

Real-Time Clinical Gait Analysis and Foot Anomalies Detection Using Pressure Sensor and Convolutional Neural Network

By

Muhammad Sayem Hasan (170021051)

Musarrat Tabassum (170021067)

Mahdi Islam (170021079)

A Thesis Submitted to the Academic Faculty in Partial Fulfillment of the Requirements for
the Degree of

Bachelor of Science in Electrical and Electronic Engineering




Department of Electrical and Electronic Engineering
Islamic University of Technology (IUT)
Gazipur, Bangladesh.

May 2022.

CERTIFICATE OF APPROVAL

The thesis titled "Real-Time Clinical Gait Analysis and Foot Anomalies Detection Using Pressure Sensor and Convolutional Neural Network" submitted by Muhammad Sayem Hasan (170021051), Musarrat Tabassum (170021067) and Mahdi Islam (170021079) has been found as satisfactory and accepted as partial fulfillment of the requirement for the degree of Bachelor of Science in Electrical and Electronic Engineering on 10th May, 2022.

Approved by:



(Signature of the Supervisor)

Mirza Muntasir Nishat

Assistant Professor

Department of Electrical and Electronic Engineering (EEE)

Islamic University of Technology (IUT)



(Signature of the Co-Supervisor)

Fahim Faisal

Assistant Professor

Department of Electrical and Electronic Engineering (EEE)

Islamic University of Technology (IUT)

Declaration of Authorship

This is to certify that the work presented in this thesis paper is the outcome of research carried out by the candidates under the supervision of Mirza Muntasir Nishat, Assistant Professor, Department of Electrical and Electronic Engineering (EEE), Islamic University of Technology (IUT). It is also declared that neither this thesis paper nor any part thereof has been submitted anywhere else for the reward of any degree or any judgment.

Authors

Sayem
10-5-22

Muhammad Sayem Hasan
ID-170021051

Musarrat
10-5-22

Musarrat Tabassum
ID-170021067

Mahdi
10-5-22

Mahdi Islam
ID-170021079

Dedicated to

*Our beloved parents & teachers whose support made it all
possible for us*

TABLE OF CONTENTS

List of Tables.....	vii
List of Figures.....	viii
List of Acronyms.....	X
Acknowledgments.....	xi
Abstract.....	xii
1. Introduction.....	01
1.1 Introduction.....	01
1.2 Symptoms & Different types of abnormal gait.....	02
1.3 Gait analysis techniques.....	07
1.4 Objective of this research.....	10
1.5 Main Contribution.....	10
1.6 Thesis Outline.....	11
2. Literature Review.....	12
2.1 Relevant Research.....	12
2.2 Comparative Analysis of Relevant Research.....	13
3. Methodology.....	14
3.1 Basic Methodology.....	14
3.2 Description of Basic Methodology.....	15
3.3 Description of Performance Matrices in ML/DL.....	15
3.4 Formula of Performance Matrices.....	16
4. Data Preprocessing.....	17
4.1 Dataset.....	17
4.2 Image Extraction.....	17
4.3 Image transformation and class expansion.....	18
4.4 Image Augmentation.....	20
5. Introduction to Algorithms.....	29
5.1 Convolutional Neural Network.....	29
5.2 Transfer Learning.....	31
5.2.1 Approaches of Transfer Learning.....	32
5.2.2 VGG16.....	34
5.2.3 ResNet50.....	35
5.2.4 MobileNet_V2.....	36

5.2.5 DenseNet.....	36
6. Result & Analysis.....	37
6.1 Implementation of Transfer Learning Algorithms.....	37
6.1.1 VGG16.....	37
6.1.2 ResNet50.....	39
6.1.3 MobileNet_V2.....	40
6.1.4 DenseNet.....	42
6.2 Comparative Analysis of Transfer Learning Algorithms.....	43
6.2.1 Training Phase (20 Epochs).....	43
6.2.2 Training Phase (30 Epochs).....	44
6.2.3 Training Phase (40 Epochs).....	45
6.2.4 Test Phase.....	46
7. Conclusion & Future works.....	48
7.1 Conclusion.....	48
7.2 Future works.....	49
References.....	52

List of Tables

Table 2.1:	<i>Comparison between relevant works.....</i>	13
Table 4.1:	<i>Number of image samples per class before image augmentation.....</i>	20
Table 4.2:	<i>Number of image samples per class after image augmentation.....</i>	28
Table 6.1:	<i>Results.....</i>	46

List of Figures

Figure 1.1:	<i>Execution of gait cycle</i>	1
Figure 1.2:	<i>Spastic gait</i>	4
Figure 1.3:	<i>Scissors gait</i>	5
Figure 1.4:	<i>Steppage gait</i>	5
Figure 1.5:	<i>Waddling gait</i>	6
Figure 1.6:	<i>Propulsive gait</i>	6
Figure 3.1:	<i>Basic Methodology Flow-chart</i>	14
Figure 4.1(a):	<i>Visual representation of Left Foot Backward Leaned</i>	18
Figure 4.1(b):	<i>Visual representation of Left Foot Forward Leaned</i>	19
Figure 4.2:	<i>Nine classes of gait disorders</i>	19
Figure 4.3:	<i>Image Rotation</i>	22
Figure 4.4:	<i>Image Flipping (a)Original Image (b)Horizontal Flip (c)Vertical Flip</i> ...22	
Figure 4.5:	<i>Image Noising</i>	23
Figure 4.6:	<i>Image Blurring</i>	24
Figure 4.7:	<i>Image Padding</i>	24
Figure 4.8:	<i>Image Scaling</i>	25
Figure 4.9:	<i>Image translation</i>	25
Figure 4.10:	<i>Image Shearing</i>	25
Figure 4.11(a):	<i>Initial Training Accuracy/Loss before augmentation</i>	27
Figure 4.11(b):	<i>Initial Training Accuracy/Loss after augmentation</i>	27
Figure 5.1:	<i>CNN architecture</i>	30
Figure 5.2:	<i>Transfer Learning</i>	31
Figure 5.3:	<i>Examples of Difference in Transfer Learning</i>	32
Figure 5.4:	<i>Transfer Learning Overview</i>	33
Figure 5.5:	<i>VGG16 architecture</i>	34
Figure 5.6:	<i>ResNet architecture</i>	35
Figure 6.1:	<i>Accuracy & Loss during training for VGG16</i>	37
Figure 6.2:	<i>Confusion Matrix of VGG16</i>	38
Figure 6.3:	<i>Accuracy & Loss during training for ResNet50</i>	39
Figure 6.4:	<i>Confusion Matrix of ResNet50</i>	40
Figure 6.5:	<i>Accuracy & Loss during training for MobileNetV2</i>	40
Figure 6.6:	<i>Confusion Matrix of MobileNetV2</i>	41
Figure 6.7:	<i>Accuracy & Loss during training for DenseNet</i>	42
Figure 6.8:	<i>Confusion Matrix of DenseNet</i>	43
Figure 6.9:	<i>Training and validation Accuracy of the four models at 20 epochs</i>	43
Figure 6.10:	<i>Training and validation Loss of the four models at 20 epochs</i>	44

Figure 6.11:	<i>Training and validation Accuracy of the four models at 30 epochs.....</i>	<i>44</i>
Figure 6.12:	<i>Training and validation Loss of the four models at 30 epochs.....</i>	<i>45</i>
Figure 6.13:	<i>Training and validation Accuracy of the four models at 40 epochs.....</i>	<i>45</i>
Figure 6.14:	<i>Training and validation Loss of the four models at 40 epochs.....</i>	<i>46</i>
Figure 7.1:	<i>Generation of piezoelectricity.....</i>	<i>49</i>
Figure 7.2:	<i>Methodology of generation of piezoelectricity.....</i>	<i>50</i>
Figure 7.3:	<i>Piezoelectric Sensor.....</i>	<i>50</i>

List of Acronyms

GDD	Gait Disorder Detection
CNN	Convolutional Neural Network
ADAM	Adaptive Moment
ReLU	Rectified Linear Unit
SGD	Stochastic Gradient Descent
VGG16	Visual Geometry Group
ResNet50	Residual Network 50
MobileNet_V2	Mobile Network Version2
DenseNet	Dense Network

ACKNOWLEDGMENTS

Foremost, we would like to express our sincere gratitude and gratefulness to the Almighty Allah; without His graces and blessings, this study would not have been possible.

Acknowledging all who helped us complete this work, we wish to compliment the university's significant role and the department that has been very amiable to us during the entire period of our research.

We are indebted to our honorable supervisor, Mirza Muntasir Nishat sir, for his selfless support, motivation, patience, enthusiasm, and extensive knowledge of the relevant fields. His continuous guidance and careful supervision kept us going even during the hardest of hours. Lastly, our warmest tribute to our parents, family members, and friends, whose moral support and well wishes benefitted us spiritually in achieving our goals.

ABSTRACT

This research presents a novel insight on gait disorder detection using transfer learning algorithms on sensor-acquired data based on the implementation of popular Convolutional Neural Network (CNN) models. So, the research basically deals with Deep Learning. The dissertation proposes the use of pressure sensors to extract heatmap images during gait, which are then trained and tested in various classification algorithms for gait abnormality diagnosis and detection. Gait is a biological and scientific study of body movement and locomotion that emphatically serves as a reliable parameter for inspecting the human body's neuromuscular and skeletal systems. To build a convenient and precise classification system for possible application, synthetic data was generated in multiple preexisting CNN models, which were then evaluated using conventional performance metrics. The proposed notion yielded experimental findings that showed higher accuracies for all transfer learning schemes tested, with the Vgg16 model achieving a notable accuracy of 97.15%. As a result, the analysis demonstrated not only a significant performance in terms of accuracy, but also reduced complexity and computing time, making the approach efficient yet effective. A detailed comparative analysis of performance with all other algorithms was carried out in terms of accuracy, precision, recall, and F-1 score.

Chapter-1

Introduction

1.1 Introduction

A hale mind and a sound human life is nothing but physical sturdiness & mental soundness. Human mobility is a prerequisite for both physical and mental health & any form of natural locomotion such as walking; jogging or running can be interpreted as human mobility or gait. This eventually results in the introduction of gait analysis which emphasizes on the observation of gait & the detection of any type of gait abnormalities, postural irregularities in addition to the evaluation of clinical mediation and rehabilitation programs [1]

The word gait refers to the motion while taking a stride, jogging or any kind of locomotion. Muscular system [2], cardiovascular system [3], nervous system is collaboratively functioning in a complex process while a human gait is in effect; so gait anomaly may indicate fault in any of these systems. The execution of a proper gait cycle requires the amalgamation of cognitive ability and muscle precision [4].

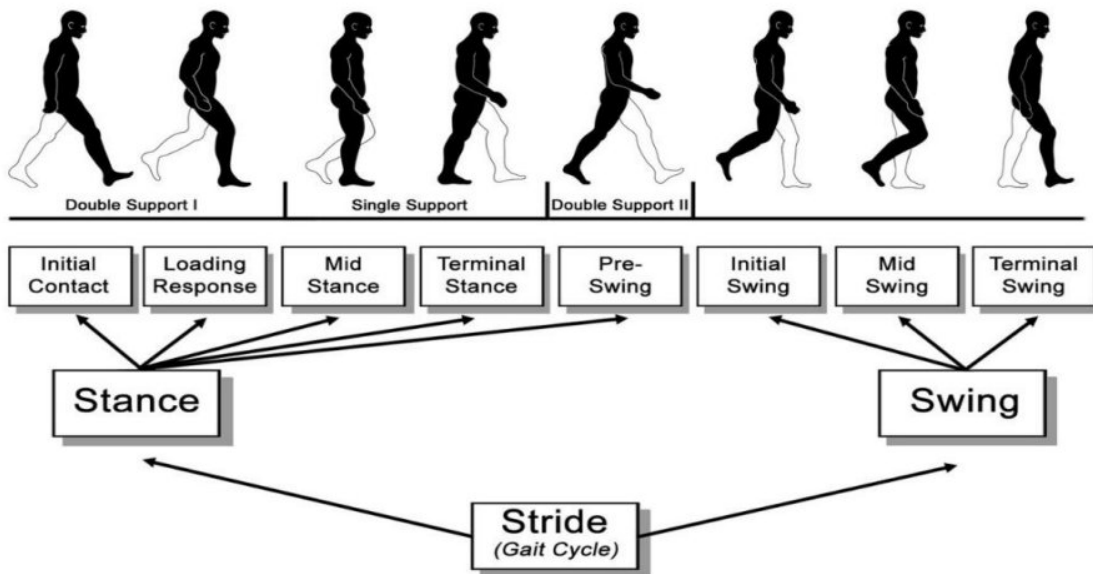


Figure 1.1- Execution of gait cycle

The graphic depicts a complete gait cycle. The phases have been called based on where the black colored leg is located. The graphic depicts the beginning phase till a similar phase is repeated. As a result, this procedure is comparable to periodic motion. (This is a simple harmonic motion.)

The black-colored leg may be seen in the illustration. The heel is barely on the ground. The term "first contact" is used to describe it. "Loading reaction," "Mid stance," "Terminal posture," and "Pre-swing" are all seen after that. The **STANCE** phase includes all of these steps. Until the phase "Terminal Stance" occurs, the black-colored leg bears the weight of the entire body. In the meantime, the rear leg of the opposite leg slides forward. Only the black

colored leg's digits are contacting the floor during the "Pre-swing" phase. The heel of the second leg, on the other hand, is on the floor. This completes the first part of the gait cycle. The **SWING** phase follows. The leg with the black color leaves the surface and moves forward faster than the other leg. In the meantime, the other leg is still on the floor. That indicates the other leg will follow the **STANCE** phase in the second phase. The black colored leg moves through three phases: "**Initial swing**," "**Mid swing**," and "**Terminal swing**" as we follow it.

The scene was witnessed from the black-colored leg's heel contacting the ground. The identical posture is now repeated following these eight steps. As a result, Gait Cycle may be defined as a periodic motion.

The efficient execution of a gait cycle necessitates the coordination of all muscles involved in the walking process. While a human is in action, the muscular system, cardiovascular system, and neural system all work together in a complicated process, therefore gait abnormalities might signal a problem with any of these systems.

Proper medication & research regarding gait disorder is imperative since gait anomaly may have repercussions like human imbalance that eventually may be the precursors of collapsing leading to severe injuries. Symptoms of gait disorder vary from person to person (mainly for the senile and athletes) however, there's a range of prevalent symptoms including strenuousness while taking strides, vertigo, dizziness etc. and these core issues directed to perform extensive research in this field & an important factor for the identification of Gait Activity Detection (GAD), Gait Event Detection (GED), Gait Disorder Detection (GDD) & so on [5]. GDD was formerly accomplished by videotaping patients' movements with computer-controlled video cameras, Inertial Measurement Unit (IMU) sensors, and numerous small body monitoring devices [6]. These were typically and manually done by medical experts which increases the complexity of the process significantly. From the previous decades the evolution of machine learning enabled the researchers to implement it on the detection of various gait disorders but this approach also had noteworthy limitations. Convolutional Neural network, which is a subsidiary of Deep Neural Network, is flexible to both big data & multidimensional data. The deployment of CNN was based on heatmap images from pressure sensors obtained from foot movement videos. This study makes a significant approach by making gait detection easier since there are no regulations of observing the whole locomotion using videos or 3D motion analysis.

1.2 Symptoms & Different types of abnormal gait

The study focuses on gait abnormalities and attempts to discover the problem in a short amount of time. Gait problem occurs when a person is unable to complete a gait cycle effectively. Gait issues can be detected by a few indicators. Let's take a look at a few of them.

Temporary gait or balance issues might be caused by a variety of factors.

