



ELYSIUMPRO

A UNIT OF ELYSIUM GROUPS

FINAL YEAR PROJECT

DIGITAL IMAGE PROCESSING 2019-2020

TITLES WITH ABSTRACTS



CALL US @

(+91) 9944 7933 98 | (+91) 452 - 424 2842, 424 2843

20 Years of Experience | Automated Services | 24/7 Help Desk Support
Advanced Technologies and Tools | Legitimate Members of all Journals
Quality Project Training | Industry Exposure



ELYSIUMPRO
INSPIRING THE LEADING EDGE TECHNOLOGIES

www.elysiumpro.in



#227, Elysium Campus, Church Road, Anna Nagar,
Madurai - 625020, Tamil Nadu, India

CALL US @

(+91) 9944 7933 98 | (+91) 452 - 424 2842, 424 2843

[f /ElysiumPro Project Center](#)

[t /ElysiumPro](#)

[in /ElysiumPro](#)

Elysium PRO



Titles with Abstracts 2019-20

**EPRO DIP
- 001**

Bilateral Two-Dimensional Matrix Regression Preserving Discriminant Embedding for Corrupted Image Recognition

Nuclear-norm-based matrix regression (NMR) methods have been successfully applied for the recognition of corrupted images. However, most of these methods do not consider the label information and are classified as unsupervised learning methods. In this paper, we propose a new regression-based algorithm, named bilateral two-dimensional matrix regression preserving discriminant embedding (B2DMRPDE). The proposed algorithm constructs the within-reconstruction graph and between-reconstruction graph using NMR. Then, B2DMRPDE aims to seek a subspace in which the within-class reconstructive residual is minimized and the between-class reconstructive residual is maximized based on Fisher's criterion. Hence, B2DMRPDE can capture the potential discriminative information for classification. To enhance the classification effectiveness, we present a new NMR-based classifier to determine the class label of the testing sample. Extensive experiments on face image databases were performed, and the results validate the effectiveness of the proposed method.

**EPRO DIP
- 002**

A Privacy-Preserving Edge Computation-Based Face Verification System for User Authentication

The face recognition has become a common means of identity authentication because of the advantages of uniqueness, non-invasive and not easy to be stolen. The outsourcing of face recognition to the service provider is a typical manner nowadays. However, it raises vital concerns about the privacy of outsourcing server due to the sensitivity of face data. Therefore, a frame of identity authentication based on the technology of privacy-preserving face recognition is presented in this paper. The convolutional neural network is used for face feature extraction. To overcome the issue of privacy leaked, a method of secure nearest neighbor that can compute the cosine similarity over encrypted feature vectors is proposed. What's more, the edge computing is introduced in our frame to increase the authentication efficiency by removing some operations from the cloud to the edge of the Internet. Moreover, we also propose a secret sharing homomorphism technology which is used for distributed computing to improve the fault-tolerance of our identity authentication system. The experimental results show that the proposed schemes are secure and effective.

**EPRO DIP
- 003**

A Unified Variational Model for Single Image Dehazing

Haze is a common weather phenomenon, which hinders many outdoor computer vision applications such as outdoor surveillance, navigation control, vehicle driving, and so on. In this paper, a simple but effective unified variational model for single image dehazing is presented based on the total variation regularization. From the perspective of the relationship between image dehazing and Retinex, the dehazing problem can be formulated as the minimization of a variational Retinex model. The proposed variational model incorporates two ℓ_1 -norm regularization terms to constrain the scene transmission and the inverted scene radiance respectively, which can be better applied into image dehazing field. Different from the conventional two-step framework, our proposed model can simultaneously obtain the accurate transmission map and the recovered scene radiance by integrating the transmission estimation stage and the image recovery stage into the unified variational model. The entire optimization of the proposed unified variational model can be solved by an alternating direction minimization scheme. The experiments on various simulated and real-world hazy images indicate that the proposed algorithm can yield considerably promising results comparative to several state-of-the-art dehazing and enhancement techniques.

**EPRO DIP
- 004**

Robust Median Filtering Forensics Using Image Deblocking and Filtered Residual Fusion

Median filtering (MF) is frequently applied to conceal the traces of forgery and therefore can provide indirect forensic evidence of tampering when investigating composite images. The existing MF forensic methods, however, ignore how JPEG compression affects median filtered images, resulting in heavy performance degradation when detecting filtered images stored in the JPEG format. In this paper, we propose a new robust MF forensic method based on a modified convolutional neural network (CNN). First, relying on the analysis of the influence on median filtered images caused by JPEG compression, we effectively suppress the interference using image deblocking. Second, the fingerprints left by MF are highlighted via filtered residual fusion. These two functions are fulfilled with a deblocking layer and a fused filtered residual (FFR) layer. Finally, the output of the FFR layer becomes input when extracting multiple features for further classification using a tailor-made CNN. The extensive experimental results show that the proposed method outperforms the state-of-the-art methods in both JPEG compressed and small-sized MF image detection.

