FACE RECOGNITION USING NEURAL NETWORK-BASED PATTERNMATCHING ALGORITHM

[1]B Anitha, [2]Sakshi Rajanlawar, [3]Shreeya Bandaru

[1]Assistant professor in Department of Information Technology Bhoj Reddy Engineering College for Women

[1]anithabejugama@gmail.com

[2,3] UG Scholars in Department of Information Technology Bhoj Reddy Engineering College for Women

[2]sakshirajanlwar@gmail.com , [3] shreeyabandaru127@gmail.com

ABSTRACT

Face recognization (FR) is characterized as the cycle through which individuals are recognized utilizing facial pictures. This innovation is applied extensively in biometrics, security data, getting to controlled regions, keeping of the law by various implementation bodies, savvy cards, and reconnaissance innovation. The facial acknowledgment framework is constructed utilizing two stages. The initial step is a cycle through which the facial elements are gotten or removed, and the subsequent step is design characterization. Profound learning, explicitly the convolutional neural network (CNN), has as of late gained excellent headway in FR innovation. This paper explores the exhibition of the pre-prepared CNN with multi-class support vector machine (SVM) classifier and the exhibition of move picking up utilizing the AlexNet model to perform order. The review considers CNN design, which has up until this point kept the best result in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) in the previous years, all the more explicitly, AlexNet and ResNet-50. To decide execution enhancement of the CNN calculation, acknowledgment precision was utilized as a determinant. Further developed characterization rates were found in the thorough analyses that were finished on the different datasets of ORL, GTAV face, Georgia Tech face, marked faces in the wild (LFW), frontalized named faces in the wild (F_LFW), YouTube face, and FEI faces. The outcome showed that our model accomplished a higher exactness contrasted with the majority of the cutting edge models. A precision scope of 94% to 100 percent for models with all data sets was gotten. Additionally, this was gotten with an improvement in acknowledgment precision up to 39%.

KEYWORDS: Face recognition, Convolutional neural network (CNN), support vector machine (SVM), ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), Deep learning

INTRODUCTION

The group global positioning frameworks are unacceptable for realtime following applications. They require discovery reactions of future casings in advance and go with huge calculation to create upgraded directions to develop longer directions; an iterative connecting process is performed until expanding the predefined affiliation cost. It infers that personalities of the tracklets can be changed by connecting results at every cycle. Then again, web based global positioning frameworks can be applied for time basic applications since they consecutively fabricate directions in light of frame by frame relationship without the iterative affiliations. In any case, the web-based frameworks are probably going

to deliver divided directions under impediments when identifications of blocked objects are not accessible or erroneous. They experience the ill effects of format float when movements and appearances quickly change. Thus, the presentationof the internet based frameworks is essentially corrupted in complex scenes where articles are much of the time impeded, and their appearances are very comparative. In this paper, we propose an internet global positioning framework, which can heartily follow numerous items even in complex scenes yet additionally be reasonable for web based following applications. Fundamentally, we foster our framework in light of the Bayesian methodology as done in past internet global positioning frameworks to successively appraise the conditions of items (for example position, size, and ID) with online gave

location at each casing.

Following of articles is significant for some PC vision applications. This is a somewhat simple errand when the items are disengaged and effortlessly recognized from the foundation. Nonetheless, in mind boggling and swarmed conditions, many items are available that might have comparable appearances, and impede each other; additionally

impediments by other scene objects are normal. This paper proposes a technique that can vigorously follow various items under such testing conditions. Conventional element based following techniques, for example, those in light of variety, striking focuses or movement masses, don't have a discriminative model that recognizes the item class of interest from others. Utilization of item indicators as discriminative models beats this restriction. It additionally empowers programmed introduction and end of directions.

