# A deep learning saliency model for exploring viewers' dwell-time distributions over Areas of Interest on webcam-based eye-tracking data

Alessandro Bruno[1], Marouane Tliba[2], Arzu Çöltekin[3]
[1]Department of Computing and Informatics at Bournemouth University, Poole, United Kingdom
[2]Institut National Des Télecommunications et TIC, Oran, Algeria
[3]University of Applied Sciences and Arts Northwestern Switzerland's Institute for Interactive Technologies, Brugg-Windisch, Switzerland

## INTRODUCTION

Visual saliency is a common computational method to detect attention-drawing regions in images, abiding by top-down and bottom-up processes of visual attention. Computer Vision methods deal with Visual Saliency by extracting the so-called Saliency Maps that need to be validated through eye-tracking sessions.

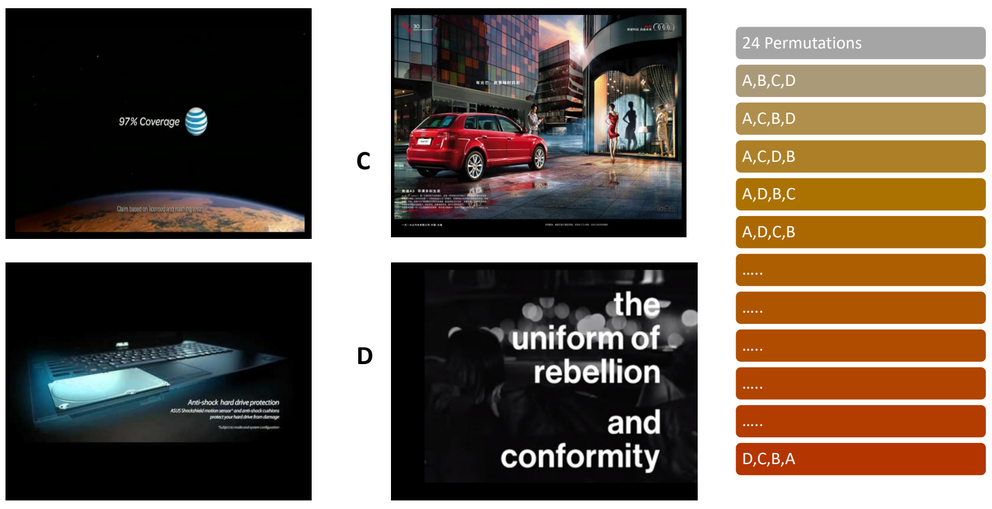
New webcam-based tools, powered by the developments in machine learning, come into play to help track down onscreen eye movements. Claimed error rates of recent webcam eye trackers can be as low as 1.05

Using webcams allows reaching a broader participant pool and collecting data over different experiments (e.g., free viewing or task-driven). We collect webcam eye-tracking data over a collection of images with 2-4 salient objects against a homogenous background. Objects within the images represent our AOIs (areas of interest).

Objectives: a) Check how eye movements vary on AOIs across all spatial permutations of the same AOI in a given image; b) Extract correlations for a given image containing N objects between viewers' eye movement dwell times over the N AOIs and the corresponding AOIs saliency maps. In our study, we focus on grid-based layouts.

## Grid-based Layout permutations

**Premise**: For a given image layout, for example a 2 by 2 grid-based layout, a number of 4! spatial permutations are given (it adds up to 24 spatial permutations).

