International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

Sentiment Classification Based on Human Behavior Using Deep Neural Model

Syed Saad Husain, Dr. Salim Y. Amdani, Dr. Suresh. S. Asole
Babasaheb Naik College of Engineering, Pusad, Maharashtra, India

Article Info
Volume 9, Issue 2
Page Number : 100-107

Publication Issue :
March-April-2022

Article History
Accepted : 20 March 2022
Published: 30 March 2022

ABSTRACT

The task of automatically analyzing sentiments from a tweet has more use now than ever due to the spectrum of emotions expressed from national leaders to the average man. Analyzing this data can be critical for any organization. Sentiments are often expressed with different intensity and topics which can provide great insight into how something affects society. Sentiment analysis in Twitter mitigates the various issues of analyzing the tweets in terms of views expressed and several approaches have already been proposed for sentiment analysis in twitter. Resources used for analyzing tweet emotions are also briefly presented in literature survey section. In this paper, hybrid combination of different model's LSTM-CNN have been proposed where LSTM is Long Short Term Memory and CNN represents Convolutional Neural Network. Due to issues such as background clutter, partial occlusion, changes in scale, viewpoint, lighting, and appearance, recognising human activities from video sequences or still, images is a difficult process. Many applications, such as video surveillance systems, human-computer interaction, and robotics for human behavior classification, necessitate multiple activity recognition systems. We provide an efficient approach for human activity classification and extraction. In our project, we looked into the flaws of existing human recognition systems. We proposed using multiple frames and averaging their averages to determine the activity label rather than using a single frame to solve these restrictions. This strategy is efficient since we are averaging n frames and considering temporal storage.

Keywords : Emotion Detection, Natural Language Processing, Sentiment Analysis, Text-Based Emotion Detection

I. INTRODUCTION

Sentiments are something that one can express through various ways as it could be verbal, written or over the internet. Natural Language Processing and Python provide a well-developed tool with which one can easily get rid of the lexical adulterations and focus on the actual context to successfully predict the sentiments for the same.

In human-to-human interaction and interpersonal relationships, Human Activity Recognition (HAR) plays an important role. It provides information about a person's identification, personality, and psychological state. The ability to perceive another person's actions is a natural human trait. Computers, on the other hand, face a difficult task. In the realm of computer vision and machine learning, HAR is the most popular research. Many applications, such as video surveillance systems, human-computer interaction, and robotics for human behavior classification, necessitate multiple activity recognition systems.

Human activity recognition (HAR) is a vast topic of study that focuses on identifying a person's individual movement or action based on sensor data. Indoors, common activities such as walking, conversing, standing, and sitting are examples of movements. They could also be more targeted actions, such as those carried out in a kitchen or on a manufacturing floor. Video, radar, or other wireless technologies can be used to record sensor data from afar. Data can also be collected directly on the subject, such as through the use of specialized hardware or smartphones with accelerometers and gyroscopes.

Among the many classification techniques, two main questions arise: "What action?" and "What action?" (i.e., the recognition issue) and "Where in the video?" (This is known as the localization problem). When attempting to recognize human activities, one must first determine a person's kinetic states so that the computer can efficiently recognise the activity. Human activities like "walking" and "running" arise naturally in everyday life and are quite simple to recognise. On the other hand, more complicated operations, such as "peeling an apricot," are more difficult to recognise. Complex tasks can be decomposed into simpler activities that are more easily recognised. Detecting objects in a scene may help to better understand human activities by providing valuable information about the upcoming event.

The goal of human activity recognition is to look at activities from video or still images. Human activity recognition systems are driven by this fact to accurately classify input data into its underlying activity category. Human activities are divided into five categories based on their complexity: I gesture; (ii) atomic actions; (iii) human-to-object or human-to-human interactions; (iv) group actions; (v) behaviours; and (vi) events.

HR has been a difficult problem to solve, but it must be done. It will mostly be employed as an assistive technology in eldercare and healthcare when combined with other technologies such as the Internet of Things (IoT). HR can be carried out with the assistance of sensors, smartphones, or images. In this paper, we present a number of state-of-the-art methods and describe each one using a literature survey. Different datasets are used for each of the methods in which data is collected using various devices such as sensors, images, accelerometers, gyroscopes, and so on, and these devices are placed in various locations. The findings of each technique and the type of data set are then compared to provide a result.

The goal of human activity recognition (HAR) is to characterize a person's actions based on a set of sensor readings. Collecting this type of information is no longer a difficult undertaking. With the rise of the Internet of Things, nearly everyone now owns a device that tracks their movements. A smartwatch, a pulsometer, or even a smartphone can be used. Typically, this is done using a fixed-length sliding window approach for feature extraction, which requires two parameters to be fixed: the window size and the shift.