- Injury
- Trauma
- Inflammation
- Pain

Muscle neurological disorders can cause long-term problems. Specific conditions can create problems involving gait, balance, and coordination, including:

1. Joint pain or conditions, such as arthritis
2. Multiple sclerosis (MS)
3. Meniere's disease
4. Brain hemorrhage
5. Brain tumor
6. Parkinson's disease
7. Chiari malformation (CM)
8. Spinal cord compression or infarction
9. Guillain-Barré syndrome
10. Peripheral neuropathy
11. Myopathy
12. Cerebral palsy (CP)
13. Gout
14. Muscular dystrophy
15. Obesity
16. Chronic alcohol misuse
17. Vitamin B-12 deficiency
18. Stroke
19. Vertigo
20. Migraine
21. Deformities
22. Certain medications, including antihypertensive drugs

Gait problems can also be caused by a restricted range of motion or weariness. Muscle weakness might affect one or both legs at the same time. Walking becomes difficult as a

result. Numbness in the feet and legs might make it difficult to tell where your feet are moving or whether they are on the ground.

Gait disorder is classified into one of five categories based on the symptoms of an individual's walk.

- **Spastic gait**

When a person's one leg is rigid and drags in a semicircular motion on the side most affected by long-term muscular tightness, this is known as spastic gait. Patients with cerebral palsy or multiple sclerosis are prone to this. Increased tone in the adductors causes the legs to cross in severe cases.



Figure 1.2- Spastic gait

- **Scissors gait**

A scissors gait is common in people whose legs fold inward. When walking with this kind, a person's legs cross and may collide. The crossing action resembles the opening and shutting

of scissors. High muscular tone (spasticity) in the hip adductors causes it. The hip adductors are primarily responsible for pulling the thighs together.

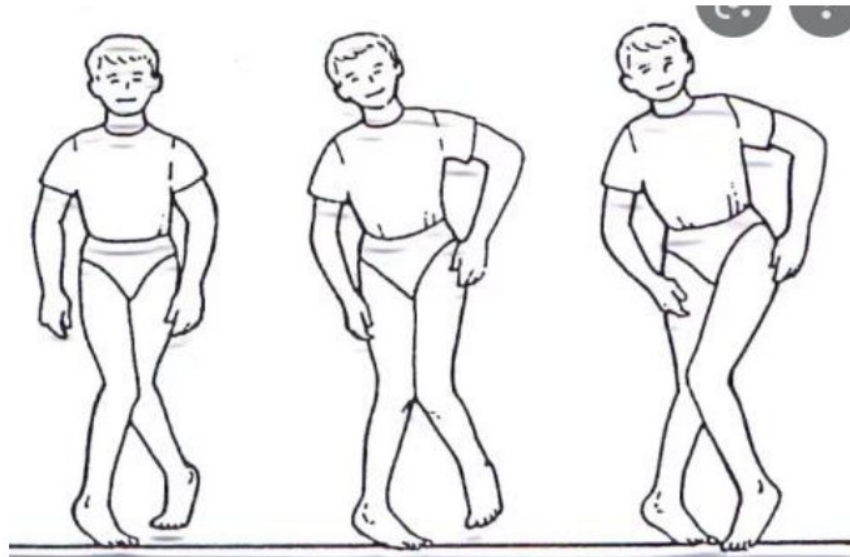


Figure 1.3- Scissors gait

- **Steppage gait**

When a person walks, their toes point to the ground, which is known as steppage gait. As a person moves forward, the toes frequently scrape against the ground. As a result of the weakening of the muscles that induce dorsiflexion of the ankle joint, this is essentially an inability to elevate the foot when walking.

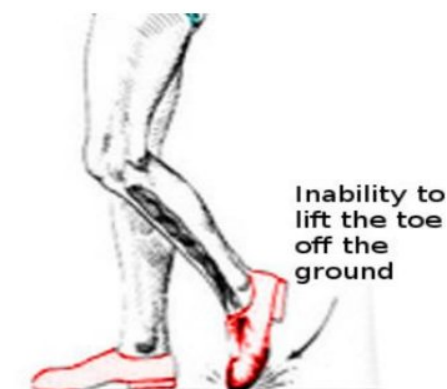


Figure 1.4- Steppage gait

- **Waddling gait**

When walking, a person with a waddling stride goes from side to side, as the term implies.



Figure 1.5- Waddling gait

This condition is more common in persons who have bilateral gluteus medius muscle weakening, which is the major hip abductor. As a result, the weakening of the hip girdle and upper thigh muscles is the primary cause of this illness.

- **Propulsive Gait**

A person with this form of gait impairment will be halted with tight posture and a forward bent head and neck. Toxins, Parkinson's disease, or specific drugs are the most common causes. The individual may appear to be tightly keeping a slouched pose.

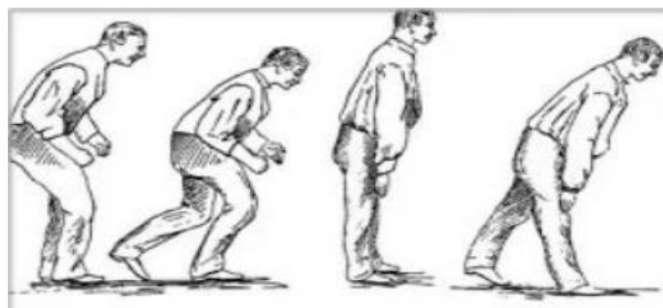


Figure 1.6- Propulsive gait

A person with a limp is also regarded to have an aberrant gait, in addition to these five categories. A limp might be transient or permanent, much like other deformities. Without medical assistance, a limp may disappear.

These are the many kinds of gait abnormalities. These forms of foot abnormalities can be inherited. However, similar injuries might occur in sportsmen or regular people during everyday activity. As a result, it is critical to initially discover these sorts of abnormalities. Then adequate care may be provided. However, if the problems are not correctly identified, the treatment may fail. So, let's have a look at several gait analysis methods.

1.3 Gait analysis techniques

In the field of research, gait analysis is becoming older. The description of the walking principle by Leonardo da Vinci, Galileo, and Newton started a scientific study of gait. Borelli, a Galileo student, outlined how balanced walking might be achieved utilizing the notion of the body's center of gravity in 1682. Edward Muybridge and Leland Stanford were the first to research the gait mechanism in 1878. Despite the fact that substantial and extensive study has previously been conducted in this area, it has yet to be properly exploited. Gait analysis research is a continual endeavor since new models and approaches are always being created. With the introduction of new computer techniques and technology in 1960, clinical gait analysis has gained traction. As a result of the ongoing research process, there are currently a variety of techniques to gait analysis. Following a thorough investigation, it has been shown that modern human gait analysis methodologies may be divided into four categories:

- **Vision based approach-**

Analog or digital cameras are used to analyze gait. This can be accomplished by 2 methods.

1. **Direct approach -**

Direct vision based technique is another name for this method. This approach is often used to determine a human's kinematics. Active or passive markers were created to carry out real-time direct vision based gait analysis procedures. An optoelectronic system is used in the direct gait analysis approach. Optoelectronic systems are designed to transform light signals

into electrical signals and track the light produced or reflected by markers. Active markers are miniature LEDs that emit a signal and may be mounted to the subject. The cameras emit infrared light signals and detect the reflected signals from the body markers. Because separate markers operate at predetermined frequencies, this signal is utilized to pinpoint the position of the marker. Given that these systems can frequently find markers uniquely at a distance as tiny as 1 mm, their precision is excellent.

This strategy, however, has a few drawbacks. The presence of wires or components that potentially impact the subject's gait pattern obstructs their natural movements. Near infrared light is utilized to illuminate the markers, which reflect light emitted by the cameras back to them. Passive markers are reflective scotchlite tape-covered spheres. They're made to reflect incident light back along the line of incidence. The researchers provide a variety of angles for observing gait abnormality, but they all agree on two optimum views for observing gait abnormality (sagittal or lateral and frontal or rear). Many camera-based motion capture systems are available on the market for 2D and 3D analysis. For two-dimensional analysis, one camera is sufficient. The camera is in a sagittal perspective, however this method has the drawback of not being able to catch all out of plane movements. This constraint can be circumvented using three-dimensional analysis, however more than one camera is required. A point of interest region of the topic should be viewed by at least two cameras at the same time to project the three-dimensional model. Leading gait laboratories often employ four to eight cameras.

2. Indirect approach -

Markers are not used in the indirect method to vision-based gait analysis. The attributes of the participants can be analyzed using the footage collected by video cameras, either based on the model, appearance, or a combination technique. This is the most popular and widely used method of person recognition in surveillance based on human gait.

- **Sensor based approach-**

Sensors can be put on the subject's body or on the floor to correctly perform gait analysis. Electromyography (EMG) and inertial systems are applied to the subject's body on a surface or middle basis. The kinetics of the subject's movement may also be obtained with Force Platform.

EMG is used to investigate muscle electrical activity while walking. Motor Unit Action Potentials (MUAPs) are recorded using needle or surface EMG electrodes. Even a single movement involves a number of muscles. The amplitude of EMG signals obtained during walking can also be used to determine relative muscle tension, although special knowledge of electrode placement is required, and they are susceptible to interference.

Gait analysis with inertial systems is also possible. The system works on the basis of motion resistance. Inertia and segment orientation are measured using accelerometers and gyros. Gyroscope sampling rates are similar to accelerometer sampling rates. To get the kinematics of a subject's movement, several researchers propose using gyroscopes with accelerators. Segment location, step detection, and stride length may all be determined using them. It is vulnerable to disturbance and has complicated algorithms.

Floor platform-based sensors are another sort of sensor. These sensors are used to determine the forces involved in the production of ground response force, force pattern, foot plantar pressure distribution, and step and gait phase detection. When the foot comes into touch with a GRF plate, the force magnitude and direction are calculated. Transducers (with steel plates) affixed to each corner of the GRF plate sense the force imparted to the plate. This force is translated to an electrical signal, which is then used to determine the center of pressure, which is one of three orthogonal force components in the subject's movement.

The last sort of sensor to discuss is the pressure sensor. These are used to obtain load information for sensors. They are inserted inside the shoe's insoles. These piezoelectric-based sensors provide an electrical signal when mechanical strain is applied to them. The health specialists will then conduct a gait analysis. However, this method has several disadvantages. This approach has a space constraint. Subjects must maintain their foot on the middle of a plate of floor platforms for correct measurement; this makes them aware, and they are unable to demonstrate their regular pattern. Another barrier is the financial aspect. Even the most well-equipped gait labs have a maximum of two or three plates. As a result, healthcare providers are unable to catch the participants' frequent patterns.

Other gait analysis systems are also available. As a human gait analysis technique, other technologically based approaches such as Electrogoniometer, magnetic, and medical imaging based systems can be applied. In a time-constrained context, this Electrogoniometer-based

technique is ineffective since it takes time to connect to a subject. Because it uses a magnetic field to track ferromagnetic markers, the magnetic system-based technique does not require a line of sight for the markers, as vision-based marker approaches do. Magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound can be used to collect movement and anatomical data of the subject's section. It's then utilized to create a subject-specific computational model to which kinematic and kinetic data may be applied. Interference affects these systems as well. A hybrid technique, which combines two or more of the above approaches, is also available. For a better understanding of human gait, scientists employed vision, EMG, and force platforms.

1.4 Objective of this research

The process of gait analysis is time consuming at times. The primary objective of this research is to develop an automated technique for detecting gait anomalies using pre-trained Convolutional Neural Networks; **Vgg16, ResNet50, MobileNet_V2, and DenseNet**. Furthermore, when compared to comparable fields of research, this analysis done on sensor images resulted in higher accuracy as well as a shorter computation time which has been possible due to the introduction of more output categories or classes. Thus, this research on GDD may ameliorate the detection process in the medical sector because of the fact that the recognition problem can effectively be solved by using only images instead of videos that eliminates the whole complexity of 3D data preprocessing.

1.5 Main Contribution

Since the main goal is to execute gait analysis properly and detect foot anomalies within a short span of time, pre-processed synthesized data have been considered. The data was basically a heat map video which was generated by a pressure sensor. Then, using various software's and editing processes, the dataset was generated. Then, Machine Learning algorithms were applied on the image datasets. The proposed analysis divides gait irregularities into nine categories, with eight of them containing distinct gait or foot anomalies and the remaining one being normal gait or foot. The accuracy, precision, recall, and f1 score obtained from assessing all of the models on the test data is observed. Various transfer learning algorithms have been applied to observe the performance parameters. Maximum accuracy was obtained in the algorithm **VGG16** and obtained an accuracy of

97.15%.

1.6 Thesis Outline

In chapter – 1, the gait cycle process and abnormalities in gait cycle are discussed. Then, symptoms and various types of gait disorders were explained. Finally, the necessity of ML & DL algorithms for increasing the accuracy of diagnostics & the main contribution of the research work is also described.

In chapter – 2, relevant research works published recently on human gait analysis & foot anomalies detection using ML & DL algorithms are analyzed comparatively in-line with this research topic of interest.

In chapter – 3, the overall workflow diagram and the methodology of the study are presented. This chapter depicts the details of the approach introduced in this research analysis.

In chapter-4, the data preprocessing steps are described in detail along with the necessity and reason of using synthetic data through data augmentation techniques are broadly discussed.

This chapter clearly shows the comparison of the number of samples due to the implementation of the data augmentation techniques. The Data preprocessing & Feature engineering process that we've performed are visually depicted, which is an integral part of any ML/DL research.

In chapter-5, introduction to all the algorithms used in this research are broadly discussed. The DL and transfer learning schemes and how they work along with the architecture of different transfer learning schemes and algorithms are explained in detail.

In chapter-6, description of the confusion & performance matrices of all the 4 algorithms are shown & compares the overall result & performance which portrays **97.15%** accuracy in the classification of the VGG16 algorithms on the used dataset.

Finally, the discussion and future prospects of our research work is concluded in chapter – 7 by mining out the ways to improve & better implement our study of interest.

Chapter-2 *Literature Review*

2.1 Relevant Research

In the recent past ample research has been done in order to unravel more developed prospects for gait anomaly detection and gait analysis. Previously, Z. Liu and S. Sarkar introduced a method for gait analysis called average silhouette representation for a sequence [17]. J. Han and B. Bhanu introduced the GEI method [18]. Average silhouette representation was shown to be more effective because it took less time to compute.

Authors also presented GPE [19] and GPPE [20]. According to H. Iwama's research, large gait datasets reduce GEI's performance level [21]. Another significant proposal given by M. Turk and A. Pentland for gait recognition is Principal component analysis to excerpt the gait's distinctive properties [22]. Though this technique failed when working with 2D dimensional data, Linear Discriminant Analysis (LDA) [23] was developed in low dimensional space segregated classes. PCA and LDA was a simple way of detection but the performance fell as a result of the translation of two-dimensional data to one dimensional data, as PCA and LDA cannot function with two-dimensional data.

Furthermore, several other writers later expanded research in this subject by developing subspace learning algorithms for gait recognition [24] and incorporating learning from features through the use of high order tensors [25]. Z. Lai, Y. Xu et al. suggested matrix-based Sparse Bilinear Discriminant method (SBDA) effective for gait detection [26] and another study used two procedures for the production of gait features: Locality Preserving Projections (LPP) [27] and Local Fisher Discriminant Analysis (LFDA) [28].

Machine learning and deep learning have recently gained popularity due to their increased accuracy and efficiency in health sectors [29-33]. The constraint of machine learning algorithms in being unable to cope with multidimensional data eventually led to the development of

alternate domains. CNN has pulled off recently and performed admirably for areas pattern classification, and computer vision, as well as visualization and visual imagery analysis. C. Szegedy proposed GoogleNet, a large network with 27 layers that operates with max, average pooling, dropout technique, and a softmax classifier [34]. This concept was not explicitly applied to gait detection, but rather was a broad strategy for classification and detection with 43.9% accuracy. Gradually authors in [35] gave an idea of CNN application for detecting periodic human actions and required 5000 epochs to train the CNN architecture. A. Krizhevsky in his study trained a large CNN collected from ImageNet [36] however that was not a direct approach of gait detection and analysis. A. Karpathy offered in his research study [37] a method to minimize the training time for large-scale video categorization. All of these researches have offered distinct applications and implementations of CNN in gait analysis, but none of them have revealed the exact method of detection.

