iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52033

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Analysis of Walking Pattern using LRCN for Early Diagnosis of Dementia in Elderly Patients

Rimsha Salahuddin[1], Pooja Khanna[2], Sachin Kumar[3], Pragya[4]
[1, 2, 3]*Amity University Uttar Pradesh, Lucknow Campus*
[4]*MVD College, Lucknow University*

Abstract: ***Analysis of walking pattern using LRCN for early diagnosis of Dementia in elderly patients has emerged as one of the most important study areas in the fields of health and human-machine interaction in recent years. Many artificial intelligence-based models for activity recognition have been developed; however, these algorithms fail to extract spatial and temporal data, resulting in poor real-world long-term HAR performance. Furthermore, in the literature, a restricted number of datasets for physical activity recognition with a smaller number of activities are publicly available. Given these constraints, we create a hybrid model for activity recognition that combines Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). Long-Short Term Memory (LSTM) networks are designed to work with sequence data, while Convolutional Neural Networks (CNN) are helpful for improving understanding of picture information. CNN can be used to extract spatial data at points in a series (video). The LSTM layer is granted authority for temporal sequence modelling with extracted spatial characteristics in situ, whereas the convolution layer is responsible for evaluating spatial features from the frames. 50 activity categories, 25 video groups per activity category, 133 medium videos per activity category, and more make up the UCF50- Action Recognition Dataset. 320 frames per second, 199 frames per video, 240 frames each for each video. Each video is an average of 26 intermediate frames per second. We determine which model performs best by analysing its Loss and Accuracy curves, then we test that model using certain testing films. The drawback of this paper is that most persons performing various duties cannot use our vision.***

Keywords: analysis of walking pattern; convolutional neural network; deep learning; long short-term memory; machine learning.

I. INTRODUCTION

This The analysis of walking patterns using LRCN for early detection of dementia in senior patients is important in people's daily lives since it can learn advanced information about human movements from raw sensor data. The rise of human-computer interaction applications has made HAR technology a prominent research direction both at home and abroad. People may be able to automatically classify the type of human motion and collect the data that the human body requires by extracting features from everyday activities, setting the framework for future intelligent applications. Automatic detection of everyday activities has the potential to aid in the treatment of ailments such as obesity, diabetes, and cardiovascular disease. Moderate to intense physical exercise, for example, relates to lower risk factors for obesity, cardiovascular and pulmonary diseases, cancer, depression, and improved bone health. As a result, proper physical activity measurement is essential for developing intervention methods. It also provides rich contextual data from which more significant data can be deduced.

The most recent research on activity recognition (AR) demonstrates that it is possible to identify human behaviours by using artificial intelligence (AI) techniques, particularly machine learning (ML) techniques, on sensor data. Using ML techniques on wearable accelerometers has been most effective, and these gadgets are probably the most developed sensors for identifying single-user basic behaviours like running, walking, standing, sitting, lying, and similar ones. Accelerometers can measure human motion, mostly by measuring linear 3D accelerations, and can also estimate body postures, primarily by measuring orientation regarding the Earth's gravity. Multi-accelerometer systems have already demonstrated their capacity to accurately identify various actions.

Deep learning (DL) has evolved as a revolutionary approach to classical machine learning in recent years. DL is capable of high-level data abstraction, allowing for robust models capable of dealing with the high noise associated with AR challenges. Each layer in a typical DL architecture combines features (output) from the preceding layer and modifies them using a non-linearity function to generate a new feature set.

©IJRASET: All Rights are Reserved

This enables the network to automatically learn the best features for a given problem domain, producing a hierarchy in which fundamental characteristics are identified in the initial layers of the network and abstract elements from previous layers are fused to build complex feature maps in the deeper layers. Where computer vision, speech recognition, and natural language processing are concerned, DL is already cutting-edge.



Figure 1. Diversified Walking Pattern

Many artificial intelligence-based models are developed for activity recognition; however, these systems perform poorly in the real world on long-term HAR due to their inability to extract spatial and temporal data. Given these constraints, we create a hybrid model for activity detection that incorporates Convolution Neural Network (CNN) and Long Short-Term Memory (LSTM), with the LSTM network learning temporal information and CNN extracting spatial data.

A. Image Recognition

Image recognition is a technique used to analyse a frame, assess, or identify the objects inside it, and associate them with the category to which they belong. When we see anything or a group with our eyes, we immediately classify and correlate objects with various scenarios. Each has its own significance. However, the most challenging task that machines can complete is visual recognition. Object recognition is the process of grouping the recognized objects into either the same category or a distinct one.

