XAI-Guided Interventional Retraining for Domain Generation

QIU Yaowen
The Hong Kong University of Science and Technology
yqiuau@connect.ust.hk

***Abstract*—Models trained in one domain usually do not perform well in other domains. For example, models trained on ImageNet have much lower accuracies on ImageNet V2. The paper proposed an idea of two-stage training consists of a) regular training and b) interventional retraining. The regular training is same as training an image classifier while the interventional retraining leverages the Grad-CAM to generate purified images with object of interest only. The model is forced to learn the correct correlation between valid feature units and classes. A feature contribution formular is defined and the heatmap visualization proves that some feature units are indeed more important than others during classification. Several experiments and a super-class binary classification are done to investigate in the paper. The result illustrates that pure interventional retraining has limit impact on reducing the accuracy gap between domains because the continuous training of baseline model will strength the gap. However, if the interventional retraining is done on a continuing trained baseline, the result will be better. And the result of binary super-class classification demonstrates that model can distinguish the super-class very successfully by keeping the unimportant feature units frozen. It shows that the general idea of our project is valid and valuable.**

Keywords—X-AI, Interventional Retraining, Grad-CAM, super-class

I. INTRODUCTION

Models trained in one domain usually do not perform well in other domains. For example, models trained on ImageNet have much lower accuracies on ImageNet V2. Domain adaptation is about learning models that generalize well across domains. One of the most important reason is the model may focus on wrong object instead of the ground truth reference. Imaging an image labelled as horse describes that a horse is on a grassland. A model with no bias will only focus on the area where the horse locates while a worse one may also focus on some parts of the grassland and associate them with the label. Thus, the worse model may wrongly predict an image as horse just because grassland is on the image when it is applied on other domains of data.

Labels often highly associated with certain group of feature units. In other words, for a certain label, some feature units are more important than others. The paper uses the term "contribution" to measure the degree of importance of a feature unit to a label. Experiments are done to demonstrate that the contribution of each pair of feature unit and label are different and a heatmap is helping to visualize it.

When classifying an image, the output depends on the object of interest and the background. The paper proposed a method for domain adaptation that consists of two training phases: (1) regular training, and (2) interventional retraining (IR). The regular training is exactly same as ordinary image classifier model training. When regular training is done, the parameters of feature extraction layers, or the backbone, is frozen and the prediction modules are activated. A subset of regular training set where the background of each image in the subset is removed is used in IR. In the second training state, the prediction module of the model will force to learn to rely on the important feature unit only. The purpose of this method is to reduce the impact of background. The subset is filtered by the perplexity, a measurement of difficulty of classification of an image.

II. RELATED WORK

A. Domain Generation

Domain adaptation is the ability to apply an algorithm trained in one or more "source domains" to a different "target domain". In domain adaptation, the source and target domains all have the same feature space but different distributions. A domain shift is a change in the data distribution between an algorithm's training dataset, and a dataset it encounters when deployed. Domain adaptation is significant in real-life problem. In many applications such as medicine development and earthquake prediction, there are usually lack of enough data to train a well-behaved model. Usually, people choose to train the model on similar domains that have enough data and perform fine-tuning. However, since the distribution is always different, it is difficult to achieve same performance as training domain.

B. Perplexity

The term perplexity is used to describe the extent of difficulty of classifying. [1] illustrates the concept of example perplexity in details. Motivated by the observation that an image is difficult to classify for humans if many people find it confusing and classify it correctly, the paper created a population of N classifiers. First, it re-sampled the training images using proportion of 25%, 50%, and 75%, respectively. With the original training images set, totally 10 different training sets are found. Then, it selected 10 main-stream CNN architectures (i.e., VGG, ResNet, etc) and trained them on aforementioned training image set and save each model in different stages. Totally $N = 500$ models are created.

