



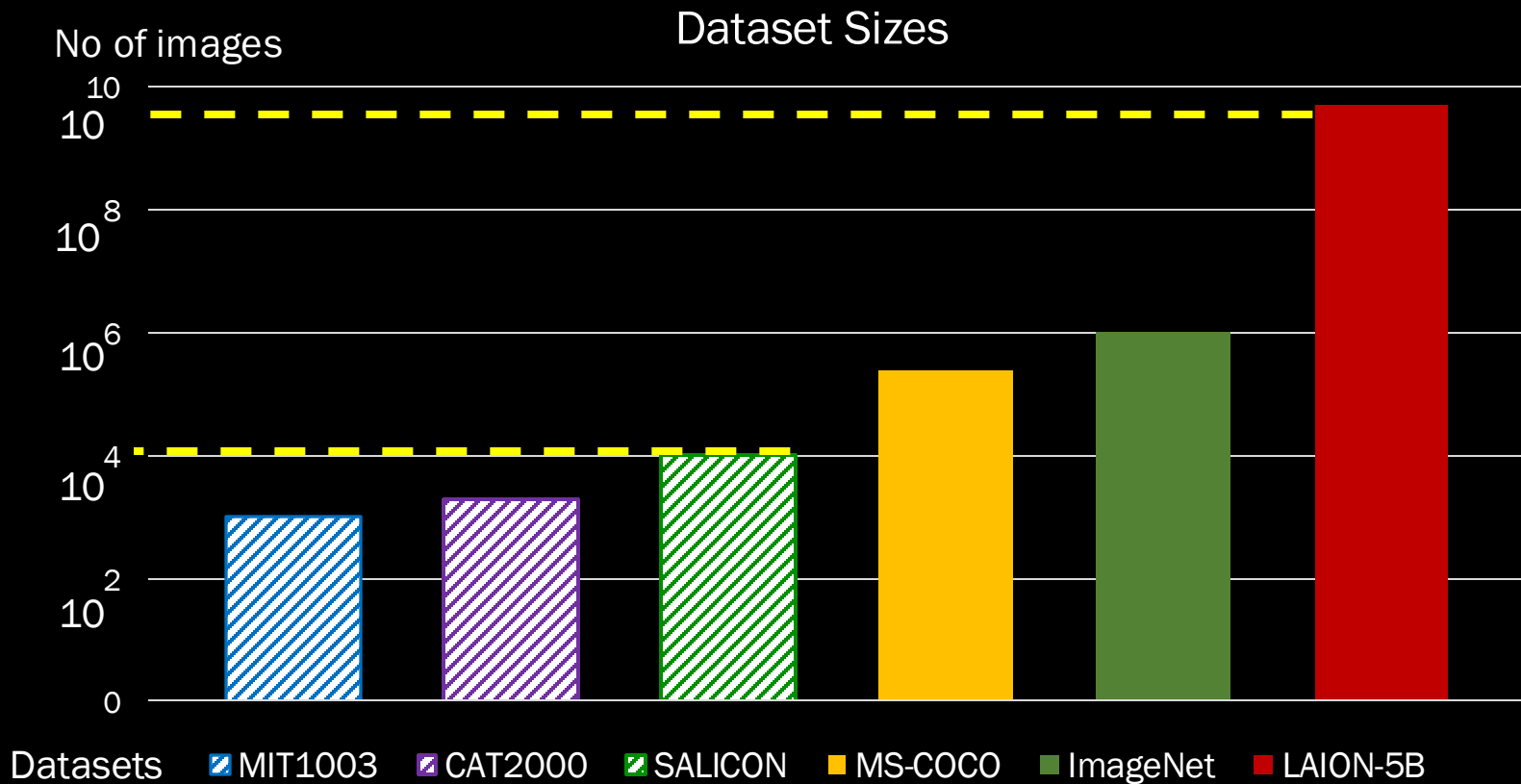
# Data Augmentation via Latent Diffusion for Saliency Prediction

Bahar Aydemir, Deblina Bhattacharjee, Tong Zhang, Mathieu Salzmann, Sabine Süsstrunk  
School of Computer and Communication Sciences, EPFL, Switzerland