Machine learning, particularly deep reading technology, has had considerable success in numerous computer vision and imaging tasks throughout the years. Many deep learning techniques that are renowned for being adaptable as In-Depth Learning algorithms are employed to achieve the best results in picture identification. Although the terms image recognition and image acquisition are synonymous, they differ technologically. After a photo or frame is taken, the spread-out objects in the image are detected by image acquisition. Face detection is an illustration of image recognition, in which algorithms are designed to find facial features in photos. Whether the results are meaningful in any manner is irrelevant if we are serious about acquisition. To differentiate one object from another and ascertain how many distinct businesses are contained inside an image is the purpose of image acquisition.

Each item has a set number of binding boxes. However, the conventional approach to computer vision calls for a high level of technical knowledge, extensive engineering time, and includes numerous parameters that must be determined in person, while the functionality of other functions is limited.

B. Image Recognition Systems in Organized Steps

1) *Dataset:* Every approach to visualization requires training data, which can either be a video or a picture. In order for neural networks to generate ideas for what particular classes look like, they need those training images from the collected database.
2) *Training of Neural Networks for Image Recognition:* The neural network algorithm includes captured frames that are inserted from organized databases. This is a comprehensive machine learning feature for building a model for photo recognition. image recognition to visualize the convolutional neural network before recognizing particular categories a training algorithm is required.

3) *Testing the AI model:* This module tests the use of images from the learning database. This is used to evaluate the model's performance, accuracy, and usefulness. As a result, between 80 and 90 percent of the entire image library is used for model training, with the remaining data being saved for testing a model. The effectiveness of a model is evaluated using a set of metrics that show the accuracy of identification, the dependability % of each image used for testing, and other factors.

C. Image Recognition Applications

AI imaging technology is becoming more and more necessary across the board. Its uses add economic value to a variety of sectors, including agriculture, security, retail, and health care.

1) *Face Analysis:* Systems can determine identity, goals, emotional states, physical health, age, or nationality using computer-assisted face recognition. Even measuring perceived levels of school attraction is a goal of some visual aides. Other facial recognition-related tasks include face recognition, face recognition photos on a website can be identified and matched using visual processing algorithms, as well as facial verification.
2) *Medical Image Analysis:* The application of medical image analysis in the medical industry is now recognized as a minor component of AI's larger advantages. For instance, melanoma, a dangerous skin cancer, is diagnosed using a variety of processes. Software for deep picture recognition enables monitoring of the tumour over time, for instance, abnormal breast cancer scan detection. Covid-19 characteristics are recognized using the in-depth image learning algorithm COVID-NET. - Using AI recognition to fight COVID.
3) *Object Detection:* AI image recognition software and video recognition software can be used to recognize characters, patterns, logos, items, places, objects, forms, and colours. Image recognition can be used while coexisting with a variety of software thanks to customization. For instance, once a popular computer vision programme in grocery stores, a photo-viewing system, is tailored to find people and can be employed by counting people.
4) *Food Image Recognition:* With computer-assisted food testing, image recognition is employed to study various food varieties in depth. By looking at food photographs captured by mobile devices, photo recognition software applications are used to improve the accuracy of original dietary measures.
5) *Image Search Recognition:* This image recognition system explores the most effective and simple methods of image detection by learning visual features using deep neural networks. Researchers have put up a sizable image dictionary that serves as a training set for neural network features to address this stretching problem.

II. RELATED WORK

Activity recognition in videos has potential applications in multiple fields such as multimedia indexing and surveillance. [1][2] Human activity recognition aids the daily activities as well as diseases like dementia [3]. Cognitive disorders are on the increase today, owing to the advancement of technology [4] [5]. According to the World Health Organization, around ten million new cases of dementia are diagnosed each year [6]. Deep learning, a method of analysing and evaluating human behaviour, has attracted a lot of attention and interest from academics as a possible solution to this problem. [8]. These data are used to train a variety of machine learning and deep learning models, including CNNs, LSTMs, ConvLSTMs, and LRCNs, to learn self-detecting behaviour and detect anomalous behaviours like fainting or not reacting when alone in a dwelling. [9][10]. After sensing the activity, the person in charge of monitoring will be able to take immediate action to hospitalize the elderly or cognitively impaired. [11]][12]. Many studies use Convolutional Neural Networks with Long Short-Term Memory (CNN-LSTM) to recognize and identify human physical actions utilizing video analysis acquired from an internal source to deliver an effective solution. Human Activity Recognition using Machine Learning Sowmya et al. [13] designed a machine learning-based system for detecting and recognizing human activity. Sandeep [14] created a method for recognizing human movement on cellphones using machine learning. He used smart phones with gyroscopes, accelerometers, GPS sensors, and compass sensors, among other motion-detecting gadgets. Subasi et al. [15] developed a system for recognizing human activities in a smart healthcare setting by using machine learning techniques. They demonstrated how these techniques are utilized to perform robust and accurate HAR on two distinct datasets. Singh et al.[16] proposed a human activity recognition using data collected from smart home sensors. They applied several machine learning algorithms such as Naive Bayes (NB), Decision Trees , HMM , CRF , Nearest Neighbor (NN) and Support Vector Machines (SVM). Abbaspour et al. [17] modeled various combinations of CNNs and RNNs to process and classify human activity data. Orozco et al [18] used a Robust CNN LSTM approach to describe Human Action Recognition in Videos. They analyzed the CNN-LSTM architectures using the KTH, UCF-11, and HMDB-51 datasets. Deotalea [19] employed Deep Learning to recognize human activity in untrimmed video for the sports domain.

