ISTANBUL COMMERCE UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

EVENT OR EMERGENCY CASE DETECTION BY HUMAN RUNNING

Mohamed Artan ABDI

Supervisor

Assist. Prof. Dr. Metin TURAN

MASTER'S THESIS
DEPARTMENT OF COMPUTER ENGINEERING
ISTANBUL – 2021

ACEPTANCE AND APPROVAL PAGE

On 25/02/2021 Mohamed Artan Abdi successfully defended the thesis, entitled "Event or Emergency Case Detection by Human Running" which he prepared after fulfilling the requirements specified in the associated legislation, before the jury members whose signatures are listed below. This thesis is accepted as a Master's thesis by Istanbul commercial university, Graduate School of Natural and Applied Science.

Approved by:

Supervisor	**Asst. Prof. Dr. Metin TURAN** İstanbul Ticaret University
Jury Member	**Asist. Prof. Dr. Feyza Merve HAFIZOĞLU** İstanbul Ticaret University
Jury Member	**Asist. Prof. Dr. Zeynep TURGUT** Haliç University

Approval date:15.03.2021

Istanbul Commerce University, Graduate School of Natural and Applied Sciences, accordance with the 1st article of the Board of Directors Decision dated 15.03.2021 and numbered 2021/308, "Mohamed Artan Abdı" (TC:99188802858.) who has determined to fulfill the course load and thesis obligation was unanimously decided to graduated.

Prof. Dr. Necip ŞİMŞEK
Head of Graduate School of Natural and Applied Science

DECLARATION OF ACADEMIC AND ETHIC INTEGRITY

I hereby declare that

- All the information in this thesis obtained and presented in accordance with academic and ethical integrity,
- I have fully cited and accurately and clearly referred to each source of inspiration and information, wherever and whenever applicable for all materials and results that are not original to this work,
- Neither whole nor any part of this document is not presented any other place previously

Mohamed Artan ABDI

15.03.2021

TABLE OF CONTENTS

Page

ABSTRACT

M.Sc. Thesis

Event or Emergency Case Detection by Human Running

Mohamed Artan ABDI

Istanbul Commerce University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

Supervisor: Assist. Prof. Dr. Metin TURAN
2021, 40 pages

Sport is the fundamental necessary of social integration; on another hand, security is one of the particular significant services that environmental need. The purpose of this study is to classify the everyday actions of the human, whether sport actions such walking and running as regular movement or running in the case of emergency or event. Classifying the common actions of the human such walking and running, it plays an important role of investigation criminals. Significant progress has been made in computer vision and machine learning recent years. CNN, a deep learning algorithm for image processing was used for the model. The dataset, a thousand of images, of the study were collected from different sites of the internet or extracted from videos. Classify frequent human movements, whether a regular walk or running action, were separated by 86.85% success in the research.

Keywords: CNN, Deep Learning, Human Activities, Investigation Criminals, Security System.

ÖZET

Yüksek Lisans Tezi

İnsan Koşusu ile Olay veya Acil Durum Tespiti

Mohamed Artan ABDI

Istanbul Ticaret Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Metin TURAN
2021, 40 sayfa

Spor, sosyal entegrasyon için temel bir gerekliliktir; diğer yandan güvenlik, çevrenin ihtiyaç duyduğu önemli hizmetlerden biridir. Bu çalışmanın amacı, ister spor yürüyüşü / düzenli hareket olarak koşma, ister acil veya olay durumunda koşma olsun, insanın günlük eylemlerini tanımaktır. İnsanın yürüme ve koşma gibi ortak eylemlerini sınıflandırmak, suçluların soruşturulmasında önemli bir rol oynayabilir. Bilgisayarla görme ve makine öğreniminde son yıllarda önemli ilerlemeler kaydedilmiştir. Model için görüntü işlemede derin öğrenme algoritması olan CNN kullanıldı. Araştırmanın veri seti, bin görsel, internetin farklı sitelerinden toplandı veya videolardan çıkarıldı. Düzenli bir yürüyüş veya koşma hareketi olsun, sık insan hareketlerini sınıflandıran araştırmada% 86.85 başarı ile ayrıldı.

Anahtar Kelimeler: CNN, Derin Öğrenme, Güvenlik Sistemi, İnsan Faaliyetleri, Suçluları Soruşturma**.**

ACKNOWLEDGEMENT

I would like to express my special thanks of gratitude to my teacher (Assist. Prof. Dr. Metin TURAN) who gave me the golden opportunity to do this wonderful project on the topic (Event or Emergency Case Detection by Human Running), which also helped me in doing a lot of research and I came to know about so many new things I am thankful. Secondly, I would also like to thank my parents and friends who helped me a lot in finalizing this project within a limited time.

Mohamed Artan ABDI
Istanbul, 2021

LIST OF FIGURES

Page

LIST OF TABLES

Page

SYMBOLS AND ABBREVIATIONS LIST

AI	Artificial Intelligent
CNNs	Convolutional Neural Networks
DL	Deep learning
DNNs	Deep Neural Networks
FFNNs	Feed Forward Networks
HL	Hidden Layer
ML	Machine learning
RNN	Recurrent Neural Network
SVM	Support Vector Machine

1. INTRODUCTION

Serving the environment is the only reason, most researchers are involved to produce useful things and services that are a benefit to the environment. Sport is a fundamental necessity of social integration on the other hand security is one of the particular significant services that environmental needs. Classifying the common actions of the human such as walking and running, can play an important role in investigating criminals (Zurlini et al, 2008).

Participating in social integration through a project that helps to facilitate the prevention of upcoming risks means this system will work to differentiate between walker and runner, which will be followed up by the person who runner was in the form of exercise or in a manner that may be detrimental to the safety of the area (Affonso et al, 2015).

Every single person in the environment is doing his/her activity such as walking, running, jumping, driving, shopping, etc. Every single activity of these no matter the reason for those actions, they are using by their own body to complete the activity (Martinez et al, 2017). The human brain can predict what this person does, we can guess with a confidence result of this person is running/walking by the emergency case or common sport, but it requires to follow this action from start to end (Wang et al, 2019; Venkatesan et al, 2017).

After technology came up to routing our life and also machines made more benefits to our modern society, a question that came up to researcher's mind. Can a machine recognize human action like the human brain? (Ajuzieogu, 2019).

1.1. Background of Machine learning

Many researchers are coming up with different ideas to build a system that can work such as the human brain. In the early 1950s was discovered the idea of machine learning, two years later, the phrase "Machine Learning", first time, Arthur Samuel came up in 1952. After that moment, it starts pioneering systems using a simple algorithm, day by day it was growing very faster than similar researches because it became a desperately needed approach (Samuel, 1959; Hofmann et al, 2008).

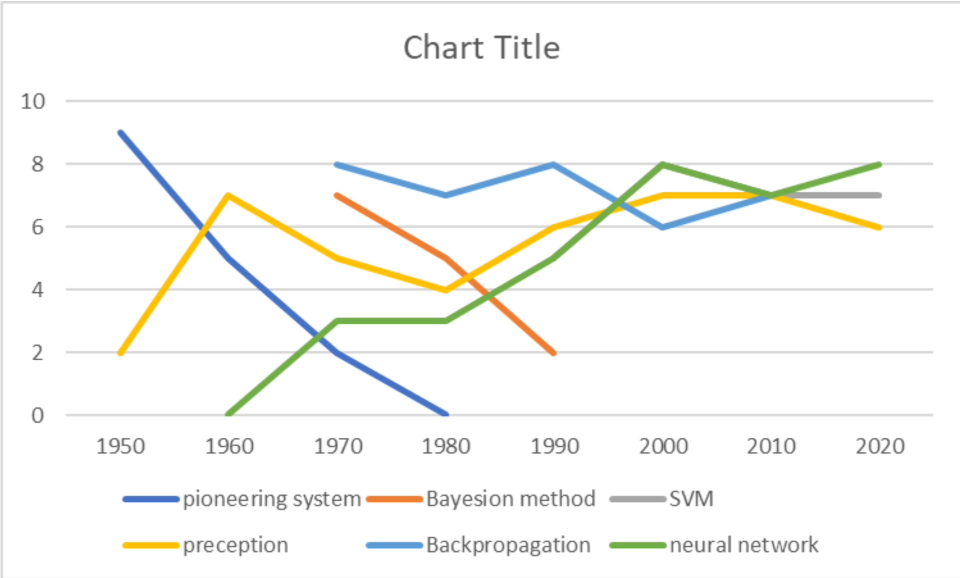


Figure 1.1. Decades of machine learning.

Many decades that through many stages as described in Figure 1.1, for inference probabilistic it was produced Bayesian methods. A while moment, basic pattern recognition started up, such as algorithms for mapping route (Shalev-Shwartz et al, 2014).

1.2. Proposed System

Our interesting system that is planned to separate running and walking activities in order to use in the case of event or emergency case.

Figure 1.2. Proposed system example

In Figure 1.2 as it can be seen that there is a person who is escaping from security inside a market. The case can be recognized by using our model in such stable environments as the street, market, etc.

Figure 1.3. Proposed system example 2

In Figure 1.3 as it has seen the female is running as a common sport, our model classifies this person as running.

1.3. Brief Overview of Tools and Technics

Producing a system that can capture a human movement then can predict what is the action. It is a quite complicated challenge the way we can deal with this new technological approach. It requires a dataset, an appropriate algorithm, and tools to produce such a system (Carrasquilla et al, 2017).

