference

Bayesian inference

\hat{x} is the test data and \hat{y} is the class label.

$$P(\hat{y} | \hat{x}) = E_{P(w|D)}[P(\hat{y} | \hat{x}, w)]$$

Variational Inference

Variational approximation to posterior distribution on weights.

$q(w | \theta)$ is *variational posterior distribution*

θ is parameters as μ (*mean*) and σ (*varaince*)

Minimize Kullback-Leibler divergence between true Bayesian posterior and variational.

$$\theta^* = \operatorname{argmin} KL[q(w | \theta) || P(w)] - E_{q(w|\theta)}[\log P(D|w)]$$

Cost function

$$F(D, \theta) = \sum_{i=1}^n \log q_{\theta}(w^i | D) - \log p(w^i) - \log p(D | w^i)$$



Experiment & Observation

- For prior, used Gaussian distribution as scale mixture model.
- μ as 0 and σ_1 and σ_2 as -1 and -6
- For variational posterior, used Gaussian distribution with
- μ as -1 and σ_1 as -4
- Adam optimizer with learning rate as 0.07



Conclusion & Future Work

Contextual details helps in getting more details about the emotional state of person.

Working on Bayesian model to improve its accuracy and making model to learn uncertainty

Working on thesis simultaneously.

Plan to publish paper.

References

- [1] A. Dhall, J. Joshi, I. Radwan, and R. Goecke. Finding happiest moments in a social context. In Asian Conference on Computer Vision, pages 613–626. Springer, 2012.
- [2] A. Dhall, R. Goecke, S. Lucey, and T. Gedeon. Acted facial expressions in the wild database. Australian National University, Canberra, Australia, Technical Report TR-CS-11, 2, 2011.
- [3] R. Kosti et al, "Context Based Emotion Recognition using EMOTIC Dataset," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1-1, 2019.
- [4] R. Kosti et al, "EMOTIC: Emotions in context dataset," in 2017, . DOI: 10.1109/CVPRW.2017.285.
- [5] C. Blundell *et al*, "Weight uncertainty in neural networks," in 2015, .
- [6] J. Lee, S. Kim, S. Kim, J. Park and K. Sohn, "Context-Aware Emotion Recognition Networks," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South), 2019, pp. 10142-10151.

Multi-modality User Interface Search

Xi Chen

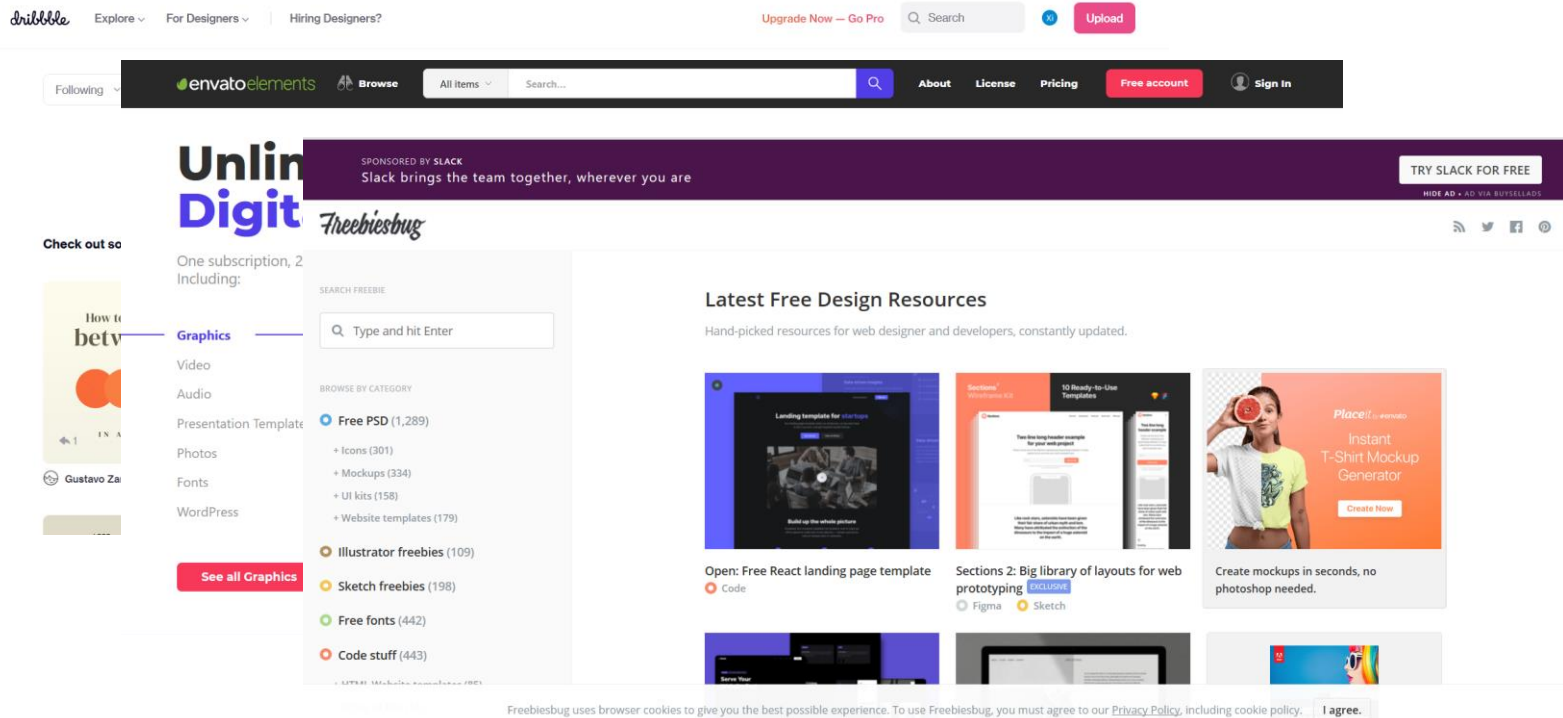
Supervisor: Zhenchang Xing

Outline

- Motivation
- Existing works
- Main tasks
- Data
- Methodology
- Current result
- Future work

