Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

Telecommunication Engineering Graduate Program

Subscription Fraud Prevention in Telecommunication using Deep Learning Approach: the case of ethio telecom

By: Gebremeskel Aregay

Advisor: Rosa Tsegaye (PHD)

A Thesis Submitted to

School of Electrical and Computer Engineering

In Partial Fulfillment of the Requirements for the Degree of Masters of Science in Telecommunications Engineering

December, 2021

Addis Ababa, Ethiopia

Addis Ababa University

Addis Ababa Institute of Technology

School of Electrical and Computer Engineering

Telecommunication Engineering Graduate Program

MSc Thesis on

Subscription Fraud Prevention in Telecommunication using Deep Learning Approach: the case of ethio telecom

By: Gebremeskel Aregay Kidane

Signed by the Examining Committee:

_____	_____
Chair/School Dean	Signature
_____	_____
Advisor	Signature
_____	_____
Examiner (1)	Signature
_____	_____
Examiner (2)	Signature
_____	_____
Director of Postgraduate Program	Signature

Declaration

I, the undersigned, declare that the thesis comprises my own work in compliance with internationally accepted practices; I have fully acknowledged and referred all materials used in this thesis work.

Gebremeskel Aregay Kidane _____

Name Signature

Abstract

With the rapid development of telecommunication services, telecommunication fraud is a challenging issue that fraudsters continuously participate in abusing the services to get undeserved benefit. For telecommunication operators who have no way of identifying mechanisms for their customers' legitimacy during service subscription, could lead to subscription fraud which is the cause for other fraud types. The purpose of the research is to prevent subscription fraud by identifying customer's face using face recognition techniques and propose the deployment scenario in telecommunication industries especially for those who have customer portraits.

This study uses deep learning algorithms to build a face recognition model using a dataset of 124 identities collected from the Internet, personal image gallery and capturing from some friends. These identities have two different datasets; with 2596 images and 1240 images. The former dataset split into 80% and 20% ratio whereas the later one is vice versa provided that each identity in the training set has two images. Two separate experiments have been conducted for the first dataset; CNN with transfer learning, and MTCNN with FaceNet and SVM, whereas in the second dataset, the experiment with better accuracy in the first dataset is further retrained.

Experimental results obtained an accuracy of 72.54% for CNN with transfer learning and 99.24% for MTCNN, FaceNet and SVM in the first dataset. The latter set of algorithms has achieved an accuracy of 98.69% in the second dataset which is recommended as a better solution. Finally, feasibility of the model deployment scenario has been analyzed and proposed by assessing the necessary requirements used for implementation.

Keywords: Telecommunication Fraud, Subscription Fraud, ethio telecom, Deep Learning, Face Recognition, CNN, Transfer Learning, MTCNN, FaceNet

Acknowledgement

First and foremost, I would like to thank the Almighty God for giving me the courage and strength to complete my thesis work in this toughest time.

I'm very grateful to my adviser, Rosa Tsegaye (PHD) for her valuable directions and comments during my thesis work. Next, I would like to express my appreciation to ethio telecom for offering me the opportunity to attend my postgraduate program in Addis Ababa University, Addis Ababa Institute of Technology (AAiT). Additionally, I would like to thank Tamirat Teshome, Derese Agonafir, Melkamu Asmare, Fantaye Mekonen, Samuel Medhin and others who contribute a lot in providing me necessary information and data that help me to accomplish my study.

Last but not least, my special gratitude goes to my wife, Rishan Birhane for her continuous support and courage to pursue my study in her difficult times. I would also like to thank my parents for what they did to me and their indispensable role throughout my life that contributed a lot for what I'm today.

Thank you!
Gebremeskel Aregay Kidane

Table of Contents

DECLARATION II

ABSTRACT III

TABLE OF CONTENTS V

LIST OF FIGURES VII

LIST OF TABLES VII

ACRONYMS VIII

CHAPTER ONE 1

INTRODUCTION 1

1.1. Background *1*

1.2. Statement of the Problem *2*

1.3. Objective *3*

1.3.1. General Objective 3

1.3.2. Specific Objectives 3

1.4. Scope and Limitation of the Study *4*

1.5. Contribution of the Study *4*

1.6. Literature Review *4*

1.7. Thesis Organization *6*

CHAPTER TWO 7

OVERVIEW OF TELECOMMUNICATIONS FRAUD 7

2.1. Telecommunication Fraud Types *8*

2.1.1. Superimposed Fraud 9

2.1.2. International Revenue Share Fraud 9

2.1.3. Subscriber Identity Module Box Fraud 9

2.1.4. Subscription Fraud 10

CHAPTER THREE 12

DEEP LEARNING ALGORITHM AND FACE RECOGNITION 12

3.1. Deep Learning *12*

3.1.1. Convolutional Neural Networks 13

3.1.2. Transfer Learning 15

3.1.3. Multi-Task Cascaded Convolutional Neural Network 17

3.1.4. FaceNet 19

3.1.5. Support Vector Machine 21

3.2. Face Recognition *21*

3.2.1. Challenging areas in Face Recognition 22

CHAPTER FOUR 23

METHODOLOGY 23

4.1. Data Collection and Preparation *23*

4.2. Model Building Environment *25*

4.3. System Model *25*
4.4. Tools Selection *28*
4.4.1. OpenCV 29
4.4.2. TensorFlow 29
4.4.3. Keras 30
4.4.4. Python 30

CHAPTER FIVE **31**

EXPERIMENTAL RESULTS AND MODEL IMPLEMENTATION 31
5.1. Results *31*
5.1.1. Using CNN with Transfer Learning 31
5.1.2. Using MTCNN and FaceNet 32
5.2. Model Implementation *34*
5.2.1. How the model prevents Subscription Fraud 34
5.2.2. Storage Requirement 35
5.2.3. Concurrency 36
5.2.4. Mode of Deployment 38
5.2.5. Cost Estimation 39

CHAPTER SIX **41**

CONCLUSION AND FUTURE WORKS 41
6.1. Conclusion *41*
6.2. Future Works *42*

REFERENCES **43**

List of Figures

FIGURE 3. 1: GENERAL CNN ARCHITECTURE [24] 15
FIGURE 3. 2: VGG16 ARCHITECTURE [11] 17
FIGURE 3. 3: MTCNN ARCHITECTURE FRAMEWORK [28] 18
FIGURE 3. 4: FACENET ARCHITECTURE [30] 20
FIGURE 3. 5: TRIPLET LOSS FUNCTION [30] 20
FIGURE 4. 1: FACE RECOGNITION PROCESS MODEL 26
FIGURE 5.1: SAMPLE OUTPUTS OF THE FACE RECOGNITION MODEL 33
FIGURE 5. 2: PROPOSED FLOW DIAGRAM OF SUBSCRIPTION FRAUD PREVENTION MODEL 35
FIGURE 5. 3: RESULT OF CONCURRENT OPERATION OF REQUESTS 37
FIGURE 5. 4: INTEGRATION OF FACE RECOGNITION FUNCTION WITHIN E-CAF SYSTEM [49] 38

List of Tables

TABLE 1.1: CFCA SUMMARY REPORT OF REVENUE LOSS DUE TO FRAUD [4] 2
TABLE 1.2: REVENUE LOSS REPORT DUE TO FRAUD IN ETHIO TELECOM [5] 2
TABLE 4. 1: DATASET SPLITTING 24
TABLE 5. 1: OUTPUT OF CNN WITH VGG16 USING DIFFERENT SCENARIOS 32
TABLE 5. 2: EXPERIMENTAL OUTPUTS FOR MTCNN AND FACENET 33
TABLE 5. 3: DETAIL COST ESTIMATION OF THE MODEL DEPLOYMENT 40

Acronyms

ABANN	AdaBoost and Artificial Neural Network
ANN	Artificial Neural Networks
AI	Artificial Intelligence
API	Application Programming Interface
AT&T	American Telephone and Telegraph Company
CDR	Call Detail Records
CFCA	Communications Fraud Control Association
CNN	Convolutional Neural Networks
CRM	Customer Relationship Management
DR	Disaster Recovery
E-CAF	Electronic Customer Acquisition Form
FC	Fully Connected
FCN	Full connected Convolution Network
FMS	Fraud Management System
GPU	Graphics Processing Unit
HA	High Availability
ICT	Information and Communication Technology
IRSF	International Revenue Share Fraud
ML	Machine Learning
MTCNN	Multi-Task Cascaded Convolutional Neural Network
NMS	Non-Maximum Suppression
O-Net	Output Network
OpenCV	Open source Computer Vision
P-Net	Proposal Network
PBX	Private Branch eXchange
PC	Personal Computer
PCA	Principal Component Analysis
PRS	Premium Rate Service
R-Net	Refine Network

ReLU	Rectified Linear Unit
SIMBox	Subscriber Identity Module Box
SMS	Short Message Service
SVM	Support Vector Machine
TPU	Tensor Processing Unit
VGG	Visual Geometry Group
VoIP	Voice over Internet Protocol

Chapter One

Introduction

1.1. Background

The rapid development of telecommunication network services create immense opportunities for both service providers and customers in terms of delivery of newly emerged technological products, getting various services and generating a significant amount of revenues. As part of the customers' endless demand, every time with the emerging of new telecommunication services, there is always an intention of using the services without or partial payment. In parallel to the expansion of telecommunication services and the huge amount of revenue generation, there is a need for fraudsters participating in telecommunication fraud activities. To keep the revenue generation as needed, in any means telecom operators have to protect their networks as well as services from fraudsters.

Telecommunication fraud is among one of the most severe threats to revenue and quality of service in telecom networks. It is a theft or deliberate abuse of services that has a remarkable impact on the company's revenue generation [1]. In ethio telecom, fraud is a serious issue and special attention has been given to manage the different fraudulent activities in the company.

In 2017, Communications Fraud Control Association (CFCA) has announced an annual fraud survey report and globally telecommunication operators were hit with $29.2 billion revenue losses [2]. Similarly in 2019, $28.3 billion revenue loss has occurred [3]. Thus, telecommunication fraud is a major issue in telecommunication service providers and their customers around the world. Subscription fraud is among one of the top telecom frauds that has significant revenue loss.

Subscription fraud is characterized by a fraudster using own, stolen or fabricated identity to get telecommunication services with no intention to pay. The theft here is plain and simple but hard to detect 'intent' at the point of sale [4]. Customers can use their mobile service numbers to commit fraud in other fraud types such as Subscriber Identity Module box (SIMBox) fraud and

roaming fraud which have a very high impact on the company's revenue loss. In ethio telecom, currently customers are forced to subscribe for a limited number of mobile service numbers, but there is a possibility that one customer can subscribe using counterfeit identities to get service numbers more than the allowed number of service numbers to use for fraud activities. Preventing subscription fraud helps to reduce committing different illegal activities even outside the company. According to [4], since 2013 subscription fraud is one of the top five fraud types that have a high contribution in revenue loss as shown in Table 1.1. In the case of ethio telecom, though the majority of the revenue loss is happening due to SIMBox fraud, subscription fraud is the root cause for this fraud type. The revenue loss in ethio telecom is increasing according to [5] shown in Table 1.2.

