International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

A Review on Cross Weather Traffic Scene Understanding Using Transfer Learning for Intelligent Transport System

Ms. Deepa Mane[1], Dr. Sandhya Arora[2], Mr. Sachin Shelke[3]
[1]Research Scholar , Comp dept, *deepabmane@gmail.com*
SKNCOE, Pune, India
[2]*Professor, Comp dept sandhya.arora@cumminscollege.in*
CCOEW, Pune, India
[3]Asst. Professor , IT dept, *sacheenshelke@gmail.com*
PICT, Pune, India

Abstract—

Intelligent transport systems (ITS) have revolutionized the transportation industry by integrating cutting-edge technologies to enhance road safety, reduce traffic congestion and optimize the transportation network. Scene understanding is a critical component of ITS that enables real-time decision-making by interpreting the environment's contextual information. However, achieving accurate scene understanding requires vast amounts of labeled data, which can be costly and time-consuming. It is quite challenging to Understand traffic scene captured from vehicle mounted cameras. In recent times, the combination of road scene-graph representations and graph learning techniques has demonstrated superior performance compared to cutting-edge deep learning methods across various tasks such as action classification, risk assessment, and collision prediction. It's a grueling problem due to large variations under different illumination conditions. Transfer learning is a promising approach to address this challenge. Transfer learning involves leveraging pre-trained deep learning models on large-scale datasets to develop efficient models for new tasks with limited data. In the context of ITS, transfer learning can enable accurate scene understanding with less data by reusing learned features from other domains.

This paper presents a comprehensive overview of the application of transfer learning for scene understanding in cross domain. It highlights the benefits of transfer learning for ITS and presents various transfer learning techniques used for scene understanding. This survey paper provides systematic review on cross domain outdoor scene understanding and transfer learning approaches from different perspective, presents information on current state of art and significant methods in choosing the right transfer learning model for specific scene understanding applications.

Keywords: Cross weather Scene understanding, Autonomous driving, deep learning, transfer learning, Traffic Scene.

I. INTRODUCTION

Scene Understanding is an essential requirement for various visual tasks, and it has received significant attention in the field of computer vision. Nonetheless, the task of aligning images captured in different environments, such as scenes under diverse weather or seasonal conditions, continues to pose a considerable challenge. Despite the notable progress made in learning-based semantic segmentation[1], specifically in accurately parsing road scenes in well-illuminated settings, effectively training a reliable model for segmenting road scenes under varying weather conditions remains extremely difficult, particularly when there is a lack of semantic labels in the training samples This is quite challenging especially under different environments like the outdoor scenes[3]. It may require Scaling up in CNN to achieve The New approaches focuses on enhancing the accuracy of deep convolutional neural networks (CNNs) by leveraging efficient neural network architectures[5]. This involves fine-tuning these networks to create a novel application specifically designed for recognizing objects and scenes in outdoor environments. The primary objective of this application is to excel in accurately identifying and understanding outdoor settings. Artificial intelligence and the computer vision community face a complex challenge in understanding and recognizing visual scenes. This problem involves two primary inputs such as images and videos[7]. Applications utilizing computer vision and AI agents have the potential to cover a wide range of areas, including autonomous robot navigation, autonomous vehicle navigation and image retrieval. Recent advancements in deep learning, specifically deep convolutional neural networks, coupled with the availability of extensive annotated datasets, have sparked a significant interest in addressing this problem. Image recognition and classification serve as fundamental tasks in understanding the content of images and visual data. Outdoor environments

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

present visually dynamic scenes with complex backgrounds, exhibiting both intra-class and inter-class variations[7]. In such settings, the utilization of artificial intelligence-based automated systems becomes crucial for solving complex computer vision problems. Automated systems play a vital role in enabling the successful and secure accomplishment of diverse tasks. They are particularly instrumental in addressing the challenge of scene recognition, which involves predicting the category of a scene based on an input image. This paper provides a comprehensive review of various scene recognition[8] applications that rely on deep convolutional neural networks. and transfer learning.

Since the 1970s, Intelligent Transportation Systems (ITS) have been continuously evolving and are considered the future direction of transportation systems. ITS integrates advanced technologies, including electronic sensor technologies, data transmission technologies, and intelligent control technologies, into the transportation infrastructure [34]. The primary goal of ITS is to enhance the services provided to drivers and passengers in transportation systems. In ITS, data can be collected from various sensors, smart cards, GPS. Effectively analysing and leveraging this seemingly disorganized data through accurate and efficient data analytics can significantly improve the quality of ITS services. As ITS continues to develop, the volume of data generated in the system has increased to the Petabytes of data. Traditional data processing systems proves to be inadequate to meet the demands of data analytics due to their inability to anticipate the rapid growth in data volume and complexity. This paper presents a comparative overview of the application of transfer learning[9] for scene understanding in cross domain. It highlights the benefits of transfer learning for ITS and presents various transfer learning techniques used for scene understanding. The paper also discusses the challenges and limitations of transfer learning for ITS and proposes potential solutions to overcome them. This survey paper provides systematic review on cross domain outdoor scene understanding and transfer learning approaches from different perspective, presents information on current state of art and significance of choosing the right transfer learning model for various scene understanding applications[10].

Figure1: Different weather condition Images of the traffic Scene[33].

II. BACKGROUND

An essential drawback of present machine learning techniques is their incapability to effectively leverage acquired knowledge from one task to aid in the learning process of a new task. Transfer learning is principally an enhancement in the feature learned in a new task through the knowledge transfer from a affiliated task that has formerly been learned. While utmost machine learning algorithms are designed in such way that single tasks is addressed, the development of algorithms that grease transfer literacy is ongoing topic of interest in the machine learning. Humans exhibit remarkable abilities in learning from limited examples and adapting to new situations by leveraging analogies[11]. To tackle this challenge, transfer learning has emerged as a field of study, aiming to understand the mechanisms of knowledge and data reuse between a source and target domain. Novel methods and algorithms in transfer learning focus on knowledge extraction[2], analogy formation, and human reasoning[12]. Experimental evaluations of transfer learning approaches involve the utilization of benchmark datasets as well as real-world scenarios encompassing both source and target domains.

Convolutional neural networks (CNNs) have demonstrated remarkable achievements in various challenging computer vision tasks, including image classification[13], object detection, and semantic segmentation. However, it is important to note that most studies and approaches in outdoor scene understanding primarily focus on images captured during daytime and under favourable conditions, characterized by adequate illumination[14] and distinct boundaries between objects. Conversely, there is a noticeable lack of exploration and research in handling images acquired under unfavourable conditions, such as nighttime[7] or foggy days, particularly in applications like scene parsing.