1.3.1. Dataset

The dataset used in this research is various collection of images were collected from the high-level of human activities on a single selected actor in the environment, that can interact with the research. The dataset is composed of two different tags, plural images and singular images that include.

- Walking
- Running
- Other actions

✓ singular image was collected from different environments to face the challenge of every single image and also to make sure it can deal with one by one.

- ✓ plural images are more than two images in one group that had the same action or same person, one action can be capturing from various viewpoints, for example, front side, angles, backside, or even on top.

The dataset was collected from the environmental on various places such as different web sites on the internet and important places of society center, such like meeting places, parks and sport centers. Instead of choosing already prepared datasets.

1.3.2. Appropriate algorithm

Developing an intelligent system, it is required artificial intelligence (AI). AI allows computers to make their own decision without explicitly being informed. Abnormal growing AI in recent years has become popular, because of the extensive amount of data generated each day with the digital transformation (Martinez et al, 2017).

Machine learning (ML) is a child that belongs to the AI, it is related to obtainable knowledge by analysis. Deep learning is a sub-child of ML methods that are shown in Figure 1.4. (Mishra et al 2017) ML may take more time to train and predict the result, but it is appropriate to our problem with a moderate amount of data. (Bini, 2018). Deep learning (DL) is the same function as ML, except it takes a large amount of data than ML and less time of training and predicting, shown in Figure 1.5, It means deep learning is a more appropriate algorithm for our challenge (Yamashita et al, 2018).

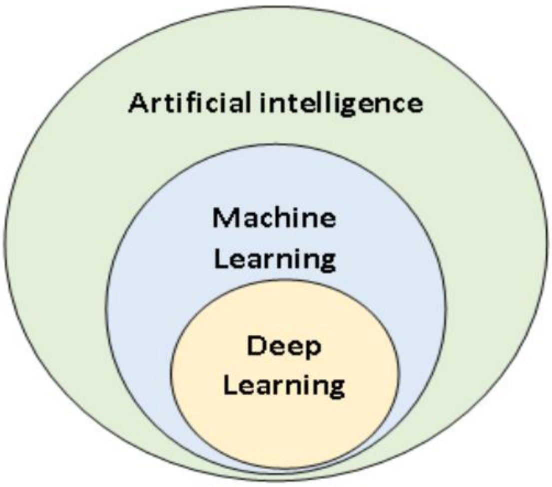


Figure 1.4. Artificial intelligence family

Deep learning is enabling neural networks with multiple layers, to increase the size of the data, also it requires more computation and bigger models, and helpful performance than traditional machine learning algorithms (Lecun et al, 2015).

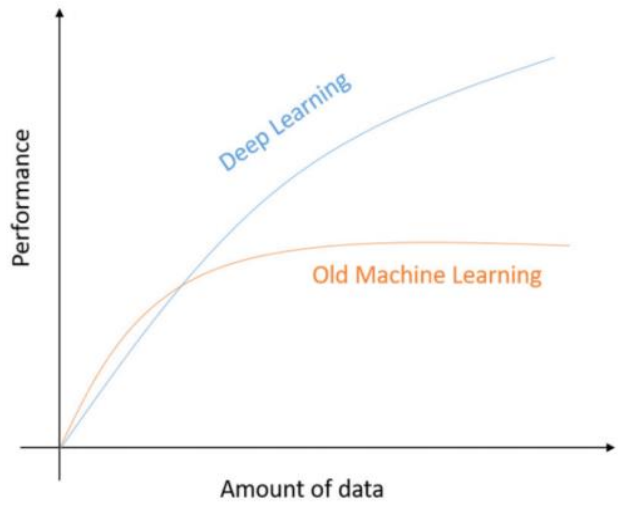


Figure 1.5. Performance comparison of ML and DL Algorithms

Deep learning contains different types of an algorithm that applies for different applications based on the requirement and the performance with different types of data: Deep Neural Network, Convolutional Neural Network, Recurrent Neural Network and more neural network (Rajalingappaa, 2018).

Convolutional neural network (CNN) is a powerful model in deep learning, which applies images to biases, various aspects, and can be able to recognize one from the other, as an input image or video. Nowadays CNN has become more popular that basis a lot of modern machine learning applications (Albawi et al, 2017).

1.3.3. Experimental tools

As above mentioned, CNN is a powerful model. We selected our experimental work because CNN is a feedforward network that flows data from input to one link direct to output as shown in Figure 1.6 (Alsallakh et al, 2018).

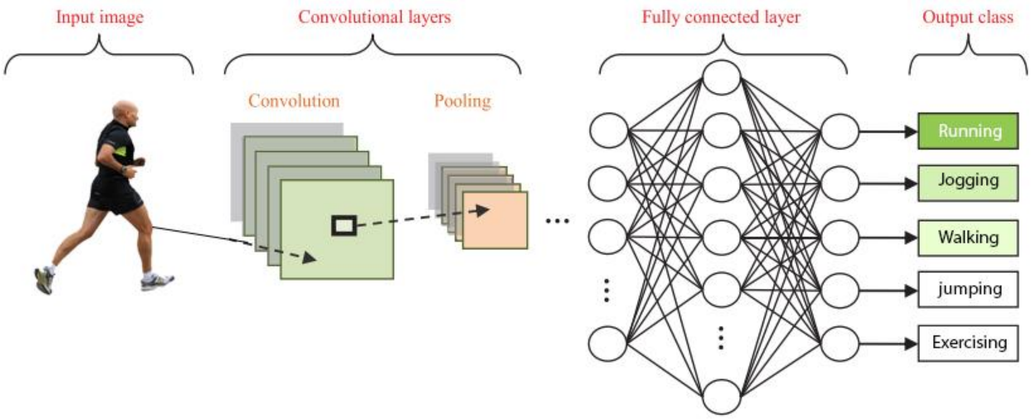


Figure 1.6. Convolutional Neural Network architecture

Selecting and evaluating an appropriate algorithm while comparing it to other algorithms approaches is well done. The reason that selected deep learning more particularly CNN is the ability to especially performance and prediction of the detecting the capture actions with a high-capacity dataset that took from the environment. Besides, it is required to declare CNN libraries. This system is developed using Python libraries such as Keras and TensorFlow.

2. LITERATURE REVIEW

Human movement recognizing and classifying actions challenge has been worked by many researchers before and recent many researchers are working on this challenge. In this section we looked at the similarity of these works and the different technics they applied, we reviewed them separately and compare them against one another evaluating their experimental.

According to the research of Kong and Fu, the interest of CNN in image classification is to study methods of automated labeling of a sequence of motion capture data as a further means of facilitating their reuse in long-term productions (Kong et al, 2018).

The walking/running image categorization technique proposed by Julieta and her colleagues seeks to bring about an improvement in the accuracy of categorizing the image files through the process of combining the information concerning the relationships (Martinez et al, 2017).

Affonso and his colleagues, research shows that the learning of statistical models concerning human movement is a challenging activity due to the increased measurement, non-linear forces at work as well as the stochastic outline of humanoid walking or running (Affonso et al, 2017).

Most researches prioritized focusing on natural activities classification Alsallakh and Jourabloo, El-Fiqi, and her colleagues, and Peng and his colleagues surveyed the recent methods, their studies, and datasets committed on regular common actions.

Furthermore, Lapuschkin and his colleagues state that classification of data entails analysis of individual candidates through the extraction of a sequence of attributes from attempting to equal them to a particular class in a collection of the closed session (a researcher) which is also supported by research undertaken.

Hou and his colleagues have been demonstrated the effectiveness of the proposed tube convolutional neural network for action detection, they evaluate T-CNN on three trimmed video datasets including UCF-Sports (Hou et al, 2017).

Du and his colleagues proposed a new approach that can recognize the action while it is streaming. In reality, it is more complicated work to recognize action during the flow. They recognize the action from multiple capturing videos. It is clearer to understand the action and it’s easier to detect action. This approach is similar to plural action detecting (Du et al, 2017).

Chen and Xue proposed an approach to the spatiotemporal localization (detection) and classification of multiple concurrent actions within temporally untrimmed videos through localizing and scoring actions from color images.

As early mentioned, many researchers had studied on human activity recognition, those approaches were designed to recognize, human body types, there is also previous works design to recognize actions of the human, such as research by D and his colleagues it can recognize actions before it capturing, on other hand Hou and Lee and their colleagues they may make the decisions as well as we did. However, even though they obtained successful results on both recognizing human movements after it happened or while it was happening in life, as far as we know they didn’t make classification of walking and running which is difficult to be categorized because of being similar actions. Our approach is to detect a person if he/she is walking regularly or he/she is running in an emergency case. Recognizing two different actions and categorizing them from the common human movement, is more challenging than recognizing only one action movement.

3. METHODOLOGY

The ultimate point of this research is to develop a technological system that can detecting by emergency or events cases by human running. Developing such technological system, it is required to demonstrate artificial intelligence approaches.

3.1. Artificial Intelligence

Artificial Intelligence is advancing the ability of digital machines, and machines can control by their own self such like robot. Phenomenon of artificial intelligence is affecting all machines and computers of today. Successful digital machines of AI such as, robot controlling himself like human being, high level games which are online or offline, military network systems, self-driving vehicles, delivery network routing in intelligence way, and etc. are become an extraordinary increasing approach.

