

Titles &
Abstract

2024-2025



ELYSIUMPRO
INSPIRING THE LEADING EDGE TECHNOLOGIES

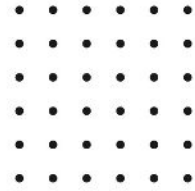
Advanced Academic Final Year Projects

IMAGE PROCESSING

IEEE Projects



elysiumpro.in



EPRO_IMG_001

BoNuS: Boundary Mining for Nuclei Segmentation with Partial Point Labels

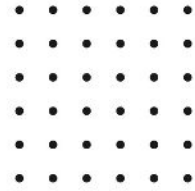
In this project, we present BoNuS, a novel approach for nuclei segmentation leveraging boundary mining with partial point labels. The proposed method aims to enhance the accuracy and efficiency of segmenting cell nuclei in histopathological images. Our approach begins with preprocessing steps, including grayscale conversion, adaptive histogram equalization, and median filtering, to improve image quality and contrast. Subsequently, we employ binarization and morphological operations to extract and refine the nuclei boundaries. To evaluate the segmentation performance, we calculate key metrics such as Intersection over Union (IoU) and Dice coefficient, which provide insights into the overlap and accuracy of the predicted masks compared to the ground truth. Our GUI application enables users to visually inspect the input images, original masks, and predicted masks, alongside the computed threshold value, IoU, and Dice coefficient for each image. The BoNuS framework demonstrates robust performance in segmenting nuclei, even with limited annotations, making it a valuable tool for facilitating histopathological analysis and advancing research in computational pathology.

Ultrasound Nodule Segmentation Using Asymmetric Learning with Simple Clinical Annotation

ratio annotations directly from ultrasound clinical diagnoses for automated nodule segmentation. Especially, an asymmetric learning framework is developed by extending the aspect ratio annotations with two types of pseudo labels, i.e., conservative labels and radical labels, to train two asymmetric segmentation networks simultaneously. Subsequently, a conservative-radicalbalance strategy (CRBS) strategy is proposed to complementally combine radical and conservative labels. An inconsistency-aware dynamically mixed pseudo-labels supervision (IDMPS) module is introduced to address the challenges of over-segmentation and under-segmentation caused by the two types of labels. To further leverage the spatial prior knowledge provided by clinical annotations, we also present a novel loss function namely the clinical anatomy prior loss. Extensive experiments on two clinically collected ultrasound datasets (thyroid and breast) demonstrate the superior performance of our proposed method, which can achieve comparable and even better performance than fully supervised methods using ground truth annotations

EPRO_IMG_002





EPRO_IMG_003

Multiple Latent Space Mapping for Compressed Dark Image Enhancement

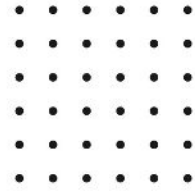
enhancing compressed dark images while avoiding compression artifacts amplification. Since texture details intertwine with compression artifacts in compressed dark images, detail enhancement and blocking artifacts suppression contradict each other in image space. Therefore, we handle the task in latent space. To this end, we propose a novel latent mapping network based on variational auto-encoder (VAE). Firstly, different from previous VAE-based methods with single-resolution features only, we exploit multiple latent spaces with multi-resolution features, to reduce the detail blur and improve image fidelity. Specifically, we train two multilevel VAEs to project compressed dark images and normal-light images into their latent spaces respectively. Secondly, we leverage a latent mapping network to transform features from compressed dark space to normal-light space. Specifically, since the degradation models of darkness and compression are different from each other, the latent mapping process is divided mapping into enlightening branch and deblocking branch. Comprehensive experiments demonstrate that the proposed method achieves state-of-the-art performance in compressed dark image enhancement.

Anomaly Detection for Medical Images Using Teacher–Student Model With Skip Connections Mingxuan Liu and Multiscale Anomaly

Anomaly detection (AD) in medical pictures identifies aberrant inputs based on normal training examples. Knowledge distillation (KD) based on the T-S model is a simple and effective method for identifying anomalies. However, its effectiveness is limited by the similarity between teacher and student network architectures (S-Nets). This article proposes a T-S model with skip connections (Skip-TS) trained using direct reverse KD (DRKD) for detecting AD in medical images. To address the issue of low sensitivity to structural similarity, we propose an encoder-decoder architecture with a pre-trained encoder (T-Net) and a randomly initialized decoder (S-Net). In our base paper use different type of dataset for checking the segmentation model here use the Breast cancer dataset for checking our S-net and T-net. First import all the packages and read the images, preprocess all the images by resizing and color conversion and then design the encoder and decoder using the up sampling and down-sampling model. Next step is to fit the model with the train and test image with mask images and then finally get the segmented portion for the selected image.

EPRO_IMG_004





EPRO_IMG_005

Learning Decentralized Traffic Signal Controllers with Multi-Agent Graph Reinforcement Learning

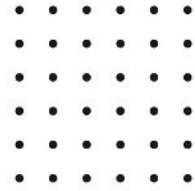
The Smart Traffic Management System using cloud computing aims to enhance urban traffic flow and reduce congestion through advanced image processing and deep learning techniques. The system begins by processing vehicle images from a dataset, performing pre-processing tasks like resizing and grey-scale conversion. Key features are extracted using statistical measures and Local Binary Pattern (LBP). These images are then split into training and testing sets to develop and evaluate a vehicle detection model using YOLO. Traffic density is computed based on vehicle counts per unit length and sent to a cloud server for storage. The server further calculates the optimal green signal time using a Signal Switching Algorithm, designed to optimize traffic flow. The system's performance is quantified through metrics such as accuracy and error rate, and results are presented in comparison graphs and tables.

Analysis of Brain Imaging Data for the Detection of Early Age Autism Spectrum Disorder Using Transfer Learning Approaches for Internet of Things

This project delves into the potential of using deep learning to enhance the diagnosis of Autism Spectrum Disorder (ASD). Functional magnetic resonance imaging (fMRI) data shows promise in revealing brain dysfunction in ASD patients. We employ deep learning techniques, specifically Convolutional Neural Networks (CNNs) and transfer learning, to analyze fMRI data from the Autism Brain Imaging Data Exchange (ABIDE) datasets. Our optimized CNN achieves an 81% accuracy in classifying autistic and typically developing brains, exceeding the performance of previous methods. This project provide that deep learning could serve as a valuable tool for ASD diagnosis. We also aim to evaluate the model by calculating its accuracy, precision, and recall for classification purposes.

