Identification of Novel Classes in Object Class Recognition

Alon Zweig, Dagan Eshar, and Daphna Weinshall

School Computer Science & Engineering, Hebrew University of Jerusalem, Jerusalem, Israel

Abstract. For novel class identification we propose to rely on the natural hierarchy of object classes, using a new approach to detect incongruent events. Here detection is based on the discrepancy between the responses of two different classifiers trained at different levels of generality: novelty is detected when the general level classifier accepts, and the specific level classifier rejects. Thus our approach is arguably more robust than traditional approaches to novelty detection, and more amendable to effective information transfer between known and new classes. We present an algorithmic implementation of this approach, show experimental results of its performance, analyze the effect of the underlying hierarchy on the task and show the benefit of using discriminative information for the training of the specific level classifier.

1 Introduction

A number of different methods have been developed to detect and recognize object classes, showing good results when trained on a wide range of publicly available datasets (see e.g. [1,3]). These algorithms are trained to recognize images of objects from a known class. Ideally, when an object from a new class appears, all existing models should reject it; this is the only indication that a new object class has been seen. However, similar indication will be obtained when no object exists in the image. And what with outliers and noisy images, low recognition likelihood by all existing models may be obtained even when a known object is seen in the image. This is one of the fundamental difficulties with the prevailing novelty detection paradigm - negative evidence alone gives rather non-specific evidence, and may be the result of many unrelated reasons.

In this paper we are interested in novel object class identification - how to detect an image from a novel class of objects? Unlike the traditional approach to novelty detection (see [4,5] for recent reviews), we would like to utilize the natural hierarchy of objects, and develop a more selective constructive approach to novel class identification. Our proposed algorithm uses a recently proposed approach to novelty detection, based on the detection of incongruencies [6]. The basic observation is that, while a new object should be correctly rejected by all existing models, it can be recognized at some more abstract level of description. Relying on positive identification at more abstract levels of representation allows for subsequent modeling using related classes, as was done e.g. in [7], where the representation of a new class was built from a learnt combination of classifiers at different levels of generality.

Specifically, in Section 2 we describe our approach to novel class identification. We consider a multi-class scenario, where several classes are already known; a sample from

an unknown class is identified based on the discrepancy between two classifiers, where one accepts the new sample and the second rejects it. The two classifiers are hierarchically related: the accepting classifier fits a general object class description, and the rejecting classifier fits the more specific level describing the known object classes. We use this property, that detection is based on rejection by only a small number of related classes, to develop a discriminative approach to the problem - the specific classifiers are trained to distinguish between the small group of related sub-classes in particular.

Our approach is general in the sense that it does not depend on the specifics of the underlying object class recognition algorithm. To investigate this point, we used in our implementation (and experiments) two very different publicly available object recognition methods [1,3]. Due to space limitations we present results only for [1], applying our approach when using the model of [3] showed similar results.

We tested our algorithms experimentally (see Section 3) on two sets of objects: a facial data set where the problem is reduced to face verification , and the set of motorbikes from the Caltech256 bench-mark dataset. We found that discriminative methods, which capture distinctions between the related known sub-classes, perform significantly better than generative methods. We also demonstrate in our experiments the importance of modeling the hierarchical relations as tightly as possible.

2 Our Approach & Algorithm

In order to identify novel classes, we detect discrepancies between two levels of classifiers that are hierarchically related. The first level consists of a single 'general category' classifier, which is trained to recognize objects from any of the known sub-classes, see Section 2.1. The second level is based on a set of classifiers, each trained to make more specific distinctions and classify objects from a small group of related known sub-classes. Using these specific classifiers, we build a single classifier which recognizes a new sample as either belonging to one of the set of known sub-classes or not, see Section 2.2.

We look for a discrepancy between the inference made by the two final classifiers, from the two different levels, where the general classifier accepts and the specific classifier rejects a new sample. This indicates that the new sample belongs to the general category but not to any specific sub-class; in other words, it is a novel class.

This algorithm is described a bit more formally in Algorithm 1, with further details given in the following subsections.