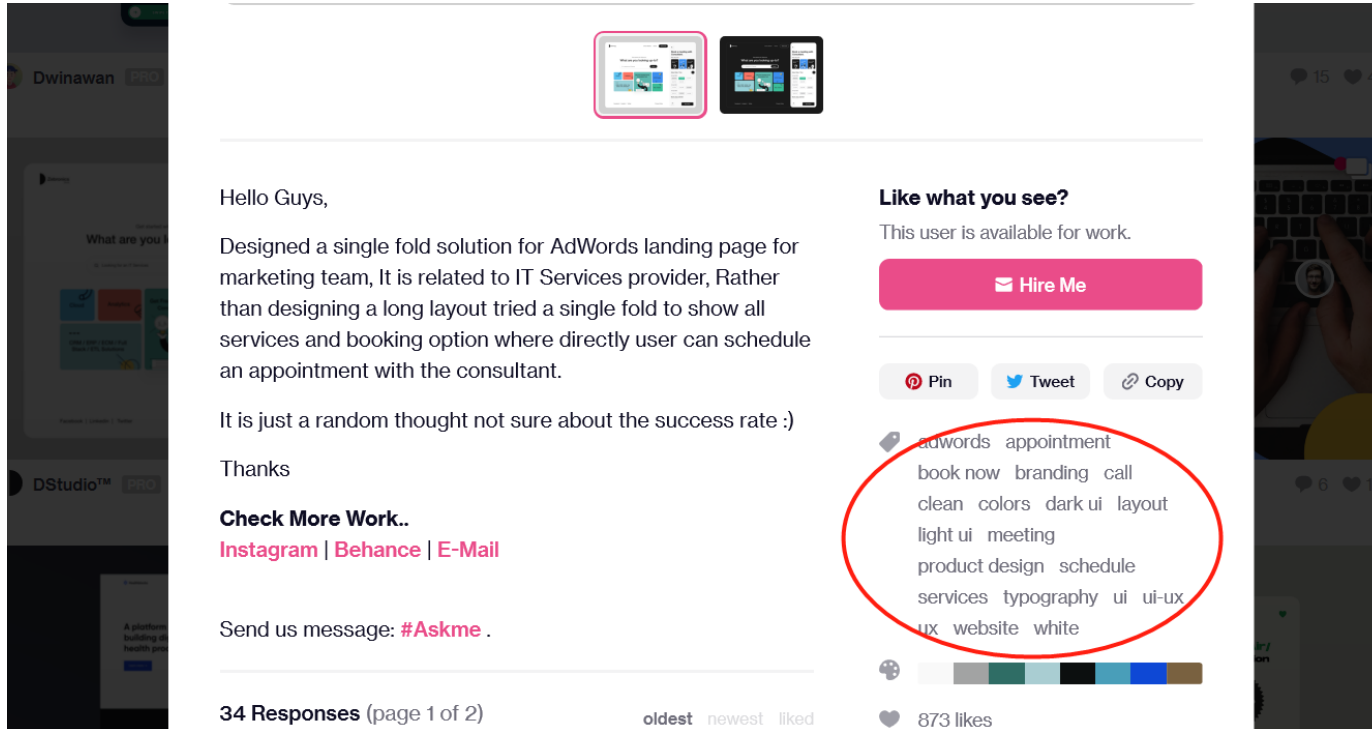
Motivation

- Well-design UI makes an application easy to use.
- A large number of UI templates are available online.



Tags in Dribbble

- Tagging designs requires manual efforts and could be subjective which may easily lead to errors and affect search results.



Dwinawan 1,234

Designed a single fold solution for AdWords landing page for marketing team, It is related to IT Services provider, Rather than designing a long layout tried a single fold to show all services and booking option where directly user can schedule an appointment with the consultant.

It is just a random thought not sure about the success rate :)

Thanks

Check More Work..
[Instagram](#) | [Behance](#) | [E-Mail](#)

Send us message: [#Askme](#).

Like what you see?
This user is available for work.
[Hire Me](#)

[Pin](#) [Tweet](#) [Copy](#)

adwords appointment
book now branding call
clean colors dark ui layout
light ui meeting
product design schedule
services typography ui ui-ux
ux website white

34 Responses (page 1 of 2) [oldest](#) [newest](#) [liked](#)

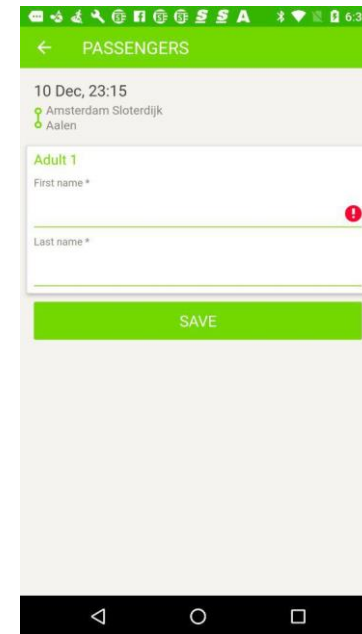
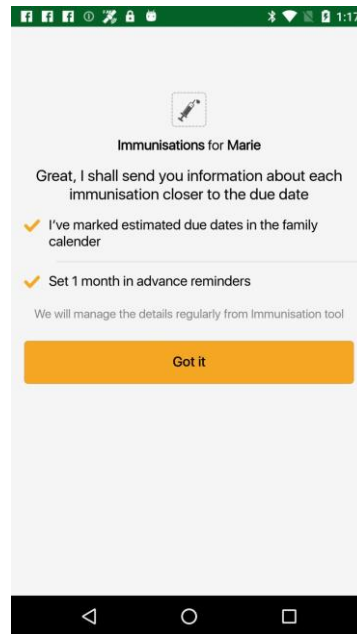
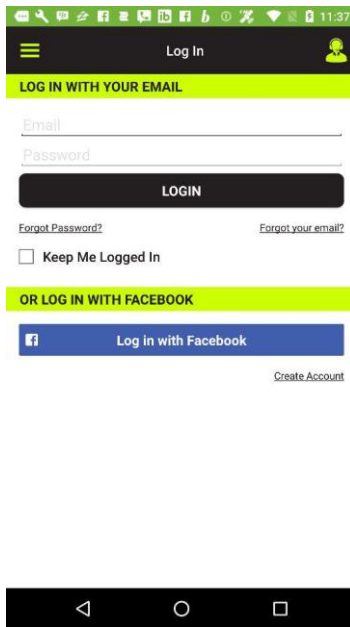
873 likes

