1st.International Conference of Information Technology to enhance E-learning and other Application,( IT-ELA 2020 ), Baghdad
College of Economic Sciences University, Baghdad, Iraq

# Content-based Recommender System Improvement using Hybrid Technique

Khamael Raqim Raheem
*Department of Software. College of Information Technology*
*University of Babylon*
Babil, Iraq
khmrak@itnet.uobabylon.edu.iq

Israa Hadi Ali
*Department of Software. College of Information Technology*
*University of Babylon*
Babil, Iraq
israa_hadi@itnet.uobabylon.edu.iq

***Abstract*— The research in recommender systems has evolved considerably over the past years; however, to date the investigation on how the emotive state of the user could be used to complement such technologies is sparse. Many systems used the emotions of the user as implicit feedbacks in the recommender systems, but limited works analyzed the emotions and using them for item profile modelling. This system extracts the affective features of the user's face using the convolution neural network with three-dimension to build the item profile. Also, the support vector machine classifier is used for user profile building and introduces recommendations. The experiments conducted on an LDOS-PerAff-1 dataset indicates that the hybrid technique contributes to improving the content-based recommender system performance using affective-automated features for item modelling instead of handcrafted features that extract using traditional techniques.**

***Keywords—content-based recommender system, affective-automated features, handcrafted features, 3D convolution neural network, support vector machine, emotions***

## I. INTRODUCTION

Recommender systems (RSs) are software tools or techniques that support the user in the decision-making process by suggesting possibilities that the system predicts [1]. RSs are associated with numerous applications such as Amazon.com, for a book recommendation, compact disks, and other items [2], MovieLens, for the movie [3], and VERIFIED technologies, for news articles recommendation [4]. There are several algorithms applied in the RS field like the content-based filtering (CBF) [5], collaborative filtering (CF) [6], and context-based filtering [7].

In 2005, the research community accepted the importance of emotions in RSs. Emotions are important to RSs because the consumption of an item is made with the expectation of encountering emotions. Emotions are fundamental to human experience and the impact on daily activities and decision making. For instance, the emotional state influences the user's decision to consume the recommended item. The effective recommender systems are applied for images [8], news[9], and movies [10].

There are limited researches used the emotions for item profile modelling in content-based recommender (CBR) systems. Group of researchers in [11] used affective metadata (metadata that portrays the user's emotions) for item profile modelling and explored the impact of this on the efficiency of RS. Also, most researches in affective RS used conventional methods to extract the emotive features of the user like the active appearance model (AAM) features, the facial fiducial points [12, 13], and Gabor features [14]. The features extracted using these methods are called handcrafted features that it is highly dependent on manual feature engineering. Recently, in terms of facial expression recognition (FER), deep learning-based approaches extract more robust features (automated features). These approaches are more robust to the environments with different elements, e.g., illumination and occlusion, which means that they can greatly outperform the conventional approaches. For example, the convolution neural network (CNN) deep learning technique used in [15] to classify facial expression in to positive or negative. Additionally, the study in [16] designed a three-dimension convolution neural network (3D CNN) for facial expression prediction from videos.

The rest portion of the paper is organized as takes after. Section II presents an overview of the related work in the recommender system field. Section III clarifies the content-based recommender algorithm. Section IV explains the 3D convolution neural network. In section V the proposed system design is presented. The results and discussion display in Section VI and finally the conclusion is presented in Section VII.

## II. RELATED WORK

There are many types of research done in the RS field that track the facial features of the user to produce recommendations. The researchers in [17] designed a search interface that applied real-time facial expression analysis using the eMotion system to aggregate information on the users' affective behaviour. They presented a multimodal recommender system to exploit information to classify the topical relevance of the perused videos, with the help of a support vector machine (SVM), and eventually enrich the profiles of RS. Comparative work was done by [18]. They created a system that extracts effective video clip labels from the facial expressions of users.

The researchers in [11] demonstrated that the emotions detected through facial expressions improved an RS performance. They investigated the influence of affective metadata (AM) of the user on the efficacy of content-based recommendations for a subset of IAPS database images [19]. They tested the output of the RS using four various classifiers: SVM, Naïve Bayes (NB), and Adaptive Boosting (AdaBoost). They measured the system performance using the metrics of precision, recall, and f-measure.