EPRO_IMG_006





EPRO_IMG_007

Development of Convolutional Neural Network to Segment Ultrasound Images of Histotripsy Ablation

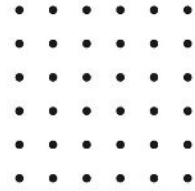
Accurate segmentation of kidney tumors is crucial for diagnosis, treatment planning, and monitoring of kidney tumour. Convolutional neural networks (CNNs) have shown promising results in medical image segmentation tasks. In this study, we propose a CNN-based approach for the segmentation of kidney tumors from abdominal ultrasound images. The proposed method involves a multi-stage CNN architecture that first localizes the kidney region and then segments the tumor within the kidney. The accurate and robust segmentation of kidney tumors can assist clinicians in early detection, treatment planning, and monitoring of kidney cancer, ultimately improving patient outcomes.

IACC: Cross-Illumination Awareness and Color Correction for Underwater Images Under Mixed Natural and Artificial Lighting

In the realm of underwater photography, the quest to overcome challenges related to limited visibility and color distortion has spurred innovative research endeavors. This study integrates cutting-edge deep learning methodologies with conventional image processing techniques, particularly emphasizing the efficacy of hybrid networks. The primary workflow involves the enhancement of low-light underwater images through the application of hybrid-CNNs. The process begins with the collection and preprocessing of a comprehensive dataset, including extraction of Red, Blue, and Green channels, resizing, and color correction. Subsequently, the dataset undergoes a split of 80:20 ratio, introducing variability. The pivotal stage employs CNN for classify the classes. Further, YOLO-v8 is applied for object detection, providing parameters for image verification through metrics such as accuracy, loss, confusion matrix, classification report. The study culminates in the evaluation of model performance, elucidated by accuracy and loss metrics, advancing the understanding and application of image enhancement in underwater photography.

EPRO_IMG_008





EPRO_IMG_009

Federated-Reinforcement Learning-Assisted IoT Consumers System for Kidney Disease Images

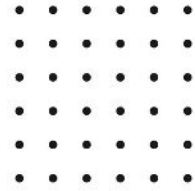
This study introduces a novel approach utilizing Support Vector Machines (SVM) and Convolutional Neural Networks (CNN) for the automated identification of kidney stones from medical images. The proposed system leverages SVM for feature extraction and classification, complemented by CNNs for deep learning-based image analysis. Integrating these algorithms enables robust prediction capabilities, distinguishing between images depicting kidney stones and normal kidney structures. The effectiveness of the approach is validated through rigorous experimentation on a dataset of kidney disease images, demonstrating high accuracy and reliability in diagnostic outcomes. This research contributes to advancing medical imaging technology by providing a feasible and efficient tool for early detection and diagnosis of kidney stones, facilitating timely medical interventions and improved patient outcomes.

A Deep Quantum Convolutional Neural Network Based Facial Expression Recognition For Mental Health Analysis

In image preprocessing, the a facial region of interest has been segmented from the input image with using classical deep feature representation and the quantum part that involves successive sets of quantum convolutional layers followed by random quantum variational circuits for feature learning . In this the proposed system has attained a faster training approach using the proposed quantum convolutional neural network approach that takes $O(\log(n))$ time. Facial expression recognition, a crucial aspect of social cognition, is one of the diagnostic abnormalities of schizophrenia. Using facial expression recognition system has been proposed that analyzes static, sequential, or video facial images from medical healthcare data to detect emotions in people's facial regions . The algorithms such as Quantum Convolutional neural network for predicting and recognizing the face emotions. Finally, the experimental results shows that the performance metrics such as accuracy, precision and recall.

EPRO_IMG_010





EPRO_IMG_011

Light-weight Retinal Layer Segmentation

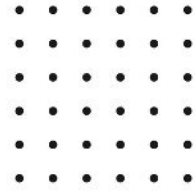
Automatic retinal layer segmentation with medical images, such as optical coherence tomography (OCT) images, serves as an important tool for diagnosing ophthalmic diseases. However, it is challenging to achieve accurate segmentation due to low contrast and blood flow noises presented in the images. In addition, the algorithm should be light-weight to be deployed for practical clinical applications. Therefore, it is desired to design a light-weight network with high performance for retinal layer segmentation. In this paper, we propose LightReSeg for retinal layer segmentation which can be applied to OCT images. Specifically, our approach follows an encoder-decoder structure, where the encoder part employs multi-scale feature extraction and a Transformer block for fully exploiting the semantic information of feature maps at all scales and making the features have better global reasoning capabilities, while the decoder part, we design a multi-scale asymmetric attention (MAA) module for preserving the semantic information at each encoder scale. The experiments show that our approach achieves a better segmentation performance compared to the current state-of-the-art method TransUnet with 105.7M parameters on both our collected dataset and two other public datasets, with only 3.3M parameters.

Deep Omni-supervised Learning for Rib Fracture Detection from Chest Radiology Images

In this study, we explore the application of deep omni-supervised learning techniques for the detection of rib fractures from chest radiology images. The primary objective is to accurately identify regions of rib fractures and predict their affected status using deep neural networks such as ResNet or MobileNet. Chest radiology plays a critical role in diagnosing rib fractures, but the process can be challenging due to variations in image quality, patient anatomy, and fracture characteristics. By leveraging deep omni-supervised learning, which combines labeled, unlabeled, and weakly labeled data, we aim to enhance the robustness and accuracy of fracture detection. The proposed approach harnesses the hierarchical features learned by models like ResNet or MobileNet, enabling effective localization and classification of rib fractures. =

EPRO_IMG_012





EPRO_IMG_013

Multi-Task Learning for Fatigue Detection and Face Recognition of Drivers via Tree-Style Space-Channel Attention Fusion Network

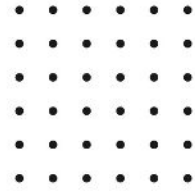
Drowsiness of drivers is one of the main reasons behind road accidents. It is natural for the drivers who take long drives to doze off behind the steering wheel. Drowsiness and fatigue detection are critical for ensuring safety in various domains, such as driving, aviation, and industrial operations. This paper presents a deep learning approach to detect drowsiness and fatigue in using LAnet, a convolutional neural network (CNN) specifically designed for analyzing facial features and eye positions. LAnet leverages advancements in image processing and deep learning to accurately identify signs of drowsiness and fatigue from images. Our method involves several stages: image loading, pre-processing, feature extraction, and classification. The core of our system, LAnet, is trained on a diverse dataset of facial images with various states of alertness, enabling it to learn and indicate of drowsiness.