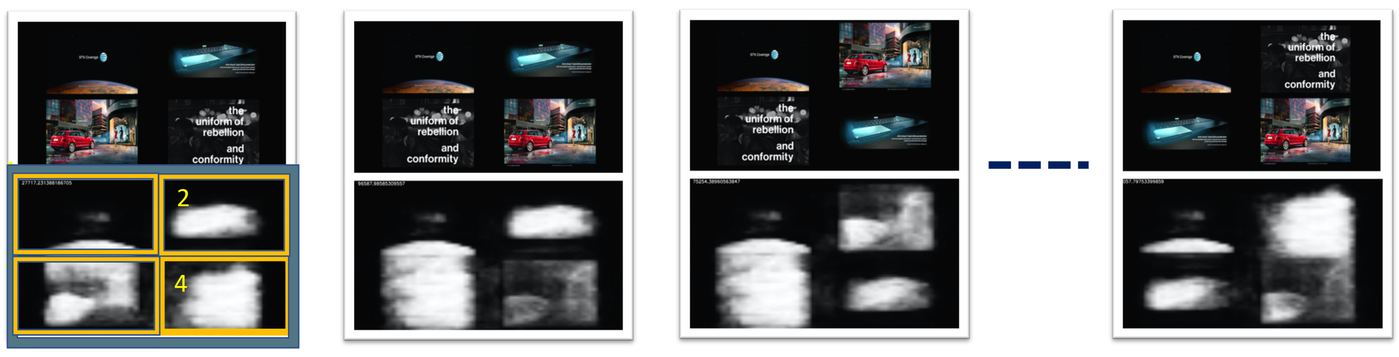



## Saliency Maps across layout permutations

Some regions, such as the one with a red car, show different local saliency maps across different spatial permutations.

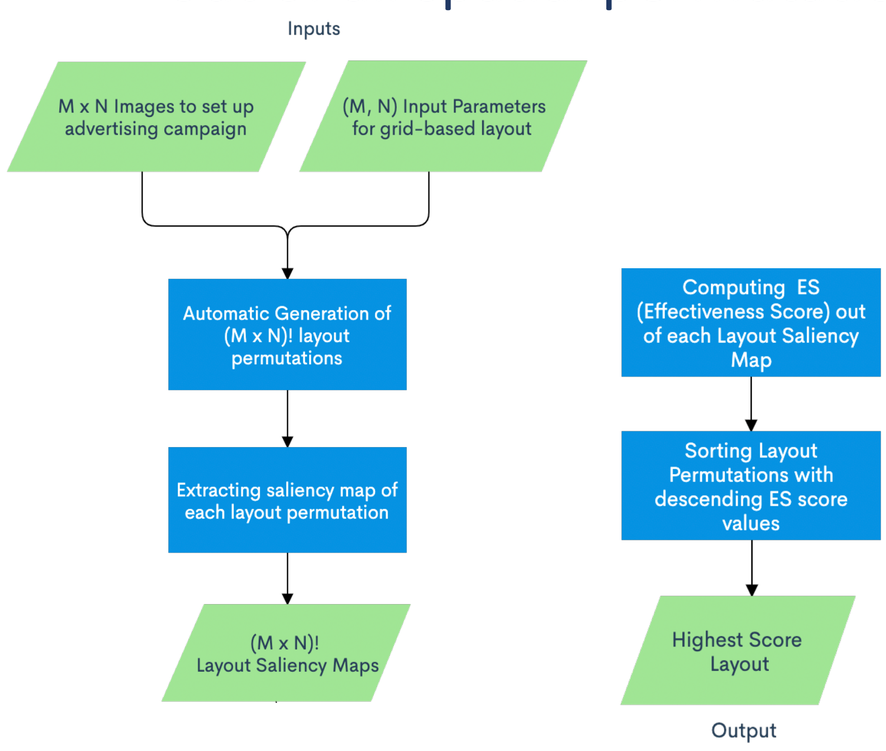
Saliency Maps are extracted by using a **deep learning-based solution**[4] evaluated over an object-oriented image and video dataset called **DAVIS** [5].

Experiments show different saliency '**behaviours**' of the same regions whose an image consists of.

