ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES

PhD THESIS

Çağlar CENGİZLER

DEVELOPMENT OF NATURE INSPIRED ALGORITHMS FOR IDENTIFICATION OF SPINE ON ULTRASOUND IMAGES IN SPINA BIFIDA CASES

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

ADANA, 2019

ÇUKUROVA UNIVERSITY

INSTITUTE OF NATURAL AND APPLIED SCIENCES

DEVELOPMENT OF NATURE INSPIRED ALGORITHMS FOR IDENTIFICATION OF SPINE ON ULTRASOUND IMAGES IN SPINA BIFIDA CASES

Çağlar CENGİZLER

PhD THESIS

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

We certify that the thesis titled above was reviewed and approved for the award of degree of the Doctor of Philosophy by the board of jury on 28/06/2019.

.....	.....	.....
Assoc. Prof. Dr. M. Kerem ÜN SUPERVISOR	Assoc. Prof. Dr. Mustafa ORAL MEMBER	Assist. Prof. Dr. Ahmet AYDIN MEMBER
.....	.....	
Assist. Prof. Dr. İbrahim Cem BAYKAL MEMBER	Assist. Prof. Dr. Esra Saraç EŞSİZ MEMBER	

This PhD Thesis is written at the Department of Institute of Natural And Applied Sciences of Çukurova University.

Registration Number:

Prof. Dr. Mustafa GÖK
Director
Institute of Natural and Applied Sciences

Note: The usage of the presented specific declerations, tables, figures, and photographs either in this thesis or in any other reference without citiation is subject to "The law of Arts and Intellectual Products" number of 5846 of Turkish Republic.

ABSTRACT

PhD THESIS

DEVELOPMENT OF NATURE INSPIRED ALGORITHMS FOR IDENTIFICATION OF SPINE ON ULTRASOUND IMAGES IN SPINA BIFIDA CASES

Çağlar CENGİZLER

ÇUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

Supervisor: Assoc. Prof. Dr. M. Kerem Ün
Year: 2019 , Pages: 88
Jury: Assoc. Prof. Dr. M. Kerem Ün
Assoc. Prof. Dr. Mustafa Oral
Assist. Prof. Dr. Ahmet Aydın
Assist. Prof. Dr. İbrahim Cem Baykal
Assist. Prof. Dr. Esra Saraç Eşsiz

Spina bifida (SP) is a spine defect observed when a baby is still in the womb. The defect is caused by unfinished closure of the embryonic neural column. This thesis proposes a combination of novel nature inspired methods for locating spinal axis on sonograms, where deformation due to spina bifida is observed. The method involves a flocking dynamics based optimization approach for reducing the size of the search space (bones on the sonogram) and a meta-heuristic evolutionary approach, where the sonogram is divided into columns and bone blobs belonging to the spine are classified. Accordingly, a specific genetic structure and fitness function is utilized and conventional genetic operators are applied to search the actual solution. Results show that both approaches are promising and proposed combination of algorithms generally can distinguish the spinal bones from others even if severe morphological defects exist on the sonograms.

Key Words: Spina Bifida, Ultrasound, Machine Learning, Flocking Dynamics

ÖZ

DOKTORA TEZİ

SPİNA BİFİDA GÖRÜLEN ÖRNEKLERDE OMURGA TESPİTİ İÇİN DOĞADAN ESİNLENİLMİŞ ALGORİTMA GELİŞTİRİLMESİ

Çağlar CENGİZLER

ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
ELEKTRİK ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

Danışman: Doç. Dr. M. Kerem Ün

Yıl: 2019 , Sayfa: 88

Jüri: Doç. Dr. M. Kerem Ün
Doç. Dr. Mustafa Oral
Dr. Öğr. Üyesi Ahmet Aydın
Dr. Öğr. Üyesi İbrahim Cem Baykal
Dr. Öğr. Üyesi Esra Saraç Eşsiz

Spina bifida (SB), bebek hala rahimdeyken ortaya çıkan, sık görülen bir omurga kusurudur. Defekt, embriyonik sinirsel kolonun tamamlanmamış şekilde kapatılmasıyla oluşmaktadır. Bu tez bahsi geçen rahatsızlığın teşhisi için destekleyici olabilecek doğadan esinlenilmiş yaklaşımlar içermektedir. Buna göre sürü dinamiğine dayalı bir optimizasyon yöntemi ve kemiklerin diskriminasyonu icin genetik algoritma temellerine dayalı bir yaklaşım önerilmiştir. Önerilen optimizasyon yaklaşımında kuş sürüleri olarak kabul edilen kemik bölgelerinin ait oldukları sınıfa göre kümelenebilmeleri için göç yolları hesaplanmaktadır. Optimizasyon sürecini takip eden tanılama basamağında görüntü dikey bloklara bölünerek kemik adacıklarının buna göre kodlanması sağlanmıştır. Kodlanan adacıklar önerilen uygunluk denklemi sayesinde genetik operatörler uygulanarak sınıflandırılabilmiş, ciddi morfolojik bozukluk söz konusu olsa bile omurga kemiklerinin ayırt edilmesi sağlanmıştır.

Anahtar Kelimeler: Spina Bifida, Ultrason, Makine Öğrenimi, Sürü Dinamiği

GENİŞLETİLMİŞ ÖZET

Spina bifida (SB), bebek hala rahimdeyken ortaya çıkan, sık görülen bir omurga kusurudur. Defekt, embriyonik sinirsel kolonun tamamlanmamış şekilde kapatılmasıyla oluşmaktadır. Serebral palsiden sonra en sık görülen sakatlıklardan birisi olmakla birlikte erken teşhis ve teşhisin doğruluğu bebeğin ve ailenin geleceği açısından kritik önem teşkil edebilmektedir. Buna göre, anomalinin omurga hattı üzerindeki konumunun ve şiddetinin doğru tespiti bebeğin doğum sonrası hayat kalitesi ve doğumun sonlandırılması gibi hayati konularda ailelere kılavuzluk edilebilmesini sağlamaktadır. Hastalığın tespiti doğumdan önce iki geleneksel tarama yöntemiyle gerçekleştirilebilmektedir. Bu yöntemlerden ilki, anne kanında alfa-fetoprotein taramasına dayanmaktadır. İkinci yaklaşım ise omurga anomalisinin ultrason görüntülerinde bir operatör aracılığı ile aranması üstüne kurgulanmıştır. Gebeliğin ikinci üç alık döneminin başlarında sonogramlarda belirginleşmeye başlayan vertebral ve kafatası ossifikasyonları olası bir anomalinin tespitini mümkün kılmaktadır. Teşhis esnasında ilgili anatomik kısımlar uzman operatör tarafından sagital, koronal ve enine eksenlerde gerçek zamanlı ve görsel olarak incelenmekte, olası anomalinin varlığı ve konumu bu sayede tanılanmaktadır. Testin görsel incelemeye dayalı doğası uzman operatörlerin öznel kararları nedeni ile hastalığın varlığı veya konumu her zaman doğru teşhis edilemeyebilmektedir. Bu noktada bilgisayar destekli teşhisin testin başarısını arttırmaya yönelik katkıda bulunabileceği muhtemeldir. Buna göre fetal sonogram görsellerinde omurgayı oluşturan kemik hattının sınıflandırılabilmesi, hastalığın otomatik teşhisi açısından önemli bir adım olarak görülebilir. Bu çalışmada ilgili anomalinin görüldüğü sonogram örneklerinin doğadan esinlenilmiş algoritmalar aracılığı ile sınıflandırılarak yüksek deformasyon görülen durumlarda bile omurgayı oluşturan

kemiklerin diğerlerinden ayırt edilebilmesi amaçlanmıştır. Olası bir bilgisayar destekli teşhis sisteminin kritik bir adımı olarak görülebilecek bu sınıflandırma probleminin çözümünün hastalığın teşhisinde önemli rol oynaması muhtemeldir. Buna göre çalışmada örnekler üç kısımda incelenmiştir. İlk kısımda örnekler bir önişlemden geçirilerek kemikleri temsil eden adacıklar ikilik görüntüler olarak elde edilmiş ve adacıkların çeşitli geometrik, biçimsel ve konumsal özellikleri incelemeye tabi tutulmuştur. Bu incelemede kemik adacıklarını temsil edecek en belirgin özelliklerin olası bir konvansiyonel deterministik sınıflayıcı için ayırt edici olup olamayacağı sınanmıştır. Sınama iki farklı deney düzeneğinde gerçekleştirilmiştir. Bu düzeneklerden ilkinde özelliklerin ayırt ediciliği istatistiki bir yaklaşım olan T-testi ile sınanmıştır. İkinci düzenekte ise yaygın kullanım bulmuş iki farklı eğitimsiz kümeleme algoritması aracılığı ile bütün özellik kombinasyonları test edilerek kümeleme algoritmalarının başarısı F-skoru cinsinden karşılaştırılmıştır. çalışmanın ilk bölümü göstermiştir ki konvansiyonel sınıflama yöntemlerinin olası şekilsel ve geometrik özellikler ile omurga kemiklerini ayırt edebilme ihtimali düşüktür ancak görülmüştür ki omurga kemiklerinin konumsal ilişkileri omurga aidiyetine dair bilgi içermektedir ve ayırt edicilik üzerinde etkisi olabilir. çalışmanın ikinci kısmında bahsi geçen sınıflama probleminin kompleksitesinin düşürülebilmesi için bir optimizasyon yöntemi önerilmiştir. önerilen optimizasyon yöntemi kuş sürüsü dinamiğine dayanmaktadır. Buna göre aynı sınıfa ait kemiklerin birbiri ile bağlanarak sınıflandırılması gereken toplam kemik bölgesi sayısının düşürülmesi bu sayede arama uzayının sınıflama öncesi optimize edilmesi hedeflenmiştir. Bağlama işlemi kuş sürüsü dinamiğini temsil eden üç ana ilke aracılığı ile gerçekleştirilmektedir. Bu ilkeler çarpmaktan kaçınma, merkezlenme ve hizalanma olarak listelenebilir. Buna göre kemik adacıkları kuş sürüleri olarak kabul edilerek her sürü merkezi ile temsil edilmiş ve ilgili sürüler merkezlerinden bağlanmıştır. Ayrıca her sürünün algı

mesafesi sınırlandırılmış, algı sınırı içindeki bağlanmaya en müsait sürü, komşu sürülerin yaptığı önceki bağlantılar ve mesafe göz önünde bulundurularak seçilmiştir. Bunlara ek olarak belirli bir çap içinde yalnız kalmış kemik adacıkları yok sayılarak öldürülmüş (silinmiş) bu işlemler sonucunda sınıflamaya tabi tutulması gereken kemik bölgesi sayısı minimuma indirilmiştir. Sonuçlar göstermektedir ki uygulanan optimizasyon yöntemi yüksek başarı ile aynı sınıftaki kemikleri birbirine bağlayabilmekte, efektif bir biçimde arama uzayını küçültebilmektedir. Buna göre uygulanan kodun örneklerin %70'inde sınıflar arası bağlantı hiç(arzu edilmeyen tipte bağlantılar) yapmadığı, bu tip bağlantıların görüldüğü örneklerde ise minimum sayıdaki bu hatanın omurga kemiği grubu üstüne oturtulan bir polinomu kabul edilebilir sınırlarda oynattığı gözlenmiştir. örneklerin hiç birinde hatalı öldürme yani omurga kemiğine tekabül eden bir kemiğin silinmesi görülmemiştir. Uygulanan deneysel meta-sezgisel ideal sınıflayıcı aracılığı ile optimize edilmiş ve edilmemiş örnekler karşılaştırılmış, sınıflayıcının optimize örneklerle daha efektif çalışma eğiliminde olduğu tespit edilmiştir. çalışmanın son kısmında örneklerin kolonlara bölünmesine dayalı bir genetik yaklaşım uygulanarak omurga kemiğini temsil etme ihtimali en yüksek kemikler zincirinin tanılanması amaçlanmıştır. Buna göre popülasyonları oluşturacak her bir bireyi temsil etmek üzere yeni bir genetik mimari tasarlanmıştır. Popülasyonları her yeni kuşakta olası doğru cevaba yaklaştırması beklenen genetik operatörler bu yeni mimari üstünde uygulanmıştır. Uygulanan genetik mimari, sütunlara bölünmüş örnek üzerinde her sütuna düşen kemik adacığı parçalarını bir merkez ile temsil etmekte, bu merkezlerin her birini sütunun seçkileri olarak kabul etmektedir. Bu sayede oluşturulan her birey sütuna başına bir seçki kullanarak bir seçkiler zinciri yaratmakta, yaratılan zincirin omurgayı temsil etme yeteneği çalışmada önerilen uygunluk denklemi aracılığı ile yargılanmaktadır. Bu noktada eklenmelidir ki her birey oluşumunu tamamladıktan sonra aynı kemik

adacığına ait seçkilerin birbirini takip etmesi sağlanmıştır. Bu, bireyin soldan sağa taranması aracılığı ile gerçekleştirilmektedir. Tarama sonucunda bir sütunu temsil etmesi beklenen seçki bir önceki sütununun seçimine göre tercih edilmiştir. çalışmada konvansiyonel genetik algoritma yaklaşımında kullanılan operatörlere ek olarak çeşitliliği olabilecek en verimli şekilde sağlamak adına bir enerji minimizasyonu operatörü uygulanmıştır. Buna göre iki seçki arasındaki mesafe ve açı (uzunluk cinsine çevrilmiş hali) toplamının o bölgedeki enerji miktarını temsil ettiği kabul edilmiş, bireylerin en yüksek enerji taşıdığı bölge tespit edilerek enerji minimize edilecek şekilde ilgili sütunlarda seçkiler yenilenmiştir. çalışmada yaratılan bireylerin omurgayı temsil etme başarısını yargılayabilmek için üç ana bileşeni olan bir uygunluk fonksiyonu önerilmiştir. Buna göre seçkiler arasındaki toplam mesafenin, ortalama kıvrımın ve seçkilerin bir polinom üzerine yerleşebilme başarısının ağırlıklı kombinasyonu uygunluk denklemi olarak kabul edilmiştir. çalışmada uygulanan genetik algoritma üç farklı ideal uygunluk denklemi (F-skoru, öklid uzaklığı cinsinden hatası, polinom üstüne oturma başarısı) ile test edilmiş ayrıca çalışmada önerilen uygunluk denklemi enerji minimizasyonu ile, kuş sürüsü yaklaşımı kullanılarak optimize edilmiş örneklerle ve optimize edilmemiş örneklerle test edilmiştir. Sonuçlar göstermektedir ki önerilen genetik yaklaşım bireyleri kuşaktan kuşağa doğru sonuca taşıyabilmektedir. Ayrıca önerilen uygunluk denkleminin doğruya yakın yargıda bulunabildiği görülmüştür. Nümerik sonuçlar enerji minimizasyonu ve kuş sürüsü dinamiğine dayalı optimizasyon yaklaşımlarının başarıyı arttırma eğiliminde olduğunu göstermektedir. ümit vaat eden karşılaştırmalı sonuçlara dayanarak çalışmada önerilen metotların ileri derecede deformasyona uğramış örneklerde bile omurgayı oluşturan kemikleri diğerlerinden ayırt etmede efektif olabileceğini öne sürmek mümkündür.

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor, Assoc. Prof. Dr. M. Kerem Ün for his valuable suggestions, patience, motivation, and immense knowledge. I am highly thankful for his guidance which helped me greatly in research and in writing of this thesis.

A very special gratitude goes out to Prof. Dr. Selim Büyükkurt for bringing this important topic to my attention. His support during the data collection phase and guidance in the continuation of this research has been invaluable.

I would like to express my appreciation to Assoc. Prof. Dr. Mustafa Oral and Assist. Prof. Dr. Ahmet Aydın for their insightful comments and encouragement during my research. Assist. Prof. Dr. İbrahim Cem Baykal and Assist. Prof. Dr. Esra Saraç Eşsiz have been kind enough to be jury members in my dissertation committee and I am thankful to them.

Last but not the least, I would like to thank my parents and my beloved wife who have supported me with their love and encouragement throughout the entire process.

