IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN2347–3657

Volume13, Issue 1,2025

A NOVEL APPROACH TO DETECTING ABNORMAL EVENTS IN FOOTPATHS VIA DEEP LEARNING

[1]GULLAPELLY NIKHITHA, [2]SABER HUSSAIN,[3]ALETI SAI ASHWITHA REDDY, [4]K BADHANAKURTHI RAKESH, [5]BALEM DEEPIKA REDDY, [6]Mr. G LACHIRAM, [7]Mr. PVV SRINIVASARAO,

[12345]Student Department of DS, Narsimha Reddy Engineering College, Maisammaguda (V), Kompally, Secunderabad, Telangana-500100.

[6]Assistant Professor, Department of CSE, Narsimha Reddy Engineering College, Maisammaguda (V), Kompally, Secunderabad, Telangana-500100.

[7]Assistant Professor, Department of Mechanical Engineering, Narsimha Reddy Engineering College, Maisammaguda (V), Kompally, Secunderabad, Telangana-500100.

Abstract—

Because it may pick up on any form of odd movement, a video surveillance system is an important part of the surveillance industry. We have utilized and analyzed the patterns of the unexpected or abnormal motions captured by the security cameras. Afterwards, the patterns that don't fit the norm are employed for capture. Bikers, skateboarders, little carts on pedestrian walkways, etc. are just a few of the ways they appear in film sequences. The primary objective of this research is to design and implement a Video Anomaly Detection (VAD) system that can effectively analyze films for anomalous movements by using deep learning and image processing methods.

Keywords: video surveillance, image processing, VAD,deep learning, patterns

INTRODUCTION

Intelligent video surveillance, a complex area of research in computer vision and machine learning systems, has lately gained increased interest due to concerns

related to the topic of global safety. Be on the lookout for suspicious behavior in crowded public places, such as shopping malls, train stations, and bus stops. More and more public and private spaces are installing surveillance cameras as a result of technology developments and falling prices. People are usually the ones that keep an eye out for bad things happening. Multiple cameras shot footage at the same time, which they analyze. Being exposed to long durations without visual stimulation impairs their ability to detect unusual events as they happen. This renders the current setup mostly ineffective outside of forensics as a recording device. Consequently, it is critical to have automated anomaly detection in place that can detect and respond to issues in real-time. As a result, we're aiming to build an anomaly system that can spot strange movement in videos. Here, the input picture dataset is analyzed using a combination of image processing methods and deep learning techniques, including recurrent neural networks and convolutional neural networks.

LITEARTURE SURVEY

When anything happens that doesn't fit the usual pattern, we call it an anomaly. The primary component of surveillance is the detection of anomalies. That being said, it is often discovered to be a difficult undertaking due to the possibility of misleading anomaly detection in certain regions. For instance, whereas shooting a pistol is considered unusual in everyday life, it is considered par for the course at a shooting range. As a result, certain occurrences are seen strange even when they conform to local norms. This research reveals a relationship between actual criminal data and community data, which was utilized to investigate violent crimes using the WAKA dataset (Waikato Environment for Knowledge research). This research used three different regression methods: additive, linear, and decision stump. Out of these three, linear regression demonstrated the power of deep learning approaches for anomaly identification by determining the event's unpredictability. the third [4] A pattern has been detected in the analysis of anomaly detection in Philadelphia, according to a research by Kim S et al.

ISSN2347–3657

Volume13, Issue 1,2025

Logistic regression, ordinal regression, decision trees, and k-nearest neighbor are some of the machine learning approaches used to train the model for anomaly detection from vast data sources. The models' accuracy rate is 69%. [5] Elharrouss et al. [6] utilized a dataset of past crime scenes to make predictions about where crimes are likely to occur. In order to make sense of the data, the levenerg-marquardt method was used. Along with this, the research also made use of the scaled approach to analyze and comprehend the data. With an accuracy of 78%, the scaled approach was determined to be the top performer. Asan added bonus, it may reduce crime by as much as 78%. [7] [8]

With data integrated into a 200x250 m grid, Sultani et al. performed a comprehensive analysis on anomaly detection in urban regions. With retrospect, it was clearly evaluated. Anomaly detection methods like neural networks and ensemble logistic regression were part of the approach they suggested. The findings indicate that anomaly prediction is more precise when executed every fourteen days as opposed to doing it monthly. [9]