Another recent study [38] focused on gait analysis and detection by predicting human pose and then synthesizing it using a convolutional neural network for classification. Though the study obtained better detection accuracy, the technique comprises two detection processes, which adds to the process's complexity.

However, the goal of this research is to detect gait abnormalities utilizing pressure sensors and various ConvNets based on the output categories of the foot abnormalities. This could be a key technique for replacing typical detection methods proposed previously.

2.2 Comparative Analysis of Relevant Research

Reference	Algorithm	Best Accuracy
[24]	Transfer Learning	90.83%
[25]	RNN Auto encoder	95.9%
[26]	FallsNet (Transfer learning)	86.4%
[27]	CNN	91.9%
[28]	FNN	96%
This Study	Transfer Learning	97.15%

Table 2.1: Comparison between relevant works

Chapter-3 Methodology

3.1 Basic Methodology

The basic working flow of our research is depicted by a flowchart below-

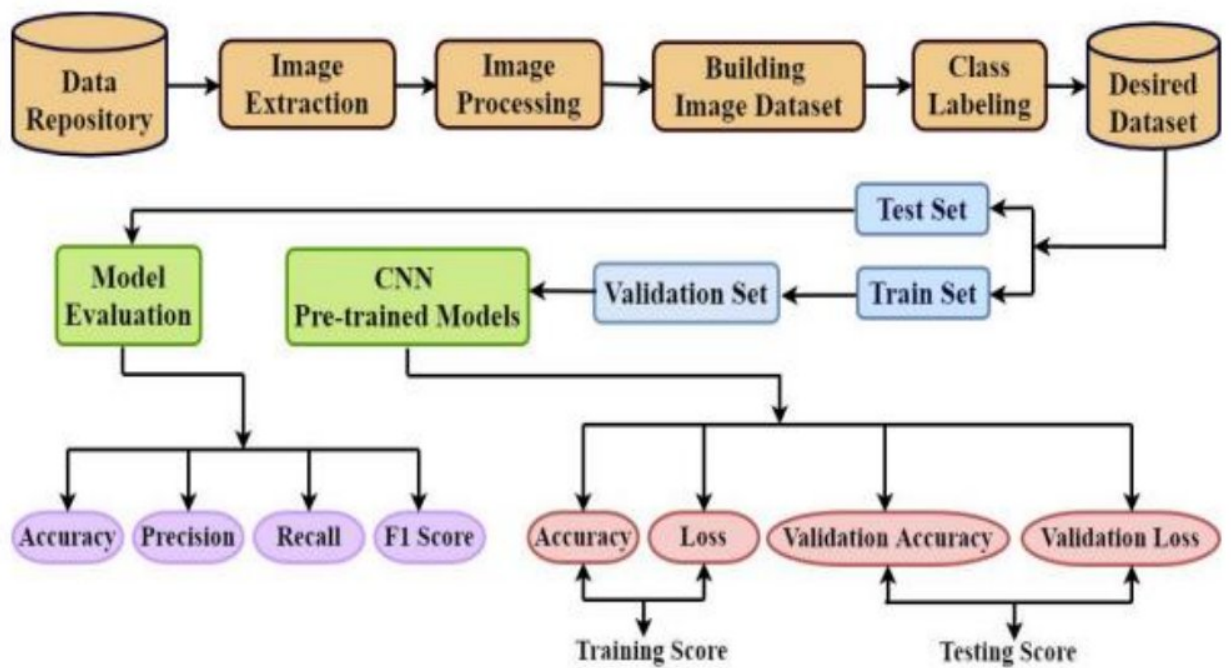


Figure 3.1- Basic Methodology Flow-chart

3.2 Description of Basic Methodology

This section describes the proposed framework to further progress with it to address the gait anomaly recognition problem using CNN based on a transfer learning scheme. Different pre trained CNN models were used to extract features from collected video frames which originated from the collected dataset. MATLAB R2021a software, a superior language for technical computation was implemented for the proposed method. This analysis required a 64-bit Windows PC with Intel ® 2GHzx64 based processor & 8GB RAM.

Initially the dataset was obtained as video files containing different classes of foot abnormalities. Then the individual images were collected from the videos using MATLAB. After the images were collected, they were further processed to obtain the desired dataset. Then the dataset was further expanded and then they were labeled accordingly. Finally, the desired dataset was collected.

The desired dataset was divided into train and test sets. Train set was then generated to CNN pre-trained model where they were divided again to training and validation sets to get the training and validation metrics. Finally after the model was trained by the train set, it was evaluated using the test set. The evaluation was measured using another set of parameters. Finally comparing the results of different CNN models the best model was chosen. The proposed approach is represented in Figure 3.1

3.3 Description of Performance Matrices in ML/DL

Performance Matrices of ML and DL come from the confusion matrices. Confusion matrices give the visualization of the performance of the ML/DL models. Each row of the matrix presents the instances of the actual class, where each column represents the instances of the predicted class or vice-versa. One example of confusion matrix is given below-

Confusion Matrix		Predicted	
		True	False
Actual	True	True Positive (TP)	False Negative (FN)
	False	False Positive (FP)	True Negative (TN)

Here, we can see four terms like True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Their description is given below -

- a) **True Positive (TP):** A class was predicted Positive, which is True actually.
- b) **True Negative (TN):** A class was predicted Negative, which is True actually.
- c) **False Positive (FP):** A class was predicted Positive, which is False actually.
- d) **False Negative (FN):** A class was predicted Negative, which is False actually.

3.4 Formula for Performance Matrices

In order to measure the performances of Machine Learning & Deep Learning models in the research are-

- a) $Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$
- b) $Precision = \frac{TP}{TP + FP}$
- c) $Recall = \frac{TP}{TP + FN}$
- d) $F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall}$

Accuracy

This is one of the parameters used in this paper for the result analysis. As our study is a multiclass classification analysis so balanced accuracy has been taken as the parameter of result determination. It is calculated from the confusion matrix and here the accuracy demonstrates the average accuracy of each nine classes.

Loss

It is simply the measure of error between actual and predicted result that depicts a clear picture of the CNN model. The aim is to get the minimal loss function that expresses how well the model got trained up and how well the algorithm worked on the dataset. The loss function has a close relationship with accuracy while measuring the performance of the neural network.

Confusion Matrix

One of the parameters of result analysis used in this study analysis is confusion matrix. Confusion matrix basically performs the comparison between actual and predicted result in the form of a matrix and the principal diagonal value of the matrix shows the correctly classified elements. Rest of the elements of the matrix demonstrates the mismatched results between actual and predicted values.

Chapter-4

Data Preprocessing

Data preprocessing was vital for this research. The dataset was initially in video format which was then transformed to the desired dataset by the following steps:

1. Image Extraction
2. Image transformation and class expansion
3. Image Augmentation

4.1 Dataset

The initial video dataset for this investigation was obtained from [1], which consisted of six classes, each with five videos that examined and measured human balancing skills using pressure sensors. The classes were:

1. Stand Still
2. Forward lean
3. Backward lean
4. Left lean
5. Right lean
6. Twisting

The aforementioned dataset contains recordings of the foot that capture anomalies in the right or left foot, resulting in a recognized gait approach.

4.2 Image Extraction

To evaluate a complete gait cycle, the video footage was preprocessed in MATLAB to collect the requisite RGB images from every frame, resulting in the collection of every conceivable frame. As a result, gait video data was transformed into an image sequence. As a consequence, the desired dataset was created by selecting 40 photos from a total of 160 for

each film, yielding 200 image samples for each class. Standstill, forward lean, backward lean, left lean, right lean, and twisting were the first six categories in the raw video collection.

4.3 Image transformation and class expansion

Classes were expanded using image samples to obtain a proper dataset for the study analysis as well as to make the dataset relevant to actual disorders:

- 1) Normal Foot
- 2) Left foot forward leaned
- 3) Right foot forward leaned
- 4) Left foot backward leaned
- 5) Right backward leaned
- 6) Left sided lean
- 7) Right sided lean
- 8) Left twist
- 9) Right twist.

Two of the classes are visually represented in terms of foot orientations in Figure 4.1. The classes are left foot backward leaned, left foot forward leaned respectively.



(a)



(b)

Figure 4.1- Visual representation of a) Left Foot Backward Leaned & b) Left Foot Forward Leaned

In Figure 4.2 all the classes are shown respectively.

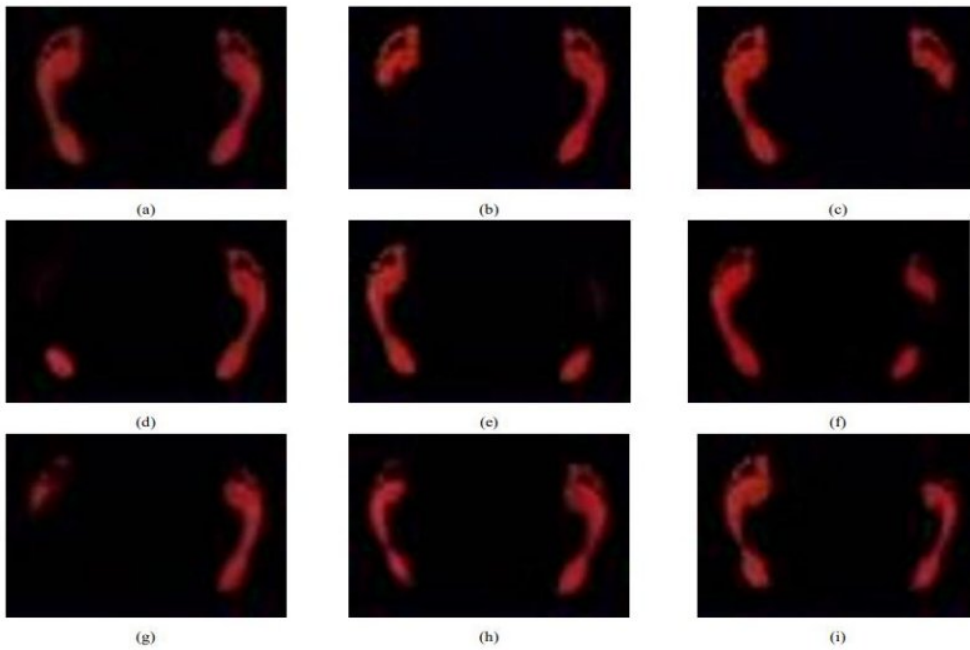


Figure 4.2- Nine classes of gait disorders; (a) Normal foot (b) Left foot forward leaned (c) Right foot forward leaned (d) Left foot backward leaned (e) Right foot backward leaned (f) Left sided lean (g) Right sided lean (h) Left foot twisted (i) Right foot twisted

Number of class	Name of class	Samples before augmentation
1	Normal Foot	200
2	Left foot forward leaned	200
3	Right food forward leaned	200
4	Left foot backward leaned	200
5	Right foot backward leaned	200
6	Left sided lean	200
7	Right sided lean	200
8	Left foot twisted	200
9	Right foot twisted	200

Table 4.1: Number of image samples per class before image augmentation

4.4 Image Augmentation

When a deep learning model is given a large amount of data, it usually performs well. In general, the more data we have, the higher the model's performance. Image augmentation is a method of modifying existing images in order to generate additional data for the model training process. In other words, it is the technique of artificially increasing the dataset available for deep learning model training. We don't need to manually collect these photos because they're all generated from training data. This increases the training sample without having to go out and gather the information. Note that all of the photos will have the same label, which is the label of the original image that was used to create them. Image augmentation is in general of two types. These are:

1. Augmented execution or augmented evaluation
2. Augmented virtuality

1. Augmented execution or augmented evaluation: In this augmentation type the augmentation result is seen physically as after the augmentation the augmented images are stored physically. As a result the dataset size increases significantly.

2. Augmented virtuality: In this augmentation the augmented images are not found physically. In this case the augmentation is performed just before generating the image in the model. As a result the image augmentation is not seen physically and they are not stored afterwards as well.

Image augmentation techniques are versatile and can be done in various ways. Some of the widely used image augmentation techniques are:

- Image Rotation
- Image Shifting
- Image Flipping
- Image Noising
- Image Blurring
- Image Padding
- Image Scaling
- Image Cropping
- Image Translation
- Image Shearing

Image Rotation: Rotates the image into desired angle. The range is 0 degree to 180 degrees in both clockwise and anti-clockwise. The information in the image remains the same before and after the rotation. It's important to keep in mind that image dimensions may not be kept after rotation. If your image is square, rotating it at right angles will keep the image size the same. If it's a rectangle, turning it 180 degrees will keep it the same size. The final image size will alter as the image is rotated at finer angles. In the next section, we'll look at how we may address this problem. Examples of square images rotated at right angles are shown below.

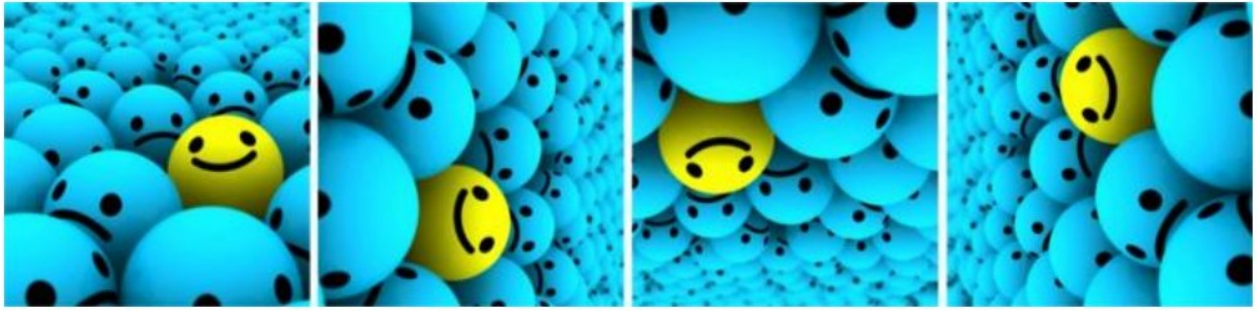


Figure 4.3: Image Rotation

Image Shifting: Image shifting is basically modifying the position of the objects in the image by changing the photos, giving the model greater variation. This could eventually lead to a more generic model. Image shift is a geometric transformation that translates the position of each object in the image to the final output image's new location. So, if an object is at (x,y) in the original image, it will be relocated to (X, Y) in the new image, as illustrated below. The corresponding shifts dx and dy , as well as the different directions, are given.

$$(x,y) \longrightarrow (X,Y)$$

$$X = x + dx$$

$$Y = y + dy$$

Image Flipping: Image flipping is basically the 180 degree rotation or the mirror image of the original image. So it can be stated as an extension of image rotation. There are two types of image flipping:

1. Horizontal Flip
2. Vertical Flip



(a)

(b)

(c)

Figure 4.4: Image Flipping (a) Original Image (b) Horizontal Flip(c) Vertical Flip

Image Noising: Over fitting occurs when your neural network tries to learn high frequency features (patterns that appear frequently) that aren't always effective. Image Noising is a popular image augmentation technique in which random noise is generated in the image with the aim to enable the model to learn how to distinguish between the signal and the noise in an image. Gaussian noise, with zero mean, has data points at all frequencies, substantially distorting high frequency information. Lower frequency components (often, your intended data) are corrupted as well, but your neural network can learn to ignore this. Adding just the appropriate amount of noise can help the model learn more effectively. The salt and pepper noise, which appears as random black and white pixels scattered throughout the image, is a toned-down variation of this. This effect is similar to that of adding Gaussian noise to an image, however it may have less information distortion. This technique also enhances the ability of resisting image alterations.