For a given example $\mathbf{x}$, $P_i(y \mid \mathbf{x})$ is the probability distribution computed by model i. The information entropy $H(P_i(y \mid \mathbf{x})) = -\sum_y P_i(y \mid \mathbf{x}) \log_2 P_i(y \mid \mathbf{x})$ is a measure of the degree of confidence of model i classifying example $\mathbf{x}$. The perplexity of the probability distribution is defined to be $2^{H(P_i(y|\mathbf{x}))}$. The C-Perplexity is defined as the geometric mean of perplexity of the probability distribution over N models, as (1) demonstrates. The higher the C-perplexity, the higher the uncertainty or the lower the confidence that a classifier tries to classify example $\mathbf{x}$.

$$\Phi_C(\mathbf{x}) = \left[\prod_{i=1}^{N} 2^{H\left(P_i(y|\mathbf{x})\right)}\right]^{\frac{1}{N}} \tag{1}$$

The definition of X-Perplexity is the fraction of misclassification over N models, as shown in (2).

$$\Phi_X(\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}(C_i(\mathbf{x}) \neq y) \tag{2}$$

C. Grad-CAM

Grad-CAM, or Gradient-weighted Class Activation Mapping, is a method using for visualizing the contribution of each pixel to the final model classification. It is a heatmap of the model's attention on the image. It locates the true label object when the model is no-bias. It is class-discriminative, which means the activation map is different for each label.

Let $A^k = [a_{ij}^k]$ be a feature map in the last convolutional layer for an input x. The quantity $\alpha_k^c = \frac{1}{Z}\sum_{i,j} \frac{\partial z_c}{\partial a_{ij}^k}$ can measure the "important" of the feature map A^k is to the score $z_c(x)$. Then, the Grad-CAM heatmap is computed as $L_{Grad-cAM}^c = \mathrm{ReL}U(\sum_k \alpha_k^c A^K)$. The $\mathrm{ReL}U$ is used since we only interested in features that have positive contribution.

III. Methodology

A. Framework

There are mainly two parts in common Convolutional Neural Network (CNN): a feature extractor, or backbone, and a prediction module. The backbone extracts both low-level and high-level features through convolutional kernels. The prediction module accepts the extracted features as input and output the probability of being classified as a label. When classifying an image, the output depends on the object of interest and the background. The domain shift problem usually causes a huge decrease in performance when a model trained on source domain is applied on the target domain. One of the main reasons is that the model fails to learn the correct relationship of between labels and object of interest. The paper proposed a method for domain adaptation that consists of two training phases: (1) regular training, and (2) interventional retraining (IR), shown in the Fig.1. The regular training is exactly the same as the way of training image classifiers. A batch of image data is fed into the model and model will output a probability vector of all classes. The loss is calculated using the ground truth and the prediction. The error is backward to each layer and the parameters of each layer will be updated using gradient descent method. This procedure is conducted several rounds, or epochs, to train a model from scratch. Since ImageNet is popular among image classification competition, there are tons of pretrained model on the internet. For the regular training, the paper directly adopts the ResNet-50 pretrained model. The purpose of the interventional training is to force the model to learn the correct correlation between the ground truth label and feature units. First, an image with object of interest only will be fed into the model. A feature map can be computed through the backbone. Then, the prediction module will give the probability vector of classes belonging according to the feature map. Once the error is computed, only the parameters of prediction module will be updated. The parameters

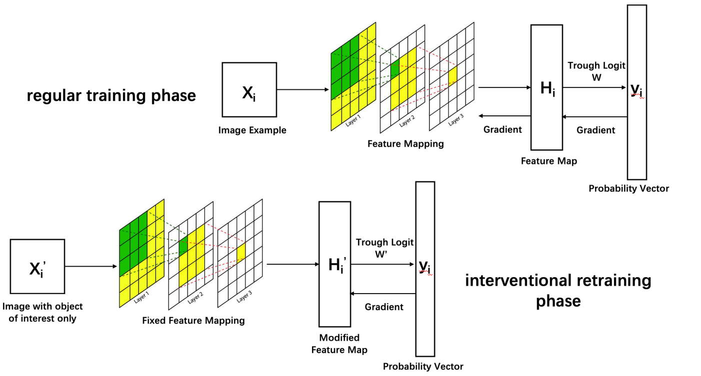

Fig.1. Framework overview. There are two training phases: (1) regular training, and (2) interventional retraining (IR)

Of the feature extractor will be fixed since we presume that it can already learn the high-level features during stage 1.

B. Locates object of interest

As mentioned above, the Grad-CAM method can locate the object of interest in an image. Suppose a model can correctly classify an image with low entropy, which mean it has great confidence. We can compute its Grad-CAM on the image and will find that the pixels with high positive contribution can depict the shape of the object of interest precisely. If a threshold is used to transform the activation heatmap as a binary mask with 0/1 value only, the pixel with value 0 will be removed and pixel with value 1 is kept. In the end, we can get the image with object of interest only. Fig.2 demonstrates the whole process.