<https://ivrl.github.io/Augsal>



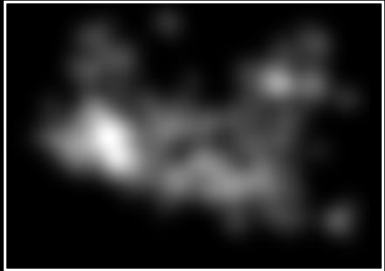
# Motivation



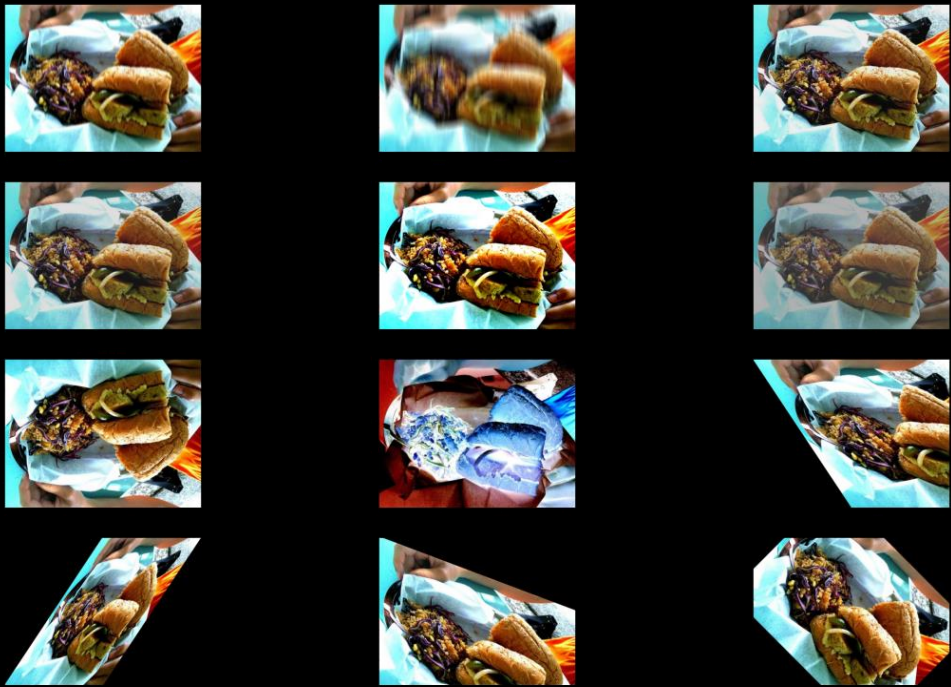
# Data Augmentation



Input image



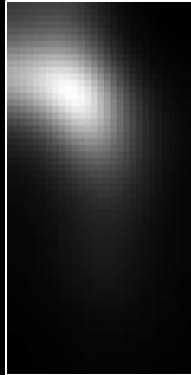
Ground Truth



Classical Augmentations

# What if we use existing augmentation techniques?

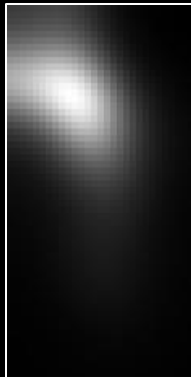
Cropping



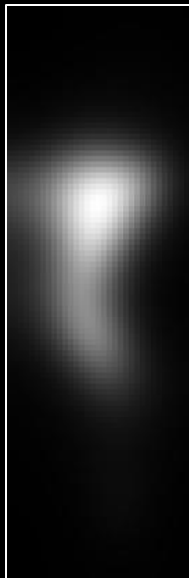
Cropped variation -1

# What if we use existing augmentation techniques?

Cropping



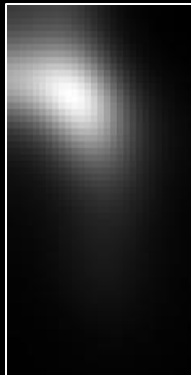
Cropped variation -1



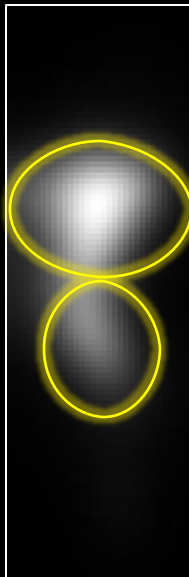
Cropped variation -2

# What if we use existing augmentation techniques?

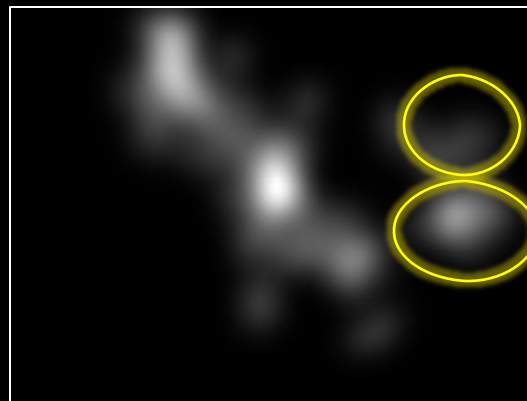
## Cropping



Cropped variation -1