The researchers in [20] presented an RS for movies, that used facial recognition with a hybrid approach, joining collaborative filtering with content-based filtering. Their goal is to create a system that does not disturb the user to provide information. However, even though they use facial recognition, they use it for different purposes. The first purpose is to detect the gender and the age of the user; the second one, to assign a genre to the movies from the emotions shown by the user while watching the trailer. They have an actual implementation of the website that they use in life for evaluating the system. The study in [21] analyzed a user's facial expressions and physiological parameters when viewing a video. They determined the emotions of a user (happy, sad, angry, surprised, scared, disgusted, and neutral) and selected rational real-time video clips. They claimed that the analysis of the facial expressions and physiological parameters of a video viewer could indicate possible offers to users for video clips that they currently prefer.

In the proposed system, an off-line hybrid recommender system will be designed for background image recommendation using the 3D CNN. The previous researches which used traditional methods had a limitation with the operation of extracting the affective features (handcrafted features) which used to build the RS. While in the proposed RS, a 3D CNN deep learning technique will use to extract more robust effective features (automated features) from the user's face while viewing the item. The contributions of the proposed RS system are:

1. We are designing new 3D CNN architecture to extract automated features from the video clip of the user's face while viewing the item.
2. We are building an item profile based on the extracted automated features.

The SVM classifier will train on each user's preferences (items likes or dislikes) to build the user profile and introduce the recommendations.

## III. CONTENT-BASED FILTERING ALGORITHM

Content-based filtering algorithm recommends items that are close to items liked by the user already. The root of the CBF is in information retrieval and information filtering [22, 23]. The user and item profiles are constructed, and a comparison made between them to detect the similarity. The system compares any unused item to those inside the user profile to produce the recommendations [24]. To build an item profile, the items must be analyzed. The items can represent by applying feature extraction to represent them as vectors. The user preferences are captured in order to build a user profile. The advantages of this type of filtering are there is no data sparsity, and there is no cold start problem[32-34].

## IV. THREE-DIMENSIONAL CONVOLUTION NEURAL NETWORK

The CNN is a form of deep neural learning and feed-forward networks that can be used for a variety of machine learning functions, for example, facial expressions recognition [25, 26]. In order to recognize emotions from videos, the changing of facial features in consecutive frames must be taken into consideration. Extracting the emotions from the video is more complicated. Most conventional methods consider structures autonomously while disregarding the temporal relationships of the sequential frames in a series that is central to the identification of unpretentious changes in facial frame appearance. Recognition of emotions in images can be visualized in a given time by displaying several frames.

Several neural network architecture deals with frame sequence classification; one of these architectures are 3D CNN. The mixture of spatial and temporal data is utilized in 3D convolution to learn about a transition in consecutive frames. The 3D convolution kernel is a 3D cube that takes into consideration local spatial area adjacent frames. If the kernel size is 2x2x2x1, the dot product considers the receptive field is 2x2 in two consecutive frames, and the fourth dimension is the colour channel (which in this case is 1 because of the use of the grayscale image). Fig. 1 portrays the 3D kernel convolution.

## V. PROPOSED SYSTEM DESIGN

In this section, important details about the dataset will explain, and the main stages for the proposed RS will illustrate. These stages are pre-processing, item profile modelling, and user profile modelling and recommendation. The Process diagram in fig. 2 illustrates the proposed RS design. In the pre-processing stage, the video clips of users are processed to be contained the same number of frames. Each video clip will convert in to sample; the sample contains eight frames to identify the emotion of the user. Also, the frames in each sample will be resized and convert to a grayscale format, and the face and eye will detect and extract from each frame.

In the item profile modelling stage, the facial features will extract from each sample. These features are obtaining from the output of the third full connection (FC) layer in the 3D CNN. The extracted features will use to calculate the affective-automated features (AAF) to build the item profile. The AAF is the mean and standard deviation values of the users that consumed the same item.

In the final stage, the SVM classifier will use for user profile building and introduce recommendations. Each user in RS has its preferences, and these preferences are represented by items that are like or dislike by the user. The SVM will be trained on the AAF vectors and corresponding ground truth ratings as classes for each user to build the user profile. Also, the recommendation will introduce based on the classification results of the SVM classifier.