Figure 4.5: Image Noising

Image Blurring: Images that are extracted can be from different sources which will result in varying quality of images. As a result the images with low quality will hinder the process of feature extraction. To tackle this problem image blurring is applied to help the model to be resilient to low quality images as well. This will help the model to learn features from images of varying qualities without affecting their performance.



Figure 4.6: Image Blurring

Image Padding: A certain CNN model needs to take images of the same dimensions as input. But due to varying sources the dimensions of the image changes as well. So by transforming the images to the same size, a large amount of information might be lost. This is why image padding is important. Popular padding techniques are zero padding, nearest mode padding, etc. The image is padded on all edges with a provided value in padding.

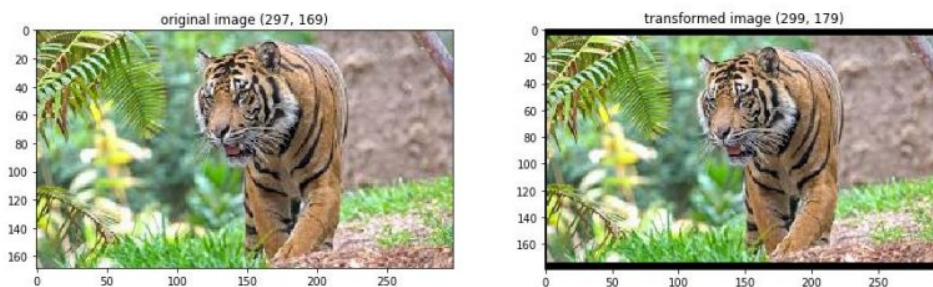


Figure 4.7: Image Padding

Image Scaling: Image scaling or resizing is vital for any model because a certain CNN model can only take images of the same dimensions as input. In image scaling the image is resized to the specified size, for example, the image's width can be doubled. Scaling the image outward or inward is possible. The final image size will be greater than the original image size after scaling outward. Most image frameworks extract a piece of the new image that is the same size as the old. The next lesson will include scaling inward, which reduces the image size and forces us to make assumptions about what lies beyond the boundary. Below are some instances of scaled images.

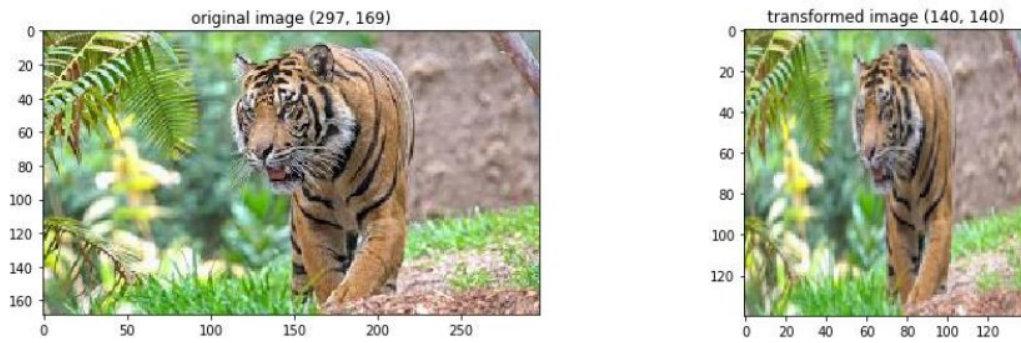


Figure 4.8: Image Scaling

Image Cropping: In cropping, a portion of the image is selected e.g. in the given example the center cropped image is returned. Unlike scaling, we just sample a piece of the original image at random. This part is then resized to the original image size. Random cropping is the name given to this technique.

Image Translation: The image is simply moved in the X or Y direction during translation (or both). We presume that the image has a black background beyond its limit in the following example, and we translate it correctly. Because most things can be found practically everywhere in the image, this type of augmentation is extremely beneficial. This makes your convolutional neural network search in all directions.



(a)

(b)

(c)

Figure 4.9: Image translation (a) Original Image (b) Translated to the right (c) Translated to the left

Image Shearing: Image shearing is a bounding box transformation we change the rectangular image into a parallelogrammed image by shearing it. This is the shearing transformation matrix.

$$\begin{vmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \end{vmatrix}$$

A horizontal shear is illustrated above. The pixel with coordinates x, y is moved to $x + \alpha \cdot y, y$ in this case, where α is the shearing factor.

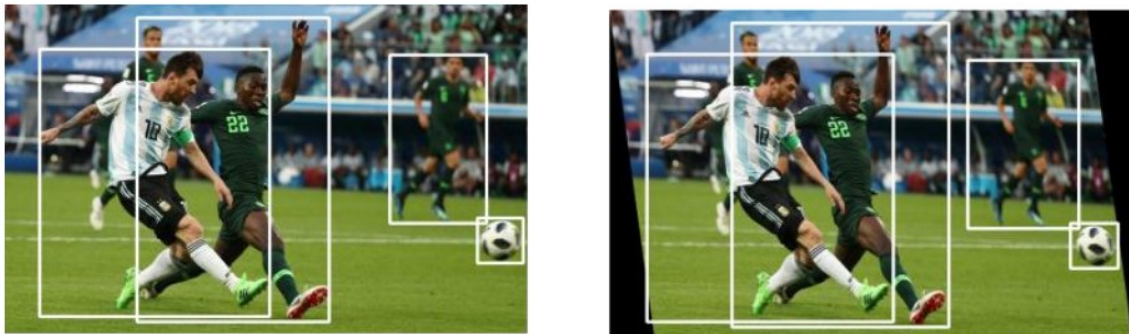


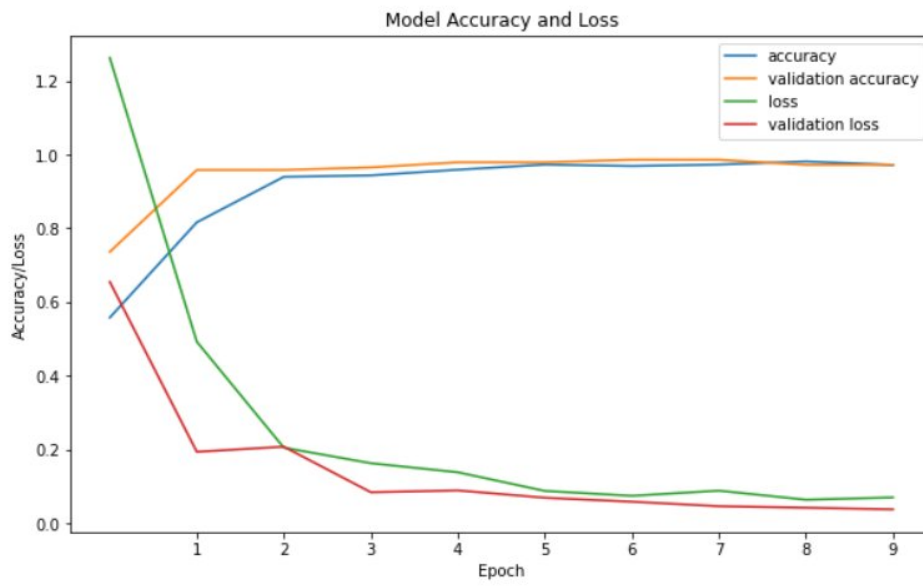
Figure 4.10: Image Shearing

The advantages of image augmentation are:

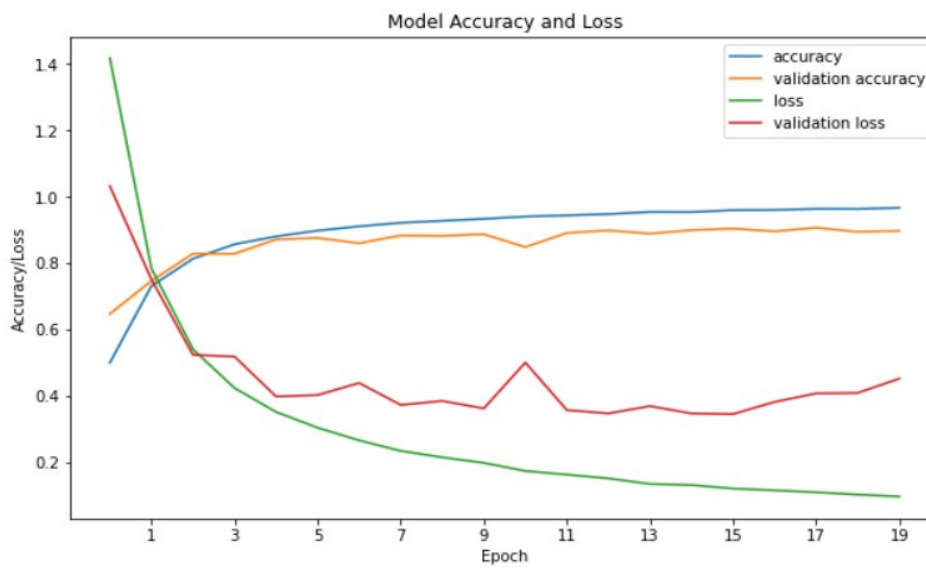
- It reduces the cost of collection of data.
- It reduces the cost of labeling data.
- It improves the model prediction accuracy.
- It prevents data scarcity.
- It frames better data models.
- It reduces data overfitting.
- It creates variability and flexibility in data models.

As stated above, image augmentation was significant in this study as well. Initially the model was trained without augmenting the dataset. The results were extremely good which could also be the effect of the model being overfit. That's why image augmentation was introduced and the initial findings after augmentation proved the fact of overfitting. Figure 4.11 shows

the comparison between the model before and after image augmentation.



a. Before augmentation



b. After augmentation

Figure 4.11(a) Initial Training Accuracy/Loss before augmentation (b) Initial Training Accuracy/Loss after augmentation

Number of class	Name of class	Samples before augmentation	Samples after augmentation
1	Normal Foot	200	3470
2	Left foot forward leaned	200	3435
3	Right food forward leaned	200	3415
4	Left foot backward leaned	200	3414
5	Right foot backward leaned	200	3444
6	Left sided lean	200	3417
7	Right sided lean	200	3452
8	Left foot twisted	200	3384
9	Right foot twisted	200	3448

Table 4.2: Number of image samples per class after image augmentation

So, image augmentation was used with the goal of making the dataset more versatile, increasing the number of samples, and proving the authenticity of the analysis, as the image dataset originally contained foot heatmaps of a single person, indicating the data's simplicity by having heatmaps of similar size, rotation, and shape. The table below shows the number of samples that are in each class after augmentation. Augmentation techniques used for the analysis are : (a) rotation (b) width shift (c) height shift (d) zoom range (e) fill mode = “nearest”

The dataset was split and distributed beforehand into training and testing sets, and then augmentation was done independently to each set to ensure that the test set did not contain any patterns that were comparable to the training sets, making the test set wholly unfamiliar to the classifier. The target of this study is to implement ConvNet classifiers that work well on generalized data of various people & will not overfit.

Chapter-5

Introduction to Algorithms

5.1. Convolutional Neural Network:

Convolutional Neural Network(CNN) based classifiers have been effectively working and can be trained to perform in various classification problems, recognition and visualization fields. CNN or convolutional neural network is a modified deep neural network used for image or other multidimensional data analysis and classification. Since the analysis of this study was done with images so the dataset is 2 dimensional that resulted in the training of the CNN classifier for further analysis. The incompetence of the traditional machine learning methods on 2 dimensional data led to the approach with CNN as it performs efficiently with 2 dimensional data. A CNN recognizes patterns in images and analyzes visual imagery whereas trains the patterns into more complex shapes in later stages. In case of pattern recognition or for any multidimensional data the implementation of CNN has been done in many recent past research works for its prospects.

CNN has the following components:

A CNN model has hidden layers called convolutional layers which detects simple patterns initially and is trained afterwards. These layers, also called **kernels**, are convolved through the image for pattern recognition. Convolution extracts particular 'features' from an input picture using a 'kernel.' A kernel is a matrix that is slid over the picture and multiplied with the input in order to improve the output in a desired way.

There are other non-convolutional layers such as **Max Pooling layers**, fully connected layers which are called **dense layers**, **flattening layers**, **padding layers** used for other purposes. For this study the CNN is composed of 2 convolutional layers, 2 fully connected dense layers. The size of the convolutional filter is 3*3. The activation function used after both of the convolutional layers is **ReLU**. Max pooling layers have been added after convolutional layers to decrease the dimensions and collect important features from the filters. The final dense layer is used for the output classification. The optimizer used is **Adam** and learning rate= .001 after some trials and errors.

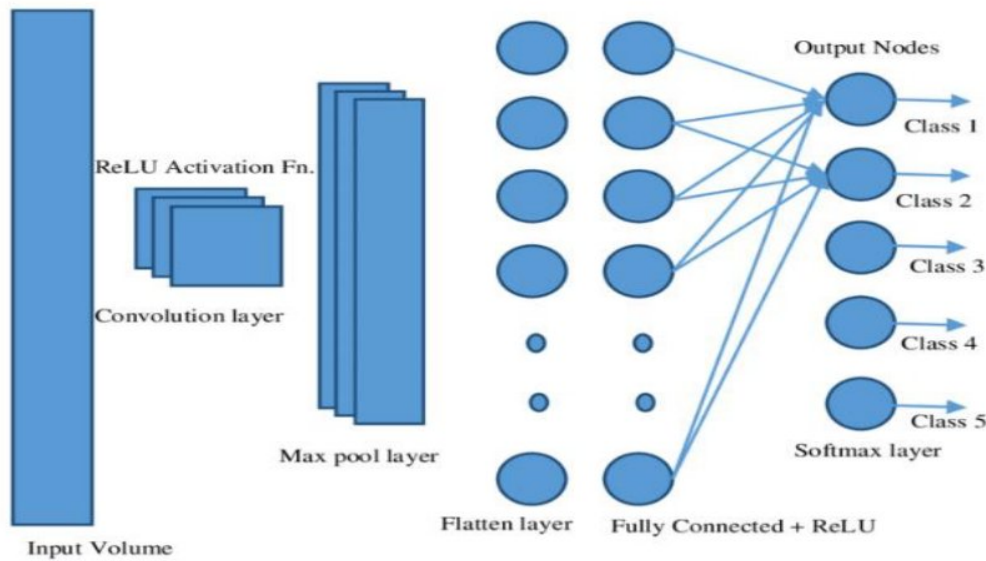


Figure 5.1: CNN architecture

Adaptive Moment Estimation is an algorithm for optimization technique for gradient descent. The method is really efficient when working with large problems involving a lot of data or parameters. It requires less memory and is efficient. Intuitively, it is a combination of the ‘gradient descent with momentum’ algorithm and the ‘RMS’ algorithm.

After each convolutional layer and fully connected layer, the activation function for faster training is used. Although the sigmoid activation function has historically produced outstanding results for neural networks, it has limitations such as sluggish convergence and vanishing gradient difficulties. Because it can offer effective training networks and solve the issue of vanishing gradients, **ReLU** has become prominent in deep learning.

In multi-class classification problems, categorical cross entropy is a loss function. These are problems in which an example may only fit into one of many categories, and the model must choose which one. With the categorical cross entropy loss function, **Softmax** is the sole activation function that is suggested. Although this is not true for tiny values, the softmax function highlights the greatest values and suppresses those that are far below the maximum value. It sums the outputs to make them immediately treatable as probabilities over the output.

It's frequently utilized as the loss function in the last layer of a classifier model with categorical cross entropy.

Let's define Gradient Descent first before moving on to Stochastic Gradient Descent (SGD). Gradient Descent is a widely used optimization approach in Machine Learning and Deep Learning, and it can be used in nearly all learning algorithms. A function's gradient is its slope. It determines how much a variable changes in reaction to changes in another variable.