These are some of the data you could use:

1. Body acceleration.
2. Gravityacceleration.
3. Body angularspeed.

The machine learning model used for activity recognition is built on top of the devices' available sensors.

II. LITERATURE REVIEW

To get a proper idea about the human activity recognition system, it was important to review some of the previous works of other researchers in the field. We studied published information in this particular subject area to expand and diversify our knowledge base of this topic.

In [1] a CNN-LSTM is offered as a comprehensive deep learning-based activity recognition architecture to minimise model complexity and enhance accuracy, while in [2,] a method to automatically pull out discriminative characteristics for recognising activities is described. Essentially, the approach described here employs CNN to capture local dependency as well as signal scale invariance.

[3] introduces the scene context characteristics that characterise the subject's environment on a global and local level. A DNN structure has been described in order to provide a high-level description of human activity that combines motion and context data. A HAR system is created in [4] using a RESNET-34 3D CNN Model. The model is trained using the Kinetics data set, which has 400 classes depicting human actions in ordinary life and work, each with 400 or more films.

[5] presents a case study in which the use of a CNN feature extractor that has already been trained is evaluated in real-world scenarios. Different topologies are first assessed in order to identify the best models for human activity recognition. In this approach, a CNN model that has been pre-trained can be obtained. The model is then used as a feature extractor to evaluate its performance on a large-scale real-world dataset. In [6,] thirty-two research publications based on sensing technology used in human activity recognition were examined, with the three main categories being RGB cameras, depth sensors, and wearable devices. There is also a full discussion of the advantages and disadvantages of the sensing technology.

A fantastic method for action detection, localization, and video matching is provided in the paper [7], which is based on a hierarchical codebook model of local Spatio-temporal video volumes. For human activity recognition, convolutional layers are combined with LSTM and a deep learning neural network in [8]. The model described here uses an automated technique to extract the characteristics, which are then further classified using model attributes.

[9] compares deep, convolutional, and recurrent techniques on three datasets. These datasets are made up of movement data collected by wearable sensors. As proposed in [10],

off-the-shelf sensors from smart phones and smart watches are coupled and used to recognize human activities. This provides the optimal balance of computational complexity and recognition accuracy for the system. Several evaluations were conducted in order to identify which classification method and features should be employed.

In [11], a survey is presented that focuses on systems that try to classify full-body motions like punching, walking, waving, and so on, and they are classed based on how they describe the spatial and temporal structure of actions. [12] offers a deep neural network that uses convolutional layers and long short-term memory (LSTM) to automatically extract activity features and classify them using only a few model parameters.

[13] proposes an effective HAR approach called Inertial sensor signal to Image(Iss2Image). It's a revolutionary encoding approach that converts an inertial sensor input into an image with minimal distortion, as well as a CNN model for activity classification based on images. [14] introduces a new way for recognizing human activities. This technique has greatly improved recognition accuracy while also reducing complexity. A two-stage end-to-end CNN and data augmentation are used in this method.

[15] describes a ConvNets-based technique for activity recognition. Multiple visual signals are combined in this method. In this research, skeleton images are generated from skeleton joint sequences using a novel approach. This is how motion data is represented. [16] develops an acceleration-based human activity recognition system using a CNN model. The convolution kernel is tweaked in this case to fit the properties of tri-axial acceleration signals. A comparison of this strategy with various other ways for achieving recognition is performed on the same dataset.

[17] uses a smartphone accelerometer sensor to collect data on three human activities: walking, running, and staying stationary. A 1D CNN-based approach is used to recognize these human actions.

Various Machine learning and Deep learning methods have been introduced for sentiment analysis of tweets. State-of-the-art systems [11,12] used approaches of incorporating various models also, applied highlight vectors including semantics, syntactic elements, and word embeddings to show tweets. There have been a great deal of manners by which specialists have used to arrange Reddit remarks as "discouraged" or "not discouraged" [13]. The paper utilizes a BERT [14] based model, and a brain network with a word inserting (CNN) model for characterization. Their outcomes showed that CNN without implanting performed better compared to the BERT based model. Lexical revisions furthermore, concentrated preprocessing with an Apache System (SVM) [15] have been utilized to form a model with 5% more prominent precision than the customary opinion investigation technique. However the research plainly makes reference to in the end that the outcome will exposed to change to a huge degree relying upon the dataset utilized.

