DOI: https://doi.org/10.24297/ijct.v22i.9219

A Comprehensive Review of Computer Vision Techniques to Interest Physicians and Surgeons, Role of A Clinical Biomechanical Engineer in Pre-Operative Surgical Planning, And Preamble To HSG-Amoeba, A New Concept of Biomedical Image Modeling Technique.

Gandhi, Harjeet Singh MD, FRCS, FRCSC, MSc (Biomed Eng.)

Orthopaedic surgeon, Clinical assistant, Hamilton Health Sciences, Hamilton, Ontario, Canada

harjeetg52@yahoo.co.uk

Abstract

Background: The science of computer vision is replication of human vision for pattern recognition and segregation of objects-of-interest at macro- and micro-level. There are numerous computer vision techniques with greater focus on deep learning utilizing artificial neural network. Only few of them can be readily applied to medical images for surgical interventions.

Study objective: As this narrative review is aimed at the medical community it is not encumbered with mathematical algorithms, albeit important. Apart from discussion on basic concepts and chronological development of the computer vision techniques the study introduces role of clinical biomechanical engineering team at the time of surgical planning.

Methodology: The study literature was searched on Google Scholar, on Google chrome, Wikipedia and cited in the reviewed articles referring to the original studies describing various computer vision techniques between 1980 to 2021.

Result: There is enormous discursive literature to read with extremely variable computer vision terminology unknown to the medical community is densely populated with advanced mathematics leading to lack of interest among majority of the physicians as the end user. There are inconsistencies in the usage of medical terminology and definitions.

Comments and conclusion: Standalone image processing and segmentation is meaningless without patient information for clinical applications in daily practice. There is a dire need for streamlining of computer vision science to teach medical community, introduction of a new breed of in-house clinical biomechanical engineers and supplementary residency program for residents to accept it as standard of patient care.

Keywords: Sternotomy, segmentation, active contour, convolutional neural network, solid modelling, patient-appropriate medicine

1.0 Introduction to computer vision

Computer vision is an engineering design to match the biological processes of human vision and its intricate neural network to read 2D and 3D images in real time. It is analogous to only part of the active process of human vision, and it carries out image processing based on mathematical operations. The enthusiasm of computer scientists has paced well to advance the computer vison techniques matching industrial development and rapidly growing medical technologies to extract essential information out of 2D and 3D digital medical images assisting medical diagnosis and therapy.

Computer vision includes image processing, image analysis, and pattern recognition techniques(Klette, 2014). The techniques of image analysis are image processing numerical tools for recognizing and differentiating regions by calculating grey-level values, size, and shape, textures, the architectural relationship between constituent tissue properties to produce quantitative data for subsequent decision making. In medical image analysis, depending on clinical objectives, one or all these steps are engaged

sequentially_ 1. image acquisition with a suitable modality; 2. image pre-processing for enhancement and inclusion of elements perceptible to human vision; 3. segmentation to identify region-of-interest; 4. feature extraction to measure grey-level values, intensity gradient, textures, geometrical properties; 5. Segmentation of one or multiple regions of an image that contains a specific object of interest; 6. higher-level processes such as image recognition and registration to classify and label various regions by sorting them into categories; and finally 7. decision making for medical diagnostics and therapeutic applications(Davies & Holloway, 2005). The main concept behind medical image processing is to enhance the visual quality of digital images to erase defects such as geometric distortions, blurring, and noise, to acquire meaningful information.

Image processing is broadly categorized into three areas(Prats-Montalbán et al., 2011), a) image compression to excise redundancies in input or original image; b) pre-processing it by reducing noise, pixel calibration, and standardization, enhancing edges of each region to produce a 'new' image by manipulating image space; and c) image analysis to provide numerical and graphical information to quantify output image characteristics and properties. In computer vision, segregating regions-of-interest (segmentation) is a mainline activity therefore there are a plethora of image segmentation methods that have evolved within image processing techniques in the last fifty years. Since the late eighties, the potential role of computer vision in medical image quantification has grown enormously and is still growing to improve the segmentation methods to increase its clinical application. Much of it has been applied successfully, particularly in the field of medical nuclear physics to devise treatment strategies and allied medical radiology imaging technologies to acquire high-resolution 2D and 3D imaging studies and temporal four-dimensional (4D) patient-based follow-up during treatment. However, there is still a lack of evidence and a significant gap between experimental and routine surgical treatment as a standard of care. There is abundant experimental literature available for review and serious consideration. It can only be achieved in a collaboration led by a surgical team and clinical biomechanical engineers as hospital-based members rather than standalone professionals in an engineering laboratory.

2.0 Study objectives

The present review is motivated by growing interest of the physicians in navigation surgery and machine learning for in-depth diagnosis and therapeutics. Considering limited mathematical adroitness of the physicians the study is configured to meet their primary needs to know and understand the 'nuts and bolts' of computer vision for emerging image-based practical applications. Therefore, this narrative review is not encumbered with mathematical algorithms, albeit important. The enormity of the subject is such that the review includes only the most relevant basic computer science concepts followed by important image processing, traditional image segmentation methods in practice since the beginning of 1980s and more recent techniques. Considering the pros and cons of fast evolving computer vision science, at the end there is preamble to a newly conceptualized physiology-based biomimetic model called hyalite sol-gel Amoeba (HSG-Amoeba) model, devised to acquiesce tissue densities for surgical application.

The computer vision science directed at processing and segmentation of medical images is a multidisciplinary undertaking, therefore the concept of clinical biomechanical engineering team within the hospital surgical department is discussed at the onset. The role of a clinical biomechanical engineer is to carryout higher numerical analysis and assist in selecting the most suitable implant design based on [1]*patient-specific* and *patient-appropriate* (Gandhi, 2019) parameters at the time surgical planning and medical treatments.

[1] The word *'specific'* refers to a precise anatomical location or an entire segment of a region or organ system, e.g., head of the humerus of the skeletal system, and *'appropriate'* means that something is suitable and favourable only for a single patient under special circumstances, including patient physique and multiple influential co-morbidities. A technique can only be called *'patient-specific'* if the collected set of data to which it is applied belongs to a specific anatomical section/segment of a bone as in the case of skeleton or a specific element of an organ for

As an example, the properties of digital image of the thorax are discussed as it is a routine and the most frequent investigation undertaken using all imaging modalities. It is a complex multilayered medical image consisting of all kinds of solids and fluids projecting radiolucent and radio-opaque tissues for computer vision applications.

Currently, there is an enormous amount of unaligned and discursive computer vision learning material, but much of it is outside the scope of the present study, therefore it is carefully handled to a limited number of computer vison techniques including only high yield segmentation methods that may be applied successfully to real-life 3D anatomy.

3.0 Role of a clinical biomechanical engineer

The science of computer vision is a challenging subject given that there are hundreds of techniques and complex mathematical algorithms. Currently, it is outside the avenue and beyond aptitude of most medical practitioners for computer vision techniques to be a standard of care without a clinical biomechanical engineer becoming established as a team member within the department of surgery to play an active role in improving patient-based outcomes. Now that high-quality 3D models can be rapidly printed with additive manufacturing techniques in a variety of biomaterials, this new technology and science of computer vision can be employed to advance surgical techniques from its current position in the engineering and biomedical engineering laboratory into clinical practice and operative surgery. The collaboration of engineers and end-users working together on a leveled platform will generate immediate stimulus and feedback. It will accelerate the operation of computer-aided technologies with the input of the interested clinicians in parallel with clinical biomechanical engineers. Thereby the process of development and application of 2D and 3D image-based technologies can become directly feasible routine for *patient-specific* biomechanical assessments as an integral part of pre-operative surgical planning to meet the requirements of *patient-appropriate medicine*(Gandhi, 2019).

The current attempt is to initiate such pre-operative surgical planning objectives that will introduce the in-house role of a clinical biomechanical engineer to undertake a surgeon-induced surgically relevant mathematical analysis based on the *patient-appropriate* clinical status of individual patients. It is realized and expected that introduction of such pre-operative planning will incur extra costs in terms of human resources, equipment, person-hour, and computational time. Overall surgical techniques and process of patient care will hardly be tempered with the introduction of clinical biomechanical engineers to the department of surgery to make engineering recommendations based on mathematical analysis of 3D patient-specific model and finite element analysis (FEA) to advise an implant design from currently available implant systems. With the presence of engineering expertise wherever necessary *patient-specific* and *patient-appropriate* implants can be readily manufactured by using additive manufacturing technology.

4.0 Methodology

Primarily, the study literature was searched on Google Scholar, Google Chrome by using keywords such as computer vision, image processing, image segmentation, image registration, medical image processing; computer vision textbooks; Wikipedia and relevant references cited in the reviewed articles. Secondarily, searched for original studies describing various techniques and methods during computer vision evolution between 1980s to 2021. The illustrations were drawn in MS Paint and artwork by hand.

5.0 Medical imaging technology

In the field of medical imaging, the x-ray recording machine is equivalent to the light-sensitive camera, which has been in use since its discovery in December 1895 by Roentgen. The x-ray beam when projected

biomechanical evaluation and higher numerical analysis for clinical application rather than to the entire organ system of that individual; without the inclusion of any kind of extraneous experimental or population based statistical values to constitute *'patient-appropriate medicine'*.

through 3D biological structures of varying radio-opaque and radiolucent densities, it is recorded as a 2D image and, when recorded with computed tomography (CT) in the form of numerous 2D slices, the set of data can be automatically organized into 3D reconstruction for storage and display. The medical image modality of nuclear magnetic resonance imaging (MRI) is based on the concentration and distribution of hydrogen ions in the tissues affected by the magnetic field to produce 2D data, which can also be transformed into 3D images. The radiographic digital image data is stored as a picture archiving and communication system (PACS). The CT and the magnetic resonance tomography set of data are stored for future display as digital imaging and communication in medicine (DICOM). Alternative to these modalities, a recently introduced electronic operating system (EOS) based radiographic recording machine can register two orthogonal views simultaneously and can automatically generate reasonable quality of 3D images for limited regional anatomy of interest. Despite all these innovative and much-preferred imaging modalities, the modality of plain radiography has been eagerly exploited by obtaining two or more projection views of the regional anatomy and synthesized by applying image registration and mathematical algorithms to create 3D embodiments for further analysis and processing.

5.1 Digital image

The digital image is a binary numerical representation of a 2D object or a scene. Technically it is described as a bitmap image, also called a raster image (Fig. 1). A raster image is an electronic file that consists of discrete picture elements, which are the smallest areas within a scene called pixels.

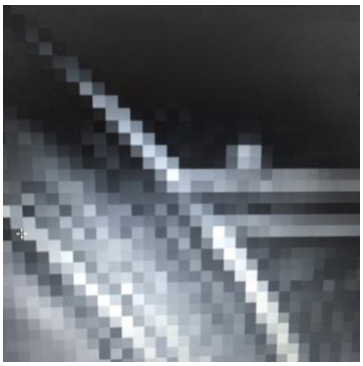


Figure 1 – Raster image showing pixels of varying grey scale intensities.

A pixel(Lyon, 2006) is the smallest luminescent unit or a cell in the raster form of a 2D digital image. The word 'pixel' is a synthetic term that originated in the early sixties, formed from 'picture' (pix) and 'element' (el). In the case of a 3D volumetric image, the element is called voxel (**vo**lume + pi**x** + **el**ement). Unless a normal digital image is magnified (zoom-in) to more than a hundred percent the pixels are too small to be seen with normal human vision. They are graphic representations of thousands and millions of photosensitive electronic tiny sensors on the full frame of a digital sensory plate of an image recording device. A one-megapixel device has one million pixels. The size of each square pixel is related to the viewing area, aspect ratio, and pixels per square inch of an image. It is approximately 0.26 X 0.26mm, 96.15 pixels per square inch with 1280 X 1024 resolution. The degree of image resolution (sharpness) is defined by the number of pixels in the grid of the digital sensory plate of the recording device.

In the computer system, it is in the form of binary (bi) digits (ts) (0, 1 or black and white) called bits, and the matrix formed by the pixels is a series of bits. A set of 4 bits arranged in a pixel has 2^4 = 16 different values/levels of grey values or brightness levels. One bit pixel means either black or white colour. The medical images have a minimum of 8 bits or one byte (bit + eight) in a pixel to provide 2^8 = 256 grey values for clarity and contrast, suitable for human visualization. The quality of the image needed for analysis and

processing is because of the black, white, and grey brightness or intensity gradient (contrast) that constitutes the greyscale of an image. The distribution of the brightness levels in an image can be shown graphically in the form of a histogram called greyscale intensity histogram, which is a kind of map of pixel values in each image.

An optimal image is expected to have greyscale values ranging between 0-255, where 0 depicts black and 255^{th} represents maximum white brightness. A highly saturated pixel value means a complete loss of information. There should be an even distribution of grey shades representing the best possible tissue material density, to provide good contrast between different tissues and organ boundaries to assist easy recognition, analysis, and least number of steps for segmentation and further processing.

6.0 Image registration, segmentation, and other operations

In an acquired image light reflected off the object is represented as a bright region often referred to as foreground and surrounding area as background. In a classic black and white image bright foreground as a region-of-interest (ROI) can be isolated by manually drawing a definitive line around it as its edges. In other words, there is a point of separation between the bright and the dark intensities where the ROI ends forming a threshold between two zones, which can be classified and labeled as foreground and background. These simple activities of edge location and the defining threshold between two different intensities divide the image into two or more distinct regions. In simple terms, this intentional process of partitioning the image into segments for further study by an observer is referred to as segmentation. These peculiar features can also be used to match similar images of different sizes and in the distant future to express the change in the image. This technique to match or register a pre-existing image to a new image is called registration. Semantic segmentation is the process of assigning a label to every pixel in the image is in stark contrast to classification, where a single label is assigned to same class as a single entity. On the other hand, instance segmentation treats multiple objects of the same class as distinct individual objects or instances.

These fundamental ideas form the basis for hundreds of registration and segmentation algorithms developed in the last fifty years during the evolution of computer vision science. The simplest way to classify segmentation and registration techniques is to recognize edges and define thresholds based on image intensities to divide an image into two or more regions somehow interactively or automatically. A compound image has a varying degree of grey levels distributed all over and a collection of intensities to give the image a kind of textured appearance, which too can be used for image segmentation and registration. Fortunately, almost all medical images presented for examination are in varying tones of black and white intensities and textures (grey-level scale) for diagnosis and treatment purposes. Unlike light reflected photography, the radiographic medical image formation is an X-ray transmission technology and represented grey-level scale intensities are because of tissue densities. It gets rather more intricate in an image like that of the chest where there are several overlapping structures forming layers of differing densities.

Thus, in computer vision, the two main activities which form the basis of greater understanding and clinical applications of medical digital images for translation of an average prototype image (example, source, reference, atlas) to individual patient image (real, target, new reference, end object, objective) are registration and segmentation. It is most important to know that both registration and segmentation methods are driven by the application of mathematical algorithms, which are constructed on well-established physics-based algebraic principles and geometric theorems.

As the main goal of registration is to achieve the best possible translation between images by an energy minimization algebraic solver (algorithm) automatically there is an additional computational cost to it. This involves a) transformation between images, b) objective similarity measures, and c) optimization(Sotiras et al., 2013; Zitová & Flusser, 2003). There are numerous parametric and geometric methods described in the

literature to carry out the registration of 2D and 3D medical images(Salvi et al., 2007; Sotiras et al., 2013; Tam et al., 2013; van Kaick et al., 2011).

6.1 Segmentation

Within the larger picture of the image there are numerous geometrical primitives, such as one-dimensional points and lines, two-dimensional squares, and circles, and when these are reconstructed into volumetric shapes form functional objects or regions of human anatomy. Each object or region may or may not have clear dividing edges and boundaries. However, in cellular structures arrangement of cells in a defined way, and more importantly, tissue density, air, and water content confer upon each tissue a distinctive physical property. It is due to these characteristics each tissue is discretely designated grey-level values and rendered in the form of image brightness or intensity in analogue and digital images.

Before any kind of treatment is implemented, it is only wise to analyze the whole image and specifically the ROI thoroughly by somehow segregating it into smaller units to increase the focus of study, and subject it to further analysis. This step of isolating the ROI is called segmentation(Del Toro et al., 2014). With the growth of computer vision and artificial intelligence, the interest has steered towards computer-based medical image analysis and processing. At the same time, there had been the development of numerous mathematical, heuristic, and metaheuristic algorithms with increasing accuracy to improve quality and cost-effectiveness in terms of computational cost and automation to save time.

As the medical images are collections of a variety of tissues forming specialized areas, each area may require attention at various times during the evolution of a disease. To increase focus, an image can be analyzed by categorizing the whole image into specific organs, subdividing to label them into various parts. And, for greater understanding, segment diseased tissues from the healthy or even decompose desired normal anatomy for finite element analysis, and in the case of skeletal system with the bone-implant interface. Recently, with the development of robotic vision and autonomous vehicles on the horizon the image processing to assess the surroundings deeply has moved to each pixel and sub-pixel level analysis as part of deep learning programme (Lecun et al., 2015a; Shotton et al., 2009).

6.2 Image transformation operations

Image processing for segmentation of ROI is performed at the level of each set of pixels. The medical images with an assortment of tissues within the same frame can be sorted by similarities and/or dissimilarities based on grey-level values of their pixels as regions, aggregate, and clusters or by some other kind of association such as colour, a mixture of tissues, air and water within the region that is of interest to physicians to perform specific procedures. To increase focus, three basic operations are generally required during one or the other type of segmentation technique to transform an input digital image into an output image or other kind of representation(Castleman, 1996). These operations are performed following the coordinates in the image space. A point output value at a specific coordinate is represented exactly to the same coordinate in the input value of the original image.

As the standard medical digital images have square pixels or cubic voxels, the images are sampled by applying a square grid over an image. The included neighbouring pixels can be arranged in a grid of either four connected pixels, above, below, right, and left of the selected pixel forming a grid of five pixels. Or the output grid can be formed with eight connected pixels all around the centrally placed selected pixel when a square sampling grid is applied to the input image (Fig. 2).