2.1 General Category Level Classifier

In order to learn the general category level classifier, we consider a small set of related known classes as being instances of a single (higher level, or more abstract) class. In accordance, all the examples from the known classes are regarded as the positive set of training examples for the more abstract class. For the negative set of examples we use either clutter or different unrelated objects (none of which is from the known siblings). As we shall see in Section 3, this general classifier demonstrates high acceptance rates when tested on the novel sub-classes.

Algorithm 1 Unknown Class Identification

Input :

$\mathbf{x}$ test image
C_G general level classifier
C_j specific level classifiers, j = 1..|$\#of$known sub-classes|
$V^c_{C_i}$ average confidence of train or validation examples classified correctly as C_i
$V^w_{C_i}$ average confidence of train or validation examples classified wrongly as C_i (zero if there are none).

1. Classify $\mathbf{x}$ using C_G
2. if accept
   Classify $\mathbf{x}$ using all C_j classifiers and obtain a set of confidence values $V_{C_j}(\mathbf{x})$
   Let $i = \arg\max_j V_{C_j}(\mathbf{x})$
   Define $S(\mathbf{x}) = (V_{C_i}(\mathbf{x}) - V^w_{C_i})/(V^c_{C_i} - V^w_{C_i})$
   (a) if $S(\mathbf{x}) > 0.5$
   label $\mathbf{x}$ as belonging to a known class
   (b) else label $\mathbf{x}$ as belonging to a novel (unknown) class
3. else label $\mathbf{x}$ as a background image

2.2 Specific Category Level Classifier

At the specific level classifier the problem is essentially reduced to the standard novelty detection task of deciding whether a new sample belongs to any of the known classes or to an unknown class. However, the situation is somewhat unique and we take advantage of this: while there are multiple known classes, their number is bounded by the degree of the hierarchical tree (they must all be sub-classes of a single abstract object). This suggests that a discriminative approach could be rather effective.

The training procedure of the specific level classifier is summarized in Algorithm 2 with details provided in subsequent subsections, while the classification process is described in step 2 of Algorithm 1 above.

Algorithm 2 Train Known Vs. Unknown Specific Class Classier

1. For each specific class, build a discriminative classifier with: *positive examples:* all images from the specific class. *negative examples:* images from all sibling classes.
2. Compute the Normalized Confidence function.
3. Choose a classification threshold for novel classes.

Step 1: Discriminative Multi-Class Classification To solve the multi-class classification problem in Step 1 of the algorithm, we train a discriminative object class classifier for each of the known specific classes and classify a new sample according to the most likely classification (max decision). We incorporate the discriminative information in

the training phase of each of the single known classes-specific classifiers. Specifically, each single class classifier is trained using all images from its siblings (other known classes under the same general level class) as negative examples. Thus the specific level object model learnt for each known class is optimized to separate the class from its siblings.

Step 2: Normalized Confidence Score For a new sample $\mathbf{x}$ which is classified as C_i, we obtain an estimate for classification confidence $V_{C_i}(\mathbf{x})$, the output of the learnt classifier. After the max operation in the Step 1, this value reflects the classification confidence of the multi-class classifier. Given this estimate, we would like to derive a more accurate measure of confidence as to whether or not the classified sample belongs to the group of known sub-classes.

To do this, we define a normalized score function, which normalizes the confidence estimate $V_{C_i}(\mathbf{x})$ relative to the confidence estimates of correct classifications and wrong classifications for the specific-class classifier, as measured during training or validation. Specifically, let $V^c_{C_i}$ denote the average confidence of train or validation examples classified correctly as C_i, and let $V^w_{C_i}$ denote the average confidence of train or validation examples from all other sub-classes classified wrongly as belonging to class C_i. The normalized score $S(\mathbf{x})$ of $\mathbf{x}$ is calculated as follows:

$$S(\mathbf{x}) = \frac{(V_{C_i}(\mathbf{x}) - V^w_{C_i})}{(V^c_{C_i} - V^w_{C_i})} \tag{1}$$

If the classes can be well separated during training, that is $V^c_{C_i} >> V^w_{C_i}$ and both groups have low variance, the normalized score provides a reliable confidence measure for the multi-class classification.