Using anomaly data from the fifteen years immediately before 2017, Rummens et al. conducted a comprehensive analysis of anomalous activity. The detection has been carried out using methods like decision trees and k-nearest neighbors. While tested on a dataset with 5,60,000 anomalous actions, it achieved an accuracy level of 39-44%. [10] the eleventh

METHODOLOGY

A kind of machine learning known as deep learning incorporates hidden layers into its architecture alongside the more traditional input and output layers. It is possible for deep learning to reproduce

any autonomous human activity. The greater the number of hidden layers, the more accurate the model becomes. It is possible to utilize a deep learning model with only one layer, but the results will be subpar. [12]
Here, the model was built using RNN and CNN, two deep learning approaches. The model receives a large quantity of images as input. Several security cameras in various locations captured these images. By using various image processing methods, the collection's images are examined. The model's structure allows it to evaluate and identify patterns. In the first step, the picture dataset is divided into two parts: the train set and the test set. With the help of the train data, the model is trained, and then its performance is evaluated with the help of the test data. In order to construct these models, we have made use of deep learning methods like convolutional neural networks (CNNs). We use the algorithm with the best accuracy in our model construction process after we personally evaluate each method for accuracy, performance, etc. [13] [14]

SYSTEM ARCHITECTURE

The system architecture is presented in fig.1.

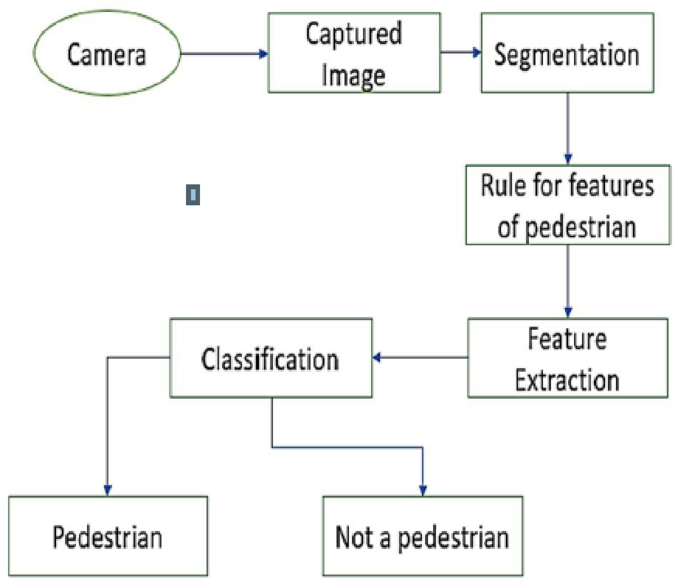


Fig.1. System architecture

Images, sounds, and videos are only few of the formats used for the input data. In this case, we eliminated duplicates using preprocessed pedestrian photo data. Consequently, the

The method of feature extraction is used. This stage involves retrieving the relevant attributes from the image. Next, the Feature Selection is executed. In this step, we use the features acquired in the previous step to choose the most relevant attributes. Now you may divide the dataset into two parts: the test set and the train set. After that, we tweak the Deep Learning model so it can predict whether

International Journal of Information Technology & Computer Engineering
ISSN2347–3657
Volume13, Issue 1,2025

the input image has an abnormal object in it. [15] [16]

IMPLEMENTATION

A. DATA COLLECTION

In this approach, we utilize films to identify anomalies in pedestrian pathways and produce a better, safer travel for pedestrians. Data may be of numerous sorts.

This task makes use of a dataset developed at UC San Diego known as UCSD anomaly detection. The dataset was generated by installing cameras along pedestrian walkways and recording footage of these pathways. Included in the dataset are two folders: ped1 and ped2. Both ped1 and ped2 include human walkers in various settings; ped1 has 34 training video samples, 36 testing video samples, 16 training films, and 12 testing films, respectively. Figures 2 and 3 display excerpts from the video dataset. [17]

Fig.2. Captured image

Fig. 3. Captured image

DATA PREPROCESSING

We can use optical flow and object identification to spot the out-of-the-ordinary things according to their motion, which includes how fast they're moving and how they seem. You may use these techniques