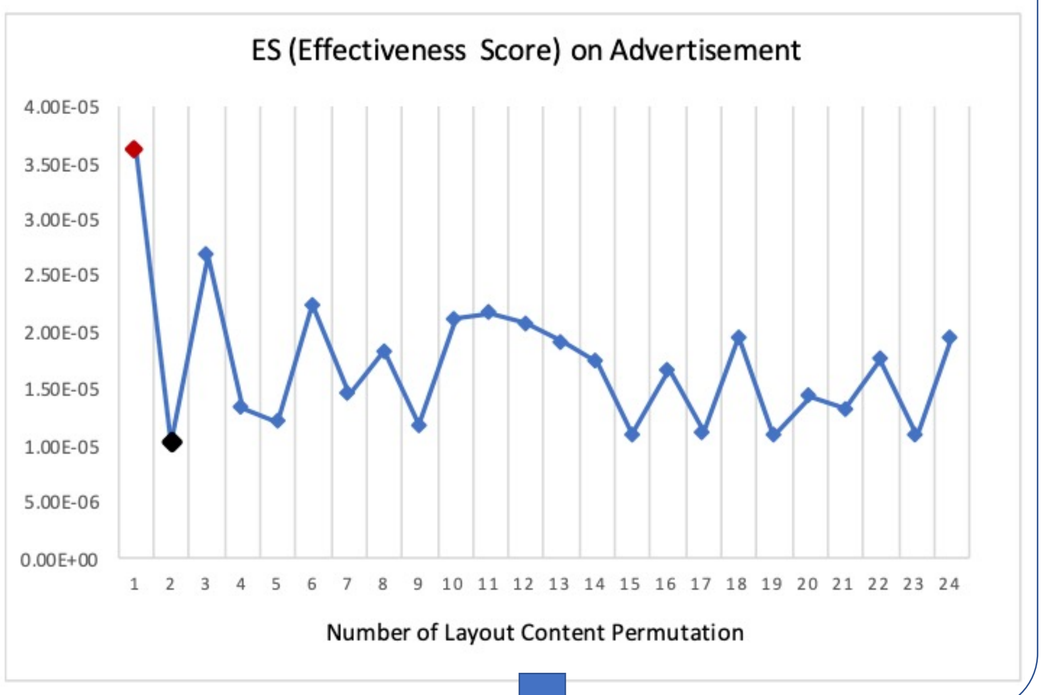



Permutation 1 | Permutation 2 | Permutation 3 | Permutation 24

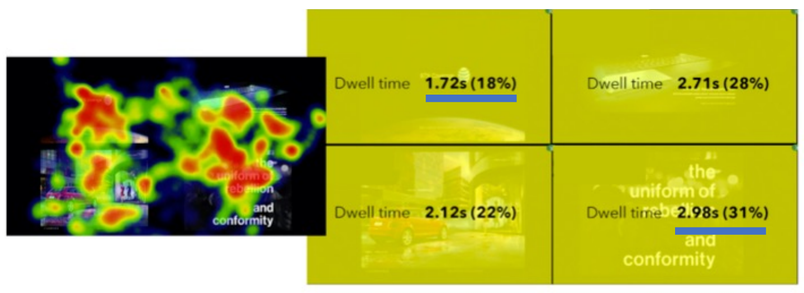
## Algorithm for generation of ES (Effectiveness Score) out of all spatial permutations



