

Visual Attention with Deep Learning

Erik Perillo¹, Esther Luna Colombini¹

¹Institute of Computing (IC) – University of Campinas (Unicamp)
Caixa Postal 6176 – 13.084-971 – Campinas – SP – Brazil

erik.perillo@gmail.com, esther@ic.unicamp.br

Abstract. *Vision is a key element in one's process of understanding the world. The high volume of sensorial data is however problematic because most of the information is often irrelevant. Humans realize sensorial filtering by what we call attention. We propose the application of Deep Learning for obtaining a visual saliency system which behaves similarly to humans. We built a new convolutional neural network with relatively simple architecture, yielding a performance level consistently among the best ten state of the art models in MIT300 benchmark.*

Resumo. *A visão é elemento-chave no processo de entender o mundo para um ser. Entretanto, a alta quantidade de dados sensoriais é problemática, havendo muitas vezes irrelevância de informação. Nos seres humanos, há um filtro sensorial realizado pela atenção. Propomos a aplicação de Deep Learning para a obtenção de um sistema de saliência visual que se comporte como o dos seres humanos. Construímos uma nova rede neural convolucional de arquitetura relativamente simples e com um desempenho que a coloca consistentemente entre os dez melhores modelos estado da arte no MIT300 benchmark.*

1. Introduction

- What's attention?
- Visual saliency
- Old models, new methods

2. Objectives

- Use old techniques to new, good model
- Try to be simple

3. Methodology

- Old method: be succinct
- New method: architecture, datasets, datapreproc, training

4. Results

- Old method
- New method, mit300 bm

5. Conclusion

- Ey, that's pretty good
- next steps

References