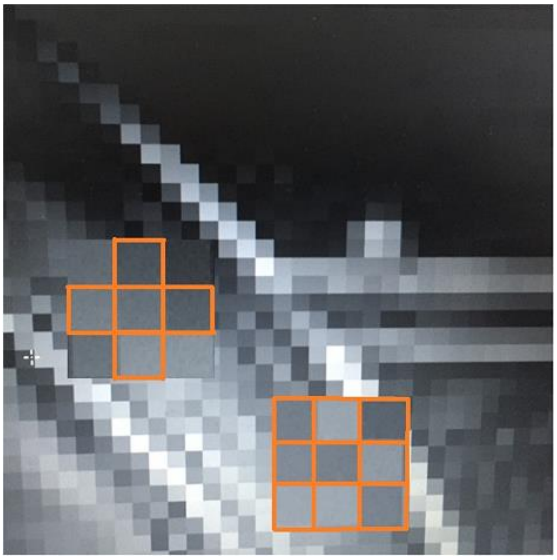


Figure 2 – Raster image showing 1 + 4 pixel grid and 1 + 8 pixel grid

6.3 Image intensity histogram

Image intensity histogram is a statistical representation of the distribution of grey levels in the spatial domain of an original image (Fig. 3). The number of pixels and how their grey-level values are distributed in a digital image are generally displayed as either a bar chart or a graph on Cartesian x-y co-ordinates(Dzung et al., 1998; Pham et al., 2000). A standard histogram displays the distribution of pixels and their numerical intensities ranging from 0 to 255. For a red, green, and blue colour image a 3D histogram

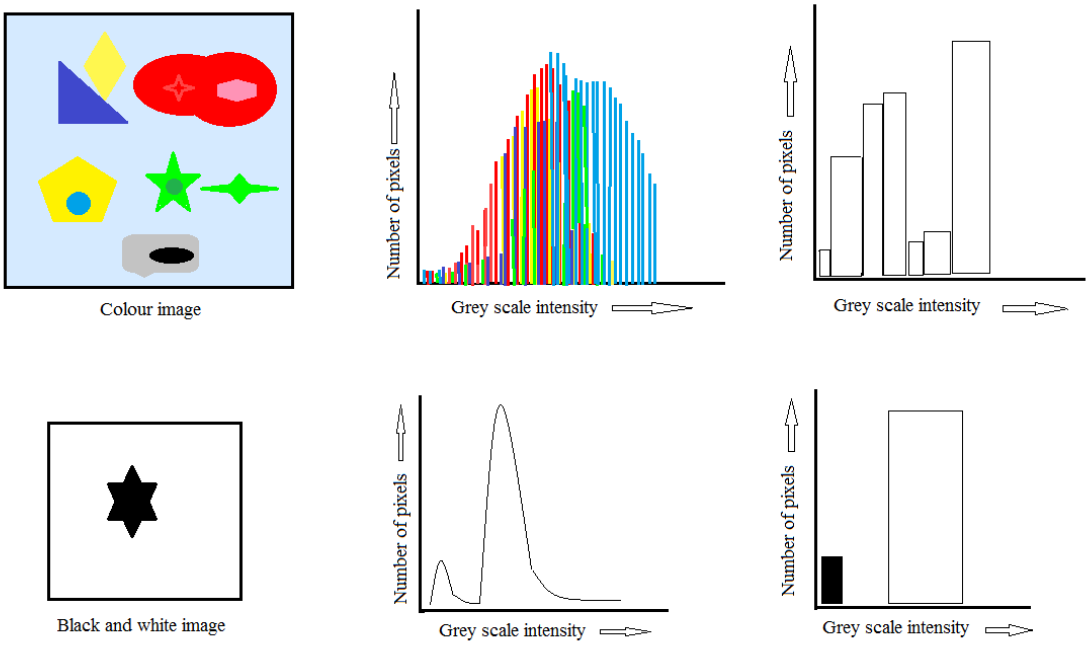


Figure 3 – Intensity histograms and graphic representations of a coloured and a black and white image.

is generated. An image histogram gets importance from its application in numerous techniques for image pre-processing to enhance quality to prepare it for various segmentation methods.

A digital image histogram is constructed by scanning the image to count the pixels and record the intensities at each pixel coordinate. Binary image (black and white) histogram will have bimodal representation, a low peak for black and a high peak for white high-intensity segment of the image. When there are numerous grey tones, the histogram shows the multi-modal distribution of the pixels. The main advantage of generating a histogram for a digital image is that an operator can alter image quality by manipulating the histogram, which can help to enhance the sharpness of the edges of the objects and make the segmentation process easier. As most of this is done automatically it becomes particularly important in the case of medical images to preserve the tissue densities and normal topography of the structures in the image during segmentation for higher quantitative analysis. Therefore, the idea of histogram equalization, where during pre-processing the pixels are distributed evenly over the whole intensity range to transform the output image, is not a great idea in medical image analysis.

6.4 Image noise

Noise signal in physics is a non-periodic undesirable disturbance produced by irregular vibrations of random non-harmonic frequencies of varying amplitudes arising out of scientific instruments and surrounding ambient environment(Mandar et al., 2015). It is the noise from various sources that contaminate and degrades the quality of medical images and makes the segmentation task difficult. In the case of medical images, the noise is random as opposed to from defined sources causing distortions such as shading and lack of focus.

Although modern-day technology has reduced noise levels associated with digital cameras and medical imaging modalities to almost negligible levels, the noise produced by photons is hard to eliminate. The problem of photon noise is because of statistical variation of photons which strike the sensitive pixel grid like a drum. It is the statistical variations associated with photon count that this finite amount of noise would be impossible to eradicate. In addition to the photon noise, thermal noise which also degrades the quality of medical images is generated from thermal energy produced when electrons are released from CCD material. The other well-recognized troublesome sources of noise worth mentioning are transistors, capacitors, and amplifiers making up the image recorders, and surrounding electromagnetic environment in all hospital workspaces.

7.0 Continuum Mechanics

Within engineering mechanics and mathematics, continuum mechanics study deals with the deformation of solids and fluids following the laws of physics("Continuum Mechanics for Engineers, Third Edition," 2009). The fundamental assumption is that almost all materials at the macroscopic level are homogeneous, isotropic, have continuous structure and when described are independent of a coordinate system. However, traditionally the motion of infinitesimal points in solids and fluids is described in terms of Lagrangian and Eulerian equations. Mathematically, both formulations when applied to appropriate models result in similar outcomes, but the Eulerian technique is practiced more often.

7.1 Application of continuum mechanics to deformable models

The concept of deformable models has been directly inherited from geometrical designs through computer graphics. In computational modeling, it becomes necessary to depict handcrafted graphs as numerical data to generate mathematical algorithms to make the virtual motifs physically active. Deformable models are generally represented with curves and surfaces. In the past, curves and surfaces had been numerically

defined by [2]splines and patches when making engineering drawings. Over time these have been refined to develop 2D and 3D computer models for application in industry and medical imaging. The curves and surfaces are represented by a set of control points and movement of these control points iteratively move them to new positions, adding and removing, and altering their weights (parameters) mathematically are employed as active curves and surfaces(Gibson & Mirtich, 1997). The reason physics-based deformable models have been developed further is that they can be handled interactively and automated with reasonable expectations of an operator, which would be otherwise impractical to implement with geometrical techniques alone.

Moreover, a computer model can be represented as a system of masses and springs in various configurations and deformed according to a mathematical algorithm. The linear springs follow well-established Hooke's law and masses as control points can be moved and tracked by choosing appropriate weights to animate the devised system. Human tissues are viscoelastic and anisotropic, which makes them nonlinear springs. But for image segmentation and further numerical analysis with tissues as homogeneous and isotropic adequacy of linear dynamic motion of control points as representative masses is considered acceptable and run-on Newton's second law. Macroscopically continuum models of real objects are solids, and the computational models are discrete points and lines simulating the real objects. Therefore, the active motion of these elements as dynamic computer models when analyzed can only provide approximate simulation on a small discrete time scale with the help of iteratively run algorithms. They need to be parameterized to define initialization, driving vectors, velocity, and position of termination(Gibson & Mirtich, 1997). To have a near real approximation the model design must be given parameters based on prior understanding of the original object and its material properties must be taken into consideration. For stability the continuum mechanical model acted on by internal and external forces must reach a state of equilibrium by minimization of its potential energy. In the case of a deformable model, the work done is the sum of applied load on the control points, the force distributed on the surface in between the points, and gravity acting on the whole body of the system. The potential energy only reaches a minimum value when the material of the whole deformed system experiences zero displacements, becoming smooth again.

7.2 Application of continuum mechanics to finite element methods

In the case of continuous solids, the process of deformation causing displacement of material, its equilibrium state is solved by a continuous differential equation. Whereas in the case of the mass-spring model with discrete control points and lines forming a mesh of finite elements the state of equilibrium is at each node of the mesh. This leads to the concept of the finite element method, which divides/discretizes a solid object into a set of discrete finite elements, where a state of equilibrium upon deformation can be solved by the partial differential equation for each element with constraints applied as boundary conditions for given material properties(Bathe, 2006; Zienkiewicz et al., 2013). For finite element analysis the elements are variously shaped, and recently in the case of a 3D image in-situ voxels are used to represent the finite elements instead of embedding a mesh of elements in an object or simply a wireframe superficially to represent deformation and calculate the influence of internal and external forces.

8.0 Concept of energy minimization

[2] Footnote: A **spline curve** is a piecewise polynomial formula to represent the shape of a mechanical spline. A hull (covering surface or frame of a body) is a deformed 3D structure bounded by several 2D squares or rectangular fasciae, called **patches**, and several of these patches put together in continuity form **hyper-patches,** which are constructed with parametric curves known as **Bezier curves** and the complete fascia forms the **Bezier surface**. A collinear set of control points joined together form a linear curve and applying varying number of weights or application of force at each point, a desired non-linear curve can be formed. When the first and the last points of it are kept fixed then in-between control points embedded in the curve or surface can be manipulated to form a smooth non-linear curve is referred to as non-uniform rational basis-splines (**NURB-Spline**), which is an intuitively driven technique as part of computer-aided design engineering.

Energy is an agency that actively sets an entity into motion, drives it ahead, and finally terminates its motion. In computer vision, all these steps are applied interactively to user-run deformable active curve, surface, and even three-dimensional volume models for processing a variety of images. In this respect, the energy source can be internal to motivate the model and external originating from pixel intensities of the image. In engineering physics, the formulation and solution are solved for a system such that it achieves minimal energy to reach a stable state at the point of termination of the desired motion. In user-run experimental physical systems such as virtual reality, the energy is created by writing suitable algorithms based on principles of continuum mechanics to determine the motion of an active model with specified compulsory constraints to initiate and inhibit the activity. The energy parameters are frequently defined to match the prior knowledge of the goals to fit objectives of image registration and segmentation. Image energy is pixel intensities as the formulated data in Cartesian coordinates to represent the difference in intensities in x and y direction according to intensity gradient. It is the magnitude and directional change in intensity of grey-level or colour in a digital image. This image property forms the basis of image processing to delineate edges to identify and classify various regions within panoramic view of multi-object multilayer radiographic images like the chest.

In mathematical formulation conservation of energy may not be entirely viewed like in physics. When developing a mathematical algorithm internal (input) and external (output) energies or forces are approximated by providing modifiable quantities as parameters and constraints such that the process can be terminated as desired without loss or waste. In terms of geometrical optimization, it is assumed that in the given location it has minimum energy to attain a stable state in that spatial domain.

9.0 Quantum mechanics in computer modeling

Quantum mechanics is the behaviour of atoms and elementary particles on the basis that whenever radiant energy is transferred, the transfer occurs in pulsations or stages rather than continuously, and the amount of energy transferred during each stage is of a definite quantity. In computer vision processes, except for techniques like convolutional neural network (CNN), the available computational models are based on quantum mechanics, a force field, which is used to calculate the potential energy of a system of atoms or items (say pixels of an image) for dynamic simulation to guess the correct geometry. An iterative algorithm is applied, where the first force acting between atoms or items is calculated, and if this force is less than the supposed force then the calculation is terminated. Otherwise, more atoms are computed until the predicted force is reduced to less than the chosen threshold. Quantum mechanics can be used to calculate energy and the gradient of the potential energy surface for the position of the atoms or items in the given domain. The optimization (method to select the best element or combination of elements from the available solutions) algorithm can be applied theoretically to minimize the forces in methods like gradient descent as in image intensity gradient and potential energy of surface curvature.

In the same way, the energy minimization concept has been applied to numerous image analyses and processes, particularly ones that are classified under active contour models, in the form of mathematical algorithms. In these cases, the virtual geometric model in the form of a curve, a surface, or a mesh during motion has higher internal energy when getting attracted towards the desired site under the influence of external energy derived from the image characteristics. When the energies approximate each other, the designed algorithm reaches the minimum energy the motion of the curve model is terminated, if the algorithm is idealized to the task.

10.0 Heuristic and Metaheuristic

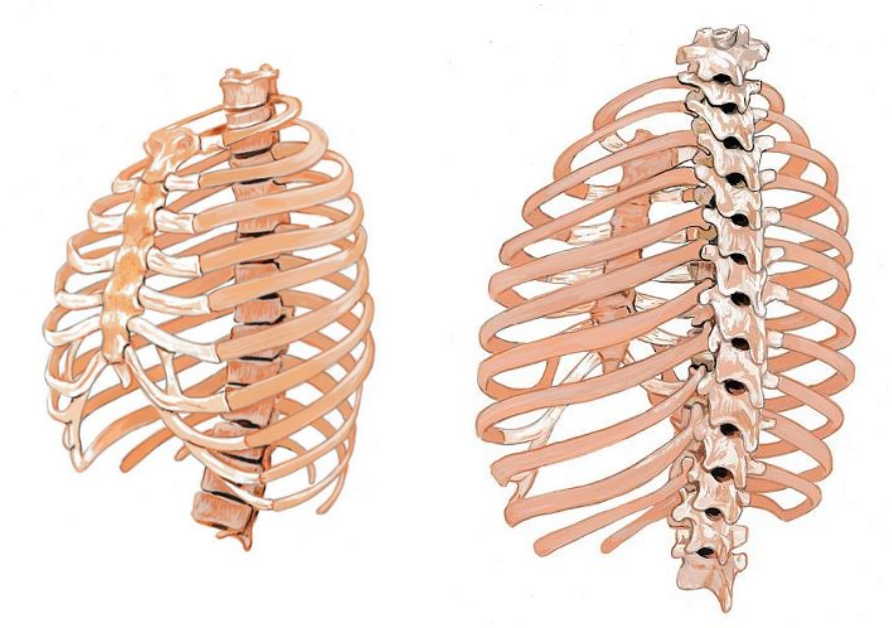
The ability of human beings to know or learn by immediate perception and understanding of truths and facts without consideration to logic and reason, particularly in the case of a new instance is often referred to as intuition. It is this intuitive power that encourages human beings and animals to undertake a personal

investigation of observed events to find the best possible solution. This kind of lateral thinking and developing collateral ideas can help understand most of the computer vision modeling techniques and how the mathematical formulations may be conceived and updated. This process of immediate acquisition of potential knowledge and understanding of any problem to find a method to begin and construct the most reasonable solution without certainty forms the concept of heuristics (Gk Heuriskein - to find). A metaheuristic is a kind of heuristic beyond intuitive learning and observations to draw inferences. It is derived from biological phenomena such as best gene selection, which can be abstract and demanding for the optimization of engineering problems.

In computer vision, heuristics do not guarantee optimal solutions, in fact, do not guarantee solutions at all. Heuristics can be applied to real or practical problems, whereas mathematical algorithms apply to abstract, theoretical problems ending in any number of satisfactory solutions(Romanycia & Pelletier, 1985). The heuristic approach taken after an exhaustive algorithm, in contrast, is supposed to deliver a guaranteed solution. The goal of modern heuristics in computer vision is to overcome the disadvantages of premature convergence to nearest local optima(Kokash, 2005). Metaheuristics often perform better through a random and thorough search and is a favoured method for generating solutions to many engineering problems(Yang, 2010) The list of metaheuristics is long and diverse(Luke, 2013). Heuristic and metaheuristic methods can handle a wide range of problems and explore large solution space but may still fail to provide optimal solutions.

11.0 Digital imaging of the thorax (chest)

The chest region is a collection of numerous soft tissue structures containing air in the respiratory tract and oesophagus; large blood vessels and heart filled with blood, and other small solid structures usually not appreciated on a plain radiograph unless calcified. The thoracic skeleton consists of twelve pairs of bilaterally arranged ribs, articulating with twelve thoracic vertebrae at the back and on the front with intervening costal cartilages are tied in the mid-line by a tie-board called sternum (Fig. 4). There are significant variations with change in posture, quiet and laboured breathing, differences among males and females, and age-related. All these variations occurring during respiration influence the relative position of



A. **B.**

Figure 4 – **A.** anterior to posterior and **B.** posterior to anterior oblique views of the thoracic skeleton.

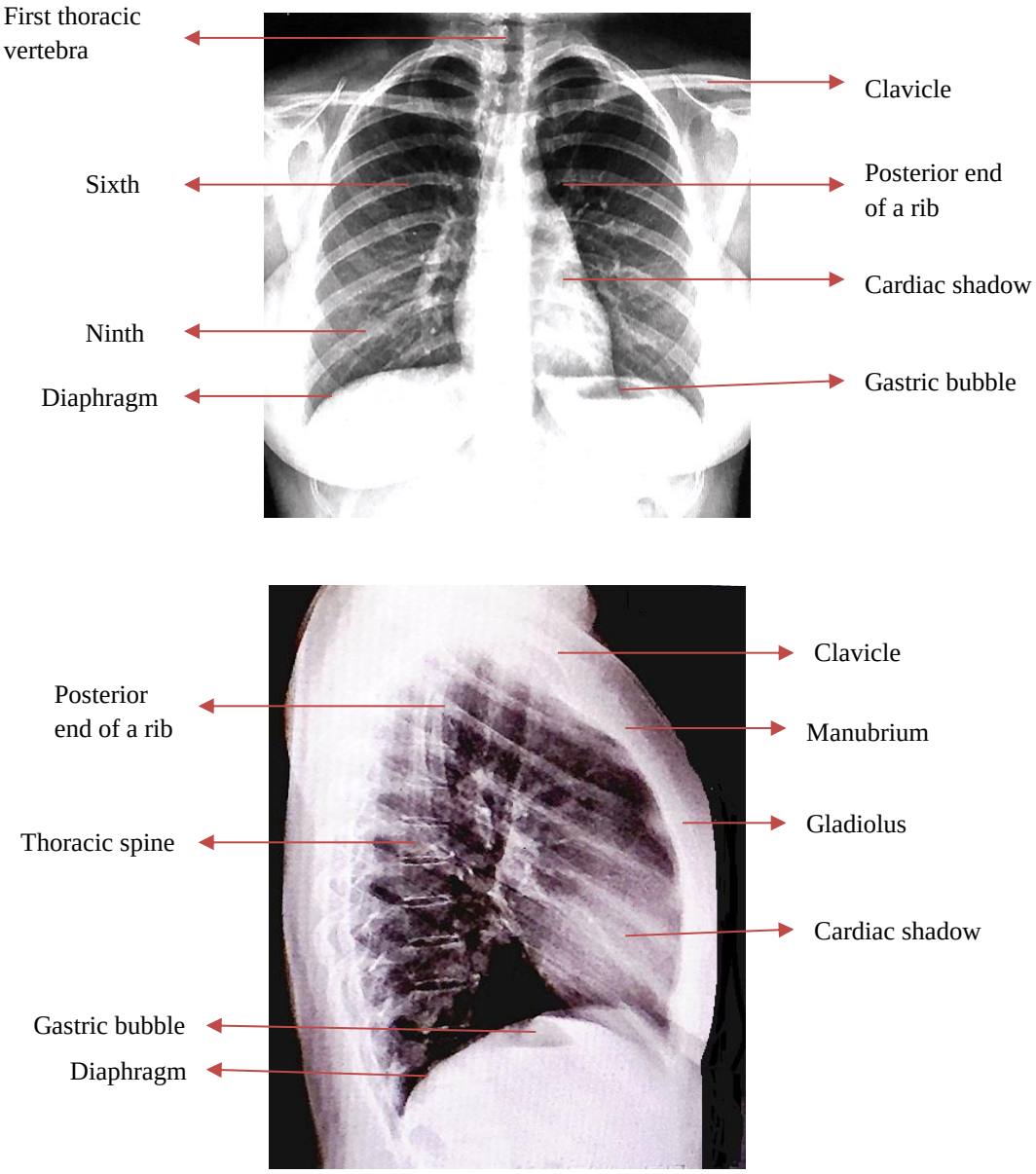


Figure 5 – A posterior-anterior (**top**) and lateral (**bottom**) radiographic views of the chest. The images modified to highlight skeletal structures of the thorax.

the ribs and vertebral column during imaging of the thorax when it comes to marking salient control points on the dynamic skeletal elements for morphometry(Kenyon et al., 1991), and segmentation during the modeling process.