Cross-domain outdoor scene understanding refers to the task of recognizing and understanding scenes in outdoor environments across different domains, such as different weather conditions, lighting, or camera viewpoints. Transfer learning is a technique that allows models trained on one domain to be adapted to another domain. There have been several studies on cross-domain outdoor scene understanding using transfer learning, with various approaches such as deep neural networks, domain adaptation, and feature transfer[15]. These techniques improving the performance of outdoor scene recognition and understanding across different domains. Some challenges in this area include the lack of large-scale and diverse datasets for training and evaluation, as well as the need for efficient and effective transfer learning methods that can generalize well to unseen domains. Outdoor scene understanding is important technology that involves the development of algorithms and models for analysing and interpreting various elements in outdoor scenes, such as objects, surfaces, and their relationships. There are many

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

approaches to outdoor scene understanding, including traditional computer vision techniques, deep learning models, and hybrid approaches[16] that combine both. A comparative analysis of outdoor scene understanding would involve comparing the strengths and weaknesses of different approaches to this problem. Some factors to consider might include the accuracy of different algorithms in recognizing objects and their relationships, the computational efficiency of different techniques, the availability and quality of training data for different models, and the scalability of different methods to larger and more complex scenes. Overall, deep learning approaches have shown great promise in outdoor scene understanding, especially for tasks like object detection and segmentation[17]. However, traditional computer vision techniques can still be effective in certain scenarios, such as when the goal is to extract specific geometric features from a scene. Hybrid approaches that combine both deep learning and traditional techniques may be the most effective for achieving both accuracy and efficiency in outdoor scene understanding.

A. Challenges

Self-driving vehicles must respond quickly to their surroundings, but in the real world, they may encounter new and complex situations that can put the vehicle in difficult positions. This uncertainty increases the risk of incorrect decisions, potentially endangering passengers and others nearby. Therefore, the model used by self-driving cars [19]must be dynamic, capable of making real-time decisions while also being aware of its own confidence level, learning from new situations, and updating its parameters accordingly. However, the decision-making process for self-driving cars is often done using black-box neural models, which raises questions about how to explain and justify the car's actions. Additionally, accurately perceiving pedestrians and vehicles and understanding complex driving scenes[20] in adverse weather conditions remain significant challenges for autonomous vehicles., which is critical to ensure the safe operation of self-driving cars in urban areas.

This survey focuses on three primary research questions. Firstly, it investigates whether the existing datasets are capable of generalizing to scene understanding in complex cross environments[21][22]. Secondly, it examines the ability of current methods to effectively segment visual scenes with uncertain elements like fog and rain, as well as unstructured elements such as rough roads and non-smooth pathways for pedestrians. Finally, the survey evaluates whether the current methods have the potential to learn attentively from ongoing scenes and incorporate event-based scene understanding[23]. The research outcomes of this survey shed light on the strengths and weaknesses of existing methods and datasets for scene understanding in complex visual environments.

B. Contributions

This survey presents a noteworthy contribution to the Intelligent Transportation Systems (ITS) community as it thoroughly assesses the most recent advancements in scene understanding through diverse techniques. The survey offers the following key contributions:

1. A comprehensive introduction to scene understanding, which outlines the general pipeline and provides a detailed explanation of each step. This serves as a useful resource for newcomers to the field who need to acquire prior knowledge in all aspects of scene understanding for Autonomous Driving.
2. This survey provides an extensive examination and critical analysis of the prominent research papers and datasets that have emerged in the field of scene understanding during the last decade.
3. This study conducts a comprehensive performance analysis of the current state-of-the-art methods, taking into account their computational resource requirements and the platforms on which they are developed. The findings of this study hold particular significance for industrialists aiming to implement more effective scene understanding strategies. Some contributions also provide open-source implementations, which are leveraged in this review by executing, analysing, and comparing the resources consumed by each method.
4. Presents a logically derived set of future research guidelines based on the analyzed literature. These guidelines effectively identify open problems, challenges, and research opportunities within the domain of scene understanding.

III. SCENE UNDERSTANDING

Scene understanding involves inferring higher-level information from a scene, such as the relationships between objects and the activities that are taking place. This can be done using deep learning models, such as graph convolutional networks and spatio-temporal transformers[24]. Scene understanding in cross-weather traffic refers to the ability of an intelligent system, such as an autonomous vehicle or a traffic management system, to comprehend and make sense of the surrounding environment during varying weather conditions, particularly in the context of traffic scenarios. Cross-weather traffic presents unique challenges for scene understanding due to the impact of weather conditions on visibility, road surface conditions[26], and the behaviour of other road users. To achieve scene understanding in such scenarios, multiple sensing modalities and advanced algorithms are employed. Here are some key components of scene understanding in cross-weather traffic: Sensor Fusion: Multiple sensors, such as cameras, lidar, radar, and thermal

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

sensors, are used to gather information about the environment[25]. Sensor fusion techniques integrate data from these sensors to obtain a more comprehensive and accurate understanding of the scene. Advanced computer vision and machine learning algorithms are employed to process the sensor data and extract relevant information. These algorithms can detect and classify objects such as vehicles, pedestrians, cyclists, and road infrastructure, even in challenging weather conditions like rain, snow, fog, or low light[27]. Weather conditions are modelled and taken into account to adapt the perception algorithms accordingly. For instance, algorithms may adjust their parameters to account for reduced visibility or changes in the appearance of objects due to rain or fog. Object Tracking and Prediction: Once objects are detected, they need to be tracked over time to understand their motion patterns and intentions. Predictive algorithms[28] can estimate the future trajectories of objects to anticipate their behaviour and make appropriate driving decisions. Semantic Understanding: Scene understanding involves not only detecting and tracking objects but also interpreting the semantic meaning of the scene. This includes understanding traffic signs, road markings, traffic rules, and the intent of other road users, which can be challenging in adverse weather conditions. Based on the understanding of the scene, decision-making algorithms can generate plans and trajectories for the autonomous vehicle to navigate safely through the cross-weather traffic. These algorithms consider factors such as road conditions, traffic flow, and potential hazards. understanding systems often incorporate machine learning techniques to continuously learn and adapt to changing weather conditions and improve their performance over time. They can leverage large datasets to train models that generalize well to different weather scenarios. Researchers have proposed various approaches to tackle this problem, including feature adaptation, model adaptation, and domain adaptation. The existing literature suggests that transfer learning can effectively improve the performance of outdoor scene understanding across different domains, such as day to night, weather changes, and seasonal changes. However, there are still challenges that need to be addressed, such as the selection of appropriate source and target domains, the design of effective adaptation strategies[29], and the evaluation of the robustness of the learned models. Some of the popular datasets used in this context include SUN, Place365, and ADE20K.

the effectiveness of the proposed framework will be verified by extensive experiments carried out using suitable data.

