International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 14 Issue 2 –APRIL 2015.

Identification and Detection of Seed Borne Diseases of Soybean Using Image Processing - A Survey

Anita Kinnikar[1], Padmashree Desai[2] and Shamarao Jahagirdar[3]

[1,2]Dept. Of CSE, B.V.B. College of Engg. & Tehnology, Hubli, Karnataka
[3]Dept. of Plant Pathology, AICRP on Soybean, University of Agricultural Sciences, Dharwad, Karnataka

***Abstract* - Soybean is an important commercial crop known as the "GOLDEN BEAN". Damage is an important quality factor for grading, marketing and end use of Soybean. Seed damage can be caused by weather, fungi, insects, artificial drying, and by mechanical damage during harvest, transportation, storage and handling. Seed-borne pathogens causes enormous losses to crops in the world as well as in India. The presence of pathogenic propagules in a seed lot is pivotal because infected seed may fail to germinate, causes infection to seedlings and growing plants. The crop currently earns about Rs. 6976 crores of foreign exchange through exports of defatted oil cake. India now ranks 4th in terms of global soybean area sown and 5th in terms of soybean production after USA, Brazil, Argentina and China. Currently, soybean yield losses due to individual disease/insect/weed species ranges from 20 to 100 per cent. However, with integrated pest management schedule, 30-35 per cent additional yield can be obtained.**

Main challenge here is without disintegration finding the defected seeds in big lot. Image processing helps to extract low level features such as colour, texture and shape to identify the diseases and classify them. The paper discusses detail study of diseases, causes and different techniques that can be used to identify and detect them. Input data set consists of 100 Soybean colour images and also microscopic images used in the research at University of Agricultural Science, Dharwad. Proposed work has greater social impact by helping farmers to identify the seed borne diseases at early stage and thereby increasing yield.

Index Terms: Detection; Features; Identify; Image processing; Seed borne;

I. Introduction

The Soybean or Soya bean (Glycine max) is a legume species native to East Asia, which is highly cultivated for its edible seed. The plant has been classified under oil producing plant for its edible seed rather than for its pulse by the Food and Agricultural Organization (FAO). Soybean is an annual legume categorized under Fabaceae family. Soybean is one of the most important bean among all in the world which provides vegetable protein for millions of human and ingredients for thousands of chemical products. It is most nutritious and easily digested food of the bean family. The soybean is considered as one of the richest and cheapest sources of protein. It is a staple in the diet of humans and animals in different corners of world today. The seed contains 17 percent oil and 63 percent meal, 50 percent of which is protein. Soybean is a good source of protein for diabetics as it contains no starch. Globally, the most important feed grain legume is soybean (Glycine max), with a total production of 216,144,262 tonnes and harvested area of 94,899,216 hectors. Today, the world's top most producers of soybean are USA, Brazil, Argentina, India and China. Approximately 85 percent of the world's soybeans are processed, or "crushed," annually for the production of soybean meal and oil and 98 percent of the soybean meal that is crushed is further processed for preparation of animal feed. About 95 percent is consumed as edible oil; the rest is used for industrial products such as in the production of fatty acids, soaps and biodiesel.

Seed-borne pathogens present a serious threat to seedling establishment. Close association with seeds facilitates the long-term survival, introduction into new areas and widespread dissemination of pathogens. Under greenhouse conditions, the risks of significant economic losses due to diseases are great because factors including high populations of susceptible plants, high relative humidity, high temperatures and overhead irrigation, promote explosive plant disease development. Under these conditions, the most effective disease management strategy is exclusion which is accomplished by using seed detection assays to screen and eliminate infested seed lots before planting. The conventional method of disease detection in plants using naked eye observation method is cubersome and is non effective. Using computer vision toolbox the disease detection in plants is efficient and is not time consuming. The process of diagnosis used in plants (i.e. recognition of symptoms and signs of diseases) is completely based on the use of scientific techniques. On the basis of symptoms of particular diseases and with the help of agricultural scientists, identification of diseases becomes easier. Plant pathologists can analyze the digital images using digital image processing toolbox in Matlab for diagnosis of plant diseases. Detection and recognition of diseases in plants using machine learning is very fruitful in providing symptoms of identifying diseases at its earliest.

This paper mainly addresses the issues related to the soybean. The issues includes, Identifying and detecting seed-borne diseases of soybean using image processing. Once the disease is detected then quantifying the effect of diseases in terms of percentrage.