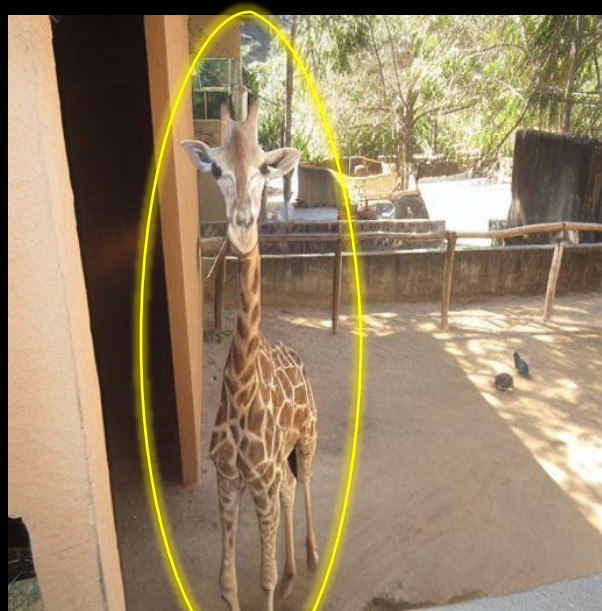


Cropped variation -2



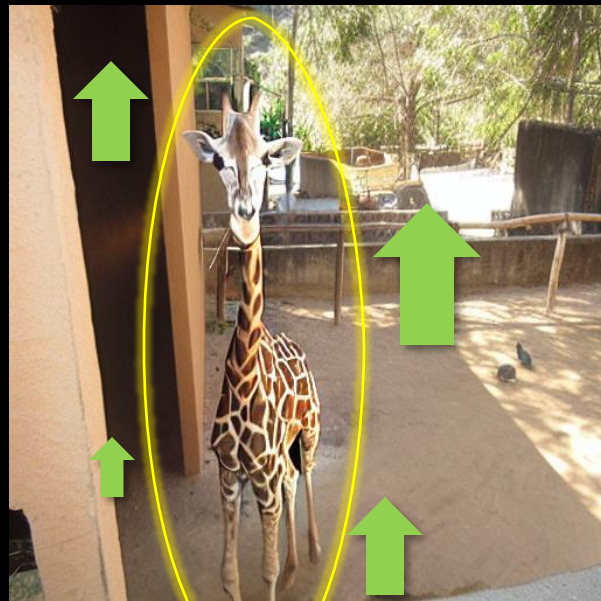
Original

# How to change saliency in a *predictable* way?



Original image

Increase  
contrast



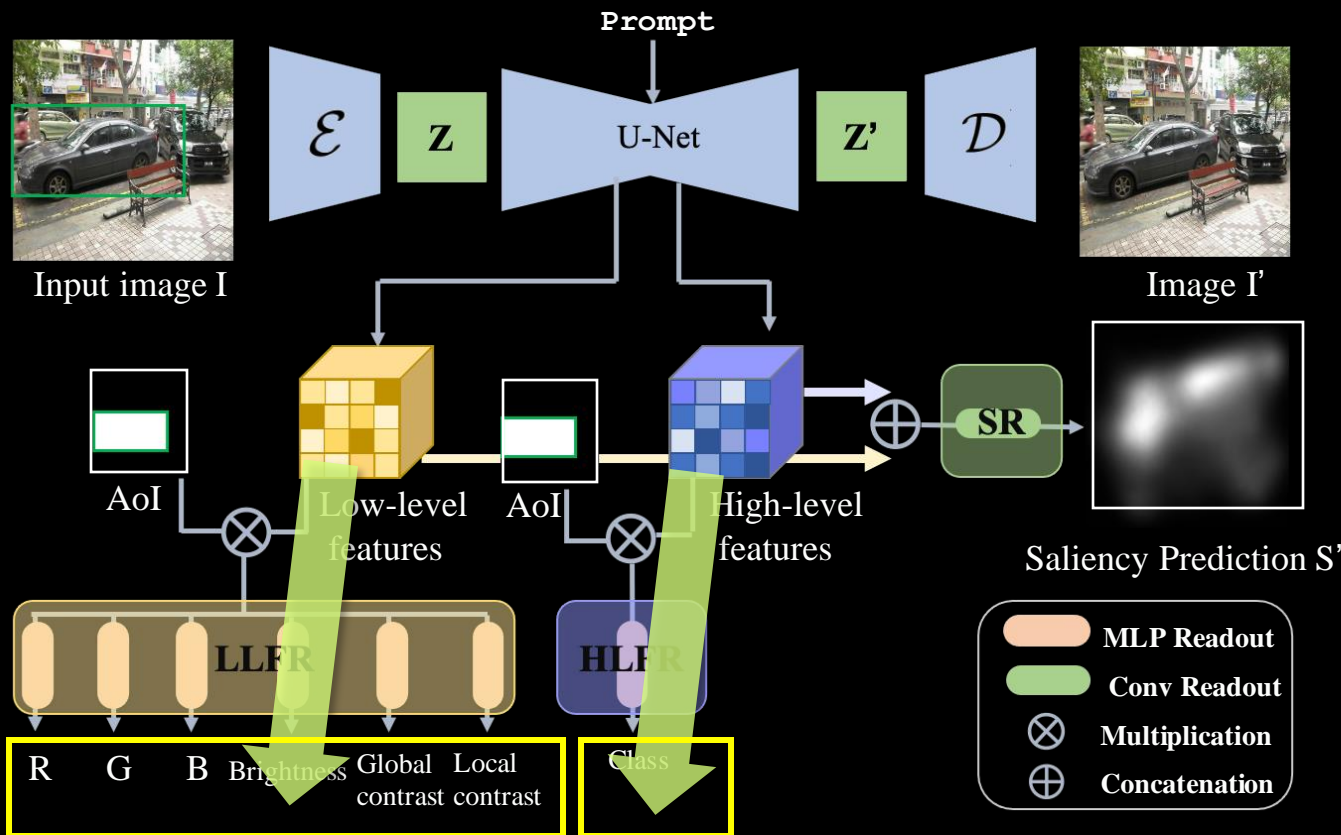
Edited image

# Method

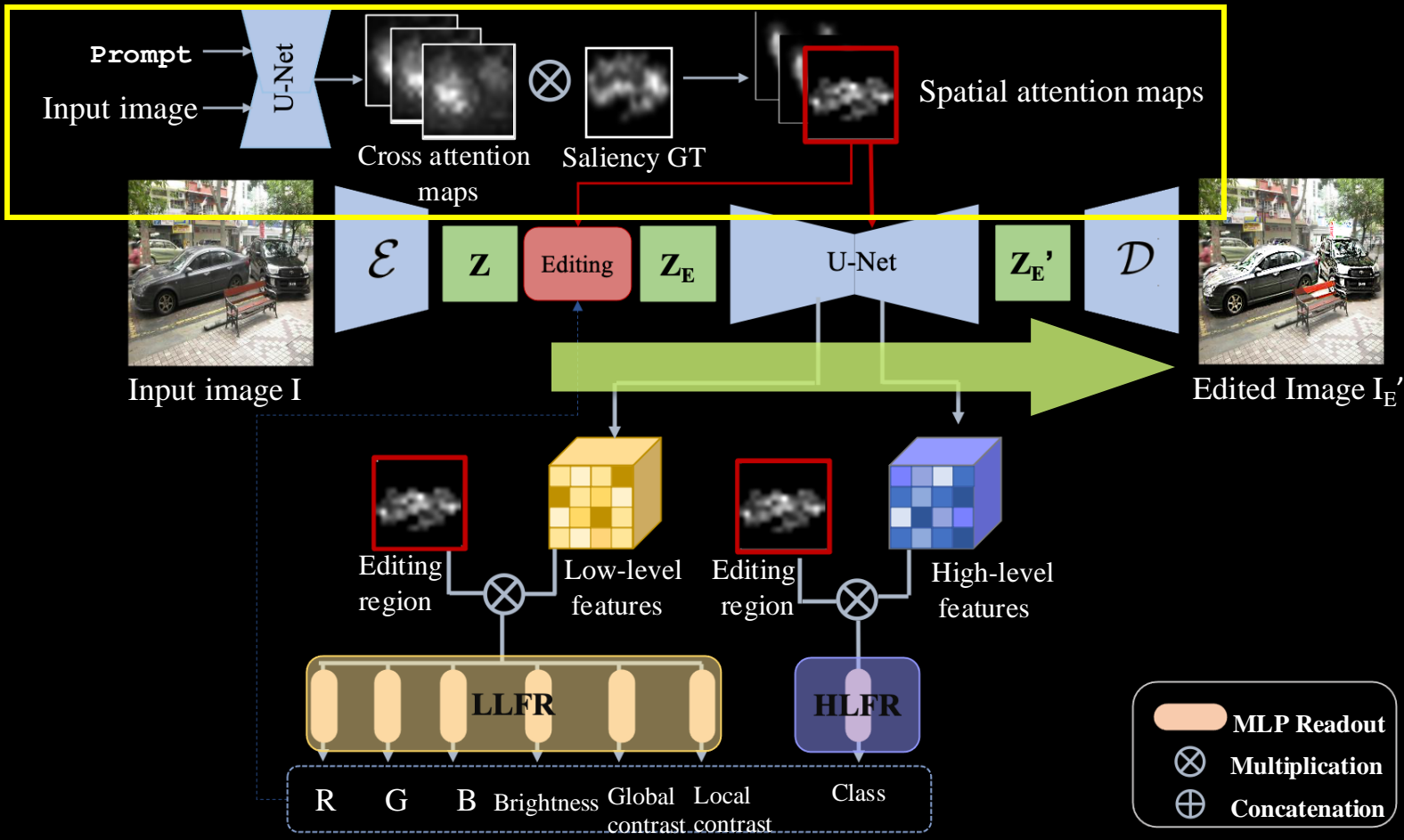




# Overview of our model



# Overview of our model

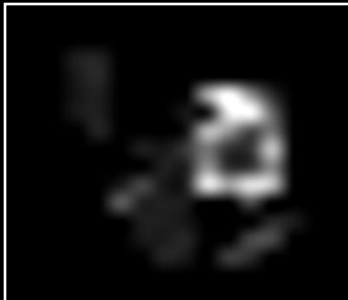


Original Images

Target Region

Intensity

Contrast



Brightness

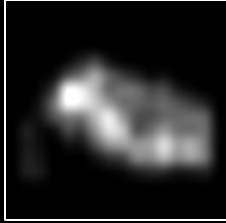


Color

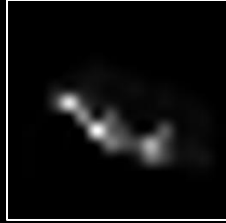




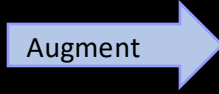
Original image



Saliency GT



Spatial Attention Map

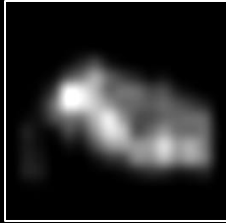


