## Abstract

In open-set environments, the introduction of unseen classes adds a third stage of data for incremental learning, alongside seen classes of old and new. Existing methods by binary optimization between old and new classes struggle to model the complex triadic relationships among old, new, and unseen classes. In this paper, we introduce the Triadic Elastic Structure Learning (TESL) framework for open-set incremental 3D object retrieval. Specifically, to overcome the Bi-Directional Catastrophic Forgetting (BDCF) problem in open-set incremental learning, we employ the Triadic Regularized Embedding (TRE) module. This module helps to globally preserve the triadic relationships among the three stages of classes, while simultaneously facilitating semantic-specific learning for new classes from a local perspective. To address the challenge of Bi-Directional Knowledge Transfer (BDKT), our method leverages the high-order correlations among objects from different stages, i.e., old classes, new classes, and unseen classes. This is achieved by constructing an elastic hypergraph based on incremental and global correlations. We construct four multi-modal datasets for this open-set incremental 3D object retrieval framework. The proposed method outperforms all other methods except for the upper-bound reference. The average forgetting is measured on the old seen classes after training on the new classes. A smaller area below the average forgetting curve indicates less forgetting. From the results in Table 1 and Figure 3, we can observe that our method outperforms all other methods except for the upper-bound reference. The better performance indicates that the proposed method can alleviate the imbalance between the triadic classes, while enhancing the representation abilities for both seen and unseen classes.

## Methodology

As shown in Figure 1, the architecture of TESL is composed of two modules: Triadic Regularized Embedding (TRE) and Elastic Structure Distillation (ESD). For objects from new classes, our framework takes the unified pre-trained features from the last task as input. The ESD module is designed to generate the embeddings of new classes while alleviating the imbalance between old, new, and unseen classes. Next, in the ESD stage, the elastic hypergraph structure is constructed based on the incremental and global correlations. Guided by this structure, hypergraph convolution and correlation distillation are adopted to leverage the high-order correlations among objects from different stages, textit{i.e.,} old classes, new classes, and unseen classes. Finally, the distilled embeddings are generated for retrieval in the current task and can also serve as pre-trained features for the next task.

## Experiments

We evaluate the open-set incremental 3D object retrieval results from TESL framework in the table. We provide the results of retrieval performance on both seen and unseen classes. Comparison results show that the proposed method outperforms all other methods except for the upper-bound reference on all datasets. In particular, on the unseen classes of OIMN and OIAB datasets, respectively. Our method achieves 0.7817/0.6302 average mAP (Avg.) with about 9.67%/10.19% improvements compared with the second-best method. We also provide the Precision-Recall Curve (PR-Curve) and average forgetting curves to evaluate the performance of incremental retrieval. For the PR-Curve, the larger area below the curve indicates better performance. The average forgetting is measured on the old seen classes after training on current new classes. A smaller area beneath the average forgetting curve indicates less forgetting. From the results in Figure 2 and Figure 3, we can observe that our method outperforms all other methods except for the upper-bound reference. The better performance indicates that by the TRE and ESD modules, the proposed method can alleviate the imbalance between the triadic classes, while enhancing the representation abilities for both seen and unseen classes.

## Table 1: Quantitative results and relative improvement of unseen classes on OIMN dataset

<table>
<thead>
<tr>
<th>Unseen</th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
<th>Task 5</th>
<th>Avg.</th>
<th>Δ (%)</th>
<th>Task 1</th>
<th>Task 2</th>
<th>Task 3</th>
<th>Task 4</th>
<th>Task 5</th>
<th>Avg.</th>
<th>Δ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper-bound Fine-tuning</td>
<td>0.7298</td>
<td>0.7047</td>
<td>0.6617</td>
<td>0.6141</td>
<td>0.5988</td>
<td>0.6618</td>
<td>-4.78</td>
<td>0.4540</td>
<td>0.4377</td>
<td>0.4161</td>
<td>0.3889</td>
<td>0.3817</td>
<td>0.4157</td>
<td>-3.62</td>
</tr>
<tr>
<td>T₀ + EWC</td>
<td>0.6144</td>
<td>0.5679</td>
<td>0.5400</td>
<td>0.5067</td>
<td>0.4760</td>
<td>0.5410</td>
<td>16.48</td>
<td>0.4024</td>
<td>0.3813</td>
<td>0.3633</td>
<td>0.3401</td>
<td>0.3211</td>
<td>0.3616</td>
<td>10.78</td>
</tr>
<tr>
<td>T₀ + NCE/EWC</td>
<td>0.6198</td>
<td>0.5737</td>
<td>0.5468</td>
<td>0.5129</td>
<td>0.4837</td>
<td>0.5474</td>
<td>15.12</td>
<td>0.4067</td>
<td>0.3854</td>
<td>0.3689</td>
<td>0.3450</td>
<td>0.3268</td>
<td>0.3666</td>
<td>9.29</td>
</tr>
<tr>
<td>T₀ + L2 Loss</td>
<td>0.6109</td>
<td>0.5865</td>
<td>0.5753</td>
<td>0.5599</td>
<td>0.5404</td>
<td>0.5746</td>
<td>9.67</td>
<td>0.3958</td>
<td>0.3878</td>
<td>0.3849</td>
<td>0.3753</td>
<td>0.3657</td>
<td>0.3819</td>
<td>4.91</td>
</tr>
<tr>
<td>T₀ + ALICE</td>
<td>0.6248</td>
<td>0.5925</td>
<td>0.5659</td>
<td>0.5506</td>
<td>0.5315</td>
<td>0.5731</td>
<td>9.96</td>
<td>0.4041</td>
<td>0.3894</td>
<td>0.3754</td>
<td>0.3660</td>
<td>0.3505</td>
<td>0.3771</td>
<td>6.25</td>
</tr>
<tr>
<td>T₀ + Ours</td>
<td>0.6117</td>
<td>0.5884</td>
<td>0.5727</td>
<td>0.5547</td>
<td>0.5422</td>
<td>0.5739</td>
<td>9.80</td>
<td>0.3952</td>
<td>0.3872</td>
<td>0.3851</td>
<td>0.3688</td>
<td>0.3615</td>
<td>0.3796</td>
<td>5.55</td>
</tr>
</tbody>
</table>

Figure 1: An overview of the proposed Triadic Elastic Structure Representation (TESR) framework for open-set incremental 3D object retrieval. Our framework comprises two modules: Triadic Regularized Embedding (TRE) and Elastic Structure Distillation (ESD), which are used for hierarchical knowledge retention and structure-aware distillation, respectively.

Figure 2: The average precision-recall curves of the unseen and seen classes on the OIMN and OIAB datasets, respectively.

Figure 3: The average forgetting evaluation of the unseen and seen classes on the OINT and OIAB datasets, respectively.