This new review [16] principally centered around twitter involves an unaided methodology for a graphical portrayal of feelings all progressively applied over scale dataset for the years 2014-2016. Sondy [17] an open source java based social elements analyzer with its primary center being client impact and occasion recognition. Lambda design is likewise programming engineering and is used alongside AI to break down enormous information streams like twitter. Lexical assets like Wilson et al. [18] named a rundown of English words in sure and negative classes. Once again [19] application for twitter has utilizes the AFINN dictionary which is a rundown of words appraised from –5 to 5 based on how positive or negative they are. Research papers as composed by broadly utilized the Stanford NLP library which produces incredible outcomes for deliberation and evacuation of information. Different Deep learning models have been used to foster start to finish frameworks in many errands including text order, discourse acknowledgment and picture characterization. Results show that these sorts of frameworks naturally separate significant level highlights from crude information [20,21].

III. METHODOLOGY

Extracting a single frame of video and then recognizing and classifying the activity in that frame is the simplest way to identify Human Activities in video segments. All of the individual frames' Human Activity can be detected, and the activity can be assigned to that frame. However, as previously stated, there is a disadvantage to this approach in that the model will not be able to achieve higher accuracy. Another disadvantage of this approach is that if we consider, for example, a person playing football on a field and a person running on a track, the activity of a person playing football on a field could be incorrectly classier as running because we are only recognising key activity.

Extracting a single frame of video and then recognizing and classifying the activity in that frame is the simplest way to identify Human Activities in video segments. All of the individual frames' Human Activity can be detected, and the activity can be assigned to that frame. However, as previously stated, there is a disadvantage to this approach in that the model will not be able to achieve higher accuracy. Another disadvantage of this approach is that if we consider, for example, a person playing football on a field and a person running on a track, the activity of a person playing football on a field could be incorrectly classified as running because we are only recognizing key activity.

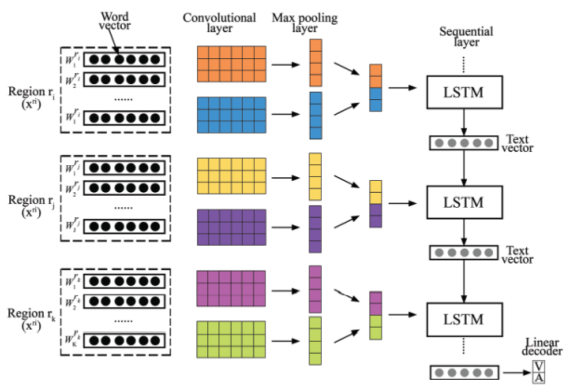


Figure 1. Deep Neural Model

Instead of classifying and showing findings for a single frame, we average outcomes over n finite frames to solve this challenge. That flashing would be effectively eliminated. We can utilise the moving average/rolling average technique with a n sized window once we've determined the value of n.

Consider,

n - Number of frames to average over

Pf - Final predicted probabilities

P - Current frame's predicted probabilities P-1 - Last frame's predicted probabilities P-2 - 2nd last frame's predicted probabilities

.

.

P-n+1 = (n-1)th last frame's predicted probabilities

We will sum the probabilities of all individual recognized activities and take the average of it. The activity with the maximum average is chosen.

For example, If we consider n as 3 and two classes i.e. [Running, Walking] then the predicted probabilities for P-2, P-1 and P are [0.95,0.05], [0.97,0.03], [0.98,0.02] respectively

Then the predicted values,

for **Running** = (0.95 + 0.97 + 0.98)/3 = **0.97,**

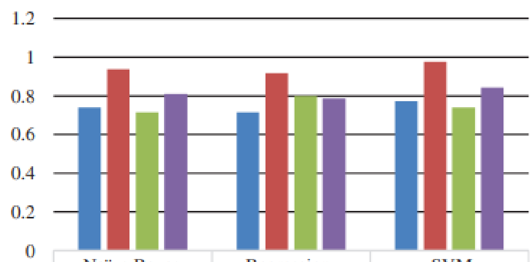
for **Walking** = (0.05 + 0.03 + 0.02)/3 = **0.03**

As 0.97 (Running score) > 0.03 (Walking score),

hence, Prediction = Running.

```
For each token t in review
    #Extract Explicit Aspects
    If (isToken_AspectTerm(t))
        aspectCategory =getAspectCategoryFromLexicon(t)
        aspectCategories.append(aspectCategory)
        aspectIndices.append(token_index)
    #Extract Implicit Aspects
    If (isToken_SentimentWord(t) AND
                getNearestAspect(token_index, aspectIndices[])==null)
        aspectCategory =getAspectCategoryFromLexicon(t)
        aspectCategories.append(aspectCategory)
        aspectIndices.append(token_index)
End For
```

Using this implementation we can get rid of flickering. Apart from averaging the n frames there is also a need to store temporal information of the sequence of frames as the model will not be able to distinguish between a similar set of actions of a person in the same environment as averaging would not be useful here if we do not know the sequence of frames ex: Action of a person sitting down and standing up on a chair will have the same values of average probability. So we also need to consider its sequence.