A. Organization of Paper

Section II provides related work done by researches. Section III discusses about the Image features, color models, color moments and types of seedborne diseases. Section IV deals with proposed approach. Section V specify experimental setup used for work. Section V summarizes the paper in the form of conclusion.

II. Related work

Marcelo de Carvalho Alves & et al[1] proposed a system Scanning electron microscopy (SEM) to detect seed-borne fungi in seeds submitted to blotter test. Two different conditions of blotter test and water restriction treatment were evaluated. In the blotter test, seeds were subject to conditions that enabled pathogen growth and expression, while the water restriction method consisted in prevent seed germination during the incubation period, resulting in the artificial inoculation of fungi. In the first condition, seeds of common bean (Phaseolus vulgaris L.), maize (Zea mays L.) and cotton (Gossypium hirsutum L.) were submitted to the standard blotter test and then prepared and observed with SEM. In the second condition, seeds of cotton (G. hirsutum), soybean (Glycine max L.) and common bean (P. vulgaris L.) were respectively inoculated with Colletotrichum gossypii var. cephalosporioides, Colletotrichum truncatum and Colletotrichum lindemuthianum by the water restriction technique, followed by preparation and observation with SEM. The standard SEM methodology was adopted to prepare the specimens. Considering the seeds submitted to the blotter test, it was possible to identify Fusarium sp. on maize, C. gossypii var. cephalosporioides and Fusarium oxysporum on cotton, Aspergillus flavus, Penicillium sp., Rhizopus sp. And Mucor sp. on common bean. Structures of C. gossypii var. cephalosporioides, C. truncatum and C. lindemuthianum were observed in the surface of inoculated seeds. Scanning electron microscopy methodology was a potential alternative for study and identification of seed-borne fungi in seeds of common bean (Phaseolus vulgaris L.), maize (Zea mays L.) and cotton (Gossypium hirsutum L.) after blotter test preparation. Scanning electron microscopy enabled to observe fungi lesions and fungi signals in seeds of cotton (Gossypium hirsutum L.), soybean (Glycine max L.) and common bean (Phaseolus vulgaris L.) inoculated by water restriction technique associated with Colletotrichum gossypii var. cephalosporioides, Colletotrichum truncatum and Colletotrichum lindemuthianum, respectively. The adoption SEM methodology in blotter test analysis could be useful to identify fungal structures that enable ensuring the implementation of correct diagnoses as well as to facilitate conduction of more detailed taxonomic classification of seed-borne fungi.

Sudhir Gupta and et al.[2], proposed a cost-effective way to segregate the seeds. Such a system would not only facilitate soybean grading but also serve as a quality control tool for processing facilities such as seed cleaning plants and oil mills. Here we incorporate image processing as a tool to obtain the results. The idea can be extended further to separate other seeds in a similar manner. Color features in the RGB (red-green-blue) color space are computed. A feed-forward neural network is trained to classify sample soybean seeds. Almost all bad seeds are correctly identified. A computer vision-based system can be developed for automated sorting and packing to speed up the distribution to the farmers and industries. The method once implemented is expected to serve as an easy and a reliable technique to achieve the separation of the good soybean seeds, assuring to be a fast and an accurate method using color image analysis. The results of this study show that color features and a properly trained neural network can effectively classify soybeans. A computer vision–based system could be developed for automated sorting. The classification accuracy was acquired under laboratory setting, so it had some limits. In future work, a large quantity of soybean seeds, not one at a time, will be investigated.

Md. Muzahid-E-Rahman and et al[3], proposed a system in which *Fusarium moniliforme* J. Sheld was observed as important seed-borne pathogen of soybean (*Glycine max*) having a great potentiality to cause seed-borne infection after establishing a rapid pathogenesis. The infection process was studied with scanning electron microscopy (SEM). Infection hyphae grew directly from the side or near the end of the conidium and entered the seed coat through cuticular cell juncture with or without forming appressorium. After invasion of the fungus, rapid degradation of cell wall occurred followed by intercellular and intracellular development of the fungus. Finally, tissues of the seed lost their integrity and identity and seemed as rotten mass covering with dense mycelium.