Gradient Descent is a mathematically defined convex function whose output is the partial derivative of a collection of input parameters. The higher the slope, the greater the gradient. The term "stochastic" refers to a system or process with a random probability. As a result, instead of selecting the whole data set for each iteration in **Stochastic Gradient Descent**, a few samples are chosen at random. The term "batch" is used in Gradient Descent to refer to the total number of samples from a dataset that are utilized to calculate the gradient for each iteration. The batch is considered to be the whole dataset in standard Gradient Descent optimization, such as Batch Gradient Descent. Although considering the entire dataset is quite valuable for reaching the minima in a less noisy and random manner, the challenge emerges as our datasets become large.

5.2 Transfer Learning:

Transfer learning, a ML scheme, is now widely used in deep learning networks, and it works by accumulating information that has been proficient in case of massive sets of data in a specific field and then transferring that information to a smaller dataset in a similar field to boost and upgrade performance.

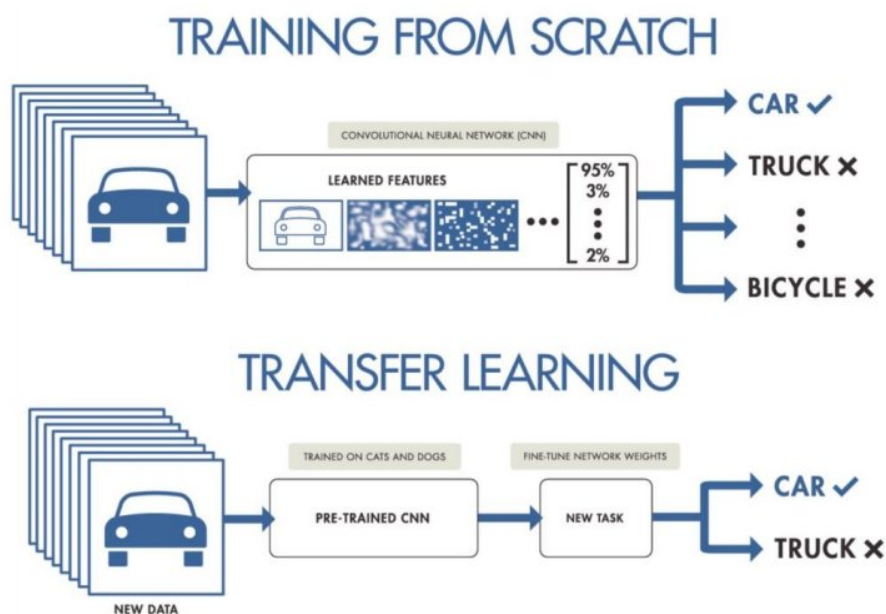


Figure 5.2: Transfer Learning

The training of CNN classifiers can be divided into two categories: training from scratch and training utilizing a transfer learning method. In a traditional CNN model, the kernels interpret simple patterns in images on the first few layers, while later stages train the patterns into more complicated structures. The training procedure is inconvenient, time-consuming, and frequently fails to deliver sufficient results. As a result, pretrained models are preferable because their base or the weights of kernels are already trained, leaving only the dense layers to be tweaked.

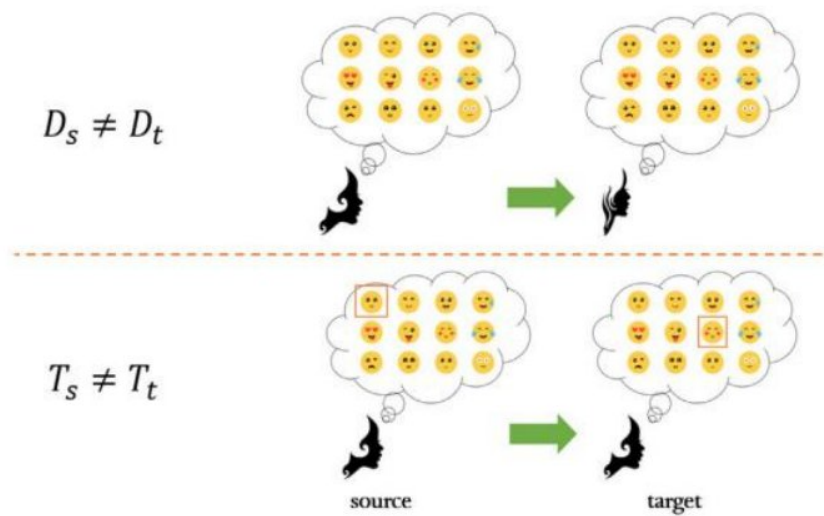


Figure 5.3: Examples of Difference in Transfer Learning

During transfer learning, the information of a previously trained machine learning model is applied to a new but closely related issue. If you trained a basic classifier to predict if a picture has a backpack, for example, you might utilize the model's training experience to identify additional things like sunglasses.

Neural networks are used in computer vision to identify edges in the first layer, shapes in the intermediate layer, and task-specific properties in the latter layers. Transfer learning uses the early and central layers, whereas the subsequent layers are just retrained. It uses the labeled data from the task on which it was trained.

5.2.1 Approaches to Transfer Learning

- **TRAINING A MODEL TO REUSE IT**

Consider the situation where there is a need to perform Task A but don't have enough data to train a deep neural network. Finding a similar task B with a lot of data is one way around this. One can utilize the deep neural network to train on task B and then use the model to solve problem A. The problem that is seeking to be addressed will determine whether there is a need to employ the entire model or just a few levels.

If they have the same input in both tasks, possibly reusing the model and making predictions for the new input is an option. Alternatively, changing and retraining different task-specific layers and the output layer is a method to explore.

- **USING A PRE-TRAINED MODEL**

The second option is to employ a model that has already been trained. There are a number of these models available, so do some research beforehand. The number of layers to reuse and retrain depends on the task.

Keras, for example, provides nine pre-trained models that can be used for transfer learning, prediction, feature extraction and fine-tuning. One can find these models, and also some brief tutorials on how to use them, here. There are also many research institutions that release trained models.

This type of transfer learning is most commonly used throughout deep learning.

- **FEATURE EXTRACTION**

Another option is to utilize deep learning to determine the optimal representation of the problem, which entails identifying the key features. This method is known as representation learning, and it may frequently produce significantly better results than hand-designed representation.

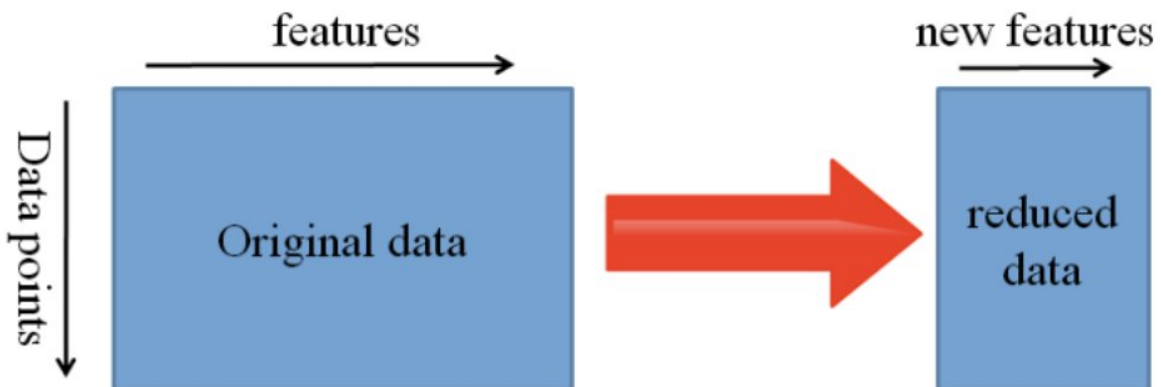


Figure 5.4: Transfer Learning Overview

Domain: The feature space X of n dimensions and the probability distribution $P(X)$ of X , where $X = \{x_1, x_2, \dots, x_n\} \in X$ make up a domain D . In transfer learning, the source domain contains known information and is generally represented by D_S , whereas the target domain contains unknown knowledge and is usually represented by D_T .

Task: A task is a learning goal that consists of the label space Y and the prediction function $f(\cdot)$, where $Y = \{y_1, y_2, \dots, y_n\} \in Y$. The label space of the source domain and the target domain are represented as Y_S and Y_T , respectively, according to the task description.

Transfer learning: When there is a disparity across domains or tasks, transfer learning may be used to transfer information from the source domain to the target domain. It's worth emphasizing that the domains or tasks must be different but comparable to some level for transfer learning to work. The information gained is frequently identified as Y in certain issues. Data in this domain, which is represented as $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, is made up of labels and their accompanying characteristics.

5.2.2 VGG16:

VGG16 is one of the most popular CNN architectures. This classification task consists of nine phases, while the initial training of VGG16 is focused on the classification of 1000 categories. As a result, the 16th or last dense layer must be replaced with a dense layer comprising seven neurons. The network contains 16 convolution layers and a receptive field of 3×3 which is referred to as the kernel. It has a max pooling layer of size 2×2 and 5 such layers in total.

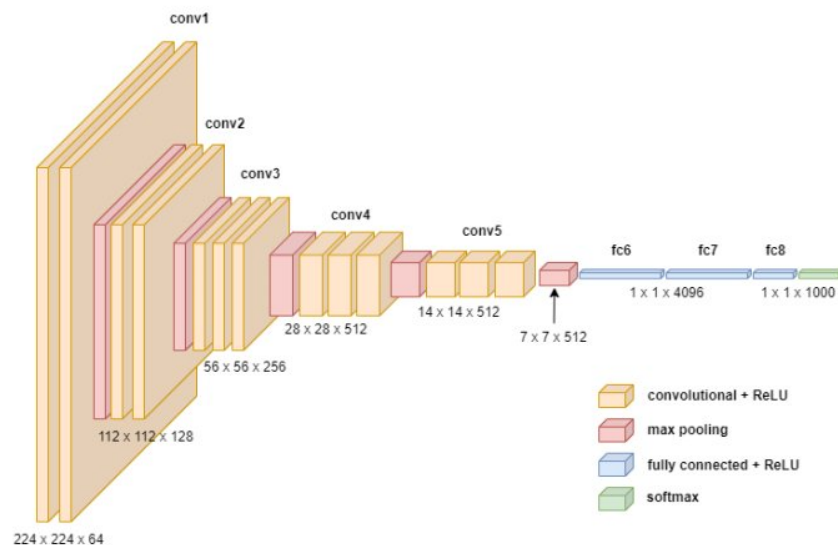


Figure 5.5: VGG16 architecture

However, the VGG16 model has several limitations in terms of applicability. First, the fully connected layer includes a large number of parameters that take up a lot of memory and processing resources, making the VGG16 model difficult to deploy in the front end. Second, the network model is single-structured, and its performance pales in comparison to certain complex advanced networks. Furthermore, VGG16 lacks an effective strategy to prevent

gradients from disappearing, and issues like sluggish convergence speed and gradient explosion are likely to arise during model training.

5.2.3 ResNet50:

ResNet50 is a 50-layer residual network in abbreviated form. The network's picture input size is 224×224 pixels. Resnet50 is similar to VGG-16, but it has the added feature of identity mapping. By allowing gradients to flow through this additional shortcut path, ResNet alleviates the problem of vanishing gradients. ResNet's identity mapping allows the model to skip a CNN weight layer if the current layer is not required. This helps to avoid the problem of overfitting the training set. ResNet aims at finishing challenging errands as well as increases exposure correctness. ResNet aids in discouraging the trials of deep CNN.

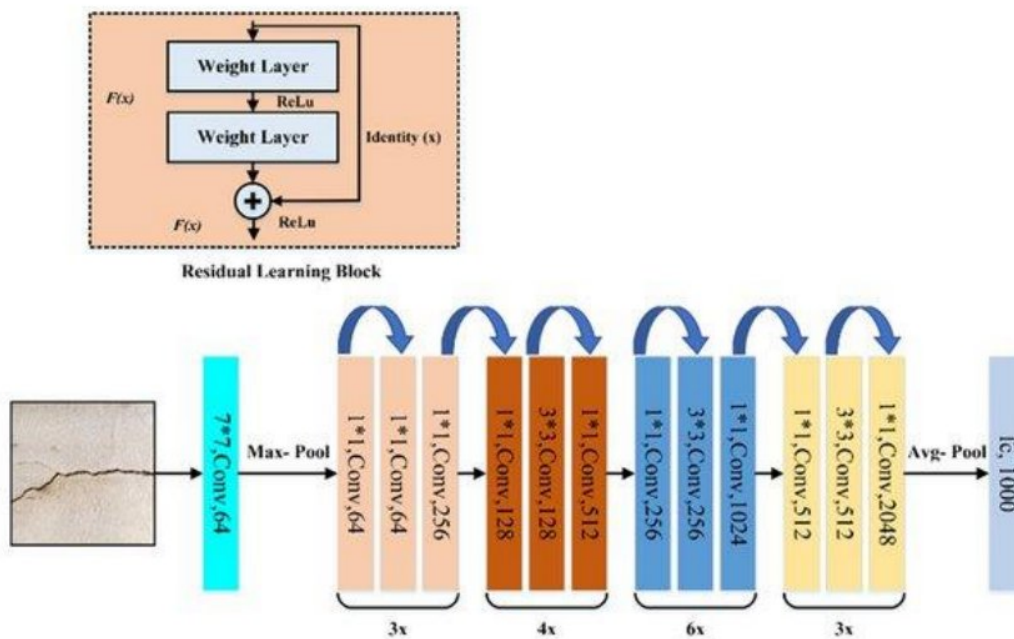


Figure 5.6: ResNet architecture

The delta necessary to achieve the final forecast from one layer to the next is predicted by ResNet. By enabling a second shortcut way for gradient to flow via, ResNet eliminates the vanishing gradient issue. ResNet's identity mapping allows the model to skip a CNN weight layer if it isn't required. This helps to prevent the problem of overfitting the training set.

5.2.4 MobileNet_V2:

The scientists chose the MobileNet V2 model because it is competent as well as beneficial. The key benefit of adopting the MobileNet architecture is that it requires significantly less computational effort than a traditional CNN model, making it ideal for use on mobile devices and computers with limited computational capabilities. Instead of employing standard 3x3 convolution filters, MobileNet splits the operation into depth-wise separable 3x3 convolution filters followed by 1x1 convolution. Abstraction layers are developed for in-depth platforms, which point through a customary, rectified linear unit (ReLU). The resolution multiplier variable is used to reduce the input image's dimensionality as well as the internal representation.

5.2.5 DenseNet

Transfer learning has been utilized aimed at the DenseNet outline to improve the recital of models. Contrary to common assumption, DenseNets require less constraint than standard CNNs because of not requiring knowledge of unnecessary feature maps. DenseNet's key notion is feature reuse, which leads to highly compact versions. Because no feature-maps are repeated, it requires fewer parameters than other CNNs. In DenseNet, each layer takes extra input from the previous levels and provides that to the next. The idea of concatenation is employed to ensure that every layer obtains united knowledge from the above levels. The DenseNet-201 architecture is used in this study analysis.

Chapter-6

Result & Analysis

6.1 Implementation of Transfer Learning Algorithms

To classify nine kinds of foot anomalies, six transfer learning techniques were used. They were first taught on the train set. The model was optimized using Adam optimizer during training. As it was a multiclass classification problem the mode that was used is called categorical cross-entropy. The research was conducted in a Python environment with the TensorFlow framework. The CNN architectures were derived from pre-existing Keras models. ResNet50, VGG16, MobileNetV2, and DenseNet201 were the functions employed in the study. The ImageNet dataset was used to pre-train these models. During the training phase the number of epochs were varied to obtain the feasible result in terms of time and performance. The epochs were kept to 20,30,40 respectively. By comparing them, it was concluded that 30 epoch has the most feasible result. After completing the train phase, the models were further evaluated by the test set. Confusion matrix was obtained for each of the models from which the other parameters such as accuracy, precision, recall, f1 score, auc score were derived. Result analysis for each model and then a comparative analysis between them is given below.