EPRO DIP
- 005

Single Image Snow Removal via Composition Generative Adversarial Networks

Snowflakes attached to the camera lens can severely affect the visibility of the background scene and compromise the image quality. In this paper, we solve this problem by visually removing snowflakes to convert the snowy image into a clean one. The problem is troublesome; the information about the background of the occluded regions is completely lost for the most part. For removing snowflakes from a single image, we proposed a composition generative adversarial network. Different from the previous generative adversarial networks, our generator network comprises clean background module and a snow mask estimate module. The clean background module aims to generate a clear image from an input snowy image, and snow mask estimate module is used to produce the snow mask in an input image. During the training step, we put forward a composition loss between the input snowy image and composition of the generated clean image and estimated snow mask. We use a dataset named Snow100K² including indoor and outdoor scenes to train and test the proposed method. The extensive experiments on both synthetic and real-world images show that our network has a good effect and it is superior to the other state-of-the-art methods.

EPRO DIP
- 006

Weighting Quantization Matrices for HEVC/H.265-Coded RGB Videos

In the HEVC/H.265 video coding standard, weighting quantization matrices (WQMs) are supported to take advantage of the characteristics of the human visual system (HVS). However, the default WQMs utilized in HEVC are developed for YCbCr videos instead of RGB videos. In this paper, a set of new WQMs is proposed for video coding in RGB color space. First, we utilize the spatial contrast sensitivity function (CSF) to model the bandpass property of HVS. To derive the parameters of the spatial CSF, a series of subjective experiments is conducted to obtain the just-noticeable distortion (JND) thresholds of several selected DCT subbands. In addition, the sensitivities of different DCT subbands in one color channel, as well as among R, G, and B channels, are considered to design the WQMs of intra-coded 8×8 blocks. Moreover, to reduce the data size of WQMs, the WQMs for other block sizes are derived from intra 8×8 WQMs. The proposed WQMs are then applied into HEVC to directly code RGB videos. The experimental results demonstrate that when the PSNRs of G, B, and R channels are combined with a ratio of 4:1:1, the proposed WQMs can achieve an average BD-rate saving of 12.64% and 20.51%, respectively, in all-intra (AI) and low-delay (LD) profiles compared to HEVC without WQMs. The proposed scheme also enjoys a better video quality metric (VQM) performance.

EPRO DIP
- 007

An Effective Gaussian Fitting Approach for Image Contrast Enhancement

Contrast enhancement plays an important role in image processing applications. This paper proposes a low-complexity automatic method for contrast enhancement. The method exploits the high-frequency distribution of an image to estimate an intensity-weighting matrix, which is then used to control the Gaussian fitting curve and shape the distribution of the contrast gain. As such, the curve can be easily designed to enhance the important details hidden in noteworthy regions. Subsequently, the proposed grayscale transformation that is obtained from the Gaussian fitting can rationally express contrast distribution. Unlike previous contrast enhancement methods, our technique is fully automatic and target-oriented in the sense that it can be directly applied to any input image without any parameter adjustment. The experimental results for some of the widely accepted criteria demonstrate the superiority of our proposed method over the conventional enhancement techniques, especially in the aspects of visual pleasure, anti-noise capability, and target-oriented contrast enhancement.

EPRO DIP
- 008

Sheep Identification Using a Hybrid Deep Learning and Bayesian Optimization Approach

Sheep are considered a necessary source of food production worldwide. Therefore, the sheep identification is vital for managing breeding and disease. Moreover, it is the only guarantee of an individual's ownership. Therefore, in this paper, sheep identities were recognized by a deep convolutional neural network using facial bio-metrics. To obtain the best possible accuracy, different neural networks designs were surveyed and tested in this paper. The Bayesian optimization was used to automatically set the parameters for a convolutional neural network; in addition, the AlexNet configuration was also examined in this paper. In this paper, the sheep recognition algorithms were tested on a data set of 52 sheep. Not more than 10 images were taken of each sheep in different postures. Thus, the data augmentation methodologies such as rotation, reflection, scaling, blurring, and brightness modification were applied; 1000 images of each sheep were obtained for training and validation. The experiments conducted in this paper achieved an accuracy of 98%. Our approach outperforms previous approaches for sheep identification.