	Naïve Bayes	Regression	SVM
Accuracy	0.7407	0.7143	0.7719
Precision	0.9375	0.9167	0.9756
Recall	0.7143	0.8	0.7407
F1 Score	0.8108	0.7857	0.8421

So, we use Single-Frame CNN Architecture in which we average the probability of n frames and also account for the sequence of frames, utilizing the temporal information in a video to solve the above issues, considering the environmental context. This approach works efficiently as we are averaging n finite frames. Hence we can also take a few spread out frames out of n frames to avoid unnecessarily classifying all n frames.

IV. CONCLUSION AND FUTURE SCOPE

Human Activity Recognition (HAR) provides information on a person's identity, personality, and psychological state which helps for sentiment classification. It is important in human-to-human interaction and interpersonal relationships. We investigated the shortcomings of existing human recognition systems in our project. Instead of utilizing a single frame to address these limitations, we recommended employing numerous frames and averaging their averages to determine the activity label. Because we are averaging n frames and considering temporal storage, this technique is efficient. As a result, we were able to propose a systematic method for the HAR system. The accuracy that our model gives is 98%.

In the future, our application can be integrated with various other applications. HAR can be used in daily life monitoring applications and elderly care applications. We can also combine HAR with smart surveillance systems, medical research, automatic sports commentary, and other applications. It can also be used to monitor social distance, monitor public places and to prevent crimes.

V. REFERENCES

[1]. Mutegeki, R., & Han, D. S. (2020). A CNN-LSTM Approach to Human Activity Recognition. 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). doi:10.1109/icaiic48513.2020.9065078

[2]. Zeng, M., Nguyen, L. T., Yu, B., Mengshoel, O. J., Zhu, J., Wu, P., & Zhang, J. (2014). Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors. Proceedings of the 6th International Conference on Mobile Computing, Applications and Services. doi:10.4108/icst.mobicase.2014.257786

[3]. Li Wei and Shishir K. Shah, Computer Science Department, University of Houston, 3551 Cullen Blvd., Houston, TX 77204, U.S.A., HumanActivity Recognition using Deep Neural Network with Contextual Information,

[4]. Ms Shikha, Rohan Kumar, Shivam Aggarwal, Shrey Jain, Human Activity Recognition, International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-9 Issue-7, May 2020

[5]. Cruciani, F., Vafeiadis, A., Nugent, C., Cleland, I., McCullagh, P., Votis, K., … Hamzaoui, R. (2020). Feature learning for Human Activity Recognition using Convolutional Neural Networks. CCF Transactions on Pervasive Computing and Interaction. doi:10.1007/s42486-020-00026-2

[6]. Ann, O. C., & Theng, L. B. (2014). Human activity recognition: A review. 2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014). doi:10.1109/iccsce.2014.7072750

[7]. Javan Roshtkhari, M., & Levine, M. D. (2013). Human activity recognition in videos using a single example. Image and Vision Computing, 31(11), 864–876. doi:10.1016/j.imavis.2013.08.005

[8]. Ankita, Shalli Rani, Himanshi Babbar, Sonya Coleman, Aman Singh and Hani Moaiteq Aljahdali, An Efficient and Lightweight Deep Learning Model for Human Activity Recognition Using Smartphones, Sensors 2021, 21, 3845. https://doi.org/10.3390/s21113845

[9]. Hammerla N, Halloran S, Plötz T. Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables. Twenty-Fifth International Joint Conference on Artificial Intelligence. 9-15 July 2016, New York: AAAI Press / International Joint Conferences on Artificial Intelligence.