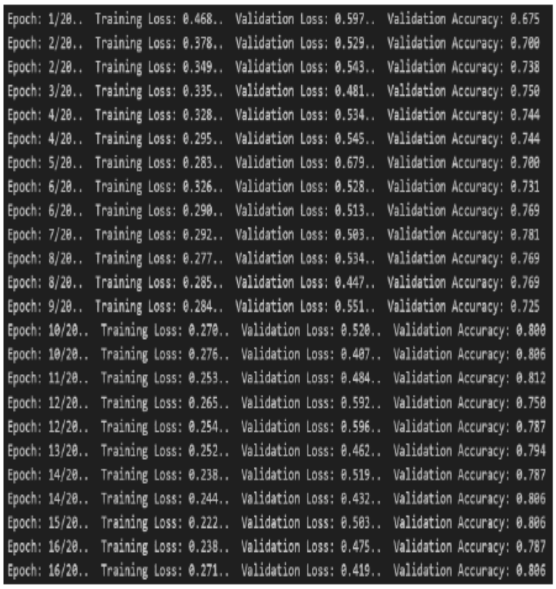
to the video frames included in the Object Detection dataset is: To recognize the item in the video frame, object detection is essential. It lets us identify things from the frames of the video, such as people or anomalies like motorcycles, automobiles, trucks, skaters, etc. The Flow of Light: Optical flow compares successive frames to ascertain the mobility of the video's objects. Using the brightness constancy assumption, which is dependent on the pixel's brightness, it may detect anomalies, or items moving quickly, in the pedestrian track. [18] The year 19 Normalization of Batches
To make the model's learning process faster, we use the batch normalization method in this project. If this layer works as promised, it should make our model easier to comprehend and the input image training process go by faster.
With this model's batch size of 32, a total of 32 picture samples are sent across the network simultaneously. We use a batch normalization layer to expedite the network's learning process using the training data. In [20],
C. Instruction and Evaluation
Different camera positions are used to create the "ped1" and "ped2" folders inside the dataset, which are then used to train and test the model. The number of video samples used for training and testing in Pred1 is 34 and in Pred2 it is 36 and 12, respectively. Every video clip is reduced to 200 frames so that the models may be trained. Pred1 and pred2 both feature binary flags that distinguish between normal and abnormal. You may also use the given masks to assess the stats. When choosing the sets, we take into account their sizes; for optimal model performance, the training data should be larger than the testing data. [21]
Predicting output is a necessary capability for any model that uses input data for training. Once the model is ready, training may commence, during which the Epoch method is used to practice the extracted pedestrian characteristics from the input picture. With respect to the training set, a "epoch" is the number of cycles that should occur. Figure 4 shows that the model was trained using 20 epochs, or iterations, of the training data. To rephrase, we used deep learning methods to train the data 20 times. Every time, we check the accuracy. The

