ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL SCIENCES

Mobile Based Expert System for Diagnosis of Cattle Skin Diseases with Image Processing Techniques

Bezawit Lake

A Thesis Submitted to the Department of Computer Science in Partial Fulfillment for the Degree of Master of Science in Computer Science

Addis Ababa Ethiopia

October 2019

ADDIS ABABA UNIVERSITY

COLLEGE OF NATURAL SCIENCES

BEZAWIT LAKE

Advisor: Fekade Getahun (PhD)

This is to certify that the thesis prepared by *Bezawit Lake*, titled: *Mobile based Expert System for Diagnosis of Cattle Skin Diseases with Image Processing Techniques* and submitted in partial fulfillment of the requirements for the Degree of Master of Science in Computer Science complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Name	Signature	Date
Advisor: Fekade Getahun (PhD)	____________	________________
Examiner: _________________________	____________	________________
Examiner: _________________________	____________	________________

Abstract

Cattle population is critical socioeconomic assets in a nation like Ethiopia where the society depends on farming and animal husbandry. However, there is huge loss of livestock population by a disease that undermines the efforts towards achieving food security and poverty reduction. Many expert systems have been developed for the diagnosis of cattle disease. The diagnosis starts by collecting information about symptoms, signs and other related issues. In most of them, this information is obtained from the person using text dialogue. Every person has different ways of expressing the same thing, which results, in the inconsistency of description lead to an incorrect diagnosis.

To address this problem, we propose an approach for cattle disease diagnosis by integrating image processing using deep learning with an expert system. The proposed system has an expert system and an image processing component. The symptom identified by naked eyes are represented using image and its category is identified by the image processing component. The image processing component consists training and classification phase. In the training phase images collected from different source are preprocessed and feed to the classification model. The classification model used is a convolutional neural network with three convolutional and two fully connected layers. In the classification phase the trained model is used to classify the input images. The expert system have reasoning, knowledgebase and user interface component. The user interface allows communication between the system and the user. The knowledgebase contains information and facts required for diagnosis. The reasoning component reaches a final diagnosis conclusion based on classification results and other related information.

The developed classification model trained on 3990 dataset collected from different sources. To increase the dataset we apply different augmentation techniques. We split the dataset into 90% for training and 10% for testing. The model classifies the input symptom image with 95 % accuracy. The entire system has been evaluated by veterinarians and people having cattle farming, the analysis shows that the system is effective to diagnosis cattle disease.

Keywords: Cattle disease diagnosis, Expert system, Image processing, deep learning, Convolutional neural network, Location information.

Acknowledgement

As always I don’t know, what I did to deserve your blessing. I don't think we should spend our life praying for things, He give us what we deserve. But I do believe we should thank God for what He's given to us. I want to thank you, Lord, for life and all that's in it. Thank you for the day and for the hour, and the minute. Thank you for giving me everything. Thank you, thank you, and thank you.

I would like to pass my sincere thank you to my advisor Dr Fekade Getahun for his unlimited support, easy working environment and his constructive advice. He was always there whenever, I had a question. He consistently allowed me to be efficient and leads me in the right direction. Thank you for everything and wish you happy life. Thank you.

I would also like to thank Mr. Bekele Worku for giving me this chance to learn. I must express my gratitude for the experts who were involved in this research Dr Bedlu Amare, Dr Ligalem Abera and Dr Abdi Mohammed.

Finally, I want to express my gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study. Gasheye thank you for your encouragement and support. My brother Wude, Hymi, my friend Alazar, this accomplishment would not have been possible without you. Thank you. My friends Engu, Hiwi thank you for being there when am feel blue. Thank you every one for being with me. Thank you.

Table of Contents

List of Figures iii

List of Tables iv

List of Algorithm v

CHAPTER 1 : INTRODUCTION 1

1.1 Background 1

1.2 Motivation 2

1.3 Statement of the problem 2

1.4 Objectives 4

1.5 Methods 4

1.6 Scope and Limitations 5

1.7 Application of results 5

1.8 Thesis Outline 6

CHAPTER 2 : LITERATURE REVIEW 7

2.1 Introduction 7

2.2 Cattle in Ethiopia 7

2.3 Skin Disease 8

2.4 Skin Disease Diagnosis Techniques 12

2.5 Expert system 13

2.5.1 Components of Expert System 14

2.5.2 Phase of building expert system 15

2.6 Digital Image Processing 18

2.6.1 Image 18

2.6.2 Steps of Digital Image Processing 21

2.6.3 Image Classification 24

CHAPTER 3 : RELATED WORK 34

3.1 Introduction 34

3.2 Expert system with text input 34

3.3 Diagnosis system with image input 36

3.4 Expert system with text and image input 37

3.5 Summary 40

CHAPTER 4 : DESIGN OF IMAGE BASED CATTLE DISEASE DIAGNOSIS APPROACH 41

4.1 Introduction........ 41
4.2 IBCDD design consideration........ 41
4.3 The IBCCD System Architecture........ 41
4.3.1 User interface........ 42
4.3.2 Location Extraction........ 43
4.3.3 Knowledge Engineering........ 43
4.3.4 Reasoning module........ 46
4.3.5 Image Analysis........ 48
4.4 Summary........ 55
CHAPTER 5 : IMPLEMENTATION AND EVALUATION........ 56
5.1 Introduction........ 56
5.2 Dataset preparation........ 56
5.3 Development Environment........ 58
5.4 System Prototype........ 58
5.5 System evaluation........ 59
5.6 Comparison with deep neural network........ 63
CHAPTER 6 : CONCLUSION AND FEATURE WORKS........ 64
6.1 Conclusion........ 64
6.2 Contribution........ 64
6.3 Feature work........ 65
REFERENCE........ 66
Appendix A: Interview Questions........ 73
Appendix B: sample rule and case used in the system........ 74
Appendix C: System evaluation questionnaires........ 75

List of Figures

Figure 2-1 Cattle infected by LSD 10
Figure 2-2 Cattle infected by Wart 11
Figure 2-3 Cattle infected by Ringworm 12
Figure 2-4 architecture of expert system 14
Figure 2-5 representation of image using coordinates 19
Figure 2-6 representation of image using matrix 19
Figure 2-7 fundamental steps of DIP 21
Figure 2-8 result of convolution operation 29
Figure 2-9 result of pooling operation 30
Figure 4-1 proposed system architecture 42
Figure 4-2 disease diagnosis conceptual model 45
Figure 4-3 database design 46
Figure 4-4 architecture of the reasoner component 47
Figure 4-5 architecture of image analysis module 49
Figure 4-6 classification model architecture 51
Figure 4-7 workflow of training algorithm 52
Figure 5-1 LSD infected cattle 57
Figure 5-2 LSD image after transformation 57
Figure 5-3 images after augmentation technique applied 58
Figure 5-4 prototype user interface 59
Figure 5-5 plot of training and testing accuracy 60
Figure 5-6 plot of training and testing loss 60
Figure 5-7 confusion matrix of the model 61
Figure 5-8 plot of DNN training and testing accuracy 63

List of Tables

Table 2:1 occurrence of Lumpy skin disease 9
Table 2:2 most common skin diseases 9
Table 2:3 Comparison of different segmentation technique 22
Table 2:4 difference between deep learning 27
Table 3:1 related work summary 38
Table 5:1 collected data before and after transformation 57
Table 5:2 records of evaluator evaluation 62

List of Algorithm

Algorithm 4-1 location extraction algorithm 43
Algorithm 4-2 case retrieval algorithm 47
Algorithm 4-3 reasoner algorithm 48
Algorithm 4-4 image resizing algorithm 49
Algorithm 4-5 algorithm for normalizing image 50
Algorithm 4-6 algorithm for filter weight initialization 53
Algorithm 4-7 algorithm for feedforward propagation 54
Algorithm 4-8 algorithm for backward propagation 55

CHAPTER 1 : INTRODUCTION

1.1 Background

Medical diagnosis is the process of determining disease based on symptoms and signs [13]. It is an essential part of disease management and prevention. Whether it is affecting a single or an entire society disease must be controlled. A small outbreak of disease can spread and affect the entire population. As a result quick detection and diagnosis is critical to prevent any disease outbreak from further spread.

Ethiopia has the highest draft livestock population in Africa continent and is ranked first in Africa and tenth in the world [10]. Ethiopia has 59.5 million heads of cattle which contributes 40% of the annual agricultural output and 15% of the total gross domestic product [10]. Livestock are critical socio-economic assets in a nation like Ethiopia where many of the population depend on farming and animal husbandry. However, potential economic benefits are constrained by many reasons.

The most important constraints to cattle productions are widespread endemic diseases including viral, bacterial, and parasitic infestation [10]. Endemic and trans-boundary cattle diseases are usually transmittable leading into an outbreak in a short period of time [1]. The outbreak of disease reduce productivity and quality of animal products (skin, hides, milk etc.). If it is not controlled it may cause international trade restriction of animal and animal products. So it is essential to know cattle diseases and solutions to increase the productivity.

Artificial intelligence (AI) is playing a significant role in medical diagnosis. There are different AI technology used in medical domain like expert system, digital image processing etc. Expert system (ES) is a computer system or software that emulates decision-making ability of human experts in a specific field [2]. ES combines knowledge base with inference engine. This program acts as an intelligent consultant or advisor in a particular environment, based on stored knowledge. Digital image processing (DIP) is a technique which manipulates images using computer system [6]. DIP start from low level processing noise removal to high level processing like classification, recognition etc. The integration of different technology from AI enable to design a system with high performance [15, 38].

1.2 Motivation

The Ethiopian government designed an export development strategy to generate foreign exchange [10]. The export commodities include in the strategy are livestock and their products. In the 1990s livestock sub-sector is the second major source of foreign currency through the export of live animals, meat, leather, hides, and skin [9]. Despite the amount of livestock, there has been a decline in export earnings from livestock products in comparison to other African countries. The highest portion of decline is directly related to skin diseases or secondary damage that occurs when the animal scratches itself to relieve the itching associated with some of these diseases [9]. From the total product, 65% of skin and hides are often rejected because of poor quality [9]. The high prevalence of diseases in cattle requires serious attention to minimize the effect of the problem it causes.

AI techniques are distinguished with their successful results in relation to human and animal health studies [1, 2, 4]. They provide a new perspective to solve problems which are known in the medical field. Therefore, the main motivation is the desire to minimize the decline by implementing AI technologies in the animal health sector in Ethiopia.

1.3 Statement of the problem

To the society whose life depends on farming and animal husbandry, the wellbeing of the animal is a major concern. Diseases are the main threats that affect normal function in animals. In Ethiopia, the diseases cause huge mortality and morbidity in the cattle population and the lost in international livestock markets is about 1.5-2.5 Billion birr annually [10].

According to UNDP (United Nations Development Program), [7] many of the diseases in Ethiopia can be controlled by vaccine technology, by timely recognition of the disease followed by the acquisition of the pharmaceuticals. However, the remoteness of the livestock holder and shortage of infrastructure facilities to support health services delivery, the diseases are not controlled easily. Consequently, in Ethiopia, the majority of disease intervention consists of mass inoculations following outbreaks rather than preventive measures.

Beyene et al. [25] studied veterinary drug prescribing practices in Adama district. They find out there is a problem of correct diagnosis and drug administration based on clinical signs observed and diseases diagnosed, which could lead to irrational drug use. The prescribing practice for antimicrobial shows deviation from the standard recommended by WHO. The findings had also shown there are problems in generic prescribing, incorrect diagnosis and standardized patient case book handling and low prescriber educational status.

Currently, many of the diagnosis systems are either text or image based on acquiring symptoms encountered. The use of text description accuracy is dependent on the person understanding and describing the conditions that occurred efficiently. The use of image description is not dependent on anyone and it is efficient form of description. Using only images as form of description lacks major diagnosis steps to follow. According to Ararsa Dugma [24] in the investigation of any animal disease, the veterinarian must undertake a careful and thorough clinical examination to recognize the nature of the affection. For completeness and accuracy of diagnosis, the following things should be considered patient data, immediate/present history, past history, management, and environmental history. We need to have a system where symptoms can be described in efficient way for correctness of diagnosis result.

There are research works that attempt to design an expert system for cattle disease diagnosis [1, 2, 4]. Their work targets a person who can understand and describe the symptoms that occurred in the cattle using text. However, the expert who understands the cases is very low in the rural area of Ethiopia in which most cattle are found. The ratio of veterinarians to animals is 1: 500,000 [34] which is not sufficient to provide an efficient and on-time diagnosis.

While the main aim of diagnosis system is giving on time and efficient diagnosis based on occurred symptoms, the works done so far are not satisfactory because:

- Not designed for novice user who are the one close to the cattle’s e.g. Farmers.
- Text based symptom description, its accuracy is dependent on understanding of the person who describes the symptoms.
- The interaction to the system is not user friendly, require many selection and form filling.
- They are not implemented for mobile which can be more feasible than computers

Diagnosing diseases through text symptom description could lead to an incorrect diagnosis and difficult for someone who doesn’t have veterinary knowledge. So, developing a mobile application that doesn’t require assistance from the expert to describe the symptoms will help the farmer, pastoralists and anyone who don’t have medical background. In this research work, we aim to integrate technology from AI like expert system, Digital image processing, and deep learning to design cattle disease diagnosis system. This is believed to have an advantage on the animal health care specifically on cattle to improve healthcare services.

1.4 Objectives

General objective

The main objective is to design a system that can diagnosis cattle disease and gives treatment recommendation based on the encountered symptoms.

Specific objectives

In order to achieve the general objective we identify the following specific objectives.

- Identify requirements for designing the diagnosing system.
- Gather knowledge from domain experts.
- Collect image data from different sources
- Analyze the data collected and choose suitable knowledge representation model.
- Study and select suitable image analyses techniques.
- Choose suitable learning method
- Design the system architecture
- Prototype development
- Test and evaluate the prototype.

1.5 Methods

The research follows a design science research methodology because it produces a system that can diagnosis cattle disease. In order to achieve the specified objectives, the following list of methodologies will be employed.

Literature Review: A detailed review and assessment will be done on the expert systems, knowledge representation, modeling, and inference engine. Techniques and methods in Digital image processing are also exploited. Review is made on cattle disease and their treatment. From the review of these works, appropriate techniques and methods will be selected.

Data collection: We will collect images from different veterinary clinics in Ethiopia and other secondary source. Other information is gathered by conducting interviews and document analysis.

Tools and Development Environments: Different free and open source tools will be used during prototype development. Java programming language with android studio will be used to implement the expert system. The image processing will be done in Anaconda Keras Environment with python programing language.

Testing and Evaluation: The proposed solution will be evaluated in terms of its goals and performance. Machine learning evaluation methods and user evaluation will be used to evaluate the system.

1.6 Scope and Limitations

This study is aimed to find out if a system, integration of different methods from AI would bring efficient system to diagnosis cattle diseases. It is going to focus on the following two aspects.

- Building deep learning classification model
- Design a way to integrate to expert system

Disease symptoms which cannot be identified by naked eyes and not possible to capture by camera is not included in this study.

1.7 Application of results

The result of this study will bring efficient and timed diagnosis for cattle population. It will benefit the society whose life depend on agriculture and pastoralist. For cattle owners living in remote areas since they don't need to take the sick animal long distance for identifying the disease. Animal health professionals who are working in rural towns can use it to give an efficient diagnosis. Investors who are investing on animal farming can use this system. Veterinary and animal science experts are also beneficiary from the system in that it can be used as a decision support tool and to train students. In addition, it will also increase the livestock export market when diseases are controlled earlier before affecting too much cattle population. It can be also be used for knowing the occurrence of epidemic diseases.

1.8 Thesis Outline

The rest of the paper is organized as following order. Chapter 2 reviews concepts and methods relevant and related to the proposed approach. Chapter 3 presents related works. Chapter 4 presents the proposed diagnosis approach. In chapter 5 implementation and experimental results are presented. Finally, we conclude the overall work presented in this research work, presents contributions of the study and draws future directions in chapter 6.

CHAPTER 2 : LITERATURE REVIEW

2.1 Introduction

This chapter attempts to give background about cattle population in Ethiopia, the disease occurred in Ethiopia cattle, expert system architecture and component, development phases of expert system and digital image processing. Different methods and techniques of digital image processing will also be discussed.

2.2 Cattle in Ethiopia

Ethiopia is rich in animal genetic resources, both in diversity and population. The agricultural sample survey report of (CSA, 2016/17) revealed that about 59.5 million heads of cattle, 60.90 million sheep and goats population are found in the country. According to Tweldemedhn mekonnen [11], 98.20% of the total cattle population in Ethiopia are local breeds while hybrid and exotic breeds accounted for 1.62 and 0.18%, respectively. Ethiopia is a place for 28 recognized indigenous cattle breeds [45] in the continent. The diverse agroecology, cultural and ethnic diversity, a long-lasting agricultural practice and farming systems in the country have contributed to the diversification of the population.

Ethiopian Indigenous cattle can be broadly classified into four categories Large East African Zebu, Small East African Zebu, Sanga, and Zenga. Large East African Zebu mostly inhabit North West, South and Central highland of Ethiopia and include Begait, Boran and Arisi cattle [45]. Small East African Zebu include Bale, Jem-Jem, Harar, Ogaden, Sheko/Smad, Adwa, Jigjiga, Goffa, Guraghe, Hammer, and Ambo cattle. The Sanga breed group includes Anuak and Raya Azebo Cattles. The Zenga breed group includes Fogera, Northern Shire, Adwa and parts of Agame cattles.

Ethiopian indigenous cattle breeds have unique morphological features that distinguish them from other cattle. These include horn shape, horn size, and body size. In addition to physical features, non-visible traits such as productivity, disease and climatic stress resistance differ among breeds. These characteristics are largely the result of natural and human selection. Some breeds are already known for their unique adaptive attributes like Sheko cattle, good economic performances of Ethiopian Boran cattle. One of the well-known outstanding features of Ethiopian cattle is Trypanosomosis resistance [45].