Year	Top five fraud total revenue loss ($Billion)	Subscription fraud loss contribution ($Billion)
2013	20.23	5.22
2015	16.68	2.55
2017	9.59	2.03

Table 1. 1: CFCA summary report of revenue loss due to fraud [4]

Year	Ethio telecom's total revenue loss ($Million)
2015/16	35.5
2016/17	52
2017/18	89

Table 1. 2: Revenue loss report due to fraud in ethio telecom [5]

1.2. Statement of the Problem

Telecommunication fraud is a major issue in telecommunication service providers as well as their customers. In ethio telecom, customers get registered at the sales office using Customer Relationship Management (CRM) and electronic Customer Acquisition Form (e-CAF) systems to get services. During the registration process, there is no way of checking mechanism that uniquely identifies the identity of the customer. Customers can subscribe to get service using different false identities and even a name of one customer can be written in different forms/alphabets so that he/she is treated as a different identity in the system. For example a

customer whose name is written in any of the local languages in his/her Identification (ID) card, it can be registered in different ways during the subscription time in the system. Because, in Ethiopia ID cards are written in local languages but in ethio telecom systems, customers register in English language. However, in e-CAF we have one thing that can uniquely identify the person, i.e. photo. The sales person always captures a photo of the customer when subscribed for the service.

Currently, ethio telecom uses Fraud Management System (FMS) to detect different telecom fraudulent activities using rule-based or threshold-based techniques. However, due to the dynamic behavior of fraudsters, it is not efficient to detect fraudulent activities using these techniques. In ethio telecom, there is no system which is capable of identifying the subscriber's identity either by using photo or any other mechanism. Therefore, to combat this challenge a dynamic system is required by building a model using deep learning algorithms to detect and prevent the fraudsters as well as to reduce the revenue loss of the company.

1.3. Objective

1.3.1. General Objective

In general, the objective of this research is to prevent fraudulent activities committed in telecom operators due to subscription fraud which is the root cause for several fraud types by using deep learning approach.

1.3.2. Specific Objectives

The specific objectives of the research work are:

- ✓ Investigate deep learning techniques that help to prevent subscription fraud
- ✓ Choose appropriate deep learning algorithm(s) and techniques for subscription fraud prevention
- ✓ Build a model with the selected relevant algorithm(s)
- ✓ Evaluate the performance of the model
- ✓ Analyze the model for real implementation
- ✓ Propose model deployment scenario including the requirements needed

1.4. Scope and Limitation of the Study

The scope of this research work is to develop a model that helps to prevent subscription fraud and propose a deployment scenario by assessing required things in ethio telecom. Nowadays there are different types of telecommunication frauds, among those subscription fraud is the scope of this thesis work. The study is limited to images dataset available in ethio telecom even though the company did not provide us the data and only limited numbers of image data were used in the experiment.

1.5. Contribution of the Study

This research work can help telecom operators to understand the techniques and methods of how to prevent subscription fraud during their service provision and delivery. The developed model can play a great role in verifying the subscriber's identity to have a real identification. The study will have a significant role in minimizing the revenue loss. On the other hand, it is also helpful for researchers to give hints for other related research topics.

1.6. Literature Review

Several researches have been conducted in the area of telecommunication fraud. Telecommunication fraud is a dynamic global issue, numerous studies employ data mining and machine learning (ML) techniques to tackle this problem.

As stated in [1], telecommunication fraud involves the theft of services or deliberate abuse of voice and data networks, and impacts on the telecom company in four ways - financially, marketing, customer relations and shareholder perceptions. It focuses on Adaptive flexible techniques using advanced data analysis like Artificial Neural Networks (ANNs) due to its inherent ability to adapt, as well as its speed and efficiency to manage and detect telecom frauds. Researchers use neural networks for the detection of telecommunication fraud using the new concept of profiling that aims to deal with the major weaknesses in the rule based techniques of dealing with large variations in genuine usage between various customers.

Finally, the network has successfully trained using simulated data resulting in a 98% detection rate when tested for call-sell fraud.

Authors [6] describe subscription fraud as the acquisition of telecom services using fraudulently obtained subscriber documents, or false identification. They have developed a decision tree model to determine the impact of subscription fraud in mobile telecommunication company focusing on post-paid subscribers. According to the researchers, the majority of previous researches focus on reducing or stopping subscription fraud and very few determine the impact it has on mobile telecommunication companies. The result of the model shows that 15.9% of subscriber churns are due to subscription fraud and that 30% revenue loss is caused by subscription fraud.

A research in [7] shows that subscription fraud has been one of the top-five ranked fraud types since 2013. It stated that out of the top-five total annual revenue loss in 2017; nearly 24% of the total loss occurred due to subscription fraud. The researcher has conducted a comparative performance of three machine learning algorithms: ANN, Support Vector Machine (SVM) and J48 to better detect subscription fraud by using Call Detail Records (CDR) data. Finally, J48 using Cross Validation becomes the best classifier algorithm by scoring 99.3% accuracy.

Researchers [8] have proposed a random rough subspace based neural network ensemble method to detect subscription fraud using CDR data in mobile telecommunication. All the aforementioned subscription fraud researches are done based on the CDR data to detect fraudsters. However, instead of detecting subscription fraudsters using CDR data, it could be better to prevent using various means such as face recognition when subscribers try to subscribe illegitimately to get telecom services. Thus, the focus of this research work is to prevent subscription fraud using deep learning algorithms that use face images as an input and recognize the subscriber's face.

The study in [9] has introduced some novel models for all steps of a face recognition system. In the face detection step, a hybrid model combining AdaBoost and ANN (ABANN) is proposed to solve the process efficiently. In the next step, labeled faces detected by ABANN have been aligned by Active Shape Model and Multi-Layer Perceptron; and in the feature extraction step, geometric feature based method and Independent Component Analysis method have been used

for improving the efficiency; whereas in the face matching step, a model combining Multi ANN has been applied for geometric features matching of human face.

Deep learning methods, especially Convolutional Neural Networks (CNN) have achieved significant success in face recognition problems [10]. CNNs are neural networks that employ convolution operation instead of matrix multiplication in the convolutional layers. The paper [10] has proposed a modified deep learning neural network to learn face representation from a smaller dataset. The network was trained using the standard AT&T face database and then substantial improvement in recognition rate was achieved.

CNN with Transfer Learning method is proposed for automated face recognition system [11]. Transfer learning is a ML method where the knowledge gained from a particular task is transferred to improve the process of learning in another related task. The CNN architectures like VGG, ResNet etc., are already trained on the huge image database of ImageNet. The paper [11] used VGG16 CNN architecture for transfer learning and conducted an experiment on two publicly available face databases: Yale and AT&T. Finally, it achieved an accuracy of 100% and 98.7% for AT&T and Yale face databases respectively.

1.7. Thesis Organization

The remaining part of this thesis work is organized as follows. Chapter 2 presents the overview of telecommunications fraud including some specific fraud types, especially subscription fraud. Chapter 3 discusses deep learning algorithms and face recognition techniques used for this thesis. The methodology followed for the model building process is covered in chapter 4. Chapter 5 discusses experimental results and model deployment. Finally, conclusions and future works are incorporated in chapter 6.

Chapter Two

Overview of Telecommunications Fraud

Telecommunications fraud is one of the most severe threats to revenue generation and quality of service in telecommunication industries. It is defined as the theft of telecommunication services or the use of telecommunication services to commit other forms of fraud [4]. It can also be described as the abuse of telecommunications products or services with the intention of illegally acquiring money from telecommunication service providers or its customers. With the continuous development of telecommunications network technology, the way that criminals commit telecommunications network fraud is becoming more hidden and realistic. Telecom fraud has a dynamic nature that means whenever fraud actors (fraudsters) feel that they will be detected or prevented, they try to find a way to bypass security measures [12]. Because fraudsters are smart enough that continuously look for exploitable weaknesses in the telecom services and networks. Part of their motivation is accounted for by the fact that once an exploit is defined, plenty of potential targets could be available.

The problem of telecommunication fraud is the huge loss of revenue that can affect the credibility and performance of telecommunication companies [12]. Fraudsters have the need to either reduce or completely avoid the charges for using the services. Globally, telecommunications fraud losses are estimated in multi-billions of U.S dollars every year [3]. Telecom fraud negatively impacts the telecommunication company financially, marketing, customer relations, and shareholder perceptions [1].

Some researchers [13] have underlined the significance of distinguishing between fraud prevention and fraud detection. Fraud prevention which is the focus of this research is the first layer of protection technique to secure the telecommunication services against fraudulent activities. Its primary purpose is to avoid fraud from occurring in the first place. Whereas, fraud detection involves identifying fraudulent activities as quickly as possible once it has been committed. Fraud detection is the second layer of protection which tries to discover and identify fraudulent activities after it has been committed.

2.1. Telecommunication Fraud Types

As per the study in [4], telecommunication frauds are categorized into four groups: known as Contractual fraud, Hacking fraud, Technical fraud, and Procedural fraud. In **Contractual fraud**, revenue is generated through the normal use of service while having no intention of paying for the usage; for example, Premium Rate service (PRS) fraud and Subscription fraud. In **Hacking fraud**, revenue is generated for the fraudster by breaking into insecure systems and exploiting on any available functionality such as Private Automatic Branch eXchange (PABX) fraud. In **Technical fraud,** all frauds involve attacks against weaknesses in the technology of the mobile system. Such frauds typically need some initial technical knowledge; although once a weakness has been exposed this information is often quickly distributed in a form that non-technical people can use. Cloning and technical internal frauds are some of the examples of this category. Whereas, in **Procedural fraud**, all frauds involve attacks against the procedures used to minimize exposure to fraud and attack the weaknesses in the business procedures used to grant access to the system. Examples of this category are Faulty vouchers, Voucher ID duplication, and Roaming fraud.

According to CFCA, till 2013 every form of fraud was considered as a fraud type. The grouping of frauds into fraud type and fraud method is done in the 2013 report [4]. Based on this report some examples of the fraud types are arbitrage, spamming, domestic/international revenue share fraud, phishing, PRS fraud, and roaming fraud. On the other hand, the fraud methods category incorporates Cramming/Slamming, PBX Hacking, SMS Phishing, Subscription Fraud, Wangiri fraud, and so on.

As per the research [14], there are many types of frauds that threaten the telecommunication sector, which are considered as the most popular fraud area. It is estimated that more than 200 variants of telecom fraud exist in the telecommunications industry. Among those, subscription fraud and super-imposed fraud are the most prevalent types of telecommunication frauds. In general, some of the prevalent fraud types and methods in the telecommunication industry are described below.

2.1.1. Superimposed Fraud

In superimposed fraud, fraudsters take over a legitimate account and use services without the necessary authority for using it and it would appear as phantom calls in the bill [14]. In such cases, the abnormal usage is superimposed upon the normal usage of the legitimate customers. There are different ways to carry out superimposed fraud, including mobile phone cloning and obtaining calling card authorization details.

2.1.2. International Revenue Share Fraud

International Revenue Share Fraud (IRSF) is a telecom fraud where perpetrators artificially inflate traffic by generating calls to different international number ranges with no intention to pay for the calls or paying where there exists some form of arbitrage opportunity [15]. IRSF is one of the most severe fraud types that cause high revenue loss among the top-five types in [2]. Fraudsters receive a share of the revenue from the terminating telecom operators that earn the charges generated due to the inflated traffic. IRSF is often carried out in conjunction with many different methods used by the fraudster to obtain access to the network services required to generate calls. Examples include subscription fraud, PBX hacking, Wangiri, fraud and SMS spamming [15].