International Journal of
Information Technology & Computer Engineering

ISSN2347–3657

Volume13, Issue 1,2025

trained model may then be tested. In this scenario, the concealed image has to be loaded in order to make a prediction. [22] is a



```
Epoch: 1/20..  Training Loss: 0.468..  Validation Loss: 0.597..  Validation Accuracy: 0.675
Epoch: 2/20..  Training Loss: 0.378..  Validation Loss: 0.529..  Validation Accuracy: 0.700
Epoch: 2/20..  Training Loss: 0.349..  Validation Loss: 0.543..  Validation Accuracy: 0.738
Epoch: 3/20..  Training Loss: 0.335..  Validation Loss: 0.481..  Validation Accuracy: 0.750
Epoch: 4/20..  Training Loss: 0.328..  Validation Loss: 0.534..  Validation Accuracy: 0.744
Epoch: 4/20..  Training Loss: 0.295..  Validation Loss: 0.545..  Validation Accuracy: 0.744
Epoch: 5/20..  Training Loss: 0.283..  Validation Loss: 0.679..  Validation Accuracy: 0.700
Epoch: 6/20..  Training Loss: 0.326..  Validation Loss: 0.528..  Validation Accuracy: 0.731
Epoch: 6/20..  Training Loss: 0.290..  Validation Loss: 0.513..  Validation Accuracy: 0.769
Epoch: 7/20..  Training Loss: 0.292..  Validation Loss: 0.503..  Validation Accuracy: 0.781
Epoch: 8/20..  Training Loss: 0.277..  Validation Loss: 0.534..  Validation Accuracy: 0.769
Epoch: 8/20..  Training Loss: 0.285..  Validation Loss: 0.447..  Validation Accuracy: 0.769
Epoch: 9/20..  Training Loss: 0.284..  Validation Loss: 0.551..  Validation Accuracy: 0.725
Epoch: 10/20..  Training Loss: 0.270..  Validation Loss: 0.520..  Validation Accuracy: 0.800
Epoch: 10/20..  Training Loss: 0.276..  Validation Loss: 0.407..  Validation Accuracy: 0.806
Epoch: 11/20..  Training Loss: 0.253..  Validation Loss: 0.484..  Validation Accuracy: 0.812
Epoch: 12/20..  Training Loss: 0.265..  Validation Loss: 0.592..  Validation Accuracy: 0.750
Epoch: 12/20..  Training Loss: 0.254..  Validation Loss: 0.596..  Validation Accuracy: 0.787
Epoch: 13/20..  Training Loss: 0.252..  Validation Loss: 0.462..  Validation Accuracy: 0.794
Epoch: 14/20..  Training Loss: 0.238..  Validation Loss: 0.519..  Validation Accuracy: 0.787
Epoch: 14/20..  Training Loss: 0.244..  Validation Loss: 0.432..  Validation Accuracy: 0.806
Epoch: 15/20..  Training Loss: 0.222..  Validation Loss: 0.503..  Validation Accuracy: 0.806
Epoch: 16/20..  Training Loss: 0.238..  Validation Loss: 0.475..  Validation Accuracy: 0.787
Epoch: 16/20..  Training Loss: 0.271..  Validation Loss: 0.419..  Validation Accuracy: 0.806
```

Fig.4. Epoch

Figures 5 and 6 show the CNN's accuracy and loss graphs, respectively. As part of this research, we train the model for 20 epochs.

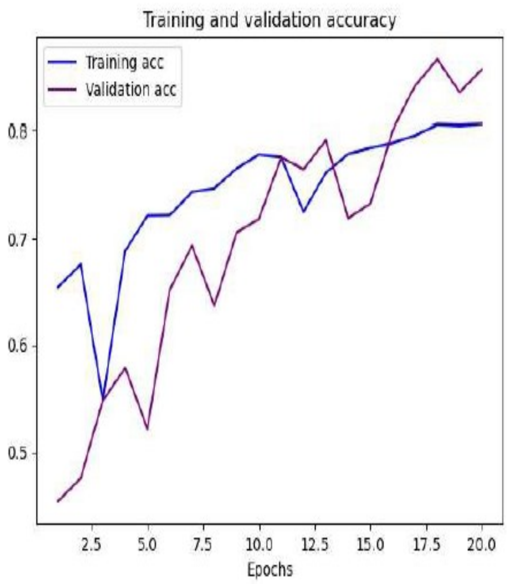


Fig.5. Accuracy

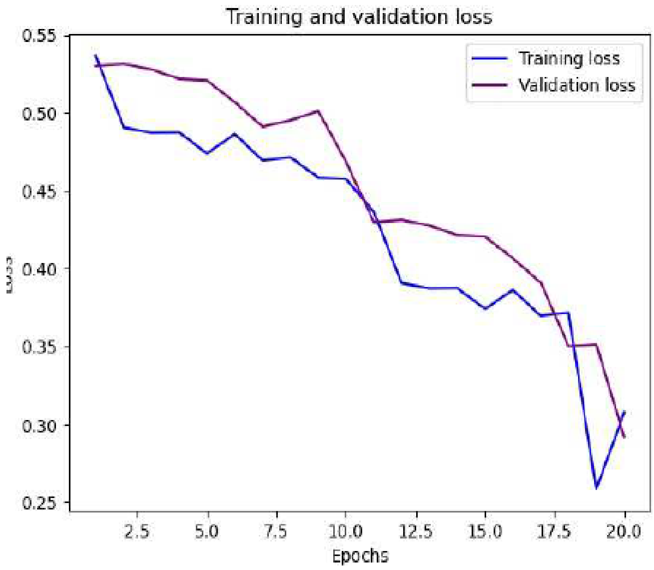


Fig.6. Loss

MODEL SELECTION

The many-to-many LSTM and convolution neural networks are two types of recurrent neural networks. This is where deep learning techniques come into play.