Despite the potentials of diversified genetic resource, the huge loss of livestock population through disease undermine the efforts towards achieving food security and poverty reduction [20]. Drug administration and control authority of Ethiopia classified livestock diseases into infectious diseases, non-infectious diseases, skin conditions, pediatric problems, obstetrics and gynecology problems, ophthalmologic and Ear, Nose and Throat (ENT) disorders as well as acute/emergency problems [44]. In this research paper we will focus on skin diseases.

2.3 Skin Disease

Skin has a complex structure, being composed of many different tissues. The most important functions are control of body temperature, protects the body from physical damage and bacterial invasion etc. The animal skin comes in different kinds of textures and forms. There is the dry warty skin of toads and crocodiles, the wet slimy skin of fish and frogs, the hard shell of tortoises and the soft supple skin of snakes. Mammalian skin is covered with hair, that of birds with feathers, and fish and reptiles have scales. Pigment in the skin, hairs or feathers can make the outer surface almost any color of the rainbow. Despite their difference in texture and forms all animals are affected by skin disease. Skin is one of the first systems affected when an animal becomes sick.

Skin disease in cattle are conditions that cause inflamed, irritated or scaly skin, hair loss, changes in pigmentation of the skin and visible growths etc. [50]. Cattles are affected by various skin problems, some of which are easy to cure while others are more complicated with zoonotic importance. Due to a wide variety of diseases in cattle, the root cause of the diseases are often complex and can evolve over time. The existence of various skin diseases has been reported from many parts of Ethiopia. Lumpy skin disease (LSD) is the major reported epidemic diseases in different parts of Ethiopia. The occurrence of LSD in different regions of Ethiopia is shown in Table 2.1. Its transmission from infected to healthy animal is high, as a result it mostly occurs as an outbreak. Because of lack of information in disease occurred in one area its epidemic capability is significant. The epizootic characteristics of this disease has close association with climatic condition like heavy rains [22].

Table 2:1 occurrence of Lumpy skin disease

Zone	Years of reported outbreaks				
	2007	2008	2009	2010	2011
Addis Ababa			3	7	1
Afar			3	2	2
Amhara	92	68	35	40	22
Ben. Gumuz	3				5
Gambela				1	9
Oromiya	95	154	219	286	160
SNNP	18	18	14	32	17
Somali			3	9	4
Tigray	7	8	2	18	13

There are other common skin disease mainly reported in different parts of Ethiopia due to their consequences [19, 20, 21, 39]. Table 2.2 shows the most common skin disease in Ethiopia.

Table 2:2 most common skin diseases

NO	Location	Skin disease
1	Hawassa	LSD, Ringworm, Acariasus, Pediculosis, Dermatophilosos
2	Ambo	Wart, Ticks, Lice, Mange mites, Dermatophilosos, Photsentization, Branding
3	Gondar	LSD, Tick, Lice, Dermatophilosos, Mange
4	Adama	Ticks, Pediculosis, Demodicosis and Dermatophilosis
5	Bure	LSD, Pediculosis
6	Tigray(Tsegede,Welkayte, Kafta-Humera)	LSD, Lice, Tick, mange, mite

Lumpy skin disease (LSD): is an infectious, eruptive, occasionally fatal disease of cattle caused by a virus of the family Poxviridae [41]. As shown in Figure 2.1, it is characterized by nodules on the skin and other parts of the body. The nodules are well circumscribed, round, slightly raised, firm, and painful. The virus has important economic implications since affected animals tend to have permanent damage to their skin, lowering the commercial value of their hide. Additionally, the disease often results in reduced milk production, poor growth, infertility, abortion, and sometimes death. The skin nodules contain a firm, creamy-gray or yellow mass of tissue [42].

In Ethiopia, LSD was first observed in 1983 in the north-western part of the country (south-west of Lake Tana) [46]. The disease has now spread to almost all of the country. Because of the wide distribution of the disease and the size and structure of the cattle population in Ethiopia, it is likely that LSD is one of the most economically important livestock diseases in the country [46].

The control of LSD can be achieved through vaccination, restriction of animal movement and eradication of infected and exposed animals. However, this requires adequate financial, infrastructural and human resources, and information systems. Under the prevailing conditions in Ethiopia it has not been possible to implement all these strategies and thus vaccination has been adopted as the most important practical approach to LSD control for many years [22].

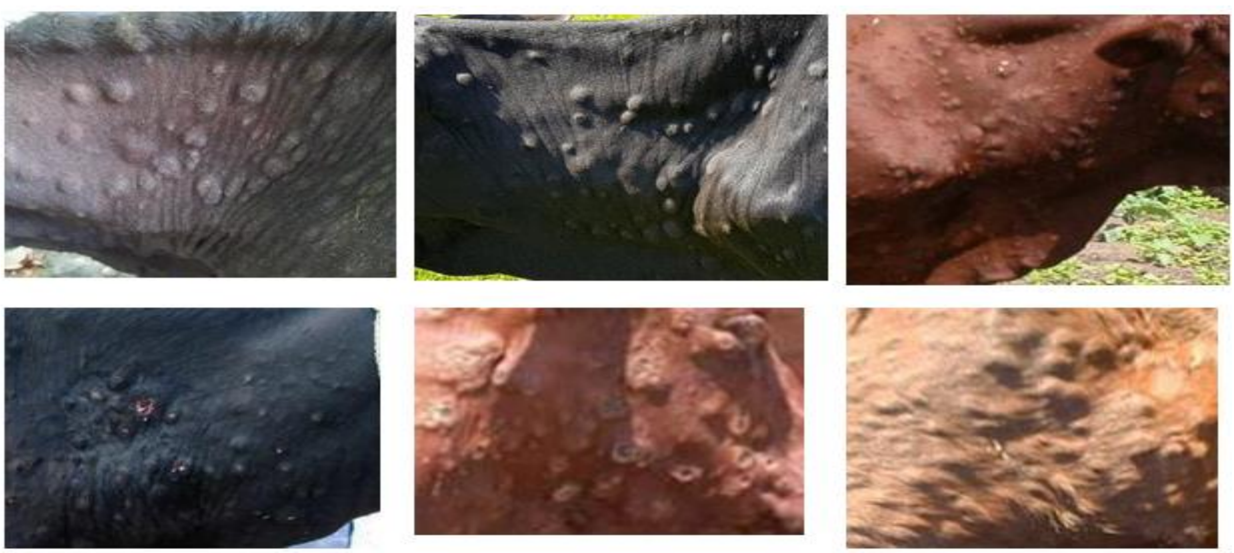

Figure 2-1 Cattle infected by LSD

Bovine Papillomatosis (Warts): are caused by an infectious and contagious virus (bovine papilloma virus, BPV) that spreads via contact from infected cattle to non-infected cattle [42]. It is a contagious neoplastic viral disease of animals characterized by the presence of multiple skin tumors or growths particularly, on head and neck areas. Although it regresses spontaneously, some cases may take a prolonged period and/or extended to malignant form. It can be exhibited as benign nodular lesions, fingerlike projections or cauliflower-like small growths on the skin arising from stratified squamous epithelium that may appear solitary or in multiples. The common sites for the development of cutaneous warts are head, eyelids, ears, neck, dewlap, brisket, shoulders and legs, occasionally on the back, para-genital region and along the lower line of the abdomen. Warts may be congenital or acquired. They are viral in origin and tend to be Species-Specific and to be most common in young animals. Young cows in winter are frequently affected in the skin of the eyelids and along the lower line of the abdomen, but the growths often drop off spontaneously from these positions when the young animals are turned out to grass in the early spring [41]. The symptoms shown when cattle is affected by Wart is shown in Figure 2.2.

Warts are the most common tumors affecting cattle, with most cases seen on the head and dewlap between 6 and 24 months of age. The causal virus can be spread by physical contact or equipment such as halters or milking machine. Lesions vary from flat, wide based warts to cauliflower-like growths. Extensive growths that fail to resolve may be seen in immune suppressed animals (e.g. persistent BVDV infection).

Figure 2-2 Cattle infected by Wart

Dermatophytosis (Ringworm): is a transmissible infectious skin disease caused by Trichophyton verrucosum, a spore forming fungi [41]. Ringworm is a fungal infection of the skin with a worldwide distribution. The disease has significant economic consequence to the farmer since growth rates are affected in the active stage of infection and the disease causes hide damage. Ringworm is common in young stock and is readily transmitted to humans (zoonosis). The symptoms of Ringworm as shown in Figure 2.3, is greyish lesions which are slightly raised, well-circumscribed, and more common on the head and neck but may extend over much of the body. In calves most commonly found around eyes, on ears and on back, in adult cattle chest and legs more common [42].

Figure 2-3 Cattle infected by Ringworm

2.4 Skin Disease Diagnosis Techniques

Veterinary disease diagnosis relies on knowledge of Anatomy, Physiology, Pathology and Animal behavior, skills in the methods and techniques of clinical examination, clinical sign and pathogenesis of the diseases [24].

Disease problems in veterinary medicine are presented to the clinician through the medium of the owner's complaint, which is a request for professional assistance by giving the infected animal information. Cattle health experts use symptoms and signs as clues that can help determine the most likely diagnosis when illness is present. In order to make a good diagnosis the doctor will go through a process that involves several steps, allowing them to gather as much information as possible. The steps of the diagnostic process fall into three broad categories. The first step is Initial Diagnostic Assessment, which the expert gather information which is relevant to diagnosis the cattle infected. The causes of various skin diseases requires a detailed history, because many skin diseases that look similar are differentiated based on interpreting clinical signs and historical patterns. A complete general history including information about prior illnesses, vaccinations, husbandry (housing, feeding practices, etc.), changes in attitude and food consumption, exposure to other animals must be obtained. The second step is diagnostic Testing, performance, interpretation, and communication of test results. Finally referral, consultation, treatment & follow-up, Physician follow-up, referrals and consults will be given based on the diagnosis result.

Clinical skin disease investigations are conducted by examination of the skin of each animal through visual inspection and palpation. The kind of symptoms shown when cattle are diagnosed with skin disease is lusterless dry brittle hair, hair loss, flaking skin, rough hair coat, matted hairs, a ring-like lesion on the skin, and crust wound on the skin, etc. In our study, we include disease which is enough to be diagnosed by shown symptoms and signs which can be captured by a camera. Palpation is a method of examination in which the examiner feels the size, shape, firmness or location of something (of body parts when the examiner is a health professional). So, for the inspection method we use image analysis techniques and for the palpation, we use text information.

2.5 Expert system

Artificial Intelligence (AI) refers to the activity of building intelligent systems. An intelligent system is a system that can reason about a particular domain, making a decision and have problem-solving capability [3]. A system or an agent can be said to be intelligent when the agent's performance cannot be distinguished from that of a human performing the same task. The core problems of artificial intelligence include programming computers for certain traits such as reasoning, problem-solving, perception, learning, planning, synthesis, and classification, etc. The first major and successful Artificial Intelligence research technologies were expert systems.

Expert system (ES) is one field of Artificial Intelligence that mimics the process of thinking and expert knowledge to solve a specific problem [2]. The term expert system could be applied to any computer program which is able to draw conclusions and make decisions, based on the knowledge it has. The problems in which expert systems are being developed are those that require considerable human expertise for their solution. Examples of such problem domains are medical diagnosis, financial advice, product design, etc. [27].

2.5.1 Components of Expert System

According to Engidu.et.al [1] an expert system may have components like knowledge base, knowledge acquisition, reasoning engine, user interface, and learning module as shown in Figure 2.4.

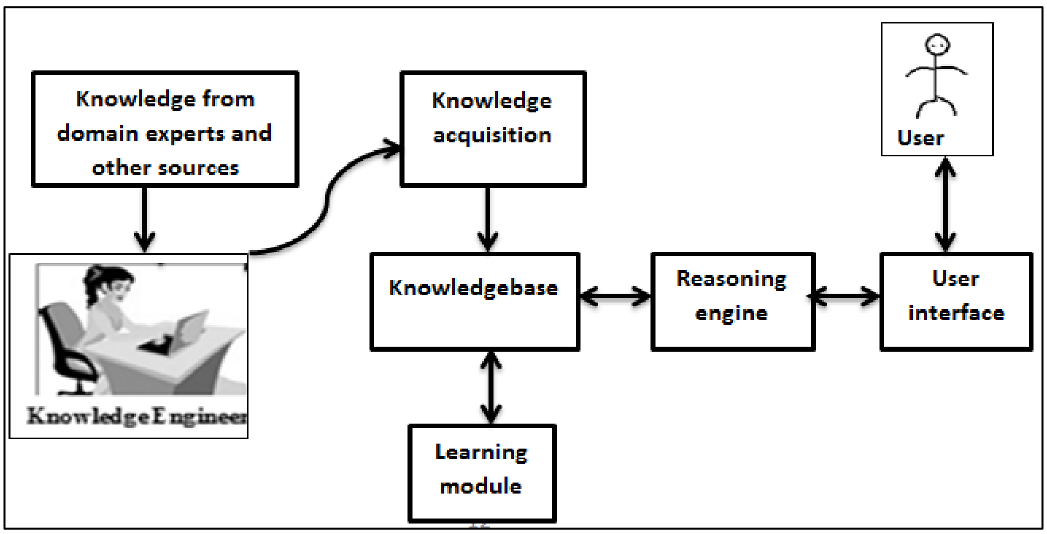

Figure 2-4 architecture of expert system

Knowledge Acquisition is a process of gathering relevant information in a specific problem domain. The knowledge can be obtained from domain experts, medical history retained in clinics, books, and documents. Knowledgebase constitutes the problem solving rules, facts, or intuition that a human expert might use in solving problems in a given problem domain.

Reasoning (Inference) Engine is a program component that activates the knowledge in the knowledge base. The function of the reasoning is to explore information and relationships from the knowledge base to provide answers, predictions and suggestions. As a result, it consists of algorithms for manipulating the knowledge represented in the knowledge base to solve a problem presented to the system. Learning module is an extension component which makes use of existing knowledge and previous experience or historical data to increase performance efficiency. User Interface is a graphical user interface which enables users to interact with the system to get or perform tasks.

2.5.2 Phase of building expert system

Building an expert system is known as knowledge engineering and its practitioners are called knowledge engineers. The knowledge engineer must make sure that the system has all the knowledge needed to solve a problem. The knowledge engineer must also decide knowledge representation form and ensure that the computer can use the knowledge efficiently by selecting suitable reasoning methods [27]. Knowledge-engineering process includes five major phases, knowledge acquisition, knowledge representation, knowledge validation, inferencing and explanation and justification [43]. In this section we briefly describe each activities in detail.

Knowledge Acquisition

An expert treating a patient uses both knowledge and data. The data is the patient's record, including patient history, environment and management condition, drugs given, response to drugs, and so on, whereas knowledge is what the expert has learned in medical school and in the years of internship, specialization, and practice. Knowledge consists of facts, prejudices, beliefs, and experience. The most fundamental distinction of knowledge type exists between factual and heuristic knowledge. Factual Knowledge is knowledge that is measurable, observable, easily articulated and understood, and heuristic knowledge is knowledge attained by intelligent guesswork (i.e. emotional intelligence) rather than by metric driven data. It is less rigorous, more experiential, more judgmental knowledge of performance. In contrast to factual knowledge, heuristic knowledge is rarely discussed, and is largely individualistic. Many knowledge bases in expert system include factual and heuristic knowledge.

The process of collecting knowledge in a problem domain is called knowledge acquisition [8]. Knowledge acquisition is the most important as well as the most difficult task in the development of expert system. The main reason for its difficulty is the communication gap between the knowledge engineer and the domain expert.

The methods of knowledge acquisition can be divided into manual, semi-automated and automated. The manual approach is a human-intensive method for knowledge acquisition, such as interviews, observations, document analysis, and brainstorming etc. Semi-automatic methods uses computer based tools to support knowledge engineers in order to facilitate the acquisition process. Automatic methods involves using computer software to automatically discover knowledge from a set of data. It includes Inductive learning, neural computing and genetic algorithms [8]. According to the problem at hand different methods of acquisition can be used. Some of the most common methods are interview, observation and document analysis.

Interview: is the most commonly used knowledge acquisition technique. It can be classified as structured and unstructured interviews. The structured interview is a formal version in which the knowledge engineer has planned the whole session whereas unstructured interview involves informal kinds of asking domain experts to elicit their knowledge. The success of an interview session is dependent on the questions asked, the ability of experts to articulate their knowledge and communication skills of a knowledge engineer to ask domain experts.

Observation: it involves observation of experts during performing their tasks. It may take different forms like on site observation, discourse analysis and active participation. Document analysis involves gathering knowledge from existing documentations. It may involve the interaction with a human expert to confirm or add to the existing information or knowledge.

Knowledge representation

Knowledge representation is the systematic means of encoding knowledge gathered from the expert. It is important and crucially affects the performance and efficiency of the system. The forms of knowledge representation often used in ES are rules, semantic nets, frames, and cases [8].

Rules: the knowledge is represented in the form of condition/action pairs: IF this condition (or premise or antecedent) occurs, THEN some action (or result or conclusion or consequence) will (or should) occur. Rules are natural way of representation and a simulation of the cognitive behavior of human experts. Rules are easy for a human expert to read, understand and maintain.

Semantic network: the knowledge is represented through the connection between objects or class of objects. It is composed of nodes and links that show hierarchical relationships between objects. It is a directed graph consisting of vertices which represent concepts and edges which represent semantic relations between the concepts. The advantage of semantic network is easy to follow hierarchy, easy to trace associations and flexible. However, semantic networks have problems in knowledge representation which include no internal structure of nodes, relationships between multiple nodes is not possible and no easy way to represent heuristic information, meaning attached to nodes might be ambiguous and exception handling is difficult.

Frames: are a natural extension of semantic networks. Knowledge represented in a frame is divided into slots. They consist of sets of slots filled by values, procedures for calculating values, or pointers to other frames. Frames can be formalized as a set of relations between entities having certain properties. It is a collection of questions to be asked about a hypothetical situation and specifies issues to be raised and methods to be used in dealing with them. The advantage of using frame based approach is its expressive power, easy to set up slots for new properties and relations, easy to include default information and easy to detect missing values. Its drawback includes difficult to program, difficult for inference, lack of inexpensive software.

