Generative Adversarial Networks (GAN) Saliency Detection, Depth Based Saliency Comparison

Hamish Sams

The University of Sheffield

3rd May 2019

What is saliency prediction?

(a) RGB (b) Depth (c) Salient

Figure: Example saliency from the NLPR dataset

(a) RGB (b) Depth (c) Salient

Figure: Example saliency from the Olesova dataset

Saliency prediction algorithms identify eyecatching objects in a scene/quantify how salient each pixel of a scene is.

What is a GAN

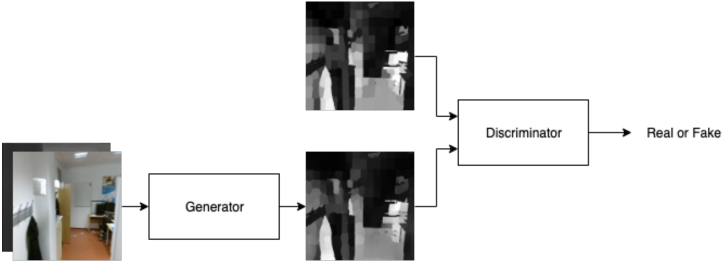

Figure: GAN System diagram

GANs are a design of neural network based on two networks:

- Generator - To create fake images based off some input
- Discriminator - To classify real and fake images

Discriminator

A discriminator is based on a convolutional neural network to classify images.

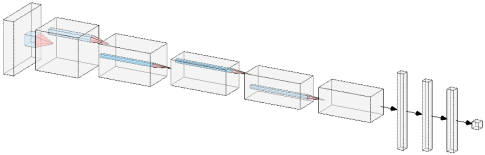

Figure: Convolutional neural network

A convolutional neural network is based on image convolution to downsample an image into a more processable amount of data. with each convolution more feature maps are generated. Once the images are small enough the data is linearised and densely connected.

CNN Example

```python
image = cv2.imread('uniLogo.png')

image=cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

image=image/255

plt.imshow(image)
plt.axis('off')
plt.show()
print("Shape: ",image.shape)
print("Size: ",image.size,image.size/image.size*100, "%")
```


```
Shape:  (225, 225, 3)
Size:  151875 100.0 %
```

Figure: Load image code

```python
#Define model
model = Sequential()
model.add(Conv2D(25,5,5,input_shape=image.shape))

#Print model
out=model.predict(image.reshape([1,225, 225, 3]))
out = out[0,:, :, :]
for i in range(0, len(out[0][0])):
    plt.subplot(5,len(out[0][0])/5,i+1)
    plt.imshow(out[:,:,i],cmap=plt.cm.binary)
    plt.axis('off')
plt.show()
print("shape: ",out.shape)
print("Size: ",out.size,out.size/image.size*100, "%")
```

```
shape:  (45, 45, 25)
Size:  50625 33.33333333333333 %
```

Figure: Tensorflow single Convolution code

Generator

A generator is based on a convolutional neural network followed by a de-convolutional neural network.

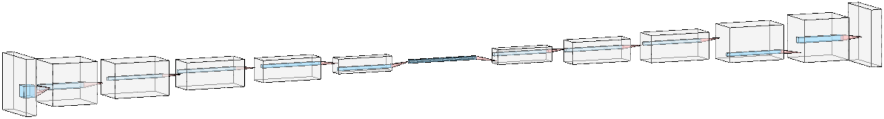

Figure: Convolutional to de-convolutional neural network

Project aims

There are two main aims for this project:

- Create a novel saliency prediction algorithm using a GAN.
- Compare saliency prediction with RGB and RGB-D

Methodology

1. Choose GAN generator and discriminator design
2. Train 3 GANs to each accept: RGB, Depth and RGB-D separately
3. Use other combination methods on RGB and depth data to compare ours to.
4. Quantify model efficiency

Results

Two different methods were used to quantify model accuracy:

1. F-measure
2. Receiver operating characteristic curve (ROC)

Both of these are based on True positive rate, false positive rate, true negative rate and false negative rate.

F-Measure

F-Measure is a numeric value calculated as the harmonic mean of precision and recall:

$$F = 2 \cdot \frac{\frac{TP}{TP+FP} \cdot \frac{TP}{TP+FN}}{\frac{TP}{TP+FP} + \frac{TP}{TP+FN}} \tag{1}$$

1. TP - True positive rate
2. FP - False positive rate
3. FN - False negative rate

F-Measure ranges from 1-0 with 1 meaning 100% true positive rate and 0% false positives/negatives.

Receiver operating characteristic

ROC is a line generated by varying a threshold on an image and comparing that to ground truth.

Figure: Varying threshold on saliency map

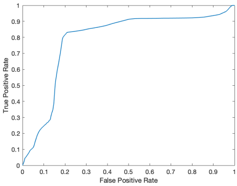

Figure: Single ROC

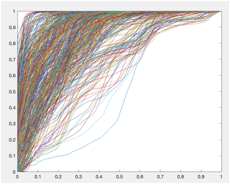

Figure: All ROC in dataset

Results

Model	F-Measure	AUC
Ours(RGB)	0.7248	0.9227
GBVS(RGB)	0.3163	0.7052
Itti(RGB)	0.3097	0.7380
LMH	0.1878	0.7646

Model	F-Measure	AUC
GP	0.5607	0.8668
GBVS (RGB)	0.5113	0.8927
Itti (RGB)	0.4221	0.8597
Ours (*)	0.3950	0.7731
LMH	0.3208	0.7921

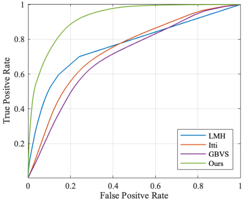

Figure: Olesova results

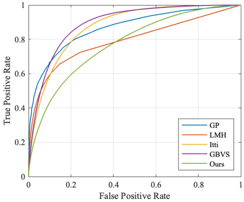

Figure: NLPR results

Results - Depth

Combination	F - Measure	AUC
RGB	0.7248	0.9227
RGBDepth	0.6866	0.8723
RGB+Depth	0.6121	0.8674
RGB*Depth	0.4144	0.8227
Depth	0.4131	0.6821

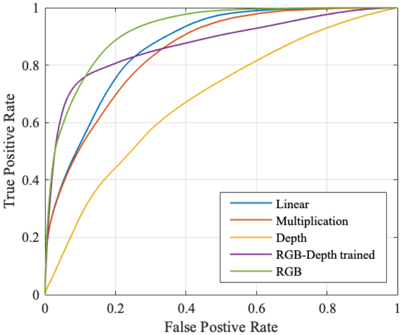

Figure: Our combination methods compared

Future Work

A few ideas:

- Design neural network to take multiple inputs (3/4 channel)
- Run network on a larger epoch system
- Train/Test on similar saliency datasets
- Implement the network in a system such as compression

Conclusion

We set out with the following aims:

- Create a novel saliency prediction algorithm using a GAN.
- Compare saliency prediction with RGB and RGB-D

Any Questions?