CONTENTS **PAGE**

ABSTRACT I
ÖZ II
GENİŞLETİLMİŞ ÖZET III
ACKNOWLEDGMENTS VII
CONTENTS VIII
LIST OF TABLES XII
LIST OF FIGURES XIV
LIST OF ABBREVIATIONS XVIII
1. INTRODUCTION 1
1.1. Spina Bifida 1
1.2. Machine Learning in Automated Analysis of Ultrasound Images . . 3
1.3. Meta-heuristic Applications of Computer Aided Medicine 6
1.3.1. Meta-heuristic Applications in Various Branches of Medicine 6
1.3.2. Meta-heuristic Applications in Radiology and Ultrasound . 7
1.4. Motivation and Contributions of the Thesis 9
2. MATERIAL AND METHODS 13
2.1. Preparation of Data-Set from Raw Ultra-Sound Images 13
2.2. Evaluation Methodology 15
2.3. Definition of Search Space 15
2.3.1. Feature Based Characteristics of Search Space 17
2.3.1.1. Area Based Features 17
2.3.1.2. Shape and Morphology Based Features 18
2.3.1.3. Location Based Features 21
2.3.1.4. Significance of the Features 22

2.4. Flocking Dynamics Based Search Space Optimization 29
2.4.1. Centering and Neighbouring 30
2.4.2. Concept of Perception 31
2.4.3. Fatal Distance 32
2.4.4. Alignment 32
2.4.5. Search Space After Flocking Based Optimization Approach 33
2.4.6. Meta-heuristic Evaluation of Flocking Based Optimization Approach 36
2.4.6.1. Evolution Mechanism 37
2.4.6.2. Fitness Criteria 38
2.4.6.3. Performance of the Meta-heuristic Classification . 39
2.5. Genetic Algorithm Based Discrimination of Flocks 41
2.5.1. Implemented Genetic Structure and Individuals 41
2.5.1.1. Propagation of Connections through Chromosomes 42
2.5.2. Implemented Evolutionary Algorithm 44
2.5.2.1. Population Diversity 46
3. RESULTS 49
3.1. Performance of Euclidean Distance Based Fitness 50
3.2. Performance of F-Score Based Fitness 52
3.3. Performance of the Cubic Spline Interpolation Based Fitness 53
3.4. Performance of Proposed Fitness Criteria 55
4. DISCUSSION 63
4.1. Discussion 63
4.2. Conclusion and Future Work 66
REFERENCES 69
CURRICULUM VITAE 75

APPENDIX 77

LIST OF TABLES **PAGE**

Table 2.1. Average values of the three area based features 17

Table 2.2. Average area and standard deviation values for spine and non-spine bone regions 18

Table 2.3. Average values of shape and morphology based features with standard deviation values 20

Table 2.4. Average values of distance to nearest neighbour with standard deviations 22

Table 2.5. Names and numbers of the examined features 23

Table 2.6. Features given on Table 2.5 sorted according to their significance with the T-test approach, where Rank 1 indicates the best and Rank 6 indicates the worst class separability 24

Table 2.7. Best resulting combinations of features given on Table 2.5 classified with K-means for ten samples 27

Table 2.8. Best feature combinations obtained with FCM for ten samples . 29

Table 2.9. Percent reduction of search space given for each sample with false connections and false kill rates 34

Table 2.10. Integrated RMSE values of hand drawn curves and estimated curve with respect to the ground truth polynomial 36

Table 2.11. Best F-scores presented for several generations 39

Table 3.1. Euclidean distance based golden fitness results (average of ten repetitions) 51

Table 3.2. Number of runs where perfect classification is achieved with Euclidean distance based golden fitness 51

Table 3.3. F-Score based golden fitness test results (average of ten repetitions) 53

Table 3.4. Number of runs where perfect classification is achieved with F-score based golden fitness 53
Table 3.5. Spline fitness based golden fitness test results (average of ten repetitions) 54
Table 3.6. Number of runs where perfect classification is achieved with spline fitness based golden fitness 55
Table 3.7. RMSE values for the curves drawn by the three experts 56
Table 3.8. Weight combinations resulting to the lowest average order error given for ten samples. 58
Table 3.9. Initial test parameters are given 58
Table 3.10. Average RMSE values of best fitting individuals for ten repetitions over 100 generations 59
Table 3.11. Percentage of best individuals that attain the best possible solution 60
Table 3.12. Minimum and maximum RMSE success over generations . . . 60

LIST OF FIGURES **PAGE**

Figure 1.1. a) A sonogram example for a healthy spine formation. (b) An SB case where spine is highly deformed 2

Figure 2.1. Fundamental steps of proposed approach 13

Figure 2.2. Stages of data visually depicted 14

Figure 2.3. Sample regions from data set and contours of extracted binary masks. Filled blobs indicate spine bones. 16

Figure 2.4. Effect of increasing eccentricity. Lower eccentricity means higher similarity to a perfect circle. 19

Figure 2.5. Equivalent diameter feature visualized on two sample bone regions. The circle equivalent to the region is shown in darker color. 19

Figure 2.6. Distance to nearest neighbour visually depicted 21

Figure 2.7. Angle defined using the nearest neighbours of a blob. Darker blob is the subject of measurement. 22

Figure 2.8. Block diagram of the utilized exhaustive search algorithm . . . 25

Figure 2.9. Two flocks (i.e. bone blobs) are connected through their centroids (indicated with circles) with a straight line to form a single object. 31

Figure 2.10. Perception zone for a sample flock. There are three immediate neighbours that the flock is able to detect. 31

Figure 2.11. Effect of fatal distance. A sample region before (a) and after (b) eliminating isolated flocks 32

Figure 2.12. Choice of best neighbour according to orientation. A search is done for red coloured flock. Blue arrow is the previous connection and green arrow indicates best connection for red coloured flock. 33
Figure 2.13. Curves hand drawn by experts shown on sample regions 35
Figure 2.14. Output of flocking shown on different samples. Filled regions are bones on the spine axis where migration paths are shown with lines. 36
Figure 2.15. Flow chart of the implemented conventional genetic algorithm . 38
Figure 2.16. Classification performance of genetic setting is given in different shades of gray for each sample. Best result of ten repetitions is given for each generation 40
Figure 2.17. Columnar genetic structure graphically depicted. (a) Previously detected bone regions (darker contours indicate spine regions). (b) Sample region is divided into columns. (c) Each centroid is marked as a selection alternative belonging to the related column. 41
Figure 2.18. Three different candidate solutions on a five column sample. Selections taken from each column are connected with a line. . . . 42
Figure 2.19. Effect of propagation is visualized on three different candidate solutions where sample region is divided into five columns . . . 43
Figure 2.20. The applied evolutionary approach given as a block diagram . . 44
Figure 2.21. Sample five column individuals as examples of low and high fitness individuals 45
Figure 2.22. From left to right, first frame indicates highest energy points. Best alternative to minimize the energy is shown in the second frame. 47

Figure 3.1. Active region within the range of actual spine bones is shown between gray regions. Bold contours indicate the spine bones. . 49
Figure 3.2. Euclidean distance based error 50
Figure 3.3. F-score based fitness concept explained on a four column search space. There are six labelled objects which are indicated on ground truth visual. Conversion table shows how individuals and ground truth are coded into binary arrays to calculate F-score . . 52
Figure 3.4. A cubic spline, drawn on centroids of actual spine bones. 54
Figure 3.5. Manually drawn curves on a sample region and the extracted points from these curves used in error measurement 56
Figure 3.6. Performance of the proposed algorithm demonstrated visually on six sample regions. Continuous curve, circles and dotted line indicate the polynomial fitted to selected blobs, centroids of the selected blobs and line segments connecting the blob centroids, respectively. 61

LIST OF ABBREVIATIONS

SB	Spina Bifida
CAD	Computer aided diagnosis
NBC	Naive Bayes classifier
SVM	Support vector machine
ANN	Artificial neural networks
BPNN	Back propagation neural networks
CNN	Convolutional neural networks
GA	Genetic algorithm
MRI	Magnetic resonance imaging
CT	Computed tomography
FCM	Fuzzy C-means
RMSE	Root mean square error

1. INTRODUCTION

1.1. Spina Bifida

Spina bifida (SB) is a developmental fetal spine defect. The defect is caused by incomplete closure of the embryonic neural column which results to a split deformation on spinal anatomy (Konur et al. 2015). (The name of this pathology translates as ”split spine”.) It is the most commonly encountered *disability* after cerebral palsy (Copp et al. 2015). Many advances have been made in past 75 years in preventing and diagnosing this non-lethal fetal anomaly (Trudell and Odibo, 2014). SB can be diagnosed with two conventional screening methods before birth. The first method involves screening mother's blood for alpha-fetoprotein and the second one is based on observing the ultrasound images to identify any existing anatomical abnormality of spine (vertebral ossifications of the fetus becomes detectable on sonograms in the early second trimester) or deviations in intracranial anatomy possibly caused by spina bifida (Buyukkurt et al, 2013). It should be noted that, while three-dimensional ultrasound and magnetic resonance imaging are emerging methods for characterization of the open spina bifida spinal lesion, ultrasound is still the gold standard diagnostic tool for spina bifida (Trudell and Odibo, 2014). In the ultrasonic spine examination, the clinician visually looks for signs of an open spine. It is possible to visually observe fetal spine in real time in sagittal, transverse and coronal planes for diagnosis (Figure 1.1).

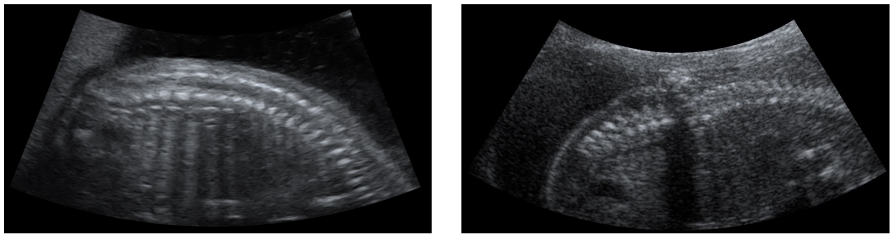


Figure 1.1. a) A sonogram example for a healthy spine formation. (b) An SB case where spine is highly deformed

The diagnosis, as well as the location of the defect, is crucial for parental decision making about possible interventions. It has been shown that if the upper pole of the defect is below the fourth lumbar vertebrae, the estimated prognosis is always better (Bruner and Tulipan, 2004). It has been recently demonstrated that 3D ultrasonography is effective in the diagnosis of the defect and the detection of the upper pole of the lesion (Ozaras, 2004; Buyukkurt et al., 2013). However, 2D or 3D ultrasonography requires a certain level of expertise. (A recent paper underlines the adjacent role of biochemical screening in the diagnosis of fetal spina bifida (Racusin et al., 2016)). Hence, a computerized support system for detecting the location of the spine on the ultrasound images may be helpful in the diagnostic process. For this, the system should be able to automatically identify the spine axis on the sonogram, which in itself would be a challenging task. Since the ultrasound is taken early in the pregnancy, the bony tissues are still in their developmental stage and not very distinguishable on the images due to the distortions and noise on the sonograms (Tolay et al., 2009). Consequently, conventional morphological image processing methods are likely to be insufficient for effectively locating the bones of the spine axis and a specific algorithm may be more appropriate for the task. In this study, to achieve this task, a novel computer aided registration method is proposed that itself involves

task-specific nature inspired mechanisms.

1.2. Machine Learning in Automated Analysis of Ultrasound Images

Ultrasound is one of the safest and most efficient medical imaging modalities for diagnosis of a number of diseases as well as prenatal screening of spinal anomalies (Drukker et al., 2002). Machine learning based methods are frequently utilized in the literature for coping with automated classification problems on ultrasound images. One of the main motivations behind these studies is the fact that manual reading of ultrasound images normally requires years of experience (Huang, et al., 2018) and any system that would facilitate this process is a welcome development in the area. It has been shown in the literature that computer aided approaches improves the accuracy of final diagnosis, reduces the required labour and decreases the amount of expertise required to complete the diagnosis task (Huang, et al., 2018).

It should be noted that a computer aided system developed for automated analysis of ultrasound images is aimed at locating specific lesions on the images and may not be a direct diagnosis tool. In other words, the main function of such a system is to reduce the observational burden. On the other hand, it is also possible that a specific system may be developed to help in or improve the diagnosis, i.e. help the clinician to take the final decision (Mansoor et al., 2015).

Most *computer aided diagnosis* (CAD) systems for ultrasound images consist of four fundamental stages no matter if they are developed for the purpose of locating or classifying the lesions. These stages are preprocessing of images, segmentation of lesions, feature extraction and classification. The study presented here also proposes a methodology which includes preprocessing, optimization and classification stages.

Many machine learning based classifiers, used in CAD systems, have been developed to work on ultrasound images. One of the most basic classifier is the *linear*

classifier, where linear combinations of features are grouped into several categories. In one study, linear discriminant analysis is utilized with four features from 400 samples of breast sonograms (Horsch, et al., 2002). The conclusion has been that the linear classifier is potentially effective for diagnosis of malignancy from ultrasound data. In another study, Fisher Linear Discriminant has been used with a series of other machine learning approaches (like support vector machines(SVM), K-Nearest-Neighbours and Bayes) to classify the different cataract severities in B-Scan ocular ultrasound images (Caixinha, et al., 2016).

Another widely used and relatively simple to utilize approach is based on Bayes' theorem, where the probability of an event is estimated according to prior knowledge of conditions which are likely to be related to the event (D'Agostini, 1995). A common statistical classifier is naive Bayes classifier (NBC). It has been also used to classify breast sonograms to distinguish healthy and cancerous tissue (Yang et al., 2012).

SVM is a popular machine learning methodology designed to form a hyperplane for creating groups in feature space. A study which includes an SVM based classifier trained to locate the heart in ultrasound images has been realized (Maraci et al., 2017). In another study, a hybrid method is presented where a combination of evolutionary computing and SVM is employed for tumor classification. The most significant features from ultrasound images are determined and used to classify the segmented tumor as benign or malignant (wu and friends).

The decision tree approach is also a machine learning method used in the analysis of ultrasound images for classification purposes. The feature space is divided into subsets which are further divided into sub-subsets to form a flowchart-like structure. For a given input data, a sequence of branches of the tree are followed to reach a point where a decision is made about the analysed item. The effectiveness of this

straightforward algorithm in a CAD system is put to test, when seven grayscale features from 232 normal and 294 pathological thyroid ultrasound images are extracted in order to identify samples with Hashimoto thyroiditis (Acharya, et al., 2014). In this study, the decision tree approach has classified the features with 73.1% accuracy, 79.3% sensitivity and 65.2% specificity. In another study, a probabilistic boosting-tree based classifier mechanism is utilized for automated segmentation of the object of interest from the background in ultrasound images. The novel method is able to learn the appearance of the object of interest, which is achieved by training of a discriminative constrained probabilistic boosting tree. Fully automatic measurement of head circumference on ultrasound images has been achieved with this algorithm (Carneiro et al., 2007).

Artificial neural network (ANN) is a popular supervised machine learning technique, developed to mimick the functioning of neurons and is nature-inspired in that sense. There are many known variants of ANN models, some of which have been utilized in connection with ultrasound images. A least-mean squares back-propagation neural network was utilized for classification of ophthalmic artery stenosis (Guler and Ubeyli, 2003). The researchers have examined ophthalmic artery Doppler signals of 105 subjects, 48 of which diagnosed positive for ophthalmic artery stenosis. The classification approach is reported to detect ophthalmic artery stenosis with up to 90.6% accuracy. In another study, Baumgartner *et al.* have proposed a computer aided screening system which is based on convolutional neural networks (CNN) (Baumgartner et al., 2017). Their automated system is able to automatically detect 13 fetal standard views in 2D ultrasound data. Moreover, the network architecture they have implemented is designed to operate in real-time and can detect standard views with an average F-score of 0.798.

1.3. Meta-heuristic Applications of Computer Aided Medicine

In this study, a meta-heuristic approach is proposed for locating deformed spinal column in ultrasound samples. Meta-heuristic is a ”higher level” heuristic method to find a quick solution to an optimization problem with minimal complexity and computation without guaranteeing feasibility or optimality (Beheshti and Shamsuddin, 2013). Meta-heuristic approaches are implemented to enhance the capability of exploring and exploiting of the search space in heuristic algorithms. One of the main components of finding an optimum through exploring the search space is the diversification of the search space via randomization. Increased diversity would prevent meta-heuristic algorithms from being trapped at a local optima. In addition to diversification, selection of the best solution among the possible candidates is crucial for converging to an optimum (Yang, 2011). In contrast to supervised machine learning methods, meta-heuristic search approach typically do not require a training stage (Yang, 2010).

1.3.1. Meta-heuristic Applications in Various Branches of Medicine

One branch of medicine where genetic algorithm (GA) based methods are of interest is oncology. Diagnosis of cervical cancer with Pap smear test is often done with a CAD system and different algorithms have been developed for automatic detection one of the most desired problems to be solved on that spot. One study proposed in 2009 has presented an effective GA scheme to determine the best subset combination of previously extracted 20 features from cervical cell images (Marinakis et al., 2009). In one particular study that has utilized GA for automated cancer diagnosis, Pena-Reyes and Sipper have proposed a combination of fuzzy system and GA. They have examined their approach on fine needle aspiration samples of the Wisconsin breast cancer diagnosis database. A total of nine visually assessed characteristics

are extracted as features defining the shape and morphology of the cells. They have utilized GA to optimize the parameters of their fuzzy system. It should be noted that, as a result of 120 evolutionary runs, classification performance of their evolutionary approach exceeds 94.5% (Pena-Reyes and Sipper, 1999). GA is also utilized for segmentation tasks of oncology problems. Cengizler and Un have proposed a GA based nuclei segmentation approach. Their methodology is based on a GA based clustering implementation where Calinski-Harabasz internal cluster validation criteria is measured as fitness (Cengizler and Un, 2017).

Cardiovascular medicine is another medical branch where meta-heuristic methods are used on radiological images. In one exemplary study, a combination of GA and SVM is proposed for electrocardiogram arrhythmia classification. The researchers have utilized GA to optimize SVM by tuning the parameters of SVM. Their results show that, their optimized version of SVM is capable to distinguish four types of arrhythmias with 93% accuracy (Nasiri et al., 2009). Additionally, a combination of GA and wavelet transform is proposed in 2011 for denoising ECG data. In this approach, GA is similarly utilized for selecting best parameters for wavelet denoising parameters (El-Dahshan, 2011). It has been reported as a result, that wavelet transform optimized with GA functions better for denoising ECG signals. Poli and friends proposed a similar GA based optimization approach where GA is utilized to find the best parameters for linear, nonlinear polynomial filters and adaptive maxima detector. Their approach optimizes the QRS detection on ECG signals (Poli et al., 1995).

1.3.2. Meta-heuristic Applications in Radiology and Ultrasound

Evolutionary methods are meta-heuristic, nature inspired algorithms that use principles of genetics for solution search (Ghaheri, et al., 2015). Accordingly, several operators like mutations and cross-over are inspired from genetic laws for explo-

ration and exploitation of the search space. When using a genetic algorithm (GA), a candidate solution is accepted as an ”individual” in a solution ”population”. The initial generation of the randomly created individuals form the first generation and pass their properties to next generations of the population by ”reproduction”. ”Fitness” is a quantitative measure defined for the closeness of an individual to the optimal solution. Solutions that are not ”fit” enough are systematically eliminated every generation. That way candidate solutions become fitter every generation. The iterative procedure converges to the optimal solution if the scheme works properly (Holland, 1992). Unsupervised segmentation of three different tissue regions from ultrasound samples is performed by employing a meta-heuristic feature selection approach (Jiang, 2001). Both the segmentation and the feature selection tasks are achieved in an unsupervised manner in this study.

There are many variants of genetic methods in the literature. Genetic algorithms are widely utilized in medical applications, in particular those related with radiology. Automated detection of the organ boundaries and morphological analyses on radiography images are mostly accomplished with meta-heuristic search methods including GA.

An algorithm has been developed for detecting thin edges in data acquired by magnetic resonance imaging (MRI), computed tomography (CT), and ultrasound modalities, where the researchers have added several enhancements for optimizing the performance of the conventional GA approach (Gudmundsson, et al., 1998). Another study has proposed a hierarchical genetic algorithm for segmenting multispectral human brain MRI. The study includes a combination of hierarchical genetic algorithm with a fuzzy learning-vector quantization network. The utilized evolutionary approach have reportedly outperformed other clustering methods in specificity (Yeh et al., 2008). One study proposed by Dehmeshki *et al.* involves the detection

of nodules with spherical elements on CT scans (Dehmeshki et al., 2007). They have proposed a shape-based genetic algorithm template-matching method. The nodules have been detected with a success rate of 90% in this study.

It is also possible to address X-ray images as subject to GA based analysis. The classic scheme of the GA has been improved by introducing novel genetic operators in one study. The efficiency and robustness of the approach is demonstrated on some artificially created and real X-ray data(Ulyanenkov and Sobolewski, 2005). In another study, wavelet analysis and GA have been employed for mammogram mass segmentation (Pereira et al., 2015), which is another radiological example where GA is put to use.