Case: based representations store a large set of previous cases with their solutions in the case library and use them whenever a similar new case occur. It avoids the personal influence of individual experts. It pass the expert and look directly at the information that allowed them to learn and acquire their problem solving capability.

Knowledge Evaluation (Validation)

Evaluation is the degree to which inferences and decisions are justified by evidence. For expert systems, this requires analyzing the decision making capabilities of a system. So, expert system evaluation process involves assessment of many aspects of expert system components like evaluating the quality of knowledge represented, quality of knowledge base and the decision process. This is done usually by test cases involving domain experts to confirm the accuracy or to check if it conforms to certain criteria [2]. The system is evaluated by comparing system decisions to the correct answers. A valid system would produce the same answers as the experts.

Inferencing (Reasoning)

The inference engine deduces facts or draws conclusions from the knowledge base based on the user input and the facts from the knowledge base. There are different reasoning methods in expert systems like rule based, case based, model based and hybrid reasoning [1].

RBR (rule based reasoning) is reasoning technique which solve input problem by, finding applicable rules by matching against the rules of the knowledge base then, intermediate results are generated by choosing one of the inference mechanism forward or backward chaining, and the process is repeated till the desired solution state is reached. RBR allow compact representation of general knowledge. It have natural way of knowledge representation and adding or removing rule is possible without affected the system. RBR encounter inference efficiency problems for very large rule bases. The drawback of RBR is conclusion can't be drawn from the rules if missing or unexpected data is found. Also, when rules increase, maintaining the knowledge base is difficult.

CBR (case based reasoning) is one of the reasoning techniques which solve input problem by, identifying and adapting similar problems stored in a library of past experiences/problems [1]. Generally case based reasoning process follow four steps, retrieve the most similar case, reuse the case, revise the proposed solution and, retain the new solution as a part of a new case. CBR have the ability to encode historical knowledge directly and allow simple additive model for knowledge acquisition. If an appropriate case can be found, new problems can often be solved in much less time. However, cases do not often include deeper knowledge of the domain. This handicaps explanation facilities, and in many situations it allows the possibility that cases may be misapplied, leading to poor quality or wrong advice. They also encounter inference efficiency problem for large case base and it is difficult to explain the reasoning steps.

Model based reasoning (MBR) is one of the reasoning technique which solve input problem by following series of steps that logically follow the model. It use functional and structural knowledge of the domain in problem solving. Model based reasoners are often built using scientific and theoretical knowledge. It allow the transfer of knowledge between tasks. MBR lack descriptive knowledge of the domain and the complexity increase because it operates at a level of detail and the output cannot be predicted on unusual situations. By combining the above reasoning methods to get a cooperative effect where the strengths of one system can compensate for the weakness of another has been also used for many expert system.

2.6 Digital Image Processing

2.6.1 Image

Image is defined as a two-dimensional function, f(x, y), where x and y are spatial coordinates, and the amplitude at any pair of coordinates (x, y) is called the intensity or gray level of the image at that point. When x, y, and the amplitude values of f are all finite, discrete quantities, the image is called a digital image [5]. Digital image composed of a finite number of elements, each of which has a particular location and value. These elements are referred to as picture elements (pixels).

There are two important ways to represent f(x, y). The first way is a plot of the function, with two axes determining spatial location and the third axis being the values of f which are also known as intensities as a function of the two spatial variables x and y, as shown below in Figure 2.5.

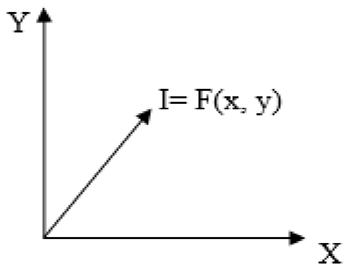

Figure 2-5 representation of image using coordinates

This representation is useful when working with sets whose elements are expressed as triplets of the form (x, y, z), where x and y are spatial coordinates and z is the value of f at coordinates (x, y). However, complex images generally are too detailed and difficult to interpret from such plots and need other representation. The second representation is displaying the numerical values of f(x, y) as an array called matrix with the representation of an M x N as shown in the Figure 2.6 [6].

$$f(x, y) = \begin{pmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{pmatrix}$$

Figure 2-6 representation of image using matrix

In this representation, the origin of the image is at the top left. Moreover, the positive x-axis extends downward and the positive y-axis extends to the right. This is a conventional representation based on the fact that many image displays, like TV monitors, sweep an image starting at the top left and moving to the right one row at a time [6].

Digital image processing (DIP) is manipulation of digital images using computers. DIP focuses on developing a computer system that is able to perform processing on an image [5]. DIP have very wide applications in medical diagnosis, remote sensing, transmission and encoding, Machine/Robot vision, pattern recognition etc. [6].

DIP has been extensively used in various (human, animal, plant) disease diagnosis approaches and assisting experts to select the right treatment. It can either be used to recognize the symptoms of a disease on the skin or even in the molecular analysis using microscope images that display the anatomy of the tissues [40].

Digital Image processing can be seen in three levels. These levels are termed as low, mid, and high level processes. Low level processes involve primitive operations such as image preprocessing to reduce noise, contrast enhancement, and image sharpening. A low level process is characterized by the fact that both its inputs and outputs are images. Mid-level processing involves tasks such as segmentation. Unlike low level processing, in mid-level its inputs are images, but its outputs are attributes extracted from those images. High level processing involves making sense of a group of recognized objects like classification, tracking etc. [6].

2.6.2 Steps of Digital Image Processing

Basically different image processing applications may follow different steps. However, the fundamental steps that every image processing applications pass through are shown in Figure 2.7.

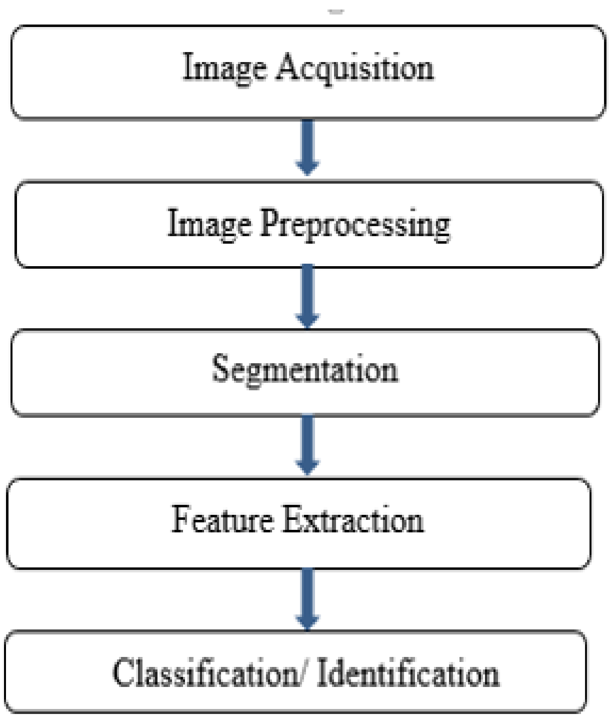

Figure 2-7 fundamental steps of DIP

Image acquisition: is used to acquire an image from a source, usually a hardware-based source. The image that is acquired is completely unprocessed and is the result of whatever hardware was used to generate it [5]. The hardware device can be anything from mobile camera to a massive optical telescope.

Image pre-processing: is low level processing to enhance the images prior to computational processing. The image pre- processing is a technique to bring out detail that is obscured, or simply to highlight certain features of interest in an image. The removal of noise, edge enhancement, restoration, resampling and the soft focus (blurring) effect are way to enhance an image presentation.

Image segmentation: is the process used to locate objects and boundaries (e.g., lines or curves) in images. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. The segmented objects are often termed the foreground and the rest of the image is the background. It is applied to remove the unwanted part from the image and find the region of interest. For any given image, we cannot generally speak of a single, correct segmentation. Rather, the correct segmentation of the image depends strongly on the types of object or region we are interested in identifying. The central question in image segmentation is what relationship a given pixel have with respect to its neighbors and other pixels in the image in order that it be assigned to one region or another.

There are many segmentation techniques, but they can be categorized into detection of discontinuities and detection of similarities. The segmentation technique based on detection of discontinuities is the process of partitioning an image based on abrupt changes in intensity. Examples of such algorithms include all edge detection algorithms. On the contrary, detection of similarities is based on continuities. These techniques divide the entire image into sub regions depending on some similarity rules. Examples of such algorithms include thresholding, region growing etc. Whether the segmentation technique used is discontinuity or similarity, the end result of any segmentation process is a binary image [6]. Table 2.3 show the segmentation techniques used commonly in image processing application.

Table 2:3 Comparison of different segmentation technique

Segmentation	Description	Advantage	Disadvantage
Thresholding	based on the histogram peaks of the image to find particular threshold values	✓ No need of previous information ✓ simplest method	✓ highly dependent on peaks ✓ spatial details are not considered
Edge based method	based on discontinuity detection	✓ good for images having better contrast between objects	✓ not suitable for wrong detected or too many edges
Region Based Method	based on partitioning image into homogeneous regions	✓ more immune to noise ✓ useful when it is easy to define similarity criteria	✓ expensive method in terms of time and memory
Clustering Method	based on division into homogeneous clusters	✓ fuzzy uses partial membership ✓ useful for real problems	✓ determining membership function is not easy
ANN Based Method	based on the simulation of learning process for decision making	✓ no need to write complex programs	✓ more wastage of time in training

Feature Extraction: Features are the information extracted from images in terms of numerical values that are difficult to understand and correlate by human. If we consider the image as data the information extracted from the data is known as features. Features extracted from an image are of much more lower dimension than the original image. The reduction in dimensionality reduces the overheads of processing bunch of images. A good feature set contains discriminating information, which can distinguish one object from other objects. The selected set of features should be a small set whose values efficiently discriminate among patterns of different classes, but are similar for patterns within the same class. There are several image features which represent an image for classification/identification systems. Most popular among them are color, texture and shape of an image.

Color Features: are defined to a particular color space. Once the color space is specified, color feature can be extracted from the image. The most and common used color features include color histogram, color moments (CM), color coherence vector (CCV) and color correlogram, etc. The efficiency of the color feature resides in the fact that it is independent and insensitive to size, rotation and the zoom of the image [65].

Texture Features: texture is a repeating pattern of local variations in image intensity. It refers to surface characteristics and appearance of an object by the size, shape, density, arrangement, proportion of its elementary parts. It is generally believed that human visual systems use texture for recognition and interpretation. Color is usually a pixel property while texture can only be measured from a group of pixels [66]. Based on the domain from which the texture feature is extracted, texture feature classified into two categories spatial texture feature extraction and spectral texture feature extraction [65]. In spatial, texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain, whereas in the spectral it transforms an image into frequency domain and then calculates feature from the transformed image. Among the current approaches used in image processing to describe texture, the so called statistical approach is the widely used because it produces good results with low computational costs. This method considers the distribution of gray levels and their interrelationship. The pixel values are used to construct numerical structures which are associated to the texture pattern of the image.

Shape Features: the shape of an object refers to its physical structure and profile. Shape features are mostly used for finding and matching shapes, recognizing objects or making measurement of shapes. Moment, perimeter, area and orientation are some of the characteristics used for shape feature extraction technique. The shape of an object is determined by its external boundary abstracting from other properties such as color, content and material composition, as well as from the object's other spatial properties. Shape features are commonly used in the object recognition and shape description. The shape features extraction techniques are classified as region based and contour based. The contour methods calculate the feature from the boundary and ignore its interior, while the region methods calculate the feature from the entire region.

2.6.3 Image Classification

Classification is the task of approximating a mapping function (f) from input variables (X) to discrete output variables (Y) [57]. In which the result is predicting the class of given data points. Classes are sometimes called as targets/ labels or categories. Classification is a two-step process, consisting training and classification step. In the training steps, a classification algorithm builds the classifier by analyzing and learning from a training dataset and their associated class labels. In the classification step the model is used to predict class labels for given data. The tasks depend on labeled datasets to learn the correlation between labels and data. It is a kind of supervised learning.

There are many classification algorithms designed by considering different assumptions [12]. Naive Bayes Classifier is a classification technique based on Bayes Theorem. A Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. Naive Bayes model is easy to build and particularly useful for very large data sets.

Nearest Neighbor is one of the classification algorithm. It takes a bunch of labelled points and uses them to learn how to label other points. To label a new point, it looks at the labelled points closest to that new point, and has those neighbors vote, so whichever label the most of the neighbors have is the label for the new point. Support Vector Machines is one of the classification algorithm. It represent the training data as points in space separated into categories by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

Artificial Neural Networks (ANN) is an information processing paradigm that is inspired by the way biological nervous systems process information. We are constantly analyzing the world around us. Without conscious effort, we make predictions about everything we see, and act upon them. When we see something, we label every object based on what we have learned in the past. Neural networks work with the same concept by adjusting the connection exist between neurons. It is composed of a large number of highly interconnected processing elements called neurons, which convert an input vector into some output.

Neural networks have the potential for solving problems in which some inputs and corresponding output values are known, but the relationship between the inputs and outputs either not well understood or difficult to translate into a mathematical function. These conditions are commonly found in tasks involving classification of agricultural and animal products, plant and animal disease identification etc. ANN is better from the above method because, it is nonlinear model that is easy to use and understand compared to statistical methods. Its non-parametric behavior allow better performance in complex problems because they don't need background of statistic.

ANN can be classified into shallow and deep learning. A shallow learning have less number of neurons compared to Deep learning (DL) [12]. Image classification has been done on both learning methods. But, there is basic difference between them. Shallow learning use one neuron for each input which increase number of parameters. For example let us consider we have an image with 224 x 224 with 3 color channels. The training contain 150,528 weight for a single image. It will cause difficulties while training and possible overfitting in the model. The problem become difficult when image size increase because the amount of weights becomes larger.

In shallow learning we lose spatial information when the image is flattened into the network. For example, if a picture of LSD affected cattle appears in the top left of the image in one picture and the bottom right of another picture, the shallow learning will try to correct itself and assume that LSD affected cattle will always appear in the bottom right of the picture. Hence shallow learning react differently to an input images and its shifted version and they are not translation invariant. But DL is translation invariant. Another difference is deep learning scale with data, whereas shallow learning converges. Shallow learning reach a stable state at a certain level of performance when you add more examples and training data to the network. But deep learning often continue to improve as the size of the data increase [14].

Another main difference is shallow learning don't extract feature by themselves. They need feature extractor. Feature extractor are problem dependent and must be rewritten for each new dataset. The efficiency of the extractor depends on the assumption of a person who design the extractor, because we simply cannot conceive every possible feature that will be useful [57]. Hence, using shallow learning for image classification task is not advisable. We choose deep learning for our research work.

Deep Learning is concerned with algorithms inspired by the structure and function of the brain called artificial neural networks. The performance of this type of model improves by training with more examples and by increasing its depth or representational capacity. In addition to scalability, another often cited benefit of deep learning models is their ability to perform automatic feature extraction from raw data [12].

There are many types of deep learning used for variety of tasks in Artificial intelligence [63] like deep neural network, deep belief network, recurrent neural networks and convolutional neural network. Deep neural network is neural network having more than two layers. It can model complex non-linear relationships. It is used for classification as well for regression. Deep neural networks use sophisticated mathematical modeling to process data in complex ways.

Deep belief network is unsupervised probabilistic deep learning algorithm. It can be defined as a stack of restricted Boltzmann machines, in which each RBM layer communicates with both the previous and subsequent layers. The nodes are connected to each other across layers, but no two nodes of the same layer are linked. That is, there is no intra-layer communication. The idea behind a DBN is that it is possible to stack more RBMs on top of each other by building a deep network, which presents the ability to extract a hierarchal representation of the input at multiple level of abstraction.

Deep Boltzmann network (DBM) is unsupervised, probabilistic, generative model with entirely undirected connections between different layers. It contains visible units and multiple layers of hidden units. Like RBM, no intra layer connection exists in DBM. Connections exists only between units of the neighboring layers network of symmetrically connected stochastic binary units DBM can be organized as bipartite graph with odd layers on one side and even layers on one side units within the layers are independent of each other but are dependent on neighboring layers. Deep Boltzmann Machine (DBM) have entirely undirected connections

Recurrent neural networks (RNNs) are a type of artificial neural network that are able to recognize and predict sequences of data such as text, genomes, handwriting, spoken word, or numerical time series data. They have loops that allow a consistent flow of information and can work on sequences of arbitrary lengths. The RNNs attempts to address the necessity of understanding data in sequences. Recurrent nets differ from feedforward nets because they include a feedback loop, whereby output from step n-1 is feedback to the net to affect the outcome of step n, and so forth for each subsequent step.

Convolutional neural network (CNN) is one of the main learning mechanisms to do images recognition, images classifications, objects detections etc. It is based on learning levels of representations. The higher lever concepts are defined from lower-lever ones, and the same lower lever concepts can help to define many higher lever concept. It learn multiple levels of representation and abstraction which helps to understand dataset such as images, audio and text. It is advantageous of simple structure, less training parameters because of shared weights and adaptability [47].

Table 2:4 difference between deep learning

No	Network type	Type of learning	Type of feature extracted
1	DNN	Supervised	Spatial Feature
2	DBN	Unsupervised	Latent Feature
3	DBM	Unsupervised	Latent Feature
4	RNN	Supervised	Temporal Feature
5	CNN	Supervised	Spatial Feature

To choose which network fit for our problem, we must consider the type of problems and the data we use. Each deep network have its own characteristics as shown in Table 2.4. Our task is classifying the input symptoms, so it is classification problem and the data we have is labeled and its supervised learning. The data types is image which are represented in pixel. One pixel is dependent on the value from its 8 neighbors, the feature is spatial feature. We choose CNN for image classification.

Convolutional Neural Network (CNN)

D.H Hubel et.al [53] in their study on the brain of mammals suggested a new model for how mammals perceive the world visually. They showed that cat and monkey visual cortexes include neurons that exclusively respond to neurons in their direct environment. They described two basic types of visual neuron cells in the brain that act in a different way simple cells (S cells) and complex cells (C cells). The simple cells activate when we identify basic shapes as lines in a fixed area and a specific angle. The complex cells have larger receptive fields and their output is not sensitive to the specific position in the field. They show cat and monkey visual cortexes contain neurons that individually respond to small regions of the visual field.