EPRO DIP
- 009

Building Recognition Based on Sparse Representation of Spatial Texture and Color Features

In this paper, we presented a novel building recognition method based on a sparse representation of spatial texture and color features. At present, the most popular methods are based on gist features, which can only roughly reflect the spatial information of building images. The proposed method, in contrast, uses multi-scale neighborhood sensitive histograms of oriented gradient (MNSHOGs) and color auto-correlogram (CA) to extract texture and color features of building images. Both the MNSHOG and the CA take spatial information of building images into account while calculating texture and color features. Then, color and texture features are combined to form joint features whose sparse representation can be dimensionally reduced by an autoencoder. Finally, an extreme learning machine is used to classify the combined features after dimensionality reduction into different classes. Several experiments were conducted on the benchmark Sheffield building dataset. The mean recognition rate of our method is much higher than that of the existing methods, which shows the effectiveness of our method

EPRO DIP
- 010

An Image Encryption Method Based on Elliptic Curve ElGamal Encryption and Chaotic Systems

Due to the potential security problem about key management and distribution for the symmetric image encryption schemes, a novel asymmetric image encryption method is proposed in this paper, which is based on the elliptic curve ElGamal (EC-ElGamal) cryptography and chaotic theory. Specifically, the SHA-512 hash is first adopted to generate the initial values of a chaotic system, and a crossover permutation in terms of chaotic index sequence is used to scramble the plain-image. Furthermore, the generated scrambled image is embedded into the elliptic curve for the encrypted by EC-ElGamal which can not only improve the security but also can help solve the key management problems. Finally, the diffusion combined chaos game with DNA sequence is executed to get the cipher image. The experimental analysis and performance comparisons demonstrate that the proposed method has high security, good efficiency, and strong robustness against the chosen-plaintext attack which make it have potential applications for the image secure communications.

**EPRO DIP
- 011**

Region-of-Interest Compression and View Synthesis for Light Field Video Streaming

Light field videos provide a rich representation of real-world, thus the research of this technology is of urgency and interest for both the scientific community and industries. Light field applications such as virtual reality and post-production in the movie industry require a large number of viewpoints of the captured scene to achieve an immersive experience, and this creates a significant burden on light field compression and streaming. In this paper, we first present a light field video dataset captured with a plenoptic camera. Then a new region-of-interest (ROI)-based video compression method is designed for light field videos. In order to further improve the compression performance, a novel view synthesis algorithm is presented to generate arbitrary viewpoints at the receiver. The experimental evaluation of four light field video sequences demonstrates that the proposed ROI-based compression method can save 5%-7% in bitrates in comparison to conventional light field video compression methods. Furthermore, the proposed view synthesis-based compression method not only can achieve a reduction of about 50% in bitrates against conventional compression methods, but the synthesized views can exhibit identical visual quality as their ground truth.

**EPRO DIP
- 012**

Human Action Recognition Using Multilevel Depth Motion Maps

The advent of depth sensors opens up new opportunities for human action recognition by providing depth information. The main purpose of this paper is to present an effective method for human action recognition from depth images. A multilevel frame select sampling (MFSS) method are proposed to generate three levels of temporal samples from the input depth sequences first. Then, the proposed motion and static mapping (MSM) method is used to obtain the representation of MFSS sequences. After that, this paper exploits the block-based LBP feature extraction approach to extract features information from the MSM. Finally, the fisher kernel representation is applied to aggregate the block features, which is then combined with the kernel-based extreme learning machine classifier. The developed framework is evaluated on three public datasets captured by depth cameras. The experimental results demonstrate the great performance compared with the existing approaches.

**EPRO DIP
- 013**

Perceptual Image Hashing Based on Weber Local Binary Pattern and Color Angle Representation

This paper proposes an efficient scheme for generating image hashing by combining the local texture and color angle features. During the stage of texture extraction, using Weber's Law, the difference ratios between the center pixels and their surrounding pixels are calculated and the dimensions of these values are further reduced by applying principal component analysis to the statistical histogram. In the stage of color feature extraction, the color angle of each pixel is computed before dimensional reduction and is carried out using a discrete cosine transform and a significant coefficients selection strategy. The main contribution of this paper is a novel construction for image hashing that incorporates texture and color features by using Weber local binary pattern and color angular pattern. The experimental results demonstrate the efficacy of the proposed scheme, especially for the perceptual robustness against common content-preserving manipulations, such as the JPEG compression, Gaussian low-pass filtering, and image scaling. Based on the comparisons with the state-of-the-art schemes, receiver operating characteristic curves and integrated histograms of normalized distances show the superiority of our scheme in terms of robustness and discrimination.