Nonetheless, however there has been critical improvement in object identification procedures, the exactness of the cutting edge object locators is still distant from great. Missed location, misleading problems and incorrect reactions are normal. The following strategy should work with such disappointments, and furthermore with the troubles because of impediments and appearance likeness among different items. In most existing affiliation based following strategies, a proclivity score between location reactions (or tracklets) is registered once and fixed for all later handling. What this paper propose is a more adaptable methodology where affiliations are made in a few levels and the proclivity measure is refined at each level in light of the information got at the past level. A scene model is likewise assessed from the tracklets and afterward used to produce the last item directions. This technique is applied to follow walkers in various recordings; the exploratory outcomes show that our technique accomplishes an enormous improvement in following precision absent a lot of expansion in computational expense.

Specifically, there is areas of strength for a for portable vision frameworks than can work in unconstrained situations of everyday human living. Building such frameworks has been a far-ultimate objective of scene understanding since the 1970ies, however it is likewise a significant necessity for some applications sooner rather than later of versatile mechanical technology and shrewd vehicles. Up until this point, nonetheless, the sheer intricacy of

numerous certifiable scenes has frequently hindered progress toward this path. In this paper, we center around a significant structure block for portable vision applications, to be specific the capacity to follow numerous individuals in occupied road scenes

as seen from a versatile spectator. This could be a versatile robot, an electric wheelchair, or a vehicle going through a packed downtown area. It has for some time been contended that scene examination in such complex settings requires the blend of and cautious exchange between a few different vision modules. In any case, it is generally hazy the way that such a blend ought to be embraced and which properties are basic for its prosperity. In this paper, we propose a particular plan how to coordinate visual odometry, profundity assessment, object location, and following, and show its relevance by and by.

Multi-target following is significant for some applications like observation and human-PC cooperation frameworks. Its point is to find the objectives, recover their directions, and keep up with their personalities through a video succession; this is an exceptionally difficult issue in jam-packed conditions when the impediments of targets are regular. Specifically, comparative appearance and convoluted cooperations between various targets frequently bring about mistaken following outcomes, for example, track discontinuity and personality switches. Discovery based following techniques have become well known because of ongoing upgrades in object location execution.

These strategies coordinate a few signs like appearance, movement, size, and area into a proclivity model to gauge similitude between discovery reactions or between tracklets in an affiliation streamlining system. While numerous calculations have been proposed for the affiliation system, there has been generally less exertion addressed to foster superior appearance models. Numerous past strategies just figure the distance between two all encompassing variety histograms for consistency estimation.

Multi-target following is a now traditional, yet troublesome undertaking in PC vision. Numerous objectives while powerfully keeping up with information affiliation stays a generally open issue. This is because of a few viewpoints. A principal trouble is the intricacy of the state space one needs to manage: the quantity of conceivable objective directions after some time is exceptionally huge (as a matter of fact endless, in the event that the area space

is persistent), and there is a direction for each of a discrete (yet frequently obscure) number of targets. Without anyone else a colossal state space need not be an issue, but rather a few actual imperatives present conditions both between various areas of similar objective and between various targets.

For example, each item's direct and precise speed should be truly conceivable, and the distance between any two articles can't turn out to be for arbitrary reasons little. Since the discrete directions are not free of one another, augmenting their joint back is overall NP-complete. That's what to intensify, between object impediments cause appearance changes and missing proof. Starting qualities acquired by an EKF (upper left) and an ILPbased tracker (base left) and following outcomes after worldwide nonstop advancement (right). Our strategy produces smooth, tireless directions and altogether diminishes the quantity of bogus up-sides and lost targets. proaches have as of late been proposed which expect to remember them for the model and track down a joint arrangement, rather than following each target exclusively.

This is normally accomplished by confining the state space to a limited arrangement of competitor areas, either by thresholding the perception probability or by consistently discretizing the area space. The discretization, along with specific improvements of the actual requirements, yields energy capabilities for which a (almost) worldwide least can be found. Albeit this property is surely appealing, the cost to pay is an energy capability which just generally approximates the basic back.