In addition to X-ray based studies, one study focused on ultrasound is proposed by Mnigot and friends which they have presented a methodology to maximized the contrast resolution without a priori information via automatically found optimal imaging wave (Mnigot and Girault, 2016). They have utilized GA for limiting the search for optimal imaging wave.

A combination of GA and active contour approach is utilized by Talebi and friends in 2011 to cope with segmentation tasks of ultrasound images. They have used GA approach to minimize the energy of contour belonging to the object desired to be segmented they have claimed that their approach outperformed the conventional active contour approach (Talebi et al., 2011)

1.4. Motivation and Contributions of the Thesis

In this thesis, two different implementations of nature inspired algorithms are simultaneously used to detect the spinal axis on sonograms obtained from fetus with diagnosed spina bifida.

The sonograms after segmentation involves an abundant number of binary

(black and white) blobs, most of which correspond to neither spine nor even another bone. All possible combinations of blobs (of bone or other) represent the search space, and the combination that includes nothing but all spine bones is the sought optimal solution. In the first stage of the methodology, the search space is optimized by combining certain adjacent blobs, that are likely to belong to the spine, into a single item. This is accomplished by accepting automatically each extracted bone blob as a bird flock and calculating its ”migration path” to other flocks for clustering. The large size of the search space is a problem characteristics that has a direct effect on the classification ability of a meta-heuristic machine learning method (Kuri-Morales and Rodrguez-Erazo, 2009). Accordingly, by combining blobs into single items, we effectively decrease the total number of blobs and total number of their possible combinations, thus reducing the size of our search space (Srinivas and Patnaik, 1991).

Second stage of the methodology is implemented for the discrimination of blobs (or combined blobs of the first stage) belonging to the spine in the optimized search space. A specific GA architecture is developed for finding the best combination of blobs representing the spine. The ultrasound image at hand is divided into columns in that approach. A bone region is assigned to each columns either as a whole or partially. Hence, each individual of each generation is formed by taking one bone blob (or combined bone blob) either as whole or partially from every column and connect them together into a chain. The spine does not have an arbitrary shape. Consequently, we require the curve that is fit to this chain to resemble a polynomial curve that can represent the spine shape. The fitness function is accordingly defined in terms of the arc length and curvature of this fitted curve. The GA is expected to find the best individual (i.e. chain of blobs) among those that form spine-like formations.

To the best of our knowledge, there is only one published work where the spine axis is identified from ultrasound images with image processing approaches (Tolay

et al., 2009). The current work is novel not only because it is the first of its kind focused on spine identification on fetal sonograms, but also because its algorithm is specifically tailored for spina bifida cases.

This thesis is organized as follows: Chapter 2 introduces the utilized data set, the search space optimization inspired from flocking dynamics approach and the evolutionary algorithm for discrimination of spinal axis. Chapter 3 presents the results. The results along with a thorough performance evaluation are discussed in Chapter 4 and concluding remarks are given in Chapter 5.

2. MATERIAL AND METHODS

Automated detection of spinal axis on fetal ultrasound samples is the objective of this study. Accordingly, two nature inspired algorithms are implemented as stages of an automated system where data originates from ultrasound. The stages of the study is graphically summarized in Figure 2.1.



Figure 2.1. Fundamental steps of proposed approach

Automation starts with extraction of bone regions from raw ultrasound images by the application of a series of filters, which is followed by search space reduction stage inspired by flocking dynamics. Final stage of the proposed methodology is evolutionary discrimination of bones belonging to the spinal axis.

2.1. Preparation of Data-Set from Raw Ultra-Sound Images

The methodology is tested on sagittal plane images of ten different fetal spines, diagnosed with SB with varying degrees of severity by an expert obstetrician. Sample images are acquired with a General Electric 4k ultrasound imaging device (General Electric, Boston, MA) by the same expert. Raw ultrasound samples are given in Appendix A.

All sample images are processed and filtered during the preprocessing stage. That stage involves basic filtering operations that suppress noise and convert raw images into binary masks where each bone region appears as a blob. It should be noted that fetal spine bones to be located are among these blobs.

The process starts by blurring gray-scale raw images with Gaussian filter. Av-

erage size of the applied Gaussian filter is 29x29 pixels for all sample images. Consecutively, the blurred images are processed with top-hat filter with a disk-shaped structuring element, where average radius of the disk equals 16 pixels for all samples. Resulting images are thresholded in order to generate a binary image. Average threshold intensity value for all images is 19.

Applied operations remove high frequency noise from the samples and allow us to deal with only bone-like structures on the image (Tolay et al., 2009). Finally, a blob size filter is applied to the image for removing small objects (Figure 2.2).

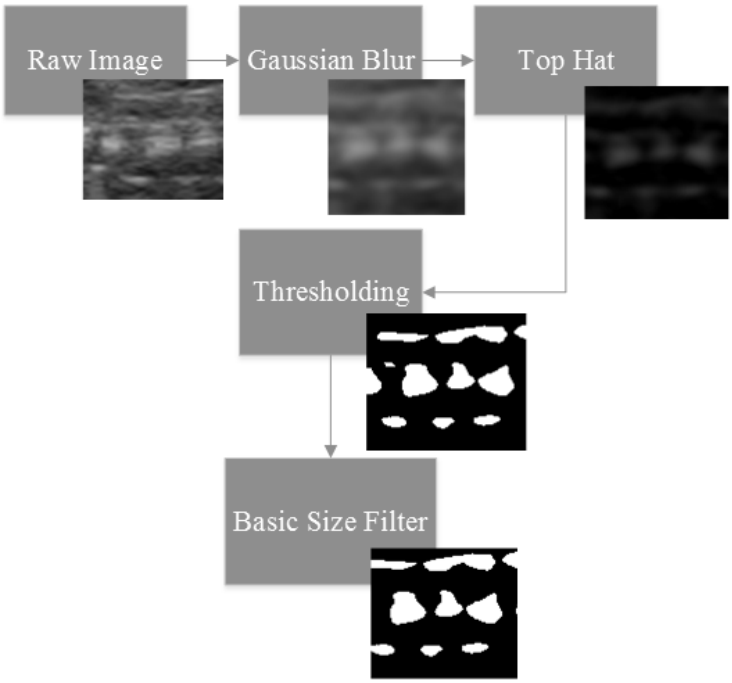


Figure 2.2. Stages of data visually depicted

Resulting images are binary masks (i.e. black and white images) where pixels associated with bone regions are white. To establish the ground truth, bone regions

belonging to the spine are marked on all images by an expert.

2.2. Evaluation Methodology

After basic filtering operations, an expert obstetrician has identified all spine bones in each of the ten sample images and marked them. The images where spine bones are correctly identified represent the *ground truth*. For evaluation purposes, a fifth order polynomial is fitted using the centroids of the spine regions on each image. We will refer to this polynomial as the *ground truth polynomial*, since it is the continuous representation of the actual spine. A fifth order polynomial is adequate to capture the shape of the actual spine without going into oscillatory behavior referred to as *Runge phenomenon* that occurs with higher order interpolations.

Simultaneously, three other expert obstetricians, who has not seen the filtered images, draw the spine axis by hand on each raw sonogram. This is achieved via a digitizer tablet with pen whose movement resolution is 2540 lpi. The ground truth, the ground truth polynomial as well as the curves drawn by the experts are utilized for evaluation purposes which will be detailed where necessary in the upcoming sections.

2.3. Definition of Search Space

A total of 682 bone regions on ten sample images are examined and discriminated in this study. 168 of these have been identified as spinal bones. Identification is made by an expert on binary blobs, which allows us to examine the search space in two clusters as spine and non-spine bones. The average number of bone regions per image is found to be 68.2, 27.7%-49.6% of which corresponds to a spine. Some of the samples and extracted masks are shown in Figure 2.3.

Preprocessing stage results to a binary bone region mask where each separate blob represents a bone. It is known that, spinal axis is most likely located on some

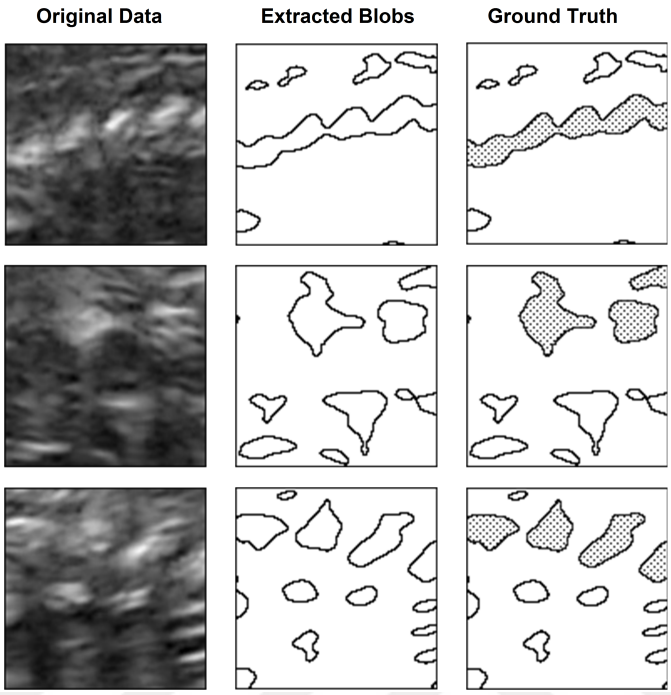


Figure 2.3. Sample regions from data set and contours of extracted binary masks. Filled blobs indicate spine bones.

of these bone representing blobs. Each possible combination of bones is a candidate solution to the problem and hence, an element of search space. Clearly, it is not known *a priori* how many spine bones are visible and distinguishable on an image. It is also not known how many connections are required to best identify the spine. Accordingly, any algorithm aimed at locating the spine should be able to eliminate bone combinations that are too far from reality and to identify those blobs that constitute the spine. Several possible features that characterize the search space are examined in this study as explained in the coming sections.

2.3.1. Feature Based Characteristics of Search Space

The above stated problem is an object classification problem and the first idea that comes to mind is to utilize conventional classification schemes to find a solution to this problem. Several prominent features are examined to see if the introduced problem would be an object to conventional deterministic methods using these features.

2.3.1.1. Area Based Features

Area distribution of the bones is one of the features that characterizes the search space. Accordingly, several area based measurements are presented to outline the search space characteristics. These are:

- Total number of pixels representing bone regions.
- Area of bounding box, framing the total area where bone regions are located.
- Percentage of number of spine bone pixels to number of total pixels inside the bone regions.

Average values of three area based features for ten samples are given in Table 2.1.

Table 2.1. Average values of the three area based features

Total Bone Area	**Bounding Box Area**	**Spine Area Percent**
26364.10	316650.4	36.07%

According to measurements presented in Table 2.1, it can be asserted that bone regions are evenly distributed throughout the image samples. However area of extracted blobs belonging to spinal axis is less than %5 of bounding box area for all samples.

In addition to cumulative measurements, the same quantities are separately evaluated for spine and non-spine bones. Area averages are given for each of these two group in Table 2.2.

Table 2.2. Average area and standard deviation values for spine and non-spine bone regions

Spine Bones		**Other Bones**	
Area	STD	Area	STD
529,39	614,61	339,88	797,51

Coefficient of variation (defined as the ratio of standard deviation to mean value) is larger than 1 for both groups (Table 2.2). In that case, it can be said that, both groups cannot be distinguished according to their area.

2.3.1.2. Shape and Morphology Based Features

Features related to shape and morphology of the blobs can affect the characteristics of the search space. Three shape-based features are extracted to evaluate the significance of shape and morphology on the separability of spine bones from non-spine bones.

One such feature is eccentricity whose value varies between 0 and 1. It equals the eccentricity of the ellipse that has the same second moment of area as the blob. Once, the equivalent ellipse is defined, the eccentricity is calculated as:

$$e = \sqrt{1 - \frac{b^2}{a^2}} \tag{2.1}$$

where a and b are lengths of semi-major and semi-minor axes of the ellipse, respectively. Effect of increasing eccentricity on shape is visually presented in Figure 2.4 with different sample blobs.

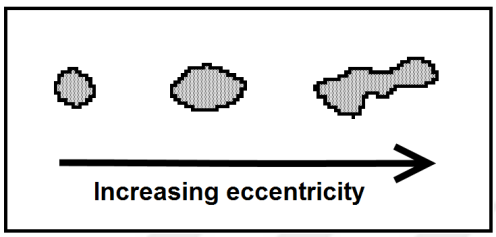


Figure 2.4. Effect of increasing eccentricity. Lower eccentricity means higher similarity to a perfect circle.

Another geometric feature that can be measured is the equivalent diameter. It is equal to the diameter of the circle which has the same area as the blob. The formula for this feature is given as:

$$EquivalentDiameter = \sqrt{\frac{4 * Area}{\pi}} \tag{2.2}$$

This feature is visually demonstrated in Figure 2.5.

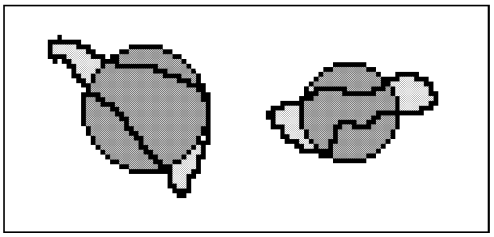


Figure 2.5. Equivalent diameter feature visualized on two sample bone regions. The circle equivalent to the region is shown in darker color.

The third geometric measure of interest is compactness. It is calculated as the ratio of the area of a blob to the area of a circle with the same perimeter, given as:

$$Compactness = \frac{4 * \pi * Area}{(Perimeter)^2} \tag{2.3}$$

Cluster average with standard deviation values of all shape and morphology based features are presented in Table 2.3.

Table 2.3. Average values of shape and morphology based features with standard deviation values

Equivalent Diameter			
Spine Bones		**Other Regions**	
Average	Std.	Average	Std.
23,57	10,90	17,35	11,48
Eccentricity			
Spine Bones		**Other Regions**	
Average	Std.	Average	Std.
0,85	0,15	0,86	0,12
Compactness			
Spine Bones		**Other Regions**	
Average	Std.	Average	Std.
0,84	0,28	0,90	0,32

2.3.1.3. Location Based Features

Features based on spatial positions of blobs also exist in the feature space as well as shape and morphology based features. It should be remembered that the spine has a typical shape for which we have some *a priori* knowledge. Hence, spine bone blobs are expected to follow a curve that represents this typical spine shape.

One location based feature is the distance of the nearest neighbouring blob to the considered blob (Figure 2.6). Accordingly, for each blob, the distance of the nearest neighbour is measured.

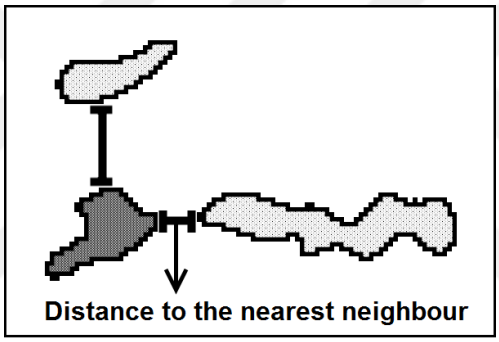


Figure 2.6. Distance to nearest neighbour visually depicted

Another related feature extracted from the images is the angular position (i.e. direction) of the nearest neighbouring blobs. If blobs are connected with a chain through their centroids, the connections of the considered blob with its two nearest neighbours form an angle, which is taken as a location based feature (Figure 2.7).

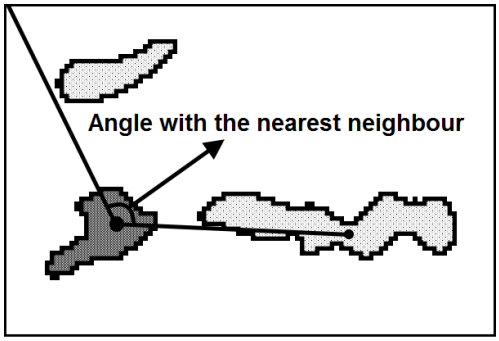


Figure 2.7. Angle defined using the nearest neighbours of a blob. Darker blob is the subject of measurement.

Average values of location based features for spine and non-spine bone regions are summarized in Table 2.4.

Table 2.4. Average values of distance to nearest neighbour with standard deviations

Distance to the closest neighbour			
Spine Bones		Other Regions	
Average	Std.	Average	Std.
11,48	5,01	20,53	22,16
Angle with the closest neighbour			
Spine Bones		Other Regions	
Average	Std.	Average	Std.
133,63	35,30	125,37	39,34

2.3.1.4. Significance of the Features

The identification task of this study is basically a two class binary classification problem. The above mentioned features are analysed with respect to their significance in classifying spine and non-spine bones. Accordingly, it is possible to

rank previously extracted features with respect to their significance for binary classification using an objective evaluation criterion. Two different approaches have been utilized to evaluate the features for this task.

T-Test Based Class Separability Approach: T-test is a common hypothesis test which checks if the difference in averages of two groups is statistically significant or insignificant. Absolute value two-sample T-test with pooled variance estimate is utilized here as class separability criterion in ten ultrasound samples. Significance of each extracted feature (the examined features are named and numbered on Table 2.6) is tested for two classes and then ranked for each sample. Due to variable nature of ultrasound samples, we have observed that the same feature is usually ranked differently on different samples (Table 2.5). This indicates that the defined features alone cannot distinguish spine from non-spine in this approach.

Table 2.5. Names and numbers of the examined features

Feature Number	**Feature Name**
1	Area
2	Eccentricity
3	Equivalent Diameter
4	Compactness
5	Nearest Blob Distance
6	Nearest Blob Angle

Table 2.6. Features given on Table 2.5 sorted according to their significance with the T-test approach, where Rank 1 indicates the best and Rank 6 indicates the worst class separability

	Rank 1	**Rank 2**	**Rank 3**	**Rank 4**	**Rank 5**	**Rank 6**
Sample 1	5	3	4	6	1	2
Sample 2	5	6	3	4	2	1
Sample 3	3	2	1	5	4	6
Sample 4	3	1	5	6	2	4
Sample 5	5	3	1	4	6	2
Sample 6	6	5	3	2	1	4
Sample 7	3	2	4	1	5	6
Sample 8	5	4	2	6	3	1
Sample 9	6	5	2	4	3	1
Sample 10	3	1	5	4	2	6

Exhaustive Approach: Exhaustive search or brute-force search is a general problem solving approach where *all* possible combinations of features are systematically evaluated. An exhaustive search approach is the second way taken to determine the significance of extracted features (Figure 2.8).