Step 3: Choosing a threshold Unlike the typical discriminative learning scenario, where positive and negative examples are given during training, in the case of novelty detection no actual negative examples are known during training. In particular, when the specific-object classifiers are trained, they are given no example of the new sub-class whose detection is the goal of the whole process. Thus it becomes advantageous to set rather conservative limits on the learnt classifiers, more so than indicated by the train set. In other words, in order to classify a new sample as known (a positive example for the final classifier), we will look for higher confidence in the classification. This is done by setting the threshold of the normalized confidence measure to reject more than originally intended during training. Since the normalized confidence measure lies in the range [0..1], we set the threshold in our experiments to 0.5.

3 Experiments

3.1 Datasets & Method

We used two different hierarchies in our experiments. In the first hierarchy, the general parent category level is the 'Motorbikes' . 22 object classes, taken from [2], were

added, in order to serve together with the original data set as the pool of object classes used both for the unseen-objects results described in section 3.3 and for the random grouping described in Section 3.5. In the second hierarchy, the general parent category level is the 'Face' level, while the more specific offspring levels are faces of six different individuals .

All experiments were repeated at least 25 times with different random sampling of test and train examples. We used 39 images for the training of each specific level class in the 'Motorbikes' hierarchy, and 15 images in the 'Faces' hierarchy. For each dataset with n classes, n conditions are simulated, leaving each of the classes out as the unknown (novel) class.

3.2 Basic Results

Figure 1 shows classification results for the discriminative approach described in Section 2. These results show the classification rates for the different types of test samples: *Known* - samples from all known classes during the training phase; *Unknown* - samples from the unknown (novel) class which belongs to the same General level as the Known classes but has been left out during training; *Background* - samples not belonging to the general level which were used as negative examples during the General level classifier training phase; and *Unseen* - samples of objects from classes not seen during the training phase, neither as positive nor as negative examples. The three possible types of classification are: *Known* - samples classified as belonging to one of the known classes; *Unknown* - samples classified as belonging to the unknown class; and *Background* - samples rejected by the General level classifier.

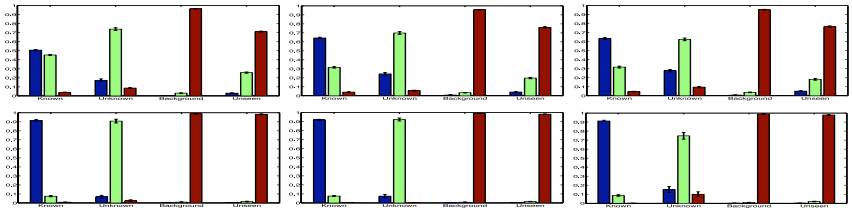


Fig. 1. Classification ratios for 4 groups of samples: Known Classes, Unknown Class, Background and sample of unseen classes. Ratios corresponding to the three possible classification rates are shown: left bar (blue) shows the known classification rate, middle bar (green) shows the unknown classification rate, and right bar (red) shows the background classification rate (rejection by the general level classifier). The top row plots correspond to the Motorbikes general level class, where the Cross (left), Sport (middle) and Road Motorbikes (right) classes are each left out as the unknown class. The bottom row plots are representative plots of the Faces general level class, where KA (left), KL (middle) and KR (right) are each left out as the unknown class.

The results in Fig. 1 show the desired effects: each set of samples - Known, Unknown and Background, has the highest rate of correct classification as its own category. As desired, we also see similar recognition rates (or high acceptance rates) of the Known and Unknown classes by the general level classifier, indicating that both are

regarded as similarly belonging to the same general level. Finally, samples from the Unseen set are rejected correctly by the general level classifier.

3.3 Discriminative Specific Classifiers Improve Performance

We checked the importance of using a discriminative approach by comparing our approach for building discriminative specific-level classifiers to non-discriminative approaches. In all variants the general level classifier remains the same.