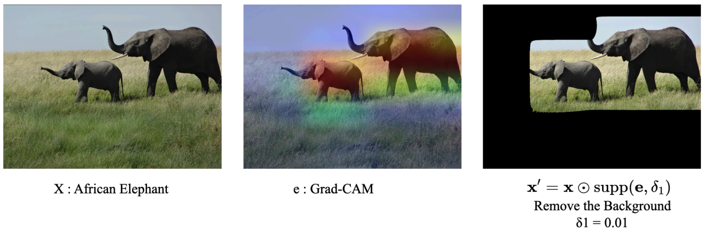

Fig.2. an illustration of remove background of an image using Grad-CAM

The value of δ_1 significantly influences the quality of the generated image. Fig.3. demonstrates the change of generated images with the change of δ_1. The ground truth label of the example image is rifle. When $\delta_1 = 0.1$, the whole part of the rifle is included in the purified image but the proportion of it is only about 50%. With the increase of δ_1, the part of rifle is kept but the proportion it takes in the image is also increased. A proper setting of δ_1 can let the model learn the correlation between the object of interest and label more precisely.

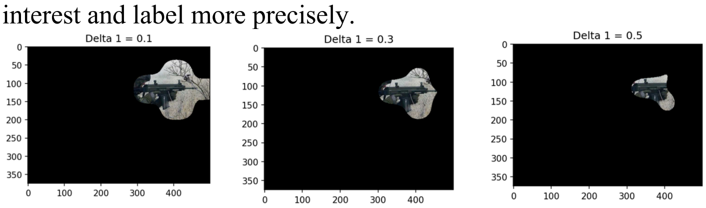

Fig.3. The value of δ_1 significantly influences the quality of generated image

C. Visualization explanation of important feature units

It is hard to prove that some feature units are indeed more important than others for each during image classification. But there is an intuitive way to visualize it.

Define a training example is $(\boldsymbol{x}_i, y_i)$, we feed it into a model m to get the heatmap $\boldsymbol{e}_i$. From $\boldsymbol{e}_i$, using the method proposed in part B, we could get a "purified" input $\boldsymbol{x}_i'$, a new image with background removed. Then, we feed $\boldsymbol{x}_i'$ into m and compute and activation h'_{ui} for each feature unit h_u. Let $a_{ui} = h'_{ui} w_{uy_i}$, it is the contribution of h_u regards to class y_i. In overall, we define the contribution of feature vector for a class c through

$$A_{uc} = \frac{\sum_{i:y_i=c} a_{ui}}{\sum_{i:y_i=c} 1} \tag{3}$$

We could get a heatmap to see which group of feature units are more important for each class. To improve the precision of the heatmap, C-perplexity is used to filter a purified training set.

Fig.3 demonstrates the overall heatmap of contribution against all feature units(y-axis) to all classes(x-axis). From the heatmap, we can observe that for each class, indeed some feature units are more important. And the set of important feature units are varied across all classes. Some feature units are more likely to be important than others.

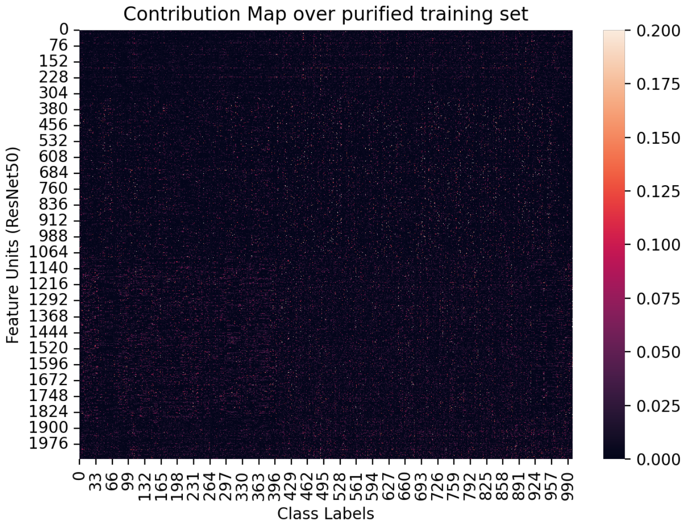

Fig .4. The contribution heatmap over purified training set.

D. Baseline Model

A baseline model should be built to compare with methods used. Create a baseline model is easy. A pretrained ResNet-50 [4] model is used as the baseline Model.