The neural network's backpropagation aids in model building, and the learning rates may be used for improved model training, all because of the network's capacity to extract features, train them with several layers, and achieve greater performance.
Deep Neural Network (CNN)
Our approach to deep learning here is similar to CNNs. It is not an easy undertaking to construct a CNN pedestrian anomaly detection model. In order to construct deep learning models more effectively, convolutional neural networks (CNNs) include the following layers. Authorized
KERNAL
By using the Kernal technique, the camera's input photos may have their information or attributes retrieved and stored as a matrix of pixels. The 3x3 matrix used in this project is created from the picture's input pixel matrix and compared to the kernel matrix, which is likewise 3x3. This indicates that the kernel size is 3x3. A feature extraction matrix is the result of multiplying the two matrices. Since we used to expand the model, the kernel size is 3x3. [23] The
SLATTER COAT
By including a flattening layer after the convolution layer, the project's multi-dimensional array of image attributes may be condensed to a single dimension. To find out whether the given

International Journal of Information Technology & Computer Engineering

ISSN2347–3657

Volume13, Issue 1,2025

image is malignant or not, the fully connected network takes this flattened array of picture properties as an input layer. Consequently, the network's overall performance and processing speed would be enhanced by adding this layer after the convolution layer and lowering the size of the features map. [24]

The Sigmoid function, however, is used to ascertain whether the picture constitutes an anomaly. The sigmoid function takes values between zero and one. In order for the network to learn complex models, this sigmoid function makes it non-linear. And therefore, the rectified Linear unit function and the Sigmoid function are both used in this project. This model auto-refines the neural network's parameters using the Adaptive Moment Estimator (Adam), which estimates the values from the network's previous layer of neurons. Both the learning rate and the accuracy of the model are enhanced with the help of this optimizer during sample training. [25] The [26]

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 254, 254, 16)	448
max_pooling2d (MaxPooling2D)	(None, 127, 127, 16)	0
conv2d_1 (Conv2D)	(None, 125, 125, 32)	4640
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 32)	0
conv2d_2 (Conv2D)	(None, 60, 60, 16)	4624
max_pooling2d_2 (MaxPooling2D)	(None, 30, 30, 16)	0
flatten (Flatten)	(None, 14400)	0
dense (Dense)	(None, 256)	3686656
dense_1 (Dense)	(None, 1)	257

Fig.7. Output of CNN

```
for batch in test.as_numpy_iterator():
    X, y = batch
    yhat = model.predict(X)
    pre.update_state(y, yhat)
    re.update_state(y, yhat)
    acc.update_state(y, yhat)
print(pre.result(), re.result(), acc.result())
```

Fig.8. Predicting model

This project's algorithm uses training examples to determine whether the loaded picture of a pedestrian has any abnormalities. We use the one with the best accuracy score.

as the ultimate grading of precision. In this project, the Convolution Neural Network (CNN) achieves an accuracy rate of 80%.

In summary,

We examined the anomaly because we care deeply about people's safety and want to make sure nothing might hurt them. Bikes, cycles, trucks, skatters, and anything else that doesn't belong in a pedestrian walkway are considered oddities. This code uses deep learning and image processing techniques to identify anomalies in the video. It starts by converting the video into frames. Then, the frames are preprocessed to extract useful features from the dataset. Noise is removed from the frames. Using motion detection and object detection, it classifies objects based on their size and how they are moving, excluding pedestrians. The video frames serve as input for this model, which may have its learning rate tweaked for faster learning. For the best possible model performance, the hyperparameters are additionally fine-tuned. When it comes to extracting better features and achieving high accuracy, deep learning approaches like recurrent neural networks and convolution neural networks are the way to go. Consequently, our model is a top-notch performer when it comes to predicting route abnormalities, with an accuracy rate of 80%. By identifying the outliers, we can enhance the safety measures for others using the route. Motion and object detection are the basis for this model's learning process. The model may be improved by adding additional parameters to this. It is possible to improve this model and get very accurate results by adjusting its parameters.