Cross weather conditions in traffic scene understanding refers to the ability of computer vision systems to recognize and classify traffic scenes under different weather conditions. This is a challenging task because weather conditions such as rain,

fog, snow, and sun glare can significantly affect the appearance of traffic scenes, making it difficult for computer

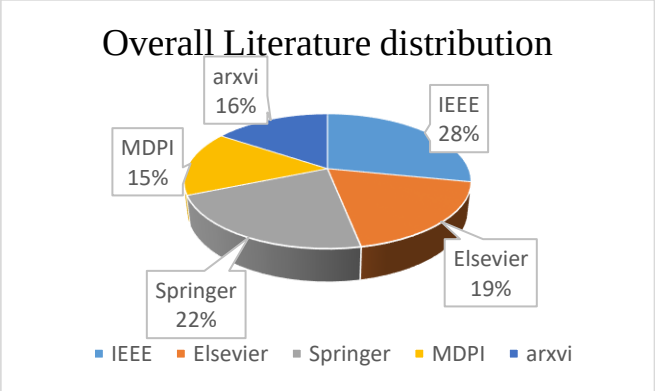

Figure2: The overall literature cross weather scene understanding methods.

vision systems to accurately identify objects and understand the overall scene.

To address this challenge, various techniques and models for cross-weather traffic scene understanding, have developed including:

1. Image enhancement: This involves improving the quality of images captured in adverse weather conditions by enhancing contrast, brightness, and colour saturation. Image enhancement techniques can help to reduce the impact of weather conditions on image quality and improve the accuracy of object detection and recognition.
2. Data augmentation: This involves artificially generating additional training data by modifying existing images to simulate different weather conditions. Data augmentation techniques[30] can help to improve the robustness of computer vision systems to variations in weather conditions.
3. Transfer learning: This involves using pretrained models trained on large datasets in other domains or weather conditions as a starting point for training new models on traffic scenes under different weather conditions. Transfer learning can help to reduce the amount of labeled data required for training and improve the generalization ability of computer vision systems.
4. Fusion of modalities: This involves combining information from multiple sensors, such as cameras and LiDAR, to improve the accuracy of object detection and recognition under adverse weather conditions. For example, LiDAR can provide depth information that can help to distinguish objects from background clutter in foggy conditions[4].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
***Article Received:** 26 August 2023 **Revised:** 20 October 2023 **Accepted:** 02 November 2023*

IV. TRANSFER LEARNING

In the era of hyperconnected Deep Learning, classifiers are now consuming labels and labeled data at unprecedented rates. Although computational power is no longer a concern in model training, the production of high-quality labeled data still requires significant time, cost, and human effort. Therefore, the scalability of DL becomes essential in numerous complex domains[14].

Machine learning and data mining methods have found extensive utilization in diverse real-world applications. Traditional machine learning approaches typically assume that training and testing data are sourced from the same domain This suggests that the input feature space and the distribution characteristics of the data are equivalent. However, this assumption doesn't always hold in real-world machine learning scenarios. In some cases, collecting training data can be challenging or expensive. Therefore, there is a growing need to develop high-performance learning models that can be trained using easily and the transfer learning solution this surveyed in this context are capable of handling big data scenarios without depending on the size of the data. Unlike traditional machine learning, which assumes that training and testing data share the same input feature space, transfer learning techniques can cope with differing input feature spaces and distribution characteristics.

If there is a disparity in the distribution of data between the training and testing datasets, the performance of a predictive model can deteriorate. In certain scenarios, obtaining training data that matches the feature space[2] and predicted data distribution characteristics of the test data can be difficult and expensive. Hence, there is a requirement to build a high-performance learning model for a target domain, which is trained using a related source domain. This is the underlying motivation for transfer learning. To address the scarcity of labeled data, researchers have utilized both Transfer Learning (TL) and Active Learning (AL) techniques [3]. TL involves leveraging pre-trained models from different domains, while AL focuses on selecting the most informative subset of instances for labeling. Particularly in the case of streaming data [4], it may not always be feasible to have labels for every instance due to factors such as high cost, rapid influx of data, or unavailability of human experts. Additionally, streaming data presents challenges like short data lifespan, inapplicability of batch methods, and issues related to concept and feature drift, as well as model switching, which can hinder the effectiveness of training. Deep networks are known for their ability to learn transferable features[21]. However, recent discoveries indicate that deep features must eventually transition from being general to specific as they propagate through the network. As a result, the transferability of features significantly decreases in higher layers, particularly when there is a growing discrepancy between the domains [5]. In simpler terms, the features in the deeper layers of a deep network heavily rely on domain-specific information.

Earlier Model Similar to Transfer Learning

1. Semi-Supervised Learning

Transfer learning aims to maximize the utilization of unlabeled data in the target task or domain. It bears similarity to semi-supervised learning, which operates within the conventional machine learning framework but assumes a limited number of labeled samples for training. Specifically, semi-supervised domain adaptation[6] can be seen as a form of semi-supervised learning in the presence of domain shift. Notably, various lessons and insights acquired from semi-supervised learning are equally relevant and applicable to the field of transfer learning.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