TABLE I. PERFORMANCE OF BASELINE MODEL

Model	V1 Top-1 Acc.	V1 Top-5 Acc.	V2 Top-1 Acc.	V2 Top-5 Acc.	Gap-1	Gap-5
ResNet-50	75.67%	92.44%	63.01%	84.14%	12.66%	8.30%

TABLE I demonstrates the performance of pretrain ResNet-50 Model. The accuracy gap between ImageNet validation set and imagenet-v2 is 12.66%.

E. Version I: Use all training examples only

First, a small experiment based on two-stage training idea is conducted. All the training examples in ILSVRC 2012 are used to generate purified examples, the image with presumed object of interest only. The parameter of last layers (i.e. prediction layer) is random initialized by Xaiver. The threshold of activation map is $\delta_1 = \max_p \mathbf{e}_i(p) \times 0.3$.

F. Version II: Use "Purified training set"

In this version, instead of using all training set in ILSVRC 2012, a "purified" dataset is used. C perplexity and X perplexity are used to filter the original training set. We calculate the 20% percentile value of overall C perplexity of all training set and set this value to the threshold to filter the training set. Images with C-perplexity larger than threshold will be removed. Furthermore, the threshold of X-perplexity is set to be 0.3. Images with X-perplexity larger than threshold will be removed, either. About 18.67% of training set is kept. he threshold of activation map is $\delta_1 = \max_p \mathbf{e}_i(p) \times 0.15$.

G. Version III: Union training and class balance of Version1

There are some drawbacks in Version1. First, the distribution of class in "purified" training set is not considered. The class distribution of original training set is balanced. However, since the perplexity of each class is quite different, the class distribution of filtered dataset is changed. Some classes are easier to classified while some classes are not. In this case, the number of examples of a certain group of classes is greater than others. To overcome it, the class discriminative filter is used. We kept the percentile of C-perplexity unchanged, while filter pure examples for each class. For each class, examples with C-perplexity smaller or equal than the local threshold are kept. It guarantees that exactly 20% of training set is kept and the class distribution remains the same. The second improve is union training. The pretrained model is used for classifying common image. In version II, only the purified images are used to re-train the model, which will add confusion to the model. The union training expands the training set with original training, which could help the model learn the positive correlation between the object of interest and ground truth label. Furthermore, weight decayed is adopted to restrict the amplitude of change of parameters of prediction layers.

H. Version IV: Adpot stronger baseline

The baseline significantly influences the quality of purified examples. Models with better performance on classification can generate clearer and directive explanation heatmap of the object of interest. If the Intersection over Unit (IOU) between Grad-CAM and the object of interest is larger, the model could learn the correlation better. In this version of training, the model InceptionV3 is used. Other training settings are same as Version 2. Furthermore, to investigate the trend of pretrained model, two contrast experiments are done. First, the pretrained model is trained in the next 5 epochs with the same hyper-parameters mentioned above. The performance of continuing pretrained model is recorded. Second, the baseline model and the continuing pretrained model will be adopted version III training.

Table II demonstrates the performance of InceptionV3 model. Compared to ResNet50, both the accuracy of validation set and imagenetv2 of InceptionV3 are higher. However, the decrease of accuracy between two domains are also higher.

TABLE II. PERFORMANCE OF VERSION IV BASELINE MODEL

Model	V1 Top-1 Acc.	V1 Top-5 Acc.	V2 Top-1 Acc.	V2 Top-5 Acc.	Gap-1	Gap-5
Inception-V3	77.90%	93.70%	63.87%	85.17%	14.03%	8.53%

I. *Super-class classification*

To give brief demonstrate and validation of whether the "important features" can be used to classify a group of classes, a tiny experiment called super-classification. It transfers the 1k-classification problem to binary classification buy picking a super-class. The task for the model is to distinguish whether the given image belongs to the super-class or not. The paper chosen car as the super-class, which contains 10 real ImageNet labels: ambulance, beach wagon, cab, convertible, jeep, limousine, minivan, model T, racer, and sport car. The proportion of the number of super-class examples and non-super-class examples is 1:99. For convenience, the super-class is considered as positive label. A two-stage training approach is used in the super-class classification. A pretrained model is used as the backbone and its parameters are frozen during training. In the last convolutional layers, the important feature units are active while the unimportant feature units are inactive (i.e. set to 0). The prediction layer is random initialized. Since the class distribution is highly imbalance, the weight of loss function is calculated differently.