Automatic Accident Detection, Segmentation and Duration Prediction Using Machine

In recent years, the application of machine learning in the field of traffic management has gained significant attention, particularly for automatic accident detection, segmentation, and duration prediction. This study explores a comprehensive approach that leverages advanced machine learning algorithms to enhance road safety and traffic flow efficiency. The proposed system integrates multiple stages: accident detection, precise segmentation of accident scenes, and accurate prediction of accident duration. Using real-time traffic data and sensor inputs, the system employs deep learning models to identify anomalies and potential accidents. Techniques such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are utilized to process and analyze vast amounts of data, ensuring prompt and reliable detection. Once an accident is detected, the system performs segmentation to isolate and identify the specific regions of interest within the scene. This involves using advanced image processing techniques and neural networks to delineate the affected areas, which aids in efficient resource allocation and emergency response.

EPRO_IMG_014





EPRO_IMG_015

Enhanced Biometric Template Protection Schemes for Securing Face Recognition in IoT Environment

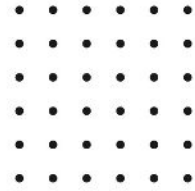
In the rapidly evolving landscape of Internet of Things (IoT) environments, securing face recognition systems against unauthorized access is paramount. This paper proposes enhanced biometric template protection schemes aimed at ensuring the integrity and privacy of facial recognition data. The objective is to develop a robust system capable of accurately determining whether a detected face is authorized. This is achieved through the utilization of machine learning techniques such as Support Vector Machines (SVM) and deep learning methods including Convolutional Neural Networks (CNN) tailored for 2D facial analysis. Once the face's authorization status is determined, the system extracts pertinent biometric information, encrypts it using RSA encryption, and securely stores it in the cloud. The proposed algorithm ensures that only authorized entities possess the capability to decrypt and access the stored data, thereby safeguarding sensitive information from unauthorized disclosure.

DiffWater: Underwater Image Enhancement Based on Conditional Denoising Diffusion Probabilistic Model

Enhancing underwater images poses significant challenges due to issues like poor white balance, low contrast, and non-uniform illumination. By incorporating advanced techniques such as white balance correction, contrast adjustment, and histogram equalization, DiffWater effectively improves the visual quality of underwater images. The core of this enhancement process is driven by a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. The CNNs are employed to capture and refine the spatial features of the image, while the LSTMs handle the temporal dependencies to model the intricate variations in underwater scenes. This hybrid algorithm not only corrects color distortions and enhances contrast but also normalizes the image histogram, providing a clearer and more accurate representation of underwater environments.

EPRO_IMG_016





EPRO_IMG_017

AIROGS: Artificial Intelligence for Robust Glaucoma Screening Challenge

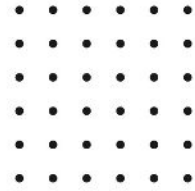
The AIROGS challenge focuses on enhancing the accuracy and robustness of glaucoma screening through advanced artificial intelligence techniques. Our approach leverages the power of two state-of-the-art convolutional neural network architectures: ResNet and MobileNet. ResNet's deep residual learning framework addresses the vanishing gradient problem, enabling the model to learn intricate features and patterns associated with glaucoma from large-scale retinal imaging datasets. This synergy enhances the robustness of the screening process, ensuring reliable detection across diverse populations and varying imaging conditions. Our solution also incorporates advanced pre-processing techniques and data augmentation to improve generalization and reduce overfitting, ensuring consistent performance in real-world clinical settings.

TRTAR: Transmissive RIS-assisted Through-the-wall Human Activity Recognition

Device-free human activity recognition plays a pivotal role in wireless sensing. However, current systems often fail to accommodate signal transmission through walls or necessitate dedicated noise removal algorithms. To overcome these limitations, we introduce TRTAR: a device-free passive human activity recognition system integrated with a transmissive reconfigurable intelligent surface (RIS). TRTAR eliminates the necessity for dedicated devices or noise removal algorithms, while specifically addressing signal propagation through walls. Unlike existing approaches, TRTAR solely employs a transmissive RIS at the transmitter or receiver without modifying the inherent hardware structure. Experimental results demonstrate that TRTAR attains an average accuracy of 98.13% when signals traverse concrete walls.

EPRO_IMG_018





EPRO_IMG_019

ChatCAD+: Towards a Universal and Reliable Interactive CAD using LLMs

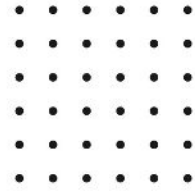
The integration of Computer-Aided Diagnosis (CAD) with Large Language Models (LLMs) presents a promising frontier in clinical applications, notably in automating diagnostic processes akin to those performed by radiologists and providing consultations similar to a virtual family doctor. Despite the promising potential of this integration, current works face at least two limitations: (1) From the perspective of a radiologist, existing studies typically have a restricted scope of applicable imaging domains, failing to meet the diagnostic needs of different patients. Also, the insufficient diagnostic capability of LLMs further undermine the quality and reliability of the generated medical reports. (2) Current LLMs lack the requisite depth in medical expertise, rendering them less effective as virtual family doctors due to the potential unreliability of the advice provided during patient consultations. To address these limitations, we introduce ChatCAD+, to be universal and reliable. Specifically, it is featured by two main modules: (1) Reliable Report Generation and (2) Reliable Interaction.

A Colorectal Coordinate-Driven Method for Colorectum and Colorectal Cancer Segmentation in Conventional CT Scans

Automated colorectal cancer (CRC) segmentation in medical imaging is the key to achieving automation of CRC detection, staging, and treatment response monitoring. Compared with magnetic resonance imaging (MRI) and computed tomography colonography (CTC), conventional computed tomography (CT) has enormous potential because of its broad implementation, superiority for the hollow viscera (colon), and convenience without needing bowel preparation. However, the segmentation of CRC in conventional CT is more challenging due to the difficulties presenting with the unprepared bowel, such as distinguishing the colorectum from other structures with similar appearance and distinguishing the CRC from the contents of the colorectum. To tackle these challenges, we introduce DeepCRC-SL, the first automated segmentation algorithm for CRC and colorectum in conventional contrast-enhanced CT scans. We propose a topology-aware deep learning-based approach, which builds a novel 1-D colorectal coordinate system and encodes each voxel of the colorectum with a relative position along the coordinate system.