Many fields of artificial intelligence are used different ways of approaches such like mathematical optimization, search engine optimization, artificial neural network and etc. in this field of AI was founded machines work with human, and thinking like human brain (Ajuzieogu, 2019).

Our ultimate point is to build a system that can work like human brain, our first step we require to select one of these approaches that we mentioned above. Since we are not dealing with searching engine or mathematical optimization, we need to select artificial neural network.

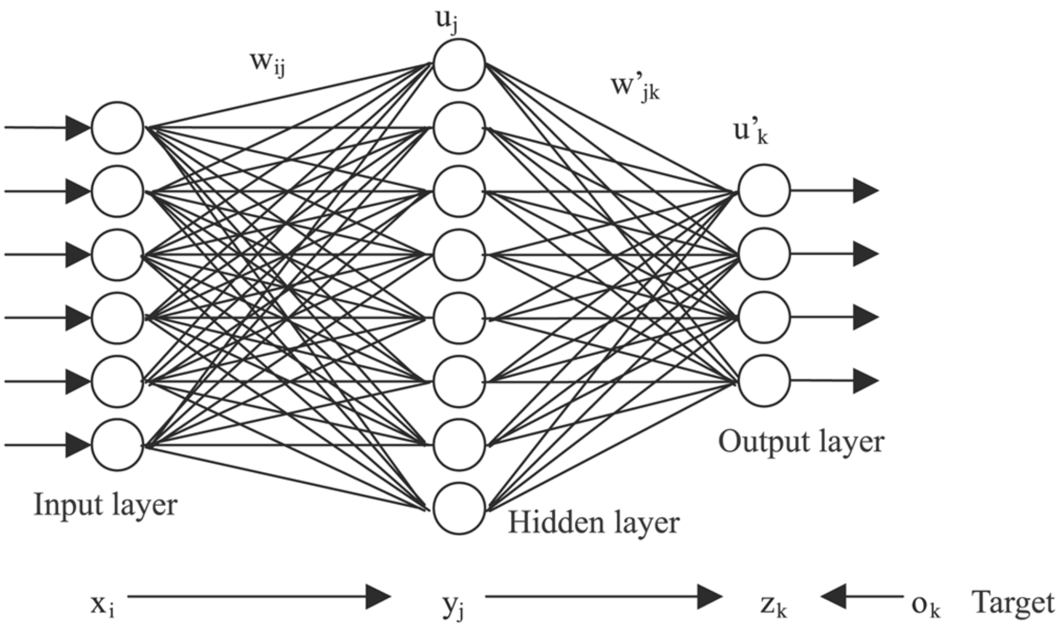


Figure 3.1. Biological way of neural network

Figure 3.1 In biological way of neural networks are number of units collected between each other or collection of nodes. This network is working like synapses structure in a biological brain (Samborska et al, 2014)

Artificial neural networks {ANNs}, is a collecting of nodes relating between each other,

3.2. Machine Learning

Machine learning is subset that belong to AI, it provides ability that can improve experience an automatically learning to the system. In order to take decision of the result, ML makes a prediction that uses a mathematical model build in algorithm called ‘training and testing’ based on sample data (Khan et al, 2006).

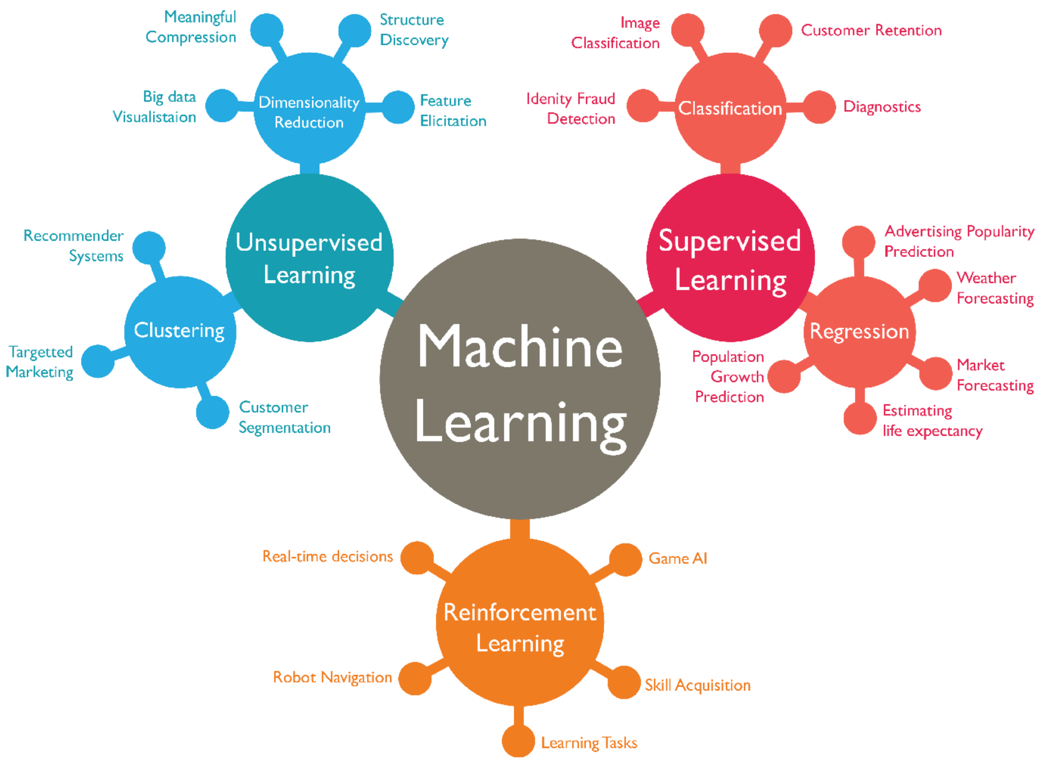


Figure 3.2. Three approaches of ML.

- ✓ Supervised learning Is one of the ML tasks that learn the function of the map from an input to output.
- ✓ Unsupervised learning is one of the ML tasks that looks for previous undetected patterns.
- ✓ Reinforcement learning is works like an agency giving space to work ML. (Abdi, 2016).

3.3. Deep Learning

Deep learning is one of the AI approaches and the mostly powerful approach that nearly work like human brain. DL is sub-child in broader family of ML based on ANN.

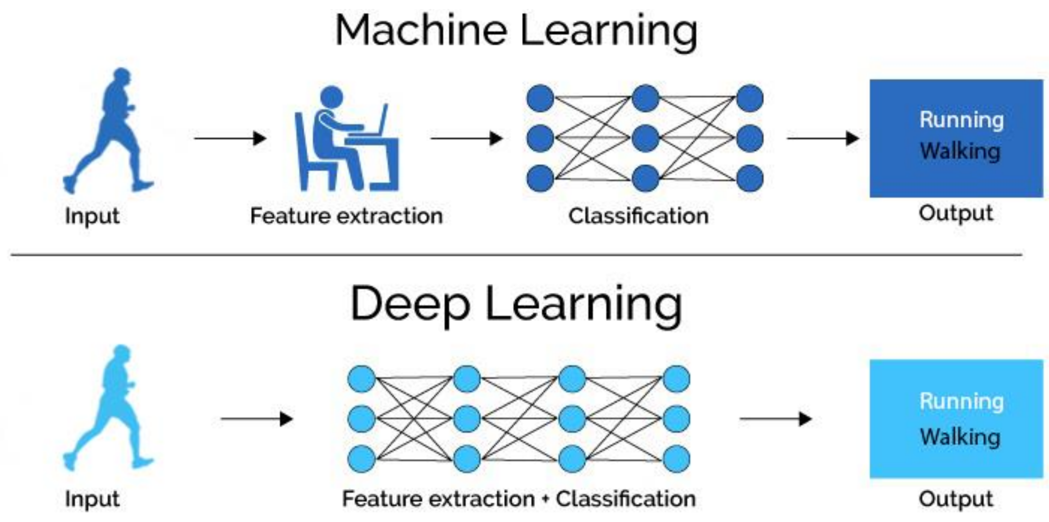


Figure 3.3. Machine Learning vs Deep Learning.

Deep learning is more powerful and useful then machine learning because of DL takes decision by it is own, shown in Figure 3.3. DL uses same learning approaches with ML supervised and unsupervised except is uses semi-supervised (Goodfellow, et al, 2016).

DL have a many different types of architectures that applies for different applications based on the requirement and performance with different types of data these are: Deep Neural Network, Convolutional Neural Network, Recurrent Neural Network and more neural network architectures.

3.4. Convolutional Neural Network (CNN)

CNN was selected, because of it is a specialized form of the neural network developed to a sequence challenges such as text and images.

In recent years CNN has been growing up faster than other approaches on deep learning, it has obtained successful results on classification of text and images (Alwzwazy et al, 2016). CNN is a specialized form of the neural networks developed to sequence challenges shown in Figure 3.4, where old fashion artificial neural network is working like feedforward. On other hand, we are dealing with two thousand frames of dataset, CNN can process large capacity data than traditional ANNs (Alsallakh et al, 2018).