IV. EXPERIMENT

A. *Dataset*

Both the training set and validation are from ILSVRC 2012. There are 1.28M training images in the training set with 1,000 categories and 50,000 validation images. In addition, researchers [2] have shown that most of classifiers overfit to existing validation set, and they have developed a novel validation set called Imagenetv2[3] which all 10,000 examples are sampled from ImageNet. Most of models drop accuracy on this version of validation set. In this project, it is used as another validation set to investigate whether methods improve the generalization ability of the model. To distinguish this validation set, uses validation set I indicates original validation set and validation set II indicates Imagenetv2.

B. *Training*

The hyper parameters of following experiments are set to be the same. The epoch is set to 5. The optimizer used is Adam with learning rate 1e-4.

V. RESULT AND ANALYSIS

A. *Version I to Version III*

Table III demonstrates the performance of Version I to Version III training. Version I and version II have similar performance on each metric. Both the top-1 accuracy and top-5 accuracy decrease in the validation set. Version II drops 0.87% accuracy on top-5 accuracy of validation set. And Version I also drop 0.64% accuracy on top-1 accuracy of validation set. Compared to the decrease on validation set, the decrease on imagenetv2 is minored. In spite of the decrease on the validation set, the gap decreases. Version I reduce the domain gap about 0.62% and version II decreases the domain gap about 0.74%. It shows that both of training method helps to solve the domain generation problem. The case of Version III training is completely different from the others. Both the top-1 accuracy and top-5 accuracy increase, but all other metrics show worse performance. One of the possible reasons is the union strength the bias on original training set, which shrinkage the effect of intervene training.

TABLE III. PERFORMANCE OF VERSION I TO VERSION III

	Validation I		Validation II		Gap (I – II)	
	Top-1	Top-5	Top-1	Top-5	Top-1	Top-5
ResNet50	75.67%	92.44%	63.01%	84.14%	12.66%	8.30%
Version I	75.03% (-0.64%)	92.01% (-0.43%)	62.99% (-0.22%)	84.01% (-0.13%)	12.04% (-0.62%)	8.00% (-0.30%)
Version II	74.89% (-0.78%)	91.57% (-0.87%)	62.55% (-0.46%)	84.01% (-0.13%)	12.34% (-0.32%)	7.56% (-0.74%)
Version III	75.16% (+0.11%)	92.59% (+0.13%)	62.98% (-0.03%)	84.13% (-0.01%)	12.78% (+0.12%)	8.46% (+0.16%)

TABLE IV. PERFORMANCE OF VERSION I TO VERSION III

	Validation I		Validation II		Gap (I – II)	
	Top-1	Top-5	Top-1	Top-5	Top-1	Top-5
Inceptionv3	77.90%	93.70%	63.87%	85.17%	14.03%	8.53%
Inceptionv3 (epoch 5)	78.15% (+0.25%)	94.10% (+0.40%)	63.90% (+0.03%)	85.19% (+0.02%)	14.25% (+0.22%)	8.91% (+0.38%)
Inceptionv3 (Version III)	78.10% (+0.20%)	93.98% (+0.28%)	63.79% (-0.08%)	85.12% (-0.05%)	14.31% (+0.28%)	8.86% (+0.33%)
Inceptionv3 (epoch 5 + Version III)	78.19% (+0.29%)	94.13% (+0.43%	63.87% (-)	85.18% (+0.01%)	14.32% (+0.29%)	8.92% (+0.39)

B. Version IV

Table IV demonstrates the result of contrast experiments. The second row illustrates the performance of continued training model. Compared to the baseline, it is clear that both of top-1 accuracy and top-5 accuracy on validation set and imagenetv2 increase while the gap also increases. This finding is interesting because few people realized that the pretrained model is not the model that achieves that best performance on validation set in fact. The continued training will larger the domain gap. The third row demonstrates the result of baseline model adopted Version III training and the fourth row demonstrates the result of continued training model adopted Version III training either. The change of data is consistent of previous finding of version III training. The baseline model with version III training has increase on validation set only. The performance on imagenetv2 and domain gap decrease. The continued training model adopted the version III training has both increase on top-1 accuracy and top-5 accuracy on validation set. Also, the gap decreases.