EPRO_IMG_020





EPRO_IMG_021

Recurrent Generic Contour-based Instance Segmentation with Progressive Learning

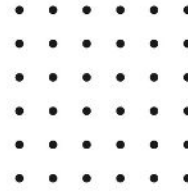
Contour-based instance segmentation has been actively studied, thanks to its flexibility and elegance in processing visual objects within complex backgrounds. In this work, we propose a novel deep network architecture, i.e., PolySnake, for generic contour-based instance segmentation. Motivated by the classic Snake algorithm, the proposed PolySnake achieves superior and robust segmentation performance with an iterative and progressive contour refinement strategy. Technically, PolySnake introduces a recurrent update operator to estimate the object contour iteratively. It maintains a single estimate of the contour that is progressively deformed toward the object boundary. At each iteration, PolySnake builds a semantic-rich representation for the current contour and feeds it to the recurrent operator for further contour adjustment. Through the iterative refinements, the contour progressively converges to a stable status that tightly encloses the object instance.

Joint Learning of Blind Super-Resolution and Crack Segmentation for Realistic Degraded Images

This article proposes crack segmentation augmented by super-resolution (SR) with deep neural networks. In the proposed method, an SR network is jointly trained with a binary segmentation network in an end-to-end manner. This joint learning allows the SR network to be optimized for improving segmentation results. For realistic scenarios, the SR network is extended from nonblind to blind for processing a low-resolution (LR) image degraded by unknown blurs. The joint network is improved by our proposed two extra paths that further encourage the mutual optimization between SR and segmentation. Comparative experiments with state-of-the-art (SoTA) segmentation methods demonstrate the superiority of our joint learning, and various ablation studies prove the effects of our contributions.

EPRO_IMG_022





EPRO_IMG_023

Light-weight Retinal Layer Segmentation with Global Reasoning

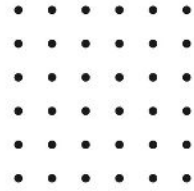
Automatic retinal layer segmentation with medical images, such as optical coherence tomography (OCT) images, serves as an important tool for diagnosing ophthalmic diseases. However, it is challenging to achieve accurate segmentation due to low contrast and blood flow noises presented in the images. In addition, the algorithm should be lightweight to be deployed for practical clinical applications. Therefore, it is desired to design a lightweight network with high performance for retinal layer segmentation. In this article, we propose LightReSeg for retinal layer segmentation which can be applied to OCT images. Specifically, our approach follows an encoder-decoder structure, where the encoder part uses multiscale feature extraction and a transformer block for fully exploiting the semantic information of feature maps at all scales and making the features have better global reasoning capabilities, while the decoder part, we design a multiscale asymmetric attention (MAA) module for preserving the semantic information at each encoder scale. The experiments show that our approach achieves a better segmentation performance compared with the current state-of-the-art method TransUnet with 105.7 M parameters on both our collected dataset and two other public datasets, with only 3.3 M parameters.

CM-MaskSD: Cross-Modality Masked\ Self-Distillation for Referring Image Segmentation

Referring image segmentation (RIS) is a fundamental vision-language task that intends to segment a desired object from an image based on a given natural language expression. Due to the essentially distinct data properties between image and text, most of existing methods either introduce complex designs towards fine-grained vision-language alignment or lack required dense alignment, resulting in scalability issues or mis-segmentation problems such as over- or under-segmentation. To achieve effective and efficient fine-grained feature alignment in the RIS task, we explore the potential of masked multimodal modeling coupled with self-distillation and propose a novel cross-modality masked self-distillation framework named CM-MaskSD, in which our method inherits the transferred knowledge of image-text semantic alignment from CLIP model to realize fine-grained patch-word feature alignment for better segmentation accuracy.

EPRO_IMG_024





EPRO_IMG_025

From Canteen Food to Daily Meals: Generalizing Food Recognition to More Practical Scenarios

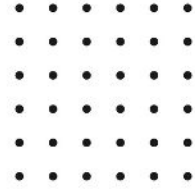
The precise recognition of food categories plays a pivotal role for intelligent health management, attracting significant research attention in recent years. Prominent benchmarks, such as Food-101 and VIREO Food-172, provide abundant food image resources that catalyze the prosperity of research in this field. Nevertheless, these datasets are well-curated from canteen scenarios and thus deviate from food appearances in daily life. This discrepancy poses great challenges in effectively transferring classifiers trained on these canteen datasets to broader daily-life scenarios encountered by humans. Toward this end, we present two new benchmarks, namely DailyFood-172 and DailyFood-16, specifically designed to curate food images from everyday meals. These two datasets are used to evaluate the transferability of approaches from the well-curated food image domain to the everyday-life food image domain.

Harvard Glaucoma Fairness: A Retinal Nerve Disease Dataset for Fairness Learning and Fair Identity Normalization

Fairness (also known as equity interchangeably) in machine learning is important for societal well-being, but limited public datasets hinder its progress. Currently, no dedicated public medical datasets with imaging data for fairness learning are available, though underrepresented groups suffer from more health issues. To address this gap, we introduce Harvard Glaucoma Fairness (Harvard-GF), a retinal nerve disease dataset including 3,300 subjects with both 2D and 3D imaging data and balanced racial groups for glaucoma detection. Glaucoma is the leading cause of irreversible blindness globally with Blacks having doubled glaucoma prevalence than other races. We also propose a fair identity normalization (FIN) approach to equalize the feature importance between different identity groups. Our FIN approach is compared with various state-of-the-art fairness learning methods with superior performance in the racial, gender, and ethnicity fairness tasks with 2D and 3D imaging data, demonstrating the utilities of our dataset Harvard-GF for fairness learning. To facilitate fairness comparisons between different models, we propose an equity-scaled performance measure, which can be flexibly used to compare all kinds of performance metrics in the context of fairness. The dataset and code are publicly accessible via <https://ophai.hms.harvard.edu/datasets/harvard-gf3300/>

EPRO_IMG_026





EPRO_IMG_027

A Review of Building Extraction From Remote Sensing Imagery: Geometrical Structures and Semantic Attributes

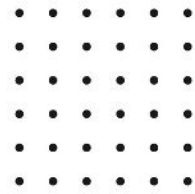
In the remote sensing community, extracting buildings from remote sensing imagery has triggered great interest. While many studies have been conducted, a comprehensive review of these approaches that are applied to optical and synthetic aperture radar (SAR) imagery is still lacking. Therefore, we provide an in-depth review of both early efforts and recent advances, which are aimed at extracting geometrical structures or semantic attributes of buildings, including building footprint generation, building facade segmentation, roof segment and superstructure segmentation, building height retrieval, building-type classification, building change detection, and annotation data correction. Furthermore, a list of corresponding benchmark datasets is given. Finally, challenges and outlooks of existing approaches as well as promising applications are discussed to enhance comprehension within this realm of research.