For Recognizing Complex Human Actions in Live Videos Hybrid FR-DL Method was suggested by Fatemeh Serpush and Mahdi Rezaei [20]. KUN XIA et al. [21] developed a classic pattern recognition approach for human activity recognition based on LSTM-CNN architecture. Their approaches rely on fixed sensors, which include acoustic sensors, radars, static cameras, and other ambient-based sensors.

III.PROPOSED SYSTEM

One Convolutional Neural Networks (CNNs) excel at picture data, while Long-Short Term Memory (LSTM) networks excel at sequential data. When the two are combined, you have the best of both worlds and can handle tough computer vision problems. as in video split. To accomplish this, we will employ two distinct TensorFlow structures and methods.

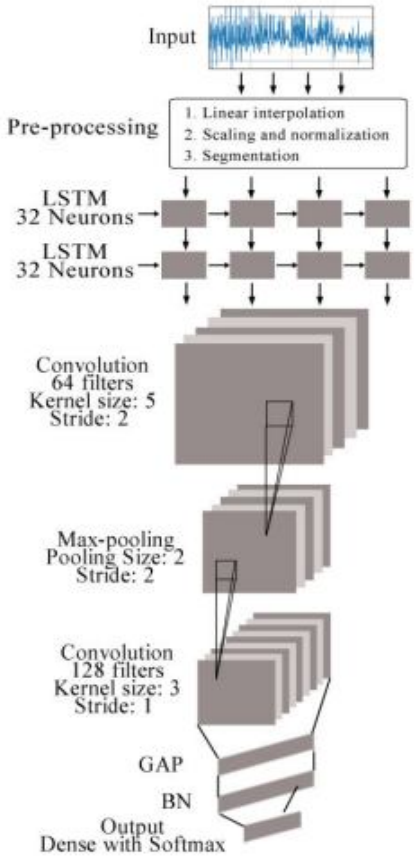


Figure 2. Frame diagram of LSTM-CNN Model

Finally, we'll use the most effective model to generate predictions on YouTube videos. In the beginning, we download and visualize the data and labels to get a sense of what we're up against. We will use UCF50 - An Activity Data Recognition Set, which includes real-time YouTube videos that distinguish this data set from many others. Systems of recognition that are not real and are used by players.

A. The UCI-HAR

UCI dataset was built from the recordings of 30 subjects aged 19-48 years. During the recording, all subjects were instructed to follow an activity protocol. And they wore a smart phone (Samsung Galaxy S II) with embedded inertial sensors around their waist. The six activities of daily living are standing (Std), laying (Lay), walking (Walk), walking downstairs (Down) and walking upstairs (Up). In addition, this dataset also includes postural transitions that occur between the static postures: standing to sitting, sitting to standing, sitting to laying, laying to sitting, standing to laying, laying to standing. Specifically, in this paper, only six basic activities were selected as input samples due to the percentage of postural transitions is small. The experiments had been video-recorded to manually label the data. Finally, data on 3-axial acceleration and 3-axial angular velocity were recorded at a fixed rate of 50Hz.

Dataset Contains

- 50 Action Classes with a total video of 6671
- 25 Video Groups per Activity Category
- 133 Medium Videos per Activity Category
- 199 Average number of frames per video
- 320 Frame Range for Each Video
- 240 Average Frame Length Each video

Medium frames 26 per second per video

B. Convolution Neural Networks

A subcategory of neural networks, convolutional neural networks exhibit all the traits of neural networks. Since an image is a nxn-dimensional matrix, the convolution is calculated using an input image of the same size and a filter that can travel along the matrix. The kernel or filter is the portion of the matrix used to generate the output that produces a 3x3 matrix for the same.

The kernel processes the input image. A stride is the name for the 1-pixel side-over that the kernel experiences. Additionally, after each slide, multiply a portion of the matrix under the kernel by the output total to get an integer that forms a component of the feature matrix. Feature Map is another name for the Convolved Feature Matrix.

The 3x3 matrix functions as both a filter and a kernel, and it is the same sliding matrix. The input image in real-world data may be fairly large and contain a great deal of information, the majority of which is unimportant for our deep learning or classification challenge. To address a issue on our hands it uses a kernel or filters.