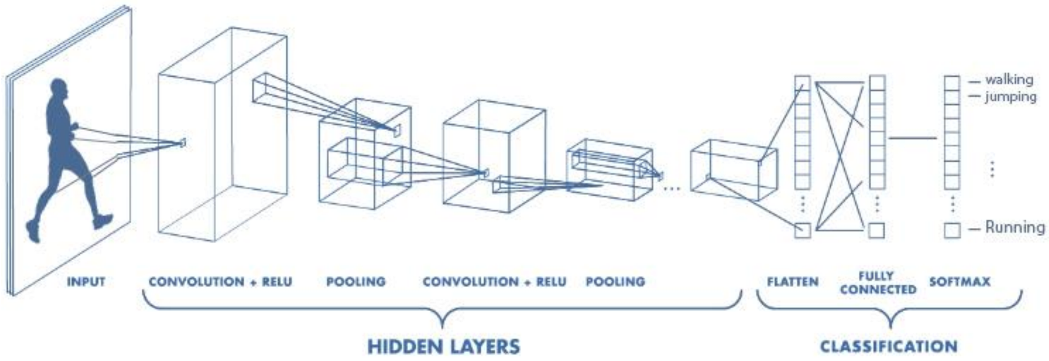


Figure 3.4. Convolutional Neural Network {CNN}

Since the work is to classifying an image from the frame to running or walking categories, CNN is one of the main approaches to develop such challenge. CNN describes at a point when in the course of training a recurrent procedure, the weights are in change and become very limited such that they have no impact towards learning of the data, shown in Figure 3.4. CNN classify such challenge by design, as all propagated information from the system should first pass-through input, processing and output. Such activated functions more so those aims to function on the data so only to permit significant evidence to sustain propagation in the course of training (El-Fiqi et al, 2018).

3.5. Models

The main performances of CNN are declaring the models and how the data is models? as shown in Figure 3.5 the model takes a frame as input data in a fixed size, then passes through the CNN model that consists of five convolutional layers with pooling operations and three fully connected layers (Wang et al, 2016). The main function of models can be summarized as estimating the parameters in a limited number of iterations so that the convergent model can map inputs to outputs (Lecun et al, 2015).

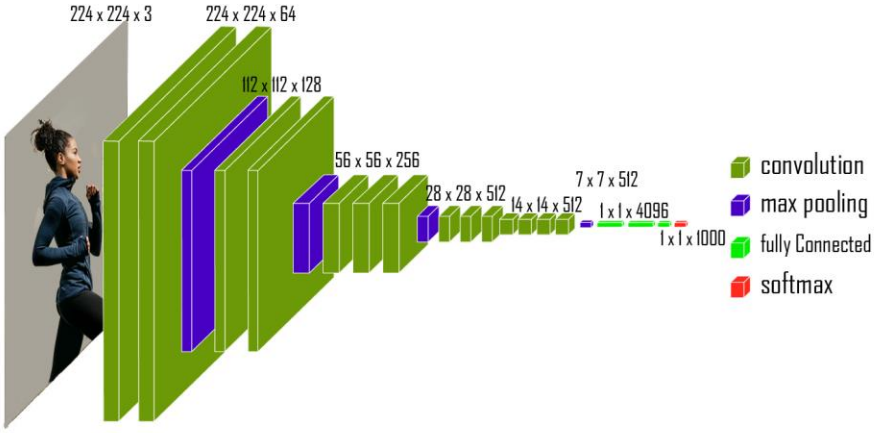


Figure 3.5. CNN Model.

3.6. Data Collection and pre-processing

Data Collection means gathering relative sources from the environment to address critical evaluation. The data used in this study was collected from a broad survey done on the internet and inside the community, the data were collected randomly from free sites (Bengio et al, 2013).

Approximately 30 percent of the dataset, were extracted from videos, it has become an integral part, which is mostly part of plural images. The details of the dataset are given in Table 3.1.

Every single image collected from the internet or extracted from videos is pre-processed in the same way and divided into sub-classes of similar size. The technique has the benefit of classifying individual files in terms of the likelihood of the data.

Table 3.1: Resource details of the dataset

Dataset type	From internet	Extracting from video	Total Dataset
Walking	240	760	1000
Running	260	740	1000

3.7. Image Representation

A single image represents is the most critical process of when the dataset is loading to the system. When the process of training and testing data is starting single image representation can be the most distinguished frames from human movements to walking and running categories, where we assumed as the practical improvement of the human brain shows to be relative to the system.

This project is using scikit-image for image represent, a collection of image processing algorithms implemented in the Python programming language. This library allows image processing to learn algorithms by adjusting and modifying.

3.8. Data Representation

There are major concerns about, how data is represented in a neural network. How the input or training data can be represented and also how the output can be structured in our case is designed such as labels for every class in the research. How feedback and data are represented in every single frame of human movement activity as signified within the Cartesian space. There are primary computational units in an artificial neural network, according to different architectures shown in Figure 3.6 (Craven et al, 1997).

These are three layers of ANN

1. Input layer
2. Hidden layers
3. Output layers

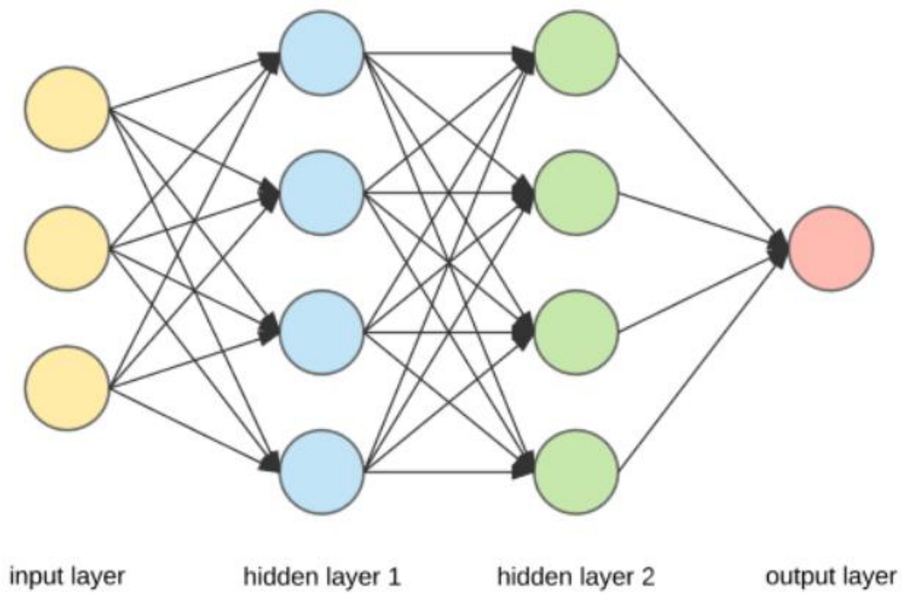


Figure 3.6. Fully Connected Neural Network Architecture

3.8.1. Input layer

Input layer is initial point that start the project, extracting data from an external source and composed input neurons into the system to processed by sequential layers of artificial neurons (Ben-Hur et al, 2001) shown in Figure 3.7.

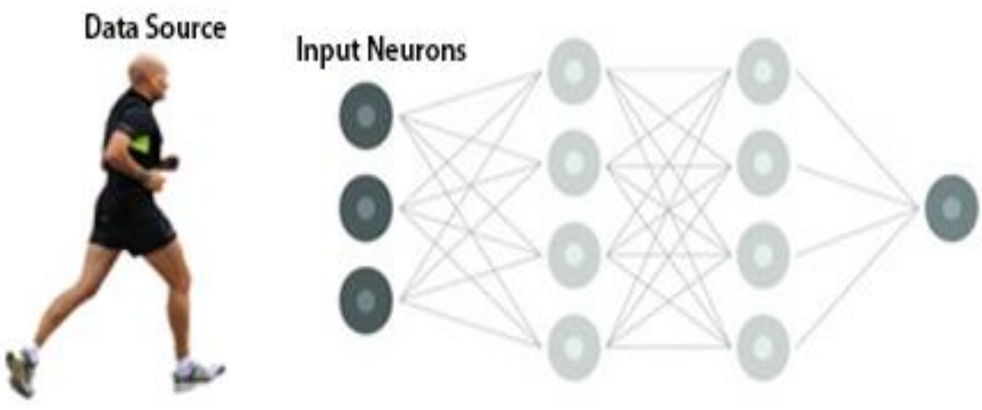


Figure 3.7. Input Data and Neurons

Since the decision has been made in CNN approach (Davis et al, 2006). Therefore, every single image collected from the internet and extracting from videos, are evaluated and divided into models of similar size and every model sent to the network for forecasting. The technique has the benefit of classifying individual files in terms of the likelihood of the data (Zhou, 2020).

3.8.2. Hidden layer

The main important section, that done the most process of the system is inside hidden layer, on other hand, hidden layer is the most critical section that effect the successful of the project (Ciresan et al, 2011).