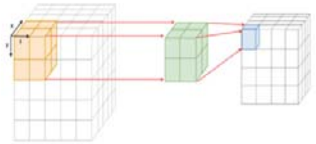


Fig. 1. 3D convolution process

## A. Dataset

The proposed system will build a CBR system using the LDOS-PerAff-1 dataset [27]. The dataset contains 52 videos, one for each user. The frame rate for each video file is 15fps. Each video split into 70 video clips.

The video clip contains a sequence of frames that reflect the user emotive response while viewing the IAPS image to select it as a background. The overall videos are 3640 video clips to the 52 users reacting to 70 diverse images from the IAPS database. Each video clip annotates with affective metadata and ground-truth rating. The effective metadata for each item profile consists of the statistical moments (mean and standard deviation) for the dimensional emotion (valence, arousal and dominance). The mean and standard deviation have represented the values of the dimensional emotion of the users that view the same item. Table I shows an example of an item profile containing the effective metadata.

## B. Pre-processing

The extraction of facial features using 3D CNN from each video clip needs to consolidate the number of frames inside each video clip. According to the results reported in [28], succinct frame fragments are used instead of a whole video. This study showed the concise 1-7 frame snippets are adequate to identify human action as expected by biological vision systems observation. In the proposed system, each video clip will convert into a sample with 8 frames to identify the affective response of the user. The frames at each sample resized to 32x32 pixels where each frame converted to a grayscale image.

The study in [29] explained that 'the positive affective responses result in positive evaluations of the focal item, whereas negative affective responses result in negative evaluations. Based on this study the sample (sequence of frames that reflect the affective response of the user) will have a positive class if the user likes the viewing item or will have a negative class if the user dislikes the viewing item. These classes will take from ground truth ratings in the dataset. Fig. 3 illustrates two samples to the user 101 while watching the images 1931 and 1300. Fig. 3 A displays the affective response of the user when giving "1" as a rating to the consumed item.

TABLE I. ITEM PROFILE WITH AFFECTIVE METADATA

| **Metadata Field (handcrafted features)** |
|---|
| Image id |
| Valence mean |
| Valence standard deviation |
| Arousal mean |
| Arousal standard deviation |
| Dominance mean |
| Dominance standard deviation |

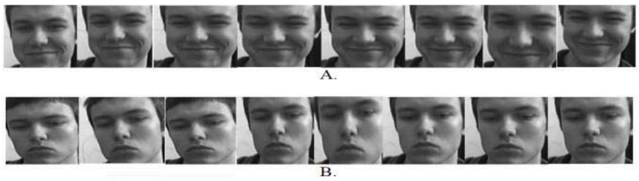


Fig. 3. A. User 101 while viewing item 1931 and give a 1 rating, B. User 101 while viewing item 1300 and give a 0 rating.

Figure 3 B displays the negative response when the user gives "0" as a rating to the consumed item. The pre-processing implemented over each frame in the sample to determine the face and eye using a digital library (DLib) face detector and Haar cascade classifier. Pre-processing outcomes are 70 samples per-user; the samples are annotated with either a positive or a negative class. The data augmentation (DA) process is applied to increase the size of the training samples for the 3D CNN. DA is the process of taking samples that are already in a training dataset and processing them to create many altered versions of the same samples. Here the vertical and horizontal flipping, blurring, and sharpening to original sequences of frames are made to generate new samples added to the training samples.

## C. Item Profile Modeling

The 3D CNN model will build based on training samples. The 3D CNN will train on samples annotated with positive and negative classes. The sequence of frames to the positive samples reflects the positive emotion the user has during item viewing. In contrast, the negative samples contain a sequence of frames that reflect the negative emotion the user has during item viewing. Fig. 4 depicted the layers of the proposed 3D CNN such that several layers with different purposes that will be used for network design like 3D convolution (3D Conv), batch normalization (Batch Norm), 3D max-pooling (3D Max Pool), fully connection, dropout, flatten, and output layer. The 3D convolution layers are used to extract the spatiotemporal features. In order to present nonlinearity in convolution layers, an activation function will use where the rectified linear unit (ReLU) activation function is used commonly.

