

Tips of Writing a \LaTeX Paper

Using VS Code & Overleaf V2 as an Example

Ming-Ming Cheng

College of Computer Science, Nankai University

Course materials: <https://mmcheng.net/writing/>

2022/05/24

Table of Contents

- 1 Overleaf for collaboration
- 2 Git and TortoiseGit
- 3 Grammarly
- 4 VS Code + TexLive
- 5 Using \LaTeX
- 6 Figures
- 7 Checklist
- 8 Rebuttal
- 9 Project page

Using Overleaf V2 for Joint Working

See also the official tutorial

<https://www.overleaf.com/learn/latex/Tutorials>

Online collaborative \LaTeX editing

- Easy to work with many co-authors

Version control using git sync

- Using GitHub for Git based version control

Comments discussion

- Comments location, author name, time, and reply

Overleaf V2 Interface

The screenshot displays the Overleaf V2 interface for a project titled "EdgeRCF". The interface is divided into several sections:

- Top Bar:** Navigation icons (back, forward, refresh, home) and the project URL: <https://www.overleaf.com/project/5bd7bea0ef398e7552cefa33>. On the right, there are icons for star, share, submit, history, and chat.
- Left Panel:** A file explorer showing a directory structure with folders like "CVPR", "Imgs", and "Other", and files like "edge.bib", "IEEETra...", and "Sh...".
- Source Editor:** A LaTeX source code editor showing the following code:


```

41 \def\state-of-the-art~}
42
43 % correct bad hyphenation here
44 \hyphenation{net-works}
45
46
47
48 \begin{document}
49 %
50 \title{Richer Convolutional Features
51 for Edge Detection}
52
53 \author{Yun Liu, Ming-Ming Cheng,
54 Xiaowei Hu, Jia-Wang Bian, Le Zhang,
55 Xiang Bai, and Jinhui Tang \thanks{*}
56 M.M. Cheng is the corresponding author.
57 URL: http://mmcheng.net/rcfedge/}
58 \IEEEcompsoctemizethanks{
59 \IEEEcompsocthanksitem Y. Liu, M.M.
60 Cheng, and J.W. Bian, are with the
61 College of Computer Science, Nankai
62 University,
63 Tianjin 300350, China.
64 \IEEEcompsocthanksitem L. Zhang is with
      
```
- Right Panel:** A preview window showing the rendered PDF. It includes a "Track changes" notification for a title change on Nov 22, 2018. The preview shows the title "Richer Convolutional Features for Edge Detection" and the author list. Below the text, there are images of edge detection results on various input images, comparing the proposed method with other state-of-the-art methods.

Fig.: Website Interface. (Overleaf V2)

Avoid Meaningless Account Names

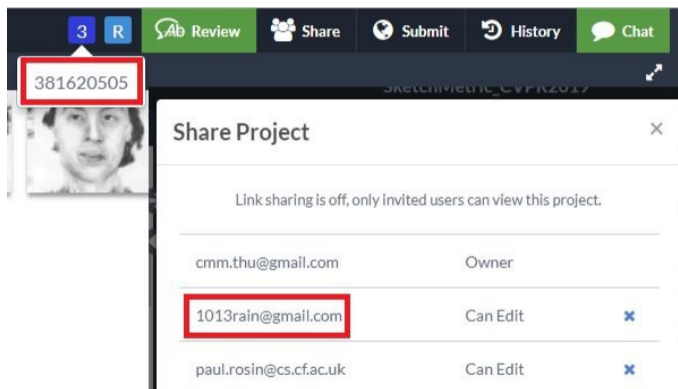


Fig.: Avoid such meaningless names during working with others, email communications, etc. Academic accounts have additional benefits/discount from commercial companies.

Direct Git access available in Overleaf V2.

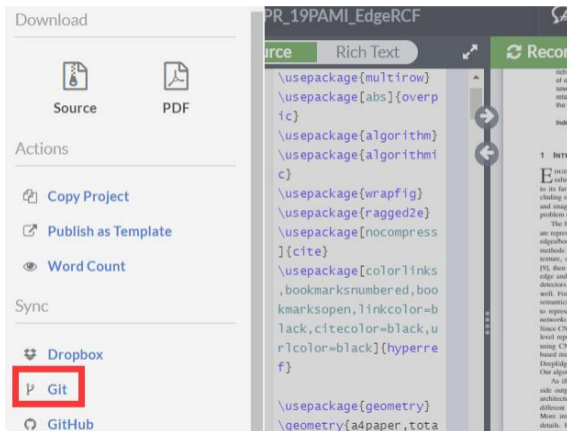
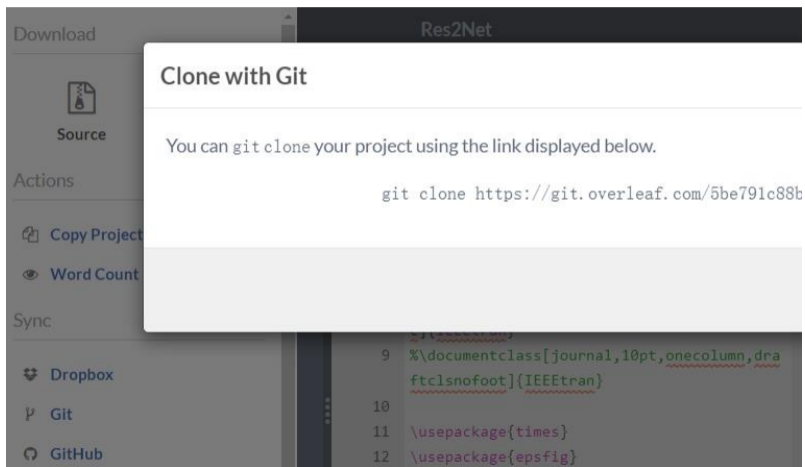


Fig.: Direct Git access is launched in Jan. 2019.

Overleaf with Git



Download

Res2Net

Clone with Git

You can git clone your project using the link displayed below.

```
git clone https://git.overleaf.com/5be791c88b
```

Source

Actions

Copy Project

Word Count

Sync

Dropbox

Git

GitHub

```
9 %\documentclass[journal,10pt,onecolumn,draft,clsnofoot]{IEEEtran}
10
11 \usepackage{times}
12 \usepackage{epsfig}
```

Fig.: Overleaf version control using Git.

TortoiseGit for Windows

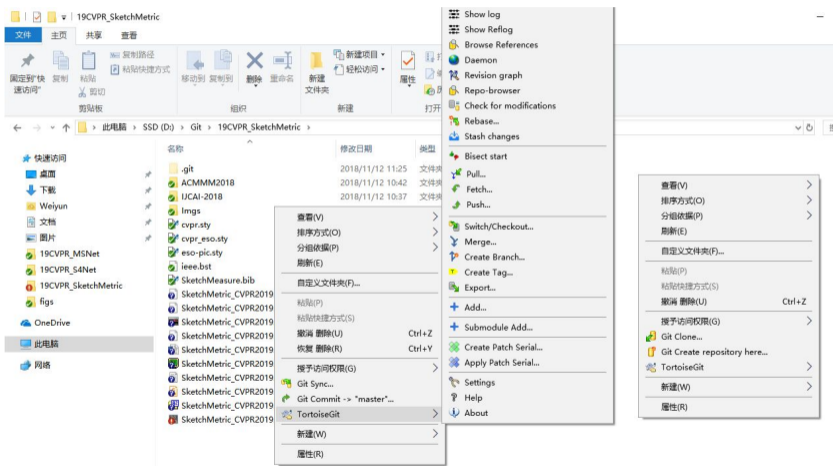


Fig.: Windows explore interface with TortoiseGit.