The multilayered complex chest anatomy of varying densities and constantly moving structures greatly influence the quality of the radiographic image. Unlike the film based x-ray plates, the digitally acquired images have much greater resolution and heterogeneous grey-level values of bones and soft tissues. Higher image resolution, variable texture, and grey-level relative to factual bone density permit interactive and automatic segmentation of the thoracic skeleton on plain radiographs possible. On CT imaging

heterogeneous density of the ribs produce an intensity gradient range from 700 to 3000 Hounsfield Units(Candemir et al., 2016). However, due to the crowding effect of the ribs and overlapping soft tissues, the cumulative density of structures in the mediastinum and thoracic vertebral column in the same sagittal plane, the sternum on posterior-anterior view is almost completely concealed (Fig. 5).

Frequently, ribs below the 6^{th} and 7^{th} pair are poorly seen on a posterior-anterior view demanding an oblique view to visualize the whole radius of the lower ribs. A 15-20 degree right oblique posterior-anterior view shows a slightly warped eccentrically placed image of the sternum over the cardiac shadow, and its edge-on profile view on a lateral view may be used for radiographic morphometry and segmentation. Overlapping of structures in the image field tend to add ambiguity to grey-level values and intensity gradient when measured in Hounsfield units on plain radiographs. Whether the adjoining structures of higher radio-density affect the HU of lower density tissue remains unclear. There is no standardization of radiographic technique in routine clinical practice to match the posterior-anterior and lateral views of the chest x-rays for grey-level and intensity gradient-based registration of the images to reconstruct a 3D image. To exclusively acquire images of thoracic skeleton intrathoracic soft tissues can be subtracted with dual-energy subtraction x-ray but overlapping of ribs due to their oblique descent from posterior to anterior show false boundaries of the ribs, and poor location of anterior ends of the ribs due to lower bone density(Candemir et al., 2016) is further masked by soft tissue organs immediately under the diaphragm, which affect their true grey-level values.

The quality of thoracic skeletal imaging may also be adversely influenced by thick subcutaneous adipose coat in cases of high BMI, and occlusion by large breasts. Alternatively, biplane electronic optical system (EOS) based digital plain radiographic scanner takes posterior-anterior and lateral views at the same time provide a 3D reconstruction of the thoracic skeleton of the chest falls short because costal cartilages and sternum are not included routinely(Bertrand et al., 2008). 3T and higher Tesla MRI imaging of the thorax is essential for building a complete articulated thoracic skeleton to include fascio-cutaneous and adipose layers, breasts, tendons, and muscles of the trunk acting on the ribs and the sternum, costal cartilages and articular cartilage, and joint capsules and ligaments if an ideal thoracic model must be constructed for realistic FE analysis and surgical application. Unfortunately, the image intensities of MRI do not relate to the tissue densities. Currently, building an ideal 3D image of the complete articulated thoracic skeleton from 2D radiographs is an ambitious undertaking.

12.0 Modeling processes in computer vision

Virtual computer models are kind of analogies as digital (numerical) representations to include salient features essentially to analyze the problem in return for a reasonable solution close to reality. A simple and compact model of a 2D pattern of an object is called a template, and in medical image processing and analysis to accommodate variations in size and shape, multiple templates or atlases forming a training set is generated to make an average atlas for copying and editing a new patient-specific medical image by registration. To overcome the computational cost of checking against each atlas in the set one at a time the process can be automated to find the best fit(Sabuncu et al., 2009).

In computer modeling, a deformable model is a virtual geometric entity or object whose shape can be changed iteratively based on heuristic knowledge from mathematics, physics, and engineering mechanics. These geometric entities are generally represented as curves, surfaces, blobs, solids, or even bio-mimetic machines. These deformable motifs are motivated and powered by mathematical principles of partial differential equations. The 2D and 3D dynamic models are frequently designed to operate in time-varying dimensions (4D). Mathematically, these entities are representative drawings (maps) of the computer models in space, and under applied force change from existing shape to a new shape (deformation).

12.1 Control points

The thought of control points and point distribution model is like pencil and paper games of dots and boxes, joining dots to make subjective contours represented in dots of 2D objects by having *a priori* knowledge. The control points pierced into a template like old-fashioned postage-stamp perforations are employed to customize the size by moving points in hand-crafts and sartorial practice.

In computer modeling, control points or landmarks are a collection of points that define the shape and salient features of an entity in an image or a geometric model to act as handles for manipulating the image dimensions during registration and segmentation. A shape defined by a set of a finite number of points distributed within static Cartesian co-ordinate space is termed point distribution model. Each of these points defining the shape is given a vector comprising x and y coordinates. These shape vectors are statistically computed to a mean shape. Several same shape medical image atlases of varying dimensions are converted to a single mean atlas and the shape variability is calculated by performing statistical shape analysis.

12.2 Principal component analysis and Procrustes analysis

Every shape has only a few defining features but for greater accuracy sometimes is represented by hundreds of control points depending on the size and complexity. To reduce the computational cost and data the reduction method used to perform statistical shape analysis is called Principal Component Analysis (PCA). It is a linear algebraic formulation that involves transformation matrices, finding Eigenvectors and Eigenvalues (Eigen from German means unique) by selecting important shape variables to produce the expected outcome. The approach to compare the shapes and scale them up or down the generalized Procrustes analysis is undertaken. The Procrustes (pro – in favour of + crusty, means a bad-tempered person with bad manners; and in a Greek legend, character Procrustes attacked travelers, who either stretched or cut off their legs to make them fit the length of his bed)(*Gage Canadian Dictionary*, 1983) analysis provides a measurement of the distance between superimposed two or more shape atlases annotated by control points to segment a new reference image.

PCA is a dimensionality reduction process to select the most important variables and let go of the least important and calculate new independent variables which are representative of all the original variables (www.towardsdatascience.com). When performing PCA it is important to preserve salient features for accuracy to maintain proportions and achieve maximum possible compression of the data.

12.3 Atlas in computer modeling

An atlas can be a single illustration or collection of illustrations representing an image of a specific normal or pathological anatomy of a patient population. Within a given population there can be significant variations in the representative organ anatomy under study that would reflect in medical images as well. Initially, performing image analysis and segmentation requires the collection of parameters from hundreds of training images to include anatomical variations, to build an atlas of a prototype model to learn a specific anatomical region. However, to recruit and collect a large population-based data is not only expensive but also a major time-consuming laborious step. The image of a new patient can be segmented by employing only a single atlas model; however, the multiple atlas approach provides much higher accuracy(Aljabar et al., 2009).

12.4 Geometry and graphics in computer modeling

In computer graphics, 3D modeling is a mathematical process to represent the surfaces of an object. The created 3D model of ROI is a representation of a physical entity with identifiable edges and boundaries within an image containing multiple entities. Solid models are employed for higher medical image analysis as they represent the volume of an object like a solid cube rather than a shell model akin to a box with corners, edges, and surfaces only. In 3D space, the edges can be a collection of points or dots connected by geometric primitives such as triangles and squares to form polygons. Vertices or corners are points where

three or more edges meet in 3D space, and if each included polygon has a face, then the collection of connected polygons forms an element. The elements can be a collection of triangles, squares, hexagons, etc. Many numerical analyses are performed on meshes made up of a variety of polygons overlaid on the region-of-interest or the whole image. One such construction of polygon mesh for analysis is called the finite element method in engineering. The points at the vertices of the mesh acting as control points (nodes) may be used for manipulating the underlying enclosed image entity to modify its geometry.

12.5 Parametric modeling

A parametric model is created from already known facts as a *priori* knowledge of an entity or object. Parameters are set of quantities as one or more independent variables, and parametric equations are commonly used to express the coordinates of the points that make up a geometric entity. In computer modeling techniques, the freedom to change parameters help interactively modify specifications of a model by the user to update dimensions for customization. The knowledge of regional human anatomy of interest is the ultimate source of *a priori* knowledge when it comes to segregating ROI in computer vision applications. Manual segmentation performed by an expert to develop a training prototype atlas from a given radiology image furnishes ground truth to validate the results of computer-generated segmentation. In mechanical engineering for constructing a computer-aided design, the parametric term is used to describe the change in the shape of 2D or 3D shell or solid model geometry by altering the values of the dimensions. Numerous open accesses and expensive for sale mechanical CAD software such as Solid works®, Inventor®, CATIA®, Mat lab®, etc. are available for parametric computer modeling.

13.0 Classification of computer modeling techniques

Almost all available medical image recording modalities deliver grey-level digital images consisting of pixels or voxels. Each of these micro-size elements has 8 or 16 bits, displaying 256 (2^8) or 65,536 (2^{16}) grey tones. Under natural ambient light, the human visual system can easily discriminate approximately 200 grey-level shades or tones(Barten, 1992). Therefore, even the liquid crystal display units with 8-bit pixels are adequate for viewing most digital medical images. With this minimum degree of image resolution, it is practically possible for a trained anatomist, radiologist, surgeon, or pathologist to manually segment a region of interest with suitable computer equipment. There is a significant degree of variability of intensity gradient, grey-level value, image texture due to image heterogeneity, the spatial organization of the tissues in the ROI and spatial-temporal (time lag as the 4th dimension) when it comes to follow-up a disease evolution to plan treatment. For these reasons the computer vision community engaged in medical image processing has put enormous effort to develop numerous algebraic, geometric, heuristic, and meta-heuristic variety of segmentation algorithms to run the process interactively and automatically. To mitigate the deficiencies of the antecedent algorithms to fulfill a specific kind of need for analysis of medical images and, other computer vision images. However, each newer algorithm is delivered with new limitations of its own, which has led to the expansion of the classification system.

A classification system makes the process of decision-making tasks easier whether it is to better define patient-specific regional anatomy or description of disease pathology from a variety of tissues in an image for patient-appropriate management. In an image, the same type of normal tissue such as fibrous, glandular, muscles, bones, etc. can be expected to have similar intensity distribution giving it self-segregating texture unless there are pathological changes in the region-of-interest.

In clinical practice, it is not a simple matter to define normal anatomy in a large busy radiographic image like that of the thorax with layers of structures. In this regard, the task of selecting the best segmentation technique from among the many is even more engaging to locate a pathological lesion. The importance of image processing lies in labeling diseased tissues reflected as pixel/voxel intensity values in two- and three-dimensional modes of medical image for segmentation for further analysis. Therefore, varying pixel

intensity, scattered noise, and textures in the image domain play a significant role when it comes to classifying and developing segmentation techniques. A simple way to classify and label ROI in an image would be to apply grey-level scale and intensity histogram for thresholding, detection of edges, and texture recognition in each region. Unfortunately, there remain many imperfections particularly when it comes to studying the extremely variable and heterogeneous nature of medical images with the scope to develop newer techniques in the future.

Here, a simplified yet practical classification of computer-based modeling techniques presented, although limited to the present task it may be valuable in many other ways (Table 1).

Table 1- A limited structural and functional classification of computer-based models.

A. Grey-level scale and intensity-based techniques a. Thresholding, edge detection, fractal dimension, region-based, clustering etc.	**B. Rigid models** a. Image matching techniques b. Atlas-based templates	**C. Free form models** a. Free form deformable models b. Extended free form deformable models c. Free form templates –rubber masks, spring-loaded models
D. Deformable (intensity gradient and energy minimization) models a. Parametric models b. Non-parametric models	**E. Parametric models** a. Active contour models – Snakes, adaptable snake, topologically adaptive snake, gradient vector flow snake, inertial snake algorithm, snake growing algorithm, directional snake algorithm, balloons, intelligent balloons, 3D morphable model, 3D surface meshes, fire propagation algorithm etc. b. Statistical shape models - Active shape models, Active appearance models c. Others – Grow and cut techniques, edge flow, live-wire, fuzzy transformation and connectedness, Fourier transformation etc.	**F. Non-parametric models** a. Level set curve and surface evolution methods – Level set method, geometric method, geodesic method b. Atlas-based method, watershed method, graph-cut techniques
G. Metaheuristic segmentation and optimization algorithms a. Deep learning Artificial neural network – fully convolutional neural network, U-Net, V-Net, Generative adversarial Network, Projective adversarial Network.	**H. Bio-mimetic models** a. Deformable organisms – with or without artificial intelligence	**I. 3D solid modeling** a. Parametric and geometric models, multilayer parametric solid modeling

b. Genetic algorithm and numerous other algorithms		

14.0 Methods for computer modeling

A computer-aided model is a virtual representation of a known object and its variations with or without added constraints by the user help to define significant features of its shape to remain true to the anatomy of the structure. Segmentation of normal and pathological anatomy helps to have a greater understanding of disease evolution. In case there is a need for biomechanical evaluation of solid weight bearing anatomical structures then the choice of physical properties such as anisotropy, elastic modulus, and similarity of material distribution of the model should match in-vivo biological properties to allow the form and shape of the model to respond congruously under applied loads. Unlike solid engineering materials, the deformation of biological materials can be highly unpredictable depending on material distribution and anatomical shape. This is true as the shape and material distribution of diseased structure changes so does its response to applied loads. In other instances, during follow-up of soft tissue pathology, there are morphologic and morphometric changes in the anatomy and appearance of tissues in acquired radiology images. Therefore, to establish computer-aided modeling of medical images, in a busy medical practice, it becomes important to have the collaboration of a clinical biomechanical engineering team to undertake the process of registration and segmentation of temporal shape for higher analysis and timely therapeutic intervention.

To achieve such goals what follows is a selection of the most common computer modeling techniques that have been applied to medical image processing and analysis. It is not exhaustive in terms of the amount of literature currently available on the subject. To keep it simple for basic understanding and within the framework of the study objectives, none of the techniques are supported by mathematical algorithms.

14.1 Edge detection and edge detectors

An image is a collection of specific and non-specific segments representing various features interrelated to each other functionally. The term edge means a set of points that distinguishes between the adjacent class of regions and feature-based variable segments that can be organized by curved lines(Marr & Hildreth, 1980). As the main objective of segmentation is to segregate a ROI the process of edge detection is the first and foremost tool to split a digital image to discriminate various zones and extract features of importance. They can be usually differentiated by their relative intensities and textures, separated by a set of featureless pixels that act as discontinuity without occluding the edge features of adjoining segments or territories. Presuming that the structural tissue of the given anatomy is homogeneous and of the same class, then the image intensity gradient can be expected to be the same as well. In that case, the region-of-interest (ROI) act as foreground, and adjacent territories of the image with relatively different grey-level scale and intensity act as the background to help detect its edge. At the edge, there is an abrupt change in pixel intensity to a higher value. The objective is to seek and identify a very precise narrow band of pixels with predefined brightness (threshold) that sharply marks the limits of the region. Each adjacent region would have its limiting edge and in-between any two edges of well-defined regions lay the boundary contour.

Automatic edge detection using differential operators is often performed by applying two-dimensional convoluting kernels or filters to the image for segmentation as an image pre-processing step. The operators can be configured to search vertical, horizontal, or diagonal edges(Senthilkumaran & Rajesh, 2009).

The first-order linear operators are Sobel, Prewitt, and Roberts (Fig. 6). The non-linear second-order operators are Laplacian, Kirsch, and Wallis. The Sobel is a discrete differential operator that calculates image luminance by considering intensity gradient. The Sobel operator filter consists of two sets of 3 X 3 matrices, which are horizontal and vertical in x and y coordinates, plotted in the plane of the image (overlying the image) to align and approximate to each pixel of the image to estimate the intensity gradient.

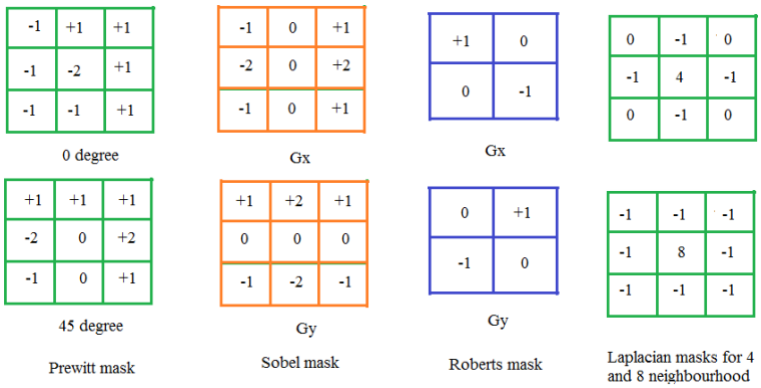


Figure 6 – First and second derivative operators – Prewitt, Sobel, Roberts and Laplacian masks.

The second derivative Laplacian operator is good at finding the fine details in an image and is applicable when the edge needs to be detected regardless of the difference in pixel grey-level values around it(Haddon, 1988). The Laplacian filter kernels are omnidirectional and extract edge information in all directions. If the sum of all convolved values is equal to zero, the filter kernel reveals all image areas having significant brightness, making it an isotropic operator. The Laplacian operator responds strongly to isolated pixels than edge line pixels making it very sensitive to noise and less effective in case the image is very noisy. Therefore, in the presence of noise, the Laplacian operator undergoes a low pass filtering step by applying a Gaussian filter as a preliminary step. This two-step process is named Laplacian of Gaussian operation(Mendoza, F. & Lu, 2015).