Edited image

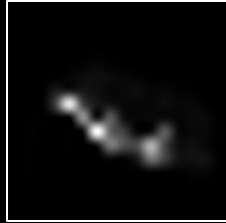
**The loss function**



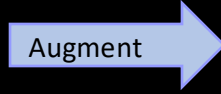
Original image



Saliency GT



Spatial Attention Map



Edited image

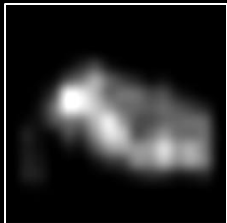


Original image

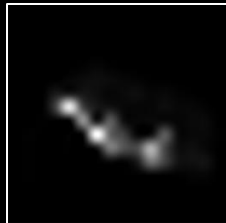
**The loss function**



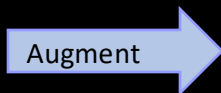
Original image



Saliency GT



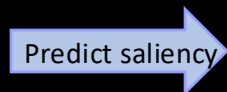
Spatial Attention Map



Edited image



Original image

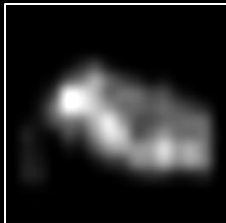


Original prediction

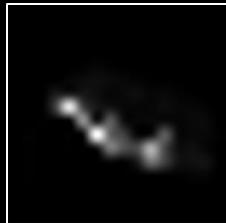
**The loss function**



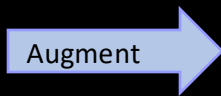
Original image



Saliency GT



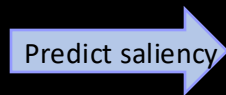