REFERENCES

[1]. "Evaluating the use of public surveillance cameras for crimecontrol and prevention." by Vigne, N.G.L., Lowry, S.S.,Markman, J.A., Dwyer, A.M. in US Department of Justice,Office of Community Oriented Policing Services. UrbanInstitute, Justice Policy Center, Washington, DC (2011)

[2]. "A world with a billion cameras watching you is just aroundthe corner." by Lin, L., Purnell, N. in Wall Str. J. (2019)

[3]. "spatio-temporal adversarial networks for abnormal eventdetection." by Lee, S., Kim, H. G., Ro, Y. M.: STAN inIEEE international conference on acoustics, speech andsignal processing (ICASSP)(2018)

[4]. "Observe locally, infer globally a space-time MRF fordetecting abnormal activities with incremental

updates." ByKim, J., Grauman, K. in IEEE Conference on ComputerVision and Pattern Recognition (2009)

[5]. "A review of video surveillance systems" by O. Elharrouss, N.Almaadeed, S. Al-Maadeed in J. Vis. Commun. ImageRepresent. (2021)
"E2E-VSDL: end-to-end video surveillance-based deeplearning model to detect and prevent criminal activities" byM.Q. Gandapur in Image Vis Comput. (2022)

[6]. "Survey on contemporary remote surveillance systems forpublic safety" by T.D. Raty in IEEE Transactions onSystems, Man and Cybernetics. (2010)

[7]. "Abnormal event detection at 150 FPS in MATLAB." by LuC., Shi, J., Jia, J. in IEEE International Conference onComputer Vision (2013)

[8]. "A unified approach to interpreting model predictions."byLundberg, S.M., Lee, S. In Proceedings of the 31st

[9]. International Conference on Neural Information ProcessingSystems(2017)

[10]. "Robust real-time unusual event detection using multiplefixed-location monitors." by Adam, A., Rivlin, E.,

[11]. Shimshoni, I., Reinitz, D. in IEEE Trans. Pattern Anal.Mach. Intell. (2008)

[12]. "Video anomaly detection with compact feature sets foronline performance." by Leyva, R., Sanchez, V., Li, C.-T.in IEEE Trans. Image Process.(2017)

[13]. "Real-world anomaly detection in surveillance videos" by W.Sultani, C. Chen, M. Shah in Proceedings of the IEEEConference on Computer Vision and Pattern Recognition.(2018)

[14]. Ravanbakhsh, M., Nabi, M., Sangineto, E., Marcenaro, L.,Regazzoni, C., Sebe, N.: Abnormal event detection invideos using generative adversarial nets. In: 2017 IEEEinternational conference on image processing (ICIP), pp.1577–1581 (2017)

[15]. "Crime analysis through machine learning" by S. Kim, P.Joshi, P.S. Kalsi, P. Taheri in IEEE 9th Annual InformaationTechnology, Electronics and Mobile CommunicationConference(IEMCON),IEEE (2018)

[16]. "Decision-theoretic saliency: computational principles,biological plausibility, and implications forneurophysiology and psychophysics." by Gao, D.,Vasconcelos, N. in Neural Comput.(2009)

[17]. "A causal framework to determine the effectiveness ofdynamic quarantine policy to mitigate COVID-19." ByKristjanpoller, W., Michell, K., Minutolo, M.C. in Appl.SoftComput. (2021)

[18]. "Hierarchical context modeling for video event recognition."by Wang, X., Ji, Q. in IEEE Trans. Pattern Anal. Mach.Intell.(2017)

[19]. "Detecting anomalies in image classification by means ofsemantic relationships." by Pasini, A., Baralis, E. in IEEESecond International Conference on Artificial Intelligenceand Knowledge Engineering (AIKE), Sardinia, Italy.(2019)

[20]. "The use of predictive analysis in spatiotemporal crimeforecasting: building and testing a model in an urban

[21]. context" by A. Rummens, W. Hardyns, L. Pauwels in Appl.Geogr. (2017)

[22]. "Segmentation of COVID-19 pneumonia lesions: a deeplearning approach." by Ghomi, Z., Mirshahi, R., Bagheri,A.K., Fattahpour, A., Mohammadiun, S., Gharahbagh,A.A., Djavadifar, A., Arabalibeik, H., Sadiq, R., Hewage,K. in Med. J. Islam. Repub. Iran(2020)