Fukushima [64] proposed a hierarchical neural network model called it Neocognitron. This model was inspired by the concepts of the simple and complex cells. The neocognitron was able to recognize patterns by learning about the shapes of objects. The neocognitron introduced the two basic types of layers in CNNs convolutional and down sampling layers.

Yann LeCun et.al [54] introduce Convolutional Neural Network. Their first Convolutional Neural Network called LeNet and was able to classify digits from hand written numbers. After this CNN was widely used in image classification tasks [57, 58, 59]. CNN automatically learn a hierarchy of features used for classification purposes. This is accomplished by successively convolving the input image with learned filters to build up a hierarchy of feature maps. The hierarchical approach allows to learn more complex, as well as translation and distortion invariant, features in higher layers [63]. For example given LSD infected cattle image, the convolution layer detects features. Then the fully connected layers act as a classifier on top of these features, and assign a probability for the input image being LSD infected. The convolution layers learn such complex features by building on top of each other. The first layers detect edges, the next layers combine them to detect shapes, to following layers merge this information to infer that this is a LSD. The CNN doesn't know what LSD is. By seeing a lot of them in images, it learns to detect the feature that uniquely describe LSD. Also it is computationally efficient. It uses special convolution and pooling operations and performs parameter sharing. This enables CNN models to run on any device, making them universally attractive.

Convolutional Neural Networks have a different architecture than other Neural Networks. In other Neural Networks each layer is fully connected to all neurons in the layer before. In convolutional Neural Networks the neurons in one layer do not connect to all the neurons in the next layer but only to a small region of it.

The basic layers in CNNs are convolutional layers, pooling layers, and fully connected layers. Convolutional layer (Conv layer) is a layer that gives the network its name. It is the first layer to extract features from an input image. It performs an operation called a convolution. In the context of CNN, a convolution is a linear operation that involves element-wise multiplication between the input image and filter [14]. The convolution operation is performed by sliding the filter over the input image. At every location, a matrix multiplication is performed and sums the result to produce feature map. The convolution operation produce the feature map. Figure 2.8 demonstrate how the operation is performed.

The main process in conv layer is convolution operation. A filter is small matrix which is used to detect pattern in the input image. The values of the matrix are initialized with random numbers using different techniques. This filter detects edges, textures specific shapes etc. The filter is placed on top of image and multiplied with value on the image. After that all results are summed and the result is written to output matrix called feature map.

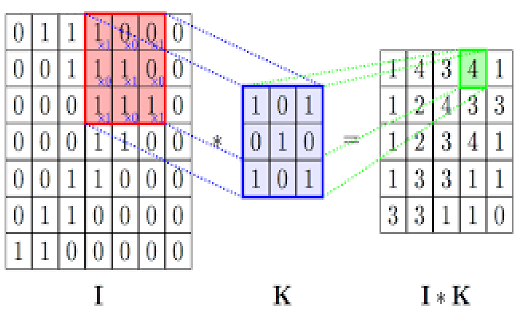

Figure 2-8 result of convolution operation

Pooling layers simplify the information obtained from the convolutional layer output. Pooling layer takes feature map output from the convolutional layer and prepares a condensed feature map. It is used to progressively reduce the size of the input representation. Hence, reduce the number of required parameters, the amount of computation required and control overfitting. The most common method of pooling is max pooling. In max pooling, the input image is represented by each area are the maximum value in each area. This makes a smaller size with fewer parameters. In Figure 2.9 the feature map is reduced into 2x2 representation by taking maximum values and moving 2 strides.

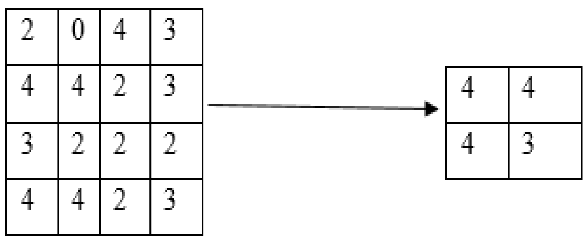

Figure 2-9 result of pooling operation

Fully connected layer is the only layer in the model where every neuron in the previous layer connects to every neuron in the next layer. It is uses the features from the output of the previous layer to classify the input image based on the training data. There can be multiple of the above layers in the model depending on the complexity of the function which is going to be mapped by the model. Having more layers will enable the model to learn complex relationships. However, it takes a lot of time to train and adjust weights.

Activation function: are mathematical equations that determine the output of a neural network. The function is attached to each neuron in the network, and determines whether it should be activated or not, based on whether each neuron’s input is relevant for the model’s prediction. It serves to introduce non linearity in the modeling capabilities of the network. A neural network without Activation function would simply be a linear regression Model, which has limited power and does not perform well. There are three types of Activation Functions binary, linear and nonlinear function.

Binary Function is a threshold based activation function. If the input value is above or below a certain threshold, the neuron is activated and sends exactly the same signal to the next layer. It does not allow multi value outputs, for example, it cannot support classifying the inputs into one of several categories.

Linear Activation Function is activation function takes the form A = cx. It takes the inputs, multiplied by the weights for each neuron, and creates an output signal proportional to the input. Linear function is better than binary function because it allows multiple outputs, not just yes and no. All layers of the neural network collapse into one with linear activation functions, no matter how many layers in the neural network. The last layer will be a linear function of the first layer because a linear combination of linear functions is still a linear function. So a linear activation function turns the neural network into just one layer. Also using backpropagation to train the model because the derivative of the function is a constant, and has no relation to the input, X. So it's not possible to go back and understand which weights in the input neurons can provide a better prediction.

Non-Linear Activation Functions allow the model to create complex mappings between the network inputs and outputs. It is essential for learning and modeling complex data, such as images, video, audio, and data sets which are non-linear or have high dimensionality. They allow backpropagation because they have a derivative function which is related to the inputs. They allow stacking of multiple layers of neurons to create a deep neural network. Multiple hidden layers of neurons are needed to learn complex data sets with high levels of accuracy.

We need a neural network model to learn and represent almost anything and any arbitrary complex function which maps inputs to outputs that is why we need non linearity. Hence using nonlinear Activation we are able to generate non-linear mappings from inputs to outputs. Some of the most common nonlinear Activation function are:

Sigmoid is activation function of form f(x) = 1 / 1 + exp (-x). It's recommended to be used only on the output layer so that we can easily interpret the output as probabilities since it has restricted output between 0 and 1. It is computationally expensive. Tanh is activation function of form f(x) = 1 — exp (-2x) / 1 + exp (-2x). It is also computationally expensive. RELU is activation function of form f(x) = max (0, x). Compared to other activation function, it is computationally expensive. There is no definitive guide for which activation function works best on specific problems. It's a trial and error process by trying different set of functions and see which one works best on the problem at hand.

Loss function: is one of the parameters required to quantify how close a particular neural network is to the ideal weight during the training process. It's a method of evaluating how well specific algorithm models the given data. If predictions deviates too much from actual results, loss function would cough up a very large number. Loss functions can be classified into two major categories depending upon the type of learning task we are dealing with Regression and Classification losses. In classification, we are trying to predict output from set of finite categorical values. Regression deals with predicting a continuous value. The choice of the best function of loss resides in understanding what type of error is acceptable for the problem in particular.

Training CNN: the problem of training is equivalent to the problem of minimizing the loss function. The procedure used to carry out the learning process in a neural network is called the optimization algorithm. There are many different optimization algorithms. All have different characteristics and performance in terms of memory requirements, speed and precision. These algorithms can gradient-based or not. The most common used algorithm is gradient based.

The aim of training is to reduce the difference between the predicted output and the actual output. Our goal is to minimize the cost function by finding the optimized value for weights. To achieve this we run multiple iterations with different weights [47]. This is what Gradient descent is about.

Gradient descent is an iterative optimization algorithm to reduce the cost function. This will help models to make accurate predictions. Gradient indicates the direction of increase, as we want to find the minimum point in the valley we need to go in the opposite direction of the gradient. We update parameters in the negative gradient direction to minimize the loss. During the training process, we change the weights of our model to minimize the loss function, and make our predictions as correct as possible.

Gradient descent uses the first derivative (gradient) of the loss function when updating the parameters. The process consists in chaining the derivatives of the loss of each hidden layer from the derivatives of the loss of its upper layer. In each of the iterations, once all the neurons have the value of the gradient of the loss function that corresponds to them, the values of the parameters are updated in the opposite direction to that indicated by the gradient. The gradient, in fact, always points in the direction in which the value of the loss function increases. Therefore, if the negative of the gradient is used, we can get the direction in which we tend to reduce the loss function.

The training consists the following steps:

- Start with values (often random) for the network parameters (wij weights and bj biases).
- Take a set of examples of input data and pass them through the network to obtain their prediction.
- Compare these predictions obtained with the values of expected labels and calculate the loss with them.
- Perform the backpropagation in order to propagate this loss to each and every one of the parameters that make up the model of the neural network.
- Use this propagated information to update the parameters of the neural network with the gradient descent in a way that the total loss is reduced and a better model is obtained.
- Continue iterating in the previous steps until we consider that we have a good model.

CHAPTER 3 : RELATED WORK

3.1 Introduction

In this chapter, we discuss research works which are related to plant and animal disease diagnosis. We reviewed diagnosis works that are related to expert system and image analysis and diagnosis which used both methods.

3.2 Expert system with text input

Engidu et.al [1] developed web based expert system with hybrid reasoning consisting of case and rule based reasoning for the diagnosis of cattle diseases in Ethiopia. The system accept text based symptoms from a user then search a solution from a similar case in the case base. When case based reasoning fails to identify the disease, it will be transferred to rule based reasoning. The system also include learning module which gives learning capability from the past decision experience. The system was evaluated through system performance testing and user acceptance testing with corresponding results of 91.67% and 82.34% respectively and learning module performance was 100%. They conclude the proposed approach has a promising result to be used for cattle diagnosis purpose.

Derejaw Lake [2] developed web based expert system to diagnosis infectious and non-infectious cattle diseases in Ethiopia using rule based reasoning. In addition to diagnosis, the system allows the users to view diseases detail, post and view current dangerous cattle diseases. Also it have Amharic and English knowledge base. The system was evaluated by domain experts (veterinarians) and animal health assistants, the average performance result was 87.2%.

Berhanu Aebissa [3] developed rule based expert system for coffee disease diagnosis. The system helps non experts to identify major disease and pests that frequently occur in coffee and provides the appropriate description, treatments and prevention. The system agreed with human expert opinions in 83.6% of the decisions.

Solomon Gebremariam [17] developed a prototype of self-learning knowledge expert system for diagnosis and treatment of diabetes. Decision tree is used as a knowledge modeling technique for the developed rule based expert system. Backward chaining inference mechanism is used in that possible solutions are given initially and supportive facts verifying the solution are selected from the knowledge base. The overall performance of the prototype was 84.2%.

Seblewongel Esseynew [16] developed a prototype expert system for anxiety mental disorder diagnosis. Decision tree was used as a knowledge modeling technique while rule based reasoning was chosen for inference mechanism. The prototype achieved 85% accuracy based on domain expert's evaluation.

Ahmad Zamsuri [4] developed android application expert system to simplify the disease detection and showing the brief information about the cattle's using rule based forward reasoning engine. The system start by asking questions about the symptoms occurred and it deal with the uncertainty values in cow disease using Certainty Factor (CF). They conclude the prototype designed was easily operated by the user or it is user friendly.

Desalegn Aweke Wako [35] developed a rule based expert system for wheat disease diagnosis. The system works by accepting symptoms of wheat plants through text and find possible cause of diseases. The system was tested by system testing and user acceptance testing. The system overall performance is 87.78%. They recommended more research work must be done to incorporate pictures which shows symptoms in order to identify the damage of the wheat crop.

Whisnu Ulinnuha Setiabudi et.al [36] developed application which can diagnosis dental diseases. The application is rule based and it can diagnose the patient based on griping of the patient about the symptoms. The accuracy of the system test is performed by 20 patients, there were 19 cases of corresponding and 1 cases that do not fit. So, from system testing performed by 20 patients resulted in a 95% accuracy rate.

Hongyan Gao et.al [30] developed equine disease diagnosis expert system. The system provide symptom based diagnosis and differential diagnosis based on disease signs in equine patients. In addition, the system provides several modules, including disease learning and expert online, which serve to provide educational opportunities for the end users. To abstract the knowledge of equine disease object oriented and ontology technology are used. The reasoning mechanism used is uncertainty inference model based on improved reasoning of evidence credibility. Based on functional evaluation of the system, the diagnostic accuracy is 88%.

3.3 Diagnosis system with image input

Shweta [28] developed a system which detect plant diseases using image processing technique. The system takes images captured by smart phones as input and image analysis are done. Image scaling is used to convert the original image into thumbnails to reduce the processing time. K-Means Clustering is used for segmentation and color, texture edges, morphology features are extracted from the image then deep learning algorithm are used for detection and classification of the diseases.

Sujatha R et.al [29] developed a diagnosis system for leaf disease using image processing. The system start by loading the leaf image into the system and convert the image from RGB into HSI. Then contrast enhancement is done to enhance the image. To identify the affected area of leaf K-Means clustering is used and classification is done using SVM.

Pravin S. Ambad [31] developed a system used for early prevention and detection of skin diseases using image analysis. The system accept the RGB image and convert into gray images. Then the pre-processing is done to enhance and remove noise using median filter. They use feature extracted from the image like entropy, standard deviation, texture index, correlation factor. The classification model used is Adaboost classifier.

Nikos Petrellis [33] developed a windows phone application capable of recognizing plant diseases using image processing. The user captures photos of plant parts with lesions then runs the plant disease recognition application. The image is converted into a grey image and threshold is used to separate the affected region or part of the plant. It extracts lesion features like number of spots, their grey level and area using histogram indicating the number of pixels that have a specific red, green or blue color level. The system have GPS localization used to determine the specific rural region where the plant exists which capture historical weather data. The application is tested for vineyard diseases a success rate higher than 90% has been achieved in the disease recognition process. But the classification method used is not mentioned.

Daniel Hailemichael Lemessa [6] developed a system capable of assessing the quality of maize sample constituents using digital image processing techniques. The system pass through preprocessing, segmentation, feature extraction and classification components. The preprocessing component removes false regions based on size. The segmentation component separate individual maize sample constituents from the background and from each other using a hybrid of color structure tensor and thresholding. Color, shape, and size are feature used in classification. Finally, the classification component uses ANN to classify constituents of maize sample. The classification use a feedforward artificial neural network classifier with backpropagation learning algorithm. The overall success rate for the classification of the system is 97.8%.

Getahun Tigistu [5] developed a system that automatically identifies flower diseases using image processing. In order to remove noises they use image enhancement and segmentation. Then, extracted the texture features of an image using Gabor feature extraction and from this texture features, we have calculated seven different measures of central tendency and dispersion of the extracted Gabor texture features. Based on the extracted features the training and testing data that are used to identify are extracted. Finally, appropriate machine learning pattern identifier is selected to classify an image in to its class of disease. The developed system can successfully identify the examined flower with an accuracy of 83.3%.

3.4 Expert system with text and image input

Amarathunga et al. [38] designed rule based expert system for diagnosis of human skin diseases. In their system the user will upload an image of infected skin and answer questions based on the skin condition to detect diseases and get a treatment. The system have two components image processing and data mining component.

- ✓ Image processing component: The image must be captured at 5cm distance between camera lens and affected skin without any light effects. To eliminate noises from the particular skin image and to get smooth image, median filtering and Gaussian is used. Threshold segmentation is used to separate healthy skin from disease. Color and shape features are used to create the classification model.
- ✓ Data mining component: information is acquired by a questionnaire which is displayed as a form on the user interface to get details about the diseases. Then based on those images extracted features (color and shape) and answer from the questionnaire the classification model is created. Multi-Layer Perceptron MLP classifiers is used.

Carl Louie Aruta et.al [60] developed a mobile application for the diagnosis of skin diseases using case-based reasoning with image processing to detect diseases by applying similar past problems.

It uses image processing and case based reasoning methods. The system receives image input with specified resolution and input from user. Then it search any previous cases 90% similar to the entered data. Texture feature is extracted using ABCD rule. Multilayer Perceptron Classifier (MLP) is used to build classification model. If the system found a similar case, it will generate a diagnostic solution. But if the system didn't found the case, it will retrieve another case from the case library. The system can successfully detect 6 different skin diseases with an accuracy of 90%. However it didn’t mention how the cases will be retrieved and how the current case is compared to the previous case.

Mohammed El helly [15] proposed an approach for integrating image analysis techniques into diagnostic expert system. They present the approach through cucumber fungal diseases as a case study. The system have image analyzer component which takes image of defected plant as input and extract discriminative features. Then those features are used to classify into specific class using ANN.

P. Spyridonos et.al [61] incorporate image analysis techniques to confine subjective evaluation of histopathological variables. The resulted diagnostic model could be seen as a two level structured diagnostic system. In the lower level an automatic image analysis system is employed to extract cell nuclei quantitative descriptors. These descriptors feed some of the concepts of the FCM model working in the upper level.