Another popular edge detection algorithm is the Canny edge detector(Canny, 1986). The algorithm begins with a Gaussian filter to smooth the image for noise reduction, followed by a computation of intensity gradient magnitude like the Sobel operator. In the next iteration, a step called non-maximum suppression is applied to remove pixels that are not part of the edge. Finally, the previously set upper and lower threshold is applied to select relevant edge pixels. The pixels below the lower threshold are discarded and one greater than the higher threshold marks the edge. The pixels in-between the two thresholds only one with the highest value closer to the upper threshold are considered for inclusion to strengthen the edge. The Sobel operator can detect edges reasonably well following edge orientations, while the Canny edge detector after removal of noise with enhanced signal to noise ratio can deliver a one-pixel wide edge in the output image.

In medical images, the edges are not always uniformly connected showing an objective boundary, which is a low intensity very narrow belt between edges of two adjacent regions. Medical images of pathological lesions are more like irregular coastlines rather than smooth coasts. Therefore, to state that the marked functional and anatomical boundary contour of the ROI is the differentiating threshold for surgical planning is never a trivial undertaking, particularly when a malignant tumour in the image is ROI for segmentation. Considering fewer effective results of the edge detection techniques, they have become important only as a pre-processing step to more recently developed more sophisticated techniques for improved segmentation.

14.2 Thresholding

In computer modeling, the process of separating two or more regions based on variable intensity and/or texture is called thresholding. The foreground as a ROI can be easily isolated by setting upper and lower limits (histogram graph valley as minima and peak as maxima) by suitably dividing the peaks of the image histogram at the depth of the valley, which divides the whole image into foreground and background bearing similar value pixels (global thresholding). Although grey-level thresholding appears simple but to select multiple correct thresholds for a heterogeneous medical image can be less than optimum for complete segmentation, more so if the image quality is marred by a low signal-to-noise ratio. In a multimodal image of numerous grey levels of low peaks, several intermediate peaks, and high peaks, generally, the upper threshold is used to initiate the marking of an edge, pixel by pixel.

In an ideal bimodal image, the threshold is selected at the bottom of the valley of the histogram to represent foreground and background(Prewitt & Mendelsohn, 1966). It is often difficult to select the threshold when the valley is flat and broad, the image is noisy, and peaks are of uneven heights. The remedy in such cases for selecting threshold the Otsu algorithm is preferred(Otsu, 1979). It is a non-parametric and unsupervised method, which selects an optimal threshold automatically but requires prior information about the ROI. It is based on the local minima property and the global property of the histogram. The selected threshold creates a binary image from grey-level pixel values. All pixels below and above the threshold value are respectively assigned zero and one value. Otsu method uses exhaustive search to evaluate the criteria for maximizing the variance between pixels above and below the threshold. In the case of medical images with multimodal threshold, the Otsu method takes a much longer time to select multilevel threshold(Sahoo et al., 1988). To improve upon the Otsu method, it is combined with other approaches. One such strategy is to combine with the K-means algorithm as both are based on criteria of minimizing within-class variance. The Otsu algorithm operates on global thresholding and K-means on local thresholding. Application of grey-level histogram is a preliminary step to the Otsu algorithm and applied in conjunction with K-means for enhanced segmentation(Vala & Baxi, 2013).

Very often the edges are enhanced by convolving the image gradient to produce a new edge. The convolution process involves implementing a filter/kernel to the original image domain in each direction. There are several such filters to emphasize the edges. However, a lack of standardization of image collection by different image modalities and inconsistent image intensity gradient can prevent precision, cause the failure of registration and segmentation algorithms to match time lag follow-up scenarios in the same patient, and during comparative clinical studies for higher analysis.

14.2.1 Fractal dimension-based edge detection and thresholding

The concept of fractals and fractal dimension was conceived by Mandelbrot in 1982(Benoit B. Mandelbrot, 1984) to characterize complex irregular objects in nature. The hierarchy of human body tissues arranged in complex fractals reflected in the heterogeneity of medical images with irregular structures is difficult to analyze with Euclidian geometry, however, it lends well to fractal mathematics for image segmentation(Vuduc & 634, 1997). The box-counting technique (counting overlying grid-squares) is used to measure the fractal dimension of medical images based on a finite set of pixel data(Alan I. Penn, 1996) and develop algorithms for medical image analysis. The equation for calculating fractal dimension (Fd) of 2D geometric shapes is expressed as Fd = Log(n) / Log(m), where n is the number of units and m is the magnification factor.

Fundamentally, it is a matter of grouping together elements of similar fractal dimensions to mark out the edges segregating ROI as a thresholding and edge detection tool. Recently, with the availability of FracLac© box-counting software application for assessing complex patterns, its role had been further exploited for

computational image processing to study the progression of malignant tissue histology(da Silva et al., 2021) and high resolution MRI images containing(Marusina et al., 2017) atypical focal lesions.

14.3 Active contour models

Currently, there are numerous classes and algorithms of deformable models utilized for image segmentation by defining the region and/or edge detection to delineate boundary contour. Edge detection segmentation techniques based on thresholding and local filters provide irregular and frequently incomplete segmentation, particularly if the image is noisy. The search into the development of dynamic deformable models started in the early 1980s to advance the state of computer vision from free-form deformable models and template-based shape-matching performed by correlation of intensities in the spatial domain and feature space. The free form deformable, and matching template atlas had limited dynamic variability to meet the demands of highly versatile dynamic changes occurring in the medical images within any given population and time-lag follow-up in the same patient. Hence, deformable active contour models evoked greater attention for medical image segmentation and higher analysis. The main thought behind dynamic deformable models is that they can be operated based on prior knowledge of the imaging anatomy and ROI, with or without predefined parameters to run and deform one-, two-, and three-dimensional models by constructing time-dependent suitable algorithms explicitly and implicitly to match the shape of a known object boundary contour. The model is called active because it can adopt and follow the given set of data with constraints, and the ability to react to laws of physics in response to image properties.

The active contour models are based on the principle of unification of physical laws and optimization theory(McInerney & Terzopoulos, 1996). An active contour model is usually represented by an arbitrary geometric curve-shaped body (a spline) that can alter its shape in space and over time, guided by the image properties and imposed constraints. The curve model acts and behaves like an elastic energy string to represent its internal parameters and each point on it moves in conjunction with each other when deformed by an applied force. The curve behaves within its constraints to conform to a new shape and is optimized according to image energy (intensity gradient) as the external parameters. At a point of energy equilibrium state, the curve model delineates the boundary contour of a ROI. The system achieves an equilibrium state when internal elastic properties of the curve model balance with external image intensity gradient as its external energy. Most of the active contour algorithm consists of three phases – a suitable initialization site of the curve in relation to boundary contour of the ROI, progressive deformation of the curve by the leading control points to energy minimization site and termination criteria to cease at the edge, converging to zero and phase out. The number of iterations required is set by segmenting operator according to size and shape of ROI. Fundamentally, a whole variety of these interactive models best operate based on prior knowledge of the anatomical space and topology, edge pixel intensity and image intensity gradient, local and global spatial information, to motivate the curve with user-dependent parameters.

For better understanding, it is important to learn that line is a one-dimensional geometric shape whose course can be defined by an equation on a graph representing statistical data. The term curve here refers to the active computer model and contour is a line that represents the edge or boundary of a structure, an object, or a region to segregate it from the rest of the image.

14.3.1 Snakes

The pioneering ‘Snakes’ algorithm introduced in 1988 is an energy minimizing spline guided by internal and external forces to detect lines, edges, and even subjective contours(Kass et al., 1988) to approach human visual perception to segment ROI. ‘Snakes’ has its roots in autonomous energy minimization models(Sperling, 1970) and shares the concept of deformable rubber mask templates(Widrow, 1973a, 1973b), to develop an interactive technique and guide the model to a local minimum as a point of energy

equilibrium state. The basic 'Snakes' model behavior depends on internal spline forces (elasticity and rigidity), which makes it act like a membrane and a thin plate, making it a controlled continuity spline(Terzopoulos, 1986). The spline is motivated by the external energy of image intensity gradient and applied constraint forces. Adjusting the weights of stretching (membrane behaviour) and bending (thin plate behaviour) of the spline controls its active behaviour to provide smoothness and regular outline to conform it to the boundary contour at the final moment of energy minimization.

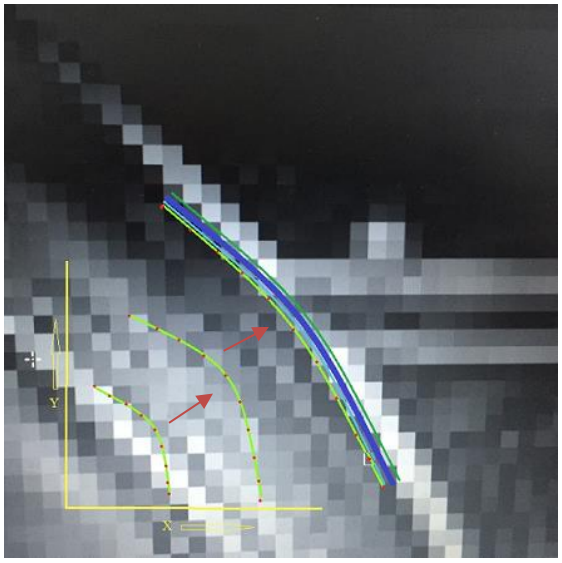


Figure 7 – Illustration showing propagation of an active curve (chartreus green with red control points) within a region-of interest towards its edge (indigo) within the boundary contour (dark blue) separating it from the edge (dark green) of an adjoining region. Arrows shows direction of progression of the curve.

The total energy of the parametric 'Snakes' spline curve is sum of internal energy at a specific point in space of arc length plus image energy at a specific point of the arc length in space plus energy constraint at a specific point in space and its distance from the specified edge (Fig. 7). The curve tends to get attracted to the brighter intensity of the closest lines and edges depending on its constraints and gets pulled to terminate only at points of energy minimization. The minimum local potential energy of the curve is equivalent to the maximum local intensity gradient of the image at the edge of desired boundary contour. The 'Snakes' algorithm iteratively by discretization of Euler-Lagrangian equation into sub-equations intends to reach equilibrium state so that curve energy is equal to image energy as its terminating criterion. In these time-dependent discretization iterations if time duration is longer than the matched traveled distance, say from one pixel to the next then the curve can transgress the intended edge resulting in unstable boundary contour. Therefore, the algorithm parameters are set to align with a high-intensity gradient so that the edge of the required segmentation region after necessary iterations ultimately marks the boundary contour.

It is sensitive to noise and due to lack of gradient in the depth of concavities prevents its convergence to delineate such boundary contours. The algorithm fails to handle the heterogeneity of the image gradient which represents the variable topology of all medical images, and it also requires recurrent additional parameters with changes in image geometry. It is these limitations and increasing demand for medical image analysis and processing, which have given a tremendous impetus to the computer vision community to make necessary modifications to already ingenious 'Snakes' and develop further new ideas. Most of the modifications within physics of energy minimization are at the level of image spatial domain, higher

dimensional model geometry and energy terms; and non-parameterization of the model to reduce the number of iterations.

14.3.2 Balloon 'Snakes'

Balloon 'Snakes' curve(Cohen, 1991) is an extension of primordial 'Snakes' discussed above. Balloon 'Snakes' was devised to overcome freezing of the 'Snakes' at local minimum image energy in presence of noise and homogeneous image intensity gradient. In this algorithm, a new energy functional term called balloon force or pressure is added to the primordial 'Snakes' that controls the amplitude and direction of the balloon curve. The control points of the balloon curve have positive and negative signs, which respectively inflates or deflates the active curve depending on the magnitude of balloon force to overcome noise and react to gradient potential referred to as distance potential force to increase capture range and refrain from leaking through weak edges. Balloon models operate better when edges are strong, or the image had been pre-processed by applying an edge detector filter. Balloons fail to recognize subjective contours and inflate beyond the expected boundary contour and may overshoot missing edges.

14.3.3 Intelligent balloon

The algorithm of the deformable intelligent balloon model is interactive, and it also operates on the energy minimization principle, with prior knowledge of the structure due for segmentation(Duan & Qin, 2001). The highlight of this balloon is that it can simultaneously recover both arbitrary shape geometry and changing topology. The algorithm begins simply with interactive seed implantation, which then begins to deform the model and grow towards the edge according to modeled dataset. Another significant feature of the algorithm is that it adaptively subdivides the model geometry locally and globally during deformation and automatically detects self-collision. It is because of this self-collision recognition the model evolves correctly to modify following the image topology and mark the boundary contour. Combination with mature polygonal mesh optimization techniques makes the process of intelligent balloon effective to generate good quality segmented regional topology.

14.3.4 Level set methods

The level set technique came to be introduced into image segmentation to resolve the issue of segmenting an image with variable topology by applying an active curve or a surface. The curve or the surface is activated and updated at each time step by forces derived from the image intensity. The level set method was developed to describe the front propagation to model ocean waves and burning flames(Osher & Sethian, 1988) based on Hamilton-Jacobi formulations. And later it was applied to medical image processing. The construction of the technique is simple, where a virtual model of curve or curve embedded in a 2D or 3D surface is laid over and above the image plane, generally referred to as hyper-surface. The 'Snakes' algorithm advances the curve perpendicular to the control points explicitly with the knowledge where the front would terminate. Upon meeting sharp corners, concavities, and changing topology, the orientation and number of control points of the evolving front must be re-parametrized interactively to keep the distance between the points smaller and smooth, which makes the process unmanageable(Sethian, 1985). Further to help the evolving curve or surface continuously flow and mark the edges of changing topology it must split, and merge frequently as required(Malladi et al., 1995).

To simplify, imagine when a large blob of honey flows under the force of gravity on two adjacent objects, for example around the shape of a honey dipper, it is perturbed at the interface and splits upon meeting the obstruction (Fig. 8). After marking the outer boundary contour of the first object the residual blob soon merges and continues to evolve and move under the same force of gravity to mark out the boundary of the next object. With the unchanging force of gravity, the continuous flow of honey with the constant speed on the changing topology of the honey dipper demonstrates the fundamental concept of the time-dependent

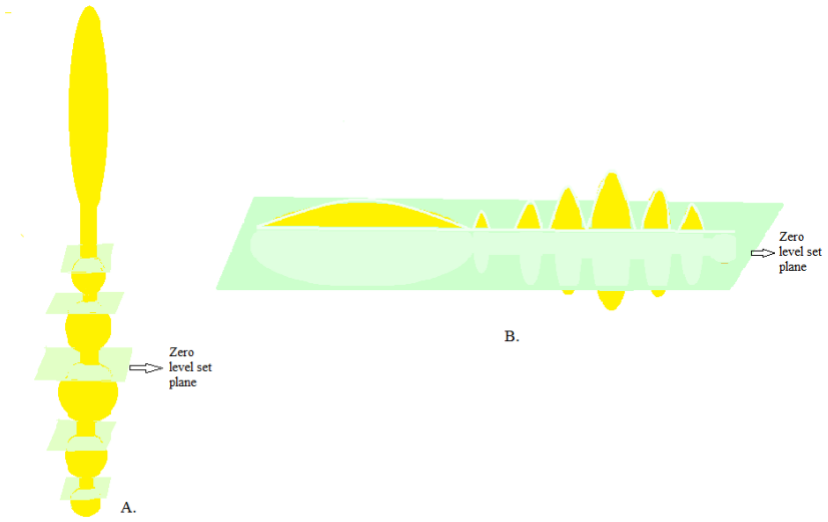


Figure 8 – Illustration of honey dippers, **A.** showing multiple zero level set planes along the path as honey flows from top to the bottom of the dipper over its variable shape and dimensions; **B.** demonstrating part segmentation (white outline) of honey dipper elements as the level set plane/hypersurface progresses over its varying topology.

implicit definition of front propagation. Similarly, in each level set state, the value is zero at any time during propagation or progressive motion of the curve or surface front with a constant speed. When the zero-thickness hypersurface is not engaged, its speed term at time zero is <0. At the interface, it is set to move at the speed term equal to zero and as it propagates it is >0. Each time the propagating surface or curve meets a new obstacle formed by the changing regional topology it is stopped in its path and reset to zero at each passing time step. The whole process is reiterated between these values. The force is high/positive inside the ROI and low at the high-intensity gradient at the edge, stopping the curve to define the boundary contour. The direction of the moving front is always normal (perpendicular) to mark out the boundary contour evenly rather than lopsided.

Although this active contour model for flexible topology does not require prior knowledge of object shape knowledge of image anatomy can be useful to introduce repelling force to prevent curve leakages at the edges of the region-of-interest.

The initial curve can be placed anywhere in the plane of an image. The embedded curve is programmed to search the boundary of the ROI either by propagating inward or outward in the normal direction. After initialization, the algorithm is automatic except to vary the smoothness of the curve to match the boundary contour. Despite its ability to manage segmentation of variable shapes and textures in the image, the level set technique has difficulty fully segmenting an edge with deep concave irregularities. This may be resolved by having an extrinsic and an intrinsic level set curve to define the boundary contour and the edge can be made stronger by combining the level set approach with an edge detector(Shah, 1996).

14.3.5 Gradient vector flow for 'Snakes'

There had been several modifications to overcome the difficulties of curve initialization and propagation into the concavities by making changes either to the forces related to the curve or the image properties. One such example is pressure force to inflate and deflate the curve as a balloon to push the curve towards the boundary contour faster and into the concavities of the edges. Too much force can cause trespassing

of the curve through areas of low-intensity gradient edges at the boundary. Alternatively, the propagation of the curve may be influenced by increasing the image-based potential forces, such as introducing dense vector fields called gradient vector fields (GVF)(Xu & Prince, 1997). In this technique, the energy minimization is achieved by solving linear partial equations which diffuse the gradient vectors of a grey-level edge map computed from the image. The active curve run on the concept of gradient vector field algorithm is called gradient vector field snakes. As a result, the GVF snakes can be initialized further away from the boundary contour of the ROI, and progress into a concave boundary contour. At least, it tends to resolve the issues of curve initialization and effectively intrude into narrow and deep spaces to mark concave boundary contour.

The formulation of the GVF field vector is such that it directs the Snake normal to the edges, varies uniformly across homogeneous regions of an image, has large magnitude only when close to the edges, and forces the Snake into the concavity instead split to skip across the aperture of the concavity. The generalized diffusion equations of applied physics are employed and solved by discretization and iteration to "fill in" the concave boundary missed by the original 'Snakes' formulation(Xu & Prince, 1997). GVF snake can be initialized inside or outside the boundary contour for segmentation and does not require prior knowledge of image structure. If the balloon model bulges through rents in the edges, then the GVF snake misses sharp corners by forming round outlines due to its chosen regularization internal forces resulting in a less-than-optimal edge map.

Like other parametric and geometric active contour models, the GVF snake too fails to automatically handle variable topology in an image(Li et al., 2005).