Existing Work

- Guigle
 - Key word based search engine
 - automatically tagging UI with
 - text displayed on a screen
 - user interface components
 - app name
 - screen colors
 - Query Search
 - Requiring information for view hierarchies

Existing Work

- Swire
 - Sketch based search engine
 - Convolutional sub-networks

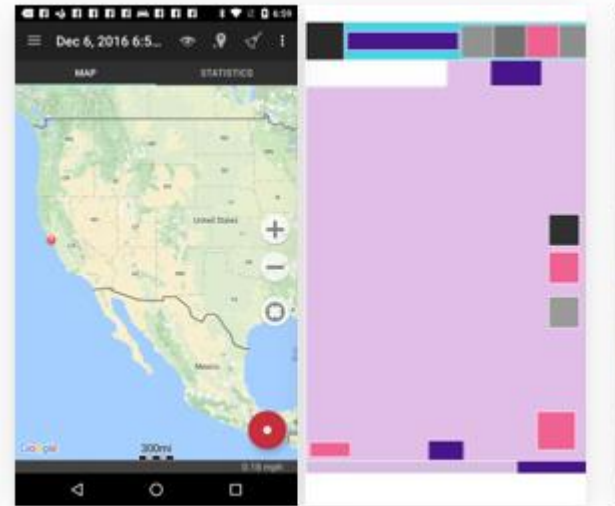
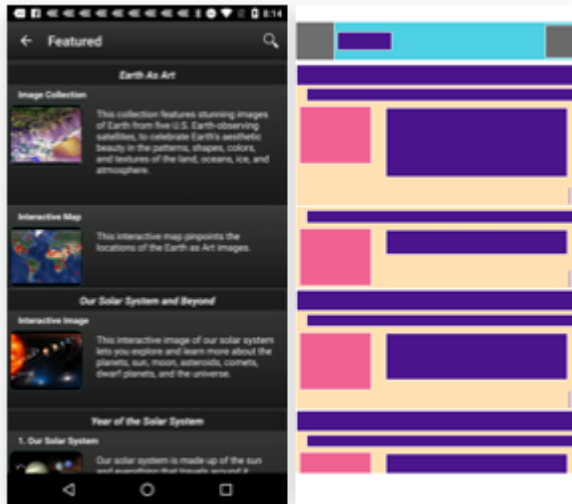


Main tasks

- Combine the relevance between high-fidelity UI screenshot and low-fidelity UI wireframe and relevance between UI screenshot and UI text
 - UI-sketch relevance
 - Convolutional encoder
 - UI-text relevance
 - Adversarial cross-modal encoder
 - Combination
 - weighted sum of both

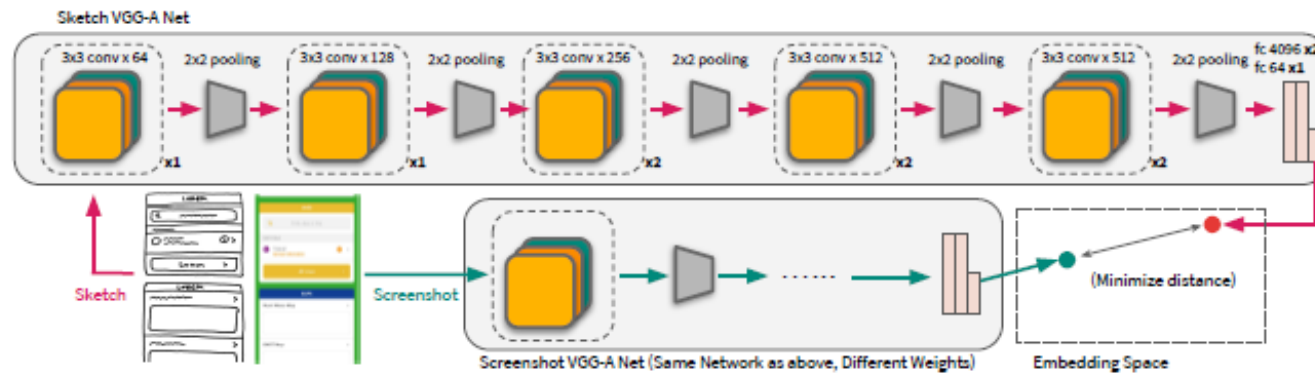
Data

- Rico – a mobile app dataset for building data-driven design applications
- 66k+ UI Screenshots and corresponding sketches



Retrieved from <http://interactionmining.org/rico#quick-downloads>

Methodology - UI-sketch relevance



Retrieved from <https://dl.acm.org/doi/fullHtml/10.1145/3290605.3300334>

Loss function:

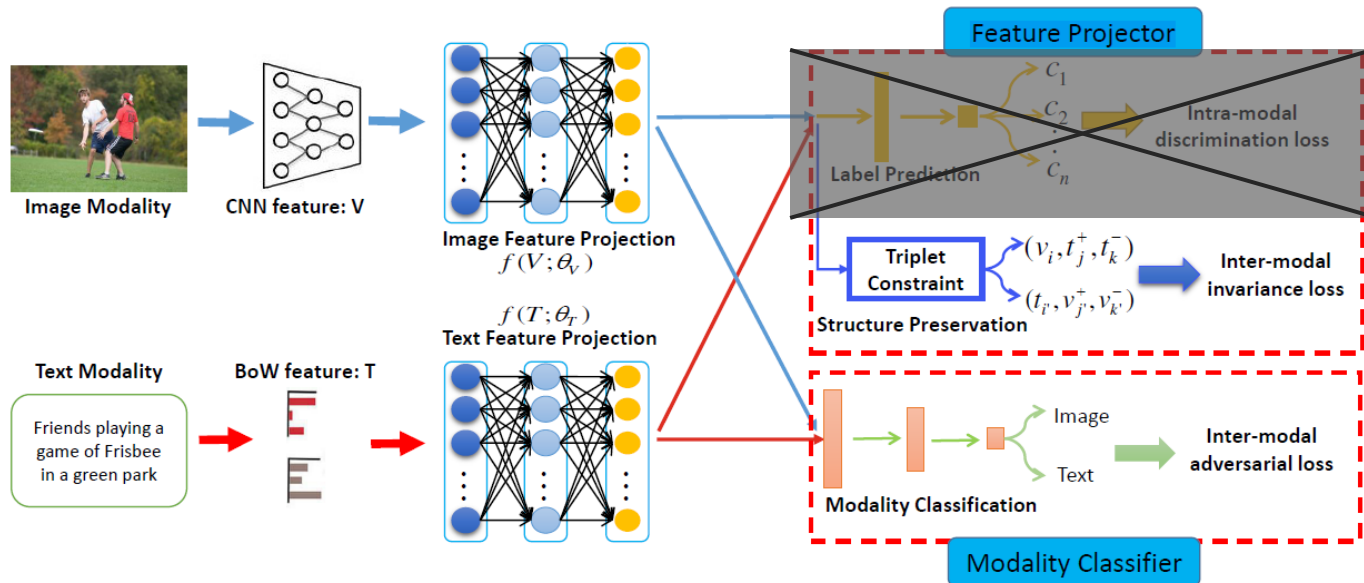
$$L = l_2(s, u^+) + \max(0, k - l_2(s, u^-))$$

$$l_2(s, u^+) = ||f_s(s) - f_u(u^+)||_2$$

$$l_2(s, u^-) = ||f_s(s) - f_u(u^-)||_2$$

s – sketch, u^+ - positive matched UI, u^- - negative unmatched UI

Methodology - UI-text relevance



Retrieved from https://blog.csdn.net/qq_33373858/article/details/81837084

Loss function:

$$L = L_{fp} - \alpha L_{mc}$$

L_{fp} = sum of the losses for two triplet constraints

L_{mc} = cross entropy loss

Current result

	Top-1 accuracy	Top-5 accuracy	Top-10 accuracy
Sketch-based	0.39	0.63	0.71
Sketch&text-based	0.44	0.66	0.71

Top-k accuracy: proportion that the right answer appears in the top k results

Future work

- Higher performance
- More baseline methods
- Better evaluation metrics
- (Generalisation)

Visual Question Answering (VQA)

Xuwei Xu

Bachelor of Advanced Computing (Honours)

Supervisor: Dr. Lars Petersson

Outline

Background

- Problem Motivation
- Problem Definition

Current Methods

- Attention Mechanism
- Fusion
- MuRel

Exploration

- Answer related
- Aggregated network

Challenge

Future Work

Motivation

Computer Vision

- Recognition
- Detection
- Classification
- Reconstruction
- Segmentation
- ...

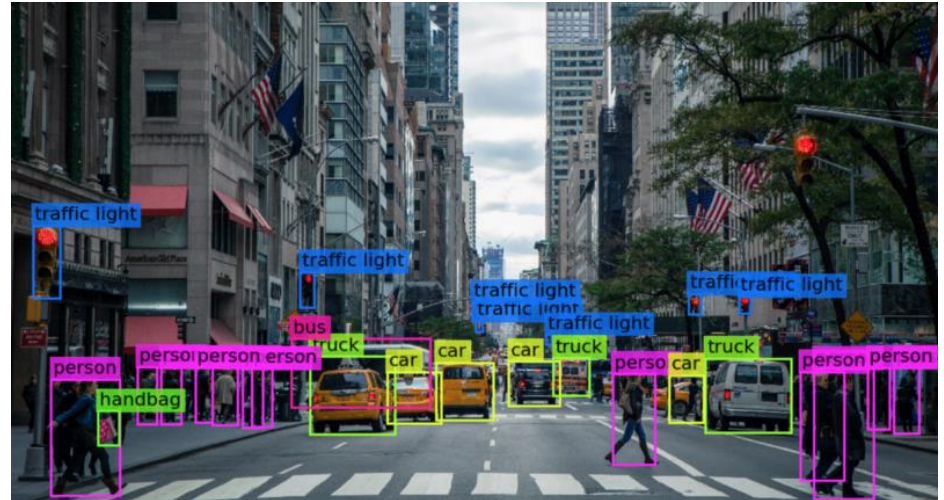


Image detection

Natural Language Processing

- Natural language understanding
- Natural language generation
- ...