TortoiseGit for Windows

Different icons means different status

- Icons changes after sync using Git Clone.
- Red: on disk version is newer, commit ASAP.
- Green: on disk version is up to date.

Sync as ASAP to avoid version conflicts

- Default commit to **local repository**, not online repository.
- Following this order: **commit, pull, push.**

Break lines to avoid version conflict

- Avoid version conflict due to simultaneous editing
- Git software discover changes line-by-line
- Also makes the writing with pretty hierarchy

```
226 As one of the most fundamental problem in computer vision,  
227 edge detection has been extensively studied for several decades.  
228 %  
229 \textit{Early pioneering methods} mainly focus on the utilization of intensity  
230 and color gradients, such as Canny \cite{canny1986computational}.  
231 However, these early methods are usually not accurate enough  
232 for real-life applications.  
233 %  
234 To this end, \textit{feature learning based methods} have been proposed.  
235 These methods, such as Pb \cite{martin2004learning},  
236 gPb \cite{arbelaez2011contour}, and SE \cite{dollar2015fast},
```

Fig.: Break lines **ASAP**.

Pay Attention to Warnings in Overleaf

Warnings are often **very helpful**

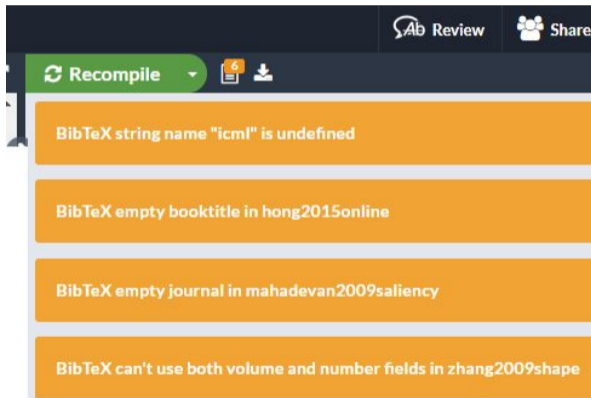


Fig.: Confirm every warning.

Pay Attention to Warnings in Overleaf

Warnings are often **very helpful**. Confirm every warning.

```
enlarge the receptive field
layers than using large kernels.

Net provides a stronger multi-scale representation model
eters.

exNet and VGGNet stack filters directly,
feature layer has a relatively fixed receptive field.
pology,
se networks can only have a relatively inflexible receptive
handling objects with a small range of scales.

Network (NIN)~\cite{lin2013network} inserts multi-layer perceptrons as
nto the large network
discriminability for local patches within the receptive field.
convolution introduced in NIN has been a popular module to fuse features.}
```

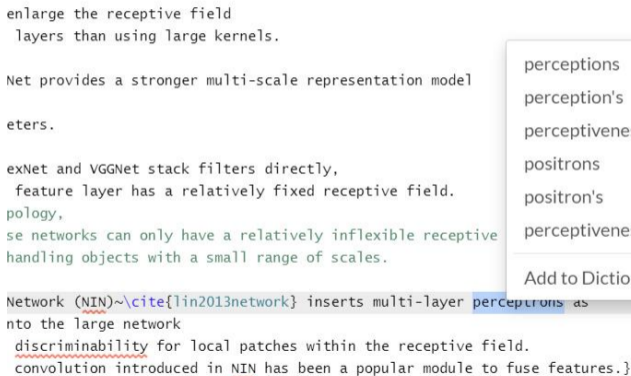
A screenshot of a LaTeX editor interface. The main text area contains several paragraphs of text with some words underlined in red, indicating warnings. The word 'perceptrons' is highlighted in blue. A context menu is open over the word, listing several suggestions: 'perceptions', 'perception's', 'perceptiveness', 'positrons', 'positron's', and 'perceptiveness's'. At the bottom of the menu is the option 'Add to Dictionary'.

Fig.: Explicitly disable miss-alarms and use "add to dictionary" if you are confident.

Grammarly

- Prepare: Trying to get a clean pdf: <https://www.lin-zheng.com/latex-pure/>

The weight in Eqa. (6) can be represented as follows:

$$\tilde{w}_p = \alpha + \psi(p, \{a_f\}) (\beta - \alpha) y_p, \quad (7)$$

where a_f means the first point in \mathcal{A}_p .

In experiments, we choose τ at 100, α at 0.8, β at 2.0.

3.3. Structural Integrity Strategy

Through experiments, we find that the prediction masks of neural networks may contain some scattered regions of wrong results. In most cases, people will prefer to get the object mask which maintains structural integrity in the task of interactive segmentation. Therefore, we propose a strategy to maintain the structural integrity of the segmentation based on interaction points.

Normally, we take 0.5 as the threshold to get the final binarized mask from the output of neural networks. Let \mathcal{P} represent these points which are predicted as foreground. We will postprocess these prediction areas according to the interaction points and get new \mathcal{P}' , which is formulated as follows: where $\sigma(p_1, p_2) = 1$ when there is an eight-connected path from point p_1 to point p_2 . The structural integrity strategy can work in most cases. The effect of it can be seen in table.

$$\mathcal{P}' = \{p \in \mathcal{P} \mid \exists_{a \in \mathcal{A}_p} \sigma(p, a) = 1\}, \quad (8)$$

where $\sigma(p_1, p_2) = 1$ when there is an eight-connected path from point p_1 to point p_2 . The structural integrity strategy can work in most cases. The effect of it can be seen in Tab. 2.

3.3. Structural Integrity Strategy

Through experiments, we find that the prediction masks of neural networks may contain some scattered regions of wrong results. In most cases, people will prefer to get the object mask which maintains structural integrity in the task of interactive segmentation. Therefore, we propose a strategy to maintain the structural integrity of the segmentation based on interaction points.

Normally, we take 0.5 as the threshold to get the final binarized mask from the output of neural networks. Let \mathcal{P} represent these points which are predicted as foreground. We will postprocess these prediction areas according to the interaction point and get new \mathcal{P}' , which is formulated as follows: where $\sigma(p_1, p_2) = 1$ when there is an eight-connected path from point p_1 to point p_2 . The structural integrity strategy can work in most cases. The effect of it can be seen in table.

3.4. Strength Analysis

Can the first click attention really improve the quality of segmentation? In this section, we will illustrate some benefits of joining the first click supervision by comparing some visual results in figure.