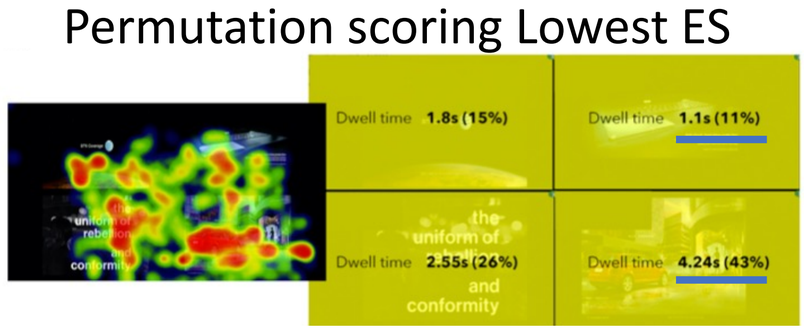

## Permutations and ES Ranking



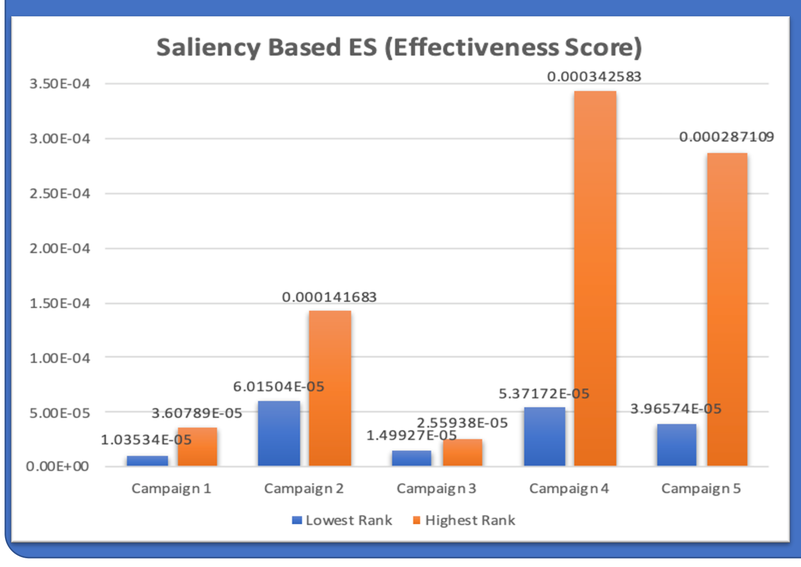

## Permutation scoring Highest ES

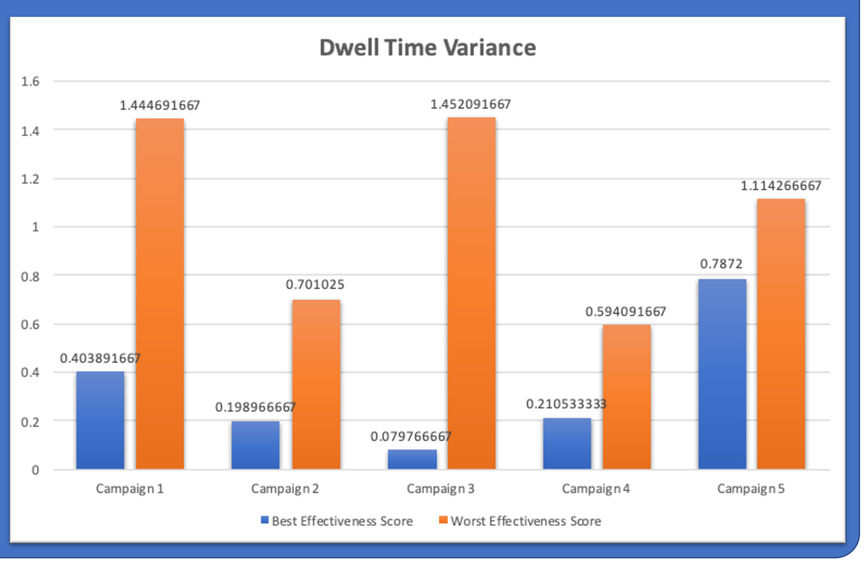



## Permutation scoring Lowest ES




## ES Score and Overall Dwell Time Variance over 5 different Graphical Campaigns




ES equations can be found in [1]

[1] Bruno, Alessandro and Lancette, Stéphane and Zhang, Jinglu and Moore, Morgan and Ward, Ville P and Chang, Jian, A Saliency-Based Technique for Advertisement Layout Optimisation to Predict Customers' Behaviour, ICPR Workshops (2), pages 495--507, 2020