C. Grad-CAM analysis and Type of Error Definitions

Fig.5. demonstrates Grad-CAM analysis on some wrong predicted images. Six examples are chosen shown in two columns. The left column demonstrates the Grad-CAM of ground truth object while the right column demonstrates the Grad-CAM of wrong prediction. The title of each sub-image includes the prediction confidence of wrong prediction. The images are sorted by the confidence in ascending order. For the images labelled with mailbox and wagon, it is clear the Grad-CAM of wrong prediction is very close to the Grad-CAM of true image. For the images labeled with banjo, ski, and mask, the location of Grad-CAM is totally different with the ground truths. For example, the label is banjo while the model focuses on the bookshelf. Thus, it gives the prediction as library. Or the ground truth is ski, but model focuses on the hand of the women. Since it detects human face and gloves on the face, it predicts as mask. In such as case, model gives wrong prediction because of wrong location of attention. The ground truth of the last image is riffle and model predict as bearskin. It is obvious that the Grad-CAM of the ground truth is also wrong. The model fails to locate the object of interest. In this case, the intervene training using the purified examples definitely has negative impact

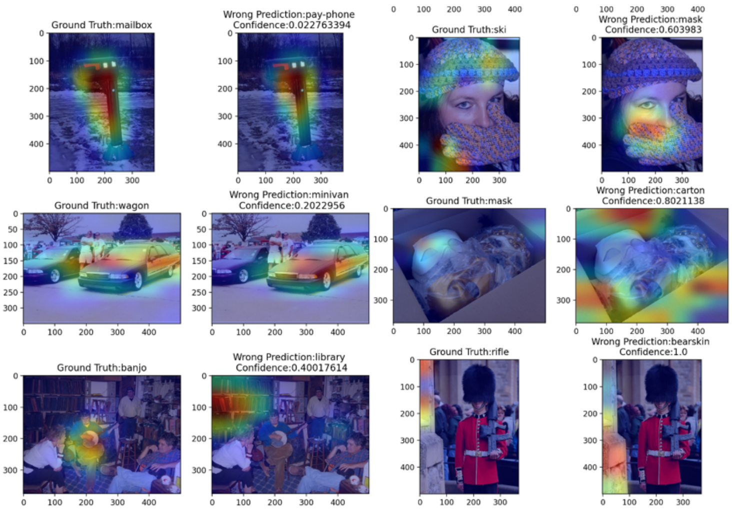

Fig .5. Grad-CAM analysis on some wrong predicted images

on the model. The performance of model will be reduced. Followed by the visualization, we proposed three types of errors: Type I Error, True Ground Truth Attention, Wrong Prediction Attention. The model gives wrong prediction purely due to the wrong location of object. It counts the largest proportion in the wrong prediction set, around 40% ~ 50%. The Type II Error is Wrong Ground Truth Attention, True Prediction Attention. In this case, the examples are labelled with wrong class. The model could give the true prediction of the images but failed to consistent with the ground truth. It happens in the validation set but are few. The type III Error is Both Ground Truth Attention and Wrong Prediction Attention are True. The model gives wrong prediction because of the class co-occurrence or class similar. Class co-occurrence refers to there are multiple objects in the same examples. The model can give true prediction on one object which is not labeled. The class similar refers to some classes are too similar to distinguish. This type of error counts second largest of proportion in the wrong prediction set. The last type of error is Both Ground Truth Attention and Wrong Prediction Attention are Wrong. In this case, it is impossible for model to make correct prediction. We should mainly Focus on Type I & Type III Error.

D. Super-class classification

Before starting the super-class classification, we use the approach described in section III.C to find the important features of car super-class. Fig. 6 demonstrates the heatmap of contribution of car superclass. It is clear that some features are indeed more important than others when an image is about to classified as car or non-car.

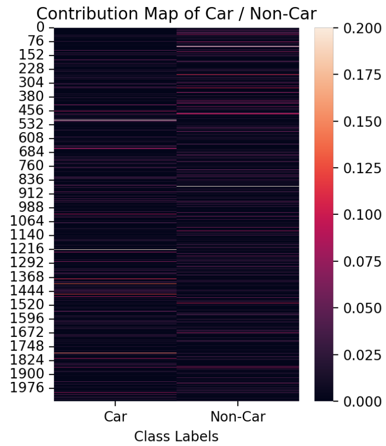

Fig .6. Heatmap of important features of super-class "car"

Table V demonstrates the result of super-class classification. Since the class distribution in the validation set is highly unbalance, robust metrics such as Area Under Curve (AUC) and precision are used to measure the performance. The result on validation set is better than imagenetv2, however, the AUC is already close to 98%, which means that even the class distribution is highly unbalance, the model can still make good classification when the intervene training is used.