Focus Invariance. We know that all positive and negative points are equally important in most methods. They take all annotated points as input to generate the final result. These positive points except the first one are often clicked for repairing local details and may be close to the boundary of the target object. If the neural network treats these points equally as the first point, it will often result in a wrong segmentation. For example, in figure (a), we want to segment the table with a white tablecloth. The first click is near the center of the table. The other positive point is used to fix a corner near the edge of the table. Without the guidance of the first point, the neural network will mistakenly segment the person in the image because it treats each point equally. With the help of our first click attention, there will be fewer wrong segmentations.

Location Guidance. Obviously, the first point guides the location of the target object. If there are multiple objects in the scene, there will be less error segmentation in local regions with the help of the first point. For example, in figure (b), we want to segment the left sheep. We click three negative points around the right sheep. Without the accurate understanding of the global location information, the network may be mistaken that there is a target object in the area surrounded by these negative points. This may cause some errors, such as the wrong prediction of the right sheep. With the first click attention, the prediction will focus on the location of the first click and get a better result.



Figure 4. Illustration for benefits of first click attention. The left and right columns show the prediction masks with and without the FCA module, respectively.

the scene, there will be less error segmentation in local regions with the help of the first point. For example, in Fig. 4 (b), we want to segment the left sheep. We click three neg-

Fig.: Remove figures, tables, big equations to fascinate pdf → docx.

Grammarly

■ Prepare: Use MS Word to open pdf file and save as docx

插入 设计 布局 引用 邮件 审阅 视图 帮助 金山PDF Acrobat 百度网盘 Grammarly 告诉我你想要做什么

Just checks All Issues Correctness Clarity Engagement Delivery Deactivated Suggestions Plagiarism Get Expert Writing Help Style Guide Business Settings Help

2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48 50 52 54 56

In the task of interactive image segmentation, users initially click one point to segment the main body of the target object and then provide more points on mislabeled regions iteratively for a precise segmentation. Existing methods treat all interaction points indiscriminately, ignoring the difference between the first click and the remaining ones. In this paper, we demonstrate the critical role of the first click about providing the location and main body information of the target object. A deep framework, named First Click Attention Network (FCA-Net), is proposed to make better use of the first click. In this network, the interactive segmentation result can be much improved with the following benefits: focus invariance, location guidance, and error-tolerant ability. We then put forward a click-based loss function and a structural integrity strategy for better segmentation effect. The visualized segmentation results and sufficient experiments on five datasets demonstrate the importance of the first click and the superiority of our FCA-Net.

1. Introduction

Interactive image segmentation aims to segment the instances of interest with minimal user input. It directly benefits many applications, e.g. image editing and medical imaging analysis. Recent years, with the popularization of data-driven deep learning techniques, the demand for mask-level annotations has increased dramatically in some fields, such as salient object detection, semantic segmentation, instance segmentation, camouflaged object detection, and image/video manipulation. Efficient interactive segmentation technologies are in urgent need to alleviate the annotating cost. Therefore, more and more researchers are carrying out extensive exploration in this field.

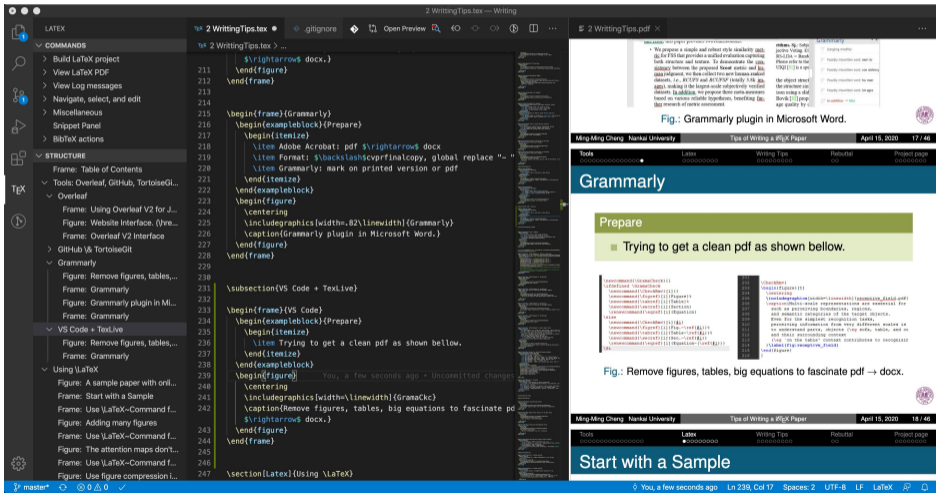
Many ways of interaction have been explored, such as bounding boxes, scribbles, and points. Drawing a bounding box as

Grammarly

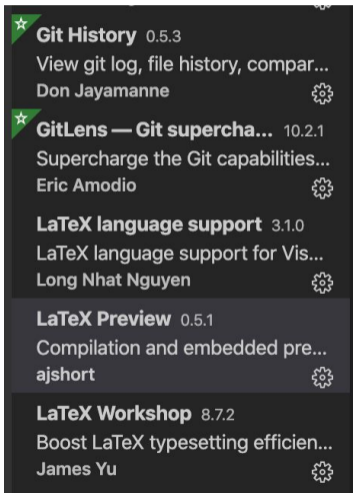
- A deep f... · Rephrase sentence
- much · Remove the phrase
- CORRECTNESS: GRAMMAR
- a better
- It seems that article use may incorrect here.
- the su... · Change the word

Fig.: Grammarly plugin in Microsoft Word.

VSCode

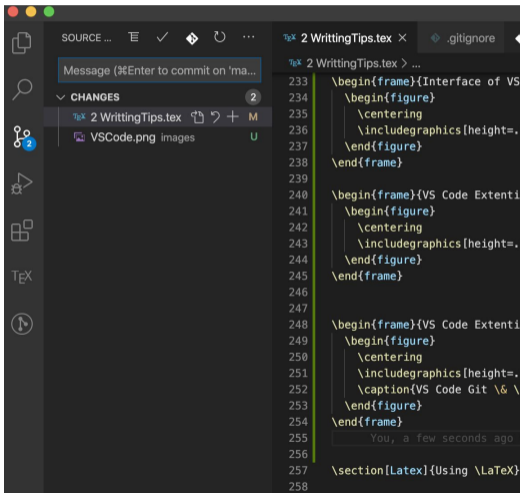


VS Code Extensions



A screenshot of the VS Code extension marketplace sidebar. It lists several extensions with their names, versions, and authors. The extensions are: Git History 0.5.3 by Don Jayamanne, GitLens — Git supercha... 10.2.1 by Eric Amodio, LaTeX language support 3.1.0 by Long Nhat Nguyen, LaTeX Preview 0.5.1 by ajshort, and LaTeX Workshop 8.7.2 by James Yu. Each extension has a gear icon for settings.