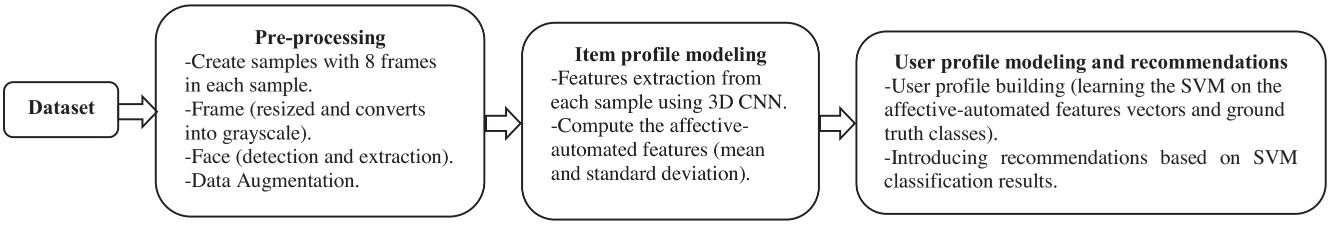


Fig.2. Process diagram of the proposed recommender system

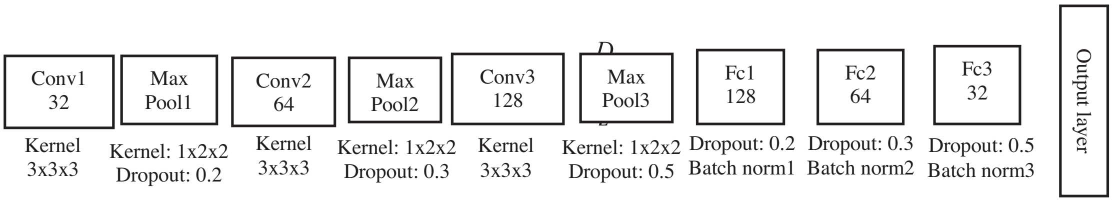


Fig. 4. Proposed 3D convolution neural network layers

ReLU is performed by basically thresholding matrix values at zero, compared to the large computation needed by sigmoid and tanh activation functions. The 3D Max Pool layer continually decreases the 3D convolution output while it retains the important features. In a small spatiotemporal window, the 3D Pool layer selects the finest representation of elements. Batch Norm is a technique that helps boost speed, performance, and reliability of CNN. It is used to adjust and scale the activation in order to normalize the input data.

The dropout in the network decreases system overfitting of overtraining samples. Too it is used to incorporate the capability for regularisation. The flatten layer is nothing but the extension to a one-dimensional array required for the fully connected layer from multidimensional inputs. In the sense of hierarchical feature extraction, the FC layers display more nonlinearity in the network. In order to build the class scores, the output layer is used where the sigmoid function used on binary classes in this layer. The researchers in [30] stated that The output values of the hidden layers could be treated as features for any other classifiers. Therefore the outputs of the third FC layer will use for item profile building in the proposed RS.

The study in [11] claimed that the selection of the first two statistical moments in the dimensional emotion is a representation of the users' common emotional response to an item. The concept is similar to the 'known collaborative tagging from the Web2.0 environment that, intuitively, would contribute more to the overall variance'. By taking advantage of the study above the mean and standard deviation will be computed in the proposed RS based on the outputs (facial features) of the third FC layer in order to build the item profile. The item profile will contain the mean and standard deviation of the effective facial features for the users that have a common emotional response to an item.

### *D. User Profile Modeling and Recommendations*

The SVM classifier will train on each user's preferences. The user's preferences are all the items that are like or dislike by the user. The items profiles are built in the previous subsection, where each item profile will contain AAF for a specific item. The user profile for each user will construct by

training of the SVM classifier on the AAF vectors (user's preferences). Each input vector of AAF has a corresponding positive or negative class. These classes represent the ground truth ratings. SVM classifier will learn the relations between the input AAF and the corresponding classes and saves these relations in the user profile that represents the learned knowledge. The results of the SVM classification will use to give recommendations for the proposed RS.

## VI. EXPERIMENTAL RESULTS AND DISCUSSION

There are two types of experiments conducted in the proposed system. In the first experiment, CBR systems are designed based on the effective metadata (handcrafted features) obtained from the dataset using SVM, NB, AdaBoost classifiers. In the second experiment, the CBR system is designed based on AAF obtained from the 3D CNN model using the SVM classifier.