The proposed framework, nonetheless, coordinates three principle parts to handle the issues. To accurately allot recognitions with tracks under fractional impediments, the visual following part connects online discoveries with existing tracks by assessing track presence probabilities as well as the probabilities of them, and updates conditions of tracks with related identifications. Notwithstanding, it is as yet hard to follow objects when no location is free for quite a while. For this situation, the track the board part ends follows low reality probabilities and partners the ended tracks with different tracks or discoveries have a place with similar items in order toconnect them. For fruitful relationship in the other two sections, the web-based model advancing part steadily learns discriminative appearance models

with refreshed following outcomes.

As like other internet following methodologies, our visual following part is planned in view of the Bayesian methodology for web based following, however a clever information relationship with a track presence likelihood is integrated to appoint identifications into tracks all the more accurately under fractional impediments. In this way, our track the board part performs track to follow relationship to connect divided tracks under long haul impediments as like tracklets relationship in clump global positioning frameworks. An endeavor to consolidate the two methodologies has given a global positioning framework discriminative part based models. Nonetheless, it is essentially not the same as our own since their framework is planned in light of the clump following structure. Accordingly, tracks are created by universally partner recognitions of the whole arrangements. It demonstrates that their framework isn't appropriate for web based following applications. The two phase global positioning framework proposed is additionally like our own. They produce privately streamlined tracks by partner perceptions with tracks and universally improved tracks by partner divided tracks. They utilize the voracious strategy for nearby affiliation, though we utilize a clever information affiliation. What's more, they utilize the predefined appearance model, however our internet advancing part refreshes discriminative appearance models with web based following outcomes. Thus, our framework can recognize various articles well, despite the fact that the appearances of the items regularly change.

RELATED WORK

In [1] Ninad Mehendale et al presents Facial appearance for feeling identification has generally been a simple undertaking for people, however accomplishing a similar errand with a PC calculation is very difficult. With the new headway in PC vision and AI, identifying feelings from images is conceivable. In this paper, we propose a clever strategy called facial feeling acknowledgment utilizing Convolutional brain organizations (FERC). The FERC depends on two-section Convolutional brain organization (CNN): The initial segment eliminates the foundation from the image, and the subsequent part focuses on the facial element vector extraction. In FERC model, expressional vector (EV) is utilized to track down the five unique kinds of customary look. Administrative information

were gotten from the put away data set of 10,000 pictures (154 people). It was feasible to feature the feeling with 96% precision, utilizing an EV of length 24qualities accurately. The two-level CNN works in series and the last layer of perceptron changes the loads and type values with every emphasis. FERC contrasts from commonly followed techniques with single-level CNN, consequently working on the precision. Moreover, an original foundationevacuation strategy applied, before the age of EV, tries not to manage different issues that might happen (for instance distance from the camera). FERC was widely tried with more than 750K pictures utilizing broadened Cohn-Kanade appearance, Caltech faces, CMU and NIST datasets. We anticipate that the FERC feeling recognition should be valuable in numerous applications like prescient learning of understudies, lie finders, and so on.

In [2] Xiangyun Zhao, Nuno Vasconcelos et al presents Objective capacities for preparing of profound organizations for face-related acknowledgment assignments, like look acknowledgment (FER), for the most part consider each example autonomously. In this work, we present an original pinnacle steered profound organization (PPDN) that utilizes an example with top articulation (simple example) to manage the middle component reactions for an example of non-top articulation (hard example) of a similar kind and from a similar subject. The articulation advancing interaction from non-top articulation to top articulation can subsequently be verifiably implanted in the organization to accomplish the invariance to articulation forces. A unique reason back-proliferation methodology, top angle concealment (PGS), is proposed for network preparing. It drives the moderate layer include reactions of non-top articulation tests towards those of the relating top articulation tests, while keeping away from the converse. This tries not to debase the acknowledgment capacity for tests of pinnacle articulation because of obstruction from their non-top articulation partners. Broad examinations on two famous FER datasets, Oulu-CASIA and CK+, exhibit the prevalence of the PPDN over cutting edge FER techniques, as well as the upsides of both the organization structure and the improvement procedure.In [3] Dumitru Erhan, Will Cukierski et al presents The ICML 2013 Workshop on Challenges in Representation Learning zeroed in on three