Yankun Peng and et al.[5] *China Agricultural University China,* proposed system uses among the optical analysis methods, near-infrared (NIR) spectroscopy is the most popular method because of its non-destructive nature, the low operating cost and the fast response times[11], and it also has been successfully applied to quality control in food[12][13], petrochemical, pharmaceutical, clinical and biomedical and environmental sectors (Ripoll et al., 2008). Near-infrared (0.7- 2.5μm; 12900-4000cm-1) spectroscopy is further classified into NIR reflectance spectroscopy and NIR transmission spectroscopy. NIR can be non-dispersive (filter-based instrumentation), dispersive and use Fourier transform-based instrumentation. All these researches have shown the possibility and reasonability for determination of pesticide concentration using NIR spectroscopy. Pesticide concentration can be readily measured with NIR spectroscopy and optical imaging technology. However the accuracy and precision could be improved. There is a need to develop rapid optical techniques for pesticide determination which could be used in the future for agro-food safety assurance. The optical technique could be one of the most useful tools along with the advancement of spectral instrument for determination of pesticide residue.

P.K.Sajeesh, M.S.L.Rao and Shamarao Jahagirdar[6] in "Molecular detection, transmission and histopathological studies of seed-borne fungal infection of soybean (glycine max (l.) Merill)", proposed system uses the early detection of seed borne infection by novel molecular techniques like Polymeric Chain Reaction (PCR) using fungal specific primer has tested here. ITS1 and ITS4 were used for the diagnostic purpose of seed-borne fungal infection. Oligonucleotide specific primers targeting the ITS region have been demonstrated to selectively detect several agriculturally important fungi. Further the amplified product sequenced and revealed the sample genome has homology with Cercospora kikuchii and Rhizoctonia bataticola genome. The pathogenic ability of seed borne fungi was proved in transmission studies carried out by seedling symptom test and pot culture studies. Finally the histopathological studies revealed the location of seed borne fungi in the infected soybean seeds. In most of the observations the fungal pathogens were occupied in the pericarp, hilum and endosperm regions.

III. Image and its features

*A. CIE L*a*b* Color Space*

Also called CIELAB color model the second uniform color space derived from CIE XYZ space in1976, with white reference point . L*a*b* color model determines the color depending on its position in a 3D color space, the L* components is the lightness of the color (when L* =0 means black and when L* =100 referred to while) and the chroma*(for positive values indicate red and for negative values indicate green) and the hue b* (positive values refer to yellow while negative values refer to blue) as illustrated in Fig. 1. CIELAB is device independent and considered very important for desktop color[8].

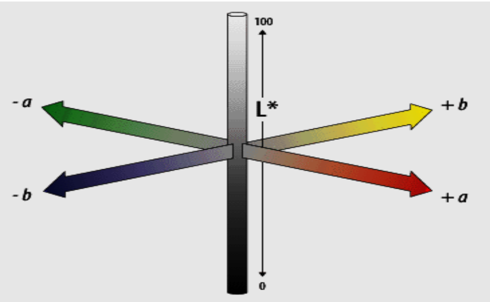


Fig. 1. CIE L*a*b* color model [8].

B. RGB color model

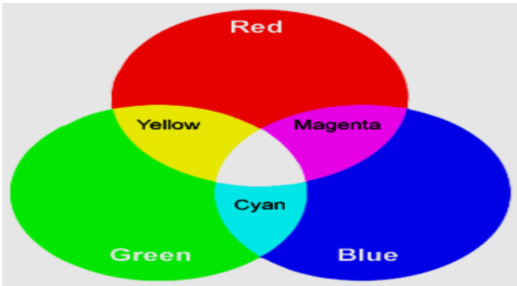
From the three additive primary colors (red, green and blue) the name of the model has been derived and in a light spectrum they combined together as one color, and can be mixed to produce new spectrum colors. The RGB color space could be represented as a cube by normalized RGB color values in the range [0,1] with gray values on the main diagonal of the black values (0,0,0) and on the opposite corner the white values (1,1,1). It is considered as the base color model for most image applications since the acquired image does not need any further transformation for displaying in the screen. RGB color model is classified into two types; Linear RGB Color Space, and Nonlinear RGB Color Space. Referring to linear RGB color model as (RGB), and to nonlinear RGB color model as (R'G'B'), which will be explained in the following subsections:[8]