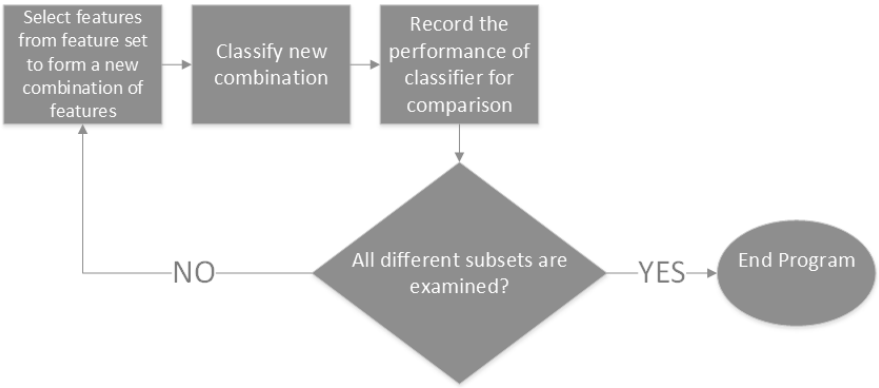


Figure 2.8. Block diagram of the utilized exhaustive search algorithm

With six examined features, classification tasks are realized with $2^6 - 1 = 63$ different feature subsets. Two well-known clustering algorithms are implemented to classify spines and non-spines using each feature set. The first one is K-means where each of the n observations (i.e. bone blobs) is divided into two clusters (spine and non-spine) by assigning them to the nearest mean obtained in the previous classification iteration (Kanungo et al., 2002).

To formulate K-means, let the arrays of observations and centers be defined, respectively, as:

$$X = \{x_1, x_2, x_3 ..., x_n\} \tag{2.4}$$

$$V = \{v_1, v_2, v_3 ..., v_c\} \tag{2.5}$$

where c is the number of clusters. Center of the jth cluster (v_j) is calculated with observations as:

$$v_j = (\frac{1}{c_j}) \sum_{i=1}^{c_j} x_i \tag{2.6}$$

where c_j stands for the number of observations in jth cluster. Clustering iteration starts with forming random clusters and finding their centers. Then, each observation is assigned to the closer one of the two centers and cluster centers are recalculated. The procedure is repeated until square sums of distance to the center of each cluster attains a minimum:

$$J(V) = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i - v_j \right\|^2 \tag{2.7}$$

where, k, n, x_i and v_j indicates number of clusters, number of observations, ith observation and centroid of the jth cluster respectively.

Classification performance of the method with each feature subset is measured with F-score, which can be expressed as:

$$Fscore = 2 * \frac{Precision * Recall}{(Precision + Recall)} \tag{2.8a}$$

where,

$$Precision = \frac{Tp}{(Tp + Fp)} \tag{2.8b}$$

$$Recall = \frac{Tp}{(Tp+Fn)} \tag{2.8c}$$

Here, Tp, Fp and Fn are the numbers of true positives, false positives and false negatives, respectively (Goutte and Gaussier, 2005). Best feature combination is chosen for each sample with respect to its F-score (Table 2.7).

Table 2.7. Best resulting combinations of features given on Table 2.5 classified with K-means for ten samples

	Max. F-Score	**Feature 1**	**Feature 2**	**Feature 3**	**Feature 4**	**Feature 5**	**Feature:6**
Sample 1	0,55	0	0	1	1	0	0
Sample 2	0,71	0	0	1	0	0	0
Sample 3	0,55	0	1	1	0	0	0
Sample 4	0,62	0	0	1	0	0	0
Sample 5	0,45	0	0	0	0	1	0
Sample 6	0,47	0	1	0	0	0	0
Sample 7	0,44	0	0	1	0	1	0
Sample 8	0,71	0	1	1	1	1	0
Sample 9	0,56	0	1	0	0	1	0
Sample 10	0,55	0	1	1	0	0	0

The second classifier implemented in the study is the fuzzy C-means (FCM) classifier. FCM algorithm runs on similar principles as the K-means algorithm. However, unlike to K-means, FCM operates on partial membership, where an observation has a degree of belongingness to each cluster (Cannon et al., 1986). Objective function of FCM with membership degree is given as:

$$j_m = \sum_{i=1}^{N} \sum_{j=1}^{C} u_{ij}^{m} \left\| x_i - c_j \right\|^2 ; 1 < m < \infty \tag{2.9}$$

where u_{ij} is the degree of membership, x_i is ith observation of data, and c_j is center of the jth cluster. m is a finite real value larger than 1 that sets the level of fuzziness. Accordingly u_{ij} is calculated by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{\left\| x_i - c_j \right\|}{\left\| x_i - c_k \right\|} \right)^{\frac{2}{m-1}}} \tag{2.10}$$

where center of the clusters are calculated by:

$$c_j = \frac{\sum_{j=1}^{N} u_{ij}^{m} . x_i}{\sum_{j=1}^{N} u_{ij}^{m}} \tag{2.11}$$

Best classification performances and feature combinations obtained with FCM are given in Table 2.8.

Table 2.8. Best feature combinations obtained with FCM for ten samples

	Max. F-Score	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature:6
Sample 1	0,53	0	0	1	1	0	0
Sample 2	0,69	0	0	1	0	1	0
Sample 3	0,54	0	1	1	0	0	0
Sample 4	0,62	0	0	1	0	0	0
Sample 5	0,48	0	0	0	0	1	0
Sample 6	0,47	0	1	0	0	0	0
Sample 7	0,50	0	1	1	0	1	0
Sample 8	0,65	0	1	0	0	1	0
Sample 9	0,57	0	1	0	0	1	0
Sample 10	0,47	0	0	1	0	0	0

Best results of both applied classifiers have been consistent with result of T-Test based class separability approach. No result has a maximum F-score exceeding 0.71 and best feature combination differs from one sample to the other. One cause for the variance in the significance of the morphology, location and area based features among samples may be originating from noise and distortions on the samples and deformed morphology of the spine. It should be noted again, that analyzed sonograms are not regular healthy samples. Consequently, it can be concluded that conventional machine learning methods may be inadequate to distinguish spine from non-spine in this particular problem.

2.4. Flocking Dynamics Based Search Space Optimization

In this work, our objective is to simplify the identification problem of the vertebrae on ultrasound samples by reducing the size of the search space (i.e. all bone blobs) of the problem. Considering the large number of bone blobs on a typical ultrasound, the optimal solution to the problem is sought in a large solution space. If

neighbouring bone regions that are likely to belong to the spine simultaneously can be combined into a single bone from an algorithmic point of view, then the number of bone blobs is effectively decreased and the search space will be significantly reduced in size. That way, the solution can be found with less computational effort. The proposed algorithm is expected to identify neighbouring spine fragments, as they appear on the sonograms, and combine them into a single bone region.

The algorithm for connecting neighbouring bone regions is based on the behaviour of bird flocks. Each bone region is taken as a bird flock. A series of calculations are applied to the flock (i.e. a bone region) to determine its migration path, which is equivalent to find an appropriate neighbour to connect the bone region to. In this study, three basic principles of flocking behaviour have been incorporated to the algorithm, namely collision avoidance, aligning and centering (Reynolds, 1997), which are explained below.

2.4.1. Centering and Neighbouring

Adjacent pixels are assumed to be in the same local flock which allows us to accept each detected blob as a flock. Accordingly, each local flock has a flock centroid. Proposed algorithm connects flocks by connecting the centroids to form flock groups. Hence, flocks are programmed to search and try to find the best neighbour to connect to. It should be noted that flocks try to find the shortest possible path, therefore, a ”migration” path is indicated by a straight line connecting two flock centroids (Figure 2.9).

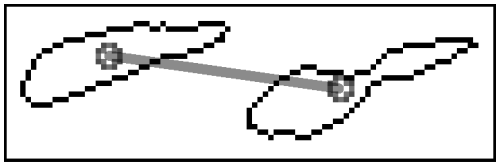


Figure 2.9. Two flocks (i.e. bone blobs) are connected through their centroids (indicated with circles) with a straight line to form a single object.

2.4.2. Concept of Perception

Bird flocks avoid to crowd by keeping a certain distance between them, which is referred to as collision avoidance principle (Cui et al., 2006). Each flock is assumed to have a limited perception zone and other flocks are visible only if they are inside this perception zone. The region that each flock perceives is determined by expanding the flock perimeter by a specified amount in the direction of the boundary normal. This means that only those flocks that are within each others perception distance can be connected (Figure 2.10).

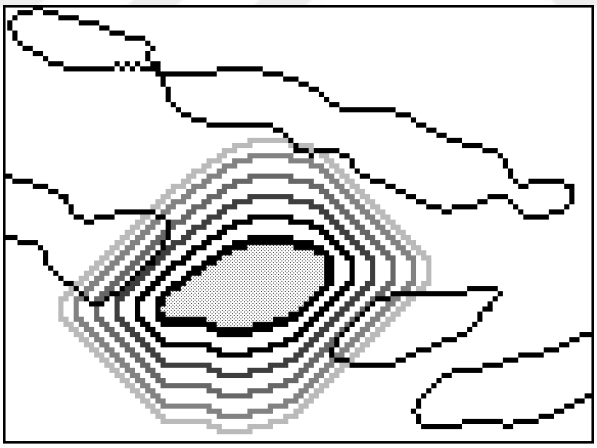


Figure 2.10. Perception zone for a sample flock. There are three immediate neighbours that the flock is able to detect.

2.4.3. Fatal Distance

There is a possibility that a flock could not detect any local flock-mates to connect to, if it is too isolated from all other flocks. A fatal distance is defined such that if the nearest flock is further than from this distance, the flock is terminated (Figure 2.11). This effectively means that a bone region which is too isolated from other bone regions is not likely to be a part of the spine.

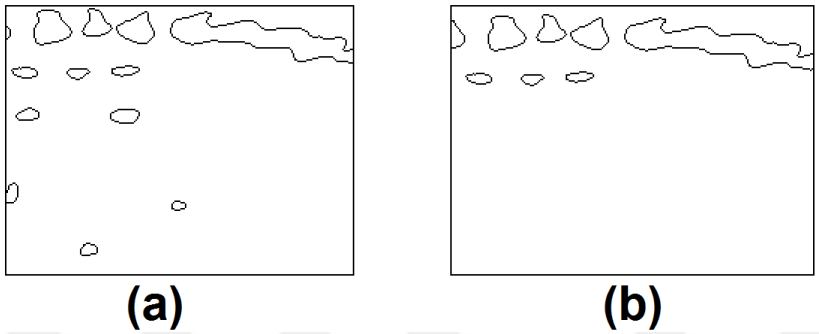


Figure 2.11. Effect of fatal distance. A sample region before (a) and after (b) eliminating isolated flocks

2.4.4. Alignment

Proposed algorithm searches for the best neighbour to connect for each flock. It should be noted that orientation of the connections that two or more flocks make with each other will affect the choice of the next neighbour to be added to this flock group (alignment principle) (Su et al., 2009). This allows local flocks that align along a smooth curve to be grouped together (i.e. aligned bone regions are likelier to belong to the spine.) Fundamental steps for determining the flocks to be connected is called Algorithm 1 which is visually depicted in Figure 2.12.

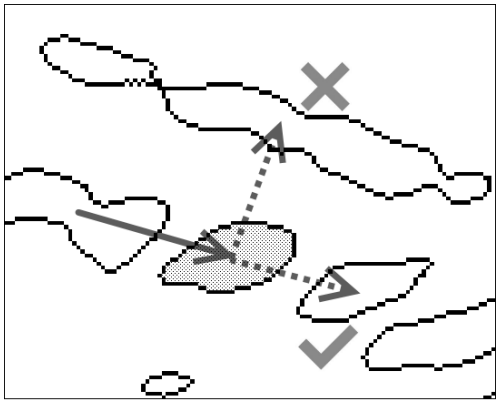


Figure 2.12. Choice of best neighbour according to orientation. A search is done for red coloured flock. Blue arrow is the previous connection and green arrow indicates best connection for red coloured flock.

2.4.5. Search Space After Flocking Based Optimization Approach

Once flock groups are formed and isolated flocks are terminated, the remaining single flocks (those that are neither eliminated nor connected) and flock groups can be part of the spine. All possible combinations of these elements form the solution space. The solution space is now smaller since it contains a smaller number of elements due to the combined and eliminated flocks. Consequently, the complexity of the problem is reduced. At that point, there would be two possible failure scenarios. One of them is the possibility of false connection. Connections of bones belonging to the correct cluster (either spine or non-spine) should be considered a true connection. For example, two or more spine regions connected or equivalently two or more non-spine regions connected should be considered optimal. The other scenario that may become reality is false elimination (deletion) of a spine bone. If any spine blob erased falsely due to its distance to a nearest neighbour, success of the utilized classifier would be compromised. The occurrence of these failure scenarios with this algorithm is summarized for the sample images in Table 2.9.

Table 2.9. Percent reduction of search space given for each sample with false connections and false kill rates

	Reduction Rate	**Number of Connections**	**False Connection Rate**	**False Kill Rate**
Sample 1	28,94%	32	0%	0%
Sample 2	32,14%	26	0%	0%
Sample 3	43,01%	44	4,54%	0%
Sample 4	26,02%	43	16,27%	0%
Sample 5	64,86%	14	0%	0%
Sample 6	33,33%	13	0%	0%
Sample 7	61,36%	26	0%	0%
Sample 8	34,88%	11	0%	0%
Sample 9	50%	23	17,39%	0%
Sample 10	33,80%	39	0%	0%

To evaluate the flocking based search space optimization algorithm, a ground truth is established where the actual spine bone regions are identified and marked by an expert obstetrician. Spine bones constitute a discrete collection. Sometimes, a seemingly large number of false connections may occur, but the estimated spine axis looks still close to the actual axis. In this case, the number of falsely connected bone regions may be misleading with respect to the success of the optimization scheme. This occurs typically when falsely connected bones are relatively small in size and located close to the actual spine axis. Error of this type is the very reason why we have decided to define a fifth order ground truth polynomial as well. After bone regions are appropriately grouped in the search space reduction, a fifth order polynomial is fitted using the group centroid coordinates and a comparison is made with the ground truth polynomial defined earlier.

Experimental results indicate that, proposed optimization approach can reduce the number of blobs (i.e. the size of the search space) up to %64. Proposed method

connects bone regions with a success rate of %100 in seven samples.

The deviation of the fitted polynomial as well as the deviation of the curves drawn by the three expert obstetrician(Figure 2.13) from the ground truth polynomial are calculated through the root mean square error, given as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - f(x_i))^2}{n}} \tag{2.12}$$

where (x_i,y_i) indicate the coordinates of the points of the estimated curve, coming either from the optimization algorithm or drawn by the three experts. $f(x_i)$ is the corresponding y value of the ground truth polynomial. The error is evaluated in an integral sense and thus n is the number horizontal pixels in the x-range of the spine image. Objective evaluation of the total error is achieved by comparing it with the error between hand drawn curves and curve fitted to actual spine bone blob centroids (Table 2.10).

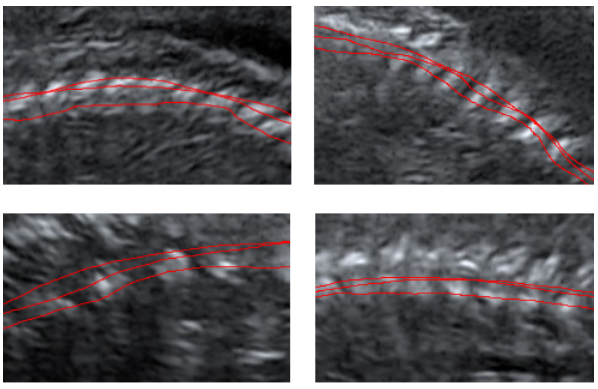


Figure 2.13. Curves hand drawn by experts shown on sample regions

Table 2.10. Integrated RMSE values of hand drawn curves and estimated curve with respect to the ground truth polynomial

	Number of False Connected Blobs	Exp. 1	Exp. 2	Exp. 3	Estimated Curve after Optimization
Sample 1	0	12,842	31,859	7,232	3,946
Sample 2	0	21,628	9,585	5,630	1,786
Sample 3	2	11,599	15,524	18,260	8,490
Sample 4	10	8,965	14,389	11,626	7,043
Sample 5	0	18,641	49,547	46,639	0,836
Sample 6	0	7,186	11,819	18,771	1,040
Sample 7	0	11,963	16,232	15,223	4,212
Sample 8	0	9,949	32,390	14,218	0,282
Sample 9	6	7,862	6,291	15,834	8,140
Sample 10	0	7,664	12,257	17,127	0,341

Typical output of flocking process is presented for six different sonogram samples in Figure 2.14.

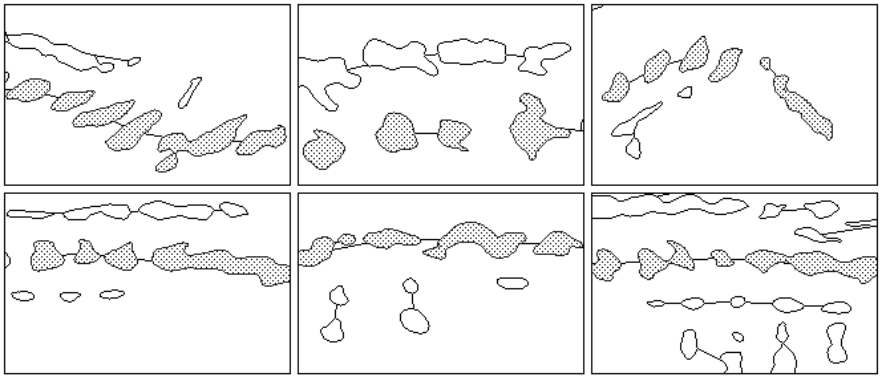


Figure 2.14. Output of flocking shown on different samples. Filled regions are bones on the spine axis where migration paths are shown with lines.

2.4.6. Meta-heuristic Evaluation of Flocking Based Optimization Approach

Before introducing our custom-made genetic algorithm in the coming sections, we evaluate the performance of the search space reduction by comparing the conver-

gence of optimized and non-optimized search spaces with a straightforward genetic algorithm. The performance of the algorithm is an indication for the effectiveness of the above-described flocking procedure.

The classification process is based on the evolution of an initial population and it is accomplished via common genetic operators. Some important concepts related to genetic algorithms are explained in connection with the current problem below:

Individuals: Each candidate solution to the problem of spine identification is an individual. Individuals consist of genes.

Genes: Each flock region is represented by a gene in the candidate solution (i.e. individual). Accordingly, for a given problem (i.e. a sonogram), each individual has a fixed number of genes (representing the number of visible spine bone regions in the sonogram)(Chang et al., 2009). A gene is a binary variable, which equals one in case the related flock belongs to spine and zero otherwise.