**EPRO DIP
- 014**

Reversible Data Hiding With Image Enhancement Using Histogram Shifting

Traditional reversible data hiding (RDH) focuses on enlarging the embedding payloads while minimizing the distortion with a criterion of mean square error (MSE). Since imperceptibility can also be achieved via image processing, we propose a novel method of RDH with contrast enhancement (RDH-CE) using histogram shifting. Instead of minimizing the MSE, the proposed method generates marked images with good quality with the sense of structural similarity. The proposed method contains two parts: the baseline embedding and the extensive embedding. In the baseline part, we first merge the least significant bins to reserve spare bins and then embed additional data by a histogram shifting approach using arithmetic encoding. During histogram shifting, we propose to construct the transfer matrix by maximizing the entropy of the histogram. After embedding, the marked image containing additional data has a larger contrast than the original image. In the extensive embedding part, we further propose to concatenate the baseline embedding with an MSE-based embedding. On the recipient side, the additional data can be extracted exactly, and the original image can be recovered losslessly. Comparing with existing RDH-CE approaches, the proposed method can achieve a better embedding payload.

**EPRO DIP
- 015**

Deep Image Compression in the Wavelet Transform Domain Based on High Frequency Sub-Band Prediction

In this paper, we propose to use deep neural networks for image compression in the wavelet transform domain. When the input image is transformed from the spatial pixel domain to the wavelet transform domain, one low-frequency sub-band (LF sub-band) and three high-frequency sub-bands (HF sub-bands) are generated. Low-frequency sub-band is firstly used to predict each high-frequency sub-band to eliminate redundancy between the sub-bands, after which the sub-bands are fed into different auto-encoders to do the encoding. In order to further improve the compression efficiency, we use a conditional probability model to estimate the context-dependent prior probability of the encoded codes, which can be used for entropy coding. The entire training process is unsupervised, and the auto-encoders and the conditional probability model are trained jointly. The experimental results show that the proposed approach outperforms JPEG, JPEG2000, BPG, and some mainstream neural network-based image compression. Furthermore, it produces better visual quality with clearer details and textures because more high-frequency coefficients can be reserved, thanks to the high-frequency prediction.

**EPRO DIP
- 016**

Robust Human Activity Recognition Using Multimodal Feature-Level Fusion

Automated recognition of human activities or actions has great significance as it incorporates wide-ranging applications, including surveillance, robotics, and personal health monitoring. Over the past few years, many computer vision-based methods have been developed for recognizing human actions from RGB and depth camera videos. These methods include space-time trajectory, motion encoding, key poses extraction, space-time occupancy patterns, depth motion maps, and skeleton joints. However, these camera-based approaches are affected by background clutter and illumination changes and applicable to a limited field of view only. Wearable inertial sensors provide a viable solution to these challenges but are subject to several limitations such as location and orientation sensitivity. Due to the complementary trait of the data obtained from the camera and inertial sensors, the utilization of multiple sensing modalities for accurate recognition of human actions is gradually increasing. This paper presents a viable multimodal feature-level fusion approach for robust human action recognition, which utilizes data from multiple sensors, including RGB camera, depth sensor, and wearable inertial sensors. We extracted the computationally efficient features from the data obtained from RGB-D video camera and inertial body sensors.

EPRO DIP
- 017

Curvature Bag of Words Model for Shape Recognition

The object shape recognition of nonrigid transformations and local deformations is a difficult problem. In this paper, a shape recognition algorithm based on the curvature bag of words (CBoW) model is proposed to solve that problem. First, an approximate polygon of the object contour is obtained by using the discrete contour evolution algorithm. Next, based on the polygon vertices, the shape contour is decomposed into contour fragments. Then, the CBoW model is used to represent the contour fragments. Finally, a linear support vector machine is applied to classify the shape feature descriptors. Our main innovations are as follows: 1) A multi-scale curvature integral descriptor is proposed to extend the representativeness of the local descriptor; 2) The curvature descriptor is encoded to break through the limitation of the correspondence relationship of the sampling points for shape matching, and accordingly it forms the feature of middle-level semantic description; 3) The equal-curvature integral ranking pooling is employed to enhance the feature discrimination, and also improves the performance of the middle-level descriptor. The experimental results show that the recognition rate of the proposed algorithm in the MPEG-7 database can reach 98.21%. The highest recognition rates of the Swedish Leaf and the Tools databases are 97.23% and 97.14%, respectively. The proposed algorithm achieves a high recognition rate and has good robustness, which can be applied to the target shape recognition field for nonrigid transformations and local deformations.