Spatial Attention Map



Edited image



Original image



Original prediction

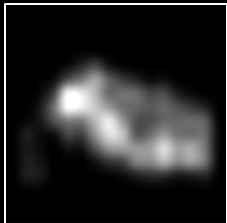


Edited image

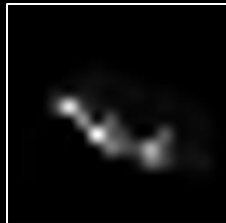
**The loss function**



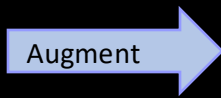
Original image



Saliency GT



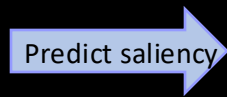
Spatial Attention Map



Edited image



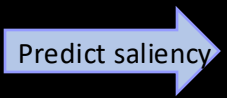
Original image



Original prediction



Edited image



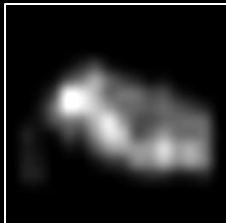
Augmented prediction

**The loss function**

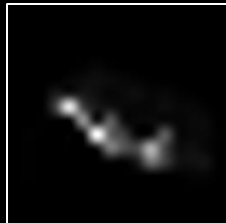




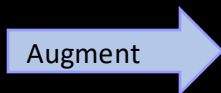
Original image



Saliency GT



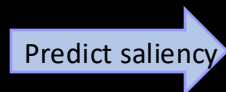
Spatial Attention Map



Edited image

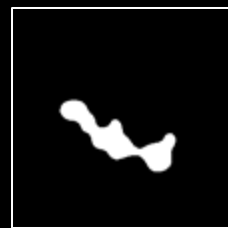


Original image



Original prediction

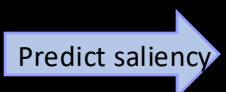
x



Binary spatial attention map

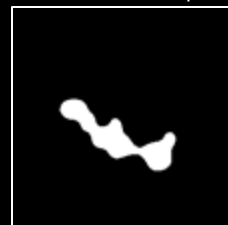


Edited image



Augmented prediction

x



Binary spatial attention map



BCE Loss

*Did you increase your saliency prediction over this region?*

# The loss function



Original image

Augment

Increase contrast



Augmented images

EPFL



<https://ivrl.github.io/Augsal>



FONDS NATIONAL SUISSE  
SCHWEIZERISCHER NATIONALFONDS  
FONDO NAZIONALE SVIZZERO  
SWISS NATIONAL SCIENCE FOUNDATION

This work was supported by the Swiss National Science Foundation via the **Sinergja grant CRSII5-180359**.