Modeling Extent-of-Texture Information for Ground Terrain Recognition

Shuvozit Ghose[a], Pinaki Nath Chowdhury[b], Partha Pratim Roy[c], Umapada Pal[b]

[a]Institute of Engineering and Management, Kolkata, India.
[b]Indian Statistical Institute, Kolkata, India.
[c]Indian Institute of Technology, Roorkee, India.

Abstract

Ground Terrain Recognition is a difficult task as the context information varies significantly over the regions of a ground terrain image. In this paper, we propose a novel approach towards ground-terrain recognition via modeling the Extent-of-Texture information to establish a balance between the order-less texture component and ordered-spatial information locally. At first, the proposed method uses a CNN backbone feature extractor network to capture meaningful information of ground terrain images, and model the extent of texture and shape information locally. Then, it encodes order-less texture information and ordered shape information in a patch-wise manner, and utilizes an intra-domain message passing mechanism to make every patch aware of each other for rich feature learning. Next, the model combines the extent of texture information with the encoded texture information and the extent of shape information with the encoded shape information patch-wise and then exploit Extent-of-texture (EoT) Guided Inter-domain Message passing module for sharing knowledge about the opposite domain to balance out the order-less texture information with ordered shape information. Finally, Bilinear model outputs a pairwise correlation between the order-less texture information and ordered shape information, and classifier classifies the ground terrain image efficiently. The experimental results indicate the superior performance of the proposed model over existing state-of-the-art techniques on DTD, MINC and GTOS-mobile datasets. The source code of the proposed system is publicly available at https://github.com/ShuvozitGhose/Ground-Terrain-EoT.