6.1.1 VGG16

Training Phase:

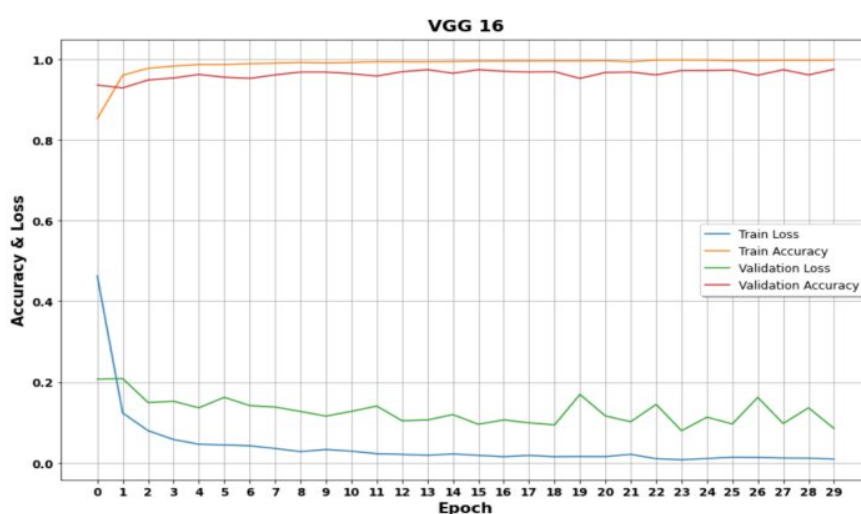


Figure 6.1: Accuracy & Loss during training for VGG 16

Max Validation Accuracy: 97.45%
 Mean Validation Accuracy: 96.23%
 Minimum Validation Loss: 7.9%
 Mean Validation Loss: 12.75%

Test Phase:

From the test phase, confusion matrix is obtained which is shown in Figure 6.2

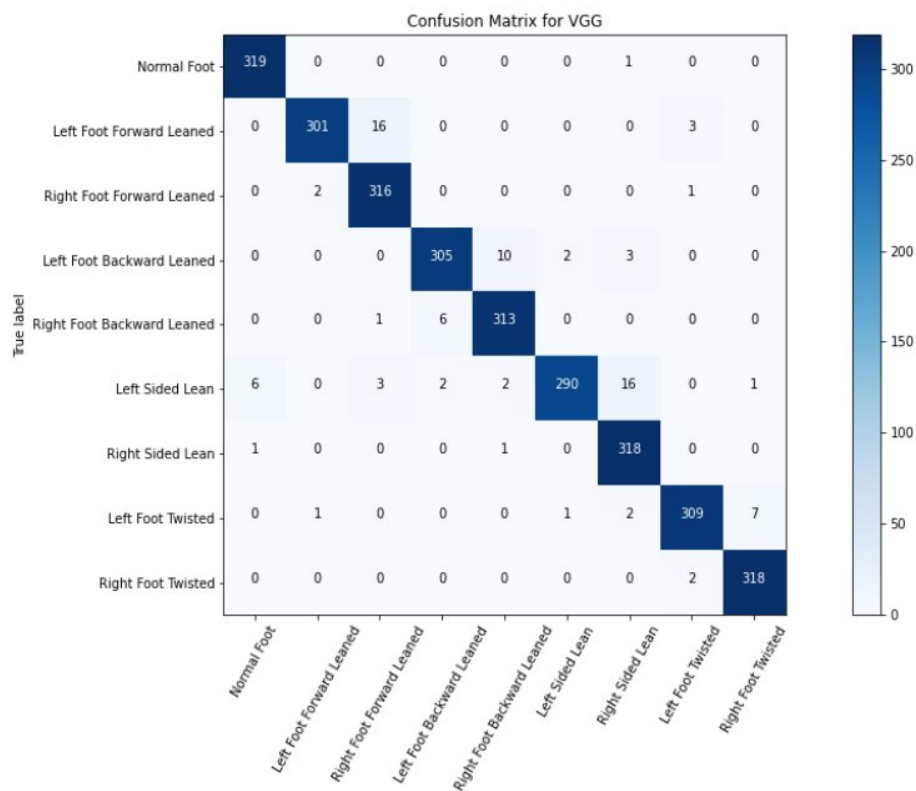


Figure 6.2: Confusion Matrix of VGG 16

6.1.2 ResNet50

Training Phase:

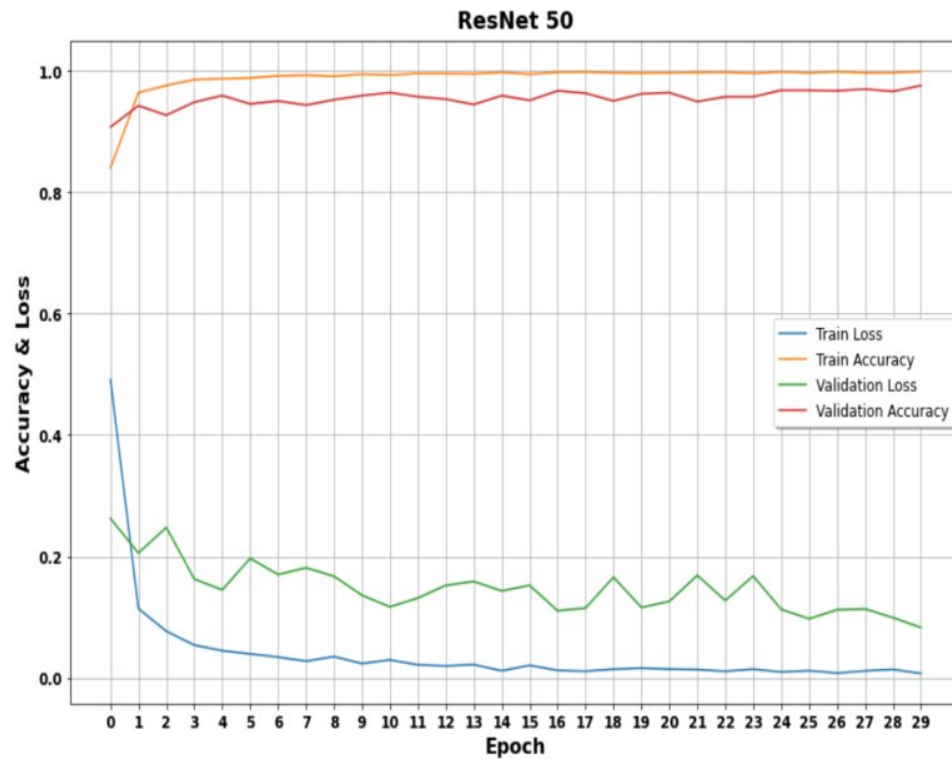


Figure 6.3: Accuracy & Loss during training for ResNet 50

Max Validation Accuracy: 97.55%
Mean Validation Accuracy: 95.47%
Minimum Validation Loss: 8.28%
Mean Validation Loss: 14.81%

Test Phase:

From the test phase , confusion matrix is obtained which is shown in Figure 6.4

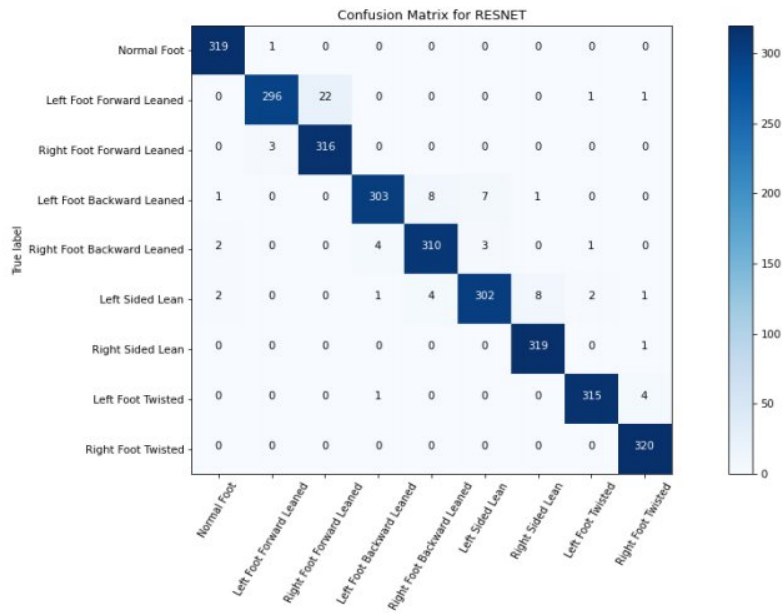


Figure 6.4: Confusion Matrix of ResNet50

6.1.3 MobileNet_V2

Training Phase:

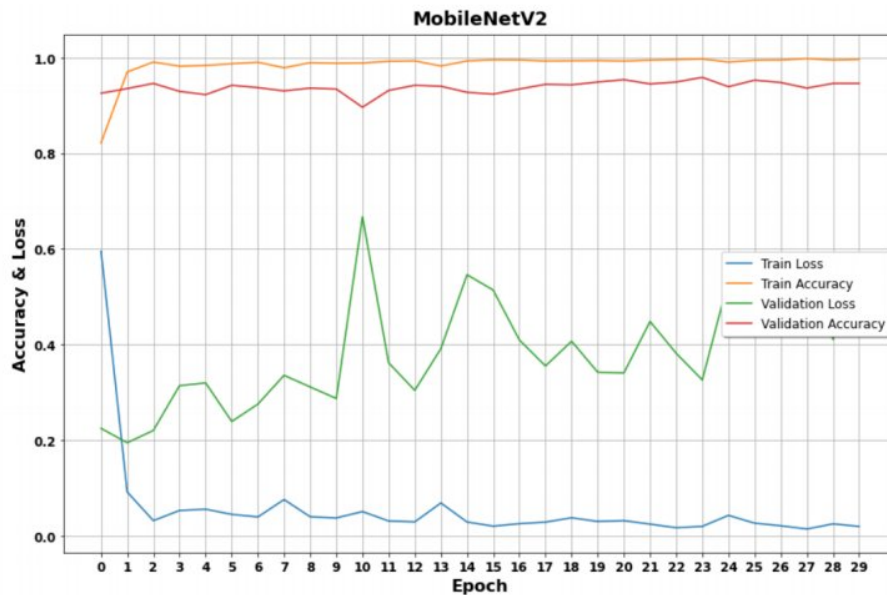


Figure 6.5: Accuracy & Loss during training for MobileNetV2

Max Validation Accuracy: 95.89%
 Mean Validation Accuracy: 93.83%
 Minimum Validation Loss: 19.47%
 Mean Validation Loss: 38.17%

Test Phase:

From the test phase, confusion matrix is obtained which is shown in Figure 6.6

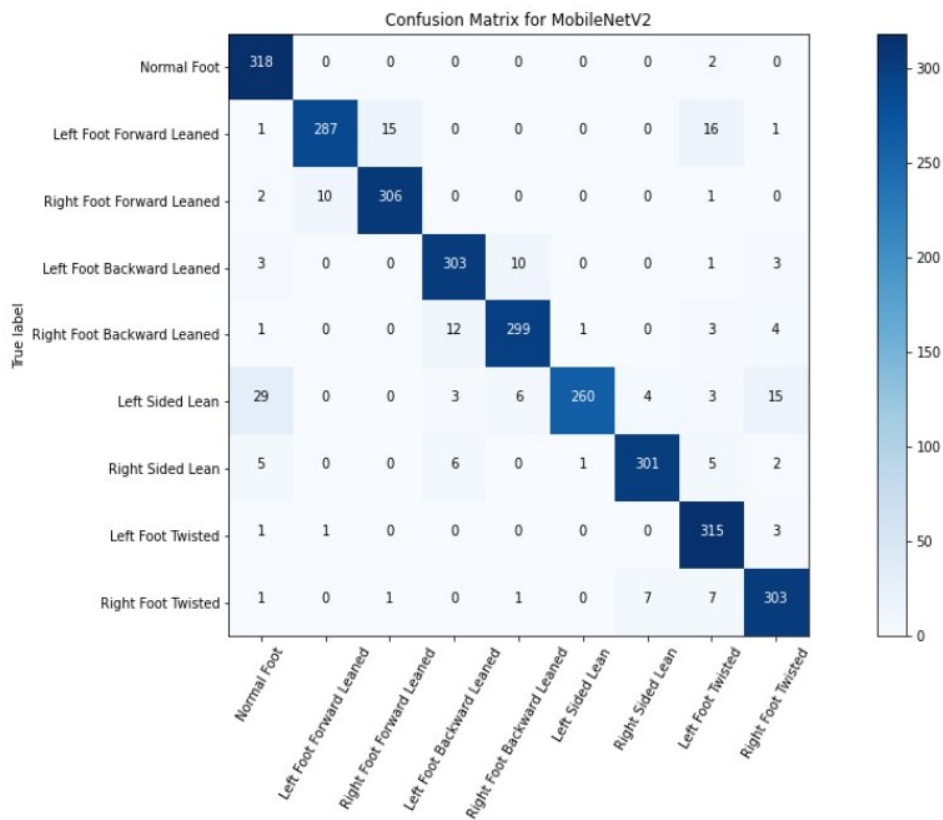


Figure 6.6: Confusion Matrix of MobileNet_V2

6.1.4 DenseNet

Training Phase:

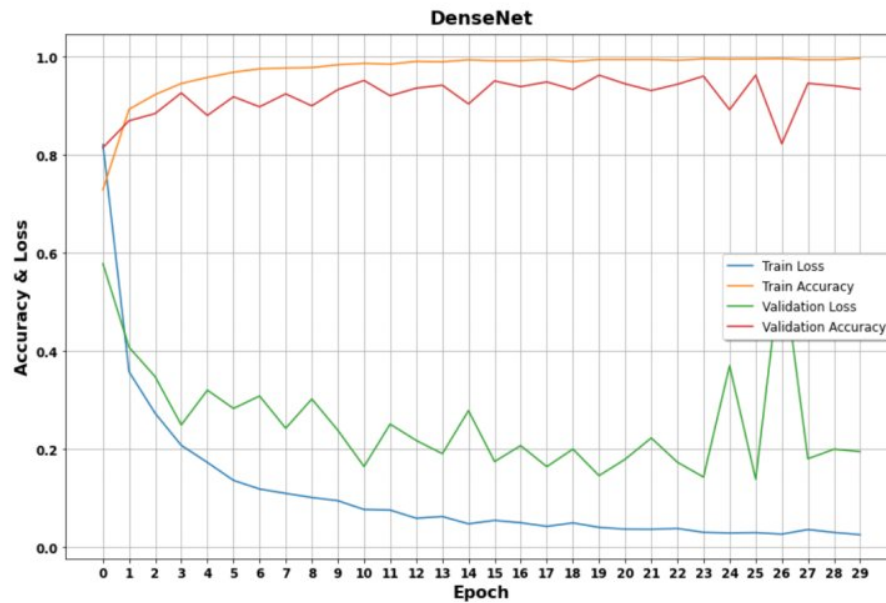


Figure 6.7: Accuracy & Loss during training for DenseNet

Max Validation Accuracy: 96.18%
Mean Validation Accuracy: 91.99%
Minimum Validation Loss: 25.47%
Mean Validation Loss: 13.82%

Test Phase:

