IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A SYSTEMATIC MACHINE LEARNING APPROACH FOR DETECTION OF HEMOGLOBIN LEVELS

Author's Information:

1. Anup Bhashkarrao Chaudhary

Student of Master of Engineering at Department of Electronics & Telecommunication, Sipna COET, Amravati

2. Dr. A. P. Thakare

Professor at Department of Electronics & Telecommunication, Sipna COET, Amravati

Abstract

Digital Image processing techniques can be used to explore primitive diagnostics methods for disease detection at early stages with limited resources and skilled labour. These techniques can also assist doctors during clinical examination without any need for invasive pathological test and this facilitates patient comfort and avoids infection during blood test. Various blood components such as hemoglobin and billirubin whose approximate measure can directly be identified by just viewing the colour of patient skin, nails, eye or any other target area can be measured and classified in terms of the colour content of the image of the targeted area. Analysis of image processing techniques in conjunction with specialized supervision can provide significant exploration in the field of biomedicine and clinical applications. This research work proposes an image processing based non-invasive method of measuring haemoglobin (Hb) concentration present in patient's blood by analyzing the color and texture of digital photographs of patient's palpebral conjunctiva. The images of patient's palpebral conjunctiva were processed and eight relevant features were extracted .Artificial neural network classifier was used to correlate the output quantity to be measured with the values of the quantity measured by the standard method as per the guidelines given by WHO. Further, based on the testing results obtained by the classifier the patients whose Hb concentration was less than 11g/dL were screened as anaemic patients. A confusion matrix was then plotted to evaluate and compare the predicted classification results with the actual value of Hb obtained from invasive test. It was found that the proposed algorithm was able to diagnose anemia with 71.42% sensitivity and 89.47% specificity. The proposed method is helpful for detection of not only severe anemia but works well in detection of moderate anemia too thus predicting the hemoglobin value to an accuracy of 81.81%.The proposed work is useful for giving assistance to medical practitioners for reliable diagnosis of anaemia in the clinic itself and in low resource settings.

Introduction:

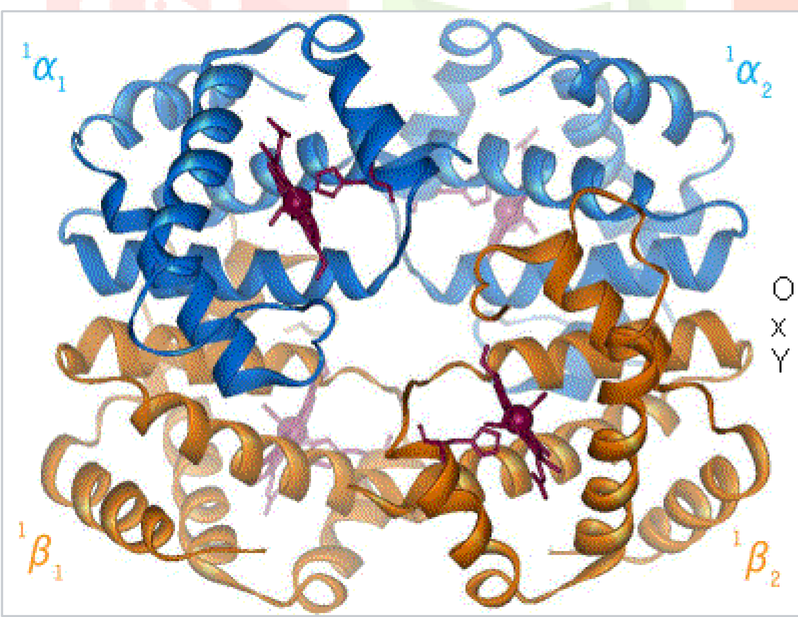
Anaemia is defined as a quantitative reduction of haemoglobin, the oxygen-carrying component of red blood cells [1]. The World Health Organisation (WHO) estimated that 24.8% of the global population was anaemic between 1993 and 2005 [2]. The gold standard for anaemia diagnosis is *ex-vivo* measurement of haemoglobin concentration in whole blood. This method requires venepuncture and specialised equipment, which may introduce delays or be unavailable in resource-poor settings [3]. Point of care testing methods for anaemia are widely available and typically involve analysis of blood by finger-prick sample. These methods are rapid and inexpensive but require liquid reagents and may expose healthcare workers to risks of blood-borne infections [4].