2.1.3. Subscriber Identity Module Box Fraud

Subscriber Identity Module Box (SIMBox) fraud is one of the most common fraud types in ethio telecom that affects the revenue and quality of service of the network. SIMBox fraud is a bypass fraud type that is used in international calls and it has emerged with the use of Voice over Internet Protocol (VoIP) technologies [16][17][18]. SIMBox frauds hijack international voice calls and transmit to a cellular device through the Internet and changes to a local call at the destination network. SIMBox devices operate with a multiple number of SIM cards both on foreign and national operators. Whenever an operator detects and deactivate the fraudulent service numbers, fraudsters can use new set of SIM card numbers until the operator detects and shutdown them [17]. In the case of ethio telecom, fraudsters use subscription fraud to acquire excess amounts of new SIM cards to replace the detected fraud SIM cards. The main driving

factor of this fraud type is the tariff difference between local and international calls, where termination costs of international calls are high [18].

2.1.4. Subscription Fraud

In the telecommunication industry, subscription fraud is one of the prevalent telecom frauds. It is the use of fake identities and/or identity theft at the point of sale that enables either the fraudulent use of telecom services or the use of such services for subsequent fraudulent activities. Once a fake subscription is done, fraudsters get access to telecom services, including data, voice, Short Message Services (SMS), and mobile financial services such as mobile payment and mobile banking [19]. The motive might be no more than being opportunistic or attempting to exploit a known vulnerability. However, this is now run by organized criminals – building multiple fraudulent identities over long periods. Furthermore, fraudsters have gained detailed fraud system knowledge and continually test the thresholds to exploit the loopholes in the systems [4].

Subscription fraud is usually the precursor to other fraud types such as PRS fraud, IRSF fraud, SIMBox fraud and Roaming fraud [7]. Subscription fraud can affect both post-paid and pre-paid subscribers.

According to [6], there are three subscription fraud methods named:

1. Subscription Fraud (Identity) - is the utilization of a real identity without the owner's knowledge to obtain goods and services with no intention to pay.
2. Subscription Fraud (Application) - is a creation of false details to gain access to goods and services with no intention to pay.
3. Subscription Fraud (Credit Mulling/Proxy) - is the utilization of real identity details to obtain goods and services with no intention to pay.

Both the first and third fraud methods are caused by identity theft while subscription fraud (application) is caused by identity fraud.

Subscription fraud resulting from inadequate identity verification procedures can take time to discover the fraudsters. As a result, telecommunication service providers are subject to

considerable financial cost loss through a single subscription such as international calls, internet, and mobile financial services. The failure to prevent forged or stolen identifications at the point of sales exposes fraudsters to commit widespread fraudulent activities. Therefore, effective identity verification service is a key technique to tackle subscription fraud that helps telecom operators to reduce revenue loss.

Chapter Three

Deep Learning Algorithm and Face Recognition

3.1. Deep Learning

Systems rely on hard-coded knowledge face challenges to solve dynamic problems. Due to this challenge, an Artificial Intelligence (AI) system is suggested to acquire knowledge dynamically by extracting patterns from raw data. AI is the simulation of human intelligence in machines that are programmed to think like human beings and imitate their activities [20]. The way of implementing AI by extracting features from the raw data is accomplished by ML. The introduction of ML allowed computers to tackle problems involving knowledge of the real world and make decisions that appear subjective. ML is the study of algorithms that automatically improve their performance with experience enriching their decisions through learning, which is attained by an iterative process. It is a good and applicable approach to have AI systems that can operate in sophisticated real world environments.

Deep learning is a subset of ML that can achieve higher flexibility of learning to represent the world as a nested hierarchy of concepts. Deep learning is considered to be a suitable way for extracting important and meaningful features from the available raw data. It does not depend on hand-crafted features like local binary patterns, a histogram of gradients, etc., and most importantly it performs a hierarchical feature extraction. It learns features layer wise which means that in initial layers it learns low-level features and as it moves up the hierarchy it starts to learn a more abstract representation of the data. On the other hand, as compared to deep learning, ML is not a good method for extracting meaningful features from the raw data. It relies on hand-crafted features as an input to perform well [21].

Deep learning has made great contributions back to other sciences by providing useful tools for processing massive amounts of data and making useful predictions in scientific fields. It is one of the areas of ML which learns representations from data with emphasis on learning successive layers of increasingly meaningful representations and works based on the concept of ANNs. The neural network basically comprises three layers: input layer, multiple hidden layers and output

layer. The number of layers which contribute to the data is called the depth of the model. There are different types of deep learning algorithms such as CNN, Long Short Term Memory Network, Recurrent Neural Network, Auto Encoder, etc [21]. All these algorithms use different types of neural networks to accomplish specific tasks. Deep learning algorithms require large amounts of computing power and information to solve complicated issues which can work with almost any kind of data.

3.1.1. Convolutional Neural Networks

Convolutional Neural Network (CNN) also called ConvNet is one of the most widely used types of deep neural networks applied to computer vision [20] that employs a mathematical operation called convolution in place of general matrix multiplication in at least one of their layers to analyze visual images by processing data with grid-like topology [22]. Convolutional networks are simply neural networks and have been tremendously successful in practical applications. In contrast to simple neural networks, CNNs consist of many layers with an input of images. Modern convolutional networks for object recognition provide a model of visual processing that neuroscientists can study. Natural images have many statistical properties that are invariant to translation. CNN is mainly used for major classification problems in which the features are automatically learnt from low level to high level on successive increasing layers of the network [11].

The operation of CNN algorithm has three fundamental stages [11]: multiple convolutions, non-linear operations and multiple pooling operations. In the first stage, multiple convolutions are performed using filters simultaneously to produce a set of linear activations. In the second stage, a nonlinear (e.g. rectified linear) activation function is performed which will directly return the input if positive and zero otherwise. On the other hand, pooling operation is used to reduce the size of the feature map at the third stage. Max pooling, average pooling, weighted average pooling are some of the pooling functions. Max pooling finds and takes the maximum value in a rectangular window, whereas average pooling takes the average of all the values in the rectangular window. At the end of CNN, the output of the last pooling layer is given as an input to the fully connected (FC) layer that each node in the FC layer is connected to all nodes in the preceding layer. FC layer is used to perform classification of the input data into different classes.

On the other hand, CNNs are trained through a process called back propagation which consists of 4 stages: forward and backward pass, loss function, and weight update [11]. The filter weights are initialized randomly. During the forward pass, the training images are passed to the network. The error rate is calculated using loss function which compares the network output with the desired output then back pass and weight update stages take place based on the error rate. The back propagation process can be repeated for several iterations until convergence occurs.

CNN is unique due to its architecture that performs segmentation, feature extraction and classification in one processing module with minimal pre-processing tasks on the input image. Minimal domain knowledge of the problem is considered to be sufficient to accomplish efficient pattern recognition tasks. This has been proven by a wide range of applications that are using CNN such as face detection, face recognition, gender recognition, object recognition, character recognition, and texture recognition [23]. CNNs involve many connections/layers including convolution, pooling and fully-connected layers, and realizes form of regularization [22]. Regularization is a function to prevent over-fitting of the model training, and it can be improved by adding a regularization strategy such as data augmentation or dropout, or by adjusting a weight decay coefficient. Figure 3.1 shows the general architecture of CNN.

3.1.1.1. Convolution Layer

Convolution layer is the main building block of a CNN that encompasses a series of filters or learnable kernels which aim at extracting distinct features from the input. Each kernel in this layer is used to calculate a feature map. The first convolutional layer extracts low-level meaningful features such as corners, edges, lines, and textures; whereas the next convolutional layer extracts higher-level features. Finally, the last layer extracts the highest-level features of the input data. Convolution layer uses an independent matrix filter that performs convolution operation to detect patterns in the image and end up with a specific feature map of the filter size. An Activation function called Rectified Linear Unit (ReLU) is applied to the convolution layer to get a rectified feature map of the input image. The convolution operation of CNN is denoted as:

$$S(t) = (x*w)(t) \quad (1)$$

Where x is the input, w is filter kernel and S is the output called feature map or kernel map.

3.1.1.2. Pooling Layer

Pooling layer is applied after each convolution block that helps to reduce the resolution of the previous feature maps through compressing features and computational complexity of the network. It adjusts the features robust to disorder, noise, and other small variations that ensures the network to focus on the most important patterns. In general, a pooling layer produces down-sampled versions of the input map and reduces the dimensionality of the feature maps used in the subsequent layers [22]. Max pooling is the most common approach in pooling layer, which reduces the size of the input image (sub-sampling) that makes the representation become approximately invariant to small translation of the input.

3.1.1.3. Fully Connected Layer

Fully connected layer is the last stage of the CNN topology that comprises of a generic multi-layer network. The last higher few layers are FC one-dimensional layers to all activations in the previous layer. After feature extraction is done by the convolution and pooling layers, we need to classify the data into various classes using FC neural network. FC layer is simply a feed forward neural networks that forms the last few layers in the network. It gets an input from the flattened output of the last pooling layer.

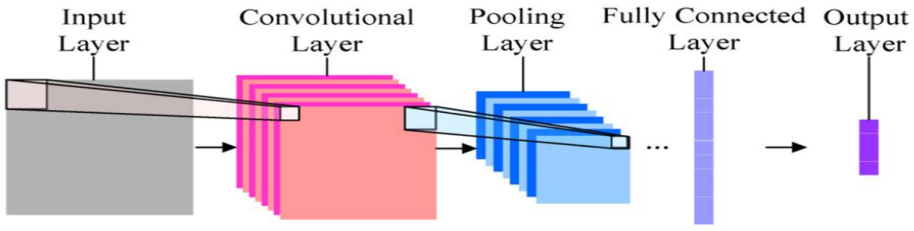


Figure 3. 1: General CNN Architecture [24]

3.1.2. Transfer Learning

In ML, a method where knowledge gained from a particular task is transferred to another related task to improve the learning process is called transfer learning [11]. Transfer learning is a

technique of using the knowledge of weights and its layers of an existing trained model to a new similar untrained model which speeds up the learning process of the new model [25]. During the training of the new model, one or more layers of the trained model are used on the problem of interest.

Transfer learning can be understood using the analogy of the Human Brain [26]. Humans keep on looking at a lot of things in their daily lives, but on seeing any new similar type of objects, we can easily identify those objects based on previous experience. Also, seeing new objects adds to our old experience to create a new experience. Thus, transfer learning is the concept of transfer of learning to learn new things on top of old experience. Another scenario, consider two people who want to learn playing piano. One person has extensive music knowledge through playing guitar, whereas the other one has no previous experience of playing music. The person with an extensive music knowledge will be able to learn piano in an efficient manner by transferring previously learned music knowledge to the task of learning to play piano [27].

To perform transfer learning, the concept of freezing some part of the layers is used to prevent the layer weights from being updated. It helps to ensure faster model training for the new data by ensuring that rather than training all the layers on getting new data, only train a specific set of layers or the newly added layers. Thus, by utilizing the experience of the pre-trained models, higher accuracy can be achieved even for smaller datasets and in a very short period of time. A technique called fine-tuning can also be used to train the model to ensure better accuracy as compared to allowing all the layers in the model to be trained from the scratch by providing initial random weights [26].

The CNN architectures like ResNet, AlexNet, VGG, and others are already trained on the huge image database of ImageNet consisting of 1000 classes that contains more than one million labeled high resolution images [11]. The knowledge gained from the already trained task is transferred to learn another new task, especially where there is lack of sufficient training data. As a result, it shows good performance in classification and the computational complexity is greatly reduced as the process didn’t start from scratch.