Enhancing Contrastive Learning With Positive Pair Mining for Few-Shot Hyperspectral Image Classification

In recent years, deep learning has emerged as the dominant approach for hyperspectral image (HSI) classification. However, deep neural networks require large annotated datasets to generalize well. This limits the applicability of deep learning for real-world HSI classification problems, as manual labeling of thousands of pixels per scene is costly and time consuming. In this article, we tackle the problem of few-shot HSI classification by leveraging state-of-the-art self-supervised contrastive learning with an improved view-generation approach. Traditionally, contrastive learning algorithms heavily rely on hand-crafted data augmentations tailored for natural imagery to generate positive pairs. However, these augmentations are not directly applicable to HSIs, limiting the potential of self-supervised learning in the hyperspectral domain. To overcome this limitation, we introduce two positive pair-mining strategies for contrastive learning on HSIs. The proposed strategies mitigate the need for high-quality data augmentations, providing an effective solution for few-shot HSI classification. Through extensive experiments, we show that the proposed approach improves accuracy and label efficiency on four popular HSI classification benchmarks. Furthermore, we conduct a thorough analysis of the impact of data augmentation in contrastive learning, highlighting the advantage of our positive pair-mining approach.

EPRO_IMG_028





EPRO_IMG_029

Mean-Weighted Collaborative Representation-Based Spatial-Spectral Joint Classification for Hyperspectral Images

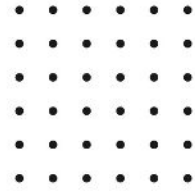
Collaborative representation (CR) models have been widely used in hyperspectral image (HSI) classification tasks. However, most CR classification models lack stability and generalization when targeting small samples as well as spatial homogeneity and heterogeneity problems. Therefore, this article proposes a mean-weighted CR classification model (MWCRC) based on the joint spatial-spectral data. It imposes mean and weighted constraints on the representation coefficients based on CR, which attenuates the noise effect and increases the distinguishability between classes. Second, a sample augmentation method based on the principle of minimizing the representation residuals is proposed. Sample augmentation is realized through initial classification and calculation of representation residuals to achieve the objective of consolidating model stability and improving classification accuracy.

Lossless Data Hiding in NTRU Cryptosystem by Polynomial Encoding and Modulation

Lossless data hiding in ciphertexts (LDH-CT) is to perform data embedding without changing their plaintexts, which can be used to transmit extra data in the applications of homomorphic encryption at little cost. In this paper, two LDH-CT algorithms named Polynomial Encoding (PE) and Polynomial Modulation (PM) are proposed for the "N-th Degree Truncated Polynomial Ring Unit" (NTRU) scheme, respectively. In the PE algorithm, a polynomial is encoded according to a string of bit values and further used to encrypt a plain-text polynomial. After decrypting the ciphertext, the encoded polynomial can be retrieved so that dozens of bit values can be extracted from it. Moreover, the PE algorithm can be combined with a polynomial partitioning strategy to achieve data extraction before decryption as well. In applying the PM algorithm, no parameter setting of an NTRU cryptosystem is changed while a cipher-text polynomial is generated by selectively sampling a polynomial to match the to-be-hidden value. Furthermore, the data hidden with the PM algorithm can be pre-chosen to be extracted without decryption or after decryption, and in each case up to 10 bit values can be hidden into one cipher-text polynomial.

EPRO_IMG_030





EPRO_IMG_031

An Interpretable and Accurate Deep-Learning Diagnosis Framework Modeled With Fully and Semi-Supervised Reciprocal Learning

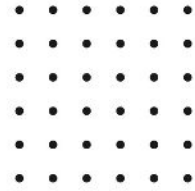
The deployment of automated deep-learning classifiers in clinical practice has the potential to streamline the diagnosis process and improve the diagnosis accuracy, but the acceptance of those classifiers relies on both their accuracy and interpretability. In general, accurate deep-learning classifiers provide little model interpretability, while interpretable models do not have competitive classification accuracy. In this paper, we introduce a new deep-learning diagnosis framework, called InterNRL, that is designed to be highly accurate and interpretable. InterNRL consists of a student-teacher framework, where the student model is an interpretable prototype-based classifier (ProtoPNet) and the teacher is an accurate global image classifier (GlobalNet). The two classifiers are mutually optimised with a novel reciprocal learning paradigm in which the student ProtoPNet learns from optimal pseudo labels produced by the teacher GlobalNet, while GlobalNet learns from ProtoPNet's classification performance and pseudo labels.

High-Frequency Space Diffusion Model for Accelerated MRI

Diffusion models with continuous stochastic differential equations (SDEs) have shown superior performances in image generation. It can serve as a deep generative prior to solving the inverse problem in magnetic resonance (MR) reconstruction. However, low-frequency regions of k -space data are typically fully sampled in fast MR imaging, while existing diffusion models are performed throughout the entire image or k -space, inevitably introducing uncertainty in the reconstruction of low-frequency regions. Additionally, existing diffusion models often demand substantial iterations to converge, resulting in time-consuming reconstructions. To address these challenges, we propose a novel SDE tailored specifically for MR reconstruction with the diffusion process in high-frequency space (referred to as HFS-SDE). This approach ensures determinism in the fully sampled low-frequency regions and accelerates the sampling procedure of reverse diffusion. Experiments conducted on the publicly available fastMRI dataset demonstrate that the proposed HFS-SDE method outperforms traditional parallel imaging methods, supervised deep learning, and existing diffusion models in terms of reconstruction accuracy and stability. The fast convergence properties are also confirmed through theoretical and experimental validation. Our code and weights are available at <https://github.com/Aboriginer/HFS-SDE>.