- Git History** 0.5.3
View git log, file history, compar...
Don Jayamanne
- GitLens — Git supercha...** 10.2.1
Supercharge the Git capabilities...
Eric Amodio
- LaTeX language support** 3.1.0
LaTeX language support for Vis...
Long Nhat Nguyen
- LaTeX Preview** 0.5.1
Compilation and embedded pre...
ajshort
- LaTeX Workshop** 8.7.2
Boost LaTeX typesetting efficien...
James Yu



A screenshot of the VS Code editor interface. The top part shows a commit message input field with the text "Message (#\$Enter to commit on 'ma...". Below this is a "CHANGES" section showing two files: "2 WritingTips.tex" and "VSCoDe.png images". The main editor area displays a LaTeX document with code snippets for creating frames and figures. The code includes commands like `\begin{frame}`, `\begin{figure}`, `\centering`, `\includegraphics`, and `\end{figure}`. The document is titled "2 WritingTips.tex" and has a ".gitignore" file open in the background.

```
233 \begin{frame}{Interface of VS...
234 \begin{figure}
235 \centering
236 \includegraphics[height=.5
237 \end{figure}
238 \end{frame}
239
240 \begin{frame}{VS Code Extensio
241 \begin{figure}
242 \centering
243 \includegraphics[height=.5
244 \end{figure}
245 \end{frame}
246
247
248 \begin{frame}{VS Code Extensio
249 \begin{figure}
250 \centering
251 \includegraphics[height=.5
252 \caption{VS Code Git \& \U
253 \end{figure}
254 \end{frame}
255 You, a few seconds ago
256
257 \section[Latex]{Using \LaTeX}
258
```


VS Code Extensions

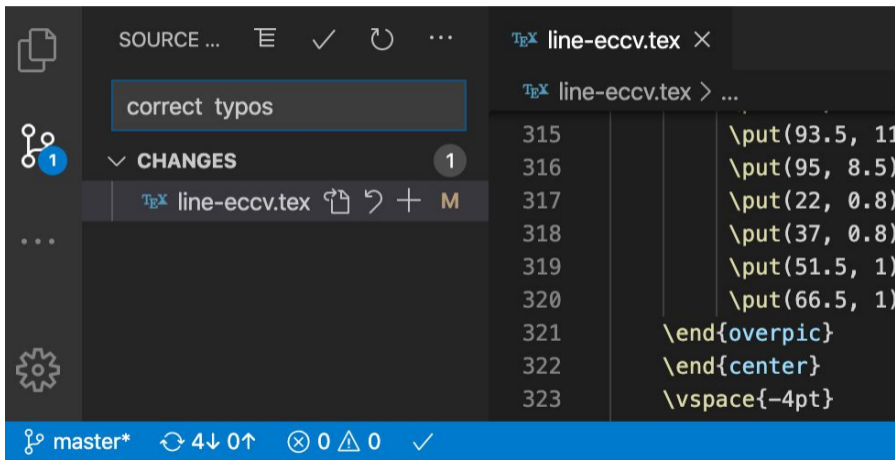


Fig.: VS Code Git & \LaTeX Extensions.

VS Code Git Extensions

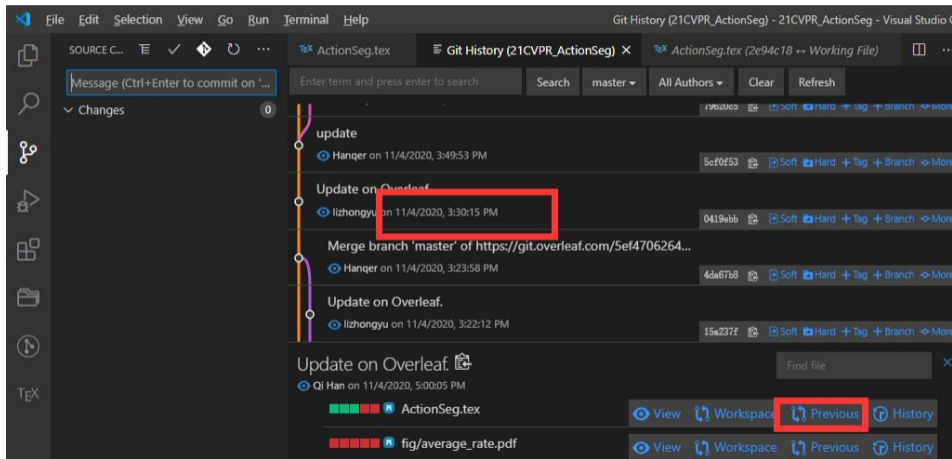
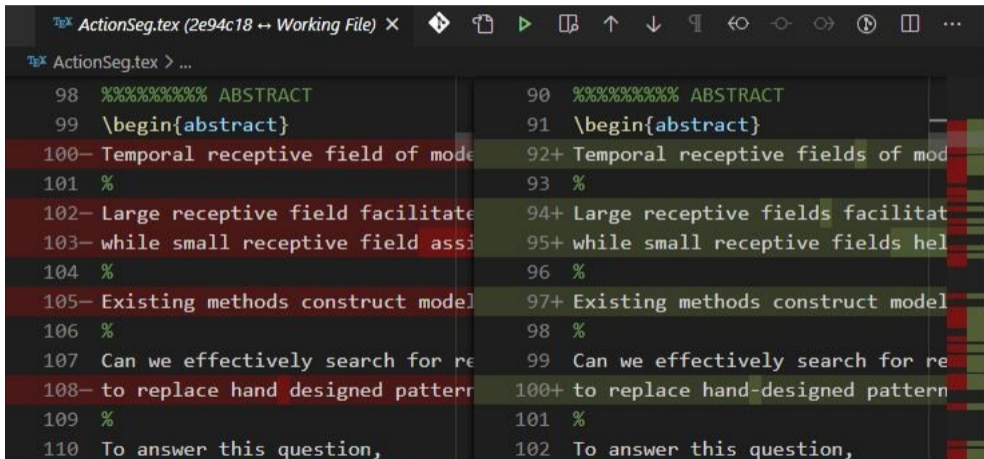


Fig.: VS Code Git extensions for version history.

VS Code Git Extensions



```
TeX ActionSeg.tex (2e94c18 ↔ Working File) ×
TeX ActionSeg.tex > ...
98 %%%%%%%%%% ABSTRACT
99 \begin{abstract}
100- Temporal receptive field of mode
101 %
102- Large receptive field facilitate
103- while small receptive field assi
104 %
105- Existing methods construct model
106 %
107 Can we effectively search for re
108- to replace hand designed pattern
109 %
110 To answer this question,
90 %%%%%%%%%% ABSTRACT
91 \begin{abstract}
92+ Temporal receptive fields of mod
93 %
94+ Large receptive fields facilitat
95+ while small receptive fields hel
96 %
97+ Existing methods construct model
98 %
99 Can we effectively search for re
100+ to replace hand-designed pattern
101 %
102 To answer this question,
```

Fig.: VS Code Git extensions for version history.

TexLive



最新TeXLive 环境的安装与配置

Start with a Sample

- You don't have to learn Latex carefully
- Simply find some useful examples and start
- You can find Latex source of many vision papers here:

<https://mmcheng.net/salobj/> or <https://www.overleaf.com/read/nkyrytybmzcm>

The screenshot displays the Overleaf online LaTeX editor interface. On the left, the source code is visible, showing the document structure with commands like `\begin{document}`, `\title{Salient Object Detection and Segmentation}`, `\title{Global Contrast based Salient Region Detection}`, and author information for Ming-Ming Cheng, Niloy J. Mitra, XiaoLei Huang, Philip H. S. Torr, and Shi-Min Hu. The right pane shows the rendered PDF output, which includes the title, authors, an abstract, and the beginning of the introduction section. The abstract describes a regional contrast-based salient object detection algorithm. The introduction starts with the sentence: "We, as humans, are experts at quickly and accurately identifying the most visually noticeable". Below the text, there are several small images illustrating salient objects in various scenes.