Population: A population is the collection of a predefined number of candidate solutions (Whitley, 1994).

Generation: Population renewed with evolutionary operators forms the next generation in the genetic algorithm (Lorena, et al., 2002).

2.4.6.1. Evolution Mechanism

The evolution of a population is realized by applying two genetic operators to the population at every iteration (representing a generation). The crossing-over operator allows the best (i.e. fittest) solutions to pass their properties to the subsequent iterations. The mutation operator increases the chance of finding fitter solutions by diversifying the individuals in the search space (Figure 2.15). Each solution in the initial population is rated according to its fitness value. From the best solutions, new solutions are created through genetic operators, which are supposed to have improved

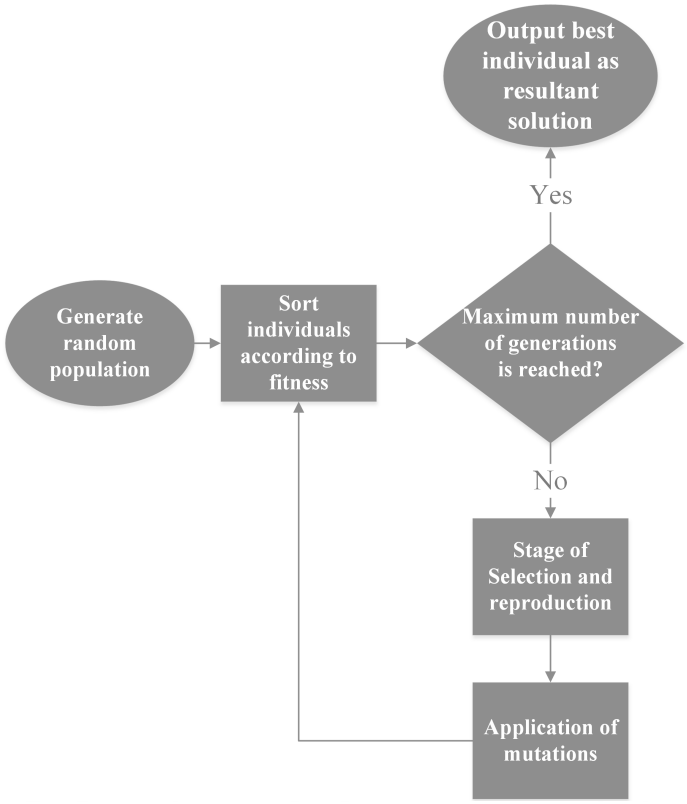


Figure 2.15. Flow chart of the implemented conventional genetic algorithm

fitness values, if the algorithm works properly.

2.4.6.2. Fitness Criteria

The fitness value is an indicator for the closeness of the individual to the actual solution (Bottaci, 2001). In many problems, utilizing genetic algorithms, the best solution (or whether there exists a best solution at all) is not known. In this work, however, the best solution (i.e. ground truth) is known and each individual of each generation is compared with the best solution and assigned a fitness value accordingly. This way, we are able to evaluate the effectiveness of the flocking algorithm.

Fitness of an individual is measured by calculating its F-score with respect to the ground truth.

Each iteration generates a new population. The best individual in each generation sets the success rate of that generation. Test problems are run up to a maximum of 100 generations.

2.4.6.3. Performance of the Meta-heuristic Classification

The genetic algorithm is tried with the ten sample images for 100 generations on both optimized (reduced number of bone regions) and non-optimized (original number of bone regions) version of each image sample. Each run is repeated ten times and maximum F-score of ten classifications are presented for comparison (Table 2.11).

Table 2.11. Best F-scores presented for several generations

	Non-optimized				**Optimized**			
	1. Gen.	**20. Gen.**	**50. Gen.**	**100. Gen.**	**1. Gen.**	**20. Gen.**	**50. Gen.**	**100. Gen.**
Sample 1	0,775	0,845	0,919	0,969	0,896	1,000	1,000	1,000
Sample 2	0,702	0,844	0,897	0,982	0,806	1,000	1,000	1,000
Sample 3	0,779	0,815	0,877	0,951	0,842	0,936	0,979	0,986
Sample 4	0,725	0,833	0,898	0,968	0,828	0,931	0,933	0,933
Sample 5	0,804	0,887	0,923	0,967	0,865	0,924	0,950	0,992
Sample 6	0,812	0,907	0,975	0,988	0,933	1,000	1,000	1,000
Sample 7	0,714	0,822	0,870	0,943	0,853	0,920	0,986	1,000
Sample 8	0,821	0,947	1,000	1,000	0,945	1,000	1,000	1,000
Sample 9	0,747	0,842	0,947	0,987	0,854	0,923	0,951	0,951
Sample 10	0,745	0,882	0,914	0,972	0,878	0,991	1,000	1,000

In all image samples, the optimized version consistently results to a better and faster convergence to the correct solution. The iteration converges to a value of one

if spinal bone regions are all correctly connected, which is the case in most samples. A value smaller than one implies that either convergence has not occurred yet or a non-spine region is connected to a bone group. The visual complexity of each sonogram image may be different. Convergence rate may differ significantly from one sample to the other. The convergence behaviour of different sonogram samples are comparatively displayed in Figure 2.16.

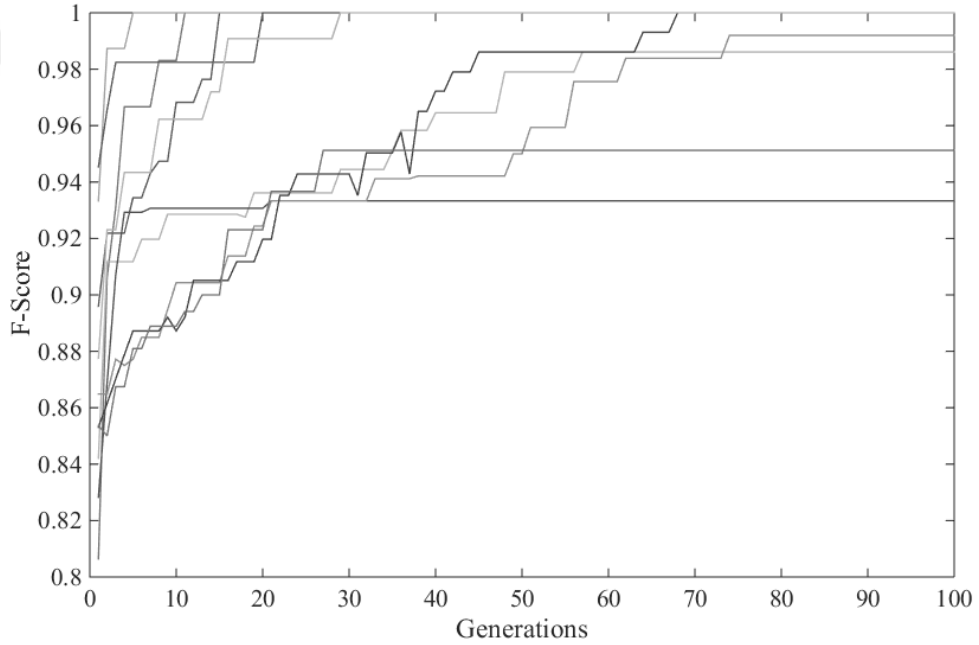


Figure 2.16. Classification performance of genetic setting is given in different shades of gray for each sample. Best result of ten repetitions is given for each generation

2.5. Genetic Algorithm Based Discrimination of Flocks

Automated detection of spine axis is the main objective of this study. Accordingly, a custom made meta-heuristic mechanism is presented for forming a chain of previously extracted and optimized (i.e. grouped) spine bone regions among spinal axis.

2.5.1. Implemented Genetic Structure and Individuals

It is possible to observe that spine bones roughly align along the horizontal on most samples. That knowledge leads us to take an approach where horizontally lined-up bone blobs are sought within columns of sample sonograms. The proposed algorithm divides each sample into columns, and blobs, grouped blobs or part of these that remains in each column are assigned to this column according to their location (Figure 2.17). Each assigned pixel group is represented by its centroid.

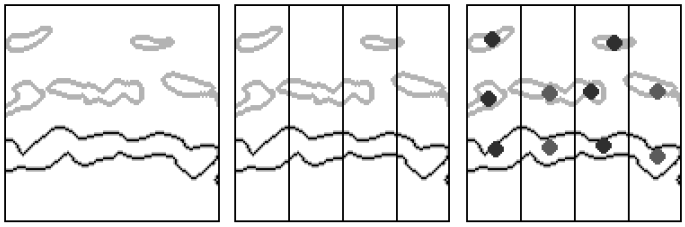


Figure 2.17. Columnar genetic structure graphically depicted. (a) Previously detected bone regions (darker contours indicate spine regions). (b) Sample region is divided into columns. (c) Each centroid is marked as a selection alternative belonging to the related column.

A column too wide may cause loss on sensitivity or too many narrow columns may increase the problem complexity. Furthermore, the presence of a bone region in a specific column is assumed if more than 100 pixels of the region falls within that

column. With these considerations, average value of column width is chosen to be 30 pixels for all samples.

An encoding scheme is implemented for generating candidate solutions compatible with introduced columnar approach. According to implementation, each candidate solution is a chain of binary blobs or blob groups selected from different columns (Figure 2.18). The selection starts with the left most column of the image and proceeds to the right. Each complete chain is accepted as an individual (i.e. a solution candidate) and individuals form a population. Accordingly, chromosomes of individuals store four values for each column. These values are column number, label of the selected blob from the column and, x and y coordinates of the centroid of the selected blob.

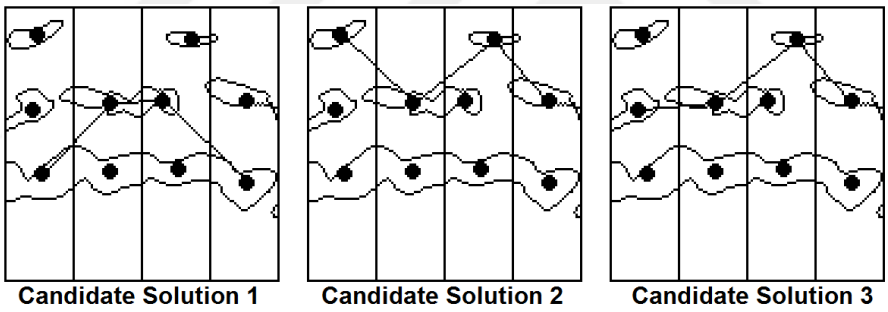


Figure 2.18. Three different candidate solutions on a five column sample. Selections taken from each column are connected with a line.

2.5.1.1. Propagation of Connections through Chromosomes

Division of the image to columns usually cause whole blobs to be sliced into multiple pieces. Since this procedure will increase the number of bone regions, it may counter the effect of previous search space reduction where bone blobs are combined in order to reduce their number. To cope with that unwanted side effect, a bone piece

is always selected if it originally belongs to the same bone blob with the previous selection. Pseudo code for the process is given below:

Algorithm 1 Propagation of connections

1: *From left to right*
2: *For*: *Each Column*
3: **if** chosen pixel group is part of another blob from neighbouring column **then**
4: *Update recent chromosome to connect pixel groups belonging to same blob*
5: Go to: Next Column

The effect of this procedure is visualized in Figure 2.19 for three different solutions on a five column sample region.

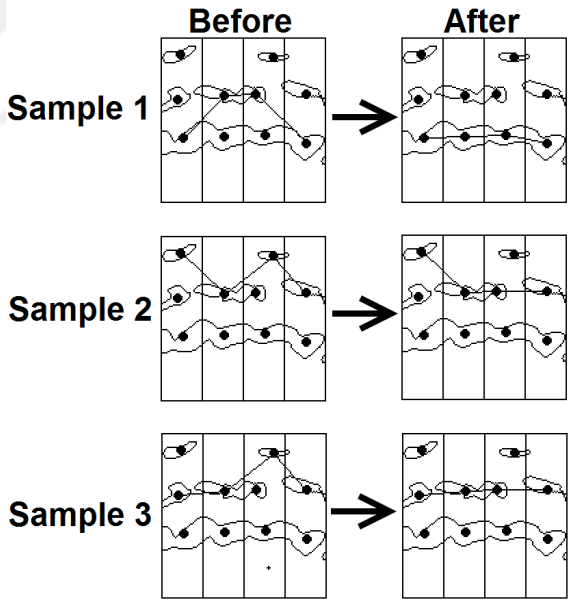


Figure 2.19. Effect of propagation is visualized on three different candidate solutions where sample region is divided into five columns

2.5.2. Implemented Evolutionary Algorithm

The evolution of populations is realized by applying several operators to each population at every iteration (i.e. generation) (Figure 2.20).

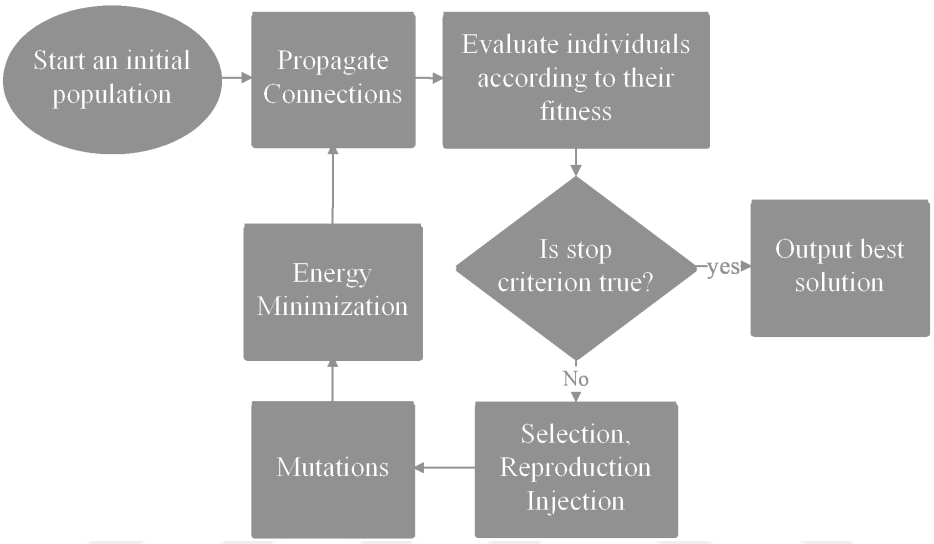


Figure 2.20. The applied evolutionary approach given as a block diagram

In this study, implemented meta-heuristic classifier follows the conventional genetic algorithm flow. However, each genetic operator that makes the individuals evolve is reimplemented to be compatible with the specific genetic structure in this problem. Explanation and implementation of the genetic algorithm blocks are explained in the coming subsections.

During the evolution process, selection of the fittest individuals is a crucial task which allows the algorithm to converge to the actual solution efficiently. Accordingly, each solution in the initial population is rated according to its fitness value. New solutions from the best solutions are created through genetic operators.

The fitness criteria involves a linear combination of three quantities that can be taken as a *cumulative feature* of an individual. The first quantity is the sum of

the distances between the centroids of consecutive regions. The second one involves the average of the curvature values measured at the centroids of selected regions (except at the first and last column). The last quantity is the average root-mean-square-distance of the centroids to a third order polynomial fitted to the bone region centroids. Fitness based on these quantities make sure that a candidate that has a certain level of smoothness (and lacks zigzaggedness just like a real spinal axis) is a fit candidate (Figure 2.21).

Let $x_i = (x_i, y_i)$ represent the centroid coordinates of the selection in the ith column

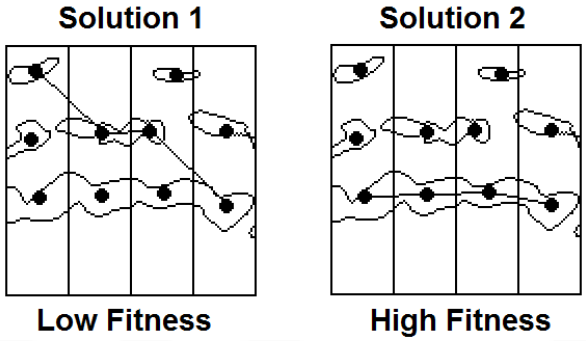


Figure 2.21. Sample five column individuals as examples of low and high fitness individuals

and $f_i = (x_i, z_i)$ be the coordinates of the points on the fitted third order polynomial that has the same x-coordinates as the centroids. Then, the fitness function is given as:

$$fitness = \frac{1}{\alpha \sum |x_{i+1} - x_i| + \beta \sum \kappa_i + \gamma \sum |f_i - x_i|} \tag{2.13}$$

where α, β and γ are the weighting parameters. κ_i is the curvature at the centroid with coordinates x_i and equals the inverse of the radius of the circle that is fit to x_{i-1}, x_i and x_{i+1}.

At each generation, after the fitness of each individual is determined, the crossover operator is applied to the population for creating the next generation from the existing fittest solutions. The operation starts with the selection of two random parent solutions from the population. To create a bias toward the selection of fittest solutions, the probability for a solution to be selected is set proportional to its fitness value. In this method, also known as *roulette wheel selection method*, a solution with a fitness value of, say, 2c is twice as likely to be selected for the crossover operation compared to a solution with a fitness of c.

After the parent solutions are picked, a new solution is formed by combining blocks of parents' chromosomes. In other words, the columns are divided into two blocks, one block is taken from one parent while the other one is taken from the other parent to create the new individual. In general, the new individual does not obtain equal amount of genetic material from the parents.

2.5.2.1. Population Diversity

Diversity of each generation is a parameter that may drastically affect the ability of the algorithm to explore through the search space. Too many similar individuals or clone individuals may cause premature convergence (Gupta and Ghafir, 2012). An inadequate level of population diversity is one of the causes that makes the algorithm stop at a suboptimal state. In that case, creation of individuals that outperform their parents is no longer possible.

Mutations: Mutation is another operator applied to increase genetic diversity in the population. Mutation is applied by randomly changing the selection in a specified number of columns in a certain percentage of individuals. Since mutation may increase or decrease the fitness of an individual, it is not applied to the fittest few individuals in a generation. Mutation also prevents population generations from getting

stuck in a state of local optimum, where fitness of individuals cannot be further improved. On the other hand, excessive levels of mutation may have a negative effect, where new individuals are generated without an apparent improvement in fitness. The level of mutation has been fine-tuned in this study to obtain the desired effect.

Injection: Injecting completely new and randomly created individuals to replace the least fit individuals and present clones at each generation is the third precaution taken to ensure diversity.

Energy Minimization: As mentioned before, the fitness ensures the smoothness of the curve that may represent the actual spine. The three quantities present in Eq.(2.13) are sometimes interpreted as energy components of a curve (Talebi et al., 2011) and a high fitness corresponds to a minimal energy.