We varied the amount of discriminative information used when building the specific level classifiers, by choosing different sets of examples as the negative training set: 1) *1vsSiblings - Exploiting knowledge of sibling relations*, the most discriminative variant, where all train samples of the known sibling classes are used as the negative set when training each specific known class classifier. 2) *1vsBck - No knowledge of siblings relations*, a less discriminative variant, where the negative set of examples is similar to the one used when training the general level classifier.

Applying these different variants when training models of [1] results in entirely different object models for which known vs. unknown classification results depict different ROC curves, as shown in Fig. 2. Specifically, when comparing the *1vsSiblings* to *1vsBck* curves in Fig. 2, we see that for all different choices of classes left out as the unknown - the corresponding ROC curve of the *1vsSiblings* method shows much better discrimination of the known and unknown classes. This demonstrates that discriminative training with the sibling classes as negative samples significantly enhances performance.

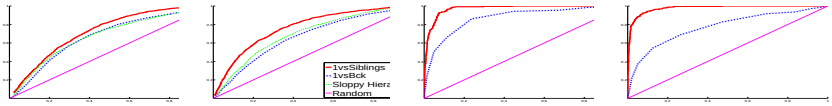


Fig. 2. ROC curves showing True-Unknown classification rate on the Y-axis vs. False-Unknown Classification rate on the X-axis. We only plot examples accepted by the General level classifier. *1vsSiblings* denotes the most discriminative training protocol, where specific class object models are learnt using the known siblings as the negative set. *1vsBck* denotes the less discriminative training protocol where the set of negative samples is the same as in the training of the General level classifier. Sloopy-Hierarchy denotes the case where the hierarchy was built using the procedure described in Section 3.5. Each plot shows results for a different class left out as the unknown, from left to right and top to bottom respectively: 'Cross-Motorbikes' , 'Sport-Motorbikes', 'KA' Face and 'KL' Face. We only show two representative cases for each dataset, as the remaining cases look very similar.

3.4 Novel Class Detector is Specific

To test the validity of our novel class detection algorithm, we need to verify that it does not mistakenly detect low quality images, or totally unrelated novel classes, as novel sub-classes. Thus we looked at two types of miss-classifications:

First, we tested our algorithm on samples of objects from classes which are not related to the general class and had not been shown during the training phase. These samples are denoted *unseen*, for which any classification other than rejection by the general level classifier is false. We expect most of these objects to be rejected, and expect the rate of false classification of unrelated classes as unknown classes to be similar to the rate of false classification as known classes. As Fig. 1 shows, by far most unseen samples are correctly rejected by the general level classifier. For the Faces data set we see that in cases of miss-classification there is no tendency to prefer the unknown class, but this is not the case with the Motorbikes data set; thus our expectation is only partially realized. Still, most of the false classification can be explained by errors inherited from the embedded classifiers.

Second, to test the recognition of low quality images, we took images of objects from known classes and added increasing amounts of Gaussian white noise to the images. As can be seen in Fig. 1, almost all the background images are rejected correctly by the general level classifier, and the addition of noise maintains this level of rejection. On the other hand the fraction of known objects classified correctly decreases as we increase the noise.

In Fig. 3 we examine the pattern of change in the misclassification of samples from the known class with increasing levels of noise - whether a larger fraction is misclassified as an unknown object class or as background. Specifically, we show the ratio (FU-FB)/(FU+FB), where FU denotes false classification as unknown class, and FB denotes false classification as background. The higher this ratio is, the higher the ratio of unknown class misclassifications to background misclassifications is. An increase in the false identification of low quality noisy images as the unknown class should correspond to an increase in this expression as the noise increases. In fact, in Fig. 3 we see the opposite - this expressions decreases with noise. Thus, at least as it concerns low quality images due to Gaussian noise, our model does not identify these images as coming from novel classes.