Table 3:1 related work summary

Category	Title	Algorithm	Authors
Expert system with Text	Web based expert system for diagnosis of cattle disease	✓ Hybrid reasoning (rule and case based reasoning ✓ Success and failure driven learning	Engidu Gebre-Amanueal, Fekade Getahun and Anteneh Assalif
	Web based expert system for cattle disease diagnose	✓ Rule based reasoning	Derejaw Lake
	Web based cattle disease expert system diagnosis with forward chaining method	✓ Rule based reasoning ✓ Forward chaining inference engine	Ahmad Zamsuri, Wenni Syafitri, Muhamad Sadar

	Development of Knowledge Based System for Wheat Disease Diagnosis: A Rule Based Approach	✓ Rule based	Desalegn Aweke Wako
	Developing a Knowledge Based System for Coffee Disease Diagnosis and Treatment	✓ Rule based	
Expert system with image	Diagnosis of fish diseases Using Artificial Neural Networks	✓ Feed-forward back-propagation neural networks	J.N.S. Lopes, A.N.A. Gonçalves
	Image processing techniques for identification of fish disease	✓ PCA machine learning algorithms ✓ HOG (Histogram of Gradient), FAST (Features from Accelerated Segment Test)	Shaveta Malik, Tapas Kumar, A.K.Sahoo
	Dermatological Disease Detection using Image Processing and Artificial Neural Network	✓ Median , smooth filter for Pre processing ✓ Sobel operator for segmentation ✓ Color and shape feature ✓ ANN	Rahat Yasir, Md. Ashiqur Rahman, and Nova Ahmed,
	Detection and Analysis of Plant Diseases Using Image Processing Technique	✓ K-Means Clustering for segmentation ✓ Color, texture and morphology feature ✓ Deep learning	Shweta R. Astonkar, V. K. Shandilya,
	Leaf disease detection using image processing	✓ Contrast enhancement ✓ K-Means clustering for segmentation ✓ SVM	Sujatha R, Y Sravan Kumar and Garine Uma Akhil
	A Image analysis System to Detect Skin Diseases	✓ Median filter to remove noise ✓ Adaboost classifier	Pravin S. Ambad, A. S. Shirsat
	A Smart Phone Image Processing Application for Plant Disease Diagnosis	✓ Threshold for segmentation ✓ GPS localization	Nikos Petrellis,
	Development of Automatic Maize Quality Assessment System Using Image Processing Techniques	✓ Threshold and color structure tensor for segmentation ✓ Color and shape feature ✓ ANN	Daniel Hailemichael Lemessa

	Automatic Flower Disease Identification using image processing	✓ Ostu threshold for segmentation ✓ Gabor feature extraction ✓ ANN	Getahun Tigistu,
Expert system with text and image	Mobile-based medical Assistance for diagnosing Different types of skin Diseases using case-based Reasoning with image Processing	✓ Case based reasoning ✓ Multilayer perceptron Classifier (MLP) ✓ ABCD for feature extraction	Carl Louie Aruta, Colinn Razen Calaguas, Jan Kryss Gameng
	Expert system for Diagnosis of skin diseases	✓ MLP ✓ Rule based reasoning	A.A.L.C. Amarathunga, E.P.W.C. Ellawala, G.N. Abeysekara, C. R. J. Amalraj

3.5 Summary

As shown in Table 3.1 many expert system are designed to diagnose diseases from observed symptoms. Most of them use symptom described using either text or image. There are also research done by combining image and text for plant and human disease diagnosis. But the image analysis part is done by shallow networks which need manual feature extraction. Using manual feature extraction is not good for classification tasks because, we specify restrictions on what features represent the input data. Generally existing work don't use automatic feature extraction in the image analysis and not tested for animal disease.

CHAPTER 4 : DESIGN OF IMAGE BASED CATTLE DISEASE DIAGNOSIS APPROACH

4.1 Introduction

Disease diagnosis is the major and first step in treatment and prevention of diseases. Using image processing for disease diagnosis has been a wide area of research topics. In this chapter, we present our image-based cattle disease diagnosis (IBCDD) approach. We will discuss the expert system and image analysis module. The components of each module are described along with relevant techniques, algorithms, and considerations in the proposed system.

4.2 IBCDD design consideration

We consider the following list while designing and developing the proposed IBCDD.

Types of disease: The approach considered disease which has significant symptoms identified by inspection (symptoms are seen by the naked eye) and can be represented through an image. Disease which doesn't have distinctive symptoms and can't be identified by inspection is not included.

Capacity of mobile phone: diagnosis is urgent or a condition that needs instant reply. The model must consider the capacity of the mobile phone to process the user query. So, the model must consider the majority population mobile phone capacity.

Area under consideration: when the image is captured only the area which is affected must be captured. Not the whole cattle body is needed or captured.

GPS Location on camera must be on: we need location information of the image. The location on the mobile phone must be on when image is captured.

4.3 The IBCCD System Architecture

The proposed system is the integration of expert system and image analysis using deep learning. The system collect information about the occurred disease through the user interface. It will reach to diagnosis conclusion based on facts and knowledge stored using the reasoning component. The system architecture for the proposed system is illustrated in Figure 4.1.

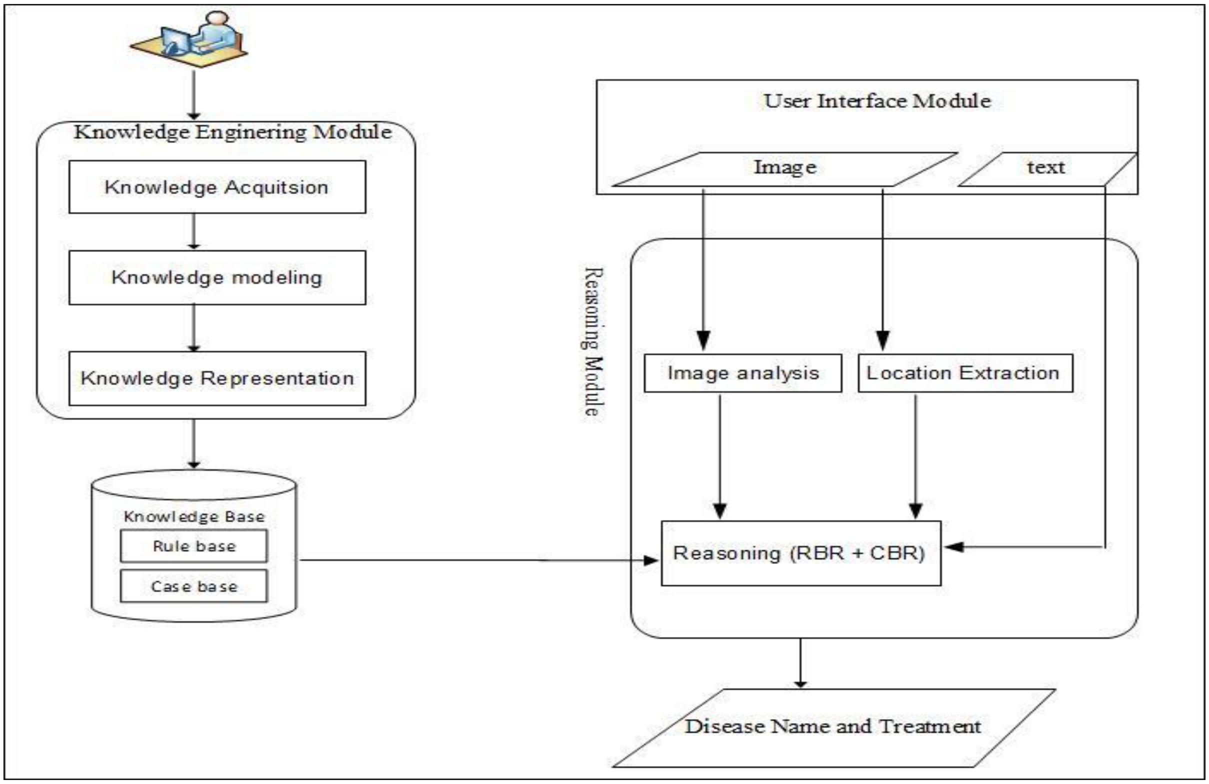

Figure 4-1 proposed system architecture

Occurred symptoms and signs are brought to the system using the UI (user interface). Symptoms identified by inspection are brought to the system using an image. Then, the category is identified by the Image analysis module. Symptoms identified by palpation are brought using text dialogue. Location, where the image is captured, is extracted from the image to guide the diagnosis procedure. Most of the cattle disease has the same symptoms and can be identified through detailed examination with more information. Hence, the diagnosis must incorporate the required information to reach diagnosis results. The reasoning component reaches a final diagnosis result by integrating classification and location extraction results and other text information obtained from the user.

4.3.1 User interface

The user interface (UI) is responsible for communication between the user and the system. It is used for acquiring necessary symptoms and signs occurred in the infected cattle. The UI component have text and image based user interface which allows the system to gather information through the two medium. The text interface allows to input symptoms which are detected by palpation. The text interface is in form of checkbox, where the user can select the appropriate symptoms. The image interface allows to input symptoms which are recognized by inspection method. The image can be captured through mobile camera and brought to the system.

4.3.2 Location Extraction

The Location extraction module is responsible for acquiring location information where the image is captured. Smartphones attach all kinds of Meta data to images. This Meta data is known as Exchangeable image file format (EXIF). It is a standard of adding metadata to image files. Depending on the camera, EXIF data will store the current state of the camera when the photo was taken including date and time, shutter speeds, focal lengths, flash, lens type, location data etc. To capture location information from image the camera must be GPS enabled. When image is captured the image will be geotagged. Geotagged means the longitude and latitude of the photo is stored in the image metadata. So, EXIF data about longitude and latitude is extracted from the image to get location information about where the image is taken to guide the diagnosis process.

Algorithm 4-1 location extraction algorithm

```
Input: Image
Output: Location //location of where the input image is captured
Begin:
  Location=Image.getEXIFdata(Longitude, Latitude)
Return Location
END
```

4.3.3 Knowledge Engineering

The knowledge engineering module is responsible to build the required knowledge base for the proposed approach. It contains KA (knowledge acquisition), KM (knowledge modeling) and KR (knowledge representation) components.

Knowledge acquisition (KA): is used to acquire knowledge and fact required for the diagnosis. The rules are extracted from the standard veterinary treatment guideline prepared by Drug Administration and Control Authority (DACA) [44], Blacks dictionary [42] and Merck veterinary manual [41]. The cases are collected from different case study papers, veterinary articles and journals.

Knowledge modeling (KM): is responsible for the acquired knowledge from KA process will be modeled to understand the whole diagnosis and treatment techniques. It demonstrate the diagnosis procedures in skin diseases. Knowledge acquired in the knowledge acquisition process about diagnosis methods, prevention and control strategies are modeled using hierarchical tree structure.

Disease diagnosis knowledge modeling

When skin diseases are reported to Experts they follow basic procedures to predict the possible diseases. They start by observing symptoms on the infected animals through Inspection which is one of the methods in physical examination. Inspection is a visual examination done some distance away from the animal or go round the animal. Other symptoms are identified by Palpation which requires touching the infected cattle by hands to detect the presence of pain in tissue by noting increased sensitivity. They continue by checking if there are endemic diseases that frequently affect cattle living in that region and considering season if there is a potential to be infected by a seasonal disease. Based on the available information they predict possible occurred diseases. Then they advise treatment and control. The proposed disease diagnosis model considered all these features as shown in Figure 4.2.

Symptoms is the first clue that reflects the presence of a certain disease in the cattle. In cattle disease, there are symptoms that can be identified by inspection means by seeing (observing by naked eyes) the infected body part. Symptoms like Lesions on the outer surface of the body like skin and coat, Nose, Mouth, Eyes, Legs, and Anus can be detected by inspection. There are symptoms that are identified by palpation. Using palpation, we can detect the presence of pain, the firmness of the lesion, consistency, and temperature of body parts, etc.

Seasonal and Endemic Disease: it is common to encounter seasonal incidence of diseases. Both dry and wet seasons trigger different diseases. The seasonal variance affects the nutrient available for food consumption, which may trigger seasonal skin diseases. Moreover, disease vectors and intermediate hosts have a seasonal effect. For example, LSD transmitter insects and ticks are found more in wet weather which increases the occurrence of LSD in this season. Therefore, knowing which disease is prone to which season is an important factor to facilitate diagnosis. Knowledge of Endemic Disease which frequently happens and massively outbreak also guide the diagnosis procedure.

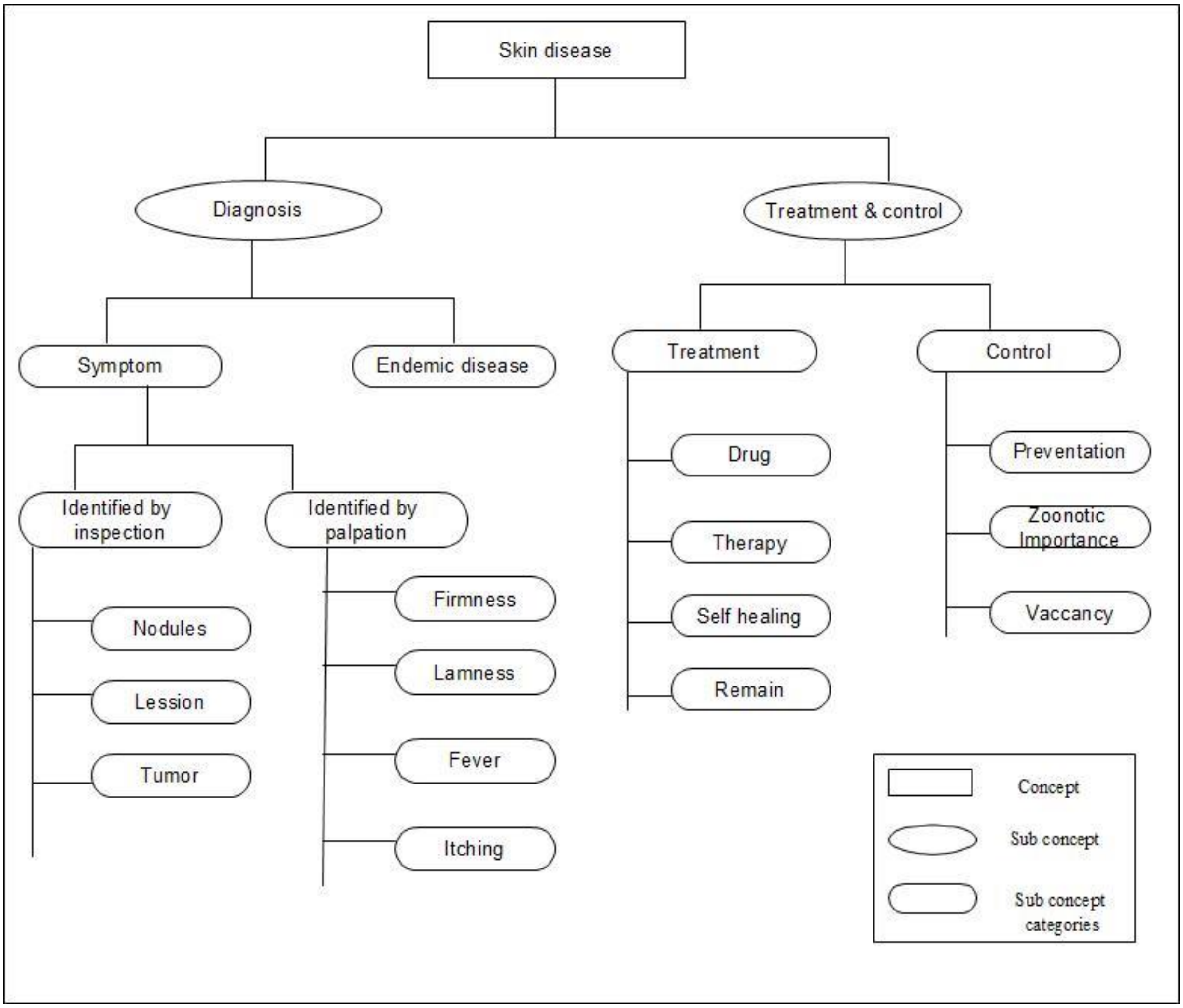

Figure 4-2 disease diagnosis conceptual model

Knowledge Representation: after the knowledge is modeled it is represented in the way which is suitable for operation in a system. The case is represented using cause and effect pairs and stored in the database. We use problem solution case representation for Location and their associated disease in that area.

Problem (Case Location): it represents the near city which the disease is occurred and contains list of city in Ethiopia where more disease is reported. Solution (Most occurred disease): it represent the disease occurred in that area. It includes the disease list occurred in that location.

For example: Case Location: Ambo

Most occurred Disease: LSD, Ringworm, Ticks

Data in the form of problem solution pair is extracted for each case location is stored in the database. Design of the databases is shown in Figure 4.3. Table CaseList contain data about location information. LocationCaseDisease table contain data about the diseases occurred in each locations. RuleMapping table contains data about each rules and for what disease they are formulated.

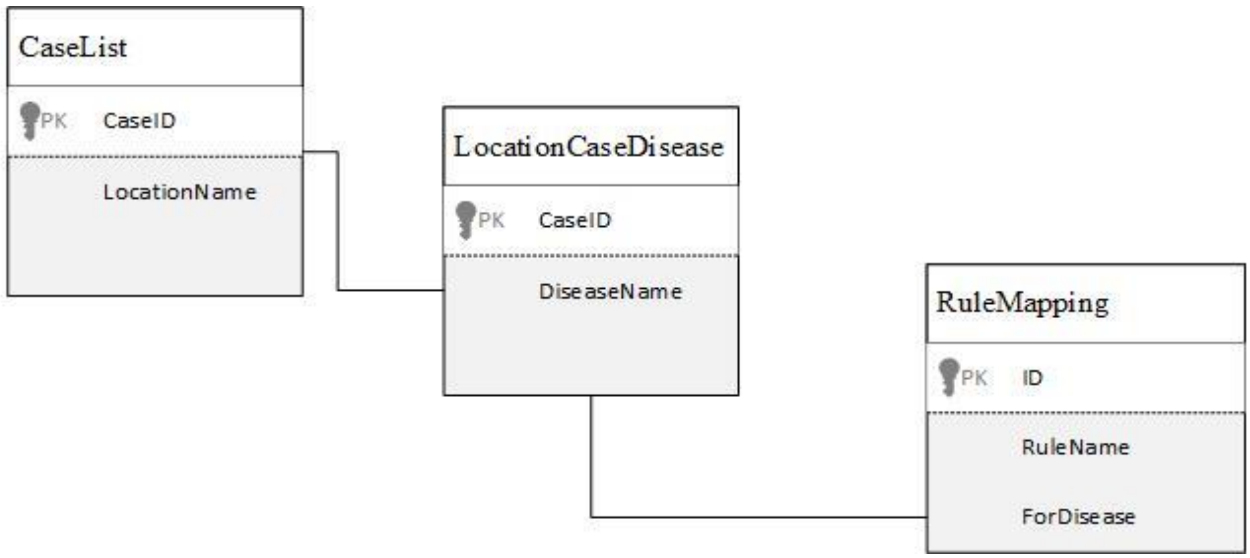

Figure 4-3 database design

The rules acquired are represented using if…Then representation, if part contains the symptoms encountered and then part contains the consequence of the symptom. The treatment rules are represented using if…then representation, if part contains the disease and then part contains the treatment of the disease.