As a part of the genetic algorithm, an energy minimization operation is also applied to better the individuals. This is accomplished by changing the selection in the column of the highest energy centroid in order to reduce individual's energy (Figure 2.22).

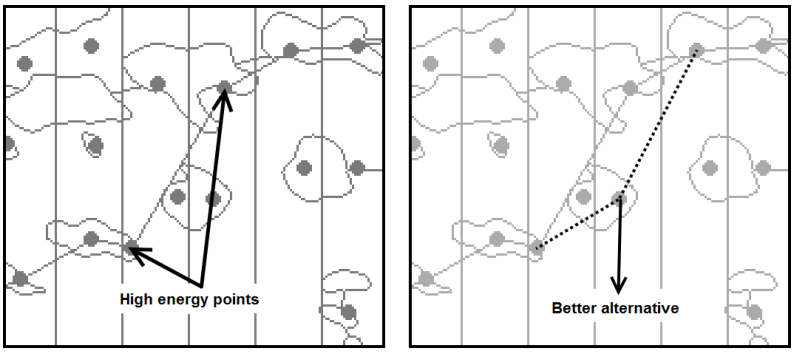


Figure 2.22. From left to right, first frame indicates highest energy points. Best alternative to minimize the energy is shown in the second frame.

It should be noted that energy minimization process would not always create

better individuals. However, it is also helpful for increasing the diversity and lower the possibility of sticking at a suboptimal state.

3. RESULTS

We first evaluate the performance of the genetic algorithm with three different fitness criteria that are based on the ground truth. In other words, the fitness of each individual at every generation is evaluated by comparing it with the ground truth. The fitness defined with a knowledge of the ground truth is referred to as *golden fitness* in this manuscript. Since the spine roughly falls in the middle of a typical sonogram, first an active region is defined on the image by ignoring the first few and the last few columns (Figure 3.1).

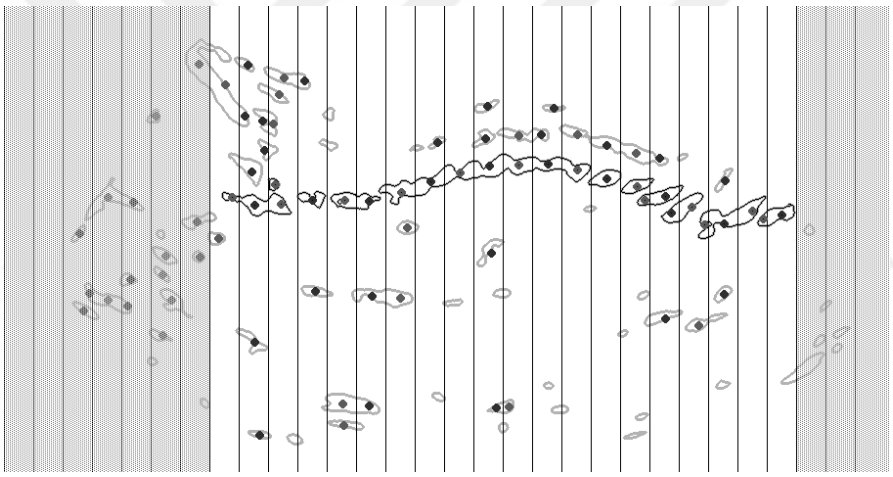


Figure 3.1. Active region within the range of actual spine bones is shown between gray regions. Bold contours indicate the spine bones.

The algorithm is run with both optimized and non-optimized search spaces. The three fitness criteria are tested with 50 individuals which evolve through 100 generations. Each test is repeated ten times.

3.1. Performance of Euclidean Distance Based Fitness

The first fitness criteria is based on the Euclidean distance (expressed in pixels) between the centroids of the selected region and the actual spine region in each column (Figure 3.2). Hence, the total fitness of any individual is calculated as:

$$Fitness = \begin{cases} 1, \text{ if Total Error=0;} \\ \dfrac{1}{\text{Total Error}}, \text{ Otherwise} \end{cases} \tag{3.1}$$

Note that, the above error is measured in pixels and is hence an integer for each individual which, in practice, cannot be smaller than a value in the order of tens. Its average value over ten repetitions can be a decimal number.

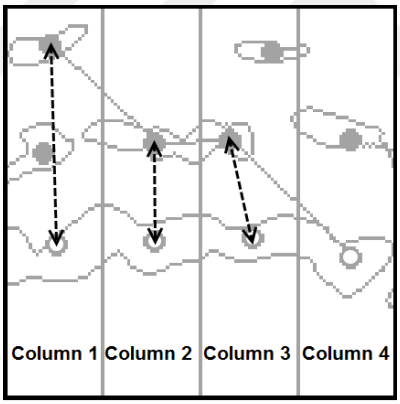


Figure 3.2. Euclidean distance based error

If there is no detected spine bone inside the column, the error is assumed to be non-present in that column. In case there are multiple spine regions in the column, the one closest to the selected region is accepted as the ground truth. Presented results in Table 3.1 are average values of ten repetitions.

Table 3.1. Euclidean distance based golden fitness results (average of ten repetitions)

	Non-optimized				Optimized			
	1th Gen	**20th Gen**	**50th Gen**	**100th Gen**	**1th Gen**	**20th Gen**	**50th Gen**	**100th Gen**
Sample 1	0,104	0,901	1,000	1,000	0,404	1,000	1,000	1,000
Sample 2	0,505	0,902	1,000	1,000	0,900	0,900	1,000	1,000
Sample 3	0,001	0,004	0,004	0,004	0,001	0,103	0,103	0,309
Sample 4	0,006	0,023	0,023	0,220	0,004	0,702	0,702	1,000
Sample 5	0,004	0,116	0,314	0,414	0,206	0,703	0,703	0,703
Sample 6	0,010	0,902	1,000	1,000	1,000	1,000	1,000	1,000
Sample 7	0,001	0,204	0,304	0,305	0,001	0,402	0,601	0,801
Sample 8	0,106	1,000	1,000	1,000	1,000	1,000	1,000	1,000
Sample 9	0,002	0,005	0,005	0,105	0,001	0,002	0,002	0,002
Sample 10	0,004	0,105	0,105	0,105	0,603	1,000	1,000	1,000

The number of perfect classifications in 1th, 20th, 50th and 100th generations are presented in Table 3.2.

Table 3.2. Number of runs where perfect classification is achieved with Euclidean distance based golden fitness

	Non-optimized				Optimized			
	1th Gen	**20th Gen**	**50th Gen**	**100th Gen**	**1th Gen**	**20th Gen**	**50th Gen**	**100th Gen**
Sample 1	1	9	10	10	4	10	10	10
Sample 2	5	9	10	10	9	9	10	10
Sample 3	0	0	0	0	0	1	1	3
Sample 4	0	0	0	2	0	7	7	10
Sample 5	0	1	3	4	2	7	7	7
Sample 6	0	9	10	10	10	10	10	10
Sample 7	0	2	3	3	0	4	6	8
Sample 8	1	10	10	10	10	10	10	10
Sample 9	0	0	0	1	0	0	0	0
Sample 10	0	1	1	1	6	10	10	10

3.2. Performance of F-Score Based Fitness

Next, F-score is used as a fitness criteria with the genetic algorithm. Accordingly, each blob on the image is labelled and a ground truth matrix is extracted which indicates the true blobs in each column. Each individual is compared with that matrix to calculate F-score (Figure 3.3). If there are multiple spine bones in any column, only one of them is accepted as true blob. Also, implemented F-score based fitness approach is binary, which means that each column is allowed to hold only one spine bone.

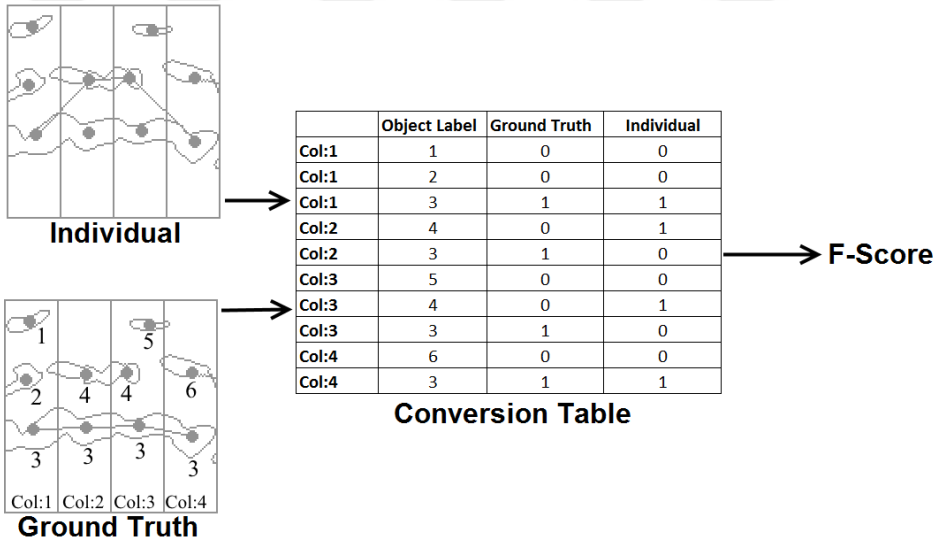


Figure 3.3. F-score based fitness concept explained on a four column search space. There are six labelled objects which are indicated on ground truth visual. Conversion table shows how individuals and ground truth are coded into binary arrays to calculate F-score

Average results and quantities of perfect classifications for F-score based fitness is given in Table 3.3 and Table 3.4 respectively.

Table 3.3. F-Score based golden fitness test results (average of ten repetitions)

	Non-optimized				Optimized			
	1th Gen	**20th Gen**	**50th Gen**	**100th Gen**	**1th Gen**	**20th Gen**	**50th Gen**	**100th Gen**
Sample 1	0,938	0,997	0,997	0,998	0,964	1,000	1,000	1,000
Sample 2	0,912	0,990	0,990	0,992	0,893	0,993	0,996	1,000
Sample 3	0,923	0,980	0,980	0,980	0,891	0,956	0,967	0,970
Sample 4	0,890	0,974	0,976	0,978	0,902	0,981	0,981	0,981
Sample 5	0,947	0,991	0,992	0,992	0,972	0,997	0,998	0,998
Sample 6	0,953	0,988	0,993	0,995	1,000	1,000	1,000	1,000
Sample 7	0,938	0,978	0,979	0,981	0,939	0,988	0,990	0,991
Sample 8	0,907	0,988	0,988	0,990	1,000	1,000	1,000	1,000
Sample 9	0,880	0,978	0,978	0,980	0,860	0,940	0,943	0,943
Sample 10	0,921	0,982	0,985	0,986	0,974	1,000	1,000	1,000

Table 3.4. Number of runs where perfect classification is achieved with F-score based golden fitness

	Non-optimized				Optimized			
	1th Gen	**20th Gen**	**50th Gen**	**100th Gen**	**1th Gen**	**20th Gen**	**50th Gen**	**100th Gen**
Sample 1	0	8	8	9	2	10	10	10
Sample 2	0	5	5	6	0	8	9	10
Sample 3	0	1	1	1	0	0	1	2
Sample 4	0	1	1	1	0	0	0	0
Sample 5	0	5	5	5	0	8	9	9
Sample 6	0	5	7	8	10	10	10	10
Sample 7	0	0	0	0	0	1	3	4
Sample 8	0	5	5	6	10	10	10	10
Sample 9	0	3	3	3	0	0	0	0
Sample 10	0	2	2	2	2	10	10	10

3.3. Performance of the Cubic Spline Interpolation Based Fitness

Cubic splines are fit both on the centroids of the ground truth blobs and the blobs of each candidate solution (Figure 3.4). Fitness is measured by calculating

RMSE between ground truth spline and candidate solution's spline in an integral sense (i.e. RMSE is measured at each pixel and summed up). Average results and the number of perfect classifications achieved for this fitness test is given in Table 3.5 and Table 3.6 respectively.

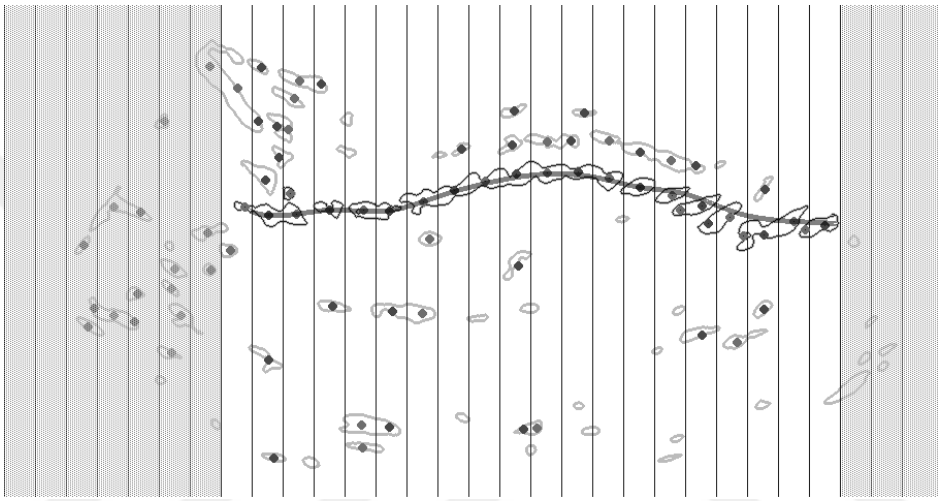


Figure 3.4. A cubic spline, drawn on centroids of actual spine bones.

Table 3.5. Spline fitness based golden fitness test results (average of ten repetitions)

	Non-optimized				**Optimized**			
	1th Gen	**20th Gen**	**50th Gen**	**100th Gen**	**1th Gen**	**20th Gen**	**50th Gen**	**100th Gen**
Sample 1	0,080	0,908	1,000	1,000	0,361	0,910	1,000	1,000
Sample 2	0,181	0,506	0,596	0,636	0,267	0,916	0,916	0,916
Sample 3	0,023	0,070	0,068	0,099	0,029	0,156	0,157	0,259
Sample 4	0,052	0,184	0,192	0,248	0,464	1,000	1,000	1,000
Sample 5	0,040	0,384	0,403	0,529	0,176	0,989	0,989	1,000
Sample 6	0,162	1,000	1,000	1,000	1,000	1,000	1,000	1,000
Sample 7	0,022	0,040	0,053	0,054	0,023	0,161	0,178	0,179
Sample 8	0,403	0,726	0,726	0,723	1,000	1,000	1,000	1,000
Sample 9	0,045	0,308	0,308	0,438	0,055	0,431	0,720	0,820
Sample 10	0,082	0,513	0,512	0,671	0,425	0,957	1,000	1,000

Table 3.6. Number of runs where perfect classification is achieved with spline fitness based golden fitness

	Non-optimized				Optimized			
	1th Gen	**20th Gen**	**50th Gen**	**100th Gen**	**1th Gen**	**20th Gen**	**50th Gen**	**100th Gen**
Sample 1	0	8	10	10	3	9	10	10
Sample 2	0	1	2	2	1	9	9	9
Sample 3	0	0	0	0	0	1	1	2
Sample 4	0	0	0	0	3	10	10	10
Sample 5	0	2	2	4	1	7	7	10
Sample 6	0	10	10	10	10	10	10	10
Sample 7	0	0	0	0	0	1	1	1
Sample 8	2	3	3	3	10	10	10	10
Sample 9	0	0	0	1	0	4	7	8
Sample 10	0	2	2	3	2	9	10	10

3.4. Performance of Proposed Fitness Criteria

The fitness defined specifically for the spine identification in this work assumes no knowledge of the ground truth as it normally is in practical problems. The best solution is compared with the ground truth after the iteration procedure is completed to evaluate the performance of the general method. The RMSE value is calculated by comparing the distance between the centroids of the selected region and the actual spine region in each column.

Additionally, the RMSE value between the curves drawn by the experts and the actual spine (Figure 3.5) is also calculated through the vertical distance between the centroid of the actual bone and the curve in each column (Table 3.7). Note that the error calculation is done in a discrete manner at each column which is different than the integral based calculation given in Section 2.4.5. The error is measured between the blob centroids of the ground truth and the corresponding points of the hand drawn curves (Figure 3.5).

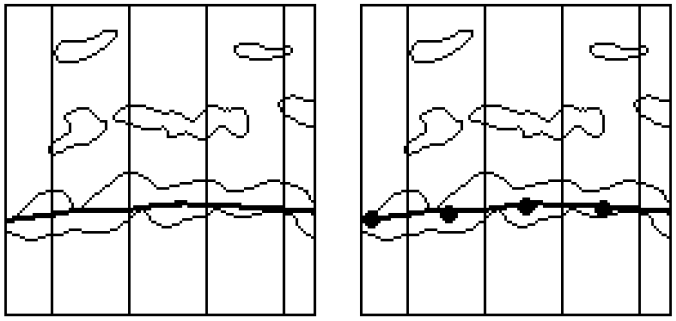


Figure 3.5. Manually drawn curves on a sample region and the extracted points from these curves used in error measurement

Table 3.7. RMSE values for the curves drawn by the three experts

	Exp.1	**Exp.2**	**Exp.3**	**Mean**
Sample 1	14,355	11,556	6,279	10,730
Sample 2	27,181	16,045	11,995	18,407
Sample 3	11,682	19,021	18,201	16,301
Sample 4	10,927	15,784	8,893	11,868
Sample 5	7,619	10,469	12,542	10,210
Sample 6	11,584	8,079	5,627	8,430
Sample 7	15,371	20,473	13,516	16,453
Sample 8	12,562	32,298	8,813	17,891
Sample 9	5,491	10,381	17,603	11,158
Sample 10	9,756	12,532	16,337	12,875

The objective function given in Eq.(2.13) is utilized in the study to evaluate the fitness of each individual. There are three different geometric characteristics (dis-

tance of adjacent centroids, closeness to a cubic polynomial of the centroids and curvature) contributing to the individual's fitness as explained earlier. Effectiveness and functionality of different linear combinations of the weighting parameters α, $\beta$and γ are examined in a grid search scheme to come up with an approximate optimal combination. To achieve that, we have taken a single, moderately evolved generation of individuals from the genetic procedure. Individuals at this generation is ranked first with respect to the proposed fitness criteria and next according to their closeness to the ground truth (which is typically not known under normal circumstances). The rank difference for each individual in both ranking approaches is averaged over all individuals.

Note that the three geometric characteristics have dimensions of length (or length inverse for curvature). To non-dimensionalize them, each quantity is divided to the population average of the same quantity.