In this research, we have used VGG16 network architecture for transfer learning-based face recognition and applied to a new untrained model on the dataset we have prepared. VGG16 is a

very popular convolution-based network model in computer vision techniques which of 16 layers as shown in Figure 3.2. During the model training we have removed the top layers of the VGG16 model to fine-tune and added other layers such as Flatten, Dense, Drop, and Softmax layers for classification. Flattening is used to convert the x-dimensional arrays into a single column linear vector that is passed to FC layer, and then the dense layer adds the FC layer to the neural network. Whereas the drop layer is used to randomly drop some value to avoid over-fitting. It is noted that the ReLU activation function has been applied in all layers except the last layer [25]. However, Softmax is used at the last layer of the model for face recognition purpose. In general, transfer learning has mainly the following advantages:

1. No need of large datasets for training, and
2. Requires less computational power as compared to training a CNN from scratch.

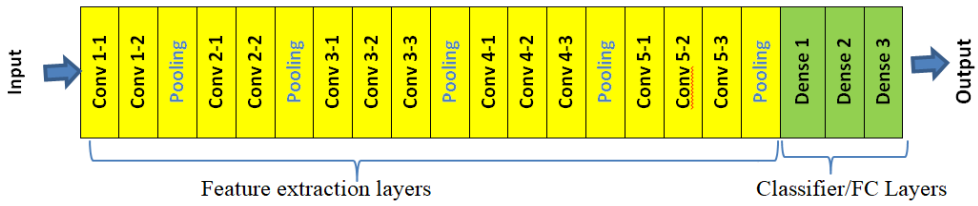


Figure 3. 2: VGG16 Architecture [11]

3.1.3. Multi-Task Cascaded Convolutional Neural Network

Multi-Task Cascaded Convolutional Neural Network (MTCNN) is a deep learning algorithm used for face detection and alignment tasks [28][29]. Both face detection and face alignment tasks play an important role in the application of face recognition. MTCNN has different advantages such as memory consumption is small, and can realize real-time face detection. It is more robust to light, angle and facial expression changes in an unconstrained environment and has a better face detection effect. When MTCNN starts processing an image, it performs image resizing operation to scale the original image to different scales to generate an image pyramid. These images with different scales are sent to the three sub-networks for training in order to detect different sizes of human faces and realize multi-scale target detection. The three MTCNN sub-networks are Proposal Network (P-Net), Refine Network (R-Net), and Output Network (O-Net) as shown in Figure 3.3.

A. P-Net is the first sub-network in the MTCNN framework with a basic architecture of a full connected convolution network (FCN). The main task of P-Net is to give a boundary box for each face bounds by generating a large number of target area frames [28]. The image pyramid constructed is used to extract the preliminary features and to ensure the accuracy of the frame through a FCN, and roughly obtain the face candidate frame and frame regression vector. Then the candidate frames are regressed by the frame, and finally the candidate frames with high coincidence are merged by non-maximum suppression (NMS) algorithm which is used to select the best bounding box from the multiple predicted bounding boxes [29].

B. R-Net is slightly similar to P-Net but includes more appropriate bounding boxes compared to P-Net. It has a more complex structure than the P-Net structure of the upper layer and helps to reject false candidates from the P-Net. R-Net makes further judgment on the output window of the upper layer, and uses border regression and NMS algorithm to exclude the face candidate frames with low score, so as to select several groups of locally optimal face candidate frames [29].

C. O-Net is a relatively more complicated convolution network than the above two sub-networks with one more convolution layer for outputting the final five facial features by supervising the face areas and regressing the facial features [29]. The five key facial landmarks produced by this sub-network are the left and right corners of the mouth, the nose, and both eyes.

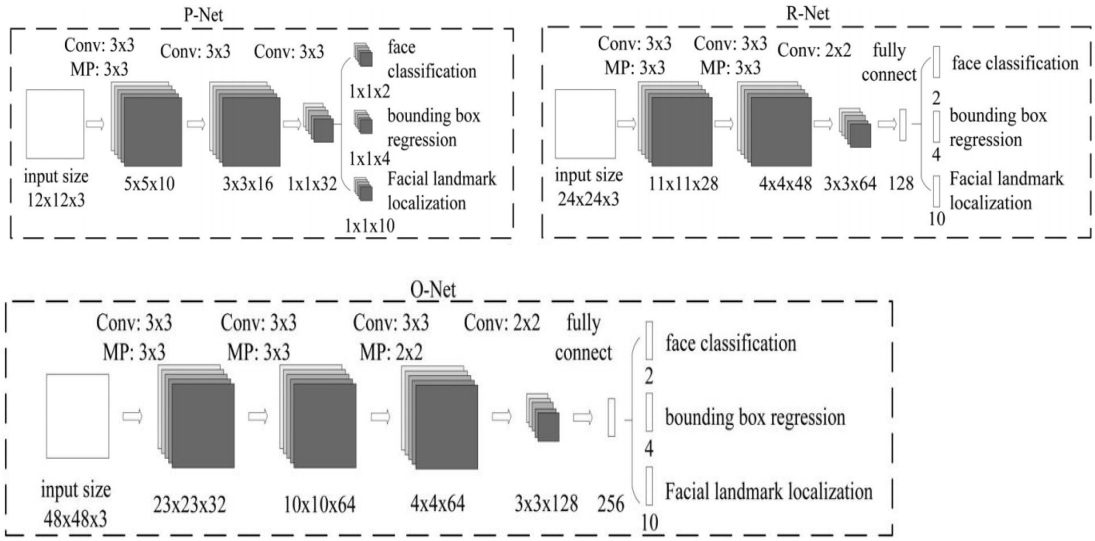


Figure 3. 3: MTCNN Architecture Framework [28]

There are three important tasks in MTCNN during the model training. These are face classification, bounding box regression, and facial landmark localization [28].

1. Face Classification: MTCNN sets the learning target as a two-class classification problem by indicating if the given image has a face or doesn't have a face.

2. Bounding Box Regression: For each candidate window, the bounding box regression technique helps to predict the offset between it and the nearest ground truth. There are four coordinates, including top, width, height, and left.

3. Facial Landmark Localization: like the previous task, facial landmark detection is formulated as a regression problem. During facial landmark localization, the five important key points such as the centers of the left and right eyes, the center of the nose, and the left and right mouth corners are detected and identified.

3.1.4. FaceNet

FaceNet algorithm, also called a model, is developed by Google researchers using Deep CNN which is used to extract high quality features from the face images [30][31]. FaceNet provides a unique architecture for performing face recognition, verification and clustering tasks. It maps each face image into Euclidean space such that the distances in that space correspond to face similarity or difference. FaceNet uses 2 types of CNNs architectures: (1) Zeiler & Fergus architecture and (2) GoogLeNet style Inception model [30]. The Zeiler & Fergus architecture is used for visualizing the training process of a CNN and introduces a novel visualization technique that gives insight into the function of intermediate layers and the operation of the classifiers. In the second architecture, multiple filters of different sizes are used simultaneously and their results are concatenated.

The FaceNet network consists of a batch input layer of a 160 x 160 pixels with 3 channels and a deep CNN followed by L2 normalization that calculates the distance of the vector coordinate from the origin of the vector space, which results in the face embedding. The face embeddings are the most important features of a person's face with 128-dimensional vector values used to train a face

identification model. Embedding is obtained from the level of similarity and differences in faces, so that if the face has a similarity the value will get closer, and if it is different the value will get far apart [30]. The architecture of FaceNet is shown in Figure 3.4.

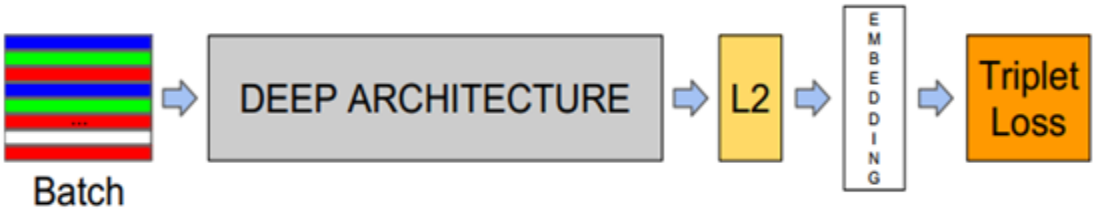


Figure 3. 4: FaceNet Architecture [30]

Another important aspect of FaceNet is its triplet-based loss function which comes after embedding during training [30]. It uses triplet loss function along with the deep CNN to learn the features of the face and achieves state-of-the-art accuracy. Triplet loss function is a function that gets closer the distance between an anchor (baseline) and a positive where the positive represents the same identity and it gets farther the distance between negative to indicate different identity in the given dataset as shown in Figure 3.5.

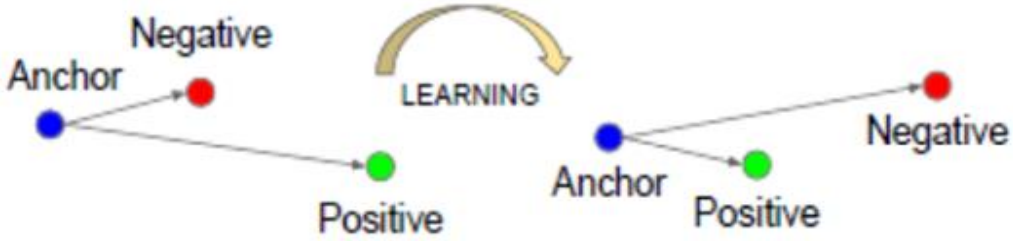


Figure 3. 5: Triplet Loss Function [30]

3.1.5. Support Vector Machine

Support Vector Machine (SVM) is a popular supervised learning algorithms used to solve classification as well as regression problems [32]. There are two types of SVM: linear and non-linear. Linear SVM is used for a dataset to classify linearly into two classes by using a single straight line, whereas non-linear SVM is used for non-linearly separated data, that could not be classified or separated by a straight line.

In practice, the SVM algorithm is implemented with a kernel that transforms an input data space into the required form. It uses a technique called the kernel trick in which the kernel takes a low dimensional input space and transforms it into a higher dimensional space. In other words, the kernel converts non-separable problems into separable problems by adding more dimensions to it. This makes SVM more powerful, flexible and accurate. SVM is widely used and has good performance in face recognition [33]. In this thesis paper we have used a linear kernel to classify our subjects after the embeddings have been obtained for facial recognition.

3.2. Face Recognition

Face recognition is a technique of identifying and verifying the identity of individuals in images using their face. It is the use of computer vision technology and related algorithms, from the pictures or videos to find faces, and then analyzing the identity [24]. Face recognition systems can be used to identify people in real-time, videos and photos. The human face is a sophisticated multidimensional structure that can convey a lot of information about the individual, including expression, feeling and facial features. It is a challenging task to analyze facial features effectively and efficiently due to its time and effort consumption. The face recognition problem can be categorized into two main phases [34]: face verification and face identification. Face identification is a one-to-many mapping for a given face against a database of known faces (i.e. who is this person?), whereas face verification is a one-to-one mapping of a given face against a known identity (i.e. Is this the person?).