Definition of VQA



- What are they doing?
- How many competitors?
- Is it raining?
- What is the color of their shirts?
-

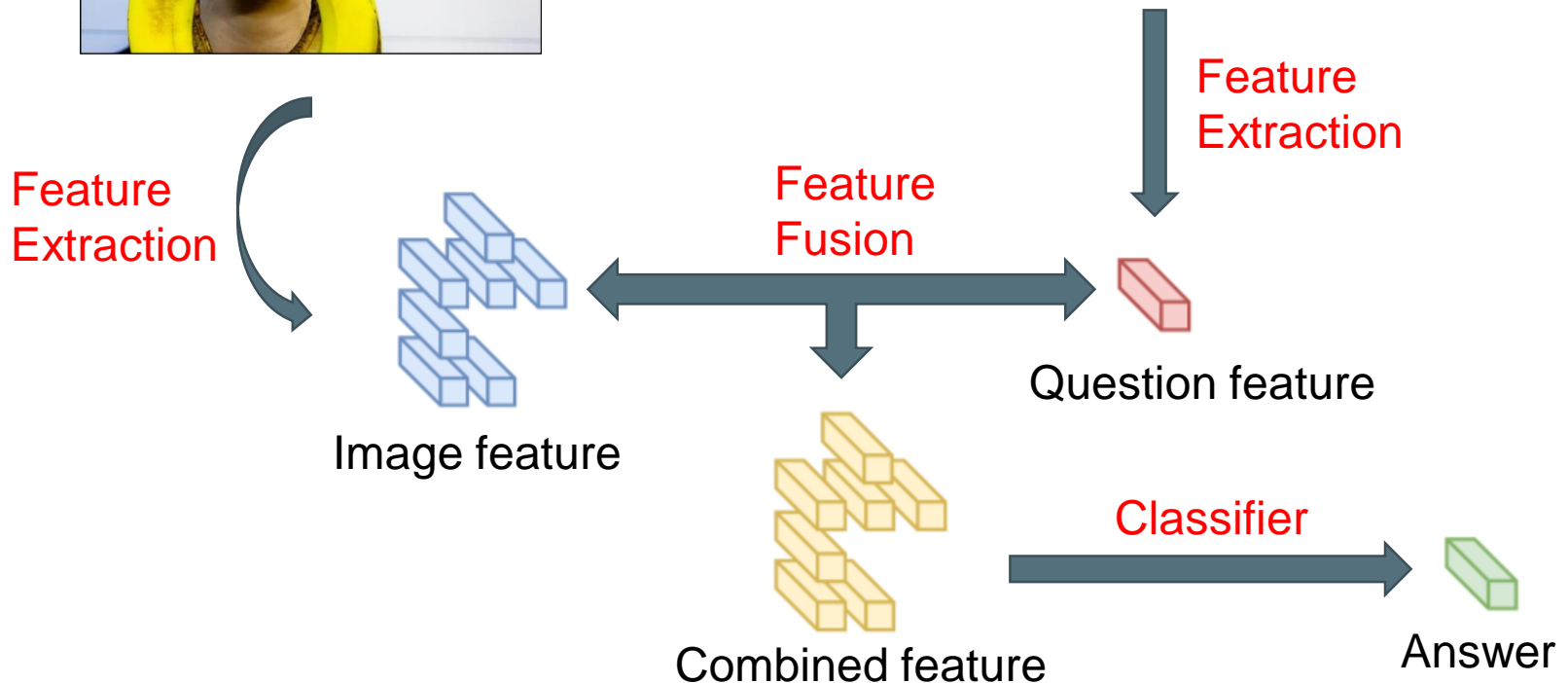
VQA

- Input: Image v + Textual question q
- Output: An textual answer a to the question
- Building a system/algorithm that takes (ideally) any image and a question asked in natural language about that image and provides a natural language answer to that question with reference to the image as the output. $VQA(v, q) = a$
- Requires:
An understanding of vision, language and commonsense knowledge to answer.

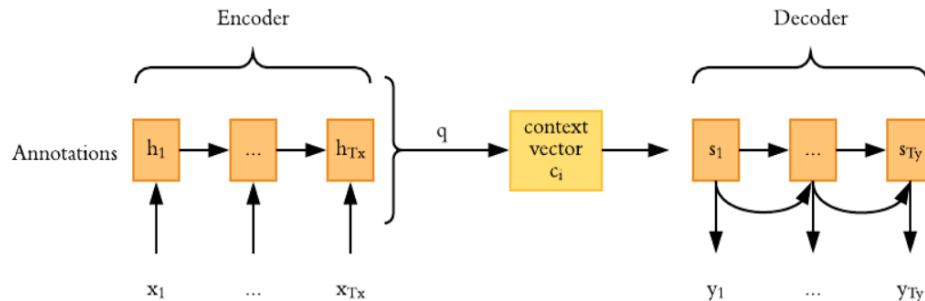
Current Methods



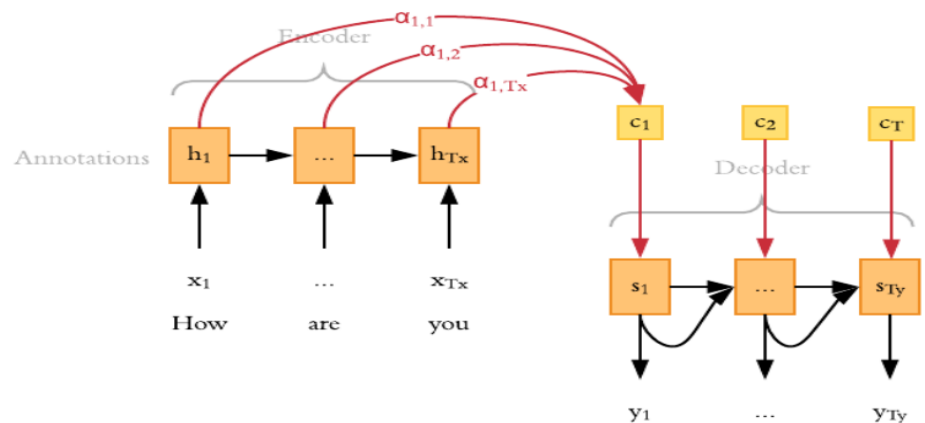
What is the mustache made of?