To rapidly screen for anaemia, clinicians often examine for conjunctival pallor. This involves subjective evaluation of the colour of the conjunctival membrane, with the presence of pallor indicating anaemia. Although this clinical sign can be useful, conjunctival pallor has a low sensitivity and specificity for prediction of anaemia and inter-observer agreement is poor [5–7]. Assessment can be improved by the use of colour-scale cards, which are compared directly to the conjunctiva, resulting in improved inter-observer agreement, sensitivity and specificity [8].

Haemoglobin predominantly absorbs green light and reflects red light, and as a result haemoglobin concentration affects tissue colour [9]. An “erythema index” (EI) has been developed to objectively quantify the degree of erythema of skin lesions, using digital photography followed by analysis of the red and green components of images [9,10].

In this study, we aimed to determine whether a conjunctival EI calculated from digital photographs taken in ambient lighting conditions correlates with haemoglobin concentration. Our goal was to develop a non-invasive method of anaemia detection using a consumer camera or smartphone. We found that EI of the palpebral conjunctiva correlated with haemoglobin concentration and compared favourably with clinician assessment. Our findings suggest that quantification of conjunctival pallor using a digital camera or smartphone has potential application as a non-invasive and affordable screening method for anaemia.

Literature Review:



A schematic visual model of oxygen-binding process, showing all four monomers and hemes, and protein chains only as diagrammatic coils, to facilitate visualization into the molecule. Oxygen is not shown in this model, but, for each of the iron atoms, it binds to the iron (red sphere) in the flat heme. For example, in the upper-left of the four hemes shown, oxygen binds at the left of the iron atom shown in the upper-left of diagram. This causes the iron atom to move backward into the heme that holds it (the iron moves upward as it binds oxygen, in this illustration), tugging the histidine residue (modeled as a red pentagon on the right of the iron) closer, as it does. This, in turn, pulls on the protein chain holding the histidine.

When oxygen binds to the iron complex, it causes the iron atom to move back toward the center of the plane of the porphyrin ring (see moving diagram). At the same time, the imidazole side-chain of the histidine residue interacting at the other pole of the iron is pulled toward the porphyrin ring. This interaction forces the plane of the ring sideways toward the outside of the tetramer, and also induces a strain in the protein helix containing the histidine as it moves nearer to the iron atom. This strain is transmitted to the remaining three monomers in the tetramer, where it induces a similar conformational change in the other heme sites such that binding of oxygen to these sites becomes easier.

As oxygen binds to one monomer of hemoglobin, the tetramer's conformation shifts from the T (tense) state to the R (relaxed) state. This shift promotes the binding of oxygen to the remaining three monomer's heme groups, thus saturating the hemoglobin molecule with oxygen.[65]

In the tetrameric form of normal adult hemoglobin, the binding of oxygen is, thus, a cooperative process. The binding affinity of hemoglobin for oxygen is increased by the oxygen saturation of the molecule, with the first molecules of oxygen bound influencing the shape of the binding sites for the next ones, in a way favorable for binding. This positive cooperative binding is achieved through stericconformational changes of the hemoglobin protein complex as discussed above; i.e., when one subunit protein in hemoglobin becomes oxygenated, a conformational or structural change in the whole complex is initiated, causing the other subunits to gain an increased affinity for oxygen. As a consequence, the oxygen binding curve of hemoglobin is sigmoidal, or S-shaped, as opposed to the normal hyperbolic curve associated with noncooperative binding.

The dynamic mechanism of the cooperativity in hemoglobin and its relation with low-frequency resonance has been discussed.

Methodology:

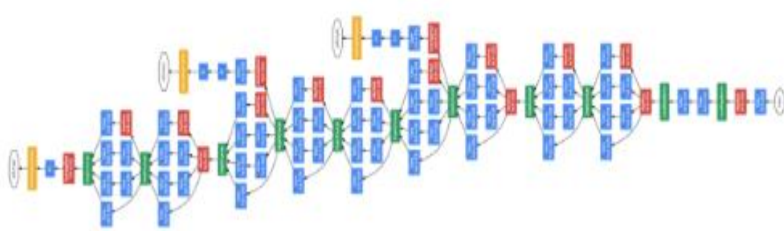
Conclusion Neutral Network

How to split a neural network

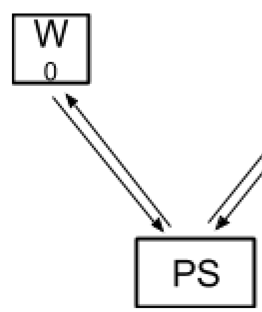
There are two ways to split load of a neural network into several machines:

Network / Model parallelism.