There are different face recognition systems which are currently employed for numerous applications such as access control, law enforcement, information security, surveillance systems,

smart cards and attendance [34]. In the face recognition process, various tasks can be performed such as pre-processing of the input images, features extraction and classification. Removing any noise content, background illumination, and normalization of the input image is part of the pre-processing step. After the face images are pre-processed, important features are extracted using different ML/deep learning algorithms. Finally, a classification is done using classifier algorithms such as SVM, Neural Networks, and CNN [11].

3.2.1. Challenging areas in Face Recognition

Both face detection and recognition have some challenges for the successful implementation of models. These challenges are aging, illumination, partial occlusion, pose variance, background and expression [34]. Aging is an inevitable natural process during the lifetime of a person as compared to other facial variations, whereas occlusion can be natural or artificial obstacles in an input image. Occlusion affects the performance of a model when people deceive it either by the use of hair, scarves, sunglasses, or placing hands in front of faces. Pose variance is another obstacle in achieving successful face recognition due to the lack of standardized rules for taking a picture that makes it more difficult to distinguish and recognize the faces from images. Illumination is an observable property and effect of light. The effect of light on the images badly affects the face recognition system.

Chapter Four

Methodology

In this chapter, we covered the methodologies we have followed to conduct the thesis paper. First, we identified the problem available in ethio telecom and reviewed related works. After that data collection is done, and algorithms and tools are selected to build the best model. In addition to the model building, we also did the model implementation and how to use it. In the coming subtopics, we presented some of the details as follows.

4.1. Data Collection and Preparation

In ethio telecom, getting customer data is too difficult because of the security policy of the company, especially the data that is highly related with customer information. Due to this, we did not get customers' photos from the company. Thus, we tried to collect all the photos of the people from different sources such as the Internet (web image gallery, social media (Facebook), and YouTube), personal photo gallery and capturing photos from some friends. When we collect the photos from Internet, almost all the people are Ethiopian celebrities and we searched them one by one using their names on the web as well as social media then took the photos by cropping using snipping tool and saved to a folder created by their individual's name with a .jpg, .png and .jpeg file extension. Additionally when we collect from YouTube videos, it is by using the same way of cropping using a snipping tool. Similar way of gathering images from the personal photo gallery source has also been done. The dataset we have obtained was not in an organized way in the source that is ready for such a purpose. All the image data we collected tried to resemble the real images in the company in order to consider the natural varieties of ours from the other world such as skin color, structure of the face and others.

During the image data collection, we have considered the pose, illumination, occlusion and facial expression status of images. In addition, we have also tried to consider the number of images per an individual, i.e. for some people, we have very few images such as two, three, and others in the training set to make it similar to the company's data, because in the company there could be many customers with few images.

We have collected a total of 2596 images dataset for 124 identities and this dataset is split into training and testing datasets. Two types of dataset splitting have been done to conduct an experiment. The first one is with a ratio of 80% and 20% for training and test dataset respectively. In this ratio, 2068 images are used for training and 528 images for testing purposes. The second type of dataset splitting mechanism is the reverse of the earlier one with a total dataset modification. In this dataset splitting mechanism, we have reduced the number of images in the training set to two in each identity. However the number of images per identity in the testing dataset is not limited to a fixed number like in the training set. This type of dataset categorization helps to know how the model will be effective in the real environment. Because customers in the real environment may not have many images, rather most of the time they could have two or three images per identity, especially when the customer is subscribing for the first time. The categorization of the dataset is shown in Table 4.1.

#	# of Training images	# of Test images	Total # of images	Ratio (Train : Test)
1	2068	528	2596	80% : 20%
2	248	992	1240	20% : 80%

Table 4. 1: Dataset splitting

After the required data is collected, it has to be prepared for the selected algorithm to develop the model. Data preparation is a way of arranging and making suitable data for the algorithm used to build the model. When the input image is forwarded to the face recognition algorithms, it has to be in the right format such as .png, .jpg or .jpeg. Since we are using deep learning algorithms to recognize the face of individuals, the required part is detected, cropped, resized and aligned by the selected algorithms as part of the pre-processing, and then necessary features extraction from the face images will be done by the algorithms. Finally, a face classification or recognition will be done to identify the identity of the photo. All these tasks from pre-processing to classification will be accomplished automatically.

Another data that we have collected for the model deployment analysis is daily sales report of newly registered subscribers and market cost assessment. We have used a two months sales report of newly registered subscribers and calculated the average daily sales to estimate the requirements that help to implement the model. Using this data requires concurrent session or recognition time,

and model size forecast is done. In addition, the estimated cost used to implement the model is done by conducting market assessment and current items cost from different sources.

4.2. Model Building Environment

In this thesis work, we have mainly used a popular and flexible deep learning open-source library known as Keras consisting of Tensorflow as background engine and python programming language. To train the model, a personal computer (PC) environment is used which has a specification of Intel® Core™ i7-8665U CPU @1.90GHz Core(s), 8GB RAM and 64-bit windows 10. The experiment conducted using CNN with VGG16 transfer learning uses different parameters such as epochs, number of epochs per step, batch size to train the model repeatedly. Whereas the experiment that uses MTCNN and FaceNet doesn't have epochs, rather it is a one-time training process. In the first experiment, the purpose of using different epochs and steps per epoch is to earn better accuracy.

On the other hand, we have also used another environment which is a Google computer engine backend (Google colab) that provides Graphics Processor Unit (GPU) for fast graphics data processing with a dynamically assigned 2GB RAM. This is to check how the training speed is fast compared to the CPU processor and it is relatively faster, however the result of the model has similar accuracy.

4.3. System Model

In general, the face recognition process is carried out with five stages consisting of face input, preprocessing, feature extraction, face matching, and face recognition results. The input image contains a single face of individuals which is ready for detection by the algorithms. The overall process of the model is shown in Figure 4.1 and subsequently with some detailed descriptions for each component of the diagram.

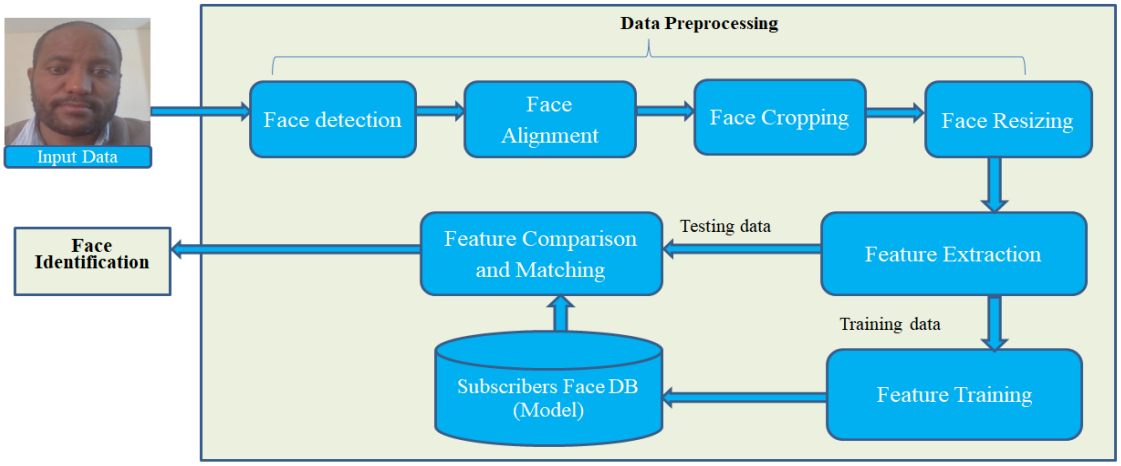


Figure 4. 1: Face Recognition Process Model

a. Face Detection

Face detection is a way of finding human faces from the given images that contain plenty of non-face objects within the image. It is a process of identifying and locating the presence of human faces in digital photos and videos. Different face detection algorithms usually begin their execution by finding out the parts of the face such as eyes, mouth, nose, and eyebrows to conclude whether a face is available in an image or not. The algorithms used in face detection determine whether the input images are positive images (i.e. face is available) or negative images (i.e. with no face). After it is identified and located, a bounding box is used to mark and select for the next action.

b. Face Alignment

The position of the input image is not always in an appropriate structure that is suitable for the algorithms to train a model. The position of the face and its landmarks has an effect on the face recognition process. Thus, a face analysis technique called face alignment is used to estimate a sparse set of specific points or landmarks that define the shape of the face. It is also defined as the process of identifying the geometric structure of faces in digital images, and attempting to obtain the alignment of the face based on translation, scale, and rotation. The estimation of landmarks may include the eyes, nose, jaw, eyebrows, mouth and others [35].

The alignment of faces could be done using the supplied algorithms, in this case the deep learning algorithms such as MTCNN.

c. Face Cropping

Another pre-processing step of the face recognition process is face cropping that helps the algorithm to work appropriately in the facial classification task. Face cropping is a process of taking only the identified face part by cutting and excluding the other parts of the image such as background, accessories and others [36]. After cropping the face from the original image, the new face image parts are then resized to a preferred size for the algorithm.

d. Face Resizing

Deep learning algorithms need all input images to have the same shape and size, because the graphics processing unit applies the same instruction to a batch of images at the same time in order to be super-fast [36]. The purpose of face resizing is to produce a smaller data size and rushes the processing time of the algorithm [37]. In OpenCV, to perform the face resizing we can use the function called resize() either to downscale or upscale the given face image so that the uniformity of the image's size will be constant throughout the dataset. For example, the input image size for VGG16 is 224 x 224, and for FaceNet is 160 x 160.

e. **Feature Extraction**

Once the pre-processing stage is complete, the feature extraction task continues to uniquely identify the image identity. Feature extraction is an important technique that highly affects the accuracy of face recognition [38]. In feature extraction, there are different challenges that inhibit the extraction of important feature components of the face such as pose, occlusion, varied facial expression, and illumination [39]. The effectiveness of the face recognition system depends upon the quality of the extracted features. This is because facial landmarks and fiducial points such as nose, eyes, mouse, etc. identified by the algorithms determine how features are presented accurately. Different algorithms use different ways of extracting the vital features of the image for further feature matching/classification purpose.

f. Feature Matching

After features are extracted from the face image, then they are going to use it to find out the matching points between the images. Image matching is the process of judging the similarity among the different input images by evaluating the resemblance and consistency of their features. Generally, there are two categories of image matching [40]: gray-based and feature-based image matching. The gray-based image matching method is simple and accurate but weaker for image changes such as illumination and scale change. Where as in feature-based image matching, features are extracted and quantified by some mathematical computations. The higher the robustness of feature extraction has the better of the correct feature matching and the accuracy of the model highly depends on the output of the matching of the image features.

In general, when we conduct an experiment of the model building, we provide an input image data to the algorithms (CNN in the first experiment and MTCNN in the second experiment) to detect the face part from the given image and then an alignment is done to an appropriate position to improve the recognition process. After the detection and alignment tasks have been done then these algorithms cut the face part and resize it according to the specific required size of each algorithm. Since all face inputs do not have the same size, thus resizing is important and mandatory. In this case, when we use VGG16 transfer learning the face is resized to 224x224 and for FaceNet it is 160x160. Finally, the feature extraction and matching tasks are performed to identify a particular person.

4.4. Tools Selection

To achieve the objective of this study, different tasks and methods are required, namely face detection, resizing, cropping, features extraction and classification. All these tasks are accomplished using various algorithms and tools to conduct a successful experiment and obtain better results. The required algorithms are described in chapter three whereas in this section, the selected tools are presented.