Attention Mechanism



Encoder-decoder

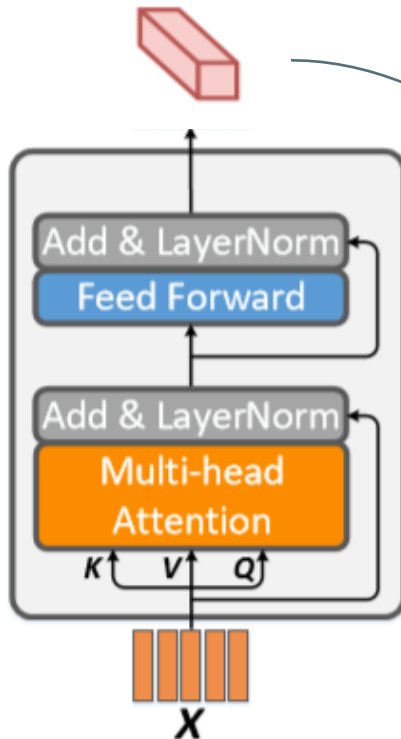


Encoder-decoder with attention

- Designed for language processing
- Weighted annotations
- Emphasize important information
- We can use image features to represent annotations and use question features to represent weights

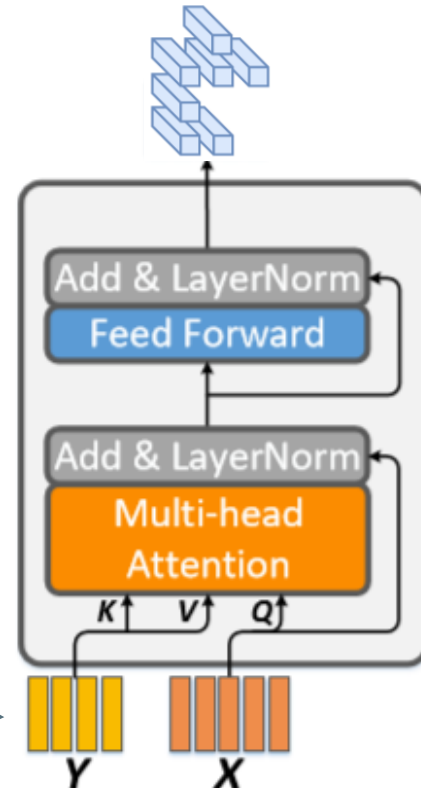
Attention Mechanism

Question attention



$$\text{Attention}(q, K, V) = \text{softmax}(qK^T)V$$

Image attention



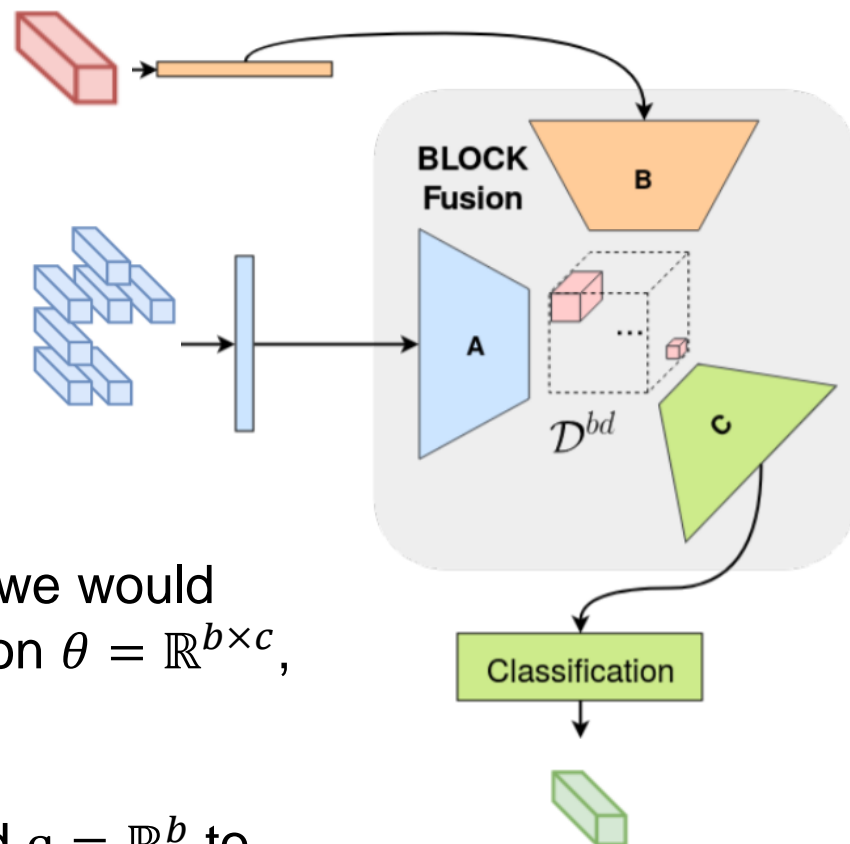
$$\text{Attention}(q, v) = \text{softmax}(vq^T)q$$

Feature Fusion

- Fully Connection
- Recurrent Spatial Attention
- Stacked Attention
- Deep Co-Attention
-
- **Bilinear Attention Fusion!**

If we want to map $x = \mathbb{R}^{a \times b}$ to $y = \mathbb{R}^{a \times c}$, we would introduce an internal tensor with dimension $\theta = \mathbb{R}^{b \times c}$, hence $y = x\theta$.