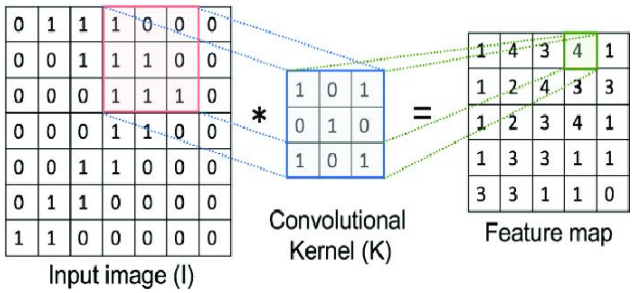


Figure 3. Working of Kernel with 3x3

C. Feature Maps

Feature Maps are the results produced by repeatedly applying convolution to the sub regions of an entire image. Initial pass's input comes from the previous layer. One pixel is moved at a time during convolution using the nxn kernel and a stride of 1. Convolution is performed over a specific sub region of the input image, which activates neurons and produces output containing information about features like edges and curves. This output is then compiled into a feature map.

Long-Short Term Memory (LSTM):

A unique type of RNN recognized for adapting long-term dependency is called as a long short-term memory network. LSTMs were created. A form of repeated set of neural network modules is present in all developing neural networks.

The repeating module of LSTMs interacts with each of the network's layers in a distinctive way.

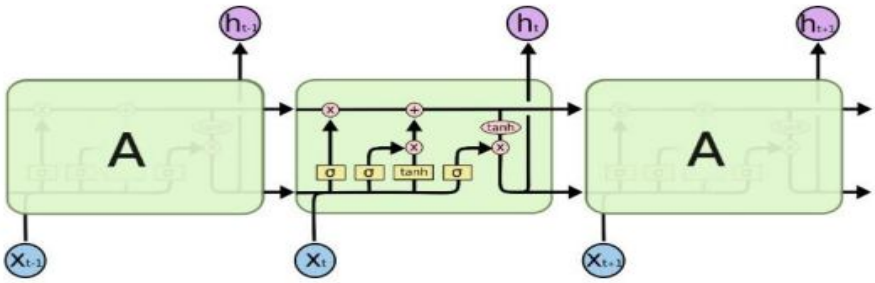


Figure 4. Communicating Layers in LSTM Cell

1) Analyse and Pre-process Data
- Load data set
- Use various charts and metrics to better understand the data.
- Data is never available in the format that is needed for training; therefore, data pre-processing is necessary.
- Potentially construct new features.
- Normalize features.
- One hot encoding of labels.

2) Separate Training and Test Data Set
- In case that data is not already separated into a training and test data set, split the data.
- Lock the test data away and do not touch it until the final evaluation.

3) Create Model/Estimators
- Decide on the type of machine learning technique to apply based on the given problem statement.
- Create the ML model/estimator.
- Define necessary hyper parameters such as learning curve, epochs, batch size, etc.
- Separate the validation data set from the training set.

4) Train Model and Evaluate
- Fit the chosen/model estimator.
- Review the performance of the training via learning curve, confusion matrix, etc.
- If necessary, modify the existing model/estimator and re-run the training.
- Continue until a good model is found which fits well regarding training and validation accuracy and loss/cost.

5) Evaluate Model Against Test Data
- In the final step, evaluate the chosen model against the hold out set/test data set.
- This step should only be executed once in order to avoid that the test set influences the definition of your target model that you derived.

Store the model for future use/ improvements.

IV. IMPLEMENTATION

To observe the first flow of the selected films written by associated labels, we will select all of the random categories in the database and the random video in each selected category. We will be able to visualize all areas of the database in this manner.

The database will next go through some preliminary processing. In order to minimize the numbers and make the data normal to distance [0-1], we will first examine video files in the database and adjust the frame size of the movies to a specific width and length. IMAGE_HEIGHT, IMAGE_WIDTH, and SEQUENCE_LENGTH values, allowing for faster integration when training the pre-specified network. The sequence length can be extended for the best outcomes, but this only affects a small portion of the problem, and extending the values will make the computer process more expensive.

A list of framed and standard frames for the video is created by the frames_extraction () function, which takes the channel as an input. Although not all frames are added to the list since we only require a small subset of them, the method will read the video file frame by frame. A length distribution across the sequence. To recover features, class labels, and video_files_paths, execute the frame_extraction () function in all video files for the selected categories after creating a create_dataset () function that will replicate in all classes listed in CLASSES_LIST on a regular basis. With a single hot code, released class labels (class references) are transformed into vectors.

We now have the required features, including label_hot_encoded_one (a NumPy array storing all class labels in one hot coding format) and a NumPy list containing all video output frames. Therefore, we will now divide our data into training and testing sets. Prior to fragmentation, we will also perform database manipulation to remove bias and identify variances that accurately reflect the distribution of the data overall.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

In the fourth stage, we use a ConvLSTM cell combination to apply the first technique. An LSTM network with convolution functions is referred to as having ConvLSTM cells. It is an architecture-centric convolutional LSTM that enables the identification of regions of data while considering temporal linkages. This technique uses video segmentation to capture local associations in individual frames and fleeting interactions over all frames. ConvLSTM may accept 3-dimensional inputs (width, length, number_ of channels) because of this transformation structure, however plain LSTM can only accept 1-dimensional inputs, which is why LSTM cannot integrate to produce a spatiotemporal data model in it. We use Keras ConvLSTM2D's repeating layers to generate the model. The number of filters and kernel size required for the convolutional operation are also recorded by the ConvLSTM2D layer. By activating the layer output, which is flat at the end, a dense layer is created. The softmax, which eliminates all potential outcomes for each action stage.