difficulties: the black box learning challenge, the look acknowledgment challenge, and the multimodal learning challenge. We depict the datasets made for these difficulties and sum up the consequences of the contests. We give ideas to coordinators of future difficulties and a few remarks on what sort of information can be acquired from AI contests. The coordinators didn't uncover the wellspring of the dataset until after the challenge was finished. To cause the test to underscore semi-administered learning, just 1,000 marked models were saved for preparing. One more 5,000 were utilized for the public chief board. For these models, the marks are not given to the contenders, however the elements are. Each group might transfer expectations for these models two times a day. The subsequent precision is distributed publicy. The public test set is along these lines a kind of approval set, yet additionally gives one's rivals data. One more 5,000 models were utilized for the private test set. The highlights for these models are given to the contenders also, however just the challenge directors see the exactness on them until after the challenge has finished. The private test set is utilized to decide the victor of the challenge.

In [4] Nitish Srivastava, Yoshua Bengio et al gives Deep brain nets countless boundaries are exceptionally strong AI frameworks. Notwithstanding, over fitting is a not kidding issue in such organizations. Enormous organizations are additionally delayed to utilize, making it hard to manage over fitting by consolidating the forecasts of a wide range of huge brain nets at test time. Dropout is a strategy for resolving this issue. The key thought is to haphazardly drop units (alongside their associations) from the brain network during preparing. This keeps units from co-adjusting excessively. During preparing, dropout tests from an outstanding number of various diminished" networks. At test time, it is not difficult to surmised the impact of averaging the forecasts of this multitude of diminished networks by essentially utilizing a solitary unthinned network that has more modest loads. This altogether lessens over fitting and gives significant upgrades over other regularization techniques. We show that dropout works on the exhibition of brain networks on regulated learning errands in vision, discourse acknowledgment, archive arrangement and computational science, gettingcutting edge outcomes on numerous benchmark informational indexes. One potential clarification for the predominance of sexual multiplication is that, over the long haul, the basis

for normal determination may not be individual wellness yet rather blend capacity of qualities. The capacity of a bunch of qualities to have the option to function admirably with one more irregular arrangement of qualities makes them more hearty. Since a quality can't depend on an enormous arrangement of accomplices to be available consistently, it should figure out how to accomplish something helpful all alone or in a joint effort with few different qualities.

In [5] Zheyi Fan, Jiao Jiang, Shuqin Weng-2016 et al presents To work on the precision of stride acknowledgment by completely using step data, a human walk acknowledgment calculation in view of the Discrete Cosine Transform (DCT) and the Linear Discriminant Analysis (LDA) is proposed in this paper. To begin with, recurrence area highlights are removed from the Gait Energy Image (GEI) by DCT, which can actually recognize different recurrence parts of human walk. Then, at that point, these recurrence highlights are additionally planned into the ideal discriminant vectors space by LDA, which can upgrade the segregation and lessen aspects of elements. At last, the character acknowledgment is carried out as per the most brief comparability distance. Probes the data set given by Chinese Academy of Sciences (Institute of Automation) exhibit that the proposed highlight extraction technique has the most elevated acknowledgment rate among the looked at strategies. recurrence abundancy range of GEI, which mirrors the recurrence appropriation of its stride picture. The low-recurrence parts situating in the upper left corner of the recurrence adequacy range are comparing to the districts with bigger dark qualities in GEI, to be specific the gradually changing pieces of human body during their strolling. In actuality, the parts in the lower right corner of the recurrence adequacy range are the high-recurrence parts, which address the quick changing pieces of GEI. They mirror the districts with more modest dim qualities in GEI and contain plentiful subtleties and edge data.

