Saliency based Domain Adaptation

Pranav Goyal
B.Tech CSE
IIIT Delhi
pranav17078@iiitd.ac.in

Vishesh Agrawal
B.Tech CSB
IIIT Delhi
vishesh18420@iiitd.ac.in

Lavanya Verma
B.Tech CSE
IIIT Delhi
vishesh18115@iiitd.ac.in

Bhavye Anand Gupta
B.Tech CSE
IIIT Delhi
bhavye17038@iiitd.ac.in

ABSTRACT

In this report we explain our CV course project on Data Augmentation and Domain Adaptation using saliency-based methods. Our primary objective was Domain Adaptation using saliency-based methods and our secondary objective was Data Augmentation using saliency-based methods. We experimented with a couple of novel approaches using saliency and finally determined the approach which worked the best. In the task of few-shot Domain Adaptation, we were able to achieve an improved accuracy of 27% on the Office-31 dataset. Further, for zero-shot Domain Adaptation we were able to achieve an improved accuracy of 5%.

1. INTRODUCTION

Deep Learning approaches have shown to be brittle to domain shifts [1], i.e., small changes in the input distribution (via noise, quality difference, visual appearance, changes in image surroundings and other measures). For this reason, they fail to perform well on shifted distributions even when they are able to perform at superhuman levels on the data with which they were trained.

Changes in data distribution often occurs in real life scenarios, and thus, researchers have started looking into this important problem of Domain Adaptation.

We came up with few promising solutions based on the novel usage of saliency for Domain Adaptation. The idea behind using saliency was that irrespective of distribution shifts the salient object, or in this case, the foreground will remain the same and hence, the saliency maps will be able to play an important role for Domain Adaptation. This is similar to how the human visual system works as it is fairly good in identifying the same objects even in changed surroundings, etc.

We considered the problem of few shot Domain Adaptation, i.e., very few seen images from the target domain.

2. METHODOLOGY

The crux of our approaches was to attempt to come up with a domain invariant feature map through which the neural network is able to classify both the source domain images as well as the target domain images.

We used three different methods for saliency prediction:

- Spectral Residual Approach in OpenCV. [2]
- Co-Saliency Detection using contrast cues and spatial cues [3]
- BASNet: Boundary-Aware Salient Object Detection. [4] This is the State of the Art method for saliency prediction.

2.1 First Approach

In our first approach we used a foreground extracted map as our input feature for training and testing purposes. This was obtained by performing otsu binarization of the saliency map to give a segmented map, followed by multiplying the segmented map with the input image.

2.2 Second Approach

In the second approach, we used saliency maps multiplied by input image as the input feature. This input feature was a smoother version of the foreground extracted images.

Figure 1: Foreground extracted map (for first approach)

2.3 Third Approach

In the third approach, we used the saliency map and segmented map as extra channels along with the R, G, B channels in the following ways:

- Using saliency map as the fourth channel along with R, G, B channels.
- Using segmented map as the fourth channel along with R, G, B channels.
- Using saliency map and segmented map as the fourth and fifth channels.

2.4 Fourth Approach

The fourth approach is centered around saliency-based patch transfer. This was proposed as a data augmentation strategy in [5]. We extended this approach for our problem of domain adaptation and came up with new augmented images by mixing source and target images. Specifically, we cropped salient patches from images of one domain and superposed them on salient regions of images of the other domain, while labelling the obtained image as the label of transferred image. We also implemented the data augmentation technique from [5] to reduce training time for achieving 100% accuracy on intra-domain classification.

We also tried a variation of salient patch transfer using image mixing by setting the proportion of transferred and transformed image as

saliency(transferred image) : 1-saliency(transformed image).

2.5 Fifth Approach

We also tried a multi-input method by taking the image and it's foreground extracted version as the two inputs.

3. EXPERIMENTAL RESULTS

Figure 2: Soft foreground extracted map (for second approach)

Figure 3: Patch transferred from source to target image (for fourth approach)

3.1 About the Dataset

We worked on the Office-31 domain adaptation dataset. This contains 2817 high quality images in the Amazon domain, 795 low quality, noisy images in the Webcam domain and 498 high resolution DSLR images. This is a multi-category classification dataset with 31 classes of objects commonly found in office workplaces, that are common across all the domains. We worked with Amazon as our source domain and Webcam as our target domain.

3.2 Experiments

We performed all our experiments using the Resnet-18 ImageNet pre-trained network. For our inter-domain task, we kept access to 10% of the target samples during training. The training was stopped at the best target domain performance. (Further training was leading to decrease in target performance). The following sets of experiments were performed:

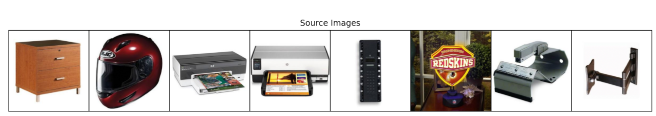

Figure 4: Office-31 images from the Amazon domain

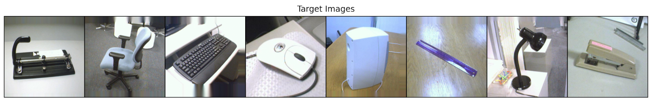

Figure 5: Office-31 images from the Webcam domain

1. Baseline Intra-domain training and testing.
2. Intra-domain training and testing with salient patch transfer data augmentation.
3. Baseline Inter-domain training and testing.
4. Inter-domain training and testing with the approaches of section 2.
5. Zero shot domain adaptation.

3.3 Results

Domain	Epochs	Train Acc.	Test Acc.
Amazon	5	100	99
Webcam	3	100	98

Table 1: Baseline Intra-Domain Classification

Domain	Epochs	Train Acc.	Test Acc.
Amazon	2	100	99
Webcam	1	100	99

Table 2: Intra-Domain Classification with data augmentation

4. CONCLUSION

We have thus shown initial results that saliency could act as a very useful tool for Domain Adaptation. We designed multiple saliency based methods and our best performing method derived from a salient patch transfer data augmentation method gave an improvement of 27% in the target domain accuracy for few shot domain adaptation in which only 10% target samples were taken for training.

We also tried our hands on zero-shot domain adaptation. While the problem was much harder, our best performing method was able to give a 5% improvement in target accuracy.

5. REFERENCES

[1] A. Farahani et al., "A Brief Review of Domain Adaptation", 2020.

[2] X. Hou et al., "Saliency Detection: A Spectral Residual Approach", 2007.

Source	Train Loss	Test Loss	Train Acc.	Test Acc.
Amazon	1.6	3.5	62	18
Webcam	1.4	3.6	68	15

Table 3: Baseline Inter-Domain Classification

Approach No.	Train Loss	Test Loss	Train Acc.	Test Acc.
1.	1.6	2.9	63	29
2.	1.6	3.1	63	25
3.	1.8	3.4	60	19
4.	**1.2**	**2.2**	**72**	**45**
5.	1.5	2.8	64	31

Table 4: Inter-Domain Classification with proposed approaches

Train Loss	Test Loss	Train Acc.	Test Acc.
1.4	5.6	67	0.5

Table 5: Baseline zero-shot domain adaptation

Train Loss	Test Loss	Train Acc.	Test Acc.
1.2	4.4	73	6

Table 6: Zero-shot domain adaptation using our best performing fourth approach

[3] H. Fu et al., ”Cluster-based Co-saliency Detection”, 2013.
[4] X. Qin et al., BASNet: Boundary-Aware Salient Object Detection”, CVPR 2019.
[5] Uddin AF et al., ”SaliencyMix: A Saliency Guided Data Augmentation Strategy for Better Regularization”, 2020.