Today's neural networks are made of many layers. Each layer requires a **set of computations that are usually represented as a graph**. The data is passed along the calculation nodes of the graph in tensors which are just multidimensional arrays. For example an image can be made of 4 color channels (R, G, B, A) each of them being a 2 dimensional array (3 dimensions). If the input is a video we have a new dimension, time. If, for performance reasons, we pass more than one sample in a batch we add another dimension. In this approach each computation can be placed on a different device, **a device can be a CPU or GPU on the same machine or in a different one**. This way it is possible to get the best of each available device but it is needed to develop carefully and assign each operation to a specific device. This means that **it is needed to know in advance the hardware the model will run in**. This can work on a lab but it may not work well in an industrial / business environment.



Data parallelism. The whole model is run on a machine. **Several machines run the whole model with different parts of the data**. Each of the machines running the model is called a worker (W). Every given amount of time (usually after a batch) the different instances of the model share the parameters and reconcile them before continuing with the training. They send the information about the parameters to a set of machines called parameter servers (PS) that merge the results and send them back to the workers. This approach **solves the problem of knowing the hardware setup in advance**. In fact it can be changed at any time. This is a plus from the resiliency point of view. This can be very important in the case of very long computations. It also gives flexibility in the sense that it may be possible to increase or reduce the number of machines (computer capacity) to make room for more jobs. It also comes with a price. New machines are needed in the form of the parameter servers that will be in charge of gathering and merging parameters. **The reconciliation process can be synchronous or asynchronous**. Both have advantages and drawbacks. The asynchronous is more resilient and allows the usage of machines with different processing power but it may waste part ot the computation time. The synchronous one instead works better if all the machines have the same processing power and are close in the communication network. But it is more sensitive to any issue with the worker nodes. It also requires extra code to do the synchronization.



Environment: In order to evaluate data parallelism the following environment has been deployed. Not all the elements needed are described here, only the ones that have effect in the performance.

Data: As explained, the data will not be stored in the processing machines, it will be sent to them. This way we simulate a probable scenario where the data is received in streaming. For testing purposes the data can be stored in files in specific (non processing) machines or generated by a botnet simulating the behaviour of customers.

This setup simplifies the distribution of the data to a variable number of processing nodes. This is of course not the fastest way to process data in a machine learning environment. This task is done using a kafka cluster. The botnet publishes the events on a topic and the processing nodes acting as consumers read them and use them as input vector for the neural network.

RESULT:

Processing: The training of the neural network in a distributed way using Tensorflow requires the use of two type of nodes. The work itself is done by the worker nodes. There is also a set of nodes (parameter servers) that receive the different versions of the parameters generated by the worker nodes and reconcile them before sending them back to the worker nodes. The parameter reconciliation is done in asynchronous mode.

Image compression

An important development in digital image compression technology was the discrete cosine transform (DCT), a lossy compression technique first proposed by Nasir Ahmed in 1972.[14] DCT compression became the basis for JPEG, which was introduced by the Joint Photographic Experts Group in 1992.[15] JPEG compresses images down to much smaller file sizes, and has become the most widely used image file format on the Internet.[16] Its highly efficient DCT compression algorithm was largely responsible for the wide proliferation of digital images and digital photos,[17] with several billion JPEG images produced every day as of 2015.[18]

Digital signal processor (DSP)

Electronic signal processing was revolutionized by the wide adoption of MOS technology in the 1970s.[19] MOS integrated circuit technology was the basis for the first single-chip microprocessors and microcontrollers in the early 1970s,[20] and then the first single-chip digital signal processor (DSP) chips in the late 1970s.[21][22] DSP chips have since been widely used in digital image processing.[21]