We decrease the number of frames, do away with unnecessary calculations, and pull layers to avoid over-modelling the data by employing maxpooling3D layers. Architecture is straightforward and has few formal requirements. This is because we are only working with a limited data set that call for a big model. To generate the necessary convlstm model, we will now use the create_convlstm_model () method developed earlier.

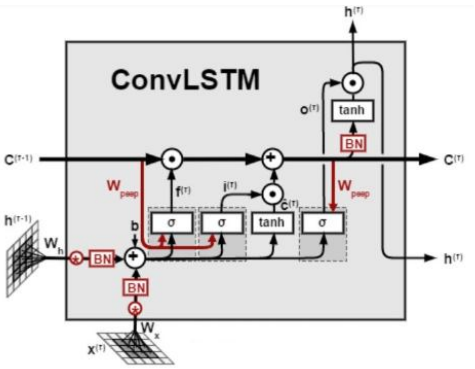


Figure 5. ConvLSTM model Network

Following that, we will add an early stopping call back to prevent overfitting and begin training after building the model with 50 epochs and a batch size of 4.

The Convolution and LSTM layers are combined into a single model using the LRCN Approach in the fifth phase. Using the CNN model with the independently trained LSTM model is another way that is similar. Local features can be extracted from data using a CNN model. While the LSTM model can predict performed action on video, frames in a video can also be used for this purpose. A pre-trained model can be used for this, but the LSTM model can be better prepared for the challenge. However, in this instance, we'll employ a different technique called the Long-term Recurrent Convolutional Network (LRCN), which combines the CNN and LSTM layers into a single model. A local element is excluded from frames using convolutional layers, and the LSTM layer is supplied extruded local features. For each step in the sequence modelling time, convolutional layers are utilized to exclude a local element from frames, and extruded local features are sent to the LSTM layer. In this manner, the network learns. Spatial-temporal features are used directly in the final training, resulting in a more robust model.

We'll also utilize the TimeDistributed wrapper layer, which allows the same layer to be applied to a full video frame independently. As a result, if the layout is initially rounded, construct a layer that can accept input (none_from_mames, width, length, number_ of channels). The layer arrangement was (width, length, number_ of channels), which is quite beneficial because it allows you to incorporate the full video in a single shot.

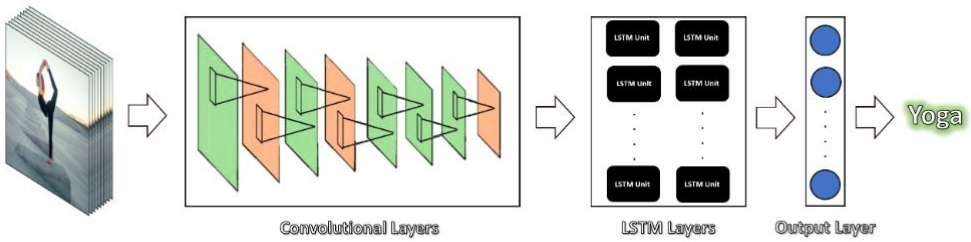


Figure 6. LRCN Model Network

To implement our LRCN architecture, we will initially employ Conv2D levels, followed by MaxPooling2D and Dropout layers. The Conv2D layers' collected features will be flattened with the Flatten layer and delivered with the LSTM layer. The output of the LSTM layer is then used by a dense layer with a softmax implementation to forecast the action to be executed. We will now utilize the previously constructed create_LRCN_model() function to generate the desired LRCN model. We will assemble and commence construction after examining the structure.

The plot_metric () function is used to rank the models after they have been tested and saved for losses and precision curves. For each model, this function will display training and certification metrics. Model loss curves are a method that employs a total loss associated with no validation left. Similarly, the models' precision curves are ranked from complete perpendicular accuracy to complete verification accuracy. The model was trained with 70 epochs and a batch size of four.