EPRO DIP
- 018

Color Image Compression-Encryption Algorithm Based on Fractional-Order Memristor Chaotic Circuit

In this paper, a fractional-order memristive chaotic circuit system is defined according to memristor circuit. The dynamic characteristics are analyzed through the phase diagram, bifurcation diagram, and Lyapunov exponent spectrum, and the randomness of the chaotic pseudo-random sequence is tested by NIST SP800-22. Based on this fractional-order memristive chaotic circuit, we propose a novel color image compression-encryption algorithm. In this algorithm, compression sensing (CS) algorithm is used for compression image, and then using Zigzag confusion, add modulus and BitCircShift diffuse encrypt the image. The theoretical analysis and simulation results indicate that the proposed compression and encryption scheme has good compression performance, reconstruction effect, and higher safety performance. Moreover, it also shows that the new algorithm facilitates encryption, storage, and transmission of image information in practical applications.

**EPRO DIP
- 019**

Exploiting EEG Signals and Audiovisual Feature Fusion for Video Emotion Recognition

External stimulation, mood swing, and physiological arousal are closely related and induced by each other. The exploration of internal relations between these three aspects is interesting and significant. Currently, video is the most popular multimedia stimuli that can express rich emotional semantics by its visual and auditory features. Apart from the video features, human electroencephalography (EEG) features can provide useful information for video emotion recognition, as they are the direct and instant authentic feedback on human perception with individuality. In this paper, we collected a total of 39 participants' EEG data induced by watching emotional video clips and built a fusion dataset of EEG and video features. Subsequently, the machine-learning algorithms, including Liblinear, REPTree, XGBoost, MultilayerPerceptron, RandomTree, and RBFNetwork were applied to obtain the optimal model for video emotion recognition based on a multi-modal dataset. We discovered that using the data fusion of all-band EEG power spectrum density features and video audio-visual features can achieve the best recognition results. The video emotion classification accuracy achieves 96.79% for valence (Positive/Negative) and 97.79% for arousal (High/Low). The study shows that this method can be a potential method of video emotion indexing for video information retrieval.

**EPRO DIP
- 020**

A Comprehensive Survey of Video Datasets for Background Subtraction

Background subtraction is an effective method of choice when it comes to detection of moving objects in videos and has been recognized as a breakthrough for the wide range of applications of intelligent video analytics (IVA). In recent years, a number of video datasets intended for background subtraction have been created to address the problem of large realistic datasets with accurate ground truth. The use of these datasets enables qualitative as well as quantitative comparisons and allows benchmarking of different algorithms. Finding the appropriate dataset is generally a cumbersome task for an exhaustive evaluation of algorithms. Therefore, we systematically survey standard video datasets and list their applicability for different applications. This paper presents a comprehensive account of public video datasets for background subtraction and attempts to cover the lack of a detailed description of each dataset. The video datasets are presented in chronological order of their appearance. Current trends of deep learning in background subtraction along with top-ranked background subtraction methods are also discussed in this paper. The survey introduced in this paper will assist researchers of the computer vision community in the selection of appropriate video dataset to evaluate their algorithms on the basis of challenging scenarios that exist in both indoor and outdoor environments.

**EPRO DIP
- 021**

Real-Time Detection of Apple Leaf Diseases Using Deep Learning Approach Based on Improved Convolutional Neural Networks

Alternaria leaf spot, Brown spot, Mosaic, Grey spot, and Rust are five common types of apple leaf diseases that severely affect apple yield. However, the existing research lacks an accurate and fast detector of apple diseases for ensuring the healthy development of the apple industry. This paper proposes a deep learning approach that is based on improved convolutional neural networks (CNNs) for the real-time detection of apple leaf diseases. In this paper, the apple leaf disease dataset (ALDD), which is composed of laboratory images and complex images under real field conditions, is first constructed via data augmentation and image annotation technologies. Based on this, a new apple leaf disease detection model that uses deep-CNNs is proposed by introducing the GoogLeNet Inception structure and Rainbow concatenation. Finally, under the hold-out testing dataset, using a dataset of 26,377 images of diseased apple leaves, the proposed INAR-SSD (SSD with Inception module and Rainbow concatenation) model is trained to detect these five common apple leaf diseases.