Fig.: A sample paper with online Latex editing.

Use \LaTeX Command for Uniform Formatting

More examples in <https://www.overleaf.com/read/zqttbzknmjrz>

Use `\newcommand{\cmd}[] {#1}`

```
\newcommand{\figref}[1]{Fig. \ref{#1}}
```

```
\newcommand{\tabref}[1]{Tab. \ref{#1}}
```

```
\newcommand{\eqnref}[1]{Eq. (\ref{#1})}
```

```
\newcommand{\secref}[1]{Sec. \ref{#1}}
```

```
\newcommand{\sArt}[1]{state-of-the-art~}
```

```
\newcommand{\AddImg}[1]{\includegraphics[height=.2\linewidth] {#1}}
```

Use L^AT_EX Command for Formatting

Adding many images: <https://www.overleaf.com/read/zqttbzknmjrz>

```
\renewcommand{\AddImg}[1]{\includegraphics[width=.245\linewidth]{samples/#1}}
\newcommand{\AddImgs}[1]{\AddImg{#1-img} \AddImg{#1-gt} \AddImg{#1-rcf}
\AddImg{#1-ucm} \vspace{.02in} }
\begin{figure*}[!t]
  \centering
  \AddImgs{bsds-97010}
  \AddImgs{bsds-368037}
  \AddImgs{nyud-5017}
  \AddImgs{nyud-6233}
  \caption{Some examples of RCF.
  \textbf{Top two rows}: BSDS500 \cite{arbelaez2011contour}.
  \textbf{Bottom two rows}: NYUD \cite{silberman2012indoor}.
  From \textbf{Left} to \textbf{Right}: origin image, ground truth,
  RCF edge map, RCF UCM map.}
  \label{fig:samples}
\end{figure*}
```

Keep proper file size for sharing

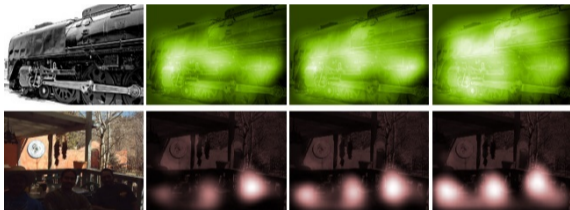


Fig.: The attention maps essentially has low resolution. Resize the length to 100px.

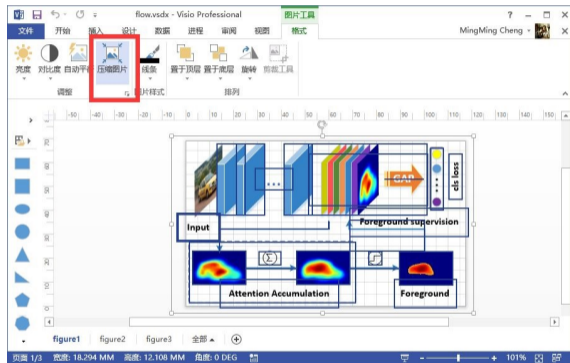


Fig.: Use figure compression in Visio to maintain proper file size.

Keep proper file size for sharing

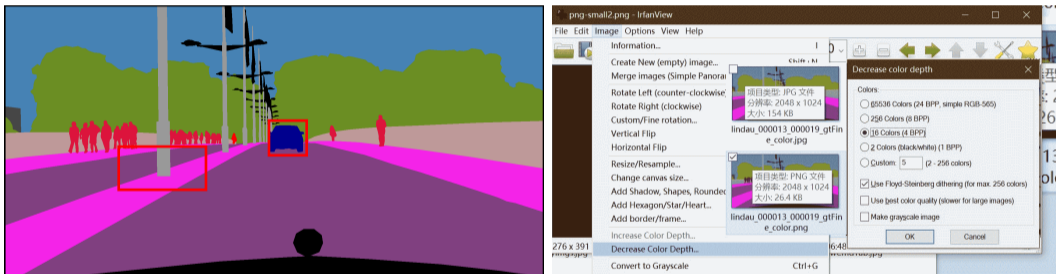


Fig.: For images with few color types, png files are smaller, otherwise, using jpg files instead.

Use \LaTeX Command for Formatting

Adding pretty table <https://www.overleaf.com/read/zqttbzknmjrz>

```

\newcommand{\gPbUCM}{gPb-UCM \cite{arbelaez2011contour}}
\newcommand{\MCG}{MCG \cite{arbelaez2014multiscale}}
\newcommand{\gPbNG}{gPb+NG \cite{gupta2013perceptual}}
\newcommand{\SENG}{SE+NG+ \cite{gupta2014learning}}
\newcommand{\COB}{COB \cite{maninis2016convolutional}}
\newcommand{\RCFResNet}{RCF-ResNet50}
\begin{figure}[!ht]
  \centering
  \begin{tabular}{c|c|c|c|c} \hline
    \multicolumn{2}{c}{Methods} & \multicolumn{2}{c}{Boundaries ( $\mathcal{S}_b$ )} \\
    & \multicolumn{2}{c}{Regions ( $\mathcal{S}_{op}$ )} \\ \cline{2-5}
    & ODS & OIS & ODS & OIS \\ \hline
    \gPbUCM & 0.631 & 0.661 & 0.242 & 0.283 \\
    \MCG & 0.651 & 0.681 & 0.264 & 0.300 \\
    \gPbNG & 0.687 & 0.716 & 0.286 & 0.324 \\
    \SENG & 0.706 & 0.734 & 0.319 & 0.359
  \end{tabular}

```

Use Strings for Journal/Conference Names

Use @String

- @string(CVPR = "IEEE CVPR")
- @string(CVPR= "IEEE Conf. Comput. Vis. Pattern Recog.")
- @string(CVPRW= "IEEE Conf. Comput. Vis. Pattern Recog. Worksh.")
- @string(NIPS= "Adv. Neural Inform. Process. Syst.")
- @string(IJCV = "Int. J. Comput. Vis.")
- @string(PAMI = "IEEE Trans. Pattern Anal. Mach. Intell.")

```
%More examples in https://www.overleaf.com/read/pcznwnthsprs
@inproceedings {arbelaez2014multiscale,
  title="Multiscale combinatorial grouping",
  author="Marques, Ferran and Malik, Jitendra",
  booktitle=CVPR,
  pages="328–335",
  year="2014"
}
```

Reference - Formats

Consist.: Use @String to ensure consistency.

Types: Use “@inproceedings” for conference, and “@Article” for journal.

Warning: **Don't** directly copy a bibtex from Google Sch.

Author: Use “and” instead of “,” to separate names.

Priority: **Important:** Journal > conference > arXiv.