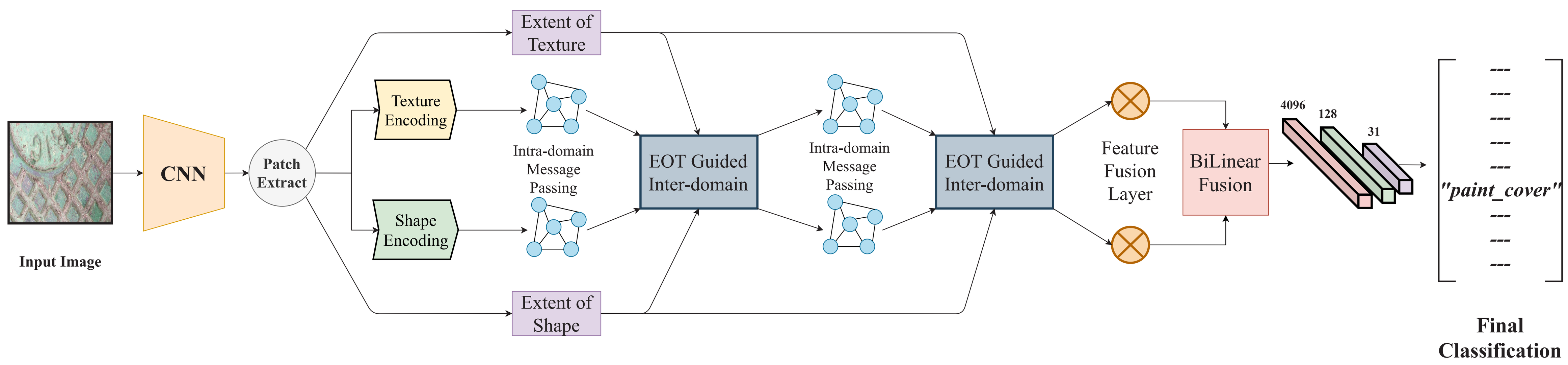
1. Problem

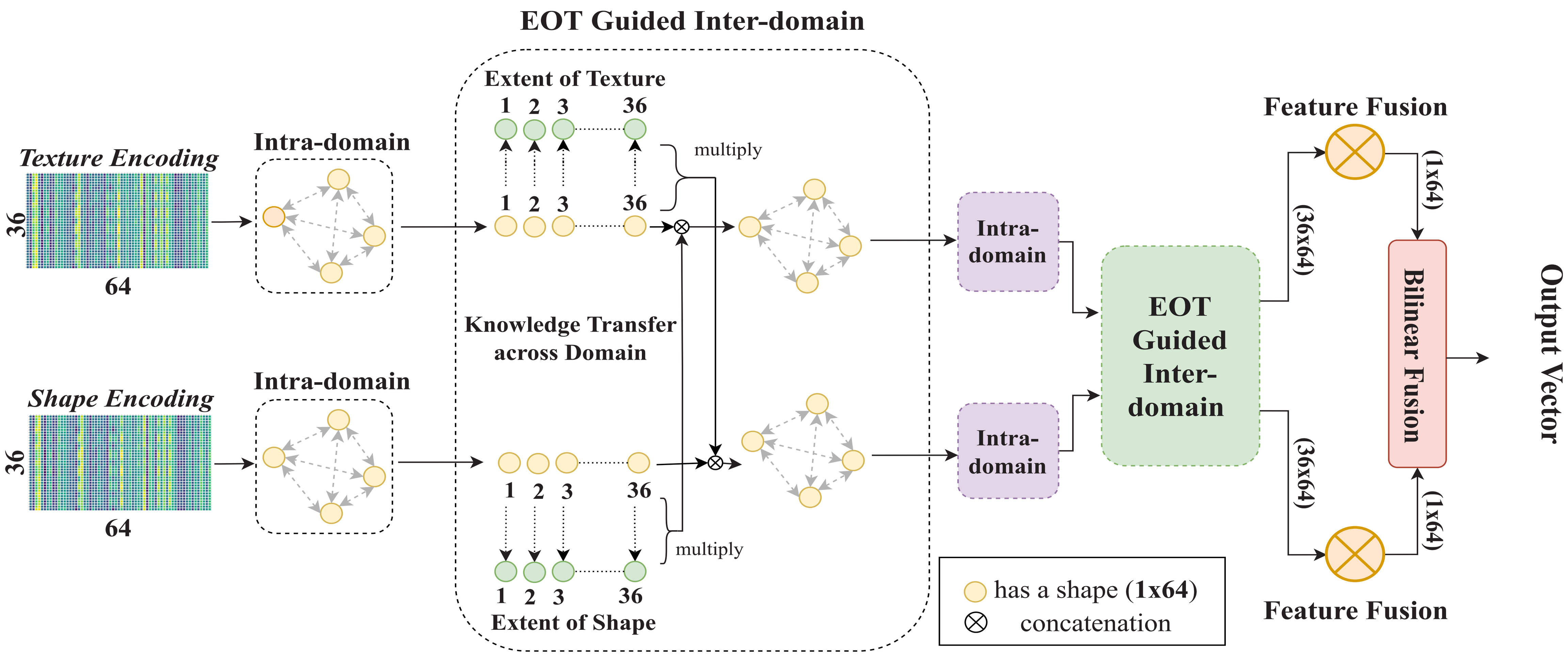
Ground Terrain Recognition is a difficult task due to various reasons. Firstly, real-world ground terrain images usually have highly complicated terrain surfaces and may not have any obvious feature or edge points. Moreover, the terrain surface may be as highly complex as the surface of Earth. Secondly, many class boundaries of the ground terrain images are ambiguous. For example, the class *"leaves"* is similar to *"grass"*, whereas the grass images contain a few leaves. Similarly, *"asphalt"* class is similar to *"stone-asphalt"* which is an aggregate mixture of stone and asphalt. Finally, the context information varies significantly over the regions of a ground terrain image, like some local regions possess significant texture information, while shape information is more dominant at some other parts.

2. Motivation

As most real-world ground terrain images show wide variations in texture and shape information at different local regions in an image, thus the classification of such realistic ground terrain images requires a more local level modeling of texture and shape information.

4. Proposed Methodology





Motivated by this observation, we propose a novel approach towards ground-terrain recognition via modeling the extent of texture information to establish a balance between the order-less texture and ordered-spatial information locally. We first use a CNN backbone feature extractor network to capture the meaningful information of the ground terrain image. Then, we model the extent of texture and shape information locally. we encode the order-less texture information and ordered-spatial information patch-wise. Next, we utilize Intra-domain Message passing mechanism to make every patch aware of each other for rich feature learning. Subsequently, we combine extent of texture information with encoded texture information, and the extent of shape information with the encoded shape information patch-wise to establish a more local level balance of the texture and shape information. Furthermore, we exploit Extent-of-Texture Guided Inter-domain Message passing module, for sharing knowledge about the other domain to balance out the order-less texture information with ordered shape information patch-wise. Next, we aggregate all the patches to get the global order-less texture and ordered shape information of the ground terrain image. Furthermore, a Bilinear model captures pairwise correlation between the order-less texture and ordered shape information. Finally, a classifier is used to classify the ground terrain image.

6. Results

Comparison of **Deep-TEN**, baseline **B1, B2, B3** and **B4** with the proposed methodology for single scale and multi scale training on GTOS-mobile [1] dataset using a pre-trained ResNet-18 module as the convolutional layer. Baseline B1 is similar to Deep Encoding Pooling Network (DEP) by Xue [1].

	Deep-TEN [2]	**B1 [1]**	**B2**	**B3**	**B4**	**Proposed Method**
Single Scale	74.22	76.07	77.81	78.55	78.93	**80.39**
Multi Scale	76.12	82.18	83.78	84.31	84.36	**85.71**

Comparing Our method with several state-of-the-art methods on Describable Textures Dataset (DTD) and Materials in Context Database (MINC)

Method	DTD [3]	MINC-2500 [5]
FV-CNN [4]	72.3	63.1
Deep-TEN [2]	69.6	80.4
DEP [1]	73.2	82.0
Proposed Method	**75.7**	**85.3**

7. Conclusion and Future Work

In this paper, we have proposed a novel approach towards ground-terrain recognition via modeling the extent of texture information to establish a balance between the order-less texture component and ordered-spatial information locally. The driving idea of our architecture is the modeling of context information locally. The proposed framework is simple and easy to implement. It is capable of detecting ground terrain in the real-world scenario. We demonstrate the effectiveness of our system by conducting experiments on publicly available ground terrain datasets.

References

[1] J. Xue et al. *Deep texture manifold for ground terrain recognition*, CVPR 2018.

[2] H. Zhang et al. *Deep TEN: Texture Encoding Network*, arXiv preprint arXiv:1612.02844 2016.

[3] M. Cimpoi, et al. *Describing textures in the wild*, CVPR 2014.

[4] M. Cimpoi et al. *Deep filter banks for texture recognition and segmentation*, CVPR 2015.

[5] S. Bell et al. *Material recognition in the wild with the materials in context database*, CVPR 2015.