**EPRO DIP
- 022**

Cryptanalysis and Enhancement of an Image Encryption Scheme Based on Bit-Plane Extraction and Multiple Chaotic Maps

Recently, an image encryption scheme combining bit-plane extraction with multiple chaotic maps (IESBC) was proposed. The scheme extracts binary bit planes from the plain-image and performs bit-level permutation and confusion, which are controlled by a pseudo-random sequence and a random image generated by the Logistic map, respectively. As the rows and columns of the four MSBPs are permuted with the same pseudo-random sequence and the encryption process does not involve the statistical characteristics of the plain-image, the equivalent secret key of IESBC can be disclosed in the scenario of known/chosen-plaintext attacks. This paper analyzes the weak points of IESBC and proposes a known-plaintext attack and a chosen-plaintext attack on it. Furthermore, we proposed an enhanced scheme to fix the shortcomings and resist the proposed plaintext attacks. The experimental simulation results demonstrated that the enhanced scheme is excellent in terms of various cryptographic metrics.

EPRO DIP
- 023

A Detection Method for Apple Fruits Based on Color and Shape Features

The skins of most mature apple fruits are incompletely red and also include green and pale yellow color, which increases the difficulty of fruit detection by machine vision. A detection method based on color and shape features is proposed for this kind of apple fruits. Simple linear iterative clustering (SLIC) is adapted to segment images taken in orchards into super-pixel blocks. The color feature extracted from blocks is used to determine candidate regions, which can filter a large proportion of non-fruit blocks and improve detection precision. Next, the histogram of oriented gradient (HOG) is adopted to describe the shape of fruits, which is applied to detect fruits in candidate regions and locate the position of fruits further. The proposed method was tested by images taken under different illuminations. The average values of recall, precision, and F_1 reach 89.80%, 95.12%, and 92.38% respectively. The performance of detecting fruits covered at different levels is also tested. The values of the recall are all more than 85%, which indicates that proposed method can detect a great part of covered fruits. Compared with pedestrian detection method and faster region-based convolutional neural network (RCNN), the proposed method has the best performance and higher than faster RCNN slightly. However, the proposed method is not robust to noise and its elapsed time of one image is 1.94 s and less than faster RCNN.

EPRO DIP
- 024

Exposure Based Multi-Histogram Equalization Contrast Enhancement for Non-Uniform Illumination Images

Non-uniform illuminated images pose challenges in contrast enhancement due to the existence of different exposure region caused by uneven illumination. Although Histogram Equalization (HE) is a well-known method for contrast improvement, however, the existing HE-based enhancement methods for non-illumination often generated the unnatural images, introduced unwanted artifacts, and washed out effect because they do not utilize the information from the different exposure regions in performing equalization. Therefore, this study proposes a modified HE-based contrast enhancement technique for non-uniform illuminated images namely Exposure Region-Based Multi-Histogram Equalization (ERMHE). The ERMHE uses exposure region-based histogram segmentation thresholds to segment the original histogram into sub-histograms. With the threshold sub-histograms, the ERMHE then uses an entropy-controlled gray level allocation scheme to allocate new output gray level range and to obtain new thresholds that will be used to repartition the histogram prior to HE process. A total of 154 non-uniform illuminated sample images are used to evaluate the application of the proposed ERMHE. By comparing ERMHE to four existing HE-based contrast enhancement namely, Global HE, Mean Preserving Bi-Histogram Equalization (BBHE), Dualistic Sub-Image Histogram Equalization (DSIHE), and Contrast Limited Adaptive Histogram Equalization (CLAHE), qualitatively.

EPRO DIP
- 025

Learning Deep Features for One-Class Classification

We present a novel deep-learning based approach for one-class transfer learning in which labeled data from an unrelated task is used for feature learning in one-class classification. The proposed method operates on top of a Convolutional Neural Network (CNN) of choice and produces descriptive features while maintaining a low intra-class variance in the feature space for the given class. For this purpose two loss functions, compactness loss and descriptiveness loss are proposed along with a parallel CNN architecture. A template matching-based framework is introduced to facilitate the testing process. Extensive experiments on publicly available anomaly detection, novelty detection and mobile active authentication datasets show that the proposed Deep One-Class (DOC) classification method achieves significant improvements over the state-of-the-art.

EPRO DIP
- 026

An Efficient Texture Descriptor for the Detection of License Plates from Vehicle Images in Difficult Conditions

This paper aims to identify the license plates under difficult image conditions, such as low/high contrast, foggy, distorted, and dusty conditions. This paper proposes an efficient descriptor, multi-level extended local binary pattern, for the license plates (LPs) detection system. A pre-processing Gaussian filter with contrast-limited adaptive histogram equalization enhancement method is applied with the proposed descriptor to capture all the representative features. The corresponding bins histogram features for a license plate image at each different level are calculated. The extracted features are used as the input to an extreme learning machine classifier for multiclass vehicle LPs identification. The dataset with English cars LPs is extended using an online photo editor to make changes on the original dataset to improve the accuracy of the LPs detection system. The experimental results show that the proposed method has a high detection accuracy with an extremely high computational efficiency in both training and detection processes compared to the most popular detection methods. The detection rate is 99.10% with a false positive rate of 5% under difficult images. The average training and detection time per vehicle image is 4.25 and 0.735 s, respectively.