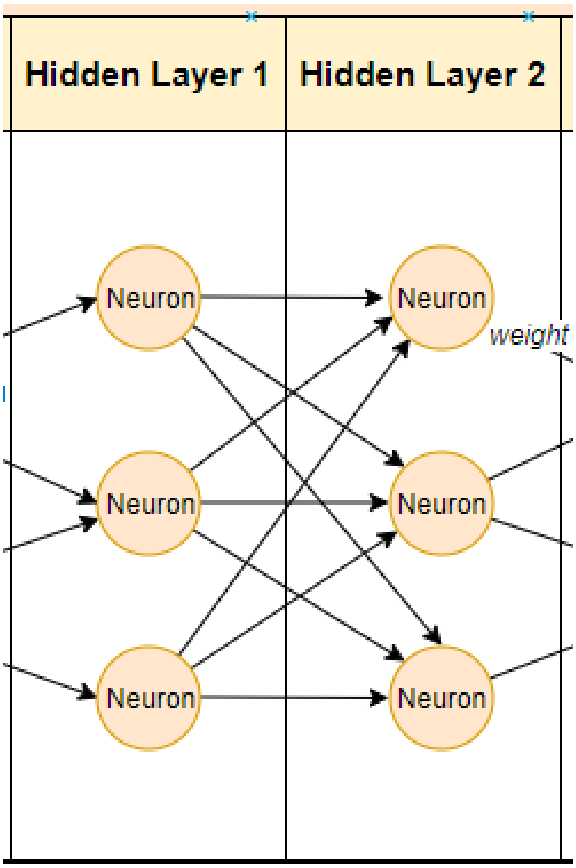


Figure 3.8. Hidden Layers

Hidden-layer (HL) reside in between Input layer and the output layer, in algorithm way. HL are mathematical unites, related each-others that transform data from input until they produce the output as explained Figure 3.8.

The number of neurons that hidden inside the HL are determined by:

Training data samples / Factory * (Input Neurons + Output Neurons) (1)

HL neurons take into consideration of the image where shows a human character start running as he/she prepares to run (Jin et al, 2017). Now imagine how the human brain would understand for either walking or running such that system would be able to realize that the specific file would be a satisfactory response in either instance. Even as the practical improvement of the human brain shows to be relative to the system, it makes confident that at least this is a hint of the model logical (Masoud et al, 2003).

3.8.3. Output layer

The output layer is an important layer, that improves the quality of the system, as early mentioned, we are using python to develop this project. Planning and structuring of the output are used confusion matrix tables and vertical/horizontal axis chart for accuracy rate (Cong et al, 2019).

3.9. Implementation

As earlier mentioned, the decision is made to implement the project in Python with the use of Keras and Tensorflow. The primary reasoning for this selection was owing to the ease availed by these tools, which made it increasingly passable for fast prototyping as shown in Figure 3.9.

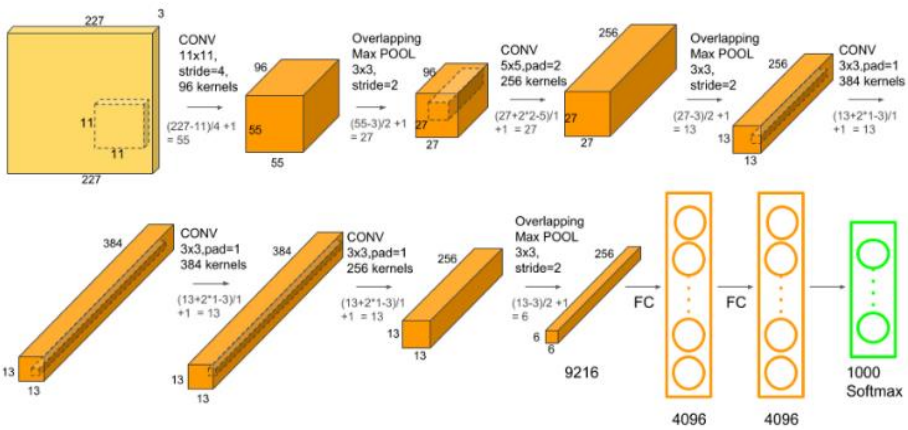


Figure 3.9. CNN parameter structure of implemented model

The answer that is applied for the discussed experimentations relies on three steps; The first step is importing libraries that using this project, we declare Tensorflow libraries from kares, and other requirement libraries to run the project (Chen et al, 2017).

The second step of the project is loading the dataset to the system, and give them batches. the data consists of two types listed in Table 3.2. unique and duplicated images. Around 50 percentages of total data are unique images (different characters), they are different scenes. Approximately 50 percentages of total data are duplicated images, where the same scene is duplicated in various angles and sides.

The proportion of data used in train and test phases are given at (Table 3.2).

Table 3.2. Proportion Details of data for Train and Test

Partition	Training	Testing
Running	80%	20%
Walking	80%	20%

After loading the dataset, the most critical step that effect the successful of the project is training process. The main important section, that improves the quality of the system is training process (Wen et al, 2017).

Once the training process is done, the model send data to tested, with test data. The amount of training and test is the most critical factor in the accuracy rate. Training and test set are divided by 70%–30%. It means that 70% of all the data will be used for training and 30% for the test, according to the dataset structure (Table 3.2).

4. RESULT

Multiple experimental tests are undertaken with the tool providing coverage of varied network architectures as well as numerous subsets of the research. Besides, this is an experimental overload to the most machine learning problem to get a better assessment of the models and architectures over the dataset. A series of images downloaded from the internet was prepared with careful editing such that they appear only a singular achievement for every file.

The convolutional neural network (CNN) is able to learn long-term dependencies amidst time steps of sequenced data. So as to input walking and running images into a neural network, the first step is to load the data into the system.

4.1. First Experimental

The table below is a summarization of the outcomes gained from the prediction of the first training and testing result. It is vital to notice that increasing the number of epochs means increasing the volume of models. In the first experimental it’s been training 10 epochs, it produces a clearly classified result, at the same time, this experimental was increasing the number of epochs up-to 100 epochs to see the capacity of the system and also, to know the number of data it is able to declare one time.

Table 4.1. First experimental result

Epoch	Accuracy	Test-accuracy
10	0.4800	0.4400
50	0.5200	0.4800
100	0.6600	0.6000

The chart below (Figure 4.1) is a comparison of the accuracy, which is to say the vertical axis gained from the first experiment with the use of varied sample sizes along the horizontal axis.

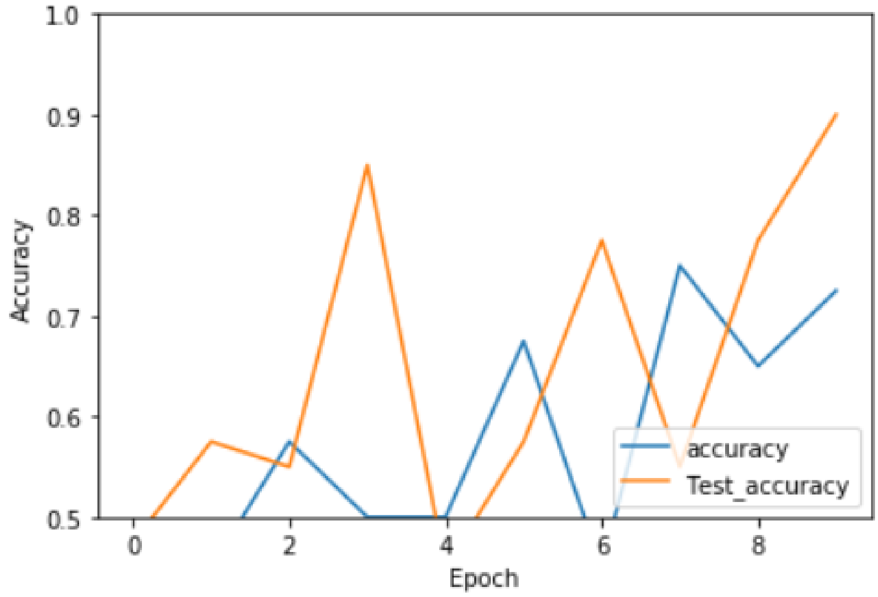
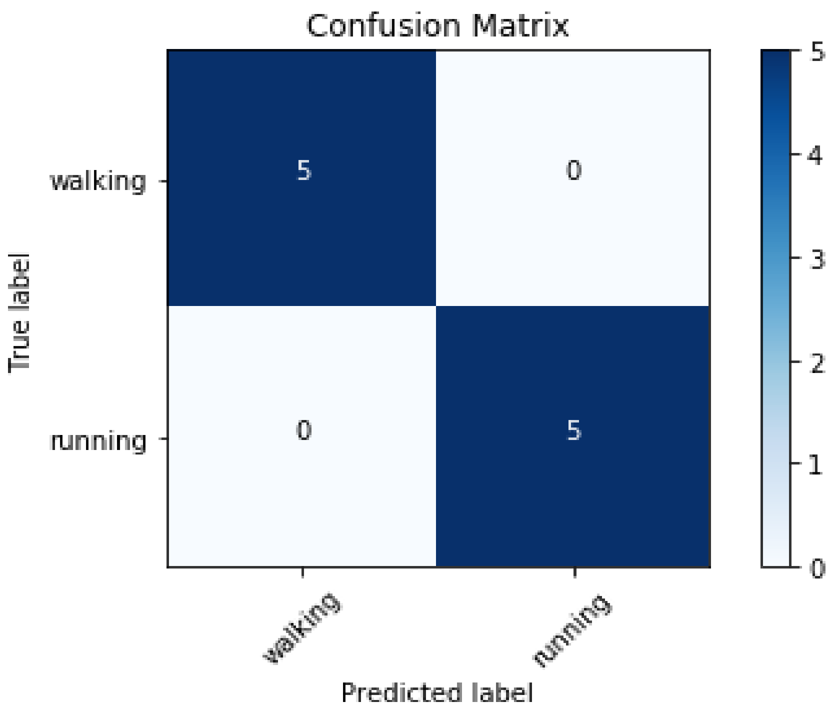


Figure 4.1. Accuracy rate of first experimental

The Table 4.2 is a representation of the confusion matrix gained by taking as an instance the best outcome of first experimental. In the analysis, it is clear that the dimmer areas of the matrix that certain classes such as walking and running have improved classified.

