

# Supplementary Materials

## Multi-Label Learning with Block Diagonal Labels

Anonymous Authors

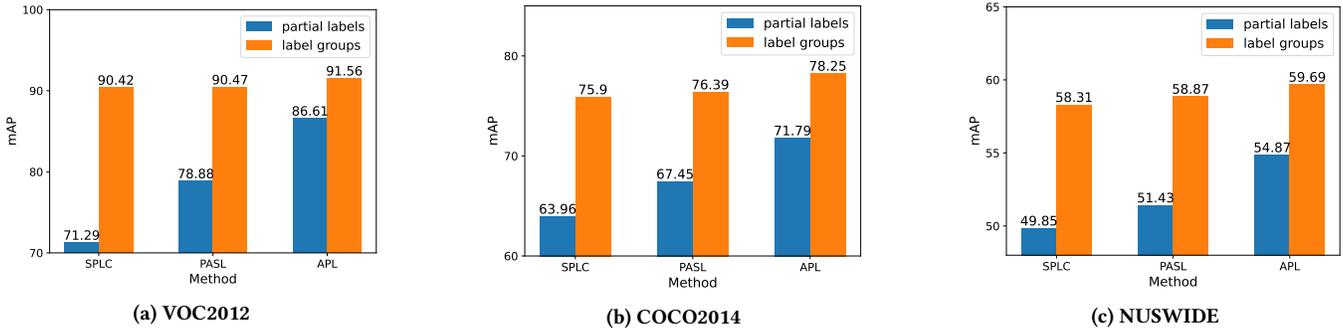


Figure 1: Experimental results of the state-of-the-art methods in *block diagonal labels* and *partial labels*. The experiments in *partial labels* are marked in blue. The experiments in *block diagonal labels* are marked in orange.

Table 1: Comparison of the proportion of annotated labels between *block diagonal labels* and *partial labels*.  $\#P_{Pos+Neg}$  indicates the proportion of all annotated labels.  $\#P_{Pos}$  indicates the proportion of positive labels.

|          | <i>partial labels</i> |             | <i>block diagonal labels</i> |             |
|----------|-----------------------|-------------|------------------------------|-------------|
|          | $\#P_{Pos+Neg}$       | $\#P_{Pos}$ | $\#P_{Pos+Neg}$              | $\#P_{Pos}$ |
| VOC2012  | 10.98%                | 0.80%       | 10.59%                       | 5.25%       |
| COCO2014 | 9.99%                 | 0.37%       | 9.31%                        | 1.64%       |
| NUSWIDE  | 15.99%                | 0.48%       | 15.60%                       | 2.01%       |

### 1 COMPARISON BETWEEN BLOCK DIAGONAL LABELS AND PARTIAL LABELS

To evaluate the superiority of *block diagonal labels*, we compare the experiments of *block diagonal labels* with those of *partial labels*. As illustrated in Table 1, the number of annotated labels is almost the same between *block diagonal labels* and *partial labels*. We generate the *block diagonal labels* datasets by randomly selecting a block of labels. For comparison between these two settings, we generate the *partial labels* datasets with a similar annotation proportion as the *block diagonal labels* ones. The positive proportions  $\#P_{Pos}$  of them differ greatly. For example,  $\#P_{Pos}$  of *block diagonal labels* is greater than that of *partial labels* by 4.45% on VOC2012. Randomly annotating a fixed percentage of labels in *partial labels*, which may generate images with all negative labels, will result in a very small number of positives.

Fig 1 shows results of the state-of-the-art methods in *block diagonal labels* and *partial labels*. The results in our *block diagonal labels* significantly surpass the results in *partial labels* on all benchmarks, which proves the superiority of our setting. Especially for SPLC, the performance gaps between these two settings are remarkable. Evident 19.13%, 11.94%, and 8.46% mAP gains are achieved on

VOC2012, COCO2014, and NUSWIDE, respectively. The recent approaches treat most unknown labels as negatives. This assumption is mostly correct due to the large number of true negatives. The effect of annotated negative labels is reduced, while the number of positive labels has a significant impact on the final performance. Therefore, the *block diagonal labels* setting not only reduces the annotation workload from both learning and annotating aspects but also increases the number of positives, leading to a notable enhancement in performance.

In addition, on all benchmarks, we find the performance gaps of our APL are relatively small compared to other methods. Our APL outperforms other methods by a large margin in all experiments. Such results prove the effectiveness and robustness of our APL.