3D CNN is a machine learning algorithm that has a training phase and testing phase. The input samples are split into an 80% training set and a 20% test set. In the training phase, 3D CNN is trained on the training samples to design a model that can extract the facial features for each sample for the testing set. The training of the network is stopped when a minimum error of the loss function obtained. Fig. 5 illustrates the binary-cross entropy loss curve for the training and testing data. The network is applied in Keras library using an adaptive learning rate optimization algorithm (Adam).

Moreover, the network is trained for 43 epochs with a batch size of 32 sample per-epoch. In the testing phase, the output of the third FC layer (32 outputs) will use for item profile building. The mean and standard deviation are computed for the users that viewing the same item based on the FC outputs. Table II illustrates the affective-automated features of the building item profile. The F1-mean feature is calculated using the first output of the FC layer to the users that viewing a specific item; also, the F1-standard deviation is computed similarly. The rest outputs of the FC layer are used to compute the other means and standard deviations.

The AAF is used as inputs to the SVM classifier. SVM with radial basis function (RBF) kernel is trained on input AAF and corresponding classes obtained from the ground truth ratings to build the user profile for each user. The 10-fold cross-validation test scheme is used to assess the performance of the CBF system [31]. In the two types of experiments the

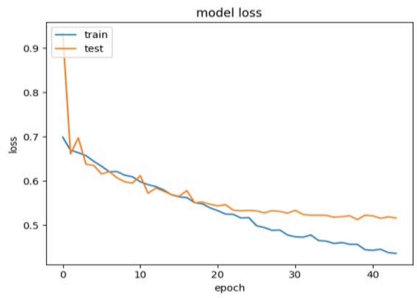


Fig. 5. The curve of the loss function

TABLE II. THE IMPROVEMENT ITEM PROFILE

| **Affective-Automated Features** |
|---|
| Image id |
| F1-mean |
| F1-standard deviation |
| .<br>. |
| F32-mean |
| F32- standard deviation |

Performance measures will compute for each user separately using cross-validation with 10-fold. The output of the classification process to each classifier is binary where the items identified as relevant (class 1) or nonrelevant (class 0) for the individual being observed. The outcome of the classification process is a confusion matrix. The item classified correctly (a true positive (TP) or true negative (TN)) or classified incorrectly (false positive (FP) or false negative (FN)).

The confusion matrix contains the number of correct and incorrect classifications. Table III depicts the confusion matrix for binary classification. The precision, recall, and f-measure are used to evaluate the RS performance.

Precision (P) is the portion of recaptured items that are relevant to the search. The recall (R) is the portion of the relevant items that are recaptured out of all of the relevant items. The f-measure (F) is a trade-off between precision and recall. The calculation for precision, recall, and f-measure is shown in “(1)”, “(2)” and “(3)” respectively.

$$P = TP / (TP+FP) \quad (1)$$

$$R=TP/ (TP+FN) \quad (2)$$

$$F=2x(PxR) / (P+R) \quad (3)$$

Finally, the averages of P, R, and F to all users are calculated to produce the performance of the proposed RS. The results obtained from the first experiment using handcrafted features compared with the second experiment using AAF show that the performance of the CBR system can be improved using AAF for item profile building. Table IV illustrates the results of the two experiments.

The two experiments applied to a portion of users in the dataset, the reason that some videos of the users contain many challenges such as the occlusion and head position changing. These challenges affect the efficiency of 3D CNN; therefore, the best videos are used in the proposed 3D CNN. In the literature, the CBR system is designed based on handcrafted features using all users in the dataset, the results for this system illustrated in table V.

TABLE III. CONFUSION MATRIX FOR BINARY CLASSIFICATION

| **Predicted Class** | **Actual Class** | |
|---|---|---|
| | ***Class 1*** | ***Class 0*** |
| ***Class 1*** | TP | FP |
| ***Class 0*** | FN | TN |

The AAF represents robust features that extracted to reflect the emotional state of the user while viewing the item, where 3D CNN is used for this goal. The utilization of spatiotemporal features joint training using 3D CNN promoted the more exact.