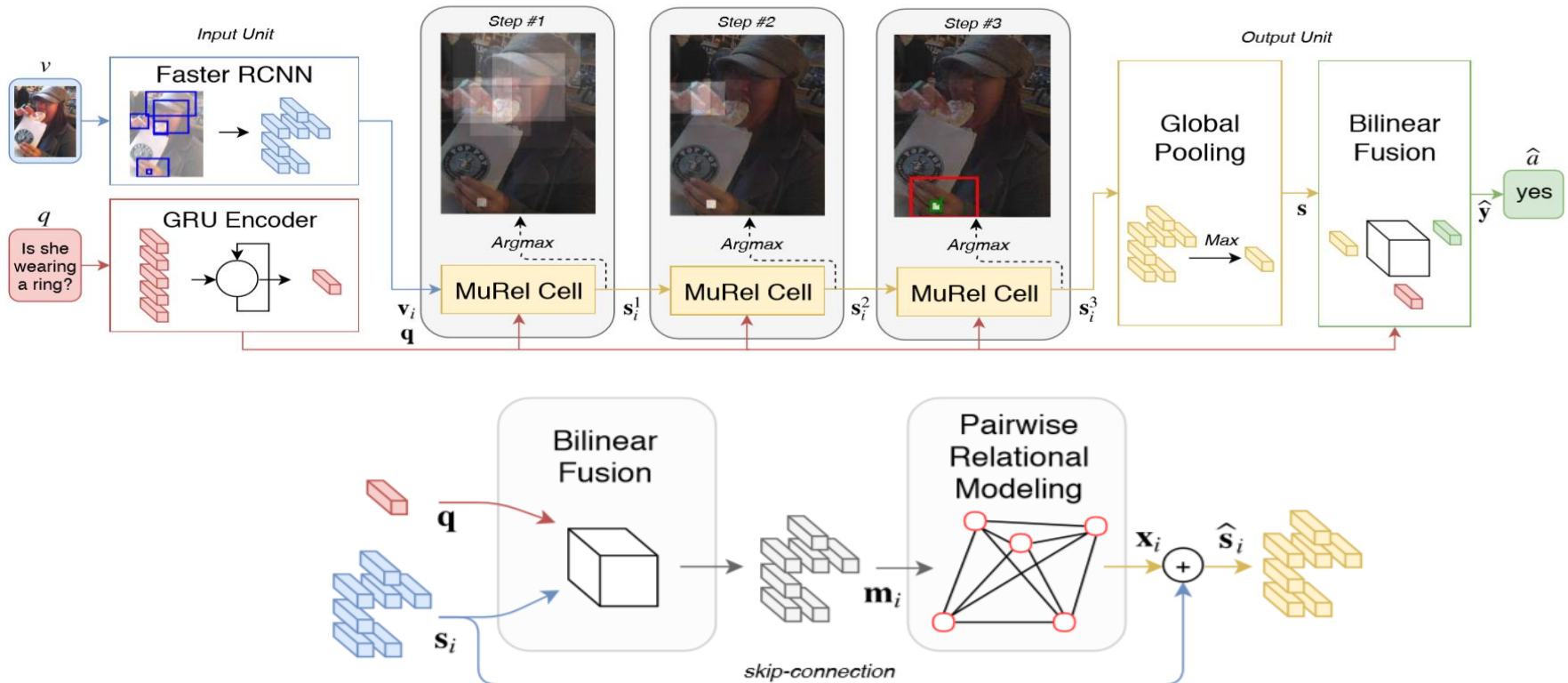
Similarly, if we want to map $v = \mathbb{R}^{n \times a}$ and $q = \mathbb{R}^b$ to $a = \mathbb{R}^{n \times c}$, we can introduce a tensor with dimension $T = \mathbb{R}^{a \times b \times c}$. Hence, $a = (T \times_1 v) \times_2 q$



MuRel

Multimodal Relational Reasoning (MuRel)

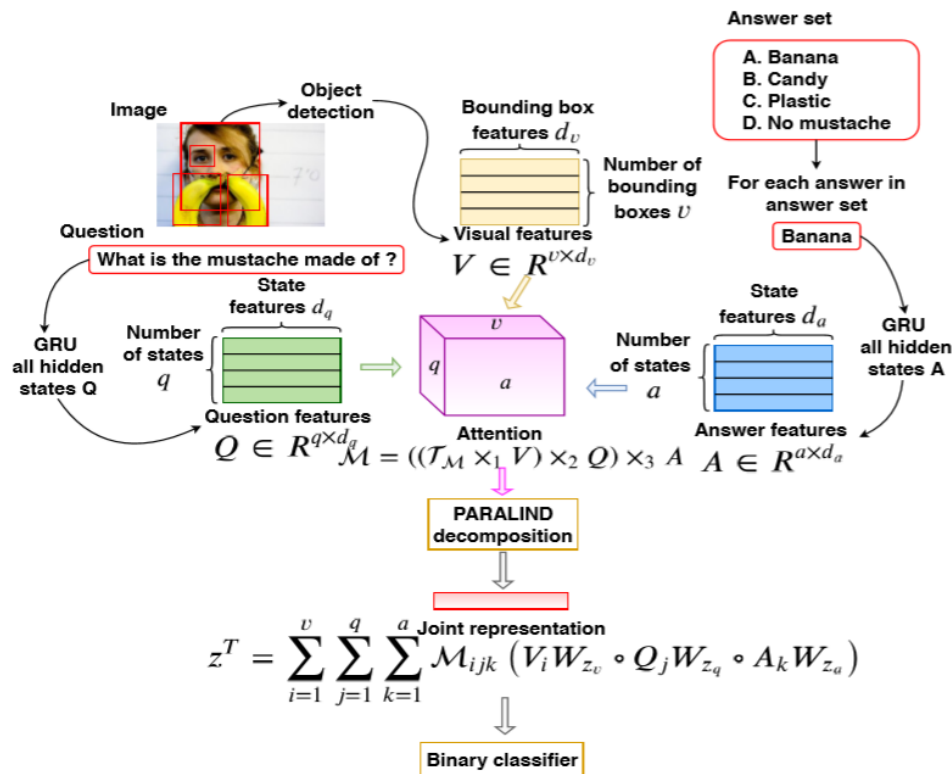
- State-of-the-art
- Combine attention mechanism and bilinear fusion
- Introduce spatial relationship