TABLE 1. COMPREHENSIVE ANALYSIS OF EXISTING LITERATURE ON CROSS DOMAIN SCENE UNDERSTANDING

Ref	Method	dataset	Remarks
[2]	A Dense correspondence-based transfer learning (DCTL)	DCTL	Constructing compact and effective representations via cross-domain metric learning and subspace alignment for cross-domain retrieval
[4]	Curriculum Model Adaptation (CMAda)	Foggy Zurich	Curriculum Model Adaptation exploits both our synthetic and our real data in a synergistic manner and significantly boosts performance on real fog without using any labeled real foggy image
[6]	Fully Convolutional Network (FCN) architecture	Cityscapes dataset	Semantic Segmentation of Urban Scenes using Convolutional Neural Networks
[7]	Unsupervised label transfer	Synthia	The encoder only extracts the common part from both domains but ignores some individual features of an input image
[9]	Topic Model	QMUL Junction dataset	scene alignment method is tractable and free from overfitting due to the consideration of the full distribution of activities in the scene and feeding the transformed topic-word vector matrix of the target domain scene into the aforementioned TextCNN for multilabel classification
[10]	Depth Transformer	KITTI Make3D	Estimate depth from a single RGB image
[12]	A benchmark dataset	1. CamVid dataset 2. Daimler urban segmentation dataset 3. street scenes dataset	Given one test image, we first need to find similar images as nearest neighbor set whose labels would be trans ferable to the test image. Then, the dense correspondences should be established between the test image and images in the nearest neighbor set.
[13]	Latent Generative Model with Intensity Consistency	BDD	Image encoding in the latent image manifold for cross-weather or cross-domain image alignment tasks.
[17]	Generative adverserial network	1. Philly-Commuting Road Scene (PRS) 2. Nordland Railroad Scene Dataset (NRS) 3. RobotCar Seasons Dataset (RCS)	The alignment problem is formulated as a constrained 2D flow optimization problem with latent encoding, which can be decoded into an intensity-constancy image on the image manifold.
[25]	Semi supervised semantic segmentation method	MIT67	This method can simultaneously optimize the standard supervised classification loss on labeled samples and the unsupervised consistency loss on unlabeled samples by using an integrated prediction technology

2. Multi-Task Learning

In transfer learning, our main objective is to achieve high performance specifically on the target task or domain. Conversely, multi-task learning aims to excel across all available tasks. By leveraging knowledge gained from learning related tasks, we can enhance performance on the target task. The primary distinction between transfer learning and multi-task learning lies in the assumption of labeled data for each task. Multi-task learning[18] typically assumes the availability of labeled data for each task, whereas transfer learning does not always rely on such labeled data. Additionally, in multi-task learning, models are commonly trained jointly on both the source and target task data, which may not be the case for all transfer learning scenarios. Nevertheless, even in the absence of target data during training, insights obtained from multi-task learning remain valuable for transfer learning. The lessons learned from analysing tasks advantageous for multi-task learning can still guide decision-making in transfer learning.

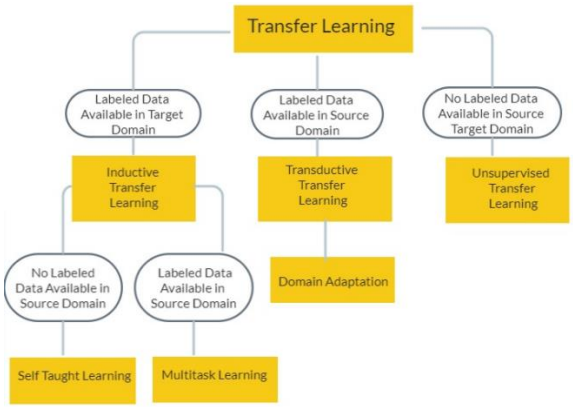

Figure3. Categorization of Transfer Learning based on data availability

The transfer process can be briefly summarized with four main questions: (a) from where to transfer, (b) what to transfer, (d) how to transfer, and (c) when to transfer[7]. The transfer process starts with (i) a target task to be learnt in a target context, (ii) a set of solutions to the source tasks already learnt in the source contexts, and (iii) the transfer (similarity) connections which are decided based on the similarity or resemblance between the target and the source problems. Identification of these assignments answers the "where" question. Even though there are many studies addressing diverse issues in transfer learning, the "where" question is mostly left unanswered. This question mainly refers to the similarity notion which can be very subjective depending on the point of view as it is mentioned in the psychological perspective. Mostly the similarity connections have to be established between the target data and the source data or solutions (models). For instance, in order to decide if a motorbike detection problem is a suitable source for a bicycle detection problem, This setting is also known as "zero shot learning"[8] and the general forms of the meta-data are class labels and associated definitions or attributes which can be obtained by both supervised and unsupervised methods. Once the source problems, target problem and transfer connections are decided, the next step is to decide what to transfer, how, and when. The "what" term generally defines the type of information transferred from the source to the target which can be a solution (model parameters) or data (instances). "How" defines the nature of the transfer, if the transferred knowledge will be transformed or transferred as it is, and how it will be used while learning the new task. The question "when" asks in which situations transfer should be performed. This question is highly related to the concept of negative transfer. If source tasks are not too similar and/or there is already adequate amount of data for learning the target. The question of "When to transfer" explores the circumstances under which skill transfer should be pursued. Similarly, we seek to understand the situations in which knowledge transfer should be avoided. In certain scenarios where the source domain and target domain lack a meaningful relationship, attempting brute-force transfer may yield unsuccessful outcomes. In the worst-case scenario, such transfer attempts can even detrimentally impact the learning performance.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

I. DATASETS

Currently, numerous datasets are accessible for tasks involving scene understanding, including those focused on outdoor scenes traffic scene specifically, the segmentation literature extensively covers representative datasets designed for advanced driver assistant System (ADAS), The subsequent sections will delve into an in-depth analysis of these datasets, providing comprehensive statistics. can be found in Table II.

A. KITTI

KITTI [10] is a dataset for 3D visual tasks, specifically focusing on outdoor scenes. This dataset comprises stereo images that capture the surrounding road environment, accompanied by matching 3D laser scans. The collection of 3D image data is accomplished using a pair of high-resolution stereo camera, one grayscale and one color. It also incorporates the advanced The OXTS RT 3003 localization system integrates GPS, GLONASS, IMU, and RTK correction signals to achieve precise positioning. Additionally, a Velodyne HDL-64E laser scanner mounted on the vehicle generates real-time 3D points for the captured scene

TABLE 2. COMPREHENSIVE ANALYSIS OF EXISTING LITERATURE ON CROSS DOMAIN SCENE UNDERSTANDING WITH REFERENCE TO DATASETS AND ADAPTATION TECHNIQUE USED