EPRO_IMG_032





EPRO_IMG_033

A Test Statistic Estimation-Based Approach for Establishing Self-Interpretable CNN-Based Binary Classifiers

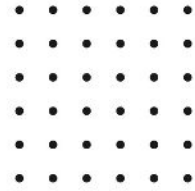
Interpretability is highly desired for deep neural network-based classifiers, especially when addressing high-stake decisions in medical imaging. Commonly used post-hoc interpretability methods have the limitation that they can produce plausible but different interpretations of a given model, leading to ambiguity about which one to choose. To address this problem, a novel decision-theory-inspired approach is investigated to establish a self-interpretable model, given a pre-trained deep binary black-box medical image classifier. This approach involves utilizing a self-interpretable encoder-decoder model in conjunction with a single-layer fully connected network with unity weights. The model is trained to estimate the test statistic of the given trained black-box deep binary classifier to maintain a similar accuracy. The decoder output image, referred to as an equivalency map, is an image that represents a transformed version of the to-be-classified image that, when processed by the fixed fully connected layer, produces the same test statistic value as the original classifier.

Multi-view learning for automatic classification of multi-wavelength auroral images

Auroral classification plays a crucial role in polar research. However, current auroral classification studies are predominantly based on images taken at a single wavelength, typically 557.7 nm. As a result, the integration of information from multiple wavelengths has received comparatively less attention, resulting in low classification rates for complex auroral patterns. Furthermore, existing studies employing traditional machine learning or deep learning approaches have not achieved an optimal balance between accuracy and speed. To overcome these challenges, this article proposes a lightweight auroral multiwavelength fusion classification network, MLCNet, based on a multiview approach. First, we develop a lightweight feature extraction backbone to improve the classification rate and effectively cope with the increasing amount of auroral observation data. Second, considering the existence of multiscale spatial structures in auroras, we design a novel multiscale reconstructed feature module. Finally, to highlight the discriminative information between auroral classes, we propose a lightweight attention feature enhancement (LAFE) module. The proposed method is validated using auroral observations from the Arctic Yellow River Station (YRS) during 2003–2004.

EPRO_IMG_034





EPRO_IMG_035

Time Efficient Ultrasound Localization Microscopy Based on A Novel Radial Basis Function 2D Interpolation

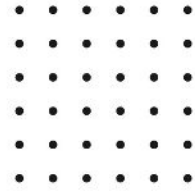
Ultrasound localization microscopy (ULM) allows for the generation of super-resolved (SR) images of the vasculature by precisely localizing intravenously injected microbubbles. Although SR images may be useful for diagnosing and treating patients, their use in the clinical context is limited by the need for prolonged acquisition times and high frame rates. The primary goal of our study is to relax the requirement of high frame rates to obtain SR images. To this end, we propose a new time-efficient ULM (TEULM) pipeline built on a cutting-edge interpolation method. More specifically, we suggest employing Radial Basis Functions (RBFs) as interpolators to estimate the missing values in the 2-dimensional (2D) spatio-temporal structures. To evaluate this strategy, we first mimic the data acquisition at a reduced frame rate by applying a down-sampling (DS = 2, 4, 8, and 10) factor to high frame rate ULM data. Then, we up-sample the data to the original frame rate using the suggested interpolation to reconstruct the missing frames.

Guided Filter of Random Patches Network and Relaxed-Collaborative-Representation-Based Hyperspectral Image

Feature extraction and accurate classification are crucial tasks in the land-cover classification of the hyperspectral image (HSI). We propose a guided filter (GF) of a random patches network (RPNet) and a relaxed collaborative representation (RCR)-based HSI classification (HSIC) method called GRR. The shallow and deep features are extracted using RPNet that requires no pretraining stage. In addition to the obtained feature set, the original HSI and extracted features are then filtered by GF to preserve the edge details. After that, all the distinct feature sets are separately concatenated with the original HSI to keep the original structure of the data. The high-dimensional feature sets are then processed by a linear discriminant analysis (LDA) to increase class separability and to select the most representative features. Since few train samples are available in the HSIC task, the efficiency of LDA is improved using superpixel segmentation to generate pseudosamples. In the final stage, the reduced-dimension feature sets are classified by the use of superpixel-guided RCR, which utilizes the resemblance and discrimination of the feature sets efficiently. The extensive experiments on the real HSIs are carried out to validate the efficacy of the proposed method.

EPRO_IMG_036





EPRO_IMG_037

Active Learning-Based Spectral–Spatial Classification for Discriminating Tree Species in Hyperspectral Images

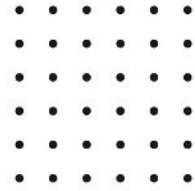
Exploiting spectral–spatial information and reducing the number of required training samples are important for improving tree species classification performance in hyperspectral images. In this article, an active learning-based spectral–spatial classification (ALSSC) model is proposed to reduce the demand for training samples while improving the classification performance. To improve classification performance, the proposed ALSSC employs two ways to exploit spectral–spatial information within the hyperspectral image: 1) features used in classification are extracted from multiscale superpixels; 2) the classification result is refined by guided filtering and subsequently employed as the input for the next round of classification. To reduce the demand for training samples, after each round of classification, active learning (AL) is adopted to select the most informative samples from the unlabeled testing set to enrich the training set. To validate the effectiveness of the proposed ALSSC, experiments are conducted using a tree species classification dataset collected by an airborne hyperspectral sensor.

Protocol-Based Distributed Security Fusion Estimation for Time-Varying Uncertain Systems Over Sensor Networks: Tackling DoS

This article studies the distributed fusion estimation (DFE) issue for networked multi-sensor systems (NMSSs) with stochastic uncertainties, bandwidth-constrained network and energy-constrained denial-of-service (DoS) attacks. The stochastic uncertainties reflected in both the state and measurement models are characterized by multiplicative noises. For reducing the communication burden, local estimation signals are subject to dimensionality reduction processing. And the improved Round-Robin (RR) protocol is used on the channels from local estimators to the fusion estimator. To reflect the actual situation, the dimensionality reduction strategy is designed from the defender's point of view in the sense of minimum fusion error covariance (FEC). And the attack strategy is designed from the attacker's point of view in the sense of maximum FEC. Then, based on a compensation model, a recursive distributed Kalman fusion estimation algorithm (DKFEA) is proposed. The stability conditions making the mean square error (MSE) for DFE bounded are derived. In the end, the validity of the presented DKFEA is verified by an illustrative example.