Using Chinese in L^AT_EX

Refer to the demo in <http://mmcheng.net/salobj/>

Menu
中文模板(MyCVPR)
Review
Share
Submit
History
Chat

Source
Rich Text
↗
Recompile
📄
📂
↗

```

62 %%%%%%%%%% TITLE
63
64 \title{基于全局对比度的显著性区域检测\thanks{本文为
CVPR'11论文
65 \cite{Cheng_SaliencyCvpr11}的中文翻译版。}}
66
67 \author{Ming-Ming Cheng{1}\quad Guo-Xin
Zhang{1}\quad Niloy J. Mitra{2}\
68 \quad Xiaolei Huang{3}\quad Shi-Min
Hu{1}\quad \\\
69 {1} Tsinghua University \quad \quad
70 {2} KAUST \quad \quad {3} Lehigh
University\\
71 }
72
73 \maketitle
74 % \thispagestyle{empty}
75
76 %%%%%%%%%% ABSTRACT
                    
```

基于全局对比度的显著性区域检测^{*}

Ming-Ming Cheng¹ Guo-Xin Zhang¹ Niloy J. Mitra² Xiaolei Huang³ Shi-Min Hu¹

¹ Tsinghua University ² KAUST ³ Lehigh University

Abstract

可靠的视觉显著性估计使得即使在没有先验知识的情况下也可以对图像进行适当的处理。因此,视觉显著性估计仍然是许多计算机视觉任务的重要步骤,其中包括:图像分割、目标识别、自适应压缩、内容敏感图像编辑、图像检索等。我们提出了一种基于区域对比度的视觉显著性区域检测算法。该方法同时考虑全局对比度和空间相干性。该算法简单、高效,并且可以产生全分辨率的显著性图。在国际上现有最大的公开测试集上,该方法的检测结果具有更高的精度和更好的召回率,明显优于现有视觉显著



图 1. 给定输入图像(上),可以通过全局对比度分析得到高分辨率的视觉显著性图(中)。这种视觉显著性图可以进一步被用来获取感兴趣物体区域(下)。

Ming-Ming Cheng Nankai University

Tips of Writing a L^AT_EX Paper

2022/05/24 29 / 49

Avoid Large Boring Figure

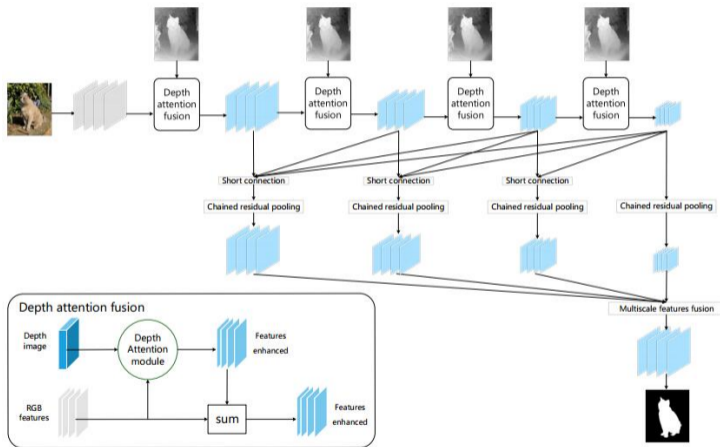
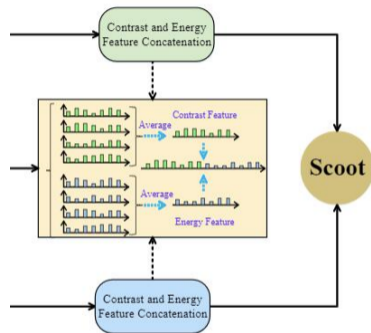


Fig.: Use smaller figure size or add more information.

Keep Figure Fonts Similar to the Main Text



pen synthesis and GT sketch. We first quantize into $K \times K$ ($K=4$ by default) blocks. Thirdly, the block-level” co-occurrence texture statistics (*e.g.*,

Fig.: Avoid too small/big fonts. Many reviewers read printed papers.

Figure Path and Intermediate Files

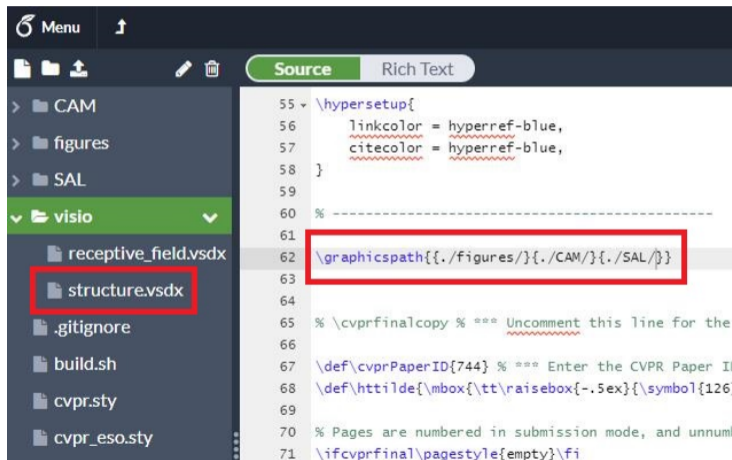


Fig.: Set figure path. Key all intermediate files for re-editing.

Avoid Descriptions without Actual Content

- Our method is better than Wang et al. [1].
‘is better’ → ‘**produces more accurate results**’.
- We propose a novel method XXX.
‘propose’ imply ‘**novel**’.
- We propose an ~~essential~~ XXX method . . .

Checklist: limited novelty

- This method is a simple combination of [A] and [B]

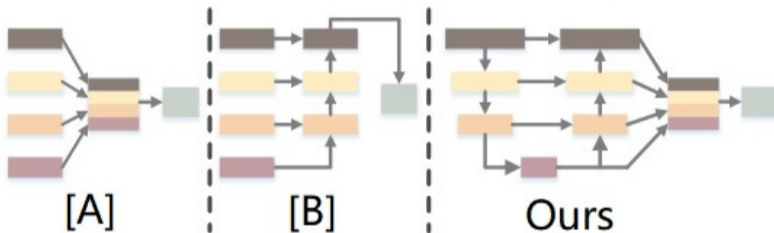
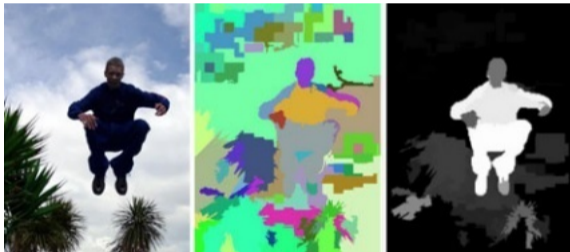


Fig.: Properly describe/illustrate your contribution.

Checklist: limited novelty

- Limited novelty, straightforward extension of xxx.



$$S(r_k) = w_s(r_k) \sum_{r_k \neq r_i} e^{-\frac{D_s(r_k, r_i)}{\sigma_s^2}} w(r_i) D_r(r_k, r_i),$$

Fig.: Explain why our **elegant/simple** approach is not **straightforward/naive** [Cheng et al., PAMI'15].