We have used different tools and open source libraries such as OpenCV, TensorFlow, Keras, and Python to build the face recognition model. These tools are freely available for all developers and researchers and they are briefly depicted in the following sub-sections.

4.4.1. OpenCV

Open Source Computer Vision (OpenCV) is a computer vision and ML library that was built to provide a common structure for computer vision applications and to influence the use of machine interpretation in the commercial products like facial recognition which is extensively used in today's world [41]. OpenCV is an image and video processing open-source library which is used for analysis like facial detection, face recognition, character recognition, license plate reading, photo editing, and advanced robotic vision. OpenCV has plenty of optimized algorithms. These algorithms can be used to detect and identify faces, objects, track moving objects in a video or in an image, follow eye movements etc. OpenCV gives support to TensorFlow, Torch, Caffe, and other deep learning frameworks [42] and has some advantages such as:

- It is fast as compared to other libraries, since it is written in C/C++
- Works better on system with lesser RAM
- Supports Operating Systems such as MacOS, Linux and Windows
- Provides communication between humans and computers

4.4.2. TensorFlow

TensorFlow is an open-source software library for ML that provides a collection of workflows to develop and train models using Python [43]. It was developed by Google Brain Team for internal use, and can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. TensorFlow helps to build and train state-of-the-art models with good performance, and it gives the flexibility and control with features like the Model Sub-classing Application Programming Interface (API) and Keras functional API for the creation of complex topologies. It has a flexible ecosystem of libraries, tools, and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML applications.

4.4.3. Keras

Keras is the most used deep learning framework open-source software library that provides a python interface for ANNs and acts as an interface for the TensorFlow library [44]. Keras is designed to enable fast experimentation with deep neural networks and focuses on being user-friendly, modular, and extensible. It is an API designed not for machines but human beings that follows best practices for reducing cognitive load: offers simple & consistent APIs, minimizes the number of user actions required for common use cases, and provides clear & actionable error messages. Keras is an industry-strength framework built on top of TensorFlow that can scale to large clusters of GPUs or TPU case.

4.4.4. Python

Python is an interpreted high-level general-purpose object-oriented programming language that has extensive scientific library [45] used for web development, ML, and complex data analysis. Python's language constructs as well as its object-oriented approach aim to help programmers write clear logical code for small and large-scale projects. It is dynamically typed and garbage-collected that supports multiple programming paradigms including structured particularly procedural, functional, and object-oriented programming. Due to its comprehensive standard library, it is often described as a "batteries included" language [42].

Chapter Five

Experimental Results and Model Implementation

In this chapter, two different separate experiments have been conducted using deep learning algorithms and results are obtained and analyzed. In the first experiment, we have used CNN with VGG16 transfer learning mechanism, and in the second experiment MTCNN, FaceNet and SVM algorithms are used to have the best model with better face identification accuracy that plays a great role in preventing subscription fraud in telecommunications industry in general and particularly in ethio telecom. In face recognition model building; image quality, pose, occlusion, and illumination contribute significant effect in the recognition performance. Different image processing algorithms could have various levels that help to improve the recognition accuracy of the face recognition model.

5.1. Results

5.1.1. Using CNN with Transfer Learning

The architecture of CNN contains in general a total of eight blocks of layers. The first five are convolutional blocks with max pooling and the remaining three are fully-connected layers with ReLU and Softmax. In transfer learning, the convolutional block layers are set to freeze. We did not train our data in these layers because the VGG16 model transfers the knowledge of the previously trained model from the ImageNet data. The only layers that are trainable are the last few layers, specially the fully-connected layers of the model.

In CNN with VGG16 transfer learning, the input image size is set to 224×224×3 pixels where the 224×224 is the size of the height and width, and the number 3 indicates the number of channels (the colored images). The size of the input image is different in each convolutional block. For example in convolutional block 1, there are two layers with image size of 224×224, and in convolutional block 2, there are two layers with 112×112 image size. Whereas in blocks 3, 4 and 5, there are three convolutional layers in each block with different image sizes. Every block layer of the convolution is followed by *Max pooling.* Finally, the fully-connected layer takes an input of

the one-dimensional input which flattens the 2-dimensional array into one-dimensional data. In VGG16 all the convolution kernels are of size 3x3 and maxpool kernels are of size 2x2 with a stride of two.

In deep learning models to work well generally there has to be a large amount of data. However, it is difficult to always get enough data for model training. Therefore to minimize such issues a data augmentation [46] technique is used which is a solution to the problem of limited data. In this experiment we have used this technique in order to increase the training dataset during training time.

We have conducted different approaches to obtain better accuracy of the face recognition by changing the number of epochs and steps per epoch in a given platform as shown in Table 5.1. When the number of epochs and/or steps per epoch increases the accuracy of the model increases to some extent but it takes more time to build the model.

#	No. of epochs	Steps per epoch	Accuracy
1	20	200	68.56%
2	30	100	66.10%
3	30	500	70.45%
4	50	65	67.23%
5	100	300	72.54%

Table 5. 1: Output of CNN with VGG16 using different scenarios

5.1.2. Using MTCNN and FaceNet

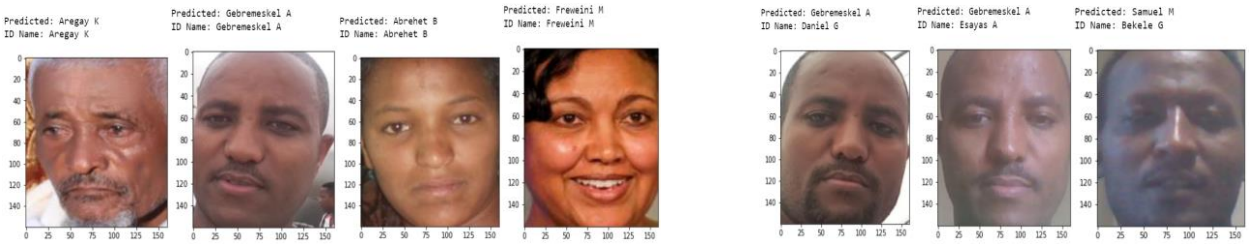
In this specific experiment, the MTCNN method is used to detect and extract the faces part from the input image and then aligns to an appropriate position which makes ready for the next step, feature extraction. Whereas the FaceNet is used to extract high-quality features from faces, called face embeddings, that can then be used to train a face recognition model. Face embeddings is a vector that represents the features extracted from the face. This can then be compared with the vectors generated for other faces. Each face is represented by 128 measurements. After all the features of the face input images are extracted the SVM classifier is used to classify people from the input image features provided by the FaceNet. In this specific experiment we have obtained different outputs depending on the dataset categorization as depicted in Table 5.2.

#	# of Training images	# of Testing images	Accuracy
1	2068	528	99.24%
2	248	992	98.69%

Table 5. 2: Experimental outputs for MTCNN and FaceNet

In the first experiment, we have used 80% of the dataset for training and the rest for testing purposes. The output of the experiment achieves an accuracy of 99.24% which is the highest accuracy from all the conducted experiments.

In the second experiment, the dataset splitting method is the reverse of the first experiment but the number of images is totally changed because the main intention of the experiment is to have every identity with few numbers of images. In this case, each identity has two images in the training dataset but different numbers of images in the testing dataset. Finally, the result of this experiment achieves 98.69% accuracy. Sample outputs of the model are shown in Figure 5.1. In Figure 5.1 (a), all the subscribers are legitimate and the model correctly predicts the service requestor, which means the name in the system is the same as the name in the ID card (Kebele ID). However, in (b) the subscribers are attempting to commit a subscription fraud but the model identifies them as they are not right subscribers.



a) Legitimate Subscribers b) Subscription Fraud Attempt

Figure 5.1: Sample outputs of the face recognition model

5.2. Model Implementation

Here in this sub-section, some of the considerations during the face recognition model implementation are discussed in addition to how the model prevents fraud. These considerations are mainly focused on ethio telecom, Ethiopia.

5.2.1. How the model prevents Subscription Fraud

The proposed model will probably have a great contribution in the process of preventing subscription fraud in telecommunications industries, especially for telecom companies that have systems that collect customer's photos. When the customer goes to the sales office for service subscription, he/she directly contacts the sales person to get the service and then the salesperson asks the customer whether he/she has an existing service number subscribed before or is new to any telecom service subscription of that specific company. If the customer is not new and is requesting additional services, the sales person requests the customer for his/her service number or account number of any service number that belongs to him/her. Then the sales person searches the customer's detail information from Customer Relationship Management (CRM) and/or e-CAF systems, and cross-checks with his/her profile information such as identification card or other authorized documents. After that according to the company's policy, the sales person could either give or deny the requested service and then end the process.

However, if the customer is claiming that he/she has never been registered for any telecom services before, the sales person is forced to check whether this customer is really new or not by applying the face recognition model. If the customer is available in the system and the information in the system is different with the profile information at hand such as identification card, the customer is really trying to commit a subscription fraud and will not get registered for any service. However, if the information in the system is the same as with the information at customer's hand, even though he/she is trying to deceive the company, it is considered as a legitimate customer and he/she could get or deny the requested services depending on the company's policy. The detail flow diagram of the process is depicted in Figure 5.2.

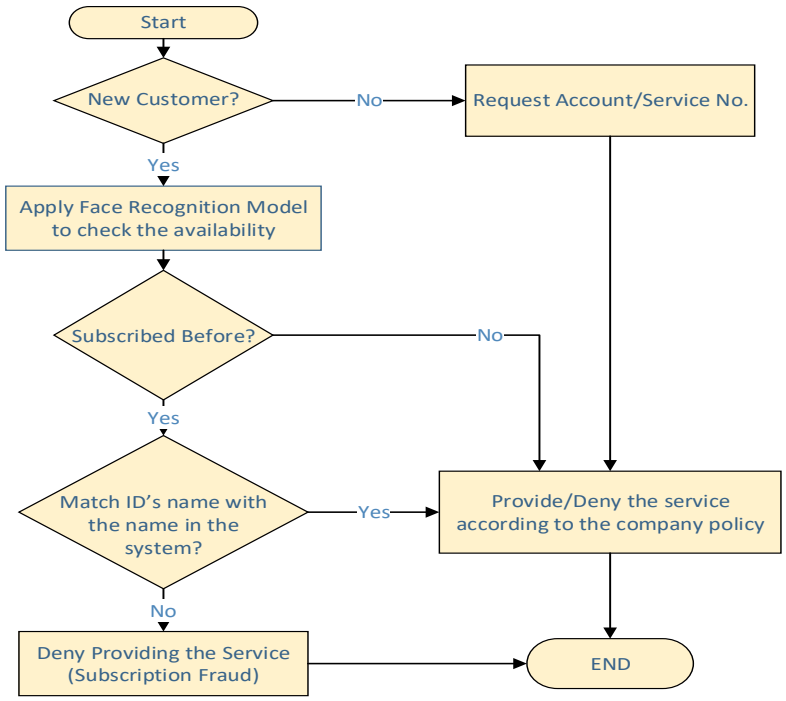


Figure 5. 2: Proposed Flow Diagram of Subscription Fraud Prevention model

5.2.2. Storage Requirement

In the case of ethio telecom, when the customer gets subscribed for a service for the first time, customer's information such as photo, signature, scanned application documents, ID card, and agent signature are stored in the system. To register for a single customer, on average a total of 150KB storage space is required. For a single customer's picture, the size ranges between 15KB and 50KB. Currently, the e-CAF system has a total capacity of 46TB data of all the registered customers. Out of this, by calculating the average capacity of the picture, the total size of the customers' photo is assumed to be around 10TB. Probably the total existing data could decrease during data cleansing before the model training, because it includes photos of fraudulent customers.