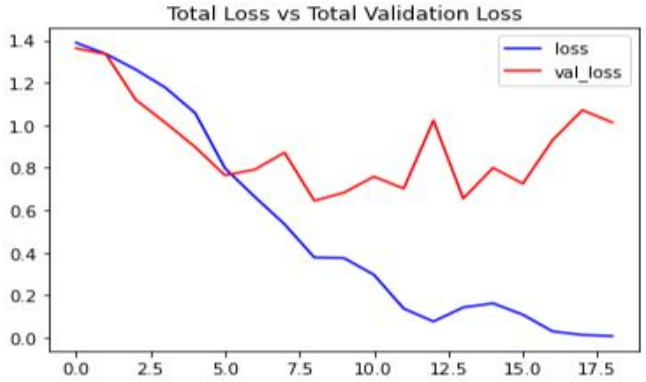


Figure 7. Total Loss Vs Total Validation Loss for ConvLSTM Model Network

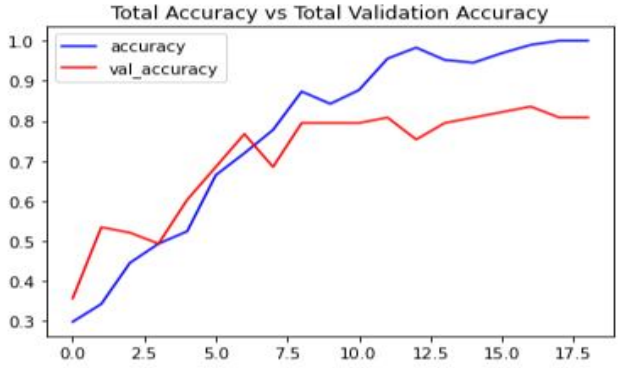


Figure 8. Total Accuracy Vs Total Validation Accuracy for ConvLSTM Model Network

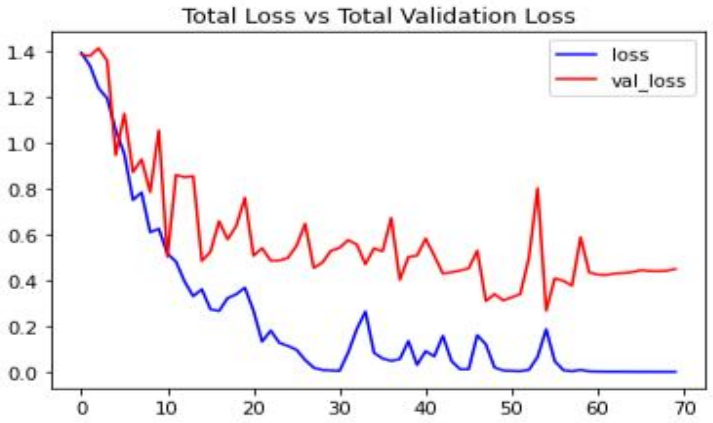


Figure 9. Total Loss Vs Total Validation Loss for LRCN Model Network

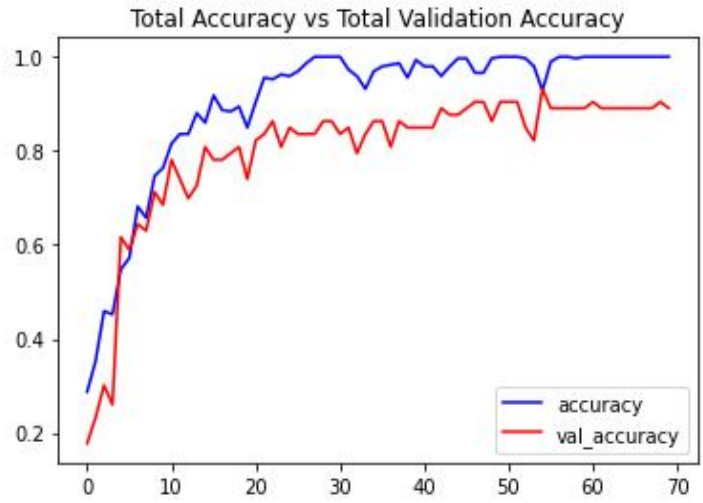


Figure 10. Total Accuracy Vs Total Validation Accuracy for LRCN Model Network

According to the structural metrics, LRCN outperformed ConvLSTM with 92% accuracy and 20% loss. While ConvLSTM achieved an accuracy of 80%. According to the results, the LRCN model performed exceptionally well in a small number of classes. As a result, in this stage, we will create an LRCN model to test in future YouTube videos. We will first construct a download function for_youtube_videos () to download YouTube videos.

The pafy library is being used. The library simply requires the URL in the video and its associated metadata as the title of the video to download it. Next, we'll write a predict_on_video () function that will automatically read the frame of the video frame in a path sent as an input, conduct action recognition on the video, and save the results.

V. CONCLUSIONS

Work suggested a CNN-LSTM technique to human activity recognition using the UCF50-Human Activity Recognition dataset in this research. This method aids in robust frame extraction from a movie by utilizing both the CNN and LSTM models, both of which perform admirably with the extraction of spatial and temporal features, respectively. When compared to other deep learning algorithms that employ films as a dataset, the ConvLSTM and LRCN models performed better. For each of these CNN-LSTM model designs, we evaluated the model metric using total loss vs total validation loss and total accuracy vs total validation accuracy. The accuracy of the LRCN model was 92%, whereas the accuracy of the ConvLSTM model was 80%. When compared to the ConvLSTM model, LRCN training took less time. This model is used with complex activities as dataset categories to tackle activity recognition.