In the grid search setting, the weighting parameters are assigned to the values 1, 2, 3 and 4 in all possible combinations. Here, the relative values of the parameters are important rather than their absolute values. Hence, combinations that are multiple of each other are considered equivalent and run only once. For each linear combination of α, β and γ, proposed genetic algorithm is iterated ten times on each sample until the 20th generation. Consequently, average of ranking order errors are calculated for ten repetitions for each sample. The best weight combination for each sample is given in Table 3.8. The parameters of the genetic algorithm are given in Table 3.9.

Table 3.8. Weight combinations resulting to the lowest average order error given for ten samples.

	Lowest Average Error	α	β	γ
Sample 1	45,621	3	1	1
Sample 2	74,187	1	4	1
Sample 3	77,910	4	1	2
Sample 4	95,375	4	1	1
Sample 5	76,481	1	1	4
Sample 6	60,280	1	3	2
Sample 7	143,184	4	1	1
Sample 8	69,157	4	1	1
Sample 9	193,268	4	4	1
Sample 10	69,244	1	1	4

Table 3.9. Initial test parameters are given

Cell Size	30
Cell Threshold	50
Number of Individuals	400
Injection Rate	10
Mutation Rate	90
Max. Number of Generations	100

Tests in three experimental settings are run with the proposed fitness criteria for evaluation purposes. The first setting involves classification with energy min-

imization performed on optimized search space. The second one is performed on optimized search space but without energy minimization applied. The final setting involves classification on non-optimized search space and without energy minimization. Results are given as the average of ten repetitions for each setting (Table 3.10). In some samples, the error does decrease below a certain level and iteration converges to a ”best possible” solution. The rate at which the best individuals reach this best possible solution is given in Table 3.11. In addition, the minimum and maximum RMSE values attained by best individuals over generations are given in Table 3.12 for all settings.

Table 3.10. Average RMSE values of best fitting individuals for ten repetitions over 100 generations

Classification of Optimized Samples with Energy Minimization Approach										
	Samp. 1	Samp. 2	Samp. 3	Samp. 4	Samp. 5	Samp. 6	Samp. 7	Samp. 8	Samp. 9	Samp. 10
Gen:50	0,000	0,000	16,788	14,630	0,000	0,000	0,000	0,000	47,619	0,000
Gen:100	0,000	0,000	1,590	14,630	0,000	0,000	0,000	0,000	47,619	0,000
Classification of Optimized Samples without Energy Minimization Approach										
	Samp. 1	Samp. 2	Samp. 3	Samp. 4	Samp. 5	Samp. 6	Samp. 7	Samp. 8	Samp. 9	Samp. 10
Gen:50	0,000	0,000	13,451	14,630	0,000	0,000	0,000	0,000	47,619	0,000
Gen:100	0,000	0,000	3,010	14,630	0,000	0,000	0,000	0,000	47,619	0,000
Classification of Non-optimized Samples without Energy Minimization										
	Samp. 1	Samp. 2	Samp. 3	Samp. 4	Samp. 5	Samp. 6	Samp. 7	Samp. 8	Samp. 9	Samp. 10
Gen:50	4,472	0,000	26,963	16,394	1,076	23,252	108,280	0,000	45,719	19,204
Gen:100	4,472	0,000	19,922	22,037	0,000	7,364	108,280	0,000	45,719	15,363

Table 3.11. Percentage of best individuals that attain the best possible solution

Classification of Optimized Samples with Energy Minimization Approach										
	Samp. 1	Samp. 2	Samp. 3	Samp. 4	Samp. 5	Samp. 6	Samp. 7	Samp. 8	Samp. 9	Samp. 10
Gen:50	100	100	30	100	100	100	100	100	100	100
Gen:100	100	100	90	100	100	100	100	100	100	100
Classification of Optimized Samples without Energy Minimization Approach										
	Samp. 1	Samp. 2	Samp. 3	Samp. 4	Samp. 5	Samp. 6	Samp. 7	Samp. 8	Samp. 9	Samp. 10
Gen:50	100	100	60	100	100	100	100	100	100	100
Gen:100	100	100	90	100	100	100	100	100	100	100
Classification of Non-optimized Samples without Energy Minimization										
	Samp. 1	Samp. 2	Samp. 3	Samp. 4	Samp. 5	Samp. 6	Samp. 7	Samp. 8	Samp. 9	Samp. 10
Gen:50	100	100	10	40	90	40	100	100	100	50
Gen:100	100	100	50	10	100	80	100	100	100	60

Table 3.12. Minimum and maximum RMSE success over generations

	Non-optimized Samples Without Energy		**Optimized Samples Without Energy**		**Optimized Samples With Energy**	
	Min. Err	Max. Err	Min. Err	Max. Err	Min. Err	Max. Err
Sample 1	4,472	4,472	0,000	0,000	0,000	0,000
Sample 2	0,000	0,000	0,000	0,000	0,000	0,000
Sample 3	0,000	46,618	0,000	30,104	0,000	15,896
Sample 4	8,102	26,359	14,630	14,630	14,630	14,630
Sample 5	0,000	0,000	0,000	0,000	0,000	0,000
Sample 6	0,000	38,976	0,000	0,000	0,000	0,000
Sample 7	108,280	108,280	0,000	0,000	0,000	0,000
Sample 8	0,000	0,000	0,000	0,000	0,000	0,000
Sample 9	45,719	45,719	47,619	47,619	47,619	47,619
Sample 10	0,000	38,409	0,000	0,000	0,000	0,000

The detected spinal axis is visually depicted for each sample in Appendix B. Some images are partially shown in Figure 3.6.

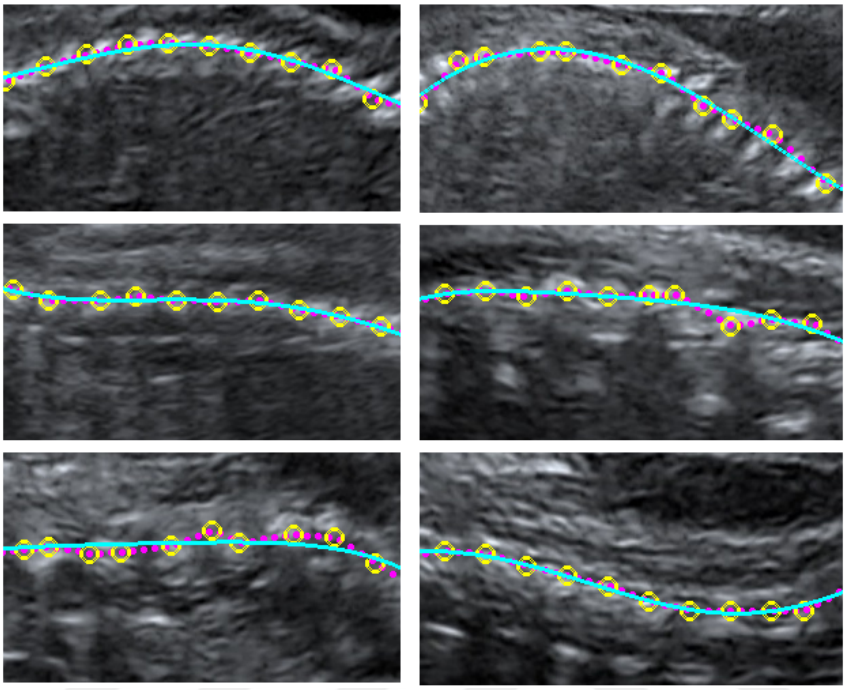


Figure 3.6. Performance of the proposed algorithm demonstrated visually on six sample regions. Continuous curve, circles and dotted line indicate the polynomial fitted to selected blobs, centroids of the selected blobs and line segments connecting the blob centroids, respectively.

4. DISCUSSION AND CONCLUSIONS

4.1. Discussion

While studies towards identification of different organs and fetal features (head radius etc.) on ultrasound images exist, identification of a deformed fetal spine on a sonogram is not much studied in the literature. In this study, we have developed a nature inspired approach for identification of spinal axis in SB cases. This procedure involves locating the spine bones among other detected bone or bone-like blobs.

Conventional techniques that are based on extraction and unsupervised clustering of salient features have shown to underperform, indicating the necessity of more sophisticated and focused methods for exploration and exploitation of the problem search space, i.e. the blobs extracted from raw sonograms via conventional filtering operations.

The necessity of specific algorithm is demonstrated by tackling the problem first with conventional techniques. The statistical approach based on T-test has revealed that the most significant feature changes from sample to sample. Similarly, exhaustive test results with K-means and FCM have shown that the best feature combination to classify bone from non-bone is in general different for each sample. The F-score for the most successful classification has turned out to be 0.71 (Table 2.7 and Table 2.8). This performance variability may be related with variable noise levels, spine deformations and distortions present in sonograms. Consequently, these observations indicate that novel search and optimization algorithms are required to handle the introduced classification problem.

The flocking dynamics based search space optimization approach has been one of the main novelties of this research. The proposed algorithm has been seen to combine bone regions into groups effectively to reduce the total number of bone

blobs forming the search space. It has been observed that the reduced search space can have as much as %64 less bone regions. The algorithm has connected blobs without any inter-class (i.e. bone with non-bone) connections in seven out of ten samples. False kill (i.e. removal of isolated spine blobs) has not been observed in any of the samples.

It is observed that, inter-class connections may form as an unwanted result of proposed optimization approach. To observe the effect of these false connections, polynomials fitted on spines that possibly involve falsely connected spine regions, have been compared with polynomials fitted on actual spine. It has been observed that falsely connected regions may not cause a significant deviation between the two fitted polynomials. The deviation between the two is found to be comparable to the deviation of the curves drawn by the expert obstetricians on the sonograms (Table 2.10).

The flocking dynamics based search space optimization has been further evaluated by identifying the spine using a conventional genetic algorithm. Accordingly, the ground truth is assumed to be known and F-score is accepted as fitness criteria in the genetic algorithm. It should be noted that such an approach will eventually find the actual spine. It has been observed that the same genetic algorithm reaches the ground truth in a fewer number iterations with an optimized search space compared to the non-optimized one (Table 2.11).

The proposed evolutionary scheme is first evaluated with three standard objective functions and found to be convergent under ideal conditions (i.e. when the ground truth is known). The columnar approach utilized in this work seem to be compatible with conventional genetic algorithm operators and capable to reach perfect solution. It has been also verified that proposed search space optimization approach leads to faster convergence.

The fitness function suggested for the introduced spine discrimination problem is first evaluated with respect to its ability to rank individuals in a generation. The grid search methodology, where the weights of three geometric characteristics (distance of adjacent centroids, closeness to a cubic polynomial of the centroids and curvature) varies between one and four in combination, has shown that all of these three components of the fitness function can be significant in selected samples (Table 3.8). With a higher value for the weight α, the fitness function seem to distinguish the "good" individuals in a higher number of samples. Nevertheless, higher values of β and γ may be necessary for better ranking of individuals in different samples. This observation allows us to conclude that, a linear combination of all three components instead of utilizing only one would increase the effectiveness of the fitness function in a bigger pool of samples.

The final test in this study has involved simultaneously the search space optimization, the proposed fitness function and the energy minimization operation applied during the evolution process. The results has shown that search space optimization and energy minimization, when used with the proposed fitness function, may result to a better performance in discriminating spine bones in the data set utilized in this research (Table 3.10). SB pathology causes irregularity usually on the spinous process (i.e. bony projection of posterior vertebra) but not on the corpus (i.e. main body of the spine). However, in some cases of SB, morphological irregularity may occur in the corpus while spinous process looks intact. Our algorithm may capture the spinous process instead of the corpus in such cases. In fact, the samples where the algorithm cannot find the ground truth are those where corpus and spinous process are partially connected during spine identification. While the ground truth is not found in those samples, the spine as a whole is nicely captured from a visual point of view. (See Appendix B.)

To our knowledge, there is only one study that concerns detection of fetal spine from ultrasound images (Tolay, et. al., 2009). In this study, healthy fetal spine axes are detected on relatively clear ultrasound images. The researchers have combined conventional image processing techniques with data clustering approach. Moreover, their work involves cubic polynomial fitting and cross correlation. Similar deterministic methods have been tried in the preliminary stages of this research, yet failed to produce a reasonable outcome. This observation has led us to conclude that presence of SB pathology and varying standards of data quality requires a more sophisticated and feature independent approach such as the one presented here.

In the literature, a CAD system for SB is proposed which is based on altered shape of the skull that occurs in this pathology (Konur, 2018; Konur et al., 2015). These studies involve image processing and machine learning tools to associate the skull shape with the pathology. While we have not proposed a CAD system here, our work is also based on analysis of pathological ultrasound samples. Additionally we have looked to the identification of the spine rather than the skull shape. Once spine is properly identified the next natural step may be classification of healthy and diseased spine.

4.2. Conclusion and Future Work

This thesis study has proposed a novel nature inspired methodology for locating the spine on ultrasound samples of SB cases. The proposed methodology is found to be effective for locating the spine in most of the sonograms. The problem has a clinical significance since SB is a relatively common pathology. Hence, we have worked together with expert obstetricians when developing and evaluating the algorithms. Moreover, the testing stage of certain parts of the algorithm has involved comparison of spines identified by our algorithm and by the experts. We have ob-

served that in most samples the spines identified by the proposed algorithm closely match those drawn by the experts on the sonograms.

In the future, we intend to further refine the presented optimization and classification approaches. More specialised and distinctive features and up-to-date machine learning methods can increase the success of the segmentation stage which would also improve the overall success of the methodology. Search space optimization, by its very nature, can be accomplished in various ways. A more effective search space optimization can be the topic of another study. Similarly, various aspects of the genetic algorithm (parameters of the fitness function, energy minimization and other operators etc.) can be optimized to handle the specific problem at hand.

While this study does not intend to propose a CAD system for SB, the proposed methodology can constitute the first natural stage of such a CAD system. Clearly, such a study constitutes a higher clinical challenge and requires a closer cooperation with clinicians.

REFERENCES

Acharya, U. R., Sree, S. V., Krishnan, M. M. R., Molinari, F., Zielenik, W., Bardales, R. H., ... & Suri, J. S. (2014). ComputerAided Diagnostic System for Detection of Hashimoto Thyroiditis on Ultrasound Images From a Polish Population. Journal of Ultrasound in medicine, 33(2), 245-253.

Baumgartner, C. F., Kamnitsas, K., Matthew, J., Fletcher, T. P., Smith, S., Koch, L. M., ... & Rueckert, D. (2017). Sononet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. IEEE transactions on medical imaging, 36(11), 2204-2215.

Beheshti, Z., & Shamsuddin, S. M. H. (2013). A review of population-based meta-heuristic algorithms. Int. J. Adv. Soft Comput. Appl, 5(1), 1-35.

Bottaci, L. (2001, April). A genetic algorithm fitness function for mutation testing. In Proceedings of the SEMINALL-workshop at the 23rd international conference on software engineering, Toronto, Canada.

Bruner, J. P., & Tulipan, N. (2004). Tell the truth about spina bifida. Ultrasound in Obstetrics & Gynecology, 24(6), 595-596.

Buyukkurt, S., Binokay, F., Seydaoglu, G., Gulec, U. K., Ozgunen, F. T., Evruke, C., & Demir, C. (2013). Prenatal determination of the upper lesion level of spina bifida with three-dimensional ultrasound. Fetal diagnosis and therapy, 33(1), 36-40.

Caixinha, M., Santos, M., & Santos, J. (2016). Automatic cataract hardness classification ex vivo by ultrasound techniques. Ultrasound in medicine & biology, 42(4), 989-998.

Cannon, R. L., Dave, J. V., & Bezdek, J. C. (1986). Efficient implementation of the fuzzy c-means clustering algorithms. IEEE transactions on pattern analysis and machine intelligence, (2), 248-255.

Carneiro, G., Georgescu, B., Good, S., & Comaniciu, D. (2007, October). Automatic

fetal measurements in ultrasound using constrained probabilistic boosting tree. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 571-579). Springer, Berlin, Heidelberg.

Cengizler C., Un M.K. (2017). Evaluation of Calinski-Harabasz Criterion as Fitness Measure for Genetic Algorithm Based Segmentation of Cervical Cell Nuclei. British Journal of Mathematics & Computer Science, 22, 1-13.

Chang, D. X., Zhang, X. D., & Zheng, C. W. (2009). A genetic algorithm with gene rearrangement for K-means clustering. Pattern Recognition, 42(7), 1210-1222.

Copp, A. J., Adzick, N. S., Chitty, L. S., Fletcher, J. M., Holmbeck, G. N., & Shaw, G. M. (2015). Spina bifida. Nature reviews Disease primers, 1, 15007.

Cui, X., Gao, J., & Potok, T. E. (2006). A flocking based algorithm for document clustering analysis. Journal of systems architecture, 52(8-9), 505-515.

D'Agostini, G. (1995). A multidimensional unfolding method based on Bayes' theorem. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 362(2-3), 487-498.

Dehmeshki, J., Ye, X., Lin, X., Valdivieso, M., & Amin, H. (2007). Automated detection of lung nodules in CT images using shape-based genetic algorithm. Computerized Medical Imaging and Graphics, 31(6), 408-417.

Drukker, K., Giger, M. L., Horsch, K., Kupinski, M. A., Vyborny, C. J., & Mendelson, E. B. (2002). Computerized lesion detection on breast ultrasound. Medical physics, 29(7), 1438-1446.

El-Dahshan, E. S. A. (2011). Genetic algorithm and wavelet hybrid scheme for ECG signal denoising. Telecommunication Systems, 46(3), 209-215.

Ghaheri, A., Shoar, S., Naderan, M., & Hoseini, S. S. (2015). The applications of genetic algorithms in medicine. Oman medical journal, 30(6), 406.

Goutte, C., & Gaussier, E. (2005, March). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In European Conference on Information Retrieval (pp. 345-359). Springer, Berlin, Heidelberg.

Gudmundsson, M., El-Kwae, E. A., & Kabuka, M. R. (1998). Edge detection in medical images using a genetic algorithm. IEEE transactions on medical imaging, 17(3), 469-474.

Guler, I., & Ubeyli, E. D. (2003). Detection of ophthalmic artery stenosis by least-mean squares backpropagation neural network. Computers in Biology and Medicine, 33(4), 333-343.

Gupta, D., & Ghafir, S. (2012). An overview of methods maintaining diversity in genetic algorithms. International journal of emerging technology and advanced engineering, 2(5), 56-60.

Holland, J. H. (1992). Genetic algorithms. Scientific american, 267(1), 66-73.

Horsch, K., Giger, M. L., Venta, L. A., & Vyborny, C. J. (2002). Computerized diagnosis of breast lesions on ultrasound. Medical Physics, 29(2), 157-164.