EPRO_IMG_038





EPRO_IMG_039

Toward Robust and Unconstrained Full Range of Rotation Head Pose Estimation

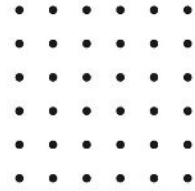
Estimating the head pose of a person is a crucial problem for numerous applications that is yet mainly addressed as a subtask of frontal pose prediction. We present a novel method for unconstrained end-to-end head pose estimation to tackle the challenging task of full range of orientation head pose prediction. We address the issue of ambiguous rotation labels by introducing the rotation matrix formalism for our ground truth data and propose a continuous 6D rotation matrix representation for efficient and robust direct regression. This allows to efficiently learn full rotation appearance and to overcome the limitations of the current state-of-the-art. Together with new accumulated training data that provides full head pose rotation data and a geodesic loss approach for stable learning, we design an advanced model that is able to predict an extended range of head orientations. An extensive evaluation on public datasets demonstrates that our method significantly outperforms other state-of-the-art methods in an efficient and robust manner, while its advanced prediction range allows the expansion of the application area. We open-source our training and testing code along with our trained models: <https://github.com/thohemp/6DRepNet360> .

Field-to-Field Coordinate-Based Segmentation Algorithm on Agricultural Harvest Implements

Establishing and maintaining farmland geometric boundaries is crucial to increasing agricultural productivity. Accurate field boundaries enable farm machinery contractors and other farm stakeholders to calculate charges, costs and to examine machinery performance. Field segmentation is the process by which agricultural field plots are geofenced into their individual field geometric boundaries. This paper presents a novel coordinate-based method to perform trajectory segmentation and field boundary detection from a tractor towing an implement. The main contribution of this research is a practical, robust algorithm which can solve for challenging field-to-field segmentation cases where the operator engages the towed implement continuously across several fields. The algorithm first isolates raw machinery trajectory data into unique job sites by using a coarse filter on geolocation data and implement power-take off activation. Next, the coordinate data is plotted and image processing techniques are applied to erode any pathway(s) that may present in job sites with adjacent working fields.

EPRO_IMG_040





EPRO_IMG_041

Ensemble Image Explainable AI (XAI) Algorithm for Severe Community-Acquired Pneumonia and COVID-19 Respiratory

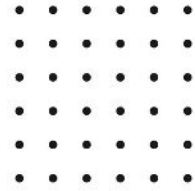
Since the onset of the COVID-19 pandemic in 2019, many clinical prognostic scoring tools have been proposed or developed to aid clinicians in the disposition and severity assessment of pneumonia. However, there is limited work that focuses on explaining techniques that are best suited for clinicians in their decision making. In this article, we present a new image explainability method named ensemble AI explainability (XAI), which is based on the SHAP and Grad-CAM++ methods. It provides a visual explanation for a deep learning prognostic model that predicts the mortality risk of community-acquired pneumonia and COVID-19 respiratory infected patients. In addition, we surveyed the existing literature and compiled prevailing quantitative and qualitative metrics to systematically review the efficacy of ensemble XAI, and to make comparisons with several state-of-the-art explainability methods (LIME, SHAP, saliency map, Grad-CAM, Grad-CAM++). Our quantitative experimental results have shown .

Edge-Computing-Enabled Deep Learning Approach for Low-Light Satellite Image Enhancement

Edge computing enables rapid data processing and decision-making on satellite payloads. Deploying deep learning-based techniques for low-light image enhancement improves early detection and tracking accuracy on satellite platforms, but it faces challenges due to limited computational resources. This article proposes an edge-computing-enabled inference model specifically designed onboard satellites. The proposed model follows an encoder-decoder architecture to generate the illumination map with low multiplication matrix complexity, 25.52 GMac of 1920*1200 image size. To reduce nanosatellite hardware consumption with a single-precision floating-point format, the edge-computing-enabled inference model proposes a quantized convolution that computes signed values. The proposed inference model is deployed on Arm Cortex-M3 microcontrollers onboard satellite payload (86.74 times faster than normal convolution model) but also has a similar quality with the low-light enhanced in full-precision computing of lightweight training model by using the peak signal-to-noise ratio (average of 28.94) and structural similarity index (average of 0.85) metrics.

EPRO_IMG_042





EPRO_IMG_043

An Efficient Method for Detecting Dense and Small Objects in UAV Image

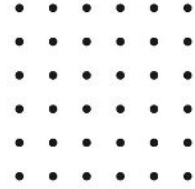
Object detection in unmanned aerial vehicle (UAV) images is an important and challenging task for many applications, which often needs highly efficient detection algorithms to meet the accuracy and real-time requirements of the applications. In this article, we investigate efficient mechanisms for detecting dense and small objects in UAV images. Specifically, 1) kernel K-means is used to obtain optimal anchors for dense and small object detection; 2) a spatial information enhancement module is proposed to improve the detection accuracy of dense objects by extracting object spatial location information; 3) a Coord_C3 module is proposed to improve the receptive field of the network and to reduce the number of network parameters; and 4) a small detection head is added in the Head of the network and skip connections are employed in the Neck of the network to improve the detection accuracy of small objects. Experimental results on the VisDrone-2019, LEVIR-ship, and Stanford Drone datasets show that our method not only has higher detection accuracy but also runs faster compared to state-of-the-art detection methods.

Conventional to Deep Ensemble Methods for Hyperspectral Image Classification: A Comprehensive Survey

Hyperspectral image classification (HSIC) has become a hot research topic. Hyperspectral imaging (HSI) has been widely used in a wide range of real-world application areas due to the in-depth spectral information stored within each pixel. Noticeably, the detailed features, i.e., a nonlinear correlation between the obtained spectral data and the correlating HSI data object, generate efficient classification results that are complex for traditional techniques. Deep learning (DL) has recently been validated as an influential feature extractor that efficiently identifies the nonlinear issues that have arisen in various computer vision challenges. This motivates using DL for HSIC, which shows promising results. This survey provides a brief description of DL for HSIC and compares cutting-edge methodologies in the field. We will first summarize the key challenges for HSIC, and then, we will discuss the superiority of DL and DL ensemble in addressing these issues. In this article, we divide state-of-the-art DL methodologies and DL with ensemble into spectral features, spatial features, and combined spatial-spectral features in order to comprehensively and critically evaluate the progress

EPRO_IMG_044





EPRO_IMG_045

Deep Evidential Remote Sensing Landslide Image Classification With a New Divergence, Multiscale Saliency and an Improved

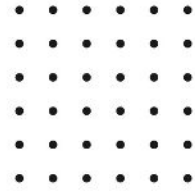
Hitherto, image-level classification on remote sensing landslide images has been paid attention to, but the accuracy of traditional deep learning-based methods still have room for improvement. The evidence theory is found efficient to boost the accuracy of neural networks, however, the present study argues three challenges that hinder the lead-in of this theory in deep landslide image classification. Aiming at the three problems, this study makes three improvements. For the interpretability and decision-invariance losses of three previous divergences, we propose a belief Jensen–Renyi divergence with properties proven. To couple the evidence theory with deep remote sensing landslide image classification, a channelwise multiscale visual saliency fusion is developed. We additionally find that the channelwise fusion is capable to reduce false recognition of networks as compared with original RGB images. To avoid decision failures in evidence-theoretic fusion process, we design an interpretability improved three-branched fusion.