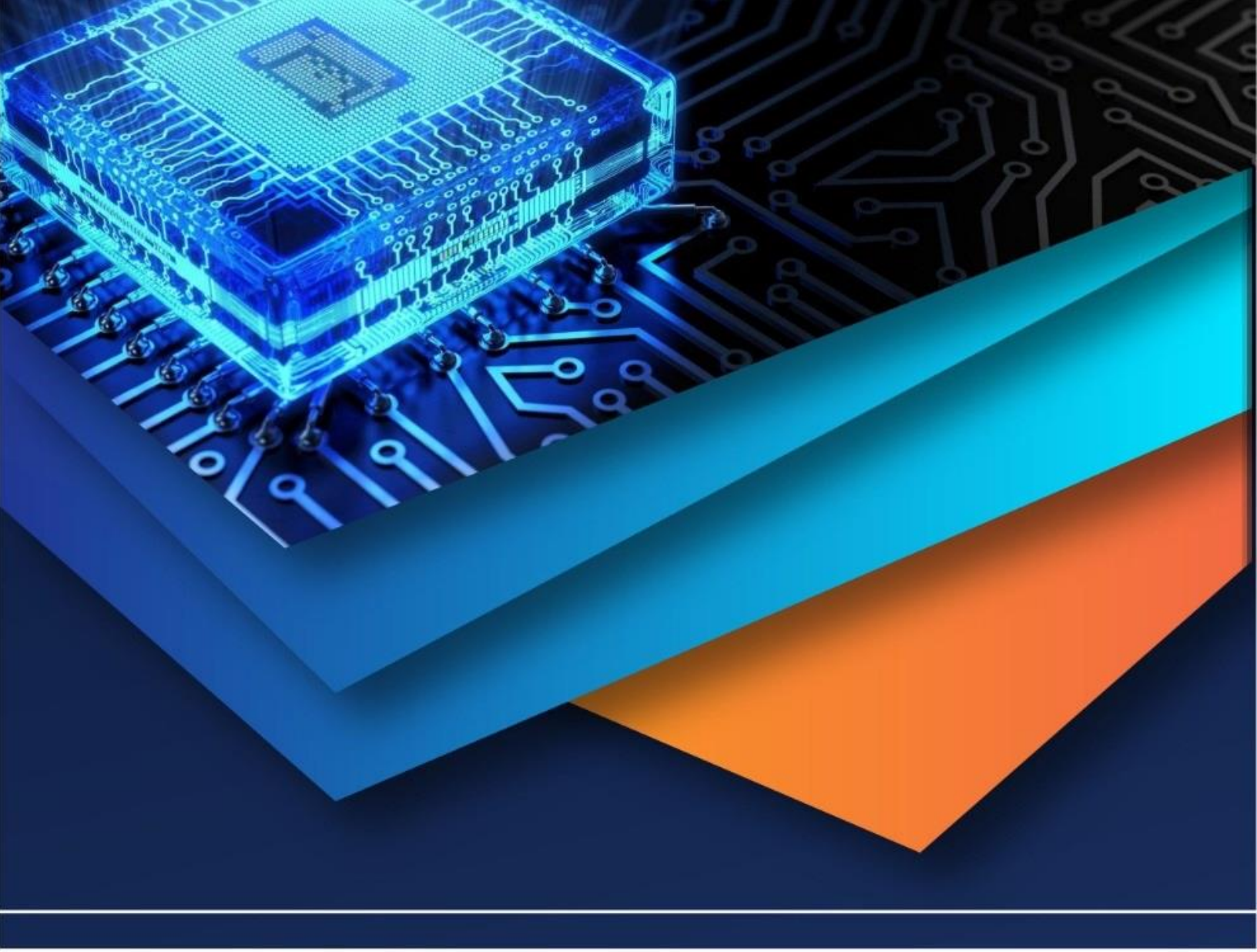
The model will be trained in the future to recognize activity on several people performing diverse tasks in the frame. The data used to train this asset in the model should be annotated for more than one person's activity and should additionally include the bounding box coordinates. Cropping out each individual and doing activity recognition separately on each person is a hacky technique to execute activity recognition separately on each person, but this will be quite expensive.