Huang, Q., Zhang, F., & Li, X. (2018). Machine Learning in Ultrasound Computer-Aided Diagnostic Systems: A Survey. BioMed research international, 2018.

Jiang, C. F. (2001, July). Autosegmentation of ultrasonic images by the genetic algorithm. In Medical Imaging 2001: Image Processing (Vol. 4322, pp. 151-158). International Society for Optics and Photonics.

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. IEEE Transactions on Pattern Analysis & Machine Intelligence, (7), 881-892.

Kuri-Morales, A., & Rodrguez-Erazo, F. (2009). A search space reduction methodology for data mining in large databases. Engineering Applications of Artificial Intelligence, 22(1), 57-65.

Konur, U., Grgen, F. S., Varol, F., & Akarun, L. (2015). Computer aided detection of spina bifida using nearest neighbor classification with curvature scale space features of fetal skulls extracted from ultrasound images. Knowledge-Based Systems, 85, 80-95.

Konur, U. (2018). Computerized detection of spina bifida using SVM with Zernike moments of fetal skulls in ultrasound screening. Biomedical Signal Processing and Control, 43, 18-30.

Lorena, L. A., Narciso, M. G., & Beasley, J. E. (2002). A constructive genetic algorithm for the generalized assignment problem. Evolutionary Optimization, 5, 1-19.

Mansoor, A., Bagci, U., Foster, B., Xu, Z., Papadakis, G. Z., Folio, L. R., ... & Mollura, D. J. (2015). Segmentation and image analysis of abnormal lungs at CT: current approaches, challenges, and future trends. RadioGraphics, 35(4), 1056-1076.

Maraci, M. A., Bridge, C. P., Napolitano, R., Papageorghiou, A., & Noble, J. A. (2017). A framework for analysis of linear ultrasound videos to detect fetal presentation and heartbeat. Medical image analysis, 37, 22-36.

Marinakis, Yannis, Georgios Dounias, and Jan Jantzen. "Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification." Computers in Biology and Medicine 39.1 (2009): 69-78.

Mnigot, S., & Girault, J. M. (2016). Optimization of contrast resolution by genetic algorithm in ultrasound tissue harmonic imaging. Ultrasonics, 71, 231-244.

Nasiri, J. A., Naghibzadeh, M., Yazdi, H. S., & Naghibzadeh, B. (2009, November). ECG arrhythmia classification with support vector machines and genetic algorithm. In Computer Modeling and Simulation, 2009. EMS'09. Third UKSim European Symposium on (pp. 187-192). IEEE.

Ozaras, N. (2015). Spina Bifida and Rehabilitation. Turkish Journal of Physical Medicine and Rehabilitation, 61(1).

Pena-Reyes, C. A., & Sipper, M. (1999). A fuzzy-genetic approach to breast cancer diagnosis. Artificial intelligence in medicine, 17(2), 131-155.

Pereira, D. C., Ramos, R. P., & Do Nascimento, M. Z. (2014). Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm. Computer methods and programs in biomedicine, 114(1), 88-101.

Poli, R., Cagnoni, S., & Valli, G. (1995). Genetic design of optimum linear and non-linear QRS detectors. IEEE Transactions on Biomedical Engineering, 42(11), 1137-1141.

Racusin, D. A., Villarreal, S., Antony, K. M., Harris, R. A., Mastrobattista, J., Lee, W., ... & Aagaard, K. M. (2016). Role of maternal serum alpha-fetoprotein and ultrasonography in contemporary detection of spina bifida. Obstetrical & Gynecological Survey, 71(4), 199-200.

Reynolds, C. W. (1987, August). Flocks, herds and schools: A distributed behavioral model. In ACM SIGGRAPH computer graphics (Vol. 21, No. 4, pp. 25-34). ACM.

Srinivas, M., & Patnaik, L. M. (1991, November). Learning neural network weights using genetic algorithms-improving performance by search-space reduction. In Neural Networks, 1991. 1991 IEEE International Joint Conference on (pp. 2331-2336). IEEE.

Su, H., Wang, X., & Lin, Z. (2009). Flocking of multi-agents with a virtual leader. IEEE Transactions on Automatic Control, 54(2), 293-307.

Talebi, M., Ayatollahi, A., & Kermani, A. (2011). Medical ultrasound image segmentation using genetic active contour. Journal of Biomedical Science and Engineering, 4(02), 105.

Tolay, P., Vajinepalli, P., Bhattacharya, P., Firtion, C., & Sisodia, R. S. (2009, February). Automated fetal spine detection in ultrasound images. In Medical Imaging 2009: Computer-Aided Diagnosis (Vol. 7260, p. 72603O). International Society for Optics and Photonics.

Trudell, A. S., & Odibo, A. O. (2014). Diagnosis of spina bifida on ultrasound: always termination?. Best Practice & Research Clinical Obstetrics & Gynaecology, 28(3), 367-377.

Ulyanenkov, A., & Sobolewski, S. (2005). Extended genetic algorithm: application to x-ray analysis. Journal of Physics D: Applied Physics, 38(10A), A235.

Whitley, D. (1994). A genetic algorithm tutorial. Statistics and computing, 4(2), 65-85.

Wu, W. J., Lin, S. W., & Moon, W. K. (2012). Combining support vector machine with genetic algorithm to classify ultrasound breast tumor images. Computerized Medical Imaging and Graphics, 36(8), 627-633.

Yang, M. C., Huang, C. S., Chen, J. H., & Chang, R. F. (2012). Whole breast lesion detection using naive bayes classifier for portable ultrasound. Ultrasound in medicine & biology, 38(11), 1870-1880.

Yang, X. S. (2010). A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization (NICSO 2010) (pp. 65-74). Springer, Berlin, Heidelberg.

Yang, X. S. (2011). Review of metaheuristics and generalized evolutionary walk algorithm. arXiv preprint arXiv:1105.3668.

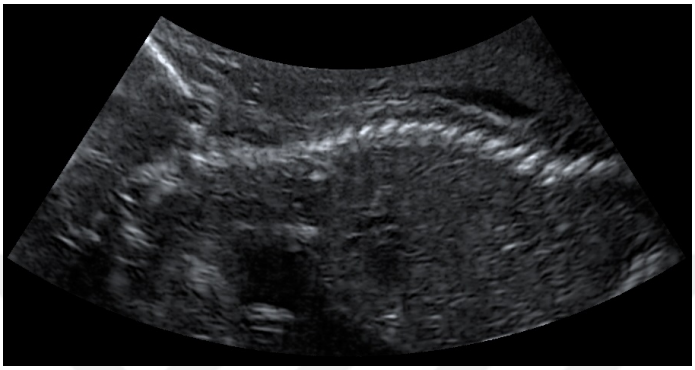
Yeh, J. Y., & Fu, J. C. (2008). A hierarchical genetic algorithm for segmentation of multi-spectral human-brain MRI. Expert Systems with Applications, 34(2), 1285-1295.

CURRICULUM VITAE

Caglar Cengizler was born on 02/08/1986 in Sivas. He received his Bachelor's degree in electrical and electronics engineering from Atilim University in 2010. He received his M.S. degree in Computer Engineering Department from Cukurova University in 2013. He still works as lecturer in Cukurova University.

APPENDIX

APPENDIX A: The raw ultrasound samples before preprocessing are given in this Appendix.



Appendix A.1. Ultrasound sample 1



Appendix A.2. Ultrasound sample 2



Appendix A.3. Ultrasound sample 3



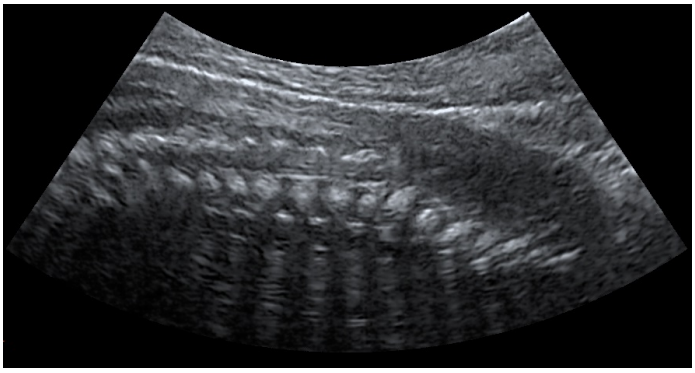
Appendix A.4. Ultrasound sample 4



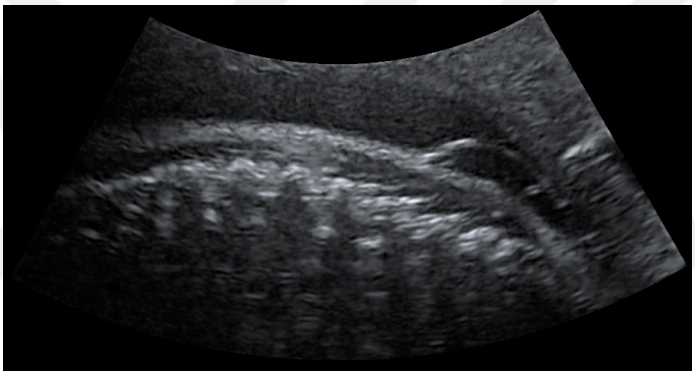
Appendix A.5. Ultrasound sample 5



Appendix A.6. Ultrasound sample 6



Appendix A.7. Ultrasound sample 7



Appendix A.8. Ultrasound sample 8

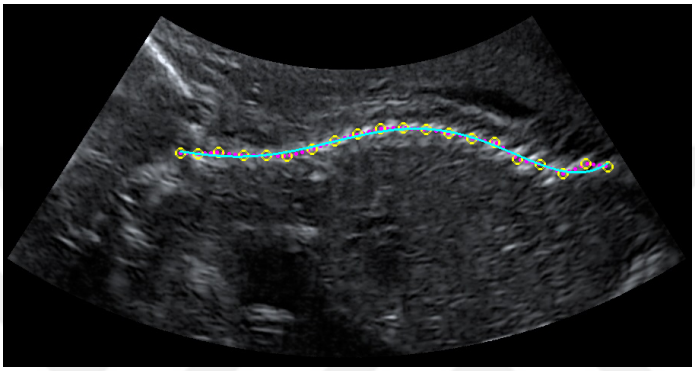


Appendix A.9. Ultrasound sample 9

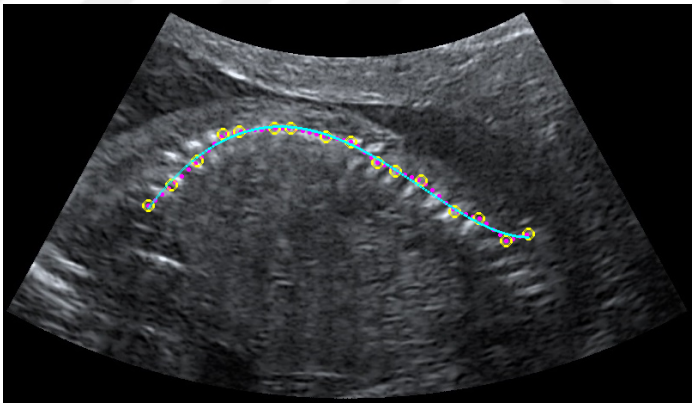


Appendix A.10. Ultrasound sample 10

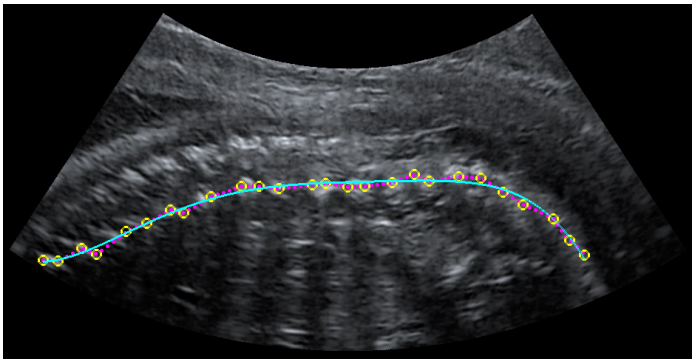
APPENDIX B: The detected spinal axis is visually depicted on each sample. Continuous curve, circles and dotted line indicate the polynomial fitted to selected blobs, centroids of the selected blobs and line segments connecting the blob centroids, respectively.



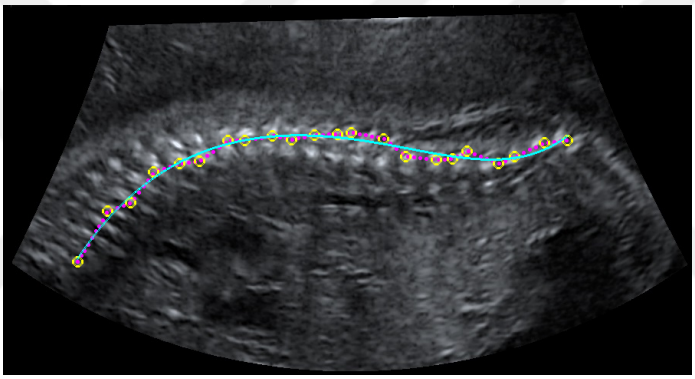
Appendix B.1. Detected axis is indicated on ultrasound sample 1



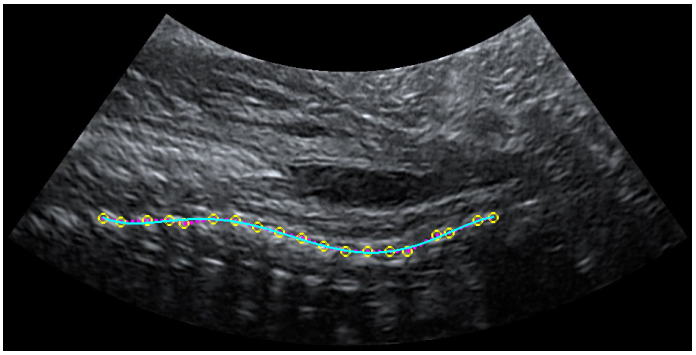
Appendix B.2. Detected axis is indicated on ultrasound sample 2



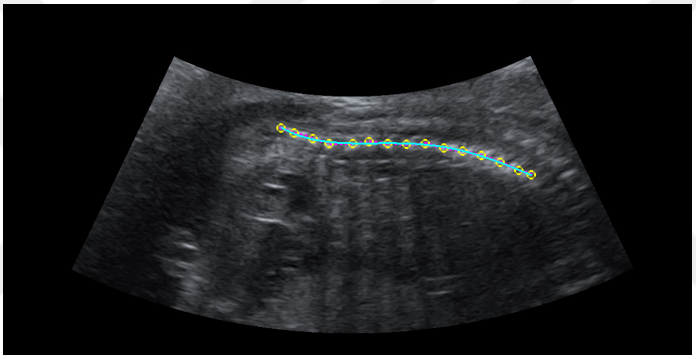
Appendix B.3. Detected axis is indicated on ultrasound sample 3



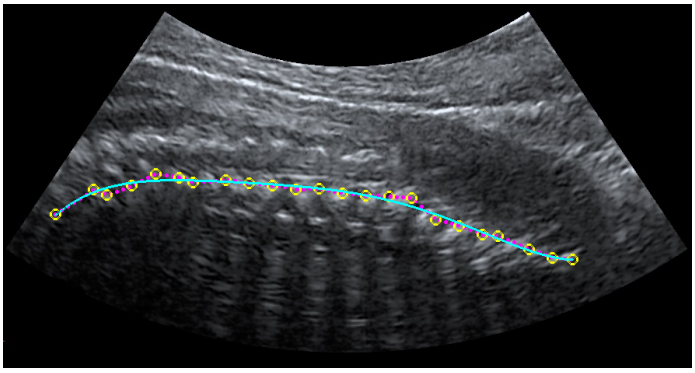
Appendix B.4. Detected axis is indicated on ultrasound sample 4



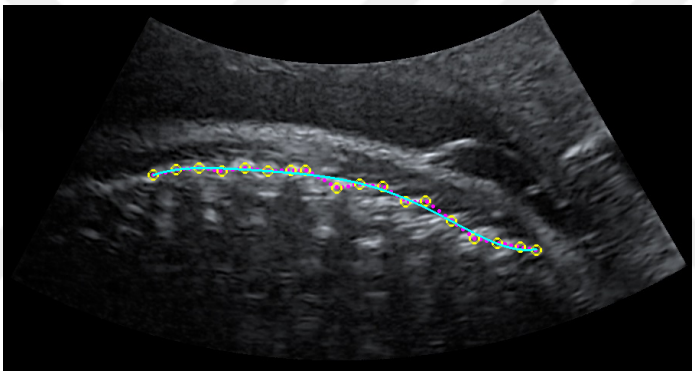
Appendix B.5. Detected axis is indicated on ultrasound sample 5



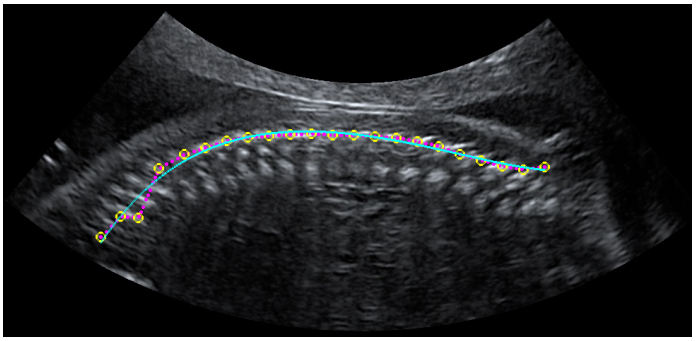
Appendix B.6. Detected axis is indicated on ultrasound sample 6



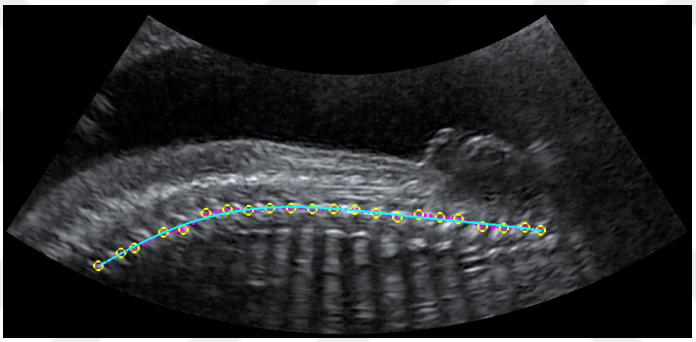
Appendix B.7. Detected axis is indicated on ultrasound sample 7



Appendix B.8. Detected axis is indicated on ultrasound sample 8



Appendix B.9. Detected axis is indicated on ultrasound sample 9



Appendix B.10. Detected axis is indicated on ultrasound sample 10