14.3.6. Chan-Vese active contour model

The Chan-Vese model based on the combination of a piecewise smoothing and termination criterion of an edge defining algorithm(Mumford & Shah, 1989) and zero level-set terms is energy minimizing active contour algorithm(Chan & Vese, 1999) independent of edge gradient function developed to segment very noisy or smooth edges of a ROI that marks both internal and external boundary contours.

The Chan-Vese active contour model algorithm directs a closed curve to effectively define the edges. It can be initialized either endogenous or exogenous to the ROI. It is provided with two constants during its evolution, each representing the averages of the image intensity gradient "inside" and "outside" the boundary contour. The endogenous area belongs to the ROI to be segmented. The algorithm is formed by "fitting energy" with two terms, one for inside and the other outside the edge. Boundary contour characteristics minimize the "fitting energy". If the curve goes outside, it means that the energy is >0 and vice-versa depending on whether the curve is endogenous or exogenous at initialization. The "fitting energy" will be minimized if the inside and outside intensities are more or less than zero. The "fitting energy" is minimized only in case the curve is on the boundary contour of the ROI. To make it robust, the curve is given additional regularizing terms, such as length of the curve and area of ROI within the boundary contour. The model algorithm segments ROI by simultaneously recognizing the intensity variations within and outside it. Thus, it can define boundary contours without edge gradient function, appreciating not the only smooth edge but even discrete points forming an edge. It is sensitive irrespective of the site of initialization and automatically detects the internal edge of a torus (doughnut shape). The curve length term scales to match the journey around the edges of the ROI and responds automatically to the change in topology as well.

14.3.7 Edge-flow technique

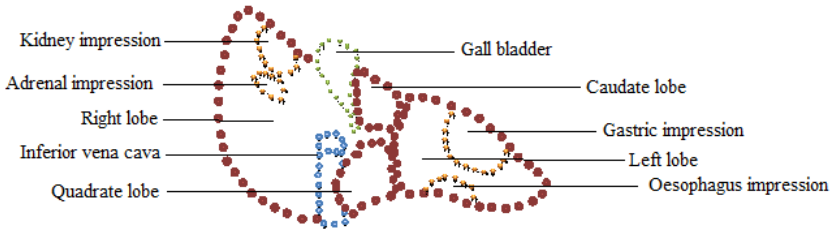
In a medical image with multiple objects as region-of-interests, there are numerous edges to segment each region. However, as the images are heterogeneous each region can be expected to have a different texture and pixel grey-level values flowing towards its edge (edge-flow) and each of the higher intensity edges, therefore, has a distinct intensity gradient. The edge-flow technique algorithm(Ma & Manjunath, 2000) is an iterative edge flow vector propagation following the texture and grey-level scale for the detection of boundary contour of a region-of-interest in image segmentation.

This technique too is based on an energy minimization algorithm, model parameters, and image properties. Like pixels, the edge-flow in the image feature domain texture elements (Texels) are defined by computed parameters depending on the heterogeneity of the pattern, size of the region, and probable regional boundaries. It is the change in local pixel and texture information that propagates the edge-flow energy. The flow of intensity gradient due to phase difference between neighbouring regions is opposite to each other upon reaching the edges as a point of zero crossing that defines the boundary contour, segmenting the two regions. Energy is required to help the flow forward depending on the pixel grey-level scale, texture scale, and intensity gradient across the boundary contour.

The Edge-flow vector is equal to its flow energy (distance × forward direction). Between two or more neighbouring regions there is a bidirectional edge flow vector, forward and backward in each region, and based on computed probabilities the flows are directed to find the nearest edge following directions of intensity gradient. At each pixel, local edge-flow energy estimates its flow direction, and it is propagated iteratively in opposite directions towards edges of adjacent regions. The energy is minimized at the zero-crossing of phase difference and upon reaching a stable state indicates the probable boundary between two edge pixels. The sum of the flow energy on either side is the boundary energy. Wherever the direction of the edge flow vector is similar propagating towards the periphery those locations are included in the local region. It is only when two opposing edge signals of the flow vector are detected the boundary is defined. A region merging algorithm is finally applied to fuse similar neighbouring regions to segment multiple regions of interest.

15.0 Statistical shape model

Human vision can expertly reveal subjective contours and recognize patterns representing a specific class of objects when marked in dots having *priori* knowledge. Like human faces, many human anatomical organs and the collection of cells forming a specific tissue tend to have some similarity and an average shape. The outline of these organs and structures has definable edges, which can be region-of-interest in an image for further examination. Technically, the shape is a property of an object which does not change during the registration process of translation, rotation, and scaling in size. When several examples of similar shapes are available a mean shape can be derived by applying statistical analysis(Cootes et al., 1992). The outline of the shape in a set of these examples called training sets can easily be marked manually by marking points or dots on the edges and salient surface features. This basic art concept and 'Snakes' to deform the connected point annotations is the basis of what has been described as statistical



A.

B.

Figure 9 – **A.** - A point cloud of undersurface of the liver showing various lobes, impressions of nearby organs, including gall bladder and inferior vena cava; **B.** - a point distribution model of an oak leaf in red, its central and peripheral veins in green and that of Sternal buckle-G (IP of the author) in blue dots placed between leaflets/ribs.

shape modeling (SSM) technique(Cootes et al., 1992). Thus, SSM is a technique to construct the mean shape (an atlas or prototype) of a specific class of object or ROI in an image from several modes or variations as examples or training samples of an anatomical structure.

A pre-requisite to construct a mean statistical model from a series of 2D shape examples in a training set is that all the annotated points on the shape examples are located at the same corresponding positions of salient features. The cloud of all the points representing the whole image or a region-of-interest called point distribution model (PDM) is the active shape model (ASM)(Cootes et al., 1995) (Figure 9). When the statistical shape model encompasses the surface features (texture) by incorporating the grey-level intensity information it is called active appearance model (AAM)(Cootes et al., 1998; Cootes & Taylor, 2002; Timothy F. Cootes, Gareth J. Edwards, 2001), which increases the strength of this technique. When the positions of relevant anatomical point annotations that describe the frame of ROI are analyzed they construct a mean statistical shape model. A computer program is generally developed to automatically perform this tedious and time-consuming task of building PDM as a few dozens to hundreds of examples may have to be produced to generate a mean statistical shape model. The cloud of a point distribution model can be denser in 3D images for segmenting ROI for surgical application.

Where there are significant variations in shapes (modes) among the examples of the training set, at the time of registration the training set is further computed to scale all the samples to minimize the differences by applying generalized Procrustes analysis(Gower, 1975). Generalized Procrustes analysis iteratively aligns, scales, and reorient the whole training set to develop the mean statistical shape by minimizing the sum of squared distances between corresponding points to fit all the included examples.

Another issue with the PDM of a very dense and variable cloud pattern is to somehow select the most important shape vectors to compute the largest variance. This is achieved by applying Principal Component Analysis (PCA)(Jolliffe, 2002). Based on the normal bell curve the first 25 percent principal components have 68.5% of the variance (one standard deviation) and the first 100 have 90.1% of the information. The PCA reduces the number of variables in the mean shape and still represents almost the same boundary contour of the ROI. To build a quality statistical mean shape model maximum number of annotation points, mode variations in the ROI and most relevant pixels at the edge are included to increase the precision, particularly when segmenting a malignant lesion for surgical intervention.

ASM reconstruction is simple and fast but less robust in managing a new image with extreme variations in texture and irregular boundary of a malignant mass. When both ASM and AAM models are combined they have a greater range of parameters for statistical analysis and accuracy to mark out the boundary contour. However, if the difference between the mean shape and the new image under study differs significantly there will be a loss of outlying points because of Procrustes and Principal Component analyses reducing the quality of the image segmentation(Heimann & Meinzer, 2009).

SSM technique is suitable for image processing of stable hard tissue anatomy and introduction of automatic annotation of points for modeling 3D image segmentation has popularized the technique(Heimann & Meinzer, 2009). The region-of-interests in medical images that change rapidly within a short period at follow-up and undergo shape transformation during registration are unsuitable for analysis with SSM technique. Hence, it is only proficient in model-specific classes and stable shapes, such as the normal anatomy of the bones and brain.

Although larger training sets can increase flexibility and universality of a mean shape but makes the overall task of population recruitment expensive. SSM technique well established in face recognition as per its origin may not be ideal for medical images in clinical practice.

16.0 Three-dimensional morphable model

The completeness of 2D images of 3D human anatomy and image quality within a specific class of image is dependent on the pose, projection, illumination (brightness), the density of tissues, the effect of occlusion by adjacent structures, the motion of anatomical region during acquisition and temporal variations on follow-up. The lesions seen within ROI in orthogonal views can change configuration relative to each other unless coupled together directly by some salient point correspondences. The maximum clinical value of any imaging technology for diagnosis and treatment planning can only be realized by 3D reconstruction. Like the SSM technique, the 3D morphable modeling approach for image segmentation requires class-specific prior knowledge for developing training sets and statistical information(Blanz & Vetter, 1999). Ideally, 3D models for medical applications require 3D training examples, which are frequently obtained from computed tomography and magnetic resonance imaging techniques.

The technique of 3D morphable models includes shape and texture vectors obtained through dense pixelwise correspondences to merge shape and texture properties into a single framework is called image vectorization(Jones & Poggio, 1998). The texture vector contains the grey-level value of each pixel in the examples of the training set and the prototype image model for correspondence to a new patient image.

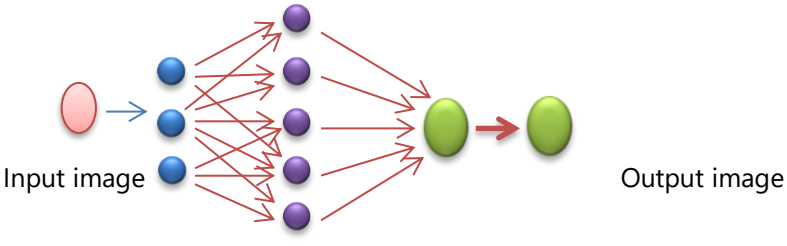
Huge variations in medical images in the same class and within the same subject over time and during image acquisition require hundreds of training set examples. There are three steps to prepare a training set and construct a 3D morphable model(Blanz & Vetter, 1999). The first step involves pre-processing multiple views of 2D radiographic images or scans of 3D reconstruction from CT or MRI images to remove any artifacts and select to be modeled. Next, pixelwise correspondences are computed between each scanned example and a reference image mesh. Finally, the example images are aligned and scaled by applying general Procrustes analysis, followed by PCA. The major disadvantage in building a 3D morphable model is

the expense and ethical issues collecting hundreds of scanned 3D training examples, and pipeline to process the raw data onto a standard 3D mesh(Blanz & Vetter, 2003).

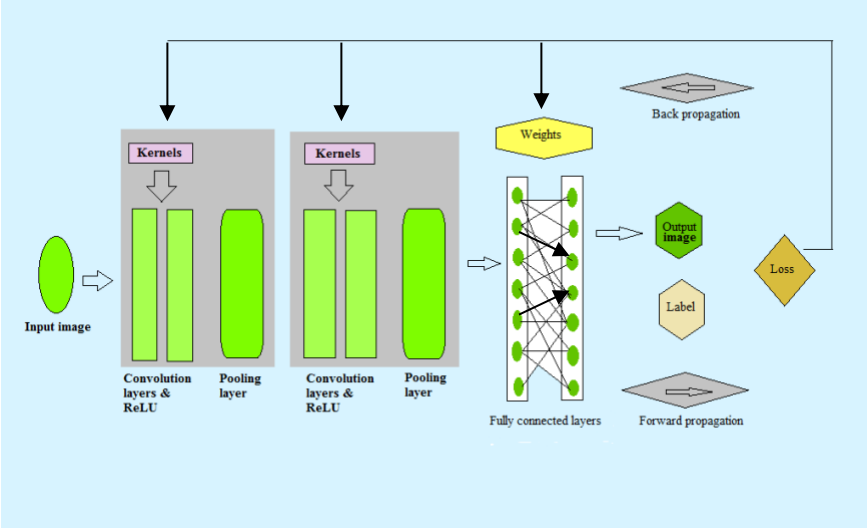
17.0 Deep learning neural networks

In the case of human vision, the moment an image is reflected on the retina, the humans can immediately label objects of interest and respond appropriately. An expert physician requires repetitive training and practice to characterize a medical image to recognize and segregate by focusing (attention mechanism) on a region of interest. It is this property of intricate operations of the multilayered anatomy and physiology of over 100 billion neurons in the brain that inspired the development of neural networks for image analysis and segmentation process. The neurons in the brain fire impulses when the retina is exposed to vertical, horizontal, and diagonal lines or edges(Hubel & Wiesel, 1962). And the idea that neurons organized in columnar architecture respond to a stimulus can be trained to learn graphic information for future applications encouraged the development of computer-based neural networks, often referred to as machine learning. It led to the development of multiple computer vision techniques including machine learning to detect features of a variety of objects in a scene. Machine learning process requires training set of specific image/s to train the network, a verification set in the form expertly segmented image as reference and parameters for training, and evaluation or test set to verify and authenticate the newly segmented patient image under examination. Deep learning is a machine learning mathematical algorithm, which has multiple layers to progressively train and learn in depth about the object of interest and refine it to extract features of interest from a busy image.

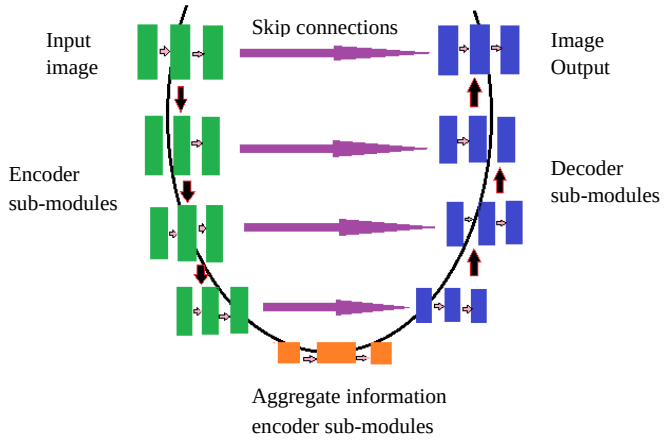
The artificial neural network (Fig. 10) is a computing system that consists of collection of artificial neurons carrying signals, process the information and pass the information further down the network. Each artificial neuron like biological neuron gets activated when signal intensity reaches certain threshold. To learn the task and train the network "weights" or parameters are added along with "bias" term as a constant value to the inputs by processing number of reference images (training atlases) through multiple successive layers. The process of network training continues through backpropagation and assessment of loss function (error) until there is minimum error by adjusting "weights" between the predicted output and the target output. The target output is either based on predetermined criteria or manual segmentation by an expert to set "ground truth" to increase robustness of the model and decrease value of loss function. Generally, the error between predicted output and expected target output never approaches zero and frequently need redesigning the network. Although it is a supervised model, however once the network is trained and optimized the specific task is automatic for segmenting and diagnosing a particular class of pathological lesion or an intended object in future.



A.



B.



C.

Figure 10 – **A.** – traditional multilayer perception neural network; **B.** - multilayered convolution neural network. The kernels/filters convolve with the image to extract various features; the linear operations of the convolution engage nonlinear activation function called rectified linear unit (ReLU) to compute function f(x) = max (0, x); the pooling layer reduce dimensionality of the network; and fully connected layer aggregates information by selecting patches from input feature maps and discard others. Back propagation algorithm trains neural network through learnable kernels and weights, which are updated based on loss function; **C. -** U-Net, arrangement of channels of down-sampling (green) and up-sampling sub-modules (blue).

The traditional multilayered perception neural networks were a crude version of the biological visual cortex with an image input layer, one or two hidden layers, and an output layer(Lecun et al., 2015a) (Fig.10A). In machine learning, the input image is preprocessed by applying a kernel/filter to remove any noise and improve contrast, followed by thresholding and edge extraction. In the next layer, regional features are

extracted by collecting certain pixels based on intensity variations, texture, and colour; and the next layers resolve horizontal and vertical lines(Coates et al., 2013; Lecun et al., 2015b). The machine learning classifier such as neural network classifies and labels target features to determine the segmentation boundary. It is the neural network classifier that is trained and learns to classify new images. This kind of simple multilayer perception model becomes active only when a certain threshold value of image pixels is reached. One significant limitation of this kind of model was that it used one perception for each input pixel in an image, when multiplied by three in the case of standard red, green, and blue images, the model failed to handle large image data. Secondly, it was trained for specific objects with defined features in a fixed coordinate and, if a given object was moved even within the same scene the network corrected the orientation of the object itself as trained and failed to accept and learn the changes. This could not work optimally for the analysis and processing of evolving heterogeneous medical images efficiently. However, with the advent of graphic processing unit (GPU) and fast computer machines for the gaming industry the interest in deep learning revived(Coates et al., 2013; Lecun et al., 2015a) and the traditional neural networks evolved to convolutional neural networks advancing the process to manage the image data in the form of multiple arrays, including coloured RBG images.

Based on image texture, the segmentation task of medical images is regarded as semantic segmentation (pixel by pixel segmentation of the same class of objects in an image). For this purpose, there are number of networks in practice, such as fully convolutional neural network (CNN), U-Net, V-Net, Non-local U-Net, generative adversarial network (GAN) projective adversarial network (PAN) is briefly discussed here.

17.1 Fully convolutional neural networks

The fully convolutional neural network(Long et al., 2014) forms the basis of what came to be known as deep learning for processing image data. The convolutional neural network (CNN) classifier can directly handle raw image data without initial preprocessing and has multiple layers(Rizwan I Haque & Neubert, 2020). It has much greater accuracy for image classification, which is an important feature to classify, and label variegated medical images, particularly when an image has multiple satellite lesions. It deals with an image layer by layer, filtering, and computing relevant pixels as part of its learning process, like some of the classical computer vision techniques. It depends on the fact that all pixels in the neighbourhood forming an object, or the same type of tissue are intimately related to each other than one at a distance.

Convolution (from Latin convolvulus means to roll together, merge, fuse, or converge) neural networks are designed to adapt to the properties of images even if the tissue elements are in a different frame/slice appear at another location and can perform pixelwise segmentation process. The fully convolutional neural network architecture is made up of interrelated three types of multiple sequential layers examining the morphology of the image to locate and align edges by recognizing associated pixels/voxels forming a hierarchical structure. Every layer transforms a 3D volume input image to a 3D volume output image. The three types of layers are convolutional layer, pooling layer, and fully connected layer(Kenji Suzuki, 2017; Yamashita et al., 2018) (Fig. 10B). In the convolutional layers, learnable filters or kernels are applied to the original image. It is in these layers that the receptive field of an input image is scanned by a series of a user-specific grid of parameters in the form of filters or kernels to extract certain image features. The number of filters, filter stride, filter size, 3 x 3 or 5 x 5 grid window are the most important parameters. To convolve with image pixel grid, the kernel scans from top left to the bottom right corner of the image. At each group of pixels on the image, based on the size of a filter window convolutional value (pixel grid grey level value and filter grid weights) is calculated and stored. The process of optimizing the values and weights of the filters decides the quality of training and learning of the network. The activation non-linear rectified linear unit (ReLU) improves the training speed. Next the pooling layer down-sample the input data, extract the maximum average value of a convolved window and reduce image dimension. There can be numerous convolutional and pooling layers to increase the capability of the network.