## VII. CONCLUSION

This paper presented a hybrid 3D CNN-SVM system for recommendations using 3D CNN with fully connected layers for spatiotemporal features extraction and the SVM classifier for user profile building and introduces recommendations. The 3D CNN used to extract robust automated features that are used for the item profile modelling. SVM classifier is used for the user profile building and introduces recommendations. The comparison CBR system using affective-automated features with the CBR systems using handcrafted features over the RS dataset shows that the affective-automated features can improve the efficiency of the CBR system. In the future, this system can extend by building an RS based on a combination of handcrafted and automated features where the handcrafted features are extracting using traditional techniques, and automated features are extracting using deep learning techniques.

TABLE IV. COMPARISON OF THE PROPOSED SYSTEM EXPERIMENTS USING HANDCRAFTED FEATURES AND AUTOMATED FEATURES

| Features | classifier | P% | R% | F% |
|---|---|---|---|---|
| affective metadata (handcrafted features) | SVM | 0.62 | 0.65 | 0.64 |
| | NB | 0.53 | 0.57 | 0.55 |
| | AdaBoost | 0.62 | 0.60 | 0.61 |
| affective-automated features | 3D CNN- SVM | 0.69 | 0.71 | 0.70 |

TABLE V. RESULTS OF THE CONTENT-BASED RECOMMENDER SYSTEMS USING HANDCRAFTED FEATURES IN [11]

| Classifier | P% | R% | F% |
|---|---|---|---|
| SVM | 0.61 | 0.55 | 0.58 |
| NB | 0.58 | 0.58 | 0.58 |
| AdaBoost | 0.57 | 0.42 | 0.48 |

## REFERENCES


[1] F. Ricci, L. Rokach, and B. Shapira, "Recommender systems: introduction and challenges," in *Recommender systems handbook*, ed: Springer, 2015, pp. 1-34.

[2] G. Linden, B. Smith, and J. York, "Amazon. Com recommendations: Item-to-item collaborative filtering," *IEEE Internet Computing,* vol. 7, pp. 76-80, 2003.

[3] B. N. Miller, I. Albert, S. K. Lam, J. A. Konstan, and J. Riedl, "MovieLens unplugged: experiences with an occasionally connected recommender system," in *Proceedings of the 8th international conference on Intelligent user interfaces*, 2003, pp. 263-266.

[4] D. Billsus, C. A. Brunk, C. Evans, B. Gladish, and M. Pazzani, "Adaptive interfaces for ubiquitous web access," *Communications of the ACM,* vol. 45, pp. 34-38, 2002.

[5] M. Soares and P. Viana, "Tuning metadata for better movie content-based recommendation systems," *Multimedia Tools and Applications,* vol. 74, pp. 7015-7036, 2015.

[6] H. Naji and G. Al-Sultany, "Collaborative filtering recommendation using personalized page rank algorithm with new personalized parameters," *J. Eng. Appl. Sci.,* vol. 12, pp. 4108-4112, 2017.

[7] E. F. ABDULLAH, G. A. Al-SULTANY, and H. N. NAWAF, "ROUGH SET BASED CONTEXT SUGGESTIONS," *Journal of Theoretical and Applied Information Technology,* vol. 96, 2018.

[8] M. Tkalcic, A. Odic, A. Kosir, and J. Tasic, "Affective labelling in a content-based recommender system for images," *IEEE transactions on multimedia,* vol. 15, pp. 391-400, 2012.

[9] R. S. Bader, "EANRS: An Emotional Arabic News Recommender System," in *2019 4th Scientific International Conference Najaf (SICN)*, 2019, pp. 139-144.

[10] C. Orellana-Rodriguez, E. Diaz-Aviles, and W. Nejdl, "Mining affective context in short films for an emotion-aware recommendation," in *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, 2015, pp. 185-194.

[11] M. Tkalčič, U. Burnik, and A. Košir, "Using effective parameters in a content-based recommender system for images," *User Modeling and User-Adapted Interaction,* vol. 20, pp. 279-311, 2010.

[12] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE transactions on pattern analysis and machine intelligence,* vol. 31, pp. 39-58, 2008.

[13] F. De la Torre and J. F. Cohn, "Facial expression analysis," in *Visual analysis of humans*, ed: Springer, 2011, pp. 377-409.