From the test phase, confusion matrix is obtained which is shown in Figure 6.8

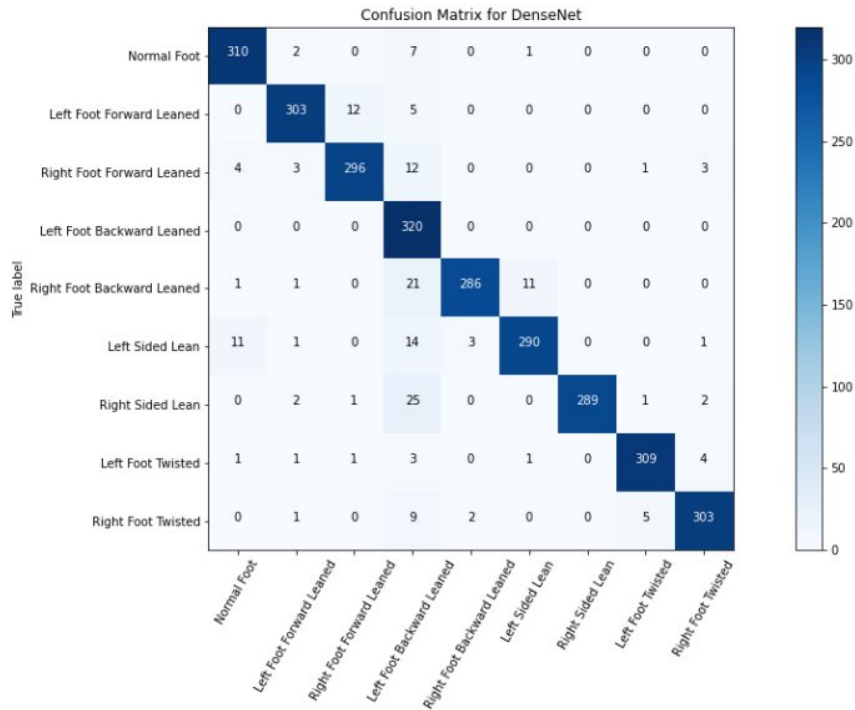


Figure 6.8: Confusion Matrix of DenseNet

6.2 Comparative Analysis of the Transfer Learning Algorithms

6.2.1 Training Phase(20 Epochs)

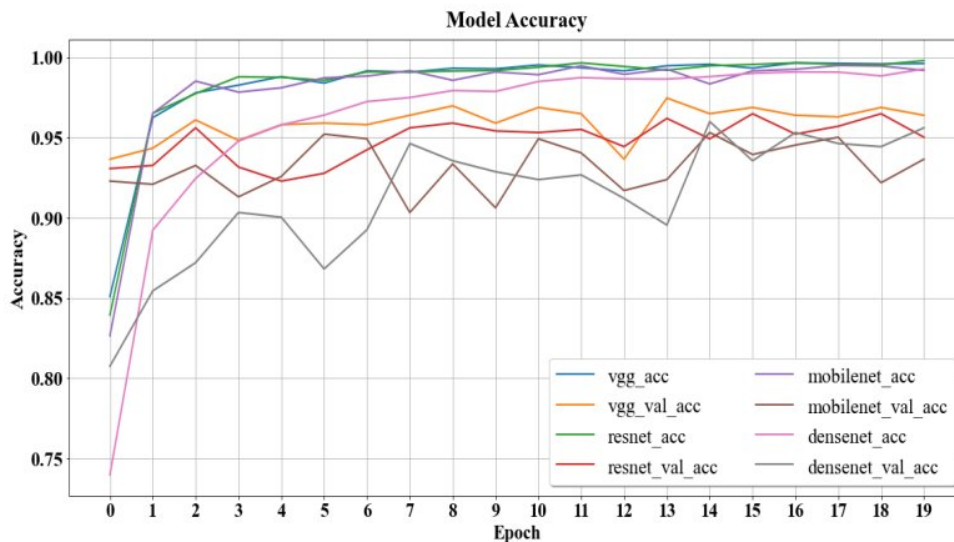


Figure 6.9: Training and validation Accuracy of the four models at 20 epochs

Max Validation Accuracy: 96.45%(VGG16)

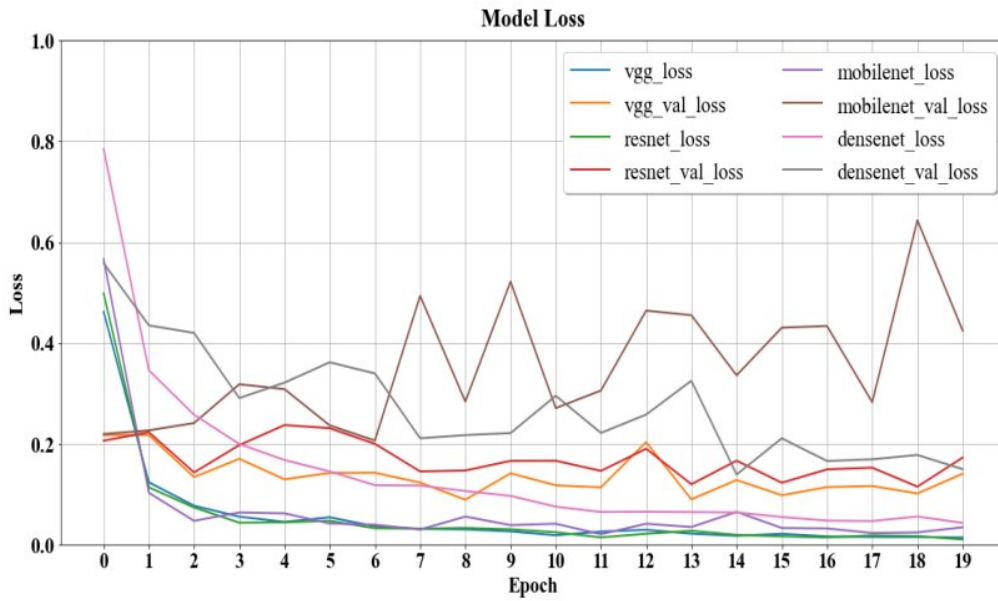


Figure 6.10: Training and validation Loss of the four models at 20 epochs

Minimum Loss: 8.9%(VGG16)

6.2.2 Training Phase(30 Epochs)

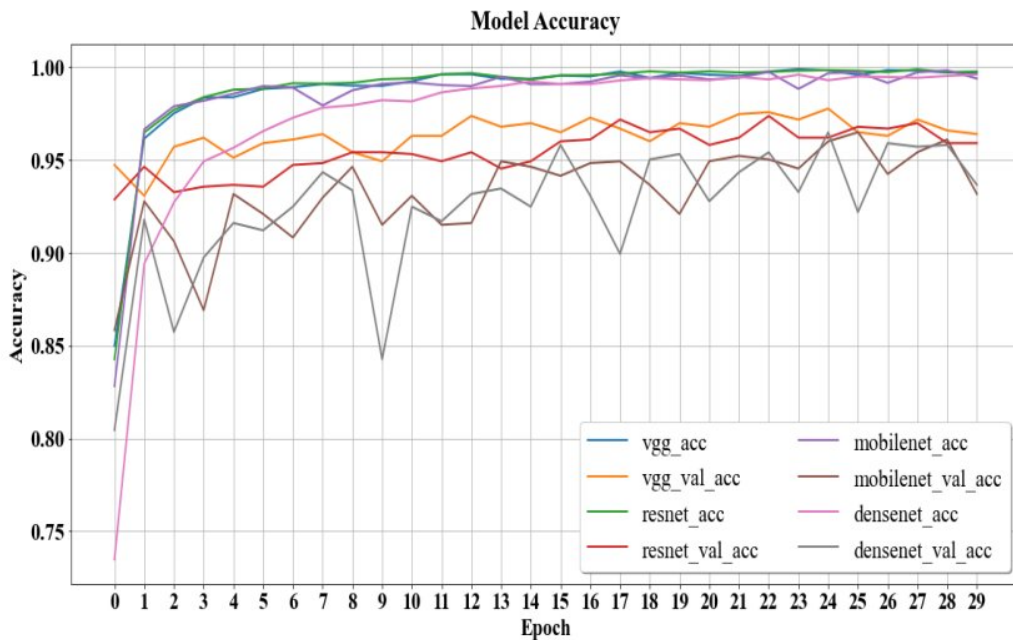


Figure 6.11: Training and validation Accuracy of the four models at 30 epochs

Max Accuracy: 97.75%(VGG16)

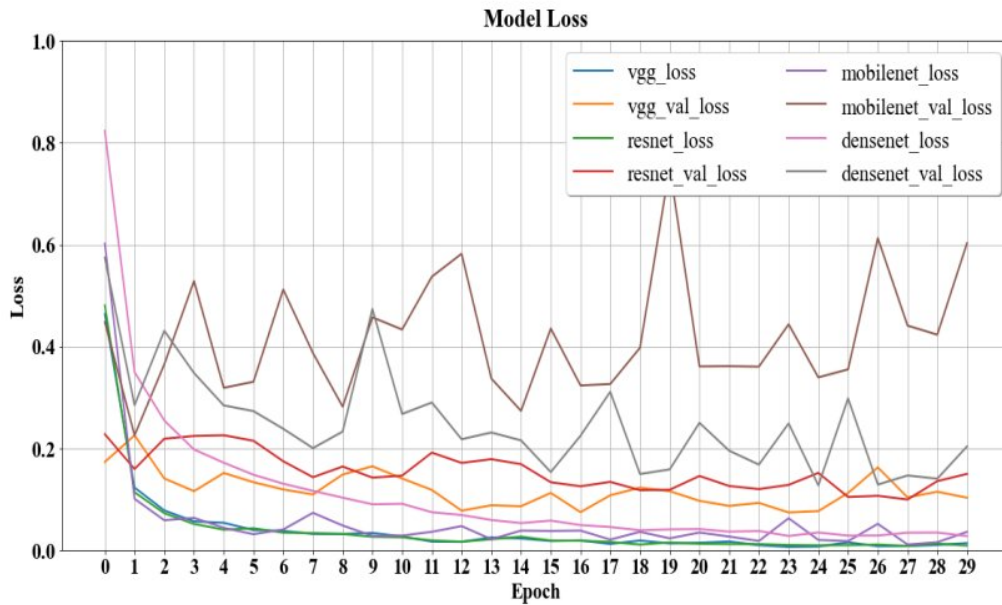


Figure 6.12: Training and validation Loss of the four models at 30 epochs

Minimum Loss: 7.3%(VGG16)

6.2.3 Training Phase(40 Epochs)

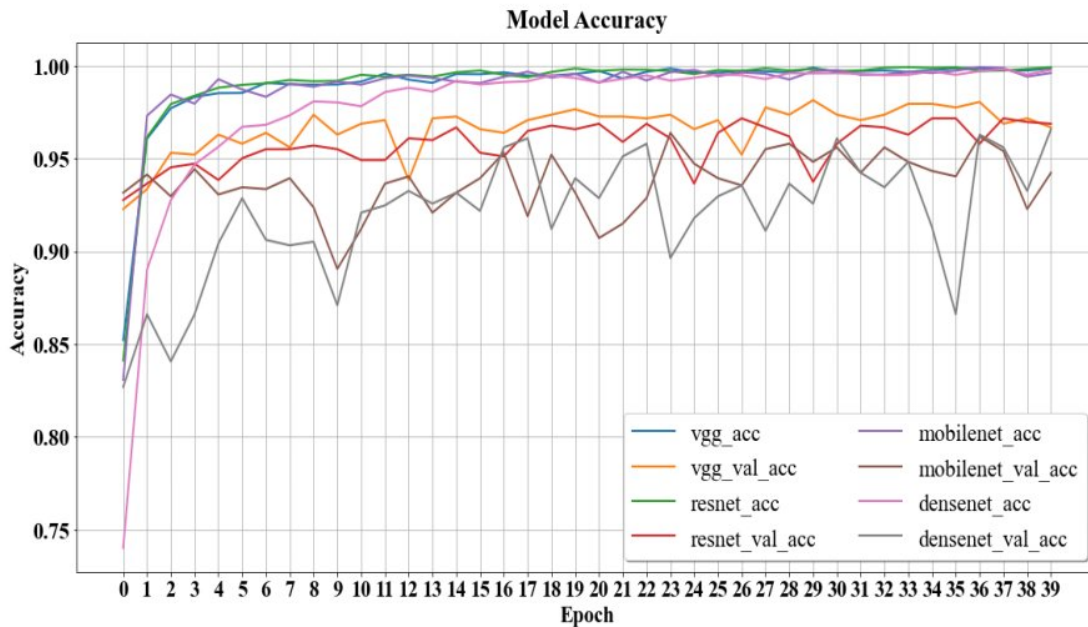


Figure 6.13: Training and validation Accuracy of the four models at 40 epochs

Max Accuracy: 98%

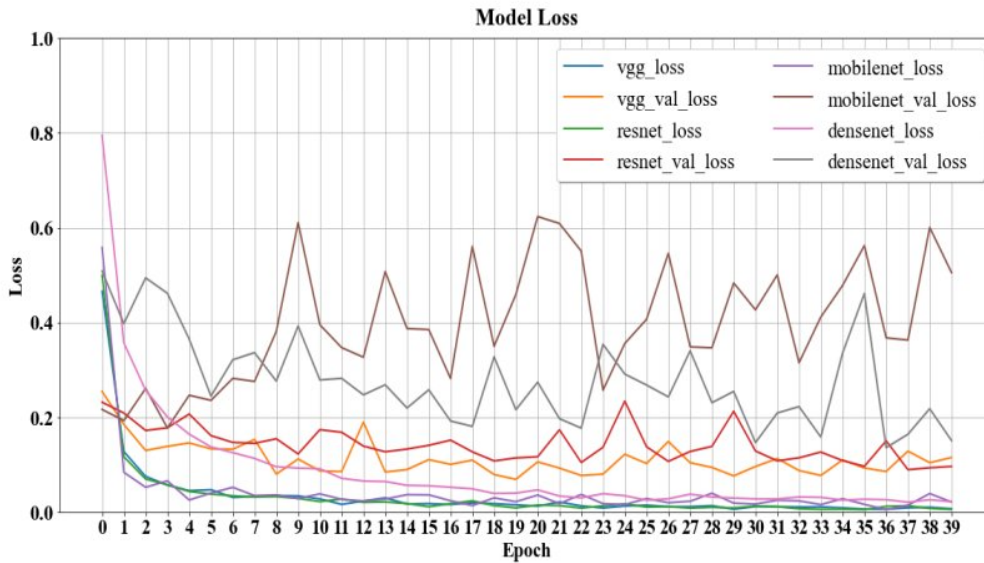


Figure 6.14: Training and validation Loss of the four models at 40 epochs

Minimum Loss: 7%

6.2.4 Test Phase

Performance Parameters	VGG16	ResNet50	MobileNet_V2	DenseNet
Accuracy	97.15	96.15	94.16	94.82
Precision	97.15	96.15	94.16	94.82
Recall	97.21	96.38	94.38	95.10
F1 Score	97.14	96.22	94.15	94.85
ROC-AUC	99.89	99.73	99.77	99.69
Log Loss	11.22	21.65	48.20	44.20

Table 6.1: Results

The best feasible result for each model was reached by varying the number of epochs and trainable layers. Every model's accuracy and loss were investigated with 20, 30 and 40 epochs, resulting in improved accuracy the dataset was used to train all of the models for nine different classes. The image dataset procured from data preprocessing is generated and labeled. Initially, the train set is sent to the Convolutional Neural Network for training the

model. Training different models generated their corresponding performance parameters. These parameters were then further reevaluated by the test set. The figures above show the training metrics at each iteration where each iteration or epoch is an estimation and update of the network parameters. From the figures it's evident that the training process of Vgg16 and ResNet50 were consistent as there were close to no spikes indicating that the models were adequately trained, which was reflected in the model's performance on the test data.