[10]. Bishoy Sefen, Sebastian Baumbach, Andreas Dengel, and Slim Abdennadher, Human Activity RecognitionUsing Sensor Data of Smartphones and Smartwatches, ICAART-1, 8th Int'l Conference on Agents and Artificial IntelligenceAt: Rome, ItalyVolume: 2

[11]. Daniel Weinland, Rémi Ronfard, Edmond Boyer. A survey of vision-based methods for action representation, segmentation and recognition. Computer Vision and Image Understanding, Elsevier, 2011, 115 (2), pp.224-241. 10.1016/j.cviu.2010.10.002 hal-00640088

[12]. Xia, K., Huang, J., & Wang, H. (2020). LSTM-CNN Architecture for Human Activity Recognition. IEEE Access, 8, 56855–56866. doi:10.1109/access.2020.2982225

[13]. Hur, T., Bang, J., Huynh-The, T., Lee, J., Kim, J.-I., & Lee, S. (2018). Iss2Image: A Novel Signal-Encoding Technique for CNN-Based Human Activity Recognition. Sensors, 18(11), 3910. doi:10.3390/s18113910

[14]. Huang, J., Lin, S., Wang, N., Dai, G., Xie, Y., & Zhou, J. (2020). TSE-CNN: A Two-Stage End-to-End CNN for Human Activity Recognition. IEEE Journal of Biomedical and Health Informatics, 24(1), 292–299. doi:10.1109/jbhi.2019.2909688

[15]. Khaire, P., Kumar, P., & Imran, J. (2018). Combining CNN streams of RGB-D and skeletal data for human activity recognition. Pattern Recognition Letters. doi:10.1016/j.patrec.2018.04.035

[16]. Chen, Y., & Xue, Y. (2015). A Deep Learning Approach to Human Activity Recognition Based on Single Accelerometer. 2015 IEEE International Conference on Systems, Man, and Cybernetics. doi:10.1109/smc.2015.263

[17]. Song-Mi Lee, Sang Min Yoon, & Heeryon Cho. (2017). Human activity recognition from accelerometer data using Convolutional Neural Network. 2017 IEEE International Conference on Big Data and Smart Computing (BigComp). doi:10.1109/bigcomp.2017.7881728

[18]. P. Goel, D. Kulshreshtha, P. Jain and K. Shukla, "Prayas at emoint 2017: An ensemble of deep neural architectures for emotion intensity prediction in tweets," in Proc. of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Copenhagen, Denmark, pp. 58–65, 2017.

[19]. V. Duppada, R. Jain and S. Hiray, "Seernet at semeval-2018 task 1: Domain adaptation for affect in tweets," in Proc. of the 12th Int. Workshop on Semantic Evaluation, The Association for Computational Linguistics, Stroudsburg, vol. 2, pp. 18–23, 2018.

[20].M. M. Tadesse, H. Lin, B. Xu and L. Yang, "Detection of depression-related posts in reddit social media forum," IEEE Access, vol. 7, pp. 44883–44893, 2019.
[21].J. Devlin, M. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," NAACL-HLT 2019, Minneapolis, Minnesota, vol. 1, pp. 4171–4186, 2019.
[22].F. Bravo, M. Mendoza and B. Poblete, "Combining strengths, emotions and polarities for boosting twitter sentiment analysis," in Proc. of the Second Int. Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM'13), Association for Computing Machinery, New York, NY, USA, pp. 1–9, 2013.
[23].N. Azzouza, K. Astouati,A. Oussalah and S. AitBachir, "A real-time twitter sentiment analysis using an unsupervised method," in Proc. of the 7th Int. Conf. on Web Intelligence, Mining and Semantics (WIMS'17), Association for Computing Machinery, New York, NY, USA, pp. 1–10, 2017.
[24].A. Guille, C. Favre, H. Hacid and D. Abdelkader, "An open-source platform for social dynamics mining and analysis," in Proc. of SONDY, United State, pp. 1005–1008, 2013.
[25].M. Wilson, W. Moss, E. Helen and V. Halen, "Perceptual distance and competition in lexical access," Journal of Experimental Psychology, Human Perception and Performance, vol. 22, pp. 1376–1392, 1997.
[26].F. Årup, "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs," in Proc. of the ESWC2011 Workshop on 'Making Sense of Microposts': Big Things Come in Small Packages 718 in CEUR Workshop Proc., Ithaca, USA, pp. 93–98, 2011.
[27].Y. Cun, Y. Bengio and G. Hinton, "Deeplearning," Nature, vol. 521, pp. 436–444, 2015.
[28].D. Tang, B. Qinand T. Liu, "Deep learning for sentiment analysis: Successful approaches and future challenges," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 5, pp. 292–303, 2015.

Cite this article as :

Syed Saad Husain, Dr. Salim Y. Amdani, Dr. Suresh. S. Asole "Sentiment Classification Based on Human Behavior Using Deep Neural Model", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 9 Issue 2, pp. 100-107, March-April 2022.
Journal URL : https://ijsrset.com/IJSRSET229229