[14] R. Zhi and Q. Ruan, "Facial expression recognition based on two-dimensional discriminant locality preserving projections," *Neurocomputing,* vol. 71, pp. 1730-1734, 2008.

[15] P. Barros, C. Weber, and S. Wermter, "Emotional expression recognition with a cross-channel convolutional neural network for human-robot interaction," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, 2015, pp. 582-587.

[16] B. Hasani and M. H. Mahoor, "Facial expression recognition using enhanced deep 3D convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 30-40.

[17] I. Arapakis, Y. Moshfeghi, H. Joho, R. Ren, D. Hannah, and J. M. Jose, "Integrating facial expressions into user profiling for the improvement of a multimodal recommender system," in *2009 IEEE International Conference on Multimedia and Expo*, 2009, pp. 1440-1443.

[18] H. Joho, J. M. Jose, R. Valenti, and N. Sebe, "Exploiting facial expressions for affective video summarization," in *Proceedings of the ACM international conference on image and video retrieval*, 2009, pp. 1-8.

[19] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (IAPS): Technical manual and affective ratings," *NIMH Center for the Study of Emotion and Attention,* vol. 1, pp. 39-58, 1997.

[20] A. Mahata, N. Saini, S. Saharawat, and R. Tiwari, "Intelligent movie recommender system using machine learning," in *International Conference on Intelligent Human-Computer Interaction*, 2016, pp. 94-110.

[21] A. Kaklauskas, R. Gudauskas, M. Kozlovas, L. Peciure, N. Lepkova, J. Cerkauskas*, et al.*, "An Affect-Based Multimodal Video Recommendation System," *Studies in Informatics and Control,* vol. 25, p. 6, 2016.

[22] R. Baeza-Yates and B. Ribeiro-Neto, *Modern information retrieval* vol. 463: ACM press New York, 1999.

[23] N. J. Belkin and W. B. Croft, "Information filtering and information retrieval: Two sides of the same coin?," *Communications of the ACM,* vol. 35, pp. 29-38, 1992.

[24] M. J. Pazzani and D. Billsus, "Content-based recommendation systems," in *The adaptive web*, ed: Springer, 2007, pp. 325-341.

[25] S. E. Kahou, C. Pal, X. Bouthillier, P. Froumenty, Ç. Gülçehre, R. Memisevic*, et al.*, "Combining modality specific deep neural networks for emotion recognition in video," in *Proceedings of the 15th ACM on International conference on multimodal interaction*, 2013, pp. 543-550.

[26] B.-K. Kim, J. Roh, S.-Y. Dong and S.-Y. Lee, "Hierarchical committee of deep convolutional neural networks for robust facial expression recognition," *Journal on Multimodal User Interfaces,* vol. 10, pp. 173-189, 2016.

[27] M. Tkalčič, A. Košir, and J. Tasič, "The LDOS-PerAff-1 corpus of facial-expression video clips with affective, personality and user-interaction metadata," *Journal on Multimodal User Interfaces,* vol. 7, pp. 143-155, 2013.

[28] K. Schindler and L. Van Gool, "Action snippets: How many frames do human action recognition require?," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008, pp. 1-8.

[29] P. H. Bloch, "Seeking the ideal form: Product design and consumer response," *Journal of Marketing,* vol. 59, pp. 16-29, 1995.

[30] X.-X. Niu, and C. Y. Suen, "A novel hybrid CNN–SVM classifier for recognizing handwritten digits," *Pattern Recognition,* vol. 45, pp. 1318-1325, 2012.

[31] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Ijcai*, 1995, pp. 1137-1145.

[32] S. B Sadkhan, A. Q Hameed, H. A Hamed, "A proposed identification method for multi-user chirp spread spectrum signals based on adaptive Neural-Fuzzy Inference System (ANFIS) ", 2016 Al-Sadeq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), PP: 1-4

[33] M. Gaata, W. Puech, S. Sadkhan, Saad Hasson, "No-reference quality metric for watermarked images based on combining of objective metrics using neural network", 2012 3rd International Conference on Image Processing Theory, Tools and Applications (IPTA), PP: 229-234

[34] S. B Sadkhan, A. Dh. Radhi, "Fuzzy Logic used in Textual Emotion Detection", 2017 Second Al-Sadiq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), PP: 242-245.