After 30 epochs, the validation accuracy of Vgg16 and ResNet50 outperformed MobileNet_V2 and DenseNet, with a max and mean value of 97.75% and 96.33% for Vgg16 and 97.36% and 95.44% for ResNet50, respectively. Despite the fact that the maximum validation accuracy of MobileNet_V2 and Densenet performed likewise, the mean of the accuracies differed somewhat, with values of 93.25% and 92.42%, respectively. This discrepancy of MobileNet_V2 and DenseNet over epochs is visible in the spikes on specific phases, which were eventually mimicked in the test results. The proposed analysis divides gait irregularities into nine categories, with eight of them containing distinct gait or foot anomalies and the remaining one being normal gait or foot. The performance scores obtained from assessing entire models on the test data are shown in the table above. Among the results obtained in the testing session for the better synthesis and analysis the performance of the models is depicted in Table 6.1.

Confusion matrix is the main parameter of finding the precision and recall which are acting as the deciding factor if the models had over-fitted the training dataset and also indicate if the models were perfectly trained by understanding the correlation provided that the accuracy of the models are high. Hence it is evident from the performances provided by the models that all the four models were consistent in classifying the categories while VGG16 and ResNet50 outperformed the others with accuracy of 97.15% and 96.21% respectively. However, ResNet and VGG16 had nearly identical results in the investigation. When there are vanishing gradient concerns, ResNet usually performs better. On the other hand, VGG16 is optimized for image recognition. However, in order to avoid any inaccuracies, to properly evaluate and analyze model performance, the f1 score value of each model is taken into account giving a consistent value of 97.14% and 96.22% accordingly.

Chapter-7

Conclusion and Future works

7.1 Conclusion

Any latent disability causes people to suffer since it goes undiscovered for a long time, which might make things worse by making it incurable and resulting in lifetime impairment; gait disorders are no different. GDs result in a loss of personal freedom, falls and injuries, as well as a significant decline in life quality. Gait disorder should be extensively evaluated to enhance patient's mobility and prevent falls furthermore to identify the underlying reasons as soon as feasible. Traditional methods of GDD were rigid and these were manually performed by medical specialists, considerably increasing the process's complexity and time. The purpose of this research is to develop a significant yet simple GDD approach based on the implementation of CNN models, specifically transfer learning, to classify a person's foot anomalies. The 97.15% accuracy achieved from the model Vgg16 in distinguishing normal and abnormal gait demonstrates the applicability of the proposed mechanism. Additionally, the future prospects of the model can be promising since there is still room for the model to evolve by generating real life data rather than the synthetic data employed in this research. As a result, this innovative insight aims to make gait anomaly detection a big change in the medical field, making it accessible to the general public.

The introduction of **transfer learning** has not only made the study effective, furthermore introducing **data augmentation** techniques significantly increased the size and complexity of the data which created opportunities to make a generalized model for various types of sensor data.

The purpose of this research is to develop a significant yet simple gait disorder detection approach based on the implementation of CNN models, specifically transfer learning, to classify a person's foot anomalies.

7.2 Future Works

Our future goal is to basically reduce the time required for gait analysis which means to execute the process with ease and comfort. Another vital prospect of the project is to detect the disorder or anomalies with clinical perfection. To obtain the proposal's main concept is to ensure the proper utilization of piezoelectricity. Piezoelectricity means electricity resulting from pressure or latent heat. The main concept is to generate electric energy converted from mechanical energy.

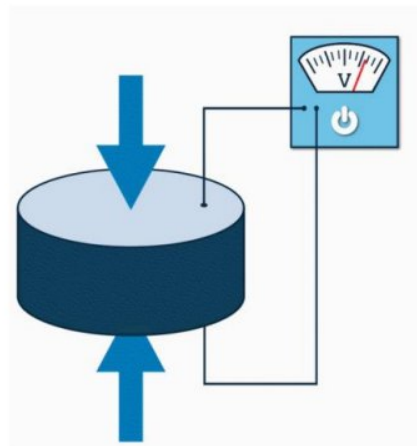


Figure 7.1: Generation of piezoelectricity

If mechanical stress is applied to certain materials, they tend to accumulate electric charges. The piezoelectric effect generally means that a voltage will be generated when any pressure is applied to any piezoelectric material. Piezoelectric materials can be used to create various sensors or actuators: applied periodic electrical signals can result in the generation of ultrasonic waves for imaging purposes.

When pressure is not applied to piezoelectric material, the molecules arrange themselves in a certain way which corresponds to an equilibrium of the matter and no electric field is found. When a pressure is applied however, those molecules change position and align into a dipolar state in which a resultant electric field is found. The strength of the field depends on the magnitude of pressure.

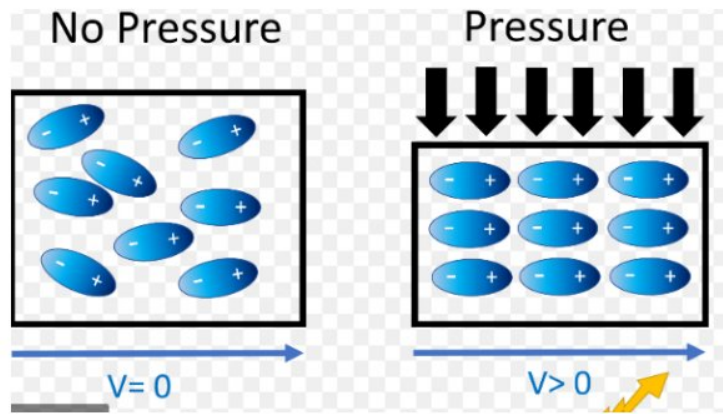


Figure 7.2: Methodology of generation of piezoelectricity

So, by using this foot sensor, heatmap images can be taken and can be implemented instantly for detecting any type of gait disorder within a very short span of time. That would complete the gait analysis process with ease.

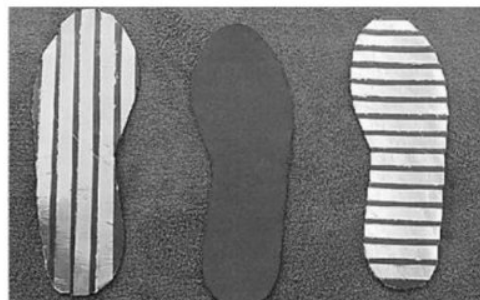


Figure 7.3: Piezoelectric Sensor

The fundamental future works includes-

- To introduce a mobile app that will detect foot anomalies through gait analysis just by evaluating the heatmap images from the pressure sensors given to the app as input.
- To increase evaluation matrices of the model for proving the strength of the model.
- To evaluate the model on real data along with the implementation of the new architectures such as convnext, coatnet, swin which will be used to compare with the existing ones.

- Hardware implementation of the model to make the detection easier in the medical sector.

Thus this innovative insight aims to make **gait anomaly detection** a big change in the medical field, making it accessible to the general public and making an impact on the upcoming **AI takeover** in the medical sector.

References

1. Cifuentes, C.A. et al. "Assistive Devices for Human Mobility and Gait Rehabilitation," In *Human-Robot Interaction Strategies for Walker-Assisted Locomotion*, pp. 1-15, Springer, Cham, 2016
2. Gómez-Cabello, A., et al. "Sitting time increases the overweight and obesity risk independently of walking time in elderly people from Spain," *Maturitas*, 73(4), pp. 337-343, 2012.
3. W. Pirker and R. Katzenschlager, "Gait disorders in adults and the elderly: A clinical guide," *Wiener klinische Wochenschrift*, vol. 129, nos. 3-4, pp. 81-95, Feb. 2017, doi: 10.1007/s00508-016-1096-4.
4. Preeti Khera and Neelesh Kumar (2020) Role of machine learning in gait analysis: a review, *Journal of Medical Engineering and Technology*, 44:8, 441-467, DOI: 10.1080/03091902.2020.1822940
5. Negi, S., Sharma, S. and Sharma, N., 2021. FSR and IMU sensor based human gait phase detection and its correlation with EMG signal for different terrain walks. *Sensor Review*.
6. Prasanth, H., Caban, M., Keller, U., Courtine, G., Ijspeert, A., Vallery, H. and Von Zitzewitz, J., 2021. Wearable sensor-based real time gait detection: A systematic review. *Sensors*, 21(8), p.2727.
7. Faisal F., et al. "An Investigation for Enhancing Registration Performance with Brain Atlas by Novel Image Inpainting Technique using Dice and Jaccard Score on Multiple Sclerosis (MS) Tissue", *Biomedical and Pharmacology Journal*, 12(3), pp. 1249-1262, 2019
8. Nishat, M. M., Faisal, F., et al. "An Investigative Approach to Employ Support Vector Classifier as a Potential Detector of Brain Cancer from MRI Dataset," *ICECIT*, pp. 1-4, IEEE, 2021, doi: 10.1109/ICECIT54077.2021.9641168
9. Nishat, M. M., Faisal, F., et al. "Performance Investigation of Different Boosting Algorithms in Predicting Chronic Kidney Disease," *2nd International Conference on STI*, pp. 1-5. IEEE, 2020.
10. Asif, M. A. A. R. et al. "Performance Evaluation and Comparative Analysis of Different Machine Learning Algorithms in Predicting Cardiovascular Disease," *Engineering Letters*, 29(2), 731-741, 2021
11. Farazi, M. R., et al. "Inpainting multiple sclerosis lesions for improving registration performance with brain atlas," *MediTec*, pp. 1- 6. IEEE, 2016 doi: 10.1109/MEDITEC.2016.7835363

12. Asif, M. A. A. R., et al. "Computer Aided Diagnosis of Thyroid Disease Using Machine Learning Algorithms," 2020 11th ICECE, pp. 222-225. IEEE, 2020. doi: 10.1109/ICECE51571.2020.9393054
13. M. M. Nishat et al., "An Investigation of Spectroscopic Characterization on Biological Tissue," iCEEICT 2018, pp. 290-295, doi: 10.1109/CEEICT.2018.8628081.
14. Faisal, F., et al. "Covid-19 and its impact on school closures: a predictive analysis using machine learning algorithms," In 2021 International Conference on Science and Contemporary Technologies (ICSCT), IEEE, 2021, DOI: 10.1109/ICSCT53883.2021.9642617
15. Rahman, A. A., et al. "Detection of Epileptic Seizure from EEG Signal Data by Employing Machine Learning Algorithms with Hyperparameter Optimization," BioSMART, IEEE 2021, DOI: 10.1109/BioSMART54244.2021.9677770
16. Rahman, A. A., et al. "Enhancing the Performance of Machine Learning Classifiers by Hyperparameter Optimization in Detecting Anxiety Levels of Online Gamers," ICCIT, IEEE, 2021 DOI: 10.1109/ICCIT54785.2021.9689911
17. Liu, Z.J., and Sarkar, S., "Simplest representation yet for gait recognition: averaged silhouette," Proceedings of the 17th International Conference on Pattern Recognition, pp. 211-214, 2004
18. J. Han and B. Bhanu, "Individual recognition using gait energy image," IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(2), pp. 316-322, 2006, doi: 10.1109/TPAMI.2006.38.
19. K. Bashir, T. Xiang, and S. Gong, "Gait recognition using gait entropy image," in Proc. 3rd Int. Conf. Imag. Crime Detection Prevention (ICDP), 2009, pp. 1-6.
20. Jeevan, M., Jain, N., Hanmandlu, M., and CHETTY, G. (2013). Gait recognition based on gait pal and pal entropy image. In D. Taudman, and M. Wu (Eds.), Proceedings 20th 2013 IEEE International Conference on Image Processing (ICIP 2013) (pp. 4195-4199). IEEE
21. H. Iwama, et al. "The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition," IEEE Trans. Inf. Forensics Security, 7(5), pp. 1511-1521, 2012
22. Matthew Turk, Alex Pentland; Eigenfaces for Recognition. J Cogn Neurosci 1991; 3 (1): 71-86.
23. P. N. Belhumeur, et al., "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), pp. 711-720, 1997
24. N. Liu, J. Lu and Y. Tan, "Joint Subspace Learning for ViewInvariant Gait Recognition," in IEEE Signal Processing Letters, vol. 18, no. 7, pp. 431-434, 2011, doi:

10.1109/LSP.2011.2157143.

25. Yan, S., et al. "Discriminant analysis with tensor representation," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 526-532

26. Z. Lai, et al, "Human Gait Recognition via Sparse Discriminant Projection Learning," IEEE Transactions on Circuits and Systems for Video Technology, 24(10), pp. 1651-1662, 2014

27. X. He and P. Niyogi, "Locality preserving projections," in Proc. NIPS, vol. 16, 2003, pp. 153–160.

28. M. Sugiyama, "Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis," J. Mach. Learn. Res., vol. 8, pp. 1027–1061, 2007.

29. T. Hasan et al., "Exploring the Performances of Stacking Classifier in Predicting Patients Having Stroke," NICS, 2021, pp. 242-247, doi: 10.1109/NICS54270.2021.9701526

30. A. A. Rahman et al., "Detection of Mental State from EEG Signal Data: An Investigation with Machine Learning Classifiers," 2022 14th International Conference on Knowledge and Smart Technology (KST), 2022, pp. 152-156, doi: 10.1109/KST53302.2022.9729084

31. Nishat, M. M., Faisal, F., et al. "A Comprehensive Analysis on Detecting Chronic Kidney Disease by Employing Machine Learning Algorithms," EAI Endorsed Transactions on Pervasive Health and Technology, phat 18: e6, 2021, DOI: 10.4108/eai.13-8-2021.170671

32. Nishat, M. M., et al. "A Comprehensive Investigation of the Performances of Different Machine Learning Classifiers with SMOTE-ENN Oversampling Technique and Hyperparameter Optimization for Imbalanced Heart Failure Dataset," Scientific Programming, <https://doi.org/10.1155/2022/3649406>

33. Rahman, M. R., et al. "CNN-based Deep Learning Approach for Micro-crack Detection of Solar Panels," 3rd International Conference on STI 4.0, 2021, pp. 1-6, doi: 10.1109/STI53101.2021.9732592

34. C. Szegedy et al., "Going deeper with convolutions," IEEE CVPR, pp. 1-9, 2015

35. E. P. Ijjina et al. "One-Shot Periodic Activity Recognition Using Convolutional Neural Networks," 13th International Conference on Machine Learning and Applications, pp. 388-391, 2014

36. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Proc. Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.

37. A. Karpathy, et al. "Large-scale video classification with convolutional neural networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 1725–1732, 2014.

38. A. Rohan, M. Rabah, T. Hosny and S. -H. Kim, "Human Pose Estimation-Based Real-Time Gait Analysis Using Convolutional Neural Network," in *IEEE Access*, vol. 8, pp. 191542-191550, 2020
39. Sadeghzadeh Yazdi, N., et al. "Acton, Modeling spatiotemporal patterns of gait anomaly with a CNN-LSTM deep neural network," *Expert Systems with Applications*, vol 185, 2021, 115582
40. K. Jun, D. Lee, K. Lee, S. Lee and M. S. Kim, "Feature Extraction Using an RNN Autoencoder for Skeleton-Based Abnormal Gait Recognition," in *IEEE Access*, vol. 8, pp. 19196-19207, 2020
41. M. Martinez et al., "Falls Risk Classification of Older Adults Using Deep Neural Networks and Transfer Learning," *IEEE Journal of Biomedical and Health Informatics*, 24(1), pp. 144-150, 2020
42. Fricke, Christopher et al. "Evaluation of Three Machine Learning Algorithms for the Automatic Classification of EMG Patterns in Gait Disorders." *Frontiers in neurology*, vol. 12, 666458, 2021.
43. Sa, T. R. D., and Carlos MS Figueiredo. "Classification of anomalous gait using Machine Learning techniques and embedded sensors." *arXiv preprint arXiv:2110.06139* (2021)