

# TAN WANG

UESTC, No.2006 Xiyuan Avenue ◇ Chengdu, Sichuan, P.R.China

(+86) 186-5602-3625 ◇ wangt97@hotmail.com ◇ Github: [Wangt-CN](#) ◇ Website: [Tan Wang](#)

## RESEARCH INTERESTS

---

**Computer Vision, Visual Reasoning, Cross-modal Retrieval**

## EDUCATION

---

**University of Electronic Science and Technology of China (UESTC)** *Sep. 2016 - Present*

Third-year undergraduate student in Electronic Information Engineering

**GPA**(Overall): **93.57**/100 or **4.0**/4.0

**Rank:** 1/284 (2018-2019) or 1/415 (first 2 years)

**Highlighted Courses:** Calculus I (96/100), Calculus II (97/100), Pattern Recognition (99/100), Probability and Mathematical Statistics (96/100), Signals and Systems (95/100), Digital Signal Processing (96/100), Analog electronic Circuits (96/100), Electromagnetic Field Theory (97/100)

**Chiba University, Japan**

*Aug. 2017*

Exchange Program

Sakura Science Club scholarship awardee. Funded by Japan Science and Technology Agency (JST).

## PUBLICATION & MANUSCRIPTS

---

**Matching Images and Text with Multi-modal Tensor Fusion and Re-ranking**

- **Tan Wang**, Xing Xu, Yang Yang, Alan Hanjalic, Heng Tao Shen and Jingkuan Song

- ACM MM2019

**Cross-Modal Attention with Semantic Consistence for Image-Text Matching**

- Xing Xu\*, **Tan Wang\***, Yang Yang, Lin Zuo, Fumin Shen and Heng Tao Shen

- *Major revision*, IEEE Transactions on Neural Networks and Learning Systems (TNNLS)

**Radial Graph Convolutional Network for Visual Question Generation**

- **Tan Wang**, Xing Xu, Yang Yang, Alan Hanjalic and Heng Tao Shen

- *Major revision*, IEEE Transactions on Neural Networks and Learning Systems (TNNLS)

## RESEARCH EXPERIENCE

---

**Nanyang Technological University (NTU)**

*Jul. 2019 - Present*

*Research Assistant, MReal Lab*

*Supervised by Prof. Zhang Hanwang.*

· *Visual Reasoning and Commonsense*

- Motivation: Recently the research area focuses from recognition to cognition. The object detection and recognition is such reliable for us to move on visual and machine reasoning. We focus on how to create more smart and exciting model based on the existed recognition technology.

**University of Electronic Science and Technology of China (UESTC)** *Mar. 2018 - Present*

*Research Assistant, Center for Future Media*

*Supervised by Prof. Xing Xu and Prof. Yang Yang. Collaborated with Prof. Alan Hanjalic.*

· *Image-text Matching with Semantic Consistence*

- Motivation: Recently the fine-grained methods for image-text matching that explore the local alignment between the image regions and the sentence words have shown advance on inferring the image-text correspondence. However, it is hard to achieve since some important image regions may be inaccurately detected or even missing. Meanwhile, some words with high-level semantics cannot be strictly corresponding to a single image region.
- We propose a novel hybrid matching approach named Cross-modal Attention with Semantic Consistence (CASC) for image-text matching, which is a joint framework that performs cross-modal attention for local alignment and multi-label prediction for global semantic consistence.
- Extensive experiments on Flickr30k and MSCOCO datasets demonstrate the effectiveness of the proposed CASC and further show its superior performance compared with 13 state-of-the-art methods.

#### · *An Efficient Image-text Matching Framework*

- Motivation: The existing two kinds of methods for image-text matching cannot well balance the matching accuracy and model complexity.
- We propose a novel framework for image text matching that achieves remarkable matching performance with acceptable model complexity and much less time consuming.
- Extensive experiments on two benchmark datasets demonstrate that our MTFN-RR consistently achieves the state-of-the-art matching performance with much less time complexity.

#### · *Visual Question Generation Task*

- Motivation: The existing approaches typically treat the VQG task as a reversed Visual Question Answer (VQA) task, requiring the exhaustive match among all the image regions and the given answer with huge time consuming.
- We propose an innovative answer-centric approach for visual question generation termed Radial Graph Convolutional Network (Radial-GCN) to focus on the relevant image regions only.
- Extensive experiments on three benchmark datasets show the effectiveness of our approach compared to the reference methods. Even on the unexplored challenging zeroshot VQA task, the synthesized questions by our method remarkably boost the performance of several best VQA methods from 0% to over 40%.

## SCHOLARSHIPS

---

<b>National Scholarship</b> (Top 2% student)	2017-2018
<b>Tang Lixin Sponsored Elite Scholarship</b> (Only 60 awardees per year in UESTC)	2017-2018
<b>Honor Student Scholarship</b> (10 awardees per year in Department)	2017-2018
<b>Outstanding Student Scholarship</b> (Top 10% student)	2017-2018
<b>National Scholarship</b> (Top 2% student)	2016-2017
<b>Best Freshman Award</b> (Only 1 awardee per year in Department)	2016-2017
<b>Outstanding Student Scholarship</b> (Top 10% student)	2016-2017

## LEADERSHIP

---

### Lecture Group of EE Department

*Founder & President*

*Oct. 2017 - Sep. 2018*

- Organized academic forum, sharing sessions, Q&A meetings more than 30 times, serving over 1000 students on studying and future planing.
- The team grows to 30 people and won the Outstanding Student Organisation prize in 2018.

### Innovative Entrepreneurship Project of UESTC

*Team Leader*

*Sep. 2017 - Mar. 2018*

- This project focus on the pedestrian detection in low-light condition with excellent conclusion. We combine the recent pedestrian detection models with the low-light image enhancement algorithm based on Laplace operator.
- Responsible for the code implementation and project promotion.

## REVIEWS

---

Reviewer for ICME 2019

Help to review for AAAI 2019, KDD 2019, Access

## SKILLS

---

**Programming Languages**

Python, Matlab, L<sup>A</sup>T<sub>E</sub>X, Shell

**Deep Learning**

Pytorch, Tensorflow

**Others**

Multisim, VerilogHDL