During the model implementation, we are going to take the existing image data capacity to have an effective face recognition system. Whatever the customers' image capacity is, the size of the model reduces amazingly during the model construction as compared to the original data size. In this thesis work, we have used customers' images with a total capacity of 47.2MB. After training

these images, the model size becomes 5.16MB. This implies that when we build a model for the currently available data size, the size of the model will be 1.09TB.

According to the company's sales report, the average daily new customer registration is 17,080. To get the average daily sales report we have used a two months sales report. In this case, for each customer at least one portrait is taken by the system and saved in the storage system. However, during model building, two images will be taken (1 captured and 1 from the ID card). This implies on average 1.06GB (i.e. 2*17,080*32.5KB) of data is stored daily to the system, where 32.5KB is the average size of a single image. Thus, the face recognition model size will be increased daily 0.12GB, and roughly it will become 43.8GB per year.

5.2.3. Concurrency

To determine the concurrent operation of the system, we have checked two important things: (1) the number of new services sold per day and (2) performance test of the model by providing different numbers of images at a time for recognition. We took a sales report of two months that contains only newly registered customers. The report has a total of 836,920 newly registered customers that took different types of services from the company. At a time a single customer can register for only one service or multiple services. If the customer is registered for more than one service, the face recognition system considers it as one identity because it checks only once before the service provisioning takes place. Thus, there is no need to use the face recognition model for each service numbers' subscription of a single subscriber at a time.

Once the customer is registered to any of the telecom services, another time when he/she comes to subscribe for additional services, the sales person requests him/her service number or account number belongs to him/her. Then the sales person could be able to search the customer's profile from the system and cross-check his/her identity against the documents that the customer provided such as ID card or other authorized information that uniquely identifies him/her. Checking the identity of the customer before getting subscribed for the service has the probability of saving time and storage space of the company's system whenever the subscriber is not new but claiming to be registered as a newcomer for any of the telecom services. That means the subscriber is trying to commit a subscription fraud by using different techniques such as counterfeit documents or other means.

As described in the above storage requirement part, the number of new subscribers that get registered daily are around 17,080. If we further divide this number of customers to per hour, it becomes 2,135. These 2,135 customers per hour means around 36 customers get the service per minute. Therefore, what we can understand from this sales report is that the model should support at least 36 customers per minute.

Performance of the model has a determinant factor for the fast delivery of the services. The model should have the capability of recognizing multiple customers at a time. The experiment is done on a local PC laptop and the output of the experiment is shown in Figure 5.3. For the model to work concurrently, an application software is required that supports users to login and work concurrently. The application software helps to access the code/script of the face recognition model.

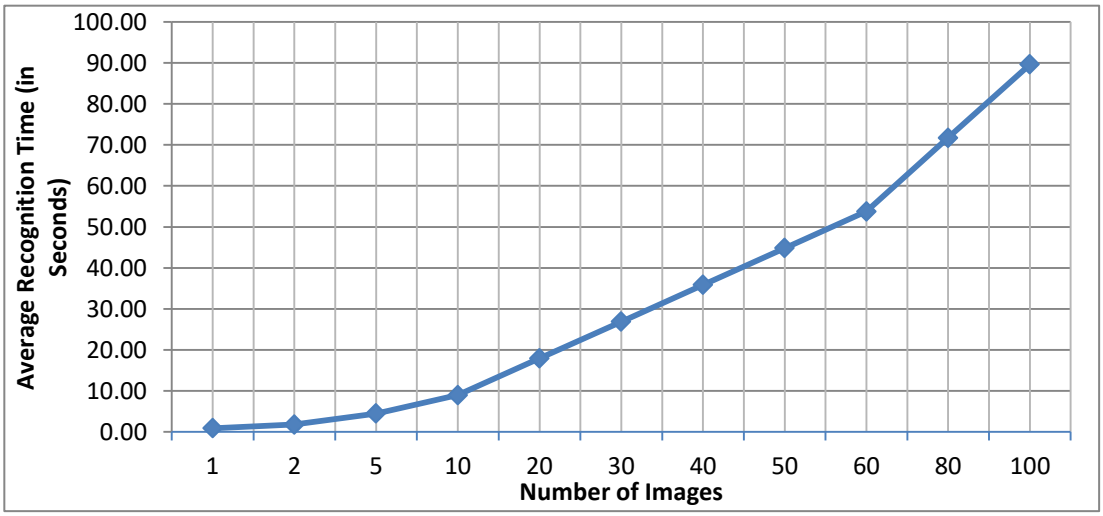


Figure 5. 3: Result of concurrent operation of requests

According to our experiment, to recognize 40 images/customers at a time 36 seconds are required. That means within less than one minute (i.e. 54 seconds), 60 images/customers are recognized. This shows that the model is sufficient enough to comply with the company's requirement. To increase the response time of the recognition, it is advisable to use higher performance servers that use GPU processors. Some studies announce that the performance of GPU is faster than the CPU and in some scenarios it is 4-5 times faster [47]. Thus, it would be more preferable to deploy the solution on a server that has a GPU.

5.2.4. Mode of Deployment

Basically, there are two types of deployment modes: client-based and server/web-based. In our case, we have proposed a web-based mode of deployment for the face recognition model. In a web-based way of deployment, the face recognition solution will be deployed at the central part of the server. Then all clients will access the system through web applications from the centralized server. The use of web-based mode of deployment has various advantages such as saves storage space and maintenance time. That is, the specified model capacity will be stored at the server side storage. But if it is put in every single client, the required storage space will be too much though the response time could be faster. The other advantage of web-based deployment mode is it will have an ease and consistent model update whenever the model update is needed. It is also easy for maintenance during system failures. However, the hardware on the server side is often more powerful than the client side.

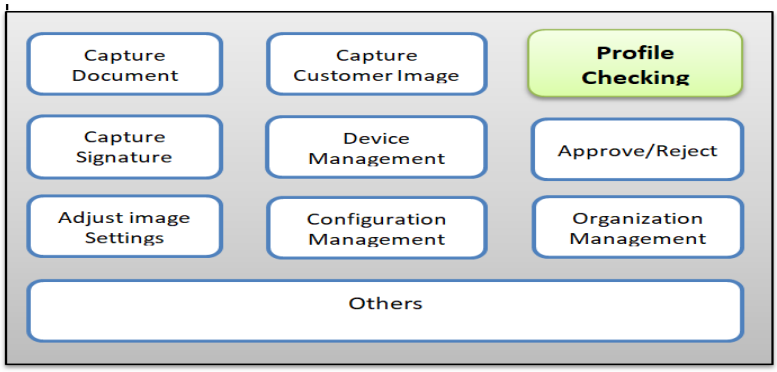


Figure 5. 4: Integration of Face Recognition function within e-CAF system

If the company decides to deploy the model, it is advisable to integrate it with the existing e-CAF system by including as one module as shown in Figure 5.4. The module can be called as "Profile Checking" and added to an e-CAF system architecture that helps to identify the subscriber's presence in the system with his/her face before any other activity. The e-CAF system is a web-based system that receives requests and response from/to different clients throughout the company. Thus, the deployment mode of the face recognition model has to be similar to the existing system.

5.2.5. Cost Estimation

Face recognition system implementation requires lots of hardware materials such as high performing servers, data storage, switches, cameras, and quality cables with sufficient bandwidth. Usually the hardware materials are the main capital as well as operational expenses in most face recognition system deployment. In the case of ethio telecom, the infrastructure that is used for customer's photo capture, transmission media/bandwidth, and image data storage are already available in the e-CAF system. Whenever the company needs to implement the face recognition system, the cost has to be considered for the items such as servers, storage, load balancer, switch, and etc. In addition to these hardware materials cost, the company should consider the face recognition software development cost as well as integration cost with existing systems such as e-CAF and other related deployment costs.

We have done an assessment on the required materials to estimate the cost of deployment from different sources such as online market [48][49], ethio telecom's current projects item price, as well as market analysis on ICT Companies to develop and deploy for the application software. The items listed in Table 5.3 are the main items used for the model implementation and there could be few additional items that their cost is not significant. According to our assessment, the rough estimated cost needed to implement the system is $993,649. This could be probably the minimum amount that the company can expense to deploy the model at this time just according to the current cost assessment. Whenever the time of deployment delays, the expenditure could vary in terms of materials cost and other related costs.

The system is proposed to be deployed in a high availability (HA) mode and will have a disaster recovery (DR) system to ensure continuity of the business without interruption whenever frequent failures or catastrophic disasters occur. We have proposed four servers: two for the main system and HA, one for DR and one for test-bed environments. Similarly, the storage will be distributed proportionally to the servers' implementation. Both the load balancers and the switches will be deployed as a redundancy. In addition to the network design of the system, a face recognition system requires at least 100 Mbps bandwidth to deploy and work efficiently [50]. However, the existing e-CAF system uses a bandwidth of 1Gbps and thus we will use the same bandwidth as the existing system which is much better than the minimum requirement.

#	Item Name	Spec	Qty	Estimated Unit Price	Estimated Total Price
1	Server	Dell PowerEdge R740 2U Server, 2x12-Core, Gold 6146, 3.2Ghz, 192GB RAM, 2x1.2TB HDD	4	$14,262	$57,048
2	Storage	OceanStor Dorado 6000 V6(2U, Dual Ctrl, SAS, AC\240V HVDC,1TB Cache,4*(4*12Gb) SAS, 25*2.5")	4	$26,472	$105,888
		Disk Enclosure: SAS Disk Enclosure (2U,AC\240V HVDC, 2.5", Expansion Module, 25 Disk Slots, Without Disk Units)	4	$1,078	$4,312
		Disks: 3.84TB SSD SAS Disk Unit(2.5")	15	$2,663	$39,945
3	Load Balancer	Load balancer: Radware D-5208-Perform, SLB+GSLB, Throughput 12Gbps, 2*10G SFP+(With Fiber Module), 8*1GE, Dual AC Power, 100-240V	2	$3,863	$7,726
4	Switch	Cisco SX550X-24F 24-PORT 10G SFP+ stackable managed switch - Managed - L3	3	$3,650	$10,950
5	Application Software	Application development + deployment cost	1	$729,167	$729,167
		Integration with e-CAF	1	$25,000	$25,000
6	Database software	Microsoft SQL Server 2019 Enterprise (2 Core)	1	$13,613	$13,613
	Total				$993,649

Table 5. 3: Detail Cost Estimation of the model deployment

Chapter Six

Conclusion and Future Works

6.1. Conclusion

Telecommunication fraud is among one of the most severe threats to revenue and quality of service in telecommunication industries. It is the deliberate abuse of telecom services to get benefit for those who did the fraud and has a great impact on the company's revenue generation. Fraudsters use various techniques to get telecommunication services without the intention of paying or with partial payment. Among the various telecom fraud types, subscription fraud is one of the well-known fraud types that negatively impacts telecommunication services. It is also an enabler for other fraud types such as SIMBox fraud, roaming fraud, and IRSF fraud.