Checklist: show results in a better way.

HPatches Results

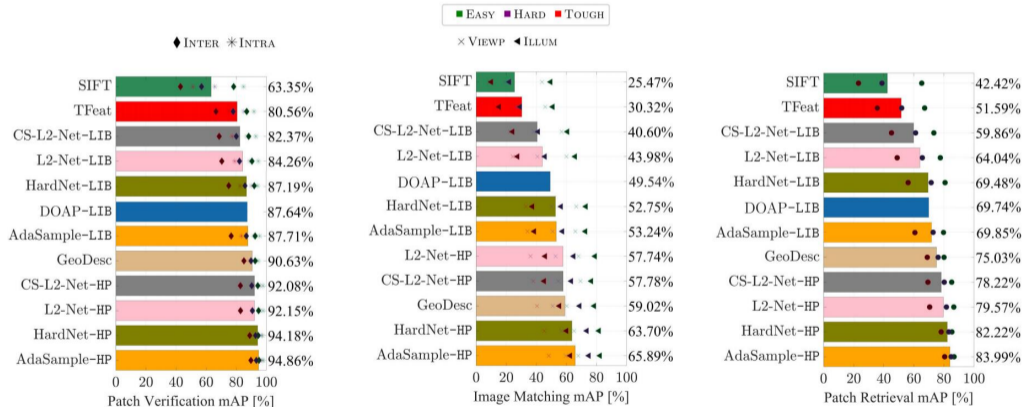


Fig.: Presents figures in a proper way.

Checklist: missing experiments to clarify

The performance gain could

- also come from ...
- due to ...
- come with the cost of ...
- how well does this and that ...
- sensitive to key parameters ...

Checklist: good practice

- Source code will be released upon paper acceptance.
- Discuss of the limitations.
- Theoretical proof regarding what could (not) happen.
- Analysis the computational/memory cost.
- Try to keep your ideas SHARP.
- Avoid defining many unnecessary terms.
- Be fair (at lease in form).
- Verified every design decision (try to keep simple).

Analysis reviewer comments

Keep claim and make a rebuttal¹

- What is peer review?
- How to rebuttal?
- Does rebuttal matter?

¹<https://zhuanlan.zhihu.com/p/104298923>

Analysis reviewer comments

- Save comments. No access after rebuttal period!
- Suggestion: size 5, Calibri, single line space, no space before&after paragraph.

1853: Shifting More Attention to Video Salient Object Detection

Reviewer #2

- **1. Summary.**
The paper makes two contributions - a large dataset of video images with salient object outlines with associated eye fixation data, and a computational model for estimating salient objects. Both contributions appear to be significantly useful to the computer vision community.
- **2. What aspects of the paper are particularly good?**
+ Contributions clearly stated and validated
+ High practical impact
+ High impact on research community
- **3. Strengths.**
+ Provides an extensive video saliency/salient object dataset with eye fixation data. This will be very useful to the attention community.
+ A method for detecting and localizing salient objects is presented which exhibits excellent performance.
- **4. What aspects of the paper most need improvement?**
- **Typos and grammar problems** make reading difficult
- **5. Weaknesses.**

shows good performance, but like all top dogs will probably fall by the wayside as progress moves on. The dataset will have a longer lasting impact.

8. Comments to author.

- What is "**catholically dynamic-viewing**"? (line 83).
- Figure 4 is labelled as "**Center bias**". Is this just the average saliency over all frames in the dataset, or is there some processing done to isolate a center bias?
- The dataset has the same issue as does DHF1K in that the fixations from different users are pooled together (even in the raw fixation maps) so the individual fixation trajectories are not available. But it is the trajectory that is important in modeling dynamic attention shifts - how the eye or attention moves from one location to the next. With a pooled map of fixations it is not possible to recover the specific transitions from one location to another. All one has is the average locations from frame to frame. Granted, even that information is useful, but it would be even better to be able to model and annotate actual scanpaths. I do not know **if individual scanpath data is**

Fig.: Highlight key issues with **bold&red** for better printing.

Project page is useful for publish & discuss

- Login: <https://mmcheng.net/salobj/>

Permalink: <https://mmcheng.net/salobj/> Edit

Global contrast based salient region detection



[Ming-Ming Cheng](#), [Niloy J. Mitra](#), [Xiaolei Huang](#), [Philip H. S. Torr](#), [Shi-Min Hu](#)



Status & Visibility ▼

🕒 5 Revisions

Permalink ▼

Categories ▼

Tags ▼

Featured Image ▲



Replace image

800x445

[Remove featured image](#)

Project page is useful for publish & discuss

- An example: <https://mmcheng.net/wp-login.php>

The screenshot shows a web browser displaying a WordPress blog post. The address bar shows the URL `https://mmcheng.net/salobj/`. The top navigation bar includes links for 'Abroad', 'Blog', 'Course', 'Graphics', 'Manuel', 'NanKai', 'Resource', 'Review', 'Vision', 'Work', '国外学者', 'Editor', and 'Shopping'. The main content area features a dark header with 'New' and 'Edit Post' buttons, both highlighted with red boxes. Below the header is a navigation menu with 'HOME', 'PEOPLE', 'RESEARCH', 'PUB', 'DATA', 'TEACHING', 'GRANTS', and 'NEWS'. The main content area displays a post titled 'Global contrast based salient region detection' under the 'Research' category. The post is dated '03/08/2013' and has '107 Comments'. The author is listed as 'Ming-Ming Cheng, Niloy J. Mitra, Xiaolei Huang, Philip H. S. Torr, Shi-Min Hu'. An 'Edit' button is highlighted with a red box. The post content includes three images: a person in a blue hood, a silhouette of an airplane, and a blue sky.

Project page is useful for publish & discuss

- An example: <https://mmcheng.net/salobj/>

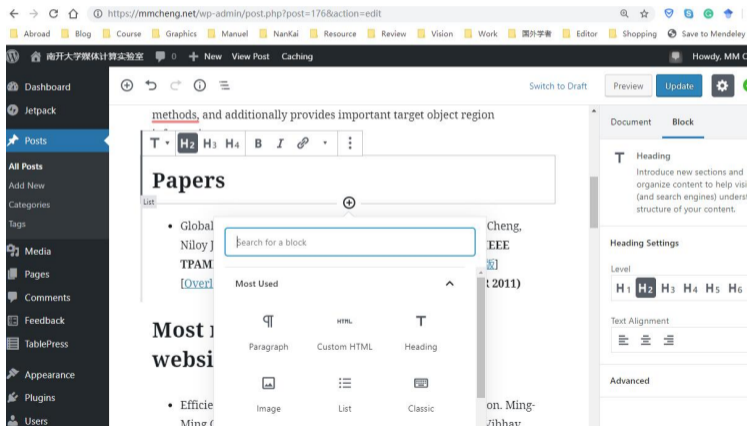


Fig.: Online editing using WordPress.

