基于蜕变测试的文本定位系统稳定性测试

晏荣杰 王思琪 闫艺宣 高红雨 严 俊

题目：Stability Evaluation for Text Localization Systems via Metamorphic Testing

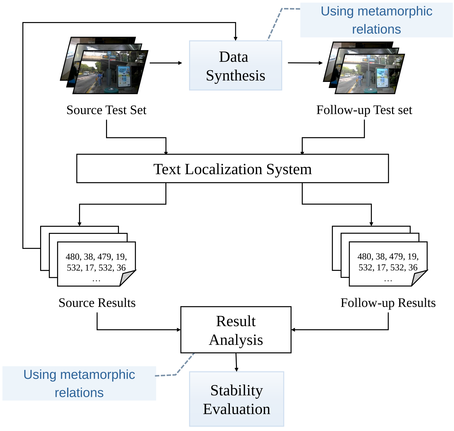
期刊：The Journal of Systems and Software

联系方式： 晏荣杰 yrj@ios.ac.cn
严俊 yanjun@ios.ac.cn

Abstract

In this paper, we propose a methodology to automatically evaluate the stability of text localization systems via metamorphic relations, where a stable system should output consistent results for similar inputs with the same text segments. We introduce six metamorphic relations that should be preserved in a stable text localization system and define the corresponding metrics for stability evaluation. With the defined metamorphic relations, we apply metamorphic testing techniques to compare the inputs and outputs to evaluate system stability, and further diagnose the causes of inconsistency. The extensive experimentation on both academic and commercial text localization systems demonstrates the effectiveness of our method on stability evaluation for such systems.

Overview



Concepts

- **Text localization systems**
  A text localization system aims to determine the positions of text segments in an image.
- **Metamorphic relation (MR)**
  If a system is correctly implemented, and the inputs of a system satisfy some relation, the outputs of the corresponding inputs should also satisfy a related relation. Such relations are the necessary properties of the system, and called MRs.
- **Stability evaluation**
  The Consistency on dataset between follow-up and source with its MR.

Metamorphic Relations

- Increasing brightness (MR_{ib})
- Decreasing brightness (MR_{db})
- Channel switch (MR_{cs})
- Perspective transformation (MR_{pt})
- Watermarking (MR_{wm})
- Masking (MR_{ma})


Source


follow-up of MR_{ib}


follow-up of MR_{db}


follow-up of MR_{mt}


follow-up of MR_{wm}


follow-up of MR_{ma}

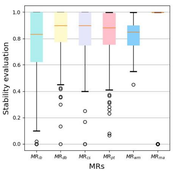

follow-up of MR_{cs}

Experimentation

- **Effectiveness of stability evaluation on text localization tools**

Stability evaluation of academic systems

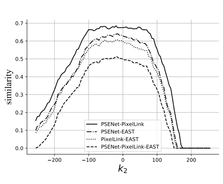
MRs	PSENet	PixelLink	EAST
MR_{ib}	0.779	0.746	0.719
MR_{db}	0.863	0.836	0.831
MR_{cs}	0.859	0.929	0.816
MR_{pt}	0.845	0.820	0.822
MR_{wm}	0.820	0.545	0.669
MR_{ma}	0.850	0.720	0.842

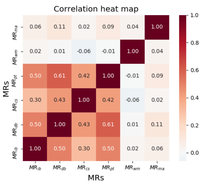


Stability evaluation of commercial systems on selected samples

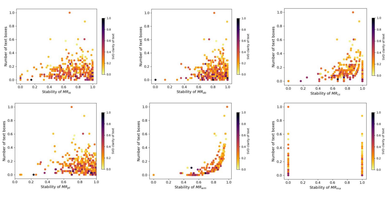
MRs	GCP	AWS	Azure	Tencent	PSENet	PixelLink	EAST
MR_{ib}	0.547	0.522	0.621	0.500	0.701	0.690	0.521
MR_{db}	0.611	0.739	0.723	0.520	0.735	0.700	0.742
MR_{cs}	0.771	0.730	0.749	0.730	0.865	0.848	0.694
MR_{pt}	0.677	0.728	0.676	0.632	0.683	0.693	0.646
MR_{wm}	0.630	0.800	0.770	0.720	0.580	0.450	0.750
MR_{ma}	0.714	0.551	0.592	0.592	0.694	0.653	1.000

- **Redundancy of metamorphic relations**





- **Usability of evaluation results**



Conclusion

- We have proposed a methodology to evaluate the stability of text localization systems with metamorphic testing techniques.
- We have introduced six MRs w.r.t. the properties of text localization systems and the feature of their inputs. Follow-ups can be generated automatically and compared with the source according to the defined metamorphic relations.
- The extensive experimentation on both academic and commercial text localization tools reveals many inconsistent outputs, and demonstrates that the methodology is effective in shooting both the advantages and disadvantages of such systems, and evaluating system stability.
- We have also investigated various image evaluation metrics to analyze the relation between the features of the 29 images and the stability of tools, which can be applied as heuristics for further diagnosis and improvement.