The objective of this thesis work was to develop a model and propose the deployment scenarios to prevent subscription fraud in the telecommunications industry, specifically in ethio telecom, Ethiopia. Even though it is too difficult to get the image data from ethio telecom due to the company's customer privacy policy, to achieve this objective a limited number of images have been collected from other different sources such as the Internet, personal photo gallery, and capturing photos from friends. These images were prepared by cropping the face parts and ready for input to the selected algorithms. Using these images a model was built, evaluated, and the one with better accuracy is proposed for implementation.

In this thesis work, two separate experiments were conducted; CNN with transfer learning as the first experiment and MTCNN, FaceNet and SVM algorithms together as the second experiment. In the first experiment, an accuracy of 72.54% was achieved and the later one has obtained an accuracy of 99.24%. Therefore, the model which is built using the second experiment is recommended for implementation in the company to prevent subscription fraud.

Face recognition system is proposed to prevent subscription fraud whenever subscribers try to deceive to get telecom services. When the subscriber requests a telecom service, the salesperson asks whether the requestor is new or a customer of the company. If the requestor acts as a new subscriber then the face recognition system will be applied, otherwise checking every requestor could be a waste of time. The system is proposed to be deployed in a web-based mode at the

server side for the current size of 1.09TB and is calculated to be increased daily by 0.12GB. Finally, a roughly estimated deployment cost is proposed based on the current market analysis and costs around $993,649 by considering the main materials in addition to the software.

6.2. Future Works

Our face recognition model could have great importance to ethio telecom in preventing subscription fraud as well as other fraud types. There could be some works for future researchers as:

- Further researches can collect large number of images from the company and train the model that may increase the accuracy of the model
- Hopefully, the images of the company could have high quality than the images we have used, so it is better to train a model using those images, especially for CNN with transfer learning algorithms
- Future researchers can find a way of how the size of the model can be further reduced and ways to increase the number of concurrent sessions with less response time
- Our work is done on a low performance PC but future researchers can prepare high performance test-bed servers to test the model's recognition response time
- Researchers can work on other specific parts of the face or other uncovered techniques
- Future researchers can work on other deep learning algorithms to get an enhanced model

References

[1] A. J. Hussain, and E. Chew, "Data Mining and Telecommunication Fraud Detection using Artificial Neural Networks," 2015.

[2] PrivSec Report, "Telecommunications: the battle against fraud," 2017, Available: https://gdpr.report/news/2017/05/29/telecommunications-battle-fraud/ [Accessed Feb 12, 2021].

[3] Global Telecom, "Communications Fraud Control Association Announces Results of 2019 Global Telecom Fraud Survey," 2019.

[4] G. Koi-akrofi et al., "Global telecommunications fraud trend analysis," International Journal of Innovation and Applied Studies, Vol. 25 No. 3 , pp. 940-947, 2019.

[5] New Business Ethiopian, "Ethiopia arrests 32 engaged in telecom fraud," Feb 24, 2019, Available: *https://newbusinessethiopia.com/crime/ethiopia-arrests-32-engaged-in-telecom-fraud/* [Accessed Jan 11, 2021].

[6] F. M. Kau and O. P. Kogeda, "Impact of Subscription Fraud in Mobile Telecommunication Companies," pp. 42–47, 2019.

[7] T. Derebe, "A Comparative Analysis of Machine Learning Algorithms for Subscription fraud Detection: The case of ethio telecom," Feb 21, 2020.

[8] I. Awoyelu et al., "Fraud Detection in Telecommunications Industry : Bridging the Gap with Random Rough Subspace Based Neural Network Ensemble Method," vol. 6, no. 10, pp. 6–15, 2015.

[9] T. H. Le, "Applying Artificial Neural Networks for Face Recognition," Hindawi Publishing Corporation Advances in Artificial Neural Systems, vol. 2011, 2011, doi: 10.1155/2011/673016.

[10] U. Aiman and V. P. Vishwakarma, "Face Recognition Using Modified deep learning neural network," 8th ICCCNT, no. 40222, pp. 3–7, 2017.

[11] R. M. Prakash et al., "Face Recognition with Convolutional Neural Network and Transfer Learning," Proc. 2nd Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2019, no. Icssit, pp. 861–864, 2019, doi: 10.1109/ICSSIT46314.2019.8987899.

[12] L. G. Kabari, D. N. Nanwin, and E. U. Nquoh, "Telecommunications Subscription Fraud Detection using Artificial Neural Networks," doi: 10.14738/tmlai.36.1695.

[13] C. M. Held, C. A. Perez, and P. A. Este, "Subscription fraud prevention in telecommunications using fuzzy rules and neural networks," vol. 31, pp. 337–344, 2006, doi: 10.1016/j.eswa.2005.09.028.

[14] A. Abdallah et al., "Fraud detection system: A survey," Journal of Network and Computer Applications, vol. 68, pp. 90–113, 2016.

[15] GSMA, "GSM Association Official Document FF.21 - Fraud Manual," Ver 9.0, 2012.

[16] A. Aljarray and A. Abouda, "Analysis and Detection of Fraud in International Calls Using Decision Tree," Almadar Journal for Communications, Information Technology and Applications AJCITA, 2015.

[17] I. Murynets et al., "Analysis and Detection of SIMbox Fraud in Mobility Networks," IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, 2014.

[18] H. Kahsu, "SIM-Box fraud detection using data mining techniques: The case of ethio telecom," p. 84, Nov. 3, 2018.

[19] Thales, "Identity fraud in Telecommunication," Available: *https://www.thalesgroup.com/en/markets/digital-identity-and-security/mobile/inspired/identity-fraud-in-telecomunication* [Accessed Jan 06, 2021].

[20] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning," Book in preparation for MIT Press, 2015.

[21] Aditya Sharma, "Differences between Machine Learning & Deep Learning," datacamp, Nov. 15, 2018, Available: https://www.datacamp.com/community/tutorials/machine-deep-learning? [Accessed Mar 02, 2021].

[22] I. Namatēvs, "Deep Convolutional Neural Networks: Structure, Feature Extraction and Training," Inf. Technol. Manag. Sci., vol. 20, no. 1, pp. 40–47, 2018, doi: 10.1515/itms-2017-0007.

[23] A. R. Syafeeza et al., "Convolutional neural network for face recognition with pose and illumination variation," Int. J. Eng. Technol., vol. 6, no. 1, pp. 44–57, 2014.

[24] Y. Li, and S. Cha, "Face Recognition system," 2018.

[25] A. K. Dubey and V. Jain, "Automatic facial recognition using VGG16 based transfer learning model," J. Inf. Optim. Sci., vol. 41, no. 7, pp. 1589–1596, 2020, doi: 10.1080/02522667.2020.1809126.

[26] Chirag Goel, "Face Recognition using Transfer Learning," June 09, 2020, Available: https://medium.com/@chiraggoelit/face-recognition-using-transfer-learning-9986728c443d [Accessed May 18, 2021].

[27] S. J. Pan and Q. Y. Fellow, "A Survey on Transfer Learning," IEEE Transactions on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345–1359, 2010, doi:10.1109/tkde.2009.191.

[28] K. Zhang et al., "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks," IEEE Signal Process. Lett., vol. 23, no. 10, pp. 1499–1503, 2016, doi: 10.1109/LSP.2016.2603342.

[29] N. Zhang et al., "Research on face detection technology based on MTCNN," Proc. - 2020 Int. Conf. Comput. Network, Electron. Autom. ICCNEA 2020, pp. 154–158, 2020, doi: 10.1109/ICCNEA50255.2020.00040.

[30] F. Schroff et al., "FaceNet: A unified embedding for face recognition and clustering," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 07-12-June-2015, pp. 815–823, 2015, doi: 10.1109/CVPR.2015.7298682.

[31] F. Cahyono et al., "Face recognition system Using FaceNet Algorithm for Employee Presence," 4th Int. Conf. Vocat. Educ. Training, ICOVET 2020, pp. 57–62, 2020, doi: 10.1109/ICOVET50258.2020.9229888.

[32] JavaTpoint, "Support Vector Machine Algorithm," Available: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm [Accessed June 02, 2021].

[33] P. Navin, Y. Singh, "Support vector machines for face recognition", IRJET Volume: 02 Issue: 08. 1521, Nov 2015.

[34] M. Lal et al., "Study of face recognition techniques: A survey," Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 6, pp. 42–49, 2018, doi: 10.14569/IJACSA.2018.090606.

[35] C. Álvarez Casado and M. Bordallo López, "Real-time face alignment: evaluation methods, training strategies and implementation optimization," *J. Real-Time Image Process.*, no. 0123456789, 2021, doi: 10.1007/s11554-021-01107-w.

[36] J. Duran, "A Quick Guide on Preprocessing Facial Images for Neural Networks using OpenCV in Python," 2019, Available: *https://medium.com/yottabytes/a-quick-guide-on-preprocessing-facial-images-for-neural-networks-using-opencv-in-python-47ee3438abd4* [Accessed October 04, 2021].

[37] N. H. Barnouti, "Improve Face Recognition Rate Using Different Image Pre-Processing Techniques," American Journal of Engineering Research (AJER), no. April, 2016.

[38] X. Li and H. Niu, "Feature extraction based on deep-convolutional neural network for face recognition," *Concurr. Comput. Pract. Exp.*, vol. 32, no. 22, p. 1, 2020, doi: 10.1002/cpe.5851.

[39] S. R. Benedict and J. S. Kumar, "Geometric Shaped Facial Feature Extraction for Face Recognition," 2016 IEEE International Conference on Advances in Computer Applications (ICACA), pp. 275–278, 2016.

[40] Y. Liu, X. Xu, and F. Li, "Image feature matching based on deep learning," *2018 IEEE 4th Int. Conf. Comput. Commun. ICCC 2018*, pp. 1752–1756, 2018, doi: 10.1109/CompComm.2018.8780936.

[41] OpenCV, "About" Available: *https://opencv.org/about/* [Accessed Sep 22, 2021].

[42] M S Sruthi et al., "A Fast and Accurate Face Recognition Security System," 2021 J. Phys.: Conf. Ser., 2021, doi: 10.1088/1742-6596/1916/1/012185.

[43] Wikipedia, "TensorFlow," Available: *https://en.wikipedia.org/wiki/TensorFlow* [Accessed July 19, 2021].

[44] Wikipedia, "Keras," Available: *https://en.wikipedia.org/wiki/Keras* [Accessed July 19, 2021].

[45] Primoz Podr, zaj Boris Kuster, "Face Detection and Face Recognition in Python Programming Language," proceedings of the 7th ICIA2018, pp. 1–7, 2018.

[46] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *J. Big Data, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.*

[47] E. Buber and B. Diri, "Performance Analysis and CPU vs GPU Comparison for Deep Learning," 2018 6th Int. Conf. Control Eng. Inf. Technol. CEIT 2018, no. October, pp. 1–6, 2018, doi: 10.1109/CEIT.2018.8751930.

[48] ebay, "Computer Servers," Available: *https://www.ebay.com/itm/401602191358?hash=item5d815b1ffe:g:LDkAAOSwfgZbo7E-* [Accessed October 03, 2021].

[49] Digital Software Market, "Microsoft SQL Server 2019 Enterprise (2 Core)," Available: *https://www.digitalsoftwaremarket.com/product/microsoft-sql-server-2019-enterprise-2-core/* [Accessed October 28, 2021].

[50] AllGoVision, "Datasheet Face Recognition," 2015.