FsrGAN: A Satellite and Radar-Based Fusion Prediction Network for Precipitation Nowcasting

Precipitation nowcasting refers to the prediction of small-scale precipitation events at minute and kilometer scales within the upcoming 0 to 2 h, which significantly impacts both human activities and daily life. However, prevailing deep learning models have primarily focused on a single radar echo data source, limiting their ability to effectively capture intricate and rapidly evolving precipitation patterns. Thus, meteorological satellite is considered to supplement radar echo data. To achieve a comprehensive integration of multisource data with enhanced details, a two-stage fusion satellite and radar GAN-based prediction network (named FsrGAN) is proposed. In the first stage, we design a satellite-radar fusion prediction network known as FsrNet. This network employs an encoder-fusion-decoder architecture, where a novel spatial-channel attention (SCA) is proposed to enhance the filtering and fusion of multisource and multiscale features. In the second stage, we introduce a GAN-based network (FusionGAN) that also mines the complementary information of satellite images to sharpen the first-stage predicted radar maps with more details. Experiments are conducted on meteorological dataset in the Yangtze River Delta (YRD) region.

EPRO_IMG_046





EPRO_IMG_047

Unsupervised Human Activity Recognition Learning for Disassembly Tasks

Large volumes of used electronics are often collected in remanufacturing plants, which requires disassembly before harvesting parts for reuse. Disassembly is mainly conducted manually with low productivity. Recently, human-robot collaboration has been considered as a solution. To assist effectively, robots should observe work environments and recognize human actions accurately. Rich activity video recording and supervised learning can be used to extract insights; however, supervised learning does not allow robots to self-accomplish the learning process. This study proposes an unsupervised learning framework for achieving video-based human activity recognition. The framework consists of two main elements: 1) a variational-autoencoder-based architecture for unlabeled data representation learning and 2) a hidden Markov model for activity state division. The complete explicit activity classification is validated against ground truth labels; here, we use a case study of disassembling a hard disk drive. The framework shows an average recognition accuracy of 91.52%, higher than competing methods.

Exploring Deep Learning and Machine Learning Approaches for Brain Hemorrhage Detection

Brain hemorrhage refers to a potentially fatal medical disorder that affects millions of individuals. The percentage of patients who survive can be significantly raised with the prompt identification of brain hemorrhages, due to image-guided radiography, which has emerged as the predominant treatment modality in clinical practice. A Computed Tomography Image has frequently been employed for the purpose of identifying and diagnosing neurological disorders. The manual identification of anomalies in the brain region from the Computed Tomography Image demands the radiologist to devote a greater amount of time and dedication. In the most recent studies, a variety of techniques rooted in Deep learning and traditional Machine Learning have been introduced with the purpose of promptly and reliably detecting and classifying brain hemorrhage. This overview provides a comprehensive analysis of the surveys that have been conducted by utilizing Machine Learning and Deep Learning. This research focuses on the main stages of brain hemorrhage, which involve preprocessing, feature extraction, and classification, as well as their findings and limitations. Moreover, this in-depth analysis provides a description of the existing benchmark datasets that are utilized for the analysis of the detection process.

EPRO_IMG_048





EPRO_IMG_049

A Novel Grading Biomarker for the Prediction of Conversion from Mild Cognitive Impairment to Alzheimer's Disease

Objective: Identifying mild cognitive impairment (MCI) subjects who will progress to Alzheimer's disease (AD) is not only crucial in clinical practice, but also has a significant potential to enrich clinical trials. The purpose of this study is to develop an effective biomarker for an accurate prediction of MCI-to-AD conversion from magnetic resonance images. **Methods:** We propose a novel grading biomarker for the prediction of MCI-to-AD conversion. First, we comprehensively study the effects of several important factors on the performance in the prediction task including registration accuracy, age correction, feature selection, and the selection of training data. Based on the studies of these factors, a grading biomarker is then calculated for each MCI subject using sparse representation techniques. Finally, the grading biomarker is combined with age and cognitive measures to provide a more accurate prediction of MCI-to-AD conversion.

Machine Learning Analysis of Human Skin by Optoacoustic Mesoscopy for Automated Extraction of Psoriasis and Aging Biomarkers

Ultra-wideband raster-scan optoacoustic mesoscopy (RSOM) is a novel modality that has demonstrated unprecedented ability to visualize epidermal and dermal structures in-vivo. However, an automatic and quantitative analysis of three-dimensional RSOM datasets remains unexplored. In this work we present our framework: Deep Learning RSOM Analysis Pipeline (DeepRAP), to analyze and quantify morphological skin features recorded by RSOM and extract imaging biomarkers for disease characterization. DeepRAP uses a multi-network segmentation strategy based on convolutional neural networks with transfer learning. This strategy enabled the automatic recognition of skin layers and subsequent segmentation of dermal microvasculature with an accuracy equivalent to human assessment. DeepRAP was validated against manual segmentation on 25 psoriasis patients under treatment and our biomarker extraction was shown to characterize disease severity and progression well with a strong correlation to physician evaluation and histology. In a unique validation experiment, we applied DeepRAP in a time series sequence of occlusion-induced hyperemia from 10 healthy volunteers.

EPRO_IMG_050





2.1L+ Customer Satisfied | **27+** Specialized Domains | **107+** Campus Tie-ups | **24/7** Support Center | **57+** Countries Covered | **2600+** Journal Access

📞 99447 93398

✉ info@elysiumpro.in

🌐 elysiumpro.in

📍 229, First Floor, A Block,
Elysium Campus,
Church Rd, Anna Nagar,
Madurai, Tamil Nadu -
625020.

Delivery Centres

Chennai | Coimbatore | Tirunelveli | Virudhunagar | Trichy