In the end, the fully connected layer has complete connectivity with all the neurons in the preceding and succeeding layer and collates the averaged information from the preceding feature maps to determine which features most correlate to a specific class in the image. The filter hierarchy in each layer learns an increasing number of features and the final layers have complete image representation to recognize the entire shape, form, and orientation of the structures in the image. To train a neural network the computer adjusts its kernels/filters in convolutional layers and filters values or weights in the fully connected layers through a training process called back-propagation(Yamashita et al., 2018). Convolutional neural network filters are given values and weights randomly at the initiation, and during the forward propagation, the output image loss function calculates the kernel values. The weights and learnable parameters are updated according to loss value through optimization back-propagation algorithm(Yamashita et al., 2018). The efficacy of an image segmentation system may be evaluated by using standard statistical methods for accuracy, precision, and reproducibility to assess for correctly predicting a total number of actual disease pixels/voxels.

17.2 U-Net, V-Net, and adversarial networks

The U-Net(Ronneberger et al., 2015) is so named because the architecture of channels, encoder contracting path and decoder expanding path submodules are arranged symmetrically with skip connections between the descending and ascending limbs (Fig. 10C). This high performance network based on fully convoluted network without a fully connected layer can be trained "end-to-end" from very few images in the training set. Also, the pooling operators are replaced by up-sampling operator to improve resolution. As there are few examples to teach the network the data is augmented by rescaling using rotation, shifting, elastic deformation and even making changes to pixel grey values. For training the input images and corresponding segmentation maps are used to train the network. The U-Net takes less than a second to segment a 512 x 512 image on a GPU. There are four encoder submodules in the descending limb and four decoder submodules in the ascending limb. The encoder has two convolutional layers for down-sampling and decoder submodules successively increase resolution by up-sampling of input image dataset and enhance accuracy of segmentation by inclusion and exclusion of relevant pixels. The network skip connections between the descending and ascending limbs for up-sampling resulting in same resolution as input in the encoder submodule as the input of the next decoder submodule, thereby it has the capability to handle simultaneously low level and high level information increasing accuracy and extracting complex features. The same group introduced 3D U-Net(Çiçek et al., 2016), where the 3D operators of the network architecture consisting of 3D down-sampling convolution, 3D up-sampling convolution, and 3D max-pooling layer is fed with 3D sparsely annotated limited number of images for training. The data is augmented by elastic deformation, and it is the weighted soft-max loss function that allows the network to learn from only labelled pixels while setting the rest to zero. Once the network is trained on sparse 2D CT data, input of individual slices arranged in a stack can provide segmentation of 3D volume. Residual U-Net and hybrid dense U-Net are weighted convolutional networks. The former network is preferred for retinal vascular tree segmentation in diabetic and hypertensive retinopathy and latter is for segmentation of liver tumour in contrast enhanced CT imaging.

The V-Net(Milletari et al., n.d.) is like 3D U-Net trained on limited number of images in a training set. The data is augmented by applying random non-linear transformation and histogram matching. To discriminate between enormous size of background and much smaller foreground voxels, the V-Net employs Dice co-efficient (calculate similarities) loss function instead of statistical mean square error (the average squared difference between the estimated value and actual value), or the cross-entropy loss function (entropy is measure of randomness in information and cross-entropy refers to difference in two random variables). The descending limb path of the V-Net compresses the data and during up-sampling gradually decompress the

signal until it reaches the original size extracting the required features for image segmentation. The architecture claims to ensure convergence in a fraction of the time compared to similar networks that do not learn residual functions.

Generative adversarial network (GAN)(Goodfellow et al., 2014) is based on game theory philosophy having two models to train simultaneously to segment the intended group of pixels. The generative component adds random noise to the input image fabricating a pseudo-image and the adversarial part of the network is the discriminator that estimates the probability to remove the random noise from the training data to segment the target pixels in the image. Projective adversarial network(Khosravan et al., 2019) (PAN) utilizes 2D CT image projections without the need for 3D reconstructed image to generate high level 3D information for segmentation. The three component architecture of the PAN consists of a fully connected network called segmentor with an encoder and decoder to handle 2D grey-scale input image to produce a pixel level probability map, and two adversarial networks to capture high level information during training phase. One of the adversarial networks has an additional attention mechanism module to focus on highly desired object pixels to increase overall performance of the PAN model.

18.0 Evolution based genetic algorithms

Evolutionary algorithms being beyond everyday simple observations are considered metaheuristic. They are based on biological concepts of natural selection and the genetic propagation of species. An evolution-based computational genetic algorithm(The Evolution of Intelligence. The Nervous System as a Model of Its Environment, n.d.; Holland, 1962) is not an image segmentation algorithm but a search-based optimization technique applied to engineering problems. Such a technique is frequently applied secondarily to find an optimal solution to difficult problems, which otherwise would take an extremely long time and require extensive computational resources to solve.

In nature, there is continuous randomization of events and physical adaptation without a fixed scheme and pattern. Over time because of this continuous change the nature is forced to produce better solutions more often than best to overcome arising adversities. The same would apply to a genetic algorithm for an engineering problem to find a better solution from among many possible alternatives to solve a given problem in terms of cost and time, rather than the best which may not be practical and cost-effective. In this regard, genetic algorithms have proven to be a reasonable metaheuristic method to provide practical near-optimal solutions for the image segmentation process within minimal computational space, cost, and time. Analogous to evolutionary genetics a set of solutions is chosen heuristically (former experience or speculation), which form a given population (the total number of solutions from which samples are selected for statistical measurement). Each solution is a representation of a chromosome and elements of that solution are the genes arranged in a specific order (Fig. 11). The values or parameters given to each gene are alleles. Within the computational space arrangement of the solutions form genotype, and upon decoding the most effective solutions from genotype space enters phenotype space (activity space) as functionally adaptable solution to solve the problem. Otherwise, to better the functional fitness of the selected solution it is encoded and returned to the genotype space for re-evaluation. The genetic operators alter the composition of the solution (chromosome) by changing the parameters through the exchange of values (alleles) between a selected pair of solutions (chromosomes) called crossover operation or introducing new parameters (mutation) within one or more solution elements (genes) of the set of solutions (chromosome population), or even make new pair of solutions to find a better solution and make the algorithm robust. Once a 'fit' solution is found to solve the problem the algorithm (evolution) is terminated. The random process of crossover of values (alleles) at one or more specific crossover points is considered one of the most significant steps in running a genetic algorithm(Goldberg, 1989) to better the solution to make it robust for solving the problem in-hand. Mutation adds randomness and more options to the current set of solutions.

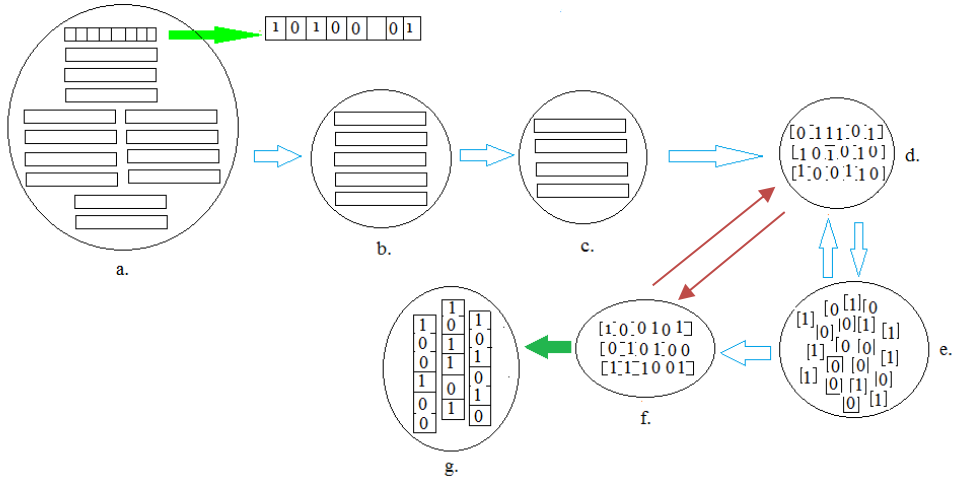


Figure 11 – The steps of genetic algorithm for optimization, **a.** each of the solutions (chromosome) is a string of real values (genes) encoding the parameters, **b.** selects a set of random strings from the old population and **c.** choose the best individual strings, **d.**, **e.**, and **f.** process of selecting the best values and discarding others, **g.** new generation of strings with best values. It is the variation operator that combines and alters best values in each string by crossover blending and random mutation.

19.0 Technique of animated deformable organism

The subject of virtual animation and artificial intelligence is a rapidly expanding field. Deformable organism model (Hamameh et al., 2001; Hamarneh et al., 2009) for medical image analysis is a parametric semiautomatic image segmentation technique based on well-established interdisciplinary specialties such as continuum mechanics, Newtonian dynamics and numerical computation, differential geometry, vector calculus, and computer graphics(Nealen et al., 2006). In principle, the deformable organism is advanced modeling of 'Snakes' and other deformable models based on time-integrated Euler-Lagrangian partial differential equation algorithms for image segmentation(Heimann & Meinzer, 2009). The deformable organism is an interactive systematically mechanized model with artificial life control algorithms given a predefined set of criteria to manipulate and deform its shape in response to image properties. The activation of model deformation and behaviour is based on the perception of the image data and prior knowledge of anatomy, topology, pathology, and any other structural landmarks. The model is programmed to be aware of itself and its environment to search intelligently by exploring the entire image space to reach correct solutions.

19.1 Deformable organism model

The deformable organism model has a hierarchical structure incorporating biomechanical principles and a higher life centre. It is organized into four main systems consisting of structural geometry, motor system, a sensory system for perception, and a cognitive centre to manage the behaviour of the organism, which defines it as an artificial intelligent 'life'. The geometry of deformable organism that forms the morphology, shape, and topological constraints is built on the medial (*median*) axis(Hamarneh et al., 2004) and medial (*median*) sheet(Hamarneh et al., 2007). Geometric deformation (stretching, shrinking, bending, thickening) manipulates the geometric structure of the model without constraints. The structure explicitly stores the

location information on its Lagrangian surface mesh and internal nodes, and in the Eulerian model, the shape is formulated to zero level-set function. The motor system simulates Hooke's spring-mass law. The actuation of the motor system at basic low-level skills brings about simple bulging and stretching, and higher-level skills are parametrized to cause rigid transformations such as translation, sweeping, bending, and smoothing the medial (*median*) axis depending on underlying geometry. Sensory perception is facilitated by providing a set of sensors to perceive the external environment to gather information regarding image properties and interact with other organisms in the image space. There are image gradient sensors, texture sensors, and edge detection sensors. These sensory organs can be trained to focus on specific image features so that non-essential non-region-of-interest is ignored and search for relevant programmed features. There are internal sensors confined to the organism and external free-floating sensors within the image domain. The sensory information is transmitted to the cognitive centre for behaviour modification of the organism. The cognitive centre responds to the internal and external sensory stimuli as predefined parameters, as well as to any user-initiated changes to redirect the behaviour of the organism to match image segmentation requirements. In return, per-programming and learning ability over time, the model behaves by appropriately deforming and translating to detect the edge.

According to prior knowledge of anatomy and segmentation requirements and, programmed parameters, the organism initiates the process, deforms, analyzes, make decisions, and segment ROI as planned. The deformable organism model has the capability of processing 3D medical images.

19.2 Artificial intelligence deformable organism model

For segmentation of MRI images, an artificial intelligence organism with a primitive brain has also been developed(McInerney et al., 2002). This kind of intelligent organism model is thought of as higher evolved versions of deformable active contour models. The structure of a deformable organism consists of a virtual body form capable of altering its shape to match the ROI for segmentation. The construction of boundary contour is achieved through sensory communication between image properties and the primitive brain of the organism. The rudimentary brain has centres for perception, motor control, and cognition. They can perform voluntary movements and respond by altering body shape based on sensory input and motor response. The organism once released in the image domain tends to search for edges of the region for segmentation.

20.0 Solid modeling

Solid is a three-dimensional object bounded by a uniform surface having a homogeneous continuous dense interior with four or more facets (polyhedrons), sphere, and cylinder without overhanging edges. Mathematically, one can derive mass and area to determine its volume. On the other hand, a wireframe is an ambiguous three-dimensional object. It is an incomplete structure unsuitable for mass and volume determination. A polyhedral object whether solid or wireframe obeys geometric principles relating to its shape and parameters. The movements of their edges and vertices respond following the Cartesian coordinates in Euler space. And, how the various parts of the object are related to each other expresses its topology. Fundamentals of constructive solid geometry in combination with elementary geometric primitives such as lines, triangles, blocks, etc. provide necessary elements for constructive solid modeling(Requicha & Voelcker, 1982).

20.1 Parametric solid modeling

A parametric solid is a constrained parametrized object whose parameters (dimensions) can be manipulated in a stepwise fashion according to the wish of an operator to achieve the required configuration instantaneously. In the computer environment, the creation and alteration of 3D shapes are achieved by changing parameters to manipulate its size and configuration to exploit the design of an object. The process of parametric solid modeling acts on the meta-structure of a shape as it can be modified to create new

instances by implementing computer-aided design software in comparison to the geometric structure which represents a specific real shape of the solid object and its constant topology.

Parametric modeling technique has the potential for application to segment and create new instances of skeletal elements as 3D models and even change dimensions of a reference prototype radiological image to match patient-specific new target image. In the case of the thoracic cage skeleton, there are multiple elements, each element possessing unique architectural geometry and topology created by their adjacency and connectivity with nearly solid cortical and trabecular structure internally. Unlike a true solid all its components are irregular having overhanging edges at places as seen on plain radiographic and computed tomographic images.

The intention to construct a parametric solid model is to develop a competent technique whereby new parametrized instances of skeletal and soft tissue structures can be built and segmented for application of surgical implants and finite elements for non-invasive analysis in a virtual environment. Application of parametric solid modeling lends itself to construct a model of 3D human vertebral column satisfactorily(Rodriguez et al., 2011). The method is history-based modeling from a 3D reference prototype CT reconstruction image to parameterize the new patient-specific image in terms of morphometric dimensions of its anatomy and bone density. To the constructed model patient-appropriate related boundary conditions and loads can be applied to reach the correct solution when FE analysis is sought. Unfortunately, as the number of overall parameters will increase, so would the computation time and cost.

Unlike plain radiography, CT can acquire 3D anatomy easily but there is concern over excessive radiation to the patient. However, from a single generic atlas of 3D reconstructed CT image parametric solid model of the same class of two or more plain radiographic images of a new patient can be constructed for surgical application. The thoracic skeletal cage has an elaborated anatomical structure of articulated twelve pairs of ribs and cartilages with the sternum on the front and twelve thoracic vertebrae on the back, all having a variety of articulations and accessory structures. Within the computer virtual environment appropriately designed or existing engineering software for parametric solid modeling may be used to re-parametrize the plain radiographic views by using the image registration method. One such open-access software called 3D Slicer (www.slicer.org) employs greyscale densities to segment skeletal parts(Rodriguez et al., 2011). Segmentation of the CT slices can be performed manually by using a thresholding tool such as a computer mouse to create a 3D labeled map. Each segmented element can be annotated in a variety of colours to identify the thoracic cage components. The soft tissue structures can also be defined if MRI slices are used for developing the model. The thoracic cage model can be iteratively smoothed if required to remove noise and edited before stereo tessellation lithography (STL) file is generated for transfer into 3D slicer or other engineering CAD software for parametric solid modeling for creating a new target 3D model.

The step in parametric solid modeling that helps to resize a solid structure to manipulate all the parts at once by changing one parameter may not apply to irregular widely varying human skeletal anatomy of the thoracic cage, which is an assembly of multiple structures. User interaction may become necessary to adapt the prototype parametric solid model for refining its application and registration to the new patient-specific image by moving specific control points at the salient anatomical vertices or an alternative technique is devised to constrain the use of automatic re-parametrization method.

The pipeline line of the parametric solid modeling (Fig. 12), say for thoracic skeletal structure begins with 0.625mm thick CT slices. Each slice after the smoothing process is segmented manually or using an automatic deformable model for each element of the articulated cage skeleton. 3D CT-based reference prototype image of the segmented structure is mapped. With the help of the 3D Slicer from the patient-specific 2D plain radiograph 3D parametric solid model is prepared by registering a 3D reference CT

prototype, followed by creating STL from the created 3D patient-specific model. The patient-specific STL file with an applied surgical implant is transferred to a finite

3-D Slicer application

2-D 0.625mm CT slices of the chest

↓

Smoothing of the skeletal elements

↓

Segmentation of the skeletal elements

↓

3-D reconstruction of the thoracic cage

↓

Mapping of the elements

↓

Prototype 3-D image of the thoracic cage

↓

Patient-specific 2-D plain radiographic views

↓

Registration of prototype to the patient-specific views

↓

3-D parametric thoracic cage model

Patient-specific STL file of the thoracic skeleton

+

Test implant STL file

↓

Sternum bone-implant interface development

↓

FE model and analysis

Figure 12 – Parametric solid modeling algorithm

element analysis space for the creation of the mesh. Patient-appropriate boundary conditions and, bone and implant material properties are applied to obtain stress-strain von Mises maps. This pipeline set up from the step after the creation of the patient-specific STL file onwards can be run in parallel on multiple computer stations to receive results from several different types of implants at the same time. And, depending on the facilities the setup may be run for as many patients as necessary for more than one elective operating list. Such a computational patient-specific and patient-appropriate operating model in practice is feasible only in the presence of a clinical biomechanical engineering team.

The single most advantage of the parametric solid model will be its ability to create all shapes and sizes of normal and anomalous sterna, ribs, and vertebral anatomy as new patient-specific instances and finite element analysis by applying patient-appropriate boundary conditions to choose a better implant for the best possible outcome in a patient.