C. Linear RGB Color Space

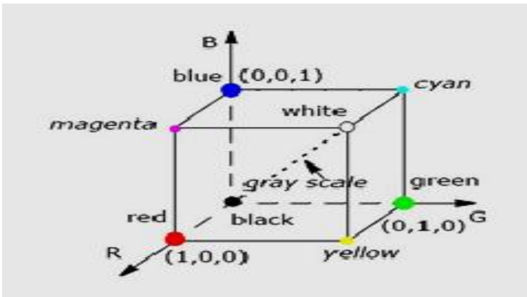
Linear RGB space attains color consistency via various appliances using color management system. Linear RGB not suitable for numerical analysis and seldom used for image representation, it is used for computer graphics applications. The mapping to nonlinear done using gamma correction factorγ of the camera or any input device, in the range [0,1] for both of the models[8].

D. Non-linear RGB Color Space

The data of input image captured with a camera or scanner are the R'G'B' values represented in the range from 0 to 255. These data then stored for using in image processing applications, JPEG, MPEJ standard. The transformation from linear to nonlinear values, and from nonlinear back to linear RGB values within the range [0, 1] is defined in definition 1and 2 [8] :



(a)



(b)

Fig. 2. RGB color Model [8].(a) Primary colors representation. (b) Primary colors cube.

E. Types of Seedborne diseases

In Table 1 we have listed the diseases which we are going to detect using image processing techniques.It describes the type of diseases and its characteristics.

Table 1. Fungi seedborne diseases

Fungus	Importance
Phomopsis	Causes pod and stem blight disease. Infection levels of 20% and above could indicate a serious effect on viability.
Cercospora	Causes purple seed stain disease. Discolors the seed and ruptures the seed coat. May cause a small reduction in germination but not as severe as Phomopsis.
Fusarium	Some evidence suggests that Fusarium may reduced seed germination. However, the significance of seed infection is not well known.
Bacillus spp	Commonly associated with reduced seed viability.
Aspergillus	Germination will be reduced if seed has been stored at too high a moisture counter for several weeks.
Penicillium	Germination will be reduced if seed has been stored at too high a moisture counter for several weeks.
Alternaria	There is no evidence to suggest a reduction in seed quality by these fungi.
Chaetomium	There is no evidence to suggest a reduction in seed quality by these fungi.
Cladosporium	There is no evidence to suggest a reduction in seed quality by these fungi.

IV. Proposed work

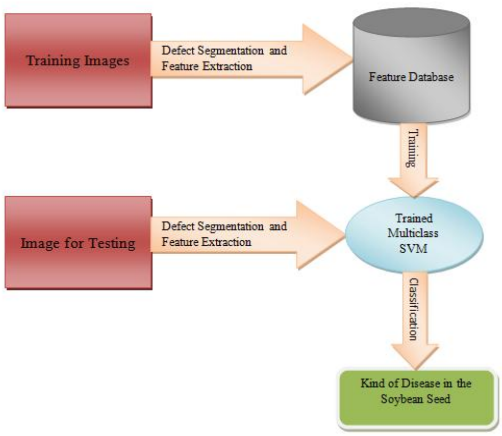


Fig. 3. Proposed Approach

Image categorization, in general, relies on combinations of structural, statistical and spectral approaches. Structural approaches describe the appearance of the object using well-known primitives, for example, patches of important parts of the object. Statistical approaches represent the objects using local and global descriptors such as mean, variance, and entropy. Finally, spectral approaches use some spectral space representation to describe the objects such as Fourier spectrum[14]. In this paper, we introduce a method which exploits statistical color and texture descriptors to identify seed borne diseases of soybean in a multi-class scenario.

The steps of the proposed approach are shown in the Fig.1. Defect segmentation, feature extraction, training and classification are the major tasks to be performed. For the seed borne disease identification problem, precise image segmentation is required; otherwise the features of the non-infected region will dominate over the features of the infected region. K-means based defect segmentation is used to detect the region of interest which is the infected part only in the image. The proposed approach operates in two phases, i.e. training and classification. Training is required to learn the system with the characteristics of each type of diseases. First we extract the feature from the segmented portion of the images that are being used for the training and store in a feature database. Then, we train support vector machine with the features stored in the feature database. Finally any input image can be classified into one of the classes using feature derived from segmented part of the input image and trained support vector machine[9].