Table 4.2. Confusion matrices for the first experimental.



The accuracy rate that achieved first experimental was very less, it produces un-expected rate. However, next experimental is required large data and increase number of epochs simultaneously.

4.2. Second Experimental

To increase these impacts, the second experimental, undertaken with the consideration of preview prediction result. The outcome that gained in this second experiment of predictions is presented in the Table 4.3.

Table 4.3. Second experimental result.

Epoch	Accuracy	Test-accuracy
200	0.6700	0.6100
400	0.7200	0.5600
600	0.7400	0.5200

Evidence shows that the increased accuracy is different on the first experiment that is attributed to the underfitting characteristic of the model. It also means the larger training dataset of the network experiences improved performance.

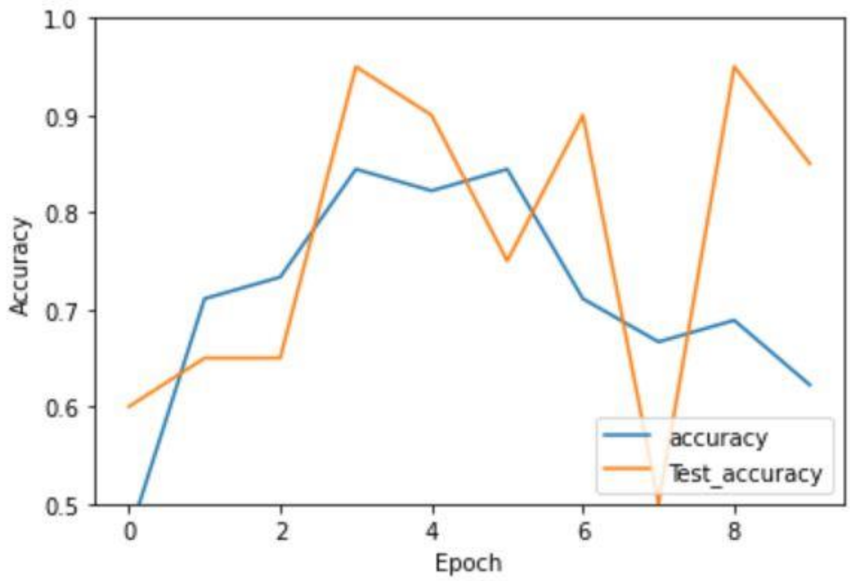
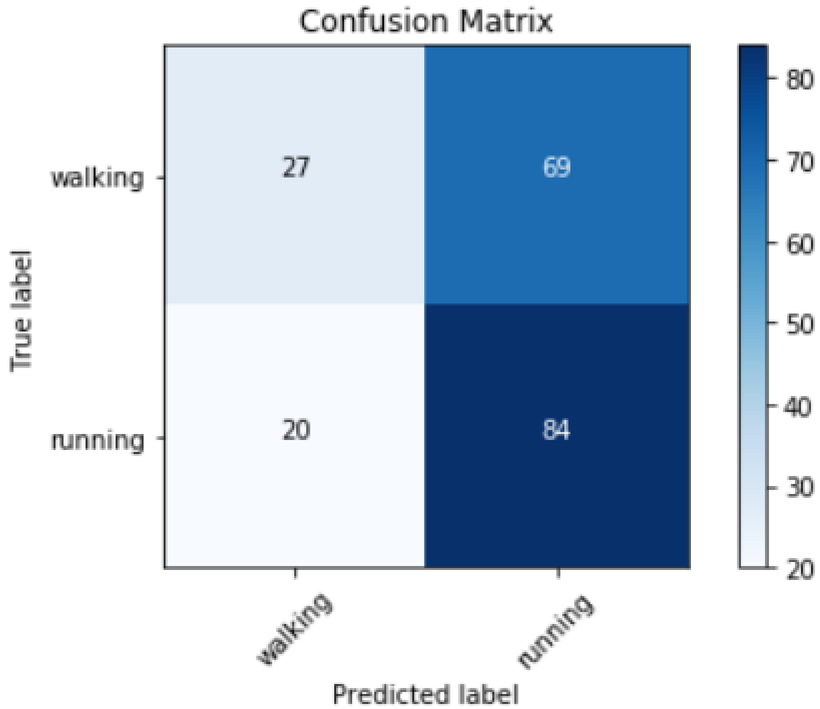


Figure 4.2. Accuracy rate of second experimental

Once more, as we take the best outcome as an instance, the accuracy rate is growing up.

Table 4.4. Confusion matrices for the second experimental



4.3. Third Experimental

Following a careful evaluation experimental. The confusion matrices gained after a combination of the first and second outcomes of two experiment, it is observed that a particular increase of dataset showed a lousy result even in the fact that those were not enough accuracy. Table 4.5 is a present after increase the number of epochs, it seems more relative to previous experimental.

Table 4.5: Third experimental result.

Epoch	Accuracy	Test-accuracy
800	0.8000	0.6400
1000	0.8400	0.6000
1500	0.8600	0.7800

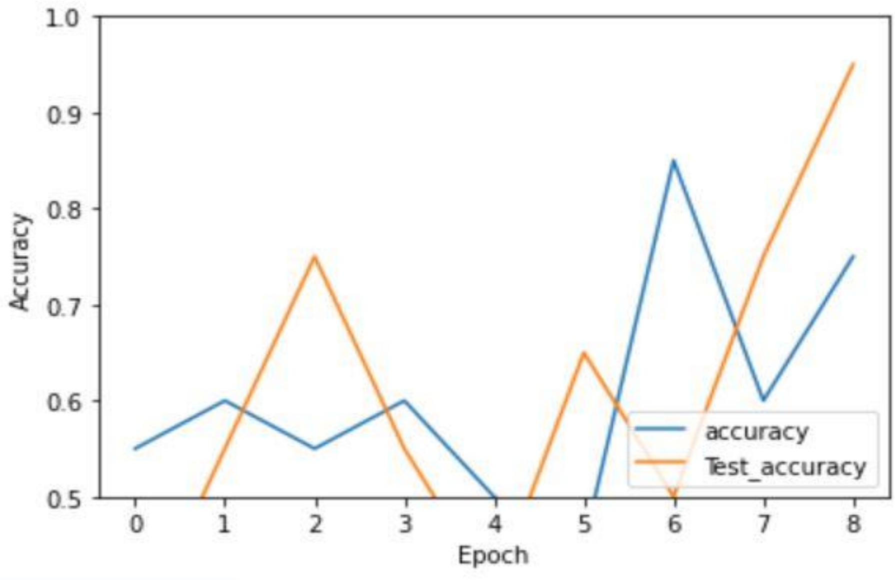
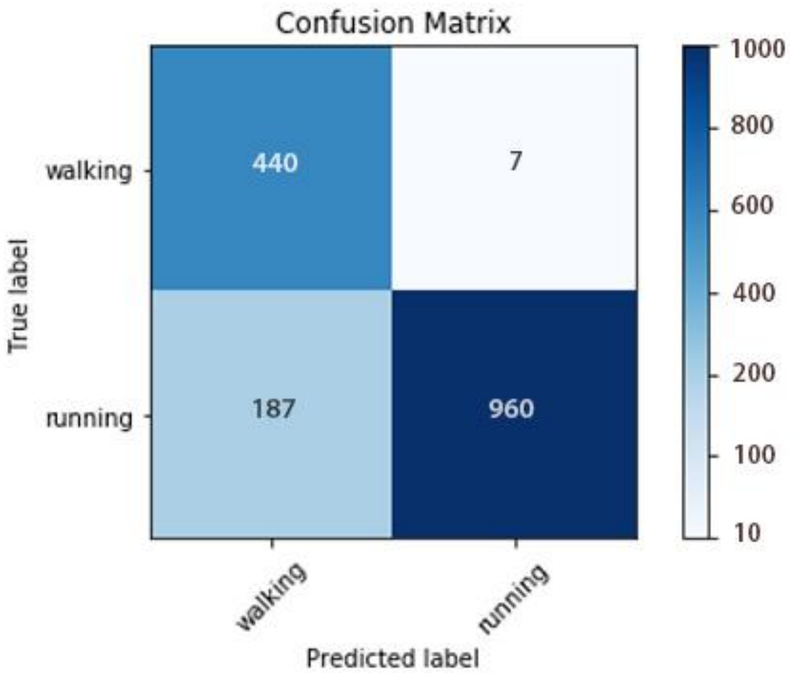


Figure 4.3: Accuracy rate of third experimental

The above chart (Figure 4.3) is a comparison of the accuracy, which is to say the vertical axis gained from the third experimentations, it is more relative to previous comparison.

Table 4.6: Confusion matrices for the third experimental.



4.4. Final Experimental

Comparing the accuracy rate of first, second and third experimentations. First and second experimental shows acceptable update and increasing ration, except third experimental is observed that a particular increase of epoch more than 1000 showed a lousy result.