20.2 Direct modeling technique

It is a technique of building and editing a model that goes together. In this computer-aided model, the operator directly manipulates the geometry of a 3D solid or a wireframe at its vertices. This model does not have a ready history of parameters to which the dimensions can be changed to fit a new shape of the same class of object. This adds significant versatility to this modeling technique. Here the designer can perform direct deformation to change the configurations of the native object elastically as a go-along design project. The ease of modification adds the capability to allow rapid iteration of the 3D base raw wireframe stock model to be converted to the new patient-specific image as an open-mesh structure initially during registration. The segmented image is sized into a 3D structure based on morphometric measurement of multiple views of 2D images by selecting salient anatomical control points and refined by iterations to smooth the whole structure. The wireframe mesh can be converted to 3D stereo-lithographic model by including the bone mineral density based on Hounsfield units recorded from the native patient-specific radiograph for each of its elements when an image structure has multiple components. The current mesh can then be transformed into a 3D fully textured voxel-based greyscale model ready for polyhedral finite element mesh to apply surgical implants and patient-appropriate boundary conditions.

Direct medical image modeling technique can open a new field for the construction of 3D solid skeletal structures from 2D multiple views of plain radiographs element-by-element to generate a completely new patient-specific 3D model without the need for a library of many atlases and a mean statistical model. Like the former history-based parametric solid modeling technique, in this history-free and edit-free direct modeling system the user may not be able to return and edit the dimensions during the post-processing phase and iterate further. This makes the direct modeling technique demanding as it requires sufficient expert knowledge of both anatomical structures and computer thresholding to take full advantage of it. It is an interactive method and completely patient-specific once the clinical biomechanical engineering team has patient-specific bone material properties and patient-appropriate clinical parameters with the collaboration of a clinical team. This modeling technique can be used for direct stereo-lithographic modeling for medical image segmentation and therapeutic applications. Most of the currently available engineering CAD systems are a hybrid of parametric solid modeling and direct modeling tools. Solidworks® comes with a direct editing option and creates a history tree with the potential to make necessary changes during the advanced modeling phase. Synchronous technology (Siemens Nx® and Solid edge®) means having a display of both parametric and direct modeling menu on the same screen environment without history-tree and the necessity to re-compute editing job.

20.3 Multilayer parametric solid modeling

The key dimensions of a skeletal region or a segment can be accurately extracted either directly from cadaveric dried bones or 3D reconstructed CT image. Then these dimensions can be used as parameters to enable customization of the same class of skeletal anatomy of a new patient from 2D multiple plain radiology images to 3D stereo images based on the multilayer parametric solid modeling technique such as for rapid shoe-last customization(Wang et al., 2011).

The multilayer parametric solid technique consists of multiple layers - feature control point layer, parametric curve layer, parametric surface layer, dimension layer, and constraint layer to obtain parametric piecewise reconstruction. An interactive draft-driven deformation method is used to customize the object mesh to match the surface curves and contours. The feature point layer represents control points marked at all the

salient features and the dimension layer to resize the dimensions of the object. The whole object is placed in the global coordinate frame such that the sagittal, coronal, and axial planes correspond to normal anatomy. This level of accuracy will help recognize anatomical details during the conversion of 2D to 3D solid images. The parametric curve layer describes concave and convex contours, and the parametric surface layer carries the mesh for piecewise interactive draft-driven deformation. The object anatomy is divided into as many regions as possible in a meaningful way representing salient features on which the technique can be conducted. When the object is divided into relevant parts it becomes easier to edit the local surface contours and edges to accurately adapt the base prototype 3D model to the 2D new patient-specific images. The Slice technique can be used to manipulate the mesh vertices and every other layer is adjusted to regularize it to the new image. In the case of the sternum, it can be divided into manubrium, gladiolus further into several sternebrae, and xiphisternum. Next mark multiple control points on the curves and surface contours and sharp corners at the attachment of costal cartilages and clavicles in both right posterior-anterior oblique and lateral views. Dense labeling of surface features points and piecewise surface division help with the topology and appearance of the final model. The robustness of the parametric curves is maintained by parametric constraints for curve stability at the intersection nodes during the deformation process.

The object mesh of the 3D prototype is encoded parametrically to preserve its geometric information and then used on new patient-specific 2D multiple view images by decoding to acquire similar geometry as long their class is the same. The encoding is built on the feature control points in the feature point layer to follow the salient features of the new images based on the prototype geometry. The relative position of each vertex of the object mesh is encoded and decoded to adapt to the salient feature points on the new image. The draft-driven deformation is performed by directly dragging the feature control points and in real-time following the curves and surface contours under coplanar, dimensional, and directional constraints at the intersection nodes.

Over time numerous atlases of the same object class with variations will build a comprehensive library whatever solid modeling technique is employed. At the same time collecting a patient-matched set of data and outcomes based on age, gender, co-morbidities, BMI, etc. can help modify and develop new implant design. These modeling techniques are adaptable to medical image modalities and the design process is visible, intuitive, and interactive as the model is being constructed.

21.0 Comments and analytic conclusion

Computer vision is an interesting science made up of a kaleidoscope of numerous algebraic, geometrical, and differential calculus-based algorithms, and recently integration of biological processes. The literature is extensive, and its jumbled evolution has occurred rapidly in the last fifty years. There are numerous synonymous terms to confuse a newcomer to this science and require a simplified organization to make its study practical for those physicians interested in medical image processing and analysis for surgical diagnosis, preoperative planning, radiotherapy, and medical therapeutics.

The major objective in medical image processing and analysis is optimal segmentation of heterogeneous images, registration across available imaging modalities, 2D to 3D transformations of human anatomy, and application of FE analysis in-silico to harvest mathematical solutions for in-vivo surgical applications. To accomplish this ambition there are numerous semiautomatic and automatic computer vision techniques for image analysis and processing in the manufacturing and media industry applicable to medical images with varying degree of success. Despite great efforts there is not a single image registration and segmentation modeling technique that can be applied confidently to the medical images in daily clinical practice for therapeutic applications. To solve any scientific problem there is a set of input values that are processed to get optimal output values. For optimization or a solution providing 'best' output values in mathematics there is an amelioration exercise by randomly selecting the parameters to maximize or minimize objective

functions by varying input parameters. The set of all available solutions or values as input parameters forms the search space and within this space is present a better solution to the problem in hand, which in the case of image processing is the quest for a segmentation algorithm. As a result, there are so many attempts to refine the segmentation process to resolve few yet significant problems of image analysis and processing leading to hundreds of techniques that the task to devise a comprehensive classification system applicable to clinical medical practice is lacking. It is expected to remain in the evolution phase to develop a technique for managing medical images as a standard of care. So far most of the refinements are occurring piecemeal and have resolved pre-existing issues only partially. Currently, deep learning has become the hub of computer scientists worldwide.

The basic issues in medical image processing are blurred edges, edge interruptions, image noise, intricate tissue textures, heterogeneous intensity gradient, anatomical variations, changing topology, and changes in pathological images over time in the same patient. Moreover the 'work-piece,' the image in medical computer vision science, there is a lack of standardization of image production by the same and different radiographic imaging modalities, which leads to variations in the viewed frame for registration/ comparison and extraction of the data for concrete decision-making on each follow-up visit in the same patient remains unresolved.

As the primary objective of introducing computer vision science into the medical field is an intensive activity of image segmentation and 3D rendering of 2D radiographic images for FE modeling and analysis, there is a need for the expertise of a dedicated clinical biomechanical engineering team within a surgical unit. The simplest way for an expert to draw a boundary contour manually to segment a region-of-interest is with a pointing device such as a computer mouse or pen. The same can be done to reconstruct a 3D computed tomography volume structure by segmenting one slice at a time and finally computing the segmented slices of a region by stacking them together. It is also convenient to modify and replace the previously drawn boundary contour on the same or new follow-up image for comparison. Although manual segmentation is used frequently it is time-consuming, expensive, and lacks consistent reproduction even in the hands of an expert(Preim & Botha, 2014). But manual segmentation, the method to establish "ground-truth," may still be preferred if the image has a significant number of artifacts or the ROI in an instance of a tumour that evolves unexpectedly rapidly where an automatic segmentation technique may provide incomplete boundary contour. Developing an automatic segmentation algorithm is challenging when it comes to selecting a mathematical model and parametric values to compute missing details for optimal results and satisfactory clinical outcomes. On the other hand, the fully automatic techniques for image segmentation comes with individual pros and cons. To overcome some of the difficulties encountered development of several semi-automatic (interactive) techniques has been encouraged throughout the evolution of computer vision science(Olabarriaga & Smeulders, 2001) because semi-automatic techniques rely on operator attended computer interaction to modify parameters allowing increased accuracy.

Although currently medical image processing and analysis has been limited to radiology, radiotherapy, and radio-surgical interventions, however, the science of computer vision is rapidly expanding into the fields of computational anatomy and biomechanics, and the study of normal and pathological tissue histology. It may have application in clinical methods for the surface examination of the human body to detect for example subtle changes in hyper- and hypopigmentation changes of skin lesions such as malignant melanoma and leukoderma respectively, at each follow up visit. The human visual system can easily segregate a well-defined island of homogeneous density with a single grey-level value in a medical image. However, when a less distinct high-intensity lesion important to the physician is embedded in a background of similar grey-level value and texture, it can be very easily overlooked and missed by even an expert radiologist. A good example of this kind of scenario is a mammogram of a condition referred to as "white dense breast", which can harbour malignant tissue. It can be extremely difficult to automatically segment malignant tissue embedded in such a condition based on grey-level scale and intensity gradient, and almost

impossible to visually isolate and excise the area confidently to include all malignant extensions at surgery. Similarly, in a multilayered posterior-anterior view of chest radiographic image with a spectrum of pixel intensities a malignant lesion hiding behind a rib or sternum overlapping cardiac shadow can be easily missed. This issue has been largely resolved with the introduction of dual-energy subtraction radiography to separate the soft tissues from the skeletal structures. Still, it is not a simple task to segment thoracic skeleton by subtracting soft tissue structures and merging two orthogonal views taken separately to reconstruct a three-dimensional model and try to estimate bone mineral density to extract bone material properties for bone-implant finite element analysis.

What can be concluded by looking at the big picture of computer vision as a practical instrument in medicine is its application to recognize tissue patterns and to distinguish between ROI from the rest of its environment for qualitative and quantitative analysis, diagnosis, and temporal evolution to make prompt decision to undertake appropriate treatment. Unfortunately, to achieve these goals fully, there is a lack of reproducibility of grey-level values and intensity gradient to consistently acquire radiology images and flawless cross-registration techniques between various radiology modalities. There is also no universal standard segmentation technique that can be applied to all kinds of medical images. Almost all the known registration and segmentation methods are far from perfect for daily clinical practice.

A medical image is a collection of orderly arranged pixels to construct a meaningful image. Most of the commonly employed segmentation techniques operate on grey-level pixel intensities and intensity gradients to describe the projected information in an image. It is inopportune that intensity gradient-based techniques are easily influenced by lack of standard luminance when the images of the same anatomy are compared in a patient over time and co-registered among different modalities. The numerous mathematical assumptions and alteration of image characteristics during pre-processing to remove artifacts can undermine the physical value of segmentation and its true clinical relevance. The process of segregating ROI containing an irregular lesion in an image speckled with noise, particularly near the discriminating edges between healthy and pathological anatomy can be a serious concern during surgical planning, when there is crossing over of true and false edges creating an ambiguous line for surgical excision, as in a case for total resection of a brain tumour such as glioblastoma.

The traditional "Snakes" and other active contour algorithms are generally based on pixel grey values and intensity gradient to delineate higher intensity edges, but no medical image is perfectly homogeneous, has a noise-free region and perfect leakproof edges. If the virtual environment and the algorithm parameters are not optimized to image features and quality the process of segmentation can be easily jeopardized mid-stream. The algorithm is overwhelmed by heterogeneity, irregular and concave edges, and a lack of prior knowledge of intended regional boundaries to assist with the initial placement site of the contour model by the user interactively or training the set of images in the case of deep learning for accurate output. The other significant issue is variable topology, which in the case of a medical image can be strikingly diverse intensity gradient due to highly variable texture of pathological tissue, occlusion of edges and merger between adjacent regions, inconstant presence of air, tissue water (oedema, necrosis, cyst formation), and fat distribution add an extra layer of complexity. It is biological fact that there are no empty spaces between juxtaposed organs abutting each other or within their normal substance carrying pathological lesions, except for cleavable dissecting planes. The merger of a variety of tissues occludes boundaries, resulting in radiology images that contain many pixels or voxels of varying grey-level values to make it difficult to assign the label and weight appropriately for thresholding and edge detection.

Now multiple imaging modalities and modes are routinely used to increase the accuracy and precision of diagnosis. The deep learning convolution neural network classifier lends itself well to facilitating fusion of CT and MRI images to improve segmentation of region-of-interest(Zhou et al., 2019). Here each modality image is employed as an input image and fused with the input image of the other. The fusion of the included

images occurs either at the input level, intermediate layer level, or decision level neural network designed to independently learn complementary representative information of individual modality to create the segmented output image. Despite such increasing interest in deep learning convolution neural networks and their application to radiology images, there are some serious challenges. Clinical application of deep learning neural networks for *patient-specific* and *patient-appropriate* analysis would require an extreme degree of robustness for higher applications.

Segmentation of medical images is dependent on the quality of the image, which in turn depends on standardization of image recording modalities, surrounding environment, and layers of organs projected as overlapping anatomy. The most thwarting of all that impair image quality is the incurable noise artifacts arising from electronics of a machine. The decision to improve image quality the intensity histogram is applied to almost all intensity gradient-based segmentation techniques sensitive to noise and image manipulation during pre-processing. It revises grey-level values altering projected original tissue properties that automatically alter Hounsfield units in the ROI, apparent bone density, and ultimately the extracted values of soft tissue and bone material properties for FE evaluation, defeating the objective of making patient-specific choices. Therefore, preservation of image features such as distribution of pixel values, intensity gradient, texture, etc. are extremely important. It is this endeavour what makes the process of segmentation an optimization dilemma still to be solved in computer vision.

No two tumours appear similar in their native anatomy nor are spherical. Tumours, whether carcinomas or sarcomas, malignant or benign, come in variety of forms and shapes and consistencies like fruits and vegetables (Fig. 13) with a difference that the tumours have intricate vascular supply and extensions embedded in the surrounding normal anatomy. Unlike human vision, the science of computer vision, computer 'eyeballing,' is unable to report confidently on the likely material properties and material qualities, whether it is soft, fluffy, firm, fluid, its compressibility, and degree of hardness for practical application directly without having prior knowledge of one or more of its salient features. A lot more research effort is needed in image processing and analysis for recognition of material properties and post-processing numerical analysis and, the ability to recognize items in an image having similar grey-level pixel values to distinguish shiny plastic from real steel and titanium. The ability to discriminate material qualities delivered

Figure 13 – Like fruits and vegetables tumours come in all kinds of forms and shapes with unpredictable extensions into important surrounding structures.

at end of the segmentation process without direct haptic experience would enormously help provide significant knowledge what to expect at surgery during pre-operative planning.

Considering some of the limitations of the current segmentation techniques outlined above the hyalite sol-gel Amoeba (HSG-Amoeba) model based on the theory of pixel grey-level values, intensity gradient, and tissue characteristics for medical image segmentation have been conceived for publication in the future. It is a deformable biomimetic model. Conceptually, the interior milieu of the HSG-Amoeba model is governed by active sol-gel models of soft matter physics(Petrie & Yamada, 2016). The workings of this biophysical model are based on the anatomy, physiology, and biomechanical principles that motivate the propagation of the Amoeba Proteus by extending multiple pseudopodia under the influence of environmental stimuli. The evolution of each stage of the model during propagation follows principles of solid mechanics, fluid mechanics, and bioelectricity. Finally, at the time of apoptosis, its protoplasm transforms from sol-gel to a solid-state. During the phase transition, the substance of the organism acquires 3D anatomy of the region-of-interest such that the resultant tissue densities are representative of original tissue intensities, whereof mechanical properties can be calculated for the higher application.

The deep learning networks seems to have become the 'panache' of the computer vision scientists for segmentation of medical images, but it cannot be called the final frontier, the 'panacea' for all kinds of hugely variable images of pathological anatomy. It is true that the physics based 'traditional' segmentation techniques cannot be directly equated to the much preferred biology based metaheuristic techniques, however there is place for amalgamation of the two to drive the segmentation techniques into the future. The deep learning models are not flawless as the fully convolutional neural network can be unresponsive to finer details resulting spatial inconsistencies in pixel selection(Liu et al., 2021). Although there are deep learning models which require limited examples to train a network, however for greater accuracy like statistical shape modeling a large training dataset is needed for successful training, greater accuracy and flawless performance of the models(Anaya-Isaza et al., 2021; Liu et al., 2021). Harvesting of large number of patient records is expensive and demands responsibility of privacy and data confidentiality, which adds another layer to a model. To resolve this issue recently devised model called asynchronized discriminator GAN has a central generator to synthesize pseudo-image of input image as in the GAN model. The discriminator learns to distinguish the actual image for segmentation resulting in an anonymous secure data which is then incorporated for further learning and training of the network for segmentation(Chang et al., 2020). In a recent systematic review and meta-analysis of two hundred and seventy peer-reviewed studies out of 11,921 deep learning studies of medical and surgical imaging the quality of segmentation for diagnostic accuracy has been questioned (Aggarwal et al., 2021)! In addition, the notable deficiencies such as heterogeneity, extensive variation in methodology, terminology and outcome measure have been highlighted that can lead to considerable uncertainties to convince surgical community to accept deep learning as a trustworthy and cost-effective tool for pre-operative planning and higher numerical analysis to put into daily practice.

The main derivative of this study is that although there are numerous segmentation algorithms and optimization approaches in computer vision applied to segregate and quantify the ROI there is no technique yet ready for standardization for medical image analysis and processing to become part of a wide spectrum of daily clinical practice. Most of the segmentation techniques applied to medical images are driven by physics-based 'energy minimization' principles and numerous mathematical algorithms, rather than physiological principles. So far there is no standardization of various radiology imaging modalities and constancy of grey-level scale for medical images based on tissue characteristics to extract material properties for in-vivo surgical application. The concept of HSG-Amoeba is an attempt to include image segmentation of 3D reconstruction from multiple views of 2D plain radiographs to extract tissue properties to formulate material properties, preferably within the same computer space.

Multiple aspects of computer vision convolved with medical and surgical specialties makes it a multi-disciplinary specialty demanding immediate dire need for positioning clinical biomechanical engineers within hospital teams, particularly affiliated to surgical department. It is expected that over time with a greater understanding of computer vision techniques and machine learning, the interest of medical professionals through appropriate educational tools the whole process will develop truly into a precision instrument for the direct patient-based application. As of now, there is limited interest projected in this field by the medical and surgical specialties.