A. Defect Segmentation

Image segmentation is a convenient and effective method for detecting foreground objects in images with stationary background. Background subtraction is a commonly used class of techniques for segmenting objects of interest in a scene. This task has been widely studied in the literature. Specular reflections, background clutter, shading and shadows are the major factors that affect the efficiency of the system. Therefore, in order to reduce the scene complexity, it might be interesting to perform image segmentation focusing on the object's description only. K-means clustering technique is used for the defect segmentation. Images are partitioned into four clusters in which one or more cluster contains only infected region of the soybean. K-means clustering algorithm was developed by J. MacQueen (1967) and later by J. A. Hartigan & M. A. Wong (179). The K-means clustering algorithms classify the objects (pixels in our problem) into K number of classes based on a set of features. The classification is carried out by minimizing the sum of squares of distances between the data objects and the corresponding cluster.

B. Feature Extraction

Different colour and texture moments can be used to extract features. The Colour moments suitable for this system are Global Color Histogram, Color Coherence Vector, Local Binary Pattern, and Completed Local Binary Pattern.

C. Global Color Histogram (GCH)

The Global Color Histogram (GCH) is the simplest approach to encode the information present in an image[14]. A GCH is a set of ordered values, for each distinct color, representing the probability of a pixel being of that color. Uniform normalization and quantization are used to avoid scaling bias and to reduce the number of distinct colors [14].

D. Color Coherence Vector (CCV)

An approach to compare images based on color coherence vectors are presented by Pass, Zabih, & Miller. They define color coherence as the degree to which image pixels of that color are members of a large region with homogeneous color. These regions are referred as coherent regions. Coherent pixels are belongs to some sizable contiguous region, whereas incoherent pixels are not. In order to compute the CCVs, the method blurs and discretizes the image‘s color-space to eliminate small variations between neighboring pixels. Then, it finds the connected components in the image in order to classify the pixels of a given color bucket is either coherent or incoherent. After classifying the image pixels, CCV computes two color histograms: one for coherent pixels and another for incoherent pixels. The two histograms are stored as a single histogram.

E. Supervised Learning and Classification

Supervised learning is a machine learning approach that aims to estimate a classification function f from a training data set. The trivial output of the function f is a label (class indicator) of the input object under analysis. The learning task is to predict the function outcome of any valid input object after having seen a sufficient number of training examples. In the literature, there are many different approaches for supervised learning such as Linear Discriminant Analysis (LDA), Support Vector Machines (SVMs), Classification Trees, Neural Networks (NNs), and Ensembles of Classifiers[16]. Recently, a unified approach is presented by[15] that can combine many features and classifiers. The author approaches the multi-class classification problem as a set of binary classification problem in such a way one can assemble together diverse features and classifier approaches custom-tailored to parts of the problem. They define a class binarization as a mapping of a multi-class problem onto two-class problems (divide-and-conquer) and referred binary classifier as a base learner. For N-class problem $NX(N-1)/2$ binary classifiers will be needed where N is the number of different classes[9].

V. Experimental setup

The experimental setup uses a Matlab Ver 7.0 onwards for identifying and detecting seed borne diseases of soybean by extracting low level features such as color, texture and shape. Input data set consists of 100 soybean digital color images and also microscopic images used the research at University of Agricultural Science, Dharwad. Here we are comparing these two types of images for better efficiency in finding seed borne diseases in earlier stage. The experimental results helps to determine normal soybean and the diseased soybean and CLBP feature shows more accurate result for the identification of seed borne diseases of soybean.

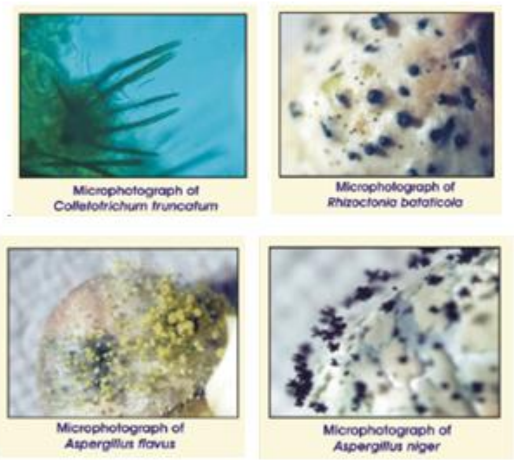


Fig. 4. The microphotograph of various fungus affected to the soybean

Fig. 5. The soybean seeds infected with different fungus.