EPRO DIP
- 027

Occlusion Aware Facial Expression Recognition Using CNN with Attention Mechanism

Facial expression recognition in the wild is challenging due to various unconstrained conditions. Although existing facial expression classifiers have been almost perfect on analyzing constrained frontal faces, they fail to perform well on partially occluded faces that are common in the wild. In this paper, we propose a convolution neural network (CNN) with attention mechanism (ACNN) that can perceive the occlusion regions of the face and focus on the most discriminative un-occluded regions. ACNN is an end-to-end learning framework. It combines the multiple representations from facial regions of interest (ROIs). Each representation is weighed via a proposed gate unit that computes an adaptive weight from the region itself according to the unobstructedness and importance. Considering different ROIs, we introduce two versions of ACNN: patch-based ACNN (pACNN) and global-local-based ACNN (gACNN). pACNN only pays attention to local facial patches. gACNN integrates local representations at patch-level with global representation at image-level. The proposed ACNNs are evaluated on both real and synthetic occlusions, including a self-collected facial expression dataset with real-world occlusions, the two largest in-the-wild facial expression datasets (RAF-DB and AffectNet) and their modifications with synthesized facial occlusions.

EPRO DIP
- 028

A novel weakly supervised Multitask Architecture for Retinal Lesions Segmentation on Fundus Images

Obtaining the complete segmentation map of retinal lesions is the first step towards an automated diagnosis tool for retinopathy that is interpretable in its decision-making. However, the limited availability of ground truth lesion detection maps at a pixel level restricts the ability of deep segmentation neural networks to generalize over large databases. In this paper, we propose a novel approach for training a convolutional multi-task architecture with supervised learning and reinforcing it with weakly supervised learning. The architecture is simultaneously trained for three tasks: segmentation of red lesions and of bright lesions, those two tasks done concurrently with lesion detection. In addition, we propose and discuss the advantages of a new preprocessing method that guarantees the color consistency between the raw image and its enhanced version. Our complete system produces segmentations of both red and bright lesions. The method is validated at the pixel level and per-image using four databases and a cross-validation strategy. When evaluated on the task of screening for the presence or absence of lesions on the Messidor image set, the proposed method achieves an area under the ROC curve of 0.839, comparable to the state of the art.

EPRO DIP
- 029

Patch-based Output Space Adversarial Learning for Joint Optic Disc and Cup Segmentation

Glaucoma is a leading cause of irreversible blindness. Accurate segmentation of the optic disc (OD) and cup (OC) from fundus images is beneficial to glaucoma screening and diagnosis. Recently, convolutional neural networks demonstrate promising progress in joint OD and OC segmentation. However, affected by the domain shift among different datasets, deep networks are severely hindered in generalizing across different scanners and institutions. In this paper, we present a novel patchbased Output Space Adversarial Learning framework (pOSAL) to jointly and robustly segment the OD and OC from different fundus image datasets. We first devise a lightweight and efficient segmentation network as a backbone. Considering the specific morphology of OD and OC, a novel morphology-aware segmentation loss is proposed to guide the network to generate accurate and smooth segmentation. Our pOSAL framework then exploits unsupervised domain adaptation to address the domain shift challenge by encouraging the segmentation in the target domain to be similar to the source ones. Since the whole-segmentationbased adversarial loss is insufficient to drive the network to capture segmentation details, we further design the pOSAL in a patch-based fashion to enable fine-grained discrimination on local segmentation details.

EPRO DIP
- 030

A Structure-Based Human Facial Age Estimation Framework under a Constrained Condition

Developing an automatic age estimation method towards human faces continues to possess an important role in computer vision and pattern recognition. Many studies regarding facial age estimation mainly focus on two aspects: facial aging feature extraction and classification/regression model learning. To set our work apart from existing age estimation approaches, we consider a different aspect -system structuring, which is, under a constrained condition: given a fixed feature type and a fixed learning method, how to design a framework to improve the age estimation performance based on the constraint? We propose a four-stage fusion framework for facial age estimation. This framework starts from gender recognition, and then go to the second phase, gender-specific age grouping, and followed by the third stage, age estimation within age groups, and finally ends at the fusion stage. In the experiment, three well-known benchmark datasets, MORPH-II, FG-NET, and CLAP2016, are adopted to validate the procedure. The experimental results show that the performance can be significantly improved by using our proposed framework and this framework also outperforms several state-of-the-art age estimation methods.