Worth taking note of is these datasets have considerably distinct training data set the scope, and still, the network was capable of training with satisfaction in them. At this point, we can hypothesize that this is due to the nature of those particular movements that substantially differ from each other, which makes it modest for the network to differentiate them from each other.

Table 4.7. Final experimental result.

Epoch	Accuracy	Test-accuracy
2000	0.8800	0.8000
3000	0.9200	0.8200
4000	0.9600	0.8400
5000	1.000	0.9400

The chart below (Figure 4.4) is a comparison of the accuracy, which is to say the vertical axis gained from the experimentations with the use of varied sample sizes along the horizontal axis. Overall, the results achieved with these experiments are taken into consideration as a promise which indicates the model's viability, as shown in the stacked Convolutional Neural Network that could learn with success how to classify the actions of humans from images downloaded from the internet or extracted from a video.

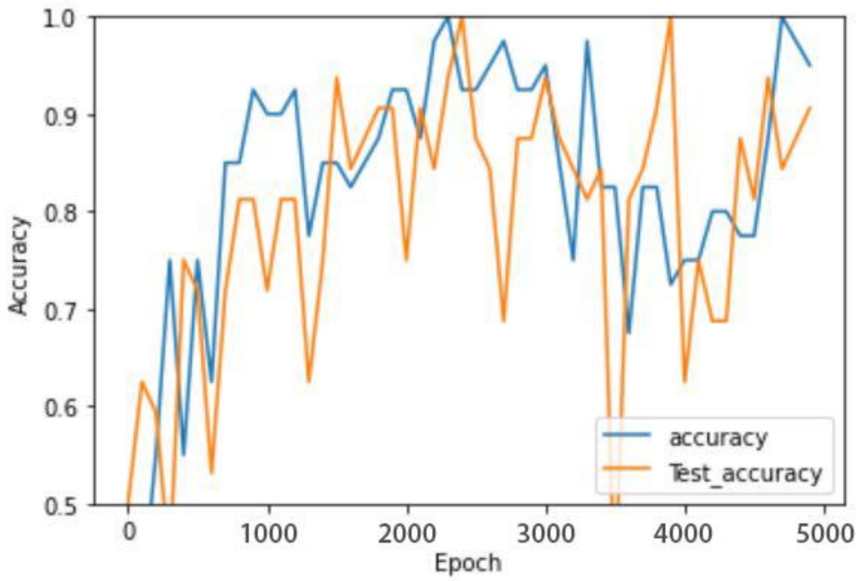
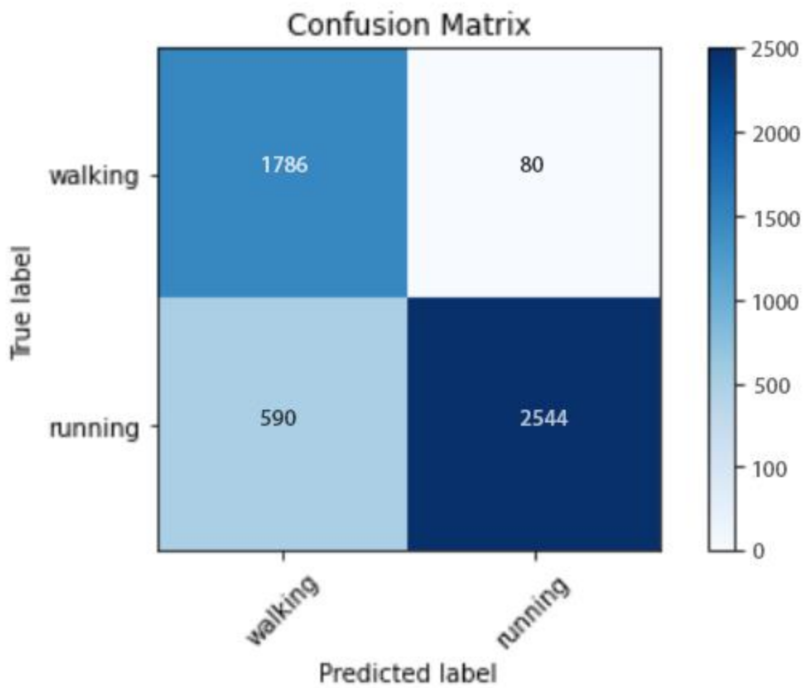


Figure 4.4. Accuracy rate of last experimental

Table 4.8. Confusion matrices for the last experimental.



The Table 4.8 is a representation of the confusion matrix gained by taking as an instance of the best outcome. When the matrix data evaluated by final experimental, labelled correctly to classes walking and running categories.

4.5. Experimental comparison

Before token multiple experimental our first experimental output result was produced a successfully classified result as shown in Figure 4.5. However, the number of epochs were very less, and the accuracy rate it produces was under 50 percent means our success ratio is very low. After that time, we decide to take number of experimental and make them comparison between each other to select the greatest one.

Figure 4.5 Classified result

The significant difference of the experimentalist is range of the success. While all undertaken experimentalists are similar in their general performance, they differ in regard to number of frames that done as well the number of epochs that undertake.

As early describe, first experimental test took 10 to 100 epochs, it produces low range classified result, it gave us confident that at least our system is working. Since our priority focus on the environmental ‘Event or Emergency Case Detection by Human Running’, we decide to increase the capacity of the research by increasing the number of frames that takes single time.

Second experimental, as we promise early, that upcoming experimental will be increasing the number of frames takes in single time as well the number of epochs to get a better accuracy rate. We took whole the dataset, then increase the number of epochs by adding extra 500 epochs which means the number of epochs in second experimental is increasing until 600 epochs.

Third experimental, faced a challenge we are not able to increase more than 1000 epochs, to get out of this predicament, we add more layers to the project, than training process was able to deal with 2500 epoch, now accuracy rate is high. However, we have not yet achieved final accuracy rate.

After that moment, we have passed second and third experimentalist and take one last experimental, as it is our final experimental, we must take the final step, the number of epochs now are doubled the previous number, on other hand single frames are taking by test in the last experimental.

The other important metrics for evaluation the model is precision (1), recall (2) and F1 (3). Evaluation is given in the Table 5 for both classes separately.

Precision = True positive / True positive + False positive (1)

Recall = True positive / True positive + False Negative (2)

F1 = 2 * Precision * Recall / Precision + Recall (3)

4.6. Limitations

At the end of the research, the system holds the limitation of the single-use image files downloaded from the internet, which means that the entire preparation and forecast files have to be of a similar dimension (N) where N is the number of samples. Besides, another limitation concerning the data is the fact that the dataset is a unique selection from the environment, which means they gained outcome is a reflection of that.

However, even as the substantial improvement of the four experiment shows to it is relative to the previous experiment, it makes confident that at least this is an indication of the model feasibility, mean this issue is not related to data gathering. Also, with bigger datasets, the systems should have a better performance concerning the accuracy classification and data, thereby solving the underfitting challenge.

5. CONCLUSION

In this research, the ultimate point to clarify an improvement in the accuracy of classifying the image files through the process of combining the actions concerning the movement. These are inclusive of the training and testing comprised within the images. This study is based on a design of a system reliant on deep learning more particularly a CNN.

Different articles in different ideas, and all have high accuracy given at Table 5.1 It seems that the accuracy obtained in this research can be considered as successful under so similar activity separation and limited data set.

Table 5.1. Comparison of Literature

Researchers	Accuracy Rate	Year	Dataset
Lee et al,	92.7%	2017	Own Dataset
Chen et al,	93.8%	2015	Samples
Huang et al,	87.9%	2019	Own dataset
El-Fiqi et al,	97.0%	2018	Collection

Our results indicate that the performance is well structured. However, the research has some shortcomings in regards to the dataset, scope and the number of epochs done on this project. At this time, we can hypothesize that this is due to the nature of those particular movements that substantially differ from each other, which makes it modest for the network to differentiate them from each other.

In a future, it is expected that training will be improved by the increasing number of epochs as well as extending the scope and the capacity of the project by adding of other actions, such as taking into consideration active postures of the body, the head of running people marked and processed for emotion analysis that could be of wonderful.

REFERENCES

Affonso, C., Rossi, A. L. D., Vieira, F. H. A., & de Leon Ferreira, A. C. P. (2017). Deep learning for biological-image classification. Expert Systems with Applications, , 85, 114-122.

Affonso, C., Sassi, R. J., Barreiros, R. M. (2015). Biological image classification using rough-fuzzy artificial neural network. Expert Systems with Applications, 42(24), 9482-9488.

Ajuzieogu, U.. (2019). The Role of Artificial Intelligence in Modern Computing and Education.

Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017). Understanding of a convolutional neural network. IEEE International Conference on Engineering and Technology (ICET), 1-6.

Alsallakh, B., Jourabloo, A., Ye, M., Liu, X., Ren, L.. (2018). Do Convolutional Neural Networks Learn Class Hierarchy?. IEEE Transactions on Visualization and Computer Graphics, 24(01), 152-162.

Alwzwazy, H.A., Albehadili, H.A., Alwan, Y.S., Islam, N.E. (2016) Handwritten digit recognition using convolutional neural networks. In: Proceedings of international journal of innovative research in computer and communication engineering, 4(2), 1101-1106.

Ben-Hur, A., Horn, D., Siegelmann, H. T., & Vapnik, V. (2001). Support vector clustering. Journal of machine learning research, 2(12), 125-137.