4.3.4 Reasoning module

The reasoning module is responsible to seek information and relationship from the knowledge base to reach conclusion for the problem at hand. This component starts with CBR to acquire a disease that occurred in the input location case and RBR follows to reach a final conclusion. The general high level architecture of the reasoning module is shown in Figure 4.4.

We have corresponding map to signify which rules are for which disease. Each rule has an associated mapping that represents the set of preconditions determining the rule's applicability to a given situation. This helps to reduce the number of rules in order to optimize the inference engine execution speed which indeed increase the performance. For example rule 1 for LSD etc. If certain antecedents (symptoms) are evaluated as true, then it logically follows the consequent are proved, and then the diagnosis result and treatment of the disease is provided.

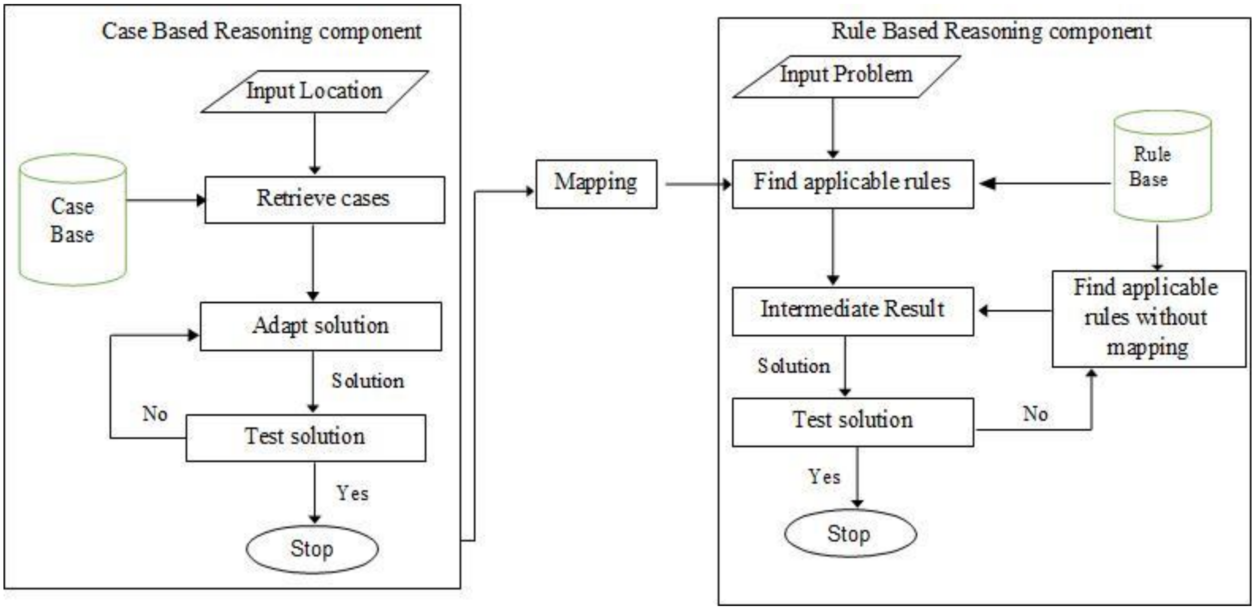

Figure 4-4 architecture of the reasoner component

Retrieve cases: case which are similar to the input cases are retrieved. Since, the case base is location case base we need exact match when cases are retrieved. The algorithm used to retrieve cases are described in Algorithm 4.2.

Algorithm 4-2 case retrieval algorithm

```
Input: input Case Cn, case stored in case base Cm=[cm1,cm2,cmk]
Output: Similarity level
Begin:
For each case in the input case
  Find the corresponding case in the stored case base
  Compare the two values to each other
      if Cn= Cmi
        Si=1
      else Si=0
END
Return Si
END
```

Adapt solution: the solution for user query is decided. If retrieval stage identifies diseases, then the solution of the retrieved cases will be used. But if CBR retrieves no case, then the query is forward to RBR component so that final decision will be inferred from rules. The work flow of the reasoning component is described in Algorithm 4.3.

Algorithm 4-3 reasoner algorithm

```
Input: input Location case, classified image, text information
Output: Possible disease
Begin:
For input Location case
 Location disease=result from CBR// list of diseases
 If (Location disease=NULL)// Location case not found in KB
    Location=input Location case
    Possible disease= fire rules without mapping//no disease is
                     stored on that location
     Store possible disease, Location in case base KB // as new
                      case for the location
 Else Fire rules with context // Location case found in KB
     Compute Mapping for Location disease with rule base
     Possible disease=fire rules with mapping
      If (Possible disease=NULL)//no solution means it is new or
                epidemic disease for that Location
           Possible disease= fire rules without mapping
           Store possible disease, location in case base KB// as
                     new case for the location
    END
```

4.3.5 Image Analysis

The image analysis module is responsible for classifying the input image into its categories. The analysis starts with preprocessing the input images and stop when the classification result found. Figure 4.5 shows the architecture of the image analysis module.

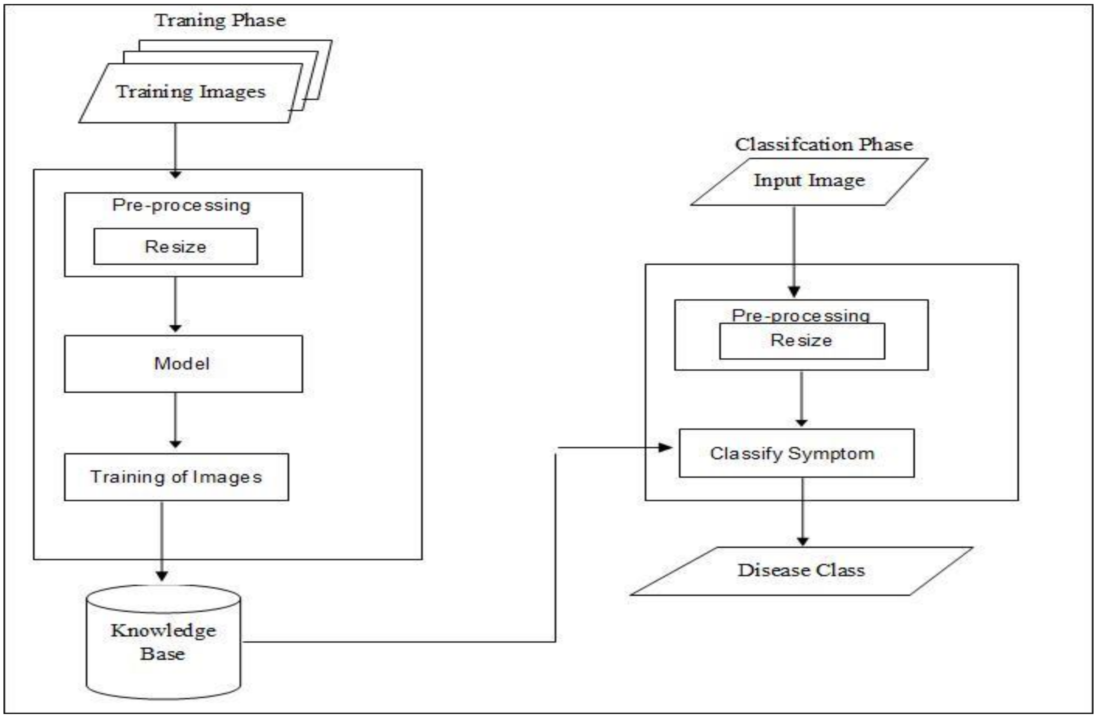

Figure 4-5 architecture of image analysis module

4.3.5.1 Pre-processing

The preprocessing component is responsible to make the images suitable for the overall image analysis activity. The main tasks done are resizing and normalizing. Resizing: is one of the preprocessing technique which brings the whole image into the same size. The image is collected from different sources and have different size. We decided a new image size that best reflects the contents of the image with less processing time. All images in our dataset are resized into 200 x 200 pixels by Algorithm 4.4.

Algorithm 4-4 image resizing algorithm

Input: Image
Output: resized Image
Begin:
For each image in dataset   Resized Image=resize(image, target_size=200,200)
Return resized Image
END

Normalizing is another preprocessing technique we use before further process. Image pixel values are integer between the ranges of 0 to 255. Although these pixel values can be presented directly to the model, it can result slower training time and overflow. Overflow is what happens when numbers get too big and the machine fails to compute correctly. So, by using Algorithm 4.5 we normalize our data values down to a decimal between 0 and 1 by dividing the pixel values with 255.

Algorithm 4-5 algorithm for normalizing image

Input: resized Image
Output: Normalized Image
Begin:

Divide_By=255.0
For each data in dataset
   Image=float32(image)
   Image=image/Divide_By
END

Return Image
END

4.3.5.2 Model

The architecture of the model compose layers which are responsible for feature extraction and classifying the input image into one of the categories. In our model 200x200 RGB image is passed through a stack of convolutional (conv) layers, where we use filters with 3 × 3 (which is the smallest size to skim pattern). The convolution stride is fixed to 1 pixel. The spatial resolution is preserved after convolution. Spatial pooling is carried out by three max pooling layers, which follow the conv layers. Max pooling is performed over a 2 × 2 pixel window, with stride 2. Figure 4.6 shows the proposed architecture of the CNN classification model.

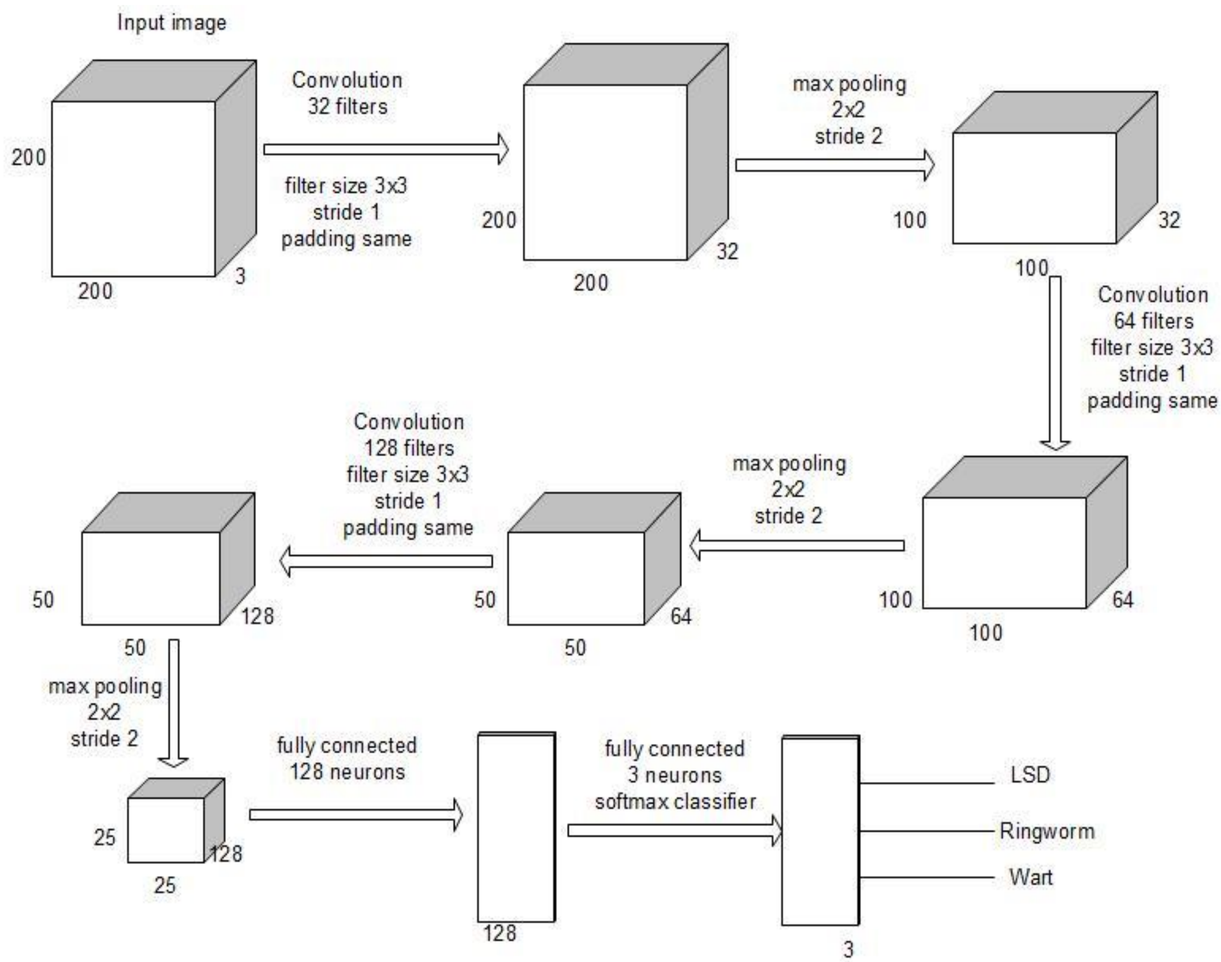

Figure 4-6 classification model architecture

In convolutional network the neurons are arranged in 3 dimensions width, height and depth. The conv layers consist of a set of learnable filters. Every filter is arranged with width and height, and extends through the full depth of the input volume. The first conv layer accepts 200x200x3 image and have filter with 3x3x3 where the last 3 represent the depth of the filter which is similar to the input image. As we slide the filter over the width and height of the input volume we will produce a 2 dimensional activation map that gives the responses of that filter at every spatial position. Then we will stack these activation maps along the depth dimension and produce the output volume. In the first conv layer 32 filters are applied which result in 200x200x32 activation map. The max polling layer will perform down sampling operation along the spatial dimensions (width, height), resulting in volume 100x100x32.

The second conv layer accept the result of the first conv layer 100x100x32 and apply its 64 filters to the input which result 100x100x64 activation map. Max poling follows to perform down sampling and result 50x50x64 volume. Its function is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network, and also control overfitting. The third conv layer apply 128 filters. The final fully connected layer compute the class scores, resulting in volume of size 1x1x3, where the 3 numbers correspond to a class score, among the 3 categories of our dataset. We observe adding more layers does not improve the performance on our dataset. It leads to overfitting, increased memory consumption and computation time. The removal of one layer from the model result poor performance because the model will not generalize enough with less number of layer in the model.

4.3.5.3 Train model

This component is where actually the model is trained for the task of image classification. The training start by defining optimizer, cost function and a metric. The workflow of the training is shown in the Figure 4.7.

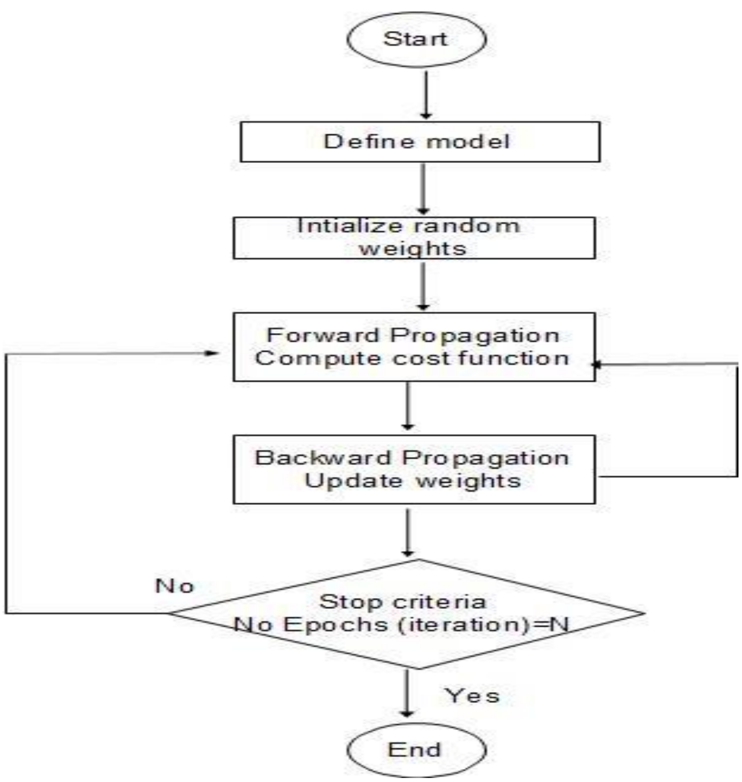

Figure 4-7 workflow of training algorithm

The goal of a convolutional layer is feature extraction. When filter weights move over an image it check for patterns in that section of the image. Filter weights change when the model is trained. In evaluation, these weights return high values if it thinks it is seeing a pattern it has seen before. The combinations of high weights from various filters let the network predict the content of an image.

The training is done using stochastic gradient descent with momentum (SGDM) optimizer, with learning rate of 0.001, mini-batch size of 64 for 50 number of epochs. The training takes nine hours with HP Intel core i5 -6200u CPU. The weight initialization is done using HE_UNIFORM initialization because it is compatible with the activation function we use RELU for its less training time. The algorithm for initializing weight is shown in Algorithm 4.6.