PROBLEM DEFINITION

Existing facial features of FER can be divided into two categories: handcrafted feature and learned feature. The design of useful handcrafted feature relies on domain knowledge, which requires a heavyworkload to be carried out manually. In the spatial domain, facial features can be calculated based on geometry and image gradient. In the frequency domain, the high-frequency components

correspond to noises and edges. On the contrary, the low-frequency components are comprehensive measurement of image intensity.

PROPOSED SYSTEM

We propose a learnable duplication piece (LMK) which fills in as a proficient channel to learn facial elements in recurrence space. In view of the LMK, we develop various increase layers for include learning. A synopsis layer is proposed following increase layers to additional yield significant level highlights. Third, in view of the property of discrete cosine change (DCT), we perform basic data extraction (CIE) for direct aspect decrease, and use the proposed layers to build the Basic-FreNet, which can yield undeniable level element on the generally utilized DCT highlight.
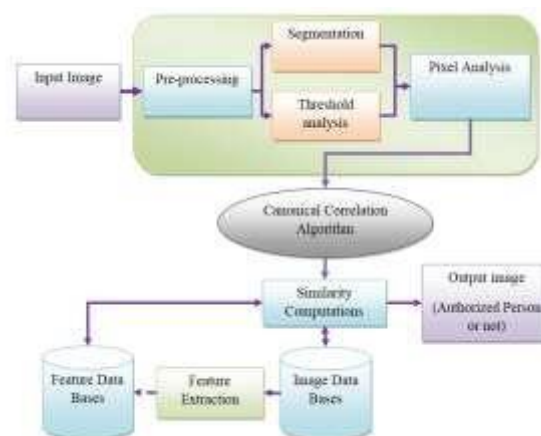
ARCHITECTURE DIAGRAM

Fig 1 Architecture diagram

PROPOSED PROCESS ANALYSIS

- Input Face Image Analysis
- Authentication module
- Training module
- Fingerprint Image
- Image preprocessing
- Fingerprint image enhancement
- Minutiae feature extraction
- Matching methods
- Extract match image from the database.

INPUT FACE IMAGE ANALYSIS

To picture investigation capability first interaction is catch the picture. And afterward contrast the picture with the information base capability and confirm the understudy participation capability

Picture examination is the extraction of significant

data from pictures; fundamentally from advanced pictures through computerized picture handling methods. Picture investigation errands can be pretty much as basic as perusing bar coded labels or as modern as distinguishing an individual from their face.

PCs are irreplaceable for the investigation of a lot of information, for undertakings that require complex calculation, or for the extraction of quantitative data. Then again, the human visual cortex is a phenomenal picture investigation device, particularly for removing more significant level data, and for some applications

— counting medication, security, and remote detecting

— human examiners actually can't be supplanted by PCs. Thus, numerous significant picture examination devices, for example, edge identifiers and brain networks are roused by human visual insight models.

A computerized video recorder DVR, now and again alluded to by the promoting term individual video recorder PVR, is purchaser hardware gadget or application programming. To records video in an advanced configuration to a circle drive, USB streak drive, SD memory card or other neighborhood or arranged mass capacity gadget. The term incorporates set-top boxes STB with direct to circle recording office, versatile media players PMP with recording, recorders PMR as camcorders that record onto Secure Digital memory cards and programming for PCs which empowers video catch and playback to and from a hard plate. This module predominantly manages recording the recordings. Utilizing these recordings the unapproved individual is followed. The recordings are caught utilizing the webcam.

Fig 2 Frame recording process

Online Learning Algorithm

There are times in a creation when the recording just exists in one configuration and the client needs to utilize it in another. A typical illustration of that is a 24p venture being given a laying out shot that was shot on digit beta. For this situation, the recording is intertwine, 60 fields at 30 casings each second while the venture type is moderate and 24

edges each second. This module manages changing the video over completely to outlines. The video is switched over completely to outlines in view of time seconds.