The discrete cosine transform (DCT) image compression algorithm has been widely implemented in DSP chips, with many companies developing DSP chips based on DCT technology. DCTs are widely used for encoding, decoding, video coding, audio coding, multiplexing, control signals, signaling, analog-to-digital conversion, formatting luminance and color differences, and color formats such as YUV444 and YUV411. DCTs are also used for encoding operations such as motion estimation, motion compensation, inter-frame prediction, quantization, perceptual weighting, entropy encoding, variable encoding, and motion vectors, and decoding operations such as the inverse operation between different color formats (YIQ, YUV and RGB) for display purposes. DCTs are also commonly used for high-definition television (HDTV) encoder/decoder chips.[23]

Medical imaging

In 1972, the engineer from British company EMI Housfield invented the X-ray computed tomography device for head diagnosis, which is what is usually called CT (computer tomography). The CT nucleus method is based on the projection of the human head section and is processed by computer to reconstruct the cross-sectional image, which is called image reconstruction. In 1975, EMI successfully developed a CT device for the whole body, which obtained a clear tomographic image of various parts of the human body. In 1979, this diagnostic technique won the Nobel Prize.[4] Digital image processing technology for medical applications was inducted into the Space Foundation Space Technology Hall of Fame in 1994.

Tasks

Digital image processing allows the use of much more complex algorithms, and hence, can offer both more sophisticated performance at simple tasks, and the implementation of methods which would be impossible by analogue means.

In particular, digital image processing is a concrete application of, and a practical technology based on:

- Classification
- Feature extraction
- Multi-scale signal analysis
- Pattern recognition
- Projection

Some techniques which are used in digital image processing include:

- Anisotropic diffusion
- Hidden Markov models
- Image editing
- Image restoration
- Independent component analysis
- Linear filtering
- Neural networks
- Partial differential equations
- Pixelation
- Point feature matching
- Principal components analysis
- Self-organizing maps
- Wavelets

CONCLUSION

The study findings proved that the nutrition ball intervention increase the haemoglobin level and reduce the symptoms of anaemia among adolescent girls with iron deficiency anaemia. The subjects who received nutrition ball had a significant improvement in haemoglobin level. There was association between post test haemoglobin level and heavy menstrual flow, duration of menstrual cycle and drink tea/coffee per day.

REFERENCES:

1. Ann Marrier Tomey, Martha Raile Alligood, (2006). Nursing theroiests and their work (6th ed).Missouri, Mosby Publication.
2. Agarwal,LP. (2006). Modern educational research, New Delhi. dominant publishers and distributors.
3. Anne waugh, Allison Grant, 2004, Anatomy and physiology in health and illness Ninth Edition, Churchill Livingstone.
4. Basavanthapp, (1998). Community health nursing. (1st ed)., New Delhi, Jaypee Publication.
5. Basvanthappa.B.T,(2006). Nursing research. New Delhi, Jaypee Brothers Medical Publishers(Pvt) Ltd.
6. Basavanthappa, B.T (2003) medical –surgical nursingJaypee brothers, medical publishers (p) Ltds Bangalore.
7. Black and Hawks,(2005). Medical Surgical Nursing 7th edition. Volume1,Saunders Publications New Delhi.
8. Brunner and Siddarth (2008) Text of Medical Surgical Nursing 11th Edition,Wolters Kluwer India Pvt.Ltd., New Delhi .
9. Burns Nancy and groove susan, (2001). Understanding Nursing research, 2nd edition, sunders, Philadelphia. 90
10. Christopher.R.W, Edward, (1995) Principles and practice of Medicine Seventeenth Edition, Churchill Livingstone .
11. Davidson's, (2006) Principles and practice of Medicine 20th Edition, Churchill Livingstone publication.
12. K.K. Gulan, (2006). Community health nursing princeples a and practice. (1st ed). New Delhi , Kumar Publishing House.
13. M.C. Gupta, B.K. Mahajan, (2003). Preventive and social medicine. (3rd ed). New Delhi, Jaypee Publication.
14. Guptha s p (2003) Statistical methods New Delhi suthan chand educational publishers india, 31st revised edition.
15. Incredibly easy,(2004) Medical & Surgical Nursing Lippin cott, William& Wilkins, A wolters khuwer company,
16. Kumar and Clark, (2002) Clinical Medicine fifth Edition W.B. Saunders company