Ref	Method	dataset	cross-domain features	Adaptation
[2]	A Dense correspondence-based transfer learning (DCTL)	DCTL	sunny day, night, snowy day, rainy night, cloudy day and foggy day	Subspace domain adaptation
[4]	Curriculum Model Adaptation (CMAda)	Foggy Zurich, Cityscape DBF	Clear weather images, Synthetic Foggy images, real foggy iamges	RefineNet, CMA2, CMA3,AdSegNet Tsai
[6]	Fully Convolutional Autoencoder Network architecture	Cityscapes dataset	NA	NA
[7]	Unsupervised label transfer	Synthia, BDD100k	daytime images , nighttime images	Shared encoder
[9]	Topic Model	QMUL Junction dataset	Activity attributes	DASVM, MDA-HS, TKL, DDA, STL, TNNAR,
[10]	Depth Transformer	KITTI Make3D	Urban raods and highways	deep segmentation models
[12]	A benchmark dataset	1. CamVid dataset 2. Daimler urban segmentation dataset 3. street scenes dataset	sunny day, night, snowy day, rainy night, cloudy day and foggy day	Subspace domain adaptation
[13]	Latent Generative Model with Intensity Consistency	1. Philly-Commuting Road Scene (PRS) 2. Nordland Railroad Scene Dataset (NRS) 3. RobotCar Seasons Dataset (RCS)	sunny, cloudy, snowy, foggy, rainy, spring, summer, falland winter	SIFT Flow, CNN Geo (affine), WAlign NN (affine+TPS), CycleGan+WAlign NN
[17]	Generative adverserial network	BDD100k	sunny day, night, snowy day, rainy night, cloudy day and foggy day	SURF, Decaf6, Decaf7
[25]	Semi supervised semantic segmentation method	MIT67, Google Earth	day, night, rainy	Data augmentation

To ensure precise ground truth data, The stereo cameras in use are meticulously calibrated and synchronized with both the localization system and the laser scanner. Within the dataset, there are a grand total of 14,999 pairs of RGB stereo images, each with a resolution of 1240 × 376 pixels. This includes both the actual image and its corresponding ground truth. The dataset is split into two main portions: a training set containing 7,841 samples and a test set comprising 7,518 samples. Within the training set, there are two distinct subsets a training subset with 3,712 samples and a validation subset with 3,769 samples. The validation subset primarily serves the purpose of validation during the training process.

B. CityScapes

CityScapes [6] is an exceptional dataset designed for high-quality pixel-level semantic segmentation, specifically tailored for understanding urban street scenes. It encompasses approximately 5,000 pixel-level annotated images captured across roughly 50 cities in Germany and neighboring countries. These images showcase intricate urban scenes, exhibiting diverse weather conditions, backgrounds, and scene layouts. The dataset stands out from previous benchmark datasets for street scene understanding due to its exceptional variety, substantial size, intricate scene complexity, and rich annotations.

In order to facilitate the differentiation of semantic representations for individual objects in the captured images, meticulous annotation is performed using 30 different categories. In order to accommodate semantic segmentation tasks, the dataset is partitioned into four distinct subsets: 2,993 images for training, 503 images for validation, 1,531 images for testing, and an auxiliary set comprising 20,021 images. The training, validation, and test image sets are paired with refined high-level annotations, whereas the auxiliary image set contains annotations of a comparatively less detailed or coarser nature.

C. DCTL

The cross-domain traffic scene dataset consists of traffic scene images collected from two separate road routes. The images from the first road route are extracted from five video sequences recorded by a testing vehicle. This subset contains a total of 1,130 traffic scene images [2], encompassing 226 unique

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

locations. The road route includes traffic scenes from both urban and highway environments. The videos were recorded on the identical road route, but they encompassed a range of weather and lighting conditions. In each specific location, images were taken under five diverse conditions: "clear day," "nighttime," "snowy day," "rainy night," and "overcast day." Every image within this subgroup maintains a consistent resolution of 856×270 pixels. The images from the second road route were sourced from two video sequences available on YouTube. This image dataset comprises a total of 698 traffic scene images taken from 349 distinct locations. In each of these locations, two distinct conditions were documented: 'sunny day' and 'rainy day'. All images within this subset maintain a consistent resolution of 640×360 pixels. Furthermore, out of the entire dataset, which includes 1,130 images obtained from the first route and an additional 100 images from the second route, all have been meticulously annotated using the LabelMe tool. These annotations encompass 13 different object categories within these images.

D. SYNTHIA

The SYNTHIA[7] dataset provides a collection of 9,400 multi-viewpoint photo-realistic frames generated from a virtual city. Each frame within the dataset is meticulously annotated at the pixel level, covering 13 different classes. The resolution of every frame is set at 1280×960 pixels.

E. Berkely Deep Drive (BDD)

The BDD dataset [17] is an extensive dataset designed for road scene understanding, encompassing diverse driving videos and GPS/IMU data. The dataset covers various tasks such as drive-able area segmentation, road object detection, instance segmentation, and lane mark detection. It consists of hours of driving footage, showcasing visuals scenes across the different cities in the USA. The dataset captures scenes in different weather and lighting conditions.

In addition to the video data, the dataset also provides GPS/IMU driving trajectories, recorded using GPS, IMU, gyroscope, and magnetometer sensors. These trajectories enable accurate location tracking. The dataset offers image-level annotations for a wide range of driving scene understanding tasks. Notable object detection annotations include traffic lights, traffic signs, buses, pedestrians, motorcycles, bicycles, trucks, and cars. The dataset captures the same scene across different domains, accounting for significant variations in appearance. To address this, transformations are learned using data from both domains.

II. ADAPTATION TECHNOQUES USED

A. Subspace-based domain adaptation

Subspace-based domain adaptation involves the projection of both source and target data into a shared subspace to ensure maximum consistency between their distributions. This approach assumes the presence of abundant labeled data in the source domain, but limited data in the target domain. The objective is to leverage information from labeled data in the source domain to adapt to new data in the target domain. In the context of traffic scene understanding [2], we consider weather or illumination conditions as distinct domains, and our focus is on cross-domain learning. Our goal is to tackle the challenge of recognizing images of the same scene across different domains, accounting for significant variations in appearance.

B. DASVM method

The existing method can only handle scenarios where the features of the source and target domains are isomorphic. In the DASVM method [10], after initializing the discriminant function with source-domain samples for the target domain problem, it iteratively eliminates the source-domain samples and gradually adjusts the discriminant function to fit the target domain instances. Each iteration of DASVM requires a time equivalent to that of supervised SVM learning.

On the other hand, the TKL method, a data-dependent spectral learning approach, focuses solely on adapting the edge distribution while disregarding the alignment of the conditional probability distribution. Therefore, further investigation into joint distribution adaptation based on spectral learning is necessary.

specifically designed. STL [10] leverages the intra-affinitys of classes to iteratively transfer knowledge within the same class. Meanwhile, TNNAR [10] performs knowledge transfer for activity recognition by selecting multiple source domains and employing deep neural networks. In CDAR, the source and target domain data share the same dimensions, labels, and potentially the same data distribution. However, applying these methods to practical cross-domain traffic scene understanding applications may not be feasible due to the inherent differences in the scene characteristics.