Fig 3 Frame conversion process

ONLINE LEARNING ALGORITHM

Internet learning, the memory expected to store the capacity stays consistent even with added pieces of information, since the arrangement processed at one stage is refreshed when another information point opens up, after which that information point can then be disposed of. For some definitions, for instance nonlinear piece strategies, genuine web based learning is absurd, however a type of half breed internet learning with recursive calculations can be utilized. For this situation, the space necessities are not generally destined to be consistent since it requires putting away all past data of interest, however the arrangement might find opportunity to process with the expansion of another data of interest, when contrasted with bunch learning procedures.

USER PROCESS

In this module the client can enlist. During Registration the client can transfer the picture. Subsequent to transferring the pictures can store on the data set. From that the face acknowledgment interaction should be possible
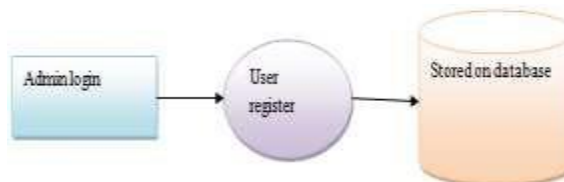

Fig 4 User process

FACE RECOGNITION

Object Tracking Algorithm-Frames put away on one drive from that it contrast and the information base and confirmation cycle to be finished. The location and following of appearances and facial highlights in video groupings is a key and testing issue in PCvision. This exploration region has numerous applications in face distinguishing proof frameworks, model-based coding, look location, human-PCassociation, remotely

coordinating, and so forth. This module manages perceiving the substance of the people in the recorded video and shows whether they are approved or unapproved people

Fig 5 face recognition

OBJECT TRACKING ALGORITHM

The issue of following an article in a video given its area in the principal outline and no other data. As of late, a class of following procedures called "following by location" has been displayed to give promising outcomes at continuous paces. These strategies train a discriminative classifier in a web-based way to isolate the item from the foundation. This classifier bootstraps itself by utilizing the ongoing tracker state to separate positive and negative models from the ongoing edge

Fingerprint image

In this part we are gathering the example information's of both living and dead fingers and put away in a data set. The pictures of both live and dead people are in the single data sets on the whole.

Fingerprint Representation:

The representations of fingerprints are four types:

- Image-based representation.
- Global ridge pattern.
- Local ridge detail.
- Intra-ridge detail.

Fingerprint image enhancement:

Unique finger impression picture quality is a significant figure the exhibition of particularsextraction and matching calculations. A decent quality unique finger impression picture Figure 1 has high difference among edges and valleys. A low quality finger impression picture Figure 2 and 3 is low interestingly, uproarious, broken, or smudgy, causing deceptive and missing particulars. Low quality can be because of cuts, wrinkles, or injuries on the outer layer of fingertip, unnecessarily wet or dry skin condition,

uncooperative disposition of subjects, harmed and messy scanner gadgets, bad quality fingers old individuals, manual specialist and different elements.

Fig 5 finger print verification

CONCLUSION

In this task, pre-prepared convolution neural network (CNN) structures were applied for face biometric framework with various methodologies. To start with, we applied the pre-prepared CNN AlexNet and ResNet-50 for extricating highlights and the help vector machine SVM for grouping. Second, we applied move gaining from the AlexNet model for extricating highlights and arrangement. In the review, we directed three analyses. To start with, we assessed the exhibition for pre-prepared convolutional brain network AlexNet for separating learned elements and utilizing a multi-class support vector machine (SVM) for the order task. Second, we assessed the exhibition for a pre-prepared CNN ResNet-50 for separating learned elements and involving a SVM as a classifier. Third, we assessed the exhibition for move gaining from pre-prepared CNN AlexNet for the order task. The examination study was directed on different datasets (Georgia Tech face dataset, FEI faces, GTAV face, YouTube face, LFW, F_LFW, ORL, and DB_Collection). The outcomes showed the precision scope of 94% to 100 percent for models with all data sets acquired. The outcomes for AlexNet with SVM affirmed that an ideal component can be extricated from 'fc7'. For the testing time, ResNet-50 with SVM took less time than different organizations with all datasets. We contrasted our model and the best in class models as far as the datasets (FEI faces, LFW, YouTube face, and ORL). The outcomes showed that our model accomplished a higher exactness than the greater part of the best in class models. Later on, we mean to additionally further develop acknowledgment and arrangement exactness. To do as such, more information bases should beincorporated for preparing our CNN models, as well as to test different