Algorithm 4-6 algorithm for filter weight initialization

```
Input: fin number of input units filter // is the number of
input units of the weight
Output: Initializer
Begin:
   Sd=sqrt(6.0/fin)
   For I in range (ni): //number of input nodes
       For J in range (nh): // number of hidden nodes
        Initializer [I,J]=np.float32(random.uniform(-sd,sd)
   Return Initializer
    END
```

The training (learning) in the model is presented as the function with the following form:

F(x) = (f_5(f_4(f_3(f_2(f_1(x))))) where: f_1(x): Function learned on first conv layer

f_2(x): Function learned on second conv layer

f_3(x): Function learned on third conv layer

f_4(x): Function learned on fourth hidden layer

f_5(x): Function learned on output layer

The function value is computed by iterative process of going and return through layers of the model. The going is a forward propagation of the information and the return is a back propagation of the information. The first phase forward propagation occurs when the network is exposed to the training data. Passing the input data through the network cause the neurons to apply their transformation to the information they receive from the neurons of the previous layer and send it to the neurons of the next layer. When the data has crossed all the layers, the final layer will reach into a result of label prediction for those input data. Algorithm 4.7 show the forward propagation process in training the model.

Algorithm 4-7 algorithm for feedforward propagation

```
Input: Training data, Initialized filter
Output: Z function Output
Begin:
Present the dataset to the network
For every data in the training dataset
 for each layer in the network
  for every node in the layer
   Calculate the weight of the inputs to the node
            Z=(x*W)+b // x is input, w is weight  , b is bias
   Calculate activation node Z=max (0,Z) // the activation RELU
   End
  End
End
Return function output
END
```

After forward propagation loss function is calculated to estimate the error. The loss function is used to compare and measure how well or bad our prediction result in relation to the correct result. Since our model task is classification the loss we use is Cross Entropy Loss which can be defined by the following equation.

$$crossEntropyLoss = -(Ty\ log(Py) + (1 - Ty)\log(1 - Py) \qquad [4.1]$$

Where Ty is the ground truth label for y,

Py is the predicted label for y.

The loss information calculated propagated backwards starting from the output layer to all neurons in the hidden layer. The neurons of the hidden layer only receive a fraction from the total loss, based on the contribution each neuron has contributed to the original output. This process is repeated, layer by layer, until all the neurons in the network have received a loss signal that describes their relative contribution to the total loss. This is done by calculating the partial derivate of the cost function with relative to the weight of each neurons. Algorithm 4.8 show backward propagation algorithm used while training the CNN model.

Algorithm 4-8 algorithm for backward propagation

```
Input: network assigned input and output
Output: weight update parameter
Begin:
Propagate the errors backward through the network
for all hidden layers
 for every node in the layer
   Calculate the nodes signal error
   Update each nodes weight in the network
       end
     end
   Return weight update parameter
    END
```

4.4 Summary

Diagnosis is a process of identifying possible diseases based on symptoms and signs. We propose a diagnosis system with integration of expert system and image processing using deep learning. The proposed system can achieve effective diagnosis using different mediums of input. For each input we follow appropriate processing methods. The images are preprocessed and classified into their category based on a trained CNN model. The CNN model constitutes a feature extractors and classifications layer. The location data extracted from the images enables the system to know the epidemic capability of the disease and guide the diagnosis procedure. The knowledge engineering component is responsible to build knowledge and fact needed for the diagnosis. Then the required knowledge and facts are represented in a way suitable for manipulation in computer system. Based on the result from the classification model and the text information the system reaches a final diagnosis conclusion. Then it will give a treatment recommendation based on the identified disease.

CHAPTER 5 : IMPLEMENTATION AND EVALUATION

5.1 Introduction

In this chapter, we discuss the experiments carried out to test the effectiveness of our proposed system. The data set used, results achieved in the classification process and the system performance will be discussed.

5.2 Dataset preparation

The process used for getting data ready for the classification model can be summarized in three steps: collect data, preprocess data and transform data. We follow this process in iterative with many loops to prepare the dataset required.

Step 1 is concerned with collecting available data needed to solve the problem. Image data is collected from D/m University, Addis Ababa veterinary school, Internet and other secondary sources. The step 2 is about getting the collected data into a form that can be easy to work. We follow formatting and cleaning steps to preprocess the data.

The formatting is about making the data selected in a format that is suitable for the work. The collected images are converted into JPEG format. JPEG format is selected because most of the collected images are in JPEG format and extraction of location information is possible with this format. The cleaning steps is about cleaning (correcting) missing data, unknown value. There are image data with unidentified labels, we remove image which we can't get the prober label for them.

In step 3 we transform the data collected. This steps is related to making the dataset suitable for the algorithm used and knowledge of the problem domain. In diagnosis we only need the part of the cattle which is infected or symptoms occurred. So, we transform the collected image by cropping the area where symptoms are present. The original images are cropped into images containing symptoms regions. The cropping follow the following rules, constitute healthy and symptom part, isolated symptoms taken individually, widespread symptoms taken both as a whole and divided into regions. Below we are demonstrating the transformation steps we take for sample image Figure 5.1.

Figure 5-1 LSD infected cattle

After transformation based on the specified rules we transform the input images into Figure 5.2 several images. In addition to transform our data, this step also increase our dataset. The data we collect before and after transformation is shown in Table 5.1.

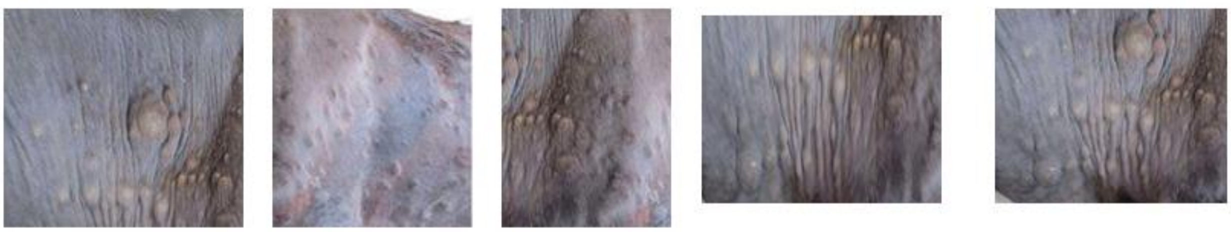

Figure 5-2 LSD image after transformation

Our model require large amount of labeled data. But getting enough data is a major problem in our cases which leads to use other techniques to expand our dataset. Data augmentation provides a means for increasing the quantity of training data available for machine learning, and is particularly relevant when training deep learning systems from scratch [61].

Table 5:1 collected data before and after transformation

No	Disease	Number of collected data	After Transformation
1	Lumpy skin disease (LSD)	84	146
2	Ringworm	57	100
3	Wart	64	124

In order to combat the high expense of collecting thousands of training images, image augmentation is used. Image Augmentation is the process of taking images that are in a training dataset and manipulating them to create many altered versions of the image. It provide more images to train and expose our classifier to a wider variety of transformed images to make the classifier more robust. It have been widely used on small datasets for combatting over-fitting [57, 62, 63].

Techniques of augmentation used in our dataset include horizontally and vertically flipping, zoom, shear and rotate. Figure 5.3 shows LSD infected cattle images after applying augmentation.

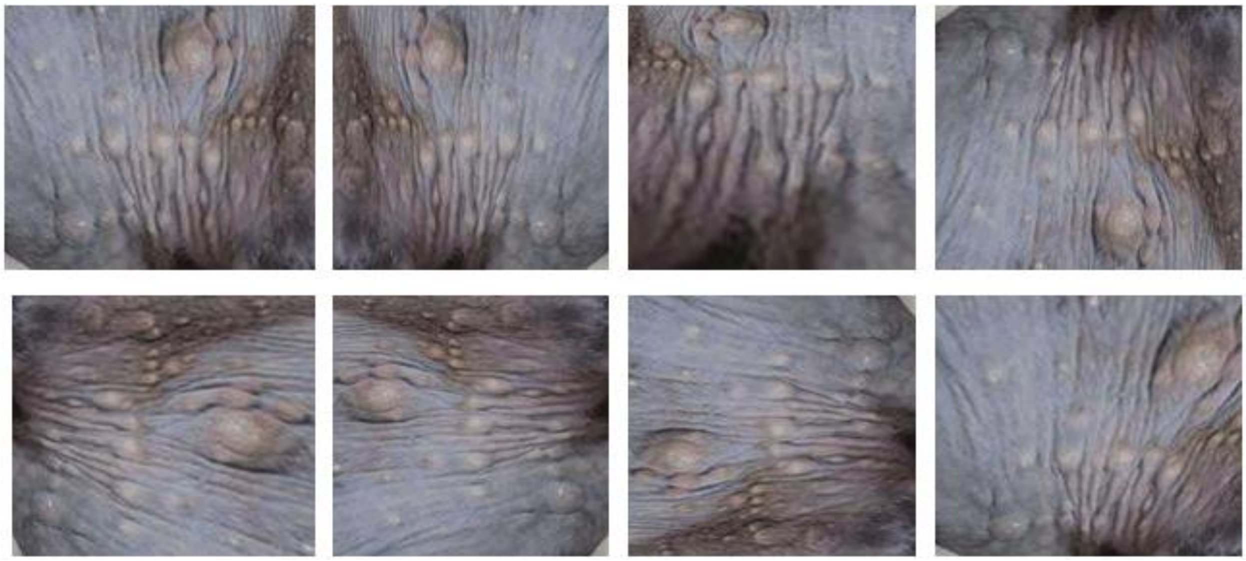

Figure 5-3 images after augmentation technique applied

After applying augmentation our dataset expand to 3990. From these images 90% of the dataset is used for training and 10% of the dataset is used for testing.

5.3 Development Environment

We use different development tools while developing the porotype. Java programming language with android development environment is used to work with user application. The expert system is also implemented by Java. The image analysis is done on Anaconda Keras. To bring the model to mobile environment we use TensorFlow. To augment image data we use Augmentor. Augmentor is a Python package made available under the terms of the MIT license. The package emphasis on providing operations that are typically used in the generation of image data for machine learning problems.

5.4 System Prototype

To demonstrate the validity of the proposed system, we develop a porotype system. The prototype developed show the user interface and output of the system.

System user Interface: it allows the user to upload a photo and fill text information to the system. After the inputs are feed into the system, location extraction is done. Then the image is passed to the processing components and diagnosis result is given to the user.

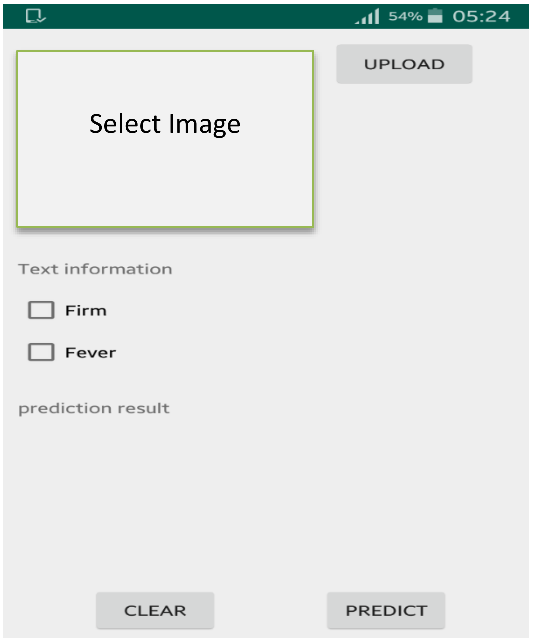

Figure 5-4 prototype user interface

5.5 System evaluation

The accuracy of the system is affected by performance of the classification model, so evaluation is done on the model. We evaluate the performance of the classification model using deep learning evaluation techniques accuracy and confusion matrix. Then accuracy of entire system is evaluated with user evaluation.

The performance of the model will be poor either by overfitting or underfitting the data. The training of the model is plotted to see possibility of overfitting and underfitting in the model. Overfitting happens when a model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. When the training accuracy is above the test accuracy it means the model is overfitting. Our model is not overfitting as shown in the Figure 5.5. There is not significant difference between the value of training and test accuracy.

Figure 5-5 plot of training and testing accuracy

Underfitting refers to a model that can neither model the training data nor generalize to new data. When validation loss is below the training loss the model is underfitting. As shown in the figure 5.6 our model is not underfitting.

The model achieve 95% accuracy in 50 epochs. Continuing the training above 50 epochs, the model try to learn the data and the noise and the performance is not changing but overfitting happen.

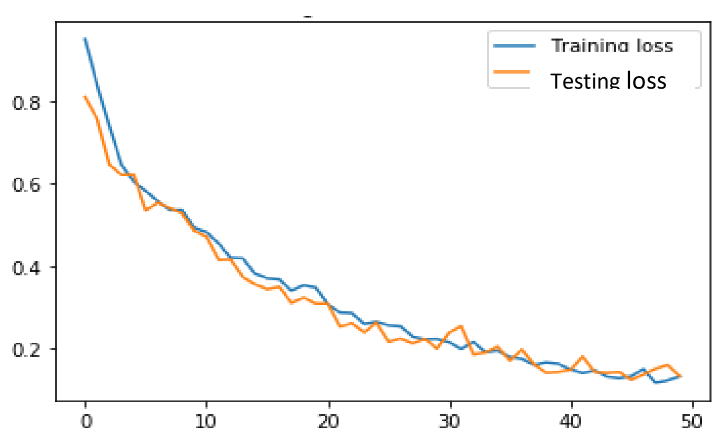

Figure 5-6 plot of training and testing loss

To summarize the performance of the model we use confusion matrix. In the confusion matrix the number of correct and incorrect predictions are summarized with count values and broken down by each class. Figure 5.7 shows how many of the images are misclassified and classified correctly.

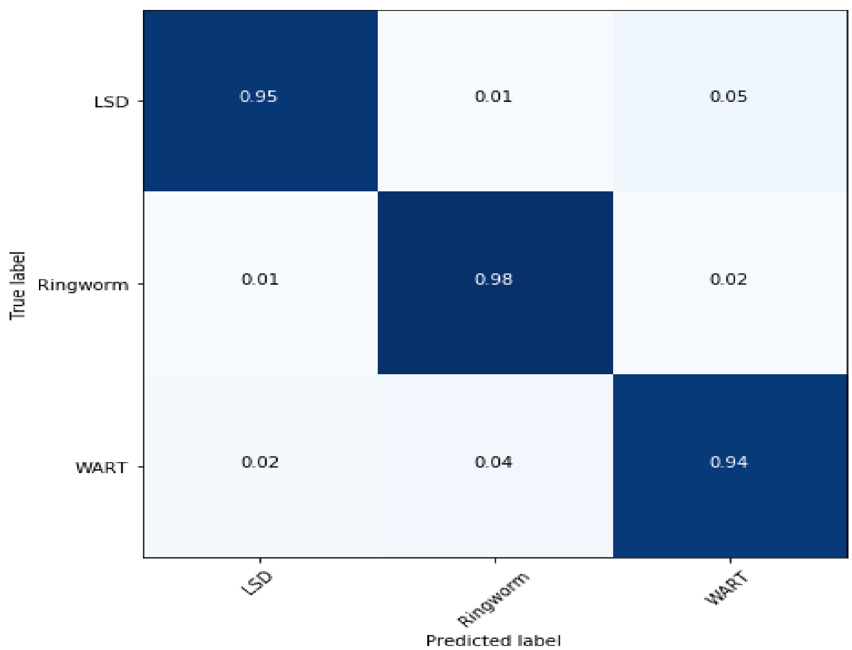

Figure 5-7 confusion matrix of the model

The ringworm classification is high compared to other, because it have easily distinctive feature from the other two class. The misclassification of LSD as Wart, Wart as LSD is noticed. Because LSD and ringworm have more similar symptoms which is sometimes difficult to distinguish distinctive feature among them.

The entire proposed system is evaluated by 4 people, we select randomly from different profession. 2 of them are veterinarians, and the rest are individuals who have cattle farming. The selection is based on the assumption that those with veterinary background can see and evaluate technical details while others may evaluate the applicability, accuracy and importance of the system. Before starting the evaluation process, the system was explained in detail to the evaluators. The questioner we use for system evaluation is shown in Appendix c where user puts the weight for each evaluation. Weight values shows the value of the evaluation, where 3 indicate the highest, 1 the lowest and 2 indicate medium.

For organizing the response of the evaluation we construct Evaluation matrix. It is a table with one row for each evaluator and columns address evaluation questions. The evaluator response is recorded in Table 5.2, where:

EQ1: the prototype system user friendly

EQ2: response time of the system

EQ3: validity of symptom description

EQ4: correctness of diagnosis conclusion

EQ5: correctness of treatment advice

EQ6: applicability of the system

EQ7: performance of the system

Table 5:2 records of evaluator evaluation

	EQ1	EQ2	EQ3	EQ4	EQ5	EQ6	EQ7
1	3	3	3	3	3	2	3
2	3	3	3	3	3	2	2
3	3	2	3	3	2	3	3
4	3	3	3	3	2	3	3

In EQ1, the evaluation result shows the prototype is user friendly. It's easy for anyone to use because the interaction doesn't need any background knowledge. The evaluation result of EQ2, shows the response time for user query is good. The response time is constrained by storage and processing capability of user phone. For all user the response time will be different. In EQ3, the evaluation result shows high value for the validity of symptom description. So representing symptoms using image is valid diagnosis method in cattle disease.

In EQ4, the diagnosis result shows high value, which validate our diagnosis approach. In EQ5, the correctness of treatment shows lower result. The veterinary recommend considering body mass index when drug is prescribed, which our system didn't include.

In EQ6 and EQ7, the diagnosis result show the performance and applicability of the proposed system high. We can conclude the integration of expert system and image processing using deep learning gives efficient and timely diagnosis for cattle diseases.

5.6 Comparison with deep neural network

Comparison is done with deep neural networks using the same dataset and parameter with architectural difference in the model. The deep neural network consist five layers as our CNN model. The model achieve 71 % accuracy at 50 epochs.

The difference in the performance of the two models came with their architecture. The CNN convolutional layer extract feature which can represent the dataset effectively. In DNN the input is flattened to feed to the system and it lose spatial information about the pixels. Also DNN poor performance came because the model over fit the training data as shown in Figure 5.8.

Figure 5-8 plot of DNN training and testing accuracy

CHAPTER 6 : CONCLUSION AND FEATURE WORKS

6.1 Conclusion

Many works are done so far in the area of cattle disease diagnosis. Almost all of them consider the description of the symptom is text. Recently some works have been done to incorporate images as symptom descriptions in plant and human disease diagnosis. But they use manual feature extraction, which is not efficient because of human constraints, and imagination put when designing the feature extractor.