Bengio, Y., Courville, A., Vincent, P. (2013). Representation learning: a review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35(8), 1798-1828.

Bini, S. A. (2018). Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?. The Journal of arthroplasty, 33(8), 2358-2361.

Carrasquilla, J., Melko, R. G. (2017). Machine learning phases of matter. Nature Physics, 13(5), 431-434.

Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L. (2018) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834-848.

Chen, Y., & Xue, Y. (2015). A deep learning approach to human activity recognition based on single accelerometer. IEEE international conference on systems, man, and cybernetics, 1488-1492.

Ciresan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011). Flexible, high performance convolutional neural networks for image classification. In Twenty-second international joint conference on artificial intelligence. 2(2), 1237-1242.

Cong, I., Choi, S., Lukin, M. D. (2019). Quantum convolutional neural networks. Nature Physics, 15(12), 1273-1278.

Craven, M. W., & Shavlik, J. W. (1997). Using neural networks for data mining. Future generation computer systems, 13(2-3), 211-229.

Davis, J. W., Tyagi, A. (2006). Minimal-latency human action recognition using reliable-inference. Image and Vision Computing, 24(5), 455-472.

Ding, C., Li, Y., Xia, Y., Wei, W., Zhang, L., & Zhang, Y. (2017). Convolutional neural networks based hyperspectral image classification method with adaptive kernels. Remote Sensing, 9(6), 618.

Du, L., Du, Y., Li, Y., Su, J., Kuan, Y. C., Liu, C. C., & Chang, M. C. F. (2017). A reconfigurable streaming deep convolutional neural network accelerator for Internet of Things. IEEE Transactions on Circuits and Systems I: Regular Papers, 65(1), 198-208.

El-Fiqi, H., Wang, M., Salimi, N., Kasmarik, K., Barlow, M., Abbass, H. (2018) Convolution Neural Networks for Person Identification and Verification Using Steady State Visual Evoked Potential. IEEE International Conference on Systems, Man, and Cybernetics (SMC), 1062-1069.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning. Cambridge, MIT press, 1(2), 326p.

Hofmann, T., Schölkopf, B., & Smola, A. J. (2008). Kernel methods in machine learning. The annals of statistics, 36(3), 1171-1220.

Hou, R., Chen, C., & Shah, M. (2017). Tube convolutional neural network (t-cnn) for action detection in videos. In Proceedings of the IEEE international conference on computer vision, 5822-5831.

Huang, J., Lin, S., Wang, N., Dai, G., Xie, Y., & Zhou, J. (2019). Tse-cnn: A two-stage end-to-end cnn for human activity recognition. IEEE journal of biomedical and health informatics, 24(1), 292-299.

Jin, K. H., McCann, M. T., Froustey, E., Unser, M. (2017). Deep convolutional neural network for inverse problems in imaging. IEEE Transactions on Image Processing, 26(9), 4509-4522.

Khan, M. N. A., & Wakeman, I. (2006). Machine learning for post-event timeline reconstruction. In First Conference on Advances in Computer Security and Forensics, Liverpool, UK.

Kong, Y., Fu, Y. (2018). Human Action Recognition and Prediction: A Survey. journal of latex class files, 13(9), 1-20.

Lapuschkin, S., Binder, A., Muller, K. R., & Samek, W. (2017). Understanding and comparing deep neural networks for age and gender classification. In Proceedings of the IEEE International Conference on Computer Vision Workshops, 1629-1638.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436–444.

Lee, S. M., Yoon, S. M., & Cho, H. (2017). Human activity recognition from accelerometer data using Convolutional Neural Network. IEEE international conference on big data and smart computing (bigcomp), 131-134.

Martinez, J., Black, M. J., and Romero, J. (2017). On Human Motion Prediction Using Recurrent Neural Networks. 4674-4683.

Masoud, O., Papanikolopoulos, N. (2003). A method for human action recognition. Image and Vision Computing, 21(8), 729-743.

Mishra, C., & Gupta, D. L. (2017). Deep machine learning and neural networks: An overview. IAES International Journal of Artificial Intelligence, 6(2), 66.

Peng, Y., & Yin, H. (2017). Markov random field based convolutional neural networks for image classification. In International Conference on Intelligent Data Engineering and Automated Learning, Springer, Cham, 387-396.

Rajalingappaa, S. (2018). Deep Learning for Computer Vision: Expert techniques to train advanced neural networks using TensorFlow and Keras. Packt Publishing Ltd, 310p, Birmingham.

Samborska, I. A., Alexandrov, V., Sieczko, L., Kornatowska, B., Goltsev, V., Cetner, M. D., & Kalaji, H. M. (2014). Artificial neural networks and their application in biological and agricultural research. J. NanoPhotoBioSciences, 2(14). 2347-7342.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. IBM Journal of research and development, 3(3), 210-229.

Shalev-Shwartz, S., & Ben-David, S. (2014). Understanding machine learning: From theory to algorithms. Cambridge university press I-Iv.

Venkatesan, R., & Li, B. (2017). Convolutional neural networks in visual computing: a concise guide. CRC Press, Taylor and Francis group, 186p, Boca Raton.

Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., & Xu, W. (2016). Cnn-rnn: A unified framework for multi-label image classification. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2285-2294.

Wang, M., El-Fiqi, H., Hu, J., Abbass, H. A. (2019). Convolutional neural networks using dynamic functional connectivity for EEG-based person identification in diverse human states. IEEE Transactions on Information Forensics and Security, 14(12), 3259-3272.

Wen, L., Li, X., Gao, L., Zhang, Y. (2017). A new convolutional neural network-based data-driven fault diagnosis method. IEEE Transactions on Industrial Electronics, 65(7), 5990-5998.

Yamashita, R., Nishio, M., Do, R. K. G., Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into imaging, 9(4), 611-629.

Zhou, D. X. (2020). Universality of deep convolutional neural networks. Applied and computational harmonic analysis, 48(2), 787-794.

Zurlini, G., & Müller, F. (2008). Environmental security. In Sven Erik Jørgensen and Brian D. Fath, Systems Ecology. 2(1-5), 1350-1356.

APPENDIX

Appendix A - Tensorflow Libraries

```
import keras
from keras.models import Sequential
from keras.layers import Activation
from keras.layers.core import Dense, Flatten
from keras.optimizers import Adam
from keras.metrics import categorical_crossentropy
from keras.preprocessing.image import ImageDataGenerator
from keras.layers.normalization import BatchNormalization
from keras.layers.convolutional import *
from sklearn.metrics import confusion_matrix
from sklearn.metrics import accuracy_score
import itertools
import matplotlib.pyplot as plt
%matplotlib inline
```

Appendix B - Loading data to system

```
train_path = 'data/train'

test_path = 'data/test'

train_batches = ImageDataGenerator().flow_from_directory(train_path,
target_size=(224,224),

    classes=['walking', 'running'], batch_size=800)

test_batches = ImageDataGenerator().flow_from_directory(test_path, target_size=(224,224),

    classes=['walking', 'running'], batch_size=300)
```

Appendix C - Training process

```
model = Sequential([
    Conv2D(32, (3, 3), activation='relu', input_shape=(224,224,3)),
    Flatten(),
    Dense(2, activation='softmax'), ])
model.compile(Adam(lr=0.0001), loss='categorical_crossentropy', metrics=['accuracy'])
history = model.fit_generator(generator=train_batches, steps_per_epoch=100,
    validation_data=test_batches, validation_steps=100, epochs=5000, verbose=2)
```

Appendix D – Output Structure

```
predictions = model.predict_generator(generator=train_batches, steps=1, verbose=0)
cm = confusion_matrix(test_labels, predictions[:,0])
def plot_confusion_matrix(cm, classes,
                          normalize=False,
                          title='Confusion matrix',
                          cmap=plt.cm.Blues): ""
    plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar()
    tick_marks = np.arange(len(classes))
    plt.xticks(tick_marks, classes, rotation=45)
    plt.yticks(tick_marks, classes)
    if normalize:
        cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]
        print("Normalized confusion matrix")
    thresh = cm.max() / 2.
    for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
        plt.text(j, i, cm[i, j],
                 horizontalalignment="center",
                 color="white" if cm[i, j] > thresh else "black")
    plt.tight_layout()
    plt.ylabel('True label')
    plt.xlabel('Predicted label')
cm_plot_labels = ['walking','running']
plot_confusion_matrix(cm, cm_plot_labels, title='Confusion Matrix')
```

CURRICULUM VITAE

Name and Surname : Mohamed Artan Abdi

Place of Birth : Mogadishu, Somalia

Date of Birth : 11/09/1997

Marital Status : Single

Foreign Languages : English (Advanced)
Turkish (Pre-Intermediate)
Arabic (Beginner)

Email : mdartan4@gmail.com

Educational Status

Bachelor's Degree : Eelo University
Faculty of Computer Science and Information
Technology 2016

Master's Degree : Ticaret University
Graduate school of natural and applied sciences
Department of Computer Engineering, 2021

Work Experience

SOON : IT officer
September 2016–2019

Tusmo Company : Chief Designer
October 2014-2015

Borama Electronic : Computer Technician
January 2013-2014

Publications

Mohamed Artan Abdi & Metin Turan (2021) 'Event or Emergency Case Detection by Human Running', 6th International Congress on Information and Communication Technology (ICICT 2021), London, Uk