C. MDA-HS method

The method trains individual SVM classifiers are generated using training data from individual source domains. The ultimate prediction for each target domain sample is determined by averaging the predictions made by all these classifiers. computational complexity of the method is primarily determined by the training of the SVM classifiers. In the MDA-HS method[9], the problem involves samples from different source domain The source domain samples are characterized by

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

diverse feature types, whereas the target domain samples encompass all types of features.

The DDA approach seamlessly conducts classifier adaptation, distribution alignment, and distinctive embedding. Various domains showcase unique feature representations. In contrast, the MFSAN technique employs multiple domain-invariant representations for training multiple domain-specific classifiers. It additionally aligns the results of these classifiers for the target samples. The training procedure encompasses refining convolutional and pooling layers and training classifier layers through backpropagation, a process that can be computationally demanding.

CWAN stands as a deep learning framework explicitly tailored to tackle the complexity of multisource heterogeneous domain adaptation (MHDA). Its uniqueness lies in its ability not just to assess the significance of distinct source domains, but also to harness the conditional distribution existing between the source and target domains, thereby facilitating a potent transfer of knowledge. However, when dealing with a larger number of source domains, the model needs to solve additional learnable parameters, resulting in a complex optimization problem.

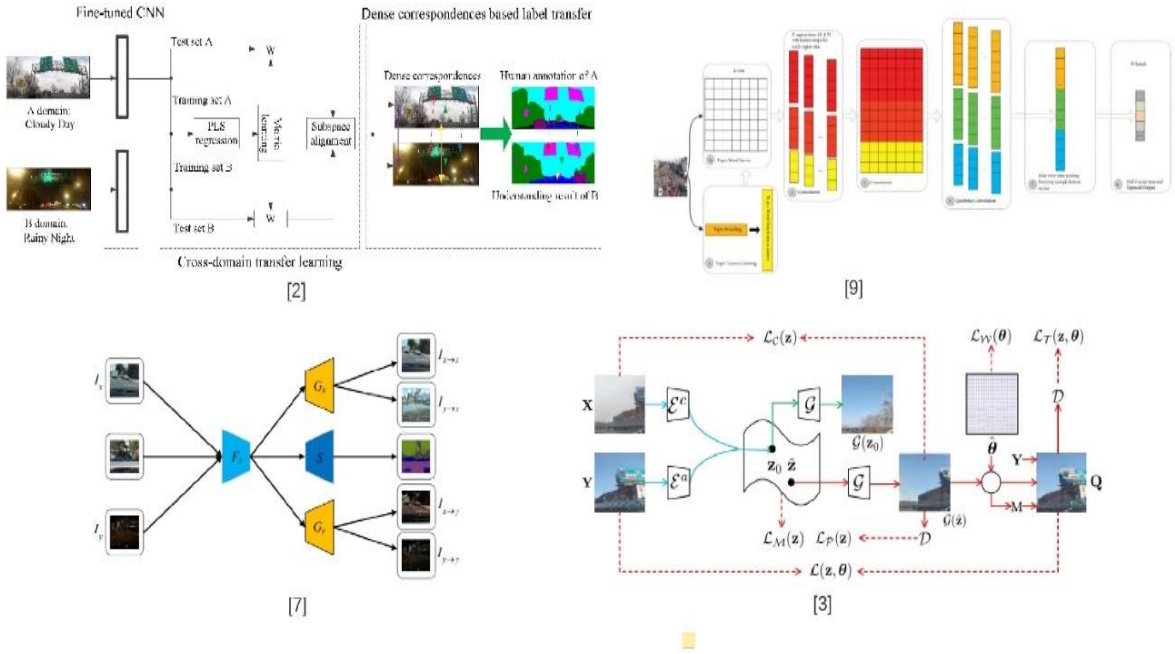

Figure 4: Various cross domain scene understanding architectures, adopted by mainstream research in baseline strategies

D. Generative Adversial Network(GAN)

During GAN training, the objective is to train a generator to deceive employ a discriminator as part of our research, with the primary goal of distinguishing between artificially generated samples and authentic ones. In our study, we predominantly utilize GAN [6] to investigate latent images existing within the latent image manifold within the desired domain. Our research is closely related to certain image-to-image translation techniques. For instance, the conditional GAN (cGAN) employs a conditional GAN framework to learn the transformation between input and output images at a "pixel-to-pixel" level. Meanwhile, CycleGAN introduces a learning method for translating an image from a source domain to a target domain without the need for paired examples. In current research, objective is to establish a dense correspondence. However, existing image-to-image translation methods fall short in generating an appropriate image for the purpose of image alignment.

Recent studies have combined GANs with Auto-encoders are designed to segregate the encoding of images into distinct appearance and content codes. In the context of registration or alignment applications. It is essential to preserve the content for spatial information tracking while adjusting the appearance to align with the pixel intensity distribution.

DeCAF

To facilitate the widespread analysis of deep convolutional features, we have developed a Python framework that enables

easy training of networks comprising various types of layers. Additionally, our framework allows efficient execution of pre-trained networks without the constraint of requiring a GPU, which can sometimes impede the deployment of trained models [18].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
Article Received: *26 August 2023* ***Revised:*** *20 October 2023* ***Accepted:*** *02 November 2023*

, utilizing C implementation for computationally intensive portions of the code that are linked to Python. In terms of computational speed, our model can process approximately 40 images per second on an 8-core standard machine when executing the CNN model in a mini-batch mode.

E. SIFT FLOW

SIFT Flow is a computer vision algorithm that combines the Scale-Invariant Feature Transform (SIFT) algorithm with optical flow techniques. It was proposed by Liu, Y., et al. in 2008 as a method for aligning and matching images with significant appearance changes, such as differences in viewpoint, scale, and illumination. The SIFT Flow algorithm aims to establish dense correspondences between pixels in two images, enabling the alignment of image pairs with large geometric and photometric variations. The process involves two primary stages: SIFT Feature Extraction, wherein the SIFT algorithm is utilized to capture unique local features from both images. SIFT features are invariant to changes in scale, rotation, and affine transformations, making them robust descriptors for matching. Flow Computation: Optical flow techniques are employed to estimate the dense correspondence field between the SIFT feature points in the two images. The computed flow field represents the pixel-level correspondences, allowing for the alignment and matching of the images. features are invariant to changes in scale, rotation, and affine transformations, making them robust descriptors for matching. Flow Computation: Optical flow techniques are employed to estimate the dense correspondence field between the SIFT feature points in the two images. The computed flow field represents the pixel-level correspondences, allowing for the alignment and matching of the images.