VI. Conclusion

Commercial production of soybean is a potential source of income to farmers, traders and government. Reliable detection of seed borne diseases in early stages is essential for economic, production and agricultural benefits. Key contribution of the paper is to identify the issues related to seedborne diseases and their root cause. An image processing based approach is proposed in this paper for seed borne disease identification and detection problem. The proposed approach is composed of mainly three steps. In the first step defect segmentation is performed using K-means clustering technique. In the second step features are extracted. In the third step training and classifications are performed on a Multiclass

SVM. We have used different types of seed borne soybean diseases as a case study. Proposed work has greater social impact by helping researchers and farmers to identify the seed borne diseases at early stage and thereby increasing yield.

References

[1] Marcelo de Carvalho Alves and Edson Ampelio Pozza(2012). “Scanning electron microscopy detection of seed-borne fungi in blotter test”. Current microscopy contributions to advances in science and technology (A. Mendez-Vilas, Ed.).

[2] Sudhir Gupta, Smitha Patil, Asha Kulkarni and Zoheb Hukkeri, “Classification of Soybean seeds by Color Image Analysis”.

[3] Md. Muzahid-E-Rahman, Md. Rustom Ali, Tapan Kumar Dey, Md. Delowar Hossain(2010). “Scanning electron microscopy of invasion process by fusarium moniliforme on soybean seed”. Journal of plant protection research.

[4] Dipali Majumder, Thangaswamy Rajesh, E. G. Suting, Ajit Debbarma(2013). “Detection of seed borne pathogens in wheat: recent trends”. Australian journal of crop science. AJCS 7(4):500-507 (2013). ISSN:1835-2707.

[5] Yankun Peng, Yongyu Li and Jingjing Chen *China Agricultural University China,* “Optical Technologies for Determination of Pesticide Residue”.

[6] P. K. Sajeesh, M. S. L. Rao and Shamarao Jahagirdar(2014). “Molecular detection, transmission and histopathological studies of seed-borne fungal infection of soybean (glycine max (l.) Merill)”. An international quarterly journal of life sciences. The Bioscan 9(1): 247-251, 2014.

[7] Sabah Bashir, Navdeep Sharma(2012). “Remote Area Plant Disease Detection Using Image Processing”. *IOSR Journal of Electronics and Communication Engineering (IOSRJECE) ISSN : 2278-2834 Volume 2, Issue 6, PP 31-34*.

[8] Noor A. Ibraheem, Mokhtar M. Hasan, Rafiqul Z. Khan, Pramod K. Mishra(2012). “Understanding Color Models: A Review”. ARPN Journal of Science and Technology. VOL. 2, NO. 3. ISSN 2225-7217.

[9] *Shiv Ram Dubey and Anand Singh Jalal, GLA University Mathura, India, “*Adapted Approach for Fruit Disease Identification using Images*”* .

[10]Anagha Deshpande(2014), “*A Review on Preharvesting Soybean Crop Pests and Detecting Techniques*”. International Journal of Advance Research in Computer Science and Management Studies. Volume 2, Issue 2. ISSN: 2321-7782.

[11] Armenta, S., S. Garrigues, and M. de la Guardia(2007). “Partial least squares-near infrared determination of pesticides in commercial formulations”. *Journal of Vibrational Spectroscopy*. 44: 273-278.

[12] Subbiah, J., C.R. Calkins, A. Samal, and G.E. Meyer(2008). “Visible/near-infrared hyperspectral imaging for beef tenderness prediction”. *Journal of Computers and Electronics in Agriculture*. 64: 225-233.

[13] Leroy, B., S. Lambotte, O. Dotreppr, H. Lecocq, L. Istasse, and A. Clinquart(2003). “Prediction of technological and organoleptic properties of beef longissimus thoracis from near-infrared reflection and transmission spectra. *Meat Science”*. 66: 45-54.

[14]Gonzalez, R., & Woods, R. (2007) *Digital Image Processing, 3rd edition*, Prentice-Hall.

[15] Rocha, A., Hauagge, C., Wainer, J., & Siome, D. (2010). “Automatic Fruit and Vegetable Classification from Images”. *Computers and Electronics in Agriculture*, 70, 96-104.

[16] Bishop, C. M.(2006). “*Pattern Recognition and Machine Learning”. 1st edition*, Springer.