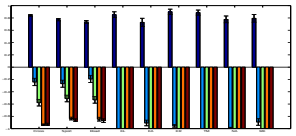


Fig. 3. This figure shows the effect of noise on the rate of false classification of samples from known classes. Each bar shows the average over all experiments of (FU-FB)/(FU+FB). Results are shown for both the Motorbikes and Faces datasets, and each group of bars shows results for a different class left out as the unknown. In each group of bars, the bars correspond to increasing levels of noise, from the left-most bar with no noise to the fifth right-most bar with the most noise.

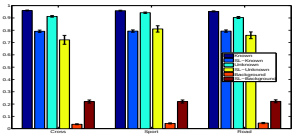


Fig. 4. General level classifier acceptance rates, with Strict and sloppy hierarchies. Six bars show, from left to right respectively: Strict hierarchy known classes ('Known'), Sloppy hierarchy ('SL-Known'), Strict hierarchy unknown class ('Unknown'), Sloppy hierarchy ('SL-Unknown'), Strict hierarchy background ('Background'). Sloppy hierarchy ('SL-Background'). Results are shown for the cases where the Cross, Sport or Road-Motorbikes are left as the unknown class, from left to right respectively

3.5 Sloppy Hierarchical Relation

In order to explore the significance of hierarchy in our proposed scheme, we followed the procedure as described in section 2 using a ”sloppy” hierarchy, thus comparing results with “strict” hierarchy to results with “sloppier” hierarchy. We only changed one thing - instead of using a group of strict hierarchically related sub-classes, we collected a random group of sub-classes; all other steps remained unchanged. The acceptance rate by the general level classifier using the strictly built hierarchy vs. sloppy hierarchy is shown in Fig. 4. Results are shown for objects belonging to the known classes, unknown-class and background images. Correct unknown classification vs. false unknown classification of samples that were accepted by the general level classifier are shown in Fig. 2 for both the strict and sloppy hierarchy.

As can be seen in Fig. 4, the general level classifier which is learnt for the sloppy hierarchy is less strict in the sense that more background images are falsely accepted and more known and unknown images are falsely rejected by the general level classifier. We also see from Fig. 2 that the distinction between known classes and an unknown class is significantly improved with the strict hierarchy as compared to the sloppy hierarchy. Combining both the general level classifier and the specific level classifier, clearly Algorithm 1 for the identification of unknown classes performs better when given access to a strict hierarchy, as compared to some sloppier hierarchy.

4 Summary

We address the problem of novel object class recognition - how to know when confronted with the image of an object from a class we have never seen before. We exploit existing hierarchical relations among the known classes, and propose a hierarchical discriminative algorithm which detects novelty based on the disagreement between two classifier: some general level classifier accepts the new image, while specific classifiers reject it. We analyze the properties of the algorithm, showing the importance of modeling the hierarchical relations as strictly as possible, and the importance of using a discriminative approach when training the specific level classifier.

References

1. A. Bar-Hillel, T. Hertz, and D. Weinshall. Efficient learning of relational object class models. *ICCV*, 2005. 1, 2, 6
2. G. Griffin, A. Holub, and P. Perona. Caltech-256 object category dataset. Technical Report UCB/CSD-04-1366, California Institute of Technology, 2007. 4
3. B. Leibe, A. Leonardis, and B. Schiele. Robust object detection with interleaved categorization and segmentation. *IJCV*, 77(1):259–289, 2008. 1, 2
4. M. Markou and S. Singh. Novelty detection: a review-part 1: statistical approaches. *Signal Processing*, 83(12):2481–2497, 2003. 1
5. M. Markou and S. Singh. Novelty detection: a review-part 2: neural network based approaches. *Signal Processing*, 83(12):2499–2521, 2003. 1
6. D. Weinshall, H. Hermansky, A. Zweig, J. Luo, H. Jimison, F. O, and M. Pavel. Beyond Novelty Detection: Incongruent Events, when General and Specific Classifiers Disagree. *NIPS*, 2008. 1
7. A. Zweig and D. Weinshall. Exploiting Object Hierarchy: Combining Models from Different Category Levels. *ICCV*, 2007. 1