Project page is useful for publish & discuss

- Useful plug-ins: e.g. table press

The screenshot displays the WordPress TablePress plugin interface. The left sidebar contains navigation options: Dashboard, Jetpack, Posts, Media, Pages, Comments, Feedback, and TablePress (highlighted). Under TablePress, there are options for 'All Tables', 'Add New Table', 'Import a Table', 'Export a Table', 'Plugin Options', and 'About TablePress'. The main content area is divided into two sections: 'Table Information' and 'Table Content'. The 'Table Information' section shows the following details:

- Table ID: 38
- Shortcode: `[table id=38 /]`
- Table Name: Saliency segmentation time
- Description: Table. Comparison of average time for different saliency segmentation methods.
- Last Modified: 01/08/2019 19:52 by MM Cheng

The 'Table Content' section displays a table with the following data:

	A	B	C	
1	<input type="checkbox"/> Method	Time(s)	Code Type	1
2	<input type="checkbox"/> FT	0.247	Matlab	2
3	<input type="checkbox"/> <u>SEG</u>	7.48	M&C	3
4	<input type="checkbox"/> CB	36.5	M&C	4

Fig.: Online editing using WordPress Table Press plug-in.

Project page is useful for publish & discuss

■ Useful plug-ins: e.g. table press

Performance: AUC & MAE

Show entries

Search:

Method	T-AUC	T-MAE	J-AUC	J-MAE	D-AUC	D-MAE	S-AUC	S-MAE	M-AUC	M-MAE	E-AUC	E-MAE
MBD	0.915	0.162	0.838	0.225	0.903	0.168	0.922	0.137	0.964	0.107	0.917	0.172
ST	0.911	0.179	0.806	0.240	0.895	0.182	0.922	0.145	0.961	0.122	0.914	0.193
QCUT	0.907	0.128	0.831	0.178	0.897	0.119	0.860	0.148	0.956	0.118	0.909	0.171
HDCT	0.878	0.177	0.771	0.209	0.869	0.164	0.898	0.162	0.941	0.143	0.866	0.199
RBD	0.887	0.15	0.826	0.212	0.894	0.144	0.899	0.13	0.955	0.108	0.894	0.173
GR	0.829	0.256	0.747	0.311	0.846	0.259	0.854	0.189	0.925	0.198	0.831	0.285
MNP	0.854	0.255	0.768	0.286	0.835	0.272	0.888	0.215	0.895	0.229	0.82	0.307
UFO	0.853	0.165	0.775	0.216	0.839	0.173	0.845	0.18	0.938	0.15	0.875	0.207
MC	0.895	0.184	0.823	0.231	0.887	0.186	0.877	0.182	0.951	0.145	0.91	0.204
DSR	0.902	0.142	0.826	0.196	0.899	0.139	0.915	0.14	0.959	0.121	0.914	0.173

Fig.: Online editing using WordPress Table Press plug-in.

Project page is useful for publish & discuss

■ Useful plug-ins: e.g. Slides Show

The image shows the WordPress Slides Show plug-in interface on the left and a diagram of a neural network architecture on the right.

WordPress Slides Show Interface:

- Title:** Hi-Fi: Hierarchical Feature I
- Description:** UCAI 2018
- Show Information?:**
 - Both title and description
 - Title only
 - Description only
 - None, do not show
- Info Opacity:** 70
- Galleries:**
 - Select All
 - Featured researches
- Image Type:**
 - Media Upload
 - Upload
- Choose Image:** [Image thumbnail] [Choose File] [https://m...]

Neural Network Diagram:

- Input:** I_{t-1} and I_t (473 x 473 x 3)
- Architecture:** ResNet50+PDC, SSLSTM, and Loss.
- Layers:** H_{t-1} , H_t , S_{t-1} , S_t , M_{t-1} , M_t , F_{t-1} , F_t , F_{t+1} .
- Operations:** dilated conv, concatenation, element-wise multiplication.
- Losses:** L^{1000} , L^{500} , L^{250} .
- Output:** Saliency shift.

Fig.: Online editing using WordPress Slides Show plug-in.

Project page is useful for publish & discuss

■ Useful plug-ins: e.g. Slides Show

HOME PEOPLE RESEARCH PUBLICATIONS CODE & DATA TEACHING GRANTS NEWS

News Research

Shifting More Attention to Video Salient Object Detection

05/05/2019 范登平 0 Comments Edit

CVPR 2019 Oral, Accept rate = 5.6% [288/5160]

Read more

Research

Geometry-aware ICP for Scene Reconstruction from RGB-D Camera

29/04/2019 Shao-Ping Lu 0 Comments Edit

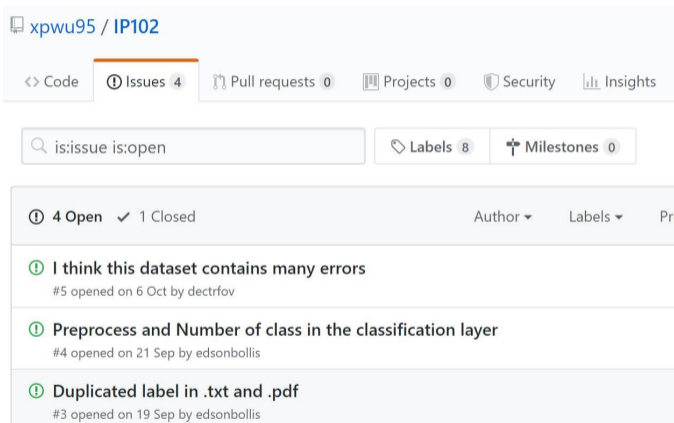
Bo Ren¹, Jia-Cheng Wu¹, Ya-Lei Lv¹, Ming-Ming Cheng¹, and Shao-Ping Lu¹
¹College of Computer Science, Nankai University, Tianjin 300350, China

Read more

Fig. 5. Illustration of critical frames using L-sharp. The top, middle and bottom rows are the inputs and reconstruction outputs using depth information at three sequential frames (200, 201) by ElasticFusion. (a) Input depth data (with RGB input at adjacent camera positions). (b) Camera pose estimations of default ICP. (c-d) Camera pose estimation applying our modification to distance metric and applying the multiplicative terms, respectively. It is clear that the default ICP behaves just as our baseline in Sub-figure 5.1, and both of our two reconstruction agents are able to provide correct camera pose estimation at this critical point.

Fig.: A good vs. bad example.

Project page is useful for publish & discuss



The screenshot shows the GitHub interface for the repository 'xpwu95 / IP102'. The navigation bar includes 'Code', 'Issues 4', 'Pull requests 0', 'Projects 0', 'Security', and 'Insights'. Below the navigation bar is a search bar with the query 'is:issue is:open' and filters for 'Labels 8' and 'Milestones 0'. The main content area displays a list of 4 open issues:

- 4 Open** ✓ 1 Closed
- I think this dataset contains many errors**
#5 opened on 6 Oct by dectrfov
- Preprocess and Number of class in the classification layer**
#4 opened on 21 Sep by edsonbollis
- Duplicated label in .txt and .pdf**
#3 opened on 19 Sep by edsonbollis

Fig.: All questions, issues, etc. **must** be replied on time.

Q & A?