VII. METHODOLOGIES AND TECHNIQUES USED

A. Evaluation Metrics

In our evaluation, we employ commonly used measures for scene understanding systems, namely the average pixel-wise recognition rates and per-class recognition rates [35]. These metrics serve as effective indicators of performance in assessing scene understanding capabilities.

B. Scene Retrieval

The objective of scene retrieval is to find a set of traffic scene images in an archived dataset that exhibit visual similarity to a given query image, denoted as I. This process involves calculating a similarity measure, denoted as m(I, Id), between the query image I and the images in the database, resulting in a set of top-N most similar images, denoted as R = {I1d, I2d, ..., Ind}[35]. In this context, our aim is to locate images from the same location when provided with an image from a particular scenario. However, as depicted in Figure 1, images from the same location can display significant variations in appearance due to factors such as illumination, rain, snow, and more. Therefore, it is crucial to employ a robust feature representation that can effectively compute the similarity measure.For this purpose, deep Convolutional Neural Network (CNN) features are utilized, which have been learned on a large and diverse dataset. These features serve as powerful descriptors that can be applied to various datasets. In this study, The image representations for all traffic scenes are derived from the final convolutional layer output of AlexNet [11], which was pretrained on the ILSVRC-2012 dataset. Consequently, when a test image from a particular scenario is introduced, the comparison between this test image and the database images takes place within the realm of deep features.

C. Label Transfer

The basic objective which is getting followed in many approaches is to transfer the labels from the retrieved similar images to the input image, necessitating the establishment of dense correspondences between them. Existing approaches such as SIFT flow, as demonstrated in [2] and [12], have shown the ability To create semantically significant correspondences, an approach involving the alignment of local SIFT representations is employed. An alternative tactic is, explored in [27], involves working with super pixels to alleviate the computational burden associated with pixel-level inferences in large-scale datasets. Once the similar images have been obtained by utilizing the dense correspondences, it becomes possible to transfer the labels obtained from the similar images retrieved from the input image.

VIII. CONCLUSIONS AND EXPECTED OUTCOMES

This survey aims to provide consolidated and summarized analysis of scene understanding in autonomous vehicles by discussing the strengths and weaknesses of existing methods in different environments. Our main finding is that scene understanding in autonomous vehicles is still far from perfect, with various limitations present in current methods. We have provided a comprehensive review of these limitations and offered relevant suggestions and outlooks.

The survey thoroughly examines foundational studies that center around deep learning models, their hierarchical structures, and the specific challenges associated with each model category. Moreover the survey delves into performance evaluation strategies that are suitable for segmentation models, explores special loss functions, and discusses widely adopted approaches in the field. datasets in the autonomous driving domain. Open challenges in scene understanding are highlighted, along with future research directions supported by recent literature references. It is evident that experts in the intelligent transportation systems (ITS) community are continuously striving to develop more effective scene understanding strategies using data. The current focus of mainstream research is primarily on improving model accuracy

by leveraging the capabilities of neural layers. However, it is essential to recognize that there are additional challenges that must be tackled to ensure the development of reliable, trustworthy, and safe autonomous driving systems. In order to improve scene understanding, there is an urgent requirement for robust models that integrate Levels of preference for segmented objects, the ability to handle coarse structural details, and the classification of potential risks. Capitalizing on these opportunities in a timely manner can propel research in Intelligent Transportation Systems (ITS) and take scene segmentation to new heights. Such advancements can facilitate the integration of autonomous vehicles into real-world environments, thereby promoting safer and more dependable travel services.

REFERENCES

[1] H. C. Shin et al., 'Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning', IEEE Trans Med Imaging, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.

[2] S. Di, H. Zhang, C. G. Li, X. Mei, D. Prokhorov, and H. Ling, 'Cross-Domain Traffic Scene Understanding: A Dense Correspondence-Based Transfer Learning Approach', IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 3, pp. 745–757, Mar. 2018, doi: 10.1109/TITS.2017.2702012.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, 'ImageNet Classification with Deep Convolutional Neural Networks', Commun ACM, vol. 60, no. 6, 2017, doi: 10.1145/3065386.

[4] D. Dai, C. Sakaridis, S. Hecker, and L. van Gool, 'Curriculum Model Adaptation with Synthetic and Real Data for Semantic Foggy Scene Understanding', Int J Comput Vis, vol. 128, no. 5, pp. 1182–1204, May 2020, doi: 10.1007/s11263-019-01182-4.

[5] J. Lu, M. Xu, R. Yang, and Z. Wang, 'Understanding and Predicting the Memorability of Outdoor Natural Scenes', IEEE Transactions on Image Processing, vol. 29, pp. 4927–4941, 2020, doi: 10.1109/TIP.2020.2975957.

[6] H. Zhou, J. Ma, C. C. Tan, Y. Zhang, and H. Ling, 'Cross-Weather Image Alignment via Latent Generative Model with Intensity Consistency', IEEE Transactions on Image Processing, vol. 29, pp. 5216–5228, 2020, doi: 10.1109/TIP.2020.2980210.

[7] C. Song, J. Wu, L. Zhu, M. Zhang, and H. Ling, 'Nighttime Road Scene Parsing by Unsupervised Domain Adaptation', IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 4, pp. 3244–3255, Apr. 2022, doi: 10.1109/TITS.2020.3033569.

[8] M. Milford et al., 'Condition-invariant, top-down visual place recognition', in 2014 IEEE International Conference on Robotics and Automation (ICRA), May 2014, pp. 5571–5577. doi: 10.1109/ICRA.2014.6907678.

[9] P. Soviany, R. T. Ionescu, P. Rota, and N. Sebe, 'Curriculum self-paced learning for cross-domain object detection', Computer Vision and Image Understanding, vol. 204, Mar. 2021, doi: 10.1016/j.cviu.2021.103166.

[10] Y. Yang, H. Dong, G. Liu, L. Zhang, and L. Li, 'Cross-Domain Traffic Scene Understanding by Integrating Deep Learning and Topic Model', Comput Intell Neurosci, vol. 2022, 2022, doi: 10.1155/2022/8884669.

[11] C. Zhang, W. Ding, G. Peng, F. Fu, and W. Wang, 'Street View Text Recognition with Deep Learning for Urban Scene Understanding in Intelligent Transportation Systems', IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 7, pp. 4727–4743, Jul. 2021, doi: 10.1109/TITS.2020.3017632.

[12] M. Naseer, S. H. Khan, and F. Porikli, 'Indoor Scene Understanding in 2.5/3D for Autonomous Agents: A Survey', Mar. 2018, doi: 10.1109/ACCESS.2018.2886133.