TABLE V. PERFORMANCE OF SUPER-CLASS CLASSIFICATION

	Validation I	Validation II
Accuracy	0.9801	0.9770
AUC	0.9951	0.9943
Precision	0.9801	0.9770
Recall	0.9801	0.9770

VI. CONCLUSION

In this independent project, a X-AI guided two-state training method is proposed to solve the domain generation problem. The first phase is regular training, and the second phase is Grad-CAM based intervene training. Before the experiments, the definition of the contribution formula again each pair of feature unit and class is given, and a heatmap is drawn to show the average contribution of each pair of feature unit and class. The visualization demonstrates that when classifying a given image, indeed some feature units are more important than others.

Guided by this finding, we hope the intervene training could reduce the domain gap. The basic idea is to force the model to learn the correct correlation between relationship and class labels. The Grad-CAM is used to locate the object of interest in the example. When the intervene training is conducting, the backbone is frozen. Only the prediction layer is active. Several Version of training methods are tested in the experiment. Furthermore, a super-class binary classification task is used to validate the model could learn to distinguish a group of classes by frozen the unimportant feature units.

However, from the result table, the intervene training only has slight effect on reducing the accuracy gap between validation set and imagenetv2. One of the possible reasons is that the continuing training of baseline model will strength the gap of two validation set, which is conflict with our purpose. By adopting the version III training on continued training model, the result is better than pure adopting the intervene training on baseline model. In the end, the model can distinguish the super-class very successfully by keeping the unimportant feature units frozen. The AUC scores achieve approximate to 0.98 even the class distribution is extremely imbalance. It shows that the general idea of our project is valid and valuable.

REFERENCES

[1] Nevin L. Z , Weiyan X., Zhi L, et al. Example Perplexity.

[2] Deng J, Dong W, Socher R, et al. Imagenet: A large-scale hierarchical image database[C]//2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009: 248-255.

[3] Recht, B., Roelofs, R., Schmidt, L., & Shankar, V. (2019, May). Do imagenet classifiers generalize to imagenet?. In International Conference on Machine Learning (pp. 5389-5400). PMLR.

[4] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision (pp. 618-626).

APPENDIX A: MINUTES

Minutes of 1st Meeting

Date: 2022-02-08
Time: 5:00 pm
Place: Zoom
Attending: QIU Yaowen, Anthony Chiu, and all students applied for Prof. Nevin L.Zhang's project
TA: XIE Weiyan
Supervisor: Prof. Nevin L.Zhang

1. Approval of minutes
   This is first formal group meeting, so there were no minutes to approve
2. Discussion Items
   - Administrative matters are announced
   - Time schedule is announced
   - Basic introduction of project
   - Q & A
3. Goals for next meeting
   Each student should give a clear description of project objective and scope

Minutes of 2nd Meeting

Date: 2022-03-01
Time: 3:40 pm
Place: Zoom
Attending: QIU Yaowen, Anthony Chiu
TA: XIE Weiyan
Supervisor: Prof. Nevin L.Zhang

4. Approval of minutes
   The minutes of the last meeting were approved without amendment
5. Discussion Items
   - Each student gave a clear description of project objective and scope
   - The supervisor gave advice to each student and answer student's questions
6. Goals for next meeting
   Each student should give a progress report of meeting

Minutes of 3rd Meeting

Date: 2022-04-12
Time: 3:40 pm
Place: Zoom
Attending: QIU Yaowen, Anthony Chiu
TA: XIE Weiyan
Supervisor: Prof. Nevin L.Zhang

7. Approval of minutes
   The minutes of the last meeting were approved without amendment
8. Discussion Items
   - Each student reported the progress of project
   - The supervisor asked questions about the detail of the progress and gave advice
9. Goals for next meeting
   Each student should give a final project report

Minutes of 4th Meeting

Date: 2022-05-13
Time: 2:00 pm
Place: Zoom

Attending: QIU Yaowen, Anthony Chiu
TA: XIE Weiyan
Supervisor: Prof. Nevin L.Zhang

10. Approval of minutes
    The minutes of the last meeting were approved without amendment
11. Discussion Items
    - Each student gave final report of the whole project
    - The supervisor asked questions in detail
12. Goals for next meeting
    N/A