convolutional brain network models for better working

REFERENCE

[1] P. Ekman and W. V. Friesen, ―Constants across cultures in the face and emotion,‖ J. Personality Social Psychol., vol. 17, no. 2, pp. 124–129, 1971.

[2] A. Kläser, M. Marszalek, and C. Schmid, ―A spatio-temporal descriptor based on 3D-gradients,‖ in Proc. Brit. Mach. Vis. Conf., 2008, pp. 99.1–99.10.

[3] P. Scovanner, S. Ali, and M. Shah, ―A 3-dimensional sift descriptor and its application to action recognition,‖ in Proc. 15th Int. Conf. Multimedia, 2007, pp. 357–360.

[4] L. Ma, ―Facial expression recognition using 2-D DCT of binarized edge images and constructive feedforward neural networks,‖ in Proc. IEEE Int. Joint Conf. Neural Netw. (IEEE World Congr. Comput. Intell.), Jun. 2008, pp. 4083–4088.

[5] S. Mohseni, H. M. Kordy, and R. Ahmadi, ―Facial expression recognition using DCT features and neural network based decision tree,‖ in Proc. ELMAR, 2013, pp. 361–364.

[6] Y. Tang, X. M. Zhang, and H. Wang, ―Geometric-convolutional feature fusion based on learning propagation for facial expression recognition,‖ IEEE Access, vol. 6, pp. 42532–42540, 2018.

[7] H. Jung, S. Lee, J. Yim, S. Park, and J. Kim, ―Joint fine-tuning in deep neural networks for facial expression recognition,‖ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 2983–2991.

[8] X. Liu, B. V. K. V. Kumar, P. Jia, and J. You, ―Hard negative generation for identity-disentangled facial expression recognition,‖ Pattern Recognit., vol. 88, pp. 1–12, Apr. 2019.

[9] G. Zhao and M. Pietikainen, ―Dynamic texture recognition using local binary patterns with an application to facial expressions,‖ IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 6, pp. 915–928, Jun. 2007.

[10] Y. Xiao, L. Ma, and K. Khorasani, ―A new facial expression recognition technique using 2-D DCT and neural networks based decision tree,‖ in Proc. IEEE Int. Joint Conf. Neural Netw., Jul. 2006, pp. 2421–2428.

[11] M. H. Siddiqi, R. Ali, A. M. Khan, Y.-T. Park, and S. Lee, ―Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields,‖ IEEE Trans. Image Process., vol. 24, no. 4, pp. 1386–1398, Apr. 2015.

[12] A. Majumder, L. Behera, and V. K. Subramanian, ―Automatic facial expression recognition system using deep network-based data fusion,‖ IEEE Trans. Cybern., vol. 48, no. 1, pp. 103–114, Jan. 2018.

[13] J. Chen, Z. Chen, Z. Chi, and H. Fu, ―Facial expression recognition in video with multiple feature fusion,‖ IEEE Trans. Affect. Comput., vol. 9, no. 1, pp. 38–50, Jan. 2018.

[14] H. Jung, S. Lee, J. Yim, S. Park, and J. Kim, ―Joint fine-tuning in deep neural networks for facial expression recognition,‖ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 2983–2991.

[15] A. Kacem, M. Daoudi, B. B. Amor, and J. C. Alvarez-Paiva, ―A novel space-time representation on the positive semidefinite cone for facial expression recognition,‖ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 3180–3189.