[13] C. Ju, D. Gao, R. Mane, B. Tan, Y. Liu, and C. Guan, Federated Transfer Learning for EEG Signal Classification. 2020. doi: 10.0/Linux-x86_64.

[14] H. Zhou, J. Ma, C. C. Tan, Y. Zhang, and H. Ling, 'Cross-Weather Image Alignment via Latent Generative Model with Intensity Consistency', IEEE Transactions on Image Processing, vol. 29, pp. 5216–5228, 2020, doi: 10.1109/TIP.2020.2980210.

[15] U. Nadeem, S. A. A. Shah, F. Sohel, R. Togneri, and M. Bennamoun, 'Deep learning for scene understanding', in Smart Innovation, Systems and Technologies, vol. 136, Springer Science and Business Media Deutschland GmbH, 2019, pp. 21–51. doi: 10.1007/978-3-030-11479-4_2.

[16] X. Li et al., 'Yolov3-Pruning(transfer): real-time object detection algorithm based on transfer learning', J Real Time Image Process, Aug. 2022, doi: 10.1007/s11554-022-01227-x.

[17] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, 'A survey of transfer learning', Journal of Big Data, vol. 3, no. 1, Dec. 2016, doi: 10.1186/s40537-016-0043-6.

[18] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell. 'DeCAF: a deep convolutional activation feature for generic visual recognition', In Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32 (ICML'14) 2014.

[19] Torrey, L. and Shavlik, J. 'Transfer Learning' Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques, IGI Global, Hershey, 242-264 (2010) https://doi.org/10.4018/978-1-60566-766-9.ch011.

[20] P. Ganesh Pawar and V. Devendran, 'Scene Understanding: A Survey to See the World at a Single Glance', 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT), 2019.

[21] M. Long, Y. Cao, Z. Cao, J. Wang, and M. I. Jordan, 'Transferable Representation Learning with Deep Adaptation Networks', IEEE Trans Pattern Anal Mach Intell, vol. 41, no. 12, pp. 3071–3085, Dec. 2019, doi: 10.1109/TPAMI.2018.2868685.

[22] A. Chakraborty, · Cosmin Anitescu, · Xiaoying Zhuang, and T. Rabczuk, 'Domain adaptation based transfer learning approach for solving PDEs on complex geometries', vol. 1, p. 3, doi: 10.1007/s00366-022-01661-2.

[23] S. J. Pan and Q. Yang, 'A survey on transfer learning', IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10. pp. 1345–1359, 2010. doi: 10.1109/TKDE.2009.191.

[24] D. P. Thesis, 'Transfer Learning for Object Category Detection'.

[25] M. Afif, R. Ayachi, Y. Said, and M. Atri, 'Deep Learning Based Application for Indoor Scene Recognition', Neural Process Lett, vol. 51, no. 3, pp. 2827–2837, Jun. 2020, doi: 10.1007/s11063-020-10231-w.

[26] R. Xiao, Y. Wang, and C. Tao, 'Fine-Grained Road Scene Understanding from Aerial Images Based on Semisupervised Semantic Segmentation Networks', IEEE Geoscience and Remote Sensing Letters, vol. 19, 2022, doi: 10.1109/LGRS.2021.3059708.

[27] K. Wang, X. Gao, Y. Zhao, X. Li, D. Dou, and C.-Z. Xu, 'Pay Attention to Features, Transfer Learn Faster CNNs', International Conference on Learning Features, 2020.

[28] 'Scene Understanding Using Deep Learning - ScienceDirect'. https://www.sciencedirect.com/science/article/pii/B978012811318900020X (accessed Aug. 25, 2022).

[29] J. Lu, N. Verma, and N. K. Jha, 'Convolutional Autoencoder-Based Transfer Learning for Multi-Task Image Inferences', IEEE Trans Emerg Top Comput, vol. 10, no. 2, pp. 1045–1057, 2022, doi: 10.1109/TETC.2021.3068063.

[30] F. Husain, B. Dellen, and C. Torras, 'Scene Understanding Using Deep Learning', Handbook of Neural Computation, pp. 373–382, Jan. 2017, doi: 10.1016/B978-0-12-811318-9.00020-X.

[31] F. Zhuang et al., 'A Comprehensive Survey on Transfer Learning', CORR Nov. 2019, Available: http://arxiv.org/abs/1911.02685

[32] S. Safavi and M. Jalali, 'DeePOF: A hybrid approach of deep convolutional neural network and friendship to Point-of-Interest (POI) recommendation system in location-based social networks', Concurr Comput, vol. 34, no. 15, Jul. 2022, doi: 10.1002/CPE.6981.

[33] Zhu L, Yu F, Wang Y, Ning B, IEEE Transactions on Intelligent Transportation Systems, (2019), 383-398, 20(1)

[34] Shuai Di, Honggang Zhang, Xue Mei, D. Prokhorov and H. Ling, "A benchmark for cross-weather traffic scene understanding," *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, Rio de Janeiro, 2016, pp. 2150-2156, doi: 10.1109/ITSC.2016.7795904.

Ms. Deepa B. Mane

Research Scholar in Computer Engineering at SPPU. Received M.E. (Information Technology) Pune Institute of computer technology, Pune, India. B.E. (Computer Engineering) from Mumbai University. Her Research interest are in Machine Learning, deep learning and Data Science. Working as assistamt Professor in SPPU. Has Received Best Paper Presented in National Conference on Technical Revolution 2016.

Dr. Mrs. Sandhya Arora

Working as Professor (Computer Engineering) at Cummins College of Engineering. She has more than 25 years of academia. Received PH.D. in Computer Engineering from Jadavpur University from Kolkata India. She has Research Interest in Soft Computing, Online Learning, Intelligent Tutoring System, Pattern Recognition, Image processing, Data analytics, machine learning. She has published many research papers in top cited journals with 736 Citations.. Her h-index is 11. She has also published books and Patents. Second Prize Winner of National Event "Teaching-Learning-Evaluation Hackathon" conducted by Symbiosis University, Pune.

Mr. Sachin D. Shelke

pursed Bachelor of Computer Engineering from Shivaji University Kolhapur, Maharashtra and Master of Information Technology from Bharti Vidyapeeth, Pune. He is currently pursuing Ph.D. and working as Assistant Professor in Department of Information Technology, PICT, Pune. He is a Life member of Computer Society of India (CSI). His main research work focuses on Data Analytics, Data Structures and Software test