In the field of surgery, if *patient-specific* and *patient-appropriate medicine* practice must succeed effectively then the role of a clinical biomechanical engineering team within the department of surgery should not be a far cry, and its institutionalization as a hospital practitioner be seriously pursued. To implement such a vision from the bottom-up, indeed there is a need for the development of computer vision and machine learning opportunities. Ideally, it can be achieved as part of a higher clinical and surgical training curriculum in the form of a well-defined residency program under the mentorship of a clinical biomechanical engineering team at university teaching centres like any other residency program amounting to an innovating doctor-scientist graduate and postgraduate certification of medical computer science.

Conflict of interest: The author declare that he has no conflicts of interest.

Funding: The study was unfunded.

References_

Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S. W., Karthikesalingam, A., King, D., Ashrafian, H., & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. In *npj Digital Medicine* (Vol. 4, Issue 1). Nature Research. https://doi.org/10.1038/s41746-021-00438-z

Alan I. Penn, M. H. L. (1996, April 16). Estimating fractal dimension of medical images. *Proceedings SPIE Volume 2710, Medical Imaging 1996: Image Processing; (1996)*.

Aljabar, P., Heckemann, R. A., Hammers, A., Hajnal, J. V., & Rueckert, D. (2009). Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2009.02.018

Anaya-Isaza, A., Mera-Jiménez, L., & Zequera-Diaz, M. (2021). An overview of deep learning in medical imaging. In *Informatics in Medicine Unlocked* (Vol. 26). Elsevier Ltd. https://doi.org/10.1016/j.imu.2021.100723

Barten, P. G. J. (1992). Physical model for the contrast sensitivity of the human eye. *Human Vision, Visual Processing, and Digital Display III*. https://doi.org/10.1117/12.135956

Bathe, K. J. (2006). Finite element procedures. Second edition. In *Mit*.

Benoit B. Mandelbrot. (1984). The fractal geometry of Nature. *The American Mathematical Monthly*, *91*(9), 594–598.

Bertrand, S., Laporte, S., Parent, S., Skalli, W., & Mitton, D. (2008). Three-dimensional reconstruction of the rib cage from biplanar radiography. *IRBM*. https://doi.org/10.1016/j.rbmret.2008.03.005

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1999*. https://doi.org/10.1145/311535.311556

Blanz, V., & Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.2003.1227983

The evolution of intelligence. The nervous system as a model of its environment, Technical report, no. 1, contract no. 477(17).

Candemir, S., Jaeger, S., Antani, S., Bagci, U., Folio, L. R., Xu, Z., & Thoma, G. (2016). Atlas-based rib-bone detection in chest X-rays. *Computerized Medical Imaging and Graphics*. https://doi.org/10.1016/j.compmedimag.2016.04.002

Canny, J. (1986). A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.1986.4767851

Castleman, K. R. (1996). *Digital image processing* (Secnond). Prentice Hall Inc.

Chan, T., & Vese, L. (1999). An active contour model without edges. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.

Chang, Q., Qu, H., Zhang, Y., Sabuncu, M., Chen, C., Zhang, T., & Metaxas, D. (2020). *Synthetic Learning: Learn From Distributed Asynchronized Discriminator GAN Without Sharing Medical Image Data*. http://arxiv.org/abs/2006.00080

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation*. http://arxiv.org/abs/1606.06650

Coates, A., Huval, B., Wang, T., Wu, D. J., Ng, A. Y., & Catanzaro, B. (2013). *Deep learning with COTS HPC systems*.

Cohen, L. D. (1991). On active contour models and balloons. *CVGIP: Image Understanding*. https://doi.org/10.1016/1049-9660(91)90028-N

Continuum Mechanics for Engineers, Third Edition. (2009). In *Continuum Mechanics for Engineers, Third Edition*. https://doi.org/10.1201/9781420085396

Cootes, T. F., Edwards, G. J., & Taylor, C. J. (1998). Active appearance models. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/BFb0054760

Cootes, T. F., & Taylor, C. J. (2002). *Using grey-level models to improve active shape model search*. https://doi.org/10.1109/icpr.1994.576227

Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1992). Training Models of Shape from Sets of Examples. In *BMVC92*. https://doi.org/10.1007/978-1-4471-3201-1_2

Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active shape models - their training and application. *Computer Vision and Image Understanding*. https://doi.org/10.1006/cviu.1995.1004

da Silva, L. G., da Silva Monteiro, W. R. S., de Aguiar Moreira, T. M., Rabelo, M. A. E., de Assis, E. A. C. P., & de Souza, G. T. (2021). Fractal dimension analysis as an easy computational approach to improve breast cancer histopathological diagnosis. *Applied Microscopy*, *51*(1). https://doi.org/10.1186/s42649-021-00055-w

Davies, B. E. R., & Holloway, R. (2005). Machine Vision: Theory , Algorithms and Practicalities , Third Edition Order from Morgan Kaufmann Publishers. *Pattern Recognition Letters*.

Del Toro, O. A. J., Goksel, O., Menze, B., Müller, H., Langs, G., Weber, M. A., Eggel, I., Gruenberg, K., Holzer, M., Jakab, A., Kontokotsios, G., Krenn, M., Fernandez, T. S., Schaer, R., Taha, A. A., Winterstein, M., & Hanbury, A. (2014). VISCERAL -VISual concept extraction challenge in RAdioLogy: ISBI 2014 challenge organization. *CEUR Workshop Proceedings*.

Duan, Y., & Qin, H. (2001). Intelligent Balloon: A subdivision-based deformable model for surface reconstruction of arbitrary topology. *Proceedings of the Symposium on Solid Modeling and Applications*.

Dzung, L., Chenyang, X., & Prince, J. L. (1998). A survey of current methods in medical image segmentation. *Department of ECE, Johns Hopkins Univ., Tech. Rep*. https://doi.org/10.1039/C5AN01075F

Gage Canadian dictionary. (1983). Gage educational publishing company.

Gandhi, H. S. (2019). Rationale and options for choosing an optimal closure technique for primary midsagittal osteochondrotomy of the sternum. Part 3: Technical decision making based on the practice of patient- appropriate medicine. *Critical Reviews in Biomedical Engineering*. https://doi.org/10.1615/CritRevBiomedEng.2019026454

Gibson, S. F. F., & Mirtich, B. (1997). A Survey of Deformable Modeling in Computer Graphics. *Merl - a Mitsubishi Electric Research Laboratory*.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Networks*. http://arxiv.org/abs/1406.2661

Gower, J. C. (1975). Generalized procrustes analysis. *Psychometrika*. https://doi.org/10.1007/BF02291478

Haddon, J. F. (1988). Generalised threshold selection for edge detection. *Pattern Recognition*. https://doi.org/10.1016/0031-3203(88)90054-4

Hamameh, G., McFnemey, T., & Terzopoulos, D. (2001). Deformable organisms for automatic medical image analysis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/3-540-45468-3_9

Hamarneh, G., Abu-Gharbieh, R., & Mcinerney, T. (2004). Medial profiles for modeling deformation and statistical analysis of shape and their use in medical image segmentation. *International Journal of Shape Modeling*. https://doi.org/10.1142/S0218654304000663

Hamarneh, G., McIntosh, C., McInerney, T., & Terzopoulos, D. (2009). Deformable organisms: An artificial life framework for automated medical image analysis. In *Computational Intelligence in Medical Imaging: Techniques and Applications*. https://doi.org/10.1201/9781420060614

Hamarneh, G., Ward, A. D., & Frank, R. (2007). Quantification and visualization of localized and intuitive shape variability using a novel medial-based shape representation. *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*. https://doi.org/10.1109/ISBI.2007.357081

Heimann, T., & Meinzer, H. P. (2009). Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*. https://doi.org/10.1016/j.media.2009.05.004

Holland, J. H. (1962). Outline for a Logical Theory of Adaptive Systems. *Journal of the ACM (JACM)*. https://doi.org/10.1145/321127.321128

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*. https://doi.org/10.1113/jphysiol.1962.sp006837

Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encyclopedia of Statistics in Behavioral Science*. https://doi.org/10.2307/1270093

Jones, M. J., & Poggio, T. (1998). Multidimensional Morphable Models: A Framework for Representing and Matching Object Classes. *International Journal of Computer Vision*. https://doi.org/10.1023/A:1008074226832

Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*. https://doi.org/10.1007/BF00133570

Kenji Suzuki. (2017). Overview of deep learning in medical imaging. *Radiol Phys Technol*, *10*(3), 257–273.

Kenyon, C. M., Pedley, T. J., & Higenbottam, T. W. (1991). Adaptive modeling of the human rib cage in median sternotomy. *Journal of Applied Physiology*, *70*(5), 2287–2302. https://doi.org/10.1152/jappl.1991.70.5.2287

Khosravan, N., Mortazi, A., Wallace, M., & Bagci, U. (2019). *PAN: Projective Adversarial Network for Medical Image Segmentation*. http://arxiv.org/abs/1906.04378

Klette, R. (2014). *Concise Computer Vision - An Introduction into Theory and Algorithms*. Springer-Verlag.

Kokash, N. (2005). An introduction to heuristic algorithms. *Department of Informatics and Telecommunications*.

Lecun, Y., Bengio, Y., & Hinton, G. (2015a). Deep learning. In *Nature*. https://doi.org/10.1038/nature14539

Lecun, Y., Bengio, Y., & Hinton, G. (2015b). Deep learning. In *Nature* (Vol. 521, Issue 7553, pp. 436–444). Nature Publishing Group. https://doi.org/10.1038/nature14539

Li, C., Liu, J., & Fox, M. D. (2005). Segmentation of edge preserving gradient vector flow: An approach toward automatically initializing and splitting of snakes. *Proceedings - 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*. https://doi.org/10.1109/CVPR.2005.314

Liu, X., Song, L., Liu, S., & Zhang, Y. (2021). A review of deep-learning-based medical image segmentation methods. *Sustainability (Switzerland)*, *13*(3), 1–29. https://doi.org/10.3390/su13031224

Long, J., Shelhamer, E., & Darrell, T. (2014). *Fully Convolutional Networks for Semantic Segmentation*. http://arxiv.org/abs/1411.4038

Luke, S. (2013). Essentials of Metaheuristics, second edition. In *Optimization*.

Lyon, R. F. (2006). A brief history of "pixel." *Digital Photography II*. https://doi.org/10.1117/12.644941

Ma, W. Y., & Manjunath, B. S. (2000). EdgeFlow: a technique for boundary detection and image segmentation. *IEEE Transactions on Image Processing*. https://doi.org/10.1109/83.855433

Malladi, R., Sethian, J. A., & Vemuri, B. C. (1995). Shape Modeling with Front Propagation: A Level Set Approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/34.368173

Mandar, M., Sontakke, D., Meghana, M., & Kulkarni, S. (2015). Different Types of Noises in Images and Noise Removing Technique. *International Journal of Advanced Technology in Engineering and Science*.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London Series B, Containing Papers of a Biological Character Royal Society (Great Britain)*.

Marusina, M. Y., Mochalina, A. P., Frolova, E. P., Satikov, V. I., Barchuk, A. A., Kuznetcov, V. I., Gaidukov, V. S., & Tarakanov, S. A. (2017). MRI image processing based on fractal analysis. *Asian Pacific Journal of Cancer Prevention*, *18*(1), 51–55. https://doi.org/10.22034/APJCP.2017.18.1.51

McInerney, T., Hamarneh, G., & Shenton, M Terzopoulos, D. (2002). Deformable organisms for automatic medical image analysis. *Med Image Anal.*, *6*(3), 251–266.

McInerney, T., & Terzopoulos, D. (1996). Deformable models in medical image analysis: A survey. *Medical Image Analysis*. https://doi.org/10.1016/S1361-8415(96)80007-7

Mendoza, F. & Lu, R. (2015). Basics of image analysis. In B. & L. R. Park (Ed.), *Hyperspectral imaging technology in food and agriculture* (pp. 9–56). Springer Science + Business Media.

Milletari, F., Navab, N., & Ahmadi, S.-A. (n.d.). *V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation*. http://promise12.grand-challenge.org/results/

Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*. https://doi.org/10.1002/cpa.3160420503

Nealen, A., Müller, M., Keiser, R., Boxerman, E., & Carlson, M. (2006). Physically based deformable models in computer graphics. *Computer Graphics Forum*. https://doi.org/10.1111/j.1467-8659.2006.01000.x

Olabarriaga, S. D., & Smeulders, A. W. M. (2001). Interaction in the segmentation of medical images: A survey. *Medical Image Analysis*. https://doi.org/10.1016/S1361-8415(00)00041-4

Osher, S., & Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*. https://doi.org/10.1016/0021-9991(88)90002-2

Otsu, N. (1979). THRESHOLD SELECTION METHOD FROM GRAY-LEVEL HISTOGRAMS. *IEEE Trans Syst Man Cybern*. https://doi.org/10.1109/tsmc.1979.4310076

Petrie, R. J., & Yamada, K. M. (2016). Multiple mechanisms of 3D migration: The origins of plasticity. In *Current Opinion in Cell Biology*. https://doi.org/10.1016/j.ceb.2016.03.025

Pham, D. L., Xu, C., & Prince, J. L. (2000). A survey in current methods in medical image processing. *Annual Review of Biomedical Engineering*. https://doi.org/10.1146/annurev.bioeng.2.1.315

Prats-Montalbán, J. M., de Juan, A., & Ferrer, A. (2011). Multivariate image analysis: A review with applications. In *Chemometrics and Intelligent Laboratory Systems*. https://doi.org/10.1016/j.chemolab.2011.03.002

Preim, B., & Botha, C. (2014). Image Analysis for Medical Visualization. In *Visual Computing for Medicine*. https://doi.org/10.1016/b978-0-12-415873-3.00004-3

Prewitt, J. M. S., & Mendelsohn, M. L. (1966). THE ANALYSIS OF CELL IMAGES. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/j.1749-6632.1965.tb11715.x

Requicha, A. A. G., & Voelcker, H. B. (1982). Solid Modeling: A Historical Summary and Contemporary Assessment. *IEEE Computer Graphics and Applications*. https://doi.org/10.1109/MCG.1982.1674149

Rizwan I Haque, I., & Neubert, J. (2020). Deep learning approaches to biomedical image segmentation. In *Informatics in Medicine Unlocked* (Vol. 18). Elsevier Ltd. https://doi.org/10.1016/j.imu.2020.100297

Rodriguez, J., Choudhury, A. A., & Dragone, B. (2011). Application of parametric solid modeling for orthopedic studies of the human spine. *ASEE Annual Conference and Exposition, Conference Proceedings*.

Romanycia, M. H. J., & Pelletier, F. J. (1985). What is a heuristic? *Computational Intelligence*. https://doi.org/10.1111/j.1467-8640.1985.tb00058.x

Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. http://arxiv.org/abs/1505.04597

Sabuncu, M. R., Balci, S. K., Golland, P., Shenton, M. E., Shenton, M. E., & Shenton, M. E. (2009). Image-Driven Population Analysis Through Mixture Modeling. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2009.2017942

Sahoo, P. K., Soltani, S., & Wong, A. K. C. (1988). A survey of thresholding techniques. In *Computer Vision, Graphics and Image Processing*. https://doi.org/10.1016/0734-189X(88)90022-9

Salvi, J., Matabosch, C., Fofi, D., & Forest, J. (2007). A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*. https://doi.org/10.1016/j.imavis.2006.05.012

Senthilkumaran, N., & Rajesh, R. (2009). Edge Detection Techniques for Image Segmentation – A Survey of Soft Computing Approaches. *International Journal of Recent Trends in Engineering*.

Sethian, J. A. (1985). Curvature and the evolution of fronts. *Communications in Mathematical Physics*. https://doi.org/10.1007/BF01210742

Shah, J. (1996). Common framework for curve evolution, segmentation and anisotropic diffusion. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Shotton, J., Winn, J., Rother, C., & Criminisi, A. (2009). Texton Boost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. *International Journal of Computer Vision*. https://doi.org/10.1007/s11263-007-0109-1

Sotiras, A., Davatzikos, C., & Paragios, N. (2013). Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*. https://doi.org/10.1109/TMI.2013.2265603

Sperling, G. (1970). Binocular Vision: A Physical and a Neural Theory. *The American Journal of Psychology*. https://doi.org/10.2307/1420686

Tam, G. K. L., Cheng, Z. Q., Lai, Y. K., Langbein, F. C., Liu, Y., Marshall, D., Martin, R. R., Sun, X. F., & Rosin, P. L. (2013). Registration of 3d point clouds and meshes: A survey from rigid to Nonrigid. In *IEEE Transactions on Visualization and Computer Graphics*. https://doi.org/10.1109/TVCG.2012.310

Terzopoulos, D. (1986). Regularization of Inverse Visual Problems Involving Discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. https://doi.org/10.1109/TPAMI.1986.4767807

Timothy F. Cootes, Gareth J. Edwards, and C. J. T. (2001). Active appearance models. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, *23*(6), 681–685.

Vala, M., & Baxi, A. (2013). A review on Otsu image segmentation algorithm. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*.

van Kaick, O., Zhang, H., Hamarneh, G., & Cohen-Or, D. (2011). A survey on shape correspondence. *Eurographics Symposium on Geometry Processing*. https://doi.org/10.1111/j.1467-8659.2011.01884.x

Vuduc, R., & 634, G. (1997). *Image Segmentation Using Fractal Dimension*. https://www.researchgate.net/publication/2441059

Wang, J., Zhang, H., Lu, G., & Liu, Z. (2011). Rapid parametric design methods for shoe-last customization. *International Journal of Advanced Manufacturing Technology*. https://doi.org/10.1007/s00170-010-3144-y

Widrow, B. (1973a). The "rubber-mask" technique-I. Pattern measurement and analysis. *Pattern Recognition*. https://doi.org/10.1016/0031-3203(73)90042-3

Widrow, B. (1973b). The "rubber-mask" technique-II. Pattern storage and recognition. *Pattern Recognition*. https://doi.org/10.1016/0031-3203(73)90043-5

Xu, C., & Prince, J. L. (1997). Gradient vector flow: A new external force for snakes. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. In *Insights into Imaging*. https://doi.org/10.1007/s13244-018-0639-9

Yang, X. S. (2010). Engineering Optimization: An Introduction with Metaheuristic Applications. In *Engineering Optimization: An Introduction with Metaheuristic Applications*. https://doi.org/10.1002/9780470640425

Zhou, T., Ruan, S., & Canu, S. (2019). A review: Deep learning for medical image segmentation using multi-modality fusion. *Array*, *3–4*, 100004. https://doi.org/10.1016/j.array.2019.100004

Zienkiewicz, O., Taylor, R., & Zhu, J. Z. (2013). The Finite Element Method: its Basis and Fundamentals: Seventh Edition. In *The Finite Element Method: its Basis and Fundamentals: Seventh Edition*. https://doi.org/10.1016/C2009-0-24909-9

Zitová, B., & Flusser, J. (2003). Image registration methods: A survey. *Image and Vision Computing*. https://doi.org/10.1016/S0262-8856(03)00137-9