EPRO DIP
- 031

Attention Residual Learning for Skin Lesion Classification

Automated skin lesion classification in dermoscopy images is an essential way to improve the diagnostic performance and reduce melanoma deaths. Although deep convolutional neural networks (DCNNs) have made dramatic breakthroughs in many image classification tasks, accurate classification of skin lesions remains challenging due to the insufficiency of training data, inter-class similarity, intra-class variation, and lack of the ability to focus on semantically meaningful lesion parts. To address these issues, we propose an attention residual learning convolutional neural network (ARL-CNN) model for skin lesion classification in dermoscopy images, which is composed of multiple ARL blocks, a global average pooling layer, and a classification layer. Each ARL block jointly uses the residual learning and a novel attention learning mechanisms to improve its ability for discriminative representation. Instead of using extra learnable layers, the proposed attention learning mechanism aims to exploit the intrinsic self-attention ability of DCNNs, i.e. using the feature maps learned by a high layer to generate the attention map for a low layer. We evaluated our ARL-CNN model on the ISIC-skin 2017 dataset. Our results indicate that the proposed ARL-CNN model can adaptively focus on the discriminative parts of skin lesions, and thus achieve the state-of-the-art performance in skin lesion classification.

EPRO DIP
- 032

Foreground Fisher Vector: Encoding Class-Relevant Foreground to Improve Image Classification

Image classification is an essential and challenging task in computer vision. Despite its prevalence, the combination of deep convolutional neural network (DCNN) and Fisher vector (FV) encoding method has limited performance, since the class-irrelevant background used in traditional FV encoding may result in less discriminative image features. In this paper, we propose the foreground FV (fgFV) encoding algorithm and its fast approximation for image classification. We try to separate implicitly the class-relevant foreground from the class-irrelevant background during the encoding process via tuning the weights of the partial gradients corresponding to each Gaussian component under the supervision of image labels, and then use only those local descriptors extracted from the class-relevant foreground to estimate FVs. We have evaluated our fgFV against the widely used FV and improved FV (iFV) under the combined DCNN-FV framework and also compared them to several state-of-the-art image classification approaches on 10 benchmark image datasets for the recognition of fine-grained natural species and artificial manufactures, categorization of course objects, and classification of scenes. Our results indicate that the proposed fgFV encoding algorithm can construct more discriminative image presentations from local descriptors than FV and iFV, and the combined DCNN-fgFV algorithm can improve the performance of image classification.

EPRO DIP
- 033

Multiview Semi-Supervised Learning Model for Image Classification

Semi-supervised learning models for multiview data are important in image classification tasks, since heterogeneous features are easy to obtain and semi-supervised schemes are economical and effective. To model the view importance, conventional graph-based multiview learning models learn a linear combination of views while assuming a priori weights distribution. In this paper, we present a novel structural regularized semi-supervised model for multiview data, termed Adaptive MULTiview SEMI-supervised model (AMUSE). Our new model learns weights from a priori graph structure, which is more reasonable than weight regularization. Theoretical analysis reveals the significant difference between AMUSE and the prior arts. An efficient optimization algorithm is provided to solve the new model. Experimental results on six real-world data sets demonstrate the effectiveness of the structural regularized weights learning scheme.

EPRO DIP
- 034

Multi-Classification of Brain Tumor Images Using Deep Neural Network

Brain tumor classification is a crucial task to evaluate the tumors and make a treatment decision according to their classes. There are many imaging techniques used to detect brain tumors. However, MRI is commonly used due to its superior image quality and the fact of relying on no ionizing radiation. Deep learning (DL) is a subfield of machine learning and recently showed a remarkable performance, especially in classification and segmentation problems. In this paper, a DL model based on a convolutional neural network is proposed to classify different brain tumor types using two publicly available datasets. The former one classifies tumors into (meningioma, glioma, and pituitary tumor). The other one differentiates between the three glioma grades (Grade II, Grade III, and Grade IV). The datasets include 233 and 73 patients with a total of 3064 and 516 images on T1-weighted contrast-enhanced images for the first and second datasets, respectively. The proposed network structure achieves a significant performance with the best overall accuracy of 96.13% and 98.7%, respectively, for the two studies. The results indicate the ability of the model for brain tumor multi-classification purposes.



THANK YOU!

Elysium PRO



Titles with Abstracts 2019-20