REFERENCES

[1] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, “Machine recognition of human activities: A survey,” IEEE Transactions on Circuits and Systems for Video technology, vol. 18, no. 11, pp. 1473–1488, 2008.
[2] M. Vrigkas, C. Nikou, and I. A. Kakadiaris, “A review of human activity recognition methods,” Frontiers in Robotics and AI, vol. 2, p. 28, 2015.
[3] S. Rong, G. Xu, B. Liu, Y. Sun, L. G. Snetselaar, R. B. Wallace, B. Li, J. Liao, and W. Bao, “Trends in mortality from parkinson disease in the united states, 1999–2019,” Neurology, vol. 97, no. 20, pp. e1986–e1993, 2021.
[4] N. Roy, A. Hassan, R. Alom, M. Rajib, and K. Al-Mamun, “The situation of alzheimer’s disease in bangladesh: Facilities, expertise, and awareness among general people,” Journal of Neurological Disorders, vol. 8:7, p. 7, 01 2021.
[5] D. Aarsland, K. Andersen, J. P. Larsen, and A. Lolk, “Prevalence and Characteristics of Dementia in Parkinson Disease: An 8-Year Prospective Study,” Archives of Neurology, vol. 60, no. 3, pp. 387–392, 03 2003. [Online]. Available: https://doi.org/10.1001/archneur.60.3.387
[6] K. N. R. Challa, V. S. Pagolu, G. Panda, and B. Majhi, “An improved approach for prediction of parkinson’s disease using machine learning techniques,” in 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES). IEEE, 2016, pp. 1446–1451.

[7] M. Rahman, M. Uddin, J. Chowdhury, and T. Chowdhury, "Effect of levodopa and carbidopa on non-motor symptoms and signs of parkinson's disease," Mymensingh medical journal : MMJ, vol. 23, no. 1, p. 18—23, January 2014. [Online]. Available: http://europepmc.org/abstract/MED/24584367
[8] V. Bianchi, M. Bassoli, G. Lombardo, P. Fornacciari, M. Mordonini, and I. De Munari, "Iot wearable sensor and deep learning: An integrated approach for personalized human activity recognition in a smart home environment," IEEE Internet of Things Journal, vol. 6, no. 5, pp. 8553–8562, 2019.
[9] D. Liciotti, M. Bernardini, L. Romeo, and E. Frontoni, "A sequential deep learning application for recognising human activities in smart homes," Neurocomputing, vol. 396, pp. 501–513, 2020.
[10] V. Jacquot, Z. Ying, and G. Kreiman, "Can deep learning recognize subtle human activities?" in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 14 244–14 253.
[11] S. Wan, L. Qi, X. Xu, C. Tong, and Z. Gu, "Deep learning models for real-time human activity recognition with smartphones," Mobile Networks and Applications, vol. 25, no. 2, pp. 743–755, 2020.
[12] F. J. Rodriguez Lera, F. Martin Rico, A. M. Guerrero Higueras, and V. M. Olivera, "A context-awareness model for activity recognition in robot-assisted scenarios," Expert Systems, vol. 37, no. 2, p. e12481, 2020.
[13] Puneeth, Sowmya, S. Ziyan, and M. R. Manu, "Human activity recognition using machine learning," International Journal of Research in Engineering, Science and Management, vol. 4, no. 7, p. 253–255, Jul. 2021. [Online]. Available: http://journals.resaim.com/ijresm/article/view/1051
[14] S. K. Polu and S. Polu, "Human activity recognition on smartphones using machine learning algorithms," International Journal for Innovative Research in Science & Technology, vol. 5, no. 6, pp. 31–37, 2018.
[15] A. Subasi, K. Khateeb, T. Brahimi, and A. Sarirete, "Human activity recognition using machine learning methods in a smart healthcare environment," in Innovation in health informatics. Elsevier, 2020, pp. 123– 144.
[16] D. Singh, E. Merdivan, I. Psychoula, J. Kropf, S. Hanke, M. Geist, and A. Holzinger, "Human activity recognition using recurrent neural networks," in Machine Learning and Knowledge Extraction, A. Holzinger, P. Kieseberg, A. M. Tjoa, and E. Weippl, Eds. Cham: Springer International Publishing, 2017, pp. 267–274.
[17] S. Abbaspour, F. Fotouhi, A. Sedaghatbaf, H. Fotouhi, M. Vahabi, and M. Linden, "A comparative analysis of hybrid deep learning models for human activity recognition," Sensors, vol. 20, no. 19, p. 5707, 2020.
[18] C. I. Orozco, E. Xamena, M. E. Buemi, and J. J. Berlles, "Reconocimiento de acciones humanas en videos usando una red neuronal cnn lstm robusta," Ciencia y tecnolog´ıa, no. 20, pp. 23–36, 2020.
[19] D. Deotale, M. Verma et al., "Human activity recognition in untrimmed video using deep learning for sports domain," 2020.
[20] F. Serpush and M. Rezaei, "Complex human action recognition in live videos using hybrid fr-dl method," arXiv preprint arXiv:2007.02811, 2020.
[21] K. Xia, J. Huang, and H. Wang, "Lstm-cnn architecture for human activity recognition," IEEE Access, vol. 8, pp. 56 855–56 866, 2020

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)