A diagnosis approach is proposed by integrating an expert system and image processing using deep learning models. Diagnosis starts by acquiring information on the occurred disease through image and text. Symptoms identified by inspection are acquired by capturing the image through mobile phones. Symptoms identified by palpation is presented to the system using text. To know the epidemic capability, location information is presented to the diagnosis system. Then image are preprocessed and class of the images is identified by the trained CNN model. The final diagnosis conclusion is drawn by the reasoner component of the expert system using image classification results, text and location information.

We have used different tools and development environments while developing the prototype. Evaluation of the system is carried out using different evaluation criteria. The evaluation result shows that the proposed approach provides an effective diagnosis for cattle disease.

6.2 Contribution

The contributions of this research work are:

- ✓ Cattle disease diagnosis Image dataset
- ✓ Image analysis and expert system integration architecture
- ✓ Image classification model
- ✓ Method to use location information to known incidence of Epidemic disease
- ✓ Prototype of image and text based cattle disease diagnosis system.

6.3 Feature work

This research work explores different areas that can be further improved as well as some components that should be implemented and integrated for better functioning of the system.

- ✓ Increasing the dataset to include other disease.
- ✓ Adding disease which are examined in laboratory.
- ✓ Design to include body mass index when treatment is recommended to the proposed architecture.

REFERENCE

[1] Engidu Gebre-Amanueal, Fekade Getahun and Anteneh Assalif, "Web Based Expert System for Diagnosis of Cattle Disease," *Invited Paper in Proceedings of the 10th International Conference on Management of Emergent Digital ecosystems (MEDES'18),* ACM, Tokyo, Japan, s2018.

[2] Derejaw Lake, 2013, "Web-based Expert System for Cattle Diseases Diagnose", Master's Thesis. Department of Computer Science, Addis Ababa University, Addis Ababa, Ethiopia.

[3] Berhanu Aebissa, 2012, "Developing a Knowledge based system for coffee disease diagnosis and treatment", Master's Thesis. Department of computer Science, Addis Ababa University, Addis Ababa, Ethiopia.

[4] Ahmad Zamsuri, Wenni Syafitri, Muhamad Sadar, "Web Based Cattle Disease Expert System Diagnosis with forward Chaining Method", *IOP Conf. Ser.: Earth Environ. Sci,* 2017.

[5] Getahun Tigistu, "Automatic Flower Disease Identification using image processing", Master's Thesis. Department of computer Science, Addis Ababa University, Addis Ababa, Ethiopia.

[6] Daniel Hailemichael Lemessa, 2015, "Development of Automatic Maize Quality Assessment System Using Image Processing Techniques", Master's Thesis. Department of computer Science, Addis Ababa University, Addis Ababa, Ethiopia.

[7] Muktar Rashid, Robert Shank, Technical Report, Rough Guide to Animal Diseases in Ethiopia, http://www.africa.upenn.edu/eue_web/diseas94.htm, retrieved on October 28.

[8] Adel Hamdan Mohammad, Nedhal Abdul Majied Al Saiyd, "A Framework for Expert Knowledge Acquisition", *International Journal of Computer Science and Network Security, VOL.10 No.11*, November 2010.

[9] Daniel Teshome and Samuel Derso, "Prevalence of Major Skin Diseases in Ruminants and its Associated Risk Factors at University of Gondar Veterinary Clinic, North West Ethiopia", J *Veterinar Sci Technol S: S13-002. Doi:10.4172/2157-7579.1000S13-002,* 2015.

[10] Tadesse Birhanu, “Prevalence of the Major Infectious Animal Diseases Affecting Livestock Trade Industry in Ethiopia”, *Journal of Biology, Agriculture and Healthcare ISSN 2224-3208* Vol.4, No.17, 2014.

[11] Teweldemedhn Mekonnen, “Characterization of productive and reproductive performances, morphometric and challenges and opportunities of indigenous cattle breeds of Ethiopia”, *Int. J. Livest. Prod Vol. 9(3), pp. 29-41, DOI: 10.5897/IJLP2017.0426*, March 2018.

[12] Jason Brownlee, “Master Machine Learning Algorithms”, 2016.

[13] https://en.wikipedia.org/wiki/Medical_diagnosis retrieved on October 12.

[14] U. Raghavendra, N.S. Bhat, A. Gudigar, U.R. Acharya, “Automated system for the detection of thoracolumbar fracture using a CNN architecture”, *Future Generation Computer Systems*, *https://doi.org/10.1016/j.future.2018.03.023*, 2018.

[15] Mohammed El-Helly, “Integrating diagnostic expert system with image processing via loosely coupled technique”.

[16] Seblewongel Esseynew, 2011, “Prototype knowledge based system for Anxiety Mental Disorder Diagnosis”, Master’s Thesis. Department of Computer Science, Addis Ababa University, Addis Ababa, Ethiopia.

[17] Solomon Gebremariam, 2013, “A Self-Learning Knowledge Based System for Diagnosis and Treatment of Diabetes”, Master’s Thesis. Department of Computer Science, Addis Ababa University, Addis Ababa, Ethiopia.

[18] Kindu Wondmnew, Mohamed Hussien, Dagnachew Endalamaw and Wosen Temesgen, “Determination of prevalence of external parasite on different species”, “*Advance Biotechnology & Microbiology*”, *Volume 10 Issue 5*, 2018.

[19] Werkineh Gelalcha Olkeba1, Edilu Jorga Sarba, Belay Abebe Deres and Endrias Zewdu Gebremedhin, “Prevalence of major skin diseases of cattle and associated risk factors around Ambo Town, Ethiopia”, 2016.

[20] Addise Ambilo, Achenef Melaku, "Major Skin Diseases of Cattle: Prevalence and Risk Factors in and around Hawassa Southern Ethiopia", "*Journal of Advanced Veterinary Research*", *Volume 3,* 2013.

[21] Aklilu Feleke, 2008, "Major Animal Health Problems Of Market Oriented Livestock Development In Bure Woreda", A thesis submitted to the Faculty of Veterinary Medicine, Haramaya University, as a partial fulfillment for the requirements of the Degree, Doctor of Veterinary Medicine (DVM).

[22] Ahmed Ali Farah Gumbe, "Review on lumpy skin disease and its economic impacts in Ethiopia", "*Journal of Dairy Veterinary and Animal Research*", *Volume 7 Issue 2,* 2018.

[23] Yogesh Kumar, Yogyata Jain, "Research Aspects of Expert System", *International Journal of Computing & Business Research,* 2012.

[24] Ararsa Duguma, "Practical Manual on Veterinary Clinical Diagnostic Approach", *J Vet Sci Technol 7: 337. Doi:10.4172/2157-7579.1000337,* 2016.

[25] Beyene T, Assefa S, Ayana D, Jibat T, Tadesse F, et al, "Assessment of Rational Veterinary Drugs Use in Livestock at Adama District Veterinary Clinic, Central Ethiopia", J Veterinar Sci Techno 7: 319. Doi:10.4172/2157-7579.1000319, 2016.

[26] Solomon, Chris, and Toby Breckon, "Fundamentals of Digital Image Processing: A practical approach with examples in Matlab". John Wiley & Sons, 2011.

[27] Bassem S. Abu-Nasser, "Medical Expert Systems Survey", *International Journal of Engineering and Information Systems (IJEAIS) Vol. 1 Issue 7,* September 2017, Pages: 218-224

[28] Shweta R. Astonkar, V. K. Shandilya, "Detection and Analysis of Plant Diseases Using Image Processing Technique", "*International Research Journal of Engineering and Technology*", *Volume: 05 Issue: 04,* 2018.

[29] Sujatha R, Y Sravan Kumar and Garine Uma Akhil, "Leaf disease detection using image processing", "*Journal of Chemical and Pharmaceutical Sciences*", March 2017.

[30] Hongyan Gao, Guimiao Jiang, Xiang Gao, Jianhua Xiao, Hongbin Wang," An equine disease diagnosis expert system based on improved reasoning of evidence credibility", *INFORMATION PROCESSING IN AGRICULTURE, https://doi.org/10.1016/j.inpa.2018.11.003*", 2018.

[31] Pravin S. Ambad, A. S. Shirsat, "A Image analysis System to Detect Skin Diseases", "*IOSR Journal of VLSI and Signal Processing (IOSR-JVSP*)", 2016.

[32] Er.Shrinidhi Gindhi, Ansari Nausheen, Ansari Zoya, Shaikh Ruhin, "An Innovative Approach for Skin Disease Detection Using Image Processing and Data Mining", "*International Journal of Innovative Research in Computer and Communication Engineering*", Vol. 5, Issue 4, April 2017.

[33] Nikos Petrellis, "A Smart Phone Image Processing Application for Plant Disease Diagnosis"," *6th International Conference on Modern Circuits and Systems Technologies (MOCAST)"*, 2017.

[34] Melaku Tefera, "Global Crisis and the Challenge of Veterinary Teaching in Ethiopia", *Global Veterinaria 5 (6): 294-301, 2010 ISSN 1992-6197*, 2010

[35] Desalegn Aweke Wako, "Development of Knowledge Based System for Wheat Disease Diagnosis: A Rule Based Approach", "*Journal of Computer Engineering & Information Technology*", 2017.

[36] Whisnu Ulinnuha Setiabudi, Endang Sugiharti, Florentina Yuni Arini, "Expert System Diagnosis Dental Disease Using Certainty Factor Method", "*Scientific Journal of Informatics Vol. 4, No. 1"*, 2017.

[37] Rahat Yasir, Md. Ashiqur Rahman, and Nova Ahmed, "Dermatological Disease Detection using Image Processing and Artificial Neural Network", "*8th International Conference on Electrical and Computer Engineering*", December, 2014.

[38] Amarathunga, E.P.W.C. Ellawala, G.N. Abeysekara, C. R. J. Amalraj, "Expert System for Diagnosis of Skin Diseases", *international journal of scientific & technology research volume 4, issue 01* January 2015.

[39] H. T. YACOB, B. NESANET, A. DINKA, "Prevalence of major skin diseases in cattle, sheep and goats at Adama Veterinary Clinic", *Oromia regional state, Ethiopia, Faculty of Veterinary Medicine, Addis Ababa University, Debre Zeit, ETHIOPIA.*

[40] Nikos Petrellis, "A Review of Image Processing Techniques Common in Human and Plant Disease Diagnosis", "*Symmetry2018, 10, 270, doi:10.3390/sym10070270*", 2018.

[41] Kenilworth, N.J., "The Merck Veterinary Manual", 2015.

[42] Edward Boden, "*Black's veterinary dictionary"*, 2005, ISBN: 9780713687637.

[43] Peter J.F Lucas and Linda C. Van der Gaag, "principles of expert system", *Center of mathematics and computer science, Amsterdam Addison Wesley*, 1991.

[44] Standard Veterinary Treatment Guidelines for Veterinary Clinics, Drug Administration and Control Authority of Ethiopia, January, 2010.

[45] Berhane Hagos, "Ethiopian Cattle Genetic Resource and Unique Characteristics under a Rapidly Changing Production Environment A Review", *International Journal of Science and Research (IJSR),* 2016.

[46] W. Molla, M. C. M. de Jong and K. Frankena, "Temporal and spatial distribution of lumpy skin disease outbreaks in Ethiopia in the period 2000 to 2015", *BMC Veterinary Research*, 2017.

[47] Tianyi Liu, Shuangsang Fang, Yuehui Zhao, Peng Wang, Jun Zhang, "Implementation of Training Convolutional Neural Networks", *University of Chinese Academy of Sciences, Beijing, China.*

[48] Munirah, M. Y., S. Suriawati, and P. P. Teresa, "Design and Development of Online Dog Diseases Diagnosing System." *International Journal of Information and Education Technology*, 2016.

[49] Patil.S. P, Kumbhar.V.P, Yadav.D.R, Ukirade.N.S, "Detection of Leaf Diseases by Image Processing", *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 4, Issue 4,* April 2015

[50] https://www.msdvetmanual.com/integumentary-system/integumentary-system, diagnosis of skin disease, retrieved on October 24.

[51] Diederik Kingma and Jimmy Ba, "Adam: A Method for Stochastic Optimization", *conference paper at ICLR* 2015.

[52] Sebastian Ruder, "An overview of gradient descent optimization algorithms", *Insight Centre for Data Analytics, NUI Galway Aylien Ltd.*, 2017.

[53] D. H. Hubel and T. N. Wiesel, "Receptive Fields, Binocular Interaction and Functional Architecture In The Cat's Visual Cortex", *J. Phyiiol. (1962),* 160. 1962.

[54] Yann Lecun, Leon Bottou, Yoshua Bengio and Patrik Haffner,"Gradient based learning applied to document recognition", *IEEE,* 1998.

[55] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", *conference paper at ICLR* 2015.

[56] Shiliang Sun, Zehui Cao, Han Zhu, and Jing Zhao, "A Survey of Optimization Methods from a Machine Learning Perspective", June 2019.

[57] Catherine R. Alimboyong, Alexander A. Hernandez, Ruji P. Medina, "Classification of Plant Seedling Images Using Deep Learning", *IEEE Region 10 Conference*, October 2018.

[58]Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks".

[59]Sajjad M, Khan S, Muhammad K, Wu W, Ullah A, Baik SW, "Multi-Grade Brain Tumor Classification using Deep CNN with Extensive Data Augmentation", *Journal of Computational Science*, 2018.

[60] Carl Louie Aruta, Colinn Razen Calaguas, Jan Kryss Gameng, "Mobile-based Medical Assistance for Diagnosing Different Types of Skin Diseases Using Case-based Reasoning with Image Processing", *International Journal of Conceptions on Computing and Information Technology Vol. 3, Issue. 3*, October 2015.

[61] P. Spyridonos, E.I. Papageorgiou, P.P. Groumpos, and G.N. Nikiforidis, "Integration of Expert Knowledge and Image Analysis Techniques for Medical Diagnosis".

[62] Artzai Picona, Aitor Alvarez-Gilaa, Maximiliam Seitzc,d, Amaia Ortiz-Barredob, Jone Echazarraa,, Alexander Johannes, "Deep convolutional neural networks for mobile capture

device-based crop disease classification in the wild”, *Computers and Electronics in Agriculture*, https://doi.org/10.1016/j.compag.2018.04.002, 2018.

[63] Muhammad Imran Razzak, Saeeda Naz and Ahmad Zaib, “Deep Learning for Medical Image Processing: Overview, Challenges and the Future, Classification in BioApps”, *Lecture Notes in Computational Vision and Biomechanics 26, https://doi.org/10.1007/978-3-319-65981-7_12*. 2018.

[64] K. Fukushima and S.miyake “necognitron: A new algorithm for pattern recognition tolerant of deformation and shifts in position”, *pattern recognition, vol 15, pp 455 469*, 1982.

[65] Seyyid Ahmed Medjahed, “A Comparative Study of Feature Extraction Methods in Images Classification”, “*I.J. Image, Graphics and Signal Processing*”, 2015.

[66] Dong ping Tian, “A Review on Image Feature Extraction and Representation Techniques”, “*International Journal of Multimedia and Ubiquitous Engineering*, Vol. 8, No. 4, July, 2013.

Appendix A: Interview Questions

Interview Questions

1. What are the main cattle skin diseases common in Ethiopia?
2. What are the main causes this disease?
3. What are the symptoms of this disease?
4. How do you differentiate clinical signs b/n skin diseases?
5. Which skin disease is commonly mistaken (difficult to identify or similar symptoms) and how do you identify it?
6. Is clinical sign seen in eyes can be sufficient to diagnosis skin disease?
7. What kind of information additional to observed symptoms is need to decide if the cattle is infected with skin disease?
8. Is there any type in LSD, wart which shows different symptoms if yes what are they?
9. How do you recognize a disease as epidemic diseases?
10. How frequent do outbreaks occur? How fast can an outbreak be controlled?
11. What is the effect of cattle skin diseases for cattle owners as well as the economy?
12. Can skin diseases transmittable to humans and to the herd? If so, which ones and how the transmission be prevented?
13. What about their curability? To what level are the diseases getting proper medication?
14. What methods can be used for prevention of skin cattle diseases?

Appendix B: sample rule and case used in the system

Sample Rules used in the system

Rule 1: IF *Image classified as = LSD* AND (OR) (the lesion is firm) AND (OR) (*discharge from the eyes and noses*); then *the disease is: Lumpy Skin disease.*

Rule 2: IF *Image classified as* = Ringworm AND (OR) (*loss of appetite OR weight loss*); then *the disease is: Ringworm.*

Rule 3: IF *Image classified as* =Wart AND interface with basic activity in cattle; then *the disease is: Wart.*

Sample case base location with corresponding disease list

No	Case Location	Possible disease
1	Hawassa	LSD, Ringworm, Acariasus, Pediculosis, Dermatophilosos
2	Ambo	Wart, Ticks, Lice, Mange mites, Dermatophilosos, Photsentization, Branding
3	Gondar	LSD, Tick, Lice, Dermatophilosos, Mange
4	Adama	LSD, Ticks, Pediculosis, Demodicosis and Dermatophilosis

Appendix C: System evaluation questionnaires

The proposed system evaluation questionnaires

Objective of the questionnaires: to evaluate the proposed Image based cattle disease diagnosis.

Instruction: Please put X for each evaluation question. If your evaluation is low= 1, medium= 2, high= 3.

	1	2	3
User friendly interface of the prototype system			
Response time of the prototype system			
Validity of symptom description			
Correctness of diagnosis result			
Correctness of treatment and recommendation			
Applicability and performance of the prototype system			

Declaration

I, the undersigned, declare that this research is my original work and has not been presented for degree in any other university, and that all sources of materials used for the research have been acknowledged.

Declared by:

Name: Bezawit Lake

Signature: ___________________

Date: _______________________

Confirmed by advisor:

Name: Fekade Getahun (PhD)

Signature: _________________

Date: ____________________

Place and date of submission: Addis Ababa University, October 2019.