My Approach

1. Introduce answers into the network

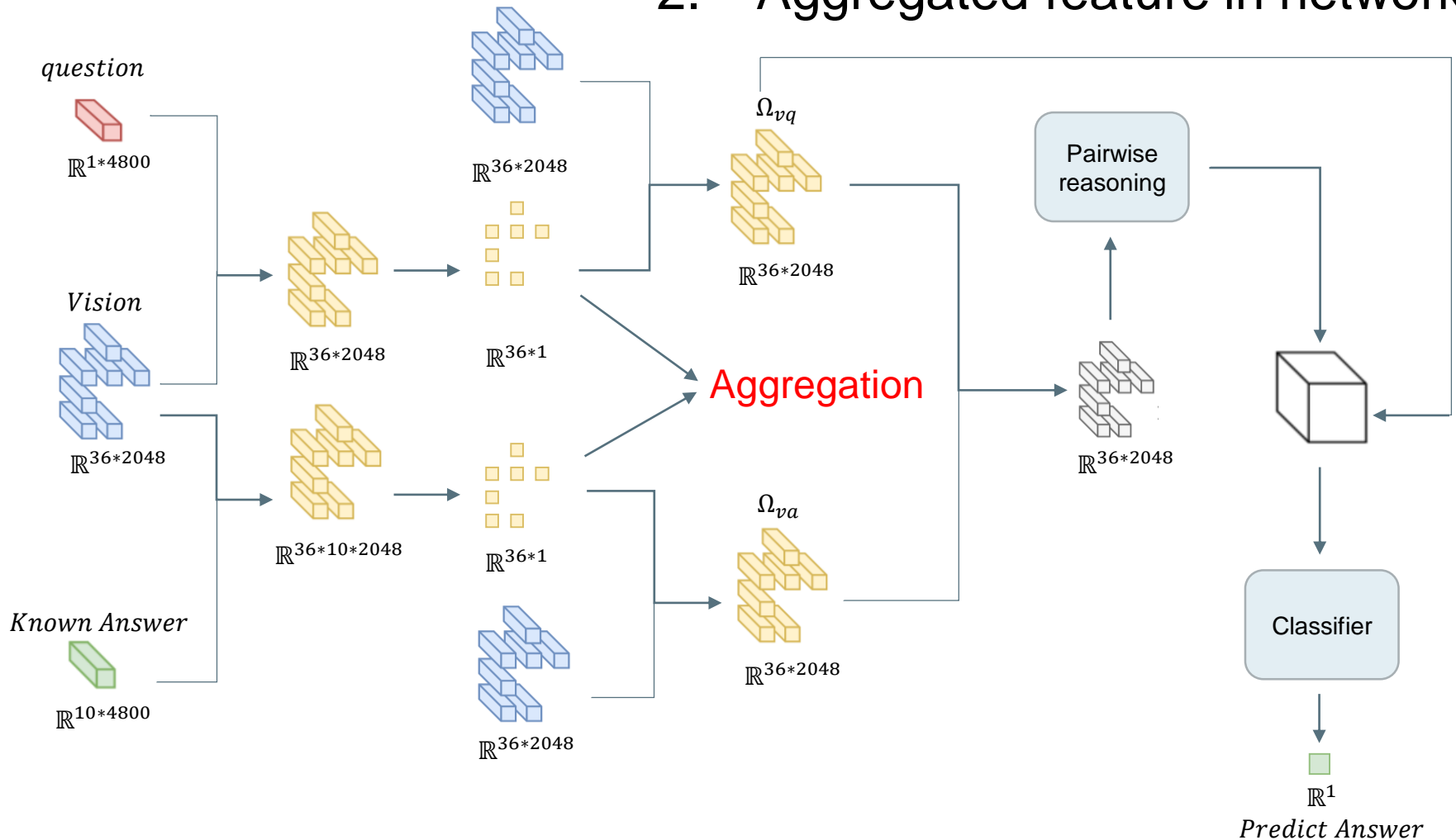
Inspired by Compact Trilinear Interaction for Visual Question Answering



- Adding answer features to the network in training
- Decompose the trilinear fusion into several bilinear fusions to reduce the computation
- Combine the answer and the question

My Approach

2. Aggregated feature in network



Challenge

Open-ended question

- Answers are not in the corpus
- Answers need extra inferences

Computation resource

- Computing bilinear or trilinear fusion consumes huge computation resource
- Could be very slow if compute from raw data
- Impossible to use on mobile device at present

Future Work

More features in the network

- Use spatial relationship as one feature in trilinear or bilinear fusion
- Use image feature to extract question attentions

General dataset

- Test performance on more general dataset, e.g. TDIUC
- Test performance on open-ended question dataset, e.g. OKVQA

Novel structure

- Design new structures to combine attentions
- Design less complex structure so that VQA system can be used on mobile or micro device in daily life.

Answer generation

- Automatically generate answer from online information

Reference

- <https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e>
- <https://medium.com/@joealato/attention-in-nlp-734c6fa9d983>
- <https://zhuanlan.zhihu.com/p/43493999>
- H. Ben-Younes, R. Cadene, N. Thome, and M. Cord. Block: Bilinear superdiagonal fusion for visual question answering and visual relationship detection. In Proceedings of the 33rd Conference on Artificial Intelligence (AAAI), 2019.
- R. Cadene, H. Ben-Younes, N. Thome, and M. Cord. MUREL: Multimodal Relational Reasoning for Visual Question Answering. Computer Vision and Pattern Recognition (CVPR), 2019.
- Z. Yu, J. Yu, J. Fan, and D. Tao. Multi-modal factorized bilinear pooling with co-attention learning for visual question answering. IEEE International Conference on Computer Vision (ICCV), pages 1839–1848, 2017.
- T. Do, T. Do, H. Tran, E. Tjiputra and Q. D. Tran. Compact Trilinear Interaction for Visual Question Answering, Computer Vision and Pattern Recognition (CVPR), 2019.